Nature Is Stingy: Universality of Scrooge Ensembles in Quantum Many-Body Systems

Wai-Keong Mok,[1] Tobias Haug,[2] Wen Wei Ho,[3,4] and John Preskill[1,5]
[1]*Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*
[2]*Quantum Research Centre, Technology Innovation Institute, Abu Dhabi, UAE*
[3]*Department of Physics, National University of Singapore, Singapore 117551*
[4]*Centre for Quantum Technologies, National University of Singapore, Singapore 117543*
[5]*AWS Center for Quantum Computing, Pasadena, California 91125, USA*

(Received 2 March 2026; revised 22 July 2026; accepted 11 August 2026; published 2 October 2026)

Recent advances in quantum simulators allow direct experimental access to ensembles of pure states generated by measuring part of an isolated quantum many-body system. These projected ensembles encode fine-grained information beyond thermal expectation values and provide a new window into quantum thermalization. In chaotic dynamics, projected ensembles exhibit universal statistics governed by maximum-entropy principles, known as deep thermalization. At infinite temperature, this universality is characterized by Haar-random ensembles. More generally, physical constraints such as finite temperature or conservation laws lead to Scrooge ensembles, which are maximally entropic distributions of pure states consistent with these constraints. Here, we introduce Scrooge k-designs, which approximate Scrooge ensembles, and we use this framework to sharpen the conditions under which Scrooge-like behavior emerges. We first show that global Scrooge designs arise from long-time chaotic unitary dynamics alone, without measurements. Second, we show that measuring a complementary subsystem of a scrambled global state drawn from a global Scrooge $2k$-design induces a local Scrooge k-design. Third, we show that a local Scrooge k-design arises from an arbitrary entangled state when the complementary system is measured in a scrambled basis induced by a unitary drawn from a Haar $2k$-design. These results show that the resources required to generate approximate Scrooge ensembles scale only with the desired degree of approximation, enabling efficient implementations. Complementing our analytical results, numerical simulations identify coherence, entanglement, nonstabilizerness, and information scrambling as essential ingredients for the emergence of local Scrooge-like behavior. Together, our findings advance theoretical explanations for maximally entropic, information-stingy randomness in quantum many-body systems.

DOI: 10.1103/tb52-jxmx

Subject Areas: Condensed Matter Physics, Quantum Information, Statistical Physics

I. INTRODUCTION

Understanding universal behavior in complex systems is a central goal of physics. In closed quantum many-body systems, generic unitary dynamics is expected to drive local subsystems toward equilibrium, so local observables are well described by generalized Gibbs ensembles determined by conserved quantities [1,2]. Explaining how irreversibility emerges from unitarity has led to powerful concepts, such as the eigenstate thermalization hypothesis [3–5] and the maximum-entropy principle of statistical mechanics [6–15].

Recently developed quantum simulators provide access to fine-grained information going beyond thermal expectation values [16–21]. In particular, one can experimentally probe the features of projected ensembles [17,22]: Given a many-body pure quantum state on the composite system AB, one measures the complementary system B in a fixed local basis, hence sampling from an ensemble of conditional pure states on A. Averaging over these pure states, weighted by the corresponding measurement outcome probabilities, yields the marginal density operator σ_A. However, the structure of the projected ensemble encodes additional properties of the composite state that are not captured by σ_A alone.

In chaotic quantum systems with no relevant conservation laws (or at very high temperatures), analytical, numerical, and experimental studies (including some rigorous mathematical proofs) have established that the projected ensemble on A becomes indistinguishable from

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

the Haar distribution at late times [17,21–45]. The Haar ensemble is a collection of states distributed uniformly at random over Hilbert space. Thus, not only is σ_A maximally mixed, but the projected ensemble is maximally random as well.

In the presence of constraints such as energy or charge conservation, the situation is much richer, as uniform Haar randomness is no longer attainable. Instead, pioneering work by Goldstein *et al.* [43,44], based on canonical typicality [46,47], argued that the resulting limiting projected ensemble is an analog of the Haar distribution, distorted to account for the constraints, called the Scrooge ensemble [48]. Among all ensembles realizing a fixed density operator σ_A, the Scrooge ensemble is the most stingy in that it minimizes the classical information that can be accessed by measuring A [49,50]. Recent work by Mark *et al.* [31] further argued that Scrooge ensembles emerge universally in broad physical settings, such as in late-time chaotic quantum dynamics, and that their emergence adheres to a generalized maximum entropy principle.

The appearance of such maximally random projected ensembles has been observed across a host of physically distinct systems like quantum spin systems [22,31,33,40], fermionic systems [30], and bosonic continuous-variable systems [28], constituting a strong form of universality that has come to be known as deep thermalization, a burgeoning and active area of research [17,21–45].

Given the apparent ubiquity of Scrooge behavior in nature, a compelling question is not simply whether such behavior can arise, in principle, but how robustly it arises and what resources are required, in practice, for the emergence of ensembles that accurately approximate Scrooge ensembles in low-complexity, experimentally accessible regimes. Existing works, however, rely on idealized assumptions, including the preparation of highly complex many-body states or measurement bases [44], or chaotic evolution over infinitely long times [31]. These assumptions limit their applicability to realistic experiments, leaving open the question of whether Scrooge behavior can emerge robustly with finite resources.

In this paper, we expand on these insightful earlier contributions by developing a unified and rigorous framework based on the notion of approximate Scrooge k-designs. State k-designs, much studied in quantum information theory [51,52], are distributions on pure states whose first k moments match those of the Haar distribution; hence, a k-design cannot be distinguished from the Haar ensemble by any measurement acting on at most k copies of the state. Likewise, a Scrooge k-design matches the first k moments of the Scrooge ensemble, and an approximate Scrooge k-design is close to an exact one in an appropriate sense. This concept provides a quantitative tool for studying the emergence of Scrooge behavior for systems that have finite size and that evolve for a finite time. Within this framework, we prove three main theorems that provide sufficient conditions for the emergence of Scrooge behavior in several different physical settings. These results are summarized in Fig. 1.

Our first theorem concerns the temporal ensemble, which describes the states reached when a closed quantum system governed by a specified Hamiltonian evolves for a random amount of time. Assuming the absence of spectral resonances, Mark *et al.* [31] showed that averaging over very long evolution times is equivalent to averaging over the random phase ensemble [53–55], consisting of superpositions of energy eigenstates with random phases, a phenomenon called Hilbert-space ergodicity [56–60]. Moreover, Ref. [31] derived limiting expressions for the statistical moments of this ensemble and showed that they assume a form similar to, though not exactly equal to, those of the Scrooge ensemble. Our Theorem 1 rigorously formalizes this closeness in the language of Scrooge designs, with corresponding quantum information-theoretic error bounds:

Theorem 1 (Global Scrooge from chaotic dynamics, informal). If the temporal ensemble generated by a generic chaotic quantum system has low purity ($k^2\|\sigma\|_2 \ll 1$, where σ is the density operator realized by the ensemble), then it forms an approximate Scrooge k-design. This result establishes that Scrooge-like universality can already arise naturally from unitary quantum dynamics, beyond the context of projected ensembles.

We next consider projected ensembles in the context of deep thermalization. Goldstein *et al.* [44] studied the projected ensemble in a bipartite system AB in two idealized settings. First, they considered global states drawn from an *exact* Scrooge ensemble and showed that, when system B is measured in a fixed basis, the projected ensemble on A converges weakly to the Scrooge ensemble in the thermodynamic limit $|B| \to \infty$. Second, they demonstrated that a Scrooge ensemble on A emerges from an arbitrary bipartite entangled state on AB if a unitary drawn from the *exact* Haar ensemble is applied to B prior to measuring.

In practice, however, exact Scrooge ensembles do not arise in finite systems or after finite-time evolution, nor is exact Haar randomness available in the laboratory. Furthermore, experimentally validating an exact Scrooge ensemble would require collective measurements on arbitrarily many copies of the system. Our framework of Scrooge k-designs allows us to extend the settings considered by Goldstein *et al.* to an experimentally relevant regime. Adopting modern quantum information-theoretic notions of convergence, we build a quantitative theory of the resources needed for deep thermalization that applies under realistic conditions. Our results are captured by the following theorems.

Theorem 2 (Local Scrooge design from global Scrooge design, informal). Consider a single global state drawn from *any* approximate Scrooge $2k$-design with small

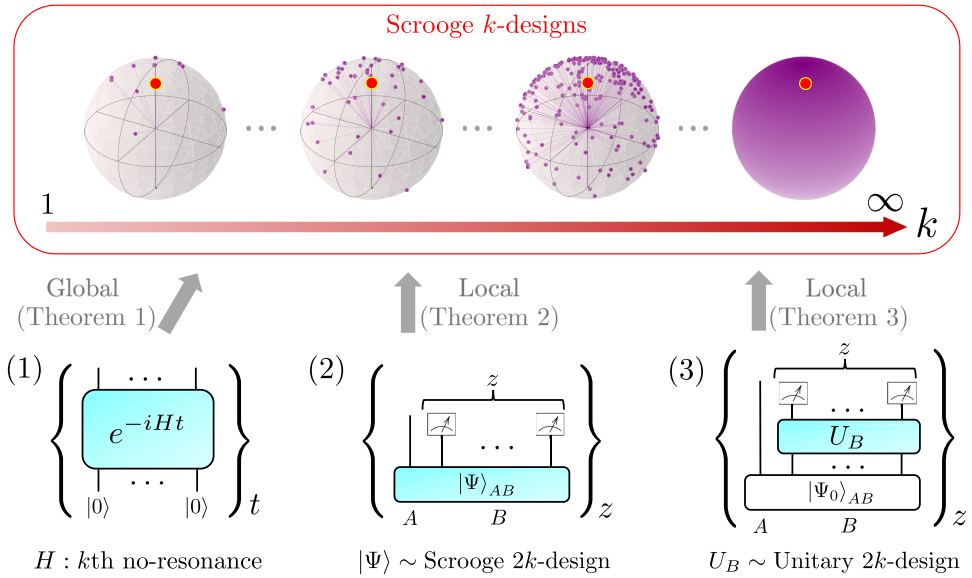

FIG. 1. Scrooge k-designs in temporal and projected ensembles. Top panel: Scrooge ensembles are distributions of random pure states that reveal minimal classical information, amongst all ensembles that realize a fixed density operator σ. In this work, we introduce Scrooge k-designs, which approximate Scrooge ensembles up to the kth moment. For the illustrative case of a single qubit shown here, it is depicted by purple points over the Bloch sphere; in all cases illustrated, the density operator σ is common and represented by a red circle. Higher Scrooge k-designs usually require more states and complexity in the ensemble, converging to the continuous Scrooge ensemble in the $k \to \infty$ limit. Bottom panel: We rigorously show that Scrooge k-designs emerge in the following physical settings, commonly considered in the literature in the context of Hilbert-space ergodicity and deep thermalization: (1) temporal ensemble $\{e^{-iHt}|0\rangle\}_{t>0}$ obtained by evolving an initial reference state $|0\rangle$ with a Hamiltonian H satisfying the kth no-resonance condition, for late times t; (2) projected ensemble generated by a state $|\Psi\rangle_{AB}$ drawn from a global Scrooge $2k$-design, with subsystem B measured in an arbitrary fixed basis; (3) projected ensemble generated by an arbitrary bipartite entangled state $|\Psi_0\rangle_{AB}$, with subsystem B scrambled by a unitary U_B drawn from a Haar $2k$-design, prior to measurements in an arbitrary fixed basis. We rigorously prove that the temporal ensemble (1) forms a global Scrooge k-design (Theorem 1), the projected ensemble in (2) forms a probabilistic mixture of local Scrooge k-designs (Theorem 2), and the projected ensemble in (3) forms a local Scrooge k-design (Theorem 3).

relative error. When B is measured in a fixed basis, the resulting projected ensemble on A is a probabilistic mixture of approximate Scrooge k-designs with high probability. Because the design order k may be regarded as a proxy for quantum complexity [61,62], this result indicates that Scrooge-like behavior of the projected ensemble is expected when the global state is sufficiently scrambled [62]. In the special case where the global ensemble approximates the Haar distribution, Theorem 2 shows that a global AB state drawn from an approximate Haar $2k$-design yields a projected ensemble on A that is an approximate Haar k-design, strengthening the results in Refs. [22,63]. Theorem 2 reduces to the results of Goldstein *et al.* [44] in the limit $k \to \infty$, where the complexity of the global state may become exponential in the size of AB.

Theorem 3 (Local Scrooge design from scrambled measurements, informal). Consider an *arbitrary* entangled state of a bipartite quantum system AB and suppose that a unitary transformation drawn from an approximate unitary $2k$-design is applied to B prior to measuring in a fixed basis. Then, the resulting projected ensemble on A forms an approximate Scrooge k-design with high probability. Complementing Theorem 2, Theorem 3 shifts the focus from the randomness of the global state to the randomness of the measurement basis. Combined with recent efficient constructions of unitary designs [61,64], this result provides a practical scheme for generating Scrooge designs using quantum circuits whose depth scales logarithmically with system size. Theorem 3 reduces to the result of Goldstein *et al.* [44] in the limit $k \to \infty$, where the quantum complexity of the scrambling unitary becomes exponential in the size of B.

Our three theorems delineate, in a unified and systematic way, several physically distinct many-body settings in which hierarchies of Scrooge behavior arise. In particular, Theorems 1 and 2 together suggest that projected ensembles formed from late-time temporal states generically exhibit Scrooge-like behavior. This physically natural conclusion was anticipated by Mark *et al.* [31] through an analysis of the distribution of unnormalized postmeasurement states but had not previously been established at the level of normalized projected ensembles. Using the techniques developed in this work, we rigorously confirm this expectation, though, strictly speaking, not directly from the theorems.

Finally, complementing our analytical theorems, we further clarify the physical resources needed for Scrooge behavior to emerge in projected ensembles. We argue that

coherence, entanglement, nonstabilizerness (magic), and information scrambling are all necessary ingredients: Removing any one of them can obstruct Scrooge-like universality. Conversely, when these resources are present, local Scrooge behavior can arise even in systems that do not thermalize dynamically. Extensive numerical simulations in several quantum many-body settings, including commuting circuits, doped Clifford circuits, and Hamiltonian ground states, are performed to illustrate these points. In particular, we demonstrate that ground states of one-dimensional integrable Hamiltonians—whose area laws and symmetries preclude thermalization—can nevertheless generate emergent Scrooge designs when measured in a random stabilizer basis. These results indicate that Scrooge universality extends beyond conventional dynamical settings such as deep thermalization or Hilbert-space ergodicity.

Together, our results provide a comprehensive and rigorous theoretical framework to analyze the emergence of maximally entropic, information-stingy randomness in quantum many-body systems—formulated in terms of Scrooge k-designs—and strongly sharpen and generalize previously known results in the literature. Beyond their conceptual significance, Scrooge designs generalize the notion of quantum randomness to physically constrained regimes. Since Haar k-designs underlie applications such as randomized benchmarking and classical shadow tomography [17,35,51,65–67], our results, by replacing Haar randomness with Scrooge designs, extend comparable performance guarantees to realistic quantum simulators operating at finite temperature or under symmetry and entanglement constraints.

This paper is organized as follows. In Sec. II, we introduce the basic notions of quantum state ensembles and settings for their emergence in natural quantum many-body systems; we also describe the key ensemble of interest —the Scrooge ensemble—studied in this work. In Sec. III, we present our main theorems, which identify general scenarios under which the Scrooge ensemble emerges. In Sec. IV, we elucidate the key physical ingredients underlying emergent Scrooge behavior in projected ensembles and discuss the physical insights provided by our theoretical results. In Sec. V, we report numerical investigations of emergent Scrooge behavior in a variety of quantum many-body systems. Finally, we conclude in Sec. VI and provide an outlook on future directions.

II. PRELIMINARIES OF QUANTUM STATE ENSEMBLES: PROJECTED AND TEMPORAL ENSEMBLES, SCROOGE ENSEMBLES, AND SCROOGE DESIGNS

We begin with a general but brief introduction to ensembles of pure quantum states and explain how to characterize them statistically and through their information-theoretic properties. We then review two physical settings where such ensembles naturally appear. The first setting is deep thermalization, which concerns projected ensembles of local postmeasurement states. The second setting is Hilbert-space ergodicity, which concerns temporal ensembles of global states generated by quantum chaotic dynamics. We then elaborate upon the Scrooge ensemble, the ensemble of interest in our paper. The informed reader may skip to our main results in Sec. III.

A. Quantum state ensembles and their information content

In this work, we are interested in ensembles of pure quantum states,

$$\mathcal{E} = \{p_i, |\psi_i\rangle\}_i. \tag{1}$$

Here, $|\psi_i\rangle$ is a quantum state supported on a D-dimensional Hilbert space $\mathcal{H}$, and p_i is that state's associated *a priori* probability. For simplicity, we assume that i is a discrete label; more generally, it can be a continuous label, in which case p_i is replaced by a probability measure. Such ensembles could arise in a multitude of physical contexts, for example, in quantum communication wherein classical information is encoded in quantum messages, or in the quantum many-body dynamical phenomena of deep thermalization and Hilbert-space ergodicity (reviewed below).

For now, we keep our discussions purely formal and explain common tools used to characterize quantum state ensembles. A standard diagnostic is to study the statistical moments of the distribution that an ensemble $\mathcal{E}$ encodes over the Hilbert space $\mathcal{H}$. To wit, for moment $k \in \mathbb{N}$, such information is encoded by the moment operator

$$\rho_{\mathcal{E}}^{(k)} := \mathop{\mathbb{E}}_{\psi\sim\mathcal{E}} (|\psi\rangle\langle\psi|)^{\otimes k} = \sum_i p_i (|\psi_i\rangle\langle\psi_i|)^{\otimes k} \tag{2}$$

supported on the symmetric subspace of $\mathcal{H}^{\otimes k}$ (note that $\rho_{\mathcal{E}}^{(k)}$ constitutes a valid density matrix on this space). We can compare the statistical similarity of one ensemble $\mathcal{E}$ with another ensemble $\mathcal{E}'$ at each moment k via the trace distance

$$\Delta^{(k)} := \frac{1}{2} \left\| \rho_{\mathcal{E}}^{(k)} - \rho_{\mathcal{E}'}^{(k)} \right\|_1, \tag{3}$$

where $\|\cdot\|_p$ is the Schatten p-norm (thus $p = 1$ above) and $\rho_{\mathcal{E}'}^{(k)}$ is the kth-moment operator of $\mathcal{E}'$, constructed analogously to Eq. (2).

A particularly useful and general lens with which one can understand the distribution that a quantum state ensemble describes in Hilbert space is through their information content, captured by the so-called accessible information $\mathcal{I}_{\text{acc}}(\mathcal{E})$ utilized in quantum information theory [68,69]. Given an ensemble $\mathcal{E}$, the accessible information quantifies the maximum amount of classical information about the label i that can be extracted by an observer

through an optimal measurement on a single copy of a state drawn from the ensemble:

$$\mathcal{I}_{\mathrm{acc}}(\mathcal{E}) \coloneqq \sup_{M \in \mathrm{POVM}} I(\mathcal{E}:M). \tag{4}$$

Here, M is a positive operator-valued measure (POVM), and $I(\mathcal{E}:M)$ is the classical mutual information between the probability distribution $\{p_i\}$ and the distribution of measurement outcomes; see Appendix A for details. For an ensemble $\mathcal{E}$ with average state $\sigma = \sum_i p_i |\psi_i\rangle\langle\psi_i| \equiv \rho_{\mathcal{E}}^{(1)}$, i.e., the density matrix, there are bounds on how small and large accessible information can be:

$$Q(\sigma) \leq \mathcal{I}_{\mathrm{acc}}(\mathcal{E}) \leq S(\sigma), \tag{5}$$

where $Q(\sigma) = -\sum_j [\lambda_j \ln \lambda_j \prod_{k\neq j} \lambda_j/(\lambda_j - \lambda_k)]$ is the so-called subentropy [49] and $S(\sigma) = -\sum_j \lambda_j \log \lambda_j$ is the von Neumann entropy, where $\{\lambda_j\}_j$ are the eigenvalues of σ. The latter is famously known as the Holevo bound [70] (here written for pure states).

The accessible information $\mathcal{I}_{\mathrm{acc}}(\mathcal{E})$ quantifies the "ergodicity" of the ensemble $\mathcal{E}$, that is, how uniformly the ensemble fills the Hilbert space, where a more ergodic ensemble has smaller accessible information and a less ergodic ensemble has larger accessible information. Indeed, the ensemble that maximizes the accessible information of a given density matrix σ is an ensemble of mutually orthogonal states, which are perfectly distinguishable by the optimal measurement. This ensemble is clearly far from ergodic—it is a discrete collection of well-separated states in Hilbert space.

The ensemble that minimizes the accessible information for a given σ is less trivial and more interesting. It is the so-called Scrooge ensemble, denoted Scrooge(σ), a continuous collection of overlapping quantum states so named because of its information-stingy property [49,50]. We give its precise definition in Sec. II C. Among all ensembles whose average state is constrained to be σ, Scrooge(σ) is the most spread out over Hilbert space, maximizing the difficulty of acquiring information about signal state $|\psi_i\rangle$. Indeed, in the special case where $\sigma = I/D$ is the maximally mixed state, Scrooge(σ) reduces to the uniformly distributed Haar ensemble such that every pure state in Hilbert space is equally likely.

B. Projected ensembles and temporal ensembles

Thus far, we have introduced quantum state ensembles as purely formal mathematical objects. Recently, two physical settings where quantum state ensembles arise have received increasing attention in the study of quantum many-body systems. One is the projected ensemble composed of local postmeasurement states, and the other is the temporal ensemble of global states generated by time evolution. Both ensembles have been found to exhibit universal features in limiting cases, embodied by the physical phenomena of "deep thermalization" and "Hilbert-space ergodicity," respectively. Here, we quickly review these topics.

1. Projected ensemble and deep thermalization

Deep thermalization concerns the ensemble of states on a local region of a quantum many-body system conditioned upon measurements of the complementary region; this is called the projected ensemble and is defined as follows [17,22].

Let $|\Psi\rangle_{AB}$ be a global "generator" D-dimensional quantum state supported on a bipartite Hilbert space AB, with subregions $A(B)$ of dimensions $D_A(D_B)$, respectively, so that $D = D_A D_B$. While not strictly necessary for this work, we will regard $|\Psi\rangle_{AB}$ as an N-qubit state with subsystems A and B comprising N_A and N_B qubits, respectively. In such a case, $D_A = 2^{N_A}$ and $D_B = 2^{N_B}$. Consider next a projective measurement performed on the "bath" subsystem B, described by a set of orthogonal rank-1 projectors $\{\Pi_z\}_{z=1}^{D_B}$ satisfying the normalization $\sum_z \Pi_z = I_B$. Measurement outcome z occurs with probability $p_z = \langle\Psi|(I_A \otimes \Pi_z)|\Psi\rangle$, upon which the associated postmeasurement state on the unmeasured subsystem A is $|\psi_z\rangle_A \propto (I_A \otimes \Pi_z)|\Psi\rangle$. The set of such (normalized) postmeasurement projected states, weighted by their probabilities, constitutes the projected ensemble

$$\mathcal{E}(\Psi) \coloneqq \{p_z, |\psi_z\rangle_A\}_{z=1}^{D_B} \tag{6}$$

[depicted in settings (2) and (3) of Fig. 1]. By capturing correlations with classical information extracted from the bath B, the projected ensemble provides a more refined description of the local subsystem A than the reduced density matrix σ_A.

The term deep thermalization, a burgeoning research topic in recent years [17,21–45], refers to the observation that, for generator states $|\Psi\rangle_{AB}$ arising from late-time dynamics of chaotic quantum many-body systems and in the thermodynamic limit $D_B \to \infty$ with D_A fixed, the projected ensemble generically approaches Scrooge(σ_A) (and related variants), which has been recognized to have maximal entropy in the sense of harboring minimal accessible information, given the reduced density matrix σ_A [31]. In particular, when there are no conservation laws constraining the dynamics such that we expect σ_A to be maximally mixed, the corresponding Scrooge ensemble is the Haar ensemble. If the dynamics conserves energy and charge such that we expect quantum thermalization to drive σ_A to a Gibbs state with a specified temperature and chemical potential [7–10], the corresponding Scrooge ensemble is then the most stingy "unraveling" of the Gibbs state in terms of constituent pure states.

2. Temporal ensemble and Hilbert-space ergodicity

Quantum state ensembles also arise naturally when we consider evolution of a global state for a random time. Suppose a D-dimensional quantum system evolves unitarily under some Hamiltonian $H(t)$, which, in general, could be time dependent, and suppose the evolution time t is sampled uniformly from the interval $[0, T]$. The resulting state ensemble

$$\mathcal{E}_{\text{Temporal}} := \{dt/T, |\Psi_t\rangle\}_{t\in[0,T]} \tag{7}$$

[depicted in setting (1) of Fig. 1] is called the temporal ensemble [31,53–55]. Here, t plays the role of the classical label i in Eq. (1). If energy is not conserved and $H(t)$ has generic time dependence, one expects the temporal ensemble to be uniformly distributed for any choice of the initial state because there is no preferred direction in Hilbert space, allowing the system to explore the entire Hilbert space without constraints. Indeed, the emergence of the Haar ensemble at late times has been demonstrated rigorously for various classes of time-dependent $H(t)$ (such as aperiodic and quasiperiodic drives consisting of sequences of kicks generated randomly or even deterministically according to the Fibonacci and Thue-Morse words), a phenomenon termed complete Hilbert-space ergodicity [56–60].

When $H(t) = H$ is time independent, energy conservation ensures that the populations $|\langle E_j|\Psi_t\rangle|^2$ in the energy basis $\{|E_j\rangle\}_j$ remain invariant in time, precluding the emergence of Haar, in general. However, this does not impose constraints on the relative phases between energy eigenstates. At late observation times T, it is natural to expect the temporal ensemble to be well described by the so-called random phase ensemble [31,53,71]

$$\mathcal{E}_{\text{random phase}} := \left\{\frac{d^D\varphi}{(2\pi)^D}, \sum_{j=1}^{D} |\langle E_j|\Psi_0\rangle| e^{i\varphi_j}|E_j\rangle\right\}, \tag{8}$$

where $|\Psi_0\rangle = \sum_{j=1}^{D}\langle E_j|\Psi_0\rangle|E_j\rangle$ is the initial state and $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_D)$ are angles uniformly distributed on the D-dimensional torus $[0, 2\pi)^D$. Indeed, the density matrix associated with the temporal ensemble is the diagonal ensemble [3,4]

$$\sigma_{\text{diag}} = \sum_{j=1}^{D} |\langle E_j|\Psi_0\rangle|^2 |E_j\rangle\langle E_j|, \tag{9}$$

a well-known equilibrium state, defined by the initial state dephased in the energy basis [46]. The random phase ensemble may be thought of as the most ergodic ensemble subject to the constraint of conservation of populations on energy eigenstates (in the sense that the relative phases $e^{i\varphi_j}$, which are the remaining degrees of freedom, are distributed with no preferred location on the torus); thus, its emergence can be understood as Hilbert-space ergodicity in the case of energy conservation.

Technically, Mark *et al.* proved that the kth moments of the random phase ensemble describe the limiting form of the temporal ensemble if the energy spectrum of H satisfies the assumption of a so-called "kth no-resonance condition" [31]: Any two subsets of energy levels $\{E_1, \ldots, E_k\}$ and $\{E'_1, \ldots, E'_k\}$ satisfy $\sum_{i=1}^{k} E_i = \sum_{i=1}^{k} E'_i$ if and only if the subsets are equivalent, up to a reordering of energy levels. This assumption describes the scenario where energy levels are generic, a physical property that is intuitively expected to hold true for quantum chaotic systems, whose energy spectra are known to be well governed by random matrix theory [6]. In particular, for $k = 1$, this condition is equivalent to having no degeneracies in the energy spectrum, a mild condition that holds generically for both interacting integrable and quantum chaotic systems alike.

C. Scrooge ensemble and Scrooge designs

We have seen how quantum state ensembles arise in many physical settings, and in certain appropriate limits (large system sizes or late times), they appear to take various simple universal limiting forms, like the Haar ensemble, the Scrooge ensemble, and the random phase ensemble. As mentioned in the Introduction, our key interest in this paper is to solidify general but precise conditions when the information-stingy Scrooge ensemble can be *rigorously* proven to appear, including but going even beyond the setting of deep thermalization, and in the context of Scrooge designs. To that end, we now formally define the Scrooge ensemble and Scrooge designs.

Definition 1 (Scrooge ensemble [49]). The Scrooge ensemble with density matrix σ supported on a D-dimensional Hilbert space, denoted Scrooge(σ), is the unique ensemble $\mathcal{E}$ of pure states satisfying

$$\mathcal{E} = \arg\min_{\mathcal{E}'} \mathcal{I}_{\text{acc}}(\mathcal{E}'), \quad \text{subject to} \quad \mathop{\mathbb{E}}_{\psi\sim\mathcal{E}'} |\psi\rangle\langle\psi| = \sigma. \tag{10}$$

An explicit construction of Scrooge(σ) is given by

$$\text{Scrooge}(\sigma) = \left\{D\langle\phi|\sigma|\phi\rangle \mathrm{d}\phi, \frac{\sqrt{\sigma}|\phi\rangle}{\|\sqrt{\sigma}|\phi\rangle\|}\right\}, \tag{11}$$

with $\mathrm{d}\phi$ the Haar measure on the unit sphere in $\mathbb{C}^D$ [49]. From the expression, one sees that the Scrooge ensemble can be thought of as a deformed version of the Haar ensemble, to account for the constraint that its mean is σ.

We are also interested in ensembles that *approximate* the Scrooge ensemble. Analogously to state k-designs routinely used in quantum information theory [51,52], which provide low-complexity approximations that capture the statistical properties of the Haar ensemble up to the kth

moment, we likewise introduce the notion of k-designs for the Scrooge ensemble:

Definition 2 (Approximate Scrooge k-designs). Let $\mathcal{E}$ be an ensemble of pure states, with the kth moment

$$\rho_{\mathcal{E}}^{(k)} = \mathop{\mathbb{E}}_{\psi\sim\mathcal{E}}(|\psi\rangle\langle\psi|)^{\otimes k}. \tag{12}$$

We say $\mathcal{E}$ is a Scrooge(σ) k-design with additive error ε if

$$\frac{1}{2}\left\|\rho_{\mathcal{E}}^{(k)} - \rho_{\text{Scrooge}}^{(k)}(\sigma)\right\|_1 \leq \varepsilon, \tag{13}$$

and a Scrooge(σ) k-design with relative error ε if [72]

$$(1-\varepsilon)\rho_{\text{Scrooge}}^{(k)}(\sigma) \preccurlyeq \rho_{\mathcal{E}}^{(k)} \preccurlyeq (1+\varepsilon)\rho_{\text{Scrooge}}^{(k)}(\sigma) \tag{14}$$

in Loewner order, where $A \preccurlyeq B$ if $B - A$ is positive semidefinite. The kth moment of Scrooge(σ) is given by

$$\rho_{\text{Scrooge}}^{(k)}(\sigma) = D \mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)}\left[\frac{(\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma})^{\otimes k}}{\langle\phi|\sigma|\phi\rangle^{k-1}}\right]. \tag{15}$$

We note that evaluating the kth moment $\rho_{\text{Scrooge}}^{(k)}(\sigma)$ in Eq. (15) involves integrating a rational function over the Haar ensemble and is thus not amenable to simple closed-form expressions using standard Weingarten calculus [73–75]. While Mark *et al.* derived an explicit expression [31], its form is complicated and unwieldy. Therefore, they also considered the closely related *unnormalized* Scrooge ensemble $\tilde{\mathcal{E}}$ composed of unnormalized states $\tilde{\mathcal{E}} = \{\sqrt{D\sigma}|\phi\rangle\}$, where $|\phi\rangle \sim \text{Haar}(D)$, with σ the D-dimensional density matrix defining the Scrooge ensemble in question. For this ensemble, the kth moment takes a simple and analytically tractable form: $\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma) = (D\sigma)^{\otimes k}\rho_{\text{Haar}}^{(k)}$, to which many of their results pertain. In our work, the unnormalized Scrooge also plays a useful albeit intermediary role, as captured in the following technical lemma we derive and whose proof we present in Appendix B.

Lemma 1 (Scrooge approximation). Consider the ensemble of unnormalized states $\tilde{\mathcal{E}} = \{\sqrt{D\sigma}|\phi\rangle\}$, where $|\phi\rangle \sim \text{Haar}(D)$, and σ is an arbitrary density matrix with dimension D. For $k^2\|\sigma\|_2 \ll 1$, $\tilde{\mathcal{E}}$ forms a Scrooge(σ) k-design with additive error $\mathcal{O}(k\|\sigma\|_2)$ and relative error $\mathcal{O}(4^k k\|\sigma\|_2)$.

Lemma 1 is the technical backbone of this paper: The kth moment of the *normalized* Scrooge ensemble (the actual object of interest) can be approximated by the simpler expression up to an error controlled by the purity of σ (see Ref. [62] for an analogous relative-error bound, controlled by $\|\sigma\|_\infty$ instead of $\|\sigma\|_2$) [76]. This result generalizes that of Ref. [31] for $k = 2$ and will be key in allowing us to make rigorous statements on the limiting form of the temporal and projected ensembles by analyzing their unnormalized counterparts.

III. EMERGENT SCROOGE DESIGNS IN QUANTUM MANY-BODY SYSTEMS

We are now ready to state our main results, which provide sufficient conditions for the rigorous emergence of Scrooge designs in quantum many-body systems.

A. Natural emergence of Scrooge designs in dynamics

Our first result pertains to the temporal ensemble for energy-conserving Hamiltonian dynamics, which, as explained in Sec. II B 2, generically matches the random phase ensemble (8) in the limit of long observation time. Reference [31] found that the kth moment of the random phase ensemble (which equals the temporal ensemble under the assumption of the kth no-resonance condition on the spectrum) is well approximated by that of the unnormalized Scrooge ensemble, in the limit of low purity. Our Lemma 1 allows us to go further, relating it directly to the kth moment of the normalized Scrooge ensemble, resulting in Theorem 1.

Theorem 1 (Scrooge designs from late-time chaotic dynamics). The ensemble of states sampled at late times in quantum chaotic dynamics, generated by a Hamiltonian obeying the kth no-resonance condition, forms an approximate Scrooge(σ_{diag}) k-design for $k^2\|\sigma_{\text{diag}}\|_2 \ll 1$, with an additive error

$$\varepsilon = \mathcal{O}(k\|\sigma_{\text{diag}}\|_2). \tag{16}$$

The proof is presented in Appendix C. Our proof works by showing that the kth moment of the random phase ensemble is close to that of the unnormalized ensemble $\{\sqrt{D\sigma_{\text{diag}}}|\phi\rangle\}$, where $|\phi\rangle \sim \text{Haar}(D)$, with a trace distance of $\mathcal{O}(k^2\|\sigma_{\text{diag}}\|_2^2)$. Lemma 1 then guarantees that the resulting ensemble forms a Scrooge(σ_{diag}) k-design, incurring an extra additive error of $\mathcal{O}(k\|\sigma_{\text{diag}}\|_2)$.

Theorem 1 offers substantial improvements compared to those arising from the results of Ref. [31]. Extracting an error bound from the expressions derived in Ref. [31] and invoking our Lemma 1 yields the much weaker $\varepsilon = \mathcal{O}(k!\exp(k)\|\sigma_{\text{diag}}\|_2^2 + k\|\sigma_{\text{diag}}\|_2)$, which contains an additional term that grows very quickly with k, though this was not explicitly stated there [77]. Note that $\|\sigma_{\text{diag}}\|_2$ is often exponentially small in N, for typical initial product states that are supported on exponentially many energy eigenstates, which holds even for integrable many-body Hamiltonians, thereby satisfying the low-purity condition. On the other hand, the kth no-resonance condition is generically expected for interacting many-body Hamiltonians, including both integrable and chaotic systems [31]. Integrable systems that are effectively noninteracting, such as the 1D transverse-field Ising model,

may violate the no-resonance conditions due to algebraic relations between the energy levels. While beyond the scope of this paper, it would be interesting to study the robustness of Theorem 1 when the no-resonance conditions are weakly violated [55,78].

Conceptually, Theorem 1 rigorizes the expectation that Scrooge-like behavior *naturally* appears already in late-time dynamics; in contrast to the case of deep thermalization, no measurements are needed. Here, "late-time" formally refers to the long-time limit $T \to \infty$. In Ref. [62], it was argued that the temporal ensemble for a fixed time-independent Hamiltonian typically requires $T \gtrsim \exp(kN)$ to form an approximate Scrooge k-design, which can be physically thought of as the timescale required to resolve exponentially small energy gaps. Hence, not only is the late-time temporal ensemble maximally ergodic in the sense that the relative phases of energy eigenstates are uniformly distributed, in addition (for the case of low purity), its accessible information is close to minimal. Thus, similar maximum entropy principles apply to both deep thermalization and Hilbert-space ergodicity, conceptually unifying these two phenomena.

B. Emergence of Scrooge designs in the projected ensemble of a sufficiently complex initial state

Our next theorem returns to the projected ensemble. Motivated by the insight from the previous result, Theorem 1—that *global* Scrooge designs emerge in late-time chaotic dynamics of a quantum many-body system—we inquire if a typical global state drawn at late times (and, more generally, from a global Scrooge ensemble) can itself produce Scrooge behavior *locally* in the setting of deep thermalization. As mentioned in the Introduction, Goldstein *et al.* [44] considered the setting where the global generator state is drawn from an *exact* Scrooge ensemble; here, we generalize this to the case of Scrooge designs.

Theorem 2 (Emergent Scrooge k-design from Scrooge $2k$-design generator). Fix $k \in \mathbb{N}$. Let $|\Psi\rangle_{AB}$ be sampled from a Scrooge(σ) $2k$-design, with relative error ε. Denote the reduced density matrices of σ on A and B by σ_A and σ_B, respectively. Let $\mathcal{E}(\Psi)$ be the projected ensemble generated from $|\Psi\rangle_{AB}$ by measuring B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$, and denote the normalized density matrix $\hat{\sigma}_{A|z} = (I_A \otimes \langle z|)\sigma(I_A \otimes |z\rangle)/\langle z|\sigma_B|z\rangle$. Then, for $\|\hat{\sigma}_{A|z}\|_2 \ll 1$ and $1 \ll D_A \leq D_B$,

$$\mathbb{E}_\Psi \left\| \rho_{\mathcal{E}}^{(k)}(\Psi) - \sum_{z=1}^{D_B} \langle z|\sigma_B|z\rangle \rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}) \right\|_1 \leq \mathcal{O}\left(\sqrt{{D_A}^k (\varepsilon + \|\sigma\|_2 + D_A \|\sigma\|_2^2)} \right). \tag{17}$$

This result is proven in Appendix D, with Lemma 1 playing a key role in the proof. Theorem 2 implies that, with high probability, the projected ensemble for $\varepsilon + \|\sigma\|_2 \ll {D_A}^{-k}$ approximates a mixture of Scrooge($\hat{\sigma}_{A|z}$), i.e., a so-called "generalized Scrooge ensemble" [31]. Note that, throughout this paper, we will suppress error terms that vanish in the asymptotic limit $D_A, D_B \to \infty$ for simplicity. Intuitively, Theorem 2 says that we can interpret each projected state $|\psi_z\rangle$ as a realization of Scrooge($\hat{\sigma}_{A|z}$), weighted by the average measurement probability $\langle z|\sigma_B|z\rangle$. This case is nontrivial since each projected state $(|\psi_z\rangle\langle\psi_z|)^{\otimes k}$ is, in general, far from $\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})$.

While Theorem 2 states that the projected ensemble is formally a mixture of exponentially many (in N_B) distinct Scrooge ensembles, in certain physical settings, it is possible to obtain a simpler, "coarse-grained" mixture with a much smaller number of distinct Scrooge ensembles. For instance, as argued for in Ref. [33], if the generator state $|\Psi\rangle$ is obtained under U(1)-symmetric scrambling dynamics, then the 2^{N_B} possible measurement outcomes can be grouped into at most $N_B + 1$ equivalence classes. Consequently, the projected ensemble is a mixture of only a polynomial number of distinct Scrooge ensembles. In the extreme case where the measurement basis is "information nonrevealing," that is, measurements do not reveal information about the local charge [31,33] so that $\hat{\sigma}_{A|z} \approx \sigma_A$, the different Scrooge($\hat{\sigma}_{A|z}$) ensembles collapse to a single ensemble Scrooge(σ_A). This process yields $\rho_{\mathcal{E}}^{(k)}(\Psi) \approx \rho_{\text{Scrooge}}^{(k)}(\sigma_A)$, i.e., a vanilla single Scrooge ensemble, where σ_A is the reduced density matrix of $|\Psi\rangle$ on A. Understanding if a similar coarse-grained description of the generalized Scrooge ensemble exists for generator states obtained under different conservation laws, like energy, would be an interesting future direction.

In the following, we discuss various examples applying Theorem 2.

1. Example: Infinite-temperature states

A special case of Theorem 2 is when the density matrix in the definition of the Scrooge generator is maximally mixed, $\sigma = I/D$. We then immediately have that local Haar k-designs emerge from a typical global Haar $2k$-design generator state in the context of deep thermalization [79]. In fact, for this special case, we are able to improve upon the error bound stated in Theorem 2, which we present as the following corollary.

Corollary 1 (Emergent k-design from $2k$-design generator). Let $|\Psi\rangle_{AB}$ be sampled from a $2k$-design, with relative error ε. Note that $\mathcal{E}(\Psi)$ is the projected ensemble generated from $|\Psi\rangle_{AB}$ by measuring B in an arbitrary orthonormal basis. For $k^2 \ll D_A$ and $D_A, D_B \gg 1$,

$$\mathbb{E}_\Psi \|\rho_{\mathcal{E}}^{(k)}(\Psi) - \rho_{\text{Haar},A}^{(k)}\|_1 \leq \sqrt{\frac{{D_A}^k}{k!}\left(\frac{1}{D_B} + \mathcal{O}(\varepsilon)\right)}. \tag{18}$$

The proof is given in Appendix D. The main improvement in Eq. (18) compared to Eq. (17) is that the latter is valid only for any fixed k independent of system size, while k can grow exponentially large in N_A in the former.

Moreover, Corollary III B 1 is a significant improvement over existing results in the literature [22], which require $|\Psi\rangle_{AB}$ to be drawn from a k'-design with $k' = \mathcal{O}(k^2 \log D_A)$ in order to guarantee that $\mathcal{E}(\Psi)$ forms an approximate k-design on A. In contrast, we only require $k' = 2k$ independent of system size. For the simplest nontrivial case of $k = 2$, Corollary III B 1 states that drawing $|\Psi\rangle_{AB}$ from *any* 4-design (with sufficiently small ε) suffices for $\mathcal{E}(\Psi)$ to form an approximate 2-design, which is, in fact, optimal: Multiqubit stabilizer states form exact 3-designs with $\varepsilon = 0$ [80–82], but the projected ensemble generated by a stabilizer state forms only a 1-design, which is far from a 2-design [41,83]. This observation highlights the essential role of nonstabilizerness (also known as magic) in deep thermalization, complementing existing works [41,42], a point we discuss further in Sec. IV and explore numerically in Sec. V.

The $2k$-design with relative error ε used in Corollary III B 1 can be prepared by applying a global unitary drawn from a unitary $2k$-design (with relative error ε) on a fixed reference state $|0\rangle_{AB}$. Such unitaries can be constructed efficiently, for example, using local random circuits with depth $\tilde{\mathcal{O}}(k \log(N/\varepsilon))$ [64] or doped Clifford circuits with $\tilde{\mathcal{O}}(Nk + \log(1/\varepsilon))$ non-Clifford gates [84,85]. We also remark that, while we use relative error in the theorem for convenience, the emergence of state designs, as implied by Corollary III B 1, also holds true if $|\Psi\rangle$ is sampled from a $2k$-design with additive error $\varepsilon \ll k!/(D_A D_B^2)^k$ [86].

2. (Almost) example: Quantum chaotic Hamiltonian dynamics

Returning to the motivation of Theorem 2, we may ask whether a typical late-time state evolved under a chaotic global quantum many-body Hamiltonian is guaranteed to generate a local projected ensemble that is approximately Scrooge. If so, this conclusion would provide a satisfying and broadly applicable explanation for deep thermalization. It is tempting to declare that the desired result follows immediately from our theorems; Theorem 1 supports the claim that chaotic time evolution produces a global state that is sufficiently scrambled, and then the Scrooge-like behavior of the projected ensemble seems to follow from Theorem 2. Unfortunately, though, this reasoning is flawed for a subtle technical reason: Theorem 2 requires a generator state drawn from an approximate Scrooge design with small *relative* error, while Theorem 1 only guarantees the weaker *additive* error.

Fortunately, the statement "dynamically generated global states yield local Scrooge behavior" can be justified nevertheless *without* invoking Theorem 2:

Proposition 1 (Emergent Scrooge k-design from late-time chaotic dynamics, informal). Let $|\Psi_t\rangle = e^{-iHt}|\Psi_0\rangle$ be drawn from the temporal ensemble (7) defined for long times T, and suppose H is an arbitrary Hamiltonian satisfying the kth no-resonance condition. In the low-purity limit, the projected ensemble $\mathcal{E}(\Psi)$ forms a generalized Scrooge ensemble with high probability.

The formal statement and proof are provided in Appendix D. While Proposition 1 does not formally utilize Theorem 2, the former's derivation is inspired by technicalities in the proof of the latter.

We note that Ref. [31] had already anticipated this result: They proved that the unnormalized projected ensemble (i.e., the collection of local postmeasurement states before normalizing) is close to the unnormalized Scrooge ensemble $\tilde{\mathcal{E}}$, as defined in Lemma 1. Using the analytical tools developed in this work, in particular, Lemma 1, our result, Proposition 1, closes this conceptual gap and proves that Scrooge behavior emerges also in the normalized projected ensemble, the physical ensemble of interest in deep thermalization.

3. Example: Canonical thermal pure quantum ensemble

An example of a set of quantum states that form a relative-error Scrooge design, thereby satisfying the conditions of Theorem 2, is the so-called "canonical thermal pure quantum" (cTPQ) ensemble, which was introduced by the quantum thermalization and thermodynamics communities [89,90] and is closely related to the notion of canonical typicality [46,47,91]. Given a time-independent Hamiltonian H and inverse temperature β, it is defined as

$$\mathcal{E}_{\mathrm{cTPQ}} = \left\{ \frac{1}{\mathcal{N}} \sum_{j=1}^{D} \xi_j e^{-\beta H/2} |j\rangle \right\}_{\xi}. \tag{19}$$

Above, $\xi = (\xi_1, \ldots, \xi_D)$, where ξ_j are taken to be independent zero-mean complex Gaussian variables [92], $\{|j\rangle\}$ is an arbitrary orthonormal basis, and $\mathcal{N}$ is the normalization factor.

The cTPQ states constitute pure-state approximations to thermal Gibbs states (at inverse temperature β), in the sense that expectation values of local observables are reproduced up to fluctuations that vanish exponentially with system size [89,90] since their associated density matrix is, by construction, exponentially close to the thermal state $\sigma_\beta \propto \exp(-\beta H)$ [89]. In fact, such states have been demonstrated numerically to also reproduce certain physical properties of scrambled many-body states at finite effective temperatures, such as entanglement entropy, thermodynamic quantities, and phase diagrams of lattice gauge theories [89,90,93].

In Appendix C, we show that $\mathcal{E}_{\mathrm{cTPQ}}$ forms a Scrooge(σ_β) k-design with relative error $\mathcal{O}(4^k k \|\sigma_\beta\|_2)$. Theorem 2 therefore immediately implies that cTPQ states deeply

thermalize, producing local Scrooge behavior with $\sigma_A = \mathrm{Tr}_B \sigma_\beta$. Since energy eigenstates of quantum many-body chaotic systems are expected to satisfy the eigenstate thermalization hypothesis and have also been numerically shown to deep thermalize locally to Scrooge [22,31], cTPQ states offer a natural framework for modeling such behavior, inspired by random matrix theory and quantum typicality.

C. Emergence of Scrooge designs in the projected ensemble for a sufficiently complex measurement basis

Thus far, we have analyzed the cases where the generator state $|\Psi\rangle_{AB}$ is assumed to be drawn from a statistical ensemble, and the measurement basis is fixed but arbitrary. A complementary scenario occurs when we allow $|\Psi\rangle_{AB}$ to be fixed but arbitrary, and the projected ensemble is generated by applying a scrambling unitary U_B to B and then measuring in a fixed basis. This setting is also considered by Goldstein *et al.* [44], wherein they considered U_B to be drawn from the *exact* Haar distribution; here, we relax this stringent assumption and take the scrambling unitary to be drawn from an approximate Haar $2k$-design, which can be implemented efficiently with local random circuits [64]. We prove that, in this scenario, the projected ensemble is universally Scrooge.

Theorem 3 (Emergent Scrooge k-design from 2k-design measurement basis). Let $|\Psi\rangle_{AB} = (I_A \otimes U_B)|\Psi_0\rangle_{AB}$, where $|\Psi_0\rangle_{AB}$ is an arbitrary bipartite state, and suppose the scrambling unitary U_B is sampled from a unitary $2k$-design with relative error ε. Denote the reduced density matrix of $|\Psi\rangle$ on A by σ_A, with effective dimension $D_{A,\mathrm{eff}} = (\|\sigma_A\|_2/\|\sigma_A\|_4)^4$. Here, $\mathcal{E}(\Psi)$ is the projected ensemble generated by $|\Psi\rangle_{AB}$ by measuring B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. Then, for $k^4 \ll D_{A,\mathrm{eff}}$ and $1 \ll D_A \le D_B$,

$$\mathbb{E}_{U_B} \left\| \rho_{\mathcal{E}}^{(k)}(\Psi) - \rho_{\mathrm{Scrooge}}^{(k)}(\sigma_A) \right\|_1 \le \sqrt{(D_A \|\sigma_A\|_2^2)^k \mathcal{O}\left(\frac{k^2}{D_{A,\mathrm{eff}}}\right) + \frac{{D_A}^k}{k!} \mathcal{O}\left(\frac{k^{2k+2}}{D_B} + \varepsilon\right)}. \quad (20)$$

This result is proven in Appendix E, with Lemma 1 being a key step in the proof. Theorem 3 refines the expectation that Scrooge behavior emerges locally when the measurement basis on the complementary region is sufficiently entangling and complex, modeling the information scrambling arising in natural chaotic many-body systems. The condition $k^4 \ll D_{A,\mathrm{eff}}$ required in Theorem 3 is typically satisfied for complex many-body states, where the effective dimension of σ_A (roughly, how many eigenvalues contribute significantly) is approximately $D_A = 2^{N_A}$ for Haar random states when $D_A \ll D_B$ (a more physical example is discussed below). In such cases, Theorem 3 implies that the projected ensemble is, with high probability, close to a Scrooge(σ_A) k-design for a sufficiently large D_B and a sufficiently small ε. This result generalizes the theorem of Wilming and Roth [25], who showed that the projected ensemble forms a Haar k-design when σ_A is close to maximally mixed, and the unitary U_B is drawn exactly from the Haar measure.

1. Example: Many-body states at thermal equilibrium

As a physical example, suppose the global state $|\Psi\rangle$ obeys quantum thermalization, i.e., the local reduced density matrix σ_A is well approximated by the local Gibbs state $\exp(-\beta H_A)/\mathrm{Tr}[\exp(-\beta H_A)]$ on region A. Here, H_A is the restriction of the global Hamiltonian H to subsystem A, and β is the inverse temperature. The local Gibbs state is often a good approximation for the marginal of the global Gibbs state $\mathrm{Tr}_B(e^{-\beta H})/\mathrm{Tr}(e^{-\beta H})$, when $N_A \gg 1$, and naturally arises from quench dynamics or from energy eigenstates of nonintegrable Hamiltonians (which are believed to obey the eigenstate thermalization hypothesis) [94,95].

Now, we can always express H_A as a linear combination of Pauli operators

$$H_A = \sum_{m=1}^{M} c_m P_m, \quad (21)$$

where c_m are real coefficients and P_m are Pauli operators on A. Assuming that the bulk spectral density of H_A is Gaussian distributed (which holds approximately for a wide range of local Hamiltonians [96,97]), we obtain (see Appendix E)

$$\|\sigma_A\|_p \approx \frac{1}{2^{N_A(1-1/p)}} \exp\left(\frac{p-1}{2}\beta^2\Delta^2\right), \quad (22)$$

where $\Delta^2 = \sum_m c_m^2 \sim M$ is the variance of the Gaussian distribution. Thus, the condition $k^4 \ll D_{A,\mathrm{eff}}$ needed in Theorem 3 becomes $k^4 \ll 2^{N_A}\exp(-4\beta^2\Delta^2)$. For Hamiltonians with geometrically local interactions, we expect $M \propto N_A$. This implies the emergence of Scrooge designs for β smaller than some constant $\beta_c \propto \sqrt{N_A/M}$, in other words, if the temperature is sufficiently high. Crucially, this does not require σ_A to be close to infinite temperature, as in previous works [25], thereby generalizing existing results to the finite-temperature scenario.

Theorem 3 also provides a protocol to sample from thermal Scrooge designs, which are potentially useful for learning properties of thermal states [98]. In this protocol, one prepares a purification of the thermal state $\sigma_A \propto \exp(-\beta H_A)$. This can be done via quantum Gibbs sampling algorithms [99,100], imaginary time evolution [101,102], or other constructions of the thermofield double state [103,104]. Then, additional ancilla qubits are introduced. The reference system and the ancilla make up

subsystem B, which is then scrambled with a unitary drawn from a $2k$-design. Finally, each projective measurement on B results in a sample of a thermal Scrooge k-design on A.

IV. PHYSICAL INGREDIENTS BEHIND EMERGENT SCROOGE BEHAVIOR IN PROJECTED ENSEMBLES

Our theorems have identified several general conditions under which quantum state ensembles with Scrooge-like behavior, as captured within the framework of Scrooge designs, can be provably shown to emerge: namely, dynamically in the temporal ensemble (Theorem 1); and via measurements in the projected ensemble, if the generator state is itself drawn from a Scrooge design (Theorem 2) or if the measurement basis is rotated by a Haar design (Theorem 3). Importantly, these theorems demonstrate that the Scrooge ensemble appears in a myriad of quantum many-body settings. On the other hand, we do not expect our theorems to exhaustively cover all possible scenarios where Scrooge behavior can arise; moreover, they do not directly inform us as to the necessary physical ingredients that underlie its emergence.

Here, we discuss the necessary ingredients that a generator state $|\Psi\rangle_{AB}$ on a bipartite system AB has to possess in order for it to exhibit emergent Haar or Scrooge-like behavior in its projected ensemble.

Interestingly, recent works have begun to shed light on this question by quantitatively studying the effect of the amount of quantum information resources [105] in governing the degree of universal randomness produced [36,41,42,45]. Recent work has shown that, for certain types of resources known as "threshold-localizable" resources, which include coherence and magic, a finite resource density is needed to achieve deep thermalization [106]. In particular, Ref. [36] studied the concept of coherence: a measure of the amount of superposition in a fixed orthonormal reference basis, chosen to be the computational basis $\{|z\rangle\}$ by convention, motivated by experimental considerations [107,108]. Intriguingly, they predicted that a generator state $|\Psi\rangle = \sum_z c_z|z\rangle$ may have near-maximal entanglement between A and B (hence its reduced density matrix is maximally mixed) yet coherence that is too low [109] such that it may fail to deeply thermalize to the Haar ensemble when measured in the computational basis.

Explicit examples of such states are furnished by so-called random subset phase states [36,38,110], recently introduced in the field of pseudoentanglement [111], with coherence tuned by the size of the subset in question. Concretely, for low-coherence states, it was argued that the resulting distribution of the projected ensemble is just a collection of (classical) computational states on the local subsystem—clearly nonergodic and having maximal accessible information instead of minimal. Further, this breaking of ergodicity happens in a robust fashion, with a phase transition separating the nonergodic distribution from a deeply thermalized Haar distribution, upon tuning the coherence density past a critical finite value. Intuitively, the reason for this is because measurements, which we take to be in the computational basis, select only those global bit-strings $z = (z_A, z_B)$ in the decomposition of the generator state $|\Psi\rangle = \sum_z c_z|z\rangle$ in which z_B agrees with the measurement outcome. If the number of such compatible z (captured precisely by coherence) is below some critical value, this may result in a vanishing fraction of states $|z_A\rangle$ that contribute to the projected state—clearly precluding it from behaving like a Haar random vector; see Ref. [36] for details. The upshot is that coherence (and, importantly, not only a nonzero value of it but a sufficiently high density of it) is a necessary ingredient for the appearance of Scrooge-like behavior in the projected ensemble.

We next focus on magic: a quantifier of the computational resources needed to describe quantum states beyond stabilizer states [112–114]. Recent works have quantified how magic governs the degree k of Haar k-designs formed in the projected ensemble [41,42,45], but here we present a crisp example showing that magic is also a necessary ingredient for the appearance of Scrooge-like behavior in the projected ensemble. Consider an N-qubit stabilizer state $|\Psi_{\text{Stab}}\rangle_{AB}$ with a reduced density matrix on a subsystem A maximally mixed, and construct the projected ensemble by measuring B in the standard computational basis (our argument in fact also applies more generally to any stabilizer basis). If the principle of maximum entropy for state ensembles applies, we should expect the projected ensemble to be well described by the Haar ensemble. Yet, as previously discussed in Sec. III, the projected ensemble is far from a 2-design, despite the fact that stabilizer states and Clifford unitaries form exact 3-designs. Our Theorems 2 and 3 (henceforth collectively referred to as "$2k \to k$" theorems), as well as Corollary III B 1, inform us that if we can augment the design properties of the initial state or measurement basis to be at least an approximate 4-design, then the projected ensemble is guaranteed to deeply thermalize to the Haar ensemble at the $k = 2$ moment; this missing ingredient is provided by magic [41,84,115]. More generally, our $2k \to k$ theorems can be harnessed to yield resource-theoretic bounds on the magic required for Scrooge universality to emerge.

However, simply adding non-Clifford elements to the system is not sufficient. To illustrate this point, consider the stabilizer state $|\Psi_{\text{Stab}}\rangle_{AB}$ of the previous example, and apply a random diagonal unitary on B, i.e., $|\Psi\rangle_{AB} = (I_A \otimes U_B^{\text{diag}})|\Psi_{\text{Stab}}\rangle_{AB}$, where $U_B^{\text{diag}} = \text{diag}(e^{i\varphi_j})$ in the computational basis, and $\varphi_j \in [0, 2\pi)$ are uniformly distributed. The resulting state $|\Psi\rangle_{AB}$ (which we call a diagonal-scrambled stabilizer state) now possesses high magic, induced by the diagonal unitary. Nevertheless, it is clear that its projected ensemble $\mathcal{E}(\Psi)$ still forms only a 1-design when B is measured in the computational basis since the unitary U_B commutes with the measurements,

both of which are diagonal in the computational basis. In contrast, if we pick U_B from a unitary $2k$-design, then Theorem 3 guarantees that the projected ensemble will now be close to a k-design.

What this example shows us is that the unitary (or dynamics) used to construct the generator state $|\Psi\rangle$ (which in the previous example consisted of the Clifford unitary used to prepare the stabilizer state followed by U_B^{diag}), starting from an unentangled product state on AB, needs to be sufficiently information scrambling (and, further, in a nonlocal way), in order to achieve emergent Scrooge universality. More precisely, we expect that if the scrambling unitary U_B hides quantum information about A (i.e., quantum correlations shared between A and B) in nonlocal degrees of freedom in B, then the emergent Scrooge projected ensemble on A also hides information about the measurement outcome on B from any measurement on A (i.e., attains minimal accessible information, precisely Scrooge behavior).

To summarize, the emergence of Scrooge behavior in projected ensembles requires the presence of coherence, magic, and information scrambling (nonlocal entanglement within the system). When any of these ingredients are absent or insufficient, obstructions to the formation of Scrooge projected ensembles can arise (see Fig. 2). This finding parallels a well-known fact in quantum complexity theory: Quantum advantage requires the coexistence of the resources of coherence, magic, and entanglement. When any of these ingredients are absent, quantum advantage is lost; i.e., efficient classical simulation becomes possible [116–118].

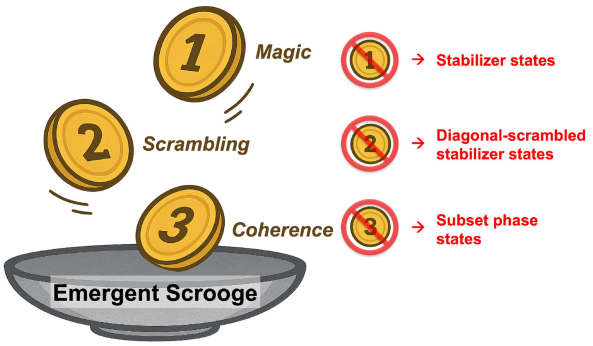

FIG. 2. Essential physical ingredients for emergent Scrooge designs. The emergence of Scrooge designs in projected ensembles requires the generator state $|\Phi\rangle_{AB}$ to exhibit magic (or nonstabilizerness), quantum information scrambling (via nonlocal entanglement), and coherence. When any of these ingredients is absent or insufficient, obstructions to Scrooge behavior can arise. Representative examples include stabilizer states, which lack magic; diagonal-scrambled stabilizer states (where a unitary diagonal in the computational basis is applied to subsystem B of a stabilizer state), which only weakly scrambles quantum information between subsystems A and B; and subset phase states, which may exhibit low coherence density depending on subset size.

In the next section, we perform numerical simulations supporting the aforementioned theoretical discussion and explore the interplay between coherence, magic, and quantum information scrambling in producing Scrooge behavior within the projected ensemble across a variety of models. These simulations encompass regimes that lie beyond the settings required of our theorems and thus complement the analytical results of Sec. III. Our simulations suggest that Scrooge designs can emerge even more broadly than in the settings for which we currently have rigorous guarantees.

V. NUMERICAL INVESTIGATIONS

In this section, we numerically study the projected ensembles generated in several quantum many-body settings in which the key physical ingredients of coherence, magic, and degree of information scrambling can be systematically controlled:

(i) Commuting quantum circuit evolution. We consider generator states produced from so-called "commuting quantum circuits" (i.e., circuits composed of mutually commuting gates, studied in the quantum computing community) [119,120]. Concretely, starting from a product state $|+\rangle^{\otimes N}$, we apply a unitary circuit that is diagonal in the measurement (computational) basis. Although such circuits can generate states with extensive entanglement and magic, by changing the regions where interactions are applied, the circuits' commuting structure allows for controlled tuning of how quantum information is delocalized across the system, such that the resulting projected ensembles can either succeed or fail to form a good state design. In addition, by rotating the measurement basis, we can inject controlled amounts of coherence into the system. We will show that this results in a coherence-induced deep thermalization transition, as predicted recently by Ref. [36].

(ii) Doped Clifford circuit evolution. Next, we study generator states produced from Clifford circuits doped with single-qubit non-Clifford gates. This construction allows information scrambling and magic to be tuned largely independently: Clifford dynamics efficiently scrambles quantum information, while the non-Clifford gates introduce controlled amounts of magic. We study how tuning these two control knobs affects the quality of the Scrooge designs formed in the projected ensemble, and we find that Scrooge behavior emerges only when both the circuit depth and density of non-Clifford gates are sufficiently large.

(iii) Ground states of 1D many-body Hamiltonians. Finally, we consider ground states of local 1D

integrable Hamiltonians, which are characterized by low entanglement and low complexity. To probe the role of the measurement basis, we apply various unitary rotations U_B to the qubits on subsystem B prior to measurement, which injects various amounts of magic and information scrambling. We find that if the rotations are insufficiently complex, no Scrooge behavior forms, while if the rotations are sufficiently complex, Scrooge behavior can emerge.

Together, our numerical investigations allow us to showcase the roles of coherence, magic, and quantum information scrambling in driving emergent local Scrooge behavior. They complement and go beyond existing numerical results that have demonstrated Scrooge behavior in late-time quenches or finite-temperature eigenstates of chaotic Hamiltonians [22,31]: While the latter examples are physically relevant for experiments, it is difficult to control and isolate the individual ingredients driving Scrooge behavior as we do here.

In our numerical simulations, we quantify the closeness of a projected ensemble $\mathcal{E}(\Psi)$ generated by $|\Psi\rangle_{AB}$ to the Scrooge(σ_A) k-designs via the kth moment trace distance

$$\Delta^{(k)}(\sigma_A) = \frac{1}{2}\|\rho_{\mathcal{E}}^{(k)} - \rho_{\text{Scrooge}}^{(k)}(\sigma_A)\|_1, \tag{23}$$

where σ_A is the reduced density matrix of $|\Psi\rangle_{AB}$ on A. By construction, the first-moment trace distance is $\Delta^{(1)} = 0$.

A. Commuting quantum circuit evolution

First, we study how different degrees of information scrambling in the circuit preparing a generator state, tuned by the support of the scrambler, affect the formation of Haar k-designs in its projected ensemble.

We initialize a tripartite system $A \cup B_1 \cup B_2$ of qubits in the state

$$|\eta\rangle = \prod_{i=1}^{N_A} \text{CZ}_{i,i+N_A} |+\rangle^{\otimes N}, \tag{24}$$

where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $\text{CZ}_{i,j} = \text{diag}(1,1,1,-1)$ is the controlled phase gate acting on the ith and jth qubits. We take the first N_A qubits to constitute subsystem A, and its complement is $B = B_1 \cup B_2$. Here, $|\eta\rangle$ describes a state composed of N_A pairs of maximally entangled states between A and B_1, with the remaining qubits in the x-polarized $|+\rangle$ state. The reduced density matrix σ_A on A is hence maximally mixed.

Next, we scramble the quantum correlations initially shared between A and B_1 into B_2. We consider applying a quantum circuit on $B = B_1 \cup B_2$, where the gates are all diagonal in the computational basis (which is also the measurement basis). As these gates are all mutually commuting, this class of circuits has been termed "commuting quantum circuits" [121,122]. We can model deep commuting circuit evolution, where the gates are randomly drawn, by a single random diagonal unitary $U_B^{\text{diag}} = \text{diag}(e^{i\varphi_j})$ applied on B, with $\varphi_j \in [0, 2\pi)$ drawn uniformly at random. Then, the generator state reads [123]

$$|\Psi\rangle_{AB} = (I_A \otimes U_B^{\text{diag}})|\eta\rangle. \tag{25}$$

Measuring B now in the computational basis to construct the projected ensemble on A, it is clear that it fails to form a state 2-design (which would be a Scrooge 2-design in this case where σ_A is maximally mixed), even though the state possesses both high entanglement and high magic. This failure occurs because U_B^{diag} commutes with the measurement operator and because $|\eta\rangle$ is a stabilizer state; hence, only a 1-design projected ensemble can result from stabilizer measurements on B.

The failure to form a k-design (for $k > 1$) can equivalently be understood by tracking the spread of quantum correlations in the Heisenberg picture: Because the circuit is commuting, an operator initially supported on B_2 never has support on A after the evolution. Thus, measurement outcomes on B_2 do not affect the projected ensemble, precluding the formation of 2-designs on A.

Now, suppose we change the scrambler U_B^{diag} into a diagonal unitary whose support *extends* to the full system AB [see Fig. 3(a)], i.e.,

$$|\Psi\rangle_{AB} = U_{AB}^{\text{diag}}|\eta\rangle. \tag{26}$$

This process yields so-called uniform random phase states [124]

$$|\Psi\rangle_{AB} = \frac{1}{2^{N/2}} \sum_j e^{-i\varphi_j}|j\rangle, \tag{27}$$

where $\{|j\rangle\}_j$ are computational basis states on AB and each φ_j is uniformly distributed. In contrast to the previous case, now an operator supported on B_2 can spread throughout the entire system, allowing measurements on B_2 to be correlated with the state on A: Quantum information is nonlocally scrambled across the entire system. As we will now show, this stronger form of information scrambling enables Scrooge behavior to emerge in the projected ensemble.

Uniform random phase states (27) are known to form approximate k-designs with additive error $\Theta(k^2/2^N)$ [124], while they are far from being a k-design in relative error for any $k > 1$; this precludes the application of Corollary III B 1. Nevertheless, we investigate the nature of the projected ensemble through numerical simulations. In Fig. 3(b), we plot $\Delta^{(2)}$ against N_A for different N_B. For finite N_B, we see $\Delta^{(2)}$ initially decays with N_A, then increases again when N_A is on the same order as N_B, where our numerics indicate that the turning point (defined to be

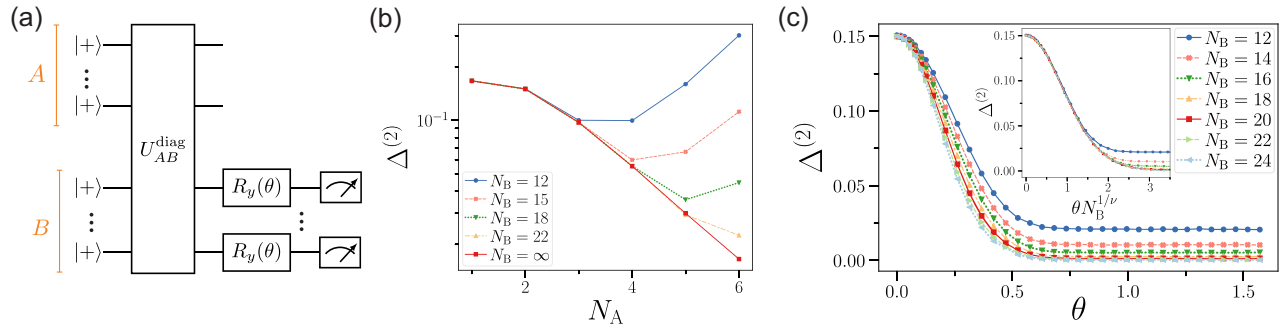

FIG. 3. Emergent 2-designs from commuting quantum circuit evolution. (a) Circuit to generate projected ensemble from commuting quantum circuit dynamics, Eq. (27). We apply a random diagonal unitary U_{AB}^{diag} on $|+\rangle^{\otimes N}$, rotate B by angle θ around the y axis with single-qubit unitaries $R_y(\theta) = \exp(-i\theta Y/2)$, and measure subsystem B in the computational basis. (b) Trace distance to Haar 2-design $\Delta^{(2)}$ against N_{A} for $\theta = 0$. For finite N_{B}, the moment operator of the projected ensemble is computed exactly (or using 3×10^6 measurement samples for $N_{\mathrm{B}} = 22$), and $\Delta^{(2)}$ itself is averaged over up to 20 random instances of the phase state. Here, $N_{\mathrm{B}} = \infty$ is computed from the ensemble of uniform random phase states over N_{A} qubits, where the moment operator is averaged over up to 10^7 random instances. (c) $\Delta^{(2)}$ against rotation angle θ for different N_{B}. The inset shows rescaled $\theta N_{\mathrm{B}}^{1/\nu}$, where $\nu = 2$. Here, $N_{\mathrm{A}} = 2$, $N = N_{\mathrm{A}} + N_{\mathrm{B}}$, and $\Delta^{(2)}$ is averaged over 10 random instances.

the minimal distance) occurs for $N_A \approx N_B/4$. This finding suggests that, as we send $N_{\mathrm{B}} \to \infty$, $\Delta^{(2)}$ should decay with N_A monotonically. We see this result in Fig. 3(b), where the numerics in the $N_B \to \infty$ limit are performed by taking the projected ensemble to be uniform random phase states, but now over N_A qubits (exact as $N_B \to \infty$). Note that the requirement of $N_A \gg 1$ arises because random phase states form a k-design with additive error $\Theta(k^2/2^{N_A})$. Thus, this regime is conceptually different from the conventional setting of deep thermalization, which concerns the behavior of the projected ensemble in the thermodynamic limit $N_B \to \infty$ at fixed N_A.

Deep thermalization behavior (local emergence of Haar designs for fixed N_A) can nevertheless be reproduced by applying a single-qubit rotation $R_y(\theta) = \exp(-i\theta Y/2)$, where Y is the Pauli-Y operator, to all qubits in subsystem B prior to measurement. For any angle $\theta \neq 0 \pmod{\pi}$ and fixed N_A, we provide numerical evidence in Fig. 3(c) that $\Delta^{(2)} \to 0$ as $N_B \to \infty$. This approach indicates that deep thermalization is obstructed for $\theta = 0$ but occurs for any nonzero rotation angle.

This observation aligns with the discussion in Ref. [36]. Those authors argued that a combination of the coherence of the initial state and the coherence of the measurement basis determines whether the projected ensemble on A is deeply thermalized or not. Now, random phase states, Eq. (27), have coherence $C = \alpha_0 N$, where $\alpha_0 = 1$ (recall Sec. IV for the definition of C) and the rotated measurement basis [the computational basis rotated by $R_y(\theta)$] has coherence $\alpha_m N_B$, where $\alpha_m = 0$ for $\theta = 0$ and α_m increases monotonically as θ increases [given by $\alpha_m \sim \theta^2 \ln(1/\theta^2)$ for small θ]. Reference [36] predicted a phase transition in the projected ensemble, with $\alpha_0 + \alpha_m < 1$ corresponding to a non-deep-thermal phase and $\alpha_0 + \alpha_m > 1$ a deep-thermal one; thus, $\alpha_0 + \alpha_m = 1$ is the critical point that maps to $\theta_c = 0$ in our setup in which $\alpha_0 = 1$. Our numerical findings in Fig. 3(c) confirm this prediction; a finite-size scaling ansatz, which fits the data well, indicates that trace distance $\Delta^{(2)}$ approaches zero for any fixed nonzero θ as N_B increases.

Thus, our numerical investigations have shown that the emergence of state designs requires sufficient scrambling (beyond just entanglement and magic) between A and B, as well as coherence, which is induced by the global diagonal unitary and coherence-injecting y rotations, respectively.

B. Doped Clifford circuits

Next, we explore the necessity of both magic and information scrambling for the projected ensemble to exhibit emergent Scrooge designs. As depicted in Fig. 4(a), we consider an initial state defined on a tripartite system $A \cup B_1 \cup B_2$ such that $B = B_1 \cup B_2$, with tunable entanglement between A and B_1:

$$|\Psi(\chi)\rangle = |\psi_{\mathrm{ent}}(\chi)\rangle^{\otimes N_{\mathrm{A}}} |0_{B_2}\rangle^{\otimes N_{\mathrm{B}} - N_{\mathrm{A}}}. \tag{28}$$

Here, $|\psi_{\mathrm{ent}}(\chi)\rangle = \cos(\chi/2)|0_A 0_{B_1}\rangle + \sin(\chi/2)|1_A 1_{B_1}\rangle$ is a two-qubit state where the first qubit is in A and the second in B_1. Note that χ controls the entanglement between A and B_1, with maximal entanglement between A and B_1 achieved at $\chi = \pi/2$ (at this point, $|\psi_{\mathrm{ent}}\rangle$ is a Bell state), resulting in the reduced density matrix σ_A being maximally mixed. To probe the emergence of Scrooge designs away from the infinite-temperature limit, we fix $\chi = \pi/6$. We next apply on B a circuit U_B composed of d layers, where each layer consists of randomly chosen single-qubit Clifford gates on each qubit and fixed CNOT gates arranged in a 1D nearest-neighbor geometry. Additionally, we dope the overall circuit with N_{T} T-gates defined as $T = \mathrm{diag}(1, e^{-i\pi/4})$ [84,125] randomly inserted over space and time. While

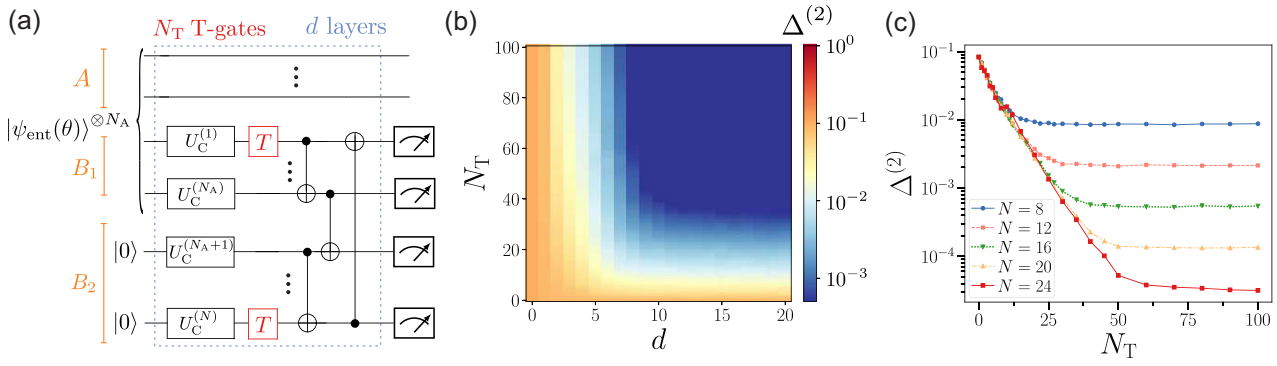

FIG. 4. Emergent Scrooge 2-designs from doped Clifford circuit evolution. (a) We consider projected ensembles from the generator state Eq. (28), where we measure in the computational basis on $B = B_1 \cup B_2$ after applying a quantum circuit composed of d layers of local Clifford gates (random single-qubit Clifford gates together with fixed CNOT gates) doped with N_T T-gates in total, randomly placed within the circuit. (b) Heat map of the trace distance to Scrooge 2-design $\Delta^{(2)}$ in the d versus N_T plane, for $\chi = \pi/6$. Here, we fixed $N_A = 1$, $N_B = N - N_A$, and $N = 20$. (c) $\Delta^{(2)}$ against N_T for different total qubit numbers N and fixed $d = 30$. We choose $N_A = 1$, and $\Delta^{(2)}$ is averaged over 500 random realizations of the circuit. Similar behavior is observed for higher moments $k > 2$ (see Appendix G).

the d layers of Clifford CNOT gates introduce an increasing amount of information scrambling on B, the N_T T-gates inject magic into the circuit, which allows us to study the effects of information scrambling and magic on $\Delta^{(2)}$ relatively independently.

In Fig. 4(b), we show a heat map of $\Delta^{(2)}$ in the d-N_T plane. We see that for small N_T or d, $\Delta^{(2)}$ is relatively large. Only with both sufficient magic and information scrambling (large N_T and d) do we achieve a Scrooge 2-design on A with small $\Delta^{(2)}$ [42], as expected. Our numerical results do not depend strongly on χ; see Appendix G.

Next, in Fig. 4(c), we study the role of magic and system size N in more detail. We plot $\Delta^{(2)}$ for different N and N_T (for large d), finding that $\Delta^{(2)}$ decreases exponentially in N_T, reaching a saturation value due to finite-size effects. The saturation value of $\Delta^{(2)}$ decreases exponentially with N, which is reached for $N_T \approx 2.5N$, closely matching the saturation transition found in magic resource theories [126,127]. This feature highlights that an extensive amount of magic is needed for deep thermalization to the Scrooge ensemble (more precisely, the emergence of Scrooge designs).

We note that similar behavior is observed when the doped Clifford circuit is applied on AB, instead of just B, which we show in Appendix H.

C. Ground states of many-body Hamiltonians

Finally, we study emergent Scrooge designs in naturally realizable many-body systems. In contrast to the usual paradigm of quantum chaotic systems where deep thermalization has been probed [22,31,35], here we consider the projected ensemble generated from ground states of highly structured, integrable many-body systems, subject to various scramblers of quantum information on the system B where measurements are taken. We find that the projected ensemble can still form Scrooge k-designs if the measurement basis is sufficiently complex, demonstrating Theorem 3 in action.

Concretely, we study the 1D transverse-field Ising Hamiltonian of N spin-$\frac{1}{2}$ particles (which we will refer to also as qubits henceforth) with external field h on a chain with periodic boundary conditions,

$$H_{\text{Ising}} = -\sum_{j=1}^{N} X_j X_{j+1} - h \sum_{j=1}^{N} Y_j, \tag{29}$$

where X_j (Y_j) is the Pauli X (Y) operator acting on the jth qubit. This model is well known to be integrable and, in fact, mappable to free fermions by the Jordan-Wigner transformation [128]. Consequently, its eigenstates are all highly structured, with low complexity, even at the quantum critical point $h = 1$.

We focus on the ground states $|\Psi\rangle_{AB}$ of H_{Ising} in the spin-flip even sector with varying h, where A is chosen to be a small contiguous region of spins and B its complement. Prior to measuring B in the computational basis to construct the projected ensemble, we apply different scrambling unitaries U_B on B, chosen from (i) the identity I_B, (ii) tensor products of N_B (independently drawn) Haar random single-qubit unitaries, (iii) random Clifford unitaries, and (iv) Haar random unitaries over N_B qubits.

In Fig. 5(a), we plot the second-moment trace distance $\Delta^{(2)}$ of the projected ensemble to the Scrooge ensemble constructed with the reduced density on A for $N_A = 1$ and $N_B = 19$, varying h and for different choices of U_B. When U_B is the identity or a product of single-qubit unitaries, we find the value of $\Delta^{(k)}$ remains high compared to the Clifford and Haar cases, even though $N_B \gg N_A$. This behavior can

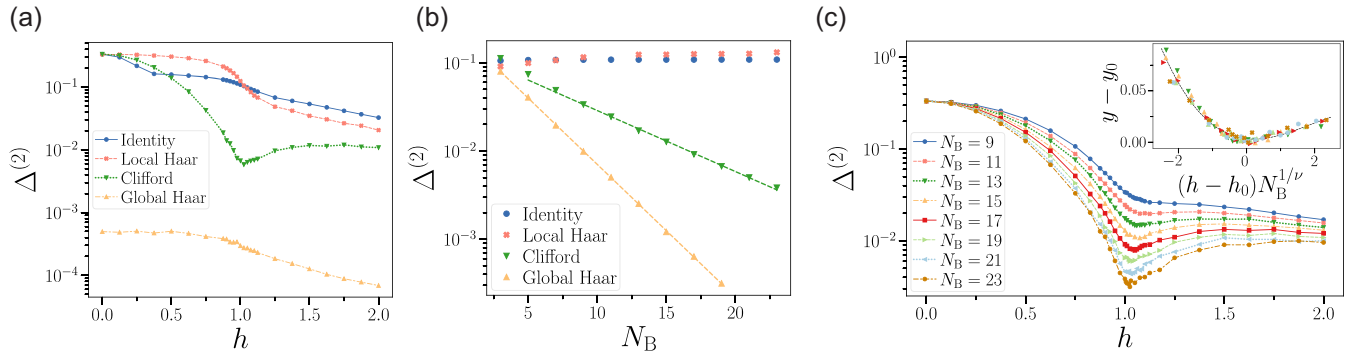

FIG. 5. Emergent Scrooge 2-designs from 1D integrable ground states. We study projected ensembles generated from the ground state of the transverse-field Ising model (29), where we measure B in various bases that arise from rotations of the computational basis by various unitaries U_B applied on B only. (a) $\Delta^{(2)}$ against field h for $N_A = 1$ qubits, with U_B chosen from the identity, single-qubit Haar random unitaries, random Clifford unitaries, or (global) unitaries drawn from the Haar measure on $N_B = 19$ spins. (b) $\Delta^{(2)}$ against N_B for different U_B and $h = 1$. Dashed lines show the fit with $\Delta^{(2)} \sim 2^{-\alpha N_B}$, where we find $\alpha_{\text{Clifford}} \approx 0.23$ and $\alpha_{\text{Haar}} \approx 0.5$. (c) $\Delta^{(2)}$ against h for U_B being random Clifford unitaries and different N_B. The inset shows the same data rescaled by defining $y = \log_2(\Delta^{(2)})/N_B$ and subtracting the field h_0 with minimal y_0 around the critical field $h_c = 1$. When rescaling the field with $N_B^{1/\nu}$ where $\nu = 1$ from the Ising universality class, the data for different N_B nearly collapse to a single curve, a hallmark of universality of the critical field $h_c = 1$ (see Appendix I). We fit the collapsed data with a third-order polynomial as a dashed line. Similar behavior is observed for higher moments $k > 2$ (see Appendix I).

presumably be attributed to the low complexity nature of the ground state (even at the critical point $h = 1$) and to the fact that single-qubit measurements do not inject additional complexity. As a result, for product unitaries, quantum correlations between A and B are not efficiently scrambled into nonlocal correlations in B.

In contrast, for the Clifford and Haar cases, the unitaries are apparently scrambling enough to yield comparatively lower $\Delta^{(2)}$, with Haar random unitaries on B being (unsurprisingly) more scrambling in nature and hence exhibiting the lowest $\Delta^{(2)}$. We note that while Clifford unitaries on their own are insufficient to guarantee Scrooge 2-designs due to an absence of magic, it is reasonable to expect that the ground state of H_{Ising}, despite its free-fermionic nature, supplies sufficient magic to enhance scrambling on B toward effective 4-design behavior, potentially allowing Theorem 3 to apply (although we do not establish this rigorously). Curiously, we find that $\Delta^{(2)}$ shows a pronounced dip at the quantum critical point $h \approx 1$, suggesting that magic is highest among all ground states there, which is in line with observations made in previous works [129].

In Fig. 5(b), we study the scaling of the trace distance $\Delta^{(2)}$ with the size N_B of the measured system at the critical point $h = 1$. We find that $\Delta^{(2)}$ is nearly independent of N_B for local measurement bases defined by rotating the computational basis by identity and single-qubit Haar unitaries, confirming that local measurement bases do not yield emergent Scrooge designs on A. In contrast, measuring in a basis that is scrambled by Cliffords or Haar-random unitaries on B yields an exponentially decaying distance $\Delta^{(2)} \sim 2^{-\alpha N_B}$, with $\alpha \approx 0.2$ for Cliffords and $\alpha \approx 0.5$ for Haar, values that may indicate their scrambling power.

Finally, in Fig. 5(c), we study the Clifford case in greater detail for different N_B. We observe that the dip near $h \approx 1$ becomes more pronounced for larger N_B. This behavior is closely tied to the fact that magic is maximized at the critical point [129] and that magic is known to enhance the scrambling power of Clifford unitaries [41,85,115]. Further, as shown in the inset, we find universal behavior around the critical point, with data for different N_B nearly collapsing to a single curve when rescaling the field with $N_B^{1/\nu}$, where $\nu = 1$, which is precisely the critical exponent of the Ising universality class [129,130]. A qualitatively similar behavior is also observed for the ground state of the 1D Heisenberg model, as we show in Appendix J.

Altogether, our numerical results show that magic and quantum information scrambling are key physical ingredients for emergent Scrooge designs in projected ensembles. These properties can manifest intrinsically in naturally occurring chaotic many-body states or be injected externally by applying a scrambling unitary on subsystem B. We provide additional numerics supporting our observations in Appendix K. There, we show that a random stabilizer state can produce emergent state designs on A if a single layer of T gates is applied to B before measuring in the X basis, implying that it suffices for magic to be injected locally [22]. Our examples extend beyond the regimes in which Scrooge behavior can be rigorously established by our theorems or in those considered by other deep thermalization works, providing evidence that the emergence of information-stingy ensembles may be a generic phenomenon in quantum many-body systems.

VI. DISCUSSION AND OUTLOOK

This work addresses a fundamental question: What universal principle governs the higher-order statistical fluctuations of quantum many-body systems? Building on prior insights into the information-stinginess of nature, we provide a unified and rigorous framework for characterizing the emergence of maximally entropic Scrooge-like behavior. Our results, by clarifying when and why Scrooge universality arises, place earlier observations on a precise quantitative footing.

Our approach also highlights the robustness of Scrooge universality across diverse physical settings. In particular, Theorem 2 shows that local Scrooge behavior emerges whenever the global generator state is drawn from a Scrooge design, regardless of the microscopic mechanism by which that state is prepared. This mechanism-agnostic perspective opens the door to identifying new physical scenarios in which Scrooge behavior may arise.

Our numerical investigations further elucidate the physical resources underlying this universality. We find that quantum coherence, entanglement, nonstabilizerness (magic), and information scrambling play essential roles: Removing any one of these ingredients can obstruct the emergence of Scrooge behavior in projected ensembles. Beyond these core ingredients, other resources may also play a role, including non-Gaussianity [131], whose absence leads to restricted forms of deep thermalization [28,30], and imaginarity [132], which can impose effective realness on projected ensembles and alter their universal structure from unitary to orthogonal symmetry classes [29]. Developing a systematic classification of such resources (and incorporating them into refined maximum-entropy principles) is an important direction for future work.

Several concrete open questions arise from our work. First, while our "$2k \to k$" theorems are optimal for $k = 2$, it remains open whether the required design order can be reduced for high moments, for example, to "$k + \mathcal{O}(1) \to k$." Second, our findings suggest a potential unifying perspective on two complementary notions of information hiding. Applying a scrambling unitary to the measurement basis on B (as in Theorem 3) hides *quantum* information about A in highly nonlocal observables on B. On the other hand, the resulting projected ensemble on A hides *classical* information about the measurement outcome on B. Whether these two mechanisms fit into a common theoretical framework remains an intriguing question. Third, the universality of Scrooge ensembles may enable new applications in quantum information science. Emergent state designs in infinite-temperature deep thermalization have already led to advances in learning and benchmarking analog quantum simulators [17,35,66,67]. Extending these ideas to Scrooge designs could yield analogous protocols operating in more general, physically constrained settings.

Finally, our results raise the broader question of whether an information-theoretic analog of the maximum entropy principle extends to even more general settings, such as incomplete measurements or noisy dynamics, where each state in a projected ensemble becomes mixed rather than pure [133,134]. An affirmative answer would point toward a unified framework for the emergence of universal randomness in both pure and mixed quantum ensembles, marking a significant conceptual advance in quantum statistical mechanics.

ACKNOWLEDGMENTS

We thank Andreas Elben, Soumik Ghosh, Hsin-Yuan Huang, Daniel Mark, Max McGinley, Alexey Milekhin, Arjun Mirani, Yihui Quek, Thomas Schuster, Federica Surace, Shreya Vardhan, and Michelle Xu for insightful discussions. We are especially grateful to Soonwon Choi for improving the presentation of our manuscript, and to Max McGinley and Thomas Schuster for sharing their preliminary results [62] with us, which were invaluable for our analysis. W. W. H. is supported by the National Research Foundation (NRF), Singapore, through the NRF Fellowship No. NRF-NRFF15-2023-0008, and through the National Quantum Office, hosted in A*STAR, under its Centre for Quantum Technologies Funding Initiative (Grant No. S24Q2d0009). J. P. acknowledges support from the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator, and the National Science Foundation (Grant No. PHY-2317110). The Institute for Quantum Information and Matter is an NSF Physics Frontiers Center.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: NOTATION AND PRELIMINARIES

In this appendix, we provide a self-contained introduction to the notation and mathematical preliminaries, which we use throughout this work.

1. Norms of operators and random variables

For an arbitrary operator A, we denote its Schatten-p norm by $\|A\|_p$, where

$$\|A\|_p = [\mathrm{Tr}(|A|^p)]^{1/p}, \qquad p \in [1, \infty), \tag{A1}$$

and $\|A\|_\infty$ is the operator norm. Schatten-p norms satisfy monotonicity: For $1 \le p \le q \le \infty$,

$$\|A\|_p \geq \|A\|_q. \tag{A2}$$

For an arbitrary random variable X, we denote its L^p-norm by $\|X\|_{L^p}$, where

$$\|X\|_{L^p} = [\mathbb{E}(|X|^p)]^{1/p}, \qquad p \in [1, \infty), \tag{A3}$$

and $\|X\|_{L^\infty} = \operatorname{ess\,sup}|X|$ is the essential supremum of X. Note that L^p norms satisfy the following inequality: For $1 \leq p \leq q \leq \infty$,

$$\|X\|_{L^p} \leq \|X\|_{L^q}. \tag{A4}$$

2. Symmetric subspace, permutation operators, and moments of the Haar ensemble

Given the k-fold Hilbert space $\mathcal{H}^{\otimes k}$, the symmetric subspace of $\mathcal{H}^{\otimes k}$, denoted $\mathcal{H}_{\mathrm{sym}}^{(k)}$, is the vector space spanned by all states that are invariant under an arbitrary permutation of the k replicas. To each permutation $\pi \in S_k$ (where S_k is the symmetric group of order k), we can associate a corresponding permutation operator $\hat{\pi} \in \mathcal{L}(\mathcal{H}^{\otimes k})$ that permutes between the k copies. The permutation operator $\hat{\pi}$ is a unitary representation of π. Thus,

$$\mathcal{H}_{\mathrm{sym}}^{(k)} = \mathrm{span}\{|\psi\rangle \in \mathcal{H}^{\otimes k} | \hat{\pi}|\psi\rangle = |\psi\rangle \ \forall \ \pi \in S_k\}. \tag{A5}$$

The dimension of the symmetric subspace $\mathcal{H}_{\mathrm{sym}}^{(k)}$ is

$$D_k \equiv \binom{D+k-1}{k}, \tag{A6}$$

where D is the dimension of the Hilbert space $\mathcal{H}$. For $k^2 \ll D$, it is useful to write

$$D_k = \frac{D^k}{k!}\left(1 + \mathcal{O}\left(\frac{k^2}{D}\right)\right). \tag{A7}$$

Thus, $D_k \approx D^k/k!$. The orthogonal projector onto the symmetric subspace has a special meaning: It is proportional to the kth moment of the Haar ensemble $\mathrm{Haar}(D)$,

$$\rho_{\mathrm{Haar}}^{(k)} = \frac{1}{D_k}\hat{P}_{\mathrm{sym}}^{(k)} = \frac{1}{k! D_k}\sum_{\pi \in S_k}\hat{\pi}, \tag{A8}$$

where

$$\hat{P}_{\mathrm{sym}}^{(k)} = \frac{1}{k!}\sum_{\pi \in S_k}\hat{\pi} \tag{A9}$$

is the orthogonal projector onto $\mathcal{H}_{\mathrm{sym}}^{(k)}$. This relationship can be derived using the Schur-Weyl duality [135]. In a slight abuse of notation, we will also use Haar (D) to denote the Haar measure on the unitary group $U(D)$. The mathematical object under consideration (pure quantum state or unitary operator) will be clear from the context.

In this paper, we often decompose the system into complementary subsystems A and B. Thus, $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, with Hilbert space dimensions D_A and D_B, respectively ($D = D_A D_B$). To disambiguate the notation, we will use subscripts to indicate the subsystem under consideration, where appropriate. For example, the dimension of the k-fold symmetric subspaces for subsystems A and B will be denoted $D_{A,k}$ and $D_{B,k}$, respectively.

3. Weingarten calculus for the unitary group

The Weingarten calculus provides a very useful tool for evaluating polynomial functions of Haar random states and unitaries. Here, we briefly introduce the results relevant for this work and establish the notation used throughout the manuscript. A detailed treatment can be found in Refs. [73–75].

For permutations $\sigma, \pi \in S_k$, we define the $k! \times k!$ Gram matrix G, with matrix elements

$$G_{\sigma\pi} = \mathrm{Tr}(\hat{\sigma}^\dagger \hat{\pi}) = D^{\#\mathrm{cycles}(\sigma^{-1}\pi)}, \tag{A10}$$

where $\hat{\sigma}$ and $\hat{\pi}$ are the permutation operators defined above. Here, $\#\mathrm{cycles}(\pi)$ counts the number of disjoint cycles in the permutation π. The inverse of G gives the Weingarten matrix, where the matrix elements

$$\mathrm{Wg}(\sigma^{-1}\pi, D) \equiv (G^{-1})_{\sigma\pi} \tag{A11}$$

are known as the Weingarten functions associated with the unitary group. Note that the Gram matrix is only invertible for $1 \leq k \leq D$, which we will assume throughout this work. The main formula for our purposes is the k-fold twirling identity,

$$\mathop{\mathbb{E}}_{U \sim \mathrm{Haar}(D)}[U^{\otimes k} A U^{\dagger \otimes k}] = \sum_{\sigma,\pi \in S_k} \mathrm{Wg}(\sigma^{-1}\pi, D)\mathrm{Tr}(A\hat{\pi}^\dagger)\hat{\sigma}. \tag{A12}$$

As a consistency check, setting $A = (|0\rangle\langle 0|)^{\otimes k}$ reproduces Eq. (A8), for any arbitrary reference state $|0\rangle$.

4. Projected ensemble

Here, we give a concise review of the projected ensemble and establish the notation used in this work. Starting from a bipartite quantum state $|\Psi\rangle_{AB}$, which we refer to as the generator state, we measure subsystem B in a complete orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. By default, we choose the measurement basis to be the computational basis, unless stated otherwise. Note that there is no loss of generality here since any rotation to the measurement basis can be

absorbed into the definition of $|\Psi\rangle_{AB}$. This process yields the projected state

$$|\psi_z\rangle = \frac{(I_A \otimes \langle z|_B)|\Psi\rangle_{AB}}{\sqrt{p_z}} \equiv \frac{|\tilde{\psi}_z\rangle}{\sqrt{p_z}} \tag{A13}$$

on subsystem A, where

$$p_z = \langle\tilde{\psi}_z|\tilde{\psi}_z\rangle = (I_A \otimes \langle z|_B)(|\Psi\rangle\langle\Psi|)(I_A \otimes |z\rangle_B) \tag{A14}$$

is the Born probability of measuring the outcome z on B. Collectively, this result defines the projected ensemble on A generated by $|\Psi\rangle_{AB}$,

$$\mathcal{E}(\Psi) = \{p_z, |\psi_z\rangle\}_{z=1}^{D_B}. \tag{A15}$$

The kth moment of the projected ensemble reads

$$\rho_{\mathcal{E}}^{(k)}(\Psi) = \sum_{z=1}^{D_B} p_z(|\psi_z\rangle\langle\psi_z|)^{\otimes k} = \sum_{z=1}^{D_B} \frac{(|\tilde{\psi}_z\rangle\langle\tilde{\psi}_z|)^{\otimes k}}{\langle\tilde{\psi}_z|\tilde{\psi}_z\rangle^{k-1}}. \tag{A16}$$

For $k = 1$, this finding is exactly the reduced density matrix of $|\Psi\rangle_{AB}$ on A. To quantify the statistical closeness between the projected ensemble and a reference ensemble $\mathcal{E}_{\text{ref}}$ (which, in the context of this work, is the Scrooge ensemble or a variant of it), we compute the trace distance of their kth moments,

$$\Delta^{(k)} = \frac{1}{2}\|\rho_{\mathcal{E}}^{(k)}(\Psi) - \rho_{\text{ref}}^{(k)}\|_1, \tag{A17}$$

where $\rho_{\text{ref}}^{(k)}$ is the kth moment of $\mathcal{E}_{\text{ref}}$. In this work, we often consider the scenario where $|\Psi\rangle_{AB}$ is drawn from some distribution. Thus, a useful metric is the average trace distance

$$\mathbb{E}_\Psi[\Delta^{(k)}] = \frac{1}{2}\mathbb{E}_\Psi\|\rho_{\mathcal{E}}^{(k)}(\Psi) - \rho_{\text{ref}}^{(k)}\|_1. \tag{A18}$$

It is important that the trace distance is computed before the average. This requirement implies that if $\mathbb{E}_\Psi[\Delta^{(k)}]$ is small, the projected ensemble $\rho_{\mathcal{E}}^{(k)}(\Psi)$ is close to $\mathcal{E}_{\text{ref}}$ up to the kth moment, for *any* generator state $|\Psi\rangle$ drawn from the distribution, with high probability. This fact can be easily derived from Markov's inequality, which gives

$$\mathbb{P}_\Psi(\Delta^{(k)} > \delta) \le \frac{\mathbb{E}_\Psi[\Delta^{(k)}]}{\delta} \tag{A19}$$

for any $\delta > 0$. Computing the trace distance analytically is often difficult. It is more analytically tractable to compute the Schatten 2-norm distance (also known as the Hilbert-Schmidt distance), which gives an upper bound for $\mathbb{E}_\Psi[\Delta^{(k)}]$:

$$\mathbb{E}_\Psi[\Delta^{(k)}] \le \frac{1}{2}\sqrt{D_{A,k}}(\mathbb{E}_\Psi[\|\rho_{\mathcal{E}}^{(k)}(\Psi) - \rho_{\text{ref}}^{(k)}\|_2^2])^{1/2}. \tag{A20}$$

5. Scrooge ensemble

The Scrooge ensemble is defined in Ref. [49] as the state ensemble $\mathcal{E}$ that attains the minimum accessible information $\mathcal{I}_{\text{acc}}(\mathcal{E})$, among all ensembles that realize a given density matrix σ. Here, we briefly review the concept of accessible information [68,69], which can be understood by the following scenario. Bob samples a state drawn from the ensemble $\mathcal{E} = \{p_z, |\psi_z\rangle\}_z$ and sends the state to Alice. The ensemble $\mathcal{E}$ is known to both parties *a priori*. Alice's goal is to determine which state Bob has sent, i.e., the classical label z, by performing a measurement (mathematically represented by the POVM M) on the state. In the context of the projected ensemble, p_z is the probability of measuring the outcome z on subsystem B, and $|\psi_z\rangle$ is the normalized state on subsystem A conditioned on the outcome z. Note that, as explained in Sec. II of the main text, we can more generally consider z in $\mathcal{E}$ to be a continuous label, upon which the probability p_z should be replaced by a probability measure $\mu(z)$.

The accessible information quantifies how much classical information about z can be gained by Alice via applying an optimal single-copy measurement on the quantum state. Thus, we define

$$\mathcal{I}_{\text{acc}}(\mathcal{E}) = \sup_M \mathcal{I}(\mathcal{E}:M), \tag{A21}$$

where $\mathcal{I}(\mathcal{E}:M)$ is the mutual information between the measurement M and the ensemble $\mathcal{E}$. This approach leads to the Scrooge ensemble (Definition 1)

$$\begin{aligned} &\text{Scrooge}(\sigma) = \arg\min_{\mathcal{E}} \mathcal{I}_{\text{acc}}(\mathcal{E}), \\ &\text{subject to} \quad \mathop{\mathbb{E}}_{\psi\sim\mathcal{E}}|\psi\rangle\langle\psi| = \sigma. \end{aligned} \tag{A22}$$

The Scrooge ensemble can be explicitly constructed as

$$\text{Scrooge}(\sigma) = \left\{D\langle\phi|\sigma|\phi\rangle \mathrm{d}\phi, \frac{\sqrt{\sigma}|\phi\rangle}{\|\sqrt{\sigma}|\phi\rangle\|}\right\}, \tag{A23}$$

with $\mathrm{d}\phi$ the Haar measure on $\mathbb{C}^D$. The kth moment of the Scrooge ensemble is computed as follows:

$$\begin{aligned} \rho_{\text{Scrooge}}^{(k)}(\sigma) &= \mathop{\mathbb{E}}_{\psi\sim\text{Scrooge}(\sigma)}[(|\psi\rangle\langle\psi|)^{\otimes k}] \\ &= \mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)}\left[D\langle\phi|\sigma|\phi\rangle \frac{(\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma})^{\otimes k}}{\langle\phi|\sigma|\phi\rangle^k}\right] \\ &= D\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)}\left[\frac{(\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma})^{\otimes k}}{\langle\phi|\sigma|\phi\rangle^{k-1}}\right]. \end{aligned} \tag{A24}$$

In the limit $\sigma \to I/D$ (with I the identity operator), $\rho^{(k)}_{\mathrm{Scrooge}}(\sigma)$ reduces to $\rho^{(k)}_{\mathrm{Haar}}$ in Eq. (A8). More generally, for any measurable function f,

$$\mathop{\mathbb{E}}_{\psi\sim\mathrm{Scrooge}(\sigma)}[f(\psi)] = D \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\left[\langle\phi|\sigma|\phi\rangle f\left(\frac{\sqrt{\sigma}|\phi\rangle}{\langle\phi|\sigma|\phi\rangle^{1/2}}\right)\right]. \tag{A25}$$

Unlike the Haar ensemble, the kth moment of Scrooge(σ) does not have a simplified form in terms of permutation operators. The main technical difficulty arises from the form of $\rho^{(k)}_{\mathrm{Scrooge}}(\sigma)$ written above, which is a rational function of the Haar random state $|\phi\rangle$. This finding implies that the standard Weingarten calculus, which is very useful in evaluating polynomial functions of Haar random states, does not apply here. Nonetheless, if σ is a low-purity state, we can obtain a simple approximation to $\rho^{(k)}_{\mathrm{Scrooge}}(\sigma)$, with an error that is controlled by the purity of σ. This case will be discussed in Appendix B.

6. Scrooge k-designs

State k-designs provide a low-order approximation of the Haar ensemble, by matching only the first k statistical moments of the Haar ensemble. We can generalize this definition to Scrooge(σ) k-designs, which match the first k moments of Scrooge(σ). The moments do not need to match exactly; for practical purposes, it suffices for the moments to be close, up to a small error ε. We reproduce Definition 2 from the main text.

Definition 3 (Approximate Scrooge k-designs). Let $\mathcal{E}$ be an ensemble of pure states, with the kth moment

$$\rho^{(k)}_{\mathcal{E}} = \mathop{\mathbb{E}}_{\psi\sim\mathcal{E}}[(|\psi\rangle\langle\psi|)^{\otimes k}]. \tag{A26}$$

Note that $\mathcal{E}$ is a Scrooge(σ) k-design with additive error ε if

$$\frac{1}{2}\left\|\rho^{(k)}_{\mathcal{E}} - \rho^{(k)}_{\mathrm{Scrooge}}(\sigma)\right\|_1 \le \varepsilon, \tag{A27}$$

and it is a Scrooge(σ) k-design with relative error ε if

$$(1-\varepsilon)\rho^{(k)}_{\mathrm{Scrooge}}(\sigma) \preccurlyeq \rho^{(k)}_{\mathcal{E}} \preccurlyeq (1+\varepsilon)\rho^{(k)}_{\mathrm{Scrooge}}(\sigma). \tag{A28}$$

Here, $\rho^{(k)}_{\mathrm{Scrooge}}(\sigma)$ is the kth moment of Scrooge(σ), given in Eq. (15).

The notation $A \preccurlyeq B$ for positive-semidefinite operators A and B is equivalent to the statement that $B - A$ is positive semidefinite.

A relative error of ε implies an additive error of ε. This error can be easily derived: Suppose $\mathcal{E}$ has relative error ε. Then,

$$\begin{aligned}\frac{1}{2}\left\|\rho^{(k)}_{\mathcal{E}} - \rho^{(k)}_{\mathrm{Scrooge}}(\sigma)\right\|_1 &= \sup_{0\preccurlyeq P\preccurlyeq I} \mathrm{Tr}[P(\rho^{(k)}_{\mathcal{E}} - \rho^{(k)}_{\mathrm{Scrooge}}(\sigma))] \\ &\le \varepsilon \sup_{0\preccurlyeq P\preccurlyeq I} \mathrm{Tr}(P\rho^{(k)}_{\mathrm{Scrooge}}(\sigma)) \\ &\le \varepsilon\left\|\rho^{(k)}_{\mathrm{Scrooge}}(\sigma)\right\|_1 \\ &= \varepsilon.\end{aligned} \tag{A29}$$

However, the converse is not true. The relative error can be arbitrarily large, even if the additive error is small. For example, if $\rho^{(k)}_{\mathcal{E}}$ has support on the null space of $\rho^{(k)}_{\mathrm{Scrooge}}(\sigma)$, the relative error is infinitely large.

For ensembles of quantum states, the additive error is a meaningful measure of statistical closeness between the ensembles. In the context above, if $\mathcal{E}$ is a Scrooge(σ) k-design with additive error ε, then any quantum measurement that can act collectively on up to k copies of the quantum state can only distinguish between $\mathcal{E}$ and Scrooge (σ) with a success probability of at most [69]

$$p_{\mathrm{succ}} = \frac{1+\varepsilon}{2}. \tag{A30}$$

In other words, if the additive error ε is small, the optimal measurement strategy is only marginally better than the naive strategy of random guessing, which has a success probability of $1/2$.

7. Approximation of the kth moment

To aid the analysis in the paper, we introduce the following technical lemma, which provides an approximation to the kth moment of an ensemble.

Lemma 2 (kth moment approximation). Consider the ensemble of pure states

$$\mathcal{E} = \left\{p_j = \langle\psi'_j|\psi'_j\rangle, |\psi_j\rangle = \frac{|\psi'_j\rangle}{\sqrt{p_j}}\right\}_{j=1,\ldots,|\mathcal{E}|} \tag{A31}$$

with kth moment

$$\rho^{(k)}_{\mathcal{E}} = \sum_{j=1}^{|\mathcal{E}|} \frac{(|\psi'_j\rangle\langle\psi'_j|)^{\otimes k}}{p_j^{k-1}}, \tag{A32}$$

where each unnormalized state $|\psi'_j\rangle$ in the ensemble is randomly drawn from a distribution $\mathcal{D}_j$. Construct the proxy

$$\tilde{\rho}^{(k)}_{\mathcal{E}} = \sum_{j=1}^{|\mathcal{E}|} \frac{(|\psi'_j\rangle\langle\psi'_j|)^{\otimes k}}{q_j^{k-1}}, \tag{A33}$$

where $q_j > 0$ are independent of $|\psi'_j\rangle$. Then,

$$\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j}\left\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1 \le \sum_{j=1}^{|\mathcal{E}|}\left(\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j} p_j^2\right)^{1/2}\left(1-\frac{2}{q_j^{k-1}}\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j} p_j^{k-1}+\frac{1}{q_j^{2k-2}}\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j} p_j^{2k-2}\right)^{1/2}. \tag{A34}$$

Proof.

$$\begin{aligned}\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j}\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\|_1 &\le \mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j}\sum_{j=1}^{|\mathcal{E}|} p_j^k\left|\frac{1}{p_j^{k-1}}-\frac{1}{q_j^{k-1}}\right| && \text{(triangle inequality)}\\ &= \sum_{j=1}^{|\mathcal{E}|}\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j} p_j\left|1-\frac{p_j^{k-1}}{q_j^{k-1}}\right| && \text{(linearity of expectation)}\\ &\le \sum_{j=1}^{|\mathcal{E}|}\left(\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j} p_j^2\right)^{1/2}\left(1-\frac{2}{q_j^{k-1}}\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j} p_j^{k-1}+\frac{1}{q_j^{2k-2}}\mathop{\mathbb{E}}_{\psi'_j\sim\mathcal{D}_j} p_j^{2k-2}\right)^{1/2} && \text{(Cauchy-Schwarz).}\end{aligned} \tag{A35}$$

■

Lemma 2 formalizes the intuition that $\rho_{\mathcal{E}}^{(k)}$ and the proxy $\tilde{\rho}_{\mathcal{E}}^{(k)}$ are close, if the random variables p_j are close to q_j. Thus, we can conveniently replace p_j (which depends on the distribution $\mathcal{D}_j$), by the weight q_j (independent of the distribution $\mathcal{D}_j$), with an error that can be controlled.

The following lemma gives us control over the unnormalized projected ensemble, generated by states drawn from distributions that are ε-close in relative error.

Lemma 3. Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be pure-state ensembles that are ε-close in relative error up to the first $2k$ moments, for some $k\in\mathbb{N}$, i.e.,

$$(1-\varepsilon)\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}(|\Psi\rangle\langle\Psi|)^{\otimes 2k} \preccurlyeq \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}(|\Psi\rangle\langle\Psi|)^{\otimes 2k} \preccurlyeq (1+\varepsilon)\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}(|\Psi\rangle\langle\Psi|)^{\otimes 2k}, \tag{A36}$$

with $\varepsilon \ge 0$. Let $\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)$ be the kth moment of the unnormalized projected ensemble generated by the bipartite pure state $|\Psi\rangle_{AB}$,

$$\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi) = \sum_{z=1}^{D_B} q_z^{1-k}[(I_A\otimes\langle z|)|\Psi\rangle\langle\Psi|(I_A\otimes|z\rangle)]^{\otimes k}, \tag{A37}$$

for an arbitrary orthonormal basis $\{|z\rangle\}$ on subsystem B, and $q_z \ge 0$ are arbitrary non-negative weights independent of $|\Psi\rangle_{AB}$. Define

$$\Delta = \left|\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\right\|_2^2 - \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\right\|_2^2\right| \tag{A38}$$

for some arbitrary fixed positive operator M. Then, Δ satisfies the inequality

$$\Delta \le \varepsilon \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\left[2\,\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M)+\left(\sum_{z=1}^{D_B} q_z^{1-k}\langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle^k\right)^2\right]. \tag{A39}$$

Proof. Expanding the definition of Δ gives

$$\begin{aligned}\Delta &= \left|\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)^2) - \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)^2) - 2\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M) + 2\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M)\right|\\ &\le 2\left|\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M) - \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M)\right| + \left|\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)^2) - \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)^2)\right|.\end{aligned} \tag{A40}$$

Since $\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)$ can be regarded as a positive linear map acting on $(|\Psi\rangle\langle\Psi|)^{\otimes k}$, the positive-semidefinite ordering (A36) is preserved, i.e.,

$$(1-\varepsilon)\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi) \preccurlyeq \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi) \preccurlyeq (1+\varepsilon)\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi). \tag{A41}$$

For a positive operator M, this finding implies

$$(1-\varepsilon)\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}\left(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M\right) \preccurlyeq \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\mathrm{Tr}\left(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M\right) \preccurlyeq (1+\varepsilon)\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}\left(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M\right). \tag{A42}$$

Thus,

$$\left|\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M) - \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M)\right| \le \varepsilon \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}\left(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M\right). \tag{A43}$$

Next, we write

$$\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)^2) = \mathrm{Tr}\left[\mathcal{N}(|\Psi\rangle\langle\Psi|^{\otimes 2k})\hat{\tau}_A\right], \tag{A44}$$

where

$$\mathcal{N}(X) = \sum_{z,z'=1}^{D_B}(q_z q_{z'})^{1-k}\left(I_A^{\otimes 2k}\otimes\langle z^{\otimes k}z'^{\otimes k}|\right)X\left(I_A\otimes|z^{\otimes k}z'^{\otimes k}\rangle\right) \tag{A45}$$

is a positive linear map, and $\hat{\tau}_A$ is a permutation that swaps between the first and second k tensor factors of subsystem A and acts identically on subsystem B. Therefore,

$$\begin{aligned}\left|\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)^2) - \mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_2}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)^2)\right| &\le \varepsilon\,\mathrm{Tr}\,\mathcal{N}\left(\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}(|\Psi\rangle\langle\Psi|)^{\otimes 2k}\right)\\ &= \varepsilon\mathop{\mathbb{E}}_{\Psi\sim\mathcal{E}_1}\sum_{z,z'}(q_z q_{z'})^{1-k}\langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle^k\langle\Psi|(I_A\otimes|z'\rangle\langle z'|)|\Psi\rangle^k\\ &= \varepsilon\left(\sum_{z=1}^{D_B}q_z^{1-k}\langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle^k\right)^2.\end{aligned} \tag{A46}$$

Combining Eqs. (A43) and (A46) gives the desired bound for Δ. ∎

APPENDIX B: SCROOGE APPROXIMATION LEMMA

Let us define the ensemble of unnormalized states

$$\tilde{\mathrm{S}}\mathrm{crooge}(\sigma) = \left\{d\phi, \sqrt{D\sigma}|\phi\rangle\right\}, \tag{B1}$$

where $d\phi$ is the Haar measure on the unit sphere in $\mathbb{C}^D$ and σ is an arbitrary density matrix with dimension D. Note that $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$ is referred to as $\tilde{\mathcal{E}}$ in the main text. The kth moment of $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$ is given by

$$\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma) = \sqrt{D\sigma}^{\otimes k}\rho_{\mathrm{Haar}}^{(k)}\sqrt{D\sigma}^{\otimes k} = (D\sigma)^{\otimes k}\rho_{\mathrm{Haar}}^{(k)}, \tag{B2}$$

where in the final equality we used the fact that $\rho_{\mathrm{Haar}}^{(k)}$ commutes with $\sigma^{\otimes k}$. While the ensemble $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$ describes unnormalized states and is unphysical, it is a useful mathematical object in proving the main results of this paper. First, we note that in the low-purity regime, the states in $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$ are close to normalized.

Lemma 4 (Moment bounds). Let σ be a positive operator with unit trace. Denote the overlap between σ and $|\phi\rangle$ by $p_\sigma(\phi)\equiv\langle\phi|\sigma|\phi\rangle$, where $|\phi\rangle$ is a Haar random state. Assuming $\binom{k}{2}\|\sigma\|_2 < 1$, the kth moment of $p_\sigma(\phi)$ satisfies the inequality

$$1+\binom{k}{2}\|\sigma\|_2^2 \le k!D_k\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}[p_\sigma(\phi)^k] \le 1+\frac{\binom{k}{2}\|\sigma\|_2^2}{1-\binom{k}{2}\|\sigma\|_2}. \tag{B3}$$

Proof. Using Eq. (A8), we can write the kth moment of $p_\sigma(\phi)$ as

$$\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}[p_\sigma(\phi)^k] = \mathrm{Tr}\left(\sigma^{\otimes k}\rho^{(k)}_{\mathrm{Haar}}\right) = \frac{1}{k!D_k}\sum_{\pi\in S_k}\mathrm{Tr}(\sigma^{\otimes k}\hat{\pi}). \tag{B4}$$

Thus, the quantity we would like to bound is

$$k!D_k \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}[p_\sigma(\phi)^k] = \sum_{\pi\in S_k}\mathrm{Tr}(\sigma^{\otimes k}\hat{\pi}). \tag{B5}$$

First, note that if π is the identity permutation, $\mathrm{Tr}(\sigma^{\otimes k}\hat{\pi}) = [\mathrm{Tr}(\sigma)]^k = 1$, since σ has unit trace. Next, consider the case where π contains a single transposition, i.e., π has a Cayley distance of 1. There are $\binom{k}{2}$ such permutations, each contributing $\mathrm{Tr}(\sigma^{\otimes k}\hat{\pi}) = \mathrm{Tr}(\sigma^2) = \|\sigma\|_2^2$. All the remaining terms omitted are positive. Thus, we have the lower bound

$$\sum_{\pi\in S_k}\mathrm{Tr}(\sigma^{\otimes k}\hat{\pi}) \geq 1 + \binom{k}{2}\|\sigma\|_2^2. \tag{B6}$$

To prove the upper bound, let us examine the contribution from all permutations with a fixed Cayley distance $l \geq 1$. There are at most $\binom{k}{2}^l$ such permutations, and the contribution of each permutation can be upper bounded by $\mathrm{Tr}(\sigma^{\otimes k}\hat{\pi}) \leq \mathrm{Tr}(\sigma^{l+1})$. This bound is attained if π contains a cycle of length $l+1$. Thus, we have

$$\sum_{\pi\in S_k}\mathrm{Tr}(\sigma^{\otimes k}\hat{\pi}) \leq 1 + \binom{k}{2}\|\sigma\|_2^2 + \sum_{l=2}^{k-1}\binom{k}{2}^l\mathrm{Tr}(\sigma^{l+1}). \tag{B7}$$

Using the inequality

$$\mathrm{Tr}(\sigma^{l+1}) \leq [\mathrm{Tr}(\sigma^2)]^{\frac{l+1}{2}} = \|\sigma\|_2^{l+1}, \tag{B8}$$

we obtain

$$\begin{aligned}\sum_{\pi\in S_k}\mathrm{Tr}(\sigma^{\otimes k}\hat{\pi}) &\leq 1 + \|\sigma\|_2\sum_{l=1}^{k-1}\left[\binom{k}{2}\|\sigma\|_2\right]^l \\ &\leq 1 + \|\sigma\|_2\sum_{l=1}^{\infty}\left[\binom{k}{2}\|\sigma\|_2\right]^l \\ &= 1 + \frac{\binom{k}{2}\|\sigma\|_2^2}{1-\binom{k}{2}\|\sigma\|_2},\end{aligned} \tag{B9}$$

where the infinite series converges because we assumed $\binom{k}{2}\|\sigma\|_2 < 1$. ∎

The above lemma tells us that, up to corrections proportional to the purity of σ,

$$\|\sqrt{D\sigma}|\phi\rangle\|^2 = D\langle\phi|\sigma|\phi\rangle \approx 1, \tag{B10}$$

where $|\phi\rangle \sim \mathrm{Haar}(D)$.

The negative moments of $\langle\phi|\sigma|\phi\rangle$ can also be upper bounded in terms of $\|\sigma\|_\infty$, as shown in Ref. [62].

Lemma 5 (Lemma 2, Ref. [62]). Let σ be an arbitrary density operator with $m = \lfloor\|\sigma\|_\infty^{-1}\rfloor$. For $0 \leq q < m$,

$$\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\langle\phi|D\sigma|\phi\rangle^{-q} \leq \exp\left(\frac{q\left(q+1\right)}{2(m-q)}\right). \tag{B11}$$

For $q \ll m$, it simplifies to

$$\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\langle\phi|D\sigma|\phi\rangle^{-q} \leq 1 + \mathcal{O}\left(\frac{q^2}{m}\right). \tag{B12}$$

We now prove Lemma 1 from the main text, which we reproduce here, stated more formally.

Lemma 6 (Scrooge approximation). Consider the ensemble of unnormalized states $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma) = \{\sqrt{D\sigma}|\phi\rangle\}$, where $|\phi\rangle \sim \mathrm{Haar}(D)$, and σ is an arbitrary density matrix with dimension D. The kth moment of $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$ is $\tilde{\rho}^{(k)}_{\mathrm{Scrooge}}(\sigma) = (D\sigma)^{\otimes k}\rho^{(k)}_{\mathrm{Haar}}$. For any $k^2\|\sigma\|_2 \ll 1$, $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$ forms a $\mathrm{Scrooge}(\sigma)$ k-design with additive error $\mathcal{O}(k\|\sigma\|_2)$, i.e.,

$$\|\rho^{(k)}_{\mathrm{Scrooge}}(\sigma) - \tilde{\rho}^{(k)}_{\mathrm{Scrooge}}(\sigma)\|_1 \leq \mathcal{O}(k\|\sigma\|_2), \tag{B13}$$

and a $\mathrm{Scrooge}(\sigma)$ k-design with relative error $\mathcal{O}(4^k k\|\sigma\|_2)$, i.e.,

$$(1-\varepsilon)\tilde{\rho}^{(k)}_{\mathrm{Scrooge}}(\sigma) \preccurlyeq \rho^{(k)}_{\mathrm{Scrooge}}(\sigma) \preccurlyeq (1+\varepsilon)\tilde{\rho}^{(k)}_{\mathrm{Scrooge}}(\sigma), \tag{B14}$$

where $\varepsilon = \mathcal{O}(4^k k\|\sigma\|_2)$.

Proof of additive error $\mathcal{O}(k\|\sigma\|_2)$.

$$\begin{aligned}\left\|\rho^{(k)}_{\mathrm{Scrooge}}(\sigma) - \tilde{\rho}^{(k)}_{\mathrm{Scrooge}}(\sigma)\right\|_1 &= D\left\|\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}(\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma})^{\otimes k}\left(\frac{1}{\langle\phi|\sigma|\phi\rangle^{k-1}} - D^{k-1}\right)\right\|_1 \\ &\leq D\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\langle\phi|\sigma|\phi\rangle^k\left|\frac{1}{\langle\phi|\sigma|\phi\rangle^{k-1}} - D^{k-1}\right|,\end{aligned} \tag{B15}$$

where we used

$$\tilde{\rho}^{(k)}_{\text{Scrooge}}(\sigma) = D^k \sqrt{\sigma}^{\otimes k} \rho^{(k)}_{\text{Haar}} \sqrt{\sigma}^{\otimes k} = D^k \mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} (\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma})^{\otimes k} \tag{B16}$$

in the first line. Defining the random variable

$$X = D\langle\phi|\sigma|\phi\rangle, \tag{B17}$$

we have

$$\begin{aligned}\left\|\rho^{(k)}_{\text{Scrooge}}(\sigma) - \tilde{\rho}^{(k)}_{\text{Scrooge}}(\sigma)\right\|_1 &\le \mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} |X - X^k| \\ &= \mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} |(1-X)(X + X^2 + \cdots + X^{k-1})| \\ &\le \left[\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} (1-X)^2\right]^{1/2} \left[\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} (X + X^2 + \cdots + X^{k-1})^2\right]^{1/2}. \end{aligned} \tag{B18}$$

Explicitly evaluating the Haar averages gives

$$\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} X = 1 \tag{B19}$$

and

$$\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} X^2 = \frac{D}{D+1}(1 + \|\sigma\|_2^2), \tag{B20}$$

which yields

$$\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} (1-X)^2 = \frac{D\|\sigma\|_2^2 - 1}{D+1}. \tag{B21}$$

For $k^2\|\sigma\|_2 \ll 1$, Lemma 4 implies that

$$\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} X^k = 1 + \mathcal{O}(k^2\|\sigma\|_2^2), \tag{B22}$$

which leads to

$$\begin{aligned}&\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} (X + X^2 + \cdots + X^{k-1})^2 \\ &\le (k-1)^2\left(1 + \mathcal{O}(k^2\|\sigma\|_2^2)\right)\end{aligned} \tag{B23}$$

since the left-hand side can be expanded as a sum of $(k-1)^2$ terms, each contributing $1 + \mathcal{O}(k^2\|\sigma\|_2^2)$. Therefore,

$$\begin{aligned}&\left\|\rho^{(k)}_{\text{Scrooge}}(\sigma) - \tilde{\rho}^{(k)}_{\text{Scrooge}}(\sigma)\right\|_1 \\ &\le (k-1)\sqrt{\frac{D\|\sigma\|_2^2 - 1}{D+1}}(1 + \mathcal{O}(k^2\|\sigma\|_2^2)).\end{aligned} \tag{B24}$$

This result vanishes exactly when $k = 1$ or when σ is the maximally mixed state (i.e., Haar limit). In general, the error bound is $\mathcal{O}(k\|\sigma\|_2)$. ∎

Proof of relative error $\mathcal{O}(4^k k\|\sigma\|_2)$. It suffices to show that $\rho^{(k)}_{\text{Haar}}$ and

$$A^{(k)}_{\text{Scrooge}}(\sigma) = \mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} \frac{(|\phi\rangle\langle\phi|)^{\otimes k}}{\langle\phi|D\sigma|\phi\rangle^{k-1}} \tag{B25}$$

are ε-close in relative error since multiplying by $\sqrt{D\sigma}^{\otimes k}$ on both sides yields $\tilde{\rho}^{(k)}_{\text{Scrooge}}(\sigma)$ and $\rho^{(k)}_{\text{Scrooge}}(\sigma)$, respectively, which does not increase the relative error. Both $\rho^{(k)}_{\text{Haar}}$ and $A^{(k)}_{\text{Scrooge}}(\sigma)$ are supported on $\mathcal{H}^{(k)}_{\text{sym}}$, the symmetric subspace of $\mathcal{H}^{\otimes k}$. For any $|x\rangle \in \mathcal{H}^{(k)}_{\text{sym}}$,

$$\begin{aligned}\left|\langle x|\rho^{(k)}_{\text{Haar}} - A^{(k)}_{\text{Scrooge}}(\sigma)|x\rangle\right| &= \left|\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} \langle x|(|\phi\rangle\langle\phi|)^{\otimes k}|x\rangle \left(\frac{1}{\langle\phi|D\sigma|\phi\rangle^{k-1}} - 1\right)\right| \\ &= \left|\mathop{\mathbb{E}}_{\phi\sim\text{Haar}} \langle x|(|\phi\rangle\langle\phi|)^{\otimes k}|x\rangle \frac{1}{\langle\phi|D\sigma|\phi\rangle^{k-1}} (1 - \langle\phi|D\sigma|\phi\rangle^{k-1})\right| \\ &\le \mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)} \langle x|(|\phi\rangle\langle\phi|)^{\otimes k}|x\rangle \frac{1}{\langle\phi|D\sigma|\phi\rangle^{k-1}} \left|1 - \langle\phi|D\sigma|\phi\rangle^{k-1}\right|.\end{aligned} \tag{B26}$$

Applying Hölder's inequality,

$$\left|\langle x|\rho^{(k)}_{\mathrm{Haar}} - A^{(k)}_{\mathrm{Scrooge}}(\sigma)|x\rangle\right| \le \left(\mathbb{E}_\phi \langle x|(|\phi\rangle\langle\phi|)^{\otimes k}|x\rangle^4\right)^{1/4}\left(\mathbb{E}_\phi \langle\phi|D\sigma|\phi\rangle^{4(1-k)}\right)^{1/4}\left[\mathbb{E}_\phi(1-\langle\phi|D\sigma|\phi\rangle^{k-1})^2\right]^{1/2}. \tag{B27}$$

Now,

$$\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} \langle x|(|\phi\rangle\langle\phi|)^{\otimes k}|x\rangle^4 = \langle x^{\otimes 4}|\rho^{(4k)}_{\mathrm{Haar}}|x^{\otimes 4}\rangle \le \left\|\rho^{(4k)}_{\mathrm{Haar}}\right\|_\infty = \frac{1}{D_{4k}}. \tag{B28}$$

Next, using Lemma 5 [62], for $k \ll m$,

$$\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} \langle\phi|D\sigma|\phi\rangle^{4(1-k)} \le 1 + \mathcal{O}\left(\frac{k^2}{m}\right), \tag{B29}$$

where $m = \lfloor \|\sigma\|_\infty^{-1} \rfloor$. Finally, for $k^2\|\sigma\|_2 \ll 1$, we have

$$\mathbb{E}_\phi(1-\langle\phi|D\sigma|\phi\rangle^{k-1})^2 = \mathcal{O}(k^2\|\sigma\|_2^2). \tag{B30}$$

Note that $k^2\|\sigma\|_2 \ll 1$ implies $k \ll m$ since $\|\sigma\|_\infty^2 \le \|\sigma\|_2$. Combining these bounds, and using the identity $1 = D_k\langle x|\rho^{(k)}_{\mathrm{Haar}}|x\rangle$, we find

$$\left|\langle x|\rho^{(k)}_{\mathrm{Haar}} - A^{(k)}_{\mathrm{Scrooge}}|x\rangle\right| \le \mathcal{O}\left(\frac{k\|\sigma\|_2}{D_{4k}^{1/4}}\right) = \mathcal{O}\left(\frac{k\|\sigma\|_2 D_k}{D_{4k}^{1/4}}\right)\langle x|\rho^{(k)}_{\mathrm{Haar}}|x\rangle = \mathcal{O}(4^k k\|\sigma\|_2)\langle x|\rho^{(k)}_{\mathrm{Haar}}|x\rangle. \tag{B31}$$

Thus, the relative error is $\mathcal{O}(4^k k\|\sigma\|_2)$. ∎

We speculate that the prefactor of 4^k in the relative error can be improved. An analogous relative-error bound $\mathcal{O}(k^2\|\sigma\|_\infty^{1/2})$, which is controlled by $\|\sigma\|_\infty$ instead of $\|\sigma\|_2$, was proven in Ref. [62]. Our relative error in Lemma 1 has a better scaling with the norm of σ since $\|\sigma\|_2 \le \|\sigma\|_\infty^{1/2}$, but it has a worse scaling in k. In essence, the Scrooge approximation lemma (Lemma 1) tells us that

$$\rho^{(k)}_{\mathrm{Scrooge}}(\sigma) \approx (D\sigma)^{\otimes k}\rho^{(k)}_{\mathrm{Haar}}, \tag{B32}$$

for an arbitrary low-purity density matrix σ. Crucially, the right-hand side involves only the kth moment of the Haar ensemble, while $\rho^{(k)}_{\mathrm{Scrooge}}(\sigma)$ is a rational function of Haar random states, which enables us to analyze the scenarios where the Haar ensemble is replaced by a k-design, with the rigorous guarantee given by Lemma 1. We will make extensive use of this approach in the following appendixes to prove the main theorems in the paper.

APPENDIX C: EXAMPLES OF SCROOGE k-DESIGNS

In the main text, we assert that the random phase ensemble,

$$\mathcal{E}_{\text{random phase}} = \left\{\sum_{j=1}^{D} |\langle E_j|\Psi_0\rangle| e^{i\varphi_j}|E_j\rangle\right\}_\varphi, \tag{C1}$$

and canonical thermal pure quantum (cTPQ) states,

$$\mathcal{E}_{\text{cTPQ}} = \left\{\frac{1}{\mathcal{N}}\sum_{j=1}^{D}\xi_j e^{-\beta H/2}|j\rangle\right\}_\xi, \tag{C2}$$

form approximate Scrooge(σ) designs, with σ given by the diagonal state σ_{diag} and thermal Gibbs state $\sigma_\beta \propto \exp(-\beta H)$, respectively. These results can be derived using Lemma 1 by relating these ensembles to the proxy $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$, as we demonstrate in this appendix.

1. Random phase ensemble

Here, we prove Theorem 1 in the main text. Since the temporal ensemble obtained by late-time Hamiltonian dynamics is described by the random phase ensemble $\mathcal{E}_{\text{random phase}}$, assuming the Hamiltonian H satisfies the kth no-resonance condition [31], it suffices to show that the random phase ensemble indeed forms a Scrooge k-design.

Theorem 4 (Random phase ensembles form Scrooge k-designs). Let $\mathcal{E}_{\text{random phase}}$ be the random phase ensemble defined in Eq. (C1). Note that $\mathcal{E}_{\text{random phase}}$ forms a Scrooge(σ_{diag}) k-design with additive error

$$\varepsilon = \mathcal{O}(k\|\sigma_{\text{diag}}\|_2), \tag{C3}$$

for $k^2\|\sigma_{\text{diag}}\|_2 \ll 1$, where $\sigma_{\text{diag}} = \sum_j |\langle E_j|\Psi_0\rangle|^2 |E_j\rangle\langle E_j|$ is the diagonal state in the energy eigenbasis.

Proof. Let $r_i \equiv |\langle E_i|\Psi_0\rangle|^2$, so that $\sigma_{\text{diag}} = \sum_i r_i |E_i\rangle\langle E_i|$. Let $\{|n\rangle\}$ be the normalized occupation number basis with respect to the energy eigenbasis $\{|E_i\rangle\}$, i.e., $|n\rangle = |n_1, n_2, \ldots, n_D\rangle$, with n_i denoting the number of copies of $|E_i\rangle$ in the symmetric state $|n\rangle$. In this basis, the random phase ensemble has the kth moment

$$\mathop{\mathbb{E}}_{\psi\sim\text{random phase}} (|\psi\rangle\langle\psi|)^{\otimes k} = \sum_{|n|=k} k! \left(\prod_{i=1}^{D} \frac{r_i^{n_i}}{n_i!}\right) |n\rangle\langle n|, \tag{C4}$$

where the sum runs over all occupation number vectors $n = (n_1, n_2, \ldots, n_D)$ such that $\sum_{i=1}^{D} n_i = k$. Let us denote

$$G^{(k)}(\sigma_{\text{diag}}) = \sigma_{\text{diag}}^{\otimes k} \sum_{\pi\in S_k} \hat{\pi} = \sum_{|n|=k} k! \left(\prod_{i=1}^{D} r_i^{n_i}\right) |n\rangle\langle n|. \tag{C5}$$

Thus,

$$\left\| \mathop{\mathbb{E}}_{\psi\sim\text{random phase}} (|\psi\rangle\langle\psi|)^{\otimes k} - G^{(k)}(\sigma_{\text{diag}}) \right\|_1 = \sum_{|n|=k} k! \left(\prod_{i=1}^{D} \frac{r_i^{n_i}}{n_i!}\right) \left(\prod_{i=1}^{D} n_i! - 1\right). \tag{C6}$$

The sum over n can be split into two contributions. The first contribution comes from the "collision-free" sector, where $n_i \in \{0, 1\}$ for all i. This contribution is zero since $n_i! = 1$. The second contribution comes from the "collision" sector, where there exists at least one i such that $n_i \geq 2$. Therefore,

$$\begin{aligned} \left\| \mathop{\mathbb{E}}_{\psi\sim\text{random phase}} (|\psi\rangle\langle\psi|)^{\otimes k} - G^{(k)}(\sigma_{\text{diag}}) \right\|_1 &\leq \sum_{\substack{|n|=k \\ \text{collision}}} k! \prod_{i=1}^{D} r_i^{n_i} \\ &= \text{Tr}\left[G^{(k)}(\sigma_{\text{diag}})\right] - \sum_{\substack{|n|=k \\ \text{collision-free}}} k! \prod_{i=1}^{D} r_i^{n_i}. \end{aligned} \tag{C7}$$

Consider k independent draws from the distribution $\{r_i\}$. The probability of a collision (i.e., at least two draws with the same outcome) is given by (from the birthday paradox)

$$\Pr(\text{collision}) = 1 - \sum_{\substack{|n|=k \\ \text{collision-free}}} k! \prod_{i=1}^{D} r_i^{n_i} \leq \binom{k}{2} \sum_{i=1}^{D} r_i^2 = \mathcal{O}(k^2\|\sigma_{\text{diag}}\|_2^2). \tag{C8}$$

This formula yields

$$\left\| \mathop{\mathbb{E}}_{\psi\sim\text{random phase}} (|\psi\rangle\langle\psi|)^{\otimes k} - G^{(k)}(\sigma_{\text{diag}}) \right\|_1 \leq \text{Tr}\left[G^{(k)}(\sigma_{\text{diag}})\right] - 1 + \mathcal{O}\left(k^2\|\sigma_{\text{diag}}\|_2^2\right) = \mathcal{O}\left(k^2\|\sigma_{\text{diag}}\|_2^2\right), \tag{C9}$$

where in the last line we used the fact that $\text{Tr}[G^{(k)}(\sigma_{\text{diag}})] = 1 + \mathcal{O}(k^2\|\sigma_{\text{diag}}\|_2^2)$ in the regime $k^2\|\sigma_{\text{diag}}\|_2 \ll 1$, which can be shown using Lemma 4. Finally, using

$$\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma_{\text{diag}}) = \frac{D^k}{k! D_k} G^{(k)}(\sigma_{\text{diag}}) = \left[1 + \mathcal{O}\left(\frac{k^2}{D}\right)\right] G^{(k)}(\sigma_{\text{diag}}) \tag{C10}$$

and Lemma 1, together with the triangle inequality, we obtain

$$\left\| \mathop{\mathbb{E}}_{\psi\sim\text{random phase}} (|\psi\rangle\langle\psi|)^{\otimes k} - \rho_{\text{Scrooge}}^{(k)}(\sigma_{\text{diag}}) \right\|_1 \leq \mathcal{O}\left(\frac{k^2}{D}\right) + \mathcal{O}\left(k^2\|\sigma_{\text{diag}}\|_2^2\right) + \mathcal{O}\left(k\|\sigma_{\text{diag}}\|_2\right) = \mathcal{O}\left(k\|\sigma_{\text{diag}}\|_2\right), \tag{C11}$$

which concludes the proof. ∎

2. Canonical thermal pure quantum states

Next, we show that the cTPQ ensemble forms a Scrooge k-design in relative error, with respect to the density matrix

$$\sigma_\beta = \frac{e^{-\beta H}}{\mathrm{Tr}(e^{-\beta H})} \tag{C12}$$

for an arbitrary Hamiltonian H.

Theorem 5 (cTPQ states form Scrooge k-designs). Let $\mathcal{E}_{\mathrm{cTPQ}}$ be the ensemble of canonical thermal pure quantum states defined in Eq. (C2). For $k^2\|\sigma_\beta\|_2 \ll 1$, $\mathcal{E}_{\mathrm{cTPQ}}$ forms a Scrooge(σ_β) k-design with relative error ε satisfying

$$1+\varepsilon = [1+\mathcal{O}(4^k k\|\sigma_\beta\|_2)]^2, \tag{C13}$$

where $\sigma_\beta = \exp(-\beta H)/\mathrm{Tr}[\exp(-\beta H)]$. For a fixed $k\in\mathbb{N}$, $\varepsilon = \mathcal{O}(4^k k\|\sigma_\beta\|_2)$.

Proof. Define $|g\rangle = \sum_{j=1}^{D}\xi_j|j\rangle$, which is a random (unnormalized) vector of D independent zero-mean complex Gaussian variables with unit variance. Writing

$$|g\rangle = \||g\rangle\| \times |\phi\rangle, \tag{C14}$$

each Gaussian vector $|g\rangle$ is associated with a Haar random state $|\phi\rangle$. Note that $\||g\rangle\|$ and $|\phi\rangle$ are statistically independent. Thus, we can write the kth moment of $\mathcal{E}_{\mathrm{cTPQ}}$ as

$$\mathop{\mathbb{E}}_{\psi\sim\mathrm{cTPQ}}(|\psi\rangle\langle\psi|)^{\otimes k} = \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\frac{(\sqrt{\sigma_\beta}|\phi\rangle\langle\phi|\sqrt{\sigma_\beta})^{\otimes k}}{\langle\phi|\sigma_\beta|\phi\rangle^k}. \tag{C15}$$

For any $|x\rangle\in\mathcal{H}^{(k)}_{\mathrm{sym}}$,

$$\begin{aligned}
&\left|\langle x|\left(\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\frac{|\phi\rangle\langle\phi|^{\otimes k}}{\langle\phi|D\sigma_\beta|\phi\rangle^k} - \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}|\phi\rangle\langle\phi|^{\otimes k}\right)|x\rangle\right| \\
&= \left|\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\langle x|(|\phi\rangle\langle\phi|)^{\otimes k}|x\rangle\langle\phi|D\sigma_\beta|\phi\rangle^{-k}(1-\langle\phi|D\sigma_\beta|\phi\rangle^k)\right| \\
&\le \left(\mathbb{E}_\phi\langle x|(|\phi\rangle\langle\phi|)^{\otimes k}|x\rangle^4\right)^{1/4}\left(\mathbb{E}_\phi\langle\phi|D\sigma_\beta|\phi\rangle^{-4k}\right)^{1/4}\left[\mathbb{E}_\phi(1-\langle\phi|D\sigma_\beta|\phi\rangle^k)^2\right]^{1/2},
\end{aligned} \tag{C16}$$

where we used Hölder's inequality in the final line. Now,

$$\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\langle x|(|\phi\rangle\langle\phi|)^{\otimes k}|x\rangle^4 = \langle x^{\otimes 4}|\rho^{(4k)}_{\mathrm{Haar}}|x^{\otimes 4}\rangle \le \left\|\rho^{(4k)}_{\mathrm{Haar}}\right\|_\infty = \frac{1}{D_{4k}}. \tag{C17}$$

Next, using Lemma 5, for $k \ll m$,

$$\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\langle\phi|D\sigma_\beta|\phi\rangle^{-4k} \le 1+\mathcal{O}\left(\frac{k^2}{m}\right), \tag{C18}$$

where $m = \lfloor\|\sigma_\beta\|_\infty^{-1}\rfloor$. Note that our assumption $k^2\|\sigma_\beta\|_2 \ll 1$ already implies $k \ll m$ since $\|\sigma_\beta\|_\infty^2 \le \|\sigma_\beta\|_2$. Finally,

$$\mathbb{E}_\phi(1-\langle\phi|D\sigma_\beta|\phi\rangle^k)^2 = \mathcal{O}(k^2\|\sigma_\beta\|_2^2). \tag{C19}$$

Combining these bounds, and using the identity $1 = D_k\langle x|\rho^{(k)}_{\mathrm{Haar}}|x\rangle$, we find

$$\left|\langle x|\left(\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\frac{|\phi\rangle\langle\phi|^{\otimes k}}{\langle\phi|D\sigma_\beta|\phi\rangle^k} - \rho^{(k)}_{\mathrm{Haar}}\right)|x\rangle\right| \le \frac{1}{D_{4k}^{1/4}}\left(1+\mathcal{O}\left(\frac{k^2}{m}\right)\right)\mathcal{O}(k\|\sigma_\beta\|_2) = \mathcal{O}(4^k k\|\sigma_\beta\|_2)\langle x|\rho^{(k)}_{\mathrm{Haar}}|x\rangle. \tag{C20}$$

Thus,

$$(1-\varepsilon')\rho^{(k)}_{\mathrm{Haar}} \preccurlyeq \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)}\frac{|\phi\rangle\langle\phi|^{\otimes k}}{\langle\phi|D\sigma_\beta|\phi\rangle^k} \preccurlyeq (1+\varepsilon')\rho^{(k)}_{\mathrm{Haar}}, \tag{C21}$$

where $\varepsilon' = \mathcal{O}(4^k k\|\sigma_\beta\|_2)$. Multiplying both sides of the operator by $\sqrt{D\sigma_\beta}^{\otimes k}$ does not change the relative error, and we obtain

$$(1-\varepsilon')\tilde{\rho}^{(k)}_{\mathrm{Scrooge}}(\sigma_\beta) \preccurlyeq \mathop{\mathbb{E}}_{\psi\sim\mathrm{cTPQ}}(|\psi\rangle\langle\psi|)^{\otimes k} \preccurlyeq (1+\varepsilon')\tilde{\rho}^{(k)}_{\mathrm{Scrooge}}(\sigma_\beta). \tag{C22}$$

Recall from Lemma 1 that $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma_\beta)$ forms a $\mathrm{Scrooge}(\sigma_\beta)$ k-design with relative error $\varepsilon'' = \mathcal{O}(4^k k\|\sigma_\beta\|_2)$. Combining the relative errors, we obtain the desired result

$$(1-\varepsilon)\rho^{(k)}_{\mathrm{Scrooge}}(\sigma_\beta) \preccurlyeq \mathop{\mathbb{E}}_{\psi\sim\mathrm{cTPQ}}(|\psi\rangle\langle\psi|)^{\otimes k} \preccurlyeq (1+\varepsilon)\rho^{(k)}_{\mathrm{Scrooge}}(\sigma_\beta), \tag{C23}$$

where $1+\varepsilon = (1+\varepsilon')(1+\varepsilon'')$. ∎

If we further assume the low-purity regime $4^k k\|\sigma_\beta\|_2 \ll 1$, we find that the cTPQ ensemble is a $\mathrm{Scrooge}(\sigma_\beta)$ k-design with relative error

$$\varepsilon = \mathcal{O}(4^k k\|\sigma_\beta\|_2). \tag{C24}$$

APPENDIX D: PROJECTED ENSEMBLE GENERATED BY STATE DRAWN FROM SCROOGE 2k-DESIGN

In this appendix, we prove Theorem 2, one of the main results of the paper. Theorem 2 says that the projected ensemble generated by a state drawn from a Scrooge $2k$-design approximates a probabilistic mixture of Scrooge k-designs (for brevity, we will refer to such a mixture as a generalized Scrooge k-design). The proof of Theorem 2 is rather lengthy and proceeds in several steps, outlined below.

(1) First, we show that for the unnormalized generator state sampled from $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$ in Eq. (B1), the projected ensemble is locally close to a generalized Scrooge k-design (Lemma 7).
(2) Next, with the help of Lemma 7, we show that the projected ensemble generated by a global state drawn from an exact Scrooge $2k$-design forms a generalized Scrooge k-design (Theorem 6).
(3) Finally, we relax the assumption that the generator state is drawn from an exact $\mathrm{Scrooge}(\sigma)$ $2k$-design and instead consider the case where the generator state is drawn from an approximate $\mathrm{Scrooge}(\sigma)$ $2k$-design. Using the results of Theorem 6, we arrive at Theorem 2 in the main text, which we reproduce here (Theorem 7).

We then specialize our results to the infinite-temperature limit where the generator state is drawn from an approximate Haar $2k$-design, giving Corollary III B 1 in the main text. In this case, we can obtain an improved bound on the trace distance between the kth moments of the projected and Haar ensembles.

Lemma 7 (Projected ensemble generated by an unnormalized Scrooge state). Let $|\Psi\rangle_{AB}$ be an unnormalized state drawn from the ensemble $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$. Denote the reduced density matrix of σ on A and B by σ_A and σ_B, respectively. Consider the projected ensemble $\mathcal{E}(\Psi)$ obtained by applying projective measurements on B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. Then, assuming that $k^2\|\hat{\sigma}_{A|z}\|_2 \ll 1$ for all z, and $1 \ll D_A \le D_B$,

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}\left\|\rho^{(k)}_{\mathcal{E}} - \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho^{(k)}_{\mathrm{Scrooge}}(\hat{\sigma}_{A|z})\right\|_1 \le \mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2\right)}\right) + \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2), \tag{D1}$$

where $\sigma_{A|z} \equiv (I_A\otimes\langle z|)\sigma(I_A\otimes|z\rangle)$, and $\hat{\sigma}_{A|z} = \sigma_{A|z}/\langle z|\sigma_B|z\rangle$ is the normalized conditional mixed state on A.

Proof. For a given generator state $|\Psi\rangle$, the projected ensemble has the kth moment

$$\rho^{(k)}_{\mathcal{E}} = \sum_{z=1}^{D_B}\frac{(|\tilde{\psi}_z\rangle\langle\tilde{\psi}_z|)^{\otimes k}}{p_z^{k-1}}, \tag{D2}$$

where

$$|\tilde{\psi}_z\rangle = (I_A\otimes\langle z|)|\Psi\rangle \tag{D3}$$

is the unnormalized projected state, and

$$p_z = \langle\tilde{\psi}_z|\tilde{\psi}_z\rangle = \langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle \tag{D4}$$

is the measurement outcome probability. Let us construct the proxy

$$\tilde{\rho}^{(k)}_{\mathcal{E}} = \sum_{z=1}^{D_B}\frac{(|\tilde{\psi}_z\rangle\langle\tilde{\psi}_z|)^{\otimes k}}{\langle z|\sigma_B|z\rangle^{k-1}}, \tag{D5}$$

where the sum over z is restricted to measurement outcomes satisfying $\langle z|\sigma_B|z\rangle > 0$. Then, from Lemma 2, we have

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}\left\|\rho^{(k)}_{\mathcal{E}} - \tilde{\rho}^{(k)}_{\mathcal{E}}\right\|_1 \le \sum_{z=1}^{D_B}(\mathbb{E}_\Psi p_z^2)^{1/2}\left[1-2\frac{\mathbb{E}_\Psi p_z^{k-1}}{\langle z|\sigma_B|z\rangle^{k-1}}+\frac{\mathbb{E}_\Psi p_z^{2k-2}}{\langle z|\sigma_B|z\rangle^{2k-2}}\right]^{1/2}. \tag{D6}$$

Now,

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\text{crooge}(\sigma)} p_z^k = \frac{D^k}{k!D_k}\sum_{\pi\in S_k}\text{Tr}\left(\sigma_{A|z}^{\otimes k}\hat{\pi}_A\right) = \left[1+\mathcal{O}\left(\frac{k^2}{D}\right)\right]\langle z|\sigma_B|z\rangle^k\sum_{\pi\in S_k}\text{Tr}\left(\hat{\sigma}_{A|z}^{\otimes k}\hat{\pi}_A\right), \tag{D7}$$

using the fact that $\text{Tr}\sigma_{A|z} = \langle z|\sigma_B|z\rangle$ and $\hat{\sigma}_{A|z} = \sigma_{A|z}/\langle z|\sigma_B|z\rangle$. For $k^2\|\hat{\sigma}_{A|z}\|_2 \ll 1$, we can use Lemma 4 to obtain

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\text{crooge}(\sigma)} p_z^k = \langle z|\sigma_B|z\rangle^k\left[1+\mathcal{O}(k^2\|\hat{\sigma}_{A|z}\|_2^2)\right]. \tag{D8}$$

Substituting this result into the upper bound above gives

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\text{crooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)} - \tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1 \le \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2). \tag{D9}$$

Next, we seek to bound

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\text{crooge}(\sigma)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)} - \mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\text{crooge}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_2^2 = \underbrace{\mathbb{E}_\Psi\text{Tr}\left(\tilde{\rho}_{\mathcal{E}}^{(k)2}\right)}_{(*)} - \underbrace{\text{Tr}\left[\left(E_\Psi\tilde{\rho}_{\mathcal{E}}^{(k)}\right)^2\right]}_{(**)}. \tag{D10}$$

We first evaluate the second term $(**)$. Note that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\text{crooge}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)} &= D^k\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle^{1-k}\mathop{\mathbb{E}}_{\phi\sim\text{Haar}(D)}[(I_A\otimes\langle z|)\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma}(I_A\otimes|z\rangle)]^{\otimes k}\\
&= \frac{D^kD_{A,k}}{D_k}\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle^{1-k}\sigma_{A|z}^{\otimes k}\rho_{\text{Haar,A}}^{(k)}\\
&= \frac{D^kD_{A,k}}{D_k}\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\hat{\sigma}_{A|z}^{\otimes k}\rho_{\text{Haar,A}}^{(k)}\\
&= \left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right)\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\tilde{\rho}_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}),
\end{aligned} \tag{D11}$$

which can be interpreted as a mixture of approximate Scrooge$(\hat{\sigma}_{A|z})$ ensembles, weighted by the average probability $\langle z|\sigma_B|z\rangle$ of measuring the outcome z. Thus,

$$\begin{aligned}
(**) &= \left(\frac{D^kD_{A,k}}{D_k}\right)^2\sum_{z,z'=1}^{D_B}\langle z|\sigma_B|z\rangle\langle z'|\sigma_B|z'\rangle\text{Tr}\left(\hat{\sigma}_{A|z}^{\otimes k}\rho_{\text{Haar},A}^{(k)}\hat{\sigma}_{A|z'}^{\otimes k}\rho_{\text{Haar},A}^{(k)}\right)\\
&= \frac{D^{2k}D_{A,k}}{D_k^2}\sum_{z,z'=1}^{D_B}\langle z|\sigma_B|z\rangle\langle z'|\sigma_B|z'\rangle\text{Tr}\left(\hat{\sigma}_{A|z}^{\otimes k}\hat{\sigma}_{A|z'}^{\otimes k}\rho_{\text{Haar},A}^{(k)}\right),
\end{aligned} \tag{D12}$$

using the fact that $\rho_{\text{Haar},A}^{(k)2} = \rho_{\text{Haar},A}^{(k)}/D_{A,k}$. The first term in Eq. (D10), denoted $(*)$, can be evaluated to

$$(*) = D^{2k}\sum_{z,z'=1}^{D_B}\langle z|\sigma_B|z\rangle^{1-k}\langle z'|\sigma_B|z'\rangle^{1-k}\,\text{Tr}\left[(I_A^{\otimes 2k}\otimes|z'^kz^k\rangle\langle z^kz'^k|)\sigma^{\otimes 2k}\rho_{\text{Haar}}^{(2k)}\hat{\tau}\right]. \tag{D13}$$

Here, we defined the shorthand notation $|z^kz'^k\rangle \equiv |z\rangle^{\otimes k}\otimes|z'\rangle^{\otimes k}$. Note that $\hat{\tau}$ is a permutation operator acting on the $2k$-fold replica Hilbert space $\mathcal{H}^{\otimes 2k}$, which simply swaps the first and second k replicas, e.g., $\hat{\tau}(|\psi_1\rangle^{\otimes k}\otimes|\psi_2\rangle^{\otimes k}) = |\psi_2\rangle^{\otimes k}\otimes|\psi_1\rangle^{\otimes k}$. Now, expand $\rho_{\text{Haar}}^{(2k)}$ in terms of the permutation operators $\hat{\pi}$, where π is an element of the symmetric group S_{2k}, to obtain

$$(*) = \frac{D^{2k}}{(2k)!D_{2k}} \sum_{z,z'=1}^{D_B} \sum_{\pi \in S_{2k}} \langle z|\sigma_B|z\rangle^{1-k} \langle z'|\sigma_B|z'\rangle^{1-k} \mathrm{Tr}[(I_A^{\otimes 2k} \otimes |z'^k z^k\rangle\langle z^k z'^k|)\sigma^{\otimes 2k} \hat{\pi}\,\hat{\tau}]. \tag{D14}$$

The sum over $\pi \in S_{2k}$ can be split into sums over two classes of permutations: (i) permutations that can be decomposed in the form $\hat{\pi} = \hat{\pi}_1 \otimes \hat{\pi}_2$, where $\pi_1 \in S_k$ and $\pi_2 \in S_k$ are arbitrary permutations acting on the first and second k replicas, respectively, and (ii) permutations that cannot be decomposed in this form, which we denote by the shorthand $\pi \neq \pi_1\pi_2$. Thus, we find

$$\begin{aligned}
(*) &= \frac{D^{2k}}{(2k)!D_{2k}} \sum_{z,z'=1}^{D_B} \sum_{\pi_1,\pi_2 \in S_k} \langle z|\sigma_B|z\rangle^{1-k} \langle z'|\sigma_B|z'\rangle^{1-k} \mathrm{Tr}[(I_A^{\otimes 2k} \otimes |z'^k z^k\rangle\langle z^k z'^k|)\sigma^{\otimes 2k} (\hat{\pi}_1 \otimes \hat{\pi}_2)\hat{\tau}] \\
&\quad + \frac{D^{2k}}{(2k)!D_{2k}} \sum_{z,z'=1}^{D_B} \sum_{\substack{\pi \in S_{2k} \\ \pi \neq \pi_1\pi_2}} \langle z|\sigma_B|z\rangle^{1-k} \langle z'|\sigma_B|z'\rangle^{1-k} \mathrm{Tr}[(I_A^{\otimes 2k} \otimes |z'^k z^k\rangle\langle z^k z'^k|)\sigma^{\otimes 2k} \hat{\pi}\,\hat{\tau}] \\
&= \frac{D^{2k}}{(2k)!D_{2k}} \sum_{z,z'=1}^{D_B} \sum_{\pi_1,\pi_2 \in S_k} \langle z|\sigma_B|z\rangle\langle z'|\sigma_B|z'\rangle \mathrm{Tr}(\hat{\sigma}_{A|z}^{\otimes k} \hat{\pi}_1 \hat{\sigma}_{A|z'}^{\otimes k} \hat{\pi}_2) \\
&\quad + \frac{D^{2k}}{(2k)!D_{2k}} \sum_{z,z'=1}^{D_B} \sum_{\substack{\pi \in S_{2k} \\ \pi \neq \pi_1\pi_2}} \langle z|\sigma_B|z\rangle^{1-k} \langle z'|\sigma_B|z'\rangle^{1-k} \mathrm{Tr}[(I_A^{\otimes 2k} \otimes |z'^k z^k\rangle\langle z^k z'^k|)\sigma^{\otimes 2k} \hat{\pi}\,\hat{\tau}] \\
&= \frac{D^{2k}}{(2k)!D_{2k}} (k!)^2 D_{A,k} \sum_{z,z'=1}^{D_B} \langle z|\sigma_B|z\rangle\langle z'|\sigma_B|z'\rangle \mathrm{Tr}(\hat{\sigma}_{A|z}^{\otimes k} \hat{\sigma}_{A|z'}^{\otimes k} \rho_{\mathrm{Haar},A}^{(k)}) \\
&\quad + \frac{D^{2k}}{(2k)!D_{2k}} \sum_{z,z'=1}^{D_B} \sum_{\substack{\pi \in S_{2k} \\ \pi \neq \pi_1\pi_2}} \langle z|\sigma_B|z\rangle^{1-k} \langle z'|\sigma_B|z'\rangle^{1-k} \mathrm{Tr}[(I_A^{\otimes 2k} \otimes |z'^k z^k\rangle\langle z^k z'^k|)\sigma^{\otimes 2k} \hat{\pi}\,\hat{\tau}].
\end{aligned} \tag{D15}$$

Substituting the expressions for $(*)$ and $(**)$ into Eq. (D10) yields

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\Psi \sim \tilde{S}\mathrm{crooge}(\sigma)} \left\| \tilde{\rho}_{\mathcal{E}}^{(k)} - \mathop{\mathbb{E}}_{\Psi \sim \tilde{S}\mathrm{crooge}(\sigma)} \tilde{\rho}_{\mathcal{E}}^{(k)} \right\|_2^2 \\
&\leq \mathcal{O}\left(\frac{k^2}{D}\right) (k!)^2 D_{A,k} \sum_{z,z'=1}^{D_B} \langle z|\sigma_B|z\rangle\langle z'|\sigma_B|z'\rangle \mathrm{Tr}\left(\hat{\sigma}_{A|z}^{\otimes k} \hat{\sigma}_{A|z'}^{\otimes k} \rho_{\mathrm{Haar},A}^{(k)}\right) \\
&\quad + \frac{D^{2k}}{(2k)!D_{2k}} \sum_{z,z'=1}^{D_B} \sum_{\substack{\pi \in S_{2k} \\ \pi \neq \pi_1\pi_2}} \langle z|\sigma_B|z\rangle^{1-k} \langle z'|\sigma_B|z'\rangle^{1-k} \mathrm{Tr}[(I_A^{\otimes 2k} \otimes |z'^k z^k\rangle\langle z^k z'^k|)\sigma^{\otimes 2k} \hat{\pi}\,\hat{\tau}].
\end{aligned} \tag{D16}$$

To bound the first term on the right-hand side of Eq. (16), we use Hölder's inequality to obtain

$$\mathrm{Tr}\left(\hat{\sigma}_{A|z}^{\otimes k} \hat{\sigma}_{A|z'}^{\otimes k} \rho_{\mathrm{Haar},A}^{(k)}\right) \leq \frac{1}{D_{A,k}} \mathrm{Tr}^k(\hat{\sigma}_{A|z}) \mathrm{Tr}^k(\hat{\sigma}_{A|z'}) = \frac{1}{D_{A,k}}. \tag{D17}$$

To bound the second term, note that it must necessarily involve cross terms like $\sigma_{A|zz'} \equiv (I_A \otimes \langle z|)\sigma(I_A \otimes |z'\rangle)$ and $\sigma_{A|z'z} = \sigma_{A|zz'}^\dagger$, due to the constraint $\pi \neq \pi_1\pi_2$. Moreover, $\sigma_{A|zz'}$ and $\sigma_{A|z'z}$ appear in pairs. Thus, using Hölder's inequality, $\|\sigma_{A|zz'}\|_1^2 \leq \|\sigma_{A|z}\|_1 \|\sigma_{A|z'}\|_1$, and

$$\begin{aligned}
\sum_{\substack{\pi \in S_{2k} \\ \pi \neq \pi_1\pi_2}} \mathrm{Tr}[(I_A^{\otimes 2k} \otimes |z'^k z^k\rangle\langle z^k z'^k|)\sigma^{\otimes 2k} \hat{\pi}\,\hat{\tau}] &\leq ((2k)! - k!^2) |\mathrm{Tr}[(I_A^{\otimes 2k} \otimes |z'^k z^k\rangle\langle z^k z'^k|)\sigma^{\otimes 2k} \hat{\pi}\,\hat{\tau}]| \\
&= ((2k)! - k!^2) \|\sigma_{A|zz'}\|_1^2 \|\sigma_{A|z}\|_1^{k-1} \|\sigma_{A|z'}\|_1^{k-1} \\
&= ((2k)! - k!^2) \|\sigma_{A|zz'}\|_1^2 \langle z|\sigma_B|z\rangle^{k-1} \langle z'|\sigma_B|z'\rangle^{k-1}.
\end{aligned} \tag{D18}$$

Substituting these bounds into Eq. (16), we obtain

$$\begin{aligned}\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_2^2 &\le \mathcal{O}\left(\frac{k^{2k+2}}{D}\right)+\frac{D^{2k}}{(2k)!D_{2k}}((2k)!-k!^2)\sum_{z,z'=1}^{D_B}\|\sigma_{A|zz'}\|_1^2\\ &\le \mathcal{O}\left(\frac{k^{2k+2}}{D}\right)+\mathcal{O}(4^k k^{2k} D_A\|\sigma\|_2^2),\end{aligned} \tag{D19}$$

where we have used the inequality

$$\sum_{z,z'=1}^{D_B}\|\sigma_{A|zz'}\|_1^2 \le D_A\sum_{z,z'=1}^{D_B}\|\sigma_{A|zz'}\|_2^2 = D_A\|\sigma\|_2^2. \tag{D20}$$

Combining all the results above, we have an upper bound on the average trace distance,

$$\begin{aligned}\left(\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1\right)^2 &\le D_{A,k}\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_2^2\\ &\le \mathcal{O}\left(\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2\right)\right).\end{aligned} \tag{D21}$$

Combining Eqs. (D9) and (D21), and using the triangle inequality, we obtain

$$\begin{aligned}\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)}-\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1 &\le \mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1+\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1\\ &\le \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2)+\mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2\right)}\right),\end{aligned} \tag{D22}$$

where $\mathbb{E}_\Psi\tilde{\rho}_{\mathcal{E}}^{(k)}$ is given by Eq. (D11), which implies that

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)}-\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 \le \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2)+\mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2\right)}\right). \tag{D23}$$

To relate this finding to a mixture of $\rho_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})$, we use Lemma 1 and the triangle inequality to obtain

$$\begin{aligned}\left\|\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle(\rho_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})-\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}))\right\|_1 &\le \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\left\|\rho_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})-\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1\\ &= \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2). \quad \text{(Lemma 1)}\end{aligned} \tag{D24}$$

Finally, applying the triangle inequality again yields the desired result

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{S}\mathrm{crooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)}-\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 \le \mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2\right)}\right)+\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2). \tag{D25}$$

■

We will also make use of the following lemma.

Lemma 8. Let $\rho_{\mathcal{E}}^{(k)}(\Psi)$ be the kth moment of the projected ensemble $\mathcal{E}(\Psi)$ on subsystem A generated by $|\Psi\rangle_{AB}$, measured in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$ on subsystem B. For any fixed operator M, and any density matrix σ with dimension D,

$$\left| \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Scrooge}(\sigma)} \left\| \rho_{\mathcal{E}}^{(k)}(\Psi) - M \right\|_1 - \mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)} \left\| \rho_{\mathcal{E}}^{(k)}(\Psi) - M \right\|_1 \right| \leq \sqrt{\frac{D\|\sigma\|_2^2 - 1}{D+1}} \|M\|_1. \tag{D26}$$

Proof. Let us write $\rho_{\mathcal{E}}^{(k)}$ explicitly as

$$\rho_{\mathcal{E}}^{(k)} = \sum_{z=1}^{D_B} \frac{[(I_A \otimes \langle z|)|\Psi\rangle\langle\Psi|(I_A \otimes |z\rangle)]^{\otimes k}}{\langle\Psi|(I_A \otimes |z\rangle\langle z|)|\Psi\rangle^{k-1}}. \tag{D27}$$

By the definitions of the ensembles $\mathrm{Scrooge}(\sigma)$ and $\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)$, we have

$$\begin{aligned} \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Scrooge}(\sigma)} \|\rho_{\mathcal{E}}^{(k)} - M\|_1 &= D \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} \langle\phi|\sigma|\phi\rangle \left\| \frac{1}{\langle\phi|\sigma|\phi\rangle} \sum_{z=1}^{D_B} \frac{[(I_A \otimes \langle z|)\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma}(I_A \otimes |z\rangle)]^{\otimes k}}{\langle\phi|\sqrt{\sigma}(I_A \otimes |z\rangle\langle z|)\sqrt{\sigma}|\phi\rangle^{k-1}} - M \right\|_1 \\ &= \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} \left\| D \sum_{z=1}^{D_B} \frac{[(I_A \otimes \langle z|)\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma}(I_A \otimes |z\rangle)]^{\otimes k}}{\langle\phi|\sqrt{\sigma}(I_A \otimes |z\rangle\langle z|)\sqrt{\sigma}|\phi\rangle^{k-1}} - D\langle\phi|\sigma|\phi\rangle M \right\|_1, \end{aligned} \tag{D28}$$

and

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)} \|\rho_{\mathcal{E}}^{(k)} - M\|_1 = \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} \left\| D \sum_{z=1}^{D_B} \frac{[(I_A \otimes \langle z|)\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma}(I_A \otimes |z\rangle)]^{\otimes k}}{\langle\phi|\sqrt{\sigma}(I_A \otimes |z\rangle\langle z|)\sqrt{\sigma}|\phi\rangle^{k-1}} - M \right\|_1. \tag{D29}$$

For convenience, let us denote

$$K(\phi) \equiv D \sum_{z=1}^{D_B} \frac{[(I_A \otimes \langle z|)\sqrt{\sigma}|\phi\rangle\langle\phi|\sqrt{\sigma}(I_A \otimes |z\rangle)]^{\otimes k}}{\langle\phi|\sqrt{\sigma}(I_A \otimes |z\rangle\langle z|)\sqrt{\sigma}|\phi\rangle^{k-1}}. \tag{D30}$$

Therefore,

$$\begin{aligned} &\left| \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Scrooge}(\sigma)} \|\rho_{\mathcal{E}}^{(k)}(\Psi) - M\|_1 - \mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)} \|\rho_{\mathcal{E}}^{(k)}(\Psi) - M\|_1 \right| \\ &= \left| \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} \left(\|K(\phi) - D\langle\phi|\sigma|\phi\rangle M\|_1 - \|K(\phi) - M\|_1 \right) \right| \\ &\leq \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} \left| \|K(\phi) - D\langle\phi|\sigma|\phi\rangle M\|_1 - \|K(\phi) - M\|_1 \right| && \text{(Jensen's inequality)} \\ &\leq \mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} \|D\langle\phi|\sigma|\phi\rangle M - M\|_1 && \text{(reverse triangle inequality)} \\ &= \left(\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} |1 - D\langle\phi|\sigma|\phi\rangle| \right) \|M\|_1 \\ &\leq \left[\mathop{\mathbb{E}}_{\phi\sim\mathrm{Haar}(D)} (1 - D\langle\phi|\sigma|\phi\rangle)^2 \right]^{1/2} \|M\|_1 && \text{(Cauchy-Schwarz)} \\ &= \sqrt{\frac{D\|\sigma\|_2^2 - 1}{D+1}} \|M\|_1. \end{aligned} \tag{D31}$$

■

Now, using Lemmas 7 and 8, we can analyze the case where the generator state is drawn from the exact $\mathrm{Scrooge}(\sigma)$.

Theorem 6 (Projected ensemble generated by a Scrooge state). Let $|\Psi\rangle_{AB}$ be a state drawn from Scrooge(σ). Denote the reduced density matrices of σ on A and B by σ_A and σ_B, respectively. Consider the projected ensemble $\mathcal{E}(\Psi)$ obtained by applying projective measurements on B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. Then, assuming that $k^2\|\hat{\sigma}_{A|z}\|_2 \ll 1$ for all z, and $1 \ll D_A \leq D_B$,

$$\begin{aligned}&\mathop{\mathbb{E}}_{\Psi\sim\text{Scrooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)} - \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 \\ &\leq \mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2\right)}\right) + \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2) + \sqrt{\frac{D\|\sigma\|_2^2-1}{D+1}},\end{aligned} \tag{D32}$$

where $\sigma_{A|z} \equiv (I_A \otimes \langle z|)\sigma(I_A \otimes |z\rangle)$, and $\hat{\sigma}_{A|z} = \sigma_{A|z}/\langle z|\sigma_B|z\rangle$ is the normalized conditional mixed state on A.

Proof. Using Lemma 8 with

$$M = \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}), \tag{D33}$$

we have

$$\mathop{\mathbb{E}}_{\Psi\sim\text{Scrooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)} - M\right\|_1 \leq \mathop{\mathbb{E}}_{\Psi\sim\tilde{\text{S}}\text{crooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)} - M\right\|_1 + \sqrt{\frac{D\|\sigma\|_2^2-1}{D+1}}\|M\|_1. \tag{D34}$$

From Lemma 7, we obtain

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{\text{S}}\text{crooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)} - M\right\|_1 \leq \mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2\right)}\right) + \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2). \tag{D35}$$

We can also bound $\|M\|_1$ via

$$\|M\|_1 \leq \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\left\|\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 = \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle \tag{D36}$$

since $\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})$ is a normalized density operator. Therefore,

$$\begin{aligned}&\mathop{\mathbb{E}}_{\Psi\sim\text{Scrooge}(\sigma)}\left\|\rho_{\mathcal{E}}^{(k)} - \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 \\ &\leq \mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2\right)}\right) + \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2) + \sqrt{\frac{D\|\sigma\|_2^2-1}{D+1}}.\end{aligned} \tag{D37}$$

■

Theorem 7 (Projected ensemble generated by an approximate Scrooge 2k-design state). Let $|\Psi\rangle_{AB}$ be sampled from an approximate Scrooge(σ) $2k$-design, with relative error ε. Denote the reduced state of σ on A and B by σ_A and σ_B, respectively. Consider the projected ensemble $\mathcal{E}$ obtained by applying projective measurements on B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. Then, assuming that $k^2\|\hat{\sigma}_{A|z}\|_2 \ll 1$ for all z, and $1 \ll D_A \leq D_B$,

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1$$
$$\leq \mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2D_B\|\sigma\|_2^2+\frac{D\varepsilon}{k^2}+\frac{4^k}{k}D\|\sigma\|_2\right)}\right)+\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}\left(\sqrt{\varepsilon+4^kk\|\sigma\|_2+k^2\|\hat{\sigma}_{A|z}\|_2^2}\right), \tag{D38}$$

where $\sigma_{A|z}\equiv(I_A\otimes\langle z|)\sigma(I_A\otimes|z\rangle)$, and $\hat{\sigma}_{A|z}=\sigma_{A|z}/\langle z|\sigma_B|z\rangle$ is the normalized conditional mixed state on A.

Proof. From Lemma 1, the proxy $\tilde{\text{Scrooge}}(\sigma)$ is an approximate Scrooge(σ) $2k$-design with relative error $\varepsilon'=\mathcal{O}(4^kk\|\sigma\|_2)$. This implies that any approximate Scrooge(σ) $2k$-design with relative error ε must also be close to $\tilde{\text{Scrooge}}(\sigma)$ up to $2k$ moments with a relative error ε'' that satisfies

$$1+\varepsilon''=(1+\varepsilon)(1+\varepsilon'), \tag{D39}$$

which gives $\varepsilon''=\varepsilon+\varepsilon'+\varepsilon\varepsilon'=\mathcal{O}(\varepsilon+\varepsilon')=\mathcal{O}(\varepsilon+4^kk\|\sigma\|_2)$. In other words, we have

$$(1-\varepsilon'')\tilde{\rho}_{\text{Scrooge}}^{(2k)}(\sigma)\preccurlyeq\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}(|\Psi\rangle\langle\Psi|)^{\otimes 2k}\preccurlyeq(1+\varepsilon'')\tilde{\rho}_{\text{Scrooge}}^{(2k)}(\sigma). \tag{D40}$$

Now, using Lemma 3, where we choose $\mathcal{E}_1$ to be $\tilde{\text{Scrooge}}(\sigma)$, and $\mathcal{E}_2$ to be the approximate Scrooge(σ) $2k$-design, and the fixed operator

$$M=\mathop{\mathbb{E}}_{\Psi\sim\tilde{\text{Scrooge}}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)}=\left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right)\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\tilde{\rho}_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}), \tag{D41}$$

where we have defined

$$\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)=\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle^{1-k}[(I_A\otimes\langle z|)|\Psi\rangle\langle\Psi|(I_A\otimes|z\rangle)]^{\otimes k} \tag{D42}$$

for the projected ensemble $\mathcal{E}$ with $q_z=\langle z|\sigma_B|z\rangle$, we obtain

$$\begin{aligned}\Delta&=\left|\mathop{\mathbb{E}}_{\Psi\sim\tilde{\text{Scrooge}}(\sigma)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\right\|_2^2-\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\right\|_2^2\right|\\&\leq\varepsilon''\mathop{\mathbb{E}}_{\Psi\sim\tilde{\text{Scrooge}}(\sigma)}\left[2\,\text{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M)+\left(\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle^{1-k}\langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle^k\right)^2\right].\end{aligned} \tag{D43}$$

Rearranging, we obtain

$$\begin{aligned}\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\right\|_2^2&\leq\mathop{\mathbb{E}}_{\Psi\sim\tilde{\text{Scrooge}}(\sigma)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\right\|_2^2\\&\quad+\varepsilon''\mathop{\mathbb{E}}_{\Psi\sim\tilde{\text{Scrooge}}(\sigma)}\left[2\text{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M)+\left(\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle^{1-k}\langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle^k\right)^2\right].\end{aligned} \tag{D44}$$

The first term on the right-hand side is given by Eq. (D19) in the proof of Theorem 6 above, with

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{\text{Scrooge}}(\sigma)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\right\|_2^2\leq\mathcal{O}\left(\frac{k^{2k+2}}{D}\right)+\mathcal{O}(4^kk^{2k}D_A\|\sigma\|_2^2). \tag{D45}$$

Next, by linearity of the expectation and trace,

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}\mathrm{Tr}(\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)M)=\mathrm{Tr}\Big(\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)\Big)^2\le\frac{D^{2k}}{D_k^2}=\mathcal{O}(k!^2),\tag{D46}$$

which can be obtained from Eqs. (D12) and (D17). Furthermore, by the Cauchy-Schwarz inequality,

$$\begin{aligned}\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}\left(\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle^{1-k}\langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle^k\right)^2&\le\left[\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle^{1-k}\Big(\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}\langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle^{2k}\Big)^{1/2}\right]^2\\&=\left(1+\mathcal{O}\left(\frac{k^2}{D}\right)\right)\sum_{z,z'=1}^{D_B}\langle z|\sigma_B|z\rangle\langle z'|\sigma_B|z'\rangle\Big(1+\mathcal{O}(k^2\|\hat{\sigma}_{A|z}\|_2^2)\Big)\\&=1+o(1)\end{aligned}\tag{D47}$$

since we assumed $k^2\|\hat{\sigma}_{A|z}\|_2\ll 1$ for all z. The notation $o(1)$ is a shorthand for terms that vanish asymptotically in the limit of large Hilbert space dimensions. Thus,

$$\begin{aligned}\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\Big\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\Big\|_2^2&\le\mathcal{O}\left(\frac{k^{2k+2}}{D}\right)+\mathcal{O}\Big(4^k k^{2k}D_A\|\sigma\|_2^2\Big)+\mathcal{O}(\varepsilon''k^{2k})\\&=\mathcal{O}\left(\frac{k^{2k+2}}{D}\left(1+\frac{4^k}{k^2}{D_A}^2D_B\|\sigma\|_2^2+\frac{D\varepsilon''}{k^2}\right)\right).\end{aligned}\tag{D48}$$

This result implies the trace distance bound

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\Big\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\Big\|_1\le\sqrt{D_{A,k}}\Big(\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\Big\|\tilde{\rho}_{\mathcal{E}}^{(k)}(\Psi)-M\Big\|_2^2\Big)^{1/2}\le\mathcal{O}\left(\sqrt{\frac{k^{k+2}{D_A}^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2D_B\|\sigma\|_2^2+\frac{D\varepsilon''}{k^2}\right)}\right).\tag{D49}$$

The next step is to show that $\rho_{\mathcal{E}}^{(k)}$ is, with high probability, close to $\tilde{\rho}_{\mathcal{E}}^{(k)}$, when $|\Psi\rangle_{AB}$ is sampled from the Scrooge(σ) $2k$-design. To this end, we use Lemma 2 to write

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\Big\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\Big\|_1\le\sum_{z=1}^{D_B}\Big(\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}p_z^2\Big)^{1/2}\left[1-2\frac{\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}p_z^{k-1}}{\langle z|\sigma_B|z\rangle^{k-1}}+\frac{\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}p_z^{2k-2}}{\langle z|\sigma_B|z\rangle^{2k-2}}\right]^{1/2}.\tag{D50}$$

Using the relative error property,

$$(1-\varepsilon'')\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}{p_z}^{\ell}\le\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}{p_z}^{\ell}\le(1+\varepsilon'')\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}{p_z}^{\ell}\tag{D51}$$

for $1\le\ell\le 2k$. From Eq. (D8),

$$\mathop{\mathbb{E}}_{\Psi\sim\tilde{\mathrm{S}}\mathrm{crooge}(\sigma)}p_z^{\ell}=\langle z|\sigma_B|z\rangle^{\ell}\Big[1+\mathcal{O}\Big(\ell^2\|\hat{\sigma}_{A|z}\|_2^2\Big)\Big].\tag{D52}$$

Therefore,

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\Big\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\Big\|_1\le\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}\left(\sqrt{\varepsilon''+k^2\|\hat{\sigma}_{A|z}\|_2^2}\right).\tag{D53}$$

By the triangle inequality, and substituting the definition of M, we obtain

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 \le \mathcal{O}\left(\sqrt{\frac{k^{k+2}{D_A}^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2+\frac{D\varepsilon''}{k^2}\right)}\right) + \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}\left(\sqrt{\varepsilon''+k^2\|\hat{\sigma}_{A|z}\|_2^2}\right). \tag{D54}$$

In the above, we have replaced $\tilde{\rho}_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})$ by $\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})$ in M because this only incurs an additional trace distance error of $\mathcal{O}(\sum_z\langle z|\sigma_B|z\rangle k\|\hat{\sigma}_{A|z}\|_2)$ by the triangle inequality. It can be verified that this recovers Theorem 6 when $\varepsilon''=0$. As stated above, $\varepsilon''=\mathcal{O}(\varepsilon+4^k k\|\sigma\|_2)$; thus,

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 \le \mathcal{O}\left(\sqrt{\frac{k^{k+2}D_A^{k-1}}{D_B}\left(1+\frac{4^k}{k^2}{D_A}^2 D_B\|\sigma\|_2^2+\frac{D\varepsilon}{k^2}+\frac{4^k}{k}D\|\sigma\|_2\right)}\right) + \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}\left(\sqrt{\varepsilon+4^k k\|\sigma\|_2+k^2\|\hat{\sigma}_{A|z}\|_2^2}\right). \tag{D55}$$

■

For a fixed $k\in\mathbb{N}$, Eq. (D55) simplifies to

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 \le \mathcal{O}\left(\sqrt{{D_A}^k(\varepsilon+\|\sigma\|_2+D_A\|\sigma\|_2^2)}\right) \tag{D56}$$

to leading order, giving Theorem 2 in the main text.

1. Projected ensemble generated by a state drawn from a Haar 2k-design

In the special case where $\sigma=I/D$ is maximally mixed, the Scrooge ensemble reduces to the Haar ensemble, giving Corollary III B 1. In this limit, it turns out that the error bound can be improved compared to Theorem 2, for technical reasons. The proof follows similarly to Theorem 7 above.

Theorem 8 (Projected ensemble generated by an approximate 2k-design state). Let $|\Psi\rangle_{AB}$ be sampled from a $2k$-design, with relative error ε. Consider the projected ensemble $\mathcal{E}(\Psi)$ obtained by applying projective measurements on B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. Then, assuming $k^2\ll D_A$, with $D_A, D_B\gg 1$,

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\rho_{\text{Haar},A}^{(k)}\right\|_1 \le \sqrt{\frac{D_{A,k}}{D_B}+\mathcal{O}\left(\frac{k^2}{D_A}+D_{A,k}\varepsilon\right)}. \tag{D57}$$

Proof. The kth moment of the projected ensemble $\mathcal{E}(\Psi)$ can be written as

$$\rho_{\mathcal{E}}^{(k)}=\sum_{z=1}^{D_B}p_z^{1-k}[(I_A\otimes\langle z|)|\Psi\rangle\langle\Psi|(I_A\otimes|z\rangle)]^{\otimes k}, \tag{D58}$$

where

$$p_z=\langle\Psi|(I_A\otimes|z\rangle\langle z|)|\Psi\rangle \tag{D59}$$

is the probability of measuring the outcome z. Let us construct the proxy

$$\tilde{\rho}_{\mathcal{E}}^{(k)}=\sum_{z=1}^{D_B}D_B^{k-1}[(I_A\otimes\langle z|)|\Psi\rangle\langle\Psi|(I_A\otimes|z\rangle)]^{\otimes k}. \tag{D60}$$

The average trace distance between $\rho_{\mathcal{E}}^{(k)}$ and $\tilde{\rho}_{\mathcal{E}}^{(k)}$ can be bounded using Lemma 2,

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1 \le \sum_{z=1}^{D_B}\left(\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}p_z^2\right)^{1/2}\left(1-2D_B^{k-1}\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}p_z^{k-1}+D_B^{2k-2}\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}p_z^{2k-2}\right)^{1/2}. \tag{D61}$$

For any $1 \le \ell \le 2k$,

$$(1-\varepsilon)\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} {p_z}^{\ell} \le \mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}} {p_z}^{\ell} \le (1+\varepsilon)\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} {p_z}^{\ell}, \tag{D62}$$

from the definition of the relative error ε. The Haar average of ${p_z}^{\ell}$ can be explicitly evaluated as

$$\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} {p_z}^{\ell} = \frac{1}{\ell! D_\ell}\sum_{\pi\in S_\ell} {D_A}^{\#\mathrm{cycles}(\pi)} = \frac{D_{A,\ell}}{D_\ell} = \frac{1}{D_B^\ell}\left(1+\mathcal{O}\left(\frac{\ell^2}{D_A}\right)\right). \tag{D63}$$

Thus,

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}} \left\|\rho_{\mathcal{E}}^{(k)} - \tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1 = \mathcal{O}\left(\sqrt{\varepsilon + \frac{k^2}{D_A}}\right). \tag{D64}$$

Next, using Lemma 3 with the choice of $q_z = 1/D_B$, $M = \rho_{\mathrm{Haar},A}^{(k)}$,

$$\begin{aligned}\Delta \equiv& \left|\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}} \left\|\tilde{\rho}_{\mathcal{E}}^{(k)} - \rho_{\mathrm{Haar,A}}^{(k)}\right\|_2^2 - \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} \left\|\tilde{\rho}_{\mathcal{E}}^{(k)} - \rho_{\mathrm{Haar,A}}^{(k)}\right\|_2^2\right| \\ &\le 2\varepsilon \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} \mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)}\rho_{\mathrm{Haar,A}}^{(k)} + \varepsilon D_B^{2k-2} \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} \left(\sum_{z=1}^{D_B} p_z^k\right)^2.\end{aligned} \tag{D65}$$

The first term can be evaluated using

$$\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} \mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)}\rho_{\mathrm{Haar,A}}^{(k)} = \frac{1}{D_{A,k}} \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} \mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)} = \frac{D_B^{k-1}}{D_{A,k}} \sum_{z=1}^{D_B} \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} p_z^k = \frac{1}{D_{A,k}}\left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right). \tag{D66}$$

The second term can be evaluated using

$$\begin{aligned} D_B^{2k-2} \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} \left(\sum_{z=1}^{D_B} p_z^k\right)^2 &= D_B^{2k-2} \sum_{z,z'=1}^{D_B} (I_A^{\otimes 2k} \otimes \langle z^k z'^k|)\rho_{\mathrm{Haar}}^{(2k)}(I_A^{\otimes 2k} \otimes |z^k z'^k\rangle) \\ &= \frac{D_B^{2k-2}}{D^{2k}}\left(1+\mathcal{O}\left(\frac{k^2}{D}\right)\right) \sum_{z,z'=1}^{D_B} \sum_{\pi\in S_{2k}} {D_A}^{\#\mathrm{cycles}(\pi)} \langle z^k z'^k|\hat{\pi}_B|z^k z'^k\rangle \\ &= \frac{1}{D_B^2 D_A^{2k}}\left(1+\mathcal{O}\left(\frac{k^2}{D}\right)\right)\left[\sum_{z=1}^{D_B}\sum_{\pi\in S_{2k}} {D_A}^{\#\mathrm{cycles}(\pi)} + \sum_{\substack{z,z'=1 \\ z\neq z'}}^{D_B}\left(\sum_{\pi\in S_k} {D_A}^{\#\mathrm{cycles}(\pi)}\right)^2\right] \\ &= \frac{1}{D_B^2 D_A^{2k}}\left(1+\mathcal{O}\left(\frac{k^2}{D}\right)\right)(D_B D_A^{2k} + D_B(D_B-1)D_A^{2k})\left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right) \\ &= 1+\mathcal{O}\left(\frac{k^2}{D_A}\right). \end{aligned} \tag{D67}$$

Therefore, $\Delta = \mathcal{O}(\varepsilon)$, and

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}} \left\|\tilde{\rho}_{\mathcal{E}}^{(k)} - \rho_{\mathrm{Haar,A}}^{(k)}\right\|_2^2 \le \mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)} \left\|\tilde{\rho}_{\mathcal{E}}^{(k)} - \rho_{\mathrm{Haar,A}}^{(k)}\right\|_2^2 + \mathcal{O}(\varepsilon). \tag{D68}$$

Now,

$$\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\rho_{\mathrm{Haar,A}}^{(k)}\right\|_2^2=\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)}\mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)2}-2\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)}\mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)}\rho_{\mathrm{Haar},A}^{(k)}+\frac{1}{D_{A,k}}$$
$$=\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)}\mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)2}-\frac{2}{D_{A,k}}\left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right)+\frac{1}{D_{A,k}},\tag{D69}$$

where

$$\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)}\mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)2}=D_B^{2k-2}\sum_{z,z'=1}^{D_B}\left(I_A^{\otimes 2k}\otimes\langle z^k z'^k|\right)\rho_{\mathrm{Haar}}^{(2k)}\left(I_A^{\otimes 2k}\otimes|z^k z'^k\rangle\right)$$
$$=\frac{1}{D_B^2D_A^{2k}}\left(1+\mathcal{O}\left(\frac{k^2}{D}\right)\right)\sum_{z,z'=1}^{D_B}\sum_{\pi\in S_{2k}}D_A^{\#\mathrm{cycles}(\pi)}\langle z^k z'^k|\hat{\pi}_B|z'^k z^k\rangle$$
$$=\frac{1}{D_B^2D_A^{2k}}\left(1+\mathcal{O}\left(\frac{k^2}{D}\right)\right)\left[D_BD_A^{2k}\left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right)+D_B(D_B-1)k!D_A^k\left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right)\right]$$
$$=\frac{1}{D_{A,k}}\left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right)+\frac{1}{D_B}.\tag{D70}$$

Thus,

$$\mathop{\mathbb{E}}_{\Psi\sim\mathrm{Haar}(D)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\rho_{\mathrm{Haar,A}}^{(k)}\right\|_2^2=\frac{1}{D_B}+\mathcal{O}\left(\frac{k^2}{D_AD_{A,k}}\right)\tag{D71}$$

and

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\rho_{\mathrm{Haar,A}}^{(k)}\right\|_2^2\le\frac{1}{D_B}+\mathcal{O}\left(\frac{k^2}{D_AD_{A,k}}+\varepsilon\right),\tag{D72}$$

which implies that

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\rho_{\mathrm{Haar,A}}^{(k)}\right\|_1\le\left(D_{A,k}\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\rho_{\mathrm{Haar,A}}^{(k)}\right\|_2^2\right)^{1/2}\le\sqrt{\frac{D_{A,k}}{D_B}+\mathcal{O}\left(\frac{k^2}{D_A}+D_{A,k}\varepsilon\right)}.\tag{D73}$$

Finally, by the triangle inequality,

$$\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\rho_{\mathrm{Haar,A}}^{(k)}\right\|_1\le\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\rho_{\mathrm{Haar,A}}^{(k)}\right\|_1+\mathop{\mathbb{E}}_{\Psi\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\rho_{\mathcal{E}}^{(k)}\right\|_1\le\sqrt{\frac{D_{A,k}}{D_B}+\mathcal{O}\left(\frac{k^2}{D_A}+D_{A,k}\varepsilon\right)},\tag{D74}$$

which yields the desired result. ∎

2. Late-time chaotic Hamiltonian dynamics

Consider a generator state $|\Psi\rangle_{AB}$ obtained by evolving the initial state $|\Psi_0\rangle_{AB}$ under an ergodic Hamiltonian H, for a late-time t. Generically, it is reasonable to expect $|\Psi\rangle_{AB}$ to be modeled by the random phase ensemble

$$\mathcal{E}_{\text{random phase}}=\left\{\frac{d^D\varphi}{(2\pi)^D},\sum_{j=1}^{D}|\langle E_j|\Psi_0\rangle|e^{i\varphi_j}|E_j\rangle\right\},\tag{D75}$$

with the diagonal density matrix (in the energy basis $\{|E_j\rangle\}_j$)

$$\sigma_{\text{diag}} = \sum_{j=1}^{D} |\langle E_j|\Psi_0\rangle|^2 |E_j\rangle\langle E_j|, \tag{D76}$$

as described in the main text. If H satisfies the kth no-resonance condition, Mark *et al.* [31] showed that the temporal ensemble

$$\mathcal{E}_{\text{temp}} = \{e^{-iHt}|\Psi_0\rangle | t \in (-\infty, \infty)\} \tag{D77}$$

is close to the random phase ensemble, up to the kth moment. While we have shown that the random phase ensemble forms an approximate Scrooge k-design (Theorem 1 in the main text), Theorem 2 does not strictly apply in this scenario since the random phase ensemble is not close to Scrooge(σ) in relative error. Nonetheless, it is interesting to ask whether the projected ensemble generated by a state drawn from the random phase ensemble forms a generalized Scrooge ensemble, as in Theorem 2. This problem is partially resolved with Theorem 4 of Ref. [31], where they showed that the kth moment of the unnormalized projected ensemble $\mathcal{E}(\Psi)$ is approximately that of the unnormalized generalized Scrooge ensemble, i.e.,

$$\tilde{\rho}_{\mathcal{E}}^{(k)} \approx \sum_{z=1}^{D_B} \langle z|\sigma_B|z\rangle \tilde{\rho}_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}), \tag{D78}$$

with an error that is expected to be exponentially small in the system size. In the above, σ_A and σ_B are the reduced density matrices of σ_{diag} on A and B, respectively. Note that $\sigma_{A|z} \equiv (I_A \otimes \langle z|)\sigma(I_A \otimes |z\rangle)$, and $\hat{\sigma}_{A|z} = \sigma_{A|z}/\langle z|\sigma_B|z\rangle$ is the normalized conditional mixed state on A. Using the technical results developed in proving Theorem 2, we close this conceptual gap by showing that the Scrooge behavior does hold true for the actual normalized projected ensemble, i.e.,

$$\rho_{\mathcal{E}}^{(k)} \approx \sum_{z=1}^{D_B} \langle z|\sigma_B|z\rangle \rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}). \tag{D79}$$

This case is stated informally in Proposition 1 of the main text, which we reproduce here, stated more formally.

Theorem 9 (Emergent Scrooge designs from late-time generator states). Let $|\Psi\rangle_{AB}$ be the generator state drawn from the random phase ensemble (D75). Denote the reduced state of σ_{diag} on A and B by σ_A and σ_B, respectively, where σ_{diag} is the diagonal ensemble given in Eq. (D76). Consider the projected ensemble $\mathcal{E}(\Psi)$ obtained by applying projective measurements on B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. Then, assuming $k^2\|\hat{\sigma}_{A|z}\|_2 \ll 1$,

$$\mathbb{E}_\Psi \left\| \rho_{\mathcal{E}}^{(k)} - \sum_{z=1}^{D_B} \langle z|\sigma_B|z\rangle \rho_{\text{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}) \right\|_1 \leq \sum_{z=1}^{D_B} \langle z|\sigma_B|z\rangle \mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2) + \mathcal{O}(\Delta_\beta^{1/2}), \tag{D80}$$

where $\sigma_{A|z} \equiv (I_A \otimes \langle z|)\sigma(I_A \otimes |z\rangle)$, and $\hat{\sigma}_{A|z} = \sigma_{A|z}/\langle z|\sigma_B|z\rangle$ is the normalized conditional mixed state on A. Above,

$$\Delta_\beta = \sum_{z=1}^{D_B} \frac{\langle z^{\otimes 2}|\text{Tr}_A(\sigma_{\text{diag}}^{(2)})|z^{\otimes 2}\rangle}{\langle z|\sigma_B|z\rangle}, \tag{D81}$$

where

$$\sigma_{\text{diag}}^{(2)} = \sum_j |\langle E_j|\Psi_0\rangle|^4 |E_j^{\otimes 2}\rangle\langle E_j^{\otimes 2}|. \tag{D82}$$

Proof. We first bound the average trace distance between $\tilde{\rho}_{\mathcal{E}}^{(k)}$ and $\rho_{\mathcal{E}}^{(k)}$. Define $X = p_z/\mathbb{E}_\varphi p_z$, where $\mathbb{E}_\varphi$ indicates the averaging over the uniformly random phases in Eq. (C1), equivalent to averaging over the generator states $\mathbb{E}_\Psi$. Averaging over the random phase ensemble, we find

$$\begin{aligned}
\mathbb{E}_\varphi \|\tilde{\rho}_{\mathcal{E}}^{(k)} - \rho_{\mathcal{E}}^{(k)}\|_1 &\leq \mathbb{E}_\varphi \sum_{z=1}^{D_B} p_z |1 - X^{k-1}| \\
&= \mathbb{E}_\varphi \sum_{z=1}^{D_B} |(1-X)(1 + X + \cdots + X^{k-2})| \\
&\leq \sum_{z=1}^{D_B} \langle z|\sigma_B|z\rangle \left(\mathbb{E}_\varphi (1-X)^2\right)^{1/2} \left(\mathbb{E}_\varphi \left(X + X^2 + \ldots + X^{k-1}\right)^2\right)^{1/2}.
\end{aligned} \tag{D83}$$

Using

$$\begin{aligned}\mathbb{E}_\varphi X^k &\le \langle z|\sigma_B|z\rangle^{-k}\sum_{\pi\in S_k}\mathrm{Tr}[\sigma^{\otimes k}(I_A\otimes|z\rangle\langle z|)^{\otimes k}\hat{\pi}]\\ &= \langle z|\sigma_B|z\rangle^{-k}\sum_{\pi\in S_k}\mathrm{Tr}\left(\sigma_{A|z}^{\otimes k}\hat{\pi}_A\right)\\ &= \sum_{\pi\in S_k}\mathrm{Tr}\left(\hat{\sigma}_{A|z}^{\otimes k}\hat{\pi}_A\right)\\ &= 1+\mathcal{O}\left(k^2\|\hat{\sigma}_{A|z}\|_2^2\right) \qquad \text{(Lemma 4)},\end{aligned} \tag{D84}$$

we have

$$\mathbb{E}_\varphi\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\rho_{\mathcal{E}}^{(k)}\right\|_1 \le \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2). \tag{D85}$$

This case implies that $\tilde{\rho}_{\mathcal{E}}^{(k)}\approx\rho_{\mathcal{E}}^{(k)}$ in the low-purity regime. Similarly, from Lemma 1, we have

$$\left\|\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle(\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})-\rho_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z}))\right\|_1 \le \sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2). \tag{D86}$$

Thus, in the low-purity regime, we have

$$\mathbb{E}_\varphi\left\|\rho_{\mathcal{E}}^{(k)}-\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\rho_{\mathrm{Scrooge}}^{(k)}(\hat{\sigma}_{A|z})\right\|_1 \le \mathcal{O}(\Delta_\beta^{1/2})+\sum_{z=1}^{D_B}\langle z|\sigma_B|z\rangle\mathcal{O}(k\|\hat{\sigma}_{A|z}\|_2), \tag{D87}$$

where Δ_β is the error term defined in Ref. [31], which is argued to be exponentially small in system size in typical many-body systems. The subscript β indicates that $|\Psi\rangle$ has an effective temperature β^{-1}. ∎

APPENDIX E: PROJECTED ENSEMBLE GENERATED BY MEASUREMENTS IN A SCRAMBLED BASIS

In this appendix, we prove Theorem 3 in the main text. We first prove the following lemmas, which will be useful in proving Theorem 3.

Lemma 9 (kth moment of unnormalized projected ensemble). Let $|\Psi\rangle_{AB}$ be an arbitrary bipartite state, with subsystem density operators denoted σ_A and σ_B, respectively. Consider the projected ensemble $\mathcal{E}$ obtained by applying a Haar random unitary U on subsystem B, followed by projective measurements on B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. The average kth moment of the unnormalized state $|\tilde{\psi}_z\rangle=(I_A\otimes\langle z|U)|\Psi\rangle$ is given by

$$\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}(|\tilde{\psi}_z\rangle\langle\tilde{\psi}_z|)^{\otimes k}=\frac{D_{A,k}}{D_{B,k}}\sigma_A^{\otimes k}\rho_{\mathrm{Haar},A}^{(k)}. \tag{E1}$$

Proof. Let us write the Schmidt decomposition of $|\Psi\rangle_{AB}$ as

$$|\Psi\rangle_{AB}=\sum_i\sqrt{\lambda_i}\,|i\rangle_A\otimes|i\rangle_B, \tag{E2}$$

where λ_i are the Schmidt coefficients and $|i\rangle$ are the corresponding Schmidt vectors. Then, we have

$$\begin{aligned}\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}(|\tilde{\psi}_z\rangle\langle\tilde{\psi}_z|)^{\otimes k} &= \sum_{\substack{i_1,\ldots,i_k\\ j_1,\ldots,j_k}}\left(\prod_{\alpha=1}^k\lambda_{i_\alpha}\lambda_{j_\alpha}\right)^{1/2}(|i_1,\ldots,i_k\rangle\langle j_1,\ldots,j_k|)_A\langle j_1,\ldots,j_k|\rho_{\mathrm{Haar},B}^{(k)}|i_1,\ldots,i_k\rangle_B\\ &= \frac{1}{k!D_{B,k}}\sum_{\substack{i_1,\ldots,i_k\\ j_1,\ldots,j_k}}\sum_{\pi\in S_k}\left(\prod_{\alpha=1}^k\lambda_{i_\alpha}\lambda_{j_\alpha}\right)^{1/2}(|i_1,\ldots,i_k\rangle\langle j_1,\ldots,j_k|)_A\langle j_1,\ldots,j_k|\hat{\pi}_B|i_1,\ldots,i_k\rangle_B.\end{aligned} \tag{E3}$$

By orthonormality of the Schmidt vectors, the term $\langle j_1,\ldots,j_k|\hat{\pi}_B|i_1,\ldots,i_k\rangle_B = \delta_{j_1,\pi(i_1)}\ldots\delta_{j_k,\pi(i_k)}$ enforces the constraints on $j_1,\ldots,j_k$, which simplifies the sum to

$$\begin{aligned}\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}(|\tilde{\psi}_z\rangle\langle\tilde{\psi}_z|)^{\otimes k} &= \frac{1}{k!D_{B,k}}\sum_{i_1,\ldots,i_k}\sum_{\pi\in S_k}\left(\prod_{\alpha=1}^{k}\lambda_{i_\alpha}\right)(|i_1,\ldots,i_k\rangle\langle i_1,\ldots,i_k|)_A\hat{\pi}_A^\dagger \\ &= \frac{D_{A,k}}{D_{B,k}}\sigma_A^{\otimes k}\rho_{\mathrm{Haar},A}^{(k)},\end{aligned} \tag{E4}$$

where we used $\sigma_A = \sum_i \lambda_i(|i\rangle\langle i|)_A$. ∎

Lemma 10 (Average k-copy overlap between projected states). Let $|\Psi\rangle_{AB}$ be an arbitrary bipartite state, with subsystem density operators denoted σ_A and σ_B, respectively. Denote the unnormalized projected state as $|\tilde{\psi}_z\rangle = (I_A\otimes\langle z|U)|\Psi\rangle$. Assume $k^2 \ll D_B$ and $k\|\sigma_A\|_4 \ll \|\sigma_A\|_2$. Then, for any pair of measurement outcomes $z\neq z'$,

$$\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}|\langle\tilde{\psi}_z|\tilde{\psi}_{z'}\rangle|^{2k} \le \frac{k!\|\sigma_A\|_2^{2k}}{D_B^{2k}}\left[1+\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)\right]+\mathcal{O}\left(\frac{k^{2k+2}}{D_B^{2k+1}}\right). \tag{E5}$$

Proof. Using Weingarten calculus, we obtain

$$\begin{aligned}\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}|\langle\tilde{\psi}_z|\tilde{\psi}_{z'}\rangle|^{2k} &= \mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}\mathrm{Tr}\left(\sigma_B^{\otimes 2k}U_B^{\dagger\otimes 2k}|z^{\otimes k}z'^{\otimes k}\rangle\langle z'^{\otimes k}z^{\otimes k}|U_B^{\otimes 2k}\right) \\ &= \sum_{\pi,\chi\in S_{2k}}\mathrm{Wg}(\chi^{-1}\pi,D_B)\mathrm{Tr}\left(\sigma_B^{\otimes 2k}\hat{\pi}_B^\dagger\right)\langle z'^{\otimes k}z^{\otimes k}|\hat{\chi}_B|z^{\otimes k}z'^{\otimes k}\rangle\end{aligned} \tag{E6}$$

in terms of permutation operators $\hat{\pi}_B$ and $\hat{\chi}_B$, and the Weingarten function $\mathrm{Wg}(\chi^{-1}\pi,D_B)$. To proceed, observe that, by orthogonality of $|z\rangle$ and $|z'\rangle$, the nonvanishing contributions to the Weingarten sum are those with χ of the form $\chi=\chi_1\chi_2\tau$, where $\chi_1,\chi_2\in S_k$ are arbitrary permutations acting on the first and second sets of k-copy replicas, and $\tau\in S_{2k}$ is a fixed permutation that swaps between the first and second sets of k-copy replicas. This approach will contract every $\langle z|$ with $|z\rangle$ and every $\langle z'|$ with $|z'\rangle$, giving $\langle z'^{\otimes k}z^{\otimes k}|\hat{\chi}_B|z^{\otimes k}z'^{\otimes k}\rangle = 1$. Thus, we rewrite the sum as

$$\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}|\langle\tilde{\psi}_z|\tilde{\psi}_{z'}\rangle|^{2k} = \sum_{\substack{\pi\in S_{2k}\\ \chi_1,\chi_2\in S_k}}\mathrm{Wg}((\chi_1\chi_2\tau)^{-1}\pi,D_B)\mathrm{Tr}\left(\sigma_B^{\otimes 2k}\hat{\pi}_B^\dagger\right). \tag{E7}$$

Next, we split the sum over $\pi\in S_{2k}$ into terms where $\pi=\chi_1\chi_2\tau$ and $\pi\neq\chi_1\chi_2\tau$, respectively. This process gives

$$\begin{aligned}\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}|\langle\tilde{\psi}_z|\tilde{\psi}_{z'}\rangle|^{2k} &= \mathrm{Wg}(1_{2k},D_B)\sum_{\chi_1,\chi_2\in S_k}\mathrm{Tr}(\sigma_B{}^{\otimes 2k}(\hat{\chi}_{1,B}\otimes\hat{\chi}_{2,B})\hat{\tau}_B)+\sum_{\substack{\pi\in S_{2k}\\ \chi_1,\chi_2\in S_k\\ \pi\neq\chi_1\chi_2\tau}}\mathrm{Wg}((\chi_1\chi_2\tau)^{-1}\pi,D_B)\mathrm{Tr}\left(\sigma_B^{\otimes 2k}\hat{\pi}_B^\dagger\right) \\ &\le \frac{1}{D_B^{2k}}\left[1+\mathcal{O}\left(\frac{k^{7/4}}{D_B^2}\right)\right]k!\sum_{\chi\in S_k}\mathrm{Tr}\left(\sigma_B^{2\otimes k}\hat{\chi}_B\right)+\sum_{\substack{\pi\in S_{2k}\\ \chi_1,\chi_2\in S_k\\ \pi\neq\chi_1\chi_2\tau}}|\mathrm{Wg}((\chi_1\chi_2\tau)^{-1}\pi,D_B)|,\end{aligned} \tag{E8}$$

where we used [73]

$$\mathrm{Wg}(1_{2k},D_B) = \frac{1}{D_B^{2k}}\left[1+\mathcal{O}\left(\frac{k^{7/4}}{D_B^2}\right)\right] \tag{E9}$$

valid for $4k^2 < D_B$. Applying Lemma 4, we obtain

$$\sum_{\chi\in S_k}\mathrm{Tr}(\sigma_B^{2\otimes k}\hat{\chi}_B) = \mathrm{Tr}^k(\sigma_B^2)\sum_{\chi\in S_k}\mathrm{Tr}\left(\frac{\sigma_B^{2\otimes k}}{\mathrm{Tr}^k(\sigma_B^2)}\hat{\chi}_B\right) = \|\sigma_A\|_2^{2k}\left[1+\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)\right], \tag{E10}$$

using the fact that σ_A and σ_B share the same nonzero eigenvalues. To bound the last term in Eq. (E8), we use

$$\begin{aligned}\sum_{\substack{\pi\in S_{2k}\\ \chi_1,\chi_2\in S_k\\ \pi\neq\chi_1\chi_2\tau}}|\mathrm{Wg}((\chi_1\chi_2\tau)^{-1}\pi,D_B)| &= \sum_{\substack{\pi\in S_{2k}\\ \chi_1,\chi_2\in S_k\\ \pi\neq 1_{2k}}}|\mathrm{Wg}(\pi,D_B)|\\ &= (k!)^2\left(\sum_{\pi\in S_{2k}}|\mathrm{Wg}(\pi,D_B)|-|\mathrm{Wg}(1_{2k},D_B)|\right)\\ &= (k!)^2\left\{\frac{(D_B-2k)!}{D_B!}-\frac{1}{D_B^{2k}}\left[1+\mathcal{O}\left(\frac{k^{7/4}}{D_B^2}\right)\right]\right\}\\ &= (k!)^2\left\{\frac{1}{D_B^{2k}}\left[1+\mathcal{O}\left(\frac{k^2}{D_B}\right)\right]-\frac{1}{D_B^{2k}}\left[1+\mathcal{O}\left(\frac{k^{7/4}}{D_B^2}\right)\right]\right\}\\ &= \mathcal{O}\left(\frac{k^{2k+2}}{D_B^{2k+1}}\right).\end{aligned} \tag{E11}$$

Substituting these formulas into Eq. (E8) yields the desired result. ∎

Now, we are ready to prove Theorem 3 in the main text, which we reproduce here for convenience.

Theorem 10. Let $|\Psi\rangle_{AB}$ be an arbitrary bipartite state, with subsystem density operators denoted σ_A and σ_B, respectively. Consider the projected ensemble $\mathcal{E}(\Psi)$ obtained by applying a unitary U_B, drawn from an approximate unitary $2k$-design with relative error ε, on subsystem B, followed by projective measurements on B in an arbitrary orthonormal basis $\{|z\rangle\}_{z=1}^{D_B}$. Then, assuming that $k \ll \|\sigma_A\|_2/\|\sigma_A\|_4$ and $1 \ll D_A \le D_B$,

$$\mathbb{E}_{U_B}\left\|\rho_{\mathcal{E}}^{(k)}-\rho_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right\|_1 \le \left[(D_A\|\sigma_A\|_2^2)^k\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)+\mathcal{O}\left(\frac{D_A^k k^{k+2}}{D_B}\right)+\mathcal{O}(D_{A,k}\varepsilon)\right]^{1/2}. \tag{E12}$$

Proof. The kth moment of the projected ensemble $\mathcal{E}$ generated by $|\Psi_U\rangle_{AB}=(I_A\otimes U_B)|\Psi\rangle_{AB}$ is denoted as

$$\rho_{\mathcal{E}}^{(k)}=\sum_{z=1}^{D_B}p_z^{1-k}[(I_A\otimes\langle z|)|\Psi_U\rangle\langle\Psi_U|(I_A\otimes|z\rangle)]^{\otimes k}, \tag{E13}$$

where $p_z=\langle\Psi_U|(I_A\otimes|z\rangle\langle z|)|\Psi_U\rangle$ is the probability of measuring the outcome z on subsystem B. Since U is sampled from an approximate $2k$-design with relative error ε, the generator states $|\Psi_U\rangle$ satisfy

$$(1-\varepsilon)\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}(|\Psi_U\rangle\langle\Psi_U|)^{\otimes 2k}\preccurlyeq\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}(|\Psi_U\rangle\langle\Psi_U|)^{\otimes 2k}\preccurlyeq(1+\varepsilon)\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}(|\Psi_U\rangle\langle\Psi_U|)^{\otimes 2k}. \tag{E14}$$

Now, using Lemma 3, with the choices

$$\tilde{\rho}_{\mathcal{E}}^{(k)}={D_B}^{k-1}\sum_{z=1}^{D_B}[(I_A\otimes\langle z|)|\Psi_U\rangle\langle\Psi_U|(I_A\otimes|z\rangle)]^{\otimes k} \tag{E15}$$

and

$$M=\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)={D_A}^k\sigma_A^{\otimes k}\rho_{\mathrm{Haar},A}^{(k)}, \tag{E16}$$

we have

$$\begin{aligned}\Delta &= \left|\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right\|_2^2-\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right\|_2^2\right|\\ &\le 2\varepsilon\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}\mathrm{Tr}\left[\tilde{\rho}_{\mathcal{E}}^{(k)}\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right]+\varepsilon{D_B}^{2k-2}\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}\left(\sum_{z=1}^{D_B}p_z^k\right)^2.\end{aligned} \tag{E17}$$

With

$$\begin{aligned}
\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)} \mathrm{Tr}[\tilde{\rho}_{\mathcal{E}}^{(k)}\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)] &= \frac{{D_B}^k D_{A,k}}{D_{B,k}}\mathrm{Tr}(\sigma_A^{\otimes k}\rho_{\mathrm{Haar},A}^{(k)}\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)) && \text{(Lemma 9)}\\
&= \frac{{D_B}^k}{D_{B,k}}\mathrm{Tr}(\sigma_A^{\otimes k}\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A))\\
&= \frac{{D_B}^k {D_A}^k}{D_{B,k}}\mathrm{Tr}(\sigma_A^{2\otimes k}\rho_{\mathrm{Haar},A}^{(k)})\\
&= k!\left(1+\mathcal{O}\left(\frac{k^2}{D_A}\right)\right)\|\sigma_A\|_2^{2k}\left(1+\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)\right) && \text{(Lemma 4)}\\
&= k!\|\sigma_A\|_2^{2k}\left(1+\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)\right) && (\|\sigma_A\|_4^4/\|\sigma_A\|_2^4 \geq 1/D_A), \qquad \text{(E18)}
\end{aligned}$$

and, using the Cauchy-Schwarz inequality, we obtain

$$\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}}\left(\sum_{z=1}^{D_B} p_z^k\right)^2 \leq \left[\sum_{z=1}^{D_B}\left(\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}} {p_z}^{2k}\right)^{1/2}\right]^2 = \frac{1}{{D_B}^{2k-2}}\left(1+\mathcal{O}(k^2\|\sigma_A\|_2^2)\right). \tag{E19}$$

Thus,

$$\Delta \leq 2\varepsilon k!\|\sigma_A\|_2^{2k}\left(1+\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)\right)+\varepsilon\left(1+\mathcal{O}(k^2\|\sigma_A\|_2^2)\right) = \mathcal{O}(\varepsilon). \tag{E20}$$

This finding implies that

$$\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right\|_2^2 \leq \mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right\|_2^2+\mathcal{O}(\varepsilon). \tag{E21}$$

To proceed, we need to evaluate

$$\begin{aligned}
\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right\|_2^2 &= \mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}\mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)2} - 2\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}\mathrm{Tr}\left[\tilde{\rho}_{\mathcal{E}}^{(k)}\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right]+\mathrm{Tr}\tilde{\rho}_{\mathrm{Scrooge}}^{(k)2}(\sigma_A)\\
&= k!\|\sigma_A\|_2^{2k}\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)+\mathcal{O}\left(\frac{k^{2k+2}}{D_B}\right). \qquad \text{(E22)}
\end{aligned}$$

For this purpose, we use

$$\mathop{\mathbb{E}}_{U\sim\mathrm{Haar}(D_B)}\mathrm{Tr}\tilde{\rho}_{\mathcal{E}}^{(k)2} = \frac{1}{D_B}\left(1+\mathcal{O}(k^2\|\sigma_A\|_2^2+k^{2k+2})\right)+k!\|\sigma_A\|_2^{2k}\left(1+\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)\right) \tag{E23}$$

from Lemma 10, and

$$\begin{aligned}
\mathrm{Tr}\tilde{\rho}_{\mathrm{Scrooge}}^{(k)2}(\sigma_A) &= D_A^{2k}\mathrm{Tr}\left(\sigma_A^{\otimes k}\rho_{\mathrm{Haar},A}^{(k)}\sigma_A^{\otimes k}\rho_{\mathrm{Haar},A}^{(k)}\right)\\
&= \frac{D_A^{2k}}{D_{A,k}}\mathrm{Tr}(\sigma_A^{2\otimes k}\rho_{\mathrm{Haar},A}^{(k)})\\
&= k!\|\sigma_A\|_2^{2k}\left(1+\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)\right) \qquad \text{(Lemma 4)}. \qquad \text{(E24)}
\end{aligned}$$

Thus,

$$\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathrm{Scrooge}}^{(k)}(\sigma_A)\right\|_2^2 \leq \mathcal{O}\left(\frac{k^{2k+2}}{D_B}\right)+k!\|\sigma_A\|_2^{2k}\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)+\mathcal{O}(\varepsilon). \tag{E25}$$

This result can be used to upper bound the trace distance, via

$$\left(\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma_A)\|_1\right)^2 \le D_{A,k}\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma_A)\right\|_2^2$$
$$= (D_A\|\sigma_A\|_2^2)^k\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)+\mathcal{O}\left(\frac{D_A^k k^{k+2}}{D_B}\right)+\mathcal{O}(D_{A,k}\varepsilon). \tag{E26}$$

The next step is to bound the average trace distance between $\rho_{\mathcal{E}}^{(k)}$ and $\tilde{\rho}_{\mathcal{E}}^{(k)}$, and the trace distance between $\rho_{\text{Scrooge}}^{(k)}(\sigma_A)$ and $\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma_A)$. From Lemma 2, we have

$$\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1 \le \sum_{z=1}^{D_B}\left(\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}p_z^2\right)^{1/2}\left(1-2D_B^{k-1}\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}p_z^{k-1}+D_B^{2k-2}\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}p_z^{2k-2}\right)^{1/2}. \tag{E27}$$

Using the fact that

$$\begin{aligned}\mathop{\mathbb{E}}_{U\sim\text{Haar}(D_B)}p_z^k &= \mathop{\mathbb{E}}_{U\sim\text{Haar}(D_B)}\langle z|U\sigma_B U^\dagger|z\rangle^k\\ &= \text{Tr}\left(\sigma_B^{\otimes k}\rho_{\text{Haar},B}^{(k)}\right)\\ &= \frac{1}{k!D_{B,k}}\left(1+\mathcal{O}(k^2\|\sigma_B\|_2^2)\right) && \text{(Lemma 4)}\\ &= \frac{1}{D_B^k}\left(1+\mathcal{O}\left(\frac{k^2}{D_B}\right)\right)\left(1+\mathcal{O}(k^2\|\sigma_A\|_2^2)\right)\\ &= \frac{1}{D_B^k}\left(1+\mathcal{O}(k^2\|\sigma_A\|_2^2)\right) && (\|\sigma_B\|_2^2=\|\sigma_A\|_2^2\ge 1/D_B),\end{aligned} \tag{E28}$$

and

$$(1-\varepsilon)\mathop{\mathbb{E}}_{U\sim\text{Haar}(D_B)}p_z^\ell \le \mathop{\mathbb{E}}_{U\sim 2k\text{-design}}p_z^\ell \le (1+\varepsilon)\mathop{\mathbb{E}}_{U\sim\text{Haar}(D_B)}p_z^\ell \tag{E29}$$

for all $1\le\ell\le 2k$, we obtain

$$\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1 \le \mathcal{O}\left(\sqrt{\varepsilon+k^2\|\sigma_A\|_2^2}\right). \tag{E30}$$

From Lemma 6,

$$\|\rho_{\text{Scrooge}}^{(k)}(\sigma_A)-\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma_A)\|_1 \le \mathcal{O}(k\|\sigma_A\|_2). \tag{E31}$$

Therefore, by the triangle inequality,

$$\begin{aligned}&\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\rho_{\text{Scrooge}}^{(k)}(\sigma_A)\right\|_1\\ &\le \mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\rho_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\mathcal{E}}^{(k)}\right\|_1+\left\|\rho_{\text{Scrooge}}^{(k)}(\sigma_A)-\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma_A)\right\|_1+\mathop{\mathbb{E}}_{U\sim 2k\text{-design}}\left\|\tilde{\rho}_{\mathcal{E}}^{(k)}-\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma_A)\right\|_1\\ &\le \mathcal{O}\left(\sqrt{\varepsilon+k^2\|\sigma_A\|_2^2}\right)+\mathcal{O}(k\|\sigma_A\|_2)+\left[(D_A\|\sigma_A\|_2^2)^k\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)+\mathcal{O}\left(\frac{D_A^k k^{k+2}}{D_B}\right)+\mathcal{O}(D_{A,k}\varepsilon)\right]^{1/2}\\ &= \left[(D_A\|\sigma_A\|_2^2)^k\mathcal{O}\left(k^2\frac{\|\sigma_A\|_4^4}{\|\sigma_A\|_2^4}\right)+\mathcal{O}\left(\frac{D_A^k k^{k+2}}{D_B}\right)+\mathcal{O}(D_{A,k}\varepsilon)\right]^{1/2},\end{aligned} \tag{E32}$$

with the dominant contribution to the error bound coming from the average trace distance between $\tilde{\rho}_{\mathcal{E}}^{(k)}$ and $\tilde{\rho}_{\text{Scrooge}}^{(k)}(\sigma_A)$. ■

1. Application: Local Hamiltonian at finite temperatures

Theorem 3 requires the condition $k \ll \|\sigma_A\|_2/\|\sigma_A\|_4$ for the projected ensemble to converge to Scrooge(σ_A). As stated in the main text, we can define the effective dimension of σ_A via

$$D_{A,\mathrm{eff}} = \left(\frac{\|\sigma_A\|_2}{\|\sigma_A\|_4}\right)^4. \tag{E33}$$

If $\sigma_A = I_A/D_A$ is the maximally mixed state, then $D_{A,\mathrm{eff}} = D_A$. Subsequently, the above condition reads $k^4 \ll D_{A,\mathrm{eff}}$.

This condition is usually satisfied in many-body quantum systems. To analyze a concrete example, let us consider the case where σ_A is the thermal Gibbs state,

$$\sigma_A = \frac{e^{-\beta H_A}}{\mathrm{Tr}(e^{-\beta H_A})}, \tag{E34}$$

where β is the inverse temperature and H_A is the Hamiltonian restricted to subsystem A.

Let us denote the spectral density of the Hamiltonian H_A by $f(E)$, which satisfies the normalization

$$\int_{-\infty}^{\infty} dE\, f(E) = 1. \tag{E35}$$

We assume that H_A has a Gaussian spectral density, i.e.,

$$f(E) = \frac{1}{\sqrt{2\pi\Delta^2}} \exp\left(-\frac{(E-\mu)^2}{2\Delta^2}\right), \tag{E36}$$

where

$$\mu = \frac{1}{D_A}\sum_{i=1}^{D_A} E_i = \frac{\mathrm{Tr} H_A}{D_A} \tag{E37}$$

is the mean energy and

$$\Delta^2 = \frac{1}{D_A}\sum_{i=1}^{D_A} E_i^2 - \mu^2 = \frac{\mathrm{Tr} H_A^2}{D_A} - \frac{\mathrm{Tr}^2 H_A}{D_A^2} \tag{E38}$$

is the variance. To convert the discrete sum over energies to an integral over the spectral density, we use $D_A^{-1}\sum_{i=1}^{D_A} \to \int dE\, f(E)$. Thus, the partition function at inverse temperature β is given by

$$\begin{aligned} Z_\beta = \mathrm{Tr}\, e^{-\beta H_A} &= \sum_{i=1}^{D_A} e^{-\beta E_i} = D_A \int_{-\infty}^{\infty} dE\, f(E) e^{-\beta E} \\ &= D_A \exp\left(-\beta\mu + \frac{1}{2}\beta^2\Delta^2\right). \end{aligned} \tag{E39}$$

Note that this gives the partition function for any H (regardless of spectral density) to quadratic order in β, i.e., in the high-temperature regime.

The generalized purities read

$$\begin{aligned} \mathrm{Tr}({\sigma_A}^q) = \frac{Z_{q\beta}}{{Z_\beta}^q} &= \frac{D_A \exp(-q\beta\mu + \frac{1}{2}q^2\beta^2\Delta^2)}{D_A^q \exp(-q\beta\mu + \frac{1}{2}q\beta^2\Delta^2)} \\ &= D_A^{1-q} \exp\left(\frac{1}{2}q(q-1)\beta^2\Delta^2\right). \end{aligned} \tag{E40}$$

For the above equation to be self-consistent, we need to impose the constraint $\mathrm{Tr}({\sigma_A}^q) \le 1$, which implies the constraint $\beta^2\Delta^2 \le (2/q) N_A \ln 2$. Outside of this range (e.g., at low temperatures), the expression becomes unphysical. This finding reflects the fact that the finiteness of the spectrum is relevant in the low-temperature regime, leading to the breakdown of the Gaussian approximation. In what follows, we will work in the self-consistent regime.

Let us expand H_A in the Pauli basis, i.e.,

$$H_A = \sum_{m=1}^{M} c_m P_m, \tag{E41}$$

where $P_m \neq I$ is a Pauli string on N qubits and $|c_m| \le 1$ are real coefficients, as explained in Eq. (21) of the main text. Then,

$$\mu = \frac{\mathrm{Tr}\, H_A}{D_A} = 0, \tag{E42}$$

and

$$\Delta^2 = \frac{\mathrm{Tr}\, H_A^2}{D_A} = \frac{1}{D_A}\sum_{m,n=1}^{M} c_m c_n \mathrm{Tr}(P_m P_n) = \sum_{m=1}^{M} c_m^2 \le M. \tag{E43}$$

For a geometrically local H_A, we expect $M \propto N_A$; thus, $\Delta^2 = O(N_A)$. For simplicity, we set $M = \alpha N_A$, where α is some constant. Note that we are using dimensionless units here, with the characteristic energy scale given by the typical magnitude of c_m. Thus, we have

$$\|\sigma_A\|_2 = \frac{1}{D_A^{1/2}} e^{\beta^2\Delta^2/2} \tag{E44}$$

and

$$\|\sigma_A\|_4 = \frac{1}{D_A^{3/4}} e^{3\beta^2\Delta^2/2}, \tag{E45}$$

which yields

$$\frac{\|\sigma_A\|_4}{\|\sigma_A\|_2} = \frac{1}{D_A^{1/4}} e^{\beta^2\Delta^2} = \exp\left(\beta^2\Delta^2 - \frac{N_A}{4}\ln 2\right). \quad \text{(E46)}$$

For $\beta < \beta_c$, where

$$\beta_c = \sqrt{\frac{N_A}{4\Delta^2}\ln 2} \geq \sqrt{\frac{N_A}{4M}\ln 2} = \sqrt{\frac{\ln 2}{4\alpha}}, \quad \text{(E47)}$$

the ratio $\|\sigma_A\|_4/\|\sigma_A\|_2$ is exponentially suppressed in N_A, and therefore $k \ll \|\sigma_A\|_2/\|\sigma_A\|_4$ is satisfied for large N_A and any subexponential k. Finally, we remark that the value of β_c lies within the range $\beta^2\Delta^2 \leq N_A \ln 2/2$, imposed by self-consistency up to $q = 4$, as explained above.

APPENDIX F: UNIFORM RANDOM PHASE STATES

Here, we present additional numerical results on the random phase states given by

$$|\Psi\rangle_{AB} = \frac{1}{2^{N/2}}\sum_j e^{-i\varphi_j}|j\rangle. \quad \text{(F1)}$$

While in the main text we consider the trace distance $\Delta^{(k)}$ for our numerical studies, now we also regard a weaker notion of distance, namely, the Hilbert-Schmidt distance. It is given by

$$\Delta_{\mathrm{HS}}^{(k)}(\sigma) = \frac{1}{2}\left\|\rho_{\mathcal{E}}^{(k)} - \rho_{\mathrm{Scrooge}}^{(k)}(\sigma)\right\|_2, \quad \text{(F2)}$$

which is a weaker notion of statistical closeness between the ensembles $\mathcal{E}$ and Scrooge(σ) because

$$\Delta_{\mathrm{HS}}^{(k)} \leq \Delta^{(k)} \leq \sqrt{D_{A,k}}\Delta_{\mathrm{HS}}^{(k)}; \quad \text{(F3)}$$

thus, a small $\Delta_{\mathrm{HS}}^{(k)}$ does not guarantee a small $\Delta^{(k)}$.

We study random phase states with computational basis measurements ($\theta = 0$). We plot the Hilbert-Schmidt distance $\Delta_{\mathrm{HS}}^{(k)}$ in Fig. 6(a) and reproduce the trace distance from the main text for reference in Fig. 6(b). We find similar behavior in both cases, although the difference between finite N_{B} and $N_{\mathrm{B}} \to \infty$ is less pronounced for $\Delta_{\mathrm{HS}}^{(k)}$ compared to $\Delta^{(k)}$.

APPENDIX G: T-DOPED CLIFFORD CIRCUITS

In this appendix, we study emergent Scrooge ensembles via rotating the measurement basis using Clifford circuits doped with T-gates. In particular, we prepare the entangled state

$$|\Psi(\chi)\rangle = |\psi_{\mathrm{ent}}(\chi)\rangle^{\otimes N_{\mathrm{A}}}|0_{B_2}\rangle^{\otimes N_{\mathrm{B}}-N_{\mathrm{A}}}, \quad \text{(G1)}$$

with $|\psi_{\mathrm{ent}}(\chi)\rangle = \cos(\chi/2)|0_A 0_{B_1}\rangle + \sin(\chi/2)|1_A 1_{B_1}\rangle$ between A and B_1. We then apply a Clifford unitary of depth d and N_{T} gates on $B = B_1 \cup B_2$ only, in the setting described in the main text. Now, in Fig. 7, we study the behavior as a function of depth d and N_{T} in more detail. First, in Fig. 7(a), we plot $\Delta^{(2)}$ against d for different N_{T}. We find that $\Delta^{(2)}$ decreases with d and converges to a limiting value for large d, indicating the need for scrambling via entangling gates to generate projected Scrooge ensembles. However, we find that the minimal $\Delta^{(2)}$ decreases with an increasing number of T-gates N_{T}, demonstrating that magic is also necessary for Scrooge ensembles. Notably, there is a number of T-gates beyond which $\Delta^{(2)}$ no longer improves.

In Fig. 7(b), we show $\Delta^{(2)}$ against d for different total qubit numbers N, where we choose $N_{\mathrm{T}} = 3N$, i.e., in the limit where magic is large such that we can converge to minimal $\Delta^{(2)}$. We find that for all N, the decrease in d is the same, until beyond a certain d, where $\Delta^{(2)}$ does not decrease further and plateaus. The plateau value of $\Delta^{(2)}$ arises from finite-size effects and decreases exponentially with N. We refer to the depth d where $\Delta^{(2)}$ stops decreasing as d_0. We find that d_0 increases with N, which we find to be approximately $d_0 \approx \frac{2}{3}N$ from our numerical study.

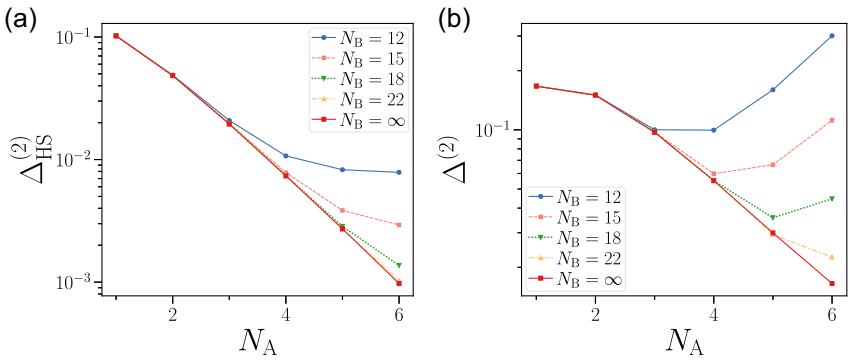

FIG. 6. Hilbert-Schmidt distance to Haar 2-design $\Delta_{\mathrm{HS}}^{(2)}$ of the projected ensemble generated from uniform random phase states (F1). We measure N_{B} qubits in the computational basis ($\theta = 0$), and show (a) $\Delta_{\mathrm{HS}}^{(2)}$ and (b) $\Delta^{(2)}$ against N_{A}.

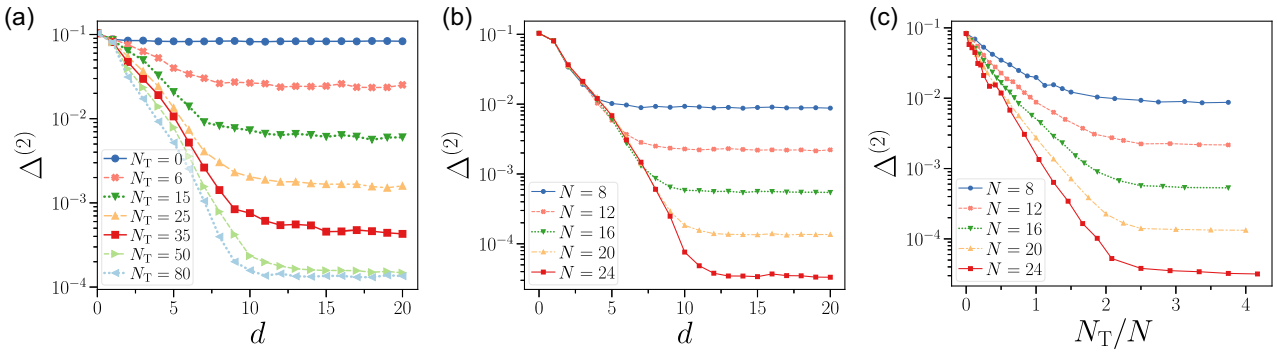

FIG. 7. Trace distance to Scrooge 2-design $\Delta^{(2)}$ of projected ensemble generated from $|\Psi(\chi)\rangle$ as defined in Eq. (G1) where we choose $\chi = \pi/6$. We then measure on B in a rotated basis via unitaries composed of d layers of local Clifford gates doped with N_T T-gates. (a) $\Delta^{(2)}$ against d for different N_T. We have $N_A = 1$, $N_B = 19$, and average over 500 random realizations of the circuit. (b) $\Delta^{(2)}$ against d for different total qubit numbers N and fixed $N_T = 3N$. (c) $\Delta^{(2)}$ against T-gate density N_T/N for different total qubit numbers N and fixed depth $d = 30$.

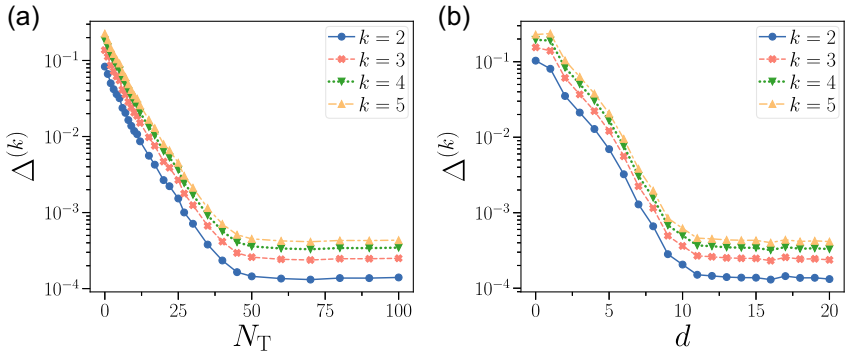

FIG. 8. Trace distance to Scrooge k-design $\Delta^{(k)}$ of projected ensemble generated from $|\Psi(\chi)\rangle$ as defined in Eq. (G1) where we choose $\chi = \pi/6$. We then measure in a transformed basis on B with d layers of local Clifford gates doped with N_T T-gates. (a) $\Delta^{(k)}$ against N_T for $d = 30$ and different k. (b) $\Delta^{(k)}$ against d for $N_T = 3N$. We have $N_A = 1$, $N_B = 19$, and average over 100 random realizations of the circuit.

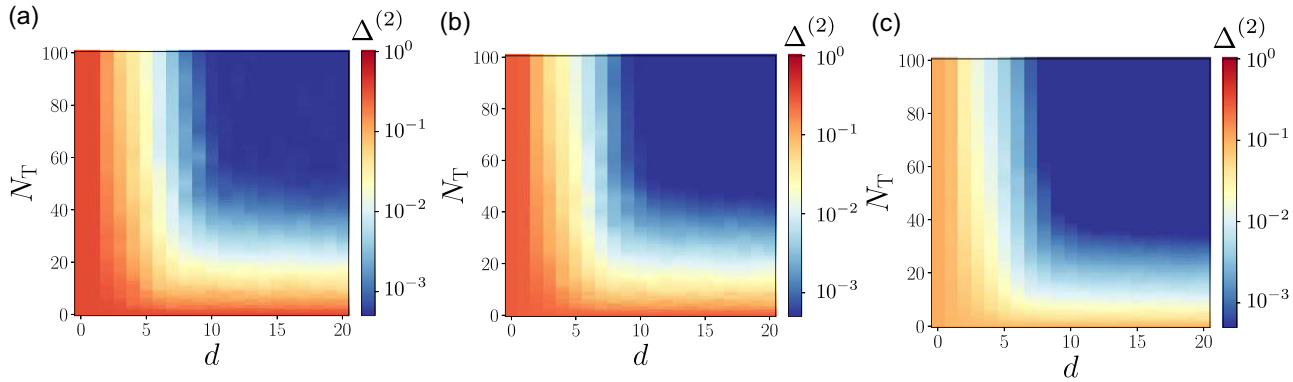

FIG. 9. Comparison of different degrees of entanglement between A and B, where we show different angles χ for the projected ensemble generated from $|\Psi(\chi)\rangle$ as defined in Eq. (G1). We measure in a transformed basis on B with d layers of local Clifford gates doped with N_T T-gates. We plot d against N_T, where we show trace distance to Scrooge 2-design $\log_{10}(\Delta^{(2)})$ in color. We have (a) $\chi = \pi/2$, (b) $\chi = \pi/3$, and (c) $\chi = \pi/6$. We fix $N_A = 1$, $N_B = 19$, and $N = 20$, and average over 100 random realizations of the circuit.

In Fig. 7(c), we show $\Delta^{(2)}$ against T-gate density N_T/N for deep circuits of depth $d = 30$. We find that $\Delta^{(2)}$ decreases with N_T/N, converging to an N-dependent minimum. We observe that this minimum is reached at approximately $N_T \approx 2.5$, indicating a transition in magic when the system becomes fully random. This finding mirrors the saturation transition in magic observed in Ref. [126].

Now, we study higher moments k of the Scrooge ensemble. In Fig. 8(a), we plot $\Delta^{(k)}$ against N_T for large d and different k. We find a qualitatively similar decay and convergence behavior for all k. Similarly, in Fig. 8(b), we plot $\Delta^{(k)}$ against d for large N_T.

Finally, we study different angles χ for the generator state $|\psi_{\text{ent}}(\chi)\rangle$. The choice of χ affects the entanglement between subsystems A and B and thus the reduced density matrix σ_A of the corresponding Scrooge ensemble. We show $\chi = \pi/2$ in Fig. 9(a), which corresponds to the case where $\sigma_A = I/D_A$ is the maximally mixed state. Then, we show $\chi = \pi/3$ in Fig. 9(b) and $\chi = \pi/6$ in Fig. 9(c). We find similar behavior as a function of N_T and d for all χ.

APPENDIX H: T-DOPED CLIFFORD GENERATOR STATES

Next, we consider T-doped Clifford circuits, where, in contrast to the main text, we apply the circuit on both subsystems A and B: We prepare a $|0\rangle^{\otimes N}$ state and then apply the doped Clifford + T circuit on all N qubits (instead of just subsystem B). Here, in Fig. 10(a), we study circuits of d layers, composed of random single-qubit Clifford gates, CNOT gates arranged in a 1D nearest-neighbor configuration, which are doped with a total of N_T T-gates placed randomly in the circuit. While the scrambling via the Clifford circuit increases with d, the magic increases with N_T. In Fig. 10(a), we show the circuit and setup in detail. Then, in Fig. 10(b), we show a 2D heat map for $\Delta^{(2)}$, with varying N_T and d. We find that, for low N_T or low d, the additive error between the projected ensemble and the corresponding Scrooge ensemble is large. Both N_T and d need to be sufficiently large to yield a small $\Delta^{(2)}$. In Fig. 10(c), we study $\Delta^{(2)}$ against N_T for different N, where we fix the circuit depth to be $d = 30$. We find that $\Delta^{(2)}$ decays exponentially with N_T and saturates to a plateau

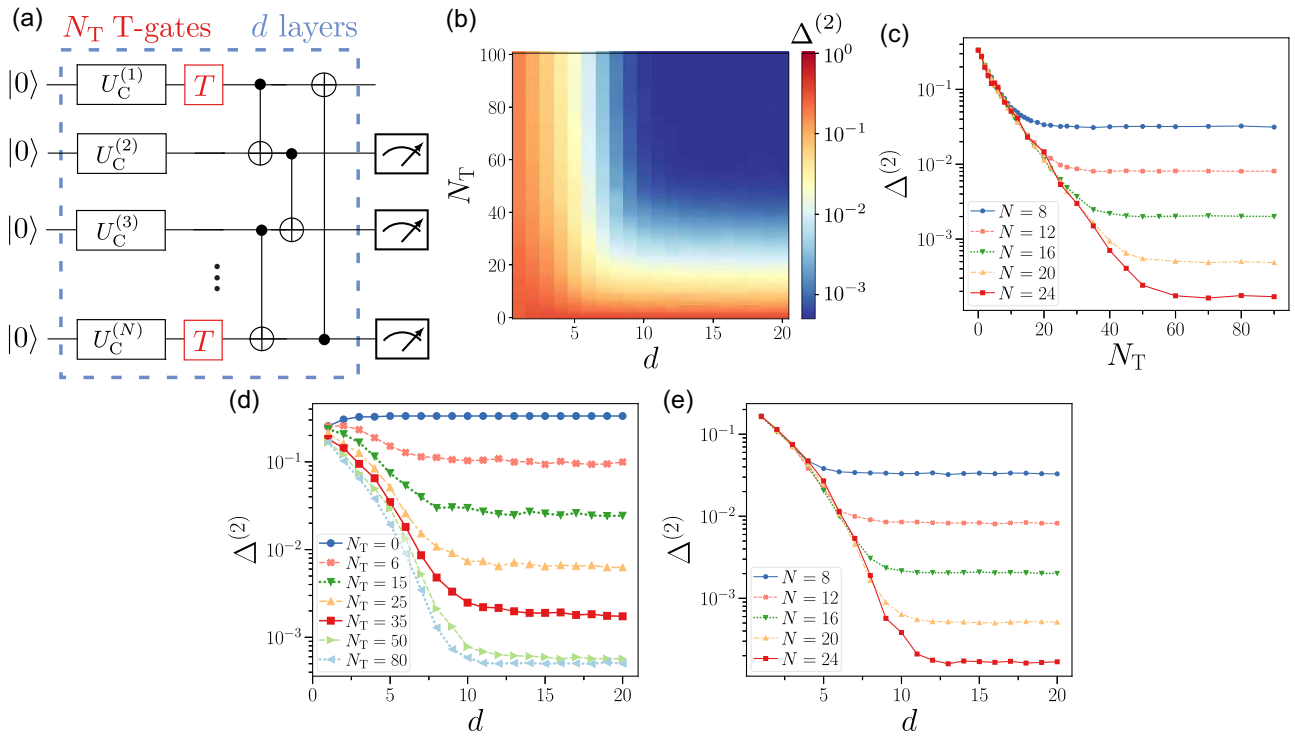

FIG. 10. Distance to Scrooge 2-design $\Delta^{(2)}$ of projected ensemble generated from $U|0\rangle^{\otimes N}$, where N-qubit unitary U consists of d layers of random single-qubit Clifford unitaries and a 1D layer of nearest-neighbor CNOT gates with periodic boundary conditions, where the circuit is doped with N_T T-gates at random positions. (a) Sketch of doped Clifford circuit with $N_A = 1$ and $N_B = 3$. (b) Plot of d (starting from $d \geq 1$) against N_T, where we show $\Delta^{(2)}$ (in logarithmic scale) on a color scale. We choose $N_B = 19$ and $N_A = 1$. (c) $\Delta^{(2)}$ against N_T for different total qubit numbers N and fixed $d = 30$. We have $N_A = 1$, $N_B = N - N_A$, and we average over 500 random realizations of the circuit. (d) $\Delta^{(2)}$ against d for different N_T. (e) $\Delta^{(2)}$ against d for different total qubit numbers N and fixed $N_T = 3N$. We average over 500 random realizations of the circuit.

value at large N_T. The plateau value is primarily due to finite-size effects, and it decays exponentially with N. In Fig. 10(d), we plot $\Delta^{(2)}$ against d for different N_T. The behavior mirrors the one observed in Fig. 7(a). In Fig. 10(e), we study $\Delta^{(2)}$ against d for different N, where we choose $N_T = 3N$, i.e., in the limit where magic is large such that we can converge to minimal distance. Notably, we find similar behavior as in the case of Fig. 7(b).

APPENDIX I: TRANSVERSE-FIELD ISING MODEL

In this appendix, we study projected ensembles generated from the ground state of the 1D transverse-field Ising model with periodic boundary conditions, as defined in the main text, in more detail. We apply a random Clifford unitary on subsystem B, then measure the qubits in B in the computational basis, and finally construct the projected ensemble over the remaining N_A qubits. In Fig. 11(a), we study the distance to the Scrooge 2-design $\Delta^{(2)}$ against N_B for different h. Notably, we find exponential decay $\Delta^{(2)} \sim 2^{-\alpha(h)N_B}$ for all $h \neq 0$, where we fit $\alpha(h)$. In Fig. 11(b), we plot $\alpha(h)$ against h, finding a pronounced maximum close to the critical point $h = 1$, indicating that the distance decays fastest at the critical point.

Next, we study in Fig. 11(c) the field h against $y = \log_2(\Delta^{(2)})/N_B$, which is the logarithm of the trace distance normalized by N_B. We define y as it converges to a nonzero constant for large N_B, as we will show below. Notably, we find that, for sufficiently large N_B, there is a pronounced dip for $h \approx 1$, matching the well-known critical point of the Ising model [130]. Now, we characterize the behavior around the critical point. First, we define the minimal distance $y_0(N_B) = \min_h y(h, N_B)$, where we perform the minimization at a small region around the critical point. In Fig. 11(d), we plot y_0 against N_B, finding that y_0 increases with N_B. Following the approach of Ref. [129], we fit the curve with $y_0 = aN_B^\gamma + y_c$, finding good agreement. The fit allows us to extract the asymptotic value $y_0(N_B \to \infty)$. In Fig. 11(e), we plot the minimal field $h_0(N_B) = \arg\min_h y(h, N_B)$ against N_B. We perform a

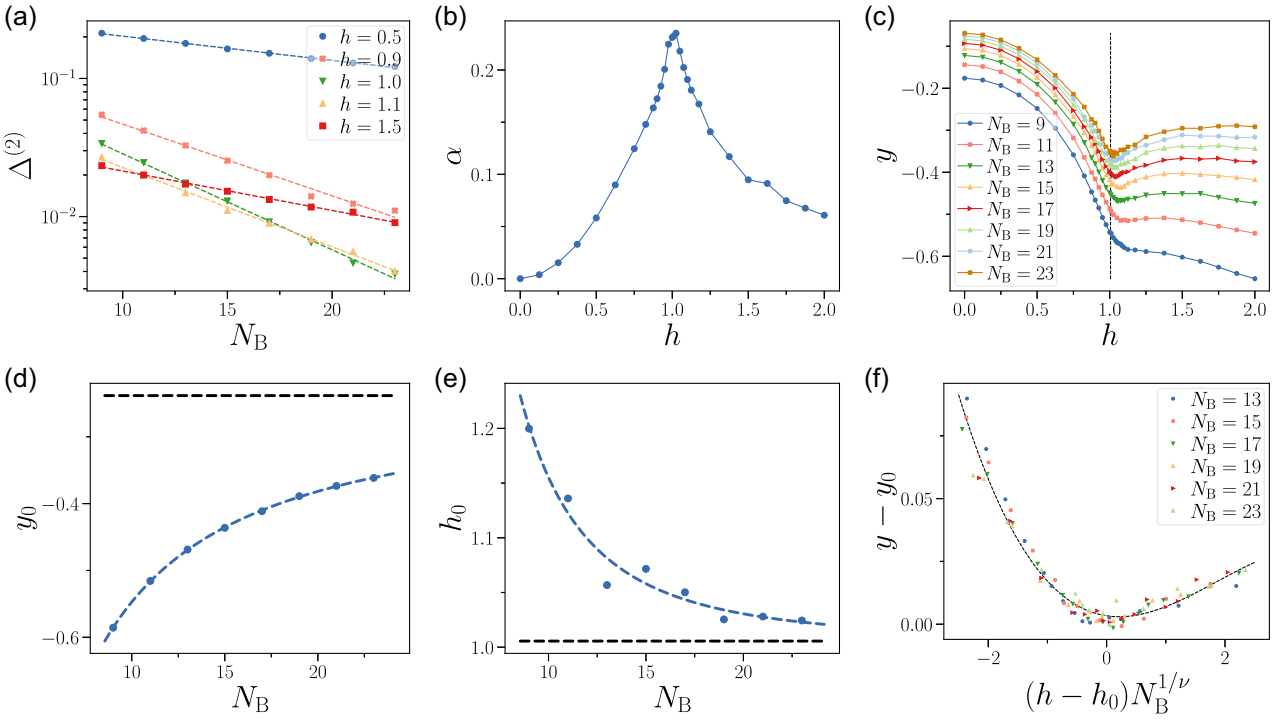

FIG. 11. Distance to Scrooge 2-design $\Delta^{(2)}$ for projected ensemble generated from the ground state of the Ising model (29). We apply unitary U_B on N_B qubits, measure N_B qubits, and gain a projected ensemble of $N_A = 1$ qubits. (a) $\Delta^{(2)}$ against N_B for different h for the case of random Clifford unitaries. Dashed lines are the fit $\Delta^{(2)} \sim 2^{-\alpha(h)N_B}$ (b) Fitted exponential decay rates $\alpha(h)$ against h. In panel (c), we define $y = \log_2(\Delta^{(2)})/N_B$ and plot against h. The dashed vertical line is the critical point $h_c = 1$ of the Ising model, which we now proceed to fit using y. In panel (d), we plot the minimal distance $y_0(N_B) = \min_h y(h, N_B)$ against N_B, where the minimization over h is performed around a small neighborhood around $h = 1$. We fit with $y_0 = aN_B^\gamma + y_c$, where we find, as the asymptotic value, $y_c \equiv y_0(N_B \to \infty) \approx -0.238(7)$. In panel (e), we plot the field with minimal distance $h_0(N_B) = \text{argmin}_h y(h, N_B)$ against N_B. We fit with $h_0 = aN_B^\gamma + h_c$, where we find the fitted critical field $h_c \equiv h_0(N_B \to \infty) \approx 1.006(22)$. In panel (f), we rescale y and h with fitted y_0 and h_0. We find that curves for different N_B collapse onto a single curve when rescaling $h - h_0$ with $N_B^{1/\nu}$, where we have $\nu = 1$ as expected for the Ising universality class. The dashed line is a third-order polynomial fit.

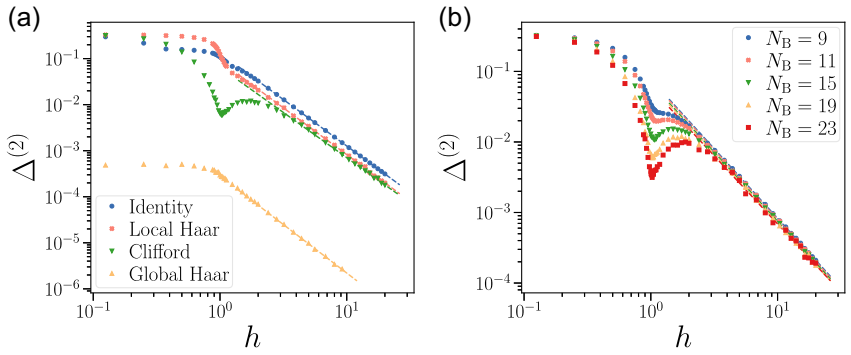

FIG. 12. Distance to Scrooge 2-design $\Delta^{(2)}$ of the projected ensemble generated from the ground state of the Ising model by applying unitary U_B on B and then measuring B in the computational basis. We have $N_A = 1$ qubits for the projected ensemble. In panel (a), we plot $\Delta^{(2)}$ against h for different types of unitaries U_B on B, which has $N_B = 19$ qubits. For large $h \gg 1$, the dashed line shows a polynomial fit $\Delta^{(2)} \sim h^{-\alpha}$, where we find $\alpha \approx 2$. In panel (b), we plot $\Delta^{(2)}$ against h for different N, where we choose random Clifford unitaries U_B.

similar polynomial fit with $h_0 = aN_B^{\gamma} + h_c$, finding the asymptotic field $h_c \equiv h_0(N_B \to \infty) \approx 1.006(22)$, which closely matches the critical point $h = 1$. Our scaling analysis shows that, within our numerical study, the dip in $\Delta^{(2)}$ indeed converges to $h = 1$ for $N_B \to \infty$ (see Appendix I). Thus, emergent Scrooge designs can provide a method to determine the critical point [129,130].

Next, in Fig. 11(f), we rescale y and h with fitted y_0 and h_0 for different N_B. We find that curves for different N_B collapse onto a single curve when rescaling $h - h_0$ with $N_B^{1/\nu}$, where we have $\nu = 1$ as expected for the Ising universality class. We also show a third-order polynomial fit as a dashed line, allowing us to predict the behavior close to the critical point for all N_B.

Next, we study the behavior of the projected ensemble generated from the ground state of the Ising model for large fields h. We transform bipartition B of the ground state with random Clifford unitaries U_B. In Fig. 12(a), we plot $\Delta^{(2)}$ against h for different types of unitaries applied on N_B. Via fitting, we find that, for all unitaries, we have a decay $\Delta^{(2)} \sim h^{-\alpha}$, with $\alpha \approx 2$. The decay results from the fact that, for $h \to \infty$, the ground state of the Ising model is a product state. As product states are trivially exact Scrooge ensembles, $\Delta^{(2)}$ must decay to zero as $h \to \infty$.

In Fig. 12(b), we plot $\Delta^{(2)}$ against h for different N_B. For large h, we find that all N_B follow the same polynomial decay as $\Delta^{(2)} \sim h^{-\alpha}$, where again we find $\alpha \approx 2$.

Finally, we study the error of Scrooge k-design $\Delta^{(k)}$ beyond $k = 2$. In Fig. 13, we plot $\Delta^{(k)}$ against h for different measurement bases (via unitary U_B and $k = 2, \ldots, 5$). Here, we have the identity applied on B in Fig. 13(a), the tensor product of single-qubit Haar random unitaries in Fig. 13(b), random Clifford unitaries in Fig. 13(c), and unitaries drawn from the Haar measure in Fig. 13(d). We find similar behavior for all k shown, indicating that the qualitative behavior for $k = 2$ holds similarly for higher k.

APPENDIX J: HEISENBERG MODEL

Next, we study the Heisenberg model with anisotropy h,

$$H_{XXZ} = \sum_{j=1}^{N} (-X_j X_{j+1} - Y_j Y_{j+1} - h Z_j Z_{j+1}), \quad \text{(J1)}$$

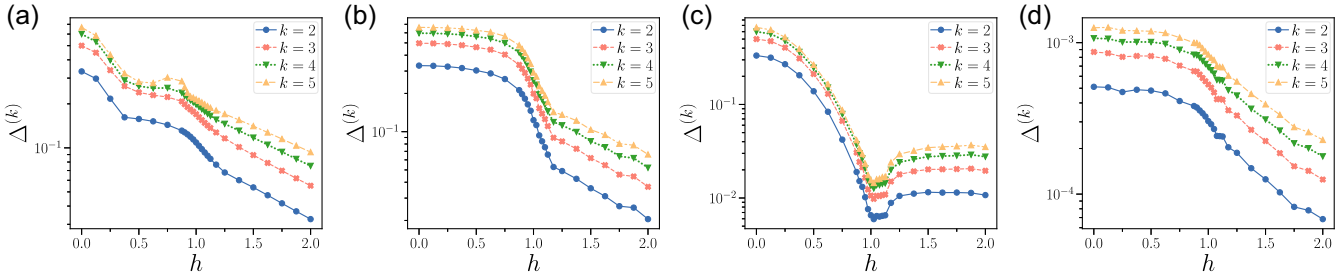

FIG. 13. Distance to Scrooge k-design $\Delta^{(k)}$ of the projected ensemble generated from the ground state of the Ising model by applying unitary U_B on B and then measuring it in the computational basis. We show $\Delta^{(k)}$ against h for different types of unitaries U_B on B with $N_B = 19, N_A = 1$. We regard U_B as (a) identity, (b) single-qubit Haar random unitaries, (c) random Clifford unitaries, and (d) unitaries drawn from the Haar measure on N_B qubits.

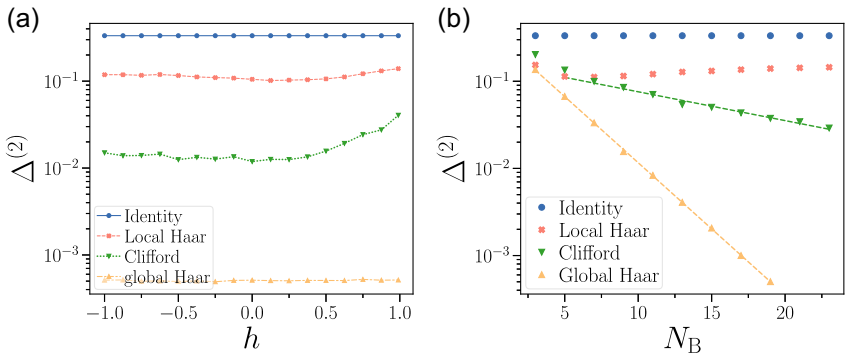

FIG. 14. Trace distance to Scrooge 2-design $\Delta^{(2)}$ for the projected ensemble generated from the ground state of the Heisenberg model (J1). We apply unitary U_B on N_B qubits, measure N_B qubits, and gain the projected ensemble of $N_A = 1$ qubits. (a) $\Delta^{(2)}$ against anisotropy h for U_B being identity, single-qubit Haar random unitaries, random Clifford unitaries, or random unitaries drawn from the Haar measure on N_B qubits. (b) $\Delta^{(2)}$ against N_B for different U_B. The dashed line is the fit with $\Delta^{(2)} \sim 2^{-\gamma N_B}$, where we find $\gamma_{\text{Clifford}} \approx 0.11$ and $\gamma_{\text{Haar}} \approx 0.5$.

where X_j, Y_j, and Z_j are the respective Pauli x, y, and z operators acting on the jth qubit. We study the emergent Scrooge designs generated from the ground state in Fig. 14. We apply unitary U_B on B and regard the projected ensemble of the remaining N_A qubits. We regard U_B as being the identity, single-qubit Haar random unitaries, or random unitaries drawn from the Haar measure on N_B qubits. In Fig. 14(a), we find that $\Delta^{(2)}$ changes slightly with anisotropy h. Notably, we find an increase in $\Delta^{(2)}$ for $h \approx 1$, which is most pronounced when we apply Clifford unitaries on B. In Fig. 14(b), we study $\Delta^{(2)}$ against N_B for different U_B for $h = 1$. We find an exponential decay for Clifford and Haar random unitaries on B, while local unitaries (i.e., single-qubit Haar or identity) yield a large $\Delta^{(2)}$ for any N_B.

APPENDIX K: STABILIZER STATES WITH DIFFERENT BASIS MEASUREMENTS

In this appendix, we study the projected ensemble generated from stabilizer states. We prepare a random N-qubit stabilizer state, apply unitaries U_B on the bipartition B, proceed to measure B in the computational basis, and study the emergent Scrooge ensemble on A. For U_B, we consider random Clifford unitaries, the tensor product of single-qubit Haar random unitaries, the tensor product of T-gates followed by the tensor product of Hadamard gates to measure in the X-basis, and N_B-qubit Haar random unitaries. In Fig. 15, we show the distance to Haar 2-design $\Delta^{(2)}$ against N_B for different classes of unitaries U_B. Notably, we find that when U_B is Clifford, $\Delta^{(2)}$ is large, which follows from the fact that the state after basis rotation is still a stabilizer state. In contrast, for random single-qubit Haar unitaries, rotation into the T-basis $T = \text{diag}(1, e^{-i\pi/4})$, or Haar random unitaries over N_B, we find exponential decay with N_B. Thus, by injecting magic into B via a unitary (which does not commute with the computational measurement basis), we can generate emergent Scrooge ensembles with low error. Notably, as the initial stabilizer state is already highly coherent and has been scrambled via Cliffords, we find that injection of magic via measurements in a local magical basis (such as single-qubit Haar random unitaries) is then sufficient to produce good Scrooge designs. This finding contrasts with the case of ground states of local Hamiltonians (such as the Ising or Heisenberg model), which are weakly scrambled, and thus local (magical) measurements are not sufficient to yield Scrooge designs.

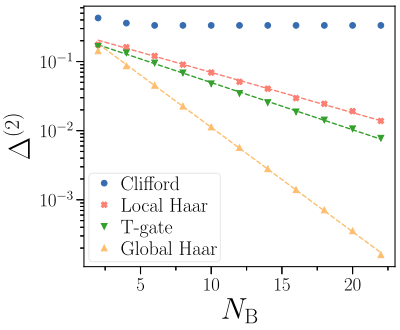

FIG. 15. Distance to Haar 2-design $\Delta^{(2)}$ for the projected ensemble of $N_A = 1$ qubits, generated from the initial stabilizer state over $N = N_A + N_B$ qubits. We apply on B either random Clifford unitaries, single-qubit Haar random unitaries, the tensor product of T-gates (followed by the tensor product of Hadamard gates), or random unitaries drawn from the Haar measure, and we measure B in the computational basis. We plot $\Delta^{(2)}$ against N_B. The dashed line is a fit with $\Delta^{(2)} \sim 2^{-\gamma N_B}$, where we find $\gamma_{\text{1-Haar}} \approx 0.19$, $\gamma_{\text{T-gate}} \approx 0.23$, and $\gamma_{\text{Haar}} \approx 0.5$. We average over 100 random initializations.

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

[2] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, *Colloquium: Many-body localization, thermalization, and entanglement*, Rev. Mod. Phys. **91**, 021001 (2019).

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

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

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

[6] 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).

[7] F. Borgonovi, F. Izrailev, L. Santos, and V. Zelevinsky, *Quantum chaos and thermalization in isolated systems of interacting particles*, Phys. Rep. **626**, 1 (2016).

[8] T. Mori, T. N. Ikeda, E. Kaminishi, and M. Ueda, *Thermalization and prethermalization in isolated quantum systems: A theoretical overview*, J. Phys. B **51**, 112001 (2018).

[9] M. Ueda, *Quantum equilibration, thermalization and prethermalization in ultracold atoms*, Nat. Rev. Phys. **2**, 669 (2020).

[10] C. Gogolin and J. Eisert, *Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems*, Rep. Prog. Phys. **79**, 056001 (2016).

[11] E. T. Jaynes, *Information theory and statistical mechanics*, Phys. Rev. **106**, 620 (1957).

[12] W. Grandy, *Principle of maximum entropy and irreversible processes*, Phys. Rep. **62**, 175 (1980).

[13] L. Martyushev and V. Seleznev, *Maximum entropy production principle in physics, chemistry and biology*, Phys. Rep. **426**, 1 (2006).

[14] J. R. Banavar, A. Maritan, and I. Volkov, *Applications of the principle of maximum entropy: From physics to ecology*, J. Phys. Condens. Matter **22**, 063101 (2010).

[15] S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, *Principles of maximum entropy and maximum caliber in statistical physics*, Rev. Mod. Phys. **85**, 1115 (2013).

[16] M. F. Parsons, A. Mazurenko, C. S. Chiu, G. Ji, D. Greif, and M. Greiner, *Site-resolved measurement of the spin-correlation function in the Fermi-Hubbard model*, Science **353**, 1253 (2016).

[17] J. Choi, A. L. Shaw, I. S. Madjarov, X. Xie, R. Finkelstein, 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, *Preparing random states and benchmarking with many-body quantum chaos*, Nature (London) **613**, 468 (2023).

[18] P. Schauß, M. Cheneau, M. Endres, T. Fukuhara, S. Hild, A. Omran, T. Pohl, C. Gross, S. Kuhr, and I. Bloch, *Observation of spatially ordered structures in a two-dimensional Rydberg gas*, Nature (London) **491**, 87 (2012).

[19] M. Foss-Feig, G. Pagano, A. C. Potter, and N. Y. Yao, *Progress in trapped-ion quantum simulation*, Annu. Rev. Condens. Matter Phys. **16**, 145 (2025).

[20] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, *Quantum supremacy using a programmable superconducting processor*, Nature (London) **574**, 505 (2019).

[21] Z. Yan, Z.-Y. Ge, R. Li, Y.-R. Zhang, F. Nori, and Y. Nakamura, *Characterizing many-body dynamics with projected ensembles on a superconducting quantum processor*, Sci. Adv. **12**, eaeb8213 (2026).

[22] J. S. Cotler, D. K. Mark, H.-Y. Huang, F. Hernández, J. Choi, A. L. Shaw, M. Endres, and S. Choi, *Emergent quantum state designs from individual many-body wave functions*, PRX Quantum **4**, 010311 (2023).

[23] M. Ippoliti and W. W. Ho, *Solvable model of deep thermalization with distinct design times*, Quantum **6**, 886 (2022).

[24] W. W. Ho and S. Choi, *Exact emergent quantum state designs from quantum chaotic dynamics*, Phys. Rev. Lett. **128**, 060601 (2022).

[25] H. Wilming and I. Roth, *High-temperature thermalization implies the emergence of quantum state designs*, arXiv:2202.01669.

[26] P. W. Claeys and A. Lamacraft, *Emergent quantum state designs and biunitarity in dual-unitary circuit dynamics*, Quantum **6**, 738 (2022).

[27] H. Shrotriya and W. W. Ho, *Nonlocality of deep thermalization*, SciPost Phys. **18**, 107 (2025).

[28] C. Liu, Q. C. Huang, and W. W. Ho, *Deep thermalization in Gaussian continuous-variable quantum systems*, Phys. Rev. Lett. **133**, 260401 (2024).

[29] T. Bhore, J.-Y. Desaules, and Z. Papić, *Deep thermalization in constrained quantum systems*, Phys. Rev. B **108**, 104317 (2023).

[30] M. Lucas, L. Piroli, J. De Nardis, and A. De Luca, *Generalized deep thermalization for free fermions*, Phys. Rev. A **107**, 032215 (2023).

[31] D. K. Mark, F. Surace, A. Elben, A. L. Shaw, J. Choi, G. Refael, M. Endres, and S. Choi, *Maximum entropy principle in deep thermalization and in Hilbert-space ergodicity*, Phys. Rev. X **14**, 041051 (2024).

[32] A. Chan and A. De Luca, *Projected state ensemble of a generic model of many-body quantum chaos*, J. Phys. A **57**, 405001 (2024).

[33] R.-A. Chang, H. Shrotriya, W. W. Ho, and M. Ippoliti, *Deep thermalization under charge-conserving quantum dynamics*, PRX Quantum **6**, 020343 (2025).

[34] N. D. Varikuti and S. Bandyopadhyay, *Unraveling the emergence of quantum state designs in systems with symmetry*, Quantum **8**, 1456 (2024).

[35] W.-K. Mok, T. Haug, A. L. Shaw, M. Endres, and J. Preskill, *Optimal conversion from classical to quantum randomness via quantum chaos*, Phys. Rev. Lett. **134**, 180403 (2025).

[36] C. Liu, M. Ippoliti, and W. W. Ho, *Coherence-induced deep thermalization transition in random permutation quantum dynamics*, Phys. Rev. Lett. **136**, 100404 (2026).

[37] B. Zhang, P. Xu, X. Chen, and Q. Zhuang, *Holographic deep thermalization for secure and efficient quantum random state generation*, Nat. Commun. **16**, 6341 (2025).

[38] S. Chakraborty, S. Choi, S. Ghosh, and T. Giurgic ă Tiron, *Fast computational deep thermalization*, Phys. Rev. Lett. **135**, 210603 (2025).

[39] M. Bejan, B. Béri, and M. McGinley, *Matchgate circuits deeply thermalize*, Phys. Rev. Lett. **135**, 020401 (2025).

[40] S. Manna, S. Roy, and G. J. Sreejith, *Projected ensemble in a system with locally supported conserved charges*, Phys. Rev. B **111**, 144302 (2025).

[41] C. Vairogs and B. Yan, *Extracting randomness from magic quantum states*, Phys. Rev. Res. **7**, L022069 (2025).

[42] H. Lóio, G. Lami, L. Leone, M. McGinley, X. Turkeshi, and J. De Nardis, *Quantum state designs via magic teleportation*, arXiv:2510.13950.

[43] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, *On the distribution of the wave function for systems in thermal equilibrium*, J. Stat. Phys. **125**, 1193 (2006).

[44] 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).

[45] N. D. Varikuti, S. Bandyopadhyay, and P. Hauke, *Deep thermalization and measurements of quantum resources*, arXiv:2512.09999.

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

[47] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, *Canonical typicality*, Phys. Rev. Lett. **96**, 050403 (2006).

[48] In Refs. [43,44] this ensemble is called the "Gaussian Adjusted Projected (GAP)" ensemble.

[49] R. Jozsa, D. Robb, and W. K. Wootters, *Lower bound for accessible information in quantum mechanics*, Phys. Rev. A **49**, 668 (1994).

[50] The term Scrooge ensemble was coined in Ref. [49]. It is most likely a reference to the fictional character Ebenezer Scrooge from Charles Dickens's 1843 novel A Christmas Carol, as this character was famous for being stingy with his money.

[51] 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).

[52] D. Gross, K. Audenaert, and J. Eisert, *Evenly distributed unitaries: On the structure of unitary designs*, J. Math. Phys. (N.Y.) **48**, 052104 (2007).

[53] Y. Nakata, P. S. Turner, and M. Murao, *Phase-random states: Ensembles of states with fixed amplitudes and uniformly distributed phases in a fixed basis*, Phys. Rev. A **86**, 012301 (2012).

[54] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, *Long-time behavior of macroscopic quantum systems*, Eur. Phys. J. H **35**, 173 (2010).

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

[56] S. Pilatowsky-Cameo, I. Marvian, S. Choi, and W. W. Ho, *Hilbert-space ergodicity in driven quantum systems: Obstructions and designs*, Phys. Rev. X **14**, 041059 (2024).

[57] S. Pilatowsky-Cameo, C. B. Dag, W. W. Ho, and S. Choi, *Complete Hilbert-space ergodicity in quantum dynamics of generalized Fibonacci drives*, Phys. Rev. Lett. **131**, 250401 (2023).

[58] S. Pilatowsky-Cameo, S. Choi, and W. W. Ho, *Critically slow Hilbert-space ergodicity in quantum morphic drives*, Phys. Rev. Lett. **135**, 140402 (2025).

[59] A. L. Shaw, D. K. Mark, J. Choi, R. Finkelstein, P. Scholl, S. Choi, and M. Endres, *Experimental signatures of Hilbert-space ergodicity: Universal bitstring distributions and applications in noise learning*, Phys. Rev. X **15**, 031001 (2025).

[60] W. Liu, Z.-W. Pan, Y. Fu, W. W. Ho, and X. Rong, *Observation of hierarchy of Hilbert space ergodicities in the quantum dynamics of a single spin*, Phys. Rev. Lett. **136**, 020401 (2026).

[61] 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).

[62] M. McGinley and T. Schuster, *The Scrooge ensemble in many-body quantum systems*, arXiv:2511.17172.

[63] S. Ghosh, A. Mirani, Y. Quek, and M. Xu, *Design boosters: From constant-time quantum chaos to ∞-designs and beyond*, arXiv:2511.08543.

[64] T. Schuster, J. Haferkamp, and H.-Y. Huang, *Random unitaries in extremely low depth*, Science **389**, 92 (2025).

[65] H.-Y. Huang, R. Kueng, and J. Preskill, *Predicting many properties of a quantum system from very few measurements*, Nat. Phys. **16**, 1050 (2020).

[66] M. McGinley and M. Fava, *Shadow tomography from emergent state designs in analog quantum simulators*, Phys. Rev. Lett. **131**, 160601 (2023).

[67] M. C. Tran, D. K. Mark, W. W. Ho, and S. Choi, *Measuring arbitrary physical properties in analog quantum simulation*, Phys. Rev. X **13**, 011049 (2023).

[68] J. Preskill, *Lecture Notes for Physics 229: Quantum Information and Computation*, Lecture Notes (California Institute of Technology, Pasadena, CA, 1998).

[69] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, England, 2011).

[70] A. S. Holevo, *Bounds for the quantity of information transmitted by a quantum communication channel*, Probl. Inf. Transm. **9**, 177 (1973).

[71] L. Mao, L. Cui, T. Schuster, and H.-Y. Huang, *Random unitaries that conserve energy*, arXiv:2510.08448 (2025).

[72] The relative error is a strictly stronger notion of approximation: A relative error ϵ implies an additive error ϵ, but the converse is not true.

[73] B. Collins and S. Matsumoto, *Weingarten calculus via orthogonality relations: New applications*, Lat. Am. J. Probab. Math. Stat. **14**, 631 (2017).

[74] B. Collins, S. Matsumoto, and J. Novak, *The weingarten calculus*, Not. Am. Math. Soc. **69**, 1 (2022).

[75] G. Köstenberger, *Weingarten calculus*, arXiv:2101.00921.

[76] In quantum many-body systems, $\|\sigma\|_2 = \sqrt{\mathrm{Tr}(\sigma^2)}$ is often exponentially small in the system size (or, equivalently, inverse power in D). Therefore, we expect the low-purity condition to be valid up to exponentially high moments k.

[77] For brevity, we suppress constants in the definition of $\exp(k)$, writing $\exp(k) = e^{c_1 k^{c_2}}$ for some constants c_1, $c_2 > 0$.

[78] J. Riddell and N. Pagliaroli, *No-resonance conditions, random matrices, and quantum chaotic models*, J. Stat. Phys. **191**, 141 (2024).
[79] This is consistent with the result of recent work by Ghosh *et al.* [63], though there they prove this for the weaker additive error and here we prove this for the stronger relative error.
[80] Z. Webb, *The Clifford group forms a unitary 3-design*, Quantum Inf. Comput. **16**, 1379 (2016).
[81] H. Zhu, *Multiqubit Clifford groups are unitary 3-designs*, Phys. Rev. A **96**, 062336 (2017).
[82] H. Zhu, R. Kueng, M. Grassl, and D. Gross, *The Clifford group fails gracefully to be a unitary 4-design*, arXiv:1609.08172.
[83] This can be derived by a counting argument: Measuring subsystem B of a stabilizer state $|\Psi\rangle_{AB}$ in a stabilizer basis will only generate at most 2^{ℓ} distinct stabilizer states on A, where $\ell \leq N_A$ is the number of ebits shared between A and B. This is insufficient to form approximate 2-designs, which require about 4^{N_A} states.
[84] J. Haferkamp, F. Montealegre-Mora, M. Heinrich, J. Eisert, D. Gross, and I. Roth, *Efficient unitary designs with a system-size independent number of non-Clifford gates*, Commun. Math. Phys. **397**, 995 (2023).
[85] L. Leone, S. F. E. Oliviero, A. Hamma, J. Eisert, and L. Bittel, *Non-Clifford cost of random unitaries*, PRX Quantum **7**, 020321 (2026).
[86] Note that in order to guarantee that the projected ensemble approximates a 2-design, the additive error must satisfy $\varepsilon \ll 1/(D_A D_B)$ since N-qubit stabilizer states form approximate 4-designs with additive error $\varepsilon = \Theta(2^{-N})$ [87,88], yet their projected ensembles only form a 1-design.
[87] R. O. P. Damanik, *Optimality in stabilizer testing, Report*, 2018.
[88] L. Bittel, J. Eisert, L. Leone, A. A. Mele, and S. F. Oliviero, *A complete theory of the Clifford commutant*, Quantum **10**, 2171 (2026).
[89] S. Sugiura and A. Shimizu, *Canonical thermal pure quantum state*, Phys. Rev. Lett. **111**, 010401 (2013).
[90] Y. O. Nakagawa, M. Watanabe, H. Fujita, and S. Sugiura, *Universality in volume-law entanglement of scrambled pure quantum states*, Nat. Commun. **9**, 1635 (2018).
[91] P. Reimann, *Typicality for generalized microcanonical ensembles*, Phys. Rev. Lett. **99**, 160404 (2007).
[92] If H has time-reversal symmetry, ξ_j is a real-valued Gaussian variable.
[93] Z. Davoudi, N. Mueller, and C. Powers, *Towards quantum computing phase diagrams of gauge theories with thermal pure quantum states*, Phys. Rev. Lett. **131**, 081901 (2023).
[94] A. Dymarsky, N. Lashkari, and H. Liu, *Subsystem eigenstate thermalization hypothesis*, Phys. Rev. E **97**, 012140 (2018).
[95] J. R. Garrison and T. Grover, *Does a single eigenstate encode the full Hamiltonian?*, Phys. Rev. X **8**, 021026 (2018).
[96] J. P. Keating, N. Linden, and H. J. Wells, *Spectra and eigenstates of spin chain Hamiltonians*, Commun. Math. Phys. **338**, 81 (2015).
[97] M. Hartmann, G. Mahler, and O. Hess, *Spectral densities and partition functions of modular quantum systems as derived from a central limit theorem*, J. Stat. Phys. **119**, 1139 (2005).
[98] L. Coopmans, Y. Kikuchi, and M. Benedetti, *Predicting Gibbs-state expectation values with pure thermal shadows*, PRX Quantum **4**, 010305 (2023).
[99] C.-F. Chen, M. J. Kastoryano, F. G. S. L. Brandão, and A. Gilyén, *Quantum thermal state preparation*, arXiv:2303.18224.
[100] C.-F. Chen, M. J. Kastoryano, and A. Gilyén, *An efficient and exact noncommutative quantum Gibbs sampler*, arXiv:2311.09207.
[101] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, *Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution*, Nat. Phys. **16**, 205 (2020).
[102] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, *Variational ansatz-based quantum simulation of imaginary time evolution*, npj Quantum Inf. **5**, 75 (2019).
[103] J. Maldacena and X.-L. Qi, *Eternal traversable wormhole*, arXiv:1804.00491.
[104] W. Cottrell, B. Freivogel, D. M. Hofman, and S. F. Lokhande, *How to build the thermofield double state*, J. High Energy Phys. 02 (2019) 058.
[105] These are certain properties of the state or evolution needed for quantum information processing to achieve quantum advantage over classical information processing.
[106] X. Feng, C. Liu, Z. Cheng, W. W. Ho, and M. Ippoliti, *Quantum resource localizability transitions in deep thermalization*, arXiv:2606.08756.
[107] T. Baumgratz, M. Cramer, and M. B. Plenio, *Quantifying coherence*, Phys. Rev. Lett. **113**, 140401 (2014).
[108] A. Streltsov, G. Adesso, and M. B. Plenio, *Colloquium: Quantum coherence as a resource*, Rev. Mod. Phys. **89**, 041003 (2017).
[109] This can be rigorously quantified by so-called relative entropy of coherence $C(|\Psi_0\rangle) = -\sum_z |c_z|^2 \ln(|c_z|^2)$, which is just the Shannon entropy of the populations in the computational basis, if the state $|\Psi_0\rangle$ is pure. Low coherence here means $C = \alpha N$ for some α less than the model-dependent critical value α^*.
[110] X. Feng and M. Ippoliti, *Dynamics of pseudoentanglement*, J. High Energy Phys. 02 (2025) 128.
[111] S. Aaronson, A. Bouland, B. Fefferman, S. Ghosh, U. Vazirani, C. Zhang, and Z. Zhou, *Quantum Pseudoentanglement*, in *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, Leibniz International Proceedings in Informatics (LIPIcs) Vol. 287, edited by V. Guruswami (Schloss Dagstuhl—Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2024), pp. 2:1–2:21.
[112] A. Y. Kitaev, *Fault-tolerant quantum computation by anyons*, Ann. Phys. (Amsterdam) **303**, 2 (2003).
[113] S. Bravyi and A. Kitaev, *Universal quantum computation with ideal Clifford gates and noisy ancillas*, Phys. Rev. A **71**, 022316 (2005).
[114] Z.-W. Liu and A. Winter, *Many-body quantum magic*, PRX Quantum **3**, 020333 (2022).

[115] G. Lami, T. Haug, and J. De Nardis, *Quantum state designs with Clifford-enhanced matrix product states*, PRX Quantum **6**, 010345 (2025).
[116] R. Jozsa and N. Linden, *On the role of entanglement in quantum-computational speed-up*, Proc. Math. Phys. Eng. Sci. **459**, 2011 (2003).
[117] S. Aaronson and D. Gottesman, *Improved simulation of stabilizer circuits*, Phys. Rev. A **70**, 052328 (2004).
[118] H. Thomas, P.-E. Emeriau, R. Mezher, E. Kashefi, H. Ollivier, and U. Chabaud, *Role of coherence for quantum computational advantage*, Phys. Rev. Lett. **135**, 150602 (2025).
[119] D. Shepherd and M. J. Bremner, *Temporally unstructured quantum computation*, Proc. R. Soc. A **465**, 1413 (2009).
[120] M. J. Bremner, R. Jozsa, and D. J. Shepherd, *Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy*, Proc. R. Soc. A **467**, 459 (2010).
[121] These are closely related to instantaneous quantum polynomial (IQP) circuits [119,120] studied in quantum complexity theory.
[122] X. Ni and M. V. den Nest, *Commuting quantum circuits: Efficient classical simulations versus hardness results*, Quantum Inf. Comput. **13**, 54 (2013).
[123] More generally, we may also apply an arbitrary unitary U_A on A, without affecting the discussion.
[124] Y. Nakata, M. Koashi, and M. Murao, *Generating a state t-design by diagonal quantum circuits*, New J. Phys. **16**, 053043 (2014).
[125] L. Leone, S. F. E. Oliviero, Y. Zhou, and A. Hamma, *Quantum chaos is quantum*, Quantum **5**, 453 (2021).
[126] T. Haug, L. Aolita, and M. Kim, *Probing quantum complexity via universal saturation of stabilizer entropies*, Quantum **9**, 1801 (2025).
[127] P. S. Tarabunga and T. Haug, *Efficient mutual magic and magic capacity with matrix product states*, SciPost Phys. **19**, 085 (2025).
[128] P. Pfeuty, *The one-dimensional Ising model with a transverse field*, Ann. Phys. (N.Y.) **57**, 79 (1970).
[129] T. Haug and L. Piroli, *Quantifying nonstabilizerness of matrix product states*, Phys. Rev. B **107**, 035148 (2023).
[130] A. Osterloh, L. Amico, G. Falci, and R. Fazio, *Scaling of entanglement close to a quantum phase transition*, Nature (London) **416**, 608 (2002).
[131] M. Walschaers, *Non-Gaussian quantum states and where to find them*, PRX Quantum **2**, 030204 (2021).
[132] K.-D. Wu, T. V. Kondra, S. Rana, C. M. Scandolo, G.-Y. Xiang, C.-F. Li, G.-C. Guo, and A. Streltsov, *Operational resource theory of imaginarity*, Phys. Rev. Lett. **126**, 090401 (2021).
[133] A. Milekhin and S. Murciano, *Observable-projected ensembles*, Quantum **9**, 1888 (2025).
[134] X.-H. Yu, W. W. Ho, and P. Kos, *Mixed state deep thermalization*, Phys. Rev. Lett. **135**, 260402 (2025).
[135] A. W. Harrow, *The church of the symmetric subspace*, arXiv:1308.6595.