Emergent Quantum State Designs from Individual Many-Body Wave Functions

Jordan S. Cotler,[1,‡] Daniel K. Mark,[2,‡] Hsin-Yuan Huang,[3,‡] Felipe Hernández,[4] Joonhee Choi,[3] Adam L. Shaw,[3] Manuel Endres,[3,*] and Soonwon Choi[2,5,†]

[1] *Society of Fellows, Harvard University, Cambridge, Massachusetts 02138, USA*
[2] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3] *California Institute of Technology, Pasadena, California 91125, USA*
[4] *Department of Mathematics, Stanford University, Stanford, California 94305, USA*
[5] *Department of Physics, University of California, Berkeley, California 94720, USA*

(Received 22 June 2021; revised 18 June 2022; accepted 21 December 2022; published 27 January 2023; corrected 17 May 2023)

Quantum chaos in many-body systems provides a bridge between statistical and quantum physics with strong predictive power. This framework is valuable for analyzing properties of complex quantum systems such as energy spectra and the dynamics of thermalization. While contemporary methods in quantum chaos often rely on random ensembles of quantum states and Hamiltonians, this is not reflective of most real-world systems. In this paper, we introduce a new perspective: across a wide range of examples, a single nonrandom quantum state is shown to encode universal and highly random quantum state ensembles. We characterize these ensembles using the notion of quantum state k-designs from quantum information theory and investigate their universality using a combination of analytic and numerical techniques. In particular, we establish that k-designs emerge naturally from generic states in a Hilbert space as well as physical states associated with strongly interacting Hamiltonian dynamics. Our results offer a new approach for studying quantum chaos and provide a practical method for sampling approximately uniformly random states; the latter has wide-ranging applications in quantum information science from tomography to benchmarking.

DOI: 10.1103/PRXQuantum.4.010311

I. INTRODUCTION

Analysis of the exact dynamics of general strongly interacting quantum many-body systems is intractable using existing analytic and numerical tools. However, there is a widely used heuristic for understanding chaotic quantum dynamics: the eigenstates and eigenvalues of chaotic Hamiltonians have properties as if they were sampled from a random ensemble. This heuristic leads one to leverage statistical approaches, such as random matrix theory [1,2], to address many physical problems and has become a foundational principle in understanding chaos and thermalization in quantum systems [3–6]. Examples of this heuristic include Berry's conjecture [7] and the eigenstate-thermalization hypothesis (ETH) [8–10], which have been supported by an overwhelming amount of numerical evidence [11–13].

While these statistical approaches hinge on the presence of random ensembles, realistic quantum systems are often described by a *fixed* Hamiltonian. Crucially, physical Hamiltonians are very special and highly structured from a mathematical and statistical point of view, as they must satisfy multiple constraints imposed by locality and energy conservation. Thus, it is a surprising empirical finding that deterministic functions of physical Hamiltonians, such as eigenstates, eigenvalues, and states evolved by Hamiltonian dynamics, can still be well characterized by typical instances of random ensembles. It remains as a fundamental question to establish a connection between isolated quantum systems and the emergence of random ensembles that dictate their statistical properties.

In this paper, we present a new perspective on the emergence of statistical behavior in chaotic quantum many-body systems: instead of imagining that a physical state is sampled from a random ensemble, we use a single many-body wave function to *generate* an ensemble of pure states on a subsystem. This approach goes beyond using the reduced density matrix, as the ensemble encodes higher-order statistical moments of the wave functions on

*mendres@caltech.edu
†soonwon@mit.edu
‡These authors contributed equally to this work.

the subsystem, as well as certain types of correlations between the subsystem and its complement. In contrast to using reduced density matrices, our approach allows the characterization of information-theoretic properties of the subsystem that cannot be otherwise described by expectation values of conventional observables.

Applying our approach to a wide class of examples, we discover a novel type of universality: under robust conditions, the generated ensemble approaches a universal value that follows the maximally entropic distribution over the subsystem Hilbert space. This universality establishes a connection between quantum many-body dynamics and information theory, with several implications. In quantum chaos and thermalization, our findings suggest that physical quantum many-body systems drive not only local observables to thermal values but also other information-theoretic properties of a subsystem (encoded in higher statistical moments) to certain universal values. This presents a generalized form of "thermalization" manifested in higher moments of quantum states. In quantum information science, the generation of maximally random pure states is known to be useful [14–19] but computationally difficult [20]. Here, such random states arise approximately in the natural setting of quantum many-body Hamiltonians with local interactions, independent of their microscopic details. Our results provide a new hardware-efficient way to generate pseudorandom quantum states. Finally, in quantum nonequilibrium dynamics, exotic phase transitions induced by measurements have been recently discovered. This manifests in higher-moment statistical properties of quantum states such as the entanglement, state purity, and variance in charge [21–24]. The statistical nature of these transitions necessitates considering an ensemble of random evolutions. The novel perspective presented here allows the analysis of statistical properties of quantum states from a single time-independent Hamiltonian evolution, providing a new framework to analyze this new physics.

Our key idea is to note that an ensemble of states can be generated from a single wave function by performing local measurements over only part of the total system. Concretely, we consider a many-body system partitioned into a subsystem A and its complement B. Performing local measurements on B, we obtain exponentially many different pure states $|\Psi_A(z_B)\rangle$ on A, each corresponding to a distinct measurement outcome z_B on B. We call the set of pure states on A, along with the associated measurement probabilities $p(z_B)$, the *projected ensemble* on A [see Figs. 1(a) and 1(b)]; see also Ref. [25].

To quantify the degree of randomness of an ensemble, we use a well-established notion from quantum information theory, namely *quantum state k-designs* [26,27]. An ensemble of pure states is said to form a quantum state k-design if its first k moments agree with those of the uniform distribution over the entire Hilbert space. Forming a higher k-design implies that the ensemble is more uniformly distributed over the Hilbert space. In addition to quantitatively characterizing the randomness of an ensemble, k-designs are also relevant for applications that only require randomness up to the first k moments [14–19].

Our claimed universality is that approximate k-designs arise from a variety of quantum many-body states. We establish two theorems showing that the projected ensemble coming from a generic many-body quantum state forms an approximate k-design as long as the size of B is sufficiently larger than the size of A.

Furthermore, a concurrent work [25] finds evidence that approximate k-designs emerge from projected ensembles in a Rydberg quantum simulator. We argue that this is a much more general phenomenon: we find strong numerical evidence across several models that approximate k-designs arise from both quantum states obtained by quenched time evolution and energy eigenstates of chaotic Hamiltonians [see Fig. 1(c)]. In the former case, we find that the degree of randomness in a projected ensemble continues to grow even after conventional local thermalization has occurred. Specifically, we establish numerically that states evolved for a longer amount of time form higher approximate k-designs. We also find that a subsystem forms a projected k-design in a time independent of global system size, suggesting that the same phenomenology persists in the thermodynamic limit. In particular, this implies that a projected k-design can form from global wave functions that are not typical instances of random states, going beyond the predictions of our rigorous theorems.

The above results hold only for the case of infinite-temperature thermalization. We then study energy eigenstates at infinite and finite effective temperatures. We find that approximate k-designs emerge from states in the middle of the energy spectrum corresponding to effective infinite temperature. For finite-temperature eigenstates, we observe that the projected ensembles converge to a universal ensemble that smoothly varies with respect to the energy density.

Our findings suggest that for a wide range of physically relevant many-body states, projected ensembles exhibit a universal form of randomness. This allows us to quantify the chaotic nature of Hamiltonian dynamics and to study the growth of complex nonlocal correlations between a subsystem and its complement beyond the conventional paradigm of quantum thermalization [8–11,13,28–30]. Postthermalization physics in quantum many-body systems exhibits interesting quantitative and qualitative differences from its classical counterpart (see, e.g., Ref. [31]) and the current work clarifies and extends this paradigm. Furthermore, our work presents a simple protocol to efficiently produce an ensemble of random pure states from fixed time-independent Hamiltonian dynamics that does not require highly fine-tuned controls. This may open up new possibilities for utilizing analog quantum simulators

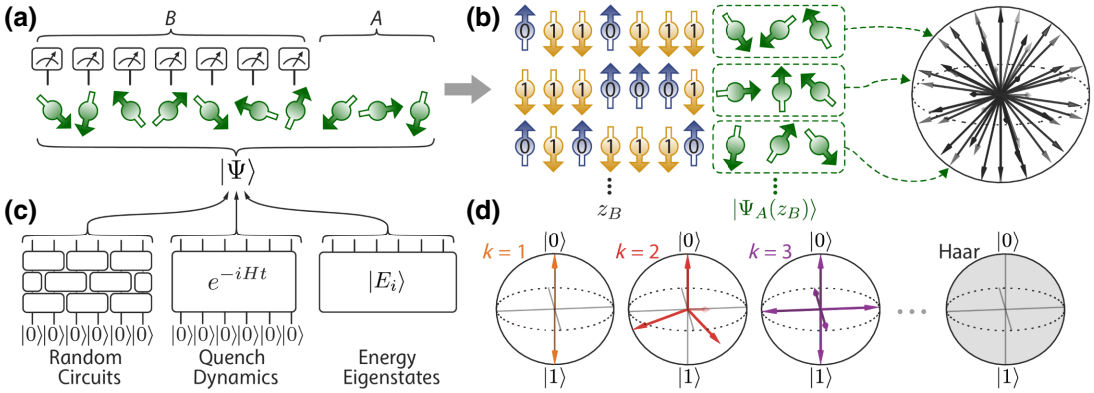

FIG. 1. The emergence of a universal quantum state ensemble from a single many-body wave function. (a) A subsystem B of a pure many-body wave function $|\Psi\rangle$ is measured in a fixed local basis; the remaining unmeasured qubits in A are in a pure state that depends on the measurement outcome on B. (b) For quantum systems consisting of qubits, the measurement on B samples random outcomes, each characterized as a bit string z_B (binary numbers in the blue and yellow arrows). Different measurement outcomes z_B occur with probability $p(z_B)$ and lead to distinct quantum states $|\Psi_A(z_B)\rangle$, forming the projected ensemble $\mathcal{E} = \{p(z_B), |\Psi_A(z_B)\rangle\}$. In the right panel, the ensemble of pure states $|\Psi_A(z_B)\rangle$ (black arrows) is randomly distributed in the Hilbert space of A (black sphere), forming an approximate quantum state k-design. (c) Examples of many-body wave functions the projected ensembles of which form approximate quantum state designs include the typical output states of random unitary circuit evolution, quantum states obtained from quenched time evolution, and energy eigenstates of a chaotic Hamiltonian at infinite temperature. (d) An illustration of minimal quantum state k-designs for a single qubit on the Bloch sphere. The formation of higher k-designs requires a larger number of pairwise nonorthogonal quantum states. In the limit of $k \to \infty$, a k-design approaches the so-called Haar ensemble, which is the uniform distribution over all pure quantum states.

in wide-ranging applications from quantum cryptography [14,15] to benchmarking [17–19,32–34]. Indeed, a new benchmarking protocol based on emergent-state 2-designs is demonstrated using a Rydberg quantum simulator in a parallel work [25].

II. QUANTUM STATE DESIGNS FROM PROJECTED ENSEMBLES

A. Ensemble of quantum states

The study of properties of random quantum states is useful because they often encode universal phenomena found in nature. Randomness is not a property of an individual wave function or density matrix but is a property of an *ensemble of quantum states*. An ensemble of states $\mathcal{E} = \{p_i, |\Psi_i\rangle\}$ is a set of quantum states $|\Psi_i\rangle$ weighted by probabilities p_i. This is simply a (discrete or continuous) probability distribution of quantum states, capturing a random process that stochastically gives pure states $|\Psi_i\rangle$.

State ensembles capture more information than the density matrix. An ensemble $\mathcal{E}$ uniquely specifies a density matrix $\rho = \sum_i p_i |\Psi_i\rangle \langle\Psi_i|$. Then, the density matrix captures the *average* of the expectation values of any observable O: $\sum_i p_i O_i = \mathrm{tr}(O\rho)$, where $O_i := \langle\Psi_i| O |\Psi_i\rangle$. However, a given density matrix ρ does not uniquely specify a state ensemble. For example, Fig. 1(d) shows several ensembles of single-qubit states that have the same density matrix ρ—the maximally mixed state $\mathbb{I}/2$—but are qualitatively different. For example, these ensembles can be distinguished through their *higher moments* of the observable expectation values O_i, e.g., the second moment $\sum_i p_i O_i^2$ is related to the variance of O_i and can be computed given $\mathcal{E}$ but not ρ. While such higher moments are difficult to measure, they can be statistically estimated with access to many (labeled) samples from the ensemble $\mathcal{E}$. The kth moment of any observable O can be calculated from the kth moment of $\mathcal{E}$:

$$\rho_{\mathcal{E}}^{(k)} = \mathbb{E}_{\Psi\sim\mathcal{E}}\left[(|\Psi\rangle\langle\Psi|)^{\otimes k}\right] = \sum_i p_i \left(|\Psi_i\rangle\langle\Psi_i|\right)^{\otimes k} . \quad (1)$$

The kth moment of O is $\sum_i p_i O_i^k = \mathrm{tr}(O^{\otimes k}\rho_{\mathcal{E}}^{(k)})$. $\rho_{\mathcal{E}}^{(1)}$ is the conventional density matrix and $\rho_{\mathcal{E}}^{(k)}$ can be viewed as a density matrix on k copies of the Hilbert space $\mathcal{H}^{\otimes k}$ describing an incoherent sum of k copies of identical states $|\Psi_i\rangle$. Just as higher moments such as variance and skewness are used to characterize a conventional probability distribution, knowledge of higher moments characterizes a state ensemble.

The higher moments of an ensemble are required to measure nonlinear quantities associated with the ensemble. For example, the Renyi entropy of a state $|\Psi\rangle$, across a bipartition A_1, A_2, is given by $S^{(n)}(|\Psi\rangle) \equiv 1/(1-n)\log\left(\mathrm{tr}_{A_1}\left[\left(\mathrm{tr}_{A_2}|\Psi\rangle\langle\Psi|\right)^n\right]\right)$. The average Renyi entropy of states in the ensemble $\mathcal{E}$ is given by $\overline{S^{(n)}} = \sum_i p_i S^{(n)}(|\Psi_i\rangle)$. This, and other nonlinear quantities, necessitates the knowledge of all kth moments $\rho_{\mathcal{E}}^{(k)}$. For a single state, the Renyi entropy is fully defined via its

reduced density matrix in a nonlinear fashion. This reduced density matrix is distinct from the density matrix of the *ensemble* $\rho_{\mathcal{E}}^{(1)}$. The average Renyi entropy $\overline{S^{(n)}}$ taken over an ensemble requires knowledge of all higher moments $\rho^{(k)}$ and hence is a nonlinear quantity of the ensemble.

Such nonlinear quantities have been the subject of recent theoretical interest. For example, recent works [21–24] have discovered a measurement-induced phase transition in quantities such as the entanglement entropy, state purity, Fisher information, or charge variance. These are all nonlinear quantities: the measurement-induced phase transition cannot be detected by any linear observable. In the cited works, the detection of these transitions requires keeping track of each measurement outcome and averaging the above nonlinear quantities over their corresponding pure states.

Lastly, we highlight a theoretically important ensemble of states: the *Haar ensemble*. This is the uniform distribution over all pure states in a Hilbert space, formally defined as the unique ensemble that is invariant under any unitary transformation U. It is the maximally entropic distribution of quantum states and we can quantify the randomness of an ensemble of states by its proximity to the Haar ensemble.

B. Quantum state designs

The randomness of an ensemble of states is quantified through its kth moments. Conventionally, the distance of an ensemble $\mathcal{E}$ from the Haar ensemble can be systematically measured by comparing their kth moments. The kth moment of the Haar ensemble $\rho_{\text{Haar}}^{(k)}$ is of interest because of its application to quantum information science [14–19]. An ensemble is called a *quantum state k-design* if its kth moment equals the kth moment of a Haar-random ensemble: $\rho_{\mathcal{E}}^{(k)} = \rho_{\text{Haar}}^{(k)}$.

While the Haar ensemble can be studied analytically using various statistical tools [35], it is extremely challenging to realize experimentally, as doing so requires an exponential amount of resources (such as the number of quantum gates or experimental operations) [20]. Instead, ensembles that form quantum state designs are considered because they mimic the Haar ensemble and can be efficiently realized in physical systems [14,36–40].

To build intuition, let us consider quantum state designs for a single qubit [see Fig. 1(d)]. For example, the ensemble of single-qubit states $\{|0\rangle, |1\rangle\}$ with equal probabilities has the same mean (first moment) as the Haar ensemble: $\rho^{(1)} = 1/2\,(|0\rangle\langle 0| + |1\rangle\langle 1|) = \mathbb{E}_{\Psi\sim\text{Haar}}[|\Psi\rangle\langle\Psi|]$, where $\mathbb{E}_{\Psi\sim\mathcal{E}}$ denotes averaging $|\Psi\rangle$ over the ensemble $\mathcal{E}$. We say that such a two-state ensemble forms a 1-design. However, the two-state ensemble does not form a 2-design because its second moment $1/2\,(|0\rangle\langle 0|\otimes|0\rangle\langle 0| + |1\rangle\langle 1|\otimes|1\rangle\langle 1|)$ differs from that of the Haar ensemble. To form a 2-design for a single qubit, one can use four distinct nonorthogonal quantum states uniformly spread over the Bloch sphere [Fig. 1(d)]. In general, for an ensemble to form a higher-order design, it must be supported over a larger number of states [41].

Formally, we say that an ensemble $\mathcal{E}$ is an ε-approximate quantum state k-design if

$$\left\| \rho_{\mathcal{E}}^{(k)} - \rho_{\text{Haar}}^{(k)} \right\|_1 \leq \varepsilon, \tag{2}$$

where $\|\cdot\|_1$ denotes the trace norm. Further details of the Haar kth moments are provided in Appendix A. This definition means that the kth moment of $\mathcal{E}$ is nearly indistinguishable from the kth moment of the Haar ensemble up to a small error ε. It can be shown that an ε-approximate k-design is also an ε-approximate j-design for any $j \leq k$ and, accordingly, larger values of k indicate that an ensemble looks more uniformly random.

Unlike the Haar ensemble, approximate designs arise in physical settings [14,36–40]. A canonical example is that of random unitary circuits [14,36,38], where a set of random two-qubit unitary gates in a certain geometric arrangement are sequentially applied to simple initial states [for an example, see Fig. 1(c)]. Then the ensemble of resulting states over different choices of unitary gates forms an ε-approximate k-design as long as the depth of the circuit is sufficiently large, scaling polynomially in k, $\log(1/\varepsilon)$, and the number of qubits N [14]. Similarly, there are a number of proposals to generate approximate designs based on time-dependent local Hamiltonian evolution [37,40].

C. Projected ensembles

In the above examples, states are sampled from approximate k-designs by realizing many distinct highly engineered quantum evolutions. By contrast, we show that approximate k-designs arise naturally from the projected ensemble of a *single* many-body wave function. More precisely, consider a many-body wave function $|\Psi\rangle$ in the Hilbert space $\mathcal{H}$ for a bipartite system consisting of N_A qubits in A and N_B qubits in B. The projected ensemble for A is generated by performing projective measurements on all N_B qubits in B in a local basis $\{|z_B\rangle\}$, where a bit string $z_B \in \{0,1\}^{N_B}$ enumerates over all 2^{N_B} measurement outcomes. After the measurement, with probability $p(z_B)$, the system is described by the normalized wave function $|\Psi_A(z_B)\rangle \otimes |z_B\rangle$, where

$$p(z_B) := \langle\Psi|\,(\mathbb{1}_A \otimes |z_B\rangle\langle z_B|)\,|\Psi\rangle \tag{3}$$

$$|\Psi_A(z_B)\rangle := \big(\mathbb{1}_A \otimes \langle z_B|\big)|\Psi\rangle/\sqrt{p(z_B)} \tag{4}$$

and $\mathbb{1}_A$ is the identity operator acting on qubits in A. This defines a *projected ensemble*:

$$\mathcal{E}_{\Psi,A} := \{p(z_B)\,,\, |\Psi_A(z_B)\rangle\}. \tag{5}$$

The ensemble consists of 2^{N_B} states that are generally not pairwise orthogonal. We call $|\Psi\rangle$ the *generator state* of $\mathcal{E}_{\Psi,A}$. Conversely, we can reconstruct $|\Psi\rangle$ from $\mathcal{E}_{\Psi,A}$, assuming that we also know the global phase of each projected state $|\Psi_A(z_B)\rangle$: $|\Psi\rangle = \sum_{zB} p(z_B)\,|\Psi_A(z_B)\rangle \otimes |z_B\rangle$.

Projected ensembles have been considered in previous works. Specifically, Refs. [42,43] use a related construction to define localizable entanglement, Refs. [44,45] discuss the projected ensemble generated by typical microcanonical states, and Refs. [46,47] discuss using a projected ensemble to generate k-designs from a fine-tuned state. In this work, we find that projected ensembles form approximate k-designs for many physically relevant many-body wave functions, including special analytically soluble cases.

D. Solvable example I: Graph states from time-independent Hamiltonian evolution

A generator state for an ε-approximate k-design can be efficiently prepared by a simple protocol, in which an initial product state is evolved by a Hamiltonian over a short time duration. The Hamiltonian can be time independent and only has nearest-neighbor Ising interactions (Fig. 2). This is possible because a universal resource state for measurement-based quantum computation can be prepared within a constant time independent of system size [46–48].

This has been explicitly shown in Ref. [47]. More specifically, let us consider $N = N_r \times N_c$ qubits arranged in a square lattice with N_r rows and N_c columns (see Fig. 2). We assume a periodic boundary condition along each column, for the concreteness and simplicity of our analysis. For later convenience, we associate every qubit in the system with one of the four distinct groups $\{a, b, c, d\}$, as indicated in Fig. 2(b). All qubits are initially prepared in $|\Psi_0\rangle = |0\rangle^{\otimes N}$ states and time evolved under the Hamiltonian

$$H_{\text{Ising}} = \sum_i h_i \sigma_i^x + \sum_{(i,j)\in E} J_0 \sigma_i^x \sigma_j^x, \tag{6}$$

for duration $\tau = \pi/(4J_0)$. Here, the on-site field strength h_i takes one of the four values $\{h_a = 0, h_b = J_0/2, h_c = -J_0/2, h_d = J_0\}$, depending on which group the qubit at site i belongs to. The Ising interactions with strength J_0 are present only for nearest-neighboring qubit pairs connected by edges of the graph shown in Fig. 2(a).

The results from Ref. [47], when viewed in the context of our currently proposed framework, show that if all $(N_c - 1)N_r$ qubits in the first $N_c - 1$ columns are measured in the computational basis, the projected ensemble for the remaining N_r qubits forms a ε-approximate k-design as long as $N_c \geq C(N_r k + \log_2(1/\varepsilon))$ with some constant C

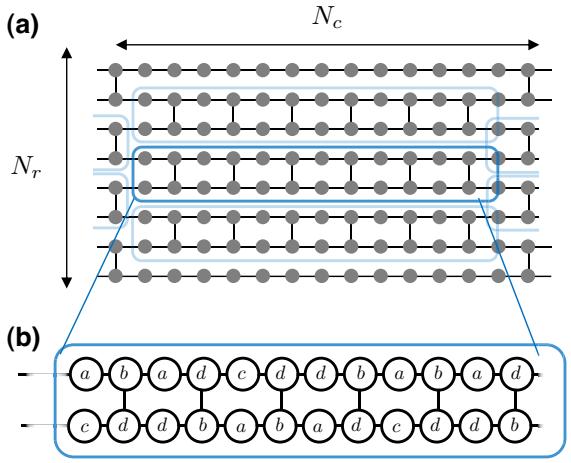

FIG. 2. A geometric arrangement of qubits for the solvable example of a generator state for an ε-approximate k-design. The generator state can be efficiently prepared by time-independent Hamiltonian evolution. (a) A graph representing the connectivity of $N = N_r \times N_c$ qubits arranged in a two-dimensional square lattice. Qubits are represented by vertices (gray circles) and a pair of neighboring qubits interact via an Ising coupling whenever an edge (black line) exists between them. The connectivity exhibits a periodicity, where the two-row 12-column unit cell (blue box) is repeated in a brick-wall layout. (b) The connectivity and grouping of qubits within a single unit cell. Each qubit belongs to one of the four distinct groups as indicated by labels $\{a, b, c, d\}$.

[48]. Note that the state-preparation time τ in this protocol is independent of the subsystem size N_r, which can be arbitrarily large as long as N_c is sufficiently larger.

The formation of the approximate k-design can be rigorously shown by using results in measurement-based quantum computation as follows [48]. Since all terms in H_{Ising} commute with one another, we can decompose the time evolution under H_{Ising} into products of two unitary evolutions, each generated by single-qubit terms and two-qubit Ising terms. When the initial state is evolved under Ising interactions without on-site field terms for duration τ, the resultant state is equivalent to the graph state [49] associated with the graph in Fig. 2(a) up to local unitary rotations along the x axis. The graph state forms a universal resource for measurement-based quantum computation: by sequentially measuring qubits column by column, one can simulate the quantum circuit dynamics of depth $N_c - 1$ consisting of N_r qubits. The output of the quantum circuit is encoded in the remaining unmeasured N_r qubits. In this simulation, the choices of quantum gates are determined by the combination of the measurement basis and the measurement outcomes. Since measurement outcomes are probabilistic, quantum gates are also random, forming an ensemble of random quantum circuits. In our case, the measurement basis is along the z direction (computational basis), which is effectively rotated along the x axis due to

the unitary evolution generated by single-qubit terms. The values of h_i are chosen such that the set of measurements represents an ensemble of quantum circuits that forms an approximate unitary k-design for large N_c [47].

E. Solvable example II: Generic quantum states

While the example in the preceding section is fine tuned, we show that, in fact, *generic* many-body wave functions are good generator states for ε-approximate k-designs.

Theorem 1. *Let* $|\Psi\rangle$ *be chosen uniformly at random from the Hilbert space* $\mathcal{H}$*. The ensemble* $\mathcal{E}_{A,\Psi}$ *forms an* ε*-approximate* k*-design with probability at least* $1-\delta$ *if*

$$N_B = \Omega\left(k N_A + \log\left(\tfrac{1}{\varepsilon}\right) + \log\log\left(\tfrac{1}{\delta}\right)\right). \tag{7}$$

Here, $\Omega(\cdot)$ denotes a lower bound up to a constant multiplicative factor and subleading terms. A proof sketch for this theorem is given in Appendix E, along with a full proof in Appendix F. This theorem establishes that all but a tiny fraction (of order approximately $1/2^{2^{N_B}}$) of the states in the Hilbert space are generator states for approximate k-designs if N_B is asymptotically larger than k times N_A. However, a quantum state randomly sampled from the entire Hilbert space is not so physical, since such a state is extremely difficult to produce experimentally [20]. To this end, we also present the following theorem.

Theorem 2. *Let* $|\Psi\rangle$ *be a state sampled from an ensemble on* $\mathcal{H}$ *that forms an* ε'*-approximate* k'*-design. Then the projected ensemble* $\mathcal{E}_{A,\Psi}$ *forms an* ε*-approximate* k*-design with probability at least* $1-\delta$ *if*

$$N_B = \Omega\left(kN_A + \log\left(\tfrac{1}{\varepsilon\delta}\right)\right), \tag{8}$$

$$k' = \Omega\left(k\left(N_B + \log\left(\tfrac{1}{\varepsilon\delta}\right)\right)\right), \tag{9}$$

$$\log\left(\tfrac{1}{\varepsilon'}\right) = \Omega\left(kN_B\left(N_B + \log\left(\tfrac{1}{\varepsilon\delta}\right)\right)\right), \tag{10}$$

$$N_A = \Omega\left(\log(N_B) + \log(k) + \log\log\left(\tfrac{1}{\varepsilon\delta}\right)\right). \tag{11}$$

The proof is given in Appendices E and F and relies on higher-order concentration of measure results [50] as well as a polynomial approximation technique used in quantum algorithms for solving linear systems [51]. This theorem shows that if the generator state is complex enough, in the sense that it is a typical state from an ε'-approximate k'-design for small ε' and large k' [52], then the projected ensemble will well approximate a quantum k-design.

Theorem 2 has implications for ongoing experiments: if a *single sample* of an approximate design is experimentally realized, our theorem states that it can be used to generate *ensembles* forming approximate designs on its subsystems. This can be achieved, for example, with a single instance of a random unitary circuit [Fig. 1(c)]. A single instance is easier to generate than an *ensemble* of random circuits, which requires fine-tuned controls to ensure that the ensemble satisfies desired statistical properties. Therefore, this approach to generating random states is extremely hardware efficient and could lead to various useful applications in quantum information science [14,15,17,18,32,33].

At a conceptual level, Theorem 2 establishes that a large class of states that can be efficiently prepared are good generators of k-designs. This raises the possibility that, even in natural chaotic quantum systems, approximate k-designs may arise from projected ensembles. We investigate this possibility in Sec. III.

III. QUANTUM STATE DESIGNS FROM CHAOTIC SYSTEMS

Motivated by Theorems 1 and 2, we numerically investigate projected ensembles that arise from chaotic Hamiltonian dynamics. In particular, we study the projected ensembles of time-evolved states and find that infinite-temperature dynamics does indeed generate projected ensembles that look Haar random in a quantifiable way, i.e., they form approximate k-designs. Crucially, we go beyond the intuition from our theorems and find that projected ensembles form in a constant time, before the global state has equilibrated, and hence very distinct from typical random states. We further investigate projected ensembles at finite temperature and find the emergence of a universal ensemble.

A. Model

We start with the paradigmatic example of the one-dimensional quantum Ising spin system with mixed fields (QIMF), described by the Hamiltonian

$$H_{\mathrm{QIMF}} = h^x \sum_{j=1}^{N} \sigma_j^x + h^y \sum_{j=1}^{N} \sigma_j^y + J \sum_{j=1}^{N-1} \sigma_j^x \sigma_{j+1}^x, \tag{12}$$

where N is the number of spins, σ_j^μ with $\mu = x, y, z$ are the Pauli operators for a spin at site j, J is the strength of the Ising interactions, and h^x and h^y are the strengths of the longitudinal and transverse fields, respectively. In the absence of the longitudinal field ($h^x = 0$), the Hamiltonian can be mapped to an integrable model of noninteracting fermions via the Jordan-Wigner transformation, leading to nonergodic dynamics. However, for any nonzero longitudinal field ($h^x \neq 0$), the Hamiltonian is ergodic and its eigenvalues and eigenvectors are expected to have properties consistent with ETH predictions. This has been explicitly checked for a specific parameter set $(h^x, h^y, J) = (0.8090, 0.9045, 1)$ [53], which we adopt for our study. We note that we do not find a qualitative differences in our results when using nearby parameter values.

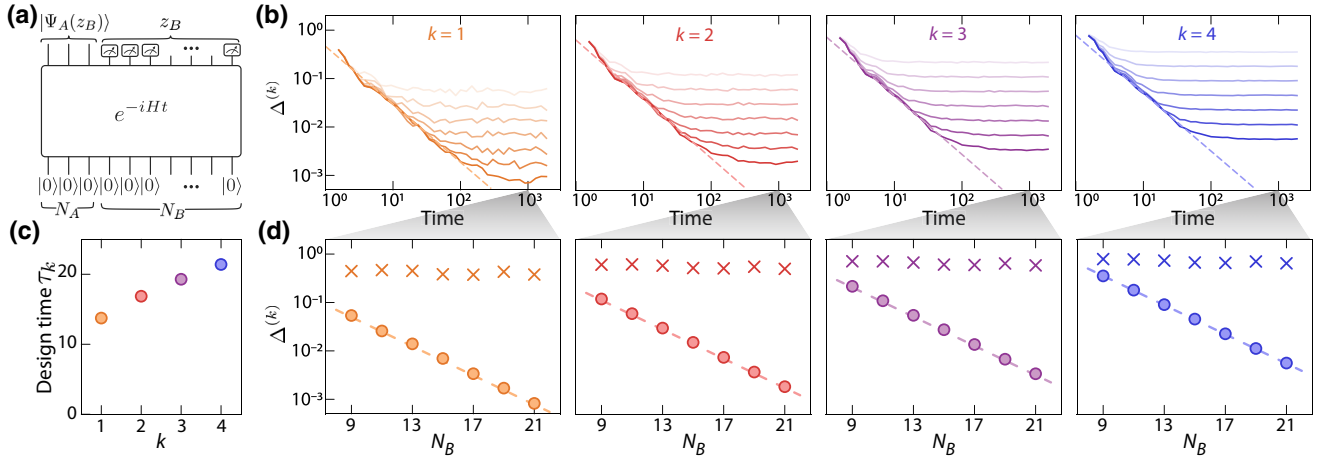

FIG. 3. Emergent quantum state designs from chaotic time evolution. (a) Quenched dynamics under a time-independent Hamiltonian H starting from an initial product state, $|0\rangle^{\otimes N}$, for a N-qubit system. The system is partitioned into two subsystems, A and B, with sizes N_A and N_B, respectively. At time t, a projective measurement in the local z basis is performed on subsystem B, resulting in a specific outcome z_B of length N_B. (b) The trace distances $\Delta^{(k)}$ between the kth moments of the Haar ensemble and projected ensembles for an $N_A = 3$ subsystem as a function of the evolution time for various total system sizes $N = 12, 14, \ldots 24$ (darker colors for increasing N). The dashed lines are a phenomenological power-law fit, yielding a scaling of $\Delta^{(k)} \sim t^{-1.2}$. (c) The design time τ_k defined by the evolution time to achieve a trace distance of $\Delta^{(k)} = 0.02$. We find that longer time evolution is required to form a higher approximate k-design. (d) Late-time trace distances at $t \approx 10^3$. For a mixed-field chaotic Hamiltonian (circles), the late-time trace distances exhibit an exponential scaling with N_B while they remain nearly constant for an integrable nonergodic Hamiltonian (crosses; time traces not shown). The dashed lines represent the trace distances $\Delta^{(k)}_{\text{em}}$ between the kth moments of the Haar ensemble and a finite set of 2^{N_B} states sampled from the Haar ensemble on A.

B. Infinite-temperature projected ensembles

We first consider the many-body state $|\Psi(t)\rangle = e^{-iH_{\text{QIMF}}t}|\Psi_0\rangle$ resulting from time evolution of the initial state $|\Psi_0\rangle = |0\rangle^{\otimes N}$ [Fig. 3(a)]. Here, $|0\rangle_j$ and $|1\rangle_j$ are the eigenstates of σ_j^z with eigenvalues $+1$ and -1, respectively. The initial product state, $|\Psi_0\rangle$, has a zero energy expectation value with respect to H_{QIMF}, corresponding to the total energy of an infinite-temperature state. This implies that local subsystems will relax to an infinite-temperature ensemble after a local thermalization time [8–10].

At any time t, the projected ensemble for a subsystem A is obtained by simulating projective measurements on the rest of the $N_B = N - N_A$ qubits in the local z basis [Fig. 3(a)] [54]. In order to check if the projected ensemble forms an approximate k-design, we compare the kth moment of the ensemble, $\rho_{\mathcal{E}}^{(k)}$ in Eq. (1), to that of the uniform ensemble $\rho_{\text{Haar}}^{(k)}$ using the trace distance $\Delta^{(k)} = 1/2 \left\| \rho_{\mathcal{E}}^{(k)} - \rho_{\text{Haar}}^{(k)} \right\|_1$. For a fixed subsystem size of $N_A = 3$, we numerically compute the trace distance $\Delta^{(k)}$ up to $k = 4$ as a function of time for various N_B [Fig. 3(b)]. In all cases, $\Delta^{(k)}$ decays in time, following a phenomenological power-law scaling $\Delta^{(k)} \sim 1/t^{1.2}$, until it saturates to a value that is governed by finite-size effects. This slow power-law decay could be due to emergent hydrodynamics associated with energy conservation [55] and is absent when the Hamiltonian is geometrically nonlocal or time dependent (see Appendix C and Fig. 6). The saturation value of $\Delta^{(k)}$ decreases exponentially with N_B [Fig. 3(d), circles], exhibiting the scaling $\Delta^{(k)} \sim 1/\sqrt{2^{N_B}}$. If the power-law scaling persists at larger system sizes, then $\Delta^{(k)}$ may decrease over an exponentially long time scale.

As a comparison, we also present $\Delta^{(k)}_{\text{em}}$, which is the trace distance between the kth moments of the Haar ensemble and the empirical Haar ensemble consisting of 2^{N_B} states sampled uniformly at random on A [Fig. 3(d), dashed lines]. Since the empirical ensemble asymptotically approaches the Haar ensemble in the limit of infinite samples, $\Delta^{(k)}_{\text{em}}$ is determined only by statistical fluctuations associated with having a finite number of quantum states (see Appendix B). Remarkably, we find that the projected ensemble obtained from the quench dynamics shows a trace distance almost identical to that of the empirical ensemble of the same size, suggesting that the former is as uniformly random as the latter. By contrast, repeating similar calculations for the integrable model ($h^x = 0$), we observe qualitatively different behavior where the trace distance to the Haar ensemble is much larger than in the nonintegrable case [Fig. 3(d), crosses]. Furthermore, there is no appreciable dependence on system size. This is expected, since integrable systems do not locally thermalize and instead relax to a generalized Gibbs ensemble

and hence will not form 1-designs at effective infinite temperature [56].

Given the emergence of k-designs in asymptotic regimes for a chaotic Hamiltonian, it is natural to ask how long it takes for a subsystem to achieve an approximate design up to a small fixed precision. To this end, we introduce a design time τ_k, defined as the time at which $\Delta^{(k)}$ becomes smaller than a certain fixed threshold ε. For a chosen threshold $\varepsilon = 0.02$, we find that the formation of higher k-designs requires longer time evolution [Fig. 3(c)]. This observation is consistent with the idea that typical quantum states from higher k-designs are more complex and hence more difficult to prepare [52,57,58].

We note that our numerical results cannot be explained by, and go beyond, our theorem in Sec. II. Naively, one could argue that the global time-evolved wave function behaves as if it is a typical instance of a random state. If so, based on our Theorem 2, its projected ensemble should form an approximate k-design. Indeed, Refs. [37,58,59] investigate whether a state design can be generated by evolving a simple state under an ensemble of different Hamiltonians at late times (scaling with system size). In contrast, our numerical simulations suggest that the projected ensemble may form very quickly, well before its global wave function behaves like a random state. Figure 3(b) suggests that for a given ε, we form an ε-approximate projected k-design in constant time, independent of the total system size N. Therefore in the thermodynamic limit, an ε-approximate projected ensemble may form even though the global state is far from a typical state. The latter is due to locality: e.g., its half-chain entanglement entropy requires at least $O(N)$ time to reach its equilibrium "typical value." In fact, the global state can never be a typical design state owing to energy conservation.

Our experimental companion paper [25] observes signatures consistent with the formation of a projected k-design in quenched states in a Rydberg quantum simulator. Namely, it has been shown that the statistical properties of measurements in the computational basis are consistent with those from a projected k-design.

Next, we investigate the properties of energy eigenstates. We repeat an analysis that is similar to the one above by replacing the time-evolved state $|\Psi(t)\rangle$ with an energy eigenstate $|E_i\rangle$ of H_{QIMF}. Figure 4 shows the trace distance $\Delta^{(k)}$ as a function of energy E_i for a projected ensemble generated from $|E_i\rangle$. We find a sharp dip at zero energy density corresponding to infinite temperature, which signifies the emergence of approximate quantum state designs [Fig. 4(a)]. At zero energy, $\Delta^{(k)}$ decreases exponentially as a function of the system size for $k = 1, 2, 3$. However, the decay rate is slightly slower than in the case of quenched dynamics [Fig. 4(b)].

In addition to the QIMF, we also study time-dependent dynamics and two other ergodic Hamiltonian models in order to corroborate the universality of our findings (see

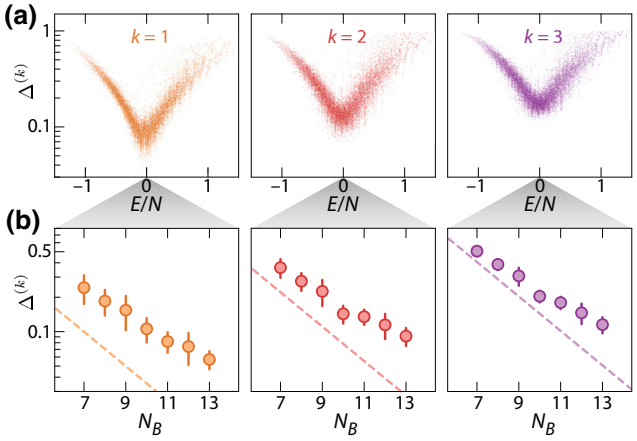

FIG. 4. Emergent quantum state designs from energy eigenstates. (a) The trace distances between the kth moments of the Haar ensemble and a projected ensemble for an $N_A = 3$ subsystem generated from the energy eigenstates of a Hamiltonian. The results are presented as a function of the energy density E/N for a total system size of $N = 14$. (b) The trace distances for the projected ensembles obtained from eigenstates near zero energy corresponding to infinite temperature. The distances exhibit an exponential decay as a function of the system size. The points are evaluated for 100 eigenstates near zero energy and the error bars denote their standard deviation. For comparison, the dashed lines represent the trace distances $\Delta^{(k)}_{\mathrm{em}}$ between the Haar ensemble and an empirical ensemble of 2^{N_B} states sampled from the Haar ensemble.

Appendix C and Fig. 8). Specifically, we consider a system of random all-to-all coupled spin-1/2 particles as well as hard-core bosons with random all-to-all hoppings with particle-number conservation. In both cases, we observe excellent convergence of projected ensembles to approximate k-designs. In the latter case, the measurement outcomes in B and the corresponding pure quantum states on A are strongly correlated owing to the particle-number-conservation symmetry; hence a naive approach based on Eq. (5) does not lead to approximate k-designs. We instead introduce symmetry-resolved projected ensembles by grouping certain subsets of measurement outcomes from B (for details, see Appendix C 3); this does lead to approximate k-designs.

C. Finite-temperature projected ensembles

For chaotic Hamiltonians, projected ensembles forming $k = 1$ designs can be anticipated from the standard picture of quantum thermalization since the first moment of a projected ensemble simply corresponds to the reduced density matrix of a subsystem. The reduced density matrix approaching the maximally mixed state, i.e., the first moment of the Haar ensemble, follows from local thermalization at infinite temperature. However, the convergence of higher moments $k \geq 2$ of the projected ensemble to

higher k-designs is nontrivial and surprising. Such convergence cannot be explained by ETH alone and suggests a new form of emergent randomness beyond the conventional framework of quantum thermalization.

A natural next step is to generalize our results to a finite-temperature setting. However, we are unaware of any straightforward generalizations of the Haar ensemble and its corresponding quantum state designs at finite temperature. Such an ensemble, if it exists, would generally depend on the system Hamiltonian and its first moment should (approximately) be a thermal state. While explicitly identifying properties of such an ensemble is an interesting future direction, here we find numerical evidence that such an ensemble exists. In Fig. 5, we compute projected ensembles for all energy eigenstates of H_{QIMF} and present the pairwise distances

$$\Delta_{ij}^{(2)} = \frac{1}{2}\left\|\rho_i^{(2)} - \rho_j^{(2)}\right\|_1, \tag{13}$$

where $\rho_i^{(2)}$ denotes the second moment of a projected ensemble generated from an eigenstate $|E_i\rangle$. We find that $\Delta_{ij}^{(2)}$ is a smooth function of energy up to small fluctuations, suggesting that the projected ensembles vary smoothly with the energy as well [Fig. 5(a)]. Also, $\Delta_{ij}^{(2)}$ is minimized when the energy difference $|E_i - E_j|$ is small [Fig. 5(b)] and in this regime $\Delta_{ij}^{(2)}$ decreases exponentially with the system size (Fig. 5(b) inset). These observations suggest that the second moment of the projected ensemble is indeed universal even at finite temperatures.

One possible choice of the universal ensemble is the *Scrooge ensemble* proposed in Ref. [60]. The Scrooge, or "Gaussian Adjusted Projected" (GAP), ensemble has been conjectured to describe the thermal equilibrium distribution and shown to arise from a projected ensembles in an certain idealized situation [44,45]: a typical microcanonical state is projectively measured in a random nonlocal basis under appropriate conditions. For a given density matrix ρ, the Scrooge ensemble is the unique distribution of pure states that has ρ as its first moment and has minimal accessible information [60]. Figure 5(c) tests this conjecture for the projected ensembles generated from natural quantum states and local measurement bases. We plot the trace distance between the second moments of the projected ensemble and corresponding Scrooge ensemble for eigenstates $|E\rangle$. Across the spectrum, the trace distance is much smaller than that against the 2-design and does not depend sensitively on the energy over a wide range $E/N \in [-0.5, 0.5]$, with deviations only near the edges of the spectrum. While this is consistent with the conjecture, it remains as an interesting future direction to systematically study when the Scrooge ensemble arises from natural many-body dynamics and, if so, to what extent.

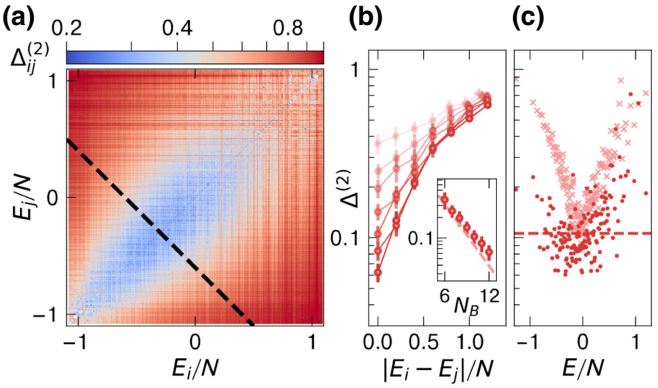

FIG. 5. A universal ensemble at finite temperature. (a) The trace distances between the second moments of the projected ensembles generated from a pair of energy eigenstates at E_i and E_j for $N_A = 3$ subsystems. We plot the pairwise distances $\Delta_{ij}^{(2)}$ for every pair of eigenstates $|E_i\rangle, |E_j\rangle$, computed for system size $N = 11$. The distances are minimized when $E_i \approx E_j$. The plot suggests the existence of a universal ensemble that depends smoothly on the energy density. The black dashed line indicates a cut defined by $(E_i + E_j)/N = -0.6$. (b) The trace distances plotted as a function of the energy-density difference, $|E_i - E_j|/N$, along the black dashed line in (a) for various systems sizes $N = 9, 10, \ldots, 15$ (darker colors for increasing N). The inset shows that the distances at zero energy difference $E_i = E_j$ decay exponentially with the system size. The trace distances from the projected ensemble are comparable to those from the finite-size ensemble, $\Delta_{\mathrm{em}}^{(2)}$, of the empirical Haar distribution (dashed line). Such an exponential scaling suggests the existence of a universal random ensemble at *finite* temperatures, which we hypothesize to be the *Scrooge ensemble*. (c) The trace distance $\Delta_E^{(2)}$ (dark red, dots) between the second moments of the projected ensemble and the corresponding Scrooge ensemble of eigenstates $|E\rangle$, for a $N_A = 3$ subsystem of a $N = 14$ system. $\Delta_E^{(2)}$ remains small for eigenstates $|E\rangle$ in a wider energy range $E/N \in [-0.5, 0.5]$, with an average value of $\overline{\Delta^{(2)}} = 0.11$ (dashed line). In contrast, the trace distance between the projected and Haar ensembles (light red, crosses) is strongly energy dependent and is large at energies away from infinite effective temperature $E = 0$.

IV. DISCUSSION

We introduce a qualitatively different study of equilibrium behavior on a subsystem. Conventional studies of ensembles such as microcanonical or canonical ensembles are fully specified by their corresponding reduced density matrices. In contrast, our formalism concerns more general statistical properties (such as higher moments) of an ensemble with a large number of pure states that are generally pairwise nonorthogonal. As such, a projected ensemble encodes additional information about a subsystem. In particular, a projected ensemble is useful when describing the information-theoretic properties of a subsystem. For example, one can ask the following question: How much information (i.e., classical bits) is required in order to specify the full wave function of a subsystem A

after its complement B is measured in the standard basis? This question is natural for a classical observer who only has direct access to the subsystem B. If the projected ensemble induced by the measurement on B was to be the Haar ensemble, an exponential number of bits would be required to specify the wave function on A. By contrast, if the projected ensemble is uniformly distributed only over computational basis states of A, a linear number of bits suffice to specify the wave function on A. Note that the former and latter projected ensembles produce the same density matrix, namely the maximally mixed state. Therefore, projected ensembles provide a novel framework to analyze the information content associated with quantum states of a subsystem and their relation with the remainder of the system.

V. OUTLOOK

Our findings open up a number of new directions in quantum chaos, thermalization, and quantum information. In particular, our numerical results demonstrate that for an initial product state evolved by a chaotic Hamiltonian, the largest k-design attained by the projected ensemble grows as a function of time; moreover, this growth persists significantly past the local thermalization time scale. While it is not presently clear how long this growth will persist, the observed growth is clearly diagnostic of the sustained development of nonlocal correlations after local thermalization has occurred (see, e.g., Refs. [31,61]). There may even be connections to the quantum complexity of a state evolving by chaotic dynamics, which likewise grows long after thermalization has occurred [52,62]. It would also be interesting to fully generalize the above to projected ensembles at finite temperature, in the presence of symmetries, quasi-integrability, or strong disorder resulting in localization [11–13].

In quantum information science, quantum state designs are valuable resources in many applications. Our work establishes projected ensembles as a practical way of sampling states from approximate designs using natural Hamiltonian evolutions of existing quantum simulators without fine control. Further, our work could lead to novel experimental quantum tomography protocols [17,34], cryptographic protocols for hiding information [14], the design of unforgeable quantum encryption [15], and also new methods for quantum device verification [18,19]. Indeed, a parallel work [25] uses projected ensembles to devise and implement a novel benchmarking protocol.

Finally, an important question at the intersection of computer science and quantum many-body physics is whether the computational complexity of simulating natural chaotic dynamics is beyond the capability of contemporary classical computers. In other words, can quantum advantage tests be performed using a fixed chaotic Hamiltonian with analog quantum simulators? Existing sampling-based quantum advantage protocols rely heavily on certain statistical and computational properties of state ensembles formed by applying random unitaries to a fixed state [63–65]. The projected ensemble emerging from generic quantum dynamics may also possess the requisite properties and our work could lead to a new approach to quantum advantage using analog quantum simulators. *Note added.*—We recently became aware of Refs. [44,45], where the projected ensemble has been studied in the context of typical microcanonical states.

ACKNOWLEDGMENTS

We thank Ehud Altman, Adam Bouland, Fernando Brandão, Aram Harrow, Wen Wei Ho, Nicholas Hunter-Jones, Vedika Khemani, Anand Natarajan, and Hannes Pichler for valuable discussions. This work was partly supported by the Institute for Quantum Information and Matter, a National Science Foundation (NSF) Physics Frontiers Center (NSF Grant No. PHY-1733907), the NSF CAREER award (1753386), the Air Force Office of Scientific Research (AFOSR) Young Investigator Program (YIP) (Grant No. FA9550-19-1-0044), the Defense Advanced Research Projects Agency (DARPA) Optimization with Noisy Intermediate-Scale Quantum devices (ONISQ) program (Grant No. W911NF2010021), the Army Research Office (ARO) Multidisciplinary University Research Initiative (MURI) program (Grant no. W911NF2010136), and the NSF Quantum Leap Challenge Institutes (QLCI) program (Grant No. 2016245). J.S.C. is supported by a Junior Fellowship from the Harvard Society of Fellows, as well as in part by the Department of Energy under Grant No. DE-SC0007870. H.H. is supported by the J. Yang & Family Foundation. F.H. is supported by the Fannie & John Hertz Foundation. J.C. acknowledges support from the IQIM postdoctoral fellowship. A.L.S. acknowledges support from the Eddleman Quantum graduate fellowship. S.C. acknowledges support from the Miller Institute for Basic Research in Science.

APPENDIX A: kTH MOMENT OF THE HAAR ENSEMBLE

To quantify the degree of randomness of our projected ensembles, we compute the trace distance $\Delta^{(k)}$ between the kth moments of the projected ensemble and the Haar ensemble:

$$\Delta^{(k)} = \frac{1}{2}\left\|\rho_{\mathcal{E}}^{(k)} - \rho_{\text{Haar}}^{(k)}\right\|_1 . \tag{A1}$$

Here, $\rho_{\text{Haar}}^{(k)}$ is the kth moment averaged over the Haar ensemble. For a Hilbert space $\mathcal{H}$ with dimension d, it has the form [35]

$$\rho_{\text{Haar}}^{(k)} = \mathbb{E}_{\Psi\sim\text{Haar}(d)}\left[(|\Psi\rangle\langle\Psi|)^{\otimes k}\right] \tag{A2}$$

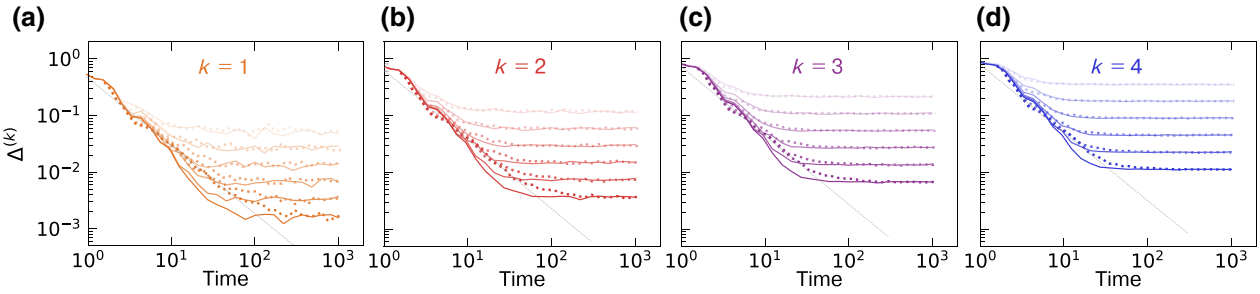

FIG. 6. Emergent quantum state designs in Floquet dynamics. The trace distance $\Delta^{(k)}$ versus the time for time evolution under the Floquet dynamics in Eq. (C1), for $k = 1, 2, 3, 4$ in (a), (b), (c), and (d), respectively. We plot the trace distance $\Delta^{(k)}$ between the kth moments of the Haar ensemble and the projected ensembles for an $N_A = 3$ subsystem and the total system sizes $N = 12, 14, \ldots, 22$ (darker colors denote increasing N). The solid lines denote the Floquet dynamics, while the dotted lines denote the time-independent dynamics. The dashed line denotes the power-law fit for time-independent dynamics. The Floquet dynamics quickly saturate to their late-time $\Delta^{(k)}$ value and deviate from power-law behavior. This suggests that the power-law decay of $\Delta^{(k)}$ is associated with hydrodynamic behavior due to energy conservation.

$$= \frac{\sum_{\pi\in S_k} \mathrm{Perm}_{\mathcal{H}^{\otimes k}}(\pi)}{d(d+1)\cdots(d+k-1)}. \tag{A3}$$

Here, S_k is the symmetric group on k elements and $\pi \in S_k$ is an element of the group. $\mathrm{Perm}_{\mathcal{H}^{\otimes k}}$ is a representation of S_k on $\mathcal{H}^{\otimes k}$, which permutes the tensor factors according to

$$\mathrm{Perm}_{\mathcal{H}^{\otimes k}}(\pi)\,|\Psi_1\rangle \otimes \cdots \otimes |\Psi_k\rangle = |\Psi_{\pi^{-1}(1)}\rangle \otimes \cdots \otimes |\Psi_{\pi^{-1}(k)}\rangle. \tag{A4}$$

The inverses in the subscripts are chosen so that $\mathrm{Perm}_{\mathcal{H}^{\otimes k}}(\pi)\cdot\mathrm{Perm}_{\mathcal{H}^{\otimes k}}(\pi') = \mathrm{Perm}_{\mathcal{H}^{\otimes k}}(\pi\circ\pi')$ (i.e., the representation is a homomorphism of S_k and not an anti-homomorphism). It is readily checked that Eq. (A2) can be written as

$$\frac{\sum_{\pi\in S_k} \mathrm{Perm}_{\mathcal{H}^{\otimes k}}(\pi)}{d(d+1)\cdots(d+k-1)} = \frac{\Pi_k}{\binom{d+k-1}{k}}, \tag{A5}$$

where Π_k is simply the projector onto the symmetric subspace of $\mathcal{H}^{\otimes k}$ (i.e., the invariant subspace under S_k); this subspace has dimension $\binom{d+k-1}{k}$.

APPENDIX B: FINITE SAMPLING ERROR FROM THE EMPIRICAL HAAR ENSEMBLE

In the main text, we construct projected ensembles of size 2^{N_B}, where N_B is the size of the complementary subsystem. In Figs. 3(d), 4(b), and 5(b), we compare the system-size scaling of $\Delta^{(k)}$ against the trace distance $\Delta^{(k)}_{\mathrm{em}}$ of the *empirical Haar ensemble*, an ensemble formed by sampling from the Haar ensemble 2^{N_B} times. In particular, we have

$$\rho^{(k)}_{\mathrm{em}} = \frac{1}{2^{N_B}}\sum_{i=1}^{2^{N_B}} (|\Psi_i\rangle\langle\Psi_i|)^{\otimes k}, \quad |\Psi_i\rangle \sim \mathrm{Haar}(d), \tag{B1}$$

$$\Delta^{(k)}_{\mathrm{em}} = \mathbb{E}_{\Psi_1,\Psi_2,\ldots\sim\mathrm{Haar}(d)}\left[\frac{1}{2}\|\rho^{(k)}_{\mathrm{em}} - \rho^{(k)}_{\mathrm{Haar}}\|_1\right]. \tag{B2}$$

This comparison is made to estimate the degree to which the error $\Delta^{(k)}$ in our projected ensemble is due to its finite size. We estimate $\Delta^{(k)}_{\mathrm{em}}$ for various values of N_B and find that $\Delta^{(k)}_{\mathrm{em}}$ scales exponentially as $1/\sqrt{2^{N_B}}$.

APPENDIX C: ADDITIONAL NUMERICAL SIMULATIONS OF ERGODIC DYNAMICS

1. Floquet evolution

Under the time-independent H_{QIMF}, the observed power-law decay of $\Delta^{(k)}$ over time evolution is likely due to emergent hydrodynamics associated with energy conservation [55]. To illustrate this, we study a time-dependent Floquet variant of H_{QIMF}:

$$H^{\mathrm{Floq.}}_{\mathrm{QIMF}}(t) = h^x(t)\sum_{j=1}^{N}\sigma^x_j + h^y\sum_{j=1}^{N}\sigma^y_j + J\sum_{j=1}^{N-1}\sigma^x_j\sigma^x_{j+1},$$
$$h^x(t) = \begin{cases}(1-\delta)h^x, & \text{if } (t \bmod 2T)\in[0,T),\\ (1+\delta)h^x, & \text{if } (t \bmod 2T)\in[T,2T);\end{cases} \tag{C1}$$

i.e., we modulate the σ^x term by $\delta\, h^x$ with period $2T$. For ease of comparison, we use the same parameters as H_{QIMF}: $(h^x, h^y, J) = (0.8090, 0.9045, 1)$ and the same initial state $|\Psi_0\rangle = |0\rangle^{\otimes N}$. In Fig. 6, we present $\Delta^{(k)}$ data for Floquet parameters $(\delta, T) = (0.5, 1)$. The late-time saturation value is the same as the time-independent case but we see deviations from power-law decay. $\Delta^{(k)}$ reaches its saturated value more quickly under Floquet dynamics.

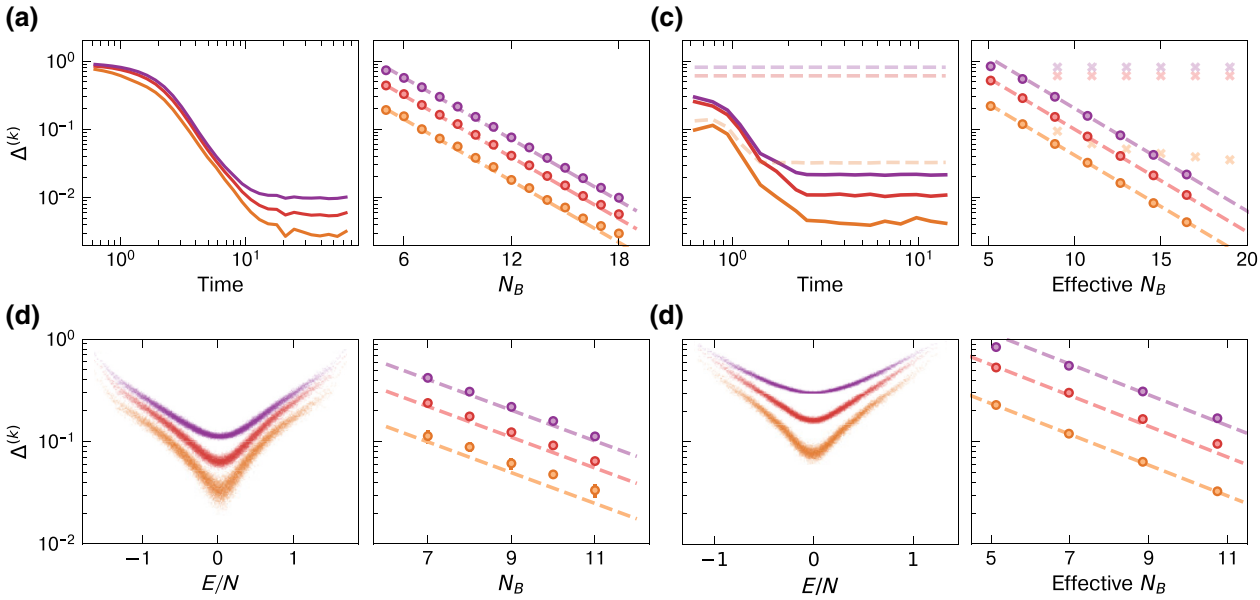

FIG. 7. Emergent quantum state designs in random-coupling and random-hopping models. (a),(b) The time-evolution (a) and eigenstate (b) data for the random all-to-all coupling model in Eq. (C2). The trace distances $\Delta^{(k)}$ for $k = 1, 2$, and 3 are plotted in orange, red, and purple, respectively, for $N_A = 3$ subsystems [(a), left]. The time-evolution data for $N = 21$ are plotted, along with the system-size scaling of the late-time $\Delta^{(k)}$ for $N = 8$ to 21 [(a), right]. We also plot $\Delta^{(k)}$ for each eigenstate for $N = 14$ [(b), left], along with the system-size scaling of the eigenstate data near $E = 0$ for $N = 10$ to 14 [(b), right]. (c),(d) The time-evolution (c) and eigenstate (d) data for the random-hopping model in Eq. (C3). Data are presented for $N_A = 5$ subsystems, in the $S^z_{\text{tot.}} = 0$ sector. In (c), time-evolution data for $N = 24$ are plotted along with system-size scaling for $N = 10, 12, \ldots, 24$. To illustrate the failure of projecting onto all strings z_B, we plot the results of this naive procedure with $N_A = 3$ in light colors and dashes. In (d), eigenstate data for $N = 16$ are plotted, along with their system-size scaling for $N = 12, 14, 16$ and 18. The system-size scaling is plotted against the effective N_B, which is $\log_2$ of the number of strings z_B and which is postselected. In each plot, the error bars are smaller than the marker sizes.

2. Random-coupling model

In the main text, we present results for the ergodic QIMF model. Here, we discuss an additional ergodic model: a spin-1/2 model with random all-to-all interactions,

$$H = \sum_{\substack{i,j=1 \\ i<j}}^{N} \sum_{\substack{\mu,\nu \in \{x,y,z\} \\ (\mu,\nu) \neq (z,z)}} J_{ij}^{\mu,\nu} \sigma_i^\mu \sigma_j^\nu , \qquad \text{(C2)}$$

where $J_{ij}^{\mu,\nu}$ are random variables drawn from independent identically distributed normal distributions: $J_{ij}^{\mu,\nu} \sim N(0, 1/N)$. Such random-coupling models are paradigmatic examples of quantum chaotic systems. The variance of the couplings is chosen so that $\text{Tr}(H^2)/2^N \sim O(N)$. Our model does not have $\sigma_i^z \sigma_j^z$ terms, so that our initial state $|\Psi_0\rangle = |0\rangle^{\otimes N}$ has zero energy and, accordingly, is regarded as being at infinite temperature.

As shown in Fig. 7(a), we see excellent convergence to k-designs for our projected ensembles constructed from time-evolved quenched states as well as eigenstates. The late-time $\Delta^{(k)}$ values are very close to $\Delta^{(k)}_{\text{em}}$. The average $\Delta^{(k)}$ values for infinite-temperature eigenstates are also close to $\Delta^{(k)}_{\text{em}}$ and show clear exponential decay [Fig. 7(b)]. Notably, we do not average over disorder realizations: each data point and time series is computed with fixed disorder realizations. This provides additional support for our conjecture that ergodic Hamiltonian systems generate approximate k-designs via time-evolved states and eigenstates at infinite temperature. Furthermore, the non-power-law behavior of $\Delta^{(k)}(t)$, along with our Floquet Hamiltonian results, supports our hypothesis that the power-law decay is due to emergent hydrodynamics associated with energy conservation.

3. Random-hopping model with U(1) symmetry

Having established that chaotic models such as the QIMF and the all-to-all random-coupling model provide projected ensembles that converge to k-designs, we next ask whether quantum models with symmetries do so as well. We specifically study a random-hopping model

$$H = \sum_{\substack{i,j=1 \\ i<j}}^{N} J_{ij}^{+} \left(\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y \right) + J_{ij}^{-} \left(\sigma_i^x \sigma_j^y - \sigma_i^y \sigma_j^x \right), \qquad \text{(C3)}$$

where $J_{ij}^{\pm}$ are random variables drawn from independent identically distributed normal distributions: $J_{ij}^{\pm} \sim$

$N(0, 1/N)$. Like the all-to-all random-coupling model, this model is also expected to exhibit chaos. However, the specific interaction terms are chosen such that this model has a $U(1)$ symmetry: there is conservation of the total magnetization $S^z_{\text{tot.}} = 1/2 \sum_j \sigma^z_j$. The model can equivalently be viewed as describing hard-core bosons with random complex all-to-all hopping amplitudes.

Any bit string in the z basis is a suitable infinite-temperature quench state. For N even, we choose the initial state $|\Psi_0\rangle = |0101\cdots01\rangle$, which lies in the $S^z_{\text{tot.}} = 0$ sector. Under time evolution by our chaotic Hamiltonian, the wave function is ergodic in the $S^z_{\text{tot.}} = 0$ Hilbert space.

Naively forming the projected ensemble, we do not find convergence to an approximate k-design. This is shown in the light curves in Fig. 7(c). With increasing system size, $\Delta^{(k)}$ for $k = 1, 2, 3$ saturate at a nonzero value. This is because of large correlations between the bit strings z_B and the projected state $|\Psi_A(z_B)\rangle$: if z_B has total magnetization s_B, $|\Psi_A(z_B)\rangle$ necessarily has total magnetization $S^z_{\text{tot.}} - s_B$ due to the global conservation law. The Hilbert space $\mathcal{H}_A$ naturally decomposes into a direct sum of multiple sectors enumerated by the magnetization s_A. Accordingly, the projected ensemble now produces multiple ensembles, one for each sector.

In order to properly account for the $U(1)$ conservation law, instead of projecting onto all strings z_B, we postselect a subset of all strings: bit strings with fixed total magnetization, e.g., $s_B = 1/2$. The projected states $|\Psi_A(z_B)\rangle$ will also have fixed magnetization and we then ask whether this ensemble forms a k-design in the subspace of $\mathcal{H}_A$ with magnetization $s_A = S^z_{\text{tot.}} - s_B$. In our numerical examples, we present results with $N_A = 5$ qubits and $s_A = -1/2$ and $s_B = 1/2$. The relevant subspace of $\mathcal{H}_A$ has dimension 10, far smaller than $2^{N_A} = 32$.

Figures 7(c) and 7(d) show the results of our symmetry resolution. We now see excellent convergence toward a k-design. As with the random-coupling model, the late-time $\Delta^{(k)}$ values are very close to $\Delta^{(k)}_{\text{em}}$. In order to make a fair comparison, we plot the late-time $\Delta^{(k)}$ against an "effective N_B," defined as $\log_2$ of the number of postselected strings z_B. $\Delta^{(k)}_{\text{em}}$ is also computed by sampling the same number of times.

Using this postselection procedure, we find excellent k-design convergence for projected ensembles from eigenstates near $E = 0$. Our results indicate that our basic approach remains valid for chaotic models, with additional symmetries if these symmetries are properly addressed.

APPENDIX D: FINITE-TEMPERATURE PROJECTED ENSEMBLES: SCROOGE ENSEMBLE

In Fig. 5(c), we compute the trace distance between the second moments of the projected ensemble and *the Scrooge ensemble*. The Scrooge, or GAP, ensemble has been introduced in Refs. [44,60]. Given a density matrix ρ, the corresponding Scrooge ensemble of states can be viewed as a "ρ distortion" of the Haar ensemble such that its first moment is exactly ρ.

In order to compute its kth moments, we use an explicit expression from Ref. [60]. The probability density of obtaining a wave function $|\psi\rangle$ is given by

$$p_{\text{Scrooge}}(\psi)d\psi = \frac{D!}{2\,\pi^D \det\rho}\,\langle\psi|\,\rho^{-1}\,|\psi\rangle^{-D-1}\,d\psi \tag{D1}$$

$$= \frac{D!}{2\,\pi^D \lambda_1\cdots\lambda_D}\,\frac{\prod_j r_j\,dr_j\,d\phi_j}{\left(|r_1|^2/\lambda_1 + \cdots |r_D|^2/\lambda_D\right)^{D+1}}\,, \tag{D2}$$

where the λ_j are the eigenvalues of ρ and $d\psi$ is the Haar measure. In Eq. (D2), we explicitly write the probability densities in the eigenbasis of ρ: $|\psi\rangle = \sum_j r_j \exp(i\phi_j)\,|\lambda_j\rangle$. This allows us to compute the kth moments $\rho^{(k)}_{\text{Scrooge}}$, in terms of its matrix elements:

$$\langle\lambda_{a_1}\cdots\lambda_{a_k}|\,\rho^{(k)}_{\text{Scrooge}}\,|\lambda_{b_1}\cdots\lambda_{b_k}\rangle = \int_{\sum_j r_j^2=1}\left(\prod_{j=1}^{D} dr_j^2 d\phi_j\right) p\left(\{r_j^2\}\right)\left(\prod_{l=1}^{k} r_{a_l} r_{b_l} e^{i(\phi_{a_l}-\phi_{b_l})}\right). \tag{D3}$$

After integrating over the phases $\{\phi_j\}$, the only nonzero matrix elements are those with $\{\lambda_{a_l}\} = \{\lambda_{b_l}\}$ or, equivalently, where the list of indices $\{a_l\}$ is a permutation of $\{b_l\}$. The remaining integral over the probability simplex $\{r_j^2 \mid r_j^2 > 0\ ,\ \sum_j r_j^2 = 1\}$ can be evaluated either numerically or analytically.

We use this expression for the second moment of the Scrooge ensemble to compute the trace distance $1/2\|\rho^{(2)} - \rho^{(2)}_{\text{Scrooge}}\|_1$ in Fig. 5(c).

APPENDIX E: PROOF SKETCH OF MAIN THEOREMS

We begin with a sketch of the proof of Theorem 1. We use streamlined notation for clarity. Recall that for a generator state $|\Phi\rangle$ on $\mathcal{H}$, the projected ensemble $\mathcal{E}_{\Phi,A} = \{p_z\ ,\ |\Phi_z\rangle\}$ has

$$p_z = \langle\Phi|P_z|\Phi\rangle \tag{E1}$$

$$|\Phi_z\rangle = \left(\mathbb{1}_A \otimes \langle z|_B\right)|\Phi\rangle/\sqrt{p_z}. \tag{E2}$$

It is convenient to define the unnormalized state

$$|\widetilde{\Phi}_z\rangle := \left(\mathbb{1}_A \otimes \langle z|_B\right)|\Phi\rangle \quad \text{for} \quad z \in \{0,1\}^{N_B} \tag{E3}$$

so that $p_z = \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$ and $|\Phi_z\rangle = |\widetilde{\Phi}_z\rangle/\sqrt{p_z}$. We note that, in our new notation, we can write

$$\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_{\Phi,A}}\left[\left(|\Psi\rangle\langle\Psi|\right)^{\otimes k}\right] \tag{E4}$$

$$= \sum_{z\in\{0,1\}^{N_B}} p_z \left(|\Phi_z\rangle\langle\Phi_z|\right)^{\otimes k} \tag{E5}$$

$$= \sum_{z\in\{0,1\}^{N_B}} \frac{\left(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|\right)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \equiv A(|\Phi\rangle). \tag{E6}$$

The key to establishing Theorem 1 is to understand the random tensor $A(|\Phi\rangle)$ when $|\Phi\rangle$ is a Haar-random state. Here, the Haar-random state is a normalized vector chosen uniformly at random in $d = d_A d_B$ complex dimensions, where $d = 2^N, d_A = 2^{N_A}, d_B = 2^{N_B}$ are the dimensions of the Hilbert space with N, N_A, and N_B qubits, respectively. Two ingredients are needed to understand $A(|\Phi\rangle)$:

(1) The expectation value of $A(|\Phi\rangle)$ over the Haar ensemble
(2) The concentration of $A(|\Phi\rangle)$ around its expectation: this tells us that, with high probability, $A(|\Phi\rangle)$ is close to its expectation

Let us discuss these two ingredients in more detail.

Using the fact that the uniform measure (i.e., the Haar measure) on the complex sphere is invariant under any unitary rotations, we can show that p_z and $|\Phi_z\rangle\langle\Phi_z|$ are independent random variables. Furthermore, $|\Phi_z\rangle\langle\Phi_z|$ is a uniform random vector from a d_A-dimensional complex sphere. Hence, we have

$$\mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)} A(|\Phi\rangle) = \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d_A)} \left(|\Psi\rangle\langle\Psi|\right)^{\otimes k}, \tag{E7}$$

which is the kth moment of the uniform ensemble. This means that the expectation of the kth moment of the projected ensemble of a randomly selected many-body wave function $|\Phi\rangle$ reproduces the kth moment of the Haar ensemble in d_A complex dimensions. Hence the expectation $\mathbb{E}_{\Phi\sim\mathrm{Haar}(d)}A(|\Phi\rangle)$ is exactly equal to the desired quantity.

To understand the concentration of $A(|\Phi\rangle)$ around its expectation, we recall the well-known result that the uniform distribution over a high-dimensional sphere has a very sharp concentration. For any Lipschitz function on the sphere, the fluctuation around its expectation value is small and the probability of having a large fluctuation decays exponentially. This is known as Levy's lemma and it allows us to upper bound the probability that $A(|\Phi\rangle)$ is far from its expectation. Together with the expectation identity, we have

$$\mathop{\mathrm{Prob}}_{\Phi\sim\mathrm{Haar}(d)}\left[\left\|A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d_A)}\left[(|\Psi\rangle\langle\Psi|)^{\otimes k}\right]\right\|_1 \geq \varepsilon\right]$$
$$\leq 2d_A^{2k}\,\exp\left(-\frac{d_B\,\varepsilon^2}{18\pi^3(2k-1)\,d_A^{4k}}\right). \tag{E8}$$

The asymptotic relation in Theorem 1 follows immediately from the above probabilistic statement. The detailed proof of Theorem 1 is given in Appendix F.

The proof of Theorem 2 is very different from that of Theorem 1. First, the expectation of $A(|\Phi\rangle)$ in Theorem 1 is computed using the invariance property of the Haar measure, which does not hold for a measure that only forms a state design. Furthermore, Levy's lemma only holds for the Haar measure, so we cannot resort to Levy's lemma to control its statistical fluctuations. To prove Theorem 2, we make use of a technique used in the context of solving linear systems on a quantum computer [51]. The key idea is to add a modulating function that approximates $A(|\Phi\rangle)$ by a polynomial function in $|\Phi\rangle$. In particular,

$$A(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \frac{\left(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|\right)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}}$$
$$\approx B(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \frac{\left(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|\right)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \mu_{k,b}(d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle), \tag{E9}$$

where the modulating function is given by

$$\mu_{k,b}(s) = (1-(1-s^{2(k-1)})^b) \quad \text{for } b \text{ even}. \tag{E10}$$

Using the binomial expansion, we can check that $B(|\Phi\rangle)$ is a polynomial function. Furthermore, $\mu_{k,b}(s)$ is very close to one when $s = d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$ is around one. Taking b larger allows us to better approximate the constant function $\mu_{k,b}(s) = 1$, which corresponds to the target expression $A(|\Phi\rangle)$. A visualization of $\mu_{k,b}(s)$ can be found in Fig. 8(a). If $|\Phi\rangle$ is sampled from the uniform measure on the quantum state space, then $s = d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$ will concentrate around one [see Fig. 8(b)]. So for $|\Phi\rangle$ sampled from the uniform measure, we can show that $A(|\Phi\rangle) \approx B(|\Phi\rangle)$. However, $|\Phi\rangle$ is sampled from a quantum state design, so there is no guarantee that $s = d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$ should concentrate around one. To address this, we utilize the following result proved in Lemma 5 (Appendix F):

$$\|A(|\Phi\rangle) - B(|\Phi\rangle)\|_1 \leq R(|\Phi\rangle), \tag{E11}$$

where $R(|\Phi\rangle)$ is a polynomial function given by

$$\sum_{z\in\{0,1\}^{N_B}} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\left(1-\mu_{k,b}\left(d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\right)\right). \tag{E12}$$

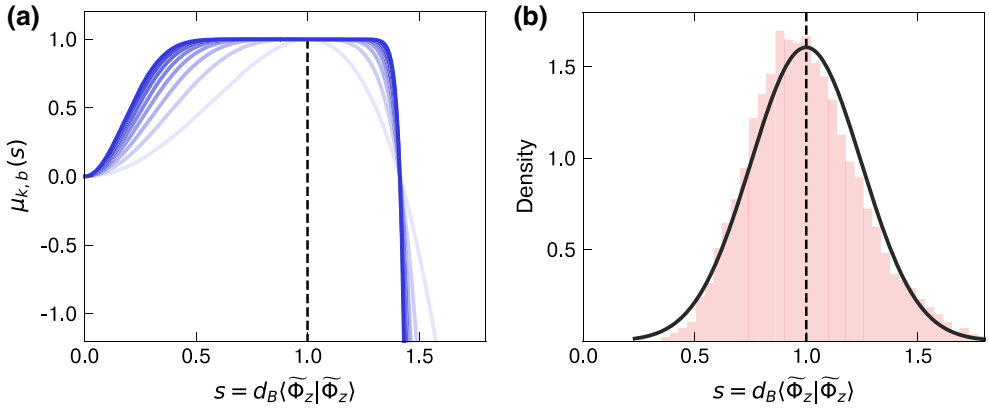

FIG. 8. A visualization of the modulating function and the concentration of $d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$. (a) The modulating function $\mu_{k,b}(s)$ defined in Eq. (E10) for varying values of b from 2 to 16. The darker colors correspond to higher b. We fix $k = 2$ in the plot. (b) The concentration of $s = d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$ for the Hilbert-space dimension of subsystem A, i.e., $d_A = 2^{N_A} = 16$, when $|\Phi\rangle$ is sampled uniformly at random from the complex sphere. We sample 10 000 different vectors $|\Phi\rangle$ to generate the histogram. The black solid line is a kernel-density-estimation fit of the histogram. If d_A is larger, then s is more concentrated.

Hence, the approximation error can be upper bounded by a polynomial function over which we have better control.

After introducing the key quantities, the proof proceeds by bounding the error between (i) $A(|\Phi\rangle)$, where $|\Phi\rangle$ is sampled from a state design, and (ii) the expectation of $A(|\Psi\rangle)$, where $|\Psi\rangle$ is from the Haar measure. We use $B(|\Phi\rangle)$ as an intermediate point of comparison:

$$\begin{aligned}&\mathop{\mathbb{E}}_{\Phi\sim(\varepsilon',k')\text{ design}}\left\|A(|\Phi\rangle)-\mathop{\mathbb{E}}_{\Psi\sim\text{Haar}}A(|\Psi\rangle)\right\|_1\\&\le\mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}}\|A(|\Phi\rangle)-B(|\Phi\rangle)\|_1\\&\quad+\mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}}\left\|B(|\Phi\rangle)-\mathop{\mathbb{E}}_{\text{Haar}}A(|\Psi\rangle)\right\|_1.\end{aligned}\tag{E13}$$

The first term can be upper bounded by $R(|\Phi\rangle)$ and we can use the fact that state designs approximate the expectation of any polynomial function under the Haar measure. Accordingly,

$$\begin{aligned}\mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}}\|A(|\Phi\rangle)-B(|\Phi\rangle)\|_1&\le\mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}}R(|\Phi\rangle)\\&\approx\mathop{\mathbb{E}}_{\text{Haar}}R(|\Phi\rangle).\end{aligned}\tag{E14}$$

In the second term, we can apply a similar idea by upper bounding the 1-norm with the 2-norm and utilizing the fact that $B(|\Phi\rangle)-\mathbb{E}_{\text{Haar}}A(|\Psi\rangle)$ is a polynomial function in $|\Phi\rangle$. This gives

$$\begin{aligned}&\mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}}\left\|B(|\Phi\rangle)-\mathop{\mathbb{E}}_{\text{Haar}}A(|\Psi\rangle)\right\|_1\\&\le\sqrt{d_A^k\mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}}\left\|B(|\Phi\rangle)-\mathop{\mathbb{E}}_{\text{Haar}}A(|\Psi\rangle)\right\|_2^2}\end{aligned}\tag{E15}$$

$$\approx\sqrt{d_A^k\mathop{\mathbb{E}}_{\text{Haar}}\left\|B(|\Phi\rangle)-\mathop{\mathbb{E}}_{\text{Haar}}A(|\Psi\rangle)\right\|_2^2}.\tag{E16}$$

We then approximate $B(|\Phi\rangle)$ by $A(|\Phi\rangle)$ in the above expression, where the error can be upper bounded by $R(|\Phi\rangle)^2$. This gives the approximate relation

$$\begin{aligned}&\mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}}\left\|B(|\Phi\rangle)-\mathop{\mathbb{E}}_{\text{Haar}}A(|\Psi\rangle)\right\|_1\\&\lesssim\sqrt{d_A^k\mathop{\mathbb{E}}_{\text{Haar}}\left(R(|\Phi\rangle)^2+\left\|A(|\Phi\rangle)-\mathop{\mathbb{E}}_{\text{Haar}}A(|\Psi\rangle)\right\|_2^2\right)}.\end{aligned}\tag{E17}$$

Using these steps, all our expressions contain only expectations over the Haar measure. More precisely,

$$\begin{aligned}&\mathop{\mathbb{E}}_{\Phi\sim(\varepsilon',k')\text{ design}}\left\|A(|\Phi\rangle)-\mathop{\mathbb{E}}_{\Psi\sim\text{Haar}}A(|\Psi\rangle)\right\|_1\\&\lesssim\mathop{\mathbb{E}}_{\text{Haar}}R(|\Phi\rangle)+\sqrt{d_A^k\mathop{\mathbb{E}}_{\text{Haar}}R(|\Phi\rangle)^2}\\&\quad+\sqrt{d_A^k\mathop{\mathbb{E}}_{\text{Haar}}\left\|A(|\Phi\rangle)-\mathop{\mathbb{E}}_{\text{Haar}}A(|\Psi\rangle)\right\|_2^2}.\end{aligned}\tag{E18}$$

The first two terms are small because

(1) $\mu_{k,b}(s)$ is close to one when s is close to one.
(2) If $|\Phi\rangle$ is sampled from the Haar measure, then $s = d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$ will concentrate around one.

See Fig. 8 for visualizations. The third term is small due to Theorem 1. Finally, we can resort to Markov's inequality

to show that with a probability of at least 0.99, we have that $\|A(|\Phi\rangle) - \mathbb{E}_{\Psi\sim\text{Haar}} A(|\Psi\rangle)\|_1$ is small. The full proof of Theorem 2 is given in Appendix F.

APPENDIX F: PROOF OF MAIN THEOREMS

1. Useful identities

We collect here several useful identities that we leverage throughout the remainder of the appendix. The first identity is

$$\mathop{\mathbb{E}}_{\Phi\sim\text{Haar}(d)}\left[(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|)^{\otimes k}\right] = \frac{\sum_{\pi\in S_k}\text{Perm}_{\mathcal{H}_A^{\otimes k}}(\pi)}{d(d+1)\cdots(d+k-1)}. \tag{F1}$$

Taking the trace, we find

$$\mathop{\mathbb{E}}_{\Phi\sim\text{Haar}(d)}\left[\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^k\right] = \frac{d_A(d_A+1)\cdots(d_A+k-1)}{d(d+1)\cdots(d+k-1)}. \tag{F2}$$

The related identity

$$\mathop{\mathbb{E}}_{\Phi\sim\text{Haar}(d)}\left[|\langle\widetilde{\Phi}_z|\widetilde{\Phi}_y\rangle|^{2k}\right] = \frac{k!\, d_A(d_A+1)\cdots(d_A+k-1)}{d(d+1)\cdots(d+2k-1)} \qquad \text{for } z\neq y \tag{F3}$$

is likewise useful.

2. Proof of Theorem 1

The main ingredient to prove Theorem 1 is Levy's lemma, which states that if a function is Lipschitz continuous, then it will concentrate around its expectation value if the variable is sampled uniformly from a high-dimensional complex sphere.

Lemma 1 (Levy's lemma). *Let $f : \mathbb{S}^{2d-1} \to \mathbb{R}$ satisfying $|f(v) - f(w)| \leq \eta\|v-w\|_2$. Then, for any $\delta \geq 0$, we have*

$$\begin{aligned}&\text{Prob}_{\Phi\sim\text{Haar}(d)}\left[\,\left|f(\Phi) - \mathbb{E}_{\Psi\sim\text{Haar}(d)}[f(\Psi)]\right| \geq \delta\right]\\ &\quad\leq 2\exp\left(-\frac{2d\,\delta^2}{9\pi^3\eta^2}\right).\end{aligned} \tag{F4}$$

Before applying Levy's lemma, we need to define a function of interest. In particular, we consider individual entries in the main quantity $\sum_{z\in\{0,1\}^{N_B}} p_z \left(|\Phi_z\rangle\langle\Phi_z|\right)^{\otimes k}$. Let $\{|i\rangle\}_{i\in\{0,1\}^{tN_A}}$ be the standard basis on $\mathcal{H}_A^{\otimes k}$ and write $|i\rangle = \bigotimes_{k=1}^t |i^{(k)}\rangle$, where each $|i^{(k)}\rangle \in \mathcal{H}_A$. Similarly, write $|j\rangle = \bigotimes_{k=1}^t |j^{(k)}\rangle$. Consider the function $f_{ij} : \mathbb{S}^{2d-1} \to \mathbb{R}$ defined by

$$f_{ij}(|\Phi\rangle) = \langle i|\left(\sum_{z\in\{0,1\}^{N_B}} \frac{\left(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|\right)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}}\right)|j\rangle \tag{F5}$$

$$= \sum_{z\in\{0,1\}^{N_B}} \frac{\prod_{\ell=1}^k \langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}}. \tag{F6}$$

A nice property of this function is that it is Lipschitz continuous and hence Levy's lemma can be applied.

Lemma 2 (Lipschitz constant). *A Lipschitz constant for f_{ij} is $\eta = 2(2k-1)$.*

Proof. Let $\vec{v}$ be an element of $\mathbb{S}^{2d-1}$. Then, $|\Phi\rangle = [\mathbb{1}\ i\mathbb{1}]\cdot\vec{v}$ is a normalized state in $\mathbb{C}^d$, where $\mathbb{1}$ denotes the $d\times d$ identity matrix. Here, we are using a mix of bra-ket notation and ordinary vector notation, where the latter is reserved for real-valued vectors. Shortly, we also use the notation $\mathbb{0}$ for the $d\times d$ matrix of all zeros. Since our f_{ij} is differentiable, we can choose any η such that

$$\eta \geq \left\|\frac{d}{d\vec{v}} f_{ij}(|\Phi\rangle)\right\|_2, \tag{F7}$$

where here, as above, $|\Phi\rangle = [\mathbb{1}\ i\mathbb{1}]\cdot\vec{v}$. We have

$$\begin{aligned}\left\|\frac{d}{d\vec{v}} f_{ij}(|\Phi\rangle)\right\|_2 &= \Bigg\|\sum_{\ell'=1}^k \sum_{z\in\{0,1\}^{N_B}} \frac{\prod_{\ell\neq\ell'}\langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \frac{d}{d\vec{v}}\langle i^{(\ell')}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell')}\rangle\\ &\quad - (k-1)\sum_{z\in\{0,1\}^{N_B}} \frac{\prod_{\ell=1}^k\langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^k}\frac{d}{d\vec{v}}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\Bigg\|_2\\ &\leq \Bigg\|\sum_{\ell'=1}^k \sum_{z\in\{0,1\}^{N_B}} \frac{\prod_{\ell\neq\ell'}\langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \frac{d}{d\vec{v}}\langle i^{(\ell')}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell')}\rangle\Bigg\|_2\\ &\quad + 2(k-1)\Bigg\|\sum_{z\in\{0,1\}^{N_B}} \frac{\prod_{\ell=1}^k\langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^k}\begin{bmatrix}P_z & \mathbb{0}\\ \mathbb{0} & P_z\end{bmatrix}\cdot\vec{v}\Bigg\|_2,\end{aligned} \tag{F8}$$

where we simply explicitly evaluate the derivatives and use the triangle inequality for the 2-norm. Let us bound each of the terms on the right-hand side in turn. For the first one, defining $M^{\pm}_{\ell'} = |i^{(\ell')}\rangle\langle j^{(\ell')}| \pm |j^{(\ell')}\rangle\langle i^{(\ell')}|$, we have

$$\left\| \sum_{\ell'=1}^{k} \sum_{z\in\{0,1\}^{N_B}} \frac{\prod_{\ell\neq\ell'}\langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \frac{d}{d\vec{v}} \langle i^{(\ell')}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell')}\rangle \right\|_2$$
$$= \left\| \sum_{\ell'=1}^{k} \sum_{z\in\{0,1\}^{N_B}} \frac{\prod_{\ell\neq\ell'}\langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \left(\begin{bmatrix} M^{+}_{\ell'} & -iM^{-}_{\ell'} \\ iM^{-}_{\ell'} & M^{+}_{\ell'} \end{bmatrix} \otimes |z\rangle\langle z|_B \right) \cdot \vec{v} \right\|_2 . \tag{F9}$$

Writing $b_{\ell',x} = (\prod_{\ell\neq\ell'}\langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle)/(\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1})$ and $M_{\ell'} = \begin{bmatrix} M^{+}_{\ell'} & -iM^{-}_{\ell'} \\ iM^{-}_{\ell'} & M^{+}_{\ell'} \end{bmatrix}$, the above evaluates to

$$\left(\sum_{\ell',p'=1}^{k} \sum_{z\in\{0,1\}^{N_B}} b^{*}_{p',x} b_{\ell',x}\ \vec{v}^{T} \cdot \left(M^{\dagger}_{p'} M_{\ell'} \otimes |z\rangle\langle z|_B \right) \cdot \vec{v} \right)^{1/2}$$
$$\leq \left(\sum_{\ell',p'=1}^{k} \sum_{z\in\{0,1\}^{N_B}} |b^{*}_{p',x}|\,|b_{\ell',x}|\,\langle\Phi| \left(|M^{\dagger}_{p'} M_{\ell'}| \otimes |z\rangle\langle z|_B \right) |\Phi\rangle \right)^{1/2}, \tag{F10}$$

where $|A| := \sqrt{A^{\dagger}A}$. But since

$$|b_{\ell',x}| = \left| \mathrm{tr}\left\{ \left(\bigotimes_{\ell\neq\ell'} |j^{(\ell)}\rangle\langle i^{(\ell)}| \right) \cdot \left(\frac{(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|)^{\otimes(k-1)}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \right) \right\} \right| \leq \left\| \bigotimes_{\ell\neq\ell'} |j^{(\ell)}\rangle\langle i^{(\ell)}| \right\|_2 \left\| \frac{(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|)^{\otimes(k-1)}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \right\|_2 = 1,$$

Eq. (F10) is less than or equal to

$$\left(\sum_{\ell',p'=1}^{k} \sum_{z\in\{0,1\}^{N_B}} \vec{v}^{T} \cdot \left(|M^{\dagger}_{p'} M_{\ell'}| \otimes |z\rangle\langle z|_B \right) \cdot \vec{v} \right)^{1/2} = \left(\sum_{\ell',p'=1}^{k} \vec{v}^{T} \cdot \left(|M^{\dagger}_{p'} M_{\ell'}| \otimes \mathbb{1}_B \right) \cdot \vec{v} \right)^{1/2}$$
$$\leq \left(\sum_{\ell',p'=1}^{k} \|M^{\dagger}_{p'} M_{\ell'}\|_{\infty} \right)^{1/2} . \tag{F11}$$

Since $\|M^{\dagger}_{p'} M_{\ell'}\|_{\infty} \leq \|M^{\dagger}_{p'}\|_{\infty} \|M_{\ell'}\|_{\infty} \leq 4$, the above is $\leq 2k$.

Now, we bound the term in the last line of Eq. (F8). Letting $c_z = (\prod_{\ell=1}^{k}\langle i^{(\ell)}|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|j^{(\ell)}\rangle)/(\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k})$, the term can be written as

$$2(k-1) \left(\sum_{z\in\{0,1\}^{N_B}} |c_z|^2\ \vec{v}^{T} \cdot \begin{bmatrix} P_z & \mathbb{0} \\ \mathbb{0} & P_z \end{bmatrix} \cdot \vec{v} \right)^{1/2} . \tag{F12}$$

Using the bound

$$|c_z| = \left| \mathrm{tr}\left\{ \left(\bigotimes_{\ell=1}^{k} |j^{(\ell)}\rangle\langle i^{(\ell)}| \right) \cdot \left(\frac{(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k}} \right) \right\} \right| \leq \left\| \bigotimes_{\ell=1}^{k} |j^{(\ell)}\rangle\langle i^{(\ell)}| \right\|_2 \left\| \frac{(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k}} \right\|_2 = 1,$$

we can upper bound Eq. (F12) by

$$2(k-1) \left(\sum_{z\in\{0,1\}^{N_B}} \vec{v}^{T} \cdot \begin{bmatrix} P_z & \mathbb{0} \\ \mathbb{0} & P_z \end{bmatrix} \cdot \vec{v} \right)^{1/2} = 2(k-1) \left(\sum_{z\in\{0,1\}^{N_B}} \vec{v}^{T} \cdot \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes P_z \right) \cdot \vec{v} \right)^{1/2} = 2(k-1). \tag{F13}$$

Putting our bounds together, we have

$$\left\| \frac{d}{d\vec{v}} f_{ij}(|\Phi\rangle) \right\|_2 \leq 2k + 2(k-1) = 2(2k-1), \tag{F14}$$

which is the desired result. ■

While the above bound on the Lipschitz constant, along with Levy's lemma, guarantees that each function f_{ij} concentrates around the its expectation value, the expectation value of f_{ij} has not yet been computed. The following lemma shows that the expectation value of f_{ij} is closely connected to the quantum state k-design on subsystem A.

Lemma 3 (**Expectation-value identity**). *We have the identity*

$$\mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)} \left[\sum_{z\in\{0,1\}^{N_B}} \frac{(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \right] = \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d_A)} \left[(|\Psi\rangle\langle\Psi|)^{\otimes k} \right]. \tag{F15}$$

Therefore, the expectation of f_{ij} is given by

$$\mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)} \left[f_{ij}(|\Phi\rangle) \right] = \langle i| \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d_A)} \left[(|\Psi\rangle\langle\Psi|)^{\otimes k} \right] |j\rangle . \tag{F16}$$

Proof. Observe that

$$\frac{(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} = (|\Phi_z\rangle\langle\Phi_z|)^{\otimes k} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle. \tag{F17}$$

Furthermore, from a technical lemma given below, we have that $|\Phi_z\rangle$ is independent from $\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$. Using this fact and the linearity of expectation, we compute

$$\begin{aligned} \mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)} \left[\sum_{z\in\{0,1\}^{N_B}} \frac{(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \right] &= \sum_{z\in\{0,1\}^{N_B}} \mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)} \left[(|\Phi_z\rangle\langle\Phi_z|)^{\otimes k} \right] \mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)} \left[\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \right] \\ &= \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d_A)} \left[(|\Psi\rangle\langle\Psi|)^{\otimes k} \right] \left(\sum_{z\in\{0,1\}^{N_B}} \frac{1}{d_B} \right), \end{aligned} \tag{F18}$$

as desired. ■

During the evaluation of the expectation value, we use a property that the normalized state $|\Phi_z\rangle\langle\Phi_z|$ and the corresponding probability p_z are independent random variables. This is proven in the following lemma.

Lemma 4. *If $|\Phi\rangle$ is a Haar-random state on $\mathcal{H}$, then the random variables $|\Phi_z\rangle\langle\Phi_z|$ and p_z are independent.*

Proof. Define the map $\mathcal{P}_z(|\Phi\rangle) := \langle\Phi|P_z|\Phi\rangle$ and let $\mathcal{R}_z$ be the map taking $|\Phi\rangle$ to the normalized state $|\Phi_z\rangle$:

$$\mathcal{R}_z(|\Phi\rangle) := \mathrm{tr}_B((\mathbb{1}_A \otimes |z\rangle\langle z|_B) \cdot |\Phi\rangle\langle\Phi|) / \mathcal{P}_z(|\Phi\rangle)). \tag{F19}$$

Let U_A be a Haar-random unitary operator on $\mathcal{H}_A$ and let $U = U_A \otimes \mathbb{1}_B$. Because U is unitary and $|\Phi\rangle$ is Haar random, the state $U|\Phi\rangle$ is also Haar random. Also note that

$$\mathcal{P}_z(U|\Phi\rangle) = \mathcal{P}_z(|\Phi\rangle)\,, \quad \mathcal{R}_z(U|\Phi\rangle) = U_A \mathcal{R}_z(|\Phi\rangle) U_A^\dagger, \tag{F20}$$

where the equivalence is in the sense of random variables. Therefore, for any functions F and G,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)} [F(\mathcal{R}_z(|\Phi\rangle))\, G(\mathcal{P}_z(|\Phi\rangle))] &= \mathop{\mathbb{E}}_{\substack{\Phi\,\sim\,\mathrm{Haar}(d)\\ U_A\,\sim\,U(d_A)}} [F(\mathcal{R}_z(U|\Phi\rangle))\, G(\mathcal{P}_z(|\Phi\rangle))] \\
&= \mathop{\mathbb{E}}_{\substack{\Phi\,\sim\,\mathrm{Haar}(d)\\ U_A\,\sim\,U(d_A)}} \left[F(U_A \mathcal{R}_z(|\Phi\rangle) U_A^\dagger)\, G(\mathcal{P}_z(|\Phi\rangle))\right] \\
&= \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d_A)} [F(|\Psi\rangle\langle\Psi|)] \mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)} [G(\mathcal{P}_z(|\Phi\rangle))]\,.
\end{aligned} \tag{F21}$$

To pass from the second line to the third, we use the fact that for any state $|\Psi\rangle$ on $\mathcal{H}_A$, the state $U_A|\Psi\rangle$ is an independent Haar-random state on $\mathcal{H}_A$. This shows that $\mathcal{R}_z(|\Phi\rangle)$ and $\mathcal{P}_z(|\Phi\rangle)$ are independent, as desired. ■

With the above lemmas, we now proceed with the proof of Theorem 1.

Proof of Theorem 1.. Let us consider the function $f_{ij}(|\Phi\rangle)$ defined in Eq. (F5). Because we have bound the Lipschitz constant $\eta \le 4k-2$, we can leverage Levy's lemma given in Lemma 1 to obtain

$$\mathrm{Prob}_{\Phi\sim\mathrm{Haar}(d)}\left[\left|f_{ij}(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)}\left[f_{ij}(|\Phi\rangle)\right]\right| \ge \varepsilon\right] \le 2\,\exp\left(-\frac{d\,\varepsilon^2}{18\pi^3(2k-1)}\right). \tag{F22}$$

Performing a union bound and rescaling $\varepsilon \to \varepsilon/d_A^{2k}$, we have

$$\mathrm{Prob}_{\Phi\sim\mathrm{Haar}(d)}\left[\left|f_{ij}(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)}\left[f_{ij}(|\Phi\rangle)\right]\right| \ge \frac{\varepsilon}{d_A^{2k}}\,,\ \text{for some}\ i,j\right] \le 2d_A^{2k}\,\exp\left(-\frac{d\,\varepsilon^2}{18\pi^3(2k-1)\,d_A^{4k}}\right). \tag{F23}$$

By comparing with the definition of f_{ij} in Eq. (F5) and using Lemma 3 to obtain the expectation value of f_{ij}, the above is equivalent to the following concentration inequality:

$$\begin{aligned}
&\mathrm{Prob}_{\Phi\sim\mathrm{Haar}(d)}\left[\left\|\sum_{z\in\{0,1\}^{N_B}} \frac{\left(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|\right)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} - \mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d)}\left[\sum_{z\in\{0,1\}^{N_B}} \frac{\left(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|\right)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}}\right]\right\|_{\text{entrywise, 1}} \ge \varepsilon\right] \\
&\le 2d_A^{2k}\,\exp\left(-\frac{d\,\varepsilon^2}{18\pi^3(2k-1)\,d_A^{4k}}\right),
\end{aligned} \tag{F24}$$

where $\|A\|_{\text{entrywise, 1}} = \sum_{i,j}|A_{ij}|$. Finally, using $\|A\|_{\text{entrywise, 1}} \ge \|A\|_1$ and applying Lemma 3, we find that

$$\mathrm{Prob}_{\Phi\sim\mathrm{Haar}(d)}\left[\left\|\sum_{z\in\{0,1\}^{N_B}} \frac{\left(|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|\right)^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} - \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d_A)}\left[(|\Psi\rangle\langle\Psi|)^{\otimes k}\right]\right\|_1 \ge \varepsilon\right] \le 2d_A^{2k}\,\exp\left(-\frac{d_B\,\varepsilon^2}{18\pi^3(2k-1)\,d_A^{4k}}\right). \tag{F25}$$

We can see that if we have

$$d_B \ge \frac{18\pi^3(2k-1)d_A^{4k}}{\varepsilon^2}\left(2k\log(d_A) + \log(2/\delta)\right), \tag{F26}$$

then the ensemble $\mathcal{E}_{\Phi,A}$ forms an ε-approximate t-design with probability at least $1-\delta$. Recall that $d_A = 2^{N_A}, d_B = 2^{N_B}$; hence taking a logarithm on both side of the above condition gives rise to

$$N_B = \Omega\left(kN_A + \log\left(\frac{1}{\varepsilon}\right) + \log\log\left(\frac{1}{\delta}\right)\right), \tag{F27}$$

which is the result stated in the main text. ■

3. Proof of Theorem 2

For convenience, we restate Theorem 2 in the following.

Theorem 3 (Restatement of Theorem 2). *Let $|\Psi\rangle$ be a state chosen from an ensemble on $\mathcal{H}$ that forms an ε'-approximate k'-design. Then the projected ensemble $\mathcal{E}_{A,\Psi}$ forms an ε-approximate k-design with probability at least $1-\delta$ if*

$$N_B = \Omega\left(kN_A + \log\left(\tfrac{1}{\varepsilon\delta}\right)\right), \tag{F28}$$

$$k' = \Omega\left(k\left(N_B + \log\left(\tfrac{1}{\varepsilon\delta}\right)\right)\right), \tag{F29}$$

$$\log\left(\tfrac{1}{\varepsilon'}\right) = \Omega\left(kN_B\left(N_B + \log\left(\tfrac{1}{\varepsilon\delta}\right)\right)\right), \tag{F30}$$

$$N_A = \Omega\left(\log(N_B) + \log(k) + \log\log\left(\tfrac{1}{\varepsilon\delta}\right)\right). \tag{F31}$$

In the following subsections, we begin with a discussion of a generalized Levy's lemma that provides sharper concentration for quadratic functions. We then give a general structure of the proof and provide the detailed proof of several technical lemmas afterward.

a. Higher-order concentration

In the proof of Theorem 1, Levy's lemma plays a crucial role in establishing the desired statement. Levy's lemma tells us about a sharp concentration when the random variables are sampled from a high-dimensional sphere. We continue to utilize concentration on high-dimensional spheres for the proof of Theorem 2. However, the original statement of Levy's lemma does not provide good bounds for Theorem 2. We instead make use of a higher-order variant of Levy's lemma. In this section, we recall a higher-order variant of Levy's lemma that has been established in Ref. [50] and use it to provide a simple proof of a concentration inequality for the quantity $\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$. The higher-order concentration results provided in Ref. [50] are a useful tool for obtaining concentration inequalities for polynomial functions on measures satisfying a log-Sobolev inequality and may prove useful in other problems arising in quantum information theory. For this work, we only need the following result from Ref. [50].

Theorem 4 (Theorem 1.13 in [50]). *Let f be a C^2-smooth function on an open neighborhood of the sphere $\mathbb{S}^{N-1}$ with $\int_{\mathbb{S}^{N-1}} f\, d\sigma_{N-1} = 0$. If*

$$\int_{\mathbb{S}^{N-1}} \|\nabla f\|_2^2\, d\sigma_{N-1} \le \frac{1}{N} \tag{F32}$$

and $\|\operatorname{Hess} f(\theta)\|_\infty \le 1$ for all $\theta \in \mathbb{S}^{N-1}$, then

$$\int_{\mathbb{S}^{N-1}} \exp((N-1)|f|/(8e))\, d\sigma_{N-1} \le 2. \tag{F33}$$

As a corollary, we obtain the following result for quadratic forms.

Corollary 1 (Concentration for quadratic forms). *Let $Q:\mathbb{R}^N \to \mathbb{R}^N$ be a linear map satisfying $\|Q\|_2 \le 1/2$ and $\|Q\|_\infty \le 1/2$ and let*

$$q_0 = \int_{\mathbb{S}^{N-1}} \theta^T Q\theta\, d\sigma_{N-1}(\theta).$$

Then,

$$\int_{\mathbb{S}^{N-1}} \exp((N-1)(\theta^T Q\theta - q_0)/(8e))\, d\sigma_{N-1} \le 2. \tag{F34}$$

Proof. Let $f(x) = x^T Qx - a_0$. Then, $\nabla f(x) = 2Qx$, so

$$\int_{\mathbb{S}^{N-1}} \|\nabla f\|_2^2\, d\sigma_{N-1} = 4\int_{\mathbb{S}^{N-1}} \theta^T Q^T Q\theta\, d\sigma_{N-1}(\theta). \tag{F35}$$

Let $\{u_j\}_{j=1}^N$ be an orthonormal basis of eigenvectors of Q^TQ with eigenvalues λ_j^2. Then, the latter integral is equal to

$$4\sum_j \lambda_j^2 \int_{\mathbb{S}^{N-1}} (\theta\cdot u_j)^2\, d\sigma_{N-1}(\theta) = \frac{4}{N}\sum_j \lambda_j^2 = 4n^{-1}\|Q\|_2^2. \tag{F36}$$

If $\|Q\|_2 \le 1/2$, then f satisfies the condition given in Eq. (F32). Moreover, $\operatorname{Hess} f = 2Q$, so $\|\operatorname{Hess} f\|_\infty = 2\|M\|_\infty$. We can therefore apply Theorem 4 to f to obtain (F34). ■

We only need to use Corollary (1) in the case in which Q is an orthogonal projection.

Corollary 2 (Concentration for quadratic forms with projectors). *Let $V \subset \mathbb{R}^N$ be a subspace of $\mathbb{R}^N$ with dimension m and let P_V be the orthogonal projection onto V. Then,*

$$\mathop{\mathbb{E}}_{\theta\sim\mathbb{S}^{N-1}}\left[\exp\left(\frac{(N-1)|\|P_V\theta\|_2^2 - m/N|}{8e\sqrt{2m}}\right)\right] \le 2 \tag{F37}$$

and, in particular, for θ sampled uniformly from the sphere,

$$\begin{aligned}&Prob_{\theta\sim\mathbb{S}^{N-1}}\left[|\|P_V\theta\|^2 - m/N| \ge \delta\right]\\ &\quad\le 2\exp\left(-\frac{1}{8e\sqrt{2}}(N-1)m^{-1/2}\delta\right).\end{aligned} \tag{F38}$$

Proof. Observe that $\|P_V\|_2^2 = m$ and $\|P_V\|_\infty = 1$, so the operator $(2m)^{-1/2}P_V$ satisfies the conditions of Corollary

1, from which Eq. (F37) follows. Then Eq. (F38) follows from Markov's inequality applied to Eq. (F37). ■

Specializing to the case in which $|\Phi\rangle$ is a Haar-random quantum state on $\mathcal{H}$ and P_V is the projector $P_z = \mathbb{1}_A \otimes |z\rangle\langle z|$, we deduce an exponential concentration inequality for $\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle$.

Corollary 3 (Concentration for probability in a projected ensemble). *We have the bound*

$$Prob_{\Phi\sim Haar(d)}\left[\left|\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle - \frac{1}{d_B}\right| \geq \delta\right] \leq 2\exp\left(-\frac{1}{8e\sqrt{2}}\, d_A^{1/2}\, d_B\, \delta\right). \tag{F39}$$

b. General structure of the proof

The proof is based on the idea of modulating the target expression, which is a rational function, with a high-degree polynomial to form a polynomial function. More precisely, we are interested in the following two functions:

$$A(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \frac{|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}}, \tag{F40}$$

$$B(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \frac{|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \times \left(1 - \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\right)^{2(k-1)}\right)^b\right), \tag{F41}$$

where $r, b > 0$ are parameters for tuning the approximation of polynomial function $B(|\Phi\rangle)$ to the target expression $A(|\Phi\rangle)$, which is a rational function. Because the 1 is cancelled in the binomial expansion

$$1 - \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\right)^{2(k-1)}\right)^b = \sum_{p=1}^{b}\binom{b}{p}(-1)^p\left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\right)^{2(k-1)p}, \tag{F42}$$

it is not hard to see that $B(|\Phi\rangle)$ is indeed a polynomial function in the real and imaginary parts of $|\Phi\rangle$.

We consider $|\Phi\rangle$ to be sampled from an (ε', k')-design. We use the basic Markov inequality to bound the concentration. A higher-order concentration inequality can also be used but would require k' to be higher. The central quantity in Markov inequality is the expectation value of the error

$$\mathop{\mathbb{E}}_{\Phi\sim(\varepsilon',k')\text{ design}} \left\| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle)\right\|_1. \tag{F43}$$

We use $B(|\Phi\rangle)$ as a surrogate to obtain an upper bound on the above quantity. This is because $B(|\Phi\rangle)$ is a polynomial function rather than a rational function in the real and imaginary parts of $|\Phi\rangle$. A triangle inequality gives the following bound:

$$\mathop{\mathbb{E}}_{\Phi\sim(\varepsilon',k')\text{ design}} \left\| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle)\right\|_1 \tag{F44}$$

$$\leq \mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}} \|A(|\Phi\rangle) - B(|\Phi\rangle)\|_1 + \mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}} \left\| B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle)\right\|_1. \tag{F45}$$

We can now analyze the two terms independently by using properties of quantum designs. For the first term, we prove in Lemma 5 that the following inequality holds:

$$\|A(|\Phi\rangle) - B(|\Phi\rangle)\|_1 \leq \sum_{z\in\{0,1\}^{N_B}} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\right)^{2(k-1)}\right)^b. \tag{F46}$$

Therefore, we define a polynomial function

$$R(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\right)^{2(k-1)}\right)^b, \tag{F47}$$

which is an upper bound on the approximation error between $A(|\Phi\rangle)$ and $B(|\Phi\rangle)$. Because a quantum design approximates any polynomial function, we have the following upper bound:

$$\mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}} \|A(|\Phi\rangle) - B(|\Phi\rangle)\|_1 \leq \mathop{\mathbb{E}}_{(\varepsilon',k')\text{ design}} R(|\Phi\rangle) \tag{F48}$$

$$\leq \mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle) + \varepsilon' \times \mathrm{E}_1, \tag{F49}$$

where the exact expression of E_1 is given in Lemma 6. We also apply a similar philosophy for bounding the second term in Eq. (F45) by first upper bounding the term by a

polynomial function that turns expectations over designs into expectations over the Haar measure:

$$\mathop{\mathbb{E}}_{(\varepsilon',k')\,\text{design}} \left\| B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \right\|_1$$

$$\leq \mathop{\mathbb{E}}_{(\varepsilon',k')\,\text{design}} \sqrt{d_A^k} \left\| B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \right\|_2 \tag{F50}$$

$$\leq \sqrt{\mathop{\mathbb{E}}_{(\varepsilon',k')\,\text{design}} d_A^k \left\| B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \right\|_2^2} \tag{F51}$$

$$\leq \sqrt{\mathop{\mathbb{E}}_{\text{Haar}} d_A^k \left\| B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \right\|_2^2 + \varepsilon' \times \mathrm{E}_2}. \tag{F52}$$

The first line follows from the relation between the 1-norm and the 2-norm. The second line applies Jensen's inequality. The third line follows from the observation that $d_A^k \, \| B(|\Phi\rangle) - \mathbb{E}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \|_2^2$ is a polynomial function in $|\Phi\rangle$. Therefore, we can turn the expectation over a quantum state design to one over the Haar measure, which incurs a small error of $\varepsilon' \times \mathrm{E}_2$. The exact expression of E_2 is given in Lemma 7. We now upper bound $\mathbb{E}_{\text{Haar}} \, \| B(|\Phi\rangle) - \mathbb{E}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \|_2^2$ by turning $B(|\Phi\rangle)$ back into $A(|\Phi\rangle)$ and incurring an additional error:

$$\mathop{\mathbb{E}}_{\text{Haar}} \left\| B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \right\|_2^2 \leq 2 \mathop{\mathbb{E}}_{\text{Haar}} \| B(|\Phi\rangle) - A(|\Phi\rangle) \|_2^2 + 2 \mathop{\mathbb{E}}_{\text{Haar}} \| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \|_2^2 \tag{F53}$$

$$\leq 2 \mathop{\mathbb{E}}_{\text{Haar}} \| B(|\Phi\rangle) - A(|\Phi\rangle) \|_1^2 + 2 \mathop{\mathbb{E}}_{\text{Haar}} \left\| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \right\|_2^2 \tag{F54}$$

$$\leq 2 \mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle)^2 + 2 \mathop{\mathbb{E}}_{\text{Haar}} \| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \|_2^2. \tag{F55}$$

The first line follows from the triangle inequality and $(a+b)^2 \leq 2(a^2+b^2), \forall a, b \in \mathbb{R}$. The second line follows from the relation between the 1-norm and 2-norm. The third line uses the inequality given in Eq. (F46), which is proved in Lemma 5. We can now combine Eqs. (F45), (F49), (F52), and (F55) to find

$$\mathop{\mathbb{E}}_{\Phi\sim(\varepsilon',k')\,\text{design}} \left\| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \right\|_1 \tag{F56}$$

$$\leq \mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle) + \varepsilon' \times \mathrm{E}_1 + \sqrt{2 \mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle)^2 + 2 \mathop{\mathbb{E}}_{\text{Haar}} \left\| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} A(|\Psi\rangle) \right\|_2^2 + \varepsilon' \times \mathrm{E}_2}. \tag{F57}$$

In Lemma 8, we give upper bounds for $\mathbb{E}_{\text{Haar}} R(|\Phi\rangle)$ and $\mathbb{E}_{\text{Haar}} R(|\Phi\rangle)^2$. Lemma 8 is the key to Theorem 2. The goodness of fit of the polynomial approximation $B(|\Phi\rangle)$ to the target expression $A(|\Phi\rangle)$ is reflected in upper bound of $\mathbb{E}_{\text{Haar}} R(|\Phi\rangle)$ and $\mathbb{E}_{\text{Haar}} R(|\Phi\rangle)^2$. By changing the polynomial approximation, one may likely obtain an improved statement of Theorem 2. We leave open the choice of the optimal polynomial approximation and focus mainly on the simple and tractable polynomial approximation given in Eq. (F41). The upper bounds for $\mathbb{E}_{\text{Haar}} R(|\Phi\rangle)$ and $\mathbb{E}_{\text{Haar}} R(|\Phi\rangle)^2$ rely on the higher-order variant of Levy's lemma given in Sec. F 3 a. In Lemma 9, we obtain an upper bound on $\mathbb{E}_{\text{Haar}} \, \| A(|\Phi\rangle) - \mathbb{E}_{\text{Haar}} A(|\Phi\rangle) \|_2^2$. Along with Lemma 6 that bounds $\varepsilon' E_1$ and Lemma 7 that bounds $\varepsilon' E_2$, we can show that for any $\varepsilon > 0$, as long as we choose $r = 2$ and b to be an even integer with

$$b = \Omega\left(N_B + \log(1/\varepsilon)\right), \tag{F58}$$

$$N_A = \Omega\left(\log(N_B) + \log(k) + \log\log(1/\varepsilon)\right), \tag{F59}$$

$$k' = \Omega\left(bk\right), \tag{F60}$$

$$\log(1/\varepsilon') = \Omega\left(\log(1/\varepsilon) + k(bN_B + \log k + N_A)\right), \tag{F61}$$

$$N_B = \Omega\left(kN_A + \log\left(\frac{1}{\varepsilon}\right)\right), \tag{F62}$$

we can obtain the following upper bounds:

$$\varepsilon' \times \mathrm{E}_1 \le \varepsilon, \qquad \varepsilon' \times \mathrm{E}_2 \le \varepsilon^2, \tag{F63}$$

$$\mathop{\mathbb{E}}_{\mathrm{Haar}} R(|\Phi\rangle) \le \varepsilon, \qquad \mathop{\mathbb{E}}_{\mathrm{Haar}} R(|\Phi\rangle)^2 \le \varepsilon^2, \tag{F64}$$

$$\mathop{\mathbb{E}}_{\mathrm{Haar}} \left\| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}} A(|\Psi\rangle) \right\|_2^2 \le \varepsilon^2. \tag{F65}$$

Together with Markov's inequality, we have the following concentration result:

$$\begin{aligned}\Pr_{\Phi\sim(\varepsilon',k')\,\mathrm{design}} \left[\left\| A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}} A(|\Psi\rangle) \right\|_1 \ge \tilde{\varepsilon} \right] \\ \times \le \frac{2\varepsilon + \sqrt{5\varepsilon^2}}{\tilde{\varepsilon}} = 5\,\frac{\varepsilon}{\tilde{\varepsilon}}\end{aligned} \tag{F66}$$

for any $\varepsilon, \tilde{\varepsilon} > 0$. Using Lemma 3 established in the proof of Theorem 1, we have

$$\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}} A(|\Psi\rangle) = \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d_A)} \left[(|\Psi\rangle\langle\Psi|)^{\otimes k} \right]. \tag{F67}$$

Therefore, the projected ensemble of a randomly sample state $|\Phi\rangle$ forms an (ε, k)-design on the local subsystem with dimension $d_A = 2^{N_A}$ under probability at least $1-\delta$ as long as the following conditions hold:

$$N_B = \Omega\left(kN_A + \log\left(\frac{1}{\varepsilon\delta}\right)\right), \tag{F68}$$

$$k' = \Omega\left(k\left(N_B + \log\left(\frac{1}{\varepsilon\delta}\right)\right)\right), \tag{F69}$$

$$\log(1/\varepsilon') = \Omega\left(kN_B\left(N_B + \log\left(\frac{1}{\varepsilon\delta}\right)\right)\right), \tag{F70}$$

$$N_A = \Omega\left(\log(N_B) + \log(k) + \log\log\left(\frac{1}{\varepsilon\delta}\right)\right). \tag{F71}$$

This concludes the proof of Theorem 2.

c. Technical lemmas

Lemma 5 (Error bound on polynomial approximation). *If b is even, then $\|A(|\Phi\rangle) - B(|\Phi\rangle)\|_1 \le R(|\Phi\rangle)$.*

Proof. Recall the following definitions:

$$A(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \frac{|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}}, \tag{F72}$$

$$\begin{aligned}B(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \frac{|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \\ \times \left(1 - \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \right)^{2(k-1)} \right)^b \right),\end{aligned} \tag{F73}$$

$$R(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \right)^{2(k-1)} \right)^b. \tag{F74}$$

We note the definition of trace norm $\|\cdot\|_1$ for Hermitian matrices:

$$\|X\|_1 = \sup_{O:\|O\|_\infty\le 1} \mathrm{tr}(OX). \tag{F75}$$

Hence, we have

$$\begin{aligned}&\|A(|\Phi\rangle) - B(|\Phi\rangle)\|_1 \\ &= \sup_{O:\|O\|_\infty\le 1} \mathrm{tr}\left(O \sum_{z\in\{0,1\}^{N_B}} \frac{|\widetilde{\Phi}_z\rangle\langle\widetilde{\Phi}_z|^{\otimes k}}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \right. \\ &\quad \left. \times \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \right)^{2(k-1)} \right)^b \right)\end{aligned} \tag{F76}$$

$$\begin{aligned}&= \sup_{O:\|O\|_\infty\le 1} \sum_{z\in\{0,1\}^{N_B}} \frac{\left(\langle\widetilde{\Phi}_z|^{\otimes k}\right) O \left(|\widetilde{\Phi}_z\rangle^{\otimes k}\right)}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \\ &\quad \times \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \right)^{2(k-1)} \right)^b\end{aligned} \tag{F77}$$

$$\begin{aligned}&\le \sup_{O:\|O\|_\infty\le 1} \sum_{z\in\{0,1\}^{N_B}} \frac{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^t}{\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^{k-1}} \\ &\quad \times \left| \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \right)^{2(k-1)} \right)^b \right|\end{aligned} \tag{F78}$$

$$\le \sup_{O:\|O\|_\infty\le 1} \sum_{z\in\{0,1\}^{N_B}} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \left(1 - \left(\frac{d_B}{r}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \right)^{2(k-1)} \right)^b \tag{F79}$$

$$= R(|\Phi\rangle). \tag{F80}$$

This concludes the proof. ■

Lemma 6 (Quantum state design on the error bound). *For $k' \ge 2b(k-1)+1$, we have*

$$\mathop{\mathbb{E}}_{(\varepsilon',k')\,\mathrm{design}} R(|\Phi\rangle) \le \mathop{\mathbb{E}}_{\mathrm{Haar}} R(|\Phi\rangle) + \varepsilon' E_1,$$

where the error term is given by

$$E_1 = \left(1 + \left(\frac{d_B}{r}\right)^{2(k-1)}\right)^b. \tag{F81}$$

In particular, if we choose $r = 2 \leq d_B$ *and recall that* $d_B = 2^{N_B}$*, then for any* $\varepsilon > 0$ *as long as*

$$k' = \Omega\left(bk\right), \tag{F82}$$

$$\log(1/\varepsilon') = \Omega\left(\log(1/\varepsilon) + bkN_B\right), \tag{F83}$$

we have the following upper bound on the error term: $\varepsilon' E_1 \leq \varepsilon$.

Proof. Letting $M_{zz} = \mathbb{1}_A \otimes |z\rangle\langle z|$, for $k' \geq \ell$, we have

$$\begin{aligned}
&\left| \sum_{z\in\{0,1\}^{N_B}} \mathop{\mathbb{E}}_{(\varepsilon',k') \text{ design}} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^\ell - \sum_{z\in\{0,1\}^{N_B}} \mathop{\mathbb{E}}_{\text{Haar}} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^\ell \right| \\
&= \left| \operatorname{tr}\left(\left(\mathbb{1}^{\otimes(k'-\ell)} \otimes \sum_{z\in\{0,1\}^{N_B}} M_{zz}^{\otimes\ell} \right) \cdot \left(\mathop{\mathbb{E}}_{(\varepsilon',k') \text{ design}} (|\Phi\rangle\langle\Phi|)^{\otimes k'} - \mathop{\mathbb{E}}_{\text{Haar}} (|\Phi\rangle\langle\Phi|)^{\otimes k'} \right) \right) \right| \\
&\leq \left\| \mathbb{1}^{\otimes(k'-\ell)} \otimes \sum_{z\in\{0,1\}^{N_B}} M_{zz}^{\otimes\ell} \right\|_\infty \left\| \mathop{\mathbb{E}}_{(\varepsilon',k') \text{ design}} (|\Phi\rangle\langle\Phi|)^{\otimes k'} - \mathop{\mathbb{E}}_{\text{Haar}} (|\Phi\rangle\langle\Phi|)^{\otimes k'} \right\|_1 \\
&\leq \varepsilon'
\end{aligned} \tag{F84}$$

where in the last inequality we use $\left\| \mathbb{1}^{\otimes(k'-\ell)} \otimes \sum_{z\in\{0,1\}^{N_B}} M_{zz}^{\otimes\ell} \right\|_\infty = 1$. Upon examining Eq. (F47), we see that for $k' \geq 2b(k-1)+1$, we can leverage (F84) to achieve

$$\mathop{\mathbb{E}}_{(\varepsilon',k') \text{ design}} R(|\Phi\rangle) \leq \mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle) + \varepsilon' \left(1 + \left(\frac{d_B}{r}\right)^{2(k-1)}\right)^b, \tag{F85}$$

which is the desired bound. ■

Lemma 7 (Quantum state design on the polynomial approximation). *For* $k' \geq 4b(k-1)+2$*,*

$$\mathop{\mathbb{E}}_{(\varepsilon',k') \text{ design}} d_A^k \left\| B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} [A(|\Psi\rangle)] \right\|_2^2 \leq \mathop{\mathbb{E}}_{\text{Haar}} d_A^k \left\| B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}} [A(|\Psi\rangle)] \right\|_2^2 + \varepsilon' E_2, \tag{F86}$$

where the error term is given by

$$E_2 = \left(2\,k! \left[-1 + \left(1 + \left(\frac{d_B}{r}\right)^{2(k-1)}\right)^b \right] + d_A^k \left[-1 + \left(1 + \left(\frac{d_B}{r}\right)^{2(k-1)}\right)^b \right]^2 \right). \tag{F87}$$

In particular, if we choose $r = 2 \leq d_B$ *and recall that* $d_B = 2^{N_B}$*, then for any* $\varepsilon > 0$ *as long as*

$$k' = \Omega\left(bk\right), \tag{F88}$$

$$\log(1/\varepsilon') = \Omega\left(\log(1/\varepsilon) + k(bN_B + \log k + N_A)\right), \tag{F89}$$

we have the following upper bound on the error term: $\varepsilon' E_2 \leq \varepsilon^2$.

Proof. Recall from Lemma 3 that

$$\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(d)}[A(|\Psi\rangle)] = \mathop{\mathbb{E}}_{\Phi\sim\mathrm{Haar}(d_A)}(|\Phi\rangle\langle\Phi|)^{\otimes k} = \frac{\Pi_{A,k}}{\binom{d_A+k-1}{k}} \tag{F90}$$

where in the last line, $\Pi_{A,k}$ is the projector onto the symmetric subspace of $\mathcal{H}_A^{\otimes k}$. Then, we can write $\|B(\Phi) - \mathbb{E}_{\Psi\sim\mathrm{Haar}}[A(|\Psi\rangle)]\|_2^2$ as

$$\left\|B(\Phi) - \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}}[A(|\Psi\rangle)]\right\|_2^2 = \mathrm{tr}(B(\Phi)^2) - \frac{2\,\mathrm{tr}(B(\Phi))}{\binom{d_A+k-1}{k}} + \frac{1}{\binom{d_A+k-1}{k}}. \tag{F91}$$

Using Eq. (F84), we have for $k' \geq 2b(k-1)+1$ that

$$d_A^k \mathop{\mathbb{E}}_{(\varepsilon',k')\,\mathrm{design}}\left(-\frac{2\,\mathrm{tr}(B(\Phi))}{\binom{d_A+k-1}{k}}\right) \leq d_A^k \mathop{\mathbb{E}}_{\mathrm{Haar}}\left(-\frac{2\,\mathrm{tr}(B(\Phi))}{\binom{d_A+k-1}{k}}\right) + \varepsilon'\cdot 2\,k!\left[-1+\left(1+\left(\frac{d_B}{r}\right)^{2(k-1)}\right)^b\right]. \tag{F92}$$

To obtain a similar bound on $d_A^k\mathbb{E}_{(\varepsilon',k')\,\mathrm{design}}\mathrm{tr}(B(\Phi)^2)$, we need the following generalization of the inequality in (F84). Letting $M_{yz} = \mathbb{1}_A \otimes |y\rangle\langle z|$, for $k' \geq 2(p+q)$ we have

$$\begin{aligned}
&\left|\sum_{y,z\in\{0,1\}^{N_B}} \mathop{\mathbb{E}}_{(\varepsilon',t')\ \mathrm{design}} |\langle\widetilde{\Phi}_z|\widetilde{\Phi}_y\rangle|^{2p}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^q\langle\widetilde{\Phi}_y|\widetilde{\Phi}_y\rangle^q - \sum_{y,z\in\{0,1\}^{N_B}} \mathop{\mathbb{E}}_{\mathrm{Haar}} |\langle\widetilde{\Phi}_z|\widetilde{\Phi}_y\rangle|^{2p}\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle^q\langle\widetilde{\Phi}_y|\widetilde{\Phi}_y\rangle^q\right| \\
&= \left|\mathrm{tr}\left(\left(\mathbb{1}^{\otimes(k'-2(p+q))}\otimes\sum_{y,z\in\{0,1\}^{N_B}} M_{yz}^{\otimes p}\otimes M_{yz}^{\otimes p}\otimes M_{zz}^{\otimes q}\otimes M_{yy}^{\otimes q}\right)\cdot\left(\mathop{\mathbb{E}}_{(\varepsilon',k')\,\mathrm{design}}(|\Phi\rangle\langle\Phi|)^{\otimes k'} - \mathop{\mathbb{E}}_{\mathrm{Haar}}(|\Phi\rangle\langle\Phi|)^{\otimes k'}\right)\right)\right| \\
&\leq \left\|\mathbb{1}^{\otimes(k'-2(p+q))}\otimes\sum_{x,y\in\{0,1\}^{N_B}} M_{yz}^{\otimes p}\otimes M_{yz}^{\otimes p}\otimes M_{zz}^{\otimes q}\otimes M_{yy}^{\otimes q}\right\|_\infty \left\|\mathop{\mathbb{E}}_{(\varepsilon',k')\,\mathrm{design}}(|\Phi\rangle\langle\Phi|)^{\otimes k'} - \mathop{\mathbb{E}}_{\mathrm{Haar}}(|\Phi\rangle\langle\Phi|)^{\otimes k'}\right\|_1 \\
&\leq \varepsilon'
\end{aligned} \tag{F93}$$

where we use that $\left\|\mathbb{1}^{\otimes(k'-2(p+q))}\otimes\sum_{y,z\in\{0,1\}^{N_B}} M_{yz}^{\otimes p}\otimes M_{yz}^{\otimes p}\otimes M_{zz}^{\otimes q}\otimes M_{yy}^{\otimes q}\right\|_\infty = 1$. Using Eq. (F93), for $k' \geq 2k + 2(2b-1)(k-1) = 4b(k-1)+2$, we have

$$d_A^k \mathop{\mathbb{E}}_{(\varepsilon',k')\,\mathrm{design}}\mathrm{tr}(B(\Phi)^2) \leq d_A^k \mathop{\mathbb{E}}_{\mathrm{Haar}}\mathrm{tr}(B(\Phi)^2) + \varepsilon'\cdot d_A^k\left[-1+\left(1+\left(\frac{d_B}{r}\right)^{2(k-1)}\right)^b\right]^2. \tag{F94}$$

Putting together Eqs. (F92) and (F94), we arrive at

$$\begin{aligned}
\mathop{\mathbb{E}}_{(\varepsilon',k')\,\mathrm{design}} d_A^k\left\|B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}}[A(|\Psi\rangle)]\right\|_2^2 &\leq \mathop{\mathbb{E}}_{\mathrm{Haar}} d_A^k\left\|B(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}}[A(|\Psi\rangle)]\right\|_2^2 \\
&+ \varepsilon'\left(2\,k!\left[-1+\left(1+\left(\frac{d_B}{r}\right)^{2(k-1)}\right)^b\right] + d_A^k\left[-1+\left(1+\left(\frac{d_B}{r}\right)^{2(k-1)}\right)^b\right]^2\right)
\end{aligned} \tag{F95}$$

for $k' \geq 4b(k-1)+2$ as claimed. ■

Lemma 8 (**Moment bounds on error function**). *For $k \geq 2, r = 1$, b even, and $d_A^{1/4} \geq 8k - 6$,*

$$\mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle) \leq \frac{d_B}{2^b} + 2d_B^{2b(k-1)+2} \exp\left(-\frac{1}{8e\sqrt{2}}\, d_A^{1/4}\right), \tag{F96}$$

$$\mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle)^2 \leq \frac{d_B^2}{2^{2b}} + 2d_B^{4b(k-1)+3} \exp\left(-\frac{1}{8e\sqrt{2}}\, d_A^{1/4}\right). \tag{F97}$$

In particular, if we fix an $\varepsilon > 0$ and recall that $d_A = 2^{N_A}, d_B = 2^{N_B}$, then as long as

$$b = \Omega\left(N_B + \log(1/\varepsilon)\right), \tag{F98}$$

$$N_A = \Omega\left(\log(N_B) + \log(k) + \log\log(1/\varepsilon)\right), \tag{F99}$$

we have the following moment bounds on the error function:

$$\mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle) \leq \varepsilon, \tag{F100}$$

$$\mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle)^2 \leq \varepsilon^2. \tag{F101}$$

Proof. Recall the following definitions with $r = 1$:

$$R(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \left(1 - \left(d_B\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle\right)^{2(k-1)}\right)^b. \tag{F102}$$

We note that $\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \in [0,1]$ and we define an associated function

$$\gamma(s) = s\left(1 - (d_B s)^{2(k-1)}\right)^b, \ \forall s \in [0,1]. \tag{F103}$$

For $s \in [(1 - 1/d_A^{1/4})/d_B, (1 + 1/d_A^{1/4})/d_B]$, we can see that $\gamma(s) \geq 0$ because b is even. We now proceed to upper bound $\gamma(s)$ in this domain. We have the following bound based on the condition that $s \leq (1 + 1/d_A^{1/4})/d_B$:

$$\begin{aligned}(d_B s)^{2(k-1)} - 1 &\leq \left(1 + \frac{1}{d_A^{1/4}}\right)^{2(k-1)} \\ - 1 &\leq \frac{2(k-1)/d_A^{1/4}}{1 - (2k-3)/d_A^{1/4}} \leq \frac{4(k-1)}{d_A^{1/4}}.\end{aligned} \tag{F104}$$

The first inequality follows from the monotonicity of $(d_B s)^{2(k-1)} - 1$. The second inequality follows from the fact that $(1+x)^n \leq 1 + nx/(1 - (n-1)x)$, $\forall z \in [-1, 1/(n-1)], n > 1$. The third inequality uses the condition that $d_A \geq 8k - 6 \geq 2(2k-3)$. We can also obtain another bound using $s \geq (1 - 1/d_A^{1/4})/d_B$,

$$1 - (d_B s)^{2(k-1)} \leq \frac{2(k-1)}{d_A^{1/4}}, \tag{F105}$$

which follows from the fact that $(1+x)^n \geq 1 + nx$, $\forall x \geq -1, n \geq 1$. Together, we have $\forall s \in [(1 - 1/d_A^{1/4})/d_B, (1 + 1/d_A^{1/4})/d_B]$, that the function $\gamma(s)$ is bounded as follows:

$$|\gamma(s)| \leq \left|(d_B s)^{2(k-1)} - 1\right|^b \leq \left(\frac{4(k-1)}{d_A^{1/4}}\right)^b \leq \frac{1}{2^b}. \tag{F106}$$

The second inequality uses the fact that $d_A^{1/4} \geq 8k - 6 \geq 8(k-1)$. If the variable s is not within that domain but is in $[0,1]$, we have

$$|\gamma(s)| \leq d_B^{2b(k-1)}. \tag{F107}$$

We consider the event G such that

$$\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle \in \left[\frac{1 - 1/d_A^{1/4}}{d_B}, \frac{1 + 1/d_A^{1/4}}{d_B}\right], \ \forall z \in \{0,1\}^{N_B}, \tag{F108}$$

which is equivalent to the event that

$$\left|\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle - \frac{1}{d_B}\right| \leq \frac{1}{d_A^{1/4} d_B}, \ \forall z \in \{0,1\}^{N_B}. \tag{F109}$$

We now utilize the concentration result given in Corollary 3,

$$\begin{aligned}&\text{Prob}_{\Phi\sim\text{ Haar}(d)}\left[\left|\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle - \frac{1}{d_B}\right| \geq \delta\right] \\ &\quad\leq 2\exp\left(-\frac{1}{8e\sqrt{2}}\, d_A^{1/2}\, d_B\, \delta\right),\end{aligned} \tag{F110}$$

to derive the probability upper bound for the complement of the event G

$$\begin{aligned}&\text{Prob}_{\Phi\sim\text{ Haar}(d)}[\,G \text{ did not happen}\,] \\ &\quad\leq 2d_B \exp\left(-\frac{1}{8e\sqrt{2}}\, d_A^{1/4}\right),\end{aligned} \tag{F111}$$

which is obtained by taking the union bound. We can now proceed to upper bound the first and second moments of $R(|\Phi\rangle)$ by noting that $R(|\Phi\rangle) = \sum_{z\in\{0,1\}^{N_B}} \gamma(\langle\widetilde{\Phi}_z|\widetilde{\Phi}_z\rangle)$.

When the event G happens, we can use Eq. (F106) to obtain

$$|R(|\Phi\rangle)| \leq \frac{d_B}{2^b}. \tag{F112}$$

If the event G did not happen, then using Eq. (F107), we have

$$|R(|\Phi\rangle)| \leq d_B^{2b(k-1)+1}. \tag{F113}$$

Together, for both $m = 1, 2$, we have

$$\mathop{\mathbb{E}}_{\text{Haar}} R(|\Phi\rangle)^m \leq \left(\frac{d_B}{2^b}\right)^m \text{Prob}_{\Phi\sim\text{Haar}(d)}[\,G \text{ happened}\,] \tag{F114}$$

$$+ d_B^{m(2b(k-1)+1)} \text{Prob}_{\Phi\sim\text{Haar}(d)}[\,G \text{ did not happen}\,] \tag{F115}$$

$$\leq \frac{d_B^m}{2^{bm}} + 2d_B^{m(2b(k-1)+1)+1} \exp\left(-\frac{1}{8e\sqrt{2}}\, d_A^{1/4}\right), \tag{F116}$$

where the last inequality uses the concentration result given in Eq. (F111). The asymptotic bounds can be obtained from the above result under suitable choices of the parameters given in the statement of this lemma. ■

Lemma 9 (Second-moment bound for projected ensemble from Haar measure).

$$\mathop{\mathbb{E}}_{\Phi\sim\text{Haar}(d)}\left[\left\|A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}(d)}[A(|\Psi\rangle)]\right\|_2^2\right] \leq 36\pi^3(2k-1)(d_A^{6k}/d_B). \tag{F117}$$

In particular, recalling that $d_A = 2^{N_A}, d_B = 2^{N_B}$, then for any $\varepsilon > 0$, as long as

$$N_B = \Omega\left(kN_A + \log\left(\frac{1}{\varepsilon}\right)\right), \tag{F118}$$

we have the upper bound

$$\mathop{\mathbb{E}}_{\Phi\sim\text{Haar}(d)}\left[\left\|A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}(d)}[A(|\Psi\rangle)]\right\|_2^2\right] \leq \varepsilon^2. \tag{F119}$$

Proof. Note that Eq. (F25) implies that

$$\text{Prob}_{\Phi\sim\text{Haar}(d)}\left[\left\|A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}(d)}[A(|\Psi\rangle)]\right\|_2^2 \geq y\right] \leq 2d_A^{2k} \exp\left(-\frac{d_B y}{18\pi^3(2k-1)d_A^{4k}}\right). \tag{F120}$$

Then, we can use the fact that $\mathbb{E}[X] = \int_0^\infty dx\ \text{Prob}[X \geq x]$ for any positive random variable X to obtain

$$\mathop{\mathbb{E}}_{\Phi\sim\text{Haar}(d)}\left[\left\|A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}(d)}[A(|\Psi\rangle)]\right\|_2^2\right] \tag{F121}$$

$$= \int_{-\infty}^{\infty} dy\ \text{Prob}_{\Phi\sim\text{Haar}(d)}\left[\left\|A(|\Phi\rangle) - \mathop{\mathbb{E}}_{\Psi\sim\text{Haar}(d)}[A(|\Psi\rangle)]\right\|_2^2 \geq y\right] \tag{F122}$$

$$\leq \int_0^\infty dy\ 2d_A^{2k} \exp\left(-\frac{d_B y}{18\pi^3(2k-1)d_A^{4k}}\right) \tag{F123}$$

$$\leq 36\pi^3(2k-1)\frac{d_A^{6k}}{d_B}, \tag{F124}$$

which establishes the bound. ■

[1] E. P. Wigner, in *The Collected Works of Eugene Paul Wigner* (Springer-Verlag Berlin, Heidelberg, Germany, 1993), p. 524.

[2] F. J. Dyson, Statistical theory of the energy levels of complex systems. i, J. Math. Phys. **3**, 140 (1962).

[3] S. Popescu, A. J. Short, and A. Winter, Entanglement and the foundations of statistical mechanics, Nat. Phys. **2**, 754 (2006).

[4] P. Reimann, Foundation of Statistical Mechanics under Experimentally Realistic Conditions, Phys. Rev. Lett. **101**, 190403 (2008).

[5] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, Phys. Rev. E **79**, 061103 (2009).

[6] A. J. Short and T. C. Farrelly, Quantum equilibration in finite time, New J. Phys. **14**, 013063 (2012).

[7] M. V. Berry, Regular and irregular semiclassical wavefunctions, J. Phys. A: Math. Gen. **10**, 2083 (1977).

[8] J. M. Deutsch, Quantum statistical mechanics in a closed system, Phys. Rev. A **43**, 2046 (1991).

[9] M. Srednicki, Chaos and quantum thermalization, Phys. Rev. E **50**, 888 (1994).

[10] M. Rigol, V. Dunjko, and M. Olshanii, Thermalization and its mechanism for generic isolated quantum systems, Nature **452**, 854 (2008).

[11] R. Nandkishore and D. A. Huse, Many-body localization and thermalization in quantum statistical mechanics, Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).

[12] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics, Adv. Phys. **65**, 239 (2016).
[13] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium: Many-body localization, thermalization, and entanglement, Rev. Mod. Phys. **91**, 021001 (2019).
[14] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, Commun. Math. Phys. **346**, 397 (2016).
[15] G. Alagic, T. Gagliardoni, and C. Majenz, in *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (Springer, Tel Aviv, Israel, 2018), p. 489.
[16] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller, Statistical correlations between locally randomized measurements: A toolbox for probing entanglement in many-body quantum states, Phys. Rev. A **99**, 052323 (2019).
[17] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).
[18] A. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller, Cross-Platform Verification of Intermediate Scale Quantum Devices, Phys. Rev. Lett. **124**, 010504 (2020).
[19] J. Carrasco, A. Elben, C. Kokail, B. Kraus, and P. Zoller, Theoretical and Experimental Perspectives of Quantum Verification, PRX Quantum **2**, 010102 (2021).
[20] E. Knill, Approximation by quantum circuits, arXiv:quant-ph/9508006 (1995).
[21] B. Skinner, J. Ruhman, and A. Nahum, Measurement-Induced Phase Transitions in the Dynamics of Entanglement, Phys. Rev. X **9**, 031009 (2019).
[22] Y. Li, X. Chen, and M. P. A. Fisher, Measurement-driven entanglement transition in hybrid quantum circuits, Phys. Rev. B **100**, 134306 (2019).
[23] M. J. Gullans and D. A. Huse, Dynamical Purification Phase Transition Induced by Quantum Measurements, Phys. Rev. X **10**, 041020 (2020).
[24] C.-M. Jian, Y.-Z. You, R. Vasseur, and A. W. W. Ludwig, Measurement-induced criticality in random quantum circuits, Phys. Rev. B **101**, 104302 (2020).
[25] J. Choi, A. L. Shaw, I. S. Madjarov, X. Xie, J. P. Covey, J. S. Cotler, D. K. Mark, H.-Y. Huang, A. Kale, H. Pichler, F. G. S. L. Brandão, S. Choi, and M. Endres, Emergent randomness and benchmarking from many-body quantum chaos, arXiv:2103.03535 (2021).
[26] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Symmetric informationally complete quantum measurements, J. Math. Phys. **45**, 2171 (2004).
[27] A. Ambainis and J. Emerson, in *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)* (IEEE, San Diego, CA, USA, 2007), p. 129.
[28] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Quantum thermalization through entanglement in an isolated many-body system, Science **353**, 794 (2016).
[29] M. Ueda, Quantum equilibration, thermalization and prethermalization in ultracold atoms, Nat. Rev. Phys. **2**, 669 (2020).
[30] K. Kaneko, E. Iyoda, and T. Sagawa, Characterizing complexity of many-body quantum dynamics by higher-order eigenstate thermalization, Phys. Rev. A **101**, 042126 (2020).
[31] J. Cotler, N. Hunter-Jones, and D. Ranard, Fluctuations of subsystem entropies at late times, arXiv:2010.11922 (2020).
[32] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Exact and approximate unitary 2-designs and their application to fidelity estimation, Phys. Rev. A **80**, 012304 (2009).
[33] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, Validating quantum computers using randomized model circuits, Phys. Rev. A **100**, 032328 (2019).
[34] H.-Y. Hu and Y.-Z. You, Hamiltonian-driven shadow tomography of quantum states, arXiv:2102.10132 (2021).
[35] A. W. Harrow, The church of the symmetric subspace, arXiv:1308.6595 (2013).
[36] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, Commun. Math. Phys. **291**, 257 (2009).
[37] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, Efficient Quantum Pseudorandomness with Nearly Time-Independent Hamiltonian Dynamics, Phys. Rev. X **7**, 021006 (2017).
[38] A. Harrow and S. Mehraban, Approximate unitary *t*-designs by short random quantum circuits using nearest-neighbor and long-range gates, arXiv:1809.06957 (2018).
[39] J. Haferkamp, F. Montealegre-Mora, M. Heinrich, J. Eisert, D. Gross, and I. Roth, Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-Clifford gates, arXiv:2002.09524 (2020).
[40] T. Farshi, D. Toniolo, C. E. González-Guillén, Á. M. Alhambra, and L. Masanes, Mixing and localization in random time-periodic quantum circuits of Clifford unitaries, J. Math. Phys. **63**, 032201 (2022).
[41] P. Delsarte, J.-M. Goethals, and J. Jacob Seidel, in *Geometry and Combinatorics* (Academic Press, Cambridge, MA, USA, 1991), p. 68.
[42] F. Verstraete, M. Popp, and J. Ignacio Cirac, Entanglement versus Correlations in Spin Systems, Phys. Rev. Lett. **92**, 027901 (2004).
[43] M. Popp, F. Verstraete, M. A. Martín-Delgado, and J. Ignacio Cirac, Localizable entanglement, Phys. Rev. A **71**, 042306 (2005).
[44] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, On the distribution of the wave function for systems in thermal equilibrium, J. Stat. Phys. **125**, 1193 (2006).
[45] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Universal probability distribution for the wave function of a quantum system entangled with its environment, Commun. Math. Phys. **342**, 965 (2016).
[46] P. S. Turner and D. Markham, Derandomizing Quantum Circuits with Measurement-Based Unitary Designs, Phys. Rev. Lett. **116**, 200501 (2016).
[47] R. Mezher, J. Ghalbouni, J. Dgheim, and D. Markham, Efficient quantum pseudorandomness with simple graph states, Phys. Rev. A **97**, 022333 (2018).
[48] R. Raussendorf and H. J. Briegel, A One-Way Quantum Computer, Phys. Rev. Lett. **86**, 5188 (2001).

[49] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. Nest, and H.-J. Briegel, Entanglement in graph states and its applications, arXiv:quant-ph/0602096 (2006).
[50] S. G. Bobkov, F. Götze, and H. Sambale, Higher order concentration of measure, Commun. Contemp. Mathematics **21**, 1850043 (2019).
[51] A. M. Childs, R. Kothari, and R. D. Somma, Quantum algorithm for systems of linear equations with exponentially improved dependence on precision, SIAM J. Comput. **46**, 1920 (2017).
[52] F. G. S. L. Brandão, W. Chemissany, N. Hunter-Jones, R. Kueng, and J. Preskill, Models of quantum complexity growth, arXiv:1912.04297 (2019).
[53] H. Kim, T. N. Ikeda, and D. A. Huse, Testing whether all eigenstates obey the eigenstate thermalization hypothesis, Phys. Rev. E **90**, 052105 (2014).
[54] We note that our measurement z basis is orthogonal to the Hamiltonian; this choice is made to ensure that the measurement outcomes are not explicitly correlated with the total energy in subsystem A.
[55] J. Lux, J. Müller, A. Mitra, and A. Rosch, Hydrodynamic long-time tails after a quantum quench, Phys. Rev. A **89**, 053608 (2014).
[56] L. Vidmar and M. Rigol, Generalized Gibbs ensemble in integrable lattice models, J. Stat. Mech.: Theory Exp. **2016**, 064007 (2016).
[57] D. A. Roberts and B. Yoshida, Chaos and complexity by design, J. High Energy Phys. **2017**, 121 (2017).
[58] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, Chaos, complexity, and random matrices, J. High Energy Phys. **2017**, 48 (2017).
[59] Y.-Z. You and Y. Gu, Entanglement features of random Hamiltonian dynamics, Phys. Rev. B **98**, 014309 (2018).
[60] R. Jozsa, D. Robb, and W. K. Wootters, Lower bound for accessible information in quantum mechanics, Phys. Rev. A **49**, 668 (1994).
[61] J. Cotler and N. Hunter-Jones, Spectral decoupling in many-body quantum chaos, J. High Energy Phys. **2020**, 1 (2020).
[62] A. R. Brown and L. Susskind, Second law of quantum complexity, Phys. Rev. D **97**, 086015 (2018).
[63] S. Aaronson and A. Arkhipov, in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, NY, USA, 2011), p. 333.
[64] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert, Architectures for Quantum Simulation Showing a Quantum Speedup, Phys. Rev. X **8**, 021010 (2018).
[65] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, On the complexity and verification of quantum random circuit sampling, Nat. Phys. **15**, 159 (2019).

Correction: The source information for Ref. [40] was not updated during the proof production cycle and has been fixed.