

Resource-efficient shadow tomography using equatorial stabilizer measurements

Guedong Park[✉], Yong Siah Teo^{✉,*}, and Hyunseok Jeong^{✉,†}*NextQuantum Innovation Research Center, Department of Physics and Astronomy,
Seoul National University, Seoul 08826, Republic of Korea*

(Received 2 April 2025; accepted 25 June 2025; published 28 July 2025)

We propose a resource-efficient shadow-tomography scheme using equatorial-stabilizer measurements generated from subsets of Clifford unitaries. For n -qubit systems, equatorial-stabilizer-based shadow-tomography schemes can estimate M observables (up to an additive error ε) using $\mathcal{O}(\log(M), \text{poly}(n), 1/\varepsilon^2)$ sampling copies for a large class of observables, including those with traceless parts possessing polynomially bounded Frobenius norms. For arbitrary quantum-state observables with a constant Frobenius norm, sampling complexity becomes n independent. Our scheme only requires an n -depth controlled-Z (CZ) circuit [$\mathcal{O}(n^2)$ CZ gates] and Pauli measurements per sampling copy. Alternatively, our scheme is realizable with $2n$ -depth circuits comprising n^2 nearest-neighboring CNOT gates, exhibiting a smaller maximal gate count relative to previously known randomized-Clifford-based proposals. We numerically confirm our theoretically derived shadow-tomographic sampling complexities with random pure states and multiqubit graph states. Finally, we demonstrate that equatorial-stabilizer-based shadow tomography is more noise tolerant than randomized-Clifford-based schemes in terms of fidelity estimation for the Greenberger–Horne–Zeilinger state and W state.

DOI: [10.1103/9pbp-jzr9](https://doi.org/10.1103/9pbp-jzr9)

I. INTRODUCTION

Quantum algorithms, which are algorithms based on quantum-mechanical principles, have been shown to outperform classical algorithms for many computational tasks [1–3]. For these purposes, certifying the quality of the prepared (noisy) quantum state is an important prerequisite. Quantum tomography [4–7] is a general method to reconstruct a typically unknown noisy n -qubit quantum state ρ (or a channel). This, however, requires exponentially many copies of ρ in n for an accurate reconstruction.

On the other hand, if we are to estimate certain classes of physical properties of unknown states, shadow tomography [8] enables us to do so with a polynomially large number of samples. The well-known randomized Pauli tomography, or six-Paulis tomography [9,10], is useful for estimating the expectation values of Hamiltonians with low-weighted Pauli observables, learning quantum channels [11] and measuring nonstabilizerness (or magic) [12,13]. This procedure uses a simple single-layer circuit structure and is therefore ineffective in estimating highly entangled properties of large quantum systems. The second protocol, randomized-Clifford shadow tomography [9], is better for estimating the fidelity between the unknown input and entangled target states for large qubit number n . This advantage originates from the fact that

Clifford operations are unitary 3-designs [14–17]. However, for large n , this protocol utilizes at most $\frac{7n^2}{2}$ neighboring CNOT gates [18,19]. Therefore, shadow tomography with arbitrary Clifford gates could lead to noisy measurement results in real experiments as the large number of noisy gates accumulates physical noise.

To cope with such problems, recent works have shown that target observables, on average over suitable ensembles of input states, possess good shadow norms even with neighboring gates of shallow depths [20–22]. However, it is uncertain whether arbitrary state inputs reside in the average case, and such a shallow shadow [20] preserves the shadow norm scaling and thus the efficiency of input-state copy number. It is known that generalized measurements [10,23,24] can reduce the gate count or sampling complexity. However, noise-tolerant implementation of such non-Clifford measurements is believed to be challenging [25–27]. There exist alternatives with reduced Clifford-gate resources [28,29]. Nevertheless, applicable input states and target observables that lead to $\text{poly}(n)$ -shadow norms are restrictive. Hence, a complete analysis of an optimal tomography scheme for broad classes of observables remains an open problem.

In this work, we introduce resource-efficient schemes for shadow tomography, catered to observable-expectation-value estimations, that offer a lower gate count and circuit depth while still being applicable to a large observable class. The key point is that it is not necessary to utilize the entire set of stabilizer states [30,31] comprising all Clifford-rotated bases. Instead, our proposed schemes employ a smaller stabilizer subset, namely, the equatorial stabilizer positive operator-valued measure (ESPOVM) [32] together with the computational basis. Additionally, it is sufficient to perform shadow tomography with an even smaller subset, namely, the real ESPOVM (RESPOVM), when the target observable is

*Contact author: ys_teo@snu.ac.kr†Contact author: h.jeong37@gmail.com

real with respect to the computational basis. Our scheme only involves a uniform sampling of commuting CZ gates as the only two-qubit gate resources.

Next, we theoretically prove shadow-tomographic capabilities for (R)ESPOVM schemes by deriving upper bounds of sampling-copy complexity for the observable-expectation-value estimation, which is directly related to shadow norm [9,20]. Specifically, we note that a sampling complexity of $\mathcal{O}(\log(M), \text{poly}(n), 1/\varepsilon^2)$ suffices to simultaneously estimate (up to an additive error ε) M complex observables belonging to a broad class, which sufficiently include observables possessing polynomially bounded Frobenius norms for their traceless parts. For arbitrary quantum-state observables (fidelity estimation), the sampling complexity is reduced to simply $\mathcal{O}(\log(M), 1/\varepsilon^2)$, which is constant in n and on par with randomized-Clifford tomography [9]. These results are promising when compared with recent results [33,34] which utilize smaller subsets of CZ-circuit-based measurements consisting of mutually unbiased bases (MUBs) [35] for estimation of observable expectation values, but has an exponentially large bound of sampling copies to estimate arbitrary state observables.

Furthermore, we shall show that our scheme can be implemented by a more simplified circuit structure, and attain improved noise tolerance. These results are derived from the fact that it involves a uniform sampling of commuting CZ gates as the only two-qubit gate resources. This enables our scheme to require only n depth of CZ gates or $2n$ depth, hence at most $n^2 < \frac{7n^2}{2}$ nearest neighboring (NN) CNOT gates [36]. Such a scaling factor optimization renders our scheme to become more robust to gate noise. Reference [20] analyzed the trade-off among the sampling, computational complexities, and required circuit-depth scaling, from constant to linear depth. On the other hand, our work aims to reduce the required depth while preserving the efficiency of sampling and computational complexity for arbitrary observables. Recent work [37] showed that small-biased shadow tomography is possible even with log-scaled depth. However, we observed that our circuit depth is still much lower in the intermediate

qubit number regime, where many practical cases of fidelity estimation reside.

In addition, we show that the estimator of (R)ESPOVM shadow tomography itself can preserve its output value from the various measurement bit-flip errors. This observation suggests another crucial structural consideration for Clifford circuits in terms of noise tolerance, aside from reducing the gate complexity. If neighboring gates have negligible noise, the state-injection-based implementation of (R)ESPOVM measurements to tailor the arbitrary noise into the dephasing noise reinforces such maintenance to the noise in long-ranged CZ gates. We provide its numerical evidence in the fidelity estimation results of Greenberger–Horne–Zeilinger (GHZ) state [38] and W states [39].

II. PRELIMINARIES

A. Clifford group and equatorial stabilizer POVM

Before presenting our main results, we define and explain several concepts and terminologies that shall be used throughout the article. We consider a source that allows one to prepare multiple copies of an n -qubit quantum state ρ . The corresponding observable O shall be a Hermitian operator and is therefore symmetric whenever its matrix elements in the computational basis are all real, at which we say O is real. We define the Clifford group Cl_n as the set of operators [30,40],

$$\text{Cl}_n = \{U \mid \forall E \in \mathcal{P}_n, UEU^\dagger \in \mathcal{P}_n\}, \quad (1)$$

where $\mathcal{P}_n$ is the Pauli group generated by tensor products of the standard Pauli operators X, Y, Z and the identity I . Next, we define the stabilizer(-states) set $\mathcal{S}_n$ as the following set of quantum states:

$$\mathcal{S}_n = \{|\psi\rangle \mid |\psi\rangle = U|0^{\otimes n}\rangle \text{ for some } U \in \text{Cl}_n\}, \quad (2)$$

where the element $|\psi\rangle \in \mathcal{S}_n$ is a (pure) stabilizer state.

We are now ready to look into subclasses of the stabilizer set $\mathcal{S}_n$. Specifically, the complex (pure) equatorial stabilizer (-states) set [32] $\mathcal{S}_n^{\text{eq}}$ and its real counterpart $\mathcal{S}_n^{\text{req}}$ are, respectively, given by

$$\mathcal{S}_n^{\text{eq}} = \left\{ |\phi_{\mathbf{A}}^{\text{eq}}\rangle \equiv \frac{1}{\sqrt{2^n}} \sum_{\mathbf{x} \in \mathbb{Z}_2^n} i^{\mathbf{x}^\top \mathbf{A} \mathbf{x}} |\mathbf{x}\rangle \mid \mathbf{A} = (a_{ij})_{i,j \in [n]} \text{ where } a_{ij} = a_{ji}, a_{ij} \in \begin{cases} \mathbb{Z}_4 & \text{if } i = j \\ \mathbb{Z}_2 & \text{if } i > j \end{cases} \right\}, \quad (3)$$

$$\mathcal{S}_n^{\text{req}} = \left\{ |\phi_{\mathbf{A}}^{\text{req}}\rangle \equiv \frac{1}{\sqrt{2^n}} \sum_{\mathbf{x} \in \mathbb{Z}_2^n} (-1)^{\mathbf{x}^\top \mathbf{A} \mathbf{x}} |\mathbf{x}\rangle \mid \mathbf{A} = (a_{ij})_{i,j \in [n]} \text{ where } a_{ij} \in \begin{cases} \mathbb{Z}_2 & \text{if } i \leq j \\ \{0\} & \text{if } i > j \end{cases} \right\}. \quad (4)$$

From Eqs. (3) and (4), we know that the number of complex equatorial stabilizer states is $|\mathcal{S}_n^{\text{eq}}| = 2^{\frac{n^2+3n}{2}}$, and that of real equatorial stabilizer states is $|\mathcal{S}_n^{\text{req}}| = 2^{\frac{n^2+n}{2}}$. As the names suggest, these are subsets of $\mathcal{S}_n$. The $\mathcal{S}_n^{\text{eq}}$ was originally used for a faster classical simulation [31,32]. More explicitly, the original norm estimation technique [31], which is a key step of classical simulation, involves the uniform sampling of stabilizer states and taking the inner product with the target magic

state. However, if we restrict the sampling set to the subset $\mathcal{S}_n^{\text{eq}} \subsetneq \mathcal{S}_n$, we may boost the speeds of sampling and inner product calculation [32].

In this work, we shall turn these two stabilizer subclasses into POVMs for shadow tomography, and consequently reveal the key advantages that are related to the exploitation of their smaller set sizes relative to $\mathcal{S}_n$. That $\mathcal{S}_n^{\text{eq}}$ and $\mathcal{S}_n^{\text{req}}$ may be readily transformed into POVMs is straightforward [41], as the corresponding sets of positive operators $\{\frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} |\phi_{\mathbf{A}}^{\text{eq}}\rangle \langle \phi_{\mathbf{A}}^{\text{eq}}|\}$

and $\{\frac{2^n}{|\mathcal{S}_n^{\text{req}}|} |\phi_A^{\text{req}}\rangle \langle \phi_A^{\text{req}}|\}$ house elements that sum to the identity (see Appendix A for the proof). We coin them (real) equatorial stabilizer POVMs, or (R)ESPOVMs for short.

B. Random Clifford tomography and classical shadows

One of the typical methods for estimating the expectation value of a given observable O is randomized-Clifford shadow tomography [9]. We briefly explain how the random Clifford operation may be used to estimate $\langle O \rangle = \text{tr}(\rho O)$ (see Ref. [9] for details). For this purpose, the algorithm using N sampling-copy number is stated below:

- (1) Take ρ as an input and uniformly choose a Clifford unitary U from Cl_n .
- (2) Take $\rho \leftarrow U \rho U^\dagger$.
- (3) Measure in the Z basis and obtain the outcome $\mathbf{p} \in \mathbb{Z}_2^n$.
- (4) Compute

$$\begin{aligned} \widehat{O} &= \text{tr}(O \mathcal{M}^{-1}(U^\dagger |\mathbf{p}\rangle \langle \mathbf{p}| U)) \\ &= (2^n + 1) \langle \mathbf{p} | U O U^\dagger | \mathbf{p} \rangle - \text{tr}(O). \end{aligned} \quad (5)$$

Repeat the aforementioned steps N times and take the average of N $\widehat{O}$'s to obtain the estimator $\langle \widehat{O} \rangle$ for $\langle O \rangle$.

Here, $\mathcal{M}^{-1}(O) \equiv (2^n + 1)O - \text{tr}(O)I$ for an arbitrary operator O is known as the classical shadow [9]. Reference [9] showed that such randomized-Clifford tomography achieves ε -accurate estimation of $\langle O \rangle$ with the sampling-complexity upper bound $\mathcal{O}(3\text{tr}(O^2)/\varepsilon^2)$. Hence, this scheme is efficient when O is a target pure state (fidelity estimation) of which $\text{tr}(O^2)$ is a constant, albeit with the assumption that every $\widehat{O}$ is efficiently computable. Reference [19] showed that we can efficiently sample a Clifford unitary U from Cl_n in $\mathcal{O}(n^2)$ time. In addition, the sampled form contains the HF—SW—H—HF sections (where small uppercase letters here refer to gate layers as opposed to the regular uppercase ones that denote gate labels, and HF and HF' are Hadamard-free sections), where the HF section is a circuit containing layers of $\text{CNOT}^\dagger$ —CZ—S—PAULI. Here, H, SW, $\text{CNOT}^\dagger$, CZ and, S, respectively, refer to pure layers of Hadamard gates, SWAP gates, CNOT gates with control qubits higher than the target ones when visualized in a standard circuit diagram, CZ gates and S gates. Naturally, SW contains many CNOT gates.

Direct fidelity-estimation scheme [28] is another conventional way for estimating fidelities. While this method invokes a lower gate count and circuit depth than our schemes, the estimable region of observables is more restricted than both the randomized-Clifford and our (R)ESPOVM schemes. As such, only the randomized-Clifford tomography schemes shall be of interest to be compared with our work.

III. EQUATORIAL-STABILIZER SHADOW TOMOGRAPHY

We now introduce the theoretical framework of the shadow tomography scheme that uses (real) equatorial stabilizer POVMs. We recall that we want to estimate the target value,

$\text{tr}(\rho O)$. We can show that it can be rewritten as

$$\begin{aligned} \text{tr}(\rho O) &= \sum_{|\phi_A^{\text{eq}}\rangle \in \mathcal{S}_n^{\text{eq}}} \frac{2^{2n}}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_A^{\text{eq}} | \rho | \phi_A^{\text{eq}} \rangle \langle \phi_A^{\text{eq}} | O | \phi_A^{\text{eq}} \rangle \\ &\quad + \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \rho | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle - \text{tr}(O). \end{aligned} \quad (6)$$

If O is a real matrix with respect to the computational basis,

$$\begin{aligned} \text{tr}(\rho O) &= \sum_{|\phi_A^{\text{req}}\rangle \in \mathcal{S}_n^{\text{req}}} \frac{2^{2n-1}}{|\mathcal{S}_n^{\text{req}}|} \langle \phi_A^{\text{req}} | \rho | \phi_A^{\text{req}} \rangle \langle \phi_A^{\text{req}} | O | \phi_A^{\text{req}} \rangle \\ &\quad + \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \rho | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle - \frac{1}{2} \text{tr}(O). \end{aligned} \quad (7)$$

A detailed proof can be seen in Appendix A. In Sec. II A, we explained that $\{\frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|\}$ and $\{\frac{2^n}{|\mathcal{S}_n^{\text{req}}|} |\phi_A^{\text{req}}\rangle \langle \phi_A^{\text{req}}|\}$ are POVMs. As a result, we obtain the following equality:

$$\sum_{\mathbf{A}} \frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_A^{\text{eq}} | \rho | \phi_A^{\text{eq}} \rangle = \sum_{\mathbf{A}} \frac{2^n}{|\mathcal{S}_n^{\text{req}}|} \langle \phi_A^{\text{req}} | \rho | \phi_A^{\text{req}} \rangle = 1. \quad (8)$$

Therefore, we can sample the $\mathbf{A}$, which represents a specific (real) equatorial stabilizer state, from the probability distributions $\{\frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_A^{\text{eq}} | \rho | \phi_A^{\text{eq}} \rangle\}$ and $\{\frac{2^n}{|\mathcal{S}_n^{\text{req}}|} \langle \phi_A^{\text{req}} | \rho | \phi_A^{\text{req}} \rangle\}$.

Now we present the estimation scheme for the ESPOVM. The RESPOVM follows in a similar manner and will be shown in Appendix A. We prepare two copies of ρ [N (even) copies in total]. The first copy is measured to obtain the outcome $\mathbf{A}$ by the distribution $\{\frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_A^{\text{eq}} | \rho | \phi_A^{\text{eq}} \rangle\}$, and the second copy is measured by the computational basis to obtain the outcome $\mathbf{p}' \in \mathbb{Z}_2^n$. Next, we take the estimator $\widehat{O}$ as

$$\widehat{O} = 2^n \langle \phi_A^{\text{eq}} | O | \phi_A^{\text{eq}} \rangle + \langle \mathbf{p}' | O | \mathbf{p}' \rangle - \text{tr}(O). \quad (9)$$

We repeat such measurements many times to have many copies of $\widehat{O}$'s and take the average to obtain the desired estimation $\langle \widehat{O} \rangle$. To further reduce the sampling copies, we can apply the median of means (MOM) estimation technique [42–44] by simply repeating the scheme and take the median of the collection of estimation values. The complete systematic procedure for (R)ESPOVM shadow tomography is shown in Algorithm 1 whose schematic illustration is given in Fig. 1. The theoretical validation of such algorithms, including the one of RESPOVM which is derived by a similar logic, will be shown in Appendix A.

Let us supply the details for the main steps. From the perspective of classical shadows [8,9], we note that our (R)ESPOVM tomography operates with the following one-to-two-copy measurement channel $\mathcal{M}$:

$$\begin{aligned} \mathcal{M}(\rho) &= \left(\sum_{\mathbf{A}} \frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_A^{\text{eq}} | \rho | \phi_A^{\text{eq}} \rangle |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|, \sum_{\mathbf{x}} \langle \mathbf{x} | \rho | \mathbf{x} \rangle |\mathbf{x}\rangle \langle \mathbf{x}| \right). \end{aligned} \quad (10)$$

The measurement channel for RESPOVM similarly follows. Hence, the scheme outputs the following respective classical

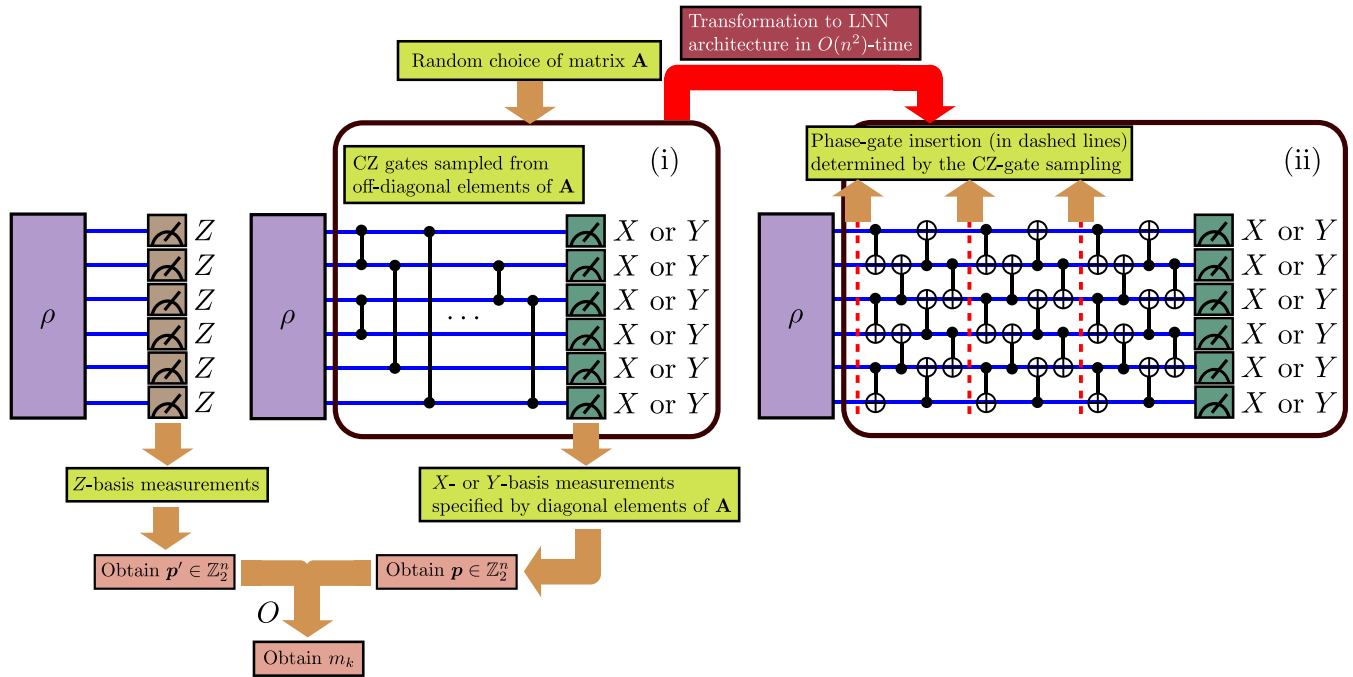


FIG. 1. Schematic for six-qubit (R)ESPOVM shadow tomography to estimate $\langle O \rangle = \text{tr}(\rho O)$ of an unknown state ρ and target observable O . Sampled CZ circuits can be transformed to a linear nearest-neighboring (LNN) architecture [36] [circuit (ii)]. Also in (ii), the inserted phase gates are single-qubit gates and the measurement is done in the same manner as in (i). Any one of the two architectures (i) and (ii) can be executed to obtain the measurement bit-string column $p \in \mathbb{Z}_2^n$. Together with the bit-string column p' acquired from the computational-basis measurement, the observable expectation value can be reconstructed after repeating N rounds of (p, p') collection.

shadows (inverted channel map $\mathcal{M}^{-1}$ [9]):

$$\begin{aligned} \mathcal{M}^{-1}(O_1, O_2) &= \begin{cases} 2^n O_1 + O_2 - I & (\text{ESPOVM} + \text{comp. basis}) \\ 2^{n-1} O_1 + O_2 - \frac{I}{2} & (\text{RESPOVM} + \text{comp. basis}). \end{cases} \end{aligned} \quad (11)$$

Such a one-to-two-copy measurement channel again implies that we need two copies of ρ for the single measurement step. We can note that indeed, $\mathcal{M}^{-1} \circ \mathcal{M}(O) = O$ (see Appendix A for details) for arbitrary (real part for RESPOVM) Hermitian operator O .

The reason why we need the computational basis measurement is that the ESPOVM is *not* informationally complete (IC), whereas the union of the ESPOVM and computational basis is IC [41,45]. In other words, along with the computational basis measurements, we are capable of reconstructing an arbitrary ρ and, thus, general observable expectation values. So, two copies of ρ are required for a single measurement trial, which would give us a bit-string column p from the ESPOVM and another bit-string column p' from the computational basis. However, in Sec. IV, we shall see that only a single copy of the input state is necessary for each trial if O is a cluster state.

From Algorithm 1, we see that ESPOVM measurement requires S , H gates before the computational basis measurements, but RESPOVM does not need S gates. Equivalently, the ESPOVM needs X , Y -basis measurements while RESPOVM only needs X -basis measurement. While both structures need

similar two-copy resources for the single copy measurement, we shall see in the next section that in many cases, the RESPOVM requires much fewer sampling-copy numbers than ESPOVM to achieve the same estimation accuracy.

We summarize the requirements for implementing Algorithm 1 for (R)ESPOVM shadow tomography in:

Theorem 1. (R)ESPOVM implementation. On an n -qubit state ρ , the implementation of n -qubit ESPOVM shadow tomography requires CZ-gate circuits of depths of at most n by Vizing's theorem [46], and n single-qubit Pauli (X, Y, Z) measurements. For RESPOVM shadow tomography, the Y -basis measurement is not needed.

IV. SAMPLING-COPY COMPLEXITY

Shadow tomography aims to estimate state properties with a sampling complexity that grows at most polynomially quickly with the qubit number n . To be more precise, let us specify that a single property estimator is accurate up to a maximal additive error of $\varepsilon > 0$. Then, the number of sampling copies N should scale as at most $\mathcal{O}(\text{poly}(n)/\varepsilon^2)$. We shall prove that our (R)ESPOVM algorithms also possess upper bounds of similar scaling behaviors with randomized-Clifford tomography.

Before doing so, we first explain the relation between required sampling copies and the variance of the estimator. We denote $\hat{O}_A^{\text{eq}} \equiv 2^n \langle \phi_A^{\text{eq}} | O | \phi_A^{\text{eq}} \rangle$, $\hat{O}_A^{\text{req}} \equiv 2^{n-1} \langle \phi_A^{\text{req}} | O | \phi_A^{\text{req}} \rangle$, and $\hat{O}_{p'}^{\text{bin}} \equiv \langle p' | O | p' \rangle$ for one observable O that are measured and obtained in each step of the shadow-tomography procedure, where the matrix A and bit string p' are sampled

ALGORITHM 1. (R)ESPOVM shadow tomography algorithm (with the MOM estimation technique).

Input: n -qubit unknown quantum state ρ , total input-state copies $N = N'K \in \mathbb{N}$ (with even N' and $K \in \mathbb{N}$), observable O , and the (R)ESPOVM scheme

Output: Estimate m

```

1 for  $z \in [K]$  do
2    $m_z \leftarrow 0$ ;  $m' \leftarrow 0$ ;
3   for  $k \in [N'/2]$  do
4     Prepare  $\rho$  as the input state;
5     Randomly choose  $|\phi_{\mathbf{A}}^{(r)\text{eq}}\rangle \in \mathcal{S}_n^{(r)\text{eq}}$ , where
6        $\mathbf{A} = (a_{ij})$ ;
7       for  $i < j \in [n]$  do
8         if  $a_{ij} = 1$  then
9           Apply CZ $_{ij}$ :  $\rho \leftarrow \text{CZ}_{ij}\rho\text{CZ}_{ij}$ ;
10      for  $i \in [n]$  do
11        if ESPOVM is used then
12          if  $a_{ii} \equiv 1 \pmod{4}$  then
13            Apply  $S^\dagger$  on qubit  $i$ ;
14          else if  $a_{ii} \equiv 2 \pmod{4}$  then
15            Apply  $Z$  on qubit  $i$ ;
16          else if  $a_{ii} \equiv 3 \pmod{4}$  then
17            Apply  $S$  on qubit  $i$ ;
18        else if  $a_{ii} \equiv 1 \pmod{2}$  (RESPOVM) then
19          Apply  $Z$  on qubit  $i$ ;
20        Apply  $H$  on qubit  $i$ ;
21      Measure to obtain bit-string  $\mathbf{p} \in \mathbb{Z}_2^n$ ;
22      for  $i \in [n]$  do
23        if ESPOVM is used then
24           $a_{ii} \leftarrow a_{ii} + 2p_i \pmod{4}$ ;
25        else if RESPOVM is used then
26           $a_{ii} \leftarrow a_{ii} + p_i \pmod{2}$ ;
27      Prepare  $\rho$  again and measure in Pauli-Z basis to
28      obtain  $\mathbf{p}' \in \mathbb{Z}_2^n$ ;
29      if ESPOVM is used then
30         $m' \leftarrow m' + 2^n \langle \phi_{\mathbf{A}}^{\text{eq}} | O | \phi_{\mathbf{A}}^{\text{eq}} \rangle$ 
31         $\quad + \langle \mathbf{p}' | O | \mathbf{p}' \rangle - \text{tr}(O)$ 
32      else if RESPOVM is used then
33         $m' \leftarrow m' + 2^{n-1} \langle \phi_{\mathbf{A}}^{\text{req}} | O | \phi_{\mathbf{A}}^{\text{req}} \rangle$ 
34         $\quad + \langle \mathbf{p}' | O | \mathbf{p}' \rangle - \frac{1}{2}\text{tr}(O)$ 
35    $m_z \leftarrow m'/(N'/2)$ ;
36  $m \leftarrow \text{median}(m_1, m_2, \dots, m_K)$ ;

```

in accordance with Algorithm 1. The first two random numbers $\hat{O}_{\mathbf{A}}^{\text{eq}}$ and $\hat{O}_{\mathbf{A}}^{\text{req}}$ refer to the respective outcome estimators for ESPOVM and RESPOVM, and $\hat{O}_{\mathbf{p}'}^{\text{bin}}$ is the outcome of a Z-basis measurement. Upon defining $\mathbb{E}(\hat{O}_{\mathbf{A},\mathbf{p}'})$ as the analytic average value of $\hat{O}_{\mathbf{A},\mathbf{p}'}$, which is defined as

$$\hat{O}_{\mathbf{A},\mathbf{p}'} \equiv \begin{cases} \hat{O}_{\mathbf{A}}^{\text{eq}} + \hat{O}_{\mathbf{p}'}^{\text{bin}} - \text{tr}(O) & (\text{ESPOVM}) \\ \hat{O}_{\mathbf{A}}^{\text{req}} + \hat{O}_{\mathbf{p}'}^{\text{bin}} - \frac{1}{2}\text{tr}(O) & (\text{RESPOVM}), \end{cases} \quad (12)$$

the estimation variance $\text{Var}(\hat{O}_{\mathbf{A},\mathbf{p}'}^2) \equiv \mathbb{E}(\hat{O}_{\mathbf{A},\mathbf{p}'}^2) - (\mathbb{E}(\hat{O}_{\mathbf{A},\mathbf{p}'}))^2$ of our two shadow-tomography schemes is shown as

$$\begin{aligned} \text{Var}^{(r)\text{eq}}(\hat{O}_{\mathbf{A},\mathbf{p}'}) &= \text{Var}(\hat{O}_{\mathbf{A}}^{(r)\text{eq}}) + \text{Var}(\hat{O}_{\mathbf{p}'}^{\text{bin}}) \\ &= \text{Var}(\hat{O}_{0\mathbf{A}}^{(r)\text{eq}}) + \text{Var}(\hat{O}_{0\mathbf{p}'}^{\text{bin}}), \end{aligned} \quad (13)$$

where $O_0 \equiv O - \text{tr}(O)\frac{I}{2^n}$ (traceless part), where the first equality is because $(\mathbf{A}, \mathbf{p}')$ sampling is independent of each other. The variance terms on the right side, after maximizing over the input state ρ , are equivalent to the respective squared shadow norms [9] of the (R)ESPOVM and Z-basis measurement. To estimate $\langle O \rangle$, we take the sample mean $\langle \hat{O} \rangle$ over the N numbers of sampled values $\hat{O}_{\mathbf{A},\mathbf{p}'}$.

The sample variance of $\langle \hat{O} \rangle$ for estimating a single property $\langle O \rangle$ is given by

$$\begin{aligned} \text{Var}(\langle \hat{O} \rangle) &= \begin{cases} \frac{\text{Var}^{\text{eq}}(\hat{O}_{\mathbf{A},\mathbf{p}'})}{N/2} & (\text{ESPOVM, comp. basis}), \\ \frac{\text{Var}^{\text{req}}(\hat{O}_{\mathbf{A},\mathbf{p}'})}{N/2} & (\text{RESPOVM, comp. basis}). \end{cases} \end{aligned} \quad (14)$$

Note that since $\langle \hat{O} \rangle$ is unbiased, $\text{Var}(\langle \hat{O} \rangle)$ is identical to the mean-squared error of $\langle \hat{O} \rangle$ that measures the estimation accuracy of $\langle \hat{O} \rangle$ with respect to $\langle O \rangle$:

$$\text{Var}(\langle \hat{O} \rangle) = \mathbb{E}((\langle \hat{O} \rangle - \langle O \rangle)^2). \quad (15)$$

Therefore, Chebyshev inequality implies that $\text{Var}^{(r)\text{eq}}(\hat{O}_{\mathbf{A},\mathbf{p}'})$ determines the required number of sampling copies $\mathcal{O}(\frac{\text{Var}^{(r)\text{eq}}(\hat{O}_{\mathbf{A},\mathbf{p}'})}{\epsilon^2\delta})$ to achieve the desired accuracy within the failure probability δ . It is known that with the aid of the MOM technique [9,44], we can reduce such complexity to $\mathcal{O}(\frac{\text{Var}^{(r)\text{eq}}(\hat{O}_{\mathbf{A},\mathbf{p}'})}{\epsilon^2} \log(\delta^{-1}))$.

The dominant term determining the scaling of $\text{Var}^{(r)\text{eq}}(\hat{O}_{\mathbf{A},\mathbf{p}'})$ is $\text{Var}(\hat{O}_{0\mathbf{A}}^{(r)\text{eq}})$, and its upper bound $\mathbb{E}((\hat{O}_{0\mathbf{A}}^{(r)\text{eq}})^2)$ is rewritten by

$$\mathbb{E}((\hat{O}_{0\mathbf{A}}^{(r)\text{eq}})^2) = \frac{2^\gamma}{|\mathcal{S}_n^{(r)\text{eq}}|} \sum_{|\phi_{\mathbf{A}}^{(r)\text{eq}}\rangle \in \mathcal{S}_n^{(r)\text{eq}}} \langle \phi_{\mathbf{A}}^{(r)\text{eq}} | \rho | \phi_{\mathbf{A}}^{(r)\text{eq}} \rangle \langle \phi_{\mathbf{A}}^{(r)\text{eq}} | O_0 | \phi_{\mathbf{A}}^{(r)\text{eq}} \rangle \langle \phi_{\mathbf{A}}^{(r)\text{eq}} | O_0 | \phi_{\mathbf{A}}^{(r)\text{eq}} \rangle \quad (16)$$

$$= 2^\gamma \text{tr}((\rho \otimes O_0 \otimes O_0) \mathbb{E}_{|\phi_{\mathbf{A}}^{(r)\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{(r)\text{eq}})} |\phi_{\mathbf{A}}^{(r)\text{eq}}\rangle \langle \phi_{\mathbf{A}}^{(r)\text{eq}}|^{\otimes 3}), \quad \gamma = \begin{cases} 3n & \text{for eq,} \\ 3n-2 & \text{for req.} \end{cases} \quad (17)$$

Here, $\mathbb{E}_{|\phi_{\mathbf{A}}^{(r)\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{(r)\text{eq}})}$ means the analytic average over uniformly chosen (real) equatorial stabilizer states. We see that $\mathbb{E}_{|\phi_{\mathbf{A}}^{(r)\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{(r)\text{eq}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|^{\otimes 3}$ is an important factor for the variance scaling. We clarify its structure in the following lemma, which will be proved in Appendix E 1.

Lemma 1.

$$\mathbb{E}_{|\phi_A^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_A^{\text{req}}\rangle \langle \phi_A^{\text{req}}|^{\otimes 3} = \frac{1}{8^n} \sum_{(x,y,z,w,s,t) \in \mathcal{K}_1(\mathbb{Z}_2^n)} |xyz\rangle \langle wst|, \quad (18)$$

$$\mathbb{E}_{|\phi_A^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|^{\otimes 3} = \frac{1}{8^n} \sum_{(x,y,z,w,s,t) \in \mathcal{K}_2(\mathbb{Z}_2^n)} |xyz\rangle \langle wst|, \quad (19)$$

where

$$\begin{aligned} \mathcal{K}_1(\mathbb{Z}_2^n) &\equiv \{v = (x, y, z, w, s, t) | x, y, z, w, s, t \in \mathbb{Z}_2^n \exists \text{ a partition of } \{1, 2, \dots, 6\} \text{ into three pairs } \{(i_1, j_1), (i_2, j_2), \\ &\quad (i_3, j_3)\} \text{ such that } v_{i_k} = v_{j_k} \text{ for } k = 1, 2, 3\}, \\ \mathcal{K}_2(\mathbb{Z}_2^n) &\equiv \{v = (x, y, z, w, s, t) | v \in \mathcal{K}_1(\mathbb{Z}_2^n), \text{ and if } \{x, y, z\} \text{ and } \{w, s, t\} \text{ share a common element, say } u, \\ &\quad \text{and the other two variables in } \{x, y, z\} \setminus \{u\} \text{ or } \{w, s, t\} \setminus \{u\} \\ &\quad \text{equal, say } u', \text{ then the remaining two variables also equal to } u'\}. \end{aligned} \quad (20)$$

For example, $\{((1, 1), (0, 0), (0, 0), (1, 1), (0, 0), (0, 0)), ((1, 0), (1, 1), (0, 0), (1, 0), (1, 1), (0, 0)), ((1, 1), (0, 1), (0, 1), (1, 1), (1, 0), (1, 0))\} \subset \mathcal{K}_1(\mathbb{Z}_2^2)$ and $\{((1, 1), (0, 0), (0, 0), (1, 1), (0, 0), (0, 0)), ((1, 0), (1, 1), (0, 0), (1, 0), (1, 1), (0, 0))\} \subset \mathcal{K}_2(\mathbb{Z}_2^2)$, but $((1, 1), (0, 1), (0, 1), (1, 1), (1, 0), (1, 0)) \notin \mathcal{K}_2(\mathbb{Z}_2^2)$.

We substitute the above lemma to Eq. (17), then we prove that $\mathbb{E}((\widehat{O}_{0A}^{(r)\text{eq}})^2)$ is bounded by $\mathcal{O}(\text{tr}(O_0^2))$. In conclusion, the variance of our (R)ESPOVM estimator achieves similar scaling with one of the random Clifford measurements [9] and a similar sampling-copy complexity. Moreover, taking the average of $\mathbb{E}((\widehat{O}_{0A}^{(r)\text{eq}})^2)$ over uniformly random input and output states, we can show that the average of RESPOVM is exactly half of one of ESPOVM and even half of one of randomized Clifford tomography. We encapsulate the following results as a theorem whose a detailed derivation will be shown in Appendixes B and E.

Theorem 2. *Sampling-complexity efficiency of (R)ESPOVM shadow tomography.* Suppose that one is able to prepare multiple copies of an unknown input quantum state ρ , and that one is given $M \geq 1$ observables $O_1, O_2, \dots, O_M$ (for RESPOVM shadow tomography, all O_j s are real).

(i) With (R)ESPOVM shadow-tomography scheme, we can estimate each $\text{tr}(\rho O_j)$ to within an additive ε -error margin with a success probability $1 - \delta$ if we repeat this scheme over $N = \mathcal{O}(\frac{\max_{1 \leq j \leq M} \{\text{tr}(O_{j0}^2)\}}{\varepsilon^2} \log(\frac{2M}{\delta}))$ copies, where O_{j0} is the traceless part, $O_j - \frac{1}{2^n} \text{tr}(O_j)$.

(ii) If $2^n \gg 1$, the averaged sampling-copy number, to achieve the estimation within an additive ε -error margin and a success probability $1 - \delta$, over uniformly distributed complex pure states ρ and real pure states σ , approaches $\frac{136}{\varepsilon^2} \log(\frac{2M}{\delta})$ for (R)ESPOVM.

(iii) If $2^n \gg 1$, the averaged sampling-copy number to achieve the estimation within an additive ε -error margin and a success probability $1 - \delta$, averaged over uniformly distributed complex pure states ρ and σ , approaches $\frac{136}{\varepsilon^2} \log(\frac{2M}{\delta})$ for ESPOVM.

Theorem 2(i) implies that our scheme is efficient when the Frobenius norm of a given observable is at most

polynomial with the qubit number n . Most notably, if O is also a quantum state σ , then $\text{tr}(O_0^2)$ becomes constant, hence the required sampling-copy number N is independent of n . When the target state is real, Theorems 2(ii) and 2(iii) imply that the average sampling-copy bound of RESPOVM is half of one of ESPOVM. Furthermore, we know that random target state $|\psi\rangle$ is completely anticoncentrated [47]. That is, $\mathbb{E}_{|\psi\rangle} \sum_{x \in \mathbb{Z}_2^n} |\langle x | \psi \rangle|^4 = \mathcal{O}(2^{-n})$. Markov's inequality implies that with high probability, $|\langle x | \psi \rangle|^2 \leq \sqrt{\frac{\text{poly}(n)}{2^n}}$ for all $x \in \mathbb{Z}_2^n$, and then we can neglect the estimator of the second computational basis measurement part for large n , saving half of the copies. Hence, we can expect RESPOVM shadow tomography to be more useful for real observable estimation.

To showcase the performance of (R)ESPOVM shadow tomography, we examine scenarios in observable-expectation value estimation between arbitrary quantum states. Figures 2(a)–2(c) illustrate three exemplifying cases: ESPOVM shadow tomography for uniformly chosen random complex input and target states, ESPOVM shadow tomography for uniformly chosen complex input and real target states, and RESPOVM shadow tomography for the same types of random states as in the second case. We observe that within $N = 2000$, the averaged squared error over randomly chosen states becomes inversely proportional to N . These results are consistent with the points of Theorem 2. Moreover, the saturated sampling copy number to achieve $\overline{\varepsilon^2} \leq 0.001$ is much lower than the upper bound shown in Theorems 2(ii) and 2(iii). Furthermore, for real target states, the saturated value of N for RESPOVM is exactly half of that for ESPOVM. Figures 2(d) and 2(e) demonstrate the success of (R)ESPOVM shadow tomography in estimating the fidelity between 50-qubit noisily prepared GHZ states [48] and pure GHZ states with reasonable N values, where one also observes that RESPOVM needs significantly fewer sampling copies than ESPOVM to achieve the same accuracy. Figure 3 brings (R)ESPOVM shadow tomography to the test through estimation of fidelity between noisy 7×7 graph states [49] and its pure counterparts that has CZ-gate connections to all neighboring qubits. Here, N is also the number of measurement trials because every

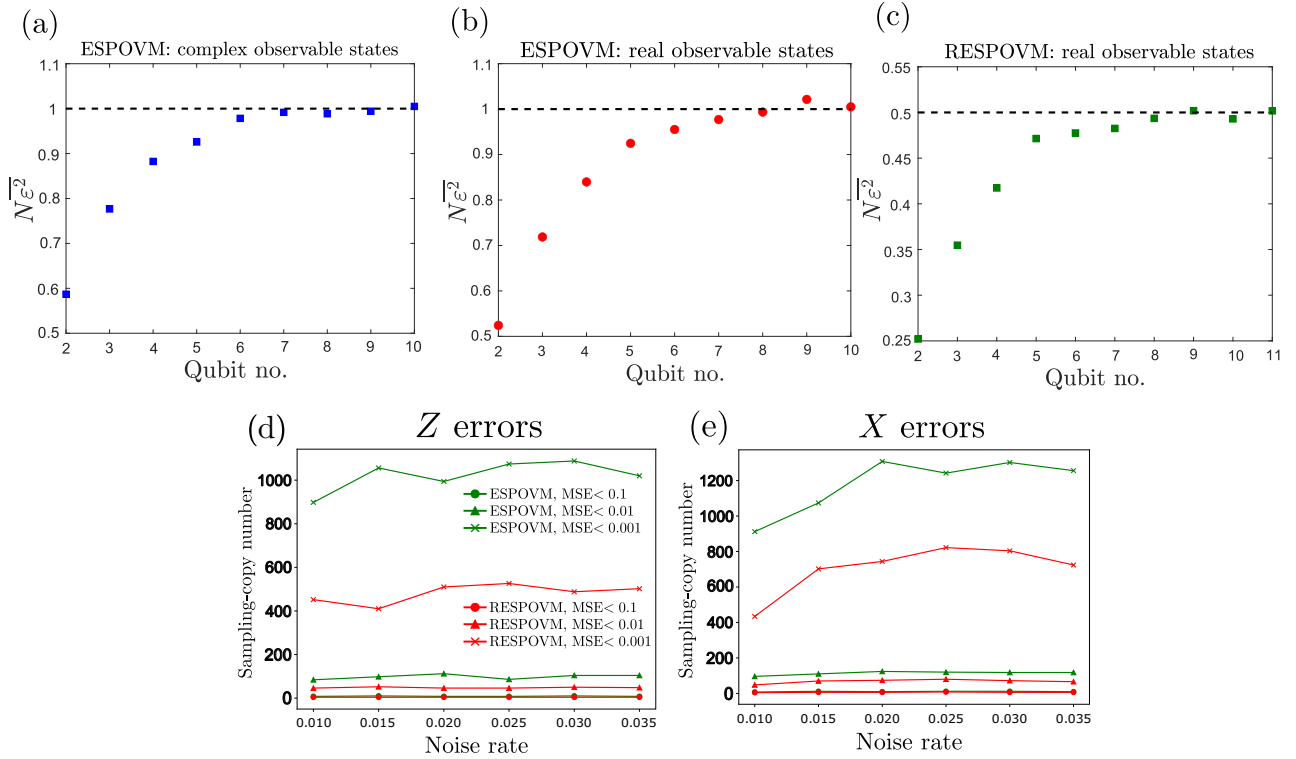


FIG. 2. (a)–(c) Sampling-copy numbers N multiplied with squared error $\overline{\epsilon^2}$ averaged on 200 experiments of different n -qubit (pure) input and target state observables. We set $N = 2000$ for (a), (b) and $N = 1000$ for (c). Also, for each experiment, we put 50 copies with given N to obtain the ϵ^2 , which is an averaged value of squared errors between the ideal expectation value of target and estimation, for one experiment. For (a), we averaged the ϵ^2 over uniform distributions of complex input and target states. For (b) and (c), the average is over uniform distributions of complex input and real target states. The horizontal dashed lines represent theoretically asymptotic values [41] of $N\epsilon^2$. (d), (e) Sampling-copy numbers N required to achieve the respective mean squared error thresholds (averaged over 400 experiments) for the fidelity estimation between a 50-qubit noisily-prepared GHZ state and pure GHZ state. We consider the i.i.d. (d) Z- and (e) X-error channels acting on every qubit with an error probability η_{prep} .

computational basis measurement outcome gives the same $\hat{O}_{p'}^{\text{bin}} \equiv \langle p' | O | p' \rangle = \frac{1}{2^n}$, so such a second measurement part may be omitted.

Alternatively, fidelity estimations of stabilizer states may be carried out using the simpler specialized quantum circuits reported in Ref. [28]. Nevertheless, we emphasize

that *all* target states achieve the same asymptotic sampling-complexity bound that is constant in n with (R)ESPOVM shadow tomography. This means that our technique clearly applies to many other interesting states, including the Dicke states [50], for which $N = \mathcal{O}(n^{2k})$ using the method in Ref. [28].

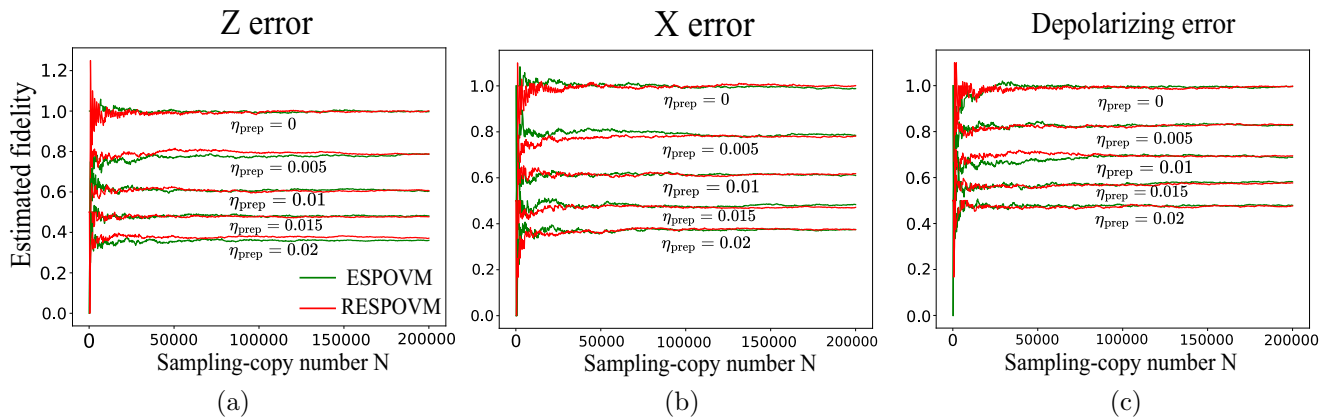


FIG. 3. The estimated fidelity with the sampling-copy number N in fidelity estimation between noisily prepared 7×7 -qubit graph states and their noiseless counterparts $|\psi\rangle$. The (a) Z, (b) X, and (c) depolarizing preparation-noise channels are i.i.d. to the individual qubits. Each estimated value of a given N is obtained from the median of all $\langle \psi | \hat{\psi} | \psi \rangle$ estimated using the (R)ESPOVM scheme over 100 experiments ($N/100$ copies for each experiment).

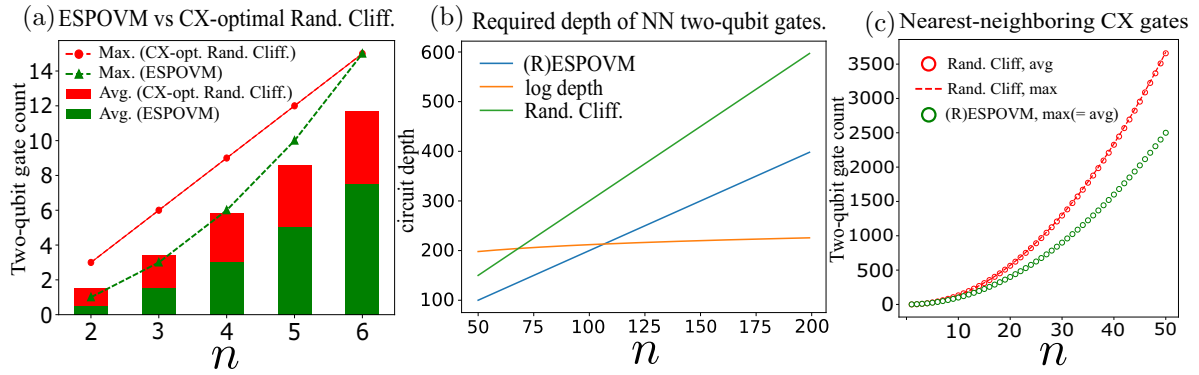


FIG. 4. (a) The average number of long-ranged two-qubit gates to implement (R)ESPOVM and randomized-Clifford circuits for low-qubit systems. The latter is obtained *via* Table 6 of Ref. [53] and is the average gate count over all equivalence classes. (b) The maximal required Clifford circuit depth for shadow tomography using neighboring gates. Using the knowledge of Ref. [18] (refer to Appendix C for a detailed elaboration), the maximal depth of random-Clifford tomography (blue) is $3n$. The orange line indicates the depth for *approximate* 3-design method [18,37,54] with some bias, $10 \times 2 \ln(\frac{9n}{\text{bias}})$. Here, we assign the allowed bias to 0.01. (c) The maximum and average numbers of NN CNOT gates (over 4000 gate samples) to implement a $2n$ -depth (R)ESPOVM and $3n$ -depth neighboring implementation of randomized-Clifford circuits [18].

V. RESOURCES FOR AND NOISE TOLERANCE OF (R)ESPOVM

Recalling Theorem 1, we note that (R)ESPOVM shadow tomography requires only CZ gates as two-qubit gate resources and hence it needs at most $\mathcal{O}(n^2)$ time. In the language of quantum circuits, Vizing's theorem [46,51,52] implies that an arbitrary CZ circuit requires a circuit depth of at most n when only CZ gates are employed. Whereas, the randomized Clifford tomography, given that HF section (see Sec. II B for its definition) can be regarded as classical postprocessing after the measurement, needs at most $2n$ depth.

Furthermore, with just NN two-qubit gates, CZ circuits with qubit reversal swapping need only $2n$ depth [36,41] of CNOT and S gates. Since we can regard qubit reversal as the reversal of measurement outcomes, (R)ESPOVM needs only $2n$ depth, which is lower than the worst-case circuit depth of $7n$ for the implementation of randomized-Clifford tomography [9,18,19,41]. In Appendix C, we further reduce the depth of randomized Clifford tomography to $3n$. This is possible since we are using only nonadaptive Clifford circuits where all CNOT gates have a control qubit on the top. These arguments are gathered into the following statements:

Theorem 3. Gate-count and depth efficiency of (R)ESPOVM shadow tomography [36]. Both ESPOVM and RESPOVM shadow tomography algorithms can be implemented by circuits of depth at most $2n$ using neighboring Clifford gates, while randomized Clifford needs at most $3n$ -depth neighboring Clifford gates. The implementation of linear nearest-neighboring (LNN) architecture takes $\mathcal{O}(n^2)$ -time for each sampling copy.

This fact, along with Theorems 1 and 2, argues that (R)ESPOVM algorithm is not a simple trade-off of Clifford circuit depth and sampling-copy number. Because scale-factor reduction of depth while preserving sampling copy scale gives significant improvement of gate noise threshold and error mitigation capability [55], it is also more compatible with noisy intermediate-scale quantum (NISQ) algorithms [56].

Figure 4 shows that our scheme significantly improved over conventional randomized-Clifford tomography regarding the required circuit depth and gate counts. The maximal gate count of our scheme matches with the random Clifford when $n = 6$. However, such CX-optimal random Clifford implementation is inefficient for larger n because a super-exponentially sized database of circuit templates is required [53]. We also compared with a recent result of Ref. [37] using a log-depth random Clifford circuit to *approximately* realize a unitary 3 design with some bias. This is because the unitary 3-design property is a sufficient condition for efficient shadow tomography with a similar scaling of the sampling complexity. Even though the depth scaling of our scheme is not logarithmic, it requires a larger depth than our scheme and even the random-Clifford scheme when n is of the order of tens. The reason is that Ref. [37] uses bilayers of alternating randomized-Clifford blocks, each of which has a qubit size occupying the whole circuit given that n is not large. In this case, the depth of the circuit is 10 times the size of a random Clifford block because the first layer needs 7 times [18,54] and the second needs 3 times ($\because$ Theorem 3) the size of a block [18].

The presence of multiple two-qubit gates in a circuit-based quantum-algorithm implementation, be it randomized-Clifford or (R)ESPOVM shadow tomography, enforces a restriction on the individual gate-noise threshold if one is to achieve a given total error rate of the output state the observable measurement is performed. In the previous section, we demonstrated that our (R)ESPOVM scheme requires a lower circuit depth than randomized Clifford sampling. Therefore, we expect that the required noise threshold for each two-qubit gate is affordably larger for (R)ESPOVMs.

Furthermore, our measurement scheme based on (R)ESPOVMs also leads to other interesting noise reduction effects. First, if the target state $\psi = \frac{1}{\sqrt{L}} \sum_{i=1}^L c_i |x_i\rangle$, then we can find the sufficient condition $\forall i \in [L], x_i \cdot \xi_p = 0 \pmod{2}$ such that the noise-corrupted output $p + \xi_p$, formed by adding the bit-flip error string ξ_p to the ideal output p of (R)ESPOVM, still gives the right estimator.

To show this, suppose we have a bit-flip error string ξ_p added to the desired outcome p . Then, from the above arguments, the sampled a_{ii} transforms to $a_{ii} + 2(p_i \oplus (\xi_p)_i)$, where $\oplus$ is the sum of modulo 2. Now, suppose the target state $|\psi\rangle = \frac{1}{\sqrt{L}} \sum_{j=1}^L c_j |\mathbf{x}_j\rangle$, where each $\mathbf{x}_j \in \mathbb{Z}_2^n$. Then $\hat{O}_A \equiv 2^n \langle \phi_A^{\text{eq}} | O | \phi_A^{\text{eq}} \rangle$ with noisy outcome $\mathbf{A}$ becomes

$$\hat{O}_A = \frac{1}{L} \left| \sum_{j=1}^L c_j \cdot i^{\mathbf{x}_j \cdot \tilde{\mathbf{a}} + \mathbf{x}_j^\top \mathbf{A} \mathbf{x}_j} \right|^2, \quad (21)$$

where $\tilde{\mathbf{a}}_i = 2(p_i \oplus (\xi_p)_i) - 2p_i = 2(\xi_p)_i \pmod{4}$. Hence, $i^{\mathbf{x}_j \cdot \tilde{\mathbf{a}}} = (-1)^{\mathbf{x}_j \cdot \xi_p}$. It means that if $\forall j \in [L]$, $\mathbf{x}_j \cdot \xi_p = 0$, then the estimator outputs the right value. The estimator can bear more bit-flip errors as the dimension of the spanning set of $\{\mathbf{x}_j\}_{j \in [L]}$ gets lower.

As typical examples, we suppose that the target state is the GHZ state $|\text{GHZ}_n\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})$ [38]. Then, the RESPOVM estimator reads

$$\hat{O} = \frac{1}{4} \left\{ 1 + (-1)^{\sum_{i=1}^n (a_i + p_i) + \sum_{i,j \in [n], i < j} a_{ij}} \right\}^2 + |\langle \text{GHZ}_n | \mathbf{p}' \rangle|^2 - \frac{1}{2}, \quad (22)$$

where $\mathbf{p}'$ is the outcome of computational basis measurement to the second copy, and a_{ij} are the elements of the sampled binary (upper triangular) matrix $\mathbf{A}$ (see Algorithm 1). In the GHZ case, even though the measurement outcome has an error, the estimator outputs the right value when the bit-flip error string ξ_p possesses an even Hamming-weight parity. In Fig. 5(b), we simulate the fidelity estimation of two pure GHZ states using noisy Clifford circuits. The closer the estimated fidelity is to 1, the more tolerant the estimation is to bit-flip noise. We see that our scheme performs much better than randomized-Clifford shadow tomography.

Second, if long-ranged CZ gates are available for our scheme, we can invoke the state-injection [57] technique to reduce the Hamming weight of the bit-flip error string ξ_p . To this end, with no loss of generality, we may suppose that the realized CZ gates are corrupted by Pauli noise, which can be obtained from twirling on an arbitrary noise channel that is not of this type [58,59]. By assuming that long-ranged gate errors are much more significant, one may set up a two-story teleportationlike circuit as shown in Fig. 5(a). It eventually has an equivalent CZ operation on the main board. Furthermore, the Pauli noise structure acting on the long-ranged CZ gate on the ancilla floor is turned into a dephasing noise. This noise transformation is carried out using neighboring CNOT gates entangling the qubits between the two floors, and postprocessing Z operations. This state injection is a linear channel and can be applied to all CZ gates. In effect, each CZ gate undergoes only dephasing noise (as a dominant noise), and such noise does not increase its weight by commuting with other CZ gates. Therefore, the weight of bit-flip errors at the final measurement part is significantly reduced.

This method will give additional noise tolerance to some symmetric states [23], given that the number of computational basis components undergoing large-weighted dephasing error is low, and hence so is the deviation of the estimation. A typical example is the W state $|\text{W}_n\rangle \equiv \frac{1}{\sqrt{n}}(|0\rangle^{\otimes(n-1)}|1\rangle + |0\rangle^{\otimes(n-2)}|1,0\rangle + \dots + |1\rangle|0\rangle^{\otimes(n-1)})$ [28] whose estimator

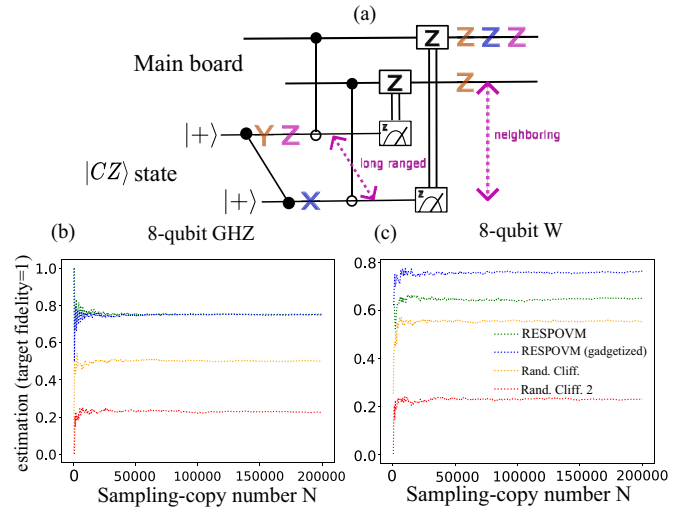


FIG. 5. (a) State-injection technique that transforms an arbitrary Pauli noise on the long-ranged gate into dephasing noise. (b), (c) CZ-based RESPOVM shadow-tomography performances in the estimation of fidelity between two pure eight-qubit GHZ (W states, respectively) states with an increasing number of input-state copies N . Here, we give the depolarizing noise, with 0.05 error rate, to each long-ranged gate [except the neighboring gates, which may also connect $(1, n)$ qubits]. The RESPOVM (gadgetized) means that for long-ranged CZ gates, we gave the dephasing error with 0.05 error rate [uniform (I, X, Y, Z) operation to each qubit, with a given error rate]. This setting is equivalent to the above state-injection technique with noiseless neighboring gates. Rand. Cliff. indicates the sampling and decomposition scheme via Ref. [19] in which noise is attributed only to the single HF section. This is because classically postprocessing of the measurement outcomes is sufficient to implement the other HF section. The Rand. Cliff. 2 method follows the sampling of Ref. [30].

reads

$$\hat{O} = \frac{\left\{ \sum_{i=1}^n (-1)^{(a_i + p_i)} \right\}^2}{2n} + |\langle \text{W}_n | \mathbf{p}' \rangle|^2 - \frac{1}{2}. \quad (23)$$

We can see the first term gets a larger deviation as the weight of ξ_p gets larger. In Fig. 5(c), we gave the initial depolarizing noise to the long-ranged CZ gates, where the dephasing error rate after the state-injection arrangement is equal to that of the initial noise rate on the ancillary CZ gate. We can see that the state-injection technique can give an additional improvement in the noise threshold to the original RESPOVM circuit.

The CZ-circuit-based architecture of our (R)ESPOVM schemes also presents an attractive feature for experimental implementation, especially with the soaring entangling-gate fidelities recently achieved in various quantum-computing platforms. While a more comprehensive survey shall be given in Appendix D, we quote some recent experimental fidelities achieved in some of these platforms: 99.5% two-qubit gate fidelity in Rydberg-atom quantum computing [60], over 99.5% for semiconductor-spin-based systems [61], and up to 99.93% with superconducting qubits [62].

VI. DISCUSSION

In this article, we propose resource-efficient methods to carry out quantum shadow tomography using equatorial-stabilizer measurements that are derived from smaller subsets of the full set of stabilizer states. We first prove that these equatorial-stabilizer measurements, together with the computational-basis measurements are indeed informationally complete and possess desirable sampling complexity that grows, in many cases, at most polynomially in the number of qubits n and reciprocally in the square of the target additive estimation error. In typical scenarios, such as rank-one observable estimation, the sampling complexity of equatorial-stabilizer shadow tomography is independent of n .

Furthermore, our shadow-tomography algorithms utilize long-ranged Clifford bases with circuit depths of n . When restricted to only LNN architectures, our algorithms are considerably efficient to implement, requiring only $2n$ -depth NN CNOT gates, owing to their simpler circuit geometry. We demonstrated that our methods require shorter circuit depth than randomized Clifford tomography for both long-ranged and LNN implementations, while still preserving the classes of applicable observables for many cases in terms of the shadow-norm or estimation variance criterion. Our depth- (or gate-) efficient shadow-tomography schemes become practically important for performing realistic quantum-computation tasks as they permit larger gate-error thresholds. Moreover, this is crucial for certain auxiliary quantum tasks, such as those in quantum error mitigation [63], for instance, where it became recently known [55] that such tasks would require sampling-copy numbers that increase exponentially with the circuit depth. We also found that along with the state injection technique, the estimators for (R)ESPOVM tomography reveal a unique tolerance over the various measurement bit-flip errors. Since state injection simplifies the CZ-gate-noise structure, postprocessing noise mitigation of shadow [64] could be applied.

Thus far, the (R)ESPOVM schemes we have found to have good shadow-tomographic properties require both computational-basis and CZ-circuit measurements that involve very different circuit depths, with the former being single layered while the latter grows linearly with n in depth. One possible future would be to investigate if there exist other such pairs of measurement circuits of comparable depths that can reduce the maximum circuit depth employed while maintaining the same estimation quality. Here, MUB subsets of the CZ-circuit measurements [33,34] could serve as an instructive starting point. Yet another desirable observation is the rather large error-threshold advantages of using equatorial-stabilizer shadow tomography on GHZ or W state-verification tasks in contrast to randomized-Clifford shadow tomography. This readily sets another research direction on fully identifying important classes of target quantum states that could lead to highly effective gate-error threshold improvements. Finally, motivated by a recent work [65], we expect that the shallow depth subset of noisy (R)ESPOVM would give another sample-optimal Pauli observable estimation method. Then we could measure the noise threshold of sampling copy improvement, which was previously considered on noisy random Clifford blocks [65], over the single-depth measurement.

Clifford circuits are well-known quantum circuits [66,67] whose noise-tolerant properties could be realizable in the near future. Therefore, we expect our work to establish practical research directions on gate-efficient observable-expectation-value estimations for noisy intermediate-scale quantum computation.

Note added. Recently, Ref. [37] appeared, discussing the possibility of approximating unitary designs with logarithmic-depth circuits. We have included an updated circuit-depth comparison between the algorithm in Ref. [37] and our (R)ESPOVM schemes. We show in Fig. 4(b) that despite the logarithmic scaling, our scheme employs relatively shorter circuit depths for moderate qubit numbers beyond 100.

ACKNOWLEDGMENTS

This work was supported by the Korean government [Ministry of Science and ICT (MSIT)], the NRF grants funded by the Korea government (MSIT) (No. RS-2023-00237959, No. RS-2024-00413957, No. RS-2024-00438415, No. RS-2025-02219034 and No. NRF-2023R1A2C1006115), the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (IITP-2025-RS-2020-II201606 and IITP-2025-RS-2024-00437191), and the Institute of Applied Physics at Seoul National University.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: ALGORITHM DETAILS FOR EQUATORIAL-STABILIZER SHADOW TOMOGRAPHY

Before we present the details of (R)ESPOVM shadow tomography, we first introduce the notation $\mathbb{E}_{\sigma \sim \mathcal{D}(A)} f(\sigma)$ that refers to the expected value of the function $f(\sigma)$ of a random variable $\sigma \in A$ that follows a uniformly random distribution $\mathcal{D}(A)$ with domain A . In particular, we call $\mathbb{E}_{\sigma \sim \mathcal{D}(A)} \sigma^{\otimes t}$ as the t th moment of A . With this, we have the following lemma [41].

Lemma A1. For the n -qubit system,

$$(i) [32] \mathbb{E}_{|\phi_A^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}| = \mathbb{E}_{|\phi_A^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_A^{\text{req}}\rangle \langle \phi_A^{\text{req}}| = \frac{I}{2^n}, \quad (A1)$$

$$(ii) [32] \mathbb{E}_{|\phi_A^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|^{\otimes 2} = \frac{1}{4^n} \left(I + \sum_{x,y \in \mathbb{Z}_2^n} |xy\rangle \langle yx| - \sum_{x \in \mathbb{Z}_2^n} |xx\rangle \langle xx| \right), \quad (A2)$$

$$(iii) \mathbb{E}_{|\phi_A^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_A^{\text{req}}\rangle \langle \phi_A^{\text{req}}|^{\otimes 2} = \frac{1}{4^n} \left(I + \sum_{x,y \in \mathbb{Z}_2^n} |xy\rangle \langle yx| + \sum_{x,y \in \mathbb{Z}_2^n} |xx\rangle \langle yy| - 2 \sum_{x \in \mathbb{Z}_2^n} |xx\rangle \langle xx| \right). \quad (A3)$$

Proof. The first and second moments of the equatorial stabilizer set are shown in Ref. [32]. So, we only consider the RESPOVM case and adopt similar strategies for ESPOVMs. All bold-faced italicized symbols refer to columns (synonymous with bit strings) and nonitalicized uppercase ones that are also bold faced refer to matrices. Note that for $\forall \mathbf{p} \in \mathbb{Z}_2^n$,

$$\begin{aligned} & Z^{\mathbf{p}} \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}| Z^{\mathbf{p}} \\ &= \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} \frac{1}{2^n} \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} (-1)^{\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{y}^T \mathbf{A} \mathbf{y} + \mathbf{p} \cdot (\mathbf{x} + \mathbf{y})} |\mathbf{x}\rangle \langle \mathbf{y}| \\ &= \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} \frac{1}{2^n} \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} (-1)^{\mathbf{x}^T \mathbf{A}' \mathbf{x} + \mathbf{y}^T \mathbf{A}' \mathbf{y}} |\mathbf{x}\rangle \langle \mathbf{y}| \\ &= \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_{\mathbf{A}'}\rangle \langle \phi_{\mathbf{A}'}| \\ &= \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_{\mathbf{A}}\rangle \langle \phi_{\mathbf{A}}|, \end{aligned} \quad (\text{A4})$$

where for $\forall i, j \in [n]$, $A'_{ii} = A_{ii} + p_i \pmod{2}$, otherwise $A'_{ij} = A_{ij}$. The last equality holds because the reparametrization of $\mathbf{A}$ does not affect its uniform-distribution first moment. Thus, $\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|$ commutes with all Pauli Z operators and is diagonal in the computational basis. Note that all diagonal elements are $\frac{1}{2^n}$. Therefore, $\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_{\mathbf{A}}\rangle \langle \phi_{\mathbf{A}}| = \frac{1}{2^n}$, which is Lemma A1(i).

Next, note that

$$\begin{aligned} & \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|^{\otimes 2} \\ &= \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} \frac{1}{4^n} \\ &\quad \times \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w} \in \mathbb{Z}_2^n} (-1)^{\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{y}^T \mathbf{A} \mathbf{y} + \mathbf{z}^T \mathbf{A} \mathbf{z} + \mathbf{w}^T \mathbf{A} \mathbf{w}} |\mathbf{x}\mathbf{y}\rangle \langle \mathbf{z}\mathbf{w}|, \\ &= \frac{1}{4^n} \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w} \in \mathbb{Z}_2^n} \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} \\ &\quad \times \{(-1)^{\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{y}^T \mathbf{A} \mathbf{y} + \mathbf{z}^T \mathbf{A} \mathbf{z} + \mathbf{w}^T \mathbf{A} \mathbf{w}}\} |\mathbf{x}\mathbf{y}\rangle \langle \mathbf{z}\mathbf{w}|. \end{aligned} \quad (\text{A5})$$

Now, we focus on terms in the curly brackets. Since the elements of $\mathbf{A}$ oscillate between 0 and 1, the nonzero coefficients of $|\mathbf{x}\mathbf{y}\rangle \langle \mathbf{z}\mathbf{w}|$ must satisfy the following equations for all $p, q \in [n] (p < q)$:

$$\begin{aligned} x_p + y_p + z_p &= w_p \\ x_p x_q + y_p y_q + z_p z_q + w_p w_q &= 0. \end{aligned} \quad (\text{A6})$$

Now, substitute the upper equation to lower one. We obtain that

$$x_p(y_q + z_q) + y_p(z_q + x_q) + z_p(x_q + y_q) = 0. \quad (\text{A7})$$

If $\mathbf{x} = \mathbf{y} = \mathbf{z}$, the above equation must hold. Suppose $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are not all the same. Then there exists $p' \in [n]$ such that two of $x_{p'}, y_{p'}, z_{p'}$ are equal but the rest are not. Also, $x_i = y_i = z_i$ for $i \leq p'$. Without loss of generality (WLOG), we may assume $x_{p'} = y_{p'} \neq z_{p'}$, then for $\forall q > p'$:

$$(x_q + y_q)(x_{p'} + z_{p'}) = 0 \rightarrow x_q = y_q \quad \because x_{p'} \neq z_{p'}. \quad (\text{A8})$$

It follows that nonzero coefficients of $|\mathbf{x}\mathbf{y}\rangle \langle \mathbf{z}\mathbf{w}|$ must be such that two of $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are equal. Based on this observation, (A6) implies several cases: (i) $\mathbf{x} = \mathbf{y}, \mathbf{z} = \mathbf{w}$; (ii)

$\mathbf{x} = \mathbf{z}, \mathbf{y} = \mathbf{w}$; and (iii) $\mathbf{x} = \mathbf{w}, \mathbf{y} = \mathbf{z}$. In all these cases, $\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} \{(-1)^{\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{y}^T \mathbf{A} \mathbf{y} + \mathbf{z}^T \mathbf{A} \mathbf{z} + \mathbf{w}^T \mathbf{A} \mathbf{w}}\} = 1$.

Hence, we obtain the following results:

$$\begin{aligned} & \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|^{\otimes 2} \\ &= \frac{1}{4^n} \left(I + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{x}| + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}| \right. \\ &\quad \left. - 2 \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}| \right), \end{aligned} \quad (\text{A9})$$

which is the result of Lemma A1(iii) and completes the proof. ■

We note that from Lemma A1(i), all elements from the respective sets $\{\frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} |\phi_{\mathbf{A}}^{\text{eq}}\rangle \langle \phi_{\mathbf{A}}^{\text{eq}}|\}$ and $\{\frac{2^n}{|\mathcal{S}_n^{\text{req}}|} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|\}$ form a POVM.

Using Lemmas A1(ii) and A1(iii), let us prove Eqs. (6) and (7), which are directly related to our tomography schemes.

First, we consider ESPOVMs. By Lemma A1:

$$\begin{aligned} & \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} (|\phi_{\mathbf{A}}^{\text{eq}}\rangle \langle \phi_{\mathbf{A}}^{\text{eq}}| \sigma |\phi_{\mathbf{A}}^{\text{eq}}\rangle \langle \phi_{\mathbf{A}}^{\text{eq}}| O |\phi_{\mathbf{A}}^{\text{eq}}\rangle) \\ &= \text{tr} \left\{ \left(\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_{\mathbf{A}}^{\text{eq}}\rangle \langle \phi_{\mathbf{A}}^{\text{eq}}|^{\otimes 2} \right) (\sigma \otimes O) \right\} \\ &= \frac{1}{4^n} \text{tr} \left\{ \left(I + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{x}| - \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}| \right) (\sigma \otimes O) \right\}, \\ &= \frac{1}{4^n} \left\{ \text{tr}(O) + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{y} \rangle \right. \\ &\quad \left. - \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle \right\}, \\ &= \frac{1}{4^n} \left\{ \text{tr}(O) + \text{tr}(O\sigma) - \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle \right\}. \end{aligned} \quad (\text{A10})$$

Hence, we obtain the following result:

$$\begin{aligned} \text{tr}(O\rho) &= \sum_{|\phi_{\mathbf{A}}^{\text{eq}}\rangle \in \mathcal{S}_n^{\text{eq}}} \frac{2^{2n}}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_{\mathbf{A}}^{\text{eq}} | \rho | \phi_{\mathbf{A}}^{\text{eq}} \rangle \langle \phi_{\mathbf{A}}^{\text{eq}} | O | \phi_{\mathbf{A}}^{\text{eq}} \rangle \\ &\quad + \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \rho | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle - \text{tr}(O). \end{aligned} \quad (\text{A11})$$

Next, we consider RESPOVMs and also assume O to be a real observable, that is, $\langle \mathbf{a} | O | \mathbf{b} \rangle \in \mathbb{R}$ for $\forall \mathbf{a}, \mathbf{b} \in \mathbb{Z}_2^n$. Then,

$$\begin{aligned} & \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} (|\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}| \sigma |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}| O |\phi_{\mathbf{A}}^{\text{req}}\rangle) \\ &= \text{tr} \left\{ \left(\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|^{\otimes 2} \right) (\sigma \otimes O) \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4^n} \text{tr} \left\{ \left(I + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{x}| + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}| \right. \right. \\
&\quad \left. \left. - 2 \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}| \right) (\sigma \otimes O) \right\} \\
&= \frac{1}{4^n} \left\{ \text{tr}(O) + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{y} \rangle \right. \\
&\quad \left. + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{x} \rangle \langle \mathbf{y} | O | \mathbf{x} \rangle - 2 \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle \right\} \\
&= \frac{1}{4^n} \left\{ \text{tr}(O) + 2 \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{y} \rangle \right. \\
&\quad \left. - 2 \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle \right\} \quad (\text{A12})
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4^n} \left\{ \text{tr}(O) + 2 \text{tr}(O\sigma) - 2 \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle \right\}. \quad (\text{A13})
\end{aligned}$$

This is equivalent to

$$\begin{aligned}
\text{tr}(O\rho) &= \sum_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \in \mathcal{S}_n^{\text{req}}} \frac{2^{2n-1}}{|\mathcal{S}_n^{\text{req}}|} \langle \phi_{\mathbf{A}}^{\text{req}} | \rho | \phi_{\mathbf{A}}^{\text{req}} \rangle \langle \phi_{\mathbf{A}}^{\text{req}} | O | \phi_{\mathbf{A}}^{\text{req}} \rangle \\
&\quad + \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \rho | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle - \frac{1}{2} \text{tr}(O). \quad (\text{A14})
\end{aligned}$$

Equation (A12) is valid when O is real. This proves Eqs. (6) and (7). In Eq. (13) of the main text, we ignore the $\text{tr}(O)$ term since it is constant and, hence, does not affect the variance.

Let us give some remarks before proceeding further. An arbitrary Hermitian input density operator ρ can be expressed as $\frac{1}{2^n} \sum_{P \in \mathcal{P}_n} \text{tr}(\rho P) P$. That is, setting O as Pauli operators, we note that ρ takes the equivalent form

$$\begin{aligned}
\rho &= \sum_{|\phi_{\mathbf{A}}^{\text{eq}}\rangle \in \mathcal{S}_n^{\text{eq}}} \frac{2^{2n}}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_{\mathbf{A}}^{\text{eq}} | \rho | \phi_{\mathbf{A}}^{\text{eq}} \rangle |\phi_{\mathbf{A}}^{\text{eq}}\rangle \langle \phi_{\mathbf{A}}^{\text{eq}}| \\
&\quad + \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \rho | \mathbf{x} \rangle |\mathbf{x}\rangle \langle \mathbf{x}| - I. \quad (\text{A15})
\end{aligned}$$

Using the POVM properties, $\sum_{\mathbf{A}} \frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_{\mathbf{A}}^{\text{eq}} | \rho | \phi_{\mathbf{A}}^{\text{eq}} \rangle = \sum_{\mathbf{x}} \langle \mathbf{x} | \rho | \mathbf{x} \rangle = 1$, we note that $\mathcal{M}^{-1} \circ \mathcal{M}(\rho) = \rho$ following the notations in Eqs. (10) and (11). Therefore, we conclude that Eq. (11) is indeed the inverse of $\mathcal{M}$. Since only the real part of ρ is expanded by only real Pauli basis operators, we can certify the RESPOVM inverse of Eq. (11) in the same manner.

From Lemma A1, Eq. (A11), and Eq. (A14), the following estimators (denoted with a caret) are constructed:

$$\widehat{\langle O \rangle} = \begin{cases} \frac{2^n}{(N/2)} \sum_{\substack{\text{sampled } \mathbf{A} \\ \text{over } N/2 \text{ copies}}} \langle \phi_{\mathbf{A}}^{\text{eq}} | O | \phi_{\mathbf{A}}^{\text{eq}} \rangle \\ + \frac{1}{(N/2)} \sum_{\substack{\text{measured } \mathbf{p}' \\ \text{over } N/2 \text{ copies}}} \langle \mathbf{p}' | O | \mathbf{p}' \rangle - \text{tr}(O) \\ \text{(ESPOVM and comp. basis),} \\ \frac{2^{n-1}}{(N/2)} \sum_{\substack{\text{sampled } \mathbf{A} \\ \text{over } N/2 \text{ copies}}} \langle \phi_{\mathbf{A}}^{\text{req}} | O | \phi_{\mathbf{A}}^{\text{req}} \rangle \\ + \frac{1}{(N/2)} \sum_{\substack{\text{measured } \mathbf{p}' \\ \text{over } N/2 \text{ copies}}} \langle \mathbf{p}' | O | \mathbf{p}' \rangle - \frac{1}{2} \text{tr}(O) \\ \text{(RESPOVM and comp. basis),} \end{cases} \quad (\text{A16})$$

where $\mathbf{A}$ and $\mathbf{p}'$ are sampled from the distribution of $\{\frac{2^n}{|\mathcal{S}_n^{(\text{r})\text{eq}}|} \langle \phi_{\mathbf{A}}^{(\text{r})\text{eq}} | \rho | \phi_{\mathbf{A}}^{(\text{r})\text{eq}} \rangle\}$ and $\{\langle \mathbf{p}' | \rho | \mathbf{p}' \rangle\}$, respectively. Note that these estimators are unbiased, i.e., data-average $\mathbb{E}(\widehat{\langle O \rangle}) = \langle O \rangle$.

We are ready to provide more details concerning the (R)ESPOVM shadow-tomography procedure, thereby explaining the validity of Algorithm 1. We shall only discuss matters related to the ESPOVM, as the dissertation for RESPOVM follows similarly. Let us assume an input state $\rho = \sum_{\mathbf{z}, \mathbf{w} \in \mathbb{Z}_2^n} |\mathbf{z}\rangle \langle \mathbf{w}| \rho | \mathbf{z}\rangle \langle \mathbf{w}|$. The procedure begins by a random selection of $|\phi_{\mathbf{A}}^{\text{eq}}\rangle \in \mathcal{S}_n^{\text{eq}}$, that is, the matrix $\mathbf{A}$ from uniformly distributed elements in the set

$$\{\mathbf{A} | \mathbf{A} = (a_{ij})_{i,j \in [n]}, a_{ij} = a_{ji}, a_{ii} \in \mathbb{Z}_4, a_{ij(i \neq j)} \in \mathbb{Z}_2\}. \quad (\text{A17})$$

Operationally, this is akin to flipping a coin $\frac{n^2-n}{2}$ times to determine all symmetric off-diagonal elements and flipping a uniform four-sided die n times to determine all diagonal elements of $\mathbf{A}$. Now, given a single-qubit (two-qubit, respectively) operation K , we define K_i (K_{ij}) as K acting on the i th (i th and j th) qubit. After choosing a random $\mathbf{A}$, we perform CZ_{ij} gate operations whenever $a_{ij} = 1$ and $i < j$, followed by single-qubit operations U_i such that

$$U_i \equiv \begin{cases} I_i & \text{if } a_{ii} = 0 \\ S^\dagger & \text{if } a_{ii} = 1 \\ Z_i & \text{if } a_{ii} = 2 \\ S_i & \text{if } a_{ii} = 3 \end{cases} \quad \text{or simply } U_i \equiv (S^\dagger)^{a_{ii}}. \quad (\text{A18})$$

Next, we apply Hadamard gates $H \equiv \bigotimes_{i=1}^n H_i$. The subsequent evolved state ρ' is written as

$$\rho' = \frac{1}{2^n} \sum_{\mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t} \in \mathbb{Z}_2^n} i^{3\mathbf{z}^\top \mathbf{A} \mathbf{z} + \mathbf{w}^\top \mathbf{A} \mathbf{w}} (-1)^{\mathbf{z} \cdot \mathbf{s} + \mathbf{w} \cdot \mathbf{t}} |\mathbf{s}\rangle \langle \mathbf{t}| \langle \mathbf{z} | \rho | \mathbf{w} \rangle. \quad (\text{A19})$$

Here, $\mathbf{a} \cdot \mathbf{b}$ refers to the binary inner product between $\mathbf{a}$ and $\mathbf{b}$. We shall now perform the Pauli Z basis measurement independently on all the qubits. The probability $\text{prob}(\mathbf{p} | \mathbf{A})$ to

obtain a specific binary string $\mathbf{p} \in \mathbb{Z}_2^n$ is

$$\begin{aligned}
 \text{prob}(\mathbf{p} | \mathbf{A}) &\equiv \langle \mathbf{p} | \rho' | \mathbf{p} \rangle \\
 &= \frac{1}{2^n} \sum_{\mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t} \in \mathbb{Z}_2^n} i^{3\mathbf{z}^\top \mathbf{A} \mathbf{z} + \mathbf{w}^\top \mathbf{A} \mathbf{w}} (-1)^{\mathbf{z} \cdot \mathbf{s} + \mathbf{w} \cdot \mathbf{t}} \\
 &\quad \times \langle \mathbf{p} | \mathbf{s} \rangle \langle \mathbf{t} | \mathbf{p} \rangle \langle \mathbf{z} | \sigma | \mathbf{w} \rangle, \\
 &= \frac{1}{2^n} \sum_{\mathbf{z}, \mathbf{w} \in \mathbb{Z}_2^n} i^{3\mathbf{z}^\top \mathbf{A} \mathbf{z} + \mathbf{w}^\top \mathbf{A} \mathbf{w}} (-1)^{\mathbf{z} \cdot \mathbf{p} + \mathbf{w} \cdot \mathbf{p}} \langle \mathbf{z} | \sigma | \mathbf{w} \rangle, \\
 &= \frac{1}{2^n} \sum_{\mathbf{z}, \mathbf{w} \in \mathbb{Z}_2^n} (-i)^{\mathbf{z}^\top \mathbf{A} \mathbf{z} - 2\mathbf{z} \cdot \mathbf{p}} i^{\mathbf{w}^\top \mathbf{A} \mathbf{w} - 2\mathbf{w} \cdot \mathbf{p}} \langle \mathbf{z} | \sigma | \mathbf{w} \rangle, \quad (\text{A20}) \\
 &= \langle \phi_{\mathbf{A}'}^{\text{eq}} | \rho | \phi_{\mathbf{A}'}^{\text{eq}} \rangle, \quad (\text{A21})
 \end{aligned}$$

where $\mathbf{A}'$ is an $n \times n$ matrix such that $a'_{ij} = a_{ij}$ for $i \neq j$, and $a'_{ii} = a_{ii} + 2p_i \pmod{4}$. The last equation can be obtained by following the definition of Eq. (3). Now we define $\mathbf{D}(\mathbf{p})$ as the $n \times n$ diagonal matrix whose i th diagonal element is equal to p_i , and also note, from Eq. (3), that

$$\begin{aligned}
 \frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_{\mathbf{A}'}^{\text{eq}} | \rho | \phi_{\mathbf{A}'}^{\text{eq}} \rangle &= \frac{1}{|\mathcal{S}_n^{\text{eq}}|} \sum_{\mathbf{z}, \mathbf{w} \in \mathbb{Z}_2^n} (-i)^{\mathbf{z}^\top \mathbf{A} \mathbf{z}} i^{\mathbf{w}^\top \mathbf{A} \mathbf{w}} \langle \mathbf{z} | \rho | \mathbf{w} \rangle \\
 &= \sum_{\mathbf{p} \in \mathbb{Z}_2^n} \frac{1}{|\mathcal{S}_n^{\text{eq}}|} \text{prob}(\mathbf{p} | \mathbf{A} - 2\mathbf{D}(\mathbf{p})), \quad (\text{A22})
 \end{aligned}$$

where the final equality is a result of Eq. (A20).

The left-hand side of Eq. (A22) corresponds to the probability distribution of $\mathbf{A}$ we want to sample from. Also, given a $\mathbf{p} \in \mathbb{Z}_2^n$, the probability to sample the matrix $\mathbf{A} - 2\mathbf{D}(\mathbf{p})$ is constant ($\frac{1}{|\mathcal{S}_n^{\text{eq}}|}$). Hence, we can interpret the rightmost side of Eq. (A22) as the probability to get the matrix $\mathbf{A}$ after we uniformly sampled the matrix $\mathbf{A}'$ and add $2\mathbf{D}(\mathbf{p})$ (see steps 21–25 in Algorithm 1). This way, the desired sample of $\mathbf{A}$ is obtained.

Furthermore, we note that one can propagate the measurement section to the leftover Hadamard gates and phase gates, becoming X - or Y -basis measurements. In that case, the CZ gates are the only remaining intermediate gates, and Vizing's theorem states that only at most n layers of CZ gates are needed to realize an arbitrary n -qubit CZ circuit, which settles part of Theorem 1. One must still carry out either the Pauli X or Y measurement depending on the phase gate acting on that qubit. However, the phase gate implementation is unnecessary and we may simply perform an X - or Y -basis measurement. The reason is that Z (S , respectively,) just flips the measurement outcome of the X (Y) measurement. For example, implementing Z and obtaining the X -measurement outcome 0 is equivalent to just measuring in the X basis, obtaining the outcome 1, which gives the same diagonal value of the sampled $\mathbf{A}$. This closes the proof of Theorem 1 for the ESPOVM scheme.

Harking back to Eq. (A11), to complete the shadow tomography protocol, one also acquires $\mathbf{p}' \in \mathbb{Z}_2^n$ following the distribution $\langle \mathbf{p}' | \rho | \mathbf{p}' \rangle$, which can be easily done by measuring ρ in the Pauli Z basis. Hence, the algorithm for an unbiased estimation of $\text{tr}(O\rho)$ can be summarized as follows:

(1) Measure one copy of ρ with the projector $|\phi_{\mathbf{A}'}^{\text{eq}}\rangle \langle \phi_{\mathbf{A}'}^{\text{eq}}|$ defined by a uniformly chosen $\mathbf{A}'$ from set $\mathcal{A}$ in (A17). This is done by (i) uniformly choosing an $\mathbf{A}' \in \mathcal{A}$, (ii) obtaining a $\mathbf{p} \in \mathbb{Z}_2^n$ from measuring this ρ copy with $|\phi_{\mathbf{A}'}^{\text{eq}}\rangle \langle \phi_{\mathbf{A}'}^{\text{eq}}|$, and (iii) and taking $\mathbf{A} = \mathbf{A}' + 2\mathbf{D}(\mathbf{p})$. The measurement outcome $\mathbf{A}$ would then follow the distribution $\frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} \langle \phi_{\mathbf{A}'}^{\text{eq}} | \rho | \phi_{\mathbf{A}'}^{\text{eq}} \rangle$.

(2) Measure one copy of ρ and obtain $\mathbf{p}' \in \mathbb{Z}_2^n$ following the distribution $\langle \mathbf{p}' | \rho | \mathbf{p}' \rangle$.

(3) Calculate $\widehat{O}_{\mathbf{A}, \mathbf{p}'} \equiv 2^n \langle \phi_{\mathbf{A}'}^{\text{eq}} | O | \phi_{\mathbf{A}'}^{\text{eq}} \rangle + \langle \mathbf{p}' | O | \mathbf{p}' \rangle - \text{tr}(O)$.

Algorithm 1 is the pseudocode encapsulation of all the arguments we made thus far.

APPENDIX B: ESTIMATION-VARIANCE ANALYSIS OF EQUATORIAL-STABILIZER SHADOW TOMOGRAPHY

We shall now sketch the proof of Theorem 2(i). We start from Eq. (13) and note that

$$\begin{aligned}
 \text{Var}(\widehat{O}_{\mathbf{A}}^{\text{eq}}) &= \mathbb{E}([2^n \langle \phi_{\mathbf{A}}^{\text{eq}} | O | \phi_{\mathbf{A}}^{\text{eq}} \rangle - \mathbb{E}(2^n \langle \phi_{\mathbf{A}}^{\text{eq}} | O | \phi_{\mathbf{A}}^{\text{eq}} \rangle)]^2) \\
 &= \mathbb{E}\left(\left[2^n \langle \phi_{\mathbf{A}}^{\text{eq}} | \left(O_0 + \frac{I}{2^n} \text{tr}(O)\right) | \phi_{\mathbf{A}}^{\text{eq}} \rangle - \mathbb{E}\left(2^n \langle \phi_{\mathbf{A}}^{\text{eq}} | \left(O_0 + \frac{I}{2^n} \text{tr}(O)\right) | \phi_{\mathbf{A}}^{\text{eq}} \rangle\right)\right]^2\right) \\
 &= \mathbb{E}([2^n \langle \phi_{\mathbf{A}}^{\text{eq}} | O_0 | \phi_{\mathbf{A}}^{\text{eq}} \rangle - \mathbb{E}(2^n \langle \phi_{\mathbf{A}}^{\text{eq}} | O_0 | \phi_{\mathbf{A}}^{\text{eq}} \rangle)]^2) \\
 &= \text{Var}(\widehat{O}_{0\mathbf{A}}^{\text{eq}}), \quad (\text{B1})
 \end{aligned}$$

where the second line is easily derived noting that O_0 is traceless part, $O - \frac{I}{2^n} \text{tr}(O)$, hence $\text{tr}(O_0) = 0$. Since we can follow the same reasoning to $\text{Var}(\widehat{O}_{\mathbf{A}}^{\text{req}})$ and $\text{Var}(\widehat{O}_{\mathbf{p}'}^{\text{bin}})$, we only need to consider the variance of the O_0 . In other words, we obtain that

$$\begin{aligned}
 \text{Var}^{(\text{r})\text{eq}}(\widehat{O}_{\mathbf{A}}^{(\text{r})\text{eq}}) &= \text{Var}^{(\text{r})\text{eq}}(\widehat{O}_{0\mathbf{A}}^{(\text{r})\text{eq}}) \\
 &= \mathbb{E}((\widehat{O}_{0\mathbf{A}}^{(\text{r})\text{eq}})^2) - \mathbb{E}(\widehat{O}_{0\mathbf{A}}^{(\text{r})\text{eq}})^2 \\
 &\leq \mathbb{E}((\widehat{O}_{0\mathbf{A}}^{(\text{r})\text{eq}})^2), \quad (\text{B2})
 \end{aligned}$$

where $\mathbb{E}$ is the expected value of $\widehat{O}_{0\mathbf{A}} \equiv 2^n \langle \phi_{\mathbf{A}}^{(\text{r})\text{eq}} | O_0 | \phi_{\mathbf{A}}^{(\text{r})\text{eq}} \rangle$. Hence, to obtain the upper bound of the estimation variance of $\widehat{O}_{0\mathbf{A}}^{(\text{r})\text{eq}}$, all we need to do is to bound $\mathbb{E}((\widehat{O}_{0\mathbf{A}}^{(\text{r})\text{eq}})^2)$ from above.

Equation (17) leads us to certify the third moment $\mathbb{E}_{|\phi_{\mathbf{A}}^{(\text{r})\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{(\text{r})\text{eq}})} |\phi_{\mathbf{A}}^{(\text{r})\text{eq}}\rangle \langle \phi_{\mathbf{A}}^{(\text{r})\text{eq}}|^{\otimes 3}$, and Lemma 1 is the answer. We assume the lemma holds (see Appendix E 1 for its proof). Then we define five more finite series, $\mathcal{C}^{(i)} \equiv \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}\mathbf{x}|$, and

$$\begin{aligned}
 \mathcal{C}^{(\text{ii})} &\equiv \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} (|\mathbf{y}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{x}\mathbf{x}| + |\mathbf{x}\mathbf{y}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{x}\mathbf{x}| + |\mathbf{x}\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{x}\mathbf{x}| \\
 &\quad + |\mathbf{y}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{y}\mathbf{x}| + |\mathbf{x}\mathbf{y}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{y}\mathbf{x}| + |\mathbf{x}\mathbf{x}\mathbf{y}\rangle \langle \mathbf{x}\mathbf{y}\mathbf{x}| \\
 &\quad + |\mathbf{y}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}\mathbf{y}| + |\mathbf{x}\mathbf{y}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}\mathbf{y}| + |\mathbf{x}\mathbf{x}\mathbf{y}\rangle \langle \mathbf{x}\mathbf{x}\mathbf{y}|), \quad (\text{B3})
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{C}^{(\text{iii})} &\equiv \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} (|\mathbf{y}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}\mathbf{y}| + |\mathbf{x}\mathbf{y}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}\mathbf{y}| + |\mathbf{x}\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{y}\mathbf{y}| \\
 &\quad + |\mathbf{y}\mathbf{y}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{x}\mathbf{x}| + |\mathbf{y}\mathbf{y}\mathbf{y}\rangle \langle \mathbf{x}\mathbf{y}\mathbf{x}| + |\mathbf{y}\mathbf{y}\mathbf{y}\rangle \langle \mathbf{x}\mathbf{x}\mathbf{y}|), \quad (\text{B4})
 \end{aligned}$$

$$\mathcal{C}^{(iv)} \equiv \sum_{x,y,z \in \mathbb{Z}_2^n} (|xyz\rangle \langle xyz| + |xyz\rangle \langle zxy| + |xyz\rangle \langle yzx| + |xyz\rangle \langle yxz| + |xyz\rangle \langle xzy| + |xyz\rangle \langle zyx|), \quad (\text{B5})$$

$$\mathcal{C}^{(v)} \equiv \sum_{x,y,z \in \mathbb{Z}_2^n} (|xxz\rangle \langle yyz| + |xxz\rangle \langle yzy| + |xxz\rangle \langle zyy| + |xxz\rangle \langle yzy| + |xxz\rangle \langle yzy| + |xxz\rangle \langle zyy| + |xxz\rangle \langle yzy| + |xxz\rangle \langle zyy|). \quad (\text{B6})$$

Following Lemma 1, the result for ESPOVMs can be rewritten as

$$\begin{aligned} & \mathbb{E}_{|\phi_A^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|^{\otimes 3} \\ &= \frac{1}{8^n} \sum_{(x,y,z,w,s,t) \in \mathcal{K}_2(\mathbb{Z}_2^n)} |xyz\rangle \langle wst| \\ &= \frac{1}{8^n} \{ \mathcal{C}^{(i)} + (\mathcal{C}^{(ii)} - 9\mathcal{C}^{(i)}) \\ &+ [\mathcal{C}^{(iv)} - 6\mathcal{C}^{(i)} - 2(\mathcal{C}^{(ii)} - 9\mathcal{C}^{(i)})] \} \\ &= \frac{1}{8^n} (4\mathcal{C}^{(i)} - \mathcal{C}^{(ii)} + \mathcal{C}^{(iv)}). \end{aligned} \quad (\text{B7})$$

Also, for RESPOVMs:

$$\begin{aligned} & \mathbb{E}_{|\phi_A^{\text{req}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{req}})} |\phi_A^{\text{req}}\rangle \langle \phi_A^{\text{req}}|^{\otimes 3} \\ &= \frac{1}{8^n} \sum_{(x,y,z,w,s,t) \in \mathcal{K}_1(\mathbb{Z}_2^n)} |xyz\rangle \langle wst| \\ &= \frac{1}{8^n} \{ \mathcal{C}^{(i)} + (\mathcal{C}^{(ii)} - 9\mathcal{C}^{(i)}) + (\mathcal{C}^{(iii)} - 6\mathcal{C}^{(i)}) \\ &+ [\mathcal{C}^{(iv)} - 6\mathcal{C}^{(i)} - 2(\mathcal{C}^{(ii)} - 9\mathcal{C}^{(i)})] \\ &+ [\mathcal{C}^{(v)} - 9\mathcal{C}^{(i)} - 3(\mathcal{C}^{(iii)} - 6\mathcal{C}^{(i)}) - (\mathcal{C}^{(ii)} - 9\mathcal{C}^{(i)})] \} \\ &= \frac{1}{8^n} (16\mathcal{C}^{(i)} - 2\mathcal{C}^{(ii)} - 2\mathcal{C}^{(iii)} + \mathcal{C}^{(iv)} + \mathcal{C}^{(v)}). \end{aligned} \quad (\text{B8})$$

Upon recalling Eq. (17),

$$\begin{aligned} & \mathbb{E}((\widehat{\mathcal{O}}_{0A}^{\text{eq}})^2) \\ &= 2^{3n} \text{tr}(\mathbb{E}_{|\phi_A^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|^{\otimes 3} (\rho \otimes O_0 \otimes O_0)). \end{aligned} \quad (\text{B9})$$

For ESPOVMs, by Eq. (B7), we can rewrite this as

$$\mathbb{E}((\widehat{\mathcal{O}}_{0A}^{\text{eq}})^2) = \text{tr}((4\mathcal{C}^{(i)} - \mathcal{C}^{(ii)} + \mathcal{C}^{(iv)})(\rho \otimes O_0 \otimes O_0)). \quad (\text{B10})$$

Similarly, we can show that from Eqs. (17) and (B8),

$$\begin{aligned} & \mathbb{E}((\widehat{\mathcal{O}}_{0A}^{\text{req}})^2) \\ &= \frac{1}{4} \text{tr}((16\mathcal{C}^{(i)} - 2\mathcal{C}^{(ii)} - 2\mathcal{C}^{(iii)} + \mathcal{C}^{(iv)} + \mathcal{C}^{(v)})(\rho \otimes O_0^{\otimes 2})). \end{aligned} \quad (\text{B11})$$

We are now ready to provide an upper bound for $\mathbb{E}((\widehat{\mathcal{O}}_{0A}^{\text{req}})^2)$ by utilizing:

Lemma B1. Given the above definitions, for arbitrary state ρ and Hermitian operator O_0 :

$$|\text{tr}(\mathcal{C}^{(i)}(\rho \otimes O_0 \otimes O_0))| \leq \text{tr}(O_0^2), \quad (\text{B12})$$

$$|\text{tr}(\mathcal{C}^{(ii)}(\rho \otimes O_0 \otimes O_0))| \leq 7\text{tr}(O_0^2), \quad (\text{B13})$$

$$|\text{tr}(\mathcal{C}^{(iii)}(\rho \otimes O_0 \otimes O_0))| \leq 6\text{tr}(O_0^2), \quad (\text{B14})$$

$$|\text{tr}(\mathcal{C}^{(iv)}(\rho \otimes O_0 \otimes O_0))| \leq 3\text{tr}(O_0^2), \quad (\text{B15})$$

$$|\text{tr}(\mathcal{C}^{(v)}(\rho \otimes O_0 \otimes O_0))| \leq 7\text{tr}(O_0^2). \quad (\text{B16})$$

See Appendix E 2 for its proof. This result implies that

$$\begin{aligned} \mathbb{E}((\widehat{\mathcal{O}}_{0A}^{\text{eq}})^2) &\leq 4|\text{tr}(\mathcal{C}^{(i)}(\rho \otimes O_0 \otimes O_0))| \\ &+ |\text{tr}(\mathcal{C}^{(ii)}(\rho \otimes O_0 \otimes O_0))| \\ &+ |\text{tr}(\mathcal{C}^{(iv)}(\rho \otimes O_0 \otimes O_0))| \leq 14\text{tr}(O_0^2). \end{aligned} \quad (\text{B17})$$

Similarly, we can show that $E_\rho((\widehat{\mathcal{O}}_{0A}^{\text{req}})^2) \leq 13\text{tr}(O_0^2)$.

Furthermore, for $\mathbb{E}((\widehat{\mathcal{O}}_{0A}^{\text{bin}})^2)$,

$$\begin{aligned} \mathbb{E}((\widehat{\mathcal{O}}_{0A}^{\text{bin}})^2) &= \sum_{x \in \mathbb{Z}_2^n} \langle x | \rho | x \rangle \langle x | O_0 | x \rangle \langle x | O_0 | x \rangle \\ &\leq \sum_{x,y \in \mathbb{Z}_2^n} \langle x | \rho | x \rangle \langle x | O_0 | y \rangle \langle y | O_0 | x \rangle \\ &\leq \text{tr}(O_0^2). \end{aligned} \quad (\text{B18})$$

With this, and the arguments leading to Eqs. (B1) and (13), Theorem 2(i) is proven. The upper bound of the sampling-copy number, that is, $\frac{(68 \times 2) \text{Var}_\rho^{(\text{req})}(\widehat{\mathcal{O}}_{0A,j})}{\varepsilon^2} \log(\frac{2M}{\delta})$ for M observables, follows from the statistical benefits of using the median-of-means estimation method [9,42–44].

Suppose that the single observable O of interest is a quantum state σ . Then $\text{tr}(O_0^2) = \text{tr}((\sigma - \frac{I}{2^n})^2) = \text{tr}(\sigma^2) - \frac{\text{tr}(\sigma)}{2^{n-1}} + \frac{1}{2^n} \leq \text{tr}(\sigma^2) \leq 1$. Hence, any quantum-state observable σ possesses a constant estimation variance and demonstrates shadow-tomographic advantage with (R)ESPOVM, provided that the trace inner product between σ and equatorial stabilizer state is computable efficiently.

The proofs of Theorems 2(ii) and 2(iii) only need an average of estimation variance over uniformly chosen input and target states. The result is shown below.

Lemma B2. (i) If $2^n \gg 1$, the averaged value $\overline{\text{Var}^{(\text{req})}(\widehat{\sigma})}^{\rho,\sigma}$ over uniformly distributed complex pure states ρ and real pure states σ approaches 1 for eq and $\frac{1}{2}$ for req. In randomized-Clifford tomography, this average variance approaches 1.

(ii) If $2^n \gg 1$, the averaged value $\overline{\text{Var}^{\text{eq}}(\widehat{\sigma})}^{\rho,\sigma}$ over uniformly distributed complex pure states ρ and σ approaches 1. In randomized-Clifford tomography, this average variance approaches 1.

Assisted by previous arguments, the proof of the above lemma naturally leads to Theorems 2(ii) and 2(iii). However, it contains more technical elements that are deferred to Appendix E 3.

Furthermore, we can see that the average estimation variance of RESPOVM is exactly half of one of ESPOVM and

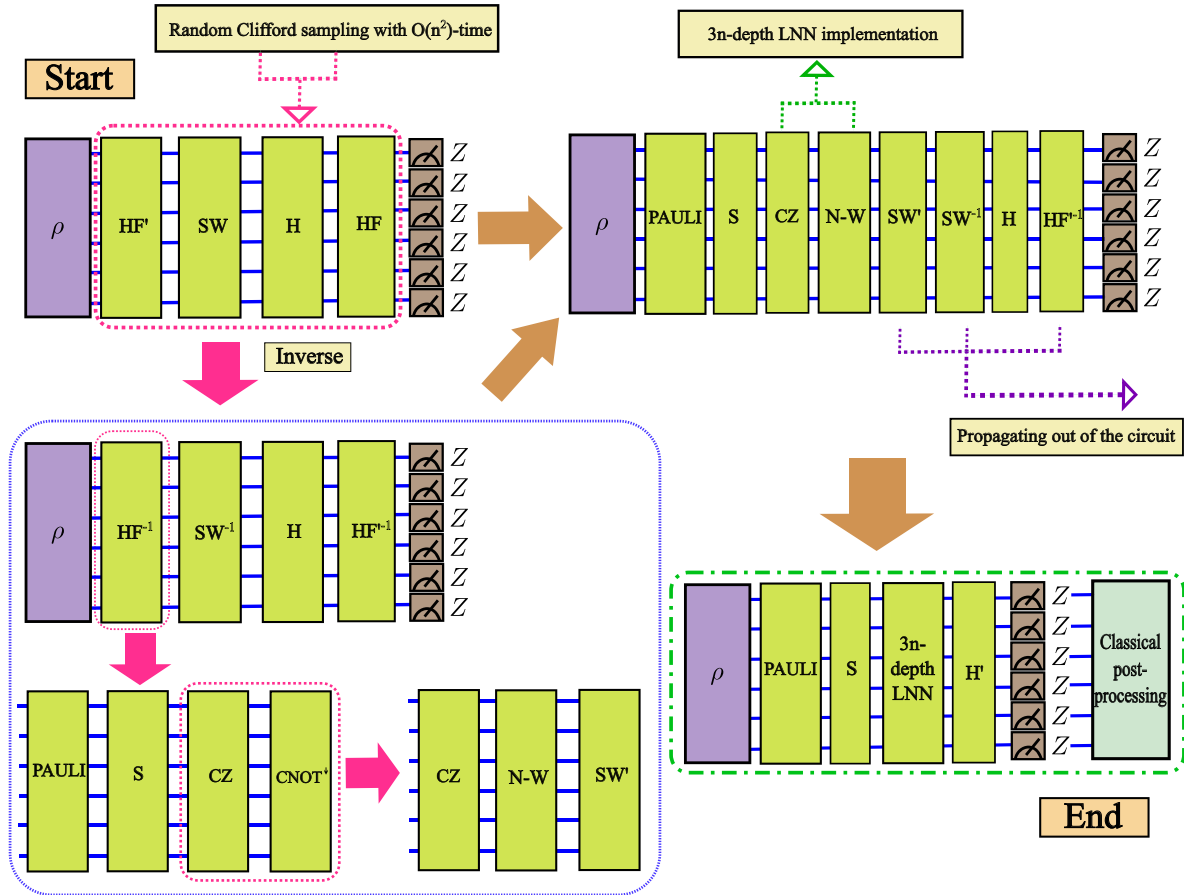


FIG. 6. The schematic diagram of uniform sampling of a $3n$ -depth Clifford-circuit measurement in the LNN architecture. Here, the N-W section is a CNOT circuit expressing the $n \times n$ northwestern matrix [18,70] for the binary strings. The final classical postprocessing step multiplies an $n \times n$ binary matrix to the measured outcome bit string, which takes $\mathcal{O}(n^2)$ time.

even one of randomized Clifford tomography. Even though RESPOVM requires two copies of input states for a single sampling trial, the total number of gates used with the same number of sampling copies is, on average, half of what is required in randomized-Clifford shadow tomography.

APPENDIX C: GATE, DEPTH, AND TIME COMPLEXITY FOR CIRCUIT IMPLEMENTATION

We now discuss the architectural requirements of (R)ESPOVM shadow tomography, which cover the main ideas of the proof for Theorem 3. In particular, this discussion focuses on the implementation of an (LNN) architecture for (R)ESPOVM tomography and compares its circuit-depth and time-complexity resources with those of randomized-Clifford shadow tomography.

We recall [19] that we can efficiently and uniformly sample a Clifford unitary U from Cl_n in $\mathcal{O}(n^2)$ time. In addition, the sampled form contains the $\text{HF}'\text{---SW---H---HF}$ sections. Each small upper case letter is explained in Sec. II B. Here, we can further reduce the depth of circuit lower than previously known results [18,46]. We note that the $\text{HF}'\text{---SW}$ sections are not needed for randomized-Clifford shadow tomography. The reason is that the measurement outcome of a uniform-Clifford measurement can be obtained from the following procedure:

One first uniformly samples a Clifford unitary operator in the above form and takes its inverse, which possesses the form $\text{HF}^{-1}\text{---H---SW}^{-1}\text{---HF}^{-1}$. Then one just measures in the Z basis after the $\text{HF}^{-1}\text{---H}$ sections and linearly transforms this intermediate bit-string outcome with the subsequent CNOT and Pauli transformations in $\text{SW}^{-1}\text{---HF}^{-1}$ to arrive at the final bitstring outcome. In $\mathcal{O}(n^3)$ time, one can optimize the maximal depth of $\text{HF}^{-1}\text{---H}$ sections with long-range gates to $n + \mathcal{O}(\log(n))$ if $n \geq 70$ [46]. Even though the corresponding asymptotic gate-count upper bound is $\mathcal{O}(\frac{n}{\log(n)})$ [68,69], it has a large constant factor [46] and hence will not improve the former limit for circuits holding hundreds of qubits. If only NN gates are available, we can express the CNOT gate layers inside the HF section in terms of a northwestern matrix [70] and SWAP gates. Here, we do not need to implement the swapping operations, since one can simply propagate all swapping operations out of the circuit. The CNOT circuit of a northwestern-matrix form after the CZ circuits can be further decomposed into $3n$ -depth NN CNOT circuits (sorting network) and phase-gate insertion layers [18,70] in $\mathcal{O}(n^3)$ -time. A schematic illustration is presented in Fig. 6.

As we saw in Theorem 1, there is an alternative depth reduction technique for CZ circuits. (R)ESPOVM shadow tomography can be implemented with at most n -depth long-ranged CZ gates (or with CNOT and H gates). This maximal

depth can be reduced to $\frac{n}{2} + \mathcal{O}(\log(n))$ given that $n \geq 39$ according to Ref. [46], which is half the gate complexity in contrast to randomized-Clifford implementation. Furthermore, with only neighboring CNOT gates and S gates, we can implement (R)ESPOVM with $2n$ -depth LNN architecture in $\mathcal{O}(n^2)$ time. To show this time-complexity bound, we may use the strategy employed in Ref. [36]. We first construct a $2n$ -depth circuit consisting of NN CNOT gates in advance following the designated rules dictated by Theorem 6 in Ref. [36]. With this, the main task for the time-complexity-bound proof is as follows: Given a CZ circuit that performs the map $|x\rangle \mapsto (-1)^{Q(x)} |x\rangle$, where $x \in \mathbb{Z}_2^n$ and Q is a quadratic binary-valued polynomial, find the values $u_j, u_{j,k}$ in $\mathbb{Z}_4$ ($j \in [n], k \in \{j+1, j+2, \dots, n\}$) such that

$$(-1)^{Q(x)} = i^{\sum_{j=1}^n u_j y_j + \sum_{j=1}^n \sum_{k=j+1}^n u_{j,k} (y_j \oplus y_k)}. \quad (\text{C1})$$

Here, $\oplus$ and $+$, respectively, refer to modulo-2 and modulo-4 additions. The y_i 's ($i \in [n]$) are defined by

$$y_i = x_1 \oplus x_2 \oplus x_3 \oplus \dots \oplus x_i. \quad (\text{C2})$$

After inserting the phase gates $S^{u_{j,k}}$ and S^{u_i} following the manual in Ref. [36], we finally realize the LNN architecture of the CZ circuit. To analyze the time complexity, we rewrite Eq. (C2) as

$$\begin{aligned} x_1 &= y_1 \\ x_i &= y_i \oplus y_{i-1} \quad (i \in \{2, 3, \dots, n\}). \end{aligned} \quad (\text{C3})$$

Now consider one off-diagonal component of $(-1)^{Q(x)}$. We have simple identity $(-1)^{x_\mu x_\nu} = i^{3x_\mu + 3x_\nu + (x_\mu \oplus x_\nu)}$. Then, by Eq. (C3),

$$(-1)^{x_\mu x_\nu} = i^{3(y_\mu \oplus y_{\mu-1}) + 3(y_\nu \oplus y_{\nu-1}) + (y_\mu \oplus y_{\mu-1} \oplus y_\nu \oplus y_{\nu-1})}, \quad (\text{C4})$$

with $y_0 \equiv 0$. Using the equality [36]

$$i^{a \oplus b \oplus c} = i^{3a + 3b + 3c + 3(a \oplus b) + 3(b \oplus c)} \quad (a, b, c \in \mathbb{Z}_2), \quad (\text{C5})$$

and setting $a = y_\mu, b = y_{\mu-1}, c = y_\nu \oplus y_{\nu-1}$, we can express $i^{y_\mu \oplus y_{\mu-1} \oplus y_\nu \oplus y_{\nu-1}}$ as a product of $i^{y_\mu}, i^{y_{\mu-1}}, i^{y_\nu \oplus y_{\nu-1}}$ and $i^{y_\mu \oplus y_{\mu-1}}, i^{y_\mu \oplus y_\nu \oplus y_{\nu-1}}, i^{y_{\mu-1} \oplus y_\nu \oplus y_{\nu-1}}$. Again, in a similar manner, we can decompose $i^{y_\mu \oplus y_\nu \oplus y_{\nu-1}}$ and $i^{y_{\mu-1} \oplus y_\nu \oplus y_{\nu-1}}$ into a product of powers of i , each with an exponent not exceeding three variables. In conclusion, there exist constant-sized sets $A, B \subset [n]$ such that

$$(-1)^{x_\mu x_\nu} = i^{\sum_{a \in A} u'_a y_a + \sum_{a, b \in B} u'_{a,b} (y_a \oplus y_b)}, \quad (\text{C6})$$

where $u'_\mu, u'_{\mu,v} \in \mathbb{Z}_4$. Therefore, upon noting that a CZ circuit has at most $\frac{n^2-n}{2}$ CZ gates, we have the following scheme for implementing any (long-ranged) CZ circuit in an LNN architecture:

(1) Prepare a $2n$ -depth NN CNOT circuit based on the prescriptions in Ref. [36].

(2) For each unitary $\text{CZ}_{j,k}$ acting on the j th and k th qubits, find the sets $A_{j,k}$ and $B_{j,k}$ such that $(-1)^{x_a x_b} = i^{\sum_{a \in A_{j,k}} u'_a y_a + \sum_{a, b \in B_{j,k}} u'_{a,b} (y_a \oplus y_b)}$. Record all u'_a and $u'_{a,b}$ corresponding to $A_{j,k}$ and $B_{j,k}$.

(3) Insert the phase gates $S^{u'_a} (a \in A_{j,k})$ and $S^{u'_{a,b}} (a, b \in B_{j,k})$ according to the insertion rules in Ref. [36].

(4) Repeat steps 2 and 3 for all $j, k \in [n]$. If there is already some phase gate acting on the location, we multiply the existing one and the gate we put this time.

The reader may consult Ref. [41] for the precise NN CNOT-circuit prescription and phase-gate insertion rule. For each $j, k \in [n]$, we note that steps 2–4 each takes constant time to execute. Since there is a total of $\mathcal{O}(n^2)$ steps to be repeated, the total time complexity is $\mathcal{O}(n^2)$.

APPENDIX D: SURVEY OF RECENT CZ-GATE IMPLEMENTATION SCHEMES

We shall list all recent fidelity values sorted according to some of these platforms. In photonic quantum computing, a fidelity of over 81% has been reported on silicon-nitride photonic chips [71], and over 99% when Zeeman coupling with quantum dots and nitrogen-vacancy centers is exploited in simulations [72]. Spin-based quantum computation allows for more than 99.5% fidelity [61, 73]. On Rydberg-interacting platforms, Refs. [74, 75] and later [60], respectively, reported gate fidelities of 89%, 97%, and 99.5% using atomic-qubit-array entanglement. Using adiabatic pulses, simulations revealed fidelity values greater than 99.9% [76]. On superconducting platforms, adiabaticity-shortcut techniques allow for over 96% fidelity [77], 97.6% with fluxonium qubits [78], and 99.1% through interfering superconducting qubits with weak anharmonicity [79]. Nonadiabatic CZ gates were also realized with 99.54% fidelity [80]. Furthermore, gmon qubits employed by Google AI Quantum demonstrated gate fidelities as high as 99.81% [81]. The more recent fidelity of 99.93% is possible with intermediate leakage control [62].

APPENDIX E: PROOFS OF OTHER LEMMAS

1. Proof of Lemma 1

We start with S_n^{req} . Similar to Eq. (A5),

$$\begin{aligned} & \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(S_n^{\text{req}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|^{\otimes 3} \\ &= \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(S_n^{\text{req}})} \frac{1}{8^n} \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{y}\mathbf{z}\rangle \langle \mathbf{w}\mathbf{s}\mathbf{t}| \\ & \quad \times (-1)^{\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{y}^\top \mathbf{A}\mathbf{y} + \mathbf{z}^\top \mathbf{A}\mathbf{z} + \mathbf{w}^\top \mathbf{A}\mathbf{w} + \mathbf{s}^\top \mathbf{A}\mathbf{s} + \mathbf{t}^\top \mathbf{A}\mathbf{t}} \\ &= \frac{1}{8^n} \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{y}\mathbf{z}\rangle \langle \mathbf{w}\mathbf{s}\mathbf{t}| \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(S_n^{\text{req}})} \\ & \quad \times \{(-1)^{\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{y}^\top \mathbf{A}\mathbf{y} + \mathbf{z}^\top \mathbf{A}\mathbf{z} + \mathbf{w}^\top \mathbf{A}\mathbf{w} + \mathbf{s}^\top \mathbf{A}\mathbf{s} + \mathbf{t}^\top \mathbf{A}\mathbf{t}}\}. \end{aligned} \quad (\text{E1})$$

By the same token, for nonzero coefficients in the summand, the following equations must hold for all $p, q \in [n]$:

$$\begin{aligned} x_p + y_p + z_p + w_p + s_p &= t_p \pmod{2} \\ x_p x_q + y_p y_q + z_p z_q + w_p w_q + s_p s_q + t_p t_q &= 0 \pmod{2}. \end{aligned} \quad (\text{E2})$$

Substituting the upper equation into the lower one, we obtain

$$\begin{aligned} 0 &= x_p(y_q + z_q + w_q + s_q) + y_p(z_q + x_q + w_q + s_q) \\ & \quad + z_p(x_q + y_q + w_q + s_q) + w_p(x_q + y_q + z_q + s_q) \\ & \quad + s_p(x_q + y_q + z_q + w_q). \end{aligned} \quad (\text{E3})$$

If $\mathbf{x} = \mathbf{y} = \mathbf{z} = \mathbf{w} = \mathbf{s}$, the above equation clearly holds. Otherwise, we have two cases:

(i) There exists $p' \in [N]$ such that two of $x_{p'}, y_{p'}, z_{p'}, w_{p'}, s_{p'}$ are 0 (1, respectively) and the rest are 1 (0), and $x_i = y_i = z_i = w_i = s_i$ for $i < p'$.

(ii) There exists $p' \in [N]$ such that four of $x_{p'}, y_{p'}, z_{p'}, w_{p'}, s_{p'}$ are 0 (1, respectively) and the other is 1 (0), and $x_i = y_i = z_i = w_i = s_i$ for $i < p'$.

First, suppose case (i). WLOG, when $x_{p'} = y_{p'} (*)$, then $z_{p'} = w_{p'} = s_{p'} = x_{p'} + 1$, which together with Eq. (E3) implies that

$$x_p y_q + x_q y_p + (x_{p'} + 1)(x_q + y_q) = x_q + y_q = 0 \pmod{2} \quad (\text{E4})$$

as $x_{p'} = y_{p'}$. So, we obtain $\mathbf{x} = \mathbf{y}$. Substituting this back to Eq. (E2),

$$\begin{aligned} z_p + w_p + s_p &= t_p \pmod{2} \\ z_p z_q + w_p w_q + s_p s_q + t_p t_q &= 0 \pmod{2}. \end{aligned} \quad (\text{E5})$$

This is exactly what we have obtained during the calculation of the second moment. Hence, we conclude the same statement: two of $\mathbf{z}, \mathbf{w}, \mathbf{s}$ are equal. By the first equation of Eq. (E5), the remaining one equals $\mathbf{t}$. Therefore, we conclude that each two of $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t}$ must be equal. Note that even if we chose two other variables, not $\mathbf{x}$ and $\mathbf{y}$ in assumption(*), we arrive at the same conclusion.

Next, let us suppose case (ii). By the first equation of Eq. (E2), $x_i = y_i = z_i = w_i = s_i = t_i$ for $i < p'$. WLOG, assume that $x_{p'} = y_{p'} = z_{p'} = w_{p'} \neq s_{p'}$. Then, we obtain $s_{p'} = t_{p'}$ from Eq. (E2). Next, we rewrite the problem as

$$\begin{aligned} y_p + z_p + w_p + s_p + t_p &= x_p \pmod{2} \\ x_p x_q + y_p y_q + z_p z_q + w_p w_q + s_p s_q + t_p t_q &= 0 \pmod{2}. \end{aligned} \quad (\text{E6})$$

Similarly, substituting the first equation into the second,

$$\begin{aligned} 0 &= t_p(y_q + z_q + w_q + s_q) + y_p(z_q + x_q + w_q + s_q) \\ &\quad + z_p(t_q + y_q + w_q + s_q) + w_p(t_q + y_q + z_q + s_q) \\ &\quad + s_p(t_q + y_q + z_q + w_q). \end{aligned} \quad (\text{E7})$$

Remember we assumed that $t_{p'} = s_{p'}$, $t_{p'} \neq w_{p'}$, $y_{p'} = z_{p'} = w_{p'}$, and $y_i = z_i = w_i = s_i = t_i$ for $i < p'$. This is exactly the case of solving under assumption (i). So, we conclude the same statement that $\mathbf{t} = \mathbf{s}$, two of $\mathbf{y}, \mathbf{z}, \mathbf{w}$ are equal, and hence the remaining one equals $\mathbf{x}$.

Throughout cases (i) and (ii), we finally conclude that the coefficient of $|\mathbf{xyz}\rangle\langle\mathbf{wst}|$ is nonzero if and only if each triplet in the variable set $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t}\}$ must be equal and in all such cases, the coefficients $\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})}\{(-1)^{\mathbf{x}^\top \mathbf{A} \mathbf{x} + \mathbf{y}^\top \mathbf{A} \mathbf{y} + \mathbf{z}^\top \mathbf{A} \mathbf{z} + \mathbf{w}^\top \mathbf{A} \mathbf{w} + \mathbf{s}^\top \mathbf{A} \mathbf{s} + \mathbf{t}^\top \mathbf{A} \mathbf{t}}\}$ all equal 1. The proof for RESPOVM is now completed.

Now, we move on to $\mathcal{S}_n^{\text{eq}}$, whence

$$\begin{aligned} &\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_{\mathbf{A}}^{\text{eq}}\rangle\langle\phi_{\mathbf{A}}^{\text{eq}}|^{\otimes 3} \\ &= \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} \frac{1}{8^n} \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t} \in \mathbb{Z}_2^n} |\mathbf{xyz}\rangle\langle\mathbf{wst}| \\ &\quad \times i^{\mathbf{x}^\top \mathbf{A} \mathbf{x} + \mathbf{y}^\top \mathbf{A} \mathbf{y} + \mathbf{z}^\top \mathbf{A} \mathbf{z} - \mathbf{w}^\top \mathbf{A} \mathbf{w} - \mathbf{s}^\top \mathbf{A} \mathbf{s} - \mathbf{t}^\top \mathbf{A} \mathbf{t}} \\ &= \frac{1}{8^n} \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t} \in \mathbb{Z}_2^n} |\mathbf{xyz}\rangle\langle\mathbf{wst}| \mathbb{E}_{|\phi_{\mathbf{A}}^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} \\ &\quad \times \{i^{\mathbf{x}^\top \mathbf{A} \mathbf{x} + \mathbf{y}^\top \mathbf{A} \mathbf{y} + \mathbf{z}^\top \mathbf{A} \mathbf{z} - \mathbf{w}^\top \mathbf{A} \mathbf{w} - \mathbf{s}^\top \mathbf{A} \mathbf{s} - \mathbf{t}^\top \mathbf{A} \mathbf{t}}\}. \end{aligned} \quad (\text{E8})$$

By the same argument, for the coefficient not to be 0, the following equation must hold. For all $p, q \in \mathbb{Z}_2^n$ ($p < q$),

$$\begin{aligned} x_p + y_p + z_p &= w_p + s_p + t_p \pmod{4} \\ x_p x_q + y_p y_q + z_p z_q + w_p w_q + s_p s_q + t_p t_q &= 0 \pmod{2}. \end{aligned} \quad (\text{E9})$$

We add $w_p + s_p$ for both sides of the first equation of (E9), and we obtain that

$$\begin{aligned} x_p + y_p + z_p + w_p + s_p &= t_p \pmod{2} \\ x_p x_q + y_p y_q + z_p z_q + w_p w_q + s_p s_q + t_p t_q &= 0 \pmod{2}. \end{aligned} \quad (\text{E10})$$

This is the equation for the $\mathcal{S}_n^{\text{req}}$ -case. Hence we obtain the same statement, each 3-couples in variable set $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t}\}$ must be equal. However, in this case, we have a stricter condition, the first equation of Eq. (E9). WLOG, we only consider two cases.

(i) $\mathbf{x} = \mathbf{y} : 2x_p + z_p = w_p + s_p + t_p \pmod{4} \rightarrow z_p = w_p + s_p + t_p \pmod{2}$. Hence two variables in $\{\mathbf{z}, \mathbf{w}, \mathbf{s}, \mathbf{t}\}$ are equal each other.

(ii) $\mathbf{x} = \mathbf{w} : y_p + z_p = s_p + t_p \pmod{4}$. If $\mathbf{y} = \mathbf{z}$, $2y_p = s_p + t_p \rightarrow \mathbf{y} = \mathbf{z} = \mathbf{s} = \mathbf{t}$. If $\mathbf{y} = \mathbf{s}$, $\mathbf{z} = \mathbf{t}$. If $\mathbf{y} = \mathbf{w}$, all cases are equivalent to previous cases. Therefore, the statement that coefficients are nonzero is equivalent to that if $|\mathbf{xyz}\rangle\langle\mathbf{wst}|$ has the two same variables where each do not locate in the same position (bra and ket) and the other two variables in bra (or ket) are equal, remaining variables in ket (or bra) must be equal to the previous two variables. Finally, in those cases, the coefficient cannot be other values but 1.

Other cases can be managed similarly and are known to have similar effects of (i) or (ii). We thus proved Lemma 1.

2. Proof of Lemma B1

First, we note that

$$\sum_{\mathbf{x} \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{xxx}\rangle\langle\mathbf{xxx}| (\sigma \otimes O_0 \otimes O_0)) = \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle\mathbf{x}|\sigma|\mathbf{x}\rangle \langle\mathbf{x}|O_0|\mathbf{x}\rangle \langle\mathbf{x}|O_0|\mathbf{x}\rangle \leq \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} \langle\mathbf{x}|\sigma|\mathbf{x}\rangle \langle\mathbf{x}|O_0|\mathbf{y}\rangle \langle\mathbf{y}|O_0|\mathbf{x}\rangle \leq \text{tr}(O_0^2). \quad (\text{E11})$$

Here, we used the fact that $\langle\mathbf{x}|O_0|\mathbf{y}\rangle \langle\mathbf{y}|O_0|\mathbf{x}\rangle \geq 0$ for $\forall \mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n$. This proves Eq. (B12).

Also, we easily note that

$$\begin{aligned}
& \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|xyz\rangle\langle xyz| + |xyz\rangle\langle xzy| + |xyz\rangle\langle yxz| + |xyz\rangle\langle zxy| + |xyz\rangle\langle zyx| + |xyz\rangle\langle yzx|)(\sigma \otimes O_0 \otimes O_0) \\
&= (\text{tr}(O_0))^2 + \text{tr}(O_0^2) + 2\text{tr}(O_0)\text{tr}(\sigma O_0) + 2\text{tr}(\sigma O_0^2) \\
&\leq (\text{tr}(O_0))^2 + \text{tr}(O_0^2) + 2\text{tr}(O_0)\|O_0\|_\infty + 2\|O_0^2\|_\infty \\
&\leq (\text{tr}(O_0))^2 + 3\text{tr}(O_0^2) + 2\text{tr}(O_0)\|O_0\|_\infty \\
&\leq 3\text{tr}(O_0^2),
\end{aligned} \tag{E12}$$

where $\|O\|_\infty \equiv \max_{|\psi\rangle} \{|\langle \psi | O | \psi \rangle|\}$. Furthermore, we can consider another case when two variables are the same in ket. Let $\sigma = \sum_\lambda \lambda |\lambda\rangle\langle \lambda|$ be spectral decomposition of σ , hence $\forall \lambda \geq 0$ and $\sum_\lambda \lambda = 1$. Then we obtain the following list of results:

$$\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|xxy\rangle\langle zzy| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle z | \sigma | x \rangle \langle z | O_0 | x \rangle \langle y | O_0 | y \rangle \right| = |\text{tr}(O_0)\text{tr}(\sigma O_0^\top)| = 0; \tag{E13}$$

$$\begin{aligned}
\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|xxy\rangle\langle zyz| (\sigma \otimes O_0 \otimes O_0)) \right| &= \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle z | \sigma | x \rangle \langle y | O_0 | x \rangle \langle z | O_0 | y \rangle \right| = |\text{tr}(\sigma(O_0^\top)^2)| \leq \|(O_0^\top)^2\|_\infty \\
&\leq \text{tr}((O_0^\top)^2) = \text{tr}(O_0^2);
\end{aligned} \tag{E14}$$

$$\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|xxy\rangle\langle yzz| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle y | \sigma | x \rangle \langle z | O_0 | x \rangle \langle z | O_0 | y \rangle \right| = |\text{tr}(\sigma O_0^\top O_0)| \leq \|O_0^\top O_0\|_\infty; \tag{E15}$$

$$\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|yxx\rangle\langle zzy| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle z | \sigma | y \rangle \langle z | O_0 | x \rangle \langle y | O_0 | x \rangle \right| = |\text{tr}(\sigma O_0 O_0^\top)| \leq \|O_0 O_0^\top\|_\infty; \tag{E16}$$

$$\begin{aligned}
\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|yxx\rangle\langle zyz| (\sigma \otimes O_0 \otimes O_0)) \right| &= \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle z | \sigma | x \rangle \langle y | O_0 | x \rangle \langle z | O_0 | y \rangle \right| = \text{tr}(\sigma(O_0^\top)^2) \leq \|(O_0^\top)^2\|_\infty \\
&\leq \text{tr}((O_0^\top)^2) = \text{tr}(O_0^2);
\end{aligned} \tag{E17}$$

$$\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|yxx\rangle\langle yzz| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle y | \sigma | y \rangle \langle z | O_0 | x \rangle \langle z | O_0 | x \rangle \right| = |\text{tr}(O_0 O_0^\top)|; \tag{E18}$$

$$\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|xyx\rangle\langle zzy| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle z | \sigma | x \rangle \langle z | O_0 | y \rangle \langle y | O_0 | x \rangle \right| \leq \text{tr}(O_0^2); \tag{E19}$$

$$\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|xyx\rangle\langle zyz| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle z | \sigma | x \rangle \langle y | O_0 | y \rangle \langle z | O_0 | x \rangle \right| \leq (\text{tr}(O_0))^2 = 0; \tag{E20}$$

$$\left| \sum_{x,y,z \in \mathbb{Z}_2^n} \text{tr}(|xyx\rangle\langle yzz| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y,z \in \mathbb{Z}_2^n} \langle y | \sigma | x \rangle \langle z | O_0 | y \rangle \langle z | O_0 | x \rangle \right| \leq \|O_0 O_0^\top\|_\infty; \tag{E21}$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|xxy\rangle\langle xyx| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle x | \sigma | x \rangle \langle y | O_0 | x \rangle \langle x | O_0 | y \rangle \right| \leq \sum_{x,y \in \mathbb{Z}_2^n} \langle y | O_0 | x \rangle \langle x | O_0 | y \rangle = \text{tr}(O_0^2) \tag{E22}$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|xxy\rangle\langle xyx| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle x | \sigma | x \rangle \langle y | O_0 | x \rangle \langle x | O_0 | y \rangle \right| \leq \sum_{x,y \in \mathbb{Z}_2^n} \langle y | O_0 | x \rangle \langle x | O_0 | y \rangle = \text{tr}(O_0^2); \tag{E23}$$

$$\begin{aligned}
\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|xxy\rangle\langle xxy| (\sigma \otimes O_0 \otimes O_0)) \right| &= \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle x | \sigma | x \rangle \langle x | O_0 | x \rangle \langle y | O_0 | y \rangle \right| \leq \sum_{x \in \mathbb{Z}_2^n} |\langle x | \sigma | x \rangle \langle x | O_0 | x \rangle \text{tr}(O_0)| \\
&= |\text{tr}(O_0)| \|O_0\|_\infty = 0;
\end{aligned} \tag{E24}$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{x}\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{x}\mathbf{x}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma^\top | \mathbf{x} \rangle \langle \mathbf{y} | O_0 | \mathbf{y} \rangle \langle \mathbf{y} | O_0 | \mathbf{x} \rangle \right| \leq \text{tr}(O_0^2) \quad [\because \text{ see Eq. (E34) below}]; \quad (\text{E25})$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{y}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}\mathbf{y}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{y} \rangle \langle \mathbf{y} | O_0 | \mathbf{x} \rangle \langle \mathbf{x} | O_0^\top | \mathbf{y} \rangle \right| \leq \sum_{y \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{y} \rangle |\langle \mathbf{y} | O_0 O_0^\top | \mathbf{y} \rangle| \leq \|O_0 O_0^\top\|_\infty; \quad (\text{E26})$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{x}\mathbf{y}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}\mathbf{y}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{x} \rangle \langle \mathbf{y} | O_0 | \mathbf{y} \rangle \langle \mathbf{y} | O_0 | \mathbf{x} \rangle \right| \leq \text{tr}(O_0^2) \quad [\because \text{ Eq. (E34)}]; \quad (\text{E27})$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{y}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}\mathbf{y}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{y} \rangle \langle \mathbf{x} | O_0 | \mathbf{x} \rangle \langle \mathbf{y} | O_0 | \mathbf{x} \rangle \right| \leq \text{tr}(O_0^2) \quad [\because \text{ Eq. (E34)}]; \quad (\text{E28})$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{y}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{x}\mathbf{x}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{y} \rangle \langle \mathbf{x} | O_0 | \mathbf{x} \rangle \langle \mathbf{x} | O_0 | \mathbf{x} \rangle \right| \leq \sum_{x,z \in \mathbb{Z}_2^n} \langle \mathbf{x} | O_0 | \mathbf{z} \rangle \langle \mathbf{z} | O_0 | \mathbf{x} \rangle \leq \text{tr}(O_0^2); \quad (\text{E29})$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{y}\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{y}\mathbf{x}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{y} \rangle \langle \mathbf{y} | O_0 | \mathbf{x} \rangle \langle \mathbf{x} | O_0 | \mathbf{x} \rangle \right| \leq \text{tr}(O_0^2) \quad [\because \text{ Eq. (E34)}]; \quad (\text{E30})$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{x}\mathbf{y}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}\mathbf{y}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O_0 | \mathbf{y} \rangle \langle \mathbf{y} | O_0 | \mathbf{x} \rangle \right| \leq \text{tr}(O_0^2); \quad (\text{E31})$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{x}\mathbf{y}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{x}\mathbf{x}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{x} | \sigma | \mathbf{x} \rangle \langle \mathbf{y} | O_0 | \mathbf{y} \rangle \langle \mathbf{x} | O_0 | \mathbf{x} \rangle \right| \leq (\text{tr}(O_0))^2 = 0; \quad (\text{E32})$$

$$\left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{x}\mathbf{y}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{x}\mathbf{x}| (\sigma \otimes O_0 \otimes O_0)) \right| = \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{x} \rangle \langle \mathbf{x} | O_0 | \mathbf{y} \rangle \langle \mathbf{x} | O_0 | \mathbf{x} \rangle \right| \leq \text{tr}(O_0^2) \quad [\because \text{ Eq. (E34)}]. \quad (\text{E33})$$

$$\begin{aligned} \left| \sum_{x,y \in \mathbb{Z}_2^n} \text{tr}(|\mathbf{x}\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{y}\mathbf{y}| (\sigma \otimes O_0 \otimes O_0)) \right| &= \left| \sum_{x,y \in \mathbb{Z}_2^n} \langle \mathbf{y} | \sigma | \mathbf{x} \rangle \langle \mathbf{y} | O_0 | \mathbf{x} \rangle \langle \mathbf{y} | O_0 | \mathbf{y} \rangle \right| \leq \sum_{y \in \mathbb{Z}_2^n} |\langle \mathbf{y} | \sigma O_0^\top | \mathbf{y} \rangle| |\langle \mathbf{y} | O_0 | \mathbf{y} \rangle| \\ &\leq \sqrt{\sum_{y \in \mathbb{Z}_2^n} |\langle \mathbf{y} | \sigma O_0^\top | \mathbf{y} \rangle|^2 \sum_{z \in \mathbb{Z}_2^n} \langle \mathbf{z} | O_0 | \mathbf{z} \rangle \langle \mathbf{z} | O_0 | \mathbf{z} \rangle} \\ &\leq \sqrt{\sum_{y \in \mathbb{Z}_2^n} \left| \sum_{\lambda} \lambda \langle \mathbf{y} | \lambda \rangle \langle \lambda | O_0^\top | \mathbf{y} \rangle \right|^2 \sum_{z,k \in \mathbb{Z}_2^n} \langle \mathbf{z} | O_0 | \mathbf{k} \rangle \langle \mathbf{k} | O_0 | \mathbf{z} \rangle} \\ &\leq \sqrt{\sum_{y \in \mathbb{Z}_2^n} |\langle \lambda_y | O_0^\top | \mathbf{y} \rangle|^2 |\langle \mathbf{y} | \lambda_y \rangle|^2 \text{tr}(O_0^2) (\lambda_y \equiv \text{argmax}_{\lambda} \{ |\langle \lambda | O_0^\top | \mathbf{y} \rangle \langle \mathbf{y} | \lambda \rangle| \})} \\ &\leq \sqrt{\sum_{y \in \mathbb{Z}_2^n} \sum_{\lambda} |\langle \mathbf{y} | \lambda \rangle|^2 \langle \lambda | O_0^\top | \mathbf{y} \rangle \langle \mathbf{y} | O_0^\top | \lambda \rangle \text{tr}(O_0^2)} \\ &\leq \sqrt{\sum_{\lambda} \langle \lambda | O_0^\top | \mathbf{y}_\lambda \rangle \langle \mathbf{y}_\lambda | O_0^\top | \lambda \rangle \text{tr}(O_0^2) (\mathbf{y}_\lambda \equiv \text{argmax}_{y \in \mathbb{Z}_2^n} \{ |\langle \lambda | O_0^\top | \mathbf{y} \rangle|^2 \})} \\ &\leq \sqrt{\sum_{\lambda} \sum_{z \in \mathbb{Z}_2^n} \langle \lambda | O_0^\top | \mathbf{z} \rangle \langle \mathbf{z} | O_0^\top | \lambda \rangle \text{tr}(O_0^2)} = \sqrt{\sum_{\lambda} \langle \lambda | (O_0^\top)^2 | \lambda \rangle \text{tr}(O_0^2)} \\ &= \sqrt{\text{tr}((O_0^\top)^2)} \sqrt{\text{tr}(O_0^2)} = \text{tr}(O_0^2). \end{aligned} \quad (\text{E34})$$

Throughout the above equations, we used the fact that $\text{tr}(O_0) = 0$. Also, Eqs. (E15), (E16), (E18), (E21), and (E26) do not contribute to $\mathcal{C}^{(i,ii,iv)}$, and are thus of no consequence to the ESPOVM analysis, for which we can assume O_0 is real and $O_0^\top = O_0$. So, we can easily note that these can also be upper bounded by $\text{tr}(O_0^2)$. Note that values of other cases can be calculated from the above-known cases, by reparametrizing or transpose-conjugating both sides, for example. After realizing that other cases are 0 or upper bounded by $\text{tr}(O_0^2)$, based on simple trigonometric inequalities, the remaining equations of Lemma B1 other than Eq. (B12) may also be proven.

3. Proof of Lemma B2

If the averaged variance $\overline{\text{Var}}^{(r)\text{eq}}(\hat{\sigma}_{\mathbf{A},p'})^{\rho,\sigma}$ over the uniformly chosen input and target state is proved to be 1 (1/2, respectively) for (R)ESPOVM when $2^n \gg 1$, by the result of Refs. [9,42–44], Theorems 2(ii) and 2(iii) will be proved.

Moreover, we recall that the mean squared error (MSE) is defined by the sample mean value of the squares of the additive error between estimation and target value. If an estimation variance is $\text{Var}(\hat{\sigma})$, then for a large sampling-copy number N , the averaged MSE $\overline{\epsilon^2}$ over many experiments of different input and target state becomes $\frac{2\overline{\text{Var}}^{(r)\text{eq}}(\hat{\sigma}_{\mathbf{A},p'})^{\rho,\sigma}}{N}$ (Note that sampling trial number is $N/2$, that is why we have the factor 2). Therefore, the asymptotic lines of Figs. 2(a)–2(c) are explained.

Now, let us assume that $2^n \gg 1$. We take the RESPOVM first. Recall Eq. (A14):

$$\begin{aligned} & \sum_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \in S_n^{\text{req}}} \frac{2^{2n-1}}{|S_n^{\text{req}}|} \langle \phi_{\mathbf{A}}^{\text{req}} | \rho | \phi_{\mathbf{A}}^{\text{req}} \rangle \langle \phi_{\mathbf{A}}^{\text{req}} | O | \phi_{\mathbf{A}}^{\text{req}} \rangle \\ &= \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \rho | \mathbf{x} \rangle \langle \mathbf{x} | O | \mathbf{x} \rangle + \frac{1}{2} \text{tr}(O) + \text{tr}(O\rho). \end{aligned} \quad (\text{E35})$$

Then we obtain that

$$\begin{aligned} \mathbb{E}(\hat{\sigma}_{\mathbf{A}}^{\text{req}})^2 &= \left(\text{tr}(\rho\sigma) + \frac{1}{2} \text{tr}(\sigma) - \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \langle \mathbf{x} | \rho | \mathbf{x} \rangle \langle \mathbf{x} | \sigma | \mathbf{x} \rangle \right)^2 \\ &= \frac{1}{4} + \text{tr}((\sigma \otimes \sigma)(\rho \otimes \rho)) + \text{tr}(\sigma\rho) \\ &\quad + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} \text{tr}((\rho^{\otimes 2} \sigma^{\otimes 2})(|\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{x}|^{\otimes 2})) \\ &\quad - 2\text{tr}(\sigma\rho) \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \text{tr}((\rho \otimes \sigma)(|\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}|)). \end{aligned} \quad (\text{E36})$$

Now, we use the following famous results:

Lemma E1 [17]. For a general N-qubit pure state $\rho = |\phi\rangle \langle \phi|$, the following identity holds:

$$\int (d\phi) |\phi\rangle \langle \phi| = \frac{I}{2^n}, \quad (\text{E37})$$

$$\int (d\phi) |\phi\rangle \langle \phi|^{\otimes 2} = \frac{1}{2^n(2^n+1)}(I + \tau), \quad (\text{E38})$$

where I is identity operator and τ is the swap operator between two tensor product parties. The volume measure $(d\phi)$ refers

to the uniform distribution (Haar measure) of all single-qubit pure states.

From this, when we averaged over all pure input states ρ , all terms except $\frac{1}{4}$ are at most the order of $\frac{1}{2^n}$ for all σ . Hence, when $2^n \gg 1$:

$$\overline{\mathbb{E}(\hat{\sigma}_{\mathbf{A}}^{\text{req}})^2}^{\rho,\sigma} \simeq \frac{1}{4}. \quad (\text{E39})$$

Next,

$$\begin{aligned} & \mathbb{E}((\hat{\sigma}_{\mathbf{A}}^{\text{req}})^2) \\ &= 2^{3n-2} \text{tr}(\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(S_n^{\text{req}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|^{\otimes 3} \rho \otimes \sigma^{\otimes 2}) \\ &\rightarrow \overline{\mathbb{E}((\hat{\sigma}_{\mathbf{A}}^{\text{req}})^2)}^{\rho} \\ &= 2^{2n-2} \text{tr}(\mathbb{E}_{|\phi_{\mathbf{A}}^{\text{req}}\rangle \sim \mathcal{D}(S_n^{\text{req}})} |\phi_{\mathbf{A}}^{\text{req}}\rangle \langle \phi_{\mathbf{A}}^{\text{req}}|^{\otimes 2} \sigma^{\otimes 2}) \\ &(\because \text{Lemma E1}). \end{aligned} \quad (\text{E40})$$

Now, since σ is pure and real, let us rewrite $\sigma = |\phi\rangle \langle \phi|$, where $|\phi\rangle = \sum_{i \in \mathbb{Z}_2^n} c_i |i\rangle$ and $c_i \in \mathbb{R}$. Then $|\phi\rangle \langle \phi|^{\otimes 2} = \sum_{i,j,k,l \in \mathbb{Z}_2^n} c_i c_j c_k c_l |ij\rangle \langle kl|$. Since $\mathbf{c}$ lies on hypersphere S^{2^n-1} , we obtain that from Lemma 1,

$$\begin{aligned} & \overline{\mathbb{E}((\hat{\sigma}_{\mathbf{A}}^{\text{req}})^2)}^{\rho,\sigma} \\ &= \frac{1}{4A} \int_{S^{2^n-1}} d^{2^n-1} \Omega \sum_{i,j,k,l} c_i c_j c_k c_l |ij\rangle \langle kl| \\ &\quad \times \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} (|\mathbf{x}\mathbf{y}\rangle \langle \mathbf{x}\mathbf{y}| + |\mathbf{x}\mathbf{y}\rangle \langle \mathbf{y}\mathbf{x}| + |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}| - 2|\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}|) \\ &= \frac{1}{4A} \sum_{\mathbf{x}, \mathbf{y}} \int_{S^{2^n-1}} d^{2^n-1} \Omega (3c_x^2 c_y^2 - 2c_x^4), \end{aligned} \quad (\text{E42})$$

where A is the surface area of S^{2^n-1} . Note that the sphere is invariant under permutation, hence

$$\begin{aligned} \overline{\mathbb{E}((\hat{\sigma}_{\mathbf{A}}^{\text{req}})^2)}^{\rho,\sigma} &= \frac{1}{4A} \left\{ 2^n \int_{S^{2^n-1}} d^{2^n-1} \Omega c_1^4 \right. \\ &\quad \left. + 3 \cdot 2^n (2^n - 1) \int_{S^{2^n-1}} d^{2^n-1} \Omega c_1^2 c_2^2 \right\}. \end{aligned} \quad (\text{E43})$$

We will separately calculate the first and second terms of the right-hand side. By noting that $c_1 = \cos(\phi_1)$,

$$\begin{aligned} & \int_{S^{2^n-1}} d^{2^n-1} \Omega c_1^4 \\ &= \oint d\phi_1 d\phi_2 \cdots d\phi_{2^n-1} \cos^4(\phi_1) \\ &\quad \times \sin^{2^n-2}(\phi_1) \sin^{2^n-3}(\phi_2) \cdots \sin(\phi_{2^n-2}) \\ &\quad (\text{where } 0 \leq \phi_1, \phi_2, \dots, \phi_{2^n-2} \leq \pi, 0 \leq \phi_{2^n-1} \leq 2\pi) \\ &= \int_{S^{2^n-1}} d^{2^n-1} \Omega \frac{\int_0^\pi d\phi_1 \cos^4(\phi_1) \sin^{2^n-2}(\phi_1)}{\int_0^\pi d\phi_1 \sin^{2^n-2}(\phi_1)} \\ &= \frac{A \int_0^{\frac{\pi}{2}} d\phi_1 \cos^4(\phi_1) \sin^{2^n-2}(\phi_1)}{\int_0^{\frac{\pi}{2}} d\phi_1 \sin^{2^n-2}(\phi_1)}. \end{aligned} \quad (\text{E44})$$

We now use the formula

$$\int_0^{\frac{\pi}{2}} \sin^m(\phi) d\phi = \begin{cases} \frac{\pi}{2} \left(\frac{m!}{2^m ((\frac{m}{2})!)^2} \right) & \text{if } m \text{ is even} \\ \frac{2^{m-1} ((\frac{m-1}{2})!)^2}{m!} & \text{if } m \text{ is odd,} \end{cases} \quad (\text{E45})$$

and the beta function $\frac{1}{2}B(p, q) = \frac{\Gamma(p)\Gamma(q)}{2\Gamma(p+q)} = \int_0^{\frac{\pi}{2}} \sin^{2p-1}(\phi) \cos^{2q-1}(\phi) d\phi$, where $\Gamma(p)$ is gamma function. Then we obtain that with Stirling's formula,

$$\begin{aligned} \int_{S^{2^n-1}} d^{2^n-1} \Omega c_1^4 &= \frac{A}{\pi} \frac{B(\frac{5}{2}, 2^{n-1} - \frac{1}{2})}{\left(\frac{(2^n-2)!}{2^{(2^n-2)((2^{n-1}-1)!)^2} \right)} \\ &\simeq \frac{A\sqrt{2\pi e}}{\pi} \Gamma\left(\frac{5}{2}\right) \frac{2^{2^n-2}}{(2^{n-1}+1)2^{n-1}} \sqrt{\frac{(2^{n-1}-\frac{3}{2})(2^{n-1}-1)}{2^n-2}} \\ &\quad \cdot \frac{(2^{n-1}-\frac{3}{2})^{2^{n-1}-\frac{3}{2}} (2^{n-1}-1)^{2^{n-1}-1}}{(2^n-2)^{2^n-2}} \\ &\simeq \frac{A\sqrt{2\pi e}}{\pi} \Gamma\left(\frac{5}{2}\right) \\ &\quad \cdot \underbrace{2^{2^n-\frac{5n}{2}-1} \frac{(2^{n-1}-\frac{3}{2})^{2^{n-1}-\frac{3}{2}} (2^{n-1}-1)^{2^{n-1}-1}}{(2^n-2)^{2^n-2}}}_{\equiv C}. \end{aligned} \quad (\text{E46})$$

Note that

$$\begin{aligned} \log_2(C) &\simeq \left(2^n - \frac{5n}{2} - 1\right) + \left(2^{n-1} - \frac{3}{2}\right)(n-1) \\ &\quad + (2^{n-1}-1)(n-1) - (2^n-2)n \\ &= -3n + \frac{3}{2}, \end{aligned} \quad (\text{E47})$$

giving us

$$\frac{1}{A} \int_{S^{2^n-1}} c_1^4 d^{2^n-1} \Omega \simeq \frac{\sqrt{2\pi e}}{\pi} 2^{-3n+\frac{3}{2}} \Gamma\left(\frac{5}{2}\right). \quad (\text{E48})$$

Let us go to the second term:

$$\begin{aligned} \int_{S^{2^n-1}} d^{2^n-1} \Omega c_1^2 c_2^2 \\ &= \oint d\phi_1 d\phi_2 \cdots d\phi_{2^n-1} \cos^2(\phi_1) \cos^2(\phi_2) \\ &\quad \times \sin^2(\phi_2) \sin^{2^n-2}(\phi_1) \sin^{2^n-3}(\phi_2) \cdots \sin(\phi_{2^n-2}) \\ &\quad (\text{where } 0 \leq \phi_1, \phi_2, \dots, \phi_{2^n-2} \leq \pi, 0 \leq \phi_{2^n-1} \leq 2\pi). \end{aligned} \quad (\text{E49})$$

This leads to

$$\begin{aligned} \int d^{2^n-1} \Omega \frac{\int_0^\pi \cos^2(\phi_1) \sin^{2^n-2}(\phi_1) d\phi_1}{\int_0^\pi \sin^{2^n-2}(\phi_1) d\phi_1} \\ \times \frac{\int_0^\pi \cos^2(\phi_2) \sin^{2^n-1}(\phi_2) d\phi_2}{\int_0^\pi \sin^{2^n-3}(\phi_2) d\phi_2} \\ = \frac{A}{2\pi} \frac{B(\frac{3}{2}, 2^{n-1} - \frac{1}{2}) B(\frac{3}{2}, 2^{n-1})}{\left(\frac{(2^n-2)!}{2^{2^n-2}((2^{n-1}-1)!)^2} \right) \left(\frac{2^{2^n-4}((2^{n-1}-2)!)^2}{(2^n-3)!} \right)} \end{aligned}$$

$$\begin{aligned} &= \frac{2A}{\pi} \frac{\Gamma(\frac{3}{2})^2 (2^{n-1} - \frac{3}{2})! (2^{n-1}-1)! (2^{n-1}-1)^2}{(2^n-2)(2^{n-1})! (2^{n-1} + \frac{1}{2})!} \\ &\simeq \frac{2A}{\pi} \Gamma\left(\frac{3}{2}\right)^2 \frac{(2^{n-1})^2}{(2^n-2)2^{n-1} (2^{n-1} + \frac{1}{2}) (2^{n-1} - \frac{1}{2})} \\ &\simeq 2^{-2n} A \left[\because \Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2} \right]. \end{aligned} \quad (\text{E50})$$

Therefore, substituting Eqs. (E46), (E47), (E39), and (E49) into Eq. (E43) leads to

$$\overline{\text{Var}(\hat{\sigma}_A^{\text{req}})^{\rho, \sigma}} = \overline{\mathbb{E}((\hat{\sigma}_A^{\text{req}})^2)^{\rho, \sigma}} - \overline{\mathbb{E}(\hat{\sigma}_A^{\text{req}})^{2\rho, \sigma}} \simeq \frac{1}{2}. \quad (\text{E51})$$

Next, let us direct our attention to ESPOVM. As we proceed in a similar manner, we may skip most of the steps but note that

$$\begin{aligned} \sum_{|\phi_A^{\text{req}}\rangle \in S_n^{\text{eq}}} \frac{2^{2n}}{|S_n^{\text{eq}}|} \langle \phi_A^{\text{req}} | \rho | \phi_A^{\text{req}} \rangle \langle \phi_A^{\text{req}} | O | \phi_A^{\text{req}} \rangle \\ = \sum_{x \in \mathbb{Z}_2^n} \langle x | \rho | x \rangle \langle x | O | x \rangle + \text{tr}(O) + \text{tr}(O\rho) \end{aligned} \quad (\text{E52})$$

and

$$\begin{aligned} \overline{\mathbb{E}((\hat{\sigma}_A^{\text{eq}})^2)^{\rho, \sigma}} &= \frac{1}{A} \int_{c \in S^{2^n-1}} d^{2^n-1} \Omega \sum_{i,j,k,l} c_i c_j c_k c_l |ij\rangle \langle kl| \\ &\quad \times \sum_{x,y \in \mathbb{Z}_2^n} (|xy\rangle \langle xy| + |xy\rangle \langle yx| - |xx\rangle \langle xx|) \\ &= \frac{1}{A} \sum_{x,y} \int_{c \in S^{2^n-1}} d^{2^n-1} \Omega (2c_x^2 c_y^2 - c_x^4). \end{aligned} \quad (\text{E53})$$

By the same method with RESPOVM, we can find that $\overline{\mathbb{E}((\hat{\sigma}_A^{\text{eq}})^2)^{\rho, \sigma}} \simeq 2$ and $\overline{\mathbb{E}(\hat{\sigma}_A^{\text{eq}})^{2\rho, \sigma}} \simeq 1$. These result in $\overline{\text{Var}(\hat{O}_A^{\text{eq}})^{\rho, \sigma}} = \overline{\mathbb{E}((\hat{\sigma}_A^{\text{eq}})^2)^{\rho, \sigma}} - \overline{\mathbb{E}(\hat{\sigma}_A^{\text{eq}})^{2\rho, \sigma}} \simeq 1$. Lastly, for both ESPOVM and RESPOVM,

$$\begin{aligned} \mathbb{E}((\hat{\sigma}^{\text{bin}})^2) &= \sum_x \langle x | \rho | x \rangle |\langle x | \sigma | x \rangle|^2 \\ &= \sum_x \text{tr}((\rho \otimes \sigma \otimes \sigma) |xxx\rangle \langle xxx|) \end{aligned} \quad (\text{E54})$$

$$\rightarrow \overline{\mathbb{E}((\hat{\sigma}^{\text{bin}})^2)^{\rho}} = \frac{1}{2^n} \sum_x (\langle x | \sigma | x \rangle)^2 \simeq 0. \quad (\text{E55})$$

Hence, even if we average over real pure states σ , the result will be nearly 0. Lastly, let us consider randomized-Clifford tomography. Then, we can refer to Eq. (S43) of Ref. [9], stating that the estimation variance is equal to the shadow norm without maximizing over ρ and it is

$$\begin{aligned} \text{Var}(\hat{\sigma}^s) &= \frac{2^n+1}{2^n+2} \left[\text{tr}(\rho) \text{tr} \left(\left(\sigma - \frac{I}{2^n} \right)^2 \right) \right. \\ &\quad \left. + 2 \text{tr} \left(\rho \left(\sigma - \frac{I}{2^n} \right)^2 \right) \right] - \left[\text{tr} \left(\rho \left(\sigma - \frac{I}{2^n} \right) \right) \right]^2. \end{aligned} \quad (\text{E56})$$

After we average over pure ρ , by Lemma E1, the second and last terms are ignored and

$$\overline{\text{Var}(\hat{\sigma}^s)}^{\rho, \sigma} \simeq \overline{\text{tr}\left(\sigma^2 - \frac{2\sigma}{2^n} + \frac{I}{4^n}\right)}^\sigma \simeq 1, \quad (\text{E57})$$

which settles the proof for Lemma B2(i). Note that RESPOVM gives half the average variance than even randomized Clifford. Since we need two input states for a single sampling trial, the number of states is not improved but we can reduce the total gate count for the whole estimation process. This is because in RESPOVM, half of the input-state copies are used for direct Z-basis measurement.

The proof of Lemma B2(ii) is much simpler. We start with

$$\mathbb{E}((\hat{\sigma}_A^{\text{eq}})^2)^\rho = 2^{2n} \text{tr}(\mathbb{E}_{|\phi_A^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|^{\otimes 2} \sigma^{\otimes 2}). \quad (\text{E58})$$

By averaging over Haar-random pure state σ , by Lemma E1 and using the fact that $\text{tr}((\rho \otimes \sigma)\tau) = \text{tr}(\rho\sigma)$, we conclude that $\mathbb{E}((\hat{\sigma}_A^{\text{eq}})^2)^{\rho, \sigma} = 2$. Also, $\mathbb{E}(\hat{\sigma}_A^{\text{eq}})^2)^{\rho, \sigma} \simeq 1$, in exactly the same way as with Eq. (E52). Hence $\overline{\text{Var}(\hat{\sigma}_A^{\text{eq}})}^{\rho, \sigma} = \mathbb{E}((\hat{\sigma}_A^{\text{eq}})^2)^{\rho, \sigma} - \mathbb{E}(\hat{\sigma}_A^{\text{eq}})^2)^{\rho, \sigma} \simeq 1$. Furthermore, similar to Eq. (E57), the average variance of random Clifford tomography is also 1. The proof is thus completed.

APPENDIX F: INFORMATIONAL COMPLETENESS OF ESPOVM + COMPUTATIONAL BASIS

In this Appendix, we prove that the ESPOVM + computational basis is IC, so this POVM can fully extract all density-matrix information of any given input state. Explicitly, the condition for informational completeness of the POVM $\{\Pi_j\}_{j \in J}$, where J is an index set and $\Pi_j > 0$ for all $j \in J$, necessitates that any quantum state be expressible as a linear combination of the elements Π_j .

We now prove the following lemmas:

Lemma F1. For a d -dimensional Hilbert space, a POVM, comprising a set of $M \geq d^2$ positive-operator elements Π_j that sum to the identity, is IC if the matrix

$$\mathbf{V} = (|\Pi_1\rangle \rangle \quad |\Pi_2\rangle \rangle \quad \cdots \quad |\Pi_M\rangle \rangle) \quad (\text{F1})$$

has rank d^2 , and only then. Here $|\Pi_j\rangle \rangle$ is a basis-dependent superket (vectorization) of Π_j .

Proof. $\mathbf{V}$ has rank $d^2 = \dim(\mathbb{C}^d \otimes \mathbb{C}^d)$ if and only if it spans a set of vectorized quantum states. By reversing the vectorization (which is an injective map), we note that all quantum state can be expressed as a linear combination of $\{\Pi_j\}$. ■

Lemma F2. A POVM $\{\Pi_j\}$ is IC if the frame operator [45]

$$\mathcal{F} \equiv \sum_{j=1}^M \frac{|\Pi_j\rangle \rangle \langle \langle \Pi_j|}{\text{tr}(\Pi_j)} \quad (\text{F2})$$

is invertible, and only then.

Proof. Suppose we have the IC POVM. Clearly from Lemma F1, $\text{rank}\{\mathbf{V}\} = d^2$. Note that $\mathcal{F} \equiv \mathbf{V} \mathbf{\Pi}^{-1} \mathbf{V}^\dagger$, where

$$\mathbf{\Pi} = \begin{pmatrix} \text{tr}(\Pi_1) & 0 & \cdots & 0 \\ 0 & \text{tr}(\Pi_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \text{tr}(\Pi_M) \end{pmatrix}. \quad (\text{F3})$$

Also, if $\mathbf{x} \in \text{Ker}(\mathbf{V})$, it is easy to see that $\mathbf{x} \in \text{Ker}(\mathcal{F})$ or $\mathbf{x}^\top \mathcal{F} \mathbf{x} = \mathbf{x}^\top \mathbf{V}^\dagger \mathbf{\Pi}^{-1} \mathbf{V} \mathbf{x} = 0 \rightarrow \|\mathbf{V} \mathbf{x}\|^2 = 0$ [the inner product is defined by the Euclidean product with the weights $1/\text{tr}(\Pi_j) > 0$ ($j \in [M]$)]. Hence, $\mathbf{V} \mathbf{x} = 0 \rightarrow \mathbf{x} \in \text{Ker}(\mathbf{V})$. So, $\text{rank}(\mathcal{F}) = M - \dim(\text{Ker}(\mathcal{F})) = M - \dim(\text{Ker}(\mathbf{V})) = \text{rank}(\mathbf{V})$. Since $\text{rank}(\mathbf{V}) = d^2$, $\text{rank}(\mathcal{F}) = d^2$. As $\mathcal{F}$ is also a $d^2 \times d^2$ matrix, $\mathcal{F}^{-1}$ exists.

Conversely, suppose that $\mathcal{F}^{-1}$ exists. Then $\text{rank}\{\mathcal{F}\} = d^2$. Since $\text{rank}\{\mathcal{F}\} \leq \text{rank}\{\mathbf{V}\} \leq d^2$, we conclude that $\text{rank}\{\mathbf{V}\} = d^2$. By Lemma F1, $\{\Pi_j\}_{j \in J}$ is IC. ■

Setting $d = 2^n$ therefore yields the desirable result.

Proposition F1. The collection of ESPOVM and computational basis forms an IC measurement. However, ESPOVM itself is not an IC POVM.

Proof. First, we calculate the frame operator of the ESPOVM:

$$\begin{aligned} \mathcal{F}_{\text{es}} &= \sum_{|\phi_A^{\text{eq}}\rangle \in \mathcal{S}_n^{\text{eq}}} \frac{2^n}{|\mathcal{S}_n^{\text{eq}}|} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|^\top \langle \phi_A^{\text{eq}}| \langle \phi_A^{\text{eq}}|^* \\ &= 2^n (\mathbb{E}_{|\phi_A^{\text{eq}}\rangle \sim \mathcal{D}(\mathcal{S}_n^{\text{eq}})} |\phi_A^{\text{eq}}\rangle \langle \phi_A^{\text{eq}}|^{\otimes 2})^\top, \end{aligned} \quad (\text{F4})$$

where $\top_2$ means partial transpose for the second part. From Lemma 1, we note that

$$\begin{aligned} \mathcal{F}_{\text{es}} &= \frac{1}{2^n} \left(1 + \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}| - \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}| \right) \\ &= \frac{1}{2^n} \left(1 + \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \sum_{\mathbf{y} \neq \mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}| \right). \end{aligned} \quad (\text{F5})$$

One can show that the operator $W = \sum_{\mathbf{x} \in \mathbb{Z}_2^n} \sum_{\mathbf{y} \neq \mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{y}\mathbf{y}|$ has at least one eigenvalue equal to -1 by inspecting its action on $|\mathbf{z}\mathbf{z}\rangle$ for any $\mathbf{z} \in \mathbb{Z}_2^n$:

$$W |\mathbf{z}\mathbf{z}\rangle = \sum_{\mathbf{x} \neq \mathbf{z} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle = \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle - |\mathbf{z}\mathbf{z}\rangle. \quad (\text{F6})$$

It follows that any ket $(|\mathbf{z}_1, \mathbf{z}_1\rangle - |\mathbf{z}_2, \mathbf{z}_2\rangle)/\sqrt{2}$ with $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{Z}_2^n$ is an eigenket in the -1 eigenvalue subspace. As a side note, it is easy to check that the superposition $|\phi\rangle \equiv \sum_{\mathbf{z} \in \mathbb{Z}_2^n} |\mathbf{z}, \mathbf{z}\rangle 2^{-n/2}$ is an eigenket that gives a positive eigenvalue of $2^n - 1$:

$$\begin{aligned} W \sum_{\mathbf{z} \in \mathbb{Z}_2^n} |\mathbf{z}\mathbf{z}\rangle 2^{-n/2} &= \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \sum_{\mathbf{z} \neq \mathbf{x} \in \mathbb{Z}_2^n} 2^{-n/2} \\ &= \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle 2^{-n/2} (2^n - 1). \end{aligned} \quad (\text{F7})$$

We may then conclude that $\mathcal{F}_{\text{es}}$ has at least one null eigenvalue. Based on Lemma F2, it follows that the ESPOVM is non-IC.

Upon recognizing that $\sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}|$ is the frame operator $\mathcal{F}_{\text{comp}}$ for the computational basis, the total frame operator may be written as

$$\begin{aligned} \mathcal{F} &= \mathcal{F}_{\text{es}} + \mathcal{F}_{\text{comp}} \\ &= \frac{1}{2^n} + |\phi\rangle \langle \phi| + \left(1 - \frac{1}{2^n}\right) \sum_{\mathbf{x} \in \mathbb{Z}_2^n} |\mathbf{x}\mathbf{x}\rangle \langle \mathbf{x}\mathbf{x}| > 0. \end{aligned} \quad (\text{F8})$$

Therefore, $\mathcal{F}$ has full rank (with all eigenvalues positive) and is invertible. By Lemma F2, the ESPOVM + computational basis is IC POVM. ■

APPENDIX G: LINEAR NEAREST-NEIGHBORING (LNN) ARCHITECTURE OF CZ CIRCUITS [36]

In this Appendix, we explain how to transform the arbitrary CZ circuit to the $2n + 2$ -depth LNN form, which has only neighboring CNOT gates. We encourage readers who are interested in Clifford circuit synthesis to read Ref. [36], from which some techniques were borrowed. This section briefly discusses some relevant background content in Ref. [36].

The actual circuit we transform is the so-called $\widehat{\text{CZ}}$ circuit. This is defined by a CZ circuit followed by qubit-reversal layers, $\text{REV} : |x_1, x_2, \dots, x_n\rangle \mapsto |x_n, x_{n-1}, \dots, x_1\rangle$. However, the CZ section is equal to CZ-REV-REV and the last REV can be ignored by reversing the measurement outcome if we only do the nonadaptive measurement. Hence, if $\widehat{\text{CZ}}$ can be transformed to a $(2n + 2)$ -depth LNN architecture, the proof is complete.

Now consider some arbitrary $\widehat{\text{CZ}}$ circuit. We will prove this only when n is odd, since that for even n follows similarly. Suppose we take $|x_1, x_2, \dots, x_n\rangle$ as the input and define the two unitaries

$$C_1 \equiv \text{CNOT}_{1 \rightarrow 2} \text{CNOT}_{3 \rightarrow 4} \dots \text{CNOT}_{n-2 \rightarrow n-1} \text{CNOT}_{3 \rightarrow 2} \text{CNOT}_{5 \rightarrow 4} \dots \text{CNOT}_{n \rightarrow n-1}, \quad (\text{G1})$$

$$C_2 \equiv \text{CNOT}_{2 \rightarrow 1} \text{CNOT}_{4 \rightarrow 3} \dots \text{CNOT}_{n-1 \rightarrow n-2} \text{CNOT}_{2 \rightarrow 3} \text{CNOT}_{4 \rightarrow 5} \dots \text{CNOT}_{n-1 \rightarrow n}, \quad (\text{G2})$$

and $C \equiv C_1 C_2$. Now, let $n = 2m + 1$. If we operate C on the input $t \leq m$ times, the output state becomes

$$\begin{aligned} C^t |x_1, x_2, \dots, x_n\rangle \\ = \bigotimes_{i=1}^n |[Pj(n-3-2(t-1)+i), Pk(2(t-1)+i)]], \end{aligned} \quad (\text{G3})$$

where $[j, k] (j \neq k) \equiv x_j \oplus x_{j+1} \oplus \dots \oplus x_k$ ($\oplus$ is the mod 2 summation), $[j, j] \equiv x_j$, and the patterns:

$$\begin{aligned} Pj \equiv (n-1, n-3, n-3, \dots, 4, 4, 2, 2, 1, 1, \\ 3, 3, \dots, n-2, n-2), \end{aligned} \quad (\text{G4})$$

$$\begin{aligned} Pk \equiv (3, 3, 5, 5, \dots, n, n, n-1, n-1, n-3, \\ n-3, \dots, 6, 6, 4, 4, 2). \end{aligned} \quad (\text{G5})$$

Note that both are of length $2n - 3$. Also, $P(j \text{ or } k)(a)$ where $a \in [2n - 3]$ means the a th element of such patterns. We note that given $t \leq m$ and $i \in [n]$, there is no case when $[Pj(n-3-2(t-1)+i), Pk(2(t-1)+i)]$ is equal with $[Pj(n-3-2(t'-1)+i'), Pk(2(t'-1)+i')]$ for the other $t' \leq m$ and $i' \in [n]$. This is schematically seen in Ref. [36]. Also, the number of tuples obtained during $t = 0, 1, \dots, m$ is $(m+1)n = \frac{n(n+1)}{2} = \binom{n}{2} + n$. These mean that after we recorded all elements of output ket at each step we operate S , for m times, we can see all $[j, k] (j, k \in [n])$. Note that $[j, j]$ can be seen at the starting point ($t = 0$).

Now we show that C^{m+1} is a qubit-reversal operation. After the m th operation, we operate C once more and calculate the p th qubit output of the output state for an even p . By the neighboring property, we just need to sum the $\{p-1, p, p+1\}$ th inputs, resulting in $[n - (2\lfloor \frac{p-1}{2} \rfloor + 1), n - (2\lfloor \frac{p-2}{2} \rfloor - 1)] \oplus [n - (2\lfloor \frac{p}{2} \rfloor + 1), n - (2\lfloor \frac{p-1}{2} \rfloor + 1)] \oplus [n - (2\lfloor \frac{p+1}{2} \rfloor + 1), n - (2\lfloor \frac{p}{2} \rfloor + 1)] = n - p + 1$, which is a reversal of p . We can also note that the result for an odd p is also a reversal. Similarly, we conclude that S^{m+1} is a qubit reversal.

Next, we recall Eq. (C2), that is, $y_i = x_1 \oplus x_2 \oplus \dots \oplus x_i$ in the main text. In this context, $y_i = [1, i]$ and $y_j \oplus y_k = [j+1, k]$. Hence, one can achieve $i^{u_j y_j}$ and $i^{u_{j,k}(y_j \oplus y_k)}$ ($u_j, u_{j,k} \in \mathbb{Z}_4$) by inserting phase gates in each section of the C -operated output or at the starting point. Also, the depth of the LNN architecture is $4(m+1) = 2n + 2$. The 1-to-1 mapping $(t, i) \leftrightarrow [Pj(n-3-2(t-1)+i), Pk(2(t-1)+i)]$ must be recorded in $\mathcal{O}(n^2)$ memory before the circuit sampling, so we do not need to be bothered with the time to find the right location for phase-gate insertion. A similar consideration applies when n is even but with slightly different C_1, C_2 , and patterns [36]. In conclusion, we can realize the $\widehat{\text{CZ}}$ operation with only nearest-neighboring CNOT and phase gates of depth $2n + 2$. Lastly, we note that Y -basis measurements can be carried out with X -basis measurements after phase gates. If these phase operations are performed before the CZ action to transform the CZ circuit into an LNN architecture, the double-layered rightmost CNOT gates can simply be regarded as classical postprocessing of the X -basis outcome of the equatorial stabilizer measurement. Therefore, only $2n$ -depth circuits are needed.

[1] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comput.* **26**, 1484 (1997).

[2] D. Deutsch and R. Jozsa, Rapid solution of problems by quantum computation, *Proc. R. Soc. London A* **439**, 553 (1992).

- [3] D. Gavinsky, M. Roetteler, and J. Roland, Quantum algorithm for the Boolean hidden shift problem, in *Computing and Combinatorics*, Lecture Notes in Computer Science Vol. 6842 (Springer, Berlin, Heidelberg, 2011), pp. 158–167.
- [4] G. M. D'Ariano, M. De Laurentis, M. G. Paris, A. Porzio, and S. Solimeno, Quantum tomography as a tool for the characterization of optical devices, *J. Opt. B* **4**, S127 (2002).
- [5] Edited by M. G. A. Paris and J. Řeháček, in *Quantum State Estimation*, Lecture Notes in Physics Vol. 649 (Springer, Berlin, 2004).
- [6] R. O'Donnell and J. Wright, Efficient quantum tomography, *Proc. Annu. ACM Symp. Theory Comput.* **899** (2016).
- [7] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, Gate set tomography, *Quantum* **5**, 557 (2021).
- [8] S. Aaronson, Shadow tomography of quantum states *Proc. Annu. ACM Symp. Theory Comput.* **325** (2018).
- [9] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, *Nat. Phys.* **16**, 1050 (2020).
- [10] A. Acharya, S. Saha, and A. M. Sengupta, Shadow tomography based on informationally complete positive operator-valued measure, *Phys. Rev. A* **104**, 052418 (2021).
- [11] H.-Y. Huang, S. Chen, and J. Preskill, Learning to Predict Arbitrary Quantum Processes, *PRX Quantum* **4**, 040337 (2023).
- [12] L. Leone, S. F. Oliviero, and A. Hamma, Stabilizer Rényi entropy, *Phys. Rev. Lett.* **128**, 050402 (2022).
- [13] S. F. Oliviero, L. Leone, A. Hamma, and S. Lloyd, Measuring magic on a quantum processor, *npj Quantum Inf.* **8**, 148 (2022).
- [14] Z. Webb, The Clifford group forms a unitary 3-design, *Quantum Inf. Comput.* **16**, 1379 (2016).
- [15] H. Zhu, Multiqubit Clifford groups are unitary 3-designs, *Phys. Rev. A* **96**, 062336 (2017).
- [16] R. Kueng and D. Gross, Qubit stabilizer states are complex projective 3-designs, *arXiv:1510.02767*.
- [17] Z. Puchała and J. Miszczak, Symbolic integration with respect to the Haar measure on the unitary groups, *Bull. Pol. Acad. Sci.* **65**, 21 (2017).
- [18] D. Maslov and W. Yang, CNOT circuits need little help to implement arbitrary Hadamard-free Clifford transformations they generate, *npj Quantum Inf.* **9**, 96 (2023).
- [19] S. Bravyi and D. Maslov, Hadamard-free circuits expose the structure of the Clifford group, *IEEE Trans. Inf. Theory* **67**, 4546 (2021).
- [20] C. Bertoni, J. Haferkamp, M. Hinsche, M. Ioannou, J. Eisert, and H. Pashayan, Shallow shadows: Expectation estimation using low-depth random clifford circuits, *Phys. Rev. Lett.* **133**, 020602 (2024).
- [21] A. A. Akhtar, H.-Y. Hu, and Y.-Z. You, Scalable and flexible classical shadow tomography with tensor networks, *Quantum* **7**, 1026 (2023).
- [22] H.-Y. Hu, S. Choi, and Y.-Z. You, Classical shadow tomography with locally scrambled quantum dynamics, *Phys. Rev. Res.* **5**, 023027 (2023).
- [23] D. Grier, H. Pashayan, and L. Schaeffer, Sample-optimal classical shadows for pure states, *Quantum* **8**, 1373 (2024).
- [24] H. C. Nguyen, J. L. Bönsel, J. Steinberg, and O. Gühne, Optimizing shadow tomography with generalized measurements, *Phys. Rev. Lett.* **129**, 220502 (2022).
- [25] S. Bravyi and A. Kitaev, Universal quantum computation with ideal Clifford gates and noisy ancillas, *Phys. Rev. A* **71**, 022316 (2005).
- [26] X. Wang, X. Zhan, Y. Li, L. Xiao, G. Zhu, D. Qu, Q. Lin, Y. Yu, and P. Xue, Generalized quantum measurements on a higher-dimensional system via quantum walks, *Phys. Rev. Lett.* **131**, 150803 (2023).
- [27] S. Bravyi and J. Haah, Magic-state distillation with low overhead, *Phys. Rev. A* **86**, 052329 (2012).
- [28] S. T. Flammia and Y.-K. Liu, Direct fidelity estimation from few pauli measurements, *Phys. Rev. Lett.* **106**, 230501 (2011).
- [29] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Practical characterization of quantum devices without tomography, *Phys. Rev. Lett.* **107**, 210404 (2011).
- [30] S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, *Phys. Rev. A* **70**, 052328 (2004).
- [31] S. Bravyi and D. Gosset, Improved classical simulation of quantum circuits dominated by Clifford gates, *Phys. Rev. Lett.* **116**, 250501 (2016).
- [32] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Simulation of quantum circuits by low-rank stabilizer decompositions, *Quantum* **3**, 181 (2019).
- [33] Q. Zhang, Q. Liu, and Y. Zhou, Minimal-clifford shadow estimation by mutually unbiased bases, *Phys. Rev. Appl.* **21**, 064001 (2024).
- [34] Y. Wang and W. Cui, Classical shadow tomography with mutually unbiased bases, *Phys. Rev. A* **109**, 062406 (2024).
- [35] T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski, On mutually unbiased bases, *Int. J. Quantum Inf.* **08**, 535 (2010).
- [36] D. Maslov and M. Roetteler, Shorter stabilizer circuits via Bruhat decomposition and quantum circuit transformations, *IEEE Trans. Inform. Theory* **64**, 4729 (2018).
- [37] T. Schuster, J. Haferkamp, and H.-Y. Huang, Random unitaries in extremely low depth, *arXiv:2407.07754*.
- [38] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going beyond Bell's theorem, *arXiv:0712.0921*.
- [39] W. Dür, G. Vidal, and J. I. Cirac, Three qubits can be entangled in two inequivalent ways, *Phys. Rev. A* **62**, 062314 (2000).
- [40] D. Gottesman, The Heisenberg representation of quantum computers, *arXiv:quant-ph/9807006*.
- [41] See Appendix for additional detailed explanations and proofs of this paper.
- [42] M. Lerasle, Lecture notes: Selected topics on robust statistical learning theory, *arXiv:1908.10761*.
- [43] C. Blair, Problem complexity and method efficiency in optimization (A. S. Nemirovsky and D. B. Yudin), *SIAM Rev.* **27**, 264 (1985).
- [44] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani, Random generation of combinatorial structures from a uniform distribution, *Theor. Comput. Sci.* **43**, 169 (1986).
- [45] A. J. Scott, Tight informationally complete quantum measurements, *J. Phys. A: Math. Theor.* **39**, 13507 (2006).
- [46] D. Maslov and B. Zindorf, Depth optimization of CZ, CNOT, and Clifford circuits, *IEEE Trans. Quantum Eng.* **3**, 1 (2022).
- [47] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits anticconcentrate in log depth, *PRX Quantum* **3**, 010333 (2022).
- [48] H. J. Briegel and R. Raussendorf, Persistent entanglement in arrays of interacting particles, *Phys. Rev. Lett.* **86**, 910 (2001).

- [49] S. Anders and H. J. Briegel, Fast simulation of stabilizer circuits using a graph-state representation, *Phys. Rev. A* **73**, 022334 (2006).
- [50] A. Bäertschi and S. Eidenbenz, *Fundamentals of Computation Theory* (Springer International Publishing, Denmark, 2019), pp. 126–139.
- [51] C. Berge and J. C. Fournier, A short proof for a generalization of Vizing’s theorem, *J. Graph. Theory* **15**, 333 (1991).
- [52] J. Misra and D. Gries, A constructive proof of Vizing’s theorem, *Inf. Process. Lett.* **41**, 131 (1992).
- [53] S. Bravyi, J. A. Latone, and D. Maslov, 6-qubit optimal Clifford circuits, *npj Quantum Inf.* **8**, 79 (2022).
- [54] R. Duncan, S. P. Aleks Kissinger, and J. van de Wetering, Graph-theoretic simplification of quantum circuits with the ZX-calculus, *Quantum* **4**, 279 (2020).
- [55] K. Tsubouchi, T. Sagawa, and N. Yoshioka, Universal cost bound of quantum error mitigation based on quantum estimation theory, *Phys. Rev. Lett.* **131**, 210601 (2023).
- [56] R. LaRose, A. Tikku, É. O’Neel-Judy, L. Cincio, and P. J. Coles, Variational quantum state diagonalization, 6-qubit optimal Clifford circuits, *npj Quantum Inf.* **5**, 57 (2019).
- [57] L. Catani and D. E. Browne, State-injection schemes of quantum computation in Spekkens’ toy theory, *Phys. Rev. A* **98**, 052108 (2018).
- [58] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, Symmetrized characterization of noisy quantum processes, *Science* **317**, 1893 (2007).
- [59] J. J. Wallman and J. Emerson, Noise tailoring for scalable quantum computation via randomized compiling, *Phys. Rev. A* **94**, 052325 (2016).
- [60] S. J. Evered, D. Bluvstein, M. Kalinowski, S. Ebadi, T. Manovitz, H. Zhou, S. H. Li, A. A. Geim, T. T. Wang, N. Maskara, H. Levine, G. Semeghini, M. Greiner, V. Vuletić, and M. D. Lukin, High-fidelity parallel entangling gates on a neutral-atom quantum computer, *Nature (London)* **622**, 268 (2023).
- [61] X. Xue, M. Russ, N. Samkharadze, B. Undseth, A. Sammak, G. Scappucci, and L. M. K. Vandersypen, Quantum logic with spin qubits crossing the surface code threshold, *Nature (London)* **601**, 343 (2022).
- [62] V. Negîrneac, H. Ali, N. Muthusubramanian, F. Battistel, R. Sagastizabal, M. S. Moreira, J. F. Marques, W. J. Vlothuizen, M. Beekman, C. Zachariadis, N. Haider, A. Bruno, and L. DiCarlo, High-fidelity controlled-Z gate with maximal intermediate leakage operating at the speed limit in a superconducting quantum processor, *Phys. Rev. Lett.* **126**, 220502 (2021).
- [63] H. Jnane, J. Steinberg, Z. Cai, H. C. Nguyen, and B. Koczor, Quantum error mitigated classical shadows, *PRX Quantum* **5**, 010324 (2024).
- [64] D. E. Koh and S. Grewal, Classical shadows with noise, *Quantum* **6**, 776 (2022).
- [65] P.-G. Rozon, N. Bao, and K. Agarwal, Optimal twirling depth for classical shadows in the presence of noise, *Phys. Rev. Lett.* **133**, 130803 (2024).
- [66] S. Shi, B. Xu, K. Zhang, G.-S. Ye, D.-S. Xiang, Y. Liu, J. Wang, D. Su, and L. Li, High-fidelity photonic quantum logic gate based on near-optimal Rydberg single-photon source, *Nat. Commun.* **13**, 4454 (2022).
- [67] T. Xie, Z. Zhao, S. Xu, X. Kong, Z. Yang, M. Wang, Y. Wang, F. Shi, and J. Du, 99.92%-fidelity CNOT gates in solids by noise filtering, *Phys. Rev. Lett.* **130**, 030601 (2023).
- [68] T. G. d. Brugière, M. Baboulin, B. Valiron, S. Martiel, and C. Allouche, Reducing the depth of linear reversible quantum circuits, *IEEE Trans. Quantum Eng.* **2**, 1 (2021).
- [69] J. Jiang, X. Sun, S.-H. Teng, B. Wu, K. Wu, and J. Zhang, Optimal space-depth trade-off of CNOT circuits in quantum logic synthesis, *Proc. Annu. ACM-SIAM Symp. Discrete Algorithms* **213** (2020).
- [70] S. A. Kutin, D. P. Moulton, and L. M. Smithline, Computation at a distance, *arXiv:quant-ph/0701194*.
- [71] J.-M. Lee, W.-J. Lee, M.-S. Kim, S. Cho, J. J. Ju, G. Navickaite, and J. Fernandez, Controlled-NOT operation of sin-photon circuit using photon pairs from silicon-photon circuit, *Opt. Commun.* **509**, 127863 (2022).
- [72] A. Russo, E. Barnes, and S. E. Economou, Photonic graph state generation from quantum dots and color centers for quantum communications, *Phys. Rev. B* **98**, 085303 (2018).
- [73] M. Rimbach-Russ, S. G. J. Philips, X. Xue, and L. M. K. Vandersypen, Simple framework for systematic high-fidelity gate operations, *Quantum Sci. Technol.* **8**, 045025 (2023).
- [74] T. M. Graham, M. Kwon, B. Grinkemeyer, Z. Marra, X. Jiang, M. T. Lichtman, Y. Sun, M. Ebert, and M. Saffman, Rydberg-mediated entanglement in a two-dimensional neutral atom qubit array, *Phys. Rev. Lett.* **123**, 230501 (2019).
- [75] H. Levine, A. Keesling, G. Semeghini, A. Omran, T. T. Wang, S. Ebadi, H. Bernien, M. Greiner, V. Vuletić, H. Pichler, and M. D. Lukin, Parallel implementation of high-fidelity multi-qubit gates with neutral atoms, *Phys. Rev. Lett.* **123**, 170503 (2019).
- [76] M. Saffman, I. I. Beterov, A. Dalal, E. J. Pérez, and B. C. Sanders, Symmetric Rydberg controlled-z gates with adiabatic pulses, *Phys. Rev. A* **101**, 062309 (2020).
- [77] T. Wang, Z. Zhang, L. Xiang, Z. Jia, P. Duan, Z. Zong, Z. Sun, Z. Dong, J. Wu, Y. Yin, and G. Guo, Experimental realization of a fast controlled-Z gate via a shortcut to adiabaticity, *Phys. Rev. Appl.* **11**, 034030 (2019).
- [78] I. A. Simakov, G. S. Mazhorin, I. N. Moskalenko, N. N. Abramov, A. A. Grigorev, D. O. Moskalev, A. A. Pishchimova, N. S. Smirnov, E. V. Zikiy, I. A. Rodionov, and I. S. Besedin, Coupler microwave-activated controlled phase gate on fluxonium qubits, *PRX Quantum* **4**, 040321 (2023).
- [79] M. A. Rol, F. Battistel, F. K. Malinowski, C. C. Bultink, B. M. Tarasinski, R. Vollmer, N. Haider, N. Muthusubramanian, A. Bruno, B. M. Terhal, and L. DiCarlo, Fast, high-fidelity conditional-phase gate exploiting leakage interference in weakly anharmonic superconducting qubits, *Phys. Rev. Lett.* **123**, 120502 (2019).
- [80] S. Li, A. D. Castellano, S. Wang, Y. Wu, M. Gong, Z. Yan, H. Rong, H. Deng, C. Zha, C. Guo, L. Sun, C. Peng, X. Zhu, and J.-W. Pan, Realisation of high-fidelity nonadiabatic CZ gates with superconducting qubits, *npj Quantum Inf.* **5**, 84 (2019).
- [81] B. Foxen, C. Neill, A. Dunsworth, P. Roushan, B. Chiaro, A. Megrant, J. Kelly, Z. Chen, K. Satzinger, R. Barends, F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, S. Boixo, D. Buell, B. Burkett, Y. Chen, R. Collins *et al.* (Google AI Quantum), Demonstrating a continuous set of two-qubit gates for near-term quantum algorithms, *Phys. Rev. Lett.* **125**, 120504 (2020).