Entanglement and spin-squeezing without infinite-range interactions

Michael Foss-Feig,[1, 2, 3] Zhe-Xuan Gong,[2, 3] Alexey V. Gorshkov,[2, 3] and Charles W. Clark[2]

[1]*United States Army Research Laboratory, Adelphi, MD 20783, USA*
[2]*Joint Quantum Institute, NIST/University of Maryland, College Park, MD 20742 USA*
[3]*Joint Center for Quantum Information and Computer Science,*
NIST/University of Maryland, College Park, MD 20742 USA

Infinite-range interactions are known to facilitate the production of highly entangled states with applications in quantum information and metrology. However, many experimental systems have interactions that decay with distance, and the achievable benefits in this context are much less clear. Combining recent exact solutions with a controlled expansion in the system size, we analyze quench dynamics in Ising models with power-law ($1/r^\alpha$) interactions in D dimensions, thereby expanding the understanding of spin squeezing into a broad and experimentally relevant context. In spatially homogeneous systems, we show that for small α the scaling of squeezing with system size is *identical* to the infinite-range ($\alpha = 0$) case. This indifference to the interaction range persists up to a critical value $\alpha = 2D/3$, above which squeezing degrades continuously. Boundary-induced inhomogeneities present in most experimental systems modify this picture, but it nevertheless remains qualitatively correct for finite-sized systems.

The dynamical growth of entanglement in far-from-equilibrium quantum systems makes them attractive for numerous applications in quantum information science, but greatly complicates their description. For sufficiently short-ranged interactions, entanglement growth after a quench is constrained to within a light cone [1], often allowing for efficient numerical simulations [2, 3]. On the other extreme, systems with infinite-range (i.e. distance independent) interactions develop entanglement on all length scales simultaneously, but are nevertheless easily characterized in terms of a small set of collective variables [4]. The intermediate territory of long but *not* infinite-ranged interactions is comparatively uncharted, and questions surrounding the growth of entanglement—and the associated computational complexity—in this regime have recently attracted significant theoretical [3, 5–10] and experimental [11–13] attention.

An important application of quench-induced entanglement is the generation of spin-squeezing [14], which has been sought as a means towards quantum-enhanced metrology [15–18], and more generally provides a theoretically tractable and experimentally accessible metric for characterizing entanglement in non-equilibrium systems [19–24]. Strategies to produce spin-squeezed states [22, 25–28] and the experiments that have pursued them [16–18, 29–31] are extremely diverse, but tend to share the idealization of infinite-range interactions and perfectly collective dynamics. Indeed, the difficulty of studying non-equilibrium systems with long (but not infinite) ranged interactions has obscured the extent to which squeezing is robust to relaxing this idealization. This question is, however, of immediate importance to many long-range interacting experimental systems, including polar molecules [11, 32], magnetic atoms [33], Rydberg atoms [34, 35], trapped ions [12, 13, 36, 37], and optical lattice clocks [38], in which interactions decay with distance.

To answer this question, we consider a generalization of the *single-axis twisting model* [14] where the spin-spin cou-
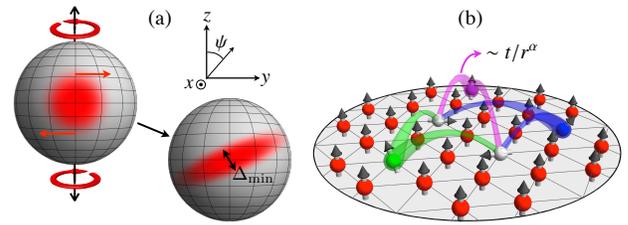

FIG. 1. (Color online) (a) Spin squeezing in the single-axis twisting model depicted on the Bloch sphere. The shaded areas represent the uncertainty due to spin projection noise. (b) A physical process contributing to the correlation between two spins (in white) via three intermediate spins. Each leg connecting a white spin to a colored one caries a factor t/r^α, with r the distance between the spins.

plings decay as a power law $1/r^\alpha$, and analytically compute how the amount of squeezing scales with the number of spins $\mathcal{N}$. Remarkably, we find that for spatially homogeneous systems, the scaling of squeezing with system size is *completely* unaffected by the decay of interactions for $\alpha < 2D/3$, with D the dimension of space. In the single-axis twisting model ($\alpha = 0$), quantum correlations between pairs of spins conspire to reduce spin projection noise [Fig. 1(a)] below the standard quantum limit by a multiplicative factor ξ, called the squeezing parameter. For $\alpha > 0$, the dynamics are not collective and the spin-state cannot be represented on a Bloch sphere, but the indifference of squeezing to interaction range can be understood by considering the physical processes that generate such correlations. An example is shown in Fig. 1(b), where correlations between two spins (in white) are mediated by three other spins (colored). Each leg represents one application of the Hamiltonian and hence caries a factor $\sim t/r^\alpha$, where t is time and r is the separation between the spins connected by that leg. For sufficiently small α we can restrict our attention to processes where the two white spins are minimally coupled—i.e. by exactly two legs—to each spin mediating their correlation, since these contributions to the correlation diverge most strongly with $\mathcal{N}$. Summing over the position of the mediating

spin we find a contribution $\sim t^2 \int d^D r / r^{2\alpha} \sim \mathcal{N} \times t^2 \mathcal{N}^{-2\alpha/D}$ (the final step here follows from dimensional analysis, and noting that the linear system size scales as $\mathcal{N}^{1/D}$). Because t always enters in this manner, all correlation functions can be written in terms of $\mathcal{N}$ and a rescaled time $\tau = t\mathcal{N}^{-\alpha/D}$. Therefore, even though ξ depends on $\mathcal{N}$, t and α, we can approximate it as a universal function $\xi(\mathcal{N}, \tau)$ of the system size and a rescaled time, with no reference to the range of interactions. It follows that, at least for sufficiently small α, the decay of interactions only affects the *time* at which the optimal squeezing occurs, but does not affect *how much* squeezing can be achieved. In what follows, this claim will be substantiated with an explicit calculation, and we will determine the critical value of α above which this simple picture breaks down. Before proceeding, we note that most experimental systems are not spatially homogeneous (owing to boundary effects), and for sufficiently large $\mathcal{N}$ such inhomogeneity will be seen to impose a finite minimal value for ξ that is independent of system size. Nevertheless, we will demonstrate that the qualitative picture just described persists in relatively large (but necessarily finite) systems.

Model and formalism.—We consider non-equilibrium spin dynamics generated by the Hamiltonian

$$H = \frac{J}{2} \sum_{i \neq j} \frac{\hat{\sigma}_i^z \hat{\sigma}_j^z}{|\mathbf{r}_i - \mathbf{r}_j|^\alpha}, \tag{1}$$

where indices i, j label sites of a regular lattice, which are located at positions $\mathbf{r}_j$. Note that this Hamiltonian reduces, in the limit $\alpha \to 0$, to the single-axis twisting model. The simplest protocol for generating spin-squeezing from the Hamiltonian in Eq. (1) is to initiate all of the spins pointing along a common direction in the xy plane; without loss of generality we choose the x-axis [Fig. 1(a)]. The variance of the total spin measured along a direction $\cos(\psi)\hat{y} + \sin(\psi)\hat{z}$ perpendicular to the x-axis is given by

$$\Delta_\psi(t)^2 = \frac{1}{4} \sum_{i,j} \langle (\hat{\sigma}_i^y \cos\psi + \hat{\sigma}_i^z \sin\psi)(\hat{\sigma}_j^y \cos\psi + \hat{\sigma}_j^z \sin\psi) \rangle,$$

where we have taken advantage of the fact that all single-spin expectation values in the yz plane vanish by symmetry. In the initial state the variance is isotropic and given by $\Delta_\psi(0) = \mathcal{N}/4$. At any later time t, the variance depends on ψ. Minimization of the variance with respect to the angle ψ at fixed t can be carried out without explicitly evaluating the correlation functions involved [14], yielding the squeezing parameter as a function of time [39],

$$\xi(t)^2 \equiv \mathcal{N}^{-1} \Delta_{\min}^2(t) = 1 + \mathcal{A}(t) - \sqrt{\mathcal{A}(t)^2 + \mathcal{B}(t)^2}, \tag{2}$$

where

$$\mathcal{A}(t) = \frac{1}{2\mathcal{N}} \sum_{i,j} \left(\langle \hat{\sigma}_i^y \hat{\sigma}_j^y \rangle - \langle \hat{\sigma}_i^z \hat{\sigma}_j^z \rangle \right), \tag{3}$$

$$\mathcal{B}(t) = \frac{1}{2\mathcal{N}} \sum_{i,j} \left(\langle \hat{\sigma}_i^y \hat{\sigma}_j^z \rangle + \langle \hat{\sigma}_i^z \hat{\sigma}_j^y \rangle \right). \tag{4}$$

In general $\xi(t)$ will decrease with time (starting at $\xi(0) = 1$), obtaining a minimum value before eventually growing large. We characterize the time $\tilde{t}$ at which ξ is minimized, and the squeezing $\tilde{\xi} = \xi(\tilde{t})$ achieved at that time, by how they scale with the system size $\mathcal{N}$ in the $\mathcal{N} \to \infty$ limit [40],

$$\tilde{\xi}^2 \sim \mathcal{N}^\nu \quad \text{and} \quad \tilde{t} \sim \mathcal{N}^\mu. \tag{5}$$

For infinite-range interactions ($\alpha = 0$), the behavior of the system simplifies because the dynamics is constrained to a small set of collective states, known as the Dicke manifold for its role in the theory of superradiance [4]. Equation (2) then takes a simple closed form, and analysis of its asymptotic behavior yields $\nu = \mu = -2/3$ [14]. Note that $\nu = 0$ and $\nu = -1$ correspond, respectively, to the standard quantum limit and the Heisenberg limit; the single-axis twisting model produces squeezed states that fall in between these two limits.

Scaling of squeezing for $\alpha \neq 0$.—For interactions that decay with distance the dynamics is no longer constrained to the Dicke manifold. However, correlation functions can still be computed using the exact solutions of Refs. [19, 41]. We find

$$\mathcal{A}(t) = \frac{1}{4\mathcal{N}} \sum_{i \neq j} \left(\prod_{k \neq i,j} \cos(\varphi_{ik} - \varphi_{jk}) - \prod_{k \neq i,j} \cos(\varphi_{ik} + \varphi_{jk}) \right),$$

$$\mathcal{B}(t) = \frac{1}{\mathcal{N}} \sum_{i \neq j} \sin(\varphi_{ij}) \prod_{k \neq i,j} \cos\varphi_{ik}, \tag{6}$$

where $\varphi_{ik} = 2Jt|\mathbf{r}_i - \mathbf{r}_k|^{-\alpha}$. To proceed we will expand these functions as series in time and analyze, at each order in time, the leading contributions in powers of the system size $\mathcal{N}$. To this end, it is helpful to give the terms in these expansions diagramatic representations. For example, Fig. 2 shows a schematic representation of the series expansion for $\mathcal{A}(t)$.

$$\mathcal{A}(t) = \frac{1}{2}\left(\bigcirc + \cdots \right)_{\text{odd}}$$
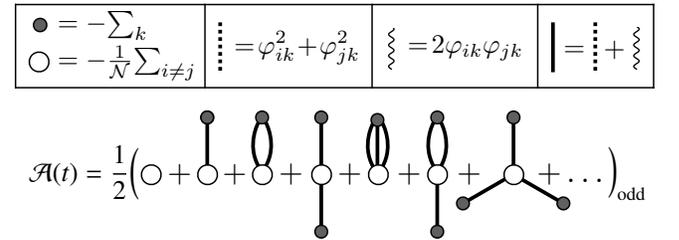
FIG. 2. Expansion of Eq. (6) [see also Eq. (S1)]. Each diagram has an associated combinatorial factor, which is obtained by multiplying the entire diagram by $1/m!$ for each group of m identical vertices, and then multiplying each black vertex by $1/(2n)!$, where n is the number of legs connected to that vertex. The subscript "odd" implies to keep only diagrams with an odd number of wavy legs.

These diagrams facilitate a sorting of contributions to $\mathcal{A}(t)$ and $\mathcal{B}(t)$ by powers of $\mathcal{N}$. For example, in the case $\alpha = 0$ it is straightforward to see that tree diagrams—those for which each black vertex is connected by a single leg to the central white vertex—contribute dominantly (in powers of $\mathcal{N}$) at any order in time.

To obtain a valid expansion for the squeezing parameter $\xi(t)$ near $t = \tilde{t}$, we assume that $\mathcal{A}(\tilde{t}) \gg \mathcal{B}(\tilde{t})$, which can be verified

at the end of the calculation. This assumption enables us to rewrite $\xi(t)$ as the following convergent series

$$\xi(t)^2 = 1 - \frac{1}{2}\frac{\mathcal{B}(t)^2}{\mathcal{A}(t)} + \frac{1}{8}\frac{\mathcal{B}(t)^4}{\mathcal{A}(t)^3} - \dots \quad (7)$$

Inspection of Eq. (7) reveals that $\xi(t) \ll 1$ requires $\delta(t) \equiv \mathcal{A}(t) - \frac{1}{2}\mathcal{B}(t)^2 > 0$ must satisfy $\delta(t) \ll \mathcal{A}(t), \frac{1}{2}\mathcal{B}(t)^2$. Plugging $\mathcal{A}(t) = \frac{1}{2}\mathcal{B}(t)^2 + \delta(t)$ into Eq. (7) and working to lowest nontrivial order in both $\delta(t)$ and $\mathcal{B}(t)^{-1}$, we then obtain

$$\xi(t)^2 \approx \frac{1 + 2\delta(t)}{\mathcal{B}(t)^2}. \quad (8)$$

Our goal is now to understand the asymptotic behavior of Eq. (8) for large $\mathcal{N}$ by expanding the correlation functions involved diagramatically.

For $0 < \alpha < D/2$, each black vertex attached by a single leg contributes at order $t^2 \mathcal{N}^{1-2\alpha/D}$ to a given diagram, as can be verified by converting the sums involved into integrals [42]. Summations over black vertices with more than one leg either converge or diverge less strongly with $\mathcal{N}$, and therefore $\mathcal{A}(t)$ is dominated by tree diagrams. Factoring out all contributions from wavy legs, we obtain

$$\mathcal{A}(t) \approx \frac{1}{2}\Big(\overset{\vdots}{\circ} + \overset{}{\circ} + \dots\Big) \times \Big(\circ + \overset{\vdots}{\circ} + \overset{\vdots}{\circ} + \overset{\vdots}{\circ} + \dots\Big).$$

Here, "$\approx$" means "equal to leading order in $\mathcal{N}$ at all orders in t", and a product of two diagrams is carried out by joining their white vertices. The two series in parentheses above can be resummed to give

$$\mathcal{A}(t) \approx \frac{1}{2\mathcal{N}}\sum_{i \neq j}\sinh(\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j)) \times \exp\Big(-\sum_k(\varphi_{ik}^2 + \varphi_{jk}^2)\Big), \quad (9)$$

where $\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j) \equiv \sum_k \varphi_{ik}\varphi_{jk}$. In the supplemental material [42], we show that for a homogeneous system the terms with one wavy leg, which can be obtained from Eq. (9) by expanding the sinh to first order, coincide with the expansion of $\frac{1}{2}\mathcal{B}(t)^2$ to leading order in $\mathcal{N}$ at all orders in time (this also holds true for $D/2 < \alpha < D$, even though non-tree diagrams must be included in this case). Thus we have

$$\frac{1}{2}\mathcal{B}(t)^2 \approx \frac{1}{2}\Big(\overset{\vdots}{\circ}\Big) \times \Big(\circ + \overset{\vdots}{\circ} + \overset{\vdots}{\circ} + \overset{\vdots}{\circ} + \dots\Big),$$

$$\delta(t) \approx \frac{1}{2}\Big(\overset{}{\circ} + \dots\Big) \times \Big(\circ + \overset{\vdots}{\circ} + \overset{\vdots}{\circ} + \overset{\vdots}{\circ} + \dots\Big).$$

The equivalence between the set of diagrams with a single wavy leg and $\frac{1}{2}\mathcal{B}(t)^2$ has a clear physical origin: In the expansion of $\mathcal{A}(t)$, these diagrams generate correlations between the
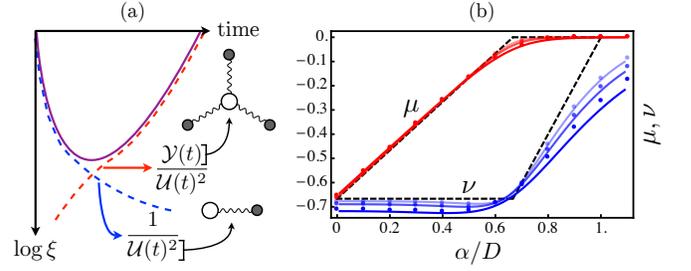

FIG. 3. (Color online) (a) Schematic plot of Eq. (11), demonstrating how the optimal squeezing results from a competition of two different diagrams. (b) Scaling of spin-squeezing with the interaction exponent α. The blue (red) points show the scaling ν (μ) for three different values of $\mathcal{N} = 10^3, 10^4, 10^5$ (from darker to lighter shades), calculated using Eq. (6). The solid blue (red) lines are the corresponding results obtained from Eq. (11). The black-dashed lines are analytic results, valid for $\mathcal{N} \to \infty$, and reported in the table on the fourth page of the manuscript.

y components of two spins at sites i and j through their mutual precession around the z-projection of a single intermediate spin k. On the other hand, that same diagram can be viewed as the product of two independent processes that generate yz correlations, either between the spins i and k or spins j and k, and therefore also provides the leading contribution in the expansion of $\mathcal{B}(t)^2$.

In the exponent of Eq. (9), we encounter sums that scale (at the time of optimal squeezing) as $\tilde{t}^2 \times \mathcal{N}^{1-2\alpha/D} \sim \mathcal{N}^{2\mu+1-2\alpha/D}$. These sums will be small in the large $\mathcal{N}$ limit as long as $\mu < \alpha/D - 1/2$, in which case we set the exponential to unity; this condition will be assumed, and can be shown to be self-consistent at the end of the calculation. Setting this exponential to unity amounts to dropping all diagrams containing dashed legs in $\mathcal{A}(t)$, $\mathcal{B}(t)^2$, and $\delta(t)$, and so at this level of approximation we have

$$\mathcal{B}(t)^2 \approx \frac{1}{\mathcal{N}}\sum_{i \neq j}\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j) = \sum_{i \neq j}\varphi_{ik}\varphi_{jk} \approx \mathcal{U}(t)^2, \quad (10)$$

where $\mathcal{U}(t) \equiv \sum_i \varphi_{ik}$. Note that the equality results from homogeneity, and the second $\approx$ implies equality to leading order in $\mathcal{N}$. Still working under the assumption $\mu < \alpha/D - 1/2$, one can check that at the time of optimal squeezing $\delta(t)$ is dominated by the lowest-order diagram (with three wavy legs), given by $\mathcal{Y}(t) = \frac{1}{\mathcal{N}}\frac{2^3}{3!}\sum_{i \neq j}\mathcal{V}(\boldsymbol{r}_i - \boldsymbol{r}_j)^3$, and therefore

$$\xi^2(t) \approx \frac{1 + \mathcal{Y}(t)}{\mathcal{U}(t)^2}. \quad (11)$$

The diagrams that represent $\mathcal{U}(t)$ and $\mathcal{Y}(t)$, and their contributions to the behavior of $\xi^2(t)$, are shown schematically in Fig. 3a. Physically, $\mathcal{Y}(t)$ corresponds to a process where the y components of two spins become correlated via their mutual interaction with three other spins (as in Fig. 1b). By converting sums into integrals and doing dimensional analysis, these diagrams can readily be seen to scale as $\mathcal{U}(t) \sim \mathcal{N}(t\mathcal{N}^{-\alpha/D})$

and $\mathcal{Y}(t) \sim \mathcal{N}^4(t\mathcal{N}^{-\alpha/D})^6$. Therefore, in terms of a rescaled time $\tau = t\mathcal{N}^{-\alpha/D}$, Eq. (11) obtains the same form that it would in terms of the bare time t at $\alpha = 0$. As an immediate consequence, the optimal squeezing exponent ν cannot depend on α and must take on its $\alpha = 0$ value of $\nu = -2/3$. Similarly, we conclude that $\tau \sim \mathcal{N}^{-2/3}$ (because this is how t scales at $\alpha = 0$), giving $\mu = -2/3 + \alpha/D$.

Equation (11) actually remains valid as α exceed $D/2$, but the calculation leading to it becomes more difficult, in part because the expansion of $\mathcal{A}(t)$ is no longer dominated by tree diagrams. A detailed analysis reveals that the leading order diagrams for $\mathcal{A}(t)$, which have one wavy leg, continue to agree with $\frac{1}{2}\mathcal{B}(t)^2$ to leading order in $\mathcal{N}$ at all orders in time [42]. In addition, $\mathcal{Y}(t)$ and $\mathcal{U}(t)^2$ continue to give the dominant contributions to $\delta(t)$ and $\mathcal{B}(t)^2$, respectively, as long as the diagram $\mathcal{Y}(t)$ diverges with $\mathcal{N}$, which persists until $\alpha = 2D/3$. Thus the result $(\nu, \mu) = (-2/3, \alpha/D - 2/3)$ holds for $0 \leq \alpha < 2D/3$. For $2D/3 < \alpha < D$, the numerator in Eq. (11) no longer diverges with the system size, while the denominator scales as $t^2\mathcal{N}^{2-2\alpha/D}$. The system-size dependence therefore factors out, giving $\xi^2 \sim \mathcal{N}^{2(\alpha/D-1)}$, causing the optimal squeezing to occur at a time that is independent of $\mathcal{N}$. For $\alpha > D$ all diagrams are convergent in the large $\mathcal{N}$ limit, and neither $\tilde{t}$ nor $\tilde{\xi}$ scale with the system size, giving $\mu = \nu = 0$. These results can be summarized as follows,

ν	$-2/3$	$2(\alpha/D-1)$	0
μ	$-2/3 + \alpha/D$	0	0

Full squeezing (obtained at later times than for $\alpha = 0$)	Partial squeezing ($\tilde{\xi}$ scales less favorably than for $\alpha = 0$)	Squeezing is independent of system size

α axis: 0 — $2D/3$ — D

To verify Eq. (11) and the associated scaling, we have computed μ and ν by direct numerical evaluation of Eq. (6) for $\mathcal{N} = 10^3, 10^4, 10^5$, and plotted them as points in Fig. 3(b). They can be seen to converge toward the asymptotic scaling results (black dashed lines) in the large system limit, and are increasingly well approximated by Eq. (11) for larger $\mathcal{N}$.

Effects of inhomogeneity.—In the presence of spatially decaying interactions, most experimental systems also lack spatial homogeneity due to boundary effects (though exceptions certainly can exist, e.g. ring traps for ions [43]). For sufficiently fast-decaying interactions, this lack of homogeneity is only expected to be relevant near the edge of the system, and thus should not be important in the large $\mathcal{N}$ limit. However, for interactions that are sufficiently long-ranged to generate significant spin-squeezing ($\alpha < D$), edge effects cannot in general be ignored (see for example Refs. [44, 45], in which boundary conditions in ion chains are shown to play an important role in local quench dynamics). In the absence of homogeneity, the cancellation of leading order diagrams between $\mathcal{A}(t)$ and $\frac{1}{2}\mathcal{B}(t)^2$ is imperfect, which results in a leading contribution to $\delta(t)$ determined by the variance of $\mathcal{U}_k(t) \equiv \sum_i \varphi_{ik}$
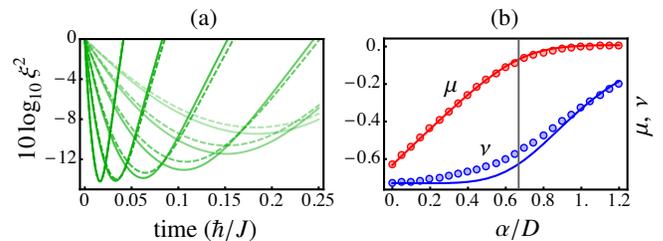

FIG. 4. (Color online) (a) Squeezing parameter as a function of time for a 1D chain with 200 spins, for a variety of exponents α ranging from 0 to $D = 1$ (darker to lighter). Solid curves are for periodic boundary conditions, while dashed curves are for open boundary conditions. (b) Scaling of squeezing with system size, again evaluated for a 200-spin 1D chain. The blue (red) line shows ν (μ) for periodic boundary conditions, while the blue (red) disks show ν (μ) for open boundary conditions.

(which is now position dependent)

$$X(t) = \frac{1}{\mathcal{N}} \sum_k \mathcal{U}_k(t)^2 - \left(\frac{1}{\mathcal{N}} \sum_k \mathcal{U}_k(t)\right)^2. \quad (12)$$

For generic open boundary conditions and $\alpha < D$, this variance will scale as $\mathcal{N}^{2-2\alpha/D}$ (the same as the lowest order contribution to $\mathcal{B}(t)^2$), and vanishes for any finite system as $\alpha \to 0$. In the large system limit for *any* $\alpha > 0$, we can thus approximate

$$\xi(t)^2 \approx \frac{1 + \mathcal{Y}(t)}{\mathcal{B}(t)^2} + \frac{X(t)}{\mathcal{B}(t)^2}. \quad (13)$$

Since the second term stays finite as $\mathcal{N} \to \infty$, this expression leads to a minimum squeezing parameter that is independent of system size.

Implications for experimental systems.—These results shed light on the importance of both translational invariance and collective dynamics on spin-squeezing, and have implications for a number of experimental systems where interactions are only approximately independent of distance. For example, in trapped ion experiments, finite detuning from the center of mass mode results in interactions that decay as a weak power law in space [36, 46]. For systems with open boundary conditions, our results demonstrate that even arbitrarily slow-decaying interactions cause the maximum achievable squeezing to be independent of system size in the large $\mathcal{N}$ limit. However, because the coefficient preceding $\mathcal{N}^{2-2\alpha/D}$ in $X(t)$ can be very small ($\mathcal{U}_k(t)$ does not fluctuate very much in space for very long-range interactions), Eq. (11) and the pursuant scaling arguments are still relevant on a qualitative level for characterizing spin squeezing with open boundary conditions, as long as the system size is not too large. In Fig. 4(a) we show that the squeezing for open and closed boundary conditions is not actually that different for a linear chain with only hundreds of spins. Moreover, in Fig. 4(b) we see that the scaling behavior predicted for periodic boundary conditions—e.g. squeezing that is weakly dependent on α for $\alpha < 2D/3$ and then falls off continuously, eventually vanishing for $\alpha > D$—remains qualitatively correct for systems with open boundary conditions.

Outlook.—Entanglement generation via spin squeezing has been shown to be robust against excursions outside of the Dicke (collective-spin) manifold. For large systems the spin squeezing parameter is, however, sensitive to the existence (or lack thereof) of boundary-induced inhomogeneities. It seems very likely that this sensitivity indicates a failure of the squeezing parameter—which is explicitly designed to capture the entanglement of permutationally (and therefore translationally) invariant spin systems—to properly capture the entanglement generated in inhomogeneous systems. From a more fundamental standpoint, it would therefore be worthwhile to explore whether the scaling with system size of quantum Fisher information [47, 48], entanglement depth [23], or any other metrics of entanglement, persists for spatially decaying interactions, even in geometries where the scaling of the spin-squeezing parameter does not.

---

[1] E. H. Lieb and D. W. Robinson, Comm. Math. Phys. **28**, 251 (1972).

[2] P. Calabrese and J. Cardy, J. Stat. Mech. **2005**, P04010 (2005).

[3] J. Schachenmayer, B. P. Lanyon, C. F. Roos, and A. J. Daley, Phys. Rev. X **3**, 031015 (2013).

[4] R. H. Dicke, Phys. Rev. **93**, 99 (1954).

[5] M. B. Hastings and T. Koma, Comm. Math. Phys. **265**, 781 (2006).

[6] P. Hauke and L. Tagliacozzo, Phys. Rev. Lett. **111**, 207202 (2013).

[7] Z.-X. Gong, M. Foss-Feig, S. Michalakis, and A. V. Gorshkov, Phys. Rev. Lett. **113**, 030602 (2014).

[8] M. Foss-Feig, Z.-X. Gong, C. W. Clark, and A. V. Gorshkov, Phys. Rev. Lett. **114**, 157201 (2015).

[9] J. Eisert, M. van den Worm, S. R. Manmana, and M. Kastner, Phys. Rev. Lett. **111**, 260401 (2013).

[10] Z. Eldredge, Z.-X. Gong, A. H. Moosavian, M. Foss-Feig, and A. V. Gorshkov, arXiv:1612.02442 (2016).

[11] B. Yan, S. A. Moses, B. Gadway, J. P. Covey, K. R. A. Hazzard, A. M. Rey, D. S. Jin, and J. Ye, Nature **501**, 521 (2013).

[12] P. Richerme, Z.-X. Gong, A. Lee, C. Senko, J. Smith, M. Foss-Feig, S. Michalakis, A. V. Gorshkov, and C. Monroe, Nature **511**, 198 (2014).

[13] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos, Nature **511**, 202 (2014).

[14] M. Kitagawa and M. Ueda, Phys. Rev. A **47**, 5138 (1993).

[15] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen, Phys. Rev. A **50**, 67 (1994).

[16] J. Esteve, C. Gross, A. Weller, S. Giovanazzi, and M. K. Oberthaler, Nature **455**, 1216 (2008).

[17] B. Lücke, M. Scherer, J. Kruse, L. Pezzé, F. Deuretzbacher, P. Hyllus, O. Topic, J. Peise, W. Ertmer, J. Arlt, L. Santos, A. Smerzi, and C. Klempt, Science **334**, 773 (2011).

[18] Z. Chen, J. G. Bohnet, S. R. Sankar, J. Dai, and J. K. Thompson, Phys. Rev. Lett. **106**, 133601 (2011).

[19] M. Foss-Feig, K. R. A. Hazzard, J. J. Bollinger, and A. M. Rey, Phys. Rev. A **87**, 042101 (2013).

[20] T. E. Lee, S. Gopalakrishnan, and M. D. Lukin, Phys. Rev. Lett. **110**, 257204 (2013).

[21] L. I. R. Gil, R. Mukherjee, E. M. Bridge, M. P. A. Jones, and T. Pohl, Phys. Rev. Lett. **112**, 103601 (2014).

[22] A. Sorensen, L. M. Duan, J. I. Cirac, and P. Zoller, Nature **409**, 63 (2001).

[23] A. S. Sørensen and K. Mølmer, Phys. Rev. Lett. **86**, 4431 (2001).

[24] Z. Zhang and L.-M. Duan, Phys. Rev. Lett. **111**, 180401 (2013).

[25] A. Kuzmich, L. Mandel, and N. P. Bigelow, Phys. Rev. Lett. **85**, 1594 (2000).

[26] M. Foss-Feig, A. J. Daley, J. K. Thompson, and A. M. Rey, Phys. Rev. Lett. **109**, 230501 (2012).

[27] E. G. Dalla Torre, J. Otterbach, E. Demler, V. Vuletic, and M. D. Lukin, Phys. Rev. Lett. **110**, 120402 (2013).

[28] E. Davis, G. Bentsen, and M. Schleier-Smith, Phys. Rev. Lett. **116**, 053601 (2016).

[29] J. Esteve, C. Gross, A. Weller, S. Giovanazzi, and M. K. Oberthaler, Nature **455**, 1216 (2008).

[30] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, Phys. Rev. Lett. **106**, 130506 (2011).

[31] J. G. Bohnet, K. C. Cox, M. A. Norcia, J. M. Weiner, Z. Chen, and J. K. Thompson, Nat. Photonics **8**, 731 (2014).

[32] K. R. A. Hazzard, B. Gadway, M. Foss-Feig, B. Yan, S. A. Moses, J. P. Covey, N. Y. Yao, M. D. Lukin, J. Ye, D. S. Jin, and A. M. Rey, Phys. Rev. Lett. **113**, 195302 (2014).

[33] A. de Paz, A. Sharma, A. Chotia, E. Maréchal, J. H. Huckans, P. Pedri, L. Santos, O. Gorceix, L. Vernac, and B. Laburthe-Tolra, Phys. Rev. Lett. **111**, 185305 (2013).

[34] M. Saffman, T. G. Walker, and K. Mølmer, Rev. Mod. Phys. **82**, 2313 (2010).

[35] P. Schauß, M. Cheneau, M. Endres, T. Fukuhara, S. Hild, A. Omran, T. Pohl, C. Gross, S. Kuhr, and I. Bloch, Nature **491**, 87 (2012).

[36] J. W. Britton, B. C. Sawyer, A. C. Keith, C. C. J. Wang, J. K. Freericks, H. Uys, M. J. Biercuk, and J. J. Bollinger, Nature **484**, 489 (2012).

[37] J. G. Bohnet, B. C. Sawyer, J. W. Britton, M. L. Wall, A. M. Rey, M. Foss-Feig, and J. J. Bollinger, Science **352**, 1297 (2016).

[38] M. J. Martin, M. Bishof, M. D. Swallows, X. Zhang, C. Benko, J. von Stecher, A. V. Gorshkov, A. M. Rey, and J. Ye, Science **341**, 632 (2013).

[39] As an entanglement witness or a measure of metrological enhancement, ξ should be normalized by twice the Bloch vector length rather than the number of particles. However, the Bloch vector length at the optimal squeezing time approaches $\mathcal{N}/2$ in the large $\mathcal{N}$ limit, and thus its omission does not affect the derived scaling exponents.

[40] These scaling exponents are defined formally by $\nu = \lim_{\mathcal{N}\to\infty} d\log\tilde{\xi}^2/d\log\mathcal{N}$, $\mu = \lim_{\mathcal{N}\to\infty} d\log\tilde{t}/d\log\mathcal{N}$. These derivatives, evaluated at finite $\mathcal{N}$, determine the curves in Fig. 3. The arbitrary choice of ξ^2 instead of ξ in the definition of ν is made so that μ and ν are equal at $\alpha = 0$.

[41] M. van den Worm, B. C. Sawyer, J. J. Bollinger, and M. Kastner, New J. Phys. **15**, 083007 (2013).

[42] See the supplemental material for details.

[43] T. Li, Z.-X. Gong, Z.-Q. Yin, H. T. Quan, X. Yin, P. Zhang, L.-M. Duan, and X. Zhang, Phys. Rev. Lett. **109**, 163001 (2012).

[44] Z.-X. Gong and L.-M. Duan, New J. Phys. **15**, 113051 (2013).

[45] B. Neyenhuis, J. Smith, A. C. Lee, J. Zhang, P. Richerme, P. W. Hess, Z.-X. Gong, A. V. Gorshkov, and C. Monroe, arXiv:1608.00681 (2016).

[46] 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, Science **340**, 583 (2013).

[47] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[48] G. Tóth and D. Petz, Phys. Rev. A **87**, 032324 (2013).

S1. SUPPLEMENTAL MATERIAL

In this supplemental material, we describe in more detail how correlation functions can be expanded in terms of the diagrams described in the paper. We also derive the scaling properties of different types of diagrams for various ranges of α, and provide some additional details on the derivation leading from Eq. (8) to Eq. (11) of the main text.

As a preliminary remark, we emphasize that the scaling of $\tilde{\xi}$ and $\tilde{t}$ in the large $\mathcal{N}$ limit cannot be understood by simply expanding $\xi(t)$ as a power series in time. The reason is that, though squeezing does occur at short times (and thus $\tilde{t}$ is indeed a small parameter), the existence of the large parameter $\mathcal{N}$ makes it impossible to assign, a priori, preference to low-order terms in this expansion. This is the fundamental reason why a careful expansion to high orders in time, which we carry out diagrammatically, is necessary. As a simple (though still subtle) example, consider the $\alpha = 0$ case, which was first studied in Ref. [14]. There, one finds that while $\mathcal{N}t^2$ is a suitable small parameter, $\mathcal{N}t$ actually becomes large (in the large $\mathcal{N}$ limit) at the time of optimal squeezing, as can be immediately verified by plugging in the scaling $\tilde{t} \sim \mathcal{N}^{-2/3}$. Therefore, a simultaneous expansion in both $\varepsilon = \mathcal{N}t^2$ *and* $\delta = 1/(\mathcal{N}t)$ is necessary to capture squeezing dynamics at low order. That such an expansion is not necessarily (and indeed turns out not to be) analytic in time confirms our initial statement that low orders in short-time perturbation theory will fail to describe spin squeezing. It is nevertheless useful to expand all two-spin correlation functions as series in time, and study the leading order contributions, in powers of the system size $\mathcal{N}$, at each order. Understanding this scaling to all orders in time will allow us to identify suitable small parameters in which to expand ξ, and to thereby identify its scaling with $\mathcal{N}$ for non all-to-all interactions.

A. Series expansion of correlation functions

To begin, we can expand the correlation functions in Eq. (6) of the manuscript. Of particular importance is the expansion for $\mathcal{A}(t)$; defining $\Upsilon_k \equiv (\varphi_{ik} + \varphi_{jk})^2$, we find

$$\mathcal{A}(t) = \frac{1}{4\mathcal{N}} \sum_{i \neq j} \left(\prod_{k \neq i,j} \cos\left(\varphi_{ik} - \varphi_{jk}\right) - \prod_{k \neq i,j} \cos\left(\varphi_{ik} + \varphi_{jk}\right) \right) \tag{S1}$$

$$= -\frac{1}{2\mathcal{N}} \sum_{i \neq j} \left(1 - \sum_k \frac{\Upsilon_k}{2!} + \sum_k \frac{\Upsilon_k^2}{4!} + \frac{1}{2!} \sum_{k \neq l} \frac{\Upsilon_k}{2!} \frac{\Upsilon_l}{2!} - \sum_k \frac{\Upsilon_k^3}{6!} - \sum_{k \neq l} \frac{\Upsilon_k}{2!} \frac{\Upsilon_l^2}{4!} - \frac{1}{3!} \sum_{k \neq l \neq m} \frac{\Upsilon_k}{2!} \frac{\Upsilon_l}{2!} \frac{\Upsilon_m}{2!} + \cdots \right)_{\text{odd}}$$

In the second line, the subscript "odd" implies that, when expanding out a product of multiple Υ's, we should only keep terms that have an odd number of cross terms of the form $\varphi_{ik}\varphi_{jk}$. In the third line, in light of the rules in Fig. 2 of the manuscript, the subscript implies that we should only keep diagrams with an odd number of wavy legs.

B. Analysis of the leading order contributions

In the case of $\alpha = 0$, each black vertex contributes a power of $\mathcal{N}$ to a diagram. Given that each leg contributes a factor of t^2, the leading diagrams at each order in time are all those with the same number of legs and black vertices. In the case of m legs, such diagrams are proportional to $t^{2m}\mathcal{N}^{1+m}$. For $0 < \alpha < D/2$, each vertex with a single leg contributes a factor of $t^2\mathcal{N}^{1-2\alpha/D}$, while any vertex with more than one leg either converges or diverges less strongly with $\mathcal{N}$. This result is easy to check for a dashed-line vertex by analyzing the sums

$$\sum_{k \neq i} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_k|^{2m\alpha}} \sim \begin{cases} \mathcal{N}^{1-2m\alpha/D} & \text{if } 2m\alpha < 1; \\ 1 & \text{if } 2m\alpha > 1. \end{cases} \tag{S2}$$

Thus for $\alpha < D/2$ a black vertex with a single dashed leg always diverges, and grows faster with system size than a black vertex with more than one leg. The contribution of a vertex with a single wavy leg depends on the positions of sites i and j, which are eventually summed over. To analyze the contribution from such vertices we must therefore understand the behavior of $\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j) = \sum_k |\boldsymbol{r}_i - \boldsymbol{r}_k|^{-\alpha}|\boldsymbol{r}_k - \boldsymbol{r}_j|^{-\alpha}$. This sum can be estimated by converting it to an integral; for example in 1D, using a continuous coordinate rescaled by the system size $z = j/\mathcal{N}$, we would have

$$\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j) \approx \mathcal{N}^{1-2\alpha} \int_0^1 \frac{dz}{|x - z|^\alpha |z - y|^\alpha}. \tag{S3}$$

This integral converges for all $x \neq y$ whenever $\alpha < D$ (the singularities at $z = x, y$ are integrable), and thus the overall scaling of $\mathcal{V}$ is determined entirely by the prefactor $\mathcal{N}^{1-2\alpha}$. For $\alpha < D/2$ the integral remains finite even as $x \to y$, which implies that $\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j)$ remains proportional to $\mathcal{N}^{1-2\alpha}$ in the large $\mathcal{N}$ limit even if the separations between sites i and j grows with the system size. Thus we find that $\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j) \sim \mathcal{N}^{1-2\alpha} f(i, j)$, where $f(i, j)$ is a function that stays finite in the limit of large (i.e. on the order of the system size) values of $|\boldsymbol{r}_i - \boldsymbol{r}_j|$. The validity of the above derivation does not rely (except in detail) on the lattice structure or dimension, and hence we find more generally that $\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j) \sim \mathcal{N}^{1-2\alpha/D} f(i, j)$. Because the function $f(i, j)$ tends to a finite value for extensively large separations, the summation over i and j yields two powers of the system size $\mathcal{N}$, and therefore the contribution of a diagram with m legs and vertices scales as $\mathcal{N}^{1+m(1-2\alpha/D)}$ regardless of the types of legs.

In light of the above discussion, and still restricting our attention to the case $\alpha < D/2$, we can approximate the correlation function $\mathcal{A}(t)$ by keeping just the leading order contributions in $\mathcal{N}$ at each order in time:

The set of diagrams with a single wavy leg can be simplified by factoring out the vertex common to all diagrams, and resumming the remaining diagrams to give

$$\frac{1}{2}\frac{1}{\mathcal{N}}\sum_{i \neq j} \mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j) \exp[-\sum_{k \neq i,j}(\varphi_{ik}^2 + \varphi_{jk}^2)] = \frac{1}{2}\sum_{i \neq j}\left(\varphi_{il}\exp(-\sum_{k \neq i,j}\varphi_{ik}^2)\right)\left(\varphi_{jl}\exp(-\sum_{k \neq i,j}\varphi_{jk}^2)\right) \tag{S4}$$

$$\approx \frac{1}{2}\mathcal{B}(t)^2, \tag{S5}$$

where the $\approx$ means "correct to leading order in $\mathcal{N}$" at each order in time. The equality in Eq. (S4) can be verified by recalling the definition of $\mathcal{V}$ and exchanging summations:

$$\sum_{i \neq j}\mathcal{V}(\boldsymbol{r}_i, \boldsymbol{r}_j) = \sum_l \sum_{i \neq j \neq l}\frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_l|^\alpha |\boldsymbol{r}_l - \boldsymbol{r}_j|^\alpha} = \mathcal{N}\sum_{i \neq j}\varphi_{il}\varphi_{jl}. \tag{S6}$$

Eq. (S5) then follows immediately by expanding the sin functions in the definition of $\mathcal{B}(t)$ to lowest order (which is correct to leading order in $\mathcal{N}$). This approximate equality is the origin of the similarity between $\mathcal{A}(t)$ and $\frac{1}{2}\mathcal{B}(t)^2$. We can also resum the entire series (including any number of wavy legs) to give

$$\mathcal{A}(t) \approx \frac{1}{2}\frac{1}{\mathcal{N}}\sum_{i\neq j}\sinh[\mathcal{V}(\boldsymbol{r}_i,\boldsymbol{r}_j)]\exp[-\sum_{k\neq i,j}(\varphi_{ik}^2 + \varphi_{jk}^2)]. \tag{S7}$$

Expanding the sinh function to first order gives the $\frac{1}{2}\mathcal{B}(t)^2$ contribution, while the 3$^{\text{rd}}$ and higher order terms give $\delta(t)$. Now that we have the desired series representation for $\delta(t)$, and can see that the leading terms at order t^{2m} in time scale as $\mathcal{N}^{1+m(1-2\alpha/D)}$. Truncating this series to lowest (6$^{\text{th}}$) order in time and plugging it into Eq. (8) of the manuscript leads immediately to Eq. (11).

The situation for $\alpha > D/2$ requires slightly more care. First of all, dashed legs ($\sim \sum_k |\boldsymbol{r}_i - \boldsymbol{r}_k|^{-2\alpha}$) now converge, and hence the diagrams with a single leg per vertex are not parametrically larger (in $\mathcal{N}$) than diagrams with the same number of legs but fewer vertices—hence we cannot select a simple set of dominant diagrams at a given order in time. Similarly, for fixed $|\boldsymbol{r}_i-\boldsymbol{r}_j|$, $\mathcal{V}(\boldsymbol{r}_i,\boldsymbol{r}_j)$ no longer diverges in the large $\mathcal{N}$ limit, as suggested by the prefactor in Eq. (S3). Indeed, for a separation $|\boldsymbol{r}_i - \boldsymbol{r}_j|$ that grows proportionally to the system size, $\mathcal{V}(\boldsymbol{r}_i,\boldsymbol{r}_j)$ actually tends to zero as a power law as $\mathcal{N} \to \infty$. An immediate consequence is that the leading order contributions to $\mathcal{A}(t)$ are diagrams with only a single wavy leg, which suggests (incorrectly) that the leading contributions to $\delta(t)$ might be of a different character than they were for $\alpha < D/2$. However, these leading order contributions are still precisely (i.e. to leading order in $\mathcal{N}$) canceled by the factor of $\frac{1}{2}\mathcal{B}(t)^2$ in the definition of $\delta(t)$. The cancelation of this more general class of diagrams, which includes those having one wavy leg and other dashed legs connected to the same vertex, is a consequence of the identity $\cos(a \pm b) = \cos(a)\cos(b) \mp \sin(a)\sin(b)$, employed inside the products of equation Eq. (6) *before* the expansion.

To determine the next-to-leading-order contributions, we note that for $\alpha > D/2$ the integral in Eq. (S3) has a power law divergence $\sim |x - y|^{1-2\alpha}$ for small $|x - y|$. In a diagram with m wavy legs, the summation over i and j can be converted into an integral

$$\sum_{i\neq j}\mathcal{V}(\boldsymbol{r}_i,\boldsymbol{r}_j)^m \approx \mathcal{N}^{1+m(1-2\alpha)}\int dr\left(\int \frac{dz}{|z|^\alpha|z-r|^\alpha}\right)^m. \tag{S8}$$

Because the inner integral has a singularity $\sim r^{1-2\alpha}$, the outer integral converges whenever $m(2\alpha - 1) < 1$, and so the overall scaling of the digram is determined by the prefactor $\mathcal{N}^{1+m(1-2\alpha)}$. For $m(2\alpha - 1) > 1$ the outer integral gains a system-size dependence (due to the non-integrable singularity at $r = 0$) $\sim \mathcal{N}^{m(2\alpha-1)-1}$, canceling the prefactor and causing the diagram to converge in the large $\mathcal{N}$ limit. From this discussion, it follows that progressively higher order diagrams converge first as α exceeds $D/2$, and the last remaining divergent diagram contributing to $\delta(t)$, namely $\mathcal{Y}(t) \sim \mathcal{N}^{4-6\alpha/D}$, converges for $\alpha > 2D/3$.