Optimizing Shadow Tomography with Generalized Measurements

H. Chau Nguyen,* Jan Lennart Bönsel,† Jonathan Steinberg,‡ and Otfried Gühne§
Naturwissenschaftlich–Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, 57068 Siegen, Germany

(Received 20 June 2022; revised 27 October 2022; accepted 31 October 2022; published 23 November 2022)

Advances in quantum technology require scalable techniques to efficiently extract information from a quantum system. Traditional tomography is limited to a handful of qubits, and shadow tomography has been suggested as a scalable replacement for larger systems. Shadow tomography is conventionally analyzed based on outcomes of ideal projective measurements on the system upon application of randomized unitaries. Here, we suggest that shadow tomography can be much more straightforwardly formulated for generalized measurements, or positive operator valued measures. Based on the idea of the least-square estimator shadow tomography with generalized measurements is both more general and simpler than the traditional formulation with randomization of unitaries. In particular, this formulation allows us to analyze theoretical aspects of shadow tomography in detail. For example, we provide a detailed study of the implication of symmetries in shadow tomography. Moreover, with this generalization we also demonstrate how the optimization of measurements for shadow tomography tailored toward a particular set of observables can be carried out.

DOI: 10.1103/PhysRevLett.129.220502

Introduction.—Quantum technology is based on our ability to manipulate quantum mechanical states of well-isolated systems: to encode, to process, and to extract information from the states of the system. Extracting information in this context means to design and perform measurements on the system so that observables or other properties of the system such as its entropy can be inferred. Naively, one may attempt to perform tomography of the state of the system. This amounts to making a sufficiently large number of different measurements on the system so that the density operator describing the state can be inferred [1–5].

However, when considering how the density operator is used later on, the entire information contained in the density operator is often not needed [6]. In fact, it is impractical to even write down the density operator when the number of qubits is large since the dimension of the many-qubit system increases exponentially. In practice, most often one is not interested in the elements of the density operator itself, but rather in certain properties of the quantum state, such as the mean values of certain observables or its entropy. Aiming at inferring directly the observables, bypassing the reconstruction of the density operator, shadow tomography has been theoretically proposed [6]. Huang *et al.* [7] thereafter suggested a practical procedure to realize this aim, which has attracted a lot of attention in the contemporary research of quantum information processing.

The idea of the protocol is simple. Traditionally, quantum state tomography is thought to be only useful once one has accurate enough statistics of measurements. However, state estimators such as the least-square estimator can actually be carried out in principle for arbitrary diluted data [8], a fact well established in data science and machine learning [9,10]. Indeed, a single data point can contribute a noisy estimate of the state, and the final estimated state is obtained by averaging over all the data points. Expectedly, when the data are diluted, the estimated quantum state can be highly noisy and far away from the targeted actual state in the high dimensional state space. This noisy estimation is, however, sufficient to predict certain observables or properties of the quantum states accurately [6,7]. Crucially, estimation of observables and certain properties of the quantum states for single data points can also be processed without explicitly writing down the density operator [7]. This endows the technique with the promise of scalability.

As for collecting data, Huang *et al.* [7] suggested to perform random unitaries from a certain chosen set of unitaries on the system and perform a standard ideal measurement afterward. This is equivalent to choosing randomly a measurement from a chosen set. Since then, various applications of the technique have been found in energy estimation [11,12], entanglement detection [13,14], metrology [15], analyzing scrambled data [16], and quantum chaos [17], to name a few. Further developments to improve the performance of the scheme [13,18–22] and generalization to channel shadow tomography have also been proposed [23,24]. In this Letter, we propose a general framework for shadow tomography with so-called generalized measurements [or positive operator valued measures (POVMs)]. This theoretical framework contains the randomization of unitaries as a special case, and at the same time allows for analysis of unavoidable noise in realistic quantum measurements [25,26], where projective

measurements may not be available. To our knowledge, there is so far a single proposed procedure for shadow tomography with generalized measurements [27]. The suggested procedure is, however, based on application of the original construction of classical shadows in Ref. [7] upon manually synthesizing the postmeasurement states for generalized measurements. On the contrary, here we show that classical shadows for generalized measurement can be derived straightforwardly from the least-square estimator [8], which requires no further assumptions on the post-measurement states and contains the framework for ideal measurements as a special case (see Appendixes A–D in Ref. [28]). In fact, our proposed framework turns out to be much more general and at the same time simpler than randomization of unitaries.

Shadow tomography with generalized measurements.—Consider a quantum system of dimension D, which can be either a single qubit or many qubits. A generalized measurement E on the system (POVM) is a collection of positive operators called effects, $E = \{E_1, E_2, \ldots, E_N\}$, summing up to the identity $\sum_{k=1}^{N} E_k = \mathbb{1}$. Each generalized measurement E defines a map Φ_E which maps a density operator ρ to a probability distribution over measurement outcomes:

$$\Phi_E(\rho) = \{\mathrm{Tr}(\rho E_k)\}_{k=1}^{N}. \tag{1}$$

When the measurement is performed, an outcome k is obtained according to this distribution.

Typical measurements in standard quantum mechanics are generalized measurements whose effects E_k are rank-1 projections, referred to as *ideal measurements*. For example, the measurement of a Pauli operator σ_x is an ideal measurement, whose effects are projections on the spin states in the x direction, $\{|x^+\rangle\langle x^+|, |x^-\rangle\langle x^-|\}$. On the other hand, randomizing three Pauli measurements σ_x, σ_y, σ_z is equivalent to a generalized measurement with effects proportional to the projections on the spin states in the x, y, and z direction, $1/3\{|x^\pm\rangle\langle x^\pm|, |y^\pm\rangle\langle y^\pm|, |z^\pm\rangle\langle z^\pm|\}$; see Appendix D in Ref. [28] for further discussion. Since these effects form an octahedron on the Bloch sphere, we also refer to it as the octahedron measurement. Generalized measurements, however, allow for much more flexible ways of extracting information from the system. For example, one can consider the generalized measurements defined by different polytopes as in Fig. 1. Generalized measurements also become indispensable in modeling realistic noise in measurement implementation.

Repeating the measurement M times on the system results in a string of outcomes $\{k_i\}_{i=1}^{M}$. Shadow tomography starts with associating a single outcome k to a distribution $\vec{q}_k = \{\delta_{kl}\}_{l=1}^{N}$ in $\mathbb{R}^N$. Such a single data point can be used to obtain a noisy estimate of ρ, called *classical shadow* [6,7], $\hat{\rho}_k = \chi(\vec{q}_k)$. As a general strategy in data science, one can require the shadow estimator χ to be the least-square estimator, of which the solution is well known [8,10] (see also Appendix A in Ref. [28]):

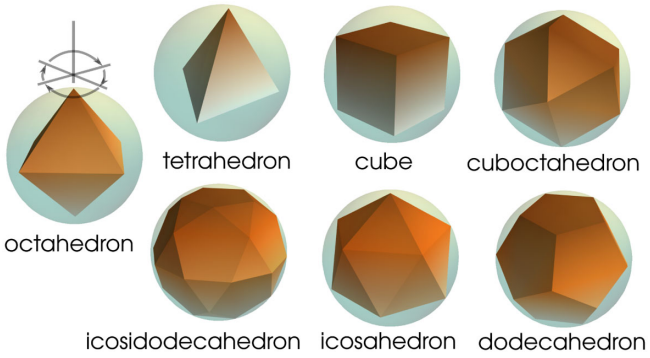

FIG. 1. Example of generalized measurements defined by polytopes on the Bloch sphere: octahedron ($N = 6$), tetrahedron ($N = 4$), cube ($N = 8$), cuboctahedron ($N = 12$), icosahedron ($N = 12$), dodecahedron ($N = 20$), icosidodecahedron ($N = 30$).

$$\chi_{\mathrm{LS}} = (\Phi_E^\dagger \Phi_E)^{-1} \Phi_E^\dagger. \tag{2}$$

Here we assume that the effects $\{E_k\}_{k=1}^{N}$ span the whole operator space, in which case E is said to be *informationally complete*, so that $\Phi_E^\dagger \Phi_E$ is invertible. Note that $\Phi^\dagger(\vec{q}_k) = E_k$, and if we define $C_E = \Phi_E^\dagger \Phi_E$, then $C_E(\rho) = \sum_{k=1}^{N} \mathrm{Tr}(\rho E_k) E_k$. The classical shadow can therefore also be written as

$$\hat{\rho}_k = C_E^{-1}(E_k). \tag{3}$$

The classical shadow in Eq. (3) resembles that in Ref. [7]. However, when the measurement is not ideal, the effects E_k do not represent the state of the system after the measurement. In fact, unlike the procedure proposed in Ref. [27], our derivation suggests that the states of the system after the measurement is here not important (see Appendixes A–C in Ref. [28]).

Estimation of observables and the shadow norm.—Each of the classical shadow (3) serves as an intermediate processed data point for further computation of observables. Given an observable X, each of the classical shadows $\hat{\rho}_k$ gives an estimate for the mean value $\langle X \rangle$ as $\hat{x}_k = \mathrm{Tr}(\hat{\rho}_k X)$. With the whole dataset $\{k_i\}_{i=1}^{M}$, the average $1/M \sum_{i=1}^{M} \hat{\rho}_{k_i}$ converges to ρ (Appendix A in Ref. [28]); therefore, $1/M \sum_{i=1}^{M} \hat{x}_{k_i}$ converges to $\langle X \rangle$. In this way, the mean value $\langle X \rangle$ can be estimated. For further refinement using the median-of-means estimation and estimation of polynomial functions of the density operator, see Ref. [7]. As also noted there, the asymptotic rate of convergence of the estimation is related to the variance of the estimator. For an observable X, the variance of the estimator can be computed as $\mathrm{var}(\hat{x}_k) = \sum_{k=1}^{N} \mathrm{Tr}(\hat{\rho}_k X)^2 \mathrm{Tr}(\rho E_k) - \langle X \rangle^2$. Ignoring the second term results in an upper bound for the variance, and finally assuming the worst case scenario,

i.e., maximization over ρ, one arrives at the definition of the shadow norm of X [7],

$$\mathrm{var}(\hat{x}_k) \le \|X\|_E^2 = \lambda_{\max}\left\{\sum_{k=1}^{N} \mathrm{Tr}(\hat{\rho}_k X)^2 E_k\right\}, \quad (4)$$

where $\lambda_{\max}\{\cdot\}$ denotes the maximal eigenvalue of the corresponding operator. The estimation procedure applies not only to an observable, but equally well to a set of observables $\mathcal{X}$. Assuming that the observables by certain normalization all have the same physical unit, the quality of shadow tomography with a generalized measurement E can be characterized by the maximal shadow norm:

$$\kappa_E^2(\mathcal{X}) = \max\{\|X\|_E^2 : X \in \mathcal{X}\}. \quad (5)$$

In the following, we would simply use κ_E^2 if the set of observables $\mathcal{X}$ is clear. Being the upper bound of the variance of the estimator, the smaller κ_E^2, the better the estimator accuracy [36].

Symmetry of generalized measurements and the computation of the classical shadows.—It has been observed that for certain classes of measurements, the inverse channel C_E^{-1} is particularly simple [7,37]. We are to show that behind this simplicity is the symmetry of the generalized measurement [38–40].

To give the simple intuition, we discuss the example of the octahedron generalized measurement over a qubit plotted in Fig. 1, leaving the general argument for high dimensional cases in Appendix E of Ref. [28]. Picking a vertex of the octahedron which corresponds to the effect E_k in Fig. 1, we consider the symmetry rotations of the octahedron that leave this vertex invariant. These are rotations by multiples of $\pi/2$ around the axis going through the chosen vertex. Noticeably, there is a single projection (and its complement) that is invariant under these rotations, which corresponds to the state of the spin pointing to the vertex itself. In other words, the effect E_k is uniquely specified by the symmetry. One can show that the corresponding classical shadow $\hat{\rho}_k$ is also invariant under these rotations, which then implies that it is a linear combination of E_k and the identity operator $\mathbb{1}$. In fact, this is a general property of the so-called uniform and rigidly symmetric measurements defined also for systems of general dimension D [28,38], which include in particular the symmetric solids in Fig. 1. In all these cases, one has

$$\hat{\rho}_k = aE_k + b\mathbb{1}. \quad (6)$$

The coefficients a and b can be explicitly computed, $a = (D\beta - \alpha^2)/(D\gamma - \alpha^3)$ and $b = (\gamma - \alpha\beta)/(D\gamma - \alpha^3)$, where $\alpha = \mathrm{Tr}(E_k)$, $\beta = \mathrm{Tr}(E_k^2)$, and $\gamma = \sum_{l=1}^{N} \mathrm{Tr}(E_k E_l)^2$ (which are all independent of k).

Effects of noise in measurements.—Measurements in realistic experimental setups are not ideal. The imperfection

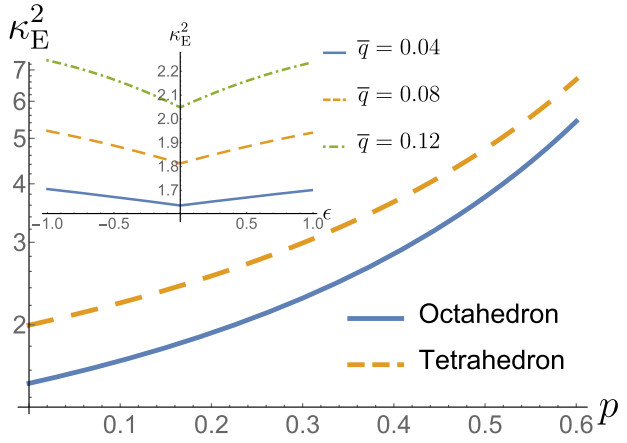

FIG. 2. Effects of depolarization noise (main) and simple readout error noise (inset) on the maximal shadow norm of 128 pure state projections distributed according to the Haar measure.

is due to various sources of noise in setting up the parameters of the measurement devices, or the resolution and the accuracy of readout signals [25,26]. As an example, suppose that the measurement E is not perfectly implemented, where the device fails to couple to the system with probability p and indicates an outcome at complete random. This can be modeled by the effects that are depolarized as $\{p\mathbb{1}/N + (1-p)E_k\}_{k=1}^N$. Another example is the readout error, which is particularly important for superconducting qubits [41,42]. In a simplified model, an outcome 0 in the computational basis is misread as 1 with probability q_+, and 1 misread as 0 also with probability q_-. The error rate averaged over the two bases is $\bar{q} = (q_+ + q_-)/2$ and the asymmetry between them is characterized by $\epsilon = (q_+ - q_-)/(q_+ + q_-)$. The measurement effects of the octahedron measurement implemented by randomization under this noise becomes $1/3\{(1-q_\pm)|t^\pm\rangle\langle t^\pm| + q_\mp|t^\mp\rangle\langle t^\mp|, t = x, y, z\}$. For further discussion, see Appendix F of Ref. [28].

Our formalism directly takes measurement error correction into account, once the noisy effects with an appropriate model are used instead of the ideal ones. To access the quality of the shadow tomography after error correction, we choose $|\mathcal{X}| = 128$ pure state projections distributed according to the Haar measure as observables. The dependence of the maximal shadow norm κ_E^2 on the noise parameters for the tetrahedron and the octahedron measurements is shown in Fig. 2. It is interesting to see that in either case the maximal shadow norm κ_E^2 depends only weakly on small error rate, showing the robustness of shadow tomography against noise.

Optimization of generalized measurement for shadow tomography.—Given a set of observables $\mathcal{X}$, one would like to find the generalized measurement E so that the maximal shadow norm is minimized:

$$E^* = \arg\min_E \kappa_E^2(\mathcal{X}). \quad (7)$$

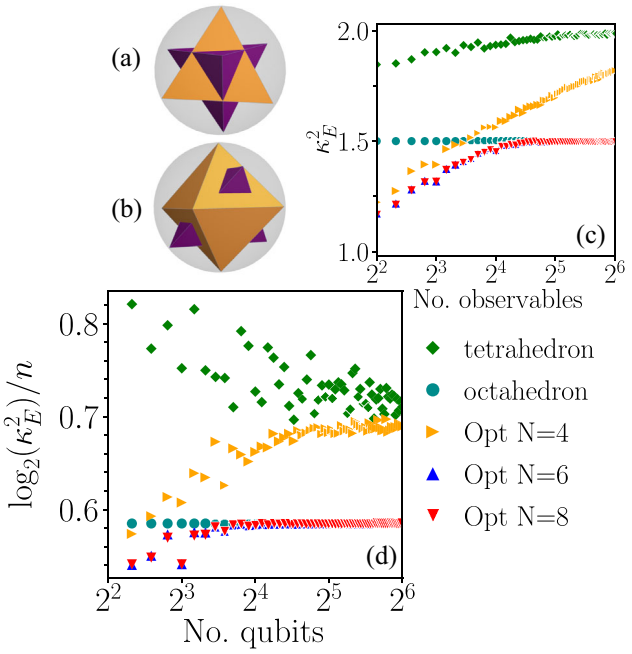

FIG. 3. Targeted observables and optimal generalized measurements. (a) For observables corresponding to four projections defined by the orange tetrahedron, the measurement corresponding to the inverted tetrahedron measurement (violet) is optimal. (b) For observables corresponding to eigenprojections of the Pauli observables σ_x, σ_y, and σ_z (orange octahedron), the violet tetrahedron measurement is optimal. (c) Optimal shadow norms given by the optimizer (labeled Opt with the number of measurement outcomes) as a function of the number of single-qubit projection observables randomly distributed according to the Haar measure. (d) Similarly, optimal shadow norms given by the optimizer as a function of the number of qubits. The observables are tensor products of single-qubit projections distributed according to the Haar measure. In panels (c) and (d), the shadow norm for the tetrahedron and octahedron measurements are also shown.

Restricted to randomized unitaries, this optimization is impractical to carry out (see Appendix B of Ref. [28]). Extending to all generalized measurements, this is simply an optimization over a convex domain. We implemented simulated annealing for the minimization and found the obtained optima to be highly reliable (see Appendix G of Ref. [28]). Below we start with discussing the case of a single qubit, which, despite being simple, is also the basis to understand the case of many qubits.

Example 1.—Consider a single qubit. As for the observables $\mathcal{X}$, we consider the following possibilities. (a) Take observables to be 4 projections corresponding to the orange tetrahedron in Fig. 3(a). The squared shadow norm κ_E^2 is 2 for the tetrahedron generalized measurement defined exactly by these 4 projections, and $3/2$ for the octahedron measurement. The optimizer suggests that the tetrahedron measurement plotted in violet in Fig. 3(a), obtained by centrally inverting the orange tetrahedron, is optimal with $\kappa_E^2 = 1$. (b) As observables consider the projections onto the eigenstates of the Pauli observables, see the orange octahedron in Fig. 3(b). The octahedron measurement itself gives $\kappa_E^2 = 3/2$. Interestingly, the optimizer shows that $\kappa_E^2 = 3/2$ can also be obtained with the tetrahedron generalized measurement of 4 outcomes indicated in violet in Fig. 3(b). (c) Lastly, as observables we consider random projections distributed according to the Haar measure on the Bloch sphere. Figure 3(c) presents the shadow norms obtained by the optimizer with respect to the number of observables. For small number of observables ($|\mathcal{X}| \lesssim 15$), the optimizer always finds measurements with a given number of outcomes significantly better than the standard tetrahedron ($N = 4$) or the octahedron measurements ($N = 6$). It is interesting to see that if the number of outcomes is fixed to be 6 or 8, the κ_E^2 converges to the octahedron measurement with the value of $3/2$.

The last example (c) hints that the octahedron measurement is somewhat special. Indeed, it turns out that the squared shadow norm with respect to the octahedron measurement for any projection is identically $3/2$. Using this fact, we show that if the targeted observables are *all* the projections on arbitrary pure states of the qubit, the optimal measurement would be the octahedron measurement assuming equal trace of the effects (see Appendix H of Ref. [28]).

Tensoring the shadow construction for many-body systems.—Shadow tomography is especially designed for the cases where the system size is large. Consider the case where the system consists of n qubits, corresponding to the total dimension of $D = 2^n$. In this case, shadow tomography can be performed by making (identical or not identical) generalized measurements $\{E^{(1)}, E^{(2)}, \ldots, E^{(n)}\}$ on each of the qubits, each described by a collection of N_i effects, $E^{(i)} = \{E_k^{(i)}\}_{k=1}^{N_i}$. Theoretically, this corresponds to a measurement of a generalized measurement E^{tot} on the whole system with each effect labeled by a string of outcomes $\boldsymbol{k} = \{k^{(1)}, k^{(2)}, \ldots, k^{(n)}\}$, $E_{\boldsymbol{k}}^{\mathrm{tot}} = E_{k^{(1)}}^{(1)} \otimes E_{k^{(2)}}^{(2)} \otimes \cdots \otimes E_{k^{(n)}}^{(n)}$. The whole general analysis above can be applied. In fact, such a string $\boldsymbol{k}$ of outcomes corresponds simply to the classical shadow,

$$\hat{\rho}_{\boldsymbol{k}}^{\mathrm{tot}} = \hat{\rho}_{k^{(1)}}^{(1)} \otimes \hat{\rho}_{k^{(2)}}^{(2)} \otimes \cdots \otimes \hat{\rho}_{k^{(n)}}^{(n)}, \tag{8}$$

where $\hat{\rho}_{k^{(i)}}^{(i)}$ is the classical shadow corresponding to the measurement $E^{(i)}$ on the ith qubit. Crucially, the typical observables of the system can be easily estimated without (impractically) explicitly computing the classical shadows in the form of a $D \times D$ matrix [7]. Indeed, an observable X on the system is often of the form $X = X^{(1)} \otimes X^{(2)} \otimes \cdots \otimes X^{(n)}$. Then, a single string of outcomes $\boldsymbol{k}$ gives rise to a single estimate of $\langle X \rangle$ as $\mathrm{Tr}[\hat{\rho}_{k^{(1)}}^{(1)} X^{(1)}] \mathrm{Tr}[\hat{\rho}_{k^{(2)}}^{(2)} X^{(2)}] \cdots \mathrm{Tr}[\hat{\rho}_{k^{(n)}}^{(n)} X^{(n)}]$. The final estimate

of $\langle X \rangle$ is as usual obtained by averaging over all data points. Observe that it is not necessary to construct the large density operator of the whole system. Moreover, the shadow norm of such a factorized observable also factorizes $\|X\|_E = \|X^{(1)}\|_{E^{(1)}} \|X^{(2)}\|_{E^{(2)}} \cdots \|X^{(n)}\|_{E^{(n)}}$.

Optimizing generalized measurements for many-body systems.—For many qubits, the number of parameters to be optimized in Eq. (7) increases exponentially. To simplify, one can assume that for a many-qubit system, the generalized measurement is factorized as a tensor product over the qubits as we discussed above. Moreover, if there is no preference among the qubits, one can also assume that $E^{(1)} = E^{(2)} = \cdots = E^{(n)}$. The complexity of the computation under these assumptions is only linear in the number of qubits and the number of observables.

Example 2.—We consider a system of up to $n = 64$ qubits. We choose $|X| = n$ observables which are products of different component observables on single qubits. The component observables on single qubits are randomly distributed according to the Haar measure. As the qubits are equivalent, one might anticipate that the optimal factorizing measurement for the qubits is similar to those that are optimized separately for each qubit. Our simulations confirm this expectation. In Fig. 3(d), for small number of qubits ($n \lesssim 10$), the optimizer with $N = 6$ and $N = 8$ gives significantly lower shadow norms for the choice of tetrahedron or octahedron measurements. On the other hand, observe that as the number of qubits increases, the obtained optimal shadow norm converges to that given by the octahedron measurement, pointing to the speciality of the octahedron measurement on qubit-based platforms as we discussed in Example 1(c).

Conclusion.—Being both more general and simpler, the formulation of shadow tomography with generalized measurements sheds light on various aspects of shadow tomography. This also opens a range of interesting questions for future research. Further analysis of realistic noise in the existing and future experiments [43,44] of shadow tomography could be considered. Extension of this framework to channel tomography is of direct interest. It would also be important to see whether the technique of derandomization [18] can also be incorporated. The optimality of the octahedron measurement for shadow tomography for a qubit-based system suggests a connection between geometry and shadow tomography. Investigation of this connection and extension for higher dimensional systems would be an interesting direction. Also the construction of optimal measurements for nonlinear functions of the density operator, or shadow tomography of a specific set of density operators, is in demand for further applications of shadow tomography.

The authors would like to thank Satoya Imai, Matthias Kleinmann, Martin Kliesch (with indirect comments), Michał Oszmaniec, Salwa Shaglel, Zhen-Peng Xu, and Benjamin Yadin for inspiring discussions and comments. The University of Siegen is kindly acknowledged for enabling our computations through the OMNI cluster. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, Projects No. 447948357 and No. 440958198), the Sino-German Center for Research Promotion (Project M-0294), the ERC (Consolidator Grant No. 683107/TempoQ), and the German Ministry of Education and Research (Project QuKuK, BMBF Grant No. 16KIS1618K). J. L. B. and J. S. acknowledge support from the House of Young Talents of the University of Siegen.

Note added.—Recently, we learned that related results to estimate noncommuting observables from a special generalized measurement have been derived in Ref. [45]. Additionally, a similar protocol for shadow tomography using the special example of a symmetric and informationally complete (SIC) POVM has been suggested and experimentally implemented in Ref. [44]. Both works do not, however, develop a framework for shadow tomography with arbitrary generalized measurements.

*chau.nguyen@uni-siegen.de
†jan.boensel@uni-siegen.de
‡steinberg@physik.uni-siegen.de
§otfried.guehne@uni-siegen.de

[1] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Measurement of the Wigner Distribution and the Density Matrix of a Light Mode using Optical Homodyne Tomography: Application to Squeezed States and the Vacuum, Phys. Rev. Lett. **70**, 1244 (1993).
[2] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Measurement of qubits, Phys. Rev. A **64**, 052312 (2001).
[3] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek al kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, Scalable multiparticle entanglement of trapped ions, Nature (London) **438**, 643 (2005).
[4] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne, Systematic Errors in Current Quantum State Tomography Tools, Phys. Rev. Lett. **114**, 080403 (2015).
[5] M. Paris and J. Řeháček, *Quantum State Estimation* (Springer, Berlin, 2004).
[6] S. Aaronson, Shadow tomography of quantum states, SIAM J. Comput. **49**, STOC18-368 (2020).
[7] H. Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).
[8] M Guţă, J. Kahn, R. Kueng, and J. A. Tropp, Fast state tomography with optimal error bounds, J. Phys. A **53**, 204001 (2020).
[9] P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias,

low-variance introduction to machine learning for physicists, Phys. Rep. **810**, 1 (2019).
[10] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006).
[11] C. Hadfield, Adaptive Pauli shadows for energy estimation, arXiv:2105.12207.
[12] C. Hadfield, S. Bravyi, R. Raymond, and A. Mezzacapo, Measurements of quantum Hamiltonians with locally-biased classical shadows, arXiv:2006.15788.
[13] A. Elben, R. Kueng, H. Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch, Mixed-State Entanglement from Local Randomized Measurements, Phys. Rev. Lett. **125**, 200501 (2020).
[14] A. Neven, J. Carrasco, V. Vitale, C. Kokail, A. Elben, M. Dalmonte, P. Calabrese, P. Zoller, B. Vermersch, R. Kueng, and B. Kraus, Symmetry-resolved entanglement detection using partial transpose moments, npj Quantum Inf. **7**, 152 (2021).
[15] A. Rath, C. Branciard, A. Minguzzi, and B. Vermersch, Quantum Fisher Information from Randomized Measurements, Phys. Rev. Lett. **127**, 260501 (2021).
[16] R. J. Garcia, Y. Zhou, and A. Jaffe, Quantum scrambling with classical shadows, Phys. Rev. Res. **3**, 033155 (2021).
[17] L. K. Joshi, A. Elben, A. Vikram, B. Vermersch, V. Galitski, and P. Zoller, Probing Many-Body Quantum Chaos with Quantum Simulators, Phys. Rev. X **12**, 011018 (2022).
[18] H. Y. Huang, R. Kueng, and J. Preskill, Efficient Estimation of Pauli Observables by Derandomization, Phys. Rev. Lett. **127**, 030503 (2021).
[19] T. Zhang, J. Sun, X. X. Fang, X. M. Zhang, X. Yuan, and H. Lu, Experimental Quantum State Measurement with Classical Shadows, Phys. Rev. Lett. **127**, 200501 (2021).
[20] S. Chen, W. Yu, P. Zeng, and S. T. Flammia, Robust shadow estimation, PRX Quantum **2**, 030348 (2021).
[21] H. Y. Hu and Y. Z. You, Hamiltonian-driven shadow tomography of quantum states, Phys. Rev. Res. **4**, 013054 (2022).
[22] H. Y. Hu, S. Choi, and Y. Z. You, Classical shadow tomography with locally scrambled quantum dynamics, arXiv:2107.04817.
[23] R. Levy, D. Luo, and B. K. Clark, Classical shadows for quantum process tomography on near-term quantum computers, arXiv:2110.02965.
[24] J. Helsen, M. Ioannous, I. Roth, J. Kitzinger, E. Onorati, A. H. Werner, and J. Eisert, Estimating gate-set properties from random sequences, arXiv:2110.13178.
[25] F. Arute, K. Arya, R. Babbush *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).
[26] Y. Chen, M. Farahzad, S. Yoo, and T.-C. Wei, Detector tomography on IBM quantum computers and mitigation of an imperfect measurement, Phys. Rev. A **100**, 052315 (2019).
[27] A. Acharya, S. Saha, and A. M. Sengupta, Shadow tomography based on informationally complete positive operator-valued measure, Phys. Rev. A **104**, 052418 (2021).
[28] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.220502 for the derivation of the classical shadow from the least-square estimator, the discussion on the relationship between randomizing unitaries with generalized measurements, the symmetry analysis of classical shadows, and the proof of the optimality of the octahedron measurement for the construction of random observables, which contains Refs. [29–35].
[29] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement* (Cambridge University Press, Cambridge, England, 2012).
[30] G. M. D'Ariano, P. Lo Presti, and P. Perinotti, Classical randomness in quantum measurements, J. Phys. A **38**, 5979 (2005).
[31] B. Nachman, M. Urbanek, W. A. de Jong, and C. W. Bauer, Unfolding quantum computer readout noise, npj Quantum Inf. **6**, 84 (2020).
[32] K. Temme, S. Bravyi, and J. M. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, Phys. Rev. Lett. **119**, 180509 (2017).
[33] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes* (Cambridge University Press, 2007).
[34] V. Granville, M. Krivanek, and J.-P. Rasson, Simulated annealing: A proof of convergence, IEEE Trans. Pattern Anal. Mach. Intell. **16**, 652 (1994).
[35] W. Fulton and J. Harris, *Representation Theory* (Springer, New York, 2004).
[36] In practice, the targeted state could be very different from the worst case scenario assumed in obtaining the shadow norm. Therefore, it might also be informative to consider the average of the variance with respect to certain ensemble of states.
[37] K. Bu, D. E. Koh, R. J. Garcia, and A. Jaffe, Classical shadows with Pauli-invariant unitary ensembles, arXiv:2202.03272.
[38] H. C. Nguyen, S. Designolle, M. Barakat, and O. Gühne, Symmetries between measurements in quantum mechanics, arXiv:2003.12553.
[39] H. Zhu and B. G. Englert, Quantum state tomography with fully symmetric measurements and product measurements, Phys. Rev. A **84**, 022327 (2011).
[40] Y. I. Bogdanov, G. Brida, I. D. Bukeev, M. Genovese, K. S. Kravtsov, S. P. Kulik, E. V. Moreva, A. A. Soloviev, and A. P. Shurupov, Statistical estimation of the quality of quantum-tomography protocols, Phys. Rev. A **84**, 042108 (2011).
[41] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, Mitigating measurement errors in multiqubit experiments, Phys. Rev. A **103**, 042605 (2021).
[42] R. Hicks, B. Kobrin, C. W. Bauer, and B. Nachman, Active readout-error mitigation, Phys. Rev. A **105**, 012419 (2022).
[43] L. E. Fischer, D. Miller, F. Tacchino, P. Kl. Barkoutsos, D. J. Egger, and I. Tavernelli, Ancilla-free implementation of generalized measurements for qubits embedded in a qudit space, Phys. Rev. Res. **4**, 033027 (2022).
[44] R. Stricker, M. Meth, L. Postler, C. Edmunds, C. Ferrie, R. Blatt, P. Schindler, T. Monz, R. Kueng, and M. Ringbauer, Experimental single-setting quantum state tomography, PRX Quantum **3**, 040310 (2022).
[45] D. McNulty, F. B. Maciejewski, and M. Oszmaniec, Estimating quantum Hamiltonians via joint measurements of noisy non-commuting observables, arXiv:2206.08912.