ABSTRACT

Title of Dissertation: DISCRETE AND CONTINUOUS VARIABLE SYSTEMS: PROPERTIES, PROTOCOLS, AND APPLICATIONS

Joseph Thomas Iosue
Doctor of Philosophy, 2025

Dissertation Directed by: Professor Alexey V. Gorshkov
Department of Physics

Professor Victor V. Albert
Department of Physics

Quantum information science is a promising, interdisciplinary field focusing on both understanding and utilizing quantum systems. Two major paradigms of quantum mechanics are discrete variable (finite dimensional) systems, such as qubits and qudits, and continuous variable (infinite dimensional) systems, such as bosonic modes. In this dissertation, we explore the properties, protocols, and applications of both discrete and continuous variable systems.

In the first part of this dissertation, we study Hilbert space structures called quantum state designs, which are small ensembles of quantum states that mimic properties of the full space. While such designs are well-studied in the discrete variable setting, we show that they can also be defined and constructed in the continuous variable setting. Using specific multimode ensembles, we demonstrate continuous variable shadow tomography protocols which allow for efficient estimation of expectation values of many observables. Additionally, we use these ensembles to define notions of average and entanglement fidelities of continuous variable quantum channels, and we derive an explicit relationship between them that resembles

the analogous relationship in the discrete variable setting.

Meanwhile, on the discrete variable side, we construct a theory of designs on the torus and find general methods for constructing them in arbitrary dimensions. Using these toric designs and their relationship to quantum state designs, we construct many new and explicit families of quantum state designs. Furthermore, we use toric designs to prove various structure theorems about complete sets of mutually unbiased bases.

In the second part of this dissertation, we examine entanglement in continuous variable systems. Specifically, we analytically derive average and typical entanglement properties, as measured by all integer Rényi-α entropies, of random ensembles of Gaussian states outputted from a Gaussian boson sampling device.

Finally, in the third part of this dissertation, we examine the use of qubit systems for resolving frequency spectrums in signal processing applications. Specifically, we show that a classical signal whose spectrum contains closely spaced frequencies can be resolved by coupling the signal to a qubit and performing a superresolution protocol. We find general conditions for a protocol to exhibit superresolution and show various analytic and numerically-optimized protocols that achieve superresolution.

DISCRETE AND CONTINUOUS VARIABLE SYSTEMS: PROPERTIES, PROTOCOLS, AND APPLICATIONS

by

Joseph Thomas Iosue

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:
Professor Nathan Schine, Chair
Professor Alexey V. Gorshkov, Co-Chair and Research Advisor
Professor Victor V. Albert, Co-Chair and Research Advisor
Professor Alexander Barg, Dean's Representative
Professor Mohammad Hafezi

Acknowledgements

I would first like to thank my advisors, Alexey Gorshkov and Victor Albert, for their guidance and support throughout my PhD journey.

Since I first met Alexey when I was in undergrad, he has been a constant source of inspiration and enthusiasm, and he is what initially drew me to the University of Maryland for graduate school. Given to his incredible breadth of knowledge and expertise, he has an uncanny ability to get me "unstuck" with a continuous stream of ideas and new ways of looking at a problem. He is an exceptional teacher and mentor, and whatever the context, he has always had my best academic and personal interests in mind. Additionally, his (apparent) appreciation for my frequent attempts at humor is always confidence-inspiring.

Victor's ability to consistenly generate new ideas and insights, often by combining seemingly disparate concepts from math, physics, and computer science, has been remarkable to witness since we began working together. Some of the highlights of my time in graduate school have been our many long whiteboard sessions, which always reminded me of what pleasure comes from working together through problems and coming to new understandings. I only wish we could have spent more time around a chalkboard instead.

I have had the pleasure of working with many other wonderful people during my PhD, including Alexander Barg, Ansgar Burchards, Yu-An Chen, Jonathan Conrad, Abhinav Deshpande, Adam Ehrenberg, Chris Fechisin, Michael Gullans, Dominik Hangleiter, Shubham Jain, Zijian Liang, Jacob Lin, Connor Mooney, Leigh Norris, Jeet Shah, Kunal Sharma, Paraj Titum, Yan-Qi Wang, Yu-Xin Wang, Braydon Ware, James Watson, Yijia Xu, Bowen

Yang, and Jason Youm, all of whom I have had the great fortune of being co-authors with on various projects. I would especially like to thank Connor, Kunal, Paraj, Yu-Xin, and Adam, and I hope to have many more collaborations with them in the future. Outside of direct co-authors, I would like to thank all of the members of the Gorshkov and Albert groups, especially Jake Bringewatt and Dhruv Devulapalli, for many fun and useful conversations.

I am tremendously grateful to my family for all of their support and for shaping who I am today. It is hard to fathom everything they have done for me and the great influence they have been and continue to be on my life. Especially to my mom, who first piqued my interest in mathematics, and to my dad, who originally sparked my curiosity in mechanics, thereby naturally leading me down the route of mathematical physics. And 對我的新家人— 謝謝你們的愛和支持。我很高興叫你們我的家人！

Finally, and most importantly, thank you to my wife, Andrea. You've had to put up with a lot these last five years, not the least being an unhealthy amount of 3AM wake-ups where I was *sure* I solved the problem I was laying in bed thinking about, only to then find me tired and cranky the next morning, at which point I had inevitably determined my subconscious solution to be flawed. But the day always became brighter with you around, and I am so fortunate to know that you will always be by my side.

Citations to Previously Published Work

With the exception of Chapter 7 and Appendix E, all the work appearing in this dissertation has either been published or appears on the preprint server *arXiv*. The work appearing in Chapter 7 and Appendix E is currently in preparation to appear on *arXiv* [1]; this work was done at and in collaboration with Johns Hopkins University Applied Physics Laboratory. Other works that I have been involved with during my graduate studies that do not appear in this thesis are listed below along with a summary of their content.

Part I: Entanglement and Gaussian boson sampling

- Chapter 2 and Appendix A. Ref. [2]. J. T. Iosue, A. Ehrenberg, D. Hangleiter, A. Deshpande, and A. V. Gorshkov, "Page curves and typical entanglement in linear optics", Quantum **7**, 1017 (2023), arXiv:2209.06838 [quant-ph]

- Chapter 3 and Appendix B. Ref. [3]. J. Youm, J. T. Iosue, A. Ehrenberg, Y.-X. Wang, and A. V. Gorshkov, "Average Rényi entanglement entropy in Gaussian boson sampling", Phys. Rev. Res. **7**, 023125 (2025), arXiv:2403.18890 [quant-ph]

Other works related to Gaussian boson sampling not included in this thesis:

- Ref. [4]. A. Ehrenberg, J. T. Iosue, A. Deshpande, D. Hangleiter, and A. V. Gorshkov, "Transition of anticoncentration in Gaussian boson sampling", Phys. Rev. Lett. **134**, 140601 (2025), arXiv:2312.08433 [quant-ph]

Studies anticoncentration in Gaussian boson sampling by using a graph theoretic approach to compute moments of relevant Hafnians. Finds a transition in anticoncentration, hinting at a transition in sample complexity.

- Ref. [5]. A. Ehrenberg, J. T. Iosue, A. Deshpande, D. Hangleiter, and A. V. Gorshkov, "Second moment of Hafnians in Gaussian boson sampling", Phys. Rev. A **111**, 042412 (2025), arXiv:2403.13878 [quant-ph]

  Companion work to Ref. [4]. Fully develops the graph theoretic approach and numerically exactly computes the expected value of the linear cross entropy benchmark in Gaussian boson sampling experiments.

Other works studying entanglement in quantum information more broadly not included in this thesis:

- Ref. [6]. Z. Liang, Y. Xu, J. T. Iosue, and Y.-A. Chen, "Extracting topological orders of generalized Pauli stabilizer codes in two dimensions", PRX Quantum **5**, 030328 (2024), arXiv:2312.11170 [quant-ph]

  Develops an algorithm for enumerating and studying the topological and anyonic properties of translationally invariant Pauli stabilizer codes on qudits of any dimension.

- Ref. [7]. Z. Liang, B. Yang, J. T. Iosue, and Y.-A. Chen, *Operator algebra and algorithmic construction of boundaries and defects in (2+1)d topological pauli stabilizer codes*, Oct. 2024, arXiv:2410.11942 [quant-ph]

  Extends the results and algorithms from Ref. [6] to systems with boundaries, studying boundary anyons and the bulk-boundary correspondence.

- Ref. [8]. J. Shah, C. Fechisin, Y.-X. Wang, J. T. Iosue, J. D. Watson, Y.-Q. Wang, B. Ware, A. V. Gorshkov, and C.-J. Lin, *Instability of steady-state mixed-state symmetry-*

protected topological order to strong-to-weak spontaneous symmetry breaking, Oct. 2024, arXiv:2410.12900 [quant-ph]

Analytically and numerically studies symmetry protected topological order and strong to weak symmetry breaking in mixed quantum states.

Part II: Designs

- Chapter 4 and Appendix C. Ref. [9]. J. T. Iosue, K. Sharma, M. J. Gullans, and V. V. Albert, "Continuous-variable quantum state designs: theory and applications", Phys. Rev. X **14**, 011013 (2024), arXiv:2211.05127 [quant-ph]

- Chapter 5. Ref. [10]. J. Conrad, J. T. Iosue, A. G. Burchards, and V. V. Albert, *Continuous-variable designs and design-based shadow tomography from random lattices*, Dec. 2024, arXiv:2412.17909 [quant-ph]

- Chapter 6 and Appendix D. Ref. [11]. J. T. Iosue, T. C. Mooney, A. Ehrenberg, and A. V. Gorshkov, "Projective toric designs, quantum state designs, and mutually unbiased bases", Quantum **8**, 1546 (2024), arXiv:2311.13479 [quant-ph]

Other works related to designs in quantum information theory not included in this thesis:

- Ref. [12]. S. P. Jain, J. T. Iosue, A. Barg, and V. V. Albert, "Quantum spherical codes", Nature Physics **20**, 10.1038/s41567-024-02496-y (2024), arXiv:2302.11593 [quant-ph]

  Generalizes classical spherical error correction codes to the bosonic quantum setting, and uses tools from designs theory to designs and analyze codes. Some of the techniques from Part II are used here.

Part III: Sensing

- Chapter 7 and Appendix E. Ref. [1]. J. T. Iosue, L. Norris, and P. Titum, *Superresolution in quantum sensing*, To appear, 2025

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Quantum mechanics is arguably the most experimentally well-tested theory in all of physics. Like any good mathematical theory, quantum physics can be derived from a small set of axioms, typically called the postulates of quantum mechanics. From just these postulates and comparable axioms for classical mechanics, it has been proven that quantum computers can solve various oracle[1] tasks exponentially faster than classical computers [13]. Even without oracle assumptions, there are quantum algorithms that offer super-polynomial speedups over the best known classical algorithms for many computational tasks. Shor's factoring algorithm, for example, could break in polynomial time the RSA protocol that encrypts the majority of financial and otherwise sensitive data online [14]. This alone demonstrates the importance of understanding the complexity of quantum theory.

Given this setting, in this thesis, we are broadly interesting in (1) understanding the complexity of quantum systems, (2) developing useful protocols for quantum devices, and (3) finding applications of quantum systems. We take these one at a time.

Understanding complex quantum systems In order to determine what types of problems quantum systems are more useful for than classical systems, it is helpful to understand *how* and *why* they are powerful. The origin of the complexity and the computational power of quantum systems is surprisingly subtle. A common misconception is to attribute too much

[1] A quantum *oracle* in this context is a black box that can coherently implement a function (*e.g.* in superposition). It is an unphysical and mathematical ideality, but is nonetheless a useful assumption for proving rigorous separations between quantum and classical computation. Ultimately, we hope to prove separations without invoking oracle assumptions, both within and between the classical and quantum complexity hierarchies, but this has thus far proven difficult in both the classical and quantum settings [15].

of the power of quantum computation to superposition. Indeed, superposition is a necessary but not sufficient condition for a quantum device to computationally outperform classical devices. The most immediate way to see that superposition is not powerful on its own is to consider a probabilistic classical computer. We can, without loss of generality [16, 17], consider reversible classical computation. A classical computation on n bits is then simply a permutation of the bits, and a probabilistic classical computation is a convex combination of permutations. To see that probabilistic classical computation acts in superposition, we can map n-bit bitstrings to 2^n dimensional vectors via a map $b \leftrightarrow |b\rangle$. A probabilistic classical computation acting on b acts as $M|b\rangle$ for some doubly stochastic (*i.e.* convex combination of permutations) matrix M. Given the expansion $M|b\rangle = \sum_{b'} p_{b'}|b'\rangle$, the output is the bitstring b' with probability $p_{b'}$. The fact that $p_{b'}$ can be any probability distribution is the statement that this computation is happening in superposition.

While superposition is not alone sufficient, it is tempting to think that the addition of interference effects gives rise to additional power. However, classical wave mechanics allows for superposition and interference, and yet it is computationally no more powerful than universal classical computation.

Notice that there is no quantum entanglement present in classical wave mechanics. Thus, we may be led to conjecture that the combination of superposition, interference, and entanglement yields quantum computation outperforming classical computation. However, the benefit of entanglement is subtle, and indeed the study of entanglement and its relation to computational complexity is the key motivation behind Part I of this thesis. Entanglement is a necessary ingredient for a quantum computational advantage over probabilistic classical computation, since quantum computations with little entanglement can be simulated efficiently classically [18–20]. On the other hand, too much entanglement can cause a state to be useless for computation. In particular, the typical (over the Haar measure) finite-dimensional quantum state is too entangled to be useful for computation [21, 22]. One particularly nice

example of a class of quantum computations with superposition, interference, and entanglement that nonetheless yields no advantage over classical computation is the Clifford group [23].

Part I of this thesis is motivated by understanding the relationship between entanglement and complexity in a specific setting, called Gaussian boson sampling (GBS), that is thought to exhibit quantum advantage—that is, up to certain complexity theoretic conjectures, it is not possible for a classical computer to efficiently sample from the output probability distribution defined by GBS [24–28]. A GBS device can be thought of as part of a restricted class of quantum computers; it is not capable of universal quantum computation or universal classical computation, but it can nonetheless efficiently perform sampling tasks that are believed to require exponentially more time with classical computers. Due to the more restrictive nature of GBS relative to universal quantum computation and its experimental feasibility [29–31], GBS offers a tractable context to analytically study the nature of quantum advantage and entanglement.

GBS experiments generate random quantum states with respect to an ensemble defined in Part I. There are states in this ensemble that are easy to efficiently classically simulate (in the sense of sampling), however the *typical* state in the ensemble is believed to be classically intractable to simulate. Thus, in order to understand the interplay between complexity and entanglement in this setting, we must first understand the *typical properties* of entanglement. In Chapter 2 of Part I, we study entanglement as measured by the Rényi-2 entropy [32]. We prove certain regimes in the parameter space of a GBS experiment where Rényi-2 entanglement is typical, and indeed we derive analytically the value that it typically takes. In Chapter 3 of Part I, we extend these results to Rényi-α entropies [32] for all integer $\alpha \geq 1$, including the von Neumann entanglement entropy $\alpha = 1$.

Prior work [33, 34] considered typical entanglement properties more analogously to the setting of thermodynamics, where one considers the interaction between a subsystem S and

a much larger bath B. That is, given a GBS experiment, one can divide the system into two parts labeled S and B. In the setting when S is *much* smaller than B, Refs. [33, 34] proved that entanglement is typical. A key takeaway of Part I of this thesis is that entanglement in GBS is typical even when S and B are the same size. Therefore, this result is potentially much more relevant as it tells us that no matter how one measures entanglement in a GBS experiment, it will always be typical.

Developing useful protocols for quantum devices In this thesis, we study both discrete variable (DV) and continuous variable (CV) quantum systems, where the former corresponds to quantum mechanics in finite dimensions and the latter to infinite dimensions. In Part I, we study entanglement in GBS, which is a CV system; in Part II, we study both CV and DV systems; and in Part III, we study a particular DV system.

There are many advantages to building quantum computers using CV resources as opposed to DV resources, such as a continuous parameter-set of transversal gates and Hamiltonian-based bias preserving gates [35]. However, due to the intricacies of infinite dimensionality, there are many protocols that have been developed for DV devices that have no analogue in CV devices. In Chapters 4 and 5 of Part II, we are broadly interested in addressing this challenge and porting over tools and protocols from DV to CV. Specifically, we focus on protocols that utilize DV quantum t-designs and extend them to CV systems.

In DV quantum information theory, t-designs come in two primary flavors—unitary t-designs and state t-designs [36–70]. The unitary group has a unique uniform probability measure called its Haar measure [71]. As with any probability space, it has moments – *e.g.* first moments are means, (central) second moments are variances, etc. A unitary t-design is an ensemble of unitary operators whose first t moments match that of the unitary group. Similarly, the set of all DV quantum states has a natural probability measure called the Fubini-Study measure [72], and a quantum state t-design is an ensemble of quantum states whose first t moments match those of the full set of quantum states. Intuitively, a

t-design on a space mimics the full space—as t increases, a t-design captures more and more features of the full space.

Quantum state and unitary designs are important tools in tomography [40, 43, 48–52], state distinction [44, 53], randomized benchmarking [47, 54–57], fidelity estimation [47, 58–63], cryptography [64, 65], sensing [66, 73], fundamental physics [45, 67–69], universality of gate sets [74], and error correction [75–78]. t-designs proved difficult to properly define and find in CV systems [79]. This means that none of the applications proven to work through the use of designs, *e.g.* quantum state fidelity relations [47, 54–57, 59, 60] and design-based tomographic protocols [49–52], carry over naturally to CV.

In Chapter 4 of Part II, we shed light on the previous difficulties by proving that CV quantum state and CV unitary t-designs cannot exist for any $t > 1$. The reason they cannot exist is a strictly infinite dimensional phenomenon stemming from properties of infinite sequences in L^p spaces. This no go theorem rules out the straightforward extension of the aforementioned DV protocols to CV systems. Nonetheless, we find that there is still a way forward. Using a mathematical construct known as a rigged Hilbert space [80], we define a new type of CV state design called a rigged t-design. We prove that such a design exists for all t, thereby bypassing the prior no go theorem. Using this construction, we develop some applications, including average fidelity calculations, entanglement verification, and shadow tomography. For the remainder of this overview, we will focus on shadow tomography.

Note that a rigged design no longer only consists of quantum states, but rather of more general elements of the rigged Hilbert space. One way of viewing these more general elements is as elements of positive operator-valued measures (POVMs), where a POVM is simply an arbitrary quantum measurement [13]. Indeed, another way of understanding rigged t-design is as POVMs that satisfy a few extra conditions. Leaving out some technical details, a rigged 1-design is a POVM, and a rigged t-design is a POVM whose second through t^{th} moments are fixed in a particular way. Thus, the $t = 1$ condition allows us to measure with respect to

a rigged design, and the $t > 1$ conditions provide us with additional information about the measurement.

Therefore, we expect that DV protocols that utilize designs as a measurement scheme *can* be ported over to CV systems. One particularly useful such protocol is shadow tomography [49–52]. In DV systems, shadow tomography at its core is an intelligent measurement scheme. Specifically, given measurement access to an unknown quantum state, one can build a *classical snapshot* of the state. An obvious goal is to find a measurement scheme such that the classical snapshot contains the most amount of information about the state as possible. In many settings, shadow tomography is this optimal measurement scheme. DV shadow tomography protocols measure with respect to POVMs defined by 2- or 3-designs and store the measurement outcomes classically. Intuitively, designs are used because they mimic the full state space, effectively "covering" the full space uniformly. With the classical snapshot, one can then classically compute the expectation value of many observables [49, 50, 52]. Furthermore, the classical snapshot provides a useful classical characterization of the state that can be fed into machine learning algorithms [51].

Using rigged t-designs, in Chapter 4 of Part II, we generalize design-based shadow tomography to the CV setting, mostly focusing on the mathematical aspects of such a procedure. Next, in Chapter 5 of Part II, we focus more carefully on the application of CV designs to shadow tomography and develop an explicit procedure. Specifically, we show that an ensemble of CV multimode states, called Gottesman-Kitaev-Preskill (GKP) states [81], forms a rigged 2-design. Furthermore, its third moments can be analytically controlled enough so that measuring with respect to the GKP POVM yields a good CV shadow tomography procedure.

GKP states are important for CV quantum error correction [81–83]. Measuring with respect to the GKP POVM is the same measurement that is performed for syndrome extraction [82]. Thus, our result shows that the tools that have been and are being developed for

GKP quantum error correction [84] can be directly used to perform CV shadow tomography.

In the final chapter of Part II, we return our attention to DV systems and study quantum state designs and other discrete structures in finite dimensions. From the perspective of quantum information theory, the most important and interesting spaces over which to form designs are (1) the set of quantum states [44], (2) the diagonal subgroup of the unitary group [85], (3) the set of CV quantum states [86] (Chapters 4 and 5), (4) the n-sphere [12], and (5) the (projective) unitary group [87]. In Chapter 6 of Part II, we add the *projective torus* to this list, where we define the projective torus to be the standard torus with a global phase redundancy removed[2]. Broadly speaking, we show for the first time that (1-4) are all intimately connected through projective toric designs. In fact, one way of understanding the rigged t-design constructions from Chapter 4 is from the perspective of an infinite dimensional projective toric t-design.

In the first part of Chapter 6 of Part II, we initiate the study of projective toric t-designs, which have not been explicitly studied before. We prove a connection between projective toric t-designs, root lattices, and crystal ball sequences, and use this connection to derive a (sometimes provably tight) lower bound on the size of t-designs for all t. The projective torus has a natural group structure, which allows us to consider projective toric t-designs that are themselves subgroups. We show that cyclic group designs are in one-to-one correspondence with objects from additive combinatorics called *difference sets* [88]. Using a class of difference sets called Sidon sets and Singer sets [89], we explicitly construct an infinite family of minimal projective toric 2-designs, and infinite families of t-designs for all t that are asymptotically minimal. Using these constructions, we further construct infinite families of t-designs on the diagonal subgroup of the unitary group.

Furthermore, we also show that projective toric designs are crucial to understanding

[2]Quantum states in n dimensions are unit vectors with a global phase redundancy removed and thus correspond to the space $\Omega_n/\mathrm{U}(1)$, where Ω_n is the unit sphere in $\mathbb{C}^n$. Analogously, we define the projective torus to be $\mathrm{U}(1)^n/\mathrm{U}(1)$.

another interesting and notoriously difficult object in quantum information theory—*complete sets of mutually unbiased bases (MUBs)* [39, 41, 90]. Complete sets of MUBs are maximal sets of orthonormal bases where measurement in one basis reveals nothing about any of the other bases. For example, measuring a qubit in the X-basis reveals nothing about the state of the qubit along the Z-basis, and vice-versa. Thus, the phenomenon of complete sets of MUBs is at the heart of quantum mechanics. Indeed, due to their relationship to the uncertainty principle, complete sets of MUBs are important in quantum key distribution [91], quantum state reconstruction [39], quantum error correction [92, 93], and entanglement detection [94].

Despite their fundamental importance to both the theory of quantum information and to applications, their existence or non-existence in non-prime-power dimensions has been an open problem for decades. Concrete statements about the structure of such MUBs are few and far in-between, and instead many conjectures stand in their place. Our work makes such a concrete statement by explicitly connecting complete sets of MUBs to projective toric 2-designs. We use this connection to prove a number of interesting statements about complete sets of MUBs. First, we disprove a conjecture by Zhu from 2015 regarding the fundamental structure of quantum state 2-designs, SIC-POVMs, and complete sets of MUBs [95]. We achieve this by constructing explicit counterexamples from the Sidon set projective toric 2-designs that yield quantum state 2-designs. Indeed, these (hundreds of) counterexamples are, to the best of our knowledge, new quantum state 2-designs. Additionally, we prove a fundamental distinction between complete sets of MUBs in prime-power dimensions (where they are better understood) and non-prime-power dimensions (where they are notoriously difficult) in terms of whether the corresponding projective toric 2-design is a group or not. The existence of complete sets of MUBs in non-prime-power dimensions is a long-outstanding problem in quantum information theory. This new distinction that we derive provides new insights into this fundamental problem.

Finding applications of quantum systems Thus far, we have been focused on understanding the computational advantage that quantum systems can provide and on useful protocols and mathematical structures for quantum systems. In the last part of this thesis, Part III, we shift our focus to quantum sensing applications of qubit quantum devices.

Quantum sensing is a broad field that has found applications everywhere from sensitive magnetometry to detection of gravitational waves [96]. A general parameter estimation problem is one where there is some unknown parameter—such as a magnetic field strength or a frequency present in a time varying signal—that we wish to measure. The goal is therefore to estimate this parameter to as high accuracy as possible. In Part I of this thesis, we are interested in understanding when and how quantum devices offer a computational advantage over their classical counterparts; similarly, in Part III of this thesis, we are interesting in understanding when and how quantum resources provide a benefit over classical resources for parameter estimation.

We consider the setting when there is an unknown parameter encoded into the frequencies present in a time varying (classical) signal. Specifically, we wish to identify *peaks* in the frequency spectrum of a stochastic signal. In Chapter 7 of Part III, we find that, because the signal is stochastic, estimating the locations of neighboring peaks with purely classical methods requires observing the signal for a time that scales inversely with the distance between them. The task of resolving two closely spaced peaks is similar to Rayleigh's curse in optics, where two light sources cannot be distinctly resolved if they are separated by less than the diffraction limit [97]. Ref. [98] designed a quantum superresolution protocol that can in principle resolve two arbitrarily close point light sources provided that the centroid of the two point sources is known, thereby beating Rayleigh's curse. Motivated by this, we study the use of quantum resources in the frequency resolution problem. By coupling the signal to a qubit, we find necessary and sufficient conditions under which control and measurement protocols result in substantially improved peak estimation. Analogously to

those in Ref. [98], our protocols require knowledge of the centroid of the two frequencies, and we quantify how accurately the centroid must be known in order for the protocols to succeed.

Reminiscent of our analysis of entanglement in Part I, in Chapter 7 of Part III, we further study improvements to these protocols when we couple many entangled qubits to the signal. Interestingly, we find that the stochasticity of the signal results in an entanglement advantage that is different in character from the more typical *Heisenberg limit* entanglement advantage that appears in more "textbook" parameter estimation problems [96]. In fact, the entanglement advantage, while still present, is less pronounced in this setting that in more standard parameter estimation problems. Thus, frequency estimation problems provide a new fruitful ground for understanding the benefit of various quantum resources to sensing protocols.

Part I

Entanglement in Gaussian boson sampling

Chapter 2: Page curves and typical entanglement in linear optics

Section 2.1: Introduction

In his pioneering paper [99], Page considered a Haar-random pure state in an N-dimensional Hilbert space. He conjectured an exact formula for the average entanglement entropy of the reduced density matrix on an rN-dimensional subspace, where $r \in [0, 1]$. His conjecture was proven soon after [100–102]. This average, which is itself a function of r and N, is now known as the *Page curve*. Page also defined the *average information of the subsystem* as the difference between the maximum of the entropy and the average entropy. This quantity has since come to be known as the *Page correction*. Since then, the Page curve and correction have found applications in a variety of areas, such as black holes [69, 103, 104], quantum information theory [105, 106], and statistical mechanics [107–114], among others. As a next step, many works considered the typical deviation of the entanglement entropy from its average. In particular, a system is said to exhibit *typical entanglement* if there is a vanishingly small probability that a random state has entanglement bounded away from the average. This phenomenon was introduced and studied in a variety of systems [21, 22, 65, 115–121].

Entanglement is a key feature of quantum physics and can be used as a resource to complete various tasks, such as teleportation, key distribution, dense coding, and many others [13, 32, 72, 122, 123]. Furthermore, entanglement is a necessary ingredient for quantum advantage, since quantum computations with little entanglement can be simulated efficiently classically [18–20]. One can define quantum advantage using the language of complexity as

using quantum resources to perform a task that is classically hard but quantumly easy. One such task is sampling from the output probability distribution of a Gaussian Boson Sampling experiment [24–28]. The general relationship between entanglement and complexity is largely unknown, but at least some entanglement is necessary for classical hardness. Similarly, in the setting of Boson Sampling, at least some amount of non-Gaussianity is necessary for sampling hardness [124]. On the other hand, too much entanglement can cause a state to be useless for computation. In particular, the typical (over the Haar measure) finite-dimensional quantum state is too entangled to be useful for computation [21, 22]. Thus, studying average and typical entanglement is necessary for learning about the useful part of entanglement and what utility random states have.

Entanglement in infinite-dimensional quantum states generated by bosonic Gaussian inputs and Gaussian operations has found direct application in areas such as quantum sensing [125–129] and quantum communication [130, Ch. 12]. Furthermore, the original motivation for studying the finite-dimensional Page curve was to study the black hole information paradox [99, 103]. Since the degrees of freedom around black holes are inherently infinite-dimensional bosonic (e.g. photonic) modes, an infinite-dimensional bosonic Page curve may help to understand black hole information dynamics [130, Ch. 14].

Our contributions. In this work, we study Page curves and the typicality of entanglement in continuous-variable bosonic Gaussian states. Specifically, we compute average and typical entanglement quantities averaged over passive (energy-conserving) Gaussian unitaries, also called linear optical unitaries, with a fixed initial product state of squeezed vacuum states on n modes. Indeed, this setup is exactly that of a Gaussian Boson Sampling experiment, which is of great recent interest due to experimental claims of quantum advantage via Gaussian Boson Sampling [29–31].

We describe this setup in more detail in Section 2.2. Then in Section 2.3, we begin by studying the regime when all the initial squeezing strengths are equal, and denote this

squeezing strength by s. We derive an analytic expression for the average Rényi-2 entropy of a subsystem of k modes as a function of s, n, and k that is exact asymptotically in n for arbitrary values of s and k. Using this expression, we exactly compute the corresponding Page correction. These results are summarized in Fig. 2.1. In Section 2.4, we then study the presence of typical entanglement for various scalings of k with n. When the distance between the entanglement of a random state and the average entanglement value vanishes additively (resp. multiplicatively), we say that entanglement is strongly (resp. weakly) typical. We prove that entanglement as measured by the Rényi-2 entropy is weakly typical for any k and strongly typical whenever $k \in o(n)$. We further show that entanglement as measured by the von Neumann entropy is weakly typical whenever $k \in o(n)$. Finally, in Section 2.5, we generalize our discussion to the regime when the initial squeezing strengths are not necessarily equal. We show that, if a certain conjecture is true, then the Rényi-2 and von Neumann entanglement entropies are both weakly typical whenever $k \in o(n)$.

Prior work. Refs. [120, 131] studied Page curves and entanglement in fermionic Gaussian states. Serafini et al. have considered typical entanglement in bosonic Gaussian states [34, 132], where they defined two measures, namely the microcanonical and canonical measures, on the set of all n-mode bosonic Gaussian states and averaged with respect to these measures. Roughly, averaging over the microcanoncial measure corresponds to integrating over all bosonic Gaussian states up to a fixed bounded total energy, and averaging over the canonical measure corresponds to integrating over all bosonic Gaussian states with a Boltzmann weight factor decaying with the energy of the state. Fukuda and Koenig generalized these results by studying entanglement averaged over passive Gaussian unitaries with a fixed initial product state of squeezed vacuum states on n modes [33]. Indeed, their setup is exactly the one we consider in this work. As noted in Ref. [33], the measure defined by fixing squeezing strengths and then applying a random passive Gaussian unitary generalizes both the microcanonical and canonical measures, as the latter two measures can be

expressed as convolutions of the former with certain distributions on the set of all squeezing configurations.

Serafini et al. and Fukuda and Koenig consider entanglement in a subsystem of k bosonic modes when $k \in o(n)$; roughly, Serafini et al. allow $k \in \mathcal{O}(1)$ and Fukuda and Koenig allow $k \in o\big(n^{1/3}\big)$. To the best of our knowledge, there are currently no results on average entanglement or typical entanglement in the regime when $k \in \Omega\big(n^{1/3}\big)$. This is what we address in our work. Concerning typical entanglement, we emphasize that our results do not supplant the results of Ref. [33] in general, but rather only in certain regimes. In particular, we primarily consider the situation of equal squeezing strengths, whereas their results pertain to the situation of arbitrary squeezing. Similarly, we primarily consider the Rényi-2 entropy, whereas their results apply to both the Rényi-2 and von Neumann entropies. We summarize our work and that of Ref. [33] on typical entanglement in bosonic Gaussian states in Table 2.1.

For general quantum states, the von Neumann entropy has certain properties, such as *strong subadditivity*, that the Rényi-α entropies do not. For other such properties, we refer to Ref. [32]. Because of this, the von Neumann entropy is generally considered a better measure of entanglement than the Rényi-α entropies. Notably, however, it has been shown that the Rényi-2 entropy is special when restricting to bosonic Gaussian states. For example, it was recently proven that for bosonic Gaussian states, the Rényi-2 entropy also obeys strong subadditivity [133, 134]. The Rényi-2 entropy is also equal, up to a constant, to the phase-space Shannon sampling entropy $-\int W(\boldsymbol{x},\boldsymbol{p}) \log W(\boldsymbol{x},\boldsymbol{p}) \, \mathrm{d}^{\mathrm{n}}\boldsymbol{x} \, \mathrm{d}^{\mathrm{n}}\boldsymbol{p}$ of the Wigner distribution $W(x,p)$ of the state for Gaussian states [133]. Here the position vector $\boldsymbol{x}$ and momentum vector $\boldsymbol{p}$ parameterize the phase space of n oscillator modes. The phase-space Shannon sampling entropy has an operational meaning in terms of sampling via homodyne detections [133, 135]. Furthermore, it has been shown that for pure tripartite Gaussian states, the Rényi-2 entropy obeys a strong subadditivity inequality that is stronger than that for the von Neumann entropy [136]. Finally, correlation measures for Gaussian states based on the

Rényi-2 entropy have been found that have no counterpart when using the von Neumann entropy [137]. As noted in Ref. [133], the aforementioned results are "planting the seeds for a full Gaussian quantum information theory based on the Rényi-2 entropy." Our work is focused primarily on entanglement in bosonic Gaussian states as measured by the Rényi-2 entropy and can thus be viewed as planting a few more seeds.

Section 2.2: Setup

In this work, we consider a linear optical system of n modes, where each mode is an independent quantum harmonic oscillator. We will restrict our attention to bosonic Gaussian states. We refer to Refs. [33, 138] for background on the theory of Gaussian states, and we provide the necessary details to understand this paper in Appendix A.1.1.

Consider an n-mode mixed state ρ in the Hilbert space of square integrable wavefunctions $\mathcal{H} = L^2(\mathbb{R})^{\otimes n}$. For each $i \in \{1, \ldots, n\}$, let $\hat{x}_i$ and $\hat{p}_i$ be the position and momentum quadrature operators on the i^{th} mode. For each i, define $\hat{r}_i \coloneqq \hat{x}_i$ and $\hat{r}_{n+i} \coloneqq \hat{p}_i$. ρ is a Gaussian state if there is a $\beta > 0$ and a Hamiltonian $\hat{H}$ that is at most quadratic in the quadrature operators such that ρ is the thermal state $\rho \propto \mathrm{e}^{-\beta \hat{H}}$. Since $\hat{H}$ contains only linear and quadratic quadrature terms in $\hat{r}_i$, ρ is fully characterized by its first and second moments, $\operatorname{Tr}(\rho \hat{r}_i)$ and $\sigma_{ij} = \frac{1}{2} \operatorname{Tr}[\rho(\hat{r}_i \hat{r}_j + \hat{r}_j \hat{r}_i)] - \operatorname{Tr}(\rho \hat{r}_i) \operatorname{Tr}(\rho \hat{r}_j)$. σ is called the *covariance matrix* of the state ρ.

A unitary U is Gaussian if there exists a Hamiltonian $\hat{H}$ that is at most quadratic in the quadrature operators such that $U = \mathrm{e}^{-\mathrm{i}\hat{H}}$. One can show that a Gaussian unitary maps Gaussian states to Gaussian states. Any Gaussian state can be generated by acting on the vacuum state with a Gaussian unitary. The set of all Gaussian unitaries is isomorphic to the real symplectic group of $2n \times 2n$ matrices $\mathrm{Sp}(2n)$. By the Euler decomposition theorem of a symplectic matrix, it follows that any pure Gaussian state can be generated by acting on an initial product state of squeezed vacuum modes with a passive (energy conserving) Gaussian

unitary. The set of all passive Gaussian unitaries is isomorphic to $\mathrm{Sp}(2n) \cap \mathrm{O}(2n) \cong \mathrm{U}(n)$, where $\mathrm{O}(2n)$ is the orthogonal group of $2n \times 2n$ matrices and $\mathrm{U}(n)$ is the unitary group of $n \times n$ matrices. $\mathrm{Sp}(2n)$ is not compact and thus does not have a finite Haar measure to average over. Physically, this is due to the fact that the Gaussian operation of squeezing can take on unbounded values in $\mathbb{R}$. However, $\mathrm{U}(n)$ is compact, and it is hence well-defined to consider uniformly sampling from $\mathrm{U}(n)$ according to the finite Haar measure.

We are therefore motivated to consider the following notion of a random pure Gaussian state. Namely, we initialize the i^{th} mode to be in a squeezed vacuum state with fixed squeezing parameter $s_i \in \mathbb{R}$ for each $i \in \{1, \ldots, n\}$. We then randomly sample a passive Gaussian unitary from $\mathrm{U}(n)$ and apply it to the n modes. This random state is thus characterized by a fixed choice of the squeezing parameters s_i and a Haar-random choice of a passive Gaussian unitary U. Understanding the properties of such random states is of great interest, particularly because Gaussian Boson Sampling experiments rely on precisely those state preparations. For squeezing parameters s_i for $i \in \{1, \ldots, n\}$, the total expected number of bosons of the state on the n modes is $\sum_{i=1}^{n} \sinh^2(s_i)$. For simplicity, we will begin by considering the case when all the squeezing parameters are equal; $s_i = s$ for each i for some fixed s. In the case of equal squeezings, the average total boson number per mode is $\sinh^2 s$. The general case of unequal squeezings will be discussed in Section 2.5.

The n modes are then partitioned into two groups — one group of $k = rn$ modes for some $0 \leq r \leq 1$, and one group of $n - k = (1 - r)n$ modes. We then compute the entropy of the reduced state of the k modes, or equivalently, since we are considering pure states, the entropy of the reduced state of the $n - k$ modes [13]. Let the density matrix of the reduced state be ρ. For the entropy function, we will be primarily focused on the Rényi-2 entropy $S_2 = -\log \operatorname{Tr} \rho^2$, although we will also prove various statements on the von Neumann entropy $S_1 = -\operatorname{Tr} \rho \log \rho$ as well. The Rényi-2 entropy takes the elegant form $S_2 = \frac{1}{2} \log \det \sigma = \frac{1}{2} \operatorname{Tr} \log \sigma$, where σ is the covariance matrix of ρ [138]. For the Gaussian

state generated by the unitary $U \in \mathrm{U}(n)$ acting on the squeezed product state with squeezing strength s on each mode, the Rényi-2 entropy of the subsystem of k modes is denoted by $S_2(U)$, and the von Neumann entropy by $S_1(U)$, where the dependence of $S_1(U)$ and $S_2(U)$ on r, n, and s is implicit. We will derive statistical properties of $S_2(U)$ and $S_1(U)$ in the asymptotic limit $n \to \infty$.

Our main results involve the Rényi-2 entropy, and the following proposition allows us to use many of the results on the Rényi-2 entropy to bound the von Neumann entropy. Furthermore, we will also make use of the maximum Rényi-2 entropy to prove our later results on the Page correction.

Proposition 2.2.1. *Let $s \in \mathbb{R}$ and $r \in [0, 1]$. Then for each $j \in \{1, 2\}$,*

$$\max_{U \in \mathrm{U}(n)} S_j(U) = n \min(r, 1-r) h_j(\cosh(2s)), \tag{2.1}$$

where $h_1(x) = \frac{x+1}{2} \log \frac{x+1}{2} - \frac{x-1}{2} \log \frac{x-1}{2}$ and $h_2(x) = \log x$. Furthermore, for any fixed $U \in \mathrm{U}(n)$, $S_1(U) \geq S_2(U)$ and

$$S_1(U) < S_2(U) + n \min(r, 1-r)(1 - \log 2). \tag{2.2}$$

The full proof of Proposition 2.2.1 is given in Appendix A.2. A tighter version of Eq. (2.2) was originally derived in Ref. [136, Eq. 15], but we will only need this weaker version. For completeness, we provide a different proof of Eq. (2.2). Eq. (2.1) is perhaps implicit in various results in Refs. [34, 132, 138–140], but we have not found it directly stated anywhere. The lower bound $S_1(U) \geq S_2(U)$ is a general property of the Rényi entropies [13, 32], whereas the upper bound holds only for Gaussian states. While trivial upper bounds also exist in the general case, Eq. (2.2) is tighter. Intuitively one may view this as an extension of the fact that for Gaussian states, the Rényi-2 and von Neumann entropies share many useful properties, such as strong subadditivity and others mentioned in Section 2.1.

Section 2.3: Expectation value

Our first results concern the Rényi-2 Page curve, which is the expectation value of the Rényi-2 entanglement entropy as a function of the partition size ratio $r = k/n$ and the squeezing strength s. Recall that the dependence of $S_2(U)$ on r, n, and s is implicit. We find an exact formula as an infinite series for the Page curve in the limit $n \to \infty$.

Theorem 2.3.1 (Rényi-2 Page curve)**.** *Fix $s \in \mathbb{R}$ and $r \in [0,1]$. Let $C_\ell := \frac{1}{\ell+1}\binom{2\ell}{\ell}$ be the ℓ^{th} Catalan number, and let ${}_2F_1$ be the hypergeometric function [141–144]. Then*

$$\begin{aligned}
&\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \frac{1}{n} S_2(U) \\
&\quad = r \log\cosh(2s) - \sum_{\ell=1}^{\infty} r^{\ell+1} \tanh^{2\ell}(2s) \frac{C_\ell}{2\ell} \, {}_2F_1(1-\ell, \ell; \ell+2; r) \qquad (2.3) \\
&\quad = r \log\cosh(2s) + \sum_{d=2}^{\infty} r^d \sum_{\ell=\lceil d/2 \rceil}^{d-1} (-1)^{d-\ell} \tanh^{2\ell}(2s) \frac{(2\ell-1)!}{d(d-1)(d-\ell-1)!(2\ell-d)!\ell!}. \qquad (2.4)
\end{aligned}$$

This function is symmetric under $r \mapsto 1-r$, and hence the formula holds when r is replaced with $\min(r, 1-r)$. Furthermore, asymptotically in n, $\mathbb{E}_{U\in \mathrm{U}(n)} S_2(U) = n\alpha(s,r) - \lambda(s,r) + o(1)$, where $\alpha(s,r)$ is precisely $\lim_{n\to\infty} \mathbb{E}_{U\in \mathrm{U}(n)} \frac{1}{n} S_2(U)$ given above, $\lambda(s,r)$ is independent of n, and $o(1)$ denotes terms that go to zero as $n \to \infty$.

We plot the analytic Page curve for $s = 3/4$ in Fig. 2.1(a), and we confirm our results numerically in Fig. 2.1(b). The proof of Theorem 2.3.1, given in Appendix A.3.1, primarily uses two ingredients. The first ingredient is the asymptotic form of the Weingarten calculus for integrating over the unitary group with respect to the Haar measure [146, 147]. From this we get an equation for the Page curve that is initially daunting. The second main ingredient is the fact that the Page curve must be symmetric under $r \mapsto 1-r$ since the global state

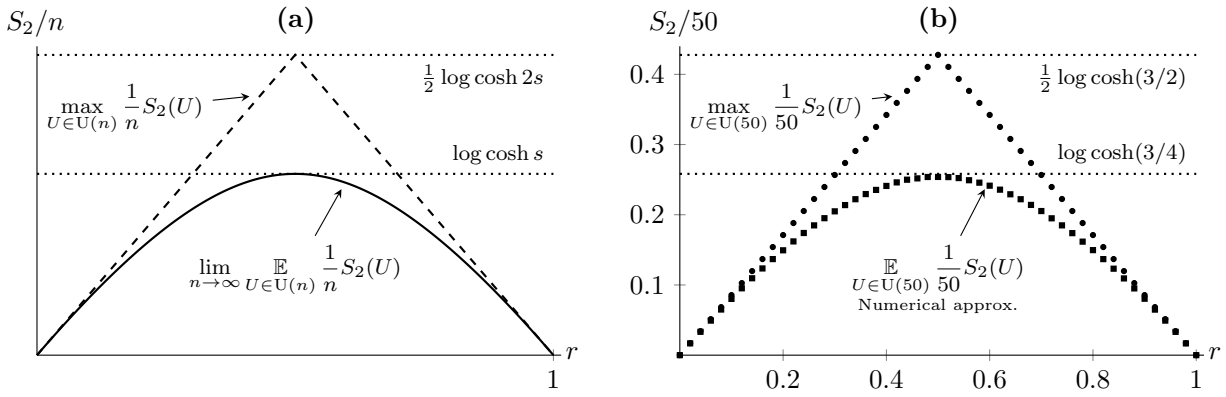

Figure 2.1: (a) Exact results for the Rényi-2 Page curve from Proposition 2.2.1, Theorem 2.3.1, and Corollary 2.3.2. (b) Numerical simulations of the Rényi-2 Page curve for $n = 50$ modes and squeezing $s = 3/4$. We plot the values for $r = k/50$ for each $k \in \{0, 1, \ldots, 50\}$. We perform the simulation by generating random unitary matrices and doing the matrix multiplication described in Appendix A.1.1. We provide the code for this simulation on GitHub [145].

on the n modes is pure [13]. Quite miraculously, this fact is enough to simplify the equation for the Page curve and arrive at Theorem 2.3.1.

The Page curve derived in Theorem 2.3.1 can be written as

$$\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in\mathrm{U}(n)} \frac{1}{n}S_2(U) = \sum_{\ell=1}^{\infty} \frac{\tanh^{2\ell}(2s)}{2\ell} G_\ell(r), \tag{2.5}$$

where $G_\ell(r) \coloneqq r - f_\ell(r)$ and $f_\ell(r)$ is a polynomial of degrees $\ell + 1$ through 2ℓ in r (given in Eq. (A.62) in Appendix A.3.1). Polynomials $G_\ell(r)$ of this form are uniquely determined by the requirement that $G_\ell(r) = G_\ell(1 - r)$, which ensures that the Rényi-2 entropy of a subsystem is equal to that of its complement since we are considering pure states. It is from this requirement that we ultimately derive the Page curve. We show that the resulting $G_\ell(r)$ can be understood as a good approximation to $m(r) \coloneqq \min(r, 1 - r)$ from below, which we will call the ℓ^{th} approximation. Indeed, the approximation is especially good near the endpoints $r = 0$ and $r = 1$, where the first ℓ derivatives of $G_\ell(r)$ match those of $m(r)$. As $\ell \to \infty$, the approximation becomes better and better such that $\lim_{\ell\to\infty} G_\ell(r) = m(r)$.

This provides an interpretation of the derived form of the Page curve. The strength of the squeezing s determines the weight that the Page curve has on the ℓ^{th} approximation to $m(r)$. For small squeezing, only low order approximations contribute, with the most dominant contribution being the parabolic shape $G_1(r) = r(1-r)$. When the squeezing is increased, there is more contribution from higher order approximations, giving the Page curve more of the triangle shape of $m(r)$. We see a manifestation of this interpretation as

$$\lim_{s\to 0}\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)}\frac{1}{s^2 n}S_2(U) = 2r(1-r), \tag{2.6}$$

$$\lim_{s\to \infty}\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)}\frac{1}{sn}S_2(U) = 2\min(r,1-r). \tag{2.7}$$

Meanwhile, from Eq. (2.1), the maximal Rényi-2 entropy is $\max_U \frac{1}{n}S_2(U) = m(r)\log\cosh(2s)$. As stated, near the endpoints $r = 0$ and $r = 1$, $G_\ell(r)$ is a very good approximation to $m(r)$. Thus, regardless of the squeezing strength, when the subsystem size $k = rn$ is small (or when its complement is small), the average entanglement is very close to maximal.

Unfortunately, we are unable to simplify the infinite sum for general r in Theorem 2.3.1 further. However, the Page curve can be fully simplified at $r = 1/2$ — which is where the maximum for a fixed s occurs — to $\log\cosh s$. Indeed, we find that

$$G_\ell(1/2) = \frac{1}{2}\left(1 - 4^{-\ell}\binom{2\ell}{\ell}\right). \tag{2.8}$$

From this and the maximum Rényi-2 entropy from Eq. (2.1), we also find the exact expression for the Page correction at $r = 1/2$ to be $\frac{1}{2}\log\left(1+\tanh^2 s\right)$. Let us formally state these results as follows.

Corollary 2.3.2. *For a fixed $s \in \mathbb{R}$ when $r = 1/2$,*

$$\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)}\frac{1}{n}S_2(U) = \log\cosh s, \tag{2.9}$$

and the Page correction is

$$\lim_{n\to\infty}\frac{1}{n}\left(\max_{U\in\mathrm{U}(n)} S_2(U) - \mathop{\mathbb{E}}_{U\in\mathrm{U}(n)} S_2(U)\right) = \frac{1}{2}\log\left(1+\tanh^2 s\right). \tag{2.10}$$

The proof of Corollary 2.3.2, given in Appendix A.3.2, is a straightforward consequence of a simplification of the hypergeometric function ${}_2F_1(a, 1-a; c; 1/2)$ in terms of gamma functions due to Bailey's theorem [141–144]. Altogether, we have derived the exact formula [Eqs. (2.3) and (2.4)] for the Rényi-2 page curve in the regime of equal squeezers as a series in $\tanh^2(2s)$ and r. In the special case when $r = 1/2$, we simplified the series to obtain an exact value of $\log\cosh s$. From this, we derived the Page correction, or information of the subsystem, at $r = 1/2$ to be exactly $\frac{1}{2}\log\left(1+\tanh^2 s\right)$. Furthermore, since the Page curve is concave in r while the maximum entropy is linear, this correction is maximized at $r = 1/2$. A summary of the Page curve results thus far is provided in Fig. 2.1.

We now shift our attention to the constant term in the Page curve. Theorem 2.3.1 states that asymptotically in n, $\mathbb{E}_{U\in\mathrm{U}(n)} S_2(U) = n\alpha(s,r) - \lambda(s,r) + o(1)$, and it provides the exact expression for $\alpha(s,r)$. We further find the following result for λ, which is proven in Appendix A.3.3.

Proposition 2.3.3. *Fix $s \in \mathbb{R}$ and $r \in [0,1]$. Then,*

$$\lambda(s,r) = -\frac{1}{8}\log\left(1 - 4r(1-r)\tanh^2(2s)\right). \tag{2.11}$$

Note that $\lambda(s,r) \geq 0$ for all s and r, and therefore $\mathbb{E}_U S_2(U) \leq n\alpha(s,r)$ asymptotically. At $r = 1/2$, this simplifies to $\lambda(s,1/2) = \frac{1}{4}\log\cosh(2s)$.

The proof of Proposition 2.3.3 is very similar to the proof of Theorem 2.3.1, again crucially using the symmetry $r \mapsto 1-r$ to simplify the expression coming from the Weingarten

function. However, there is one substantial difficulty in the proof of Proposition 2.3.3 that does not occur in that of Theorem 2.3.1. In this case, the symmetry $r \mapsto 1 - r$ is *almost* enough to fully determine λ; however, there are constants that cannot be determined by the symmetry alone. We derive expressions for these constants that are complicated sums involving permutations and Catalan numbers. Using objects that arise in bioinformatics when studying gene orders, called *breakpoint graphs* [148], Ref. [149] evaluates these sums, hence completing the proof of Proposition 2.3.3. These constants are discussed more in Appendix A.3.3.

In summary, we have derived an explicit form of the asymptotic Rényi-2 entropy Page curve for all r and s up to corrections that vanish as $n \to \infty$.

Section 2.4: Variance and typicality

Next, we shift our attention to the variance of the Rényi-2 entropy so as to make statements about typicality of entanglement; that is, how different a random state's entanglement is from the average. Using the results for the Rényi-2 entropy, we will also be able to prove some weaker results for the von Neumann entropy. The typicality results presented below are summarized in Table 2.1.

Typicality is of interest because it characterizes the applicability of statistical averages. Indeed, statistical mechanics often relies on quantities being typical so that thermodynamic average quantities, such as average energy and average pressure, can accurately represent their true values. In order to quantify the deviation from average, we consider two measures of deviation corresponding to multiplicative and additive distance. If the multiplicative distance between a quantity and its average vanishes in the thermodynamic limit, then that quantity is called *weakly typical*. If the additive distance vanishes in this limit, then that quantity is called *strongly typical*. With this intuition, we now formally define weak and strong typicality following Ref. [121].

Definition 2.4.1 (Typicality)**.** *Let S be a nonnegative random variable on the unitary group* $\mathrm{U}(n)$*, and denote its value at $U \in \mathrm{U}(n)$ by $S(U)$. S is called* weakly typical *if for any constant* $\epsilon > 0$*,*

$$\lim_{n\to\infty} \Pr_{U\in\mathrm{U}(n)}\left[\left|\frac{S(U)}{\mathbb{E}_{V\in\mathrm{U}(n)} S(V)} - 1\right| < \epsilon\right] = 1. \tag{2.12}$$

S is called strongly typical *if for any constant $\epsilon > 0$,*

$$\lim_{n\to\infty} \Pr_{U\in\mathrm{U}(n)}\left[\left|S(U) - \mathop{\mathbb{E}}_{V\in\mathrm{U}(n)} S(V)\right| < \epsilon\right] = 1. \tag{2.13}$$

		$k \in \Theta(n)$	$k \in o(n)$	$k \in o\left(n^{1/3}\right)$ [33]
Equal squeezing	*Rényi-2*	weak	strong	strong
	von Neumann	?	weak	strong
Unequal squeezing	*Rényi-2*	?	weak*	strong
	von Neumann	?	weak*	strong

Table 2.1: A summary of the current status of rigorous results on typical entanglement in Gaussian bosonic systems, where in this figure we assume single-mode squeezing parameters that are independent of n. Strong and weak typicality are defined in Definition 2.4.1. Note that "weak*" indicates that the result is not fully proven, but depends on Conjecture 2.5.2. Where we say "weak", we have not ruled out the possibility that the typicality is also strong. The total number of modes is denoted by n, and $0 \leq k \leq n$ is the number of modes in the subsystem. "Equal squeezing" refers to the case when each mode is initially squeezed with the same strength, whereas "unequal squeezing" refers to the general case when each mode can be squeezed independently. The two leftmost columns come from Corollaries 2.4.3 and 2.4.4 and Remark 2.5.3. The rightmost column all follows from the results of Ref. [33]. Prior to Ref. [33], Refs. [34, 132] proved strong typicality in the regime $k \in \mathcal{O}(1)$.

If $\mathbb{E}_{U\in\mathrm{U}(n)} S(U)$ does not decay as $n \to \infty$, then strong typicality clearly implies weak typicality. We will be concerned with typicality of S_1 and S_2, the von Neumann and Rényi-2 entropy respectively. One can compute the variance of the entropy over the Haar measure and then apply Chebyshev's inequality to obtain typicality results. Therefore, we now focus on the variance. We first find the general form of the asymptotic Rényi-2 variance in the equal squeezing regime and show that it is independent of n. Recall that the dependence of

$S_2(U)$ on r, n, and s is implicit.

Theorem 2.4.2 (Rényi-2 variance)**.** *Let $s \in \mathbb{R}$ and $r \in [0,1]$. Then*

$$\lim_{n\to\infty} \operatorname*{Var}_{U\in \mathrm{U}(n)} S_2(U) = \sum_{d=2}^{\infty} \omega^{(d)} \tanh^{2d}(2s) \left(r(1-r)\right)^d , \tag{2.14}$$

where $\omega^{(d)} \in \mathbb{Q}$ is some number that depends only on d. In particular, $\omega^{(2)} = 1/2$.

The proof of this theorem, given in Appendix A.4, is very similar to the proof of Theorem 2.3.1. To prove this theorem, we again crucially use that $\mathrm{Var}_U\, S_2(U)$ must be symmetric under $r \mapsto 1-r$ since the full state on the n modes is pure. Interestingly, in contrast to Theorem 2.3.1 where we found that the Page curve grows linearly with n, the asymptotic variance is independent of n. Indeed, Theorem 2.4.2 is the bosonic analogue of the result that the variance for fermionic Gaussian states is asymptotically constant [120].

From Theorem 2.4.2, we find typicality in certain regimes as an immediate corollary. In particular, since the variance is independent of n while the average grows with n, the entanglement is always weakly typical. Furthermore, the variance is asymptotically zero if $s \in o(1)$ and/or $r \in o(1)$, and typicality is therefore strong in those regimes. Altogether, Theorem 2.4.2 and Chebyshev's inequality lead to the following corollary.

Corollary 2.4.3. *The Rényi-2 entropy is weakly typical for any $s \in \mathbb{R}$ and $r \in [0,1]$. Furthermore, if the subsystem size $k = rn$ scales as $k \in o(n)$, or if the squeezing scales as $s \in o(1)$, then the Rényi-2 entropy is strongly typical, and $\Pr_{U\in \mathrm{U}(n)}[|S_2(U) - \mathbb{E}_V\, S_2(V)| < \epsilon] \geq 1 - \frac{\Omega(1/n)}{\epsilon^2}$. On the other hand, if r and s are constant in n, then $\lim_{n\to\infty} \Pr_{U\in \mathrm{U}(n)}[|S_2(U) - \mathbb{E}_V\, S_2(V)| < \epsilon] \geq 1 - \frac{c(r,s)}{\epsilon^2}$, where $c(r,s)$ is a constant independent of n (but depends on r and s).*

In summary, we have shown weak typicality in the Rényi-2 entropy for all s and k and strong typicality whenever $k \in o(n)$. On the other hand, if s and r do not tend to zero with

increasing n, the variance converges to a constant value independent of n. Hence, we cannot use Chebyshev's inequality to prove strong typicality in this case. Notably, an asymptotically constant variance does not necessarily imply an absence of strong typicality either — that is, the probability that the entropy deviates from its average can scale as $1/n^2$, but the entropy can deviate by an amount proportional to n thus resulting in a constant variance. We are therefore unable to make a definitive statement about strong typicality in the $r \in \Theta(1)$ and $s \in \Omega(1)$ case.

In the next corollary, we will use these results to address typicality as measured by the von Neumann entropy. Specifically, we will use Proposition 2.2.1 to show weak typicality of the von Neumann entropy as long as the subsystem size scales sublinearly with the system size.

Corollary 2.4.4. *Let $s \in \mathbb{R}$ and $r \in [0,1]$. If the subsystem size $k = rn$ scales as $k \in o(n)$, then the von Neumann entropy is weakly typical.*

Proof. Using Proposition 2.2.1 to upper bound $\mathbb{E}_U \frac{1}{n^2} S_1(U)^2$ and to lower bound $\left(\mathbb{E}_U \frac{1}{n} S_1(U)\right)^2$, we find that

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n^2} \operatorname*{Var}_{U\in \mathrm{U}(n)} S_1(U) < \lim_{n\to\infty} \frac{1}{n^2} \operatorname*{Var}_{U\in \mathrm{U}(n)} S_2(U) \\ + 2r \log(\mathrm{e}/2) \lim_{n\to\infty} \frac{1}{n} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} S_2(U) \\ + r^2 \log^2(\mathrm{e}/2). \end{aligned} \tag{2.15}$$

From Theorem 2.3.1, the first term in the right-hand side is always zero. Furthermore, from Theorem 2.4.2, the second term is $\mathcal{O}(r^2)$. Hence, $\operatorname{Var}_U S_1(U)/n^2 \in \mathcal{O}(r^2)$, which is zero as $n \to \infty$ when $r \in o(1)$. □

We again emphasize that our typicality results thus far, which are summarized in Table 2.1, only apply to the case when the initial squeezing strength on each mode is the same. On the other hand, Ref. [33] proves strong typicality when $k \in o\left(n^{1/3}\right)$ for both the von Neumann and Rényi-2 entropies in the general case when squeezing strengths can be different

on different modes. It is for this reason that our results do not supplant those of Ref. [33] in general, but rather only in certain regimes. To the best of our knowledge, our results are the first to address the $k \in \Omega\big(n^{1/3}\big)$ regime.

Ultimately, we would like to determine in exactly which regimes strong and weak typicality occur and do not occur. Our results thus far almost complete the story for the regime of equal squeezing when typicality is measured with the Rényi-2 entropy, since we have proven strong typicality when $k \in o(n)$ and weak when $k \in \Theta(n)$; the missing piece is whether or not typicality is strong when $k \in \Theta(n)$. However, the story is even more incomplete for the regime of equal squeezing when typicality is measured with the von Neumann entropy, though we made some progress by proving that typicality is at least weak whenever $k \in o(n)$. In this regime, the best known result for strong typicality is when $k \in o\big(n^{1/3}\big)$ as proven in Ref. [33]. Indeed, this is also currently the best known result in the regime of unequal squeezing. In Section 2.5, we will use our results on equal squeezing typicality to add to the story for unequal squeezing.

Section 2.5: Generalizing to unequal squeezing

So far, we have considered the restricted setting where each mode i is initially squeezed with strength $s_i = s$ for some $s \in \mathbb{R}$. We now generalize by allowing the squeezing strengths to be different on each mode. As such, for the remainder of this section, the squeezings will be $(s_1, \ldots, s_n)$, where each $s_i \in \mathbb{R}$, and $s_{\max}$ and $s_{\min}$ are defined as $s_{\max} \coloneqq \max_i |s_i|$ and $s_{\min} \coloneqq \min_i |s_i|$.

To begin, we focus on the Rényi-2 Page curve in the regime of unequal squeezing. When squeezing is small, we can utilize the equal squeezing Page curve in Theorem 2.3.1 to compute the Page curve for unequal squeezing. Recall that the dependence of $S_2(U)$ on r, n, and s is implicit.

Corollary 2.5.1. *Let each $s_i \in \mathbb{R}$ and $r \in [0,1]$. Then, as $n \to \infty$,*

$$\mathop{\mathbb{E}}_{U \in \mathrm{U}(n)} S_2(U) = 2r(1-r) \sum_{i=1}^{n} s_i^2 + \mathcal{O}\big(rns_{\max}^4\big) . \tag{2.16}$$

Proof. The Rényi-2 entropy of a Gaussian state with covariance matrix σ is proportional to $\log\det\sigma$. log is real analytic and $\det\sigma$ is analytic in the parameters of the initial covariance matrix σ_0. Hence, S_2 can be written as a power series in s_i. There exists a passive Gaussian unitary, specifically a product of one-mode phase shifters, that acts on σ_0 via the transformation $s_i \to -s_i$. Hence, by the translational invariance of the Haar measure — that is, the invariance of the Haar measure under the application of any fixed unitary — the power series must be an expansion in s_i^2. Furthermore, there exists a passive Gaussian unitary, specifically a product of beamsplitters, that acts on σ_0 via the transformation $s_i \to s_{\tau(i)}$ for some permutation τ of n elements. Therefore, again by the translational invariance of the Haar measure, the power series must be symmetric under $s_i \to s_{\tau(i)}$. It follows that the $\mathcal{O}(s_i^2)$ term must be of the form $g(r,n)\sum_i s_i^2$ for some function $g(r,n)$. When all s_i are equal, the power series must reduce to Theorem 2.3.1, which fixes $g(r,n)$ to be $2r(1-r)$.

The next term is $\mathcal{O}(s_{\max}^4)$, but what is the k dependence? We will show that the k dependence is at most linear, proving that the remaining terms in the power series are $\mathcal{O}(ks_{\max}^4)$. Recall that $S_2(U) = \frac{1}{2}\operatorname{Tr}\log\sigma(U)$ where $\sigma(U)$ is the covariance matrix for the state generated by U from the initial product squeezed state σ_0. Since the log function is concave, Jensen's inequality implies that $\mathbb{E}_U S_2(U) \le \frac{1}{2}\operatorname{Tr}\log\mathbb{E}_U\sigma(U)$. $\mathbb{E}_U\sigma(U)$ is calculated in Eq. (A.5) in Appendix A.1.1 to be $\nu\mathbb{I}_{2k\times 2k}$, where $\nu = \frac{1}{n}\sum_i \cosh(2s_i) \le \cosh(2s_{\max})$. Therefore, $\mathbb{E}_U S_2(U) \in \mathcal{O}(k\log\nu)$. □

One particularly interesting application of this corollary is when each $s_i \in \mathcal{O}(1/\sqrt{n})$. In this case, the average total number of bosons in the n modes is $N = \sum_i \sinh^2(s_i) \in \mathcal{O}(1)$. Thus, when one considers a constant number N of bosons in the system as the number of

modes is taken to infinity, one finds the Page curve to be $2r(1-r)N$.

We now shift our focus to entanglement typicality in the regime of unequal squeezing. The results described in the remainder of this section are summarized in Table 2.1. Ideally, we would like to make further statements about entanglement in the regime of unequal squeezing by using our previous results on equal squeezing. One way to potentially proceed is to use the equal squeezing results to bound the unequal squeezing quantities. Intuitively, we expect that $\mathbb{E}_U S_2(U)$ for arbitrary unequal squeezing is upper bounded by $\mathbb{E}_U S_2(U)|_{\text{all } s_i = s_{\max}}$ and lower bounded by $\mathbb{E}_U S_2(U)|_{\text{all } s_i = s_{\min}}$. In other words, by *increasing* all of the squeezing strengths until they all are equal, the average entanglement will *increase*. In this spirit, we make the following conjecture.

Conjecture 2.5.2. *Let $r \in [0,1]$, $s_i \in \mathbb{R}$, and $n \in \mathbb{N}$. Then, for any i,*

$$0 \leq \frac{\partial}{\partial(s_i^2)} \left(\mathop{\mathbb{E}}_{U \in \mathrm{U}(n)} \frac{1}{n} S_2(U) \right), \quad \textit{and} \tag{2.17}$$

$$0 \leq \frac{\partial}{\partial(s_i^2)} \left(\mathop{\mathbb{E}}_{U \in \mathrm{U}(n)} \frac{1}{n^2} S_2(U)^2 \right). \tag{2.18}$$

Conjecture 2.5.2 seems intuitive — by increasing the magnitude of any individual squeezing strength, the number of bosons in the system increases, and therefore it would seem surprising for the average entanglement to decrease. We note that, somewhat counterintuitively, one can find explicit unitaries and squeezing configurations for which $\frac{\partial}{\partial(s_i^2)} S_2(U) < 0$ (we provide an example in our code repository [145]), and hence the presence of the $\mathbb{E}_U$ is necessary for the conjecture. Despite its intuitiveness, we have been unable to rigorously prove Conjecture 2.5.2. The derivative of $S_2(U) = \frac{1}{2} \log\det \sigma(U)$ is $\frac{1}{2} \operatorname{Tr}\left(\sigma(U)^{-1} \frac{\partial \sigma(U)}{\partial (s_i)^2} \right)$. The difficulty in computing the expectation value over $U \in \mathrm{U}(n)$ arises due to the presence of the inverse $\sigma(U)^{-1}$.

Nonetheless, under the assumption that Conjecture 2.5.2 is true, we can immediately upper and lower bound the Rényi-2 Page curve for an arbitrary squeezing configuration

$(s_1, \ldots, s_n)$ by using Theorem 2.3.1 with $s = s_{\max}$ and $s = s_{\min}$ respectively,

$$\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \frac{1}{n} S_2(U) \geq \lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \frac{1}{n} S_2(U)\big|_{s_i=s_{\min}}, \tag{2.19}$$

$$\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \frac{1}{n} S_2(U) \leq \lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \frac{1}{n} S_2(U)\big|_{s_i=s_{\max}}. \tag{2.20}$$

Furthermore, the conjecture implies that

$$\begin{aligned} &\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \frac{1}{n^2} S_2(U)^2 \\ &\qquad \leq \lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \frac{1}{n^2} S_2(U)^2\big|_{s_i=s_{\max}}. \end{aligned} \tag{2.21}$$

From this, we can also make statements on weak typicality for unequal squeezers by bounding the variance. The variance is $\mathbb{E}_U S_2(U)^2 - (\mathbb{E}_U S_2(U))^2$. We can therefore upper bound the variance by upper bounding the first term with Eq. (2.21) and lower bounding the second term with Eq. (2.19).

Remark 2.5.3. Let $r \in [0, 1]$ and $s_i \in \mathbb{R}$. If Conjecture 2.5.2 is true and if the subsystem size $k = rn \in o(n)$ or $s_{\max} \in o(1)$, then the Rényi-2 entropy is weakly typical. Similarly, if $k \in o(n)$, then the von Neumann entropy is weakly typical.

Proof. We can bound the variance of the Rényi-2 entropy as

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n^2} \operatorname*{Var}_U S_2(U) \leq \lim_{n\to\infty} \frac{1}{n^2} \\ \times \left[\mathop{\mathbb{E}}_U \left(\frac{1}{n} S_2(U)|_{\text{all } s_i=s_{\max}} \right)^2 \right. \\ \left. - \left(\mathop{\mathbb{E}}_U \frac{1}{n} S_2(U)|_{\text{all } s_i=s_{\min}} \right)^2 \right]. \end{aligned} \tag{2.22}$$

From Theorem 2.4.2, the $\mathbb{E}_U$ can be brought inside the parentheses in the first term. Then, the right hand side can be computed using Theorem 2.3.1, which gives $\lim_{n\to\infty} \frac{1}{n^2} \operatorname{Var}_U S_2(U) \in \mathcal{O}(r^2 s_{\max}^4)$. Hence S_2 is weakly typical if $r \in o(1)$ or $s_{\max} \in o(1)$.

For the von Neumann entropy, we again make use of Eq. (2.15), which gives $\lim_{n\to\infty} \frac{1}{n^2} \operatorname{Var}_U S_1(U) \in \mathcal{O}(r^2 s_{\max}^2)$, which goes to zero if $r \in o(1)$. □

In summary, we have used our results from Sections 2.3 and 2.4 on entanglement in the equal squeezing regime to prove various statements in the unequal squeezing regime. In particular, under the assumption that Conjecture 2.5.2 is true, we prove both the Rényi-2 and von Neumann entropies are weakly typical whenever the subsystem size k scales as $k \in o(n)$. The best known result for the presence of strong typicality in the unequal squeezing regime is when $k \in o\big(n^{1/3}\big)$ as proven in Ref. [33]. The current status of rigorous results on typical entanglement in Gaussian bosonic systems is summarized in Table 2.1.

Section 2.6: Conclusion

In this work, we studied the average and variance of the Rényi-2 and von Neumann entropies in random bosonic Gaussian systems. We computed the Rényi-2 Page curve and Page correction when all the initial squeezing strengths are equal, and we proved various results on the typicality of the Rényi-2 and von Neumann measures of entanglement. Given that the Rényi-2 entropy is a function of only the purity, it is often tractable to measure experimentally. It would be interesting to compare the analytic formula in Theorem 2.3.1 to an experimental Gaussian Boson Sampling device to determine how well it is generating and maintaining bipartite entanglement.

We have identified several open problems that would generalize and expand our results. One such open problem is to prove Conjecture 2.5.2, which would allow our results on the Page curve to apply more generally. Perhaps the most important remaining task is to complete Table 2.1 by proving typicality of entanglement in the remaining regimes, such as the regime of unequal squeezing and the von Neumann entropy.

For the latter, we note two potentially fruitful avenues. The first comes from the formula

for the von Neumann entropy of a Gaussian state with covariance matrix σ given in Ref. [150]. Let $D_\sigma \coloneqq \sqrt{\det \sigma} = \mathrm{e}^{S_2}$, where S_2 is the Rényi-2 entropy of σ. Ref. [150] derives expressions for all the Rényi-α entropies, including the von Neumann entropy, as functions of D_σ. In our work, we found an expression for $\mathbb{E}_U\, S_2(U)$ in the regime of equal squeezers by expanding $S_2(U)$ in a power series and exactly computing asymptotic expectation values over the unitary Haar measure. To find the Rényi-α entropy $\mathbb{E}_U\, S_\alpha(U)$, one could similarly attempt to expand in powers of D_σ, and therefore compute $\mathbb{E}_U\, S_\alpha(U)$ by computing $\mathbb{E}_U\, \mathrm{e}^{jS_2(U)}$ for various values of j. A second potential way of computing $\mathbb{E}_U\, S_1(U)$ is similar, where one could use the formula

$$S_1 = \frac{1}{2}\log\det\left[\frac{\sigma + \mathrm{i}\Omega}{2}\right] + \frac{1}{2}\operatorname{Tr}[\operatorname{arccoth}(\mathrm{i}\Omega\sigma)\mathrm{i}\Omega\sigma]\,, \tag{2.23}$$

where Ω is the $2n \times 2n$ symplectic form given in Eq. (A.1) [138, 151, 152]. The equations resulting from using these methods with the Weingarten calculus may potentially be too difficult to simplify at first glance, as was the case in this work. However, it would be interesting to see if the presence of the $r \mapsto 1 - r$ symmetry is enough, as it was in this work, to reduce the complicated Weingarten expressions to something much more tractable and simple.

In this paper, we have only considered truly Haar-random unitaries. One important question concerns how these results translate to the case where one uses random local passive Gaussian gates to generate random unitary circuits of finite depth. Indeed, an interesting open problem is to determine the depth dependence of entanglement, sampling complexity, and gate complexity in linear optical circuits. Sampling complexity refers to the classical complexity of generating samples from the output probability distribution defined by a fixed depth linear optical circuit, and gate complexity refers to the minimum number of nearest-neighbor beamsplitters required to generate the probability distribution. Currently, the precise relationship between entanglement and complexity is largely unknown. Numerical analyses of entanglement dynamics in linear optical circuits have been reported in Refs. [153,

154]. Partial analytical work was done in Ref. [155], but only in the regime where one or a small number of modes are not initially vacuum. On the contrary, a typical Gaussian Boson Sampling experiment initially squeezes many or all of the modes. Information on such entanglement growth could yield insights on implementations of Gaussian Boson Sampling experiments as well as the complexity of computing output probabilities from such experiments. On the complexity side, many recent works have studied classical simulation and classical sampling complexity of linear optical circuits in certain regimes of low depth and the phase transition at which the complexity passes from easy to hard [156–160]. Indeed, both entanglement and complexity are expected grow with depth, and further study may reveal that the relationship is even more intimate.

In this work, we have characterized the entanglement properties of Gaussian states such as they arise in Gaussian Boson Sampling. In this setting, we also know that sampling from Fock basis measurements of the Gaussian state is computationally intractable. It remains an exciting question to better understand the role that entanglement plays in this context. An important aspect of this direction is to understand how entanglement and measurement bases interact. After all, some form of non-Gaussianity is crucial to generate complexity in bosonic computations [124].

Chapter 3: Average Rényi entanglement entropy in Gaussian boson sampling

Section 3.1: Introduction

Major advances in quantum computing over the past decade have opened up potential breakthroughs in various fields of science [161, 162]. One such advancement is Gaussian Boson Sampling (GBS), an experimental framework for demonstrating quantum advantage that produces a sample from a distribution that is classically hard to compute [24–26, 163]. In this framework, squeezed photons are transformed by a Haar-random, passive Gaussian unitary consisting of beamsplitters and phase shifters. Although GBS is not computationally universal, it is nonetheless crucial for understanding the leverage of quantum devices over their classical counterparts at performing specific tasks.

The preparation of squeezed states is a much easier task than that of single photons used in Fock-state boson sampling, hence making GBS an experimentally favorable avenue toward demonstrating quantum advantage. Indeed, such demonstrations have already been realized in the lab [29–31].

There is strong evidence that classically sampling from the output probability distribution of a GBS experiment is hard in the complexity-theoretic sense [27, 29–31, 164], whereas such sampling can be done efficiently on a quantum device. Moreover, entanglement is a critical asset across quantum computing, with additional applications in teleportation, dense coding, and quantum communication [123, 165–167]. Therefore, quantifying the behavior of

entanglement in Gaussian bosonic systems has been of recent interest [29, 31], especially to sharpen the connection between entanglement and computational complexity. Indeed, the relationship between entanglement and complexity is generally unknown. Classical hardness requires at least some entanglement, but too much entanglement can also cause a state to be useless for computation. In fact, a randomly chosen finite-dimensional quantum state generally has too much entanglement to be computationally useful [21, 22]. This prompts our study of average and typical entanglement, and motivates further study into the general connection between entanglement and complexity.

In 1993, Page conjectured an explicit formula for the average entanglement entropy over all Haar-random pure states in an N-dimensional Hilbert space of the reduced density matrix on an rN-dimensional subsystem for $r \in [0, 1]$. This conjecture was proven soon after in the $N \to \infty$ limit, and the average entanglement entropy as a function of r is now known as the *Page curve* [99–101]. The Page curve was initially studied in the context of information dynamics in and around black holes [168], and has since found various applications in quantum information and condensed matter theory [169]. As a natural follow-up, many works investigated the deviation of entanglement entropy from its expected value. A system exhibits *typical entanglement* if the probability that a random state has entanglement bounded away from the average is small. This behavior, along with strong and weak classes of entanglement typicality, has been studied in black holes, thermodynamics, and other systems [65, 105, 118, 121, 169–171].

Leaving the realm of standard finite-dimensional systems, the Page curve for fermionic Gaussian states has been extensively analyzed [120, 131, 172–176]. A natural next step, especially given the relevance to GBS, is to study bosonic Gaussian states. The typicality of entanglement in random bosonic Gaussian states was studied in Refs. [2, 33, 34, 132], and the Page curve for Gaussian bosonic systems was studied in Ref. [2]. However, Ref. [2] only computed the Page curve for the Rényi-2 entropy.

In this work, we derive analytic expressions for the Rényi-α Page curves of bosonic Gaussian systems for all $\alpha \in \mathbb{Z}_{>0}$, with special considerations for the von Neumann (Rényi-1) Page curve. The entropy of a mixed state is an indicator of its entanglement with the rest of the universe [13], and indeed the Rényi-α Page curve is a measure of the entanglement of an average state in a system as measured by the Rényi-α entropy, for $\alpha \in \mathbb{R}^+$. The Rényi-α entropies offer a quantitative way to compare the amount of entanglement in two given quantum states, in that if the Rényi-α entanglement entropy of state $|\psi_1\rangle$ is greater than $|\psi_2\rangle$ for any given α, then one cannot convert $|\psi_2\rangle$ into $|\psi_1\rangle$ via local operations and classical communication. Each Rényi-α entropy has its own unique mathematical features, and notable examples are the von Neumann/Shannon entropy ($\alpha = 1$), collision entropy ($\alpha = 2$), and min-entropy ($\alpha \to \infty$). In particular, the min-entropy, the entropy defined by $-\log\max(p_i)$ for each p_i a probability of occurrence of outcome i when working with a discrete random variable, is a useful expression for working with extracting randomness from variables when lower entropies (e.g. von Neumann) do not suffice [177]. Therefore, it is natural to analyze the Rényi-α entropies in bosonic systems and extract their unique properties.

Our expressions are asymptotically exact in the number of modes n. Using these expressions, we prove various results about typicality of entanglement with respect to these Rényi-α entropies. A surprising feature that our analysis reveals is that Rényi-α entropies all concentrate for integer α, which suggests that the entanglement spectrum of the states generated by Gaussian boson sampling also concentrate. This result could be useful in theoretical analysis of the performance of Gaussian boson sampling circuits.

Our setup is exactly that of a GBS experiment, and is the same as the setups considered in Refs. [2, 33]. A pure input quantum state with n squeezed vacuum modes is acted on by a Haar-random, linear unitary (passive Gaussian) circuit. The Page curve, as a function of r, is the average Rényi-α entanglement between a group of rn modes and the remaining system

of $(1-r)n$ modes, where $r \in [0,1]$. We expound our setup in greater detail in Section 3.2.

In Section 3.3, we present our results for the von Neumann entropy. Under an equal squeezing regime, we first present an explicit formula for the von Neumann Page curve in Section 3.3.1 in terms of the partition ratio r and initial squeezing strength s, which is exact up to terms that decay to zero as $n \to \infty$. We subsequently discuss the von Neumann entropy under squeezing limits $s \to 0$ and $s \to \infty$ in Section 3.3.2.

We then generalize our analyses to the Rényi-α entropies for integers $\alpha \geq 2$ in Section 3.4. In Section 3.4.1, we derive the Rényi-α Page curve in terms of the partition ratio r and initial squeezing strength s, with additional terms that vanish to zero as $n \to \infty$. We calculate squeezing limits of the Rényi-α entropy, this time with the extra dependence on α, in Section 3.4.2. Finally, we evaluate the Rényi-α entropy under unequal, small initial squeezing strengths in Section 3.4.3 and derive typicality of entanglement results in Section 3.4.4.

Section 3.2: Setup

In this section, we present our setup. We consider a bosonic system of n modes. For a pedagogical introduction to the concepts discussed in the section, we refer to Ref. [138]. For each $i \in \{1, 2, \ldots, n\}$, define $r_i \coloneqq x_i$ to be the position operator and $r_{n+i} \coloneqq p_i$ to be the momentum operator. A density matrix ρ describes a Gaussian state if there exists a $\beta > 0$ and a Hamiltonian H that is at most quadratic in the quadrature operators r_i such that ρ is the thermal state $\rho \propto \mathrm{e}^{-\beta H}$. When ρ is Gaussian, it is fully characterized by its first and second moments $\mathrm{Tr}(\rho r_i)$ and $\sigma_{ij} = \frac{1}{2}\mathrm{Tr}[\rho(r_i r_j + r_j r_i)] - \mathrm{Tr}(\rho r_i)\mathrm{Tr}(\rho r_j)$. σ is called the *covariance matrix* of the state ρ.

In this work, we investigate the properties of bosonic states after the application of a linear optical unitary. Linear optical unitaries generate passive (boson-number conserving, or equivalently energy conserving) Gaussian unitaries. The set of passive Gaussian unitaries is isomorphic to $\mathrm{Sp}(2n) \cap O(2n) \cong \mathrm{U}(n)$, where $\mathrm{Sp}(2n)$ is the *symplectic group* of $2n \times 2n$

matrices, $O(2n)$ is the *orthogonal group* of $2n \times 2n$ matrices, and $\mathrm{U}(n)$ is the *unitary group* of $n \times n$ matrices [71].

We now define our notion of a random, pure Gaussian state on n modes. We initialize the i^{th} mode to be in a squeezed vacuum state with fixed squeezing parameter $s_i \in \mathbb{R}\ \forall\, i \in \{1, \ldots, n\}$. We then randomly sample a passive Gaussian unitary $U \in \mathrm{U}(n)$ according to the Haar measure [71] and apply it to the n modes. The output state is characterized by the input squeezing strengths s_i and the unitary U. For squeezing parameters $s_1, s_2, \ldots, s_n$, the total expected number of bosons in the state on the n modes is $\sum_{i=1}^{n} \sinh^2(s_i)$. In this work, we primarily consider the case when all the squeezing parameters are equal, $s_i = s$. Then, the average total number of bosons per mode is $\sinh^2(s)$.

We partition the n output modes into two groups: one group of $k = rn$ modes for some $0 \leq r \leq 1$, and the other group of $n - k = (1 - r)n$ modes. We calculate the Rényi-α entropy of the reduced state of the k modes. Because we are considering pure states, this is equivalent to the entropy of the reduced state of the $n - k$ modes [13].

For the density matrix $\rho(U)$ on the k modes, we denote its Rényi-α entropy by $S_\alpha(U) \coloneqq \frac{1}{1-\alpha} \log \operatorname{Tr}(\rho(U)^\alpha)$. The limit as $\alpha \to 1$ yields the von Neumman entropy $S_1(U) \coloneqq -\operatorname{Tr} \rho(U) \log \rho(U)$. We note that the dependence of these quantities on r, n, and s is implicit. Furthermore, the average value of $S_\alpha(U)$ over $\mathrm{U}(n)$ with respect to the Haar measure is defined as the Rényi-α Page curve.

We will frequently utilize $\sigma(U)$, the covariance matrix of $\rho(U)$. We also use the *symplectic eigenvalues* of $\sigma(U)$, defined as the eigenvalues of the matrix $i\Omega\sigma(U)$ where

$$\Omega \coloneqq \begin{pmatrix} 0_{n\times n} & \mathbb{I}_{n\times n} \\ -\mathbb{I}_{n\times n} & 0_{n\times n} \end{pmatrix}. \tag{3.1}$$

Section 3.3: Von Neumann Entropy

In this section, we calculate an explicit formula for the von Neumann Page curve as a function of equal squeezing strength s and the partition size ratio r up to terms that decay to zero as n increases. Then, we compute and interpret the Page curve in the squeezing limits $s \to 0$ and $s \to \infty$. Next, we discuss an approach to the case of small, unequal squeezing strengths. Finally, we comment on difficulties in deriving the variance and entanglement typicality of the von Neumann entropy in this setting.

Subsection 3.3.1: Explicit formula

The von Neumann entropy is a function of the squeezing strength s (assuming all modes are equally squeezed) and the partition size ratio $r = \frac{k}{n}$, up to terms that decay to zero as $n \to \infty$. We explicitly calculate this function, which also leads us to the von Neumann Page curve as $n \to \infty$.

Theorem 3.3.1 (von Neumann entropy)**.** *Choose any $s \in \mathbb{R}$ and $r \in [0,1]$. Let $C_i = \frac{1}{i+1}\binom{2i}{i}$ represent the ith Catalan number and ${}_2F_1$ the hypergeometric function [141]. Then, the average von Neumann entropy over all unitaries $U \in U(n)$ is*

$$\begin{aligned} \mathop{\mathbb{E}}_{U\in U(n)} S_1(U) = \sum_{i=1}^{\infty} \left[\frac{1}{2i} - \frac{1}{3}\operatorname{sech}^2(2s)\tanh^{2i}(2s)\,{}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, \operatorname{sech}^2(2s)\right)\right] \\ \times (nG_i(r) - H_i(r) + o(1)), \end{aligned} \tag{3.2}$$

where

$$G_i(r) = r - r^{i+1}C_i\,{}_2F_1(1-i, i, 2+i, r), \tag{3.3}$$

$$H_i(r) = 4^{i-1}(r(1-r))^i \tag{3.4}$$

are functions independent of n detailed in Ref. [2], and $o(1)$ denotes terms that vanish as $n \to \infty$.

The derivation for this formula is detailed in Appendix B.1.1. In the derivation we use a formula for the von Neumann entropy in terms of the covariance matrix of the subsystem of k modes. Using the decomposition of the covariance matrix derived in Ref. [2], we Taylor expand about the identity matrix and utilize the negative binomial theorem. Finally, we use formulas from Ref. [2] regarding expectation values over $\mathrm{U}(n)$ of certain functions of matrix elements to arrive at the final expression.

The von Neumann Page curve as $n \to \infty$ is derived from Theorem 3.3.1 by dividing by n. The $H_i(r)$ and $o(1)$ terms on the RHS subsequently vanish. One important note is that the Page curve must be symmetric under $r \mapsto 1 - r$ since the global state on the n modes is pure [13]. Since $G_i(r)$ and $H_i(r)$ are symmetric about $r = \frac{1}{2}$ [2], the von Neumann Page curve derived from Theorem 3.3.1 is indeed symmetric about $r = \frac{1}{2}$. This symmetry allows one to replace r with $\min(r, 1 - r)$.

Subsection 3.3.2: Squeezing limits

Theorem 3.3.1 applies for all equal squeezing strengths s. We find considerable simplifications when examining the limiting behavior as a function of s. We begin with the $s \to 0$ limit, which corresponds to photons in their near-vacuum states. One might expect the von Neumann Page curve to grow with the total number of bosons, $\sum_{i=1}^{n} \sinh s^2 \approx n s^2$ for small s_i. Indeed, this behavior occurs for the rest of the Rényi entropies (with integer $\alpha \geq 2$) as shown in Section 3.4.2. However, the pertinent limit $\lim_{s\to 0} \lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{s^2 n} S_1(U)$ diverges to infinity. In Appendix B.1.2, we prove this result and demonstrate the relevant scaling to be $\mathop{\mathbb{E}}_{U} S_1(U) \sim s^2 \log(1/s^2)$, as formulated in the following theorem.

Theorem 3.3.2. *For the squeezing strength limit $s \to 0$,*

$$\lim_{s\to 0}\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in U(n)}\frac{1}{(s^2\log(1/s^2))\,n}S_1(U) = r(1-r). \tag{3.5}$$

We note that of course the entanglement must be zero when $s = 0$ simply because there are no bosons present when $s = 0$. Theorem 3.3.2 addresses how the entanglement approaches zero as $s \to 0$. While the derivation of Theorem 3.3.2 is fairly straightforward, its significance is more interesting. The von Neumann Page curve's growth in order $s^2\log(1/s^2)$ for small s can be attributed to the unique definition of the von Neumann entropy $S_1(\rho) = -\operatorname{Tr}(\rho\log\rho)$, where ρ is the density matrix of the quantum state. Therefore, the symplectic eigenvalues ν_j of the covariance matrix would also scale the von Neumann entropy in order $\nu_j\log\nu_j$, warranting the $s^2\log(1/s^2)$ dependence (as $\nu_j \sim s^2$). Furthermore, the von Neumann entropy is non-analytic unlike the rest of the Rényi entropies. Therefore, the argument presented in Ref. [2] that the Rényi entropy can be expanded in lowest order s^2 doesn't apply to the von Neumann case. More specifically, the deviation from analytic behavior arises from the von Neumman entropy formula in terms of the covariance matrix σ [138, 152, 178]:

$$S_1 = \frac{1}{2}\log\det\left(\frac{\sigma+\mathrm{i}\Omega}{2}\right) + \frac{1}{2}\operatorname{Tr}\left(\coth^{-1}(\mathrm{i}\Omega\sigma)\mathrm{i}\Omega\sigma\right), \tag{3.6}$$

where Ω is the symplectic form defined in Eq. (3.1). Since $\coth^{-1}(x)$ is not a real analytic function, the resulting Page curve is not analytic, either.

In a similar vein, we compute the von Neumann Page curve as $s \to \infty$, assuming an equal squeezing regime.

Theorem 3.3.3. *In the strong squeezing limit $s \to \infty$,*

$$\lim_{s\to\infty}\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in U(n)}\frac{1}{sn}S_1(U) = 2\min(r, 1-r). \tag{3.7}$$

Here, the von Neumann Page curve scales with s as $s \to \infty$. This is because, for large absolute squeezing values, higher orders in r dominate the expansion of Theorem 3.3.1, which scale with s. Interestingly, Theorem 3.3.3 remains the exact same for the Rényi-2 entropy [2], and indeed we will see in Sec. 3.4.2 that it is the same for all the Rényi-α entropies for integral α.

We now compare this result to the case of averaging over *all* n-mode bosonic states with a fixed boson number constraint. In particular, the volume (*i.e.* proportional to n) scaling of the corresponding Page curve is [179]

$$\lim_{n\to\infty} \frac{1}{n} S_1^{(n_b)} = ((r_b+1)\log(1+r_b) - r_b \log(r_b)) \min(r, 1-r), \tag{3.8}$$

where $r_b n = n_b$ the number of bosons. In this case, the Hilbert space is finite-dimensional with dimension $\binom{n+n_b-1}{n_b}$. As $n_b \to \infty$, this yields $\frac{1}{n} S_1^{(n_b)} \sim \log(r_b) \min(r, 1-r)$

In our Gaussian setup, the expected number of bosons is $\langle n_b \rangle = n \sinh^2(s)$. In this case, the Hilbert space in infinite-dimensional, but the fixed squeezing constraint gives an energy constraint. Using Theorem 3.3.3, we then see that as $n_b \to \infty$, $\frac{1}{n} S_1 \sim \log(\langle r_b \rangle) \min(r, 1-r)$.

We therefore see that, asymptotically, the average entanglement scales proportional to the volume of the system. Given a fixed but asymptotically large energy (*i.e.* number of bosons), the average entanglement over *Gaussian* states equals the average entanglement over *all* bosonic states.

Finally, we also note that when $r \to 0$, all of the Page curves that we have computed, and indeed all of the Page curves computed in Ref. [179], scale proportionally to r. In particular, we can see from Ref. [179, Fig. 2, Eq. (23)] that the fixed particle number Page curves for (a) fermions, (b) hardcore bosons, and (c) (multi-species) bosons all have Page curves that are proportional to r as $r \to 0$. Similarly, Page curves in finite-dimensional spaces without constraints are also proportional to r as $r \to 0$ [65, 99–101, 105, 118, 121, 169–171]. In the Gaussian case that we have computed in this work, the Page curve is also proportional to r

in this limit, as can be seen in Theorems 3.3.1 and 3.4.1.

Subsection 3.3.3: Unequal squeezing

In Ref. [2], the expected value of the Rényi-2 entropy for small, unequal squeezing strengths was calculated through a power-series analysis. The analyticity of the Rényi-2 function was what allowed writing the entropy as a power series in s_i, where s_i was the squeezing strength of the ith mode. Using this power series expansion, along with the translational invariance of the Haar measure, Ref. [2] derived a formula for unequal squeezing in order s^2, which reduces to the $s \to 0$ squeezing limit for small, equal s_i.

However, due to the von Neumann entropy not being analytic around $s = 0$, we cannot perform the same analysis to derive a formula for unequal squeezing strengths. Furthermore, directly approaching the unequal squeezing regime via the Taylor expansion of the von Neumann entropy runs into another problem, since the separation between s and r dependence present in the Rényi-2 case presented in Ref. [2] no longer applies here. Specifically, the matrix W described in Appendix B.1.1 is no longer independent of s. Therefore, carrying out the Taylor expansion would yield complicated matrix expressions in σ that do not simplify appreciably at first glance. Due to these difficulties, we are unable to derive the unequal squeezing case for the von Neumann entropy. It is nonetheless possible, though, that executing the complicated matrix Taylor expansion and simplifying terms yields a closed-form result for the von Neumann entropy.

Section 3.4: General Rényi-α Entropies

In this section, we extend much of our analysis with the von Neumann entropy in Section 3.3 to Rényi-α entropies for integer $\alpha \geq 2$. We also derive results for the typicality of entanglement and the small, unequal squeezing case. First, we present the Rényi-α Page curve

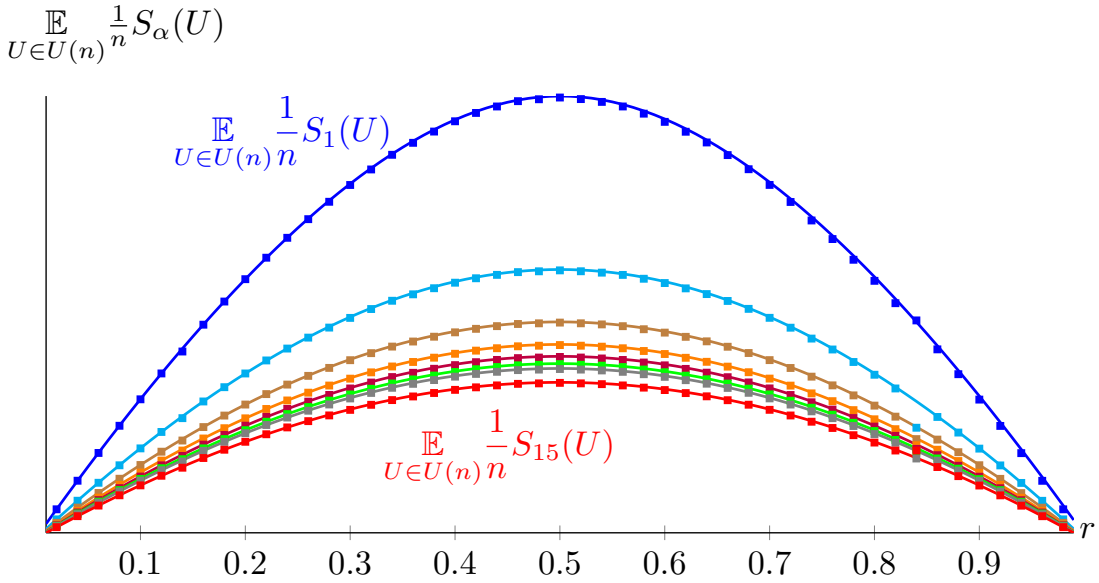

Figure 3.1: Simulated Page curves for Rényi-α entropies for $\alpha \in \{1, 2, 3, 4, 5, 6, 7, 15\}$ in dots (smaller Page curves indicate larger α), overlaid on the analytic expressions derived in Theorem 3.3.1 and Theorem 3.4.1. The simulated values were run 250 times. The number of modes is $n = 400$ with equal squeezing strength $s = 0.5$.

$\forall \alpha \in \mathbb{Z}, \alpha \geq 2$. Next, we investigate squeezing limits as $s \to 0$ and $s \to \infty$, as well as consider the Rényi-α Page curve for small, unequal squeezing strengths. We also study the variance of the entanglement and prove various regimes of entanglement typicality. Finally, we show some numerical simulations confirming our results.

Subsection 3.4.1: Explicit formula

As before, we consider a system of n modes, each with equal initial squeezing strength s.

Theorem 3.4.1 (Rényi-α Page curve)**.** *For all integer $\alpha \geq 2$, define $a = \left\lfloor \frac{\alpha-1}{2} \right\rfloor$. Let $\zeta = 1$ if α is even and $\zeta = 0$ if α is odd. Then, the Rényi-α entropy is given by*

$$\mathop{\mathbb{E}}_{U\in U(n)} S_\alpha(U) = \frac{\zeta}{\alpha-1} \mathop{\mathbb{E}}_{U\in U(n)} S_2 + \frac{1}{\alpha-1}\sum_{i=1}^{\infty}\left[\sum_{m=1}^{a} \frac{\sinh^{2i}(2s)}{i(\cosh^2(2s)+\cot^2\frac{\pi m}{\alpha})^i}\right](nG_i(r)-H_i(r)+o(1))\,, \tag{3.9}$$

where $\mathop{\mathbb{E}}_{U\in U(n)} S_2$ *is the expected value of the Rényi-2 entropy as a function of* s *and* r *given as*

$$\mathop{\mathbb{E}}_{U\in U(n)} S_2 = \sum_{i=1}^{\infty} \frac{\tanh^{2i}(2s)}{2i}(nG_i(r)-H_i(r)+o(1))\,, \tag{3.10}$$

and where $G_i(r)$ *and* $H_i(r)$ *are defined in Eq.* (3.3).

The proof of Theorem 3.4.1 is shown in Appendix B.2.1. In the proof, we use a formula for the Rényi-α entropy in terms of the symplectic eigenvalues of σ, the covariance matrix of the reduced state of k modes (see, e.g., Ref. [138]). We convert this formula into a matrix expression in terms of Ω and σ, and perform significant simplifications. Finally, to compute the expectation value over the unitary group, we again utilize averaged matrix quantities derived in Ref. [2].

Since $G_i(r)$ is symmetric under $r \to 1-r$, we see that the Page curves are also manifestly symmetric under $r \to 1-r$. Various analytic Rényi-α Page curves are plotted in Fig. 3.1. We also plot numerically-simulated values of the Page curve, where we generate Haar-random unitary matrices and compute the entropies using the positive symplectic eigenvalues ν_j of each covariance matrix σ. Then, we compute the respective Rényi-α entropy as $S_\alpha(U) = \sum_{j=1}^{n} \frac{1}{1-\alpha}\ln\left(\frac{2^\alpha}{(\nu_j+1)^\alpha-(\nu_j-1)^\alpha}\right)$ [138]. The numerical and analytical values overlap, indicating the validity of Theorem 3.3.1 and Theorem 3.4.1.

Subsection 3.4.2: Squeezing limits

We now investigate the Rényi-α Page curve for all integer $\alpha \geq 2$ in the squeezing limits $s \to 0$ and $s \to \infty$. We find that, just as in the Rényi-2 case, the Page curve scales with s^2 when s approaches zero and scales with s when $s \to \infty$.

Theorem 3.4.2. *For integer $\alpha \geq 2$, the Page curves of the Rényi-α entropy as the equal squeezing strength s approaches 0 and ∞, respectively, are given by*

$$\lim_{s\to 0}\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in U(n)}\frac{1}{ns^2}S_\alpha(U) = \frac{\alpha}{\alpha-1}r(1-r), \tag{3.11}$$

$$\lim_{s\to \infty}\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in U(n)}\frac{1}{ns}S_\alpha(U) = 2\min(r,1-r). \tag{3.12}$$

In general, the Rényi entropies satisfy $S_\alpha \geq S_{\alpha'}$ whenever $\alpha \leq \alpha'$. This behavior matches with our formulas for the Rényi-α limits in Theorem 3.4.2, as the Page curve decreases when α increases. From Theorem 3.4.2, it is furthermore evident that $S_2(U) = 2S_\infty(U)$ when $s \to 0$. This relation between two Rényi entropies saturates the case of a discrete random variable with outcome probabilities $p_1, p_2, \ldots, p_m$, where

$$S_2 = -\log\sum_i p_i^2 \leq -\log\sup p_i^2 = -2\log\sup p_i = 2S_\infty. \tag{3.13}$$

In the infinite squeezing case, the Page curve remains independent of α in order s as $\alpha \to \infty$. This indicates that, as $s \to \infty$, the Page curve approaches the curve of maximum Gaussian entanglement (that is, the maximum is taken over all Gaussian states with a fixed initial squeezing configuration).

		$k = \Theta(n)$	$k = o(n)$ [2]	$k = o(n^{1/3})$ [33]
Equal squeezing	*Rényi-α ($\alpha \geq 2$)*	weak	strong	strong
	von Neumann	?	weak	strong
Unequal squeezing	*Rényi-α ($\alpha \geq 2$)*	?	weak*	strong
	von Neumann	?	weak*	strong

Table 3.1: A summary of the current status of rigorous results on entanglement typicality in Gaussian bosonic systems. This table is the same as the one provided in Ref. [2] and Table 2.1, but the Rényi-2 case has been generalized for all Rényi-α entropies where $\alpha \in \mathbb{Z}^+$. Note that "weak*" indicates that the result is not fully proven, but is based on Conjecture 10 in Ref. [2]. Furthermore, wherever "weak" typicality is displayed, we have not ruled out the possibilty of strong typicality. We present typicality results when k, the number of modes in the subsystem, scales with $\Theta(n)\,, o(n)\,,$ and $o\big(n^{1/3}\big)\,,$ respectively. Note that we only take the "worst-case" typicality; if there exists both weak and strong typicality for one scaling, we display only the weakly typical case. We present typicality results separately for the Rényi-α entropies and the von Neumann entropy. The leftmost column ($k = \Theta(n)$) stems from our findings in Section 3.4.4, while the middle column ($k = o(n)$) follows from Ref. [2]. The rightmost column ($k = o\big(n^{1/3}\big)$) follows from the results of Ref. [33]. Refs. [34, 132] proved strong typicality in the regime $k = \mathcal{O}(1)$.

Subsection 3.4.3: Unequal squeezing

We now consider the Rényi-α entropy when the initial squeezing strengths are not necessarily equal, but are close to zero. We let the squeezing strength on the ith mode be s_i for some $s_i \in \mathbb{R}$. Define $s_{\max} \coloneqq \max\left(|s_1|, |s_2|, \ldots, |s_n|\right)$.

Theorem 3.4.3. *For integer $\alpha \geq 2$, the Rényi-α entropy for small, unequal squeezing strengths $s_1, s_2, \ldots, s_n$ is given by*

$$\mathop{\mathbb{E}}_{U \in \mathrm{U}(n)} S_\alpha(U) = \frac{\alpha}{\alpha - 1} r(1-r) \sum_{i=1}^{n} s_i^2 + \mathcal{O}\big(r n s_{\max}^4\big)\,. \tag{3.14}$$

To prove Theorem 3.4.3, note that the Rényi-α entropy is real analytic for positive integer α. This is because the Rényi-α expression, when written in terms of the covariance matrix σ, contains only analytic expressions in the vicinity of $s = 0$, and thus the final expression is also analytic in the small-squeezing limit [138]. Thus, we can write the entropy as a power

series in s_i. Furthermore, there exists a passive, local (i.e. acting independently on each mode) Gaussian unitary that acts on the initial covariance matrix σ_0 with the transformation $s_i \to -s_i$. Since this unitary can absorbed into the Haar random unitary U, the power series can be expanded in s_i^2. Note that when the absolute squeezing strengths are close to zero, the first statement of Theorem 3.4.2 is a special case of Theorem 3.4.3.

There exists another passive Gaussian unitary that acts on σ_0 via the transformation $s_i \to s_{\tau(i)}$ for any permutation τ of n elements. Since this unitary can also be absorbed into the Haar random unitary U, the power series is symmetric under $s_i \to s_{\tau(i)}$. Thus, the s_i^2 term in the power series expansion is of the form $g(r,n)\sum_{i=1}^n s_i^2$ for some function $g(r,n)$. The equal-squeezing result in Theorem 3.4.2 then immediately implies $g(r,n) = \frac{\alpha}{\alpha-1} r(1-r)$.

Subsection 3.4.4: Typicality of entanglement

Next, we investigate the variance of the Rényi-α Page curve. Chebyshev's inequality allows one to relate the variance with the *typicality of entanglement*, where typicality describes how quickly the difference between some quantity (in this case, entanglement) and its average vanishes. This analysis is useful in statistical mechanics, where typicality is useful for thermodynamic variables to accurately represent their mean values.

The formal definition of typicality as described in Ref. [2] is as follows: a unitary-dependent nonnegative random variable $S(U)$ is *weakly typical* if, for any constant $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr_{U\in U(n)} \left[\left| \frac{S(U)}{\mathop{\mathbb{E}}_{V\in U(n)} S(V)} - 1 \right| < \epsilon \right] = 1. \tag{3.15}$$

Similarly, $S(U)$ is *strongly typical* if, for any constant $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr_{U\in U(n)} \left[\left| S(U) - \mathop{\mathbb{E}}_{V\in U(n)} S(V) \right| < \epsilon \right] = 1. \tag{3.16}$$

Perhaps contrary to intuition, a weakly typical variable has its multiplicative distance vanish in the thermodynamical limit, while a strongly typical variable has its additive distance vanish. This occurs because the entropy goes to infinity as $n \to \infty$. In this regard, strong typicality implies weak typicality.

We are able to prove typicality results based on a crucial invariance described in Ref. [2]. Specifically, Ref. [2] derives that the variance of the Rényi-2 entropy is constant as $n \to \infty$. We extend this to general integer $\alpha \geq 2$ using a Cauchy-Schwarz argument in Appendix B.2.4. Using a bound for the variance via Chebyshev's inequality, the Rényi-α entropy exhibits weak typicality for $k = \Theta(n)$ (again, for all integer $\alpha \geq 2$). We present our results, as well as past analyses of entanglement typicality in Gaussian bosonic systems, in Table 3.1.

Section 3.5: Conclusions and open questions

In this paper, we have carefully studied the Rényi and von Neumann entropies of random bosonic Gaussian states. Such states are exactly the output states of a GBS experiment. Specifically, we have found closed formulas for the Rényi-α entanglement entropy in terms of the squeezing strength s and the mode partition ratio r that are exact asymptotically in n, where $\alpha \in \mathbb{Z}^+$. Using this, we calculated the Rényi-α Page curve and various results stemming from it, such as $s \to 0$ and $s \to \infty$ squeezing limits, and, for $\alpha \geq 2$, variance and typicality of entanglement.

We have not reached a conclusion regarding the typicality of entanglement for the von Neumann entropy as k scales with $\Theta(n)$. We have that $\mathrm{Var}_U\, S_1(U) = \mathbb{E}_U\, S_1(U)^2 - (\mathbb{E}_U\, S_1(U))^2$, which we have not been able to simplify appreciably when using the von Neumann entropy formula from Theorem 3.3.1. Specifically, the squaring of hypergeometric functions yields terms in the variance expansion that we have not been able to simplify. More research should be done to determine the typicality of entanglement for the von Neumann entropy. However, since we've proven in Appendix B.2.4 that the Rényi-α entropies for integer $\alpha \geq 2$ are only

weakly typical, the von Neumann entropy is unlikely to be strongly typical. It remains to be rigorously proven, perhaps through examining special cases of the von Neumann entropy, such as when $r = \frac{1}{2}$ or $s \to 0$, and attempting to extract the variance from these cases.

Another point of research is in generalizing the Rényi-α entropy in Theorem 3.4.1 to non-integer α. In our work, we only considered $\alpha \in \mathbb{Z}^+$. Extending to non-integer α would allow a formula for the Rényi entropies that is continuous in α, and could be useful to examine the behavior of the Rényi entropy as $\alpha \to 1$.

We also propose other areas of further research. Thus far, we have only considered Haar-random unitaries that describe the linear optical circuit. However, an interesting point of research could be in examining the entanglement associated with random linear optical circuits of finite depth, that is, the *depth dependence* of entanglement entropy. Currently, the precise relationship between entanglement and circuit depth is largely unknown [180]. There have been some numerical studies [153], but only specialized cases where most, but not all, initial modes are vacuum are understood analytically [155].

Another related notion deals with *sampling complexity*, the overall complexity of classical computers producing a sample from approximately the same distribution as the quantum system [181]. Quantifying the relationship between entanglement and complexity, and studying the depth dependence of entanglement and complexity, could yield crucial new insights into the classical hardness of simulating quantum experiments such as GBS and, more generally, into the relation between entanglement and complexity.

Another problem to tackle deals with unequal squeezing. There are two primary avenues of future research. First, for small, varying squeezing strengths $s_1, s_2, \ldots, s_n$, what is the average von Neumann entanglement entropy of the subsystem? We have outlined a method in Section 3.3.3 of tackling this problem, but initial attempts have proved futile due to the heavy presence of inverse matrices. Nonetheless, learning the behavior of the von Neumann entropy for unequal squeezing regimes is important from a practical standpoint, since many

GBS experiments are conducted with one set of modes at zero squeezing strength and the other at some varied squeezing strength close to zero. Second, it is interesting to consider Rényi-α entropies for systems with varying squeezing strengths that are not necessarily close to zero. The argument given in Section 3.4.3 revolves around the entropy expression being analytic around $s = 0$ and vanishing in higher orders of $s_{\max}$. However, for the most general case of $s_{\max} \in \mathbb{R}$, one cannot simply extend the $s \to 0$ squeezing limit for unequal squeezing. Nonetheless, deriving such an expression would be important for a wider range of bosonic systems.

Part II

Quantum state designs and their application

Chapter 4: Continuous-variable quantum state designs: theory and application

Section 4.1: Introduction & summary

It is useful in a wide variety of fields to be able to efficiently calculate uniform averages of polynomial functions over points in a space. Prominent examples include Gaussian quadrature rules [182] and spherical designs [183, 184], which reduce integrals of polynomials to weighted sums of polynomial values at particular points. More generally, a *t-design* over a space is a set of points picked in such a way that averaging any polynomial of degree $\leq t$ over the design is equivalent to uniformly averaging the same polynomial over the space. Gaussian quadrature rules and spherical designs are t-designs over the hypercube and hypersphere, respectively, and closely related ideas can be formulated for simplices and tori [185–191] as well as general topological spaces [192].

Designs also have a number of important applications in quantum theory. A quantum state t-design is an ensemble of quantum states such that expectation values of homogeneous polynomials of degree t or less in the amplitudes of quantum states are the same whether the averaging is performed uniformly over all states or over only the states in the design [36–46]. State, unitary, and spherical [44, 47] designs are important tools in tomography [40, 43, 48–52], state distinction [44, 53], randomized benchmarking [47, 54–57], fidelity estimation [47, 58–63], cryptography [64, 65], sensing [66, 73], fundamental physics [45, 67–69], and error correction [75–78].

Both the original formulation of designs and its quantum counterparts hold only for finite-dimensional spaces. This means that none of the applications proven to work through the use of designs, e.g., quantum state fidelity relations [47, 54–57, 59, 60] and design-based tomographic protocols [49–52], carry over naturally to countably *infinite*-dimensional spaces. Such spaces are important for quantum applications because they describe quantum systems whose natural degrees of freedom are continuous variables (CVs), e.g., electromagentic modes of optical or microwave cavities, or mechanical modes of harmonic oscillators.

Formulating a notion of designs would unlock important abilities for CV systems. We proceed to do so in this paper, summarizing both our formalism and several fleshed-out applications below.

Non-existence of CV designs A first attempt to define state t-designs for CV systems was made in Ref. [79]. The authors showed that a particular set of CV states — the Gaussian states [193] — does not form a CV 2-design. This is perhaps surprising since Gaussian unitaries are the infinite-dimensional analog of finite-dimensional Clifford unitaries, which themselves can form 2-designs [194–197]. Similarly, Ref. [153] defined the notion of CV unitary t-designs and argued that Gaussian unitaries do not form a 2-design. These results leave open the question of whether CV state (unitary) designs require non-Gaussian states (unitaries).

In this work, we answer this open question and prove that CV state and unitary t-designs do not exist for any $t \geq 2$. Our results hold for any separable, infinite-dimensional Hilbert space, not just the space $L^2(\mathbb{R})$ associated with CV quantum systems. Thus, even the inclusion of non-Gaussian states and unitaries does not help in defining t-designs over CV systems.

Our proof relies on the connection between state designs and simplex designs. We first show that infinite-dimensional simplex t-designs do not exist for $t \geq 2$. Then, using the simple fact that the complex probability amplitudes of any pure quantum state can be

parameterized by a simplex (for the moduli) and a torus (for the phases), we show by contradiction that infinite-dimensional state designs do not exist either.

Rigged designs We show that removing the requirement for states to be normalizable yields a meaningful extension of the notion of designs. We define *rigged t-designs* that utilize states in a rigged Hilbert space, the Hilbert space populated by, e.g., the non-normalizable eigenstates of the oscillator position and momentum operators. We construct several examples for rigged 2-designs, thus proving that rigged state designs exist even though CV state ($t \geq 2$)-designs do not.

In particular, it is well-known that there is no notion of *uniform* integration over $L^2(\mathbb{R})$, and our proof that CV t-designs do not exist for $t \geq 2$ proves that there is no form of integration over $L^2(\mathbb{R})$ that has even basic qualities that *mimick* uniform integration. Rigged designs get around this shortcoming by expanding the integration space to the set of all non-normalizable states in a rigged Hilbert space—specifically, the space $S(\mathbb{R})'$ of *tempered distributions*. We construct a measure on $S(\mathbb{R})'$ that mimicks the qualities of a uniform measure over infinite-dimensional quantum states, and we then construct designs on this space.

Our first rigged design consists of Fock states as well as the phase states, which form a well-known positive operator-valued measure (POVM) that is optimal for measuring the angle of rotation induced on a mode [198–204] (see [205, Sec. 3.9] for an exposition). The other examples combine Fock states with the cosine and sine states (and rotated states thereof), close relatives of the phase states [199]. In all cases, an extra parameter is induced on the phase states via evolution by a "Kerr" Hamiltonian $\hat{n}^2$, with $\hat{n}$ the occupation number operator [206].

Design-based shadows The ability to use rigged t-designs as POVMs lends itself to a natural extension of shadow tomography [49–52] to CV systems. In finite-dimensional

versions of such protocols, one generates a classical snapshot of an unknown quantum state by performing random measurements according to the states from a 2- or 3-design. Then the expectation values over several observables can be efficiently and accurately estimated using these classical snapshots [50].

We propose a CV shadow tomography protocol based on the Kerred-phase-state and Fock-state rigged 2-design. The advantage of our protocol is that it maintains the key feature of the original qubit shadow protocols; namely, the ability to efficiently measure many observables using only a set of "shadow" snapshots of a particular form. This protocol can be generalized to an efficient multi-mode protocol using a recent result [207]. We discuss how our rigged CV shadows can be used for CV entanglement verification.

Although our design-based shadow protocol is more experimentally taxing than, e.g., CV shadows based on conventional homodyne or photon parity measurements [207], it can be implemented by combining and improving previously demonstrated experimental techniques. In order to utilize our first (second, third) rigged two-design as a POVM in the lab, one needs to be able to evolve the system under a Kerr Hamiltonian and then apply the phase (cosine, sine) state POVM. In addition, one needs to alternatively measure in the Fock-state basis. All three aspects of this protocol — CV phase measurements [208, 209], photon-number resolution (e.g., [210]), and engineered Kerr evolution [211–213] — have been realized in some form in microwave cavities coupled to superconducting qubits [206]. Providing an experimentally realizable implementation of our protocol that can achieve the same scaling as our predicted sample complexity is an interesting avenue for future work.

Approximate CV designs Another natural question to ask is whether *approximate* CV state designs exist in $L^2(\mathbb{R})$. Or, can the notion of designs be defined over the CV states that satisfy some energy constraints? We provide a solution to this problem by regularizing the rigged CV designs.

To approximate our rigged designs with sets of normalized states, we use operators called

regularizers, which correspond to different cutoffs of the infinite-dimensional space. For example, a regularizer that projects onto a low-energy finite-dimensional subspace corresponds to a hard cutoff, i.e., a maximum-energy constraint. A regularizer that smoothly decays with increasing energy but has support on the full infinite-dimensional space corresponds to a soft cutoff, i.e., an average-energy constraint. By analogy to numerical quadrature rules on the real line, a sharp cutoff is akin to restricting the domain of integration to a compact interval, while a smooth cutoff is akin to endowing the line with a Gaussian measure. Moreover, certain regularizers allow us to extend the notion of a frame potential [40, 41, 214] to infinite dimensions.

Regularizers (*a.k.a.* cooling or damping operators) and related ideas have been employed in works on CV quantum error-correcting codes [215, 216][217, Appx. B], uniform continuity for quantum entropies [218], energy-constrained capacities [219] and distances [220, 221] of CV channels, and CV cryptographic protocols [222].

Average CV fidelity Armed with regularized-rigged designs, we extend the well-known notion of average fidelity (over all states) of a quantum channel from finite-dimensional [47, 54–57] to CV systems. In previous such extensions, systems were limited to the setting where the average fidelity between operations is estimated over an ensemble of coherent states [223–231]. Other approaches to benchmarking CV operations rely on witnesses that are lower bounds to the true average fidelity over an ensemble of Gaussian states [229, 230], while energy-constrained diamond-distance based performance estimates require knowledge of the noise model in experimental approximations and are often computationally taxing [220, 231–236].

We provide two different definitions of the average fidelity of a CV quantum channel. These formulas can be directly employed to estimate the average fidelity between CV quantum gates and their experimental approximations [232]. Our formulation yields an experimental procedure to estimate the average fidelity of an arbitrary CV quantum gate without

requiring the knowledge of the noise involved in experimental implementations. As a concrete example, we estimate the average fidelity between an ideal displacement operation and its experimental approximation [232], suggesting that an average over coherent states only is not a good approximation to an average over all CV states.

Average-to-entanglement fidelity relation Another interesting open question in CV information theory is to establish a relation between the average channel fidelity and the entanglement fidelity, similar to the finite-dimensional setting [59, 60]. In this work, we solve this open problem and establish connections between average and entanglement fidelities for CV operations.

We utilize the conventional notion of single-mode CV entanglement fidelity, namely, the fidelity over a two-mode squeezed vacuum state [138, 237]. We then evaluate our average fidelity over states in the corresponding regularized-rigged design. Combining these two fidelity formulas, we establish a simple relation between the average gate fidelity and the entanglement fidelity for CV operations.

Relating designs to MUBs As an additional result of independent interest, we find a relationship between torus 2-designs and complete sets of mutually unbiased bases [238], and we prove that the condition of mutually unbiasedness can be replaced by a torus 2-design condition.

Outline The rest of the paper is meant to succinctly relay the results and is structured as follows. In Section 4.2, we introduce finite-dimensional designs. In Section 4.3, we develop the notion of infinite-dimensional designs and prove that CV state and unitary t-designs do not exist for any $t \geq 2$. In Section 4.4, we then define rigged designs and provide explicit constructions for rigged 2-designs. In Section 4.5, we introduce regularized rigged designs. In Section 4.6, we study applications of rigged and regularized rigged designs. In particular,

in Section 4.6.1, we introduce the shadow tomography formalism to CV quantum states. In Section 4.6.2, we discuss how such rigged CV shadows can be used for CV entanglement verification. Next, in Section 4.6.3, we define various notions of the average fidelity of a CV quantum channel using regularized rigged 2-designs. We then prove a relationship between the CV entanglement and average fidelities. Finally, in Section 4.7, we conclude with a brief summary and discuss open questions.

Section 4.2: Finite dimensional designs

In this section, we review relevant prior results on finite-dimensional state designs, making contact with designs on simplices and tori.

Quantum state designs reduce integrals of polynomials over all quantum states to averages over a discrete set. Let $\mathbb{C}^d$ denote a d-dimensional Hilbert space with orthonormal basis $\{|n\rangle\}_{n=0}^{d-1}$. Due to their normalization and global-phase redundancy, quantum states in this space correspond to points in the complex-projective space $\mathbb{CP}^{d-1}$ [72, 239]. A non-trivial complex-projective t-design is a set of states $X \subsetneq \mathbb{CP}^{d-1}$, sampled according to some probability measure μ, satisfying [36–45]

$$\mathop{\mathbb{E}}_{\psi\in X} f(\psi) = \int_{\mathbb{CP}^{d-1}} f(\psi)\,\mathrm{d}\psi \tag{4.1}$$

for any polynomial $f(\psi)$ of degree t or less in the amplitudes of ψ and degree t or less in the conjugate amplitudes. The canonical measure $\mathrm{d}\psi$ on the set of such quantum states, called the *Fubini-Study measure* [72, 239], is the unique unit-normalized volume measure that is invariant under the action of the unitary group $\mathrm{U}(d)$ (see Appendix C.2.1 for more details).

The above conventional relation can be lifted into a relation between particular operators by using the fact that polynomials of degree up to t in state degrees of freedom can be expressed as expectation values of operators with t copies of the state.

Consider, for example, $t = 2$ and an arbitrary polynomial $f(\psi) = \sum_{j,k,l,m=0}^{d-1} f_{jklm} \bar{\psi}_j \bar{\psi}_k \psi_l \psi_m$ in the amplitudes $\psi_j \coloneqq \langle j|\psi\rangle$ and their conjugates $\bar{\psi}_j$, with complex coefficients f_{jklm}. This polynomial can equivalently be expressed as an expectation value of a bipartite operator $\hat{f}$ with respect to two copies of $|\psi\rangle\langle\psi|$,

$$f(\psi) = \sum_{j,k,l,m=0}^{d-1} f_{jklm} \langle l|\psi\rangle\langle m|\psi\rangle\langle\psi|j\rangle\langle\psi|k\rangle \tag{4.2a}$$

$$= \mathrm{Tr}\left(\hat{f}\ |\psi\rangle\langle\psi|^{\otimes 2}\right)\ , \tag{4.2b}$$

where $\hat{f} = \sum_{j,k,l,m=0}^{d-1} f_{jklm} |j\rangle|k\rangle\langle l|\langle m|$. Using this relation, we see that X is a t-design if and only if

$$\mathop{\mathbb{E}}_{\psi\in X} (|\psi\rangle\langle\psi|)^{\otimes t} = \int_{\mathbb{CP}^{d-1}} (|\psi\rangle\langle\psi|)^{\otimes t}\, \mathrm{d}\psi\ . \tag{4.3}$$

Next, we can use representation theory (see Appendix C.3.3 for details) to solve the integral on the right-hand side, yielding

$$\int_{\mathbb{CP}^{d-1}} (|\psi\rangle\langle\psi|)^{\otimes t}\, \mathrm{d}\psi = \frac{\Pi_t^{(d)}}{\mathrm{Tr}\,\Pi_t^{(d)}}\ , \tag{4.4}$$

where $\Pi_t^{(d)}$ is the projector onto the permutation-invariant (*a.k.a.* symmetric [240]) subspace of $(\mathbb{C}^d)^{\otimes t}$, the t-fold tensor product of the original space, and Tr is the trace function. When $t = 1$, this integral reduces to a resolution of the identity. For higher t, the resolution can only be of the symmetric subspace since the t-fold tensor product of any state is symmetric under all permutations (see Appendix C.2.2 for details).

Combining the above manipulations yields the following "operator-level" definition of a complex-projective t-design,

$$\mathop{\mathbb{E}}_{\psi\in X} (|\psi\rangle\langle\psi|)^{\otimes t} = \frac{\Pi_t^{(d)}}{\mathrm{Tr}\,\Pi_t^{(d)}}\ . \tag{4.5}$$

Designs can be obtained via the convenient parameterization of pure states in terms of a

simplex and a torus. State amplitudes can be written as

$$\langle j|\psi\rangle = \sqrt{p_j}\ \mathrm{e}^{\mathrm{i}\phi_j}\ , \tag{4.6}$$

where the probabilities p_j add up to one due to normalization, and the phases ϕ_j are 2π-periodic (with ϕ_0 set to zero to remove global-phase redundancy). By definition, the probability distribution defined by p_j is a point on the $(d-1)$-simplex,

$$\Delta^{d-1} \coloneqq \left\{ (p_0, \ldots, p_{d-1}) \in [0,1]^d \;\middle|\; \sum_{j=0}^{d-1} p_j = 1 \right\}, \tag{4.7}$$

while the vector of phases parameterizes a $(d-1)$-torus T^{d-1}. Hence, volume integration over all states is reduced to volume integration over the simplex and the torus [72, 239] (see Appendix C.3.4 for details). This naturally makes contact with simplex and torus designs.

Simplex and torus designs are defined in similar fashion to complex-projective designs. A set $X \subset \Delta^m$ of probability vectors is an m-simplex t-design if for all tuples $a = (a_1, \ldots, a_t) \in \{0, 1, \ldots, m\}^t$,

$$\mathop{\mathbb{E}}_{q \in X} \prod_{i=1}^{t} q_{a_i} = \int_{\Delta^m} \prod_{i=1}^{t} p_{a_i}\, \mathrm{d}p, \tag{4.8}$$

where $\mathrm{d}p$ is the standard measure on the simplex. A set of angles $X \subset T^m$ is an m-torus t-design if for all tuples $a = (a_1, \ldots, a_t) \in \{1, 2, \ldots, m\}^t$ and $b = (b_1, \ldots, b_t) \in \{1, 2, \ldots, m\}^t$,

$$\mathop{\mathbb{E}}_{\theta \in X} \prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\theta_{a_i} - \theta_{b_i})} = \int_{T^m} \prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\phi_{a_i} - \phi_{b_i})}\, \mathrm{d}\phi, \tag{4.9}$$

where $\mathrm{d}\phi$ is the standard measure on the torus. We discuss various constructions of simplex and torus designs in Appendix C.3.1 and Appendix C.3.2, respectively.

There is a bilateral connection between complex-projective designs and designs on the corresponding simplices and tori. Denoting π as the "Born-rule" map that produces the vector of probabilities $(p_n)_{n=0}^{d-1}$ from a state $|\psi\rangle$, the set $\pi(X)$ is a simplex t-design for any

complex-projective t-design X [70, 189] (see Appendix C.3.5 for details). On the other hand, a combination of a simplex and a torus t-design of appropriate dimensions yields a complex-projective t-design [189]. We provide a proof of these latter connections and present various combinations that yield complex-projective 2-designs for all d in Appendix C.3.4.

Our simplex designs from Eq. (4.8) are more commonly referred to as simplex positive, interior (or boundary) cubature rules [185–190]. Our torus t-designs from Eq. (4.9) closely resemble *trigonometric cubature rules* [186], but the two are not equivalent. In Appendix C.6, we show that torus designs are equivalent to a special case of torus cubature rules from Ref. [189]. We then establish a connection between torus 2-designs and mutually unbiased bases (MUBs), which might be of independent interest. To the best of our knowledge, this connection has not been previously discussed.

Section 4.3: Continuous-variable designs

In this section, we develop the notion of continuous-variable (CV) designs and present our main results in Theorem 4.3.2 and Corollary 4.3.3.

Let $L^2(\mathbb{R})$ denote an infinite-dimensional, separable Hilbert space of square-integrable functions on the real line, with a countable *Fock-state* (*a.k.a.* photon number-state or occupation number-state) basis $\{|n\rangle \mid n \in \mathbb{N}_0\}$, where $\mathbb{N}_0$ denotes the natural numbers including zero. We note that all separable Hilbert spaces are isomorphic to $L^2(\mathbb{R})$. We call unit-norm vectors in $L^2(\mathbb{R})$ CV quantum states.

The right-hand side of Eq. (4.4) is straightforward to generalize to infinite dimensions. Let Π_t denote the projector onto the symmetric subspace of t copies of $L^2(\mathbb{R})$ (see Appendix C.2.2). For any tuples $a = (a_1, \ldots, a_t) \in \mathbb{N}_0^t$ and $b = (b_1, \ldots, b_t) \in \mathbb{N}_0^t$,

$$\Pi_t(a; b) \coloneqq \left(\bigotimes_{i=1}^{t} \langle a_i| \right) \Pi_t \left(\bigotimes_{i=1}^{t} |b_i\rangle \right) \tag{4.10}$$

denotes the matrix elements of Π_t. The trace of this projector, $\operatorname{Tr}\Pi_t$, is infinite, but we can simply omit it from the equation.

The left-hand side of Eq. (4.4) is unfortunately impossible to generalize to infinite dimensions [241]. Since $L^2(\mathbb{R})$ is infinite dimensional, there is no finite Haar measure on its corresponding unitary group $\mathrm{U}(L^2(\mathbb{R}))$ [242, Sec. 5]. Therefore, there is no natural unitarily invariant volume measure on the set of CV quantum states. However, if one *could* define the unitarily invariant volume integration over all CV states, Schur's lemma *would* imply that the resulting integration would be proportional to Π_t. Therefore, in principle, infinite-dimensional state designs can be defined using the definition of complex-projective designs in Eq. (4.5), but without the $\operatorname{Tr}\Pi_t$ term.

An infinite-dimensional design may be parameterized by points in a noncompact space with a non-normalizable measure. To accommodate this, we relax the assumption that the parameter space of a design is a probability space and instead assume it is a generic measure space — a triple consisting of $X \subset L^2(\mathbb{R})$, a collection Σ of all reasonable subsets of X called a σ-algebra, and a measure μ (see Appendix C.4 for details). The only difference from a probability space is that $\mu(X)$, the measure on the entire space, no longer has to be finite.

Combining the above ideas, we define CV designs as abstract measure spaces that average to the unnormalized symmetric-subspace projector.

Definition 4.3.1. *Let $X \subset L^2(\mathbb{R})$. The measure space (X, Σ, μ) is a* continuous-variable t-design *if*

$$\int_X (|\psi\rangle\langle\psi|)^{\otimes t} \, \mathrm{d}\mu(\psi) = \Pi_t, \tag{4.11}$$

where we use the weak (Pettis) integral. [1] *In other words, for all tuples* $a, b \in \mathbb{N}_0^t$,

$$\int_X \left(\prod_{i=1}^{t} \langle a_i|\psi\rangle\langle\psi|b_i\rangle \right) \mathrm{d}\mu(\psi) = \Pi_t(a;b) \ . \tag{4.12}$$

Definition 4.3.1 is a formalized version of the definition of CV state t-designs given in Ref. [79]. We note that Definition 4.3.1 bypasses the issue of defining a volume measure on the set of CV quantum states. We do not perform any integration on the set of all states and instead *require* a design to match the projector onto the symmetric subspace. This construction is illustrated in Fig. 4.1.

There is an alternative motivation for Definition 4.3.1 that we describe in detail in Appendix C.4.5. It is based on the following observation in finite dimensions. Integration over the set $\mathbb{CP}^{d-1}$ of d-dimensional quantum states is equivalent to integration over $\mathbb{C}^d$ with d independent zero-mean, unit-variance complex Gaussian distributions. The integration is over each of the d amplitudes of the quantum state with respect to the Gaussian measure, and the resulting state is then normalized. We can similarly put an infinite product of Gaussian measures on the space $\mathbb{C}^\infty$ and then define a CV t-design to be a measure space over $L^2(\mathbb{R})$ that matches integration over $\mathbb{C}^\infty$ for polynomials of degree t or less. We show in Appendix C.4.5 that this definition is equivalent to Definition 4.3.1.

Since $\Pi_1 = \mathbb{I}$, where $\mathbb{I}$ denotes the infinite-dimensional identity operator, any orthonormal basis for $L^2(\mathbb{R})$ or POVM is a CV 1-design. For example, the photon-number basis $|n\rangle$ satisfies $\sum_{n\in\mathbb{N}_0} |n\rangle\langle n| = \mathbb{I}$, which corresponds to a photon counting measurement. Coherent states $\{|\alpha\rangle\}$ also form a 1-design as they satisfy $\int_{\mathbb{C}} |\alpha\rangle\langle\alpha| \frac{\mathrm{d}^2\alpha}{\pi} = \mathbb{I}$, which corresponds to a heterodyne measurement. Finally, the eigenstates of $\cos(\phi)\hat{x} + \sin(\phi)\hat{p}$ form a 1-design, which corresponds to a homodyne measurement.

[1] The use of the weak integral in the definition of CV designs is well-motivated. For the purposes of designs, the weak (Pettis) integral is more natural than the strong (Bochner) integral because we are generally interested in averaged functions of ψ. Ultimately, we will prove that CV t-designs do not exist for $t \geq 2$, which immediately implies the result for the case of the strong integral as well.

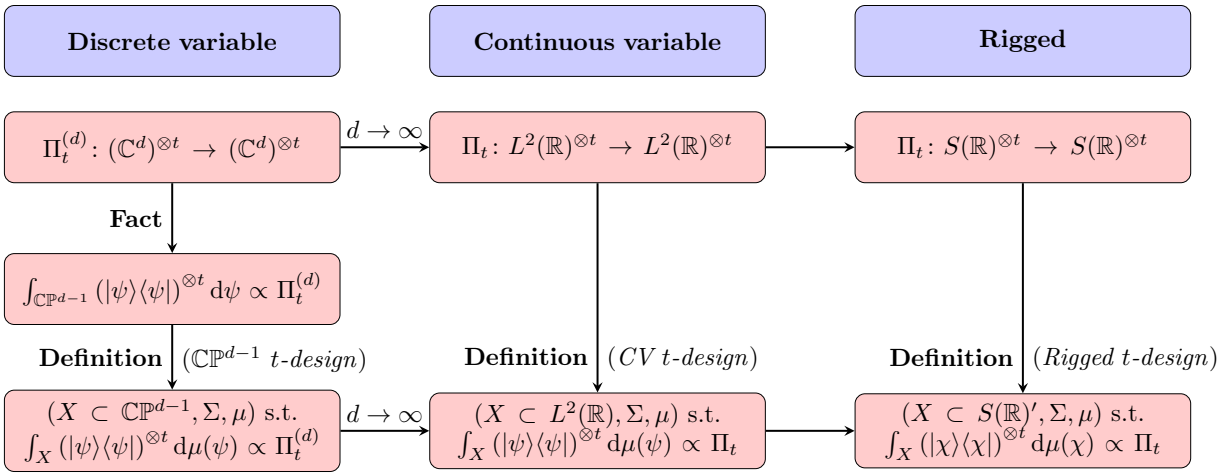

Figure 4.1: Sketch of definitions of finite-dimensional designs, continuous-variable (CV) designs and rigged designs. The key point is the absence of the middle block in the middle and right columns. A generalization of the middle block to the continuous-variable case is ill-defined, as discussed in Section 4.3. Therefore, to define CV designs, we simply skip the middle step, as discussed in Definition 4.3.1. An alternative characterization/definition of CV and rigged designs is described in Appendix C.4.5.

In Section 4.2, we argued that a complex-projective design on $\mathbb{C}^d$ gives rise to a simplex design. Similarly, in Appendix C.4.1, we prove that the existence of CV t-designs implies the existence of infinite-dimensional simplex t-designs. Here, we define a infinite-dimensional simplex design by starting with a finite-dimensional simplex integration over the unit-normalized Lebesgue measure and then removing the normalization requirement of the measure as we take the dimension to infinity.

By construction, a CV 1-design induces an infinite dimensional simplex 1-design by converting the amplitudes of a quantum state to probabilities via the Born rule. For example, the simplex design induced by the CV 1-design $\{|n\rangle \mid n \in \mathbb{N}_0\}$ is a set of probability distributions $\{p^{(n)} \mid n \in \mathbb{N}_0\}$. Here $p^{(n)} = (p_0^{(n)}, p_1^{(n)}, \dots)$ is a probability distribution over $\mathbb{N}_0$ defined as $p_i^{(n)} = \delta_{in}$.

As for $t > 1$ designs, we prove that *no set of CV states*, Gaussian or not, forms a CV t-design for any $t \geq 2$ (see Appendix C.4 for proofs).

Theorem 4.3.2. *For any $t \geq 2$, continuous-variable state t-designs do not exist.*

The non-existence of state ($t \geq 2$)-designs immediately implies non-existence of unitary ($t \geq 2$)-designs because their existence would imply the existence of state designs.

Corollary 4.3.3. *For any $t \geq 2$, continuous-variable unitary t-designs do not exist.*

To prove Theorem 4.3.2, we show that infinite-dimensional simplex t-designs do not exist for $t \geq 2$, and then invoke the connection between state and simplex designs described in Section 4.2. The non-existence of infinite-dimensional simplex designs can be understood as follows. All simplex ($t \geq 2$)-designs require at least one point near the centroid of the simplex. The centroid of a finite-dimensional simplex Δ^{d-1} is the point $(1/d, \ldots, 1/d)$. However, for the infinite-dimensional case, the centroid is no longer a valid point on the probability simplex. In the context of quantum states, this translates to the fact that uniform superpositions of all Fock states are not normalizable. We are therefore motivated to remove the requirement that elements of CV t-designs are normalized states.

Section 4.4: Rigged designs: definition & constructions

The non-existence of CV t-designs for $t > 1$ stems from the requirement that elements of said designs, according to Definition 4.3.1, belong to $L^2(\mathbb{R})$ and thus should have finite norm. We are therefore motivated to develop a new notion of CV designs that allows for non-normalizable states.

To include non-normalizable states in a CV design, we need to consider a set larger than $L^2(\mathbb{R})$. We consider the space of *tempered distributions*, denoted as $S(\mathbb{R})' \supset L^2(\mathbb{R})$, which contains infinitely squeezed position or momentum states as well as oscillator phase states [198–202, 205]. Despite being awkwardly called "states", these and other distributions may not be normalizable.

The use of distributions, whether for our purposes or for CV measurement protocols such as homodyne detection [226], is only well-defined for those CV states for which inner products

with tempered distributions are finite. This class consists of those states which admit finite expectation values of all powers of the occupation number operator $\hat{n} = \sum_{n\in\mathbb{N}_0} n\,|n\rangle\langle n|$, making up the *Schwartz space* $S(\mathbb{R}) \subset L^2(\mathbb{R})$ [243]. Together, the three spaces of interest make up the Gelfand triple $S(\mathbb{R}) \subset L^2(\mathbb{R}) \subset S(\mathbb{R})'$, the standard *rigged Hilbert space* for a quantum harmonic oscillator [80, 244, 245].

We modify Definition 4.3.1 to include tempered distributions. The motivation for our modification is summarized in Fig. 4.1.

Definition 4.4.1. *Let $X \subset S(\mathbb{R})'$. The measure space (X, Σ, μ) is called a* rigged t-design *if*

$$\int_X (|\chi\rangle\langle\chi|)^{\otimes t'} \,\mathrm{d}\mu(\chi) = \alpha_{t'}\Pi_{t'} \tag{4.13}$$

for all positive integers $t' \leq t$, where $\alpha_{t'} \in (0, \infty)$, where we use the weak (Pettis) integral. In other words, for all tuples $a, b \in \mathbb{N}_0^{t'}$,

$$\int_X \left(\prod_{i=1}^{t'} \langle a_i|\chi\rangle\langle\chi|b_i\rangle\right) \mathrm{d}\mu(\chi) = \alpha_{t'}\Pi_{t'}(a; b) \tag{4.14}$$

for $t \leq t'$.

Analogously to what is discussed below Definition 4.3.1, there is an alternative motivation Definition 4.4.1 that we describe in Appendix C.4.5. Recall that we described an equivalent definition of CV designs to be measure spaces over $L^2(\mathbb{R})$ that match integration over $\mathbb{C}^\infty$ with an infinite product of Gaussian measures. In Appendix C.4.5, we further show $\mathbb{C}^\infty \setminus S(\mathbb{R})'$ has measure zero in $\mathbb{C}^\infty$, so that the integration over $\mathbb{C}^\infty$ is equivalent to integration over $S(\mathbb{R})'$. It follows therefore that rigged t-designs exist for any $t \in \mathbb{N}$, since we can simply take the aforementioned measure space over $S(\mathbb{R})'$ to be our design. This design is however not desirable since it involves infinite-dimensional integration. We thus look for more manageable measure spaces that form rigged designs.

Inclusion of distributions circumvents the no-go Theorem 4.3.2 and allows us to construct

several examples of rigged 2-designs. Our first example consists of Fock states $\{|n\rangle\}_{n\in\mathbb{N}_0}$ and a family of distributions that we call *Kerred phase states* $|\theta\rangle_\varphi$ — tempered distributions defined informally as

$$|\theta\rangle_\varphi := \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{N}_0} \exp\left[\mathrm{i}(\theta n + \varphi n^2)\right] |n\rangle\,, \tag{4.15}$$

and formally as functionals mapping $|\psi\rangle \in S(\mathbb{R})$ to

$$\psi(\theta,\varphi) := {}_\varphi\langle\theta|\psi\rangle = \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{N}_0} \exp\left[-\mathrm{i}(\theta n + \varphi n^2)\right] \langle n|\psi\rangle. \tag{4.16}$$

The Kerred phase "states" consist of oscillator phase states [198–205], evolved up to some "time" θ under a Hamiltonian $\hat{n}^2$ associated with the optical Kerr effect. In Appendix C.4.3, we prove that

$$\frac{1}{2}\sum_{n\in\mathbb{N}_0} (|n\rangle\langle n|)^{\otimes t} + \frac{1}{2}\int_{-\pi}^{\pi} \mathrm{d}\theta \int_{-\pi}^{\pi} \mathrm{d}\varphi \left({}_\varphi|\theta\rangle\langle\theta|_\varphi\right)^{\otimes t} = \alpha_t \Pi_t \tag{4.17}$$

for $t = 1$ and $t = 2$, where $\alpha_1 = \pi + 1/2$ and $\alpha_2 = 1$.

To show that the above set is a design, we extend simplex and torus 2-designs to the rigged regime (see Appendix C.4.3 for details). The integration over the two phases $\{\theta, \varphi\}$ corresponds to a torus 2-design. The Fock states $|n\rangle$ correspond to extremal points of a simple simplex 2-design consisting of extremal points and the centroid in the finite-dimensional case, with the centroid vanishing in the infinite-dimensional case (as discussed in Section 4.3). By removing the normalization condition, we define an "non-normalizable centroid", which corresponds to a uniform superposition of Fock states $|\theta = 0\rangle_{\varphi=0}$. Combining such a state with the aforementioned torus 2-design gives the Kerred phase states.

Oscillator phase states are (left) eigenstates of the oscillator phase operator $Z = \sum_{n\in\mathbb{N}_0} |n+1\rangle\langle n|$ [198–205], an analogue of the oscillator raising operator but without the square-root factor. Both the phase and raising operators do not admit right eigenstates, but ±-

superpositions of each operator with its adjoint yield (anti-)Hermitian operators that admit well-known distributions as eigenstates. Superpositions of lowering and raising operators admit position and momentum states as eigenstates, respectively, while superpositions of the phase operator and its adjoint admit the *cosine* and *sine states*, respectively [199]. In Appendix C.4.3, we show that these two sets of states, when evolved under the Kerr Hamiltonian and combined with Fock states, make up two more examples of rigged 2-designs. More generally, since Z is unitarily related to $Ze^{i\omega}$ via a Fock-space rotation $e^{i\omega\hat{n}}$, eigenstates of a linear combination of $Ze^{i\omega}$ and its conjugate should similarly yield a distinct set of designs for any ω.

We do not provide constructions of useful rigged 3-designs. As shown with an example in Appendix C.4.3, not all simplex 2-designs can be extended to infinite dimensions. Thus, the difficulty in constructing a rigged 3-design lies is finding a simplex 3-design that is well-behaved enough to be extended to infinite dimensions. We leave this exciting open question for future work.

Section 4.5: Regularized Rigged Designs

Our rigged designs consist of non-normalizable states, but some applications require approximate versions of such designs that consist of physical quantum states. One way to approximate is to simply truncate the Fock space, corresponding to a *hard* or *maximum-energy cutoff*. This brings us back to finite dimensions, reducing rigged designs to ordinary quantum state designs. Another way, possible only with our infinite-dimensional formulation, is to impose a *soft* or *average-energy cutoff* that maintains the ability for states to have infinite support in Fock space. Both cutoffs can be encompassed in a general regularization formalism.

Let the *regularizer* R be a positive-semidefinite operator that yields a corresponding

"regularized projector"

$$\Pi_t^{(R)} \coloneqq R^{\otimes t}\Pi_t R^{\otimes t} \quad \text{such that} \quad \operatorname{Tr}\Pi_t^{(R)} < \infty \,. \tag{4.18}$$

The two aforementioned energy cutoffs correspond, respectively, to regularizers

$$R = \begin{cases} P_d \coloneqq \sum_{n=0}^{d-1} |n\rangle\langle n| & \text{hard cutoff, } d \in \mathbb{N}_0 \\ R_\beta \coloneqq \mathrm{e}^{-\beta\hat{n}} & \text{soft cutoff, } \beta > 0 \end{cases} , \tag{4.19}$$

but our formalism allows for more general R. We construct regularized designs by applying a regularizer to elements of a rigged design.

Suppose (X, Σ, μ) is a rigged t-design satisfying $\alpha_t = 1$. Regularization by an appropriate regularizer, such as R_β and P_d, converts X into a set of normalized states

$$Y \coloneqq \{|\psi\rangle = R\,|\chi\rangle \,/\, \|R\,|\chi\rangle\| \mid |\chi\rangle \in X\} \,, \tag{4.20}$$

with corresponding σ-algebra Σ_Y and measure

$$\mathrm{d}\nu(|\psi\rangle) = \mathrm{d}\mu(\chi) \cdot \|R\,|\chi\rangle\|^{2t} \,/\, \operatorname{Tr}\Pi_t^{(R)} \,. \tag{4.21}$$

These regularized designs average to $\Pi_t^{(R)} / \operatorname{Tr}\Pi_t^{(R)}$ instead of Π_t,

$$\begin{aligned} \int_Y (|\psi\rangle\langle\psi|)^{\otimes t}\,\mathrm{d}\nu(\psi) &= \int_X (R\,|\chi\rangle\langle\chi|\,R)^{\otimes t}\,\frac{\mathrm{d}\mu(\chi)}{\operatorname{Tr}\Pi_t^{(R)}} \\ &= \frac{\Pi_t^{(R)}}{\operatorname{Tr}\Pi_t^{(R)}} \,. \end{aligned} \tag{4.22}$$

The use of normalized states allows us to promote Y to a probability space. By taking the trace of both sides of Eq. (4.22) and applying assumption (4.18), we see that the measure ν is automatically normalized, $1 = \int_Y \langle\psi|\psi\rangle^t\,\mathrm{d}\nu(\psi) = \nu(Y)$. This allows us to express

$\int_Y (\cdot)\,\mathrm{d}\nu(\psi)$ as a statistical expectation $\mathbb{E}_{\psi\in Y}(\cdot)$ of states in Y sampled according to the distribution defined by ν (see Appendix C.4.4 for details). This yields the definition below, with a related definition of regularized CV unitary designs provided in Appendix C.4.7.

Definition 4.5.1. *Let $Y \subset L^2(\mathbb{R})$. The probability space (Y, Σ_Y, ν) is called an* R-regularized rigged t-design *if*

$$\mathop{\mathbb{E}}_{\psi\in Y} (|\psi\rangle\langle\psi|)^{\otimes t} = \frac{\Pi_t^{(R)}}{\operatorname{Tr}\Pi_t^{(R)}} \ . \tag{4.23}$$

Analogous to the discussion below Definitions 4.3.1 and 4.4.1, there is again an alternative motivation for Definition 4.5.1 that we detail in Appendix C.4.5. Recall that an infinite product of zero-mean, unit-variance Gaussian measures on $\mathbb{C}^\infty$ forms a rigged t-design. We show in Appendix C.4.5 that if the variance of the i^{th} measure is instead λ_i such that the diagonal operator $R_{ii} = \lambda_i$ is trace class ($\sum_i \lambda_i < \infty$), then the resulting measure space is a $\sqrt{R}$-regularized rigged t-design for any $t \in \mathbb{N}$. Importantly, with this measure, $\mathbb{C}^\infty \setminus L^2(\mathbb{R})$ has measure zero in $\mathbb{C}^\infty$ so that the design is a measure space over $L^2(\mathbb{R})$ as desired.

We now consider regularizing the Fock-state and Kerred phase-state design (4.17) with the soft-energy cutoff $R = R_\beta = \mathrm{e}^{-\beta\hat{n}}$ (4.19). Denote the regularized Kerred phase states (*a.k.a.* phase coherent states [246]) as

$$|\tilde{\theta}\rangle_\varphi := \frac{R_\beta\,|\theta\rangle_\varphi}{\|R_\beta\,|\theta\rangle_\varphi\|} = \sqrt{1-\mathrm{e}^{-2\beta}}\sum_{n=0}^{\infty} \mathrm{e}^{-\beta n + \mathrm{i}\theta n + \mathrm{i}\varphi n^2}\,|n\rangle \ , \tag{4.24}$$

such that $\||\tilde{\theta}\rangle_\varphi\| = 1$. Then it follows that

$$\sum_{n\in\mathbb{N}_0} w_n\,|n\rangle\langle n|^{\otimes 2} + f_\beta \int {}_\varphi|\tilde{\theta}\rangle\langle\tilde{\theta}|_\varphi^{\otimes 2}\,\mathrm{d}\theta\,\mathrm{d}\varphi = \frac{\Pi_2^{(R_\beta)}}{\operatorname{Tr}\Pi_2^{(R_\beta)}}, \tag{4.25}$$

where the limit of integration for both θ and φ is $[-\pi,\pi]$, $f_\beta := \cosh\beta/(\mathrm{e}^\beta(2\pi)^2)$, and $w_n = \frac{4\sinh^2\beta\cosh\beta}{\mathrm{e}^{\beta(4n+3)}}$.

Given a fixed average-energy constraint E, it is natural to define an energy-constrained

state design consisting of states $\{\psi\}$, such that each state in the design satisfies $\mathrm{Tr}(\hat{n}\psi) \leq E$. Our regularized-rigged design does not satisfy this condition explicitly as it contains Fock states $|n\rangle$ with $n > E$, as shown in Eq. (4.25). However, the contribution of large n terms is suppressed by the w_n coefficient in Eq. (4.25), which decays exponentially with n. Thus, our regularized-rigged designs are good approximations to energy-constrained state designs. It is an interesting open question to further develop the framework for energy-constrained state designs; we make some headway in this direction by formulating constrained integration in Appendix C.3.6, albeit for the finite-dimensional case.

As another example, we show in Appendix C.4.6 that displaced Fock states form regularized 2-designs for which the regularizer is the maximum-energy cutoff from Eq. (4.19), granted that we are allowed to use negative weights in the combination.

An important feature not inherited from the finite-dimensional case is that, in general, an R-regularized rigged t-design is not an R-regularized rigged $(t-1)$-design. For example, if Y is an R-regularized rigged 2-design, then

$$\mathop{\mathbb{E}}_{\psi\in Y} |\psi\rangle\langle\psi| = \frac{\Pi_1^{(R)}}{2\,\mathrm{Tr}\,\Pi_2^{(R)}} \left((\mathrm{Tr}\,R^2)\mathbb{I} + R^2\right) \neq \frac{\Pi_1^{(R)}}{\mathrm{Tr}\,\Pi_1^{(R)}}, \tag{4.26}$$

violating Eq. (4.23) for $t = 1$. Similarly, if Y is an R-regularized rigged 3-design, then

$$\begin{aligned}\mathop{\mathbb{E}}_{\psi\in Y} (|\psi\rangle\langle\psi|)^{\otimes 2} =& \frac{\Pi_2^{(R)}}{3\,\mathrm{Tr}\,\Pi_3^{(R)}} \\ &\times \Big((\mathrm{Tr}\,R^2)\mathbb{I}\otimes\mathbb{I} + \mathbb{I}\otimes R^2 + R^2\otimes\mathbb{I}\Big)\end{aligned} \tag{4.27}$$

instead of $\Pi_2^{(R)}/\,\mathrm{Tr}\,\Pi_2^{(R)}$.

Notice that, as R gets closer to the identity in Eqs. (4.26) and (4.27), $\mathrm{Tr}\,R^2$ dominates the remaining terms. This behavior holds for general t. As described further in Appendix C.4.4, if Y is an R-regularized rigged t-design, then it is *almost* an R-regularized rigged $(t-1)$-design

in the sense that

$$\mathop{\mathbb{E}}_{\psi\in Y}(|\psi\rangle\langle\psi|)^{\otimes(t-1)} \approx \frac{\Pi_{t-1}^{(R)}}{\operatorname{Tr}\Pi_{t-1}^{(R)}}\left(1+\mathcal{O}\left(1/\operatorname{Tr}R^2\right)\right). \tag{4.28}$$

We conclude this section by generalizing the *frame potential* from finite dimensions [40, 41, 214] to regularized rigged t-designs. For a positive definite (and therefore invertible) regularizer R, we define the frame potential of an ensemble $\mathcal{G}$ over unit vectors in $L^2(\mathbb{R})$ to be

$$V_t^{(R)}(\mathcal{G}) \coloneqq \mathop{\mathbb{E}}_{\psi,\phi\in\mathcal{G}} \left|\langle\psi| R^{-1} |\phi\rangle\right|^{2t}. \tag{4.29}$$

In Appendix C.4.4, we prove the following proposition regarding the frame potential.

Proposition 4.5.2. *Let R be positive definite. For any ensemble $\mathcal{G}$,*

$$V_t^{(R)}(\mathcal{G}) \geq \frac{1}{\operatorname{Tr}\Pi_t^{(R)}}, \tag{4.30}$$

with equality if and only if $\mathcal{G}$ is an R-regularized rigged t-design.

Note the presence of the R^{-1} in Eq. (4.29). We will see something similar in Section 4.6.3, where we find that finite-dimensional formulas nicely generalize to infinite-dimensions by introducing factors of R^{-1} to R-regularized rigged designs.

Section 4.6: Applications of rigged designs

In Section 4.6.1, we develop a shadow tomography protocol for CV systems based on rigged CV designs. In Section 4.6.2, we show how such CV rigged shadows can be used for entanglement verification. In Section 4.6.3, we develop the notion of the average fidelity of a CV quantum channel by using regularized-rigged 2-designs, relate this fidelity to the CV entanglement fidelity, and compare various fidelities for the case of the pure loss channel.

Subsection 4.6.1: Design-based CV shadows

The main idea behind finite-dimensional shadow tomography protocols is to perform random measurements of an unknown state to create classical snapshots through which many properties of the same unknown state can be efficiently predicted [49–52]. One can perform $\mathcal{O}(\log M)$ random measurements of an unknown state ρ to accurately predict the expectation values of M different observables with high probability. Each measurement for one such protocol yields a *shadow* of the form $3|e\rangle\langle e| - I$ on each qubit of the system, where e is an eigenstate of one of the qubit Pauli matrices, and I is the two-by-two identity. The number of measurements needed is independent of the dimension of the Hilbert space, a property that can be proven using designs [50].

Shadow tomography can be framed in terms of informationally-complete positive operator-valued measures (POVMs), which include quantum state $(t \geq 2)$-designs [52]. The concept of POVMs extends to infinite dimensions in such a powerful way that POVM elements can even be tempered distributions [205], [222, Appx. A]. Such POVMs are widely used. For example, homodyne measurements correspond to measurements in the position-state POVM or its rotated counterparts [226], while measuring in the phase-state POVM is optimal for determining the angle induced by a phase-space rotation [205, Sec. 3.9].

Utilizing rigged designs as infinite-dimensional POVMs, we develop a CV shadow tomography protocol (see Appendix C.5.1 for more details). Here, our goal is to determine $\langle \mathcal{O}_j \rangle := \mathrm{Tr}(\rho \mathcal{O}_j)$ for a collection of M single-mode observables $\mathcal{O}_1, \ldots, \mathcal{O}_M$, where ρ is an unknown infinite-dimensional state which we can access on a quantum device. We first describe a protocol utilizing a rigged 3-design, and then describe a protocol utilizing a rigged 2-design such as the one constructed in Eq. (4.17). The former case is slightly more general and easier to describe, but we have not yet constructed useful rigged 3-designs. We leave this question for future work.

CV shadows with rigged three-designs

Let (X, Σ, μ) denote a rigged 3-design, which implies that $\int_X (|\chi\rangle\langle\chi|)^{\otimes t} \, \mathrm{d}\mu(\chi) = \alpha_t \Pi_t$ for $t \in \{1, 2, 3\}$ and $\alpha_t \in (0, \infty)$. Without loss of generality, let $\alpha_1 = 1$, rescaling the measure μ if necessary. Then, it follows that the design resolves the identity,

$$\int_X |\chi\rangle\langle\chi| \, \mathrm{d}\mu(\chi) = \mathbb{I} \, , \tag{4.31}$$

and therefore, $\nu \colon A \mapsto \int_A |\chi\rangle\langle\chi| \, \mathrm{d}\mu(\chi)$ is a POVM.

Recall that a POVM maps subsets, which correspond to collections of measurement outcomes, to bounded, nonnegative self-adjoint operators (see Appendix C.2.1 for a measure theory review and Appendix C.5.1 for a short review on POVMs). Sampling from such a POVM results in sampling measurement outcomes from the probability measure $\mu' \colon A \mapsto \mathrm{Tr}(\rho\nu(A))$. We denote the measurement outcome corresponding to χ as $c(\chi)$ that we then store on a classical computer.

Suppose that we measure N times from μ', resulting in outputs $\{c(\chi_1), \ldots, c(\chi_N)\}$. Each of these outputs corresponds to a CV shadow

$$\hat{\rho}_i \coloneqq \frac{2}{\alpha_2} |\chi_i\rangle\langle\chi_i| - \mathbb{I} \, . \tag{4.32}$$

Note that $|\chi_i\rangle$ is not generally a physical quantum state but instead a tempered distribution. Fortunately, this is unimportant, since we are simply storing a description of $|\chi_i\rangle$ on a classical computer.

Using the classical snapshot and the classical description of observables $\mathcal{O}_j$, one can compute

$$\hat{o}_j \coloneqq \frac{1}{N} \sum_{i=1}^{N} \mathrm{Tr}(\hat{\rho}_i \mathcal{O}_j) \, . \tag{4.33}$$

On average, this yields the right answer: by the rigged 2-design property of X, $\mathbb{E}[\hat{o}_j] = \langle\mathcal{O}_j\rangle$,

where the expectation value is taken over measurement outcomes. Moreover, convergence to the right answer depends only on the features of $\mathcal{O}_j$: using the rigged 3-design property of X, we find that $\mathrm{Var}(\hat{o}_j) = \mathcal{O}\left(\frac{(\mathrm{Tr}|\mathcal{O}_j|)^2}{N}\right)$ in the large-N limit (see Appendix C.5.1 for details). We perform the aforementioned procedure K times, resulting in a collection $C_j = \{\hat{o}_j^{(1)}, \ldots, \hat{o}_j^{(K)}\}$. Following Ref. [50, Thm. 1], for each j, the median of C_j is within ε of $\langle \mathcal{O}_j \rangle$ with probability at least $1 - \delta$ provided that

$$N = \mathcal{O}\left(\frac{1}{\varepsilon^2}(\max_j \mathrm{Tr}\,|\mathcal{O}_j|)^2\right), \tag{4.34a}$$

$$K = \mathcal{O}(\log(M/\delta))\,. \tag{4.34b}$$

In other words, using a shadow tomography procedure with a rigged 3-design, we can accurately determine the expectation values of M observables using only $\sim \log M$ measurements, provided that each observable $\mathcal{O}_j$ is reasonably well-behaved; that is, provided that $\max_j \mathrm{Tr}\,|\mathcal{O}_j|$ is not too large.

CV shadows with rigged two-designs

If we had only used a rigged 2-design in the above protocol, we would still have that $\mathbb{E}[\hat{o}_j] = \langle \mathcal{O}_j \rangle$. For certain observables $\mathcal{O}_j$, we can show that a rigged 2-design is sufficient to give reasonable bounds on the variance by following an analogous result in finite dimensions from Ref. [52].

As before, suppose we have a collection of N shadows $\hat{\rho}_1, \ldots, \hat{\rho}_N$ sampled from the POVM defined by the rigged 2-design, yielding estimates $\hat{o}_j$ (4.33). We pick observables that satisfy

$$c < \mathrm{Tr}(\hat{\rho}_i \mathcal{O}_j) < d \tag{4.35}$$

for some $c < d \in \mathbb{R}$ almost surely for every shadow $\hat{\rho}_i$. Then, to achieve a success probability

of at least $1-\delta$ and maximum additive error ε, we need only

$$N \geq \log\left(\frac{2M}{\delta}\right)\frac{(d-c)^2}{2\varepsilon^2} \tag{4.36}$$

shadows to determine $\langle \mathcal{O}_j \rangle$ for each $1 \leq j \leq M$.

For concreteness, we consider a simple example of the rigged 2-design shadow protocol. Let each observable $\mathcal{O}_j$ be of the form $\mathcal{O}_j = |a_j\rangle\langle b_j| + |b_j\rangle\langle a_j|$ for $a_j, b_j \in \mathbb{N}_0$. We use the rigged 2-design from Eq. (4.17) consisting of Fock states and Kerred phase states. The explicit sampling step for this procedure is worked out in Appendix C.5.1. Using the explicit form of $|\theta\rangle_\varphi$, it follows that for any possible shadow $\hat{\rho}_i$ coming from this design, $|\operatorname{Tr}(\hat{\rho}_i \mathcal{O}_j)| < 1/5$. Therefore, to determine the M observables $\{\mathcal{O}_j\}$ to a maximum additive error ε with success probability at least $1-\delta$, we need only

$$N \geq \log\left(\frac{2M}{\delta}\right)\frac{2}{25\varepsilon^2} \tag{4.37}$$

measurements.

Subsection 4.6.2: Entanglement verification

In finite dimensions, classical shadows of a quantum state allows for the checking of many entanglement witnesses on that state [50]. Indeed, the same result holds for design-based CV shadows.

From Ref. [247, Thm. 2.2], for infinite-dimensional states ρ, ρ is entangled if and only if there exists a finite-rank operator A and a real number α such that $\alpha + \operatorname{Tr}(\rho A) < 0$ and $\alpha + \operatorname{Tr}(\sigma A) \geq 0$ for all separable states σ. Since A is finite rank, the expectation value of A with respect to a rigged shadow is finite even though the rigged shadow is not a normalizable quantum state. Hence, the use of rigged shadows (obtained from very few measurements of ρ) allows one to test many candidate witnesses A in order to determine if ρ is entangled.

Subsection 4.6.3: Fidelities of CV quantum channels

We develop the notions of the average fidelity of a continuous-variable (CV) quantum channel as well as their relationship to the CV entanglement fidelity. Such notions require approximate (i.e., regularized) versions of our rigged designs. We work out the case of a general positive semi-definite regularizer R, but note that the reader should keep in mind the two physically relevant hard- and soft-energy cases (4.19), corresponding to $R = P_d \coloneqq \sum_{n=0}^{d-1} |n\rangle\langle n|$ and $R = R_\beta \coloneqq \mathrm{e}^{-\beta\hat{n}}$, respectively. Finally, we benchmark the performance of a displacement operation by evaluating various fidelities for the case of the loss channel in Section 4.6.3.

Average fidelity of CV quantum channels

In a d-dimensional Hilbert space, quantum states belong to a compact space $\mathbb{CP}^{d-1}$. Therefore, one can define quantities that are averaged over all quantum states. In particular, for a quantum channel $\mathcal{D}$, the *average channel fidelity* is defined as [47, 58–62]

$$\overline{F}(\mathcal{D}) \coloneqq \int_{\mathbb{CP}^{d-1}} \langle\psi|\, \mathcal{D}(|\psi\rangle\langle\psi|)\, |\psi\rangle \, \mathrm{d}\psi, \tag{4.38}$$

quantifying how close $\mathcal{D}$ is to an identity channel on average. Due to non-existence of a standard measure on infinite-dimensional space, as discussed in Section 4.3, this formula cannot be extended to CV systems.

Since there are exactly two copies of $|\psi\rangle\langle\psi|$ in the integrand for the average fidelity, the integral over all states can be substituted with an average over any state 2-design X using Eqs. (4.4-4.5),

$$\overline{F}(\mathcal{D}) = \mathop{\mathbb{E}}_{\psi\in X} \langle\psi|\, \mathcal{D}(|\psi\rangle\langle\psi|)\, |\psi\rangle \ . \tag{4.39}$$

The design provides a more manageable sample of states that is useful for estimating the average fidelity of quantum operations [47, 54–57]. This formula *can* be extended to infinite dimensions using normalized (i.e., regularized) versions of our rigged designs from Section 4.5.

Let Y denote a regularized-rigged 2-design with a general positive semi-definite regularizer R. There is more than one way to generalize the average fidelity from the finite-dimensional case, and we consider two average-fidelity quantities defined for a CV channel $\mathcal{D}$,

$$\overline{F}_1^{(R)}(\mathcal{D}) \coloneqq N_R \mathop{\mathbb{E}}_{\psi\in Y} \langle\psi| R^+\mathcal{D}(\psi)R^+ |\psi\rangle\,, \tag{4.40a}$$

$$\overline{F}_2^{(R)}(\mathcal{D}) \coloneqq \mathop{\mathbb{E}}_{\psi\in Y} \langle\psi| \mathcal{D}(\psi) |\psi\rangle \ \ , \tag{4.40b}$$

where we use the short-hand notation $\mathcal{D}(\psi) = \mathcal{D}(|\psi\rangle\langle\psi|)$, and where the constant $N_R = \frac{\operatorname{Tr} R^4+(\operatorname{Tr} R^2)^2}{\operatorname{Tr} R^2+(\operatorname{Tr} R)^2}$. The second quantity faithfully uses two copies of the normalized state projections $|\psi\rangle\langle\psi|$ sampled from the design, while the first can revert one copy back to its non-normalizable version using the Moore-Penrose inverse R^+ of the regularizer.

As a sanity check, let us employ a hard-energy cutoff and plug in the regularizer $R = P_d = \sum_{n=0}^{d-1} |n\rangle\langle n|$ from Eq. (4.19) into Eq. (4.40). This essentially recovers the finite-dimensional case. Since the Moore-Penrose inverse of a projector is itself, the two average-fidelity quantities are equal for this case. Moreover, if $\mathcal{D}$ is trace-preserving for states within the subspace defined by P_d, then $\overline{F}_1^{(P_d)}(\mathcal{D}) = \overline{F}_2^{(P_d)}(\mathcal{D}) = \overline{F}(\mathcal{D})$, recovering the finite-dimensional design-based average fidelity from Eq. (4.39).

In a setting relevant to CV states enjoying infinite support, such as coherent or squeezed states, one should consider a regularizer with no zero eigenvalues. We prove in Appendix C.5.2 that an R-regularized rigged 2-design is informationally-complete for states on the entire Fock space whenever R is invertible. Therefore, choosing $R = P_d$ may not be a good approximation of average fidelity over all CV states.

Average-to-entanglement fidelity relation

In the finite-dimensional case, the *entanglement fidelity* for a quantum channel $\mathcal{D}$ on $\mathbb{C}^d$ is [47, 58–60]

$$F_e(\mathcal{D}) \coloneqq \langle\phi| (\mathcal{I} \otimes \mathcal{D})(\phi) |\phi\rangle \ , \tag{4.41}$$

where $|\phi\rangle \coloneqq \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} |n\rangle \otimes |n\rangle$ denotes a maximally entangled state, and $\mathcal{I}$ is the identity channel. This fidelity quantifies how well entanglement with a reference system is preserved by $\mathcal{D}$. We refer the reader to Ref. [248, Ap. A] for a nice review of the utility of the entanglement fidelity. The entanglement fidelity is related to the average fidelity by the following simple formula [59, 60]

$$\overline{F}(\mathcal{D}) = \frac{dF_e(\mathcal{D}) + 1}{d+1} \ . \tag{4.42}$$

We can similarly relate our average-fidelity relations (4.40) to a CV version of entanglement fidelity.

Maximally entangled states become non-normalizable as $d \to \infty$, meaning that CV versions of such states also have to be regularized in order to define an analogous fidelity. We require that R be diagonal in the Fock-state basis and define the regularized state

$$|\phi_R\rangle \coloneqq \frac{1}{\sqrt{\operatorname{Tr} R}} (R^{1/4} \otimes R^{1/4}) \sum_{n \in \mathbb{N}_0} |n\rangle \otimes |n\rangle \ , \tag{4.43}$$

a purification of the *regularizer state*

$$\rho_R \coloneqq \operatorname{Tr}_2(|\phi_R\rangle\langle\phi_R|) = R/\operatorname{Tr} R \ . \tag{4.44}$$

The *R-regularized CV entanglement fidelity* of a channel $\mathcal{D}$ is then

$$F_e^{(R)}(\mathcal{D}) \coloneqq \langle\phi_R| (\mathcal{I} \otimes \mathcal{D})(\phi_R) |\phi_R\rangle \, . \tag{4.45}$$

In Appendix C.5.2, we show that both CV average-fidelity quantities from Eq. (4.40) are related to the CV entanglement fidelity (4.45) as

$$\overline{F}_1^{(R)}(\mathcal{D}) = \frac{d_R F_e^{(R)}(\mathcal{D}) + \mathrm{Tr}[\mathcal{D}(\rho_{R^2})RR^+]}{d_R+1}, \tag{4.46a}$$

$$\overline{F}_2^{(\sqrt{R})}(\mathcal{D}) = \frac{d_R F_e^{(R)}(\mathcal{D}) + d_R\,\mathrm{Tr}[\mathcal{D}(\rho_R)\rho_R]}{d_R+1}. \tag{4.46b}$$

Since we are assuming R is diagonal, RR^+ is simply a projector onto the subspace for which R has support. For invertible R, this subspace is the whole space so that $RR^+ = \mathbb{I}$, and therefore Eq. (4.46a) yields a CV generalization of the finite-dimensional average-to-entanglement fidelity relation (4.42):

$$\overline{F}_1^{(R)}(\mathcal{D}) = \frac{d_R F_e^{(R)}(\mathcal{D}) + 1}{d_R+1}, \tag{4.47}$$

where the *effective dimension* dictated by the regularizer is the inverse purity of the regularizer state (4.44),

$$d_R \coloneqq 1/\,\mathrm{Tr}\,\rho_R^2 = (\mathrm{Tr}\,R)^2/\,\mathrm{Tr}\,R^2\ . \tag{4.48}$$

This effective dimension in the infinite-dimensional case plays the role of, and reduces to, the actual dimension in the finite-dimensional case.

The above general formulation reduces to a more physically relevant one when the soft-energy cutoff $R = R_\beta = \mathrm{e}^{-\beta\hat{n}}$ (4.19) is used as the regularizer. The state $|\phi_R\rangle$ (4.43) becomes a Gaussian two-mode squeezed vacuum state (*a.k.a.* thermofield double) with squeezing parameter $r = \log[\frac{1+\mathrm{e}^{-\beta/2}}{\sqrt{1-\mathrm{e}^{-\beta}}}]$ [138], while the regularizer state (4.44) becomes a thermal state whose "inverse temperature" $\beta > 0$ sets the energy scale of states involved in the regularization. The effective dimension (4.48) becomes

$$d_R = 2\,\mathrm{Tr}(\rho_R\hat{n}) + 1\ , \tag{4.49}$$

directly related to the mean energy of the thermal state. Similar energy-dependent factors also replace dimensions in studies of uniform continuity for quantum entropies [218] as well as bounds on energy-constrained capacities of Gaussian channels [219]. Since $|\phi_R\rangle$ is a Gaussian state, the corresponding CV entanglement fidelity may be extractable via reasonable experimental protocols [228, 229].

As for more general R, we emphasize that Eqs. (4.46) and (4.47) only hold as is when R is diagonal in the $\{|n\rangle\}$ basis even though Eq. (4.40) is well-defined even when R is not diagonal. Of course, one has the freedom to arbitrarily choose the basis with respect to which the CV entanglement fidelity is defined, so R being diagonal is not a substantial restriction.

Recall in Eq. (4.29) and Proposition 4.5.2, we saw that introducing factors of R^{-1} into a definition of frame potential resulted in finite-dimensional formulae nicely generalizing to infinite-dimensions. We again see this effect present in Eq. (4.47). The definition of $\overline{F}_1^{(R)}$ utilizes factors of R^{-1} while $\overline{F}_2^{(R)}$ does not. As a consequence, the finite-dimensional relation (4.42) involving $\overline{F}$ and F_e very closely matches the infinite-dimensional relation (4.47) involving $\overline{F}_1^{(R)}$ and $F_e^{(R)}$, whereas the relation involving $\overline{F}_2^{(R)}$ and $F_e^{(R)}$ (4.46b) contains a factor not present in the finite-dimensional case.

Fidelity benchmarks for displacement operations

We compare the fidelity quantities introduced in this section to known quantities for the case of the pure loss channel, $\mathcal{D} = \mathcal{L}^\kappa$ [249], with transmissitivity $\kappa \in [0, 1]$. This case is relevant to benchmarking the performance of displacement operations that are implemented via a non-ideal two-mode beam-splitter, with the transmissivity characterizing the degree of nonideality [232, 250]. All quantities described below are computed analytically in Appendix C.5.2.

In order to put all quantities on as equal of a footing, we set them to be a function of a fixed energy scale $\bar{n}$ using the following convention (with other choices possible). For the soft- and hard-energy regularizers, $R_\beta = e^{-\beta\hat{n}}$ and $P_d = \sum_{n=0}^{d-1} |n\rangle\langle n|$ (4.19), respectively, we

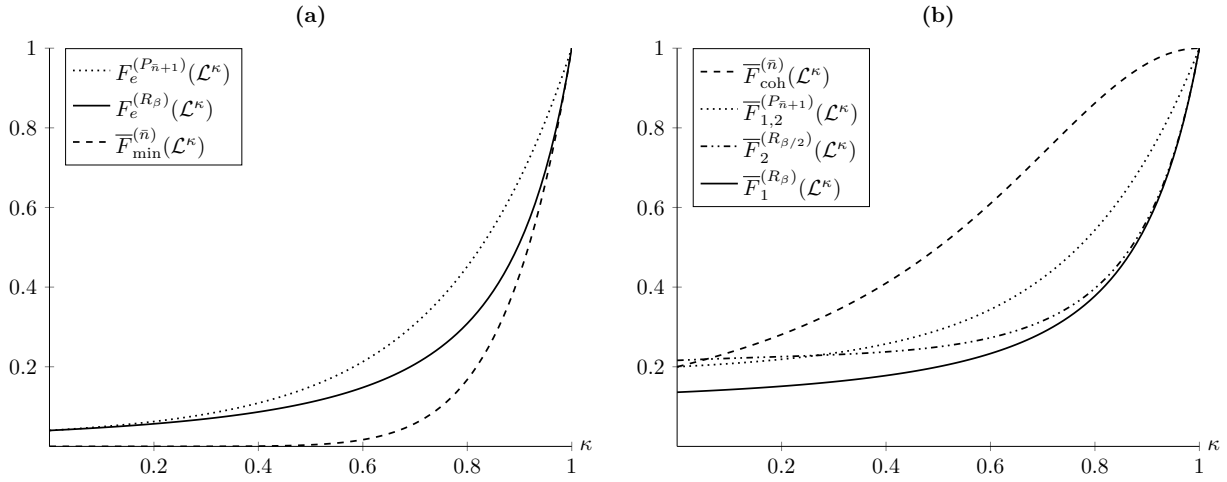

Figure 4.2: Various fidelity benchmarks for the pure-loss channel $\mathcal{L}^\kappa$ plotted vs the channel's transmissitivity κ, with the energy-constrained parameter $\bar{n} = 4$, and all other fidelity parameters being functions of $\bar{n}$ according to Eq. (4.50). **(a)** Comparison of fidelities that utilize a reference mode: the CV entanglement fidelity F_e (4.45) with soft- and hard-energy constraints (4.19) as well as the minimum energy-constrained entanglement fidelity $F_{\min}$ (4.51). **(b)** Comparison of our three average-fidelity quantities — the soft-energy constrained average fidelities $\overline{F}_1^{(R_\beta)}$ (4.46a) and $\overline{F}_2^{(R_{\beta/2})}$ (4.46b) as well as the hard-energy constrained case $\overline{F}_{1,2}^{(P_d)}$ — with the fidelity $\overline{F}_{\text{coh}}^{\bar{n}}$ (4.52) calculated by averaging over an ensemble of coherent states. The qualitatively different behavior of the coherent-state fidelity suggests that it may not be a good approximation to averages over CV states.

set

$$\beta = \log(1 + 1/\bar{n}) \qquad \text{and} \qquad d = \lfloor \bar{n} \rfloor + 1 \ . \tag{4.50}$$

The soft-energy cutoff then corresponds to an average energy of $\bar{n}$ for the regularizer thermal state (4.44) and an effective dimension $d_R = 2\bar{n} + 1$ (4.49).

Our first comparison is between all fidelities that utilize a reference mode. This comparison is between the CV entanglement fidelity (4.45), with either soft- or hard-energy regularization, and its minimum energy-constrained version [232, 251]

$$\overline{F}^{(\bar{n})}_{\min}(\mathcal{L}^\kappa) \coloneqq \min_{\psi_{EA}:\mathrm{Tr}(\hat{n}_A \psi_A) \leq \bar{n}} \langle \psi | \, \mathcal{L}^\kappa(\psi_{EA}) \, | \psi \rangle_{EA} \ , \tag{4.51}$$

consisting of an optimization of the CV entanglement fidelity over all input states whose average energy on the mode acted on by the channel is bounded by $\bar{n}$.

The three reference-mode fidelities $\{F_e^{(P_d)}, F_e^{(R_\beta)}, F^{(\bar{n})}_{\min}\}$ are plotted for $\bar{n} = 4$ and all transmissivities $\kappa \in [0, 1]$ in Fig. 4.2(a). All quantities decrease in similar fashion with decreasing transmissitivy, with the soft-energy fidelity following the scaling of the minimum case slightly better than the hard-energy fidelity near unity transmissitivity. Due to the parameterization picked in Eq. (4.50), the entanglement fidelities for the two energy constraints are equal for zero transmissivity, $F_e^{(P_d)} = F_e^{(R_\beta)} = 1/(\bar{n} + 1)^2$ at $\kappa = 0$.

Our second comparison is between fidelities that do not utilize a reference mode. This set includes both of our CV average fidelities from (4.40), each with either a soft- or a hard-energy constraint. These are related to the entanglement fidelity of a CV channel via Eqs. (4.46a) and (4.46b), respectively. Since the pure-loss channel is trace preserving on the subspace defined by P_d, two of these four fidelities are equal in the case of the hard-energy constraint, $\overline{F}_1^{(P_d)} = \overline{F}_2^{(P_d)} =: \overline{F}_{1,2}^{(P_d)}$. This comparison also includes the average fidelity of the

pure-loss channel over an ensemble of coherent states,

$$\overline{F}^{(\bar{n})}_{\text{coh}}(\mathcal{L}^{\kappa}) \coloneqq \int_{\mathbb{C}} p(\alpha) \langle\alpha| \mathcal{L}^{\kappa}(\alpha) |\alpha\rangle \, \mathrm{d}^2\alpha, \tag{4.52}$$

where $|\alpha\rangle$ denotes the coherent state specified by $\alpha \in \mathbb{C}$. We choose the density function to be $p(\alpha) = \frac{1}{\pi\bar{n}} \mathrm{e}^{-|\alpha|^2/\bar{n}}$ to ensure that the average occupation number of the ensemble of coherent states is $\int_{\mathbb{C}} p(\alpha) |\alpha|^2 \, \mathrm{d}^2\alpha = \bar{n}$.

The four average-fidelity quantities $\{\overline{F}^{(P_d)}_{1,2}, \overline{F}^{(R_\beta)}_{1}, \overline{F}^{(R_{\beta/2})}_{2}, \overline{F}^{(\bar{n})}_{\text{coh}}\}$ are plotted for $\bar{n} = 4$ and all transmissivities κ in Fig. 4.2(b). Note that the average fidelity over an ensemble of coherent states does not qualitatively match the other fidelities. In particular, the concavity of $\overline{F}_{\text{coh}}$ near unity transmissivity is different from the other fidelity quantities. This may be related to the fact that an ensemble of coherent states only forms a CV 1-design, whereas the other fidelities are defined with respect to various notions of 2-designs. This result suggests that the coherent-state average may not be a useful approximation for an average over all CV states.

Section 4.7: Conclusion

In this work, we study quantum state designs in finite and infinite dimensions. In finite dimensions, we review a method for constructing complex-projective designs using simplex and torus designs. In particular, we establish a relationship between torus designs and complete sets of mutually unbiased bases.

We then prove a no-go theorem implying that a naïve extension of the definition of state designs to infinite dimensions fails. Similarly, we prove that CV unitary t-designs do not exist for any $t \geq 2$. Prior to our work, it was proven [79] (argued [153]) that Gaussian resources are not sufficient to form CV state (unitary) designs. Our no-go theorem establishes a stronger result implying that even non-Gaussian resources are not sufficient to form CV designs.

The lack of CV designs is due to a restriction to using only normalizable states. We successfully extend the notion of state designs to infinite dimensions by proposing a new definition of CV state designs using non-normalizable states. These non-normalizable states belong to a rigged Hilbert space, and we provide various constructions of such rigged 2-designs consisting of Fock states and oscillator phase states [198–202, 205] subject to Kerr-Hamiltonian evolution.

As an application of rigged designs, we extended the formalism of shadow tomography [49–52] to CV systems. We show that our rigged 2-designs and, if useful ones exist, rigged 3-designs can yield efficient shadow-based protocols. It is an interesting direction to experimentally implement our design-based CV shadow tomography protocol based on rigged 2-designs and compare it with other protocols based on homodyne or heterodyne measurements [226], which can also be formulated within a shadow-like framework (albeit without the use of designs) [207]. The POVMs defined by the rigged 2-designs that we constructed are highly non-Gaussian. It is an exciting open theoretical and experimental direction to develop techniques to measure from such POVMs.

We construct approximate CV designs by regularizing the elements of rigged designs. These regularized-rigged designs consist of physical quantum states and therefore can be used to define information-theoretic quantities, such as fidelities, for CV quantum channels. In particular, we define various notions of the average fidelity of a CV channel. We then establish a relation between the average fidelity and the entanglement fidelity of a CV channel. Our result is a natural generalization of finite-dimensional formulas [13], where the dimension is replaced by the effective dimension that depends on the mean energy of the input state to the channel. It is an interesting open question to develop efficient experimental methods to prepare states belonging to regularized-rigged designs introduced in our work. On the theory side, it may be interesting to determine a relationship between the energy-constrained diamond distance [220, 221] and the average fidelity introduced in our work.

As discussed in Section 4.6.3, an important application of regularized-rigged designs is to estimate the average fidelity between an ideal unitary and its experimental approximations. We emphasize again that our results are applicable directly when analytical expressions of an ideal unitary gate and its experimental approximation are known. Instead of estimating the average fidelity over a subset of states such as coherent states, one can calculate a good approximation of the average fidelity over *all* states using our regularized-rigged designs.

We construct rigged and regularized-rigged CV state 2-designs, leaving the interesting question of constructing useful CV state t-designs for $t \geq 3$ to future work. Another interesting direction is to develop the notion of energy-constrained CV state designs, where each state in the design satisfies a fixed energy constraint. Our regularized-rigged state designs are good approximations of energy-constrained CV state designs.

Our rigged designs are defined on the Hilbert space $L^2(\mathbb{R})$ of a single mode, but can formally be mapped into any other countably infinite Hilbert space because all such spaces are isomorphic. A mapping like this from the single-mode space to the space $L^2(\mathbb{R}^n)$ of multiple modes is likely to be physically obscure. An interesting future topic would be to develop designs for other spaces, such as multiple modes, rotors and rigid bodies [216], using states natural to those spaces. For example, we anticipate that designs similar to our Kerred phase-state designs can be formulated for the space of the planar rotor, $L^2(\mathrm{U}(1))$ [216, Sec. IV.B]. Similarly, cross-Kerr interactions [206] may provide a recipe for rigged designs for multiple modes.

We also prove that CV unitary t-designs do not exist for any $t \geq 2$. A natural research question is whether, similar to rigged CV state designs, there exists a reasonable notion of CV operator designs. We introduce one such notion in this work, leaving the interesting and important question of how to construct such designs to future work.

Finally, another interesting avenue to explore is that of designs for function spaces. In Appendix C.4.5, we showed how our rigged designs can be interpreted as designs over infinite-

dimensional function spaces. Can this theory be further generalized to other functional integrals, such as e.g. path integrals? In particular, in field theories, one is typically interested in correlators (i.e. polynomials in the fields) of various degrees; a t-design is therefore a space of fields that match all correlators up to degree t. Can designs be defined and used in this context? Refs. [252–255], which contain a small number of cubature rules for Weiner integrals, may be a useful place to start.

Chapter 5: Continuous-variable designs and design-based shadow tomography from random lattices

Section 5.1: Introduction

In Chapter 4, we defined rigged t-designs and found some single mode examples of rigged 2-designs. In this chapter, we find a natural class of states that form multimode rigged 2-designs.

We denote the ensemble of all n-mode Gottesman-Kitaev-Preskill (GKP) states [215, 256, 257] for all lattices by

$$\mathcal{X}_n = \{|\Lambda; \boldsymbol{\alpha}\rangle = D(\boldsymbol{\alpha})\,|\Lambda\rangle\,,\ \Lambda \in Y_n,\ \boldsymbol{\alpha} \in \mathcal{P}(\Lambda)\}\,. \tag{5.1}$$

Here, $|\Lambda\rangle$ is an equal superposition of coherent states at all points in the symplectically self-dual lattice $\Lambda = \Lambda^\perp \subset \mathbb{R}^{2n}$, and Y_n is the set of all such lattices with its unique unit-normalized Haar measure. This state is displaced by all possible complex vectors $\boldsymbol{\alpha}$ in the lattice's fundamental domain $\mathcal{P}(\Lambda)$ — the unit cell centered at the origin — via displacement operators $D(\boldsymbol{\alpha})$. Each lattice is defined by a symplectic matrix, and the space $Y_n = \mathrm{Sp}_{2n}(\mathbb{Z})\backslash\mathrm{Sp}_{2n}(\mathbb{R})$ of all lattices is parameterized by all such matrices, up to basis transformations done by integer-valued symplectic matrices. This space happens to be compact and comes with its own Haar measure, $\mu(\Lambda)$ [258, 259].

In Sections 5.2 and 5.3, we show two different ways of proving the following theorem.

Theorem 5.1.1 (GKP states form a rigged 2-design)**.** *The ensemble $\mathcal{X}_n$ forms a rigged*

2-*design with respect to the Haar measure over the space of symplectic lattices* Y_n *and the uniform measure over the fundamental domain* $\mathcal{P}(\Lambda)$.

A key necessary lemma is the following.

Lemma 5.1.2 (Symplectic mean value formula [258], [260, Thm. 2])**.** *Let* $f : \mathbb{R}^{2n} \to \mathbb{R}$ *be a (Riemann) integrable function such that the following quantities converge absolutely, and let* $\Lambda \in Y_n$ *denote a symplectic lattice. For*

$$F(\Lambda) = \sum_{\boldsymbol{\lambda} \in \Lambda \setminus \{0\}} f(\boldsymbol{\lambda}), \tag{5.2}$$

it holds that

$$\int_{Y_n} d\mu(\Lambda)\, F(\Lambda) = \int_{\mathbb{R}^{2n}} d\boldsymbol{x}\, f(\boldsymbol{x}). \tag{5.3}$$

Then, in Section 5.4, we discuss a generalization of shadow tomography similar to that discussed in Section 4.6.1 and Appendix C.5.1, which shows how the displaced GKP rigged 2-design can be used for shadow tomography.

Section 5.2: Continuous variable frame potentials

Building on Chapter 4, we examine the ensemble of GKP states defined in the main text through the lens of continuous variable *frame potentials*. Frame potentials are an important tool in the analysis of evenly distributed states and we discuss their interpretation further below. To maintain well-defined expressions throughout, we need to work with appropriately regularized quantities, regularized through a Hermitian regularization operator R. The end of this section contains a general discussion on the frame potential and its interpretation.

Let Π_t be the projector onto the symmetric subspace of $L^2(\mathbb{R}^n)^{\otimes t}$, and for a positive definite Hermitian trace-class operator R, define $\Pi_t^{(R)} = R^{\otimes t} \Pi_t R^{\otimes t}$. For any measure space

$\mathcal{X}_n$ over $S'(\mathbb{R}^n)$, define

$$W_t^{(R)}(\mathcal{X}_n) \coloneqq \int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X)\, d\mu_{\mathcal{X}_n}(Y)\ |\langle X|R|Y\rangle|^{2t} \tag{5.4a}$$

$$X_t^{(R)}(\mathcal{X}_n) \coloneqq \int_{\mathcal{X}_n} d\mu_{\mathcal{X}_n}(X)\, \langle X|R^2|X\rangle^t \tag{5.4b}$$

$$\mathcal{F}_t^{(R)}(\mathcal{X}_n, a_t) \coloneqq W_t^{(R)}(\mathcal{X}_n) - 2a_t X_t^{(R)}(\mathcal{X}_n) + a_t^2 \operatorname{Tr}\left[\Pi_t^{(R)}\right]. \tag{5.4c}$$

The following proposition is a modification of Proposition 4.5.2.

Proposition 5.2.1. *For any ensemble $\mathcal{X}_n$ and any positive real number a_t, $\mathcal{F}_t^{(R)}(\mathcal{X}_n, a_t) \geq 0$. Furthermore, there exists positive real numbers $a_{t'}$ such that $\mathcal{F}_{t'}^{(R)}(\mathcal{X}_n, a_{t'}) = 0$ for each $t' = 1, 2, \ldots, t$ if and only if $\mathcal{X}_n$ is a rigged t-design.*

There are a few easy facts that we need:

$$\operatorname{Tr}\left[\Pi_1^{(R)}\right] = \operatorname{Tr}\left[R^2\right] \tag{5.5a}$$

$$\operatorname{Tr}\left[\Pi_2^{(R)}\right] = \frac{1}{2}\left(\operatorname{Tr}\left[R^2\right]^2 + \operatorname{Tr}\left[R^4\right]\right) \tag{5.5b}$$

$$\operatorname{Tr}\left[\Pi_3^{(R)}\right] = \frac{1}{6}\left(\operatorname{Tr}\left[R^2\right]^3 + 3\left(\operatorname{Tr}\left[R^4\right]\right)\left(\operatorname{Tr}\left[R^2\right]\right) + 2\operatorname{Tr}\left[R^6\right]\right). \tag{5.5c}$$

Finally, for the operator R, let $c_R(\boldsymbol{\alpha}) \coloneqq \operatorname{Tr}\left[D^\dagger(\boldsymbol{\alpha})\, R\right]$ denote its characteristic function. We have

$$c_{R^2}(0) = \operatorname{Tr}\left[R^2\right] = \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}\ |c_R(\alpha)|^2 \tag{5.6a}$$

$$\begin{aligned} c_{R^4}(0) &= \operatorname{Tr}\left[R^4\right] \\ &= \int_{\mathbb{R}^{2n}} d\boldsymbol{\alpha}_1 d\boldsymbol{\alpha}_2 d\boldsymbol{\alpha}_3\, c_R(\boldsymbol{\alpha}_1)\, c_R(\boldsymbol{\alpha}_2)\, c_R(\boldsymbol{\alpha}_3)\, c_R^*(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3)\, e^{-i\pi\left(\boldsymbol{\alpha}_1^T J \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_1^T J \boldsymbol{\alpha}_3 + \boldsymbol{\alpha}_2^T J \boldsymbol{\alpha}_3\right)}. \end{aligned} \tag{5.6b}$$

In a very long calculation we will now proceed to show the following theorem, using the proposition above. This calculation serves as independent verification that the ensemble $\mathcal{X}_n$

of GKP states discussed in the main text indeed yields a rigged 2-design.

Theorem 5.2.2. *The ensemble of GKP states $\mathcal{X}_n \subset \mathcal{S}'(\mathbb{R}^n)$ defined in the main text forms a rigged 2-design with $a_1 = 1$ and $a_2 = 2$.*

Proof. Using the frame potential, we can check if our ensemble $\mathcal{X}_n$ defined in the main text is a $t = 1, 2$-design. Specifically, we need to express $X_t^{(R)}(\mathcal{G})$ and $W_t^{(R)}(\mathcal{G})$ for each t in terms of $\mathrm{Tr}[R^2]$ and $\mathrm{Tr}[R^4]$.

Recall that elements of $\mathcal{X}_n$ are of the form $D(\boldsymbol{\alpha})\,|\Lambda\rangle$, where $|\Lambda\rangle\langle\Lambda| = \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} D(\boldsymbol{\lambda})$. In general, our computations will include expressions of the form

$$\begin{aligned}
\langle\Lambda|D^\dagger(\boldsymbol{\alpha})\,RD(\boldsymbol{\alpha}')\,|\Lambda'\rangle &= \sum_{\boldsymbol{\lambda}\in\Lambda}\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}\, c_R(\boldsymbol{\gamma})\, e^{i\Phi(\boldsymbol{\lambda})}\,\mathrm{Tr}\big[D(\boldsymbol{\lambda})\,D^\dagger(\boldsymbol{\alpha})\,D(\boldsymbol{\gamma})\,D(\boldsymbol{\alpha}')\big] && \text{(5.7a)}\\
&= \sum_{\boldsymbol{\lambda}\in\Lambda}\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}\, c_R(\boldsymbol{\gamma})\, e^{i\Phi(\boldsymbol{\lambda})} e^{i\pi\boldsymbol{\lambda}^T J\boldsymbol{\alpha} - i\pi\boldsymbol{\gamma}^T J\boldsymbol{\alpha}'}\underbrace{\mathrm{Tr}[D(\boldsymbol{\lambda}-\boldsymbol{\alpha})\,D(\boldsymbol{\gamma}+\boldsymbol{\alpha}')]}_{\delta(\boldsymbol{\gamma}-\boldsymbol{\alpha}+\boldsymbol{\alpha}'+\boldsymbol{\lambda})} && \text{(5.7b)}\\
&= \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} c_R(\boldsymbol{\alpha}-\boldsymbol{\alpha}'-\boldsymbol{\lambda})\, e^{i\pi\boldsymbol{\lambda}^T J\boldsymbol{\alpha} - i\pi(\boldsymbol{\alpha}-\boldsymbol{\alpha}'-\boldsymbol{\lambda})^T J\boldsymbol{\alpha}'}. && \text{(5.7c)}
\end{aligned}$$

If $\boldsymbol{\alpha} = \boldsymbol{\alpha}'$ this is

$$\langle\Lambda|D^\dagger(\boldsymbol{\alpha})\,RD(\boldsymbol{\alpha})\,|\Lambda\rangle = \sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} c_R(-\boldsymbol{\lambda})\, e^{i2\pi\boldsymbol{\lambda}^T J\boldsymbol{\alpha}}. \tag{5.8a}$$

We begin with $t = 1$. We compute

$$\begin{aligned}
X_1^{(R)}(\mathcal{X}_n) &= \int_{Y_n} d\mu(\Lambda)\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\langle\Lambda|D^\dagger(\boldsymbol{\alpha})\,R^2 D(\boldsymbol{\alpha})\,|\Lambda\rangle && \text{(5.9a)}\\
&= \int_{Y_n} d\mu(\Lambda)\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}\sum_{\boldsymbol{\lambda}\in\Lambda} e^{i\Phi(\boldsymbol{\lambda})} e^{-i2\pi\boldsymbol{\alpha}^T J\boldsymbol{\lambda}} c_{R^2}(\boldsymbol{\gamma})\,\mathrm{Tr}[D(\boldsymbol{\lambda})\,D(\boldsymbol{\gamma})] && \text{(5.9b)}\\
&= c_{R^2}(0) = \mathrm{Tr}\big[R^2\big]. && \text{(5.9c)}
\end{aligned}$$

Similarly, we have

$$W_1^{(R)}(\mathcal{X}_n) = \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \int_{\mathcal{P}(\Lambda)'} d\boldsymbol{\alpha}'\, \langle \Lambda | D^\dagger(\boldsymbol{\alpha})\, RD(\boldsymbol{\alpha}')\, |\Lambda'\rangle \langle \Lambda' | D^\dagger(\boldsymbol{\alpha}')\, RD(\boldsymbol{\alpha})\, |\Lambda\rangle \tag{5.10a}$$

$$\begin{aligned} &= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \sum_{\boldsymbol{\lambda}\in\Lambda} \sum_{\boldsymbol{\lambda}'\in\Lambda'} \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \\ &\quad \times \int_{\mathcal{P}(\Lambda)'} d\boldsymbol{\alpha}'\, e^{i\Phi(\boldsymbol{\lambda})} e^{i\Phi(\boldsymbol{\lambda}')} e^{-i2\pi\boldsymbol{\alpha}^T J\boldsymbol{\lambda} - i2\pi\boldsymbol{\alpha}'^T J\boldsymbol{\lambda}'} \operatorname{Tr}[D(\boldsymbol{\lambda})\, RD(\boldsymbol{\lambda}')\, R] \end{aligned} \tag{5.10b}$$

$$= \operatorname{Tr}\left[R^2\right], \tag{5.10c}$$

where we used

$$\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} e^{-i2\pi\boldsymbol{\alpha}^T J\boldsymbol{\lambda}} = \delta_{\mathbf{0},\boldsymbol{\lambda}},\ \boldsymbol{\lambda} \in \Lambda. \tag{5.11}$$

Altogether, this yields

$$\mathcal{F}_1^{(R)}(\mathcal{X}_n, a_1) = \operatorname{Tr}\left[R^2\right] - 2a_1 \operatorname{Tr}\left[R^2\right] + a_1^2 \operatorname{Tr}\left[R^2\right] \tag{5.12}$$

We see that, with $a_1 = 1$, we obtain $\mathcal{F}_1^{(R)}(\mathcal{X}_n, a_1 = 1) = 0$.

Now, on to $t = 2$. Here we have

$$X_2^{(R)}(\mathcal{X}_n) = \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \langle \Lambda | D^\dagger(\boldsymbol{\alpha})\, R^2 D(\boldsymbol{\alpha})\, |\Lambda\rangle^2 \tag{5.13a}$$

$$= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda},\boldsymbol{\lambda}'\in\Lambda} e^{i\Phi(\boldsymbol{\lambda}) - i\Phi(\boldsymbol{\lambda}')} c_{R^2}(-\boldsymbol{\lambda})\, c_{R^2}(\boldsymbol{\lambda}') \underbrace{\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha}\, e^{i2\pi(\boldsymbol{\lambda}-\boldsymbol{\lambda}')^T J\boldsymbol{\alpha}}}_{\delta_{\boldsymbol{\lambda},\boldsymbol{\lambda}'}} \tag{5.13b}$$

$$= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}\in\Lambda} |c_{R^2}(\boldsymbol{\lambda})|^2 \tag{5.13c}$$

$$= |c_{R^2}(0)|^2 + \int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda}\, |c_{R^2}(\boldsymbol{\lambda})|^2 \tag{5.13d}$$

$$= \operatorname{Tr}\left[R^2\right]^2 + \operatorname{Tr}\left[R^4\right]. \tag{5.13e}$$

We also need to compute

$$W_2^{(R)}(\mathcal{X}_n) = \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \int_{\mathcal{P}(\Lambda)'} d\boldsymbol{\alpha}' \, \left|\langle \Lambda | D^\dagger(\boldsymbol{\alpha})\, RD(\boldsymbol{\alpha}')\, |\Lambda'\rangle\right|^4 \tag{5.14a}$$

$$= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 d\boldsymbol{\gamma}_4\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R(\boldsymbol{\gamma}_4) \sum_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\in\Lambda;\boldsymbol{\lambda}_1',\boldsymbol{\lambda}_2'\in\Lambda'} \tag{5.14b}$$

$$\int_{\mathcal{P}(\Lambda)} d\boldsymbol{\alpha} \int_{\mathcal{P}(\Lambda)'} d\boldsymbol{\alpha}' e^{-i2\pi\boldsymbol{\alpha}'^T J\left(\boldsymbol{\lambda}_1'+\boldsymbol{\lambda}_2'\right) - i2\pi\boldsymbol{\alpha}^T J(\boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2)} \tag{5.14c}$$

$$\mathrm{Tr}[D(\boldsymbol{\gamma}_1)\, D(\boldsymbol{\lambda}_1')\, D(\boldsymbol{\gamma}_2)\, D(\boldsymbol{\lambda}_1)\, D(\boldsymbol{\gamma}_3)\, D(\boldsymbol{\lambda}_2')\, D(\boldsymbol{\gamma}_4)\, D(\boldsymbol{\lambda}_2)] \tag{5.14d}$$

$$= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 d\boldsymbol{\gamma}_4\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R(\boldsymbol{\gamma}_4) \sum_{\boldsymbol{\lambda}\in\Lambda;\boldsymbol{\lambda}'\in\Lambda'} \tag{5.14e}$$

$$\mathrm{Tr}\left[D(\boldsymbol{\gamma}_1)\, D(\boldsymbol{\lambda}')\, D(\boldsymbol{\gamma}_2)\, D(\boldsymbol{\lambda})\, D(\boldsymbol{\gamma}_3)\, D^\dagger(\boldsymbol{\lambda}')\, D(\boldsymbol{\gamma}_4)\, D^\dagger(\boldsymbol{\lambda})\right] \tag{5.14f}$$

$$= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 d\boldsymbol{\gamma}_4\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R(\boldsymbol{\gamma}_4) \sum_{\boldsymbol{\lambda}\in\Lambda;\boldsymbol{\lambda}'\in\Lambda'} \tag{5.14g}$$

$$\underbrace{\mathrm{Tr}[D(\boldsymbol{\gamma}_1)\, D(\boldsymbol{\gamma}_2)\, D(\boldsymbol{\gamma}_3)\, D(\boldsymbol{\gamma}_4)]}_{e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{-i\pi\boldsymbol{\gamma}_3^T J\boldsymbol{\gamma}_4}\delta(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3+\boldsymbol{\gamma}_4)}\, e^{-i2\pi\boldsymbol{\lambda}^T J(\boldsymbol{\gamma}_3+\boldsymbol{\gamma}_4-\boldsymbol{\lambda}')} e^{-i2\pi\boldsymbol{\lambda}'^T J(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3)} \tag{5.14h}$$

$$= \int_{Y_n} d\mu(\Lambda)\, d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3\, c_R(\boldsymbol{\gamma}_1)\, c_R(\boldsymbol{\gamma}_2)\, c_R(\boldsymbol{\gamma}_3)\, c_R^*(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3) \sum_{\boldsymbol{\lambda}\in\Lambda;\boldsymbol{\lambda}'\in\Lambda'} \tag{5.14i}$$

$$e^{-i\pi\boldsymbol{\gamma}_1^T J\boldsymbol{\gamma}_2} e^{-i\pi\boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1-\boldsymbol{\gamma}_2)} e^{i2\pi\boldsymbol{\lambda}^T J(\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2+\boldsymbol{\lambda}')} e^{-i2\pi\boldsymbol{\lambda}'^T J(\boldsymbol{\gamma}_2+\boldsymbol{\gamma}_3)} \tag{5.14j}$$

We have

$$\int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-i2\pi\boldsymbol{\lambda}^T J\boldsymbol{x}} = 1 + \delta(\boldsymbol{x}), \tag{5.15}$$

such that we can further simplify

$$\ldots = \int_{Y_n} d\mu(\Lambda') \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 \, c_R(\boldsymbol{\gamma}_1) \, c_R(\boldsymbol{\gamma}_2) \, c_R(\boldsymbol{\gamma}_3) \, c_R^*(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3) \tag{5.16a}$$

$$\sum_{\boldsymbol{\lambda}' \in \Lambda'} e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{-i\pi \boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2)} \left\{1 + \delta(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\lambda}')\right\} e^{-i2\pi \boldsymbol{\lambda}'^T J(\boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3)} \tag{5.16b}$$

$$= \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 \, c_R(\boldsymbol{\gamma}_1) \, c_R(\boldsymbol{\gamma}_2) \, c_R(\boldsymbol{\gamma}_3) \, c_R^*(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3) \tag{5.16c}$$

$$e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{-i\pi \boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2)} \left\{1 + \delta(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2)\right\} \tag{5.16d}$$

$$+ \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 \, c_R(\boldsymbol{\gamma}_1) \, c_R(\boldsymbol{\gamma}_2) \, c_R(\boldsymbol{\gamma}_3) \, c_R^*(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3) \tag{5.16e}$$

$$e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{-i\pi \boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2)} \int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda}' \left\{1 + \delta(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\lambda}')\right\} e^{-i2\pi \boldsymbol{\lambda}'^T J(\boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3)} \tag{5.16f}$$

$$= \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 \, c_R(\boldsymbol{\gamma}_1) \, c_R(\boldsymbol{\gamma}_2) \, c_R(\boldsymbol{\gamma}_3) \, c_R^*(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3) \tag{5.16g}$$

$$e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{-i\pi \boldsymbol{\gamma}_3^T J(-\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2)} \left\{1 + \delta(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2)\right\} \tag{5.16h}$$

$$+ \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 \, c_R(\boldsymbol{\gamma}_1) \, c_R(\boldsymbol{\gamma}_2) \, c_R(\boldsymbol{\gamma}_3) \, c_R^*(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3) \tag{5.16i}$$

$$e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{i\pi \boldsymbol{\gamma}_3^T J(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2)} \left\{\delta(\boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3) + e^{i2\pi(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2)^T J(\boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3)}\right\} \tag{5.16j}$$

This final sum has 4 parts, the first equals $\mathrm{Tr}[R^4]$. The second is

$$\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_3 \, |c_R(\boldsymbol{\gamma}_1)|^2 |c_R(\boldsymbol{\gamma}_3)|^2 = \mathrm{Tr}\left[R^2\right]^2, \tag{5.17a}$$

same as the third

$$\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 |c_R(\boldsymbol{\gamma}_1)|^2 |c_R(\boldsymbol{\gamma}_2)|^2 = \mathrm{Tr}\left[R^2\right]^2. \tag{5.18a}$$

Finally, the last term is computed as

$$\int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_3 \, c_R(\boldsymbol{\gamma}_1) \, c_R(\boldsymbol{\gamma}_2) \, c_R(\boldsymbol{\gamma}_3) \, c_R^*(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3) \, e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{i\pi \boldsymbol{\gamma}_3^T J(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2)} e^{i2\pi(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2)^T J(\boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3)} \tag{5.19a}$$

$$= \int_{\mathbb{R}^{2n}} d\boldsymbol{\gamma}_1 d\boldsymbol{\gamma}_2 d\boldsymbol{\gamma}_4 \, c_R(\boldsymbol{\gamma}_1) \, c_R(\boldsymbol{\gamma}_2) \, c_R^*(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_4) \, c_R(\boldsymbol{\gamma}_4) \, e^{-i\pi \boldsymbol{\gamma}_1^T J \boldsymbol{\gamma}_2} e^{-i\pi \boldsymbol{\gamma}_4^T (\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2)} = \mathrm{Tr}\left[R^4\right]. \tag{5.19b}$$

Thus in total, we have

$$W_2^{(R)}(\mathcal{X}_n) = 2\,\mathrm{Tr}\left[R^4\right] + 2\,\mathrm{Tr}\left[R^2\right]^2 \tag{5.20}$$

and

$$\mathcal{F}_2^{(R)}(\mathcal{X}_n, a_2) = 2\,\mathrm{Tr}\left[R^4\right] + 2\,\mathrm{Tr}\left[R^2\right]^2 - 2a_2 \left\{\mathrm{Tr}\left[R^2\right]^2 + \mathrm{Tr}\left[R^4\right]\right\} + \frac{a_t^2}{2}\left\{\mathrm{Tr}\left[R^2\right]^2 + \mathrm{Tr}\left[R^4\right]\right\}. \tag{5.21}$$

From this equation, we can observe that we have, with $a_2 = 2$, $\mathcal{F}_2^{(R)}(\mathcal{X}_n, a_2 = 2) = 0$. Under the premise of Proposition 5.2.1 we hence have verified that $\mathcal{X}_n$ is a rigged 2-design. □

The meaning of the frame potential

We now give a basic interpretation of the frame potential in *finite-dimensional* systems. The frame potential has its roots in Welch bounds [261, 262]. There are many great expositions of the frame potential, its relation to designs, mutually unbiased bases, and scrambling and chaos [41, 45, 214, 263, 264]. Here, we give a slightly different argument for its applicability.

Ultimately, we want to characterize some ensemble $\mathcal{X}_n$ of states via a number $F(\mathcal{X}_n) \in \mathbb{R}$. Because states are elements of projective space, this function F must be a function of projectors rather than states. Additionally, there is no ordering on $\mathcal{X}_n$, meaning $F(\mathcal{X}_n)$ should be independent of reordering or relabeling the elements of $\mathcal{X}_n$. We are thus led to

$$F_t^{(k)}(\mathcal{X}_n) = \mathrm{Tr} \mathop{\mathbb{E}}_{\psi_1,\ldots,\psi_k \in \mathcal{X}_n} (|\psi_1\rangle\langle\psi_1|)^{\otimes t} \ldots (|\psi_k\rangle\langle\psi_k|)^{\otimes t} = \mathop{\mathbb{E}}_{\psi_1,\ldots,\psi_k \in \mathcal{X}_n} \langle\psi_1|\psi_2\rangle^t \langle\psi_2|\psi_3\rangle^t \ldots \langle\psi_k|\psi_1\rangle^t, \tag{5.22}$$

where we can tune t to probe different moments of $\mathcal{X}_n$.

We then notice that $F_t^{(1)}(\mathcal{X}_n) = 1$ for any $\mathcal{X}_n$ (because in finite-dimensions we can always assume the states are unit-normalized), and so is trivial. The first nontrivial function that can characterize the t^{th} moment of $\mathcal{X}_n$ is thus $F_t^{(2)}(\mathcal{X}_n)$. $F_t^{(2)}(\mathcal{X}_n)$ is in fact the standard frame potential for finite-dimensional systems, and as such we will notate $F_t(\mathcal{X}_n) \equiv F_t^{(2)}(\mathcal{X}_n)$.

If we are interested in measuring how distributed an ensemble of states is (*i.e.* how much it "covers" the full space of states and how spread it is over the space), we want the average magnitude of the overlap between the states in the ensemble to be small. Therefore, the $F_t(\mathcal{X}_n)$ being small is a measure of how evenly spread $\mathcal{X}_n$ is. This then gives some intuition as to why the frame potential shows up as a measure of scrambling and chaos.

When thinking about t-designs, we need ensembles $\mathcal{X}_n$ that maximally cover the whole space of states in order for them to mimic the whole space itself. Therefore, $\mathcal{X}_n$ is a t-design if and only if $F_t(\mathcal{X}_n)$ is minimized (the proof of this is analogous to the proof of Proposition 5.2.1). When $\mathcal{X}_n$ does mimic the t^{th} moment of the whole space of states, then clearly $F_t^{(k)}(\mathcal{X}_n)$ is determined for all k. Thus, we have found that even though $F_t^{(k)}(\mathcal{X}_n)$ in principle contains more information about the t^{th} moment of $\mathcal{X}_n$ than just $F_t(\mathcal{X}_n) = F_t^{(2)}(\mathcal{X}_n)$, $F_t(\mathcal{X}_n)$ alone already characterizes how far $\mathcal{X}_n$ is from being sufficiently spread over the whole space of states. In particular, via some continuity analysis, $F_t^{(k)}(\mathcal{X}_n)$ can be bounded in terms of how far $F_t(\mathcal{X}_n)$ is from its minimum, and this bound gets tighter as $F_t(\mathcal{X}_n)$ approaches its minimum.

In the CV case, $W_t^{(R)}(\mathcal{X}_n)$ is the analogue of $F_t(\mathcal{X}_n)$. The function $X_t^{(R)}(\mathcal{X}_n)$ is an artifact of the fact that the states are not normalized. In the case of regularized rigged designs, $X_t^{(R)}(\mathcal{X}_n)$ can in fact be gotten rid of as in Proposition 4.5.2, making the analogy between the finite- and infinte-dimensional frame potentials more direct.

Section 5.3: Diagonal coherent state basis

In this appendix, we provide another proof that the displaced GKP ensemble forms a rigged 2-design by working in the coherent state basis (i.e. the P-representation). We then perform the analogous 3-design calculation, but are unfortunately unable to complete it.

The key point is that the diagonal elements of an operator in the coherent state basis fully determine the operator [265, 266]. Hence, in order to check if an ensemble $\mathcal{X}_n$ is a rigged t-design, we only need to check that

$$\int_{\mathcal{X}_n} dX \prod_{i=1}^{t} |\langle \boldsymbol{\alpha}_i | X \rangle|^2 = a_t \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t | \Pi_t | \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t \rangle \tag{5.23}$$

for some $a_t > 0$. For $\mathcal{X}_n$ the ensemble of GKP codes defined in the main text, this amounts to showing that

$$g(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t) = a_t \left\langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t \right| \Pi_t \left| \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t \right\rangle , \tag{5.24}$$

where

$$g(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t) \coloneqq \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\beta} \prod_{i=1}^{t} \mathrm{Tr}\left[D(\boldsymbol{\beta}) \left|\Lambda\right\rangle\!\left\langle\Lambda\right| D(-\boldsymbol{\beta}) D(\boldsymbol{\alpha}_i) \left|0\right\rangle\!\left\langle 0\right| D(-\boldsymbol{\alpha}_i) \right] \tag{5.25a}$$

$$= \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\beta} \prod_{i=1}^{t} \sum_{\boldsymbol{\lambda}_i \in \Lambda} e^{i\Phi(\boldsymbol{\lambda}_i)} \mathrm{Tr}\left[D(\boldsymbol{\beta}) D(\boldsymbol{\lambda}_i) D(-\boldsymbol{\beta}) D(\boldsymbol{\alpha}_i) \left|0\right\rangle\!\left\langle 0\right| D(-\boldsymbol{\alpha}_i) \right] \tag{5.25b}$$

$$= \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\beta} \prod_{i=1}^{t} \sum_{\boldsymbol{\lambda}_i \in \Lambda} e^{i\Phi(\boldsymbol{\lambda}_i)} \left\langle 0\right| D(-\boldsymbol{\alpha}_i) D(\boldsymbol{\beta}) D(\boldsymbol{\lambda}_i) D(-\boldsymbol{\beta}) D(\boldsymbol{\alpha}_i) \left|0\right\rangle \tag{5.25c}$$

$$= \int_{Y_n} d\mu(\Lambda) \int_{\mathcal{P}(\Lambda)} d\boldsymbol{\beta} \prod_{i=1}^{t} \sum_{\boldsymbol{\lambda}_i \in \Lambda} e^{i\Phi(\boldsymbol{\lambda}_i)} e^{2\pi i \boldsymbol{\lambda}_i^T J (\boldsymbol{\beta} - \boldsymbol{\alpha}_i)} \left\langle 0\right| D(\boldsymbol{\lambda}_i) \left|0\right\rangle \tag{5.25d}$$

$$(\text{Eq. (5.11)}) = \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_t \in \Lambda \\ \text{s.t. } \boldsymbol{\lambda}_1+\cdots+\boldsymbol{\lambda}_t=0}} e^{i\sum_{j=1}^t \Phi(\boldsymbol{\lambda}_j)} e^{2\pi i \sum_{j=1}^t \boldsymbol{\alpha}_j^T J \boldsymbol{\lambda}_j} \prod_{j=1}^t \langle 0|\boldsymbol{\lambda}_j\rangle \tag{5.25e}$$

$$= \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_t \in \Lambda \\ \text{s.t. } \boldsymbol{\lambda}_1+\cdots+\boldsymbol{\lambda}_t=0}} e^{i\sum_{j=1}^t \Phi(\boldsymbol{\lambda}_j)} e^{2\pi i \sum_{j=1}^t \boldsymbol{\alpha}_j^T J \boldsymbol{\lambda}_j} e^{-\frac{1}{2}\sum_{j=1}^t \|\boldsymbol{\lambda}_j\|^2}. \tag{5.25f}$$

The RHS of Eq. (5.23) yields

$$\Pi_1 = I \qquad \Longrightarrow \quad \langle \boldsymbol{\alpha}| \Pi_1 |\boldsymbol{\alpha}\rangle = 1 \tag{5.26a}$$

$$\Pi_2 = \frac{1}{2}(I \otimes I + \text{SWAP}) \qquad \Longrightarrow \quad \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2| \Pi_2 |\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\rangle = \frac{1}{2} + \frac{1}{2}|\langle \boldsymbol{\alpha}_1|\boldsymbol{\alpha}_2\rangle|^2. \tag{5.26b}$$

Furthermore, using $\Pi_3 = \frac{1}{6}\sum_{\sigma\in S_3} W_\sigma$, where W_σ permutes the tensor factors, we also have

$$\begin{aligned}\langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3| \Pi_3 |\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\rangle = \frac{1}{6}\Big[1 + |\langle \boldsymbol{\alpha}_1|\boldsymbol{\alpha}_2\rangle|^2 + |\langle \boldsymbol{\alpha}_1|\boldsymbol{\alpha}_3\rangle|^2 + |\langle \boldsymbol{\alpha}_2|\boldsymbol{\alpha}_3\rangle|^2 \\ + 2\,\text{Re}\left(\langle \boldsymbol{\alpha}_1|\boldsymbol{\alpha}_2\rangle\langle \boldsymbol{\alpha}_2|\boldsymbol{\alpha}_3\rangle\langle \boldsymbol{\alpha}_3|\boldsymbol{\alpha}_1\rangle\right)\Big].\end{aligned} \tag{5.27}$$

For $t = 1$, we have

$$g(\boldsymbol{\alpha}) = \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}\in\Lambda \\ \text{s.t. } \boldsymbol{\lambda}=0}} e^{i\Phi(\boldsymbol{\lambda})} e^{2\pi i \boldsymbol{\alpha}^T J \boldsymbol{\lambda}} e^{-\|\boldsymbol{\lambda}\|^2/2} \tag{5.28a}$$

$$= 1 \tag{5.28b}$$

$$= \langle \boldsymbol{\alpha}| \Pi_1 |\boldsymbol{\alpha}\rangle, \tag{5.28c}$$

therefore showing that the displaced GKP ensemble forms a 1-design with $a_1 = 1$.

For $t = 2$, we have

$$g(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) = \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2 \in \Lambda \\ \text{s.t. } \boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2=0}} e^{i\sum_{j=1}^2 \Phi(\boldsymbol{\lambda}_j)} e^{2\pi i \sum_{j=1}^2 \boldsymbol{\alpha}_j^T J \boldsymbol{\lambda}_j} e^{-\frac{1}{2}\sum_{j=1}^2 \|\boldsymbol{\lambda}_j\|^2} \tag{5.29a}$$

$$= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}\in\Lambda} e^{2\pi i \boldsymbol{\lambda}^T J(\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1)} e^{-\|\boldsymbol{\lambda}\|^2} \tag{5.29b}$$

$$= 1 + \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}\in\Lambda\setminus\{0\}} e^{2\pi i \boldsymbol{\lambda}^T J(\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1)} e^{-\|\boldsymbol{\lambda}\|^2} \tag{5.29c}$$

$$\text{(Lemma 5.1.2)} = 1 + \int_{\mathbb{R}^{2n}} d\boldsymbol{\lambda}\; e^{2\pi i \boldsymbol{\lambda}^T J(\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1)} e^{-\|\boldsymbol{\lambda}\|^2} \tag{5.29d}$$

$$= 1 + e^{-\|\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1\|^2} \tag{5.29e}$$

$$= 1 + \left|\langle \boldsymbol{\alpha}_1 | \boldsymbol{\alpha}_2 \rangle\right|^2 \tag{5.29f}$$

$$= 2 \left\langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \right| \Pi_2 \left| \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \right\rangle . \tag{5.29g}$$

This therefore shows that the displaced GKP ensemble forms a rigged 2-design with $a_2 = 2$.

The analogous calculation for $t = 3$ requires a second moment analogue of Lemma 5.1.2 beyond the extension provided in Ref. [267]. For $t = 3$, we have

$$g(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3) = \int_{Y_n} d\mu(\Lambda) \sum_{\substack{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3\in\Lambda \\ \text{s.t. } \boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2+\boldsymbol{\lambda}_t=0}} e^{i\sum_{j=1}^t \Phi(\boldsymbol{\lambda}_j)} e^{2\pi i \sum_{j=1}^t \boldsymbol{\alpha}_j^T J \boldsymbol{\lambda}_j} e^{-\frac{1}{2}\sum_{j=1}^3 \|\boldsymbol{\lambda}_j\|^2} \tag{5.30a}$$

$$= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\in\Lambda} e^{2\pi i\left(\boldsymbol{\alpha}_1^T J\boldsymbol{\lambda}_1 + \boldsymbol{\alpha}_2^T J\boldsymbol{\lambda}_2 - \boldsymbol{\alpha}_3^T J(\boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2)\right)} e^{-\|\boldsymbol{\lambda}_1\|^2/2 - \|\boldsymbol{\lambda}_2\|^2/2 - \|\boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2\|^2/2} \tag{5.30b}$$

$$= \int_{Y_n} d\mu(\Lambda) \sum_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\in\Lambda} e^{2\pi i\left(\boldsymbol{\lambda}_1^T J(\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_1) + \boldsymbol{\lambda}_2^T J(\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1)\right)} e^{-\|\boldsymbol{\lambda}_1\|^2/2 - \|\boldsymbol{\lambda}_2\|^2/2 - \|\boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2\|^2/2}. \tag{5.30c}$$

To complete this calculation, we need an expression for the second moment integral over Y_n. Once this is known, we would need to compare the result to Eq. (5.27).

Section 5.4: Global and local design-based shadow tomography for 3-designs

In the previous sections, we have leveraged the rigged 2-design property of our ensemble $\mathcal{X}_n$ together with knowledge about the specific structure of the ensemble and/or additional physical assumptions in order to bound the variance scaling. While we have not been able to show whether a stronger 3-design property also holds for our ensemble, we explore what this property would imply. The following discussion applies to design-based classical shadow tomography protocols for either finite or inifinite systems (via the definition of rigged designs).

Throughout this section, we will always assume that O is traceless in order to simplify calculations. Assume that we have a measure space $(\mathcal{X}, \mu)$ that is a 3-design. In other words, for all $t \leq 3$, $\int_{\mathcal{X}} X^{\otimes t} d\mu(X) = a_t \Pi_t$, where each $X \in \mathcal{X}$ is a positive operator, $a_1 = 1$, and $a_t > 0$. A rigged 3-design in infinite-dimensions and a quantum state 3-design in finite dimensions satisfies exactly this criteria. In particular, in d dimensions for d finite, if we choose $a_t = d/\operatorname{Tr}(\Pi_t)$ (ie $a_2 = 2/(d+1)$ and $a_3 = 6/(d+1)(d+2)$) and require each X to be rank 1, then we get the standard definition of a quantum state 3-design. In infinite-dimensions, if we allow a_2, a_3 to be arbitrary constants and require each X to be rank 1, then we get the definition of a rigged 3-design.

Subsection 5.4.1: Global shadows

Define

$$M(\rho) = \int_{\mathcal{X}} \operatorname{Tr}[\rho X] X \, \mathrm{d}\mu(X) = \frac{a_2}{2}(I + \rho). \tag{5.31}$$

Suppose that we can sample from the probability distribution $dP(X, \rho) = \operatorname{Tr}(\rho X) \, \mathrm{d}\mu(X)$. Given a sample X from this distribution, we have the estimator $o(X) = \frac{2}{a_2} \operatorname{Tr}[XO]$. The

expectation value of our estimator is

$$\mathbb{E}_{X\sim dP}[o(X)] = \int_{\mathcal{X}} o(X)\ dP(X) \tag{5.32a}$$
$$= \frac{2}{a_2}\int_{\mathcal{X}} \mathrm{Tr}[OX]\,\mathrm{Tr}[\rho X]\ d\mu(X) \tag{5.32b}$$
$$= 2\,\mathrm{Tr}[\Pi_2(\rho\otimes O)] \tag{5.32c}$$
$$= \mathrm{Tr}[\rho O] \tag{5.32d}$$

The variance of our estimator is

$$\mathrm{Var}_{X\sim dP}[o(X)] = \int_{\mathcal{X}} o(X)^2\ dP(X)\ - \mathrm{Tr}[\rho O]^2 \tag{5.33a}$$
$$= \frac{4}{a_2^2}\int_{\mathcal{X}} \mathrm{Tr}[XO]^2\,\mathrm{Tr}[\rho X]\ d\mu(X)\ - \mathrm{Tr}[\rho O]^2 \tag{5.33b}$$
$$= \frac{4a_3}{a_2^2}\,\mathrm{Tr}[\Pi_3(O\otimes O\otimes\rho)] - \mathrm{Tr}[\rho O]^2 \tag{5.33c}$$
$$\leq \frac{2a_3}{a_2^2}\,\mathrm{Tr}[O^2], \tag{5.33d}$$

where the last step is executed using Eq. (E9) in Chapter 4. We see that the efficiency of a global shadow tomography protocol with a 3-design depends on $2a_3/a_2^2$. In finite dimension d, when using quantum state 3-designs, the scaling dependencies of this fraction on d cancel and it follows that the variance of the corresponding estimator is $\sim \mathrm{Tr}[O^2]$. In infinite dimensions, a rigged 3-design satisfies $a_2, a_3 \sim 1$, and thus again we obtain a variance scaling of $\sim \mathrm{Tr}[O^2]$.

We highlight that the factor a_2 alone is not important, but rather the ratio $2a_3/a_2^2$. In general, even in the displaced GKP case (where we are unsure if the ensemble forms a rigged 3-design), the fact that $a_2 \sim 1$ (as opposed to $a_2 \sim 2^{-n}$ as is typically the case in qubit protocols) is not the key difference dictating the variance scaling of the corresponding shadow tomography protocol. Ultimately, what matters is the ratio $2a_3/a_2^2$. For finite d and

utilizing quantum state 3-designs, we have

$$a_2 = \frac{2}{d+1} \quad \text{and} \quad a_3 = \frac{6}{(d+1)(d+2)}, \tag{5.34}$$

yielding $\frac{2a_3}{a_2^2} \xrightarrow{d\to\infty} 3$, which is precisely the scaling observed in Ref. [50].

Subsection 5.4.2: Local shadows

In this section, we will assume that our observable O is of form $O = O_1 \otimes \cdots \otimes O_n$, and we assume that $\mathrm{Tr}[O] = 0$ (for example, in finite dimensions, O can be a Pauli observable). We expand our state as $\rho = \sum_\alpha c_\alpha \rho_1^{(\alpha)} \otimes \cdots \otimes \rho_n^{(\alpha)}$, with $\mathrm{Tr}[\rho_i^{(\alpha)}] = 1$ for each i and α. We sample from the ensemble $\mathcal{X} = \mathcal{X}_1^{\otimes n}$ with measure $dP(X_1 \otimes \cdots \otimes X_n)$ defined by

$$dP(X_1 \otimes \cdots \otimes X_n) = \mathrm{Tr}[\rho(X_1 \otimes \cdots \otimes X_n)]\ d\mu(X_1) \ldots d\mu(X_n). \tag{5.35}$$

Our estimator is

$$o(X_1 \otimes \cdots \otimes X_n) = \left(\frac{2}{a_2}\right)^n \mathrm{Tr}[X_1 O_1] \ldots \mathrm{Tr}[X_n O_n]. \tag{5.36}$$

The expectation value of our estimator is then

$$\begin{aligned}
\mathbb{E}_{X_1 \otimes \cdots \otimes X_n \sim dP}[o(X_1 \otimes \cdots \otimes X_n)] &= \int_{\mathcal{X}_1^{\otimes n}} o(X_1 \otimes \cdots \otimes X_n)\ dP(X_1 \otimes \cdots \otimes X_n) && \text{(5.37a)} \\
&= \left(\frac{2}{a_2}\right)^n \sum_\alpha c_\alpha \prod_{i=1}^n \int_{\mathcal{X}_1} \mathrm{Tr}[X_i O_i]\, \mathrm{Tr}[X_i \rho_i^{(\alpha)}]\ d\mu(X_i) && \text{(5.37b)} \\
&= \sum_\alpha c_\alpha \prod_{i=1}^n \mathrm{Tr}[O_i \rho_i^{(\alpha)}] && \text{(5.37c)} \\
&= \mathrm{Tr}[\rho O] && \text{(5.37d)}
\end{aligned}$$

The variance of our estimator is

$$\begin{aligned}
\mathrm{Var}_{X_1\otimes\cdots\otimes X_n\sim dP}[o(X_1\otimes\cdots\otimes X_n)] &\leq \int_{\mathcal{X}_1^{\otimes n}} o(X_1\otimes\cdots\otimes X_n)^2\; dP(X_1\otimes\cdots\otimes X_n) && \text{(5.38a)}\\
&= \sum_\alpha c_\alpha \prod_{i=1}^n \frac{4}{a_2^2}\int_{\mathcal{X}_1} \mathrm{Tr}[X_i O_i]^2\,\mathrm{Tr}[X_i\rho_i^{(\alpha)}]\; d\mu(X_i) && \text{(5.38b)}\\
&= \sum_\alpha c_\alpha \prod_{i=1}^n \frac{2a_3}{a_2^2}\,\mathrm{Tr}[O_i^2] && \text{(5.38c)}\\
&= \left(\frac{2a_3}{a_2^2}\right)^n \prod_{i=1}^n \mathrm{Tr}[O_i^2] && \text{(5.38d)}\\
&= \left(\frac{2a_3}{a_2^2}\right)^n \mathrm{Tr}[O^2] && \text{(5.38e)}
\end{aligned}$$

Again we see that the relevant factor is $\frac{2a_3}{a_2^2}$. For fixed $d = 2$, using Eq. (5.34), this value is $2a_3/a_2^2 = 9/4 < 3$.

If the design was such that $2a_3 \leq a_2^2$, this design-based shadow tomography protocol is expected attain an asymptotic variance of 0 for *any* observable. This suggests that in general a parameter tradeoff $2a_3 \geq a_2^2$ must hold and the optimal ensemble is the one that minimizes this inequality. Observe that here we have used the *local shadow tomography* protocol to derive a statement about the global shadow as well as general parameter tradeoffs between the factors a_t of any 3-design.

Chapter 6: Projective toric designs, quantum state designs, and mutually unbiased bases

Section 6.1: Introduction

Given a measure space (M, μ) and a set of polynomials on M, a *t-design* on M is a measure space $(X \subset M, \nu)$ satisfying $\int_X f \, \mathrm{d}\nu = \int_M f \, \mathrm{d}\mu$ for all polynomials f of degree $\leq t$ [182–192, 268, 269]. Classic examples are Gaussian quadrature rules [182] and spherical designs [183, 184], where the measure space M is the hypercube and hypersphere, respectively. Typically, one is interested in finding designs where X is a discrete measure space such that the integral over X with respect to ν reduces to a weighted sum that is often simpler to compute. However, this is not always possible; in the case of rigged designs (defined below), it is often crucial that X be a non-discrete measure space [9].

Specific forms of t-designs for particular choices of measure spaces M have found a plethora of uses in the field of quantum information theory [36–70]. In particular, complex projective space $\mathbb{CP}^{d-1}$ describes the space of d-dimensional quantum states [72], so t-designs on $M = \mathbb{CP}^{d-1}$ are called *complex-projective* or *quantum state t-designs*. These quantum state designs also relate to other mathematical objects such as symmetric, informationally complete positive operator valued measures (SIC-POVMs) and complete sets of mutually unbiased bases (MUBs), which themselves are conjectured to relate to finite projective geometry. Finite-dimensional quantum state designs can be generalized to designs on infinite-dimensional, or continuous-variable, quantum systems by defining *rigged quantum*

state t-designs, which are designs on the space of tempered distributions $M = S(\mathbb{R})'$ [9]. The (projective) unitary group $\mathrm{PU}(d)$ describes the space of noiseless dynamics of quantum states, and these too admit constructions of *unitary t-designs*. Finally, the space of mixed quantum states with the Hilbert-Schmidt volume measure allows for *mixed-state t-designs* [70]. Therefore, a better understanding of various kinds of t-designs can also lead to deep insights about quantum information.

Consider the complex sphere Ω_d; that is, the set of unit vectors in $\mathbb{C}^d$. Any vector in Ω_d can be written (non-uniquely) as $|q, \phi\rangle := \sum_{n=1}^{d} \sqrt{q_n} \mathrm{e}^{\mathrm{i}\phi_n} |n\rangle$, where $\{|n\rangle\}_{n=1}^{d}$ forms an orthonormal basis, $q = (q_n)_{n=1}^{d}$ is a discrete probability distribution ($\sum_n q_n = 1$), and $\phi = (\phi_n)_{n=1}^{d}$ is a set of phases. Therefore, q belongs to the $(d-1)$-simplex Δ^{d-1} and ϕ to the d-torus T^d. Via this mapping $\Delta^{d-1} \times T^d \to \Omega_d$, one can combine simplex designs and toric designs to form complex spherical designs [189]. Identifying $\mathbb{CP}^{d-1}$ with $\Omega_d/\mathrm{U}(1)$ (that is, quantum states are complex unit vectors with a global phase redundancy), we have a similar mapping $\Delta^{d-1} \times P(T^d) \to \mathbb{CP}^{d-1}$ defined as $(q, [\phi]) \mapsto [|q, \phi\rangle]$, where $P(T^d) = T^d/\mathrm{U}(1)$ is the projective torus (see Definition 6.2.4) and $[\cdot]$ denotes equivalence classes in the respective quotient spaces. In a similar way as before, via this mapping one can combine simplex designs and *projective toric designs* (see Definition 6.2.5) to form quantum state designs [9, 189].

In what follows, we flesh out and formalize this argument. Specifically, we formalize the notion of projective toric designs—both finite- and infinite-dimensional—and provide various constructions thereof. We discuss the connection between projective toric designs and difference sets [88, 270, 271], and use this correspondence to construct more projective toric designs, including some minimal ones. We illustrate the connection to quantum state designs and various other mathematical objects. Using minimal projective toric 2-designs, we construct an infinite family of almost-minimal complex-projective 2-designs. Finally, we discuss many exciting open questions regarding projective toric designs, some of which are deeply

connected to long-outstanding conjectures in mathematics, such as some conjectures relating to finite affine and projective spaces. In particular, we construct explicit counterexamples to a conjecture by Zhu [95, Conj. 1] on the relationship between uniformly-weighted quantum state 2-designs and complete sets of MUBs, and we prove a fundamental distinction between complete sets of MUBs in prime-power dimensions versus in dimension 6.

Relation to prior work. Toric designs have been considered before. Trigonometric cubature rules are such designs on the torus [186, 268, 269]. Ref. [189] generalized the idea of trigonometric cubature to more general algebraic tori. Ref. [9] studied designs on projective tori and showed an equivalence to a specific case of Ref. [189], and further showed that such projective toric designs are related to complete sets of MUBs [238]. However we believe the presentation given in Section 6.2 gives new clarity and focus on the subject. Furthermore, Section 6.2.1 compiles, to the best of our knowledge, all previously known constructions of projective toric t-designs[1], and indeed generalizes some of these constructions.

The main novel contributions of our work lie in Sections 6.2.2, 6.3 and 6.4. In Section 6.2.2, we prove a general lower bound on the size of projective toric t-designs for all dimensions and all t by relating these designs to the crystal ball sequence corresponding to the root lattice A_{n-1} [272, 273]. In Section 6.3, we relate difference sets to projective toric designs. We show how the former can be used to construct the latter. Using the connection between difference sets and projective toric designs, we furthermore relate dense difference sets to the crystal ball sequence mentioned above, and derive new (to the best of our knowledge) bounds on the size of B_t mod m sets (*cf.* Corollary 6.3.6). In Section 6.3.1, using our construction of projective toric t-designs for all t and dimensions n, we construct corresponding toric t-designs for t and n (where recall that a toric design is also a design on the diagonal subgroup of the unitary group). In Section 6.4.1, we describe the relationship between projective toric designs and quantum state designs. This relationship was first

[1]Of course, many toric designs are known, and these always project to projective toric designs. Such constructions are not compiled in this manuscript.

noted in Refs. [9, 189], though we believe that Section 6.4.1 greatly clarifies the details of this connection. In Section 6.4.1, we also construct an infinite family of *almost-minimal* quantum state 2-designs—that is, quantum state 2-designs of size exactly one more than minimal. While these specific almost-minimal designs have been noted before in Ref. [274], we arrive at the construction via a different route that utilizes projective toric designs, which we believe may have a better hope of generalizing to other infinite families and $t > 2$.

In Section 6.4.2, we use projective toric 2-designs from our difference set construction to yield uniformly-weighted quantum state 2-designs in dimension d of size exactly $d(d+1)$ that do *not* form complete sets of MUBs, thereby disproving a conjecture by Zhu that has been open for nine years [95, Conj. 1] (see also Ref. [275] for a discussion on Zhu's conjecture). In Section 6.4.3, we further characterize the relationship between projective toric 2-designs and complete sets of MUBs by proving (*cf.* Proposition 6.4.6) that the phases involved in any complete set of MUBs in dimension 6 must form a non-group projective toric 2-design. In particular, this highlights a fundamental distinction between all known constructions of complete sets of MUBs in prime-power dimensions versus any potential construction in dimension 6 (and, we conjecture, in all non-prime-power dimensions). We then discuss one possible modification of Zhu's conjecture relating to this fundamental distinction and discuss potential paths towards proving this new conjecture (*cf.* Conjecture 6.4.8). Finally, Section 6.5 compiles a number of new interesting open problems involving projective toric designs, highlighting their connection to a number of other open problems in mathematics.

Section 6.2: Theory of projective toric designs

We begin with some basic definitions that are used throughout the rest of the paper.

Definition 6.2.1 (Torus)**.** *Let* $T := \mathbb{R}/2\pi\mathbb{Z}$*. When* $n \in \mathbb{N}$*, let* $I_n := \{1, 2, \ldots n\}$*; when* $n = \infty$*, let* $I_n = I_\infty := \mathbb{N}$*. For such* n*, let* $T^n := \prod_{i \in I_n} T$ *with the product topology. Define the projection maps* $p_i \colon T^n \to T$ *as* $(\phi_j)_{j \in I_n} \mapsto \phi_i$*. For all* $n \in \mathbb{N} \cup \{\infty\}$*, let* μ_n *denote* T^n*'s*

unit-normalized Haar measure.

Note that by Tychonoff's theorem, T^∞ is compact. For all n, T^n is therefore a compact abelian group and thus has a unique unit-normalized Haar measure.

By definition, the product topology on T^∞ is the coarsest topology such that the projection maps p_i are continuous. Similarly, we endow T^∞ with the smallest σ-algebra such that the projections p_i are measurable. This σ-algebra is generated by sets of the form $A = \prod_{i\in\mathbb{N}} A_i$, where each A_i is a measurable subset of T and all but finitely many A_i are equal to T. Define a measure μ' on T^∞ by $\mu'(A) = \prod_{i\in\mathbb{N}} \mu_1(A_i)$. From Ref. [276, Thm. 10.6.1] (or Ref. [277] for a shorter proof), this definition of μ' on such subsets uniquely determines μ' on the whole space. Clearly μ' is transitionally-invariant and unit-normalized, and therefore $\mu' = \mu_\infty$.

We now define trigonometric cubature rules [186, 268, 269], which are designs on the torus. To match the general terminology of this paper, we prefer to use the term *toric design*.

Definition 6.2.2 (Toric design)**.** *A T^n t-*design *(or trigonometric cubature rule of dimension n and degree t [186, 268, 269]) is a measure space $(X \subset T^n, \Sigma, \nu)$ such that*

$$\int_X \exp\left(\mathrm{i}\sum_{j=1}^n \alpha_j\phi_j\right) \mathrm{d}\nu(\phi) = \int_{T^n} \exp\left(\mathrm{i}\sum_{j=1}^n \alpha_j\phi_j\right) \mathrm{d}\mu_n(\phi) \tag{6.1}$$

for all $\alpha \in \mathbb{Z}^n$ satisfying $\sum_{j=1}^n |\alpha_j| \leq t$.

The torus T^n is the same as the maximal torus of the unitary group $T(\mathrm{U}(n))$ [71], and indeed a T^n design is a design on $T(\mathrm{U}(n))$ [189]. Since $T(\mathrm{U}(n))$ is the diagonal subgroup of the unitary group $\mathrm{U}(n)$, we see that toric designs can equivalently be thought of as designs on the diagonal subgroup of $\mathrm{U}(n)$. Such designs are of interest in quantum information theory [85].

Example 6.2.3. In this example, we consider $n = 1$ and $t = 2$. Let X be the discrete, uniformly-weighted measure space $X = \{0, 2\pi/3, 4\pi/3\} \subset T^1$. Then, for every integer $-2 \leq \alpha \leq 2$,

$$\frac{1}{|X|} \sum_{\phi \in X} \mathrm{e}^{\mathrm{i}\alpha\phi} = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{\mathrm{i}\alpha\theta} \, \mathrm{d}\theta = \begin{cases} 1 & \text{if } \alpha = 0 \\ 0 & \text{if } 1 \leq |\alpha| \leq 2. \end{cases} \tag{6.2}$$

Hence, X is a T^1 2-design. ◇

We now consider the projective torus, an important object in the study of quantum mechanics because it removes a global phase redundancy (see Section 6.4).

Definition 6.2.4 (Projective torus)**.** *Let $P(T^n)$ denote the* projective torus $P(T^n) \coloneqq T^n/T$*, where here T denotes the inclusion $T \hookrightarrow T^n$ by $T \ni \theta \mapsto (\theta, \theta, \dots) \in T^n$.*

In other words, $P(T^n)$ is the set points in T^n identified up to a constant additive factor. Clearly, for any $f : T^n \to \mathbb{C}$ to descend to a well-defined function on $P(T^n)$ it must be constant on the cosets of the diagonal subgroup; in other words, it must satisfy $f(\mathrm{e}^{\mathrm{i}\phi_1 + \mathrm{i}\theta}, \mathrm{e}^{\mathrm{i}\phi_2 + \mathrm{i}\theta}, \dots) = f(\mathrm{e}^{\mathrm{i}\phi_1}, \mathrm{e}^{\mathrm{i}\phi_2}, \dots)$ for all $\theta \in T$. Hence, when studying designs on $P(T^n)$, we need only consider monomials on T^n where the degree and conjugate degree are equal. A degree t monomial on $P(T^n)$ therefore lifts to $\exp\bigl(\mathrm{i} \sum_{k=1}^t (\phi_{a_k} - \phi_{b_k})\bigr)$ for $a, b \in I_n^t$. We are thus now in a position to define a $P(T^n)$ t-design.

Definition 6.2.5 (Projective toric design)**.** *Fix an $n \in \mathbb{N} \cup \{\infty\}$ and $t \in \mathbb{N}$. Let $X \subset P(T^n)$ and (X, Σ, ν) be a measure space. X is called a $P(T^n)$* t-design *if for all $a, b \in I_n^t$,*

$$\int_X \exp\left(\mathrm{i} \sum_{j=1}^t (\phi_{a_j} - \phi_{b_j})\right) \mathrm{d}\nu(\phi) = \int_{P(T^n)} \exp\left(\mathrm{i} \sum_{j=1}^t (\phi_{a_j} - \phi_{b_j})\right) \mathrm{d}\mu_{n-1}(\phi). \tag{6.3}$$

Here we denote the unit-normalized Haar measure on $P(T^n)$ as simply μ_{n-1} since $P(T^n) \cong T^{n-1}$. X is called discrete *if ν is a counting measure, and is called* finite *if it is discrete and $|X| < \infty$. If X is finite, then $|X|$ is called the* size *of X.*

We note that, in the language of Ref. [189], a $P(T^n)$ design is a design on the maximal torus of the projective unitary group $T(\mathrm{PU}(n))$. It was shown in Ref. [9] that the two notions coincide[2]. Clearly a $P(T^n)$ t-design is also a $(t-1)$-design, since we can let $a_t = b_t$ and have the integrand become an arbitrary degree $(t-1)$ monomial. Additionally, a $P(T^n)$ t-design is also a $P(T^{n-1})$ t-design, as can be seen by picking a subset of indices.

Example 6.2.6. In this example, we consider $n = 2$ and $t = 1$. For any point in $P(T^2)$, which is itself an equivalence class, we choose a representative of the equivalence class to be zero in the first entry of the tuple. In other words, since the equivalence relation $\sim$ that we quotient T^n by to get $P(T^n) = {}^{T^n}\!/_{\sim}$ is $(\vartheta, \varphi) \sim (\vartheta + \theta, \varphi + \theta)$, we can always choose θ such that the first entry in the tuple is 0.

Let X be the discrete, uniformly-weighted measure space $X = \{(0,0), (0, 2\pi/3), (0, 4\pi/3)\} \subset P(T^2)$. Then,

$$\frac{1}{|X|} \sum_{\phi \in X} \mathrm{e}^{\mathrm{i}(\phi_a - \phi_b)} = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(\theta_a - \theta_b)} \, \mathrm{d}\theta_2 = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b. \end{cases} \tag{6.4}$$

Note that, since we fix the first entry of any element of $P(T^2)$ to be 0, $\theta_1 = 0$ and the Haar measure on $P(T^2)$ is $\mathrm{d}\theta_2$. Hence, we see that X is a $P(T^2)$ 1-design. ◇

Throughout this work, we use double braces to denote multisets, whereas single braces denote sets as usual; that is, $\{\!\{1, 2, 2\}\!\} = \{\!\{2, 1, 2\}\!\} \neq \{\!\{1, 2\}\!\}$, whereas $\{1, 2, 2\} = \{1, 2\} = \{2, 1\}$. Since the integrand in Eq. (6.3) contains only a finite number of projection maps, we can use Fubini's theorem to compute the integral on the right-hand side. By choosing a set of representatives of $P(T^n)$ to be those phases ϕ for which $p_1(\phi) = \phi_1 = 0$, we can think of $P(T^n)$ as $\{0\} \times T^{n-1}$. In this way, we have that $p_1(\phi) = 0$ for all ϕ. It follows that

[2]We note that, in contrast to our manuscript, Ref. [9] refers to $P(T^n)$ designs as T^n designs and refers to T^n designs as trigonometric cubature rules.

$X \subset \{0\} \times T^{n-1}$ is a $P(T^n)$ t-design if

$$\int_X \exp\left(\mathrm{i}\sum_{j=1}^{t}(\phi_{a_j} - \phi_{b_j})\right) \mathrm{d}\nu(\phi) = \int_{\{0\}\times T^{n-1}} \exp\left(\mathrm{i}\sum_{j=1}^{t}(\phi_{a_j} - \phi_{b_j})\right) \mathrm{d}\mu_{n-1}(\phi) \tag{6.5a}$$

$$= \begin{cases} 1 & \text{if } \{\!\{a_i \mid i \in \{1,\dots,t\}\}\!\} = \{\!\{b_i \mid i \in \{1,\dots,t\}\}\!\} \\ 0 & \text{otherwise} \end{cases}. \tag{6.5b}$$

Suppose that we set each $b_j = 1$. It follows that X must match integration of polynomials on T^{n-1} of degree t and conjugate degree 0 (because $\phi_{b_j} = 0$). Similarly, we can set each $a_j = 1$, and thus X must match integration of degree 0 and conjugate degree t. More generally, we see that it must match on monomials on T^{n-1} of degree (t_1, t_2) whenever $t_1 \leq t$ and $t_2 \leq t$. It follows that a T^{n-1} $(2t)$-design is a $P(T^n)$ t-design, and a $P(T^n)$ t-design is a T^{n-1} t-design. The reverse implications however do not hold in general.

By linearity, a $P(T^n)$ t-design exactly integrates all polynomials on $P(T^n)$ of degree t or less. It is the projective nature of the polynomials that we are integrating that give projective toric designs their interesting structure that is quite different than the structure of toric designs. For example, as we will see, for finite n, $P(T^n)$ 2-designs must be of size at least $n(n-1)+1$, and indeed this can be saturated for many n; in contrast, it is known that a T^n 4-design requires size at least $2n^2$, 3-design requires at least $4n$ points (which can often be achieved), and 2-design requires at least $2n$ points (and $2n+1$ can often be achieved) [269]. Indeed, the difference between toric designs (*i.e.* trigonometric cubature rules) and projective toric designs is analogous to the difference between (complex) spherical designs and (complex) projective designs.

Subsection 6.2.1: Constructions of projective toric designs

In this section, we present a few simple constructions in order to get a handle on projective toric designs. Later, in Section 6.3, we construct more (and smaller) projective toric t-designs by utilizing difference sets and Sidon sets from additive combinatorics [88]. Throughout this section, we write points in $P(T^n)$ as representatives in T^n with the first entry set to 0.

Our first example is a $P(T^n)$ 2-design of size n^2 whenever n is prime, and slightly larger when n is not prime. Note that this construction can be generalized to be size n^2 whenever n is a prime power, but we do not do this here. The generalized construction can be seen in the phases in the complete set of MUBs in prime-power dimensions given in Ref. [39].

Theorem 6.2.7 (Thm. C.9 of [9])**.** *Let $n \in \mathbb{N}$. Define p to be the smallest prime number strictly larger than* $\max(2, n)$ *(by the prime number theorem, $p \in \mathcal{O}(n + \log n)$). Let $X \subset T^n$ be the set*

$$X = \left\{ \left(0, 2\pi(q_1 + q_2)/p, 2\pi(2q_1 + 4q_2)/p, \ldots, 2\pi((n-1)q_1 + (n-1)^2 q_2)/p\right) \mid q_1 \in \mathbb{Z}_p, q_2 \in \mathbb{Z}_p \right\} \tag{6.6}$$

and v the constant map[3] *$v(\phi) = 1/|X|$. Then X with the counting measure weighted by v is a $P(T^n)$ 2-design.*

We can easily write out the construction for $n = 2$, where we have $p = 3$, and therefore $X = \{(0,0), (0, 2\pi/3), (0, 4\pi/3)\}$ with weight $v(\phi) = 1/3$ is a $P(T^2)$ 2-design. We show the construction in Fig. 6.1 for this example of $n = 2$ with $p = 3$ as well as for $n = 3$ with $p = 5$.

We can extend this construction to the case when $n = \infty$.

Theorem 6.2.8. *Let $X \subset T^\infty$ be the set*

$$X = \left\{ \left(0, \vartheta + \varphi, 2\vartheta + 4\varphi, \ldots, j\vartheta + j^2\varphi, \ldots\right) \mid \vartheta, \varphi \in [0, 2\pi] \right\} \tag{6.7}$$

[3] We corrected a minor error in Thm. C.9 of [9]. Namely, the map v was stated as $v(\phi) = 1/p^2$. This is correct for all $n > 2$, as $|X| = p^2$. However, when $n = 2$, $|X| = p = 3$.

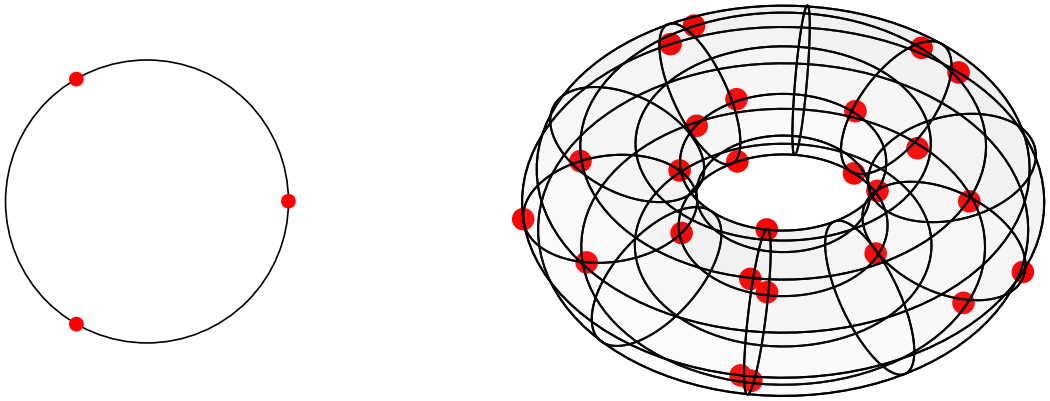

Figure 6.1: The construction of the 2-design in Theorem 6.2.7 for (left) $n = 2$ with $p = 3$ and (right) $n = 3$ with $p = 5$. Note we are representing points in $P(T^n)$ here as points in T^{n-1} by discarding the first coordinate which we fix to 0. The number of points in the design for (left) $n = 2$ is p and for (right) $n = 3$ is $p^2 = 25$.

and ν the unit normalized Lebesgue measure on $[0, 2\pi]^2$ (i.e. $\mathrm{d}\nu = \mathrm{d}\vartheta\,\mathrm{d}\varphi/(2\pi)^2$). Then X is a T^∞ 2-design.

Proof. For any $a, b, c, d \in \mathbb{N}$,

$$\int_X \exp(\mathrm{i}(\phi_a + \phi_b - \phi_c - \phi_d))\,\mathrm{d}\nu(\phi) = \int_{[0,2\pi]^2} \exp\big(\mathrm{i}\vartheta(a + b - c - d) + \mathrm{i}\varphi(a^2 + b^2 - c^2 - d^2)\big)\,\frac{\mathrm{d}\vartheta\,\mathrm{d}\varphi}{(2\pi)^2} \tag{6.8a}$$

$$= \begin{cases} 1 & \text{if } a + b = c + d \ \wedge\ a^2 + b^2 = c^2 + d^2 \\ 0 & \text{otherwise} \end{cases} \tag{6.8b}$$

$$= \begin{cases} 1 & \text{if } \{\!\{a, b\}\!\} = \{\!\{c, d\}\!\} \\ 0 & \text{otherwise} \end{cases}, \tag{6.8c}$$

where in the last line we used [9, Lem. C.10]. □

We now consider a construction for arbitrary t.

Theorem 6.2.9 (Thm. C.8 of [9]). *Let $n, t \in \mathbb{N}$, and $X \subset T^n$ be the set*

$$X = \left\{(0, 2\pi d_1/(t+1), 2\pi d_2/(t+1), \ldots, 2\pi d_{n-1}/(t+1)) \mid d \in \mathbb{Z}_{t+1}^{n-1}\right\}, \tag{6.9}$$

and v be the constant map $v(\phi) = (t+1)^{-(n-1)}$. Then X with the counting measure weighted by v is a $P(T^n)$ t-design.

Example 6.2.10 ($n = 2$, $t = 3$). We have

$$\begin{aligned} X = \Bigg\{ &(0,0,0), \left(0,0,2\pi\frac{1}{3}\right), \left(0,2\pi\frac{1}{3},0\right), \left(0,0,2\pi\frac{2}{3}\right), \left(0,2\pi\frac{2}{3},0\right), \\ &\left(0,2\pi\frac{1}{3},2\pi\frac{2}{3}\right), \left(0,2\pi\frac{2}{3},2\pi\frac{1}{3}\right), \left(0,2\pi\frac{1}{3},2\pi\frac{1}{3}\right), \left(0,2\pi\frac{2}{3},2\pi\frac{2}{3}\right) \Bigg\}, \end{aligned} \tag{6.10}$$

with $v(\phi) = 1/9$, is a $P(T^3)$ 2-design. ◇

We now extend this construction to $n = \infty$.

Theorem 6.2.11. *Let $t \in \mathbb{N}$ and $X_1 \subset T$ be the discrete probability space $X_1 = \{2\pi d/(t+1) \mid d \in \mathbb{Z}_{t+1}\}$. Let $X = \prod_{i\in\mathbb{N}} X_1$ and its σ-algebra be generated by sets of the form $\prod_{i\in\mathbb{N}} A_i$ where each A_i in the power set $A_i \in \mathcal{P}(X_1)$ and for all but finitely many i we have $A_i = X_1$. Define ν by its action $\nu(A) = \prod_{i\in\mathbb{N}}(|A_i| / |X_1|)$, and note that ν uniquely extends to a measure on X [276, Thm. 10.6.1]. Then X is a T^∞ t-design.*

Proof. Let $m = \max(\max_j a_j, \max_j b_j)$. Since t is finite, we are only ever dealing with a finite number of projection maps p_i in the integrand. Therefore, we can apply Fubini's theorem to separate the integral $\int_X$ into a product of an integral over X_1^m and an integral over the rest of the space. Hence,

$$\int_X \exp\left(\mathrm{i}\sum_{j=1}^{t}(\phi_{a_j} - \phi_{b_j})\right) \mathrm{d}\nu(\phi) = \frac{1}{\left|\mathbb{Z}_{t+1}^m\right|} \sum_{d\in\mathbb{Z}_{t+1}^m} \exp\left(\frac{2\pi\mathrm{i}}{t+1}\sum_{j=1}^{t}(d_{a_j} - d_{b_j})\right) \tag{6.11a}$$

$$= \begin{cases} 1 & \text{if } \{\!\{a_j \mid j \in \{1,\dots,t\}\}\!\} = \{\!\{b_j \mid j \in \{1,\dots,t\}\}\!\} \\ 0 & \text{otherwise} \end{cases}. \tag{6.11b}$$

□

Finally, for completeness, we note the asymptotic existence theorem proven in Ref. [189].

Theorem 6.2.12 (Thm. 3.3 and Cor. 5.4 of [189]). *Asymptotically in $n \to \infty$ but for finite n, a $P(T^n)$ t-design must have size at least $\frac{n^t(1-o(1))}{\lceil t/2\rceil!\lfloor t/2\rfloor!}$ and there exists t-designs of size $n^t(1+o(1))$.*

Subsection 6.2.2: Minimal projective toric designs

A very natural question that one can ask is *what is the size of the smallest projective toric t-design?* We call such designs *minimal*. Ref. [9, Prop. C.11] proved a lower bound on the size of minimial projective toric 2-designs. In this section, we generalize this bound and prove a lower bound on the size of minimal projective toric t-designs for all t. In the case when t is even, we conjecture that this bound is tight. In Section 6.3, we show that the $t = 2$ bound can be saturated in many dimensions.

We begin by, for all $n, s \in \mathbb{N}$, defining the set

$$P_s^{(n)} \coloneqq \left\{ \mathbf{q} - \mathbf{r} \;\middle|\; \mathbf{q}, \mathbf{r} \in \mathbb{N}_0^n,\ \sum_{i=1}^n q_i = \sum_{i=1}^n r_i = s \right\}. \tag{6.12}$$

An element $\mathbf{q} - \mathbf{r} \in P_s^{(n)}$ corresponds to a monomial $\exp(\mathrm{i}\sum_{j=1}^n (q_j - r_j)\phi_j)$ on $P(T^n)$. We show that $|P_s^{(n)}|$ is the s^{th} element of the crystal ball sequence corresponding to the root lattice $A_{n-1} \coloneqq \{\boldsymbol{v} \in \mathbb{Z}^n \mid \sum_{i=1}^n v_i = 0\}$ [272, 273], and therefore arrive at the explicit formula for $|P_t^{(n)}|$ given in Eq. (6.13). We begin by defining the crystal ball sequence of A_{n-1}. Let $S_{n-1}(t)$ denote the number of vertices of A_{n-1} a distance t away from some fixed vertex, where

we define distance for the lattice A_{n-1} as follows: letting $\mathcal{R} \coloneqq \{\mathbf{e}_i - \mathbf{e}_j \mid i, j \in \{1, \ldots, n\}\}$ be the roots of A_{n-1}, the distance between $\mathbf{x}, \mathbf{y} \in A_{n-1}$ is the smallest d such that $\mathbf{x} - \mathbf{y} \in d\mathcal{R}$, where $d\mathcal{R} := \mathcal{R} + \mathcal{R} + \cdots + \mathcal{R}$ is the d-fold set sum of $\mathcal{R}$. The sequence $(S_{n-1}(t))_{t\in\mathbb{N}_0}$ is the *coordination sequence* of A_{n-1} [272]. The *crystal ball numbers* are the partial sums $G_{n-1}(s) = \sum_{x=0}^{s} S_{n-1}(x)$ [272]. The explicit formula for $G_{n-1}(s)$ is [272, 273]

$$G_{n-1}(s) = {}_3F_2(1-n, -s, n; 1, 1; 1) = \sum_{i=0}^{s} \binom{n-1}{i}^2 \binom{n-i+s-1}{s-i}, \tag{6.13}$$

where ${}_3F_2$ denotes the generalized hypergeometric function [141–144]. We can easily see that $P_s^{(n)} = s\mathcal{R}$, and furthermore $G_{n-1}(s) = |s\mathcal{R}|$ by definition since it is precisely the set of all points that are reachable within a path of at most s edges. It therefore follows that

$$|P_s^{(n)}| = G_{n-1}(s). \tag{6.14}$$

We recall that Ref. [9] showed the equivalence of $P(T^n)$ designs and designs on the algebraic torus $T(\mathrm{PSU}(n))$ as defined in Ref. [189]. Ref. [189] further explored the connection between such designs and the root lattice of $\mathrm{PSU}(n)$, which is A_{n-1}. This gives a hint as to why A_{n-1} shows up in the analysis of projective toric designs. Indeed, each point in A_{n-1} corresponds to a monomial on $P(T^n)$. $P_s^{(n)}$ is precisely all points on A_{n-1} a distance of less than or equal to s from the origin. Since the origin corresponds to the constant monomial (*i.e.* degree 0), $P_s^{(n)}$ corresponds to all monomials of degree less than or equal to s.

We now prove a lower bound on the size of projective toric designs. We note that this bound is compatible with the asymptotic bound given in Theorem 6.2.12. One can see this by using the asymptotic expansion of the binomial coefficients in Eq. (6.13).

Proposition 6.2.13. *Let $n \in \mathbb{N}$ and (X, Σ, ν) be a finite $P(T^n)$ t-design. Then $|X| \geq G_{n-1}(\lfloor t/2 \rfloor)$, where $G_{n-1}(s)$ is given in Eq. (6.13).*

Proof. We prove the bound for even t. The bound for odd t is then automatically valid since

the minimal size of a $(t+1)$-design is at least as large as the minimal size of a t-design. We therefore restrict our attention to even t for the rest of the proof.

Since X is a finite, discrete measure space, we can rewrite $\int_X (\cdot)\,\mathrm{d}\nu$ as $\sum_{\phi\in X} v(\phi)(\cdot)$. The projective toric t-design condition can be expressed as follows. Let each $\phi \in X$ label a basis element of $V \coloneqq \mathbb{C}^{|X|}$ so that $\{|\phi\rangle \mid \phi \in X\}$ is an orthonormal basis of V. Then for $\mathbf{k} \in P_{t/2}^{(n)}$, define $|\mathbf{k}\rangle = \sum_{\phi\in X} \sqrt{v(\phi)}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\phi}\,|\phi\rangle$. The t-design condition is equivalently stated as $\langle \mathbf{k}|\mathbf{k}'\rangle = \delta_{\mathbf{k},\mathbf{k}'}$. Hence, $\{|\mathbf{k}\rangle \mid \mathbf{k} \in P_{t/2}^{(n)}\}$ must be orthonormal in V, meaning that $|P_{t/2}^{(n)}| \leq \dim V = |X|$. The proposition then follows from Eq. (6.14). □

Furthermore, we can prove that a minimal t-design for even t must be uniformly weighted.

Proposition 6.2.14. *Let $X \subset P(T^n)$ and let $v\colon X \to (0,\infty)$ define a weighted discrete measure on X. Suppose the measure space defined by X and v is a minimal t-design with t even. Then $v(\theta) = 1/|X|$.*

Proof. This proof essentially follows that of Ref. [269, Thm. 2.2]. The $P(T^n)$ t-design condition is written as $MM^\dagger = \mathbb{I}_{|P_{t/2}^{(n)}|\times|P_{t/2}^{(n)}|}$, where $M_{\mathbf{k},\theta} = \sqrt{v(\theta)}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\theta}$. If X is minimal—that is, if $|X| = |P_{t/2}^{(n)}|$—then M is a square matrix so that $MM^\dagger = \mathbb{I}$ if and only if $M^\dagger M = \mathbb{I}$. From the latter condition, it follows that $\delta_{\theta,\theta'} = \sqrt{v(\theta)v(\theta')}\sum_{k\in P_{t/2}^{(n)}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\theta-\theta')}$. When $\theta = \theta'$, we therefore find that $v(\theta) = 1/|P_{t/2}^{(n)}| = 1/|X|$. □

Finally, we conjecture that the bound given in Proposition 6.2.13 is tight for even t.

Conjecture 6.2.15. *When t is even, the bound given in Proposition 6.2.13 is tight in the sense that there are infinitely many dimensions n for which the bound is saturable.*

In Section 6.3, we show how minimal t-designs are related to difference sets. Using this connection, we construct an infinite family of minimal 2-designs that indeed saturate the bound given in Proposition 6.2.13, and we derive a bound on the size of dense difference sets.

Section 6.3: Relation to difference sets

We say that $X \subset P(T^n)$ is a *group toric t-design* if X is a t-design and also inherits group structure from $P(T^n)$. In this section, we consider the case when X is a cyclic group for finite n and a circle group for $n = \infty$. In this case, we find connections to Sidon sets and difference sets [88]. Using this, in Section 6.3.1, we construct minimal $P(T^n)$ 2-designs whenever $n-1$ is a prime power, and more generally we construct t-designs of size $\frac{(n-1)^{t+1}-1}{n-2}$ whenever $n-1$ is a prime power.

We begin with the infinite case. Suppose that $X \subset P(T^\infty)$ is a t-design and isomorphic to the circle group U(1). Then there is a single element $z \in \mathbb{Z}^\infty$ such that $X = \{\theta z = (\theta z_1, \theta z_2, \dots) \mid \theta \in [0, 2\pi]\}$. In order for X to be a design, it must be that

$$\int_0^{2\pi} \exp\left(\mathrm{i}\theta \sum_{j=1}^{t} (z_{a_j} - z_{b_j})\right) \frac{\mathrm{d}\theta}{2\pi} = \begin{cases} 1 & \text{if } \{\!\{a_j \mid j \in \{1,\dots,t\}\}\!\} = \{\!\{b_j \mid j \in \{1,\dots,t\}\}\!\} \\ 0 & \text{otherwise} \end{cases} \tag{6.15}$$

for all $a, b \in \mathbb{N}^t$. It follows that z must satisfy

$$\left(\sum_{j=1}^{t} z_{a_j} = \sum_{j=1}^{t} z_{b_j}\right) \iff \left(\{\!\{a_j \mid j \in \{1,\dots,t\}\}\!\} = \{\!\{b_j \mid j \in \{1,\dots,t\}\}\!\}\right). \tag{6.16}$$

In other words, the sum of any t elements of z must be unique. If we restrict z to be in $\mathbb{Z}^\infty_{\geq 0}$, then Eq. (6.16) is exactly the definition for z to be a *B_t set* [88, Def. 4.27]. In the special case of $t = 2$, we need to find a $z \in \mathbb{Z}^\infty_{\geq 0}$ such that $z_a + z_b = z_c + z_d$ if and only if $\{\!\{a, b\}\!\} = \{\!\{c, d\}\!\}$. Such a z is called a *Sidon set* [88].

Definition 6.3.1 (B_t and Sidon sets [88])**.** *A B_t set*[4] *is an element $z \in \mathbb{Z}^\infty_{\geq 0}$ satisfying Eq. (6.16) for all $a, b \in \mathbb{N}^t$. A* Sidon set *is a B_2 set.*

[4]Note that we are considering z to be a tuple and yet calling it a difference "set". It is understood that we are talking about the set $\{z_a \mid a \in \mathbb{N}\}$.

We have therefore proven the following proposition.

Proposition 6.3.2. *Group $P(T^\infty)$ t-designs isomorphic to the circle group are in one-to-one correspondence with B_t sets.*

We next give a simple example of a B_t set.

Example 6.3.3 (Exponential B_t set)**.** Let $z \in \mathbb{Z}^\infty$ be defined by $z_a = t^a$. In this case, z_a written in base t is $100\ldots0$, a 1 followed by a 0s. It follows easily that every sum is unique up to reordering. ◇

We now discuss finite n. Suppose that $X \subset P(T^n)$ is a t-design and isomorphic to the cyclic group $\mathbb{Z}_m$. It follows that X is a size m t-design and is generated by a fixed $z \in \mathbb{Z}_m^n$. In order for X to be a design, it must be that

$$\sum_{d=0}^{m-1} \exp\left(\frac{2\pi \mathrm{i} d}{m} \sum_{j=1}^{t} (z_{a_j} - z_{b_j})\right) = \begin{cases} 1 & \text{if } \{\!\{a_j \mid j \in \{1,\ldots,t\}\}\!\} = \{\!\{b_j \mid j \in \{1,\ldots,t\}\}\!\} \\ 0 & \text{otherwise} \end{cases} \tag{6.17}$$

for all $a, b \in I_n^t$, where recall that $I_n = \{1, 2, \ldots, n\}$. It follows that z must satisfy

$$\left(\sum_{j=1}^{t} z_{a_j} \equiv \sum_{j=1}^{t} z_{b_j} \pmod m\right) \iff \left(\{\!\{a_j \mid j \in \{1,\ldots,t\}\}\!\} = \{\!\{b_j \mid j \in \{1,\ldots,t\}\}\!\}\right). \tag{6.18}$$

In other words, the sum of any t elements of z must be unique, or equivalently,

$$\left|\left\{\sum_{j=1}^{t} z_{a_j} \bmod m \mid a \in I_n^t\right\}\right| = \binom{n+t-1}{t}. \tag{6.19}$$

Eq. (6.18) is precisely the definition for z to be a *B_t mod m set of size n* [88].

Definition 6.3.4 (Modular B_t and Sidon sets [88])**.** *A B_t* mod m set of size n *is an element $z \in \mathbb{Z}_m^n$ satisfying Eq. (6.18) for all $a, b \in I_n^t$. A* Sidon set of size n mod m *is a B_2 mod m set of size n.*

We have therefore shown the following proposition.

Proposition 6.3.5. *Group $P(T^n)$ t-designs isomorphic to the cyclic group $\mathbb{Z}_m$ are in one-to-one correspondence with B_t mod m sets of size n.*

Given Proposition 6.3.5 and the bound in Proposition 6.2.13, we immediately arrive at the following corollary.

Corollary 6.3.6. *Any B_t mod m set must have size n satisfying $m \geq G_{n-1}(\lfloor t/2 \rfloor)$, where $G_{n-1}(s)$ is given in Eq. (6.13). Furthermore, if Conjecture 6.2.15 is true, then this bound is tight for even t in the sense that there are infinitely many dimensions n for which the bound is saturable.*

We have been unable to find the bound in Corollary 6.3.6 in the existing literature on difference sets. If this bound is indeed new, it illustrates the utility of studying projective toric designs due to the many interesting mathematical objects to which they relate.

In the special case of $t = 2$, $B_{t=2}$ mod m sets are called a *Sidon sets of size n mod m*. Notably, by a simple counting argument, any Sidon set of size n mod m must satisfy $m \geq n(n-1)+1$.[5] Further, for many but not all n, this bound can be saturated, as we discuss later. When the bound is saturated, we say the Sidon set is *dense*. Hence, for every n for which there is a Sidon set of size n mod $n(n-1)+1$, there is a *minimal* $P(T^n)$ 2-design—that is, a $P(T^n)$ 2-design of size $n(n-1)+1$, hence saturating the lower bound from Proposition 6.2.13.

For one example of a dense Sidon set, consider $n = 6$ and $m = G_{n-1}(1) = n(n-1)+1 = 31$. Then one can easily check that $z = (0, 1, 3, 8, 12, 18)$ is a Sidon set and thus gives rise to a $P(T^6)$ 2-design of size 31. A simple numerical search however reveals that there does not exist a Sidon set of size 7 mod $7(7-1)+1 = 43$. Furthermore, by the classification of finite

[5] The Sidon set condition can be restated as stipulating that $z_a - z_c \equiv z_d - z_b$ if and only if $\{\!\{a, b\}\!\} = \{\!\{c, d\}\!\}$. We therefore need $z_a - z_c$ to be unique for all a and c. First choose an $a \in I_n$ and then choose a $c \in I_n$ with $c \neq a$. This gives us $n(n-1)$ distinct values. Further, we have one more value—namely 0—coming from when $a = c$.

abelian groups, any group of order 43 must be isomorphic to $\mathbb{Z}_{43}$. Therefore, we have the following corollary.

Corollary 6.3.7. *Either there are no $P(T^7)$ 2-designs of size saturating the lower bound given in Proposition 6.2.13, or such a saturating design cannot be isomorphic to a group.*

Subsection 6.3.1: Explicit families of designs from Singer sets

There is a general construction of dense Sidon sets—called Singer sets—whenever $n-1$ is a prime power [89]. Thus, with this, we have constructed minimal $P(T^n)$ 2-designs whenever $n-1$ is a prime power, and these designs are isomorphic to the cyclic group $\mathbb{Z}_{n(n-1)+1}$. For completeness, we review the Singer set construction in Appendix D.1. However, we note that the details of the Singer set construction are not necessary to understand for our work. Indeed, our results only use that such a construction *exists*. For reference, we provide code for constructing Singer sets [278].

Indeed more generally, we review Singer's construction in Lemma D.1.2 of B_t mod $\frac{(n-1)^{t+1}-1}{n-2}$ sets of size n whenever $n-1$ is a prime power. Using Proposition 6.3.5, we have therefore constructed explicit $P(T^n)$ t-designs of size $\frac{(n-1)^{t+1}-1}{n-2}$ whenever $n-1$ is a prime power, and these designs are isomorphic to the cyclic group $\mathbb{Z}_{\frac{(n-1)^{t+1}-1}{n-2}}$. Furthermore, since the restriction of a $P(T^m)$ t-design to $P(T^n)$ for $n \leq m$ is still a t-design, it follows that for all n we have constructed explicit $P(T^n)$ t-designs of size $\frac{(m-1)^{t+1}-1}{m-2}$, where m is the smallest integer greater than or equal to n such that $m-1$ is a prime power.

Finally, we recall that a $P(T^n)$ t-design is a T^{n-1} t-design (see the discussion below Eq. (6.5)). We also note that a $P(T^n)$ t-design be made into a T^n $(2t)$-design by twirling over a $T^1 = S^1$ $(2t)$-design. For example, if a set $\Phi \subset \{0\} \times T^{n-1}$ is a $P(T^n)$ t-design, then the set $\{\phi + (\theta, \dots, \theta) \mid \phi \in \Phi, \theta \in \Theta\}$ is a T^n $(2t)$-design when Θ is a S^1 $(2t)$-design. We

can see this as follows. Suppose $\sum_{j=1}^{n} |\alpha_j| \leq 2t$. Then,

$$\frac{1}{|\Phi| \cdot |\Theta|} \sum_{\phi \in \Phi} \sum_{\theta \in \Theta} \mathrm{e}^{\mathrm{i} \sum_{j=1}^{n} \alpha_j (\phi_j + \theta)} = \int_{T^1} \mathrm{e}^{\mathrm{i}\theta \sum_{j=1}^{n} \alpha_j} \, \mathrm{d}\mu_1(\theta) \times \frac{1}{|\Phi|} \sum_{\phi \in \Phi} \mathrm{e}^{\mathrm{i} \sum_{j=1}^{n} \alpha_j \phi_j} \tag{6.20a}$$

$$= \delta_{\sum_{j=1}^{n} \alpha_j = 0} \frac{1}{|\Phi|} \sum_{\phi \in \Phi} \mathrm{e}^{\mathrm{i} \sum_{j=1}^{n} \alpha_j \phi_j} \tag{6.20b}$$

$$= \delta_{\sum_{j=1}^{n} \alpha_j = 0} \int_{P(T^n)} \mathrm{e}^{\mathrm{i} \sum_{j=1}^{n} \alpha_j \phi_j} \, \mathrm{d}\mu_{n-1}(\phi) \tag{6.20c}$$

$$= \prod_{j=1}^{n} \delta_{\alpha_j, 0}, \tag{6.20d}$$

where in Eq. (6.20c) we used that $\mathrm{e}^{\mathrm{i} \sum_{j=1}^{n} \alpha_j \phi_j}$ is at most a degree t monomial on $P(T^n)$ when $\sum_{j=1}^{n} \alpha_j = 0$. The set $\Theta = \left(\frac{2\pi}{2t+1}\right) \mathbb{Z}_{2t+1}$ is a S^1 $(2t)$-design. We have therefore constructed explicit T^n $(2t)$-designs of size $(2t+1) \times \frac{(m-1)^{t+1}-1}{m-2}$ for all n.

Section 6.4: Relation to quantum state designs and MUBs

Projective toric designs are closely connected to complex-projective designs [36–47], continuous-variable (CV) rigged designs [9], and complete sets of mutually unbiased bases (MUBs) [238]. These connections arise by concatenating projective toric and simplex designs in order to generate elements in complex-projective space, which in turn satisfy the design condition. We discuss the connection here. In Section 6.4.1, we set up the connection between projective toric designs and quantum state designs and use it to construct *almost minimial* quantum state 2-designs (*ie.* quantum state 2-designs in d dimensions of size d^2+1). Using this connection, in Sections 6.4.2 and 6.4.3, we find a close connection between projective toric designs and MUBs, and we use this connection to prove various results. Namely, in Section 6.4.2, we disprove Zhu's conjecture, and in Section 6.4.3, we characterize a fundamental difference between complete sets of MUBs in prime-power dimensions vs in dimension 6 in terms of the group structure of the associated projective toric designs.

Subsection 6.4.1: Quantum state designs from projective toric designs

Denote the complex unit sphere by $\Omega_d = \{z \in \mathbb{C}^d \mid \sum_{i=1}^d |z_i|^2 = 1\}$, which can be identified with S^{2d-1}. Let $\mathbb{CP}^{d-1}$ be complex-projective space $\Omega_d/\mathrm{U}(1)$. Pick an orthonormal basis $\{|n\rangle \mid n \in \{1, \dots, d\}\}$ of $\mathbb{C}^d$. A polynomial f on Ω_d descends to a well-defined polynomial on $\mathbb{CP}^{d-1}$ if and only if it is invariant under the action of $\mathrm{U}(1)$—that is, $f(\mathrm{e}^{\mathrm{i}\theta}|\psi\rangle) = f(|\psi\rangle)$ for all θ and $|\psi\rangle \in \Omega_d$. It follows that all degree t monomials on $\mathbb{CP}^{d-1}$ are of the form $\prod_{i=1}^t \langle a_i|\psi\rangle\langle\psi|b_i\rangle$ for $a, b \in I_d^t$ (recall that $I_d = \{1, 2, \dots, d\}$). A $\mathbb{CP}^{d-1}$ t-design is thus a measure space (X, Σ, ν) such that, for all $a, b \in I_d^t$,

$$\int_X \left(\prod_{i=1}^t \langle a_i|\psi\rangle\langle\psi|b_i\rangle\right) \mathrm{d}\nu(\psi) = \int_{\mathbb{CP}^{d-1}} \left(\prod_{i=1}^t \langle a_i|\psi\rangle\langle\psi|b_i\rangle\right) \mathrm{d}\psi = \frac{\Pi_t^{(d)}(a;b)}{\operatorname{Tr}\Pi_t^{(d)}}, \tag{6.21}$$

where $\Pi_t^{(d)}$ is the projector onto the symmetric subspace of $(\mathbb{C}^d)^{\otimes t}$,

$$\Pi_t^{(d)}(a;b) \coloneqq \left(\bigotimes_{i=1}^t \langle a_i|\right) \Pi_t^{(d)} \left(\bigotimes_{i=1}^t |b_i\rangle\right), \tag{6.22}$$

and $\mathrm{d}\psi$ denotes the Fubini-Study volume measure on $\mathbb{CP}^{d-1}$. The last equality is a simple consequence of Schur's lemma and the unitary invariance of $\mathrm{d}\psi$ [43, 45][9, Ap. C3].

Let $\Delta^{d-1} = \{p \in [0,1]^d \mid \sum_{i=1}^d p_i = 1\}$ denote the $(d-1)$-dimensional simplex. Simplex t-designs have analogous definitions to those of toric and complex-projective designs [185, 187–190]. Any vector $|\psi\rangle \in \Omega_d$ can be represented as $|p, \phi\rangle \coloneqq \sum_{n=1}^d \sqrt{p_n}\mathrm{e}^{\mathrm{i}\phi_n}|n\rangle$ for some (not necessarily unique) $p \in \Delta^{d-1}$ and $\phi \in T^d$. For a complex unit vector $|\psi\rangle \in \Omega_d$, let $[|\psi\rangle]$ denote the equivalence class corresponding to a point in $\mathbb{CP}^{d-1}$. Let $\pi\colon \Delta^{d-1} \times P(T^d) \to \mathbb{CP}^{d-1}$ be defined by $(p, \phi) \mapsto [|p, \phi\rangle]$, where ϕ is any representative of an equivalence class in $T^d/T = P(T^d)$. The pullback of the Fubini-Study volume form along π is precisely the Lebesgue measure on Δ^{d-1} times the Lebesgue measure on $P(T^d)$ (see Appendix D.2). Together, this implies that the concatenation of a Δ^{d-1} t-design and a $P(T^d)$ t-design yields

a $\mathbb{CP}^{d-1}$ t-design [9, 189].

We note that the analogous result holds for the complex sphere Ω_d; namely, concatenation of a Δ^{d-1} t-design and a toric $(2t)$-design (see Definition 6.2.2) yields a Ω_d t-design. The reason that we only need a projective toric design in the $\mathbb{CP}^{d-1}$ case, as opposed to a full toric design as in the Ω_d case, is because polynomials on $\mathbb{CP}^{d-1}$ are more restricted than on Ω_d. On Ω_d, $z_1 z_2 \bar{z}_3$ is a valid monomial. On the other hand, this is an invalid monomial on $\mathbb{CP}^{d-1} = \Omega_d / \mathrm{U}(1)$ since it varies under the action of $\mathrm{U}(1)$.

One particularly nice simplex 2-design contains the extremal points and the centroid (see e.g. [9, Thm. C4]), which we show in the following proposition (see Example 6.4.2 for a simple example).

Proposition 6.4.1. *Let $c = (1/d, \dots, 1/d) \in \Delta^{d-1}$ be the centroid of the simplex and $e^{(1)} = (1, 0, \dots, 0)$, …, $e^{(d)} = (0, 0, \dots, 1)$ be the extremal points, and define $D = \{c, e^{(1)}, \dots, e^{(d)}\}$. Define the weight function $w \colon D \to [0, 1]$ by $w(e^{(j)}) = \frac{1}{d(d+1)}$ and $w(c) = \frac{d}{d+1}$. The discrete probability space defined by (D, w) is a Δ^{d-1} 2-design.*

When concatenating the extremal points $e^{(j)}$ of the simplex with a projective toric design, we get the basis vectors $[|j\rangle] \in \mathbb{CP}^{d-1}$, since $[|e^{(j)}, \phi\rangle] = [|j\rangle]$ for any ϕ. When concatenating the centroid with a finite-sized projective toric design X, we get a collection of points $\{[|c, \phi\rangle] \in \mathbb{CP}^{d-1} \mid \phi \in X\}$. Hence, the total number of points in the resulting complex-projective design is $d + |X|$. Recalling Proposition 6.2.13, we have that $|X| \geq d(d-1) + 1$. Furthermore, from Section 6.3, we found an explicit construction using Singer sets of these minimal projective toric designs whenever $d+1$ is a prime power. It follows that the resulting complex-projective 2-design is of size $d^2 + 1$ (note that this complex-projective design is not uniformly-weighted). Interestingly, the smallest possible complex-projective 2-design—also called a SIC-POVM—has size d^2. The existence of SIC-POVM's in all dimensions d is still an open problem.

These *almost-minimal* $\mathbb{CP}^{d-1}$ 2-designs that we just constructed using Singer sets—$\mathbb{CP}^{d-1}$

2-designs of size d^2+1—were first constructed in Ref. [274]. Notably, however, our utilization of projective toric designs indicates a possible path toward extending such constructions to higher t-designs.

Example 6.4.2 ($d = 3$)**.** We construct the above almost-minimal $\mathbb{CP}^{d-1}$ 2-design in the case of $d = 3$. Let us utilize the minimal $P(T^3)$ 2-design given by the mod 7 Sidon set $(0, 1, 3)$. The corresponding projective toric design is given by the phases

$$\left\{ \left(0, \frac{2\pi k}{7}, \frac{2\pi k}{7} \times 3 \right) \mid k \in \mathbb{Z}_7 \right\}, \tag{6.23}$$

where we understand $(0, \theta, \phi) \in T^3$ to be a representative of an equivalence class in $P(T^3)$. Denote by (p_0, p_1, p_2) an element of Δ^2. Consider the Δ^2 2-design from Proposition 6.4.1 given by the centroid $(1/3, 1/3, 1/3)$ weighted by $3/4$ and the extremal points $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ each weighted by $1/12$. Finally, denote points in $\mathbb{CP}^2$ by $[|\psi\rangle]$ for $|\psi\rangle$ a unit vector in $\mathbb{C}^3$, and fix an orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$. Let

$$X = \{[|0\rangle], [|1\rangle], [|2\rangle]\} \cup \left\{ [|\psi_k\rangle] := \left[\frac{1}{\sqrt{3}} \left(|0\rangle + \mathrm{e}^{2\pi \mathrm{i} k/7} |1\rangle + \mathrm{e}^{2\pi \mathrm{i} k \times 3/7} |2\rangle \right) \right] \mid k \in \mathbb{Z}_7 \right\}. \tag{6.24}$$

Turn X into a discrete measure space by weighting $[|0\rangle], [|1\rangle]$, and $[|2\rangle]$ each by $1/12$ and each $[|\psi_k\rangle]$ by $3/(4 \times 7)$. The resulting measure space is a $\mathbb{CP}^2$ 2-design of size 10. ◇

Similar to the finite-dimensional case above, we see analogous results in infinite-dimensional quantum systems. Ref. [79] introduced the notion of a continuous variable (CV) t-design. Ref. [9] proved that such designs do not exist and therefore introduced *rigged* CV t-designs. A simplex design can be generalized to the unnormalized infinite-dimensional simplex. It then follows that the concatenation of an infinite-dimensional simplex t-design and a $P(T^\infty)$ t-design yields a rigged CV t-design. We therefore see that designs on the infinite-dimensional projective torus $P(T^\infty)$ are closely related to designs on other infinite-dimensional spaces.

Subsection 6.4.2: MUBs and quantum state designs: counterexamples to Zhu's conjecture

In this subsection, we explicitly derive the relationship between complete sets of MUBs and projective toric 2-designs. We then use this relationship to disprove Zhu's conjecture regarding the structure of MUBs [95, Conj. 1].

The non-existence of complete sets of MUBs in non-prime-power dimensions has been a well-known and long-outstanding question in quantum information theory. Since the question of non-existence has proven incredibly difficult, there have been many related conjectures made in hopes of making some progress. One such conjecture is as follows.

> **Zhu's conjecture** [95, Conj. 1]. Any (uniformly-weighted) quantum state 2-design in d dimensions of size no more than $d(d+1)$ is either a complete set of MUBs or a SIC-POVM.

See also Ref. [275] for a discussion on Zhu's conjecture. In this subsection, we disprove this conjecture by explicitly constructing counterexamples that utilize the difference set construction of projective toric designs. We begin by recalling the definition of complete sets of MUBs. We then show the relationship of MUBs to projective toric designs, and we restate Zhu's conjecture in terms of projective toric designs. We then show explicit examples of projective toric designs that violate Zhu's conjecture. We then prove a nice characterization of complete sets of MUBs in dimension 6 in terms of non-group projective toric 2-designs (*cf.* Proposition 6.4.6). Finally, we discuss one possible modification of Zhu's conjecture and discuss potential paths towards proving this new conjecture (*cf.* Conjecture 6.4.8).

For brevity, we take Ref. [41, Thms. 3, 4] as our definition[6] of a complete set of MUBs.

Definition 6.4.3. *A set $\mathcal{M} \subset \mathbb{CP}^{d-1}$ is a complete set of MUBs if and only if*

[6] The second condition in Definition 6.4.3 is actually implied by the first and third [279] (we thank Daniel McNulty for pointing this out).

1. $|\mathcal{M}| = d(d+1)$,

2. $\mathcal{M}$ *is a (uniformly-weighted) quantum state* 2*-design (ie. a* $\mathbb{CP}^{d-1}$ 2*-design), and*

3. *for every* $|\psi\rangle \neq |\varphi\rangle \in \mathcal{M}$, $|\langle\psi|\varphi\rangle|^2 \in \{0, 1/d\}$.

In Ref. [9, App. F], it was shown that projective toric designs are closely related to complete sets of MUBs. For completeness, we state this relationship (modified from Ref. [9] to use the terminology of our paper) as a theorem and sketch the proof of the direction that is most important for us in this work.

Theorem 6.4.4 (App. F of Ref. [9])**.** *A complete set of MUBs* $\mathcal{M}$ *exists in dimension* d *if and only if there exists a uniformly-weighted* $P(T^d)$ 2*-design* X *of size* $|X| = d^2$ *satisfying*

$$\forall \phi \neq \theta \in X : \quad \left| \sum_{j=1}^{d} \mathrm{e}^{\mathrm{i}(\phi_j - \theta_j)} \right|^2 \in \{0, d\} . \tag{6.25}$$

Proof sketch. The "only if" direction is proven in Ref. [9, Lem. F.2] (in Ref. [9], a projective toric design is referred to as a torus design).

We now sketch the proof of the "if" direction. Let X be a uniformly-weighted $P(T^d)$ 2-design. As discussed in Section 6.4.1, we can concatenate X with the simplex 2-design given in Proposition 6.4.1 to yield the $\mathbb{CP}^{d-1}$ 2-design $\mathcal{M}$ consisting of the elements

1. $|n\rangle$ for $n \in \{0, \ldots, d-1\}$, each weighted by $\frac{1}{d(d+1)}$; and

2. $|c, \phi\rangle$ for $\phi \in X$, each weighted by $\frac{d}{d+1} \times \frac{1}{|X|} = \frac{1}{d(d+1)}$,

where recall that $c = (1/d, \ldots, 1/d) \in \Delta^{d-1}$ is the centroid and $|p, \phi\rangle = \sum_{n=1}^{d} \sqrt{p_n} \mathrm{e}^{\mathrm{i}\phi_n}$. Hence, $\mathcal{M}$ is a uniformly-weighted 2-design of size $\mathcal{M} = d(d+1)$. To show that $\mathcal{M}$ is a complete set of MUBs, the only thing left to show is condition (3) of Definition 6.4.3. The

only nontrivial overlaps to consider are $\langle c, \theta | c, \phi \rangle$ for $\phi \neq \theta \in X$, giving

$$|\langle c, \theta | c, \phi \rangle|^2 = \left| \frac{1}{d} \sum_{j=1}^{d} \mathrm{e}^{\mathrm{i}(\phi_j - \theta_j)} \right|^2 \in \{0, 1/d\}\,, \tag{6.26}$$

which comes by assumption of Eq. (6.25). □

Using Theorem 6.4.4 and the fact that (as we saw in the proof) any uniformly-weighted $P(T^d)$ 2-design yields a uniformly-weighted quantum state 2-design of size $d(d+1)$ via concatenation with the simplex design given in Proposition 6.4.1, we see that we can now rephrase part of Zhu's conjecture as follows.

> **Rephrasing of (part of) Zhu's conjecture** [95, Conj. 1]. Any uniformly-weighted $P(T^d)$ 2-design X of size $|X| = d^2$ must satisfy Eq. (6.25).

We note that we are rephrasing Zhu's conjecture as it pertains to uniformly-weighted quantum state 2-designs of size $d(d+1)$, as this is the case of interest for complete sets of MUBs and for the remainder of this manuscript. We do not make any statements regarding the existence of uniformly-weighted quantum state 2-designs of size $\lneq d(d+1)$.

In the language of our paper, Zhu's conjecture can be seen as conjecturing that a complete set of MUBs exists in dimension d if and only if there exists a uniformly-weighted $P(T^d)$ 2-design X of size $|X| = d^2$. Given the Sidon set construction of $P(T^d)$ 2-designs from Section 6.3, we see that we can disprove Zhu's conjecture by finding a Sidon set of size d mod d^2 that does not satisfy Eq. (6.25). Via brute force numerical searches for Sidon sets of size d mod d^2 for small d, we can find many counterexamples. Indeed, there are 288 such Sidon sets when $d = 6$ (see our code [278]), hence yielding 288 counterexamples to Zhu's conjecture in dimension 6. The simplest counterexample however occurs in dimension $d = 3$, as we show in the following example.

Example 6.4.5 (Counterexample to Zhu's conjecture in dimension 3)**.** Utilizing $\{0, 1, 3\}$, which is a Sidon set of size 3 mod 9, we arrive at the following uniformly-weighted quantum

state 2-design of size exactly $d(d+1) = 12$ that is *not* a complete set of MUBs:

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \bigcup \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ e^{\frac{2\pi i k}{9}} \\ e^{\frac{6\pi i k}{9}} \end{pmatrix} \;\middle|\; k \in \{0, 1, \ldots, 8\} \right\}. \tag{6.27}$$

◇

Subsection 6.4.3: MUBs and group designs

In this subsection, we consider general group projective toric 2-designs and when they can yield complete sets of MUBs. We find that while such designs can yield complete sets of MUBs is prime-power dimensions, they cannot in dimension in 6. This illustrates a fundamental difference between the structure of MUBs in prime-power versus non-prime-power dimensions.

Consider the following general parameterization of a subgroup of $P(T^d)$. Suppose that the subgroup is isomorphic to $\mathbb{Z}_{\alpha_1} \times \cdots \times \mathbb{Z}_{\alpha_k}$. We can generate each factor $\mathbb{Z}_{\alpha_j}$ by $z^{(\alpha_j)} \in \mathbb{Z}_{\alpha_j}^d$, where, since we are considering the projective torus, we fix $z_1^{(\alpha_j)} = 0$. We call the subgroup $X\left(\alpha_1, \ldots, \alpha_k; z^{(\alpha_1)}, \ldots, z^{(\alpha_k)}\right)$, and we have that

$$X\left(\alpha_1, \ldots, \alpha_k; z^{(\alpha_1)}, \ldots, z^{(\alpha_k)}\right) = \left\{ \sum_{j=1}^{k} \frac{2\pi n_j}{\alpha_j} z^{(\alpha_j)} \;\middle|\; n_1 \in \mathbb{Z}_{\alpha_1}, \ldots, n_k \in \mathbb{Z}_{\alpha_k} \right\}. \tag{6.28}$$

For example, when d is prime, the $P(T^d)$ 2-design in the standard MUB construction [39] (Theorem 6.2.7) is

$$X\left(d, d; (0, 1, 2, 3, \ldots, d-1), \left(0, 1, 4, 9, \ldots, (d-1)^2\right)\right). \tag{6.29}$$

In the case of prime-power dimensions, the construction is a generalization of Theorem 6.2.7

that uses the field theoretic trace [39]. Since the trace is linear, one can easily verify the design to be a group. Thus, for all prime-power d, there are group 2-designs of size d^2 satisfying Eq. (6.25).

However, via a numerical bruteforce search (see our code [278]), we find that there are no $P(T^6)$ group 2-designs of size $6^2 = 36$ satisfying Eq. (6.25). The search is done by recognizing that, given the classification of finite abelian groups, there are only four α to consider:

$$\alpha = (\alpha_1, \alpha_2) = (4, 9), \tag{6.30a}$$

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) = (3, 3, 4), \tag{6.30b}$$

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) = (2, 2, 9), \tag{6.30c}$$

$$\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (2, 2, 3, 3). \tag{6.30d}$$

We can then explicitly check every generator $z^{(\alpha_j)}$ for each of these group structures. We have therefore proven (via exhaustive numerical search) the following proposition, which comes as a corollary of the exhaustive search and Theorem 6.4.4.

Proposition 6.4.6. *If a complete set of MUBs exists in dimension 6, then there is a uniformly-weighted $P(T^6)$ 2-design X of size $|X| = 36$ such that X is not a subgroup of $P(T^6)$.*

Indeed, we believe this proposition highlights why finding complete sets of MUBs in dimension 6 is difficult: the projective toric designs that form the set must be something more complicated than a group. We state the generalization of Proposition 6.4.6 to all non-prime-power dimensions as a conjecture.

Conjecture 6.4.7. *Let d be a non-prime-power. If a complete set of MUBs exists in dimension d, then there is a uniformly-weighted $P(T^d)$ 2-design X of size $|X| = d^2$ such that X is not a subgroup of $P(T^d)$.*

Proposition 6.4.6 and Conjecture 6.4.7 highlight a fundamental distinction between complete sets of MUBs in prime-power dimensions versus those in non-prime-power dimensions. To the best of our knowledge, all currently known complete sets of MUBs in prime-power dimensions come from projective toric designs that are groups. On the contrary, we showed that this is not possible in $d = 6$, and conjecture it more generally.

Given this result, we make the following conjecture, which can be viewed as a modification of Zhu's conjecture that evades our earlier counterexamples.

Conjecture 6.4.8. *A complete set of MUBs exists in dimension d that is not a prime power if and only if there is a uniformly-weighted $P(T^d)$ 2-design X of size $|X| = d^2$ such that X is not a subgroup of $P(T^d)$.*

Given Conjecture 6.4.8, we naturally conjecture that there do not exist non-group, uniformly-weighted $P(T^6)$ 2-designs of size 36. More generally, we conjecture that, in *any* dimension, a complete set of MUBs must come from a group projective toric 2-design. We phrase this conjecture in terms of Eq. (6.25).

Conjecture 6.4.9. *Let d be any dimension and let X be a $P(T^d)$ 2-design of size $|X| = d^2$. If X is not a subgroup of $P(T^d)$, then X does not satisfy Eq.* (6.25).

If this conjecture is true, then Proposition 6.4.6 would prove the nonexistence of complete sets of MUBs in dimension 6.

Section 6.5: Conclusion and open questions

In this work, we have developed the theory of projective toric designs and their relation to various other objects in and areas of mathematics and physics. There is still much unknown and we believe there are still many exciting connections to be made. We now discuss various future research directions relating to projective toric designs.

Complete sets of mutually unbiased bases In this work, we proved (*cf.* Proposition 6.4.6) that the phases in a complete set of MUBs in dimension 6 must form a uniformly-weighted $P(T^6)$ 2-design of size 36 that is *not* a subgroup of $P(T^6)$. Using projective toric designs that *are* subgroups, we constructed families of quantum state 2-designs of size exactly $d(d+1)$ that are not complete sets of MUBs, thereby disproving Zhu's conjecture [95, Conj. 1]. Given this result, we proposed a modified verison of Zhu's conjecture (*cf.* Conjecture 6.4.8) regarding the relationship between complete sets of MUBs and projective toric designs. An obvious interesting open problem is to prove this conjecture. Another interesting direction is to try to prove that non-group $P(T^6)$ 2-designs of size 36 do not exist. If our conjecture is true, then this result would prove the long-outstanding problem regarding the existence of complete sets of MUBs in dimension 6. Finally, projective toric designs may have a close connection to Hadamard matrices, since the latter are also related to MUBs [72]. We leave this interesting question to future work.

Minimal projective toric designs In this work, we showed that if X is a $P(T^n)$ 2-design, then $|X| \geq n(n-1)+1$. Furthermore, using Sidon sets, we showed that the bound can be saturated when $n-1$ is a prime power. However, we also showed that the bound cannot always be satisfied using the Sidon set construction; for example, when $n=7$, the Sidon set construction does not yield a minimal projective toric 2-design. We thus have the following open question: do projective toric 2-designs saturating the bound exist for all n?

We showed that if the t-design is a cyclic group, then the constructions are in one-to-one correspondence with B_t mod $|X|$ sets. In the case of e.g. $n=7$ and $t=2$, $n(n-1)+1=43$ is prime so that the only group design could be a cyclic group. Therefore, if one can prove that a minimal design must be a group, then one would prove that the $t=2$ bound cannot be saturated for all n. Must the minimal design be a group?

We further proved that if X is a $P(T^n)$ t-design, then $|X| \geq G_{n-1}(\lfloor t/2 \rfloor)$. We conjectured that the bound is tight when t is even. Can this conjecture be proven? Can the bound be

tightened for odd t? Can one construct saturating designs? As we saw in Proposition 6.2.13, the lower bound on the size of projective toric 2-designs matches the lower bound on the size of dense modular Sidon sets. We believe that the analogous statement holds for all t. Using the connection between difference sets and projective toric designs, we related dense B_t mod m sets to the root lattice A_{n-1} and proved a bound relating the size n of the set and the value of m. This connection seems to be a fruitful area to continue exploring.

Connection to affine/projective planes A *finite projective plane* is a tuple (P, L) of a finite set of points P and lines $L \subseteq 2^P$ (where 2^P means the power set of set P, i.e. the set of all subsets of P) such that:

1. Any two points are elements of a unique common line
2. Any two lines intersect at a unique point
3. There exist four points in P such that no line contains more than two of them.

Affine planes are defined similarly. A tuple (P, L) can only be a finite projective plane if there exists some $d \in \mathbb{N}$ such that $|P| = |L| = d^2 + d + 1$. However, finite projective planes have only been constructed for d a prime power, and are known to *not* exist if d is both not the sum of two squares and $d \equiv 1$ or 2 mod 4. These numeric similarities, along with deep connections between combinatorial designs and finite geometry, hint at a deeper connection between projective toric designs and finite projective planes. In addition, projective planes appear in the construction of Sidon sets, and are conjectured to correspond to dense ones [280].

Further, a complete set of MUBs yields a finite projective plane, while a SIC-POVM in prime power dimensions yields a finite affine plane [90, 238, 281]. As mentioned above, MUBs are closely related to projective toric designs, while SIC-POVMs are minimal complex-projective designs. All of this circumstantial evidence begs the question: are there interesting

direct connections one can make between projective toric designs and finite planes, either projective or affine?

Connection to other designs Recall that complex projective designs can be constructed by concatenating simplex and projective toric designs. Similarly, rigged continuous variable t-designs can be constructed in an analogous way by using $P(T^\infty)$ designs. One can ask: how much can this result be generalized? Can we use similar constructions for toric varieties and flag varieties? Indeed $\mathbb{CP}^n$ is a toric variety with moment map to the associated polytope being the simplex Δ^n. The moment map allows us to project $\mathbb{CP}^n$ designs to Δ^n designs. Projective toric designs allow us to pullback along the moment map and build $\mathbb{CP}^n$ designs from Δ^n designs. How much more general can this result be made?

New families of quantum state designs Using the families of $P(T^n)$ t-designs constructed in Section 6.3.1, can we generate new interesting families of quantum state t-designs? To do this, we need to find families of simplex t-designs. In the $t = 2$ case, we used a particularly nice simplex 2-design that allowed us to construct almost-minimal quantum state 2-designs from minimal projective toric 2-designs. Can we find similarly nice simplex t-designs for $t > 2$? Our construction in Section 6.3.1 of $P(T^d)$ t-designs of size (asymptotically in d) $\approx d^t$ yields quantum state t-designs of size $\approx |D|\, d^t$, where D is an Δ^{d-1} t-design. It is an interesting question to study simplex t-designs to arrive at potentially new explicit constructions of quantum state t-designs.

Approximate designs One can consider approximate projective toric t-design, which are points on the projective torus that integrate monomials of degree $\leq t$ up to an error of ε. How does the size of the minimal approximate t-design depend on t and ε? If one takes an ε_1-approximate simplex t-design and ε_2-approximate projective toric design and concatenates them, what is the ε with which we get an ε-approximate complex-projective t-design? In

Appendix D.3, we take the first steps to study such approximate designs. In particular, we define ε-approximate $P(T^n)$ t-designs, and we provide an upper bound on the minimum number of points drawn uniformly randomly from $P(T^n)$ needed to form an ε-approximate $P(T^n)$ t-design with probability $1-\delta$. We show that the resulting bound depends on the crystal ball sequences of the root lattices A_{n-1} [272, 273] given in Eq. (6.13). Using the discussion around Eq. (6.20), these approximate $P(T^n)$ designs can be lifted to approximate T^n designs, which we recall are then approximate designs on the diagonal subgroup $T(\mathrm{U}(n))$ of $\mathrm{U}(n)$, which are of inherent interest in quantum information theory [85]. Ref. [85] also constructs approximate designs on $T(\mathrm{U}(n))$, which can of course be projected to designs on $P(T^n)$.

Part III

Superresolution in quantum sensing

Chapter 7: Superresolution in quantum sensing

Section 7.1: Introduction

Estimating the separation between peaks in a spectrum is a problem that occurs in many applications, such as nuclear magnetic resonance spectroscopy for the identification of molecules and chemical structure [282] and magnetometry for measurement and detection of weak magnetic fields [283]. The task of resolving two closely spaced peaks is similar to Rayleigh's curse in optics, where two light sources cannot be distinctly resolved if they are separated by less than the diffraction limit [97]. Ref. [98] designed a quantum superresolution protocol that can in principle resolve two arbitrarily close point light sources, thereby beating Rayleigh's curse.

In this work, we consider a similar problem. Given a signal with a frequency spectrum that contains closely spaced peaks, resolving the different peaks typically requires observing the signal for a time that scales inversely with the peak separation. We design quantum superresolution protocols, and indeed a general method of finding and optimizing such protocols, that can resolve arbitrarily closely spaced frequency peaks in finite time. Our work is related to that of Ref. [284], where they provide sufficient conditions on a quantum state to exhibit superresolution. In this work, we provide conditions in terms of the protocol and controls, thus giving conditions and procedures for designing superresolution protocols. We summarize and describe our setting and results in Section 7.1.1.

Subsection 7.1.1: Model & problem description

We begin with a signal $\gamma(t)$ which is itself a stochastic Gaussian process. We assume that the signal has zero mean $\langle\gamma(t)\rangle = 0$ and that the signal is stationary $\langle\gamma(t)\gamma(t')\rangle = \langle\gamma(t+\tau)\gamma(t'+\tau)\rangle$, where we denote averaging over the stochasticity by $\langle\cdot\rangle$. It follows that $\langle\gamma(t)\gamma(t+\tau)\rangle$ is a function $C(\tau)$ independent of t so that the spectral density

$$S(\omega) \coloneqq \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\omega\tau} C(\tau)\,\mathrm{d}\tau \tag{7.1}$$

is a function of only the frequency ω. Notice that under our given assumptions on γ, $\gamma(t)$ is completely determined by $S(\omega)$.

We consider the case when $S(\omega > 0) \sim \delta(\omega - \omega_1) + \delta(\omega - \omega_2)$, where $\delta(x)$ is the Dirac delta function. That is, the spectral density of the signal γ is infinitely peaked at two distinct frequencies ω_1 and ω_2. One such example is the signal

$$\begin{aligned}\gamma(t) = g\big(&A_1\cos(\omega_1 t) + A_2\cos(\omega_2 t)\\ &+ B_1\sin(\omega_1 t) + B_2\sin(\omega_2 t)\big),\end{aligned} \tag{7.2}$$

where A_1, A_2, B_1 and B_2 are i.i.d. normal random variables with mean 0 and variance 1. Such a signal is considered in Ref. [284]. Denote the centroid and frequency difference by $\omega_c \coloneqq (\omega_1 + \omega_2)/2$ and $\Delta\omega \coloneqq \omega_2 - \omega_1$ respectively. Our task is to determine the frequency difference $\Delta\omega$ up to relative error δ — that is, absolute error $\delta\Delta\omega$. Specifically, we are interested in the limiting case when $\Delta\omega \to 0$. We assume that the centroid ω_c is known; later we will generalize to when the centroid is only approximately known.

In the setting of metrology, this task falls naturally within the framework of parameter estimation. Specifically, given access to some form of measurements $\boldsymbol{\xi}$ of the signal $\gamma(t)$, there is some probability distribution $P(\boldsymbol{\xi}|\Delta\omega)$ over measurement outcomes $\boldsymbol{\xi}$ for a fixed value of the parameter of interest $\Delta\omega$. The fundamental limit on the amount of information

about $\Delta\omega$ that one can extract given access to $P(\boldsymbol{\xi}|\Delta\omega)$ is given by the Fisher information

$$\mathrm{FI}_{\Delta\omega}(P) \coloneqq \sum_{\boldsymbol{\xi}} \frac{1}{P(\boldsymbol{\xi}|\Delta\omega)} \left(\frac{\partial P(\boldsymbol{\xi}|\Delta\omega)}{\partial \Delta\omega} \right)^2. \tag{7.3}$$

The Cramér-Rao bound gives a fundamental bound in terms of the Fisher information on how accurately one can estimate a parameter: given N measurements from the probability distribution P, the variance of any unbiased estimator of $\Delta\omega$ is greater than $\frac{1}{N\,\mathrm{FI}_{\Delta\omega}(P)}$. Thus, the goal is as follows: given access to the signal $\gamma(t)$, design an experiment that generates a probability distribution P that maximizes $\mathrm{FI}_{\Delta\omega}(P)$.

In this setting, we now formally define superresolution in terms of the Fisher information. A protocol exhibits *superresolution* if the associated probability distribution $P(\boldsymbol{\xi}|\Delta\omega)$ satisfies

$$\lim_{\Delta\omega\to 0} \mathrm{FI}_{\Delta\omega}(P) > 0. \tag{7.4}$$

We further require that the protocol takes finite time even in the limit as $\Delta\omega \to 0$.

We now couple the signal to a qubit via the Hamiltonian $H(t) = \gamma(t)\sigma^z + c(t)\sigma^x$, where σ^z and σ^x are the Pauli-Z and Pauli-X operators and $c(t)$ is a control form. More generally, we will also consider the affect of a stochastic dephasing noise process $\lambda(t)$, thus resulting in the Hamiltonian

$$H(t) = (\gamma(t) + \lambda(t))\sigma^z + c(t)\sigma^x \tag{7.5}$$

After a total time $T = \kappa\tau$, with $\tau \coloneqq 2\pi/\omega_c$, we measure the qubit, resulting in a probability distribution P. Our task is therefore to choose a κ and control form $c(t)$ to maximize $\lim_{\Delta\omega\to 0} \mathrm{FI}_{\Delta\omega}(P)$.

Subsection 7.1.2: Outline & summary of results

To begin, we restrict our attention to instantaneous controls — these are controls that are a sum of Dirac delta functions, $c(t) = \frac{\pi}{2}\sum_{i=1}^{M}\delta(t-t_i)$, hence corresponding a sequence of instantaneous Pauli-X gates performed at times specified by t_i. The first main result of our work, described in Section 7.2, is a set of necessary and sufficient conditions on the controls $c(t)$ and the total time $\kappa\tau$ to achieve superresolution. The conditions are expressed in terms of the filter function, which itself depends *only* on $c(t)$ and κ. This characterization elucidates a key distinction between our work and that of Ref. [284]; namely, we provide sufficient conditions for superresolution in terms of conditions on the *controls* (*i.e.* conditions on the *filter function*), whereas Ref. [284] provides conditions on the quantum state that is being sampled from.

Next, in Section 7.3, we derive two families of instantaneous control superresolution protocols; namely, free controls with $c(t) = 0$ and CPMG controls with $c(t)$ consisting of CPMG sequences. We carefully analyze their performance and error bounds, both in the noiseless ($\lambda(t) = 0$) and noisy ($\lambda(t) \neq 0$) case. In the noisy case, we will specifically analyze Lorentzian noise, which is a common noise model for quantum sensors such as NV defects in diamonds. We find that the overlap of the filter function with the noise spectrum has a large effect on the performance of the protocols, and therefore the CPMG superresolution protocol can substantially outperform the free superresolution protocol in the presence of Lorentzian noise.

In Section 7.4, we generalize to continuous control protocols. By analyzing its filter function, we find explicit controls $c(t)$ that outperform both the free and CPMG superresolution protocols in certain regimes. Furthermore, a major contribution of our work is a well-defined numerical procedure to optimize controls to achieve the best results subject to various conditions of interest, such as noise strength and type, limitations on the length and width of control pulses, etc. We demonstrate this numerical procedure in Section 7.4.1, where we

find numerically optimized controls that outperform all of the aforementioned protocols in certain regimes.

In Section 7.5, we show a straightforward method of taking any of superresolution protocol and modifying it to work with entangled qubits, resulting in the need for less resources. Finally, in Section 7.6, we compare our superresolution protocols to traditional methods in quantum noise spectroscopy. We also compare to information-theoretically optimal classical protocols by studying the classical Fisher information in various settings. These comparisons help to elucidate when superresolution protocols should be used, and when other methods are more favorable.

Section 7.2: Superresolution conditions

Throughout this work, we will fix the initial qubit state to be $|+\rangle$ and fix the measurement basis to be $\{|+\rangle, |-\rangle\}$, where $|+\rangle$ ($|-\rangle$ resp.) is the $+1$ (-1 resp.) eigenvector of σ^x. Indeed, we show in Appendix E.2 that this can be done without loss of generality. We will begin by considering instantaneous controls, so that $c(t) = \frac{\pi}{2}\sum_{i=1}^{M}\delta(t-t_i)$ for some choices of t_i and M. We return to the more general case of continuous controls in Section 7.4. Such an instantaneous control sequence can be described by its *switching function* $f(t)$, where $f(t) = -1$ if $t_{i-1} < t < t_i$ and $M - i$ is odd, and $f(t) = 1$ otherwise.

We set the noise $\lambda(t) = 0$, and return to the noisy case in Section 7.2.2. After evolving the initial $|+\rangle$ state for a time $\kappa\tau$, we denote by P the probability of measuring $|+\rangle$. We are then interested in the Fisher information of the probability distribution $\langle P\rangle$, where $\langle\cdot\rangle$ denote the average over the stochastic signal. Given our signal in Eq. (7.2), $\langle\cdot\rangle$ denotes averaging over A_1, A_2, B_1, and B_2.

$\langle P\rangle$ can be expressed in terms of the signal's spectrum $S(\omega)$ and a *filter function* $F(\omega, t)$ that depends only on the switching function $f(t)$ [285–288]. Indeed, we review filter functions

in Appendix E.1, and show that

$$\begin{aligned} \langle P\rangle &= \frac{1}{2} + \frac{1}{2}\mathrm{e}^{-\chi}, \qquad \text{where} \\ \chi &= \frac{1}{\pi}\int_{-\infty}^{\infty} S(\omega)F(\omega,\kappa\tau)\,\mathrm{d}\omega, \\ F(\omega,t) &= \left|\int_0^t f(t')\mathrm{e}^{\mathrm{i}\omega t'}\,\mathrm{d}t'\right|^2. \end{aligned} \tag{7.6}$$

The key point is that the dependence on the controls and the signal is separated in terms of the filter function and the spectrum.

The signal in Eq. (7.2) has spectrum

$$\begin{aligned} S(\omega) = \pi g^2\big[&\delta(\omega_1-\omega)+\delta(\omega_1+\omega) \\ &+\delta(\omega_2-\omega)+\delta(\omega_2+\omega)\big], \end{aligned} \tag{7.7}$$

making χ simply a sum of $F(\pm\omega_1,\kappa\tau)$ and $F(\pm\omega_2,\kappa\tau)$. Taylor expanding around small $\Delta\omega$ and using Eq. (7.3), we find that

$$\mathrm{FI}_{\Delta\omega}(\langle P\rangle) = \begin{cases} g^2F''(\omega_c,\kappa\tau)+\mathcal{O}(\Delta\omega^2) & \text{if } F(\omega_c,\kappa\tau)=0 \\ \mathcal{O}(\Delta\omega^2) & \text{otherwise.} \end{cases} \tag{7.8}$$

where $F''(\omega_c,\kappa\tau)$ denotes the second derivative of the filter function with respect to ω evaluated at ω_c and $\kappa\tau$. Thus, a protocol, which is defined by filter function and a total time $\kappa\tau$, exhibits superresolution if $F(\omega_c,\kappa\tau)=0$ and $F''(\omega_c,\kappa\tau)>0$.

We note that the superresolution conditions are specific to the signal spectrum, however our procedure can in principle be adapted to other signals as well. The exact conditions on the filter function to ensure superresolution for different signal spectrums may differ, but such conditions can nonetheless be found.

Subsection 7.2.1: Basic analysis

We now describe a basic analysis of the superresolution protocol. Because everything is symmetric under $\Delta\omega \to -\Delta\omega$, in the limit of small $\Delta\omega$, the probability distribution takes the form $\langle P\rangle = a - b\Delta\omega^2 + \mathcal{O}(\Delta\omega^4)$. From Eqs. (7.6) and (7.7), $a = 1 - g^2 F(\omega_c, \kappa\tau)$ and $b = \frac{g^2}{4} F''(\omega_c, \kappa\tau)$. With this and Eq. (7.3), the Fisher information takes the values

$$\mathrm{FI}_{\Delta\omega}(\langle P\rangle) = \begin{cases} 4b + \mathcal{O}(\Delta\omega^2) & \text{if } a \in \{0, 1\} \\ \frac{4b^2}{a(1-a)}\Delta\omega^2 + \mathcal{O}(\Delta\omega^4) & \text{if } 0 < a < 1. \end{cases} \tag{7.9}$$

By sampling from the qubit at the end of the protocol N times, we output an estimate $\tilde{P} = N_+/N$ of $\langle P\rangle$, where N_+ is the number of times the qubit was measured to be in the $|+\rangle$ state. This results in an estimate $\widetilde{\Delta\omega} = \sqrt{(\tilde{P} - a)/b}$ of $\Delta\omega$. Therefore, an estimation error of $2b\Delta\omega\varepsilon$ in $\langle P\rangle$ corresponds to an estimation error of $\sim \varepsilon$ in $\Delta\omega$. By Chebyshev's inequality with the variance $\langle P\rangle(1 - \langle P\rangle)/N$ of the binomial distribution, we find that

$$\begin{aligned} &\Pr\left[\left|\tilde{P} - \langle P\rangle\right| \geq 2b\Delta\omega\varepsilon\right] \\ &\qquad \leq \frac{a(1-a) + |b(1-2a)|\,\Delta\omega^2}{4Nb^2\Delta\omega^2\varepsilon^2} + \mathcal{O}\big(\Delta\omega^4\big)\,. \end{aligned} \tag{7.10}$$

Hence, in order to achieve a fixed absolute error ε in our estimate of $\Delta\omega$, we need

$$N \sim \begin{cases} \frac{1}{4b\varepsilon^2} & \text{if } a \in \{0, 1\} \\ \frac{a(1-a)}{4b^2\Delta\omega^2\varepsilon^2} & \text{if } 0 < a < 1 \end{cases} \tag{7.11}$$

measurements. Notice that in both cases, from Eq. (7.9), the number of samples required is $N \sim 1/(\varepsilon^2\,\mathrm{FI}_{\Delta\omega}(\langle P\rangle))$, which is simply a manifestation of the Cramér-Rao bound.

It is more natural to work in terms of relative error $\delta = \varepsilon/\Delta\omega$ since we are considering the limit as $\Delta\omega \to 0$. Hence, in the superresolution case $a \in \{0, 1\}$, to achieve a fixed relative

accuracy δ, it is sufficient to sample

$$N \sim \frac{1}{4b\delta^2\Delta\omega^2} = \frac{1}{g^2F''(\omega_c,\kappa\tau)\delta^2\Delta\omega^2} \tag{7.12}$$

times from $\langle P\rangle$ when $\Delta\omega \to 0$, whereas in the non-superresolution case, one needs to sample $N \sim 1/\Delta\omega^4$ times to achieve the same relative accuracy.

Thus far in our analysis, we have assume that we sample from the probability distribution $\langle P\rangle$ N times in order to get the estimate $\tilde{P}$. However, in reality, we will be sampling from P for many different realizations of the parameters A_1, A_2, B_1, and B_2. Fortunately, this is not an issue due to the concentration of P around $\langle P\rangle$. In particular, we show in Appendix E.1.2 that $\sigma^2 = \langle P^2\rangle - \langle P\rangle^2 = 2\langle P\rangle^2(1-\langle P\rangle)^2$. For superresolution protocols where $a \in \{0,1\}$, this variance becomes $\sigma^2 = 2b^2\Delta\omega^4 + \mathcal{O}(\Delta\omega^6)$. Then Bennett's inequality implies that our estimate $\widetilde{\langle P\rangle}$ of $\langle P\rangle$ after N measurements will satisfy

$$\begin{gathered}\Pr\Big[\Big|\widetilde{\langle P\rangle} - \langle P\rangle\Big| > \xi\Big] \le \exp\big[-N\sigma^2 h(\xi/\sigma^2)\big] \\ h(x) = (1+x)\log(1+x) - x.\end{gathered} \tag{7.13}$$

Recall that we require $\xi \le 2b\delta\Delta\omega^2$ in order to achieve a relative error of δ in our estimate of $\Delta\omega$. It follows that one needs to sample $N \sim \big(b\delta\Delta\omega^2\log\big(\frac{\delta}{b\Delta\omega^2}\big)\big)^{-1}$ times in order to effectively sample from $\langle P\rangle$. From Eq. (7.12), our superresolution protocols require $N \sim \frac{1}{b\delta^2\Delta\omega^2}$. Hence, the superresolution protocols already fulfill the necessary condition to combat this, and superresolution is therefore practically unaffected by it.

Subsection 7.2.2: Nonzero noise

We now consider the case when the noise $\lambda(t)$ is nonzero. Recall from Eq. (7.1) that $S(\omega)$ is the spectrum of the signal $\gamma(t)$, and we similarly define $S_\lambda(\omega)$ to be the spectrum of the noise $\lambda(t)$. In this case, Eq. (7.6) gets modified via $\chi \to \chi + \chi_\lambda$, where $\chi_\lambda = \frac{1}{\pi}\int S_\lambda(\omega)F(\omega,\kappa\tau)\,\mathrm{d}\omega$.

Given a superresolution protocol, corresponding to $a = 1$ in the language of the previous section, the effect is therefore $a \to 1 - \frac{1}{2}\chi_\lambda$. It follows that superresolution is not robust to noise, in the sense that $\lim_{\Delta\omega\to 0} \mathrm{FI} = 0$ due to $a < 1$.

Nonetheless, for finite but small $\Delta\omega$, superresolution protocols can still yield good performance. In particular, we can design controls such that χ_λ is minimized. In general, we are therefore interested in controls such that $F(\omega_c, \kappa\tau) = 0$, $F''(\omega_c, \kappa\tau)$ is maximized, and χ_λ is minimized. These criteria are well-suited for numerical optimization, which we explore in Section 7.4.1.

Section 7.3: Instantaneous control superresolution protocols

While *exactly instantaneous* controls are not possible, approximately instantaneous controls are accessible [286]. Throughout this work, we will consider two instantaneous control superresolution protocols: free evolution and CPMG evolution. Free evolution refers to *no controls*, so that the corresponding switching function is $f(t) = 1$ for all t. CPMG evolution refers to a sequence of $\kappa/2$ CPMG sequences. Specifically, letting κ be an even integer, we apply instantaneous σ^x gates at times $\tau/2, 3\tau/2, \ldots, (2\kappa - 1)\tau/2$. From Eq. (7.6), it immediately follows that in both cases, the superresolution conditions are satisfied with

$$\mathrm{FI}_{\Delta\omega\to 0} = \begin{cases} \frac{8\pi^2 g^2 \kappa^2}{\omega_c^4} & \text{free (if } \kappa \in \mathbb{N}) \\ \frac{32\pi^2 g^2 \kappa^2}{\omega_c^4} & \text{CPMG (if } \kappa \in 2\mathbb{N}). \end{cases} \tag{7.14}$$

We can easily compare, for example, (FE-SR) the free evolution protocol with $\kappa = 2$, (FE-Non-SR) the free evolution protocol with $\kappa = 5/2$, and (CPMG-SR) the CPMG protocol with $\kappa = 2$. Using Eq. (7.12), we see that in order the achieve a fixed relative estimation error, CPMG-SR requires 4 times *fewer* measurements than FE-SR. Additionally, FE-Non-SR requires $\sim \Delta\omega^{-2}$ times *more* measurements than the two superresolution protocols FE-SR

and CPMG-SR, despite the fact that FE-Non-SR uses a larger evolution time.

Thus, designing controls that satisfy the superresolution criteria yields a huge improvement over non-superresolution protocols. Furthermore, we can see differences between different superresolution protocols stemming from how their filter function derivatives at the centroid compare. We can see even larger differences emerge between superresolution protocols by introducing noise. Recall that in order the minimize the effect of a noise spectrum $S_\lambda(\omega)$, we want to design controls such that the corresponding filter function has minimal overlap with S_λ. In Fig. 7.1, we plot the filter functions for FE-SR and CPMG-SR with $\kappa = 2$ alongside a Lorentzian noise spectrum

$$S_\lambda(\omega) = \frac{g_\lambda^2}{\pi} \frac{2W}{4\omega^2 + W^2}, \tag{7.15}$$

where W is the full width at half maximum (FWHM) and g_λ is the strength of the noise. As can be seen in the figure, the CPMG-SR filter function has as substantially smaller overlap with the noise than does FE-SR, and we therefore suspect that the benefit of CPMG-SR over FE-SR will become even more pronounced in the presence of Lorentzian noise (and more generally any low frequency noise).

In fact, in Appendix E.3 we perform detailed analysis of CPMG-SR and FE-SR with the addition of Lorentzian noise, and we find that, given the same number of measurements, CPMG-SR will achieve a relative error of $\sim \sqrt{\frac{W}{\kappa^2 \omega_c}}$ times the relative error that FE-SR achieves. For example, when $\kappa = 2$, CPMG-SR achieves a relative error of less than $\sqrt{\frac{W}{15\omega_c}}$ times the relative error that FE-SR achieves. Note that as the FWHM of the noise grows, the overlap of the CPMG filter function increases more and more so that at very large FHWM, the advantage is lost. However, for low frequency noise, we see a large benefit stemming from the filter function having low weight on low frequencies.

The effects discussed above can be seen clearly in numerical simulations shown in Fig. 7.2. The simulations are performed by letting all three of the protocols have access to $N =$

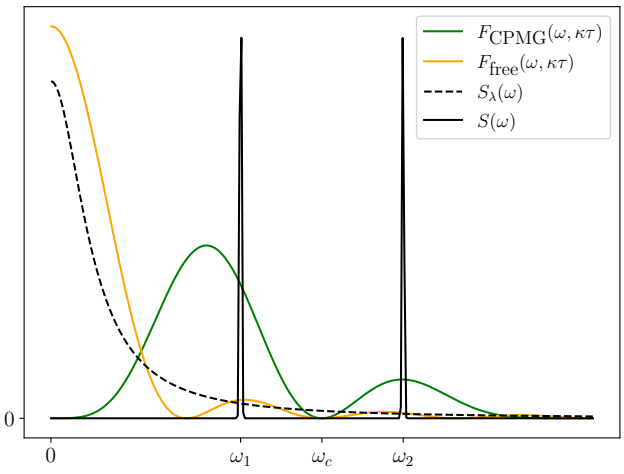

Figure 7.1: Illustrations of the filter functions corresponding to the free evolution and CPMG superresolution protocols, an example Lorentzian noise spectrum $S_\lambda(\omega)$, and the (approximate) signal spectrum $S(\omega)$. In the presence of noise, we desire to minimize the overlap integral between the noise spectrum and the filter function while satisfying the superresolution criteria $F(\omega_c, \kappa\tau) = 0$ and $F''(\omega_c, \kappa\tau) > 0$.

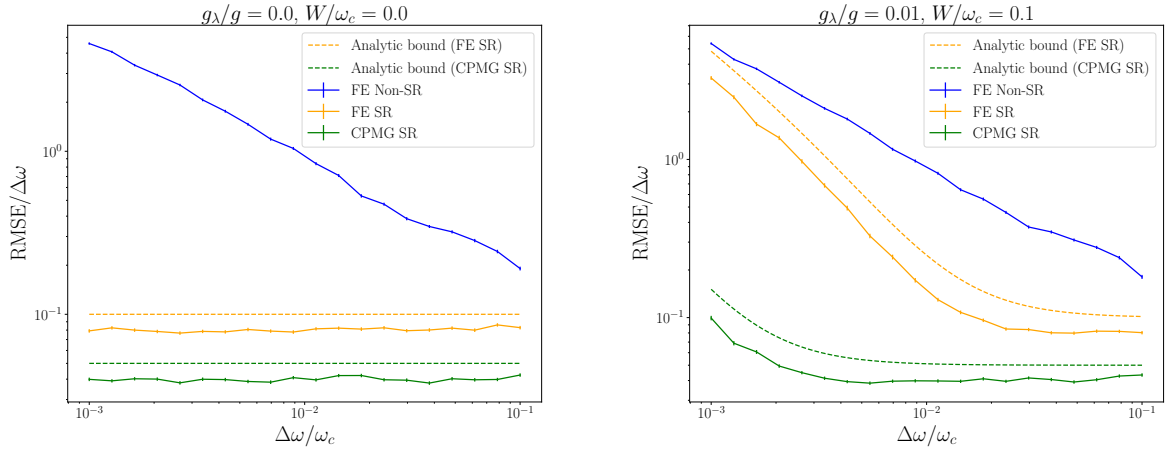

Figure 7.2: Given the parameters $g = 0.1$ MHz, $\omega_c = 1$ MHz, and desired relative error $\delta = 0.1$, we numerically simulate the free evolution protocol with $\kappa = 5/2$ which is *not* a superresolution protocol (FE Non-SR), the free evolution superresolution protocol with $\kappa = 2$ (FE SR), and the CPMG superresolution protocol with $\kappa = 2$ (CPMG SR). For each of the protocols, we utilize N_{free} samples as given in Eq. (7.11). **(Left)**: Noiseless case; the Lorentzian noise parameters are set to $g_\lambda = W = 0$. As $\Delta\omega \to 0$, the SR protocols achieve the desired relative accuracy while the non-SR protocol fails. **(Right)**: Noisy case; the Lorentzian noise parameters are set to $g_\lambda = 0.001$ MHz and FWHM $W = 0.1$ Hz. The CPMG superresolution protocol performs significantly better due to the fact that the corresponding filter function overlaps with the noise spectrum significantly less. In both figures, the analytic error bounds are from Appendix E.3.

$\frac{\omega_c^4}{8\pi^2 g^2 \kappa^2 \delta^2 \Delta\omega^2}$ measurements, with $\kappa = 2$. From Eq. (7.12), it follows that in the noiseless case, FE-SR should achieve a relative error of at most δ and CPMG-SR a relative error of at most $\delta/4$, while FE-Non-SR should achieve a relative error that grows as $\Delta\omega$ decreases. This is confirmed numerically in Fig. 7.2. Furthermore, Appendix E.3 shows that in the noisy case with small FWHM, we see FE-SR initially achieves a relative error of δ, but the relative error grows as $\Delta\omega$ decreases due to the noise. On the other hand, CPMG-SR remains at a relative error of $\sim \delta/4$ for much smaller values of $\Delta\omega$ (due to the $\sqrt{\frac{W}{15\omega_c}}$ factor derived above), until eventually it also must grow as $\Delta\omega \to 0$.

Finally, we mention one peculiarity in Fig. 7.2. Recall that each of the methods, including FE-Non-SR, has access to $N \sim \Delta\omega^{-2}$ measurements. Therefore, if we are sampling from the probability distribution defined by Eq. (7.6) for FE-Non-SR, the relative error should scale as $\Delta\omega^{-2}$ as $\Delta\omega \to 0$ in both the noisy and noiseless cases. However, the protocols are not sampling from $\langle P \rangle$, but rather P. We showed around Eq. (7.13) that this distinction is unimportant for superresolution protocols, but the same analysis with a non-superresolution protocol reveals that the distinction *is* important unless $N \sim \Delta\omega^{-4}$. This finite sampling error results in estimates $\widetilde{\langle P \rangle}$ of $\langle P \rangle$ that can sometimes yield estimates of $\Delta\omega^2$ that are negative. In the numerical simulations presented in Fig. 7.2, we combat this effect by simply taking the absolute value of the estimate. Nonetheless, this approach affects the expected scaling, and indeed the FE-Non-SR relative error does not increase exactly as $\Delta\omega^{-2}$ as $\Delta\omega \to 0$.

In summary, we have seen that choosing superresolution controls whose filter function has minimal overlap with the noise in a system can yield dramatic improvements. It is therefore natural to numerically search for controls that satisfy the superresolution criteria and minimize overlap with the noise. In the next section, we generalize our discussion from instantaneous controls to continuous controls and show how this optimization can yield substantial benefits even over CPMG-SR.

Section 7.4: Continuous control superresolution protocols

Thus far, we have used filter functions to study instantaneous control methods, yielding Eq. (7.6). We now allow the controls $c(t)$ to be arbitrary, and define the integrated controls

$$\theta_c(t) = \int_0^t c(t')\,\mathrm{d}t'. \tag{7.16}$$

In Appendix E.1, we review the filter function formalism for continuous controls by following the Magnus expansion [289]. The upshot is that with a modification $F^{(2)}(\omega, t)$ of the filter function $F(\omega, t)$, Eq. (7.6) holds up to order $\mathcal{O}(g^4)$, and therefore the superresolution criteria derived in Section 7.2 are still valid with respect to $F^{(2)}$ as long as $g^2/\Delta\omega^2$ is small.

More specifically, as $\Delta\omega \to 0$, the Magnus expansion [289] with our signal Eq. (7.2) gives

$$\begin{aligned}\langle P\rangle = 1 - g^2 &\left(2F^{(2)}(\omega_c, \kappa\tau) + \frac{2g^2}{3}F^{(4)}(\omega_c, \omega_c, \kappa\tau)\right) \\ &+ g^2\Delta\omega^2\left(\frac{1}{4}F^{(2)\prime\prime}(\omega_c, \kappa\tau) + \frac{g^2}{3}F^{(4)\prime\prime}(\omega_c, \omega_c, \kappa\tau)\right) + \mathcal{O}\left(\Delta\omega^4, g^6\right).\end{aligned} \tag{7.17}$$

Here, $F^{(2)}$ is a generalization of F in Eq. (7.6) given by

$$\begin{aligned}F^{(2)}(\omega, t) = &\left|\int_0^t \cos(2\theta_c(t'))\mathrm{e}^{\mathrm{i}\omega t'}\,\mathrm{d}t'\right|^2 \\ &+ \left|\int_0^t \sin(2\theta_c(t'))\mathrm{e}^{\mathrm{i}\omega t'}\,\mathrm{d}t'\right|^2,\end{aligned} \tag{7.18}$$

and $F^{(4)}$ is a higher order filter function given by Eq. (E.21) in Appendix E.1, and $F^{(4)\prime\prime}$ denotes the second derivative with respect to the first argument. When $c(t)$ is an instantaneous control, then $\sin(2\theta_c(t)) = 0$ and $\cos(2\theta_c(t)) = f(t)$, therefore reducing to Eq. (7.6) as expected.

If $F^{(2)}$ satisfies the superresolution criteria $F^{(2)}(\omega_c, \kappa\tau) = 0$ and $F^{(2)\prime\prime}(\omega_c, \kappa\tau) > 0$, then Eq. (7.17) implies a region of $\Delta\omega$ where the corresponding protocol effectively behaves as a

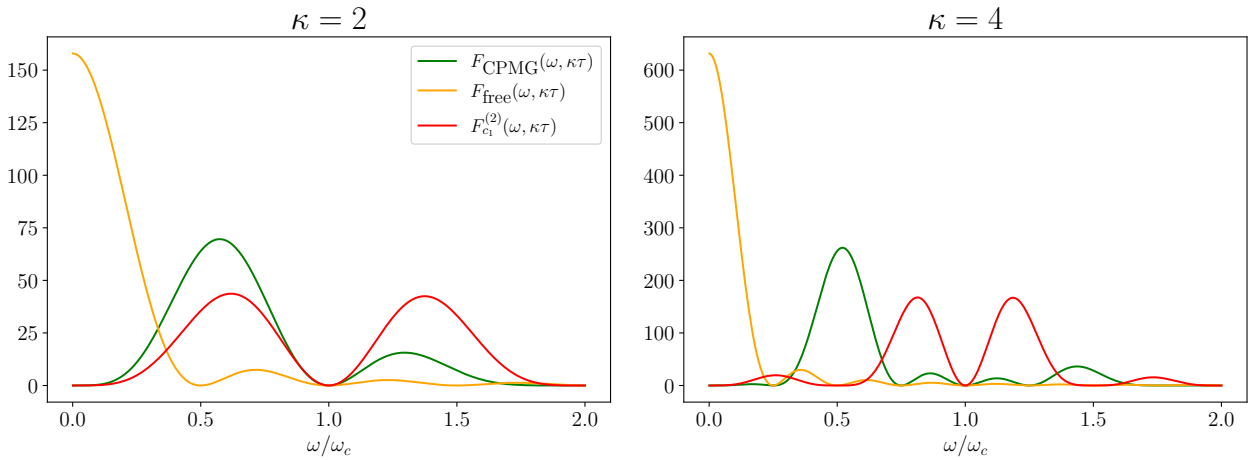

Figure 7.3: The filter functions for the $c_1(t)$ protocol (continuous controls) and for the CPMG and free superresolution protocols (instantaneous controls). Note that we are plotting the lowest order filter function. For the instantaneous control protocols, the lowest order filter function is the only one that matters; however, for continuous controls protocols, higher order filter functions become important. (Left) Filter functions for $\kappa = 2$. (Right) Filter functions for $\kappa = 4$.

superresolution protocol. In particular, the region is given by all values $\Delta\omega$ satisfying

$$\Delta\omega^2 \gtrsim \frac{8g^2 F^{(4)}(\omega_c, \omega_c, \kappa\tau)}{3F^{(2)\prime\prime}(\omega_c, \kappa\tau)}. \tag{7.19}$$

One major benefit of considering continuous controls is the ability to numerical optimize over a continuous space to find particularly good protocols. Indeed in Section 7.4.1, we specifically discuss numerically optimizing to find protocols that attempt to maximize this region by minimize $F^{(4)}(\omega_c, \omega_c, \kappa\tau)$.

Before numerically optimizing, though, we analytically study a specific continuous control to demonstrate that continuous controls can be beneficial as compared to the free and CPMG instantaneous controls. We define the controls

$$c_1(t) = -\frac{1}{2} + \frac{\pi}{2}\delta\left(t - \frac{\pi\kappa}{\omega_c}\right), \tag{7.20}$$

which represents a continuous constant drive with a single instantaneous σ^x gate halfway

through. When κ is an integer, the filter function $F_{c_1}^{(2)}$ for c_1 obeys the superresolution criteria. *If* the higher filter functions were also to vanish at the centroid, then c_1 would yield superresolution with the Fisher information

$$\lim_{\Delta\omega\to 0} \mathrm{FI}_{\Delta\omega}^{(c_1)} = \frac{\pi^4 g^2 \kappa^4}{\omega_c^4} \qquad (\text{if } \kappa \in \mathbb{N}). \tag{7.21}$$

Eq. (7.19) then implies that the c_1 protocol behaves as a superresolution protocol with this Fisher information in the region $\Delta\omega \gtrsim \sqrt{\frac{3g^2}{2\pi^2}} \sim g/3$. We will hence refer to the c_1 protocol as c_1-SR

We can immediately see a benefit of c_1 as compared to FE-SR and CPMG-SR by comparing their Fisher informations. Specifically, the Fisher information for FE-SR and CPMG-SR increase quadratically with the total time, whereas the Fisher information for c_1-SR increases quartically. Furthermore, comparing their filter functions shown in Fig. 7.3, c_1-SR will be effected less by low frequency noise, especially as κ increases, due to its lower weight on low frequencies.

Indeed, we see this effect in numerical simulations presented in Fig. 7.5. We can explicitly see c_1-SR perform better as g decreases, and the gap between the achieved relative error of CPMG-SR and c_1-SR increase as the Lorentzian noise strength increases when g is small.

Subsection 7.4.1: Numerical optimization of superresolution protocols

One major benefit of considering the superresolution criteria for continuous controls is that it easily allows for numerical optimization. In particular, in this section, we will begin with c_1-SR as an initial seed and optimize to find $c_1^{(\mathrm{opt})}$-SR.

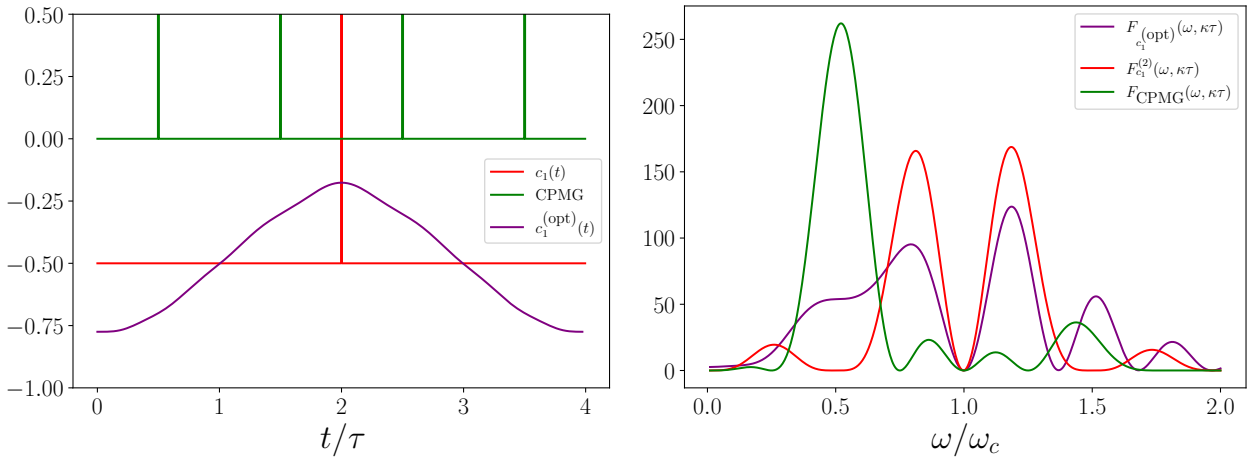

Figure 7.4: For $\kappa = 4$. (Left): c_1, $c_1^{(\mathrm{opt})}$, and CPMG control sequences. (Right): Corresponding filter functions. The green and red filter functions are the same as in the right of Fig. 7.3.

We minimize the following Lagrangian,

$$\begin{aligned}\mathcal{L} = -F^{(2)\prime\prime}(\omega_c,\kappa\tau) + \mu_1 \int S_\lambda(\omega) F^{(2)}(\omega,\kappa\tau)\,\mathrm{d}\omega + \mu_2 F^{(2)}(\omega_c,\kappa\tau) \\ + \mu_3 \left\|c(t)\right\|_2^2 + \mu_4 F^{(4)}(\omega_c,\omega_c,\kappa\tau) + \mu_5 \left\|c'(t)\right\|_2^2 .\end{aligned} \tag{7.22}$$

The first term represents maximizing the Fisher information (proxied by the second derivative of the filter function at the centroid). The μ_2 term enforces the superresolution condition that the filter function vanishes at the centroid – in the case of instantaneous controls, this *exactly* results in superresolution, whereas in the case of continuous controls, this results in superresolution up to order g^4. The μ_1 term means that we are attempting to minimize the overlap between the filter function and the Lorentzian noise, thus reducing the effect of the noise on our protocols. The μ_3 term encourages the resulting controls to have small amplitude. The μ_4 term tries to minimize the effect of the higher order filter functions. Finally, the μ_5 term enforces that the resulting controls are smooth.

We emphasize that the numerical optimization framework is flexible, and there are many different terms that one could add to the Lagrangian depending on the setting and task. The Lagrangian we have chosen for illustration purposes attempts to find superresolution

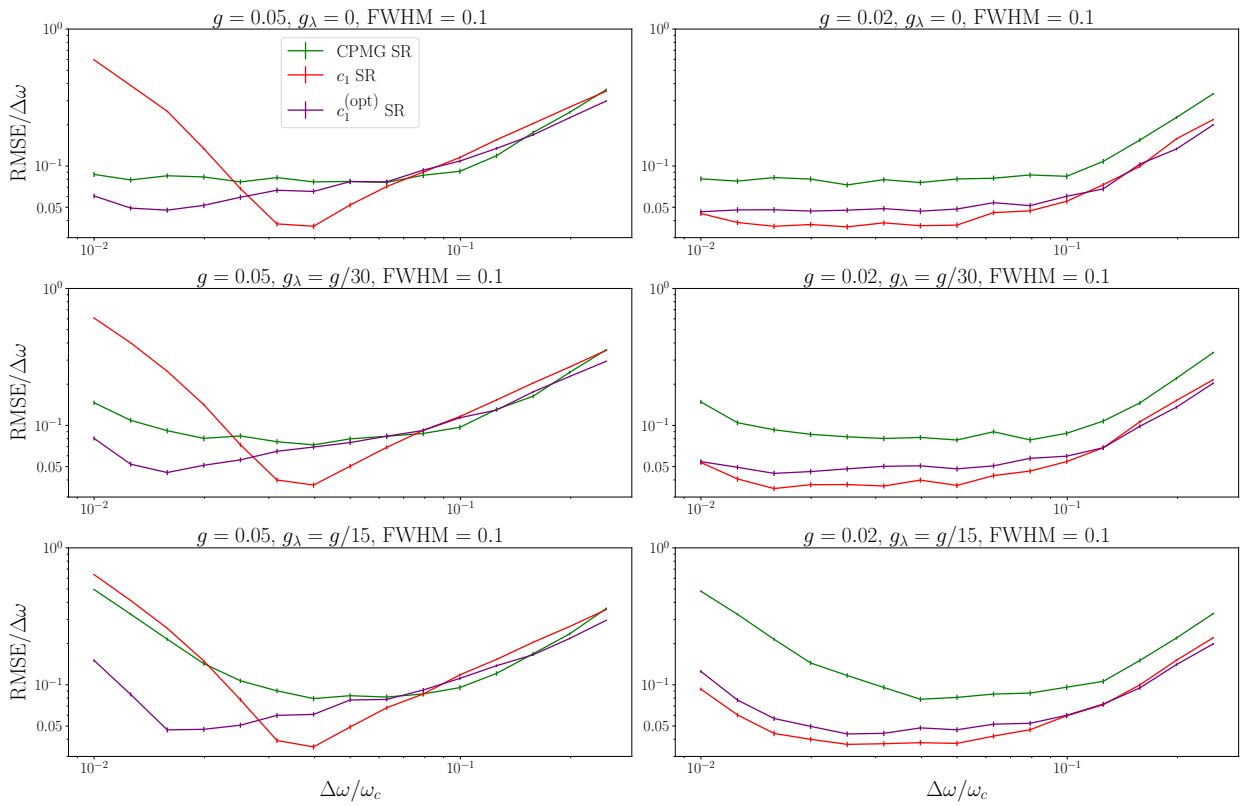

Figure 7.5: This plot is analogous to Fig. 7.2, except we now also consider continuous control protocols. We numerically simulate the CPMG protocol, the c_1 protocol, and the optimized $c_1^{(\mathrm{opt})}$ protocol (where recall we optimized for a small and smooth control amplitude with small fourth filter function), all with $\kappa = 4$. For each of the protocols, we utilize N_{CPMG} samples for a desired noiseless relative error of $\delta = 0.1$ given in Eq. (7.11). Note that all the protocols fail at large enough $\Delta\omega$ because we are using a second order Taylor expansion approximation for $\langle P \rangle$. (Top left): Noiseless case ($g_\lambda = W = 0$) and $g = 0.06$. As $\Delta\omega \to 0$, the instantaneous control protocols achieve the desired relative accuracy while the c_1 protocol fails due to the higher order filter functions. However, there is a range of $\Delta\omega$ values where the c_1 protocol outperforms the instantaneous control protocols due to their filter functions having a larger second derivative at the centroid. Meanwhile, due to its small fourth filter function, the c_1^{opt} protocol performs well for smaller values of $\Delta\omega$. (Top right): Noiseless case ($g_\lambda = W = 0$) and $g = 0.02$. This is similar to the top left case, except that now because g is smaller, the effect of the higher order filter functions is reduced, resulting in a larger range of values of $\Delta\omega$ for which the continuous control protocols outperform the instantaneous control protocols. (Middle left): Noisy case ($g_\lambda = g/30$, FWHM $W = 0.1$) and $g = 0.06$. (Middle right): Noisy case ($g_\lambda = g/30$) and $g = 0.02$. This case is the same as the Middle left, except again because of the smaller g value, the continuous control protocols perform better for longer. (Bottom left): Noisy case ($g_\lambda = g/15$, FWHM $W = 0.1$) and $g = 0.06$. (Bottom right): Noisy case ($g_\lambda = g/15$) and $g = 0.02$. Same as the bottom left, except due to a smaller g value, the continuous control protocols perform better.

controls that are the most robust to noise, are small and smoothly varying, and have the largest regime of validity according to Eq. (7.19). For example, if instead of Lorentzian noise, the relevant noise model is $1/f$ noise as is common in superconducting qubit platforms [290], then the form of $\mathcal{L}$ is the same except that S_λ now becomes the spectrum of $1/f$ noise.

We numerically optimize $\mathcal{L}$ with respect to the controls $c(t)$. For illustration, we seed our optimizer with the c_1 protocol given in Eq. (7.20), though we replace the instantaneous pulse in the middle simply by a random small number. We will fix $\kappa = 4$. After optimizing, we find new controls, which we call $c_1^{(\mathrm{opt})}(t)$. The resulting control sequence is shown in Fig. 7.4, along with the control sequence for c_1 for comparison. Also in Fig. 7.4, we show the corresponding filter functions for $c_1^{(\mathrm{opt})}$, c_1, and CPMG. Note that one will get different results depending on what is being optimized for. In this example, we want $\max_t |c_1^{(\mathrm{opt})}(t)|$ to be small and we want $c_1^{\mathrm{opt}}(t)$ to be smooth. We see that $\max_t |c_1^{(\mathrm{opt})}(t)| \sim 3/4$ and $c_1^{\mathrm{opt}}(t)$ is quite smooth, whereas the ideal CPMG and c_1 protocols have infinite amplitude jumps. Despite the restriction on the amplitude, the numerical optimization is still able to find a good control sequence that performs well. Specifically, the $c_1^{(\mathrm{opt})}(t)$ and c_1 protocols both have a much larger value of $F''(\omega_c, \kappa\tau)$ and have smaller overlap with low frequency Lorentzian noise, and $c_1^{(\mathrm{opt})}(t)$ is able to achieve this even with small amplitude controls. Additionally, because we optimized c_1^{opt} to have a small value of $F^{(4)}(\omega_c, \omega_c, \kappa\tau)$, c_1^{opt} will outperform c_1 at larger values of g where higher filter functions matter.

We simulate the protocols to produce the analogous plot as Fig. 7.2, and all of these effects are shown in Fig. 7.5. We see that when $\mathcal{O}(g^4)$ terms are irrelevant, both the c_1 and $c_1^{(\mathrm{opt})}$ protocol outperform CPMG, both in the noiseless and noisy cases.

Section 7.5: Entanglement enhancement

In this section, we show that entangled qubits and measurements can take any superresolution protocol that utilizes a single qubit and N repetitions and convert it to a superres-

olution protocol that achieves the same relative error δ that utilizes $N_e \sim \frac{R}{2\delta^{1/2}}$ qubits and $N_r = \frac{2N}{N_e\sqrt{\delta}R}$ repetitions, with $R = \max(\sqrt{\log\left(\frac{\delta}{b\Delta\omega^2}\right)}, 1)$ and $b = \frac{g^2}{4}F''(\omega_c, \kappa\tau)$. Thus, as the relevant figure of merit is to compare N to N_eN_r, we see that utilizing entanglement requires a factor of $N_eN_r/N \sim 2\sqrt{\delta}/R$ fewer resources. Importantly, this comes with the caveat that the entangled qubits must be evolved with the Hamiltonian $H_{N_e}(t) = \sum_{i=1}^{N_e}(\gamma(t)\sigma_i^z + c(t)\sigma_i^x)$. This means that the N_e entangled qubits must be evolved by the signal with the *same* realization of A_1, A_2, B_1, and B_2.

For simplicity, we will derive this entanglement advantage by utilizing GHZ states, but note that one can get the same advantage (up to a constant factor) with other by using other states, such as Hamming weight $N_e/2$ Dicke states. We therefore let the initial state be $|\Psi_0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle^{\otimes N_e} + |1\rangle^{\otimes N_e}\right)$, evolve under $H_{N_e}(t)$ for a time $\kappa\tau$, and measure the overlap with $|\Psi_0\rangle$. Thus, if $U(t)$ is the time evolution operator under $H(t)$, then $U(t)^{\otimes N_e}$ is the time evolution operator under $H_{N_e}(t)$. We are therefore interested in the expected probability $\langle P_{N_e}\rangle$ with

$$P_{N_e} = \left\langle \left| \langle\Psi_0| U(\kappa\tau)^{\otimes N_e} |\Psi_0\rangle \right|^2 \right\rangle. \tag{7.23}$$

Restricting our attention to instantaneous controls, following the $N_e = 1$ calculation in Appendix E.1, we find that $\langle P_{N_e}\rangle = \frac{1}{2} + \frac{1}{2}\mathrm{e}^{-N_e^2\chi}$, with χ given in Eq. (7.6). Computing the Fisher information then yields

$$\lim_{\Delta\omega\to 0} \mathrm{FI}_{\Delta\omega}(\langle P_{N_e}\rangle) = N_e^2 \lim_{\Delta\omega\to 0} \mathrm{FI}_{\Delta\omega}(\langle P\rangle) = 4N_e^2 b. \tag{7.24}$$

In other words, the Fisher information for N_e entangled qubits is N_e^2 times that of a single qubit. It follows from the Cramér Rao bound that in order to achieve a relative error of ε, the protocol must repeated

$$N_r \sim \frac{1}{4N_e^2 b\delta^2\Delta\omega^2} \tag{7.25}$$

times.

Recall from below Eq. (7.13) that we require $N_r \gtrsim \frac{1}{b\delta\Delta\omega^2 R^2}$. Fixing N_r in this way yields $N_e \sim \frac{R}{2\sqrt{\delta}}$ and this a total resource requirement of $N_e N_r \sim \frac{1}{2Rb\delta^{3/2}\Delta\omega^2}$. Comparing this to the number of repetitions required when $N_e = 1$, $N \sim \frac{1}{4b\delta^2\Delta\omega^2}$, we see that utilizing $\frac{R}{2\sqrt{\delta}}$ entangled qubits requires a factor of $N_e N_r/N \sim 2\sqrt{\delta}/R$ fewer resources.

Finally, the same result holds for continuous controls superresolution protocols in their regime of applicability – that is, when the higher order filter functions can be neglected.

We note an interesting feature of the nature of the entanglement advantage in these superresolution protocols. If we were sampling from $\langle P_{N_e} \rangle$ rather than from different realizations of P_{N_e} (that is, with respect to different A_1, A_2, B_1, B_2 coefficients), then we could set $N_r = 1$. In this case, Eq. (7.25) is the standard square-root quantum entanglement advantage, because N_e can then be chosen to be $\sqrt{N}$. However, sampling from P_{N_e} gives the requirement on N_r coming from Eq. (7.13) in order to effectively sample from $\langle P_{N_e} \rangle$, therefore limiting the entanglement advantage to only $2\sqrt{\delta}/R$, which is essentially only a constant factor due to the logarithm in R.

Section 7.6: Comparison to other methods

In this section, we will compare superresolution protocols to other quantum and classical protocols. In particular, in Section 7.6.1, we will describe standard quantum noise spectroscopy (QNS) techniques for estimating $\Delta\omega$, and in Section 7.6.2, we will perform a classical Fisher information analysis on a few different settings, therefore analyzing the *best possible* classical methods for estimating $\Delta\omega$ in these settings.

Subsection 7.6.1: Quantum noise spectroscopy

In our superresolution protocols, we look for filter functions that vanish at the centroid ω_c but have a large second derivative. Traditional QNS methods, however, aim for filter functions with a large a narrow peak at a fixed frequency ω^*. By sliding such a filter function along

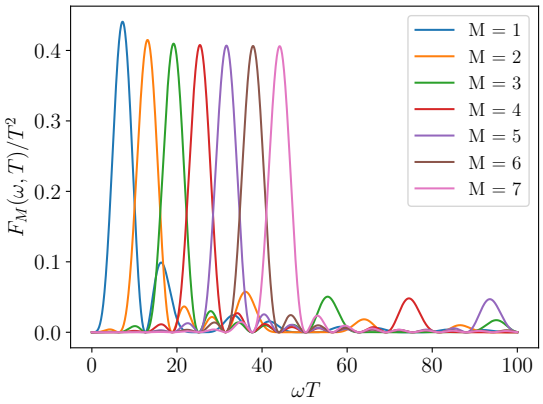

Figure 7.6: The filter functions given in Eq. (7.26) for a sequence of M CPMG sequences in a total time T. As M increases, the primary peaks slide along the frequency axis; as T increases the primary peaks become narrower in frequency.

the ω axis, the overlap between the filter function and the signal $S(\omega)$ will vary, being largest when ω^* is near ω_1 or ω_2. In this way, one can estimate $\Delta\omega = \omega_2 - \omega_1$. In stark contrast to superresolution protocols, these QNS methods require $\kappa \to \infty$ as $\Delta\omega \to 0$ in order for the peak in the filter function to be narrow than $\sim \Delta\omega$ around ω^*. Thus, superresolution protocols fundamentally outperform traditional QNS methods at estimating $\Delta\omega$ as they only require the signal to be observed for a fixed, finite amount of time.

To be more specific, we describe one particular traditional QNS method that utilizes CPMG sequences. Given a total evolution time of T, we consider the filter function arising from applying M CPMG sequences – in this case, the controls consist of instantaneous σ^x gates at time $T/(4M), 3T/(4M), 5T/(4M), \ldots, (4M-1)T/(4M)$. The analytic form of the resulting filter function is

$$F_M(\omega, T) = \frac{16}{\omega^2} \sec^2\left(\frac{\omega T}{4M}\right) \sin^2\left(\frac{\omega T}{2}\right) \sin^4\left(\frac{\omega T}{8M}\right), \tag{7.26}$$

and the resulting filter function is shown in Fig. 7.6.

A key feature of $F_M(\omega, T)$ is that it has a large peak at some ω^* that depends on M, and

the width of this peak scales inversely with T. Thus, given the signal $\gamma(t)$, we can evolve a qubit under $\gamma(t)$ and the M CPMG sequences for a time T and measure the expectation value of σ^x. This will yield the value of the overlap between the $F_M(\omega, T)$ and $S(\omega)$. By performing this protocol for many different values of M, we can estimate ω_1 and ω_2 based on when these expectation values are maximal. However, the resulting resolution along the ω-axis only scales as $1/T$, so that T must scale at least as $1/\Delta\omega$ as $\Delta\omega \to 0$.

Subsection 7.6.2: Classical methods

Throughout this work, we have considered the signal $\gamma(t)$ given in Eq. (7.2). A superresolution protocol begins with the quantum state $|+\rangle$, evolves with the Hamiltonian in Eq. (7.5) for a time $\kappa\tau$, and then measures in the σ^+ basis. This protocol is repeated N times, where the $\boldsymbol{a} = (A_1, A_2, B_1, B_2)$ in $\gamma(t)$ are *different* i.i.d. random normal coefficients for each of the N runs. Such a procedure is applicable the scenario when the coherence time of the signal is small, so that one is not able to make multiple measurements of the signal with the *same* $\boldsymbol{a}$ coefficients. In this section, we consider the case when the coherence time is long enough such that one is able to make M measurements of $\gamma(t)$ with the same coefficients. This procedure is then repeated N/M times, resulting in N measurements of the signal. We want to determine what is the best possible accuracy a classical algorithm can achieve in its estimate of $\Delta\omega$ when given access to these N measurements.

Before performing the analysis, we summarize the results. When $M = 1$, the Fisher information vanishes, so that there is no algorithm that is able to achieve any accuracy in estimating $\Delta\omega$. Note that our quantum superresolution protocols also operate in the $M = 1$ regime and make N measurements; however, the key difference is that although the quantum procedures only measure once for a fixed realization of the $\boldsymbol{a}$ coefficients just as the classical procedure, the quantum protocol can have nonzero Fisher information because the qubit is subjected to time evolution corresponding to the signal for *all* times from 0 to $\kappa\tau$ with a

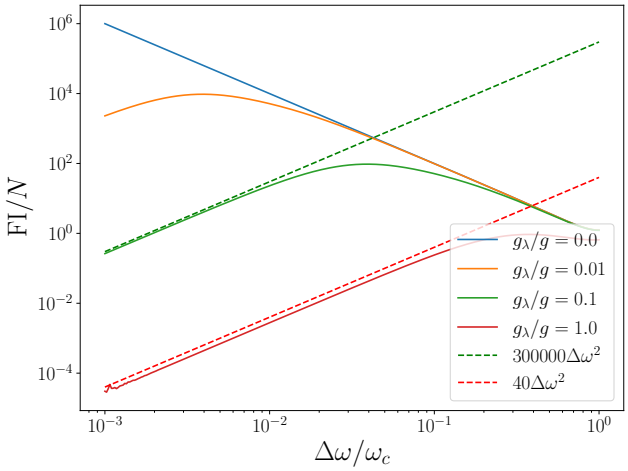

Figure 7.7: We plot FI$/N$ for the $M = 4$ measurement block with Lorentzian noise with FWHM $W = 1/10$ and strength g_λ, $\omega_c = 1$, and $\kappa = 1$. For this plot, the measurement times were chosen to be $t_m = (m-1)\kappa\tau/M$ for $m = 1, 2, 3, 4$.

fixed $\boldsymbol{a}$.

On the other hand, we also consider the case when the coherence is long enough so that $M > 1$ measurements are able to be made with a fixed realization of $\boldsymbol{a}$. We show that if $M \geq 4$ and there is no noise ($\lambda(t) = 0$), the classical Fisher information scales as $\sim \frac{N}{\Delta\omega^2}$ as $\Delta\omega \to 0$, and hence a classical procedure is able to estimate $\Delta\omega$ arbitrarily well as $\Delta\omega \to 0$. We therefore see that the quantum superresolution protocols are best suited for when the coherence time is short (so that $M \geq 4$ is not practically achievable) and $\Delta\omega$ is small.

Finally, we also consider the classical $M \geq 4$ case when subjected to Lorentzian noise. We find that the Fisher information scales as $\sim \frac{N}{\Delta\omega^2}$ for large $\Delta\omega$ and as $\sim N\Delta\omega^2$ as $\Delta\omega \to 0$, which is in fact the scaling that the quantum superresolution protocols achieve despite them only having access to a single measurement per realization of $\boldsymbol{a}$. The Fisher information as a function of $\Delta\omega$ can be seen in Fig. 7.7 for various values of the noise strength.

Fisher information analysis

Because the Fisher information is additive, we can simply consider $N = M$, and then multiply the end result by a factor of N/M. Given $N = M$, we consider measuring the signal $\gamma(t)$ at

times $0 \leq t_1 < t_2 < \cdots < t_M \leq \kappa\tau$. The result is a vector $\boldsymbol{s} = (s_1, \ldots, s_M)$ of measurement outcomes. Specifically, from Eq. (7.2), in the noiseless case, we will have $\boldsymbol{s} = gD\boldsymbol{a}$, where D is the $M \times 4$ matrix

$$D = \begin{pmatrix} \cos(\omega_1 t_1) & \cos(\omega_2 t_1) & \sin(\omega_1 t_1) & \sin(\omega_2 t_1) \\ \vdots & \vdots & \vdots & \vdots \\ \cos(\omega_1 t_M) & \cos(\omega_2 t_M) & \sin(\omega_1 t_M) & \sin(\omega_2 t_M) \end{pmatrix}. \tag{7.27}$$

In the presence of noise λ, we instead have $\boldsymbol{s} = gD\boldsymbol{a} + g_\lambda \boldsymbol{\lambda}$, where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_M)$, and $g_\lambda \lambda_m$ is the value of the noise at time t_m.

In order the calculate the Fisher information, we are first interested in the probability density $\rho(\boldsymbol{s}|\Delta\omega)$ that quantifies the probability of measuring a given $\boldsymbol{s}$ over the random values of $\boldsymbol{a}$ and the random values of $\boldsymbol{\lambda}$. The probability distribution over $\boldsymbol{a}$ is simply a zero-mean, unit-variance independent Gaussian distribution on each factor, while the distribution over $\boldsymbol{\lambda}$ is zero-mean multivariate Gaussian distribution with $M \times M$ covariance matrix V^{-1} with entries $(V^{-1})_{ij} = \exp[-W\pi\,|t_i - t_j|]$. Thus, letting $\delta^{(M)}$ denote the Dirac delta function on M elements, we have that

$$\begin{aligned} \rho(\boldsymbol{s}|\Delta\omega) = \frac{\sqrt{|\det V|}}{(2\pi)^2(2\pi)^{M/2}} \int_{\mathbb{R}^4} \mathrm{d}\boldsymbol{a} \int_{\mathbb{R}^M} \mathrm{d}\boldsymbol{\lambda} \\ \times \exp\left[-\frac{1}{2}(\boldsymbol{a}^T\boldsymbol{a} + \boldsymbol{\lambda}^T V \boldsymbol{\lambda})\right] \delta^{(M)}(gD\boldsymbol{a} + g_\lambda\boldsymbol{\lambda} - \boldsymbol{s})\,. \end{aligned} \tag{7.28}$$

We compute the Fisher information in two cases, $M = 1$ and $M = 4$. When $M = 1$, a straightforward Gaussian integral shows that $\rho(\boldsymbol{s}|\Delta\omega)$ is independent of $\Delta\omega$ even if $g_\lambda = 0$. Therefore, the Fisher information is zero for all $\Delta\omega$ and N. Intuitively, this is due to the randomness in the $\boldsymbol{a}$ coefficients; measuring the signal N times with N different $\boldsymbol{a}$ yields a sequence of random measurements with no information about the underlying signal.

Next, we consider the $M = 4$ case. We assume that $t_1, \ldots, t_4$ have been chosen so that

D is an invertible 4×4 matrix. For example, $t_m = m\kappa\tau/4$ yields an invertible D unless κ/ω_c is an even integer. In this case, the $\delta^{(4)}$ function exactly cancels the integral over $\boldsymbol{a}$, and the $\boldsymbol{\lambda}$ integral can be performed as a standard multivariate Gaussian integral. We then find that $\rho(\boldsymbol{s}|\Delta\omega)$ is the probability density function of a multivariate zero-mean Gaussian with covariance matrix Σ^{-1}, where $\Sigma = \sigma - g_\lambda^2\sigma(V + g_\lambda^2\sigma)^{-1}\sigma$ and $\sigma = (D^{-1})^T D^{-1}$.

The Fisher information for a multivariate Gaussian distribution whose covariance matrix depends on an unknown parameter is well-known and takes the form [291, Eq. (6)]

$$\mathrm{FI} = \frac{N}{8}\,\mathrm{Tr}\left[\Sigma(\Sigma^{-1})'\Sigma(\Sigma^{-1})'\right], \tag{7.29}$$

where $(\Sigma^{-1})'$ denotes the entrywise derivative of Σ^{-1} with respect to $\Delta\omega$, and we put back the factor of N/M. Using that for any matrix A, $(A^{-1})' = -A^{-1}A'A^{-1}$, we can substitute for $\frac{\mathrm{d}}{\mathrm{d}\Delta\omega}(V + g_\lambda^2\sigma)^{-1}$ and $\frac{\mathrm{d}}{\mathrm{d}\Delta\omega}\sigma = \frac{\mathrm{d}}{\mathrm{d}\Delta\omega}(D^{-1T}D^{-1})$ in order to arrive at a numerically computable form of the Fisher information. The result is shown in Fig. 7.7.

In fact, in the noiseless case when $g_\lambda = 0$, the exact Fisher information can be computed using a computer symbolic algebra software. For example, when $\kappa = 1$, we find that

$$\mathrm{FI}_{\kappa=1} = \frac{N\pi^2}{8\omega_c^2}\left(2\csc^2\left(\frac{\pi\Delta\omega}{2\omega_c}\right) - 1\right) = \frac{N}{\Delta\omega^2} + \mathcal{O}(1)\,. \tag{7.30}$$

To see why the Fisher information diverges as $\Delta\omega \to 0$ in the noiseless case, we consider a simple protocol when $N = M = 5$. Note that of course the Fisher information can only increase with increasing M, and therefore the $M = 5$ Fisher information must also diverge as $\Delta\omega \to 0$. To be concrete, we will let $\kappa = 1$, and by assumption there is a single realization of $\boldsymbol{a}$. We measure the signal at times $0, \tau/5, 2\tau/5, 3\tau/5, 4\tau/5$. Then the signal that we measure will be $\boldsymbol{s} = D\boldsymbol{a}$. This time, because $M = 5$, D is 5×4; As $\Delta\omega \to 0$, $D = D_0 + \Delta\omega D_1 + \Delta\omega^2 D_2 + \mathcal{O}(\Delta\omega^3)$. Hence, for small $\Delta\omega$, after measuring the signal five times and getting $\boldsymbol{s}$, we have a system of five equations for five unknowns $(A_1, A_2, B_1, B_2$

and $\Delta\omega$), $D_0\boldsymbol{a} + \Delta\omega D_1\boldsymbol{a} + \Delta\omega^2 D_2\boldsymbol{a} = \boldsymbol{s}$. From here, we can simply solve for the unknowns in order to determine $\Delta\omega$. As $\Delta\omega \to 0$, the Taylor expansion of D becomes more and more acccurate, meaning our estimate of $\Delta\omega$ will also become more and more accurate. Therefore, in this limit, we can achieve arbitrarily good accuracy in our estimate with only five measurements. Recall the Cramér Rao bound says that the optimal estimation error ε^2 behaves as $\frac{1}{N\,\mathrm{FI}}$. Because it is possible to achieve arbitrarily good accuracy (*i.e.* $\varepsilon \to 0$) with $N = M = 5$ as $\Delta\omega \to 0$, we see that the Fisher information must diverge in this limit.

In conclusion, we see that if the coherence time of the signal is long enough that multiple measurements of the signal are able to be made with a single realization of $\boldsymbol{a}$, then simple classical methods will outperform quantum superresolution protocols. On the other hand, in the short coherence time setting, quantum protocols are the better choice.

Section 7.7: Conclusion

In this work, we studied quantum superresolution for frequency resolution and estimation. We found explicit conditions on the filter function defined solely in terms of the control Hamiltonian for a protocol to exhibit superresolution. Using these conditions, we find many such protocols, as well as show how to numerically optimize control sequences to find controls that exhibit superresolution and satisfy other conditions, such as robustness to a specified noise model, smoothness of controls, etc. We further showed that any superresolution protocol can be improved with entanglement. We compared quantum superresolution methods to standard QNS methods and other classical methods in order to understand the regime in which quantum superconducting protocols will outperform other methods.

Part IV

Appendices

Appendix A: Technical details for Chapter 2

Section A.1: Preliminaries

In this preliminary appendix, we will establish some notation and equations that will be used throughout the rest of the appendices. In particular, in Appendix A.1.1, we review bosonic Gaussian states and describe our setup. In Appendix A.1.2, we describe integration over the unitary group with the Weingarten calculus. Finally, in Appendix A.1.3, we restrict our attention to the case when all initial squeezing strengths are equal and derive a series formula for the Rényi-2 entropy that is used in many of our proofs.

Subsection A.1.1: Bosonic Gaussian states

Here, we describe the setup and fix the notation required for the proofs of our main results. We consider a very similar setup as the one described in Ref. [33] and use much of the same notation as them. For a review of bosonic Gaussian states, we recommend Ref. [138]. Since we are only interested in entanglement properties, the first moments — displacements — of the Gaussian states will be irrelevant, and we will ignore them.

We consider a system of n bosonic modes. Each mode $1 \leq i \leq n$ is initially in a squeezed state with squeezing strength $s_i \in \mathbb{R}$. Define the diagonal matrix $Z = \mathrm{diag}(\mathrm{e}^{2s_1}, \ldots, \mathrm{e}^{2s_n})$. The initial state can be represented by the covariance matrix $\sigma_0 = Z \oplus Z^{-1}$. Define $A = \frac{1}{2}(Z - Z^{-1})$ and $B = \frac{1}{2}(Z + Z^{-1})$.

The set of all passive Gaussian unitaries — that is, energy-conserving unitaries — acting

on n modes is $\mathrm{Sp}(2n) \cap \mathrm{O}(2n)$ which acts on the covariance matrix by conjugation. Here $\mathrm{O}(2n)$ is the orthogonal group of $2n \times 2n$ matrices, and $\mathrm{Sp}(2n)$ is the real symplectic group of $2n \times 2n$ matrices defined with respect to the symplectic form

$$\Omega \coloneqq \begin{pmatrix} 0_{n\times n} & \mathbb{I}_{n\times n} \\ -\mathbb{I}_{n\times n} & 0_{n\times n} \end{pmatrix}. \tag{A.1}$$

Let $\eta\colon \mathrm{U}(n) \to \mathrm{Sp}(2n) \cap \mathrm{O}(2n)$ be the isomorphism

$$\eta(U) = \begin{pmatrix} \mathrm{Re}(U) & \mathrm{Im}(U) \\ -\mathrm{Im}(U) & \mathrm{Re}(U) \end{pmatrix}. \tag{A.2}$$

We evolve the initial state with covariance matrix σ_0 by a passive Gaussian unitary, which corresponds to a $U \in \mathrm{U}(n)$. The resulting state is $\tilde{\sigma}(U) \coloneqq \eta(U)\sigma_0\eta(U^\dagger) = \eta(U)\sigma_0\eta(U)^T$.

Define the $k \times n$ matrix P and the $n \times n$ projector Π as

$$P \coloneqq \begin{pmatrix} \mathbb{I}_{k\times k} & 0_{k\times(n-k)} \end{pmatrix}, \qquad \Pi \coloneqq \begin{pmatrix} \mathbb{I}_{k\times k} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & 0_{(n-k)\times(n-k)} \end{pmatrix} = \mathbb{I}_{k\times k} \oplus 0_{(n-k)\times(n-k)}. \tag{A.3}$$

Then let $\hat{P} \coloneqq P \oplus P$ and $\hat{\Pi} = \Pi \oplus \Pi$. The covariance matrix corresponding to the reduced state on the first $k \leq n$ modes is $\sigma(U) \coloneqq \hat{P}\tilde{\sigma}(U)\hat{P}^T$. Denote the element-wise complex conjugate of the unitary U by $\bar{U}$, and the conjugate transpose by $U^\dagger$. By simply doing the matrix multiplication, one finds that

$$\sigma(U) = \frac{1}{2}\begin{pmatrix} P\left[UBU^\dagger + \bar{U}B\bar{U}^\dagger + UA\bar{U}^\dagger + \bar{U}AU^\dagger\right]P^T & -\mathrm{i}P\left[UBU^\dagger - \bar{U}B\bar{U}^\dagger - UA\bar{U}^\dagger + \bar{U}AU^\dagger\right]P^T \\ \mathrm{i}P\left[UBU^\dagger - \bar{U}B\bar{U}^\dagger + UA\bar{U}^\dagger - \bar{U}AU^\dagger\right]P^T & P\left[UBU^\dagger + \bar{U}B\bar{U}^\dagger - UA\bar{U}^\dagger - \bar{U}AU^\dagger\right]P^T \end{pmatrix}. \tag{A.4}$$

Note that $\sigma(U)$ is a covariance matrix on k modes and is correspondingly a positive $2k \times 2k$ matrix. Throughout this paper, we define $r \coloneqq k/n$. One can derive from Eq. (A.6) that the

average over the Haar measure is $\mathbb{E}_U UBU^\dagger = \mathbb{E}_U \bar{U}B\bar{U}^\dagger = \frac{\operatorname{Tr} B}{n}\mathbb{I}_{n\times n}$, whereas all the other terms have expectation value 0 since they do not contain an even number of U's and $\bar{U}$'s. Therefore,

$$\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \sigma(U) = \frac{\operatorname{Tr} B}{n}\mathbb{I}_{2k\times 2k}. \tag{A.5}$$

The *symplectic eigenvalues* of $\sigma(U)$ are the positive eigenvalues of $\mathrm{i}\Omega\sigma(U)$. There are k symplectic eigenvalues labeled as ν_i for $1 \leq i \leq k$. Let the von Neumann entropy of the reduced state be $S_1(U)$, and the Rényi-2 entropy of the reduced state be $S_2(U)$. Then $S_j(U) = \sum_{i=1}^k h_j(\nu_i)$, where $h_1(x) = \frac{x+1}{2}\log\frac{x+1}{2} - \frac{x-1}{2}\log\frac{x-1}{2}$ and $h_2(x) = \log x$ [138]. The Rényi-2 entropy of the reduced state takes a particularly nice form in terms of the standard eigenvalues of $\sigma(U)$, namely $S_2(U) = \frac{1}{2}\log\det\sigma(U) = \frac{1}{2}\operatorname{Tr}\log\sigma(U)$. The dependence of $S_1(U)$ and $S_2(U)$ on r, n, and s is implicit.

Subsection A.1.2: Weingarten calculus

Since we are interested in average entanglement, we will be averaging over the unitary group $\mathrm{U}(n)$ with respect to the unique unit normalized Haar measure. To do so, we will use the *Weingarten calculus* [146, 147]. For a matrix U, let U_{ij} denote the entry in row i and column j. Then,

$$\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} U_{i_1 j_1}\dots U_{i_q j_q}\bar{U}_{i'_1 j'_1}\dots\bar{U}_{i'_q j'_q} = \sum_{\sigma,\tau\in S_q} \delta_{i_1 i'_{\sigma(1)}}\dots\delta_{i_q i'_{\sigma(q)}}\delta_{j_1 j'_{\tau(1)}}\dots\delta_{j_q j'_{\tau(q)}}\mathrm{Wg}(\sigma\tau^{-1}, n), \tag{A.6}$$

where S_q denotes the permutation group on q elements. Wg is called the *Weingarten function*. In our proofs, we will need the asymptotic form of the Weingarten function, which is given by

$$\mathrm{Wg}(\sigma, n) = \frac{1}{n^{q+|\sigma|}}\prod_i (-1)^{|c_i|-1}C_{|c_i|-1} + \mathcal{O}\left(n^{-q-|\sigma|-2}\right), \tag{A.7}$$

where $|\sigma|$ denotes the minimum number of transpositions needed to generate the permutation σ, $C_m = (2m)!/m!(m+1)!$ is the m^{th} Catalan number, and σ is a product of cycles c_i of length $|c_i|$.

Subsection A.1.3: Series formula for the Rényi-2 entropy

In this subappendix, we Taylor expand $S_2(U) = \frac{1}{2}\operatorname{Tr}\log\sigma(U)$ and derive a series formula for the Rényi-2 entropy when all the initial squeezing values are equal. Hence, for each $1 \leq i \leq n$, we set $s_i = s$. Crucially, the resulting formula is a series in the squeezing strength s, and the effect of the unitary U is separated from that of the squeezing strength s.

We would like to apply the Taylor series for the matrix logarithm, and hence must first consider its convergence. We find that

$$\begin{aligned}
\|\sigma(U) - \mathbb{I}_{2k\times 2k}\| &= \left\|\hat{P}\eta(U)\left(\sigma_0 - \mathbb{I}_{2n\times 2n}\right)\eta(U)^T\hat{P}^T\right\| && \text{(A.8)}\\
&= \left\|\hat{\Pi}\eta(U)\left(\sigma_0 - \mathbb{I}_{2n\times 2n}\right)\eta(U)^T\hat{\Pi}^T\right\| && \text{(A.9)}\\
&\leq \left\|\hat{\Pi}\right\|^2\left\|\eta(U)\left(\sigma_0 - \mathbb{I}_{2n\times 2n}\right)\eta(U)^T\right\| && \text{(A.10)}\\
&= \left\|\eta(U)\left(\sigma_0 - \mathbb{I}_{2n\times 2n}\right)\eta(U)^T\right\| && \text{(A.11)}\\
&= \|\sigma_0 - \mathbb{I}_{2n\times 2n}\| && \text{(A.12)}\\
&= \max\left\{\left|\mathrm{e}^{2s} - 1\right|, \left|\mathrm{e}^{-2s} - 1\right|\right\} && \text{(A.13)}\\
&= \mathrm{e}^{2|s|} - 1, && \text{(A.14)}
\end{aligned}$$

and therefore the Taylor series for log,

$$\log\sigma(U) = -\sum_{j=1}^{\infty}\frac{(-1)^j}{j}\left(\sigma(U) - \mathbb{I}_{2k\times 2k}\right)^j, \tag{A.15}$$

converges for all $|s| < R \coloneqq \frac{1}{2}\log 2$. Since $\sigma(U)$ is a positive, real symmetric matrix, this expression is indeed real and nonnegative. To make this work for all $s \in \mathbb{R}$, we can let $N \geq 1$

and consider

$$S_2(U) = \frac{1}{2}\log\det\sigma(U) \tag{A.16}$$

$$= \frac{1}{2}\log\left(N^{2k}\det\frac{\sigma(U)}{N}\right) \tag{A.17}$$

$$= k\log N + \frac{1}{2}\operatorname{Tr}\log\left(\frac{1}{N}\sigma(U)\right), \tag{A.18}$$

which follows from the fact that the determinant of products is equal to the product of determinants. For any given $s \in \mathbb{R}$, we can choose N large enough such that $\left\|\frac{1}{N}\sigma(U) - \mathbb{I}\right\| < 1$ and therefore the Taylor series for log can be used. We therefore find that for large enough N,

$$S_2(U) = rn\log N - \frac{1}{2}\operatorname{Tr}\sum_{j=1}^{\infty}\frac{(-1)^j}{j}\left(\frac{1}{N}\sigma(U) - \mathbb{I}_{2k\times 2k}\right)^j. \tag{A.19}$$

When all $s_i = s$, A simplifies to $A = \sinh(2s)\mathbb{I}_{n\times n}$ and B to $B = \cosh(2s)\mathbb{I}_{n\times n}$, and therefore $\sigma(U)$ simplifies to $\sigma(U) = \cosh(2s)\mathbb{I}_{2k\times 2k} + \sinh(2s)M$, where

$$M := \begin{pmatrix} P\operatorname{Re}(\bar{U}U^\dagger)P^T & P\operatorname{Im}(\bar{U}U^\dagger)P^T \\ P\operatorname{Im}(\bar{U}U^\dagger)P^T & -P\operatorname{Re}(\bar{U}U^\dagger)P^T \end{pmatrix}. \tag{A.20}$$

With Eq. (A.19),

$$S_2(U) = nr\log N - \frac{1}{2}\operatorname{Tr}\sum_{j=1}^{\infty}\frac{(-1)^j}{j}\left(\left[\frac{1}{N}\cosh(2s) - 1\right]\mathbb{I}_{2k\times 2k} + \frac{1}{N}\sinh(2s)M\right)^j \tag{A.21}$$

$$= nr\log N - \frac{1}{2}\operatorname{Tr}\sum_{j=1}^{\infty}\frac{(-1)^j}{j}\sum_{\ell=0}^{j}\binom{j}{\ell}\frac{1}{N^\ell}\sinh^\ell(2s)\left[\frac{1}{N}\cosh(2s) - 1\right]^{j-\ell}M^\ell \tag{A.22}$$

$$\begin{aligned} = nr\log N &- rn\sum_{j=1}^{\infty}\frac{(-1)^j}{j}\left[\frac{1}{N}\cosh(2s) - 1\right]^j \\ &- \frac{1}{2}\operatorname{Tr}\sum_{j=1}^{\infty}\frac{(-1)^j}{j}\sum_{\ell=1}^{j}\binom{j}{\ell}\frac{1}{N^\ell}\sinh^\ell(2s)\left[\frac{1}{N}\cosh(2s) - 1\right]^{j-\ell}\operatorname{Tr}M^\ell \end{aligned} \tag{A.23}$$

$$= nr \log N + nr \log(\cosh(2s)/N) - \frac{1}{2} \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sum_{\ell=1}^{j} \binom{j}{\ell} \frac{1}{N^\ell} \sinh^\ell(2s) \left[\frac{1}{N} \cosh(2s) - 1 \right]^{j-\ell} \operatorname{Tr} M^\ell \tag{A.24}$$

$$= nr \log \cosh(2s) - \frac{1}{2} \sum_{\ell=1}^{\infty} \frac{1}{N^\ell} \sinh^\ell(2s) \operatorname{Tr} M^\ell \sum_{j=\ell}^{\infty} \frac{(-1)^j}{j} \binom{j}{\ell} \left[\frac{1}{N} \cosh(2s) - 1 \right]^{j-\ell} \tag{A.25}$$

$$= nr \log \cosh(2s) - \frac{1}{2} \sum_{\ell=1}^{\infty} \frac{1}{N^\ell} \sinh^\ell(2s) \operatorname{Tr} M^\ell \frac{(-1)^\ell}{\ell} \left(N \operatorname{sech}(2s) \right)^\ell \tag{A.26}$$

$$= nr \log \cosh(2s) - \frac{1}{2} \sum_{\ell=1}^{\infty} \frac{(-1)^\ell}{\ell} \tanh^\ell(2s) \operatorname{Tr} M^\ell. \tag{A.27}$$

Of course, $S_2(U)$ is independent of the choice of N, and hence the N dependence has dropped out.

The only thing left to compute is $\operatorname{Tr} M^\ell$. Since we are now only dealing with traces, we can replace P with Π in M. This nicely simplifies some formulas, since Π is a square matrix, and indeed a projector, whereas P is a rectangular matrix. Henceforth, we will therefore let

$$M = \begin{pmatrix} \Pi R \Pi & \Pi I \Pi \\ \Pi I \Pi & -\Pi R \Pi \end{pmatrix}, \tag{A.28}$$

where $R = \operatorname{Re}(\bar{U} U^\dagger)$ and $I = \operatorname{Im}(\bar{U} U^\dagger)$. Then, doing the matrix multiplication,

$$M^2 = \begin{pmatrix} \Pi R \Pi R \Pi + \Pi I \Pi I \Pi & \Pi R \Pi I \Pi - \Pi I \Pi R \Pi \\ \Pi I \Pi R \Pi - \Pi R \Pi I \Pi & \Pi R \Pi R \Pi + \Pi I \Pi I \Pi \end{pmatrix}. \tag{A.29}$$

We then notice that $M^2 = \begin{pmatrix} D & E \\ -E & D \end{pmatrix}$, where $D = \operatorname{Re}(W)$, $E = \operatorname{Im}(W)$, and $W = \Pi U \bar{U}^\dagger \Pi \bar{U} U^\dagger \Pi$. It is then easy to verify that $M^{2j} = \begin{pmatrix} D_j & E_j \\ -E_j & D_j \end{pmatrix}$, where $D_{j+1} = D_j D - E_j E$,

$E_{j+1} = D_j E + E_j D$, $D_1 = D$, and $E_1 = E$. It is then also easy to verify that this recurrence relation is solved by $D_j = \mathrm{Re}(W^j)$ and $E_j = \mathrm{Im}(W^j)$. Finally, $\mathrm{Tr}\, M^{2j} = 2\,\mathrm{Tr}\, D_j$.

Using this recurrence, we can do the matrix multiplication $M^{2j+1} = M^{2j} M$ to find that $\mathrm{Tr}\, M^{2j+1} = 0$. Hence, we only need to worry about even powers of M, giving

$$S_2(U) = nr \log\cosh(2s) - \frac{1}{2}\sum_{\ell=1}^{\infty} \frac{1}{2\ell} \tanh^{2\ell}(2s)\, \mathrm{Tr}\, M^{2\ell}. \tag{A.30}$$

We then use that $\mathrm{Tr}\, M^{2j} = 2\,\mathrm{Tr}\, D_j$ to find that

$$S_2(U) = nr \log\cosh(2s) - \sum_{\ell=1}^{\infty} \frac{1}{2\ell} \tanh^{2\ell}(2s)\, \mathrm{Tr}\,\mathrm{Re}\, W^{\ell}. \tag{A.31}$$

Furthermore, since W is Hermitian, $\mathrm{Tr}\, W^j = \mathrm{Tr}\, \bar{W}^j$. Therefore, we arrive at

$$S_2(U) = nr \log\cosh(2s) - \sum_{\ell=1}^{\infty} \frac{1}{2\ell} \tanh^{2\ell}(2s)\, \mathrm{Tr}\, W^{\ell} \tag{A.32}$$

$$= n \sum_{\ell=1}^{\infty} \frac{\tanh^{2\ell}(2s)}{2\ell} \left(r - \frac{1}{n}\, \mathrm{Tr}\, W^{\ell} \right), \tag{A.33}$$

where in the last step we used the Taylor expansion of $\log\cosh(\mathrm{arctanh}\, t)$ in the variable $t = \tanh(2s)$.

Eqs. (A.32) and (A.33) hint at why equal initial squeezings simplify the problem of studying averaged entanglement properties. Specifically, the contribution from the squeezing strength and the contribution from the unitary are separated. Thus, to determine averaged entanglement properties, we only need to deal with the matrix $W = \Pi U U^T \Pi \bar{U} \bar{U}^T \Pi$.

Section A.2: Rényi-2 and von Neumann entropies — Proof of Proposition 2.2.1

In this appendix, we prove Proposition 2.2.1. We derive the maximum of the Rényi-2 and von Neumann entropies, and the former will be of use later when we derive the Rényi-2 Page correction. Furthermore, we prove that the von Neumann entropy can be bounded by the Rényi-2 entropy. In this way, our results on the Rényi-2 Page curve can be used to bound the von Neumann Page curve.

We begin by proving that the maximum of the Rényi-2 entropy is

$$\max_{U\in \mathrm{U}(n)} S_2(U) = n \min(r, 1-r) \log\cosh(2s). \tag{A.34}$$

Proof. Recall that $W = F^\dagger F$, where $F = \Pi \bar{U} U^\dagger \Pi$. Therefore, W is a nonnegative operator, and $\operatorname{Tr} W^\ell \geq 0$. The proposition then immediately follows from Eq. (A.32) if we can show that for every $r \leq 1/2$, there exists a unitary such that $W = 0$. The case when $r > 1/2$ is taken care of by the fact that the Rényi-2 entropy is symmetric under $r \mapsto 1-r$ since the global state on the n modes is pure. Therefore, we now assume that $r \leq 1/2$, and we show that there are unitaries $U \in \mathrm{U}(n)$ such that $W = 0$.

Since $F^\dagger F$ is nonnegative, we need to prove that there exists a unitary such that $F = \Pi \bar{U} U^\dagger \Pi = 0$. Hence we must prove that there exists a U such that $(\bar{U}U^\dagger)_{ji} = 0$ for all $1 \leq i, j \leq k = rn$. Equivalently, we can conjugate the expression, giving $\overline{(\bar{U}U^\dagger)}_{ji} = \sum_{a=1}^n U_{ia} U_{ja} = 0$. Therefore, we just need to find a set of $k = rn$ orthonormal vectors $\mathcal{S} = \{|\psi_1\rangle, \ldots |\psi_k\rangle\}$ in $\mathbb{C}^n$ such that $\langle \bar{\psi_i} | \psi_j \rangle = 0$. Let $|\psi_i\rangle = |R_i\rangle + \mathrm{i}\,|I_i\rangle$ for real vectors $|R_i\rangle$ and $|I_i\rangle$. Since $\mathcal{S}$ is orthonormal, we find that

$$\delta_{ij} = \langle \psi_i | \psi_j \rangle \tag{A.35}$$

$$= (\langle R_i| - \mathrm{i}\,\langle I_i|)(|R_j\rangle + \mathrm{i}\,|I_j\rangle) \tag{A.36}$$

$$= \langle R_i|R_j\rangle + \langle I_i|I_j\rangle + \mathrm{i}\left(\langle R_i|I_j\rangle - \langle I_i|R_j\rangle\right). \tag{A.37}$$

The condition that $\langle\bar{\psi}_i|\psi_j\rangle = 0$ implies that

$$0 = \langle\bar{\psi}_i|\psi_j\rangle \tag{A.38}$$

$$= (\langle R_i| + \mathrm{i}\langle I_i|)(|R_j\rangle + \mathrm{i}|I_j\rangle) \tag{A.39}$$

$$= \langle R_i|R_j\rangle - \langle I_i|I_j\rangle + \mathrm{i}\left(\langle R_i|I_j\rangle + \langle I_i|R_j\rangle\right). \tag{A.40}$$

Hence, we just need to find k vectors $|R_i\rangle \in \mathbb{R}^n$ and k vectors $|I_i\rangle \in \mathbb{R}^n$ satisfying

1. $\langle R_i|R_j\rangle + \langle I_i|I_j\rangle = \delta_{ij}$,
2. $\langle R_i|R_j\rangle - \langle I_i|I_j\rangle = 0$,
3. $\langle R_i|I_j\rangle + \langle I_i|R_j\rangle = 0$, and
4. $\langle R_i|I_j\rangle - \langle I_i|R_j\rangle = 0$,

for all $i, j \in \{1, \ldots, k\}$. This is trivial since $k \leq n/2$, and we give one construction here. Let $\mathcal{R} = \{|R_1\rangle, \ldots, |R_k\rangle\}$ be an orthogonal set satisfying $\langle R_i|R_j\rangle = \delta_{ij}/2$. Since $\dim[\mathrm{span}(\mathcal{R})^{\perp}] = n - k \geq k$, we can choose another orthogonal set $\mathcal{I} = \{|I_1\rangle, \ldots, |I_k\rangle\}$ satisfying $\langle I_i|I_j\rangle = \delta_{ij}/2$ such that $\mathrm{span}(\mathcal{I}) \cap \mathrm{span}(\mathcal{R}) = \{0\}$. This immediately means that condition 3 and 4 are satisfied, because $\langle R_i|I_j\rangle = 0$ for all i and j. Furthermore, condition 1 is satisfied because $\langle R_i|R_j\rangle + \langle I_i|I_j\rangle = \delta_{ij}/2 + \delta_{ij}/2 = \delta_{ij}$. Finally, condition 2 is satisfied because $\langle R_i|R_j\rangle - \langle I_i|I_j\rangle = \delta_{ij}/2 - \delta_{ij}/2 = 0$. □

Next, we use Eq. (A.34) to prove that the maximum of the von Neumann entropy is

$$\max_{U\in\mathrm{U}(n)} S_1(U) = n\min(r, 1-r)h_1(\cosh(2s)). \tag{A.41}$$

Proof. From Ref. [150], the Rényi-α entropies for $\alpha \geq 1$ are all increasing with increasing $\det \sigma$. Therefore, when S_2 increases, so does S_α. It follows that the unitary that maximizes S_2 also maximizes S_1. Then, when $r \leq 1/2$, Eqs. (A.20) and (A.34) together imply that S_1 is maximized when $\sigma = \cosh(2s)\mathbb{I}_{2k\times 2k}$ and thus when $\nu_i = \cosh(2s)$ for each i, yielding $\max_U S_1(U) = nrh_1(\cosh(2s))$. The case when $r > 1/2$ is taken care of by the symmetry of the von Neumann entropy when $r \mapsto 1 - r$ since we are dealing with a bipartite system [13]. □

Finally, we prove the bound

$$S_1(U) < S_2(U) + n \min(r, 1-r)(1 - \log 2), \tag{A.42}$$

which was first derived in Ref. [136, Eq. 15].

Proof. Recall that $S_j = \sum_{i=1}^k h_j(\nu_i)$, where $\nu_i \geq 1$ are the symplectic eigenvalues of the covariance matrix σ, $h_2(x) = \log x$, and $h_1(x) = \frac{x+1}{2}\log\frac{x+1}{2} - \frac{x-1}{2}\log\frac{x-1}{2}$ (see Appendix A.1.1). We will prove that $h_1(x) < h_2(x) + 1 - \log 2$, which proves the claim. We do this by noting that $h_1(x) - h_2(x)$ is monotonically increasing, so that for any x, $h_1(x) - h_2(x) < \lim_{y\to\infty}(h_1(y) - h_2(y))$. A simple calculation shows that this limit is equal to $1 - \log 2$. □

Section A.3: Rényi-2 Page curve

In this appendix, we prove our main results on the Rényi-2 Page curve. In Appendix A.3.1, we prove that asymptotically in n, $\mathbb{E}_{U\in \mathrm{U}(n)} S_2(U) = n\alpha(s,r) - \lambda(s,r) + o(1)$. We then derive the exact formula for the linear term $\alpha(s,r)$. Then, in Appendix A.3.2, we simplify $\alpha(s,1/2)$ and use the result to prove a simple expression for the Rényi-2 Page correction at $r = 1/2$. Finally, in Appendix A.3.3, we derive a formula for the constant term $\lambda(s,r)$ up to constant factors $a^{(\ell)}$. We then derive the formula for $a^{(\ell)}$, simplify it for $\ell \in \{1,2,3,4,5\}$,

and conjecture the exact value of $a^{(\ell)}$ for all ℓ.

Subsection A.3.1: Linear term — Proof of Theorem 2.3.1

Theorem 2.3.1 concerns the expectation value of $S_2(U)$ over $\mathrm{U}(n)$ when all the initial squeezing values are equal. From Eq. (A.32), we see that it only remains to compute $\mathbb{E}_U \operatorname{Tr} W^\ell$. Writing the matrix multiplication of $W^\ell = (\Pi U U^T \Pi \bar{U} \bar{U}^T \Pi)^\ell$ in terms of the matrix entries of U and Π and simplifying, we find that

$$\begin{aligned}\operatorname{Tr} W^\ell = \sum_{i_1,\ldots,i_{2\ell}=1}^{k} \sum_{i'_1,\ldots,i'_{2\ell}=1}^{k} \sum_{j_1,\ldots,j_{2\ell}=1}^{n} \sum_{j'_1,\ldots,j'_{2\ell}=1}^{n} \\ \delta_{i'_{2\ell},i_1}\delta_{i'_1,i_2}\delta_{i'_2,i_3}\delta_{i'_3,i_4}\ldots\delta_{i'_{2\ell-1},i_{2\ell}} \\ \times \delta_{j_1,j_2}\delta_{j'_1,j'_2}\delta_{j_3,j_4}\delta_{j'_3,j'_4}\ldots\delta_{j_{2\ell-1},j_{2\ell}}\delta_{j'_{2\ell-1},j'_{2\ell}} \\ \times U_{i_1,j_1}\ldots U_{i_{2\ell},j_{2\ell}}\bar{U}_{i'_1,j'_1}\ldots\bar{U}_{i'_{2\ell},j'_{2\ell}}.\end{aligned} \tag{A.43}$$

We note that this is a simple result of matrix multiplication. The restriction on the i and i' indices to $\{1,\ldots,k\}$ is a result of the fact that $\Pi_{a,a} = 0$ for all $a > k$. Applying Eq. (A.6), we immediately find that

$$\begin{aligned}\mathop{\mathbb{E}}_{U\in\mathrm{U}(n)} \operatorname{Tr} W^\ell = \sum_{i_1,\ldots,i_{2\ell}=1}^{k} \sum_{i'_1,\ldots,i'_{2\ell}=1}^{k} \sum_{j_1,\ldots,j_{2\ell}=1}^{n} \sum_{j'_1,\ldots,j'_{2\ell}=1}^{n} \sum_{\sigma,\tau\in S_{2\ell}} \mathrm{Wg}(\sigma\tau^{-1},n) \\ \times \delta_{i'_{2\ell},i_1}\delta_{i'_1,i_2}\delta_{i'_2,i_3}\ldots\delta_{i'_{2\ell-1},i_{2\ell}} \\ \times \delta_{j_1,j_2}\delta_{j'_1,j'_2}\ldots\delta_{j_{2\ell-1},j_{2\ell}}\delta_{j'_{2\ell-1},j'_{2\ell}} \\ \times \delta_{i_1,i'_{\sigma(1)}}\ldots\delta_{i_{2\ell},i'_{\sigma(2\ell)}} \\ \times \delta_{j_1,j'_{\tau(1)}}\ldots\delta_{j_{2\ell},j'_{\tau(2\ell)}}.\end{aligned} \tag{A.44}$$

Simplifying Eq. (A.44) at first seems impossible, but it will turn out that we do not need to. All we need to learn from it is the following lemma.

Lemma A.3.1. *Fix a positive integer ℓ. There exist coefficients $\alpha_d^{(\ell)}$ for $d \in \{\ell+1, \ell+2, \ldots, 2\ell\}$ such that*

$$f_\ell(r) \coloneqq \lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in \mathrm{U}(n)} \frac{1}{n} \operatorname{Tr} W^\ell = \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d. \tag{A.45}$$

Proof. The proof will proceed as follows. First, we will prove that $\operatorname{Tr} W^\ell$ contains a term proportional to n and no terms proportional to n^a for any $a > 1$. Therefore, $f_\ell(r)$ is indeed independent of n. Next, we will prove that $f_\ell(r)$ has no terms r^a for $a > 2\ell$ and $a \leq \ell$. Throughout this proof, we interpret the delta functions in Eq. (A.44) as constraints on the summations. Different permutations on the indices result in a different number of constraints and hence terms with different powers of n and k.

Recall that $W = F^\dagger F$ where $F = \Pi \bar{U} U^\dagger \Pi$. Therefore, $\operatorname{Tr} W$ is equal to the Frobenius norm $\|F\|_F^2$, which is the sum of the square absolute values of the entries of F. Thus, by removing the projector Π from W (i.e. setting it to $\mathbb{I}$), the trace cannot decrease. It follows that the presence of Π cannot increase the trace $\operatorname{Tr} W^\ell$. Getting rid of the Π from W and using the cyclic nature of the trace, we see that $\operatorname{Tr} W^\ell \leq n$. Furthermore, consider Eq. (A.44) with $\sigma = \tau$ defined by $\sigma(1) = 2\ell$, $\sigma(2) = 1$, $\sigma(3) = 2$, $\ldots$, $\sigma(2\ell) = 2\ell - 1$. Then, $\sigma\tau^{-1}$ is the identity, $\mathrm{Wg}(\sigma\tau^{-1}, n)$ contributes a factor of $n^{-2\ell}$. With this σ, the sum over the i and i' yields a factor of $k^{2\ell}$. Finally, with this chosen τ, the sum over j will yield $\delta_{j_1', j_2'} \ldots \delta_{j_{2\ell-1}', j_{2\ell}'} \delta_{j_1', j_{2\ell}'} \delta_{j_2', j_1'} \ldots \delta_{j_{2\ell}', j_{2\ell-1}'}$. Then summing over j', we get a single factor of n. Hence, the term with the specific permutation described above yields a term of the form $nk^{2\ell} n^{-2\ell} = nr^{2\ell}$.

We have shown that there is a term proportional to n and that there are no terms proportional to n^2, n^3, etc. Since we are working asymptotically in n, we can therefore ignore all terms proportional to $\frac{1}{n^a}$ for every nonnegative a. This proves that $\lim_{n\to\infty} \mathbb{E}_{U\in \mathrm{U}(n)} \frac{1}{n} \operatorname{Tr} W^\ell$ is independent of n and only depends on r, which justifies the definition of the function $f_\ell(r)$. The only thing left to show is that $f_\ell(r)$ has only terms $r^{\ell+1}$ through $r^{2\ell}$. So we only need to show that there are no terms r^a for $a > 2\ell$ and $a \leq \ell$. We begin with the former.

To look at powers of r, it is sufficient to look at the powers of k. We therefore restrict our attention to the sum over i and i' in Eq. (A.44). The sum over i_1, $\sum_{i_1=1}^{k}$, will give either a factor of 1 or a factor of k depending on how the index i_1 is constrained by the Kronecker delta functions, and similarly for i_2 through $i_{2\ell}$. In order to get the highest power of k, we require the fewest constraints on i and i' (i.e. the fewest *distinct* Kronecker deltas). Hence, we require σ to be the permutation satisfying

$$\delta_{i'_{2\ell},i_1}\delta_{i'_1,i_2}\delta_{i'_2,i_3}\dots\delta_{i'_{2\ell-1},i_{2\ell}} = \delta_{i_1,i'_{\sigma(1)}}\dots\delta_{i_{2\ell},i'_{\sigma(2\ell)}}. \tag{A.46}$$

This permutation is $\sigma(1) = 2\ell$, $\sigma(2) = 1$, $\sigma(3) = 2$, $\dots$, $\sigma(2\ell) = 2\ell - 1$. With this σ, we see that a sum over i and i' will give a factor of $k^{2\ell}$. Hence, 2ℓ is the highest power of k that can be achieved.

Next we need to show that $\ell + 1$ is the lowest power of k that can be achieved. The sum over j and j' can give at most a factor of n^ℓ. This is because the first line of delta functions, $\delta_{j_1,j_2}\delta_{j'_1,j'_2}\dots\delta_{j_{2\ell-1},j_{2\ell}}\delta_{j'_{2\ell-1},j'_{2\ell}}$, reduces the sum over 2ℓ indices j and 2ℓ indices j' down to just a sum over ℓ indices j and ℓ indices j'. The second line of delta functions, $\delta_{j_1,j'_{\tau(1)}}\dots\delta_{j_{2\ell},j'_{\tau(2\ell)}}$, cannot be made equivalent to the first line by any choice of τ; in fact, the second line imposes all new constraints. Therefore, this line further reduces the sum over ℓ indices j and ℓ indices j' down to just a sum over ℓ indices j (or ℓ indices j', but not both). Hence, the highest power of n that we get from the summations over j and j' is n^ℓ. Putting this together with the fact that asymptotically $\mathrm{Wg}(\pi, n)$ is at most $n^{-2\ell}$, we find that any term coming from Eq. (A.44) is at most $n^{-\ell} \times$ (dependence on k). Therefore, any powers of k that are less than $\ell + 1$ can be ignored; if the sum over i and i' yields a term that is k^a for some $a \leq \ell$, then that term will be constant or decreasing with n. But from above, we already have terms that are proportional to n resulting from Eq. (A.44), and so terms that are constant or decreasing can be ignored. □

As alluded to in the proof of the lemma, the asymptotic form of $\mathbb{E}_U \operatorname{Tr} W^\ell$ will be a term linear in n times a function of r, plus a term constant in n times a function of r, plus terms that decay to zero asymptotically with n. Hence, $\mathbb{E}_U S_2(U) = n\alpha(s,r) - \lambda(s,r) + o(1)$. In this section, we are interested only in the linear term $\alpha(s,r)$ because we are computing $\mathbb{E}_U S_2(U)/n$. However, in the next section, we prove Proposition 2.3.3 which provides the form of $\lambda(s,r)$.

Actually computing $\alpha_d^{(\ell)}$ from Eq. (A.44) seems challenging. However, there is a nice workaround that uses what we know about the Rényi-2 entropy being symmetric under $r \mapsto 1-r$, as the total state on the n modes is pure.

Lemma A.3.2. *Let $\alpha_d^{(\ell)}$ be as in Lemma A.3.1. Then*

$$\alpha_d^{(\ell)} = 2(-1)^{d-\ell-1} \binom{2\ell-1}{\ell-1} \binom{\ell}{d-\ell-1} \frac{2\ell-d+1}{(d-1)d}. \tag{A.47}$$

Note that this corresponds to the sequence A062991 on OEIS [292].

Proof. From Lemma A.3.1, asymptotically $\mathbb{E}_U \operatorname{Tr} W^\ell = n f_l(r)$ for a polynomial f_ℓ that is a sum over terms of degree $\ell+1$ through 2ℓ in r. From Eq. (A.33), we know that

$$\lim_{n\to\infty} \mathop{\mathbb{E}}_U \frac{1}{n} S_2(U) = \sum_{\ell=1}^{\infty} \frac{t^{2\ell}}{2\ell} \left(r - f_\ell(r)\right), \tag{A.48}$$

where $-1 < t = \tanh(2s) < 1$. This whole function must be symmetric under $r \to 1-r$ *at every order in t*. The reason it must be symmetric at every order is as follows. Suppose we choose s such that $t \in \mathcal{O}\big(n^{-1/2}\big)$. Then $\mathbb{E}_U S_2(U) = \frac{1}{2} n t^2 (r - f_1(r)) + o(1)$. Hence, $r - f_1(r)$ must be symmetric since S_2 is symmetric. We then choose s such that $t \in \mathcal{O}\big(n^{-1/4}\big)$ and consider $\mathbb{E}_U \big[S_2(U) - \frac{1}{2} n t^2 (r - f_1(r))\big]$, which will be equal to $\frac{1}{4} n t^4 (r - f_2(r)) + o(1)$. Since S_2 is symmetric and $r - f_1(r)$ is symmetric, it follows that $r - f_2(r)$ is symmetric. We then continue inductively like this to prove that $r - f_\ell(r)$ must be symmetric for all ℓ.

We therefore find that for all ℓ, $f_\ell(r)$ must satisfy $r - f_\ell(r) = 1 - r - f_\ell(1-r)$, or

$$f_\ell(r) - f_\ell(1-r) = 2r - 1. \tag{A.49}$$

We then plug in the form of $f_\ell(r) = \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d$. Eq. (A.49) must hold for every r, and therefore we can equate the coefficients in front of each r^d term to get a system of equations that can be solved for the values of $\alpha_d^{(\ell)}$. We then find that

$$\begin{aligned}
2r - 1 &= \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} \left(r^d - (1-r)^d\right) && \text{(A.50)}\\
&= \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d - \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} \sum_{j=0}^{d} \binom{d}{j} (-r)^j && \text{(A.51)}\\
&= \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d - \sum_{j=0}^{2\ell} (-r)^j \sum_{d=\max(j,\ell+1)}^{2\ell} \alpha_d^{(\ell)} \binom{d}{j} && \text{(A.52)}\\
&= \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d - \sum_{d=0}^{2\ell} (-r)^d \sum_{j=\max(d,\ell+1)}^{2\ell} \alpha_j^{(\ell)} \binom{j}{d} && \text{(A.53)}\\
&= -\sum_{d=0}^{\ell} (-r)^d \sum_{j=\ell+1}^{2\ell} \alpha_j^{(\ell)} \binom{j}{d} + \sum_{d=\ell+1}^{2\ell} r^d \left(\alpha_d^{(\ell)} - (-1)^d \sum_{j=d}^{2\ell} \alpha_j^{(\ell)} \binom{j}{d}\right). && \text{(A.54)}
\end{aligned}$$

Equating the degrees in r, we see the following conditions;

1. $1 = \sum_{j=\ell+1}^{2\ell} \alpha_j^{(\ell)}$,
2. $2 = \sum_{j=\ell+1}^{2\ell} j\alpha_j^{(\ell)}$,
3. For $2 \le d \le \ell$, $0 = \sum_{j=\ell+1}^{2\ell} \binom{j}{d} \alpha_j^{(\ell)}$,
4. For $\ell + 1 \le d \le 2\ell$, $\alpha_d^{(\ell)} = (-1)^d \sum_{j=d}^{2\ell} \binom{j}{d} \alpha_j^{(\ell)}$.

Note that one can derive equivalent conditions by requiring that the polynomial $(x+1/2) - f_\ell(x+1/2)$ is even in x. Condition 2 and condition 3 together constitute a linear system of ℓ linearly independent equations. To verify this, one must show that $\det C \neq 0$ where C

is the $\ell \times \ell$ matrix with entries $C_{ij} = \binom{\ell+i}{j}$. As noted by Benoit Cloitre in OEIS sequence A000984 [292], $\det C = \binom{2\ell}{\ell}$. For a proof, see Ref. [293].

Since we have ℓ linearly independent equations for ℓ variables α, if there is a solution to the four conditions then there is a *unique* solution. One can then verify that a solution to the four conditions, and therefore *the* solution, is given by Eq. (A.47).

When verifying the four conditions, one finds that the right hand side of all four of the conditions can be simplified in terms of the hypergeometric function ${}_2F_1$ defined as

$$ {}_2F_1(-m, b; c; r) = \sum_{a=0}^{m} (-1)^a \binom{m}{a} \frac{(c-1)!(a+b-1)!}{(b-1)!(a+c-1)!} r^a \tag{A.55} $$

when m is nonnegative [141–144]. For example, condition 4, written as $(-1)^d = \sum_{j=d}^{2\ell} \binom{j}{d} (\alpha_j^{(\ell)} / \alpha_d^{(\ell)})$, reduces to $(-1)^d = {}_2F_1(d-2\ell, d-1; d-\ell; 1)$. Define the Pochhammer symbol as $(x)_a := \frac{\Gamma(x+a)}{\Gamma(x)}$. When $a \in \mathbb{N}$, ${}_2F_1(-a, b; c; 1) = \frac{(c-b)_a}{(c)_a}$ [141–144]. We therefore find that

$$ {}_2F_1(d-2\ell, d-1; d-\ell; 1) = \frac{(1-\ell)_{2\ell-d}}{(d-\ell)_{2\ell-d}} \tag{A.56} $$

$$ = \frac{\Gamma(d-\ell)\Gamma(1-d+\ell)}{\Gamma(\ell)\Gamma(1-\ell)} \tag{A.57} $$

$$ = \frac{\Gamma(d-\ell)\Gamma(1-\ell)(-d+\ell)(-d+\ell-1)\dots(1-\ell)}{\Gamma(\ell)\Gamma(1-\ell)} \tag{A.58} $$

$$ = (-1)^{1+(1-\ell)-(-d+\ell)} \frac{\Gamma(d-\ell)(d-\ell)(d-\ell+1)\dots(\ell-1)}{\Gamma(\ell)} \tag{A.59} $$

$$ = (-1)^d \frac{\Gamma(d-\ell)(\ell-1)!}{\Gamma(\ell)(d-\ell-1)!} \tag{A.60} $$

$$ = (-1)^d, \tag{A.61} $$

hence proving condition 4. Conditions 1, 2, and 3 are similar. □

Plugging in this result, we find that

$$f_\ell(r) \coloneqq \frac{1}{n} \mathop{\mathbb{E}}_U \operatorname{Tr} W^\ell = 2\binom{2\ell-1}{\ell-1} \sum_{i=0}^{\ell-1} (-1)^i r^{i+\ell+1} \binom{\ell}{i} \frac{\ell-i}{(\ell+i)(\ell+i+1)} \tag{A.62}$$

$$= r^{\ell+1} C_\ell \, {}_2F_1(1-\ell, \ell; \ell+2; r), \tag{A.63}$$

where ${}_2F_1$ is the hypergeometric function and $C_\ell \coloneqq \frac{1}{\ell+1}\binom{2\ell}{\ell}$ is the ℓ^{th} Catalan number. The simplification from the first to second line follows from the definition of the hypergeometric function given in Eq. (A.55). Plugging this exact formula for $f_\ell(r) = \lim_{n\to\infty} \mathbb{E}_U \frac{1}{n} \operatorname{Tr} W^\ell$ into Eq. (A.32), and swapping sums, we arrive precisely at Theorem 2.3.1. We have therefore completed the proof.

To get a sense for f_ℓ, we list a few here.

$$f_1(r) = r^2$$

$$f_2(r) = -r^4 + 2r^3$$

$$f_3(r) = 2r^6 - 6r^5 + 5r^4$$

$$f_4(r) = -5r^8 + 20r^7 - 28r^6 + 14r^5$$

$$f_5(r) = 14r^{10} - 70r^9 + 135r^8 - 120r^7 + 42r^6$$

$$f_6(r) = -42r^{12} + 252r^{11} - 616r^{10} + 770r^9 - 495r^8 + 132r^7$$

$$f_7(r) = 132r^{14} - 924r^{13} + 2730r^{12} - 4368r^{11} + 4004r^{10} - 2002r^9 + 429r^8$$

$$f_8(r) = -429r^{16} + 3432r^{15} - 11880r^{14} + 23100r^{13} - 27300r^{12} + 19656r^{11} - 8008r^{10} + 1430r^9.$$

Recall that we required $G_\ell(r) \coloneqq r - f_\ell(r)$ to be symmetric under $r \mapsto 1-r$. We show some of the resulting plots in Fig. A.1. Intuitively, one can understand $G_\ell(r)$ as being an approximation to the function $m(r) \coloneqq \min(r, 1-r)$; as ℓ increases, this approximation gets better and better.

More specifically, we can interpret the polynomials $G_\ell(r)$ in multiple ways. Once the form

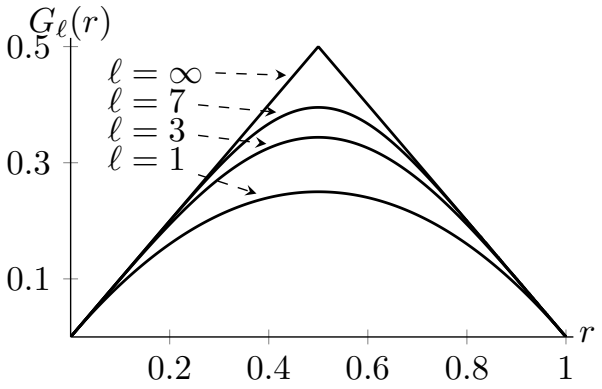

Figure A.1: The plots of $G_\ell(r) \coloneqq r - f_\ell(r)$ for various values of ℓ. $f_\ell(r)$ is given in Eq. (A.62). In our proof, we crucially used that $G_\ell(r)$ is symmetric under $r \mapsto 1 - r$.

of $f_\ell(r)$ is fixed by Lemma A.3.1, we showed in Lemma A.3.2 that G_ℓ is uniquely determined by the symmetry requirement that $G_\ell(r) = G_\ell(1-r)$. In the proof of Lemma A.3.2, we showed that G_ℓ is uniquely determined by conditions 1, 2, and 3. Each of these conditions has a simple interpretation. Condition 1, $1 = \sum_{j=\ell+1}^{2\ell} \alpha_j^{(\ell)}$, is the condition that $G_\ell(1) = 0$. Condition 2, $2 = \sum_{j=\ell+1}^{2\ell} j\alpha_j^{(\ell)}$, is the condition that the derivative $G'_\ell(1) = -1$. Condition 3, $0 = \sum_{j=\ell+1}^{2\ell} \binom{j}{d} \alpha_j^{(\ell)}$ for each $2 \leq d \leq \ell$, is the condition that the d^{th} derivative $G_\ell^{(d)}(1) = 0$. Hence, condition 1, 2, and 3 are imposing that $G_\ell^{(d)}(0) = m^{(d)}(0)$ and $G_\ell^{(d)}(1) = m^{(d)}(1)$ for all $0 \leq d \leq \ell$. Indeed, from the derived form of f_ℓ, we find that $G_\ell(r)$ is an approximation from below to $m(r)$, and it is an especially good approximation near the endpoints $r = 0$ and $r = 1$. We show some examples of this in Fig. A.1.

We will call $G_\ell(r)$ the ℓ^{th} order approximation to $m(r)$. From Eq. (A.33), $\lim_{n\to\infty} \mathbb{E}_U \frac{1}{n} S_2(U) = \sum_{\ell=1}^{\infty} \frac{t^{2\ell}}{2\ell} G_\ell(r)$. Thus, $t = \tanh(2s)$ is weighting how relevant each approximation is. For small squeezing, most of the weight is concentrated on low-order approximations. The lowest order approximation is $G_1(r) = r(1-r)$ resulting in a parabolic shaped Page curve. When the squeezing is very large, more and more weight is placed on high-order approximations so that the Page curve begins to resemble the triangle $G_\infty(r) = m(r)$. We see a manifestation

of this interpretation as

$$\lim_{s\to 0}\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)}\frac{1}{s^2 n}S_2(U) = 2r(1-r), \tag{A.64}$$

$$\lim_{s\to \infty}\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)}\frac{1}{sn}S_2(U) = 2\min(r,1-r), \tag{A.65}$$

where the latter comes from the full expression in Theorem 2.3.1. Meanwhile, the maximal Rényi-2 entropy is $\max_U \frac{1}{n}S_2(U) = m(r)\log\cosh(2s)$ from Eq. (A.34). As stated, near the endpoints $r = 0$ and $r = 1$, $G_\ell(r)$ is a very good approximation to $m(r)$. Thus, regardless of the squeezing strength, when the subsystem size $k = rn$ is small (or when its complement is small), the average entanglement is very close to maximal.

Subsection A.3.2: Maximum value — Proof of Corollary 2.3.2

In Appendix A.3.1, we derived the exact formula for the Rényi-2 Page curve as an infinite series. Here we will show that the series can be completely simplified when $r = 1/2$. Bailey's theorem says that [141–144]

$${}_2F_1(a, 1-a; c; 1/2) = \frac{\Gamma(c/2)\Gamma((1+c)/2)}{\Gamma((a+c)/2)\Gamma((1+c-a)/2)}. \tag{A.66}$$

Plugging this into the Page curve in Theorem 2.3.1 at $r = 1/2$ and simplifying with the duplication formula $\Gamma(\ell+1/2) = \frac{\sqrt{\pi}(2\ell-1)!}{(\ell-1)!2^{2\ell-1}}$ [294],

$$\lim_{n\to\infty}\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)}\frac{1}{n}S_2(U) = \frac{1}{2}\log\cosh(2s) - \frac{1}{4}\sum_{\ell=1}^{\infty}\left(\tanh^2(2s)/4\right)^\ell \frac{1}{\ell}\binom{2\ell}{\ell}. \tag{A.67}$$

We find that the second term is $\frac{1}{4}\int_0^{t/4}\frac{f(x)}{x}\,\mathrm{d}x$, where $t = \tanh^2(2s)$ and $f(x)$ is the generating function of the central binomial coefficients $f(x) = \sum_{\ell=1}^{\infty}\binom{2\ell}{\ell}x^\ell$. Via the generalized binomial theorem, one finds the generating function evaluates to $f(x) = (1-4x)^{-1/2}-1$ [292, A000984].

Performing the integral, we find the second term to be $\frac{1}{4}\log\left(\frac{4-4\sqrt{1-t}}{t+t\sqrt{1-t}}\right)$, which simplifies to $\frac{1}{2}\log\cosh(2s) - \log\cosh s$. Subtracting the second term from the first yields $\log\cosh s$ as desired.

Finally, using Eq. (A.34) at $r = 1/2$, we find the Page correction to be $\frac{1}{2}\log\cosh(2s) - \log\cosh s$, which simplifies to $\frac{1}{2}\log(1+\tanh^2 s)$.

Subsection A.3.3: Constant term — Proof of Proposition 2.3.3

In Appendix A.3.1, we found that when all the initial squeezers are equal to s, $\mathbb{E}_U S_2(U) = n\alpha(s,r) - \lambda(s,r) + o(1)$, and we explicitly computed $\alpha(s,r)$. In this section, we determine $\lambda(s,r)$ up to a set of constants and conjecture an explicit value of the constants.

In Appendix A.3.1, we found that, asymptotically in n, $\mathbb{E}_U \operatorname{Tr} W^\ell = nf_\ell(r) + g_\ell(r) + o(1)$, where f_ℓ and g_ℓ are functions of r. From Eq. (A.32), it follows that $\lambda(s,r) = \sum_{\ell=1}^{\infty} \frac{1}{2\ell}\tanh^{2\ell}(2s) g_\ell(r)$. Furthermore, in Lemma A.3.1, we found that f_ℓ takes the form $f_\ell(r) = \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d$. Indeed, the proof of Lemma A.3.1 applies almost identically to $g_\ell(r)$, except that because we are interested now in the constant term in $\mathbb{E}_U \operatorname{Tr} W^\ell$ instead of the term linear in n, g_ℓ takes the form $g_\ell(r) = \sum_{d=\ell}^{2\ell} \beta_d^{(\ell)} r^d$. The extra term $\propto r^\ell$ in the polynomial g_ℓ does not occur in f_ℓ.

The Rényi-2 entropy must be symmetric under $r \mapsto 1-r$ at every order in s and n, since the full state on the n modes is pure. Therefore, $\lambda(s,r)$ must be symmetric under $r \mapsto 1-r$ at every order in $\tanh(2s)$, meaning that $g_\ell(r) = g_\ell(1-r)$. The following lemma will therefore be useful here, and we will also find use of it in Appendix A.4.

Lemma A.3.3. *Fix a polynomial $g_\ell(r) = \sum_{d=\ell}^{2\ell} \beta_d^{(\ell)} r^d$. If $g_\ell(r) = g_\ell(1-r)$, then $g_\ell(r) = \beta_{2\ell}^{(\ell)} (-1)^\ell \left(r(1-r)\right)^\ell$.*

Proof. Simplifying $g_l(r) - g_l(1-r) = 0$ with the binomial theorem, we find

$$0 = \sum_{d=\ell}^{2\ell} \beta_d^{(\ell)} \left(r^d - (1-r)^d\right) \tag{A.68}$$

$$= -\sum_{d=0}^{\ell-1} (-r)^d \sum_{j=\ell}^{2\ell} \beta_j^{(\ell)} \binom{j}{d} + \sum_{d=\ell}^{2\ell} r^d \left(\beta_d^{(\ell)} - (-1)^d \sum_{j=d}^{2\ell} \beta_j^{(\ell)} \binom{j}{d}\right). \tag{A.69}$$

Equating all degrees of r, we find

1. For $0 \leq d \leq \ell - 1$, $0 = \sum_{j=\ell}^{2\ell} \beta_j^{(\ell)} \binom{j}{d}$;
2. For $\ell \leq d \leq 2\ell$, $\beta_d^{(\ell)} = (-1)^d \sum_{j=d}^{2\ell} \beta_j^{(\ell)} \binom{j}{d}$.

Condition 1 is a system of ℓ linearly independent equations. To verify this, one must show that $\det D \neq 0$, where D is the $(\ell+1) \times (\ell+1)$ matrix with entries $D_{ij} = \binom{i+\ell-1}{j-1}$ for $1 \leq i \leq \ell+1$ and $1 \leq j \leq \ell$, and entries $D_{i,\ell+1} = \delta_{i,\ell+1}$. In other words, the rightmost column is all zeros except for the entry on the diagonal. Inserting this rightmost column is equivalent to adding one new equation to the linear system, where this new equation simply fixes the value of one of the variables. Since the rightmost column is all zeros except for the $(\ell+1, \ell+1)$ entry, we use Laplace's expansion to find that $\det D = \det D'$, where D' is the $\ell \times \ell$ upper left submatrix of D. One can prove that $\det D' \neq 0$ in a similar way as shown in Ref. [293]. Alternatively, one can use Corollary 11 from Ref. [295], which shows that the matrix with entries $\binom{a_i}{j-1}$ has zero determinant if and only if there are indices $i \neq j$ such that $a_i = a_j$. For the matrix D', $a_i = \ell + i - 1$, and therefore the determinant is nonzero.

In summary, we have a system of ℓ linearly independent equations for $\ell + 1$ variables. We can therefore uniquely express the solution by fixing one of the variables. Suppose we know the value of $\beta_{2\ell}^{(\ell)}$. Then one can verify that conditions 1 and 2 are satisfied by $\beta_{\ell+i}^{(\ell)} = (-1)^{\ell+i} \binom{\ell}{i} \beta_{2\ell}^{(\ell)}$, for $0 \leq i \leq \ell$. With the binomial theorem, this simply becomes $g_\ell(r) = \beta_{2\ell}^{(\ell)} (-1)^\ell \left(r(1-r)\right)^\ell$.

When verifying that the two conditions are satisfied by $\beta_{\ell+i}^{(\ell)} = (-1)^{\ell+i} \binom{\ell}{i} \beta_{2\ell}^{(\ell)}$, one finds

that the right hand side of both conditions can be simplified in terms of the hypergeometric function via Eq. (A.55). For example, condition 2, written as $(-1)^d = \sum_{j=d}^{2\ell} \binom{j}{d} (\beta_j^{(\ell)}/\beta_d^{(\ell)})$, reduces to $(-1)^d = {}_2F_1(d-2\ell, d+1; d-\ell+1; 1)$. As in the proof of Lemma A.3.2, we find that

$$
\begin{aligned}
{}_2F_1(d-2\ell, d+1; d-\ell+1; 1) &= \frac{(-\ell)_{2\ell-d}}{(d-\ell+1)_{2\ell-d}} && \text{(A.70)}\\
&= \frac{\Gamma(d-\ell+1)\Gamma(\ell-d)}{\Gamma(-\ell)\Gamma(\ell+1)} && \text{(A.71)}\\
&= \frac{\Gamma(d-\ell+1)\Gamma(-\ell)(\ell-d-1)(\ell-d-2)\dots(-\ell)}{\Gamma(-\ell)\Gamma(\ell+1)} && \text{(A.72)}\\
&= (-1)^{1+(\ell-d+1)-(-\ell)} \frac{\Gamma(d-\ell+1)(d+1-\ell)(d+2-\ell)\dots(\ell)}{\Gamma(\ell+1)} && \text{(A.73)}\\
&= (-1)^d \frac{\Gamma(d-\ell+1)\ell!}{\Gamma(\ell+1)(d-\ell)!} && \text{(A.74)}\\
&= (-1)^d, && \text{(A.75)}
\end{aligned}
$$

which proves condition 2. The proof of condition 1 is similar. □

For convenience, we define $a^{(\ell)} := \beta_{2\ell}^{(\ell)}$. It follows from Lemma A.3.3 that

$$
\lambda(s,r) = \sum_{\ell=1}^{\infty} \frac{(-1)^\ell}{2\ell} a^{(\ell)} \tanh^{2\ell}(2s) \left(r(1-r)\right)^\ell . \tag{A.76}
$$

Note that if $a^{(\ell)} = (-1)^\ell 4^{\ell-1}$, then the Taylor series of log implies

$$
\lambda(s,r) = -\frac{1}{8} \log\bigl(1 - 4r(1-r)\tanh^2(2s)\bigr) , \tag{A.77}
$$

which is precisely Proposition 2.3.3. Hence, the only thing left to prove is the following lemma.

Lemma A.3.4. *For all $\ell \in \mathbb{N}$, $a^{(\ell)} = (-1)^\ell 4^{\ell-1}$.*

To make progress on Lemma A.3.4, we write the formula for $a^{(\ell)}$. Recall that $\beta_{2\ell}^{(\ell)}$ is the coefficient in front of the $r^{2\ell}$ term in $g_\ell(r)$. Looking at Eq. (A.44), we see that the only way to get an $r^{2\ell}$ term (i.e. a $k^{2\ell}$ term) is if σ is the permutation $\sigma(1) = 2\ell$, $\sigma(2) = 1$, $\sigma(3) = 2$, $\ldots$, $\sigma(2\ell) = 2\ell - 1$. Therefore, we are interested in the constant term (the term asymptotically independent of n) in the following term of Eq. (A.44):

$$\begin{aligned} & n^{2\ell} \sum_{j_1,\ldots,j_{2\ell}=1}^{n} \sum_{j'_1,\ldots,j'_{2\ell}=1}^{n} \sum_{\tau \in S_{2\ell}} \mathrm{Wg}(\sigma\tau^{-1}, n) \\ & \qquad \times \delta_{j_1,j_2}\delta_{j'_1,j'_2} \ldots \delta_{j_{2\ell-1},j_{2\ell}}\delta_{j'_{2\ell-1},j'_{2\ell}} \\ & \qquad \times \delta_{j_1,j'_{\tau(1)}} \ldots \delta_{j_{2\ell},j'_{\tau(2\ell)}}, \end{aligned} \tag{A.78}$$

which simplifies to

$$\begin{aligned} & = n^{2\ell} \sum_{j_1,\ldots,j_{2\ell}=1}^{n} \sum_{j'_1,\ldots,j'_{2\ell}=1}^{n} \sum_{\tau \in S_{2\ell}} \mathrm{Wg}(\tau^{-1}, n) \\ & \qquad \times \delta_{j_1,j_2}\delta_{j'_1,j'_2} \ldots \delta_{j_{2\ell-1},j_{2\ell}}\delta_{j'_{2\ell-1},j'_{2\ell}} \\ & \qquad \times \delta_{j_1,j'_{\tau(2\ell)}}\delta_{j_2,j'_{\tau(1)}}\delta_{j_3,j'_{\tau(2)}} \ldots \delta_{j_{2\ell},j'_{\tau(2\ell-1)}} \end{aligned} \tag{A.79}$$

$$\begin{aligned} & = n^{2\ell} \sum_{j_2,j_4,\ldots,j_{2\ell}=1}^{n} \sum_{j'_1,\ldots,j'_{2\ell}=1}^{n} \sum_{\tau \in S_{2\ell}} \mathrm{Wg}(\tau, n) \\ & \qquad \times \delta_{j'_1,j'_2} \ldots \delta_{j'_{2\ell-1},j'_{2\ell}} \\ & \qquad \times \delta_{j_2,j'_{\tau(2\ell)}}\delta_{j_2,j'_{\tau(1)}}\delta_{j_4,j'_{\tau(2)}}\delta_{j_4,j'_{\tau(3)}} \ldots \delta_{j_{2\ell},j'_{\tau(2\ell-2)}}\delta_{j_{2\ell},j'_{\tau(2\ell-1)}} \end{aligned} \tag{A.80}$$

$$\begin{aligned} & = n^{2\ell} \sum_{j'_1,\ldots,j'_{2\ell}=1}^{n} \sum_{\tau \in S_{2\ell}} \mathrm{Wg}(\tau, n) \\ & \qquad \times \delta_{j'_1,j'_2} \ldots \delta_{j'_{2\ell-1},j'_{2\ell}} \\ & \qquad \times \delta_{j'_{\tau(2\ell)},j'_{\tau(1)}}\delta_{j'_{\tau(2)},j'_{\tau(3)}} \ldots \delta_{j'_{\tau(2\ell-2)},j'_{\tau(2\ell-1)}}. \end{aligned} \tag{A.81}$$

Note we used that $\mathrm{Wg}(\tau, n) = \mathrm{Wg}(\tau^{-1}, n)$ from Eq. (A.7). Let $\#(\tau)$ be the number of cycles in the disjoint cycle decomposition of τ and let $\{c_i^{(\tau)} \mid i \in \{1, \ldots, \#(\tau)\}\}$ be the cycle

ℓ	$a^{(\ell)}$	
1	$-C_1$	$= -1$
2	$-4C_0^2C_1 + 4C_0C_2$	$= 4$
3	$-9C_0^4C_1 + 15C_0^2C_1^2 - C_1^3 + 18C_0^3C_2 - 6C_0C_1C_2 - 9C_0^2C_3$	$= -16$
4	$-16C_0^6C_1 + 80C_0^4C_1^2 - 40C_0^2C_1^3 + 48C_0^5C_2 - 112C_0^3C_1C_2$ $+16C_0C_1^2C_2 + 8C_0^2C_2^2 - 48C_0^4C_3 + 24C_0^2C_1C_3 + 16C_0^3C_4$	$= 64$
5	$-25C_0^8C_1 + 250C_0^6C_1^2 - 380C_0^4C_1^3 + 80C_0^2C_1^4 - C_1^5 + 100C_0^7C_2$ $-600C_0^5C_1C_2 + 440C_0^3C_1^2C_2 - 20C_0C_1^3C_2 + 130C_0^4C_2^2$ $-50C_0^2C_1C_2^2 - 150C_0^6C_3 + 320C_0^4C_1C_3 - 60C_0^2C_1^2C_3 - 40C_0^3C_2C_3$ $+100C_0^5C_4 - 60C_0^3C_1C_4 - 25C_0^4C_5$	$= -256$

Table A.1: A table showing the first five values of $a^{(\ell)}$ from Eq. (A.88), which matches Lemma A.3.4.

decomposition. Then, from Eq. (A.7), asymptotically,

$$\mathrm{Wg}(\tau, n) = \frac{1}{n^{2\ell+|\tau|}} \prod_{i=1}^{\#(\tau)} (-1)^{|c_i^{(\tau)}|-1} C_{|c_i^{(\tau)}|-1}, \tag{A.82}$$

where recall C_m is the m^{th} Catalan number. Therefore, $a^{(\ell)}$ is the constant term independent of n in

$$\begin{aligned} \sum_{j_1,\dots,j_{2\ell}=1}^{n} \sum_{\tau \in S_{2\ell}} \frac{1}{n^{|\tau|}} \prod_{i=1}^{\#(\tau)} (-1)^{|c_i^{(\tau)}|-1} C_{|c_i^{(\tau)}|-1} \\ \times \delta_{j_1,j_2} \dots \delta_{j_{2\ell-1},j_{2\ell}} \\ \times \delta_{j_{\tau(2\ell)},j_{\tau(1)}} \delta_{j_{\tau(2)},j_{\tau(3)}} \dots \delta_{j_{\tau(2\ell-2)},j_{\tau(2\ell-1)}}. \end{aligned} \tag{A.83}$$

Define $\xi \colon S_{2\ell} \to \mathbb{N}$ such that

$$n^{\xi(\tau)} = \sum_{j_1,\dots,j_{2\ell}=1}^{n} \delta_{j_1,j_2}\delta_{j_3,j_4} \dots \delta_{j_{2\ell-1},j_{2\ell}} \times \delta_{j_{\tau(2\ell)},j_{\tau(1)}} \delta_{j_{\tau(2)},j_{\tau(3)}} \dots \delta_{j_{\tau(2\ell-2)},j_{\tau(2\ell-1)}} \tag{A.84}$$

$$= \sum_{j_1,\dots,j_{\ell}=1}^{n} \delta_{j_{\lceil \tau(2\ell)/2 \rceil},j_{\lceil \tau(1)/2 \rceil}} \delta_{j_{\lceil \tau(2)/2 \rceil},j_{\lceil \tau(3)/2 \rceil}} \dots \delta_{j_{\lceil \tau(2\ell-2)/2 \rceil},j_{\lceil \tau(2\ell-1)/2 \rceil}}. \tag{A.85}$$

Note that the definition of ξ is independent of the value of n, and so we can equivalently

define ξ as

$$\xi(\tau) = \log_2 \sum_{j_1,\dots,j_\ell=1}^{2} \delta_{j_{\lceil \tau(2\ell)/2 \rceil}, j_{\lceil \tau(1)/2 \rceil}} \delta_{j_{\lceil \tau(2)/2 \rceil}, j_{\lceil \tau(3)/2 \rceil}} \cdots \delta_{j_{\lceil \tau(2\ell-2)/2 \rceil}, j_{\lceil \tau(2\ell-1)/2 \rceil}}. \tag{A.86}$$

To get the constant term, we need $\frac{n^{\xi(\tau)}}{n^{|\tau|}} = 1$. Therefore,

$$a^{(\ell)} = \sum_{\substack{\tau \in S_{2\ell} \text{ s.t.} \\ \xi(\tau) = |\tau|}} \prod_{i=1}^{\#(\tau)} (-1)^{|c_i^{(\tau)}|-1} C_{|c_i^{(\tau)}|-1}. \tag{A.87}$$

Finally, since the sum of the lengths of the cycles of a permutation $\tau \in S_{2\ell}$ is always 2ℓ, we find

$$a^{(\ell)} = \sum_{\substack{\tau \in S_{2\ell} \text{ s.t.} \\ \xi(\tau) = |\tau|}} (-1)^{\#(\tau)} \prod_{i=1}^{\#(\tau)} C_{|c_i^{(\tau)}|-1}. \tag{A.88}$$

From this equation, we can exactly compute $a^{(\ell)}$ on a computer for small values of ℓ. Table A.1 shows the first five of these, which all match Lemma A.3.4 saying that $a^{(\ell)} = (-1)^\ell 4^{\ell-1}$. Note that if we change the condition of $\xi(\tau) = |\tau|$ in Eq. (A.88) to $\xi(\tau) = |\tau|+1$, then this gives us the term that is linear in n and hence is the equation for $\alpha_{2\ell}^{(\ell)}$ from Lemma A.3.2, which is $\frac{(-1)^{\ell+1}}{2\ell-1}\binom{2\ell-1}{\ell-1}$.

To complete the proof of Lemma A.3.4, we need to evaluate Eq. (A.88) for all $\ell \in \mathbb{N}$. This is done in Ref. [149] using objects called *breakpoint graphs* that arise in the study of gene orders in bioinformatics [148]. Roughly, ξ has an interpretation in terms of cycles of breakpoint graphs.

Section A.4: Variance of the Rényi-2 entropy — Proof of Theorem 2.4.2

In this section, we shift our attention away from $\mathbb{E}_U S_2(U)$ and instead to $\mathrm{Var}_U S_2(U) = \mathbb{E}_U S_2(U)^2 - (\mathbb{E}_U S_2(U))^2$, and we will prove Theorem 2.4.2. We are again interested in the case where all the initial squeezers are equal to s. Using Eq. (A.32), this becomes

$$\operatorname*{Var}_{U\in\mathrm{U}(n)} S_2(U) = \sum_{\ell,\ell'=1}^{\infty} \frac{1}{(2\ell)(2\ell')} \tanh^{2\ell+2\ell'}(2s) \left(\mathop{\mathbb{E}}_U (\operatorname{Tr} W^\ell)(\operatorname{Tr} W^{\ell'}) - \left(\mathop{\mathbb{E}}_U \operatorname{Tr} W^\ell\right)\left(\mathop{\mathbb{E}}_U \operatorname{Tr} W^{\ell'}\right)\right) \tag{A.89}$$

$$= \sum_{d=2}^{\infty} \tanh^{2d}(2s) \sum_{\ell=1}^{d-1} \frac{1}{4\ell(d-\ell)} \left(\mathop{\mathbb{E}}_U (\operatorname{Tr} W^\ell)(\operatorname{Tr} W^{d-\ell}) - \left(\mathop{\mathbb{E}}_U \operatorname{Tr} W^\ell\right)\left(\mathop{\mathbb{E}}_U \operatorname{Tr} W^{d-\ell}\right)\right). \tag{A.90}$$

As a direct consequence of Lemma A.3.1, we find that asymptotically

$$\left(\mathop{\mathbb{E}}_U \operatorname{Tr} W^\ell\right)\left(\mathop{\mathbb{E}}_U \operatorname{Tr} W^{\ell'}\right) = n^2 p_{\ell,\ell'}(r) + n q_{\ell,\ell'}(r) + t_{\ell,\ell'}(r) + o(1)\,, \tag{A.91}$$

where $p_{\ell,\ell'}$ is a polynomial of degrees $\ell+\ell'+2$ through $2\ell+2\ell'$ in r, $q_{\ell,\ell'}$ is a polynomial of degrees $\ell+\ell'+1$ through $2\ell+2\ell'$ in r, and $t_{\ell,\ell'}$ is a polynomial of degrees $\ell+\ell'$ through $2\ell+2\ell'$ in r.

Furthermore, we find an analogous result for $\mathbb{E}_U(\operatorname{Tr} W^\ell)(\operatorname{Tr} W^{\ell'})$. Let $L \coloneqq 2\ell+2\ell'$. Using

Eq. (A.6), we find that

$$\begin{aligned}\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)}(\operatorname{Tr} W^{\ell})(\operatorname{Tr} W^{\ell'}) = \sum_{i_1,\dots i_L=1}^{k}\sum_{i'_1,\dots i'_L=1}^{k}\sum_{j_1,\dots j_L=1}^{n}\sum_{j'_1,\dots j'_L=1}^{n}\sum_{\sigma,\tau\in S_L}\mathrm{Wg}(\sigma\tau^{-1},n)\\ \times \delta_{i'_{2\ell},i_1}\delta_{i'_1,i_2}\delta_{i'_2,i_3}\dots\delta_{i'_{2\ell-1},i_{2\ell}}\\ \times \delta_{i'_L,i_{2\ell+1}}\delta_{i'_{2\ell+1},i_{2\ell+2}}\delta_{i'_{2\ell+2},i_{2\ell+3}}\dots\delta_{i'_{L-1},i_L}\\ \times \delta_{j_1,j_2}\delta_{j'_1,j'_2}\dots\delta_{j_{2\ell-1},j_{2\ell}}\delta_{j'_{2\ell-1},j'_{2\ell}}\\ \times \delta_{j_{2\ell+1},j_{2\ell+2}}\delta_{j'_{2\ell+1},j'_{2\ell+2}}\dots\delta_{j_{L-1},j_L}\delta_{j'_{L-1},j'_L}\\ \times \delta_{i_1,i'_{\sigma(1)}}\dots\delta_{i_L,i'_{\sigma(L)}}\\ \times \delta_{j_1,j'_{\tau(1)}}\dots\delta_{j_L,j'_{\tau(L)}},\end{aligned} \tag{A.92}$$

and the asymptotic form of Wg function is given in Eq. (A.7). In a similar proof to Lemma A.3.1 but with Eq. (A.92) instead of Eq. (A.44), we analogously find that asymptotically

$$\mathop{\mathbb{E}}_{U\in \mathrm{U}(n)}(\operatorname{Tr} W^{\ell})(\operatorname{Tr} W^{\ell'}) = n^2 P_{\ell,\ell'}(r) + nQ_{\ell,\ell'}(r) + T_{\ell,\ell'}(r) + o(1)\,, \tag{A.93}$$

where $P_{\ell,\ell'}$ is a polynomial of degrees $\ell+\ell'+2$ through $2\ell+2\ell'$ in r, $Q_{\ell,\ell'}$ is a polynomial of degrees $\ell+\ell'+1$ through $2\ell+2\ell'$ in r, and $T_{\ell,\ell'}$ is a polynomial of degrees $\ell+\ell'$ through $2\ell+2\ell'$ in r. For completeness, we prove this result for the $P_{\ell,\ell'}$ term in the following lemma. The proofs for the $Q_{\ell,\ell'}$ and $T_{\ell,\ell'}$ terms follow from trivially tweaking the final part of the proof.

Lemma A.4.1. *Fix positive integers ℓ and ℓ'. There exist coefficients $\beta_d^{(\ell,\ell')}$ for $d\in\{\ell+\ell'+2,\dots,2\ell+2\ell'\}$ such that*

$$P_{\ell,\ell'}(r) \coloneqq \lim_{n\to\infty}\frac{1}{n^2}\mathop{\mathbb{E}}_{U}(\operatorname{Tr} W^{\ell})(\operatorname{Tr} W^{\ell'}) = \sum_{d=\ell+\ell'+2}^{2\ell+2\ell'}\beta_d^{(\ell,\ell')}r^d. \tag{A.94}$$

Proof. Much of the details of this proof are the same as in the proof of Lemma A.3.1. We will use the asymptotic form of the Wg function, which is written in Eq. (A.7). The proof will

proceed as follows. First, we will prove that $(\operatorname{Tr} W^\ell)(\operatorname{Tr} W^{\ell'})$ contains a term proportional to n^2 and no terms proportional to n^a for any $a > 2$. Therefore, $P_{\ell,\ell'}(r)$ is indeed independent of n. Next, we will prove that $P_{\ell,\ell'}(r)$ has no terms r^a for $a > L$ and $a \leq \ell + \ell' + 1$. Throughout this proof, we interpret the delta functions in Eq. (A.92) as constraints on the summations. Different permutations on the indices result in a different number of constraints and hence terms with different powers of n and k.

Recall that Π can only decrease the trace. Getting rid of the Π in W and using the cyclic nature of the trace, we see that $(\operatorname{Tr} W^\ell)(\operatorname{Tr} W^{\ell'}) \leq n^2$. Furthermore, consider Eq. (A.92) with $\sigma = \tau$ defined by $\sigma(1) = 2\ell$, $\sigma(2) = 1$, $\ldots$, $\sigma(2\ell) = 2\ell - 1$ and $\sigma(2\ell + 1) = L$, $\sigma(2\ell + 2) = 2\ell + 1$, $\ldots$, $\sigma(L) = L - 1$. Then, $\sigma\tau^{-1}$ is the identity, and so $\operatorname{Wg}(\sigma\tau^{-1}, n)$ contributes a factor of n^{-L}. With this σ, the sum over the i and i' yields a factor of k^L. Finally, the sum over j yields

$$\delta_{j'_1,j'_2}\delta_{j'_3,j'_4}\ldots\delta_{j'_{L-1},j'_L}\delta_{j'_1,j'_{2\ell}}\delta_{j'_2,j'_1}\ldots\delta_{j'_{2\ell},j'_{2\ell-1}}\delta_{j'_{2\ell+1},j'_L}\delta_{j'_{2\ell+2},j'_{2\ell+1}}\ldots\delta_{j'_L,j'_{L-1}}. \tag{A.95}$$

Then summing over j', we get two factors of n. Hence, the term with the specific permutation described above yields a term of the form $n^2 k^L n^{-L} = n^2 r^L$.

We have shown that there is a term proportional to n^2 and that there are no terms proportional to n^a for $a > 2$. Since we are working asymptotically in n, we can therefore ignore all terms proportional to n^a for every $a < 2$. This proves that $\lim_{n\to\infty} \frac{1}{n^2}(\operatorname{Tr} W^\ell)(\operatorname{Tr} W^{\ell'})$ is independent of n and only depends on r, which justifies the definition of the function $P_{\ell,\ell'}(r)$. The only thing left to show is that $P_{\ell,\ell'}(r)$ has only terms $r^{\ell+\ell'+2}$ through $r^{2\ell+2\ell'} = r^L$. So we only need to show that there are no terms r^a for $a > L$ and $a \leq \ell + \ell' + 1$. We begin with the former.

To look at powers of r, it is sufficient to look at powers of k. We therefore restrict our attention to the sum over i and i' in Eq. (A.92). In order to get the highest power of k, we require the least constraints on i and i' (i.e. the least distinct Kronecker deltas). We

therefore require σ to be the permutation so that

$$\delta_{i'_{2\ell},i_1}\delta_{i'_1,i_2}\delta_{i'_2,i_3}\dots\delta_{i'_{2\ell-1},i_{2\ell}}\delta_{i'_L,i_{2\ell+1}}\delta_{i'_{2\ell+1},i_{2\ell+2}}\delta_{i'_{2\ell+2},i_{2\ell+3}}\dots\delta_{i'_{L-1},i_L} = \delta_{i_1,i'_{\sigma(1)}}\dots\delta_{i_L,i'_{\sigma(L)}}. \quad \text{(A.96)}$$

This permutation σ is exactly the σ described above that gave the term proportional to $n^2 r^L$. Hence, L is the highest power of r that can be achieved.

Next we need to show that $\ell + \ell' + 2$ is the lowest power of k that can be achieved. The sum over j and j' can give at most a factor of $n^{\ell+\ell'}$. This is because the first line of delta functions, $\delta_{j_1,j_2}\delta_{j'_1,j'_2}\dots\delta_{j_{L-1},j_L}\delta_{j'_{L-1},j'_L}$, reduces the sum over the L indices j and the L indices j' down to just a sum over $\ell + \ell'$ indices j and $\ell + \ell'$ indices j'. The second line of delta functions, $\delta_{j_1,j'_{\tau(1)}}\dots\delta_{j_L,j'_{\tau(L)}}$, cannot be made equivalent to the first line by any choice of τ; in fact, the second line imposes all new constraints. Therefore, this line further reduces the sum over the $\ell+\ell'$ indices j and the $\ell+\ell'$ indices j' to just a sum over the $\ell+\ell'$ indices j (or the $\ell + \ell'$ indices j', but not both). Hence the highest power of n that we can get from the summations over the j and j' is $n^{\ell+\ell'}$. Putting this together with the fact that asymptotically $\mathrm{Wg}(\pi, n)$ is at most $n^{-L} = n^{-2\ell-2\ell'}$, we find that any term coming from Eq. (A.92) is at most $n^{-\ell-\ell'} \times$ (dependence on k). Therefore, any powers of k that are less than $\ell + \ell' + 2$ can be ignored; if the sum over i and i' yields a term that is k^a for some $a \leq \ell + \ell' + 1$, then that term will scale linearly or less with n. But from above, we already have terms that are proportional to n^2, and so terms that are linear or less in n can be ignored. □

Therefore, from Eq. (A.90), we find that asymptotically

$$\operatorname*{Var}_{U\in\mathrm{U}(n)} S_2(U) = \sum_{d=2}^{\infty} \tanh^{2d}(2s)\left(n^2 p_d(r) + n q_d(r) + t_d(r)\right) + o(1)\,, \quad \text{(A.97)}$$

where p_d is a polynomial of degrees $d+2$ through $2d$ in r, q_d is a polynomial of degrees $d+1$ through $2d$ in r, and t_d is a polynomial of degrees d through $2d$ in r. The variance must be symmetric under $r \mapsto 1 - r$ at every order in s and n. Therefore each p_d, q_d, and t_d must

themselves be symmetric under $r \mapsto 1-r$. It then immediately follows from Lemma A.3.3 that $t_d(r) = \omega^{(d)} (r(1-r))^d$ for some constants $\omega^{(d)}$. In the following lemma, we show that p_d and q_d must be the zero polynomial.

Lemma A.4.2. *Let d be a positive integer, and $f(r) = \sum_{j=d+1}^{2d} \gamma_j r^j$. If $f(r) = f(1-r)$, then $f(r) = 0$.*

Proof. Expanding with the binomial theorem, we find

$$
\begin{aligned}
0 &= f(r) - f(1-r) && \text{(A.98)}\\
&= \sum_{j=d+1}^{2d} \gamma_j \left(r^j - (1-r)^j\right) && \text{(A.99)}\\
&= \sum_{j=d+1}^{2d} \gamma_j r^j - \sum_{j=d+1}^{2d} \gamma_j \sum_{i=0}^{j} (-r)^i \binom{j}{i} && \text{(A.100)}\\
&= \sum_{j=d+1}^{2d} \gamma_j r^j - \sum_{i=0}^{2d} (-r)^i \sum_{j=\max(i,d+1)}^{2d} \gamma_j \binom{j}{i} && \text{(A.101)}\\
&= -\sum_{j=0}^{d} (-r)^j \sum_{i=d+1}^{2d} \gamma_i \binom{i}{j} + \sum_{j=d+1}^{2d} r^j \left(\gamma_j - (-1)^j \sum_{i=j}^{2d} \gamma_i \binom{i}{j}\right). && \text{(A.102)}
\end{aligned}
$$

This must be true to all orders in r, and hence we equate each degree of r to zero. This gives us the conditions:

1. For $0 \le j \le d$, $0 = \sum_{i=d+1}^{2d} \gamma_i \binom{i}{j}$;
2. For $d+1 \le j \le 2d$, $\gamma_j = (-1)^j \sum_{i=j}^{2d} \gamma_i \binom{i}{j}$.

Choosing any d of the equations from the first condition gives a linear system of d linearly independent equations. To verify this, one must show that $\det C \neq 0$, where C is the $d \times d$ matrix with entries $C_{ij} = \binom{d+i}{j}$. This was shown in the proof of Lemma A.3.2.

Therefore, we have d linearly independent equations for d variables. Hence, if there is a solution, then there is one *unique* solution. We easily see that $\gamma_j = 0$ is a solution, and therefore it is the solution. This gives $f(r) = 0$, completing the proof. □

Since p_d and q_d are zero, we have therefore found that

$$\lim_{n\to\infty} \operatorname*{Var}_{U\in \mathrm{U}(n)} S_2(U) = \sum_{d=2}^{\infty} \omega^{(d)} \tanh^{2d}(2s) \left(r(1-r)\right)^d , \tag{A.103}$$

proving the first part of Theorem 2.4.2. Equating this equation to Eq. (A.90), we find that

$$\omega^{(d)} = \lim_{n\to\infty} \left(r(1-r)\right)^{-d} \sum_{\ell=1}^{d-1} \frac{1}{4\ell(d-\ell)} \left(\mathop{\mathbb{E}}_{U} (\operatorname{Tr} W^\ell)(\operatorname{Tr} W^{d-\ell}) - \left(\mathop{\mathbb{E}}_{U} \operatorname{Tr} W^\ell \right) \left(\mathop{\mathbb{E}}_{U} \operatorname{Tr} W^{d-\ell} \right) \right) \tag{A.104}$$

$$= \lim_{n\to\infty} \left(r(1-r)\right)^{-d} \sum_{\ell=1}^{d-1} \frac{1}{4\ell(d-\ell)} \operatorname*{Cov}_{U\in \mathrm{U}(n)} \left(\operatorname{Tr} W^\ell, \operatorname{Tr} W^{d-\ell} \right) , \tag{A.105}$$

where Cov is the covariance, and we know that $\omega^{(d)}$ is independent of r and n. From our expressions for $\operatorname{Tr} W^\ell$ in terms of the Weingarten calculus, it follows that $\omega^{(d)} \in \mathbb{Q}$. Recall that W has two factors of U and two factors of $\bar{U}$. Hence, we can exactly compute $\omega^{(2)}$ by integrating over fourth moments of the Haar measure on the unitary group. To do this, we use the Mathematica package `RTNI` that symbolically computes expressions over the Haar measure [296]. This Mathematica package precomputes the symbolic expressions for $\mathrm{Wg}(\sigma, n)$ for $\sigma \in S_4$. From this, $\omega^{(2)}$ is a sum over powers of $\operatorname{Tr} \Pi = k$ with coefficients depending on n. One can then simplify this expression and take the limit $n \to \infty$ to find that $\omega^{(2)} = 1/2$. Our Mathematica code for this calculation is provided on GitHub [145].

Appendix B: Technical details for Chapter 3

Section B.1: von Neumann Entropy

In this section of the appendix, we show derivations of our results regarding the von Neumann entropy. We cover the derivation of the von Neumann entropy presented in Theorem 3.3.1, show squeezing limits as $s \to 0$ and $s \to \infty$, (Theorem 3.3.2 and Theorem 3.3.3), and discuss a potential avenue for solving the unequal squeezing case.

Subsection B.1.1: Page Curve for the von Neumann Entropy

Here, we derive the von Neumann entanglement entropy between the subsystems with k and $n-k$ modes, in terms of the partition size ratio $r = \frac{k}{n}$ and the squeezing strength s. We use much of the notation from Ref. [2], and the background on bosonic Gaussian states and operations can be found in Ref. [138].

Consider a system of n bosonic modes, where each mode $1 \leq i \leq n$ is prepared in a single-mode squeezed state with squeezing strength s_i. This n-mode Gaussian state is acted on by a passive linear optical unitary $U \in U(n)$.

We divide the output modes into two groups, with one group bearing $0 \leq k \leq n$ modes. Let σ be the $2k \times 2k$ covariance matrix of the reduced state on these k modes. As shown in Refs. [138, 152, 178], the von Neumann entanglement entropy $S_1(U)$ of the whole system is

given by

$$S_1(U) = \frac{1}{2}S_1^{(1)} + \frac{1}{2}S_1^{(2)} \quad \text{with} \quad \begin{cases} S_1^{(1)} = \log\det\left(\frac{\sigma+\mathrm{i}\Omega}{2}\right) \\ S_1^{(2)} = \mathrm{Tr}\left(\coth^{-1}(\mathrm{i}\Omega\sigma)\mathrm{i}\Omega\sigma\right) \end{cases}, \tag{B.1}$$

where $\coth^{-1}(x)$ is the inverse hyperbolic cotangent function, and the symplectic form Ω equals

$$\Omega := \begin{pmatrix} 0_{n\times n} & \mathbb{I}_{n\times n} \\ -\mathbb{I}_{n\times n} & 0_{n\times n} \end{pmatrix}. \tag{B.2}$$

We aim to derive a formula for S_1 solely in terms of r, n, and s using Eq. (B.1).

Define the $k \times n$ matrix P and the $n \times n$ projector Π such that

$$P := \begin{pmatrix} \mathbb{I}_{k\times k} & 0_{k\times(n-k)} \end{pmatrix}, \tag{B.3}$$

$$\Pi := \begin{pmatrix} \mathbb{I}_{k\times k} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & 0_{(n-k)\times(n-k)} \end{pmatrix} = \mathbb{I}_{k\times k} \oplus 0_{(n-k)\times(n-k)}. \tag{B.4}$$

From Ref. [2], when $s_i = s_j \; \forall i, j \in \{1, 2, \ldots, n\}$, the covariance matrix σ of the reduced state after the application of U can be expressed as

$$\sigma = \cosh(2s)\mathbb{I} + \sinh(2s)M, \quad \text{where} \quad M := \begin{pmatrix} P\,\mathrm{Re}(\bar{U}U^\dagger)\,P^T & P\,\mathrm{Im}(\bar{U}U^\dagger)\,P^T \\ P\,\mathrm{Im}(\bar{U}U^\dagger)\,P^T & -P\,\mathrm{Re}(\bar{U}U^\dagger)\,P^T \end{pmatrix}, \tag{B.5}$$

and $\bar{U}, U^T, U^\dagger$ represent the conjugate, transpose, and conjugate transpose of U, respectively. Two useful and easily checkable properties of M are that $\Omega M = -M\Omega$ and $\mathrm{Tr}(\Omega M^j) = \mathrm{Tr}(M^{2j-1}) = 0, \; \forall j \in \mathbb{Z}^+$.

Using Eq. (B.5) and the properties above, $S_1^{(1)}$ from Eq. (B.1) is expressed as

$$
\begin{aligned}
S_1^{(1)} &= \operatorname{Tr}\log\left(\frac{\sigma+\mathrm{i}\Omega}{2}\right) \qquad \text{(B.6)}\\
&= \sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}\operatorname{Tr}\left(\frac{\sigma+\mathrm{i}\Omega}{2}-\mathbb{I}\right)^{j}\\
&= \sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j2^j}\operatorname{Tr}\left((\cosh(2s)-2)\mathbb{I}+\sinh(2s)M+\mathrm{i}\Omega\right)^{j}\\
&= \operatorname{Tr}\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j2^j}\sum_{\ell=0}^{j}\binom{j}{\ell}\left(\sinh(2s)M+\mathrm{i}\Omega\right)^{\ell}\left(\cosh(2s)-2\right)^{j-\ell}\\
&= \sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j2^j}\left(2k(\cosh(2s)-2)^{j}+\operatorname{Tr}\sum_{\ell=1}^{j}\binom{j}{\ell}\left(\sinh(2s)M+\mathrm{i}\Omega\right)^{\ell}\left(\cosh(2s)-2\right)^{j-\ell}\right)\\
&= 2k\log\left(\frac{\cosh 2s}{2}\right)-\operatorname{Tr}\sum_{\ell=1}^{\infty}\left(\sinh(2s)M+\mathrm{i}\Omega\right)^{\ell}\sum_{j=\ell}^{\infty}\frac{(-1)^{j}}{j2^j}\binom{j}{\ell}\left(\cosh(2s)-2\right)^{j-\ell}\\
&= 2k\log\left(\frac{\cosh 2s}{2}\right)-\operatorname{Tr}\sum_{\ell=1}^{\infty}\frac{(-1)^{\ell}}{\ell}\left(\sinh(2s)M+\mathrm{i}\Omega\right)^{\ell}\cosh^{-\ell}(2s)\\
&= 2k\log\left(\frac{\cosh 2s}{2}\right)-\operatorname{Tr}\sum_{\ell=1}^{\infty}\frac{1}{2\ell}\left(\sinh(2s)M+\mathrm{i}\Omega\right)^{2\ell}\cosh^{-2\ell}(2s)\\
&= 2k\log\left(\frac{\cosh 2s}{2}\right)-\operatorname{Tr}\sum_{\ell=1}^{\infty}\frac{1}{2\ell}\cosh^{-2\ell}(2s)\left(\sinh^2(2s)M^2+\mathbb{I}\right)^{\ell}. \qquad \text{(B.7)}
\end{aligned}
$$

As discussed in Ref. [2], let $W=\Pi U\bar{U}^{\dagger}\Pi\bar{U}U^{\dagger}\Pi$. Then, $\operatorname{Tr}(M^{2i})=2\operatorname{Tr}(W^i)\ \forall i\in\mathbb{Z}^+$. We emphasize that the dependence on U and r is implicit in W. Importantly, though, there is no s dependence in W. We also note that, since now everything is written in terms of traces, we are able to replace P with Π so that all matrices are square. Incorporating this definition into Eq. (B.7), we have

$$
\begin{aligned}
S_1^{(1)} &= 2k\log\left(\frac{\cosh 2s}{2}\right)-2\sum_{\ell=1}^{\infty}\frac{1}{2\ell}\cosh^{-2\ell}(2s)\sum_{j=0}^{\ell}\binom{\ell}{j}\sinh^{2j}(2s)\operatorname{Tr}W^{j} \qquad \text{(B.8)}\\
&= 2k\log\left(\frac{\cosh 2s}{2}\right)-2\sum_{\ell=1}^{\infty}\frac{1}{2\ell}\cosh^{-2\ell}(2s)\left(k+\sum_{j=1}^{\ell}\binom{\ell}{j}\sinh^{2j}(2s)\operatorname{Tr}W^{j}\right)
\end{aligned}
$$

$$= 2k \log\left(\frac{\cosh 2s}{2}\right) + k \log\left[\tanh^2(2s)\right] - 2\sum_{j=1}^{\infty} \sinh^{2j}(2s) \operatorname{Tr} W^j \sum_{\ell=j}^{\infty} \frac{1}{2\ell \cosh^{2\ell}(2s)} \binom{\ell}{j}$$

$$= k \log\left[\cosh^2(s) \sinh^2(s)\right] - 2\sum_{j=1}^{\infty} \sinh^{2j}(2s) \operatorname{Tr} W^j \frac{1}{2j \sinh(2s)^{2j}}$$

$$= k \log\left[\frac{\sinh^2(2s)}{4}\right] - \sum_{j=1}^{\infty} \frac{1}{j} \operatorname{Tr} W^j. \tag{B.9}$$

We use a similar approach to simplify $S_1^{(2)}$. Notably, we use the Taylor expansion

$$\coth^{-1}(x) = \sum_{j=0}^{\infty} \frac{x^{-2j-1}}{2j+1} \quad \text{for} \quad |x| \geq 1. \tag{B.10}$$

The $|x| \geq 1$ condition is met, as the eigenvalues of $i\Omega\sigma$ come in positive-negative pairs, each with magnitude greater than (or equal to) 1. Thus,

$$\begin{aligned}
S_1^{(2)} &= \operatorname{Tr}(\coth^{-1}(\mathrm{i}\Omega\sigma)\mathrm{i}\Omega\sigma) \\
&= \sum_{j=0}^{\infty} \frac{1}{2j+1} \operatorname{Tr}\left((i\Omega\sigma)^{-2j}\right) \\
&= \sum_{j=0}^{\infty} \frac{(-1)^j}{2j+1} \operatorname{Tr}\left((\Omega\sigma)^{-2j}\right) \\
&= \sum_{j=0}^{\infty} \frac{(-1)^j}{2j+1} \operatorname{Tr}\left((\Omega(\cosh(2s)\mathbb{I} + \sinh(2s)M)\right)^{-2j}\right) \\
&= \sum_{j=0}^{\infty} \frac{(-1)^j}{2j+1} \operatorname{Tr}\left((\cosh(2s)\Omega + \sinh(2s)\Omega M)^{-2j}\right) \\
&= \sum_{j=0}^{\infty} \frac{(-1)^j}{2j+1} \operatorname{Tr}\left(\left(\sinh^2(2s)M^2 - \cosh^2(2s)\mathbb{I}\right)^{-j}\right).
\end{aligned} \tag{B.11}$$

We use the negative binomial theorem to expand Eq. (B.11). For non-negative integers j and real numbers x and a, we have

$$(x+a)^{-n} = \sum_{k=0}^{\infty} (-1)^k \binom{n+k-1}{k} x^k a^{-n-k}, \tag{B.12}$$

as discussed in Ref. [297]. This equation can be extended to matrices if x and a commute. Using $x = \sinh^2(2s)M^2$ and $a = -\cosh^2(2s)\mathbb{I}$, we have

$$S_1^{(2)} = 2k + \sum_{j=1}^{\infty} \frac{(-1)^j}{2j+1} \operatorname{Tr}\left(\sum_{i=0}^{\infty} (-1)^i \binom{j+i-1}{i} \sinh^{2i}(2s) M^{2i} (-\cosh^2(2s))^{-j-i} \right) \tag{B.13}$$

$$\begin{aligned} &= 2k + \sum_{j=1}^{\infty} \left(\frac{1}{2j+1} \cdot \sum_{i=0}^{\infty} \binom{j+i-1}{i} \tanh^{2i}(2s) \operatorname{sech}^{2j}(2s) \operatorname{Tr}(M^{2i}) \right) \\ &= 2k \cosh(2s) \tanh^{-1}(\operatorname{sech}(2s)) + \sum_{i=1}^{\infty} \tanh^{2i}(2s) \operatorname{Tr}(M^{2i}) \sum_{j=1}^{\infty} \frac{\operatorname{sech}^{2j}(2s)}{2j+1} \binom{j+i-1}{i} \\ &= 2k \cosh(2s) \tanh^{-1}(\operatorname{sech}(2s)) \\ &\quad + \frac{2}{3} \operatorname{sech}^2(2s) \sum_{i=1}^{\infty} \tanh^{2i}(2s) \, {}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, \operatorname{sech}^2(2s) \right) \operatorname{Tr}(W^i). \end{aligned} \tag{B.14}$$

Therefore,

$$\begin{aligned} \mathop{\mathbb{E}}_{U \in U(n)} S_1(U) &= \frac{k}{2} \log\left[\frac{\sinh^2(2s)}{4} \right] - \sum_{j=1}^{\infty} \frac{1}{2j} \mathop{\mathbb{E}}_{U} \operatorname{Tr} W^j \\ &\quad + k \cosh(2s) \tanh^{-1}(\operatorname{sech}(2s)) \\ &\quad + \frac{\operatorname{sech}^2(2s)}{3} \sum_{i=1}^{\infty} \tanh^{2i}(2s) \, {}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, \operatorname{sech}^2(2s) \right) \mathop{\mathbb{E}}_{U} \operatorname{Tr}(W^i) \end{aligned} \tag{B.15}$$

$$\begin{aligned} &= nr \left(\frac{1}{2} \log\left[\frac{\sinh^2(2s)}{4} \right] + \cosh(2s) \tanh^{-1}[\operatorname{sech}(2s)] \right) \\ &\quad + \frac{\operatorname{sech}^2(2s)}{3} \sum_{i=1}^{\infty} \left[\tanh^{2i}(2s) \, {}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, \operatorname{sech}^2(2s) \right) - \frac{3\cosh^2(2s)}{2i} \right] \mathop{\mathbb{E}}_{U} \operatorname{Tr} W^i \end{aligned} \tag{B.16}$$

$$= \sum_{i=1}^{\infty} \left[\frac{1}{2i} - \frac{1}{3} \operatorname{sech}^2(2s) \tanh^{2i}(2s) \, {}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, \operatorname{sech}^2(2s) \right) \right] (nG_i(r) - H_i(r) + o(1)), \tag{B.17}$$

where the step from Eq. (B.16) to Eq. (B.17) arises from substituting $\mathop{\mathbb{E}}_{U} \operatorname{Tr} W^i = nr - (nG_i(r) - H_i(r) + o(1))$ as derived in Ref. [2] and the subsequent cancellation of the constant term. Finally, dividing by n and noting the vanishing behavior of $H_i(r)$ and $o(1)$ as $n \to \infty$,

the von Neumann Page curve is

$$\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{n} S_1(U) = \sum_{i=1}^{\infty} \left[\frac{1}{2i} - \frac{1}{3}\mathrm{sech}^2(2s)\tanh^{2i}(2s)\, {}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, \mathrm{sech}^2(2s)\right)\right] G_i(r), \tag{B.18}$$

which is given in Theorem 3.3.1 in the main text.

Subsection B.1.2: Limit as $s \to 0$

In this section, we show the validity of Theorem 3.3.2. Moreover, we elucidate the odd behavior of the von Neumann Page curve scaling with $s^2\log(1/s^2)$ instead of the traditional s^2 seen in the other Rényi-α entropies with integer $\alpha \geq 2$. We also give our motivation for investigating the $s^2\log(1/s^2)$ dependence.

For small squeezing strengths, one might expect the Rényi entanglement entropy to scale with the number of bosons in all modes. Furthermore, since the total expected number of bosons in the state of n modes is $\sum_{i=1}^{n}\sinh^2 s \approx ns^2$ for equal squeezing, one may guess that all Rényi-α Page curves scale with s^2.

However,

$$\lim_{s\to 0}\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{s^2 n} S_1(U) = \infty. \tag{B.19}$$

One difficulty arises in evaluating the limit as $s \to \infty$ of the first term in Eq. (B.16). Notably,

$$\frac{1}{2s^2}\log\left[\frac{\sinh^2(2s)}{4}\right] = \frac{\log s}{s^2} + \frac{2}{3} - \frac{4}{45}s^2 + \frac{64}{2835}s^4 - \frac{32}{4725}s^6 + \frac{1024}{467775}s^8 - \ldots, \tag{B.20}$$

$$\frac{1}{s^2}\cosh(2s)\tanh^{-1}(\operatorname{sech}(2s)) = -\frac{\log s}{s^2} + \frac{1-6\log s}{3} + \frac{53-60\log s}{90}s^2 + \frac{251-252\log s}{2835}s^4 + \dots. \tag{B.21}$$

Summing Eqs. (B.20) and (B.21) yields a complicated series involving $\log s$, which diverges as $s \to 0$. Therefore, we try evaluating $\lim_{s\to 0}\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{(s^2\log 1/s^2)n} S_1(U)$ in hopes of achieving a convergent result. Using the entanglement entropy expression derived in Eq. (B.16),

$$\begin{aligned}\lim_{s\to 0}\lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{(s^2\log 1/s^2)n} S_1(U) = \frac{r}{s^2\log 1/s^2}\cdot\left(\frac{1}{2}\log\left[\frac{\sinh^2(2s)}{4}\right] + \cosh(2s)\tanh^{-1}[\operatorname{sech}(2s)]\right) \\ + \frac{\operatorname{sech}^2(2s)}{3s^2\log(1/s^2)}\sum_{i=1}^{\infty}\left[\tanh^{2i}(2s)\,{}_2F_1\left(\frac{3}{2},1+i,\frac{5}{2},\operatorname{sech}^2(2s)\right) - \frac{3\cosh^2(2s)}{2i}\right] \\ \cdot r^{i+1}C_i\,{}_2F_1(1-i;i;i+2;r).\end{aligned} \tag{B.22}$$

Note that $\frac{1}{s^2\log(1/s^2)}\cdot r\left(\frac{1}{2}\log\left[\frac{\sinh^2(2s)}{4}\right] + \cosh(2s)\tanh^{-1}[\operatorname{sech}(2s)]\right)$, the constant term in Eq. (B.22), approaches r as $s \to 0$. Now, we need to find the limit as $s \to 0$ of the second term in Eq. (B.22):

$$\begin{aligned}&= \lim_{s\to 0}\frac{1}{s^2\log(1/s^2)}\frac{\operatorname{sech}^2(2s)}{3} \\ &\quad\times\sum_{i=1}^{\infty}\left[\tanh^{2i}(2s)\,{}_2F_1\left(\frac{3}{2},1+i,\frac{5}{2},\operatorname{sech}^2(2s)\right) - \frac{3\cosh^2(2s)}{2i}\right] r^{i+1}C_i\,{}_2F_1(1-i;i;i+2;r)\end{aligned} \tag{B.23}$$

$$\begin{aligned}&= \sum_{i=1}^{\infty}\lim_{s\to 0}\frac{1}{s^2\log(1/s^2)}\left(\frac{1}{3}\operatorname{sech}^2(2s)\tanh^{2i}(2s)\,{}_2F_1\left(\frac{3}{2},1+i,\frac{5}{2},\operatorname{sech}^2(2s)\right) - \frac{1}{2i}\right) \\ &\quad\times r^{i+1}C_i\,{}_2F_1(1-i;i;i+2;r) \\ &= \sum_{i=1}^{\infty}\lim_{s\to 0}\frac{1}{s^2\log(1/s^2)}\left(\frac{1}{3}\cdot 1\cdot(2s)^{2i}\,{}_2F_1\left(\frac{3}{2},1+i,\frac{5}{2},(1-2s^2)^2\right) - \frac{1}{2i}\right) \\ &\quad\times r^{i+1}C_i\,{}_2F_1(1-i;i;i+2;r).\end{aligned} \tag{B.24}$$

Note that, in the $i = 1$ case, letting $F(s) = \frac{1}{s^2 \log(1/s^2)} \frac{1}{3} \cdot 1 \cdot (2s)^{2i}$, the left-hand derivative $\frac{dF}{ds}_{s=0^-} = -\infty$, while the right-hand derivative $\frac{dF}{ds}_{s=0^+} = \infty$. This behavior deviates from all other values of $i \geq 1$, where $\frac{dF}{ds} = 0$ evaluated at $s = 0$. Combined with the dependence of ${}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, (1-2s^2)^2\right)$ on i, one can explicitly calculate that Eq. (B.23) evaluates to $-r^2$, as the $i = 1$ case yields $-r^2$ and subsequent i yield 0 in the sum. Thus,

$$\lim_{s\to 0} \lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{(s^2 \log(1/s^2))n} S_1(U) = r(1-r). \tag{B.25}$$

It can easily be seen from Eq. (B.16) that the result is the same if we swap the limits, so that by the Moore-Osgood theorem [298],

$$\lim_{\substack{n\to\infty \\ s\to 0}} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{(s^2 \log(1/s^2))n} S_1(U) = r(1-r). \tag{B.26}$$

as displayed in Theorem 3.3.2 in the main text.

Subsection B.1.3: Limit as $s \to \infty$

In a similar vein as Appendix B.1.2, we also compute the von Neumann Page curve for equal squeezing as $s \to \infty$. The Page curve scales with s in the limit $s \to \infty$, so we compute $\lim_{s\to\infty} \lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{sn} S_1(U)$. We once again use the expression found in Eq. (B.16) in our evaluations.

$$\begin{aligned}
\lim_{s\to\infty} \lim_{n\to\infty} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{sn} S_1(U) &= \lim_{s\to\infty} A_1 + \lim_{s\to\infty} A_2, \\
\text{where } A_1 &= \frac{r}{s}\left(\frac{1}{2}\log\left[\frac{\sinh^2(2s)}{4}\right] + \cosh(2s)\tanh^{-1}[\operatorname{sech}(2s)]\right), \\
A_2 &= \frac{1}{s}\sum_{i=1}^{\infty}\left[\frac{1}{3}\operatorname{sech}^2(2s)\tanh^{2i}(2s)\,{}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, \operatorname{sech}^2(2s)\right) - \frac{1}{2i}\right]\cdot f_i(r), \\
f_i(r) &= r^{i+1} C_i\, {}_2F_1(1-i; i; i+2; r).
\end{aligned} \tag{B.27}$$

It turns out that

$$\lim_{s\to\infty} A_1 = \lim_{s\to\infty}\left(\frac{r}{s}\frac{1}{2}\log\left[\frac{\sinh^2(2s)}{4}\right]\right) + \lim_{s\to\infty}\left(\frac{r}{s}\cdot\cosh(2s)\tanh^{-1}[\mathrm{sech}(2s)]\right) \tag{B.28}$$

$$= \frac{r}{2}\lim_{s\to\infty}\left(\frac{1}{s}\log\left[\frac{\sinh^2(2s)}{4}\right]\right) + 0$$

$$= 2\min(r, 1-r). \tag{B.29}$$

Since the Page curve is symmetric about $r = \frac{1}{2}$, Eq. (B.29) contains a $\min(r, 1-r)$ term (as opposed to a simple r-dependence). Next, we calculate $\lim_{s\to\infty} A_2$:

$$\lim_{s\to\infty} A_2 = \lim_{s\to\infty}\frac{1}{s}\sum_{i=1}^{\infty}\left[\frac{1}{3}\mathrm{sech}^2(2s)\tanh^{2i}(2s)\,{}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, \mathrm{sech}^2(2s)\right) - \frac{1}{2i}\right]\cdot f_i(r) \tag{B.30}$$

$$= \lim_{s\to\infty}\frac{1}{s}\sum_{i=1}^{\infty}\left[\frac{1}{3}\cdot 0\cdot 1\,{}_2F_1\left(\frac{3}{2}, 1+i, \frac{5}{2}, 0\right) - \frac{1}{2i}\right]\cdot f_i(r) \tag{B.31}$$

$$= 0. \tag{B.32}$$

Therefore,

$$\lim_{\substack{n\to\infty\\ s\to\infty}} \mathop{\mathbb{E}}_{U\in U(n)} \frac{1}{sn} S_1(U) = 2\min(r, 1-r), \tag{B.33}$$

where we again used the same discussion as around Eq. (B.26)

Eq. (B.33) is verified in Fig. B.2, alongside other Rényi-α Page curves in the limit as $s \to \infty$. Note that both the von Neumann and Rényi Page curves approach the same limit of $2\min(r, 1-r)$ (see Appendix B.2.3 for the relevant derivation). This indicates that, at large squeezing values, the von Neumann entropy behaves similarly to the larger class of Rényi entropies.

Section B.2: General Rényi-α Entropies

In this section, we extend our derivation in Appendix B.1 from the von Neumann entropy to all general Rényi-α entropies, where α is a positive integer greater than 1. We find a closed formula for the Rényi-α Page curve as a function of equal squeezing strength s and partition ratio r in the asymptotic limit $n \to \infty$. Similar to the von Neumann case, we calculate squeezing limits of the Rényi-α entropies as $s \to 0$ and $s \to \infty$ and then discuss typicality of entanglement based on the variance.

Subsection B.2.1: Page Curve for the Rényi-α entropy

Here, we present our full derivation for the Page curve of the Rényi-α entropy, where α is an integer great than 1, in terms of the partition ratio r and equal squeezing strength s.

Our starting point is the general expression relating the Rényi-α entropy with the symplectic eigenvalues ν_j of the covariance matrix σ (recall that σ is precisely the second moment of ρ [138]):

$$\mathcal{S}_\alpha(\rho) = \frac{1}{\alpha - 1} \sum_{j=1}^{k} \log[(\nu_j + 1)^\alpha - (\nu_j - 1)^\alpha] - \frac{\alpha k \log 2}{\alpha - 1}. \tag{B.34}$$

If α is a positive integer, then this function has $\alpha - 1$ singularities. We can solve for those singularities explicitly:

$$(\nu_0 + 1)^\alpha = (\nu_0 - 1)^\alpha \Rightarrow \nu_0 + 1 = (\nu_0 - 1) e^{i\frac{2\pi\ell}{\alpha}}, \tag{B.35}$$

$$\Rightarrow \nu_0 = -i \cot \frac{\pi \ell}{\alpha}, \quad \ell = 1, \ldots, \alpha - 1. \tag{B.36}$$

Note that we exclude the $\ell = \alpha$ case as Eq. (B.35) does not have a root in that case. Thus, the polynomial $(\nu_j + 1)^\alpha - (\nu_j - 1)^\alpha$ can be rewritten in terms of a product of factors as

follows:

$$(\nu_j+1)^\alpha-(\nu_j-1)^\alpha=2\alpha\nu_j^\zeta\prod_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}(\nu_j^2+\cot^2\frac{\pi m}{\alpha}), \tag{B.37}$$

where we define $\zeta:=1-(\alpha \bmod 2)$. The Rényi entropy is in turn given by

$$\mathcal{S}_\alpha(\rho)=\frac{1}{\alpha-1}\sum_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}\sum_{j=1}^{k}\log(\nu_j^2+\cot^2\frac{\pi m}{\alpha})+\frac{\zeta\sum_{j=1}^k\log\nu_j^2}{2(\alpha-1)}+\frac{k\log\alpha}{\alpha-1}-k\log 2. \tag{B.38}$$

To work with the $\log\nu_j^2$ terms, note that the squares of the symplectic eigenvalues of σ are precisely the eigenvalues of $(i\Omega\sigma)^2$ by matrix properties. Furthermore, since the eigenvalues of $i\Omega\sigma$ come in plus-minus pairs but the symplectic eigenvalues only consider positive numbers, the factor of $\frac{1}{2}$ in Eq. (B.38) arises. Taking $(i\Omega\sigma)^2=(i\Omega(\mathbb{I}\cosh(2s)+M\sinh(2s)))^2=\cosh^2(2s)-M^2\sinh^2(2s)$ into the above equation, we thus have

$$\begin{aligned}\mathcal{S}_\alpha(\rho)=&\frac{1}{2(\alpha-1)}\sum_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}\operatorname{Tr}\log(\cot^2\frac{\pi m}{\alpha}+\cosh^2(2s)-M^2\sinh^2(2s))\\&+\frac{\zeta}{4(\alpha-1)}\operatorname{Tr}\log(\cosh^2(2s)-M^2\sinh^2(2s))+\frac{k\log(\alpha)}{\alpha-1}-k\log 2.\end{aligned} \tag{B.39}$$

Let us now expand all the logarithmic functions in the sum in Eq. (B.39) in orders of M^2. Noting that $M^2\leq 1$, one could show that each of the following series always converge (for any integer ℓ):

$$\operatorname{Tr}\log(\cosh^2(2s)-M^2\sinh^2(2s))=4k\log\cosh(2s)-\sum_{\ell=1}^{\infty}\frac{\sinh^{2\ell}(2s)}{\ell\cosh^{2\ell}(2s)}\operatorname{Tr}(M^{2\ell}), \tag{B.40}$$

$$\operatorname{Tr}\log(\cot^2\frac{\pi m}{\alpha}+\cosh^2(2s)-M^2\sinh^2(2s))=2k\log(\cot^2\frac{\pi m}{\alpha}+\cosh^2(2s))\\ -\sum_{\ell=1}^{\infty}\frac{\sinh^{2\ell}(2s)}{\ell(\cosh^2(2s)+\cot^2\frac{\pi m}{\alpha})^{\ell}}\operatorname{Tr}(M^{2\ell}). \tag{B.41}$$

Note that the product in Eq. (B.37) vanishes when substituting $\alpha=2$. Therefore, one gets from Eq. (B.38) that $S_2=\sum_{j=1}^{k}\log\nu_j^2$. Combining the expansion in Eq. (B.40), we get that $S_2=k\log\cosh(2s)-\sum_{\ell=1}^{\infty}\frac{\tanh^{2\ell}(2s)}{2\ell}\operatorname{Tr}(M^{2\ell})$. This expression for S_2 is utilized in the following steps to formally express the Rényi entropy with a positive integer α as

$$\begin{aligned}\mathcal{S}_\alpha(\rho)=&k\sum_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}\frac{\log(\cot^2\frac{\pi m}{\alpha}+\cosh^2(2s))}{\alpha-1}+k\frac{\zeta\log\cosh(2s)+\log\alpha}{\alpha-1}-k\log 2 \qquad \text{(B.42)}\\ &-\frac{1}{\alpha-1}\sum_{\ell=1}^{\infty}\left[\sum_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}\frac{\sinh^{2\ell}(2s)}{\ell(\cosh^2(2s)+\cot^2\frac{\pi m}{2\alpha})^{\ell}}+\zeta\frac{\sinh^{2\ell}(2s)}{2\ell\cosh^{2\ell}(2s)}\right]\frac{\operatorname{Tr}(M^{2\ell})}{2}\\ =&k\sum_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}\frac{\log(\cot^2\frac{\pi m}{\alpha}+\cosh^2(2s))}{\alpha-1}+\frac{k\log\alpha}{\alpha-1}-k\log 2+\frac{\zeta}{\alpha-1}S_2\\ &-\frac{1}{\alpha-1}\sum_{\ell=1}^{\infty}\left[\sum_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}\frac{\sinh^{2\ell}(2s)}{\ell(\cosh^2(2s)+\cot^2\frac{\pi m}{\alpha})^{\ell}}\right]\frac{\operatorname{Tr}(M^{2\ell})}{2} \qquad \text{(B.43)}\\ =&\frac{\zeta}{\alpha-1}S_2+\frac{1}{\alpha-1}\sum_{\ell=1}^{\infty}\left[\sum_{m=1}^{a}\frac{\sinh^{2\ell}(2s)}{\ell(\cosh^2(2s)+\cot^2\frac{\pi m}{\alpha})^{\ell}}\right]\left[k-\operatorname{Tr}(W^{\ell})\right], \qquad \text{(B.44)}\end{aligned}$$

where $a\equiv\lfloor\frac{\alpha-1}{2}\rfloor=\frac{\alpha-1-\zeta}{2}$. We have used $\operatorname{Tr}M^{2\ell}=2\operatorname{Tr}W^{\ell}$ in Eq. (B.44). The step from Eq. (B.43) to Eq. (B.44) utilizes the relation presented in Eq. (B.41), where M is substituted with the identity matrix. We can then use the expression for $\mathbb{E}_{U\in\mathrm{U}(n)}\operatorname{Tr}W^{\ell}$ from Ref. [2] to derive the Page curve.

Subsection B.2.2: Limit as $s \to 0$

We now analyze the Rényi-α entropy as the squeezing strength approaches zero. This time, due to the analyticity of the Rényi entropies for integer $\alpha \geq 2$, the Page curve scales with the number of bosons $\sim s^2$ instead of the $s^2 \log(1/s^2)$ dependence explained in Appendix B.1.2.

Recall that $G_\ell(r)$ is a function symmetric about $r = \frac{1}{2}$ defined as $r - r^{\ell+1} C_\ell \, {}_2F_1(1-\ell; \ell; \ell+2; r)$. Using Eq. (B.44) and substituting $k - \mathrm{Tr}(W^\ell) = nG_\ell(r)$ to investigate the $s \to 0$ limit,

$$\lim_{s\to 0}\lim_{n\to\infty} \frac{1}{s^2 n} \mathop{\mathbb{E}}_{U} S_\alpha(U) = \lim_{s\to 0} \frac{1}{s^2} \frac{1}{\alpha - 1} \sum_{\ell=1}^{\infty} \frac{1}{\ell} (2s)^{2\ell} G_\ell(r) \sum_{m=1}^{a} \left(1 + \cot^2\left(\frac{\pi m}{\alpha}\right)\right)^{-\ell} + \frac{\zeta}{\alpha - 1} 2r(1-r) \tag{B.45}$$

$$\begin{aligned}
&= \lim_{s\to 0} \frac{1}{s^2} \frac{1}{\alpha - 1} (2s)^2 G_1(r) \sum_{m=1}^{a} \left(\cot^2\left(\frac{\pi m}{\alpha}\right) + 1\right)^{-1} + \frac{\zeta}{\alpha - 1} 2r(1-r) \\
&= \frac{4}{\alpha - 1} r(1-r) \sum_{m=1}^{a} \left(\cot^2\left(\frac{\pi m}{\alpha}\right) + 1\right)^{-1} + \frac{\zeta}{\alpha - 1} 2r(1-r) \\
&= \frac{r(1-r)}{\alpha - 1} \left(4 \sum_{m=1}^{a} \left(\cot^2\left(\frac{\pi m}{\alpha}\right) + 1\right)^{-1} + 2\zeta\right) \\
&= \frac{r(1-r)}{\alpha - 1} \left(\left(-\csc\left(\frac{\pi}{\alpha}\right) \sin\left(\frac{2\pi a + \pi}{\alpha}\right) + 2a + 1\right) + 2\zeta\right).
\end{aligned} \tag{B.46}$$

Next, we use that $a = \lfloor (\alpha - 1)/2 \rfloor = (\alpha - 1 - \zeta)/2$. Then

$$\lim_{s\to 0}\lim_{n\to\infty} \frac{1}{s^2 n} \mathop{\mathbb{E}}_{U} S_\alpha(U) = \frac{r(1-r)}{\alpha - 1} \left(\alpha + \zeta - \frac{\sin(\pi\zeta/\alpha)}{\sin(\pi/\alpha)}\right) \tag{B.47}$$

$$= \frac{\alpha}{\alpha - 1} \cdot r(1-r). \tag{B.48}$$

As in the other cases, this results holds as a double limit $\lim_{\substack{n\to\infty \\ s\to 0}}$. Evidently, the Rényi-$\alpha$ Page Curve with small squeezing approaches $r(1-r) \cdot s^2$ as $\alpha \to \infty$. We demonstrate Eq. (B.48) numerically in Fig. B.1.

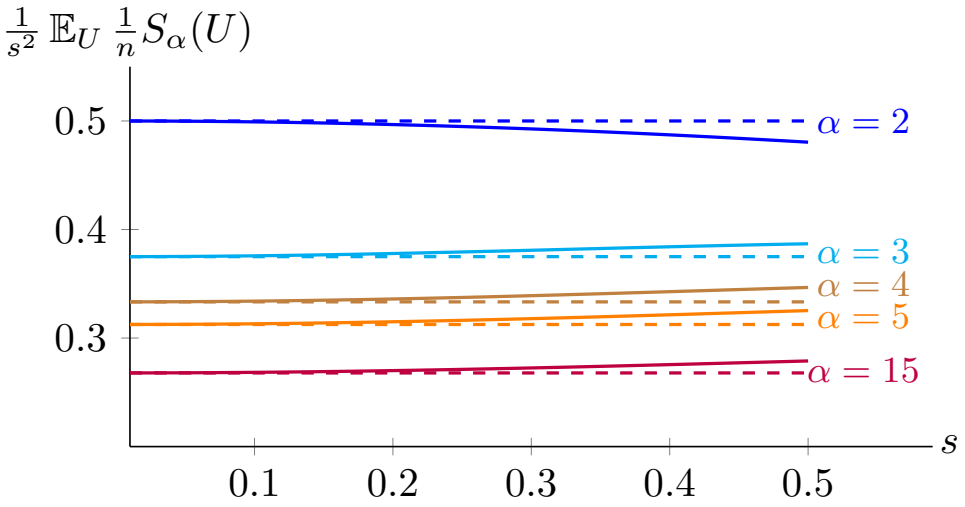

Figure B.1: The Rényi-α Page curves, for $\alpha \in \{2, 3, 4, 5, 15\}$, divided by s^2, as a function of the squeezing strength s. These functions are displayed as the solid lines. The Rényi-α Page curve limit as $s \to 0$, presented in Eq. (B.48), is shown by the dashed lines in corresponding color. As expected, the Rényi-α Page curve approaches $\frac{\alpha}{\alpha-1}r(1-r)$ in order s^2, where we take $r = 0.5$ and $n = 400$.

Subsection B.2.3: Limit as $s \to \infty$

Now, we investigate the squeezing limit $s \to \infty$.

We have

$$\lim_{s\to\infty}\lim_{n\to\infty}\frac{1}{sn}\mathop{\mathbb{E}}_{U} S_\alpha(U) = \lim_{s\to\infty}\frac{1}{s}\frac{1}{\alpha-1}\sum_{\ell=1}^{\infty}\frac{1}{\ell}\sinh^{2\ell}(2s)G_\ell(r)\sum_{m=1}^{a}\left(\cot^2\left(\frac{\pi m}{\alpha}\right)+\cosh^2(2s)\right)^{-\ell} + \frac{\zeta}{\alpha-1}2\,\min(r,1-r) \tag{B.49}$$

$$\begin{aligned}
&= \lim_{s\to\infty}\frac{1}{s}\frac{2a}{\alpha-1}\left(\sum_{\ell=1}^{\infty}\frac{\tanh^{2\ell}(2s)}{2\ell}G_\ell(r)\right)+\frac{\zeta}{\alpha-1}2\,\min(r,1-r)\\
&= \frac{2a}{\alpha-1}2\,\min(r,1-r)+\frac{\zeta}{\alpha-1}2\,\min(r,1-r)\\
&= \frac{4a+2\zeta}{\alpha-1}\,\min(r,1-r)\\
&= 2\,\min(r,1-r).
\end{aligned} \tag{B.50}$$

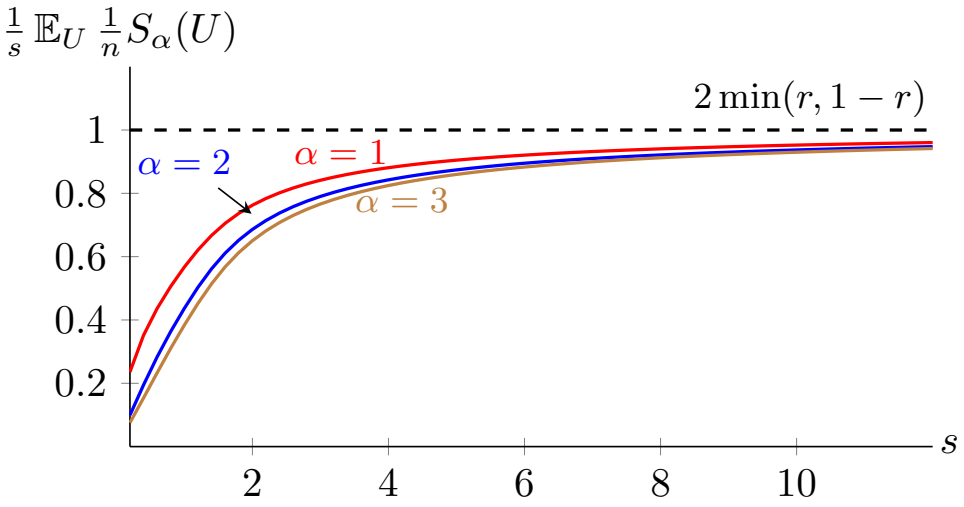

Figure B.2: Plots of the von Neumann, Rényi-2, and Rényi-3 Page curves divided by s, as a function of the squeezing strength s. These functions, which are presented more generally in Theorem 3.3.3 and Theorem 3.4.2, are shown by the solid red, blue, and brown lines, respectively. The von Neumann and Rényi Page curves in the limit as squeezing strength approaches infinity, derived in Appendix B.1.3 and Appendix B.2.3, have coinciding values of $2\min(r, 1-r)$ displayed as the black dotted line. We take $r = 0.5$.

As in the other cases, this results holds as a double limit $\lim_{\substack{n\to\infty \\ s\to 0}}$.

We numerically validate Eq. (B.50) and Eq. (B.33) in Fig. B.2 below. As shown, both the von Neumann and Rényi-α entropies indeed approach $2\min(r, 1-r)$ in order s as $s \to \infty$.

Subsection B.2.4: Variance and typicality of entanglement

In this section, we prove weak typicality of entanglement described in Section 3.4.4. We consider the general expression for the Rényi-α entropy as found in Eq. (B.44):

$$\mathcal{S}_\alpha(\rho) = \frac{\zeta}{\alpha-1}S_2 + \frac{1}{\alpha-1}\sum_{\ell=1}^{\infty}\left[\sum_{m=1}^{a}\frac{\sinh^{2\ell}(2s)}{\ell(\cosh^2(2s)+\cot^2\frac{\pi m}{\alpha})^\ell}\right]\left[k - \mathrm{Tr}(W^\ell)\right] \tag{B.51}$$

$$= \frac{1}{\alpha-1}\sum_{\ell=1}^{\infty}\left[\sum_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}\frac{\sinh^{2\ell}(2s)}{\ell(\cosh^2(2s)+\tan^2\frac{\pi m}{2\alpha})^\ell} + \zeta\frac{\sinh^{2\ell}(2s)}{2\ell\cosh^{2\ell}(2s)}\right]\left[k - \mathrm{Tr}(W^\ell)\right], \tag{B.52}$$

with the same definition $a := \lfloor \frac{\alpha-1}{2} \rfloor = \frac{\alpha-1-\zeta}{2}$. More specifically, the $\alpha = 2$ Rényi entropy is given by

$$\mathcal{S}_{\alpha=2}(\rho) = \sum_{\ell=1}^{\infty} \frac{\sinh^{2\ell}(2s)}{2\ell \cosh^{2\ell}(2s)} \left[k - \mathrm{Tr}(W^\ell)\right]. \tag{B.53}$$

From Ref. [2], we know that the following covariance is a non-zero constant independent of n:

$$\mathcal{V}_d \equiv \sum_{\ell=1}^{d-1} \frac{1}{\ell(d-\ell)} (\overline{\mathrm{Tr}(W^\ell)\,\mathrm{Tr}(W^{d-\ell})} - \overline{\mathrm{Tr}(W^\ell)} \cdot \overline{\mathrm{Tr}(W^{d-\ell})}). \tag{B.54}$$

Then the variance of $\mathcal{S}_{\alpha=2}(\rho)$ can also be proved to be a constant. Letting the fluctuation in the Rényi entropy be defined as $\delta\mathcal{S}_{\alpha=2} = \mathcal{S}_{\alpha=2}(\rho) - \overline{\mathcal{S}_{\alpha=2}(\rho)}$, we get $\mathrm{Var}\, S_{\alpha=2}(\rho) = \overline{\delta\mathcal{S}^2_{\alpha=2}}$ as

$$\overline{\delta\mathcal{S}^2_{\alpha=2}} = \frac{1}{4} \sum_{d=2}^{\infty} \tanh^{2d}(2s) \sum_{\ell=1}^{d-1} \frac{1}{\ell(d-\ell)} (\overline{\mathrm{Tr}(W^\ell)\,\mathrm{Tr}(W^{d-\ell})} - \overline{\mathrm{Tr}(W^\ell)} \cdot \overline{\mathrm{Tr}(W^{d-\ell})}) \tag{B.55}$$

$$= \frac{1}{4} \sum_{d=2}^{\infty} \tanh^{2d}(2s) \mathcal{V}_d. \tag{B.56}$$

On the other hand, we can also expand the Rényi entropy as

$$\mathcal{S}_\alpha(\rho) = \frac{1}{\alpha-1} \sum_{m=1}^{\lfloor \frac{\alpha-1}{2} \rfloor} \mathcal{C}_m(\alpha) + \frac{\zeta}{\alpha-1} \mathcal{S}_2(\rho) \tag{B.57}$$

$$= \frac{1}{\alpha-1} \left[\sum_{m=1}^{\lfloor \frac{\alpha-1}{2} \rfloor} \sum_{\ell=1}^{\infty} \left(\frac{\sinh^2(2s)}{\cosh^2(2s) + \cot^2(\frac{\pi m}{2\alpha})} \right)^\ell \frac{\left[k - \mathrm{Tr}(W^\ell)\right]}{\ell} + \frac{\zeta}{2} \sum_{\ell=1}^{\infty} \tanh^{2\ell}(2s) \frac{\left[k - \mathrm{Tr}(W^\ell)\right]}{\ell} \right], \tag{B.58}$$

where

$$\mathcal{C}_m(\alpha) = \sum_{\ell=1}^{\infty} \left(\frac{\sinh^2(2s)}{\cosh^2(2s) + \cot^2(\frac{\pi m}{2\alpha})} \right)^\ell \frac{\left[k - \mathrm{Tr}(W^\ell)\right]}{\ell}, \quad m = 1, 2, \ldots, \left\lfloor \frac{\alpha-1}{2} \right\rfloor. \tag{B.59}$$

Note that all quantities $\mathcal{C}_m(\alpha)$ are well-defined and converge for all integers m, α, and s, and, by definition, $\mathcal{C}_0(\alpha) = \mathcal{S}_{\alpha=2}(\rho)$. Further, for each $\mathcal{C}_m(\alpha)$, if we introduce its fluctuation $\delta\mathcal{C}_m(\alpha) = \mathcal{C}_m(\alpha) - \overline{\mathcal{C}_m(\alpha)}$, then

$$\overline{\delta\mathcal{C}_m^2(\alpha)} = \frac{1}{4}\sum_{d=2}^{\infty}\left(\frac{\sinh^2(2s)}{\cosh^2(2s)+\cot^2(\frac{\pi m}{2\alpha})}\right)^d \sum_{\ell=1}^{d-1}\frac{1}{\ell(d-\ell)}(\overline{\mathrm{Tr}(W^\ell)\,\mathrm{Tr}(W^{d-\ell})} - \overline{\mathrm{Tr}(W^\ell)}\cdot\overline{\mathrm{Tr}(W^{d-\ell})}) \tag{B.60}$$

$$= \frac{1}{4}\sum_{d=2}^{\infty}\left(\frac{\sinh^2(2s)}{\cosh^2(2s)+\cot^2(\frac{\pi m}{2\alpha})}\right)^d \mathcal{V}_d \tag{B.61}$$

$$\leq \frac{1}{4}\sum_{d=2}^{\infty}\tanh^{2d}(2s)\mathcal{V}_d,$$

meaning that each C_m has smaller variance than that of C_0. Furthermore, by the Cauchy-Schwarz inequality, we have that

$$\overline{\delta\mathcal{C}_m(\alpha)\delta\mathcal{C}_{m'}(\alpha)} \leq \sqrt{\overline{\delta\mathcal{C}_m(\alpha)\delta\mathcal{C}_m(\alpha)}\cdot\overline{\delta\mathcal{C}_{m'}(\alpha)\delta\mathcal{C}_{m'}(\alpha)}} \quad \forall m,m' \in 0,1,\ldots,\left\lfloor\frac{\alpha-1}{2}\right\rfloor, \tag{B.62}$$

so that covariances of the form $\overline{\delta\mathcal{C}_m(\alpha)\delta\mathcal{C}_{m'}(\alpha)}$ $\forall m,m'$ must also be bounded by a constant as n increases. Finally, since we have found the variance of $\mathcal{C}_m(\alpha)$ and $S_2(\rho)$ to be constant in n, we can prove the concentration of Rényi entropy with integer $\alpha \geq 2$:

$$\mathcal{S}_\alpha(\rho) = \frac{1}{\alpha-1}\sum_{m=1}^{\lfloor\frac{\alpha-1}{2}\rfloor}\mathcal{C}_m(\alpha) + \frac{\zeta}{\alpha-1}\mathcal{S}_2(\rho) \Rightarrow \overline{\delta\mathcal{S}_\alpha^2} = \text{const.} \tag{B.63}$$

Appendix C: Technical details for Chapter 4

Section C.1: Pointers to Appendix C

In Appendices C.2 to C.6, we provide proofs of our main results and summarize relevant background material on continuous-variable (CV) information theory. Appendix C.2 covers relevant definitions from measure theory and properties of projectors onto the symmetric subspace of a separable Hilbert space. Appendices C.3.1 and C.3.2 review finite-dimensional simplex and torus designs. Appendices C.3.3 to C.3.5 review complex-projective designs and their relationship to simplex and torus designs. Using simplex and torus designs, we develop a design formalism for constrained complex-projective integration in Appendix C.3.6. To the best of our knowledge, the formalism developed in Appendix C.3.6 is novel. Similarly, to the best of our knowledge, Definition 4.3.1, Theorem C.3.9, and Proposition C.3.11 from Appendix C.3.2 are new, though we prove in Appendix C.6 that Definition 4.3.1 is equivalent to a previous definition given in Ref. [189]. A relationship between simplex, torus, and complex-projective designs was described Ref. [189]. We further extend on this relationship in Appendices C.3.3 to C.3.5.

Readers who are familiar with finite-dimensional complex-projective designs may wish to begin directly from Appendix C.4. Appendices C.4 and C.5 discuss the main results of this paper. In Appendices C.4.1 and C.4.2, we prove that CV state and unitary t-designs do not exist for $t > 1$. In Appendices C.4.3 and C.4.4, we define and construct rigged and regularized rigged designs, which are generalizations of CV state designs. In Appendix C.4.5, we discuss an alternative characterization of CV, rigged, and reguarized rigged designs based

on integration over infinite-dimensional Gaussian measures. In Appendix C.4.7, using regularized rigged designs, we propose a new definition of an approximate CV unitary t-design. In Appendix C.5.1, we develop the formalism for CV shadows based on rigged designs. We then define the average fidelity of a CV channel based on regularized rigged designs in Appendix C.5.2. Finally, in Appendix C.6, we establish a relationship between torus 2-designs and complete sets of mutually unbiased bases that, to the best of our knowledge, had not been previously established.

Section C.2: Preliminaries

In this section, we summarize some definitions and prior results relevant for the rest of the appendix. We point readers to [138, 237] and [299] for background on continuous-variable information theory and measure theory, respectively.

Throughout this manuscript, $\mathbb{N}$ and $\mathbb{N}_0$ denote the sets of positive and non-negative integers, respectively. A t-fold Cartesian product $\mathbb{N}_0 \times \cdots \times \mathbb{N}_0$ will be denoted by $\mathbb{N}_0^t$. $\mathbb{Z}_d$ will be the integers modulo d, $\mathbb{Z}_d = \{0, \ldots, d-1\}$.

States: We will consider continuous-variable states (normalized vectors) in the separable infinite dimensional Hilbert space $\mathcal{H} = L^2(\mathbb{R})$. Separable Hilbert spaces, by definition, have a *Schauder* or *Hilbert Space basis*; any vector in a separable Hilbert space can be written as $\sum_{n=0}^{\infty} \alpha_n |v_n\rangle$ for some Schauder basis $\{v_n\}$ which is always guaranteed to exist [300, ch. 17.1]. For concreteness, when discussing an explicit basis, we will use the standard Fock basis on $L^2(\mathbb{R})$, denoted by $\{|n\rangle \mid n \in \mathbb{N}_0\}$. In the position representation, a Fock state $|n\rangle$ is $\psi_n(x) = \langle x|n\rangle = \frac{\pi^{-1/4}}{\sqrt{n!2^n}} \mathrm{e}^{-x^2/2} H_n(x)$, where H_n is the n^{th} Hermite polynomial. We will also consider d dimensional qudit states, where the Hilbert space is $\mathbb{C}^d$. We will fix an orthonormal basis of $\mathbb{C}^d$ and denote it as $\{|0\rangle, \ldots, |d-1\rangle\}$. Qudit states belong to complex-projective space $\mathbb{CP}^{d-1}$, which is described more in Appendix C.2.1.

m-torus and m-simplex: The m-torus is denoted by $T^m \cong [0, 2\pi)^m = (\mathbb{R}/2\pi\mathbb{Z})^m$. The unit-normalized Lebesgue measure on T^m is given by $\mathrm{d}\phi \coloneqq \frac{1}{(2\pi)^m}\,\mathrm{d}\phi_1 \ldots \mathrm{d}\phi_m$. Moreover, the m-simplex is defined as

$$\Delta^m = \left\{ (p_0, \ldots, p_m) \in [0,1]^{m+1} \middle| \sum_{i=0}^{m} p_i = 1 \right\}. \tag{C.1}$$

Any integration over Δ^m can be defined using the unit-normalized Lebesgue measure on Δ^m as follows:

$$\int_{\Delta^m} f(p)\,\mathrm{d}p = m! \int_{[0,1]^{m+1}} f(p_0, \ldots, p_m)\delta(1 - p_0 - \cdots - p_m)\,\mathrm{d}p_0 \ldots \mathrm{d}p_m, \tag{C.2}$$

where δ is the Dirac delta function and $f(p)$ is any function over p.

Subsection C.2.1: Measure theory

In this section, we summarize definitions and key theorems from measure theory. We point readers to [299] for more details. For a concise introduction to basic concepts in measure theory, we recommend video lectures in [301], which serves as much of the inspiration for our summary below.

For a finite set X, the most natural way to assign a measure (i.e. "size" or "volume") to subsets of X is by cardinality. However, for many applications, this method breaks down for infinitely large sets. Intuitively speaking, measure theory is a way to generalize the notion of determining the size of a subset to infinitely large sets. To begin, fix a possibly infinite set X. We will denote the power set of X by $\mathcal{P}(X)$.

To assign generalized "volumes" to subsets of X, we are looking for a map $\mu\colon \Sigma \to [0, \infty]$, where $\Sigma \subseteq \mathcal{P}(X)$ is some collection of subsets of X. For a subset $A \subseteq X$, let $A \in \Sigma$. We assign the volume, or measure, of A in X to be $\mu(A)$. Notice that the codomain of μ is the positive extended real line $[0, \infty]$, which we define to be $[0, \infty) \cup \{\infty\}$. This notation signifies

$[0, \infty)$ as the standard nonnegative part of $\mathbb{R}$, and $\{\infty\}$ as the set containing the symbol ∞. In other words, we include ∞ in the codomain of the measure μ. For all $r \in [0, \infty]$, the symbol ∞ is defined by the following three rules:

$$r + \infty \coloneqq \infty \ , \qquad r \cdot \infty \coloneqq \begin{cases} 0 & \text{if } r = 0 \\ \infty & \text{otherwise} \ , \end{cases} \qquad \infty - \infty \text{ undefined} \ . \tag{C.3}$$

The domain Σ of μ is the collection of all measurable subsets of X, where a measure is assigned to each element of Σ by μ. In particular, the collection of measurable sets should satisfy:

1. $\emptyset, X \in \Sigma$, i.e., a volume can be assigned to the empty set and the whole set X.
2. If $A \in \Sigma$, then the complement of A, $A^c = X \setminus A$, should also be in Σ, i.e., if A is measurable, the complement of A should also be measurable.
3. If a countable collection of sets A_i are in Σ, then their union $\bigcup_i A_i$ should also be in Σ.

A set $\Sigma \subseteq \mathcal{P}(X)$ satisfying these aforementioned properties is called a *σ-algebra*.

Given a set X and a σ-algebra Σ on X, one can then formally define a measure $\mu \colon \Sigma \to [0, \infty]$. μ should generalize the properties of volume, and therefore must satisfy the following two conditions:

1. $\mu(\emptyset) = 0$, i.e., the empty set has zero volume.
2. For any countable collection of pairwise disjoint sets $A_j \in \Sigma$, $\mu(\bigcup_j A_j) = \sum_j \mu(A_j)$, i.e., the volume of a region is the sum of the volumes of its constituents.

The triplet (X, Σ, μ) is called a *measure space*. A measure μ on X is called *σ-finite* if X is the union of at most countably many subsets of finite measure. In other words, if their exist

a countable collection $A_1, A_2, \cdots \in \Sigma$ such that $\bigcup_j A_j = X$ and each A_j satisfies $\mu(A_j) < \infty$, then μ is σ-finite. For example, consider $X = \mathbb{R}$, and $A_j = (j - 1.1, j + 0.1) \cup (-j - 0.1, -j + 1.1)$. The length of each A_j is 2.4 which is finite, and the countable union $\cup_j A_j = \mathbb{R}$. Hence, $\mathbb{R}$ with the measure $\mu((a, b)) = b - a$ is σ-finite.

One can show that for many cases of interest, not all subsets can be measurable (i.e. $\Sigma \neq \mathcal{P}(X)$) if the measure is desired to satisfy certain properties. For example, in the case of $X = \mathbb{R}$, we desire the measure to have the properties $\mu(r + A) = \mu(A)$ for all $r \in \mathbb{R}$ and $A \subseteq X$, and $\mu([a, b]) = b - a$. One can prove that such a $\mu \colon \mathcal{P}(X) \to [0, \infty]$ cannot exist. Hence, in general, one must restrict the σ-algebra Σ to not be the entire power set. The most important σ-algebra on $\mathbb{R}$ is the Borel σ-algebra, which is the smallest σ-algebra that contains all open sets in $\mathbb{R}$ equipped with the standard topology. The most important measure on $\mathbb{R}$ is the Lebesgue measure, which satisfies the two properties above. Given the product of two σ-finite measure spaces, one can define a unique product measure space. Using this construction, one can construct the Lebesgue measure on $\mathbb{R}^2$, and this can be reinterpreted as a Lebesgue measure on $\mathbb{C}$.

A crucial feature of measure spaces is the concept of *μ-almost-everywhere*, often abbreviated μ-a.e., or just a.e. if the measure is clear. A property is said to hold μ-a.e. if it is true everywhere except on a subset that is contained inside a subset of measure zero. For example, the rationals $\mathbb{Q}$ are contained within a measurable subset of measure zero in the reals $\mathbb{R}$ with respect to the Lebesgue measure. In fact, $\mathbb{Q}$ is itself measurable. Therefore, the property that "$r \in \mathbb{R}$ is irrational" holds a.e. with respect to the Lebesgue measure. One important property that will show up often is $f \leq g$ a.e. for two measurable functions f, g. This means that the set $\{x \in X \mid g(x) > f(x)\}$ is contained within a measurable set with measure zero.

Between two measure spaces (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2), a map $f \colon X_1 \to X_2$ is called *measurable* if the preimage of measurable sets is measurable, meaning that $f^{-1}(A_2) \in \Sigma_1$

for all $A_2 \in \Sigma_2$. For $X_2 = \mathbb{R}$ with the standard Lebesgue measure, if a map $f\colon X \to \mathbb{R}$ is measurable then for every $r \in \mathbb{R}$ the preimage $f^{-1}(\{r\}) = \{x \in X_1 \mid f(x) = r\}$ is measurable [299, p. 359-360]. Intuitively, in order to integrate over a function, we must be able to determine the measure of the domain for which that function takes a certain value. For example, in the case of a bump function $f\colon \mathbb{R} \to \mathbb{R}$ where $f(r) = c$ whenever $r \in A$ and zero otherwise, the integral of f is defined as $c\mu_1(A) = c\mu_1(f^{-1}(\{c\}))$. Therefore, we must require that $f^{-1}(\{c\})$ be measurable.

We will now briefly describe the intuition for Lebesgue integration. For a measurable set $A \in \Sigma$, the indicator function $\mathbb{1}_A(x)$ is defined to be 1 if $x \in A$ and 0 otherwise. A *simple function* is any function of the form $f_{\text{sim}} = \sum_{i=1}^{m} \alpha_i \mathbb{1}_{A_i}$ for each $\alpha_i \in \mathbb{R}$ and $A_i \in \Sigma$. The Lebesgue integral of f_{sim} is defined as $\int_X f_{\text{sim}} \, \mathrm{d}\mu \coloneqq \sum_{i=1}^{m} \alpha_i \mu(A_i)$. Let $\mathcal{S}$ denote the set of all simple functions. For a non-simple, nonnegative function f, the Lebesgue integral of f is defined as a supremum over all simple functions

$$\int_X f \, \mathrm{d}\mu \coloneqq \sup_{\substack{f_{\text{sim}} \in \mathcal{S} \\ f_{\text{sim}} \leq f \text{ a.e.}}} \int_X f_{\text{sim}} \, \mathrm{d}\mu. \tag{C.4}$$

Finally, for a general measurable function f, the Lebesgue integral of f is defined in terms of the integral of nonnegative functions by $\int_X f \, \mathrm{d}\mu = \int_X \max(0, f) \, \mathrm{d}\mu - \int_X \max(0, -f) \, \mathrm{d}\mu$, and is hence defined only if both $\max(0, f)$ and $\max(0, -f)$ are integrable since $\infty - \infty$ is undefined.

A measurable function f is said to be *integrable* if $\int_X |f| \, \mathrm{d}\mu$ is finite. One basic fact about Lebesgue integration is that if $f \leq g$ almost everywhere, then $\int_X f \, \mathrm{d}\mu \leq \int_X g \, \mathrm{d}\mu$. Also, the integral is linear, so that $\int_X (f + g) \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu + \int_X g \, \mathrm{d}\mu$. Oftentimes, we will include an integration parameter for clarity. We define the notation

$$\int_X f(x) \, \mathrm{d}\mu(x) \coloneqq \int_X f \, \mathrm{d}\mu \ . \tag{C.5}$$

The space $L^t(X, \Sigma, \mu)$ is the set measurable functions (identified if they agree almost everywhere) $f\colon X \to \mathbb{R}$ for which $\int_X |f|^t \, \mathrm{d}\mu < \infty$. Define the t-norm to be $\|f\|_t \coloneqq \left(\int_X |f|^t \, \mathrm{d}\mu\right)^{1/t}$. With respect to the t-norm, $L^t(X, \Sigma, \mu)$ is a Banach space. $L^2(X, \Sigma, \mu)$ is a Hilbert space with respect to the inner product $\langle f, g\rangle = \int_X fg \, \mathrm{d}\mu$ in the real case, and similarly in the complex case but with $g \to \bar{g}$. A bounded sequence in $L^t(X, \Sigma, \mu)$ is a sequence of measurable maps $(f_i)_{i\in\mathbb{N}}$ for which $\|f_i\|_t < M$ for some finite number $M \in \mathbb{R}$. When the σ-algebra and measure are clear from context, we denote $L^t(X, \Sigma, \mu)$ as $L^t(X)$. For example, when $L^2(\mathbb{R})$ is written, the σ-algebra and measure are assumed to be the standard Borel σ-algebra and Lebesgue measure on $\mathbb{R}$.

Theorems and lemmas

We now state and discuss various theorems that are used in our proofs. First, we review the Lebesgue Dominated Convergence Theorem, which provides a condition under which a limit can be brought inside of an integral.

Theorem C.2.1 (Lebesgue Dominated Convergence Theorem [299, Chapter 18.3])**.** *Let (X, Σ, μ) be a measure space and $(f_n)_{n\in\mathbb{N}}$ a sequence of measurable functions on X for which $f_n \to f$ pointwise almost everywhere on X and the function f is measurable. Assume there is a nonnegative function g that is integrable over X and dominates the sequence $(f_n)_{n\in\mathbb{N}}$ on X in the sense that $|f_n| \leq g$ almost everywhere on X for all n. Then f is integrable over X and $\lim_{n\to\infty} \int_X f_n \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu$.*

As a simple example of the Lebesgue Dominated Convergence Theorem, consider the sequence of functions $f_n(x) = \mathrm{e}^{-nx^2}$ on $\mathbb{R}$. $(f_n)_{n\in\mathbb{N}}$ converges pointwise to the zero function, because for every fixed x, $\lim_{n\to\infty} f_n(x) = 0$. Every f_n is bounded above by $g(x) = \mathrm{e}^{-x^2}$ for all x, and the integral of g over $\mathbb{R}$ is finite. Hence, $\lim_{n\to\infty} \int_{\mathbb{R}} f_n \, \mathrm{d}\mu = \int_{\mathbb{R}} \lim_{n\to\infty} f_n \, \mathrm{d}\mu = 0$. If we instead just compute the integral, we find that $\int_{\mathbb{R}} f_n \, \mathrm{d}\mu = \sqrt{\pi/n}$, which indeed goes to zero as $n \to \infty$.

Next, we state the Riesz Weak Compactness Theorem, which forms the backbone of our proof that continuous-variable state t-designs do not exist for $t \geq 2$.

Theorem C.2.2 (Riesz Weak Compactness Theorem [299, Chapter 19.5]). *Let (X, Σ, μ) be a σ-finite measure space. Let $1 < t < \infty$ and t' such that $1/t + 1/t' = 1$. If $(p_i)_{i=0}^{\infty}$ is a bounded sequence in $L^t(X) = L^t(X, \Sigma, \mu)$, then there exist a subsequence $(p_{i_k})_{k=0}^{\infty}$ of $(p_i)_{i=0}^{\infty}$ and a function $q \in L^t(X)$ for which*

$$\forall h \in L^{t'}(X)\colon \quad \lim_{k\to\infty} \int_X p_{i_k} h \,\mathrm{d}\mu = \int_X qh \,\mathrm{d}\mu. \tag{C.6}$$

Indeed, this theorem will be the main ingredient in our proof of the non-existence of continuous-variable t-designs. Notice that this theorem does not hold for $t = 1$, but rather $t > 1$. This will ultimately be the reason why our proof of non-existence of continuous-variable state t-designs only holds for $t \geq 2$. This is a nice sanity check, since Example C.4.4 shows explicit examples of continuous-variable state 1-designs. The proof of the Riesz Weak Compactness Theorem uses that $L^t(X)$ is a reflexive Banach space for all $1 < t < \infty$. For each such t, $L^{t'}(X)$ is naturally isomorphic to the dual space of $L^t(X)$. However, $L^1(X)$ is not the dual of $L^\infty(X)$. Since the Riesz Weak Compactness Theorem is so important for this work, we present a simple example to help understand the theorem.

Example C.2.3. Suppose $(f_n)_{n\in\mathbb{N}}$ is a sequence of functions $f_n\colon [0,1] \to \mathbb{R}$ defined by $f_n = \sqrt{n}\mathbb{1}_{[0,1/n]} \in L^2([0,1])$, where the indicator function $\mathbb{1}_{[a,b]}(x)$ on an interval $[a,b]$ is 1 for any x in the interval, and 0 elsewhere. The norm of f_n is $\|f_n\|_2^2 = \int_{[0,1]} |f_n|^2 \,\mathrm{d}\mu = n\mu([0,1/n]) = 1$. Hence, $(f_n)_{n\in\mathbb{N}}$ is a bounded sequence in $L^2([0,1])$. Therefore, there is a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ and a function $q \in L^2([0,1])$ for which

$$\forall h \in L^2([0,1])\colon \quad \lim_{k\to\infty} \int_{[0,1]} f_{n_k} h \,\mathrm{d}\mu = \int_{[0,1]} qh \,\mathrm{d}\mu. \tag{C.7}$$

Consider the constant function $h = 1 \in L^2([0,1])$. We can explicitly compute the left-

hand side to be $\lim_{k\to\infty} \int_{[0,1]} \sqrt{n_k} \mathbb{1}_{[0,1/n_k]} = \lim_{k\to\infty} \sqrt{n_k}/n_k = 0$. Similarly, one can consider functions $h_{a,b} = \mathbb{1}_{[a,b]}$ for any $a \leq b \in [0,1]$ and compute the left-hand side to be zero for all choice of a and b. Therefore, it must be that $\int_{[0,1]} q \mathbb{1}_{[a,b]} \, \mathrm{d}\mu = 0$ for all a and b, meaning that q must be zero almost everywhere.

One says that the subsequence (f_{n_k}) converges *weakly* to 0 in $L^2([0,1])$. This is to be contrasted with strong convergence. If (f_{n_k}) were to converge strongly to 0 in $L^2([0,1])$, then $\lim_{k\to\infty} \int_{[0,1]} |f_{n_k} - 0|^2 \, \mathrm{d}\mu = 0$, which is clearly not the case. It is not hard to see that the full sequence (f_n) also converges weakly to 0.

Consider instead the sequence $(f_n)_{n\in\mathbb{N}}$ defined by $f_n = (-1)^n$. The sequence is bounded in $L^2([0,1])$. Therefore, there is a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ that converges weakly to a function $q \in L^2([0,1])$. This example shows why the Riesz Weak Compactness Theorem only proves that a *subsequence* weakly converges to q, as opposed to the whole sequence. In this case, one can take the subsequence of even n so that $f_n = 1$ or odd n so that $f_n = -1$. These subsequences then converge weakly (and strongly) to 1 and -1 respectively, but the full sequence $(f_n)_{n\in\mathbb{N}}$ does not converge weakly to anything in $L^2([0,1])$. ◇

Notice that the Riesz Weak Compactness Theorem requires a σ-finite measure space. Thus, in order to use the theorem, we will need to be able to ensure that our measure space is σ-finite. The following lemma will allow us to do this.

Lemma C.2.4 ([299, Ch. 18.2, Prop. 9])**.** *Let (X, Σ, μ) be a measure space and $f\colon X \to [0,\infty]$ a nonnegative integrable function on X. Then f is finite almost everywhere and the set $\{x \in X \mid f(x) > 0\}$ is σ-finite.*

Haar measure

Suppose $(G, \cdot)$ is a compact Hausdorff topological group. Let Σ be the Borel σ-algebra on G; that is, the smallest σ-algebra that contains all open sets of G. A measure μ is called *left-invariant* if $\mu(A) = \mu(gA)$ for all $g \in G$ and $A \in \Sigma$. The Haar measure on G is the

unique left-invariant measure satisfying $\mu(G) = 1$. The finite dimensional unitary group $G = \mathrm{U}(d)$ is compact, as therefore can be equipped with the measure μ_{Haar} satisfying

$$\int_{\mathrm{U}(d)} \mathrm{d}\mu_{\mathrm{Haar}}(U) = 1, \qquad \int_{\mathrm{U}(d)} f(U)\,\mathrm{d}\mu_{\mathrm{Haar}}(U) = \int_{\mathrm{U}(d)} f(VU)\,\mathrm{d}\mu_{\mathrm{Haar}}(U) = \int_{\mathrm{U}(d)} f(UV)\,\mathrm{d}\mu_{\mathrm{Haar}}(U) \tag{C.8}$$

for any $V \in \mathrm{U}(d)$.

The Haar measure on $\mathrm{U}(d)$ induces a unitarily invariant measure on complex-projective space $\mathbb{CP}^{d-1}$. The construction is summarized as follows [72, 239, 302]. The unitary group is defined on $\mathbb{C}^d$ with respect to an inner product. The unit sphere $S^{2d-1} \subset \mathbb{C}^d$ can be viewed as an embedding into $\mathbb{C}^d$, and consists of all unit normalized vectors in $\mathbb{C}^d$. The inner product on $\mathbb{C}^d$ remains defined on S^{2d-1}. The set of all quantum states in $\mathbb{C}^d$ is $\mathbb{CP}^{d-1} := S^{2d-1}/\mathrm{U}(1)$. Modding out by $\mathrm{U}(1)$ represents the irrelevance of a global phase factor. In particular, S^{2d-1} can be viewed as a fiber-bundle, with $\mathbb{CP}^{d-1}$ the base space and $\mathrm{U}(1)$ the fiber on top of each point in $\mathbb{CP}^{d-1}$. The bundle projection $\pi\colon S^{2d-1} \to \mathbb{CP}^{d-1}$ induces a map between the tangent spaces $\pi_*\colon T_p S^{2d-1} \to T_{\pi(p)}\mathbb{CP}^{d-1}$ (the pushforward map). With some care, π_* can be used to construct a Hermitian metric on $T\mathbb{CP}^{d-1}$ via the inner product on S^{2d-1}. The real part of such a metric defines a Riemannian metric g, which can then be used to define a volume form, called the Fubini-Study volume form, on $\mathbb{CP}^{d-1}$ in the usual way. By construction, the Fubini-Study volume form is unitarily invariant with respect to the definition of unitary via the inner product on S^{2d-1}. In particular, π is a Riemannian submersion and the resulting metric is the unique unitarily invariant metric up to scaling. In this way, we have defined a unitarily invariant measure on $\mathbb{CP}^{d-1}$. Integrals with respect to this measure are denoted $\int_{\mathbb{CP}^{d-1}} f(\psi)\,\mathrm{d}\psi$. By unitary invariance, $\int_{\mathbb{CP}^{d-1}} f(U\psi)\,\mathrm{d}\psi = \int_{\mathbb{CP}^{d-1}} f(\psi)\,\mathrm{d}\psi$ for all $U \in \mathrm{U}(d)$.

A (nonfinite) Haar measure can also be defined on non-compact groups provided that they are locally compact. We will not discuss this fact much here, other than to say that the unitary group on $L^2(\mathbb{R})$, $\mathrm{U}(L^2(\mathbb{R}))$, does not have a Haar measure since it is not locally compact [242, sec. 5]. Hence, there is no natural way to integrate over all unitaries acting

on the space of continuous-variable quantum states.

Subsection C.2.2: Projector onto symmetric subspace

In this section, we summarize the analytical form of projectors onto the symmetric subspace of a separable Hilbert space $\mathcal{H}$, either finite- or infinite dimensional. Let S_t denote a group of permutations of t elements. For any $\sigma \in S_t$, let $W_\sigma \colon \mathcal{H}^{\otimes t} \to \mathcal{H}^{\otimes t}$ denote a unitary operator that transforms the Fock basis as follows:

$$W_\sigma \left|n_1\right\rangle \otimes \cdots \otimes \left|n_t\right\rangle = \left|n_{\sigma^{-1}(1)}\right\rangle \otimes \cdots \otimes \left|n_{\sigma^{-1}(t)}\right\rangle \ . \tag{C.9}$$

Let $\Pi_t \colon \mathcal{H}^{\otimes t} \to \mathcal{H}^{\otimes t}$ denote the projector onto the symmetric subspace of $\mathcal{H}^{\otimes t}$, i.e., the subspace isomorphic to the quotient space $\mathcal{H}^{\otimes t}/\{v - W_\sigma v \mid v \in \mathcal{H}^{\otimes t}, \sigma \in S_t\}$. Using W_σ, Π_t can be defined as follows.

Claim C.2.5. *For each $t \in \mathbb{N}$, let Π_t denote a projector onto the symmetric subspace of $\mathcal{H}^{\otimes t}$. Then*

$$\Pi_t = \frac{1}{|S_t|} \sum_{\sigma \in S_t} W_\sigma \ . \tag{C.10}$$

For completeness, we outline an algebraic proof of Π_t. It can also be proven using group-theoretic tools, as is shown in e.g. [45, sec. 2], and using linear algebra, as shown in [240, prop. 1]. For a complete discussion on the symmetric projector, we refer to Harrow's "The Church of the Symmetric Subspace" [240].

Proof. We denote the set of permutations of a vector $v \in \mathcal{H}^{\otimes t}$ as $P(v) \coloneqq \{W_\sigma v \mid \sigma \in S_t\}$, where W_σ is given by Eq. (C.9). Let $\mathcal{B} = \{|i\rangle \mid i \in \{1, \ldots, \dim \mathcal{H}\}\}$ be an orthonormal basis of $\mathcal{H}$, and $\mathcal{B}^{\otimes t}$ the corresponding orthonormal basis of $\mathcal{H}^{\otimes t}$.

As an example, suppose $t = 4$ and $v = |1\rangle \otimes |2\rangle \otimes |1\rangle \otimes |1\rangle \in \mathcal{B}^{\otimes t}$. We will use the notation $v!$ to mean $v! = 3! \cdot 1!$, since $|1\rangle$ occurs three times and $|2\rangle$ one time. Similarly, suppose $t = 5$

and $v = |5\rangle^{\otimes 3} \otimes |1\rangle^{\otimes 2}$. Then $v! = 3! \cdot 2!$. One can then verify that

$$\mathcal{B}' = \left\{ \frac{1}{\sqrt{v!\,|S_t|}} \sum_{\sigma \in S_t} W_\sigma v \mid v \in \mathcal{B}^{\otimes t} \right\} , \tag{C.11}$$

is an orthonormal basis of the symmetric subspace of $\mathcal{H}^{\otimes t}$. For any $v \in \mathcal{B}^{\otimes t}$ and $u \in \mathcal{B}'$, the following holds:

$$\langle v|u\rangle = \sqrt{\frac{v!}{|S_t|}} \delta_{u \in \mathrm{span}(P(v))} = \begin{cases} \sqrt{\frac{v!}{|S_t|}} & \text{if } u \in \mathrm{span}(P(v)) \\ 0 & \text{otherwise.} \end{cases} \tag{C.12}$$

Using one dimensional subspaces spanned by the basis vectors in $\mathcal{B}'$, the projector onto the symmetric subspace of $\mathcal{H}^{\otimes t}$ can be represented as

$$\Pi_t = \sum_{u \in \mathcal{B}'} |u\rangle\langle u| \ . \tag{C.13}$$

Next, we determine the matrix elements of Π_t in the basis $\mathcal{B}$. Let $v, w \in \mathcal{B}$. Then

$$\begin{aligned} \langle v| \Pi_t |w\rangle &= \sum_{u \in \mathcal{B}'} \langle v|u\rangle\langle u|w\rangle && \text{(C.14a)} \\ &= \frac{\sqrt{v!w!}}{|S_t|} \sum_{u \in \mathcal{B}'} \delta_{u \in \mathrm{span}(P(v))} \delta_{u \in \mathrm{span}(P(w))} && \text{(C.14b)} \\ &= \frac{v!}{|S_t|} \delta_{v \in P(w)}. && \text{(C.14c)} \end{aligned}$$

Finally, consider the matrix elements of $\frac{1}{|S_t|} \sum_{\sigma \in S_t} W_\sigma$:

$$\begin{aligned} \langle v| \left(\frac{1}{|S_t|} \sum_{\sigma \in S_t} W_\sigma \right) |w\rangle &= \frac{1}{|S_t|} \sum_{\sigma \in S_t} \langle v| W_\sigma |w\rangle && \text{(C.15a)} \\ &= \frac{v!}{|S_t|} \delta_{v \in P(w)}, && \text{(C.15b)} \end{aligned}$$

which proves the claim. □

We now define some more notation for the matrix elements of the symmetric projector.

Definition C.2.6. *For any tuples $a, b \in \mathbb{N}_0^t$, define*

$$\Pi_t(a;b) \equiv \Pi_t(a_1,\ldots,a_t;b_1,\ldots,b_t) \tag{C.16a}$$

$$\coloneqq \left(\bigotimes_{i=1}^{t} \langle a_i|\right) \Pi_t \left(\bigotimes_{i=1}^{t} |b_i\rangle\right), \tag{C.16b}$$

and define

$$\Lambda_t(a) \equiv \Lambda_t(a_1,\ldots,a_t) \coloneqq \Pi_t(a;a). \tag{C.17}$$

For this appendix, we will need some properties of Λ. Clearly, $\Lambda_t(a) > 0$ for all tuples $a \in \mathbb{N}_0^t$. Similarly, $\Lambda_t(a) = \Lambda_t(\sigma(a))$ for any $\sigma \in S_t$. Additionally, $\Lambda_t(a_1,\ldots,a_t) = \Lambda_t(a_1+1,\ldots,a_t+1)$. Finally, for a tuple $b \in \mathbb{N}_0^{t-1}$ and a number $i \in \mathbb{N}_0$, denote the direct sum as the tuple $b \oplus (i) \coloneqq (b_1,\ldots,b_{t-1},i)$. For any fixed $b \in \mathbb{N}_0^{t-1}$, there exists an $N \in \mathbb{N}$ such that for all $m, m' > N$: $\Lambda_t(b \oplus (m)) = \Lambda_t(b \oplus (m'))$. This means that, for example

$$\lim_{m\to\infty} \Lambda_t(0,\ldots,0,m) > 0. \tag{C.18}$$

To get a handle on these definitions and properties, consider the example for $t = 2$.

Example C.2.7. In the case of $t = 2$,

$$\Pi_2 = \frac{1}{2}\left(W_{(1)(2)} + W_{(12)}\right) = \frac{1}{2}\left(\mathbb{I} + S\right), \tag{C.19}$$

where S is the SWAP operator. We used cyclic notation for permutations, so that $(1)(2)$ is the identity permutation and (12) is the other permutation in S_2. One can also find Π_2 by

summing over projectors onto an orthonormal set of symmetric states, as

$$\begin{aligned}\Pi_2 = &|0\rangle|0\rangle\langle 0|\langle 0| \\ &+ \left(\frac{|0\rangle|1\rangle + |1\rangle|0\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0|\langle 1| + \langle 1|\langle 0|}{\sqrt{2}}\right) \\ &+ |1\rangle|1\rangle\langle 1|\langle 1| \\ &+ \left(\frac{|0\rangle|2\rangle + |2\rangle|0\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0|\langle 2| + \langle 2|\langle 0|}{\sqrt{2}}\right) + \dots .\end{aligned} \tag{C.20}$$

Therefore,

$$\Pi_2(a_1, a_2; b_1, b_2) = \frac{1}{2}\left(\delta_{a_1 b_1}\delta_{a_2 b_2} + \delta_{a_1 b_2}\delta_{a_2 b_1}\right), \tag{C.21}$$

and $\Lambda_2(a_1, a_2) = \frac{1}{2}(1 + \delta_{a_1 a_2})$. ◇

Section C.3: Finite dimensional designs

Subsection C.3.1: Simplex designs

A simplex t-design, more commonly referred to as a *(positive, interior) simplex cubature rule* in the literature [185, 187–190, 303], is a set of points on the simplex and a weight function that exactly integrates polynomials of degree t or less.

Definition C.3.1 (Simplex design)**.** *Let $P \subset \Delta^m$ be a finite set, and $u\colon P \to \mathbb{R}_{>0}$ be a weight function on P. Let* dp *denote the standard unit normalized Lebesgue measure on the simplex. The pair (P, u) is called an m dimensional* simplex t-design *if, for all tuples $a = (a_1, \dots, a_t) \in \{0, 1, \dots, m\}^t$,*

$$\sum_{q \in P} w(q) \prod_{i=1}^{t} q_{a_i} = \int_{\Delta^m} \prod_{i=1}^{t} p_{a_i}\, \mathrm{d}p. \tag{C.22}$$

The pair (P, u) defines a probability ensemble, and we will therefore define $\mathbb{E}_{q \in P}\, g(q) \coloneqq \sum_{q \in P} w(q) g(q)$ for any function g.

Since the coordinates of a point on the simplex sum to 1, by summing over one of the a_i on both sides, we find that a simplex t-design is automatically a simplex $(t-1)$-design. The measure $\mathrm{d}p$ is proportional to $\delta(1-p_0-\cdots-p_n)\,\mathrm{d}p_0\ldots\mathrm{d}p_m$. For $\beta\in\mathbb{N}_0^{m+1}$ and $p\in\Delta^m$, define $p^\beta \coloneqq \prod_{i=0}^m p_i^{\beta_i}$. For example, if $m=2$ and $\beta=(0,2,1)$, then $p^\beta=p_1^2p_2$. Then one can compute the moments from the Dirichlet distribution [185]

$$\int_{\Delta^m} p^\beta\,\mathrm{d}p = \frac{m!}{(m+\beta_0+\cdots+\beta_m)!}\prod_{i=0}^m \beta_i!. \tag{C.23}$$

We will list various simplex t-designs. We will use the notation

$$f^{(i)} = (\underbrace{0,\ldots,0}_{i},1,\underbrace{0,\ldots,0}_{m-i}), \tag{C.24}$$

so that $f_j^{(i)}=\delta_{ij}$. In this way, a point $p\in\Delta^m$ is written as $p=\sum_{i=0}^m p_i f^{(i)}$. Denote the centroid of the simplex by $c=\frac{1}{m+1}\sum_{i=0}^m f^{(i)}$.

Theorem C.3.2 (Extremal points of the unit m-simplex form a one-design)**.** *Let P be the set $P=\{f^{(i)}\mid i\in\{0,\ldots,m\}\}$, and u the constant map $w(f^{(i)})=1/(m+1)$. The pair (P,u) is a simplex one-design.*

Proof. We must prove that $\frac{1}{m+1}\sum_{i=0}^m g(f^{(i)}) = n!\int_{\Delta^m} g(p)\,\mathrm{d}p$ for any linear polynomial $g(p)=p_j$. The left-hand side is then $\frac{1}{m+1}\sum_{i=0}^m\delta_{ij}=1/(m+1)$, and the right-hand side is $\frac{1}{m+1}$ by Eq. (C.23). □

Theorem C.3.3 (Centroid of the unit m-simplex forms a one-design)**.** *Let P be the set $P=\{c\}$ and u the map $u(c)=1$. Then the pair (P,u) is a simplex one-design.*

Proof. Clearly $g(c)=1/(m+1)$ for any linear polynomial $g(p)=p_j$. □

Theorem C.3.4 (Extremal points plus the centroid of the unit m-simplex form a two-design)**.** *Let P be the set $P=\{c\}\cup\{f^{(i)}\mid i\in\{0,\ldots,m\}\}$, and u the map defined by $u(c)=\frac{m+1}{m+2}$ and $u(f^{(i)})=\frac{1}{(m+1)(m+2)}$. Then the pair (P,u) is a simplex two-design.*

Proof. It suffices to prove that

$$\frac{m+1}{m+2}g(c) + \frac{1}{(m+1)(m+2)}\sum_{i=0}^{m} g(f^{(i)}) = m!\int_{\Delta^m} g(p)\,\mathrm{d}p \tag{C.25}$$

for any quadratic polynomial $g(p) = p_j p_k$. By Eq. (C.23), the right-hand side equals $\frac{1+\delta_{jk}}{(m+1)(m+2)}$. The left-hand side is

$$\frac{m+1}{m+2}g(c) + \frac{1}{(m+1)(m+2)}\sum_{i=0}^{m} g(f^{(i)}) = \frac{1}{(m+1)(m+2)} + \frac{1}{(m+1)(m+2)}\sum_{i=0}^{m}\delta_{ij}\delta_{ik} \tag{C.26a}$$

$$= \frac{1+\delta_{jk}}{(m+1)(m+2)}, \tag{C.26b}$$

as desired. □

Theorem C.3.5 (Simplex two-design [187, thm. 2], [188, cor. 4.1])**.** *Let $r = 1/\sqrt{m+2}$. Let $v^{(i)} = rf^{(i)} + (1-r)c$. Let P be the set $P = \{v^{(i)} \mid i \in \{0,\dots,m\}\}$, and u the constant map $u(v^{(i)}) = 1/(m+1)$. Then the pair (P,u) is a simplex two-design.*

The simplex 2-design in Theorem C.3.5 utilizes $m+1$ points in Δ^m, which is in fact the best that can be done [303, tab. 1].

Subsection C.3.2: Torus designs

We define torus designs analogously to simplex designs. We let $T = [0, 2\pi)$.

Definition C.3.6. *Let $S \subset T^m$ be a finite set, and $v\colon S \to \mathbb{R}_{>0}$ be a weight function on S. Let $\mathrm{d}\phi$ denote the standard unit normalized Lebesgue measure on the torus. The pair (S,v) is called an m dimensional* torus t-design *if, for all tuples $a = (a_1,\dots,a_t) \in \{1,2,\dots,m\}^t$*

and $b = (b_1, \ldots, b_t) \in \{1, 2, \ldots, m\}^t$,

$$\sum_{\theta \in S} v(\theta) \prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\theta_{a_i} - \theta_{b_i})} = \int_{T^m} \prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\phi_{a_i} - \phi_{b_i})} \, \mathrm{d}\phi. \tag{C.27}$$

The pair (S, v) *defines a probability ensemble, and we will therefore define* $\mathbb{E}_{\theta \in S} \, g(\mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_m}) := \sum_{\theta \in S} v(\theta) g(\mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_m})$ *for any function* g.

It follows from the definition that a torus t-design is always a torus $(t-1)$-design. For example, suppose $t = 2$, and let $a = (1, j)$ and $b = (1, k)$ for any j and k. Then it is clear that the 2-design (S, v) also satisfies the 1-design condition. By definition, a torus t-design must match integration on polynomials $g(s) = g(s_1, \ldots, s_m)$ that are degree t in s and degree t in degree $\bar{s}$. One could generalize the definition to match integration on polynomials that are degree t in s and degree t' in $\bar{s}$. We call the corresponding sets (t, t') torus designs. In this way, a torus t-design is a shorthand notation for a (t, t)-design.

The definition of a torus t-design closely resembles the definition of a *trigonometric cubature rule* [303], however they are not equivalent. To the best of our knowledge, the notion of general torus cubature was first proposed in Ref. [189], where it was formulated as a generalization of trigonometric cubature rules in terms of algebraic tori. Our definition of a T^m t-design corresponds to the definition in Ref. [189] of an order t cubature rule on the maximal torus $T(\mathrm{PSU}(m+1)) \cong T^m$ with an algebraic structure given by a faithful orbit of its linear action by conjugation on the vector space of $(m+1) \times (m+1)$ complex matrices. Here, $\mathrm{PSU}(m+1)$ is the projective special unitary group, which is the special unitary group $\mathrm{SU}(m+1)$ modulo its center. In Appendix C.6, we show the equivalence of the two definitions, as well as comment on the relationship to standard trigonometric cubature and to complete sets of mutually unbiased bases.

We now construct various torus designs.

Theorem C.3.7 (1-design on the m-torus)**.** *Let S be the set*

$$S = \{(0, 2\pi q/m, 2\pi 2q/m, \ldots, 2\pi(n-1)q/m) \mid q \in \mathbb{Z}_m\}, \tag{C.28}$$

and v the constant map $v(\phi) = 1/m$. Then the pair (S, v) is an m-torus 1-design.

Proof. It is sufficient to check for $g(s) = s_a \bar{s}_b$.

$$\frac{1}{m} \sum_{q \in \mathbb{Z}_m} \mathrm{e}^{2\pi \mathrm{i} aq/m} \mathrm{e}^{-2\pi \mathrm{i} bq/m} = \frac{1}{m} \sum_{q \in \mathbb{Z}_m} \mathrm{e}^{2\pi \mathrm{i} q(a-b)/m} = \delta_{ab}. \tag{C.29}$$

Meanwhile,

$$\int_{T^m} \mathrm{e}^{\mathrm{i}\phi_a} \mathrm{e}^{-\mathrm{i}\phi_b} \, \mathrm{d}\phi = \frac{1}{(2\pi)^2} \int_{T^2} \mathrm{e}^{\mathrm{i}(\phi_a - \phi_b)} \, \mathrm{d}\phi_a \, \mathrm{d}\phi_b = \delta_{ab}. \tag{C.30}$$

□

Theorem C.3.8 (t-design on the m-torus (concatenation of t-designs on each factor of S^1))**.** *Let S be the set*

$$S = \left\{(2\pi d_1/(t+1), 2\pi d_2/(t+1), \ldots, 2\pi d_m/(t+1)) \mid d \in \mathbb{Z}_{t+1}^m\right\}, \tag{C.31}$$

and v the constant map $v(\phi) = (t+1)^{-m}$. Then the pair (S, v) is an m-torus t-design.

Proof. It is sufficient to check for $g(s) = s_{a_1} \ldots s_{a_t} \bar{s}_{b_1} \ldots s_{b_t}$.

$$\frac{1}{(t+1)^m} \sum_{d \in \mathbb{Z}_{t+1}^m} \exp\left[\frac{2\pi \mathrm{i}}{t+1}(d_{a_1} + \cdots + d_{a_t})\right] \exp\left[-\frac{2\pi \mathrm{i}}{t+1}(d_{b_1} + \cdots + d_{b_t})\right] = \begin{cases} 1 & \text{if } a \text{ is a permutation of } b \\ 0 & \text{otherwise.} \end{cases} \tag{C.32}$$

Meanwhile,

$$\int_{T^m} \mathrm{e}^{\mathrm{i}(\phi_{a_1} + \cdots + \phi_{a_t})} \mathrm{e}^{-\mathrm{i}(\phi_{b_1} + \cdots + \phi_{b_t})} \, \mathrm{d}\phi = \begin{cases} 1 & \text{if } a \text{ is a permutation of } b \\ 0 & \text{otherwise.} \end{cases}. \tag{C.33}$$

□

Theorem C.3.9 (Efficient 2-design on the m-torus)**.** *Define p to be the smallest prime number strictly larger than $\max(2, m)$ (by the prime number theorem, $p \in \mathcal{O}(m + \log m)$). Let S be the set*

$$S = \left\{ \left(0, 2\pi(q_1 + q_2)/p, 2\pi(2q_1 + 4q_2)/p, \ldots, 2\pi((m-1)q_1 + (m-1)^2 q_2)/p\right) \mid q_1 \in \mathbb{Z}_p, q_2 \in \mathbb{Z}_p \right\} \tag{C.34}$$

and v the constant map $v(\phi) = 1/p^2$. Then the pair (S, v) is an m-torus 2-design.

Proof. It suffices to prove that

$$\int_{T^m} \mathrm{e}^{\mathrm{i}(\phi_i + \phi_j - \phi_k - \phi_l)} \, \mathrm{d}\phi = \frac{1}{p^2} \sum_{\theta \in S} \mathrm{e}^{\mathrm{i}(\theta_i + \theta_j - \theta_k - \theta_l)}. \tag{C.35}$$

The right-hand side is

$$\frac{1}{p^2} \sum_{\theta \in S} \mathrm{e}^{\mathrm{i}(\theta_i + \theta_j - \theta_k - \theta_l)} = \frac{1}{p^2} \sum_{q_1, q_1 \in \mathbb{Z}_p} \mathrm{e}^{\frac{2\pi \mathrm{i}}{p} q_1 (i + j - k - l)} \mathrm{e}^{\frac{2\pi \mathrm{i}}{p} q_2 (i^2 + j^2 - k^2 - l^2)} \tag{C.36a}$$

$$= \delta_{i+j, k+l} \delta_{i^2 + j^2, k^2 + l^2} \tag{C.36b}$$

$$= \begin{cases} 1 & \text{if } i = l \wedge j = k \text{ or } i = k \wedge j = l \\ 0 & \text{otherwise,} \end{cases} \tag{C.36c}$$

where we used Lemma C.3.10 in the last line. The left-hand side is

$$\int_{T^m} \mathrm{e}^{\mathrm{i}(\phi_i + \phi_j - \phi_k - \phi_l)} \, \mathrm{d}\phi = \frac{1}{(2\pi)^m} \int_{[0, 2\pi]^m} \mathrm{e}^{\mathrm{i}(\phi_i + \phi_j - \phi_k - \phi_l)} \, \mathrm{d}\phi_1 \ldots \mathrm{d}\phi_m \tag{C.37a}$$

$$= \begin{cases} 1 & \text{if } i = l \wedge j = k \text{ or } i = k \wedge j = l \\ 0 & \text{otherwise,} \end{cases} \tag{C.37b}$$

which is equal to the right-hand side. □

The torus 2-design in Theorem C.3.9 utilizes what Ref. [304] calls the "$ax^2 + bx$ construc-

tion" that is utilized in constructions of complete sets of mutually unbiased bases. From Ref. [41], it is known that such sets form complex-projective 2-designs. Hence, we can now understand the "$ax^2 + bx$ construction" as a torus 2-design. The "$ax^2 + bx$" construction utilizes the following Diophantine system.

Lemma C.3.10. *Let $\mathbb{F}_p$ be the finite field with p elements for an odd prime p. Let F be either $\mathbb{F}_p$ or $\mathbb{Z}$, and let addition, multiplication, and equality be with respect to F (e.g. for $F = \mathbb{F}_p$, $a = b$ is the same as $a \equiv b \pmod{p}$). The Diophantine system of equations*

$$a + b = c + d, \quad a^2 + b^2 = c^2 + d^2 \tag{C.38}$$

is solved only by solutions of the form

$$(a = c) \wedge (b = d), \qquad \text{or} \quad (a = d) \wedge (b = c). \tag{C.39}$$

Proof. Plugging the first equation into the second equation, we find that $(c + d - a)^2 + b^2 = c^2 + d^2$. Simplifying yields $ab = cd$. If $b = 0$, then either c or d must equal zero, so that the solution is of the desired form. If $b \neq 0$, then

$$ab + b^2 = bc + bd \tag{C.40a}$$

$$cd + b^2 = bc + bd \tag{C.40b}$$

$$(d - b)(c - b) = 0. \tag{C.40c}$$

Therefore, either $b = c$ or $b = d$. Along with $a + b = c + d$, this proves the claim. □

Theorem C.3.9 can be generalized to the case where we allow p to be any positive integer power of a prime, because Lemma C.3.10 can be generalized to the case of any Galois (finite) field.

An m-torus 1-design trivially requires at least m elements. To conclude this subsection, we show that an m-torus 2-design requires at least $m(m-1)+1$ elements.

Proposition C.3.11. *Let (S, v) be a T^m 2-design. Then $|S| > m(m-1)$.*

Proof. The torus 2-design condition can be expressed as follows. Let

$$\Gamma = \{(0,\dots,0),(1,-1,0,\dots,0),(1,0,-1,0,\dots,0),\dots,(-1,1,0,\dots,0),\dots,(0,\dots,0,-1,1)\}, \tag{C.41}$$

so that $|\Gamma| = m(m-1)+1$. Let each $\phi \in S$ label a basis element of $V := \mathbb{C}^{|S|}$ so that $\{|\phi\rangle \mid \phi \in S\}$ is an orthonormal basis of V. Then for $k \in \Gamma$, define $|k\rangle = \sum_{\phi\in S} \sqrt{v(\phi)} e^{ik\cdot\phi} |\phi\rangle$. The 2-design condition is summed up by $\langle k|k'\rangle = \delta_{kk'}$. Hence, $\{|k\rangle \mid k \in \Gamma\}$ must be orthonormal in V, meaning that $|\Gamma| \leq \dim V = |S|$. □

Subsection C.3.3: Complex-projective Haar integral

For integration over the set of d dimensional qudit states $\mathbb{CP}^{d-1}$, one finds [43, 45]

$$\int_{\mathbb{CP}^{d-1}} (|\psi\rangle\langle\psi|)^{\otimes t} \, d\psi = \frac{\Pi_t^{(d)}}{\operatorname{Tr} \Pi_t^{(d)}}, \tag{C.42}$$

where $d\psi$ denotes the unitarily invariant Fubini-Study volume form on the complex-projective space $\mathbb{CP}^{d-1}$, and $\Pi_t^{(d)}\colon (\mathbb{C}^d)^{\otimes t} \to (\mathbb{C}^d)^{\otimes t}$ is the projector onto the symmetric subspace of $(\mathbb{C}^d)^{\otimes t}$ defined in Definition C.2.6 [72, sec. 4.5, 4.7, 7.6] [239, ex. 8.8]. We will begin by showing this equality.

Each $|\psi\rangle$ lives in the finite dimensional Hilbert space $\mathcal{H} = \mathbb{C}^d$. For any integer t, the tensor product $\mathcal{H}^{\otimes t}$ splits up into a direct sum of the symmetric and antisymmetric subspaces of $\mathcal{H}^{\otimes t}$, so that $\mathcal{H}^{\otimes t} \cong \mathcal{H}_t^{\text{sym}} \oplus \mathcal{H}_t^{\text{asym}}$. Consider the representation of the group of unitaries acting on $\mathcal{H}$, $\rho\colon \mathrm{U}(\mathcal{H}) \to \mathrm{U}(\mathcal{H}^{\otimes t})$, defined by $U \mapsto U^{\otimes t}$. The subspaces $\mathcal{H}_t^{\text{sym}}$ and $\mathcal{H}_t^{\text{asym}}$ are invariant under ρ. One can see this by noting that for any unitary U, $\rho(U)\mathcal{H}_t^{(\text{a})\text{sym}} = \mathcal{H}_t^{(\text{a})\text{sym}}$

since $U^{\otimes t}$ acts symmetrically on the tensor product factors. Therefore, the representation ρ can be decomposed into a direct sum of irreducible representations on $\mathcal{H}_t^{\text{sym}}$ and $\mathcal{H}_t^{\text{asym}}$.

We can now invoke Schur's lemma, which states that if a nonzero operator M on an irrep space commutes with every element of that irrep, then M is proportional to the identity on that irrep space. In our case, $M = \int_{\mathbb{CP}^{d-1}} (|\psi\rangle\langle\psi|)\, \mathrm{d}\psi$. The irrep space of interest is $\mathcal{H}_t^{\text{sym}}$. The elements of the irrep are unitaries $U^{\otimes t}$. Due to the unitary invariance of the Fubini-Study metric, one finds that M commutes with all unitaries of the form $U^{\otimes t}$. Therefore, by Schur's lemma, M must be proportional to the identity on $\mathcal{H}_t^{\text{sym}}$, which is precisely $\Pi_t^{(d)}$. Finally, the Fubini-Study volume measure is normalized such that the volume of $\mathbb{CP}^{d-1}$ is unity. Hence, $\operatorname{Tr} M = 1$, meaning that the proportionality constant must be $1/\operatorname{Tr} \Pi_t^{(d)}$.

Next, we discuss integration over $\mathbb{CP}^m$ with respect to the Fubini-Study volume form where $m = d-1$, and show that it can be expressed as integration over a flat simplex and a flat torus. For a formal treatment of this fact, see [72, sec. 4.5, 4.7, 7.6] [239, ex. 8.8]. One first constructs the Fubini-Study volume form (see Appendix C.2.1). Then, one constructs a coordinate transformation mapping the simplex cross the torus to a coordinate patch of $\mathbb{CP}^m$. Pulling back the volume form along this coordinate transformation yields the volume form on the simplex cross the torus.

Here, we will instead give an informal treatment. Define $p_0 = 1 - \sum_{i=1}^m p_i$ and $\phi_0 = 0$. Then the p_i and ϕ_i parameterize a quantum state $\sqrt{p_0}\,|0\rangle + \sum_{i=1}^m \sqrt{p_i} \mathrm{e}^{\mathrm{i}\phi_i}\,|i\rangle$. To define a valid state, the p are elements of the probability simplex $\Delta^m = \{p = (p_0, \ldots, p_m) \in [0,1]^{m+1} \mid \sum_{i=0}^m p_i = 1\}$ and the ϕ are elements of the torus $T^m = [0, 2\pi)^m$. We'll denote the Lebesgue measure on Δ^m by $\mathrm{d}p = \prod_i \mathrm{d}p_i$, and on T^m by $\mathrm{d}\phi = \prod_{i=1}^m \mathrm{d}\phi_i$. One can easily perform the integration over the simplex and torus to find that $\frac{m!}{(2\pi)^m}\,\mathrm{d}p\,\mathrm{d}\phi$ is a normalized volume measure such that $\operatorname{vol}(\Delta^m \times T^m) = 1$. Consider a quantum state in $\mathbb{C}^{m+1}$ parameterized by $\alpha_i \in \mathbb{C}$ as $|\psi\rangle = \sum_{n=0}^m \alpha_n\,|n\rangle$, the natural measure is $\mathrm{d}^2\alpha_0 \ldots \mathrm{d}^2\alpha_m$. Applying the polar coordinate transformation $\alpha_n = \sqrt{p_n}\mathrm{e}^{\mathrm{i}\phi_n}$ and keeping track of Jacobian factors, the measure becomes

proportional to $\mathrm{d}p\,\mathrm{d}\phi$.

In conclusion, we have determined that

$$\int_{\mathbb{CP}^m} (|\psi\rangle\langle\psi|)^{\otimes t}\,\mathrm{d}\psi = \frac{m!}{(2\pi)^m}\int_{\Delta^m\times T^m} (|p,\phi\rangle\langle p,\phi|)^{\otimes t}\,\mathrm{d}p\,\mathrm{d}\phi, \tag{C.43}$$

where

$$|p,\phi\rangle = \sqrt{p_0}\,|0\rangle + \sum_{j=1}^{m}\sqrt{p_j}\mathrm{e}^{\mathrm{i}\phi_j}\,|j\rangle\,. \tag{C.44}$$

Subsection C.3.4: Complex-projective designs from simplex and torus designs

For finite d, an ensemble $\mathcal{E}$ over $\mathbb{CP}^{d-1}$ is a complex-projective t-design if

$$\mathop{\mathbb{E}}_{|\psi\rangle\in\mathcal{E}}\left[(|\psi\rangle\langle\psi|)^{\otimes t}\right] = \int_{\mathbb{CP}^{d-1}} (|\psi\rangle\langle\psi|)^{\otimes t}\,\mathrm{d}\psi. \tag{C.45}$$

Again let $m = d-1$. The characterization of the integral over $\mathbb{CP}^m$ given in Eq. (C.43) motivates the construction of complex-projective designs via constructions of simplex and torus designs. Such a construction was also noted in [189, thm. 4.1]. In particular, Eq. (C.43) consists of a product of integrals of the form $m!\int_{\Delta^m}\sqrt{\prod_{i=1}^{2t}p_{j_i}}\,\mathrm{d}p$ and $\frac{1}{(2\pi)^m}\int_{T^m}\exp\left[\mathrm{i}\sum_{i=1}^{t}(\phi_{j_i}-\phi_{j_{t+i}})\right]\mathrm{d}\phi$. The latter integral can be evaluated by a t-design on the torus, and is equal to 1 whenever $\sum_{i=1}^{t}(\phi_{j_i}-\phi_{j_{t+i}}) = 0$ regardless of ϕ, and zero otherwise. In other words, it is only nonzero when the j_i's are paired. But when the j_i's are paired, the term $\sqrt{\sum_{i=1}^{2t}p_{j_i}}$ becomes a monomial of degree t in p. Hence, the resulting integral can be evaluated with a simplex t-design. We summarize with the following theorem.

Theorem C.3.12. *Let P be a t-design on the m-simplex Δ^m, meaning that P is an ensemble over Δ^m such that*

$$\mathop{\mathbb{E}}_{q\in P}[g(q)] = \int_{\Delta^m} g(p)\,\mathrm{d}p \tag{C.46}$$

for any polynomial $g(p) = g(p_0,\ldots,p_m)$ of degree less than or equal to t. Similarly, let S be

a t-design on the m-torus T^m, meaning that S is an ensemble over T^m such that

$$\mathop{\mathbb{E}}_{\theta\in S}\left[g(\mathrm{e}^{\mathrm{i}\theta_1},\dots,\mathrm{e}^{\mathrm{i}\theta_m})\right] = \int_{T^m} g(\mathrm{e}^{\mathrm{i}\phi_1},\dots,\mathrm{e}^{\mathrm{i}\phi_m})\,\mathrm{d}\phi \tag{C.47}$$

for any polynomial $g(s) = g(s_1,\dots,s_m)$ of degree t in s and degree t in $\bar{s}$. Then $D = P \times S$ is a t-design on $\mathbb{CP}^m$, meaning that

$$\mathop{\mathbb{E}}_{(p,\phi)\in D}\left[(|p,\phi\rangle\langle p,\phi|)^{\otimes t}\right] = \int_{\mathbb{CP}^m} (|\psi\rangle\langle\psi|)^{\otimes t}\,\mathrm{d}\psi, \tag{C.48}$$

with $|p,\phi\rangle \coloneqq \sum_{j=0}^m \sqrt{p_j}\mathrm{e}^{\mathrm{i}\phi_j}\,|j\rangle$.

We can state this in terms of weight functions as follows. Let (P,u) be a Δ^{d-1} t-design, and (S,v) be a T^d t-design. Define $D \coloneqq \{|p,\phi\rangle \mid p\in P, \phi\in S\}$, and

$$w(|p,\phi\rangle) \coloneqq u(p) \sum_{\substack{\phi'\in S \text{ st}\\ |p,\phi\rangle=|p,\phi'\rangle}} v(\phi'). \tag{C.49}$$

Then (D,w) is a $\mathbb{CP}^{d-1}$ t-design. Morally, $w(|p,\phi\rangle)$ is essentially $u(p)v(\phi)$. However, the map $(p,\phi)\mapsto|p,\phi\rangle$ is not bijective; specifically, if p is on the boundary $\partial\Delta^{d-1}$, then for any ϕ there are many ϕ' satisfying $|p,\phi\rangle = |p,\phi'\rangle$. Therefore, the definition of w must be modified accordingly, as is done in Eq. (C.49).

We will now construct explicit complex-projective designs by concatenating simplex and torus designs given in Appendices C.3.1 and C.3.2. For this subsection, we will use the following notation for complex-projective t-designs. Fix a set $D\subset\mathbb{CP}^m$ of points in $\mathbb{CP}^m$, and let $w\colon D\to\mathbb{R}_{>0}$ be a weight function. The pair (D,w) is a complex-projective t-design if

$$\sum_{|\xi\rangle\in D} w(|\xi\rangle)\,(|\xi\rangle\langle\xi|)^{\otimes t} = \int_{\mathbb{CP}^m} (|\psi\rangle\langle\psi|)^{\otimes t}\,\mathrm{d}\psi. \tag{C.50}$$

Construction 1: Combining the simplex 2-design from Theorem C.3.4 and the torus

2-design from Theorem C.3.9, we find that for any $m \in \mathbb{N}_0$, the pair (D, w) is a complex-projective 2-design, where p is the smallest prime number strictly larger than $\max(2, m)$, D is the set

$$D = \{|i\rangle \mid i \in \{0, \ldots, m\}\} \cup \{|q_1, q_2\rangle \mid q_1, q_2 \in \mathbb{Z}_p\}\,, \tag{C.51}$$

and $w\colon D \to \mathbb{R}_{>0}$ is the map defined by $w(|i\rangle) = \frac{1}{(m+1)(m+2)}$ and $w(|q_1, q_2\rangle) = \frac{m+1}{(m+2)p^2}$. By the prime number theorem, $p \in \mathcal{O}(m + \log m)$. Here $|q_1, q_2\rangle := \frac{1}{\sqrt{m+1}} \sum_{j=0}^{m} \mathrm{e}^{2\pi \mathrm{i}(q_1 j + q_2 j^2)/p} |j\rangle$. When $m+1$ is prime, this reduces to the well-known complete set of mutually unbiased bases given in [39] (indeed this can be generalized to whenever $d = m+1$ is a prime power). For prime d, this complex-projective design is uniformly weighted. However, for non-prime d, the weights are not uniform.

Construction 2: We can construct a uniformly weighted complex-projective 2-design for all m that uses $p^2(m+1)$ points by combining the simplex 2-design from Theorem C.3.5 and the torus 2-design from Theorem C.3.9. Define $r = 1/\sqrt{m+2}$ and the state

$$|\ell, q_1, q_2\rangle := \sqrt{\frac{1+rm}{m+1}} \mathrm{e}^{2\pi \mathrm{i}(q_1 \ell + q_2 \ell^2)/p} |\ell\rangle + \sqrt{\frac{1-r}{m+1}} \sum_{j \neq \ell} \mathrm{e}^{2\pi \mathrm{i}(q_1 j + q_2 j^2)/p} |j\rangle\,, \tag{C.52}$$

the set

$$D = \{|\ell, q_1, q_2\rangle \mid \ell \in \{0, \ldots, m\}\,, q_1, q_2 \in \mathbb{Z}_p\}\,, \tag{C.53}$$

and the constant map $w(|\ell, q_1, q_2\rangle) = (n+1)^{-1} p^{-2}$. Then the pair (D, w) is a complex-projective 2-design.

One can also construct a complex-projective $\geq$ 3-design for all n by combining the simplex designs given in [189, 190] with the torus design given in Theorem C.3.8.

We note that if one relaxes the requirement that the weights be nonnegative, then one can construct signed complex-projective designs by using signed simplex and torus designs. For example, simple and explicit simplex signed t-designs are given for all odd t in [305,

thm. 4]. We leave this for future work.

Subsection C.3.5: Simplex designs from complex-projective designs

In this subsection, we will discuss the opposite direction to Theorem C.3.12; namely, that complex-projective t-designs give rise to simplex t-designs via the projection $\pi\colon \mathbb{CP}^m \to \Delta^m$ defined by $|\psi\rangle \mapsto (|\langle 0|\psi\rangle|^2, \dots, |\langle m|\psi\rangle|^2)$. Such a construction was also pointed out in [189, 306]. This will be the first step in our proof of the nonexistence of continuous-variable ($t \geq 2$)-designs. We will show that a continuous-variable design gives rise to an infinite dimensional analogue of a simplex design via a lemma analogous to Lemma C.3.13, and then show that such infinite dimensional simplex designs do not exist for $t \geq 2$. Hence, it is useful to discuss the finite dimensional case first.

Lemma C.3.13. *Let D be a t-design on $\mathbb{CP}^m$. Then $\pi(D)$ is a t-design on Δ^m.*

Proof. Since D is a design on $\mathbb{CP}^m$, which satisfies by definition

$$\mathop{\mathbb{E}}_{|\xi\rangle\in D}\left[(|\xi\rangle\langle\xi|)^{\otimes t}\right] = \int_{\mathbb{CP}^m} (|\psi\rangle\langle\psi|)^{\otimes t}\,\mathrm{d}\psi \tag{C.54a}$$

$$= \int_{\Delta^m\times T^m} (|p,\phi\rangle\langle p,\phi|)^{\otimes t}\,\mathrm{d}p\,\mathrm{d}\phi, \tag{C.54b}$$

where the last line comes from Eq. (C.43). Sandwiching this equation by $\langle a_1|\dots\langle a_t|$ and $|a_1\rangle\dots|a_t\rangle$, we find

$$\mathop{\mathbb{E}}_{|\xi\rangle\in D}\left[|\langle a_1|\xi\rangle|^2\dots|\langle a_t|\xi\rangle|^2\right] = \int_{\Delta^m\times T^m} p_{a_1}\dots p_{a_t}\,\mathrm{d}p\,\mathrm{d}\phi, \tag{C.55}$$

and hence

$$\mathop{\mathbb{E}}_{q\in\pi(D)}[q_{a_1}\dots q_{a_t}] = \int_{\Delta^n} p_{a_1}\dots p_{a_t}\,\mathrm{d}p. \tag{C.56}$$

Therefore, the ensemble $\pi(D)$ matches the integral over Δ^n for degree t monomials, and thus by linearity matches for all polynomials of degree t or less. □

In terms of weight functions, we can write this as follows. Let π^{-1} denote the preimage of π. If (D, w) is a $\mathbb{CP}^m$ t-design, then $(\pi(D), u)$ is a Δ^m t-design, where

$$\begin{aligned} u\colon \pi(D) &\to \mathbb{R}_{>0} \\ p &\mapsto \sum_{\psi \in \pi^{-1}(p)} w(\psi). \end{aligned} \tag{C.57}$$

Lemma C.3.13 tells us that

$$\begin{aligned} \int_{\Delta^n} p_{a_1} \dots p_{a_t} \, \mathrm{d}p &= \frac{1}{\operatorname{Tr} \Pi_t^{(d)}} \langle a_1| \dots \langle a_t| \, \Pi_t^{(d)} \, |a_1\rangle \dots |a_t\rangle \\ &= \frac{1}{\operatorname{Tr} \Pi_t^{(d)}} \Lambda_t^{(d)}(a_1, \dots, a_t), \end{aligned} \tag{C.58}$$

where recall that $\Lambda_t^{(d)}$ and $\Pi_t^{(d)}$ are defined in Definition C.2.6. One can then define an infinite dimensional simplex design analogously to how we define continuous-variable designs in Definition C.4.2. In particular, to get something well-defined in the infinite limit, we remove the $\operatorname{Tr} \Pi_t$ normalization, and we replace the $\mathbb{E}$ by an integral over an arbitrary measure space.

Definition C.3.14 (Infinite dimensional simplex t-design)**.** *Let (X, Σ, μ) be a measure space, and fix an integer $t \in \mathbb{N}$. Let $p = (p_i)_{i \in \mathbb{N}_0}$ be a sequence of measurable maps $p_i\colon X \to [0, 1]$. If*

$$\sum_{i \in \mathbb{N}_0} p_i(x) = 1 \quad \mu\text{-a.e. in } X, \tag{C.59}$$

and

$$\forall a \in \mathbb{N}_0^t\colon \quad \int_X \prod_{j=1}^t p_{a_j}(x) \, \mathrm{d}\mu(x) = \Lambda_t(a), \tag{C.60}$$

then $[(X, \Sigma, \mu), p]$ is an infinite dimensional simplex t-design.

In this definition, Λ_t is defined in terms of $\Pi_t\colon \mathcal{H}^{\otimes t} \to \mathcal{H}^{\otimes t}$ given in Definition C.2.6, and $\mathcal{H}$ is a infinite dimensional separable Hilbert space, e.g. $L^2(\mathbb{R})$. In the next subsection, we will show in Lemma C.4.6 that without loss of generality, the measure space for an infinite

dimensional simplex-design design can be taken to be σ-finite. Then in Lemma C.4.7, we will show that no infinite simplex t-designs exist for any $t \geq 2$.

Example C.3.15 (Infinite dimensional simplex 1-design)**.** When $t = 1$, we have that $\Lambda_1(a) = 1$ for any $a \in \mathbb{N}_0$. We have many infinite dimensional simplex 1-designs. For example, let $X = \mathbb{N}_0$, $\Sigma = \mathcal{P}(X)$, and μ be the standard counting measure $\mu(A) = |A|$. Finally, for $x \in X$, let $p_a(x) = \delta_{ax}$. Then

$$\int_X p_a(x)\,\mathrm{d}\mu(x) = \sum_{x \in \mathbb{N}_0} \delta_{ax} = 1 = \Lambda_1(a), \tag{C.61}$$

as desired. ◇

Subsection C.3.6: Constrained complex-projective integration

We now briefly describe one consequence of the formalism developed so far. This subsection is essentially unrelated to the rest of the paper, but interesting nonetheless. We will sketch the consequence with an example using the number operator, though we note that it can be generalized.

Define the number operator $\hat{n}$ by $\hat{n}\,|n\rangle = n\,|n\rangle$. Consider the constraint on $|\psi\rangle \in \mathbb{CP}^{d-1}$ that $\langle\psi|\,\hat{n}\,|\psi\rangle = \mathcal{N}$ for some constant $\mathcal{N}$. Since the constraint is diagonal in the chosen basis, it only acts on the simplex part of $\mathbb{CP}^{d-1}$. In particular, while integration over $\mathbb{CP}^{d-1}$ involves integration over the simplex Δ^{d-1}, integration over $\mathbb{CP}^{d-1}$ with the constraint that $\langle\psi|\,\hat{n}\,|\psi\rangle = \mathcal{N}$ involves integration over the simplex $\tilde{\Delta}^{d-2}$, where

$$\tilde{\Delta}^{d-2} := \left\{ p \in \Delta^{d-1} \mid \sum_{n=0}^{d-1} n p_n = \mathcal{N} \right\}. \tag{C.62}$$

Recall from the Krein–Milman theorem that any compact convex subset of Euclidean space is the convex hull of its extremal points. The simplex Δ^{d-1} is the convex hull of its d extremal points $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0)$, $\ldots$, $(0, 0, \ldots, 1)$. The simplex $\tilde{\Delta}^{d-2}$ is also the convex hull

of its $d-1$ extremal points, but its extremal points are more complicated and depend on $\mathcal{N}$. In particular, we let $b^{(i)}$ denote the i^{th} extremal point of $\tilde{\Delta}^{d-2}$, so that $b^{(i)} = (b_0^{(i)}, \ldots, b_{d-1}^{(i)})$. Then it is easy to check that the extremal points are

$$b_j^{(i)} = \begin{cases} \left(1 - \frac{\mathcal{N}}{i+1}\right) \delta_{j0} + \frac{\mathcal{N}}{i+1} \delta_{j,i+1} & \text{if } i+1 \geq \mathcal{N} \\ \left(1 - \frac{\mathcal{N}-i-1}{d-i-2}\right) \delta_{j,i+1} + \frac{\mathcal{N}-i-1}{d-i-2} \delta_{j,d-1} & \text{if } i+1 \leq \mathcal{N}, \end{cases} \tag{C.63}$$

where $i \in \{0, \ldots, d-2\}$ and $j \in \{0, \ldots, d-1\}$.

It then follows, analogously to Theorem C.3.12, that a t-design on the constrained $\mathbb{CP}^{d-1}$ space can be constructed from t-designs on $\tilde{\Delta}^{d-2}$ and T^d. Furthermore, the simplex $\tilde{\Delta}^{d-2}$ can be parameterized via baryocentric coordinates in terms of the standard simplex Δ^{d-2}. In particular, a point in Δ^{d-2} defines a particular convex combination of the extremal points of $\tilde{\Delta}^{d-2}$, which gives a point in $\tilde{\Delta}^{d-2}$. Therefore, one can compute the integral $\int_{\mathbb{CP}^{d-1}} (|\psi\rangle\langle\psi|)^{\otimes t} \, \delta(\mathcal{N} - \langle\psi| \hat{n} |\psi\rangle) \, \mathrm{d}\psi$ up to proportionality by using simplex and torus designs. We note that such a construction does not work if the $\delta(\mathcal{N} - \langle\psi| \hat{n} |\psi\rangle)$ constraint is replaced with $\Theta(\mathcal{N} - \langle\psi| \hat{n} |\psi\rangle)$, where Θ is the Heaviside step function. This is for a slightly subtle reason. The δ constraint results in a measure on $\tilde{\Delta}^{d-2}$ that is, up to proportionality, the standard Lebesgue measure. On the other hand, the Θ constraint results in a more complicated measure, and indeed this measure mixes the contributions of the torus and the simplex in the integral. As such, the resulting integral is no longer over a simple product of a simplex and torus, but rather over a more complicated combination of the two.

The δ constraint that fixed $\langle\psi| \hat{n} |\psi\rangle = \mathcal{N}$ is interesting nonetheless. By using any of the simplex 1- and 2-designs from Appendix C.3.1 and any of the torus 1- and 2-designs from Appendix C.3.2, we can compute the following integrals, up to proportionality, in terms of

the extremal points $b^{(i)}$:

$$\int_{\mathbb{CP}^{d-1}} |\psi\rangle\langle\psi|\, \delta(\mathcal{N} - \langle\psi|\, \hat{n}\, |\psi\rangle)\, \mathrm{d}\psi \propto \frac{1}{d-1} \sum_{k=0}^{d-1} |k\rangle\langle k| \sum_{j=0}^{d-2} b_k^{(j)}, \tag{C.64}$$

$$\int_{\mathbb{CP}^{d-1}} (|\psi\rangle\langle\psi|)^{\otimes 2} \delta(\mathcal{N} - \langle\psi|\, \hat{n}\, |\psi\rangle)\, \mathrm{d}\psi \propto \frac{1}{d(d-1)} \sum_{i,j=1}^{d-2} (1+\delta_{ij}) \Bigg[\sum_{k_1,k_2=1}^{d-1} b_{k_1}^{(i)} b_{k_2}^{(j)} \left(|k_1\rangle\, |k_2\rangle\langle k_1|\, \langle k_2| + |k_1\rangle\, |k_2\rangle\langle k_2|\, \langle k_1|\right) + \sum_{k=1}^{d-1} b_k^{(i)} b_k^{(j)}\, |k\rangle\, |k\rangle\langle k|\, \langle k| \Bigg]. \tag{C.65}$$

If we, for example, fix $\mathcal{N} = 1$, then the result is

$$\int_{\mathbb{CP}^{d-1}} |\psi\rangle\langle\psi|\, \delta(1 - \langle\psi|\, \hat{n}\, |\psi\rangle)\, \mathrm{d}\psi \propto \left(1 - \frac{H_{d-1}}{d-1}\right) |0\rangle\langle 0| + \frac{1}{d-1} \sum_{k=1}^{d-1} \frac{1}{k}\, |k\rangle\langle k|\,, \tag{C.66}$$

where $H_{d-1} = \sum_{k=1}^{d-1} 1/k$ is the $(d-1)^{\text{th}}$ harmonic number.

Section C.4: Continuous-variable designs

In extending the definition of complex-projective designs to the infinite dimensional case of continuous-variables, one encounters the issue that $\operatorname{Tr} \Pi_t$ is not finite. Hence, in accordance with the definition of continuous-variable designs given in [79], we remove the trace in the denominator of $\Pi_t / \operatorname{Tr} \Pi_t$ and replace the equality with a proportionality. By a simple rescaling of the ensemble, the proportionality constant can be made arbitrary. Thus, we can in fact keep the equality. We are therefore tempted to define a continuous-variable t-design as an ensemble $\mathcal{E}$ satisfying

$$\text{“ } \mathop{\mathbb{E}}_{\mathcal{E}} \left[(|\psi\rangle\langle\psi|)^{\otimes t} \right] = \Pi_t \text{.”} \tag{C.67}$$

However, since $\operatorname{Tr} \Pi_t$ is infinite, it follows that the ensemble $\mathcal{E}$ must not be compact, making $\mathbb{E}_{\mathcal{E}}$ ill-defined. We therefore replace the expectation value with an integral over an arbitrary

measure space (X, Σ, μ). Here X is a set, Σ a σ-algebra on X, and $\mu\colon \Sigma \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ is a measure on X. Finally, we arrive at the precise definition of a continuous-variable t-design on $L^2(\mathbb{R})$.

Definition C.4.1 (Continuous-variable state t-design)**.** *Let $X \subset L^2(\mathbb{R})$, (X, Σ, μ) be a measure space, and fix a positive integer $t \in \mathbb{N}$. Let $\Pi_t\colon L^2(\mathbb{R})^{\otimes t} \to L^2(\mathbb{R})^{\otimes t}$ be as in Definition C.2.6. If*

$$\int_X (|\psi\rangle\langle\psi|)^{\otimes t} \, \mathrm{d}\mu(\psi) = \Pi_t, \tag{C.68}$$

where we use the weak (Pettis) integral, then (X, Σ, μ) is a continuous-variable state t-design*. Hence, we require a design to satisfy*

$$\int_X \left(\prod_{i=1}^{t} \langle a_i|\psi\rangle\langle\psi|b_i\rangle \right) \mathrm{d}\mu(\psi) = \Pi_t(a_1, \ldots, a_t; b_1, \ldots, b_t) \tag{C.69}$$

for all tuples $a, b \in \mathbb{N}_0^t$.

The motivation for this definition of continuous-variable state designs is summarized in Fig. 4.1. An alternative characterization of continuous-variable state designs is given in Appendix C.4.5. If one is familiar with weighted complex-projective designs, as defined in e.g. Ref. [43], then one can imagine that the measure μ is a Lebesgue-Stieltjes measure coming from a weight function. For the purposes of designs, the weak (Pettis) integral is more natural than the strong (Bochner) integral because we are generally interested in averaged functions of ψ. Ultimately, we will prove that continuous-variable t-designs do not exist for $t \geq 2$, which immediately implies the result for the case of the strong integral as well.

By parameterizing states in $L^2(\mathbb{R})$ with polar coordinates, one can arrive at an equivalent definition of continuous-variable t-designs.

Definition C.4.2 (Continuous-variable state t-design)**.** *Let X be an arbitrary set, (X, Σ, μ) be a measure space, and fix an integer $t \in \mathbb{N}$. Let $p = (p_i)_{i \in \mathbb{N}_0}$ and $\phi = (\phi_i)_{i \in \mathbb{N}_0}$ be sequences*

of measurable maps $p_i \colon X \to [0,1]$ *and* $\phi_i \colon X \to \mathbb{R}$ *satisfying* $\sum_{i \in \mathbb{N}_0} p_i(x) = 1$ *for almost all* $x \in X$. *Define the state* $|p(x), \phi(x)\rangle \in L^2(\mathbb{R})$ *by*

$$|p(x), \phi(x)\rangle \coloneqq \sum_{n \in \mathbb{N}_0} \sqrt{p_n(x)} \mathrm{e}^{\mathrm{i}\phi_n(x)} |n\rangle \,. \tag{C.70}$$

Let $\Pi_t \colon L^2(\mathbb{R})^{\otimes t} \to L^2(\mathbb{R})^{\otimes t}$ *be as in Definition C.2.6. If*

$$\int_X \left(|p(x), \phi(x)\rangle\langle p(x), \phi(x)|\right)^{\otimes t} \mathrm{d}\mu(x) = \Pi_t, \tag{C.71}$$

where we use the weak (Pettis) integral, then $[(X, \Sigma, \mu), p, \phi]$ *is a* continuous-variable state t-design. *Hence, we require a design to satisfy*

$$\int_X \left(\prod_{i=1}^{t} \langle a_i | p(x), \phi(x)\rangle\langle p(x), \phi(x) | b_i \rangle \right) \mathrm{d}\mu(x) = \Pi_t(a_1, \ldots, a_t; b_1, \ldots, b_t) \tag{C.72}$$

for all tuples $a, b \in \mathbb{N}_0^t$.

Definition C.4.2 will be a more operationally useful definition for our purposes, but we emphasize that Definitions C.4.1 and C.4.2 are equivalent definitions, where the latter is simply a different parameterization of the former.

Proposition C.4.3. *Definitions C.4.1 and C.4.2 are equivalent definitions.*

Proof. For any $|\psi\rangle$ coming from the first definition, we get the sequences p and ϕ for the second definition as $p_n = |\langle n|\psi\rangle|^2$ and $\phi_n = \arg\langle n|\psi\rangle$. One can then normalize each p_i by $p_i(x) \to p_i(x) / \sum_i p_i(x)$, and then absorb a factor of $\left(\sum_i p_i(x)\right)^t$ into the measure.

Conversely for any measure space and sequences p and ϕ coming from the second definition, we get the measure space $(X \subset L^2(\mathbb{R}), \Sigma, \mu)$ for the first definition since the parameterization defines states in $L^2(\mathbb{R})$. □

We include Definition C.4.1 since it closer matches the standard definition of a weighted

complex-projective design. In light of Proposition C.4.3, henceforth we will use Definition C.4.2. To become better acquainted with this definition, consider the following example of a continuous-variable state 1-design.

Example C.4.4 (Continuous-variable state 1-designs)**.** Consider the measure space where $X = \mathbb{N}_0$, Σ is the power-set $\mathcal{P}(X)$, and μ is the standard counting measure on $\mathbb{N}_0$. Let $p_n \colon x \mapsto \delta_{nx}$, and $\phi_n \colon x \mapsto 0$. Then

$$\int_X |p(x), \phi(x)\rangle\langle p(x), \phi(x)| \, \mathrm{d}\mu(x) = \sum_{x \in \mathbb{N}_0} |n = x\rangle\langle n = x| = \Pi_1, \tag{C.73}$$

where note that $\Pi_1 = \mathbb{I}$. Hence, this is an example of a continuous-variable state 1-design.

Similarly, consider $\mathbb{R}_{>0}$ with the standard Borel σ-algebra and Lebesgue measure. Consider also $[0, 2\pi)$ with the normalized Lebesgue measure. Let (X, Σ, μ) be the unique product measure space for $X = \mathbb{R}_{>0} \times [0, 2\pi)$. For an element $x \in X$, notate $x = (r, \theta)$, for $r \in \mathbb{R}_{>0}$ and $\theta \in [0, 2\pi)$. Let $p_n \colon (r, \theta) \mapsto e^{-r} r^n / n!$ and $\phi_n \colon (r, \theta) \mapsto \theta n$. Then

$$\begin{aligned}
&\int_X |p(x), \phi(x)\rangle\langle p(x), \phi(x)| \, \mathrm{d}\mu(x) \\
&\quad = \frac{1}{2\pi} \sum_{n,m \in \mathbb{N}_0} |n\rangle\langle m| \int_0^\infty \mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \; \mathrm{e}^{\mathrm{i}\theta(n-m)} \mathrm{e}^{-r} \frac{r^{n/2+m/2}}{\sqrt{n!m!}} && \text{(C.74a)} \\
&\quad = \sum_{n \in \mathbb{N}_0} |n\rangle\langle n| \int_0^\infty \mathrm{d}r \; \mathrm{e}^{-r} \frac{r^n}{n!} && \text{(C.74b)} \\
&\quad = \sum_{n \in \mathbb{N}_0} |n\rangle\langle n| = \Pi_1, && \text{(C.74c)}
\end{aligned}$$

giving another example of a 1-design. This 1-design is more commonly written as

$$\int |\alpha\rangle\langle\alpha| \, \frac{\mathrm{d}^2\alpha}{\pi} = \mathbb{I}, \tag{C.75}$$

where $\mathrm{d}^2\alpha = \mathrm{dRe}\,\alpha \, \mathrm{dIm}\,\alpha$ and $|\alpha\rangle$ is a coherent state. Namely, coherent states form an overcomplete frame. ◇

We emphasize that this definition completely sidesteps the issue that one cannot define a finite Haar measure on $\mathrm{U}(L^2(\mathbb{R}))$ since it is not a compact group, and indeed not even a locally compact group [242, sec. 5]. See Fig. 4.1 for a visualization. The issue is sidestepped by never considering the integral over all states in $L^2(\mathbb{R})$. Instead, we note that the integral over all states for a finite dimensional Hilbert space gives a finite dimensional Π_t, and we extend the definition of a design to the infinite dimensional case by extending Π_t to the infinite dimensional space $L^2(\mathbb{R})$. This is exactly the approach that was taken in [79, 153]. Alternatively, in Appendix C.4.5, we *do* explicitly consider integration over the infinite dimensional space $\mathbb{C}^\infty \supset L^2(\mathbb{R})$.

Subsection C.4.1: Non-existence of continuous-variable state designs

It has been shown that no set of Gaussian states can form a continuous-variable 2-design [79]. We extend this result to show that there do not exist continuous-variable t-designs for any $t > 1$. We emphasize that our proof in fact works for any separable Hilbert space $\mathcal{H}$, not just $L^2(\mathbb{R})$, since it only assumes the existence of a countable orthonormal basis.

Theorem C.4.5. *No continuous-variable state or unitary t-designs exist for any integer $t \geq 2$.*

Theorem C.4.5 is an immediate consequence of Lemma C.4.6 and Lemma C.4.7 below. Fig. C.1 provides an overview of the proof. We recommend reading this section first keeping in mind the specific example of $t = 2$, where the explicit form of Π_2 and Λ_2 are given in Example C.2.7. After understanding this case, the extension to arbitrary $t \geq 2$ is straightforward.

To begin, we will show that existence of continuous-variable state t-designs implies existence of simplex t-designs.

Lemma C.4.6. *If a continuous-variable t-design exists, then there exists a σ-finite measure*

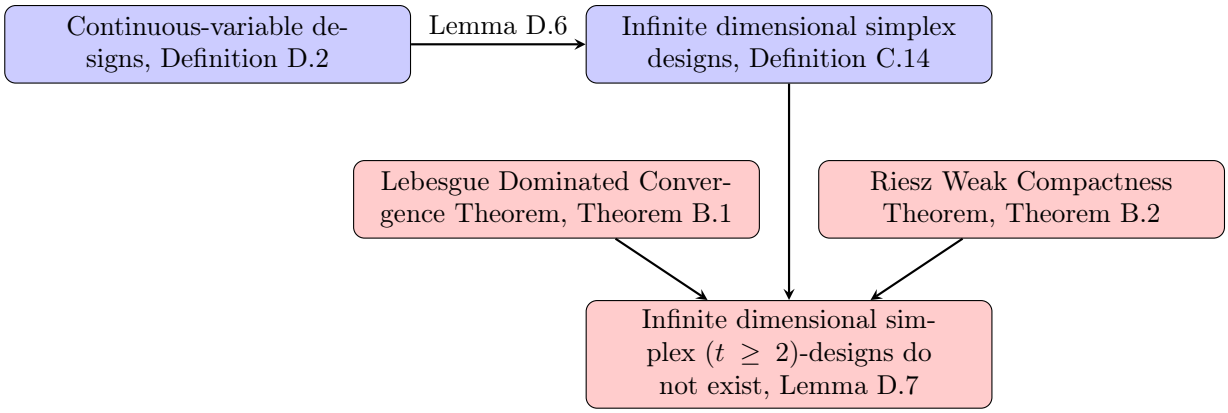

Figure C.1: An outline of the proof of the non-existence of continuous-variable t-designs for $t \geq 2$.

space (X, Σ, μ) *and a sequence* $p = (p_i)_{i=0}^{\infty}$ *of measurable maps* $p_i \colon X \to [0,1]$ *satisfying*

$$\sum_{i=0}^{\infty} p_i(x) = 1 \quad \mu\text{-a.e. in } X, \tag{C.76}$$

and

$$\forall a \in \mathbb{N}_0^t \colon \int_X \prod_{i=1}^{t} p_{a_i}(x) \, \mathrm{d}\mu(x) = \Lambda_t(a). \tag{C.77}$$

Proof. Suppose a continuous-variable state t-design exists. Then Eq. (C.72) holds for all tuples $a, b \in \mathbb{N}_0^t$, and p satisfies Eq. (C.76) by Definition C.4.2. Indeed, Eq. (C.76) is simply the requirement that the quantum states be normalized. Plugging in $a = b$ and $\Lambda_t(a) = \Pi_t(a; a)$ by definition, we get

$$\Lambda_t(a) = \int_X \left(\prod_{i=1}^{t} \langle a_i | p(x), \phi(x) \rangle \langle p(x), \phi(x) | a_i \rangle \right) \mathrm{d}\mu(x) \tag{C.78a}$$

$$= \int_X \prod_{i=1}^{t} |\langle a_i | p(x), \phi(x) \rangle|^2 \, \mathrm{d}\mu(x) \tag{C.78b}$$

$$= \int_X \prod_{i=1}^{t} p_{a_i}(x) \, \mathrm{d}\mu(x). \tag{C.78c}$$

Therefore, the measure space and sequence p satisfy Eq. (C.76) and Eq. (C.77). The only

remaining thing to show is that X can be σ-finite.

Consider the function $f = p_i^t$ whose codomain is clearly $[0,1]$. By Eq. (C.77), $0 < \int_X f \, \mathrm{d}\mu < \infty$. Hence, by Lemma C.2.4, the preimage $f^{-1}((0,1]) = p_i^{-1}((0,1])$ is a σ-finite set. Since a countable union of σ-finite sets is σ-finite, it must be that $Y \coloneqq \bigcup_{i=0}^{\infty} p_i^{-1}((0,1])$ is σ-finite (also recall that any σ-finite set is measurable). The set $X \setminus Y$ is equal to $\bigcap_{i=0}^{\infty} p_i^{-1}(\{0\})$. Eq. (C.76) is required to hold almost everywhere in X. This means that the set of points for which it does not hold is contained within a measure zero subset of X. Clearly, Eq. (C.76) does not hold when $x \in X \setminus Y$. Therefore, $X \setminus Y$ is contained within a measure zero subset of X. Thus, if Eq. (C.77) holds on X, then it also holds on Y, and of course the same is true for Eq. (C.76).

Hence we have determined if Eqs. (C.76) and (C.77) are satisfied by the measure space (X, Σ, μ), then they are also satisfied by the measure space $(Y, \Sigma|_Y, \mu|_Y)$, where $|_Y$ denotes the restriction to the subset $Y \subseteq X$. To see that $(Y, \Sigma|_Y, \mu|_Y)$ is a valid measure space, recall that we have already shown that $Y \in \Sigma$. Then one can straightforwardly check that $\Sigma|_Y$ is a σ-algebra of Y and $\mu|_Y$ is a valid measure with respect to $\Sigma|_Y$, so that the restriction of (X, Σ, μ) to Y is a measure space (see e.g. [299, Ch. 17.1, exercise 6]). We have also shown that $(Y, \Sigma|_Y, \mu|_Y)$ is σ-finite. In summary, we have shown that if a continuous-variable t-design exists, then there exists a measure space satisfying Eqs. (C.76) and (C.77). We then showed that the existence of this measure space implies the existence of a σ-finite measure space satisfying Eqs. (C.76) and (C.77), hence completing the proof. □

As we commented in Lemma C.3.13, a complex-projective design gives rise to a simplex design. At a high level, Lemma C.4.6 is extending this fact to the continuous-variable regime. Similar to how we extended the definition of a complex-projective design to infinite dimensions, the analogous extension of a simplex design to infinite dimensions is the conditions in Eqs. (C.76) and (C.77), as in Definition C.3.14. The extra bit about X being σ-finite is just a technical point needed so that Theorem C.2.2 can be used in the next lemma.

Given Lemma C.4.6, we immediately see that if no σ-finite measure space and sequence p can satisfy Eqs. (C.76) and (C.77), then no continuous-variable t-designs can exist. This is what we show in the following lemma.

Lemma C.4.7. *No (X, Σ, μ) and $(p_i)_{i=0}^{\infty}$ exist satisfying the conditions of Lemma C.4.6 for any $t \in \mathbb{N}_{\geq 2}$.*

Proof. Assume by way of contradiction that such (X, Σ, μ) and (p_i) exist. Because of Eq. (C.76), it must be that for almost all $x \in X$, $\lim_{i\to\infty} p_i(x) = 0$. Since the sequence (p_i) converges, it must be the case that every subsequence $(p_{i_k})_{k=0}^{\infty}$ of (p_i) also converges to the same point; $\lim_{k\to\infty} p_{i_k}(x) = 0$ for almost all x. For any tuple $j \in \mathbb{N}_0^t$, define $g(x) = \prod_{l=1}^{t} p_{j_l}(x)$, which is in $L^1(X) = L^1(X, \Sigma, \mu)$ (i.e. $\int_X g \, \mathrm{d}\mu < \infty$) by Eq. (C.77). Consider the sequence $(f_{i_k})_{k=0}^{\infty}$ where $f_{i_k}(x) = p_{i_k}(x)g(x)$ for any $j \in \mathbb{N}_0^t$. Then f_{i_k} converges pointwise to the zero function $f(x) = 0$ almost everywhere as $k \to \infty$, and f is obviously measurable. Since the codomain of p_{i_k} is $[0, 1]$, it follows that $f_{i_k} \leq g$ for all i_k. Therefore, (f_{i_k}) is a sequence in $L^1(X)$ and is dominated by a nonnegative integrable g. Hence, we can apply the Dominated Convergence Theorem C.2.1 to swap the limit and the integral and find that $\lim_{k\to\infty} \|f_{i_k} - f\|_1 = 0$, giving

$$\forall j \in \mathbb{N}_0^t \colon \quad \lim_{k\to\infty} \int_X p_{i_k}(x) \prod_{l=1}^{t} p_{j_l}(x) \, \mathrm{d}\mu(x) = 0. \tag{C.79}$$

Next we consider the sequence $(p_i)_{i=0}^{\infty}$, which is a bounded sequence in $L^t(X)$ since $\int_X p_i^t \, \mathrm{d}\mu < \infty$ by Eq. (C.77). Therefore, we can apply the Riesz Weak Compactness Theorem from Theorem C.2.2 to find a subseqence $(p_{i_k})_{k=0}^{\infty}$ and a function q for which for all $h \in L^{t'}(X)$,

$$\lim_{k\to\infty} \int_X p_{i_k}(x)h(x) \, \mathrm{d}\mu(x) = \int_X q(x)h(x) \, \mathrm{d}\mu(x), \tag{C.80}$$

where $t' = t/(t-1)$. Now we must prove that q is zero almost everywhere.

First we show that q must be nonnegative almost everywhere. Heuristically, this is

because q is being substituted for a limit of probabilities, which themselves are always nonnegative. More technically, let $\mathbb{1}_A$ be the indicator function, so that $\mathbb{1}_A(x)$ is 1 if $x \in A$, and zero otherwise. Since X is σ-finite, there exists a sequence (A_j) where A_j is measurable $A_j \in \Sigma$, $\mu(A_j) < \infty$, and $X = \bigcup_{j=0}^{\infty} A_j$. Since A_j has finite measure, $\mathbb{1}_{A_j} \in L^{t'}(X)$. Plugging $h = \mathbb{1}_{A_j}$ into Eq. (C.80), we find that $\int_X q(x)\mathbb{1}_{A_j}(x)\,\mathrm{d}\mu(x) \geq 0$ for all j. Therefore, $\int_A q\,\mathrm{d}\mu \geq 0$ for every $A \in \Sigma$ of finite measure, and we can build up X from such A's. This tells us that $q \geq 0$ almost everywhere.

Next we show that q must be the zero function almost everywhere. For some $j \in \mathbb{N}_0^t$, we plug $h(x) = \prod_{l=1}^{t} p_{j_l}(x) \in L^{t'}(X)$ into Eq. (C.80). Using Eq. (C.79) for the left hand side of Eq. (C.80), this tells us that

$$\forall j \in \mathbb{N}_0^t\colon \quad \int_X q(x) \prod_{l=1}^{t} p_{j_l}(x)\,\mathrm{d}\mu(x) = 0. \tag{C.81}$$

Along with the fact that q must be nonnegative almost everywhere, this implies that $q(x)$ must be zero almost everywhere whenever $p_j(x) \neq 0$ for any j. As such, q must be zero almost everywhere on the set $\bigcup_{j=0}^{\infty} p_j^{-1}((0,1])$. But we showed in the proof of Lemma C.4.6 that $X \setminus \bigcup_{j=0}^{\infty} p_j^{-1}((0,1])$ is contained within a measure zero subset of X.

We have shown that q is the zero function almost everywhere on X. Hence, Eq. (C.80) becomes

$$\forall h \in L^{t'}(X)\colon \quad \lim_{k\to\infty} \int_X p_{i_k}(x)h(x)\,\mathrm{d}\mu(x) = 0. \tag{C.82}$$

Plugging $h = p_0^{t-1}$ (which is in $L^{t'}(X)$ by Eq. (C.77)) into Eq. (C.82), we arrive at $\lim_{k\to\infty} \int_X p_0^{t-1} p_{i_k}\,\mathrm{d}\mu = 0$. But Eq. (C.77) tells us that $\lim_{k\to\infty} \int_X p_0^{t-1} p_{i_k}\,\mathrm{d}\mu = \lim_{k\to\infty} \Lambda_t(0,\ldots,0,i_k)$, which, from the definition of Λ_t in terms of Π_t, is strictly positive as shown in Eq. (C.18). We've reached a contradiction, hence completing the proof. □

The non-existence of state designs statement of Theorem C.4.5 follows as an immediate corollary of Lemmas C.4.6 and C.4.7. The non-existence of unitary designs follows straight-

forwardly from the non-existence of state designs, as we explain in the next subsection. Furthermore, Theorem C.4.5 still holds even in the case when one allows (X, Σ, μ) to be a *signed* measure space by a simple appeal to the Hahn Decomposition Theorem [307]. Indeed, using this theorem, one simply splits the signed measure space into two nonnegative measure spaces and then proceeds with the proof of Theorem C.4.5.

Subsection C.4.2: Non-existence of continuous-variable unitary designs

Theorem C.4.5 extends the results from [79], where it is shown that the set of Gaussian states does not form a state 2-design, and the results from [153], where it is shown that the set of Gaussian unitaries does not form a unitary 2-design. The non-existence of continuous-variable state t-designs for $t > 1$ immediately implies the non-existence of continuous-variable unitary t-designs for $t > 1$, since any unitary design gives rise to a state design by *twirling* a fiducial state. To be clear, we consider the definition of a continuous-variable unitary 2-design given in [153]. Namely, a unitary 2-design is any ensemble $\mathcal{E}$ of unitaries satisfying

$$\mathop{\mathbb{E}}_{\mathcal{E}}\left[(U \otimes U)A(U \otimes U)^\dagger\right] \propto \frac{1}{2}\left(\mathbb{I}\,\mathrm{Tr}[A] + S\,\mathrm{Tr}[SA]\right) \tag{C.83}$$

for any operator trace-class operator A, where S is the SWAP operator that swaps the elements of the tensor product space. Since this should hold for any A, we can substitute $A = (|\phi\rangle\langle\phi|)^{\otimes 2}$ for any fiducial state $|\phi\rangle$ (e.g. the zero Fock state $|0\rangle$). We can then define a new ensemble over states $\mathcal{E}' = \{U|\phi\rangle \mid U \in \mathcal{E}\}$. The result is

$$\mathop{\mathbb{E}}_{\mathcal{E}'}\left[(|\psi\rangle\langle\psi|)^{\otimes 2}\right] \propto \frac{1}{2}\left(\mathbb{I} + S\right) = \Pi_2, \tag{C.84}$$

which precisely matches the definition of a continuous-variable state 2-design given in [79], and the definition we use in Definition C.4.2. Hence, by contraposition, if a state design does not exist, then a unitary design necessarily does not exist.

This result holds generally for the definition of continuous-variable unitary t-designs given in [153, footnote 89]. Specifically, a unitary t-design is an ensemble $\mathcal{E}$ satisfying

$$\mathop{\mathbb{E}}_{\mathcal{E}}\left[U^{\otimes t} A U^{\dagger\,\otimes t}\right] \propto \frac{1}{|S_t|} \sum_{\sigma \in S_t} W_\sigma \operatorname{Tr}(W_\sigma^{-1} A) \tag{C.85}$$

for any trace-class operator A. Substituting $A = (|\phi\rangle\langle\phi|)^{\otimes t}$ for some fiducial state $|\phi\rangle$. Then define a new ensemble over states $\mathcal{E}' = \{U|\phi\rangle \mid U \in \mathcal{E}\}$. The result is then

$$\mathop{\mathbb{E}}_{\mathcal{E}'}\left[(|\psi\rangle\langle\psi|)^{\otimes t}\right] \propto \frac{1}{|S_t|} \sum_{\sigma \in S_t} W_\sigma = \Pi_t, \tag{C.86}$$

meaning that $\mathcal{E}'$ is a continuous-variable state t-design according to Definition C.4.2. By the nonexistence of continuous-variable $(t \geq 2)$-designs then, such a unitary design does not exist for $t \geq 2$.

Subsection C.4.3: Rigged continuous-variable state designs

The result of Theorem C.4.5 is that no continuous-variable $(t \geq 2)$-designs exist. The main hindrance to the construction of continuous-variable state $(t \geq 2)$-designs is the requirement that the states be normalized. In particular, the proof did not rely on exactly what the states were normalized to, only that $\lim_{i\to\infty} p_i = 0$. Hence, the requirement that the states belong to $L^2(\mathbb{R})$ inhibits the existence of continuous-variable designs. This motivates the approach taken in this section, where we construct *rigged continuous-variable state designs* by relaxing the normalization condition, thus allowing unphysical states such as the infinite superposition state $\sum_{n\in\mathbb{N}_0} |n\rangle$. Specifically, the we will use elements of the standard rigged Hilbert space on top of $L^2(\mathbb{R})$ to reconstruct Π_t. These elements are called *tempered distributions* (see below); some familiar tempered distributions are the position eigenstates $|x\rangle$ and the momentum eigenstates $|p\rangle$. In the next section, we will reintroduce normalization via a soft energy cutoff, hence making the rigged continuous-variable designs a type of approximate

continuous-variable design.

When we remove the normalization condition on the states, many of the finite dimensional complex-projective designs still do not naturally extend to the continuous-variable regime. For example, consider the complex-projective design given in Eq. (C.53). We begin by making the states $|m, q_1, q_2\rangle$ non-normalizable by multiplying through by $\sqrt{n+1}$, and then take the $n \to \infty$ limit. In this limit, $r = 1/\sqrt{n+2} \to 0$, and hence $\sqrt{n+1}\,|m, q_1, q_2\rangle = \sqrt{n+1}\,|m', q_1, q_2\rangle$ for all m and m'. One can straightforwardly check that these states do not reconstruct Π_2. In particular, these states only form a 1-design as the underlying simplex design is only a 1-design.

However, some finite dimensional complex-projective designs do extend to the continuous-variable regime when normalization is removed. For example, consider the following $\mathbb{CP}^{d-1}$ 2-design, when d is prime, given in Eq. (C.51), which also happens to be a maximal set of mutually unbiased bases [39, 41]. Define the state $|q_1\rangle_{q_2} \equiv {}_{q_2}|q_1\rangle \coloneqq \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} \exp\left[\frac{2\pi \mathrm{i}}{d}\left(q_1 n + q_2 n^2\right)\right] |n\rangle$. One can straightforwardly show that for each q_2, $\left\{ |q_1\rangle_{q_2} \mid q_1 \in \{0, \ldots, d-1\} \right\}$ is an orthonormal basis, and that

$$\Pi_2 = \frac{1}{2} \sum_{n=0}^{d-1} \left(|n\rangle\langle n|\right)^{\otimes 2} + \frac{1}{2} \sum_{q_1, q_2 = 0}^{d-1} \left({}_{q_2}|q_1\rangle\langle q_1|_{q_2}\right)^{\otimes 2}. \tag{C.87}$$

The phases involved in this design utilize the so-called "$ax^2 + bx$ construction" described in [304]. In the language of our paper, the "$ax^2 + bx$ construction" is alluding to a particular torus 2-design construction, namely given in Theorem C.3.9.

This design cannot be extended to a continuous-variable design because the states $|q_1\rangle_{q_2}$ are unphysical when d is infinite. If we relax the normalization condition, however, we can reconstruct the infinite dimensional symmetric projector with an analogous design.

Theorem C.4.8. *Define the non-normalizable state*

$$|\theta\rangle_\varphi \equiv {}_\varphi|\theta\rangle := \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{N}_0} \exp\left[\mathrm{i}\left(\theta n + \varphi n^2\right)\right] |n\rangle\,. \tag{C.88}$$

Then

$$\Pi_2 = \frac{1}{2}\sum_{n\in\mathbb{N}_0} \left(|n\rangle\langle n|\right)^{\otimes 2} + \frac{1}{2}\int_{-\pi}^{\pi}\mathrm{d}\varphi \int_{-\pi}^{\pi}\mathrm{d}\theta \left({}_\varphi|\theta\rangle\langle\theta|_\varphi\right)^{\otimes 2}. \tag{C.89}$$

Proof.

$$\int_{-\pi}^{\pi}\mathrm{d}\varphi \int_{-\pi}^{\pi}\mathrm{d}\theta \langle a|\theta\rangle_\varphi \langle b|\theta\rangle_\varphi\, {}_\varphi\langle\theta|c\rangle\, {}_\varphi\langle\theta|d\rangle \tag{C.90a}$$

$$= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\mathrm{d}\varphi \int_{-\pi}^{\pi}\mathrm{d}\theta \mathrm{e}^{\mathrm{i}\theta(a+b-c-d)}\mathrm{e}^{\mathrm{i}\varphi(a^2+b^2-c^2-d^2)} \tag{C.90b}$$

$$= \begin{cases} 1 & \text{if } a+b=c+d \text{ and } a^2+b^2=c^2+d^2 \\ 0 & \text{otherwise} \end{cases} \tag{C.90c}$$

$$= \begin{cases} 1 & \text{if } a=c\neq b=d \text{ or } a=d\neq b=c \text{ or } a=b=c=d \\ 0 & \text{otherwise} \end{cases} \quad \text{from Lemma C.3.10} \tag{C.90d}$$

$$= \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc} - \delta_{ab}\delta_{ac}\delta_{ad} \tag{C.90e}$$

$$= 2\Pi_2(a,b;c,d) - \delta_{ab}\delta_{ac}\delta_{ad} \tag{C.90f}$$

$$= 2\Pi_2(a,b;c,d) - \sum_{n\in\mathbb{N}_0}\langle a|n\rangle\langle b|n\rangle\langle n|c\rangle\langle n|d\rangle. \tag{C.90g}$$

□

Strictly speaking, Eq. (C.89) should say $\Pi_2\big|_{S(\mathbb{R})}$, since the $\langle\theta|_\varphi$ are only defined on Schwartz space $S(\mathbb{R}) \subset L^2(\mathbb{R})$. However, its action can be uniquely extended to all of $L^2(\mathbb{R})$. This will be formalized later on in this section.

We can find analogous results using cos and sin states from [199]. We note that looking

at the $|\theta\rangle_\varphi$ non-normalizable states was motivated by lifting the finite dimensional complex-projective design to the continuous-variable case. These states are Kerred phase states, which form a continuous projection-valued measure (PVM) [199]. From these states, we are then motivated to define the $|\cos\theta\rangle_\varphi$ and $|\sin\theta\rangle_\varphi$ states, which are defined in [199], since they are similar to the Kerred phase states, but nicer in many ways. In particular, for each φ, they form an generalized orthogonal basis.

Theorem C.4.9. *Define the non-normalizable state*

$$|\cos\theta\rangle_\varphi \equiv {}_\varphi|\cos\theta\rangle \coloneqq \sqrt{\frac{2}{\pi}} \sum_{n\in\mathbb{N}_0} \mathrm{e}^{\mathrm{i}\varphi n^2} \sin((n+1)\theta)\,|n\rangle\,. \tag{C.91}$$

Then

$$\Pi_2 = \frac{1}{4}\sum_{n\in\mathbb{N}_0} (|n\rangle\langle n|)^{\otimes 2} + \frac{1}{4}\int_{-\pi}^{\pi}\mathrm{d}\varphi \int_0^\pi \mathrm{d}\theta \left({}_\varphi|\cos\theta\rangle\langle\cos\theta|_\varphi\right)^{\otimes 2} \tag{C.92}$$

Proof.

$$\int_{-\pi}^{\pi}\mathrm{d}\varphi\int_0^\pi \mathrm{d}\theta \langle a|\cos\theta\rangle_\varphi\langle b|\cos\theta\rangle_\varphi\, {}_\varphi\langle\cos\theta|c\rangle\, {}_\varphi\langle\cos\theta|d\rangle \tag{C.93a}$$

$$= \frac{4}{\pi^2}\int_{-\pi}^{\pi}\mathrm{d}\varphi\int_0^\pi \mathrm{d}\theta\ \mathrm{e}^{\mathrm{i}\varphi(a^2+b^2-c^2-d^2)} \sin((a+1)\theta)\sin((b+1)\theta)\sin((c+1)\theta)\sin((d+1)\theta) \tag{C.93b}$$

$$= \frac{8}{\pi}\delta_{a^2+b^2,c^2+d^2}\int_0^\pi \mathrm{d}\theta \sin((a+1)\theta)\sin((b+1)\theta)\sin((c+1)\theta)\sin((d+1)\theta) \tag{C.93c}$$

$$\begin{aligned} = \frac{1}{\pi}\delta_{a^2+b^2,c^2+d^2}\int_0^\pi \mathrm{d}\theta\big[&\cos(\theta(a+b+c+d+4)) - \cos(\theta(-a+b+c+d+2)) \\ &- \cos(\theta(a-b+c+d+2)) - \cos(\theta(a+b-c+d+2)) + \cos(\theta(a-b-c+d)) \\ &- \cos(\theta(a+b+c-d+2)) + \cos(\theta(a-b+c-d)) + \cos(\theta(a+b-c-d))\big] \end{aligned} \tag{C.93d}$$

$$= \delta_{a^2+b^2,c^2+d^2}\big[\operatorname{sinc}(a+b+c+d+4) - \operatorname{sinc}(-a+b+c+d+2) - \operatorname{sinc}(a-b+c+d+2)$$
$$- \operatorname{sinc}(a+b-c+d+2) + \operatorname{sinc}(a-b-c+d) - \operatorname{sinc}(a+b+c-d+2)$$
$$+ \operatorname{sinc}(a-b+c-d) + \operatorname{sinc}(a+b-c-d)\big] \tag{C.93e}$$

$$= \delta_{a^2+b^2,c^2+d^2}\big[- \delta_{a,b+c+d+2} - \delta_{b,a+c+d+2}$$
$$- \delta_{c,a+b+d+2} - \delta_{d,a+b+c+2} + \delta_{a+d,b+c} + \delta_{a+c,b+d} + \delta_{a+b,c+d}\big], \tag{C.93f}$$

where we used that $\int_0^\pi \cos(x\theta)\,\mathrm{d}\theta = \pi \operatorname{sinc} x$, and $\operatorname{sinc} x = \frac{\sin(\pi x)}{\pi x}$ when $x \neq 0$, and 1 when $x = 0$. One can easily verify that there are no integer solutions to the Diophantine system $a^2 + b^2 = c^2 + d^2$ and $a = b + c + d + 2$. Thus,

$$\int_{-\pi}^{\pi} \mathrm{d}\varphi \int_0^{\pi} \mathrm{d}\theta \langle a|\cos\theta\rangle_\varphi \langle b|\cos\theta\rangle_\varphi \,{}_\varphi\langle\cos\theta|c\rangle\,{}_\varphi\langle\cos\theta|d\rangle = \delta_{a^2+b^2,c^2+d^2}\left(\delta_{a+d,b+c} + \delta_{a+c,b+d} + \delta_{a+b,c+d}\right). \tag{C.94}$$

We now focus on the three terms individually. The third term is solved in Lemma C.3.10 a $\delta_{a^2+b^2,c^2+d^2}\delta_{a+b,c+d} = \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc} - \delta_{ab}\delta_{ac}\delta_{ad}$. The first term is only nonzero when a, b, c, d solve

$$a^2 + b^2 = c^2 + d^2, \quad \text{and} \quad a + d = b + c. \tag{C.95}$$

Plugging $a = b+c-d$ in, we find $(b+c-d)^2+b^2 = c^2+d^2$, or equivalently $(b+c)(b-d) = 0$. Therefore, the first term is only nonzero when $b = d$ and $a = c$. A similar analysis holds for the second term, where we find that it only nonzero when $b = c$ and $a = d$. Hence, we find that

$$\int_{-\pi}^{\pi} \mathrm{d}\varphi \int_0^{\pi} \mathrm{d}\theta \langle a|\cos\theta\rangle_\varphi \langle b|\cos\theta\rangle_\varphi \,{}_\varphi\langle\cos\theta|c\rangle\,{}_\varphi\langle\cos\theta|d\rangle \tag{C.96a}$$
$$= \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc} + (\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc} - \delta_{ab}\delta_{ac}\delta_{ad}) \tag{C.96b}$$
$$= 2(\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}) - \delta_{ab}\delta_{ac}\delta_{ad} \tag{C.96c}$$

$$= 4\Pi_2(a,b;c,d) - \sum_{n\in\mathbb{N}_0} \langle a|n\rangle\langle b|n\rangle\langle n|c\rangle\langle n|d\rangle, \tag{C.96d}$$

proving the result. □

Theorem C.4.10. *Define the non-normalizable state*

$$|\sin\theta\rangle_\varphi \equiv {}_\varphi|\sin\theta\rangle := \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{N}_0} \mathrm{e}^{\mathrm{i}\varphi n^2} \left(\mathrm{e}^{\mathrm{i}(n+1)\theta} - \mathrm{e}^{-\mathrm{i}(n+1)(\theta-\pi)}\right) |n\rangle\,. \tag{C.97}$$

Then

$$\Pi_2 = \frac{1}{4}\sum_{n\in\mathbb{N}_0} \left(|n\rangle\langle n|\right)^{\otimes 2} + \frac{1}{4}\int_{-\pi}^{\pi} \mathrm{d}\varphi \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta \left({}_\varphi|\sin\theta\rangle\langle\sin\theta|_\varphi\right)^{\otimes 2} \tag{C.98}$$

Proof.

$$\int_{-\pi}^{\pi}\mathrm{d}\varphi\int_{-\pi/2}^{\pi/2}\mathrm{d}\theta \langle a|\sin\theta\rangle_\varphi\langle b|\sin\theta\rangle_\varphi\, {}_\varphi\langle\sin\theta|c\rangle\, {}_\varphi\langle\sin\theta|d\rangle \tag{C.99a}$$

$$= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\mathrm{d}\varphi\int_{-\pi/2}^{\pi/2}\mathrm{d}\theta\, \mathrm{e}^{\mathrm{i}\varphi(a^2+b^2-c^2-d^2)} \left(\mathrm{e}^{\mathrm{i}(a+1)\theta} - \mathrm{e}^{-\mathrm{i}(a+1)(\theta-\pi)}\right)\left(\mathrm{e}^{\mathrm{i}(b+1)\theta} - \mathrm{e}^{-\mathrm{i}(b+1)(\theta-\pi)}\right)$$
$$\times \left(\mathrm{e}^{-\mathrm{i}(c+1)\theta} - \mathrm{e}^{\mathrm{i}(c+1)(\theta-\pi)}\right)\left(\mathrm{e}^{-\mathrm{i}(d+1)\theta} - \mathrm{e}^{\mathrm{i}(d+1)(\theta-\pi)}\right) \tag{C.99b}$$

$$= \frac{1}{2}\delta_{a^2+b^2,c^2+d^2}\Big[\left((-1)^{a+b-c-d}+1\right)\mathrm{sinc}\left((a+b-c-d)/2\right)$$
$$+ (-1)^{b-c}\left((-1)^{a-b+c-d}+1\right)\mathrm{sinc}\left((a-b+c-d)/2\right)$$
$$+ (-1)^{a-c}\left((-1)^{a-b-c+d}+1\right)\mathrm{sinc}\left((a-b-c+d)/2\right)$$
$$+ \text{terms that are always zero when } a^2+b^2=c^2+d^2\Big] \tag{C.99c}$$

$$= \delta_{a^2+b^2,c^2+d^2}\left(\delta_{a+b,c+d} + (-1)^{b-c}\delta_{a+c,b+d} + (-1)^{a-c}\delta_{a+d,b+c}\right). \tag{C.99d}$$

From here we continue exactly as in Theorem C.4.9, and find that

$$\int_{-\pi}^{\pi}\mathrm{d}\varphi\int_{-\pi/2}^{\pi/2}\mathrm{d}\theta \langle a|\sin\theta\rangle_\varphi\langle b|\sin\theta\rangle_\varphi\, {}_\varphi\langle\sin\theta|c\rangle\, {}_\varphi\langle\sin\theta|d\rangle = 4\Pi_2(a,b;c,d) - \sum_{n\in\mathbb{N}_0}\langle a|n\rangle\langle b|n\rangle\langle n|c\rangle\langle n|d\rangle \tag{C.100}$$

completing the proof. □

We now state some facts about the non-normalizable states used above, the proof of which can be found in [199, sec. 6]. Our definition of these states differs from the definition in [199] in that we have added an additional phase factor $\mathrm{e}^{\mathrm{i}\varphi\hat{n}^2}$, but this does not affect any of the following facts. For the following, we will use δ to denote the Dirac delta function on the interval $[-\pi, \pi]$, namely $\delta(\theta) \equiv \delta_{[-\pi,\pi]}(\theta) = \frac{1}{2\pi}\sum_{j\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\theta j}$.

$$ {}_{\varphi}\langle \cos\theta | \cos\theta' \rangle_{\varphi} = \delta(\theta - \theta') \tag{C.101a} $$

$$ {}_{\varphi}\langle \sin\theta | \sin\theta' \rangle_{\varphi} = \delta(\theta - \theta') \tag{C.101b} $$

$$ \int_0^{\pi} {}_{\varphi}|\cos\theta\rangle\langle\cos\theta|_{\varphi} \, \mathrm{d}\theta = \mathbb{I} \tag{C.101c} $$

$$ \int_0^{\pi} {}_{\varphi}|\sin\theta\rangle\langle\sin\theta|_{\varphi} \, \mathrm{d}\theta = \mathbb{I} \tag{C.101d} $$

$$ \int_0^{\pi} {}_{\varphi}|\theta\rangle\langle\theta|_{\varphi} \, \mathrm{d}\theta = \mathbb{I}. \tag{C.101e} $$

From these, we see that the cos and sin states form a generalized orthogonal basis for each φ, and the cos, sin, and θ states form a continuous PVM for each φ. More generally, one can consider γ-rotated sine/cosine states

$$ |\theta\rangle_{\varphi,\gamma} \coloneqq \frac{1}{\sqrt{8}} \sum_{n=0}^{\infty} \mathrm{e}^{\mathrm{i}\varphi n^2} \left(\mathrm{e}^{\mathrm{i}(n+1)\theta} - \mathrm{e}^{-\mathrm{i}(n+1)(\theta-\gamma)} \right) |n\rangle \,, \tag{C.102} $$

where $|\theta\rangle_{\varphi,0} = |\cos\theta\rangle_{\varphi}$ and $|\theta\rangle_{\varphi,\pi} = |\sin\theta\rangle_{\varphi}$. Similar to above, for any fixed γ, summing over Fock states and integrating the $|\theta\rangle_{\varphi,\gamma}$ states over θ and φ yields a rigged 2-design.

Next, we restrict our attention to the $|\cos\theta\rangle_{\varphi}$, since a similar analysis holds for $|\sin\theta\rangle_{\varphi}$ and $|\theta\rangle_{\varphi}$. Let $\hat{a}$ and $\hat{a}^\dagger$ be the standard annihilation and creation operators so that $\hat{n} = \hat{a}^\dagger\hat{a}$. The elements of $\{|\cos\theta\rangle_{\varphi} \mid \theta \in (0,\pi)\}$ are generalized eigenvectors – or more precisely tempered distributions – of the operator $\widehat{\cos(\theta)}_{\varphi} \coloneqq \frac{1}{2}\hat{a}^\dagger(\hat{n}+1)^{-1/2}\mathrm{e}^{\mathrm{i}\varphi(2\hat{n}+1)} + \text{h.c.}$, where h.c. denotes the Hermitian conjugate of the first term. This can be concisely expressed in

terms of the Susskind-Glogower phase operator [198, 202] $\widehat{\mathrm{e}^{\mathrm{i}\theta}} \coloneqq \sum_{n\in\mathbb{N}_0} |n+1\rangle\langle n|$, yielding $\widehat{\cos(\theta)}_\varphi = \frac{1}{2}\widehat{\mathrm{e}^{\mathrm{i}\theta}}\mathrm{e}^{\mathrm{i}\varphi(2\hat{n}+1)} + \mathrm{h.c.}$.

The standard position state $\langle x|$ is to be understood as a distribution in that it is a continuous linear functional on a subset of $L^2(\mathbb{R})$ that will be described below. It is defined via the relation $\langle x|\psi\rangle \coloneqq \psi(x) = \int_{\mathbb{R}} \psi(x')\delta(x'-x)\,\mathrm{d}x'$. Similarly, the momentum states $\langle p|$ is understood as a distribution defined by $\langle p|\psi\rangle \coloneqq \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}px}\psi(x)\,\mathrm{d}x$.

The ${}_\varphi\langle\cos\theta|$ can be understood analogously; namely, ${}_\varphi\langle\cos\theta|$ is defined by

$$ {}_\varphi\langle\cos\theta|\psi\rangle \coloneqq \sqrt{\frac{2}{\pi}}\sum_{n\in\mathbb{N}_0}\sin((n+1)\theta)\mathrm{e}^{-\mathrm{i}\varphi n^2}\langle n|\psi\rangle, \tag{C.103}$$

where $\langle n|\psi\rangle$ is the standard inner product $\int_{\mathbb{R}} \bar{\psi}_n(x)\psi(x)\,\mathrm{d}x$ with $\psi_n(x) = \langle x|n\rangle$ the Fock state wavefunctions.

We now formalize this intuitive understanding of the $\langle\cos\theta|_\varphi$ states as distributions. References for this discussion are [244] for a formal treatment, and [80, 245] for a broad overview. We have found that the combination of Fock states with $|\cos\theta\rangle_\varphi$, $|\sin\theta\rangle_\varphi$, or $|\theta\rangle_\varphi$ distributions is enough to reconstruct Π_2. We call such designs *rigged* designs, since the latter states live in the rigged Hilbert space on top of $L^2(\mathbb{R})$.

The standard rigged Hilbert space of the harmonic oscillator is the Gelfand triple $S(\mathbb{R}) \subset L^2(\mathbb{R}) \subset S(\mathbb{R})'$. $S(\mathbb{R})$ is called *Schwartz space*, and as a topological vector space it has a continuous dual space. $S(\mathbb{R})'$ is called the space of *tempered distributions*, and is the continuous dual space of $S(\mathbb{R})$. For physical quantum states, we often desire that they have finite position, momentum, and energy moments. The first part of the Gelfand triple is the set of all such states, which for the harmonic oscillator is hence

$$ S(\mathbb{R}) = \bigcap_{\alpha,\beta\in\mathbb{N}_0} \mathcal{D}(\hat{x}^\alpha\hat{p}^\beta), \tag{C.104}$$

where $\mathcal{D}(M)$ denotes the maximal domain of the operator M. Since the domains of $\hat{x}$ and $\hat{p}$

are dense in $L^2(\mathbb{R})$, it follows that $S(\mathbb{R})$ is dense in $L^2(\mathbb{R})$.

$S(\mathbb{R})$ is a Fréchet space, meaning that it is a topological vector space with a topology induced by a countable family of seminorms $\|\cdot\|_{\alpha,\beta}$ defined by

$$\|f\|_{\alpha,\beta} \coloneqq \sup_{x\in\mathbb{R}} \left| x^\alpha \, \frac{\mathrm{d}^\beta f}{\mathrm{d}x^\beta} \right| \, . \tag{C.105}$$

An equivalent condition for a function f to belong to $S(\mathbb{R})$ is that $\|f\|_{\alpha,\beta} < \infty$ for all $\alpha, \beta \in \mathbb{N}_0$. The topology induced by the seminorms is equivalent to the topology induced by the metric [244, pg. 29]

$$d(f,g) = \sum_{\alpha,\beta\in\mathbb{N}_0} 2^{-\alpha-\beta} \frac{\|f-g\|_{\alpha,\beta}}{1+\|f-g\|_{\alpha,\beta}}. \tag{C.106}$$

Equipped with the metric, we can check continuity of a map $T\colon S(\mathbb{R}) \to \mathbb{C}$ in the usual way. T is continuous if for all sequences $(f_n)_{n\in\mathbb{N}}$, $f_n \xrightarrow{n\to\infty} f$ implies $T(f_n) \xrightarrow{n\to\infty} T(f)$. Strictly speaking, this is the definition of sequentially continuous, but continuity and sequential continuity are equivalent on metric spaces. Here $f_n \xrightarrow{n\to\infty} f$ means that $\forall \epsilon > 0, \exists N \in \mathbb{N}$ such that $\forall n \geq N\colon \; d(f_n, f) < \epsilon$, and $T(f_n) \xrightarrow{n\to\infty} T(f)$ means similarly but with the metric on $\mathbb{C}$.

We are now interested in characterizing the continuous dual of $S(\mathbb{R})$, denoted by $S(\mathbb{R})'$, which is a subset of the algebraic dual. Hence, we restrict our attention to linear maps $T\colon S(\mathbb{R}) \to \mathbb{C}$. When T is linear, $T(f) - T(g) = T(f-g)$. We also notice that $d(f,g) = 0$ if and only if $\|f-g\|_{\alpha,\beta} = 0$ for all α, β. Therefore, the condition that T be a tempered distribution, meaning that $T \in S(\mathbb{R})'$, is that it is linear and satisfies

$$\left(\forall \alpha, \beta \in \mathbb{N}_0\colon \lim_{m\to\infty} \|f_m\|_{\alpha,\beta} = 0\right) \implies \left(\lim_{m\to\infty} |T(f_m)| = 0\right) \tag{C.107}$$

for any sequence of functions $(f_m)_{m\in\mathbb{N}} \subset S(\mathbb{R})$.

As described above, a rigged Hilbert space is a triplet $S(\mathbb{R}) \subset L^2(\mathbb{R}) \subset S(\mathbb{R})'$. $S(\mathbb{R})'$ is the "bra space" of tempered distributions. One can analogously construct the "ket space" of *anti*linear continuous functionals on $S(\mathbb{R})$. This space is often denoted as $S(\mathbb{R})^\times$.

We will now revisit the $\langle \cos\theta|_\varphi$, $\langle \sin\theta|_\varphi$, and $\langle \theta|_\varphi$ states and show that they each belong to $S(\mathbb{R})'$, while their ket counterparts belong to $S(\mathbb{R})^\times$. As before, we will restrict our attention to the $\langle \cos\theta|_\varphi$ states, as the others are analogous. By construction, $\langle \cos\theta|_\varphi$ is clearly linear. We now show that it is continuous. We use Eq. (C.107), and compute

$$\begin{aligned}
\lim_{m\to\infty} \left| {}_\varphi\langle \cos\theta | f_m\rangle \right| &= \lim_{m\to\infty} \left| \sqrt{\frac{2}{\pi}} \sum_{n\in\mathbb{N}_0} \mathrm{e}^{-\mathrm{i}\varphi n^2} \sin((n+1)\theta) \langle n|f_m\rangle \right| && \text{(C.108a)} \\
&\leq \lim_{m\to\infty} \sum_{n\in\mathbb{N}_0} |\langle n|f_m\rangle| && \text{(C.108b)} \\
&= \lim_{m\to\infty} \sum_{n\in\mathbb{N}_0} \frac{1}{\sqrt{n!}} |\langle 0|\, \hat{a}^n\, |f_m\rangle| && \text{(C.108c)} \\
&= \lim_{m\to\infty} \sum_{n\in\mathbb{N}_0} \frac{1}{\sqrt{n!}} \left| \int_{\mathbb{R}} \psi_0(x) \left(x + \frac{\mathrm{d}}{\mathrm{d}x}\right)^n f_m(x)\, \mathrm{d}x \right| && \text{(C.108d)} \\
&\leq \lim_{m\to\infty} \sum_{n\in\mathbb{N}_0} \frac{1}{\sqrt{n!}} \left(\int_{\mathbb{R}} |\psi_0(x)|\, \mathrm{d}x \right) \left(\sup_x \left| \left(x + \frac{\mathrm{d}}{\mathrm{d}x}\right)^n f_m(x) \right| \right) && \text{(C.108e)} \\
&\propto \lim_{m\to\infty} \sup_x \left| \left(x + \frac{\mathrm{d}}{\mathrm{d}x}\right)^n f_m(x) \right| && \text{(C.108f)} \\
&= 0. && \text{(C.108g)}
\end{aligned}$$

The last line comes by assumption from Eq. (C.107). The second to last line comes from the facts that $\sum_n 1/\sqrt{n!} < \infty$ and that $\int_{\mathbb{R}} |\psi_0(x)|\, \mathrm{d}x < \infty$, where $\psi_0(x)$ is the position representation of the lowest Fock state as described in Appendix C.2. Hence, $\langle \cos\theta|_\varphi$ is a tempered distribution.

From Theorem C.4.5, we know that CV 2-designs do not exist. However, Theorems C.4.8 to C.4.10 show that rigged CV 2-designs do indeed exist, where we define a rigged CV design analogously to a standard CV design with the additional feature that tempered distributions

are allowed.

Subsection C.4.4: Regularized rigged state designs – making rigged state designs physical

Suppose we have a construction of Π_t in terms of unphysical states, so that $\Pi_t = \int_X (|\chi\rangle\langle\chi|)^{\otimes t} \, \mathrm{d}\mu$, where X is some measure space with measure μ. We use χ to denote possibly non-normalizable states, and ψ to denote properly normalized states. Define a Hermitian operator R which we'll call the regularizer. For example, R could be $\mathrm{e}^{-\beta\hat{n}}$, where $\hat{n}$ is the number operator diagonal in the Fock basis $\hat{n}\,|i\rangle = i\,|i\rangle$. Then,

$$\Pi_t^{(R)} \coloneqq R^{\otimes t} \Pi_t R^{\otimes t} \tag{C.109a}$$

$$= \int_X (R\,|\chi\rangle\langle\chi|\,R)^{\otimes t} \, \mathrm{d}\mu. \tag{C.109b}$$

As long as the amplitudes of each $|\chi\rangle$ do not grow too fast (indeed their growth is constrained by the condition that $|\chi\rangle$ be a tempered distribution; see below), the states $R\,|\chi\rangle$ will be normalizable. Define the normalized state corresponding to the tempered distribution χ as $|\psi\rangle \coloneqq R\,|\chi\rangle \,/\, \|R\,|\chi\rangle\|$. Then

$$\Pi_t^{(R)} = \int_X (|\psi\rangle\langle\psi|)^{\otimes t} \, \|R\,|\chi\rangle\|^{2t} \, \mathrm{d}\mu. \tag{C.110}$$

One can then define a new measure ν which is μ weighted by the positive factor $\|R\,|\chi\rangle\|^{2t} \,/\, \operatorname{Tr} \Pi_t^{(R)}$ (one can imagine using a Lebesgue-Stieltjes measure construction), thus giving

$$\frac{\Pi_t^{(R)}}{\operatorname{Tr} \Pi_t^{(R)}} = \int_X (|\psi\rangle\langle\psi|)^{\otimes t} \, \mathrm{d}\nu. \tag{C.111}$$

The first thing to note is that by taking the trace of both sides one finds that $\nu(X) = 1$. Hence the measure space defined by X and ν is a proper probability space. Next, suppose

that $R = \mathrm{e}^{-\beta\hat{n}}$. The parameter β is an inverse energy. $1/\beta$ fixes an energy scale of the states involved in the design. As $1/\beta \to \infty$, the energy of the states becomes infinite, and $\Pi_t^{(R)}$ looks more and more like $\Pi_t = \Pi_t^{(\mathbb{I})}$. When β is exactly zero, the equation becomes uninteresting since $\Pi_t / \operatorname{Tr} \Pi_t$ is just the zero operator.

Nevertheless, β is a parameter that one can tune that enforces a soft energy cutoff. The smaller one tunes β, the more the ensemble resembles a continuous-variable state t-design. The soft energy cutoff $\mathrm{e}^{-\beta\hat{n}}$ was chosen to ensure physicality of the resulting states. In particular, $\mathrm{e}^{-\beta\hat{n}}$ will always take a tempered distribution to a state in $L^2(\mathbb{R})$, whereas, for example, a soft-cutoff of the form $(\hat{n}+1)^{-b}$ for some $b > 0$ will not always achieve this. We therefore use $\mathrm{e}^{-\beta\hat{n}}$ to make any rigged design into an approximate design composed of physical states. This is formalized in the following proposition.

Proposition C.4.11. *If $|\chi\rangle$ is a tempered distribution, then $\mathrm{e}^{-\beta\hat{n}}|\chi\rangle$ is a state in $L^2(\mathbb{R})$ for any $\beta > 0$.*

Proof sketch. From [308, Thm. 3], any tempered distribution can be expressed as $|\chi\rangle = \sum_{n\in\mathbb{N}_0} a_n |n\rangle$. We first calculate the norm of $\mathrm{e}^{-\beta\hat{n}}|\chi\rangle$;

$$\langle\chi|\,\mathrm{e}^{-2\beta\hat{n}}\,|\chi\rangle = \sum_{n\in\mathbb{N}_0} |a_n|^2\,\mathrm{e}^{-2\beta n}. \tag{C.112}$$

We therefore find that $\mathrm{e}^{-\beta\hat{n}}|\chi\rangle \in L^2(\mathbb{R})$ as long as $|a_n|$ grows with n asymptotically slower than exponential. Hence, to prove the proposition, we need to show that if a_n grows exponentially or faster in n, then $|\chi\rangle$ is not a tempered distribution. This is proven in [308, Thm. 3]. For completeness, we show it here as well. We will use Eq. (C.107) to show this.

Fix some sequence $(f_m)_{m\in\mathbb{N}}$ of states $f_m \in S(\mathbb{R})$ satisfying

$$\forall \alpha,\beta \in \mathbb{N}_0 \colon \lim_{m\to\infty} \|f_m\|_{\alpha,\beta} = 0. \tag{C.113}$$

Specifically, let $|f_m\rangle = \mathrm{e}^{-m}\sum_{n\in\mathbb{N}_0}\mathrm{e}^{-\varepsilon n}|n\rangle$ for some arbitrarily small $\varepsilon > 0$. Then, assuming the best case where a_n grows exponentially as $a_n = \mathrm{e}^{\mathrm{i}\theta_n}\mathrm{e}^{\gamma n}$ for some $\gamma > 0$,

$$\lim_{m\to\infty}|\langle\chi|f_m\rangle| = \lim_{m\to\infty}\left|\sum_{n\in\mathbb{N}_0}\bar{a}_n\mathrm{e}^{-m-\varepsilon n}\right| \tag{C.114a}$$

$$= \lim_{m\to\infty}\mathrm{e}^{-m}\left|\sum_{n\in\mathbb{N}_0}\mathrm{e}^{-\mathrm{i}\theta_n}\mathrm{e}^{n(\gamma-\varepsilon)}\right|. \tag{C.114b}$$

Since ε can be arbitrarily small, we can always choose it so that $\gamma-\varepsilon > 0$, and therefore the sum diverges no matter the choices of the phases θ_n. Hence,

$$\lim_{m\to\infty}|\langle\chi|f_m\rangle| \neq 0, \tag{C.115}$$

proving, by Eq. (C.107), that $|\psi\rangle$ is not a tempered distribution. □

This proposition justifies our choice $\mathrm{e}^{-\beta\hat{n}}$ as the soft energy cutoff, since a cutoff such as $(1+\hat{n})^{-b}$ does not satisfy the proposition for any b. However, there do exist rigged designs for which $(1+\hat{n})^{-b}$ *is* sufficient. For example, the $|\theta\rangle_\varphi$, $|\cos\theta\rangle_\varphi$, and $|\sin\theta\rangle_\varphi$ are all tempered distributions that generate rigged 2-designs, and $(1+\hat{n})^{-2}|\theta\rangle_\varphi, (1+\hat{n})^{-2}|\cos\theta\rangle_\varphi, (1+\hat{n})^{-2}|\sin\theta\rangle_\varphi \in L^2(\mathbb{R})$. Hence, one may suggest that for these rigged designs, one should use $(1+\hat{n})^{-b}$ as a soft energy cutoff in place of $\mathrm{e}^{-\beta\hat{n}}$. However, one desirable property of physical quantum states is that all position, momentum, and energy moments are finite. In other words, one may desire that the states belong to $S(\mathbb{R}) \subset L^2(\mathbb{R})$. One can straightforwardly show that, for example, $(1+\hat{n})^{-b}|\theta\rangle_\varphi \notin S(\mathbb{R})$ for any b, whereas $\mathrm{e}^{-\beta\hat{n}}|\theta\rangle_\varphi \in S(\mathbb{R})$. This is another justification for the use of $\mathrm{e}^{-\beta\hat{n}}$.

Example C.4.12. Consider, for example, the rigged design given Theorem C.4.8. Sandwiching the design with R results in the normalized states $\sqrt{1-\mathrm{e}^{-2\beta}}\sum_{n\in\mathbb{N}_0}\mathrm{e}^{-\beta n+\mathrm{i}\theta n+\mathrm{i}\varphi n^2}|n\rangle$. Each of these states has energy $\coth(\beta)/2-1/2$. The design also still consists of the original

Fock states $|n\rangle$, but the weight in front of each Fock state decays exponentially with n as $\sim \mathrm{e}^{-\beta n}$. Thus, despite the fact that the design uses arbitrarily high energy states (i.e. $|n\rangle$ for all natural numbers n), the weight factor in front of these high energy states is exponentially small in the energy. Therefore, the design effectively uses states finitely upper bounded in energy, where the bound is tuned by β. We refer to [246, sec. 5] for a review of these states, which are related to so-called phase coherent states. ◇

We consider now an R-regularized rigged t-design $\mathcal{G}$, which satisfies

$$\mathop{\mathbb{E}}_{\psi\in\mathcal{G}}(|\psi\rangle\langle\psi|)^{\otimes t} = \frac{\Pi_t^{(R)}}{\operatorname{Tr}\Pi_t^{(R)}} \,. \tag{C.116}$$

By tracing out e.g. the last factor, we find $\mathbb{E}_{\psi\in\mathcal{G}}(|\psi\rangle\langle\psi|)^{\otimes(t-1)} \propto \operatorname{Tr}_t \Pi_t^{(R)}$. Recall from Appendix C.2.2 that $\Pi_t = \frac{1}{t!}\sum_{\sigma\in S_t} W_\sigma$. Consider a permutation $\sigma \in S_t$ that leaves the last factor fixed. Let $\pi \in S_{t-1}$ be the permutation with the same cyclic decomposition as σ. For example, when $t = 3$ and $\sigma = (12)(3)$ is the permutation swapping 1 and 2 and leaving 3 fixed, then we set $\pi = (12)$. We see that for such a σ, $\operatorname{Tr}_t(R^{\otimes t} W_\sigma R^{\otimes t}) = (\operatorname{Tr} R^2) R^{\otimes(t-1)} W_\pi R^{\otimes(t-1)}$. Hence, the sum over all such permutations results in $(\operatorname{Tr} R^2)\sum_{\pi\in S_{t-1}} R^{\otimes(t-1)} W_\pi R^{\otimes(t-1)} = (\operatorname{Tr} R^2)\Pi_{t-1}^{(R)}$. For all other permutations τ that do not leave the t^{th} factor fixed, $\operatorname{Tr}_t W_\tau$ does not pick up a factor of $(\operatorname{Tr} R^2)$. We have hence found that $\mathbb{E}_{\psi\in\mathcal{G}}(|\psi\rangle\langle\psi|)^{\otimes(t-1)} \sim (\operatorname{Tr} R^2)\Pi_{t-1}^{(R)} +$ (terms without $(\operatorname{Tr} R^2)$). Assuming that the regularizer R is close to the identity so that $(\operatorname{Tr} R^2)$ is large and applying the above arguments to both the numerator and denominator of Eq. (C.116), we have thus found that an R-regularized rigged t-design $\mathcal{G}$ satisfies

$$\mathop{\mathbb{E}}_{\psi\in\mathcal{G}}(|\psi\rangle\langle\psi|)^{\otimes(t-1)} = \frac{\Pi_{t-1}^{(R)}}{\operatorname{Tr}\Pi_{t-1}^{(R)}}\left(1 + \mathcal{O}\big(1/\operatorname{Tr} R^2\big)\right). \tag{C.117}$$

It is in this sense that an R-regularized rigged t-design is *almost* an R-regularized rigged $(t-1)$-design up to factors of $1/\operatorname{Tr} R^2$.

In the special case when $R = P_d = \sum_{n=0}^{d-1} |n\rangle\langle n|$, a P_d-regularized rigged t-design $\mathcal{G}$ is

simply a $\mathbb{CP}^{d-1}$ t-design, and hence it is also *exactly* a $(t-1)$-design. However, when R is an invertible operator, the result is only a $(t-1)$-design up to terms of order $1/\operatorname{Tr} R^2$.

Frame potential

In this section, we will generalize the well-known *frame potential* from finite-dimensional state designs [41] to regularized rigged designs. Specifically, for a positive definite regularizer R, we define the frame potential of an ensemble $\mathcal{G}$ (i.e. a probability space over unit vectors in $L^2(\mathbb{R})$) to be

$$V_t^{(R)}(\mathcal{G}) \coloneqq \mathop{\mathbb{E}}_{\psi,\phi\in\mathcal{G}} \left|\langle\psi| R^{-1} |\phi\rangle\right|^{2t}. \tag{C.118}$$

We prove the following proposition regarding R-regularized rigged t-designs and the frame potential.

Proposition C.4.13. *Let R be positive definite. For any ensemble $\mathcal{G}$,*

$$V_t^{(R)}(\mathcal{G}) \geq \frac{1}{\operatorname{Tr}\Pi_t^{(R)}}, \tag{C.119}$$

with equality if and only if $\mathcal{G}$ is an R-regularized rigged t-design.

Proof. This proof is a modification of that of Ref. [41, Eq. (3)]. Let $E \coloneqq \mathbb{E}_{\psi\in\mathcal{G}}(|\psi\rangle\langle\psi|)^{\otimes t}$ and $\xi \coloneqq (R^{-1})^{\otimes t}E - \Pi_t^{(\sqrt{R})}/\operatorname{Tr}\Pi_t^{(R)}$. By recalling the definitions of regularized-rigged designs and of the symmetric projector (C.10), we see that $\mathcal{G}$ is an R-regularized rigged t-design if and only if $\xi = 0$, or equivalently, $\operatorname{Tr}\xi^2 = 0$. We find that

$$0 \leq \operatorname{Tr}\xi^2 \tag{C.120a}$$

$$= \operatorname{Tr}\left[(R^{-1})^{\otimes t}E(R^{-1})^{\otimes t}E\right] + \frac{\operatorname{Tr}\left[(\Pi_t^{\sqrt{R}})^2\right]}{\left(\operatorname{Tr}\Pi_t^{(R)}\right)^2} - \frac{2}{\operatorname{Tr}\Pi_t^{(R)}}\operatorname{Tr}\left[E\Pi_t^{(\sqrt{R})}(R^{-1})^{\otimes t}\right] \tag{C.120b}$$

$$= V_t^{(R)}(\mathcal{G}) + \frac{\operatorname{Tr}\Pi_t^{(R)}}{(\operatorname{Tr}\Pi_t^{(R)})^2} - \frac{2}{\operatorname{Tr}\Pi_t^{(R)}}\operatorname{Tr}[E\Pi_t] \tag{C.120c}$$

$$= V_t^{(R)}(\mathcal{G}) + \frac{1}{\operatorname{Tr}\Pi_t^{(R)}} - \frac{2}{\operatorname{Tr}\Pi_t^{(R)}} \tag{C.120d}$$

$$= V_t^{(R)}(\mathcal{G}) - \frac{1}{\operatorname{Tr}\Pi_t^{(R)}}, \tag{C.120e}$$

with equality if and only if $\mathcal{G}$ is an R-regularized rigged t-design. In the second to last line, we used that $E\Pi_t = E$ and that $\operatorname{Tr} E = 1$. □

If R is instead only positive *semi*-definite and not invertible, then we can modify the definition of the frame potential to utilize the Moore-Penrose inverse R^+ in place of the inverse R^{-1}. The proposition then still holds as is, with the addition of the assumption that $RR^+\mathcal{G} = \mathcal{G}$, where recall RR^+ is a projector onto the support of R.

Notice the presence of the R^{-1} in the definition of the frame potential. We will also see such a presence in Appendix C.5.2 when generalizing fidelity quantities to infinite-dimensional spaces.

Subsection C.4.5: Alternative characterization of continuous-variable, rigged, and regularized rigged designs

To generate a random state $|\psi\rangle \in \mathbb{CP}^{d-1}$, one can equivalently choose d amplitudes $\{\alpha_i \in \mathbb{C} \mid i \in 0, \ldots, d-1$ where each α_i is drawn independently from the unit variance normal distribution $\mathcal{N}(0,1)$. The state $\frac{\sum_i \alpha_i |i\rangle}{\left\|\sum_i \alpha_i |i\rangle\right\|}$ is then a random state drawn from $\mathbb{CP}^{d-1}$.

Motivated by this and by Ref. [309, Sec. 4.1], we consider integration on the Frechét space $\mathbb{C}^\infty = \prod_{i\in\mathbb{N}_0} \mathbb{C}$ with the product topology. Define $\delta_j \colon \mathbb{C}^\infty \to \mathbb{C}$ to be the projections $\delta_j(x) = x_j$. Let Σ be the smallest σ-algebra on $\mathbb{C}^\infty$ such that δ_j is measurable for every j. Note that this corresponds to the Borel σ-algebra; that is, the product topology and the σ-algebra are both generated by sets of the form $A = \prod_{i\in\mathbb{N}_0} A_i$, where each A_i is an open subset of $\mathbb{C}$ and only finitely many A_i are proper.

Let $\mathcal{N}(0,\lambda_i)$ be the Gaussian measure on $\mathbb{C}$ with mean 0 and variance λ_i. Define the

measure $\mu\colon \Sigma \to [0,\infty]$ by $\mu \coloneqq \bigotimes_{j\in\mathbb{N}_0} \mathcal{N}(0,\lambda_j)$, where each $\lambda_j \in (0,\infty)$. The construction for such a measure is as follows. For $A = \prod_{i\in\mathbb{N}_0} A_i$ where all but finitely many A_i satisfy $A_i = \mathbb{C}$, define $\mu(A) = \prod_{i\in\mathbb{N}_0} \mathcal{N}(0,\lambda_i)(A_i)$. For every i for which $A_i = \mathbb{C}$, $\mathcal{N}(0,\lambda_i)(A_i) = 1$. Hence, μ is well-defined on such sets A since the product is finite. From its definition on such sets A, μ can be uniquely extended to all of Σ [276, Thm. 10.6.1].

Let $\{|n\rangle \mid n \in \mathbb{N}_0\}$ be a basis for $L^2(\mathbb{R})$. For $z \in \mathbb{C}^\infty$, let $|z\rangle \coloneqq \sum_{n\in\mathbb{N}_0} z_n |n\rangle$. Any tempered distribution can be expressed as $|z\rangle$ for some $z \in \mathbb{C}^\infty$ satisfying certain conditions [308, Thm. 3]. We therefore define the following subsets of $\mathbb{C}^\infty$:

$$S \coloneqq \{z \in \mathbb{C}^\infty \mid |z\rangle \in S(\mathbb{R})\} \tag{C.121a}$$

$$\ell^2_{\mathbb{C}}(\mathbb{N}_0) \coloneqq \left\{z \in \mathbb{C}^\infty \mid |z\rangle \in L^2(\mathbb{R})\right\} \tag{C.121b}$$

$$S' \coloneqq \{z \in \mathbb{C}^\infty \mid |z\rangle \in S(\mathbb{R})'\}\,. \tag{C.121c}$$

Lemma C.4.14. *Suppose that $\lambda_i = 1$ for all $i \in \mathbb{N}_0$. Then $\mu(S') = 1$ and therefore $\mu(\mathbb{C}^\infty \setminus S') = 0$.*

Proof. Note that

$$\int_{\mathbb{C}^\infty} \left\| \frac{1}{\hat{n}+1} |z\rangle \right\|^2 \mathrm{d}\mu(z) = \sum_{n\in\mathbb{N}_0} \frac{1}{(n+1)^2} \int_{\mathbb{C}^\infty} |z_n|^2 \,\mathrm{d}\mu(z) = \sum_{n\in\mathbb{N}_0} \frac{1}{(n+1)^2} < \infty. \tag{C.122}$$

Therefore, $\left\|\frac{1}{\hat{n}+1}|z\rangle\right\|^2 < \infty$ for almost all z. For any $z \in \mathbb{C}^\infty$, if $\left\|\frac{1}{\hat{n}+1}|z\rangle\right\| < \infty$, then $|z\rangle \in S(\mathbb{R})'$ [308, Thm. 3]. Hence $|z\rangle \in S(\mathbb{R})'$ μ-a.e. □

Through an analogous calculation with the integrand being $\||z\rangle\|^2$, one finds that if $\sum_{i\in\mathbb{N}_0} \lambda_i < \infty$, then $\mu(\ell^2_{\mathbb{C}}(\mathbb{N}_0)) = 1$ [309, Rem 4.1.2]. Define R to be the diagonal matrix with diagonal entries λ_i. If R is the identity, then $\mu(S') = 1$, while if R is trace-class, then $\mu(\ell^2_{\mathbb{C}}(\mathbb{N}_0)) = 1$.

It follows that if R is the identity, integrals over the measure space $(\mathbb{C}^\infty, \Sigma, \mu)$ are equal

to integrals over the restricted measure space $(S', \Sigma|_{S'}, \mu|_{S'})$. Similarly, if R is trace-class, integrals over the measure space $(\mathbb{C}^\infty, \Sigma, \mu)$ are equal to integrals over the restricted measure space $(\ell^2_{\mathbb{C}}(\mathbb{N}_0), \Sigma|_{S'}, \mu|_{S'})$.

Next, we show that when R is the identity, $(S', \Sigma|_{S'}, \mu|_{S'})$ is a rigged t-design for any $t \in \mathbb{N}$, and when R is trace-class, $(\ell^2_{\mathbb{C}}(\mathbb{N}_0), \Sigma|_{\ell^2_{\mathbb{C}}(\mathbb{N}_0)}, \mu|_{\ell^2_{\mathbb{C}}(\mathbb{N}_0)})$ is an R-regularized rigged t-design for any $t \in \mathbb{N}$. [1]

Given the construction of our measure space over $\mathbb{C}^\infty$, integrals over polynomials in z reduce to simple finite-dimensional Gaussian integration. For the purposes of designs, we are only interested in such polynomials. Consider

$$\int_{\mathbb{C}^\infty} \prod_{i=1}^{t} \langle a_i|z\rangle\langle z|b_i\rangle \, \mathrm{d}\mu(z) = \int_{\mathbb{C}^\infty} \prod_{i=1}^{t} z_{a_i} \bar{z}_{b_i} \, \mathrm{d}\mu(z). \tag{C.123}$$

for $a, b \in \mathbb{N}_0^t$. Since the integrand depends only on at most $2t$ elements of z, we can use Fubini's theorem so that the integral reduces to an integral over $\mathbb{C}^{2t}$ with the measure $\bigotimes \mathcal{N}(0, \lambda_i)$. Then, one can easily check by induction (or just by using standard properties of Gaussian integrals) that the integral equals $\Pi_t^{(\sqrt{R})}(a; b)$, and therefore

$$\int_{S'} \prod_{i=1}^{t} |z\rangle\langle z| \, \mathrm{d}\mu(z) = \Pi_t^{(\sqrt{R})} \tag{C.124}$$

in the weak sense. When R is the identity, $(S', \Sigma|_{S'}, \mu|_{S'})$ is a rigged t-design (for all $t \in \mathbb{N}_0$), and when R is trace-class, $(\ell^2_{\mathbb{C}}(\mathbb{N}_0), \Sigma|_{\ell^2_{\mathbb{C}}(\mathbb{N}_0)}, \mu|_{\ell^2_{\mathbb{C}}(\mathbb{N}_0)})$ is a $\sqrt{R}$-regularized rigged t-design (for all $t \in \mathbb{N}_0$).

[1] More accurately, the image of $(S', \Sigma|_{S'}, \mu|_{S'})$ under the map $z \mapsto |z\rangle$ – which is a measure space over $S(\mathbb{R})'$ – is a rigged t-design, and the image of $(\ell^2_{\mathbb{C}}(\mathbb{N}_0), \Sigma|_{\ell^2_{\mathbb{C}}(\mathbb{N}_0)}, \mu|_{\ell^2_{\mathbb{C}}(\mathbb{N}_0)})$ – which is a measure space over $L^2(\mathbb{R})$ – is a regularized rigged t-design.

Subsection C.4.6: Displaced Fock states as negative-weight approximate designs

The projection onto the two-body symmetric subspace is (see Appendix C.2.2)

$$\Pi_2 = \frac{1}{2}\left(\mathbb{I} + \mathrm{e}^{\mathrm{i}\frac{\pi}{2}\left(a^\dagger - b^\dagger\right)(a-b)}\right), \tag{C.125}$$

where the second operator in the parentheses is the SWAP operator, and a (b) represents the lowering operator for the first (second) mode. To simplify calculations, we apply the beam-splitter operation

$$U = \exp\left[\frac{\pi}{4}\left(a^\dagger b - a b^\dagger\right)\right], \qquad \text{acting as} \qquad U^\dagger \begin{pmatrix} a \\ b \end{pmatrix} U = \frac{1}{\sqrt{2}} \begin{pmatrix} a+b \\ b-a \end{pmatrix}, \tag{C.126}$$

which is equivalent to partitioning the two-mode Hilbert space into a tensor product of a center-of-mass $L^2(\mathbb{R})$ factor whose corresponding coordinate is symmetric under SWAP and an anti-symmetric factor whose coordinate is anti-symmetric [79, Sec. III]. In the Fock-space picture, this results in

$$U\Pi_2 U^\dagger = \mathbb{I} \otimes \frac{1 + \mathrm{e}^{\mathrm{i}\pi b^\dagger b}}{2} = \sum_{n\in\mathbb{N}_0} |n\rangle\langle n| \otimes \sum_{p\in\mathbb{N}_0} |2p\rangle\langle 2p|, \tag{C.127}$$

which now projects onto the entire symmetric factor and the even Fock-state subspace of the anti-symmetric factor.

We now determine what happens if one sums up two copies of all displaced versions of a particular Fock state $|\ell\rangle$. Using the fact that SWAP acts on displacements as $UD_\alpha^{\otimes 2}U^\dagger = D_\alpha \otimes \mathbb{I}$ and the fact that displacements form a unitary 1-design, we have

$$U\left(\int \frac{\mathrm{d}^2\alpha}{\pi} D_\alpha^{\otimes 2} |\ell\ell\rangle\langle\ell\ell| D_{-\alpha}^{\otimes 2}\right) U^\dagger = \int \frac{\mathrm{d}^2\alpha}{\pi} \left(D_\alpha \otimes \mathbb{I}\right) U|\ell\ell\rangle\langle\ell\ell|U^\dagger \left(D_\alpha^\dagger \otimes \mathbb{I}\right) \tag{C.128a}$$

$$= \mathrm{Tr}_1\left(U|\ell\ell\rangle\langle\ell\ell|U^\dagger\right) , \tag{C.128b}$$

where Tr_1 is the partial trace over the first factor.

We next write out U as a direct sum of irreducible representations of SU(2), with each representation acting on a sector of fixed total occupation number. Irreducible representations of SU(2) are known exactly in terms of the Wigner-D matrices [310], and the matrix elements we will need are

$$c_n^{(\ell)} = |\langle 2\ell - n, n|U|\ell, \ell\rangle|^2 = \left|D^\ell_{2n-\ell,0}\left(0, -\frac{\pi}{2}, 0\right)\right|^2 = \frac{(2\ell-2n)!\,(2n)!}{4^\ell\left[n!\,(\ell-n)!\right]^2} . \tag{C.129}$$

Plugging this in yields

$$U\left(\int \frac{\mathrm{d}^2\alpha}{\pi} D_\alpha^{\otimes 2}|\ell\ell\rangle\langle\ell\ell|D_{-\alpha}^{\otimes 2}\right)U^\dagger = \mathbb{I}\otimes\sum_{n=0}^{\ell} c_n^{(\ell)}|2n\rangle\langle 2n| . \tag{C.130}$$

When $\ell = 0$, we have $c_0^{(0)} = 1$, corroborating the result from [79, Sec. III]. For general ℓ, this result yields nonzero coefficients $c_n^{(\ell)}$ for all Fock states $\leq 2\ell$ in the anti-symmetric factor.

We now linearly combine instances of Eq. (C.130) with ℓ from zero to some $\ell_{\max}$ and compensate the $c_n^{(\ell)}$ using weights b_ℓ in front of each Fock state. This yields

$$U\left(\sum_{\ell=0}^{\ell_{\max}} b_\ell \int \frac{\mathrm{d}^2\alpha}{\pi} D_\alpha|\ell\ell\rangle\langle\ell\ell|D_\alpha^\dagger\right)U^\dagger = \mathbb{I}\otimes\sum_{n=0}^{\ell_{\max}}|2n\rangle\langle 2n| , \tag{C.131a}$$

$$b_\ell = \frac{1 - \sum_{p=\ell+1}^{\ell_{\max}} b_p c_\ell^{(L)}}{c_\ell^{(\ell)}} , \tag{C.131b}$$

which yields Π_2 up to the Fock state 2ℓ in the anti-symmetric factor. However, some of the b_ℓ's are negative, meaning that the right-hand side of Eq. (C.131a) cannot be treated as an expectation value of operators sampled according to a probability distribution. The ensemble can be formulated in terms of a measure space with a signed measure, and there

may be schemes to sample from such an ensemble [311]. Thus, displaced Fock states form a hard-energy regularized 2-design with a signed measure.

One may be tempted to take ℓ_{max} to infinity. In this case, the coefficient $b_{\ell_{max}} \to \infty$, showing that this regularized design does not yield a CV design and corroborating the no-go Theorem C.4.5 (recall we extended Theorem C.4.5 to the case of signed measure spaces at the end of Appendix C.4.1).

Subsection C.4.7: Approximate continuous-variable unitary designs

In finite dimensions, a unitary design reconstructs the superoperator

$$\mathcal{P}_t = \frac{1}{|S_t|} \sum_{\sigma \in S_t} \|W_\sigma\rangle\langle W_\sigma\|. \tag{C.132}$$

In this way, when acting on a fiducial state $\rho = |\phi\rangle\langle\phi|$, one finds

$$\begin{aligned}
\mathcal{P}_t \|\rho^{\otimes t}\rangle &= \frac{1}{|S_t|} \sum_{\sigma \in S_t} \|W_\sigma\rangle\langle W_\sigma\|\rho^{\otimes t}\rangle && \text{(C.133a)} \\
&= \frac{1}{|S_t|} \sum_{\sigma \in S_t} \|W_\sigma\rangle \operatorname{Tr}\left[W_\sigma^{-1}\rho^{\otimes t}\right] && \text{(C.133b)} \\
&= \frac{1}{|S_t|} \sum_{\sigma \in S_t} \|W_\sigma\rangle && \text{(C.133c)} \\
&= \Pi_t. && \text{(C.133d)}
\end{aligned}$$

From above, we have states in $L^2(\mathbb{R})$ that construct the normalized symmetric projector $\Pi_t^{(R)} = R^{\otimes t}\Pi_t R^{\otimes t}$. In a similar way, let's normalize the superoperator $\mathcal{P}_t$. Define

$$\mathcal{P}_t^{(R)} \coloneqq \frac{1}{|S_t|} \sum_{\sigma \in S_t} \|R^{\otimes t}W_\sigma\rangle\langle R^{\otimes t}W_\sigma\|. \tag{C.134}$$

Then, when acting on a fiducial state $\rho = |\phi\rangle\langle\phi|$, one finds

$$\mathcal{P}_t^{(R)} \| \rho^{\otimes t} \rangle = \frac{1}{|S_t|} \sum_{\sigma \in S_t} \| R^{\otimes t} W_\sigma \rangle \langle R W_\sigma \| \rho^{\otimes t} \rangle \tag{C.135a}$$

$$= \frac{1}{|S_t|} \sum_{\sigma \in S_t} \| R^{\otimes t} W_\sigma \rangle \mathrm{Tr}[R^{\otimes t} W_\sigma^{-1} \rho^{\otimes t}] \tag{C.135b}$$

$$= \mathrm{Tr}[(R\rho)^{\otimes t}] \Pi_t^{(R)} \tag{C.135c}$$

$$\propto \Pi_t^{(R)}. \tag{C.135d}$$

With this, we now define an approximate continuous-variable unitary t-design to be a collection of unitaries $U_i \colon L^2(\mathbb{R}) \to L^2(\mathbb{R})$ that satisfy

$$\sum_i \left(\| U_i \rangle \langle U_i \| \right)^{\otimes t} = \mathcal{P}_t^{(R)}. \tag{C.136}$$

As with rigged state designs, the parameterization i of the unitaries, represented here heuristically as a sum, may constitute a measure space. We leave determination of existence of such designs to future work.

Section C.5: Applications of rigged and regularized rigged designs

Subsection C.5.1: Continuous-variable shadows

In this subsection, we use rigged designs to construct infinite-dimensional classical shadows of a quantum state ρ. With these shadows, one can for example efficiently compute the expectation value of many observables. Ref. [52] phrased shadow tomography from Ref. [50] in terms of informationally-complete POVMs. We will generalize their discussion to infinite dimensions.

Specifically, suppose that the measure space (X, Σ, μ) is a rigged 3-design. In other words,

$$\int_X (|\chi\rangle\langle\chi|)^{\otimes t} \, \mathrm{d}\mu(\chi) = \alpha_t \Pi_t \tag{C.137}$$

for each $t \in \{1, 2, 3\}$, where $\alpha_1, \alpha_2, \alpha_3 \in (0, \infty)$ are some numbers. We assume without loss of generality that $\alpha_1 = 1$ (if not, just rescale the measure). Recall that we will use $|\chi\rangle$ to denote tempered distributions and $|\psi\rangle$ to denote physical quantum states.

Let $\nu \colon \Sigma \to P(\mathcal{H})$, where $P(\mathcal{H})$ denotes the set of nonnegative operators on an underlying separable, infinite dimensional Hilbert space $\mathcal{H}$, and define

$$\nu(A) \coloneqq \int_A |\chi\rangle\langle\chi| \, \mathrm{d}\mu(\chi) \ . \tag{C.138}$$

This map is a positive operator-valued measure (POVM) because it satisfies the axioms

1. $\nu(X) = \mathbb{I}$,
2. $\nu(\emptyset) = 0$, and
3. $\nu(\bigcup_i A_i) = \sum_i \nu(A_i)$ for countable collections of disjoint $A_i \in \Sigma$.

The first axiom is satisfied since X is a rigged 1-design with $\alpha_1 = 1$. The second axiom is trivially satisfied. The third axiom follows from the σ-additivity of the measure μ.

We can therefore measure a state ρ with respect to the POVM ν. As usual, associated to the POVM is a standard probability measure μ' defined by $\mu'(A) = \mathrm{Tr}[\rho\nu(A)]$. When measuring the state ρ with the POVM ν, we sample outcomes labeled by $\chi \in X$ from the probability measure μ'. Indeed, we have the freedom to label the outcomes however we choose. In particular, suppose that to each tempered distribution (i.e. non-normalizable, and therefore unphysical, quantum state) $|\chi\rangle \in X$, we associate a physical state $|\psi_\chi\rangle \in \mathcal{H}$ of unit norm. Then the measurement channel representing the POVM ν is $\rho \mapsto \int_X |\psi_\chi\rangle\langle\psi_\chi| \, \mathrm{d}\mu'(\chi)$.

In the realm of shadow tomography, however, we have even more freedom than this. Once we measure from the POVM, we store a shadow on a classical computer and never need to physically prepare the shadow. Therefore, we do not *need* to associate physical states $|\psi_\chi\rangle$ to the measurement outcomes of the POVM; we are free to associate the unphysical tempered distributions $|\chi\rangle$ to the measurement outcome corresponding to χ. The resulting map representing the measurement process is then

$$\mathcal{M}(\rho) = \int_X |\chi\rangle\langle\chi| \, \mathrm{d}\mu'(\chi). \tag{C.139}$$

Since μ' is a probability measure, we will define the notation $\mathbb{E}_{\chi\in X'}(\cdot) \coloneqq \int_X (\cdot)\, \mathrm{d}\mu'(\chi)$. Hence,

$$\mathcal{M}(\rho) = \mathop{\mathbb{E}}_{\chi\in X'} |\chi\rangle\langle\chi| \,. \tag{C.140}$$

$\mathcal{M}$ is not a physical quantum channel; indeed, $\operatorname{Tr}\mathcal{M}(\rho)$ is not finite. However, $\mathcal{M}$ represents the process of measuring ρ with respect to the physical POVM ν and storing the result classically. This part of the formalism, namely associating the infinite-trace operator $|\chi\rangle\langle\chi|$ to the measurement outcome χ, is the only part that differs from the finite dimensional case. In the finite dimensional case, the designs contain only physical states $|\psi\rangle$, and a physical density matrix $|\psi\rangle\langle\psi|$ is associated to the measurement outcome ψ. Ultimately, since this part of the procedure is being done classically, this difference is inconsequential, and we continue exactly as we would in the finite dimensional case.

Using the fact that X is a rigged 2-design, we can evaluate

$$\begin{aligned}
\mathcal{M}(\rho) &= \int_X |\chi\rangle\langle\chi| \, \mathrm{d}\mu'(\chi) && \text{(C.141a)}\\
&= \int_X |\chi\rangle\langle\chi| \operatorname{Tr}[\rho\,|\chi\rangle\langle\chi|] \, \mathrm{d}\mu(\chi) && \text{(C.141b)}\\
&= \operatorname{Tr}_1\left[(\rho\otimes\mathbb{I}) \int_X (|\chi\rangle\langle\chi|)^{\otimes 2} \, \mathrm{d}\mu(\chi)\right] && \text{(C.141c)}
\end{aligned}$$

$$= \alpha_2 \operatorname{Tr}_1[(\rho \otimes \mathbb{I})\Pi_2] \tag{C.141d}$$

$$= \frac{\alpha_2}{2} \operatorname{Tr}_1[(\rho \otimes \mathbb{I})(\mathbb{I} \otimes \mathbb{I} + S)] \tag{C.141e}$$

$$= \frac{\alpha_2}{2}(\mathbb{I} + \rho), \tag{C.141f}$$

where S is the SWAP operator. Hence, $\rho = \mathbb{E}_{\chi \in X'}\left[\frac{2}{\alpha_2}|\chi\rangle\langle\chi| - \mathbb{I}\right]$, and therefore for any observable $\mathcal{O}$,

$$\langle \mathcal{O} \rangle \coloneqq \operatorname{Tr}(\rho \mathcal{O}) = \mathop{\mathbb{E}}_{\chi \in X'} \operatorname{Tr}\left[\left(\frac{2}{\alpha_2}|\chi\rangle\langle\chi| - \mathbb{I}\right)\mathcal{O}\right]. \tag{C.142}$$

Suppose that we make N measurements. The output of the i^{th} measurement is a label χ_i. We store the classical shadow $\hat{\rho}_i \coloneqq \frac{2}{\alpha_2}|\chi_i\rangle\langle\chi_i| - \mathbb{I}$ on a classical computer. Therefore, after N measurements, we have a classical collection $\{\hat{\rho}_1, \ldots, \hat{\rho}_N\}$. Given sufficient information about our design and the observable, one can classically compute $\operatorname{Tr}(\hat{\rho}_i \mathcal{O})$. Define

$$\hat{o} \coloneqq \frac{1}{N} \sum_{i=1}^{N} \operatorname{Tr}(\hat{\rho}_i \mathcal{O}). \tag{C.143}$$

By construction, $\mathbb{E}[\hat{o}] = \langle \mathcal{O} \rangle$, where the expectation is taken over possible measurement outcomes. By Chebychev's inequality, $\Pr[|\hat{o} - \mathbb{E}[\hat{o}]| \geq \epsilon] \leq \operatorname{Var}(\hat{o})/\epsilon^2$, where

$$\operatorname{Var}(\hat{o}) = \sum_{i=1}^{N} \operatorname{Var}\left(\frac{1}{N} \operatorname{Tr}[\hat{\rho}_i \mathcal{O}]\right) \tag{C.144a}$$

$$= \frac{1}{N^2} \sum_{i=1}^{N} \operatorname{Var}(\operatorname{Tr}[\hat{\rho}_i \mathcal{O}]) \tag{C.144b}$$

$$= \frac{1}{N^2} \sum_{i=1}^{N} \mathop{\operatorname{Var}}_{\chi \in X'}\left(\operatorname{Tr}\left[\frac{2}{\alpha_2}|\chi\rangle\langle\chi|\,\mathcal{O}\right] - \operatorname{Tr}\mathcal{O}\right) \tag{C.144c}$$

$$= \frac{1}{N} \mathop{\operatorname{Var}}_{\chi \in X'}\left(\operatorname{Tr}\left[\frac{2}{\alpha_2}|\chi\rangle\langle\chi|\,\mathcal{O}\right] - \operatorname{Tr}\mathcal{O}\right) \tag{C.144d}$$

$$= \frac{1}{N} \mathop{\mathbb{E}}_{\chi \in X'}\left(\operatorname{Tr}\left[\frac{2}{\alpha_2}|\chi\rangle\langle\chi|\,\mathcal{O}\right] - \operatorname{Tr}\mathcal{O}\right)^2 - \frac{1}{N}\langle \mathcal{O} \rangle^2 \tag{C.144e}$$

$$= \frac{1}{N} \mathop{\mathbb{E}}_{\chi \in X'} \left(\operatorname{Tr}\left[\frac{2}{\alpha_2} |\chi\rangle\langle\chi| \, \mathcal{O} \right]^2 + (\operatorname{Tr} \mathcal{O})^2 - 2 \operatorname{Tr}\left[\frac{2}{\alpha_2} |\chi\rangle\langle\chi| \, \mathcal{O} \right] (\operatorname{Tr} \mathcal{O}) \right) - \frac{1}{N} \langle \mathcal{O} \rangle^2 \tag{C.144f}$$

$$\begin{aligned} &= \frac{4}{N\alpha_2^2} \int_X \operatorname{Tr}[|\chi\rangle\langle\chi| \, \mathcal{O}]^2 \operatorname{Tr}(\rho \, |\chi\rangle\langle\chi|) \, \mathrm{d}\mu(\chi) \\ &\quad + \frac{1}{N} (\operatorname{Tr} \mathcal{O})^2 \int_X \operatorname{Tr}(\rho \, |\chi\rangle\langle\chi|) \, \mathrm{d}\mu(\chi) \\ &\quad - \frac{4}{N\alpha_2} (\operatorname{Tr} \mathcal{O}) \int_X \operatorname{Tr}[|\chi\rangle\langle\chi| \, \mathcal{O}] \operatorname{Tr}(\rho \, |\chi\rangle\langle\chi|) \, \mathrm{d}\mu(\chi) \\ &\quad - \frac{1}{N} \langle \mathcal{O} \rangle^2 . \end{aligned} \tag{C.144g}$$

Using that X is a rigged 1-, 2-, and 3-design, we find

$$\operatorname{Var}(\hat{o}) = \frac{4\alpha_3}{N\alpha_2^2} \operatorname{Tr}[(\rho \otimes \mathcal{O} \otimes \mathcal{O}) \Pi_3] + \frac{1}{N} (\operatorname{Tr} \mathcal{O})^2 - \frac{4}{N\alpha_2} (\operatorname{Tr} \mathcal{O}) \operatorname{Tr}[(\rho \otimes \mathcal{O}) \Pi_2] - \frac{1}{N} \langle \mathcal{O} \rangle^2 \tag{C.145a}$$

$$= \frac{2\alpha_3}{3N\alpha_2^2} \left[(\operatorname{Tr} \mathcal{O})^2 + 2(\operatorname{Tr} \mathcal{O}) \langle \mathcal{O} \rangle + \operatorname{Tr} \mathcal{O}^2 + 2 \left\langle \mathcal{O}^2 \right\rangle \right] \tag{C.145b}$$

$$\begin{aligned} &\quad + \frac{1}{N} (\operatorname{Tr} \mathcal{O})^2 - \frac{2}{N\alpha_2} (\operatorname{Tr} \mathcal{O})(\operatorname{Tr} \mathcal{O} + \langle \mathcal{O} \rangle) - \frac{1}{N} \langle \mathcal{O} \rangle^2 \\ &\in \mathcal{O}\left(\frac{(\operatorname{Tr} |\mathcal{O}|)^2}{N} \right). \end{aligned} \tag{C.145c}$$

It then follows that

$$\Pr[|\hat{o} - \mathbb{E}[\hat{o}]| \geq \epsilon] \in \mathcal{O}\left(\frac{(\operatorname{Tr} |\mathcal{O}|)^2}{N\epsilon^2} \right). \tag{C.146}$$

Consider computing the expectation value of M observables $\mathcal{O}_1, \ldots, \mathcal{O}_M$ using the same N shadows $\hat{\rho}_i$ as above, and let $\hat{o}_i$ be the same as $\hat{o}$ from above but corresponding to $\mathcal{O}_i$. Then, applying the union bound, we find

$$\Pr\left[\max_i |\hat{o}_i - \mathbb{E}[\hat{o}_i]| \geq \epsilon \right] \in \mathcal{O}\left(\frac{M \max_i (\operatorname{Tr} |\mathcal{O}_i|)^2}{N\epsilon^2} \right). \tag{C.147}$$

Hence, to achieve a failure probability of at most δ, we need $N \in \mathcal{O}\left(\frac{M \max_i (\operatorname{Tr}|\mathcal{O}_i|)^2}{\delta\epsilon^2} \right)$.

Using median-of-means

We can do much better than this by using the median-of-means estimator as described in [50, Thm. 1] where we compute the median of K sample means and each mean is taken with N samples. Indeed, their theorem applies immediately, and we instead find that

$$N \in \mathcal{O}\left(\frac{1}{\epsilon^2} \max_i (\operatorname{Tr}|\mathcal{O}_i|)^2\right) \qquad \text{and} \qquad K \in \mathcal{O}(\log(M/\delta)) \tag{C.148}$$

suffices to estimate each $\langle \mathcal{O}_i \rangle$ to maximum additive error ϵ with success probability at least $1-\delta$. Thus, the total number of samples from ρ needed to accurately predict $\langle \mathcal{O}_1 \rangle, \ldots, \langle \mathcal{O}_M \rangle$ scales as $\log M$.

Unfortunately, we have not yet found a useful rigged 3-design (a rigged 3-design is described in Appendix C.4.5, but it involves infinite-dimensional integration). The 3-design condition was used to compute the variance $\operatorname{Var}(\hat{o})$.

One may wonder how well a rigged 2-design works for shadow tomography. Since the variance calculation requires three copies of $|\chi\rangle\langle\chi|$, the variance depends on the specific rigged 2-design that is used. Here we will compute the variance with respect to the rigged 2-design that uses the Kerred phase states; namely,

$$\frac{1}{2\pi+1} \sum_{n \in \mathbb{N}_0} (|n\rangle\langle n|)^{\otimes t} + \frac{2\pi}{2\pi+1} \int_{[0,2\pi]^2} ({}_\varphi|\theta\rangle\langle\theta|_\varphi)^{\otimes t} \frac{\mathrm{d}\theta\, \mathrm{d}\varphi}{2\pi} = \alpha_t \Pi_t \tag{C.149}$$

for $t \in \{1,2\}$, where $\alpha_1 = 1$ and $\alpha_2 = 1/(\pi + 1/2)$. The only term in the variance that is different is

$$\begin{aligned} &\int_X \operatorname{Tr}[|\chi\rangle\langle\chi|\,\mathcal{O}]^2 \operatorname{Tr}(\rho\,|\chi\rangle\langle\chi|)\,\mathrm{d}\mu(\chi) \\ &\qquad = \frac{1}{2\pi+1} \sum_{n \in \mathbb{N}_0} (\langle n|\,\mathcal{O}\,|n\rangle)^2 \langle n|\,\rho\,|n\rangle + \frac{1}{2\pi+1} \int_{[0,2\pi]^2} ({}_\varphi\langle\theta|\mathcal{O}\,|\theta\rangle_\varphi)^2\, {}_\varphi\langle\theta|\rho\,|\theta\rangle_\varphi \,\mathrm{d}\theta\,\mathrm{d}\varphi \end{aligned} \tag{C.150a}$$

$$= \frac{1}{2\pi+1}\sum_{n\in\mathbb{N}_0}\mathcal{O}_{n,n}^2\rho_{n,n} + \frac{1}{(2\pi+1)(2\pi)^3}\sum_{n_1,n_2,n_3\in\mathbb{N}_0}\sum_{m_1,m_2,m_3\in\mathbb{N}_0}\mathcal{O}_{n_1,m_1}\mathcal{O}_{n_2,m_2}\rho_{n_3,m_3}$$
$$\times\int_{[0,2\pi]^2}\mathrm{e}^{\mathrm{i}\theta(n_1+n_2+n_3-m_1-m_2-m_3)}\mathrm{e}^{\mathrm{i}\varphi(n_1^2+n_2^2+n_3^2-m_1^2-m_2^2-m_3^2)}\,\mathrm{d}\theta\,\mathrm{d}\varphi \tag{C.150b}$$

$$= \frac{1}{2\pi+1}\sum_{n\in\mathbb{N}_0}\mathcal{O}_{n,n}^2\rho_{n,n} + \frac{1}{(2\pi+1)2\pi}\sum_{n_1,n_2,n_3\in\mathbb{N}_0}\sum_{m_1,m_2,m_3\in\mathbb{N}_0}\mathcal{O}_{n_1,m_1}\mathcal{O}_{n_2,m_2}\rho_{n_3,m_3}$$
$$\times\,\delta_{n_1+n_2+n_3,m_1+m_2+m_3}\delta_{n_1^2+n_2^2+n_3^2,m_1^2+m_2^2+m_3^2}. \tag{C.150c}$$

Unfortunately, there is no obvious closed form simplification. We can however investigate specific cases. For example, consider the case when $\mathcal{O}$ is diagonal in the Fock basis $\{|n\rangle\}$. Then this term simply becomes $\frac{1}{2}\langle\mathcal{O}^2\rangle + \frac{1}{4\pi}(\operatorname{Tr}\mathcal{O})^2$. Hence, if we have a collection of M observables $\mathcal{O}_1,\ldots,\mathcal{O}_M$ that are each diagonal in the Fock state basis, then one needs only $\sim\log(M)\max_i(\operatorname{Tr}|\mathcal{O}_i|)^2$ measurements of ρ from the POVM defined by the rigged 2-design to estimate $\langle\mathcal{O}_1\rangle,\ldots,\langle\mathcal{O}_M\rangle$.

Perhaps a more interesting case is when $\mathcal{O}=|a\rangle\langle b|+|b\rangle\langle a|$ for positive integers a and b. Assume that $b>a$ and define $\Delta\coloneqq b-a>0$. In this case, the term above becomes

$$= \frac{1}{(2\pi+1)2\pi}\sum_{n_1,n_2,m_1,m_2\in\{a,b\}}\sum_{n_3,m_3\in\mathbb{N}_0}\mathcal{O}_{n_1,m_1}\mathcal{O}_{n_2,m_2}\rho_{n_3,m_3} \tag{C.151a}$$
$$\times\,\delta_{n_1+n_2+n_3,m_1+m_2+m_3}\delta_{n_1^2+n_2^2+n_3^2,m_1^2+m_2^2+m_3^2}$$
$$= \frac{1}{(2\pi+1)2\pi}\sum_{n_3,m_3\in\mathbb{N}_0}\big[\mathcal{O}_{a,b}^2\rho_{n_3,m_3}\delta_{2a+n_3,2b+m_3}\delta_{2a^2+n_3^2,2b^2+m_3^2} \tag{C.151b}$$
$$+\,\mathcal{O}_{b,a}^2\rho_{n_3,m_3}\delta_{2b+n_3,2a+m_3}\delta_{2b^2+n_3^2,2a^2+m_3^2}$$
$$+\,2\mathcal{O}_{a,b}\mathcal{O}_{b,a}\rho_{n_3,m_3}\delta_{n_3,m_3}\big]$$

$$= \frac{1}{(2\pi+1)2\pi}\Big[\sum_{n_3\in\mathbb{N}_0} \rho_{n_3,2(a-b)+n_3}\delta_{2a^2+n_3^2,2b^2+(2a+n_3-2b)^2}$$

$$+ \sum_{n_3\in\mathbb{N}_0} \rho_{n_3,2(b-a)+n_3}\delta_{2b^2+n_3^2,2a^2+(2b+n_3-2a)^2} \tag{C.151c}$$

$$+ 2\operatorname{Tr}\rho\Big]$$

$$= \frac{1}{\pi(2\pi+1)} + \frac{1}{(2\pi+1)2\pi}\Big[\sum_{n_3\geq 2\Delta} \rho_{n_3,n_3-2\Delta}\delta_{2a^2+n_3^2,2b^2+(n_3-2\Delta)^2}$$

$$+ \sum_{n_3\in\mathbb{N}_0} \rho_{n_3,n_3+2\Delta}\delta_{2b^2+n_3^2,2a^2+(n_3+2\Delta)^2}\Big] \tag{C.151d}$$

$$= \frac{1}{\pi(2\pi+1)} + \frac{1}{(2\pi+1)2\pi}\delta_{3b\geq a}\delta_{3a\geq b}\left[\rho_{(3b-a)/2,(3a-b)/2} + \rho_{(3a-b)/2,(3b-a)/2}\right] \tag{C.151e}$$

$$= \frac{1}{\pi(2\pi+1)}\left(1 + \delta_{3b\geq a}\delta_{3a\geq b}\operatorname{Re}(\rho_{(3b-a)/2,(3a-b)/2})\right) \tag{C.151f}$$

$$\leq \frac{2}{\pi(2\pi+1)}. \tag{C.151g}$$

Hence, if we have a collection of M observables of the form $\mathcal{O}_i = |a_i\rangle\langle b_i| + |b_i\rangle\langle a_i|$, then we can accurately determine each $\langle\mathcal{O}_i\rangle$ with only $\sim \log M$ measurements of ρ using the rigged 2-design.

Using Hoeffding's inequality

Again motivated by Ref. [52], we consider using Hoeffding's inequality and only using the 2-design property. Hence, this section applies to rigged 2-designs, of which we have constructed several. Specifically, suppose that we again consider estimating $\langle\mathcal{O}_j\rangle$ with N shadows by $\hat{o}_j = \frac{1}{N}\sum_{i=1}^N \operatorname{Tr}(\hat{\rho}_i\mathcal{O}_j)$. If $-\infty < c < \operatorname{Tr}(\hat{\rho}_i\mathcal{O}_j) < d < \infty$ almost surely for each shadow $\hat{\rho}_i$, then Hoeffding's inequality immediately implies that

$$\Pr[|\hat{o}_j - \mathbb{E}[\hat{o}_j]| \geq \epsilon] \leq 2\exp\left[-\frac{2N\epsilon^2}{(d-c)^2}\right]. \tag{C.152}$$

Then, applying the union bound,

$$\Pr\left[\max_j |\hat{o}_j - \mathbb{E}[\hat{o}_j]| \geq \epsilon\right] \leq 2M \exp\left[-\frac{2N\epsilon^2}{(d-c)^2}\right]. \tag{C.153}$$

Therefore, to achieve a failure probability of at most δ, we need

$$N \geq \log\left(\frac{2M}{\delta}\right) \frac{(d-c)^2}{2\epsilon^2} \tag{C.154}$$

to compute the M observables to additive accuracy ϵ.

For instance, we consider the example from above where the observables are $\mathcal{O}_j = |a_j\rangle\langle b_j| + |b_j\rangle\langle a_j|$ and we perform the shadows procedure with the rigged 2-design given in Eq. (C.149). One easily finds that $-2/\pi(\pi+1/2) \leq \mathrm{Tr}(\hat{\rho}_i \mathcal{O}_j) \leq 2/\pi(\pi+1/2)$. Hence, we can determine the expectation value of the M observables with error ϵ and failure probability at most δ with only $N \geq \log\left(\frac{2M}{\delta}\right) \frac{8}{\pi^2(\pi+1/2)^2\epsilon^2}$ measurements.

Worked example

We now work through a simple, explicit example of using shadow tomography with the rigged 2-design in Eq. (C.149) to determine the expectation value of M observables with $\log M$ measurements. We let each observable be $\mathcal{O}_j = |a_j\rangle\langle b_j| + |b_j\rangle\langle a_j| + |c_j\rangle\langle c_j|$ for arbitrary nonnegative integers a_j, b_j, c_j. Suppose that we have access to a blackbox quantum device that prepares ρ, but we know nothing else about it.

Generate shadows. The first step is to describe a procedure to generate a classical shadow. Recall that single-qubit "local Clifford" shadows [50] consist of choosing randomly between measuring in three different POVMs — the three Bloch-sphere axes — each yielding a binary outcome. In our case, for a single mode, we choose between measuring in either the discrete Fock-space POVM or a continuum of phase-state POVMs which differ by how much they have evolved under the Kerr Hamiltonian (quantified by φ). Each POVM has an

infinite number of outcomes: the Fock-state POVM admits a countable infinity of outcomes indexed by Fock-state occupation number n, while the phase-state POVMs have a compact continuous set of outcomes indexed by phase-state index θ.

From Eq. (C.149), we generate a shadow as follows. First we draw a random number x between 0 and 1. If $x \leq 1/(2\pi+1)$, then we measure ρ in the Fock state basis $\{|n\rangle \mid n \in \mathbb{N}_0\}$. The result will be an integer $n \in \mathbb{N}_0$ and the classical shadow is then a classical label representing the operator $\hat{\rho}^{(n)} \coloneqq (2\pi+1)\,|n\rangle\langle n| - \mathbb{I}$. If, on the other hand, $x > 1/(2\pi+1)$, then we draw a random number φ between 0 and 2π and measure ρ with the continuous POVM defined by the operators $\{{}_{\varphi}|\theta\rangle\,\langle\theta|_{\varphi} \mid \theta \in [0, 2\pi)\}$ and the measure $\mathrm{d}\theta$. The output of such a measurement is an angle $\theta \in [0, 2\pi)$ and the classical shadow is then a classical label representing the operator $\hat{\rho}^{(\theta,\varphi)} \coloneqq (2\pi+1)\,{}_{\varphi}|\theta\rangle\,\langle\theta|_{\varphi} - \mathbb{I}$.

Classically compute expectation values w.r.t. shadows. For the shadow $\hat{\rho}^{(n)}$, we easily see that

$$\mathrm{Tr}[\hat{\rho}^{(n)}\mathcal{O}_j] = \mathrm{Tr}[((2\pi+1)\,|n\rangle\langle n| - \mathbb{I})\mathcal{O}_j] = (2\pi+1)\delta_{n,c_j} - 1. \tag{C.155}$$

For the shadow $\hat{\rho}^{(\theta,\varphi)}$, we compute

$$\begin{aligned}
\mathrm{Tr}[\hat{\rho}^{(\theta,\varphi)}\mathcal{O}_j] &= \mathrm{Tr}\Big[((2\pi+1)\,{}_{\varphi}|\theta\rangle\langle\theta|_{\varphi} - \mathbb{I})\mathcal{O}_j\Big] && \text{(C.156a)}\\
&= \frac{2\pi+1}{2\pi}\sum_{n,m\in\mathbb{N}_0} \mathrm{e}^{\mathrm{i}\theta(n-m)+\mathrm{i}\varphi(n^2-m^2)}\,\langle m|\,\mathcal{O}_j\,|n\rangle - 1 && \text{(C.156b)}\\
&= \frac{2\pi+1}{2\pi}\left(\mathrm{e}^{\mathrm{i}\theta(a_j-b_j)+\mathrm{i}\varphi(a_j^2-b_j^2)} + \mathrm{e}^{\mathrm{i}\theta(b_j-a_j)+\mathrm{i}\varphi(b_j^2-a_j^2)} + 1\right) - 1 && \text{(C.156c)}\\
&= (2+1/\pi)\cos(\theta(a_j-b_j)+\varphi(a_j^2-b_j^2)) + \frac{1}{2\pi}. && \text{(C.156d)}
\end{aligned}$$

Choose the number of shadows to generate. We see that for every possible shadow $\hat{\rho}$ and every observable $\mathcal{O}_j$, $-2 - 1/\pi + 1/2\pi \leq \mathrm{Tr}[\hat{\rho}\mathcal{O}_j] \leq 2\pi$. Therefore, from Eq. (C.154),

we set

$$N = \left\lceil \log\left(\frac{2M}{\delta}\right) \frac{(2\pi + 2 + 1/\pi - 1/2\pi)^2}{2\epsilon^2} \right\rceil \approx \frac{36}{\epsilon^2} \log\left(\frac{2M}{\delta}\right). \tag{C.157}$$

Estimate expectation values w.r.t. state. With all this in place, we can now classically compute each $\langle \mathcal{O}_j \rangle$ to a maximum additive error of ϵ with success probability at least $1 - \delta$. First, generate N shadows with the procedure described above. Then, with those N shadows, classically compute the mean expectation value of each observable $\mathcal{O}_j$ over the N shadows using the expressions derived above for $\mathrm{Tr}[\hat{\rho}^{(n)}\mathcal{O}_j]$ and $\mathrm{Tr}[\hat{\rho}^{(\theta,\varphi)}\mathcal{O}_j]$. With probability at least $1 - \delta$, all of these M means will be within ϵ of the true expectation values with respect to ρ.

Subsection C.5.2: Fidelity calculations

In this subsection, we derive the calculations shown in Section 4.6.3. Throughout this subsection, we let $\mathcal{E}$ denote an R-regularized rigged 2-design, meaning that $\mathcal{E}$ is an ensemble over unit-normalized quantum states satisfying $\mathbb{E}_{\psi\in\mathcal{E}} \left(|\psi\rangle\langle\psi|\right)^{\otimes 2} = \frac{\Pi_2^{(R)}}{\mathrm{Tr}\,\Pi_2^{(R)}}$. We assume that R is positive semi-definite. Recall then that $\Pi_2^{(R)} = (R \otimes R)\Pi_2(R \otimes R)$, and $\Pi_2 = \frac{1}{2}(\mathbb{I} + S)$. Therefore, $2\,\mathrm{Tr}\,\Pi_2^{(R)} = (\mathrm{Tr}\,R^2)^2 + \mathrm{Tr}\,R^4$. From this characterization, one easily computes that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\psi\in\mathcal{E}} \langle\psi| A |\psi\rangle |\psi\rangle\langle\psi| &= \mathrm{Tr}_1\left[(A \otimes \mathbb{I}) \mathop{\mathbb{E}}_{\psi\in\mathcal{E}} |\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|\right] && \text{(C.158a)} \\
&= \frac{1}{2\,\mathrm{Tr}\,\Pi_2^{(R)}} \mathrm{Tr}_1\left[(A \otimes \mathbb{I})\left(R^2 \otimes R^2 + (R^2 \otimes R^2)S\right)\right] && \text{(C.158b)} \\
&= \frac{1}{2\,\mathrm{Tr}\,\Pi_2^{(R)}} \left[R^2\,\mathrm{Tr}(RAR) + R^2AR^2\right] && \text{(C.158c)} \\
&= \frac{R^2\,\mathrm{Tr}(RAR) + R^2AR^2}{(\mathrm{Tr}\,R^2)^2 + \mathrm{Tr}\,R^4}. && \text{(C.158d)}
\end{aligned}$$

Furthermore,

$$\mathop{\mathbb{E}}_{\psi\in\mathcal{E}} \langle\psi| A |\psi\rangle\langle\psi| B |\psi\rangle = \operatorname{Tr}\left[B \mathop{\mathbb{E}}_{\psi\in\mathcal{E}} \langle\psi| A |\psi\rangle |\psi\rangle\langle\psi|\right] \tag{C.159a}$$

$$= \operatorname{Tr} \frac{BR^2 \operatorname{Tr}(RAR) + BR^2AR^2}{(\operatorname{Tr} R^2)^2 + \operatorname{Tr} R^4} \tag{C.159b}$$

$$= \frac{\operatorname{Tr}(RBR) \operatorname{Tr}(RAR) + \operatorname{Tr}(RBR^2AR)}{(\operatorname{Tr} R^2)^2 + \operatorname{Tr} R^4}. \tag{C.159c}$$

We now study definitions of fidelity. We now assume that R is diagonal in the $\hat{n}$ basis. We define a continuous-variable version of a maximally-entangled state as [138]

$$|\phi_R\rangle \coloneqq \frac{1}{\sqrt{\operatorname{Tr} R}}(R^{1/4} \otimes R^{1/4}) \sum_{n=0}^{\infty} |n\rangle \otimes |n\rangle . \tag{C.160}$$

When $R = \mathrm{e}^{-\beta\hat{n}}$, $|\phi_R\rangle$ is a two-mode squeezed vacuum state; when $R = P_d$, $|\phi_R\rangle$ is a finite dimensional maximally entangled state. Define its reduced state on one mode by

$$\rho_R \coloneqq \operatorname{Tr}_2 |\phi_R\rangle\langle\phi_R| = R/\operatorname{Tr} R. \tag{C.161}$$

Let $\mathcal{D}$ be a quantum channel with Kraus operators K so that $\mathcal{D}(\rho) = \sum_K K\rho K^\dagger$. In analogy with the finite dimensional case, define the entanglement fidelity as

$$F_e^{(R)}(\mathcal{D}) \coloneqq \langle\phi_R| (\mathcal{I} \otimes \mathcal{D})(\phi_R) |\phi_R\rangle \tag{C.162a}$$

$$= \frac{1}{(\operatorname{Tr} R)^2} \sum_K \sum_{n,m,j,k} (\langle n| \otimes \langle n| R^{1/2})(\mathbb{I} \otimes K)(|m\rangle \otimes R^{1/2} |m\rangle) \times (\langle j| \otimes \langle j| R^{1/2})(\mathbb{I} \otimes K^\dagger)(|k\rangle \otimes R^{1/2} |k\rangle) \tag{C.162b}$$

$$= \frac{1}{(\operatorname{Tr} R)^2} \sum_K \sum_{n,j} \langle n| R^{1/2} K R^{1/2} |n\rangle\langle j| R^{1/2} K^\dagger R^{1/2} |j\rangle \tag{C.162c}$$

$$= \sum_K |\operatorname{Tr}(\rho_R K)|^2 . \tag{C.162d}$$

Furthermore, in analogy with the finite dimensional case, we define two "average fidelity" quantities,

$$\overline{F}_1^{(R)}(\mathcal{D}) \coloneqq \frac{\operatorname{Tr} R^4 + (\operatorname{Tr} R^2)^2}{\operatorname{Tr} R^2 + (\operatorname{Tr} R)^2} \mathop{\mathbb{E}}_{\psi\in\mathcal{E}} \langle\psi| R^+\mathcal{D}(\psi)R^+ |\psi\rangle\,, \tag{C.163a}$$

$$\overline{F}_2^{(R)}(\mathcal{D}) \coloneqq \mathop{\mathbb{E}}_{\psi\in\mathcal{E}} \langle\psi| \mathcal{D}(\psi) |\psi\rangle\,. \tag{C.163b}$$

We again emphasize that these definitions are independent of which R-regularized rigged 2-design $\mathcal{E}$ is used since they involve only two copies of $|\psi\rangle$ and two copies of $\langle\psi|$. Notice that when $R = P_d$, since the Moore-Penrose inverse R^+ of a projector is itself, we find that $\overline{F}_1^{(P_d)} = \overline{F}_2^{(P_d)}$.

By Eq. (C.159c), we immediately find that

$$\overline{F}_1^{(R)}(\mathcal{D}) = \frac{\operatorname{Tr} R^4 + (\operatorname{Tr} R^2)^2}{\operatorname{Tr} R^2 + (\operatorname{Tr} R)^2} \sum_K \mathop{\mathbb{E}}_{\psi\in\mathcal{E}} \langle\psi| R^+K |\psi\rangle\langle\psi| K^\dagger R^+ |\psi\rangle \tag{C.164a}$$

$$= \frac{\operatorname{Tr} R^4 + (\operatorname{Tr} R^2)^2}{\operatorname{Tr} R^2 + (\operatorname{Tr} R)^2} \sum_K \frac{\operatorname{Tr}(RK^\dagger R^+R)\operatorname{Tr}(RR^+KR) + \operatorname{Tr}(RK^\dagger R^+R^2R^+KR)}{(\operatorname{Tr} R^2)^2 + \operatorname{Tr} R^4} \tag{C.164b}$$

$$= \frac{\sum_K \frac{d_R}{(\operatorname{Tr} R)^2} \operatorname{Tr}(RK^\dagger R^+R)\operatorname{Tr}(RR^+KR) + \operatorname{Tr}(R^+R^2R^+\mathcal{D}(\rho_{R^2}))}{d_R + 1}, \tag{C.164c}$$

where we define an effective dimension $d_R \coloneqq (\operatorname{Tr} R)^2/\operatorname{Tr} R^2$. Since we are assuming R to be diagonal, $RR^+ = R^+R$. Furthermore, by definition of the Moore-Penrose inverse, $RR^+R = R$ and $R^+RR^+ = R^+$. Therefore,

$$\overline{F}_1^{(R)}(\mathcal{D}) = \frac{d_R \sum_K |\operatorname{Tr}(\rho_R K)|^2 + \operatorname{Tr}(RR^+\mathcal{D}(\rho_{R^2}))}{d_R + 1} \tag{C.165a}$$

$$= \frac{d_R F_e^{(R)}(\mathcal{D}) + \operatorname{Tr}(RR^+\mathcal{D}(\rho_{R^2}))}{d_R + 1}. \tag{C.165b}$$

We perform a similar calculation for $\overline{F}_2^{(R)}$,

$$\overline{F}_2^{(\sqrt{R})}(\mathcal{D}) = \sum_K \mathop{\mathbb{E}}_{\psi\in\mathcal{E}} \langle\psi| K |\psi\rangle\langle\psi| K^\dagger |\psi\rangle \tag{C.166a}$$

$$= \sum_K \frac{\operatorname{Tr}(R^{1/2}K^\dagger R^{1/2})\operatorname{Tr}(R^{1/2}KR^{1/2}) + \operatorname{Tr}(R^{1/2}K^\dagger RKR^{1/2})}{(\operatorname{Tr} R)^2 + \operatorname{Tr} R^2} \tag{C.166b}$$

$$= \frac{\sum_K |\operatorname{Tr}(RK)|^2 + \sum_K \operatorname{Tr}(KRK^\dagger R)}{(\operatorname{Tr} R)^2 + \operatorname{Tr} R^2} \tag{C.166c}$$

$$= \frac{d_R \sum_K |\operatorname{Tr}(\rho_R K)|^2 + d_R \sum_K \operatorname{Tr}(K\rho_R K^\dagger \rho_R)}{d_R + 1} \tag{C.166d}$$

$$= \frac{d_R F_e^{(R)}(\mathcal{D}) + d_R \operatorname{Tr}(\mathcal{D}(\rho_R)\rho_R)}{d_R + 1}. \tag{C.166e}$$

When $R = P_d$ is the projector and $\mathcal{D}$ is trace-preserving on the restricted d-dimensional subspace, both relations reduce to the finite dimensional relation. When R is invertible, such as the case when $R = \mathrm{e}^{-\beta\hat{n}}$, we find

$$\overline{F}_1^{(R)}(\mathcal{D}) = \frac{d_R F_e^{(R)}(\mathcal{D}) + 1}{d_R + 1}. \tag{C.167}$$

Loss channel

We now compute the various average fidelity quantities for the pure-loss channel $\mathcal{L}^\kappa$ defined in Section 4.6.3 and shown in Fig. 4.2. From [249, eq. 4.6], the Kraus operators for $\mathcal{L}^\kappa$ are

$$K_i = \sum_{m=0}^{\infty} \sqrt{\binom{m+i}{i}} (1-\kappa^2)^{i/2} \kappa^m |m\rangle\langle m+i| \tag{C.168}$$

for $i \in \mathbb{N}_0$.

We begin with $\overline{F}_{\mathrm{coh}}^{(\bar{n})}(\mathcal{L}^\kappa)$. Let $|\alpha\rangle$ be the coherent state specified by $\alpha \in \mathbb{C}$. Then, as calculated in [232],

$$\overline{F}_{\mathrm{coh}}^{(\bar{n})} = \frac{1}{\pi\bar{n}} \int_{\mathbb{C}} \mathrm{e}^{-|\alpha|^2/\bar{n}} \langle\alpha| \mathcal{L}^\kappa(|\alpha\rangle\langle\alpha|) |\alpha\rangle \, \mathrm{d}^2\alpha \tag{C.169a}$$

$$= \frac{1}{1+\bar{n}(1-\kappa)^2}. \tag{C.169b}$$

Next, we consider the entanglement fidelity $F_e^{(R)}(\mathcal{L}^\kappa)$. Let $|\phi_R\rangle = \frac{1}{\sqrt{\operatorname{Tr} R}}(R^{1/4}\otimes R^{1/4})\sum_{n=0}^{\infty}|n\rangle\otimes|n\rangle$ and assume that R is diagonal in the $|n\rangle$ basis. Then from Eq. (C.162d),

$$F_e^{(R)}(\mathcal{L}^\kappa) = \frac{1}{(\operatorname{Tr} R)^2}\sum_{i=0}^{\infty}\left|\operatorname{Tr}(RK_i)\right|^2. \tag{C.170}$$

When $R = R_\beta = \mathrm{e}^{-\beta\hat{n}}$, one easily finds this to be $\frac{(\mathrm{e}^\beta-1)^2}{(\mathrm{e}^\beta-\kappa)^2}$. Recall that in Section 4.6.3 we required that $d_{R_\beta} \coloneqq (\operatorname{Tr} R_\beta)^2/\operatorname{Tr} R_\beta^2 = 1+2\bar{n}$. Solving for β, we find that $\mathrm{e}^\beta = 1+1/\bar{n}$, and therefore

$$F_e^{(R_\beta)}(\mathcal{L}^\kappa) = (1+\bar{n}(1-\kappa))^{-2}. \tag{C.171}$$

On the other hand, when $R = P_d = \sum_{n=0}^{d-1}|n\rangle\langle n|$, we find

$$F_e^{(P_d)}(\mathcal{L}^\kappa) = \frac{\left(1-\kappa^d\right)^2}{(1-\kappa)^2 d^2}. \tag{C.172}$$

From Eq. (C.167), $\overline{F}_1^{(R_\beta)}(\mathcal{L}^\kappa)$ is the same as $F_e^{(R_\beta)}(\mathcal{L}^\kappa)$ up to an offset. However, $\overline{F}_2^{(R_\beta)}(\mathcal{L}^\kappa)$ is not as simple. Indeed, from Eq. (C.166e), we must compute

$$\operatorname{Tr}[\mathcal{L}^\kappa(\rho_{R_\beta})\rho_{R_\beta}] = \frac{1}{(\operatorname{Tr} R_\beta)^2}\sum_i \operatorname{Tr}(K_i R_\beta K_i^\dagger R_\beta) \tag{C.173a}$$

$$= (1=\mathrm{e}^{-\beta})^2\sum_i\sum_{a,b}\mathrm{e}^{-\beta(a+b)}\langle a|K_i|b\rangle\langle b|K_i^\dagger|a\rangle \tag{C.173b}$$

$$= (1-\mathrm{e}^{-\beta})^2\sum_{a\le b}\mathrm{e}^{-\beta(a+b)}\binom{b}{b-a}(1-\kappa^2)^{b-a}\kappa^{2a} \tag{C.173c}$$

$$= \frac{e^\beta-1}{e^\beta+\kappa^2}. \tag{C.173d}$$

Therefore, from Eq. (C.166e),

$$\overline{F}_2^{(R_\beta)}(\mathcal{L}^\kappa) = \frac{d_{R_{2\beta}}}{d_{R_{2\beta}}+1}\left[d_{R_{2\beta}}F_e^{(R_{2\beta})}(\mathcal{L}^\kappa) + \frac{e^\beta - 1}{e^\beta + \kappa^2}\right] \tag{C.174a}$$

$$= \frac{1}{d_{R_{2\beta}}+1}\left[\frac{(\mathrm{e}^{2\beta}-1)^2}{(\mathrm{e}^{2\beta}-\kappa)^2} + \frac{e^{2\beta}-1}{e^{2\beta}+\kappa^2}\right] \tag{C.174b}$$

$$= \frac{1}{\tanh\beta + 1}\left[\frac{(\mathrm{e}^{2\beta}-1)^2}{(\mathrm{e}^{2\beta}-\kappa)^2} + \frac{e^{2\beta}-1}{e^{2\beta}+\kappa^2}\right]. \tag{C.174c}$$

Requiring that $d_{R_\beta} = 1 + 2\bar{n}$ yields $\mathrm{e}^\beta = 1 + 1/\bar{n}$, giving

$$\overline{F}_2^{(R_{\beta/2})}(\mathcal{L}^\kappa) = \frac{(2\bar{n}+1)\left((1-\kappa)^2\bar{n}+2\right)}{2((1-\kappa)\bar{n}+1)^2\left((\kappa^2+1)\,\bar{n}+1\right)}. \tag{C.175}$$

Finally, we compute $\overline{F}_{1,2}^{(P_d)}(\mathcal{L}^\kappa) \equiv \overline{F}_1^{(P_d)}(\mathcal{L}^\kappa) = \overline{F}_2^{(P_d)}(\mathcal{L}^\kappa)$. From Eq. (C.165b), it only remains to compute $\mathrm{Tr}(\mathcal{L}^\kappa(P_d/d)P_d)$, which is

$$\mathrm{Tr}(\mathcal{L}^\kappa(P_d/d)P_d) = \frac{1}{d}\sum_{i=0}^{\infty}\sum_{a,b=0}^{d-1}\langle a|\,K_i\,|b\rangle\langle b|\,K_i^\dagger\,|a\rangle \tag{C.176a}$$

$$= \frac{1}{d}\sum_{a\le b=0}^{d-1}\binom{b}{b-a}(1-\kappa^2)^{b-a}\kappa^{2a} \tag{C.176b}$$

$$= 1. \tag{C.176c}$$

Indeed, this is 1 since $\mathcal{L}^\kappa$ does not take a state that is defined in the subspace P_d out of that subspace. Using $d = 1 + \bar{n}$, we find

$$F_{1,2}^{(P_{\bar{n}+1})}(\mathcal{L}^\kappa) = \frac{\left(1-\kappa^{\bar{n}+1}\right)^2}{(1-\kappa)^2(\bar{n}+1)(\bar{n}+2)} + \frac{1}{\bar{n}+2}. \tag{C.177}$$

The plots of all of these fidelities as functions of κ are shown in Fig. 4.2.

Section C.6: Torus designs, trigonometric cubature, and mutually unbiased bases

In this section, we prove the equivalence between our definition of a torus t-design (c.f. Definition C.3.6) and the definition given in Ref. [189]. In Ref. [189], Kuperberg defines a general notion of torus cubature that generalizes the more established theory of trigonometric cubature [269, 303]. We will be interested in one of the cases of his definition; namely, our definition of a T^{n+1} t-design is equivalent to his definition of a positive degree t cubature rule on $T(\mathrm{PSU}(n+1))$. After showing this, we will compare a torus design to the more standard trigonometric cubature rules and find that a torus t-design lies somewhere between a degree t and degree $2t$ positive trigonometric cubature rule. Finally, we prove a relationship between torus 2-designs and complete sets of mutually unbiased bases.

Subsection C.6.1: Equivalence to Kuperberg's definition

We begin by describing Kuperberg's definition. Consider a group $\mathcal{T}$ that is isomorphic to the torus $\mathcal{T} \cong T^n = (S^1)^n$. Suppose $\rho\colon \mathcal{T} \to \mathrm{GL}(V)$ is a free linear representation, with V a real vector space $V \cong \mathbb{R}^N$. Since ρ is free it follows that there is one or many faithful orbits $\mathcal{O}$. Suppose $u \in V$ such that $\mathcal{O} = \{\rho(g)u \mid g \in \mathcal{T}\}$ is a faithful orbit. Since $\mathcal{O}$ is faithful, $\mathcal{T}$ can be identified with $\mathcal{O}$ via $\mathcal{T} \ni g \leftrightarrow \rho(g)u \in \mathcal{O}$. With this identification, $\mathcal{T}$ inherits an algebraic structure, since it is well-defined to consider addition such as $\rho(g)u + \rho(h)u \in V$ for every $g, h \in \mathcal{T}$. With this structure, along with the unit normalized Haar measure on $\mathcal{T}$ (since $\mathcal{T}$ is compact), we can define cubature on $\mathcal{T}$ as follows. A set $S \subset \mathcal{T}$ and weight function $v\colon S \to \mathbb{R}_{>0}$ is a (positive) cubature rule of degree t on $\mathcal{T}$ if

$$\sum_{h \in S} v(h) f(\rho(h)u) = \int_{\mathcal{T}} f(\rho(g)u)\, \mathrm{d}\mu(g) \tag{C.178}$$

for any polynomial $f\colon V \to \mathbb{R}$ of degree t or less. Since $V \cong \mathbb{R}^N$, we can therefore view f as a being a function of the entries of the vectors $\rho(g)u$. Kuperberg also states that this definition is independent of u for generic choices of u as long as the resulting $\mathcal{O}$ is faithful [2].

We now apply this to $\mathcal{T} = T(\mathrm{PSU}(n+1))$. $\mathrm{PSU}(n+1)$ is the projective special unitary group of $(n+1)\times(n+1)$ matrices defined by $\mathrm{SU}(n+1)/\mathrm{U}(1)$. Then $\mathcal{T} = T(\mathrm{PSU}(n+1))$ is a maximal torus (maximal, compact, connected, abelian Lie subgroup) of $\mathrm{PSU}(n+1)$, which is the group of diagonal unitary matrices with determinant 1 modulo the center of $\mathrm{SU}(n+1)$ (i.e. modulo global phases). For a unitary $U \in \mathcal{T}$, let U_{ij} denote the entry in the i^{th} row and j^{th} column. The determinant condition implies that $U_{n+1,n+1}$ is uniquely determined by U_{ii} for $i = 1, \ldots, n$. We therefore see that $\mathcal{T} \cong T^n$. We can also take an alternative view of $\mathcal{T}$; we can view $\mathcal{T}$ as the group of diagonal unitary matrices with $U_{n+1,n+1} = 1$ modulo the center. This is the view we will take. Below we will consider the adjoint action of this group, and therefore we do not have to worry about modding out the center; ultimately, we will just end up integrating out global phases.

We let $N = 2(n+1)^2$ and identify V with the vector space of $(n+1)\times(n+1)$ complex matrices $\mathbb{C}^{(n+1)\times(n+1)}$. We consider a linear action of $\mathcal{T}$ defined by conjugation on V; in other words, $\rho(g)$ is defined by $A \mapsto gAg^\dagger$. As mentioned, we can pick any $u \in V$ as a base point as long as the resulting orbit,

$$\mathcal{O} = \left\{ \begin{pmatrix} e^{i\phi_1} & & & \\ & \ddots & & \\ & & e^{i\phi_n} & \\ & & & 1 \end{pmatrix} u \begin{pmatrix} e^{-i\phi_1} & & & \\ & \ddots & & \\ & & e^{-i\phi_n} & \\ & & & 1 \end{pmatrix} \mid \phi_1, \ldots, \phi_n \in [-\pi, \pi) \right\}, \quad \text{(C.179)}$$

[2] Personal communication with Greg Kuperberg.

is faithful. We pick u to be the matrix of all 1's, $u = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$. One can then easily check that

$$\mathcal{O} = \left\{ v \text{ where } v_{ij} = \mathrm{e}^{\mathrm{i}(\phi_i - \phi_j)} \mid \phi_1, \dots, \phi_n \in [-\pi, \pi), \phi_{n+1} = 0 \right\}. \tag{C.180}$$

A degree t positive cubature rule on $\mathcal{T}$ is a set $S \subset \mathcal{T}$ and weight function $v \colon S \to \mathbb{R}_{>0}$ that satisfies

$$\sum_{h \in S} v(h) f(huh^\dagger) = \int_{[-\pi,\pi)^n} f(gug^\dagger) \frac{\mathrm{d}\phi_1 \dots \mathrm{d}\phi_n}{(2\pi)^n}. \tag{C.181}$$

f is a polynomial of degree at most t in the entries. By linearity, we can consider f to be a monomial. From $\mathcal{O}$, we consider monomials of degree $\leq t$ in the variables

$$\left\{ \mathrm{e}^{\mathrm{i}(\phi_i - \phi_j)} \mid i, j = 1, \dots, n+1 \right\} \tag{C.182}$$

where recall that $\phi_{n+1} = 0$. It follows that an equivalent definition of a degree t positive cubature rule on $\mathcal{T}$ is as follows. Let $S \subset [-\pi, \pi)^n$ and $v \colon S \to \mathbb{R}_{>0}$. Hence, for each $\phi \in S$, $\phi_i \in [-\pi, \pi)$, and we define $\phi_{n+1} = 0$. Then (S, v) must satisfy

$$\forall j_1, \dots, j_t, k_1, \dots, k_t \in \{1, \dots, n+1\} : \ \sum_{\theta \in S} v(\theta) \mathrm{e}^{\mathrm{i}(\theta_{j_1} + \dots + \theta_{j_t} - \theta_{k_1} - \dots - \theta_{k_t})} = \int_{[-\pi,\pi)^n} \mathrm{e}^{\mathrm{i}(\phi_{j_1} + \dots + \phi_{j_t} - \phi_{k_1} - \dots - \phi_{k_t})} \frac{\mathrm{d}\phi_1}{(} \tag{C.183}$$

Notice that this takes care of *all* monimials of degree t or less. For example, consider the monomial $\mathrm{e}^{\mathrm{i}\phi_1}$. This is taken care of by setting $j_1 = 1$ and $j_2 = \dots = j_t = k_1 = \dots = k_t = n+1$.

We easily see that the right hand side (i.e. the integral) does not change if we integrate over ϕ_{n+1} instead of just fixing it to be 0. Similarly, on the left hand side, for every $\theta \in S$, we can shift each θ_i by a constant $\theta_i \mapsto \theta_i + c$ without changing anything. Therefore, we

can remove the definition that $\theta_{n+1} = 0$, and instead allow θ_{n+1} to be arbitrary. Thus, we arrive at an equivalent definition of a degree t positive cubature rule on $\mathcal{T}$ as follows. Let $S \subset [-\pi, \pi)^{n+1}$ and $v \colon S \to \mathbb{R}_{>0}$. Then (S, v) must satisfy

$$\sum_{\theta \in S} v(\theta) \prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\theta_{j_i} - \theta_{k_i})} = \int_{T^m} \prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\phi_{j_i} - \phi_{k_i})} \frac{\mathrm{d}\phi_1 \dots \mathrm{d}\phi_{n+1}}{(2\pi)^{n+1}}. \tag{C.184}$$

Notice that this is exactly our definition of an $(n+1)$-torus t-design per Definition C.3.6.

In conclusion, our definition of a T^{n+1} t-design is equivalent to Kuperberg's definition of a degree t positive cubature rule on $T(\mathrm{PSU}(n+1))$.

Subsection C.6.2: Comparison to standard trigonometric cubature

A degree t positive trigonometric cubature rule (S, v) on T^n must satisfy

$$\sum_{\theta \in S} v(\theta) \prod_{i=1}^{n} \mathrm{e}^{\mathrm{i}\alpha_i \theta_i} = \int_{T^n} \prod_{i=1}^{n} \mathrm{e}^{\mathrm{i}\alpha_i \phi_i} \, \mathrm{d}\phi \tag{C.185}$$

whenever $\sum_{i=1}^{n} |\alpha_i| \leq t$. We see that our T^n t-designs lie somewhere between a degree t and degree $2t$ trigonometric cubature rule. To see the former, we show that a torus t-design must also be a degree t trigonometric cubature rule. From the definition of torus designs,

$$\sum_{\theta \in S} v(\theta) \prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\theta_{a_i} - \theta_{b_i})} = \int_{T^n} \prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\phi_{a_i} - \phi_{b_i})} \, \mathrm{d}\phi. \tag{C.186}$$

Suppose we consider a monimial $\prod_{i=1}^{n} \mathrm{e}^{\mathrm{i}\alpha_i \theta_i}$. If $\sum_i |\alpha_i| \leq t$, then we can generate the monomial via a choice of a_i and b_i. Indeed, recall that without loss of generality we can assume that $\theta_n = 0$. Consider as an example $t = 3$, $n = 4$, and the task of generating the monoimal defined by $\alpha = (2, 0, -1, 0)$. Then we set $a_1 = 1$, $a_2 = 1$, $a_3 = 4$, $b_1 = 3$, and $b_2 = b_3 = 4$. Then $\prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\theta_{a_i} - \theta_{b_i})} = \mathrm{e}^{\mathrm{i}(\theta_1 + \theta_1 + \theta_4 - \theta_3 - \theta_4 - \theta_4)}$, which is exactly $\prod_i \mathrm{e}^{\mathrm{i}\alpha_i \theta_i}$ since $\theta_4 = 0$. Hence, by a proper choice of a_i and b_i, any monomial of degree t or less can be generated by

$\prod_{i=1}^{t} \mathrm{e}^{\mathrm{i}(\theta_{a_i} - \theta_{b_i})}$, and hence a torus t-design is also a degree t trigonometric cubature rule.

On the contrary, if we allow $\sum_i |\alpha_i| > t$, we find that there are some monomials that cannot be generated by a sufficient choice of a_i and b_i. So even though a torus t-design involves monomials of degree up to $2t$, it is not in general a degree $2t$ trigonometric cubature rule. However, since torus t-designs only involve certain monomials up to degree $2t$, a trigonometric cubature rule of degree $\geq 2t$ is a torus t-design.

Subsection C.6.3: Relation to MUBs

We begin by recalling the definition of a complete set of mutually unbiased bases (MUBs) [238].

Definition C.6.1 (Complete set of MUBs)**.** *Suppose that $B_0, \ldots, B_n$ are each orthonormal bases of $\mathbb{C}^n$. B_i and B_j are called* mutually unbiased *if*

$$\forall\, |\psi\rangle \in B_i, |\phi\rangle \in B_j : \quad |\langle\psi|\phi\rangle|^2 = 1/n. \tag{C.187}$$

The collection $B_0, \ldots, B_n$ is called a complete set of MUBs *if the bases are pairwise mutually unbiased. This can be equivalently stated in term of the phases $\theta^i_{j,k}$ involved in the bases (see below);*

1. *(Orthonormality).* $\forall i, j, k \in \{0, \ldots, n-1\} : \ \frac{1}{n} \sum_{l=0}^{n-1} \mathrm{e}^{\mathrm{i}(\theta^i_{j,l} - \theta^i_{k,l})} = \delta_{jk}$;
2. *(Mutually unbiasedness).* $\forall i \neq j, k, m \in \{0, \ldots, n-1\} : \ \left| \sum_{l=0}^{n-1} \mathrm{e}^{\mathrm{i}(\theta^i_{k,l} - \theta^j_{m,l})} \right|^2 = n.$

We now show the relationship between complete sets of MUBs and torus 2-designs. Recall the matrix Π_2 that, for any orthonormal basis $\{|0\rangle, \ldots, |n-1\rangle\}$, has matrix elements

$$\Pi_2(a, b; c, d) \coloneqq \langle a| \otimes \langle b|\, \Pi_2\, |c\rangle \otimes |d\rangle = \frac{1}{2}(\delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc}). \tag{C.188}$$

By simply doing the integration, one finds that S is an equal weight (i.e. $v(\theta) = 1/|S|$) T^n 2-design if and only if

$$\frac{1}{|S|}\sum_{\theta\in S} \mathrm{e}^{\mathrm{i}(\theta_a+\theta_b-\theta_c-\theta_d)} = 2\Pi_2(a,b;c,d) - \delta_{ab}\delta_{ac}\delta_{ad}. \tag{C.189}$$

Here we show a connection between equal weight torus 2-designs and complete sets of MUBs.

Lemma C.6.2. *The phases of a complete set of MUBs on $\mathbb{C}^n$ form an equal weighted n-torus 2-design of size n^2.*

Proof. Without loss of generality, we can assume that one of the bases is the computational basis. So assume that $B_0 = \{|0\rangle, \ldots, |n-1\rangle\}$. Then in order for $|\langle\psi|j\rangle| = 1/\sqrt{n}$ for each $j \in \{0,\ldots,n-1\}$, $|\psi\rangle \in B_i$, and $i \in \{1,\ldots,n\}$, it must be that each other basis B_i must only involve uniform superposition states over the computational basis. With this in mind, define $|\psi_j^i\rangle$ so that

$$B_i = \{|\psi_j^i\rangle \mid j \in \{1,\ldots,n\}\}. \tag{C.190}$$

Define $\theta_{j,k}^i$ so that

$$|\psi_j^i\rangle = \frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} \mathrm{e}^{\mathrm{i}\theta_{j,k}^i}|k\rangle. \tag{C.191}$$

From [41], we know that the complete set of MUBs forms a complex-projective 2-design. Therefore, $D = \{|0\rangle, \ldots, |n-1\rangle\} \cup \{|\psi_j^i\rangle \mid i,j \in \{1,\ldots,n\}\}$ is a complex-projective 2-design. We therefore find that

$$\frac{1}{n(n+1)}\left(\sum_{k=0}^{n-1}(|k\rangle\langle k|)^{\otimes 2} + \sum_{i,j=1}^{n}(|\psi_j^i\rangle\langle\psi_j^i|)^{\otimes 2}\right) = \frac{2}{n(n+1)}\Pi_2. \tag{C.192}$$

Let $a,b,c,d \in \{0,\ldots,n-1\}$. Applying $\langle a|\otimes\langle b|$ on the left hand side and $|c\rangle\otimes|d\rangle$ on the

right hand side, we find

$$\delta_{ab}\delta_{ac}\delta_{ad} + \frac{1}{n^2}\sum_{i,j=1}^{n} \mathrm{e}^{\mathrm{i}(\theta^i_{j,a}+\theta^i_{j,b}-\theta^i_{j,c}-\theta^i_{j,d})} = 2\Pi_2(a,b;c,d). \tag{C.193}$$

Per the definition of an n-torus 2-design from above, we see that the angles $\theta^i_{j,k}$ form an n-torus 2-design with size n^2. □

Furthermore, Theorem 3.3 from Ref. [312] states that if $B_0, \ldots, B_m$ are each orthonormal bases of $\mathbb{C}^n$ and $\bigcup_i B_i$ is an unweighted complex-projective 2-design, then $m = n$ only if the bases are mutually unbiased. The following lemma therefore follows.

Lemma C.6.3. *If an equal weighted n-torus 2-design exists such that the phases in the design define n orthonormal bases, then there exists a complete set of MUBs in $\mathbb{C}^n$.*

Therefore, we have an if and only if.

Corollary C.6.4. *There exists a complete set of MUBs in $\mathbb{C}^n$ if and only if there exists an equal weighted n-torus 2-design such that the phases in the design define n orthonormal bases. Concretely, there exists a complete set of MUBs in $\mathbb{C}^n$ if and only there exists angles $\theta^i_{j,k}$ such that*

1. $\forall i,j,k \in \{0,\ldots,n-1\}: \ \frac{1}{n}\sum_{l=0}^{n-1} \mathrm{e}^{\mathrm{i}(\theta^i_{j,l}-\theta^i_{k,l})} = \delta_{jk}$;
2. $\forall a,b,c,d \in \{0,\ldots,n-1\}: \ \frac{1}{n^2}\sum_{i,j=0}^{n-1} \mathrm{e}^{\mathrm{i}(\theta^i_{j,a}+\theta^i_{j,b}-\theta^i_{j,c}-\theta^i_{j,d})} = \begin{cases} 1 & \text{if } (a=c \text{ and } b=d) \text{ or } (a=d \text{ and } b \\ 0 & \text{otherwise} \end{cases}$

In summary, the definition of a complete set of MUBs has two conditions: orthonormality and mutually unbiasedness. We have shown that the mutually unbiased condition can be replaced with the condition that the phases must form a torus 2-design of size exactly n^2.

Appendix D: Technical details for Chapter 6

Section D.1: Singer sets

In this appendix, we review Singer's construction of Sidon sets of size p^m+1 for cyclic groups of size $(p^m)^2+(p^m)+1$ with p a prime [89, p. 380-381] [270, Sec. 3.5] [271]. The existence of these Singer sets implies that there is a $P(T^n)$ 2-design of size $(n-1)^2+n = n^2-n+1$, i.e., a minimal one, whenever $n-1$ is prime-power. More generally, we review Singer's construction of $B_t \pmod{m \coloneqq \frac{(n-1)^{t+1}-1}{n-2}}$ sets (*cf.* Lemma D.1.2), which yield $P(T^n)$ t-designs of size m whenever $n-1$ is a prime power for any t. We emphasize that everything in this appendix is review. We also provide code for constructing Singer sets [278].

Let θ be the generator of $\mathbb{F}^{\times}_{(n-1)^{t+1}}$, and then let

$$T_t := \{0\} \cup \{a \in [(n-1)^{t+1}-1] \mid (\theta^a - \theta) \in \mathbb{F}_{n-1} \subset \mathbb{F}_{(n-1)^{t+1}}\}. \tag{D.1}$$

The inclusion $\mathbb{F}_{n-1} \hookrightarrow \mathbb{F}_{(n-1)^{t+1}}$ is done by identifying the generator of $\mathbb{F}^{\times}_{(n-1)}$ with $\theta^{\frac{(n-1)^{t+1}-1}{n-2}}$, which makes sense as for any finite field $\mathbb{F}_q$, $|\mathbb{F}_q^{\times}| = q-1$, and $\mathbb{F}_q^{\times}$ is cyclic.

Further, note that $\mathbb{F}_{(n-1)^{t+1}}$ is a $(t+1)$-dimensional $\mathbb{F}_{n-1}$-vector space. Thus, $\{\theta^b\}_{b=0}^{t}$ is a $\mathbb{F}_{n-1}$-basis of $\mathbb{F}_{(n-1)^{t+1}}$. This means that all $\theta^a = \sum_{i=0}^{t} k_i\theta^i$ for some unique $k_i \in \mathbb{F}_{(n-1)}$. However, if $\frac{(n-1)^{t+1}-1}{n-2} | a$, we know all $i \geq 1$ have $k_i = 0$.

Then, let

$$S_t((n-1),\theta) := \left\{ l \in \mathbb{Z}_{\frac{(n-1)^{t+1}-1}{n-2}} \mid l \equiv a \bmod \left(\frac{(n-1)^{t+1}-1}{n-2}\right), a \in T_t \right\} \tag{D.2}$$

be the residues of T_t mod $\frac{(n-1)^{t+1}-1}{n-2}$. We now recount proofs of some of $S_t((n-1),\theta)$'s properties.

Lemma D.1.1. $|S_t((n-1),\theta)| = n$.

Proof. First we note there are n distinct elements of $\mathbb{F}_{(n-1)^{t+1}}$ of the form $\theta + \gamma_a$, $\gamma_a \in \mathbb{F}_{n-1}$ by the $\mathbb{F}_{n-1}$-linear independence of θ and 1. As all elements of $\mathbb{F}_{(n-1)^{t+1}}$ equal θ^a for some unique $a \in [(n-1)^{t+1}-1]$, we see that $|T_t| = n$. Now, we must show that every element of T_t has a different residue modulo $\frac{(n-1)^{t+1}-1}{n-2}$.

Suppose $a, a' := a + k\frac{(n-1)^{t+1}-1}{n-2} \in T$, $k \in \mathbb{Z}_{>0}$. Then $r := \theta^{a'}/\theta^a = \theta^{k\frac{(n-1)^{t+1}-1}{n-2}} \in \mathbb{F}_{n-1}$. But by definition of T_t, $\theta^a = \theta + \gamma_a$, $\theta^{a'} = \theta + \gamma_{a'}$. But

$$\theta^{a'} = r\theta^a = r\theta + r\gamma_a. \tag{D.3}$$

Thus, $r = 1$, meaning $(n-2)|k$, which means that only a can be in $[(n-1)^{t+1}-1]$, and thus that no two elements of T_t can have the same residue modulo $\frac{(n-1)^{t-1}+1}{n-2}$. □

Lemma D.1.2. *$S_t((n-1),\theta)$ is a B_t (mod $\frac{(n-1)^{t+1}-1}{n-2}$) set.*

Proof. Recall that $\{\theta^i\}_{i=0}^t$ is a $\mathbb{F}_{n-1}$-basis of $\mathbb{F}_{(n-1)^{t+1}}$. In other words, there exist no non-elementwise-zero tuples $(c_i)_{i=0}^t \in \mathbb{F}_{n-1}^{t+1}$ such that

$$\sum_{i=0}^{t} c_i\theta^i = 0. \tag{D.4}$$

Equivalently, θ cannot be the root of any polynomial of degree $\leq t$ with $\mathbb{F}_{n-1}$-coefficients.

Now, consider two multisets A, B, $|A| = |B| \leq t$, taking entries from $S_t((n-1),\theta)$. Then, by the definition of $S_t((n-1),\theta)$ and T_t, we see that for all $a \in A \cup B$

$$\theta^a = \alpha_a(\theta + \gamma_a) \tag{D.5}$$

for some $\alpha_a \in \mathbb{F}_{n-1}$. Now, consider $\Pi_A := \prod_{a\in A} \theta^a$ and $\Pi_B := \prod_{b\in B} \theta^b$. It is clear that $\Pi_B/\Pi_A \in \mathbb{F}_{n-1}$ and only if

$$\sum_{a\in A} a \equiv \sum_{b\in B} b \pmod{\frac{(n-1)^{t+1}-1}{n-2}}. \tag{D.6}$$

Thus, $\Pi_A - \beta_{A,B}\Pi_B = 0$ for some $\beta_{A,B} \in \mathbb{F}_{n-1}$ if and only if Eq. (D.6) holds. However, for any $\beta \in \mathbb{F}_{n-1}$, we see that $\Pi_A - \beta\Pi_B$ is a degree-t polynomial equation in θ with $\mathbb{F}_{n-1}$ coefficients, meaning it cannot have any solutions, meaning the $B_t \pmod{\frac{(n-1)^{t+1}-1}{n-2}}$ condition is satisfied. □

Subsection D.1.1: Explicit example of dense modular Sidon set

In this appendix, we work through an explicit example of the construction of the Sidon set $S_{t=2}((n-1),\theta)$ for $n = 5 = 2^2 + 1$. We begin by constructing T_t. Consider the field $\mathbb{F}_{(n-1)^{t+1}} = \mathbb{F}_{4^3} = \mathbb{F}_{2^6}$. With the irreducible polynomial $f(x) = 1 + x^5 + x^6 \in \mathbb{F}_2[x]$, we work in the polynomial representation $\mathbb{F}_{2^6} \cong \mathbb{F}_2[x]/(f(x))$.

One can check that the generator θ of the multiplicative group $\mathbb{F}_{2^6}^\times$ is x in this representation—in other words, $|\{x^m \bmod f(x) \mid m \in \mathbb{Z}_{63}\}| = 63$. We identify $\mathbb{F}_{n-1} = \mathbb{F}_{2^2} \subset \mathbb{F}_{2^6}$ via generating $\mathbb{F}_{2^2}^\times$ with

$$y = x^{\frac{(n-1)^{t+1}-1}{n-2}} = x^{21}, \tag{D.7}$$

so that $\mathbb{F}_{2^2} = \{0\} \cup \left\{y^k \mid k \in \mathbb{Z}_3\right\}$. Then

$$T_{t=2} = \{0\} \cup \{a \in \mathbb{Z}_{4^3-1} \setminus \{0\} \mid (x^a - x) \pmod{f(x)} \in \mathbb{F}_{2^2}\}. \tag{D.8}$$

Clearly, $1 \in T_{t=2}$. With that out of the way, we can rephrase this as

$$T_{t=2} = \{0,1\} \cup \left\{a \in \mathbb{Z}_{4^3-1} \setminus \{0,1\} \mid \exists k \in \mathbb{Z}_3 \colon\ x^a - x \equiv y^k \pmod{f(x)}\right\}. \tag{D.9}$$

One can straightforwardly numerically verify that $T_2 = \{0, 1, 14, 25, 58\}$. To ensure understanding of the construction, we work through why $14 \in T_2$. We need to show that $x^{14} - x \equiv y^k \pmod{f(x)}$ for $k = 0, 1$ or 2. It turns out that $k = 2$ satisfies this equation. In particular,

$$(x^{14} - x) \pmod{f(x)} = x^3 + x^4 + x^5 = y^2 \pmod{f(x)} = x^{42} \pmod{f(x)}, \tag{D.10}$$

where recall we're working with polynomials over the field $\mathbb{F}_2$. Similarly, for 25,

$$(x^{25} - x) \pmod{f(x)} = 1 + x^3 + x^4 + x^5 = y^1 \pmod{f(x)} = x^{21} \pmod{f(x)}, \tag{D.11}$$

and for 58,

$$(x^{58} - x) \pmod{f(x)} = 1 = y^0 \pmod{f(x)}. \tag{D.12}$$

Hence, we have found that $T_2 = \{0, 1, 14, 25, 58\}$. To get our Sidon set, we compute the residues $S_2 = T_2 \bmod \frac{(n-1)^{t+1}-1}{n-2} = T_2 \bmod 21$, giving

$$S_2 = \{0, 1, 14, 4, 16\} = \{0, 1, 4, 14, 16\}\,. \tag{D.13}$$

One can easily confirm that this is a Sidon set mod 21. In particular, the set of all sums $a + b \bmod 21$ for $a, b \in S_2$ is $\{0, 1, 2, 4, 5, 7, 8, 9, 11, 14, 15, 16, 17, 18, 20\}$, which has size $15 = \binom{n+t-1}{t} = \binom{6}{2}$, which is the maximal possible size.

Section D.2: Pullback of the Fubini-Study volume form

It is shown in Ref. [72, Sec. 4.5, 4.7, 7.6] that the volume measure on complex projective space is the product of the flat measure on the simplex and the flat measure on the torus. For completeness, in this appendix, we show the same result via a different method.

Let $[Z_0 : \cdots : Z_n]$ be homogeneous coordinates on $\mathbb{CP}^n$. Consider the coordinate patches $C_0, \ldots, C_n$ on $\mathbb{CP}^n$, where $C_i = \{[Z_0 : \cdots : Z_n] \mid Z_i \neq 0\}$. The volume of $\mathbb{CP}^{d-1} \setminus C_0$ is zero, and therefore for the purposes of volume integration we can restrict our attention to C_0. On C_0, we use the coordinates $z_i \coloneqq Z_i/Z_0$ for $i = 1, \ldots, n$. The (unnormalized) Fubini-Study volume form ω can then be written as

$$\omega = \frac{1}{(1 + \sum_{i=1}^n |z_i|^2)^{n+1}} \, \mathrm{d}z_1 \wedge \mathrm{d}\bar{z}_1 \wedge \ldots \mathrm{d}z_n \wedge \mathrm{d}\bar{z}_n. \tag{D.14}$$

We can write $Z_i = \sqrt{p_i} \mathrm{e}^{\mathrm{i}\phi_i}$ for $i = 0, \ldots, n$ and $\sum_{i=0}^n p_i = 1$. In other words, p is a point on the simplex $p \in \Delta^n \coloneqq \{p \in [0,1]^n \mid \sum_i p_i \leq 1\}$ (with $p_0 \coloneqq 1 - \sum_{i=1}^n p_i$) and ϕ is a point on the projective torus $\phi \in P(T^{n+1})$ (*e.g.* we can choose a representative with $\phi_0 = 0$). Therefore, $z_i = \sqrt{\frac{p_i}{p_0}} \mathrm{e}^{\mathrm{i}\phi_i - \mathrm{i}\phi_0}$.

Consider the map $\pi \colon \tilde{\Delta}^n \times P(T^{n+1}) \to C_0$, where $\tilde{\Delta}^n$ is all $p \in \Delta^n$ satisfying $p_0 > 0$. The map is $\pi^i(p, \phi) = \sqrt{\frac{p_i}{p_0}} \mathrm{e}^{\mathrm{i}\phi_i - \mathrm{i}\phi_0}$.

Proposition D.2.1. *The pullback $\pi^*\omega$ is*

$$\pi^*\omega = (-1)^{n/2} \, \mathrm{d}p_1 \wedge \ldots \mathrm{d}p_n \wedge \mathrm{d}\phi_1 \wedge \ldots \mathrm{d}\phi_n. \tag{D.15}$$

It follows from this proposition that the unit-volume normalized volume measure on $\mathbb{CP}^n$ is equal to the product of the Lebesgue measure on the simplex Δ^n and the Lebesgue measure on $P(T^{n+1})$ (where recall the latter is equal to the Lebesgue measure on T^n).

Proof of the proposition. We can without loss of generality fix $\phi_0 = 0$. We can rewrite

$$\omega = p_0^{n+1} \, \mathrm{d}z_1 \wedge \mathrm{d}\bar{z}_1 \wedge \ldots \mathrm{d}z_n \wedge \mathrm{d}\bar{z}_n. \tag{D.16}$$

Therefore,

$$\pi^*\omega = p_0^{n+1} \det(J) \, \mathrm{d}p_1 \wedge \ldots \mathrm{d}p_n \wedge \mathrm{d}\phi_1 \wedge \ldots \mathrm{d}\phi_n, \tag{D.17}$$

where

$$J = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{D.18}$$

is the Jacobian with

$$A_{ij} = \frac{\partial \pi^i}{\partial \phi_j}, \quad B_{ij} = \frac{\partial \pi^i}{\partial p_j}, \quad C_{ij} = \frac{\partial \bar{\pi}^i}{\partial \phi_j}, \quad D_{ij} = \frac{\partial \bar{\pi}^i}{\partial p_j}. \tag{D.19}$$

We can check that

$$\frac{\partial \pi^i}{\partial p_j} = \frac{1}{2}\pi^i(p,\phi)\left(\frac{\delta_{ij}}{p_i} + \frac{1}{p_0}\right), \qquad \frac{\partial \pi^i}{\partial \phi_j} = \mathrm{i}\delta_{ij}\pi^i(p,\phi). \tag{D.20}$$

Therefore, A and C are diagonal and thus commute, meaning that $\det(J) = \det(AD - CB)$. The matrix elements are $(AD - CB)_{ij} = \frac{\mathrm{i}}{p_0}\left(\delta_{ij} + \frac{p_i}{p_0}\right)$.

By the matrix determinant lemma [313], $\det(M + uv^T) = \det(M)(1 + v^T M^{-1} u)$ with $M_{ij} = \frac{\mathrm{i}}{p_0}\delta_{ij}$ and $u_i = \mathrm{i}/p_0$ and $v_i = p_i/p_0$, we find that

$$\det(J) = \left(\frac{\mathrm{i}}{p_0}\right)^n \left(1 + \sum_{i=1}^n \frac{p_i}{p_0}\right) = \left(\frac{i}{p_0}\right)^n \frac{1}{p_0} = \frac{(-1)^{n/2}}{p_0^{n+1}}. \tag{D.21}$$

The proposition follows. □

Section D.3: Approximate projective toric designs

Throughout this appendix, we are concerned with uniform finite $P(T^n)$ t-designs; that is, $P(T^n)$ t-designs X that are finite and the measure space $(X, \Sigma = \mathcal{P}(X), \nu)$ is such that $\nu(A) = |A| / |X|$. We restrict to finite n. We define approximate projective toric designs and prove a loose bound on the number $M(t, \varepsilon, \delta)$ of uniformly random points needed to form such a design.

For $\mathbf{p} \in \mathbb{N}_0^n$, let $f_{\mathbf{p}}(\phi)$ denote the monomial $\prod_{i=1}^n \mathrm{e}^{i\phi_i p_i}$. Notice that $\bar{f}_{\mathbf{p}}(\phi) = f_{\mathbf{p}}(-\phi) =$

$f_{-\mathbf{p}}(\phi)$.

Definition D.3.1. *We say that $C \subset P(T^n)$ is a (uniform) ε-approximate projective toric t-design if, for all $\mathbf{p} \in P_t^{(n)}$,*

$$\left| \frac{1}{|C|} \sum_{\phi \in C} f_{\mathbf{p}}(\phi) - \int_{P(T^n)} f_{\mathbf{p}} \, \mathrm{d}\mu_{n-1} \right| = \left| \frac{1}{|C|} \sum_{\phi \in C} f_{\mathbf{p}}(\phi) - \delta_{\mathbf{p},\mathbf{0}} \right| \leq \varepsilon. \tag{D.22}$$

Here, $P_t^{(n)}$ is the set defined in Eq. (6.12),

$$P_t^{(n)} \coloneqq \left\{ \mathbf{q} - \mathbf{r} \;\middle|\; \mathbf{q}, \mathbf{r} \in \mathbb{N}_0^n, \; \sum_{i=1}^n q_i = \sum_{i=1}^n r_i = t \right\}. \tag{D.23}$$

Note of course that with $\varepsilon = 0$ we recover the definition of an (exact) projective toric design. There is redundancy in $P_t^{(n)}$. Indeed, if Eq. (D.22) is satisfied for $\mathbf{p}$, then it is automatically satisfied for $-\mathbf{p}$. Furthermore, Eq. (D.22) is trivially satisfied for any C when $\mathbf{p} = \mathbf{0}$. Hence, we are in fact interested in the set $S_t^{(n)}$ defined by $S_t^{(n)} \coloneqq (P_t^{(n)} \setminus \{\mathbf{0}\})/\mathbb{Z}_2$, where $\mathbb{Z}_2$ denotes the group action $\mathbf{p} \mapsto \pm\mathbf{p}$. Therefore, $C \subset P(T^n)$ is an ε-approximate t-design if and only if, for all $\mathbf{p} \in S_t^{(n)}$,

$$\left| \frac{1}{|C|} \sum_{\phi \in C} f_{\mathbf{p}}(\phi) \right| \leq \varepsilon. \tag{D.24}$$

Define the probability space $\mathcal{C}_M$ to be the ensemble over subsets of $P(T^n)$ of size M. Specifically, to draw a random $C \subset P(T^n)$ from $\mathcal{C}_M$, we simply draw M uniformly random points from $P(T^n)$ with respect to the Haar measure. We often denote a subset $C \subset P(T^n)$ of size M by $C = \{\phi^{(1)}, \ldots, \phi^{(M)}\}$, where each $\phi^{(i)} \in P(T^n)$.

Definition D.3.2. *Let $M(t, \varepsilon, \delta)$ denote the **minimum** M such that a random C drawn*

from $\mathcal{C}_M$ is an ε-approximate $P(T^n)$ t-design with probability $1-\delta$. In other words,

$$\begin{aligned} M(t,\varepsilon,\delta) = \min_{M\in\mathbb{N}} & M \\ s.t. \quad & \Pr_{C\in\mathcal{C}_M}[C \text{ is an } \varepsilon\text{-approx } t\text{-design on } P(T^n)] \geq 1-\delta. \end{aligned} \tag{D.25}$$

In the following, we find an upper bound on $M(t,\varepsilon,\delta)$. This tells us that for any $M \geq M(t,\varepsilon,\delta)$, $C \in \mathcal{C}_M$ is an ε-approximate t-design with probability $\geq 1-\delta$.

Theorem D.3.3. *$M(t,\varepsilon,\delta) \leq \frac{G_{n-1}(t)-1}{2\delta\varepsilon^2}$, where $G_{n-1}(t)$ given in Eq. (6.13).*

Proof. Recall that $|P_t^{(n)}| = 2|S_t^{(n)}| + 1$, and from Eq. (6.14) $|P_t^{(n)}| = G_{n-1}(t)$. We will therefore prove that $M(t,\varepsilon,\delta) \leq \frac{|S_t^{(n)}|}{\delta\varepsilon^2}$ Define the following notation:

$$\mathop{\mathbb{E}}_{\phi\in P(T^n)} f(\phi) = \int_{P(T^n)} f \,\mathrm{d}\mu_{n-1} \tag{D.26a}$$

$$\mathop{\mathbb{E}}_{\phi\in C} f(\phi) = \frac{1}{|C|}\sum_{\phi\in C} f(\phi) \tag{D.26b}$$

$$\mathop{\mathbb{E}}_{C\in\mathcal{C}_M} = \mathop{\mathbb{E}}_{\{\phi^{(1)},\ldots,\phi^{(M)}\}\in P(T^n)^M}. \tag{D.26c}$$

For $\mathbf{p} \in S_t^{(n)}$, define

$$\Delta(C,\mathbf{p}) \coloneqq \left|\mathop{\mathbb{E}}_{\phi\in C} f_{\mathbf{p}}(\phi) - \mathop{\mathbb{E}}_{\phi\in P(T^n)} f_{\mathbf{p}}(\phi)\right|^2 = \left|\mathop{\mathbb{E}}_{\phi\in C} f_{\mathbf{p}}(\phi)\right|^2. \tag{D.27}$$

We compute the mean,

$$\mathop{\mathbb{E}}_{C\in\mathcal{C}_M} \Delta(C,\mathbf{p}) = \frac{1}{M^2} \mathop{\mathbb{E}}_{\{\phi^{(1)},\ldots,\phi^{(M)}\}\in P(T^n)^M} \sum_{i,j=1}^{M} f_{\mathbf{p}}(\phi^{(i)})\bar{f}_{\mathbf{p}}(\phi^{(j)}) \tag{D.28a}$$

$$= \frac{1}{M^2}\left[M \mathop{\mathbb{E}}_{\phi\in P(T^n)} f_{\mathbf{p}}(\phi)\bar{f}_{\mathbf{p}}(\phi) + M(M-1)\left(\mathop{\mathbb{E}}_{\phi\in P(T^n)} f_{\mathbf{p}}(\phi)\right)\left(\mathop{\mathbb{E}}_{\theta\in P(T^n)} \bar{f}_{\mathbf{p}}(\theta)\right)\right] \tag{D.28b}$$

$$= \frac{1}{M}. \tag{D.28c}$$

Meanwhile, we have that

$$\Pr_{C\in\mathcal{C}_M}\left[C \text{ is an } \varepsilon\text{-approx } t\text{-design on } P(T^n)\right] \tag{D.29a}$$

$$= \Pr_{C\in\mathcal{C}_M}\left[\forall \mathbf{p}\in S_t^{(n)} : \Delta(C,\mathbf{p}) \le \varepsilon^2\right] \tag{D.29b}$$

$$= 1 - \Pr_{C\in\mathcal{C}_M}\left[\exists \mathbf{p}\in S_t^{(n)} : \Delta(C,\mathbf{p}) > \varepsilon^2\right] \tag{D.29c}$$

$$(\textit{union bound}) \ge 1 - \sum_{\mathbf{p}\in S_t^{(n)}} \Pr_{C\in\mathcal{C}_M}\left[\Delta(C,\mathbf{p}) > \varepsilon^2\right] \tag{D.29d}$$

$$(\textit{Markov's inequality}) \ge 1 - \sum_{\mathbf{p}\in S_t^{(n)}} \frac{\mathbb{E}_{C\in\mathcal{C}_M}\Delta(C,\mathbf{p})}{\varepsilon^2} \tag{D.29e}$$

$$= 1 - \frac{|S_t^{(n)}|}{M\varepsilon^2}. \tag{D.29f}$$

Thus, we require that

$$1 - \frac{1}{M\varepsilon^2}|S_t^{(n)}| \ge 1 - \delta. \tag{D.30}$$

It follows that any $M \ge \frac{|S_t^{(n)}|}{\delta\varepsilon^2}$ satisfies, so that $M(t,\varepsilon,\delta) \le \frac{|S_t^{(n)}|}{\delta\varepsilon^2}$. □

Appendix E: Technical details for Chapter 7

Section E.1: Filter function formalism

Subsection E.1.1: Review of instantaneous control filter functions

In this section, we review filter functions [285–288] for instantaneous controls in order to derive Eq. (7.6). We are considering a qubit coupled via the Hamiltonian in Eq. (7.5) when the controls $c(t)$ are instantaneous. Specifically, suppose that we initialize the qubit in the $|+\rangle \coloneqq \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ state and subject it to the Hamiltonian $H(t) = (\gamma(t) + \lambda(t))\sigma^z$ while applying instantaneous σ^x gates at times $t_1 < \cdots < t_M$. Define $t_0 = 0$. For notational convenience, define $\gamma(t) \coloneqq \gamma(t) + \lambda(t)$. Then the state after a time $t > t_M$ is

$$|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}\int_{t_M}^{t}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'}\prod_{j=M}^{1}\left(\sigma^x\mathrm{e}^{-\mathrm{i}\int_{t_{j-1}}^{t_j}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'}\right)|+\rangle \tag{E.1a}$$

$$= \mathrm{e}^{-\mathrm{i}\int_{t_M}^{t}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'}\sigma^x\mathrm{e}^{-\mathrm{i}\int_{t_{M-1}}^{t_M}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'}\sigma^x\mathrm{e}^{-\mathrm{i}\int_{t_{M-2}}^{t_{M-1}}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'}\prod_{j=M-2}^{1}\left(\sigma^x\mathrm{e}^{-\mathrm{i}\int_{t_{j-1}}^{t_j}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'}\right)|+\rangle \tag{E.1b}$$

$$= \mathrm{e}^{-\mathrm{i}\int_{t_M}^{t}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'+\mathrm{i}\int_{t_{M-1}}^{t_M}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'-\mathrm{i}\int_{t_{M-2}}^{t_{M-1}}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'}\prod_{j=M-2}^{1}\left(\sigma^x\mathrm{e}^{-\mathrm{i}\int_{t_{j-1}}^{t_j}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'}\right)|+\rangle \tag{E.1c}$$

$$= \ldots \tag{E.1d}$$

$$= \exp\left[-\mathrm{i}\int_{t_M}^{t}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t' + \mathrm{i}\sum_{j=1}^{M}(-1)^{M-j}\int_{t_{j-1}}^{t_j}\tilde{\gamma}(t')\sigma^z\,\mathrm{d}t'\right]|+\rangle\,, \tag{E.1e}$$

where in the last line, we used that if M is odd, we can still sandwich that term with σ^x since $\sigma^x |+\rangle = |+\rangle$. Finally, we define the *switching function* $f(t)$ by

$$f(t) = \begin{cases} 1 & \text{if } t_M < t \text{ or } t_{j-1} < t < t_j \text{ for } M-j \text{ even} , \\ -1 & \text{if } t_{j-1} < t < t_j \text{ for } M-j \text{ odd} . \end{cases} \tag{E.2}$$

Then

$$|\psi(t)\rangle = \exp\left[-\mathrm{i}\int_0^t f(t')\tilde{\gamma}(t')\sigma^z \,\mathrm{d}t'\right] |+\rangle . \tag{E.3}$$

The probability $P(t) \coloneqq |\langle +|\psi(t)\rangle|^2$ of measuring 1 in the σ^x basis at time t is then

$$P(t) = \cos^2\left(\int_0^t f(t')\tilde{\gamma}(t') \,\mathrm{d}t'\right) . \tag{E.4}$$

We recall that $\gamma(t)$ is a random variable. The expectation $\langle P(t)\rangle$ of $P(t)$ is

$$\begin{aligned}
\langle P(t)\rangle &= \frac{1}{2} + \frac{1}{4}\left(\left\langle \mathrm{e}^{2\mathrm{i}\int_0^t f(t')\tilde{\gamma}(t')\,\mathrm{d}t'}\right\rangle + \left\langle \mathrm{e}^{-2\mathrm{i}\int_0^t f(t')\tilde{\gamma}(t')\,\mathrm{d}t'}\right\rangle\right) && \text{(E.5a)} \\
&= \frac{1}{2} + \frac{1}{2}\mathrm{e}^{-2\int_0^t\int_0^t f(t')f(t'')\langle\tilde{\gamma}(t')\tilde{\gamma}(t'')\rangle\,\mathrm{d}t'\,\mathrm{d}t''} && \text{(E.5b)} \\
&= \frac{1}{2} + \frac{1}{2}\mathrm{e}^{-2\int_0^t\int_0^t f(t')f(t'')(C(t',t'')+C_\lambda(t',t''))\,\mathrm{d}t'\,\mathrm{d}t''} && \text{(E.5c)} \\
&= \frac{1}{2} + \frac{1}{2}\mathrm{e}^{-2\int_0^t\int_0^t f(t')f(t'')(C(0,t''-t')+C_\lambda(0,t''-t'))\,\mathrm{d}t'\,\mathrm{d}t''} && \text{(E.5d)} \\
&= \frac{1}{2} + \frac{1}{2}\mathrm{e}^{-\frac{1}{\pi}\int_{-\infty}^{\infty}\mathrm{d}\omega(S(\omega)+S_\lambda(\omega))\int_0^t\int_0^t f(t')f(t'')\mathrm{e}^{\mathrm{i}\omega(t''-t')}\,\mathrm{d}t'\,\mathrm{d}t''} && \text{(E.5e)} \\
&= \frac{1}{2} + \frac{1}{2}\mathrm{e}^{-\frac{1}{\pi}\int_{-\infty}^{\infty}(S(\omega)+S_\lambda(\omega))F(\omega,t)\,\mathrm{d}\omega}, && \text{(E.5f)}
\end{aligned}$$

where we defined the C and C_λ to be the correlations for the Gaussian processes γ and λ, used that $\langle\gamma(t)\rangle = 0$ and that for any Gaussian random variable X, $\left\langle \mathrm{e}^{\alpha X}\right\rangle = \mathrm{e}^{\alpha\langle X\rangle+\alpha^2\langle X^2\rangle/2}$. Finally, we defined the *filter function* [285–288]

$$F(\omega,t) = \left|\int_0^t f(t')\mathrm{e}^{\mathrm{i}\omega t'}\,\mathrm{d}t'\right|^2 . \tag{E.6}$$

We define $\chi(t) \coloneqq \frac{1}{\pi}\int S(\omega)F(\omega,t)\,\mathrm{d}\omega$, and similarly define $\chi_\lambda(t)$, so that $\langle P(t)\rangle = \frac{1}{2} + \frac{1}{2}\mathrm{e}^{-\chi(t)-\chi_\lambda(t)}$, thus reproducing Eq. (7.6) as desired. The key point of the filter function formalism is that the details of the signal γ are entirely encoded by $S(\omega)$, the noise λ by $S_\lambda(\omega)$, and the details of our control sequence are entirely encoded by the filter function $F(\omega,t)$.

$\langle P(t)\rangle$ encodes the parameter $\Delta\omega$. Then we can determine the Fisher information $\mathrm{FI}_{\Delta\omega}(\langle P(t)\rangle)$ of $\langle P(t)\rangle$ with respect to $\Delta\omega$. The Fisher information tells us how accessible the information about the parameter is given access to the distribution. The Fisher information in the noiseless case is

$$\mathrm{FI}_{\Delta\omega}(\langle P(t)\rangle) = \frac{1}{\langle P(t)\rangle}\left(\frac{\partial\langle P(t)\rangle}{\partial\Delta\omega}\right)^2 + \frac{1}{1-\langle P(t)\rangle}\left(\frac{\partial(1-\langle P(t)\rangle)}{\partial\Delta\omega}\right)^2 \tag{E.7a}$$

$$= \left(\frac{\partial\langle P(t)\rangle}{\partial\Delta\omega}\right)^2 \frac{1}{\langle P(t)\rangle\left(1-\langle P(t)\rangle\right)} \tag{E.7b}$$

$$= \frac{1}{2}\left(\coth\chi(t)-1\right)\left(\frac{\partial\chi(t)}{\partial\Delta\omega}\right)^2. \tag{E.7c}$$

Given the spectrum given in Eq. (7.7) for the signal $\gamma(t)$,

$$\chi(t) = g^2\left(F(\omega_1,t)+F(\omega_2,t)+F(-\omega_1,t)+F(-\omega_2,t)\right). \tag{E.8}$$

With this signal, we are interested in the parameter $\Delta\omega = \omega_2 - \omega_1$. Assuming $\Delta\omega$ is small, this can be expanded around the centroid $\omega_c = (\omega_1+\omega_2)/2$ as

$$\chi(t) = g^2\left(4F(\omega_c,t) + \frac{1}{2}\Delta\omega^2 F''(\omega_c,t)\right) + \mathcal{O}\left(g^2\Delta\omega^4\right), \tag{E.9}$$

where we used that $F(\omega,t) = F(-\omega,t)$, and we denote by F'' the second derivative of F

with respect to ω. It follows that

$$\mathrm{FI}_{\Delta\omega}(\langle P(t)\rangle) = \begin{cases} g^2 F''(\omega_c, t) + \mathcal{O}(\Delta\omega^2) & \text{if } F(\omega_c, t) = 0 \\ \frac{1}{2} g^4 F''(\omega_c, t)^2 \left[\coth(4g^2 F(\omega_c, t)) - 1\right] \Delta\omega^2 + \mathcal{O}(\Delta\omega^4) & \text{if } F(\omega_c, t) \neq 0 \end{cases}. \tag{E.10}$$

Recall that we formally defined a procedure to exhibit *superresolution* if $\lim_{\Delta\omega\to 0} \mathrm{FI}_{\Delta\omega}(\langle P(t)\rangle) > 0$. We have therefore found a characterization of superresolution in terms of only the controls (*i.e.* the filter function). A sequence of control σ^x gates characterized by the switching function $f(t)$ gives rise to superresolution if and only if the filter function vanishes as the centroid, $F(\omega_c, t) = 0$. In summary, our goal in designing a superresolution protocol is to design a control sequence such that $F(\omega_c, t) = 0$ and $F''(\omega_c, t)$ is maximized.

We immediately find such superresolution protocols by consider free and CPMG evolution. Define the characteristic timescale $\tau \coloneqq 2\pi/\omega_c$. In the *free evolution* protocol, we apply no controls, so that the switching function $f_{\text{free}}(t) = 1$ for all t. From Eq. (E.6), the filter function is therefore

$$F_{\text{free}}(\omega, \kappa\tau) = \frac{4}{\omega^2} \sin^2\left(\frac{\pi\kappa\omega}{\omega_c}\right). \tag{E.11}$$

In the *CPMG protocol*, we assume that κ is an even integer, and our control is a sequence of $\kappa/2$ CPMG sequences; that is, we apply σ^x gates at times $\tau/2, 3\tau/2, 5\tau/2, \ldots, (2\kappa-1)\tau/2$. The switching function is thus $f_{\text{CPMG}}(t) = \operatorname{sgn}\cos(\pi t/\tau)$, where sgn is the sign function. From Eq. (E.6), the filter function is

$$F_{\text{CPMG}}(\omega, \kappa\tau) = \frac{1}{\omega^2}\left|(1 - \mathrm{e}^{\mathrm{i}\omega\tau/2}) + (\mathrm{e}^{\mathrm{i}\omega(2\kappa-1)\tau/2} - \mathrm{e}^{\mathrm{i}\omega\kappa\tau}) + \sum_{j=1}^{\kappa-1}(-1)^j(\mathrm{e}^{\mathrm{i}\omega(2j-1)\tau/2} - \mathrm{e}^{\mathrm{i}\omega(2j+1)\tau/2})\right|^2 \tag{E.12a}$$

$$= \frac{16}{\omega^2} \sin^4\left(\frac{\pi\omega}{2\omega_c}\right) \sec^2\left(\frac{\pi\omega}{\omega_c}\right) \sin^2\left(\frac{\pi\kappa\omega}{\omega_c}\right) \tag{E.12b}$$

$$= 4F_{\text{free}}(\omega, \kappa\tau) \sin^4\left(\frac{\pi\omega}{2\omega_c}\right) \sec^2\left(\frac{\pi\omega}{\omega_c}\right). \tag{E.12c}$$

We therefore see that the free evolution protocol exhibits superresolution when $\kappa \in \mathbb{Z}$ and the CPMG protocol exhibits superresolution when $\kappa \in 2\mathbb{Z}$ because $F_*(\omega_c, \kappa\tau) = 0$. We then find that the Fisher informations are exactly as in Eq. (7.14).

Subsection E.1.2: Concentration of the probability distribution

Recall that $P = \cos^2\phi$, where $\phi = \int_0^{\kappa\tau} f(t')\gamma(t')\,\mathrm{d}t'$. Furthermore, from Eq. (7.6), $\langle\phi^2\rangle = \frac{1}{2}\chi$. Therefore, using that for any Gaussian random variable X, $\left\langle \mathrm{e}^{\alpha X}\right\rangle = \mathrm{e}^{\alpha\langle X\rangle + \alpha^2\langle X^2\rangle/2}$, we have

$$\begin{aligned}
\langle P^2\rangle - \langle P\rangle^2 &= \left\langle \frac{3}{8} + \frac{1}{4}\mathrm{e}^{-2\mathrm{i}\phi} + \frac{1}{4}\mathrm{e}^{2\mathrm{i}\phi} + \frac{1}{16}\mathrm{e}^{-4\mathrm{i}\phi} + \frac{1}{16}\mathrm{e}^{4\mathrm{i}\phi}\right\rangle - \langle P\rangle^2 && \text{(E.13a)}\\
&= \frac{3}{8} + \frac{1}{2}\mathrm{e}^{-2\langle\phi^2\rangle} + \frac{1}{8}\mathrm{e}^{-4\langle\phi^2\rangle} - \langle P\rangle^2 && \text{(E.13b)}\\
&= \frac{3}{8} + \frac{1}{2}\mathrm{e}^{-\chi} + \frac{1}{8}\mathrm{e}^{-4\chi} - \langle P\rangle^2 && \text{(E.13c)}\\
&= \frac{3}{8} + \frac{1}{2}\mathrm{e}^{-\chi} + \frac{1}{8}\mathrm{e}^{-4\chi} - \left(\frac{1+\mathrm{e}^{-\chi}}{2}\right)^2 && \text{(E.13d)}\\
&= \frac{1}{8}\left(1 + \mathrm{e}^{-4\chi} - 2\mathrm{e}^{-2\chi}\right) && \text{(E.13e)}\\
&= 2\langle P\rangle^2\left(1 - \langle P\rangle\right)^2. && \text{(E.13f)}
\end{aligned}$$

This concentration equality is important to derive the conclusion around Eq. (7.13) that superresolution procedures are practically unaffected by the fact that the analyses assume sampling from $\langle P\rangle$ whereas the procols are in fact sampling from P for different realizations of the random variables A_1, A_2, B_1, and B_2.

Subsection E.1.3: Review of continuous control filter functions

In Appendix E.1.1, we reviewed the filter function formalism for a qubit coupled to a time-dependent signal and instantaneous σ^x π-pulse controls. Using the Magnus expansion from Ref. [289], we will now analyze the case when controls are allowed to be more general. For

simplicity, we will consider the noiseless case $\lambda(t) = 0$ in this derivation. We discuss the addition of noise at the end.

As before, our signal plus noise is coupled to a qubit via $H_0(t) = \gamma(t)\sigma^z$. We suppose furthermore that we supply a control Hamiltonian $H_c(t) = c(t)\sigma^x$, so that the full Hamiltonian is $H(t) = H_0(t) + H_c(t)$, as in Eq. (7.5). In the case of instantaneous controls as above, $c(t)$ is simply a sum of delta functions. Our goal is, as before, to find $P(t) \coloneqq |\langle +| U(t) |+\rangle|^2$, where $U(t) = \mathcal{T}\exp\left[-\int_0^t H(t')\,\mathrm{d}t'\right]$, with $\mathcal{T}\exp$ denoting the time ordered exponential. We introduce the control propagator

$$U_c(t) = \exp\left(-\int_0^t H_c(t')\,\mathrm{d}t'\right) = \cos(\theta_c(t))\mathbb{I} - \mathrm{i}\sin(\theta_c(t))\sigma^x, \qquad \theta_c(t) = \int_0^t c(t')\,\mathrm{d}t', \quad \text{(E.14)}$$

meaning that $U(t) = U_c(t)\tilde{U}(t)$, where

$$\tilde{U}(t) = \mathcal{T}\exp\left[-\int_0^t \tilde{H}(t')\,\mathrm{d}t'\right], \qquad \tilde{H}(t) = U_c(t)^\dagger H_0(t) U_c(t) = \gamma(t)\left(\cos(2\theta_c(t))\sigma^z + \sin(2\theta_c(t))\sigma^y\right). \quad \text{(E.15)}$$

Define the vector $\boldsymbol{a}(t) = a(t)\hat{\boldsymbol{a}}(t)$, where $\hat{\boldsymbol{a}}$ is a unit vector, such that $\tilde{U}(t) = \mathrm{e}^{-\mathrm{i}a(t)\hat{\boldsymbol{a}}(t)\cdot\boldsymbol{\sigma}}$. Then we have

$$\begin{aligned}
P(t) &= |\langle +| U_c(t)\tilde{U}(t) |+\rangle|^2 && \text{(E.16a)}\\
&= |\langle +| \tilde{U}(t) |+\rangle|^2 && \text{(E.16b)}\\
&= |\cos a(t) - \mathrm{i}\sin a(t)\hat{\boldsymbol{a}}(t)\cdot\langle +|\boldsymbol{\sigma}|+\rangle|^2 && \text{(E.16c)}\\
&= |\cos a(t) - \mathrm{i}\sin a(t) a_x(t)|^2 && \text{(E.16d)}\\
&= \cos^2 a(t) + a_x(t)^2\sin^2 a(t), && \text{(E.16e)}
\end{aligned}$$

where we defined $\hat{\boldsymbol{a}}(t) = (a_x(t), a_y(t), a_z(t))$.

Thus, the remaining task is to calculate $\boldsymbol{a}(t)$. We use the results from Ref. [289] to

compute $a(t)$. With a first order Magnus expansion, we find that

$$\boldsymbol{a}^{(1)}(t) = \int_0^t \gamma(t') \left(0, \sin(2\theta_c(t')), \cos(2\theta_c(t'))\right) \mathrm{d}t'. \tag{E.17}$$

The second order term in the Magnus expansion yields

$$\boldsymbol{a}^{(2)}(t) = \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \; \gamma(t_1)\gamma(t_2) \left(\sin(2\theta_c(t_1) - 2\theta_c(t_2)), 0, 0\right). \tag{E.18}$$

Thus, up to order g^4, we have that [1]

$$\langle P(t)\rangle = 1 - \langle a(t)^2\rangle + \frac{1}{3}\langle a(t)^4\rangle + \langle a_x(t)^2 a(t)^2\rangle = 1 - \langle \|\boldsymbol{a}^{(1)}\|^2\rangle - \langle \|\boldsymbol{a}^{(2)}\|^2\rangle + \frac{1}{3}\langle \|\boldsymbol{a}^{(1)}\|^4\rangle$$

(E.19a)

$$\begin{aligned}
&= 1 - \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \langle \gamma(t_1)\gamma(t_2)\rangle \cos(2\theta_c(t_1) - 2\theta_c(t_2)) \\
&\quad - \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^t \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_4 \; \langle \gamma(t_1)\gamma(t_2)\gamma(t_3)\gamma(t_4)\rangle \sin(2\theta_c(t_1) - 2\theta_c(t_2)) \sin(2\theta_c(t_3) - 2\theta_c(t_4)) \\
&\quad + \frac{1}{3}\int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \int_0^t \mathrm{d}t_3 \int_0^t \mathrm{d}t_4 \langle \gamma(t_1)\gamma(t_2)\gamma(t_3)\gamma(t_4)\rangle \cos(2\theta_c(t_1) - 2\theta_c(t_2)) \cos(2\theta_c(t_3) - 2\theta_c(t_4))
\end{aligned} \tag{E.19b}$$

[1] Note that odd order terms have vanished because γ has zero mean, and we ignore all terms of order g^6 or smaller.

$$
\begin{aligned}
&= 1 - \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 C(0, t_2 - t_1) \cos(2\theta_c(t_1) - 2\theta_c(t_2)) \\
&\quad - \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^t \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_4 \\
&\qquad (C(0, t_2 - t_1)C(0, t_4 - t_3) + C(0, t_3 - t_1)C(0, t_4 - t_2) + C(0, t_4 - t_1)C(0, t_3 - t_2)) \\
&\qquad \sin(2\theta_c(t_1) - 2\theta_c(t_2)) \sin(2\theta_c(t_3) - 2\theta_c(t_4)) \\
&\quad + \frac{1}{3} \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \int_0^t \mathrm{d}t_3 \int_0^t \mathrm{d}t_4 \\
&\qquad (C(0, t_2 - t_1)C(0, t_4 - t_3) + C(0, t_3 - t_1)C(0, t_4 - t_2) + C(0, t_4 - t_1)C(0, t_3 - t_2)) \\
&\qquad \cos(2\theta_c(t_1) - 2\theta_c(t_2)) \cos(2\theta_c(t_3) - 2\theta_c(t_4))
\end{aligned}
\tag{E.19c}
$$

$$
\begin{aligned}
&= 1 - \frac{1}{2\pi} \int \mathrm{d}\omega S(\omega) \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \mathrm{e}^{\mathrm{i}\omega(t_2 - t_1)} \cos(2\theta_c(t_1) - 2\theta_c(t_2)) \\
&\quad - \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^t \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_4 \\
&\qquad (C(0, t_2 - t_1)C(0, t_4 - t_3) + C(0, t_3 - t_1)C(0, t_4 - t_2) + C(0, t_4 - t_1)C(0, t_3 - t_2)) \\
&\qquad \sin(2\theta_c(t_1) - 2\theta_c(t_2)) \sin(2\theta_c(t_3) - 2\theta_c(t_4)) \\
&\quad + \frac{1}{3} \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \int_0^t \mathrm{d}t_3 \int_0^t \mathrm{d}t_4 C(0, t_2 - t_1)C(0, t_4 - t_3) \Bigg[\\
&\qquad \cos(2\theta_c(t_1) - 2\theta_c(t_2)) \cos(2\theta_c(t_3) - 2\theta_c(t_4)) + \cos(2\theta_c(t_1) - 2\theta_c(t_3)) \cos(2\theta_c(t_2) - 2\theta_c(t_4)) \\
&\qquad + \cos(2\theta_c(t_1) - 2\theta_c(t_4)) \cos(2\theta_c(t_2) - 2\theta_c(t_3)) \Bigg]
\end{aligned}
\tag{E.19d}
$$

$$
= 1 - \frac{1}{2\pi} \int \mathrm{d}\omega S(\omega) F^{(2)}(\omega, t) - \frac{1}{12\pi^2} \int \mathrm{d}\omega_1 \int \mathrm{d}\omega_2 S(\omega_1) S(\omega_2) F^{(4)}(\omega_1, \omega_2, t),
\tag{E.19e}
$$

where we defined

$$
F^{(2)}(\omega, t) = \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \mathrm{e}^{\mathrm{i}\omega(t_2 - t_1)} \cos(2\theta_c(t_1) - 2\theta_c(t_2))
\tag{E.20a}
$$

$$= \left| \int_0^t \cos(2\theta_c(t'))\, e^{i\omega t'}\, dt' \right|^2 + \left| \int_0^t \sin(2\theta_c(t'))\, e^{i\omega t'}\, dt' \right|^2 \tag{E.20b}$$

as in Eq. (7.18), and

$$\begin{aligned} F^{(4)}(\omega_1, \omega_2, t) = -\int_0^t dt_1 \int_0^t dt_2 \int_0^t dt_3 \int_0^t dt_4 e^{i\omega_1(t_2-t_1)} e^{i\omega_2(t_4-t_3)} \Big[\\ \cos(2\theta_c(t_1) - 2\theta_c(t_2)) \cos(2\theta_c(t_3) - 2\theta_c(t_4)) + \cos(2\theta_c(t_1) - 2\theta_c(t_3)) \cos(2\theta_c(t_2) - 2\theta_c(t_4)) \\ + \cos(2\theta_c(t_1) - 2\theta_c(t_4)) \cos(2\theta_c(t_2) - 2\theta_c(t_3)) \Big] \\ + 3 \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^t dt_3 \int_0^{t_3} dt_4 e^{i\omega_1(t_2-t_1)} e^{i\omega_2(t_4-t_3)} \Big[\\ \sin(2\theta_c(t_1) - 2\theta_c(t_2)) \sin(2\theta_c(t_3) - 2\theta_c(t_4)) + \sin(2\theta_c(t_1) - 2\theta_c(t_3)) \sin(2\theta_c(t_2) - 2\theta_c(t_4)) \\ + \sin(2\theta_c(t_1) - 2\theta_c(t_4)) \sin(2\theta_c(t_2) - 2\theta_c(t_3)) \Big]. \end{aligned} \tag{E.21}$$

The upshot is then Eq. (E.19e), which Taylor expands to Eq. (7.17).

We can now see that this indeed reproduces the instantaneous case from above; when the controls are instantaneous, $c(t)$ takes the form $c(t) = \frac{\pi}{2} \sum_{i=1}^{M} \delta(t - t_i)$. In this case, $\theta_c(t) = \frac{\pi}{2}$(num of pulses before time t), so that $\sin(2\theta_c(t')) = 0$ and $\cos(2\theta_c(t')) = (-1)^{\text{num of pulses before time } t}$. Note that $\cos(2\theta_c(t'))$ is exactly the switching function $f(t)$. We see that $F^{(2)}(\omega, t)$ exactly matches Eq. (E.6), as expected. Similarly, $F^{(4)}(\omega_1, \omega_2, t) = 3F^{(2)}(\omega_1, t)F^{(2)}(\omega_2, t)$.

Finally, we consider adding a nonzero noise $\lambda(t) \neq 0$. Exactly as in the instantaneous control case, the lowest order addition to Eq. (E.19e) is simply $\frac{1}{2\pi} \int d\omega S_\lambda(\omega) F^{(2)}(\omega, t)$. At higher order, we have $\frac{-1}{12\pi^2} \int d\omega_1 \int d\omega_2 S_\lambda(\omega_1) S_\lambda(\omega_2) F^{(4)}(\omega_1, \omega_2, t)$ along with the cross term $\frac{-1}{12\pi^2} \int d\omega_1 \int d\omega_2 S_\lambda(\omega_1) S(\omega_2) F^{(4)}(\omega_1, \omega_2, t)$.

Subsection E.1.4: Local optimality of instantaneous controls for superresolution

In this short section, we show that optimizing the Fisher information by tuning the filter functions results in instantaneous controls being *local* optima. In particular, we would like to minimize the Lagrangian

$$\mathcal{L} = -F^{(2)\prime\prime}(\omega_c, \kappa\tau) + \lambda F^{(2)}(\omega_c, \kappa\tau) \tag{E.22a}$$

$$= \int_0^{\kappa\tau} \mathrm{d}t_1 \int_0^{\kappa\tau} \mathrm{d}t_2 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \cos(2\theta_c(t_2) - 2\theta_c(t_1)) \,, \tag{E.22b}$$

where we are assuming that we are in the regime where $F^{(4)}$ can be ignored, and λ is a Lagrange multiplier enforcing that the filter function vanishes at the centroid $\omega = \omega_c$. Because $\frac{\mathrm{d}}{\mathrm{d}t}\theta_c(t) = c(t)$, we can consider minimizing $\mathcal{L}$ with respect to $\theta_c(t)$. Performing a variation $\delta\theta_c(t)$, we find

$$\delta\mathcal{L} = \int_0^{\kappa\tau} \mathrm{d}t_1 \int_0^{\kappa\tau} \mathrm{d}t_2 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \\ [\cos(2\theta_c(t_2) - 2\theta_c(t_1) + 2\delta\theta_c(t_2) - 2\delta\theta_c(t_1)) - \cos(2\theta_c(t_2) - 2\theta_c(t_1))] \tag{E.23a}$$

$$= -2 \int_0^{\kappa\tau} \mathrm{d}t_1 \int_0^{\kappa\tau} \mathrm{d}t_2 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \sin(2\theta_c(t_2) - 2\theta_c(t_1)) \left(\delta\theta_c(t_2) - \delta\theta_c(t_1)\right) \tag{E.23b}$$

$$= -2 \int_0^{\kappa\tau} \mathrm{d}t_1 \int_0^{\kappa\tau} \mathrm{d}t_2 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \sin(2\theta_c(t_2) - 2\theta_c(t_1)) \, \delta\theta_c(t_2) \\ + 2 \int_0^{\kappa\tau} \mathrm{d}t_1 \int_0^{\kappa\tau} \mathrm{d}t_2 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \sin(2\theta_c(t_2) - 2\theta_c(t_1)) \, \delta\theta_c(t_1) \tag{E.23c}$$

$$= -2 \int_0^{\kappa\tau} \mathrm{d}t_1 \int_0^{\kappa\tau} \mathrm{d}t_2 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \sin(2\theta_c(t_2) - 2\theta_c(t_1)) \, \delta\theta_c(t_2) \\ + 2 \int_0^{\kappa\tau} \mathrm{d}t_1 \int_0^{\kappa\tau} \mathrm{d}t_2 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \sin(2\theta_c(t_1) - 2\theta_c(t_2)) \, \delta\theta_c(t_2) \tag{E.23d}$$

$$= 4 \int_0^{\kappa\tau} \mathrm{d}t_1 \int_0^{\kappa\tau} \mathrm{d}t_2 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \sin(2\theta_c(t_1) - 2\theta_c(t_2)) \, \delta\theta_c(t_2) \tag{E.23e}$$

$$= 4 \int_0^{\kappa\tau} \mathrm{d}t_2 \left[\int_0^{\kappa\tau} \mathrm{d}t_1 \left((t_2 - t_1)^2 + \lambda\right) \cos(\omega_c(t_2 - t_1)) \sin(2\theta_c(t_1) - 2\theta_c(t_2)) \right] \delta\theta_c(t_2). \tag{E.23f}$$

Therefore, in order for the controls $c(t)$ to be a local optimum, we need $F(\omega_c, \kappa\tau) = 0$ and

$$\forall t: \quad \int_0^{\kappa\tau} \mathrm{d}t_1 \left((t - t_1)^2 + \lambda\right) \cos(\omega_c(t - t_1)) \sin(2\theta_c(t_1) - 2\theta_c(t)) = 0. \tag{E.24}$$

The latter condition is *always* satisfied for instantaneous controls, because for instantaneous controls, $2\theta_c(t) \in \pi\mathbb{Z}$ for all t.

Section E.2: Robustness of the superresolution criteria

In this section, we show that there is no measurement basis or initial state such that superresolution can be achieved without $F(\omega_c, \kappa\tau) = 0$. In other words, the superresolution criteria that the filter function vanishes at the centroid is robust.

Suppose we have an instantaneous protocol defined by the switching function $f(t)$, and suppose that the noise is $\lambda(t) = 0$. From Eq. (E.3), the state at time $\kappa\tau$ is

$$|\psi(\kappa\tau)\rangle = \exp[-\mathrm{i}\phi\sigma^z] \, |+\rangle \, , \tag{E.25}$$

where $\phi = \int_0^{\kappa\tau} f(t')\gamma(t')\sigma^z \, \mathrm{d}t'$. We pick a measurement basis defined by $|\theta, \phi\rangle = \cos\theta \, |+\rangle + \mathrm{e}^{\mathrm{i}\phi} \sin\theta \, |-\rangle$ and any vector orthogonal to it. We then consider

$$\langle P_{\theta,\phi}(\kappa\tau)\rangle = \left\langle |\langle\theta, \phi|\psi(\kappa\tau)\rangle|^2 \right\rangle \tag{E.26a}$$

$$= \left\langle \left| \cos\theta\cos\phi - \mathrm{i}\mathrm{e}^{-\mathrm{i}\phi} \sin\theta \sin\phi \right|^2 \right\rangle \tag{E.26b}$$

$$= \cos^2\theta \left\langle \cos^2\phi \right\rangle + \sin^2\theta \left\langle \sin^2\phi \right\rangle - \sin\phi\cos\theta\sin\theta \left\langle \sin(2\phi) \right\rangle \tag{E.26c}$$

$$= \sin^2\theta + \cos(2\theta) \left\langle \cos^2\phi \right\rangle - \sin\phi\cos\theta\sin\theta \left\langle \sin(2\phi) \right\rangle . \tag{E.26d}$$

We proceed as in Eq. (E.5) to find that

$$\langle P_{\theta,\phi}(\kappa\tau) \rangle = \sin^2\theta + \cos(2\theta) \langle P_+(\kappa\tau) \rangle , \tag{E.27}$$

where $P_+(t)$ is, as before, $|\langle +|\psi(t)\rangle|^2$. Repeating the calculation done at Eq. (E.7) and below, we again find that the Fisher information does not vanish as $\Delta\omega \to 0$ only if $F(\omega_c, \kappa\tau) = 0$.

If on the other hand, we had considered $P = |\langle\theta, 0| \exp[-\mathrm{i}\phi\sigma^z] |\theta, 0\rangle|^2$, then we find $\langle P \rangle = \sin^2(2\theta) + \cos^2(2\theta) P_+(\kappa\tau)$. And again we can only achieve superresolution if $F(\omega_c, \kappa\tau) = 0$.

Section E.3: Error analysis of noisy superresolution protocols

In this section, we fix the noise $\lambda(t)$ to have a Lorentzian spectrum as in Eq. (7.15) with strength g_λ and full width at half maximum (FWHM) W. The corresponding contribution of this noise is $\chi_\lambda(t) = \frac{1}{\pi} \int S_\lambda(\omega) F(\omega, t)\, \mathrm{d}\omega$. This section also serves as a rigorous analysis of the noise-free case by setting $g_\lambda = 0$.

We consider the free $\chi_\lambda^{(\mathrm{free})}$ and CPMG $\chi_\lambda^{(\mathrm{CPMG})}$ protocols. Define $\tilde{W} \coloneqq \pi W/\omega_c$. Using a sequence of change of variables, we find that

$$\chi_\lambda(\kappa\tau) = \frac{8\tilde{W} g_\lambda^2}{\omega_c^2} f_{\kappa,\tilde{W}}, \tag{E.28}$$

where

$$f_{\kappa,\tilde{W}}^{(\mathrm{free})} = \int_{-\infty}^{\infty} \frac{\sin^2(\kappa\omega)}{\omega^2(4\omega^2 + \tilde{W}^2)}\, \mathrm{d}\omega \le \frac{\pi\kappa^2}{2\tilde{W}} \tag{E.29a}$$

$$f_{\kappa,\tilde{W}}^{(\mathrm{CPMG})} = 4\int_{-\infty}^{\infty} \frac{\sin^4(\omega/2)\sec^2(\omega)\sin^2(\kappa\omega)}{\omega^2(4\omega^2 + \tilde{W}^2)}\, \mathrm{d}\omega \le f_{\kappa,0}^{(\mathrm{CPMG})}. \tag{E.29b}$$

Note that $f^{(\text{free})}_{\kappa,\tilde{W}}$ goes to ∞ as $\tilde{W} \to 0$, whereas $f^{(\text{CPMG})}_{\kappa,\tilde{W}}$ does not. For example, $f^{(\text{CPMG})}_{2,\tilde{W}} \leq f^{(\text{CPMG})}_{2,0} = \pi/6$, $f^{(\text{CPMG})}_{4,0} = \pi/3$, and $f^{(\text{CPMG})}_{6,0} = \pi/2$.

Recall that $\langle P \rangle = \frac{1}{2} + \frac{1}{2} \mathrm{e}^{-\chi(\kappa\tau) - \chi_\lambda(\kappa\tau)}$. For this analysis, we assume direct access to the probability distribution $\langle P \rangle$ because, as shown around Eq. (7.13), sampling from realizations of P rather than $\langle P \rangle$ introduces negligable error for superresolution protocols. We can easily check that

$$P = 1 - \frac{\alpha \pi^2 g^2 \kappa^2 \Delta\omega^2}{\omega_c^4} - \frac{4 g_\lambda^2 \tilde{W}}{\omega_c^2} f_{\kappa,\tilde{W}} + \mathcal{O}\big(\Delta\omega^4\big)\,. \tag{E.30}$$

where $\alpha^{(\text{free})} = 2$ and $\alpha^{(\text{CPMG})} = 8$.

In our superresolution protocols, we sample N times from the binomial distribution defined by P in order to estimate P, and we estimate $\Delta\omega$ as $\widetilde{\Delta\omega} = \sqrt{\frac{1-P}{\alpha\pi^2 g^2 \kappa^2/\omega_c^4}}$. The resulting bias is

$$b(\Delta\omega) \coloneqq \sqrt{\frac{1-P}{\alpha\pi^2 g^2 \kappa^2/\omega_c^4}} - \Delta\omega \qquad \Longrightarrow \frac{b(\Delta\omega)}{\Delta\omega} = \sqrt{1+\mathcal{W}} - 1 + \mathcal{O}\big(\Delta\omega^2\big)\,, \tag{E.31}$$

where we defined

$$r_{g_\lambda} \coloneqq \sqrt{\frac{1}{\pi}} \cdot \frac{g_\lambda}{g} \cdot \frac{\omega_c}{\Delta\omega}, \qquad \mathcal{W} \coloneqq \frac{4 f_{\kappa,\tilde{W}} \kappa r_{g_\lambda}^2 \tilde{W}}{\pi \alpha \kappa^2}. \tag{E.32}$$

Then, using Chebyshev's inequality,

$$\begin{aligned}
\Pr\Big[\Big|\widetilde{\Delta\omega} - \Delta\omega\Big| > b(\Delta\omega) + \varepsilon\Big] &= \Pr\Big[\Big|\widetilde{\Delta\omega} - \Delta\omega - b(\Delta\omega)\Big| > \varepsilon\Big] && \text{(E.33a)}\\
&\leq \frac{\operatorname{Var} \widetilde{\Delta\omega}}{\varepsilon^2} && \text{(E.33b)}\\
&\leq \frac{P(1-P)/N}{4\alpha\pi^2 g^2 \kappa^2 \varepsilon^2 (1-P)/\omega_c^4} && \text{(E.33c)}\\
&\leq \frac{1}{4N\alpha\pi^2 g^2 \kappa^2 \varepsilon^2/\omega_c^4}, && \text{(E.33d)}
\end{aligned}$$

where we used that the variance of the binomial distribution after N measurements is $P(1-P)/N$ and that $\operatorname{Var} f(P) \leq (f'(\mathbb{E}\,P))^2 \operatorname{Var} P$. Hence, with probability at least $1-p$, our

estimate of $\Delta\omega$ will be within $b(\Delta\omega) + \frac{1}{2\sqrt{p\alpha N}\pi g\kappa/\omega_c^2}$ of $\Delta\omega$.

In Figs. 7.2 and 7.5, in order to compare the different protocols on equal footing, we always use $N = N_{\text{free}} = \frac{\omega_c^4}{8\pi^2 g^2\kappa^2\delta^2\Delta\omega^2}$ measurements. Using this results in, with probability at least $1-p$, the relative error bound

$$\frac{\text{Error}}{\Delta\omega} \leq \frac{b(\Delta\omega)}{\Delta\omega} + \frac{\delta}{\sqrt{p\alpha/2}} = \sqrt{1+\mathcal{W}} - 1 + \frac{\delta}{\sqrt{p\alpha/2}} + \mathcal{O}\big(\Delta\omega^2\big)\,. \tag{E.34}$$

Plugging in the values for $\mathcal{W}$ and α yields

$$\frac{\text{Error}^{(\text{free})}}{\Delta\omega} \leq \sqrt{1+r_{g_\lambda}^2} - 1 + \frac{\delta}{\sqrt{p}} + \mathcal{O}\big(\Delta\omega^2\big) \tag{E.35a}$$

$$\frac{\text{Error}^{(\text{CPMG})}}{\Delta\omega} \leq \sqrt{1+r_{g_\lambda}^2\cdot\frac{W f_{\kappa,0}}{2\kappa^2\omega_c}} - 1 + \frac{\delta}{2\sqrt{p}} + \mathcal{O}\big(\Delta\omega^2\big)\,. \tag{E.35b}$$

To recover the error bounds for the noise-free case, r_{g_λ} and W can be set to 0. We note that the $\mathcal{O}(\Delta\omega^2)$ terms become important for large $\Delta\omega$; the resulting error is due to the Taylor series estimation no longer being a valid approximation.

As $\Delta\omega \to 0$ and when $\kappa = 2$, we see that CPMG outperforms free evolution by a factor of more than $\sqrt{15\omega_c/W}$. This result is very intuitive. In particular, CPMG outperforms free evolution because of the FWHM W, not because of the strength g_λ. The smaller W/ω_c is, the less overlap the CPMG filter function has with the noise. The effect of g_λ on the error shows up in the same way in both CPMG and free evolution. When $\omega_c = 1$ and $W = 0.1$, we see that CPMG should achieve a lower relative error than free evolution by a factor of more than ≈ 12 as $\Delta\omega \to 0$, which we indeed see in the numerical simulations presented in Fig. 7.2.

Subtracting off the bias. Suppose we somehow know (even *exactly*) the noise strength g_λ and FWHM W. Can we then subtract off the noise and regain perfect superresolution even as $\Delta\omega \to 0$? The answer is no, as we show now.

Given that we know the noise, our estimator $\widetilde{\Delta\omega}$ becomes $\sqrt{\frac{1-P-8g_\lambda^2\tilde{W}f_{\kappa,\tilde{W}}/\omega_c^2}{\alpha\pi^2 g^2\kappa^2/\omega_c^4}}$. It follows that the bias $b(\Delta\omega)$ goes to zero as $\Delta\omega \to 0$. Then, proceeding as in Eq. (E.33) and working only to order $\Delta\omega^2$, we have

$$\Pr\left[\left|\widetilde{\Delta\omega} - \Delta\omega\right| > \varepsilon\right] \leq \frac{\operatorname{Var}\widetilde{\Delta\omega}}{\varepsilon^2} \tag{E.36a}$$

$$\leq \frac{P(1-P)/N}{4\alpha\pi^2 g^2\kappa^2\varepsilon^2(1-P-8g_\lambda^2\tilde{W}f_{\kappa,\tilde{W}}/\omega_c^2)/\omega_c^4} \tag{E.36b}$$

$$\leq \frac{\frac{\alpha\pi^2 g^2\kappa^2\Delta\omega^2}{\omega_c^4} + \frac{8g_\lambda^2\tilde{W}}{\omega_c^2}f_{\kappa,\tilde{W}}}{4N\alpha\pi^2 g^2\kappa^2\varepsilon^2(\frac{\alpha\pi^2 g^2\kappa^2\Delta\omega^2}{\omega_c^4})/\omega_c^4}. \tag{E.36c}$$

It follows that in order to estimate $\Delta\omega$ to a relative error of $\delta = \varepsilon/\Delta\omega$ with high probability as $\Delta\omega \to 0$, we need $N \sim \frac{1}{\Delta\omega^4}$.

Thus, we see that even if we knew the noise exactly, subtracting off the noise gets rid of the bias but does not strictly allow us to regain superresolution. This comes down to sampling from the binomial distribution defined by $P = a - b\Delta\omega^2$, as discussed in the main text. When $a \notin \{0, 1\}$, the variance of the binomial distribution prevents superresolution.

Bibliography

[1] J. T. Iosue, L. Norris, and P. Titum, *Superresolution in quantum sensing*, To appear, 2025.

[2] J. T. Iosue, A. Ehrenberg, D. Hangleiter, A. Deshpande, and A. V. Gorshkov, "Page curves and typical entanglement in linear optics", Quantum **7**, 1017 (2023), arXiv:2209.06838 [quant-ph].

[3] J. Youm, J. T. Iosue, A. Ehrenberg, Y.-X. Wang, and A. V. Gorshkov, "Average Rényi entanglement entropy in Gaussian boson sampling", Phys. Rev. Res. **7**, 023125 (2025), arXiv:2403.18890 [quant-ph].

[4] A. Ehrenberg, J. T. Iosue, A. Deshpande, D. Hangleiter, and A. V. Gorshkov, "Transition of anticoncentration in Gaussian boson sampling", Phys. Rev. Lett. **134**, 140601 (2025), arXiv:2312.08433 [quant-ph].

[5] A. Ehrenberg, J. T. Iosue, A. Deshpande, D. Hangleiter, and A. V. Gorshkov, "Second moment of Hafnians in Gaussian boson sampling", Phys. Rev. A **111**, 042412 (2025), arXiv:2403.13878 [quant-ph].

[6] Z. Liang, Y. Xu, J. T. Iosue, and Y.-A. Chen, "Extracting topological orders of generalized Pauli stabilizer codes in two dimensions", PRX Quantum **5**, 030328 (2024), arXiv:2312.11170 [quant-ph].

[7] Z. Liang, B. Yang, J. T. Iosue, and Y.-A. Chen, *Operator algebra and algorithmic construction of boundaries and defects in (2+1)d topological pauli stabilizer codes*, Oct. 2024, arXiv:2410.11942 [quant-ph].

[8] J. Shah, C. Fechisin, Y.-X. Wang, J. T. Iosue, J. D. Watson, Y.-Q. Wang, B. Ware, A. V. Gorshkov, and C.-J. Lin, *Instability of steady-state mixed-state symmetry-protected topological order to strong-to-weak spontaneous symmetry breaking*, Oct. 2024, arXiv:2410.12900 [quant-ph].

[9] J. T. Iosue, K. Sharma, M. J. Gullans, and V. V. Albert, "Continuous-variable quantum state designs: theory and applications", Phys. Rev. X **14**, 011013 (2024), arXiv:2211.05127 [quant-ph].

[10] J. Conrad, J. T. Iosue, A. G. Burchards, and V. V. Albert, *Continuous-variable designs and design-based shadow tomography from random lattices*, Dec. 2024, arXiv:2412.17909 [quant-ph].

[11] J. T. Iosue, T. C. Mooney, A. Ehrenberg, and A. V. Gorshkov, "Projective toric designs, quantum state designs, and mutually unbiased bases", Quantum **8**, 1546 (2024), arXiv:2311.13479 [quant-ph].

[12]S. P. Jain, J. T. Iosue, A. Barg, and V. V. Albert, "Quantum spherical codes", Nature Physics **20**, 10.1038/s41567-024-02496-y (2024), arXiv:2302.11593 [quant-ph].

[13]M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, 10th anniversary (Cambridge University Press, 2010).

[14]P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer", SIAM Journal on Computing **26**, 1484–1509 (1997), arXiv:quant-ph/9508027 [quant-ph].

[15]S. Arora and B. Barak, *Computational complexity: a modern approach* (Cambridge University Press, 2009).

[16]C. H. Bennett, "Logical reversibility of computation", IBM Journal of Research and Development **17**, 525–532 (1973).

[17]T. Toffoli, "Reversible computing", in Automata, languages and programming, edited by J. de Bakker and J. van Leeuwen (1980), pp. 632–644.

[18]R. Jozsa and N. Linden, "On the role of entanglement in quantum-computational speed-up", Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **459**, 2011–2032 (2003).

[19]G. Vidal, "Efficient Classical Simulation of Slightly Entangled Quantum Computations", Physical Review Letters **91**, 147902 (2003).

[20]F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, "Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems", Physical Review Letters **93**, 207204 (2004).

[21]D. Gross, S. T. Flammia, and J. Eisert, "Most Quantum States Are Too Entangled To Be Useful As Computational Resources", Physical Review Letters **102**, 190501 (2009).

[22]M. J. Bremner, C. Mora, and A. Winter, "Are Random Pure States Useful for Quantum Computation?", Physical Review Letters **102**, 190502 (2009).

[23]D. Gottesman, *The Heisenberg representation of quantum computers*, July 1998, arXiv:quant-ph/9807006 [quant-ph].

[24]A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O'Brien, and T. C. Ralph, "Boson sampling from a Gaussian state", Phys. Rev. Lett. **113**, 100502 (2014).

[25]C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, "Gaussian boson sampling", Phys. Rev. Lett. **119**, 170501 (2017).

[26]R. Kruse, C. S. Hamilton, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, "Detailed study of Gaussian boson sampling", Phys. Rev. A **100**, 032326 (2019).

[27]A. Deshpande, A. Mehta, T. Vincent, N. Quesada, M. Hinsche, M. Ioannou, L. Madsen, J. Lavoie, H. Qi, J. Eisert, D. Hangleiter, B. Fefferman, and I. Dhand, "Quantum computational advantage via high-dimensional Gaussian boson sampling", Sci. Adv. **8**, eabi7894 (2022).

[28]D. Grier, D. J. Brod, J. M. Arrazola, M. B. D. A. Alonso, and N. Quesada, "The Complexity of Bipartite Gaussian Boson Sampling", Quantum **6**, 863 (2022).

[29]H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, "Quantum computational advantage using photons", Science **370**, 1460–1463 (2020).

[30]L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, A. E. Lita, T. Gerrits, S. W. Nam, V. D. Vaidya, M. Menotti, I. Dhand, Z. Vernon, N. Quesada, and J. Lavoie, "Quantum computational advantage with a programmable photonic processor", Nature **606**, 75–81 (2022).

[31]H.-S. Zhong, Y.-H. Deng, J. Qin, H. Wang, M.-C. Chen, L.-C. Peng, Y.-H. Luo, D. Wu, S.-Q. Gong, H. Su, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, J. J. Renema, C.-Y. Lu, and J.-W. Pan, "Phase-programmable Gaussian boson sampling using stimulated squeezed light", Phys. Rev. Lett. **127**, 180502 (2021).

[32]M. Wilde, *Quantum information theory*, Second edition, OCLC: ocn973404322 (Cambridge University Press, 2017).

[33]M. Fukuda and R. Koenig, "Typical entanglement for Gaussian states", J. Math. Phys. **60**, 112203 (2019).

[34]A. Serafini, O. C. O. Dahlsten, D. Gross, and M. B. Plenio, "Canonical and micro-canonical typical entanglement of continuous variable systems", J. Phys. A **40**, 9551 (2007).

[35]V. V. Albert, *Bosonic coding: introduction and use cases*, Nov. 2022, arXiv:2211.05714 [quant-ph].

[36]S. G. Hoggar, "T-Designs in Projective Spaces", European Journal of Combinatorics **3**, 233–254 (1982).

[37]S. G. Hoggar, "Parameters of t-Designs in FPd-1", European Journal of Combinatorics **5**, 29–36 (1984).

[38]E. Bannai and S. G. Hoggar, "On tight t-designs in compact symmetric spaces of rank one", Proceedings of the Japan Academy, Series A, Mathematical Sciences **61**, 10.3792/pjaa.61.78 (1985).

[39]W. K. Wootters and B. D. Fields, "Optimal state-determination by mutually unbiased measurements", Annals of Physics **191**, 363–381 (1989).

[40]J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, "Symmetric informationally complete quantum measurements", Journal of Mathematical Physics **45**, 2171–2180 (2004).

[41]A. Klappenecker and M. Rotteler, "Mutually unbiased bases are complex projective 2-designs", in Proceedings. international symposium on information theory, 2005 (2005), pp. 1740–1744, arXiv:quant-ph/0502031 [quant-ph].

[42]C. Dankert, *Efficient simulation of random quantum states and operators*, 2005, arXiv:0512217 [quant-ph].

[43]A. J. Scott, "Tight informationally complete quantum measurements", Journal of Physics A: Mathematical and General **39**, 13507–13530 (2006).

[44]A. Ambainis and J. Emerson, "Quantum t-designs: t-wise independence in the quantum world", in Twenty-second annual ieee conference on computational complexity (ccc'07) (2007), pp. 129–140.

[45]D. A. Roberts and B. Yoshida, "Chaos and complexity by design", Journal of High Energy Physics **2017**, 1–64 (2017).

[46]R. Kueng and D. Gross, "Qubit stabilizer states are complex projective 3-designs", arXiv preprint arXiv:1510.02767 (2015), arXiv:1510.02767.

[47]C. Dankert, R. Cleve, J. Emerson, and E. Livine, "Exact and approximate unitary 2-designs and their application to fidelity estimation", Physical Review A **80**, 012304 (2009).

[48]S. J. van Enk and C. W. J. Beenakker, "Measuring Trρ^n on single copies of ρ using random measurements", Phys. Rev. Lett. **108**, 110503 (2012).

[49]S. Aaronson, "Shadow tomography of quantum states", in Proc. 50th annu. acm sigact symp. theory comput. (June 2018), pp. 325–338, arXiv:1711.01053.

[50]H.-Y. Huang, R. Kueng, and J. Preskill, "Predicting many properties of a quantum system from very few measurements", Nature Physics **16**, 1050–1057 (2020).

[51]H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, *Provably efficient machine learning for quantum many-body problems*, Feb. 2022, arXiv:2106.12627 [quant-ph].

[52]A. Acharya, S. Saha, and A. M. Sengupta, *Informationally complete POVM-based shadow tomography*, May 2021, arXiv:2105.05992.

[53]R. Kueng, H. Zhu, and D. Gross, *Distinguishing quantum states using clifford orbits*, 2016, arXiv:1609.08595 [quant-ph].

[54]J. Emerson, R. Alicki, and K. Życzkowski, "Scalable noise estimation with random unitary operators", Journal of Optics B: Quantum and Semiclassical Optics **7**, S347–S352 (2005).

[55]E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, "Randomized benchmarking of quantum gates", Physical Review A **77**, 012307 (2008).

[56]E. Magesan, J. M. Gambetta, and J. Emerson, "Scalable and robust randomized benchmarking of quantum processes", Phys. Rev. Lett. **106**, 180504 (2011).

[57]A. W. Cross, E. Magesan, L. S. Bishop, J. A. Smolin, and J. M. Gambetta, "Scalable randomised benchmarking of non-clifford gates", npj Quantum Information **2**, 1–5 (2016).

[58]M. A. Nielsen, "The entanglement fidelity and quantum error correction", arXiv, 10.48550/arxiv.quant-ph/9606012 (1996).

[59]M. Horodecki, P. Horodecki, and R. Horodecki, "General teleportation channel, singlet fraction, and quasidistillation", Physical Review A **60**, 1888–1898 (1999).

[60]M. A. Nielsen, "A simple formula for the average gate fidelity of a quantum dynamical operation", Physics Letters A **303**, 249–252 (2002).

[61]E. Magesan, R. Blume-Kohout, and J. Emerson, "Gate fidelity fluctuations and quantum process invariants", Physical Review A **84**, 012309 (2011).

[62]D. Lu, H. Li, D.-A. Trottier, J. Li, A. Brodutch, A. P. Krismanich, A. Ghavami, G. I. Dmitrienko, G. Long, J. Baugh, and R. Laflamme, "Experimental Estimation of Average Fidelity of a Clifford Gate on a 7-Qubit Quantum Processor", Physical Review Letters **114**, 140505 (2015).

[63]S. Bravyi, A. Chowdhury, D. Gosset, and P. Wocjan, "Quantum hamiltonian complexity in thermal equilibrium", Nature Physics **18**, 1367–1370 (2022).

[64]A. Ambainis and A. Smith, "Small pseudo-random families of matrices: derandomizing approximate quantum encryption", in *Approximation, randomization, and combinatorial optimization. algorithms and techniques* (Springer, 2004), pp. 249–260.

[65]P. Hayden, D. Leung, P. W. Shor, and A. Winter, "Randomizing quantum states: constructions and applications", Communications in Mathematical Physics **250**, 371–391 (2004).

[66]S. Kimmel and Y.-K. Liu, "Phase retrieval using unitary 2-designs", in 2017 international conference on sampling theory and applications (sampta) (2017), pp. 345–349.

[67]X. Mi, P. Roushan, C. Quintana, S. Mandra, J. Marshall, C. Neill, F. Arute, K. Arya, J. Atalaya, R. Babbush, J. C. Bardin, R. Barends, A. Bengtsson, S. Boixo, A. Bourassa, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Z. Chen, B. Chiaro, R. Collins, W. Courtney, S. Demura, A. R. Derk, A. Dunsworth, D. Eppens, C. Erickson, E. Farhi, A. G. Fowler, B. Foxen, C. Gidney, M. Giustina, J. A. Gross, M. P. Harrigan, S. D. Harrington, J. Hilton, A. Ho, S. Hong, T. Huang, W. J. Huggins, L. B. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, J. Kelly, S. Kim, A. Kitaev, P. V. Klimov, A. N. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, E. Lucero, O. Martin, J. R. McClean, T. McCourt, M. McEwen, A. Megrant, K. C. Miao, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, M. Newman, M. Y. Niu, T. E. O'Brien, A. Opremcak, E. Ostby, B. Pato, A. Petukhov, N. Redd, N. C. Rubin, D. Sank, K. J. Satzinger, V. Shvarts, D. Strain, M. Szalay, M. D. Trevithick, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, I. Aleiner, K. Kechedzhi, V. Smelyanskiy, and Y. Chen, "Information Scrambling in Computationally Complex Quantum Circuits", Science **374**, 1479–1483 (2021), arXiv:2101.08870 [cond-mat, physics:hep-th, physics:quant-ph].

[68]Y. Sekino and L. Susskind, "Fast scramblers", Journal of High Energy Physics **2008**, 065 (2008).

[69]P. Hayden and J. Preskill, "Black holes as mirrors: quantum information in random subsystems", Journal of high energy physics **2007**, 120 (2007).

[70]J. Czartowski, D. Goyeneche, M. Grassl, and K. Zyczkowski, "Isoentangled mutually unbiased bases, symmetric quantum measurements, and mixed-state designs", Phys. Rev. Lett. **124**, 090503 (2020), arXiv:1906.12291 [quant-ph].

[71]B. C. Hall, *Lie groups, lie algebras, and representations: an elementary introduction*, Second (Springer International Publishing, 2015).

[72]I. Bengtsson and K. Zyczkowski, *Geometry of quantum states: an introduction to quantum entanglement* (Cambridge University Press, Cambridge, 2008).

[73]M. Oszmaniec, R. Augusiak, C. Gogolin, J. Kołody ński, A. Acín, and M. Lewenstein, "Random bosonic states for robust quantum metrology", Phys. Rev. X **6**, 041044 (2016).

[74]A. Sawicki, L. Mattioli, and Z. Zimborás, "Universality verification for a set of quantum gates", Phys. Rev. A **105**, 052602 (2022), arXiv:2111.03862 [quant-ph].

[75]J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices, and groups* (Springer, New York, 1999).

[76]Y. Wu and S. Verdu, "The impact of constellation cardinality on Gaussian channel capacity", in 2010 48th annu. allert. conf. commun. control. comput. (Sept. 2010), pp. 620–628.

[77]F. Lacerda, J. M. Renes, and V. B. Scholz, "Coherent state constellations for Bosonic Gaussian channels", 2016 IEEE Int. Symp. Inf. Theory, 2499–2503 (2016).

[78]F. Lacerda, J. M. Renes, and V. B. Scholz, "Coherent-state constellations and polar codes for thermal Gaussian channels", Phys. Rev. A **95**, 062343 (2017), arXiv:1603.05970.

[79]R. Blume-Kohout and P. S. Turner, "The curious nonexistence of Gaussian 2-designs", Communications in Mathematical Physics **326**, 755–771 (2014).

[80]R. d. l. Madrid, "The role of the rigged Hilbert space in quantum mechanics", European Journal of Physics **26**, 287–312 (2005).

[81]D. Gottesman, A. Yu. Kitaev, and J. Preskill, "Encoding a qubit in an oscillator", Phys. Rev. A **64**, 012310 (2001).

[82]J. Conrad, "Twirling and Hamiltonian engineering via dynamical decoupling for Gottesman-Kitaev-Preskill quantum computing", Phys. Rev. A **103**, 022404 (2021).

[83]J. Conrad, J. Eisert, and F. Arzani, "Gottesman-Kitaev-Preskill codes: A lattice perspective", Quantum **6**, 648 (2022).

[84]A. L. Grimsmo and S. Puri, "Quantum error correction with the Gottesman-Kitaev-Preskill code", PRX Quantum **2**, 020101 (2021).

[85]J. Haferkamp, *On the moments of random quantum circuits and robust quantum complexity*, June 2023, arXiv:2303.16944 [quant-ph].

[86]J. T. Iosue, K. Sharma, M. J. Gullans, and V. V. Albert, "Continuous-variable quantum state designs: theory and applications", Physical Review X **14**, 10.1103/physrevx.14.011013 (2024).

[87]A. Roy and A. J. Scott, "Unitary designs and codes", Designs, Codes and Cryptography **53**, 13–31 (2009), arXiv:0809.3813 [math.CO].

[88]T. Tao and V. Vu, *Additive combinatorics* (Cambridge University Press, Cambridge, 2006).

[89]J. Singer, "A theorem in finite projective geometry and some applications to number theory", Transactions of the American Mathematical Society **43**, 377–85 (1938).

[90]M. Saniga, M. Planat, and H. Rosu, "Mutually unbiased bases and finite projective planes", Journal of Optics B: Quantum and Semiclassical Optics **6**, L19–L20 (2004).

[91]M. Planat, H. C. Rosu, and S. Perrine, "A survey of finite algebraic geometrical structures underlying mutually unbiased quantum measurements", Foundations of Physics **36**, 1662–1680 (2006), arXiv:quant-ph/0409081 [quant-ph].

[92]A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum error correction and orthogonal geometry", Phys. Rev. Lett. **78**, 405–408 (1997), arXiv:quant-ph/9605005 [quant-ph].

[93]D. Gottesman, "Class of quantum error-correcting codes saturating the quantum Hamming bound", Phys. Rev. A **54**, 1862–1868 (1996), arXiv:quant-ph/9604038 [quant-ph].

[94]C. Spengler, M. Huber, S. Brierley, T. Adaktylos, and B. C. Hiesmayr, "Entanglement detection via mutually unbiased bases", Phys. Rev. A **86**, 022311 (2012), arXiv:1202.5058 [quant-ph].

[95]H. Zhu, "Mutually unbiased bases as minimal Clifford covariant 2-designs", Phys. Rev. A **91**, 060301 (2015), arXiv:1505.01123 [quant-ph].

[96]C. L. Degen, F. Reinhard, and P. Cappellaro, "Quantum sensing", Rev. Mod. Phys. **89**, 035002 (2017), arXiv:1611.02427 [quant-ph].

[97]M. Born, E. Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*, 7th ed. (Cambridge University Press, 1999).

[98]M. Tsang, R. Nair, and X.-M. Lu, "Quantum theory of superresolution for two incoherent optical point sources", Phys. Rev. X **6**, 031033 (2016), arXiv:1511.00552 [quant-ph].

[99]D. N. Page, "Average entropy of a subsystem", Phys. Rev. Lett. **71**, 1291–1294 (1993).

[100]S. K. Foong and S. Kanno, "Proof of Page's conjecture on the average entropy of a subsystem", Phys. Rev. Lett. **72**, 1148–1151 (1994).

[101]J. Sánchez-Ruiz, "Simple proof of Page's conjecture on the average entropy of a subsystem", Phys. Rev. E **52**, 5653–5655 (1995).

[102]S. Sen, "Average Entropy of a Quantum Subsystem", Physical Review Letters **77**, 1–3 (1996).

[103]D. N. Page, "Information in black hole radiation", Physical Review Letters **71**, 3743–3746 (1993).

[104]E. Bianchi, T. De Lorenzo, and M. Smerlak, "Entanglement entropy production in gravitational collapse: covariant regularization and solvable models", Journal of High Energy Physics **2015**, 180 (2015).

[105]P. Hayden, D. W. Leung, and A. Winter, "Aspects of generic entanglement", Commun. Math. Phys. **265**, 95–117 (2006).

[106]P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, "Chaos in quantum channels", Journal of High Energy Physics **2016**, 4 (2016).

[107]H. Fujita, Y. O. Nakagawa, S. Sugiura, and M. Watanabe, "Page curves for general interacting systems", Journal of High Energy Physics **2018**, 112 (2018).

[108]T.-C. Lu and T. Grover, "Renyi entropy of chaotic eigenstates", Physical Review E **99**, 032111 (2019).

[109]Y. O. Nakagawa, M. Watanabe, H. Fujita, and S. Sugiura, "Universality in volume-law entanglement of scrambled pure quantum states", Nature Communications **9**, 1635 (2018).

[110]L. Vidmar and M. Rigol, "Entanglement entropy of eigenstates of quantum chaotic hamiltonians", Physical Review Letters **119**, 220603 (2017).

[111]L. Vidmar, L. Hackl, E. Bianchi, and M. Rigol, "Entanglement entropy of eigenstates of quadratic fermionic hamiltonians", Physical Review Letters **119**, 020601 (2017).

[112]L. Hackl, L. Vidmar, M. Rigol, and E. Bianchi, "Average eigenstate entanglement entropy of the XY chain in a transverse field and its universality for translationally invariant quadratic fermionic models", Physical Review B **99**, 075123 (2019).

[113]S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, "Canonical typicality", Physical Review Letters **96**, 050403 (2006).

[114]L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics", Advances in Physics **65**, 239–362 (2016).

[115]B. Collins, C. E. Gonzalez-Guillen, and D. Pérez-Garcia, "Matrix Product States, Random Matrix Theory and the Principle of Maximum Entropy", Communications in Mathematical Physics **320**, 663–677 (2013).

[116]M. B. Hastings, "Random MERA States and the Tightness of the Brandao-Horodecki Entropy Bound", arXiv.1505.06468, 10.48550/arXiv.1505.06468 (2015).

[117]S. Garnerone, T. R. de Oliveira, and P. Zanardi, "Typicality in random matrix product states", Physical Review A **81**, 032336 (2010).

[118]S. Popescu, A. J. Short, and A. Winter, "Entanglement and the foundations of statistical mechanics", Nat. Phys. **2**, 754–758 (2006).

[119]E. Bianchi and P. Donà, "Typical entanglement entropy in the presence of a center: Page curve and its variance", Physical Review D **100**, 105010 (2019).

[120]E. Bianchi, L. Hackl, and M. Kieburg, "Page curve for fermionic Gaussian states", Phys. Rev. B **103**, 10.1103/physrevb.103.1241118 (2021).

[121]O. C. O. Dahlsten, C. Lupo, S. Mancini, and A. Serafini, "Entanglement typicality", J. Phys. A **47**, 363001 (2014).

[122]L. Amico, R. Fazio, A. Osterloh, and V. Vedral, "Entanglement in many-body systems", Reviews of Modern Physics **80**, 517–576 (2008).

[123]R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement", Rev. Mod. Phys. **81**, 865–942 (2009).

[124]U. Chabaud and M. Walschaers, "Resources for Bosonic Quantum Computational Advantage", Physical Review Letters **130**, 090602 (2023).

[125]Q. Zhuang, Z. Zhang, and J. H. Shapiro, "Distributed quantum sensing using continuous-variable multipartite entanglement", Physical Review A **97**, 032329 (2018).

[126]E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, "Photonic quantum metrology", AVS Quantum Science **2**, 024703 (2020).

[127]C. Oh, C. Lee, S. H. Lie, and H. Jeong, "Optimal distributed quantum sensing using Gaussian states", Physical Review Research **2**, 023030 (2020).

[128]M. Malitesta, A. Smerzi, and L. Pezzè, "Distributed Quantum Sensing with Squeezed-Vacuum Light in a Configurable Network of Mach-Zehnder Interferometers", arXiv:2109.09178, 10.48550/ARXIV.2109.09178 (2021).

[129]M. Barbieri, "Optical Quantum Metrology", PRX Quantum **3**, 010202 (2022).

[130]G. Adesso, *Entanglement of Gaussian states*, 2007, arXiv:quant-ph/0702069 [quant-ph].

[131]B. Bhattacharjee, P. Nandy, and T. Pathak, "Eigenstate capacity and Page curve in fermionic Gaussian states", Phys. Rev. B **104**, 214306 (2021), arXiv:2109.00557 [quant-ph].

[132]A. Serafini, O. C. O. Dahlsten, and M. B. Plenio, "Teleportation fidelities of squeezed states from thermodynamical state space measures", Phys. Rev. Lett. **98**, 170501 (2007).

[133]G. Adesso, D. Girolami, and A. Serafini, "Measuring Gaussian Quantum Information and Correlations Using the Rényi Entropy of Order 2", Physical Review Letters **109**, 190502 (2012).

[134]G. Camilo, G. T. Landi, and S. Eliëns, "Strong subadditivity of the Rényi entropies for bosonic and fermionic Gaussian states", Physical Review B **99**, 045155 (2019).

[135]V. Bužek, C. H. Keitel, and P. L. Knight, "Sampling entropies and operational phase-space measurement. I. General formalism", Physical Review A **51**, 2575–2593 (1995).

[136]G. Adesso and R. Simon, "Strong subadditivity for log-determinant of covariance matrices and its applications", Journal of Physics A: Mathematical and Theoretical **49**, 34LT02 (2016).

[137]L. Lami, C. Hirche, G. Adesso, and A. Winter, "Schur Complement Inequalities for Covariance Matrices and Monogamy of Quantum Correlations", Physical Review Letters **117**, 220502 (2016).

[138]A. Serafini, *Quantum continuous variables* (CRC Press, July 2017).

[139]F. C. Khanna, J. M. C. Malbouisson, A. E. Santana, and E. S. Santos, "Maximum entanglement in squeezed boson and fermion states", Physical Review A **76**, 022109 (2007).

[140]S. Stanisic, N. Linden, A. Montanaro, and P. S. Turner, "Generating entanglement with linear optics", Physical Review A **96**, 043861 (2017).

[141]M. Petkovšek, H. S. Wilf, and D. Zeilberger, $A{=}B$ (A K Peters, 1996).

[142]W. Bailey, *Generalized Hypergeometric Series, By W.N. Bailey*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 32 (Camrbridge University Press, 1964).

[143]W. Zudilin, "Hypergeometric heritage of W. N. Bailey", Notices of the International Congress of Chinese Mathematicians **7**, 32–46 (2019).

[144]L. J. Slater, *Generalized hypergeometric functions* (Cambridge Univ. Press, Cambridge, 1966).

[145]Joseph T. Iosue, *GLO*, GitHub, https://github.com/jtiosue/GLO, 2022.

[146]D. Weingarten, "Asymptotic behavior of group integrals in the limit of infinite rank", Journal of Mathematical Physics **19**, 999–1001 (1978).

[147]B. Collins, "Moments and Cumulants of Polynomial random variables on unitary groups, the Itzykson-Zuber integral and free probability", arXiv.math-ph/0205010, 10.48550/arXiv.math-ph/0205010 (2002).

[148]N. Alexeev, A. Pologova, and M. A. Alekseyev, "Generalized Hultman Numbers and Cycle Structures of Breakpoint Graphs", Journal of Computational Biology **24**, 93–105 (2017).

[149]M. Alekseyev, A. Ehrenberg, J. T. Iosue, and A. V. Gorshkov, *Calculating entanglement in linear optics via breakpoint graphs*, In preparation.

[150]I. Kim, "Rényi-α entropies of quantum states in closed form: Gaussian states and a class of non-Gaussian states", Physical Review E **97**, 062141 (2018).

[151]Mark Wilde and K. Sharma, *PHYS 7895: Gaussian Quantum Information, Lecture 10*, https://markwilde.com/teaching/2019-spring-gqi/scribe-notes/lecture-10-scribed.pdf, 2019.

[152]L. Hackl and E. Bianchi, "Bosonic and fermionic Gaussian states from Kähler structures", SciPost Phys. Core **4**, 025 (2021).

[153]Q. Zhuang, T. Schuster, B. Yoshida, and N. Y. Yao, "Scrambling and complexity in phase space", Phys. Rev. A **99**, 062334 (2019).

[154]T. Zhou and X. Chen, "Non-unitary Entanglement Dynamics in Continuous Variable Systems", Physical Review B **104**, L180301 (2021), arXiv: 2103.06507.

[155]B. Zhang and Q. Zhuang, "Entanglement formation in continuous-variable random quantum networks", npj Quantum Inf. **7**, 33 (2021).

[156]A. Deshpande, B. Fefferman, M. C. Tran, M. Foss-Feig, and A. V. Gorshkov, "Dynamical Phase Transitions in Sampling Complexity", Physical Review Letters **121**, 030501 (2018).

[157]G. Muraleedharan, A. Miyake, and I. H. Deutsch, "Quantum computational supremacy in the sampling of bosonic random walkers on a one-dimensional lattice", New Journal of Physics **21**, 055003 (2019).

[158]C. Oh, Y. Lim, B. Fefferman, and L. Jiang, "Classical simulation of bosonic linear-optical random circuits beyond linear light cone", arXiv:2102.10083, 10.48550/arXiv.2102.10083 (2021).

[159]N. Maskara, A. Deshpande, A. Ehrenberg, M. C. Tran, B. Fefferman, and A. V. Gorshkov, "Complexity Phase Diagram for Interacting and Long-Range Bosonic Hamiltonians", Physical Review Letters **129**, 150604 (2022).

[160]C. Oh, Y. Lim, B. Fefferman, and L. Jiang, "Classical Simulation of Boson Sampling Based on Graph Structure", Physical Review Letters **128**, 190501 (2022).

[161]M. Möller and C. Vuik, "On the impact of quantum computing technology on future developments in high-performance scientific computing", Ethics and Information Technology **19**, 253–269 (2017), arXiv:1705.07413 [cs.CE].

[162]B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, "Quantum algorithms for quantum chemistry and quantum materials science", Chem. Rev. **120**, PMID: 33090772, 12685–12717 (2020).

[163]J. F. F. Bulmer, B. A. Bell, R. S. Chadwick, A. E. Jones, D. Moise, A. Rigazzi, J. Thorbecke, U.-U. Haus, T. V. Vaerenbergh, R. B. Patel, I. A. Walmsley, and A. Laing, "The boundary for quantum advantage in Gaussian boson sampling", Sci. Adv. **8**, eabl9236 (2022).

[164]J. Preskill, *Quantum computing and the entanglement frontier*, 2012, arXiv:1203.5813 [quant-ph].

[165]Y. Yu, "Advancements in applications of quantum entanglement", J. Phys. Conf. Ser. **2012**, 012113 (2021).

[166]W. K. Wootters and W. S. Leng, "Quantum entanglement as a quantifiable resource [and discussion]", Philosophical Transactions: Mathematical, Physical and Engineering Sciences **356**, 1717–1731 (1998).

[167]N. Zou, "Quantum entanglement and its application in quantum communication", J. Phys. Conf. Ser. **1827**, 012120 (2021).

[168]A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, "The entropy of Hawking radiation", Rev. Mod. Phys. **93**, 035002 (2021), arXiv:2006.06872 [hep-th].

[169]E. Bianchi, L. Hackl, M. Kieburg, M. Rigol, and L. Vidmar, "Volume-law entanglement entropy of typical pure quantum states", PRX Quantum **3**, 030201 (2022), arXiv:2112.06959 [quant-ph].

[170]J. Bae, M. Tiersch, S. Sauer, F. de Melo, F. Mintert, B. Hiesmayr, and A. Buchleitner, "Detection and typicality of bound entangled states", Phys. Rev. A **80**, 022317 (2009).

[171]H.-H. Lai and K. Yang, "Entanglement entropy scaling laws and eigenstate typicality in free fermion systems", Phys. Rev. B **91**, 081110 (2015).

[172]P. Łydżba, M. Rigol, and L. Vidmar, "Eigenstate entanglement entropy in random quadratic hamiltonians", Phys. Rev. Lett. **125**, 180604 (2020), arXiv:2006.11302 [cond-mat.stat-mech].

[173]P. Łydżba, M. Rigol, and L. Vidmar, "Entanglement in many-body eigenstates of quantum-chaotic quadratic hamiltonians", Phys. Rev. B **103**, 104206 (2021), arXiv:2101.05309 [cond-mat.stat-mech].

[174]I. Ul čakar and L. Vidmar, "Tight-binding billiards", Phys. Rev. E **106**, 034118 (2022), arXiv:2206.07078 [cond-mat.stat-mech].

[175]X.-H. Yu, Z. Gong, and J. I. Cirac, "Free-fermion Page curve: canonical typicality and dynamical emergence", Phys. Rev. Res. **5**, 013044 (2023).

[176]S. Murciano, P. Calabrese, and L. Piroli, "Symmetry-resolved Page curves", Phys. Rev. D **106**, 046015 (2022).

[177]R. Konig and R. Renner, "Sampling of min-entropy relative to quantum knowledge", IEEE Trans. Inf. Theory **57**, 4760–4787 (2011).

[178]M. M. Wilde and K. Sharma, *Phys 7895: gaussian quantum information lecture 10*, https://markwilde.com/teaching/2019-spring-gqi/scribe-notes/lecture-10-scribed.pdf, 2019.

[179]Y. Yauk, R. Patil, Y. Zhang, M. Rigol, and L. Hackl, "Typical entanglement entropy in systems with particle-number conservation", Phys. Rev. B **110**, 10.1103/PhysRevB.110.235154 (2024), arXiv:2310.19862 [quant-ph].

[180]D. J. Brod, "Complexity of simulating constant-depth BosonSampling", Phys. Rev. A **91**, 10.1103/physreva.91.042316 (2015).

[181]D. Hangleiter and J. Eisert, "Computational advantage of quantum random sampling", Rev. Mod. Phys. **95**, 10.1103/revmodphys.95.035001 (2023).

[182]C. F. Gauss, "Methodus nova integralium valores per approximationem inveniendi", in *Werke* (Cambridge University Press, Nov. 1866), pp. 165–196.

[183]P. Delsarte, J. M. Goethals, and J. J. Seidel, "Spherical codes and designs", Geometriae Dedicata **6**, 363–388 (1977).

[184]R. H. Hardin and N. J. A. Sloane, "McLaren's improved snub cube and other new spherical designs in three dimensions", Discrete & Computational Geometry **15**, 429–441 (1996).

[185]A. H. Stroud, *Approximate calculation of multiple integrals* (Prentice-Hall, 1971).

[186]R. Cools, "Constructing cubature formulae: the science behind the art", Acta Numerica **6**, 1–54 (1997).

[187]P. C. Hammer and A. H. Stroud, "Numerical integration over simplexes", Mathematical tables and other aids to computation **10**, 137–139 (1956).

[188]M. S. Baladram, "On explicit construction of simplex t-designs", Interdisciplinary Information Sciences **24**, 181–184 (2018).

[189]G. Kuperberg, "Numerical cubature from Archimedes' hat-box theorem", SIAM Journal on Numerical Analysis **44**, 908–935 (2006), arXiv:math/0405366 [math.NA].

[190]G. Kuperberg, "Numerical cubature using error-correcting codes", arXiv:math/0402047 (2004).

[191]N. Victoir, "Asymmetric cubature formulae with few points in high dimension for symmetric measures", SIAM Journal on Numerical Analysis **42**, 209–227 (2004).

[192]P. D. Seymour and T. Zaslavsky, "Averaging sets: a generalization of mean values and spherical designs", Advances in Mathematics **52**, 213–240 (1984).

[193]C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information", Rev. Mod. Phys. **84**, 621–669 (2012).

[194]Z. Webb, *The Clifford group forms a unitary 3-design*, Nov. 2016, arXiv:1510.02769 [quant-ph].

[195]H. Zhu, R. Kueng, M. Grassl, and D. Gross, *The Clifford group fails gracefully to be a unitary 4-design*, Sept. 2016, arXiv:1609.08172 [quant-ph].

[196]H. Zhu, "Multiqubit Clifford groups are unitary 3-designs", Physical Review A **96**, 062336 (2017).

[197]M. A. Graydon, J. Skanes-Norman, and J. J. Wallman, *Clifford groups are not always 2-designs*, Aug. 2021, arXiv:2108.04200 [quant-ph].

[198]L. Susskind and J. Glogower, "Quantum mechanical phase and time operator", Physics Physique Fizika **1**, 49–61 (1964).

[199]P. Carruthers and M. M. Nieto, "Phase and Angle Variables in Quantum Mechanics", Reviews of Modern Physics **40**, 411–440 (1968).

[200]C. W. Helstrom, "Quantum detection and estimation theory", Journal of Statistical Physics **1**, 231–252 (1969).

[201]A. S. Holevo, "Covariant measurements and imprimitivity systems", in *Quantum Probability and Applications to the Quantum Theory of Irreversible Processes*, Vol. 1055, edited by L. Accardi, A. Frigerio, and V. Gorini (Springer Berlin Heidelberg, Berlin, Heidelberg, 1984), pp. 153–172.

[202]J. Bergou and B.-G. Englert, "Operators of the phase. Fundamentals", Annals of Physics **209**, 479–505 (1991).

[203]P. M. Mathews and K. Eswaran, "Simultaneous uncertainties of the cosine and sine operators", Il Nuovo Cimento B (1971-1996) **19**, 99–104 (1974).

[204]J. H. Shapiro and S. R. Shepard, "Quantum phase measurement: A system-theory perspective", Physical Review A **43**, 3795–3818 (1991).

[205]A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Edizioni della Normale, Pisa, 2011).

[206]S. M. Girvin, "Circuit qed: superconducting qubits coupled to microwave photons", Quantum machines: measurement and control of engineered quantum systems, 113–256 (2014).

[207]S. Gandhari, V. V. Albert, T. Gerrits, J. M. Taylor, and M. J. Gullans, *Continuous-Variable Shadow Tomography*, Nov. 2022, arXiv:2211.05149 [quant-ph].

[208]H. M. Wiseman, "Adaptive phase measurements of optical modes: going beyond the marginal Q distribution", Phys. Rev. Lett. **75**, 4587–4590 (1995).

[209]L. S. Martin, W. P. Livingston, S. Hacohen-Gourgy, H. M. Wiseman, and I. Siddiqi, "Implementation of a canonical phase measurement with quantum feedback", Nature Physics **16**, 1046–1049 (2020).

[210]D. I. Schuster, A. A. Houck, J. A. Schreier, A. Wallraff, J. M. Gambetta, A. Blais, L. Frunzio, J. Majer, B. Johnson, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Resolving photon number states in a superconducting circuit", Nature **445**, 515–518 (2007).

[211]E. T. Holland, B. Vlastakis, R. W. Heeres, M. J. Reagor, U. Vool, Z. Leghtas, L. Frunzio, G. Kirchmair, M. H. Devoret, M. Mirrahimi, and R. J. Schoelkopf, "Single-Photon-Resolved Cross-Kerr Interaction for Autonomous Stabilization of Photon-Number States", Phys. Rev. Lett. **115**, 180501 (2015).

[212]M. Elliott, J. Joo, and E. Ginossar, "Designing Kerr interactions using multiple superconducting qubit types in a single circuit", New Journal of Physics **20**, 023037 (2018).

[213]Y. Zhang, J. C. Curtis, C. S. Wang, R. J. Schoelkopf, and S. M. Girvin, "Drive-induced nonlinearities of cavity modes coupled to a transmon ancilla", Physical Review A **105**, 022423 (2022).

[214]A. J. Scott, "Optimizing quantum process tomography with unitary 2-designs", Journal of Physics A: Mathematical and Theoretical **41**, 10.1088/1751-8113/41/5/055308 (2008), arXiv:0711.1017 [quant-ph].

[215]D. Gottesman, A. Kitaev, and J. Preskill, "Encoding a qubit in an oscillator", Phys. Rev. A **64**, 012310 (2001).

[216]V. V. Albert, J. P. Covey, and J. Preskill, "Robust Encoding of a Qubit in a Molecule", Physical Review X **10**, 031050 (2020), arXiv:1911.00099.

[217]N. C. Menicucci, "Fault-Tolerant Measurement-Based Quantum Computing with Continuous-Variable Cluster States", Phys. Rev. Lett. **112**, 120504 (2014).

[218]A. Winter, "Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints", Communications in Mathematical Physics **347**, 291–313 (2016).

[219]K. Sharma, M. M. Wilde, S. Adhikari, and M. Takeoka, "Bounding the energy-constrained quantum and private capacities of phase-insensitive bosonic gaussian channels", New Journal of Physics **20**, 063025 (2018).

[220]A. Winter, *Energy-constrained diamond norm with applications to the uniform continuity of continuous variable channel capacities*, 2017, arXiv:1712.10267 [quant-ph].

[221]M. E. Shirokov, "On the energy-constrained diamond norm and its application in quantum information theory", Problems of Information Transmission **54**, 20–33 (2018).

[222]E. Culf, T. Vidick, and V. V. Albert, *Group coset monogamy games and an application to device-independent continuous-variable qkd*, 2022, arXiv:2212.03935 [quant-ph].

[223]S. L. Braunstein, C. A. Fuchs, and H. J. Kimble, "Criteria for continuous-variable quantum teleportation", Journal of Modern Optics **47**, 267–278 (2000).

[224]A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik, "Unconditional quantum teleportation", Science **282**, 706–709 (1998).

[225]R. Namiki, M. Koashi, and N. Imoto, "Fidelity criterion for quantum-domain transmission and storage of coherent states beyond the unit-gain constraint", Phys. Rev. Lett. **101**, 100502 (2008).

[226]A. I. Lvovsky and M. G. Raymer, "Continuous-variable optical quantum-state tomography", Rev. Mod. Phys. **81**, 299–332 (2009).

[227]G. Chiribella and J. Xie, "Optimal design and quantum benchmarks for coherent state amplifiers", Phys. Rev. Lett. **110**, 213602 (2013).

[228]G. Bai and G. Chiribella, "Test one to test many: a unified approach to quantum benchmarks", Phys. Rev. Lett. **120**, 150502 (2018).

[229]Y.-D. Wu and B. C. Sanders, "Efficient verification of bosonic quantum channels via benchmarking", New Journal of Physics **21**, 073026 (2019).

[230]R. M. S. Farias and L. Aolita, "Certification of continuous-variable gates using average channel-fidelity witnesses", Quantum Science and Technology **6**, 035014 (2021).

[231]K. Sharma, B. C. Sanders, and M. M. Wilde, "Optimal tests for continuous-variable quantum teleportation and photodetectors", Physical Review Research **4**, 023066 (2022).

[232]K. Sharma and M. M. Wilde, "Characterizing the performance of continuous-variable Gaussian quantum gates", Physical Review Research **2**, 013126 (2020).

[233]S. Becker and N. Datta, "Convergence rates for quantum evolution and entropic continuity bounds in infinite dimensions", Communications in Mathematical Physics **374**, 823–871 (2020).

[234]S. Becker, N. Datta, L. Lami, and C. Rouzé, "Energy-constrained discrimination of unitaries, quantum speed limits, and a Gaussian Solovay-Kitaev theorem", Phys. Rev. Lett. **126**, 190504 (2021).

[235]L. Lami, "Quantum data hiding with continuous-variable systems", Phys. Rev. A **104**, 052428 (2021).

[236]H. K. Mishra, S. K. Oskouei, and M. M. Wilde, "Optimal input states for quantifying the performance of continuous-variable unidirectional and bidirectional teleportation", Phys. Rev. A **107**, 062603 (2023).

[237]A. S. Holevo, *Quantum systems, channels, information*, de Gruyter Studies in Mathematical Physics (Book 16) (de Gruyter, Nov. 2012), p. 349.

[238]T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski, "On mutually unbiased bases", International Journal of Quantum Information **08**, 535–640 (2010).

[239]M. Nakahara, *Geometry, topology and physics* (CRC press, 2018).

[240]A. W. Harrow, "The Church of the Symmetric Subspace", arXiv:1308.6595 [quant-ph] (2013).

[241]B. R. Hunt, T. Sauer, and J. A. Yorke, "Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces", Bulletin of the American Mathematical Society **27**, 217–238 (1992).

[242]R. Grigorchuk and P. de la Harpe, *Amenability and ergodic properties of topological groups: from Bogolyubov onwards*, July 2015, arXiv:1404.7030 [math].

[243]J. Becnel and A. Sengupta, "The Schwartz Space: Tools for Quantum Mechanics and Infinite Dimensional Analysis", Mathematics **3**, 527–562 (2015).

[244]W. Rudin, *Functional analysis*, 2nd ed, International series in pure and applied mathematics (McGraw-Hill, New York, 1991).

[245]F. Gieres, "Mathematical surprises and Dirac's formalism in quantum mechanics", Reports on Progress in Physics **63**, 1893–1931 (2000).

[246]V. V. Dodonov, "'Nonclassical' states in quantum optics: a 'squeezed' review of the first 75 years", Journal of Optics B: Quantum and Semiclassical Optics **4**, R1–R33 (2002).

[247]J. Hou and X. Qi, "Constructing entanglement witnesses for infinite-dimensional systems", Physical Review A **81**, 062351 (2010).

[248]V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, "Performance and structure of single-mode bosonic codes", Phys. Rev. A **97**, 032346 (2018).

[249]J. S. Ivan, K. K. Sabapathy, and R. Simon, "Operator-sum representation for bosonic Gaussian channels", Physical Review A **84**, 042311 (2011).

[250]M. G. Paris, "Displacement operator by beam splitter", Physics Letters A **217**, 78–80 (1996).

[251]R. Nair, "Quantum-limited loss sensing: multiparameter estimation and bures distance between loss channels", Phys. Rev. Lett. **121**, 230801 (2018).

[252]R. H. Cameron, "A "Simpson's rule" for the numerical evaluation of Wiener's integrals in function space", Duke Mathematical Journal **18**, 10.1215/S0012-7094-51-01810-8 (1951).

[253]I. M. Gel'fand and A. M. Yaglom, "Integration in Functional Spaces and its Applications in Quantum Physics", Journal of Mathematical Physics **1**, 48–69 (1960).

[254]S. G. Brush, "Functional Integrals and Statistical Physics", Reviews of Modern Physics **33**, 79–92 (1961).

[255]A. G. Konheim and W. L. Miranker, "Numerical evaluation of Wiener integrals", Mathematics of Computation **21**, 49–65 (1967).

[256]V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, "Performance and structure of single-mode bosonic codes", Phys. Rev. A **97**, 032346 (2018).

[257]V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, "Real-time quantum error correction beyond break-even", Nature **616**, 50–55 (2023).

[258]P. Sarnak and P. Buser, "On the period matrix of a Riemann surface of large genus (with an Appendix by J. H. Conway and N. J. A. Sloane)", Inv. Math. **117**, 27–56 (1994).

[259]J. Conrad, J. Eisert, and J.-P. Seifert, "Good Gottesman-Kitaev-Preskill codes from the NTRU cryptosystem", Quantum **8**, 1398 (2024).

[260]M. Moskowitz and R. Sacksteder, "An extension of a theorem by Hlawka", Mathematika **56**, 203–216 (2010).

[261]L. R. Welch, "Lower bounds on the maximum cross correlation of signals", IEEE Transactions on Information Theory **20**, 397–399 (1974).

[262]S. Datta, S. Howard, and D. Cochran, *Geometry of the Welch bounds*, Sept. 2009, arXiv:0909.0206 [cs.IT].

[263]A. Klappenecker and M. Rötteler, "Constructions of mutually unbiased bases", in Finite fields and applications, edited by G. L. Mullen, A. Poli, and H. Stichtenoth (2004), pp. 137–144.

[264]A. Belovs and J. Smotrovs, *A criterion for attaining the Welch bounds with applications for mutually unbiased bases*, Feb. 2008, arXiv:0802.0855 [quant-ph].

[265]K. E. Cahill and R. J. Glauber, "Ordered expansions in boson amplitude operators", Phys. Rev. **177**, 1857–1881 (1969).

[266]C. Gerry and P. Knight, *Introductory quantum optics* (Cambridge University Press, Oct. 2004).

[267]D. Kelmer and S. Yu, "The second moment of the siegel transform in the space of symplectic lattices", International Mathematics Research Notices **2021**, 5825–5859 (2019).

[268]M. Beckers and R. Cools, "A relation between cubature formulae of trigonometric degree and lattice rules", in *Numerical integration iv: proceedings of the conference at the mathematical research institute, oberwolfach, november 8–14, 1992*, edited by H. Brass and G. Hämmerlin (Birkhäuser Basel, Basel, 1993), pp. 13–24.

[269]R. Cools and I. H. Sloan, "Minimal cubature formulae of trigonometric degree", Mathematics of Computation **65**, 1583–1600 (1996).

[270]K. O'Bryant, "A Complete Annotated Bibliography of Work Related to Sidon Sequences", The Electronic Journal of Combinatorics, 10.37236/32 (2004).

[271]R. C. Bose and S. Chowla, "Theorems in the additive theory of numbers", Commentarii Mathematici Helvetici **37**, 141–147 (1962).

[272]J. H. Conway and N. J. A. Sloane, "Low-dimensional lattices. VII. Coordination sequences", Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **453**, 2369–2389 (1997).

[273]O. F. Inc., *The On-Line Encyclopedia of Integer Sequences*, https://oeis.org/A108625.

[274]B. G. Bodmann and J. Haas, *Achieving the orthoplex bound and constructing weighted complex projective 2-designs with Singer sets*, Sept. 2015, arXiv:1509.05333 [math-ph].

[275]V. G. Avella, J. Czartowski, D. Goyeneche, and K. Zyczkowski, *Cyclic measurements and simplified quantum state tomography*, Apr. 2024, arXiv:2404.18847 [quant-ph].

[276]D. L. Cohn, *Measure theory*, Second edition, Birkhäuser Advanced Texts (Birkhäuser, Boston, 2013).

[277]S. Saeki, "A Proof of the Existence of Infinite Product Probability Measures", The American Mathematical Monthly **103**, 682–683 (1996).

[278]Joseph T. Iosue, *ToricDesigns*, GitHub, https://github.com/jtiosue/ToricDesigns, 2024.

[279]M. Matolcsi and M. Weiner, "A rigidity property of complete systems of mutually unbiased bases", Open Systems & Information Dynamics **28**, 2150012 (2021), arXiv:2112.00090 [quant-ph].

[280]S. Eberhard and F. Manners, "The apparent structure of dense Sidon sets", The Electronic Journal of Combinatorics **30**, 10.37236/11191 (2023).

[281]W. K. Wootters, "Quantum measurements and finite geometry", 10.48550/arXiv.quant-ph/0406032 (2004), arXiv:quant-ph/0406032.

[282]N. Aslam, M. Pfender, P. Neumann, R. Reuter, A. Zappe, F. F. de Oliveira, A. Denisenko, H. Sumiya, S. Onoda, J. Isoya, and J. Wrachtrup, "Nanoscale nuclear magnetic resonance with chemical resolution", Science **357**, 67–71 (2017).

[283]S. Magaletti, L. Mayer, J.-F. Roch, and T. Debuisschert, "A quantum radio frequency signal analyzer based on nitrogen vacancy centers in diamond", Communications Engineering **1**, 19 (2022), arXiv:2206.06734 [quant-ph].

[284]T. Gefen, A. Rotem, and A. Retzker, "Overcoming resolution limits with quantum sensing", Nature Communications **10**, 4992 (2019), arXiv:1811.01762 [quant-ph].

[285]A. G. Kofman and G. Kurizki, "Unified theory of dynamically suppressed qubit decoherence in thermal baths", Phys. Rev. Lett. **93**, 130406 (2004).

[286]Ł. Cywiński, R. M. Lutchyn, C. P. Nave, and S. Das Sarma, "How to enhance dephasing time in superconducting qubits", Phys. Rev. B **77**, 174509 (2008), arXiv:0712.2225 [cond-mat.mes-hall].

[287]G. A. Paz-Silva and L. Viola, "General transfer-function approach to noise filtering in open-loop quantum control", Phys. Rev. Lett. **113**, 250501 (2014), arXiv:1408.3836 [quant-ph].

[288]H. Ball and M. J. Biercuk, "Walsh-synthesized noise filters for quantum logic", EPJ Quantum Technology **2**, 11 (2015), arXiv:1410.1624 [quant-ph].

[289]T. J. Green, J. Sastrawan, H. Uys, and M. J. Biercuk, "Arbitrary quantum control of qubits in the presence of universal noise", New Journal of Physics **15**, 095004 (2013), arXiv:1211.1163 [quant-ph].

[290]E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, "$1/f$ noise: implications for solid-state quantum information", Rev. Mod. Phys. **86**, 361–418 (2014), arXiv:1304.7925 [cond-mat.mes-hall].

[291]B. Porat and B. Friedlander, "Computation of the exact information matrix of Gaussian time series with stationary random components", IEEE Transactions on Acoustics, Speech, and Signal Processing **34**, 118–130 (1986).

[292]OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, Published electronically at http://oeis.org, 2022.

[293]E. Delanoy, *Determinant of matrix defined by binomial coefficient*, Mathematics Stack Exchange, https://math.stackexchange.com/q/2277633, 2017.

[294]M. Abramowitz and I. A. Stegun, eds., *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Dover books on mathematics (Dover Publ, 2013).

[295]D. Grinberg, *A hyperfactorial divisibility*, Darij Grinberg homepage.

[296]M. Fukuda, R. König, and I. Nechita, "RTNI—A symbolic integrator for Haar-random tensor networks", Journal of Physics A: Mathematical and Theoretical **52**, 425303 (2019).

[297]E. W. Weisstein, *Negative binomial series*, https://mathworld.wolfram.com/NegativeBinomialSerie html, 2024.

[298]A. Taylor, *General theory of functions and integration*, Dover Books on Mathematics Series (Dover Publications, Incorporated, 2012).

[299]H. L. Royden and P. Fitzpatrick, *Real analysis*, Vol. 32 (Pearson, 1988).

[300]P. Blanchard and E. Brüning, *Mathematical methods in physics: distributions, Hilbert space operators, variational methods, and applications in quantum physics*, 2. ed, Progress in mathematical physics Vol. 69 (Birkhäuser, Cham Heidelberg, 2015).

[301]J. P. Grossmann, *Measure Theory YouTube playlist*, 2019.

[302]S. Sternberg, *Group theory and physics*, Transferred to digital printing (Cambridge Univ. Press, Cambridge, 2003).

[303]R. Cools, "Constructing cubature formulae: the science behind the art", Acta Numerica **6**, 1–54 (1997).

[304]W. van Dam and A. Russell, "Mutually unbiased bases for quantum states defined over p-adic numbers", arXiv:1109.0060 [quant-ph] (2011).

[305]A. Grundmann and H.-M. Möller, "Invariant integration formulas for the n-simplex by combinatorial methods", SIAM Journal on Numerical Analysis **15**, 282–290 (1978).

[306]J. Czartowski, D. Goyeneche, M. Grassl, and K. Zyczkowski, "Isoentangled Mutually Unbiased Bases, Symmetric Quantum Measurements, and Mixed-State Designs", Physical Review Letters **124**, 090503 (2020).

[307]J. Benedetto and W. Czaja, *Integration and modern analysis*, Birkhäuser Advanced Texts (Birkhäuser, Boston, 2009).

[308]B. Simon, "Distributions and Their Hermite Expansions", Journal of Mathematical Physics **12**, 140–148 (1971).

[309]A. Lundari, M. Miranda, and D. Pallara, "Infinite dimensional analysis", https://www.mathematik.tu-darmstadt.de/media/analysis/lehrmaterial_anapde/hallerd/Lectures.pdf (2015).

[310]D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, 1988).

[311]K. Temme, S. Bravyi, and J. M. Gambetta, "Error Mitigation for Short-Depth Quantum Circuits", Physical Review Letters **119**, 180509 (2017).

[312]A. Roy and A. J. Scott, "Weighted complex projective 2-designs from bases: optimal state determination by orthogonal measurements", Journal of Mathematical Physics **48**, 072110 (2007), arXiv:quant-ph/0703025.

[313]D. A. Harville, *Matrix Algebra From a Statistician's Perspective* (Springer, New York, NY, 1997).