Supplemental Material for "Complexity phase diagram for interacting and long-range bosonic Hamiltonians"

Nishad Maskara,[1,2,*] Abhinav Deshpande,[2,3,4,*] Adam Ehrenberg,[2,3]
Minh C. Tran,[2,3,5] Bill Fefferman,[2,6,7] and Alexey V. Gorshkov[2,3]

[1]*Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA*
[2]*Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, MD 20742, USA*
[3]*Joint Quantum Institute, NIST/University of Maryland, College Park, MD 20742, USA*
[4]*Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125, USA*
[5]*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA*
[6]*Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, USA*
[7]*Department of Computer Science, University of Chicago, Chicago, Illinois 60637, USA*

In this Supplemental Material, we present additional details regarding our main results, including full proofs. In Sec. S1, we expand upon the applicability of our model to various situations in condensed-matter systems and quantum information. In Section S2 and S3, we present proofs of the easiness results and analyze the easiness timescales in detail. We also show in Sec. S4 that the transition is coarse in 1D. In Section S5 and S6, we prove the hardness results for interacting and free bosons, respectively. Finally, in Sec. S7, we define a quantity that serves as an order parameter for the complexity transition.

S1. RELATED MODELS

In this section, we show that our results can be easily adapted to a wide range of experimentally and theoretically interesting Hamiltonians.

Fermionic systems with nearest-neighbor interactions can be incorporated into our model by performing the mapping described in Ref. [S1]. Our model is also relevant to cold atom experiments that have been proposed as candidates for observing quantum computational supremacy [S2–S4], especially in the nearest-neighbor limit.

The power-law hopping $1/r^\alpha$ can be engineered to directly implement the classes of Hamiltonians we study. This can be done by virtually coupling the band of interest to another with a quadratic band edge to implement exponentially decaying hopping [S5–S7]. Doing this simultaneously with multiple detunings approximates a power-law with high accuracy as a sum of exponentials [S8].

In the hardcore limit, the long-range hops translate to long-range interactions between spins, which model quantum-computing platforms such as Rydberg atoms and trapped ions [S9–S13]. Therefore, the Hamiltonian we study models various physically interesting situations, both in the several limiting cases ($\alpha \to \infty$, $V \to 0$, $V \to \infty$) as well as in the general case of finite nonzero α and V.

Our methods are fairly general for lattice models with long-range power-law decaying interactions. However, we do require number-conservation so the local Hilbert space dimension can be bounded at short times. As an example of how to extend our results to different models, we can straightforwardly incorporate long-range density-density interactions $K_{ij}(t)\, n_i n_j$. The only effect on the easiness times is to modify the Lieb-Robinson velocity to $v = O(b^2)$, where b is the maximum number of bosons per cluster. This is an overall constant that does not affect the exponent.

Our model can also be used to describe a distributed modular quantum network when the Hubbard interaction V can vary spatially. Specifically, a module of qubits can be represented by hardcore bosons ($V \to \infty$), while photonic communication channels linking distant modules can be represented by sites with $V = 0$ separating the modules. As in quantum networks, our hardness times in the nearest-neighbor regime are dominated by gates between nodes, while operations within a single node are free.

We also comment on the experimentally accessible timescales in these systems. In systems with ultracold atoms in optical lattices, tunneling times of $\tau = h/J \approx 4.3$ ms are achievable, with an achievable interaction strength of $V \approx 2J$ [S14]. These systems remain coherent for evolution times up to 100τ, meaning that a system with $L \approx 100$ can be studied. Similarly, in Rydberg atom setups, the gate time for an entangling gate between two neighboring atoms (which is relevant for the hardness timescale) is of the order of 190 ns [S15]. The lifetime of atoms excited to Rydberg levels is limited by spontaneous emission and is of the order of 50 μs [S10]. This means that one can experimentally access a system with $L \approx 50/0.19 \approx 250$.

---

* The two authors contributed equally.

S2. APPROXIMATION ERROR UNDER HHKL DECOMPOSITION

In this section, we analyze the HHKL decomposition from the main text and prove that it has low error.

We first argue why it is possible to apply the HHKL decomposition lemma to the Hamiltonian H' with a Lieb-Robinson velocity of order $O(1)$. As mentioned in the main text, H' is a Hamiltonian that lives in the truncated Hilbert space of at most $b+1$ bosons per cluster. Let Q be a projector onto this subspace. Then $H' = QHQ$. Time-evolution under this modified Hamiltonian H' keeps a state within the subspace since $\left[e^{-iQHQt}, Q\right] = 0$.

The Lieb-Robinson velocity only depends on the norm of terms in the Hamiltonian which couple lattice sites. On-site terms do not contribute, which can be seen by moving to an interaction picture [S16, S17]. Therefore, since no state has more than $b+1$ bosons on any site within the image of Q, the maximum norm of coupling terms in H' is $\left\|Qa_i^\dagger a_j Q\right\| \leq b+1$. Therefore, the Lieb-Robinson velocity is at most $O(b)$ instead of $O(n)$, and we can apply the HHKL decomposition to the evolution generated by the truncated Hamiltonian H'. We now prove that the error made by decomposing the evolution due to H' is small.

Lemma 3 (Decomposition error for H')**.** *For all V and $\alpha > D+1$, the error incurred (in 2-norm) by decomposing the evolution due to H' for time t is*

$$\epsilon(t) \leq O\left(K(e^{vt_1}-1)(\ell^{-\alpha+D+1}+e^{-\ell})\sum_{j=0}^{N-1}(r_0+j\ell)^{D-1}\right), \tag{S1}$$

where $N = t/t_1$ and $\ell \leq L/N$ can be chosen to minimize the error, and r_0 is the radius of the smallest sphere containing the initially occupied bosons in a cluster.

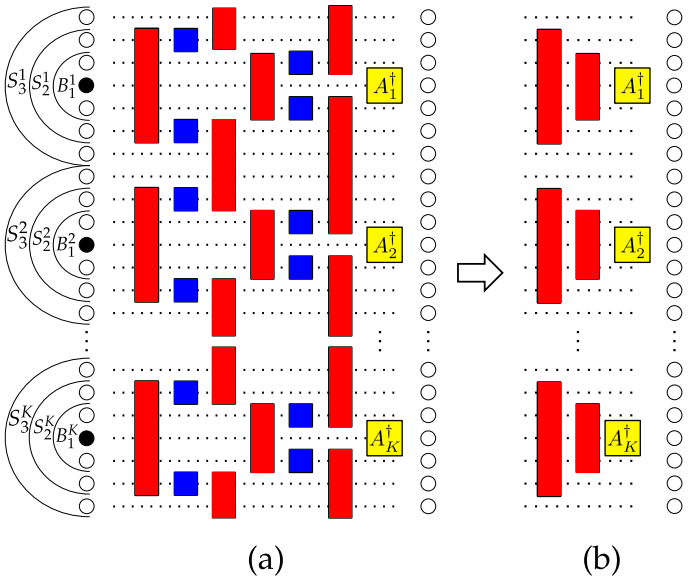


FIG. S1 (Color online). (a) Decomposition of the first two steps of the unitary evolution followed by (b) pushing the commuting terms past $A_i^\dagger$ (the product of all initial creation operators in a cluster i) to the vacuum. Red boxes represent forward evolution and blue boxes backward evolution in time.

The sketch of the proof is as follows: recall that within each cluster C_i, there is a group of bosons initially separated from the edge of the cluster by a region of width L_i. Naive application of the HHKL decomposition for the long-range case results in a timescale $t_{\text{easy}} \sim \log(L)$, because of the exponential factor $(e^{vt}-1)$ in the error. To counter this, we apply the HHKL decomposition in small time-steps t_1. Thus, within each time-step, the exponential factor can be approximated as $e^{vt_1}-1 \approx vt_1$, turning this exponential dependence into a polynomial one at the cost of an increased number of time-steps.

The first two time-steps are depicted pictorially in Fig. S1, and illustrate the main ideas. The full propagator acting on the entire lattice is decomposed by applying the HHKL decomposition K times, such that two of every three forward and reverse time-evolution operators commute with all previous operators by virtue of being spatially disjoint, allowing them to be pushed through and act identically on the vacuum. The remaining forward evolution operator effectively spreads out the bosonic operators by distance ℓ. The error per time-step is polynomially suppressed by $O(\ell^{-\alpha+D+1}+e^{-\gamma\ell})$.

While it reduces the exponential factor to a polynomial one, using time-slices comes at the cost of extra polynomial factors, originating from the sum over boundary terms $\sum_{j=0}^{N-1}(r_0+j\ell)^{D-1}$.

Proof of Lemma 3. Let the initial positions of the bosons be denoted by $(\mathsf{in}_1, \ldots, \mathsf{in}_n)$. The initial state is $|\psi(0)\rangle = a^\dagger_{\mathsf{in}_1} \ldots a^\dagger_{\mathsf{in}_n} |0\rangle$ As before, the first two time-steps are illustrated in Fig. S1. Within each cluster C_i, there is a group of bosons initially separated from the edge of the cluster by a region of width L_i. Let $A^\dagger_i(0) = \prod_{\mathsf{in}_j \in C_i} a^\dagger_{\mathsf{in}_j}$ be the creation operator for the group of bosons in the cluster C_i, so that the initial state can also be expressed as $|\psi(0)\rangle = \prod_{i=1}^K A^\dagger_i(0)\,|0\rangle$. The forward-time propagator on a region R is $U^R_{t_0,t_1} = \mathcal{T}\exp\left(-i\int_{t_0}^{t_1} H_R(s)ds\right)$. When evolved for short times, each creation operator $a^\dagger_{\mathsf{in}_i}(t)$ is mostly supported over a small region around its initial position. Therefore, as long as these regions do not overlap, each operator approximately commutes, and the state is approximately separable.

Let A^i be the smallest ball upon which $A^\dagger_i(0)$ is supported. Let $B^i_0 = A^i$ and denote its radius r^i_0, and define $r_0 = \max r^i_0$. B^i_k is a ball of radius $r^i_0 + k\ell$ containing A^i, where ℓ will be chosen to minimize the error. S^i_k is the shell $B^i_k \setminus B^i_{k-1}$ (see Fig. S1). The complement of a set X is denoted as X^c. We divide the evolution into N time steps between $t_0 = 0$ and $t_N = t$, and first show that the evolution is well-controlled for one time step from 0 to $t_1 = t/N$. We apply this decomposition K times, once for each cluster, letting $X = B^i_0$, $Y = S^i_1$ and Z be everything else:

$$U_{0,t_1} \approx U^{B^1_1}_{0,t_1}(U^{S^1_1}_{0,t_1})^\dagger U^{(B^1_0)^c}_{0,t_1} \tag{S2}$$

$$\approx U^{B^1_1}_{0,t_1}(U^{S^1_1}_{0,t_1})^\dagger U^{B^2_1}_{0,t_1}(U^{S^2_1}_{0,t_1})^\dagger U^{(B^1_0 B^2_0)^c}_{0,t_1} \tag{S3}$$

$$\approx U^{B^1_1}_{0,t_1}(U^{S^1_1}_{0,t_1})^\dagger \ldots U^{B^K_1}_{0,t_1}(U^{S^K_1}_{0,t_1})^\dagger U^{(B^1_0 \ldots B^K_0)^c}_{0,t_1}. \tag{S4}$$

The total error is $O\left(\sum_{i=1}^K (e^{vt_1}-1)\Phi(B^i_0)(\ell^{-\alpha+D+1}+e^{-\gamma\ell})\right) = O\left(K(e^{vt_1}-1)r_0^{D-1}(\ell^{-\alpha+D+1}+e^{-\gamma\ell})\right)$. Applying the decomposed unitary to the initial state and pushing commuting terms through to the vacuum state, we get

$$U_{0,t_1}|\psi(0)\rangle \approx U^{B^1_1}_{0,t_1}A^\dagger_1 \ldots U^{B^K_1}_{0,t_1}A^\dagger_K|0\rangle = \left(\prod_{i=1}^K U^{B^i_1}_{0,t_1}A^\dagger_i\right)|0\rangle.$$

We can repeat the procedure for the unitary U_{t_1,t_2}, where $t_2 = 2t_1$. Now, the separating region Y will be S^i_2, so that $S^i_2 \cap B^i_1 = \emptyset$. Each such region still has width ℓ, but now the boundary of the interior is $\Phi(B^i_1) = O((r_0+\ell)^{D-1})$. We get

$$U_{t_1,t_2} \approx \left(\prod_{i=1}^K U^{B^i_2}_{t_1,t_2}(U^{S^i_2}_{t_1,t_2})^\dagger\right) U^{(B^1_1 \ldots B^K_1)^c}_{t_1,t_2}, \tag{S5}$$

with error $O(K(e^{vt_1}-1)(r_0+l)^{D-1}(\ell^{-\alpha+D+1}+e^{-\gamma\ell}))$. The unitaries supported on S^i_2 and $(B^1_1 \ldots B^K_1)^c$ commute with all the creation operators supported on sites B^i_1, giving $|\psi(t_2)\rangle \approx U^{B^1_2}_{t_1,t_2}U^{B^1_1}_{0,t_1} \ldots U^{B^K_2}_{t_1,t_2}U^{B^K_1}_{0,t_1}|\psi(0)\rangle$. By applying this procedure a total of N times, once for each time step, we get the approximation $U_{0,t_N}|\psi(0)\rangle \approx U^{B^1_N}_{t_{N-1},t_N} \ldots U^{B^1_1}_{0,t_1} \ldots U^{B^K_N}_{t_{N-1},t_N} \ldots U^{B^K_1}_{0,t_1}|\psi(0)\rangle$. The total error in the state (in 2-norm) is

$$\epsilon \le O\left(K(e^{vt_1}-1)(\ell^{-\alpha+D+1}+e^{-\gamma\ell})\sum_{j=0}^{N-1}(r_0+j\ell)^{D-1}\right) \tag{S6}$$

$$= O\left(n(e^{vt_1}-1)(\ell^{-\alpha+D+1}+e^{-\gamma\ell})NL^{D-1}\right), \tag{S7}$$

proving Lemma 3. The last inequality comes from the fact that $K \le n$ and that $r_0 + (N-1)\ell \le \min L_i = L$. The latter condition ensures that the decomposition of the full unitary is separable on the clusters.

In the regime $\alpha > 2D + D/(\beta-1)$, t_{easy} is optimized by choosing a fixed time-step size $t_1 = O(1)$. Then, the number of steps N scales as the evolution time $N = t/t_1$. By the last few time-steps, the bosonic operators have spread out and have a boundary of size L^{D-1}, so the boundary terms contribute $O(NL^{D-1})$ in total. In the regime $D+1 < \alpha \le 2D + D/(\beta-1)$, the boundary contribution outweighs the suppression $\ell^{-\alpha+D+1}$. Instead, we use a single time-step in this regime, resulting in $t_{\text{easy}} = \Omega(\log n)$ when $\beta > 1$.

□

S3. CLOSENESS OF EVOLUTION UNDER H AND H'.

In this section, we show that the states evolving due to H and H' are close, owing to the way the truncation works. This will enable us to prove that the easiness timescale for H is the same as that of H'.

Suppose that an initial state $|\psi(0)\rangle$ evolves under two different Hamiltonians $H(t)$ and $H'(t)$ for time t, giving the states $|\psi(t)\rangle = U_t |\psi(0)\rangle$ and $|\psi'(t)\rangle = U'_t |\psi(0)\rangle$, respectively. Define $|\delta(t)\rangle = |\psi(t)\rangle - |\psi'(t)\rangle$ and switch to the rotating frame, $|\delta^r(t)\rangle = U_t^\dagger |\delta(t)\rangle = |\psi(0)\rangle - U_t^\dagger U'_t |\psi(0)\rangle$. Now taking the derivative,

$$i\partial_t |\delta^r(t)\rangle = 0 + U_t^\dagger H(t) U'_t |\psi(0)\rangle - U_t^\dagger H'(t) U'_t |\psi(0)\rangle \tag{S8}$$

$$= U_t^\dagger (H(t) - H'(t)) |\psi'(t)\rangle . \tag{S9}$$

The first line comes about because $i\partial_t U'_t = H'(t)U'_t$ and $i\partial_t U_t^\dagger = -U_t^\dagger H(t)$, owing to the time-ordered form of U_t.

Now, we can bound the norm of the distance, $\delta(t) := \||\delta(t)\rangle\| = \||\delta^r(t)\rangle\|$.

$$\delta(t) \leq \delta(0) + \int_0^t d\tau \|(H(\tau) - H'(\tau)) |\psi'(\tau)\rangle\| \tag{S10}$$

$$= \int_0^t d\tau \|(H(\tau) - H'(\tau)) |\psi'(\tau)\rangle\|, \tag{S11}$$

since $\delta(0) = 0$.

The next step is to bound the norm of $(H - H') |\psi'(\tau)\rangle$ (we suppress the time label τ in the argument of H and H' here and below). We use the HHKL decomposition: $|\psi'(\tau)\rangle = |\phi(\tau)\rangle + |\epsilon(\tau)\rangle$, where the state $|\phi(\tau)\rangle$ is a product state over clusters, and $|\epsilon(\tau)\rangle$ is the error induced by the decomposition. We first show that $(H - H') |\phi(\tau)\rangle = 0$. Since $|\phi(\tau)\rangle$ is a product state of clusters, each of which is time-evolved separately, boson number is conserved within each cluster. Therefore, each cluster has at most b bosons, and $Q |\phi(\tau)\rangle = |\phi(\tau)\rangle$. Furthermore, only the hopping terms in H can change the boson number distribution among the different clusters, and these terms move single bosons. This implies that $H |\phi(\tau)\rangle$ has at most $b + 1$ bosons per cluster, and remains within the image of Q, denoted $\mathrm{im}\ Q$. Combining these observations, we get $H' |\phi(\tau)\rangle = QHQ |\phi(\tau)\rangle = H |\phi(\tau)\rangle$. This enables us to say that $(H - H') |\phi(\tau)\rangle = (H - QHQ) |\phi(\tau)\rangle = 0$. Equation (S11) gives us

$$\delta(t) \leq \int_0^t d\tau \|(H(\tau) - H'(\tau))(|\phi(\tau)\rangle + |\epsilon(\tau)\rangle)\| \tag{S12}$$

$$= \int_0^t d\tau \|(H(\tau) - H'(\tau)) |\epsilon(\tau)\rangle\|, \tag{S13}$$

$$\leq \max_{\tau, |\eta\rangle \in \mathrm{im}\ Q} \|(H(\tau) - H'(\tau)) |\eta\rangle\| \int_0^t d\tau \||\epsilon(\tau)\rangle\|. \tag{S14}$$

In the last inequality, we have upper bounded $\|(H(\tau) - H'(\tau)) |\epsilon(\tau)\rangle\|$ by $\max_{|\eta\rangle \in \mathrm{im}\ Q} \|(H - H') |\eta\rangle\| \times \epsilon(\tau)$, where $\epsilon(\tau) := \||\epsilon(\tau)\rangle\|$. The quantity $\max_{|\eta\rangle \in \mathrm{im}\ Q} \|(H - H') |\eta\rangle\|$ can be thought of as an operator norm of $H - H'$, restricted to the image of Q. It is enough to consider a maximization over states $|\eta\rangle$ in the image of Q because we know that the error term $|\epsilon(\tau)\rangle$ also belongs to this subspace, as $|\psi'(\tau)\rangle$ belongs to this subspace. Further, we give a uniform (time-independent) bound on this operator norm, which accounts for the maximization over times τ.

Lemma 4. $\max_{|\eta\rangle} \|(H - QHQ) |\eta\rangle\| \leq \left\| \sum_{i \in C_k, j \in C_l} J_{ij} a_i^\dagger a_j \right\| \leq O(bL^{D-\alpha})$.

Proof. Notice that for each term H_i in the Hamiltonian, the operator $H - QHQ$ contains $H_i - QH_iQ$, where the rightmost Q can be neglected since $Q |\eta\rangle = |\eta\rangle$. The on-site terms $\sum_i J_{ii} a_i^\dagger a_i + V n_i(n_i - 1)/2$ do not change the boson number. Therefore, they cannot take $|\eta\rangle$ outside the image of Q, and do not contribute to $(H - QHQ) |\eta\rangle$. The only contribution comes from hopping terms that change boson number, which we bound by

$$\left\| \sum_{i \in C_k, j \in C_l} J_{ij} a_i^\dagger a_j \right\|, \tag{S15}$$

where the sum is over sites i and j in distinct clusters C_k and C_l, respectively. This is because only hopping terms that connect different clusters can bring $|\eta\rangle$ outside the image of Q, since hopping terms within a single cluster maintain the number of bosons per cluster.

For illustration, let us focus on terms that couple two clusters C_1 and C_2. The distance between these two clusters is denoted L_{12}. For any coupling J_{ij} with $i \in C_1$ and $j \in C_2$, we can bound $|J_{ij}| \leq L_{12}^{-\alpha}$ by assumption. Let

$$H_{12}^{\mathrm{hop}} = \sum_{i \in C_1} \sum_{j \in C_2} J_{ij} a_i^\dagger a_j + \mathrm{h.c.} \tag{S16}$$

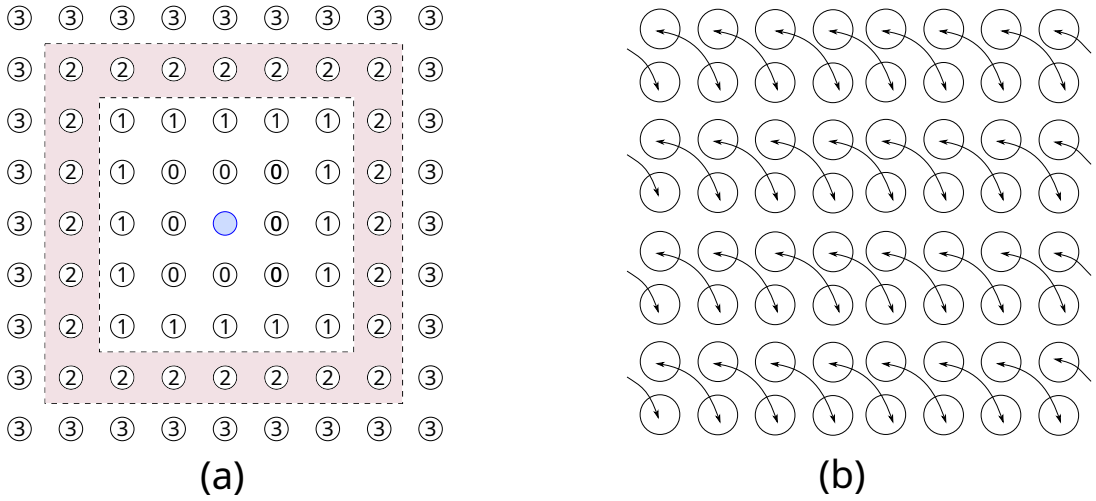


FIG. S2 (Color online). (a) Cluster distances between the blue cluster in the center and nearby clusters. The total number of clusters at cluster distance $l = 2$ (pink background) is given by $(2l+3)^2 - (2l+1)^2 = 24$. (b) Non-overlapping pairing between clusters separated by a diagonal. The distance between these clusters is $l = 0$, since they share a boundary and contain adjacent sites.

denote the sum over all such pairs of sites. Then, we can bound $\left\| H_{12}^{\text{hop}} |\eta\rangle \right\| \leq O(b)$. To see this, diagonalize $H_{12}^{\text{hop}} = \sum_i w_i b_i^\dagger b_i$. Since H_{12}^{hop} only acts on two clusters, each normal mode contains up to $2b$ bosons. The maximum eigenvalue of H_{12}^{hop} is bounded by $2b \max_i w_i$, where w_i is the maximum normal mode frequency, given by the eigenvalue of the matrix $J_{ij} : i \in C_1, j \in C_2$. We now apply the Gershgorin circle theorem, which states that the maximum eigenvalue of J is bounded by the quantity $\max_i(\sum_j |J_{ij}|) \leq L^D L_{12}^{-\alpha}$.

Taking advantage of the fact that the clusters form a cubic lattice in D dimensions, we can group pairs of clusters by their relative distances. If we label clusters i by their D-dimensional coordinate $i_1, i_2, ..., i_D$, then we can define the cluster distance l between i and j as $l+1 = \max_d |i_d - j_d|$. Cluster distance l corresponds to a minimum separation $l \times L$ between sites in different clusters. With this definition, there are $((2l+3)^D - (2l+1)^D) \approx 2^D D l^{D-1}$ clusters at a cluster distance l from any given cluster (Fig. S2(a)), and $K \times 2^D D l^{D-1}$ total pairs of clusters at cluster distance l. Notice that for a given separation vector, $K/2$ pairs of clusters (K total) can be simultaneously coupled without overlap (Fig. S2(b)). Therefore, there are approximately $2^{D+1} D l^{D-1}$ non-overlapping groupings per distance l. The sum over these non-overlapping Hamiltonians $H_{a_1 b_1}^{\text{hop}} + ... + H_{a_{K/2} b_{K/2}}^{\text{hop}}$ for each grouping is block diagonal. Therefore, the spectral norm (maximum eigenvalue) of the total Hamiltonian is equal to the maximum of the spectral norm over all irreducible blocks. Putting all this together, as long as $D - 1 - \alpha < -1$ the bound becomes

$$\max_{|\eta\rangle \in \text{im } Q} \|(H - QHQ)|\eta\rangle\| \leq \sum_{l=0}^{l_{\max}} O(2^{D+1} D l^{D-1})(2b) L^D (lL)^{-\alpha} = O(bL^{D-\alpha}). \tag{S17}$$

□

We are now in a position to prove our main easiness result.

Theorem 5 (Easiness result). *For $\alpha > D+1$, and for all V, including $V = o(1)$ and $V = \omega(1)$, we have $t_{\text{easy}} = \Omega(n^{\gamma_{\text{easy}}})$, with*

$$\gamma_{\text{easy}} = \frac{\beta - 1}{D} \times \frac{\alpha - 2D}{\alpha - D} - \frac{1}{\alpha - D}, \tag{S18}$$

and $t_{\text{easy}} = \Omega(\log n)$ if $\gamma_{\text{easy}} < 0$.

Proof. There are two error contributions, ϵ and δ, to the total error. The HHKL error ϵ is given by evaluation of Eq. (S7), which is minimized by either choosing $N = 1$ or $N = t/t_1$ with t_1 a small fixed constant. This leads to three regimes with errors

$$\epsilon \leq O(1) \times \begin{cases} ne^{vt-L}, & \alpha \to \infty \\ \frac{nt^{\alpha-D}}{L^{\alpha-2D}}, & 2D + \frac{D}{\beta-1} < \alpha < \infty \\ \frac{n(e^{vt}-1)}{L^{\alpha-D-1}}, & D+1 < \alpha \leq 2D + \frac{D}{\beta-1}. \end{cases} \tag{S19}$$

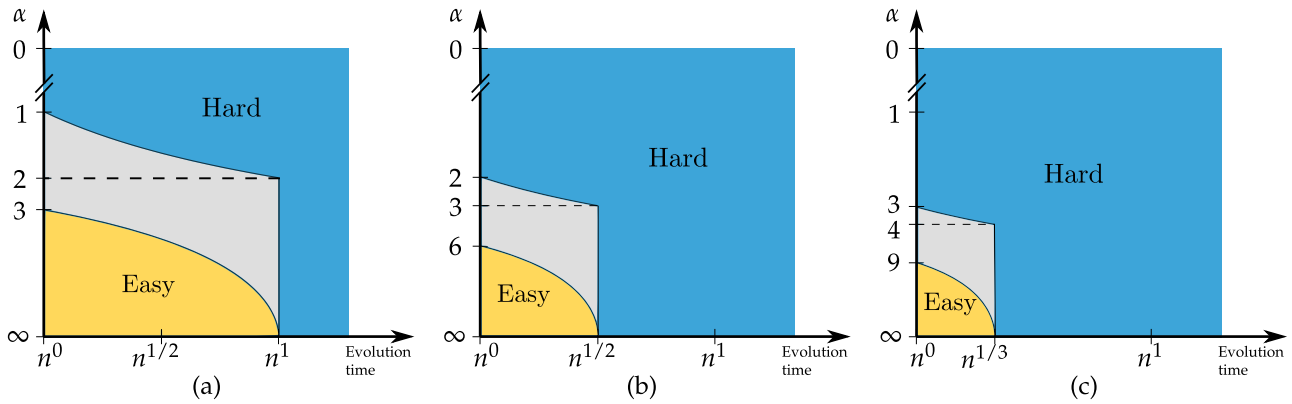


FIG. S3 (Color online). Slices of the complexity phase diagram for the long-range bosonic Hamiltonian in a) 1D, b) 2D, and c) 3D with n bosons when the number of sites is $m = \Theta(n^2)$. Colors represent whether the sampling problem is easy (yellow), hard (blue), or not currently known (gray). The X-axis parametrizes the evolution time as a polynomial function of n, and the Y-axis is α, the exponent characterizing the long-range nature of the hopping Hamiltonian (with scale $y = 1/\sqrt{\alpha}$ except for the point $\alpha = 0$).

Regime	Error	$t_{\text{easy}}(n, L)$	$t_{\text{easy}}(\rho)/t_{\text{easy}}(\rho = 1)$
$\alpha \to \infty$	ne^{vt-L}	L	$1/\rho$
$2D + \frac{D}{\beta-1} < \alpha < \infty$	$\frac{nvt^{\alpha-D}}{L^{\alpha-2D}}$	$n^{\frac{-1}{\alpha-D}} L^{\frac{\alpha-2D}{\alpha-D}}$	$\rho^{\frac{-2}{\alpha-D}}$
$D + 1 < \alpha \leq 2D + \frac{D}{\beta-1}$	$\frac{n(e^{vt}-1)}{L^{\alpha-D-1}}$	$(\alpha - D - 1)\log L - \log n$	$1/\rho$

TABLE I (Color online). Summary of easiness timescales in the different regimes. Timescales follow from the error and are presented first as a function of n and L, which are the relevant physical scales of the problem. We study the effect of the density by performing the scaling $n \to \rho n, L \to L, b \to \rho b, m \to m$. The last column shows the timescale as a function of ρ in terms of the timescale when $\rho = 1$, namely $t_{\text{easy}}(\rho = 1)$.

The truncation error, arising from using H' rather than H in the first step, is given by

$$\delta(t) \leq O(bL^{D-\alpha}) \int_0^t d\tau \epsilon(\tau). \tag{S20}$$

Therefore, we can upper bound $\delta(t)$ by ϵ times an additional factor. This factor is $bL^{D-\alpha}t$, meaning $\delta(t) = bL^{D-\alpha}t\epsilon$, when $\epsilon(\tau) = \text{poly}(\tau)$ ($\alpha < \infty$). The factor is $L^{D-\alpha}$, meaning $\delta(t) = L^{D-\alpha}\epsilon$, when $\epsilon(\tau) = \exp(v\tau)$ ($\alpha \to \infty$). Our easiness results only hold for $\alpha > D + 1$, so the L-dependent factor serves to suppress the truncation error in the asymptotic limit. Although the additional factor of t could cause $\delta(t) > \epsilon$ at late times, by this time, $\epsilon > \Omega(1)$ and we are no longer in the easy regime. Therefore, the errors presented in Eq. (S19) can be immediately applied to calculate the timescales from the main text. □

The resulting timescales are summarized in Table I, which highlights the scaling of the timescale with respect to different physical parameters. The easiness exponents in Tables I and II of the main text can be obtained from the easiness timescales t_{easy} in Table I after substituting $L = \Theta(n^{(\beta-1)/D})$. We also consider the scaling of the easiness timescales when the density of the bosons increases by a factor ρ. In our setting, we implement this by scaling the number of bosons by ρ while keeping the number of lattice sites and the number of clusters (and their size) fixed. The effect of this is to increase the Lieb-Robinson velocity: $v \to v\rho$. For all three cases, the net effect of increasing the density by a factor ρ is to decrease the easiness timescale. We also present phase diagrams for the specific cases of $D = 1, 2, 3$ in Fig. S3, which illustrates the dimensional dependence of the easy and hard regions.

S4. EXTENDED EASINESS TIMESCALE FOR 1D

In this section, we prove that 1D systems with nearest-neighbor interactions ($\alpha \to \infty$) can be simulated for longer times by using matrix product states (MPS).

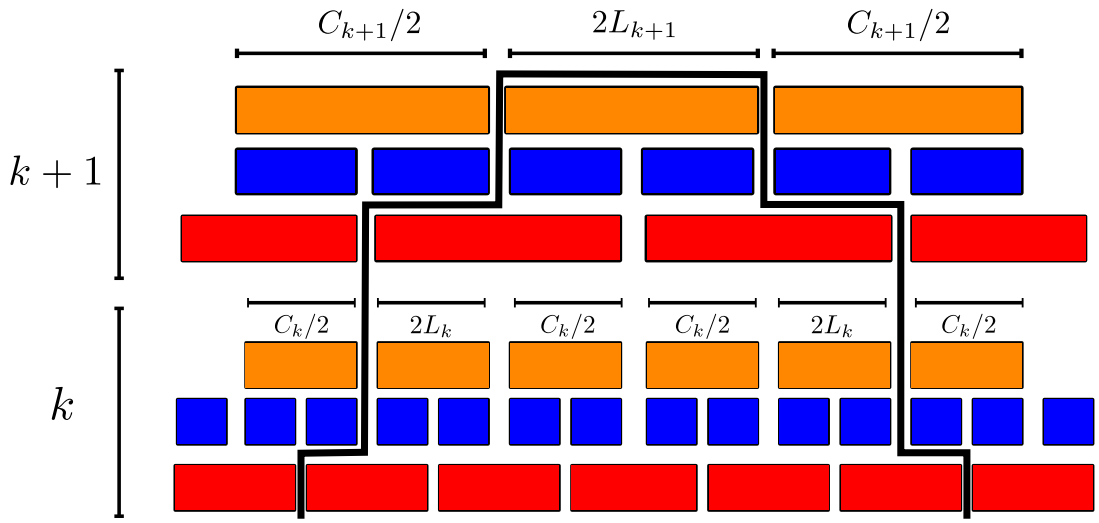


FIG. S4 (Color online). Schematic of the HHKL decomposition used to extend the classical simulation algorithm to longer times. Red and orange blocks denote forward time evolution, and blue blocks denote backwards time evolution. Two timesteps are depicted, k and $k+1$. The past causal lightcone of one of the blocks is outlined in black. The lightcone dimensions and HHKL block size double during subsequent timesteps, leading to an exponentially large causal region.

Specifically, we show that approximate sampling up to time $t = cL$ for any constant c is easy. This shows sampling is easy for any timescale $t = O(L)$. When combined with our hardness result for timescales $t = O(L^{1+\delta})$ for any $\delta > 0$, this proves the transition is coarse.

The first step is to show how time evolution up to time $t = cL$ can be approximated with low error using the HHKL decomposition. Earlier, we proved that for $t = L/v$, we can self-consistently truncate the local Hilbert space dimension while incurring little error. Then, the sampling algorithm relies on the separability of the approximate wavefunction.

For $c > 1$, this separability no longer holds. Nevertheless, we can use a similar argument to self-consistently truncate the local Hilbert space. However, due to the lack of separability, we only have an efficient simulation algorithm in $D = 1$, where tensor network representations of low-entanglement (area law) states can be efficiently contracted.

We divide the time evolution into time intervals, where we have already proved sampling is easy in the first interval up to time $t = t_1 = L/v_0$. Recall that during this first time interval, we applied the HHKL decomposition with a block size L. For the next time interval, we decompose the unitary U_{t_2,t_1} choosing an HHKL block size $2L$. By looking at the past causal lightcone of a single site, it is clear that the boson number per site is bounded by $4b$. Therefore, the Lieb-Robinson velocity during this time interval is four times larger that the first time interval, $v_1 = 4v_0$. Setting the approximation error $O\left(ne^{v_1(t_2-t_1)-2L}\right)$ to be $o(1)$, we can determine that for $t_2 - t_1 = O(\frac{L}{2v_0})$ the approximation errors are well controlled. Next, we need to generalize the argument to k time evolution steps.

We prove that we can define a self-consistent Hilbert space truncation scheme and an HHKL decomposition of the time-evolution unitary. During the k-th time interval, we choose the HHKL block size to be $L_k = 2^k L$. To estimate the maximum local boson number, we look at the spatial extent of the past causal lightcone. Let $C_k + 2L_k$ denote the spatial extent of the past causal lightcone. Here C_k measures how far sideways the lightcone extends (see Fig. S4). Looking at the lightcone, we can write down a recurrence relation for C_k.

$$2L_{k+1} + C_{k+1} = 4L_{k+1} + C_k \tag{S21}$$

$$C_{k+1} = 2L_{k+1} + C_k = 2^{k+2}L + C_k \tag{S22}$$

$$\implies C_k = 4(2^k - 1)L, \tag{S23}$$

where we have applied the initial condition $C_0 = 0$. Combined with the boson density $\rho = \frac{1}{2L}$, the maximum boson number after step k is proportional to the size of the lightcone:

$$N_k = \rho(C_k + 2L_k) \tag{S24}$$

$$= \rho L(6 \times 2^k - 4) = 3 \times 2^k - 2. \tag{S25}$$

This tells us we can choose timesteps $t_k - t_{k-1} = O(L/3v_0)$ for the error to be $o(1)$.

To complete the proof, we need to argue that the HHKL time-evolution can be simulated efficiently. The simplest way to do this is to fix the evolution time $t = cL/3v_0$. Then, we can think of the HHKL decomposed unitary as a finite-depth unitary

circuit, which generates constant (system-size independent) entanglement across any cut. Therefore, standard MPS algorithms can be applied with cost polynomial in the system size, the local Hilbert space dimension (physical index dimension), and the entanglement across any cut (virtual index dimension) [S18]. Although both the physical and virtual dimensions scale exponentially or faster with c, since we have fixed c to be system-size independent, we have an efficient sampling algorithm. Furthermore, this algorithm works for any fixed c, proving that sampling cannot be hard for evolution times $t = O(L)$ and that the transition is coarse.

S5. HARDNESS TIMESCALE FOR INTERACTING BOSONS

In this section, we provide more details about how to achieve the timescales in our hardness results, which we state in more detail first.

Theorem 6 (Hardness result)**.** *When $\alpha \geq D/2$, $V = \Omega(1)$, and $D \geq 2$, the hardness timescale is $t_{\text{hard}} = O(n^{\gamma^{\text{I}}_{\text{hard}}})$, where*

$$\gamma^{\text{I}}_{\text{hard}} = \begin{cases} \frac{\beta-1}{D} \min[1, \alpha - D], & \alpha > D \\ 0, & \alpha \in [\frac{D}{2}, D]. \end{cases} \tag{S26}$$

In all other cases, i.e. when at least one of the following cases holds: $\alpha < D/2$, $D = 1$, or nearly free bosons $V = o(1)$, the timescale is $t_{\text{hard}} = O(n^{\gamma^{\text{II}}_{\text{hard}}})$, where

$$\gamma^{\text{II}}_{\text{hard}} = \delta + \begin{cases} \frac{\beta-1}{D} \min\left[1 + \frac{O(\log(V+1))}{\log n}, \alpha - D\right], & \alpha > D \\ 0, \quad \alpha \in [\frac{D}{2}, D] \\ \frac{\beta}{D}\left(\alpha - \frac{D}{2}\right), & \alpha < D/2 \end{cases} \tag{S27}$$

for an arbitrarily small $\delta > 0$.

Almost any bosonic interaction is universal for BQP [S19] and hence these results are applicable to general on-site interactions $f(n_i)$. Reference [S20] also answers the questions of what additional gates or Hamiltonians can make linear optics universal. We first describe how a bosonic system with fully controllable local fields $J_{ii}(t)$, hoppings $J_{ij}(t)$, and a fixed Hubbard interaction $\frac{V}{2}\sum_i \hat{n}_i(\hat{n}_i - 1)$ can implement a universal quantum gate set. To simulate quantum circuits, which act on two-state spins, we use a dual-rail encoding. Using $2n$ bosonic modes, and n bosons, n logical qubits are defined by partitioning the lattice into pairs of adjacent modes, and a boson is placed in each pair. Each logical qubit spans a subspace of the two-mode Hilbert space. Specifically, $|0\rangle_L = |10\rangle, |1\rangle_L = |01\rangle$. We can implement any single qubit (2-mode) unitary by turning on a hopping between the two sites (X-rotations) or by applying a local on-site field (Z-rotations). To complete a universal gate set, we need a two-qubit entangling gate. This can be done, say, by applying a hopping term between two sites that belong to different logical qubits [S21]. All these gates are achievable in $O(1)$ time when $V = \Theta(1)$. In the limit of large Hubbard interaction $V \to \infty$, the entangling power of the gate decreases as $1/V$ [S21] and one needs $O(V)$ repetitions of the gate in order to implement a standard entangling gate such as the CNOT.

For hardness proofs that employ postselection gadgets, we must ensure that the gate set we work with comes equipped with a Solovay-Kitaev theorem. This is the case if the gate set is closed under inverse, or contains an irreducible representation of a non-Abelian group [S22]. In our case, the gate set contains single-qubit Paulis and hence has a Solovay-Kitaev theorem, which is important for the postselection gadgets to work as intended.

We will specifically deal with the scheme proposed in Ref. [S4]. It applies a constant-depth circuit on a grid of $\sqrt{n} \times \sqrt{n}$ qubits in order to implement a random IQP circuit [S23, S24] on $\sqrt{n}$ effective qubits. This comes about because the cluster state, which is a universal resource for measurement-based quantum computation, can be made with constant depth on a two-dimensional grid.

For short-range hops ($\alpha \to \infty$), we implement the scheme in four steps as shown in Fig. S5. In each step, we move the logical qubits to bring them near each other and make them interact in order to effect an entangling gate. For short-range hopping, the time taken to move a boson to a far-off site distance L away dominates the time taken for an entangling gate. The total time for an entangling gate is thus $O(L) + O(1) = O(L)$.

For long-range hopping, we use the same scheme as in Fig. S5, but we use the long-range hopping to speed up the movement of the logical qubits. This is precisely the question of state transfer using long-range interactions/hops [S25–S27]. In the following we give an overview of the best known protocol for state transfer, but first we should clarify the assumptions in the model. The Hamiltonian is a sum of $O(m^2)$ terms, each of which has norm bounded by at most $1/d(i,j)^\alpha$. Since we assume we can apply any Hamiltonian subject to these constraints, in particular, we may choose to apply hopping terms across all possible edges. This model makes it possible to go faster than the circuit model if we compare the time in the Hamiltonian model with depth in the

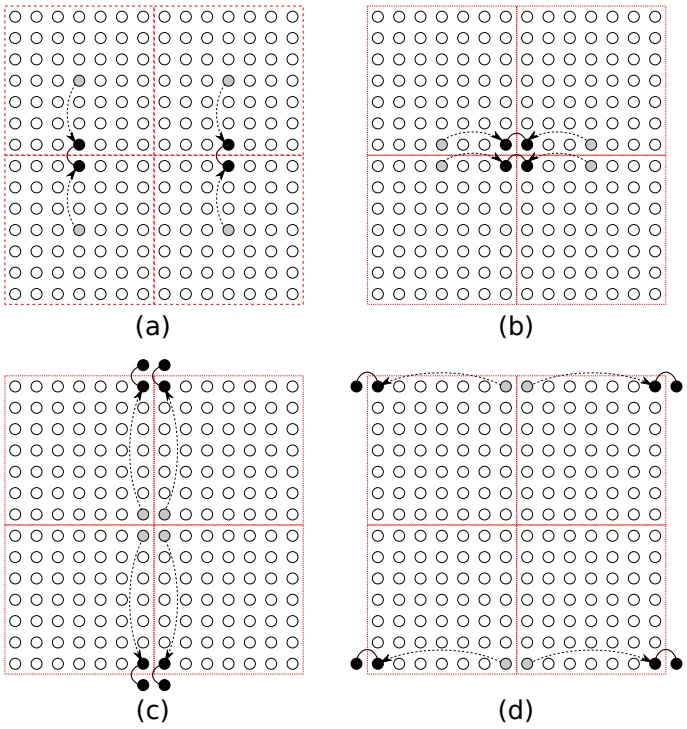


FIG. S5 (Color online). A protocol that implements the logical circuit of Ref. [S4]. Each subfigure shows the location of the site that previously encoded the $|1\rangle$ state in gray. The current site that encodes the $|1\rangle$ state is in black. The site that encodes $|0\rangle$ is not shown but moves similarly as the $|1\rangle$ state. The distance traversed by each qubit is $L + L + 2L + 2L = 6L$.

circuit model. This power comes about because of the possibility of allowing simultaneous noncommuting terms to be applied in the Hamiltonian model.

The state transfer protocols in Ref. [S26, S27] show such a speedup for state transfer. The broad idea in both protocols is to apply a map $|1\rangle_1 \to |1\rangle_A := \sum_{j\in A} \frac{1}{\sqrt{|A|}} |1\rangle_j$, followed by the steps $|1\rangle_A \to |1\rangle_B$ and $|1\rangle_B \to |1\rangle_2$, where A and B are regions of the lattice to be specified. In the protocol of Ref. [S26], which is faster than that of Ref. [S27] for $\alpha \leq D/2$, $A = B = \{j : j \neq 1, 2\}$ and each step takes time $O(L^\alpha/\sqrt{N-2})$, where $N-2$ is the number of ancillas used and L is the distance between the two furthest sites. In the protocol of Ref. [S27], which is faster for $\alpha \in (D/2, D+1]$, A and B are large regions around the initial and final sites, respectively. This protocol takes time $O(1)$ when $\alpha < D$, $O(\log L)$ when $\alpha = D$, and $O(L^{\alpha-D})$ when $\alpha > D$.

In our setting, we use the state transfer protocols to move the logical qubit faster than time $O(L)$ in each step of the scheme depicted in Fig. S5. If $\alpha < D/2$, we use all the ancillas in the entire system, giving a state transfer time of $O(m^{\alpha/D-1/2}) = O(n^{\beta(\frac{\alpha}{D}-\frac{1}{2})})$. If $\alpha > D/2$, we only use the empty sites in a cluster as ancillas in the protocol of Ref. [S27], giving the state transfer time mentioned above. This time is faster than $O(L)$, the time it would take for the nearest-neighbor case, when $\alpha < D + 1$. Therefore, for 2D or higher and $\alpha \geq D/2$, the total time it takes to implement a hard-to-simulate circuit is $\min[L, L^{\alpha-D} \log L] + O(1)$, proving Theorem 6 for interacting bosons. When $\alpha < D/2$, the limiting step is dominated by the entangling gate, which takes time $O(1)$. Therefore for this case we only get fast hardness through boson sampling, which is discussed in Section IV. Note that when $t = o(1)$ and interaction strength is $V = \Theta(1)$, the effect of the interaction is governed by $Vt = o(1)$, which justifies treating the problem for short times as a free-boson problem.

A. One dimension

In 1D with nearest-neighbor hopping, we cannot hope to get a hardness result for simulating constant depth circuits, which is related to the fact that one cannot have universal measurement-based quantum computing in one dimension. We change our strategy here. The overall goal in 1D is to still be able to simulate the scheme in Ref. [S4] since it provides a faster hardness time (at the cost of an overhead in the qubits). The way this is done is to either (i) implement $O(n)$ SWAPs in 1D in order to implement an IQP circuit [S23], or (ii) use the long-range hops to directly implement gates between logical qubits at a distance L away.

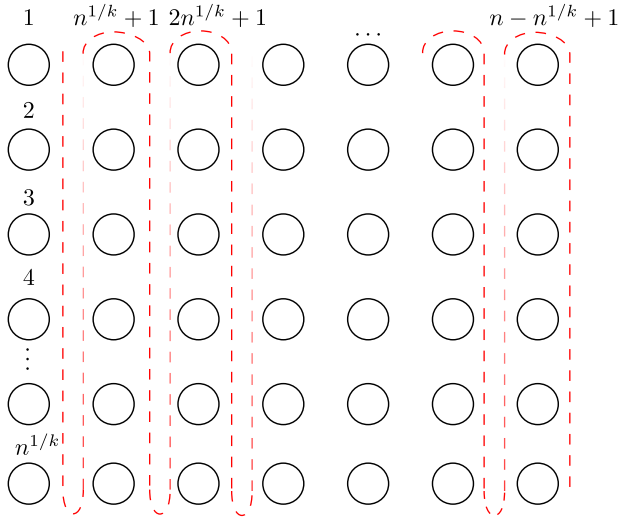


FIG. S6 (Color online). A snaking scheme to assign indices to qubits in 2D for a $n^{1/k} \times n^{1-1/k}$ grid, which is used in mapping to 1D.

For the first method, we use state transfer to implement a SWAP by moving each boson within a cluster a distance $\Theta(L)$. This takes time $O(t_s(L))$, where $t_s(L)$ is the time taken for state transfer over a distance L and is given by

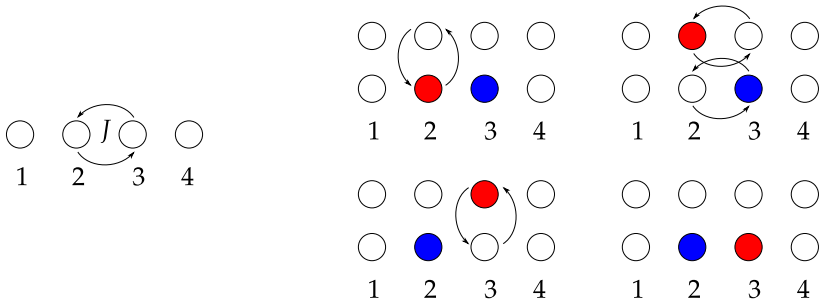
$$t_s(L) = c \times \begin{cases} L, & \alpha > 2 \\ L^{\alpha-1}, & \alpha \in (1,2] \\ \log L, & \alpha = 1 \\ 1, & \alpha \in [\frac{1}{2}, 1) \\ L^{\alpha-1/2}, & \alpha < \frac{1}{2}. \end{cases} \tag{S28}$$

We write this succinctly as $O\left(\min[L, L^{\alpha-1}\log L + 1, L^{\alpha-1/2}]\right)$. The total time for n SWAPs is therefore $O\left(n \times \min[L, L^{\alpha-1}\log L + 1, L^{\alpha-1/2}]\right)$.

The second method relies on the observation that when $\alpha \to 0$, the distinction between 1D and 2D becomes less clear, since at $\alpha = 0$, the connectivity is described by a complete graph and all hopping strengths are equal. Let us give some intuition for the $\alpha \to 0$ case. One would directly "sculpt" a 2D grid from the available graph, which is a complete graph on n vertices (one for every logical qubit) with weights w_{ij} given by $d(i,j)^{-\alpha}$. If we want to arrange qubits on a 1D path, we can assign an indexing to qubits in the 2D grid and place them in the 1D path in increasing order of their index. One may, in particular, choose a "snake-like indexing" depicted in Fig. S6. This ensures that nearest-neighbor gates along one axis of the 2D grid map to nearest-neighbor gates in 1D. Gates along the other axis, however, correspond to nonlocal gates in 1D. Suppose that the equivalent grid in 2D is of size $n^{1/k} \times n^{1-1/k}$. The distance between two qubits that have to participate in a gate is now marginally larger ($O(Ln^{1/k})$ instead of $O(L)$), but the depth is greatly reduced: it is now $O(n^{1/k})$ instead of $O(n)$. We again use state transfer to move close to a far-off qubit and then perform a nearest-neighbor entangling gate. This time is set by the state transfer protocol, and is now $t_s(n^{1/k}L) = O\left(n^{1/k} \times \min[L, L^{\alpha-1}\log L + 1, L^{\alpha-1/2}]\right)$. For large $k = \Theta(1)$, this gives us the bound $O\left(\min[L^{1+\delta}, L^{\alpha-1+\delta} + L^{\Theta(\delta)}, L^{\alpha-1/2+\delta}]\right)$ for any $\delta > 0$, giving a coarse transition. Notice, however, that faster hardness in 1D comes at a high cost– the effective number of qubits on which we implement a hard circuit is only $\Theta(n^{1/k}) = n^{\Theta(\delta)}$, which approaches a constant as $\delta \to 0$.

This example of 1D is very instructive– it exhibits one particular way in which the complexity phase transition can happen. As we take higher and higher values of k, the hardness time would decrease, coming at the cost of a decreased number of effective qubits. This smoothly morphs into the easiness regime when $\alpha \to \infty$ since in this regime both transitions happen at $t = \Theta(L)$.

If the definition of hardness is more stringent (in order to link it to fine-grained complexity measures such as explicit quantitative lower bound conjectures), then the above mentioned overhead is undesirable. In this case we would adopt the first strategy to implement SWAPs and directly implement a random IQP circuit on all the n qubits. This would increase the hardness time by a factor n.



(a) (b)

FIG. S7 (Color online). (a) A hopping between sites 2 and 3 that implements the mode unitary $\begin{pmatrix} \cos(|J|t) & -i\sin(|J|t)J/|J| \\ -i\sin(Jt)J^*/|J| & \cos(|J|t) \end{pmatrix} = e^{-it(\mathrm{Re}\{J\}X+\mathrm{Im}\{J\}Y)}$. When $|J|t = \pi$, this is a SWAP between two modes with phases $(-iJ/|J|, -iJ^*/|J|)$ that depends on $\arg J$, the argument of J. (b) A "physical" SWAP between sites 2 and 3 by using ancilla sites available whenever the system is not nearest-neighbor in 1D. The colors are used to label the modes and how they move, and do not mean that both sites are occupied. The total hopping phase incurred when performing the physical SWAP can be set to be $(+i, -i)$, which cannot be achieved with just the hopping term shown in (a).

B. Hardcore limit

In the hardcore limit $V \to \infty$, the strategy is modified. Let us consider a physical qubit to represent the presence ($|1\rangle$) or absence ($|0\rangle$) of a boson at a site. A nearest-neighbor hop translates to a term in the Hamiltonian that can be written in terms of the Pauli operators as $XX + YY$. Further, an on-site field $J_{ii}a_i^\dagger a_i$ translates to a term $\propto Z$. There are no other terms available, in particular single-qubit rotations about other axes X or Y. This is because the total boson number is conserved, which in the spin basis corresponds to the conservation of $\sum_i Z_i$. This operator indeed commutes with both the allowed Hamiltonian terms specified above.

Let us now discuss the computational power of this model. When the physical qubits are constrained to have nearest-neighbor interactions in 1D, this model is nonuniversal and classically simulable. This can be interpreted due to the fact that this model is equivalent to matchgates on a path (i.e. a 1D nearest-neighbor graph), which is nonuniversal for quantum computing without access to a SWAP gate. Alternatively, one can apply the Jordan-Wigner transformation to map the spin model onto free fermions. One may then use the fact that fermion sampling is simulable on a classical computer [S28].

When the connectivity of the qubit interactions is different, the model is computationally universal for BQP. In the matchgate picture, this result follows from Ref. [S29], which shows that matchgates on any graph apart from the path or the cycle are universal for BQP in an encoded sense. In the fermion picture, the Jordan-Wigner transformation on any graph other than a path graph would typically result in nonlocal interacting terms that are not quadratic in general. Thus, the model cannot be mapped to free, quadratic fermions and the simulability proof from Ref. [S28] breaks down.

Alternatively, a constructive way of seeing how we can recover universality is as follows. Consider again the dual rail encoding and two logical qubits placed next to each other as in Fig. S7. Apply a coupling $J(a_2^\dagger a_3 + a_3^\dagger a_2)$ on the modes 2 and 3 for time $t = \frac{\pi}{2J}$. This effects the transition $|10\rangle_{23} \to -i\,|01\rangle_{23}$ and $|01\rangle_{23} \to -i\,|10\rangle_{23}$, while leaving the state $|11\rangle_{23}$ the same. Now we swap the modes 2 and 3 using an ancilla mode that is available by virtue of having either long-range hopping or having $D > 1$. This returns the system back to the logical subspace of exactly one boson in modes 1 & 2, and one boson in modes 3 & 4, and effects a unitary locally equivalent to $\mathrm{diag}\{1, 1, 1, -1\}$ in the (logical) computational basis. This is an entangling gate that can be implemented in $O(1)$ time and thus the hardness timescale for hardcore interactions is the same as that of Hubbard interactions with $V = \Theta(1)$.

We finally discuss the case when V is polynomially large. Using the dual-rail encoding and implementing the same protocol as the non-hardcore case now takes the state $|11\rangle_{23}$ to $\lambda\,|11\rangle_{23} + \mu\frac{|20\rangle_{23}+|02\rangle_{23}}{\sqrt{2}}$, with $\mu \propto \frac{J}{\sqrt{8J^2+V^2}} \sin\left(\frac{t\sqrt{8J^2+V^2}}{2}\right)$. When $|\mu| \neq 0$, we get an error because the state is outside the logical subspace. The probability with which this action happens is suppressed by $1/V^2$, however, which is polynomially small when $V = \mathsf{poly}(n)$.

However, one can do better: by carefully tuning the hopping strength $J \in [0, 1]$ and the evolution time t, one can always achieve the goal of getting $\mu = 0$ exactly and implementing an operation $\exp\left[-i\frac{\pi}{2}X\right]$ in the $|10\rangle_{23}, |01\rangle_{23}$ subspace. This requires setting $t\sqrt{2J^2 + \frac{V^2}{4}} = m\pi$ and $t = \frac{2\pi}{J}$ for integer m. This can be solved as follows: set $m = \lceil\sqrt{8+V^2}\rceil$, and $J = \frac{V}{\sqrt{m^2-8}}$ (which is ≤ 1 since $m \geq \sqrt{8+V^2}$). The time is set by the condition $t = \frac{2\pi}{J}$, which is $\Theta(1)$. This effects a logical CPHASE$[\phi]$ gate with angle $\phi = -\pi V/J$.

Finally, the above parameters that set μ exactly to zero work even for exponentially large $V = \Omega(\exp(n))$, but this requires exponentially precise control of the parameters J and t, which may not be physically feasible. In this case, we simply observe

that $|\mu|^2$, the probability of going outside the logical subspace and hence making an error, is $O(1/V^2)$, which is exponentially small in n. Therefore, in this limit, the gate we implement is exponentially close to perfect, and the complete circuit has a very small infidelity as well.

S6. HARDNESS TIMESCALE FOR FREE BOSONS

In this section, we review Aaronson and Arkhipov's method of creating a linear optical state that is hard to sample from [S30]. We then give a way to construct such states in time $\tilde{O}\left(nm^{\alpha/D-1/2}\right)$ with high probability in the Hamiltonian model, and prove Theorem 6 for free bosons.

For free bosons, in order to get a state that is hard to sample from, we need to apply a Haar-random linear-optical unitary on m modes to the state $|1,1,\ldots 1,0,0,\ldots 0\rangle$. Aaronson and Arkhipov gave a method of preparing the resulting state in $O(n\log m)$ depth in the circuit model. Their method involves the use of ancillas and can be thought of as implementing each column of the Haar-random unitary separately in $O(\log m)$-depth. Here we mean that we apply the map $|1\rangle_j \to \sum_{i\in\Lambda} U_{ij}|1\rangle_i$ to "implement" the column i of the linear-optical unitary U. In the Hamiltonian model, we can apply simultaneous, non-commuting terms of a Hamiltonian involving a common site. The only constraint is that each term of the Hamiltonian should have a bounded norm of $1/d(i,j)^\alpha$. In this model, when α is small, it is possible to implement each unitary in a time much smaller than $O(\log m)$– indeed, we show the following:

Lemma 7. *Let U be a Haar-random unitary on m modes. Then with probability $1-\frac{1}{\mathrm{poly}(m)}$ over the Haar measure, each of the first n columns of U can be implemented in time $O\left(\frac{\sqrt{\log m}}{m^{1/2-\alpha/D}}\right)$.*

To prove this, we will need an algorithm that implements columns of the unitary. For convenience, let us first consider the case $\alpha = 0$. The algorithm involves two subroutines, which we call the single-shot and state-transfer protocols. Both protocols depend on the following observation. If we implement a Hamiltonian that couples a site i to all other sites $j\neq i$ through coupling strengths J_{ij}, then the effective dynamics is that of two coupled modes $a_i^\dagger$ and $b^\dagger = \frac{1}{\omega}\sum_{j\neq i} J_{ij}a_j^\dagger$, where $\omega = \sqrt{\sum_{j\neq i} J_{ij}^2}$. The effective speed of the dynamics is given by ω– for instance, the time period of the system is $\frac{2\pi}{\omega}$.

The single-shot protocol implements a map $a_i^\dagger \to \gamma_i a_i^\dagger + \sum_{j\neq i}\gamma_j a_j^\dagger$. This is done by simply applying the Hamiltonian $H \propto a_i^\dagger(\sum_{j\neq i}\gamma_j a_j) + \text{h.c.}$ for time $t = \frac{1}{\omega}\cos^{-1}|\gamma_i|$. In the case $\alpha = 0$, we can set the proportionality factor equal to $1/\max|\gamma_j|$. This choice means that the coupling strength between i and the site k with maximum $|\gamma_k|$ is set to 1 (the maximum), and all other couplings are equal to $|\frac{\gamma_j}{\gamma_k}|$.

The other subroutine, the state-transfer protocol is also an application of the above observation and appears in Ref. [S26]. It achieves the map $a_i^\dagger \to \gamma_i a_i^\dagger + \gamma_j a_j^\dagger$ via two rounds of the previous protocol. This is done by first mapping site i to the uniform superposition over all sites except i and j, and then coupling this uniform superposition mode to site j. The time taken for this is $\frac{1}{\omega}\left(\frac{\pi}{2}+\cos^{-1}|\gamma_i|\right)$. Since $\omega = \sqrt{m-2}$ (all $m-2$ modes are coupled with equal strength to modes i or j), this takes time $O\left(\frac{1}{\sqrt{m}}\right)$.

These subroutines form part of Algorithm 1. It can be seen that Algorithm 1 implements a map $a_j^\dagger \to U_{jj}a_j^\dagger + \sum_{i\neq j} U_{ij}a_i^\dagger$, as

Algorithm 1: Algorithm for implementing one column of a unitary

Input: Unitary U, column index j

1. Reassign the mode labels for modes $i\neq j$ in nonincreasing order of $|U_{ij}|$.
2. Implement the state-transfer protocol to map the state $a_j^\dagger|\text{vac}\rangle$ to $U_{jj}a_j^\dagger|\text{vac}\rangle + \sqrt{1-|U_{jj}|^2}a_1^\dagger|\text{vac}\rangle$. Skip this step if $|U_{jj}|\geq|U_{j1}|$ already.
3. Use the single-shot protocol between site 1 and the rest ($i\neq 1,j$) to map $a_1^\dagger \to \frac{U_{1j}}{\sqrt{1-|U_{jj}|^2}}a_1^\dagger + \sum_{i\neq 1,j}\frac{U_{ij}}{\sqrt{1-|U_{jj}|^2}}a_i^\dagger$.

desired. To prove Lemma 7 we need to examine the runtime of the algorithm when U is drawn from a Haar-random distribution.

Proof of Lemma 7. First, notice that since the Haar measure is invariant under the action of any unitary, we can in particular apply a permutation map to argue that the elements of the i'th column are drawn from the same distribution as the first column. Next, recall that one may generate a Haar-random unitary by first generating m uniform random vectors in $\mathbb{C}^m$ and then performing a Gram-Schmidt orthogonalization. In particular, this means that the first column of a Haar-random unitary may be generated by generating a uniform random vector with unit norm. This implies that the marginal distribution over any column of a unitary drawn from the Haar measure is simply the uniform distribution over unit vectors, since we argued above that all columns are drawn from the same distribution.

Now, let us examine the runtime. The first step (line 2 of the algorithm) requires time $t = O\left(\frac{1}{\sqrt{m}}\right)$ irrespective of U_{jj} because the total time for state-transfer is $\frac{1}{\omega}\left(\frac{\pi}{2} + \cos^{-1} U_{jj}\right) \leq \frac{\pi}{\omega} = \frac{\pi}{\sqrt{m-2}}$. Next, the second step takes time $t = \frac{1}{\omega}\cos^{-1}\left(\frac{U_{1j}}{\sqrt{1-|U_{1j}^2|}}\right) = O(\frac{1}{\omega})$. Now,

$$\omega = \sqrt{1^2 + \frac{|U_{3j}|^2/(1-|U_{jj}|^2)}{|U_{2j}|^2/(1-|U_{jj}|^2)} + \frac{|U_{4j}|^2}{|U_{2j}|^2} + \dots} \tag{S29}$$

$$= \sqrt{\frac{\sum_{i=2, i\neq j}^{m} |U_{ij}|^2}{|U_{2j}|^2}} = \sqrt{\frac{1-|U_{1j}|^2-|U_{jj}|^2}{|U_{2j}|^2}} \tag{S30}$$

Now in cases where $|U_{jj}| \leq |U_{1j}|$ (where $|U_{1j}|$ is the maximum absolute value of the column entry among all other modes $i \neq j$), which happens with probability $1 - \frac{1}{m}$, we will have $\omega^2 \geq \frac{1-2|U_{1j}|^2}{|U_{2j}|^2}$. In the other case when $|U_{jj}| \geq |U_{1j}|$, meaning that the maximum absolute value among all entries of column j is in row j itself, we again have $\omega^2 \geq \frac{1-2|U_{jj}|^2}{|U_{2j}|^2}$. Both these cases can be written together as $\omega^2 \geq \frac{1-2|U_{1j}|^2}{|U_{2j}|^2}$, where we now denote U_{1j} as the entry with maximum absolute value among *all* elements of column j. The analysis completely hinges on the typical ω we have, which in turn depends on $|U_{1j}|$. We will show $\Pr\left(\omega^2 \geq \frac{cm}{\log m}\right) \geq 1 - \frac{1}{\mathsf{poly}(m)}$, which will prove the claim for $\alpha = 0$.

$$\Pr\left(\omega^2 \geq \frac{cm}{\log m}\right) \geq \Pr\left(1 - 2|U_{1j}|^2 \geq c_1 \;\&\; |U_{2j}|^2 \leq \frac{c_1 \log m}{cm}\right) \tag{S31}$$

since the two events on the right hand side suffice for the first event to hold. Further,

$$\Pr\left(1 - 2|U_{1i}|^2 \geq c_1 \;\&\; |U_{2j}|^2 \leq \frac{c_1 \log m}{cm}\right) \geq \Pr\left(|U_{1j}|^2 \leq \frac{c_1 \log m}{cm}\right) \tag{S32}$$

for large enough m with some fixed $c_1 = 0.99$ (say), since $|U_{2j}|^2 \leq |U_{1j}|^2$ and $1 - 1.98 \log m/m \geq 0.99$ for large enough m.

To this end, we refer to the literature on order statistics of uniform random unit vectors $(z_1, z_2, \dots z_m) \in \mathbb{C}^m$ [S31]. This work gives an explicit formula for $F(x, m)$, the probability that all $|z_j|^2 \leq x$. We are interested in this quantity at $x = c_1 \log m/(cm)$, since this gives us the probability of the desired event ($\omega^2 \geq cm/\log m$). We have

$$\Pr\left(\frac{1}{k+1} \leq x \leq \frac{1}{k}\right) = \sum_{l=0}^{k} \binom{m}{l} (-1)^l (1-lx)^{m-1}. \tag{S33}$$

It is also argued in Ref. [S31] that the terms of the series successively underestimate or overestimate the desired probability. Therefore we can expand the series and terminate it at the first two terms, giving us an inequality:

$$\Pr\left(\frac{1}{k+1} \leq x \leq \frac{1}{k}\right) = 1 - m(1-x)^{m-1} + \frac{m^2}{2}(1-2x)^{m-1} - \dots \tag{S34}$$

$$\geq 1 - m(1-x)^{m-1}. \tag{S35}$$

Choosing $c = c_1/4 = 0.2475$, we are interested in the quantity when $k = \lfloor \frac{m}{4 \log m} \rfloor$:

$$\Pr(x \leq 4\log m/m) \geq 1 - m(1 - 4\log m/m)^{m-1} \geq 1 - \frac{1}{m^{3-4/m}}, \text{since} \tag{S36}$$

$$(1 - 4\log m/m)^{m-1} = \exp\left[(m-1)\log\left(1 - \frac{4\log m}{m}\right)\right] \leq \exp\left[-4(m-1)\frac{\log m}{m}\right] = m^{-4(1-1/m)}. \tag{S37}$$

This implies that the time for the single-shot protocol is also $t = O(\frac{1}{\omega}) = O(\sqrt{\frac{\log m}{m}})$ for a single column. Notice that we can make the polynomial appearing in $\Pr\left(\omega^2 \geq cm/\log m\right) \geq 1 - 1/\mathsf{poly}(m)$ as small as possible by suitably reducing c. To extend the proof to all columns, we use the union bound. In the following, let t_j denote the time to implement column j.

$$\Pr\left(\exists j : t_j > \sqrt{\frac{\log m}{cm}}\right) \leq \sum_j \Pr\left(t_j > \sqrt{\frac{\log m}{cm}}\right) \tag{S38}$$

$$\leq m \times \frac{1}{\mathsf{poly}(m)} = \frac{1}{\mathsf{poly}(m)} \tag{S39}$$

when the degree in the polynomial is larger than 1, just as we have chosen by setting $c = 0.2475$. This implies

$$\Pr\left(\forall j : t_j \leq \sqrt{\frac{\log m}{cm}}\right) = 1 - \Pr\left(\exists j : t_j > \sqrt{\frac{\log m}{cm}}\right) \geq 1 - \frac{1}{\mathsf{poly}(m)}. \tag{S40}$$

This completes the proof in the case $\alpha = 0$. When $\alpha \neq 0$, we can in the worst-case set each coupling constant to a maximum of $O(m^{-\alpha/D})$, which is the maximum coupling strength of the furthest two sites separated by a distance $O(m^{1/D})$. This factor appears in the total time for both the state-transfer [S26] and single-shot protocols, and simply multiplies the required time, making it $O\left(\sqrt{\frac{\log m}{m}} \times m^{\alpha/D}\right) = O\left(\frac{\sqrt{\log m}}{m^{1/2-\alpha/D}}\right)$. Finally, if there are any phase shifts that need to be applied, they can be achieved through an on-site term $J_{ii}a_i^\dagger a_i$, whose strength is unbounded by assumption and can thus take arbitrarily short time. □

The total time for implementing boson sampling on n bosons is therefore $O\left(n\frac{\sqrt{\log m}}{m^{1/2-\alpha/D}}\right) = \tilde{O}\left(n^{1+\beta(\frac{\alpha}{D}-\frac{1}{2})}\right)$, since we should implement n columns in total.

A. Optimizing hardness time

We can optimize the hardness time by implementing boson sampling not on n bosons, but on n^δ of them, for any $\delta \in (0, 1]$. The explicit lower bounds on running time of classical algorithms we would get assuming fine-grained complexity-theoretic conjectures is again something like $\exp\left[n^{\mathsf{poly}(\delta)}\right]$ for any $\delta \in (0, 1]$. This grows very slowly with n, but it still qualifies as subexponential, which is not polynomial or quasipolynomial (and, by our definition, would fall in the category "hard"). This choice of parameters allows us to achieve a smaller hardness timescale at the cost of getting a coarse (type-II) transition. We analyze this idea in three cases: $\alpha \leq D/2$, $\alpha \in (\frac{D}{2}, D]$ and $\alpha > D$.

When $\alpha \leq D/2$, we perform boson sampling on the nearest set of n^δ bosons with the rest of the empty sites in the lattice as target sites. In terms of the linear optical unitary, the unitary acts on $m - n^\delta = \Theta(m)$ sites in the lattice, although only the n^δ columns corresponding to initially occupied sites are relevant. Using the protocol in Lemma 7, the total time to implement n^δ columns of an $m \times m$ linear optical unitary is $O(n^\delta m^{\alpha/D-1/2} \log n) = \tilde{O}(n^\delta n^{\frac{\beta}{D}(\alpha-D/2)})$.

When $\alpha \in (\frac{D}{2}, D]$, the strategy is modified. We first move the nearest set of n^δ bosons into a contiguous set of sites within a single cluster. This takes time $O(n^\delta)$, since each boson may be transferred in time $O(1)$. We now perform boson sampling on these n^δ bosons with the surrounding $n^{2\delta}$ sites as targets, meaning that the effective number of total sites is $m_{\text{eff}} = O(n^{2\delta})$, as required for the hardness of boson sampling. Applying Lemma 7, the time required to perform hard instances of boson sampling is now $O(n^\delta n^{2\delta(\alpha/D-1/2)} \log n) = n^{O(\delta)}$ for arbitrarily small $\delta > 0$.

Lastly, when $\alpha > D$, we use the same protocol as above. The time taken for the state transfer is now $n^\delta \times \min[L, L^{\alpha-D}]$. Once state transfer has been achieved, we use nearest-neighbor hops instead of Lemma 7 to create an instance of boson sampling in time $O(n^{2\delta/D})$. Since state transfer is the limiting step, the total time is $n^\delta \times \min[L, L^{\alpha-D}]$. The hardness timescale is obtained by taking the optimum strategy in each case, giving the hardness timescale $t_{\text{hard}} = \tilde{O}(n^{\gamma^{\text{II}}_{\text{hard}}})$, where

$$\gamma^{\text{II}}_{\text{hard}} = \delta + \begin{cases} \frac{\beta-1}{D}\min[1, \alpha - D] & \alpha > D \\ 0 & \alpha \in (\frac{D}{2}, D] \\ \frac{\beta}{D}\left(\alpha - \frac{D}{2}\right) & \alpha < \frac{D}{2} \end{cases} \tag{S41}$$

for an arbitrarily small $\delta > 0$. This proves Theorem 6 for free bosons and for interacting bosons in the case $\alpha < D/2$. When we compare with Ref. [S32], which states a hardness result for $\alpha \to \infty$, we see that we have almost removed a factor of n from the timescale coming from implementing n columns of the linear optical unitary. Our result here gives a coarse hardness timescale of $\Theta(L)$ that matches the easiness timescale of L. More importantly, this makes the noninteracting hardness timescale the same as the interacting one.

B. Almost free bosons, $V = o(1)$

When the interaction strength satisfies $V = o(1)$ or when the bosons are almost free, we can treat the evolution as being close to that of free bosons. The total variation distance error between the actual distribution and the distribution modeled by free bosons can be upper bounded by $\left\| |\psi(t)\rangle - \left|\tilde{\psi}(t)\right\rangle \right\|_2 =: \delta(t)$, where we take H and $|\psi\rangle$ to be the actual Hamiltonian and state

and $\tilde{H}$ and $\left|\tilde{\psi}\right\rangle$ to be their respective free-bosonic approximations. Therefore, by the same logic leading up to Eq. (S11), we have the same expression here for $\delta(t)$:

$$\delta(t) \leq \int_0^t \mathrm{d}\tau \left\| \left(H(\tau) - \tilde{H}(\tau)\right) \left|\tilde{\psi}(\tau)\right\rangle \right\|_2 \tag{S42}$$

$$\leq \int_0^t \mathrm{d}\tau \left\| \sum_i f(n_i) \left|\tilde{\psi}(\tau)\right\rangle \right\|_2 . \tag{S43}$$

Just as before, we use the fact that at short times, the boson number in each cluster (and hence on each site) is bounded. Specifying to the case of Bose-Hubbard interactions, we have

$$\delta \leq \int_0^t \mathrm{d}\tau \frac{V}{2} \left\| \sum_i n_i(n_i-1) \left|\tilde{\psi}(\tau)\right\rangle \right\|_2 \tag{S44}$$

$$\leq \int_0^t \mathrm{d}\tau \frac{V}{2} \left\| \sum_i n_i(n_i-1) \right\| \left\| \left|\tilde{\psi}(\tau)\right\rangle \right\|_2 \tag{S45}$$

$$\leq O(Vtn^2). \tag{S46}$$

We observe that by taking $V = 1/\mathsf{poly}(n)$ for a sufficiently large polynomial, the quantity δ is polynomially small in n at the hardness time for free bosons. While this does not show that the hardness time is the same for all $V = O(1/\mathsf{poly}(n))$ or for all $V = o(1)$, it shows that the hardness time for free bosons is robust to perturbations. We thank one of the anonymous Referees for pointing out this case.

S7. TYPES OF TRANSITIONS

In this section, we study in more detail the number of encoded logical qubits in the system inherent in the hardness proof, which we define more carefully below. Our definition is valid for any family of Hamiltonians and the definition for circuit architectures is analogous.

Given constraints on the evolution time t (taken to be the depth in case of circuits), we would like to reduce the problem of simulating arbitrary circuits of nearest-neighbor gates in one dimension acting on q qubits for depth q to the problem of simulating postselected Hamiltonian evolution for time t under the given Hamiltonian. We call this reduction exploiting postselection an "embedding". For a given reduction, we define the quantity $q(t,n)$ to be the number of qubits m that we can embed into a Hamiltonian evolution on n qubits/particles evolving for time t. The motivation for defining this quantity is that if the function $q(t,n)$ is at least some (possibly sublinear) polynomial in n, then simulating time evolution for time t under the Hamiltonian is hard modulo widely-held conjectures in complexity theory.

A. Defining phase transitions

The function $q(t,n)$ with respect to a certain reduction is a non-decreasing function of the time t for a fixed n, since given more time, the number of qubits that can be embedded cannot decrease. At the transition timescale t_*, the quantity q transitions from being subpolynomial in n to polynomial in n as the system transitions from being easy to simulate to being hard to simulate. The 2D and 1D phase diagrams in the main text can also be viewed in terms of the function q of the corresponding family of Hamiltonians.

Figure S8 shows the behavior of $q(t,n)$ under our hardness reductions for the two types of transitions. We can see from the figure that if these reductions are optimal, sharp transitions are akin to I-order phase transitions, and coarse transitions akin to II-order transitions. Specifically, if the exponent of n in $q(t,n)$ has a discontinuity with respect to the exponent of n in the evolution time t, we have a I-order transition. For II-order transitions, on the other hand, there is no discontinuity in the exponent of n in $q(t,n)$. The exponent may be defined as $\lim_{n\to\infty}(\log q(t,n)/\log n)$. Therefore, the exponent and hence the transition are only well-defined in the thermodynamic limit $n \to \infty$, just as in regular phase transitions where non-analyticities are only visible in the thermodynamic limit.

Let us look at how we obtained the lines in Fig. S8. First, the region before exponent $\frac{\beta-1}{D}$ corresponding to time $o(L)$ is easy in all dimensions from our easiness results, meaning q necessarily scales slower than any polynomial in n in the blue region. Also, we know from our hardness results that near β/D (i.e. time $t = O(Ln^{1/D})$) we have enough time to "touch" all n qubits, giving $q(t,n) = n$ for all dimensions. This can be extended to hardness for $q(t,n) = n^\delta$ for all $\delta \in (0,1]$ in all dimensions,

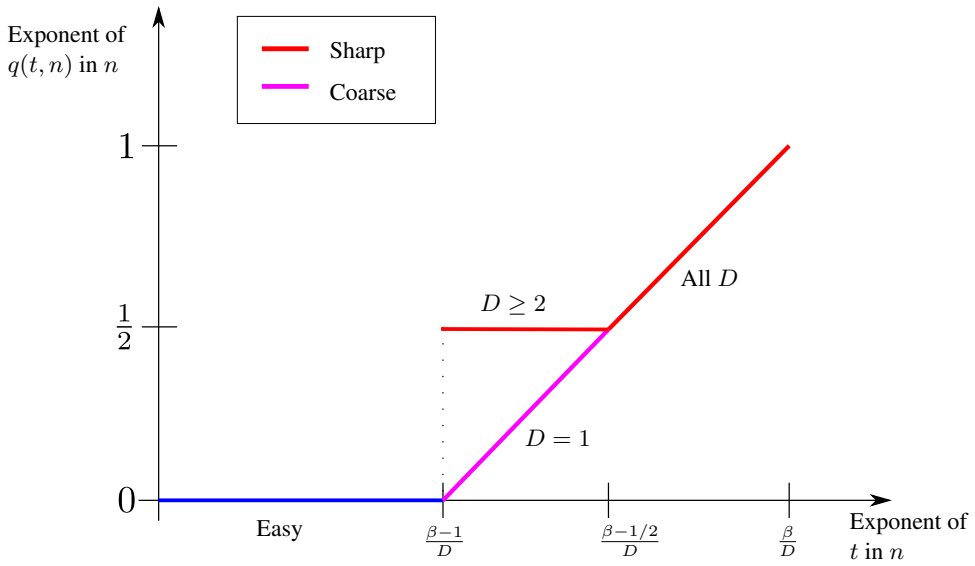


FIG. S8 (Color online). Schematic of sharp and coarse complexity phase transitions for interacting bosons. The points on the X-axis, $(\beta-1)/D$ and β/D, are different for different dimensions and therefore the lines do not really intersect in the way depicted here.

with the corresponding time $Ln^{\delta/D} = n^{(\beta-1+\delta)/D}$. For $D \geq 2$ on the other hand, we can do better: we can always encode $\sqrt{n}$ qubits even if the time is $\Theta(L)$. This strategy is better than the first strategy for time $t < n^{(\beta-1/2)/D}$. It is not known if these curves for the function q are optimal, since for optimality we should rule out other reductions that might achieve a better scaling.

To sum up, considering the quantity $q(t,n)$ gives a complementary way of looking at complexity phase transitions of the sort we study. This is a fine-grained view of looking at complexity phase transitions, since we care not just about whether $q(t,n)$ is polylog (easy) or poly (hard), but how exactly it scales with n as a function of the evolution time.

[S1] F. Verstraete and J. I. Cirac, Mapping local Hamiltonians of fermions to local Hamiltonians of spins, J. Stat. Mech. **2005**, P09012 (2005).

[S2] G. Muraleedharan, A. Miyake, and I. H. Deutsch, Quantum computational supremacy in the sampling of bosonic random walkers on a one-dimensional lattice, New J. Phys. **21**, 055003 (2018).

[S3] M. A. Norcia, A. W. Young, and A. M. Kaufman, Microscopic Control and Detection of Ultracold Strontium in Optical-Tweezer Arrays, Phys. Rev. X **8**, 041054 (2018).

[S4] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert, Architectures for Quantum Simulation Showing a Quantum Speedup, Phys. Rev. X **8**, 021010 (2018).

[S5] J. S. Douglas, H. Habibian, C.-L. Hung, A. V. Gorshkov, H. J. Kimble, and D. E. Chang, Quantum many-body models with cold atoms coupled to photonic crystals, Nat. Photonics **9**, 326 (2015).

[S6] S.-K. Chu, G. Zhu, J. R. Garrison, Z. Eldredge, A. V. Curiel, P. Bienias, I. B. Spielman, and A. V. Gorshkov, Scale-Invariant Continuous Entanglement Renormalization of a Chern Insulator, Phys. Rev. Lett. **122**, 120502 (2019).

[S7] B. Gadway, An atom optics approach to studying lattice transport phenomena, Phys. Rev. A **92**, 043606 (2015).

[S8] G. M. Crosswhite, A. C. Doherty, and G. Vidal, Applying matrix product operators to model systems with long-range interactions, Phys. Rev. B **78**, 035116 (2008).

[S9] M. Saffman, T. G. Walker, and K. Mølmer, Quantum information with Rydberg atoms, Rev. Mod. Phys. **82**, 2313 (2010).

[S10] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Probing many-body dynamics on a 51-atom quantum simulator, Nature **551**, 579 (2017).

[S11] D. Barredo, S. de Léséleuc, V. Lienhard, T. Lahaye, and A. Browaeys, An atom-by-atom assembler of defect-free arbitrary two-dimensional atomic arrays, Science **354**, 1021 (2016).

[S12] S. Korenblit, D. Kafri, W. C. Campbell, R. Islam, E. E. Edwards, Z.-X. Gong, G.-D. Lin, L.-M. Duan, J. Kim, K. Kim, and C. Monroe, Quantum simulation of spin models on an arbitrary lattice with trapped ions, New J. Phys. **14**, 095024 (2012).

[S13] R. Islam, C. Senko, W. C. Campbell, S. Korenblit, J. Smith, A. Lee, E. E. Edwards, C.-C. J. Wang, J. K. Freericks, and C. Monroe, Emergence and Frustration of Magnetism with Variable-Range Interactions in a Quantum Simulator, Science **340**, 583 (2013).

[S14] A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, V. Khemani, J. Léonard, and M. Greiner, Probing entanglement in a many-body–localized system, Science **364**, 256 (2019).

[S15] D. Bluvstein, H. Levine, G. Semeghini, T. T. Wang, S. Ebadi, M. Kalinowski, A. Keesling, N. Maskara, H. Pichler, M. Greiner, V. Vuletic, and M. D. Lukin, A quantum processor based on coherent transport of entangled atom arrays, Nature **604**, 451 (2022).

[S16] M. Foss-Feig, Z.-X. Gong, C. W. Clark, and A. V. Gorshkov, Nearly Linear Light Cones in Long-Range Interacting Quantum Systems, Phys. Rev. Lett. **114**, 157201 (2015).

[S17] M. B. Hastings and T. Koma, Spectral Gap and Exponential Decay of Correlations, Commun. Math. Phys. **265**, 781 (2006).

[S18] T. J. Osborne, Efficient Approximation of the Dynamics of One-Dimensional Quantum Spin Systems, Phys. Rev. Lett. **97**, 157202 (2006).

[S19] A. M. Childs, D. Gosset, and Z. Webb, Universal Computation by Multiparticle Quantum Walk, Science **339**, 791 (2013).

[S20] M. Oszmaniec and Z. Zimborás, Universal Extensions of Restricted Classes of Quantum Operations, Phys. Rev. Lett. **119**, 220502 (2017).

[S21] M. S. Underwood and D. L. Feder, Bose-Hubbard model for universal quantum walk-based computation, Phys. Rev. A **85**, 052314 (2012).

[S22] A. Bouland and M. Ozols, Trading Inverses for an Irrep in the Solovay-Kitaev Theorem, in *Proc. 13th Conf. Theory Quantum Comput. Commun. Cryptogr. TQC 2018*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 111 (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2018) pp. 6:1–6:15.

[S23] M. J. Bremner, R. Jozsa, and D. J. Shepherd, Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy, Proc. R. Soc. Math. Phys. Eng. Sci. **467**, 459 (2011).

[S24] M. J. Bremner, A. Montanaro, and D. J. Shepherd, Average-case complexity versus approximate simulation of commuting quantum computations, Phys. Rev. Lett. **117**, 080501 (2016).

[S25] Z. Eldredge, Z.-X. Gong, J. T. Young, A. H. Moosavian, M. Foss-Feig, and A. V. Gorshkov, Fast Quantum State Transfer and Entanglement Renormalization Using Long-Range Interactions, Phys. Rev. Lett. **119**, 170503 (2017).

[S26] A. Y. Guo, M. C. Tran, A. M. Childs, A. V. Gorshkov, and Z.-X. Gong, Signaling and Scrambling with Strongly Long-Range Interactions, (2019), arXiv:1906.02662 [quant-ph].

[S27] M. C. Tran, C.-F. Chen, A. Ehrenberg, A. Y. Guo, A. Deshpande, Y. Hong, Z.-X. Gong, A. V. Gorshkov, and A. Lucas, Hierarchy of linear light cones with long-range interactions, Phys. Rev. X **10**, 031009 (2020).

[S28] B. M. Terhal and D. P. DiVincenzo, Classical simulation of noninteracting-fermion quantum circuits, Phys. Rev. A **65**, 032325 (2002).

[S29] D. J. Brod and A. M. Childs, The computational power of matchgates and the XY interaction on arbitrary graphs, Quantum Inf. Comput. **14**, 0901 (2013).

[S30] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, in *Proc. Forty-Third Annu. ACM Symp. Theory Comput.* (ACM Press, New York, New York, USA, 2011) p. 333.

[S31] A. Lakshminarayan, S. Tomsovic, O. Bohigas, and S. N. Majumdar, Extreme statistics of complex random and quantum chaotic states, Phys. Rev. Lett. **100**, 044103 (2008).

[S32] A. Deshpande, B. Fefferman, M. C. Tran, M. Foss-Feig, and A. V. Gorshkov, Dynamical Phase Transitions in Sampling Complexity, Phys. Rev. Lett. **121**, 030501 (2018).