Measuring energy by measuring any other observable

Dominik Šafránek [1,*] and Dario Rosa [1,2,†]

[1]*Center for Theoretical Physics of Complex Systems, Institute for Basic Science, Daejeon 34126, Korea*
[2]*Basic Science Program, Korea University of Science and Technology, Daejeon 34113, Korea*

We present a method to estimate the probabilities of outcomes of a quantum observable, its mean value, and higher moments by measuring any other observable. This method is general and can be applied to any quantum system. In the case of estimating the mean energy of an isolated system, the estimate can be further improved by measuring the other observable at different times. Intuitively, this method uses interplay and correlations between the measured observable, the estimated observable, and the state of the system. We provide two bounds: one that is looser but analytically computable and one that is tighter but requires solving a nonconvex optimization problem. The method can be used to estimate expectation values and related quantities such as temperature and work in setups where performing measurements in a highly entangled basis is difficult, finding use in state-of-the-art quantum simulators. As a demonstration, we show that in Heisenberg and Ising models of ten sites in the localized phase, performing two-qubit measurements excludes 97.5% and 96.7% of the possible range of energies, respectively, when estimating the ground-state energy.

I. INTRODUCTION

Expectation values are ubiquitous in quantum physics, characterizing different types of behaviors of quantum systems. They are used as both descriptive and predictive tools and include the following examples. Mean values of generic local observables classify many-body systems according to how well they thermalize [1,2]. Vanishing total magnetization identifies a quantum phase transition [3–5]. The mean value of homodyne measurement [6–9] is evaluated in magnetic resonance imaging [10] and quantum cryptography protocols [11], while its variance is used to prove squeezing [12,13], an essential resource for quantum sensors [14]. Variances also appear in Heisenberg's uncertainty principle [15–17]. Expectation values are of interest in quantum field theory [18,19] and in nuclear physics [20,21].

Moments of energy are somewhat special due to their wide range of applications. The mean energy determines the thermodynamic entropy of the system [22–25] and its temperature [26–28]. Its change may represent heat and work [29–34] and its difference defines a measure of extractable work called ergotropy [35–37]. Variance in energy determines the precision in estimating both the time [38] and temperature [39] parameters. Both moments, when combined, provide a tight bound on the characteristic timescale of a quantum system [40–42].

Given the breadth of applications, it is clear that measuring and estimating expectation values are of essential importance, which may however be challenging. For example, in quantum many-body systems, the mean energy is considerably

difficult to measure, with only a few architecture-specific proposals [43] and experiments [44] known. This is because energy eigenstates are typically highly entangled. In quantum simulators, measuring in an entangled basis is performed by combining several elementary gates. Each gate has a fixed fidelity, and when many of such gates are combined, the fidelity is diminished, making such measurements unreliable [45–48]. Additionally, experimental setups may allow measurement only of a close but not exact observable we are interested in. This is the case, for example, in the aforementioned homodyne detection with a finite, instead of infinite, oscillator field strength [7,49].

In this paper we show that performing any measurement bounds the probabilities of outcomes, the mean value, and higher moments of any other observable. This means that, quite counterintuitively, measurements carry more information than previously known. Any observable yields some information on any other observable. The method uses correlations between the measured, the estimated observable, and the state of the system. It is precisely this interplay that allows us to bound the probability of outcomes of the estimated observable and from those its mean value and higher moments. These results immediately ameliorate the issue mentioned above: Even in experimental systems in which we have only a limited ability to measure, we can perform the best possible measurement, and this is enough to estimate the probability distribution of outcomes and the mean value of an observable that we are truly interested in measuring.

The derived bounds are further tightened by measuring in different bases and, in the case of estimating the mean energy, by measuring at different times. After some preliminaries, we show how measurement in any basis bounds the probabilities of the system to have a certain mean energy. From this we derive two bounds on the mean value of energy: One analytic,

*dsafranekibs@gmail.com
†dario_rosa@ibs.re.kr

which is easy to compute, and one tighter, which leads to an optimization problem. We discuss situations in which the analytic bound becomes relatively tight. Then we describe a few differences when bounding the mean values of observables other than energy. We illustrate this method on several experimentally relevant models. Finally, we discuss the advantages and drawbacks of this method, possible applications, and future directions.

II. RESULTS

A. Setup

Consider any quantum system and measurement given by the measurement basis $\mathcal{C} = \{|i\rangle\}$. Label i is the outcome of the measurement, and the probability of obtaining the outcome at time t is $p_i = |\langle i|\psi_t\rangle|^2$, where $|\psi_t\rangle$ is the state of the system at time t. If we create many realizations of the same experiment by repeating the prepare-and-measure sequence, we can build the statistics of outcomes and determine the probability distribution $\{p_i\}$. Thus, these probabilities are experimentally accessible.

Next we consider a Hamiltonian of the system, with spectral decomposition in terms of its eigenvalues and eigenvectors being $\hat{H} = \sum_E E|E\rangle\langle E|$. The probability of finding the system having energy E is given by $p_E = |\langle E|\psi_t\rangle|^2$. We assume to know the Hamiltonian and its spectral decomposition. However, we presume that we are unable to measure it experimentally. In other words, we cannot perform the measurement in the energy eigenbasis. As we will show next, this does not stop us from estimating its probability of outcomes and from those also the mean value of energy. Proofs for the following bounds can be found in Appendix A.

B. Bounds on energy probabilities

The key result of this paper is that we find the probability of the state having energy E between two bounds

$$a_E \leqslant p_E \leqslant b_E, \quad (1)$$

where we define

$$a_E = \max\left\{2\sum_i c_{iE}^2 - b_E, \ 0\right\},$$

$$b_E = \left(\sum_i c_{iE}\right)^2, \quad (2)$$

and $c_{iE} = \sqrt{p_i}|\langle i|E\rangle|$ (see Fig. 1). The last element contains both the probability of an outcome p_i and the correlation between the measured and the estimated observable, given by overlap $|\langle i|E\rangle|$. Thus, the above inequality connects the probability of the estimated observable to the probability of the measured observable, through the correlations between their eigenbases. We can easily derive that $b_E \leqslant 1$ from the Cauchy-Schwarz inequality. Thus, the upper bound on the energy probability is always nontrivial.

If the system is isolated, it evolves unitarily with the time-independent Hamiltonian $\hat{H}$, and energy probabilities p_E are also time independent. In contrast, probabilities p_i are time dependent if the measurement basis does not commute with
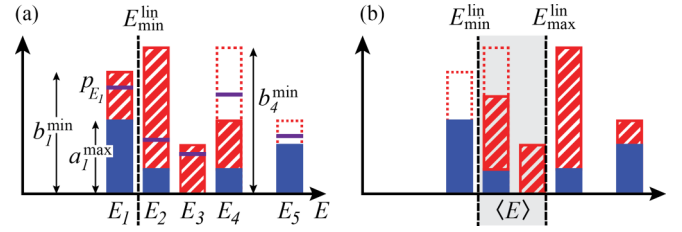

FIG. 1. Bounds on energy probabilities and the mean energy. Each energy probability p_{E_l} (purple lines) is bounded by $a_l^{\max}$ from below (blue bar) and by $b_l^{\min}$ (red bar) from above [see Eq. (3)]. The analytical bound on the mean energy $E_{\min}^{\lin} \leqslant \langle E \rangle \leqslant E_{\max}^{\lin}$ [Eq. (8)] is computed as follows. Imagine a bottle of probabilities with volume 1. To obtain (a) the lower bound on the mean energy, we pour the probabilities on the graph from the bottle and fill the minimal probability of each energy given by $a_l^{\max}$ (blue solid bar). We will have some probabilities left in the bottle, so we pour the rest and top the red bars up to their maximum $b_l^{\min}$ (red striped bar), starting from the lowest to the highest energy, until we run out of probabilities in the bottle, i.e., until all the probabilities on the graph sum up to 1. Taking the mean value of such a distribution will yield the lower bound [see Eq. (11)]. (b) The upper bound is obtained by the same method but going from the highest to the lowest energy [see Eq. (13)]. The mean energy $\langle E \rangle$ lies somewhere in the shaded region.

the Hamiltonian, and so are the bounds a_E and b_E. This leads to an interesting observation that measuring at different times can make the bound tighter. Quantitatively, we have

$$a_E^{\max} \leqslant p_E \leqslant b_E^{\min}, \quad (3)$$

where $a_E^{\max} = \max_{t \in [0,T]} a_E$ and $b_E^{\min} = \min_{t \in [0,T]} b_E$.

Let us discuss situations in which the bound becomes tight. First, assume that one of the measurement basis vectors is an energy eigenvector, i.e., $|i\rangle = |E\rangle$ for some i and E. Then the bound gives $p_E = p_i$ for this specific E, as intuitively expected. Second, consider a situation in which we always obtain a single outcome when measuring at a specific time, i.e., $p_i = 1$ for some i at some time t. Doing this is akin to identifying the state of the system as being equal to $|i\rangle$. As a result, the bound gives an identity $p_E = |\langle i|E\rangle|$ for all energies E so the entire energy distribution is determined exactly.

The two extreme cases just discussed suggest two possible scenarios in which the bounds perform well. The bounds are relatively tight when either the measurement basis resembles the eigenbasis of the estimated observable (in this case, the Hamiltonian) or the state of the system comes close to one of the measurement basis vectors during its time evolution.

In addition to optimization over time, the inequalities can be further tightened by performing measurements in different bases. Defining a set of performed measurements $\mathcal{M} = \{\mathcal{C}_m\}$, each measurement bounds the p_E independently, so we can take

$$a_E^{\max} = \max_{\mathcal{C}_m \in \mathcal{M}, t \in [0,T]} a_E^m, \quad b_E^{\min} = \min_{\mathcal{C}_m \in \mathcal{M}, t \in [0,T]} b_E^m \quad (4)$$

for the bound (3). The $a_E^m(t)$ and $b_E^m(t)$ are defined by Eqs. (2) for each measurement $\mathcal{C}_m$. This may be helpful when there are limits on the types of measurements we can perform. For example, we can be experimentally limited to using only one-

and two-qubit gates due to many-qubit gates having a low fidelity.

C. Bounds on collections of energy probabilities

Additionally, we derive the following collective bounds on the energy probabilities:

$$\sqrt{\boldsymbol{p}}^T \boldsymbol{A}_i \sqrt{\boldsymbol{p}} \leqslant p_i \leqslant \sqrt{\boldsymbol{p}}^T \boldsymbol{B}_i \sqrt{\boldsymbol{p}}. \qquad (5)$$

The left and right-hand sides are time-independent quadratic forms, which are defined by their elements as

$$
\begin{aligned}
(\boldsymbol{A}_i)_{EE'} &= (-1)^{1+\delta_{EE'}} |\langle E|i\rangle\langle i|E'\rangle|, \\
(\boldsymbol{B}_i)_{EE'} &= |\langle E|i\rangle\langle i|E'\rangle|.
\end{aligned}
\qquad (6)
$$

They are applied on the vector of the square root of energy probabilities $\sqrt{\boldsymbol{p}}_E = \sqrt{p_E}$, which are those that we would like to estimate.

For the Hamiltonian evolution, extremizing over times of measurement yields tighter bounds

$$\sqrt{\boldsymbol{p}}^T \boldsymbol{A}_i \sqrt{\boldsymbol{p}} \leqslant p_i^{\min} \leqslant p_i^{\max} \leqslant \sqrt{\boldsymbol{p}}^T \boldsymbol{B}_i \sqrt{\boldsymbol{p}}, \qquad (7)$$

where $p_i^{\min} = \min_{t\in[0,T]} p_i$ and $p_i^{\max} = \max_{t\in[0,T]} p_i$. These collective inequalities are generally nonlinear in p_E and neither convex nor concave. There are as many as the number of measurement outcomes. While they do not bound each energy probability separately, they provide relationships between their respective sizes. For example, we can derive quantitative statements of the following type: If p_{E_1} is high, then p_{E_2} must be low. They might require numerical methods to work with due to their nonlinearity. However, they can provide a robust improvement in estimating energy in some cases. See Appendix B for such an example of coarse-grained energy measurements.

Similar to Eq. (4), we can employ measurements in different bases. These generate more conditions for probabilities p_E of type (7), thus making quantitative relations between them stricter.

D. Bounds on the mean energy

Given the derived bounds on the probability distribution of energy, we can bound the mean energy of the system as

$$E_{\min}^{\text{lin}} \leqslant E_{\min} \leqslant \langle E \rangle \leqslant E_{\max} \leqslant E_{\max}^{\text{lin}}. \qquad (8)$$

The inner bound is tighter but may be challenging to compute. The outer bound is looser, but it is analytically computable.

The inner bound is computed by optimizing the mean value of energy $\langle E \rangle = \sum_E E p_E$ as

$$E_{\min} = \min_{\{p_E\}\in S} \langle E \rangle, \quad E_{\max} = \max_{\{p_E\}\in S} \langle E \rangle \qquad (9)$$

over the set of probability distributions consistent with our observations, i.e., over the set that satisfies all the required inequalities

$$S = \left\{ \{p_E\} \middle| \sum_E p_E = 1, \text{ Eq. (3), and Eq. (7)} \right\}. \qquad (10)$$

The mean energy itself is a linear function. While $\sum_E p_E = 1$ and Eq. (3) are linear constraints, Eq. (7) is in general nonlinear. Computing $E_{\min}$ and $E_{\max}$ is therefore a nonlinear

constrained optimization problem. These problems are considered to be computationally demanding, although they are difficult to characterize within computational complexity theory [50]. They can be solved only approximately by various methods.[1] The time to find an exact solution typically scales exponentially with the number of variables, in our case, the dimension of the system.

However, we can solve an easier problem by including only linear constraints in the optimization, i.e., by removing the requirement for satisfying Eq. (7). This makes the bound looser but allows for solving this optimization problem analytically. The reasoning behind the following derivation is explained in Fig. 1 and performed in detail in Appendix A. We assume that energy eigenvalues are ordered in increasing order as $E_j \leqslant E_{j+1}$, with E_1 representing the ground-state energy. We have

$$E_{\min}^{\text{lin}} = \sum_{j=1}^N E_j u_j, \qquad (11)$$

where probabilities $u_j \equiv u_{E_j}$ are computed recursively starting from the ground state as

$$u_1 = \min\left\{ b_1^{\min}, 1 - \sum_{l=2}^N a_l^{\max} \right\},$$

$$u_j = \min\left\{ b_j^{\min}, 1 - \sum_{l=1}^{j-1} u_l - \sum_{l=j+1}^N a_l^{\max} \right\}, \quad 2 \leqslant j \leqslant N. \qquad (12)$$

(We simplified lower indices as $l \equiv E_l$ and the dimension of the system is N.) Similarly, we obtain

$$E_{\max}^{\text{lin}} = \sum_{j=1}^N E_j w_j, \qquad (13)$$

where, starting from the highest-energy state, we have

$$w_N = \min\left\{ b_N^{\min}, 1 - \sum_{l=1}^{N-1} a_l^{\max} \right\},$$

$$w_j = \min\left\{ b_j^{\min}, 1 - \sum_{l=j+1}^N w_l - \sum_{l=1}^{j-1} a_l^{\max} \right\},$$

$$1 \leqslant j \leqslant N - 1. \qquad (14)$$

Bounds for the higher moments are computed by replacing eigenvalues E with E^k in Eq. (9).

E. Tightness of the analytic bound on the mean value of energy

Below Eq. (3), we discussed two cases where the bound on the energy probability distribution becomes tight. Now we

---

[1] These are, for example, the Nelder-Mead algorithm [51], random search [52], differential evolution [53], machine learning methods [54], simulated [55] or quantum [56] annealing, and modified linear programming methods [57].

show that the same arguments can also be extended to discuss the tightness of the bound on the mean energy [Eq. (8)].

The first case is when the state of the system comes close to one of the measurement basis vectors during its time evolution. If the state becomes exactly one of the basis vectors, i.e., $p_i(t) = 1$ at some time t, we can identify the state exactly as $|i\rangle$ and so all of its properties, energy included. In this case, bounds (3) become tight and we obtain $\langle E \rangle = E_{\max}^{\text{lin}} = E_{\min}^{\text{lin}}$. The most informative times of measurement are those of a low value of observational entropy [25,58–62], due to the state wandering into a small subspace of the Hilbert space [25,63] recognizable by the measurement. This can be advantageous for energy estimation in systems exhibiting recurrences and Loschmidt echo [64–68], which return close to their original state after some time.

The second case is when the measurement is close to the energy measurement itself. This happens, for example, in the localized phase of many-body localized systems, in which the energy eigenvectors tend to localize in small portions of the Fock space [69]. Thus, measuring local particle numbers is almost as good as measuring the energy itself. This is mathematically justified below Eq. (3).

F. Choosing the time interval and times of measurements

In experimental settings, the system can be evolved only over a finite time. Within this time interval, a measurement can be performed only a finite number of times. Thus, it is useful to specify the criteria for which time T the system should be evolved together with the corresponding times of measurement, for the time optimization (4) to work at its best.

We can address this heuristically given the points introduced in the preceding section. Generally, the ideal number and times of measurements depend on the initial state: If the state does not evolve much, or at all, which is the case for any energy eigenstate (such as the ground state), then only a single measurement is required. Additional measurements will not yield any improvement.

On the other hand, if a nontrivial evolution occurs, then more times of measurement might be advantageous. The rule of thumb is to measure for as long as the observational entropy related to the measurement grows until it reaches its equilibrium value. This is because bigger dips in observational entropy give more information, while small dips do not provide as much. The same criterion could be applied to identify the times of measurement within this interval. There should be as many as to reproduce the medium-size dips in the observational entropy evolution.

G. Bounding the mean values of observables other than energy

The derivations and results above can be repeated as they are for any observable that commutes with the Hamiltonian. In that context, E would denote an eigenvalue of an observable other than energy. For observables that do not commute with the Hamiltonian, the procedure can be repeated, but it must be performed only at a fixed time t (extremization over time [Eqs. (3) and (7)] is not possible). Extremization over different measurements at a fixed time [Eqs. (4)] can be employed. To summarize, the estimation of observables other than energy

works exactly the same as estimating energy, with the only difference that optimization over time can be performed only when the observable commutes with the Hamiltonian. If the observable does not commute with the Hamiltonian, then also its expectation value changes in time, so only a specific time must be chosen, but everything else proceeds identically.

H. Demonstration on experimentally relevant many-body systems

We numerically demonstrate this method on the paradigmatic example of the one-dimensional disordered Heisenberg model [70–72]. Numerical experiments for other experimentally achieved models, i.e., Ising [73–76], XY [77–80], and PXP models [81–83], are presented in Appendix E. A simple analytical example is presented in Appendix D.

The Hamiltonian is given by

$$\hat{H} = \sum_i \left(\hat{\sigma}_i^x \hat{\sigma}_{i+1}^x + \hat{\sigma}_i^y \hat{\sigma}_{i+1}^y + \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right) + \sum_i h_i \hat{\sigma}_i^z, \quad (15)$$

where $\hat{\sigma}_i^a$, $a = x, y, z$, are the Pauli operators acting at the site i. The constants h_i are randomly extracted within the interval $[-W, W]$, with W the disorder strength. We show the case $W = 0.5$ for the chaotic (delocalized) regime and $W = 10$ for the localized regime [71,72]. See Appendix E for the Bethe integrable regime $W = 0$ [84].

We choose a complete measurement in the local number basis,

$$\mathcal{C} = \{|i_1\rangle \otimes \cdots \otimes |i_L\rangle\}_{i_1,\ldots,i_L}, \quad (16)$$

for small systems simulations. There $i_j = 0, 1$ for all j and L is the length of the chain (such that the dimension of the Hilbert space is $N = 2^L$). For example, for the chain of $L = 3$ sites, the measurement basis is

$$\mathcal{C} = \{|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle\}. \quad (17)$$

This is an example of a one-local measurement, meaning that the measurement basis does not consist of states entangled between two or more sites. For large-system simulations, we also add optimized k-local measurements; k-local measurements are those that project onto states that are allowed to be entangled between k neighboring sites. An example of a two-local measurement on the chain of $L = 4$ sites is the measurement in the local Bell basis,

$$\mathcal{C} = \{|\Phi^+\rangle|\Phi^+\rangle, |\Phi^+\rangle|\Phi^-\rangle, |\Phi^+\rangle|\Psi^+\rangle, |\Phi^+\rangle|\Psi^-\rangle, \ldots\}, \quad (18)$$

where $|\Phi^\pm\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$ and $|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$.

We consider three types of initial states. First is the ground state (G). The second is a pure thermal state (C),

$$|\psi_\beta\rangle = \frac{1}{\mathcal{N}} \sum_E e^{-\beta E/2} |E\rangle, \quad (19)$$

where $\mathcal{N}$ is the normalization factor and the inverse temperature is chosen as $\beta = 6/(E_N - E_1)$. We take this choice of β to imitate a cold state. Third, we randomly choose a pure state from the Hilbert space with the Haar measure (H), imitating an
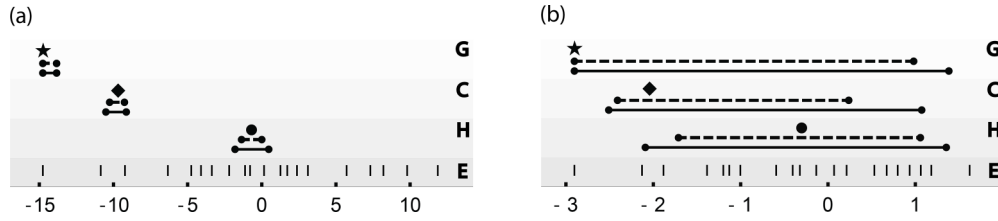
FIG. 2. Heisenberg model in the localized ($W = 0.5$) and the delocalized ($W = 10$) phases, with three particles on six sites. The energy is estimated by measuring in the local number basis (corresponding to $k = 0$ in Fig. 3). The initial state is either a ground state (G), a cold state (C), Eq. (19), or a hot state (H). The graphs show the true mean energy $\langle E \rangle$ (single symbols: Star, diamond, and disk for G, C, and H, respectively) and intervals of analytic bounds $[E_{\min}^{\lin}, E_{\max}^{\lin}]$ (solid lines) and numerical bounds $[E_{\min}, E_{\max}]$ (dashed lines) for each state. We also plot the list of energy eigenvalues at the very bottom (E). In the localized phase, energy eigenstates have a large overlap with the local number basis, making the energy estimation significantly more precise. (a) Localized (b) Delocalized.

infinite-temperature state. To compute the bounds, we evolve them with the Hamiltonian for the total time of $T = 160$.

In Fig. 2 we show estimates of energy in small systems, taking three particles on the chain of $L = 6$ sites. The Heisenberg Hamiltonian is particle conserving, so the initial state explores only a subspace of the full Hilbert space. We analytically solve for the looser bound using Eqs. (11) and (13). This solution serves as a starting point for the COBYLA optimization algorithm [57,85] to compute the tighter bound (9). In Fig. 3 we plot estimates of energy in a large system, with

five particles on $L = 10$ sites, in which computing the numeric bound is prohibitively difficult. Instead, we add more measurements and calculate the corresponding analytical bounds in the increasing degree of nonlocality.

I. Generalization to mixed states and positive-operator-valued measures

Most general quantum measurements are represented by the positive-operator-valued measure (POVM) $\mathcal{C} = \{\hat{\Pi}_i\}_i$
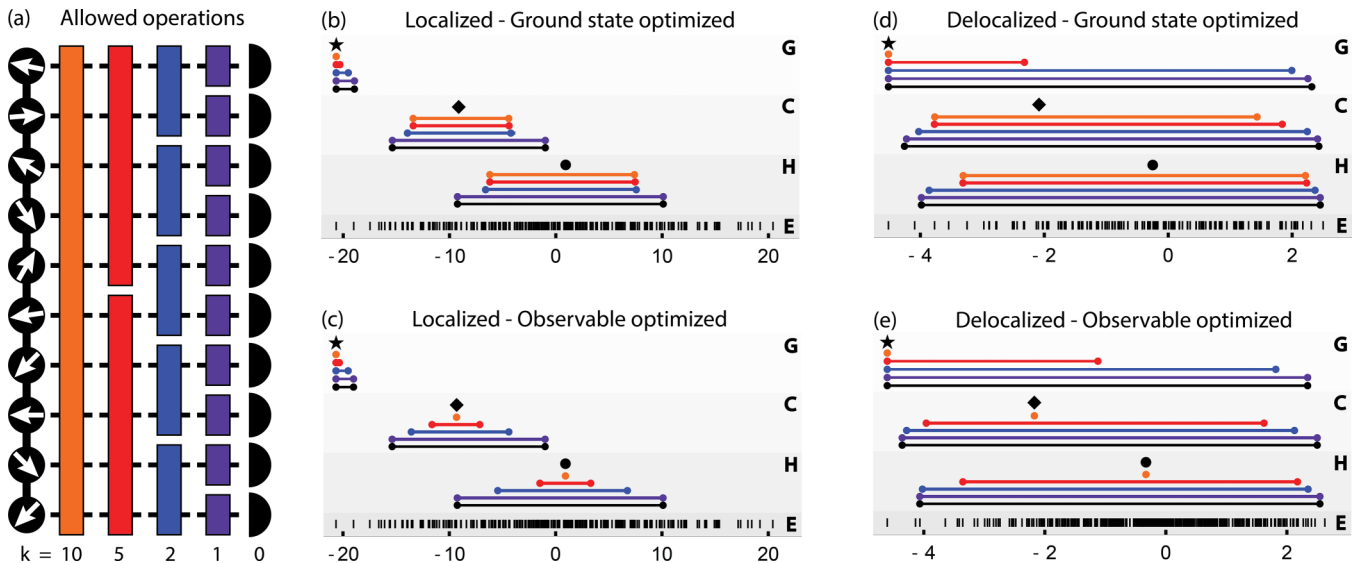

FIG. 3. Heisenberg model in (b) and (c) the localized ($W = 0.5$) and (d) and (e) the delocalized ($W = 10$) phases, with five particles on ten sites. The energy is estimated with progressively added optimized k-local measurements, with the same initial states as in Fig. 2 (G, C, and H) and energy eigenvalues plotted at the very bottom (E). In large systems, we can no longer employ numerical methods to improve the analytic bound. Instead, we employ different measurements. (a) Sketch of allowed operations. Here $k = 0$ (black) corresponds to measuring in the local number basis; $k = 1$ (purple) allows for using single-site unitary operators before measuring in the local number basis, which leads to a general single-site measurement; $k = 2$ (blue) allows using two-site local operators, etc. (b) and (d) In the ground-state-optimized method, the measurement basis is given by the eigenbasis of the k-local reduced state of the ground state [see Eqs. (A1) and (A2)]. Intervals of analytic bounds $[E_{\min}^{\lin(k)}, E_{\max}^{\lin(k)}]$ are labeled with the colors matching the allowed operations. (c) and (e) In the observable-optimized method, the experimenter measures k-local blocks of the Hamiltonian. When $k = 10$, the observable-optimized method achieves perfect energy estimation for all initial states ($k = 10$ corresponds to measuring the Hamiltonian itself). In contrast, when $k = 10$, the ground-state-optimized method achieves that only for the ground state ($k = 10$ corresponds to measuring in a basis that contains the ground state). Using two-qubit measurements ($k = 2$) in the localized phase, the bound excludes 97.5% of the possible range of energies when estimating the ground-state energy. See Appendix A for the descriptions of the ground-state-optimized and observable-optimized methods. See Appendix E for details and specifically Tables II and III for additional simulations of experimentally achievable XY, PXP, and Ising models.

satisfying the completeness relation $\sum_i \hat{\Pi}_i = \hat{I}$, where $\hat{\Pi}_i$ is a positive-semidefinite operator called a POVM element. For a density matrix $\hat{\rho}$ representing the state of the system, the probability of obtaining measurement outcome i is given by $p_i = \text{tr}(\hat{\Pi}_i \hat{\rho})$.

Positive-operator-valued measure elements admit a spectral decomposition $\hat{\Pi}_i = \sum_k \gamma_i^k |i^k\rangle\langle i^k|$. There $0 < \gamma_i^k \leqslant 1$ and $|i^k\rangle$ are orthogonal to each other for different k's. We define its volume as $V_i = \text{tr}(\hat{\Pi}_i)$. We further define $x_i^E = \min_k \gamma_i^k |\langle i^k|E\rangle|^2$, $y_i^E = \max_k |\langle i^k|E\rangle|^2$, and $\gamma_i = \min_k \gamma_i^k$. Note that these extrema are taken only over k for which γ_i^k is positive, i.e., nonzero.

The results of this paper generalize to mixed states and general measurements by taking

$$a_E = \max\left\{ \sum_i p_i\left(x_i^E + \gamma_i y_i^E\right) - \left(\sum_i \sqrt{p_i y_i^E V_i}\right)^2, 0 \right\},$$

$$b_E = \left(\sum_i \sqrt{p_i \langle E|\hat{\Pi}_i|E\rangle}\right)^2 \qquad (20)$$

in Eq. (1) and by taking

$$(A_i)_{EE'} = (-1)^{1+\delta_{EE'}}|\langle E|\hat{\Pi}_i|E'\rangle|,$$

$$(B_i)_{EE'} = |\langle E|\hat{\Pi}_i|E'\rangle| \qquad (21)$$

in Eq. (6). (See Appendix A for the proofs.) Results that come after do not depend on the specific form of the bounds. Thus these proceed identically. For a complete projective measurement, we have $\hat{\Pi}_i = |i\rangle\langle i|$, which implies $x_i^E = y_i^E = |\langle i|E\rangle|^2$ and $\gamma_i = V_i = 1$, from which the initial bounds easily follow.

III. DISCUSSION

Quantum measurements provide more information than one would initially think. We developed a method that allows us to measure one observable and predict bounds on the distribution of outcomes and expectation values of every other observable. In this method, it is assumed that we have enough copies of the initial state so we can determine the entire probability distribution of outcomes of the measured observable. The method works well either when the measured and the estimated observables resemble each other or when the system state is close to one of the measurement basis states. In those cases, the bounds will be very tight. On the other hand, if the measurement cannot distinguish between two eigenstates with very different eigenvalues and the system state has considerable overlap with one of them, the method naturally cannot give a good estimate. However, this can be overcome by combining measurements in different bases. Additionally, when estimating conserved quantities, better estimates are obtained by measuring at different times.

It is interesting to compare the presented method with the recent work of [86]. There an algorithm was provided in which it is possible to approximate the mean value of an observable with high probability by applying random measurements, called classical shadows. This idea has been extended in subsequent literature both theoretically [87–96] and experimentally [97,98].

In classical shadow literature, it is assumed that performing any type of measurement is possible. These measurements are sampled randomly from a tomographically complete set, meaning that with this set of measurements, quantum tomography is possible. The goal is to estimate the expectation value of an observable and achieve an error lower than ϵ, with as few measurements as possible. In other words, it is assumed that only a limited number of copies of the state are available to be measured.

In contrast, in the method presented here, it is assumed that only a single type of measurement can be performed. This measurement has been chosen from a limited set of measurements that are experimentally available. At the same time, it is assumed that infinitely many copies of the initial state are available. Thus, we can perform as many repetitions of the same measurement as necessary to fully specify its probability distribution.

In our method, since we do not sample from a tomographically complete set, we always have a finite error in estimating the expectation value. This error comes from the misalignment of the chosen measurement basis and the estimated observable and the eigenbasis of the density matrix.

Thus, while attempting to address a very similar goal, the two approaches are different in their assumptions and outcomes. Our method shines in exactly those situations in which not every measurement can be performed. This is motivated by the experimental capabilities of current state-of-the-art quantum simulators, which allow for the application of only one and possibly two-qubit gates. For this reason, we focused on local measurements. Of course, the method will work better with the improvement of the experimental capabilities. Its main strength, though, is to be able to give a prediction of the mean value of observables together with its error even in cases when the experimental capabilities are very low and other methods cannot be used.

We argued for using this method for estimating moments of energy, which have a wide range of applications while being difficult to measure directly in many-body systems. For instance, using this method, one can bound the characteristic timescale through the Mandelstam-Tamm and Margolus-Levitin bounds [40–42] to estimate the amount of extractable work from an unknown source of states [37] or to estimate temperature. Moreover, the latest can be used to benchmark the cooling function of quantum annealers [99–102] and adiabatic quantum computers [103]. This could be particularly suitable for systems with area-law entanglement scaling [69,104], in which local measurements should be more powerful given the absence of long-range correlations in the eigenstates of such systems. We confirmed this numerically in two gapped models, which are proven to have area-law ground states [105]. In these, estimating the ground-state energy using only local measurements works especially well. The method can be used equally well and proceeds identically for estimating observables other than energy, with the only exception that if such an observable does not commute with the Hamiltonian, then also its mean value is time dependent and therefore one has to pick a specific time for the analysis. (In contrast, when measuring energy, we could improve the bounds by measuring at different times, using that its mean value is conserved.) The method can also be

used to prove entanglement through an entanglement witness (operator $\hat{A}$), without measuring the witness itself: To prove entanglement, it is enough to show that the expectation value of this operator is negative [106–108].

On a theoretical ground, this research instigated possible paths of exploration. How to choose a measurement, given some restriction, for example, on its locality or the number of elementary gates it consists of, that leads to the tightest possible bound? Is it possible to apply machine learning models to find this optimal strategy? Will the bound give an exact value when the set of measurements is tomographically complete? Can this method be modified to identify the properties of channels instead of states? Given that this method bounds the entire probability distribution of outcomes, is it possible to modify it to calculate estimates of functions of a state other than expectation values, such as entanglement entropy?

Data and the code used to generate data for this paper are available from the corresponding authors upon reasonable request.

APPENDIX A: METHODS AND PROOFS

1. Ground-state-optimized and observable-optimized methods for finding appropriate *k*-local measurements

We choose a chain of length $L = 10$ and two-local ($k = 2$) measurements to illustrate.

a. Ground-state-optimized method

This method is inspired by the matrix product state ansatz [109] and by the correspondence between observational and entanglement entropy [110]. We choose a pure state (in our case, the ground state) $|\psi_0\rangle$ that we want to optimize. We divide the chain into L/k local parts and generate the local basis as an eigenbasis of the reduced state. This corresponds to the local Schmidt basis. For example, for the first two sites referred to as subsystem A_1 (while denoting the full system by A), we have

$$\hat{\rho}_{A_1} = \text{tr}_{A \setminus A_1} |\psi_0\rangle\langle\psi_0|, \qquad \text{(A1)}$$

from which we compute eigenvectors $\{|\psi_{i_1}^{A_1}\rangle\}$. Equation (A1) is what we refer to as the *k*-local reduced state of the ground state. The final ground-state-optimized two-local measurement basis is given as a product of such a generated local basis,

$$\left\{ |\psi_{i_1}^{A_1}\rangle \otimes |\psi_{i_2}^{A_2}\rangle \otimes |\psi_{i_3}^{A_3}\rangle \otimes |\psi_{i_4}^{A_4}\rangle \otimes |\psi_{i_5}^{A_5}\rangle \right\}. \qquad \text{(A2)}$$

This method ensures that for the full system ($k = L$), the ground state is one of the measurement basis states.

b. Observable-optimized method

This method is a *k*-local optimization for a specific observable, in our case the Hamiltonian. The basis is generated by the eigenbasis of the Hamiltonian where interaction terms between the local parts have been taken out. For example, in the Heisenberg model, the two-local measurement basis is given as the eigenbasis of the Hamiltonian

$$\hat{H}_2 = \sum_{i=1,3,5,7,9} \left(\hat{\sigma}_i^x \hat{\sigma}_{i+1}^x + \hat{\sigma}_i^y \hat{\sigma}_{i+1}^y + \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right) + \sum_{i=1}^{10} h_i \hat{\sigma}_i^z. \qquad \text{(A3)}$$

This method ensures that the measurement basis for the full system ($k = L$) is the same as the eigenbasis of the observable we optimize for. (See Appendix E 2 for details.)

2. Upper bound on energy probabilities

First, we prove an upper bound on the energy probability

$$p_E \leqslant \left(\sum_i \sqrt{p_i} \sqrt{\langle E|\hat{\Pi}_i|E\rangle} \right)^2, \qquad \text{(A4)}$$

where the right-hand side defines a_E. This is the second half of Eq. (20).

Proof. We express the spectral decomposition of the density matrix as $\hat{\rho} = \sum_m \lambda_m |\psi_m\rangle\langle\psi_m|$, where λ_m are its eigenvalues corresponding to eigenvectors $|\psi_m\rangle$. We find that $p_i = \text{tr}(\hat{\Pi}_i \hat{\rho})$ translates to $p_i = \sum_m \lambda_m \langle\psi_m|\hat{\Pi}_i|\psi_m\rangle$. A series of inequalities follows:

$$p_E = \langle E|\hat{\rho}|E\rangle = \sum_m \lambda_m |\langle E|\psi_m\rangle|^2 = \sum_m \lambda_m \left| \sum_i \langle E|\hat{\Pi}_i|\psi_m\rangle \right|^2$$

$$\leqslant \sum_m \lambda_m \left(\sum_i |\langle E|\hat{\Pi}_i|\psi_m\rangle| \right)^2$$

$$= \sum_m \lambda_m \left(\sum_i |\langle E|\sqrt{\hat{\Pi}_i^\dagger}\sqrt{\hat{\Pi}_i}|\psi_m\rangle| \right)^2$$

$$\leqslant \sum_m \lambda_m \left(\sum_i \sqrt{\langle E|\hat{\Pi}_i|E\rangle \langle\psi_m|\hat{\Pi}_i|\psi_m\rangle} \right)^2$$

$$= \sum_m \lambda_m \|\boldsymbol{x}^m\|_{1/2}$$

$$\leqslant \left\| \sum_m \lambda_m \boldsymbol{x}^m \right\|_{1/2}$$

$$= \left(\sum_i \sqrt{\sum_m \lambda_m \langle E|\hat{\Pi}_i|E\rangle \langle \psi_m|\hat{\Pi}_i|\psi_m\rangle} \right)^2$$

$$= \left(\sum_i \sqrt{p_i} \sqrt{\langle E|\hat{\Pi}_i|E\rangle} \right)^2. \tag{A5}$$

The first inequality is the triangle inequality, the second the Cauchy-Schwarz inequality, and the third Jensen's theorem applied on the ($p = \frac{1}{2}$)-seminorm

$$\|\mathbf{x}\|_{1/2} = \left(\sum_i \sqrt{x_i} \right)^2. \tag{A6}$$

Here $\mathbf{x}^m = (x_1^m, x_2^m, \dots)$ is a vector of positive entries $x_i^m = \langle E|\hat{\Pi}_i|E\rangle \langle \psi_m|\hat{\Pi}_i|\psi_m\rangle \geqslant 0$. In order to apply Jensen's theorem, we need first to confirm that the ($p = \frac{1}{2}$)-seminorm is a concave function. It is concave when restricted on vectors with positive entries $\| \|_{1/2} : \mathbb{R}_+^N \to \mathbb{R}_+$, which is indeed the case here. This follows from the reverse Minkowski inequality

$$\|\mathbf{x} + \mathbf{y}\|_p \geqslant \|\mathbf{x}\|_p + \|\mathbf{y}\|_p \tag{A7}$$

(where it is assumed that $x_i \geqslant 0$ and $y_i \geqslant 0$), which holds for all ($p < 1$)-seminorms. Taking $0 \leqslant q \leqslant 1$, we have

$$\|q\mathbf{x} + (1-q)\mathbf{y}\|_p \geqslant \|q\mathbf{x}\|_p + \|(1-q)\mathbf{y}\|_p$$
$$= q\|\mathbf{x}\|_p + (1-q)\|\mathbf{y}\|_p, \tag{A8}$$

so $\| \|_{1/2}$ is indeed concave. ∎

3. Lower bound on energy probabilities

Next we prove a lower bound on the energy probabilities

$$p_E \geqslant \max \left\{ \sum_i p_i (x_i^E + \gamma_i y_i^E) - \left(\sum_i \sqrt{p_i y_i^E V_i} \right)^2, 0 \right\}, \tag{A9}$$

where the right-hand side defines b_E. This is the first half of Eq. (20), above which x_i^E, y_i^E, and V_i are defined. We drop superscripts E to keep the notation cleaner, i.e., we write $x_i \equiv x_i^E$ and $y_i \equiv y_i^E$.

Proof. It is clear that $p_E \geqslant 0$. To derive the first inequality, we start by expressing p_E more conveniently as

$$p_E = \langle E|\hat{\rho}|E\rangle = \sum_m \lambda_m \left| \sum_i \langle E|\hat{\Pi}_i|\psi_m\rangle \right|^2$$

$$= \sum_m \lambda_m \left| \sum_{i,k} \gamma_i^k \langle E|i^k\rangle \langle i^k|\psi_m\rangle \right|^2$$

$$= \sum_{m,i,k} \lambda_m |\gamma_i^k \langle E|i^k\rangle \langle i^k|\psi_m\rangle|^2$$

$$+ \sum_{m,i\neq i',k,k'} \lambda_m \gamma_i^k \gamma_{i'}^{k'} \langle i'^{k'}|E\rangle \langle E|i^k\rangle \langle i^k|\psi_m\rangle \langle \psi_m|i'^{k'}\rangle$$

$$+ \sum_{m,i,k\neq k'} \lambda_m \gamma_i^k \gamma_i^{k'} \langle i^{k'}|E\rangle \langle E|i^k\rangle \langle i^k|\psi_m\rangle \langle \psi_m|i^{k'}\rangle$$

$$\geqslant \sum_{m,i,k} \lambda_m |\gamma_i^k \langle E|i^k\rangle \langle i^k|\psi_m\rangle|^2 + \sum_{m,i,k} \lambda_m |\gamma_i^k \langle i^k|\psi_m\rangle|^2 y_i - w$$

$$\geqslant \sum_{m,i,k} \lambda_m \gamma_i^k |\langle i^k|\psi_m\rangle|^2 (x_i + \gamma_i y_i) - w$$

$$= \sum_i p_i (x_i + \gamma_i y_i) - w. \tag{A10}$$

The first inequality follows from the definition of absolute value and the second from definitions of x_i and γ_i. We define

$$w = \sum_{m,i\neq i',k,k'} \lambda_m \gamma_i^k \gamma_{i'}^{k'} |\langle i'^{k'}|E\rangle \langle E|i^k\rangle \langle i^k|\psi_m\rangle \langle \psi_m|i'^{k'}\rangle|$$

$$+ \sum_{m,i,k\neq k'} \lambda_m \gamma_i^k \gamma_i^{k'} |\langle i^{k'}|E\rangle \langle E|i^k\rangle \langle i^k|\psi_m\rangle \langle \psi_m|i^{k'}\rangle|$$

$$+ \sum_{m,i,k} \lambda_m |\gamma_i^k \langle i^k|\psi_m\rangle|^2 y_i$$

$$\leqslant \sum_{m,i,i',k,k'} \lambda_m \gamma_i^k \gamma_{i'}^{k'} |\langle i^k|\psi_m\rangle \langle \psi_m|i'^{k'}\rangle| \sqrt{y_i y_{i'}}$$

$$= \sum_{i,i'} \sqrt{y_i y_{i'}} \sum_m \lambda_m \left(\sum_k \gamma_i^k |\langle i^k|\psi_m\rangle| \right)$$

$$\times \left(\sum_{k'} \gamma_{i'}^{k'} |\langle i'^{k'}|\psi_m\rangle| \right)$$

$$\leqslant \sum_{i,i'} \sqrt{y_i y_{i'}} \sum_m \lambda_m \left(\sqrt{\sum_k \gamma_i^k |\langle i^k|\psi_m\rangle|^2} \sqrt{\sum_k \gamma_i^k} \right)$$

$$\times \left(\sqrt{\sum_{k'} \gamma_{i'}^{k'} |\langle i'^{k'}|\psi_m\rangle|^2} \sqrt{\sum_{k'} \gamma_{i'}^{k'}} \right)$$

$$= \sum_{i,i'} \sqrt{y_i y_{i'} V_i V_{i'}} \sum_m \lambda_m ||\vec{a}_i^m||_2 ||\vec{a}_{i'}^m||_2$$

$$\leqslant \sum_{i,i'} \sqrt{y_i y_{i'} V_i V_{i'}} \sqrt{\sum_m \lambda_m ||\vec{a}_i^m||_2^2} \sqrt{\sum_{m'} \lambda_{m'} ||\vec{a}_{i'}^{m'}||_2^2}$$

$$= \sum_{i,i'} \sqrt{p_i p_{i'} y_i y_{i'} V_i V_{i'}} = \left(\sum_i \sqrt{p_i y_i V_i} \right)^2, \tag{A11}$$

where $V_i = \sum_k \gamma_i^k$, $p_i = \sum_m \lambda_m ||\vec{a}_i^m||_2^2$, $\vec{a}_i^m = (\sqrt{\gamma_i^1} |\langle i^1|\psi_m\rangle|, \sqrt{\gamma_i^2} |\langle i^2|\psi_m\rangle|, \dots)$, and $|| ||_2$ is the two-norm. The first inequality follows from the definition of y_i. We use the Cauchy-Schwarz inequality of type $|\sum_i a_i b_i| \leqslant \sqrt{\sum_i a_i^2} \sqrt{\sum_{i'} b_{i'}^2}$ in the second and the third inequality. Combining the two bounds, we obtain Eq. (A9). ∎

4. Proof of collective bounds

Finally, we prove

$$\sqrt{\mathbf{p}}^T A_i \sqrt{\mathbf{p}} \leqslant p_i, \tag{A12}$$

where $\sqrt{\mathbf{p}}_E = \sqrt{p_E}$ and $(A_i)_{EE'} = (-1)^{1+\delta_{EE'}} |\langle E|\hat{\Pi}_i|E'\rangle|$. Expanding this gives $\sum_{E,E'} \sqrt{p_E} (A_i)_{EE'} \sqrt{p_{E'}} \leqslant p_i$.

Proof. We have

$$
\begin{aligned}
p_i = \mathrm{tr}(\hat{\Pi}_i \hat{\rho}) &= \sum_{E,E'} \mathrm{tr}(|E\rangle\langle E|\hat{\Pi}_i|E'\rangle\langle E'|\hat{\rho}) \\
&= \sum_{E,E'} \langle E|\hat{\Pi}_i|E'\rangle\langle E'|\hat{\rho}|E\rangle \\
&= \sum_E \langle E|\hat{\Pi}_i|E\rangle p_E + \sum_{E\neq E'} \langle E|\hat{\Pi}_i|E'\rangle\langle E'|\hat{\rho}|E\rangle \\
&\geqslant \sum_E \langle E|\hat{\Pi}_i|E\rangle p_E - \sum_{E\neq E'} |\langle E|\hat{\Pi}_i|E'\rangle||\langle E'|\hat{\rho}|E\rangle| \\
&\geqslant \sum_E \langle E|\hat{\Pi}_i|E\rangle p_E - \sum_{E\neq E'} |\langle E|\hat{\Pi}_i|E'\rangle|\sqrt{p_E p_{E'}} \\
&= 2\sum_E \langle E|\hat{\Pi}_i|E\rangle p_E - \sum_{E,E'} |\langle E|\hat{\Pi}_i|E'\rangle|\sqrt{p_E p_{E'}} \\
&= \sum_{E,E'} \sqrt{p_E}(\boldsymbol{A}_i)_{EE'}\sqrt{p_{E'}},
\end{aligned}
\tag{A13}
$$

where $p_E = \langle E|\hat{\rho}|E\rangle$. For the last inequality, we use the fact that $\hat{\rho}$ is positive semidefinite and therefore in accordance with Sylvester's criterion for positive-semidefinite matrices [111], which says that all submatrices must have a non-negative determinant (i.e., all the principal minors are non-negative). For 2×2 submatrices this means

$$
\langle E|\hat{\rho}|E\rangle\langle E'|\hat{\rho}|E'\rangle - \langle E|\hat{\rho}|E'\rangle\langle E'|\hat{\rho}|E\rangle \geqslant 0, \tag{A14}
$$

and thus $|\langle E'|\hat{\rho}|E\rangle|^2 \leqslant p_E p_{E'}$. The second inequality $\sqrt{\boldsymbol{p}}^T \boldsymbol{B}_i \sqrt{\boldsymbol{p}} \geqslant p_i$ is proved analogously. ∎

5. Derivation of the analytic bound on the mean energy

Here we derive the recurrence formula for the computation of the lower bound upper values of the bound

$$
E_{\min}^{\lin} \leqslant \langle E \rangle \leqslant E_{\max}^{\lin}. \tag{A15}
$$

We do this with the lower value $E_{\min}^{\lin}$; the formula for the upper value follows analogously. The lower bound is given by

$$
E_{\min}^{\lin} = \sum_{l=1}^{N} E_l u_l \tag{A16}
$$

[Eq. (11)], where u_l follows a recurrence relation that we derive next. We simplify the lower indices as $l \equiv E_l$.

The idea behind the recurrence relation is described in Fig. 1. Having bounds

$$
a_l^{\max} \equiv a_{E_l}^{\max} \leqslant p_{E_l} \leqslant b_{E_l}^{\min} \equiv b_l^{\min} \tag{A17}
$$

[Eq. (3)], we find the lower bound on the mean energy by filling the minimal probability of each energy given by $a_l^{\max}$ and topping it up to the maximum $b_l^{\min}$, from the lowest to the highest energy, until the probabilities sum to one.

What does this mean mathematically? Let us think of a "bottle" of probabilities with the total volume equal to one, $V = 1$. We start with all u_l initialized at zero. We pour the minimal required amount given by $a_E^{\max} \leqslant p_E$ into each

probability u_l. After this, we have

$$
\begin{aligned}
u_1 &= a_1^{\max}, \\
u_2 &= a_2^{\max}, \\
&\vdots \\
u_N &= a_N^{\max}
\end{aligned}
\tag{A18}
$$

and the remaining volume in the bottle is $V = 1 - \sum_{l=1}^{N} a_l^{\max}$. We start topping each u_l up to its maximum value, from the lowest- to the highest-energy eigenvalue, until we run out of the probability in the bottle. For $l = 1$, the two cases can occur: Either there is enough probability in the bottle to fill u_l up to its maximum allowed value $b_1^{\min}$ or there is not. Mathematically, this topping up is expressed as

$$
u_1 = a_1^{\max} + \min\left\{ b_1^{\min} - a_1^{\max}, 1 - \sum_{l=1}^{N} a_l^{\max} \right\}. \tag{A19}
$$

The $a_1^{\max}$ can be subtracted, which gives

$$
u_1 = \min\left\{ b_1^{\min}, 1 - \sum_{l=2}^{N} a_l^{\max} \right\}. \tag{A20}
$$

The remaining volume in the bottle is $V = 1 - u_1 - \sum_{l=2}^{N} a_l^{\max}$.

Next we top up the second probability, which gives

$$
\begin{aligned}
u_2 &= a_2^{\max} + \min\left\{ b_2^{\min} - a_2^{\max}, 1 - u_1 - \sum_{l=2}^{N} a_l^{\max} \right\} \\
&= \min\left\{ b_2^{\min}, 1 - u_1 - \sum_{l=3}^{N} a_l^{\max} \right\}.
\end{aligned}
\tag{A21}
$$

The remaining volume is $V = 1 - \sum_{l=1}^{2} u_l - \sum_{l=3}^{N} a_l^{\max}$. We continue up to the maximal index N, deriving the full recursive relation [Eq. (12)]. The recursive relation for $E_{\max}^{\lin}$ is derived analogously.

APPENDIX B: POWERFUL IMPROVEMENT FROM COLLECTIVE BOUNDS

Here we show an example in which the collective bounds. (7), where

$$
p_i^{\min} = \min_{t\in[0,T]} p_i(t), \quad p_i^{\max} = \max_{t\in[0,T]} p_i(t), \tag{B1}
$$

provide a considerable improvement in the estimation of energy probabilities in comparison with using just the linear bound (3). Consider a coarse-grained energy measurement $\mathcal{C} = \{\hat{P}_{\tilde{E}}\}$ given by the coarse-grained energy projectors

$$
\hat{P}_{\tilde{E}} = \sum_{E\in[\tilde{E}, \tilde{E}+\Delta E)} |E\rangle\langle E|, \tag{B2}
$$

where ΔE denotes the resolution in measuring energy. Then $\boldsymbol{A}_{\tilde{E}} = \boldsymbol{B}_{\tilde{E}}$ are diagonal and the inequalities (7) yield

$$
p_{\tilde{E}} = p_{\tilde{E}}^{\max} = p_{\tilde{E}}^{\min} = \sum_{E\in[\tilde{E}, \tilde{E}+\Delta E)} p_E. \tag{B3}
$$

This upper bounds the sum of energy probabilities, making the determination of the mean energy much more precise. This is a stark difference from Eq. (3), which yields much less restrictive $p_E \leqslant p_{\tilde{E}}$ for each $E \in [\tilde{E}, \tilde{E} + \Delta E)$.

To give an example, consider a Hamiltonian $\hat{H} = \sum_{i=1}^{8} E_i |E_i\rangle\langle E_i|$ with the spectrum

$$\{E_1, \dots, E_8\} = \{0, 1, 2, 2.5, 3, 3.3, 3.7, 4\}. \quad \text{(B4)}$$

Consider a fixed resolution in measuring energy to be $\Delta E = 1$. This results in coarse-grained energy projectors (B2), representing the coarse energy measurement $\tilde{\mathcal{C}} = \{\hat{P}_{\tilde{E}_j}\}_{j=1}^5$ as

$$\begin{aligned}
\hat{P}_{\tilde{E}_1} &= |0\rangle\langle 0|, \\
\hat{P}_{\tilde{E}_2} &= |1\rangle\langle 1|, \\
\hat{P}_{\tilde{E}_3} &= |2\rangle\langle 2| + |2.5\rangle\langle 2.5|, \\
\hat{P}_{\tilde{E}_4} &= |3\rangle\langle 3| + |3.3\rangle\langle 3.3| + |3.7\rangle\langle 3.7|, \\
\hat{P}_{\tilde{E}_5} &= |4\rangle\langle 4|.
\end{aligned} \quad \text{(B5)}$$

This indicates that the measurement device cannot distinguish between energy states $|2\rangle$ and $|2.5\rangle$, for example, because they are too close in energy.

Consider an initial state

$$|\psi\rangle = (|2\rangle + |3\rangle)/\sqrt{2}. \quad \text{(B6)}$$

Knowing this state, we can compute

$$\begin{aligned}
p_0 = 0, \quad p_1 = 0, \quad p_2 = \tfrac{1}{2}, \quad p_{2.5} = 0, \\
p_3 = \tfrac{1}{2}, \quad p_{3.3} = 0, \quad p_{3.7} = 0, \quad p_4 = 0.
\end{aligned} \quad \text{(B7)}$$

However, the experimenter performing a coarse-grained measurement $\tilde{\mathcal{C}}$ on many copies of this initial state does not know that. Instead, using Eq. (3), they derive

$$\begin{aligned}
p_0 = 0, \quad p_1 = 0, \quad p_2 \leqslant \tfrac{1}{2}, \quad p_{2.5} \leqslant \tfrac{1}{2}, \\
p_3 \leqslant \tfrac{1}{2}, \quad p_{3.3} \leqslant \tfrac{1}{2}, \quad p_{3.7} \leqslant \tfrac{1}{2}, \quad p_4 = 0.
\end{aligned} \quad \text{(B8)}$$

The right-hand sides are obtained from the outcomes of the coarse-grained measurement. This is because half of the time, they obtain measurement outcome $\tilde{E}_3$ and the other half they get $\tilde{E}_4$. If they were to estimate the mean energy of the state only from these equations, they would obtain

$$2 \times \tfrac{1}{2} + 2.5 \times \tfrac{1}{2} = 2.25 \leqslant \langle E \rangle \leqslant 3.5 = 3.3 \times \tfrac{1}{2} + 3.7 \times \tfrac{1}{2}. \quad \text{(B9)}$$

However, from Eq. (B3), which follows from the collective bounds, they obtain an additional set of equations

$$\tfrac{1}{2} = p_2 + p_{2.5}, \quad \tfrac{1}{2} = p_3 + p_{3.3} + p_{3.7}. \quad \text{(B10)}$$

The left-hand sides are obtained from the experimental outcomes. Using this additional set of equations, the experimenter is able to derive a noticeably tighter bound on the mean energy,

$$2 \times \tfrac{1}{2} + 3 \times \tfrac{1}{2} = 2.5 \leqslant \langle E \rangle \leqslant 3.1 = 2.5 \times \tfrac{1}{2} + 3.7 \times \tfrac{1}{2}. \quad \text{(B11)}$$

This improvement will be dramatic in systems with many energy eigenstates, leading to much coarser projectors.

APPENDIX C: QUALITY FACTORS

We can employ two quality factors to assess the performance of the method for bounding the mean energy: the first,

$$Q_1 = \left(1 - \frac{E_{\max} - E_{\min}}{E_N - E_1}\right) \times 100\%, \quad \text{(C1)}$$

which measures the range of excluded energy, and the second,

$$Q_2 = \left(1 - \frac{N_{[E_{\min}, E_{\max}]}}{N}\right) \times 100\%, \quad \text{(C2)}$$

which measures the percentage of excluded energy eigenstates. Here $N_{[E_{\min}, E_{\max}]}$ denotes the number of energy eigenstates with energy between $E_{\min}$ and $E_{\max}$ and N is the dimension of the Hilbert space.

APPENDIX D: SIMPLE EXAMPLE

Consider a Hamiltonian given by the Pauli z matrix

$$\hat{H} = \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{(D1)}$$

which has energy eigenvalues $E_0 = -1$ and $E_1 = 1$, corresponding to eigenstates $|0\rangle$ and $|1\rangle$, respectively. The task is to estimate the energy of a general pure qubit state,

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle, \quad \text{(D2)}$$

where $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \phi \leqslant 2\pi$.

We consider two different two-outcome measurements to estimate energy. We will use the combined bound for the energy probabilities

$$\left| p_E - \sum_i p_i |\langle E|i\rangle|^2 \right| \leqslant b_E - \sum_i p_i |\langle E|i\rangle|^2, \quad \text{(D3)}$$

where $b_E = (\sum_i \sqrt{p_i} |\langle E|i\rangle|)^2$, which is easily derived from Eq. (1).

First, consider measuring in the z basis, i.e., measuring in the eigenbasis of the operator $\hat{M} = \hat{\sigma}_z$. This defines the measurement $\mathcal{C} = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$. Because the measurement basis is the same as the eigenbasis of the Hamiltonian, we are measuring the energy directly, so we expect the exact result. From the bound above, we have

$$\begin{aligned}
|p_{E_0} - p_0| \leqslant 0, \\
|p_{E_1} - p_1| \leqslant 0,
\end{aligned} \quad \text{(D4)}$$

independent of the initial state $|\psi\rangle$. Clearly, $p_{E_0} = p_0$, $p_{E_1} = p_1$, and $E_{\min}^{\lin} = \langle E \rangle = E_{\max}^{\lin}$, as expected.

Second, consider measuring in the x basis, i.e., measuring in the eigenbasis of the operator $\hat{M} = \hat{\sigma}_x$. This defines $\mathcal{C} = \{|+\rangle\langle +|, |-\rangle\langle -|\}$. We have

$$\begin{aligned}
\left|p_{E_0} - \tfrac{1}{2}\right| \leqslant \tfrac{1}{2}[(\sqrt{p_+} + \sqrt{p_-})^2 - 1], \\
\left|p_{E_1} - \tfrac{1}{2}\right| \leqslant \tfrac{1}{2}[(\sqrt{p_+} + \sqrt{p_-})^2 - 1].
\end{aligned} \quad \text{(D5)}$$

This means that if the state is aligned with the x axis, for example, $p_+ = 1$, then $p_{E_0} = p_{E_1} = \tfrac{1}{2}$ and we can determine the energy exactly as $\langle E \rangle = 0$. On the contrary, if the state is aligned with the z axis, implying $p_+ = p_- = \tfrac{1}{2}$, then
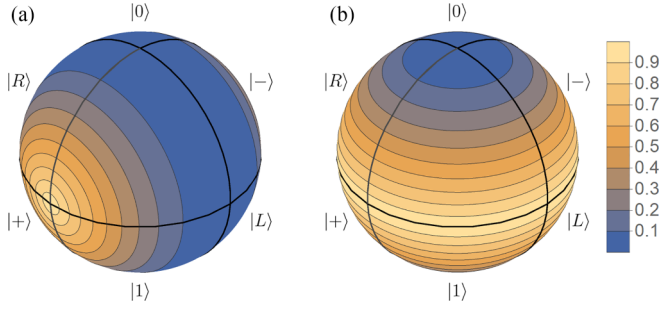
FIG. 4. Illustration of the performance when estimating energy by local Pauli x measurements in a simple qubit example. We show the quality Q_1 factor (percentage of excluded energies) when estimating energy given by the Hamiltonian $\hat{H} = \hat{\sigma}_z$ by measuring in the $\hat{\sigma}_x$ basis, for different initial states on the Bloch sphere. (a) Quality factor of the initial state. The value $Q_1 = 1$ (light orange, which happens for state $|\psi\rangle = |+\rangle$) corresponds to perfect identification of the energy, while the value $Q_1 = 0$ (dark blue, which happens for state $|\psi\rangle = |0\rangle$) corresponds to failure in identifying the energy. (b) Quality factor over initial states if we allow time evolution of the state with the Hamiltonian and measure at different times $t \in [0, \pi/2]$ to build up statistics of outcomes at each time of this interval. The Hamiltonian revolves the state around the z axis, so during the time evolution, the state picks up the best quality factor at that latitude from (a).

$|p_{E_0} - \frac{1}{2}| \leqslant \frac{1}{2}$ and $|p_{E_1} - \frac{1}{2}| \leqslant \frac{1}{2}$, which we can rewrite as $0 \leqslant p_{E_0} \leqslant 1$ and $0 \leqslant p_{E_1} \leqslant 1$. Thus we obtain a trivial bound

$$-1 = E_{\max}^{\lin} \leqslant \langle E \rangle \leqslant E_{\max}^{\lin} = 1. \tag{D6}$$

Generally, we have

$$p_+ = \tfrac{1}{2}(1 + \cos\phi \sin\theta), \quad p_- = \tfrac{1}{2}(1 + \cos\phi \sin\theta), \tag{D7}$$

which gives

$$|p_{E_0} - \tfrac{1}{2}| \leqslant \tfrac{1}{2}\sqrt{1 - \cos^2\phi \sin^2\theta},$$
$$|p_{E_1} - \tfrac{1}{2}| \leqslant \tfrac{1}{2}\sqrt{1 - \cos^2\phi \sin^2\theta}. \tag{D8}$$

This yields

$$-\sqrt{1 - \cos^2\phi \sin^2\theta} = E_{\max}^{\lin} \leqslant \langle E \rangle \leqslant E_{\max}^{\lin}$$
$$= \sqrt{1 - \cos^2\phi \sin^2\theta}. \tag{D9}$$

We visualize the corresponding quality factor [the percentage of excluded energies; see Eq. (C1)]

$$Q_1 = 1 - \sqrt{1 - \cos^2\phi \sin^2\theta} \tag{D10}$$

for this general case on the Bloch sphere in Fig. 4.

Next we consider time evolution. We have

$$|\psi_t\rangle = e^{-i\hat{H}t}|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i(\phi - 2t)}\sin\frac{\theta}{2}|1\rangle. \tag{D11}$$

Measuring $\hat{\sigma}_x$ at time t bounds the energy probabilities as

$$|p_{E_0} - \tfrac{1}{2}| \leqslant \tfrac{1}{2}\sqrt{1 - \cos^2(\phi - 2t)\sin^2\theta},$$
$$|p_{E_1} - \tfrac{1}{2}| \leqslant \tfrac{1}{2}\sqrt{1 - \cos^2(\phi - 2t)\sin^2\theta}. \tag{D12}$$

If we measure at different times during $t \in [0, \pi/2)$, we manage to tighten these bounds and obtain

$$|p_{E_0} - \tfrac{1}{2}| \leqslant \tfrac{1}{2}|\cos\theta|,$$
$$|p_{E_1} - \tfrac{1}{2}| \leqslant \tfrac{1}{2}|\cos\theta|, \tag{D13}$$

which leads to bounds on energy

$$-|\cos\theta| = E_{\max}^{\lin} \leqslant \langle E \rangle \leqslant E_{\max}^{\lin} = |\cos\theta|. \tag{D14}$$

The corresponding quality factor $Q_1 = 1 - |\cos\theta|$ is again plotted in Fig. 4.

APPENDIX E: ESTIMATION OF THE MEAN ENERGY IN EXPERIMENTALLY RELEVANT MODELS

Here we show the simulations for energy estimation using measurements in a local number basis and then using optimized k-local measurements. This is a continuation of numerical simulations shown in the main text, which contained a part of the results obtained for the Heisenberg model.

We simulate a number of models, including the Heisenberg model [70–72], which is a paradigmatic model to study for many-body localization, and then several other experimentally relevant models. These are the Ising model [73–76], known for its frequent use in quantum simulators and quantum annealers; the XY model [77–80], which is a type of nonintegrable long-range model; and the PXP model [81–83], an archetypal model for many-body quantum scars.

The Hamiltonians and the corresponding parameters are given in Table I. The simulations of energy estimation using local particle number measurements and optimized k-local measurements are shown in Tables II and III.

The bulk of the explanation necessary to understand these numerical experiments has been given in the main text. Below we give details on the types of Hilbert space considered in our simulations and we discuss methods that we designed for the optimization over the k-local measurements to estimate energy.

1. Hilbert space

In our simulations, we choose to work in a different type of Hilbert space for each model, depending on the conservation laws, and to match the experimental setups (see Table I): In the Heisenberg and the XY models, the full Hilbert space splits into subspaces, each of them characterized by a definite value of the total spin along the z axes, i.e., $\hat{S}^z = \sum_i \hat{\sigma}_i^z$. We work in the largest subspace, characterized by the value $\hat{S}^z = 0$. This conservation is also why the actual value of B in the XY model is irrelevant. For the Ising model, the total spin along the z axes is not conserved; only the parity of the total spin is conserved. In this case, we work on the parity-even subspace, which contains the Néel state. For the PXP model, the situation is more intricate: The presence of the projector operators $\hat{I} - \hat{\sigma}_i^z$ in the Hamiltonian introduces a nontrivial local constraint [82]. Consequently, the full Hilbert space shatters in many different subspaces, dynamically disconnected and having various dimensions [112]. Inside each subspace, the dynamics is generically chaotic with the presence of many-body scars [82,113,114]. However, our goal in using this model is to study the effect of the Hilbert space

TABLE I. Models used in our simulations and the Hilbert spaces considered with dimension D. Here $\hat{\sigma}_i^x$ denotes the Pauli x matrix at site i, and similarly for Pauli y and Pauli z matrices; $\hat{\sigma}_i^+$ and $\hat{\sigma}_i^-$ denote the spin creation and annihilation operators, respectively. The Ising and XY models have a nonlocal interaction of the form $J_{ij} = J_0/|i-j|^\alpha$. The parameters were taken to match those employed in experiments.

Model	Characteristics	Conservation	Hamiltonian	Parameters
Heisenberg (integrable, delocalized, (localized))	many-body localizing	total spin conserving	$\hat{H} = \sum_i \left(\hat{\sigma}_i^x \hat{\sigma}_{i+1}^x + \hat{\sigma}_i^y \hat{\sigma}_{i+1}^y + \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right) + \sum_i h_i \hat{\sigma}_i^z$	$h_i \in [-W, W]$ drawn randomly $(W=0, W=0.5, W=10)$
Ising (delocalized, localized)	gapped and many-body localizing	parity conserving	$\hat{H} = \sum_{i<j} J_{ij} \hat{\sigma}_i^x \hat{\sigma}_j^x + \frac{1}{2} \sum_i (B + h_i) \hat{\sigma}_j^z$	$J_0 = 1, \alpha = 1.13, B = 4,$ $h_i \in [-W, W]$ drawn randomly $(W=0, W=8)$
XY	gapped and long range	total spin conserving	$\hat{H} = \sum_{i<j} J_{ij}(\hat{\sigma}_i^+ \hat{\sigma}_j^- + \hat{\sigma}_i^- \hat{\sigma}_j^+) + B \sum_i \hat{\sigma}_i^z$	$J_0 = 1, \alpha = 1.24, B = 0$
PXP	quantum scars	Hilbert space shattering	$\hat{H} = \frac{\Omega}{4} \sum_i (\hat{I} - \hat{\sigma}_i^z) \hat{\sigma}_{i+1}^x (\hat{I} - \hat{\sigma}_{i+2}^z)$	$\Omega = 1$

Sizes and Hilbert spaces considered in the numerical experiments				
Small systems			Large systems	
Heisenberg	six sites, three particles	$D = 20$	ten sites, five particles	$D = 252$
Ising	six sites, even-parity subspace	$D = 32$	ten sites, even-parity subspace	$D = 512$
XY	six sites, three particles	$D = 20$	ten sites, five particles	$D = 252$
PXP	five sites, full Hilbert space	$D = 32$	ten sites, full Hilbert space	$D = 1024$

shattering on the quality of the energy estimation. Therefore, we work in the full Hilbert space.

2. Methods for optimizing local measurements

We introduce three methods of analytical optimization for the k-local measurements. The sketch of k-local measurements is shown in Fig. 5, which corresponds to Fig. 3(a). The k-local measurement consists of applying a k-local unitary operator and then measuring in the computational basis.

a. Ground-state-optimized measurements

First, we introduce a method that is k-local optimization for a specific state, in our case, the ground state. This method
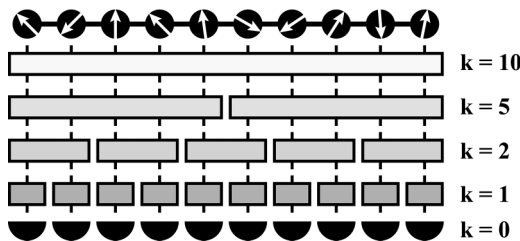

FIG. 5. Sketch of allowed operations, i.e., k-local measurements. Here $k = 0$ corresponds to measuring in the local number basis, $k = 1$ allows for using single-site unitary operators leading to a general single-site measurement, $k = 2$ for two-site local operators, etc.

is inspired by the matrix product state ansatz [109] and by the correspondence between observational and entanglement entropy [110].

The logic of the motivation is as follows: Low observational entropy means that the system state wandered into one of the small subspaces macrostates given by the measurement [60,63]. This means that we can estimate the maximal and the minimal value of the estimated observable in that subspace, which in turn translates into estimating these bounds for the system state itself. In other words, lower observational entropy means better estimates. According to Ref. [110], observational entropy minimized over local coarse grainings leads to entanglement entropy, and the minimum is achieved when the local coarse grainings are given by the Schmidt basis. This means that the measuring in the Schmidt basis will yield small observational entropy and in turn a better estimate of the mean value of observable. Thus, we need to find k-local measurements that reflect the Schmidt basis, which are expected to perform well in the estimation.

We do this as follows. Assuming we have a chain of length L (for example, $L = 10$) and some divisor $k \leqslant 10$ (for example, $k = 2$), we first divide the chain into two parts: one, system A_1, of length k [i.e., sites (1,2)] and the other, system B_1, of length $L - k$ [i.e., sites (3,4,5,6,7,8,9,10)]. We define $\hat{\rho}_0$ as the ground state (or any other state we want to optimize for). We compute the reduced density matrix

$$\hat{\rho}_{A_1} = \text{tr}_{B_1} \hat{\rho}_0 \qquad (E1)$$

TABLE II. Estimating energy with the local number and k-local optimized measurements (see Fig. 5) for various Hamiltonians and sizes of Hilbert space delineated in Table I. The initial state is either a ground state (G), a pure thermal state (C for cold), Eq. (19), or a state drawn randomly from the Hilbert space with the Haar measure (H for hot). For small systems (left panel) the graphs show the true mean energy $\langle E \rangle$ (single symbol) and intervals $[E_{\min}^{\text{lin}}, E_{\max}^{\text{lin}}]$ (solid line) and $[E_{\min}, E_{\max}]$ (dashed line) for each state ordered from top to bottom and the list of energy eigenvalues at the very bottom. For large systems the graph shows the true mean energy $\langle E \rangle$ (single symbol) and intervals $[E_{\min}^{\text{lin}(k)}, E_{\max}^{\text{lin}(k)}]$ (solid lines), denoting analytic bounds computed for k-local measurements $k = 0, 1, 2, 5, 10$ (see Fig. 4), using the ground-state-optimized method (middle panel) and observable-optimized type-1 method (right panel). (See the observable-optimized type-2 method for the PXP model in Fig. 6.) We make the following observations. (i) There is little difference between the integrable and delocalized phases of the Heisenberg model in estimating energy. Bethe integrability does not seem to play a role. (ii) Estimation of energy in the localized phase of the Heisenberg model works well for both small and large systems. It manages to exclude $Q_1 = 97.5\%$ of the range of energies when estimating the ground-state energy using two-qubit ($k = 2$) measurements. This is due to a large overlap between energy eigenstates and the local number basis. (iii) Estimation of the ground-state energy in the Ising model works well in both the localized ($Q_1 = 96.7\%$ for $k = 2$) and delocalized ($Q_1 = 92.9\%$ for $k = 2$) phases. This is due to the low entanglement in the ground state. (See Table III for a continuation of this table.)
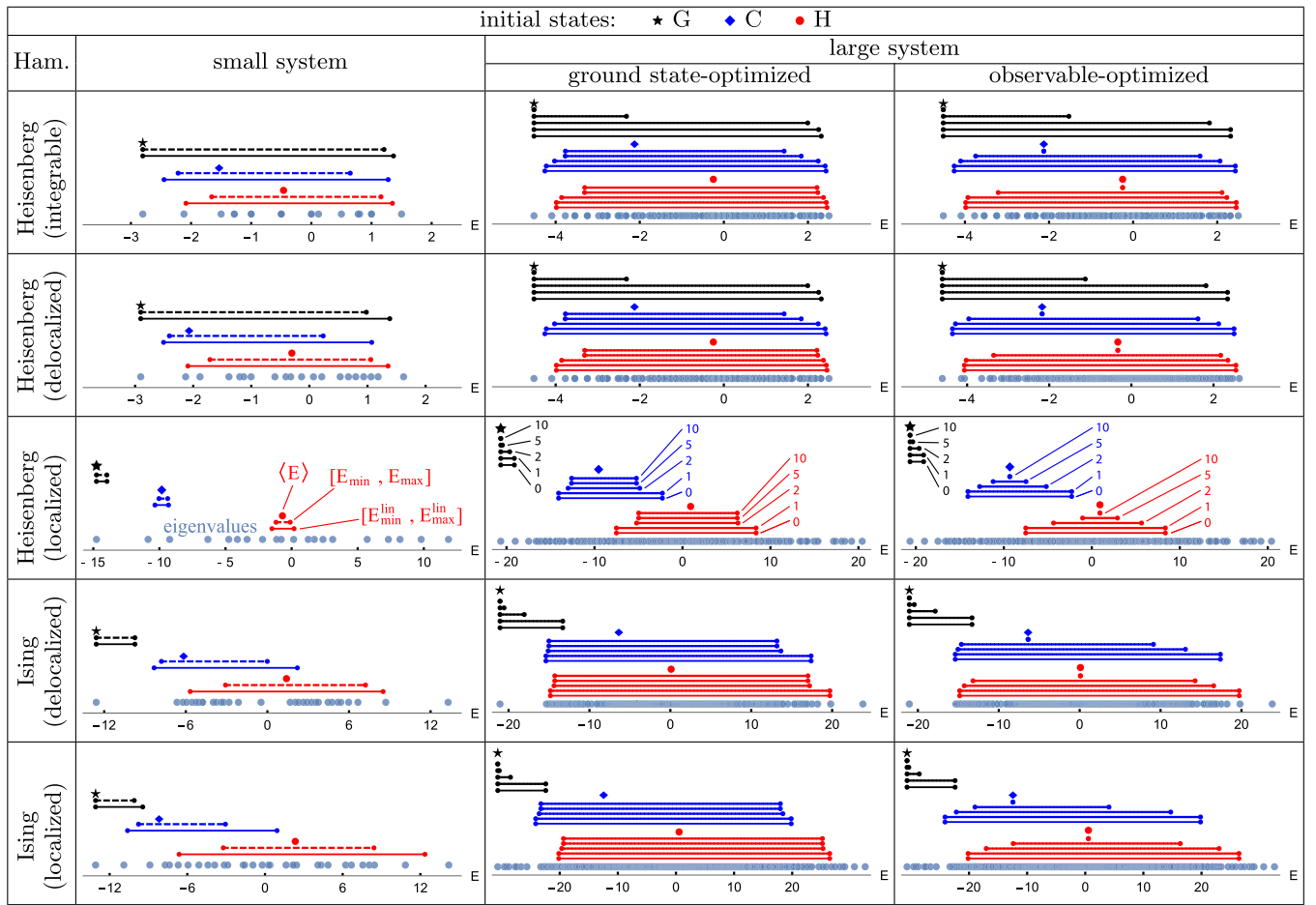
and diagonalize it. The eigenbasis of the reduced density matrix is, by definition, the system A_1-local part of the Schmidt basis. We denote this basis by $\{|\psi_{i_1}^{A_1}\rangle\}$.

Then we move on to the next k sites. We divide the system into A_2 [sites (3,4)] and B_2 [sites (1,2,5,6,7,8,9,10)]. Again, we compute the reduced density matrix

$$\hat{\rho}_{A_2} = \text{tr}_{B_2}\hat{\rho}_0 \qquad (E2)$$

and find its eigenvectors, which we denote by $\{|\psi_{i_2}^{A_2}\rangle\}$.

We continue dividing the system until the end (in our example, we go up to A_5). The final, ground-state-optimized

k-local measurement basis is then given by

$$\left\{ \left|\psi_{i_1}^{A_1}\right\rangle \otimes \left|\psi_{i_2}^{A_2}\right\rangle \otimes \left|\psi_{i_3}^{A_3}\right\rangle \otimes \left|\psi_{i_4}^{A_4}\right\rangle \otimes \left|\psi_{i_5}^{A_5}\right\rangle \right\}. \qquad (E3)$$

b. Observable-optimized measurements (type 1)

Second, we introduce a method of k-local optimization for a specific observable, in our case, the Hamiltonian. The motivation behind this optimization is that the estimation of the mean value of the observable works better the more the measurement resembles the eigenbasis of the observable. Thus, we create a procedure that generates a measurement that somewhat resembles the estimated observable.

TABLE III. Continuation of Table II. (iv) In the XY and PXP models, the local particle number measurement ($k = 0$) does not determine the ground-state energy well. This is because low- and high-energy eigenstates produce similar, or the same in the case of PXP, distribution of outcomes. (v) Estimating the hot state energy in the PXP model works significantly better than estimating the ground-state energy. This is because of the high degeneracy in the middle of the spectrum. (vi) The ground-state-optimized method performs better for ground states than the observable-optimized method but performs worse on average.
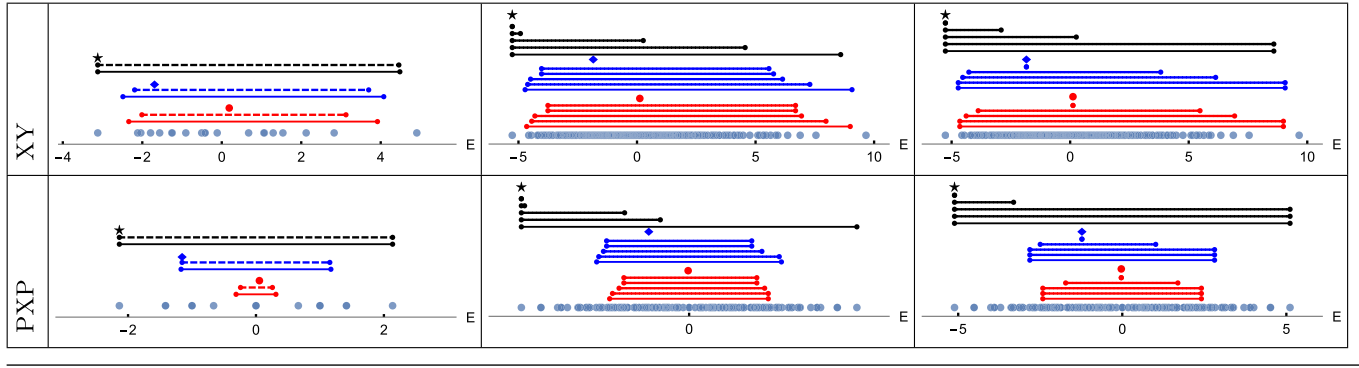
We illustrate this method on the Heisenberg model, assuming $L = 10$, which is given by the Hamiltonian (assuming hard-wall boundary conditions)

$$\hat{H} = \sum_{i=1}^{9} \left(\hat{\sigma}_i^x \hat{\sigma}_{i+1}^x + \hat{\sigma}_i^y \hat{\sigma}_{i+1}^y + \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right) + \sum_{i=1}^{10} h_i \hat{\sigma}_i^z. \quad \text{(E4)}$$

For ($k = 1$)-local measurement, we remove all the terms spanning more than a single site. This leads to a modified Hamiltonian

$$\hat{H}_1 = \sum_{i=1}^{10} h_i \hat{\sigma}_i^z. \quad \text{(E5)}$$

We call the eigenbasis of this Hamiltonian the ($k = 1$)-local observable optimized measurement for the Hamiltonian. Incidentally, in this case, this basis is precisely the same as the computational basis.

For ($k = 2$)-local measurement, we divide the lattice into blocks of two sites and remove all the terms that cross those blocks. This leads to

$$\hat{H}_2 = \sum_{i=1,3,5,7,9} \left(\hat{\sigma}_i^x \hat{\sigma}_{i+1}^x + \hat{\sigma}_i^y \hat{\sigma}_{i+1}^y + \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right) + \sum_{i=1}^{10} h_i \hat{\sigma}_i^z. \quad \text{(E6)}$$

The eigenbasis of this Hamiltonian is the ($k = 2$)-local observable optimized measurement for the Hamiltonian.

For ($k = 5$)-local measurement, we divide the lattice into two blocks of five sites and remove all the terms that cross those blocks. This leads to

$$\hat{H}_5 = \sum_{i=1,\dots,4,6,\dots,9} \left(\hat{\sigma}_i^x \hat{\sigma}_{i+1}^x + \hat{\sigma}_i^y \hat{\sigma}_{i+1}^y + \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right) + \sum_{i=1}^{10} h_i \hat{\sigma}_i^z. \quad \text{(E7)}$$

The eigenbasis of this Hamiltonian is the ($k = 5$)-local observable optimized measurement for the Hamiltonian.

In the case of ($k = 10$)-local measurement, the corresponding Hamiltonian is the original Hamiltonian itself,

$$\hat{H}_{10} = \hat{H}. \quad \text{(E8)}$$

Measuring in the basis of this Hamiltonian is the same as measuring the Hamiltonian itself, which yields perfect precision in estimating its mean value.

c. Observable-optimized measurements (type 2)

Alternatively, one can consider a different way of finding observable-optimized measurements, somewhat similar to the ground-state optimization method. For $k = 2$ and $L = 10$, divide the system into system A_1 [the first two sites (1,2)] and system B_1 [the last $L - k$ sites (3,4,5,6,7,8,9,10)]. Then compute the reduced Hamiltonian

$$\hat{H}^{A_1} = \text{tr}_{B_1} \hat{H} \quad \text{(E9)}$$

and diagonalize it, obtaining its eigenbasis $\{|\psi_{i_1}^{A_1}\rangle\}$. Continue analogously as in the ground-state optimization method to generate the global observable-optimized basis for $k = 2$,

$$\left\{ \left| \psi_{i_1}^{A_1} \right\rangle \otimes \left| \psi_{i_2}^{A_2} \right\rangle \otimes \left| \psi_{i_3}^{A_3} \right\rangle \otimes \left| \psi_{i_4}^{A_4} \right\rangle \otimes \left| \psi_{i_5}^{A_5} \right\rangle \right\}. \quad \text{(E10)}$$

In our numerical experiments, due to the specific forms of the Hamiltonian, type-1 and type-2 observable-optimized measurements differ only in the PXP model (see Fig. 6).
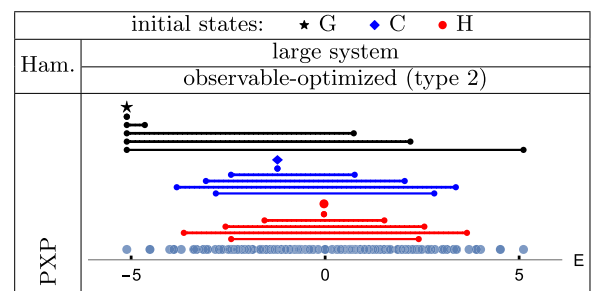

FIG. 6. Observable-optimized type-2 method for the large system of the PXP model. The PXP model is the only one of our considered models in which type 1 and type 2 differ. Note the worse performance for $k = 1$ compared to $k = 0$ for the cold and hot states and better performance for $k = 1, 2$ compared to the type-1 method for the ground state.

d. Generating the unitary

Finally, we want to transform the optimized measurement and express it as a combination of a unitary operator applied to the system's state and after that measuring in the computational basis. This is illustrated in Fig. 5. Assuming our example of $k = 2$ and $L = 10$ again, we start deriving the formula for U^{A_1} by requiring that the probability of an outcome is the same in both situations

$$\langle j_1, j_2 | U^{A_1} \hat{\rho} U^{A_1 \dagger} | j_1, j_2 \rangle \overset{!}{=} \langle \psi_{i_1}^{A_1} | \hat{\rho} | \psi_{i_1}^{A_1} \rangle, \tag{E11}$$

where $|j_1, j_2\rangle$ is the computational basis vector, $|\psi_{i_1}^{A_1}\rangle$ is an optimized basis vector on the first two sites, and we want to match each couple j_1, j_2 to one i_1. A sufficient condition is

$$|\psi_{i_1}^{A_1}\rangle = U^{A_1 \dagger} |j_1, j_2\rangle. \tag{E12}$$

Assuming that $|\psi_{i_1}^{A_1}\rangle$ is a column vector written in the computational basis and that each site is a qubit (which leads to $i_1 = 1, 2, 3, 4$), we have

$$U^{A_1 \dagger} = \begin{pmatrix} |\psi_1^{A_1}\rangle & |\psi_2^{A_1}\rangle & |\psi_3^{A_1}\rangle & |\psi_4^{A_1}\rangle \end{pmatrix}. \tag{E13}$$

This means that

$$U^{A_1} = \begin{pmatrix} |\psi_1^{A_1}\rangle^{\dagger} \\ |\psi_2^{A_1}\rangle^{\dagger} \\ |\psi_3^{A_1}\rangle^{\dagger} \\ |\psi_4^{A_1}\rangle^{\dagger} \end{pmatrix} = \begin{pmatrix} \langle\psi_1^{A_1}| \\ \langle\psi_2^{A_1}| \\ \langle\psi_3^{A_1}| \\ \langle\psi_4^{A_1}| \end{pmatrix}. \tag{E14}$$

The global optimized measurement then consists of applying the unitary operation

$$U = U^{A_1} \otimes \cdots \otimes U^{A_5} = \begin{pmatrix} \langle\psi_1^{A_1}| \\ \langle\psi_2^{A_1}| \\ \langle\psi_3^{A_1}| \\ \langle\psi_4^{A_1}| \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} \langle\psi_1^{A_5}| \\ \langle\psi_2^{A_5}| \\ \langle\psi_3^{A_5}| \\ \langle\psi_4^{A_5}| \end{pmatrix} \tag{E15}$$

on the system and then measuring in the computational basis.

[1] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, Adv. Phys. **65**, 239 (2016).

[2] J. M. Deutsch, Rep. Prog. Phys. **81**, 082001 (2018).

[3] M. Vojta, Rep. Prog. Phys. **66**, 2069 (2003).

[4] Z.-Y. Sun, Y.-Y. Wu, J. Xu, H.-L. Huang, B.-F. Zhan, B. Wang, and C.-B. Duan, Phys. Rev. A **89**, 022101 (2014).

[5] T. Tian, H.-X. Yang, L.-Y. Qiu, H.-Y. Liang, Y.-B. Yang, Y. Xu, and L.-M. Duan, Phys. Rev. Lett. **124**, 043001 (2020).

[6] H. P. Yuen and V. W. S. Chan, Opt. Lett. **8**, 177 (1983).

[7] T. Tyc and B. C. Sanders, J. Phys. A: Math. Gen. **37**, 7341 (2004).

[8] Y. Shaked, Y. Michael, R. Z. Vered, L. Bello, M. Rosenbluh, and A. Pe'er, Nat. Commun. **9**, 609 (2018).

[9] F. Raffaelli, G. Ferranti, D. H. Mahler, P. Sibson, J. E. Kennard, A. Santamato, G. Sinclair, D. Bonneau, M. G. Thompson, and J. C. F. Matthews, Quantum Sci. Technol. **3**, 025003 (2018).

[10] D. Noll, D. Nishimura, and A. Macovski, IEEE Trans. Med. Imag. **10**, 154 (1991).

[11] P. Voss, Optical homodyne detection and applications in quantum cryptography, Ph.D. thesis, TELECOM ParisTech, 2009.

[12] L. Davidovich, Rev. Mod. Phys. **68**, 127 (1996).

[13] Y. Takeno, M. Yukawa, H. Yonezawa, and A. Furusawa, Opt. Express **15**, 4321 (2007).

[14] B. J. Lawrie, P. D. Lett, A. M. Marino, and R. C. Pooser, ACS Photon. **6**, 1307 (2019).

[15] W. Heisenberg, in *Original Scientific Papers Wissenschaftliche Originalarbeiten*, edited by W. Blum, H. Rechenberg, and H. P. Dürr (Springer, Berlin, 1985), pp. 478–504.

[16] H. P. Robertson, Phys. Rev. **34**, 163 (1929).

[17] P. Busch, T. Heinonen, and P. Lahti, Phys. Rep. **452**, 155 (2007).

[18] A. S. Wightman, Phys. Rev. **101**, 860 (1956).

[19] M. Srednicki, *Quantum Field Theory* (Cambridge University Press, Cambridge, 2007).

[20] C. Bunge, J. Barrientos, and A. Bunge, At. Data Nucl. Data Tables **53**, 113 (1993).

[21] A. N. Ikot, U. S. Okorie, R. Sever, and G. J. Rampho, Eur. Phys. J. Plus **134**, 386 (2019).

[22] J. M. Deutsch, New J. Phys. **12**, 075021 (2010).

[23] R. H. Swendsen, Phys. Rev. E **92**, 052110 (2015).

[24] L. F. Santos, A. Polkovnikov, and M. Rigol, Phys. Rev. Lett. **107**, 040601 (2011).

[25] D. Šafránek, A. Aguirre, J. Schindler, and J. M. Deutsch, Found. Phys. **51**, 101 (2021).

[26] K. V. Hovhannisyan and L. A. Correa, Phys. Rev. B **98**, 045101 (2018).

[27] V. Mukherjee, A. Zwick, A. Ghosh, X. Chen, and G. Kurizki, Commun. Phys. **2**, 162 (2019).

[28] M. F. B. Cenni, L. Lami, A. Acin, and M. Mehboudi, Quantum **6**, 743 (2022).

[29] A. Engel and R. Nolte, Europhys. Lett. **79**, 10003 (2007).

[30] S. Alipour, F. Benatti, F. Bakhshinezhad, M. Afsary, S. Marcantoni, and A. T. Rezakhani, Sci. Rep. **6**, 35568 (2016).

[31] R. Modak and M. Rigol, Phys. Rev. E **95**, 062145 (2017).

[32] J. Goold, F. Plastina, A. Gambassi, and A. Silva, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, edited by F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso (Springer, Cham, 2018), pp. 317–336.

[33] G. De Chiara, P. Solinas, F. Cerisola, and A. J. Roncaglia, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, edited by F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso (Springer, Cham, 2018), pp. 337–362.

[34] A. D. Varizi, A. P. Vieira, C. Cormick, R. C. Drumond, and G. T. Landi, Phys. Rev. Res. **2**, 033279 (2020).

[35] A. E. Allahverdyan, R. Balian, and T. M. Nieuwenhuizen, Europhys. Lett. **67**, 565 (2004).

[36] R. Alicki and M. Fannes, Phys. Rev. E **87**, 042123 (2013).

[37] D. Šafránek, D. Rosa, and F. Binder, Phys. Rev. Lett. **130**, 210401 (2023).

[38] M. G. A. Paris, Int. J. Quantum Inf. **7**, 125 (2009).

[39] L. A. Correa, M. Mehboudi, G. Adesso, and A. Sanpera, Phys. Rev. Lett. **114**, 220405 (2015).

[40] L. Mandelstam and I. Tamm, in *Selected Papers*, edited by B. M. Bolotovskii, V. Y. Frenkel, and R. Peierls (Springer, Berlin, 1991), pp. 115–123.

[41] N. Margolus and L. B. Levitin, Physica D **120**, 188 (1998).

[42] S. Deffner and S. Campbell, J. Phys. A: Math. Theor. **50**, 453001 (2017).

[43] L. Villa and G. De Chiara, Quantum **2**, 42 (2018).

[44] J. L. Jiménez, S. P. G. Crone, E. Fogh, M. E. Zayed, R. Lortz, E. Pomjakushina, K. Conder, A. M. Läuchli, L. Weber, S. Wessel, A. Honecker, B. Normand, C. Rüegg, P. Corboz, H. M. Rønnow, and F. Mila, Nature (London) **592**, 370 (2021).

[45] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2002), Chap. 9.3.

[46] D. M. Reich, G. Gualdi, and C. P. Koch, Phys. Rev. Lett. **111**, 200401 (2013).

[47] R. Harper and S. T. Flammia, Quantum Sci. Technol. **2**, 015008 (2017).

[48] W. Huang, C. H. Yang, K. W. Chan, T. Tanttu, B. Hensen, R. C. C. Leon, M. A. Fogarty, J. C. C. Hwang, F. E. Hudson, K. M. Itoh, A. Morello, A. Laucht, and A. S. Dzurak, Nature (London) **569**, 532 (2019).

[49] J. Combes and A. P. Lund, Phys. Rev. A **106**, 063706 (2022).

[50] D. S. Hochbaum, Ann. Oper. Res. **153**, 257 (2007).

[51] M. A. Luersen and R. Le Riche, Comput. Struct. **82**, 2251 (2004).

[52] W. Price, J. Optim. Theory Appl. **40**, 333 (1983).

[53] R. Storn, *Proceedings of the North American Fuzzy Information Processing, Berkeley, 1996* (IEEE, Piscataway, 1996), pp. 519–523.

[54] S. Sivanandam and S. Deepa, *Introduction to Genetic Algorithms* (Springer, Berlin, 2008), pp. 165–209.

[55] D. Bertsimas and J. Tsitsiklis, Stat. Sci. **8**, 10 (1993).

[56] *Quantum Annealing and Related Optimization Methods*, edited by A. Das and B. K. Chakrabarti (Springer Science + Business Media, New York, 2005), Vol. 679.

[57] M. J. Powell, Acta Numer. **7**, 287 (1998).

[58] J. von Neumann, Eur. Phys. J. H **35**, 201 (2010).

[59] D. Šafránek, J. M. Deutsch, and A. Aguirre, Phys. Rev. A **99**, 010101(R) (2019).

[60] D. Šafránek, J. M. Deutsch, and A. Aguirre, Phys. Rev. A **99**, 012103 (2019).

[61] P. Strasberg and A. Winter, PRX Quantum **2**, 030202 (2021).

[62] F. Buscemi, J. Schindler, and D. Šafránek, New J. Phys. **25**, 053002 (2023).

[63] D. Šafránek and J. Thingna, arXiv:2007.07246.

[64] G. Usaj, H. M. Pastawski, and P. R. Levstein, Mol. Phys. **95**, 1229 (1998).

[65] C. Sánchez, P. Levstein, L. Buljubasich, H. Pastawski, and A. Chattah, Philos. Trans. R. Soc. A **374**, 20150155 (2016).

[66] H. M. Pastawski, P. R. Levstein, and G. Usaj, Phys. Rev. Lett. **75**, 4310 (1995).

[67] P. R. Levstein, A. K. Chattah, H. M. Pastawski, J. Raya, and J. Hirschinger, J. Chem. Phys. **121**, 7313 (2004).

[68] B. Rauer, S. Erne, T. Schweigler, F. Cataldini, M. Tajik, and J. Schmiedmayer, Science **360**, 307 (2018).

[69] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Rev. Mod. Phys. **91**, 021001 (2019).

[70] D. Porras and J. I. Cirac, Phys. Rev. Lett. **92**, 207901 (2004).

[71] A. Pal and D. A. Huse, Phys. Rev. B **82**, 174411 (2010).

[72] D. J. Luitz, N. Laflorencie, and F. Alet, Phys. Rev. B **91**, 081103(R) (2015).

[73] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, Nat. Phys. **12**, 907 (2016).

[74] P. Jurcevic, H. Shen, P. Hauke, C. Maier, T. Brydges, C. Hempel, B. P. Lanyon, M. Heyl, R. Blatt, and C. F. Roos, Phys. Rev. Lett. **119**, 080501 (2017).

[75] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z. X. Gong, and C. Monroe, Nature (London) **551**, 601 (2017).

[76] N. S. Bingham, S. Rooke, J. Park, A. Simon, W. Zhu, X. Zhang, J. Batley, J. D. Watts, C. Leighton, K. A. Dahmen, and P. Schiffer, Phys. Rev. Lett. **127**, 207203 (2021).

[77] B. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. Buyskikh, A. Daley, M. Cramer *et al.*, Nat. Phys. **13**, 1158 (2017).

[78] N. Friis, O. Marty, C. Maier, C. Hempel, M. Holzäpfel, P. Jurcevic, M. B. Plenio, M. Huber, C. Roos, R. Blatt, and B. Lanyon, Phys. Rev. X **8**, 021012 (2018).

[79] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Science **364**, 260 (2019).

[80] C. Maier, T. Brydges, P. Jurcevic, N. Trautmann, C. Hempel, B. P. Lanyon, P. Hauke, R. Blatt, and C. F. Roos, Phys. Rev. Lett. **122**, 050501 (2019).

[81] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Nature (London) **551**, 579 (2017).

[82] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Nat. Phys. **14**, 745 (2018).

[83] G.-X. Su, H. Sun, A. Hudomal, J.-Y. Desaules, Z.-Y. Zhou, B. Yang, J. C. Halimeh, Z.-S. Yuan, Z. Papić, and J.-W. Pan, Phys. Rev. Res. **5**, 023010 (2023).

[84] H. Bethe, Z. Phys. **71**, 205 (1931).

[85] M. J. Powell, in *Advances in Optimization and Numerical Analysis: Mathematics and Its Applications*, edited by S. Gomez and J. P. Hennart (Springer, Dordrecht, 1994), Vol. 275, pp. 51–67.

[86] H.-Y. Huang, R. Kueng, and J. Preskill, Nat. Phys. **16**, 1050 (2020).

[87] A. Zhao, N. C. Rubin, and A. Miyake, Phys. Rev. Lett. **127**, 110504 (2021).

[88] C. Hadfield, S. Bravyi, R. Raymond, and A. Mezzacapo, Commun. Math. Phys. **391**, 951 (2022).

[89] K. Bu, D. Enshan Koh, R. J. Garcia, and A. Jaffe, arXiv:2202.03272.

[90] S. H. Sack, R. A. Medina, A. A. Michailidis, R. Kueng, and M. Serbyn, PRX Quantum **3**, 020365 (2022).

[91] M. Ippoliti, arXiv:2305.10723.

[92] H.-Y. Hu and Y.-Z. You, Phys. Rev. Res. **4**, 013054 (2022).

[93] H.-Y. Hu, S. Choi, and Y.-Z. You, Phys. Rev. Res. **5**, 023027 (2023).

[94] A. Gresch and M. Kliesch, arXiv:2301.03385.

[95] A. A. Akhtar, H.-Y. Hu, and Y.-Z. You, Quantum **7**, 1026 (2023).

[96] A. Seif, Z.-P. Cian, S. Zhou, S. Chen, and L. Jiang, PRX Quantum **4**, 010303 (2023).

[97] T. Zhang, J. Sun, X.-X. Fang, X.-M. Zhang, X. Yuan, and H. Lu, Phys. Rev. Lett. **127**, 200501 (2021).

[98] G. I. Struchalin, Y. A. Zagorovskii, E. V. Kovlakov, S. S. Straupe, and S. P. Kulik, PRX Quantum **2**, 010307 (2021).

[99] M. Benedetti, J. Realpe-Gómez, R. Biswas, and A. Perdomo-Ortiz, Phys. Rev. A **94**, 022308 (2016).

[100] M. Pino and J. J. García-Ripoll, Phys. Rev. A **101**, 032324 (2020).

[101] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Rep. Prog. Phys. **83**, 054401 (2020).

[102] T. Imoto, Y. Seki, Y. Matsuzaki, and S. Kawabata, arXiv:2102.05323.

[103] N. Mohseni, M. Narozniak, A. N. Pyrkov, V. Ivannikov, J. P. Dowling, and T. Byrnes, npj Quantum Inf **7**, 71 (2021).

[104] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).

[105] M. B. Hastings, J. Stat. Mech. (2007) P08024.

[106] B. M. Terhal, Phys. Lett. A **271**, 319 (2000).

[107] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G. M. D'Ariano, and C. Macchiavello, Phys. Rev. Lett. **91**, 227901 (2003).

[108] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000).

[109] R. Orús, Ann. Phys. (NY) **349**, 117 (2014).

[110] J. Schindler, D. Šafránek, and A. Aguirre, Phys. Rev. A **102**, 052407 (2020).

[111] K. Swamy, IEEE Trans. Automat. Control **18**, 306 (1973).

[112] V. Khemani, M. Hermele, and R. Nandkishore, Phys. Rev. B **101**, 174204 (2020).

[113] M. Serbyn, D. A. Abanin, and Z. Papić, Nat. Phys. **17**, 675 (2021).

[114] A. Chandran, T. Iadecola, V. Khemani, and R. Moessner, Annu. Rev. Condens. Matter Phys. **14**, 443 (2023).