Parameters of pseudorandom quantum circuits

Yaakov S. Weinstein,[1,*] Winton G. Brown,[2,†] and Lorenza Viola[2,‡]
[1]*Quantum Information Science Group, MITRE, 260 Industrial Way West, Eatontown, New Jersey 07224, USA*
[2]*Department of Physics and Astronomy, Dartmouth College, Hanover, New Hampshire 03755, USA*
(Received 26 August 2008; published 21 November 2008)

Pseudorandom circuits generate quantum states and unitary operators which are approximately distributed according to the unitarily invariant Haar measure. We explore how several design parameters affect the efficiency of pseudorandom circuits, with the goal of identifying relevant tradeoffs and optimizing convergence. The parameters we explore include the choice of single- and two-qubit gates, the topology of the underlying physical qubit architecture, the probabilistic application of two-qubit gates, as well as circuit size, initialization, and the effect of control constraints. Building on the equivalence between pseudorandom circuits and approximate *t*-designs, a Markov matrix approach is employed to analyze asymptotic convergence properties of pseudorandom second-order moments to a 2-design. Quantitative results on the convergence rate as a function of the circuit size are presented for qubit topologies with a sufficient degree of symmetry. Our results may be useful towards optimizing the efficiency of random state and operator generation.

DOI: 10.1103/PhysRevA.78.052332 PACS number(s): 03.67.Bg, 03.67.Lx, 05.40.−a

I. INTRODUCTION

Random pure states play a prominent role in quantum information processing. Random states are defined with respect to the unitarily invariant (so-called Fubini-Study) measure on the space of unit vectors in the Hilbert space of the system [1]. Not only do random states possess the remarkable feature of saturating the classical communication capacity of a noisy quantum channel [2], they also are an enabling resource for protocols including superdense coding of quantum states [3], approximate quantum encryption, and quantum data hiding [4,5]. Random quantum states can also be used for unbiased sampling, and the characterization of both "typical" bipartite and multipartite entanglement in random states has long been the focus of extensive investigation [6,7], recent results including estimates of moments of the subsystem purity distribution [8,9], relationships to state localization properties [10], and exact expressions for the probability distribution of entanglement measures such as *G*-concurrence [11] and purity [12].

How does one generate random quantum states? One way is to apply a random unitary transformation to initial computational basis states. Similar to random states, random unitary operators are drawn uniformly from the unique unitarily invariant measure—the Haar measure on the unitary group $\mathrm{U}(N)$, where N denotes the Hilbert space dimension, $N=2^n$ for a multipartite n-qubit system [13]. Random unitaries themselves are useful for a number of quantum protocols ranging from remote state preparation [14] to efficient error characterization [15,16] and selective process tomography [17]. Unfortunately, implementing random unitaries as a sequence of one- and two-qubit gates on a quantum computer is inefficient: the required number of quantum gates grows quadratically with N, that is, exponentially with the number of qubits, n. Therefore, researchers are left with the challenge of constructing suitable pseudorandom (PR) substitutes.

A promising approach to efficiently implement an ensemble of unitaries so that the resulting moments approximate those induced by the Haar measure is provided by *PR circuits*, introduced in Ref. [15]. These circuits consist of an iterated set of single- and two-qubit gates where certain specifications are chosen at random. As the set of gates is iterated (using different single-qubit gates for each qubit and at each time step), the statistical properties of the implemented unitary operators and the resulting output states compare more and more favorably to the properties of random unitaries and states. Additional studies of PR circuits have focused on analyzing both analytically and numerically the convergence to the desired distribution [18–20], identifying optimal two-qubit gates, as well as elucidating some aspects related to the influence of qubit topology [21], and quantifying the ability to efficiently generate states with generic entanglement [22]. PR algorithms have also been formulated for cluster-state quantum computation (QC) in [23], allowing an optimal single-gate distribution to be identified. A yet different realization via local measurements on weighted graph states has been proposed in [24].

Our goal in this paper is to quantitatively investigate a number of parameters that affect the convergence of PR circuits to the Haar distribution. After providing the necessary background on the mathematical framework employed to characterize PR behavior in Sec. II, we proceed in Sec. III to assess the influence of three main design parameters in a PR circuit on a fixed number of qubits: the choice of single- and two-qubit gate distributions; the influence of different qubit topologies; and the effect of probabilistic versus deterministic two-qubit gates. Not surprisingly, these parameters are intertwined, causing the optimal choice for any one parameter to depend on one or several other parameters in complex ways. The influence of circuit size is addressed in Sec. IV. Remarkably, explicit scaling predictions turn out to be possible based on simple expressions which are consistent with existing numerical evidence and analytical results. Sections V and VI are devoted to analyze the effect of limited (non-

*weinstein@mitre.org
†Winton.G.Brown@Dartmouth.EDU
‡Lorenza.Viola@Dartmouth.EDU

selective) control and of different initial states, respectively. In Sec. VII we make some final remarks and conclude. Additional considerations on optimizing cluster-state PR circuits are included in Appendix A, whereas Appendix B discusses a proposed convergence improvement in the specific yet important case of a PR circuit implementing an approximate Clifford twirl.

II. QUANTIFYING PSEUDORANDOMNESS

We begin by discussing possible ways of quantifying the distance between the ensemble generated by PR circuits and the Haar-distributed ensemble. It is important to realize that any probability distribution over n-qubit quantum states or unitary transformations requires a number of parameters exponentially growing with n to specify. Therefore, it is impractical to gauge how well an ensemble of quantum states or unitary transformations resembles the uniform Haar ensemble based on a full characterization of the distribution.

A. t-designs

For several tasks which utilize random states and unitary transformations, the details of the full distribution are not relevant, in the sense that only statistical moments up to a *finite* order of the ensemble from which the random states or unitaries are drawn need to coincide with Haar-induced moments. That is, it suffices that the relevant probability distribution be indistinguishable from the uniform Haar distribution as long as a finite number of moments is given. This is captured by the concept of a *quantum t-design* [25,26]. Formally, an ensemble of states, $[p(\alpha)d\alpha, |\psi(\alpha)\rangle]$, or, respectively, unitary transformations, $[p(\alpha)d\alpha, U(\alpha)]$, is a state (unitary) t-design if for any polynomial f of order (t,t) of the state-vector components (or matrix elements),

$$\int p(\alpha) f(|\psi(\alpha)\rangle) d\alpha = \int f(|\psi(\alpha)\rangle) d\alpha,$$

or

$$\int p(\alpha) f(U(\alpha)) d\alpha = \int f(U(\alpha)) d\alpha,$$

where integration is over the invariant measure (Fubini-Study or Haar, respectively). A (t,t) polynomial is a linear combination of terms consisting of products of up to t variables and their t complex conjugates.

In order to assess how well a PR circuit approximates a t-design, an appropriate norm must be defined on the space of (t,t) polynomials describing moments of the circuit. We note that the concept of an ϵ-approximate t-design for quantum states was introduced in [25] as an ensemble of states $[p(\alpha)d\alpha, |\psi(\alpha)\rangle]$ satisfying

$$(1-\epsilon)\int f(|\psi(\alpha)\rangle) d\alpha \leqslant \int p(\alpha) f(|\psi(\alpha)\rangle) d\alpha \leqslant (1+\epsilon)\int f(|\psi(\alpha)\rangle) d\alpha,$$

for all (t,t) polynomials f. This notion induces a norm on the space of all (t,t) polynomials (specifically the l_∞ norm), which can quantify the distance of a distribution approximating an exact t-design. In our study, we shall focus on asymptotic convergence *rates* of arbitrary (2,2) polynomials to their expected value under the Haar measure as a function of PR circuit depth. Note that Ref. [27] has previously shown that a large class of PR circuits are efficient approximate unitary 2-designs, with respect to the diamond norm.

B. Markov chain analysis

The case $t=2$ is especially relevant, since it includes the majority of known protocols which utilize random states. Notable examples include unbiased noise estimation [16,17], and the generation of states with typical entanglement [22]. It is thus important to determine the behavior of (2,2) polynomials (second moments) of the state components under the action of PR circuits. As shown in [22], the evolution of the second moments can be mapped to an appropriately defined classical Markov chain. Thus, convergence properties of the PR circuit to a 2-design may be directly established by exploiting properties of the corresponding Markov chain.

To obtain the mapping, the density operator describing the pure quantum state being evolved by the PR circuit is written in the Pauli basis

$$\rho = |\psi\rangle\langle\psi| = \sum_\nu c_\nu P_\nu,$$

where $P_\nu = \sigma_1^{\nu_1} \otimes \ldots \otimes \sigma_n^{\nu_n}$ is a tensor-product string of single-qubit identity and Pauli operators, specified by the collective index $\nu \in \mathcal{I} = \{0,x,y,z\}^n$. Let $\mathrm{PR}(\ell)$ be the family of PR circuits of depth ℓ. We shall be interested in the evolution of the second order moments $\{\mathbb{E}_{\mathrm{PR}(\ell)}(c_{\nu,\ell}c_{\mu,\ell})\}$ as a function of the depth ℓ of the PR circuit. We will show by construction that the moments $\{\mathbb{E}_{\mathrm{PR}(\ell)}(c_{\nu,\ell}^2)\}$ follow a discrete Markov chain on $\mathcal{I}$. That is, the evolution of $\mathbb{E}_{\mathrm{PR}(\ell)}(c_{\nu,\ell}^2)$ satisfies

$$\mathbb{E}_{\mathrm{PR}(\ell+1)}(c_{\nu,\ell+1}^2) = \sum_{\nu\in\mathcal{I}} M_{\mu\nu}\mathbb{E}_{\mathrm{PR}(\ell)}(c_{\nu,\ell}^2) = \sum_{\nu\in\mathcal{I}} M_{\mu\nu}^\ell c_{\nu,0}^2,$$

where $M=\{M_{\mu\nu}\}$ is a Markov matrix. For the random circuits examined here, the remaining second moments $\{\mathbb{E}_{\mathrm{PR}(\ell)}(c_{\nu,\ell}c_{\mu,\ell})\}$, with $\nu\neq\mu$, vanish for $\ell>0$.

The class of PR circuits we are interested in have the following structure: In a single time step, local (single-qubit) gates are applied to each qubit in parallel, followed by commuting two-qubit gates belonging to the Clifford group between a specified set of neighboring qubits. Typically, the two-qubit gates will be conditional phase gates $\mathrm{CZ}=|0\rangle_1\langle 0|_1 \otimes \mathbb{1}_2 + |1\rangle_1\langle 1|_1 \otimes \sigma_z^2$. Under a single-qubit gate, each nontrivial Pauli operator transforms as

$$\sigma_a \mapsto R(\sigma_a) = \sum_{a,b} x_{ab}\sigma_b, \quad a,b \in \{x,y,z\},$$

where $R=\{x_{ab}\} \in \mathrm{SO}(3)$ depends on the applied rotation. We will proceed by determining the transformation matrix, $\bar{R}$, which describes the evolution of second-order moments of local Pauli operators (including the identity), and from this construct the transformation matrix describing the evolution

of second-order moments on the full multiqubit space. We will restrict to local gate distributions obeying

$$\mathbb{E}(x_{ca}x_{cb}) = 0, \quad \forall\, a,b,c.$$

Then $\bar{R}$ is determined by the remaining nonzero moments, specifically, $\bar{R} \equiv \mathbb{1} \oplus \mathbb{E}(x_{ab}^2)$. That $\bar{R}$ is a Markov matrix follows from the normalization condition $\Sigma_b x_{ab}^2 = 1$ on the columns of R. Assuming that each local gate is *selected independently* from the same distribution (see Sec. V for a different setting), the transformation resulting from simultaneous single-qubit rotations within a PR iteration is the n-fold tensor product of these single-qubit transformations, $L = \bar{R}^{\otimes n}$. Since each two-qubit gate is a member of the Clifford group, it simply acts (up to irrelevant phases) as a permutation on the columns of L. Thus, the full 4^n-dimensional transformation is a Markov matrix provided that $\bar{R}$ is.

Once the mapping is established, basic properties of Markov chains may be used to analyze the evolution of the second moments of the PR circuit. Specifically, if the Markov chain is *ergodic*, then the corresponding PR circuit converges to a 2-design. A Markov chain is ergodic if it is *irreducible* and *aperiodic*. That is, every state must be reached from every other state and the recurrence times to any state must not be multiples of some period $k > 1$. A consequence of ergodicity is that for any initial distribution v_o over $\mathcal{I}$, there exists a unique asymptotic distribution, v_∞, towards which the chain evolves. To illustrate this, we may expand the initial state in terms of the (linearly independent) eigenvectors $\{e_n\}$ of the Markov matrix, $v_\tau = M^\tau v_o = \Sigma_n \lambda_n^\tau c_n e_n$. For an ergodic chain, there is only one eigenvalue equal to 1, whereas all other eigenvalues have magnitude less than 1. Thus, the contribution to v_τ from each eigenvector (other than the ergodic eigenvector that is equal to 1) decays to zero exponentially at a rate governed by the corresponding eigenvalue. For times $\tau \gg \ln(\lambda_k/\lambda_1)$, the contribution from the subdominant eigenvalue, λ_1, will be the dominant nonergodic contribution, leaving

$$\|v_\tau - v_\infty\|_\infty \sim e^{-\Gamma\tau}, \tag{1}$$

where $\Gamma = -\ln(\lambda_1)$, and $\Delta = 1 - \lambda_1$ is the spectral gap of the Markov chain. Since any (2,2) polynomial may be expressed in terms of the moments $\{\mathbb{E}_{\mathrm{PR}(\ell)}(c_{\nu,\ell}c_{\mu,\ell})\}$ it follows that at sufficiently long times, τ, the difference between the expected value of any (2,2) polynomial of the state vector components over the t-design distribution and the Haar-induced expected value obeys

$$|f_{2\text{-design}}(\tau) - f_{\mathrm{Haar}}(\tau)| \sim e^{-\Gamma\tau}. \tag{2}$$

As we shall see, certain entanglement measures which serve as useful test functions for PR circuits are expressible as (2,2) polynomials and will thus converge exponentially to their Haar-expected value at a rate of Γ.

A Markov chain may be reduced by identifying a partition of the state space, $\mathcal{I}$, into subsets, $\mathcal{J}_u \subset \mathcal{I}$, such that the coarse-grained probability distribution obtained by summing over each subset follows a reduced Markov chain M'. A necessary and sufficient condition for reducing the chain is that the sum of the transition probabilities from a member of a subset, $\mu \in \mathcal{J}_u$, to all members of any subset, $\nu \in \mathcal{J}_v$, is the same for each member μ, that is, $\Sigma_{\mu,\nu\in\mathcal{J}_v} M_{\mu\nu} = M'_{uv}$ must be the same for all $u \in \mathcal{J}_u$ and all subsets $\mathcal{J}_v$. It is important to note that each eigenvalue of M' is an eigenvalue of M.

We employ a reduced representation by averaging over local X and Y Pauli operators. Since conditional phase (CZ) gates are invariant under rotations along the z axis, we may restrict our choice of single-qubit gate distributions to those which initially randomize states in the x-y plane. Now let P be a Pauli string containing at least one X_i or Y_i, and let P' be any string obtained from P by permuting X_i with Y_i. Since M randomizes X_i and Y_i, $M(P - P') = 0$. This defines the kernel of M, which may be removed by defining new variables $\Xi_i^\pm \equiv X_i \pm Y_i$. Chain states including Ξ_i^- may be discarded, whereas transitions within $\mathcal{I}' = \{0, z, \xi\}^n$, where ξ stands for either x or y, are described by M'. There is now only one parameter left to characterize the local gate distribution. Let $c \in [0,1]$ parametrize the extent to which the z axis is left invariant. We shall refer to c as the *local gate parameter* henceforth. The single-qubit gate contribution to M' is then fully described by

$$\bar{R}(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & \frac{1-c}{2} \\ 0 & 1-c & \frac{1+c}{2} \end{pmatrix}.$$

Haar-distributed rotations in SU(2) correspond to $c = 1/3$. So-called *HZ gates*, single-qubit Hadamard gates followed by a random rotation about the z axis, correspond to $c = 0$. The latter were identified in [23] as optimal single-qubit gates when using two-qubit CZ gates.

In some instances, it is possible to further reduce the Markov chain by taking into account qubit-permutation symmetries in the construction of PR circuits (see Sec. IV). The corresponding Markov chain will admit a reduced chain by forming equivalence classes of Pauli strings under symmetry operations. In the case of *full permutation symmetry*, equivalence classes may be labeled by the number of X's and Z's in each Pauli string. Accordingly, the size of the reduced state space grows only quadratically in the number of qubits n, allowing scaling behavior of Markov matrix properties with circuit size to be determined numerically. Note that while the reduced representation enabled by the initial randomization of x and y contains all of the nonzero eigenvalues of the full Markov chain, reduced representations induced by qubit permutations discard eigenvalues associated with asymmetry with respect to the permutations. Thus, the behavior of test functions which do not possess the symmetry cannot be predicted if the initial state does not also share the symmetry which permits the reduction.

C. Entanglement measures

As remarked, entanglement properties of random pure states have been extensively studied. In particular, bipartite-entanglement across a partition is known to be nearly maximal for random states. As observed in [22], linear entropy, a

measure of bipartite entanglement, may be expressed as a (2,2) polynomial in the state components, allowing the rate at which bipartite entanglement approaches its expected Haar value to be determined by Markov analysis. Both features—that the asymptotic value of the entanglement is nearly maximal, and that the simplest measure of the entanglement is a (2,2) polynomial in the state vector components—are not limited to bipartite entanglement but are representative of a more general behavior which may be appreciated by invoking the *generalized entanglement* (GE) approach [28].

The basic idea of GE is to abstract the notion of entanglement from a preferred subsystem decomposition and instead tie it to a distinguished set of observables to which one has access. Loosely speaking, a pure state $|\psi\rangle$ is defined as "generalized unentangled" relative to a distinguished set of observables, h, depending on whether or not the "reduced state" with respect to those observables is pure (extremal), and generalized entangled otherwise. Under appropriate mathematical assumptions on h [29], the simplest way to quantify GE is by taking the square length of the projection of $|\psi\rangle\langle\psi|$ onto h. Formally, this yields the so-called h purity, which may be computed as $P_h(|\psi\rangle)=\kappa\Sigma_i\langle\psi|A_i|\psi\rangle^2$, where $\kappa>0$ is a normalization constant and $\{A_i\}$ is an orthonormal (with respect to the trace norm) basis of h. Since GE relative to h may be naturally quantified as $\mathrm{GE}_h=1-P_h$, this results in a wide class of functionals which are *quadratic* in the state-vector components, hence describable by Markov analysis. The global entanglement measure of Meyer and Wallach is the simplest GE representative in the conventional multipartite setting [28–31],

$$Q(|\psi\rangle)=2-\frac{2}{n}\sum_{j=1}^{n}\mathrm{Tr}[\rho_j^2], \tag{3}$$

where ρ_j is the reduced density matrix of qubit j. The expectation of Q over random pure states is [8,10]

$$\langle Q_R\rangle=\frac{N-2}{N+1}, \tag{4}$$

the difference between this value and the average entanglement of PR output states, $|\langle Q\rangle-\langle Q_R\rangle|$, providing a useful indicator for asymptotic convergence analysis. While restricting to the average global entanglement in our present numerical studies is motivated by the possibility to make direct comparison with previous work [15], we believe it is conceptually important to highlight the broader context that these results exemplify.

D. State-vector element distribution

The other method we use to assess the quality of our PR circuits is based on the distribution of state vector probability in the computational basis. Let c_k^l denote the kth component of the lth state randomly chosen from the set of all pure states, with $\eta=|c_k^l|^2$ being the corresponding probability. The distribution of η for random states is given by the Porter-Thomas (PT) distribution [1,32],

$$\tilde{P}(\eta)=(N-1)(1-\eta)^{N-2},$$

where N is the Hilbert space dimension. In the limit $N\to\infty$, and upon rescaling to unit mean, the distribution becomes

$$P_{PT}(y)=e^{-y}, \quad y=N\eta.$$

As our second randomness indicator, we shall examine the l_2 distance (simply denoted by $\|\cdot\|_2\equiv|\cdot|$ henceforth) between the distribution of output probabilities at each PR iteration and the random (PT) distribution. Note that because this test function depends on higher order moments of the PR distribution it is not *a priori* described by the above-described Markov chain analysis and thus can yield additional insight.

III. OPTIMIZING PSEUDO-RANDOM CIRCUITS

Any universal set of gates used in a PR circuit will eventually generate Haar-distributed states and operators [18]. Our goal here is to explore how various design parameters affect PR circuits, in an effort to boost efficiency by optimizing time and qubit resources. By focusing on the network model of QC, we first characterize the optimal choice of single- and two-qubit gates for an open-chain topology as considered in the "standard" PR architecture of Ref. [15]. The effect of different qubit topologies and probabilistic gate application is addressed next, whereas, for clarity, a similar analysis for cluster-state QC is sketched in Appendix A.

A. Optimal gates for standard circuits

PR circuits were introduced in the context of liquid-state nuclear magnetic resonance quantum information processing. In that context it was natural to look at an open chain of qubits and to use two-qubit Ising (ZZ) gates between nearest neighbor (NN) qubits,

$$ZZ=\exp\left[-i\frac{\pi}{4}\sigma_z^j\sigma_z^{j+1}\right],$$

as the relevant nonlocal resource. Subsequently, other two-qubit gates have been suggested for use within PR algorithms [22]. A comprehensive study was recently undertaken [21] to determine the two-qubit gate for which a PR circuit with random single-qubit gates will converge most quickly to the Haar expected value of bipartite entanglement as quantified by linear entropy. Recall that this is essentially equivalent to determining convergence rates of second order moments of the PR circuit. The protocol studied calls for a single two-qubit gate to be applied per time interval to pairs of qubits which are selected at random from all allowable neighboring pairs according to either a closed or open chain [33]. It was determined that the so-called *XY gate* is the two-qubit gate which leads to the most efficient PR circuit of this type. The XY gate between qubits j and k is given by

$$XY=\exp\left[-i\frac{\pi}{4}(\sigma_x^j\sigma_x^k+\sigma_y^j\sigma_y^k)\right].$$

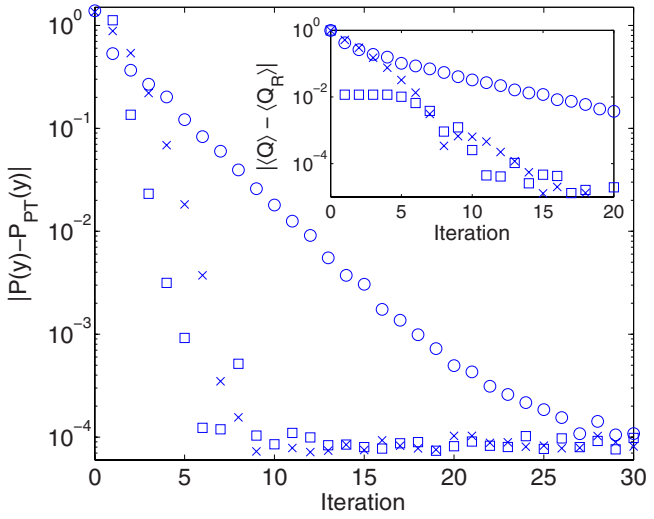

FIG. 1. (Color online) l_2 distance between PR distribution of squared moduli of components in the computational basis and the PT distribution (inset: distance of average global entanglement from random value) for CZ (○) and XY (×) two-qubit gates using single-qubit random rotations, and CZ gates using single-qubit HZ gates (□) (8 qubits, 100 implementations, all computational basis states). When using random single-qubit gates, both indicators come within 10^{-4} much more quickly for XY two-qubit gates than for CZ gates. Modifying the CZ-based algorithm to use single-qubit HZ gates boosts the convergence rate so much that it even outperforms the XY-based algorithm with random single-qubit rotations.

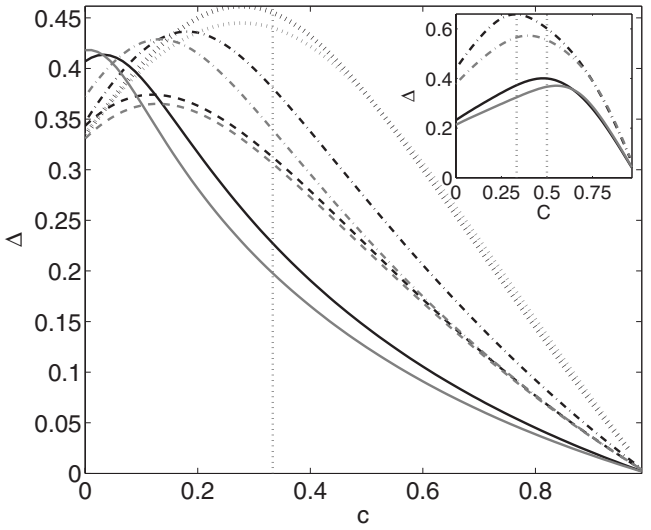

FIG. 2. Markov spectral gap, Δ, versus local gate parameter, c, for various qubit topologies using two-qubit CZ (main panel) and XY gates (inset). Dark lines: $n=6$; light lines: $n=8$. For CZ gates, the open chain topology (solid lines) has a larger gap than any of the other topologies for single-qubit HZ gates, $c=0$, but the smallest gap for single-qubit random gates, $c=1/3$ (vertical dotted line). The HZ gate is the optimal single-qubit gate for the open chain topology. The AA topology (…) has the smallest or second smallest gap (depending on the number of qubits) for HZ gates but the largest for random gates. Random gates are close to optimal single-qubit rotations for AA topology. The closed-chain topology (–·) has the second largest gap for both random rotations and HZ gates. The optimal single-qubit rotation for this topology would be $c\approx 0.18$. The star topology (dashed line) has the second smallest gap for both random and HZ single-qubit rotations and the optimal single-qubit gate (for eight qubits) is also $c\approx 0.18$. For random single-qubit gates, going from 6 to 8 qubits decreases the size of the gap for all topologies. For single-qubit HZ gates, the gap increases when going from 6 to 8 qubits for the open- and closed-chain topologies but decreases for the star topology. Inset: Δ versus c for two iterations (such that all couplings have been utilized) of PR algorithms using two-qubit XY gates. The closed-chain topology (–·) has a significantly larger gap than the open chain (solid line). The optimal single-qubit gate for eight qubits is the random rotation, $c\approx 1/3$, for the closed chain, but $c\approx 1/2$ for the open chain. Note the smaller gap (hence slower convergence rate) of $n=8$ as compared to $n=6$.

However, the use of only a single two-qubit gate per time interval introduces an unnecessary inefficiency since qubits not involved in the gate do nothing for that time interval and, in any experimental implementation, must be protected from error.

We wish to examine to what extent the number of applied gates can be traded for the number of time steps by applying gates in parallel. One cannot apply all NN XY gates during one time interval since they do not commute. Nevertheless, a more efficient PR circuit using XY gates can be obtained if, at each time interval, half the NN XY gates are applied. Figure 1 compares the state vector element distribution distance and average entanglement distance as a function of iteration for a PR algorithm which uses conditional phase (CZ) gates (which are ZZ gates to within single-qubit z rotations) between all NN qubits, and a PR algorithm which uses XY gates between half the NN pairs at each iteration (say, 1-2, 3-4, etc. at odd iterations and 2-3, 4-5, etc. at even iterations). In both cases, the single-qubit rotations are drawn randomly from SU(2). It takes significantly more time for the CZ-based PR algorithm to be within 10^{-4} of typical entanglement values and PT distribution than the XY-based PR algorithm, despite the fact that all NN CZ gates are applied at every iteration.

In all the work mentioned above, single-qubit rotations were taken to be random with respect to the Haar measure on SU(2). Based on studies of cluster-state PR algorithms, however, we have recently shown [23] that when using two-qubit CZ gates a *restricted set of single-qubit rotations*, HZ gates, allows for faster convergence with respect to test functions which are second-order polynomials. This is confirmed through the Markov analysis reported in Fig. 2 which shows that, for sufficiently large n, the optimal single-qubit gates to use in conjunction with the CZ gate is indeed at $c=0$, the HZ gate. Using these restricted random gates we see in Fig. 1 that the CZ-based PR circuit converges much faster than the CZ-based circuit with random SU(2) single-qubit gates.

What single-qubit gate distribution is optimal for use with XY gates? From the inset of Fig. 2, we see that the answer depends on the qubit topology. For an open chain, as in standard PR circuits, the optimal single-qubit gate is at $c\simeq 0.5$. For a closed-chain topology, SU(2) random single-qubit rotations appear to be optimal.

B. Effect of different qubit topologies

1. Modeling strategies and preliminaries

As mentioned, a standard PR circuit [15] applies two-qubit gates between NN qubits on an open chain. Recent studies of PR algorithms have implemented two-qubit gates between NN qubits on a closed chain (i.e., subject to periodic boundary conditions) [21], or between randomly chosen qubit pairs [22,34]. This alters the convergence rate of the algorithm by effectively changing the qubit topology. In this section, we investigate the convergence rate of the following alternate topologies which employ two-qubit CZ gates: (i) NN qubits on a closed chain; (ii) a star formation where each qubit is connected to a central one; (iii) an all-to-all (AA) topology where CZ gates are performed between every possible pairing of qubits (note that all such gates commute and can thus be performed simultaneously). Our strategy is to rely on Markov chain analysis to explore how different single-qubit gate distributions affect the convergence rate for different topologies, thereby identifying the optimal single-qubit gate for each topology. For PR circuits employing *XY* gates, open- and closed-chain NN topologies are contrasted, as a function of c.

The circuits we address in this section aim at *maximizing parallelism*, by applying as many CZ gates as possible per iteration. In this case, the closed-chain topology is *not* ergodic over the entire state space. This can be seen by noticing that each qubit undergoes exactly two CZ gates per iteration (one with each NN). Under the action of pairs of CZ gates, the number of nonidentity single Pauli operators in the Pauli string describing the state will remain even or odd. Within each of these fixed-parity subspaces, the closed-chain topology is ergodic and the PR circuit will induce exponential convergence to the corresponding ergodic state, with a rate calculated from the gap of the Markov matrix. The asymptotic entanglement for an n-qubit closed chain is also affected by the presence of the two distinct subspaces. Specifically, the total probability that the state of the Markov chain is in each subspace remains constant, and is determined by the projection of the initial state into each subspace. Under the action of the Markov chain, whatever total probability is found in each subspace is uniformly distributed within each subspace, allowing for the asymptotic closed-chain (cc) global entanglement to be calculated. For initial computational basis states, we find

$$Q_{\mathrm{cc}} = 1 - 3\frac{\sum_{k\ \mathrm{odd}}^{n}\binom{n}{k}}{\sum_{k\ \mathrm{odd}}^{n} 3^{k}\binom{n}{k}}. \tag{5}$$

For $n=8$, $Q_{\mathrm{cc}}=0.988\,235$, whereas the expected global entanglement value for the Haar distribution [from Eq. (4)] is given by $Q_R=0.988\,327$. Therefore, when we compare the size of the gap (hence the convergence rate) for the different topologies it is important to remember that the closed chain is converging to a slightly different asymptotic state than the other topologies.

Note that the same conservation of even or odd number of nonidentity Pauli operators occurs in the AA topology for odd n. This is because each qubit is involved in an even number of CZ gates. In this case, however, initial computational basis states evenly populate the two subspaces, therefore the final state is the ergodic state.

2. Numerical results

Figure 2 summarizes the behavior of the Markov spectral gap, Δ, for each topology as a function of c. The optimal single-qubit gate depends both on the topology and number of qubits. A certain single-qubit rotation may be exceedingly good for one topology, but extremely poor for another topology. Increasing the number of qubits affects the convergence of different topologies in different ways: For certain topologies, an increase in the number of qubits causes the gap to decrease, while for other topologies the gap increases for certain single-qubit rotations but decreases for other single-qubit rotations.

We first address the effect of topology on PR algorithms using random single-qubit rotations, $c=1/3$. The Markov analysis for $n=8$ suggests that the AA topology should converge the fastest, followed by the closed-chain topology (which, as mentioned above, converges to a different state). The fast convergence of the AA topology is not surprising given that this topology allows for simultaneous interaction between all qubit pairs. The open chain has the slowest convergence rate. The star topology is slower than the closed chain, despite the fact that each qubit is only two couplings away from every other qubit. The need to traverse the central qubit is likely to be the bottleneck to the spread of entanglement. This conjecture is supported by our later explorations of probabilistic application of two-qubit gates. For all topologies, increasing n from 6 to 8 decreases the gap—meaning that it takes longer for the algorithm to converge to random states.

If random single-qubit gates are replaced by HZ gates, $c=0$, the open-chain topology exhibits the largest gap, followed by the closed chain. In both cases, the convergence rate is faster than that of random single-qubit gates, and grows when going from 6 to 8 qubits. The HZ gates also provide better PR algorithm convergence for the star topology, though the rate is smaller for $n=8$ than $n=6$ in this case. Such a decrease in convergence when going from 6 to 8 qubits is more pronounced for single-qubit HZ gates than for random ones. Among the topologies we have examined, the AA topology is the only one which demonstrates better behavior for random single-qubit rotations than for the HZ gates.

The convergence to random state-vector element distribution and entanglement for PR circuits using random and HZ single-qubit gates is illustrated in Fig. 3 and Fig. 4, respectively. For both quantities in both cases, the results agree with the predictions based on Markov analysis. In particular, one sees that when using HZ gates, the AA and star topologies become the slowest, and the open- and closed-chain topologies have similar convergence rates. Also notice that both the open- and closed-chain topologies exhibit the entanglement "cutoff phenomenon" described in [23]. For the

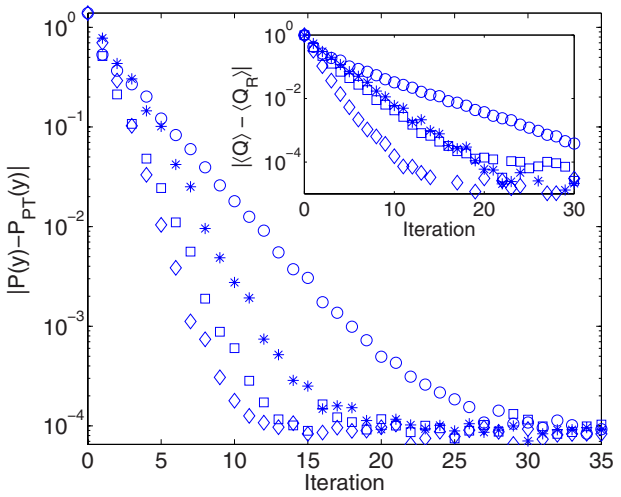

FIG. 3. (Color online) l_2 distance between PR distribution of squared moduli of components in the computational basis and the PT distribution (inset: distance of average global entanglement from random value) for different qubit topologies using CZ two-qubit gates and random single-qubit rotations (8 qubits, 100 implementations, all computational basis states). The qubits have been arranged in an open chain (○), a star formation (∗), a closed chain (□), and in such a way that all the qubits are connected to each other (◇). The convergence rate as a function of topology agrees with Markov analysis, see Fig. 2.

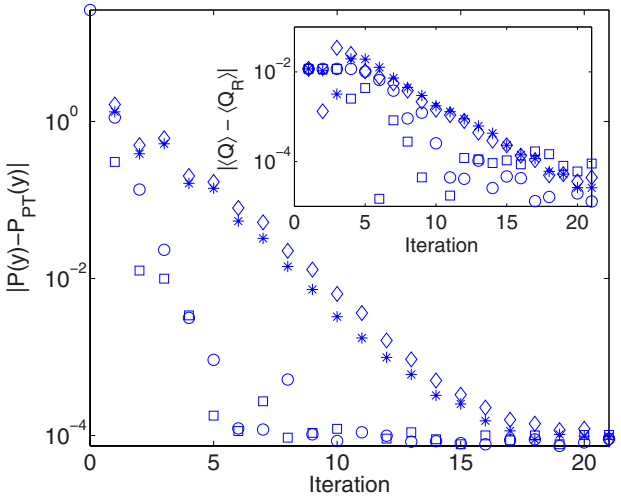

FIG. 4. (Color online) l_2 distance between PR distribution of squared moduli of components in the computational basis and the PT distribution (inset: distance of the average global entanglement from the random value) for different gate topologies using CZ gates and single-qubit HZ gates (eight qubits, 100 implementations, all computational basis states). The qubits have been arranged in an open chain (○), closed chain (□), a star formation (∗), and in such a way that all the qubits are connected to each other (◇). The decay rates of the star and AA topologies are now slower than that of the open and closed chain, as expected from the Markov analysis, see Fig. 2.

open (closed) chain, every realization of the PR algorithm for $n/2$ $(n/2-1)$ iterations is maximally entangled. Only after this point does exponential decay to random entanglement values set in [35]. Results on the convergence of output state element distribution and entanglement for PR circuits employing XY gates are given in Fig. 5. As above, in this case the PR circuit implements every other NN coupling for one iteration. Similar to CZ gates, the closed-chain topology does not result in convergence to the random state. Nevertheless, the convergence rate of the closed chain to random entanglement values and element distributions is faster than that of the open chain.

In summary, the optimal single-qubit distribution for each of the topologies exploiting two-qubit CZ gates is as follows: for the open chain, the HZ gates are optimal, $c=0$. Remarkably, these gates are naturally motivated by cluster-state QC [23]. For the AA topology, the random single-qubit rotation is optimal, whereas for the closed-chain and star topologies the optimal single-qubit rotations correspond to $c\approx 0.18$. For chain topologies, the optimal c value is also dependent on the number of qubits. If two-qubit XY gates are used, gate distributions with $c\approx 1/2$ and $c\approx 1/3$ are found to be optimal for open and closed chains, respectively.

C. Probabilistic gates

When PR algorithms employ two-qubit CZ gates, commutation allows any number of such gates to be performed in parallel. Therefore, until this point, we have constructed algorithms that maximize the number of possible CZ gates based on the qubit topology per iteration. We proceed to explore what happens if each CZ gates is applied probabilistically, that is, with probability p. As we shall see, for certain topologies lowering p *increases* the convergence rate of the PR circuit.

We begin by considering a probabilistic open chain topology. As illustrated in Fig. 6, the Markov spectral gap (hence

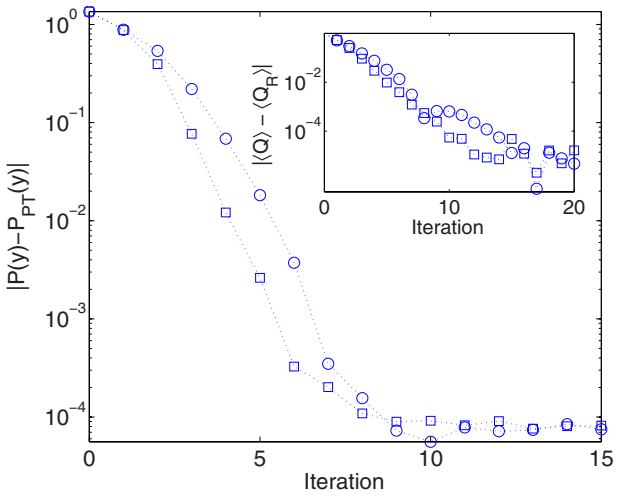

FIG. 5. (Color online) l_2 distance between PR distribution of squared moduli of components in the computational basis and the PT distribution (inset: distance of average global entanglement from random value) for different gate topologies using XY two-qubit gates and random single-qubit rotations (eight qubits, 100 implementations, all computational basis states). The qubits have been arranged in an open chain (○) and closed chain (□). As expected, the convergence rate of the closed chain is faster than that of the open chain.

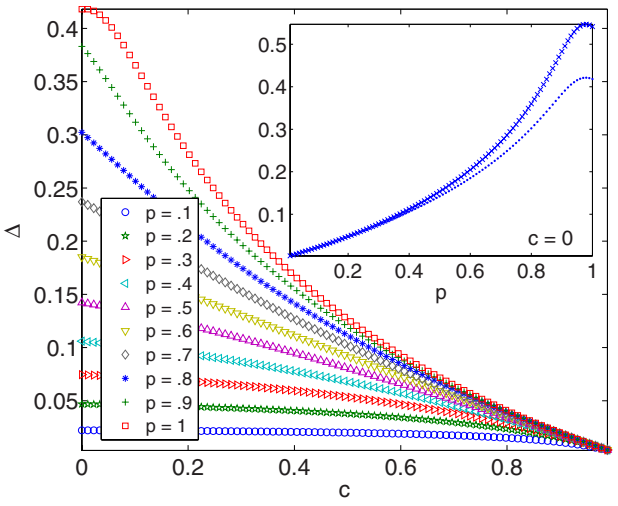

FIG. 6. (Color online) Markov spectral gap, Δ, versus local gate parameter, c, for an eight-qubit open-chain topology for different probabilities, p, that each CZ gate will be implemented. The gap increases with p for all values of c. Inset: Gap $(\cdot)$ and convergence rate $(\times)$ for HZ gates as a function of p. The HZ gate $(c=0)$ is the optimal single-qubit gate for all p until $p=0.98$.

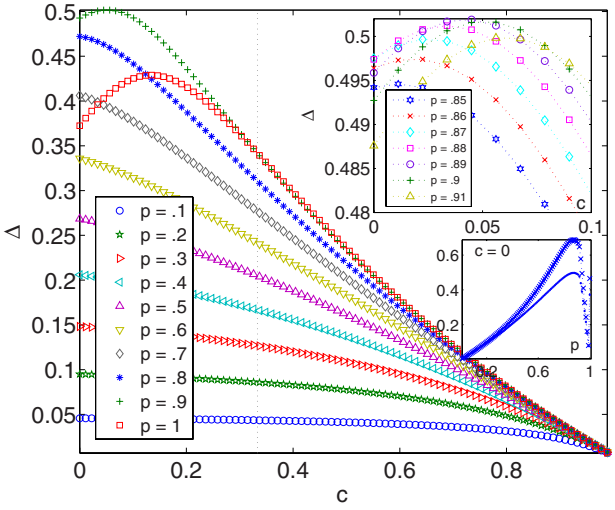

FIG. 7. (Color online) Markov spectral gap, Δ, versus local gate parameter, c, for an eight-qubit closed-chain topology with probability p that each CZ gate will be implemented. The HZ gate is optimal until $p \simeq 0.85$. Above that, a hump grows towards greater c with increasing p. The height of the hump (size of the gap) reaches a maximum at $p \simeq 0.89$, $c \simeq 4/90$ (see upper inset). As p continues to grow, the peak of the hump moves to higher values of c but now Δ shrinks. Lower inset: Gap $(\cdot)$ and convergence rate $(\times)$ for HZ gates as a function of p. For $0.9 < p < 1$, the decreased gap size is due to an eigenvalue representing coupling between the two parity subspaces of the system (see text) and thus is not reliable to fully determine the actual convergence rate of the circuit.

the convergence rate) increases as p is varied over $[0,1]$, $p=1$ recovering the deterministic case. The HZ single-qubit gate turns out to be optimal for almost all p, the improvement in the convergence rate being steady until p is close to unity. We expect this since HZ gates are the "maximally noninvariant" single-qubit gates with respect to the z axis, which is preferred by CZ gates. As $p \to 1$, a rounded hump is visible whose peak value for $n=8$ is found at $c=0.02$ but whose roundedness allows for a range of c values giving nearly optimal single-qubit gates. As the number of qubits increases, the peak of the hump shifts towards $c=0$, indicating that the nonoptimality of the HZ gates is due to the relatively small Hilbert space dimension. The gap size and convergence rate of PR algorithms using HZ gates as a function of p are explicitly shown in the inset of Fig. 6.

The behavior of a probabilistic closed-chain topology is shown in Fig. 7. Initially, as p increases so does the gap and the rate of convergence to random. The HZ single-qubit gate is optimal until $p \simeq 0.85$ (for $n=8$). As p increases further, a hump develops whose maximum increases until $p \simeq 0.89$, and decreases afterwards. The maximum of the hump shifts towards larger values of c as p approaches one, resulting in optimized convergence at $p \simeq 0.89$, $c \simeq 4/90$. Observe that the optimal probability, $p \simeq 0.89$, is close to $p=0.875=7/8$: For $n=8$, this is the value for which, on average, there is one CZ gate per iteration that is not implemented. In other words, for this p value the closed chain is similar to an open chain with $p=1$, except that the opening, corresponding to the missed CZ gate, changes position randomly at each iteration. This variable opening in the qubit chain is what increases the convergence rate to the point where the maximum attainable convergence rate on an open chain is outperformed by a closed chain with $p > 0.7$. As n increases, the hump is pushed to higher values of p, consistent with the fact that as n grows, the chances of one two-qubit gate not being applied, i.e., that there is a break in the chain, increases accordingly.

As noted above, the closed chain with $p=1$ does not converge to random due to the conservation of even or odd labeled nonidentity Pauli operators. However, for $p<1$ this conservation rule does not hold and ergodicity is restored. Nevertheless, based on calculations which directly compare the state of the system after many iterations of the Markov matrix to the random state, it appears that a transition occurs at $p \simeq 0.8$, roughly the value where the gap begins to decrease. Below this point, the mixing between the two parity subspaces is *on par* with the convergence to random within each subspace. For $p \gtrsim 0.8$, the relaxation between the subspaces decreases—eventually increasing the time for the PR circuit to reach the final random state. This is demonstrated in the lower inset of Fig. 7, which shows the gap size and convergence rate of PR algorithms using HZ single-qubit gates within the closed chain topology as a function of p. For $0.9 < p < 1$, the gap plummets and then goes back up at $p=1$. This small gap is due to the eigenvalue corresponding to the rate of transfer between the two subspaces of the system. In terms of entanglement distribution, a fast convergence to the value given by Eq. (5) occurs first, followed by a slower convergence to the expected Haar value at a rate determined by the small gap.

Figure 8 shows the Markov matrix gap for a probabilistic star topology. For $p \lesssim 0.7$, the HZ is the optimal single-qubit gate, while for $p \gtrsim 0.7$ a hump forms such that the maximum gap is no longer at $c=0$. Unlike the closed-chain topology, the peak of the hump is now always lower than the maximum overall gap, which is found at $c=0$, $p \simeq 0.7$ for $n=8$. Note that for $0.6 \lesssim p \lesssim 0.8$, the convergence rate is faster than

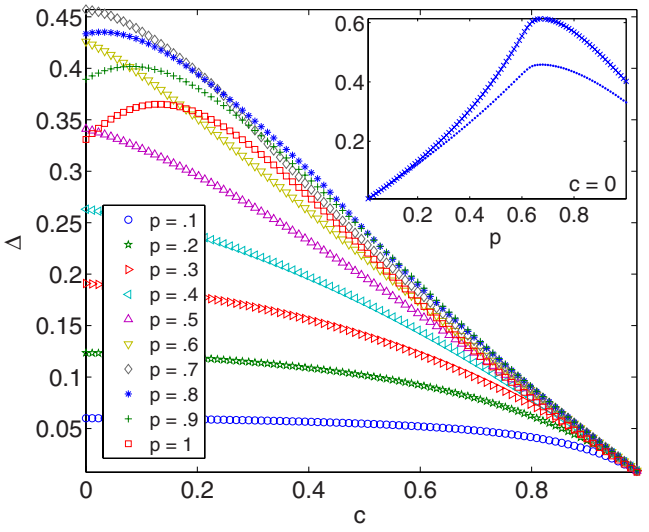

FIG. 8. (Color online) Markov spectral gap, Δ, versus local gate parameter, c, for an eight-qubit star topology with probability p that a given CZ gate will be implemented. The gap increases with p and the HZ gate ($c=0$) is the optimal single-qubit gate for $p \gtrsim 0.7$. Inset: Gap ($\cdot$) and convergence rate ($\times$) for HZ gates as a function of p.

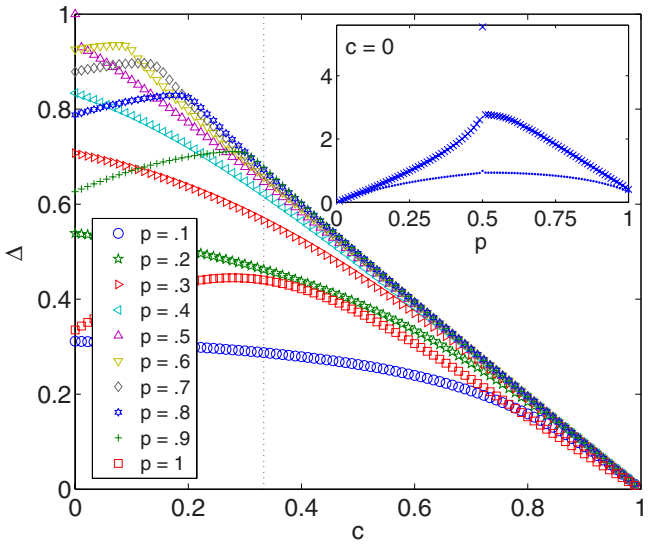

FIG. 9. (Color online) Markov spectral gap, Δ, versus local gate parameter, c, for an eight-qubit AA topology with probability p that a given CZ gate will be implemented. The largest gap appears as a singularity at $p=0.5$ and $c=0$. For $p<0.5$, the gap increases with decreasing c, the HZ gate ($c=0$) being optimal. For $p>0.5$, the largest possible gap decreases and the optimal single-qubit gate approaches random, $c=1/3$, with increasing p. In addition, the size of the gap as a function of c exhibits a sharp change of behavior after reaching its maximum. Inset: Gap ($\cdot$) and convergence rate ($\times$) for HZ gates as a function of p. The singularity at $p=0.5$ is easily seen, as is the difference in the behavior of the gap size as a function of p above and below this singularity.

that of the open-chain topology with $p=1$. The star topology has $(n-1)$ possible CZ gates thus $p=0.7$ means that there are about two inactive couplings per iteration. The curves for different p appear to be generally independent of n.

Lastly, a probabilistic AA topology is analyzed in Fig. 9. The observed gap values are much larger than any we have seen to this point. The optimal single-qubit rotation once again depends on p. For $p<0.5$, the optimal single-qubit rotation is the HZ gate, whereas for larger p the optimal choice of c increases until reaching $c \simeq 1/3$ when $p=1$. Unlike the other topologies, the gap behavior as a function of c for $p>0.5$ exhibits a sharp change after reaching a maximum. At $p=0.5$ and $c=0$, a sharp singularity occurs in both c and p, resulting in an extremely large gap, $\Delta=0.9961$. Remarkably, the size of this gap approaches one exponentially as the number of qubits increases. We will explore this specific case in more detail in the next section.

IV. SCALING BEHAVIOR

In this section we focus on the effect of scaling on PR circuits with certain qubit topologies. Specifically, symmetries in the AA and star topologies allow for the computation of Markov matrix gaps for PR circuits with a relatively large number of qubits. In addition, based on numerical evidence for the AA topology with probabilistically applied two-qubit gates, we provide a simple formula which captures the scaling of the spectral gap for different classes of qubit topologies based on their *degree of connectivity*—as measured by how many qubits are on average connected to each other via CZ gates. Our results are consistent with scaling results known to date, and complement exact results which have been recently established (notably, for NN coupling topologies) by a suitable mapping to quantum spin chains [36].

A. All-to-all topology

Each time step of a PR circuit using an AA topology consists of independent local gates on each qubit, followed by CZ gates between arbitrary pairs of qubits. We apply these gates probabilistically, in such a way that for each possible pairing of qubits a CZ gate is performed independently with probability p.

In order to access larger n, we construct a reduced representation by considering equivalence classes of Pauli operators under qubit permutations. Thus, we label the states of the reduced Markov matrix by the number of Z- and X-Pauli operators they contain. As noted previously, although each eigenvalue of the reduced chain occurs in the full representation, some eigenvalues of the full chain do not occur in the reduced chain, allowing for the possibility that the largest eigenvalue could, in principle, be removed. In practice, we verified that for up to $n=10$ where comparisons with the full representation are tractable, the eigenvalue that determines the gap occurs in both representations. Furthermore, the eigenvalue that governs the decay of computational basis states must occur in the reduced representation, because the second moments of computational basis states are invariant under qubit permutations.

For fixed p, *numerical results* show that the scaling behavior of the gap upon increasing n differs depending on whether $c=0$ or not. For $c>0$, the gap appears to converge at an exponential rate to $1-c$,

$$\Delta^{(c>0)}(p;n) = (1-c) - e^{-a(p)n}, \tag{6}$$

with a convergence rate $a(p)$ which is peaked at $p=0.5$. When $c=0$, however, the gap scales as

$$\Delta^{(c=0)}(p \leqslant 0.5;n) = 1 - \frac{1}{2^{\alpha(p)pn+\beta(p)}}, \tag{7}$$

where α and β are p-dependent factors, which are found to be of order unity for $p \leqslant 0.5$. Thus, in this case the gap converges exponentially to one.

Extremely interesting behavior of the gap occurs when $c=0$ and $p=0.5$. Remarkably, in this case the eigenvalues of the Markov chain can be obtained by exact methods and exhibit a very simple structure [37]: aside from the ergodic eigenvector with eigenvalue 1, one finds a single eigenvalue of -2^{-n}, multiply degenerate eigenvalues of $2^{-n/2}$ and $-2^{-n/2}$, all remaining eigenvalues being zero. Thus, the gap is given by

$$\Delta^{(c=0)}(p=0.5;n) = 1 - \frac{1}{2^{n/2}}. \tag{8}$$

This holds for *generic* initial states. Initial computational basis states, however, have a nonzero component only along the ergodic eigenvector, and the eigenvector with eigenvalue $1/2^n$. This further leads to a gap of

$$\Delta^{(c=0;\mathrm{comp})}(p=0.5;n) = 1 - \frac{1}{2^n}, \tag{9}$$

which is the fastest convergence rate for any PR circuit we have examined. The size of the gap as a function of n for $p \leqslant 0.5$ is shown in Fig. 10.

We can use Eq. (7) to gain unified insight into the gap scaling behavior for different classes of PR circuits. The known facts about the gap scaling with regard to the number of two-qubit gates applied *per time step* may be summarized as follows:

(i) For a PR circuit where a single two-qubit gate uniformly drawn from the Haar distribution on U(4) is implemented per time step between a pair of qubits selected at random, an asymptotic scaling of

$$\Delta^{[1]}(n) = \Theta(1/n)$$

was proved by Harrow and Low in [27]. In addition, numerical as well as analytical results by Znidaric [21,36] demonstrated the above $1/n$ scaling law to be *typical* for a large class of two-qubit gates used in conjunction with random SU(2) single-qubit gates.

(ii) For PR circuits where $\sim n$ commuting two-qubit gates are performed in parallel per time step, numerical evidence suggests that the gap tends to a constant value as n [15,23],

$$\Delta^{[n]}(n) = \mathrm{const},$$

the value of such a constant depending in general on the type of gates and qubit topology.

(iii) For the AA topology under discussion here for fixed p, $\sim n^2$ commuting two-qubit gates are performed in parallel per time step and, according to Eq. (7), our numerical data for $p<1$ and $c=0$ yield a scaling behavior of the form

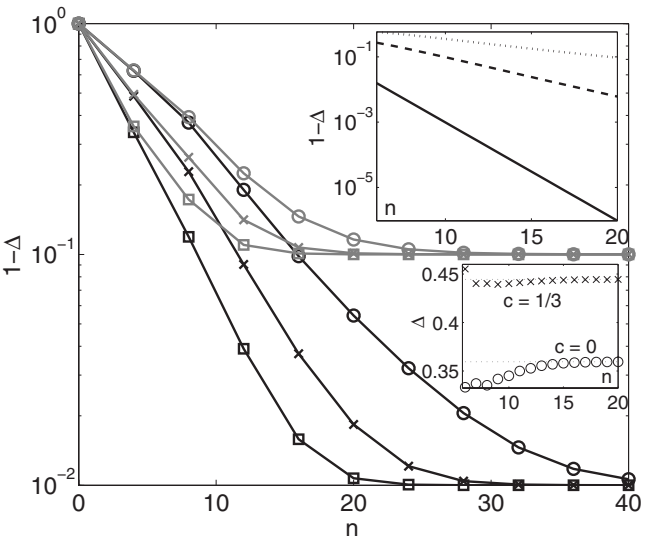

FIG. 10. $1-\Delta$, versus number of qubits for the AA topology using probabilistically applied CZ gates. Dark curves: $c=0.01$ and $p=0.25$ (○), 0.35 (×), and 0.45 (□). Light curves: $c=0.1$ using the same values of p. In both cases, the gap approaches $1-c$ for increasing n, according to Eq. (6). Upper inset: $1-\Delta$ versus number of qubits for the AA topology when $c=0$ and $p=0.2$ (dotted line), 0.4 (dashed line), and 0.5 (solid line). The gap approaches one at an exponential rate for all p, with the fastest convergence at $p=0.5$, according to Eqs. (7) and (9). Lower inset: Gap versus number of qubits for deterministic AA topology, $p=1$, for single-qubit HZ (○) and random (×) gates. The gap grows with increased n until saturating at about 0.3596 for the HZ gates and 0.4444 for the random gates.

$$\Delta^{[n^2]}(n) = 1 - \frac{1}{2^{\kappa n+\gamma}}, \quad \kappa > 0,$$

and constant scaling otherwise.

We shall now show that the convergence rates for each of the above cases are consistent when viewed in terms of the number of two-qubit gates applied. Let N_2 denote the number of two-qubit gates applied per time step. In order to compare the gap, $\Delta=1-\lambda_1$, for a protocol in which $N_2=1$ to one where $N_2 \sim n$ (such as the open-chain, closed-chain, and star topologies) or $N_2 \sim n^2$ (such as in the AA topology), we postulate an effective gap,

$$\Delta_{\mathrm{eff}}^{[N_2]} = 1 - \lambda_{\mathrm{eff}}^{[N_2]},$$

corresponding to a Markov matrix obtained by raising the original Markov matrix for which there is only one two-qubit gate per time step, to the power N_2, so that the same number of two-qubit gates are considered. Under the assumption that the typical gap for a protocol with $N_2=1$ scales as $1/n$, it follows that for $N_2 \sim n$, in the limit where n is sufficiently large,

$$\lambda_{\mathrm{eff}}^{[n]} = \lim_{n\to\infty}\left(1 - \frac{\kappa}{n}\right)^n = e^{-\kappa}.$$

This yields a gap of $\Delta_{\mathrm{eff}}^{[n]}=1-e^{-\kappa}$, which is a constant. For $N_2 \sim n^2$, the effective gap goes to $\Delta_{\mathrm{eff}}^{[n^2]}=1-e^{-\kappa n}$, which as-

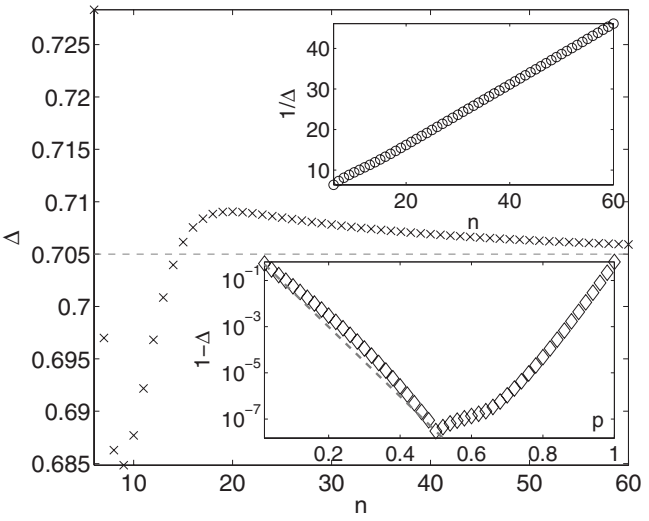

FIG. 11. Spectral gap, Δ, of a PR circuit using an AA topology as a function of the number of qubits, n, with two-qubit gates applied with probability $p=\frac{2}{n-1}$ such that at each iteration the number of gates performed is on average n. The gap asymptotically approaches a constant value of ≈ 0.705 (dashed line). The upper inset shows the inverse gap, $1/\Delta$, as a function of n for $p=\frac{2}{n(n-1)}$ so that on average one two-qubit gate is performed per iteration. The gap scales as $\propto 1/n$. The lower inset shows $(1-\Delta)$ for 50 qubits as a function of p. The gap is largest, and hence convergence quickest, at $p=0.5$. The approximate formula for the gap when $p<0.5$ is given by $\Delta=1-1/2^{np}$ and is shown as a dashed line.

ymptotically approaches 1, consistent with our numerical results.

We further note that the above characteristic scaling behaviors may be realized by allowing p, the probability of performing a two-qubit gate between a given (arbitrary) pair of qubits, to be a function of n within the AA topology. We consider the special case of $c=0$, and summarize the data in Fig. 11. First, set

$$p^{[1]}=\frac{2}{n(n-1)},$$

so that on average one CZ gate is performed per time step. Then as one may see, the gap is given asymptotically by $\Delta^{[1]}(n)=\frac{1.34}{n-1.55}$, as expected. Next, set

$$p^{[n]}=\frac{2}{n-1},$$

so that an average of n CZ gates are applied per time step. For large n, this results in a constant gap of $\Delta^{[n]}=0.705$. Finally, for fixed $p\leqslant 0.5$, the gap is seen to approach one, again as expected and consistent with Eq. (7).

In summary, we note that for a broad class of random circuits the gap scales as $\Delta=1-\exp(-\alpha N_2/n)$, $\alpha>0$. Since the convergence rate $\Gamma=-\ln(1-\Delta)=\alpha N_2/n$, this implies that *the convergence rate scales asymptotically as* $\sim 1/n$ *per two-qubit gate*. Such a scaling appears to be generic for typical PR circuits where $N_2=1$ to $\sim n$, and for *some* types of PR circuits with $N_2\sim n^2$—specifically when $p<1$ and $c=0$ for the AA topology. This is important because if one wishes to parallelize by simultaneously performing commuting two-qubit gates, then it is possible to perform the maximum $N_2\sim n^2$ gates without sacrificing the typical $1/n$ scaling per two-qubit gate in the convergence rate.

B. Star topology

For qubits arranged in the star topology, CZ gates may be applied only between the central qubit and the $(n-1)$ outer qubits. With the proper choice of parameters, this structure also allows (as the AA topology with $c=0$, $p=0.5$) for an exact determination of the Markov matrix eigenvalues [37]. Let single-qubit gates be drawn from a distribution with $c=1/3$ and each CZ gate be applied with a probability $p=3/4$. Then, the eigenvalues of the Markov matrix consist of a (2^n-2)-fold degenerate eigenvalue equal to $2/3$, and a pair of eigenvalues at $\frac{1}{3}(1\pm\frac{1}{2^{n-1}})$. The remaining eigenvalues are all zero, leading to a gap of

$$\Delta^{(c_{\text{all}}=1/3)}(p=0.75,n)=1/3.$$

We can also examine a situation where $p=3/4$ but we draw the single qubit rotation on the central qubit from a different distribution than that of the other qubits. If, for example, we draw the noncentral qubit rotations from a random distribution and the central qubit rotation from a distribution parametrized by c, the corresponding Markov matrix develops a (2^n-2)-fold degenerate eigenvalue equal to $\frac{1+c}{2}$, a pair of eigenvalues at $c\pm\frac{1-c}{2^n}$, and the remaining eigenvalues are zero. The gap is now maximized at 0.5 when $c=0$ for the local gate distribution of the central qubit,

$$\Delta^{(c_{\text{central}}=0,c_{\text{outer}}=1/3)}(p=0.75,n)=1/2.$$

Interestingly, this PR circuit is very similar in structure and performance to the approximate 2-design outlined in [25], as discussed in Appendix B.

V. EFFECT OF COLLECTIVE ROTATIONS

In all of the above analysis we have assumed PR circuits where the single-qubit rotations applied to each qubit are drawn independently from some specified distribution at each time interval. What would happen if we limit the number of times we draw from the given distribution? For example, will a PR circuit that applies the same random rotation to every qubit (or every other qubit) during a given iteration converge more slowly than if every qubit rotation is independent?

Constructing a PR circuit where we decrease the number of independently applied single-qubit random rotations may be viewed as an exercise in generating random states (or operators) with as few degrees of freedom as possible. It is well known that quantum chaotic systems with only one or two degrees of freedom can mimic several properties of random systems [38]. Exploring PR circuits by reducing the number of degrees of freedom goes one step further, in that it addresses a similar question within a general mathematical framework rather then looking at specific examples of what may be extreme cases. This aspect of PR circuits has been previously explored in [39] with respect to the unitary opera-

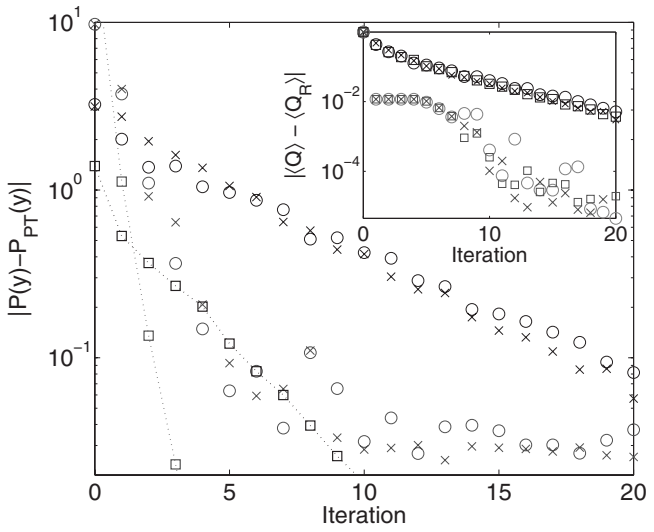

FIG. 12. l_2 distance between PR distribution of squared moduli of components in the computational basis and the PT distribution (inset: distance of average global entanglement from random value) versus number of iterations for open-chain topologies (eight qubits, 100 implementations) in which each qubit undergoes equivalent random rotations (○) or every other qubit undergoes equivalent random rotations (×). Both random single-qubit gates (dark) and HZ gates (light) were used. These are compared to circuits in which each qubit undergoes a different rotation (□). Interestingly, using different rotations on all odd and all even qubits does not improve the convergence rate beyond the one achievable by using only one collective rotation for all qubits.

tors applied by the algorithm. Here, we look at some further statistics of the generated PR states and compare them to the same characteristics of random states described above.

Specifically, let us consider an open chain of qubits with CZ gates between NN qubits. For the single-qubit gates, instead of applying different random or HZ rotations to every qubit at every PR iteration, we apply the same *collective* random or HZ rotation to all qubits for a given iteration. We also look at a case where, instead of one single-qubit rotation per iteration, we apply two different rotations, the first to odd qubits and the second to even qubits. We compare the convergence of the element distribution and entanglement for these two cases to that of an open chain where different rotations (random or HZ, respectively) are applied to each qubit. As shown in Fig. 12, limiting the diversity of single-qubit rotations appears to have little effect on the convergence rate of entanglement. The matrix element distribution, however, is strongly affected by the symmetries, and the convergence to the PT distribution is significantly slowed down by reducing the number of free parameters in the PR algorithm. Were a quantum protocol dependent on randomness in the state element distribution, these PR circuits would require significantly more steps than PR circuits with different random rotations on each qubit. It is worth noting that for the case in which every qubit undergoes the same local gate, the PR gate set is not universal over the full Hilbert space because it preserves parity with respect to reversing the order of the qubits along the chain. Nevertheless, similar to previous cases, the asymptotic decay rate is not heavily influenced by this fact, although the final asymptotic values of the test functions under consideration may differ slightly from those expected for the Haar distribution.

VI. INITIAL STATE DEPENDENCE

One of the challenges in constructing PR states is creating sufficient entanglement using simple one- and two-qubit gates. If, however, the initial state already has some entanglement the absolute convergence (although not the rate of convergence) to randomness may proceed more quickly. To demonstrate this, we apply PR circuits to initial states of the form

$$|\psi(a)\rangle = \frac{1}{K(a)}[(1-a)(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}) + a(|0\rangle + |1\rangle)^{\otimes n}], \tag{10}$$

where $K(a)$ is the normalization factor. When $a=0$, the state $|\psi(0)\rangle$ is a generalized Greenberger-Horne-Zeilinger (GHZ) state which is maximally entangled relative to arbitrary local observables, and $K(a)=1/\sqrt{2}$. When $a=1$, $|\psi(1)\rangle$ is completely separable, and $K(a)=1/\sqrt{N}$. By applying the PR circuit to initial states with different values of a, we can explore the number of iterations gained by using an initial entangled state. Of course, starting with an initially entangled state has a cost. The cost is minimal, however, since the above states are easily generated via a series of controlled σ_x rotations by angles which are dependent on a. These gates can be implemented simultaneously since all gates share the same control qubit.

Figure 13 shows the difference in state element distribution and average entanglement between random states and states produced from the PR circuit with different initial states as a function of iteration. All states approach the random state entanglement at the same exponential rate. Not too surprisingly, however, initial states with entanglement closer to that of random states approach random state entanglement faster than other states. For $n=8$, the entanglement of the state $a \simeq 0.02337$ equals that of random states, and has the fastest convergence to random entanglement (that the state with entanglement equal to that of random states converges the fastest is more clearly seen for smaller numbers of qubits). As n grows, the random entanglement becomes closer to maximum, therefore the GHZ state provides optimal convergence. Perhaps more surprisingly, the states with more initial entanglement also converge to the random state element distribution more quickly than their less entangled counterparts. This is in line with the notion that, in general, characteristic entanglement and state element distribution are correlated [40]. Figure 13 also allows us to quantify the advantage of starting with entangled states in terms of the number of iterations by noting where the different curves cross a given distance from random entanglement level. For example, the $a=0$ state gives an advantage of 4–5 iterations over the $a=1$ state.

To further test the dependence of a PR algorithm convergence on the initial state, we have also used as the initial state the open-chain cluster state (still utilizing a circuit

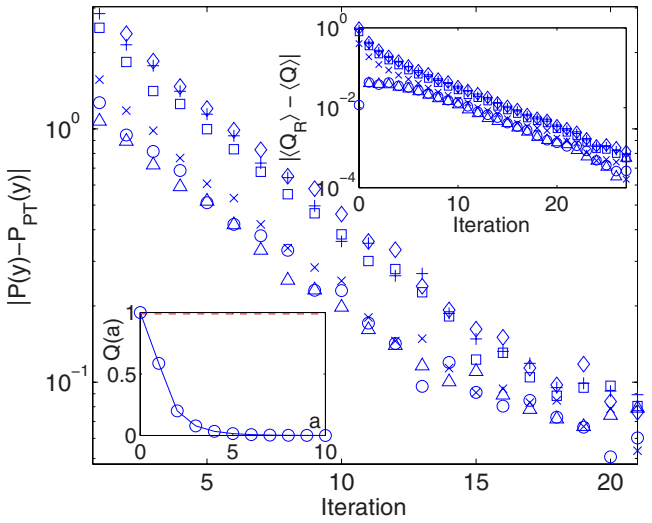

FIG. 13. (Color online) l_2 distance between PR distribution of squared moduli of components in the computational basis and the PT distribution (inset: distance of average global entanglement from random value) versus number of iterations of a PR algorithm for eight qubits (500 samples) and initial states $a=0$ (○), $a\simeq 0.02337$, corresponding to a state with random entanglement (△),0.1 (×), 0.2 (□), 0.3 (+), and 1 (◇). All states converge to random at the same rate. Note that the $a=0$ case has a four or five iteration advantage over the $a=1$ initial state. Lower inset: Entanglement of the initial states versus a. The random entanglement value is $\simeq 0.9883$.

model PR algorithm). We find that the results are equivalent to that of the generalized GHZ state. Note that with respect to the multipartite entanglement measure we are using, both these states are maximally entangled. However, the apparent equivalence between the GHZ state and cluster state does *not* hold for all topologies. For example, the convergence of a PR algorithm with the qubits in the AA topology does not benefit from an initial GHZ state, whereas it does benefit from an initial cluster state.

VII. CONCLUSION

We have studied the efficiency of PR circuits as a function of single-qubit and two-qubit gates, qubit topology, the probability of applying the two-qubit gates, and system size, as well as addressed the effect of collective rotations and different initialization. Using a Markov chain analysis, we have analyzed the strong interdependence of the choice of single- and two-qubit gates. The optimal PR algorithm will consist of a distinct pair of single- and two-qubit gates. The optimal choice of gates is also dependent on the qubit topology. Different topologies will favor one or another single- or two-qubits gates. Also dependent on the topology, the efficiency of the PR algorithm may increase if the two-qubit gates are applied probabilistically. In particular, we noticed extremely fast convergence for the AA topology with $p=0.5$.

As one of the main implications of our work, we have demonstrated how, as the number of qubits is increased, the asymptotic convergence rate of second-order moments of the PR circuit scales as $1/n$ per two-qubit gate implemented per time step. This may prove useful to quantitatively characterize PR circuits in terms of t-designs. In general, obtaining a deeper understanding of the behavior of PR algorithms as approximate higher-order t-designs appears to be an important next step toward harnessing quantum pseudorandomness.

ACKNOWLEDGMENTS

It is a pleasure to thank David G. Cory, C. Stephen Hellberg, and Cecilia C. Lopez for valuable discussions and input. W.G.B. gratefully acknowledges partial support from Constance and Walter Burke through their Special Projects Fund in QIS. Y.S.W. acknowledges support from the MITRE Technology Program under MTP Grant No. 07MSR205.

APPENDIX A: CLUSTER-STATE PARAMETERS

Several potential applications for PR circuits are in the area of quantum communications, which is naturally suited for photonic implementations. Current analyses suggest that linear optics QC will require too many resources, in terms of amount of optical hardware, for practical implementations [41]. The cluster-state approach to QC, however, may be less stringent on resource requirements. A cluster state [42] can be created by first rotating all qubits into the state $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$. Qubits are then entangled by applying CZ gates between pairs of qubits j and k. In a graphical picture of a cluster state, qubits are represented by circles and pairs of qubits that have been entangled via a CZ gate are connected by a line. A cluster state with qubits arranged in a two-dimensional lattice such that each qubit has been entangled with four NNs suffices for universal QC [43]. For photonic realizations of cluster-state QC, where two-qubit gates can only be applied probabilistically, one assumes an unlimited number of maximally entangled Einstein-Podolsky-Rosen (EPR) pairs (attained, for example, via spontaneous parametric down conversion) which can be used to construct larger cluster states using so-called fusion operations [44].

Once the desired cluster state has been constructed, any QC algorithm can be implemented using only single-qubit measurements in the x-y plane. The inherent entanglement in the cluster state means that measurement on one qubit may effect the state of the remaining qubits. Thus, one can view each row of the cluster-state lattice as the evolution of a single-qubit in time. Processing measurements are performed by column from left to right until the last column (left unmeasured), which contains the output state of the quantum algorithm, to be extracted by a last readout measurement. For a one-dimensional cluster chain, the logical operation implemented by measurement along an angle ϕ in the x-y plane is $X(\pi m)HZ(\phi)$, where H is the Hadamard gate and $Z(\alpha)$ $(X(\alpha))$ is a z (x) rotation by an angle α [45]. The dependence of the logical operation on the outcome of the measurement is manifest in $m=0,1$ for measurement outcome $-1,+1$. Any arbitrary rotation can be implemented via three logical single-qubit rotations of the above sort yielding

$$HZ(\alpha+\pi m_\alpha)X(\beta+\pi m_\beta)Z(\gamma+\pi m_\gamma), \tag{A1}$$

where (α,β,γ) are the Euler angles of the rotation. By drawing the Euler angles according to the Haar measure, a ran-

dom single-qubit rotation can be implemented. Because the Haar distribution on SU(2) is invariant under π rotations, the measurement-dependent π rotations may be ignored. Two-qubit gates are performed via a connection between two rows of the cluster state. CZ gates, in particular, are "built in" to the cluster state and simple measurement automatically implements the gate.

PR state generation via the cluster state model of QC was first explored in [23]. The algorithm was initially performed using three qubits in a chain to implement random rotations, $c=1/3$, and a connection between rows of the cluster state on every third qubit implemented the CZ gates. In this way, ℓ iterations of an n-qubit PR circuit required a lattice of $n\times 3\ell+1$ qubits (where the extra "1" comes from the final, unmeasured column). However, by filling in the extra vertical connections, thus making the single-qubit rotations HZ gates with $c=0$, such that all qubits were connected to their four NNs, the convergence rate of the PR algorithm increased.

1. Cluster topologies

We explore cluster state topologies where the columns of the two-dimensional lattice are connected as circles (making the cluster state a cylinder), where all qubits in a "column" are connected to each other (AA) and where each qubit is attached only to one central qubit (stars). Both the AA shape and the star shape are graph representations of generalized GHZ states differing only by single-qubit rotations [46]. Within the star topology, we look at two different measurement strategies—measuring the central qubit first, and measuring the central qubit last. Figure 14 shows that the GHZ-type states converge more slowly than the two-dimensional lattice and cylindrical topologies with the cylindrical topology converging fastest. This is similar to the circuit model for PR algorithms employing two-qubit CZ gates and one-qubit HZ gates where the AA topology converges the slowest.

2. Probabilistically applied cluster gates

Photonic implementations of cluster state QC require construction of the appropriate cluster state using operations that work probabilistically. Thus, any cluster state has an associated cost in the amount of resources required to, on average, build the given state. Much work has gone into finding algorithms that will lower the cost of construction for cluster states of two dimensions, i.e., in which each qubit has four NNs [47]. In these works, cost is generally associated with the number of EPR pairs needed for state construction. One conclusion that emerges from this work is that the cost of constructing one-dimensional chains is relatively low and it is joining these chains (such that CZ gates can be applied) which is the majority of the cost. For this analysis we assume that these connections can be attempted without breaking the chain (even if the operation is not successful).

We have already noted that a full two-dimensional cluster state (with every vertical connection filled in) is optimal for PR cluster circuits. However, what if we were to apply the cluster CZ gates only probabilistically? While we know from our exploration of circuit model PR algorithms using an open chain topology that the number of necessary iterations would increase, a probabilistic application of these gates would cut the construction cost of the initial cluster state.

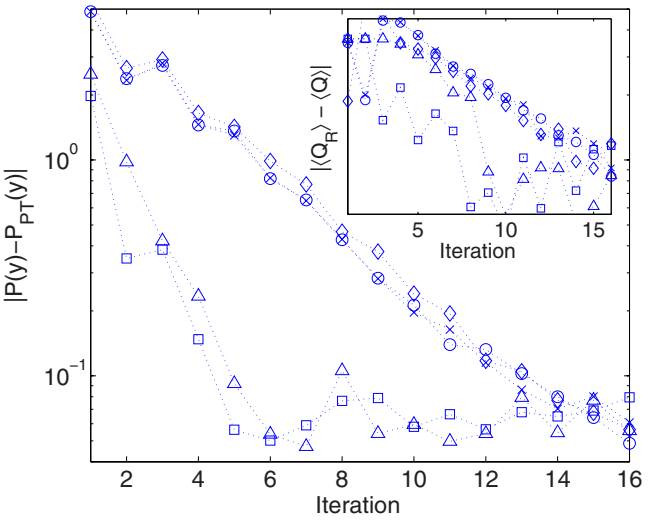

FIG. 14. (Color online) l_2 distance between PR distribution of squared moduli of components in the computational basis and the PT distribution (inset: distance of average global entanglement from random value) versus number of iterations for different cluster state topologies (eight qubits, 500 implementations): two-dimensional lattice (△), cylinder (□), star topology measuring the central qubit first (×), star topology measuring the central qubit last (○), and AA (◇). The cylindrical lattice converges most quickly. The GHZ-type topologies (star and AA) converge more slowly and are, in this way, similar to the AA topologies of circuit-model PR algorithms using two-qubit CZ and one-qubit HZ gates. The open and closed chain exhibit cutoff behavior in the entanglement convergence [23].

Let us assume that an attempt to link two cluster chains together succeeds with probability p'. We also assume the availability of infinitely long cluster chains. As an example let $p'=1/2$ and we attempt to construct a cluster state which can implement C iterations of the cluster PR circuit in which the two-qubit gates are implemented with probability p. On average, construction of such a cluster state would require $2Cpn$ probabilistic operations, where n is the number of rows (logical qubits) in the cluster lattice. Thus, for an eight-qubit cluster PR algorithm, to construct a cluster capable of implementing a PR circuit with a convergence rate of 0.547 (which is at $p=0.98$ as shown in Fig. 6) requires on average $1.96Cn$ probabilistic operations (recall that the circuit PR algorithm with HZ gates is equivalent to the cluster-state PR algorithm). A PR circuit with half that convergence rate would require $2C$ iterations but could use an algorithm with $p\simeq 0.705$. This would require on average $\simeq 2.82Cn$ attempts. Thus, for this case, it is better to construct a cluster state capable of higher p value PR circuits. However, further analysis is necessary in the more realistic case where all resources are accurately counted (including losses taken upon failure of fusion operations), for different topologies, and for other values of p and p'.

APPENDIX B: TWIRLING AS A PSEUDORANDOM CIRCUIT

Clifford twirls [48], or 2-designs, play an important role in a number of quantum information protocols. A construction for an approximate Clifford twirl has recently been suggested in [25] (see also [49] for related work). The procedure detailed in [25] consists of a sequence of twirling operations. A twirl consists of (i) performing one operation out of a specified set of operations; (ii) allowing the superoperator, Λ, to occur; (iii) undoing the original operation. The net effect of the twirl is to convert Λ into a new superoperator Λ'. In the case of a Clifford twirl, in which the set of operations is the Clifford group, any superoperator is converted into a depolarizing channel.

The effect of the sequence of twirling operators on the set of Pauli strings can be represented as a Markov matrix, identical to those that describe the decay of the second moments of a PR circuit. Thus, a one-to-one mapping exists between PR circuit constructions and approximate twirling operations. When the procedure in [25] is considered as a PR circuit, it bears a close resemblance to PR circuits in a star topology, with probabilistic CZ gates, and random local gates applied to the noncentral qubits. However, the procedure of [25] involves three time steps each of which is counted as an application of simultaneous probabilistic CZ gates. An approximate calculation of the convergence rate is made in [25] in order to determine scaling properties. We find that an exact calculation of the gap of the corresponding Markov matrix gives $\Delta_{\text{Clifford}}(c=1/3)=5/6$, irrespective of the number of qubits. We can improve the convergence rate of this procedure marginally by using the following, simpler construction: perform a PR circuit with the star topology with $c=0$ for the central qubit, while keeping $c=1/3$ for the noncentral qubits and a gate probability of $p=3/4$. We find a gap of $\Delta=1/2$. Since this must be applied three times to perform the same number of gates as in [25], this improved construction yields an effective gap of $\Delta_{\text{Clifford}}(c=0)=1-1/2^3=7/8$.

[1] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, New York, 2006).
[2] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).
[3] A. Harrow, P. Hayden, and D. Leung, Phys. Rev. Lett. **92**, 187901 (2004).
[4] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Commun. Math. Phys. **250**, 371 (2004).
[5] A. Ambainis and A. Smith, e-print arXiv:quant-ph/0404075, in *Proceedings of RANDOM 2004*, Cambridge, MA (unpublished).
[6] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993).
[7] S. K. Foong and S. Kanno, Phys. Rev. Lett. **72**, 1148 (1994).
[8] A. Scott and C. M. Caves, J. Phys. A **36**, 9553 (2003).
[9] O. Giraud, J. Phys. A **40**, 2793 (2007).
[10] L. Viola and W. G. Brown, J. Phys. A **40**, 8109 (2007).
[11] V. Cappellini, H-.J. Sommers, and K. Zyczkowski, Phys. Rev. A **74**, 062322 (2006).
[12] P. Facchi, G. Florio, and S. Pascazio, Phys. Rev. A **74**, 042331 (2006).
[13] M. L. Mehta, *Random Matrices* (Academic Press, New York, 1991).
[14] C. H. Bennett, P. Hayden, D. Leung, P. W. Shor, and A. Winter, IEEE Trans. Inf. Theory **51**, 56 (2005).
[15] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Science **302**, 2098 (2003).
[16] B. Levi, C. C. Lopez, J. Emerson, and D. G. Cory, Phys. Rev. A **75**, 022314 (2007); J. Emerson *et al.*, Science **317**, 1893 (2007).
[17] A. Bendersky, F. Pastawski, and J. P. Paz, Phys. Rev. Lett. **100**, 190403 (2008).
[18] J. Emerson, E. Livine, and S. Lloyd, Phys. Rev. A **72**, 060302(R) (2005).
[19] Y. S. Weinstein and C. S. Hellberg, Phys. Rev. Lett. **95**, 030501 (2005).
[20] L. Arnaud and D. Braun, e-print arXiv:quant-ph/0807.0775.
[21] M. Znidaric, Phys. Rev. A **76**, 012318 (2007).
[22] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio, Phys. Rev. Lett. **98**, 130502 (2007); O. C. O. Dahlsten, R. Oliveira, and M. B. Plenio, J. Phys. A **40**, 8081 (2007).
[23] W. G. Brown, Y. S. Weinstein, and L. Viola, Phys. Rev. A **77**, 040303(R) (2008).
[24] A. D. K. Plato, O. C. Dahlsten, and M. B. Plenio, e-print arXiv:quant-ph/0806.3058.
[25] C. Dankert, R. Cleve, J. Emerson, and E. Livine, e-print arXiv:quant-ph/0606161; A. Ambainis and J. Emerson, e-print arXiv:quant-ph/0701126.
[26] D. Gross, K. Audenaert, and J. Eisert, J. Math. Phys. **48**, 052104 (2007).
[27] A. Harrow and R. Low, e-print arXiv:quant-ph/0802.1919.
[28] H. Barnum, E. Knill, G. Ortiz, and L. Viola, Phys. Rev. A **68**, 032308 (2003); H. Barnum, E. Knill, G. Ortiz, R. Somma, and L. Viola, Phys. Rev. Lett. **92**, 107902 (2004); R. Somma, G. Ortiz, H. Barnum, E. Knill, and L. Viola, Phys. Rev. A **70**, 042311 (2004).
[29] H. Barnum, G. Ortiz, R. Somma, and L. Viola, Int. J. Theor. Phys. **44**, 2127 (2005).
[30] D. A. Meyer and N. R. Wallach, J. Math. Phys. **43**, 4273 (2002).
[31] G. K. Brennen, Quantum Inf. Comput. **3**, 619 (2003).
[32] K. Zyczkowski and M. Kus, J. Phys. A **27**, 4235 (1994); M. Pozniak, K. Zyczkowski, and M. Kus, *ibid.* **31**, 1059 (1998); F. Haake and K. Zyczkowski, Phys. Rev. A **42**, 1013 (1990).
[33] If no restriction is placed on allowable pairs, CZ and *XY* gates yield an equivalent convergence rate when used in conjuction with SU(2) single-qubit gates.
[34] Y. Most, Y. Shimoni, and O. Biham, Phys. Rev. A **76**, 022328 (2007).
[35] The data are noisy due to the twofold degeneracy of λ_2, the largest nontrivial eigenvalue of M'. When smoothed, the decay is exponential as expected.
[36] M. Znidaric, Phys. Rev. A **78**, 032324 (2008).

[37] The method by which the eigenvalues of the Markov matrix for $c=0$, $p=0.5$ may be exactly determined is part of a more broadly applicable perturbative approach which will be reported in a separate forthcoming analysis.
[38] F. Haake, *Quantum Signatures of Chaos* (Springer, New York, 2001).
[39] Y. S. Weinstein and C. S. Hellberg, Phys. Rev. A **69**, 062301 (2004); Y. S. Weinstein and C. S. Hellberg, *ibid.* **71**, 014303 (2005).
[40] Y. S. Weinstein and C. S. Hellberg, Phys. Rev. A **72**, 022331 (2005).
[41] M. A. Nielsen, Phys. Rev. Lett. **93**, 040503 (2004).
[42] H. J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).
[43] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).
[44] D. E. Browne and T. Rudolph, Phys. Rev. Lett. **95**, 010501 (2005).
[45] R. Raussendorf, D. E. Browne, and H. J. Briegel, Phys. Rev. A **68**, 022312 (2003).
[46] M. Hein, J. Eisert, and H. J. Briegel, Phys. Rev. A **69**, 062311 (2004).
[47] M. A. Nielsen, Phys. Rev. Lett. **93**, 040503 (2004); D. E. Browne and T. Rudolph, *ibid.* **95**, 010501 (2005); L. M. Duan and R. Raussendorf, *ibid.* **95**, 080503 (2005); Q. Chen, J. Cheng, K. L. Wang, and J. Du, Phys. Rev. A **73**, 012303 (2006); G. Gilbert, M. Hamrick, and Y. S. Weinstein, *ibid.* **73**, 064303 (2006); D. Gross, K. Kieling, and J. Eisert, *ibid.* **74**, 042343 (2006).
[48] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, IEEE Trans. Inf. Theory **48**, 580 (2002).
[49] G. Toth and J. J. Garcia-Ripoll, Phys. Rev. A **75**, 042311 (2007).