

Classical simulation of noisy quantum circuits via locally entanglement-optimal unravelings

Simon Cichy ^{1,*}, Paul K. Faehrmann ¹, Lennart Bittel ¹, Jens Eisert ^{1,2,3,†} and Hakop Pashayan ^{1,4,‡}

¹*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

²*Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner-Platz 1, 14109 Berlin, Germany*

³*Fraunhofer Heinrich Hertz Institute, 10587 Berlin, Germany*

⁴*Hon Hai (Foxconn) Research Institute, Taipei 114699, Taiwan*



(Received 15 October 2025; revised 22 June 2026; accepted 9 July 2026; published 17 August 2026)

Classical simulations of noisy quantum circuits are instrumental to our understanding of the behavior of real-world quantum systems and the identification of regimes where one expects quantum advantage. In this work, we present a highly parallelizable tensor-network-based classical algorithm, equipped with rigorous accuracy guarantees, for simulating n -qubit quantum circuits with arbitrary single-qubit noise. Our algorithm represents the state of a noisy quantum system by a particular ensemble of matrix product states from which we stochastically sample. Each pure state evolved under a single-qubit noise process is then represented by the ensemble of states that achieves the minimal average entanglement (the entanglement of formation) between the noisy qubit and the remainder. This approach lets us use a more compact representation of the quantum state for a given accuracy requirement and noise level. For a given maximum bond dimension χ and circuit, our algorithm comes with an upper bound on the simulation error, runs in $\text{poly}(n, \chi)$ -time, and improves upon related prior work (1) in scope by extending analytic methods from the three commonly considered noise models to general single-qubit noise, (2) in performance by deriving an exact, closed-form solution to the local entanglement minimization problem, previously approached via variational heuristics, thereby guaranteeing local optimality without numerical optimization overhead, and (3) in conceptual contribution by showing that the fixed unraveling used in prior work becomes equivalent to our choice of unraveling in the special case of depolarizing, dephasing, and amplitude damping noise acting on a maximally entangled state.

DOI: [10.1103/jj2r-y777](https://doi.org/10.1103/jj2r-y777)

I. INTRODUCTION

Quantum computers have the potential to outperform classical computers in certain structured computational problems. Significant effort has been directed towards delineating the boundary between those that admit an efficient classical solution and those that can be used to demonstrate quantum advantage. Some important results in this direction have focused on the specific structure of a problem necessary for achieving a quantum advantage [1–4]. Others have shown the hardness of efficient classical simulation [5–11] by relying on plausible complexity-theoretic assumptions and an anticoncentration property of certain classes of quantum circuits [11]. However, the bulk of scientific effort has focused on the classical simulation of quantum computation.

Classical simulation has served as an invaluable tool for the prediction and study of quantum phenomena as

well as the precise exploration of the quantum advantage boundary. Efficient classical simulation algorithms for stabilizer circuits [12,13], matchgate circuits [14,15] (also known as noninteracting fermions), and nonentangling circuits identify necessary resources for quantum computational speedup. Extensions of these classical simulation techniques beyond the polynomial runtime regime have established operationally meaningful quantifications of these resources. Prominent examples include the stabilizer rank [16,17], the stabilizer extent [18,19] which has been generalized to the mixed-state setting [20], extended to measures of nonmatchgateness [21–23] and measures related to entanglement such as a matrix product state's (MPS) [24–26] bond dimension in a tensor network description [27–29]. As the quantum computational task moves away from the islands of classically efficiently simulable quantum circuits, these resources quantifiably bound the exponential runtime growth of classical simulators, breaking the direct exponential dependence on the number of qubits.

Particularly relevant to the recent thrust of quantum technological development is the deep and fascinating nuance introduced by the presence of *noise* in quantum computation. At sufficiently low noise levels, in principle, quantum error correction addresses the challenges posed by the inevitable presence of quantum noise, but it introduces significant overheads [30]. Without error correction, the stochastic nature of noise can diminish quantum computational resources such

*Contact author: simon.cichy@fu-berlin.de

†Contact author: jense@zedat.fu-berlin.de

‡Contact author: hakop.pashayan@foxconn.com

as nonstabilizerness (also known as magic) [20,31,32] and entanglement [31].

Growing evidence suggests that, at logarithmic depth or beyond, typical quantum circuits in certain families that are hard to efficiently classically simulate become efficiently classically simulable in the presence of local depolarizing noise [33–38], so that an eventual quantum advantage would disappear. However, these results are limited in scope, both in the types of circuits and noise that they can accommodate. Indeed, they apply to typical random quantum circuits, limiting the insight one can draw about what quantum circuit properties computationally challenge a classical approach. Also, beyond the standard depolarizing noise model, even less is known about how noise affects the hardness of classical simulation.

Generalization to the rich landscape of local noise models is crucial for turning a largely numerical and model-specific picture into a unified and principled theory that delineates the classical-quantum boundary. First, this landscape exhibits stark phenomenological differences. Indeed, while local depolarizing noise appears to aid classical simulability, local coherent noise can hinder it.¹ Thus, in the absence of this generalization, a quantitative or even qualitative characterization of classical simulability in the presence of noise is incomplete. Second, such generalization can be invaluable in connecting the well-understood pockets of parameter space and facilitating the transfer of tools and observations between previously disconnected settings.

In this work, we focus on and build on matrix-product-state-based (MPS-based) unraveling simulators from an entirely new perspective. This is a classical simulation technique that is highly parallelizable and amenable to large-scale simulation [40,41]. Importantly, this approach makes explicit use of the fact that noise can decay entanglement, making it particularly well suited to simulation via tensor network techniques.

This work, and MPS-based unraveling simulators more generally, focus on utilizing the presence of noise to find a low-entanglement representation of the quantum state. This approach is underpinned by the observation that, in contrast to pure quantum states, a mixed quantum state ρ can be nonuniquely decomposed into a convex combination of pure quantum states $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ also known as an unraveling. In some sense, this degree of freedom in the decomposition of ρ “grows” as the purity of ρ decreases, and this can be utilized to reduce entanglement.² However, this beautiful idea brings with it significant implementation challenges. In an n -qubit system, it is not obvious how to choose a suitable entanglement-based cost function that both lends itself to optimization and connects to meaningful prop-

erties of a classical simulator, such as runtime or simulation accuracy. This is exacerbated by the vast degree of freedom in the mixed-state decompositions. Prior works have predominantly dealt with this complexity either by simplifying the system [42–45] or resorting to heuristic optimization algorithms [45–47]. These simplifications involve restricting consideration to specific noise models combined with simpler systems, such as translationally invariant systems or random circuits. The alternative approach also suffers from significant drawbacks, as heuristic optimization algorithms such as gradient descent can be slow and lack a proof showing that local minima can be avoided.

Our work turns the central promise of MPS-based unraveling simulators into a rigorous local optimization principle: minimizing the entanglement between the target qubit for the noise channel and the rest of the system. While Ref. [45] identified this objective and approached it using variational heuristics, we tackle the problem analytically. We make the crucial observation that an n -qubit mixed state generated in this way can be mapped, via an isometry, to a two-qubit mixed state where, by utilizing Ref. [48], we can compute closed-form solutions that easily map back to the provably entanglement-optimal unraveling of the n -qubit system. With this approach, we achieve two highly desirable and competing goals. First, and most crucially, we attain a highly flexible simulation algorithm where we do not need to compromise the generality of the noise model or impose restrictions on the class of circuits (except one-dimensional spatial locality). Second, we attain closed-form analytical solutions that not only make performance improvements upon the prior art but also lend themselves to analytical investigation that can aid future progress on this topic. We demonstrate this by using the conceptual insights gained from our analysis to connect to and generalize previous results for random circuits.

Our work makes a conceptually enlightening connection to other important works on this topic. Our entanglement-optimal mixed-state decompositions depend not only on the single-qubit noise channel and its target location but also on the state of the n -qubit system immediately prior to the action of the noise channel. We show that when this state is fixed to be Haar random, then our unraveling reduces to that of prior works [44,45]. These works focus on random circuits and specific single-qubit noise models and claim optimality (with respect to the Rényi 2-quasientropy) in this setting. Additionally, we present another conceptually meaningful cost function that is minimized by our unraveling applied to Haar random states. In a precise sense, this new cost function minimizes the unitarity of the Kraus decomposition, favoring Kraus operators that act like rank-1 operators that disentangle the target qubit from the rest of the system. This generalizes fixed unraveling strategies for random circuits to arbitrary single-qubit noise while unifying two conceptually very distinct approaches.

Our final significant contribution is to provide circuit and bond dimension-specific upper bounds on the trace distance error between the final state of the true system and that of any given MPS-based unraveling simulator. Our error bounds are not generic. Rather, they are sensitive to the MPS-based unraveling simulators and apply to the previous body of work on the topic [42–47]. This contribution addresses a

¹For example, Ref. [39] has tackled the nontrivial task of classically simulating stabilizer circuits in the presence of *low levels* of single-qubit coherent noise.

²An extreme but illustrative example is to note that the maximally mixed state on two qubits can be decomposed as a uniform mixture of four Bell states (each of which has maximal entanglement); however, it can also be decomposed as a uniform mixture of computational basis states (each of which has zero entanglement).

shortcoming of all MPS-based unraveling simulators and may be of potential interest beyond this setting.

This work is structured as follows. In Sec. II, we introduce key theoretical background covering unravelings, entanglement measures, and matrix product states. In Sec. III, we focus on performance guarantees. There, we characterize a broad class of simulators, state our theorem bounding the errors committed by such algorithms, and connect this to bounds on errors in sampling and expectation estimate type classical simulation tasks. In Sec. IV, we present our locally entanglement-optimal unraveling algorithm and compare it to relevant prior work. In Sec. V, we consider our unraveling in restricted settings, such as noisy random circuits and Lindbladian systems. There, we also discuss our “least unitary” unraveling. In Sec. VI, we present numerical results comparing our MPS-based unraveling simulator to other similar work. We also touch on conditions for efficient classical simulability and discuss other simulation methods (including other tensor network methods and Pauli propagation methods) in comparison to MPS-based unraveling simulators. In Sec. VII, we discuss some limitations of our work and conclude in Sec. VIII.

II. PRELIMINARIES

In this section, we recap a few notions that help understand the principles of quantum trajectory unraveling. We introduce the tool set that is necessary to understand how to unravel open system dynamics into trajectories.

A. Mixed-state decompositions

A quantum state which is not pure can always be written as a convex combination of normalized pure states

$$\rho = \sum_{i=1}^r p_i |\phi_i\rangle\langle\phi_i|. \quad (1)$$

We call this an ensemble decomposition of the state ρ . A familiar example is the eigendecomposition, which also satisfies the orthonormality condition $\langle\phi_i|\phi_j\rangle = \delta_{i,j}$. However, a richer family of decompositions can be found when an orthogonality condition is not imposed. These decompositions are related by a unitary. Given the eigendecomposition $\{\lambda_i, |v_i\rangle\}_{i=1}^r$ of ρ , by choosing any $r' \in \{r, r+1, \dots\}$ and unitary $U \in \mathbb{U}(r')$, one can construct another decomposition $\{p_{i'}, |\phi_{i'}\rangle\}_{i'=1}^{r'}$ of ρ defined by the equation

$$\sqrt{p_{i'}} |\phi_{i'}\rangle = \sum_{i=1}^r U_{i,i'} \sqrt{\lambda_i} |v_i\rangle \quad \forall i' \in \{1, \dots, r'\}. \quad (2)$$

This also implies that any two decompositions of ρ are also related by a unitary in a similar way.

B. Quantum channel descriptions

Any completely positive map (also called a quantum channel) has a Kraus representation. That is a set of operators $\{K_i\}_{i=1}^r$ such that

$$\mathcal{N} : \rho \mapsto \sum_{i=1}^r K_i \rho K_i^\dagger. \quad (3)$$

This is trace preserving if and only if $\sum_{i=1}^r K_i^\dagger K_i = \mathbb{1}$. Such a Kraus decomposition induces an ensemble decomposition of the output state. Say we apply a quantum channel $\mathcal{N}$ to some state vector $|\psi\rangle$. Then the outcome is given by

$$\mathcal{N}[|\psi\rangle\langle\psi|] = \sum_{i=1}^r K_i |\psi\rangle\langle\psi| K_i^\dagger = \sum_{i=1}^r p_i |\phi_i\rangle\langle\phi_i|, \quad (4)$$

where $\{p_i\}_i$ forms a valid probability distribution with $p_i = \|K_i |\psi\rangle\|_2^2 = \langle\psi| K_i^\dagger K_i |\psi\rangle$ and the normalized state vectors $|\phi_i\rangle$ are $|\phi_i\rangle = p_i^{-1/2} K_i |\psi\rangle$. Those are the pure-state trajectories when unraveling the noisy evolution, and they are not unique. Like ensemble decompositions, Kraus decompositions have a unitary (or rather, isometric) degree of freedom. Given two decompositions $\{K_i\}_{i=1}^r$ and $\{K_{i'}'\}_{i'=1}^{r'}$ of the same channel, there exists a unitary such that

$$K_{i'}' = \sum_{i=1}^r U_{i,i'} K_i \quad \forall i' \in \{1, \dots, r'\}. \quad (5)$$

This freedom, which we will also refer to as the choice of unraveling, is essentially synonymous with the steering of quantum channels.

Definition 1 (Unraveling). Given a channel $\mathcal{N}$ and an initial state vector $|\psi\rangle$, we call $\{p_i, |\phi_i\rangle\}_i$ an unraveling of the evolution of $|\psi\rangle$ under $\mathcal{N}$ when Eq. (4) is satisfied.

We will employ this freedom in the choice of unraveling to aid classical simulation by reducing the entanglement. The choice of Kraus decomposition in Eq. (4) that minimizes entanglement will generally depend on the state $|\psi\rangle$. By slight abuse of nomenclature, we sometimes refer to any specific Kraus decomposition of a channel $\mathcal{N}$ as an unraveling of this channel.

Consider the example of the dephasing channel. It is often described by its mixed-unitary representation

$$\mathcal{N}[\rho] = \left(1 - \frac{p}{2}\right) \rho + \frac{p}{2} Z \rho Z \quad (6)$$

with Kraus operators $\{K_i\}_{i=1}^2 = \{\sqrt{1-p/2} \mathbb{1}, \sqrt{p/2} Z\}$ and noise rate $p \in [0, 1]$. This is a probabilistic application of a unitary gate. For a given input $|\psi\rangle$, the mixed-unitary unraveling results in the two state vectors $|\psi\rangle$ and $Z|\psi\rangle$ with probabilities $1 - \frac{p}{2}$ and $\frac{p}{2}$, respectively. The same channel can also be written in terms of projectors on the computational basis

$$\mathcal{N}[\rho] = (1-p)\rho + p|0\rangle\langle 0| \rho |0\rangle\langle 0| + p|1\rangle\langle 1| \rho |1\rangle\langle 1| \quad (7)$$

with $\{K_i'\}_{i=1}^3 = \{\sqrt{1-p} \mathbb{1}, \sqrt{p} |0\rangle\langle 0|, \sqrt{p} |1\rangle\langle 1|\}$. In this case, the outcome of the noise process is unraveled into $\{|\psi\rangle, |0\rangle, |1\rangle\}$ with probabilities $\{1-p, p|\langle 0|\psi\rangle|^2, p|\langle 1|\psi\rangle|^2\}$. This second decomposition can be obtained from the first through the unitary

$$\begin{pmatrix} K_1' \\ K_2' \\ K_3' \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{2-2p}{2-p}} & 0 & \sqrt{\frac{p}{2-p}} \\ \frac{\sqrt{p}}{\sqrt{2(2-p)}} & \frac{1}{\sqrt{2}} & \sqrt{\frac{1-p}{2-p}} \\ \frac{\sqrt{p}}{\sqrt{2(2-p)}} & -\frac{1}{\sqrt{2}} & \sqrt{\frac{1-p}{2-p}} \end{pmatrix} \begin{pmatrix} K_1 \\ K_2 \\ 0 \end{pmatrix}. \quad (8)$$

Note that the study of the specific unraveling of the dephasing channel into the $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ projectors, when applied

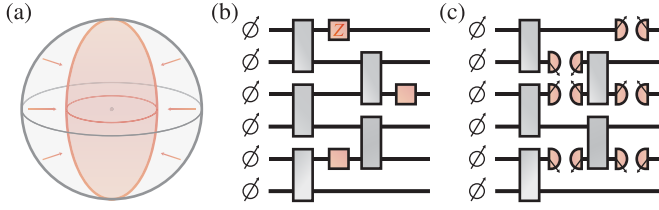


FIG. 1. Effect of single-qubit dephasing noise on the Bloch sphere (a) and two interpretations of the unraveling of single-qubit dephasing noise into pure-state trajectories. Circuit corresponding to the unraveling of one trajectory in the (b) mixed-unitary picture and (c) projective picture. The trajectories (and their entanglement) have very different behaviors in both.

to random circuits, is known as the field of measurement-induced phase transitions (see, e.g., the references within Refs. [43–46]). A representation of the loss of phase information in the computational basis, as well as examples of circuits resulting from the unraveling of single trajectories, are represented in Fig. 1. Any other unraveling (i.e., any Kraus decomposition) can also be interpreted as the coupling of the target qubit to an auxiliary system followed by a measurement thereof.

In the Supplemental Material [49] (Sec. I, including Refs. [50–57]), we discuss more details and a simple example where the unitary degree of freedom of unravelings affects entanglement. We also demonstrate how the relations above are derived and discuss some implications.

C. Entanglement measures

Here we define notions of entanglement that are central to our work. For pure states, given a partition $A : B$ of the system, it can be quantified by the von Neumann entanglement entropy

$$E(|\psi\rangle_{A:B}) = -\text{Tr}[\rho_A \log \rho_A] \quad \text{where} \quad \rho_A = \text{Tr}_B[|\psi\rangle\langle\psi|]. \quad (9)$$

In the case of mixed states, for each decomposition of the form (1), one can define an ensemble-averaged entanglement entropy

$$E_{\text{av}}(\{p_i, |\phi_i\rangle\}) = \sum_i p_i E(|\phi_i\rangle_{A:B}) \quad \text{with} \quad \sum_i p_i |\phi_i\rangle\langle\phi_i| = \rho. \quad (10)$$

These values can vary vastly, and the minimal reachable one over all decompositions is the entanglement of formation of the state, defined as

$$E_{\text{of}}(\rho_{A:B}) = \inf_{\{p_i, |\phi_i\rangle\}: \sum_i p_i |\phi_i\rangle\langle\phi_i| = \rho} \sum_i p_i E(|\phi_i\rangle_{A:B}) \quad (11)$$

as the convex hull of the reduced entropy function, also known as the convex roof extension of the entanglement entropy [53,58]. Note that although one can construct ensemble decompositions with an arbitrary number of elements, it is sufficient to consider decompositions with r^2 elements where r is the rank of ρ [58–60].

D. Matrix product states

A matrix product state (MPS) [24–26,61] is a representation of a quantum state where the coefficient of each element (usually expressed in the computational basis) of the state vector is given by a sequence of matrix products as

$$|\psi\rangle = \sum_{\sigma \in \{0,1\}^n} A_{\sigma_1}^{[1]} A_{\sigma_2}^{[2]} \dots A_{\sigma_n}^{[n]} |\sigma\rangle. \quad (12)$$

When viewing the basis label as an open index, an MPS is thus a linear tensor network where the contraction of the tensor train gives the full state vector. Note that all $A^{[k]}$ are rank-3 tensors $A^{[k]} \in \mathbb{C}^{2 \times d_k \times d_{k+1}}$ except for $A^{[1]}$ and $A^{[n]}$ which are rank-2 tensors.³ In the worst case, the bond dimensions [dimensions of the contracted indices, i.e., dimensions of the matrices in Eq. (12)] can scale exponentially with $d_k \leq 2^{\min(k, n-k)}$. Still, these constructions are commonly used when dealing with states with low entanglement [62,63]. As one is most often given a fixed computational budget, the matrices are truncated. The error committed in this procedure can be bounded using the singular values of the tensors. Indeed, the Schmidt decomposition of the state for any cut $1 : k | k+1 : n$ of the line can be obtained via an adequate singular value decomposition of one tensor. Then the truncation error when retaining the d_k largest singular values (which are the Schmidt coefficients) s_i (from $d_{k,\text{max}} = 2^{\min(k, n-k)}$) is given by

$$\epsilon_k(d_k) = \sum_{i=d_k+1}^{d_{k,\text{max}}} s_i^2. \quad (13)$$

Truncating all bonds for all bipartitions along the line, the 2-norm difference between the true state and the approximation is upper bounded by

$$\| |\psi\rangle - |\psi'\rangle \|_2^2 \leq 2 \sum_{k=1}^{n-1} \epsilon_k(d_k), \quad (14)$$

where $|\psi'\rangle$ denotes the (subnormalized) approximate version of $|\psi\rangle$ where each bond has been truncated (represented graphically in Fig. 2, top right).

For simplicity, let us consider that all bonds are truncated to a maximum bond dimension χ . MPS then allow for the efficient computation in $\text{poly}(n, \chi)$ time of many quantities of interest. These include inner products with other MPS of bond dimension $\leq \chi$, expectation values of MPO observables, index marginalization, and, importantly, computing the Schmidt decomposition across a cut in the MPS. The latter implies the computation of the entanglement entropy across any cut of the linear structure, as the squared singular values over the cut (after adequately shifting the orthogonality center) are the eigenvalues of the reduced density matrix obtained from tracing out the left or right part of the cut.

³Here the word rank is used for the number of indices of a tensor, so a scalar is a rank-0 tensor, a vector is rank 1, and a matrix is rank 2. It does not refer to concepts related to the rank of a matrix (maximal number of linearly independent columns).

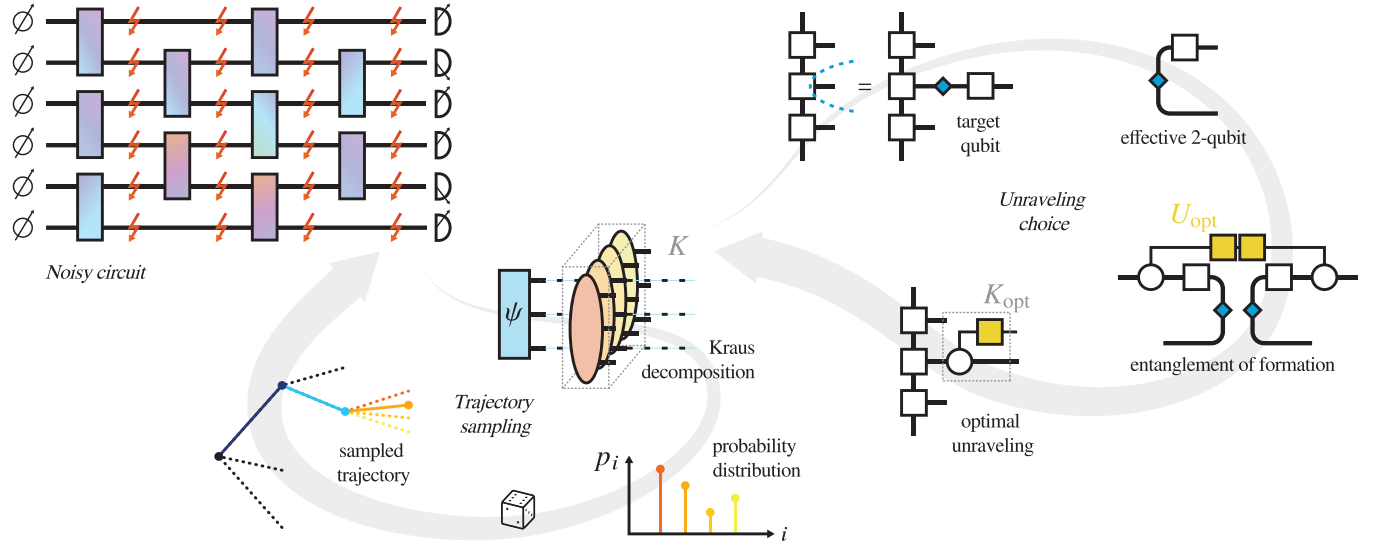


FIG. 2. Graphical representation of the unraveling of quantum trajectories based on matrix product states, for noisy circuits. Consider the setting of one-dimensional noisy circuits, composed of (two-)local unitary gates, followed by single-qubit noise channels (top left). The unitary gates are dealt with by the standard MPS contraction. The noise results in a statistical mixture of pure states. This mixture is then unraveled to sample one trajectory, and the resulting state is truncated to the allowed bond dimension. The trajectory sampling results from computing the probabilities of each path and sampling one of them, thus applying the corresponding Kraus operator (bottom). Before that, the unraveling is chosen. In our case, we do so based on the computation of the entanglement of formation of the affected effective two-qubit state (top right).

III. SIMULATION THROUGH UNRAVELED QUANTUM TRAJECTORY SAMPLING

This section introduces a general class of MPS-based classical simulation algorithms that contains our proposed algorithm as a special case. We call these trajectory sampling simulators, but these connect to long-established techniques [64–70] also referred to as quantum-jump methods or Monte Carlo wave-function methods. Trajectory sampling simulators have a controlled runtime but an uncontrolled simulation error parameter $\epsilon \in [0, 1]$. That is, the runtime of the algorithm can be directly controlled by an initial parameter choice that, in the special case of our algorithm, is the MPS bond dimension. In contrast, the simulation error is not directly controlled. In our algorithm, the error is indirectly dependent on the MPS bond dimension and approaches zero at or before the MPS bond dimension approaches its maximal value of $2^{n/2}$.

Algorithms in this class perform a task we call truncated trajectory sampling to varying degrees of error ϵ . In this section, we state one of our main results. We show that all such simulators can be used to compute an *a posteriori* probabilistic upper bound on this, circuit-dependent, simulation error. Our error bound is based on an estimator computed during the simulation procedure, which has a mean that can vary significantly depending on the specific properties of the simulator. Here, we also relate the output of this class of algorithms to more commonly used notions of classical simulation, such as simulators that sample from the output distribution of a quantum circuit and simulators that estimate expectation values with respect to a given observable.

After defining the concept of truncated trajectory sampling in Sec. III A and describing an abstract class of algorithms performing such a simulation in Sec. III B, we give our bounds

on the errors committed by such algorithms in Sec. III C and show how it results in bounds on errors in simulation tasks like sampling from the Born distribution or estimating expectation values of observables, Sec. III D.

A. Truncated trajectory sampling task

At a high level, trajectory sampling simulators break state evolution into layers and perform the following process at the level of each layer:

- (1) Use an appropriate data structure to classically represent the quantum state inputted into the computation layer.
- (2) Apply the quantum channel associated with the computation layer.
- (3) Decompose the (mixed) state into a convex combination of (usually pure) states.
- (4) Sample a state from the decomposition.
- (5) Approximate the sampled state with a nearby state that has preferable memory and runtime properties, and use this approximation as input to the next computation layer.

This procedure is represented graphically in Fig. 3. Our work focuses on utilizing the MPS data structure as described in point 1 above. Matrix product states are a powerful tool to simulate pure quantum states with little entanglement, in particular in one spatial dimension (1D). This provides a natural data structure for the trajectory state vectors. As discussed in Sec. II, mixed states provide an isometric degree of freedom in the choice of unraveling or decomposition in point 3 above. Since, typically, states with low entanglement can be approximated more accurately by a low bond dimension MPS, we employ this isometric degree of freedom to choose unravelings that lead to lower entanglement states in point 4 (and hence also incur small approximation errors when approximated by a low bond dimension MPS in point 5). We

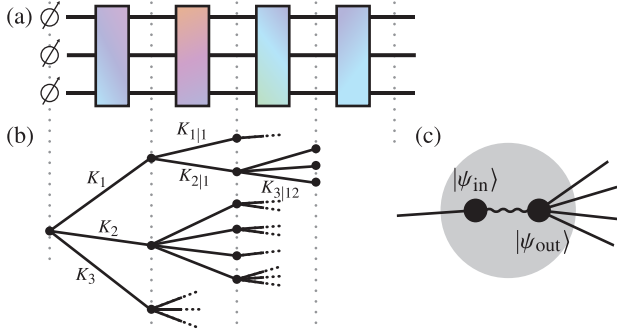


FIG. 3. (a) Simple layered process of quantum channels (represented as colored boxes, although they may be nonunitary) and (b) the corresponding Kraus tree, with Kraus elements of each edge possibly depending on the history of the path. (c) Each node has an input and an output state, related by some approximation procedure (represented as a wiggly line).

now properly define what we mean by an algorithm performing truncated trajectory sampling.

Let $\mathcal{D}$ be a classical description of a noisy quantum circuit on n qubits consisting of

- (i) a qubit ordering: a linear ordering of the n -qubits by labeling them from 1 to n ;
- (ii) an initial state: an initial state vector $|\psi^{(0)}\rangle$ given as an MPS (e.g., $|0\rangle^{\otimes n}$) with respect to the given qubit ordering;
- (iii) and a string of transformations: a sequence of efficiently representable and applicable quantum channels on n qubits (for instance, each comprised of one or two qubit unitary gates on disjoint sets of qubits composed with single-qubit noise channels).

We use ρ to denote the final quantum state associated with $\mathcal{D}$. Given $\chi \in \mathbb{N}$ and a classical description $\mathcal{D}$, the goal is to sample from an ensemble of MPS states with bond dimension at most χ such that the expected state with respect to this ensemble approximates ρ in trace distance.

To be more precise, we first define a key concept that we then use to define truncated trajectory sampling.

Definition 2 ((ϵ, χ) -convex MPS). Let ρ be an n -qubit quantum state with a specified qubit ordering. Let $\epsilon \geq 0$ and $\chi \in \mathbb{N}$. Then we say that ρ is an (ϵ, χ) -convex MPS with respect to its qubit ordering iff there exists a probability distribution $\{q_i\}_i$ and a corresponding set of pure quantum states $\{|\psi_i\rangle\}_i$ such that

- (1) for all i , $|\psi_i\rangle$ admits a classical description in the form of an MPS with bond dimension $\leq \chi$ with respect to the specified qubit ordering, and
- (2) the ensemble $\{q_i, |\psi_i\rangle\}_i$ satisfies

$$\left\| \rho - \sum_i q_i |\psi_i\rangle\langle\psi_i| \right\|_{\text{Tr}} \leq \epsilon. \quad (15)$$

We call any such ensemble, $\{q_i, |\psi_i\rangle\}_i$, an (ϵ, χ) -convex MPS witness for ρ .

Note that each MPS representation has an inherent qubit ordering, and different choices of qubit ordering are possible. Further, the minimal bond dimension of the MPS representations of the same state with respect to different qubit orderings can be extremely different. For this reason, we specify the

qubit ordering of the MPS representations. In many settings, however, a natural ordering is already imposed by the system's geometry. That is the case when studying one-dimensional quantum circuits or open systems with nearest-neighbor interactions on a line.

Definition 3 (Truncated Trajectory Sampling). Given a noisy circuit description $\mathcal{D}$ on n qubits (as specified above), let ρ be the final quantum state. Given a maximal bond dimension χ , we say that an algorithm performs ϵ -TTS($\mathcal{D}, \chi$) iff ρ is an (ϵ, χ) -convex MPS, and for some $\{q_i, |\psi_i\rangle\}_i$ that is an (ϵ, χ) -convex MPS witness for ρ , with probability q_i , the algorithm returns a classical description of a state vector $|\psi_i\rangle$ in the form of an MPS with bond dimension at most χ .

Below we define a general class of algorithms that perform truncated trajectory sampling to varying degrees of accuracy $\epsilon \in [0, 1]$. We discuss how these algorithms are constructed and how to bound ϵ for any given algorithm, χ , and $\mathcal{D}$.

B. Class of truncated trajectory sampling algorithms

We now describe a general class of classical simulation algorithms that perform truncated trajectory sampling. The simulators in this class use three key subprocedures, which we refer to as the layering, Kraus tree, and the approximation subprocedures. Although other implementation details and inputs may vary, conditioned on the details of these subprocedures, these algorithms all perform an equivalent high-level procedure, which we will describe shortly.

(1) Layering subprocedure: This algorithm specifies how the circuit appearing in the circuit description $\mathcal{D}$ is to be broken up into a string of transformations (unitary gates or noisy processes) that will be applied, in order, to a quantum state representation.

(2) Kraus tree subprocedure: Each time a transformation is applied, this can be represented by one of (potentially) many Kraus decompositions.⁴ We will consider a simulation procedure where a Kraus element is sampled from a particular Kraus decomposition associated with each noisy transformation. After the first Kraus operator is sampled from the decomposition of the first noisy transformation, subsequent Kraus decompositions may depend on the sampled Kraus element. This defines a tree structure where the following apply:

- (i) Each node is associated with an input and output state (for now, these can be treated identically; this distinction will become relevant when we introduce the approximation subprocedure), with the initial state of the system equal to the input state of the root node.
- (ii) Each edge from a node to one of its children is associated with exactly one Kraus element, where the choice of Kraus decomposition defines the set of all such edges. The latter can depend on the unique path from the root to the parent node. The update of the state along an edge follows Eq. (4). Given the output state of the parent node, $|\psi\rangle$, and a Kraus element K_i associated with an edge to a child node, the probability of that edge is $p_i = \|K_i |\psi\rangle\|_2$ and a state

⁴We note that unitary transformations are represented as a single-element Kraus decomposition.

$p_i^{-1/2} K |\psi\rangle$ is associated with the input state of the child node.

We refer to the tree labeled with Kraus operators, input and output states, and transition probabilities as the Kraus tree. Although the number of nodes in this tree may be growing exponentially in the circuit gate count, it never needs to be explicitly constructed and can be fully specified by an algorithmic procedure that, given a node and the output state associated with that node, determines how to choose the Kraus decomposition for the next transformation to be applied. We refer to this as the Kraus tree subprocedure. As a simple example, one can consider Kraus tree subprocedures, where the Kraus decomposition is independent of the node and state, depending only on the next transformation to be applied, with each type of transformation in the circuit assigned a fixed Kraus decomposition.

(3) Approximation subprocedure: This subprocedure determines the mapping from the input state to the output state of each node of the Kraus tree. This subprocedure is intended to be used to simplify a state's representation by employing a suitable approximation. In our setting, where an MPS data structure is employed, a natural example is an approximation subprocedure that truncates the input state to one with a maximum permitted bond dimension.

We consider the set of all algorithms that specify a layering, Kraus tree, and approximation subprocedure and implement a simulation by sampling a complete path from the root to a leaf of the Kraus tree based on the transition probabilities assigned to the edges. These algorithms ultimately return the output state at the leaf of the sampled branch of the Kraus tree. We call this one (truncated) *trajectory*.

We observe that if the approximation subprocedure is chosen to be the trivial identity map (i.e., the state is always represented exactly), then the expectation value over trajectories is exactly the (mixed) state ρ produced immediately prior to measurement. However, if nontrivial approximations are made at any point in the Kraus tree, the output state ρ' will generally differ from ρ .

In the Supplemental Material [49] (Sec. II.B, including Ref. [71]), we show how to construct a probabilistic upper bound for the total trace distance between ρ and ρ' using only a polynomial number of samples from branches of the Kraus tree. At a high level, we sample a branch from the Kraus tree and consider the sum of trace distances between the input and output states of each node along the branch. We show that the expectation value of this random variable (and related constructions) acts as an upper bound on the trace distance between ρ and ρ' . This permits us to estimate an upper bound of the expectation value as the mean of a number of samples of these random variables. To produce a probabilistic upper bound, we add a small buffer term to this mean to ensure that, with high probability, we are overestimating the expectation. This buffer term can be made smaller by increasing the number of samples used.

In this work, we focus on the setting of the MPS data structure where the approximation subprocedure is used to impose an upper bound on the bond dimension of the output states at all nodes. With this additional restriction, the set of algorithms consistent with the above description perform truncated trajectory sampling to varying degrees of accuracy

$\epsilon \in [0, 1]$. Theorem 1 adapts our probabilistic upper bound on the accuracy ϵ to the truncated trajectory sampling setting. We emphasize that although our accuracy upper bounds apply to a broad class of algorithms, the size of the upper bound is strongly dependent on the specific details of the algorithm. Hence, sharper algorithms can achieve a smaller upper bound on the trace distance between ρ and ρ' . In Sec. IV B, we present our algorithm, a special case of the set of algorithms described here. This provides a specific example of how unraveling can be used to reduce the accuracy parameter ϵ associated with truncated trajectory sampling. But first, we introduce some more commonly used notions of classical simulation and discuss their connection to truncated trajectory sampling.

C. Upper bounds on the simulation errors

In this section, we present our main contribution to the rigorous analysis of the errors committed in such quantum trajectory protocols (Theorem 1). Let us start by giving labels to the key quantities, and then state the main result.

Each trajectory $|\hat{\psi}_i\rangle$ is drawn from the decomposition of some state $\sigma = \sum_i q_i |\psi_i\rangle\langle\psi_i|$. This equates to saying $\mathbb{E}[|\hat{\psi}_i\rangle\langle\hat{\psi}_i|] = \sigma$. In the case of no approximations, one has $\sigma = \rho$, the exact state resulting from the noisy evolution, while $\|\rho - \sigma\|_{\text{Tr}} \geq 0$ otherwise. And although the error is hard to compute exactly efficiently, given N trajectories sampled independently, there is an estimator $\hat{\epsilon}_N$ that can be computed in $\text{poly}(n, \chi, N)$ time. This is associated with the accuracy of trajectory sampling simulation algorithms.

For a trajectory $|\hat{\psi}_i\rangle$ with $\epsilon_{\text{approx}}(\hat{\psi}_i, \ell)$ the approximation error at layer ℓ , we can define a function of the errors at each layer bounding the cumulated error of the trajectory. This (bounded) total error of the trajectory is the sum of errors $\epsilon_{\text{bound}}(\hat{\psi}_i) = \sum_{\ell=1}^L \epsilon_{\text{approx}}(\hat{\psi}_i, \ell)$ if this sum is smaller than some threshold $\epsilon_{\text{max}} - 2$ and is set to ϵ_{max} otherwise. With this, averaging over the sampled trajectories, one can compute an estimator of the error on the simulated state

$$\hat{\epsilon}_N = \frac{1}{N} \sum_{i=1}^N \epsilon_{\text{bound}}(\hat{\psi}_i). \quad (16)$$

Here, ϵ_{max} is the maximum tolerated approximation error in a single trajectory, which can be freely chosen between 2 and $2L$. This allows us to state our first main result (selecting $\epsilon_{\text{max}} = 2$).

Theorem 1 (Truncated trajectory sampling error). Given a noisy circuit description $\mathcal{D}$ with a final state ρ and a maximal bond dimension $\chi \in \mathbb{N}$, there exist an $\epsilon \geq 0$ such that with runtime $\text{poly}(n, \chi, N)$ the trajectory sampling algorithm returns samples from an (ϵ, χ) -convex MPS witness for ρ . Furthermore, for any $\delta > 0$, with probability at least $1 - \delta$,

$$\epsilon \leq \hat{\epsilon}_N + \sqrt{\frac{2}{N} \log \left(\frac{1}{\delta} \right)}. \quad (17)$$

We note that when $\chi \geq 2^{n/2}$, for all $n = 1, 2, \dots$, $\hat{\epsilon}_N = 0$.

This bound is composed of a systematic error and a statistical error. First, the systematic error is due to the truncations. Even in the hypothetical case where one would compute all (exponentially many) possible trajectories, each would

be only an approximate representation of the real trajectories. The difference between the real noisy state ρ and the mixture of *all* unraveled trajectories $\sigma = \mathbb{E}[|\hat{\psi}_i\rangle\langle\hat{\psi}_i|]$ is the systematic error of the simulation. This error can be bounded by $\lim_{N \rightarrow \infty} \hat{\epsilon}_N$. The resulting accuracy can be improved by reducing the approximation errors of each trajectory, i.e., by increasing the bond dimension of the MPS. Second, there is the uncertainty due to finite sampling, the statistical error. The systematic error is constructed from the sampled trajectories, as a cumulation of the errors in each. The exact computation would require computing all trajectories, which can not be done efficiently. As a result, one needs to estimate the systematic error itself. The bound $\hat{\epsilon}_N$ obtained by averaging the errors of the sampled trajectories holds with high probability when adding the “statistical buffer” $\sqrt{[\epsilon_{\max}^2/(2N)] \log(1/\delta)}$. The tightness of the bound (or the probability of failure) can be reduced by increasing the number of samples. When using MPS, the approximation error $\epsilon_{\text{approx}}(\hat{\psi}_i, \ell)$ of the trajectory $|\hat{\psi}_i\rangle$ at depth ℓ due to the truncations can be computed using that $\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_{\text{Tr}} = \sqrt{1 - |\langle\psi|\phi\rangle|^2}$ at the cost of having to also compute the nontruncated evolution, or using the bound from Eq. (14), resulting in

$$\epsilon_{\text{approx}}(\hat{\psi}_i, \ell) \leq 4 \sqrt{2 \sum_k \epsilon_k(\chi)}. \quad (18)$$

The full proof of Theorem 1 and details about the MPS-based errors can be found in the Supplemental Material [49] (Sec. II).

Note that, in general, the bound from Theorem 1 is not tight. For some noise models of interest, it can be made tighter using concentration arguments. Then, the error contribution of each layer is scaled by a factor decaying exponentially with the depth of the circuit. In this case, we obtain a new bound (omitting the bounding with $\epsilon_{\max} - 2$ for simplicity)

$$\hat{\epsilon}_N = \frac{1}{N} \sum_{i=1}^N \left(\sum_{\ell=1}^L \alpha^{L-\ell} \epsilon_{\text{approx}}(\hat{\psi}_i, \ell) \right) \quad (19)$$

for some noise-dependent factor $\alpha < 1$. See the Supplemental Material [49] (Sec. II.D) for a presentation of the argument and examples of specific values of α .

We are now ready to introduce some more commonly used notions of classical simulation and demonstrate how our results on truncated trajectory sampling result in bounds on simulation errors.

D. Target classical simulation tasks

In this section, we define two commonly used notions of classical simulation, which we call output distribution sampling and observable estimation, and show their connection to truncated trajectory sampling.

Given

(1) a classical description $\mathcal{D}$ of a noisy quantum circuit on n qubits,

(2) an accuracy parameter $\epsilon > 0$ and a failure probability $\delta > 0$,

the goal is to perform the following tasks with probability at least $1 - \delta$, over the randomness of the algorithm:

(i) Output distribution sampling: Sample from any probability distribution within the ϵ -ball (in total variation distance) of the outcome distribution

$$P(x | \rho) := \langle x | \rho | x \rangle \quad (20)$$

associated with measuring the final state ρ in the computational basis $\{|x\rangle\}_x$.

(ii) Observable estimation: Given an observable O in the form of an MPO of bond dimension upper bounded by $\text{poly}(\log n)$, estimate $\langle O \rangle = \text{Tr}[O\rho]$ up to $\pm\epsilon$ in additive error.

We denote an instance of each of these problems using $\text{SAMP}(\mathcal{D}, \epsilon, \delta)$ and $\text{OESTIM}(\mathcal{D}, O, \epsilon, \delta)$, respectively.

In general,⁵ we cannot expect to find an efficient classical algorithm for these simulation tasks that works for all instances. This is because solving the output distribution sampling task in $\text{poly}(n, \epsilon^{-1}, \log \delta^{-1})$ time is as hard as SAMPBQP , the class of sampling problems efficiently solvable by a (noiseless) universal quantum computer. It is also easy to show that solving the observable estimation task in $\text{poly}(n)$ time for any $\epsilon + \delta < 1/6$ and $O = |0\rangle\langle 0| \otimes I^{\otimes n-1}$ is BQP-hard .⁶ Note that the estimation of Born probabilities and marginals (the probabilities of outcomes in the computational basis) is a special case of observable estimation. For a detailed discussion of how these notions of classical simulation relate to each other, see Ref. [8].

Before introducing our algorithm in more detail, we outline how algorithms for truncated trajectory sampling can be used to perform output distribution sampling and observable estimation.

Lemma 1 (Output distribution sampling from trajectory sampling). An algorithm performing ϵ -TTS($\mathcal{D}, \chi$) with probability $1 - \delta$ can, with $\text{poly}(n, \chi)$ -time classical processing, produce a solution to $\text{SAMP}(\mathcal{D}, \frac{1}{2}\epsilon, \delta)$.

Lemma 2 (Expectation value estimation from trajectory sampling). For any $\eta > 0$, $N = 2\|O\|_{\text{op}}^2 \eta^{-2} \log(2\delta'^{-1})$ independent outputs of an algorithm performing ϵ -TTS($\mathcal{D}, \chi$) with probability $1 - \delta$ can, with $\text{poly}(n, \chi)$ -time classical processing, produce a solution to

- (i) $\text{OESTIM}(\mathcal{D}, O, \epsilon', \delta')$ where $\epsilon' = [\epsilon + 2\delta]\|O\|_{\text{op}} + \eta$ or
- (ii) $\text{OESTIM}(\mathcal{D}, O, \epsilon', \delta + \delta')$ where $\epsilon' = \epsilon\|O\|_{\text{op}} + \eta$.

Lemma 1 derives from the fact that each trajectory, independent of the choice of unraveling, is a pure state drawn from a valid ensemble decomposition of the approximate output state of the circuit. With this, the measurement statistics averaged over the elements of the decomposition match those of the mixed state. Leveraging the fact that any bound on the trace distance between states is a bound on the TV distance between the respective distributions leads to the result above.

For Lemma 2, one also has to take into account that the finite sampling of trajectories does not fully reconstruct the approximate state. Since not all trajectories are sampled but only a subset, one can only guarantee that the unraveled mixed

⁵Without placing specific restrictions on the level and nature of the noise model, this class of circuits includes universal noiseless circuits.

⁶BQP is the class of decision problems efficiently solvable by a (noiseless) universal quantum computer.

state is within a ball around the target state with high probability. The incurred statistical error is reflected in the additional term. The detailed proofs of these statements can be found in the Supplemental Material [49] (Sec. III, including Ref. [72]).

With this, Theorem 1 allows us to state bounds on the errors of the simulation tasks presented above.

Corollary 1 (Output distribution sampling). Given the setting of Theorem 1, N samples from the trajectory sampling algorithm allow for computing N samples from a probability distribution $\mathcal{P}'$ which is close to $\mathcal{P}$,

$$TV(\mathcal{P}, \mathcal{P}') \leq \hat{\epsilon}_N + \sqrt{\frac{\epsilon_{\max}^2}{2N} \log\left(\frac{1}{\delta}\right)}, \quad (21)$$

with probability $1 - \delta$ where $\mathcal{P} := \{P(x|\rho)\}_x$ is the distribution of outcomes generated when the noisy circuit described above is applied to the $|\psi^{(0)}\rangle$ state and all qubits are subsequently measured in the computational basis.

Proof. This is a direct application of Lemma 1 to the result of Theorem 1. ■

Corollary 2 (Expectation value estimation). Given the setting of Theorem 1 and a bounded observable O with $\|O\|_{\text{op}} \leq 1$, N samples from the trajectory sampling algorithm can be used to construct an estimator $\hat{O}$ for the expectation value $\langle O \rangle$ with

$$|\hat{O} - \langle O \rangle| \leq \hat{\epsilon}_N + \sqrt{\frac{\epsilon_{\max}^2}{2N} \log\left(\frac{1}{\delta}\right)} + \sqrt{\frac{2}{N} \log\left(\frac{2}{\delta'}\right)} \quad (22)$$

with probability at least $1 - \delta - \delta'$.

Proof. This is a direct application of Lemma 2 to the result of Theorem 1. ■

We can now focus entirely on how MPS-based quantum trajectories allow us to compute solutions to truncated trajectory sampling tasks. We argue that our algorithm achieves a state-of-the-art performance among unraveling-based simulators of noisy quantum circuits.

IV. OUR ALGORITHM

In this section, we will dive into the details of our algorithm. For this, we review MPS-based trajectory sampling algorithms (as presented in Sec. III B and studied in Refs. [42–47]) in Sec. IV A. We then present our locally entanglement-optimal choice of unraveling and how it is derived in Sec. IV B. In Sec. IV C, we give an overview and comparison of state-of-the-art choices of unraveling for MPS-based trajectory sampling.

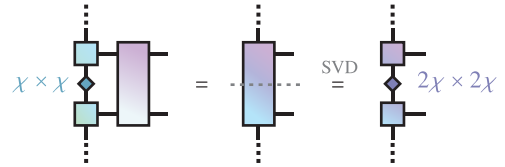
A. Details of MPS-based trajectory sampling

The central idea of quantum trajectory unraveling is to interpret open system evolution as a stochastic process over (nonunitary) normalized pure-state evolutions. For each layer, starting from a state vector $|\psi\rangle$ the effect of the computational layer Λ is unraveled by specifying a particular Kraus representation $\{K_i\}_i$ of the channel and obtaining the induced ensemble decomposition as per Eq. (4). Put into words, Eq. (4) is a statement that the output state is a probabilistic average over the distribution $\mathbf{p} = \{p_i\}_i$ of normalized state vectors $\{|\phi_i\rangle = p_i^{-1/2} K_i |\psi\rangle\}_i$. This procedure is represented graphi-

cally in Fig. 2 (bottom left). Note that one has the freedom to choose the Kraus decomposition of the channel. This different choice of unraveling is synonymous with a different U in Eq. (5), and it is precisely this freedom we will leverage to optimize the algorithm in Sec. IV B.

We now expose how to perform the update efficiently when using MPS and few-qubit unitaries, followed by single-qubit noise channels. From here on out, we consider 1D architectures of geometrically local, few-qubit, unitary gates composed with local (single-qubit) noise channels, like for instance the brickworklike circuit structure shown in Fig. 2 (top left). This setting could be made more general, but it already contains all the key elements to set the stage for our locally entanglement-optimal unraveling. Let us describe each block individually.

Given a pure state represented as an MPS of bond dimension χ , the action of a unitary gate is obtained by contracting the (four-legged) tensor with the affected MPS tensors and then recovering the MPS structure through singular value decomposition (SVD):



One layer of two-qubit, nearest-neighbor, unitaries acting on disjoint sets of qubits will result in a new MPS with increased bond dimension of at most 2χ .

Each single-qubit noise channel is also applied locally. They can be represented by a three-legged tensor $K = \sum_i K_i \otimes |i\rangle$, then Eq. (3) can be formulated as

$$\mathcal{N}[\rho] = K\rho K^\dagger = \text{[Diagram: A circle representing } \rho \text{ with a line entering from the left labeled } K \text{ and a line exiting to the right labeled } K^\dagger \text{]}.$$

After selecting the Kraus decomposition (the unraveling), the probabilities of each path (at most four in our choice of unraveling) are computed

$$p_i = \langle K_i^\dagger K_i \rangle = \langle \psi | K_i^\dagger K_i | \psi \rangle, \quad (23)$$

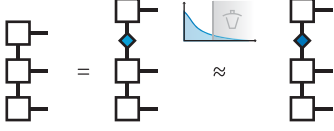
$$= \text{[Diagram: A tensor network for } p_i \text{ showing a unitary gate and two Kraus operators } K_i \text{ and } K_i^\dagger \text{]} = \text{[Diagram: A simplified tensor network for } p_i \text{]} \quad (24)$$

and one of them is sampled. The nonunitary evolution is then implemented by contraction of the respective Kraus operator and renormalization by the corresponding probability:

$$|\psi'\rangle = \frac{1}{\sqrt{p_i}} K_i |\psi\rangle = \text{[Diagram: A tensor network for } |\psi'\rangle \text{ showing a unitary gate and a Kraus operator } K_i \text{]} = \text{[Diagram: A simplified tensor network for } |\psi'\rangle \text{]} = \text{[Diagram: A final tensor network for } |\psi'\rangle \text{]}.$$

This sampling is repeated for each qubit, giving a valid trajectory of the full first layer.

The MPS may then be truncated to remain within the maximum allowed bond dimension. This is done by discarding the smallest singular values of each bond (i.e., the smallest Schmidt coefficients of each bipartition) [62]:



One is thus given a tradeoff between cost and simulation error. The approximation error for the layer, computed efficiently, is added to the cumulated error of simulating the trajectory. Note that in practical implementations, not focused on deterministic calculations, one could variationally look for the closest MPS with the target bond dimension and compute an approximation error from the resulting overlap [41,73].

Repeating this procedure (unitary evolution, noise process, truncation) for every layer returns one of the (at most 4^{nL}) trajectories of the full noisy circuit, as visualized in Fig. 2. An algorithmic description of the MPS-based unraveling of a single noise channel is presented in Algorithm 1.

ALGORITHM 1. Channel unraveling subprocedure.

```

1: INPUT
2: initial state vector  $|\psi\rangle$  as an MPS with bond dimension  $\chi$ 
3: noise channel description  $\mathcal{N}$ 
4: location of qubit affected by noise  $q$ 
5: OUTPUT
6: state  $|\psi'\rangle$  as an MPS with bond dimension  $\chi'$ 
7: PROCEDURE
8:   choose Kraus decomposition:
      $\{K_i\}_i \leftarrow \text{Kraus Unraveling}(\mathcal{N}, |\psi_0\rangle, q)$ 
9:   compute path probabilities:  $\mathbf{p} \leftarrow \{\langle\psi|K_i^\dagger K_i|\psi\rangle\}_i$ 
10:  sample a trajectory:  $i \leftarrow i \sim \mathbf{p}$ 
11:  apply Kraus operator:  $|\psi'\rangle \leftarrow \frac{1}{\sqrt{p_i}} K_i |\psi\rangle$ 

```

Note that the layering process also affects the accuracy and cost of the simulation. Although not central to our work, we outline two methods for interleaving unitary action, noise process, and truncation to illustrate the concept. One layering option is to define the unitary layer as $\lfloor n/2 \rfloor$ two-qubit unitary gates, the noise layer as the tensor product of n single-qubit noise channels, and then a truncation on all bonds

$$\text{Trunc}^{\otimes n-1} \circ \mathcal{N}^{\otimes n} \circ \mathcal{U}^{\otimes \lfloor n/2 \rfloor}. \quad (25)$$

Alternatively, one can choose to perform the unitary, noise, and truncation locally on the two qubits affected by the unitary

$$(\text{Trunc}^{\otimes 1} \circ \mathcal{N}^{\otimes 2} \circ \mathcal{U})^{\otimes \lfloor n/2 \rfloor}. \quad (26)$$

The first strategy will result in lower truncation errors, as the noise on the last qubit may affect the first bond. However, it requires an intermediate representation of the state as an MPS with bond dimension 2χ on every second bond. The first approach does not leverage the effect of the noise on all qubits to reduce the bond dimension before truncating, but never needs to construct an MPS with a bond dimension larger than χ . In general, the intermediate bond dimension will depend on the interplay between the channel layers and the truncation

imposed by the choice of layering. This results in a tradeoff between computational cost (bond dimension) and precision (truncation error).

B. Locally entanglement-optimal unraveling

Like previous works, we aim to minimize the entanglement in the trajectories to reduce the computational cost of simulating those with MPS. In this section, we present the method for obtaining a locally entanglement-optimal unraveling. We will specify the optimization problem we aim to solve and state our result on the optimality of our unraveling in Sec. IV B 1. The remainder of this section will serve to elucidate the steps of the argument by showing how to map the n -qubit problem to a two-qubit problem (Sec. IV B 2), for which one can obtain the optimal decomposition of the mixed state using long-standing results (Sec. IV B 3) and finally how to recover the optimal unitary (the optimal unraveling) from that (Sec. IV B 4).

1. Problem statement and main result

As their name suggests, bipartite measures of entanglement, as defined in Sec. II, are always relative to a partition of the state. One could pick any bipartition, e.g., the one between the target qubit and the next, or the half-system partition. Both of these translate to optimizing one specific bond of the MPS. In this work, we target the minimization of the entanglement between the noisy qubit and the rest of the state (both right and left). This local optimization objective was previously independently proposed in Ref. [45] in the context of a heuristic variational search. Here, we show that this minimization problem admits an exact analytical solution.

Formally, given an initial state vector $|\psi\rangle$ and a Kraus description of a single-qubit noise channel $\{K_i\}_i$, we find the optimal unitary U_{opt} for the noisy state $\rho = \sum_i K_i |\psi\rangle\langle\psi| K_i^\dagger$ such that

$$E_{\text{av}}(\{p_i^{\text{opt}}, |\phi_i^{\text{opt}}\rangle_{\text{target:rest}}\}) = E_{\text{of}}(\rho_{\text{target:rest}})$$

with $|\phi_i^{\text{opt}}\rangle = \frac{1}{\sqrt{p_i^{\text{opt}}}} K_i^{\text{opt}} |\psi\rangle$ and $p_i^{\text{opt}} = \|K_i^{\text{opt}} |\psi\rangle\|_2^2$

for $K_i^{\text{opt}} = \sum_j [U_{\text{opt}}]_{j,i}^* K_j$. (27)

Theorem 2. Given some state vector $|\psi\rangle$, in a representation where computing its Schmidt decomposition is efficient (e.g., an MPS),

$$|\psi\rangle = s |u_0\rangle_{\text{target}} \otimes |v_0\rangle_{\text{rest}} + \sqrt{1-s^2} |u_1\rangle_{\text{target}} \otimes |v_1\rangle_{\text{rest}} \quad (28)$$

and a single-qubit noise channel $\mathcal{N}[\cdot] = \sum_i K_i(\cdot) K_i^\dagger$, there exists an efficient and deterministic algorithm that returns the von Neumann optimal local unraveling. That is, it returns the unitary U_{opt} transforming the Kraus representation [according to Eq. (5)] such that

$$U_{\text{opt}} = \arg \min_U \left\{ \sum_{i=1}^{r'} p'_i E \left(\frac{1}{\sqrt{p'_i}} K'_i |\psi\rangle_{\text{target:rest}} \right) \right\}, \quad (29)$$

where E is the von Neumann entanglement entropy.

This unraveling is optimal in that it minimizes the ensemble-averaged von Neumann entanglement entropy between the noisy qubit and the rest of the state by reaching the entanglement of formation. To build up to this result, we will first show how to reduce the problem stated in Eq. (27) to an effective two-qubit problem. With this observation and some long-standing results on the computation of the entanglement of formation at hand, we show that this problem can be solved exactly, resulting in the optimal choice for the unitary freedom.

2. Reduction to the effective two-qubit problem

We start from an arbitrary pure state on n qubits and compute the Schmidt decomposition with respect to the cut isolating the target noisy qubit from the remaining $n - 1$ qubits, i.e.,

$$|\psi\rangle = s |u_0\rangle_{\text{target}} \otimes |v_0\rangle_{\text{rest}} + \sqrt{1 - s^2} |u_1\rangle_{\text{target}} \otimes |v_1\rangle_{\text{rest}}. \quad (30)$$

Here, we have that $|u_i\rangle \in \mathbb{C}^2$ and $|v_i\rangle \in \mathbb{C}^{2^{n-1}}$ for all i . We are then interested in the action of $\mathcal{N}_{\text{target}} \otimes \mathcal{I}_{\text{rest}}$ on $|\psi\rangle\langle\psi|$, where $\mathcal{I}$ is the identity channel. Since the noisy map only affects the first qubit, the resulting noisy state remains within Hilbert space of $\mathbb{C}^2 \otimes \text{span}\{|v_0\rangle, |v_1\rangle\}$, which is four dimensional. This is a consequence of the Schmidt decomposition having (at most) two nonzero singular values. With this, we can interpret the process as an effective two-qubit evolution. We make it explicit by constructing the isometry which rotates the single-qubit states $\{|0\rangle, |1\rangle\}$ to $\{|v_0\rangle, |v_1\rangle\}$ as

$$V = |v_0\rangle\langle 0| + |v_1\rangle\langle 1|. \quad (31)$$

With that, we can write

$$|\psi\rangle = (\mathbb{1} \otimes V) \underbrace{(s |u_0\rangle \otimes |0\rangle + \sqrt{1 - s^2} |u_1\rangle \otimes |1\rangle)}_{\text{two-qubit state}}. \quad (32)$$

This isometry could be completed to form a unitary, but it is unnecessary and would even be detrimental to our analysis. Note, however, that for any state in $\mathbb{C}^2 \otimes \text{span}\{|v_0\rangle, |v_1\rangle\}$, the operation $(\mathbb{1} \otimes V)(\mathbb{1} \otimes V^\dagger)$ leaves the state invariant. It acts as the identity on the subspace of interest.

Now let us define the channels $\mathcal{V}[\cdot] = V(\cdot)V^\dagger$ and $\mathcal{V}^\dagger[\cdot] = V^\dagger(\cdot)V$ which respectively map from $\{|0\rangle, |1\rangle\}$ to $\{|v_0\rangle, |v_1\rangle\}$ and back. Let us also slightly abuse notation and use $\mathcal{V}$ for $(\mathcal{I}_{\text{target}} \otimes \mathcal{V}_{\text{rest}})$ and $\mathcal{N}$ for $(\mathcal{N}_{\text{target}} \otimes \mathcal{I}_{\text{rest}})$ when dealing with either the two- or n -qubit state. Then, the noise channel and the isometric transformation act nontrivially on disjoint sets of qubits (the noise on the first qubit and the isometry on the rest) and therefore commute. Consequently,

$$\mathcal{V} \circ \mathcal{N} \circ \mathcal{V}^\dagger[|\psi\rangle\langle\psi|] = \mathcal{V} \circ \mathcal{V}^\dagger \circ \mathcal{N}[|\psi\rangle\langle\psi|] = \mathcal{N}[|\psi\rangle\langle\psi|]. \quad (33)$$

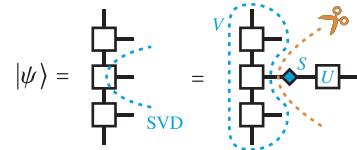
The first equality is due to the commutation relation, and the second is due to the restriction to the four-dimensional subspace. We can now see how this helps us find the optimal decomposition of the n -qubit state when the noise channel acts only on a single qubit. Defining the effective two-qubit state

vector $|\psi^{\text{eff}}\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ as

$$|\psi^{\text{eff}}\rangle = (\mathbb{1} \otimes V^\dagger) |\psi\rangle = s |u_0\rangle \otimes |0\rangle + \sqrt{1 - s^2} |u_1\rangle \otimes |1\rangle, \quad (34)$$

the optimal unraveling of the channel $\mathcal{N}$ for $|\psi^{\text{eff}}\rangle$ is also optimal for $|\psi\rangle$.

Now, V is in principle a $2 \times 2^{n-1}$ matrix, thus not practical. We can, however, take advantage of the matrix product state structure. When starting with an MPS, obtaining the isometry V is almost free: one only has to sever a bond and later stitch it back together. Graphically, on an MPS, the Schmidt decomposition above is given by the SVD of a single tensor between the physical and the bond indices (up to bending some wires):



$$|\psi\rangle = \text{[MPS diagram]} = \text{[MPS diagram with U, S, V tensors]} \quad (35)$$

Here $U = |u_0\rangle\langle 0| + |u_1\rangle\langle 1|$, $S = s |0\rangle\langle 0| + \sqrt{1 - s^2} |1\rangle\langle 1|$, and $V = |v_0\rangle\langle 0| + |v_1\rangle\langle 1|$. The state vector $|\psi^{\text{eff}}\rangle$ is then simply obtained by keeping only the U and S tensors, and bending the index between S and V to act as the effective second qubit



$$|\psi^{\text{eff}}\rangle = (\mathbb{1} \otimes V^\dagger) |\psi\rangle = \text{[Diagram of U and S tensors]} \quad (36)$$

In that sense, the isometry V is already defined by the rest of the matrix product state and does not need to be calculated (nor applied) explicitly.

3. Wootters' optimal decomposition

The optimal unraveling can then be obtained using Wootters' method for computing optimal convex decompositions of two-qubit states [48]. Given a two-qubit state in its eigendecomposition $\rho = \sum_j \lambda_j |\lambda_j\rangle\langle\lambda_j|$, Wootters shows how to find the optimal decomposition $\rho = \sum_i p_i^{\text{opt}} |\phi_i^{\text{opt}}\rangle\langle\phi_i^{\text{opt}}|$ such that

$$E_{\text{av}}(\{p_i^{\text{opt}}, |\phi_i^{\text{opt}}\rangle\}) = E_{\text{OF}}(\rho) \quad (37)$$

as well as the unitary as per Eq. (2) that maps from one to the other. See the Supplemental Material [49] (Sec. VI, including Refs. [74–76]) for an introduction to Wootters' decomposition and a practical (numerically friendly) implementation. The idea of how to use this to find unravelings of channels is straightforward. We begin by rotating the non-noisy part from the pure state to its single-qubit equivalent subspace, obtaining $|\psi^{\text{eff}}\rangle$ as per Eq. (36). Then, we apply the noise to the target qubit, resulting in the mixed state

$$\rho^{\text{eff}} = (\mathcal{N} \otimes \mathcal{I})[|\psi^{\text{eff}}\rangle\langle\psi^{\text{eff}}|]. \quad (38)$$

From this, one can apply Wootters' method to obtain the optimal convex decomposition

$$\rho^{\text{eff}} = \sum_{i=1}^4 p_i^{\text{opt}} |\phi_i^{\text{eff}}\rangle\langle\phi_i^{\text{eff}}|, \quad (39)$$

where $|\phi_i^{\text{eff}}\rangle$ are two-qubit states and this decomposition minimizes the average entanglement entropy. Rotating back gives

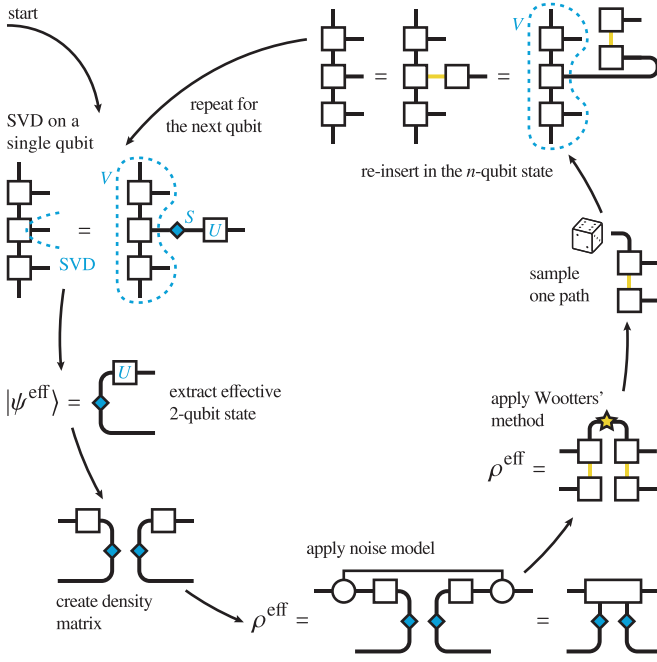


FIG. 4. Graphical representation of the procedure to find the optimal single-qubit unraveling.

an n -qubit state

$$\begin{aligned} \rho &= (\mathbb{1} \otimes V) \rho^{\text{eff}} (\mathbb{1} \otimes V^\dagger) \\ &= \sum_{i=1}^4 p_i^{\text{opt}} (\mathbb{1} \otimes V) |\phi_i^{\text{eff}}\rangle \langle \phi_i^{\text{eff}}| (\mathbb{1} \otimes V^\dagger). \end{aligned} \quad (40)$$

Since the rotation is a local operation with respect to the bipartition of interest (it does not affect the target qubit in any way), it also does not change the entanglement between the two subsystems. We have therefore found an optimal decomposition of the noisy n -qubit state when the noise channel is only applied to a single qubit.

In practice, on MPS, after applying the single-qubit noise model and finding the optimal decomposition, one can sample one state from the decomposition (39) and then contract the nonphysical index back with the rest of the MPS from Eq. (35). This is sampling one state vector $|\phi_i^{\text{eff}}\rangle$ and retrieving the sample of the n -qubit evolution by computing $|\phi_i^{\text{opt}}\rangle = (\mathbb{1} \otimes V) |\phi_i^{\text{eff}}\rangle$. A graphical overview of the complete procedure is shown in Fig. 4.

4. Optimal unraveling unitary

The procedure presented is fully functional and implemented for the numerical simulations shown in Sec. VIA. However, Wootters' method for computing the optimal decomposition of a two-qubit state takes the eigendecomposition as input, not any arbitrary decomposition. That is, it returns a unitary U_{Wootters} such that the decomposition

$$\sqrt{p_i^{\text{opt}}} |\phi_i^{\text{opt}}\rangle = \sum_{j=1} [U_{\text{Wootters}}]_{j,i}^* \sqrt{\lambda_j} |\lambda_j\rangle \quad (41)$$

is optimal. This is no major issue since the eigendecomposition of a two-qubit state can be computed, but it does not fully solve the problem stated in (27) of finding the optimal unitary U_{opt} since, in general, the induced decomposition from Eq. (4) will not be the eigendecomposition. Thus, we want to find the unitary that maps the induced decomposition to the eigendecomposition

$$\begin{aligned} \rho &= \sum_i K_i |\psi\rangle \langle \psi| K_i^\dagger \\ &= \sum_i p_i |\phi_i\rangle \langle \phi_i| \xrightarrow{U_0} \rho = \sum_j \lambda_j |\lambda_j\rangle \langle \lambda_j|. \end{aligned} \quad (42)$$

We will show here that this can also be done efficiently. Doing so, we will obtain the optimal unitary to apply directly to the Kraus decomposition of the noise channel instead of only one valid sample from the optimal distribution, thus allowing us to implement the procedure depicted in Fig. 2.

For the rest of this section, we will use the matrix notation of ensemble decompositions presented in the Supplemental Material [49] (Sec. V). Wootters' method starts with the eigendecomposition of some input state ρ , i.e.,

$$\rho = \sum_i \lambda_i |\lambda_i\rangle \langle \lambda_i| = v v^\dagger \quad \text{with} \quad v = \sum_i \sqrt{\lambda_i} |\lambda_i\rangle \langle i|. \quad (43)$$

However, in general, we are given an arbitrary decomposition of the state

$$\rho = \sum_i p_i |\phi_i\rangle \langle \phi_i| = \phi \phi^\dagger \quad \text{with} \quad \phi = \sum_i \sqrt{p_i} |\phi_i\rangle \langle i|, \quad (44)$$

and want to apply Wootters' method. The question now is, how we can find a mapping from the arbitrary decomposition ϕ to the eigendecomposition v , i.e., some unitary U_0 , such that

$$v = \phi U_0^\dagger. \quad (45)$$

To this end, we identify

$$v v^\dagger = V \Lambda V^\dagger \quad \text{with} \quad V = \sum_{i=1}^r |\lambda_i\rangle \langle i| \quad \text{and} \quad \Lambda = \sum_{i=1}^r \lambda_i |i\rangle \langle i| \quad (46)$$

and find $v = V \sqrt{\Lambda}$. Combining this with the target unitary freedom (45) results in the equation to satisfy

$$V \sqrt{\Lambda} = \phi U_0^\dagger \Leftrightarrow \phi = V \sqrt{\Lambda} U_0, \quad (47)$$

which resembles a singular value decomposition of ϕ . Indeed, we can associate

$$\text{SVD}(\phi) = U_\phi \Sigma_\phi V_\phi^\dagger = \underbrace{V \sqrt{\Lambda}}_v U_0 \Rightarrow \begin{cases} V = U_\phi, \\ \Lambda = \Sigma_\phi^2, \\ U_0 = V_\phi^\dagger. \end{cases} \quad (48)$$

From this, we find that the optimal unitary freedom to apply to the Kraus decomposition of the channel to solve (27) is given by $U_{\text{opt}} = U_{\text{Wootters}} U_0$.

TABLE I. Comparison overview of the state-of-the-art choices of unraveling based on several criteria: System: (restricted vs general): the kind of system the results apply to time-independent translationally invariant Lindbladian systems (TITIL), random circuits (RC), or arbitrary systems. Adaptive (fixed vs adaptive): whether the unraveling is fixed or adaptive (state dependent). Analytical (numerical vs analytical): whether the unraveling optimization (if any) is computed using analytic or numerical methods. Cost function: what measure of quality of the unraveling is used and optimized (for entropy-based measures, the bipartition is specified: MC for middle cut or NQR for noisy qubit vs rest of the system). Noise model: to which single-qubit noise models the results apply (DP: depolarizing, DF: dephasing, AD: amplitude damping, +: applicable even to nonfixed noise models, possibly changing for each qubit and each depth). Note that Ref. [45] appears twice: once for their analytical fixed unraveling (spin-model mapping) and once for their adaptive heuristic strategy (numerical optimization).

Work			System	Adaptive	Analytical	Cost function	Noise model
Vovk and Pichler No. 1	2022	[42]	TITIL	✓	✓	entropy rate (MC)	any
Kolodrubetz	2023	[43]	RC	✗	✗	N/A	DF
Cheng and Ippoliti	2023	[44]	RC	✗	✓	quasientropy	unital & AD
Chen <i>et al.</i>	2024	[45]	RC	✗	✓	quasientropy	DF & AD
Chen <i>et al.</i>	2024	[45]	any	✓	✗	entropy (NQR)	any ⁺
Vovk and Pichler No. 2	2024	[46]	any	✓	✗	entropy (MC)	any ⁺
Daraban <i>et al.</i>	2025	[47]	any	✓	✗	“unitarity”	any ⁺
This work			any	✓	✓	entropy (NQR)	any ⁺

C. Comparing unraveling strategies

In recent years, several works have explored trajectory sampling algorithms, aiming to reduce computational cost and ideally find or improve noise thresholds, resulting in efficient classical simulations. As mentioned above, sampling a trajectory from an evolution under a quantum channel admits a degree of freedom in the decomposition to sample from (the choice of unraveling) and the precise value of the heralded error from Theorem 1 strongly depends on this choice. For the same evolution, some unravelings result in large truncation errors while others produce states that can be well represented as MPS. Some motivating examples of the impact of this choice of unraveling are presented in the Supplemental Material [49] (Sec. I.D). In this section, we present and compare the state-of-the-art choices of unraveling. An overview of the unravelings considered here is presented in Table I.

We say that statements about an unraveling are *general* if they hold for any evolution within the setting considered in Sec. IV A, while they are *system restricted* if further restricting assumptions are made. That would be the case for average-case results over brickwork circuits of Haar random gates, or if the statements only hold for translationally invariant systems. We call an unraveling choice *fixed* if it is independent of the current state of the system. Nonetheless, one is generally not restricted to unraveling the noise in the same way at every application of the noise. An unraveling choice is said to be *adaptive* if it depends on the state to which the noise is applied. We also distinguish methods that deterministically find a cost-function optimal unraveling from those obtained from a numerical optimization without guarantees (e.g., gradient descent). This is particularly relevant for the adaptive unravelings, where the optimization happens at every application of the noise. We also emphasized whether the obtained unraveling is computed by *analytical* methods or by numerical optimization and describe the target cost function used to optimize the unraveling. Finally, we specify the noise models that can be handled by each of the unraveling methods.

Physically motivated or handpicked unravelings have been studied, in particular in the setting of time-independent trans-

lationally invariant Lindbladian systems [42] or noisy random circuits [43]. Different fixed unravelings have been shown numerically to result in significantly different performance, in particular different growths of average entanglement. To provide analytical insights on how to choose a good fixed unraveling, approximate mappings of random circuits to classical spin models have been leveraged [44,45], resulting in fixed unravelings with improved performance for random circuits. However, these methods really optimize a quasientropy which is only indirectly related to the cost of the simulation. On the other hand, adaptive unravelings have first been considered in restricted settings (time-independent translationally invariant Lindbladian systems), where it was shown that to minimize the (continuous) entanglement change rate it is sufficient to consider only two choices (homodyne and number measurements), resulting in an efficient greedy optimization algorithm [42]. Adaptive unravelings can also be computed with numerical methods, like gradient descent on the unitary freedom in Eq. (5). Versions thereof have been implemented to pick adaptive unravelings minimizing different cost functions, like entanglement entropies [44,46] or measures of nonunitarity [47].

Adaptive unravelings offer greater freedom than fixed ones, potentially leading to better performance. In particular, in addition to their analytical fixed unraveling, Ref. [45] also presents an adaptive heuristic strategy that employs the same cost function as our work but uses numerical optimization to compute the unraveling. However, numerical approaches such as gradient descent are often nondeterministic, have the risk of not finding a good minimum for their cost function, and can be significantly slower to implement on each of the numerous calls of the unraveling subprocedure. Thus, it would be valuable from both practical and conceptual perspectives to find analytically motivated choices of adaptive unraveling. Our unraveling achieves precisely that. We provide an efficient adaptive unraveling strategy with analytical guarantees for any single-qubit noise channel. Our unraveling can be computed deterministically in constant time [as opposed to the $O(\chi^2)$ or $O(\chi^3)$ scaling of previous adaptive unravelings]

and optimally disentangles the target qubit from the rest of the state.

V. APPLICATION TO SPECIFIC SYSTEMS

In this section, we discuss the application of trajectory sampling methods and the impact of our results on two frequently considered settings: time-independent Lindbladian systems (Sec. V A) and noisy random circuits (Sec. V B). In particular, we define a measure of unitarity of (Kraus) operators, demonstrate how to compute an unraveling that minimizes this measure, and show that this fixed unraveling is near optimal for random circuits.

A. Time-independent Lindbladian systems

Our method can also be applied to Lindblad equations with single-qubit jump operators. In this section, we discuss how to construct a quantum circuit that matches the language above, starting with Trotterizing the evolution of the open system. Then, the choice of unraveling can be directly applied to the resulting Kraus operators.

The Lindbladian description of the evolution of open quantum systems is given by equations of motion of the form

$$\frac{\partial \rho}{\partial t} = -i[H, \rho] + \sum_j \gamma_j \left(c_j \rho c_j^\dagger - \frac{1}{2} \{c_j^\dagger c_j, \rho\} \right), \quad (49)$$

where H is the Hamiltonian governing the coherent evolution and $\{c_j\}_j$ are the jump operators. We explicitly show in the Supplemental Material [49] (Sec. IX.A) how to compute the noisy quantum circuit corresponding to a Trotterization of the evolution above exactly. For a coherent part composed of two-local nearest-neighbor interactions

$$H = \sum_{i=1}^n h_{i,i+1} \quad \text{with} \quad h_{i,i+1} \in \mathbb{C}^{4 \times 4} \quad \forall i, \quad (50)$$

the gates from the resulting brickwork circuit are simply $U_{\text{coh}} = e^{-iHdt}$. For the incoherent part, the Kraus operators of the corresponding noise channel are obtained by computing the eigendecomposition of the Choi state of the local jump process. The Choi state in this setting is given by a reshaping of the vectorized jump superoperator $e^{L_j dt}$ with

$$L_j = \gamma_j (c_j \otimes c_j^* - \frac{1}{2} c_j^\dagger c_j \otimes \mathbb{1} - \frac{1}{2} \mathbb{1} \otimes c_j^\top c_j^*). \quad (51)$$

This standard procedure returns an exact (and orthogonal) Kraus decomposition of the noise channel resulting from the Trotterized jump process.

We also show in the Supplemental Material [49] (Sec. IX.B) that, for a jump rate γ and Trotterization time step dt , choosing $c = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ results in a dephasing channel with rate $p^{\text{DF}} = 1 - e^{-\gamma dt/2}$ and choosing $c = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ results in an amplitude damping channel with noise rate $\gamma^{\text{AD}} = 1 - e^{-\gamma dt}$. Additionally, the resulting Kraus decomposition coincides with the decomposition typically found in textbooks (labeled in the remainder as mixed-unitary decomposition for dephasing and orthogonal decomposition for amplitude damping, see also Table I).

B. Noisy random circuits

In this section, we look at the example of random quantum circuits. This setting has been heavily studied in the trajectory unraveling literature and the related field of measurement-induced phase transitions. We show in Sec. V B 1 how our unraveling fits in this literature by explaining how previous unravelings can be shown to be near optimal using our arguments based on entanglement of formation. This is done by observing that, when restricting to random circuits, our protocol can be simplified by approximating the effective two-qubit states with a Bell state (maximally entangled state on two qubits). With this, we give analytical footing to the proxy proposed in Ref. [43] and obtain that the unraveling choices presented by Chen *et al.* [45] and Cheng *et al.* [44] happen to be optimal in our sense for the noisy Bell state. We also address one open question from Ref. [44] about the improvement of thresholds when using locally and adaptively optimized unravelings. In the setting of Haar random circuits, such improvement appears to be minimal, as the two strategies converge. In Sec. V B 2, we then relate this optimality statement to the “projectiveness” or “unitarity” of the obtained Kraus decompositions.

1. Near-optimal fixed unraveling for random circuits

This argument is based on the observation that states in random circuits are locally approximately maximally entangled. As presented in Sec. IV B, the problem of finding a good unraveling for single-qubit channels can be reduced to a two-qubit problem. The Schmidt coefficients in Eq. (28) characterize the entanglement of the initial effective two-qubit state.

If the singular values are equal (thus both $1/\sqrt{2}$), then the state is effectively maximally entangled. This case may seem very specific and unlikely, but it is close to what happens for Haar random states. Indeed, it is already well known that, for sufficiently large Haar random states, any subset of qubits is “more or less maximally entangled” with the rest of the system. For a Haar random state, the distribution of eigenvalues of the reduced density matrix (squared singular values) follows a Marchenko-Pastur distribution [77,78]. For states drawn from a 2-design ensemble [79, Example 50] and a single target qubit, we have

$$\mathbb{E}_{|\psi\rangle \sim \nu} [\text{Tr}[\rho_1^2]] = \frac{2 + 2^{n-1}}{2^n + 1} = \frac{1}{2} + O\left(\frac{1}{2^n}\right) \quad (52)$$

where ν is a 2-design ensemble and $\text{Tr}[\rho_1^2] = \lambda_1^2 + \lambda_2^2$ is the purity of the reduced density matrix on that qubit. Now, since $\text{Tr}[\rho_1] = \lambda_1 + \lambda_2 = 1$, we have $\lambda_i = \frac{1}{2} + O(\frac{1}{\sqrt{2^n}})$. Since the eigenvalues of the reduced density matrix are the squared Schmidt coefficients,

$$s_i = \sqrt{\lambda_i} = \frac{1}{\sqrt{2}} + O\left(\frac{1}{\sqrt{2^n}}\right). \quad (53)$$

Interpreting the rest of the $n - 1$ qubits as a single effective qubit as we did in Sec. IV B 2, we obtain

$$|\psi^{\text{eff}}\rangle \approx \frac{1}{\sqrt{2}}(|u_1\rangle \otimes |v_1\rangle + |u_2\rangle \otimes |v_2\rangle), \quad (54)$$

which is a maximally entangled state. Furthermore, in the case of two degenerate Schmidt coefficients $s_1 = s_2 = 1/\sqrt{2}$, the local basis is irrelevant and one can enforce that one of the two Schmidt bases is the computational basis. Using this, we choose the form $|\psi\rangle \approx \frac{1}{\sqrt{2}}(|0\rangle \otimes |w_1\rangle + |1\rangle \otimes |w_2\rangle)$. This can be seen from the singular value decomposition of the matrix of coefficients (from which the Schmidt decomposition is computed), as the matrix of singular values is proportional to the identity, and thus commutes with the two basis-defining unitaries. With this, the case $s_1 = s_2 = 1/\sqrt{2}$ reduces to considering $|\psi^{\text{eff}}\rangle = |\Phi^+\rangle$:

$$|\psi\rangle = \begin{array}{c} \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \end{array} = \frac{1}{\sqrt{2}} \begin{array}{c} \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \end{array} = \frac{1}{\sqrt{2}} \begin{array}{c} \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \end{array} = \begin{array}{c} \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \end{array} \Rightarrow |\psi^{\text{eff}}\rangle = \frac{1}{\sqrt{2}} \begin{array}{c} \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \text{---} \end{array}$$

When dealing with random circuits, one can simplify the unraveling procedure: computing the optimal unraveling of the noise channel of interest for the Bell state once. We call this unraveling Haar optimal. This yields a near-optimal unraveling in a state-independent manner. We show in the Supplemental Material [49] (Sec. VII.A) that the unravelings proposed in Refs. [45] and [44] are optimal in this sense. Our argument offers a different intuition for why the Kraus decompositions found in the literature perform well in unraveling noisy random circuits based on MPS. Additionally, if one lacks information about the state and consequently assumes it to be random, this decomposition is the best estimate. In a sense, they are the best state-agnostic choice of unraveling.

2. Least unitary unraveling

As a consequence of the result in the previous section, we also find the (state-independent) unraveling, which can be interpreted as being as far as possible from a mixed-unitary unraveling. Finding the optimal unraveling for the Bell state (and thus also for random circuits) is equivalent to computing the optimal decomposition of the Choi state of the noise channel [see the Supplemental Material [49] (Sec. VIII) for details]. One can then ask about input-independent properties of unravelings. In particular, single-qubit channels are special in that they cannot create entanglement and only reduce entanglement by projecting out part of the state. Multiqubit channels, on the other hand, can disentangle states in a unitary way. We ask if one can find the Kraus decomposition of a single-qubit channel that is as far from being mixed unitary as possible, i.e., where each Kraus operator is as close to rank-1 as possible.

Corollary 3 (Least unitary unraveling). Let us quantify the unitarity of an operator K by the uniformity of its normalized singular values $\{\sigma_i/\|K\|_F\}$, measured by the Shannon entropy

$$\text{unitarity}(K) = - \sum_{i=1}^d \left(\frac{\sigma_i}{\|K\|_F} \right)^2 \log_2 \left(\frac{\sigma_i}{\|K\|_F} \right)^2. \quad (55)$$

Then, given a single-qubit channel as in Eq. (3), the locally entanglement-optimal unraveling for the Bell state is also the Kraus decomposition which minimizes the average unitarity

of the operators

$$\text{unitarity}_{\text{av}}(\{K_i\}) = \frac{1}{d} \sum_i \|K_i\|_F^2 \text{unitarity}(K_i). \quad (56)$$

The normalization with the Frobenius norm guarantees that we compute the entropy of a valid probability distribution. Note that the unitarity, as defined above, reaches its maximal value of 1 for unitary operators and the minimum 0 for rank-1 operators, e.g., rank-1 projectors. A longer reflection on the choice of definition as well as the full derivation of this result can be found in the Supplemental Material [49] (Sec. VIII). During the writing of this paper, we were made aware of Ref. [47] where a different notion of (non) unitarity is introduced, which is then optimized numerically for an adaptive unraveling.

VI. PERFORMANCE

In this section, we discuss the performance of our unraveling choice relative to other quantum trajectory sampling methods and in the broader context of classical simulation algorithms for noisy quantum dynamics. We first present numerical results for Lindbladian systems and random circuits and compare the numerical performance of our unraveling to alternative unravelings. To convey a sense of the scale of noise rate sufficient for efficient simulability using our algorithm, we discuss known noise thresholds from the trajectory sampling and measurement-induced phase transition literature. Finally, we compare the family of methods based on sampling of trajectories with other simulation methods, including the Pauli path method and other tensor network methods based on matrix product operators and locally purified density operators.

A. Numerical performance comparison

To verify the performance of our algorithm, we present numerical simulations comparing our choice of unraveling to other unraveling algorithms under various settings. In Fig. 5, we considered an open quantum system evolving under a particular Hamiltonian, while in Fig. 6, we present aggregated results based on many randomly generated Hamiltonians. In line with the literature on the simulation of noisy quantum circuits, in Fig. 7 we simulated random one-dimensional quantum circuits with independent local noise and variants thereof.

The various noisy quantum dynamics are simulated using our locally entanglement-optimal unraveling algorithm and three other types of unraveling algorithms that we refer to as the orthogonal, projective, and Haar optimal unravelings. Here, our numerical comparison focuses on three commonly considered noise models (depolarizing, dephasing, and amplitude damping), Table II provides a Kraus decomposition associated with these unravelings [see the Supplemental Material [49] (Sec. I.E) for more details]. We point out that this comparison alone does not fully capture the performance advantages of our unraveling, which applies to arbitrary single-qubit noise models. In contrast, an explicit construction of the Haar optimal unraveling, the more competitive of the

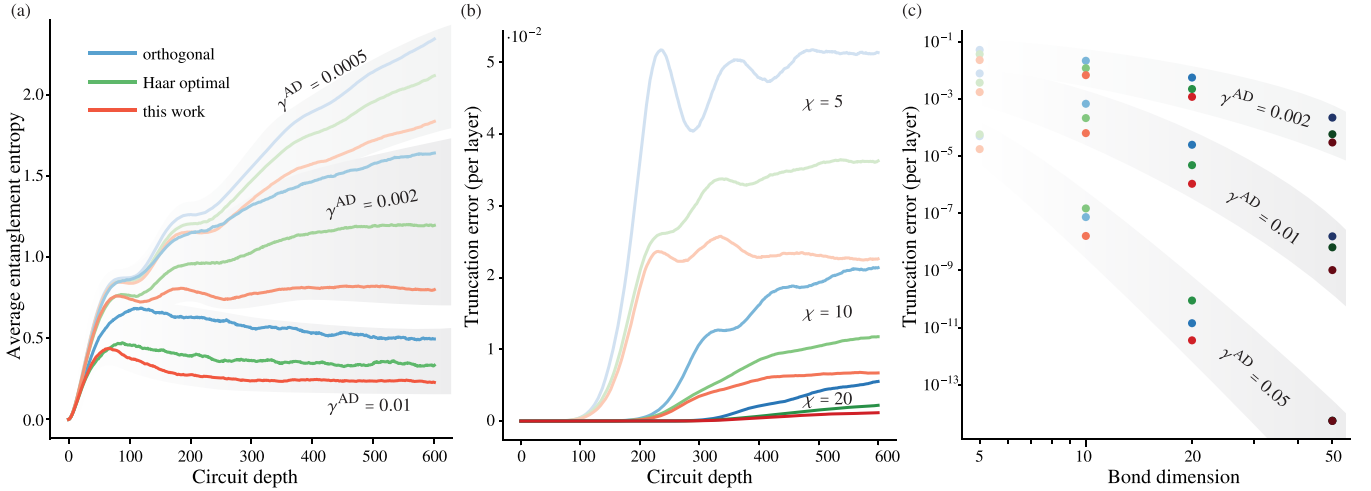


FIG. 5. Simulation of a 16-qubit system evolving under a Heisenberg Hamiltonian $H = \sum_i Y_i Y_{i+1} + 0.35X_i + 0.35Y_i + 0.5Z_i$ (open boundary conditions) with amplitude damping noise. All results are an average of 500 independently sampled trajectories and repeated for the orthogonal and *Haar* optimal fixed unravelings (see Table II) as well as our locally entanglement-optimal unraveling. (a) Average entanglement as a function of circuit depth plotted for various noise rates and unravelings in a simulation with bond dimension $\chi = 50$. (b) Average upper bound on truncation error [Eq. (14)] as a function of circuit depth, plotted for various bond dimensions in a simulation with noise rate $\gamma^{\text{AD}} = 0.002$. (c) Scaling of the truncation error (at depth 600) with bond dimension for various noise rates.

three alternatives, is only known for specific noise channels and lacks this level of generality.

Kraus decompositions can always be constructed from ensemble decompositions of the Choi state of the completely positive and trace-preserving (CPTP) map [80, Sec. 6.4], [81, Sec. 2.2]. The orthogonal Kraus decomposition is defined as the set of operators obtained from the eigendecomposition, which are necessarily orthogonal. In the case of unital noise channels such as the depolarizing and dephasing noise, the eigenoperators are unitary and hence we call these

mixed-unitary unravelings. In the case of amplitude damping noise, the eigenoperators are not unitary. This is the representation typically given in standard expositions of quantum noise for the three noise channels considered here.

We also consider the *Haar* optimal unraveling, obtained by rotating the orthogonal Kraus representation using the Hadamard matrix as the unitary freedom in Eq. (5). This unraveling is derived in Refs. [44,45] via an approximate mapping of random circuits to classical spin models. As discussed in Sec. VB, our locally entanglement-optimal unraveling also

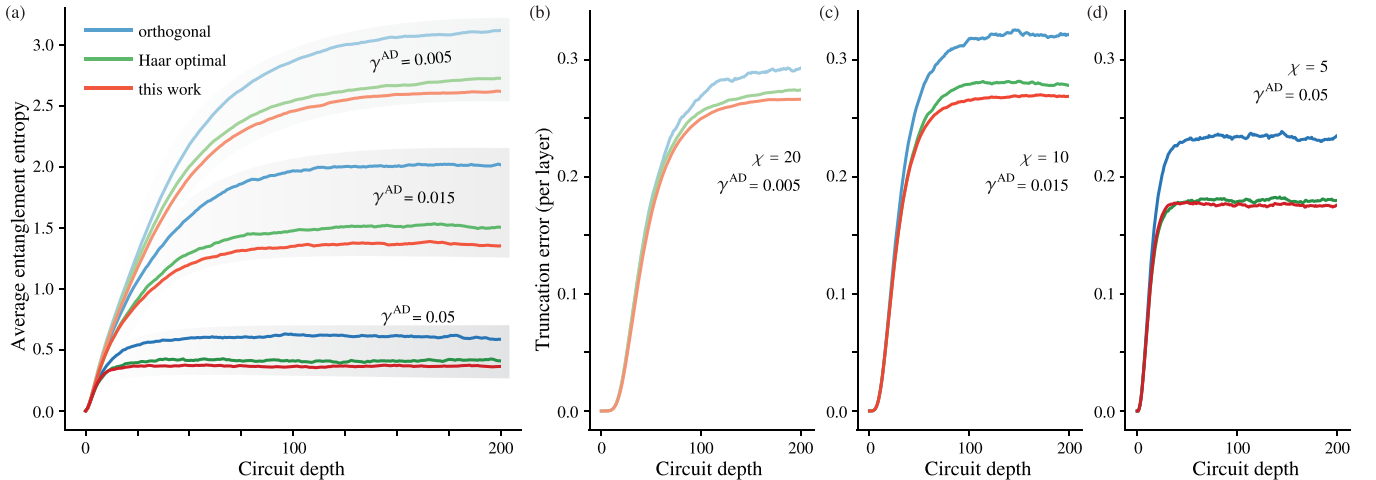


FIG. 6. Simulation of a 12-qubit open quantum system under random translationally invariant and time-independent Lindbladian evolutions with amplitude damping noise. For each random evolution, we consider a brickwork circuit composed of a fixed gate repeated at each site. The gate itself is obtained by a short-time unitary evolution $U = e^{-iHdt}$ of a random Hamiltonian H (a random Hermitian matrix with normally distributed real and imaginary parts of each entry) and a time step $dt = 0.01$. Each trace is the average of 1000 independently computed trajectories truncated to a fixed bond dimension. (a) Average entanglement as a function of circuit depth plotted for various noise rates and unravelings in a simulation with bond dimension $\chi = 20$. (b)–(d) Average upper bound on truncation error [Eq. (14)] as a function of circuit depth, plotted for various combinations of bond dimensions and noise rates.

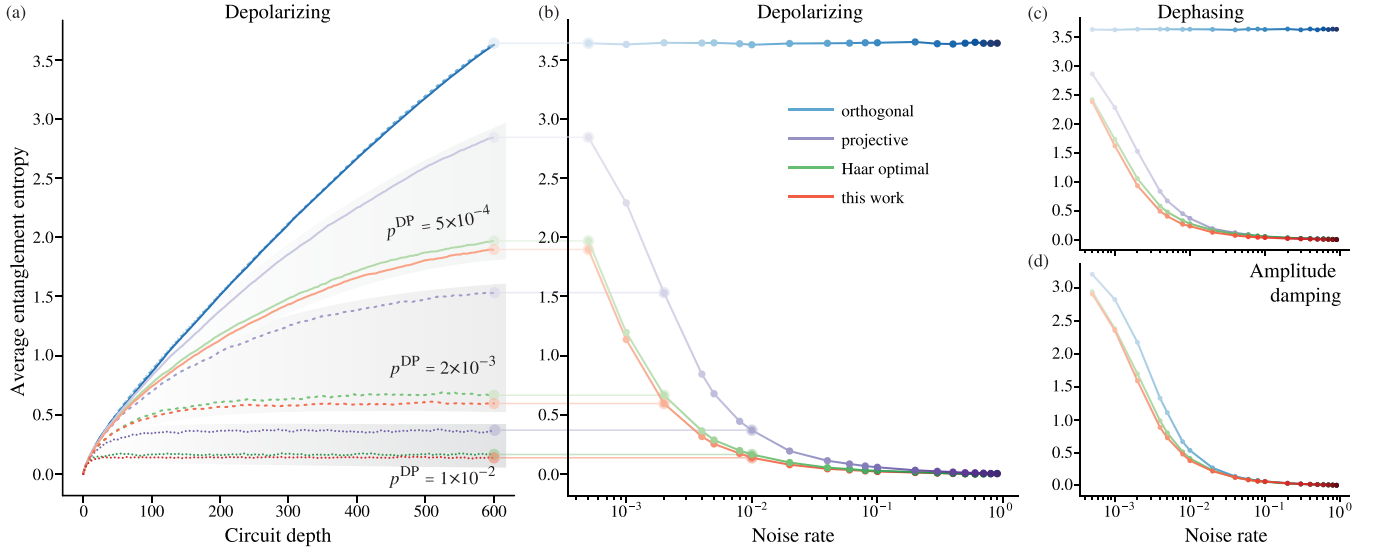


FIG. 7. Average entanglement entropy across the middle bond of the MPS in the numerical implementation of the trajectory unraveling algorithm to different random circuits on 12 qubits with noise for different choices of unraveling. Each run results from one unraveled trajectory of a newly sampled circuit, and each data point is the average over 1000 runs. For all settings, we computed the average entanglement after each layer. (a) Shows the growth of the entanglement with circuit depth for some noise rates of depolarizing noise on low-entangling random circuits. Note that the shaded area labeling noise rates excludes the mixed-unitary unraveling, as those curves overlap for all noise rates. (b)–(d) Show the value of the entropy after the last layer (600) as a function of noise rates for (b) low-entangling random circuits with depolarizing noise [as in (a)], (c) dephasing noise, and (d) low-entangling random circuits with local random rotations with amplitude damping noise. In these settings, the final value at depth 600 is representative of the relation between the different unravelings for all depths, as in (a). The entangling angle is set at 0.05 for all simulations.

becomes the Haar optimal unraveling when we restrict to Haar random states. Finally, in the case of depolarizing and dephasing noise, we also consider the projective unravelings defined in Table II.

While Ref. [45] also introduces an adaptive heuristic strategy based on minimizing local entanglement, the source code for their optimization routine is not publicly available and we do not include a direct numerical comparison here. However, because our method provides the exact analytical solution to the local entanglement minimization problem approximated by their heuristic, we expect the resulting trajectories to be identical (assuming the heuristic successfully finds the optimum). Our analytical approach finds the optimum deterministically with $O(1)$ overhead, avoiding the convergence uncertainties and computational cost of numer-

ical optimization. We focus our comparison to analytically optimal unraveling methods including the Haar optimal unraveling which is state of the art for the noise models we consider.

1. Time-independent Lindbladian systems

In Fig. 5 we demonstrate that there are settings where our locally entanglement-optimal unraveling results in a significant reduction of entanglement cost and truncation errors. Using a Heisenberg Hamiltonian $H = \sum_i Y_i Y_{i+1} + 0.35X_i + 0.35Y_i + 0.5Z_i$ with amplitude damping noise, we observe a marked performance improvements when going from the orthogonal to the Haar optimal unraveling and again going from the Haar optimal to our unraveling. We see this pattern more

TABLE II. Overview of some unravelings considered in the literature and implemented in the numerical simulations. We also call the orthogonal unraveling of the depolarizing and dephasing channels mixed unitary. The Haar optimal unraveling is obtained by rotating the orthogonal unraveling according to Eq. (5) using the specified unitary, where H is the Hadamard matrix $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. More information can be found in the Supplemental Material [49] (Sec. I.E).

Noise model	Orthogonal	Projective	Haar optimal
Dephasing	$\left\{ \sqrt{1 - \frac{p}{2}} \mathbb{1}, \sqrt{\frac{p}{2}} Z \right\}$	$\{ \sqrt{1 - p} \mathbb{1}, \sqrt{p} 0\rangle\langle 0 , \sqrt{p} 1\rangle\langle 1 \}$	$U = H$
Depolarizing	$\left\{ \sqrt{1 - \frac{3p}{4}} \mathbb{1}, \sqrt{\frac{p}{2}} X, \sqrt{\frac{p}{2}} Y, \sqrt{\frac{p}{2}} Z \right\}$	$\{ \sqrt{1 - p} \mathbb{1}, \sqrt{\frac{p}{2}} 0\rangle\langle 0 , \sqrt{\frac{p}{2}} 0\rangle\langle 1 , \sqrt{\frac{p}{2}} 1\rangle\langle 0 , \sqrt{\frac{p}{2}} 1\rangle\langle 1 \}$	$U = H \otimes H$
Amplitude damping	$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1 - \gamma} \end{pmatrix}, \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix} \right\}$	–	$U = H$

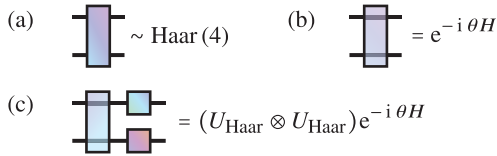


FIG. 8. Construction of the two-qubit unitaries for the three kinds of random circuits considered in the numerical simulations. (a) Two-qubit unitaries drawn from the Haar distribution. (b) Little-entangling gates $U(H, \theta) = e^{-i\theta H}$ where H is a Hermitian matrix with real and imaginary parts of the entries drawn from a normal distribution with mean 0 and standard deviation 1. The rotation angle θ is picked to be small, e.g., 0.1. (c) Little-entangling gates with local rotation $U_{\text{tot}} = (U_{\text{Haar}} \otimes U_{\text{Haar}})U(H, \theta)$ where the local (single-qubit) unitaries are drawn from the Haar distribution while the entangling gate is constructed like in (b).

generally across many instances of Hamiltonians, although sometimes the separations can be less striking. In Fig. 6, we demonstrate this by considering the entanglement and truncation error behavior of competing unravelings when averaged over many randomly selected Hamiltonians. As per Fig. 5, we observe a hierarchy of performance where our unraveling improves on the Haar optimal one, which itself outperforms the orthogonal one. We note that performance improvements within this hierarchy tend to be less marked for extremal levels of noise and bond dimension. As expected, our performance improvement over the Haar optimal unraveling diminishes in settings with low noise and strongly coupled nonintegrable Hamiltonians where the local entanglement structure of the system tends to be similar to that of Haar random states.

We note that for some instances, our unraveling does not result in the best performance. This may be explained by the greediness of our algorithm. As we will discuss in Sec. VII, a somewhat suboptimal local sampling (such as given by the fixed Haar optimal unraveling) could result in more favorable initial conditions for the unraveling in the next time step.

2. Noisy random circuits

In the case of noisy quantum circuits, we consider one-dimensional circuits with brickwork layouts of two-qubit unitary gates, each followed by single-qubit noise channels on the affected qubits. The unitary gates, as shown in Fig. 8, are either drawn from the Haar random distribution, small rotations around the axis defined by a random Hermitian matrix, or the latter followed by local Haar random unitaries. A graphical representation of such circuits can be found in Fig. 2 (top left). We simulated systems of 12 qubits with 600 layers of unitary gates and noise channels. For each noise rate, we show the average of 1000 trajectories. We repeated such simulations for depolarizing, dephasing, and amplitude-damping noise models.

Let us first discuss the case where each gate is drawn randomly from the Haar distribution. As discussed in Sec. VB, for Haar random states, our unraveling and the Haar optimal unraveling become equivalent. As a consequence, in this Haar random gates setting we expect and observe these two unravelings to perform similarly. For all three noise models consid-

ered here, the simulations yielded similar ensemble-averaged entanglement entropy values for our adaptive unraveling and the Haar optimal one [see the Supplemental Material [49] (Sec. X)]. For the two unital channels (depolarizing and dephasing), the orthogonal unraveling is a mixed-unitary unraveling and hence does not affect the entanglement growth, while our unravelings perform best. As previously shown in Refs. [43–45], we observed that the projective unravelings result in smaller values of ensemble-averaged entanglement entropy than the mixed-unitary unravelings but underperform against our unraveling. For amplitude damping, our unraveling also outperforms the orthogonal one.

We now consider the case of small rotations of random circuits, where each gate is a small-angle rotation around a randomly generated Hermitian matrix (as per Fig. 8). For such circuits, when the acting noise is depolarizing or dephasing, our unraveling shows consistently lower values of entanglement entropies than all other unravelings, including the Haar optimal one, as shown in Figs. 7(a)–7(c). First, note that, again, the mixed-unitary unraveling exhibits no dependence on noise rate. Then, in most simulations, we observed that the Haar optimal unraveling outperformed the projective unraveling (or orthogonal unraveling for amplitude damping). In particular, note that the saturation value of the entanglement when using the projective unraveling for a depolarizing rate of $\gamma = 5 \times 10^{-3}$ is nearly the same as that of the Haar optimal unraveling for $\gamma = 2 \times 10^{-3}$. In this case, an evolution with two times less noise can be simulated at the same “entanglement cost.” The Haar optimal unraveling itself is being outperformed by our locally entanglement-optimal unraveling.

Surprisingly, the improvement from the locally entanglement-optimal over Haar optimal unraveling was not to be found for amplitude damping in the low-entangling random circuits [see the Supplemental Material [49] (Sec. X)]. We associate this observation with the fact that the Haar optimal unraveling is not only optimal for Bell states under amplitude-damping noise but for any state whose Schmidt decomposition is in the computational basis [see the Supplemental Material [49] (Sec. VII.B)]. Since the small rotation unitaries do not rotate the state much, and the noise pushes it towards the all-zero state, the states remain close to states with Schmidt decompositions in the computational basis. This hypothesis is confirmed by slightly changing the circuit definition to include Haar random single-qubit rotations after each two-qubit small-angle unitary, as shown in Fig. 8(c). The single-qubit gates do not affect the entanglement between qubits, but they randomize the basis of the Schmidt decomposition. For random circuits composed of such gates, our unraveling again outperforms the Haar optimal unraveling, as shown in Fig. 7(d).

We also note that the depth of the simulation is insufficient to reach the saturation value of the average entanglement for the smaller noise rates. This is visible in Fig. 7(a) for the curves corresponding to $\gamma = 5 \times 10^{-4}$ as opposed to those for $\gamma = 2 \times 10^{-3}$ or $\gamma = 10^{-2}$ which saturated (for all except the mixed-unitary unraveling). For that reason, Figs. 7(b)–7(d) should be seen rather as an indication of the behavior. Still, our results indicate that the trend of separations should be conserved.

B. Simulability thresholds

In simulations of noisy random 1D circuits based on unraveling into trajectories, the average growth of entanglement with system size can undergo a noise rate-dependent transition from a volume law to an area law. Based on the relation between entanglement scaling and representability using MPS, the entanglement phase transition is interpreted as a threshold between nonsimulable and *efficiently* simulable noise regimes (on average over random circuits). The exact value of the threshold depends, however, on the choice of unraveling, and is thus used as a criterion for the quality of an unraveling choice. This directly connects to the study of measurement-induced phase transitions (MIPT) [82–86]. Indeed, in the setting of noisy random quantum circuits, MIPT can be understood as one particular choice of unraveling (the projective unraveling) of the dephasing noise channel into projectors onto the computational basis (see the example in Sec. II).

To convey a sense of the scale of noise rate sufficient for efficient simulability using our algorithm, we discuss known noise thresholds from the trajectory sampling and MIPT literature. Since we expect and numerically observe a fairly robust performance improvement from our unraveling, we expect our unraveling to achieve noise thresholds matching or below the MIPT and Haar optimal unraveling thresholds discussed below.

Mixed-unitary unravelings admit no noise threshold, as randomly inserted single-qubit unitaries do not affect the average entanglement dynamics of random circuits. In contrast, projective unraveling of the dephasing channel was shown to have a threshold around 0.2 [43,82,83]. The value of the thresholds was reduced further using the Haar optimal unraveling 0.09 [44,45]. The threshold is expected to be improvable further using multiqubit unravelings.

In the random circuit setting, Ref. [45] numerically compared their fixed Haar optimal unraveling to their adaptive numerical optimization unraveling (a heuristic version of our analytical unraveling method). Somewhat surprisingly at the time, they estimated similar thresholds for the two unravelings. As a consequence of Sec. VB, we now understand why the adaptive methods should reduce to the fixed Haar optimal unraveling in the random circuits setting. It is, however, to be expected that our strategy would lead to an improvement of thresholds for other systems compared to the fixed Haar optimal unraveling. Because our method provides the exact analytical solution to the local entanglement minimization problem approximated by the heuristic in Ref. [45], we expect the resulting simulability thresholds to be identical in the limit where the heuristic successfully finds the global optimum. Our contribution is thus unlikely to result in a significant shift in the physical threshold itself, but instead furnishes the computation of this optimum deterministically with $O(1)$ overhead, avoiding the convergence uncertainties and computational cost of variational optimization.

C. Comparison to other simulation methods

1. Tensor network methods

When trying to generalize tensor network representations for one-dimensional quantum systems from MPS to mixed states, one is given several options. The first approach is to use matrix product operators (MPOs), in which one adds an open leg to each tensor for the dual space [87–89]. MPOs represent the density matrix as a train of rank-4 tensors. In addition to the potential computational overhead associated with the quadratic increase in dimension, this approach is known to suffer from technical issues limiting the effectiveness of SVD based truncation. Indeed, truncations based on the efficiently implementable SVD⁷ minimize the 2-norm error of the post-truncation state. For vector representations of states such as MPS, this is the relevant metric. However, for matrix representations of states such as MPOs, the relevant metric is the trace norm. Furthermore, truncation of MPOs may result in negative eigenvalues. As checking whether an MPO has negative eigenvalues is computationally hard [90], one may also obtain nonphysical results. MPO representations are also known to suffer from an entanglement barrier when used to simulate open system evolutions [38,44,91]. It has been shown that, in certain settings, simulation based on unraveled trajectories can lead to an exponential reduction in simulation cost compared to MPO simulation and thereby avoid the entanglement barrier plaguing MPOs [46]. Such results are, however, very sensitive to the specific system being simulated, the rate of decoherence, and the initial state, among others, leading to mixed affirmations on the comparison of the efficiency of both methods [92–95]. However, to the best of our knowledge, none of these comparisons have been performed with the improved unraveling strategies including the Haar optimal unraveling from Refs. [44,45].

An alternative representation of mixed states, which is guaranteed to preserve positivity and which can be truncated like an MPS, is obtained by (locally) purifying the density matrix. This gives rise to locally purified density operators (LPDOs) [89,96–100]⁸. These, however, result in looped tensor network structures (when tracing out the auxiliary systems) which can not be optimized locally and for which there is no entanglement optimal form [101,102]. Additionally, the required bond dimensions may be unboundedly larger than for the MPO counterpart [103] (although, albeit only for exact representations [104]). In this language, trajectory sampling methods can be interpreted as measuring out the auxiliary systems, thus destroying the purifying bond of the LPDO. LPDOs are, however, strictly harder to optimize than trajectory methods, as one needs to minimize the local entanglement both within the state and between the state and the environment.

2. Pauli propagation methods

Another increasingly prominent method for the simulation of noisy quantum circuits is the Pauli propagation method

⁷For 1D tensor trains, SVDs are efficiently implementable in the dimension of the local tensors.

⁸Note that the naming of this tensor network construction is highly inconsistent in the literature. Some authors refer to them as matrix product density operators (which others use for MPOs with trace 1, independently of positivity), locally purified density operators, or purified matrix product states.

(also known as the Pauli path method or sparse Pauli dynamics) [32,35,36,105–118]. Pauli propagation and trajectory sampling methods are expected to perform best in different settings. While MPS-based trajectory methods are tailored to systems with low entanglement, Pauli propagation is rather related to the amount of magic, as they can represent evolutions with (approximately) sparse decompositions in the Pauli basis. Both methods can be used to perform both observable estimation and output distribution sampling (see Sec. III D). However, efficient Pauli path based output distribution sampling algorithms assume anticoncentration of the circuit output distribution [36] with some special exceptions [119,120].

The Pauli path method has proven to be a great tool for proving the existence of efficient classical observable estimation algorithms for typical noisy random quantum circuits. A key strength of this method is that it comes with *a priori* performance guarantees. However, this method is more limited in scope, applying to specific types of noise in specific location and to random circuits. However, progress continues on this front [32,36]. For example, Ref. [36] relaxed the location restrictions on the depolarizing noise model they consider and concurrently improved the random circuit setting to a random computational basis state initialization. Additionally, in the observable estimation setting, Ref. [32] removed the circuit randomization entirely for 2D Clifford plus T circuits and depolarizing noise.

Several other subtle differences differentiate the applicability of each method. While Pauli propagation methods consider observables that are at most combinations of polynomially many Pauli strings, MPS trajectories allow for the computation of expectation values for any observable that can be represented as a low bond dimension MPO. On the other hand, MPS trajectory simulations require initial states that can be represented as MPS states of small bond dimension, whereas Pauli propagation only requires efficient expectation value calculation for single Pauli strings. However, most commonly, the classical simulation literature focuses on simulating circuits with simple input states, like the all-zero state.

VII. LIMITATIONS

A. Local vs global optimization

The most desirable setting would be to have the ability to jointly optimize the unitary freedom over all channels, in both time and space. Indeed, given two single-qubit channels, optimizing over the unitary freedom of the joint two-qubit channel can lead to better results than optimizing them individually. Such global optimization, however, is not achievable in practice, as it involves, in the most general case, an optimization over a 4^{2nL} -dimensional unitary. Our approach optimizes each channel individually and can be seen as a time and space-greedy optimization method. In-between, one can consider time- or space-only greedy optimization, but it is unclear if such optimizations are feasible. Take the case of time-only greedy optimization. Finding the entanglement-optimal decomposition of the joint n -qubit channel is at least as hard as computing the entanglement of formation of an n -qubit mixed state, which is known to be prohibitively hard if one can not

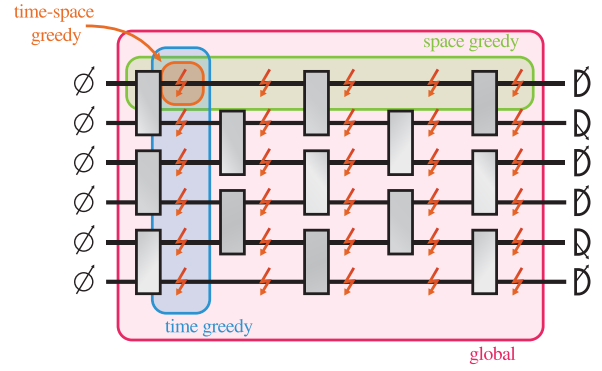


FIG. 9. Various possible choices of joint optimization of single-qubit noise channels.

leverage some symmetry [121–125]. See Fig. 9 for a graphical representation of the different types of “greediness” referred to above.

Our unraveling will, in general, perform worse than joint optimizations would. There is also no guarantee that such a greedy optimization will work at all, as one can construct instances where an “optimal” choice in an early step results in disadvantageous initial conditions in a later step. We found instances where using a fixed unraveling choice at all steps (in time and space) outperformed our greedy optimization. It would be interesting to understand in which settings such greediness performs well and in which it is disadvantageous. For instance, in the setting of time-independent Lindbladian evolutions, one may expect the greedy unraveling to fluctuate less and less as the evolution thermalizes. An example where a seemingly irrelevant change in basis significantly affects the performance of the Haar optimal unraveling is presented in the Supplemental Material [49] (Sec. XI).

B. Von Neumann entanglement entropy as the choice of cost function

It is often stated that MPS offer efficient and good approximations of states with low entanglement, while they fail at faithfully representing highly entangled states. This statement, however, is not formally correct. Let us first recall that entanglement can be measured by the more general class of *Rényi entropies*

$$S_\alpha(\rho) = \frac{1}{1-\alpha} \log \text{Tr} [\rho^\alpha], \quad 0 \leq \alpha \leq \infty. \quad (57)$$

They generalize the von Neumann entanglement entropy used in Eq. (9), which is recovered in the limit of α tending to 1. It has been shown that classes of states which have entanglement growing logarithmically with system size as measured by some Rényi entropy with $\alpha < 1$ can be represented efficiently with MPS [62] while those with Rényi $\alpha > 1$ entropies that grow superlogarithmically with system size ($\sim n^\kappa$ even for $\kappa < 1$) cannot [63]. For von Neumann entropies, the nonapproximability result holds for linearly scaling entropy, but the remaining guarantees are lost. For constant and logarithmically scaling von Neumann entropies, there are even examples of both classes of states, representable states and nonrepresentable states.

Our algorithm is devised to minimize the ensemble-averaged (von Neumann) entanglement entropy in the hope of reducing simulation costs. Since it minimizes the von Neumann entanglement entropy, there will be instances where these differences in MPS representability emerge. Although one expects the different Rényi entropies to have similar behavior in typical settings, there will be examples where optimization of the von Neumann entanglement entropy, some Rényi entropy with $\alpha < 1$, or the truncation error will give widely different optimal unravelings. In such a setting, optimizing the von Neumann entanglement entropy could lead to poor performance.

VIII. CONCLUSION AND DISCUSSION

In this work, we have aimed to substantially contribute to the body of work on the classical simulation of noisy quantum circuits with trajectory sampling by putting such studies onto a significantly more quantitative and rigorous underpinning. Concretely, we improve existing methods in two ways. We derive rigorous error bounds on the performance of such algorithms based on MPS, particularly for the relevant tasks of expectation value estimation and measurement outcome distribution sampling. We also propose ways of computing good decompositions of noise channels. Our first choice of unraveling is provably optimal in a state-dependent, greedy fashion by maximally disentangling the target noisy qubit. At the cost of computational effort similar to that of an SVD of a 4×4 matrix, we can minimize the average entanglement between the noisy qubit and the rest of the state. This unraveling is not tailored to random circuits but is applicable to any circuit. It is also not limited to specific noise channels, as we enable the deterministic treatment of systems where each qubit is affected by a potentially different noise process.

By readily lending itself to analytic investigation, the same tools allow for the construction of an even simpler, state-independent unraveling for random circuits. In this setting, it provides a conceptually significant perspective on what is being optimized. It also reproduces the previous results on fixed unravelings for random circuits and extends these to arbitrary noise models.

Our unraveling strategy has repercussions of a practical and conceptual nature. It translates into improved performance and perhaps, most importantly, an extended scope by permitting locally optimal unravelings with respect to arbitrary 1D local circuits and single-qubit noise. We believe the latter is an important extension for future progress relating to quantum advantage and the noise-driven decay of quantum computational resources.

While our results comprehensively answer questions about the optimal (greedy) single-qubit unraveling, many questions remain open, and we raise new ones. The problem of analytically motivated unravelings for multiqubit noise processes remains untackled. Although in the single-qubit setting, one can focus on disentangling one qubit, which can only be done

using nonunitary operators, the multiqubit case is more nuanced. There, one could minimize the entanglement between the target qubits and the rest of the state or within the target qubits themselves. The first approach bears some similarities with our approach, although there are barely any results on the computation of the entanglement of formation for more than two qubits. The second opens the door for entanglement reduction through unitary Kraus operators. Furthermore, it would be beneficial to lift the limitations of our approach. Developing unravelings based on Rényi entropies or even the truncation error would lead to stronger analytical performance guarantees. Some works also focused on the scaling factors of the computational cost and, in particular, on how to reduce the variance in the trajectories. It would be interesting to study our unraveling from this perspective.

With the flexibility and generality of our approach, we make progress towards a coherent theory of quantum noise and its impact on quantum advantage. Over the years, in the quest for finding quantum advantages, it has been an enormously fruitful endeavor to run quantum platforms against classical simulations of the same task, usually by resorting to tensor network models, actually improving both research fields at the same time [126–130]. For this reason, having strong predictive power and knowing what the precise limitations are for the classical simulation is important, and hence our work substantially contributes here. Our work lends itself to more detailed numerical studies investigating specific classes of circuits and their behavior in the presence of single-qubit noise. Our work provides *a posteriori* guarantees on simulation accuracy and, since these can in principle take all circuit-level data into account, generally speaking, one might expect such guarantees to be tighter than coarse-grained *a priori* guarantees. However, *a priori* performance guarantees are very important for identifying regimes or circuit classes that admit efficient classical simulability. This motivates the development of new techniques that provide *a priori* performance guarantees on our classical simulation algorithm.

ACKNOWLEDGMENTS

We would like to thank F. vom Ende for useful discussions. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (The Berlin Mathematics Research Center MATH + EXC-2046/1, Project ID No. 390685689, the ML4Q EXC 2004/1, Project ID No. 390534769) the CRC 183, SPP 2514, the BMFTR (MuniQC-Atoms, QuSol, QPIC-1), the Munich Quantum Valley, Berlin Quantum, the Quantum Flagship (PasQuans2, Millenion), and the European Research Council (DebuQC).

DATA AVAILABILITY

The code for the simulations that support the findings of this article is openly available [131].

[1] S. Aaronson and A. Ambainis, The need for structure in quantum speedups *Theory Comput.* **10**, 133 (2014).

[2] S. Ben-David and S. Podder, How symmetric is too symmetric for large quantum speedups? *arXiv:2001.09642*.

- [3] A. M. Childs and D. Wang, Can graph properties have exponential quantum speedup? [arXiv:2001.10520](#).
- [4] S. Ben-David, A. M. Childs, A. Gilyén, W. Kretschmer, S. Podder, and D. Wang, in *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Piscataway, NJ, 2020), pp. 649–660.
- [5] S. Aaronson and A. Arkhipov, The computational complexity of linear optics [Theory Comput. **9**, 143 \(2013\)](#).
- [6] M. J. Bremner, A. Montanaro, and D. J. Shepherd, Average-case complexity versus approximate simulation of commuting quantum computations, [Phys. Rev. Lett. **117**, 080501 \(2016\)](#).
- [7] A. Bouland, J. F. Fitzsimons, and D. E. Koh, in *33rd Computational Complexity Conference (CCC 2018)*, edited by R. A. Servedio, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 102 (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2018), pp. 21:1–21:25.
- [8] H. Pashayan, S. D. Bartlett, and D. Gross, From estimation of quantum probabilities to simulation of quantum circuits, [Quantum **4**, 223 \(2020\)](#).
- [9] M. Yoganathan, R. Jozsa, and S. Strelchuk, Quantum advantage of unitary Clifford circuits with magic state inputs, [Proc. R. Soc. A **475**, 20180427 \(2019\)](#).
- [10] M. Oszmaniec, N. Dangniam, M. E. S. Morales, and Z. Zimborás, Fermion sampling: A robust quantum computational advantage scheme using fermionic linear optics and magic input states, [PRX Quantum **3**, 020328 \(2022\)](#).
- [11] D. Hangleiter and J. Eisert, Computational advantage of quantum random sampling, [Rev. Mod. Phys. **95**, 035001 \(2023\)](#).
- [12] D. Gottesman, The Heisenberg representation of quantum computers, in *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, edited by S. P. Corney, R. Delbourgo, and P. D. Jarvis (International Press, Cambridge, MA, 1999), pp. 32–43.
- [13] S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, [Phys. Rev. A **70**, 052328 \(2004\)](#).
- [14] L. G. Valiant, in *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing, STOC '01* (Association for Computing Machinery, New York, 2001), pp. 114–123.
- [15] B. M. Terhal and D. P. DiVincenzo, Classical simulation of noninteracting-fermion quantum circuits, [Phys. Rev. A **65**, 032325 \(2002\)](#).
- [16] S. Bravyi and D. Gosset, Improved classical simulation of quantum circuits dominated by Clifford gates, [Phys. Rev. Lett. **116**, 250501 \(2016\)](#).
- [17] H. Qassim, H. Pashayan, and D. Gosset, Improved upper bounds on the stabilizer rank of magic states, [Quantum **5**, 606 \(2021\)](#).
- [18] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Simulation of quantum circuits by low-rank stabilizer decompositions, [Quantum **3**, 181 \(2019\)](#).
- [19] A. Heimendahl, F. Montealegre-Mora, F. Vallentin, and D. Gross, Stabilizer extent is not multiplicative, [Quantum **5**, 400 \(2021\)](#).
- [20] J. R. Seddon, B. Regula, H. Pashayan, Y. Ouyang, and E. T. Campbell, Quantifying quantum speedups: Improved classical simulation from tighter magic monotones, [PRX Quantum **2**, 010345 \(2021\)](#).
- [21] O. Reardon-Smith, M. Oszmaniec, and K. Korzekwa, Improved simulation of quantum circuits dominated by free fermionic operations, [Quantum **8**, 1549 \(2024\)](#).
- [22] B. Dias and R. Koenig, Classical simulation of non-Gaussian fermionic circuits, [Quantum **8**, 1350 \(2024\)](#).
- [23] O. Reardon-Smith, The fermionic linear optical extent is multiplicative for 4 qubit parity eigenstates, [arXiv:2407.20934](#).
- [24] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Matrix product state representations [Quantum Inf. Comput. **7**, 401 \(2007\)](#).
- [25] G. Vidal, Efficient classical simulation of slightly entangled quantum computations, [Phys. Rev. Lett. **91**, 147902 \(2003\)](#).
- [26] G. Vidal, Efficient simulation of one-dimensional quantum many-body systems, [Phys. Rev. Lett. **93**, 040502 \(2004\)](#).
- [27] R. Orús, Tensor networks for complex quantum systems, [Nat. Rev. Phys. **1**, 538 \(2019\)](#).
- [28] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, [Rev. Mod. Phys. **93**, 045003 \(2021\)](#).
- [29] J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area laws for the entanglement entropy*, [Rev. Mod. Phys. **82**, 277 \(2010\)](#).
- [30] E. T. Campbell, B. M. Terhal, and C. Vuillot, Roads towards fault-tolerant universal quantum computation, [Nature \(London\) **549**, 172 \(2017\)](#).
- [31] S. Virmani, S. F. Huelga, and M. B. Plenio, Classical simulability, entanglement breaking, and quantum computation thresholds, [Phys. Rev. A **71**, 042328 \(2005\)](#).
- [32] G. González-García, J. I. Cirac, and R. Trivedi, Pauli path simulations of noisy quantum circuits beyond average case, [Quantum **9**, 1730 \(2025\)](#).
- [33] D. S. Franca and R. Garcia-Patron, Limitations of optimization algorithms on noisy quantum devices, [Nat. Phys. **17**, 1221 \(2021\)](#).
- [34] A. Deshpande, P. Niroula, O. Shtanko, A. V. Gorshkov, B. Fefferman, and M. J. Gullans, Tight Bounds on the Convergence of Noisy Random Circuits to the Uniform Distribution, [PRX Quantum **3**, 040329 \(2022\)](#).
- [35] D. Aharonov, X. Gao, Z. Landau, Y. Liu, and U. Vazirani, A polynomial-time classical algorithm for noisy random circuit sampling, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC 2023)* (Association for Computing Machinery, New York, NY, 2023), pp. 945–957.
- [36] T. Schuster, C. Yin, X. Gao, and N. Y. Yao, A polynomial-time classical algorithm for noisy quantum circuits, [Phys. Rev. X **15**, 041018 \(2025\)](#).
- [37] A. A. Mele, A. Angrisani, S. Ghosh, S. Khatri, J. Eisert, D. S. França, and Y. Quek, Noise-induced shallow circuits and the absence of barren plateaus, [Nat. Phys. **22**, 751 \(2026\)](#).
- [38] K. Noh, L. Jiang, and B. Fefferman, Efficient classical simulation of noisy random quantum circuits in one dimension, [Quantum **4**, 318 \(2020\)](#).
- [39] R. S. Bennink, E. M. Ferragut, T. S. Humble, J. A. Laska, J. J. Nutaro, M. G. Pleszkoch, and R. C. Pooser, Unbiased simulation of near-Clifford quantum circuits, [Phys. Rev. A **95**, 062337 \(2017\)](#).
- [40] S. V. Isakov, D. Kafri, O. Martin, C. V. Heidweiller, W. Mruczkiewicz, M. P. Harrigan, N. C. Rubin, R. Thomson, M. Broughton, K. Kissell, E. Peters, E. Gustafson, A. C. Y. Li, H. Lamm, G. Perdue, A. K. Ho, D. Strain, and S. Boixo, Simulations of quantum circuits with approximate noise using qsim and Cirq, [arXiv:2111.02396](#).

- [41] A. Sander, M. Fröhlich, M. Eigel, J. Eisert, P. Gelß, M. Hintermüller, R. M. Milbradt, R. Wille, and C. B. Mendl, Large-scale stochastic simulation of open quantum systems, *Nat. Commun.* **16**, 11074 (2025).
- [42] T. Vovk and H. Pichler, Entanglement-optimal trajectories of many-body quantum Markov processes, *Phys. Rev. Lett.* **128**, 243601 (2022).
- [43] M. Kolodrubetz, Optimality of Lindblad unfolding in measurement phase transitions, *Phys. Rev. B* **107**, L140301 (2023).
- [44] Z. Cheng and M. Ippoliti, Efficient sampling of noisy shallow circuits via monitored unraveling, *PRX Quantum* **4**, 040326 (2023).
- [45] Z. Chen, Y. Bao, and S. Choi, Optimized trajectory unraveling for classical simulation of noisy quantum dynamics, *Phys. Rev. Lett.* **133**, 230403 (2024).
- [46] T. Vovk and H. Pichler, Quantum trajectory entanglement in various unravelings of Markovian dynamics, *Phys. Rev. A* **110**, 012207 (2024).
- [47] R. Daraban, F. Salas-Ramírez, and J. Schachenmayer, Non-unitarity maximizing unraveling of open quantum dynamics, *SciPost Phys.* **18**, 048 (2025).
- [48] W. K. Wootters, Entanglement of formation of an arbitrary state of two qubits, *Phys. Rev. Lett.* **80**, 2245 (1998).
- [49] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/jj2r-y777> for further preliminaries and notation, visualizing examples, detailed proofs of the results and additional numerical results.
- [50] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, *Phys. Lett. A* **183**, 14 (1993).
- [51] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, *Phys. Rev. A* **40**, 4277 (1989).
- [52] H. Azuma and M. Ban, Another convex combination of product states for the separable Werner state, *Phys. Rev. A* **73**, 032315 (2006).
- [53] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824 (1996).
- [54] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, *Phys. Rev. A* **53**, 2046 (1996).
- [55] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81**, 865 (2009).
- [56] X.-N. Zhu and S.-M. Fei, Improved lower and upper bounds for entanglement of formation, *Phys. Rev. A* **86**, 054301 (2012).
- [57] J. S. Kim, Entanglement of formation and monogamy of multi-party quantum entanglement, *Sci. Rep.* **11**, 2364 (2021).
- [58] A. Uhlmann, Roofs and convexity, *Entropy* **12**, 1799 (2010).
- [59] A. Uhlmann, Entropy and optimal decompositions of states relative to a maximal commutative subalgebra, *Open Syst. Inf. Dyn.* **5**, 209 (1998).
- [60] K. Audenaert, F. Verstraete, and B. De Moor, Variational characterizations of separability and entanglement of formation, *Phys. Rev. A* **64**, 052304 (2001).
- [61] U. Schollwoeck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.* **326**, 96 (2011).
- [62] F. Verstraete and J. I. Cirac, Matrix product states represent ground states faithfully, *Phys. Rev. B* **73**, 094423 (2006).
- [63] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Entropy scaling and simulability by matrix product states, *Phys. Rev. Lett.* **100**, 030504 (2008).
- [64] P. Zoller, M. Marte, and D. F. Walls, Quantum jumps in atomic systems, *Phys. Rev. A* **35**, 198 (1987).
- [65] R. Dum, P. Zoller, and H. Ritsch, Monte Carlo simulation of the atomic master equation for spontaneous emission, *Phys. Rev. A* **45**, 4879 (1992).
- [66] H. Carmichael, *An Open Systems Approach to Quantum Optics: Lectures Presented at the Université Libre de Bruxelles October 28 to November 4, 1991*, Lecture Notes in Physics Monographs, Vol. 18 (Springer, Berlin, 1993).
- [67] K. Mølmer, Y. Castin, and J. Dalibard, Monte Carlo wavefunction method in quantum optics, *J. Opt. Soc. Am. B* **10**, 524 (1993).
- [68] K. Mølmer and Y. Castin, Monte Carlo wavefunctions in quantum optics, *Quantum Semiclass. Opt.* **8**, 49 (1996).
- [69] M. B. Plenio and P. L. Knight, The quantum-jump approach to dissipative dynamics in quantum optics, *Rev. Mod. Phys.* **70**, 101 (1998).
- [70] A. J. Daley, Quantum trajectories and open many-body quantum systems, *Adv. Phys.* **63**, 77 (2014).
- [71] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Am. Stat. Assoc.* **58**, 13 (1963).
- [72] A. J. Ferris and G. Vidal, Perfect sampling with unitary tensor networks, *Phys. Rev. B* **85**, 165146 (2012).
- [73] B. Vanhecke, M. Van Damme, J. Haegeman, L. Vanderstraeten, and F. Verstraete, Tangent-space methods for truncating uniform MPS, *SciPost Phys. Core* **4**, 004 (2021).
- [74] S. A. Hill and W. K. Wootters, Entanglement of a pair of quantum bits, *Phys. Rev. Lett.* **78**, 5022 (1997).
- [75] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. (Cambridge University Press, Cambridge, 2012).
- [76] M. Houde, W. McCutcheon, and N. Quesada, Matrix decompositions in quantum optics: Takagi/Autonne, Bloch-Messiah/Euler, Iwasawa, and Williamson, *Can. J. Phys.* **102**, 497 (2024).
- [77] V. A. Marčenko and L. A. Pastur, Distribution of eigenvalues for some sets of random matrices, *Math. USSR-Sbornik* **1**, 457 (1967).
- [78] M. Žnidarič, Entanglement of random vectors, *J. Phys. A: Math. Theor.* **40**, F105 (2007).
- [79] A. A. Mele, Introduction to Haar measure tools in quantum information: A beginner's tutorial, *Quantum* **8**, 1340 (2024).
- [80] J. M. Renes, *Quantum Information Theory* (De Gruyter Oldenbourg, Munich, 2022).
- [81] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, 2018).
- [82] B. Skinner, J. Ruhman, and A. Nahum, Measurement-induced phase transitions in the dynamics of entanglement, *Phys. Rev. X* **9**, 031009 (2019).
- [83] A. Zabalo, M. J. Gullans, J. H. Wilson, S. Gopalakrishnan, D. A. Huse, and J. H. Pixley, Critical properties of the measurement-induced transition in random quantum circuits, *Phys. Rev. B* **101**, 060301(R) (2020).

- [84] Y. Bao, S. Choi, and E. Altman, Theory of the phase transition in random unitary circuits with measurements, *Phys. Rev. B* **101**, 104301 (2020).
- [85] C.-M. Jian, Y.-Z. You, R. Vasseur, and A. W. W. Ludwig, Measurement-induced criticality in random quantum circuits, *Phys. Rev. B* **101**, 104302 (2020).
- [86] R. Suzuki, J. Haferkamp, J. Eisert, and P. Faist, Quantum complexity phase transitions in monitored random circuits, *Quantum* **9**, 1627 (2025).
- [87] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, Matrix product operator representations, *New J. Phys.* **12**, 025012 (2010).
- [88] J. Guth Jarkovský, A. Molnár, N. Schuch, and J. I. Cirac, Efficient description of many-body systems with matrix product density operators, *PRX Quantum* **1**, 010304 (2020).
- [89] H. Weimer, A. Kshetrimayum, and R. Orús, Simulation methods for open quantum many-body systems, *Rev. Mod. Phys.* **93**, 015008 (2021).
- [90] M. Kliesch, D. Gross, and J. Eisert, Matrix-product operators and states: NP-hardness and undecidability, *Phys. Rev. Lett.* **113**, 160503 (2014).
- [91] D. Wellnitz, G. Preisser, V. Alba, J. Dubail, and J. Schachenmayer, Rise and fall, and slow rise again, of operator entanglement under dephasing, *Phys. Rev. Lett.* **129**, 170401 (2022).
- [92] L. Bonnes and A. M. Läuchli, Superoperators vs. trajectories for matrix product state simulations of open quantum system: A case study, [arXiv:1411.4831](https://arxiv.org/abs/1411.4831).
- [93] S. Wolff, A. Sheikhan, and C. Kollath, Numerical evaluation of two-time correlation functions in open quantum systems with matrix product state methods: A comparison, *SciPost Phys. Core* **3**, 010 (2020).
- [94] G. Preisser, D. Wellnitz, T. Botzung, and J. Schachenmayer, Comparing bipartite entropy growth in open-system matrix-product simulation methods, *Phys. Rev. A* **108**, 012616 (2023).
- [95] EPL. van Nieuwenburg, J. Y. Malo, A. J. Daley, and M. H. Fischer, Dynamics of many-body localization in the presence of particle loss, *Quantum Sci. Technol.* **3**, 01LT02 (2018).
- [96] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Matrix product density operators: Simulation of finite-temperature and dissipative systems, *Phys. Rev. Lett.* **93**, 207204 (2004).
- [97] M. Zwolak and G. Vidal, Mixed-state dynamics in one-dimensional quantum lattice systems: A time-dependent superoperator renormalization algorithm, *Phys. Rev. Lett.* **93**, 207205 (2004).
- [98] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, Positive tensor network approach for simulating open quantum many-body systems, *Phys. Rev. Lett.* **116**, 237201 (2016).
- [99] S. Cheng, C. Cao, C. Zhang, Y. Liu, S.-Y. Hou, P. Xu, and B. Zeng, Simulating noisy quantum circuits with matrix product density operators, *Phys. Rev. Res.* **3**, 023005 (2021).
- [100] Y. Guo and S. Yang, Locally purified density operators for noisy quantum circuits, *Chin. Phys. Lett.* **41**, 120302 (2024).
- [101] G. D. las Cuevas, N. Schuch, D. Pérez-García, and J. I. Cirac, Purifications of multipartite states: Limitations and constructive methods, *New J. Phys.* **15**, 123021 (2013).
- [102] G. De las Cuevas, T. S. Cubitt, J. I. Cirac, M. M. Wolf, and D. Pérez-García, Fundamental limitations in the purifications of tensor networks, *J. Math. Phys.* **57**, 071902 (2016).
- [103] I. Glasser, R. Sweke, N. Pancotti, J. Eisert, and I. Cirac, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., Red Hook, NY, 2019), Vol. 32.
- [104] G. De las Cuevas, A. Klingler, and T. Netzer, Approximate tensor decompositions: Disappearance of many separations, *J. Math. Phys.* **62**, 093502 (2021).
- [105] P. Rall, D. Liang, J. Cook, and W. Kretschmer, Simulation of qubit quantum circuits via Pauli propagation, *Phys. Rev. A* **99**, 062337 (2019).
- [106] Q. Huang and C. B. Mendl, Classical simulation of quantum circuits using a multiqubit Bloch vector representation of density matrices, *Phys. Rev. A* **105**, 022409 (2022).
- [107] T. Begušić, J. Gray, and G. K.-L. Chan, Fast and converged classical simulations of evidence for the utility of quantum computing before fault tolerance, *Sci. Adv.* **10**, eadk4321 (2024).
- [108] T. Begušić, K. Hejazi, and G. K.-L. Chan, Simulating quantum circuit expectation values by Clifford perturbation theory, *J. Chem. Phys.* **162**, 154110 (2025).
- [109] E. Fontana, M. S. Rudolph, R. Duncan, I. Rungger, and C. Cîrstoiu, Classical simulations of noisy variational quantum circuits, *npj Quantum Inf.* **11**, 84 (2025).
- [110] M. S. Rudolph, E. gottes, Z. Holmes, and L. Cincio, Classical surrogate simulation of quantum systems with LOWESA, [arXiv:2308.09109](https://arxiv.org/abs/2308.09109).
- [111] I. Ermakov, O. Lychkovskiy, and T. Byrnes, Unified framework for efficiently computable quantum circuits, [arXiv:2401.08187](https://arxiv.org/abs/2401.08187).
- [112] Y. Shao, F. Wei, S. Cheng, and Z. Liu, Simulating noisy variational quantum algorithms: A polynomial approach, *Phys. Rev. Lett.* **133**, 120603 (2024).
- [113] A. Angrisani, A. Schmidhuber, M. S. Rudolph, M. Cerezo, Z. Holmes, and H.-Y. Huang, Classically estimating observables of noiseless quantum circuits, *Phys. Rev. Lett.* **135**, 170602 (2025).
- [114] C. Cîrstoiu, A Fourier analysis framework for approximate classical simulations of quantum circuits, [arXiv:2410.13856](https://arxiv.org/abs/2410.13856).
- [115] S. Lerch, R. Puig, M. S. Rudolph, A. Angrisani, T. Jones, M. Cerezo, S. Thanasilp, and Z. Holmes, Efficient quantum-enhanced classical simulation for patches of quantum landscapes, *PRX Quantum* **7**, 020359 (2026).
- [116] A. Angrisani, A. A. Mele, M. S. Rudolph, M. Cerezo, and Z. Holmes, Simulating quantum circuits with arbitrary local noise using Pauli propagation, *PRX Quantum* **7**, 020313 (2026).
- [117] V. Martinez, A. Angrisani, E. Pankovets, O. Fawzi, and D. Stilck França, Efficient simulation of parametrized quantum circuits under nonunitary noise through Pauli backpropagation, *Phys. Rev. Lett.* **134**, 250602 (2025).
- [118] M. S. Rudolph, T. Jones, Y. Teng, A. Angrisani, and Z. Holmes, Pauli propagation: A computational framework for simulating quantum systems, [arXiv:2505.21606](https://arxiv.org/abs/2505.21606).
- [119] J. Nelson, J. Rajakumar, D. Hangleiter, and M. J. Gullans, Polynomial-time classical simulation of noisy quantum circuits with naturally fault-tolerant gates, [arXiv:2411.02535](https://arxiv.org/abs/2411.02535).

- [120] J. Rajakumar, J. D. Watson, and Y.-K. Liu, in *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, SODA '25 (SIAM, Philadelphia, 2025), pp. 1037–1056.
- [121] D. P. DiVincenzo, B. M. Terhal, and A. V. Thapliyal, Optimal decompositions of barely separable states, *J. Mod. Opt.* **47**, 377 (2000).
- [122] B. M. Terhal and K. G. H. Vollbrecht, Entanglement of formation for isotropic states, *Phys. Rev. Lett.* **85**, 2625 (2000).
- [123] K. G. H. Vollbrecht and R. F. Werner, Why two qubits are special, *J. Math. Phys.* **41**, 6772 (2000).
- [124] K. G. H. Vollbrecht and R. F. Werner, Entanglement measures under symmetry, *Phys. Rev. A* **64**, 062307 (2001).
- [125] S. Gharibian, Strong NP-hardness of the quantum separability problem, *Quantum Inf. Comput.* **10**, 343 (2010).
- [126] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas, *Nat. Phys.* **8**, 325 (2012).
- [127] Y. Kim, A. Eddins, S. Anand, K. Wei, E. Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, *Nature (London)* **618**, 500 (2023).
- [128] J. Tindall, M. Fishman, E. M. Stoudenmire, and D. Sels, Efficient tensor network simulation of IBM's Eagle kicked Ising experiment, *PRX Quantum* **5**, 010308 (2024).
- [129] Y. Zhou, E. M. Stoudenmire, and X. Waintal, What limits the simulation of quantum computers? *Phys. Rev. X* **10**, 041038 (2020).
- [130] S. Patra, S. S. Jahromi, S. Singh, and R. Orús, Efficient tensor network simulation of IBM's largest quantum processors, *Phys. Rev. Res.* **6**, 013326 (2024).
- [131] Locally entanglement optimal code repository, <https://github.com/SimonCichy/locally-entanglement-optimal-unraveling>.