

Optimal quantum state tomography with local informationally complete measurements

Casey Jameson,^{1,*} Zhen Qin,² Alireza Goldar,³ Michael B. Wakin,³ Zhihui Zhu,² and Zhexuan Gong^{1,4,†}

¹*Department of Physics, Colorado School of Mines, Golden, Colorado 80401, USA*

²*Department of Computer Science and Engineering, The Ohio State University, Columbus, OH 43210, USA*

³*Department of Electrical Engineering, Colorado School of Mines, Golden, Colorado 80401, USA*

⁴*National Institute of Standard and Technology, Boulder, Colorado 80305, USA*

Quantum state tomography (QST) remains the gold standard for benchmarking and verification of quantum computers and quantum simulators. The current scale of experimental quantum devices already makes direct quantum state tomography intractable due to the exponentially large number of parameters in a generic quantum many-body state. However, most physical quantum states are structured and can often be represented by a much smaller number of parameters, making efficient QST potentially possible. A prominent example is a matrix product state (MPS) or a matrix product density operator (MPDO) with a small matrix dimension, which is believed to represent most physical states generated by one-dimensional (1D) quantum devices. We study whether a general MPS/MPDO state can be recovered with bounded errors using only a number of state copies polynomial in the number of qubits, which is necessary for efficient QST. To make this question practically interesting, we assume only local measurements of qubits directly on the target state. By using a local symmetric informationally complete positive operator-valued measurement (SIC-POVM) that requires only a single measurement setting, we provide a positive answer to the above question for a variety of common many-body quantum states, including typical short-range entangled states, random MPS/MPDO states, and thermal states of one-dimensional Hamiltonians. In addition, we also provide an affirmative no answer for certain long-range entangled states such as a family of generalized GHZ states, but with the exception of target states that are known to have real-valued wavefunctions. Our answers are supported by a near-perfect agreement between an efficient calculation of the Cramer-Rao bound that rigorously bounds the sample complexity and numerical optimization results using a machine learning assisted maximal likelihood estimation (MLE) algorithm. This agreement also leads to an optimal QST protocol using local SIC-POVM that can be practically implemented on current quantum hardware and is highly efficient for most 1D physical states. Our results also indicate that long-range entangled quantum states are in general not efficiently recoverable even if they are efficiently representable.

I. INTRODUCTION

With quantum computers rapidly scaling in size from the small-scale devices composed of only a handful of qubits to the current intermediate-scale devices composed of tens or even hundreds of qubits [1–3], the challenge of validating the performance of those devices scales at a similarly rapid pace. While a number of methods to verify the quantum state generated by intermediate-scale quantum devices have been proposed [4–8], the gold standard remains quantum state tomography (QST) [9–13], as it provides complete information about the experimental state. However, for a generic and possibly mixed N -qubit state, the number of state copies required to recover the state scales as $\mathcal{O}(2^N)$ even with the optimal measurement setting [14, 15]. Therefore, QST for generic quantum states is not scalable and only affordable on small-scale quantum devices.

In recent years, significant effort has been made toward the development of scalable QST methods by exploiting the fact that most physical states are structured in certain ways, allowing efficient classical representations through models such as matrix product states (MPS) [16], matrix product operators (MPO) [17], tensor network states [18], and quantum neural networks [19–21]. However, efficient representation of a state is in general not sufficient for efficient QST, which

requires the number of state copies, the number of elementary quantum gates needed for measurement, and the classical post-processing time to be all polynomial in N . While finding sufficient conditions for efficient QST is still a major challenge in quantum information theory, a practically more important challenge is to find efficient QST protocols that work for most physical states obtained on a given experimental quantum hardware.

In this paper, we attempt to address the above challenges with a focus on quantum states efficiently representable by MPS/MPO. These states incorporate most of the physical states generated by one-dimensional quantum computers or quantum simulators, including those with noise and errors [22]. An efficient representation here implies that the dimension of the matrices (often known as the bond dimension) in the MPS/MPO is either a constant or grows slowly with N , such that only a number of parameters polynomial in N is used to represent the state. Previous studies [9, 11, 12] show that efficient QST is possible if the MPS/MPO is locally reducible, meaning that the full system's state can be reconstructed with bounded errors using only reduced states on subsystems made of a small, N -independent number of qubits. Here efficient QST also assumes that the target state is reconstructed by an MPS/MPO with a small bond dimension. While the full system's wavefunction or density matrix cannot be reconstructed efficiently, most physically interesting properties, including non-local ones such as the entanglement entropy or topological order parameters, can be efficiently computed using the MPS/MPO representation of the state. In this regard, MPS/MPO based QST is more powerful than shadow

* cwjameson@mines.edu

† gong@mines.edu

tomography or randomized measurement toolbox [23, 24] that can usually only predict few-body correlations efficiently.

However, there are many physically interesting quantum states that are not locally reducible, such as certain long-range entangled states [25] and topologically ordered states [26, 27]. More importantly, mixed states, ubiquitous experimentally, are in general not locally reducible since those reduced states often do not uniquely determine the full system's state [9]. A complicated mathematical condition is needed to determine if a mixed state represented by an MPO is locally reducible [28]. Therefore, it remains unclear if efficient QST is possible for MPS/MPO states that are not locally reducible. In a recent work [29], we show that if one can measure all qubits in a global, Haar random basis, then any MPS/MPO with a finite bond dimension can be recovered with bounded errors using a number of state copies polynomial in N . This, however, does not imply efficient QST since the global Haar random measurement cannot be implemented efficiently. Therefore, here we restrict to only local measurement, meaning that each qubit is measured independently and individually by some POVM. Such measurement can be readily implemented in most current quantum hardware.

The first main contribution of this work is the development of an efficient method to identify the minimum number of state copies M required to perform QST on any given target state with an efficient MPS/MPO representation for a local, informationally complete POVM. This method is based on the Cramer-Rao bound in estimation theory [30] that lower bounds the statistical error of estimating parameters in a given probability distribution. While the Cramer-Rao bound has been widely used in evaluating the performance of various QST protocols [31–34], it is usually only evaluated for very small quantum systems. We show that by utilizing the structure of MPS/MPO, the Cramer-Rao bound can be computed efficiently via a novel autoregressive Monte Carlo sampling method for a large quantum many-body system. With this method, we show that typical short-range entangled states [35] and random MPS/MPO states can indeed be recovered efficiently, using a number of state copies proportional to the number of independent parameters in the target state. On the other hand, we reveal that there are certain long-range entangled states, e.g. a family of generalized GHZ states, that requires an exponentially large (in N) number of state copies to be reconstructed with a finite error. But surprisingly, if the target state is known to be represented by a real-valued MPS/MPO, this exponential scaling vanishes, despite the fact that the target state is still not locally reducible. Therefore, we argue that local reducibility does not by itself determine whether efficient QST via MPS/MPO is possible.

The second main contribution of our work is the development of a new QST protocol that combines the MPS/MPO model, local SIC-POVM, and a machine learning assisted MLE algorithm. We show that this protocol is optimal as it saturates the Cramer-Rao bound for a variety of physical target states within statistical errors. Theoretically, this optimality comes from the fact that MLE saturates the Cramer-Rao bound asymptotically in the large sample size limit [36–40]. There are also several practical advantages of our protocol:

First, the use of local SIC-POVM allows us to measure the quantum state in a single measurement setting/basis for any system size [41–44], in contrast to the more common local Pauli measurement that requires 3^N different measurement settings (i.e. measuring $\sigma_i^{x,y,z}$ for each qubit i). This removes the need for individual addressing of the qubits (individual readout of the qubits is still required), which can be significantly beneficial for certain quantum simulation platforms. For generic states, QST using SIC-POVM also requires less number of state copies than that based on the Pauli measurement for the same accuracy [15]. Moreover, for efficient QST, one cannot implement all 3^N measurement settings and the use of local Pauli measurement thus does not guarantee information completeness. While one can randomly pick the measurement basis [34] in hope of gaining information completeness statistically, this approach does not perform as well as our protocol and makes it challenging to evaluate the Cramer-Rao bound due to the randomness in the measurement settings. Experimentally, local SIC-POVM can be performed by mapping each qubit to a ququart using either ancilla qubits or extra states outside the qubit subspace, which has been recently demonstrated using trapped ions in a scalable way [45].

Second, a single measurement setting also allows us to evaluate physically interesting quantities in real time as experimental measurement samples are collected. Since our protocol does not rely on local reductions, one can estimate global quantities such as entanglement entropy or topological order parameters. The experimental state preparation and measurement cycles can be stopped once a desired convergence is reached [45]. Finally, by using state-of-the-art machine-learning methods, our MLE algorithm has an affordable classical runtime with efficient calculations of the gradients. It also works well for both pure and mixed target states, in contrast to many MPS based QST algorithms that cannot be easily adapted to mixed states [9, 34]. We have also benchmarked our algorithm using synthetic measurement data on a variety of common 1D physical states including typical short-range entangled states, random MPS/MPO states, and thermal states of 1D Hamiltonians. In most cases, the number of state copies needed to achieve high-fidelity QST is proportional to the number of parameters in the MPS/MPO representation of the target state, which ensures efficient QST.

The remainder of this paper is organized as follows: In Section II, we introduce the measurement setup (local SIC-POVM) used in this work and the model (MPS/MPO) used to reconstruct the target state. In Section III, we introduce the Cramer-Rao bound and show how to calculate it efficiently to lower bound the number of state copies needed to reconstruct a given target state. We compute the Cramer-Rao bound for various many-body entangled states with efficient MPS representations explicitly, revealing that efficient QST is possible for some target states but not the others. In Section IV, we introduce our MLE algorithm and benchmark it for a range of physical states in one dimension, showing that the fidelity of QST obtained using synthetic measurement samples closely matches that predicted by the Cramer-Rao bound. The paper concludes with an outlook on the remaining open questions.

II. MODEL AND MEASUREMENT SETTING

We start by introducing the measurement setting used throughout this paper, which is a local SIC-POVM on each qubit. SIC-POVM is a particular type of generalized quantum measurement that is highly symmetric, information complete, and having the minimal number of outcomes compatible with informational completeness. For a single qubit, a commonly used example of SIC-POVM is defined by four positive operators $M_i = |\phi_i\rangle\langle\phi_i|$ ($i = 1, 2, 3, 4$) with

$$\begin{aligned} |\phi_1\rangle &= \frac{1}{\sqrt{2}}|0\rangle & (1) \\ |\phi_2\rangle &= \frac{1}{\sqrt{6}}|0\rangle + \frac{1}{\sqrt{3}}|1\rangle \\ |\phi_3\rangle &= \frac{1}{\sqrt{6}}|0\rangle + \frac{1}{\sqrt{3}}e^{i\frac{2\pi}{3}}|1\rangle \\ |\phi_4\rangle &= \frac{1}{\sqrt{6}}|0\rangle + \frac{1}{\sqrt{3}}e^{-i\frac{2\pi}{3}}|1\rangle \end{aligned}$$

It is easy to check that $\sum_i M_i = \hat{1}$ and that $\{M_i\}$ spans the space of any single-qubit density operator. Therefore, $\{M_i\}$ form an informationally complete (IC) POVM. The symmetry of this IC-POVM manifests in $\text{Tr}(M_i M_j) = \frac{d\delta_{ij} + 1}{d^2(d+1)}$ for any i and j , where $d = 2$ for qubit.

Experimentally, the above SIC-POVM can be implemented by mapping the state of a qubit (with basis states $|0\rangle$ and $|1\rangle$) to a state of a ququart (with basis states $|0\rangle$, $|1\rangle$, $|2\rangle$, and $|3\rangle$) and performing a standard projective measurement for the ququart. The extra basis states ($|2\rangle$ and $|3\rangle$) can either come from additional states in the physical system carrying the qubit [45], or from an auxiliary qubit [35, 46]. The mapping from qubit to ququart can be done by applying the following unitary transformation on the ququart, starting from any initial state in the qubit subspace

$$U_{\text{SIC}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & e^{-i\frac{2\pi}{3}} & e^{i\frac{2\pi}{3}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & -\frac{1}{\sqrt{6}} \end{pmatrix}. \quad (2)$$

This unitary can be achieved by a combination of pulses driving the $|0\rangle - |1\rangle$, $|0\rangle - |2\rangle$, and $|0\rangle - |3\rangle$ transitions respectively [45]. Next, we measure the ququart projectively with its state collapsing to either $|0\rangle$, $|1\rangle$, $|2\rangle$, or $|3\rangle$. Denote the initial state of the qubit by $\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}$, it's not hard to see that the probabilities for the ququart measurement are given by

$$\begin{aligned} P(|0\rangle) &= \frac{1}{2}\rho_{00} & (3) \\ P(|1\rangle) &= \frac{1}{6} \left(\rho_{00} + 2\rho_{11} + \sqrt{2}\rho_{01} + \sqrt{2}\rho_{10} \right) \\ P(|2\rangle) &= \frac{1}{6} \left(\rho_{00} + 2\rho_{11} + \sqrt{2}e^{i\frac{2\pi}{3}}\rho_{01} + \sqrt{2}e^{-i\frac{2\pi}{3}}\rho_{10} \right) \\ P(|3\rangle) &= \frac{1}{6} \left(\rho_{00} + 2\rho_{11} + \sqrt{2}e^{-i\frac{2\pi}{3}}\rho_{01} + \sqrt{2}e^{i\frac{2\pi}{3}}\rho_{10} \right) \end{aligned}$$

The qubit's state ρ can thus be reconstructed from these four probabilities directly by solving the above set of linear equations. Since most quantum hardware platforms have single-shot readout of the qubit's state, a single-shot ququart readout can be experimentally achieved by three sequential qubit readouts with the help of π -pulses between the states $|2\rangle$ and $|3\rangle$ and the state $|0\rangle$ or $|1\rangle$ [45].

For multiple qubits, we perform the above SIC-POVM for each qubit independently, either in parallel or sequentially (since the measurements on different qubits mutually commute). The above unitary U_{SIC} for mapping each qubit to a ququart can also be implemented in parallel without individual addressing of the qubits. Therefore, our measurement setup requires only a constant circuit depth and is scalable.

Next, we introduce the class of quantum states we are targeting for our QST protocol. We consider a general target state represented by the density operator $\rho = \sum_{s,s'} \rho_{s,s'} |s\rangle\langle s'|$ expanded in the standard qubit basis $|s\rangle \equiv |s_1 s_2 \cdots s_N\rangle$ with $s_i = 0, 1$ for $i = 1, 2, \dots, N$. The class of states we focus on in this work takes the form of an MPO:

$$\rho_{s,s'} = \text{Tr}(A_1^{s_1 s'_1} A_2^{s_2 s'_2} \cdots A_N^{s_N s'_N}) \quad (4)$$

where each $A_i^{s_i s'_i}$ is a square matrix of dimension η (the assumption of square matrix here is more for convenience and can be relaxed). For the state to have an efficient MPO representation, we require η to be increasing at most polynomially in N . For simplicity, we assume η is a constant for the rest of the paper, as recent work suggests that states generated by a one-dimensional noisy quantum circuit can indeed be well approximated by MPOs with η independently of N [47]. Our results below can be straightforwardly generalized to the cases where η scales with N .

To make sure the MPO in the above Eq. (4) represents a physical density operator that is always Hermitian and semi-positive definite, we further assume that each matrix $A_i^{s_i s'_i}$ can be expressed as

$$A_i^{s_i s'_i} = \sum_{k=1}^{\kappa} C_{i,k}^{s_i} \otimes (C_{i,k}^{s'_i})^* \quad (5)$$

where each $C_{i,k}^{s_i}$ is a square matrix of dimension $\chi = \sqrt{\eta}$, which is often known as the *bond dimension* of an MPS. The upper limit of the summation κ is known as the *Kraus dimension*. Physically, we can view the state formed by $\{C_{i,k}^{s_i}\}$ as an MPS for a system where each qubit is coupled to an auxiliary qudit of dimension κ . The MPO formed by $\{A_i^{s_i s'_i}\}$ is then a reduced state of this MPS for the subsystem of N qubits and is therefore always physical. An MPO satisfying Eq. (5) is also called a matrix product density operator (MPDO) [17] or a 1D locally purified tensor network (LPTN) [48].

When the target state is a pure state $|\psi\rangle = \sum_s \psi_s |s\rangle$, the MPDO given by Eq. (5) reduces to an MPS formed by $\{C_i^{s_i}\}$ with the Kraus dimension being 1 and

$$\psi_s = \text{Tr}(C_1^{s_1} C_2^{s_2} \cdots C_N^{s_N}) \quad (6)$$

For a pure state, the MPS representation is preferred since the matrices in the MPO have a large dimension ($\eta = \chi^2$).

The MPS/MPO has a nice structure that any local unitary transformation on the state only changes the corresponding matrices locally. Since the local SIC-POVM measurement setup we introduced earlier only involves local unitaries, we can represent the probability $P(\mathbf{m})$ of getting a specific measurement outcome of the ququarts $\mathbf{m} = (m_1, m_2, \dots, m_N)$ with $m_i = 0, 1, 2, 3$ also in an MPO format:

$$P(\mathbf{m}) = \text{Tr} (B_1^{m_1} \dots B_N^{m_N}) \quad (7)$$

$$B_i^0 = \frac{A_i^{00}}{2}$$

$$B_i^1 = \frac{1}{6} \left(A_i^{00} + 2A_i^{11} + \sqrt{2}A_i^{01} + \sqrt{2}A_i^{10} \right)$$

$$B_i^2 = \frac{1}{6} \left(A_i^{00} + 2A_i^{11} + \sqrt{2}e^{i\frac{2\pi}{3}}A_i^{01} + \sqrt{2}e^{-i\frac{2\pi}{3}}A_i^{10} \right)$$

$$B_i^3 = \frac{1}{6} \left(A_i^{00} + 2A_i^{11} + \sqrt{2}e^{-i\frac{2\pi}{3}}A_i^{01} + \sqrt{2}e^{i\frac{2\pi}{3}}A_i^{10} \right)$$

Performing the local SIC-POVM on all qubits in the target state provides unbiased sampling of the above probability distribution $P(\mathbf{m})$. To achieve QST, we want to estimate the matrices $\{A_i^{s_i s'_i}\}$ that determine the distribution $P(\mathbf{m})$. This is however not an easy task because in practice we can only obtain a small number of measurement samples (at most polynomial in N) compared to the total number of \mathbf{m} values (which is 4^N). Therefore, the estimation is in general not perfect and it is not obvious how we can perform the most accurate estimation based on a given number of samples. Directly solving the matrices $\{A_i^{s_i s'_i}\}$ from the above nonlinear equations is generally not possible for a large N , and an efficient classical algorithm is needed. We address these challenges in the next two sections.

III. CRAMER-RAO BOUNDS

QST can be generally treated as a parameter estimation problem. Our goal is to estimate the density matrix (the collection of its matrix elements) for a target state ρ_0 using some model (also known as ansatz in quantum physics). Denote the estimation of ρ_0 by ρ , we can quantify the QST error using

$$\mathcal{R} = \|\rho - \rho_0\|_2^2 \equiv \text{Tr}[(\rho - \rho_0)^2] \quad (8)$$

If we know the target state is a pure state $|\psi_0\rangle$, we can use a pure state $|\psi\rangle$ to approximate $|\psi_0\rangle$ and the above quantity \mathcal{R} reduces to

$$\mathcal{R} = 2(1 - F) \quad (9)$$

where $F \equiv |\langle\psi|\psi_0\rangle|^2$ is the fidelity between $|\psi\rangle$ and $|\psi_0\rangle$. If both ρ and ρ_0 are mixed, \mathcal{R} is usually normalized by the purity of the target state $\text{Tr}(\rho_0^2)$ to avoid \mathcal{R} being artificially small due to a low purity of the target state [11, 12].

We assume that the model state ρ used to estimate the target state is a function of a list of independent real parameters $\boldsymbol{\theta} \equiv (\theta_1, \dots, \theta_N)$. In the current case of MPS/MPDO-based QST, $\boldsymbol{\theta}$ is the list of all matrix elements (including both

their real and imaginary parts) of $\{C_{i,k}^{s_i}\}$ in the MPDO representation of the state [see Eqs. (4)-(5)]. Note that there is no constraint on the MPS/MPDO parameters since ρ is guaranteed Hermitian and semi-positive definite. The normalization condition $\text{Tr}(\rho) = 1$ is not necessary since one can divide the above distance metric \mathcal{R} by $\text{Tr}(\rho)^2$ instead, and the quantity $\text{Tr}(\rho)$ can be computed efficiently due to the MPS/MPDO representation of ρ [16, 17]. The target state ρ_0 is obtained by setting $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Here we assume the model state's MPDO has the same bond dimension as the target state's for simplicity. In practice, the target state is usually unknown and one need to make a guess of the model state's bond dimension. This guess can be refined by increasing or decreasing the bond dimension until a convergence of the MLE cost function is reached.

Define ρ_k to be the k^{th} element of the matrix ρ in the computational basis where $k = 1, 2, \dots, 4^N$ is a linearized index of (s, s') , and $\tilde{\rho} \equiv \rho - \rho_0$. We now invoke the Cramer-Rao bound to lower bound the covariance between $\tilde{\rho}_k$ and $\tilde{\rho}_l$ when M observations of the target state is made [30, 34]:

$$\text{Cov}(\tilde{\rho}_k, \tilde{\rho}_l) \geq \frac{1}{M} \sum_{\alpha, \beta} \frac{\partial \rho_k}{\partial \theta_\alpha} (I^{-1})_{\alpha \beta} \frac{\partial \rho_l^*}{\partial \theta_\beta} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \quad (10)$$

where $\alpha = 1, 2, \dots, N$ is an index of the parameter set $\boldsymbol{\theta}$. $\text{Cov}(\tilde{\rho}_k, \tilde{\rho}_l)$ denotes the statistical average of $\tilde{\rho}_k \tilde{\rho}_l$ where the statistical fluctuations are due to a finite number (denoted by M) of random samples used in estimating ρ_0 . We note that the above Cramer-Rao bound requires ρ to be an unbiased estimator of ρ_0 , meaning that expectation value of $\tilde{\rho}$ should be a zero matrix, which may not be true. However, recent work [31, 32] shows that the Cramer-Rao bound in Eq. (10) is asymptotically correct (for a large sample size M) as long as ρ is a well-behaved estimator of ρ_0 , which is indeed the case here since each element of ρ is an analytic function of the parameter set $\boldsymbol{\theta}$ and that $\rho = \rho_0$ upon setting $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

In the above Eq. (10), I is the Fisher information matrix whose elements are defined as

$$I_{\alpha \beta} = \sum_m P(\mathbf{m}) \frac{\partial \log P(\mathbf{m})}{\partial \theta_\alpha} \frac{\partial \log P(\mathbf{m})}{\partial \theta_\beta} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \quad (11)$$

which depends on the probability $P(\mathbf{m})$ we are sampling from experimentally. Since $P(\mathbf{m})$ is a real-valued function of a set of real parameters $\boldsymbol{\theta}$, it's not hard to see that I is real, symmetric, and semi-positive definite.

Noting that $\mathbb{E}(\mathcal{R}) = \sum_k \text{Cov}(\tilde{\rho}_k, \tilde{\rho}_k)$ where \mathbb{E} denotes the statistical average (over many different sets of measurement samples), the Cramer-Rao bound in Eq. (10) reduces to [34]

$$\mathbb{E}(\mathcal{R}) \geq \frac{1}{M} \text{Tr}(K I^{-1}) \quad (12)$$

$$K_{\alpha \beta} \equiv \sum_k \frac{\partial \rho_k}{\partial \theta_\alpha} \frac{\partial \rho_k^*}{\partial \theta_\beta} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \quad (13)$$

where K is a square matrix of dimension N that is always Hermitian and semi-positive definite. Here I^{-1} denotes the pseudo-inverse of I . A matrix inverse of I usually does not exist since $P(\mathbf{m})$ can be independent of certain linear combinations of the parameters $(\theta_1, \dots, \theta_N)$, which leads to a null

space of I . This happens when there are gauge degrees of freedom of the MPS/MPDO. For example, adding a phase on each element of a $C_{i,k}^{s_i}$ matrix for both $s_i = 0$ and $s_i = 1$ will leave the MPDO in Eq. (5) invariant, and thus $P(\mathbf{m})$ is independent of such phase. The pseudo-inverse on I avoids this problem and allows for a proper calculation of the Cramer-Rao bound, because the K matrix will share the same null space with I . A more physical way to see this is to note that $\text{Tr}(KI^{-1})$ bounds the distance metric \mathcal{R} that has no gauge degree of freedom since $P(\mathbf{m})$ cannot be identical for two states ρ and ρ_0 that are not identical due to information completeness.

A key contribution made in this paper is to show that the K matrix can be evaluated efficiently while the I matrix can be estimated efficiently when the target state ρ_0 takes the MPDO form in Eq. (5). Let us first try to evaluate the matrix element $K_{\alpha\beta}$ using the structure of the MPO. For convenience, we index each real parameter θ_α in the matrix $C_{i,k}^s$ by $\alpha = (i, j, k, s, \xi)$ where $j = 1, 2, \dots, \chi^2$ denotes the linear index of the corresponding matrix element. $\xi = 0$ (1) if θ_α is the real (imaginary) part of the matrix element. One can then calculate $\frac{\partial \rho_k}{\partial \theta_\alpha}$ analytically as

$$\frac{\partial \rho_{\mathbf{s}, \mathbf{s}'}}{\partial \theta_\alpha} = \text{Tr} \left(A_1^{s_1 s'_1} \dots D_{i,\alpha}^{s_i s'_i} \dots A_N^{s_N s'_N} \right) \quad (14)$$

$$D_{i,\alpha}^{s_i s'_i} = i^\xi \delta_{s_i, s} E_j \otimes \left(C_{i,k}^{s'_i} \right)^* + i^{-\xi} \delta_{s'_i, s} C_{i,k}^{s_i} \otimes E_j$$

where E_j denotes a $\chi \times \chi$ matrix of all zeros except that its j^{th} element is 1.

Eq. (14) is in an MPO format and we can also treat it as an MPS for N ququarts. Therefore, we can evaluate $K_{\alpha\beta} = \sum_k \frac{\partial \rho_k}{\partial \theta_\alpha} \frac{\partial \rho_k^*}{\partial \theta_\beta} \Big|_{\theta=\theta_0}$ by contracting two such MPS, which can be performed efficiently on a classical computer [16].

For calculating the Fisher information matrix I , we first calculate the derivative $\frac{\partial P(\mathbf{m})}{\partial \theta_\alpha}$ used in $I_{\alpha\beta}$ analytically as

$$\frac{\partial P(\mathbf{m})}{\partial \theta_\alpha} = \text{Tr} \left(B_1^{m_1} \dots G_{i,\alpha}^{m_i} \dots B_N^{m_N} \right) \quad (15)$$

$$G_{i,\alpha}^0 = \frac{1}{2} D_{i,\alpha}^{00}$$

$$G_{i,\alpha}^1 = \frac{1}{6} \left(D_{i,\alpha}^{00} + 2D_{i,\alpha}^{01} + \sqrt{2}D_{i,\alpha}^{01} + \sqrt{2}D_{i,\alpha}^{10} \right)$$

$$G_{i,\alpha}^2 = \frac{1}{6} \left(D_{i,\alpha}^{00} + 2D_{i,\alpha}^{01} + \sqrt{2}e^{i\frac{2\pi}{3}} D_{i,\alpha}^{01} + \sqrt{2}e^{-i\frac{2\pi}{3}} D_{i,\alpha}^{10} \right)$$

$$G_{i,\alpha}^3 = \frac{1}{6} \left(D_{i,\alpha}^{00} + 2D_{i,\alpha}^{01} + \sqrt{2}e^{-i\frac{2\pi}{3}} D_{i,\alpha}^{01} + \sqrt{2}e^{i\frac{2\pi}{3}} D_{i,\alpha}^{10} \right)$$

which are identical to Eq. (7) upon replacing $A_i^{s_i s'_i}$ by $D_{i,\alpha}^{s_i s'_i}$ and $B_i^{m_i}$ by $G_{i,\alpha}^{m_i}$.

While $\frac{\partial P(\mathbf{m})}{\partial \theta_\alpha}$ is also in an MPO format, we cannot evaluate $I_{\alpha\beta}$ efficiently because $I_{\alpha\beta} = \sum_{\mathbf{m}} \frac{1}{P(\mathbf{m})} \frac{\partial P(\mathbf{m})}{\partial \theta_\alpha} \frac{\partial P(\mathbf{m})}{\partial \theta_\beta}$ [see Eq. (11)] contains an extra weight $\frac{1}{P(\mathbf{m})}$. However, we can estimate $I_{\alpha\beta}$ using Monte Carlo sampling by using the fact that both $P(\mathbf{m})$ and $\frac{\partial \log P(\mathbf{m})}{\partial \theta_\alpha} = \frac{1}{P(\mathbf{m})} \frac{\partial P(\mathbf{m})}{\partial \theta_\alpha}$ can be computed efficiently for a given \mathbf{m} . More interestingly, we can perform direct sampling over the probability distribution $P(\mathbf{m})$

instead of the commonly used Metropolis-Hastings algorithm in variational Monte Carlo method that may suffer from poor sampling quality. This direct sampling is achieved by utilizing a special nature of the MPO representation of $P(\mathbf{m})$ similar to autoregressive artificial neural networks [49]. In particular, we can calculate the following series of conditional probabilities efficiently:

$$P(m_1) = \text{Tr} (B_1^{m_1} \mathcal{B}_2 \dots \mathcal{B}_N) \quad (16)$$

$$P(m_2|m_1) = \text{Tr} (B_1^{m_1} B_2^{m_2} \dots \mathcal{B}_N)$$

$$\vdots$$

$$P(m_N|m_{N-1} \dots m_1) = \text{Tr} (B_1^{m_1} \dots B_N^{m_N})$$

where we defined $\mathcal{B}_i = \sum_{m_i} B_i^{m_i}$. We can then perform the sampling of $P(\mathbf{m})$ efficiently (which we call autoregressive sampling) as follows:

1. Calculate $P(m_1)$ for $m_1 = 0, 1, 2, 3$, then sample from this distribution [normalized by $\sum_{m_1} P(m_1)$ if $P(\mathbf{m})$ is not normalized] to get a random value of $m_1 \in \{0, 1, 2, 3\}$.
2. Depending on the value of m_1 sampled in the previous step, calculate $P(m_2|m_1)$ for $m_2 = 0, 1, 2, 3$ and sample from this distribution (again normalized) to get a random value of m_2 .
3. Iterate this process to determine m_3, \dots, m_N .

Note that the above process can be done in parallel to create many samples of $\mathbf{m} = (m_1, \dots, m_N)$ efficiently on a classical computer. We then compute $\frac{\partial \log P(\mathbf{m})}{\partial \theta_\alpha} \frac{\partial \log P(\mathbf{m})}{\partial \theta_\beta}$ for all α, β indices with each sampled \mathbf{m} and average the obtained Fisher information matrix I over all samples to get an estimation of I . We increase the number of samples \mathcal{M} until I converges. For most physical target states we have studied, we find $\mathcal{M} = 10^4 - 10^5$ sufficient to ensure accurate estimation of I for $N \leq 40$ qubits, and the calculation of the entire I matrix takes only a few minutes on a modern PC for a small bond dimension χ .

With both K and I efficiently computable, we can now compute $\text{Tr}(KI^{-1})$ in the Cramer-Rao bound [Eq. (12)] for a variety of different target states. For simplicity, we focus on pure target states as exact MPS representations exist for a variety of many-body entangled states. This also allows us to better interpret the Cramer-Rao bound using the structure of entanglement, since characterizing entanglement is much harder for a mixed state. Moreover, for a pure state we can directly lower bound the infidelity $(1 - F)$ of the QST using $\text{Tr}(KI^{-1})$ and the sample size M using Eq. (9):

$$\mathbb{E}(1 - F) \geq \frac{1}{2M} \text{Tr}(KI^{-1}) \quad (17)$$

We start by calculating $\text{Tr}(KI^{-1})$ for a set of random MPS states where the target state is expressed by Eq. (6) with each $C_i^{s_i}$ being a random $\chi \times \chi$ real matrix. The elements of each random matrix is drawn uniformly between -1 and 1 . The model state is also an MPS with bond dimension χ , and we

allow each matrix element in the MPS to be an independent parameter that is either real or complex, such that there are in total $2N\chi^2$ or $4N\chi^2$ real parameters. We compute $\text{Tr}(KI^{-1})$ for N up to 40 and χ up to 10. The results shown in Fig. 1(a) indicate that $\text{Tr}(KI^{-1}) \approx N\chi^2$ for a real MPS model and $\text{Tr}(KI^{-1}) \approx 2.2N\chi^2$ for a complex MPS model.

It may be counter-intuitive that the complex MPS model leads to a larger Cramer-Rao bound than the real MPS model for the same target state, as it appears that the complex model is just an over-parameterization. However, we note that the extra parameters in the complex model are actually fully identifiable by our informationally-complete measurements, in contrast to the gauge degrees of freedom which are not identifiable. Therefore, these extra parameters will contribute to the Cramer-Rao bound and actually increase the statistical fluctuations for the estimation of the target state with a finite number of measurement samples. In both the real and the complex cases, we find that $\text{Tr}(KI^{-1})$ is proportional to the number of independent parameters describing the target state, consistent with a rough estimate of $\text{Tr}(KI^{-1})$ obtained in Ref. [34] for random local basis measurement (which should be similar to the local SIC-POVM when the number of random bases is large). Therefore, for random MPS states, we expect efficient QST using the local SIC-POVM.

Next, we change our target state to a 1D cluster state, which is a typical short-range entangled state [25]. This state is translationally invariant (TI) and described by the following MPS of bond dimension $\chi = 2$: [50]

$$C_{\text{cluster}}^0 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \quad C_{\text{cluster}}^1 = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad (18)$$

Physically, this state can be generated by applying controlled-Z gates on all neighboring pairs of qubits on a 1D ring starting from an initial product state of all qubits in the state $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$. It is short-range entangled as it can be generated by a single-layer local quantum circuit.

As shown by Fig. 1(c), we find that for the cluster state, $\text{Tr}(KI^{-1})$ scales linearly in N with the complex model again having its value roughly twice of the real model, similar to the results for the random MPS state [see Fig. 1(a-b)]. Note that the linear in N scaling here is due to the fact that the model state is not TI and has $\mathcal{O}(N)$ parameters that are identifiable by the measurements. A TI model would lead to $\text{Tr}(KI^{-1})$ independent of N . As a result, we see that the number of state copies needed to reconstruct a given target state within a given error depends on how many parameters are used in the model. We can use less number of parameters if we have more *a priori* knowledge of the target state, such as efficient MPS/MPO representation, real vs complex wavefunctions, TI vs non-TI, etc. In other words, the knowledge we have for the target state leads to a constraint on the estimation problem, which can significantly lower the Cramer-Rao bound [51].

We now consider a paradigmatic example of a long-range entangled state, a GHZ state $|\psi_{\text{GHZ}}\rangle = |00\cdots 0\rangle + |1\cdots 1\rangle$ (not normalized for convenience). This state requires a local quantum circuit of depth $\mathcal{O}(N)$ to be created from a product

state. It is described by a TI MPS with

$$C_{\text{GHZ}}^0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad C_{\text{GHZ}}^1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (19)$$

Since both matrices are diagonal, we can in fact calculate the K and I matrices semi-analytically with either a TI or non-TI MPS model. We defer the detailed analytical calculations to the Appendix A, and summarize the results below.

If we assume the model MPS contains only real parameters, then $\text{Tr}(KI^{-1})$ quickly converges to $\frac{7}{2}$ for a TI model and $2N + \frac{3}{2}$ for a non-TI model [see Fig. 2(a)]. However, if we assume each matrix element can be complex, then $\text{Tr}(KI^{-1})$ quickly converges to 2^{N-1} for both a TI and a non-TI model [see Fig. 2(b)]. The real MPS model behaves similarly to the target states we studied previously, where $\text{Tr}(KI^{-1})$ scales with the number of parameters in the model. The complex MPS model leads to a sharply different scenario, with $\text{Tr}(KI^{-1})$ scales exponentially in N .

This exponential scaling is due to the difficulty in determining the relative phase between the two component states $|0\cdots 0\rangle$ and $|1\cdots 1\rangle$. To illustrate this, we consider a minimal model for estimating the GHZ state with just two real parameters ϕ_1 and ϕ_2 :

$$C_{\phi_1}^0 = \begin{pmatrix} e^{i\phi_1} & 0 \\ 0 & 0 \end{pmatrix}, \quad C_{\phi_2}^1 = \begin{pmatrix} 0 & 0 \\ 0 & e^{i\phi_2} \end{pmatrix} \quad (20)$$

It's easy to see that this model represents a modified GHZ state $|\phi_{\text{GHZ}}\rangle = e^{iN\phi_1}|0\cdots 0\rangle + e^{iN\phi_2}|1\cdots 1\rangle$. The analytical calculations described in Appendix A allow us to find the K and I matrices explicitly:

$$K = \frac{N^2}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad I = N^2 \delta \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad (21)$$

where $\delta \approx \frac{1}{2^N}$ and $\text{Tr}(KI^{-1}) \approx 2^{N-1}$ (asymptotically exact in the large N limit), same as that using a general complex MPS model with $\chi = 2$ (see Fig. 2(b) and Appendix A). There are two shared eigenvectors of K and I in Eq. (21). One has eigenvalue 0, corresponding to $\phi_1 + \phi_2$ which only changes the global phase of the state $|\phi_{\text{GHZ}}\rangle$. The other has eigenvalue $\delta \approx 2^{-N}$ for the I matrix, corresponding to $\phi_1 - \phi_2$, which controls the relative phase between the two component states in $|\psi_{\text{GHZ}}\rangle$. An exponentially small eigenvalue of the Fisher information matrix denotes that the measurement we performed gives little information for $\phi_1 - \phi_2$, and thus requires an exponentially large sample size in N to obtain a good estimate.

The real MPS model cannot estimate this relative phase and hence does not lead to an exponential scaling in the Cramer-Rao bound. The real MPS model implies a constraint on the target state, requiring it to be represented only by a real MPS. While such a constraint is expected to lower the Cramer-Rao bound, it is surprising that this constraint leads to an exponentially smaller bound. It should also be pointed out that constraining the target state to have real wavefunctions in the computational basis (up to arbitrary single-qubit rotations) does not make QST easy or trivial since there are many interesting and complicated many-body entangled states that are

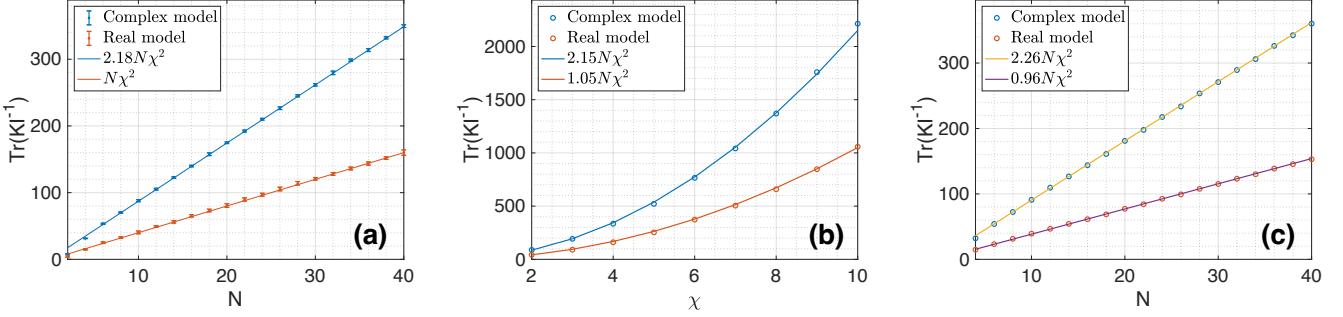


Figure 1. $\text{Tr}(KI^{-1})$ calculated for the random MPS target states as a function of N for $\chi = 2$ (a), as a function of χ for $N = 10$ (b), and for the 1D cluster state (c). We calculate $\text{Tr}(KI^{-1})$ for 10 different random MPS target states and show their mean and standard deviation (as error bars) in (a). For each target state, we use a model MPS state with the same bond dimension and assume the matrices in the model state to be either real (blue) or complex (red). The solid lines in each plot are linear fits to the calculated $\text{Tr}(KI^{-1})$ values. $\mathcal{M} = 10^5$ samples is used for the Monte Carlo sampling calculation of the Fisher information matrix. Further increasing \mathcal{M} does not lead to visible differences for all the plots.

real. For example, the eigenstates of any real Hamiltonian are real. Therefore, the real MPS model used to achieve efficient QST for the GHZ state is also capable of reconstructing any quantum many-body state represented by a real MPS (noting that the bond dimension of the MPS model can be increased without changing the Cramer-Rao bound for the GHZ state). We thus argue that local reducibility is not necessary for efficient QST using MPS/MPO, since a real GHZ state is not locally reducible. We conjecture that efficient QST is possible for any target state with an efficient real MPS representation.

Finally, we have also considered a generalized GHZ state $|\Psi_{\text{GHZ}}\rangle = U_{\gamma}^{\otimes N}|\psi_{\text{GHZ}}\rangle$ that differs from the standard GHZ state by a global $\text{SO}(2)$ rotation, where $U_{\gamma} = \begin{pmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{pmatrix}$. $|\Psi_{\text{GHZ}}\rangle$ is represented by the TI-MPS

$$C_{\gamma}^0 = \begin{pmatrix} \cos \gamma & 0 \\ 0 & \sin \gamma \end{pmatrix}, \quad C_{\gamma}^1 = \begin{pmatrix} -\sin \gamma & 0 \\ 0 & \cos \gamma \end{pmatrix} \quad (22)$$

Interestingly, a complex MPS model now yields $\text{Tr}(KI^{-1}) \sim \zeta^N$ where the base of the exponential ζ depends on the rotation angle γ . As shown in Fig. 2(c), the original GHZ state (where $\gamma = 0$) in fact has the largest ζ ($\zeta_{\text{max}} = 2$), while $\zeta_{\text{min}} \approx 1.24$ is achieved for $\gamma \approx \frac{\pi}{3}$. This implies that there also exists an optimal local SIC-POVM, one that differs from our choice [Eq. (1)] by the local unitary transformation U_{γ} . While we believe that the exponential scaling of $\text{Tr}(KI^{-1})$ applies to any local SIC-POVM, an optimal choice could lead to a much smaller infidelity for QST. However, finding such optimal measurement setting would require *a priori* knowledge of the target state and is not preferred for a universal QST protocol.

IV. MLE ALGORITHM AND BENCHMARK RESULTS

The Cramer-Rao bound studied in the previous section provides a lower bound on the expectation value of the QST infidelity (for pure states) for a given number of state copies

and a specific target state. Whether this lower bound can be achieved is a separate question. It is well known that the MLE algorithm can saturate the Cramer-Rao bound [36] theoretically in the large sample size limit. However, in practice the MLE algorithm may not be able to find the global optimum when the number of parameters is large and thus the saturation of the Cramer-Rao bound is not guaranteed.

Here we describe the MLE algorithm we used and benchmark its performance using synthetic experimental data and compare it with the Cramer-Rao bound calculated using the method developed in the previous section. Specifically, we define a cost function that is the negative log-likelihood over all M samples obtained during the local SIC-POVM measurements for the model MPS/MPDO state:

$$D_{\text{NLL}} = -\frac{1}{M} \sum_{(m_1, \dots, m_N) \in \mathcal{V}} \log (\text{Tr}[B_1^{m_1} \cdots B_N^{m_N}]) \quad (23)$$

where \mathcal{V} is the set of all M samples (each sample is a N -quart string representing a measurement outcome). Our goal is to minimize D_{NLL} by optimizing the matrices $\{C_{i,k}^{s_i}\}$ in the model state with a guessed bond dimension and Kraus dimension. The matrices $\{B_i^{m_i}\}$ in Eq. (23) are determined from $\{C_{i,k}^{s_i}\}$ using Eq. (5) and Eq. (7).

The optimization of a cost function such as the D_{NLL} above that depends on many parameters in a complicated, nonlinear way is a central goal in modern machine learning tasks. Here we use the well-known PyTorch library to perform such optimization with a stochastic gradient descent (SGD) optimizer or an Adam optimizer. To better achieve the global minimum of the cost function D_{NLL} , we have also used several different methods for scheduling the learning rate of the optimizer, including a Nesterov momentum method and a cosine annealing scheduler. In addition, each optimization is performed over multiple random initial guesses of the matrices $\{C_{i,k}^{s_i}\}$, and we pick the run with the lowest converged cost function D_{NLL} and then calculate the distance metric \mathcal{R} [see in Eq. (8)] between the corresponding model state and the target state. We note that \mathcal{R} can be computed efficiently as long as the

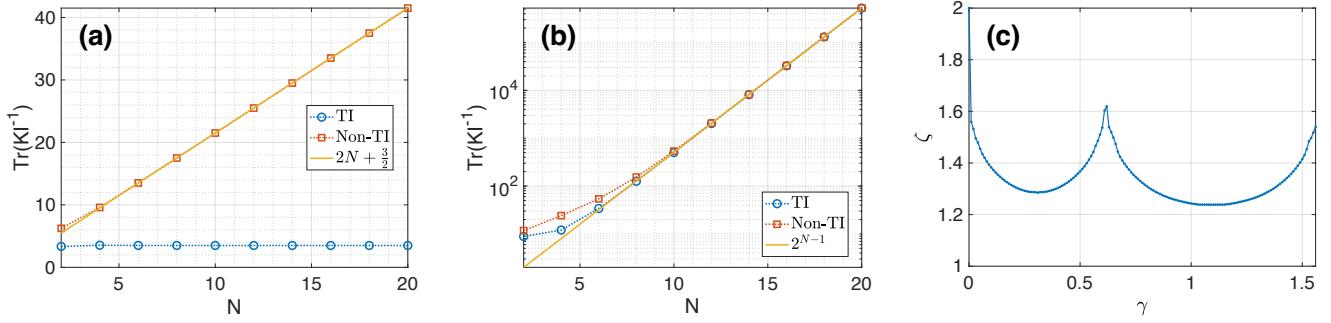


Figure 2. (a) $\text{Tr}(KI^{-1})$ calculated semi-analytically using both a translationally invariant (TI) and a non-translationally invariant (Non-TI) real MPS model for the GHZ state in Eq. (19). (b) Same as (a) but with a complex MPS model. (c) $\text{Tr}(KI^{-1})$ for the generalized GHZ state in Eq. (22) is found to be proportional to ζ^N for large N . The plot shows the value of ζ as a function of the rotation angle γ obtained via a fit of the semi-analytically computed $\text{Tr}(KI^{-1})$ values for $N = 20, 22, \dots, 60$.

model state and target state are both efficiently represented using MPS/MPO.

The input data to the optimizer is the sample set \mathcal{V} which can be generated experimentally. Here we generate the sample set synthetically for a given MPS/MPDO target state using the autoregressive sampling method described in Section III. Since the sample size M is finite, we shall repeat the entire optimization for multiple sets (10 for each of the calculations below) of random samples with the same size and average over the calculated \mathcal{R} to obtain $\mathbb{E}(\mathcal{R})$.

We first benchmark our MLE algorithm using the random MPS state, the 1D cluster state, and a generalized GHZ state as target states. For the random MPS state, we assume both the target state and the model state to be made of complex matrices of bond dimension 2. For the cluster state, we use a non-TI complex MPS model of bond dimension 2, and for the generalized GHZ state, we use a complex TI MPS model of bond dimension 2, with the target state given by Eq. (22) with the rotation angle $\gamma = \frac{\pi}{3}$. In all three cases, we see that the infidelity between the target state and the state obtained via the MLE algorithm closely match the minimum infidelity predicted by the Cramer-Rao bound [see Eq. (17)], as shown in Fig. 3. Small differences between the two are likely due to statistical fluctuations (since we have only performed the MLE algorithm for 10 different sets of random measurement samples) and limited sample size (since the MPS/MPDO model is likely not an unbiased estimator and the Cramer-Rao bound could be inaccurate if the sample size is not large enough).

We have also benchmarked our MLE algorithm using mixed target states. The infidelity between the target state ρ_0 and the state reconstructed by the MLE algorithm (denoted by ρ) cannot be computed efficiently, but we can compute the following distance metric between the two states

$$\mathcal{D} = \|\rho - \rho_0\|_2^2 / \|\rho_0\|_2^2 = \mathcal{R} / \text{Tr}(\rho_0^2), \quad (24)$$

which is the aforementioned distance metric \mathcal{R} normalized by the purity of the target state that can also be computed efficiently due to the efficient MPO representations of ρ_0 .

We first test our MLE algorithm for reconstructing random MPDO states where each $\{C_{i,k}^{s_i}\}$ matrix in Eq. (5) is a $\chi \times \chi$

random complex matrix. We choose a model MPDO state with the same bond dimension and Kraus dimension as the target state. Second, we choose target states that are thermal states of the following 1D quantum Ising model

$$H_I = \sum_{j=1}^{N-1} \sigma_j^z \sigma_{j+1}^z + B \sum_{j=1}^N \sigma_j^x. \quad (25)$$

We set $B = 1$, which makes the model at its quantum criticality in the thermodynamic limit. The thermal state is defined as $\rho_T = \frac{e^{-H_I/T}}{\text{Tr}(e^{-H_I/T})}$. To obtain ρ_T for a large system size, we use an imaginary time TEBD method [52] to approximate ρ_T using an MPO. The MLE algorithm is then used to reconstruct the target MPO state as an MPDO.

Fig. 4 indicates that for both the random MPDO states and the thermal states, efficient QST is possible as the distance metric \mathcal{D} scales approximately linearly in N . We also point that the performance of our QST protocol that combines the MLE algorithm with the local SIC-POVM compares favorably to existing QST protocols using MPS or MPO [9, 11, 12, 34], as the distance metric we achieved is close to the theoretical limit imposed by the Cramer-Rao bound. Importantly, our QST protocol is fully general and works for any target state, with the number of measurement samples expected to scale only linearly in N for most target states represented by MPS/MPDO with a finite bond dimension.

V. CONCLUSION AND OUTLOOK

In this work, we have studied the tomography of a quantum many-body state using a single measurement setting made of local SIC-POVM. This measurement setting can be readily implemented in most current quantum hardware and removes the need of individual qubit controls. We apply the local SIC-POVM measurement to a class of target states that have efficient MPS/MPDO representations and are ubiquitous in one-dimensional physical systems. For most target states in this class, we find efficient QST is possible, with the number of

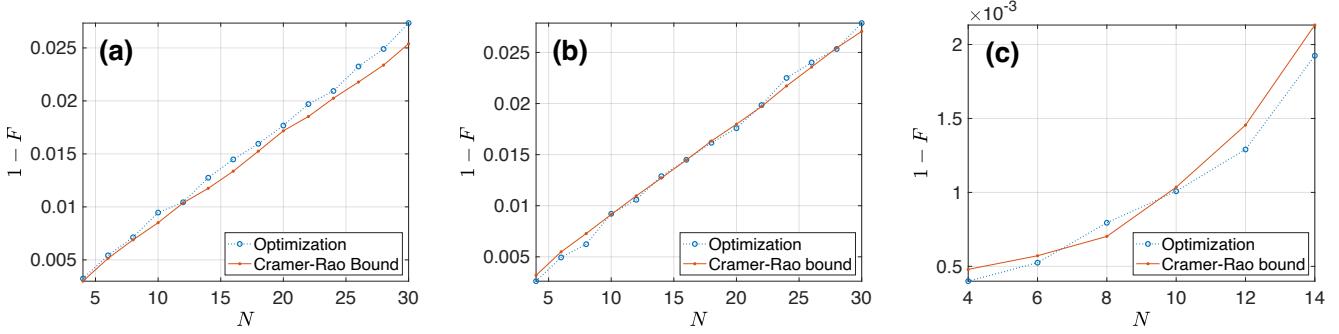


Figure 3. Results of the MLE algorithm and their comparison with the Cramer-Rao bound for (a) a 1D random complex MPS state with a bond dimension $\chi = 2$ and a sample size $M = 5 \times 10^3$, (a) a 1D cluster state defined by Eq. (18) with a complex MPS model and a sample size $M = 5 \times 10^3$, and (c) a generalized GHZ state with $\gamma = \pi/3$, defined by Eq. (22), with a complex TI MPS model and a sample size $M = 10^4$. The model MPS has a bond dimension $\chi = 2$ for all target states, and the specified sample size is used for both the MLE algorithm and the prediction of QST infidelity based on the Cramer-Rao bound.

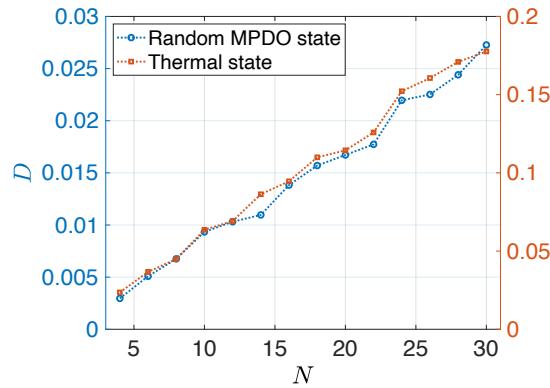


Figure 4. The distance measure D obtained by the MLE algorithm as a function of the number of qubits N for a random MPDO state with $M = 10^4$ measurement samples for each system size and a thermal state of an 1D quantum Ising model (Eq. (25) with $B = 1$ and $T = 2$) with $M = 1.5 \times 10^4$ measurement samples. The random MPDO state has a bond dimension 2 and a Kraus dimension 6, and the model MPDO state shares the same dimensions. The thermal state is represented by an MPO with a maximum matrix dimension of 7, while the model MPDO used to reconstruct it has a Kraus dimension of 4 and a bond dimension of 3.

state copies needed to scale linearly with the number of independent parameters in the MPS/MPDO. This conclusion is supported by an efficient calculation of the Cramer-Rao bound and an MLE algorithm using synthetic measurement samples that saturates the Cramer-Rao bound, hence making the algorithm optimal. We also identify counter examples of efficient QST where the target states belong to a family of GHZ states that are long-range entangled. The number of state copies needed to reconstruct such GHZ states scales exponentially with the system size. Interestingly, one can avoid this exponential scaling by constraining the model and the target state to be both real. This is in contrast to QST protocols based on local reductions, where an ad hoc N -qubit string operator needs to be measured to reconstruct the GHZ state [9, 12].

Here we do not require the measurement of any non-local operator, since our measurement setting is informationally complete. We have also shown that our MLE algorithm works for both pure target states and mixed target states, making our QST protocol universal.

While these results collectively suggest that the QST protocol proposed in this work can be used towards scalable quantum state tomography for a wide range of physical states in one dimension, a number of open questions remain. First, what is the sufficient and necessary condition for an MPS/MPDO state to be efficiently recoverable under local measurement? We suspect that the answer is closely related to whether the target state is short-range or long-range entangled, as well as to whether we restrict to real wavefunctions. Second, it is not clear whether the MPS/MPDO model provides an unbiased estimation of the target state for a finite sample size. While the Cramer-Rao bound is expected to hold for any well-behaved estimator for a sufficiently large sample size [31, 32], how to quantify the corrections to the Cramer-Rao bound for a finite sample size remains unclear. Third, can we identify a Cramer-Rao-type bound for a given target state using any local POVM? As we show in the discussion of the generalized GHZ state, for a given target state there may exist an optimal choice of a local POVM to minimize the Cramer-Rao bound. Fourth, we find that the number of samples required to reconstruct a target state accurately is closely related to the number of samples needed to compute the Fisher information matrix I accurately. Could this connection be used to identify states that are not efficiently recoverable by examining the convergence of the Fisher information matrix under an increasing number of Monte Carlo samples? Finally, can the results obtained here be applied to target states with other types of efficient classical representation, such as a higher-dimensional tensor network or an artificial neural network? We believe the study of scalable QST protocols for structured quantum states, and especially the establishment of rigorous bounds on the accuracy of such protocols, is still far from exhausted.

This work is supported by the NSF Grants Nos. PHY-2112893, CCF-2106834 and CCF-2241298, as well as the W. M. Keck Foundation. We thank the HPC center at Colorado

School of Mines and the Rocky Mountain Advanced Computing Consortium for providing computational resources needed in carrying out this work.

Appendix A: Analytical calculation of the Cramer-Rao bound for the GHZ state

In this appendix, we will calculate the Cramer-Rao bound for the GHZ state analytically. For simplicity, we start with a TI MPS model with the following diagonal matrices:

$$C^0 = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \quad C^1 = \begin{pmatrix} a_3 & 0 \\ 0 & a_4 \end{pmatrix} \quad (\text{A1})$$

The use of diagonal matrices here is justified as the inclusion of off-diagonal matrix elements does not change the resulting Cramer-Rao bound at all. The above model state achieves the target state $|\psi_{\text{GHZ}}\rangle = |0\cdots 0\rangle + |1\cdots 1\rangle$ upon setting $a_{1,4} = 1$ and $a_{2,3} = 0$.

Since A^0 and A^1 commute, we can explicitly calculate the wavefunction of the model MPS state in the computational basis as

$$\psi_s = \text{Tr}(C^{s_1} \dots C^{s_N}) = a_1^{N_0} a_3^{N-N_0} + a_2^{N_0} a_4^{N-N_0} \quad (\text{A2})$$

where N_0 denotes the number of 0s in the bit string s . Let us first assume of $a_{1,2,3,4}$ are real parameters. We can then explicitly evaluate the derivates of ψ_s over $a_{1,2,3,4}$ as

$$\begin{aligned} \frac{\partial \psi_s}{\partial a_1} &= N \delta_{N_0, N}, & \frac{\partial \psi_s}{\partial a_2} &= \delta_{N_0, 1} \\ \frac{\partial \psi_s}{\partial a_3} &= \delta_{N_0, N-1}, & \frac{\partial \psi_s}{\partial a_4} &= N \delta_{N_0, 0} \end{aligned} \quad (\text{A3})$$

where from now on we implicitly assume that all derivatives are further evaluated using the values of $a_{1,2,3,4}$ in the target state. Define the fictitious states $|\Psi_\alpha\rangle = \sum_s \frac{\partial \psi_s}{\partial a_\alpha} |s\rangle$ ($\alpha = 1, 2, 3, 4$), we find that

$$\begin{aligned} |\Psi_1\rangle &= N|0\cdots 0\rangle \\ |\Psi_2\rangle &= |10\cdots 0\rangle + |01\cdots 0\rangle + |0\cdots 01\rangle \\ |\Psi_3\rangle &= |01\cdots 1\rangle + |10\cdots 1\rangle + |1\cdots 10\rangle \\ |\Psi_4\rangle &= N|1\cdots 1\rangle \end{aligned} \quad (\text{A4})$$

which are mutually orthogonal. Next, we notice that Eq. (13) can be rewritten as

$$\begin{aligned} K_{\alpha\beta} &= \sum_{s,s'} \frac{\partial(\psi_s \psi_{s'}^*)}{\partial \theta_\alpha} \frac{\partial(\psi_s^* \psi_{s'})}{\partial \theta_\beta} \\ &= 2 \text{Re}(\langle \Psi_\alpha | \Psi_\beta \rangle \langle \psi | \psi \rangle + \langle \psi | \Psi_\alpha \rangle \langle \psi | \Psi_\beta \rangle) \end{aligned} \quad (\text{A5})$$

It's then not hard to find that K is a 4×4 matrix with

$$\begin{aligned} K_{11} &= K_{44} = 2(2N^2 + N^2) = 6N^2 \\ K_{22} &= K_{33} = 4N \\ K_{14} &= K_{41} = 2N^2 \end{aligned} \quad (\text{A6})$$

and all other matrix elements of K being zero.

Now we can further assume $a_{1,2,3,4}$ are complex parameters. Note that for any analytic function $f(z)$ with $z = x + iy$, $\frac{df}{dx} = \frac{df}{dz}$ and $\frac{df}{dy} = i \frac{df}{dz}$. As a result, $\frac{\partial \psi_s}{\partial \text{Re } a_\alpha}$ should be equal to Eq. (A3) while $\frac{\partial \psi_s}{\partial \text{Im } a_\alpha}$ should be equal to i times Eq. (A3). If α and β correspond to a parameter's real part and a parameter's imaginary part respectively, $K_{\alpha\beta} = 0$ as we are taking the real part of a purely imaginary number in Eq. (A5). If α and β both correspond to the imaginary parts, the expression of $K_{\alpha\beta}$ is modified to $K_{\alpha\beta} = 2 \text{Re}(\langle \Psi_\alpha | \Psi_\beta \rangle \langle \psi | \psi \rangle - \langle \psi | \Psi_\alpha \rangle \langle \psi | \Psi_\beta \rangle)$. Treating K now as a 8×8 matrices, we have

$$\begin{aligned} K_{55} &= K_{88} = 2(2N^2 - N^2) = 2N^2 \\ K_{66} &= K_{77} = 4N \\ K_{58} &= K_{85} = -2N^2 \end{aligned} \quad (\text{A7})$$

Noting that $|\psi_{\text{GHZ}}\rangle$ is not properly normalized, we find the normalized K matrix to be:

$$K = \begin{pmatrix} \frac{3}{2}N^2 & 0 & 0 & \frac{1}{2}N^2 & 0 & 0 & 0 & 0 \\ 0 & N & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & N & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2}N^2 & 0 & 0 & \frac{3}{2}N^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}N^2 & 0 & 0 & -\frac{1}{2}N^2 \\ 0 & 0 & 0 & 0 & 0 & N & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & N & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2}N^2 & 0 & 0 & \frac{1}{2}N^2 \end{pmatrix} \quad (\text{A8})$$

Next, we will try to calculate the Fisher information matrix for $|\psi_{\text{GHZ}}\rangle$. We can first write $P(\mathbf{m})$ in Eq. (7) as $\Phi(\mathbf{m})\Phi(\mathbf{m})^*$ where

$$\Phi(\mathbf{m}) = \text{Tr}[(F_0)^{N_0} (F_1)^{N_1} (F_2)^{N_2} (F_3)^{N_3}] \quad (\text{A9})$$

$$\begin{aligned} F^0 &\equiv \frac{1}{\sqrt{2}} C^0 \\ F^1 &\equiv \frac{1}{\sqrt{6}} C^0 + \frac{1}{\sqrt{3}} C^1 \\ F^2 &\equiv \frac{1}{\sqrt{6}} C^0 + \frac{e^{i2\pi/3}}{\sqrt{3}} C^1 \\ F^3 &\equiv \frac{1}{\sqrt{6}} C^0 + \frac{e^{-i2\pi/3}}{\sqrt{3}} C^1 \end{aligned}$$

Here N_q ($q = 0, 1, 2, 3$) denotes the number of q s in the quart string $\mathbf{m} = (m_1, m_2, \dots, m_N)$. We can evaluate Φ and its derivatives over $a_{1,2,3,4}$ (assuming $a_{1,2,3,4}$ are real for now) in the target state ($a_1 = a_4 = 1$ and $a_2 = a_3 = 0$) analytically with the help of Mathematica. The Fisher information matrix can then be evaluated by rewriting Eq. (11) as

$$I_{\alpha\beta} = 2 \text{Re} \sum_{\mathbf{m}} \left(\frac{\partial \Phi^*}{\partial \theta_\alpha} \frac{\partial \Phi^*}{\partial \theta_\beta} \frac{\Phi}{\Phi^*} + \frac{\partial \Phi^*}{\partial \theta_\alpha} \frac{\partial \Phi}{\partial \theta_\beta} \right) \quad (\text{A10})$$

Importantly, the sum over \mathbf{m} can be rewritten into the sums over $N_{0,1,2,3}$ that appeared in $\Phi(\mathbf{m})$ as

$$\sum_{N_0=0}^N \sum_{N_1=0}^{N-N_0} \sum_{N_2=0}^{N-N_0-N_1} \binom{N}{N_0} \binom{N-N_0}{N_1} \binom{N-N_0-N_1}{N_2} \quad (\text{A11})$$

where we shall set $N_3 = N - N_0 - N_1 - N_2$ in the summand of Eq. (A10). This allows us to perform the summation over \mathbf{m} efficiently.

Now assuming a_α ($\alpha = 1, 2, 3, 4$) is complex, similar to how we deal with the K matrix, here we find that $\frac{\partial \Phi}{\partial \text{Re } a_\alpha}$ is the same as $\frac{\partial \Phi}{\partial a_\alpha}$ with a_α assumed real, while $\frac{\partial \Phi}{\partial \text{Im } a_\alpha} = i \frac{\partial \Phi}{\partial \text{Re } a_\alpha}$. With the help of these analytical derivations, we can compute the Fisher information matrix numerically. After a proper normalization for the target state $|\psi_{\text{GHZ}}\rangle$, we obtain

$$I = \begin{pmatrix} \gamma N^2 & 0 & 0 & \delta N^2 & 0 & 0 & 0 & 0 \\ 0 & N & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & N & 0 & 0 & 0 & 0 & 0 \\ \delta N^2 & 0 & 0 & \gamma N^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta N^2 & 0 & 0 & -\delta N^2 \\ 0 & 0 & 0 & 0 & 0 & N & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & N & 0 \\ 0 & 0 & 0 & 0 & -\delta N^2 & 0 & 0 & \delta N^2 \end{pmatrix} \quad (\text{A12})$$

where $\gamma \approx 2$ and $\delta \approx 1/2^N$ are asymptotically exact in the large N limit. If the parameters in the MPS model are assumed real, we can just restrict the calculated K and I matrices [Eq. (A8) and Eq. (A12)] to the first four rows and columns, which correspond to the real part of the $a_{1,2,3,4}$.

Next, we generalize the above calculations to a non-TI MPS model, where

$$C_i^0 = \begin{pmatrix} a_{i,0} & 0 \\ 0 & b_{i,0} \end{pmatrix} \quad C_i^1 = \begin{pmatrix} a_{i,1} & 0 \\ 0 & b_{i,1} \end{pmatrix} \quad (\text{A13})$$

which reduces to the target state upon setting $a_{i,0} = b_{i,1} = 1$

and $b_{i,0} = a_{i,1} = 0$ for all i . It's easy to see that

$$\psi_s = \text{Tr}(A_1^{s_1} \dots A_N^{s_N}) = \prod_{i=1}^N a_{i,s_i} + \prod_{i=1}^N b_{i,s_i} \quad (\text{A14})$$

Following similar derivations as the TI case above, we can obtain the normalized K matrix in a block form as

$$K = \begin{pmatrix} \frac{3}{2}\mathbf{1} & \mathbf{0} & \mathbf{0} & \frac{1}{2}\mathbf{1} \\ \mathbf{0} & \mathbb{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{I} & \mathbf{0} \\ \frac{1}{2}\mathbf{1} & \mathbf{0} & \mathbf{0} & \frac{3}{2}\mathbf{1} \end{pmatrix} \quad (\text{A15})$$

where $\mathbf{1}$ is an $N \times N$ matrix of all 1s, $\mathbf{0}$ an $N \times N$ matrix of all 0s, and \mathbb{I} an $N \times N$ identity matrix. Similarly, we can obtain the Fisher information matrix in a block form as:

$$I = \begin{pmatrix} \gamma\mathbf{1} & \mathbf{0} & \mathbf{0} & \delta\mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} & \delta\mathbb{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \delta\mathbb{I} & \mathbb{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \delta\mathbf{1} & \mathbf{0} & \mathbf{0} & \gamma\mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \delta\mathbf{1} & \mathbf{0} & \mathbf{0} & -\delta\mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbb{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\delta\mathbf{1} & \mathbf{0} & \mathbf{0} & \delta\mathbf{1} \end{pmatrix} \quad (\text{A16})$$

Here γ and δ take exactly the same N -dependent values as the TI case. If the parameters in the model are assumed real, we can again simply restrict the K and I matrices obtained to the first four rows and columns in the above block forms.

With these analytical derivations, one can now efficiently calculate $\text{Tr}(KI^{-1})$ for both the TI and the non-TI MPS model and with either real or complex parameters. A similar but more involved calculation can be done for the generalized GHZ state in Eq. (22). The results of these calculations are shown in Fig. 2 of the main text.

-
- [1] J. Preskill, *Quantum* **2**, 79 (2018).
[2] L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, and C. Monroe, *Nature* **598**, 500 (2021).
[3] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, *Nature* **618**, 500 (2023).
[4] S. Aaronson, *SIAM Journal on Computing* **49**, STOC18 (2020).
[5] H.-Y. Huang, R. Kueng, and J. Preskill, *Nature Physics* **16** (2020).
[6] C. Harney, S. Pirandola, A. Ferraro, and M. Paternostro, *New Journal of Physics* **22**, 045001 (2020).
[7] Y. Zhu, Y.-D. Wu, G. Bai, D.-S. Wang, Y. Wang, and G. Chiribella, *Nature Communications* **13** (2022).
[8] K. Bu, D. E. Koh, R. J. Garcia, and A. Jaffe, *npj Quantum Information* **10** (2024).
[9] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, *Nature communications* **1**, 1 (2010).
[10] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, *New Journal of Physics* **14**, 095022 (2012).
[11] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, *Phys. Rev. Lett.* **111**, 020401 (2013).
[12] T. Baumgratz, A. Nüßeler, M. Cramer, and M. B. Plenio, *New Journal of Physics* **15**, 125004 (2013).
[13] D. S. França, F. G. L. Brandão, and R. Kueng, in *16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 197 (2021) pp. 7:1–7:13.
[14] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, *IEEE Transactions on Information Theory* **63**, 5628 (2017).
[15] M. Guta, J. Kahn, R. Kueng, and J. A. Tropp, *Journal of Physics A: Mathematical and Theoretical* **53**, 204001 (2020).
[16] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, *Quantum Inf. Comput.* **7**, 401 (2007).
[17] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, *Phys. Rev. Lett.* **93**, 207204 (2004).
[18] P. Silvi, F. Tschirsich, M. Gerster, J. Jünemann, D. Jaschke,

- M. Rizzi, and S. Montangero, *SciPost Phys. Lect. Notes* , **8** (2019).
- [19] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [20] T. Schmale, M. Reh, and M. Gärttner, *npj Quantum Information* **8** (2022).
- [21] N. Innan, O. I. Siddiqui, S. Arora, T. Ghosh, Y. P. Koçak, D. Paragis, A. A. O. Galib, M. A.-Z. Khan, and M. Bennai, *Quantum Machine Intelligence* **6** (2024).
- [22] K. Noh, L. Jiang, and B. Fefferman, *Quantum* **4**, 318 (2020).
- [23] H.-Y. Huang, R. Kueng, and J. Preskill, *Nat. Phys.* **16**, 1050 (2020).
- [24] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, *Nat Rev Phys* **5**, 9 (2022), arXiv:2203.11374 [quant-ph].
- [25] X. Chen, Z.-C. Gu, and X.-G. Wen, *Phys. Rev. B* **82**, 155138 (2010).
- [26] S. Bravyi, M. B. Hastings, and F. Verstraete, *Phys. Rev. Lett.* **97**, 050401 (2006).
- [27] M. C. Tran, C.-F. Chen, A. Ehrenberg, A. Y. Guo, A. Deshpande, Y. Hong, Z.-X. Gong, A. V. Gorshkov, and A. Lucas, *Phys. Rev. X* **10**, 031009 (2020).
- [28] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, *Phys. Rev. Lett.* **111**, 020401 (2013).
- [29] Z. Qin, C. Jameson, Z. Gong, M. B. Wakin, and Z. Zhu, *IEEE Transactions on Information Theory* **70**, 5030 (2024).
- [30] R. Pintelon, J. Schoukens, T. McKelvey, and Y. Rolain, *IEEE Transactions on Automatic Control* **41**, 719 (1996).
- [31] R. D. Gill and S. Massar, *Phys. Rev. A* **61**, 042312 (2000).
- [32] K. Usami, Y. Nambu, Y. Tsuda, K. Matsumoto, and K. Nakamura, *Phys. Rev. A* **68**, 022314 (2003).
- [33] R. Kosut, I. A. Walmsley, and H. Rabitz, arXiv:quant-ph/0411093.
- [34] J. Wang, Z.-Y. Han, S.-B. Wang, Z. Li, L.-Z. Mu, H. Fan, and L. Wang, *Phys. Rev. A* **101**, 032321 (2020).
- [35] P.-X. Chen, J. A. Bergou, S.-Y. Zhu, and G.-C. Guo, *Phys. Rev. A* **76**, 060303 (2007).
- [36] S. L. Braunstein, *Journal of Physics A: Mathematical and General* **25**, 3813 (1992).
- [37] T. Sugiyama, P. S. Turner, and M. Murao, *Phys. Rev. A* **83**, 012105 (2011).
- [38] M. Christandl and R. Renner, *Phys. Rev. Lett.* **109**, 120403 (2012).
- [39] B. Qi, Z. Hou, L. Li, D. Dong, G. Xiang, and G. Guo, *Scientific Reports* **3** (2013).
- [40] T. L. Scholten and R. Blume-Kohout, *New Journal of Physics* **20**, 023050 (2018).
- [41] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, *Journal of Mathematical Physics* **45**, 2171 (2004).
- [42] A. J. Scott, *Journal of Physics A: Mathematical and General* **39**, 13507 (2006).
- [43] W. Słomczyński and A. Szymusiak, arXiv:1402.0375.
- [44] M. Yoshida and G. Kimura, *Phys. Rev. A* **106**, 022408 (2022).
- [45] R. Stricker, M. Meth, L. Postler, C. Edmunds, C. Ferrie, R. Blatt, P. Schindler, T. Monz, R. Kueng, and M. Ringbauer, *PRX Quantum* **3**, 040310 (2022).
- [46] G. García-Pérez, M. A. Rossi, B. Sokolov, F. Tacchino, P. K. Barkoutsos, G. Mazzola, I. Tavernelli, and S. Maniscalco, *PRX Quantum* **2**, 040342 (2021).
- [47] K. Noh, L. Jiang, and B. Fefferman, *Quantum* **4**, 318 (2020).
- [48] A. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, *Phys. Rev. Lett.* **116**, 237201 (2016).
- [49] O. Sharir, Y. Levine, N. Wies, G. Carleo, and A. Shashua, *Phys. Rev. Lett.* **124**, 020503 (2020).
- [50] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, *Quantum Info. Comput.* **7**, 401–430 (2007).
- [51] J. Gorman and A. Hero, *IEEE Transactions on Information Theory* **36**, 1285 (1990).
- [52] D. Jaschke, S. Montangero, and L. D. Carr, *Quantum Sci. Technol.* **4**, 013001 (2018).