Approximate inverse measurement channel for shallow shadows

Riccardo Cioli[1], Elisa Ercolessi[1], Matteo Ippoliti[2], Xhek Turkeshi[3], and Lorenzo Piroli[1]

[1]Dipartimento di Fisica e Astronomia, Università di Bologna and INFN, Sezione di Bologna, via Irnerio 46, I-40126 Bologna, Italy

[2]Department of Physics, The University of Texas at Austin, Austin, TX 78712, USA

[3]Institut für Theoretische Physik, Universität zu Köln, Zülpicher Strasse 77a, 50937 Köln, Germany

Classical shadows are a versatile tool to probe many-body quantum systems, consisting of a combination of randomised measurements and classical post-processing computations. In a recently introduced version of the protocol, the randomization step is performed via unitary circuits of variable depth t, defining the so-called shallow shadows. For sufficiently large t, this approach allows one to get around the use of non-local unitaries to probe global properties such as the fidelity with respect to a target state or the purity. Still, shallow shadows involve the inversion of a many-body map, the measurement channel, which requires non-trivial computations in the post-processing step, thus limiting its applicability when the number of qubits N is large. In this work, we put forward a simple approximate post-processing scheme where the infinite-depth inverse channel is applied to the finite-depth classical shadows and study its performance for fidelity and purity estimation. The scheme allows for different circuit connectivity, as we illustrate for geometrically local circuits in one and two spatial dimensions and geometrically non-local circuits made of two-qubit gates. For the fidelity, we find that the resulting estimator coincides with a known linear cross-entropy, achieving an arbitrary small approximation error δ at depth $t = O(\log(N/\delta))$ (independent of the circuit connectivity). For the purity, we show that the estimator becomes accurate at a depth $O(N)$. In addition, at those depths, the variances of both the fidelity and purity estimators display the same scaling with N as in the case of global random unitaries. We establish these bounds by analytic arguments and extensive numerical computations in several cases of interest. Our work extends the applicability of shallow shadows to large system sizes and general circuit connectivity.

1 Introduction

The successful implementation of many ideas in quantum information science depends on our ability to probe certain properties of many-body quantum states efficiently. Yet, given a large quantum system of N qubits, performing complete tomography of its state is often unfeasible, requiring a number of measurements scaling exponentially in N [1–3]. As larger devices and platforms become available within existing quantum technology [4–13], it is thus very important to find efficient protocols to probe selected properties of many-body states, without making use of full-state tomography.

Much recent progress in this direction has followed the development of the so-called randomised-measurement (RM) toolbox [14–17]. A prominent element in this toolbox is the classical-shadow approach [15]. It consists of applying random unitary operators, drawn from some suitable ensemble, to each copy of an unknown quantum state ρ, which is then measured on the computational basis. Notably, the measurements need not be tailored to a specific property of the system. Rather, the measurement outcomes can be processed to build estimators for different quantities of interest [18–21].

The ensemble of random unitaries determines the number of measurements needed to estimate a

given quantity accurately [16]. The two original versions of the protocol [15] involve either random single-qubit rotations, defining the so-called random Pauli measurements, or global unitaries drawn randomly from the set of Clifford operators [22]. While the former is suited to probe local properties such as few-qubit correlations [15, 23], the use of global Clifford unitaries outperforms the local Pauli measurements to estimate global quantities. Two important examples are the fidelity with respect to a given target state, an essential building block for quantum-state certification [24–30], and the purity [31]. Unfortunately, realizing global unitaries is nontrivial in practice and represents a bottleneck for current noisy intermediate-scale quantum (NISQ) devices [6].

From the experimental point of view, a less demanding protocol can be obtained by replacing the global operators with quantum circuits made of random two-qubit gates, yielding the so-called shallow shadows [32–35]. For sufficiently large depth t, quantum circuits allow one to get around the use of global unitaries to estimate global quantities, making them appealing for current experimental implementation [36, 37]. For instance, it was shown that this scheme allows us to accurately estimate the fidelity with $O(1)$ measurements when the circuit depth scales as $O(\log N)$ [32]. Yet, the protocol requires non-trivial computations at the post-processing step, as one needs to invert a many-body map, the *measurement channel*. While efficient tensor-network-based algorithms for this task were provided for shallow one-dimensional (1D) circuits [32, 33], their implementation is technically nontrivial, and no efficient algorithms are known beyond the 1D architecture, limiting the method's applicability. Other strategies to overcome the inversion of the measurement channel have been developed, see for example Refs. [38, 39].

In this work, we put forward an approximate inversion of the measurement channel, consisting of applying the infinite-depth inverse channel to the finite-depth classical shadows. We focus on the problem of estimating the global properties of the system, for which global Cliffords provide an advantage over random Pauli measurements, studying, in particular, the fidelity and the purity. The scheme involves simple post-processing computations and allows for different circuit connectivity, as we illustrate for geometrically local circuits in one and two spatial dimensions and geometrically non-local circuits made of two-qubit gates.

For the fidelity, we find that the resulting estimator coincides with a known linear cross entropy [40, 41], achieving an arbitrary small approximation error δ at depth $t = O(\log(N/\delta))$ – for all the circuit connectivities we considered. For the purity, we show that the estimator becomes accurate at a depth $O(N)$. In addition, the estimator variance for both the fidelity and purity at those depths displays the same scaling with N as that obtained via global random Cliffords. We establish these bounds by analytic arguments and extensive numerical computations in several cases of interest. Our results extend the applicability of shallow shadows to large system sizes and general circuit connectivity, with potential practical applications in current quantum computing platforms.

The rest of this work is organised as follows. We begin in Sec. 2 by reviewing classical shadows based on shallow circuits and presenting the approximate inversion formula. In Sec. 3, we study the resulting approximate fidelity estimator and the corresponding quantum-certification protocol, while in Sec. 4, we report our analysis of the approximate purity estimator. We summarize our conclusions and outline directions for future work in Sec. 6. The most technical aspects of our work are consigned to several appendices.

2 Approximate inversion formula

2.1 Classical shadows and the measurement channel

This section recalls some aspects of classical shadows used in our work. We consider a system of N qubits and denote by $|0\rangle_j$, $|1\rangle_j$ the elements in the local computational basis associated with spin j. Given M sequential experimental runs preparing a copy of ρ, we apply to each of them a random unitary operator sampled from a suitable ensemble $\mathcal{U}$, followed by a projective measurement in the computational basis $|b\rangle = \otimes_{j=1}^{N} |b_j\rangle_j$, with $b_j = 0, 1$. We will consider the ensemble of quantum circuits made of two-qubit Clifford gates. Namely, we will choose $U_t = V_t V_{t-1} \cdots V_1$, where t is the circuit depth, while each "layer" V_j contains quantum gates acting on disjoint pairs

of qubits. Each gate is drawn from the uniform distribution on the set of two-qubit Clifford unitaries [22, 42].

We consider two circuit architectures: geometrically local brickwork structures both in 1D and 2D and geometrically non-local structures. The former case consists of gates pairing qubits which are nearest neighbors on a regular D-dimensional lattice. For $D = 1$, the circuit alternates layers of gates acting on pairs of qubits at positions $(j, j+1)$, with j even and odd. For $D = 2$, we choose the same connectivity of Ref. [43], where each layer contains gates that pair a qubit with one of its four neighbors, cycling through the four neighbors over four layers. The geometrically non-local quantum circuits are instead constructed as follows: at each time step, we randomly partition the set of qubits into disjoint pairs and apply a random gate to each pair[1].

For each run of the experiment, one constructs a classical "snapshot" of the quantum state, $U_t^\dagger |b\rangle$, based on the choice of unitary U_t and the measurement outcome $b \in \{0,1\}^N$. Averaging over such snapshots yields the measurement channel

$$\mathcal{M}_t(\rho) = \mathop{\mathbb{E}}_{U_t \sim \mathcal{U}} \left[\sum_{b \in \{0,1\}^N} \left\langle b \left| U_t \rho U_t^\dagger \right| b \right\rangle U_t^\dagger |b\rangle\langle b| U_t \right] . \tag{1}$$

Assuming knowledge of $\mathcal{M}_t$ and (crucially) of its inverse $\mathcal{M}_t^{-1}$, one can then define the classical shadows

$$\rho_t^{(r)} = \mathcal{M}_t^{-1} \left[(U_t^{(r)})^\dagger |b^{(r)}\rangle\langle b^{(r)}| U_t^{(r)} \right] , \tag{2}$$

while $r \in \{1, \dots M\}$ labels the experimental runs.

Classical shadows are stored and processed classically to obtain an estimate of different quantities of interest. As mentioned, we will focus on two global quantities. First, we consider the fidelity with respect to a pure target state $|\psi\rangle$, for which we construct the estimator

$$\mathcal{F}_t^{(e)} = \frac{1}{M} \sum_{r=1}^{M} f_t^{(r)}, \tag{3}$$

where

$$f_t^{(r)} = \langle \psi | \rho_t^{(r)} | \psi \rangle . \tag{4}$$

[1] These circuits are sometimes called Brownian random circuits [44, 45]

Second, we study the purity of the system state ρ, corresponding to the estimator

$$\mathcal{P}_t^{(e)} = \frac{1}{M(M-1)} \sum_{r \neq r'} \mathrm{Tr}\left(\rho_t^{(r)} \rho_t^{(r')} \right) . \tag{5}$$

Using the definition of $\mathcal{M}_t$ in Eq. (1), it is straightforward to see that $\mathcal{F}_t^{(e)}$ and $\mathcal{P}_t^{(e)}$ are faithful, namely

$$\mathbb{E}[\mathcal{F}_t^{(e)}] = \mathcal{F}_\psi(\rho) = \langle \psi | \rho | \psi \rangle , \tag{6}$$

$$\mathbb{E}[\mathcal{P}_t^{(e)}] = \mathcal{P}(\rho) = \mathrm{Tr}\left[\rho^2 \right] , \tag{7}$$

where the average is over the random unitaries and the projective measurement outcomes.

The value of M guaranteeing a small deviation of $\mathcal{F}_t^{(e)}$ from $\mathcal{F}_\psi(\rho)$ is determined by the variance of the random variable $f^{(r)}$. In turn, the latter is bounded by the square of the *shadow norm* $||O||_{\mathrm{sh}}$ [15] of the traceless operator

$$O = |\psi\rangle\langle\psi| - \frac{\mathbb{1}}{2^N} , \tag{8}$$

which reads

$$||O||_{\mathrm{sh}}^2 = \max_{\rho \text{ state}} \left\{ \mathop{\mathbb{E}}_{U_t} \sum_{b \in \{0,1\}^N} \left[\left\langle b \left| U_t \rho U_t^\dagger \right| b \right\rangle \times \left\langle b \left| U_t \mathcal{M}_t^{-1}(O) U_t^\dagger \right| b \right\rangle^2 \right] \right\} . \tag{9}$$

Fixing O as in Eq. (8), the shadow norm depends on the unitary ensemble $\mathcal{U}$: $||O||_{\mathrm{sh}}^2$ grows exponentially in N in the case of local Pauli measurements, while $||O||_{\mathrm{sh}}^2 = O(1)$ for global Clifford unitaries [15]. Finite-depth circuits interpolate between these two extremes, recovering the global Clifford bound for $t \to \infty$. In fact, both numerical and analytical evidence show that the shadow norm associated with the fidelity estimator becomes $O(1)$ already at "shallow depth", namely $t \sim \log N$ [32–34].

Similarly, the number of measurements needed to achieve an accurate description of the purity depends on the variance of the estimator $\mathcal{P}_t^{(e)}$. In this case, explicit bounds are known for the two limiting cases of Pauli measurements (corresponding to $t = 0$) and global Cliffords ($t = \infty$). In the former case, we have [14, 19, 46]

$$\mathrm{Var}\left[\mathcal{P}_0^{(e)} \right] \leq 4 \left(\frac{2^N \mathcal{P}(\rho)}{M} \right) + 2 \left(\frac{2^{2N}}{M-1} \right)^2 , \tag{10}$$

while in the latter case [15]

$$\mathrm{Var}\left[\mathcal{P}_{\infty}^{(e)}\right] \leq \left(\frac{12\mathcal{P}(\rho)}{M}\right) + \frac{2\left(9\times 2^{2N}\right)}{(M-1)^2}\,, \quad (11)$$

where we have neglected terms that are exponentially small in N. We see that the bound on the variance of $\mathcal{P}_{\infty}^{(e)}$ grows exponentially in N even for global random Cliffords. Still, the leading term for a fixed M grows as $O(2^{2N})$, providing an advantage over the scaling $O(2^{4N})$ obtained for random Pauli measurements.

Crucially, shallow shadows require the computation of the inverse channel $\mathcal{M}_t^{-1}$. For global Clifford unitaries, a simple analytic formula is known [15]

$$\mathcal{M}_{\infty}^{-1}[\rho] = (2^N+1)\rho - \mathrm{Tr}(\rho)\mathbb{1}\,. \quad (12)$$

However, no analogous result is available for finite $t > 2$ [32, 34, 47]. So far, numerical algorithms based on tensor networks (TNs) [48] were developed to compute $\mathcal{M}_t^{-1}$ for 1D circuits, with a computational cost growing polynomially in N and exponentially in t. At the same time, the problem remains intractable in higher dimensions or more general circuit connectivity. Overall, computing $\mathcal{M}_t^{-1}$ makes the post-processing step highly nontrivial [32–34], limiting the scope and applicability of shallow shadows.

2.2 Approximate inversion formula

In this section, we describe our proposed scheme to get around the challenging problem of computing and inverting $\mathcal{M}_t$, thus significantly expanding the power of shallow shadows. The main conceptual ingredient of our protocol is to simply replace the inverted measurement channel $\mathcal{M}_t^{-1}$ by its infinite-depth limit, $\mathcal{M}_{\infty}^{-1}$. That is, we introduce the approximate shadows

$$\tilde{\rho}_t^{(r)} = \mathcal{M}_{\infty}^{-1}\left[(U_t^{(r)})^{\dagger}|b^{(r)}\rangle\langle b^{(r)}|U_t^{(r)}\right]\,. \quad (13)$$

This substitution gives rise to estimators which are not faithful for finite t, at which

$$\mathbb{E}[\tilde{\rho}_t^{(r)}] \neq \rho\,. \quad (14)$$

However, since we must recover $\mathcal{M}_{\infty}^{-1}[\mathcal{M}_t(\rho)] \simeq \rho$ for large t, we expect that the modified estimators become increasingly accurate as t increases. The question is then what is the minimum depth t^* at which the error is below some threshold δ.

Since the measurement channel depends on the unitary ensemble only through its second moment $\mathbb{E}[U_t\otimes U_t^*\otimes U_t\otimes U_t^*]$, one may guess that t^* is the time taken for the second moment of the shallow-circuit ensemble to equilibrate to that of the global random Clifford ensemble. The latter is the same as the second moment of the Haar distribution due to the 2-design property of the Clifford group [49, 50]. Thus, t^* would coincide with the depth at which random circuits become a 2-design [51, 52]. In the following sections, we will clarify this intuition. We will show in particular that the time t^* heavily depends on the quantity under consideration, the fidelity and purity estimation requiring a depth $O(\log N)$ and $O(N)$, respectively.

Before leaving this section, it is important to comment on the fidelity estimator, which is obtained using the approximate inversion formula. Combining Eqs. (3) and (13), we obtain

$$\tilde{\mathcal{F}}_t^{(e)} = \frac{1}{M}\sum_{r=1}^{M}\tilde{f}_t^{(r)}\,, \quad (15)$$

where

$$\tilde{f}_t^{(r)} = \langle\psi|\mathcal{M}_{\infty}^{-1}\left[(U_t^{(r)})^{\dagger}|b^{(r)}\rangle\langle b^{(r)}|U_t^{(r)}\right]|\psi\rangle\,. \quad (16)$$

Interestingly, we can show that the expectation value of $\tilde{f}_t^{(r)}$ is closely related to a known function called linear cross entropy [40, 41]. To this end, we denote by $p_{\rho,U_t}(b)$ and $p_{\psi,U_t}(b)$, respectively, the probabilities to obtain the outcomes b when the system is initialized in the states ρ and $|\psi\rangle$ and evolved by U_t. The average of $\tilde{f}_t^{(r)}$ can then be rewritten as

$$\mathbb{E}\left[\tilde{f}_t^{(r)}\right] = \mathbb{E}_{U_t\in\mathcal{U}}H_{\mathrm{lin}}[p_{\rho,U_t}, p_{\psi,U_t}]\,, \quad (17)$$

where

$$H_{\mathrm{lin}}[p,q] = 2^N\sum_b p(b)q(b) - 1 \quad (18)$$

is the linear cross-entropy. The linear cross entropy lies at the basis of the linear-cross-entropy benchmark (XEB). In this setting, one typically considers a trivial initial product state $|0\rangle$ and wants to certify the quantum circuit applied to it, which is assumed to be noisy [40, 41, 53–58]. Namely, XEB is typically used as a means to certify the fidelity of gates, rather than to perform quantum-state certification. In this work,

instead, we assume that the gates are noiseless and Eq. (16) is thus viewed as a tool to estimate the fidelity between the physical and target states.

We also mention an interesting connection with the protocol of filtered randomized benchmarking [59]. Ref. [60], in particular, used the inverse of the Haar-random measurement channel to define an estimator for the so-called filtered randomized benchmarking signal, which was shown to coincide with the linear-cross entropy.

Finally, we mention that $f_t^{(r)}$ is also related to another estimator considered in the literature. In particular, Ref. [56] introduced the normalized estimator

$$F_c = 2\frac{\sum_z p_{\rho,U_t}(z)p_{\psi,U_t}(z)}{\sum_z p_{\rho,U_t}(z)^2} - 1\,, \qquad (19)$$

see also [56, 61, 62]. F_c was employed in the context of analog chaotic dynamics and it was demonstrated to estimate the fidelity at infinite effective temperature and short quench evolution times.

3 Approximate fidelity estimator

We now study the performance of the approximate fidelity estimator. To be concrete, we consider a quantum-certification protocol where the system is prepared in an unknown quantum state ρ. Given the target state $|\psi\rangle$, and any certification threshold $0 < \varepsilon < 1$, we want to determine whether $\mathcal{F}_\psi(\rho)$ is larger or smaller than $1-\varepsilon$.

3.1 Gate complexity

In order to establish the usefulness of the estimator $\tilde{\mathcal{F}}_t^{(e)}$, we first need to determine the depth at which the approximation error becomes small. To build intuition, we start from the ideal situation where the lab and target states coincide, $\rho = |\psi\rangle\langle\psi|$, and ask what is the minimum depth $t_0^*(N,\delta)$ such that

$$|\mathbb{E}[\tilde{\mathcal{F}}_t^{(e)}] - 1| \leq \delta\,, \qquad (20)$$

for all $t \geq t_0^*(N,\delta)$, where $\delta > 0$ is an arbitrary number that should be taken smaller than the certification threshold ε. This problem can be solved analytically when $|\psi\rangle$ is sufficiently simple, and in particular when $|\psi\rangle = \otimes_j |\phi_j\rangle$ is a product state. To this end, we use the 2-design property of the Clifford group [63], stating that the averages over random Clifford and Haar random gates are equal. Then, exploiting the invariance of the Haar measure with respect to single-qubit rotations, we derive in Appendix A.1 the following result

$$\mathbb{E}\left[\tilde{\mathcal{F}}_t^{(e)}\right] = (2^N+1)2^N Z_t - 1\,, \qquad (21)$$

where $Z_t = \mathbb{E}_{U_t}\left[|\langle 0|U_t|0\rangle|^4\right]$, while $|0\rangle = |0\rangle^{\otimes N}$ and the average is over Haar random circuits of depth t.

The average Z_t can be analyzed employing recently-introduced mappings to statistical-mechanics partition functions [64, 65]. In fact, Z_t has been previously studied in the context of anticoncentration of random quantum circuits [66, 67]. In particular, Ref. [67] provided rigorous lower and upper bounds on Z_t which directly imply

$$1 \leq \mathbb{E}\left[\tilde{\mathcal{F}}_t^{(e)}\right] \leq 1 + 2e^{-a(t-T_N)} \qquad (22)$$

where the constant a and T_N (which depends on N) are determined by the circuit architecture. For 1D brickwork circuits, $a = \log(5/4)$ and $T_N = [\log N + \log(e-1)]/a + 1$, which finally implies $t_0^*(N,\delta) \leq g(N,\delta)$, where

$$g(N,\delta) = \frac{\log(N/\delta)}{\log(5/4)} + \frac{\log[2(e-1)]}{\log(5/4)} + 1\,, \qquad (23)$$

yielding the anticipated logarithmic dependence on N. The same scaling $T_N \sim \log N$ (with a different prefactor) follows from analogous bounds on Z_t, which are proven rigorously for non-local quantum circuits and conjectured for more general architecture (including the $2D$ case) with minimal connectivity requirements [67].

In principle, one could extend the derivation of Ref. [66,67] to study more structured target states $|\psi\rangle$, but this is technically challenging. Here, we follow a different approach and provide a simple argument to establish a bound on $t^*(N,\delta)$ for *typical* target states.

Focusing first on 1D architectures, we model a typical target state as $|\psi\rangle = W_\tau|0\rangle$, where W_τ is a 1D circuit of depth τ. The value of τ is arbitrary, but we assume that the circuit architecture is the same and that the gates are drawn from a Haar random distribution. This model is expected to capture the behavior of generic many-body quantum states away from fine-tuning.

For fixed W_τ, the computations leading to Eq. (21) yield

$$\mathbb{E}_{U_t}\left[\tilde{\mathcal{F}}_t^{(e)}\right] = (2^N+1)2^N\mathbb{E}\left[|\langle 0|U_tW_\tau|0\rangle|^4\right] - 1\,. \tag{24}$$

We extract the typical behavior by averaging over the circuits W_τ defining the target state and obtain

$$\mathbb{E}\left[\tilde{\mathcal{F}}_t^{(e)}\right] = (2^N+1)2^N Z_{t+\tau} - 1\,. \tag{25}$$

Denoting by $t^*_\tau(N,\delta)$ the time at which $|\mathbb{E}[\tilde{\mathcal{F}}_t^{(e)}] - 1| \leq \delta$ for the target state $|\psi\rangle = W_\tau|0\rangle$, Eq. (25) immediately implies that $t^*_\tau(N,\delta) \leq g(N,\delta) - \tau$, where g is defined in Eq. (23). That is, for typical states of constant complexity τ (modeled by Haar-random brickworks of depth τ acting on a product state), the bound on t^* is smaller by an additive constant compared to that of product states. This is intuitive: we have related the accuracy of the estimator to the value of the anti-concentration parameter Z_t of the evolved state—the estimator becomes accurate when the evolved state's wavefunction is spread uniformly over the computational basis [67]; but since typical product states are more concentrated than typical time-evolved states, more "scrambling" is needed to achieve this condition when starting from product states. We can repeat this discussion for more general architectures, finding the same result.

The previous arguments apply when the lab and target states coincide, but this is never realised in practice. In fact, ρ will generally be mixed with $\mathcal{F}_\psi(\rho) < 1$, making the analysis more complicated. In addition, we have always considered typical target states, but one could wonder whether our bounds apply to structured or fine-tuned target states. We tackle these questions numerically, and in Sec. 3.3 we provide evidence showing that our bounds remain indeed valid in these situations and that the depth t^* displays a logarithmic scaling for all target states and circuit architectures.

We conclude this section with an important remark. Although our post-processing scheme gets around the inversion problem, evaluating the estimator (16) still requires one to compute amplitudes of the form $\langle\psi|(U_t^{(r)})^\dagger|b^{(r)}\rangle$. For generic states, the cost of this classical computation grows exponentially in N. However, the estimator can be computed efficiently (namely, in a time growing only polynomially in N) for different classes of target states. For 1D connectivity, this can be done, for instance, for TN states including Matrix Product States (MPS) [68–70] or tree tensor network states [48]. In addition, in all circuit architectures, the amplitude can be computed efficiently for stabilizer states exploiting the classical simulability of Clifford circuits [22, 42]. This class includes the so-called Greenberger-Horne-Zeilinger (GHZ) [71] and cluster state [72] as special examples. We stress that the restriction to certain classes of target states is a common feature of efficient quantum certification protocols [73].

3.2 Sample complexity

In addition to the error due to the approximate inversion formula, one also needs to take into account the fluctuations of $\tilde{f}_t^{(r)}$ or, equivalently, of the estimator $\tilde{\mathcal{F}}_t^{(e)}$ (the variances of the two random variables are related by a factor $1/M$). In particular, the efficiency of the method depends on the possibility of bounding the variance of $\tilde{\mathcal{F}}_t^{(e)}$ by a constant independent of N.

In standard classical shadow protocols, the estimator variances are bounded by the corresponding shadow norm squared (9). This object may be hard to study in general because it involves the third moment of the distribution over the random unitaries U_t. A more practical quantity is obtained by replacing the maximization in (9) by the average over all states ρ, with a uniform measure (or any 1-design measure). This defines the *locally-scrambled* shadow norm [47, 74], which is used as a proxy for $||O||_{\text{sh}}$. For shallow shadows based on 1D circuits, the locally-scrambled shadow norm was studied in Refs. [32–34] and was proven to be $O(1)$ for circuits of depth $O(\log N)$ [32].

In Appendix A.2, we show that the variance of $\tilde{f}_t^{(r)}$ is bounded by an expression obtained replacing the exact inversion channel in (9) by $\mathcal{M}_\infty^{-1}$. Following Refs. [32, 33, 47, 74], we then study its locally-scrambled version

$$||O||^2_{\text{sh,LS}}(t) = \frac{1}{2^N}\mathop{\mathbb{E}}_{U_t}\sum_{b\in\{0,1\}^N}\left[\left\langle b\left|U_t\mathcal{M}_\infty^{-1}(O)U_t^\dagger\right|b\right\rangle^2\right]. \tag{26}$$

This expression does not depend on ρ and can be studied by repeating the arguments reported in

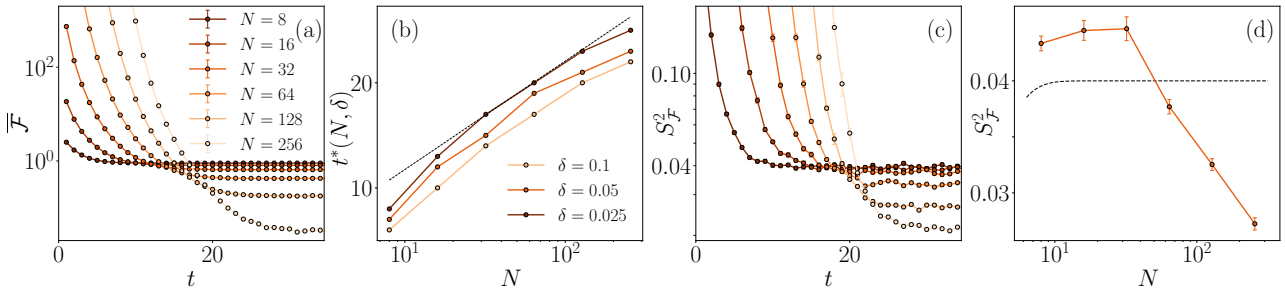

Figure 1: Numerical simulation of the certification protocol for 1D local circuits. (a): Averaged fidelity estimator $\overline{\mathcal{F}} = \mathbb{E}[\tilde{\mathcal{F}}_t^{(e)}]$ as a function of the circuit depth, for increasing N. Here, the target state $|\psi\rangle$ is the GHZ state and $\rho = \mathcal{E}_p^{\otimes N}(|\psi\rangle\langle\psi|)$, where $\mathcal{E}_p$ is the depolarizing channel with $p = 0.02$. Error bars associated with statistical fluctuations are small and not visible in the scales of the plots. (b): Depth $t^*(N,\delta)$ (defined in the main text) as a function of the system size N. The dashed line is the function in Eq. (23), shifted by a negative constant. (c): Variance of fidelity estimator $S^2_{\mathcal{F}} = \mathbb{E}[(\tilde{\mathcal{F}}_t^{(e)} - \mathcal{F}_\psi(\rho))^2]$ as a function of the depth. (d): The variance $S^2_{\mathcal{F}}$ estimated at the depth $t^*(N,\delta)$, for $\delta = 0.05$. The dashed line is the exact variance for the global Clifford protocol when the lab state coincides with the target state.

Sec. 2.2.

Let us first consider the case where the target state is a product state, $|\psi\rangle = \otimes_j |\phi_j\rangle_j$. We can repeat the derivation leading to Eq. (21), and by averaging over the random circuits U_t, we arrive at

$$||O||^2_{\text{sh,LS}}(t) = (2^N+1)^2\left[Z_t - \frac{1}{2^{2N}}\right]. \tag{27}$$

Therefore, $||O||_{\text{sh,LS}}$ depends on the same term Z_t appearing in the average in Eq. (21). Now, recalling that $t_0^*(N,\delta)$ is defined as the minimum depth at which Eq. (20) is satisfied, it follows from Eq. (21) that

$$||O||^2_{\text{sh,LS}}(t^*(N,\delta)) \leq \frac{3}{2}(2+\delta) = O(1). \tag{28}$$

Following the previous section, we can finally invoke a typicality argument and provide a similar bound for generic states of the form $|\psi\rangle = W_\tau|0\rangle$ where W_τ is a random quantum circuit of depth τ.

As mentioned, the locally-scrambled shallow norm is only a proxy for $||O||_{\text{sh}}$. However, we expect that the $O(1)$ bound found above also holds for the actual shadow norm and thus for the variance of our estimator $\tilde{\mathcal{F}}_t^{(e)}$ whenever $t \geq t^*$. We substantiate this claim via extensive numerical evidence for different circuit architectures and target states below.

3.3 Numerical results

In this section, we present numerical evidence supporting our claims and illustrate our scheme with explicit examples. We focus on two classes of states for which we can perform large-scale computations: 1D states and stabilizer states, both in one and two dimensions. For 1D states, we use an efficient method to compute the exact average over infinitely many measurement outcomes. This is based on the statistical mechanics mapping explained in Refs. [66, 67] and numerical MPS computations. For stabilizer states, we use the classical simulability of Clifford circuits [22, 42, 75] to simulate the full quantum-certification protocol. We refer to Appendix A.3 for details on the numerical implementations, while we report representative examples of our data in Fig. 1–2: the data shown in these figures are obtained from Clifford simulations and checked independently against the MPS algorithm when possible.

We first consider the case where the lab state is a mixed state $\rho \neq |\psi\rangle\langle\psi|$. In this case, we define analogously $t^*(N,\delta)$ as the minimum depth such that for $t \geq t^*$ one has $\Delta F_{\psi,\rho} \leq \delta$, where

$$\Delta F_{\psi,\rho}(t) = |\mathbb{E}[\tilde{\mathcal{F}}_t^{(e)}] - \mathcal{F}_\psi(\rho)|. \tag{29}$$

In Fig. 1, the target state is the GHZ state,

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}), \tag{30}$$

while the lab state ρ is taken to be $\rho = \mathcal{E}_p^{\otimes N}(|\psi\rangle\langle\psi|)$ where

$$\mathcal{E}_p(\sigma) = (1-p)\sigma + p\text{Tr}[\sigma]\frac{\mathbb{1}}{2}, \tag{31}$$

is the depolarizing channel. Fig. 1(a) shows $\mathbb{E}[\tilde{\mathcal{F}}_t^{(e)}]$, for increasing N, while we report in Fig. 1(b) the value $t^*(N,\delta)$ at which $\Delta F_{\psi,\rho}(t) \leq \delta$, where $\Delta F_{\psi,\rho}(t)$ is defined in Eq. (29). We note that the GHZ state is a fine-tuned state, displaying long-range correlations and thus outside the scope of our analytical arguments. Yet, our numerical results confirm a logarithmic scaling of $t^*(N,\delta)$, for arbitrary small δ. In fact, Fig. 1(b) displays the asymptotic scaling of t^* predicted by Eq. (23) (shifted by a negative constant), showing that it upper bounds $t^*(N,\delta)$ even though $|\psi\rangle$ is not a product state.

Going beyond the mean of the estimator, Fig. 1(c) and (d) show, respectively, the variance

$$S^2_{\mathcal{F}} = \mathbb{E}\left[\left(\tilde{\mathcal{F}}_t^{(e)} - \mathcal{F}_\psi(\rho)\right)^2\right], \tag{32}$$

as a function of circuit depth t and its value at the depth $t^*(N,\delta)$. The data clearly show that the variance quickly drops below an N-independent value. Interestingly, the latter is quantitatively close to that obtained in the protocol based on global Clifford unitaries, as we show in Appendix A.2. In the 1D case, we have further studied short-ranged correlated states, finding similar results, cf. Appendix A.3 for additional numerical data.

Next, we report data for circuit architectures beyond 1D in Fig. 2. Figs. 2(a,b) focus on local circuits on a 2D square lattice of side L. The target state is the cluster state,

$$|\psi\rangle = \prod_{\langle i,j\rangle} CZ_{i,j} \, |+\rangle^{\otimes N}, \tag{33}$$

where the product is over nearest-neighbor qubits on the $2D$ lattice, while $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and we assume $\rho = |\psi\rangle\langle\psi|$. We are able to perform simulations for square circuits up to $N = L^2 = 784$ qubits. The depth $t^*(N,\delta)$ appears almost independent of the system size. We conjecture that this is a finite-size effect and that a logarithmic growth would be eventually observed at even larger values of N. Nevertheless, it is remarkable that an accuracy above 95% can be obtained for $N \sim 800$ qubits by depth $t \lesssim 10$.

Finally, in Figs. 2(c,d) we study the protocol based on geometrically non-local quantum circuits, where the target state is the GHZ state. The plots clearly show that $t^*(N,\delta)$ increases at most logarithmically in N. As expected, we find that the non-local circuit requires a smaller depth in order to achieve the same accuracy for the GHZ state, compared to 1D local circuits. Again, this is intuitive, as the validity of our method relies on anticoncentration in quantum circuits, which happens faster on all-to-all connected circuit architectures [67]. For both the 2D and non-local circuits, the plots show that the variance approaches a finite constant at a depth $t^*(N,\delta)$ for arbitrary δ. In summary, all cases that we have investigated fully support the claims presented in Sec. 2.2 about the efficiency of the protocol in terms of both gate complexity and sample complexity, even beyond the domain of validity of our analytical arguments.

In all our simulations, our estimator (3) is built choosing $M = 50$, while its average and variance are estimated sampling over $R \simeq 10000$ independent values. To be more precise, the total number of simulated experimental runs is MR. The outcomes are organized into R subsets containing M shadows each, which are used to compute an instance of the estimator (3). We stress that number of measurements performed in our simulations is of the same order as, or smaller than, the typical number of measurements performed in present-day trapped ion and superconducting qubit experiments [46, 76–80], thus substantiating our claims on the practical feasibility of our scheme.

4 Approximate purity estimator

In this section, we study the performance of the approximate inversion formula for purity estimation. We consider, in particular, the estimator

$$\tilde{\mathcal{P}}_t^{(e)} = \frac{1}{M(M-1)} \sum_{r \neq r'} \mathrm{Tr}\left(\tilde{\rho}_t^{(r)} \tilde{\rho}_t^{(r')}\right), \tag{34}$$

where $\tilde{\rho}_t^{(r)}$ is defined in Eq. (13). Being a quadratic function of $\tilde{\rho}_t$, this estimator is harder to treat analytically than the fidelity estimator, and its average cannot straightforwardly be related to the anti-concentration term Z_t defined in Eq. (21). Therefore, we study its behavior numerically, restricting for simplicity to the case of 1D and all-to-all circuits. As for the fidelity, we need to consider both the approximation error given by the approximate inversion formula and the statistical error associated with the random

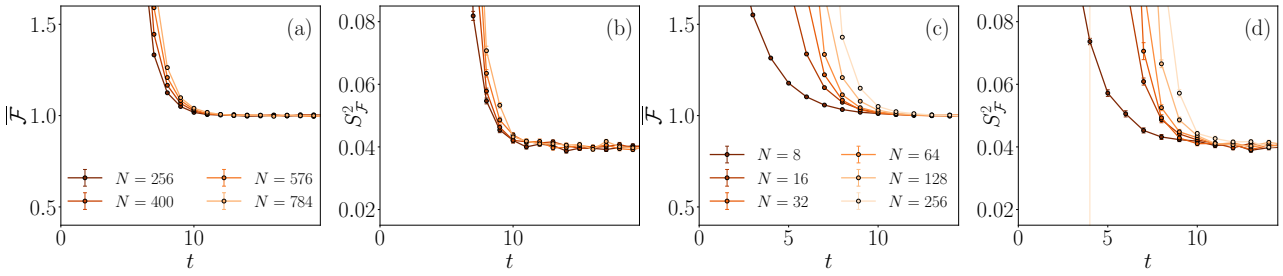

Figure 2: Numerical simulation of the certification protocol for 2D local circuits [(a)-(b)] and non-local quantum circuits [(c)-(d)]. (a): Averaged fidelity estimator $\overline{\mathcal{F}} = \mathbb{E}[\tilde{\mathcal{F}}]$ as a function of the circuit depth, for increasing N. Here, the target state $|\psi\rangle$ is the $2D$ cluster state state and $\rho = |\psi\rangle\langle\psi|$. The geometry is that of a square of side $L = \sqrt{N}$. (b): Variance of the fidelity estimator $S^2_{\mathcal{F}} = \mathbb{E}[(\tilde{\mathcal{F}} - \mathcal{F}(\rho))^2]$ as a function of the depth for the same state. (c): Averaged fidelity estimator $\overline{\mathcal{F}} = \mathbb{E}[\tilde{\mathcal{F}}]$ as a function of the circuit depth, for increasing N. Here, the target state $|\psi\rangle$ is the GHZ state and $\rho = |\psi\rangle\langle\psi|$. (d): Variance of the fidelity estimator $S^2_{\mathcal{F}} = \mathbb{E}[(\tilde{\mathcal{F}} - \mathcal{F}(\rho))^2]$ as a function of circuit depth, for the same state.

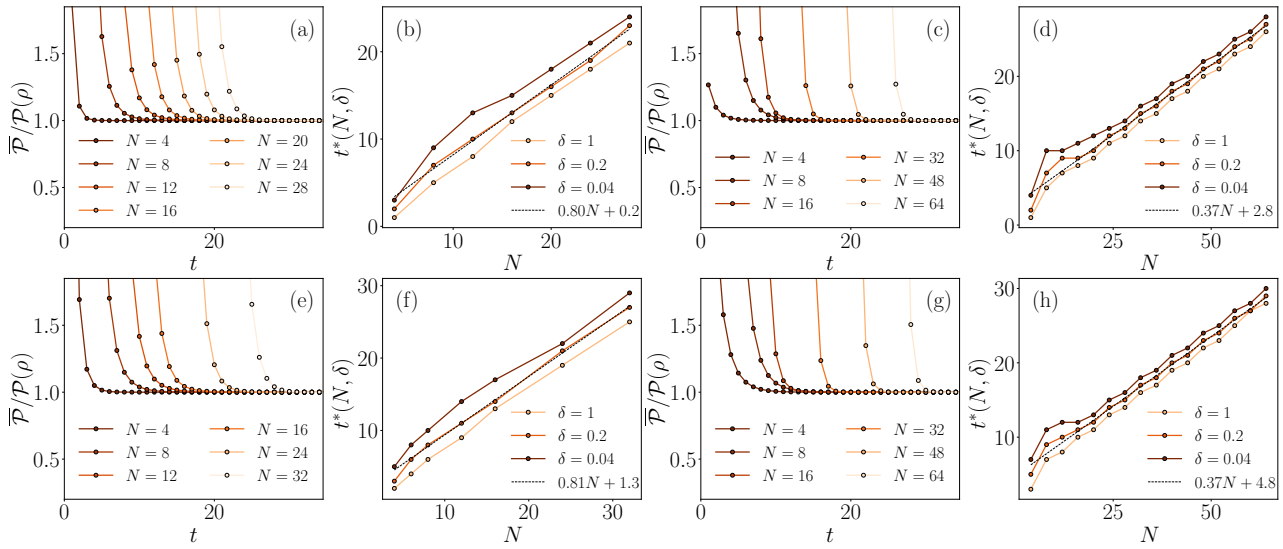

Figure 3: Numerical results for the averaged purity estimator for 1D local [(a),(b),(e),(f)] and non-local circuits [(c),(d),(g),(h)]. In panels [(a)-(d)] the state ρ is the GHZ state (30), while in [(e)-(h)] it is the state given in Eqs. (37), (38) with $\mu = 0.05$. Panels (a), (c), (e) and (g) show the ratio between the expected value of the purity estimator $\overline{\mathcal{P}}$ and the exact value of the purity $\mathcal{P}(\rho)$ as a function of circuit depth, for increasing N. Panels (b), (d), (f) and (h) show the depth $t^*(N,\delta)$ defined in (36).

measurement outcomes. We present our results separately in the following.

The error due to the approximate inversion formula depends on the average

$$\overline{\mathcal{P}}_t = \mathbb{E}\left[\tilde{\mathcal{P}}_t^{(e)}\right], \tag{35}$$

or equivalently on $\mathbb{E}[\mathrm{Tr}(\tilde{\rho}_t^{(r)}\tilde{\rho}_t^{(r')})]$. Since the purity $\mathcal{P}(\rho)$ is typically exponentially small in the system size N, it is natural to quantify the deviation of our estimator from the true value via the relative error. Accordingly, we define $t^*(N,\delta)$ as the minimum depth such that

$$\left|\frac{\overline{\mathcal{P}}_t}{\mathcal{P}(\rho)} - 1\right| \leq \delta\,, \tag{36}$$

for all $t \geq t^*(N,\delta)$.

We compute $\overline{\mathcal{P}}_t$ using numerically exact methods yielding the average over infinitely many shadows. As explained in Appendix B, these methods are based on a standard replica formalism [64, 65], mapping the problem onto a non-unitary dynamics over N qubits. In the 1D geometry, the latter dynamics can be computed efficiently using TN algorithms. In the non-local geometry, there is an emergent permutation symmetry, allowing us to perform computations up to very large system sizes, cf. Appendix B.2. In the following, we report the data obtained using these methods. We also mention that, when

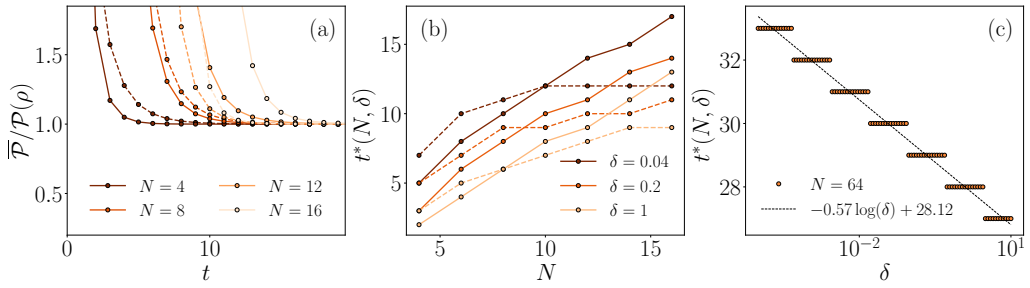

Figure 4: Comparison of the 1D local architecture (solid lines) with the non-local one (dashed lines) for small system sizes N. Here the state ρ is the product state given in Eqs. (37), (38) with $\mu = 0.05$. Panel (c) shows the scaling of $t^*(N,\delta)$ with δ for $N = 64$.

possible, we have checked these results against a direct simulation of the full purity-estimation protocol using the stabilizer formalism (namely, collecting a finite number of shadows and averaging over them). However, since the variance of the purity estimator grows exponentially in the number of qubits N, cf. (11), this approach is limited to rather small system sizes.

We study two examples. First, we consider the GHZ state (30), for which the purity is one, and the product state

$$\rho = \bigotimes_{j=1}^{N} \rho_j \,, \tag{37}$$

where

$$\rho_j = \cos^2(\mu)|0\rangle\langle 0| + \sin^2(\mu)|1\rangle\langle 1| \,, \tag{38}$$

for which the purity is

$$\mathcal{P}(\rho) = (\cos^4(\mu) + \sin^4(\mu))^N \,. \tag{39}$$

In Fig. 3 we show our results for the GHZ and product state (37), respectively in panels (a)-(d) and (e)-(h). In each row, the first two panels correspond to the 1D local architecture, while the last two to the non-local circuits. Panels (a) and (c), (e) and (g) show the ratio between the expectation value of the purity estimator and the true value as a function of the circuit depth for increasing values of N. Interestingly, we find that for both circuit architectures and all considered states, $t^*(N,\delta)$ is linear in N, with a coefficient that is fit to 0.81 ± 0.02 for 1D local circuits and 0.37 ± 0.01 for non-local ones, for both the GHZ state and the product state with $\mu = 0.05$ (for increasing - but small - values of μ we find that the linear coefficient increases slightly). This implies that the non-local architecture performs better for large values of N, as expected. For small N, however, we show in Fig. 4 that this is not necessarily true. We stress that in all cases the leading term in N of $t^*(N,\delta)$, which is proportional to N, appears to be independent of δ, at least for the values of N that we could simulate. More quantitatively, Fig. 4(c) shows the dependence of $t^*(\delta,N)$ on δ (for fixed N). The plot is consistent with the scaling $t^*(N,\delta) \propto \log(1/\delta)$.

Next, we study the variance of the purity estimator

$$S^2_{\mathcal{P}} = \mathbb{E}[(\tilde{\mathcal{P}}^{(e)}_t - \mathbb{E}[\tilde{\mathcal{P}}^{(e)}_t])^2] \,, \tag{40}$$

which determines the sample complexity of the protocol. For both the GHZ and the product state (37), we simulate the sampling of shadows using efficient simulations of Clifford circuits and report our results in Fig. 5. For each row, the first two panels correspond to the 1D local architecture, while the last two to the non-local circuits. Panels (a), (c), (e) and (g) show $S^2_{\mathcal{P}}$ as a function of circuit depth, for increasing values of N, while in (b), (d), (f) and (h) we report the same variance evaluated at $t^*(N,\delta)$, as a function of N. The plots show that in both cases, the variance approaches the global-Clifford value $\mathrm{Var}[\tilde{\mathcal{P}}^{(e)}_\infty]$ at the depth $t^*(N,\delta)$, for arbitrary δ. We can compute this variance analytically, improving the bound (11) derived for general observables in Ref. [15], cf. Appendix B.3. In the large-N limit, it simplifies to

$$\mathrm{Var}[\tilde{\mathcal{P}}^{(e)}_\infty] \sim \frac{2(2^{2N})}{M(M-1)} \,. \tag{41}$$

As already anticipated, the variance grows exponentially in the system size, but it is still a quadratic improvement over the $O(2^{4N})$ scaling of the random Pauli measurements protocol. This makes our protocol better suited to estimate subsystem purities, which have been used to detect

pure state entanglement [46,76–79] for small system sizes.

It is also worth mentioning that our approximate method could even be useful to measure entanglement-related quantities for large system sizes, under some additional assumptions. For instance, Ref. [81] introduced a protocol to measure the bipartite entanglement of pure states in the absence of long-range correlations, using only polynomially many measurements. In that protocol, the global purity is reconstructed starting from the purities of finite subregions. Ref. [81] proposed to measure such purities using local Pauli measurements, but one could instead use our approximate scheme to make the protocol even more efficient (the number of measurements would remain polynomial, but with an improved scaling).

In all the Clifford simulations, the estimator (5) is constructed choosing $M = 50$, while its average and variance are estimated by sampling over $R \approx 90000$ independent values. This specific value of R was chosen to guarantee a relative error on the purity estimator smaller than 10% for all system sizes up to $N = 10$.

5 Application to Pauli observables

Finally, we briefly discuss how the approximate inversion protocol performs in computing expectation values of Pauli observables. For a given Pauli P, we construct the approximate estimator

$$\tilde{P}_t^{(e)} = \frac{1}{M}\sum_{r=1}^{M} \mathrm{Tr}\left(P\tilde{\rho}_t^{(r)}\right), \tag{42}$$

with $\tilde{\rho}_t^{(r)}$ defined in Eq. (13). We consider the case where the lab state is the GHZ state, and study the behaviour of the protocol numerically using Clifford-circuit simulations. In the standard shadow protocols, it is known that qualitative differences in the scaling of the variance arise depending on the locality of the Pauli operators. Therefore, we study to examples of local and non-local observables, namely Z_1Z_2 and $X^{\otimes N}$, respectively. Our results are shown in Fig. 6. Once again, the first two panels of each row correspond to the 1D local architecture, while the last two to non-local circuits. Panels (a), (c), (e) and (g) show the behaviour of the expectation value of the approximate estimator as a function of circuit depth t and for increasing number of qubits N, while panels (b), (d), (f) and (h) show its variance.

In all cases, for large enough depths we find that the variance approaches the global-Clifford value of

$$\mathrm{Var}[\tilde{P}_\infty^{(e)}] = \frac{2^N + 1 - \langle P\rangle_\psi^2}{M}. \tag{43}$$

The $t \to \infty$ value of the variance was computed following the same steps as in Appendix A.2. Since the variance is exponentially large in N, we are limited to rather small system sizes. Therefore, we are not always able to carry out a precise analysis of the scaling of the circuit depth $t^*(N,\delta)$ at which the estimator becomes unbiased. Yet, our data show clearly that $t^*(N,\delta)$ is bounded by $O(N)$, which is consistent with the expected depth required for the random circuits to become a 2-design.

In fact, our data are consistent with a linear growth in N of $t^*(N,\delta)$ in the case of Z_1Z_2. We stress that for local observables the protocol involving local Pauli measurements is exponentially better than that of global random Cliffords, so we do not expect our inverse approximate method to be useful in this case. Conversely, for the non-local observable $X^{\otimes N}$, the asymptotic scaling of $t^*(N,\delta)$ is not clear from our limited system sizes. Yet, our data appear to be consistent with a sub-linear growth of $t^*(N,\delta)$ with N. In this case, the variance associated with the global Clifford protocol grows as $O(2^N)$, which is an improvement over the $O(3^N)$ scaling of local Pauli measurements, so our method could be practically useful in this case.

The estimator (42) is constructed with $M = 50$, while its average and variance are estimated sampling over $R = 36000$ independent values, such that the relative error on the Pauli estimator is smaller than 10% for all system sizes up to $N = 14$.

6 Outlook

We have introduced a simple approximate inversion formula for shallow shadows, and studied its performance for the global fidelity and purity estimation. Through analytic arguments and numerical computations, we have shown that the fidelity and purity estimators become accurate at depths scaling as $O(\log N)$ and $O(N)$, respectively. In addition, in both cases, we have shown

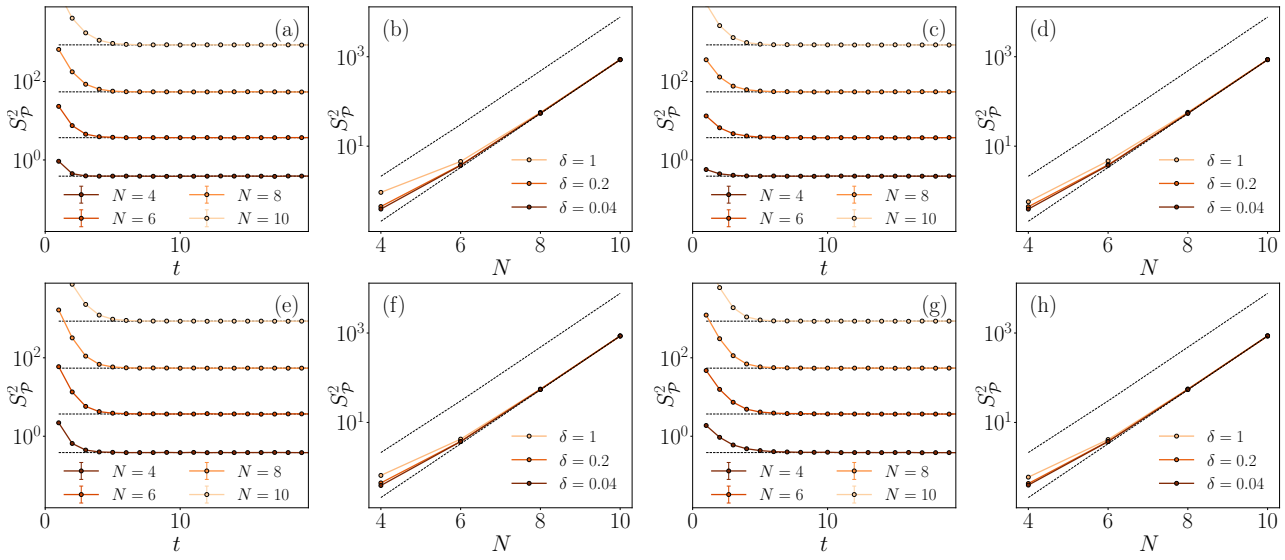

Figure 5: Numerical computation of the variance of the purity estimator for 1D local circuits [(a),(b),(e),(f)] and non-local ones [(c),(d),(g),(h)]. In panels [(a)-(d)] the state ρ is the GHZ state (30), while in [(e)-(h)] it is the state given in Eqs. (37), (38) with $\mu = 0.05$. Panels (a), (c), (e) and (g) show the sample variance as a function of circuit depth for increasing N. The dashed lines correspond to the $t \to \infty$ value in Eq. (86). Panels (b), (d), (f) and (h) display the variance at depth $t^*(N, \delta)$ for increasing N, for different values of δ. The lower dashed line represents the scaling in Eq. (41), while the upper one shows the upper bound in Eq. (11).

that the estimator variances at those depths display the same scaling with N as those obtained via global random unitaries. We have performed extensive numerical simulations showing that a modest amount of resources suffices in practice in many cases of interest, making the protocol relevant for current quantum computing platforms. We also mention that our approximate scheme may be useful beyond the applications illustrated in this work. For instance, it could be immediately combined with a recent scheme for efficient purity estimation in particular classes of many-body states [81] to reduce the sample complexity further.

Our work naturally raises the question of whether improved approximate inversion formulas can be found, possibly depending on the circuit architecture or the quantity to be estimated. Similarly to the one studied in our paper, two sensible requirements are that the approximate inversion formula can be easily computed in practice and that the corresponding estimators become accurate at modest circuit depth. It would be very interesting if such approximate formulas could be derived.

Finally, in order to further extend the applicability of our protocol, a natural direction would be to study its compatibility with available error-mitigation techniques [82–91]. In fact, we expect that relatively simple error mitigation strategies can be incorporated to take into account the effect of noise in the shallow quantum circuits used to perform the classical shadow measurements. We leave this interesting topic for future work.

Note added. While finalising this draft, Refs. [92] and [93] appeared on the arXiv, which rigorously show that quantum circuits form approximate unitary designs over N qubits in $\log(N)$-depth. An immediate consequence of this result is that the approximate estimator for the fidelity studied in Sec. 3 becomes accurate in depth $O(\log N)$, which is consistent with our findings. The approximation measure used in Ref. [92, 93] (relative error) does not directly constrain the purity estimator, so the results are also consistent with our findings in Sec. 4.

Acknowledgements

We are very grateful to Benoit Vermersch for valuable feedback on the manuscript. M.I. acknowledges helpful discussions with Robert Huang, Nick Hunter-Jones, and Vedika Khemani. X.T. acknowledges DFG under Germany's Excellence Strategy – Cluster of Excellence Matter

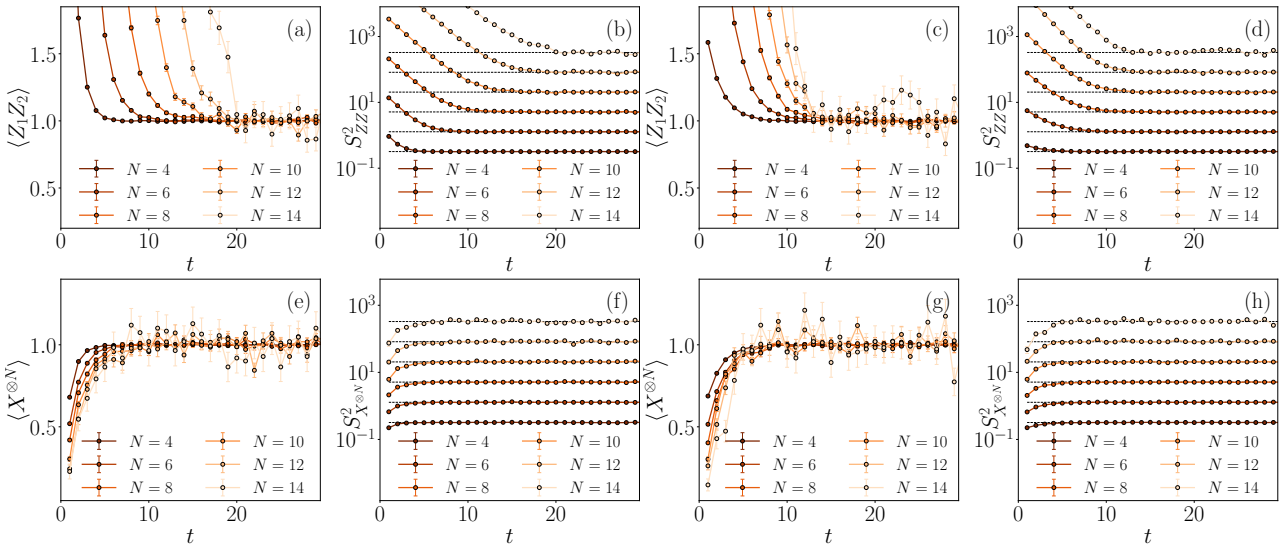

Figure 6: Numerical computation of the average of the estimator for expectation values of Pauli operators, for 1D local circuits [(a),(b),(e),(f)] and non-local ones [(c),(d),(g),(h)]. The initial state is the GHZ state (30). Panels (a), (c), (e) and (g) show the average value as a function of circuit depth for increasing N. Panels (b), (d), (f) and (h) show the variance for the same depths and sizes. The dashed lines correspond to the $t \to \infty$ value in Eq. (43).

and Light for Quantum Computing (ML4Q) EXC 2004/1 – 390534769, and DFG Collaborative Research Center (CRC) 183 Project No. 277101999 - project B01. The research of R.C. and E.E. is partially funded by INFN (project "QUANTUM"). E.E. also acknowledges financial support from the PRIN Project 2022H77XB7 "Hybrid algorithms for quantum simulators". This work was co-funded by the European Union (ERC, QUANTHEM, 101114881). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Data Availability Statement.— The source code and the data have been deposited in the Zenodo public folder [94].

A Details on the approximate fidelity estimator

This Appendix provides additional details on the derivations presented in Ref. 3.

A.1 Derivation of Eq. (21)

First, we note that the average of the fidelity estimator (15) and the variable $\tilde{f}_t^{(r)}$ in Eq. (16) coincide and we may therefore consider the latter. Using the definition of the inverse channel (12), we can rewrite

$$\tilde{f}_t^{(r)} = \langle\psi|(2^N+1)\left[(U_t^{(r)})^\dagger|b^{(r)}\rangle\langle b^{(r)}|U_t^{(r)}\right]|\psi\rangle - 1\,, \tag{44}$$

where $b^{(r)} = 0, ..., 2^N - 1$ is a bitstring sampled with probability $p_{b,U}^{(r)} = \langle b^{(r)}|U_t^{(r)}\rho(U_t^{(r)})^\dagger|b^{(r)}\rangle$ and $U_t^{(r)}$ a depth-t random circuit made of two-qubit Clifford gates. We consider the case in which the lab state and the target state coincide, $\rho = |\psi\rangle\langle\psi|$, and $|\psi\rangle$ is a product state, $|\psi\rangle = \otimes_j |\phi\rangle_j$.

The expectation value reads

$$\mathbb{E}\left[\tilde{f}_t^{(r)}\right] = (2^N+1)\sum_{b^{(r)}=0}^{2^N-1}\int d\mu(U_t^{(r)})|\langle b^{(r)}|U_t^{(r)}|\psi\rangle|^2 p_{b,U}^{(r)} - 1\,, \tag{45}$$

where we replaced the sum over 2-qubit Clifford gates by the average over Haar-random gates due to the 2-design property of the Clifford unitaries [63]. Since $|\psi\rangle$ and $|b^{(r)}\rangle$ are product states, we can express them as products of single-qubit unitaries acting on $|0\rangle = \otimes_j |0\rangle_j$. By invoking the invariance of the Haar measure for left and right multiplication, we then obtain the expression

$$\mathbb{E}\left[\tilde{\mathcal{F}}_t^{(e)}\right] = (2^N+1)2^N \int d\mu(U_t) |\langle 0|U_t|0\rangle|^4 - 1 = (2^N+1)2^N Z_t - 1\,, \tag{46}$$

where we introduced $Z_t = \mathbb{E}_{U_t}\left[|\langle 0|U_t|0\rangle|^4\right]$.

The term Z_t was studied extensively in Ref. [67]. There, a notion of collision probability $\sum_b |\langle b|U_t|0\rangle|^4$ was introduced, expressing the probability that two independent measurements of two copies of the state $U_t|0\rangle$ yield the same outcome. Averaging over the random circuit instances U_t we obtain Z_t (up to a 2^N coefficient), which has a minimum when the probability distribution $p_{b,U} = |\langle b|U_t|0\rangle|^2$ is uniform. This property was used to define the notion of anti-concentration: a random circuit U_t is said to anti-concentrate if Z_t differs at most by a constant value from its minimum.

Crucially, Ref. [67] also provided upper and lower bounds on the collision probability for different circuit architectures. For a general depth-t random circuit with two-qubit gates, the following bounds were found [67]

$$2 \leq 2^N(2^N+1)Z_t \leq 2 + 2e^{-a(t-T_N)}\,, \tag{47}$$

where a and T_N depend on the circuit architecture. Thus, the inequality $\mathbb{E}\left[\tilde{\mathcal{F}}_t^{(e)}\right] - 1 \leq \delta$ (for $\delta > 0$), is equivalent to

$$\mathbb{E}\left[\tilde{\mathcal{F}}_t^{(e)}\right] - 1 = (2^N+1)2^N Z_t - 2 \leq 2e^{-a(t-T_N)} \leq \delta\,. \tag{48}$$

This inequality immediately implies that the error on the fidelity becomes smaller than δ at a time $t_0^*(N,\delta) \leq g(N,\delta)$ with

$$g(N,\delta) = T_N + a^{-1}\log(2/\delta). \tag{49}$$

In the case of 1D brickwork circuits, $a = \log(5/4)$ and $T_N = [\log(N) + \log(e-1)]/a + 1$ [67], which finally implies Eq. (23).

A.2 Variance of the approximate estimator

We start from the definition of the locally-scrambled shadow norm for the operator O, which is obtained by replacing the maximization in Eq. (9) with the average over the ensemble $\{V\rho V^\dagger | V = \otimes_j V_j \in \mathrm{U}(2)^{\otimes N}\}$

$$\begin{aligned} ||O||^2_{\mathrm{sh,LS}} &= \mathop{\mathbb{E}}_{V,U_t} \sum_{b\in\{0,1\}^N} \left[\left\langle b\middle|U_t V\rho V^\dagger U_t^\dagger\middle|b\right\rangle \left\langle b\middle|U_t \mathcal{M}_\infty^{-1}(O) U_t^\dagger\middle|b\right\rangle^2\right] \\ &= \frac{1}{2^N}\mathop{\mathbb{E}}_{U_t} \sum_{b\in\{0,1\}^N} \left[\left\langle b\middle|U_t \mathcal{M}_\infty^{-1}(O) U_t^\dagger\middle|b\right\rangle^2\right], \end{aligned} \tag{50}$$

where we used the fact that $\mathbb{E}_V\left[V\rho V^\dagger\right] = \mathbb{1}/2^N$. From here (Eq. (26)) we can replace $\mathcal{M}_\infty^{-1}(O) = (2^N+1)O$ and evaluate

$$\begin{aligned} ||O||^2_{\mathrm{sh,LS}} &= \frac{(2^N+1)^2}{2^N}\int d\mu(U_t) \sum_{b=0}^{2^N-1} \left\langle b\middle|U_t\left(|\psi\rangle\langle\psi| - \frac{\mathbb{1}}{2^N}\right)U_t^\dagger\middle|b\right\rangle^2 \\ &= (2^N+1)^2 \int d\mu(U_t)\left[|\langle 0|U_t|\psi\rangle|^4 - \frac{2}{2^N}|\langle 0|U_t|\psi\rangle|^2 + \frac{1}{2^{2N}}\right] \\ &= (2^N+1)^2\left(\int d\mu(U_t)|\langle 0|U_t|\psi\rangle|^4 - \frac{1}{2^{2N}}\right) = (2^N+1)^2\left(Z_t - \frac{1}{2^{2N}}\right), \end{aligned} \tag{51}$$

where the last equality (Eq. (27)) holds if $|\psi\rangle$ is a product state. As before, after a depth $t^*(\delta,N)$ we have $2 \leq (2^N+1)2^N Z_{t^*} \leq 2+\delta$ and thus $||O||^2_{\mathrm{sh,LS}} \leq (1+2^{-N})(2+\delta) \leq 3(2+\delta)/2$.

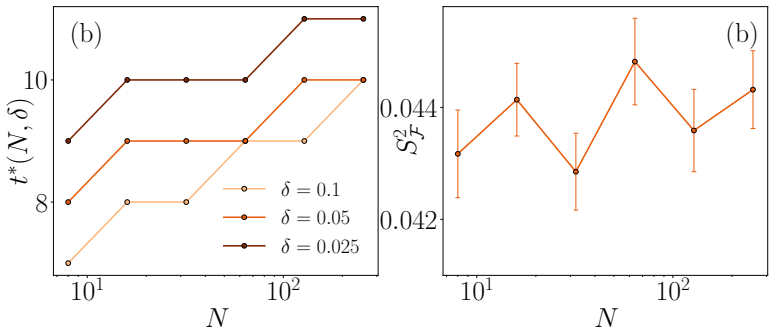

Figure 7: Numerical simulation of the certification protocol for 1D non-local circuits, where both the lab state and the target state are the GHZ state. (a): Depth $t^*(N,\delta)$ after which $\Delta F_{\psi,\psi} \leq \delta$. (b): The variance $S^2_{\mathcal{F}}$ estimated at the depth $t^*(N,\delta)$, for $\delta = 0.05$.

We can also compute explicitly the variance of $\tilde{\mathcal{F}}_t^{(e)}$ (Eq. (15)) at late times. We start from

$$\mathbb{E}\left[(\tilde{\mathcal{F}}_t^{(e)})^2\right] = \mathbb{E}\left[\frac{(2^N+1)^2}{M^2}\sum_{r,s=1}^{M} |\langle b^{(r)}|U_t^{(r)}|\psi\rangle|^2 |\langle b^{(s)}|U_t^{(s)}|\psi\rangle|^2 - 2\frac{(2^N+1)}{M}\sum_{r=1}^{M}|\langle b^{(r)}|U_t^{(r)}|\psi\rangle|^2 + 1\right]. \tag{52}$$

The expectation value of the second term when $t \to \infty$ is just $-2(\mathbb{E}[\tilde{\mathcal{F}}_t^{(e)}]+1) = -4$. The first term can be rewritten as

$$\frac{(2^N+1)^2}{M^2}\sum_r \mathbb{E}\left[|\langle b^{(r)}|U_t^{(r)}|\psi\rangle|^4\right] + \frac{(2^N+1)^2}{M^2}\sum_{r\neq s}\mathbb{E}\left[|\langle b^{(r)}|U_t^{(r)}|\psi\rangle|^2\right]\mathbb{E}\left[|\langle b^{(s)}|U_t^{(s)}|\psi\rangle|^2\right], \tag{53}$$

where in turn the second term is aymptotically

$$\frac{1}{M^2}\sum_{r\neq s}(\mathbb{E}[\tilde{f}_t^{(r)}]+1)^2 = \frac{4(M-1)}{M}, \tag{54}$$

and the first term

$$\frac{(2^N+1)^2}{M^2}\sum_{r=1}^{M}\int d\mu(U_t)\sum_{b^{(r)}=0}^{2^N-1}|\langle b^{(r)}|U_t^{(r)}|\psi\rangle|^6 = \frac{2^N(2^N+1)^2}{M}\int d\mu(U_t)\mathrm{Tr}(|0\rangle\langle 0|^{\otimes 3}U_t^{\otimes 3}|\psi\rangle\langle\psi|^{\otimes 3}(U_t^\dagger)^{\otimes 3}) \tag{55}$$

can be computed at late times by replacing the $d\mu(U_t)$ with the Haar measure over the full $U(2^N)$ unitary group. In this case, we have

$$\int d\mu_H(U)U^{\otimes 3}|\psi\rangle\langle\psi|^{\otimes 3}(U^\dagger)^{\otimes 3}) = \frac{1}{2^N(2^N+1)(2^N+2)}\sum_{\sigma\in S_3}P_\sigma, \tag{56}$$

where P_σ is the permutation operator implementing the permutation σ on the three replicas of the system. Thus the above equation yields $\frac{1}{M}\frac{6(2^N+1)}{(2^N+2)}$. Putting everything together, we get

$$\mathrm{Var}\left[\tilde{\mathcal{F}}_t^{(e)}\right] = \mathbb{E}\left[(\tilde{\mathcal{F}}_t^{(e)})^2\right] - 1 = \frac{1}{M}\frac{6(2^N+1)}{(2^N+2)} + \frac{4(M-1)}{M} - 4 = \frac{1}{M}\frac{2(2^N-1)}{2^N+2}, \tag{57}$$

which is a factor $1/M$ smaller than the variance of $\tilde{f}_t^{(r)}$.

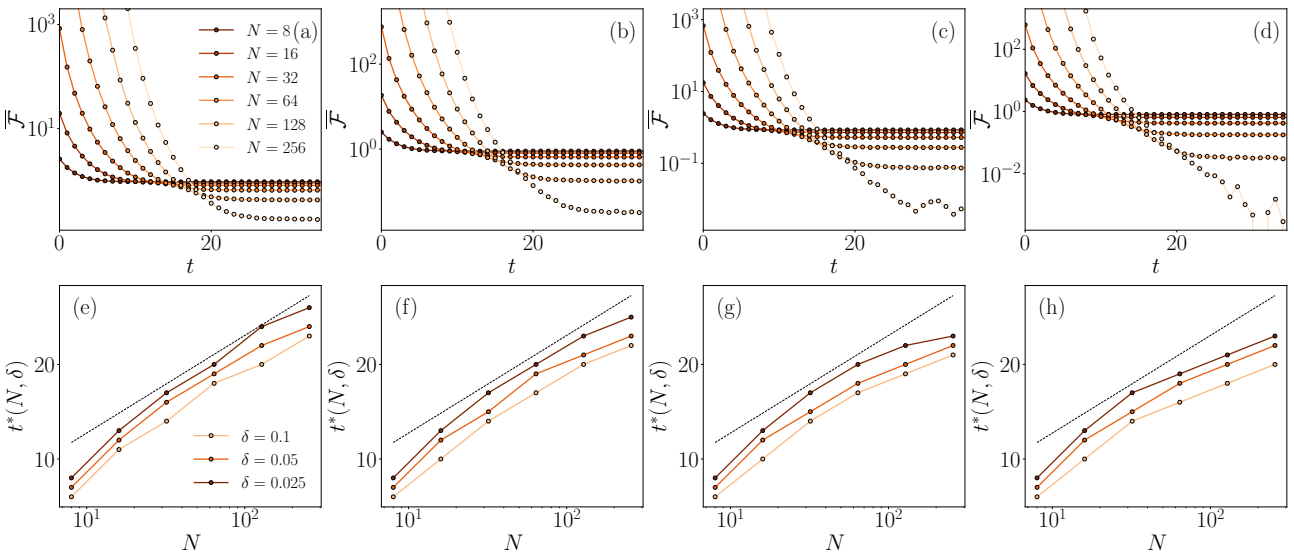

Figure 8: Numerical simulation of the certification protocol for 1D local circuits. The target state $|\psi\rangle$ is the GHZ state, the lab state $\rho = \mathcal{E}_p^{\otimes N}(|\psi\rangle\langle\psi|)$, where $\mathcal{E}_p$ is the depolarizing channel. Each column represents a different value of p, increasing linearly from $p = 0.01$ to $p = 0.04$. (a)-(d): Averaged fidelity estimator $\overline{\mathcal{F}} = \mathbb{E}[\tilde{\mathcal{F}}]$ as a function of the circuit depth, for increasing N. (e)-(h): Depth $t^*(N, \delta)$ after which $\Delta F_{\psi,\rho} \leq \delta$.

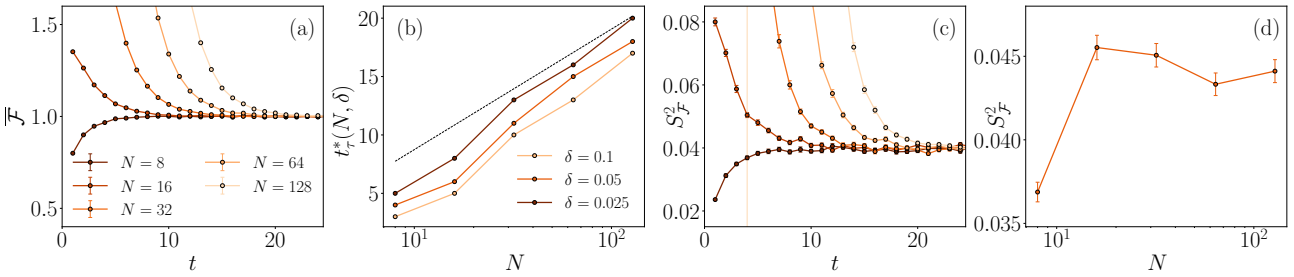

Figure 9: Numerical simulation of the certification protocol for 1D local circuits, where both the lab state and the target state are equal to a single instance of a random short-range entangled stabilizer state. (a): Averaged fidelity estimator $\overline{\mathcal{F}} = \mathbb{E}[\tilde{\mathcal{F}}]$ as a function of the circuit depth, for increasing N. (b): Depth $t^*_\tau(N, \delta)$ after which $\Delta F_{\psi,\psi} \leq \delta$, where $\tau = 3$ for the chosen state. The dashed line is the function in Eq. (23), shifted by a negative constant. (c): Variance of fidelity estimator $S^2_{\mathcal{F}}$ as a function of time. (d): The variance $S^2_{\mathcal{F}}$ estimated at the depth $t^*(N, \delta)$, for $\delta = 0.05$.

A.3 Numerical methods and additional results

In this appendix, we provide details on the numerical methods employed in our work and additional data.

Stabilizer circuit simulations were employed for 1D states, both pure (GHZ state and random short-range entangled states) and mixed (through the application of depolarizing channels), and 2D states (cluster state). For this purpose, we used the STIM library [95].

All our results are obtained by constructing a suitable initial state and extracting $M = 50$ classical shadows with the shallow-circuit protocols described in Sec. 2, for increasing circuit depth. The shadows are then post-processed to construct a single realization of our modified estimator $\tilde{\mathcal{F}}^{(s)}(t)$ for the fidelity. We then average over $R = 9600$ independent realizations of this estimator to obtain $\bar{\mathcal{F}}(t)$, and compute the sample variance $S^2_{\mathcal{F}}(t) = \sum_s (\tilde{\mathcal{F}}^{(s)}(t) - \bar{\mathcal{F}}(t))^2/R$. The error on the mean is then $\delta\mathcal{F}(t) = \sqrt{S^2_{\mathcal{F}}(t)/R}$. As for the sample variance, we construct 100 bootstrap samples and compute the variance over each of them. The standard deviation of the set of bootstrap variances is then used as the error on $S^2_{\mathcal{F}}(t)$.

In Fig. 7 we show two plots regarding non-local quantum circuits (extending Fig. 2(c) and Fig. 2(d)

of the main text), once again in the case where both the lab state and the target state are the GHZ state. Fig. 7(a) shows the values $t^*(N,\delta)$ after which $\Delta F_{\psi,\psi} \leq \delta$ for different values of δ. In agreement with the bounds found in Ref. [67], for this circuit architecture the scaling is also $a^{-1}\log(N)$, albeit with a prefactor which is smaller than the one in the local circuits counterpart. In Fig. 7(b) we plot the variance $S^2_{\mathcal{F}}(t)$ at $t^*(N,\delta)$ for $\delta = 0.05$, showing that it is bounded by a number independent of the system's size.

Going back to 1D local circuits, we consider again the case where the lab state is $\rho = \mathcal{E}_p^{\otimes N}(|\psi\rangle\langle\psi|)$, with $|\psi\rangle$ being the GHZ state. In Figs. 8(a)-(d) we extend the plot of Fig. 1(a), showing the averaged fidelity estimator $\bar{\mathcal{F}} = \mathbb{E}[\tilde{\mathcal{F}}]$ as a function of the circuit depth, for increasing N, and for increasing depolarizing probability p. All plots share the same behaviour, meaning that the convergence of the fidelity estimator to its true value does not depend on the amount of noise affecting the lab state. This is also shown in Figs. 8(e)-(h), where we plot the depth $t^*(N,\delta)$ after which $\Delta F_{\psi,\rho} \leq \delta$. We note that for increasing noise (and thus lower purities) the convergence is obtained at earlier depths.

Finally, we consider the case where the lab state $|\psi\rangle$ (and the target state) is a random short-range entangled stabilizer state, obtained by evolving the $|0\rangle$ state via a depth-3 shallow circuit $|\psi\rangle = U_3|0\rangle$ where all the two-qubit Clifford unitaries are equal to the same random two-qubit Clifford. In Fig. 9 we show the results when the latter is $(Z\otimes Z)\mathrm{CNOT}_{1,0}(\mathbb{1}\otimes H)\mathrm{CNOT}_{0,1}(HS\otimes S)\mathrm{SWAP}$. We plot once again the averaged fidelity (a) and variance (c) as functions of circuit depth for increasing N. Fig. 9(b) displays the depth $t^*_\tau(N,\delta)$ after which $\Delta F_{\psi,\psi} \leq \delta$, where $t^*_\tau(N,\delta) = t^*_0(N,\delta) - \tau$ as implied by Eq. 25. Comparing this with Fig. 1, we observe the same behaviour up to a shift of a constant $\tau = 3$, finding agreement with the analytical arguments presented in the main text. Finally, Fig. 9(d) shows the variance $S_{\mathcal{F}^2}$ at depth $t^*_\tau(N,\delta)$, for $\delta = 0.05$, which is again bounded by a number independent of N.

B Details on the approximate purity estimators

In this section we provide further detail on the computation of the approximate purity estimator.

B.1 Mapping of the average estimators onto non-unitary dynamics

We first explain the method to compute the average of the approximate estimator, which is based on a standard replica formalism [64, 65].

Consider the expectation value of the purity estimator for an initial state ρ:

$$\mathbb{E}\left[\tilde{\mathcal{P}}_t^{(e)}\right] = 2^{2N}(2^N+1)^2\int d\mu(U_t)d\mu(V_t)|\langle 0|U_tV_t^\dagger|0\rangle|^2\langle 0|U_t\rho U_t^\dagger|0\rangle\langle 0|V_t\rho V_t^\dagger|0\rangle - 2^N - 2\,. \tag{58}$$

It is convenient to view operators as states in a doubled Hilbert space, $\rho \to |\rho\rangle\!\rangle$, so that we can identify $U\rho U^\dagger \to U\otimes U^*|\rho\rangle\!\rangle$. The term inside the integral can be rewritten as the expectation value of the four-replicas operator $O = |\rho\rangle\!\rangle\langle\!\langle\rho| \otimes I_{2^{2N}}$, on the state given by

$$\mathbb{E}_{U_t}\left[U_t\otimes U_t^*\otimes U_t\otimes U_t^*\right]|0^{\otimes 4}\rangle\!\rangle\,. \tag{59}$$

Since U_t is a quantum circuit where each gate is drawn randomly and independently, we can take the average over each gate individually see Ref. [65] for more detail. In turn, since random Cliffords form a two-design, these averages can be performed analytically. Denoting by $U_{j,k}$ the two-qubit gate, we get

$$\begin{aligned}\mathcal{U}_{jk} &= \mathbb{E}_{U_{j,k}}\left[U_{j,k}\otimes U_{j,k}^*\otimes U_{j,k}\otimes U_{j,k}^*\right]\\ &= \frac{1}{d^4-1}\left[|\mathcal{I}^+_{jk}\rangle\!\rangle\langle\!\langle\mathcal{I}^+_{jk}| + |\mathcal{I}^-_{jk}\rangle\!\rangle\langle\!\langle\mathcal{I}^-_{jk}| - \frac{1}{d^2}\left(|\mathcal{I}^+_{jk}\rangle\!\rangle\langle\!\langle\mathcal{I}^-_{jk}| + |\mathcal{I}^-_{jk}\rangle\!\rangle\langle\!\langle\mathcal{I}^+_{jk}|\right)\right],\end{aligned} \tag{60}$$

where $|\mathcal{I}^\pm_{jk}\rangle\!\rangle = |\mathcal{I}^\pm\rangle\!\rangle_j \otimes |\mathcal{I}^\pm\rangle\!\rangle_k$, and the superstates $|\mathcal{I}^\pm\rangle\!\rangle$ act as $\langle\!\langle A|\mathcal{I}^+\rangle\!\rangle = \mathrm{Tr}(AS_{1,2}S_{3,4})$ and $\langle\!\langle A|\mathcal{I}^-\rangle\!\rangle = \mathrm{Tr}(AS_{1,4}S_{2,3})$ for any four-replicas operator A, where $S_{a,b}$ is the two-replicas swap operator.

In summary, computing the average (58) requires evolving an initial state $|0\rangle\!\rangle^{\otimes 4}$ by a non-unitary circuit made of the gates $\mathcal{U}_{j,k}$. Importantly, we see that $\mathcal{U}_{j,k}$ projects onto the local sub-subspace spanned by $|\mathcal{I}^{\pm}_{jk}\rangle\!\rangle$. Therefore, the dynamics takes place in an effective chain of N two-level systems.

For the 1D architecture, we have computed the evolved state using standard TN computations, using the fact that each layer of gates can be represented trivially as a matrix-product operator. Our implementation is in the ITensor library [96, 97]. While we find that the bipartite entanglement entropy grows very slowly with the number of layers, the contraction of tensors produces very small numerical values, so that the intermediate computations must be carried out at high numerical precision (recall that the dynamics is not unitary and does not preserve norms). This effectively limits the system sizes we can study to $N \lesssim 32$. In the case of non-local circuits, we can use a numerically exact method for larger sizes, as we detail in the following.

B.2 Averaged estimator for non-local circuits

For the non-local architecture considered in this work, U_t is a non-local quantum circuit composed of t "layers" $U^{(j)}$, where each layer is composed of 2-qubit Clifford unitaries acting on $N/2$ pairs of qubits chosen at random, $U^{(j)} = U_{j_1 j_2} ... U_{j_{N-1} j_N}$. Averaging over the random pairs and the unitaries, we have

$$\mathbb{E}_{U^{(j)}} \left[U^{(j)} \otimes (U^{(j)})^* \otimes U^{(j)} \otimes (U^{(j)})^* \right]^t |0^{\otimes 4}\rangle\!\rangle = \left(\frac{1}{N!} \sum_{j \in \mathrm{S}_N} \mathcal{U}_{j_1 j_2} ... \mathcal{U}_{j_{N-1} j_N} \right)^t |0^{\otimes 4}\rangle\!\rangle = \mathcal{L}^t |0^{\otimes 4}\rangle\!\rangle, \quad (61)$$

where $\mathcal{U}_{jk}$ is given in Eq. (60), while S_N is the permutation group. Thus the expectation value reads

$$\mathbb{E}\left[\tilde{\mathcal{P}}_t^{(e)}\right] = 2^{2N}(2^N+1)^2 \langle\!\langle 0^{\otimes 4} | \mathcal{L}^t O \mathcal{L}^t | 0^{\otimes 4} \rangle\!\rangle - 2^N - 2\,. \quad (62)$$

To compute it, we first note that the $\mathcal{L}$ super-operator is a projector onto the Hilbert space spanned by $\{|\mathcal{I}^+\rangle\!\rangle, |\mathcal{I}^-\rangle\!\rangle\}^{\otimes N}$, that is, onto an N-fold tensor product of 2-dimensional Hilbert spaces. Since $\langle\!\langle \mathcal{I}^+ | \mathcal{I}^+ \rangle\!\rangle = \langle\!\langle \mathcal{I}^- | \mathcal{I}^- \rangle\!\rangle = d^2$ and $\langle\!\langle \mathcal{I}^+ | \mathcal{I}^- \rangle\!\rangle = d$, we can construct an orthonormal basis $\{|\tilde{0}\rangle\!\rangle, |\tilde{1}\rangle\!\rangle\}$ for each local Hilbert space, such that

$$|\mathcal{I}^+\rangle\!\rangle = d|\tilde{0}\rangle\!\rangle, \qquad |\mathcal{I}^-\rangle\!\rangle = |\tilde{0}\rangle\!\rangle + \sqrt{d^2-1}|\tilde{1}\rangle\!\rangle\,. \quad (63)$$

Next, we note that both the initial state $|0^{\otimes 4}\rangle$ and the super-operators $\mathcal{L}$ and O are invariant under permutation of any of the N sites. Thus we can construct a basis $\{|W_n\rangle\!\rangle\}_{n=0}^N$ for the $N+1$-dimensional permutation invariant subspace:

$$|W_n\rangle\!\rangle = \frac{1}{\sqrt{N!(N-n)!n!}} \sum_\pi \pi |\tilde{1}\rangle\!\rangle^{\otimes n} |\tilde{0}\rangle\!\rangle^{\otimes N-n}\,. \quad (64)$$

Here and in the following, we denote by π an arbitrary permutation operator over the N two-level systems. Note that for convenience this basis can also be written in terms of the $\{|\mathcal{I}^+\rangle\!\rangle, |\mathcal{I}^-\rangle\!\rangle\}$ basis, as

$$|W_n\rangle\!\rangle = \frac{1}{\sqrt{N!(N-n)!n!}} \frac{1}{d^N (d^2-1)^{\frac{n}{2}}} \sum_\pi \left(d|\mathcal{I}^-\rangle\!\rangle - |\mathcal{I}^+\rangle\!\rangle \right)^{\otimes n} |\mathcal{I}^+\rangle\!\rangle^{\otimes N-n} = \mathcal{N}^{-1} \sum_\pi |v\rangle\!\rangle^{\otimes n} |\mathcal{I}^+\rangle\!\rangle^{\otimes N-n}, \quad (65)$$

where we introduced $|v\rangle\!\rangle = d|\mathcal{I}^-\rangle\!\rangle - |\mathcal{I}^+\rangle\!\rangle$.

We are now ready to compute $\mathcal{L}|W_n\rangle\!\rangle$. Fix the state $|v\rangle\!\rangle^{\otimes n} |\mathcal{I}^+\rangle\!\rangle^{\otimes N-n}$. The action of each 2-qubit channel $\mathcal{U}_{jk}$ depends on whether the qubits j,k are both in the $|v\rangle\!\rangle$ state, both in the $|\mathcal{I}^+\rangle\!\rangle$ state, or in opposite states. For a given sequence $\mathcal{U}_{j_1 j_2} \ldots \mathcal{U}_{j_{N-1} j_N}$ (which identifies $N/2$ pairs), we denote the number of pairs in each of the three configurations by $r \le n/2$, $s \le (N-n)/2$ and $N/2 - s - r$

respectively. Averaging over all permutations to obtain $\mathcal{L}$, we obtain

$$\begin{aligned}\mathcal{L}\sum_{\pi}\pi|v\rangle\!\rangle^{\otimes n}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-n} &= \frac{1}{N!}\sum_{\pi}\pi\sum_{j\in \mathrm{S}_N}\frac{(d^2-1)^{N/2-s-r}}{(d^4-1)^{N/2-s}}\left[|v\rangle\!\rangle^{\otimes 2}+2|v\rangle\!\rangle|\mathcal{I}^{+}\rangle\!\rangle\right]^{\otimes r}\\ &\quad\times|\mathcal{I}^{+}\rangle\!\rangle^{\otimes 2s}\left[|v\rangle\!\rangle^{\otimes 2}+2|v\rangle\!\rangle|\mathcal{I}^{+}\rangle\!\rangle\right]^{\otimes N/2-s-r}\\ &= \sum_{\pi}\pi\frac{1}{N!}\sum_{j\in \mathrm{S}_N}\frac{(d^2-1)^{n}}{(d^4-1)^{n-r}}\left[|v\rangle\!\rangle^{\otimes 2}+2|v\rangle\!\rangle|\mathcal{I}^{+}\rangle\!\rangle\right]^{\otimes r}\\ &\quad\times|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-2n+2r}\left[|v\rangle\!\rangle^{\otimes 2}+2|v\rangle\!\rangle|\mathcal{I}^{+}\rangle\!\rangle\right]^{\otimes n-2r}, \end{aligned}\tag{66}$$

where r and s depend on permutation j. Note that in the second line we used the fact that r and s are related by $n-2r=N-n-2s$, and thus $s=N/2-n+r$. Now we can expand the tensor products and reorder each term:

$$\sum_{\pi}\pi\sum_{j\in \mathrm{S}_N}\left[|v\rangle\!\rangle^{\otimes 2}+2|v\rangle\!\rangle|\mathcal{I}^{+}\rangle\!\rangle\right]^{\otimes r}=\sum_{\pi}\pi\sum_{j\in \mathrm{S}_N}\sum_{k=0}^{r}\binom{r}{k}|v\rangle\!\rangle^{\otimes r+k}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes r-k}2^{r-k},\tag{67}$$

and analogously for $\left[|v\rangle\!\rangle^{\otimes 2}+2|v\rangle\!\rangle|\mathcal{I}^{+}\rangle\!\rangle\right]^{\otimes n-2r}$. We arrive at

$$\begin{aligned}\mathcal{L}\sum_{\pi}\pi|v\rangle\!\rangle^{\otimes n}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-n} &= \frac{(d^2-1)^{n}}{N!}\sum_{\pi}\pi\sum_{j\in \mathrm{S}_N}\sum_{k=0}^{r}\sum_{l=0}^{n-2r}\frac{2^{n-r-(k+l)}}{(d^4-1)^{n-r}}\binom{r}{k}\binom{n-2r}{l}\\ &\quad\times|v\rangle\!\rangle^{\otimes n-r+(k+l)}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-n+r-(k+l)}.\end{aligned}\tag{68}$$

Next, we set $m=n-r+(k+l)$, yielding

$$\sum_{k=0}^{r}\sum_{l=0}^{n-2r}\frac{2^{n-r-(k+l)}}{(d^4-1)^{n-r}}\binom{r}{k}\binom{n-2r}{l}|v\rangle\!\rangle^{\otimes n-r+(k+l)}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-n+r-(k+l)}=\sum_{m=0}^{N}c_m^{(r)}|v\rangle\!\rangle^{\otimes m}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-m},\tag{69}$$

where, using the Vandermonde's identity,

$$c_m^{(r)}=\sum_{k=0}^{r}\sum_{l=0}^{n-2r}\frac{2^{n-r-(k+l)}}{(d^4-1)^{n-r}}\binom{r}{k}\binom{n-2r}{l}\delta_{m,n-r+(k+l)}=\frac{2^{2(n-r)-m}}{(d^4-1)^{n-r}}\binom{n-r}{m-n+r}.\tag{70}$$

Therefore

$$\begin{aligned}\mathcal{L}\sum_{\pi}\pi|v\rangle\!\rangle^{\otimes n}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-n} &= \frac{(d^2-1)^{n}}{N!}\sum_{\pi}\pi\sum_{j\in \mathrm{S}_N}\sum_{m=0}^{N}c_m^{(r)}|v\rangle\!\rangle^{\otimes m}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-m}\\ &= \frac{(d^2-1)^{n}}{N!}\sum_{j\in \mathrm{S}_N}\sum_{m=0}^{N}c_m^{(r)}\sum_{\pi}|v\rangle\!\rangle^{\otimes m}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-m}\\ &= \frac{(d^2-1)^{n}}{N!}\sum_{j\in \mathrm{S}_N}\sum_{m=0}^{N}(N!(N-m)!m!)^{\frac{1}{2}}d^N(d^2-1)^{\frac{m}{2}}c_m^{(r)}|W_m\rangle\!\rangle.\end{aligned}\tag{71}$$

We can now perform the sum over permutations $j\in \mathrm{S}_N$ by counting the number of permutations N_r with fixed r, and then summing over r. The reasoning is the following. We have $N/2$ pairs divided in the following way: r pairs of type $|v\rangle\!\rangle^{\otimes 2}$, s pairs of type $|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-n}$, and $N/2-s-r$ pairs of type $|v\rangle\!\rangle|\mathcal{I}\rangle\!\rangle$ or $|\mathcal{I}\rangle\!\rangle|v\rangle\!\rangle$. Thus we have $\binom{N/2}{r}\binom{N/2-r}{s}$ different ways to order them. We are then free to permute the indices $j_1...j_N$ in a way that preserves this order. Thus there are $n!$ permutations of the $|v\rangle\!\rangle$ states, $(N-n)!$ permutations of the $|\mathcal{I}^{+}\rangle\!\rangle$ states, and $2^{N/2-s-r}$ ways to order the states in the mixed pairs. Knowing that $s=N/2-n+r$, the total is

$$N_r=\binom{N/2}{r}\binom{N/2-r}{N/2-n+r}n!(N-n)!2^{n-2r}=\binom{N/2}{r,n-2r,N/2-n+r}n!(N-n)!2^{n-2r}.\tag{72}$$

Summing over r, we arrive at

$$\mathcal{L}\sum_{\pi}\pi|v\rangle\!\rangle^{\otimes n}|\mathcal{I}^{+}\rangle\!\rangle^{\otimes N-n}=\frac{d^{N}(d^{2}-1)^{n}}{\binom{N}{n}}\sum_{m=0}^{N}(d^{2}-1)^{\frac{m}{2}}\sqrt{N!m!(N-m)!}|W_{m}\rangle\!\rangle \times\sum_{r=0}^{n/2}\binom{N/2}{r,n-2r,N/2-n+r}2^{n-2r}c_{m}^{(r)} \tag{73}$$

and thus, finally

$$\mathcal{L}|W_{n}\rangle\!\rangle=\frac{(d^{2}-1)^{\frac{n}{2}}}{\binom{N}{n}}\sum_{m=0}^{N}(d^{2}-1)^{\frac{m}{2}}\sqrt{\frac{m!(N-m)!}{n!(N-n)!}}d_{m}^{(n)}|W_{m}\rangle\!\rangle\,, \tag{74}$$

where

$$d_{m}^{(n)}=\sum_{r=0}^{n/2}\binom{N/2}{r,n-2r,N/2-n+r}2^{n-2r}c_{m}^{(r)}\,. \tag{75}$$

This expression can now be used to numerically compute $\mathcal{L}^{t}$ in the permutation-invariant basis $\{|W_{n}\rangle\!\rangle\}$. This makes the computation efficient, because the corresponding Hilbert space only grows linearly in N. It only remains to expand both the initial state $|0^{\otimes 4}\rangle\!\rangle$ and the super-operator $O=|\rho\rangle\!\rangle\langle\!\langle\rho|\otimes I_{2^{2N}}$ in the permutation invariant basis $\{|W_{n}\rangle\!\rangle\}_{n=0}^{N}$.

The initial state is a product state $|0^{\otimes 4}\rangle\!\rangle=\bigotimes_{j=1}^{N}|0^{\otimes 4}\rangle\!\rangle_{j}$, so we can compute its projection on the local basis $\{|\tilde{0}\rangle\!\rangle,|\tilde{1}\rangle\!\rangle\}$ to see that

$$|0^{\otimes 4}\rangle\!\rangle=\frac{1}{d^{N}}\left(|\tilde{0}\rangle\!\rangle+\sqrt{\frac{d-1}{d+1}}|\tilde{1}\rangle\!\rangle\right)^{\otimes N}=\frac{1}{d^{N}}\sum_{n=0}^{N}\left(\frac{d-1}{d+1}\right)^{\frac{n}{2}}\binom{N}{n}^{\frac{1}{2}}|W_{n}\rangle\!\rangle\,. \tag{76}$$

For the super-operator, we focus on the case where ρ is a product state, $\rho=\bigotimes_{j=1}^{N}\rho_{j}$. Then we can compute the matrix elements of $O_{j}=\rho_{j}\otimes I_{4}$ over $\{|\tilde{0}\rangle\!\rangle,|\tilde{1}\rangle\!\rangle\}_{j}$, and obtain

$$O_{j}|\tilde{0}\rangle\!\rangle_{j}=\frac{1}{d}|\tilde{0}\rangle\!\rangle_{j},\qquad O_{j}|\tilde{1}\rangle\!\rangle_{j}=\frac{1}{d^{2}-1}\left(\mathrm{tr}\rho_{j}^{2}-\frac{1}{d}\right)|\tilde{1}\rangle\!\rangle_{j}, \tag{77}$$

and thus

$$\langle\!\langle W_{m}|O|W_{n}\rangle\!\rangle=\delta_{m,n}\frac{1}{d^{N}}\frac{1}{(d^{2}-1)^{n}}\binom{N}{n}^{-1}\sum_{j\in\tilde{\pi}}\left(d\mathrm{tr}\rho_{j_{1}}^{2}-1\right)...\left(d\mathrm{tr}\rho_{j_{n}}^{2}-1\right)1_{j_{n+1}...j_{N}}, \tag{78}$$

where the sum is over all the different subsets of n elements $\{\rho_{j_{1}},...,\rho_{j_{n}}\}_{j}$ of the set of single-qubit states $\{\rho_{k}\}_{k=1}^{N}$. If the state is translational invariant, the expression finally simplifies to

$$\langle\!\langle W_{m}|O|W_{n}\rangle\!\rangle=\delta_{m,n}\frac{1}{d^{N}}\frac{1}{(d^{2}-1)^{n}}\left(d\mathrm{tr}\rho_{j}^{2}-1\right)^{n}\,. \tag{79}$$

B.3 Variance of the global Clifford purity estimator

We want to compute the variance of the purity estimator $\tilde{\mathcal{P}}_{t}^{(e)}$ defined in Eq. (34) when $t\to\infty$, that is, when our estimator reduces to the global Clifford estimator. Following [15], we have

$$\begin{aligned}\mathrm{Var}\left[\tilde{\mathcal{P}}_{\infty}^{(e)}\right]&=\binom{M}{2}^{-2}\sum_{r<r'}\sum_{s<s'}\left(\mathbb{E}\left[\mathrm{Tr}(\tilde{\rho}_{\infty}^{(r)}\tilde{\rho}_{\infty}^{(r')})\mathrm{Tr}(\tilde{\rho}_{\infty}^{(s)}\tilde{\rho}_{\infty}^{(s')})\right]-\mathcal{P}(\rho)^{2}\right)\\&=\binom{M}{2}^{-1}\mathrm{Var}\left[\mathrm{Tr}(\tilde{\rho}_{\infty}^{(1)}\tilde{\rho}_{\infty}^{(2)})\right]+\binom{M}{2}^{-1}2(M-2)\mathrm{Var}\left[\mathrm{Tr}(\tilde{\rho}_{\infty}^{(1)}\rho)\right]\,.\end{aligned} \tag{80}$$

The first term can be written explicitly as

$$\begin{aligned}\mathrm{Var}\left[\mathrm{Tr}(\tilde{\rho}_\infty^{(1)}\tilde{\rho}_\infty^{(2)})\right] &= \mathop{\mathbb{E}}_{U,V\in\mathrm{Cl}(2^N)}\mathop{\mathbb{E}}_{b,d}\mathrm{Tr}\left[\left((2^N+1)U^\dagger|b\rangle\langle b|U-\mathbb{1}\right)\left((2^N+1)V^\dagger|d\rangle\langle d|V-\mathbb{1}\right)\right]^2-\mathcal{P}(\rho)^2\\ &= \int d\mu_H(U)d\mu_H(V)\left\{2^{2N}\left[(2^N+1)^2|\langle 0|UV^\dagger|0\rangle|^2-(2^N+2)\right]^2\right.\\ &\quad\left.\times\langle 0|U\rho U^\dagger|0\rangle\langle 0|V\rho V^\dagger|0\rangle\right\}-\mathcal{P}(\rho)^2\\ &= 2^{2N}(2^N+1)^4\int d\mu_H(U)d\mu_H(V)|\langle 0|UV^\dagger|0\rangle|^4\langle 0|U\rho U^\dagger|0\rangle\langle 0|V\rho V^\dagger|0\rangle\\ &\quad-(2^N+2+\mathcal{P}(\rho))^2\,,\end{aligned}\tag{81}$$

where in the first line we used the fact that for $t\to\infty$ the random circuit in our estimator reduces to a random global Clifford unitary $U\in\mathrm{Cl}(2^N)$, and in the second line we invoked the 3-design property of Clifford unitaries to replace the sum with an integral over Haar measure $d\mu_H(U)$. The first term in the last line can then be rewritten as

$$2^{2N}(2^N+1)^4\int d\mu_H(U)\langle 0|U\rho U^\dagger|0\rangle\,\mathrm{Tr}\left\{\left((U^\dagger)^{\otimes 2}|0\rangle\langle 0|^{\otimes 2}U^{\otimes 2}\otimes\rho\right)\int d\mu_H(V)(V^\dagger)^{\otimes 3}|0\rangle\langle 0|^{\otimes 3}V^{\otimes 3}\right\},\tag{82}$$

and now the last integral can be computed using (56), yielding

$$\begin{aligned}&= \frac{2^N(2^N+1)^3}{2^N+2}\int d\mu_H(U)\langle 0|U\rho U^\dagger|0\rangle\left(2+4\langle 0|U\rho U^\dagger|0\rangle\right)\\ &= \frac{2^N(2^N+1)^3}{2^N+2}\left(\frac{2}{2^N}+4\frac{1+\mathcal{P}(\rho)}{2^N(2^N+1)}\right)=\frac{2(2^N+1)^2}{2^N+2}\left(2^N+3+2\mathcal{P}(\rho)\right),\end{aligned}\tag{83}$$

thus arriving at

$$\mathrm{Var}\left[\mathrm{Tr}(\tilde{\rho}_\infty^{(1)}\tilde{\rho}_\infty^{(2)})\right]=\frac{2(2^N+1)^2}{2^N+2}\left(2^N+3+2\mathcal{P}(\rho)\right)-\left(2^N+2+\mathcal{P}(\rho)\right)^2.\tag{84}$$

Following similar steps, the second term in Eq. (80) can be written as

$$\begin{aligned}\mathrm{Var}\left[\mathrm{Tr}(\tilde{\rho}_\infty^{(1)}\rho)\right] &= \mathop{\mathbb{E}}_{U\in\mathrm{Cl}(2^N)}\mathop{\mathbb{E}}_{b}\mathrm{Tr}\left[\left((2^N+1)U^\dagger|b\rangle\langle b|U-\mathbb{1}\right)\rho\right]^2-\mathcal{P}(\rho)^2\\ &= 2^N\int d\mu_H(U)\left((2^N+1)^2\langle 0|U\rho U^\dagger|0\rangle^2-2(2^N+1)\langle 0|U\rho U^\dagger|0\rangle+1\right)\langle 0|U\rho U^\dagger|0\rangle-\mathcal{P}(\rho)^2\\ &= \frac{2^N+1}{2^N+2}\left(1+3\mathcal{P}(\rho)+2\mathrm{Tr}(\rho^3)\right)-2\left(1+\mathcal{P}(\rho)\right)+1-\mathcal{P}(\rho)^2\\ &= \frac{2^N+1}{2^N+2}\left(1+3\mathcal{P}(\rho)+2\mathrm{Tr}(\rho^3)\right)-\left(1+\mathcal{P}(\rho)\right)^2.\end{aligned}\tag{85}$$

Putting everything together, we finally arrive at

$$\begin{aligned}\mathrm{Var}\left[\tilde{\mathcal{P}}_\infty^{(e)}\right] = \binom{M}{2}^{-1}\left\{\frac{2(2^N+1)^2}{2^N+2}\left(2^N+3+2\mathcal{P}(\rho)\right)-\left(2^N+2+\mathcal{P}(\rho)\right)^2\right.\\ \left.+2(M-2)\left[\frac{2^N+1}{2^N+2}\left(1+3\mathcal{P}(\rho)+2\mathrm{Tr}(\rho^3)\right)-\left(1+\mathcal{P}(\rho)\right)^2\right]\right\}.\end{aligned}\tag{86}$$

Thus the leading term is $\mathrm{Var}[\tilde{\mathcal{P}}_\infty^{(e)}]\sim\frac{2(2^N)}{M(M-1)}$.

References

[1] Steven T Flammia, David Gross, Yi-Kai Liu, and Jens Eisert. "Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators". New J. Phys. **14**, 095022 (2012).

[2] Jeongwan Haah, Aram W Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. "Sample-optimal tomography of quantum states". In Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. Pages 913–925. (2016).

[3] Ryan O'Donnell and John Wright. "Efficient quantum tomography". In Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. Pages 899–912. (2016).

[4] Rainer Blatt and Christian F Roos. "Quantum simulations with trapped ions". Nature Phys. **8**, 277–284 (2012).

[5] Christian Gross and Immanuel Bloch. "Quantum simulations with ultracold atoms in optical lattices". Science **357**, 995–1001 (2017).

[6] John Preskill. "Quantum computing in the nisq era and beyond". Quantum **2**, 79 (2018).

[7] Florian Schäfer, Takeshi Fukuhara, Seiji Sugawa, Yosuke Takasu, and Yoshiro Takahashi. "Tools for quantum simulation with ultracold atoms in optical lattices". Nature Rev. Phys. **2**, 411–425 (2020).

[8] Morten Kjaergaard, Mollie E Schwartz, Jochen Braumüller, Philip Krantz, Joel I-J Wang, Simon Gustavsson, and William D Oliver. "Superconducting qubits: Current state of play". Ann. Rev. Cond. Matt. Phys. **11**, 369–395 (2020).

[9] M Morgado and S Whitlock. "Quantum simulation and computing with rydberg-interacting qubits". AVS Quantum Science **3**, 023501 (2021).

[10] C. Monroe, W. C. Campbell, L.-M. Duan, Z.-X. Gong, A. V. Gorshkov, P. W. Hess, R. Islam, K. Kim, N. M. Linke, G. Pagano, P. Richerme, C. Senko, and N. Y. Yao. "Programmable quantum simulations of spin systems with trapped ions". Rev. Mod. Phys. **93**, 025001 (2021).

[11] Yuri Alexeev, Dave Bacon, Kenneth R. Brown, Robert Calderbank, Lincoln D. Carr, Frederic T. Chong, Brian DeMarco, Dirk Englund, Edward Farhi, Bill Fefferman, et al. "Quantum computer systems for scientific discovery". PRX Quantum **2**, 017001 (2021).

[12] Emanuele Pelucchi, Giorgos Fagas, Igor Aharonovich, Dirk Englund, Eden Figueroa, Qihuang Gong, Hübel Hannes, Jin Liu, Chao-Yang Lu, Nobuyuki Matsuda, et al. "The potential and global outlook of integrated photonics for quantum technologies". Nature Rev. Phys. **4**, 194–208 (2022).

[13] Guido Burkard, Thaddeus D. Ladd, Andrew Pan, John M. Nichol, and Jason R. Petta. "Semiconductor spin qubits". Rev. Mod. Phys. **95**, 025003 (2023).

[14] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller. "Statistical correlations between locally randomized measurements: A toolbox for probing entanglement in many-body quantum states". Phys. Rev. A **99**, 052323 (2019).

[15] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Predicting many properties of a quantum system from very few measurements". Nature Phys. **16**, 1050–1057 (2020).

[16] Andreas Elben, Steven T Flammia, Hsin-Yuan Huang, Richard Kueng, John Preskill, Benoît Vermersch, and Peter Zoller. "The randomized measurement toolbox". Nature Rev. Phys. **5**, 9–24 (2023).

[17] Paweł Cieśliński, Satoya Imai, Jan Dziewior, Otfried Gühne, Lukas Knips, Wiesław Laskowski, Jasmin Meinecke, Tomasz Paterek, and Tamás Vértesi. "Analysing quantum systems with randomised measurements". Physics Reports **1095**, 1–48 (2024).

[18] S. J. van Enk and C. W. J. Beenakker. "Measuring $\mathrm{Tr}\rho^n$ on single copies of ρ using random measurements". Phys. Rev. Lett. **108**, 110503 (2012).

[19] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller. "Rényi entropies from random quenches in atomic hubbard and spin models". Phys. Rev. Lett. **120**, 050406 (2018).

[20] Lukas Knips, Jan Dziewior, Waldemar Kłobus, Wiesław Laskowski, Tomasz Paterek, Peter J Shadbolt, Harald Weinfurter, and Jasmin DA Meinecke. "Multipartite entanglement analysis from random correlations". npj Quantum Inf. **6**, 51 (2020).

[21] Andreas Ketterer, Nikolai Wyderka, and Otfried Gühne. "Characterizing multipartite entanglement with moments of random correlations". Phys. Rev. Lett. **122**, 120505 (2019).

[22] Daniel Gottesman. "Stabilizer codes and quantum error correction". Ph.D. thesis, Caltech (1997), quant-ph/9705052 (1997).

[23] Matteo Ippoliti. "Classical shadows based on locally-entangled measurements". Quantum **8**, 1293 (2024).

[24] Martin Kliesch and Ingo Roth. "Theory of quantum system certification". PRX Quantum **2**, 010201 (2021).

[25] Jose Carrasco, Andreas Elben, Christian Kokail, Barbara Kraus, and Peter Zoller. "Theoretical and experimental perspectives of quantum verification". PRX Quantum **2**, 010102 (2021).

[26] Marcus P. da Silva, Olivier Landon-Cardinal, and David Poulin. "Practical characterization of quantum devices without tomography". Phys. Rev. Lett. **107**, 210404 (2011).

[27] Steven T. Flammia and Yi-Kai Liu. "Direct fidelity estimation from few pauli measurements". Phys. Rev. Lett. **106**, 230501 (2011).

[28] Leandro Aolita, Christian Gogolin, Martin Kliesch, and Jens Eisert. "Reliable quantum certification of photonic state preparations". Nature Comm. **6**, 8498 (2015).

[29] M. Gluza, M. Kliesch, J. Eisert, and L. Aolita. "Fidelity witnesses for fermionic quantum simulations". Phys. Rev. Lett. **120**, 190501 (2018).

[30] Yuki Takeuchi and Tomoyuki Morimae. "Verification of many-qubit states". Phys. Rev. X **8**, 021060 (2018).

[31] Michael A Nielsen and Isaac Chuang. "Quantum computation and quantum information". Cambridge University Press. (2002).

[32] Christian Bertoni, Jonas Haferkamp, Marcel Hinsche, Marios Ioannou, Jens Eisert, and Hakop Pashayan. "Shallow shadows: Expectation estimation using low-depth random clifford circuits". Phys. Rev. Lett. **133**, 020602 (2024).

[33] Ahmed A Akhtar, Hong-Ye Hu, and Yi-Zhuang You. "Scalable and flexible classical shadow tomography with tensor networks". Quantum **7**, 1026 (2023).

[34] Mirko Arienzo, Markus Heinrich, Ingo Roth, and Martin Kliesch. "Closed-form analytic expressions for shadow estimation with brickwork circuits". QIC **23**, 961 (2023).

[35] Matteo Ippoliti, Yaodong Li, Tibor Rakovszky, and Vedika Khemani. "Operator relaxation and the optimal depth of classical shadows". Phys. Rev. Lett. **130**, 230403 (2023).

[36] Hong-Ye Hu, Andi Gu, Swarnadeep Majumder, Hang Ren, Yipei Zhang, Derek S. Wang, Yi-Zhuang You, Zlatko Minev, Susanne F. Yelin, and Alireza Seif. "Demonstration of robust and efficient quantum property learning with shallow shadows". Nature Communications **16**, 2943 (2025).

[37] Renato M S Farias, Raghavendra D Peddinti, Ingo Roth, and Leandro Aolita. "Robust ultra-shallow shadows". Quantum Science and Technology **10**, 025044 (2025).

[38] Joonas Malmi, Keijo Korhonen, Daniel Cavalcanti, and Guillermo García-Pérez. "Enhanced observable estimation through classical optimization of informationally overcomplete measurement data: Beyond classical shadows". Physical Review A**109** (2024).

[39] Stefano Mangini and Daniel Cavalcanti. "Low-variance observable estimation with informationally-complete measurements and tensor networks" (2024). arXiv:2407.02923.

[40] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. "Quantum supremacy using a programmable superconducting processor". Nature **574**, 505–510 (2019).

[41] Charles Neill, Pedran Roushan, K Kechedzhi, Sergio Boixo, Sergei V Isakov,

V Smelyanskiy, A Megrant, B Chiaro, A Dunsworth, K Arya, et al. "A blueprint for demonstrating quantum supremacy with superconducting qubits". Science **360**, 195–199 (2018).

[42] Daniel Gottesman. "Theory of fault-tolerant quantum computation". Phys. Rev. A **57**, 127–137 (1998).

[43] Piotr Sierant, Marco Schirò, Maciej Lewenstein, and Xhek Turkeshi. "Entanglement growth and minimal membranes in $(d+1)$ random unitary circuits". Phys. Rev. Lett. **131**, 230403 (2023).

[44] Hrant Gharibyan, Masanori Hanada, Stephen H Shenker, and Masaki Tezuka. "Onset of random matrix behavior in scrambling systems". JHEP **2018**, 1–62 (2018).

[45] Lorenzo Piroli, Christoph Sünderhauf, and Xiao-Liang Qi. "A random unitary circuit model for black hole evaporation". JHEP **2020**, 1–35 (2020).

[46] Tiff Brydges, Andreas Elben, Petar Jurcevic, Benoit Vermersch, Christine Maier, Ben P Lanyon, Peter Zoller, Rainer Blatt, and Christian F Roos. "Probing rényi entanglement entropy via randomized measurements". Science **364**, 260–263 (2019).

[47] Kaifeng Bu, Dax Enshan Koh, Roy J Garcia, and Arthur Jaffe. "Classical shadows with pauli-invariant unitary ensembles". npj Quantum Inf. **10**, 6 (2024).

[48] Pietro Silvi, Ferdinand Tschirsich, Matthias Gerster, Johannes Jünemann, Daniel Jaschke, Matteo Rizzi, and Simone Montangero. "The Tensor Networks Anthology: Simulation techniques for many-body quantum lattice systems". SciPost Phys. Lect. NotesPage 8 (2019).

[49] Joseph M Renes, Robin Blume-Kohout, Andrew J Scott, and Carlton M Caves. "Symmetric informationally complete quantum measurements". J. Math. Phys. **45**, 2171–2180 (2004).

[50] Andris Ambainis and Joseph Emerson. "Quantum t-designs: t-wise independence in the quantum world". In Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07). Pages 129–140. (2007).

[51] Aram W Harrow and Saeed Mehraban. "Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates". Comm. Math. Phys. **401**, 1531–1626 (2023).

[52] Shivan Mittal and Nicholas Hunter-Jones. "Local random quantum circuits form approximate designs on arbitrary architectures" (2023). url: https://arxiv.org/abs/2310.19355.

[53] Sergio Boixo, Sergei V Isakov, Vadim N Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J Bremner, John M Martinis, and Hartmut Neven. "Characterizing quantum supremacy in near-term devices". Nature Physics **14**, 595–600 (2018).

[54] Yunchao Liu, Matthew Otten, Roozbeh Bassirianjahromi, Liang Jiang, and Bill Fefferman. "Benchmarking near-term quantum computers via random circuit sampling" (2021). url: https://arxiv.org/abs/2105.05232.

[55] Yulin Wu, Wan-Su Bao, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan, Ming Gong, Cheng Guo, Chu Guo, Shaojun Guo, Lianchen Han, Linyin Hong, He-Liang Huang, Yong-Heng Huo, Liping Li, Na Li, Shaowei Li, Yuan Li, Futian Liang, Chun Lin, Jin Lin, Haoran Qian, Dan Qiao, Hao Rong, Hong Su, Lihua Sun, Liangyuan Wang, Shiyu Wang, Dachao Wu, Yu Xu, Kai Yan, Weifeng Yang, Yang Yang, Yangsen Ye, Jianghan Yin, Chong Ying, Jiale Yu, Chen Zha, Cha Zhang, Haibin Zhang, Kaili Zhang, Yiming Zhang, Han Zhao, Youwei Zhao, Liang Zhou, Qingling Zhu, Chao-Yang Lu, Cheng-Zhi Peng, Xiaobo Zhu, and Jian-Wei Pan. "Strong quantum computational advantage using a superconducting quantum processor". Phys. Rev. Lett. **127**, 180501 (2021).

[56] Joonhee Choi, Adam L Shaw, Ivaylo S Madjarov, Xin Xie, Ran Finkelstein, Jacob P Covey, Jordan S Cotler, Daniel K Mark, Hsin-Yuan Huang, Anant Kale, et al. "Preparing random states and benchmarking with many-body quantum chaos". Nature **613**, 468–473 (2023).

[57] Xun Gao, Marcin Kalinowski, Chi-Ning Chou, Mikhail D. Lukin, Boaz Barak, and Soonwon Choi. "Limitations of linear cross-entropy as a measure for quantum advantage". PRX Quantum **5**, 010334 (2024).

[58] T. I. Andersen, N. Astrakhantsev, A. H. Karamlou, J. Berndtsson, J. Motruk, A. Szasz, J. A. Gross, A. Schuckert, T. Westerhout, Y. Zhang, E. Forati, et al. "Thermalization and criticality on an analogue–digital quantum simulator". Nature **638**, 79–85 (2025).

[59] J. Helsen, I. Roth, E. Onorati, A.H. Werner, and J. Eisert. "General framework for randomized benchmarking". PRX Quantum **3**, 020357 (2022).

[60] Markus Heinrich, Martin Kliesch, and Ingo Roth. "Randomized benchmarking with random quantum circuits" (2022). url: https://arxiv.org/abs/2212.06181.

[61] Adam L. Shaw, Zhuo Chen, Joonhee Choi, Daniel K. Mark, Pascal Scholl, Ran Finkelstein, Andreas Elben, Soonwon Choi, and Manuel Endres. "Benchmarking highly entangled states on a 60-atom analogue quantum simulator". Nature **628**, 71–77 (2024).

[62] Xueyue Zhang, Eunjong Kim, Daniel K Mark, Soonwon Choi, and Oskar Painter. "A superconducting quantum simulator based on a photonic-bandgap metamaterial". Science **379**, 278–283 (2023). url: https://doi.org/10.1126/science.ade7651.

[63] Huangjun Zhu, Richard Kueng, Markus Grassl, and David Gross. "The clifford group fails gracefully to be a unitary 4-design" (2016). url: https://arxiv.org/abs/1609.08172.

[64] Andrew C Potter and Romain Vasseur. "Entanglement dynamics in hybrid quantum circuits". In Entanglement in Spin Chains: From Theory to Quantum Technology Applications. Pages 211–249. Springer (2022).

[65] Matthew P.A. Fisher, Vedika Khemani, Adam Nahum, and Sagar Vijay. "Random quantum circuits". Annual Rev. Cond. Matt. Phys.**14** (2023).

[66] Xhek Turkeshi and Piotr Sierant. "Hilbert space delocalization under random unitary circuits". Entropy**26** (2024). url: https://www.mdpi.com/1099-4300/26/6/471.

[67] Alexander M. Dalzell, Nicholas Hunter-Jones, and Fernando G. S. L. Brandão. "Random quantum circuits anticoncentrate in log depth". PRX Quantum **3**, 010333 (2022).

[68] Mark Fannes, Bruno Nachtergaele, and Reinhard F Werner. "Finitely correlated states on quantum spin chains". Comm. Math. Phys. **144**, 443–490 (1992).

[69] D Perez-Garcia, F Verstraete, MM Wolf, and JI Cirac. "Matrix product state representations". Quantum Inf. Comp. **7**, 401–430 (2007). url: https://arxiv.org/abs/quant-ph/0608197.

[70] J Ignacio Cirac, David Perez-Garcia, Norbert Schuch, and Frank Verstraete. "Matrix product density operators: Renormalization fixed points and boundary theories". Ann. Phys. **378**, 100–149 (2017).

[71] Daniel M Greenberger, Michael A Horne, and Anton Zeilinger. "Bell's theorem, quantum theory, and conceptions of the universe, edited by m. kafatos". Kluwer Dordrecht. (1989).

[72] Hans J. Briegel and Robert Raussendorf. "Persistent entanglement in arrays of interacting particles". Phys. Rev. Lett. **86**, 910–913 (2001).

[73] Hsin-Yuan Huang, John Preskill, and Mehdi Soleimanifar. "Certifying almost all quantum states with few single-qubit measurements". In 2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS). Pages 1202–1206. (2024).

[74] Hong-Ye Hu, Soonwon Choi, and Yi-Zhuang You. "Classical shadow tomography with locally scrambled quantum dynamics". Phys. Rev. Res. **5**, 023027 (2023).

[75] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". Phys. Rev. A **70**, 052328 (2004).

[76] K. J. Satzinger, Y.-J Liu, A. Smith, C. Knapp, M. Newman, C. Jones, Z. Chen, C. Quintana, X. Mi, A. Dunsworth, et al. "Realizing topologically ordered states on a quantum processor". Science **374**, 1237–1241 (2021).

[77] Roman Stricker, Michael Meth, Lukas Postler, Claire Edmunds, Chris Ferrie, Rainer Blatt, Philipp Schindler, Thomas Monz, Richard Kueng, and Martin Ringbauer. "Experimental single-setting quantum state tomography". PRX Quantum **3**, 040310 (2022).

[78] D. Zhu, Z. P. Cian, C. Noel, A. Risinger, D. Biswas, L. Egan, Y. Zhu, A. M. Green, C. Huerta Alderete, N. H. Nguyen, et al. "Cross-platform comparison of arbitrary quantum states". Nature Comm.**13** (2022). url: https://doi.org/10.1038%2Fs41467-022-34279-5.

[79] J. C. Hoke, M. Ippoliti, E. Rosenberg, D. Abanin, R. Acharya, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, et al. "Measurement-induced entanglement and teleportation on a noisy quantum processor". Nature **622**, 481–486 (2023).

[80] Vittorio Vitale, Aniket Rath, Petar Jurcevic, Andreas Elben, Cyril Branciard, and Benoît Vermersch. "Robust estimation of the quantum fisher information on a quantum processor". PRX Quantum **5**, 030338 (2024).

[81] Benoît Vermersch, Marko Ljubotina, J. Ignacio Cirac, Peter Zoller, Maksym Serbyn, and Lorenzo Piroli. "Many-body entropies and entanglement from polynomially many local measurements". Phys. Rev. X **14**, 031035 (2024).

[82] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. "Error mitigation for short-depth quantum circuits". Phys. Rev. Lett. **119**, 180509 (2017).

[83] Sam McArdle, Xiao Yuan, and Simon Benjamin. "Error-mitigated digital quantum simulation". Phys. Rev. Lett. **122**, 180501 (2019).

[84] Bálint Koczor. "Exponential error suppression for near-term quantum devices". Phys. Rev. X **11**, 031057 (2021).

[85] Piotr Czarnik, Andrew Arrasmith, Patrick J Coles, and Lukasz Cincio. "Error mitigation with clifford quantum-circuit data". Quantum **5**, 592 (2021).

[86] Armands Strikis, Dayue Qin, Yanzhu Chen, Simon C. Benjamin, and Ying Li. "Learning-based quantum error mitigation". PRX Quantum **2**, 040330 (2021).

[87] Christophe Piveteau, David Sutter, and Stefan Woerner. "Quasiprobability decompositions with reduced sampling overhead". npj Quantum Information **8**, 12 (2022).

[88] Yuchen Guo and Shuo Yang. "Quantum error mitigation via matrix product operators". PRX Quantum **3**, 040313 (2022).

[89] Ewout Van Den Berg, Zlatko K Minev, Abhinav Kandala, and Kristan Temme. "Probabilistic error cancellation with sparse pauli–lindblad models on noisy quantum processors". Nature Phys. **19**, 1116–1121 (2023).

[90] Sergei Filippov, Matea Leahy, Matteo AC Rossi, and Guillermo García-Pérez. "Scalable tensor-network error mitigation for near-term quantum computing" (2023). url: https://arxiv.org/abs/2307.11740.

[91] Stefano Mangini, Marco Cattaneo, Daniel Cavalcanti, Sergei Filippov, Matteo A. C. Rossi, and Guillermo García-Pérez. "Tensor network noise characterization for near-term quantum computers". Phys. Rev. Res. **6**, 033217 (2024).

[92] Thomas Schuster, Jonas Haferkamp, and Hsin-Yuan Huang. "Random unitaries in extremely low depth" (2024). url: https://arxiv.org/abs/2407.07754v1.

[93] Nicholas LaRacuente and Felix Leditzky. "Approximate unitary k-designs from shallow, low-communication circuits". 2407.07876 (2024). url: https://arxiv.org/abs/2407.07876.

[94] Xhek Turkeshi and Riccardo Cioli. "Data and code for "approximate inverse measurement channel for shallow shadows"" (2025). https://doi.org/10.5281/zenodo.15118326.

[95] Craig Gidney. "Stim: a fast stabilizer circuit simulator". Quantum **5**, 497 (2021).

[96] Matthew Fishman, Steven R. White, and E. Miles Stoudenmire. "The ITensor Software Library for Tensor Network Calculations". SciPost Phys. Codebases-Page 4 (2022).

[97] Matthew Fishman, Steven R. White, and E. Miles Stoudenmire. "Codebase release 0.3

for ITensor". SciPost Phys. CodebasesPages 4–r0.3 (2022).