Phase-random states: ensembles of states with fixed amplitudes and uniformly distributed phases in a fixed basis

Yoshifumi Nakata,1 Peter S. Turner,1 and Mio Murao1,2

1 Department of Physics, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan
2 Institute for Nano Quantum Information Electronics, University of Tokyo, Tokyo 153-8505, Japan

(Dated: July 10, 2012)

Motivated by studies of typical properties of quantum states in statistical mechanics, we introduce phase-random states, an ensemble of pure states with fixed amplitudes and uniformly distributed phases in a fixed basis. We first give a sufficient condition for canonical states to typically appear in subsystems of phase-random states, which reveals a trade-off relation between the initial state in the bounded energy subspace and the energy eigenstates that define that subspace. We then investigate the simulatability of phase-random states, which is directly related to that of time evolution in closed systems, by studying their entanglement properties. We find that starting from a separable state, time evolutions under Hamiltonians composed of only separable eigenstates generate extremely high entanglement and are difficult to simulate with matrix product states. We also show that random quantum circuits consisting of only two-qubit diagonal unitaries can generate an ensemble with the same average entanglement as phase-random states.

PACS numbers: 03.67.Mn, 03.67.Bg, 05.70.-a

I. INTRODUCTION

One of the goals of quantum many-body physics is to be able to compute properties such as expectation values of observables, entanglement and explicit state descriptions for physical systems composed of many particles. These properties depend upon the parameters of the quantum states involved, and calculations are made difficult by the fact that the number of these parameters grows exponentially with the number of particles. One way around this problem is to consider ensembles of states as opposed to individual states, as many of those parameters are then averaged out.

The most natural and well-studied ensemble of states is that of random states, the set of pure states in Hilbert space selected randomly from the unitarily invariant distribution, used in many areas of quantum physics and quantum information science [1, 2]. The entanglement of random states is an area of particular recent interest, where it has been shown that in large systems, the average amount of entanglement is nearly maximal according to several measures [3]. In the context of quantum statistical mechanics, the high entanglement of random states has been shown to lead to reduced states that are thermal, under their restriction to a subspace constrained by the amount of total energy [4].

Symmetry is often used as a constraint, however there are other ways to place useful restrictions on an ensemble. If states are restricted to a certain subspace of Hilbert space, an ensemble of states could be described by random states in that subspace. In this paper, we address another type of restriction, which leads to ensembles of states in a subset of a Hilbert space. A simple example of a subset of states that is not a subspace is the set of product states, as they are not closed under superpositions. A more interesting example is a state evolving under a time-independent Hamiltonian. Since the time evolution changes only the phases of the expansion coefficients in the Hamiltonian’s eigenbasis, states reached during the time evolution form a subset rather than a subspace. Our aim is to study random states in such a subset, namely, an ensemble of states where the randomness is restricted to the phase of the complex expansion coefficients in a given basis, which we call phase-random states.

Phase-random states are closely connected to studies of typical properties in statistical mechanics [1, 2]. States in which the phases are the pertinent degree of freedom appear in quantum information theory such as in instantaneously quantum polynomial time (IQP) circuits [5], and locally maximally entangled (LME) states [6]. Phase-random states also describe situations where the accessible information is limited to the amplitudes in a certain basis, for instance, those where a unique rank-1 measurement is allowed.

Motivated by these considerations, we investigate statistical properties of phase-random states, which clearly depend on the amplitudes of the coefficients as well as the expansion basis. We first demonstrate the thermalization of their reduced states, which implies potential uses of phase-random ensembles to realize thermal states in subsystems. Then, regarding phase-random states as typical states during a Hamiltonian dynamics, we discuss simulating them with matrix product states (MPSs) [7] by deriving the average amount of entanglement of phase-random states. Moreover, applications of phase-random states in IQP circuits and as LME states lead us to develop a scheme for generating an ensemble of states simulating the entanglement of phase-random states by a quantum circuit composed of only diagonal two qubit unitaries, which we call a phase-random circuit.

The paper is organized as follows. In Sec. II, we define phase-random states and show how they can be used to study thermalization in statistical mechanics. In Sec. III, we derive explicit formula for the average amount of en-
tanglement of phase-random states. Using the formula, we investigate the simulatability of Hamiltonian dynamics by MPSs in Sec. [V]. Finally, we introduce and analyze phase-random circuits in Sec. [V]

II. PHASE-RANDOM STATES AND THERMALIZATION

Given a Hilbert space $\mathcal{H}$, we denote an ensemble of pure states $|\psi\rangle \in \mathcal{H}$ distributed according to some measure $d\mu$ by $\Upsilon = \{|\psi\rangle\}_d\mu$. The ensemble of random states is written $\Upsilon_{\text{rand}} = \{|\psi\rangle\}_{d\psi}$, where $|\psi\rangle$ is an arbitrary state and $d\psi$ is the unitarily invariant normalized Haar measure. For a Hilbert space of $N$ qubits, consider states of the form

$$|\phi\rangle = \sum_{n=1}^{2^N} r_n e^{i\phi_n} |u_n\rangle,$$

with both the amplitudes $\{r_n\}, \sum_{n=1}^{2^N} r_n^2 = 1, 0 \leq r_n \leq 1$ and orthonormal basis $\{|u_n\rangle\}$ fixed. By phase-random states, we mean the ensemble $\Upsilon_{\text{phase}} = \{|\phi\rangle\}_d\phi$, where the phases $\phi_n$ are distributed according to the normalized Lebesgue measure given by

$$d\phi = \frac{d\phi_1}{2\pi} \cdots \frac{d\phi_{2^N}}{2\pi},$$

on $[0, 2\pi]^{2^N}$. This ensemble clearly depends on the choice of amplitudes and basis, which we write $\Upsilon_{\text{phase}}(\{r_n, |u_n\rangle\}_I_n)$ when there is need to be explicit. Note that the ensemble of phase-random ensembles with appropriately distributed amplitudes is the ensemble of random states.

We first point out that studies of thermalization in closed systems [1, 2] are special instances of the study of phase-random states. To see this, consider a Hilbert space $\mathcal{H}$ with dimension $d$, and a Hamiltonian $H = \sum_{n=1}^{d} e_n |e_n\rangle \langle e_n|$. The state at time $t$ is given by

$$|\phi(t)\rangle = \sum_{n} r_n e^{-i e_n t / \hbar + i\phi_n} |e_n\rangle,$$

where $r_n e^{i\phi_n} = \langle e_n | \phi_0 \rangle$ with $r_n \geq 0$ and $|\phi_0\rangle$ is an initial state. Then, a time averaged thermodynamical quantity is often considered by assuming phase ergodicity in the sense that the distribution of phases $e^{-i e_n t / \hbar + i\phi_n}$ are uniform in $[0, 2\pi]$ in the long-time limit. Due to this identification, all studies addressing the time average are equivalent to investigations of statistical properties of the corresponding phase-random states $\Upsilon_{\text{phase}}(\{r_n, |e_n\rangle\}_I_n)$.

For example, in [1] it was proven that time evolution typically gives rise to canonical distributions in subsystems. In this case, we consider a Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_E$ where $\mathcal{H}_S$ ($\mathcal{H}_E$) represents a system (environment) with dimension $d_S$ ($d_E$), and $\mathcal{H}_R$ is a restricted subspace constrained by the energy defined by

$$\mathcal{H}_R = \text{span}\{|e_\alpha\rangle | -\delta e < e_\alpha < e + \delta e\} \subset \mathcal{H}_S \otimes \mathcal{H}_E.$$
will involve a sum over four basis labels \( n, m, l, k \) where the only phase integral that occurs is \( \{ \} \), which, recall, is a function of \( \sum \tilde{\phi} \). Defining the computational basis for subsystem \( A \), \( \Phi \) is a function of \( r_n, u_n \). Thus we have

\[
\langle E_{L}^{(A)} \rangle_{\text{phase}} = \sum_{n, m, l} \sum_{a, b} \langle \Phi_n | \Phi_m \rangle \langle \phi_n | \phi_m \rangle - \sum_{n} \sum_{a} \langle \Phi_n | \Phi_m \rangle \langle \phi_n | \phi_m \rangle.
\]

Equation (4) simplifies the dependence of the average amount of entanglement over phases is given by

\[
\langle E_{L}^{(A)} \rangle_{\text{phase}} = I_{L}^{(A)} (\Phi) - \sum_{n=1}^{2N} E_{L}^{(A)} (|u_n\rangle).
\]

This shows that \( \langle E_{L}^{(A)} \rangle_{\text{phase}} \) is a decreasing function of the basis entanglement, \( E_{L}^{(A)} (|u_n\rangle) \). Hence ensembles that also have a separable basis \( \{ |u_n^{\text{sep}}\rangle \} \), denoted by \( \gamma_{\text{phase}}^{\text{sep}} = \gamma_{\text{phase}}(\{ 2^{-N/2}, |u_n^{\text{sep}}\rangle \}) \), give the maximum,

\[
\langle E_{L}^{(A)} \rangle_{\text{phase}} = 1 - \frac{2N_A + 2N_A - 1}{2N_{\text{ph}}},
\]

This value is greater than that of random states given by Eq. (2), see also Fig. 1. For \( \gamma_{\text{phase}}^{\text{sep}} = 2^{-N/2} \), applying the concentration of measure to \( \Delta E_{L}^{(A)} (|\phi_{\text{phase}}^{\text{sep}}\rangle) = E_{L}^{(A)} (|\phi_{\text{phase}}^{\text{sep}}\rangle) - E_{L}^{(A)} (|\phi_{\text{phase}}\rangle) \), we find

\[
\text{Prob} \left[ \Delta E_{L}^{(A)} (|\phi_{\text{phase}}^{\text{sep}}\rangle) > 2/2N + \epsilon \right] \leq e^{-c2^N},\]

where \( c = 1/(211\pi^2) \). The proof is similar to that in Ref. [6] (see Appendix C for details). Thus the entanglement of phase-random states \( \gamma_{\text{phase}}^{\text{sep}} \) is highly concentrated around the average, demonstrated in Fig. 1.

States with equal amplitudes in a separable basis are also known as LME states [9]. These are the class of multipartite states that are maximally entangleable with local auxiliary systems by only local operations. In Ref. [6], it is mentioned that the LME states should exhibit high entanglement. Our result proves this statement is true in the sense that the uniform ensemble of LME states
achieves a higher average amount of entanglement, in terms of linear entropy, than that of random states.

On the other hand, for separable-basis ensembles defined by \( \Upsilon_{\text{sep}} = \Upsilon_{\text{phase}} \{ r_n, |u_n^{\text{sep}} \rangle \} \), an upper bound of the average amount of entanglement is given by

\[
0 \leq \langle E_L^{(A)} \rangle_{\Upsilon_{\text{sep}}} \leq S_L(\hat{\Phi}^{\text{sep}}),
\]

where \( \hat{\Phi}^{\text{sep}} = \sum_n r_n^2 |u_n^{\text{sep}} \rangle \langle u_n^{\text{sep}} | \). When the number of non-zero \( r_n \) is \( R \), \( S_L(\hat{\Phi}^{\text{sep}}) \) is bounded by \( 1 - 1/R \) from above. If \( R \) is small, (for instance, if \( R = \text{poly}(N) \)), the average amount of entanglement cannot be as large as that of random states. It is therefore necessary for the basis to be entangled in order to generate a large amount of entanglement on average when \( R \) is small.

IV. SIMULATABILITY OF HAMILTONIAN DYNAMICS

We now interpret our results in the context of time-independent Hamiltonian dynamics and consider the simulatability of the state during the time evolution by assuming phase ergodicity. We consider the area law of entanglement, which states that the von Neumann entropy of entanglement of a large subsystem is at most proportional to its boundary. Since the breakdown of the area law indicates that the states cannot be simulated by MPSs with a constant matrix size \( \Gamma \), the area law gives insight into the simulatability of the state. The area law is often studied for ground states of spin systems. It is also known that, initial states that do not violate the area law will not do so over a certain time scale evolving under a local Hamiltonian \( \Gamma \).

Applying our results to a lattice of qubits, we consider the long-time average of the von Neumann entropy of entanglement generated by a time-independent Hamiltonian dynamics \( \langle E(A) \rangle_{T:\infty} \). Using the facts that the von Neumann entropy is lower bounded by the linear entropy and the concavity of the logarithm, we obtain a lower bound on \( \langle E(A) \rangle_{\Upsilon_{\text{phase}}} \), which can be identified with \( \langle E(A) \rangle_{T:\infty} \) under phase ergodicity. Thus, we have

\[
\langle E(A) \rangle_{T:\infty} \geq - \log[1 - \langle E_L^{(A)} \rangle_{\Upsilon_{\text{phase}}}],
\]

where \( \langle E_L^{(A)} \rangle_{\Upsilon_{\text{phase}}} \) is the maximum of \( \langle E(A) \rangle_{T:\infty} \). By applying Eq. 1, we can check the area law in the long-time average.

In particular, we consider Hamiltonians composed of separable eigenstates, which are often referred to as semi-classical. When the initial state is a superposition of separable eigenstates with equal amplitudes, the initial state is also separable, and the corresponding phase-random states are \( \Upsilon_{\text{phase}} \) which obtains the maximum of Eq. 1. Thus we obtain

\[
\langle E(A) \rangle_{T:\infty} \geq N_A - \log(1 + 2^{2N_A-N} - 2^{N_A-N}),
\]

which grows in proportion to the volume \( N_A \) of the subsystem \( A \) when \( N_A \ll N \), and not with any boundary size, and the area law is broken. Since entanglement concentrates around its average during the time evolution as in Eq. 6, the states are not simulatable by MPSs with a constant matrix size for most times. This is surprising at first because all eigenstates as well as the initial state are separable, however the dynamics generate extremely high entanglement and, thus, is difficult to simulate. In Ref. 12, timescales necessary for breaking the area law by time evolutions with local Hamiltonians have been studied. Combined with our result, we can explicitly estimate the timescale necessary for satisfying phase ergodicity when the Hamiltonian is composed of separable eigenstates.

V. PHASE-RANDOM CIRCUIT

We present a phase-random circuit generating an ensemble of states \( \Upsilon_{\text{phase}} \) that provides the same average entanglement as the phase-random ensemble \( \Upsilon_{\text{comp}} = \Upsilon_{\text{phase}} \{ r_n, |a_i \rangle \} \), where \( |\phi_0\rangle = \sum r_n e^{i\phi_n} |a\rangle \) is the input to the circuit and \( \{ |a\rangle \}_{n=1}^{2^N} \) is the computational basis.

A phase-random circuit is similar to those considered in 13 [14]. We consider a circuit composed of \( T \) iterations of two-qubit unitaries diagonal in the computational basis denoted by \( W_t \), where the subscript \( t \) denotes the \( t \)-th iteration \( (t = 1, 2, \ldots, T) \). For each iteration \( t \), the two-qubit unitary \( W_t \) acts on a pair of qubits \( i, j (j \neq i) \) randomly chosen uniformly from \{1, 2, \ldots, \( N \)\}, and is written

\[
W_t = CZ_{i,j}P_t, (\alpha_t)P_t, (\beta_t),
\]

where \( CZ_{ij} = \text{diag}(1, 1, 1, -1) \) is a controlled-Z operation on qubits \( i \) and \( j \), \( P_k (\theta) = \text{diag}(1, e^{i\theta}) \) denotes a phase gate on the qubit \( k \), and the two angles \( \alpha, \beta \) are randomly chosen uniformly from the interval \( [0, 2\pi] \). A specific instance of the circuit is described by the set \( \{ i_t, j_t, \alpha_t, \beta_t \}_{t=1}^T \), and the corresponding output state after \( T \) iterations of \( W_t \) is given by

\[
|\phi_T\rangle = W_T W_{T-1} \cdots W_1 |\phi_0\rangle,
\]

where \( |\phi_0\rangle = \sum r_n e^{i\phi_n} |a\rangle \) is an input state in the computational basis \( \{ |a\rangle \} \), defining the ensemble \( \Upsilon_{\text{comp}} \).
A. Summary of results

Here we state the main results, with the details of the proof to follow. Denote by $\mathbf{E}[E_L^{(A)}(|\varphi_T\rangle)]$ the expectation value of $E_L^{(A)}(|\varphi_T\rangle)$ taken over the uniform distribution of $\{i_t, j_t, \alpha_t, \beta_t\}_{t=1}^T$. We will prove the following two theorems regarding the ability of $\mathbf{E}[E_L^{(A)}(|\varphi_T\rangle)]$ to equal the average of phase random states $\langle E_L^{(A)} \rangle_{\text{phase}}^\text{comp}$ after sufficiently many iterations, and about the required number of iterations.

**Theorem 1** With the preceding definitions and notations,

$$\lim_{T \to \infty} \mathbf{E}[S_L^{(A)}(|\varphi_T\rangle)] = \langle E_L^{(A)} \rangle_{\text{phase}}^\text{comp}, \quad (8)$$

with

$$\langle E_L^{(A)} \rangle_{\text{phase}}^\text{comp} = 1 - \sum_{a,b} r_a^2 r_b^2 \Big( \prod_{i \in A} \delta_{a_i, b_i} + \prod_{i \in A} \delta_{a_i, b_i} \Big) + \sum_{a} r_a^4, \quad (9)$$

where $a_1 a_2 \cdots a_N (b_1 b_2 \cdots b_N) \in \{0, 1\}^N$ is a binary representation of $a - 1 (b - 1)$.

**Theorem 2** Let $T_{\text{mix}}(\epsilon)$ be the number of iterations required to achieve Eq. (8) with error $\epsilon$, namely,

$$\forall T > T_{\text{mix}}(\epsilon), \left| \mathbf{E}[S_L^{(A)}(|\varphi_T\rangle)] - \langle E_L^{(A)} \rangle_{\text{phase}}^\text{comp} \right| < \epsilon.$$ 

For $\Gamma \subset \{1, \cdots, N\}$, define $\kappa^{(\Gamma)}(|\varphi_0\rangle)$ such that

$$\kappa^{(\Gamma)}(|\varphi_0\rangle) = \sum_{a \neq b} r_a^2 r_b^2 \prod_{i \in \Gamma} (1 - \delta_{a_i, b_i}) \prod_{i \not\in \Gamma} \delta_{a_i, b_i}. \quad (10)$$

Then, if $\max_{\Gamma} \kappa^{(\Gamma)}(|\varphi_0\rangle) = O(2^{-N})$, $T_{\text{mix}}(\epsilon)$ is polynomial in the system size $N$ for any $A$. In particular, for $r_a \sim p_a(N)2^{-N/2}$ where $p_a(N)$ are polynomial functions of $N$, $T_{\text{mix}}(\epsilon)$ is poly($N$).

These results are especially interesting if we consider an ensemble $\Upsilon_{\text{phase}}^{\text{pseudo}}$ simulating the average amount of entanglement of $\Upsilon_{\text{phase}}^{\text{eq, sep}}$. Since the average entanglement of $\Upsilon_{\text{phase}}^{\text{eq, sep}}$ violates the area law, most states in $\Upsilon_{\text{phase}}^{\text{pseudo}}$ do also. Hence, $\Upsilon_{\text{phase}}^{\text{pseudo}}$ are not simulatable by MPSs although they are generated by a quantum circuit with a polynomial number of elementary gates.

Here, we have focused on the generation of the average amount of entanglement of phase-random states. In a separate paper [12], it is shown that phase-random circuits can approximately generate an ensemble simulating the states themselves.

In the following, we prove Theorems 1 and 2 by adapting the method developed in [13, 14] to the phase-random case. In this method, the key technique is to map the evolution of the states in the phase-random circuit to a Markov chain, and so we first briefly review Markov processes in Subsection V.B. In Subsection V.C, we present the map to a Markov chain, and then investigate its stationary distribution. As, contrary to [13, 14], the Markov chain is not irreducible in our case, we first decompose it into irreducible Markov chains in Subsection V.D. In order to calculate the average amount of entanglement, it is sufficient to consider reduced Markov chains, which are presented in Subsection V.E. By investigating the stationary distribution of the reduced Markov chain, we finally obtain the average amount of entanglement after $T$ steps in Subsection V.F.

The mixing time $T_{\text{mix}}(\epsilon)$ for achieving Eq. (8) is treated in Subsection V.C.

B. Introduction of a Markov chain

A Markov chain is a sequence of random variables that take values in a set of states $S = \{s\}$, indexed in our case by discrete steps $t$. The Markov property is that the probability of $s_{t+1}$ occurring depends only on $s_t$, and is independent of previous states. We can define at any step $t$ a probability distribution $\Pi_t$ over the states space $S$. The Markov property then ensures that subsequent distributions are related only to the previous distribution, and that this dependence can be given in the form of a step-independent, stochastic transition matrix $\mathcal{P}$, with matrix elements denoted by $\mathcal{P}(s, s')$. Thus, the probability distribution at step $t$ is given by $\Pi_t = \mathcal{P}^t \Pi_0$, where $\Pi_0$ is an initial distribution.

When a Markov chain is irreducible and aperiodic, the probability distribution on each state converges after sufficiently many steps. That is, for all $s$, there exists a unique $\Pi_\infty(s) = \lim_{t \to \infty} \Pi_t(s)$ that is independent of the initial probability distribution. Irreducibility is a property of the transition matrix implying that any state $s$ can transition to any other state in a finite number of steps, that is, for all $s$ and $s'$, there exists a $t$ such that $\mathcal{P}^t(s, s') > 0$. Aperiodicity implies that, for all states $s$, there exists a non-zero probability to remain in that state, namely, $\mathcal{P}(s, s) > 0$ for all $s$. A sufficient condition for a distribution to be stationary is given by the detailed balance condition

$$\Pi(s)\mathcal{P}(s, s') = \Pi(s')\mathcal{P}(s', s), \quad \text{for all } s, s' \in S.$$

When a Markov chain satisfies the detailed balance equations, it is referred to as reversible.

Next, we define the mixing time, which is the number of Markov chain steps required for the distance between the actual distribution and the stationary distribution to be small, where we define the distance between two probability distributions as follows. Let $\delta(s_0)$ be an initial probability distribution of a Markov chain with value 1 at $s_0$ and zero elsewhere on the state space $S$. Let us denote the sum of the probabilities of a distribution over a subset of states $S'$ by $\Pi(S') = \sum_{s \in S'} \Pi(s)$, and by
\[ \Pi_t(S'|\delta(s_0)) \] such a sum at step \( t \) of a Markov chain that initialized with the distribution \( \delta(s_0) \). The variation distance after \( t \)-steps is defined by

\[ \Delta_{s_0}(t) := \max_{S' \subseteq S} |\Pi_t(S'|\delta(s_0)) - \Pi_\infty(S')|. \]

The mixing time \( T_{\text{mix}}(\epsilon) \) is then defined for any \( \epsilon > 0 \) by

\[ T_{\text{mix}}(\epsilon) := \min\{ t \mid \max_{s_0 \in S} \Delta_{s_0}(t') \leq \epsilon \quad \text{for all} \quad t' \geq t \}. \]

This is the number of steps it would take to get \( \epsilon \)-close to the stationary distribution in the worst case. In practice, we do not actually use this definition of the mixing time, but rather the following Theorem and Corollary regarding the transition matrix.

For a transition matrix \( P \) of a reversible Markov chain, let us label the eigenvalues of \( P \) in decreasing order such that

\[ 1 = \lambda_1 > \lambda_2 > \cdots. \]

Then, \( \eta := 1 - \lambda_2 \) is called its absolute spectral gap. The absolute spectral gap \( \eta \) gives an upper bound on the mixing time as stated in the following theorem.

**Theorem 3 (Theorem 12.3 in [16])** Let \( P \) be the transition matrix of a reversible Markov chain on \( S \), and let \( \Pi(\text{min}) := \min_{s \in S} \Pi(s) \). Then

\[ T_{\text{mix}}(\epsilon) \leq \log(\frac{1}{\epsilon \Pi(\text{min})}) \frac{1}{\eta}. \]

Moreover, a lower bound on the absolute spectral gap \( \eta \) is obtained by the canonical path method. Viewing a reversible transition matrix \( P \) as a graph with vertex set \( S \), define the edge set \( E = \{(s, s')|P(s, s') > 0\} \). A canonical path from \( s \) to \( s' \) is a sequence \( \mathcal{E}_{s,s'} = (e_1, \cdots, e_m) \) of edges in \( E \) such that \( e_1 = (s, s_1), e_2 = (s_1, s_2), \cdots, e_m = (s_{m-1}, s') \) for vertices \( s_i, i = 1, 2, \cdots, m \). We have the following Corollary.

**Corollary 1 (Corollary 4 in [17])** For a given transition matrix \( P \), let \( Q(s, s') := \Pi_\infty(s)P(s, s') \) and

\[ \rho := \max_{e \in E} \frac{1}{Q(e)} \sum_{s, s'} \Pi_\infty(s)\Pi_\infty(s'). \]  

Then

\[ \frac{1}{8\rho^2} \leq \eta. \]

By combining Theorem and Corollary an upper bound on the mixing time can be obtained.

C. Map to a Markov chain

We will now show that the change in the state \( |\phi_t\rangle \rightarrow |\phi_{t+1}\rangle \) upon the application of the two-qubit unitary \( W_{t+1} \) defined by Eq. (7) can be formulated in terms of a transition matrix action on the indices of expansion coefficients of the state in the basis of local Pauli operators. The hermiticity of this basis ensures that the coefficients are real, and hence their square gives a valid probability distribution, while its locality ensures that we can focus on the qubits \( i \) and \( j \) where \( W_{t+1} \) acts, eventually simplifying the calculation of the linear entropy.

Let us consider the expansion of \( |\phi_t\rangle\langle\phi_t| \) given by

\[ |\phi_t\rangle\langle\phi_t| = \frac{1}{2N/2} \sum_{q_1, \cdots, q_N} \xi_t(q_1, \cdots, q_N)\sigma_{q_1} \otimes \cdots \otimes \sigma_{q_N}, \]

where \( q_i \in \{0, x, y, z\} \) and \( \sigma_{q_i} \) are Pauli operators. We denote \( (q_1, \cdots, q_N) \) by the vector \( q \). We construct a Markov chain defined on \( \{q\} \) in which the probability distribution is given by the expectation value of \( \xi_t^q(q) \) over \( \alpha_i \) and \( \beta_i \), which is denoted by \( E[\xi_t^q(q)] \). For this purpose, we first examine \( E[\xi_t^q(q)] \) and then construct the Markov chain. For simplicity, hereafter we omit the step indices on qubits and write \( (i, j) \).

By applying \( W_{t+1} \) on a randomly chosen pair of qubits \( (i, j) \), the coefficients \( \xi_{t+1}(q) \) of the state \( |\phi_{t+1}\rangle\langle\phi_{t+1}| \) become

\[ \xi_{t+1}(p) = \frac{1}{4} \sum_{q_i, q_j} \xi_t(p_{i \rightarrow q_i, p_j \rightarrow q_j}) \times \]

\[ \text{Tr}[\sigma_{p_i} \otimes \sigma_{p_j} W_{t+1} \sigma_{q_i} \otimes \sigma_{q_j} W_{t+1}^\dagger], \]

where \( p_{i \rightarrow q_i, p_j \rightarrow q_j} \) is \( p \) but with components \( (p_i, p_j) \) replaced by \( (q_i, q_j) \). Squaring this to arrive at a probability distribution, we have

\[ \xi_{t+1}^2(p) = \frac{1}{4} \sum_{q_i, q_j} \xi_t(p_{i \rightarrow q_i, p_j \rightarrow q_j}) \times \]

\[ \xi_t(p_{i \rightarrow q_i', p_j \rightarrow q_j'}) G_{t+1}(p, q, q'), \]

(11)

where

\[ G_{t+1}(p, q, q') := \]

\[ \text{Tr}[\sigma_{p_i} \otimes \sigma_{p_j} W_{t+1} \sigma_{q_i} \otimes \sigma_{q_j} W_{t+1}^\dagger] \times \text{Tr}[\sigma_{p_i} \otimes \sigma_{p_j} W_{t+1} \sigma_{q_i'} \otimes \sigma_{q_j'} W_{t+1}^\dagger]. \]

In order to see that \( G_{t+1} \) defines a transition matrix, it is important to recognize that it treats the sets of Pauli indices \( \{0, z\} \) and \( \{x, y\} \) equivalently. We write \( w_{0z} \) and \( w_{xy} \) for arbitrary elements of each set respectively, and we define an involution \( \sim \) as \( \sim 0 = z \) and \( \sim x = y \). Averaging
TABLE I: Table of $E[\xi_{t+1}^2(p) | \phi_t]$ as a function of $p_i$ and $p_j$.

| $p_i$ | $p_j$ | $E[\xi_{t+1}^2]$ |
|-------|-------|-----------------|
| 0     | A     | $A_t C_t(0) C_t(0) A_t$ |
|       | x     | $B_t(0) D_t D_t B_t(z)$ |
|       | y     | $B_t(0) D_t D_t B_t(z)$ |
|       | z     | $A_t C_t(z) C_t(z) A_t$ |

TABLE II: Transition probabilities.

| $(q_i, q_j)$ | $(p_i, p_j)$ | Probability |
|--------------|-------------|-------------|
| $(w_{02}, w_{02})$ | $(q_i, q_j)$ | 1 |
| $(w_{02}, w_{xy})$ | $(-q_i, x)$ | 1/2 |
| $(w_{02}, w_{xy})$ | $(-q_i, y)$ | 1/2 |
| $(w_{xy}, w_{02})$ | $(x, -q_j)$ | 1/2 |
| $(w_{xy}, w_{02})$ | $(y, -q_j)$ | 1/2 |
| $(w_{xy}, w_{xy})$ | $(x, x)$ | 1/4 |
| $(w_{xy}, w_{xy})$ | $(x, y)$ | 1/4 |
| $(w_{xy}, w_{xy})$ | $(y, x)$ | 1/4 |
| $(w_{xy}, w_{xy})$ | $(y, y)$ | 1/4 |

For a given $t$, the set of all possible $\Pi_t$ comprise the probability simplex in $V_S$ defined by $\sum \Pi_t(p) = 1$. The transition rules given in Tables I and II define a transition matrix $P$ on $V_S$ with matrix elements written as $P(q, p)$.

We are now prepared to define a Markov chain:

**Definition 1 (Markov chain $M$)** Let $M$ be a Markov chain on a set $S = \{0, x, y, z\}^N = \{q\}$. The transition process is described as follows. In each step, $i$ and $j$ are randomly chosen from $\{1, \cdots, N\}$ and the transition from $q \in S$ to $p \in S$ occurs probabilistically according to Table II. The transition probability from $q$ to $p$ and the probability distribution over $p$ after $t$ steps are denoted by $P(q, p)$ and $\Pi_t(p)$, respectively. The initial distribution $\Pi_0(p)$ is identified with $\xi_0^2(p)$.

**Proposition 1** The probability distribution $\Pi_t(p)$ of the Markov chain $M$ coincides with $E[\xi_{t+1}^2(p) | \phi_t]$. Since the initial distribution of the Markov chain $M$ is given by $\xi_0^2(p)$,

$\Pi_1(p) = \sum \Pi_s(p) \Pi_0(r)$

$= \sum \Pi_s(p) E[\xi_{s+1}^2(r) | \phi_0]$.

where the last equation is obtained using Table I with the definition of the Markov chain $M$. By induction on $t$, Proposition 1 is proven. For example

$\Pi_2(p) = \sum \Pi_s(p) \Pi_1(r)$

$= \sum \Pi_s(p) E[\xi_{s+1}^2(r) | \phi_0]$

$= E\{\sum \Pi_s(p) \xi_{s+1}^2(r) | \phi_0\}$

$= E[\xi_{t+1}^2(p) | \phi_0]$. We recall that a probability distribution $\Pi_t$ can be viewed as a vector in a $4N$-dimensional space, which we’ll call $V_S$, where $S = \{0, x, y, z\}^N$. For a given $t$, the set of all possible $\Pi_t$ comprise the probability simplex in $V_S$ defined by $\sum \Pi_t(p) = 1$. The transition rules given in Tables I and II define a transition matrix $P$ on $V_S$ with matrix elements written as $P(q, p)$.

**D. Irreducible decomposition of the Markov chain**

In this subsection, we give the irreducible decomposition of $V_S$. By the definition of the Markov chain $M$, it is obvious that the number of $x$ and $y$ in $q$ is invariant under the action of the transition matrix $P$. Thus we obtain the irreducible decomposition of $V_S$ given by Proposition 2 (see also Fig. 3).

**Proposition 2** (Irreducible decomposition of $S$)
For $q = q_1 q_2 \cdots q_N \in S = \{0, x, y, z\}^N$, let $X(q)$ be the sequence $i \in \{0, x, y, z\}^N$ and let $S(\Gamma)$ be the set defined by

$S(\Gamma) := \{q | X(q) = \Gamma\}$,
where $\Gamma$ is any subset of $\{1,2,\cdots,N\}$. Then, for the Markov chain $\mathcal{M}$, the irreducible decomposition of $V_S$ is given by

$$V_S = \bigoplus_{q \in S(\emptyset)} V(q) \oplus V_S(\Gamma),$$

where $V_S'$ is the vector space defined by the subset $S'$.

Since $V_S(q)$ is always one dimensional by definition, we have that $\Pi_i(q) = \Pi_0(q)$ for all $q \in S(\emptyset)$ and for all $t$. Thus $\Pi_i(q \in S(\emptyset))$ is given by

$$\Pi_i(q) = \Pi_0(q) = \xi_0(q)^2 = 2^{-N} \langle \phi_0 | \sigma_q | \phi_0 \rangle^2$$

$$= 2^{-N} \sum_{a,b} \rho_{a,b}^2 \prod_{i=1}^{N} \left( \delta_{q_i,0} + \delta_{q_i,1}(1-2a_i)(1-2b_i) \right),$$

(13)

where we have used the fact that, for $q \in S(\emptyset)$, $q_i \in \{0,1\}$ for all $i$ and $\sigma_q := \sigma_{q_1} \otimes \cdots \otimes \sigma_{q_N}$.

### E. Reduction of the Markov chain

In order to describe the evolution of $E_L^{(A)}(\rho_t)$, a full investigation of the Markov chain $\mathcal{M}$ is not necessary due to the definition of the linear entropy $S_L(\rho) = 1 - \text{Tr} \rho^2$. This can be seen by considering the reduced density matrix of $|\phi_t\rangle$ on a subsystem $A$.

$$\hat{\rho}_A^{(t)} = \text{Tr}_{\bar{A}} |\phi_t\rangle \langle \phi_t|$$

$$= \frac{1}{2^{N/2}} \sum_q \xi(t) \langle q | \text{Tr}_{\bar{A}} \sigma_q,$$

and $\text{Tr} \hat{\rho}_A^{(t)^2}$ is given by

$$\text{Tr} \hat{\rho}_A^{(t)^2} = 2^{N_A} \sum_{q \text{ s.t. } q_i=0, i \in \bar{A}} \xi(t) \langle q | \sigma_q.$$

Hence, its expectation value is

$$E[\text{Tr} \hat{\rho}_A^{(t)^2} | \phi_0] = 2^{N_A} \sum_{q \text{ s.t. } q_i=0, i \in \bar{A}} \Pi_i(q).$$

Thus, it is sufficient to investigate $\Pi_i(q)$ for $q$ such that $q_i = 0$ for $i \in \bar{A}$. The only important property is the number of non-zero terms in $q$. For this reason, let us define the set $\chi^{(t)}(q)$ as

$$\chi^{(t)}(q) := \{ i | i \in \{1,\cdots,N\} | q_i \neq 0, q_i \in S(\emptyset) \},$$

which indicates the positions of non-zero terms in $q \in S(\emptyset)$. Using this notation, the expectation value is written by

$$E[\text{Tr} \hat{\rho}_A^{(t)^2} | \phi_0] = 2^{N_A} \sum_{\Gamma \subset \mathcal{A}} \sum_{q \text{ s.t. } |\chi^{(t)}(q)| = \Gamma} \Pi_i(q).$$

(14)

Since $\Pi_i(q)$ for $q \in S(\emptyset)$ is already given by Eq. (13), we consider only $q \in S(\Gamma)$ for $\Gamma \neq \emptyset$.

For this reason, we can reduce the Markov chain $\mathcal{M}$ to a simpler Markov chain $\tilde{\mathcal{M}}$. For a given number of $x$ or $y$ transitions $\gamma := |\Gamma|$ in $q$, the number of non-zero elements $|\chi^{(t)}(q)|$ can take values $\{\gamma,\gamma+1,\cdots,N\}$. The new Markov chain is a drunkard’s walk on this set (see Fig. 4), with transition probabilities given by the following proposition.

#### Proposition 3

For the Markov chain $\tilde{\mathcal{M}}$ defined on $\{i \in \{\gamma,\cdots,N\}\}$, a transition from $i$ to $j$ occurs with probability,

$$p^{(t)}(i,j = j + 1) = \frac{2\gamma(N-i)}{N(N-1)},$$

$$p^{(t)}(i,j = i - 1) = \frac{2\gamma(i-\gamma)}{N(N-1)}$$

$$p^{(t)}(i,j = i) = \frac{\gamma(\gamma-1) + (N-\gamma)(N-\gamma-1)}{N(N-1)}.$$

This is directly induced from the definitions of the Markov chain $\mathcal{M}$ and $\tilde{\mathcal{M}}$. For instance, the transition $i \rightarrow i+1$ in $\tilde{\mathcal{M}}$ occurs if and only if $(q_i,q_j) = (w_{x\beta},0)$.
or \((q_i, q_j) = (0, w_{xy})\) in \(M\). As the number of zeroes is \(N - i\) and the number of \(w_{xy} = \gamma\), its probability is given by \(\frac{\gamma(N - \gamma)}{N(N - 1)/2}\).

As stated in the introductory subsection, since the Markov chain \(M_\Gamma\) is irreducible and aperiodic, it has a unique stationary distribution \(\Pi_{\infty}^{(\Gamma)}\), which is determined by the detailed balance condition and the normalization. The detailed balance condition gives the equation

\[
\Pi_{\infty}^{(\Gamma)}(i) P^{(\Gamma)}(i, i + 1) = \Pi_{\infty}^{(\Gamma)}(i + 1) P^{(\Gamma)}(i + 1, i).
\]

Using this equation, we obtain

\[
\Pi_{\infty}^{(\Gamma)}(i) = \left( \frac{N - \gamma}{i - \gamma} \right) \Pi_{\infty}^{(\Gamma)}(\gamma). \tag{15}
\]

The normalization in \(S(\Gamma)\) depends on the input state \(|\phi_0\rangle\) as

\[
\sum_{i=\gamma}^{N} \Pi_{\infty}^{(\Gamma)}(i) = \sum_{q \in S(\Gamma)} \Pi_0(q) = \sum_{q \in S(\Gamma)} \xi_0^2(q). \tag{16}
\]

On the other hand, Eq. (15) gives

\[
\sum_{i=\gamma}^{N} \Pi_{\infty}^{(\Gamma)}(i) = 2^{N-\gamma} \Pi_{\infty}^{(\Gamma)}(\gamma).
\]

Hence, the stationary distribution is given by

\[
\Pi_{\infty}^{(\Gamma)}(i) = \frac{1}{2^{N-\gamma}} \left( \frac{N - \gamma}{i - \gamma} \right) \sum_{q \in S(\Gamma)} \xi_0^2(q).
\]

Recalling that \(\xi_0^2(q) = \langle \phi_0 | \sigma_q | \phi_0 \rangle\), it is not difficult to compute \(\sum_{q \in S(\Gamma)} \xi_0^2(q)\), which gives

\[
\sum_{q \in S(\Gamma)} \xi_0^2(q) = \sum_{a \neq b} r_a^2 r_b^2 \prod_{i \in \Gamma} (1 - \delta_{a,b_i}) \prod_{i \notin \Gamma} \delta_{a,b_i}.
\]

Thus the stationary distribution for a given subset \(\Gamma\) is

\[
\Pi_{\infty}^{(\Gamma)}(i) = \frac{1}{2^{N-\gamma}} \left( \frac{N - \gamma}{i - \gamma} \right) \sum_{a \neq b} r_a^2 r_b^2 \prod_{i \in \Gamma} (1 - \delta_{a,b_i}) \prod_{i \notin \Gamma} \delta_{a,b_i}. \tag{17}
\]

### F. Calculation of \(\lim_{T \to \infty} \mathbb{E}[S_L^{(A)}(|\phi_T\rangle)]\)

We will now calculate the large time limit of the expectation value of the amount of entanglement \(\lim_{T \to \infty} \mathbb{E}[S_L^{(A)}(|\phi_T\rangle)]\) using the results of the previous two subsections. From Eq. (17) we have

\[
\mathbb{E}[\text{Tr}(\hat{\phi}_A^{(T)})^2 | |\phi_0\rangle\rangle] = 2^{N_A} \sum_{\Gamma \subseteq A} \sum_{q \text{ s.t. } \chi^{(\Gamma)}(q) \subseteq A} \Pi_T(q)
\]

\[
= 2^{N_A} \left[ \sum_{q \text{ s.t. } A \supseteq \Gamma \neq \emptyset} \sum_{q \text{ s.t. } \chi^{(\Gamma)}(q) \subseteq A} \right] \Pi_T(q). \tag{18}
\]

The first term in Eq. (18) is calculated from Eq. (13) as

\[
\sum_{q \text{ s.t. } \chi^{(\Gamma)}(q) \subseteq A} \chi^{(\Gamma)}(q) = 2^{-\gamma} \sum_{a \neq b} r_a^2 r_b^2 \prod_{r \in A} (1 - \delta_{a,b_i}) \prod_{i \notin \Gamma} \delta_{a,b_i},
\]

where the last expression is derived from the relation

\[
\sum_{\Gamma \subseteq A} \chi^{(\Gamma)}(q) = 2^{N_A} \sum_{i \in A} \delta_{a,b_i}.
\]

The second term in Eq. (18) is obtained from the stationary distributions \(\Pi_{\infty}^{(\Gamma)}(i)\) given by Eq. (17). From the definition of the Markov chain \(M_\Gamma\), for \(q, q' \in S(\Gamma)\), if the number of \(z\) in \(q\) is equal to that in \(q'\), \(\Pi_{\infty}(q) = \Pi_{\infty}(q')\), so that

\[
\sum_{A \supseteq \Gamma \neq \emptyset} \sum_{q \text{ s.t. } \chi^{(\Gamma)}(q) \subseteq A} \Pi_{\infty}(q) = \sum_{A \supseteq \Gamma \neq \emptyset} \sum_{q \text{ s.t. } \chi^{(\Gamma)}(q) \subseteq A} \Pi_{\infty}^{(\Gamma)}(i)
\]

\[
= \sum_{A \supseteq \Gamma \neq \emptyset} \sum_{q \text{ s.t. } \chi^{(\Gamma)}(q) \subseteq A} \Pi_{\infty}^{(\Gamma)}(i)
\]

\[
= \sum_{A \supseteq \Gamma \neq \emptyset} \sum_{q \text{ s.t. } \chi^{(\Gamma)}(q) \subseteq A} N_A^{\chi^{(\Gamma)}(q)} \prod_{i \in \Gamma} \delta_{a,b_i} \prod_{i \notin \Gamma} (1 - \delta_{a,b_i})
\]

\[
= 2^{-N_A} \sum_{a \neq b} r_a^2 r_b^2 \prod_{i \in \Gamma} \delta_{a,b_i},
\]

where we have used the relation

\[
\sum_{A \supseteq \Gamma \neq \emptyset} \prod_{i \in \Gamma} (1 - \delta_{a,b_i}) = -\delta_{a,b} + \prod_{i \in A} \delta_{a,b_i}.
\]

Combining the two we arrive at the final expression

\[
\lim_{T \to \infty} \mathbb{E}[\text{Tr}(\hat{\phi}_A^{(T)})^2 | |\phi_0\rangle\rangle] = \sum_{a,b} r_a^2 r_b^2 \prod_{i \in A} \delta_{a,b_i} + \prod_{i \in A} \delta_{a,b_i} - \sum_a r_a^4,
\]

and since \(E^{(A)}_{L}(|\phi_T\rangle) = 1 - \text{Tr}(\hat{\phi}_A^{(T)})^2\), Eq. (1) is obtained.

#### G. Mixing time

In this final subsection we bound the mixing time of the Markov chain \(M_\Gamma\) using Theorem 3 and Corollary 1.
From Eq. (14), the stationary distribution in $S(\Gamma)$ is given by

$$\Pi_\infty^{(i)}(i) = \frac{1}{2N-\gamma} \left( \frac{N-\gamma}{i-\gamma} \right) \kappa_i^{(T)}(|\phi_0\rangle),$$

where we have introduced the notation $\kappa_i^{(T)}(|\phi_0\rangle) := \sum_{a \neq \gamma} r_{2\gamma}^2 \prod_{i \in T} (1 - \delta_{a, b_i}) \prod_{i \notin T} \delta_{a, b_i}$. Thus $\Pi_\infty^{(i)}$ is given by

$$\Pi_\infty^{(i)}(i) = \frac{1}{2N-\gamma} \kappa_i^{(T)}(|\phi_0\rangle).$$

Let $\rho^{(T)}$ be the expression defined by Eq. (10) for the Markov chain $\hat{M}_T$. An upper bound of $\rho^{(T)}$ for $\hat{M}_T$ is then given by

$$\rho^{(T)} \leq \frac{\max_{e \in E} \sum_{i,j} \Pi_\infty^{(i)}(i) \Pi_\infty^{(j)}(j)}{\min_{e \in E} Q^{(T)}(\epsilon)}.$$

Since the graph of the Markov chain $\hat{M}_T$ is linear, as shown in Fig. 8, an upper bound on the maximum of $\sum_{i,j} \Pi_\infty^{(i)}(i) \Pi_\infty^{(j)}(j)$ is given by

$$\sum_{i,j} \Pi_\infty^{(i)}(i) \Pi_\infty^{(j)}(j) = \left( \frac{\kappa_i^{(T)}(|\phi_0\rangle)}{2N-\gamma} \right)^2 \max_{i \in \{1, \ldots, N\}} \sum_{j=1}^{N} \frac{(N-\gamma)}{i-\gamma} \sum_{j=1}^{N} (N-j)^{-\gamma} \leq \left( \frac{\kappa_i^{(T)}(|\phi_0\rangle)}{2N-\gamma} \right)^2 \max_{i \in \{1, \ldots, N\}} \sum_{j=0}^{N} \frac{(N-j)}{i-\gamma} \sum_{j=0}^{N} (N-j)^{-\gamma} = (\kappa_i^{(T)}(|\phi_0\rangle))^2.$$

On the other hand, the $\min_{e \in E} Q^{(T)}(\epsilon)$ factor can be computed as follows. In the Markov chain $\hat{M}_T$, edges are of the form $(i, i+1)$ or $(i, i-1)$. By the symmetry of the linear graph, one sees that $Q^{(T)}(N+\gamma-i, N+\gamma-i-1) = Q^{(T)}(i, i+1)$, and it is sufficient to consider the minimum of $Q^{(T)}(i, i+1)$, which is given by

$$\min_i Q^{(T)}(i, i+1) = \frac{2\gamma(N-\gamma)}{2N-\gamma N(N-1)} \kappa_i^{(T)}(|\phi_0\rangle).$$

Thus $\rho^{(T)}$ is bounded from above as

$$\rho^{(T)} \leq \frac{2N-\gamma N(N-1)}{2\gamma(N-\gamma)} \kappa_i^{(T)}(|\phi_0\rangle).$$

In order to achieve

$$\forall T > T_{\text{mix}}(\epsilon), \left| E[\dot{E}_L^{(A)}(|\phi_T\rangle)] - \langle \dot{E}_L^{(A)} \rangle_{\text{T-phase comp}} \right| < \epsilon$$

for all $A$, it is sufficient for each Markov chain $\hat{M}_T$ to converge with error $\epsilon' := \epsilon/2^N$ since the linear entropy is the sum of the stationary distributions in $\hat{M}_T$ as shown by Eq. (15). Therefore, from Theorem 3 and Corollary 1, we obtain an upper bound on $T_{\text{mix}}(\epsilon)$ given by

$$T_{\text{mix}}(\epsilon) \leq \max_T \left[ \frac{N(N-1)}{2\gamma(N-\gamma)} \frac{2N-\gamma \kappa_i^{(T)}(|\phi_0\rangle)}{2N-\gamma N(N-1)} \right]^2 \times \left[ N-\gamma - \log \left( \frac{\epsilon}{2N} \cdot \kappa_i^{(T)}(|\phi_0\rangle) \right) \right].$$

This is dominated by the factor $2^{N-\gamma \kappa_i^{(T)}(|\phi_0\rangle)}$. Thus, $\max_T \kappa_i^{(T)}(|\phi_0\rangle) = O(2^{-N})$ is sufficient for $T_{\text{mix}}(\epsilon)$ to be a polynomial in $N$.

This concludes the proof.

VI. SUMMARY

We have defined phase-random states as an ensemble of states with fixed amplitudes and with uniformly distributed phases in a fixed basis. We have discussed their use for the realization of canonical distributions in statistical mechanics. We then derived a general formula for the average amount of entanglement of phase-random states. Applying these results, we have argued for the simulatability of time evolving states by a Hamiltonian dynamics, and have shown the difficulty of their simulation for semi-classical Hamiltonian systems by MPSs. Finally, we have proven that an ensemble of states that provides the same average entanglement of phase-random states can be generated efficiently by a phase-random circuit composed of relatively simple gates.

We acknowledge V. Vedral for useful comments. This work is supported by Project for Developing Innovation Systems of MEXT, Japan and JSPS by KAKENHI (Grant No. 222812, No. 23540463 and 23240001).

Appendix A: Concentration of measure

In this appendix, we show that for the ensemble of phase-random states with equal-amplitudes in a separable basis, $T_{\text{phase}}^{\text{eq sep}}$, the amount of entanglement is highly concentrated around the average. Formally, by defining

$$\Delta E_L^{(A)}(|\phi_{\text{eq, sep}}\rangle) := | E_L^{(A)}(|\phi_{\text{eq, sep}}\rangle) - \langle E_L^{(A)} \rangle_{\text{T-phase comp}} |$$

where $|\phi_{\text{eq, sep}}\rangle = 2^{-N/2} \sum_{a=1}^{2^N} e^{i\theta_a} |u_{\text{a}}\rangle$, we prove that

$$\text{Prob} \left( \Delta E_L^{(A)}(|\phi_{\text{eq, sep}}\rangle) > 2^{-N/2} + \epsilon \right) \leq \exp \left( -c \epsilon^4 2^{-N} \right)$$

where $c = 1/(2^{10} \pi^2)$.

First, for two states in the ensemble $\text{T-phase}$ denoted by $|\phi\rangle = 2^{-N/2} \sum_{a=1}^{2^N} e^{i\theta_a} |u_{\text{a}}\rangle$ and $|\phi'\rangle = 2^{-N/2} \sum_{a=1}^{2^N} e^{i\theta'_a} |u_{\text{a}}\rangle$, let us define the distance $d(\phi, \phi')$ between them in the parameter space $[0, 2\pi)^{2N}$ by

$$d(\phi, \phi') = \frac{1}{2\pi} \sum_{a=1}^{2^N} |\theta_a - \theta'_a|.$$
Then, using the theorems in Appendix C of Ref. [1], we obtain the upper bound of the concentration function $\alpha_d(r)$ by

$$\alpha_d(r) \leq \exp\left[-\frac{r^2}{8}2^N\right], \quad (A1)$$

where the concentration function $\alpha_d(r)$ implies that, for any subset $A \in [0,2\pi]^2N$ with measure $1/2$, its $r$-neighborhood $A_r$ with respect to the metric $d$ has measure at least $1 - \alpha_d(r)$.

Now, we evaluate the amount of the change in the parameter space necessary to change $\Delta E_L^{(A)}(\phi^{eq,sep})$ more than $\epsilon$, which is obtained from the following proposition;

**Proposition 4** For $|\phi\rangle$ and $|\phi'\rangle \in \Upsilon^{eq,sep}$,

$$|\Delta E_L^{(A)}(|\phi\rangle) - \Delta E_L^{(A)}(|\phi'\rangle)| \leq 4\sqrt{\pi}d(\phi,\phi'). \quad (A2)$$

**Proof 1** Using the notation $\hat{\phi}_A = \text{Tr}_A|\phi\rangle\langle\phi|$, we calculate

$$|\Delta E_L^{(A)}(|\phi\rangle) - \Delta E_L^{(A)}(|\phi'\rangle)| \leq |\text{Tr}\hat{\phi}_A - \text{Tr}\hat{\phi}'_A|$$

$$= |\text{Tr}(\hat{\phi}_A - \hat{\phi}'_A)(\hat{\phi}_A + \hat{\phi}'_A)|$$

$$\leq \|\hat{\phi}_A - \hat{\phi}'_A\|_2\|\hat{\phi}_A + \hat{\phi}'_A\|_2$$

$$\leq 2D_{HS}(\hat{\phi}_A,\hat{\phi}'_A)$$

$$\leq 2D_{HS}(|\phi\rangle\langle\phi|,|\phi'\rangle\langle\phi'|)$$

$$\leq 2\|\phi - \phi'\|_2,$$  \quad (A3)

where $|A|_2 := \sqrt{\text{Tr}AA^t}$ is the Hilbert-Schmidt norm and $D_{HS}(A,B) = \|A - B\|_2$. Inequalities $|A\rangle$ and $|A\rangle$ are obtained using Cauchy-Schwarz and Kadison’s inequalities [19] respectively. Since $|\langle\phi| - \langle\phi'| \leq 4\sqrt{\pi}d(\phi,\phi')$, we obtain Eq. (A2).

Hence, in order to change $\Delta E_L^{(A)}(\phi^{eq,sep})$ more than $\epsilon$, $d(\phi,\phi')$ must be changed more than $\epsilon^2$. Combining this with the concentration of measure given by (A1), we obtain

$$\text{Prob}(\Delta E_L^{(A)}(\phi^{eq,sep})) > \mu_M(\Delta E_L^{(A)}(\phi^{eq,sep})) + \epsilon \leq \exp[-c\epsilon^42^N], \quad (A5)$$

where $\mu_M$ represents the median and $c = 1/(2^{11}\pi^2)$. By using Markov’s inequality and the convexity of $\sqrt{x}$, the median is bounded from above such that

$$\mu_M(\Delta E_L^{(A)}(\phi^{eq,sep})) \leq 2(\Delta E_L^{(A)})^2_{\Upsilon^{eq,sep}}$$

$$= 2\left(\sqrt{(\Delta E_L^{(A)})^2_{\Upsilon^{eq,sep}}}\right)^2$$

$$\leq 2\left(\Delta E_L^{(A)}(\phi^{eq,sep})\right)^2_{\Upsilon^{eq,sep}}$$

$$= 2\sigma E_L^{(A)},$$

where $\sigma E_L^{(A)}$ is the standard deviation of $E_L^{(A)}$. Since the standard deviation $\sigma E_L^{(A)}$ for $\Upsilon^{eq,sep}$ can be directly calculated and is upper bounded by $2^{-N}$, we obtain

$$\text{Prob}\left[\Delta E_L^{(A)}(\phi^{eq,sep}) > \frac{2}{2^N} + \epsilon\right] \leq \exp[-c\epsilon^42^N].$$

[1] N. Linden, S. Popescu, A. J. Short and A. Winter, Phys. Rev. E 79, 061103 (2009).
[2] P. Hayden and J. Preskill, JHEP 0709:120 (2007).
[3] C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor and A. Winter, IEEE Trans. Inform. Theory, vol. 51, no. 1, pp 56-74 (2005); B. M. Terhal, David P. DiVincenzo and D. W. Leung, Phys. Rev. Lett. 86, 5807-5810 (2001); P. Hayden, D. Leung, P. W. Shor and A. Winter Commun. Math. Phys. 250(2):371-391(20040).
[4] T. N. Ikeda, Y. Watanebe and M. Ueda, Phys. Rev. E 84, 021130 (2011).
[5] E. Lubkin, J. Math. Phys. 19 1028 (1978); D. N. Page, Phys. Rev. Lett., 71:1291, (1993); S. K. Foong and S. Kanno, Phys. Rev. Lett. 72, 1148 (1994); P. Hayden, D. W. Leung and A. Winter, Comm. Math. Phys. Vol. 265, No. 1, pp. 95-117 (2006).
[6] S. Goldstein, J. L, Lebowitz, R. Tumulka and N. Zanghii Phys. Rev. Lett. 96, 050403 (2006); S. Popescu, A. J. Short and A. Winter, Nature Physics, 2:754-758 (2006).
[7] S. Goldstein, J. L, Lebowitz, R. Tumulka and N. Zanghii, European Phys. J. H 35: 173-200 (2010).
[8] D. Shepherd and M. J. Bremmer, Proc. R. Soc. A 465, 1413-1439 (2009); M. J. Bremner, R. Jozsa and D. J. Shepherd, Proc. R. Soc. A 467, 459 (2011).
[9] C. Kruszynska and B. Kraus, Phys. Rev. A 79, 052304 (2009).
[10] J. Eisert, M. Cramer and M.B. Plenio, Rev. Mod. Phys. 82, 277 (2010).
[11] D. A. Meyer and N. R. Wallach, J. Math. Phys. 43 4273 (2002).
[12] S. Bravyi, M. B. Hastings and F. Verstraete, Phys. Rev. Lett. 97, 050401 (2006).
[13] R. Oliveira, O. C. O. Dahlsten and M. B. Plenio, Phys. Rev. Lett. 97, 050401 (2006).
[14] R. Oliveira, O. C. O. Dahlsten and M. B. Plenio, Phys. Rev. Lett. 97, 050401 (2006).
[15] V. Nakata and M. Murao, arXiv:1206.4451 (2012).
[16] D. A. Levin, Y. Peres and E. L. Wilmer, Markov Chains and Mixing Times, American Mathematical Society, Providence (2009).
[17] A. Sinclair, Comb. Prob. Comp. 1, 351-370 (1992).
[18] M. Leudox, The Concentration of Measure Phenomenon, AMS Monographs, Providence, RI, Vol. 89 (2001).
[19] R. V. Kadison, Ann. of Math. 56 (1952) 494.