Probing scrambling and operator size distributions using random mixed states and local measurements

Philip Daniel Blocher,[1,*] Karthik Chinni,[2,1,†] Sivaprasad Omanakuttan,[1,‡] and Pablo M. Poggi[3,1,§]

[1]*Center for Quantum Information and Control, Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131, USA*

[2]*Department of Engineering Physics, École Polytechnique de Montréal, Montréal, Quebec H3T 1J4, Canada*

[3]*Department of Physics, SUPA and University of Strathclyde, Glasgow G4 0NG, United Kingdom*

(Received 26 June 2023; accepted 29 January 2024; published 21 March 2024)

The dynamical spreading of quantum information through a many-body system, typically called scrambling, is a complex process that has proven to be essential to describe many properties of out-of-equilibrium quantum systems. Scrambling can, in principle, be fully characterized via the use of out-of-time-ordered correlation functions, which are notoriously hard to access experimentally. In this work, we put forward an alternative toolbox of measurement protocols to experimentally probe scrambling by accessing properties of the operator size probability distribution, which tracks the size of the support of observables in a many-body system over time. Our measurement protocols require the preparation of separable mixed states together with local operations and measurements, and combine the tools of randomized operations, a modern development of near-term quantum algorithms, with the use of mixed states, a standard tool in NMR experiments. We demonstrate how to efficiently probe the probability-generating function of the operator distribution and discuss the challenges associated with obtaining the moments of the operator distribution. We further show that manipulating the initial state of the protocol allows us to directly obtain the individual elements of the distribution for small system sizes.

DOI: 10.1103/PhysRevResearch.6.013309

I. INTRODUCTION

Understanding how quantum information spreads across the degrees of freedom of a quantum system is a key part of developing a comprehensive picture of nonequilibrium quantum many-body physics. In this context, the notion of scrambling has attracted much attention over the past years due to its relevance in the study of closed-system thermalization [1], quantum chaos [2], information retrieval in black holes [3,4], and quantum algorithms [5]. Scrambling refers to the dynamical delocalization of quantum information [6] and can be diagnosed by the generation of entangled states from initially separable ones, or. from the growth of initially local operators [7].

An approach to characterizing scrambling from the unitary evolution of an operator $W(t)$ in a many-body system is to analyze the dynamics of so-called operator size distributions $\{P_k(t)\}$ [8–10]. These can be obtained by coarse-graining the expansion of $W(t)$ in a complete operator basis, the choice of which depends on the nature of the system. In the case of systems of N spin-$\frac{1}{2}$ particles, a natural operator basis is the set of multibody Pauli operators $\mathcal{P}_N = \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}^{\otimes N}/\sqrt{2^N}$, which has dimension $D = 4^N$ and forms an orthonormal basis. In the Heisenberg picture an operator W may at time t be written as

$$W(t) = \mathcal{U}^\dagger(t)\, W\, \mathcal{U}(t) = \sum_{j=0}^{D-1} f[\Lambda_j; W(t)]\Lambda_j, \qquad (1)$$

where $\mathcal{U}(t)$ is the unitary time evolution operator from the initial time $t = 0$ to time t and $\Lambda_j \in \mathcal{P}_N$. Given this expansion of $W(t)$, the operator size distribution is constructed by grouping the elements of the exponentially large Pauli basis according to their size $s(\Lambda)$. Here the operator size corresponds to the number of nonidentity operators in the Pauli string (i.e., its Hamming weight) and thus $1 \leqslant s(\Lambda) \leqslant N$ such that $\mathcal{P}_N = \cup_{k=1}^N C_k$, where $C_k = \{\Lambda \mid s(\Lambda) = k\}$. The resulting operator size distribution reads

$$P_k(t) = \frac{1}{\mathrm{Tr}[W^\dagger W]} \sum_{\Lambda \in C_k} |f[\Lambda; W(t)]|^2, \qquad (2)$$

and measures the size of the support of $W(t)$. It is easy to see that $\sum_{k=1}^N P_k(t) = 1$ for all times t, and hence the operator size distribution can be regarded as a coarse-grained probability distribution in the expansion coefficients of $W(t)$.

Of particular interest is the case where $W(0)$ is a size-one (i.e., single-body) Pauli operator such that $P_k(0) = \delta_{k,1}$. As the operator grows and information becomes scrambled, the

*blocher@unm.edu
†karthik.chinni@polymtl.ca
‡somanakuttan@unm.edu
§pablo.poggi@strath.ac.uk

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

distribution shifts to higher values of k and grows in variance. The dynamics of $P_k(t)$ have been studied for various many-body models [8,10], and it holds a close connection to the Krylov picture of operator growth [7,11].

While the scrambling dynamics in the system may be described using the operator size distribution, it is not obvious how to access this distribution experimentally. A complete characterization of $P_k(t)$ for all k requires an exponential amount of resources; this is clearly seen if one reconstructs $P_k(t)$ from two-point correlation functions to obtain each of the coefficients $f[\Lambda; W(t)] = \mathrm{Tr}[\Lambda W(t)]$ in Eq. (1). A partial workaround is given by considering out-of-time-ordered correlators (OTOCs) which take the form

$$\mathcal{F}(W(t), R) = \langle W^\dagger(t) R^\dagger(0) W(t) R(0) \rangle \,. \tag{3}$$

It has been shown that moments of $P_k(t)$ can be computed by constructing averages of OTOCs over appropriate sets of operators $\{R_i\}$ [8,10,12,13]. However, accessing even a single OTOC is often challenging in experiments due to the out-of-time-ordered nature of Eq. (3) and typically requires the use of many-body time-reversal operations [14–18] or auxiliary systems [19]. In cases where OTOCs can be accessed without these tools (see, for instance, the method in Ref. [20]), reconstructing the moments of $P_k(t)$ requires the measurement of a large number of OTOCs and the task becomes unfeasible for high-order moments [10].

In this article we propose an alternative set of tools to experimentally probe the operator size distribution which circumvents the use of OTOCs completely. We combine the use of ensembles of initially separable mixed states with local random operations and local measurements at the final time in order to access a quantity $G(t)$ which is explicitly *not* an OTOC, and hence we name this quantity a "NOTOC." We show that our measurement protocol, depending on the choice of initial state ensemble, probes either the probability generating function (PGF) $F(x,t) = \sum_k P_k(t) x^k$ of the operator size distribution (method A) or its elements $\{P_k(t)\}$ directly (method B), as illustrated in Fig. 1. We demonstrate our methods numerically and show that the operator size distribution can be accurately probed even when accounting for statistical noise stemming from averaging over random operations and from experimental imperfections. In our analysis of method A, we discuss inherent shortcomings of the problem of inverting the PGF to obtain $P_k(t)$ such as its sensitivity to noise, and we show for the case of the Ising model that the PGF can itself be seen as a good indicator of the presence of scrambling in the system. For method B we discuss the efficiency of the method as the system size N increases and show that individual probabilities $P_k(t)$ can be reliably obtained as long as $k \ll N$.

Our proposed measurement protocol connects to previous works which focused on experimental schemes to diagnose scrambling. In particular, method A recovers a procedure put forward in Ref. [12] to measure operator growth in the special case when the initial states are pure. Likewise, the use of randomized operations to access properties of the operator size distribution makes this proposal complementary to the one in Ref. [20], where similar tools were used to measure OTOCs instead. Finally, the NOTOC proposed here can be seen as a generalization of the fidelity OTOC [1,14,15], which

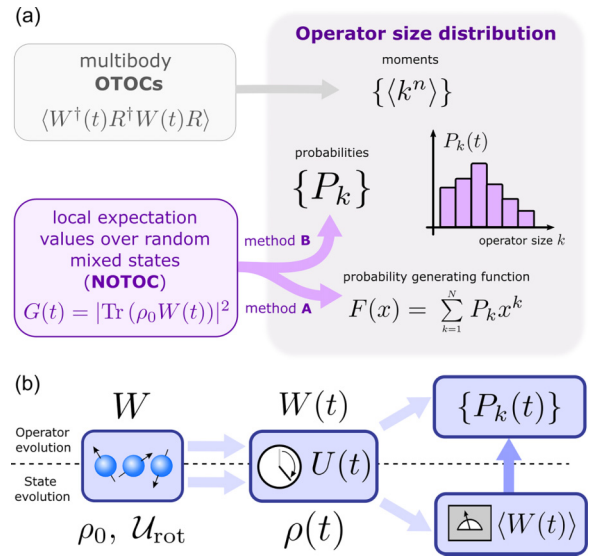

FIG. 1. (a) Relationship between the operator size distribution $P_k(t)$ of Eq. (2), out-of-time-ordered correlation functions (OTOCs), and the NOTOC approach proposed in this article using local expectation values over random mixed states. OTOCs allow (indirect) access of the moments of the operator distribution, while our proposal allows to probe the probability generating function $F(x)$ or the individual probabilities directly, depending on the choice of initial state ρ_0. (b) Illustration of the operator evolution (above the dashed line) and the proposed measurement protocol (below the dashed line). The operator size distribution $P_k(t)$ is a property of an operator in the Heisenberg picture (and thus independent of a choice of state). Our protocol proposes to access this property by measuring expectation values $\langle W(t) \rangle = \mathrm{Tr}[\rho(t) W]$ after the time evolution of a suitably chosen set of initial states ρ_0.

has been widely studied due to its ease of accessibility in experiments [21]. Our analysis reveals that this quantity is inherently connected to the operator size distribution, although in a way that is fundamentally different from regular OTOCs and that had not been revealed up to now to the best of our knowledge.

The structure of this article is as follows: In Sec. II we present the main tools to be used to probe the operator size distribution $P_k(t)$, including the choice of initial states, randomized operations and local measurements. Sections III and IV discuss two methods of implementing the necessary state preparation for our measurement protocol, with the method of Sec. III using partially polarized qubit states to obtain the probability-generating function for the operator size distribution, and the method of Sec. IV tailoring separable states to obtain the probability distribution elements directly. In Sec. V we compare our measurement protocol to previous results and proposals in the literature, extend the NOTOC measurement protocol to collective spin systems, and provide a general discussion of the NOTOC toolbox. Finally, we conclude on our work in Sec. VI and discuss possible future extensions.

II. NOTOC MEASUREMENT PROTOCOL

Our measurement protocol draws inspiration from the so-called fidelity OTOCs, which are OTOCs of the form Eq. (3)

where one chooses the operator $R = \rho_0$ in Eq. (3) to be the projector onto the pure initial state ρ_0 [1,14,15]. Fidelity OTOCs are experimentally accessible quantities as [21]

$$\langle W^\dagger(t)R^\dagger(0)W(t)R(0)\rangle \equiv |\langle W(t)\rangle|^2, \tag{4}$$

and hence they require only the measurement of a single-time expectation value. However, this identity holds only for pure initial states, and the fidelity OTOCs for mixed initial states takes on a more elaborate form, as discussed in Ref. [21].

Inspired by the simple form of the fidelity OTOC in Eq. (4) as a squared expectation value of an operator W for a pure state, we define the quantity of interest for our protocol as

$$G(t) := |\langle W(t)\rangle|^2 = |\mathrm{Tr}[\rho_0 W(t)]|^2, \tag{5}$$

with ρ_0 being a generic mixed state. The quantity $G(t)$ defined in Eq. (5) is equivalent to an OTOC only for pure states [22], and to emphasize that $G(t)$ is related to fidelity OTOCs while itself being a time-ordered correlation function at all times and for all initial states, we will refer to $G(t)$ as a "NOTOC" throughout this article. In the remainder of this section we will demonstrate the utility of defining $G(t)$ of the form Eq. (5), and in Sec. V we will comment further on its relation to fidelity OTOCs.

Using Eq. (1) the NOTOC $G(t)$ may be written as

$$G(t) = \sum_i |f[\Lambda_i; W(t)]|^2 \langle\Lambda_i\rangle^2 + \sum_{i\neq j} f[\Lambda_i; W(t)] f[\Lambda_j; W(t)]^* \langle\Lambda_i\rangle\langle\Lambda_j\rangle^*. \tag{6}$$

We now consider a generic mixed state of the form

$$\rho_0 = \frac{1}{d}\mathbb{1} + \sum_i r_i \Lambda_i, \tag{7}$$

for which $\langle\Lambda_i\rangle = r_i$. We recall from Sec. I that $C_k = \{\Lambda \mid s(\Lambda) = k\}$ is the set of Pauli operators of size k. If we were to engineer an ensemble of initial states $\{\rho_0\}$ of the form Eq. (7) such that the state coefficients $\{r_i\}$ were independent, identically distributed random variables with vanishing mean $\overline{r_i} = 0$, the second term of Eq. (6) would vanish under averaging over this ensemble. Furthermore, if we require the state coefficients of the engineered initial state to have finite variance $\overline{r_i^2} = \Delta_k$ for $\Lambda_i \in C_k$, then the averaged quantity then reads

$$\overline{G(t)} = \sum_i |f[\Lambda_i; W(t)]|^2 \overline{r_i^2} + \sum_{i\neq j} f[\Lambda_i; W(t)] f[\Lambda_j; W(t)]^* \overline{r_i r_j^*} \tag{8}$$

$$= \sum_{k=1}^{N} \sum_{\Lambda_i \in C_k} |f[\Lambda_i; W(t)]|^2 \Delta_k \tag{9}$$

$$= \sum_{k=1}^{N} \Delta_k \mathrm{Tr}[W^\dagger W] P_k(t). \tag{10}$$

Equation (10) reveals that $\overline{G(t)}$ is a linear combination of the elements of the probability distribution $\{P_k(t)\}$, with coefficients proportional to the variance Δ_k times the 2-norm $\mathrm{Tr}[W^\dagger W] = \|W\|_2^2$ of the operator W. The probability distribution $\{P_k(t)\}$ may be extracted from Eq. (10) using several methods, and we present two different measurement protocols for systems of spin-$\frac{1}{2}$ particles in the following Secs. III and IV. Our measurement protocols provide experimental access to the averaged NOTOC $\overline{G(t)}$ and the probability distribution $\{P_k\}$ through engineering of the initial state ρ_0 and subsequent measurement of the expectation value $\langle W(t)\rangle = \mathrm{Tr}[\rho(t)W]$ at the final time t. The main challenge thus lies in the preparation of random initial states ρ_0 whose coefficients $\{r_i\}$ must have appropriate statistics.

The operator size distribution is an operator property independent of the initial state of the system, and hence the Heisenberg picture lends itself nicely to the analysis of the operator size distribution's evolution in time. However, in the NOTOC measurement protocol outlined above we probe the operator size distribution using expectation values $\langle W(t)\rangle_{\rho_0} \equiv \mathrm{Tr}[\rho_0 W(t)]$, which depend on the choice of initial state ρ_0. In this way, the initial state is a control knob used by this protocol to access properties of the operator. As expectation values are quantities independent of the choice of picture, it is more natural to describe our proposed experimental measurement protocol in the Schrödinger picture as the time evolution of an initial state $\rho_0 \to \rho(t)$, with which we first evaluate the expectation value $\langle W(t)\rangle_{\rho_0} \equiv \mathrm{Tr}[\rho(t)W]$, subsequently calculate the NOTOC Eq. (5), and finally recover the averaged NOTOC Eq. (10) by appropriate averaging over initial states. The NOTOC measurement protocol is thus illustrated in this way in Fig. 1(b), where the Heisenberg operator evolution is shown above the dashed line, and our proposed measurement protocol is illustrated below the dashed line.

III. METHOD A: ACCESSING THE PROBABILITY GENERATING FUNCTION OF THE PROBABILITY DISTRIBUTION

In this section we present an experimentally relevant measurement protocol for obtaining the squared expectation value Eq. (10) using mixed states similar to those used in NMR together with random local operations [23]. The starting point of our measurement protocol is the preparation of the product state

$$\rho_{\mathrm{ini}} = \left(\frac{\mathbb{1} + \varepsilon\sigma_z}{2}\right)^{\otimes N} \tag{11}$$

$$= \frac{1}{d}\left(\mathbb{1} + \varepsilon\sum_i \sigma_z^i + \varepsilon^2 \sum_{i<j} \sigma_z^i \sigma_z^j + \cdots\right) \tag{12}$$

$$= \frac{1}{d}\mathbb{1} + \sum_{k=1}^{N} \frac{\varepsilon^k}{\sqrt{d}} \sum_{\Lambda\in C_k^z} \Lambda, \tag{13}$$

where each qubit is in a statistical mixture of being maximally mixed $\mathbb{1}$ and polarized along the z axis, with the parameter ε controlling the amount of polarization. In Eq. (13) we have defined the subset $C_k^z \subset C_k$ of size-k Pauli operators that consist of only σ_z terms (e.g., $\mathbb{1}\otimes\sigma_z\otimes\mathbb{1}$). The polarization parameter ε takes values $|\varepsilon| \leqslant 1$, and we note that for $\varepsilon = \pm 1$ the initial state ρ_{ini} is a pure state.

To create a random state ρ_0 for the experiment whose expansion coefficients r_i in Eq. (7) satisfy the appropriate statistics for Eq. (10), we apply random local rotations to the initial state Eq. (13) via the unitary $\mathcal{U}_{\rm rot} = \bigotimes_{i=1}^N \mathcal{U}_{\rm rot}^{(i)}$,

$$\rho_0 = \mathcal{U}_{\rm rot}\,\rho_{\rm ini}\,\mathcal{U}_{\rm rot}^\dagger \tag{14}$$

$$= \frac{1}{d}\mathbb{1} + \sum_{k=1}^N \frac{\varepsilon^k}{\sqrt{d}}\left(\mathcal{U}_{\rm rot}\sum_{\Lambda\in C_k^z}\Lambda\,\mathcal{U}_{\rm rot}^\dagger\right) \tag{15}$$

$$= \frac{1}{d}\mathbb{1} + \sum_{k=1}^N \frac{\varepsilon^k}{\sqrt{d}}\sum_{Q\in C_k} q_Q\,Q, \tag{16}$$

where $\mathcal{U}_{\rm rot}^{(i)}$ is a random rotation of the ith qubit that we will discuss momentarily. In the last equality of Eq. (16) we expanded $\mathcal{U}_{\rm rot}\sum_{\Lambda\in C_k^z}\Lambda\,\mathcal{U}_{\rm rot}^\dagger$ on the Pauli operators $Q\in C_k$, as the random local rotations do not change the operator size. Comparing the form of Eq. (16) to that of Eq. (7), we make the identification $r_Q = q_Q\varepsilon^k/\sqrt{d}$ for $s(Q)=k$. We choose the random local rotations $\mathcal{U}_{\rm rot}$ such that the coefficients $\{q_Q\}$ are independent, identically distributed random variables with vanishing mean $\overline{q_Q}=0$ and finite variance $\overline{q_Q^2}=\Delta_k$ for $s(Q)=k$. This yields $\overline{r_Q}=0$ and $\overline{r_Q^2}=\varepsilon^{2k}\Delta_k/d$, which is consistent with the assumptions made in Sec. II. Substituting this back into Eq. (10)—and choosing W to be a nonidentity observable with trace $\mathrm{Tr}[W^2]=d$ [24]—we thus find averaged squared expectation value,

$$\overline{G(\varepsilon,t)} = \sum_{k=1}^N P_k(t)\,\Delta_k\,\varepsilon^{2k}. \tag{17}$$

Equation (17) is an (at-most) Nth degree polynomial in ε^2 with coefficients proportional to elements $P_k(t)$ of the probability distribution of interest. This form is reminiscent of a PGF of the probability distribution $\{P_k(t)\}$ [25], and we now show that Eq. (17) is indeed a PGF by introducing the explicit form of Δ_k.

To obtain the correct statistics for the coefficients r_Q, as well as to ensure that all operators $Q\in C_k$ are sampled for all k, we propose to take the single-qubit rotation operator $\mathcal{U}_{\rm rot}^{(i)}$ to be sampled from a uniform distribution over SU(2). In each random instance the local rotation transforms the ith site Pauli-Z as

$$\sigma_z^i \to \mathcal{U}_{\rm rot}^{(i)}\sigma_z^i\mathcal{U}_{\rm rot}^{(i)\dagger} = \sum_\alpha n_\alpha^{(i)}\sigma_\alpha^i, \tag{18}$$

with $\alpha = x, y, z$, with

$$\mathbf{n}^{(i)} = \cos(\phi_i)\sin(\theta_i)\hat{\mathbf{x}} + \sin(\phi_i)\sin(\theta_i)\hat{\mathbf{y}} + \cos(\theta_i)\hat{\mathbf{z}}. \tag{19}$$

The polar angle θ_i and azimuthal angle ϕ_i are thus random variables taking values in the interval $[0,\pi)$ and $[0,2\pi)$, respectively.

The coefficients q_Q of Eq. (16) are then expressible as products of these random numbers $n_\alpha^{(i)}$, with all factors being independent of each other thanks to the local rotations being uncorrelated. Taking each pair of angles (θ_i,ϕ_i) to be uniformly distributed over the sphere, one readily obtains that $\overline{n_\alpha}=0$ leading to $\overline{r_Q}=0$ as required by our protocol. Due to symmetry the variance Δ_k is expected to be independent of $\alpha = x, y, z$, and hence we can compute it for any component. We find that

$$\overline{n_z^2} = \int d\Omega P(\theta,\phi)\cos^2(\theta) = \frac{1}{3}, \tag{20}$$

and thus we get a factor of a $1/3$ for each nonidentity operator in a given multibody Pauli operator Q. This leads to the variance $\Delta_k = 1/3^k$ which in turn implies that

$$\overline{r_Q^2} = \frac{1}{d}\frac{\varepsilon^{2k}}{3^k}. \tag{21}$$

Using the result Eq. (21) and letting $x := \epsilon^2/3$ for notational convenience, Eq. (17) may be rewritten as

$$F(x,t) := \overline{G(\sqrt{3x},t)} = \sum_{k=1}^N P_k(t)\,x^k, \tag{22}$$

which is the PGF for the probability distribution $\{P_k(t)\}$ [25]. From the PGF one may extract information about the corresponding probability distribution, including the elements and moments of the probability distribution.

We point out that the uniform sampling of the continuous group SU(2) is not strictly necessary, as it suffices to sample over a finite set of rotations given the correct first and second moments. This is equivalent to constructing a unitary 2-design and sampling the local operations from it and can be done by choosing $\mathcal{U}_{\rm rot}$ such that each qubit takes on one of the values $\pm X$, $\pm Y$, and $\pm Z$ in each shot, with a total of 6^N unique rotation unitaries $\mathcal{U}_{\rm rot}$ needed to obtain Eq. (17) without approximation. This shortcut to uniform averaging also results in $\Delta_k = 1/3^k$ and $\overline{r_i}=0$, and thus Eq. (22) is unchanged when using this method of averaging over a discrete set of rotations.

A. Application of the measurement protocol

In the remainder of this section we demonstrate via numerical simulations our measurement protocol for the case of the one-dimensional (1D) tilted field Ising model. We present results on how to obtain the PGF approximately using a subset of random rotations and show that the PGF may be used as an indicator of quantum information scrambling. Finally, we discuss an experimentally relevant method of extracting the elements of the probability distribution from the PGF and discuss its sensitivity to noise.

The Hamiltonian for the 1D tilted field Ising model is given by

$$H_{\rm Ising}(\theta) = J\sum_{i=1}^{N-1}\sigma_z^i\sigma_z^{i+1} + B\sum_{i=1}^N\left(\sigma_x^i\cos\theta + \sigma_z^i\sin\theta\right), \tag{23}$$

where the operators σ_α^i are the usual Pauli operators on site i with $\alpha = x, y, z$. The model describes N spin-$\frac{1}{2}$ particles interacting in one dimension via nearest-neighbor interactions in the presence of an external magnetic field with both a transverse and longitudinal component which are parameterized by the angle $0\leqslant\theta\leqslant\pi/2$. This model is a well-known platform for studies of many-body quantum chaos [10,26] since it is nonintegrable for generic choices of θ and presents two integrable limits at $\theta=0$ (where the model reduces to the usual transverse field Ising model) and $\theta=\pi/2$ (where the Hamiltonian is diagonal in the z basis). The scrambling

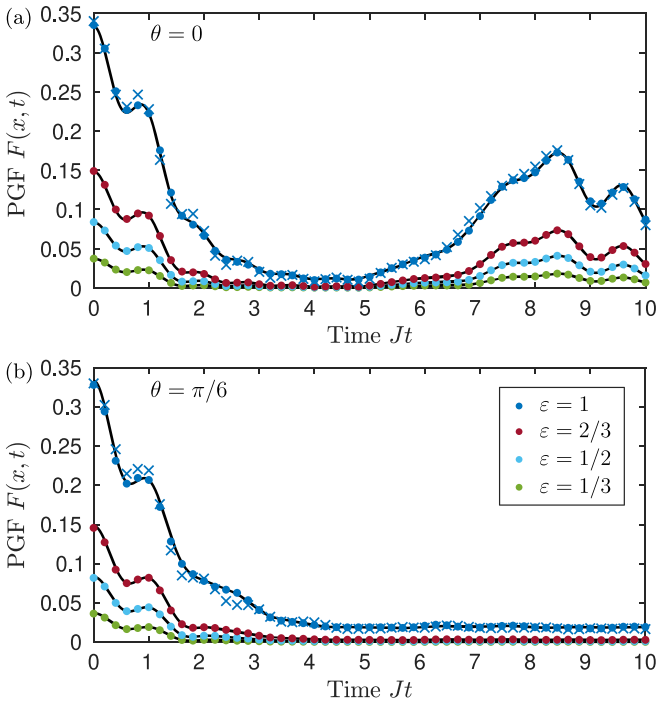

FIG. 2. Exact PGF approximated using a randomized subset of rotations for $W(0) = \sigma_y^{(1)}$. Solid black lines are the exact PGF, while colored dots denote the PGF obtained using a subset of 1000 random rotations. Results are shown for several choices of the state parameter ε [see the legend in (b)] and for both the integrable case $\theta = 0$ (a) and the highly chaotic case $\theta = \pi/6$ (b). For the case $\varepsilon = 1$ the blue crosses show the PGF obtained using a subset of 100 random rotations.

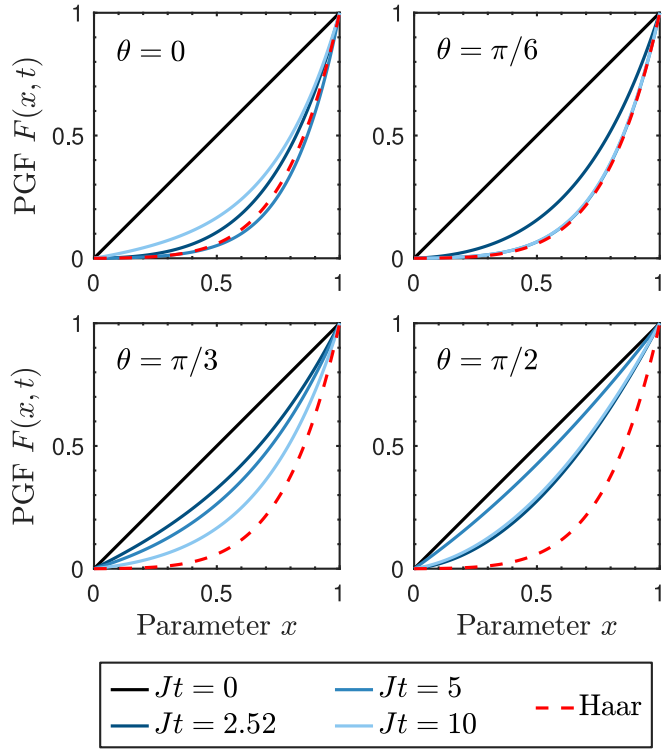

FIG. 3. Exact PGF $F(x,t)$ illustrated for several times t and B-field angles θ as a function of the parameter x. Solid lines are the PGF for $N = 6$ qubits and, for comparison, the red dashed line is the PGF for a Haar-random probability distribution. For $\theta = \pi/6$ the curves for $Jt = 5$ and $Jt = 10$ coincide.

properties of the dynamics generated by the Hamiltonian Eq. (23) have also been studied in relation to their quantum chaos characteristics. For instance, it has been established that even the integrable limit $\theta = 0$ can lead to scrambling, and in this case the mean operator size typically presents long-lived oscillations for finite system sizes. The system typically shows the highest degree of chaoticity for $\theta \simeq \pi/6$, where the operator sizes grow quickly and then equilibrate, and temporal fluctuations are suppressed [10]. Throughout the following we will consider the case $J = B$ with $N = 6$ qubits.

In Figs. 2(a) and 2(b) the solid lines behind the colored dots show the exact PGF $F(x,t)$ for the edge site operator $W(0) = \sigma_y^{(1)} \equiv \sigma_y \otimes \mathbb{1} \otimes \ldots \otimes \mathbb{1}$ for several choices of the state parameter ε. We consider the integrable case $\theta = 0$ (a) and the chaotic case $\theta = \pi/6$ (b) of the tilted field Ising model Eq. (23). For all t the two limiting cases $F(x = 1, t) = 1$ and $F(x = 0, t) = 0$, with $x = \varepsilon^2/3$, follow from the definition of the PGF Eq. (22). We observe in Fig. 2 that the primary effect of varying the state parameter ε in the considered parameter range is a change of amplitude of the PGF.

The colored dots (blue crosses) in the two panels of Fig. 2 show the PGF extracted using a random subset of 1000 (100) rotations from the $6^N = 46656$ rotations needed for the exact result. Figure 2 demonstrates that we may obtain the PGF to very good accuracy using a heavily reduced number of rotations compared to the exact result both in the integrable case $\theta = 0$ and the chaotic case $\theta = \pi/6$. This is encouraging for the experimental feasibility of implementing the present protocol. See Appendix C 1 for further details on the error induced by evaluating the PGF for a random subset of rotations.

In Ref. [10] moments of the operator size probability distribution were used as signatures of quantum information scrambling in the tilted field Ising model, with the integrable case $\theta = 0$ leading to oscillatory dynamics of the mean operator size, whereas in the chaotic case $\theta = \pi/6$ the mean operator size grew and saturated only after initial oscillations. When comparing the curves for $\theta = 0$ with the corresponding curves in $\theta = \pi/6$ in Fig. 2, we see a clear difference in behavior for $Jt \geqslant 4$, with $\theta = 0$ curves at later times exhibiting oscillations that are not present for $\theta = \pi/6$. In the following we thus explore whether the PGF may be used as an indicator of scrambling, similar to the mean operator size used in Ref. [10].

Figure 3 illustrates the PGF $F(x,t)$ for different B-field angles θ and times t as a function of the PGF parameter x. For the initial time $t = 0$ the PGF is linear in x independent of the choice of θ; however, at finite times t the behavior of the PGF is significantly different across different θs. In the chaotic case $\theta = \pi/6$ the PGF $F(x,t)$ approximately coincides with the result obtained from a Haar random state (dashed red line) for all displayed times greater than $Jt = 2.52$, indicating a fast equilibration. For $\theta = \pi/3$ the solid curves have yet to equilibrate at $Jt = 10$ (light blue curve). The two integrable cases $\theta = 0$ and $\theta = \pi/2$ display oscillatory behavior in the PGF $F(x,t)$ for a given x, similarly to what was observed in

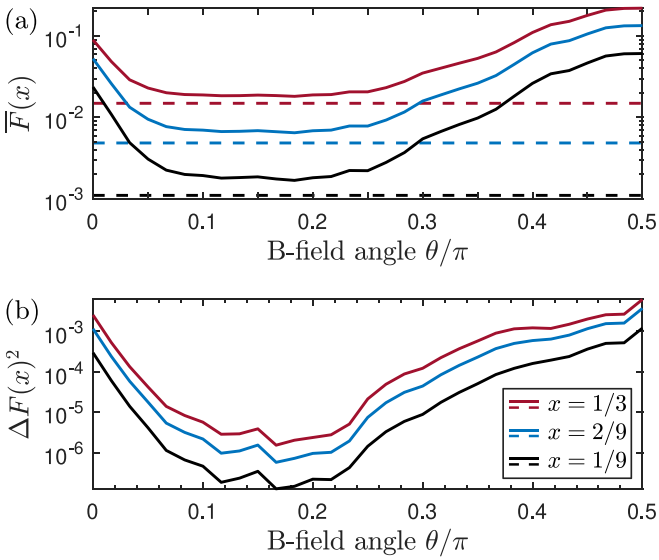

FIG. 4. (a) Time average $\overline{F}(x)$ and (b) fluctuations $\Delta F(x)^2$ of the PGF $F(x,t)$ as a function of the B-field angle θ, plotted for three values of x corresponding to $\varepsilon = 1$ (red), $\varepsilon = \sqrt{2/3}$ (blue), and $\varepsilon = \sqrt{1/3}$ (black curve). The time average and fluctuations are calculated for $5 \leqslant Jt \leqslant 80$ to exclude the initial transient dynamics from the results. In (a) we compare the time average $\overline{F}(x)$ (solid lines) with the PGF value for a Haar-random probability distribution (dashed lines).

Fig. 2, e.g., the dark blue $Jt = 2.52$ curve is below both the black $Jt = 0$ curve and the light blue $Jt = 10$ curve for all x in both cases. We also do not observe an equilibration of the PGF to the result for the Haar random state.

While the PGF $F(x,t)$ primarily serves as a quantity from which one may extract information about the corresponding probability distribution, we now illustrate how the PGF itself may be used to characterize quantum information scrambling. We propose to do this by analyzing the time average of the PGF for a fixed argument x,

$$\overline{F}(x) = \frac{1}{t_f - t_i}\int_{t_i}^{t_f} F(x,t)dt', \tag{24}$$

and the time-averaged temporal fluctuations,

$$\Delta F(x)^2 = \frac{1}{t_f - t_i}\int_{t_i}^{t_f} \left[F(x,t') - \overline{F(x)}\right]^2 dt'. \tag{25}$$

In Ref. [10] we studied analog constructions for the mean operator size (i.e., the first moment of the operator size distribution) and found that they allowed to distinguish different scrambling and quantum chaos regimes of this model. In Fig. 4(a) we show the time average $\overline{F}(x)$ (solid lines) as a function of the magnetic field angle θ. For different accessible values of x we see that that $\overline{F}(x)$ dips in the highly chaotic regime and grows near the integrable limits. This behavior originates in the fact that the chaotic case shows quick scrambling and subsequent equilibration to the Haar-random value (dashed lines), cf. Eq. (35), where the PGF is closer to 0, while the integrable cases show the largest typical values of the PGF, as was noted for $\theta = 0$ in the discussion of Fig. 2. A similar functional form is observed for the time-averaged temporal fluctuations $\Delta F(x)^2$ which we show in Fig. 4(b). This indicates that temporal fluctuations of the PGF are suppressed in the chaotic and enhanced in the integrable cases, a behavior also observed also for the mean operator size in Fig. 3 of Ref. [10].

Our present findings thus show that it is not necessary to extract the probability distribution $\{P_k(t)\}$ from the PGF $F(x,t)$ in order to characterize the scrambling of quantum information for the considered Ising model. Rather, the time average and time fluctuations of the PGF are themselves good indicators of the presence of scrambling. As demonstrated previously in Fig. 2 the PGF can be approximated to good accuracy using a heavily reduced number of random rotations, and as the PGF is not itself sensitive to the noise unlike quantities extracted from the PGF (which we will comment on in the following Sec. III B), the PGF provides an experimentally relevant quantity for characterizing quantum information scrambling. The results of Figs. 3 and 4 further emphasize the utility of the PGF itself as a quantifier of quantum information scrambling.

B. Extracting probability distribution elements and moments from the PGF

In the absence of noise on the PGF $F(x,t)$, e.g., in a numerical study of the exact PGF, the corresponding probability distribution $\{P_k(t)\}$ can be extracted from the PGF using Eq. (22) in the following way. One first calculates the PGF $F(x,t)$ for N different values of x and then inverts the system of equations Eq. (22) to obtain the elements $P_k(t)$ of the probability distribution. However, in the presence of a noisy PGF due to, e.g., a finite number of sampled rotations or experimental imperfections, we find that the procedure of inverting Eq. (22) leads to high sensitivity to noise, which we associate with a poorly conditioned linear system of equations.

For a more systematic approach we can use well-known properties of the PGF to recover the full probability distribution. From the definition of the PGF Eq. (22), one finds that [25]

$$P_k(t) = \frac{1}{k!}\frac{\partial}{\partial x}F(x=0,t), \tag{26}$$

i.e., the elements $\{P_k(t)\}$ of the probability distribution are accessible through the evaluation of derivatives of the PGF with respect to the parameter x at $x = 0$. Likewise the moments of the probability distribution may be accessed through derivatives of the PGF at $x = 1$:

$$\mathbb{E}[X^k(t)] = \left(x\frac{\partial}{\partial x}\right)^k F(x=1,t). \tag{27}$$

In the following we will focus on the extraction of the elements of the probability distribution via Eq. (26). We leave the discussion of the moment extraction for later in this section.

The nth derivative of the PGF $F(x,t)$ at $x = 0$ may be implemented for both numerical studies and in experiments using a forward-only finite-difference method of the form [27]

$$F^{(n)}(x=0,t) = (\Delta x)^{-n}\sum_{m=0}^{n+a-1} c_m^{(n,a)} F(m\,\Delta x, t) + \mathcal{O}(\Delta x^a), \tag{28}$$

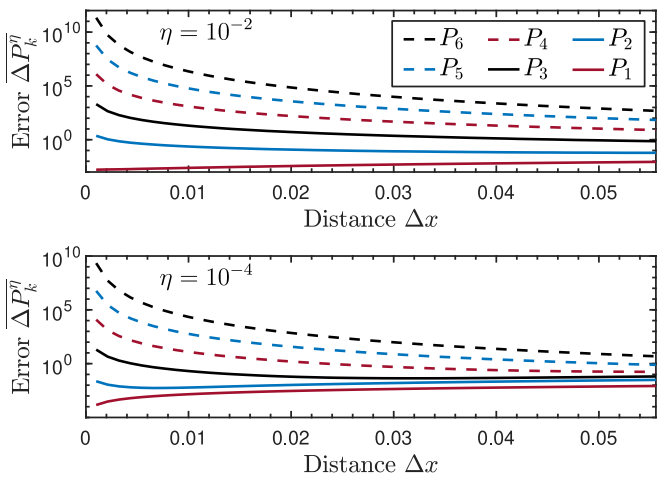

FIG. 5. Time-averaged error $\overline{\Delta P_k^\eta}$ [see Eq. (30)] for the elements of the probability distribution $\{P_k\}$, averaged over $0 \leqslant Jt \leqslant 10$. We have used multiplicative noise with strength $\eta = 10^{-2}$ (upper panel) and $\eta = 10^{-4}$ (lower panel), added to the PGF $F(x,t)$ before using finite-difference derivatives with step size Δx and accuracy $a = 2$ to extract the noisy probability distribution elements $\{P_k^\eta(t)\}$. In both panels the curves from top to bottom at $\Delta x = 0.02$ are the time-averaged errors $\overline{\Delta P_k^\eta}$ for P_6 (black dashed), P_5 (blue dashed), P_4 (red dashed), P_3 (black solid), P_2 (blue solid), and P_1 (red solid), averaged over 100 noise realizations.

where a is the accuracy of the finite-difference method. We leave the details of our implementation of the forward-only finite-difference method to Appendix A. In practice, applying this method implies preparing several states of the form Eq. (13) with various levels of purity (controlled by the parameter ε). To explore the experimental feasibility of extracting the probability distribution elements from the PGF, in the following we characterize the sensitivity of the method to noise. The noise may have origin in an approximative PGF due to choosing a reduced subset of rotations, as discussed in Sec. III A, or stem from experimental imperfections. We simulate the effect of the noise in the following way. First, the exact PGF $F(x,t)$ is calculated, on top of which we add noise. The noise is assumed to be Gaussian and multiplicative in nature [28], and hence we write the noisy PGF as

$$F^\eta(x,t) = [1 + \delta(\eta)]F(x,t), \tag{29}$$

where $\delta(\eta) \sim N(0,\eta)$ is a normal-distributed number with vanishing mean and variance η^2. From the noisy PGF we then use Eq. (28) to extract the elements $\{P_k^\eta(t)\}$ of the probability distribution in the presence of noise.

We now consider the $N = 6$ transverse field Ising model with $J = B$ and $\theta = 0$. In Fig. 5 we show for the operator $W(0) = \sigma_y^{(1)}$ the time-averaged error of the extracted probability distribution elements

$$\overline{\Delta P_k^\eta} = \frac{1}{T}\int_0^T dt \left|P_k^\eta(t) - P_k^0(t)\right|, \tag{30}$$

as a function of the finite-difference step size Δx. The error is calculated by comparing the elements $P_k^\eta(t)$ extracted from the noisy PGF with the exact elements $P_k^0(t)$ and has been averaged over 100 realizations of the Gaussian distributed noise $\delta(\eta)$. An accuracy of $a = 1$ was used for the finite-difference method; see details in Appendix A.

When using finite-difference methods, there are two major sources of error: truncation errors associated with the accuracy a of the method and rounding errors due to uncertainty on the input function (in the present case the rounding error is due to the noise added to the PGF) [27]. The former error prefers Δx as small as possible, while the latter prefers larger Δx. Hence, in choosing the step size Δx, one has to balance the contributions from the two errors. We observe for the case $\eta = 10^{-2}$ in Fig. 5 that there exists step sizes Δx for which the first two elements P_1 and P_2 can be extracted with reasonable errors; for P_1 the time-averaged error is much smaller than unity for all choices of distance Δx. For higher k the errors exceed 10^{-1} for all choices of Δx and is dominated by rounding errors [27]. We note here that the step size Δx is upper bounded by $(1/3)/(n + a - 1)$, as we need to sample $(n + a - 1)$ points between $x = 0$ and $x = 1/3$.

By decreasing the noise amplitude by two orders of magnitude to $\eta = 10^{-4}$, we observe in the lower panel of Fig. 5 that choosing $\Delta x \geqslant 0.03$ yields time-averaged errors smaller than 10^{-1} for the $k \leqslant 3$ elements, while the $k \geqslant 4$ elements of the probability distribution still carry too-large errors to be usable for the analysis of scrambling. We thus see that the above derivative-based method [Eq. (26)] of extracting the elements of the probability distribution becomes increasingly sensitive to noise with the degree of the derivative one needs to evaluate, and even for highly precise measurements of the PGF $F(x,t)$ the method Eq. (26) of extracting the probability distribution from the PGF is too sensitive to noise to allow us access to P_k for $k \geqslant 4$.

As a consequence of the high sensitivity to noise—whether the noise originates in experimental imperfections or an approximate PGF due to a finite number of sampled rotations—it does not seem experimentally feasible to extract the probability distribution elements from the PGF. Instead, one should use a protocol dedicated to the extraction of the probability distribution elements. To this end we present in Sec. IV an experimentally relevant protocol that is able to access the elements of the probability distribution without using the PGF.

Finally, we comment on the extraction of the moments of the probability distribution using Eq. (27). Accessing the moments of the probability distribution using the previously described finite-difference method requires one to create states with $x > 1/3$ to approximate the derivative. However, with the constraint of $|\varepsilon| \leqslant 1$, it is impossible to access $x > 1/3$ through the procedure outlined in this section. If one could engineer an ensemble of correlated initial states, then the variance Δ_k could possibly be manipulated sufficiently to allow access to values of $x > 1/3$. However, in the present work we do not investigate this further and leave it for possible future extensions of this work.

IV. METHOD B: ACCESSING ELEMENTS OF THE PROBABILITY DISTRIBUTION FOR SMALL OPERATOR SIZES

In the previous Sec. III we presented a measurement protocol that can access the operator size probability distribution $\{P_k(t)\}$ by measuring the PGF $F(x,t)$ in Eq. (22). Here we

return to the general result in Eq. (10) and develop an alternate NOTOC measurement protocol to access the individual probabilities. This is achieved by employing a different choice of initial state that replaces the one in Eq. (13). The protocol will provide a direct way of obtaining the probabilities and bypasses the sensitivity issues encountered when trying to invert the PGF.

In order to directly access P_k for a given $1 \leqslant k \leqslant N$, we choose a subset $\mathcal{M}_k$ of k particles in the system and prepare them the classically correlated state

$$\rho_Z^{(k)} = \frac{1}{2^k}\left(\mathbb{I}_k + \bigotimes_{i=1}^{k} Z_i\right), \tag{31}$$

while the other $N-k$ particles are left in a maximally mixed state. The resulting state is clearly unentangled as it can be produced as a statistical mixture of product states in the computational basis. Crucially, it has the desired property of being written solely in terms of weight-k operators, as discussed in Sec. II. Thus we may proceed in a similar fashion to method A: We apply a round of uniformly chosen random rotations on each of the particles in $\mathcal{M}_k$. The resulting full state takes the form

$$\rho_0 = \frac{1}{d}\left[\mathbb{1} + \sum_{Q\in C_k(\mathcal{M}_k)} q_Q Q\right], \tag{32}$$

where the modified set $C_k(\mathcal{M}_k)$ is composed of all k-body Pauli operators acting on the particles in $\mathcal{M}_k$. The application of random, local rotations ensures that the coefficients q_Q are independent, identically distributed random variables with mean $\overline{q_Q} = 0$ and variance $\overline{q_Q^2} = 1/3^k$. It follows from Eq. (9) that

$$\overline{G(t)}_{\mathcal{M}_k} = \frac{1}{3^k}\sum_{Q\in C_k(\mathcal{M}_k)} |f[Q;W(t)]|^2, \tag{33}$$

and thus we find

$$P_k(t) = 3^k \sum_{\mathcal{M}_k} \overline{G(t)}_{\mathcal{M}_k}. \tag{34}$$

In order to obtain $P_k(t)$ exactly one needs to obtain $G(t)$ for all possible subsets of $\mathcal{M}_k$ of k particles, thus requiring $\binom{N}{k}$ repetitions of the protocol. While this is in general inefficient, it can be feasible as long as k and N are not too large. In practice, the number of initial states can be reduced in at least two ways. One option is to approximate the sum in Eq. (34) by randomly sampling a reduced number of sets $\mathcal{M}_k$. Alternatively, one can exploit symmetries of the system. For instance, the Hamiltonian in Eq. (23) has a reflection symmetry with respect to the middle of the chain which can be easily leveraged to reduce the state count by a factor of 2. If the system has a full translational symmetry, then the state count can be reduced by a factor of N. The procedure to achieve this is described in the Appendix B.

In the following we present numerical results that demonstrate this method's ability to reconstruct $P_k(t)$ for various instances of the tilted field Ising model introduced in Sec. III A with $N = 6$ particles. We choose the initial operator to be $W(0) = \sigma_x^{(1)}$. Results are obtained by using all possible choices of $\mathcal{M}_k$ for each k and by sampling over $M_{\rm rot}$ realizations of randomized rotations. In Fig. 6 it can be seen that the dynamics of each size probability $P_k(t)$ is faithfully reproduced by the present method, in some cases using as little as $M_{\rm rot} = 100$ rotations. For larger operator sizes $k \geqslant 4$ the effects of finite sampling are more pronounced; however, this is to be expected as the typical probabilities are also smaller. The cases displayed in the figure correspond to $\theta = 0$ (integrable transverse field model, blue curves and symbols) and $\theta = \pi/6$ (chaotic model, red curves and symbols), and it can be readily seen that the protocol accesses the typical features expected in both cases, namely a fast spread of the operators (as seen in the decay of $P_1(t)$ at short times), a subsequent oscillatory behavior for $\theta = 0$, and equilibration for $\theta = \pi/6$.

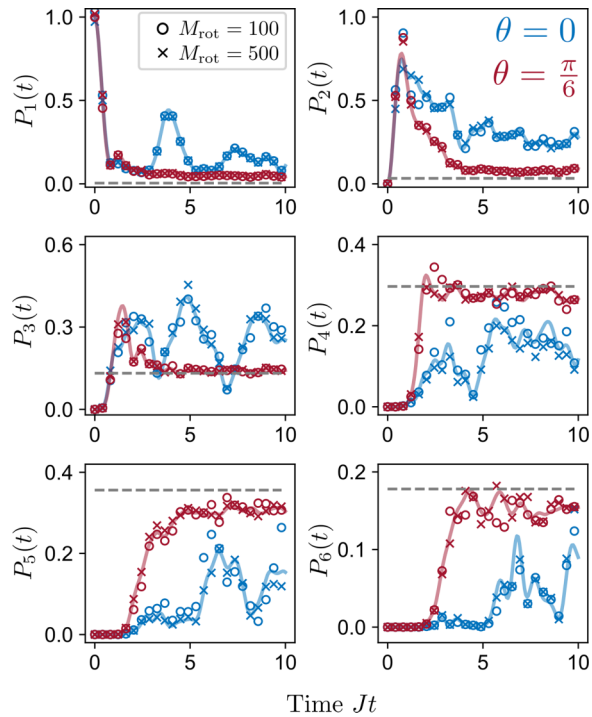

FIG. 6. Application of Method B to obtaining the operator size distribution $\{P_k(t)\}$ for the case of the Ising model of Eq. (23) with $N = 6$ and $J/B = 1$. Each panel shows a different value of $k = 1, \ldots, 6$. Full lines correspond exact numerical results obtained by solving the Heisenberg evolution of the initial operator, here chosen to be $W(0) = \sigma_x^{(1)}$. Symbols correspond to numerical simulations of Method B (Sec. IV) with different choices of the number of sampled rotations $M_{\rm rot} = 100$ (circles) and $M_{\rm rot} = 500$ (crosses). Different colors denote different regimes of the model: integrable case $\theta = 0$ (blue) and chaotic case ($\theta = \pi/6$) (red).

We note that the method presented in this section will not be feasible to obtain the full probability distribution for large system sizes $N \gg 1$, since at some point an exponential number of the subsets $\mathcal{M}_k$'s would be required to exactly recover them from Eq. (34). However, this method can be used to verify whether a system a system has scrambled to k-body operators as a long as $k \sim O(1)$. An alternative benchmark is to compare the values of the obtained P_k's with the ones corresponding to Haar-random evolution, i.e.,

$$P_k^{\rm Haar} = \binom{N}{k}\frac{3^k}{d^2-1}, \tag{35}$$

where $d = 2^N$ (see Refs. [10,12] for additional details). We show these values as gray dashed lines in each of the plots of Fig. 6, where we observe that the observable tends to equilibrate to these values in the chaotic regime ($\theta = \pi/6$) of the model.

V. COMPARISON AND DISCUSSION

A. Extensions to other systems

In this section we present an extension of the NOTOC measurement protocol introduced in Sec. II to collective spin systems, where every spin interacts with every other spin in the system. This class of systems preserves $J^2 = J_x^2 + J_y^2 + J_z^2$, where $J_i = \sum_k \sigma_i^{(k)}$ is the collective spin operator. Some well-known Hamiltonians of this type include p-spin models ($p = 2$ case is referred to as the Lipkin-Meshkov-Glick model) and the quantum kicked top model [29–31]. The dynamics of these models is often studied in the symmetric subspace, the subspace of the whole Hilbert space associated with $J = N/2$ quantum number where N is the number of spins. The dimension of this subspace increases linearly with the number of spins in the system, $d_{ss} = 2J + 1 = N + 1$.

It is natural to study scrambling in this type of system by decomposing the time-evolved operators in a basis associated with the polynomials of collective spin operators, referred to as the spherical tensor operator basis [10,32]. A spherical tensor operator $T_q^{(k)}$ is an operator that transforms under rotations in the same manner as spherical harmonics [33],

$$\mathcal{D}(\alpha,\beta,\gamma)T_q^{(k)}\mathcal{D}^\dagger(\alpha,\beta,\gamma) = \sum_{q'=-k}^{k} \mathcal{D}_{q'q}^{(k)}(\alpha,\beta,\gamma)T_{q'}^{(k)}, \quad (36)$$

where $\mathcal{D}(\alpha,\beta,\gamma) = e^{-iJ_z\alpha}e^{-iJ_y\beta}e^{-iJ_z\gamma}$ and therefore $\mathcal{D}_{q'q}^{(k)}(\alpha,\beta,\gamma) \equiv \langle k, q'|\mathcal{D}(\alpha,\beta,\gamma)|k, q\rangle$. The explicit form of the spherical tensor operators is given by [32]

$$T_q^{(k)}(J) = \sqrt{\frac{2k+1}{2J+1}} \sum_{m,m'=-J}^{J} C_{J\,m;k\,q}^{J\,m'} |J, m'\rangle\langle J, m|, \quad (37)$$

where $C_{J\,m;k\,q}^{J\,m'} = \langle J, m'|J, m; k, q\rangle$ is a Clebsch–Gordan coefficient, and $q = \{-k, -k+1, \ldots, k\}$ for a given rank of the tensor operator $k = \{0, 1, \ldots, N\}$. These operators form an orthonormal basis $\mathrm{Tr}[(T_{q'}^{(k')})^\dagger T_q^{(k)}] = \delta_{k,k'}\delta_{q,q'}$ that spans the Hilbert space associated with the symmetric subspace. Note that a kth rank tensor operator is simply a kth-order polynomial of collective spin operators, and hence the basis consists of polynomials of collective spin operators ranging from order 0 to order N. The rank of these operators can be used to construct the operator size distribution for this class of systems, similarly to the operator size (Hamming weight) used for the Pauli basis. It is then natural to analyze the dynamics of the system by considering an operator of a particular rank k at the initial time and characterizing the scrambling dynamics of the system by considering for each tensor rank k the time evolution of the element $P_k(t)$ of the operator size distribution.

In the following we describe the details of a NOTOC measurement protocol for collective spin systems that can access the elements $P_k(t)$ of the operator size distribution for small operator sizes. As in Sec. II we require the initial state to be prepared from an ensemble that satisfies $\overline{r_i} = 0$ and $\overline{r_i^2} = \Delta_k$ in Eq. (7). The nonzero coefficients r_i are here associated with a particular rank k in the expansion of the density operator in the spherical tensor operator basis $\{T_q^{(k)}\}$. To obtain states that satisfy these properties for the expansion coefficients, we start with a mixed state given by

$$\tilde{\rho}_0 = \frac{1}{d_{ss}}\left[\mathbb{1} + \frac{1}{e_0}T_0^{(k)}\right], \quad (38)$$

where $k > 0$; e_0 is the absolute value of the smallest eigenvalue of $T_0^{(k)}$, which is used to ensure the semidefiniteness of the density operator. Note that $T_0^{(k)}$ is a kth-order polynomial in J_z. For instance, $T_0^{(1)} = c_0^{(1)}J_z$ and $T_0^{(2)} = c_2^{(0)}[3J_z^2 - J(J+1)]$ where $c_0^{(1)}$ and $c_0^{(2)}$ are normalization coefficients. For $k = 1$ the state $\tilde{\rho}_0$ is a thermal equilibrium state of a system placed in an external magnetic field at high temperatures, $\rho = \frac{1}{d}(1 + \epsilon J_z)$ where $\epsilon \ll 1$ [34]. Higher-order states could potentially be prepared as equilibrium states of a system with Hamiltonian consisting of higher-order J_z terms.

Applying global rotations of the form $D(\phi,\theta) = \mathcal{D}(\alpha = \phi, \beta = \theta, \gamma = 0) = e^{-iJ_z\phi}e^{-iJ_y\theta}$ on the state $\tilde{\rho}_0$ leads to

$$\begin{aligned}\rho_0 &= D(\phi,\theta)\,\tilde{\rho}_0\,D^\dagger(\phi,\theta)\\ &= \frac{1}{d_{ss}}\left[\mathbb{1} + \frac{1}{e_0}\sum_{q'=-k}^{k}\langle k, q'|D(\phi,\theta)|k, 0\rangle\, T_{q'}^{(k)}\right]\\ &\equiv \frac{1}{d_{ss}}\mathbb{1} + \sum_{q'=-k}^{k} r_{k,q'}\, T_{q'}^{(k)},\end{aligned} \quad (39)$$

where $r_{k,q'} = (e_0 d_{ss})^{-1}\langle k, q'|D(\phi,\theta)|k, 0\rangle$, and the angles θ and ϕ are sampled randomly from a uniform distribution on the surface of a sphere. This state has zero mean and nonzero variance [35] as expected,

$$\overline{r_{k,q'}} = \int d\Omega\, r_{k,q'}(\phi,\theta) = \frac{1}{e_0 d_{ss}}\overline{\langle k, q'|D(\phi,\theta)|k, 0\rangle} = 0, \quad (40)$$

$$\begin{aligned}\overline{r_{k,q'}^2} &= \int d\Omega\, r_{k,q'}^2(\phi,\theta) = \frac{1}{e_0^2 d_{ss}^2}\overline{|\langle k, q'|D(\phi,\theta)|k, 0\rangle|^2}\\ &= \frac{1}{e_0^2 d_{ss}^2}\frac{4\pi}{2k+1},\end{aligned} \quad (41)$$

where $d\Omega = \sin\theta d\theta d\phi$. Equation (41) is analogous to Eq. (21) in that the expression for the variance depends only on a coarse-grained property of the operator basis element, namely the operator size for the Pauli basis and the operator rank for the spherical tensor basis. Hence we have

$$P_k(t) = \frac{1}{\|W\|_2^2\,\overline{r_{k,q'}^2}}\overline{|\langle W(t)\rangle|^2}, \quad (42)$$

where W is the operator of interest. This result shows that method B of Sec. IV can be naturally adapted to collective spin systems. In general, the procedure shown in this section illustrates how to choose a combination of initial states and randomized operations tailored to the choice of operator basis such that the average NOTOC connects to operator size

distributions. Notice that $N+1$ repetitions of the above protocol for different k will provide us probabilities associated with all operator sizes. However, the state in Eq. (38) is not easily accessible for higher values of k, so this protocol might only be suitable for small system sizes.

B. Relation to previous proposals

The toolbox presented in Sec. II presents some noteworthy connections with previous works which have studied how to diagnose complex many-body dynamics in different settings. For instance, Qi *et al.* [12] proposed a method to probe the growth of an operator O in quantum quench experiments using pure product states of qudits (of local dimension d_L) followed by random local operations. The authors showed that the variance $\delta O(t)^2$ of the expectation value $\langle O(t)\rangle$ over the random realizations yields

$$\delta O(t)^2 = \frac{\mathrm{Tr}(O^2)}{d_L^N} F\left(x = \frac{1}{d_L+1}, t\right), \tag{43}$$

where $F(x,t)$ is the probability-generating function associated with the operator size distribution of $O(t)$. For qubits $d_L = 2$ and so the method probes the PGF at $x = 1/3$. This is exactly the case for the NOTOC when choosing states of the form proposed in method A (Sec. III) using pure states ($\varepsilon = 1$), see Eq. (22). Therefore, our proposed method A recovers the protocol of Ref. [12] for the case of pure states and generalizes it by showing that the PGF can be sampled in a continuum of values $x \leqslant 1/3$ by using mixed states of qubits. It remains to be studied whether this generalization carries over to the case of qudits with $d_L > 2$. Additionally, these methods do not allow one to access $x > 1/3$ directly and it is unclear whether the toolbox of Sec. II provides a way around this by using a clever choice of initial states. Regarding this aspect, we point out that a recent work proposes an alternative method to access the PGF $F(x,t)$ (in principle for any x) using a single-particle mixed states [similar in form to Eq. (31) when $k = 1$] and resorting to forward $U(t)$ *and* backward $U^\dagger(t)$ evolution. Importantly, the analysis we presented in Sec. III B concerning the large sensitivity to noise of the process of obtaining the operator distribution $\{P_k(t)\}$ from its PGF $F(x,t)$ applies to all methods that aim at obtaining the PGF. Our findings indicate that obtaining the distribution from the PGF might be unfeasible in experiments, but also show that properties of the PGF itself could be used a probe for scrambling directly. More detailed work should be carried out to explore this further.

The idea of using mixed states to probe properties of operator evolution was also used recently by Peng *et al.* in Ref. [23], where the goal was to measure single-site two-time correlation functions of the form $\sum_i \mathrm{Tr}[\sigma_z^{(i)}(0)\,\sigma_z^{(i)}(t)]$ in an NMR experiment. This measurement was carried out by first preparing the weakly polarized initial state $\rho_0 \propto [\mathbb{1} + \varepsilon \sum_i \sigma_z^{(i)}]$, then allowing this initial state to become locally randomized by the effect of on-site disorder, and ultimately measuring a tunable observable using inductive measurements, rotations, and on-site disorder. An analogous method for measuring two-site two-time correlation functions was also proposed by the authors. The measurement protocol proposed in Sec. III of the present work shares significant overlap in methodology with that of Ref. [23], however the goals of the two measurement protocols differ and thus the prepared random mixed states and ultimate measurements are also different. For the purpose of extracting the operator size distribution, we note that by using the two-time correlation function studied in Ref. [23] one will have to extend the method of Ref. [23] to all m-site correlation functions, with $m \leqslant N$. This will likely not be feasible due to the nonlocal nature of the operators to be measured, as well as the issue of the exponentially growing number of operators one needs to measure which was discussed in Sec. IV for our proposed measurement protocol.

Finally, we comment on the connection between our proposal and that of Vermersch *et al.* [20], where the authors propose a way to measure OTOCs without using time-reversal operations or auxiliary systems. Instead, their proposal is based on performing randomized unitaries on a set of initial states and extracting the OTOCs from the statistical correlations between the measurement results. In principle, this method allows one to reconstruct the operator size distribution if one repeats the procedure for (exponentially many) choices of the operator R in Eq. (3). This can be achieved by using averages of OTOCs to obtain the moments of the $\{P_k(t)\}$ distribution, as outlined in Sec. VI of Ref. [10]. In contrast, our method can be seen as a way of using similar tools (i.e., preparation of product states, randomized local operations, and local measurements) to probe the operator size distribution *directly*, thus bypassing the calculation of OTOCs.

C. Relation to fidelity OTOCs

Finally, we discuss the physical interpretation of the different tools and quantities used to study quantum information scrambling. The first aspect is related to the fidelity OTOCs, introduced in Sec. II, which are a class of correlation functions obtained from the usual OTOCs of the form in Eq. (3) by choosing the early-time operator R to be the projector onto the initial state $R = \rho_0$. The use of fidelity OTOCs attracted widespread attention in the community because they are *technically* an OTOC but can be measured as a single expectation value of a (often) local operator W when the initial state is pure, $\rho_0 = |\psi_0\rangle\langle\psi_0|$ [21]. Fidelity OTOCs have interesting connections to quantities like the quantum Fisher information [1] and the Loschmidt echo [36].

The fact that typically ρ_0 is a nonlocal operator makes the fidelity OTOC relinquish the usual interpretation of OTOCs as measures of information scrambling. In particular, the relation between OTOCs involving Pauli operators and moments of the operator size distributions [6,10] does not apply to fidelity OTOCs. However, our present work shows that fidelity OTOCs are indeed connected fundamentally to operator size distributions if one takes the initial pure state $|\psi_0\rangle$ to be a product state and then considers the average fidelity OTOC over many realizations of local random rotations on the initial state.

VI. CONCLUSION

In this article we have proposed a measurement protocol for probing quantum information scrambling by measuring operator size distributions. Our measurement protocol requires the preparation of separable mixed states followed by

local operations and final-time measurements of local operators, and circumvents the typical use of OTOCs to probe scrambling properties [37,38]. We have demonstrated that the choice of initial separable mixed states in our measurement protocol provides multiple ways to access the operator size distribution, and we comment on the experimental feasibility for two particular methods, one based on first extracting the probability-generating function for the operator size distribution, and the other focused on obtaining the elements of the operator size distribution directly.

The application of our measurement protocol is illustrated in detail for the 1D tilted field Ising model, a well-known platform for studying many-body quantum chaos [26,39,40]. We numerically demonstrate the characterization of quantum information scrambling for this model using our proposed measurement protocol. We have related our proposed protocol with well-established methods of characterizing scrambling, including those based on OTOCs, and were able to establish connections between the well-established fidelity OTOC and operator size distributions using our results presented in this article. Finally, we have exemplified the extension of our measurement protocol to other types of quantum systems by considering our protocol for the case of collective spin systems. The collective spin case emphasizes further the role that state preparation plays in our measurement protocol.

In the discussion of our proposed measurement protocol's connection to the PGF, we found that the extraction of moments of the operator size distribution was not possible due to constraints on the prepared initial states that prevent us from accessing $x > 1/3$ in the probability-generating function $F(x)$. From preliminary numerical results we expect that the extraction of the first and second moments of the operator size distribution would be both resilient to noise and reasonable to implement in experiment if one could obtain values of $x > 1/3$, hence finding a way to prepare initial states allowing the extraction of moments of the operator size distribution would be an obvious extension of the present work. This potential extension should be compared to the proposal in Ref. [9] where the probability-generating function can be accessed for any x, but at the expense of requiring the implementation of time-reversal of the many-body evolution.

In Sec. V A we presented the extension of our NOTOC measurement protocol to collective spin systems. The extension of the measurement protocol to many-body qudit systems and other systems of interest for studying the nature of scrambling would be an interesting task that we leave for future work. We point out that studies of operator size distributions for many-body systems beyond qubits are also scarce, with some exceptions [6,41]. In particular, we note the case where the system of interest interacts with the environment, thus forcing one to probe scrambling and the operator size distribution in the presence of decoherence [9]. The NOTOC measurement protocol may be extended to open quantum systems using the analysis presented in Sec. II B of Ref. [21], from which a detailed analysis of the effect of decoherence may be carried out. Crucially one should revisit the definition of the coarse-grained operator size distribution Eq. (2) and consider the effect of decoherence on, e.g., the normalization of this distribution. Additionally, the relationship between the fidelity OTOC Eq. (4) and the NOTOC Eq. (5) becomes more complicated for open quantum systems [21].

Finally, we remark that the choice of measurement protocol is informed by the quantity of interest. For the operator size distribution there are three main types of quantities of interest, as illustrated in Fig. 1(a): elements of the distribution, moments of the distribution, and the probability-generating function. The elements of the distribution present the most fine-grained information about the operator size distribution, while the probability-generating function contains the least fine-grained information. Even so, our analysis indicates that the probability-generating function may be sufficient for diagnosing scrambling and quantum chaos. Method B of our proposed measurement protocol is able to obtain the elements of the distribution directly, while method A provides access to the probability-generating function. Complimentary to our proposed measurement protocol, in Ref. [10] we discussed how to obtain moments of the operator size distribution through appropriate averaging over multibody OTOCs. For a given physical system, one of these approaches may prove advantageous in exploring the properties of the operator size distribution. Thus, one should let the choice of measurement protocol be informed by the physical system, the task at hand, and the experimental capabilities.

ACKNOWLEDGMENTS

The authors acknowledge Ivan H. Deutsch, Lorenza Viola, and Andrew Zhao for insightful discussions. This work is supported by a collaboration between the U.S. DOE and other agencies. This work is based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator. P.D.B. acknowledges support from the U.S. National Science Foundation through the FRHTP Grant No. PHY-2116246. K.C. acknowledges support from Ministère de l'Économie et de l'Innovation du Québec and the Natural Sciences and Engineering Research Council of Canada. Work at the University of Strathclyde was supported by AFOSR Grant No. FA9550-181-1-0064.

APPENDIX A: FINITE-DIFFERENCE METHOD DETAILS FOR METHOD A

The nth derivative at $x = 0$ of the probability-generating function $F(x,t)$ defined in Eq. (22) is implemented in the present work as a forward-only finite-difference method [27], which takes the form Eq. (28) and is repeated here for convenience:

$$F^{(n)}(x=0,t) = (\Delta x)^{-n} \sum_{m=0}^{n+a-1} c_m^{(n,a)} F(m\,\Delta x, t) + \mathcal{O}(\Delta x^a). \tag{A1}$$

Here a is the accuracy of the finite-difference method, and the step size Δx is the distance between the $n+a$ points $\{x = m\Delta x\}_{m=0}^{n+a-1}$ used to approximate the derivative. The forward-only finite-difference coefficients $c_m^{(n,a)}$ are given in Table I for $a = 1$ and in Table II for $a = 2$ and were obtained using Ref. [42]. Figure 5 in Sec. III B was created using the forward-only finite-difference method Eq. (28) with the coefficients of Table I.

TABLE I. Forward-only finite-difference coefficients $c_m^{(n,a)}$ for derivatives of degree $n \leqslant 6$ using accuracy $a = 1$. The finite-size-difference coefficients were obtained using Ref. [42].

n	$c_0^{(n,1)}$	$c_1^{(n,1)}$	$c_2^{(n,1)}$	$c_3^{(n,1)}$	$c_4^{(n,1)}$	$c_5^{(n,1)}$	$c_6^{(n,1)}$
1	−1	1	–	–	–	–	–
2	1	−2	1	–	–	–	–
3	−1	3	−3	1	–	–	–
4	1	−4	6	−4	1	–	–
5	−1	5	−10	10	−5	1	–
6	1	−6	15	−20	15	−6	1

While in theory one can increase the accuracy a arbitrarily to minimize the error $\mathcal{O}(\Delta x^a)$ of the approximation for any $\Delta x \ll 1$, in practice the method becomes increasingly sensitive to noise as we increase the accuracy a. This increased sensitivity to noise is due to the number of points sampled $n + a$ limiting the values that Δx can take, and in the vicinity of vanishing Δx the dominant error is not the approximation of the derivative in Eq. (A1), but rather rounding errors due to the presence of the noise [27].

Figure 7 displays a comparison between the extraction of the elements of the probability distribution for accuracies $a = 1$ (solid curves) and $a = 2$ (dashed curves). Plotted is the time-averaged error $\overline{\Delta P_k^\eta}$ as defined in Eq. (30), averaged over times $Jt \in [0, 10]$ for each realization of the noise Eq. (29) and subsequently averaged over 100 noise realizations. We are considering the same system parameters and operator as in Sec. III B. We observe for the higher noise amplitude $\eta = 10^{-2}$ (upper panel) that the lower accuracy method generally outperforms the higher accuracy method for the considered step sizes Δx, the exception being for P_1 when considering $\Delta x \geqslant 0.02$. The crossover between the $a = 1$ and $a = 2$ curves at $\Delta x \approx 0.02$ for P_1 indicates a transition from being dominated by the sensitivity of the method to the added noise Eq. (29) for smaller Δx values, to being dominated by the error due to the finite-difference method's approximations for larger Δx values.

For the case of $\eta = 10^{-4}$ (lower panel in Fig. 7) we clearly see that the $a = 2$ method outperforms the $a = 1$ method for P_1, and a clear crossover point is also observed for P_2. Therefore, when working with a sufficient small noise amplitude such as $\eta = 10^{-4}$, it may be beneficial to use a higher accuracy for the finite-difference method.

TABLE II. Forward-only finite-difference coefficients $c_m^{(n,a)}$ for derivatives of degree $n \leqslant 6$, using accuracy $a = 2$. The finite-size-difference coefficients were obtained using Ref. [42].

n	$c_0^{(n,2)}$	$c_1^{(n,2)}$	$c_2^{(n,2)}$	$c_3^{(n,2)}$	$c_4^{(n,2)}$	$c_5^{(n,2)}$	$c_6^{(n,2)}$	$c_7^{(n,2)}$
1	−3/2	2	−1/2	–	–	–	–	–
2	2	−5	4	−1	–	–	–	–
3	−5/2	9	−12	7	−3/2	–	–	–
4	3	−14	26	−24	11	−2	–	–
5	−7/2	20	−95/2	60	−85/2	16	−5/2	–
6	4	−27	78	−125	120	−69	22	−3

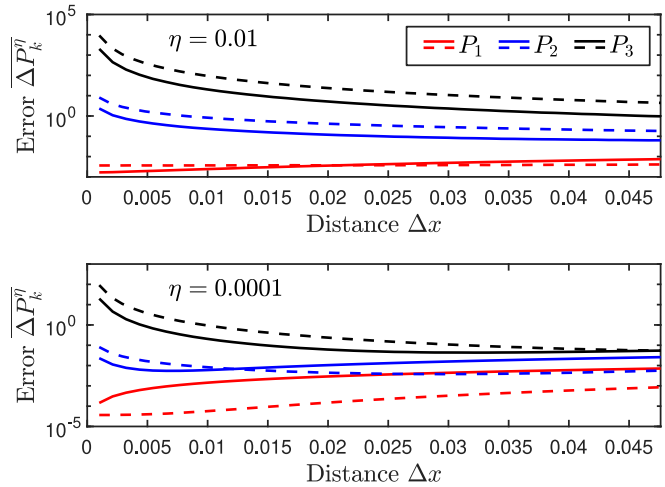

FIG. 7. Comparison of time-averaged error $\overline{\Delta P_k^\eta}$ [Eq. (30)] for accuracy $a = 1$ (solid curves) and $a = 2$ (dashed curves). The used parameters are identical to those used for Fig. 5. Shown are the time-averaged errors for P_1 (red), P_2 (blue), and P_3 (black curves), for noise amplitude $\eta = 10^{-2}$ (upper panel) and $\eta = 10^{-4}$ (lower panel).

APPENDIX B: REDUCING STATE COUNTS IN METHOD B BY EXPLOITING SYMMETRIES

Here we show how to reduce the number of states required to produce in Method B of Sec. IV by exploiting symmetries. In many cases, we are interested in the operator size distribution of an operator W which is a single site Pauli operator. For concreteness, let us assume this operator acts on site 1, $W \equiv W_1$. The NOTOC of Eq. (17) is constructed by measuring the expectation value of W_1 at the end of the protocol. In principle, however, there is no cost associated to measuring expectation values on different sites and using those get more refined information about the operator dynamics. Suppose the system has a translational invariance described by an operator T_l,

$$T_l^\dagger \left(\bigotimes_j q_j \right) T_l = \bigotimes_j q_{(j+l \mod N)}, \tag{B1}$$

such that, for instance, $T_1^\dagger (A \otimes B \otimes C) T_1 = C \otimes A \otimes B$, etc. Consider our system's evolution given by $U(t)$ such that $T_l^\dagger U(t) T_l = U(t)$. Then we have

$$W_1(t) = \sum_Q f[Q; W_1(t)] Q, \tag{B2}$$

$$T_l^\dagger U(t)^\dagger W_1 U(t) T_l = \sum_Q f[Q; W_1(t)] T_l^\dagger Q T_l, \tag{B3}$$

$$W_{1+l}(t) = \sum_Q f\left[T_l Q_l T_l^\dagger, W_1(t)\right] Q_l, \tag{B4}$$

when we have defined $Q_l \equiv T_l^\dagger Q T_l$ and used the translation invariance property. By definition, the LHS of the last equation equals

$$W_{1+l}(t) = \sum_{Q'} f\left[Q', W_{1+l}(t)\right] Q', \tag{B5}$$

and so we have that

$$f[Q, W_{1+l}(t)] = f[T_l Q T_l^\dagger, W_1(t)]. \tag{B6}$$

Suppose we start from a fixed initial state $\rho_1 = \frac{1}{d}(\mathbb{I} + \sum'_{Q\in C_k} r_Q Q)$, where the sum is over only a subset of operators of weight k, starting at position 1 up to k. If we measure $\langle W_{1+l}(t)\rangle$ for $l = 0, \ldots, N-1$, then

$$\langle W_{1+l}(t)\rangle = \sum_Q f[Q; W_{1+l}]\langle Q\rangle_{\rho_1} \tag{B7}$$

$$= \sum_Q f\big[T_l Q T_l^\dagger; W_1(t)\big]\mathrm{Tr}(\rho_1 Q) = \sum_Q f\big[T_l Q T_l^\dagger; W_1(t)\big]\mathrm{Tr}\big(T_l \rho_1 T_l^\dagger T_l Q T_l^\dagger\big) \tag{B8}$$

$$= \sum_{\tilde{Q}} f\big[\tilde{Q}; W_1(t)\big]\langle \tilde{Q}\rangle_{T_l\rho_1 T_l^\dagger}. \tag{B9}$$

In conclusion, that means that measuring $\langle W_{1+l}(t)\rangle$ yields the same result as having done the protocol starting from $T_l\rho_1 T_l^\dagger$, thus reducing the state count by a factor of N provided one can measure expectation values in all sites in the case of full translational invariance. The procedure works similarly if one only has a reflection symmetry (i.e., open boundary conditions).

APPENDIX C: APPROXIMATING PROPERTIES OF OPERATOR DISTRIBUTIONS WITH LIMITED NUMBER OF ROTATIONS

1. Method A: PGF

In Sec. III A we studied method A, the NOTOC measurement protocol aiming to access the PGF of the operator size distribution. In particular, in Fig. 2 we illustrated the PGF as a function of time for several values of the polarization parameter ε and for both the integrable case with magnetic field angle $\theta = 0$ and the chaotic case with $\theta = \pi/6$. For each choice of ε in Fig. 2, we calculated the exact PGF (the solid black curves) as well as an approximate PGF obtained using a subset of $M_{\mathrm{rot}} = 1000$ randomly selected rotations (solid dots). Furthermore, for the case $\varepsilon = 1$ we also plotted the approximate PGF obtained using a subset of only $M_{\mathrm{rot}} = 100$ randomly selected rotations (blue crosses). Figure 2 demonstrates that the PGF is well approximated using a small subset of rotations compared to the $6^N = 46\,656$ rotations needed to obtain the exact PGF.

In this Appendix we provide an in-depth study of the error introduced by sampling only a subset of the 6^N possible rotations, complimentary to the results found in the main text. Due to limitations on simulating larger system sizes exactly, we focus on errors for small system sizes $4 \leqslant N \leqslant 6$ for which the exact PGF is readily simulable. We quantify the error between the exact PGF F and the approximate PGF $\widetilde{F}$ by calculating the root-mean-squared error

$$\Delta F_{\mathrm{RMS}} = \sqrt{\frac{1}{t_f - t_i}\int_{t_i}^{t_f} dt[F(x,t) - \widetilde{F}(x,t)]^2} \approx \sqrt{\frac{1}{K}\sum_{j=1}^{K}[F(x,t_j) - \widetilde{F}(x,t_j)]^2}, \tag{C1}$$

between the exact and approximate PGF over the K timesteps t_j in the interval $0 \leqslant Jt \leqslant 10$ (the time interval shown in Fig. 2). Here the approximate PGF $\widetilde{F}$ is obtained by sampling only a limited number of rotations $M_{\mathrm{rot}} \leqslant 6^N$. Figure 8 displays the error as a function of the number of rotations M_{rot} sampled, where the rotations in each subset are chosen uniformly at random. We have here focused on the case $\varepsilon = 1$, corresponding to $x = 1/3$, which is the PGF with the highest visibility in Fig. 2. Three system sizes are plotted, with the black dashed $N = 6$ curve being representative of the case studied in the main text.

For all system sizes we observe an overall trend of the error decreasing with the number of rotations included in the sampling, with errors typically being an order of magnitude smaller than the PGF value. There is no discernible dependence of the error on the system size for the sizes considered here. We note that the case $N = 4$ requires $6^4 = 1296$ rotations to obtain the exact PGF, hence the $M_{\mathrm{rot}} = 1000$ rotations considered for the right-most $N = 4$ datapoint constitutes a majority of the rotations needed for the exact result.

The nonmonotonicity observed in the error as a function of the number of rotations sampled is to be expected, as the error for each datapoint in Fig. 8 is calculated for one particular randomly selected subset of rotations. One could average the error over different randomly selected subsets to alleviate the nonmonotonic behavior; however, the results presented in Fig. 8 are sufficient to draw conclusions about the errors induced by choosing only a subset of rotations in approximating the PGF.

2. Method B: Elements of the probability distribution

Along similar lines, we analyze the effect of reconstructing each element of P_k with a limited number M_{rot} of rotations. In Fig. 6 we showed that for a system size $N = 6$, using $M_{\mathrm{rot}} = 100$ already gives qualitative agreement with the exact result. In order to systematically study this aspect, we consider the total variational distance (TVD) which is defined as

$$D_{\mathrm{TV}}(P, Q) = \sup_k |P_k - Q_k|, \tag{C2}$$

between the ideal probability distribution $\{P_k(t)\}$ and the one reconstructed with our method $\{\tilde{P}_k(t)\}$. We choose to quantify

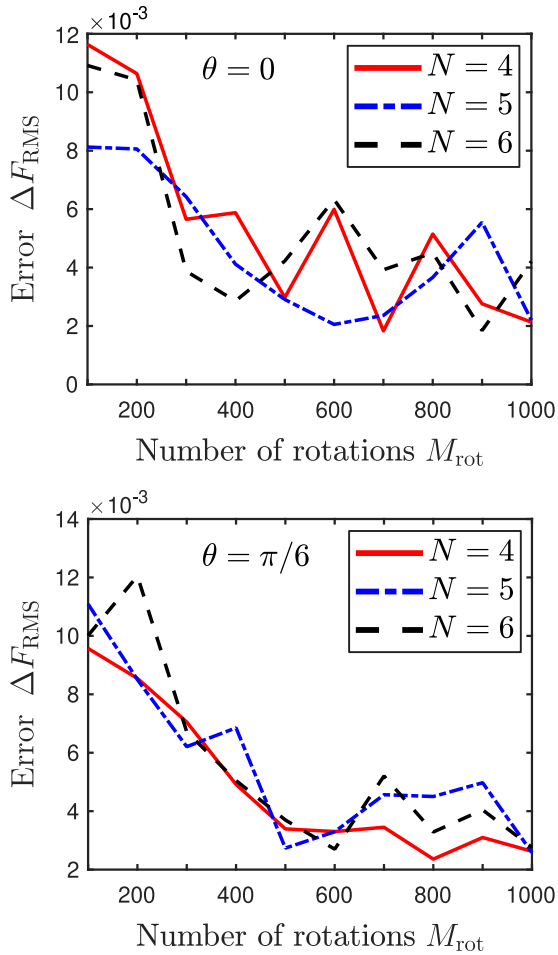

FIG. 8. Root-mean-square error ΔF_{RMS} between the exact PGF $F_{\mathrm{exact}}(x, t_j)$ and the approximate PGF $F_{\mathrm{approx.}}$, as a function of the number of rotations used for the approximation. The upper panel shows results for the integrable case $\theta = 0$, whereas the lower panel shows results for the chaotic case $\theta = \pi/6$. In both cases we see a general trend of the error decreasing with the number of rotations sampled for all system sizes N considered, although some nonmonotonic behavior is present. This nonmonotonicity is due to the random selection of rotations for each datapoint.

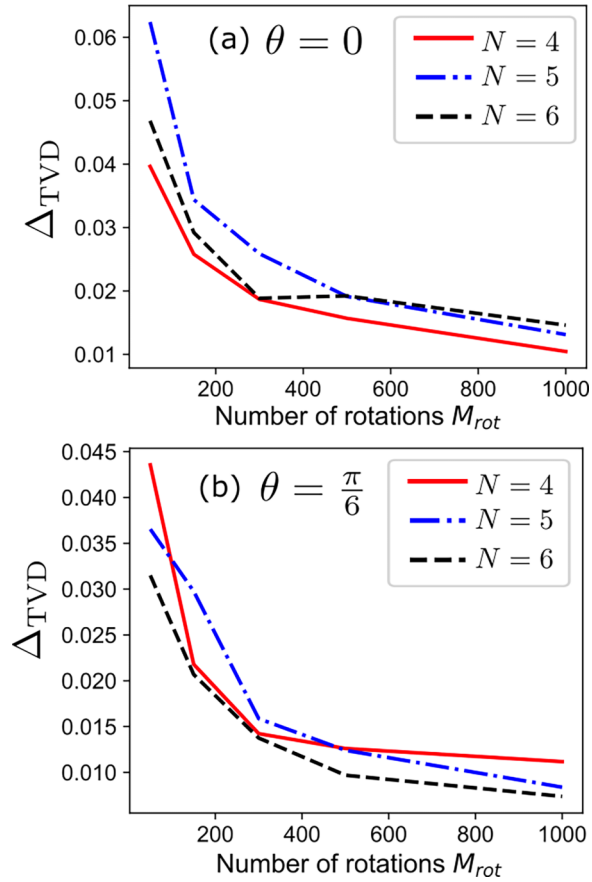

FIG. 9. Time-averaged total variational distance, defined in Eq. (C4), between the exact operator size distribution $\{P_k(t)\}$ and the approximate one obtained by Method B $\{\tilde{P}_k\}$, as a function of the number of rotations used by the method, M_{rot}. Panel (a) shows results for the integrable Ising model $\theta = 0$, while (b) corresponds to the chaotic case $\theta = \pi/6$. We observe that the error in reconstructing the full probability distribution decays with M_{rot}, with no strong dependence on system size N. General evolution parameters are the same as used in Fig. 6.

the error via the time-averaged TVD,

$$\Delta_{\mathrm{TVD}} = \frac{1}{t_f - t_i} \int_{t_i}^{t_f} dt\, D_{\mathrm{TV}}(P(t), \tilde{P}(t)) \tag{C3}$$

$$\simeq \frac{1}{K} \sum_{j=1}^{K} D_{\mathrm{TV}}(P(t_j), \tilde{P}(t_j)), \tag{C4}$$

where again t_j is in the interval $0 \leqslant Jt \leqslant 10$ and we consider the cases $N = 4, 5, 6$. We show results for the time-averaged TVD as a function of the number of rotations in Fig. 9. It is clearly seen that the error in reconstructing the full probability distribution decreases with the number of rotations used, and that deviations of 0.02–0.05 are readily available with a few hundred rotations. We observe also that the overall behavior does not depend strongly on system size for the cases studied.

[1] R. J. Lewis-Swan, A. Safavi-Naini, J. J. Bollinger, and A. M. Rey, Unifying scrambling, thermalization and entanglement through measurement of fidelity out-of-time-order correlators in the dicke model, Nat. Commun. **10**, 1581 (2019).

[2] E. M. Fortes, I. García-Mata, R. A. Jalabert, and D. A. Wisniacki, Gauging classical and quantum integrability through out-of-time-ordered correlators, Phys. Rev. E **100**, 042201 (2019).

[3] Y. Sekino and L. Susskind, Fast scramblers, J. High Energy Phys. 10 (2008) 065.

[4] P. Hayden and J. Preskill, Black holes as mirrors: quantum information in random subsystems, J. High Energy Phys. 09 (2007) 120.

[5] M. Dupont, N. Didier, M. J. Hodson, J. E. Moore, and M. J. Reagor, Entanglement perspective on the quantum approximate optimization algorithm, Phys. Rev. A **106**, 022423 (2022).

[6] Q. Zhuang, T. Schuster, B. Yoshida, and N. Y. Yao, Scrambling and complexity in phase space, Phys. Rev. A **99**, 062334 (2019).

[7] D. E. Parker, X. Cao, A. Avdoshkin, T. Scaffidi, and E. Altman, A universal operator growth hypothesis, Phys. Rev. X **9**, 041017 (2019).

[8] D. A. Roberts, D. Stanford, and A. Streicher, Operator growth in the SYK model, J. High Energy Phys. 06 (2018) 122.

[9] T. Schuster and N. Y. Yao, Operator growth in open quantum systems, Phys. Rev. Lett. **131**, 160402 (2023).

[10] S. Omanakuttan, K. Chinni, P. D. Blocher, and P. M. Poggi, Scrambling and quantum chaos indicators from long-time properties of operator distributions, Phys. Rev. A **107**, 032418 (2023).

[11] J. D. Noh, Operator growth in the transverse-field Ising spin chain with integrability-breaking longitudinal field, Phys. Rev. E **104**, 034112 (2021).

[12] X.-L. Qi, E. J. Davis, A. Periwal, and M. Schleier-Smith, Measuring operator size growth in quantum quench experiments, arXiv:1906.00524.

[13] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, Chaos in quantum channels, J. High Energy Phys. 02 (2016) 004.

[14] M. Gärttner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, and A. M. Rey, Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet, Nat. Phys. **13**, 781 (2017).

[15] M. Gärttner, P. Hauke, and A. M. Rey, Relating out-of-time-order correlations to entanglement via multiple-quantum coherences, Phys. Rev. Lett. **120**, 040402 (2018).

[16] J. Li, R. Fan, H. Wang, B. Ye, B. Zeng, H. Zhai, X. Peng, and J. Du, Measuring out-of-time-order correlators on a nuclear magnetic resonance quantum simulator, Phys. Rev. X **7**, 031011 (2017).

[17] K. X. Wei, P. Peng, O. Shtanko, I. Marvian, S. Lloyd, C. Ramanathan, and P. Cappellaro, Emergent prethermalization signatures in out-of-time ordered correlations, Phys. Rev. Lett. **123**, 090605 (2019).

[18] N. Y. Yao, F. Grusdt, B. Swingle, M. D. Lukin, D. M. Stamper-Kurn, J. E. Moore, and E. A. Demler, Interferometric approach to probing fast scrambling, arXiv:1607.01801.

[19] K. A. Landsman, C. Figgatt, T. Schuster, N. M. Linke, B. Yoshida, N. Y. Yao, and C. Monroe, Verified quantum information scrambling, Nature (London) **567**, 61 (2019).

[20] B. Vermersch, A. Elben, L. M. Sieberer, N. Y. Yao, and P. Zoller, Probing scrambling using statistical correlations between randomized measurements, Phys. Rev. X **9**, 021061 (2019).

[21] P. D. Blocher, S. Asaad, V. Mourik, M. A. I. Johnson, A. Morello, and K. Mølmer, Measuring out-of-time-ordered correlation functions without reversing time evolution, Phys. Rev. A **106**, 042429 (2022).

[22] For pure states Eq. (5) is equivalent to a fidelity OTOC, which in turn is a special case of the more general OTOC in Eq. (3) when one chooses $R(0) = \rho_0 = |\Psi_0\rangle\langle\Psi_0|$.

[23] P. Peng, B. Ye, N. Y. Yao, and P. Cappellaro, Exploiting disorder to probe spin and energy hydrodynamics, arXiv:2209.09322.

[24] This assumption still allows us to choose W to be any unnormalized Pauli operator from the set $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}^{\otimes N}$, which would be the natural operators to measure for an N-qubit system.

[25] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate Discrete Distributions*, 3rd ed. (John Wiley & Sons, New York, 2005).

[26] J. Karthik, A. Sharma, and A. Lakshminarayan, Entanglement, avoided crossings, and quantum chaos in an ising model with a tilted magnetic field, Phys. Rev. A **75**, 022304 (2007).

[27] K. Vuik, F. Vermolen, and M. van Gijzen, *Numerical Methods for Ordinary Differential Equations* (TU Delft Open, Delft, 2023).

[28] If one instead assumes that the noise is additive, i.e., $F_\eta(x,t) = F(x,t) + \delta(\eta)$, then one finds an increased sensitivity to noise compared to the multiplicative case due to the value of the PGF being close to zero whenever the parameter x is also close to zero, and hence the impact of additive noise on the PGF is higher than multiplicative.

[29] N. Meshkov, A. J. Glick, and H. J. Lipkin, Validity of many-body approximation methods for a solvable model: (ii). linearization procedures, Nucl. Phys. **62**, 199 (1965).

[30] P. Ribeiro, J. Vidal, and R. Mosseri, Thermodynamical limit of the Lipkin-Meshkov-Glick model, Phys. Rev. Lett. **99**, 050402 (2007).

[31] F. Haake, M. Kuś, and R. Scharf, Classical and quantum chaos for a kicked top, Z. Phys. B **65**, 381 (1987).

[32] A. B. Klimov and J. L. Romero, A generalized Wigner function for quantum systems with the SU(2) dynamical symmetry group, J. Phys. A: Math. Theor. **41**, 055303 (2008).

[33] L. E. Ballentine, *Quantum Mechanics: A Modern Development* (World Scientific, Singapore, 2014).

[34] K. X. Wei, C. Ramanathan, and P. Cappellaro, Exploring localization in nuclear spin chains, Phys. Rev. Lett. **120**, 070501 (2018).

[35] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988).

[36] T. Gorin, T. Prosen, T. H. Seligman, and M. Žnidarič, Dynamics of loschmidt echoes and fidelity decay, Phys. Rep. **435**, 33 (2006).

[37] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, Measuring the scrambling of quantum information, Phys. Rev. A **94**, 040302(R) (2016).

[38] J. Marino and A. M. Rey, Cavity-qed simulator of slow and fast scrambling, Phys. Rev. A **99**, 051803(R) (2019).

[39] T. Prosen, General relation between quantum ergodicity and fidelity of quantum dynamics, Phys. Rev. E **65**, 036208 (2002).

[40] C. Mejía-Monasterio, G. Benenti, G. G. Carlo, and G. Casati, Entanglement across a transition to quantum chaos, Phys. Rev. A **71**, 062324 (2005).

[41] M. S. Blok, V. V. Ramasesh, T. Schuster, K. O'Brien, J.-M. Kreikebaum, D. Dahlen, A. Morvan, B. Yoshida, N. Y. Yao, and I. Siddiqi, Quantum information scrambling on a superconducting qutrit processor, Phys. Rev. X **11**, 021010 (2021).

[42] C. R. Taylor, Finite difference coefficients calculator, https://web.media.mit.edu/crtaylor/calculator.html.