Random Regular Graph States Are Complex at Almost Any Depth

Soumik Ghosh,[1] Dominik Hangleiter,[2,*] and Jonas Helsen[3]

[1] *Department of Computer Science, University of Chicago, Chicago, Illinois 60637, USA*
[2] *Simons Institute for the Theory of Computing, University of California at Berkeley, Berkeley, California 94720, USA*
[3] *QuSoft and CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands*

(Received 14 January 2025; revised 13 June 2025; accepted 22 October 2025; published 24 November 2025)

Graph states are fundamental objects in the theory of quantum information due to their simple classical description and rich entanglement structure. They are also intimately related to instantaneous quantum polynomial-time (IQP) circuits, which have applications in quantum pseudorandomness and quantum advantage. For us, they are a toy model to understand the relation between circuit connectivity, entanglement structure, and computational complexity. In the worst case, a strict dichotomy in the computational universality of such graph states appears as a function of the degree d of a regular graph state [Ghosh *et al.*, Phys. Rev. Lett. 131, 030601 (2023)]. In this paper, we study the average-case complexity of simulating random graph states of varying degree when measured in random product bases and give distinct evidence that a similar complexity-theoretic dichotomy exists in the average case. Specifically, we consider random d-regular graph states and prove three distinct results: First, we show two families of IQP circuits of depth d and show that they anticoncentrate for any $2 < d = o(n^{1/2})$ when measured in a random X-Y plane product basis. This implies anticoncentration for random constant-regular graph states. Second, in the regime $d = \Theta(n^c)$ with $c \in (0,1)$, we prove that random d-regular graph states contain polynomially large grid graphs as induced subgraphs with high probability. This implies that they are universal resource states for measurement-based computation. Third, in the regime of high degree ($d \sim n/2$), we show that random graph states are not sufficiently entangled to be trivially classically simulable, unlike Haar-random states. Proving the three results requires different techniques—the analysis of a classical statistical-mechanics model using Krawtchouck polynomials, graph-theoretic analysis using the switching method, and analysis of the ranks of submatrices of random adjacency matrices, respectively.

DOI: 10.1103/52xz-3hpc

I. INTRODUCTION

Graph states play a fundamental role in the theory of quantum computation and communication [1] as well as the study of the complexity of physical systems [2]. They are arguably the simplest quantum states, with a classical description in terms of simple graphs, exhibiting rich quantum phenomena. From the perspective of multipartite quantum communication, they are interesting because local operations can transform the global graph topology and thus allow flexible routing [3–5]. From the perspective of many-body physics, they are interesting since they relate to computationally distinct phases of matter [2].

In our work, we consider graph states from the perspective of understanding quantum properties that lead to computational speedups. Graph states are prime candidates to study a specific quantum phenomenon, namely, entanglement, in terms of how it relates to computational complexity. On the one hand, measuring graph states in adaptive single-qubit bases allows the execution of arbitrary quantum computations through measurement-based quantum computing [6,7]. These measurements can be restricted to the X-Y plane of the Bloch sphere [8]. On the other hand, graph states exhibit a rich multipartite entanglement structure [9], which is required for quantum speedups [10]. Understanding which properties of graph states make them generically hard to simulate classically can therefore yield insights into the mechanisms underlying quantum speedups.

An insightful model to study these properties is the family of *regular graph states*, i.e., graph states whose underlying graph on n vertices is d-regular for some $0 \leq d \leq n$. These states form a family of graph states with a well-controlled connectivity structure, which relates to

[*] Contact author: mail@dhangleiter.eu

their entanglement and classical simulability. The degree of the underlying graph also connects to the circuit depth required to prepare them, since they are prepared by applying controlled-phase gates to an initial $|+\rangle^{\otimes n}$ state—all d-regular graph states can be prepared in depth at most $d+1$. d-regular graph states are therefore amenable to preparation on near-term devices with long-range connectivity such as reconfigurable atom arrays [11] and trapped ions [12]. A recent experiment demonstrated quantum advantage based on random universal circuits on regular graphs [12]. In contrast to universal random circuits, graph states can be implemented via naturally fault-tolerant operations in certain stabilizer codes [13,14] and are therefore amenable to early fault-tolerant implementations [11] as well as tailored error mitigation [15]. Graph state preparations can also be efficiently verified with the use of high-quality single-qubit measurements [16]. Thus, random regular graph states enable compelling experiments demonstrating noise-robust quantum advantage [17], efficient verification and benchmarking of generic structured circuits in long-range-connected architectures [13], and the first experimental preparations of highly structured states with applications beyond quantum advantage [18].

In the *worst case* Ghosh *et al.* [19] gave a tight connection between simulability and entanglement: when measured in an arbitrary X-Y plane product basis, d-regular graph states are hard to simulate and highly entangled if and only if $2 < d < n-3$. But both for demonstrating quantum advantage and when one is aiming to understand the intrinsic relation between hardness of simulation, multipartite entanglement, and device connectivity, it is crucial to study *generic* states from the family, i.e., random d-regular graph states [20,21]. This ensures that we are not drawing conclusions from isolated points in the family and makes random regular graphs a Goldilocks model to study those fundamental and practical questions alike [22].

In this paper, we therefore consider the *average-case complexity* of simulating uniformly random d-regular graph states when measured in an arbitrary X-Y plane product basis. In doing so, we study structured randomness in order to understand the relation between entanglement and complexity. From a fundamental perspective, this provides an alternative to fully random quantum states. From an experimental perspective, it provides an opportunity for demonstrating the computational capabilities of quantum processors for more structured circuits, as well as the potential for fault-tolerant implementations. Importantly, the setting we consider covers both sampling problems and universality for measurement-based quantum computing. We give evidence for the average-case complexity of these two problems in three distinct regimes of the regularity parameter: (1) the regime of $2 < d = O(1)$, (2) the regime of $d \in \Theta(n^c)$ for any $1/2 < c < 1$, and (3) the regime of $d \sim n/2$. To the best of our knowledge, we thus give the first ensemble of circuits which are classically intractable for *any depth* (in particular low depth) above a constant threshold. The best lower bound we had for such a threshold was logarithmic depth [13,23–25].

We first give evidence for average-case complexity in the regime of $2 < d = O(1)$. To this end, we first show that the output distributions of instantaneous quantum polynomial-time (IQP) circuits on random highly connected graphs constructed from d random matchings measured in a random X-Y plane basis have the anticoncentration property [26,27] at any depth $2 < d = o(n^{1/2})$. Combined with the fact that for each value of d there are worst-case hard instances in this family, this provides evidence for hardness of simulation on the same level of rigor as is known for other discrete families of IQP circuits [20,23]. This is the first family of quantum circuits that we are aware of for which there is evidence of simulation hardness at any sublinear depth above a constant threshold, and thus is a result of independent interest. We use this result to characterize the complexity of random regular graph states, noting that a random circuit from this matching model (strictly speaking a slight variation called the “pairing model’) at a fixed depth d yields a uniformly random d-regular graph state with probability $O(2^{-d^2})$. This implies the anticoncentration property of constant-regularity graph states and thus gives evidence for their classical intractability, showing the first main result.

In our second result, we show that a random d-regular graph with $d = \Theta(n^c)$ for any $1/2 < c < 1$ has a grid graph of polynomial size as an induced subgraph (resulting from deletion of some of the vertices). Since the grid graph is a universal resource for measurement-based quantum computation, this shows that an algorithm for sampling from the output distribution of those graph states in *any* local basis would imply a collapse to the polynomial hierarchy up to the average-case #P-hardness of a certain approximation problem [as well as being, in principle, universal resources for measurement-based quantum computing (MBQC)]. To the best of our knowledge, this is the strongest known average-case hardness result for any ensemble of graph states and the corresponding ensemble of IQP circuits.

Finally, in our third result, we show that uniformly random graph states have geometric entanglement bounded significantly away from the maximum. This makes them nontrivial and in particular not amenable to the trivial simulation algorithm of Ref. [28], giving some evidence for their average-case complexity when measured in any local basis. We believe similar results will hold for random cn-regular graphs with $0 < c < 1/2$.

Altogether, our results for the average-case behavior turn out to be qualitatively similar to those for the worst-case behavior. However, they are rather more difficult to obtain, and substantial technical work using entirely different tools for each of the results is required, pointing to different properties of the three regimes: The first result requires the analysis of a statistical-mechanics mapping [13], which we show can be reduced to asymptotic properties of the Krawtchouk polynomials [29], which have previously appeared in a variety of contexts [30]. The second result requires graph-theoretic tools to show properties of subgraphs of random d-regular graphs. The final result requires the analysis of the geometric entanglement entropy, which can be reduced to the study of extremal probability problems related to the rank of submatrices of uniformly random adjacency matrices.

Our results demonstrate that computationally universal or complex states can arise naturally from constrained randomness, and that this constraint can give rise to more complexity than less structured or completely unstructured randomness. Indeed, compared with the paradigmatic setting of quantum circuits composed of parallel Haar-random two-qubit gates, two of our results are particularly striking. First, our result that random graph states which can be prepared in constant depth exhibit anticoncentration is provably not true for constant-depth Haar-random circuits [25,31,32]. These may still be average-case hard to simulate, but numerical evidence in low dimensions points against a low depth threshold for classical intractability [24]. Second, our result that uniformly random graph states are not too entangled to be useful for measurement-based computation is also provably not true for Haar-uniformly random states [28]. Thus, our results suggests that while quantum states generated by Haar-random circuits are complex only in a limited depth regime, graph states generated by random controlled-Z (CZ) circuits are complex at almost any depth.

A. Guide for readers

Although the paper is rather long, it is composed of three parts, each covering a different main result in a fairly self-contained manner (depending on which result they want to read about, readers may jump to the relevant section without needing to read the other parts of the paper):

(1) *Anticoncentration results for IQP circuits and constant-degree regular graph states*. A summary of the results and proof techniques is given in Secs. I B 1, I B 2, and I C 1. The results are discussed in detail and proven in Sec. III, with preliminaries in Secs. II A–II C.

(2) *Universality results for random regular graph states of intermediate degree*. A summary of the results and proof techniques is given in Secs. I B 3 and I C 2. The results are proven in Sec. IV.

(3) *Absence of a geometric entanglement barrier for random graph states of high degree*. The main results and proof techniques are summarized in Secs. I B 4 and I C 3. The results are proven in Sec. V, with preliminaries in Sec. II D.

Not to be missed in any case is the discussion in Sec. I E.

B. Results

We present results on uniformly random d-regular graph states covering three different ranges for the regularity parameter d. We deal, in order, with $2 < d = O(1)$, $d = \Theta(n^c)$ for $c \in (0, 1)$, and $d \sim n/2$. Along the way we will also prove several results for different random graph models, such as the pairing and matching models (described in Sec. II B) and the uniformly random graph model. We believe these results to be of independent interest.

1. Anticoncentration of a family of random pairing and matching IQP circuits

Our first set of results deals with the average-case hardness of two families of IQP circuits of depth d satisfying $d > 2$ and $d = o(n^{1/2})$. These IQP circuits prepare (not always regular) graph states of degree at most d. Average-case hardness of constant-regularity graph states will follow from those results. To give evidence for average-case hardness of these depth-d circuits we prove that their outcome distribution, when measured in a random local basis in the X-Y plane, *anticoncentrates*. Moreover, they contain worst-case hard instances.

We state our theorems in terms of the (normalized) second moment

$$m_2(G, \theta) = 2^n \sum_{x \in \{0,1\}^n} p_{G,\theta}(x)^2$$

of the outcome distribution $p_{G,\theta}$ of the graph state on G measured in X-Y-plane angles $\theta \in [0, 2\pi)^n$. We say that the output distribution anticoncentrates if $m_2 \in O(1)$, since in that case a constant fraction of the output probabilities must be on the order of $1/2^n$. Averaging over graphs G and measurement angles θ, we find the average second moment $\overline{m}_2 = \mathbb{E}_{G,\theta}[m_2(G, \theta)]$ gives evidence for the #P-hardness of approximating the outcome probabilities of measuring states corresponding to random graphs G and random angles θ up to constant relative error [20]. Intuitively, anticoncentration prohibits an efficient average-case simulator which uses trivial approximations by zero to most probabilities of most graph states in its simulation. In this sense, it gives evidence for the average-case complexity of random graph states. Assuming additional well-founded complexity-theoretic assumptions, in conjunction with the worst-case hardness results of Ref. [19], this gives evidence for the hardness of sampling from such graph states to the same level of confidence as we have for other discrete families of circuits. The technical argument is reviewed in Ref. [20] and, roughly speaking, goes as follows: The circuit family we consider contains instances that are provably hard to strongly simulate up to relative error. Anticoncentration shows that most outcome probabilities have a similar order of magnitude and hence there is no detectable structure in the output distribution that would make a trivial simulation algorithm work. Finally, the ensembles we consider do not appear to have any exploitable structure that would help an algorithm designer to simulate a random instance compared with the worst-case instance; see also Ref. [33].

Specifically, we consider the following two ensembles of depth-d IQP circuits, which generate graph states of degree d. We call the first ensemble the “random pairing model.” This model is obtained by our choosing a uniformly random matching on nd vertices and identifying d vertices with a qubit. A CZ gate is applied to every edge in the resulting multigraph. This model (which is well studied in graph theory [34]) is motivated by the fact that, conditioning on the graph being simple, it generates a uniformly random d-regular graph state. We associate (simple) graph states with the multigraphs chosen in these ways by deleting double edges and self-loops. In the second model, the *random matching model*, a graph state is generated by application of CZ gates on d independent, uniformly random matchings. In this model, for any constant d, a regular graph state is generated with constant probability. However, not all d-regular graph states admit a perfect matching [35] (and some admit many, making this distribution different from the uniformly random regular graphs, even when conditioning on simplicity). Thus, the regular graph states generated by perfect matchings are exactly those graph states with $dn/2$ edges which can be prepared in optimal depth d.

We show that the output distribution of both models in a random X-Y basis anticoncentrates for any d satisfying $d > 2$ and $d = o(n^{1/2})$.

Theorem 1 (Anticoncentration of random pairing and matching graph states). Consider random graph states $|G\rangle$ on n vertices drawn from (1) the random pairing model on nd vertices or (2) d uniformly random matchings on n vertices. Then for $2 < d = o(n^{1/2})$

$$\mathbb{E}_{G,\theta}[m_2(G,\theta)] = o(1) + \begin{cases} 2 & \text{if } d \text{ is odd,} \\ 3 & \text{if } d \text{ is even.} \end{cases} \tag{1}$$

We also show the exact value of $\mathbb{E}_{G,\theta}[m_2(G,\theta)]$ for different values of n, d and both models in Fig. 1. We observe not only that the d-dependence of our results is tight, but also that the convergence to the asymptotic value occurs rapidly.

The results of Ref. [19] imply that for each value of d the random pairing model contains worst-case hard graphs. The random matching model trivially contains such graphs since for every odd d it contains the hexagonal lattice, and for every even d it contains the square lattice, both of which are universal for measurement-based quantum computation [36]. Thus, anticoncentration gives evidence for the average-case hardness of approximating the outcome probabilities in both cases. To the best of our knowledge, Theorem 1 is thus the first to give rigorous evidence for the hardness of simulating constant-depth circuits with random connectivity. In particular, it is the first result for constant-depth circuits which goes beyond resource states for MBQC [37–40]. At constant degree, we do not believe that random regular graphs contain (as induced subgraphs) large two-dimensional (2D) graphs (with high probability), which are usually required in MBQC constructions [36].

2. Anticoncentration of constant-degree regular graph states

Finally we can use our results on the pairing model to address the average-case complexity of random regular graph states of constant degree d. Specifically, Theorem 1 implies that random regular graphs with *any constant degree* $d > 2$ anticoncentrate, giving evidence

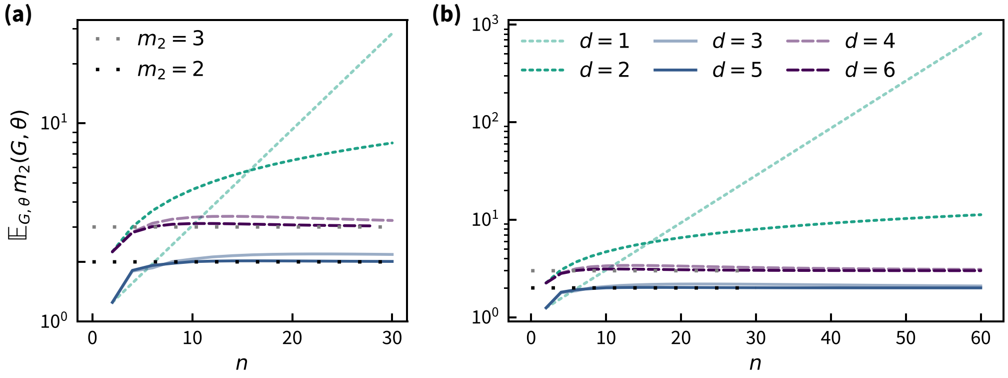

FIG. 1. Exact values of $\mathbb{E}_{G,\theta}[m_2(G,\theta)]$ [evaluated via Eq. (105)] for graph states drawn from (a) the random pairing ensemble $\mathcal{G}_p(d,n)$ and (b) the random matching ensemble $\mathcal{G}_m(d,n)$ for various values of d as a function of n. The opaque (transparent) loosely dotted lines denote the asymptotic values of $m_2 = 2$ ($m_2 = 3$) of the average second moment. Solid (dashed) lines represent odd (even) values of d. Dotted lines represent $d = 1, 2$ for which there is no anticoncentration.

for their average-case hardness. This statement follows immediately from the following more general corollary.

Corollary 1 (Anticoncentration of random d-regular graph states). Consider the uniform measure $\mathcal{G}_r(n)$ on d-regular graphs on n vertices, and uniformly random angles θ. Then, for $2 < d = o(n^{1/2})$,

$$\mathbb{E}_{G\sim\mathcal{G}_r(d,n),\theta}[m_2(G,\theta)] \leq (3+o(1))2^{d^2}. \tag{2}$$

This corollary is a straightforward consequence of the fact that conditioning on simplicity in the pairing model gives rise to uniformly random regular graphs (see Sec. II B). Corollary 1, together with the worst-case hardness result of Ref. [19], gives evidence that approximating the outcome probabilities of regular graph states of constant degree measured in random X-Y plane bases is #P-hard on average. We strongly believe that random regular graph states also anticoncentrate at super-constant depth, but leave showing that to future work.

We note that some existing results about IQP circuits can be rephrased in terms of random graph state models, yielding results similar to Corollary 1 for different models of random graphs. In particular, Ref. [13] shows that IQP circuits composed of g, CZ, and Z gates applied uniformly at random to n qubits anticoncentrate if $g = \Omega(n \log n)$ but fail to anticoncentrate for any $g = O(n)$. This directly translates to the Erdős-Rényi random graph model, in which an edge is contained in the graph with probability $p = g/\binom{n}{2}$, measured in the $\pm X$ basis. Reference [13] gives evidence that Erdős-Rényi random graphs are hard to simulate for $p \in \Omega(\log n/n)$ and suggests that they are not generically hard to simulate below that threshold. This is consistent with graph percolation, where famously Erdős and Rényi [41] showed that if $p < 1/2n$, a random graph's largest connected component has size $O(\log n)$, which implies that it is efficiently simulatable with the use of tensor-network techniques [10].

3. Universality of regular graphs of intermediate degree

Our second set of results is about the universality of random graph states, when the regularity of the graph scales quite strongly with n. In particular, when $d = \Theta(n^c)$ for $c \in (0,1)$ we argue that large grid graphs can be found as induced subgraphs in random regular graphs with high probability. This implies that the associated graph states can be turned into universal resource states with the use of only computational basis measurements. We have the following theorem:

Theorem 2 (Induced grid graphs). Let G be a random d-regular graph on n vertices, with $d = n^c$, where $0.5 < c < 1$. Then, with probability $1 - o(1)$, it contains a square grid graph on v vertices, for any $v = o(n^k)$, with $k = \min\{(1-c)/2, c/3\}$, as an induced subgraph.

This theorem is proven with use of the switching technique due to McKay and Wormald [42]. This is a standard technique in random regular graph theory, but is usually applied in situations where the subgraph to be found is of constant size (whereas for us it must grow reasonably fast with n). Because of this, and because these techniques are not widespread in the quantum computing literature, we give explicit switching calculations in Sec. IV A.

A limitation of this result is the fact that $c > 0.5$. This seems to be an unavoidable fact, as the expected number of grid graphs goes to zero whenever $c < 0.5$. This threshold behavior is observed even for grid graphs of constant size [43]. We can get around this limitation by considering *sparsified* grid graphs, constructing them by replacing every edge in an $L \times L$ grid graph by L edges and $L-1$ vertices in a line. The resulting graphs are very linelike asymptotically yet are still universal (with a polynomial space overhead): we can recover the $L \times L$ grid graph as a *vertex minor* by measuring all the added vertices in the Y basis. For these graphs we can prove the following stronger statement:

Corollary 2 (Sparsified induced grid graphs). Let G be a random d-regular graph on n vertices, with $d = n^c$, where $0.5 < c < 1$. Then, with probability $1 - o(1)$, it contains a sparsified square grid graph on v vertices, for any $v = o(n^k)$, with $k = \min\{(1-c)/2, c/3\}$, as an induced subgraph.

Corollary 9 implies that random n^c-regular graphs of intermediate degree $c \in (0,1)$ are universal resources for measurement-based quantum computing when measured in an arbitrary product basis. Note, however, that *finding* the induced sparsified grid graph may be (and probably is) a computationally difficult problem. Therefore, the standard reduction from sampling to computing probabilities of postselected polynomial-time quantum computations may not be possible in polynomial time. However, since identifying a grid graph can be done with access to an NP oracle, the reduction is possible in the polynomial hierarchy. Thus, while an efficient exact sampling algorithm in an arbitrary basis from a grid graph collapses the polynomial hierarchy to the third level [20], it still collapses it to the fourth level for graphs containing induced sparsified grid graphs.

Finally it must be noted that the proofs of the above results explicitly break down in the regime where $d = cn$ with $c \in (0,1)$. More strongly, it is known that in this regime one cannot find large (much larger than $\log(n)$-sized) nontrivial *induced subgraphs* with more than negligible probability [44]. This leads one to suspect that graph states of linear degree are not universal. We suspect that it is the case that they are universal on average, but we have no proof of this. We do provide some evidence in this direction, which we discuss in the next section.

4. Absence of a geometric entanglement barrier for uniformly random graph states

Entanglement is usually considered to be a necessary condition for universal measurement-based quantum computation. However, it was proven in Ref. [28] that a state can be “too entangled” to be used as a resource state (and also that—under the Haar measure—most states are too entangled in this way). This was done by arguing that if a resource state has high geometric entanglement, defined as

$$E_g(|\psi\rangle) = -\log\left(\max_{\alpha\in\mathrm{PROD}_n} |\langle\alpha|\psi\rangle|^2\right), \tag{3}$$

where PROD_n is the set of product states, then any NP-problem solved by MBQC with $|\psi\rangle$ as a resource state can be solved by a classical computer (with access to randomness) in time $O\left(\mathrm{poly}(n)2^{n-E_g(\psi)}\right)$. In light of the failure of the arguments above in the $d=\Theta(n)$ regime one can wonder whether this entanglement barrier shows up for random graph states [of degree cn with $c\in(0,1)$]. We give evidence that this is not the case by proving an upper bound on the geometric entanglement of uniformly random graph states.

Theorem 3 (Geometric entanglement upper bound for graph states). Choose a graph state $|G\rangle$ uniformly at random. There exist constants c and C such that

$$\mathbb{P}\left[E_g(|G\rangle) \geq n - cn^{1/4}/\log(n)\right] \leq C. \tag{4}$$

The proof of this theorem establishes a connection between the geometric entanglement of random graph states and the behavior of the ranks of the principal submatrices of random adjacency matrices, which can subsequently be analyzed with the use of ideas from extremal probability and Markov chain Monte Carlo techniques [45]. This upper bound is strong enough to exclude the simulation algorithm given in Ref. [28], which now runs in time $\Omega(2^{n^{1/4}})$. We note that this upper bound does not directly translate to a similar bound on d-regular graphs with $d=cn$. However, these distributions are very similar. For instance, the sandwich theorem [46] tells us that uniformly cn-regular graphs can be closely related to Erdős-Rényi random graphs with constant acceptance probability p (our result can be interpreted as evaluating the $p=1/2$ point). Furthermore, uniformly random graphs are with high probability quite regular, with most vertices having degree $n/2\pm O(\sqrt{n})$. We leave making these connections fully rigorous for future work.

We also provide strong evidence that this upper bound is almost tight, by providing a lower bound on the geometric entanglement of random stabilizer states (which are equivalent to graph states up to local Clifford operations).

Theorem 4 (Geometric entanglement lower bound for stabilizer states). Choose a stabilizer state $|S\rangle$ uniformly at random. There exists a constant c such that

$$\mathbb{P}\left[E_g(|S\rangle) \leq n - c\sqrt{n}\log(n)\right] \leq O(2^{-\sqrt{n}}). \tag{5}$$

This theorem is a straightforward application of the representation theory of the Clifford group. Because a constant fraction of stabilizer states are graph states (up to linear phases, which leave the geometric entanglement unchanged) this leads to a similar statement for graph states.

Corollary 3 (Geometric entanglement lower bound for graph states). Choose a graph state $|G\rangle$ uniformly at random. There exists a constant c such that

$$\mathbb{P}\left[E_g(|G\rangle) \leq n - c\sqrt{n}\log(n)\right] \leq O(2^{-\sqrt{n}}). \tag{6}$$

We believe [ignoring the $\log(n)$ factors] that the $O(\sqrt{n})$ deviation in Corollary 3 is accurate, and the slightly weaker $\Omega(n^{1/4})$ scaling in Theorem 3 is a consequence of the proof technique. We leave closing the gap between the upper and lower bounds for future work.

C. Proof ideas

We presented results in three different regimes of regularity. The proof techniques in these three regimes are all quite different, drawing on results in combinatorics, random graph theory, and random matrix theory. Here we outline, organized by regime, the techniques used in this paper.

1. Constant-degree graph states and random pairing and matching IQP circuits

The key result on graph states of constant degree is Theorem 1, which concerns random IQP circuits generated from uniformly random matchings. This result implies anticoncentration of the X-Y plane output distribution of constant-degree regular graph states. To show this theorem, we use a combination of techniques. First, we reduce the problem to a purely graph-combinatorial problem using a statistical-mechanics interpretation of the expected second moment of the output distribution described in Ref. [13]. Importantly, this model is distinct from similar models for quantum circuits with Haar-random single-qubit gates [47–49] and is significantly more involved to analyze. The resulting combinatorial problem, which amounts to counting the number of matchings which have an even number of edges crossing between two arbitrary subsets of vertices, can be further interpreted as a sum over so-called Krawtchouck polynomials [29]. These have seen use in coding theory [30], and good bounds are available [50], which allow us (with a substantial amount of combinatorial elbow grease) to provide bounds on the expected second moment.

2. Intermediate-degree graph states

The main results on intermediate-degree graphs (Theorems 2 and 2) are essentially about proving bounds on the appearance of induced subgraphs of random d-regular graphs. Proving such bounds is a well-studied problem in graph theory, both in the case of d-regular graphs or Erdős-Rényi graphs and in the induced and standard subgraph cases (see, e.g., Ref. [43] or the book by Bollobás [51]). However, in the literature usually only the case of constant-sized subgraphs is treated explicitly, whereas we require bounds for the appearance of induced subgraphs that grow quite fast with n. Thus, we prove Theorem 2 by a careful application of existing combinatorial methods, in particular the method of switchings, introduced by McKay and Wormald [42], which is particularly effective at analyzing expectation values of random variables induced by random d-regular graphs, and the second moment method, to convert expectation value estimates to statements that hold with high probability. The difficulty here lies almost entirely in the care required to get nontrivial estimates for induced subgraphs growing in size with n.

3. High-degree graph states

The main results on graph states of high degree are Theorems 3 and 4, providing lower and upper bounds on the geometric entanglement. Theorem 4 follows a relatively standard path, approximating the continuous optimization in the geometric entanglement by a discrete one through an ϵ-net, followed by a union bound and a tail bound on the overlaps with fixed product states. There is some subtlety in that we require a rather small ϵ-net for the union bound to be nontrivial. To that end we extend a nice trick from random matrix theory to the geometric entanglement in Lemma 13, proving that we can obtain a multiplicative approximation to the maximum in the geometric entanglement using a relatively small ϵ-net. The subsequent tail bound is then provided by the moments of random stabilizer states, developed in Ref. [52]. Some care must be taken to choose the right moment here, as the moments of random stabilizer states grow too fast for a straightforward exponential generating function approach to work [53].

Proving Theorem 3 is substantially more complicated. We restrict the optimization in the geometric entanglement to a special subset of product states for which we can characterize the overlap purely in terms of the corank of principal submatrices of the adjacency matrix of the graph state. This reduces bounding the geometric entanglement to an extremal probability problem on random symmetric binary matrices. Inspired by similar arguments in the theory of Gaussian processes [54], we then bound the correlation between the coranks of different principal submatrices of a single random adjacency matrix. We prove that if the overlap between the two matrices is small, then their coranks are approximately independent. This is then enough to prove a lower bound on the maximal corank, via the second moment method (in particular we use the Bonferroni inequalities). Proving this approximate independence is done by reducing the problem to a Markov chain on the integers $\mathbb{N}$, for which we then prove precise (nonspectral) upper bounds on the convergence to the stationary state.

D. Context and prior work

Our work builds on first steps made in Ref. [19] that classify the simulation complexity and entanglement properties in the *worst case* over the choice of d-regular graphs. There, it was shown that as d is increased, the simulation complexity and entanglement properties undergo two phase transitions, providing a tight link between complexity and entanglement: for $d \leq 2$ and $d \geq n-3$ the entanglement is low and simulating single-qubit measurements in any basis is classically easy for all d-regular graph, while for any other value of d there is a d-regular graph for which simulations are classically intractable and the multipartite entanglement is high. But are these hard instances isolated in their respective regularity class or are *most* instances of a class hard? In many cases, average-case complexity significantly differs from worst-case complexity, most famously so for NP-problems, where often most instances are efficiently solvable.

Thus, our results on *average-case hardness* of random d-regular graph states significantly strengthen the connection between the connectivity of a graph and the entanglement properties of the corresponding state and provide complexity-theoretic evidence that classical intractability is a generic feature of multipartite entangled states. Generally speaking, though, there are only a few techniques to address average-case complexity such as random self-reducibility—these techniques primarily involve reducing the average-case problem to proving average-case hardness of computing the permanent of a matrix (see, e.g., Refs. [26,55,56])—which work for Haar-random gate sets since they are continuously parameterized but are not applicable to the discrete randomness of random graphs. This is why the average-case complexity of random graph states is a qualitatively different question from both their worst-case complexity and the average-case complexity of continuous ensembles.

From a different perspective, our work can be viewed as a first attempt to answer the question of whether random regular graph states are resources for MBQC [and in the regime of $d = \Theta(n^c)$, we answer in the affirmative]. Assessing which graphs form resource states for MBQC has a rich history [1,7,36,57,58] and has motivated the idea of computational phases of matter [59–65]. This question has also arisen in the study of the impact of noise on computational power. Concretely, Browne *et al.* [66] studied the impact of erasure noise on MBQC on grid graphs in terms of percolation phenomena. Here, at each lattice site, there is a finite probability that the qubit on that lattice site is erased, resulting in a random cluster state as originally introduced by Briegel and Raussendorf [58]. Below a certain value of that probability—known as the percolation threshold—noisy grid graphs can then be classically simulated, whereas above that probability, noisy grid graphs remain universal resource states. Equivalently, we can view this setting as the random graph state ensemble given by random subgraphs of the grid graph [58], which is an ensemble rather different from the ones studied in this paper.

E. Discussion and outlook

To our knowledge, our results are the first to explore the average-case complexity of random graph states with bounded degree. They give evidence toward the average-case complexity of simulating graph states of any degree $2 < d \lesssim n/2$ using different types of results relating to the hardness and simulability of random graph states. In particular, our results interpolate between these degree regimes. This interpolation runs via the degree of regular graphs and thus extends the results of Ref. [19] to the average case.

However, they can by no means be said to be the final word on the average-case complexity of regular graph states, or low-depth quantum circuits. Thus, they raise a number of interesting questions. In particular, they have interesting consequences when related to a variety of different themes in the study of simulating sampling from quantum circuits. In our discussion, we will discuss each regularity regime in turn and formulate a number of open questions and conjectures.

1. Constant-degree graph states and random matching IQP circuits

Theorem 1 gives evidence for the average-case hardness of simulating a family of random quantum circuits with random connectivities of any depth larger than 2 and scaling slower than the system size n. It is the first circuit family for which anticoncentration has been shown extending from any constant depth to sublinear depth.

Similar constructions of universal random circuits based on regular graphs with perfect matchings were previously studied numerically and experimentally in Ref. [12]. Since commuting IQP circuits have often been precursors to results for random circuits (e.g., in terms of complexity [33,67] and noisy simulation [23,68,69]), our result may thus also help in the rigorous study of constant-depth random circuits in arbitrary geometries. In particular, it is worth considering our results in relation to the results obtained by Napp *et al.* [24], who give evidence that average-case hardness fails for constant-depth Haar-random circuits, as well as to the results obtained by DeCross *et al.* [12], who give evidence that the simulation complexity of random circuits on different connectivities remains bounded at very low depths. To see why our results are different, we observe that both those studies consider random circuits in which arbitrary single-qubit rotation gates can be applied throughout the circuit. It may indeed be that in this case average-case hardness requires a strictly super-constant circuit depth. At least depth $\Omega(\log\log n)$ is required for anticoncentration in models that are invariant under Haar-random single-qubit gates [25]. Our results circumvent this lower bound by restricting the local bases we measure in to the X-Y plane. Hence, the fact that we measure only in the X-Y plane is critical to the anticoncentration in constant depth we show. It remains an interesting question, however, how the simulation complexity of such random circuits depends on the circuit depth. As a first step toward this, it would be interesting to consider the random matching IQP circuits with measurements in a Haar-random single-qubit basis. By the result in Ref. [25], anticoncentration will fail for constant depth, and thus this model would be helpful to understand the mechanism governing the arising lower bound.

It is also interesting to consider the relation of our results to recent results for simulating noisy IQP circuits [70]. In Ref. [70] it is shown that noisy IQP circuits at any depth larger than a noise-dependent constant can be efficiently classically simulated, and one might think that our results might yield some leeway to circumvent these results. However, note that a previous noisy simulation algorithm due to Bremner *et al.* [23] will work for any IQP circuits which anticoncentrate, and therefore also applies to the circuit families we consider here. Noisy constant-depth IQP circuits might also be interesting to consider in the context of the recently discovered transition in the cross-entropy benchmark (XEB) versus the fidelity [71–74]. While IQP circuits giving rise to Erdős-Rényi random graph states with some edge probability p exhibit the transition in the XEB [13], this transition may be understood as occurring because in such graphs the probability that no gate is applied to a particular qubit vanishes only exponentially in d rather than nd. In contrast, for IQP circuits giving rise to d-regular graph states, entangling gates are guaranteed to be applied to every qubit, and thus the transition may disappear. This would allow them to be reliably and (sample-)efficiently benchmarked with the XEB.

2. Intermediate-degree graph states and universality for measurement-based computation

Our results focus on the question of whether random graph states can be classically simulated, which is *a priori* a question which is distinct from the question of whether random graph states are universal resources for measurement-based quantum computation. While some of our results (in the intermediate regime) explicitly make use of universal resources, it is interesting to ask in what sense or to what extent random graph states in different regimes can function as universal resource states. To answer this question, it must first be made clear that the notion of a "universal resource" can have very different meanings. A natural notion of universality is one in which we just ask that for any number of qubits n, the family of states contains a state on $m \geq n$ qubits such that with use of measurements and feedforward an arbitrary unitary can be implemented on n of the m qubits up to some precision threshold [36]. Note that this notion is not concerned with efficiency, and hence the graphs containing a 2D grid graph in the intermediate regularity regime are certainly examples of universal resources in this sense. The universal resource is even efficient in the sense that $m = \mathsf{poly}(n)$. However, note that these graphs cannot be efficiently found in general, and finding them may require the solution of an NP-problem (showing this is an interesting open question). To computationally exploit a universal resource, it is a prerequisite that a sufficiently large 2D subgraph can also be efficiently found, since the only known efficient constructions make use of those graphs. This raises the question of whether in the intermediate-degree regime certain random ensembles contain grid graphs as subgraphs that can also be efficiently found.

In the regimes of very high and very low degree, whether or not random regular graph states are universal resources remains an interesting open questions. In particular, in those regimes, we do not expect there to be large grid graphs or hexagonal graphs as induced subgraphs of many regular graphs. But all known ways of compiling a quantum computation in a measurement-based way make explicit use of the presence of such a subgraph. We believe the region in which explicit large grid graphs can be found can be expanded (e.g., into the $d = cn$ regime) but only by going beyond induced subgraphs. Ideally, one would like to characterize which regular graphs have large grid graphs as *vertex minors*, which is the graph-theoretic notion capturing the action of local Clifford gates and Pauli measurements. The vertex minor problem is NP-hard in the worst case [75], but little is known about its average-case behavior. Even to understand whether grid graphs on an exponentially small or even constant subset of qubits can be found remains an interesting open question. There might also be resources for measurement-based quantum computation which are not equivalent to hexagonal or grid graphs. We think this is a fruitful avenue of future research.

3. High-degree graph states

In the regime of high degree we also lack a full characterization of the complexity of graph states, in particular the question of classical simulability is not fully settled. While we have excluded the simulation algorithm given in Ref. [28], we cannot exclude the existence of more tailor-made classical algorithms for MBQC with random linear-degree graph states. In Sec. V C we describe a simulator where all Pauli measurements are explicitly calculated and only non-Pauli measurements are simulated. We believe that even this more sophisticated algorithm fails, but cannot prove it as of yet. We leave this as a conjecture. It would also be interesting to prove variants of our results for random cn-regular graph states with $0 < c < 1/2$, which we believe can be done through the sandwich theorems from random graph theory [46].

4. Toward analyzing noisy, architecture-constrained graph states

Aside from the different regimes of the regularity parameter just discussed, an interesting open question raised by our work is the complexity of graph states that can be naturally realized in different quantum computing architectures. The arbitrary connectivity required to generate the random regular graphs may be difficult to realize in practice. For instance, while in principle reconfigurable atom or ion arrays [11,12] arbitrary connectivity between qubits can be realized, there may be more natural random ensembles of graph states, for instance, subgraphs of a high-dimensional hypercube [13]. In other architectures, such as superconducting qubits, even more constrained low-dimensional lattice geometries are imposed [76].

A further interesting question in the context of more realistic circuit ensembles is to what extent noise affects the output distribution. In the literature on sampling from the output distributions of random quantum circuits, a prominent quantity is the XEB score, which generalizes anticoncentration to noisy circuits and serves as a measure of the quality of the sampled distribution [67,76,77]. An important result in this context is that for local noise rates on the order of $\lesssim 1/n$ and where the ideal output distribution anticoncentrates, the XEB can be used as a proxy of the many-body fidelity of the premeasurement state [71,74,78], and moreover, that this is the regime in which there are no exploits that classical algorithms might use to achieve a high XEB score significantly more efficiently than brute-force simulations. In contrast, in the regime where the noise rate is larger than $1/n$, such exploits exist [73]. Similar statements can also be made for IQP circuits [11,13], and it is an interesting open question to analyze the XEB for noisy random graph state ensembles.

II. PRELIMINARIES

In this section we recall some assorted facts about graphs, graph states, and stabilizer states that we will need throughout the rest of the paper. This is by no means meant to be an exhaustive introduction. For graph theory (with a focus on random graphs) we recommend the classic textbook by Bollobás [51], and good introductions to graph states and stabilizer states can be found in Refs. [9,79].

We will also require properties of Krawtchouk polynomials as well as bounding techniques for the convergence of Markov chains, which we also recall in this section.

A. Graphs, graph states, and stabilizer states

We begin by reviewing some standard graph-theoretic notions. A graph G is a set of vertices V (usually the set $[n]$, and a set of vertices $E \subset V \times V$ connecting them). We denote by $\overline{G}$ the complement graph, which has the same vertex set, and the complementary set $\overline{E} \in V \times V$ of edges. We will occasionally be somewhat sloppy in notation and write $e \in G$ ($e \notin G$) to indicate that e is (not) an edge in G [and thus $e \in E$ ($e \notin E$)]. Similarly we will sometimes write subsets of edges as $S \subseteq G$. A subgraph H of G is obtained from G by considering a subset of the edges $E' \subseteq E$, and an induced subgraph H is obtained by considering a subset of the *vertices* $V' \subseteq V$, which has edges $E' = E \cup V' \times V'$.

The symbol $G_{n,d}$ refers to an n-vertex d-regular graph, and $\overline{G}_{n,d}$ refers to the complement of the same graph . We drop the subscripts when the number of vertices and the regularity parameter are clear from the context. For a graph G, we define $m(G)$, the density of the graph, as

$$m(G) = \max\left\{\frac{|E|_{H'}}{|V|_{H'}}, H' \subseteq G, |V|_{H'} > 0\right\}. \tag{7}$$

An n-qubit graph state $|G\rangle$ is defined in terms of an n-vertex graph G as

$$|G\rangle = \prod_{(i,j):U_G[i,j]=1} CZ_{i,j} \left|+^n\right\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{x^T U_G x} |x\rangle , \tag{8}$$

where U_G is the upper triangle of the adjacency matrix A_G of G and the inner product is taken over $\mathbb{F}_2$. Graph states are a type of stabilizer state (which we denote by $|S\rangle$), which means they are the joint $+1$ eigenvectors of a set of 2^n mutually commuting Pauli matrices. There are $2^{n(n-1)/2}$ graph states and $2^n \prod_{i=1}^{n}(2^i+1)$ stabilizer states. A key property of stabilizer states is the following expectation value (taken uniformly over the stabilizer states), first derived in Ref. [52]:

$$\mathbb{E}_{|S\rangle} |S\rangle\langle S|^{\otimes t} = \frac{1}{2^n \prod_{i=0}^{t-2}(2^n+2^i)} \sum_{T\in\Sigma_{t,t}} R(T), \tag{9}$$

where $\Sigma_{t,t}$ is the set of Lagrangian subspaces of $\mathbb{F}_2^{2t}$ with respect to a particular generalized quadratic form, and $R(T)$ is an nt-qubit representation of this set (which can be given a semigroup structure). We will need very few properties of this set (see Ref. [52] for an exhaustive treatment), only that

$$|\Sigma_{t,t}| = \prod_{i=0}^{t-1}(2^i+1), \tag{10}$$

$$\mathrm{tr}(|\beta\rangle\langle\beta|^{\otimes t} R(T)) \leq 1 \tag{11}$$

for all states $|\beta\rangle$. The latter statement follows by combining Proposition 56 and Theorem 72 in Ref. [80] (a more direct statement, using a different proof technique, can also be found as Corollary 6.11 in Ref. [81]).

B. Random graph models

We will consider three different models of random regular graphs (with regularity parameter d), each with slightly different properties.

Definition 1 (Uniformly random d-regular graphs). $G(n,d)$ is the distribution over graphs generated by choosing d-regular graphs uniformly at random.

Definition 2 (Matching model of d-regular multigraphs). $G_m(d,n)$ for n even is the distribution over regular multigraphs of degree d generated by choosing d matchings independently uniformly at random and composing the result.

Definition 3 (Pairing model of d-regular multigraphs). $G_p(d,n)$ for dn even is the distribution over regular multigraphs of degree d generated by choosing a uniformly random matching M on the set $\{(i,j) \| i \in [n], j \in [d]\}$ and adding an edge between $i, \hat{i}$ whenever (i,j), $(\hat{i},\hat{j})$ is an edge in M.

All of these models can be efficiently sampled from, even when d is relatively large [82]. Note that the pairing model allows for both multiple edges between vertices and self-edges (loops). For d constant, the pairing model and the uniformly random graph model are related. The following lemmas are well known in the graph theory literature [34]:

Lemma 1. The probability that the pairing model yields a simple d-regular graph is bounded from below as

$$\mathbb{P}\left[G \text{ is simple}\right] \geq 2^{-d^2}. \tag{12}$$

Lemma 2. Conditioned on being simple, the graphs obtained from the configuration model are *uniformly distributed* d-regular graphs.

We have a similar simplicity condition in the matching model.

Lemma 3. The probability that the random matching model yields a simple d-regular graph is bounded from below as

$$\mathbb{P}\left[G \text{ is simple}\right] \geq 2^{-d}. \tag{13}$$

However, in this model, conditioning on simplicity does not yield a uniform distribution on d-regular graphs. This can be easily seen by noting that there exist d-regular graphs that do not contain a perfect matching. The conditional distribution is equivalent to uniformly random d-regular graphs in a weaker sense, called "contiguity" [35], meaning that an event that happens with high probability in one distribution also happens with high probability in the other. We will not make use of this connection in this paper and consider the matching model to be interesting in its own right.

C. Krawtchouck polynomials

Krawtchouk polynomials are a family of polynomials that prominently appear in classical error correction codes and Boolean analysis. We will need them in the proof of Theorem 6. We will briefly recap their definition here as well as a number of upper bounds. The (binary) Krawtchouk polynomial of degree i and size N is defined as

$$K_i^N(x) := \sum_{q=0}^{i} (-1)^q \binom{x}{q}\binom{N-x}{i-q}. \tag{14}$$

These polynomials are orthogonal under a binomial measure (see, e.g., Ref. [30]), i.e.,

$$\sum_{t=0}^{N} \binom{N}{t} K_i^N(t) K_j^N(t) = 2^N \binom{N}{i} \delta_{ij}. \tag{15}$$

This immediately implies a pointwise upper bound for integer evaluations of the polynomial of the form

$$|K_i^N(t)| \leq 2^{N/2} \binom{N}{i}^{1/2} \binom{N}{t}^{-1/2}. \tag{16}$$

This is a rather straightforward upper bound but it will service almost all of our needs (it is also surprisingly close to being tight, see Ref. [83]). However, to cover certain parameter regimes we will also need a more sophisticated bound from Corollary 16 in Ref. [50] of the form

$$|K_i^N(t)| \leq \binom{N}{i} \left(\frac{i}{N} + \frac{(N-t)^2}{N^2}\right)^{i/2}. \tag{17}$$

D. Drift and minorization of Markov chains

Consider a Markov chain P on a (possibly countably infinite) state space $\mathcal{X}$. We will maintain that a Markov matrix acts from the left, i.e., $P(x,x') := \mathbb{P}(X_{t+1} = x' | X_t = x)$ for $x, x' \in \mathcal{X}$. If the Markov chain is irreducible, it will have a unique stationary distribution π. Bounding the convergence time of the Markov chain toward π on unbounded systems is generally tricky. In this paper we use the drift and minorization method [84], which can bound convergence from a fixed starting state to the stationary distribution for an irreducible Markov chain in a way that does not depend directly on the size of $\mathcal{X}$. This method consists of two steps, first bounding the time it takes for the Markov chain to pool into a "small set," and then bounding thermalization within that set. The first requirement (the "drift") is the existence of a *drift* function V which controls the convergence to a *small set* C:

Definition 4 (Drift toward a small set). A function $V : \mathcal{X} \to [0,\infty]$ is a drift function (toward a set $C \subset \mathcal{X}$) for a Markov chain P if there exists a constant $0 < \lambda < 1$ and a constant $b < \infty$ such that

$$PV(x) \leq \lambda V(x) + bI_C, \tag{18}$$

where I_c is the indicator function on the set C and $PV(x) := \sum_{x' \in \mathcal{X}} P(x,x')V(x'))$.

The second requirement is a minorization condition on the small set C

Definition 5 (Minorization on a small set). A Markov chain P satisfies a minorization condition on the set C if there exists a probability distribution ν on C and a constant $\delta > 0$ such that for all $x \in C$ we have

$$P(x,x') \geq \delta \nu(x'). \tag{19}$$

Finally, the drift and minorization conditions are called "compatible" if there exists a constant $d > 2b/(1-\lambda)$ such that the level set $\{x \in \mathcal{X} \parallel V(x) \leq d\}$ is included in the small set C.

If a Markov chain has a drift function and satisfies the minorization condition on a small set C in a compatible way, we can bound the convergence of the Markov chain to its stationary distribution in an exponential way. We will use the following theorem due to Rosenthal (Theorem 12 in Ref. [85]):

Theorem 5. Let P be a Markov chain on a (countable) state space $\mathcal{X}$, compatibly satisfying a drift and minorization condition with function $V(x)$, small set C, and parameters λ, b, δ, d. Let ν be an initial distribution and let

π be the stationary distribution of P. We now have for all $0 < r < 1$ that

$$\left\|P^k\nu - \pi\right\|_{TV} \leq (1-\delta)^{rk} + (1 + 2b/(1-\lambda) + \mathbb{E}_\nu(V)) \times \left[\left(\frac{1+2b+\lambda d}{1+d}\right)^{1-r}(1+2\lambda d + b)^r\right]^k, \tag{20}$$

where $\|\cdot\|_{TV}$ is the total variation distance.

Due to the many free parameters this theorem is quite flexible, but also rather difficult to use (and gives quite conservative bounds). We will use it in the proof of Theorem 3.

III. ANTICONCENTRATION OF GRAPH STATES OF CONSTANT DEGREE

In this section we compute the average (normalized) second moment for two ensembles of random graphs with degree d which induce measures on d-regular graphs.

First, we compute the average second moment for a random multigraph in the random pairing model. Conditioning on simplicity will then give results for uniformly random d-regular (simple) graphs, as discussed in Sec. II B. To obtain the result in Theorem 1 for the random matching model, we will need to slightly adapt the proof of this more complicated case.

Theorem 6 (Anticoncentration of random pairing model graph states: restatement of part (1) of Theorem 1). Consider the uniform measure $\mathcal{G}_p(n,d)$ on d-regular multigraphs chosen from the pairing model. Then, for any $2 < d = o(n^{1/2})$, we have

$$\mathbb{E}_{G\sim\mathcal{G}_P(n,d),\theta}(m_2(G,\theta)) = o(1) + \begin{cases} 2 & \text{if } d = 1 \mod 2, \\ 3 & \text{if } d = 0 \mod 2. \end{cases} \tag{21}$$

Corollary 4 (Restatement of Corollary 1). Consider the uniform measure $\mathcal{G}_r(d,n)$ on d-regular graphs on n vertices, and uniformly random angles θ. Then for any constant d

$$\mathbb{E}_{G\sim\mathcal{G}_r(d,n),\theta}[m_2(G,\theta)] \leq (3+o(1))2^{d^2}. \tag{22}$$

Proof. The corollary follows directly from Theorem 6, observing that the probability that a random multigraph in the pairing model $G \leftarrow \mathcal{G}_p(d,n)$ is simple, is lower bounded by $1/2^{d^2}$, and hence, conditioning on the event that a graph in $\mathcal{G}_p(d,n)$ is simple yields the statement. ■

A. Anticoncentration of a fixed graph state

The proof of Theorem 6 is based on a series of lemmas. As a first step, we find an expression for the second moment of a fixed graph state, averaged over the random choice of measurement angles. This follows more or less directly from the discussion in Appendix E of Ref. [13].

Lemma 4. Consider a graph G and a uniformly random choice of X-Y-plane measurement angles θ. Then we have

$$\mathbb{E}_\theta(m_2(G,\theta)) = 2^{-n}\sum_{L,R\subset[n],L\cap R=\emptyset}(-1)^{|A_G[L,R]|}, \tag{23}$$

where A_G is the adjacency matrix of G. $|A_G[L,R]|$ denotes the sum of the entries of the submatrix $A_G[L,R]$ of A_G corresponding to rows in L and columns in R, and counts the number of edges crossing between L and R.

Proof. To show the lemma, we use the statistical-mechanics mapping of second moments of IQP circuits derived in Ref. [13], extended to arbitrary single-qubit Z rotations. We keep the discussion of this model brief here and refer the reader to Appendix E of Ref. [13] for the derivation of the statistical-mechanics mapping. We start by observing that the second moment operator

$$M_2 = \mathbb{E}_\psi \left|\psi\right\rangle\left\langle\psi\right| \otimes \left|\psi\right\rangle\left\langle\psi\right| \tag{24}$$

of a ensemble of random states $|\psi\rangle$ determines the second moment of its outcome distribution p_ψ as $\mathbb{E}_\psi p_\psi(x)^2 = \langle x|^{\otimes 2} M_2 |x\rangle^{\otimes 2}$, and thus computing the average second moment can be reduced to computing the second moment operator of the underlying state ensemble. For a random product rotation around the Z axis applied to the $|+\rangle^{\otimes n}$ state, the second moment is given by

$$\mathbb{E}_\theta\left[\mathrm{e}^{-i\sum_j\theta_j Z_j}\left|+^n\right\rangle\!\left\langle+^n\right|\mathrm{e}^{i\sum_j\theta_j Z_j}\right] = \frac{1}{4^n}(\mathsf{I}+\mathsf{S}+\mathsf{X})^{\otimes n} = \frac{1}{4^n}\sum_{Q\in\mathcal{S}^{\otimes n}} Q, \tag{25}$$

where $\mathsf{I} = |01\rangle\langle 01| + |10\rangle\langle 10|$, $\mathsf{S} = |01\rangle\langle 10| + |10\rangle\langle 01|$, $\mathsf{X} = |00\rangle\langle 00| + |11\rangle\langle 11|$, and we let $\mathcal{S} = \{\mathsf{I},\mathsf{S},\mathsf{X}\}$ be the set of possible "states" of the statistical-mechanics model. Next, we observe that

$$CZ^{\otimes 2}(P\otimes Q)CZ^{\otimes 2} = \begin{cases} -P\otimes Q & \text{if } P = \mathsf{I}, Q = \mathsf{S} \wedge P = \mathsf{S}, Q = \mathsf{I}, \\ P\otimes Q & \text{else}, \end{cases} \tag{26}$$

and that $\sum_{x\in\{0,1\}} \langle xx| H^{\otimes 2} Q H^{\otimes 2} |xx\rangle = 1, \ \ \forall Q = \mathsf{I},\mathsf{S},\mathsf{X}$. Hence, the sign of a particular "state" $Q \in \mathcal{S}^{\otimes n}$ on a graph $|G\rangle$ is given by the parity of the CZ gates that act on an IS or SI pair.

Thus, the second moment of the outcome distribution $p_{G,\theta}$ of measuring a graph state $|G\rangle$ in the X-Y plane angles θ, averaged over the random choice of θ, can be written as

$$\mathbb{E}_\theta[m_2(G,\theta)] = \frac{1}{2^n} \sum_{Q\in\mathcal{S}^{\otimes n}} (-1)^{N(Q,G)}, \tag{27}$$

where for $G = (V, E)$ with edge set E we let $N(Q,G) = |\{(e_0, e_1) \in E : Q_{e_0} \otimes Q_{e_1} \in \{\mathsf{I}\otimes\mathsf{S}, \mathsf{S}\otimes\mathsf{I}\}\}|$ be the number of edges in the graph G coinciding with an IS or SI pair in the state Q. Since this expression depends only on the locations of the graph edges and I and S states, we can rewrite it in the form of Eq. (23). ■

Importantly, this expression can also be interpreted for multigraphs, where the elements of the adjacency matrix are now integers. The crucial number $|A_G[L,R]|$ featuring in Eq. (23) thus still counts the number of edges crossing between the sets L and R.

B. Averaging over matchings

Next, we need to evaluate the average over the random choice of (multi)graphs. Since our graphs are defined in terms of uniformly random matchings, we begin by analyzing the average parity of edges in a random matching crossing between two disjoint sets $L, R \subset [n]$. We first note some symmetry properties of the average parity in Lemma 5, and then express the average parity in terms of Krawtchouk polynonmials in Lemma 6. In the following, let $\mathcal{M}(n)$ denote the uniform distribution over perfect matchings of n vertices.

Lemma 5. Consider a random matching M on n vertices and three disjoint sets $L, R, T \subseteq [n]$ covering $[n] = R \cup L \cup T$. The value of the average parity of the matching between two sets is invariant under the interchange of R, L, T up to factors of ± 1 as

$$\mathbb{E}_{M\sim\mathcal{M}(n)}(-1)^{|A_M(L,R)|} = \mathbb{E}_{M\sim\mathcal{M}(n)}(-1)^{|A_M(R,L)|}, \tag{28}$$

$$\mathbb{E}_{M\sim\mathcal{M}(n)}(-1)^{|A_M(L,R)|} = (-1)^{|L|}\mathbb{E}_{M\sim\mathcal{M}(n)}(-1)^{|A_M(L,T)|}. \tag{29}$$

Proof. Clearly the average parity is invariant under the interchange of the arguments, i.e.,

$$\mathbb{E}_M(-1)^{|A_M(L,R)|} = \mathbb{E}_M(-1)^{|A_M(R,L)|}. \tag{30}$$

To see Eq. (29), consider a particular matching that has q edges going from L to R [and thus contributes $(-1)^q$ to the average]. This matching also has $0 \le q' \le n-q$ edges going from L to T. By construction the remaining vertices in L must be matched to each other. Hence, $q + q' = |L| \mod 2$ and therefore $(-1)^q = (-1)^{q'}(-1)^{|L|}$. Since this is true for every matching the claim follows. ■

Lemma 6. Consider two disjoint sets $L, R \subseteq [n]$ such that $|L|, |R| \le n - |L| - |R|$. Then the average parity of the number of edges of a random matching between L and R is given by

$$\mathbb{E}_{M\sim\mathcal{M}(n)}(-1)^{|A_M(L,R)|}$$
$$= \sum_{\substack{i=0 \\ i=|L| \mod 2}}^{|L|} K_i^{n-|L|}(|R|)\frac{|L|!(n-|L|-i-1)!!}{(|L|-i)!!(n-1)!!}, \tag{31}$$

where K_i is the Krawtchouk polynomial of degree i.

Proof. Denote $T = [n]\backslash(R\cup L)$. Consider a matching that has q edges between the sets L and R and q' edges between L and T. Note that this is possible only if $q + q' = |L| \mod 2$ and $q + q' \le |L|$. There are $\binom{|L|}{q}\binom{|R|}{q}\cdot q!$ ways to choose the first q edges and $\binom{|L|-q}{q'}\binom{|T|}{q'}\cdot q'!$ ways to choose the second set of edges. Once these edges are fixed, there are $(|L| - q - q' - 1)!!(n - |L| - q - q' - 1)!!$ ways to complete this edge set to a full matching of $[n]$. Hence, we can write

$$\mathbb{E}_M(-1)^{|A_M(L,R)|} = \frac{1}{(n-1)!!} \sum_{\substack{q,q'\ge 0 \\ q+q'\le|L| \\ q+q'=|L| \mod 2 \\ q\le|R|}} (-1)^q q! \binom{|L|}{q}\binom{|R|}{q} q'! \binom{|L|-q}{q'}\binom{|T|}{q'}$$
$$\times (|L| - q - q' - 1)!!(n - |L| - q - q' - 1)!!, \tag{32}$$

since there are $(n-1)!!$ matchings in total. Since $\binom{|R|}{q} = 0$, if $|R| < q$ we can drop the $q \le |R|$ constraint going forward. We can expand the binomials involving $|L|$ and use the identity $(a-1)!!a!! = a!$ to rewrite this as

$$\mathbb{E}_M(-1)^{|A_M(L,R)|} = \sum_{\substack{q,q'\geq 0\\ q+q'\leq |L|\\ q+q'=|L| \mod 2}} (-1)^q \binom{|R|}{q}\binom{|T|}{q'} \frac{|L|!(n-|L|-q-q'-1)!!}{(|L|-q-q')!!(n-1)!!}. \tag{33}$$

Changing variables $q,q' \to q, i = q+q'$, we can rewrite this further as

$$\mathbb{E}_M(-1)^{|A_M(L,R)|} = \sum_{\substack{i\geq 0\\ i=|L| \mod 2}}^{|L|} \sum_{q=0}^{i} (-1)^q \binom{|R|}{q}\binom{|T|}{i-q} \frac{|L|!(n-|L|-i-1)!!}{(|L|-i)!!(n-1)!!}, \tag{34}$$

in which we recognize the definition [Eq. (14)] of the Krawtchouck polynomial. ■

C. Proof of Theorem 6

We can now provide a proof of the core result of this section (Theorem 6). This proof is long, but much of the conceptual work has been done in previous subsections (with “merely” some counting remaining).

Proof of Theorem 6. Lemmas 5 and 6 imply that all that is relevant for the average parity of the edges between two subsets $L, R \subseteq [n]$ is the size of the subsets. Recall that these sets correspond to I and S states on a subset of the vertices and let us think of the different states on the vertices as one of three colors ($L \leftrightarrow$ I, $R \leftrightarrow$ S, $T \leftrightarrow$ X). To analyze the random pairing model, we observe the following: for a fixed choice of $L, R \subseteq [n]$ in the sum given by Eq. (23), the average parity $\mathbb{E}_{G_M}[(-1)^{|A_{G_M}[L,R]|}]$ over graphs G_M induced by a matching M (the number of edges between I-colored vertices and S-colored vertices) is equal to the average parity $\mathbb{E}_M[(-1)^{|A_M[L_d,R_d]|}]$ of two sets $L_d, R_d \subseteq [dn]$ of size $|L_d| = |L|d, |R_d| = |R|d$ over the inducing matching. This is because we can “split up” every vertex in G_M into d vertices of the same color. All that matters for the average parity is the crossings between colors blowing up the sets $L \to L_d, R \to R_d$. ■

Hence, it follows from Lemmas 4–6 that

$$\begin{aligned}
&\mathbb{E}_{G_M\sim\mathcal{G}_p(n,d),\theta}(m_2(G_M,\theta)\\
&\quad = 2^{-n}\sum_{L,R\subset[n],L\cap R=\emptyset}\mathbb{E}_{G_M\sim\mathcal{G}_p(n,d)}(-1)^{|A_{G_M}[L,R]|} && (35)\\
&\quad = 2^{-n}\sum_{L,R\subset[n],L\cap R=\emptyset}\mathbb{E}_{M\sim\mathcal{M}(nd)}(-1)^{|A_M[L_d,R_d]|}. && (36)
\end{aligned}$$

Let us denote (counterintuitively, we are following the notation in Ref. [13]) $|L| = k, |R| = l$ (correspondingly $|L_d| = dk, |R_d| = dl$). Then we can use Lemma 6 to get

$$\mathbb{E}_{G_M\sim\mathcal{G}_p(n,d),\theta}[m_2(G_M,\theta)] = 2^{-n}\sum_{L,R\subset[n]L\cap R=\emptyset}\ \sum_{\substack{i=0\\ i=dk \mod 2}}^{dk} K_i^{d(n-k)}(dl)\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!} \tag{37}$$

$$= 2^{-n}\sum_{k=0}^{n}\sum_{l=0}^{n-k}\binom{n}{k}\binom{n-k}{l}\underbrace{\left[\sum_{\substack{i=0\\ i=dk \mod 2}}^{dk} K_i^{d(n-k)}(dl)\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}\right]}_{=:M_p(n,k,l,d)}, \tag{38}$$

where the second equality counts the number of possible choices of $L, R \subseteq [n]$ with fixed sizes k, l and in the last equality we have implicitly defined $M_p(n,k,l,d)$. To get a feeling for what we need to show, consider Fig. 2. There, we show the term $M_p(n,k,l,d)$ for fixed n, d as a function of k, l. There is a clear separation of what is happening. At the boundaries, i.e., for $k = 0$ or $l = 0$ or $k = l$ the terms are largest, and they decay toward the center.

This suggests dividing up the sum according to this behavior. We can also see the symmetry under the

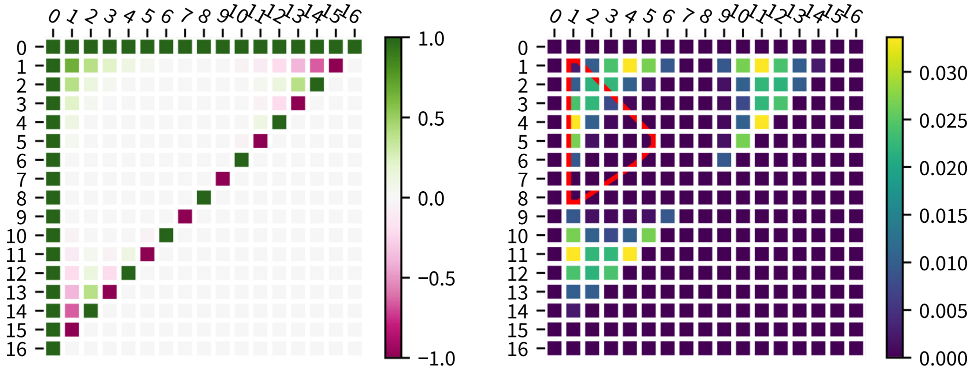

FIG. 2. For $n = 16, d = 3$, we show the dependence on k, l of $M_p(n, k, l, d)$ (left) and $2^{-n}\binom{n}{k}\binom{n-k}{k}|M_p(n, k, l, d)|$ (right) for $k, l > 0$ and $k + l < n$. The red area corresponds to the region $1 \le k \le l \le n - k - l$ in which we need to achieve a nontrivial bound.

exchange $k \leftrightarrow l$, which we will exploit later. Specifically, we will divide the sum over k, l into four cases. The only non-negligible one will be case 0 (corresponding to $k = 0$, $l = 0$, or $k + l = n$).

1. Case 0

Let us begin by considering the case in which one of L, R, T is the empty set, i.e., $k = 0 \vee l = 0 \vee n - k - l = 0$. We begin by treating the $k = 0$ subcase (we label the associated term in the overall sum $T_{k=0}$); the others will follow by symmetry:

$$2^n T_{k=0} = \sum_{R\subset[n]} \mathbb{E}_{M\sim\mathcal{M}(nd)}(-1)^{|A_M[\emptyset,R_d]|} + \sum_{L\subset[n]} \mathbb{E}_{M\sim\mathcal{M}(nd)}(-1)^{|A_M[L_d,\emptyset]|} + \sum_{\substack{L,R\subset[n]\\ L\cap R=\emptyset, L\cup R=[n]}} \mathbb{E}_{M\sim\mathcal{M}(nd)}(-1)^{|A_M[L_d,R_d]|} \tag{39}$$

$$= \sum_{R\subset[n]} \left(2\cdot \mathbb{E}_{M\sim\mathcal{M}(nd)}(-1)^{|A_M[\emptyset,R_d]|} + (-1)^{|R_d|}\mathbb{E}_{M\sim\mathcal{M}(nd)}(-1)^{|A_M[\emptyset,R_d]|}\right). \tag{40}$$

Observe that $|R_d| = ld$ and therefore $(-1)^{ld} = 1$ for even d and $(-1)^{ld} = (-1)^l$ for odd d. By Lemma 5 $\mathbb{E}_{G_M\sim\mathcal{G}_p(n,d)}(-1)^{|A_{G_M}[\emptyset,R]|} = \mathbb{E}_{G_M\sim\mathcal{G}_p(n,d)}(-1)^{|A_{G_M}[\emptyset,[n]\setminus R]|}$ the second term thus vanishes for odd d and is equal to the first term for even d.

We can now simplify the remaining sum to $|L| \le n - |L|$ using Lemma 5, incurring a factor of 2 and counting the number of choices of $L \subset n$:

$$T_{k=0} = 2^{-n} \sum_{l\le n-l} \binom{n}{l} \mathbb{E}_{M\sim\mathcal{M}(nd)}(-1)^{|A_M[\emptyset,R_d]|} \cdot \begin{cases} 4 & \text{if } d = 1 \mod 2, \\ 6 & \text{if } d = 0 \mod 2. \end{cases} \tag{41}$$

Furthermore, we can use Lemma 6 to get

$$\mathbb{E}_{M\sim\mathcal{M}(nd)}(-1)^{|A_M[\emptyset,R_d]|} = K_0^{dn}(dl)\frac{(dn-1)!!}{(dn-1)!!}, \tag{42}$$

and the fact that $K_0^{dn}(dl) = 1$ [see Eq. (14)] to get

$$T_{k=0} = 2^{-n} \sum_{0\le l\le n} \binom{n}{l} \cdot \begin{cases} 4/2 & \text{if } d = 1 \mod 2 \\ 6/2 & \text{if } d = 0 \mod 2. \end{cases} = \begin{cases} 2 & \text{if } d = 1 \mod 2, \\ 3 & \text{if } d = 0 \mod 2. \end{cases} \tag{43}$$

It remains to bound the remaining terms. We begin by defining the sets

$$W_n^* := \{(k,l) \in [n]^{\times 2} : 0 < k \le l \le n-k-l\}, \tag{44}$$

$$\mathcal{S} := \left\{(k,l) \in W_n^* \,|\, 2^{-n}\binom{n}{k}\binom{n-k}{l} \ge n^{-3}\right\}. \tag{45}$$

The set $\mathcal{S}$ captures the terms in the sum where we need to achieve a nontrivial bound. This is because the sum over k,l runs over at most n^2 terms and the average over the parity [the remaining part of Eq. (46)] is bounded by 1, so the sum over all $(k,l) \in W_n^* \setminus \mathcal{S}$ is bounded by $O(1/n)$. Using Lemma 5, we can always assume that $k \le l \le n-k-l$ (this gives a factor of $3! = 6$ in the expression) and thus compute the remaining $k \ge 1$ terms of Eq. (37) as

$$T_{k\ge 1} = 6 \sum_{1\le k\le l\le n-k-l} 2^{-n}\binom{n}{k}\binom{n-k}{l} M_p(n,k,l,d) \tag{46}$$

$$= O(1/n) + 6\sum_{(k,l)\in\mathcal{S}} 2^{-n}\binom{n}{k}\binom{n-k}{l} M_p(n,k,l,d). \tag{47}$$

We will use different bounds in different regimes of the sum over $k,l \in \mathcal{S}$, decomposing $T_{k\ge 1} = O(1/n) + T_{\text{case 1}} + T_{\text{case 2}} + T_{\text{case 3}}$ according to the following three remaining cases:

(1) Case 1: $1 \le k \le 90\log n$.
(2) Case 2: $k \ge 90 \log n$ and $2^{-n}\binom{n}{k}\binom{n-k}{l} \ge n^8$.
(3) Case 3: $k \ge 90 \log n$ and $n^{-3} \le 2^{-n}\binom{n}{k}\binom{n-k}{l} \le n^8$.

We will now treat each of these cases in turn.

2. Case 1 ($1 \le k < 90 \log n$)

For the first case, we will use the upper bound for Krawtchouck polynomials due to Derksen, given in Eq. (17), in order to bound $M_p(n,k,l,d)$:

$$M_p(n,k,l,d) = \sum_{\substack{i=0\\ i=dk \bmod 2}}^{dk} K_i^{d(n-k)}(dl)\frac{dk!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}$$

$$\le \sum_{\substack{i=0\\ i=dk \bmod 2}}^{dk} \binom{d(n-k)}{i}\left(\frac{i}{d(n-k)} + \frac{(n-k-2l)^2}{(n-k)^2}\right)^{i/2} \times \frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}. \tag{48}$$

To further bound this, we consider the factors of Eq. (48) individually:

(1) Factor 1: For the first factor, we have the crude bound

$$\binom{d(n-k)}{i} \le d^i(n-k)^i. \tag{49}$$

(2) Factor 2: To bound the second factor, we need a lower bound on l. We get this by noting that $k,l \in \mathcal{S}$ implies that

$$2^{(n-k)H(l/(n-k))} \ge \binom{n-k}{l} \ge \frac{2^n}{n^3\binom{n}{k}} \ge 2^n n^{-3-k}. \tag{50}$$

Taking the logarithm of both sides and using $H(x) \le 2\sqrt{x(1-x)}$, we obtain

$$\frac{l}{n-k}\left(1-\frac{l}{n-k}\right) \ge \left(\frac{n-(3+k)\log(n)}{2(n-k)}\right)^2. \tag{51}$$

Solving the above quadratic inequality, using that $k \le 90 \log n$, we get the bound

$$\frac{n-k}{2} - C_0\sqrt{n}\log n \le l \le \frac{n-k}{2} \tag{52}$$

for some constant C_0 and all n larger than some constant n_0. Since $l \le n/2$, we thus have

$$\left(\frac{n-k-2l}{n-k}\right)^2 \le \left(1-\frac{1}{2}+\frac{C_0\sqrt{n}\log n}{n-k}\right)^2$$
$$\le \left(\frac{1}{2}+\frac{C_0\log n}{\sqrt{n}-90\log n/\sqrt{n}}\right)^2$$
$$\le C_1\log^2(n)$$

for all $n \ge n_0$ and some constant C_1. Moreover,

$$\frac{i}{d(n-k)} \le \frac{90\log(n)}{n-90\log n} \le 1 \le C_2\log^2(n). \tag{53}$$

Altogether, we thus get for all $n \ge n_1 = \text{const}$

$$\frac{i}{d(n-k)} + \left(\frac{n-k-l}{n-k}\right)^2 \le C_2\log^2(n), \tag{54}$$

with constant C_2.

(3) Factor 3: For the last factor we have the bound

$$\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}$$
$$\le (dk)!(d(n-k)-i-1)^{-(dk+i)/2} \tag{55}$$
$$\le (dk)!(dn-2dk-1)^{-(dk+i)/2}. \tag{56}$$

Putting the bounds for factors 1–3 together, we thus find

$$M_p(n,k,l,d) \leq \sum_{\substack{i=0\\ i=dk \mod 2}}^{dk} d^{(i-dk)/2}(n-k)^i(n-2k-1/d)^{-(dk+i)/2}(dk)!(C_2\log n)^i \tag{57}$$

$$\leq \sum_{\substack{i=0\\ i=dk \mod 2}}^{dk} (n-k)^{i/2}(n-2k-1)^{-dk/2}(dk)!(C_2\log n)^i. \tag{58}$$

The dominant term in the sum is the $i = dk$ term, for which we have

$$\left(\frac{n-k}{n-2k-1}\right)^{dk/2} = \left(\frac{1}{1-\frac{k-1}{n-k}}\right)^{dk/2} \in O(n)^{dk/2}. \tag{59}$$

This gives us

$$M_p(n,k,l,d) \leq dk(dk)!C_3^{dk}n^{-dk/2}\log(n)^{dk} \tag{60}$$

for some constant C_3. Now we can go back to upper-bounding the full expression given by Eq. (46). Since $l \leq (n-k)/2$ by assumption, we also note that the set $S_k = \{l \mid (k,l) \in \mathcal{S}\}$ has size $|S_k| \leq \sqrt{n}\log(n)C_0$. Moreover, we observe that

$$\binom{n}{k}\binom{n-k}{l} \leq n^k 2^{n-k}\frac{\sqrt{8}}{\sqrt{n-k}}, \tag{61}$$

and thus obtain with a constant C_4

$$2^n|T_{\text{case 1}}| \leq 6\sum_{k=1}^{90\log(n)} 2^{n-k}\frac{n^{k+1/2}}{\sqrt{n-k}}dk(dk)!C_4^{dk}\log(n)^{dk+1}n^{-dk/2}$$

$$\leq 6\sum_{k=1}^{90\log(n)} 2^{n-k}\frac{n^{k+1/2}}{\sqrt{n-k}}(C_4)^{dk}\log(n)^{dk+1}(dk)^{dk+1}n^{-dk/2}. \tag{62}$$

Since $k^{d+1}\cdot 90^{d+1}\log(n)^{d+1} \leq n^{d/2}$ for $d > 2$ and sufficiently large n, the terms in the sum are strictly decreasing in k. Hence, the sum is dominated by the $k = 1$ term and we get for any $2 < d = o(n^{1/2})$

$$|T_{\text{case 1}}| \in O\left(d^{d+1}\log(n)^{d+3}n^{-d/2+1}\right) = o(1), \tag{63}$$

which is all we need.

3. Case 2 [$k \geq 90\log n$ and $2^{-n}\binom{n}{k}\binom{n-k}{l} \geq n^8$]

From the orthogonality of the Krawtchouk polynomials we have the bound

$$\left(K_i^{d(n-k)}(dl)\right)^2 \leq 2^{d(n-k)}\binom{d(n-k)}{i}\binom{d(n-k)}{dl}^{-1}. \tag{64}$$

Inserting this into the expression for $M_p(n,k,l,d)$ defined in Eq. (38) and using the triangle inequality, we get

$$M_p(n,k,l,d) = \sum_{\substack{i=0\\ i=dk \mod 2}}^{dk} |K_i^{d(n-k)}(dl)|\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}$$

$$\leq \sum_{\substack{i=0\\ i=dk \mod 2}}^{dk} 2^{d(n-k)/2}\binom{d(n-k)}{i}^{1/2}\binom{d(n-k)}{l}^{-1/2}\frac{(dk)!}{(dk-i)!!}\frac{(d(n-k)-i-1)!!}{(dn-1)!!}. \tag{65}$$

To obtain an upper bound, we first upper-bound the sum over i. To this end, we use the following facts about double factorials and binomials (for n even):

$$(n-1)!! = \frac{n!}{2^{n/2}(n/2)!} \quad \Rightarrow \quad \frac{\sqrt{n!}}{(n-1)!!} = 2^{n/2}\binom{n}{n/2}^{-1/2} \tag{66}$$

$$\frac{2^{nH(p)}}{n+1} \leq \frac{2^{nH(p)}}{\sqrt{8p(1-p)n}} \leq \binom{n}{pn} \leq \frac{2^{nH(p)}}{\sqrt{2\pi p(1-p)n}} \leq 2^{nH(p)}, \tag{67}$$

where $p \in [0,1]$ and $H(p) = -p\log(p) - (1-p)\log(1-p)$ is the binary entropy with base 2. We then obtain

$$\binom{d(n-k)}{i}^{1/2}\frac{(d(n-k)-i-1)!!}{(dk-i)!!} = \left(\frac{(d(n-k))!}{(d(n-k)-i)!i!}\right)^{1/2}\frac{(d(n-k)-i-1)!!}{(dk-i)!!} \tag{68}$$

$$\leq 2^{-(d(n-k)-i)/2}\binom{d(n-k)-i}{(d(n-k)-i)/2}^{1/2}\sqrt{\frac{(d(n-k))!}{i!(dk-i)!}} \tag{69}$$

$$\leq 2^{-(d(n-k)-i)/2}\frac{2^{(d(n-k)-i)/2}}{\sqrt{\pi(d(n-k)-i)}}\sqrt{\frac{(d(n-k))!}{i!(dk-i)!}} \tag{70}$$

$$\leq \sqrt{(d(n-k))!}\frac{1}{\sqrt{i!(dk-i)!}}, \tag{71}$$

where we used $(dk-i)!! \geq \sqrt{(dk-i)!}$ and $h(1/2) = 1$. We observe that this term is maximized at $i = \lfloor dk/2 \rfloor$, since $i!(dk-i)!$ is minimized at $i = \lfloor dk/2 \rfloor$. We thus get

$$\binom{d(n-k)}{i}^{1/2}\frac{(d(n-k)-i-1)!!}{(dk-i)!} \leq \frac{((d(n-k))!)^{1/2}}{\lfloor dk/2 \rfloor!}. \tag{72}$$

Hence, we can upper-bound

$$M_p(n,k,l,d) = \sum_{\substack{i=0 \\ i=dk \mod 2}}^{dk} |K_i^{d(n-k)}(dl)|\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}$$

$$\leq 2^{d(n-k)/2}\binom{d(n-k)}{dl}^{-1/2} dk\frac{(dk)!}{\lfloor dk/2 \rfloor!}\frac{((d(n-k))!)^{1/2}}{(dn-1)!!}. \tag{73}$$

We can insert this back into Eq. (46) and get

$$2^n|T_{\text{case }2}| \leq 6 \sum_{\substack{(k,l)\in\mathcal{S}, \\ 8\log(n)\leq k\leq l\leq n-k-l}} \binom{n}{k}\binom{n-k}{l}\left[\sum_{\substack{i=0 \\ i=dk \mod 2}}^{dk} |K_i^{d(n-k)}(dl)|\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}\right] \tag{74}$$

$$\leq \sum_{\substack{(k,l)\in\mathcal{S}, \\ 8\log(n)\leq k\leq l\leq n-k-l}} 6dk2^{d(n-k)/2}\binom{n}{k}\binom{n-k}{l}\binom{d(n-k)}{dl}^{-\frac{1}{2}}\frac{(dk)!}{\lfloor dk/2 \rfloor!}\frac{((d(n-k))!)^{\frac{1}{2}}}{(dn-1)!!} \tag{75}$$

$$\leq \sum_{\substack{(k,l)\in\mathcal{S},\\ 8\log(n)\leq k\leq l\leq n-k-l}} 6d^{3/2}k\sqrt{n}2^{d(n-k)/2}\binom{n}{k}\binom{n-k}{l}\binom{d(n-k)}{dl}^{-\frac{1}{2}}\binom{dn}{dk}^{-\frac{1}{2}} \tag{76}$$

$$\leq \sum_{\substack{(k,l)\in\mathcal{S},\\ 8\log(n)\leq k\leq l\leq n-k-l}} Cd^{3/2}k\sqrt{n}2^{dn/2}\binom{n}{k}\binom{n-k}{l}\binom{d(n-k)}{dl}^{-\frac{1}{2}}\binom{dn}{dk}^{-\frac{1}{2}} \tag{77}$$

$$\leq \sum_{\substack{(k,l)\in\mathcal{S},\\ 8\log(n)\leq k\leq l\leq n-k-l}} Cd^{3/2}k\sqrt{n}2^{dn/2}\binom{n}{k}\binom{n-k}{l}\binom{(n-k)}{l}^{-\frac{d}{2}}\binom{n}{k}^{-\frac{d}{2}}, \tag{78}$$

where we used $(dn-1)!! \geq \sqrt{(dn-1)!} = (dn)^{-1/2}\sqrt{dn}$ and $C > 6$ is a constant, and $\binom{dn}{dk} \geq \binom{n}{k}^d$ in the final line. By construction, and for $d \geq 3$, we now have

$$2^{-n(1-d/2)}\binom{(n-k)}{l}^{1-\frac{d}{2}}\binom{n}{k}^{1-\frac{d}{2}} \leq n^{-8d/2+8} \leq n^{-4}. \tag{79}$$

This makes the total sum

$$|T_{\text{case 2}}| \in O(n^{-1}) = o(1). \tag{80}$$

4. Case 3 [$k \geq 90\log(n)$ and $n^{-3} \leq 2^{-n}\binom{n}{k}\binom{n-k}{l} \leq n^8$]

This case is subtler and corresponds to the dominant terms of the sum over k, l, which lie in a “ring” of intermediate values; see Fig. 2(right). The first thing we note is that within this region there is a maximal value k^* of k, achieved when $l = k$ and

$$2^{-n}\binom{n}{k^*}\binom{n-k^*}{k^*} = n^8. \tag{81}$$

We can (numerically) invert this equation to show that

$$k^* \leq cn + C_1\log n \leq c^*n \tag{82}$$

for some constant C_1 and $c_* := 0.113 < c < 0.114 =: c^*$ and sufficiently large n. Similarly we can show that there is a minimal l_* given by

$$l_* \geq cn - C_2\log n \geq c_*n \tag{83}$$

for some constant C_2 and sufficiently large n. The minimal $k_* = 90\log(n)$ further implies a maximal $l^* = (n - 90\log(n))/2$.

Now we can use Derksen’s bound:

$$M_p(n,k,l,d) = \sum_{\substack{i=0\\ i=dk \mod 2}}^{dk} |K_i^{d(n-k)}(dl)|\frac{dk!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}$$
$$\leq \sum_{\substack{i=0\\ i=dk \mod 2}}^{dk} \binom{d(n-k)}{i}\left(\frac{i}{d(n-k)}+\frac{(n-k-2l)^2}{(n-k)^2}\right)^{i/2}\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}. \tag{84}$$

This equation is manifestly monotonously decreasing in l, since $l \leq (n-k)/2$ for any pair (k, l). Hence, we can insert l_* to obtain an upper bound on Eq. (84). We further split the sum over i into two cases: (1) $i \geq 130\log(n)$ and (2) $i \leq 130\log(n)$.

For case (1) we can use Hölder's inequality

$$\begin{aligned}&\sum_{\substack{i\geq 130\log n\\ i=dk \mod 2}}^{dk}\binom{d(n-k)}{i}\left(\frac{i}{d(n-k)}+\frac{(n-k-2l)^2}{(n-k)^2}\right)^{i/2}\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}\\ &\leq \max_{\substack{130\log(n)\leq i\leq dk\\ i=dk \mod 2}}\left(\frac{i}{d(n-k)}+\frac{(n-k-2l_*)^2}{(n-k)^2}\right)^{i/2}\sum_{\substack{i\geq 130\log(n)\\ i=dk \mod 2}}^{dk}\binom{d(n-k)}{i}\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}. \end{aligned}\tag{85}$$

We can upper-bound the second factor by 1 by noting that $\binom{d(n-k)}{i}=K_i^{d(n-k)}(0)$ and adding back in the $i\leq 130\log(n)$ terms to the sum. All of these terms are non-negative since the only nonzero term at $x=0$ in the definition [Eq. (14)] of the Krawtchouk polynomials is the $q=0$ term. This allows us to recognize the formula for the average parity over all matchings from Lemma 6 with $l=0$, which trivially evaluates to 1. For the maximization we note that $0\leq i/d\leq k\leq k^*=c^*n$ and $l_*\geq c_*n$, telling us that

$$\begin{aligned}\left(\frac{i}{d(n-k)}+\frac{(n-k-2l_*)^2}{(n-k)^2}\right)^{i/2}&\leq\left(\frac{i}{d(n-k^*)}+\frac{(n-2l_*)^2}{(n-k^*)^2}\right)^{i/2}\\ &\leq\left(\frac{c^*}{1-c^*}+\frac{(1-2c_*)^2}{(1-c^*)^2}\right)^{i/2}\leq 0.892^{i/2}\leq n^{-10.5},\end{aligned}\tag{86}$$

where the last bound follows from the fact that $0.892^{i/2}$ decreases in i, and hence we can use the lower bound $i\geq 130\log n$ to obtain the upper bound.

Finally we deal with the term where $i\leq 130\log(n)$. This subcase closely resembles case 1 ($k\leq 90\log n$). We will use the straightforward bound

$$\binom{d(n-k)}{i}\leq(d(n-k))^i,\tag{87}$$

and furthermore bound

$$\begin{aligned}(d(n-k)-i-1)!!&=(d(n-k)-1)!!\prod_{t=0}^{i/2}(d(n-k)-2t-1)^{-1}\qquad(88)\\ &\leq\frac{(d(n-k)-1)!!}{(d(n-k)-i-1)^{i/2}}\qquad(89)\\ &\leq\frac{\sqrt{(d(n-k))!}}{(d(n-k))^{i/2}}(1+o(1)),\qquad(90)\end{aligned}$$

using $i\leq 130\log n$ and therefore using Taylor's theorem $(d(n-k)-i-1)^{-i/2}\leq(d(n-k))^{-i/2}(1+O(\log^2(n)/(dn)))\leq(d(n-k))^{-i/2}(1+o(1))$. Analogously, we find

$$(dk-i)!!\geq dk!!\prod_{t=1}^{i/2}(dk-2t)^{-1}\geq(dk)!!(dk)^{-i/2}\geq\sqrt{(dk)!}(dk)^{-i/2}\tag{91}$$

Filling these into Eq. (84), we obtain for case (2)

$$\begin{aligned}
&\sum_{\substack{i=0\\ i=dk \mod 2}}^{\lfloor 130\log(n)\rfloor} \binom{d(n-k)}{i}\left(\frac{i}{d(n-k)}+\frac{(n-k-2l_*)^2}{(n-k)^2}\right)^{i/2}\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}\\
&\le \frac{\sqrt{(dk)!(d(n-k))!}}{\sqrt{(dn-1)!}}\sum_{\substack{i=0\\ i=dk \mod 2}}^{\lfloor 130\log(n)\rfloor}\left[\frac{(dk)(d(n-k))^2}{d(n-k)}\left(\frac{i}{d(n-k)}+\frac{(n-k-2l_*)^2}{(n-k)^2}\right)\right]^{i/2}(1+o(1))\\
&\le \frac{(dn)\sqrt{(dk)!(d(n-k))!}}{\sqrt{(dn)!}}\sum_{\substack{i=0\\ i=dk \mod 2}}^{\lfloor 130\log(n)\rfloor}\left[d^2k(n-k)\left(\frac{i}{d(n-k)}+\frac{(n-k-2l_*)^2}{(n-k)^2}\right)\right]^{i/2}(1+o(1))]. \qquad (92)
\end{aligned}$$

For the second factor in the sum we can argue that

$$\left(\frac{i}{d(n-k)}+\frac{(n-k-2l_*)^2}{(n-k)^2}\right)^{i/2}\le\left[o(1)+\left(\frac{1-2c_*}{1-c^*}\right)^2\right]^{i/2}\le 1, \qquad (93)$$

for sufficiently large n using the upper bounds on i and k and the lower bound on l. Furthermore the first factor is clearly bounded by

$$(d^2k(n-k))^{i/2}\le (dn)^i\le (dn)^{130\log(n)}. \qquad (94)$$

Combining all of these, we get

$$\begin{aligned}
&\sum_{\substack{i=0\\ i=dk \mod 2}}^{\lfloor 130\log(n)\rfloor} \binom{d(n-k)}{i}\left(\frac{i}{d(n-k)}+\frac{(n-k-2l_*)^2}{(n-k)^2}\right)^{i/2}\frac{(dk)!(d(n-k)-i-1)!!}{(dk-i)!!(dn-1)!!}\\
&\le \binom{dn}{dk}^{-1/2}(dn)^{130\log(n)+1/2}(1+o(1)). \qquad (95)
\end{aligned}$$

We can further bound this as

$$\begin{aligned}
&\log\left(\binom{dn}{dk}^{-1/2}(dn)^{130\log(n)+1/2}(1+o(1))\right) && (96)\\
&\le -dnH(k_*/n)/2+130(\log^2(n)+\log n\log d)+o(1) && (97)\\
&\le -dk_*\log(n/k_*)/2+130(\log^2(n)+\log n\log d)+o(1) && (98)\\
&\le \log^2(n)(-45d+130+o(1)) && (99)\\
&\le -5\log^2(n)(1-o(1)), && (100)
\end{aligned}$$

where we have used that $1/\binom{dn}{dk}$ is monotonously decreasing in k, so that we can bound the binomial using the lower bound $k_*=90\log(n)$, and the last bound follows from the assumption $2<d=o(n^{1/2})$. Hence, we have

$$|T_{\text{case }3}|\le n^2n^8\cdot O(n^{-10.5}+n^{-\Omega(\log n)})\le O(n^{-1/2})=o(1), \qquad (101)$$

which completes the final case.

Putting the three bounds together, we have our result that for $2<d\in o(n^{1/2})$

$$\mathbb{E}_{G_M\sim\mathcal{G}_p(n,d),\theta}[m_2(G_M,\theta)]=o(1)+\begin{cases}2 & \text{if } d=1 \mod 2,\\ 3 & \text{if } d=0 \mod 2.\end{cases} \qquad (102)$$

Sketch of the proof for the random matching case of Theorem 1. Let us conclude this section by sketching the proof of the random matching part of Theorem 1. The proof proceeds along the same lines as the proof of Theorem 6 but is somewhat simpler. We restrict ourselves to outlining the essential difference and leaving the details of adapting the proof as an exercise for the reader.

The key difference in the proof is the starting point, where instead of averaging over one large matching of dn vertices, we average over d independent matchings of n vertices:

$$\begin{aligned}\mathbb{E}_{G\sim\mathcal{G}_m(n,d),\theta}(m_2(G,\theta) &= 2^{-n}\sum_{L,R\subset[n],L\cap R=\emptyset}\mathbb{E}_{G\sim\mathcal{G}_m(n,d)}(-1)^{|A_G[L,R]|} &(103)\\ &= 2^{-n}\sum_{L,R\subset[n],L\cap R=\emptyset}\mathbb{E}_{M_1,\ldots,M_d\sim\mathcal{M}(n)}(-1)^{|A_{M_1}[L,R]|}\cdots(-1)^{|A_{M_d}[L,R]|}\\ &= 2^{-n}\sum_{L,R\subset[n],L\cap R=\emptyset}\left(\mathbb{E}_{M\sim\mathcal{M}(n)}(-1)^{|A_M[L,R]|}\right)^d, &(104)\end{aligned}$$

and we thus get by Lemmas 5 and 6

$$\mathbb{E}_{G\sim\mathcal{G}_m(n,d),\theta}(m_2(G,\theta) = 2^{-n}\sum_{L,R\subset[n],L\cap R=\emptyset}\left(\sum_{\substack{i=0\\ i=|L| \mod 2}}^{|L|}K_i^{n-|L|}(|R|)\frac{|L|!(n-|L|-i-1)!!}{(|L|-i)!!(n-1)!!}\right)^d. \qquad (105)$$

In the subsequent analysis many aspects of the proof simplify since the d-dependence is isolated to an overall power of the sum over i. ■

IV. UNIVERSALITY OF REGULAR GRAPHS OF INTERMEDIATE DEGREE

In this section, we prove that most regular graphs of sufficiently high regularity are resources for universal quantum computation. We prove this by showing that we can find sufficiently large grid graphs embedded in these graphs as induced subgraphs.

Our proof consists of a series of lemmas. We start with an exposition of the "switching method" of regular graph theory—a counting technique first introduced by McKay and Wormald [42]—as it is applicable in our context. We use the switching method to compute the expected number of grid graphs in a random regular graph. Then we compute the associated variance. Finally, we apply Chebyshev's inequality to make a typicality statement. Our proofs follow the arguments for the estimation of the probability of subgraphs [43], and induced subgraphs [86] *of constant size* of random regular graphs. However, we require estimates for induced subgraphs that grow in size with n, which means we need to be substantially more careful in our estimations.

A. An exposition of forward and reverse switching

In our proofs in Sec. IV, we make heavy use of a counting technique called "switching," which was introduced by McKay and Wormald [42]. This is a standard technique in the regular-graph theory literature (and combinatorics more generally); see, e.g., Ref. [87] for a general discussion. Because we need somewhat precise estimates, and this technique is not very well known in the quantum information literature, we do several key calculations explicitly here (specifically for our problem).

Let $\varepsilon, \varepsilon'$ be two collections of edges on the vertex set $[n]$, with $\varepsilon \cap \varepsilon' = \emptyset$ and $|\varepsilon| = s, |\varepsilon'| = s'$. Let uw be an edge in ε, and $u'w'$ be an edge in ε'. Define two sets:

$$\mathscr{L} = \left\{G_{n,d} \mid \varepsilon \subseteq G_{n,d}; \varepsilon' \subseteq \overline{G}_{n,d}\right\}, \qquad (106)$$

$$\mathscr{M} = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}, uw \notin G_{n,d}; \varepsilon' \subseteq \overline{G}_{n,d}\right\}. \qquad (107)$$

The goal of the switching method is to estimate $|\mathscr{L}|/|\mathscr{M}|$. This is done by setting up a relation R between $\mathscr{L}, \mathscr{M}$, and then estimating how many elements of $\mathscr{M}$ are related to a uniformly random element of $\mathscr{L}$ (call this expectation d_1). We will also estimate how many elements of $\mathscr{L}$ are related to a uniformly random element of $\mathscr{M}$ (call this d_2). By a basic double counting argument one can see that

$$d_1|\mathscr{L}| = |R| = d_2|\mathscr{M}|. \qquad (108)$$

This means we can estimate the ratio by estimating d_1, d_2 (note that this part is not specific to the sets in question).

We now set up the relation (which is specific to these sets). We do this by relating a graph $G' \in \mathscr{M}$ to a graph $G \in \mathscr{L}$ if G' can be reached from G by a *forward switching*. Equivalently, we will see that this means G' can be reached from G by *reverse switching*.

Given a graph $G \in \mathscr{L}$, we will choose two edges u_1w_1 and u_2w_2 of $G\backslash\varepsilon$, delete these edges together with the edge uw, and insert the new edges wu_1, w_1u_2, w_2u. This produces a graph G'. We will choose u_1w_1 and u_2w_2 of $G\backslash\varepsilon$ such that all six endpoints of the edges are distinct and, in addition, wu_1, w_1u_2, w_2u are not edges of G and are not in ε'.

We can provide an estimate for how many ways this forward switching can be performed (and thus estimate d_1).

Lemma 7. Given $\varepsilon, \varepsilon'$ defined as above, there are

$$d^2(n-\mathcal{O}(d))^2 \cdot \left(1 - \frac{\mathcal{O}(s)}{(n-\mathcal{O}(d))\cdot d} - \frac{\mathcal{O}(s')}{n-\mathcal{O}(d)}\right) \tag{109}$$

ways to perform a forward switching from $\mathscr{L}$ to $\mathscr{M}$.

Proof. The number of choices of u_1w_1 is

$$2\left(\frac{nd}{2} - \#\text{forbidden cases}\right), \tag{110}$$

where $nd/2$ is the total number of edges of the graph and we multiply by 2 because u_1w_1 and w_1u_1 represent two separate cases when the new edges are inserted. The forbidden cases are summarized as follows:

(1) There are at most s choices of u_1w_1 such that u_1w_1 is in ϵ.
(2) There are at most $(d-1)$ choices such that $w = u_1$; that is, the endpoints will not be distinct.
(c) There are at most

$$\mathcal{O}\left((d-1)^2 + s'd\right) \tag{111}$$

choices of u_1w_1 such that wu_1 and w_1u_2 are either in G or in ϵ'.

Hence,

$$\begin{aligned}\#\text{forbidden cases} &= \mathcal{O}\left(s + (d-1) + (d-1)^2 + s'd\right)\\ &= \mathcal{O}\left(s + d^2 + s'd\right).\end{aligned} \tag{112}$$

Putting everything together, we find the number of choices of u_1w_1 is

$$nd - \mathcal{O}\left(s + d^2 + s'd\right). \tag{113}$$

By a similar argument, the number of choices of u_2w_2 is also given by Eq. (113). Hence, the total number of choices of u_1w_1 and u_2w_2 is

$$\begin{aligned}&\left(d(n-\mathcal{O}(d)) - \mathcal{O}(s+s'd)\right)^2\\ &\quad= d^2(n-\mathcal{O}(d))^2 \cdot \left(1 - \frac{\mathcal{O}(s+s'd)}{(n-\mathcal{O}(d))\cdot d}\right.\\ &\qquad\left. + \frac{\mathcal{O}(s+s'd)^2)}{(n-\mathcal{O}(d))^2\cdot d^2}\right)\\ &\quad= d^2(n-\mathcal{O}(d))^2 \cdot \left(1 - \frac{\mathcal{O}(s)}{(n-\mathcal{O}(d))\cdot d} - \frac{\mathcal{O}(s')}{n-\mathcal{O}(d)}\right).\end{aligned} \tag{114}$$

■

We still need to estimate d_2. To do this we define an inverse operation called "reverse switching," mapping graphs in $\mathscr{M}$ to graphs in $\mathscr{L}$.

Starting from a graph $G' \in \mathscr{M}$ we delete edges wu_1, w_1u_2, w_2u of $G'\backslash\varepsilon$ and insert edges uw, u_1w_1, u_2w_2. Again, we allow only switchings for which all six vertices are distinct, and u_1w_1, u_2w_2 are not edges of G' and are not in ε'. This operation produces a graph which belongs to $\mathscr{L}$. Note also that the existence of a reverse switching from G' to G implies the existence of a forward switching from G to G'. Hence, counting the number of reverse switchings for a random $G' \in \mathscr{M}$ is equivalent to calculating d_2.

We can again provide an estimate of the number of ways in which this reverse switching can be performed.

Lemma 8. Given $\varepsilon, \varepsilon'$ defined as above, there are

$$d^3(n-\mathcal{O}(d))\left(1 - \mathcal{O}\left(\frac{s}{d}\right) - \frac{\mathcal{O}(s')}{n-\mathcal{O}(d)}\right) \tag{115}$$

ways to perform a reverse switching from $\mathscr{L}_3$ to $\mathscr{L}_1$.

Proof. Note that u and w are vertices in G'. Hence, the number of choices of wu_1 and w_2u such that they are in $G'\backslash\varepsilon$ is

$$(d-1-\mathcal{O}(s))^2 = (d-\mathcal{O}(s))^2. \tag{116}$$

The number of choices of w_1u_2 is

$$2\left(\frac{nd}{2} - \#\text{forbidden cases}\right). \tag{117}$$

The forbidden cases are summarized as follows:

(1) There are at most s choices of w_1u_2 such that u_1w_1 is in ϵ.
(2) Our having already picked w, u_1, w_2, and u, there are $\mathcal{O}(d-1)$ choices such that the endpoints of w_1u_2 will not be distinct.

(3) There are at most

$$\mathcal{O}((d-1)^2 + s'd) \tag{118}$$

choices of w_1u_2 such that u_1w_1 and u_2w_2 are either in G' or in ϵ'.

Hence,

$$\begin{aligned}\#\text{forbidden cases} &= \mathcal{O}\left(s + (d-1) + (d-1)^2 + s'd\right)\\ &= \mathcal{O}\left(s + d^2 + s'd\right).\end{aligned} \tag{119}$$

Taken together, the total number of choices of wu_1, w_1u_2, and w_2u is

$$\begin{aligned}&(d - \mathcal{O}(s))^2\left(d(n - \mathcal{O}(d)) - \mathcal{O}(s + s'd)\right)\\ &\quad= (d - \mathcal{O}(s))^2\left(d(n - \mathcal{O}(d)) - \mathcal{O}(s + s'd)\right)\\ &\quad= d^3(n - \mathcal{O}(d))\left(1 - \frac{\mathcal{O}(s - 2ss')}{d\cdot(n - \mathcal{O}(d))} - \frac{\mathcal{O}(s')}{n - \mathcal{O}(d)}\right.\\ &\qquad - \mathcal{O}\left(\frac{s}{d}\right) - \frac{\mathcal{O}(s's^2 - s^2)}{d^2\cdot(n - \mathcal{O}(d))} + \mathcal{O}\left(\frac{s^2}{d^2}\right)\\ &\qquad\left. - \frac{\mathcal{O}(s^3)}{d^3\cdot(n - \mathcal{O}(d))}\right)\\ &\quad= d^3(n - \mathcal{O}(d))\left(1 - \mathcal{O}\left(\frac{s}{d}\right) - \frac{\mathcal{O}(s')}{n - \mathcal{O}(d)}\right).\end{aligned} \tag{120}$$

■

From the description of the processes, the following corollary is immediate.

Corollary 5. G is reachable from G' by forward switching if and only if G' is reachable from G by reverse switching.

We can also estimate the relative sizes of two other sets given by

$$\mathscr{L} = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}; \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}, u'w' \in G_{n,d}\right\}, \tag{121}$$

$$\mathscr{M} = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}; \varepsilon' \subseteq \overline{G}_{n,d}\right\} \tag{122}$$

using forward and reverse switching. Just as before, given a graph $G \in \mathscr{L}$, we will choose two edges u_1w_1 and u_2w_2 of $G\backslash\varepsilon$ such that all six endpoints u', w', u_1, w_1, u_2, and w_2 are distinct and, in addition, $w'u_1, w_1u_2, w_2u'$ are not edges of G and are not in ε'. Then we will forward-switch to get a graph G' in $\mathscr{M}$. Similarly, we can reverse-switch to go from a graph in $\mathscr{M}$ to a graph in $\mathscr{L}$.

B. Expected number of induced subgraphs

We plan to use the second moment method to prove Theorem 2. To do this we first calculate the expected number of graphs isomorphic to a fixed graph H of a given size that show up as induced subgraphs of a randomly sampled d-regular graph. In the next subsection we then specialize to grid graphs. Throughout, this graph will be represented by a set of edges ε and a set of nonedges ε'. This latter part is required because we want to find *induced* subgraphs.

To compute the expected number of induced subgraphs we first need to prove a few subsidiary lemmas. The first allows us to "peel off" edges from the sets $\varepsilon, \varepsilon'$, without changing the probability of observing these sets too much. This lemma explicitly relies on the switching technique:

Lemma 9. Let $d = n^c$ for any choice of constant $0 < c < 1$, and let $\varepsilon, \varepsilon'$ be two collections of edges on the vertex set $[n]$, with $\varepsilon \cap \varepsilon' = \emptyset$ and $|\varepsilon| = s, |\varepsilon'| = s'$. Let uw be an edge in ε, and let $u'w'$ be an edge in ε'. Then we have

$$\begin{aligned}&\mathbb{P}\left[\varepsilon \subseteq G_{n,d}, \varepsilon' \subseteq \overline{G}_{n,d}\right]\\ &\quad= \left(\frac{n\alpha}{n - d(1-\alpha)}\right)\left(\frac{d}{n}\right)\left(1 - \frac{d}{n}\right)\\ &\qquad\times \mathbb{P}\left[\varepsilon\backslash\{uw\} \subseteq G_{n,d}, \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}\right],\end{aligned} \tag{123}$$

where

$$\alpha = \frac{1 - \mathcal{O}\left(\frac{s}{d}\right) - \frac{\mathcal{O}(s')}{n - \mathcal{O}(d)}}{1 - \frac{\mathcal{O}(s)}{(n - \mathcal{O}(d))\cdot d} - \frac{\mathcal{O}(s')}{n - \mathcal{O}(d)}}. \tag{124}$$

Proof. Note that $uw \in \epsilon$ and $u'w' \in \epsilon'$. Let

$$\mathscr{L}_1 = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}; \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}; uw \in G_{n,d}; u'w' \in \overline{G}_{n,d}\right\}, \tag{125}$$

$$\mathscr{L}_2 = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}; \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}; uw \in G_{n,d}; u'w' \notin \overline{G}_{n,d}\right\}, \tag{126}$$

$$\mathscr{L}_3 = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}; \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}; uw \notin G_{n,d}; u'w' \in \overline{G}_{n,d}\right\}, \tag{127}$$

$$\mathscr{L}_4 = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}; \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}; uw \notin G_{n,d}; u'w' \notin \overline{G}_{n,d}\right\}. \tag{128}$$

By suppressing the redundancies in notation, we can rewrite the same equations more succinctly as

$$\mathscr{L}_1 = \left\{G_{n,d} \mid \varepsilon \subseteq G_{n,d}; \varepsilon' \subseteq \overline{G}_{n,d}\right\}, \tag{129}$$

$$\mathscr{L}_2 = \left\{G_{n,d} \mid \varepsilon \subseteq G_{n,d}; \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}, u'w' \in G_{n,d}\right\}, \tag{130}$$

$$\mathscr{L}_3 = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}, uw \notin G_{n,d}; \varepsilon' \subseteq \overline{G}_{n,d}\right\}, \tag{131}$$

$$\mathscr{L}_4 = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}, uw \notin G_{n,d}; \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}, u'w' \in G_{n,d}\right\}. \tag{132}$$

By definition, we have

$$\frac{\mathbb{P}\left[\varepsilon \subseteq G_{n,d}, \varepsilon' \nsubseteq G_{n,d}\right]}{\mathbb{P}\left[\varepsilon\backslash\{uw\} \subseteq G_{n,d}, \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}\right]} = \frac{|\mathscr{L}_1|}{|\mathscr{L}_1| + |\mathscr{L}_2| + |\mathscr{L}_3| + |\mathscr{L}_4|} \tag{133}$$

$$= \frac{\frac{|\mathscr{L}_1|}{|\mathscr{L}_1|+|\mathscr{L}_3|}}{1 + \frac{|\mathscr{L}_2|+|\mathscr{L}_4|}{|\mathscr{L}_1|+|\mathscr{L}_3|}}. \tag{134}$$

To prove the lemma, it is enough to show that

$$\frac{|\mathscr{L}_1|}{|\mathscr{L}_3|} = \frac{\alpha \cdot d}{n-d} \quad \text{and} \quad \frac{|\mathscr{L}_2| + |\mathscr{L}_4|}{|\mathscr{L}_1| + |\mathscr{L}_3|} = \frac{\alpha \cdot d}{n-d}. \tag{135}$$

We first prove that $|\mathscr{L}_1|/|\mathscr{L}_3| = \alpha d/(n-d)$. Given a graph $G \in \mathscr{L}_1$, we will perform forward switching to get a graph in $\mathscr{L}_3$. We choose two edges u_1w_1 and u_2w_2 of $G\backslash\varepsilon$, delete these edges together with the edge uw, and insert the new edges wu_1, w_1u_2, w_2u. We will choose u_1w_1 and u_2w_2 of $G\backslash\varepsilon$ such that all six endpoints of the edges are distinct and, in addition, wu_1, w_1u_2, w_2u are not edges of G and are not in ε'. Then it is easy to see that the graph obtained from G by a forward switching belongs to $\mathscr{L}_3$. By Eq. (114), the number of possible forward switchings is

$$d(n - \mathcal{O}(d)) \cdot \left(1 - \frac{\mathcal{O}(s)}{(n-\mathcal{O}(d)) \cdot d} - \frac{\mathcal{O}(s')}{n - \mathcal{O}(d)}\right). \tag{136}$$

Similarly, by applying reverse switching, we can go from $\mathscr{L}_3$ to $\mathscr{L}_1$. By Eq. (120), the number of possible reverse switchings is

$$d^3(n - \mathcal{O}(d))\left(1 - \mathcal{O}\left(\frac{s}{d}\right) - \frac{\mathcal{O}(s')}{n-\mathcal{O}(d)}\right). \tag{137}$$

Since $\mathscr{L}_2 \cap \mathscr{L}_4 = \emptyset$ and $\mathscr{L}_1 \cap \mathscr{L}_3 = \emptyset$, the ratio of $|\mathscr{L}_2| + |\mathscr{L}_4|$ and $|\mathscr{L}_1| + |\mathscr{L}_3|$ is

$$\frac{|\mathscr{L}_2| + |\mathscr{L}_4|}{|\mathscr{L}_1| + |\mathscr{L}_3|} = \frac{|\mathscr{L}_2 \cup \mathscr{L}_4|}{|\mathscr{L}_1 \cup \mathscr{L}_3|}. \tag{138}$$

Moreover,

$$\mathscr{L}_2 \cup \mathscr{L}_4 = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}; \varepsilon'\backslash\{u'w'\} \subseteq \overline{G}_{n,d}, u'w' \notin \overline{G}_{n,d}\right\}, \tag{139}$$

$$\mathscr{L}_1 \cup \mathscr{L}_3 = \left\{G_{n,d} \mid \varepsilon\backslash\{uw\} \subseteq G_{n,d}; \varepsilon' \subseteq \overline{G}_{n,d}\right\}. \tag{140}$$

Once again, by the forward and reverse switching arguments as in Sec. IV A 1, we can show that

$$\frac{|\mathscr{L}_2 \cup \mathscr{L}_4|}{|\mathscr{L}_1 \cup \mathscr{L}_3|} = \frac{\alpha d}{n-d}. \tag{141}$$

Hence, the lemma follows. ■

Next we need a useful intermediate result regarding the probability of finding a subgraph (not induced) in a random regular graph. The proof is a modification of that in Ref. [43]. We will use this result in the proof of our main result.

Lemma 10. Let $d = n^c$ for any choice of $0 < c < 1$. Let ε be a fixed disjoint collection of edges on the vertex set $[n]$ of size s and let uw be an edge in ϵ. Then

$$\mathbb{P}[\varepsilon \subseteq G_{n,d}] = \left(\frac{n\alpha}{n - d(1-\alpha)}\right)\left(\frac{d}{n}\right)\mathbb{P}[\varepsilon\backslash\{uw\} \subseteq G_{n,d}], \tag{142}$$

where

$$\alpha = \frac{1 - \mathcal{O}\left(\frac{s}{d}\right)}{1 - \frac{\mathcal{O}(s)}{(n-\mathcal{O}(d))\cdot d}}. \tag{143}$$

Proof. Let us define two sets

$$\mathscr{L}_0 = \{\varepsilon \subseteq G_{n,d}\}, \tag{144}$$

$$\mathscr{L}_1 = \{\varepsilon\backslash\{uw\} \subseteq G_{n,d} : \{uw\} \notin G_{n,d}\}. \tag{145}$$

Given $G \in \mathscr{L}_0$, we apply a forward switching operation to get a graph in $\mathscr{L}_1$, and given a graph $G \in \mathscr{L}_1$, we apply reverse switching to get a graph in $\mathscr{L}_0$. By similar arguments as in Sec. IV A,

$$\frac{|\mathscr{L}_0|}{|\mathscr{L}_1|} = \frac{\alpha \cdot d}{n-d}. \tag{146}$$

Now

$$\begin{aligned}\frac{\mathbb{P}\left[\varepsilon \subseteq G_{n,d}\right]}{\mathbb{P}\left[\varepsilon \backslash \{uw\} \subseteq G_{n,d}\right]} &= \frac{|\mathscr{L}_0|}{|\mathscr{L}_0| + |\mathscr{L}_1|} = \frac{\frac{|\mathscr{L}_0|}{|\mathscr{L}_1|}}{1 + \frac{|\mathscr{L}_0|}{|\mathscr{L}_1|}} \\ &= \frac{d}{n}\left(\frac{n\alpha}{n - d(1-\alpha)}\right).\end{aligned} \tag{147}$$

From Eq. (147), the lemma follows. ■

Corollary 6. Let $d = n^c$ for any choice of $0 < c < 1$. Let uw be an edge. Then

$$\mathbb{P}\left[uw \notin G_{n,d}\right] = \left(1 + \frac{o(d)}{n}\right)\left(1 - \frac{d}{n}\right). \tag{148}$$

Proof. The proof follows from Lemma 10 by setting $s = 1$. ■

With Lemmas 9 and 10 we can estimate the probability that we observe an induced subgraph on a fixed set of vertices.

Lemma 11. Let $d = n^c$ for any choice of $0 < c < 1$. Let $\varepsilon, \varepsilon'$ be fixed disjoint collections of edges on the vertex set $[n]$, of size s, s' respectively. If we assume that $s \cdot s' = o(d)$, then

$$\mathbb{P}\left[\varepsilon \subseteq G_{n,d}, \varepsilon' \subseteq \overline{G}_{n,d}\right] = (1 - o(1))\left(\frac{d}{n}\right)^{s}\left(1 - \frac{d}{n}\right)^{s'}. \tag{149}$$

Proof. Without loss of generality, let $s < s'$. By recursively applying Lemma 9, we have

$$\begin{aligned}&\mathbb{P}\left[\varepsilon \subseteq G_{n,d}, \varepsilon' \subseteq \overline{G}_{n,d}\right] \\ &\quad = \left(\frac{n\alpha}{n - d(1-\alpha)}\right)^{s}\left(\frac{d}{n}\right)^{s}\left(1 - \frac{d}{n}\right)^{s} \\ &\qquad \times \mathbb{P}\left[\varepsilon' \backslash \{u'_1 w'_1, u'_2 w'_2, \ldots, u'_s w'_s\} \subseteq \overline{G}_{n,d}\right]\end{aligned} \tag{150}$$

for choices of edges

$$\{u'_1 w'_1, u'_2 w'_2, \ldots, u'_s w'_s\} \in \epsilon' \tag{151}$$

that are picked at each round, where

$$\alpha = \frac{1 - \mathcal{O}\left(\frac{s}{d}\right) - \frac{\mathcal{O}(s')}{n - \mathcal{O}(d)}}{1 - \frac{\mathcal{O}(s)}{(n - \mathcal{O}(d)) \cdot d} - \frac{\mathcal{O}(s')}{n - \mathcal{O}(d)}}. \tag{152}$$

When $s \cdot s' = o(d)$,

$$\left(\frac{n\alpha}{n - d(1-\alpha)}\right)^{s} = (1 - o(1)). \tag{153}$$

Let the remaining edges in ϵ' be

$$\{u'_{s+1} w'_{s+1}, \ldots u'_{s'} w'_{s'}\}. \tag{154}$$

So, by Corollary 6,

$$\begin{aligned}&\mathbb{P}\left[\{u'_{s+1} w'_{s+1}, \ldots, u'_{s'} w'_{s'}\} \in \overline{G}_{n,d}\right] \\ &\quad = \left(1 + \frac{o(d \cdot (s' - s))}{n}\right)\left(1 - \frac{d}{n}\right)^{s'-s}.\end{aligned} \tag{155}$$

When $ss' = o(d)$,

$$\left(1 + \frac{o(d \cdot (s' - s))}{n}\right) = 1 + o(1). \tag{156}$$

From this observation, the lemma follows. ■

With these lemmas, we can finally prove the main result of this subsection, namely, an estimate of the expected number of induced subgraphs isomorphic to a fixed graph H (the size of which possibly grows with n).

Corollary 7. Let $d = n^c$ for any choice of $0 < c < 1$. Let H be a fixed graph with e edges and v vertices with $e\left(\binom{v}{2} - e\right) = o(d)$. Let Y_H denote the number of induced copies of H in $G_{n,d}$ and let $\mathrm{aut}(H)$ be the number of automorphisms of H. Then

$$\begin{aligned}\mathbb{E}(Y_H) &= (1 - o(1))\frac{\binom{n}{v} v!}{\mathrm{aut}(H)}\left(\frac{d}{n}\right)^{e}\left(1 - \frac{d}{n}\right)^{\binom{v}{2} - e} \\ &= \Theta\left(n^{v-e} d^{e}\left(1 - \frac{d}{n}\right)^{\binom{v}{2} - e}\right).\end{aligned} \tag{157}$$

Proof. For each copy H' of H, define the indicator random variable $J_{H'}$ such that $J_{H'} = 1$ if and only if H' is an induced subgraph of $G_{n,d}$. By Lemma 11,

$$\mathbb{E}(J_{H'}) = (1 - o(1))\left(\frac{d}{n}\right)^{e}\left(1 - \frac{d}{n}\right)^{\binom{v}{2} - e}. \tag{158}$$

Moreover, there are exactly $\binom{n}{v} v!/\mathrm{aut}(H)$ copies of H, and by the linearity of expectation, Eq. (157) holds. ■

C. Expected number of grid graphs

Here we specialize the results of the previous subsection to grid graphs. For a square grid on v vertices (where we assume v to be a square), the number of edges

$$e = 2v - 2\sqrt{v}. \tag{159}$$

The number of edges in the complement of the grid graph is given by

$$\bar{e} = \frac{v^2 - 5v}{2} + 2\sqrt{v}. \tag{160}$$

Hence, the condition

$$\left(\binom{v}{2} - e\right) e = o(d) \tag{161}$$

of Corollary 7 implies $v = o(d^{1/3})$.

Corollary 8. Let $d = n^c$ for any choice of $0 < c < 1$. Let $G_{n,d}$ be a random d-regular graph on n vertices and let H be a grid graph having v vertices for any v satisfying $v^2 \cdot d = \mathcal{O}(n)$ and $v = o(d^{1/3})$. Let Y_H denote the number of induced copies of H in $G_{n,d}$. Then

$$\mathbb{E}(Y_H) = \Theta\left(\left(\frac{d^2}{n}\right)^v\right). \tag{162}$$

Proof. By applying Corollary 7, we have

$$\mathbb{E}(Y_H) = \Theta\left(n^{v-e} d^e \left(1 - \frac{d}{n}\right)^{\binom{v}{2}-e}\right). \tag{163}$$

The proof follows from combining the following observations: (1) $v - e = -v + 2\sqrt{v} = -\Theta(v)$, (2) $n^{v-e} = n^{-\Theta(v)}$, (3) $d^e = d^{\Theta(2v)}$, and (4) $(1 - (d/n))^{\binom{v}{2}-e} = e^{-\Theta(v^2 d/n)}$. ■

Observe that $\mathbb{E}(Y_H)$ in Eq. (162) is greater than 1 when $d = \omega(\sqrt{n})$.

D. Expected number of sparsified square grid graphs

In the previous section we noted that grid graphs are expected to appear whenever $d = \omega(\sqrt{n})$. We can push this lower bound on the degree further down by considering instead *sparsified* grid graphs. These are constructed from regular $L \times L$ grid graphs by replacing each edge with $L - 1$ vertices and L edges connected in a line. These graphs, on $2L(L-1)^2 + L^2$ vertices, are still universal resource states, as measuring all but one of the qubits on each line in the Y basis, which is equivalent to applying the local complementation operation on that vertex, gives back the $L \times L$ grid graph. If we choose the number of vertices $v = 2L(L-1)^2 + L^2$, the number of edges in a sparsified grid graph is given by $e = 2L^2(L-1)$. The number of edges in the complement is given by

$$\bar{e} = \binom{2L(L-1)^2 + L^2}{2} - 2L^2(L-1). \tag{164}$$

With this information we can reprove Corollary 8 for sparsified grid graphs:

Corollary 9. Let $d = n^c$ for any choice of $0 < c < 1$. Let $G_{n,d}$ be a random d-regular graph on n vertices and let H be a sparsified grid graph having v vertices for any v satisfying $v^2 \cdot d = \mathcal{O}(n)$ and $v = o(d^{1/3})$. Let Y_H denote the number of induced copies of H in $G_{n,d}$. Then

$$\mathbb{E}(Y_H) = \Theta\left(\left(\frac{d^{1+1/n}}{n}\right)^v\right). \tag{165}$$

Hence, for any $c > 0$ such that $d = \Theta(n^c)$ the expectation value of $\mathbb{E}(Y_H)$ is asymptotically growing.

E. Upper bound on the variance for grid graphs

We currently have a good estimate of the expected number of induced subgraphs isomorphic to a square grid. However, we would like to show that a large fraction of graphs includes at least one such induced graph. We can do this with the second moment method (which is really just Chebychev's inequality).

For each copy H' of H in K_n, we define the indicator random variable $J_{H'} = 1$ if and only if H' is an induced subgraph of $G_{n,d}$. Then

$$Y_H = \sum_{H'} J_{H'} \tag{166}$$

and

$$\begin{aligned}\mathrm{Var}(Y_H) &= \sum_{H',H''} \mathrm{Cov}(J_{H'}, J_{H''}) \\ &= \sum_{H',H''} \left(\mathbb{E}(J_{H'} J_{H''}) - \mathbb{E}(J_{H'})\mathbb{E}(J_{H''})\right). \end{aligned} \tag{167}$$

We can now calculate the variance:

Lemma 12. Let $d = n^c$ for any choice of $0 < c < 1$. Let $G_{n,d}$ be a random d-regular graph on n vertices and let H be a grid graph having v vertices for any v satisfying $v^2 \cdot d = \mathcal{O}(n)$, $v = o(d^{1/3})$, and $d = \omega(n^{0.5})$. Let Y_H denote the number of induced copies of H in $G_{n,d}$. Then

$$\mathrm{Var}(Y_H) = o\left(\mathbb{E}(Y_H)^2\right). \tag{168}$$

After proving Lemma 12, we extend it to the case of the sparsified grid, as defined in Sec. IV D. This immediately leads to the main theorem:

Theorem 7 (Restatement of Theorem 2). Let G be a random d-regular graph on n vertices, with $d = n^c$, where $0.5 < c < 1$. Then, with probability $1 - o(1)$, it contains a square grid graph on v vertices, for any $v = o(n^k)$, with $k = \min\{(1-c)/2, c/3\}$, as an induced subgraph.

Proof. Let Y_H be the number of copies of H in $G_{n,d}$. Observe that $v = o(n^k)$, with $k = \min\{(1-c)/2, c/3\}$, satisfies $v^2 \cdot d = \mathcal{O}(n)$ and $v = o(d^{1/3})$.

From the calculations in Corollary 8 and Lemma 12, by applying Chebyshev's inequality, we have

$$\mathbb{P}\left(Y_H = 0\right) \leq \frac{\mathrm{Var}(Y_H)}{\mathbb{E}(Y_H)^2} = o(1). \tag{169}$$

Hence, with probability $1 - o(1)$, Y_H is nonzero, where the probability is taken over the randomness in the choice of $G_{n,d}$. This completes the proof. ■

Now we prove Lemma 12.

Proof of Lemma 12. We estimate the variance of Y_H by dividing it into three cases, depending on what the overlap of H' and H'' looks like:

(1) *One common vertex.* Let H' and H'' have at most one vertex in common. Then, by Lemma 11,

$$\begin{aligned}\mathbb{E}(J_{H'}J_{H''}) &= (1-o(1))\left(\frac{d}{n}\right)^{2e}\left(1-\frac{d}{n}\right)^{2\binom{v}{2}-2e}\\ &= (1-o(1))\,\mathbb{E}(J_{H'})\mathbb{E}(J_{H''}).\end{aligned} \tag{170}$$

Now we calculate

$$\begin{aligned}&\sum_{|V(H')\cap V(H'')|\leq 1}\left(\mathbb{E}(J_{H'}J_{H''}) - \mathbb{E}(J_{H'})\mathbb{E}(J_{H''})\right)\\ &\leq \sum_{|V(H')\cap V(H'')|\leq 1} o(\mathbb{E}(J_{H'})\mathbb{E}(J_{H''}))\\ &\leq \sum_{|V(H')\cap V(H'')|\leq 1} o\left(\left(\frac{d}{n}\right)^{2e}\left(1-\frac{d}{n}\right)^{2\binom{v}{2}-2e}\right)\\ &= o\left(n^{2v}\left(\frac{d}{n}\right)^{2e}\left(1-\frac{d}{n}\right)^{2\binom{v}{2}-2e}\right)\\ &= o\left(\mathbb{E}(Y_H)^2\right).\end{aligned} \tag{171}$$

The second line follows from Eq. (170). The third line follows from Lemma 11. The fourth line follows from the fact that there are

$$\begin{aligned}&\left(\binom{v}{1}\binom{n-1}{v-1}\cdot\binom{n-v}{v-1} + \binom{n}{v}\cdot\binom{n-v}{v}\right)\frac{(v!)^2}{(\mathrm{aut}(H))^2}\\ &= o\left(n^{2v}\right)\end{aligned} \tag{172}$$

choices of H' and H'' such that $|V(H')\cap V(H'')| \leq 1$.

(2) *Shared induced subgraph.* For the second case, let the intersection of H' and H'' be a nonempty graph F with v_F vertices and e_F edges. For a fixed F, there are at most

$$\binom{v}{v_F}\cdot\binom{n-v_F}{v-v_F}\cdot\binom{n-v}{v-v_F}\cdot(v!)^2 = o\left(n^{2v-v_F}\right) \tag{173}$$

choices of H' and H''. Hence,

$$\begin{aligned}\sum_{E(H')\cap E(H'')\neq\emptyset}\left(\mathbb{E}(J_{H'}J_{H''}) - \mathbb{E}(J_{H'})\mathbb{E}(J_{H''})\right) &\leq \sum_{E(H')\cap E(H'')\neq\emptyset}\mathbb{E}(J_{H'}J_{H''})\\ &\leq \sum_{F\subseteq H, e_F>0} o\left(n^{2v-v_F}\right)\mathbb{E}(J_{H'}J_{H''})\\ &= o\left(\sum_{F\subseteq H, e_F>0} n^{2v-v_F}\left(\frac{d}{n}\right)^{2e-e_F}\left(1-\frac{d}{n}\right)^{2\binom{v}{2}-2e+e_F-\binom{v_F}{2}}\right)\\ &= o\left(n^{2v}\left(\frac{d}{n}\right)^{2e}\left(1-\frac{d}{n}\right)^{\binom{v}{2}-2e}\sum_{F\subseteq H, e_F>0}\frac{n^{e_F-v_F}}{d^{e_F}}\left(1-\frac{d}{n}\right)^{e_F-\binom{v_F}{2}}\right)\\ &= o\left(\mathbb{E}(Y_H)^2\right)\end{aligned} \tag{174}$$

when

$$\sum_{F\subseteq H, e_F>0} \frac{n^{e_F-v_F}}{d^{e_F}}\left(1-\frac{d}{n}\right)^{e_F-\binom{v_F}{2}} = o(1). \tag{175}$$

The fourth line follows from Eq. (157). Note that the number of subgraphs of F is given by

$$2^{e_F} = n^{e_F/\log n}. \tag{176}$$

Now

$$\sum_{F\subseteq H, e_F>0} \frac{n^{e_F-v_F}}{d^{e_F}}\left(1-\frac{d}{n}\right)^{e_F-\binom{v_F}{2}}$$
$$= n^{e_F/\log n}\cdot\frac{n^{e_F-v_F}}{d^{e_F}}\left(1-\frac{d}{n}\right)^{e_F-\binom{v_F}{2}} = o(1) \tag{177}$$

whenever

$$\frac{n^{e_F-v_F}}{d^{e_F}} = o\left(\frac{1}{n^{e_F/\log n}}\right). \tag{178}$$

This estimation holds whenever

$$d > n^{1+1/\log n - v_F/e_F}. \tag{179}$$

By the definition of the density $m(H)$, one sufficient condition for Eq. (179) to hold is

$$d > n^{1+1/\log n - 1/m(H)}. \tag{180}$$

Note that the other term

$$\left(1-\frac{d}{n}\right)^{e_F-\binom{v_F}{2}} = o(1). \tag{181}$$

If we choose H to be a grid graph, it is easy to see that $m(H)$ is at most 2. Therefore,

$$d = \omega(n^{0.5}) \tag{182}$$

makes the overall term negligible.

(3) *Empty shared induced subgraph.* Now we consider the case when H' and H'' have t vertices in common but no edges in common. There are

$$\binom{v}{t}\cdot\binom{n-t}{v-t}\cdot\binom{n-v}{v-t}\cdot\left(\frac{(v-t)!}{\text{aut}(H)}\right)^2 = o(n^{2v-t}). \tag{183}$$

such cases. The second inequality of Eq. (170) follows from the following estimations:

$$\binom{n}{t} = o(n^t),\quad \binom{n-t}{v-t} = o(n^{v-t}),\quad \binom{n-v}{v-t} = o(n^{v-t}). \tag{184}$$

Again by Lemma 11, we have

$$\mathbb{E}(J_{H'}J_{H''}) = (1+o(1))\left(\frac{d}{n}\right)^{2e}\left(1-\frac{d}{n}\right)^{2\binom{v}{2}-2e-\binom{t}{2}}. \tag{185}$$

In Eq. (185), we subtract $\binom{t}{2}$ edges from the exponent of the term on the extreme right because we are counting them twice—once in the complement set of H' and again in the complement set of H''. Hence,

$$\sum_{\substack{E(H')\cap E(H'')=\emptyset\\ |V(H')\cap V(H'')|\le t}} \left(\mathbb{E}(J_{H'}J_{H''}) - \mathbb{E}(J_{H'})\mathbb{E}(J_{H''})\right)$$
$$\le \sum_{\substack{E(H')\cap E(H'')=\emptyset\\ |V(H')\cap V(H'')|\le t}} \mathbb{E}(J_{H'}J_{H''})$$
$$= o\left(\sum_{t\ge 2} n^{2v-t}\left(\frac{d}{n}\right)^{2e}\left(1-\frac{d}{n}\right)^{2\binom{v}{2}-2e-\binom{t}{2}}\right)$$
$$= o\left(n^{2v}\left(\frac{d}{n}\right)^{2e}\left(1-\frac{d}{n}\right)^{2\binom{v}{2}-2e}\sum_{t\ge 2} n^{-t}\left(1-\frac{d}{n}\right)^{-\binom{t}{2}}\right)$$
$$= o\left(\mathbb{E}(Y_H)^2\right). \tag{186}$$

The last line follows from the fact that

$$\sum_{t\ge 2} n^{-t}\left(1-\frac{d}{n}\right)^{-\binom{t}{2}} = o(1). \tag{187}$$

Therefore, putting everything together, we have

$$\text{Var}(Y_H) = o\left(\mathbb{E}(Y_H)^2\right), \tag{188}$$

which is the lemma statement ■

Corollary 10. Let $G_{n,d}$ be a random d-regular graph on n vertices and let H be a sparsified square grid graph having v vertices for any v satisfying $v^2\cdot d = \mathcal{O}(n)$, $v = o(d^{1/3})$, and $d = n^c$ for any constant $0 < c < 1$. Let Y_H denote the number of induced copies of H in $G_{n,d}$. Then

$$\text{Var}(Y_H) = o\left(\mathbb{E}(Y_H)^2\right). \tag{189}$$

Proof. The proof is the same as that of Lemma 12 with the observation that in the second case, if H is the sparsified square grid graph, the local density is bounded as

$$m(H) \leq \frac{\sqrt{v}}{\sqrt{v}-1}. \tag{190}$$

Hence,

$$n^{1+1/\log n - 1/m(H)} \leq n^{1/m + 1/\sqrt{v}} < n^c = d \tag{191}$$

for any constant $0 < c < 1$. ■

V. GEOMETRIC ENTANGLEMENT OF HIGH-DEGREE GRAPH STATES

In this section we investigate the computational complexity of random regular graph states of high degree, i.e., $d = cn$ with $c \in (0, 1)$. We will use *uniformly* random graphs as a proxy for graphs of high degree. We suspect that these graphs are universal with high probability. However, the proofs given in the previous section explicitly break down in the regime where $d = cn$. More strongly, it is known that in this regime one cannot find large [much larger than $\log(n)$-sized] nontrivial *induced subgraphs* with more than negligible probability [44]. This leads one to suspect the contrary of our earlier assertion, namely, that graph states of high degree are almost never universal resources. Indeed this is the case for *Haar-random states* [28], which with high probability have geometric entanglement so high that MBQC measurements can be effectively simulated by coin flips. In this section we show that this is not the case for random graph states by providing (almost matching) upper and lower bounds on the (expected) geometric entanglement. This is not proof positive of universality, but at least we avoid one known barrier. At the end of the section we discuss extensions to this simulation barrier specific to stabilizer states, conjecturing that this too can be avoided.

A. Lower bound on geometric entanglement

In this section we prove Theorem 4 and Corollary 3. We begin by extending a nice trick from compressed sensing (see, e.g., Lemma 4.4.1 in Ref. [88]) on approximating extremal singular values of matrices through ϵ-nets (we found essentially this argument in Ref. [89] but it is probably folklore in the tensor community).

Lemma 13. Consider an n-qubit state $|\psi\rangle$ and let $A_{\epsilon/n}$ be a $(\ln(3/2)/n)$-net of the set of single-qubit states. We then have

$$\max_{|\alpha\rangle\in \mathrm{PROD}_n} |\langle\alpha|\psi\rangle|^2 \leq 2 \max_{|\beta\rangle\in A_{\ln(3/2)/n}^{\otimes n}} |\langle\beta|\psi\rangle|^2. \tag{192}$$

Proof. Consider a state $|\alpha^*\rangle$ such that $\max_{|\alpha\rangle\in\mathrm{PROD}} |\langle\alpha|\psi\rangle|^2 = |\langle\alpha^*|\psi\rangle|^2$. Writing $|\alpha^*\rangle = \bigotimes_{i=1}^n |\alpha_i^*\rangle$, we choose for each $|\alpha_i^*\rangle$ a state $|\beta_i\rangle \in A_{\ln(3/2)/n}$ such that $\| |\alpha_i^*\rangle - |\beta_i\rangle \| \leq \ln(3/2)/n$. We can write

$$\begin{aligned}|\langle\alpha^*|\psi\rangle| &= |\langle\bigotimes_{i=1}^n (\langle\beta_i| + \langle\alpha_i^*| - \langle\beta_i|)|\psi\rangle| \\ &\leq |\langle\beta|\psi\rangle| + \sum_{i=1}^n \binom{n}{i} (\ln(3/2)/n)^i |\langle\alpha^*|\psi\rangle|\end{aligned} \tag{193}$$

using the fact that $|\alpha^*\rangle$ maximizes the overlap with $|\psi\rangle$ over normalized product states so that $|\langle\alpha^*|\psi\rangle - \langle\beta|\psi\rangle| \leq \| |\alpha\rangle - |\beta\rangle \| \, |\langle\alpha^*|\psi\rangle|$. The bound then follows from applying this fact for every tensor factor and $\| |\alpha_i^*\rangle - |\beta_i\rangle \| \leq \ln(3/2)/n$. Now note that

$$\begin{aligned}&\sum_{i=1}^n \binom{n}{i} (\ln(3/2)/n)^i \\ &\quad = (1 + \ln(3/2)/n)^n - 1 \leq e^{\ln(3/2)} - 1 = \frac{1}{2}.\end{aligned} \tag{194}$$

We can thus invert the relation above to get the lemma statement. ■

This lemma allows us to get a multiplicative approximation to the maximum overlap using a relatively weak (and thus small) ϵ-net for the product states.

Recall that we are trying to lower-bound the geometric entanglement. This is equivalent to upper-bounding the maximization over product states. The probability that the geometric entanglement is small is equal to the probability that the maximization of the overlap is large. Now, we upper-bound this probability. Combining the above argument with the union bound and the formula for average stabilizer states Eq. (9), we obtain the following theorem:

Theorem 8 (Restatement of Theorem 4). Choose a stabilizer state $|S\rangle$ uniformly at random. There exists a constant c such that

$$\mathbb{P}\left[E_g(|S\rangle) \leq n - c\sqrt{n}\log(n)\right] \leq O(2^{-\sqrt{n}}). \tag{195}$$

Proof. We begin by noting that

$$\begin{aligned}&\mathbb{P}\left[E_g(|S\rangle) \leq n - \delta\right] \\ &\quad = \mathbb{P}\left[-\log\left(\max_{\alpha\in\mathrm{PROD}_n} |\langle\alpha|\psi\rangle|^2\right) \leq n - \delta\right] \\ &\quad = \mathbb{P}\left[\max_{|\alpha\rangle\in\mathrm{PROD}_n} |\langle\alpha|S\rangle|^2 \geq 2^{-n+\delta}\right] \\ &\quad \leq \mathbb{P}\left[\max_{|\beta\rangle\in A_{1/n}^{\otimes n}} |\langle\beta|S\rangle|^2 \geq 2^{-n+\delta-1}\right],\end{aligned} \tag{196}$$

where $A_{\ln(3/2)/n}$ is a $\ln(3/2)/n$-net for the set of single-qubit states. We know that there exists such a net of size $|A_{\ln(3/2)/n}| \leq (5n/\ln(3/2))$ [90]. Using the union bound and Markov's inequality, we can upper-bound this latter quantity as

$$\mathbb{P}\left[\max_{|\beta\rangle\in A^{\otimes n}_{\ln(3/2)/n}} |\langle\beta|\psi\rangle|^2 \leq 2^{-n+\delta-1}\right] \leq 2^{n-\delta+1}\mathbb{E}_S\left[\sum_{|\beta\rangle\in A^{\otimes n}_{\ln(3/2)/n}} |\langle\beta|S\rangle|^{2t}\right]^{1/t} \qquad (197)$$

for some integer $t > 1$ (we will specify this later) Next we use the concavity of $x^{1/t}$ and the duality formula described in Eq. (9) to obtain

$$2^{n-\delta-1}\mathbb{E}_S\left[\sum_{|\beta\rangle\in A^{\otimes n}_{\ln(3/2)/n}} |\langle\beta|S\rangle|^{2t}\right]^{1/t} \leq 2^{n-\delta-1}|A_{\ln(3/2)/n}|^{n/t}2^{-n/t}(2^{t-1}+1)^{1/t}\left[\prod_{i=0}^{t-2}\frac{(2^i+1)}{(2^n+2^i)}\right]^{1/t}. \qquad (198)$$

Choosing $t = \sqrt{n}$ and working out, we obtain

$$2^{n-\delta+1}\mathbb{E}_S\left[\sum_{|\beta\rangle\in A^{\otimes n}_{\ln(3/2)/n}} |\langle\beta|S\rangle|^{2t}\right]^{1/t} \leq 5^{\sqrt{n}}2^{-\delta}2^{\log(n)\sqrt{n}}2^{\sqrt{n}}. \qquad (199)$$

Thus, setting $\delta = 2\log(n)\sqrt{n}$, we obtain what we set out to prove. ■

The above statement holds for uniformly random stabilizer states. We can lift it to a statement about graph states by noting that the probability that a uniformly random stabilizer state has full support on the computational basis states is bounded from below by a constant C. This is a standard fact, but we prove it here for completeness:

Corollary 11 (Restatement of Corollary 3). Choose a graph state $|G\rangle$ uniformly at random. There exists a constant c such that

$$\mathbb{P}\left[E_g(|G\rangle) \leq n - c\sqrt{n}\log(n)\right] \leq O(2^{-\sqrt{n}}). \qquad (200)$$

Proof. Consider the stabilizer states of full support:

$$\left|S_{U,b}\right\rangle := 2^{-n/2}\sum_{x\in\{0,1\}^n}(-1)^{x^TUx}\, i^{x^Tb}\,|x\rangle\,, \qquad (201)$$

with U a binary upper triangular matrix and $b \in \mathbb{Z}_4^n$. We will argue that all pairs U, b correspond to different quantum states. For two such states we can compute the inner product:

$$\langle S_{V,a}|S_{U,b}\rangle = 2^{-n}\sum_{x\in\{0,1\}^n}(-1)^{x^T(U+V)x}i^{x^T(a+b)}. \qquad (202)$$

The matrix $U + V \mod 2$ is again upper triangular, and hence a theorem due to Dickson (Theorem 4, p. 438, Chapter 5, paragraph 2 [91]) tells us that there exists an invertible binary matrix P such that $P(U+V)P^{-1} = D$ with

$$x^TDx = \sum_{i=1}^{\operatorname{rank}(U+V)/2} x_{2i-1}x_{2i}. \qquad (203)$$

Note that the rank of $U + V$ is always even, so this is sensible. Absorbing the map P into the summation, we can write

$$\langle S_{V,a}|S_{U,b}\rangle = 2^{-n}\sum_{x\in\{0,1\}^n}(-1)^{x^TDx}i^{(Px)^T(a+b)}. \qquad (204)$$

For this overlap to be 1, we must have

$$(-1)^{x^TDx}i^{(Px)^T(a+b)} = 1 \qquad (205)$$

for all $x \in \{0,1\}^n$. This immediately implies that $i^{(Px)^T(a+b)} = \pm 1$ and hence that $a + b = 2z$ for some $z \in \{0,1\}^n$. Furthermore, we must have $x^TDx + x^T(P^Tz) = 0$ for all $x \in \{0,1\}^n$. Setting $P^Tz = z'$, we can write this as

$$\begin{aligned} 0 = &\sum_{i=0}^{r/2} x_{2i}x_{2i+1} + x_{2i}z'_{2i} + x_{2i+1}z'_{2i+1} \\ &+ \sum_{i=r/2+1}^{n/2} x_{2i}z'_{2i} + x_{2i+1}z'_{2i+1}, \end{aligned} \qquad (206)$$

where r is the rank of $D = U + V$. Clearly $z'_{2i} = z'_{2i-1} = 0$ for $i > r/2$. Moreover, if $r > 1$, it is easy to see (by explicit enumeration) that for any choice of z'_0, z'_1 one of the evaluations of $x_0x_1 + x_0z'_0 + x_1z'_1$ is always nonzero. Hence, we must have $r = 0$, which implies $D = 0$. This also implies $a + b = 0$, since P is invertible. Hence, we must have $U = V$ and $a = b$, which is what we set out to argue. With this it is clear that the set $\{|S_{U,b}\rangle\}_{A,b}$ is of size $2^{n(n-1)/2} \times 2^{2n}$. Comparing this with the total number of stabilizer states $2^n \prod_{i=1}^n (2^i + 1)$, we see that

$$\begin{aligned} 2^{\frac{n(n-1)}{2}+2n} & \left[2^n \prod_{i=1}^{n}(2^i+1)\right]^{-1} \\ &= 2^{\frac{n(n-1)}{2}+2n-n-\frac{n(n+1)}{2}} \prod_{i=1}^{n}(1+2^{-i})^{-1} \\ &\geq \prod_{i=1}^{\infty}(1+2^{-i})^{-1} = C, \end{aligned} \tag{207}$$

where $C^{-1} \approx 2.3842$ (by numerical evaluation). Now we can prove the actual corollary statement by noting that (1) the geometric entanglement of $|S\rangle_{U,b}$ is independent of b and (2) the set $\{|S\rangle_{U,b}\}_b$ contains exactly one graph state (namely, $b = 0$). From this we can see that

$$\mathbb{P}_{|G\rangle}\left[E_g(|G\rangle) \leq n - \delta\right] = \mathbb{P}_{|S_{U,b}\rangle}\left[E_g(|S_{U,b}\rangle) \leq n - \delta\right], \tag{208}$$

where on the right-hand side we choose U, b uniformly at random. Since conditioning on inclusion in a subset under a uniform distribution yields a uniform distribution on that subset, we can say that

$$\mathbb{P}_{|S_{U,b}\rangle}\left[E_g(|S_{U,b}\rangle) \leq n - \delta\right] \leq C^{-1}\mathbb{P}_{|S\rangle}\left[E_g(|S\rangle) \leq n - \delta\right], \tag{209}$$

which is what we wanted to show. ■

B. Upper bound on geometric entanglement

The main goal of this section is to provide a proof of the upper bound in Theorem 3. The proof of this statement is substantially more difficult than the proof of the associated lower bound in Corollary 3. We begin by proving a lemma that exactly characterizes the maximal overlap of a random graph state with the set of real stabilizer product states: $R_n = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}^{\otimes n}$. This is a relatively small subset of all product states, and therefore gives a lower bound on the maximal overlap, which translates into an upper bound for the expected geometric entanglement since the overlap enters with a minus sign.

Lemma 14. Consider the set of real stabilizer product states $R_n = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}^{\otimes n}$. We then have

$$\begin{aligned} \mathbb{E}_G(E_g(|G\rangle)) &\leq \mathbb{E}_G\left[-\log\left(\max_{|s\rangle\in R_n} |\langle G|\, s\rangle|^2\right)\right] \\ &= n - \mathbb{E}_G\left[\max_{S\subseteq[n]}(|S| - \operatorname{rank}(A_G[S]))\right], \end{aligned} \tag{210}$$

where $A_G[S] \equiv A_G[S,S]$ is the adjacency matrix of G restricted to the index set S.

Proof. The first inequality is obvious from the definition of the geometric entanglement E_g. Now note that we can specify any state $|s\rangle \in R_n$ by a set $S \subset [n]$ indicating the Hadamard basis part of the state and two bitstrings $x_S \in \{0,1\}^{|S|}$ and $y_{\bar{S}} \in \{0,1\}^{n-|S|}$ encoding the phase and computational basis states in $S, \bar{S}$ so that $|s\rangle = 2^{-|S|/2} \sum_{z\in\{0,1\}^{|S|}} (-1)^{x_S^T z} |z\rangle\, |y_{\bar{S}}\rangle$. To ease the notation, let us assume without loss of generality that $S = \{1, \ldots, |S|\}$, i.e., the submatrix $A_G[S]$ is just the top-left corner of A_G. Then

$$\begin{aligned} \max_{|s\rangle\in R_n} |\langle G|\, s\rangle|^2 = 2^{-n-|S|} & \max_{S\subset[n]} \max_{x_S\in\{0,1\}^{|S|}} \max_{y_{\bar{S}}\in\{0,1\}^{n-|S|}} \\ & \times \left|\sum_{z\in\{0,1\}^{|S|}} (-1)^{(z,y_{\bar{S}})^T U_G (z,y_{\bar{S}}) + z^T x_S}\right|^2, \end{aligned} \tag{211}$$

where we recall that U_G is the adjacency matrix of G with the lower triangular part set to zero. Splitting the matrix U_G into submatrices $U_{SS}, U_{S\bar{S}}, U_{\bar{S}S}, U_{\bar{S}\bar{S}}$, we can write the phase inside the summation as

$$\begin{aligned} &(z, y_{\bar{S}})^T U (z, y_{\bar{S}}) + x_S^T z \\ &\quad = z^T U_{SS} z + \left(x_S^T + y_{\bar{S}}^T U_{S\bar{S}}^T + y_{\bar{S}}^T U_{\bar{S}S}\right) z + y_{\bar{S}}^T U_{\bar{S}\bar{S}} y_{\bar{S}}. \end{aligned} \tag{212}$$

The final term in this equation contributes only a global phase, while the middle one can be absorbed in the maximization over x_S. Hence, we get

$$\begin{aligned} \max_{|s\rangle\in R_n} |\langle G|\, s\rangle|^2 = 2^{-n-|S|} & \max_{S\subset[n]} \max_{x_S\in\{0,1\}^{|S|}} \\ & \times \left|\sum_{z\in\{0,1\}^{|S|}} (-1)^{z^T U_{SS} z + x_S^T z}\right|^2. \end{aligned} \tag{213}$$

Next we use a theorem due to Dickson (Theorem 2 in Chap. 15 in Ref. [91]), which tells us that there exists an invertible binary matrix P_{SS} such that $P_{SS} U_{SS} P_{SS}^{-1} = D_{SS}$ with

$$z^T D_{SS} z = \sum_{i=1}^{\operatorname{rank}(U_{SS})/2} z_{2i-1} z_{2i}. \tag{214}$$

Note that the rank of U_{SS} is always even, so this equation makes sense. Absorbing P_{SS} in the sum over z and subsequently into the maximization over x_S, we obtain

$$\max_{|s\rangle\in R_n} |\langle G|s\rangle|^2 = \max_{S\subset[n]} \max_{x_S\in\{0,1\}^{|S|}} \left| \sum_{a_S\in\{0,1\}^S} (-1)^{z^T D_{SS} z + x_S^T z} \right|^2 . \quad (215)$$

Using the definition of D_{SS}, we see that this equation factorizes, and we obtain

$$\max_{|s\rangle\in R_n} |\langle G|s\rangle|^2 = 2^{-n-|S|} \max_{S\subset[n]} \left(\max_{y_1,y_2\in\{0,1\}} \left| \sum_{a_1,a_2\in\{0,1\}} (-1)^{a_1a_2+y_1a_1+y_2a_2} \right|^2 \right)^{\mathrm{rank}(U_{SS})/2} . \quad (216)$$

The inner maximization can easily be solved to obtain

$$\max_{|s\rangle\in R_n} |\langle G|s\rangle|^2 = 2^{-n-|S|}\, 2^{\mathrm{rank}(U_{SS})}. \quad (217)$$

Noting that $\mathrm{rank}(U_{SS}) = \mathrm{rank}(A_G[S])$, we obtain the lemma statement. ■

This means that the expected geometric entanglement can be controlled by the average rank of submatrices of random adjacency matrices. To determine this, we will need two facts from probability theory and classical coding theory:

Fact 1 (Bonferroni inequalities). Let $\{E_i\}_{i\in\Omega}$ be a countable set of events. Then

$$\mathbb{P}\left(\bigcup_{i\in\Omega} E_i\right) \le \sum_{i\in\Omega} \mathbb{P}(E_i), \quad \text{(union bound)}$$

$$\mathbb{P}\left(\bigcup_{i\in\Omega} E_i\right) \ge \sum_{i\in\Omega} \mathbb{P}(E_i) - \frac{1}{2} \sum_{i,j\in\Omega, i\neq j} \mathbb{P}\left(E_i \cup E_j\right). \quad \text{(Bonferroni)}$$

Our goal will be to use the second inequality to lower-bound the probability that any submatrix $A_G[S]$ of a random adjacency matrix A_G has excessive rank deficit. We can characterize the distribution of this rank for a fixed S exactly.

Fact 2 (Random adjacency matrices). Consider the set of $n \times n$ binary symmetric matrices with zeros on the diagonal (i.e., adjacency matrices). Choosing a matrix A uniformly from this set, we have $\mathbb{P}(\mathrm{rank}(A_G) = 2h + 1) = 0$) and

$$\mathbb{P}(\mathrm{rank}(A_G) = 2h) = 2^{-n^2/2+n/2} \prod_{i=1}^{h} \frac{2^{2i-2}}{2^{2i}-1} \prod_{i=0}^{2h-1} (2^{n-i}-1), \quad (218)$$

with $h \in [0, \lfloor n/2 \rfloor]$.

The formula above, found in Theorem 2 in Chap. 15 in Ref. [91] (see Ref. [92]), can be both upper-bounded and lower-bounded by a Gaussian. We have the following lemma, which follows from a straightforward calculation:

Lemma 15. Consider a uniformly random adjacency matrix A as in Fact 2. We have the following approximation:

$$\frac{e^{-2}}{4} 2^{-\frac{(n-2h)^2}{2}} 2^{\frac{(n-2h)}{2}} \le \mathbb{P}(\mathrm{rank}(A) = 2h) \le e^{2/3} 2^{-\frac{(n-2h)^2}{2}} 2^{\frac{(n-2h)}{2}}, \quad (219)$$

with $h \in [0, \lfloor n/2 \rfloor]$.

This immediately translates into a tail bound on the *rank deficiency* of a submatrix $A_G[S]$ for a fixed S.

Lemma 16. Consider a uniformly random $n \times n$ adjacency matrix A and a set of indices $S \subseteq [n]$ and define the event $E_S(t) = \{|S| - \mathrm{rank}(A) \ge t\}$ with $t \in [0, |S|]$. We have

$$\mathbb{P}(E_S(t)) \ge \frac{e^{-2}}{4} 2^{-\frac{t^2}{2}-\frac{t}{2}}. \quad (220)$$

This allows us to control the first term in the Bonferroni inequality. To characterize the second term will we need to do substantially more work. First we note that the rank of two submatrices $A[S], A[S']$ is independent conditioned on the rank of the intersection matrix $A[S \cap S']$.

Lemma 17. Consider a uniformly random $n \times n$ adjacency matrix A and two sets of indices $S, S' \subseteq [n]$, with $S \cap S' =: I$. Conditioned on the rank of the intersection ($\mathrm{rank}(A[I])$) the ranks of $A[S]$ and $A[S']$ are independent, i.e.,

$$\mathbb{P}\left(\operatorname{rank}(A[S]) = i, \operatorname{rank}\big(A[S']\big) = i'\right)$$
$$= \sum_{j=0}^{\lfloor |I|/2 \rfloor} \mathbb{P}\left(\operatorname{rank}(A[S]) = i \,|\, \operatorname{rank}(A[I]) = j\right) \mathbb{P}\left(\operatorname{rank}\big(A[S']\big) = i' \,|\, \operatorname{rank}(A[I]) = j\right) \mathbb{P}\left(\operatorname{rank}(A[I]) = j\right). \tag{221}$$

Proof. Certainly $A[S]$ and $A[S']$ are independent conditioned on the intersection $A[I]$, since the matrix elements $A[S]$ and $A[S']$ outside the intersection are independent. It is thus sufficient to establish that the rank of $A[S]$ depends only on the rank of $A[I]$ (and similarly for $A[S']$). Consider two matrices B_I, B'_I with $\operatorname{rank}(B_I) = \operatorname{rank}\big(B'_I\big)$. From Dickson's theorem we know there exists an invertible matrix R_I such that $B'_I = R_I B_I R_I^T$. Now consider the induced distributions $\mathbb{P}\left(A[S] \,|\, A[I] = B_I\right)$ and $\mathbb{P}\left(A[S'] \,|\, A[I] = B'_I\right)$. We have $\mathbb{P}\left(A[S] \,|\, A[I] = B_I\right) = \mathbb{P}\left((\mathbb{1} \oplus R_I^{-1})A[S'](\mathbb{1} \oplus R_I^{T^{-1}}) \,|\, A[I] = B_I\right)$. Clearly $\operatorname{rank}\left((\mathbb{1} \oplus R_I^{-1})A[S'](\mathbb{1} \oplus R_I^{T^{-1}})\right) = \operatorname{rank}\big(A[S']\big)$, which gives us the desired result. ■

With the above we can control the joint probability of the rank of two submatrices $A[S], A[S']$ in terms of only their intersection. Next we will show that if $|S|$ is much larger than $|I|$, then the rank of $A[S]$, conditioned on the rank of $A[I]$, is close to maximal with high probability (provided that the rank of $A[I]$ is not too small). We do this by a reduction of the problem to an infinite-dimensional Markov chain together with precise bounds on its convergence rate.

Lemma 18. Consider a uniformly random $n \times n$ adjacency matrix A and two sets of indices $S, I \subseteq [n]$, with $I \subseteq S$. Also consider integers i,j such that $|S| - i, |I| - j$ are even. There exists a constant $\rho < 1 - 10^{-5}$ such that

$$\mathbb{P}\left(\operatorname{rank}(A[S]) \leq |S| - i \,|\, \operatorname{rank}(A[I]) = |I| - j\right)$$
$$\leq e^{2/3} 2^{-\frac{(i)^2}{2} - \frac{(i)}{2}} + 7 \cdot 2^j \rho^{|S|-|I|} \tag{222}$$

for $i \geq 0$ and $\alpha > 1$.

Proof. The main strategy of this proof is to rewrite the left-hand side of Eq. (222) in terms of the convergence properties of a (formally infinite) Markov chain P and then use classical Markov chain bounding techniques (in particular the drift and minorization method [84,85]) to provide bounds on these convergence properties.

To construct the Markov chain, consider an $m \times m$ symmetric Boolean matrix A of rank r (with zeros on the diagonal). We will now symmetrically add a vector $v \in \{0,1\}^m$ to the rows and columns. If we add the column first, the probability that $\operatorname{rank}\big(\begin{pmatrix} v & A \end{pmatrix}\big) = r$ is 2^{r-m} and the probability that $\operatorname{rank}\big(\begin{pmatrix} v & A \end{pmatrix}\big) = r + 1$ is $1 - 2^{r-m}$. In the second case it is clear that

$$\operatorname{rank}\left(\begin{pmatrix} 0 & v^T \\ v & A \end{pmatrix}\right) = r + 2$$

since column and row rank are always the same. In the first case we note that there exists an x subject to $Ax = v$. This immediately implies that

$$\begin{pmatrix} v^T \\ A \end{pmatrix} x = \begin{pmatrix} v^T x \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{223}$$

because $v^T x = x^T A x = 0$ by the symmetry of A and the fact that we are working over the field $\mathbb{F}_2$. This implies that

$$\operatorname{rank}\left(\begin{pmatrix} 0 & v^T \\ v & A \end{pmatrix}\right) = r$$

(this elegant argument is due to Sloane and MacWilliams; see Lemma 3 in Chap. 15 in Ref. [91]). The sequential adding of random k columns (and rows) to a matrix A of rank r, and considering their rank, thus induces a sequence of random variables $R_0 = r, R_1, \ldots, R_k$. From the above discussion, this sequence is a (time-dependent) Markov chain with transitions $r \xrightarrow{2^{r-m}} r$, $r \xrightarrow{1-2^{r-m}} r + 2$ (where m takes the role of time).

Changing variables from the rank R_k to the rank deficiency $D_k = m + k - R_k$, we obtain another Markov chain, whose transition probabilities no longer depend on the ambient matrix dimension m (it is now homogeneous and formally infinite dimensional). The associated Markov generator is given explicitly as

$$M(j, i) := \mathbb{P}\left(D_{k+1} = j \,\middle|\, D_k = i\right)$$
$$= \begin{cases} 2^{-i} & \text{if } j = i + 1, \\ 1 - 2^{-i} & \text{if } j = i - 1, \\ 0 & \text{if } |j - i| > 1 \end{cases} \tag{224}$$

for $i,j \in \mathbb{N}$. This Markov chain is irreducible, but is periodic with period 2. Hence, it is natural to consider $Q = M^2$, which will be aperiodic, but decomposes into even and odd irreducible aperiodic subchains. We will now bound the convergence of the Markov chain Q with initial state e_r (the unit vector with 1 on the rth position). We will treat the even subchain in detail (with the odd subchain being analogous), so we assume that r is even. The stationary state π of M^2 on the even subspace can be found by appropriately taking the limit of Eq. (218) to $n \to \infty$:

$$\pi : \mathbb{N}/2 \to \mathbb{R} : i \to C^{-1} 2^{-\frac{i^2}{2}+\frac{i}{2}} \prod_{t=\frac{i}{2}+1}^{\infty} \frac{2^{2t-i}}{2^{2i-t}-1} \times \prod_{t=i+1}^{\infty} (1-2^{-t}), \tag{225}$$

where $C \geq 1$ is some appropriate normalization. The above makes intuitive sense, because it corresponds to the situation where $|S|$ is much larger than $|I|$, and thus the rank distribution of $A[|S|]$ stops depending on I. One can also explicitly verify that π is an eigenvector of M^2 with eigenvalue 1. We will use the drift and minorization method (see Theorem 5) to bound convergence to this distribution. For this bounding method we need to provide a drift function $V : \mathbb{N}/2 \to \mathbb{R}$ and a small set $C \subset \mathbb{N}/2$. We will choose $V(i) = 2^i$ and $C = \{0, 2\}$. It is tedious but straightforward to check that P^2 satisfies the conditions of Theorem 5 with parameters $\lambda = 0.55, b = 2, \delta = 0.2, d = 9, r = 0.001$ (all the difficulty lies in choosing the parameters). This bounds the convergence of the Markov chain in total variation distance as

$$\left\| \pi - M^{2k} e_r \right\|_{TV} \leq \rho_1^{2k} + \rho_2^{2k}\left(6 + 2^r\right) \leq 7\rho^k 2^r, \tag{226}$$

with $\rho = \max\{\rho_1, \rho_2\}$ and $\rho_1 < 1 - 10^{-5}, \rho_2 < 1 - 10^{-3}$. We obtain the same expression for the odd subchain.

Mapping back to our original question, we see that

$$\begin{aligned} &\mathbb{P}\left(\operatorname{rank}(A[S]) \leq |S| - i \,|\, \operatorname{rank}(A[I]) = |I| - j\right) \\ &\quad = \mathbb{P}\left(D_{|S|-|I|} \geq i \,|\, \operatorname{rank}(A[I]) = |I| - j\right) \\ &\quad \leq \sum_{t=i}^{\infty} \left(M^{|S|-|I|} e_{|I|-j}\right)_t, \end{aligned} \tag{227}$$

where we recall that e_i is ith unit vector on $\mathbb{R}^{\mathbb{N}}$ and $(\cdot)_t$ denotes the tth element. Using the triangle inequality on $(M^{|S|-|I|} e_{|I|-j} - \pi) + \pi$, we obtain

$$\begin{aligned} &\mathbb{P}\left(\operatorname{rank}(A[S]) \leq |S| - i \,|\, \operatorname{rank}(A[I]) = |I| - j\right)] \\ &\quad \leq 7\rho^{|S|-|I|} 2^j + \sum_{t=i}^{\infty} \pi_t = 7\rho^{|S|-|I|} 2^j + 22^{-\frac{i^2}{2}+\frac{i}{2}} \end{aligned}$$

using the definition of the distribution π and the basic upper bound $\prod_{t=(i/2)+1}^{\infty} (2^{2t-i}/2^{2i-t}) \leq 2$. ■

With this lemma under our belt it is finally time to prove the main theorem of this section.

Theorem 9 (Restatement of Theorem 3). Choose a graph state $|G\rangle$ on n qubits uniformly at random, where we assume $n = k^2$ for some integer k. We have

$$\mathbb{E}_G\left(E_g(|G\rangle)\right) \leq n - \Omega(n^{1/4}/\log(n)). \tag{228}$$

Proof. We begin by upper-bounding the geometric entanglement of a graph state $|G\rangle$ in terms of the maximal rank deficiency of the principal submatrices of the adjacency matrix of G. Lemma 14 tells us that

$$\mathbb{E}_G\left(E_g(|G\rangle)\right) \leq n - \mathbb{E}_G\left[\max_{S \subseteq [n]} \left(|S| - \operatorname{rank}(A_G[S])\right)\right]. \tag{229}$$

We can further upper-bound this by maximizing only over sets S that are pairwise far away in edit distance [93]. Define the set of sets $\mathcal{Q}$ by dividing $[n]$ up into intervals of size $\sqrt{n}$ (which is an integer by assumption) and taking all sets $S \subset [n]$ that contain $\sqrt{n}/2$ such intervals. This implies that all $S \in \mathcal{Q}$ have size $|S| = n/2$, that $|\mathcal{Q}| = \binom{\sqrt{n}}{\sqrt{n}/2}$, and that all sets in $\mathcal{Q}$ are pairwise distant (at least $\sqrt{n}$) in edit distance.

Since $|S| = n/2$ for $S \in \mathcal{Q}$, we can lower-bound the expected maximal rank deficiency (over $\mathcal{Q}$) by defining the events $E_S(t) = \{\operatorname{rank}(A) \leq n/2 - t\}$; recall Lemma 16. From the Markov and Bonferroni inequalities (Fact 1) we see

$$\mathbb{E}_G\left[\max_{S \in \mathcal{Q}} \left(|S| - \operatorname{rank}(A_G[S])\right)\right] \geq t\mathbb{P}\left(\bigcup_{S \in \mathcal{Q}} E_S(t)\right) \tag{230}$$

$$\geq t\left[\underbrace{\sum_{S \in \mathcal{Q}} \mathbb{P}\left(E_S(t)\right)}_{(1)} - \frac{1}{2} \underbrace{\sum_{S, S' \in \mathcal{Q}, S \neq S'} \mathbb{P}\left(E_S(t) \cup E_{S'}(t)\right)}_{(2)}\right]. \tag{231}$$

We can give a lower bound of the first term (1) using Lemma 16:

$$(1) = \sum_{S \in \mathcal{Q}} \mathbb{P}\left(E_S(n/2 - t)\right) \geq \frac{1}{4} e^{-2} \binom{\sqrt{n}}{\sqrt{n}/2} e^{-\frac{t^2+t}{2}}. \tag{232}$$

It remains to upper-bound the second term (2). To this end, we first recall Lemma 17, which implies that the joint probability $\mathbb{P}\left(E_S(t) \cup E_{S'}(t)\right)$ depends only on the size of the intersection $I = S \cap S'$. The size of this intersection can take values $|I| = 0, \sqrt{n}, \ldots, n/2 - \sqrt{n}$. Given a set $S \in \mathcal{Q}$, for each value of $w \in [0, \sqrt{n}/2 - 1]$ there are $\binom{\sqrt{n}}{w}\binom{\sqrt{n}}{\sqrt{n}-w}$ sets $S' \in \mathcal{Q}$ satisfying $|S \cap S'| = w\sqrt{n}$.

Let us choose representatives $S_0 := [n/2]$, $S'_w := [n/2 - \sqrt{n}w, n - \sqrt{n}w]$ with $I_w := S_0 \cap S'_w$ of size $|I_w| = w\sqrt{n}$. We can then write

$$(2) = \sum_{S,S' \in \mathcal{Q},\, S \neq S'} \mathbb{P}\left(E_S(t) \cup E_{S'}(t)\right) \tag{233}$$

$$= \sum_{S,S' \in \mathcal{Q},\, S \neq S'} \sum_{i,i'=0}^{t} \mathbb{P}\left(\mathrm{rank}(A[S]) = i \vee \mathrm{rank}\big(A[S']\big) = i'\right) \tag{234}$$

$$= \sum_{S,S' \in \mathcal{Q},\, S \neq S'} \sum_{j=0}^{|S \cap S'|} \sum_{i,i'=j}^{t} \mathbb{P}\left(\mathrm{rank}(A[S]) = i \,|\, \mathrm{rank}(A[I]) = j\right) \mathbb{P}\left(\mathrm{rank}(A[S]) = i \,|\, \mathrm{rank}(A[I]) = j\right) \mathbb{P}\left(\mathrm{rank}(A[I]) = j\right) \tag{235}$$

$$= \binom{\sqrt{n}}{\sqrt{n}/2} \sum_{w=0}^{\sqrt{n}/2-1} \binom{\sqrt{n}}{w} \binom{\sqrt{n}}{\sqrt{n}/2 - w} \sum_{j=0}^{w\sqrt{n}/2} \sum_{i,i'=j}^{t} \mathbb{P}\left(\mathrm{rank}(A[S_0]) = i \,|\, \mathrm{rank}(A[I_w]) = j\right)$$

$$\times\, \mathbb{P}\left(\mathrm{rank}\big(A[S'_w]\big) = i' \,|\, \mathrm{rank}(A[I_w]) = j\right) \mathbb{P}\left(\mathrm{rank}(A[I_w]) = j\right) \tag{236}$$

$$= \binom{\sqrt{n}}{\sqrt{n}/2} \sum_{w=0}^{\sqrt{n}/2-1} \binom{\sqrt{n}/2}{w} \binom{\sqrt{n}/2}{\sqrt{n}/2 - w} \sum_{j=0}^{w\sqrt{n}/2} \left(\sum_{i=j}^{n/2-t} \mathbb{P}\left(\mathrm{rank}(A[S_0]) = i \,|\, \mathrm{rank}(A[I_w]) = j\right) \right)^2 \mathbb{P}\left(\mathrm{rank}(A[I_w]) = j\right). \tag{237}$$

In the last line we used that since $|S_0| = |S'_w|$ we have

$$\mathbb{P}\left(\mathrm{rank}(A[S_0]) = i \,|\, \mathrm{rank}(A[I_w]) = j\right) = \mathbb{P}\left(\mathrm{rank}\big(A[S'_w]\big) = i \,|\, \mathrm{rank}(A[I_w]) = j\right). \tag{238}$$

Next we use Fact 2, which tells us the rank of $A[I_w]$ is likely nearly maximal. To make this precise we introduce a constant $\alpha \geq 1$ which will be precisely determined later, and split the sum over j into $j < |I_w| - \alpha t$ and $j \geq |I_w| - \alpha t$. Focusing on this sum, we see

$$\sum_{j=0}^{\lfloor\sqrt{n}w/2\rfloor} \left[\sum_{i=j}^{n/2-t} \mathbb{P}\left(\mathrm{rank}(A[S_0]) = i \,|\, \mathrm{rank}(A[I_w]) = j\right) \right]^2 \mathbb{P}\left(\mathrm{rank}(A[I_w]) = j\right) \tag{239}$$

$$\leq \sum_{j=\alpha t}^{\lfloor\sqrt{n}w/2\rfloor} \left[\sum_{i=j}^{n/2-t} \mathbb{P}\left(\mathrm{rank}(A[S_0]) = i \,|\, \mathrm{rank}(A[I_w]) = j\right) \right]^2$$

$$+ e^{2/3} \sum_{j=0}^{\lfloor\sqrt{n}w/2\rfloor - \alpha t - 1} 2^{-\frac{(\lfloor\sqrt{n}w/2\rfloor - j)^2}{2}} 2^{\frac{(\lfloor\sqrt{n}w/2\rfloor - j)}{2}} \left[\sum_{i=j}^{n/2-t} \mathbb{P}\left(\mathrm{rank}(A[S_0]) = i \,|\, \mathrm{rank}(A[I_w]) = j\right) \right]^2 \tag{240}$$

$$\leq \sum_{j=\alpha t}^{\lfloor\sqrt{n}w/2\rfloor} \left[\sum_{i=j}^{n/2-t} \mathbb{P}\left(\mathrm{rank}(A[S_0]) = i \,|\, \mathrm{rank}(A[I_w]) = j\right) \right]^2 + e^{2/3} \sum_{j=0}^{\lfloor\sqrt{n}w/2\rfloor - \alpha t - 1} 2^{-\frac{(\lfloor\sqrt{n}w/2\rfloor - j)^2}{2}} 2^{\frac{(\lfloor\sqrt{n}w/2\rfloor - j)}{2}} \tag{241}$$

$$\leq \sum_{j=\alpha t}^{\lfloor\sqrt{n}w/2\rfloor} \left[\sum_{i=j}^{n/2-t} \mathbb{P}\left(\mathrm{rank}(A[S_0]) = i \,|\, \mathrm{rank}(A[I_w]) = j\right) \right]^2 + e^{2/3} \sum_{j=0}^{\lfloor\sqrt{n}w/2\rfloor - \alpha t - 1} 2^{-\frac{(\lfloor\sqrt{n}w/2\rfloor - j)^2}{2}} 2^{\frac{(\lfloor\sqrt{n}w/2\rfloor - j)}{2}}, \tag{242}$$

where we have used Fact 2 to upper-bound $\mathbb{P}\left(\mathrm{rank}(A[I]) = j\right)$ in the second term, and have trivially bounded $\mathbb{P}\left(\mathrm{rank}(A[I]) = j\right) \leq 1$ in the first. It remains to use Lemma 18 to bound the first term. We obtain

$$(2) \leq \binom{\sqrt{n}}{\sqrt{n}/2} \sum_{w=0}^{\sqrt{n}/2-1} \binom{\sqrt{n}/2}{w} \binom{\sqrt{n}/2}{\sqrt{n}/2-w} \left[\sum_{j=\alpha t}^{w\sqrt{n}/2} \left(\sum_{i=j}^{n/2-t} 2 \cdot 2^{-\frac{(2i)^2}{2}+\frac{(2i)}{2}} + 7 \cdot 2^{2i\alpha} \rho^{(\sqrt{n}/2-w)\sqrt{n}} \right)^2 + 2 \cdot 2^{-\frac{(\alpha t-1)^2}{2}} \right]. \tag{243}$$

Setting $t = 5\sqrt{\log\left[\binom{\sqrt{n}}{\sqrt{n}/2}\right]}$ and $\alpha = 10$, we can evaluate the resulting expression (packing all numbers into a constant $C > 0$) and obtain

$$\begin{aligned} &\mathbb{E}_G \left(\max_{S \in \mathcal{Q}} \frac{n}{2} - \operatorname{rank}(A_G[S]) \right) \\ &\geq \binom{\sqrt{n}}{\sqrt{n}/2} \frac{1}{4} e^{-2} e^{-\frac{t^2+t}{2}} - \binom{\sqrt{n}}{\sqrt{n}/2} \sum_{w=0}^{\sqrt{n}/2-1} \binom{\sqrt{n}/2}{w} \binom{\sqrt{n}/2}{\sqrt{n}/2-w} \left[\sum_{j=\alpha t}^{\lfloor \sqrt{n} w/2 \rfloor} \left(\sum_{i=j}^{n/2-t} 2 \cdot 2^{-\frac{(2i)^2}{2}+\frac{(2i)}{2}} \right.\right. \\ &\quad \left.\left. + \; 7 \cdot 2^{2i\alpha} \rho^{(\sqrt{n}/2-w)\sqrt{n}} \right)^2 + 2 \cdot 2^{-\frac{(\alpha t-1)^2}{2}} \right] \\ &\geq C\sqrt{\log\left[\binom{\sqrt{n}}{\sqrt{n}/2}\right]} = \Omega\left(n^{(1/4)}/\log(n)\right), \end{aligned} \tag{244}$$

which is what we set out to prove. ■

C. Structured simulation algorithms

The upper bound on the geometric entanglement derived above is strong enough to break the simulation algorithm derived in Ref. [28]. However, better simulation algorithms might be found by exploiting the extra structure that graph states provide. In particular, we can think of any MBQC procedure as a sequence of measurements in the eigenbases of the $X, Y, X - Y, X + Y$ operators [36]. On a graph state the first three measurements can be simulated classically in polynomial time. We can thus envision an improved simulation algorithm where we classically simulate the X, Y, Z measurements and simulate the $X + Y, X - Y$ measurements by coin flips. The efficacy of this algorithm depends critically on how the overlap with a random graph state fluctuates with respect to tensor products of the eigenstates of $X + Y, X - Y$. As these states are highly magical, it is possible that their overlap with stabilizer states fluctuates much less than the maximal overlap with arbitrary stabilizer states (which, as we saw above, is dominated by contributions from product *stabilizer* states). We believe that this is not the case (and thus that this algorithm does not work), but cannot prove it as of yet. We think this is an interesting question in its own right, so we leave it as a conjecture.

Conjecture 1. If we label the set of eigenstates of $X + Y, X - Y$ as $\mathcal{T}$, there exist $c, C > 0$ subject to

$$\mathbb{P}\left[\max_{|\alpha\rangle \in \mathcal{T}^{\otimes n}} |\langle \alpha | G \rangle|^2 \geq 2^{-n+n^c} \right] \geq C, \tag{245}$$

with the probability taken uniformly over graph states.

ACKNOWLEDGMENTS

J.H. thanks Ingo Roth and Jop Briet for useful suggestions on the proof of Theorem 3, and Michael Walter and Marcel Hinsche for helpful discussions on the Clifford group. D.H. thanks Michael Gullans for many inspiring discussions on statistical-mechanics mappings and IQP circuits. We also thank Daniel Grier, Jackson Morris, and Debbie Leung for discussions on vertex minors and simulation algorithms. We gratefully acknowledge the hospitality in the spring of 2024 of the Simons Institute for the Theory of Computing, supported by DOE Quantum Systems Accelerator Grant No. FP00010905, where part of this work was conducted. D.H. acknowledges funding from the Simons Institute for the Theory of Computing, supported by DOE Quantum Systems Accelerator. J.H. acknowledges funding from the Dutch Research Council (NWO) through a Veni grant (Grant No. VI.Veni.222.331) and the Quantum Software Consortium (NWO Zwaartekracht Grant No. 024.003.037).

DATA AVAILABILITY

No data were created or analyzed in this study.

[1] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. Van den Nest, and H.-J. Briegel, Entanglement in graph states and its applications, in *Proceedings of the International School of Physics "Enrico Fermi"* (2006), Vol. 162, pp. 115–218.

[2] Robert Raussendorf, Cihan Okay, Dong-Sheng Wang, David T. Stephen, and Hendrik Poulsen Nautrup, A computationally universal phase of quantum matter, Phys. Rev. Lett. **122**, 09051 (2018).

[3] Axel Dahlberg, Jonas Helsen, and Stephanie Wehner, How to transform graph states using single-qubit operations: Computational complexity and algorithms, Quantum Sci. Technol. **5**, 045016 (2020).

[4] Frederik Hahn, Anna Pappa, and Jens Eisert, Quantum network routing and local complementation, npj Quantum Inf. **5**, 76 (2019).

[5] M. Van den Nest, J. Dehaene, and B. De Moor, Graphical description of the action of local Clifford transformations on graph states, Phys. Rev. A **69**, 022316 (2004).

[6] Robert Raussendorf and Hans J. Briegel, A one-way quantum computer, Phys. Rev. Lett. **86**, 5188 (2001).

[7] R. Raussendorf, D. E. Browne, and H. J. Briegel, Measurement-based quantum computation on cluster states, Phys. Rev. A **68**, 022312 (2003).

[8] Atul Mantri, Tommaso F. Demarie, and Joseph F. Fitzsimons, Universality of quantum computation with cluster states and (x,y)-plane measurements, Sci. Rep. **7**, 42861 (2017).

[9] M. Hein, J. Eisert, and H. J. Briegel, Multiparty entanglement in graph states, Phys. Rev. A **69**, 062311 (2004).

[10] M. Van den Nest, W. Dür, G. Vidal, and H. J. Briegel, Classical simulation versus universality in measurement based quantum computation, Phys. Rev. A **75**, 012337 (2007).

[11] Dolev Bluvstein, *et al.*, Logical quantum processor based on reconfigurable atom arrays, Nature **626**, 58 (2024).

[12] Matthew DeCross, *et al.*, Computational power of random quantum circuits in arbitrary geometries, Phys. Rev. X **15**, 021052 (2025).

[13] D. Hangleiter, M. Kalinowski, D. Bluvstein, M. Cain, N. Maskara, X. Gao, A. Kubica, M. D. Lukin, and M. J. Gullans, Fault-tolerant compiling of classically hard instantaneous quantum polynomial circuits on hypercubes, PRX Quantum **6**, 020338 (2025).

[14] Louis Paletta, Anthony Leverrier, Alain Sarlette, Mazyar Mirrahimi, and Christophe Vuillot, Robust sparse IQP sampling in constant depth, Quantum **8**, 1337 (2024).

[15] Simon Martiel and Ali Javadi-Abhari, Low-overhead error detection with spacetime codes, ArXiv:2504.15725.

[16] Martin Ringbauer, Marcel Hinsche, Thomas Feldker, Paul K. Faehrmann, Juani Bermejo-Vega, Claire Edmunds, Lukas Postler, Roman Stricker, Christian D. Marciniak, Michael Meth, Ivan Pogorelov, Rainer Blatt, Philipp Schindler, Jens Eisert, Thomas Monz, and Dominik Hangleiter, Verifiable measurement-based quantum random sampling with trapped ions, Nat. Commun. **16**, 106 (2025).

[17] Thiago Bergamaschi, Chi-Fang Chen, and Yunchao Liu, Quantum computational advantage with constant-temperature Gibbs sampling, in *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Chicago, IL, USA, 2024).

[18] Examples of such applications include secret sharing [94] and multiparty computation [95]; see Ref. [1] for an early overview.

[19] S. Ghosh, A. Deshpande, D. Hangleiter, A. V. Gorshkov, and B. Fefferman, Complexity phase transitions generated by entanglement, Phys. Rev. Lett. **131**, 030601 (2023).

[20] D. B. Hangleiter and J. Eisert, Computational advantage of quantum random sampling, Rev. Mod. Phys. **95**, 035001 (2023).

[21] Random instances from this family can be efficiently sampled [96].

[22] Another, unrelated, setting in which random graph states arise naturally is measurement-based computation based on cluster states [6] with independent and identically distributed erasure noise on all qubits [97], or nondeterministic gates [98]. In these scenarios there are percolation thresholds between easy-to-simulate and hard-to-simulate phases.

[23] Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd, Achieving quantum supremacy with sparse and noisy commuting quantum computations, Quantum **1**, 8 (2017).

[24] J. C. Napp, R. L. La Placa, A. M. Dalzell, F. G. S. L. Brandão, and A. W. Harrow, Efficient classical simulation of random shallow 2D quantum circuits, Phys. Rev. X **12**, 021021 (2022).

[25] Thomas Schuster, Jonas Haferkamp, and Hsin-Yuan Huang, Random unitaries in extremely low depth, Science **389**, 92 (2025).

[26] Scott Aaronson and Alex Arkhipov, The computational complexity of linear optics, Theory Comput. **9**, 143 (2013).

[27] Dominik Hangleiter, Juan Bermejo-Vega, Martin Schwarz, and Jens Eisert, Anticoncentration theorems for schemes showing a quantum speedup, Quantum **2**, 65 (2018).

[28] D. Gross, S. T. Flammia, and J. Eisert, Most quantum states are too entangled to be useful as computational resources, Phys. Rev. Lett. **102**, 190501 (2009).

[29] Mykhailo Krawtchouk, Sur une généralisation des polynomes d'Hermite, C. R. Math. **189**, 620 (1929).

[30] Philip Feinsilver and Jerzy Kocik, Krawtchouk Polynomials and Krawtchouk Matrices, in *Recent Advances in Applied Probability*, edited by R. Baeza-Yates, J. Glaz, H. Gzyl, J. Hüsler, and J. L. Palacios (Springer, Boston, MA, 2005), p. 115.

[31] Alexander M. Dalzell, Nicholas Hunter-Jones, and Fernando G. S. L. Brandão, Random quantum circuits anticoncentrate in log depth, PRX Quantum **3**, 010333 (2022).

[32] Abhinav Deshpande, Pradeep Niroula, Oles Shtanko, Alexey V. Gorshkov, Bill Fefferman, and Michael J. Gullans, Tight bounds on the convergence of noisy random circuits to the uniform distribution, PRX Quantum **3**, 040329 (2022).

[33] 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).

[34] Nicholas C. Wormald, Generating random regular graphs, J. Algorithms **5**, 247 (1984).

[35] Svante Janson, Random regular graphs: Asymptotic aistributions and contiguity, Comb. Probab. Comput. **4**, 369 (1995).

[36] Maarten Van den Nest, Akimasa Miyake, Wolfgang Dür, and Hans J. Briegel, Universal resources for measurement-based quantum computation, Phys. Rev. Lett. **97**, 150504 (2006).

[37] 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).

[38] Xun Gao, Sheng-Tao Wang, and L.-M. Duan, Quantum supremacy for simulating a translation-invariant Ising spin model, Phys. Rev. Lett. **118**, 040502 (2017).

[39] J. Haferkamp, D. Hangleiter, A. Bouland, B. Fefferman, J. Eisert, and J. Bermejo-Vega, Closing gaps of a quantum advantage with short-time Hamiltonian dynamics, Phys. Rev. Lett. **125**, 250501 (2020).

[40] Barbara M. Terhal and David P. DiVincenzo, Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games, Quantum Inf. Comput. **4**, 134 (2004).

[41] P. Erdős and A. Renyi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).

[42] Brendan McKay and Nicholas Wormald, Uniform generation of random regular graphs of moderate degree, J. Algorithms **11**, 52 (1990).

[43] Jeong Han Kim, Benny Sudakov, and Van Vu, Small subgraphs of random regular graphs, Discrete Math. **307**, 1961 (2007).

[44] Andrzej Ruciński, Induced Subgraphs in a Random Graph, in *North-Holland Mathematics Studies* (Elsevier, 1987), Vol. 144, pp. 275–296.

[45] Gregory Rosenthal, Efficient quantum state synthesis with one query, in *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (2024), pp. 2508–2534.

[46] Pu Gao, Mikhail Isaev, and Brendan D. McKay, in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (SIAM, 2020), p. 690.

[47] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits anticoncentrate in log depth, PRX Quantum **3**, 010333 (2022).

[48] Nicholas Hunter-Jones, Unitary designs from statistical mechanics in random quantum circuits, arXiv:1905.12053 [quant-ph].

[49] Tianci Zhou and Adam Nahum, Emergent statistical mechanics of entanglement in random unitary circuits, Phys. Rev. B **99**, 174205 (2019).

[50] Harm Derksen, Peter Ivanov, Chin Ho Lee, and Emanuele Viola, Pseudorandomness, symmetry, smoothing: I., ArXiv: 2405.13143.

[51] Béla Bollobás, Random Graphs, in *Modern Graph Theory. Graduate Texts in Mathematics* (Springer, New York, NY, 1998), Vol. 184, pp. 215–252.

[52] David Gross, Sepehr Nezami, and Michael Walter, Schur–Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations, Commun. Math. Phys. **385**, 1325 (2021).

[53] Jonas Helsen and Michael Walter, Thrifty shadow estimation: Reusing quantum circuits and bounding tails, Phys. Rev. Lett. **131**, 240602 (2023).

[54] Michel Talagrand, *Upper and Lower Bounds for Stochastic Processes* (Springer Cham, 2021), Vol. 60.

[55] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani, On the complexity and verification of quantum random circuit sampling, Nat. Phys. **15**, 159 (2018).

[56] L. G. Valiant, The complexity of computing the permanent, Theor. Comput. Sci. **8**, 189 (1979).

[57] H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest, Measurement-based quantum computation, Nat. Phys. **5**, 19 (2009).

[58] H. J. Briegel and R. Raussendorf, Persistent entanglement in arrays of interacting particles, Phys. Rev. Lett. **86**, 910 (2001).

[59] T. Chung, S. D. Bartlett, and A. C. Doherty, Characterizing measurement-based quantum gates in quantum many-body systems using correlation functions, Can. J. Phys. **87**, 219 (2009).

[60] A. C. Doherty and S. D. Bartlett, Identifying phases of quantum many-body systems that are universal for quantum computation, Phys. Rev. Lett. **103**, 020506 (2009).

[61] A. S. Darmawan, G. K. Brennen, and S. D. Bartlett, Measurement-based quantum computation in a two-dimensional phase of matter, New J. Phys. **14**, 013023 (2012).

[62] D. V. Else, S. D. Bartlett, and A. C. Doherty, Symmetry protection of measurement-based quantum computation in ground states, New J. Phys. **14**, 113016 (2012).

[63] D. V. Else, I. Schwarz, S. D. Bartlett, and A. C. Doherty, Symmetry-protected phases for measurement-based quantum computation, Phys. Rev. Lett. **108**, 240505 (2012).

[64] A. Miyake, Quantum computation on the edge of a symmetry-protected topological order, Phys. Rev. Lett. **105**, 040501 (2010).

[65] Robert Raussendorf, Wang Yang, and Arnab Adhikary, Measurement-based quantum computation in finite one-dimensional systems: String order implies computational power, Quantum **7**, 1215 (2023).

[66] Daniel E. Browne, Matthew B. Elliott, Steven T. Flammia, Seth T. Merkel, Akimasa Miyake, and Anthony J. Short, Phase transition of computational power in the resource states for one-way quantum computation, New J. Phys. **10**, 023010 (2008).

[67] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J. Bremner, John M. Martinis, and Hartmut Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[68] Dorit Aharonov, Xun Gao, Zeph Landau, Yunchao Liu, and Umesh Vazirani, A Polynomial-Time Classical Algorithm for Noisy Random Circuit Sampling, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*

(STOC 2023) (Association for Computing Machinery, New York, NY, USA, 2023), pp. 945–957.

[69] Xun Gao and Luming Duan, Efficient classical simulation of noisy quantum computation, arXiv:1810.03176 [quant-ph].

[70] Joel Rajakumar, James D. Watson, and Yi-Kai Liu, Polynomial-Time Classical Simulation of Noisy IQP Circuits with Constant Depth, in *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, p. 1037–1056.

[71] Alexander M. Dalzell, Nicholas Hunter-Jones, and Fernando G. S. L. Brandão, Random quantum circuits transform local noise into global white noise, Commun. Math. Phys. **405**, 78 (2024).

[72] A. Deshpande, P. Niroula, O. Shtanko, A. V. Gorshkov, B. Fefferman, and M. J. Gullans, Tight bounds on the convergence of noisy random circuits to the uniform distribution, PRX Quantum **3**, 040329 (2022).

[73] X. Gao, M. Kalinowski, C. N. Chou, M. D. Lukin, B. Barak, and S. Choi, Limitations of linear cross-entropy as a measure for quantum advantage, PRX Quantum **5**, 010334 (2024).

[74] Brayden Ware, Abhinav Deshpande, Dominik Hangleiter, Pradeep Niroula, Bill Fefferman, Alexey V. Gorshkov, and Michael J. Gullans, A sharp phase transition in linear cross-entropy benchmarking, arXiv:2305.04954 [quant-ph].

[75] Axel Dahlberg, Jonas Helsen, and Stephanie Wehner, The complexity of the vertex-minor problem, Inf. Process. Lett. **175**, 106222 (2022).

[76] Frank Arute, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[77] M. Kliesch and I. Roth, Theory of quantum system certification, PRX Quantum **2**, 010201 (2021).

[78] A. Morvan, *et al.*, Phase transitions in random circuit sampling, Nature **634**, 328 (2024).

[79] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, New York, 2010), 10th Anniversary ed.

[80] Lennart Bittel, Jens Eisert, Lorenzo Leone, Antonio A. Mele, and Salvatore F. E. Oliviero, A complete theory of the Clifford commutant, ArXiv:2504.12263.

[81] Marcel Hinsche, Zongbo Bao, Philippe van Dordrecht, Jens Eisert, Jop Briët, and Jonas Helsen, Clifford testing: Algorithms and lower bounds, ArXiv:2510.07164.

[82] Pu Gao and Nicholas Wormald, Uniform generation of random regular graphs, SIAM J. Comput. **46**, 1395 (2017).

[83] Naomi Kirshner and Alex Samorodnitsky, A moment ratio bound for polynomials and some extremal properties of Krawchouk polynomials and Hamming spheres, IEEE Trans. Inf. Theory **67**, 3509 (2021).

[84] Sean P. Meyn and Robert L. Tweedie, Computable bounds for geometric convergence rates of Markov chains, Ann. Appl. Probab. **4**, 981 (1994).

[85] Jeffrey S. Rosenthal, Minorization conditions and convergence rates for Markov chain Monte Carlo, J. Am. Stat. Assoc. **90**, 558 (1995).

[86] Lan Xiao, Guiying Yan, Yuwen Wu, and Wei Ren, Induced subgraph in random regular graph, J. Syst. Sci. Complex. **21**, 645 (2008).

[87] Mahdieh Hasheminezhad and Brendan D. McKay, Combinatorial estimates by the switching method, Contemp. Math. **531**, 209 (2010).

[88] Roman Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science* (Cambridge University Press, 2018), Vol.47.

[89] Ryota Tomioka and Taiji Suzuki, Spectral norm of random tensors, ArXiv:1407.1870.

[90] Patrick Hayden, Debbie Leung, Peter W. Shor, and Andreas Winter, Randomizing quantum states: Constructions and applications, Commun. Math. Phys. **250**, 371 (2004).

[91] Florence MacWilliams and Neil Sloane, *The Theory of Error-Correcting Codes* (Elsevier Science Publishers BV, 1977), Vol. 2, p. 39, ISBN: 978-0-444-85009-6.

[92] Note, however, that this equation in Ref. [91] contains a typographical error; we provide the correct formula.

[93] This is a trick to make the rest of the proof easier. This is also the reason we obtain the exponent $1/4$ instead of the expected $1/2$. We think the proof can be done without this trick (to obtain a tighter bound), but at the cost of a substantial increase in combinatorial complexity.

[94] D. Markham and B. C. Sanders, Graph states for quantum secret sharing, Phys. Rev. A **78**, 042309 (2008).

[95] Elham Kashefi and Anna Pappa, Multiparty delegated quantum computing, Cryptography **1**, 12 (2017).

[96] Pu Gao and Nicholas Wormald, Uniform generation of random regular graphs, ArXiv:1511.01175.

[97] Daniel E. Browne, Matthew B. Elliott, Steven T. Flammia, Seth T. Merkel, Akimasa Miyake, and Anthony J. Short, Phase transition of computational power in the resource states for one-way quantum computation, New J. Phys. **10**, 023010 (2008).

[98] K. Kieling, T. Rudolph, and J. Eisert, Percolation, renormalization, and quantum computing with non-deterministic gates, Phys. Rev. Lett. **99**, 130501 (2007).