

Certifying Entanglement Dimensionality by k -Reduction Moments

Changhao Yi^{1,2,3,4,*†}, Xiaodi Li^{5,2,3,4,†‡} and Huangjun Zhu^{2,3,4}

¹Department of Physics, [Shanghai University](#), Shanghai 200444, China

²State Key Laboratory of Surface Physics, Department of Physics, and Center for Field Theory and Particle Physics, [Fudan University](#), Shanghai 200433, China

³Institute for Nanoelectronic Devices and Quantum Computing, [Fudan University](#), Shanghai 200433, China

⁴Shanghai Research Center for Quantum Sciences, Shanghai 201315, China

⁵iFLYTEK Research, Hefei 230088, China



(Received 8 February 2025; revised 26 August 2025; accepted 9 February 2026; published 20 March 2026)

In this paper, we combine the k -reduction map, the moment method, and randomized measurements into a practical protocol for certifying the entanglement dimensionality. Our approach is based on the observation that a state with entanglement dimensionality at most k must stay positive under the action of the k -reduction map. The core of our protocol utilizes the moment method to determine whether the k -reduced operator, i.e., the operator obtained after applying the k -reduction map on a quantum state, contains a negative eigenvalue or not. Notably, we propose a systematic method for constructing reduction moment criteria that apply to a broader range of states, including k -unfaithful states, compared with fidelity-based methods. The performance of our approach gets better and better with the moment order employed, which is corroborated by extensive numerical simulations. To apply our approach, it suffices to implement a unitary 3-design instead of a 4-design, which is more feasible in practice than the correlation matrix method. In the course of study, we show that the k -reduction negativity, the absolute sum of the negative eigenvalues of the k -reduced operator, is monotonic under local operations and classical communication for pure states.

DOI: [10.1103/cc1n-gmj1](https://doi.org/10.1103/cc1n-gmj1)

I. INTRODUCTION

Entanglement is a characteristic feature of quantum mechanics and a vital resource in quantum information processing [1]. Consequently, certifying entanglement is crucial for many practical applications and has been a cornerstone of quantum information science since its early days [2,3]. Over the past three decades, numerous entanglement criteria have emerged. Among these, the *positive partial transpose* (PPT) criterion stands out due to its simplicity and powerful detection capability [4–6]. However, the PPT criterion demands complete knowledge of the density matrix, which is rarely available. To remedy this problem, more practical approaches based on partial transpose moments and *randomized measurements* [7,8] have

recently been developed [7,9–11]. Similar approaches can also estimate Rényi entanglement entropy [12], entanglement entropy [13], and entanglement negativity [14].

For many quantum information processing tasks, including quantum computation and quantum simulation, the mere presence of entanglement is usually insufficient. Instead, specific forms of entanglement possessing desired properties, such as high-dimensional entanglement, are often required to fully benchmark the capabilities of a quantum platform. In this context, high-dimensional entanglement is typically quantified by the *Schmidt number* [15]. With the rapid advancement of quantum technologies, preparing sophisticated entangled states has evolved from theoretical imagination to experimental reality. Consequently, the efficient certification of high-dimensional entanglement has become a pressing challenge [16], one that remains difficult even when the density matrix is fully known.

Schmidt-number witnesses [15,17], analogous to entanglement witnesses, are popular tools for certifying the entanglement dimensionality. Many Schmidt-number witnesses are based on fidelities with certain target states, leading to *fidelity-based criteria*. A state is called *k-unfaithful* [18] if it cannot be certified by such

*Contact author: cyl@shu.edu.cn

†These authors contributed equally to this work.

‡Contact author: lixiaodi@fudan.edu.cn

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

Schmidt-number- k witnesses; otherwise, it is k -faithful. So far most laboratory protocols rely on fidelity-based criteria [16,17,19,20], which necessitate significant prior information about the prepared states and require the target state itself be k -faithful. These drawbacks impose severe limitations on practical applications and call for more powerful alternatives.

An ideal method for certifying high-dimensional entanglement should be invariant under local unitary transformations and require a measurement cost no greater than that of full state tomography. Recently, the *correlation matrix* (CM) method was proposed to address this challenge [21–24]. However, estimating the moments of the CM typically requires randomized measurements based on unitary 4-designs. Unfortunately, unitary 4-designs cannot be realized by discrete groups except for one special case, and most 4-designs known in the literature are not amenable to experimental realization [25–27]. Although a CM-based protocol has been experimentally realized [28], its applications in higher dimensions remain challenging.

In this work, we propose a simple yet powerful approach for certifying the entanglement dimensionality that is more amenable to experimental realization. Our approach leverages the observation that the k -reduction map [15,29], a typical positive map, preserves the positivity of any density operator with Schmidt number up to k . A violation of this positivity condition implies that the Schmidt number of the target state is at least $k + 1$. By combining the k -reduction map and the *moment method* [30,31], we develop a practical approach for certifying the entanglement dimensionality by virtue of the k -reduction moments, the moments of the k -reduced operator. These moments can be efficiently estimated using randomized measurements based on unitary 3-designs, such as the *Clifford group* [32,33], which is much more practical than previous approaches relying on unitary 4-designs. Furthermore, our approach applies to both k -faithful and k -unfaithful states, offering a significantly broader scope of applications compared with fidelity-based methods.

When the first $(2D - 1)$ k -reduction moments are available for quantum states on a D -dimensional Hilbert space, our criterion is equivalent to the k -reduction criterion and thus can certify the Schmidt numbers of all pure states. Moreover, the Schmidt numbers of *isotropic states* [15]

can be certified by the first three k -reduction moments. The superior performance of our approach is supported by rigorous theoretical analysis and extensive numerical experiments. When collective measurements are available, the k -reduction moments featured in our approach can be estimated much more efficiently with constant sample complexity.

The rest of this paper is organized as follows. In Sec. II, we review the definition of the Schmidt number and give an overview of the CM method. In Sec. III, we clarify the properties of the k -reduction map, k -reduction negativity, and k -reduction criterion in comparison with the CM criterion. In Sec. IV, we review the basic idea of the moment method and propose our certification criteria using k -reduction moments. In Sec. V, we demonstrate the performance of our certification criteria through numerical simulations. In Sec. VI, we present two methods for estimating reduction moments and discuss the sample complexities. Finally, we conclude in Sec. VII. A summary of the main results is presented in Table I. Technical proofs are relegated to the Appendices.

II. BACKGROUND

Given a finite-dimensional Hilbert space $\mathcal{H}$, denote by $\mathcal{L}(\mathcal{H})$, $\mathcal{H}^{\mathcal{H}}(\mathcal{H})$, and $\mathcal{L}_0^{\mathcal{H}}(\mathcal{H})$ the set of linear operators, the set of Hermitian operators, and the set of traceless Hermitian operators on $\mathcal{H}$, respectively. The spectrum (with multiplicities taken into account) and the minimum eigenvalue of a Hermitian operator O are denoted by $\sigma(O)$ and $\sigma_{\min}(O)$, respectively.

A. Schmidt number and Schmidt-number witnesses

Consider a bipartite system with Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let $d_A \equiv \dim(\mathcal{H}_A)$, $d_B \equiv \dim(\mathcal{H}_B)$, $D \equiv d_A d_B$, and $d \equiv \min\{d_A, d_B\}$. Without loss of generality, we assume $d_A \leq d_B$ throughout this paper, which means $d = d_A$. The *Schmidt number* (also known as Schmidt rank) of a pure state $|\psi\rangle$ in $\mathcal{H}_{AB}$ is defined as

$$\text{SN}(|\psi\rangle) \equiv \text{rank}(\rho_A), \quad (1)$$

where $\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|)$. Note that $\text{SN}(|\psi\rangle) = 1$ iff $|\psi\rangle$ is a product state. The *Schmidt spectrum* of $|\psi\rangle$, denoted by

TABLE I. Comparison between different Schmidt-number certification protocols for quantum states with total dimension D . The parameter χ in full state tomography means the algorithm is only guaranteed to work when the state has rank no larger than χ . This parameter is irrelevant to the Schmidt number of the state.

Method	Require knowledge about the target state	Order of unitary designs	Sample complexity
Fidelity-based [17]	Yes	/	$\mathcal{O}(1)$
Correlation matrix [22,23]	No	≥ 4 on $\mathcal{H}_A, \mathcal{H}_B$	Unknown
Full state tomography of rank- χ states [34]	No	4 on $\mathcal{H}_{AB}$	$\mathcal{O}(D^2 \chi)$
k -reduction moments (This work)	No	≥ 3 on $\mathcal{H}_{AB}$	$\mathcal{O}(D)$

$\sigma(\psi)$, is the set of eigenvalues of ρ_A (with multiplicities taken into account) [35], that is, $\sigma(\psi) \equiv \sigma(\rho_A)$.

Denote by $\{|i\rangle_A\}_i$ and $\{|i\rangle_B\}_i$ the computational bases on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. If $|\psi\rangle$ has Schmidt decomposition

$$|\psi\rangle = \sum_{i=0}^{d-1} \sqrt{\lambda_i} |i\rangle_A \otimes |i\rangle_B, \quad (2)$$

then its Schmidt spectrum reads $\sigma(\psi) = \{\lambda_i\}_{i=0}^{d-1}$. In addition, the vector $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_{d-1})$ is called the Schmidt vector of $|\psi\rangle$. This vector can be regarded as a point on the $(d-1)$ -dimensional probability simplex

$$\Delta_d \equiv \left\{ (x_0, x_1, \dots, x_{d-1}) : \sum_{i=0}^{d-1} x_i = 1, x_i \geq 0 \right\}. \quad (3)$$

If $\lambda_0 = \dots = \lambda_{r-1} = 1/r$ and $\lambda_r = \dots = \lambda_{d-1} = 0$, then we get the maximally entangled state with Schmidt number r ,

$$|+_r\rangle \equiv \frac{1}{\sqrt{r}} \sum_{i=0}^{r-1} |i\rangle_A \otimes |i\rangle_B. \quad (4)$$

The concept of Schmidt number can be generalized to mixed states. Denote by $\mathcal{S}_r$ the set of quantum states on $\mathcal{H}_{AB}$ that can be expressed as convex combinations of pure states with Schmidt numbers no larger than r . By definition we have the following hierarchy relation:

$$\mathcal{S}_1 \subset \mathcal{S}_2 \subset \dots \subset \mathcal{S}_d = \mathcal{S}(\mathcal{H}_{AB}), \quad (5)$$

where $\mathcal{S}(\mathcal{H}_{AB})$ denotes the set of all density operators on $\mathcal{H}_{AB}$. The Schmidt number $\text{SN}(\rho)$ of $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ is defined as the smallest integer r such that $\rho \in \mathcal{S}_r$. Alternatively, $\text{SN}(\rho)$ can be expressed as follows [15]:

$$\text{SN}(\rho) \equiv \min_{\mathbb{D}(\rho)} \max_{|\phi_i\rangle \in \mathbb{D}(\rho)} \text{SN}(|\phi_i\rangle), \quad (6)$$

where $\mathbb{D}(\rho)$ denotes a pure-state decomposition of ρ . A quantum state with Schmidt number r can be prepared from $|+_r\rangle$ using local operations and classical communication (LOCC), but cannot be prepared from any state with a lower Schmidt rank. Hence, the Schmidt number of a mixed state ρ quantifies its entanglement dimension.

Similar to entanglement witnesses, an observable O_W is termed a *Schmidt-number- k* witness ($k \geq 2$) [29] if $\text{Tr}(\rho O_W) \geq 0$ for all $\rho \in \mathcal{S}_{k-1}$, and there exists at least one $\rho \in \mathcal{S}_k$ such that $\text{Tr}(\rho O_W) < 0$. The most common Schmidt-number- k witness is

$$\frac{k-1}{d} I - |+_d\rangle\langle+_d|. \quad (7)$$

For more general constructions, we have the following well-known result [15,18]. Suppose $|\psi\rangle \in \mathcal{H}_{AB}$. Then

$cI - |\psi\rangle\langle\psi|$ is a Schmidt-number- k witness iff

$$\text{SN}(|\psi\rangle) \geq k, \quad \sum_{i=0}^{k-2} \lambda_i^\downarrow \leq c < \sum_{i=0}^{k-1} \lambda_i^\downarrow, \quad (8)$$

where $\lambda_i^\downarrow$ is the i th largest element in $\sigma(\psi)$.

B. Schmidt-number certification via the correlation matrix

Here we review the CM method for certifying the Schmidt number [21–23], assuming that $d_A = d_B = d$. The CM criterion is a weaker but more concise version of the covariance matrix criterion [24]. Let $\{P_i^{(A)}\}_{i=1}^{d^2-1}$ and $\{P_i^{(B)}\}_{i=1}^{d^2-1}$ be orthogonal operator bases of $\mathcal{L}_0^H(\mathcal{H}_A)$ and $\mathcal{L}_0^H(\mathcal{H}_B)$, respectively, that satisfy the conditions

$$\begin{aligned} \text{Tr}(P_i^{(A)}) &= \text{Tr}(P_i^{(B)}) = 0, \\ \text{Tr}(P_i^{(A)} P_j^{(A)}) &= \text{Tr}(P_i^{(B)} P_j^{(B)}) = d \delta_{ij}. \end{aligned} \quad (9)$$

The CM of ρ with respect to these bases is the $(d^2-1) \times (d^2-1)$ matrix with entries

$$T_{ij} \equiv \frac{1}{d} \text{Tr} \left[\rho \left(P_i^{(A)} \otimes P_j^{(B)} \right) \right]. \quad (10)$$

Suppose T has singular values $\{v_i\}_{i=1}^{d^2-1}$. Then its Schatten p -norm reads

$$\|T\|_p \equiv \left(\sum_{i=1}^{d^2-1} v_i^p \right)^{\frac{1}{p}} = \left\{ \text{Tr} \left[(TT^\top)^{\frac{p}{2}} \right] \right\}^{\frac{1}{p}}. \quad (11)$$

The CM criterion states that [22,23]

$$\text{if } \text{SN}(\rho) \leq k, \text{ then } \|T\|_1 \leq k - \frac{1}{d}. \quad (12)$$

Note that this criterion is independent of the specific choices of local operator bases. If the second inequality in Eq. (12) is violated, then we can conclude that $\text{SN}(\rho)$ is larger than k .

In practice, it is difficult to estimate the singular values of T and the 1-norm $\|T\|_1$. Nevertheless, the second and fourth moments of the singular values, that is, $(\|T\|_2^2, \|T\|_4^4)$, can be estimated from the second and fourth order Haar-randomized correlators $\mathcal{C}^{(n)}$ constructed using

a special observable O ,

$$\mathcal{C}^{(n)} \equiv \int dU_A \int dU_B \left\langle U_A O U_A^\dagger \otimes U_B O U_B^\dagger \right\rangle_\rho^n, \quad (13)$$

where $\langle O \rangle_\rho \equiv \text{Tr}(\rho O)$. More precisely, we have the following relations [21]:

$$\begin{aligned} \mathcal{C}^{(2)} &= \frac{d^2}{(d-1)^2} \|T\|_2^2, \\ \mathcal{C}^{(4)} &= \frac{2d^4}{3(d-1)^4} \|T\|_4^4 + \frac{1}{3} (\mathcal{C}^{(2)})^2. \end{aligned} \quad (14)$$

Define the region

$$\mathcal{T}_k \equiv \{(\|T\|_2^2, \|T\|_4^4) : \|T\|_1 \leq k - d^{-1}\}, \quad (15)$$

whose boundary can be determined by solving an optimization problem. Then a moment-based CM criterion can be formulated as follows:

$$\text{if } \text{SN}(\rho) \leq k, \text{ then } (\|T\|_2^2, \|T\|_4^4) \in \mathcal{T}_k. \quad (16)$$

This criterion is referred to as the *moment-based CM criterion* henceforth.

In addition, by virtue of the Hölder inequality we can deduce that

$$\begin{aligned} \|T\|_2^2 &= \sum_{i=1}^{d^2-1} v_i^{\frac{2}{3}} v_i^{\frac{4}{3}} \leq \left(\sum_{i=1}^{d^2-1} v_i \right)^{\frac{2}{3}} \left(\sum_{i=1}^{d^2-1} v_i^4 \right)^{\frac{1}{3}} \\ &= \|T\|_1^{\frac{2}{3}} \|T\|_4^{\frac{4}{3}}, \end{aligned} \quad (17)$$

which implies that

$$\|T\|_1 \geq \frac{\|T\|_2^3}{\|T\|_4^2}. \quad (18)$$

In conjunction with Eq. (12) we can deduce an alternative moment-based criterion

$$\text{if } \text{SN}(\rho) \leq k, \text{ then } \frac{\|T\|_2^3}{\|T\|_4^2} \leq k - \frac{1}{d}, \quad (19)$$

which is potentially weaker but simpler than the counterpart in Eq. (16).

Currently, it is still quite challenging to apply the moment-based criteria in real experiments, although it is easier than the original CM criterion. To estimate the fourth order correlator $\mathcal{C}^{(4)}$ experimentally, we need to implement unitary transformations in a unitary 4-design. Unfortunately, the Clifford group widely used in quantum information processing fails to be a unitary 4-design

[26,32,33], and it is very challenging to sample from a unitary 4-design in experiments. Therefore, it is desirable to construct alternative certification protocols that do not rely on unitary 4-designs, which is a main motivation behind the current work.

III. SCHMIDT-NUMBER CERTIFICATION VIA THE k -REDUCTION MAP

A. k -positive maps

Given a finite-dimensional Hilbert space $\mathcal{H}$, an operator $O \in \mathcal{L}(\mathcal{H})$ is positive (semidefinite), denoted as $O \geq 0$, if $\langle \psi | O | \psi \rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$. A linear map $\mathcal{M} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$ is *positive* if

$$\mathcal{M}(O) \geq 0 \quad \forall O \geq 0. \quad (20)$$

A linear map $\mathcal{M} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$ is *k -positive* if the map $\mathcal{I}_k \otimes \mathcal{M}$ is positive, where $\mathcal{I}_k$ is the identity map on an auxiliary Hilbert space of dimension k [36–39].

Next, we turn to a bipartite system with Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and explain the applications of k -positive maps in certifying high-dimensional entanglement. Let $\mathcal{I}_A$ be the identity map on $\mathcal{L}(\mathcal{H}_A)$. A linear map $\mathcal{M}$ on $\mathcal{L}(\mathcal{H}_B)$ is *k -positive* iff $\mathcal{I}_A \otimes \mathcal{M}(|\psi\rangle\langle\psi|)$ is positive for every pure state $|\psi\rangle \in \mathcal{H}_{AB}$ with $\text{SN}(|\psi\rangle) \leq k$ [15,40]. Then we also have $\mathcal{I}_A \otimes \mathcal{M}(\rho) \geq 0$ for every mixed state ρ with $\text{SN}(\rho) \leq k$, but the converse is not necessarily true. Therefore, if $\mathcal{I}_A \otimes \mathcal{M}(\rho)$ is not positive, then we can conclude that $\text{SN}(\rho) > k$. This fact is instrumental for constructing protocols for certifying high-dimensional entanglement.

B. k -reduction map and k -reduction criterion

One of the most important and common k -positive maps is the k -reduction map [37,38], which is defined as follows:

$$\mathcal{R}_k(O) \equiv k \text{Tr}(O) I - O. \quad (21)$$

Here we are mainly interested in the map $\mathcal{I}_A \otimes \mathcal{R}_k$ acting on $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$. To simplify the following discussion, the expression $\mathcal{I}_A \otimes \mathcal{R}_k(\rho)$ will be abbreviated as $\mathcal{R}_k(\rho)$ and called the *k -reduced operator* associated with ρ . When $\rho = |\psi\rangle\langle\psi|$ is a pure state, $\mathcal{R}_k(\rho)$ can further be abbreviated as $\mathcal{R}_k(\psi)$. Our approach for certifying the Schmidt number is based on the following well-known result [15]:

Proposition 1. If $\text{SN}(\rho) \leq k$, then

$$\mathcal{R}_k(\rho) = k \rho_A \otimes I_B - \rho \geq 0. \quad (22)$$

When ρ is a pure state, the condition $\text{SN}(\rho) \leq k$ is also necessary to guarantee Eq. (22).

If the condition in Eq. (22), known as the *k -reduction condition*, is violated, then we can conclude

that $\text{SN}(\rho) > k$. The basic idea of this criterion is analogous to the famous PPT criterion. By considering the expectation value of the k -reduced operator with respect to the maximally entangled state $|+_d\rangle$ we can deduce that

$$\text{if } \text{SN}(\rho) \leq k, \quad \text{then } \langle +_d | \rho | +_d \rangle \leq \frac{k}{d}, \quad (23)$$

which reproduces a fidelity-based criterion.

By analogy with the entanglement negativity [40], here we define the k -reduction negativity of ρ , denoted by $\mathcal{N}_k(\rho)$, as the absolute value of the sum of the negative eigenvalues of $\mathcal{R}_k(\rho)$. When $\rho = |\psi\rangle\langle\psi|$ is a pure state, $\mathcal{N}_k(\rho)$ is abbreviated as $\mathcal{N}_k(\psi)$. Alternatively, $\mathcal{N}_k(\rho)$ can be expressed as follows:

$$\begin{aligned} \mathcal{N}_k(\rho) &= \frac{1}{2} (\|\mathcal{R}_k(\rho)\|_1 - \text{Tr}[\mathcal{R}_k(\rho)]) \\ &= \frac{1}{2} (\|\mathcal{R}_k(\rho)\|_1 - kd_B + 1). \end{aligned} \quad (24)$$

It measures the degree to which the k -reduction criterion is violated. With this definition, the k -reduction criterion can also be formulated as follows:

$$\text{if } \text{SN}(\rho) \leq k, \quad \text{then } \mathcal{N}_k(\rho) = 0. \quad (25)$$

In other words, if $\mathcal{N}_k(\rho) > 0$, then we can deduce that $\text{SN}(\rho) > k$.

C. k -reduction negativity of pure states

The k -reduction negativity plays an important role in certifying high-dimensional entanglement as we have seen in Sec. III B. Here we determine the k -reduction negativity of a general bipartite pure state $|\psi\rangle$ in $\mathcal{H}_{AB}$ and clarify its basic properties. To this end, we need to introduce some additional notations and results.

Given two real vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^d$, we can rearrange the entries of each vector in nonincreasing order to obtain $\mathbf{x}^\downarrow = (x'_0, x'_1, \dots, x'_{d-1})$ and $\mathbf{y}^\downarrow = (y'_0, y'_1, \dots, y'_{d-1})$. If

$$\sum_{i=0}^{\ell-1} x'_i \geq \sum_{i=0}^{\ell-1} y'_i \quad \forall \ell, \quad (26)$$

and the inequality is saturated when $\ell = d$, then we say $\mathbf{x}$ majorizes $\mathbf{y}$, denoted by $\mathbf{x} \succ \mathbf{y}$ or $\mathbf{y} \prec \mathbf{x}$. A function f on a subset of $\mathbb{R}^d$ is *Schur concave* if the condition $\mathbf{x} \succ \mathbf{y}$ implies that $f(\mathbf{x}) \leq f(\mathbf{y})$ [35, 41].

Given $\lambda \in \Delta_d$, let $\Omega_k(\lambda)$ be the $d \times d$ matrix with entries

$$\Omega_k(\lambda)_{ij} = k\lambda_i\delta_{ij} - \sqrt{\lambda_i\lambda_j} \quad (27)$$

and define

$$\theta_k(\lambda) \equiv \max\{0, -\sigma_{\min}(\Omega_k(\lambda))\}. \quad (28)$$

This function is interesting because of its close relation with the k -reduction negativity as we shall see shortly.

The basic properties of $\Omega_k(\lambda)$ and $\theta_k(\lambda)$ are summarized in Lemmas A1 and A2 in Appendix A.

Now, we can clarify the basic properties of the k -reduced operator $\mathcal{R}_k(\psi)$ as summarized in the following theorem and proved in Appendix A. Part of the results has been proved in previous works [15, 38].

Theorem 1. Suppose $|\psi\rangle \in \mathcal{H}_{AB}$ is a pure state with Schmidt vector λ , Schmidt number r , and $\tilde{r}$ distinct nonzero Schmidt coefficients. Then the k -reduced operator $\mathcal{R}_k(\psi)$ has spectrum

$$\sigma(\mathcal{R}_k(\psi)) = \sigma(\Omega_k(\lambda)) \cup [k\sigma(\psi)]^{\times(d_B-1)}, \quad (29)$$

where the superscript $d_B - 1$ denotes the multiplicity. In addition, $\mathcal{R}_k(\psi)$ has $2\tilde{r}$ distinct nonzero eigenvalues when $k \neq r$ and $2\tilde{r} - 1$ distinct nonzero eigenvalues when $k = r$. If $k \geq r$, then $\Omega_k(\lambda) \geq 0$ and $\mathcal{R}_k(\psi) \geq 0$. If instead $1 \leq k < r$, then both $\Omega_k(\lambda)$ and $\mathcal{R}_k(\psi)$ have exactly one negative eigenvalue.

Theorem 1 implies that

$$\mathcal{N}_k(\psi) = \max\{0, -\sigma_{\min}(\mathcal{R}_k(\psi))\} = \theta_k(\lambda). \quad (30)$$

In conjunction with Lemma A1 in Appendix A, we can further clarify the properties of $\mathcal{N}_k(\psi)$. The following theorem is also proved in Appendix A.

Theorem 2. Suppose $|\psi\rangle \in \mathcal{H}_{AB}$ is a pure state with Schmidt vector λ . Then $\theta_k(\lambda)$ is Schur concave in λ and nonincreasing in k . If $\text{SN}(|\psi\rangle) = r$, then $\mathcal{N}_k(\psi) \leq \max\{0, 1 - k/r\}$ and the inequality is saturated when $|\psi\rangle$ is a maximally entangled state with $\text{SN}(|\psi\rangle) = r$.

Suppose $|\psi\rangle$ and $|\psi'\rangle$ are two pure states in $\mathcal{H}_{AB}$. According to Nielsen's majorization criterion [42], $|\psi\rangle$ can be transformed into $|\psi'\rangle$ under LOCC iff the Schmidt vector of $|\psi\rangle$ is majorized by that of $|\psi'\rangle$. In conjunction with Theorem 2, we can deduce that the k -reduction negativity is nonincreasing under LOCC and thus can serve as a high-dimensional entanglement monotone for pure states.

By virtue of Theorem 2 we can further derive tight bounds on the k -reduction negativity $\mathcal{N}_k(\rho)$ and eigenvalues of $\mathcal{R}_k(\rho)$ for a general mixed state ρ .

Corollary 1. Suppose $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ with $d_A = d_B = d$. Then

$$\mathcal{N}_k(\rho) \leq 1 - \frac{k}{d}, \quad \left(\frac{k}{d} - 1\right)I \leq \mathcal{R}_k(\rho) \leq kI. \quad (31)$$

D. k -reduction negativity of pure states with depolarizing noise

As a generalization, in this section we determine the k -reduction negativity of an arbitrary pure state with depolarizing noise. Suppose $|\psi\rangle \in \mathcal{H}_{AB}$ is a pure state with $\text{SN}(|\psi\rangle) = r$ and ρ has the form

$$\rho = (1 - \varepsilon)|\psi\rangle\langle\psi| + \varepsilon \frac{I}{d_A d_B}, \quad \varepsilon \in [0, 1]. \quad (32)$$

Then

$$\begin{aligned} \mathcal{R}_k(\rho) &= (1 - \varepsilon)\mathcal{R}_k(\psi) + \varepsilon \mathcal{R}_k\left(\frac{I}{d_A d_B}\right) \\ &= (1 - \varepsilon)\mathcal{R}_k(\psi) + \frac{\varepsilon(kd_B - 1)}{d_A d_B} I, \end{aligned} \quad (33)$$

and there exists a simple linear relation between the spectrum of $\mathcal{R}_k(\rho)$ and the spectrum of $\mathcal{R}_k(\psi)$. Notably, the minimal eigenvalue of $\mathcal{R}_k(\rho)$ equals

$$(1 - \varepsilon)\sigma_{\min}(\mathcal{R}_k(\psi)) + \frac{\varepsilon(kd_B - 1)}{d_A d_B}. \quad (34)$$

In conjunction with Theorems 1 and 2, it is now straightforward to deduce the following result.

Proposition 2. The k -reduction negativity of ρ in Eq. (32) reads

$$\mathcal{N}_k(\rho) = \begin{cases} 0, & \varepsilon \geq \varepsilon^*, \\ (1 - \varepsilon)\mathcal{N}_k(\psi) - \frac{\varepsilon(d_B k - 1)}{d_A d_B}, & \varepsilon < \varepsilon^*, \end{cases} \quad (35)$$

where

$$\varepsilon^* \equiv \frac{d_A d_B \mathcal{N}_k(\psi)}{kd_B - 1 + d_A d_B \mathcal{N}_k(\psi)}. \quad (36)$$

Next, we turn to the special case in which $|\psi\rangle$ is a maximally entangled state of Schmidt rank r . Define

$$\rho_{\varepsilon,r} \equiv (1 - \varepsilon)|+_r\rangle\langle+_r| + \varepsilon \frac{I}{d_A d_B}, \quad \varepsilon \in [0, 1]. \quad (37)$$

The dependence of $\mathcal{N}_k(\rho_{\varepsilon,r})$ on k and ε is illustrated in Fig. 1. Define $\varepsilon_c^{\text{RM}}$ as the minimum value of $\varepsilon \geq 0$ such that $\mathcal{N}_{r-1}(\rho_{\varepsilon,r}) = 0$. By virtue of Proposition 2 with $\mathcal{N}_{r-1}(+_r) = 1/r$ we can deduce that

$$\varepsilon_c^{\text{RM}} = \left(1 + \frac{r^2 - r}{d_A} - \frac{r}{d_A d_B}\right)^{-1}. \quad (38)$$

As the local dimensions d_A, d_B increase, $\varepsilon_c^{\text{RM}}$ approaches 1, so the k -reduction criterion can certify the Schmidt number of $\rho_{\varepsilon,r}$ with arbitrary noise strength asymptotically.

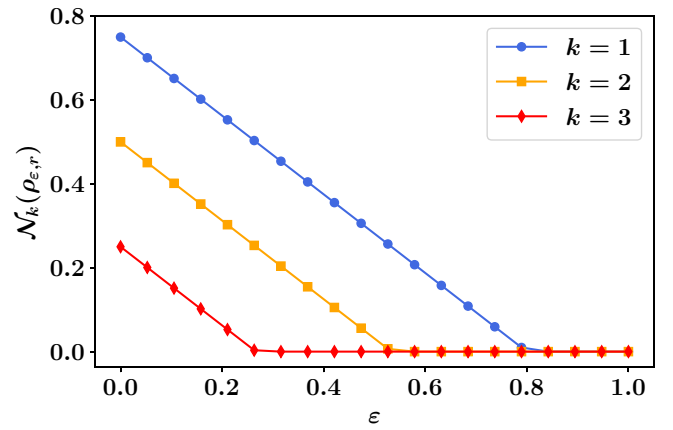


FIG. 1. The k -reduction negativity $\mathcal{N}_k(\rho_{\varepsilon,r})$ as a function of ε for $k = 1, 2, 3$. Here we set $r = d_A = d_B = 4$. The k -reduction negativity is nonincreasing in k and ε as predicted in Proposition 2.

Using the k -reduction criterion, we can further derive the following informative lower and upper bounds for the Schmidt number of $\rho_{\varepsilon,r}$ as proved in Appendix A 4.

Proposition 3. The Schmidt number of $\rho_{\varepsilon,r}$ in Eq. (37) satisfies

$$\left\lceil \frac{(1+u)r}{1+du} \right\rceil \leq \text{SN}(\rho_{\varepsilon,r}) \leq \left\lceil \frac{(1+u)r}{1+r^2 u} \right\rceil, \quad (39)$$

where $u = \varepsilon / [(1 - \varepsilon)d_A d_B]$. If $r \leq \sqrt{d}$ and $\varepsilon < 1/2$, then $\text{SN}(\rho_{\varepsilon,r}) = r$.

E. Comparison with the CM criterion

In this section we compare the k -reduction criterion in Eq. (22) with the CM criterion in Eq. (12) under various situations. From Proposition 1, we know that the k -reduction criterion can certify the Schmidt number of any pure state. By contrast, the CM criterion cannot. For example, consider the following state with Schmidt rank $r = 4$:

$$\sqrt{\frac{4}{5}}|00\rangle + \sqrt{\frac{1}{15}}|11\rangle + \sqrt{\frac{1}{15}}|22\rangle + \sqrt{\frac{1}{15}}|33\rangle, \quad (40)$$

assuming that $d_A = d_B = d = 4$. The 1-norm of its CM is approximately 2.7231, which is smaller than the upper bound $r - 1 - d^{-1} = 11/4$ in Eq. (12) with $k = r - 1$. Thus, from the CM criterion, we can only conclude that its Schmidt number is at least 3, instead of the actual Schmidt number 4.

Next, we consider the state $\rho_{\varepsilon,r}$ defined in Eq. (37). According to the discussion in Sec. III D, the k -reduction criterion can certify that the Schmidt number of $\rho_{\varepsilon,r}$ is r when ε is smaller than the threshold $\varepsilon_c^{\text{RM}}$ in Eq. (38). As

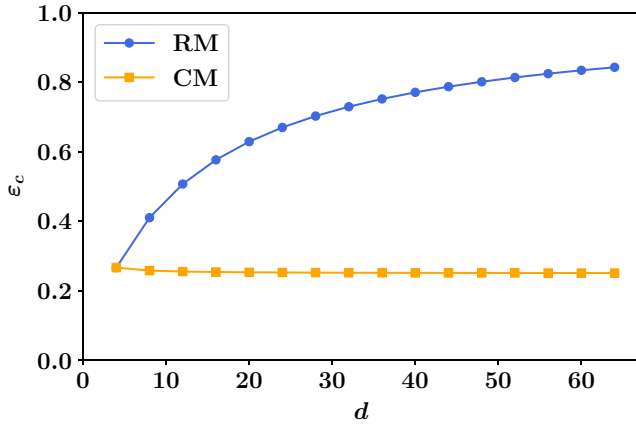


FIG. 2. Noise thresholds $\varepsilon_c^{\text{RM}}$ and $\varepsilon_c^{\text{CM}}$ associated with the k -reduction criterion and correlation matrix criterion in certifying the Schmidt number of the state $\rho_{\varepsilon,r}$ in Eq. (37) with $r = 4$. Here $d_A = d_B = d$, and $\varepsilon_c^{\text{RM}}$ is given in Eq. (38).

an analogy, we define $\varepsilon_c^{\text{CM}}$ as the maximum value of ε satisfying $\|T\|_1 \geq r - 1 - d^{-1}$. The relations between the two thresholds and the local dimension d are illustrated in Fig. 2, which implies that the k -reduction criterion can tolerate much higher noise strength than the CM criterion.

We then consider Haar-random pure states with depolarizing noise. Define the ensemble

$$\mathcal{E}_{\text{Haar},\varepsilon} \equiv \left\{ (1 - \varepsilon)|v\rangle\langle v| + \varepsilon \frac{I}{d^2} : |v\rangle \in \mathcal{H}_{\text{AB}} \right\}, \quad (41)$$

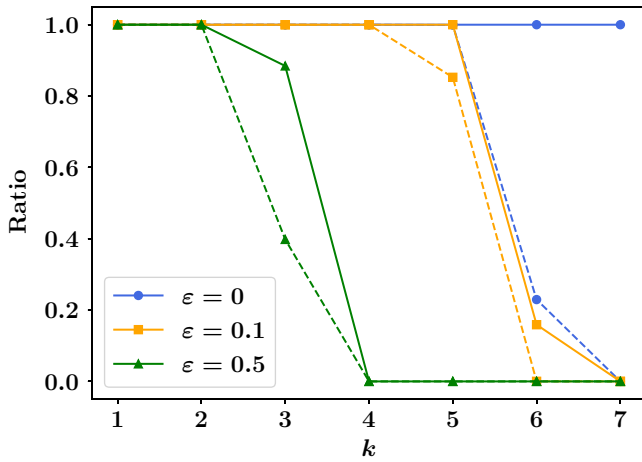


FIG. 3. Detection ratios of the k -reduction criterion and correlation matrix criterion for the ensemble of states in Eq. (41). The solid lines correspond to the k -reduction criterion in Eq. (22), while the dashed lines correspond to the correlation matrix criterion in Eq. (12). Here $d_A = d_B = 8$ and three different noise strengths are indicated in the legend. The performance of the k -reduction criterion is better than the correlation matrix criterion in all cases under consideration.

TABLE II. Detection ratio of the k -reduction criterion in Eq. (22) for the ensemble $\mathcal{E}_{256,K}$. Here $d_A = d_B = 16$.

k	K				
	2	3	4	5	6
2	1	1	1	1	1
3	1	1	1	1	0.9977
4	1	1	1	0.0168	0
5	1	1	0.0009	0	0
6	1	0.1606	0	0	0
7	1	0	0	0	0
8	1	0	0	0	0
9	0.0010	0	0	0	0

where $d_A = d_B = d$ and $0 \leq \varepsilon \leq 1$. When ε is sufficiently small, a random sample ρ from this ensemble almost surely has the maximum Schmidt number, that is, $\text{SN}(\rho) = d$. To compare the performance of the k -reduction criterion and CM criterion, Fig. 3 illustrates their detection ratios determined by numerical simulations as functions of k for three different noise strengths ε . Here the detection ratio of a criterion denotes the percentage of quantum states in a given ensemble that is certified to have $\text{SN}(\rho) > k$ by this criterion. The k -reduction criterion has higher detection ratios than the CM criterion in all cases under consideration. In other words, the k -reduction criterion is more effective in certifying the Schmidt numbers of pure states with depolarizing noise.

Finally, we consider random mixed states with induced measures [43]. Let $\mathcal{H}_C$ be a K -dimensional Hilbert space. Denote by $\mathcal{E}_{D,K}$ the ensemble of mixed states on $\mathcal{H}_{\text{AB}}$ induced by Haar-random pure states on $\mathcal{H}_{\text{AB}} \otimes \mathcal{H}_C$. Tables II and III summarize the detection ratios of the k -reduction criterion and CM criterion for various choices of the parameters K and k as determined by numerical simulations with $d_A = d_B = 16$. According to the two tables, the CM criterion is stronger than the k -reduction criterion for large K , in which case the states in the ensemble $\mathcal{E}_{D,K}$ tend to have low purities. In the special case $K = 2$, the two criteria have comparable performance: both of them can certify that almost all states in the ensemble $\mathcal{E}_{256,2}$ have Schmidt numbers at least 9.

TABLE III. Detection ratio of the correlation matrix criterion in Eq. (12) for the ensemble $\mathcal{E}_{256,K}$. Here $d_A = d_B = 16$.

k	K				
	2	3	4	5	6
5	1	1	1	1	1
6	1	1	1	0	0
7	1	1	0	0	0
8	1	0	0	0	0
9	0.0051	0	0	0	0

F. k -reduction Schmidt-number witnesses

Here we introduce a family of Schmidt-number- $(k + 1)$ witnesses by virtue of the k -reduction map, which will be useful for certifying the entanglement dimensionality of $(k + 1)$ -unfaithful states as shown in the next subsection.

Lemma 1. Suppose $\rho, \varrho \in \mathcal{S}(\mathcal{H}_{AB})$; then we have $\text{Tr}[\mathcal{R}_k(\rho)\varrho] = \text{Tr}[\rho\mathcal{R}_k(\varrho)]$. If $\text{Tr}[\mathcal{R}_k(\rho)\varrho] < 0$, then both ρ and ϱ have Schmidt numbers at least $k + 1$; both $\mathcal{R}_k(\rho)$ and $\mathcal{R}_k(\varrho)$ are Schmidt-number- $(k + 1)$ witnesses, which can certify ϱ and ρ , respectively.

Lemma 1 follows from the facts that the k -reduction map is self-adjoint and that $\mathcal{R}_k(\rho) \geq 0$ whenever $\rho \in \mathcal{S}_k$. Now, suppose $\mathcal{R}_k(\rho) \not\geq 0$, then any state supported in a subspace associated with negative eigenvalues of $\mathcal{R}_k(\rho)$ has Schmidt number at least $k + 1$. In particular, any eigenstate of $\mathcal{R}_k(\rho)$ with a negative eigenvalue has Schmidt number at least $k + 1$. The following theorem is a simple corollary of this observation.

Theorem 3. Suppose $\rho \in \mathcal{S}(\mathcal{H}_{AB})$, $\mathcal{R}_k(\rho) \not\geq 0$, and $|\Psi\rangle$ is an eigenstate of $\mathcal{R}_k(\rho)$ with a negative eigenvalue. Then $\mathcal{R}_k(\rho)$ and $\mathcal{R}_k(\Psi)$ are Schmidt-number- $(k + 1)$ witnesses, which can certify $|\Psi\rangle$ and ρ , respectively.

A Schmidt-number- $(k + 1)$ witness of the form $\mathcal{R}_k(\rho)$ with $\mathcal{R}_k(\rho) \not\geq 0$ is henceforth called a k -reduction Schmidt-number witness. Thanks to Lemma 1 and Theorem 3, any state ρ certifiable by the k -reduction map can also be certified by a k -reduction Schmidt-number witness. Note that the witness based on the maximally entangled state $|+_d\rangle$ in Eq. (7) can also be regarded as a $(k - 1)$ -reduction Schmidt-number witness. In addition, by applying the witness $\mathcal{R}_k(\rho)$ to the state ρ itself we can obtain a purity-based criterion

$$\text{if } \text{SN}(\rho) \leq k, \quad \text{then } \text{Tr}(\rho_A^2) \geq \frac{1}{k} \text{Tr}(\rho^2), \quad (42)$$

with $\rho_A = \text{Tr}_B(\rho)$.

G. Certification of entanglement dimensionality of k -unfaithful states

A quantum state $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ is k -unfaithful if $\text{Tr}(\rho O_W) \geq 0$ for all Schmidt-number- k witnesses of the form $O_W = cI - |\psi\rangle\langle\psi|$ [18]. By definition, any state with Schmidt number at most $k - 1$ is k -unfaithful. By contrast, a state ρ is k -faithful if there exists a Schmidt-number- k witness of the form $O_W = cI - |\psi\rangle\langle\psi|$ that satisfies $\text{Tr}(\rho O_W) < 0$; in other words, the state ρ is k -faithful iff

there exists $|\psi\rangle \in \mathcal{H}_{AB}$ with $\text{SN}(|\psi\rangle) \geq k$ such that

$$\langle\psi|\rho|\psi\rangle > \sum_{i=0}^{k-2} \sigma(\psi)_i^\downarrow. \quad (43)$$

By convention, 2-unfaithful (faithful) is also referred to as unfaithful (faithful). It is known that the state ρ is faithful iff there exist local unitary operators U_A, U_B such that [44]

$$\langle+_d| \left(U_A^\dagger \otimes U_B^\dagger \right) \rho (U_A \otimes U_B) |+_d\rangle > \frac{1}{d}. \quad (44)$$

However, a similar result does not hold when $k \geq 3$. The following proposition highlights the power of the 1-reduction map and 1-reduction Schmidt-number witnesses in certifying faithful states.

Proposition 4. Any faithful state in $\mathcal{S}(\mathcal{H}_{AB})$ can be certified by the 1-reduction map and a 1-reduction Schmidt-number witness.

Proof. Suppose $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ is faithful. Then, by Eq. (44), there exists a state $|\Psi\rangle = (U_A \otimes U_B) |+_d\rangle$ in $\mathcal{H}_{AB}$ such that

$$\langle\Psi|\mathcal{R}_1(\rho)|\Psi\rangle = \frac{1}{d} - \langle\Psi|\rho|\Psi\rangle < 0, \quad (45)$$

which implies that $\mathcal{R}_1(\rho) \not\geq 0$. So ρ can be certified by the 1-reduction map and a 1-reduction Schmidt-number witness by Theorem 3. ■

Next, we show that the 1-reduction map and 1-reduction Schmidt-number witnesses can certify unfaithful states. Suppose $\mathcal{H}_{AB}$ is a Hilbert space of local dimensions $d_A = d_B = d$, $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ is a Bell state in $\mathcal{H}_{AB}$, and

$$\rho = x|\Phi^+\rangle\langle\Phi^+| + (1-x)\frac{I}{d^2}, \quad x \in [0, 1]. \quad (46)$$

Then the fidelity between ρ and any maximally entangled state in $\mathcal{H}_{AB}$ is upper bounded by

$$\frac{2x}{d} + \frac{1-x}{d^2}. \quad (47)$$

Therefore, ρ is unfaithful when $x \leq (d - 1)/(2d + 1)$ according to Eq. (44). In addition, the minimum eigenvalue

of $\mathcal{R}_1(\rho)$ reads

$$-\frac{1}{2}x + \frac{(1-x)(d-1)}{d^2}. \quad (48)$$

If $x > 2/(d+2)$, then this eigenvalue is negative, which means $\mathcal{R}_1(\rho) \not\geq 0$. Now, suppose $d > 4$ and

$$x \in \left(\frac{2}{d+2}, \frac{d-1}{2d+1} \right]; \quad (49)$$

then the state ρ is unfaithful but certifiable by the 1-reduction map and some 1-reduction Schmidt-number witnesses.

In general, k -unfaithfulness can be certified via semidefinite programming [18]. To be specific, a state $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ is k -unfaithful if there exist positive operators $E_A \in \mathcal{L}(\mathcal{H}_A)$, $E_B \in \mathcal{L}(\mathcal{H}_B)$ and non-negative numbers μ_A, μ_B fulfilling the following conditions:

$$\begin{aligned} E_A \otimes I_B + I_A \otimes E_B &\geq \rho, \\ \mu_A + \mu_B &= 1, \\ \text{Tr}(E_A) &= \mu_A(k-1), \\ \text{Tr}(E_B) &= \mu_B(k-1), \\ E_A &\leq \mu_A I_A, \quad E_B \leq \mu_B I_B. \end{aligned} \quad (50)$$

For instance, the following state is 3-unfaithful:

$$\begin{aligned} \rho_{\text{UF}} &\equiv \frac{1}{2}|+_3\rangle\langle+_3| + \frac{1}{2}|\Phi_2\rangle\langle\Phi_2|, \\ |\Phi_2\rangle &\equiv \frac{1}{\sqrt{2}}(|23\rangle + |32\rangle). \end{aligned} \quad (51)$$

As a short proof, the conditions in Eq. (50) can be fulfilled by the following choices:

$$\mu_A = \mu_B = \frac{1}{2}, \quad E_A = E_B = \frac{1}{4} \sum_{i=0}^3 |i\rangle\langle i|. \quad (52)$$

Moreover, the state ρ_{UF} can be certified by a 2-reduction Schmidt-number witness introduced in Sec. III F. Direct calculation shows that $\mathcal{R}_2(\rho_{\text{UF}})$ has a unique negative eigenvalue, and the corresponding eigenstate reads

$$\begin{aligned} |\Psi\rangle &= \alpha|00\rangle + \alpha|11\rangle + \sqrt{1-2\alpha^2}|22\rangle, \\ \alpha &= \frac{\sqrt{3} + \sqrt{2}}{2\sqrt{3} + \sqrt{6}} \approx 0.67. \end{aligned} \quad (53)$$

In addition, $\text{Tr}[\rho_{\text{UF}} \mathcal{R}_2(\Psi)] = \text{Tr}[\mathcal{R}_2(\rho)|\Psi\rangle\langle\Psi|] < 0$, so the 2-reduction Schmidt-number witness $\mathcal{R}_2(\Psi)$ can certify that the state ρ has Schmidt number at least 3. In experiments, the expectation value of the witness $\mathcal{R}_2(\Psi)$ can be determined by measuring $\langle\Psi|\rho_{\text{UF}}|\Psi\rangle$ and $\text{Tr}[\text{Tr}_B(\Psi)\text{Tr}_B(\rho_{\text{UF}})]$.

IV. SCHMIDT-NUMBER CERTIFICATION VIA k -REDUCTION MOMENTS

In this section we systematically develop *the moment method*, which is designed to characterize the subset of positive operators within a given set of Hermitian operators, and apply this method to certifying the Schmidt number.

A. PT moments and k -reduction moments

Here we give an overview of how the moment method is utilized in entanglement certification [9], which has connection but is different from the method in Ref. [45].

The well-known PPT criterion states that if the partial transpose $\rho^\top$ (with respect to either party) of a bipartite state ρ is not a positive operator, then ρ is necessarily entangled. This criterion is proved to be useful for a large class of quantum states. However, noncomplete positive maps, including the partial transpose and k -reduction map, are nonphysical. We cannot implement them directly in the laboratory. One solution to this problem is quantum tomography. After reconstructing the entire density matrix from the outcomes of experimental measurements, we can directly test the PPT criterion or k -reduction criterion using a classical computer. Unfortunately, full quantum state tomography requires a huge measurement budget [46–48], which is too prohibitive for practical applications.

Recently, researchers have found more resource-efficient variants of the PPT criterion [13, 49, 50]. To apply the PPT criterion, we need to determine whether the spectrum of $\rho^\top$ contains a negative eigenvalue, which can be certified by the partial transpose (PT) moments,

$$p_n^\top \equiv \text{Tr}[(\rho^\top)^n], \quad n = 1, 2, \dots \quad (54)$$

The criterion that employs the first n PT moments is referred to as the $p_n^\top$ -PPT criterion [9–11]. For example, the $p_3^\top$ -PPT criterion states that

$$\text{if } \rho \text{ is separable, then } p_3^\top \leq (p_2^\top)^2. \quad (55)$$

In practice, $p_n^\top$ can be estimated efficiently using randomized measurements [7].

Likewise, we can devise protocols for certifying the Schmidt number using the following sequence of k -reduction moments:

$$Q \equiv (q_n)_{n \in \mathbb{N}_0}, \quad q_n \equiv \text{Tr}[\mathcal{R}_k(\rho)^n]. \quad (56)$$

According to Theorem 1, if ρ is a pure state with Schmidt vector $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_{d-1})$, then

$$q_n = \text{Tr}[\Omega_k(\lambda)^n] + (d_B - 1)k^n \sum_{i=0}^{d-1} \lambda_i^n, \quad (57)$$

where $\Omega_k(\lambda)$ is defined in Eq. (27). We can certify that $\text{SN}(\rho) > k$ if this moment sequence satisfies certain conditions as detailed below.

B. The moment method

Here we recapitulate several standard results on the moment method [30]. Denote by $\mathbb{N}$ the set of natural numbers (positive integers) and by $\mathbb{N}_0$ the set of non-negative integers.

The *moment problem* concerns the following question: given a real sequence $S = (s_n)_{n \in \mathbb{N}_0}$ and a closed subset $\mathbb{S} \subset \mathbb{R}$, when does there exist a Radon measure μ such that

$$s_n = \int_{\mathbb{S}} x^n d\mu(x) \quad \forall n \in \mathbb{N}_0. \quad (58)$$

Here the integration can be replaced by a summation when the measure μ is supported on discrete points. If such a measure exists, then μ is called a *representing measure* of S . The moment method refers to a systematic approach for solving the moment problem. The sequence S is called an $[a, b]$ -*moment sequence* if $\mathbb{S} = [a, b]$. Accordingly, the moment problem is called an $[a, b]$ -moment problem. When the sequence is finite, the corresponding moment problem is called a *truncated $[a, b]$ -moment problem*. Given $N, N_1, N_2 \in \mathbb{N}_0$ with $N_2 \geq N_1$, we shall use the following notation to represent a finite sequence truncated from an infinite sequence S :

$$S_{N_1, N_2} \equiv (s_{N_1}, s_{N_1+1}, \dots, s_{N_2}), \quad S_N \equiv S_{0, N}. \quad (59)$$

The arithmetic operations between sequences are defined by the corresponding operations on each entry,

$$(S \pm S')_i \equiv S_i \pm S'_i. \quad (60)$$

An important tool for solving the moment problem is the *Hankel matrix*. Given a non-negative integer n and a finite sequence S_{2n} , the Hankel matrix $H(S_{2n})$ is defined as follows:

$$[H(S_{2n})]_{ij} \equiv s_{i+j} \quad i, j = 0, 1, \dots, n. \quad (61)$$

By definition, $H(S_{2n})$ is an $(n+1) \times (n+1)$ real symmetric matrix. The positivity of a Hankel matrix is determined by the corresponding moment sequence. The following lemma is a standard result on the moment problem.

Lemma 2 (Theorems 10.1 and 10.2 in Ref. [30]). Suppose $N \in \mathbb{N}$ is even. Then S_N is a truncated $[a, b]$ -moment

sequence iff

$$H(S_N) \geq 0, \quad H(\bar{S}_{N-2}) \geq 0, \quad (62)$$

where

$$\bar{S}_N \equiv (a+b)S_{1, N+2} - S_{2, N+3} - abS_{N+1}. \quad (63)$$

Suppose $N \in \mathbb{N}$ is odd. Then S_N is a truncated $[a, b]$ -moment sequence iff

$$H(S_{1, N} - aS_{N-1}) \geq 0, \quad H(bS_{N-1} - S_{1, N}) \geq 0. \quad (64)$$

C. k -reduction moment criteria

To formulate the certification problem as a moment problem, we need to delve into the moment sequence $Q = (q_n)_{n \in \mathbb{N}_0}$ generated by the moments of the k -reduced operator $\mathcal{R}_k(\rho)$ as defined in Eq. (56). Thanks to Corollary 1, this sequence can be regarded as a $[-1, k]$ -moment sequence, so Q_N is a truncated $[-1, k]$ -moment sequence. If in addition $\mathcal{R}_k(\rho) \geq 0$, then Q_N is a truncated $[0, k]$ -moment sequence. Now, it remains to address the technical problem: given a truncated $[-1, k]$ -moment sequence, when can we conclude that it is not a truncated $[0, k]$ -moment sequence?

Define

$$B_N[\rho, k] \equiv \begin{cases} H(Q_{1, N}), & N \text{ odd}, \\ H(kQ_{1, N-1} - Q_{2, N}), & N \text{ even}, \end{cases} \quad (65)$$

which can be abbreviated as B_N if there is no danger of confusion. The following theorem proved in Appendix B 2 is the basis of our approach for certifying the entanglement dimensionality.

Theorem 4. Suppose $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ and $\text{SN}(\rho) \leq k$. Then $B_N[\rho, k] \geq 0$ for all N .

Remark 1. Note that the condition $B_N \geq 0$ comes directly from the standard result Lemma 2, but only contains the second condition in Eq. (62) and the first one in Eq. (64). The condition $H(Q_N) \geq 0$ holds for any moment sequence of the form in Eq. (56), and the condition $H(kQ_{N-1} - Q_{1, N}) \geq 0$ holds automatically because $\mathcal{R}_k(\rho) \leq kI$ by Corollary 1.

If $B_N \not\geq 0$ for some positive integer N , then we can conclude that Q_N is not a truncated $[0, k]$ -moment sequence, which means $\mathcal{R}_k(\rho) \not\geq 0$ and $\text{SN}(\rho) > k$. In this way, we can construct a sequence of criteria for certifying the Schmidt number. The N th order moment-based k -reduction criterion can be formulated as follows:

$$\text{if } \text{SN}(\rho) \leq k, \quad \text{then } B_N[\rho, k] \geq 0. \quad (66)$$

Note that k -reduction moments are invariant under local unitary transformations, so our criteria are also invariant

ALGORITHM 1. Certification of Schmidt number.

Input: ρ , N^* , and k .
Output: if return yes, then $\text{SN}(\rho) \geq k$.

- 1: Let $q_1 = (k-1)d_B - 1$.
- 2: Estimate the second moment q_2 of $\mathcal{R}_{k-1}(\rho)$.
- 3: **for** $N = 3, \dots, N^*$ **do**
- 4: Estimate the N -th moment q_N of $\mathcal{R}_{k-1}(\rho)$.
- 5: Construct the Hankel matrix $B_N[\rho, k-1]$ with $\{q_n\}_{n=1}^N$.
- 6: If $B_N[\rho, k-1] \not\geq 0$, return yes.
- 7: **end for**
- 8: Return no.

under local unitary transformations, in sharp contrast with fidelity-based methods.

The conditions $B_1 \geq 0, B_2 \geq 0$ are trivial because $B_1 = q_1$ and $B_2 = kq_1 - q_2$ contain only one entry each, so we usually start with $N = 3$. The simplest odd-order condition is thus

$$B_3 = \begin{pmatrix} q_1 & q_2 \\ q_2 & q_3 \end{pmatrix} \geq 0, \quad (67)$$

and the simplest even-order condition is

$$B_4 = \begin{pmatrix} kq_1 - q_2 & kq_2 - q_3 \\ kq_2 - q_3 & kq_3 - q_4 \end{pmatrix} \geq 0. \quad (68)$$

Note that B_N is a principal submatrix of B_{N+2} , which means $B_N \geq 0$ whenever $B_{N+2} \geq 0$, so the $(N+2)$ th order moment-based k -reduction criterion is usually stronger than the N th order moment-based criterion.

By virtue of Theorem 4, we can devise a simple algorithm (Algorithm 1) for certifying whether $\text{SN}(\rho) \geq k$. Here we set a truncation value for N because it is in general more difficult to determine higher-order moments accurately. On this basis, we can further devise an algorithm (Algorithm 2) for constructing the best lower bound for the Schmidt number $\text{SN}(\rho)$ based on k -reduction moments.

 ALGORITHM 2. Optimal certification of Schmidt number by the N^* th order moment.

Input: ρ , N^* , and r .
Output: a lower bound for $\text{SN}(\rho)$.

- 1: $s = 1$.
- 2: **for** $k = 2, \dots, r$ **do**
- 3: Input (ρ, N^*, k) to Algorithm 1.
- 4: **if** output is no, **then**
- 5: Return s .
- 6: **else** $s = k$.
- 7: **end if**
- 8: **end for**
- 9: Return s .

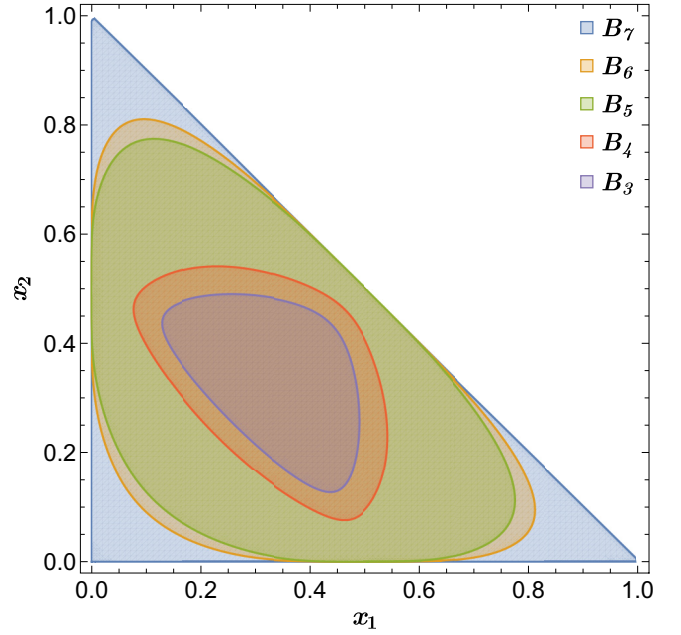


FIG. 4. Detectable regions of k -reduction moment criteria with $k = 2$. Each point represents a state with Schmidt number 3 in the form of Eq. (69). The region of states whose Schmidt numbers can be detected by the N th order moment-based criterion (that is, the set of states satisfying $B_N[\psi, 2] \geq 0$) is illustrated in the plot. The higher the order is, the larger the detectable region becomes. When $N = 7$, almost all states can be detected.

To illustrate the detection capabilities of moment-based k -reduction criteria, it is instructive to consider the following family of two-qutrit pure states:

$$\sqrt{x_1}|00\rangle + \sqrt{x_2}|11\rangle + \sqrt{1-x_1-x_2}|22\rangle, \quad (69)$$

where $x_1, x_2 > 0$ and $x_1 + x_2 < 1$; all these states have Schmidt numbers equal to 3. Figure 4 illustrates the detectable region of the N th order moment-based k -reduction criterion with $k = 2$ and $N = 3, 4, 5, 6, 7$. Note that the detection capability gets stronger and stronger as N increases. When $N = 7$, the moment-based criterion can certify the Schmidt numbers of almost all two-qutrit pure states.

Next, we show that the N th order moment-based k -reduction criterion is equivalent to the original k -reduction criterion when N is sufficiently large. The following theorem is proved in Appendix B 2.

Theorem 5. Suppose $\rho \in \mathcal{S}(\mathcal{H}_{AB})$, $\mathcal{R}_k(\rho)$ has χ distinct nonzero eigenvalues, and $N \geq 2\chi - 1$. Then $B_N[\rho, k] \geq 0$ iff $\mathcal{R}_k(\rho) \geq 0$.

Theorem 5 applies to arbitrary quantum states in $\mathcal{S}(\mathcal{H}_{AB})$. See Sec. IV E for additional results on pure states.

D. The third order moment-based criterion

In this subsection, we provide more details on the third order moment-based k -reduction criterion, that is, $N = 3$. Note that $q_1 = kd_B - 1 > 0$, so $B_3 \geq 0$ iff $\det(B_3[\rho, k]) \geq 0$. Accordingly, the criterion in Eq. (66) can be simplified as follows:

$$\text{if } \text{SN}(\rho) \leq k, \quad \text{then } \det(B_3[\rho, k]) \geq 0. \quad (70)$$

Define

$$\begin{aligned} p_n &\equiv \text{Tr}(\rho^n), \quad a_n \equiv \text{Tr}(\rho_A^n), \\ t_2 &\equiv \text{Tr}[\rho_A \text{Tr}_B(\rho^2)]. \end{aligned} \quad (71)$$

Then $\det(B_3[\rho, k])$ can be expressed as a polynomial of k as follows:

$$\det(B_3[\rho, k]) = \beta_4 k^4 + \beta_3 k^3 + \beta_2 k^2 + \beta_1 k + \beta_0, \quad (72)$$

where

$$\begin{aligned} \beta_4 &= d_B^2(a_3 - a_2^2), \\ \beta_3 &= -4d_B(a_3 - a_2^2), \\ \beta_2 &= d_B(3t_2 - 2a_2p_2) - 4a_2^2 + 3a_3, \\ \beta_1 &= 4a_2p_2 - d_Bp_3 - 3t_2, \\ \beta_0 &= p_3 - p_2^2. \end{aligned} \quad (73)$$

Hence, to apply the criterion in Eq. (70), it remains to estimate the moments p_2, p_3, a_2, a_3, t_2 .

E. k -reduction moment criteria for pure states

Recall that the k -reduction criterion can certify the Schmidt number of any pure state. Here we clarify the performance of moment-based k -reduction criteria. The following theorem proved in Appendix C 1 shows that the N th order moment-based k -reduction criterion can certify the Schmidt number of any pure state in $\mathcal{H}_{AB}$ when $N \geq 4d - 1$, where $d = \min\{d_A, d_B\}$. In addition, it is necessary to employ the N th order moment-based k -reduction criterion with $N \geq 2d$ for certain pure states in $\mathcal{H}_{AB}$.

Theorem 6. Suppose $|\psi\rangle \in \mathcal{H}_{AB}$ has Schmidt rank r and $\tilde{r}$ distinct nonzero Schmidt coefficients denoted by $\{\ell_j\}_{j=0}^{\tilde{r}-1}$. If $N \geq 4\tilde{r} - 1$, then the N th order moment criterion can detect the state $|\psi\rangle$ as

$$B_N[\psi, k] \not\geq 0 \quad \forall k = 1, 2, \dots, r-1. \quad (74)$$

Let $L = (l_n)_{n \in \mathbb{N}_0}$, $l_n = \sum_{j=0}^{\tilde{r}-1} \ell_j^n$, and

$$A_0 \equiv \begin{cases} H(L_{1,N}), & N \text{ odd}, \\ H[(r-1)L_{1,N-1} - L_{2,N}], & N \text{ even}. \end{cases} \quad (75)$$

If instead $N \leq 2\tilde{r}$ and

$$d_B > 1 + \frac{r\tilde{r}\mathcal{N}_{r-1}(\psi)}{\sigma_{\min}(A_0)}, \quad (76)$$

then the N th order moment criterion cannot detect the state $|\psi\rangle$ as

$$B_N[\psi, r-1] \geq 0. \quad (77)$$

When $|\psi\rangle = |+_r\rangle$ is the maximally entangled state of Schmidt rank r , we have $\tilde{r} = 1$. In this case, it suffices to apply the third order moment-based k -reduction criterion because Theorem 6 implies the following proposition.

Proposition 5. Given $|+_r\rangle$ defined in Eq. (4) and $1 \leq k < r$, we have $B_N[+_r, k] \not\geq 0$ for all $N \geq 3$.

F. Certification of entanglement dimensionality of isotropic states

Given a Hilbert space $\mathcal{H}_{AB}$ with $d_A = d_B = d$, an isotropic state on $\mathcal{H}_{AB}$ is a state of the form

$$\rho_F \equiv \frac{1-F}{d^2-1}I + \frac{d^2F-1}{d^2-1}|+_d\rangle\langle+_d|, \quad F \in [0, 1]. \quad (78)$$

It has been proved that $\text{SN}(\rho_F) = \lceil dF \rceil$, and the k -reduction map with $k = \lceil dF \rceil - 1$ can certify the isotropic state [15]. In the next proposition, we prove that ρ_F can be certified by fidelity-based method as well.

Proposition 6. The isotropic state ρ_F is $\lceil dF \rceil$ -faithful.

Proof. Note that

$$\text{SN}(|+_d\rangle) = d \geq \lceil dF \rceil, \quad (79)$$

$$\langle+_d|\rho_F|+_d\rangle = F > \frac{1}{d}(\lceil dF \rceil - 1), \quad (80)$$

$$\sigma_i^\downarrow(+_d) = \frac{1}{d} \quad \forall i. \quad (81)$$

In conjunction with the sufficient condition of k -faithfulness in Eq. (43), the proposition is proved. ■

Here we show that the third order moment-based criterion is sufficient for certifying all isotropic states as well. Meanwhile, the CM criterion and its moment-based variants in Eqs. (16) and (19) are equally effective. Propositions 7 and 8 below are proved in Appendices C and D, respectively.

Proposition 7. Suppose ρ_F is the isotropic state defined in Eq. (78) and $1 \leq k \leq \lceil dF \rceil - 1$. Then $B_3[\rho_F, k] \not\geq 0$.

Proposition 8. Suppose ρ_F is the isotropic state defined in Eq. (78), and T is its CM with respect to given local operator bases that satisfy the conditions in Eq. (9) in Sec. II B. Then all $d^2 - 1$ singular values of T are equal to $(d^2 F - 1)/[d(d^2 - 1)]$ and

$$\begin{aligned} \|T\|_1 &= dF - d^{-1}, \quad \|T\|_2^2 = \frac{(d^2 F - 1)^2}{d^2(d^2 - 1)}, \\ \|T\|_4^4 &= \frac{(d^2 F - 1)^4}{d^4(d^2 - 1)^3}. \end{aligned} \quad (82)$$

In conjunction with the fact $\text{SN}(\rho_F) = \lceil dF \rceil$ and Eq. (12) we can conclude that the CM criterion can certify the Schmidt number of any isotropic state as well. Even if only the second and fourth moments are available, we can still deduce that $\|T\|_1 \geq \|T\|_2^3 / \|T\|_4^2 = dF - d^{-1}$. Therefore, the moment-based criterion in Eq. (19) is equally effective, and so is the alternative in Eq. (16).

G. Certification of entanglement dimensionality of k -unfaithful states

In this subsection, we show that the k -reduction moment method can certify entanglement dimensionality of some $(k + 1)$ -unfaithful states.

First, we construct a family of k -unfaithful states on $\mathcal{H}_{AB}$, assuming that $d_A = d_B = d \geq k$. Let χ be a positive integer that satisfies $1 \leq \chi \leq \min\{d, 2k - 2\}$ and let π and τ be two permutations on $\{0, \dots, d - 1\}$. Consider the following state on $\mathcal{H}_{AB}$:

$$\begin{aligned} \rho &= \frac{1}{2} |+_k\rangle \langle +_k| + \frac{1}{2} |\Phi\rangle \langle \Phi|, \\ |\Phi\rangle &= \sum_{j=0}^{\chi-1} \sqrt{\phi_j} |\pi(j)\rangle \otimes |\tau(j)\rangle, \end{aligned} \quad (83)$$

where $\sum_{j=0}^{\chi-1} \phi_j = 1$. Note that the state ρ can be regarded as a generalization of ρ_{UF} in Eq. (51). The following proposition proved in Appendix E is helpful for constructing k -unfaithful states of the form Eq. (83) that have Schmidt number at least k .

Proposition 9. Suppose the state $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ has the form Eq. (83), where $1 \leq \chi \leq \min\{d, 2k - 2\}$, $\pi(j) \neq \tau(j)$ for each $j = 0, \dots, \chi - 1$, and the following index sets

$$\begin{aligned} \mathcal{I}_L &\equiv \{j\}_{j=0}^{\chi-1} \cup \{\pi(j)\}_{j=0}^{\chi-1}, \\ \mathcal{I}_R &\equiv \{j\}_{j=0}^{\chi-1} \cup \{\tau(j)\}_{j=0}^{\chi-1} \end{aligned} \quad (84)$$

satisfy $|\mathcal{I}_L|, |\mathcal{I}_R| \leq 2k - 2$. Then ρ is k -unfaithful. If in addition

$$\sum_{i=0}^{k-1} \frac{1}{1 + k\phi_{\pi^{-1}(i)}} > k - 1, \quad (85)$$

then $\mathcal{R}_{k-1}(\rho) \not\geq 0$.

Now, suppose $k \geq 3$, $d > 2k - 2$, $\chi = 2$, $\phi_j = 1/2$ for $j = 0, 1$, and

$$\pi(j) = k - 1 + j, \quad \tau(j) = k - j \quad \forall j, \quad (86)$$

where the arithmetic is modulo d . Then the state ρ in Eq. (83) simplifies to

$$\begin{aligned} \rho_{UF,k} &\equiv \frac{1}{2} |+_k\rangle \langle +_k| + \frac{1}{2} |\Phi_k^+\rangle \langle \Phi_k^+|, \\ |\Phi_k^+\rangle &\equiv \frac{1}{\sqrt{2}} (|k - 1, k\rangle + |k, k - 1\rangle), \end{aligned} \quad (87)$$

and it satisfies all the conditions in Proposition 9 given that

$$\sum_{j=0}^{k-1} \frac{1}{1 + k\phi_{j+1-k}} = k - 1 + \frac{1}{1 + k/2} > k - 1. \quad (88)$$

So $\rho_{UF,k}$ has Schmidt number k but is k -unfaithful.

Next, we use Algorithm 2 to certify the Schmidt number of the state $\rho_{UF,k}$ with $d_A = d_B = 16$. In Table IV we summarize the $(k - 1)$ -reduction negativity of the state $\rho_{UF,k}$ and the minimum moment order N such that $B_N[\rho_{UF,k}, k - 1] \not\geq 0$ for $k = 3, 4, \dots, 10$. These results highlight the versatility of k -reduction moments in certifying the entanglement dimensionality of k -unfaithful states in addition to k -faithful states.

TABLE IV. Certification of the entanglement dimensionality of the k -unfaithful state $\rho_{UF,k}$ using k -reduction moments. The second column shows the $(k - 1)$ -reduction negativity of $\rho_{UF,k}$, and the third column shows the minimum order N such that $B_N[\rho_{UF,k}, k - 1] \not\geq 0$, which can be determined by virtue of Algorithm 2.

k	$\mathcal{N}_{k-1}(\rho_{UF,k})$	Minimum order
3	0.0749	7
4	0.0449	7
5	0.0301	7
6	0.0216	7
7	0.0163	9
8	0.0128	9
9	0.0103	9
10	0.0085	9

V. NUMERICAL RESULTS ON DETECTION RATIOS

To complement the theoretical analysis in Sec. IV, here we provide extensive numerical results on the detection ratios of the N th order moment-based k -reduction criterion in Eq. (66) and moment-based CM criterion in Eq. (16). Given an ensemble $\mathcal{E}$ of quantum states, the *detection ratio* of the N th order moment-based k -reduction criterion is defined as

$$\Pr \{B_N[\rho, k] \not\geq 0 : \rho \sim \mathcal{E}\}. \quad (89)$$

Similarly, the detection ratio of the moment-based CM criterion is defined as

$$\Pr \{(\|T\|_2^2, \|T\|_4^4) \notin \mathcal{T}_k : \rho \sim \mathcal{E}\}. \quad (90)$$

Not surprisingly, the detection ratios may strongly depend on the ensemble $\mathcal{E}$ under consideration. Nevertheless, for any given ensemble, a higher detection ratio is an indication of a higher performance. In the rest of this section we shall consider several concrete ensembles of quantum states, including the ensemble of pure states with a fixed Schmidt number, the ensemble of Haar-random pure states with depolarizing noise, and the ensemble of mixed states with an induced measure.

A. Ensemble of pure states with a fixed Schmidt number

Define the ensemble

$$\mathcal{E}_r \equiv \left\{ |\psi\rangle = \sum_{i=0}^{r-1} \sqrt{\lambda_i} |i\rangle_A \otimes |i\rangle_B \right\}, \quad (91)$$

where $\{|i\rangle_A\}, \{|i\rangle_B\}$ are elements of two orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, and $\{\lambda_i\}_{i=0}^{r-1}$ is uniformly distributed in the interior of a $(r-1)$ -dimensional probability simplex [cf. Eq. (3)], which is also known as the Dirichlet distribution. By definition, all states in $\mathcal{E}_r$ have Schmidt numbers equal to r . Here it is not necessary to consider local unitary transformations because the moment-based k -reduction criteria and CM criterion are both invariant under local unitary transformations. In numerical simulations we choose $d = d_A = d_B$.

Figures 5 and 6 illustrate the detection ratio of the N th order moment-based k -reduction criterion in Eq. (66) as a function of N and k for the ensemble $\mathcal{E}_6$ with two different local dimensions, namely, $d = 8$ and $d = 16$. As expected the detection ratio decreases monotonically with k , but increases monotonically with the order N and eventually approaches 1 when N is sufficiently large.

Now, we compare the detection ratios of the N th order moment-based k -reduction criterion in Eq. (66) and

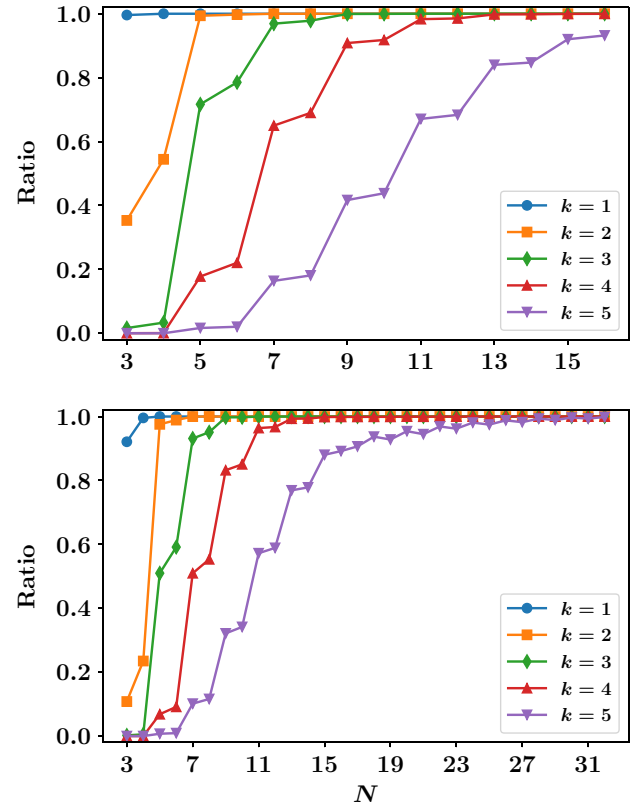


FIG. 5. Detection ratio of the N th order moment-based k -reduction criterion in Eq. (66) for the pure state ensemble $\mathcal{E}_6$ in Eq. (91) as a function of N . Here $d = 8$ (upper plot) and $d = 16$ (lower plot). The detection ratio increases monotonically with N , but decreases monotonically with k .

moment-based CM criterion in Eq. (16). Figure 7 illustrates the results on the ensemble $\mathcal{E}_6$ as functions of the order N and local dimension d , where $k = 5$. According to this figure, the detection ratio of the moment-based CM criterion is almost independent of the local dimension, but the detection ratios of moment-based k -reduction criteria tend to decrease with the local dimension. Nevertheless, the N th order moment-based k -reduction criterion has a larger detection ratio when N is sufficiently large.

B. Ensemble of Haar-random pure states with depolarizing noise

Next, we turn to the ensemble $\mathcal{E}_{\text{Haar},\varepsilon}$ of Haar-random states with depolarizing noise as defined in Eq. (41), where the parameter ε characterizes the noise strength. Figures 8–10 illustrate the detection ratios of the N th order moment-based k -reduction criterion in Eq. (66) and moment-based CM criterion in Eq. (16) for the ensemble $\mathcal{E}_{\text{Haar},\varepsilon}$ with two different local dimensions and three different noise strengths. Not surprisingly, the detection ratios increase monotonically with N but decrease monotonically

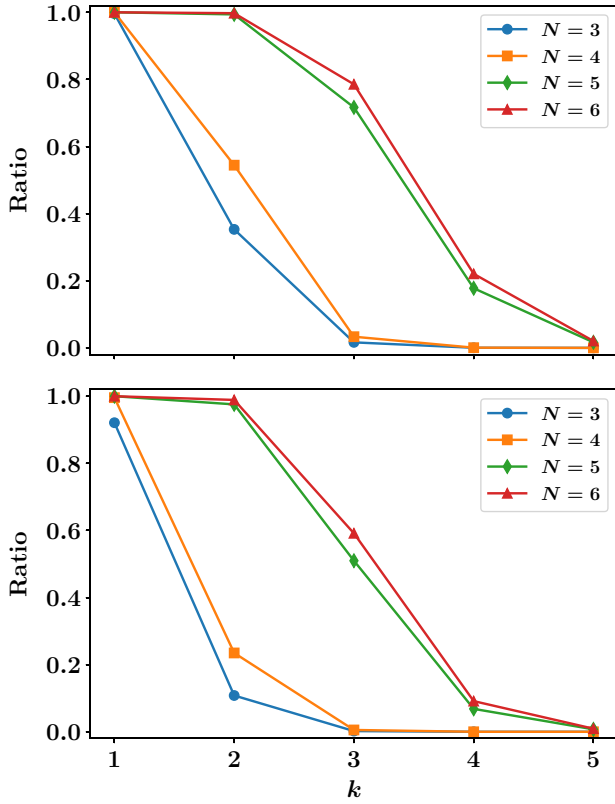


FIG. 6. Detection ratio of the N th order moment-based k -reduction criterion for the pure state ensemble $\mathcal{E}_6$ as a function of k . Here $d = 8$ (upper plot) and $d = 16$ (lower plot) as in Fig. 5.

with k and ε . When $d = 8$ and $\varepsilon = 0, 0.1$, for example, the seventh order moment-based k -reduction criterion has a larger detection ratio than the moment-based CM criterion, but the situation is different when $d = 16$.

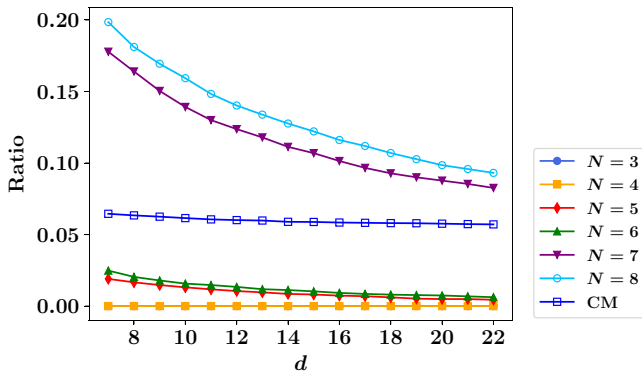


FIG. 7. Detection ratios of the N th order moment-based k -reduction criterion in Eq. (66) and moment-based correlation matrix criterion in Eq. (16) with $k = 5$ for the ensemble $\mathcal{E}_6$ in Eq. (91) with $7 \leq d \leq 22$. The N th order moment-based k -reduction criterion has a larger detection ratio than the moment-based correlation matrix criterion when $N \geq 7$.

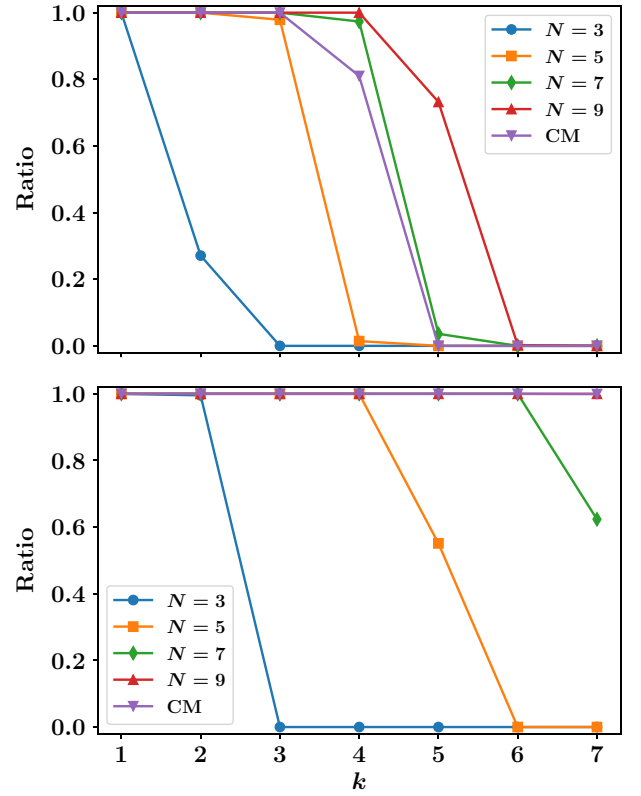


FIG. 8. Detection ratios of different k -reduction moment criteria in Eq. (66) and moment-based correlation matrix criterion in Eq. (16) for the Haar-random ensemble $\mathcal{E}_{\text{Haar},0}$ in Eq. (41). Here $d = 8$ (upper plot) and $d = 16$ (lower plot) as in Fig. 5.

C. Ensemble of mixed states with an induced measure

Finally, we consider the ensemble $\mathcal{E}_{D,K}$ of mixed states with an induced measure, which is defined in Sec. III E. In numerical simulations we choose $D = d^2 = 16^2$ and $K = 2, 3, 4, 5, 6$. Tables V and VI show the detection ratios of states with various lower bounds for the Schmidt numbers. The detection ratios in Table V are determined by moment-based k -reduction criteria described in Algorithm 2 with maximal order $N^* = 16$ and $r = 16$. The detection ratios in Table VI are determined by moment-based CM criterion in Eq. (16). From the numerical results, we can see that the k -reduction moment criteria are weaker than moment-based CM criterion for $K = 4, 5, 6$, but become stronger for $K = 3$.

VI. ESTIMATION OF k -REDUCTION MOMENTS

To certify the Schmidt number using moment-based k -reduction criteria, we need to estimate the k -reduction moments $q_n = \text{Tr}[R_k(\rho)^n]$ that appear in the entries of the Hankel matrices. Simple analysis shows that these moments can be expressed as follows:

$$\text{Tr}[\text{poly}(\rho, \rho_A \otimes I_B)]. \quad (92)$$

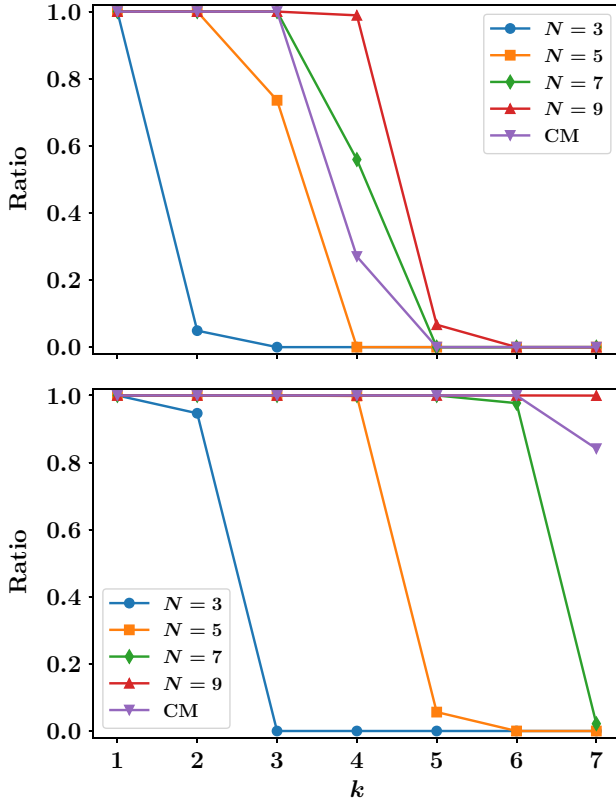


FIG. 9. Detection ratios of different k -reduction moment criteria in Eq. (66) and moment-based correlation matrix criterion in Eq. (16) for the Haar-random ensemble $\mathcal{E}_{\text{Haar},0.1}$ in Eq. (41). Here $d = 8$ (upper plot) and $d = 16$ (lower plot) as in Fig. 5.

In this section, we introduce two approaches for estimating these moments.

A. Moment estimation by randomized measurements

Here we discuss moment estimation based on the classical shadow [13], which is a prototypical method based on randomized measurements. This method consists of two steps, namely, data acquisition and prediction.

Suppose $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ and $\dim(\mathcal{H}_{AB}) = D$. In the phase of data acquisition, we first randomly sample a unitary operator U from a given unitary ensemble $\mathcal{U}$ on the Hilbert space $\mathcal{H}_{AB}$ and apply it to the state ρ under consideration. Then we perform a measurement in the computational basis and record the measurement outcome b . Based on U and b , we can create a classical shadow of the unknown state as follows:

$$\hat{\rho} = \mathcal{M}_{\text{cs}}^{-1}(U^\dagger |b\rangle\langle b| U). \quad (93)$$

Here $\mathcal{M}_{\text{cs}}$ is the measurement channel, which depends on the ensemble $\mathcal{U}$, and $\mathcal{M}_{\text{cs}}^{-1}$ is the inverse of $\mathcal{M}_{\text{cs}}$ and is known as the reconstruction map. If $\mathcal{U}$ corresponds to the

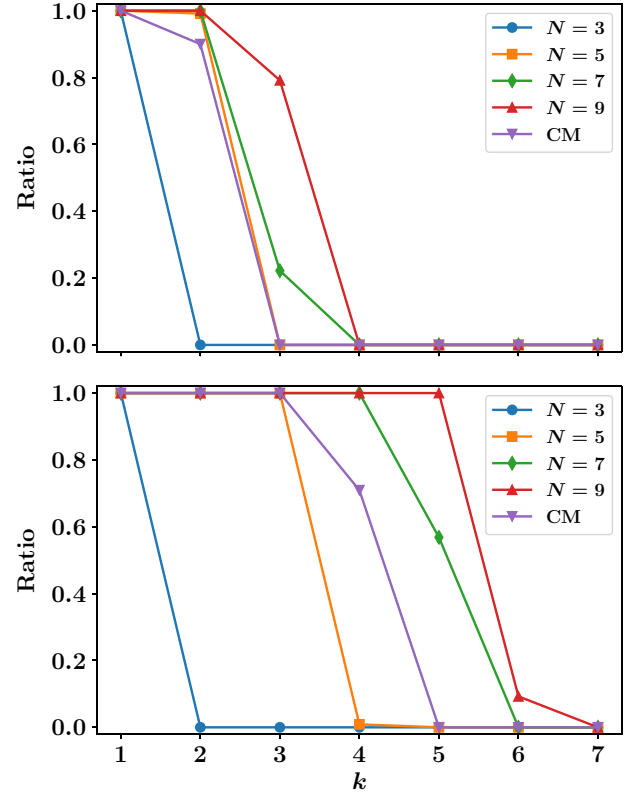


FIG. 10. Detection ratios of different k -reduction moment criteria in Eq. (66) and moment-based correlation matrix criterion in Eq. (16) for the Haar-random ensemble $\mathcal{E}_{\text{Haar},0.5}$ in Eq. (41). Here $d = 8$ (upper plot) and $d = 16$ (lower plot) as in Fig. 5.

global Clifford group, then the above equation reduces to

$$\hat{\rho} = (D+1)U^\dagger |b\rangle\langle b| U - I. \quad (94)$$

In the prediction phase, we need to estimate the expectation value of a given observable O (generalization to more observables is straightforward). Given M samples of the classical shadow we can construct an estimator based on

TABLE V. Detection ratio of the N th order moment-based k -reduction criterion in Eq. (66) with $N = 16$ (corresponding to Algorithm 2 with $N^* = 16$) for the ensemble $\mathcal{E}_{D,K}$ with $d_A = d_B = 16$ and $D = d_A d_B = 256$.

k	K				
	2	3	4	5	6
2	1	1	1	1	1
3	1	1	1	1	0.9559
4	1	1	1	0.0049	0
5	1	1	0.0001	0	0
6	1	0.0146	0	0	0
7	0.9999	0	0	0	0
8	0.0441	0	0	0	0

TABLE VI. Detection ratio of the moment-based correlation matrix criterion in Eq. (16) for the ensemble $\mathcal{E}_{D,K}$. Here $d_A = d_B = 16$ and $D = d_A d_B = 256$ as in Table V.

k	K				
	2	3	4	5	6
4	1	1	1	1	1
5	1	1	0.9337	0	0
6	0.9996	0.0016	0	0	0
7	0.0059	0	0	0	0

the empirical mean as follows:

$$\hat{o} = \frac{1}{M} \sum_{m=1}^M \text{Tr}(\hat{\rho}_m O). \quad (95)$$

The variance of this estimator is controlled by the *shadow norm* introduced in Ref. [13]. The worst-case variance is minimized when the ensemble $\mathcal{U}$ forms a unitary 3-design. Alternatively, we can employ the median-of-means to suppress the probability of large deviation.

In addition, the classical shadow method can be applied to estimating quadratic functions of ρ that have the form $o = \text{Tr}[O(\rho \otimes \rho)]$, where O is now an operator on $\mathcal{H}_{AB}^{\otimes 2}$. Here is a specific estimator [13],

$$\hat{o} = \frac{1}{M(M-1)} \sum_{m=1}^M \sum_{n=1, n \neq m}^M \text{Tr}[O(\hat{\rho}_m \otimes \hat{\rho}_n)]. \quad (96)$$

For example, the purity $\text{Tr}(\rho^2)$ can be estimated by

$$\begin{aligned} & \frac{1}{M(M-1)} \sum_{m=1}^M \sum_{n=1, n \neq m}^M \text{Tr}[\mathbb{W}_{(12)}(\hat{\rho}_m \otimes \hat{\rho}_n)] \\ &= \frac{1}{M(M-1)} \sum_{m=1}^M \sum_{n=1, n \neq m}^M \text{Tr}(\hat{\rho}_m \hat{\rho}_n), \end{aligned} \quad (97)$$

where $\mathbb{W}_{(12)}$ is the swap operator on $\mathcal{H}_{AB}^{\otimes 2}$. Other nonlinear functions can be estimated in a similar way.

Next, we turn to the estimation of reduction moments that are required for certifying the Schmidt number. The following theorem clarifies the sample complexities of estimating the moments involved in the third order moment-based k -reduction criterion; see Appendix F for a proof.

Theorem 7. Suppose $\rho \in \mathcal{S}(\mathcal{H}_{AB})$, $d_A = d_B = \sqrt{D}$, and we are given access to a unitary 3-design. Then the sample complexity of estimating the moments p_2, p_3, a_2, a_3, t_2 defined in Eq. (71) up to accuracy $\mathcal{O}(1)$ using the classical shadow method is $\mathcal{O}(D)$.

Note that the randomized measurements required to apply our certification protocol can be realized using a unitary 3-design, say the global Clifford group [32,33], which is quite appealing to practical applications. By contrast, to apply the CM protocol [22,23], it is necessary to use unitary 4-designs on $\mathcal{H}_A$ and $\mathcal{H}_B$; moreover, a unitary 8-design is required to bound the variance.

B. Moment estimation by permutation tests

Next, we show that the sample complexity of estimating the k -reduction moments can be reduced to $\mathcal{O}(1)$ if permutation tests based on collective measurements are accessible. Suppose $\rho_1, \rho_2, \dots, \rho_N \in \mathcal{S}(\mathcal{H}_{AB})$; then $\bigotimes_{n=1}^N \rho_n \in \mathcal{S}(\mathcal{H}_{AB}^{\otimes N})$. Let $\mathbb{W}_\pi$ be the unitary operator tied to the standard cyclic permutation. Add an ancillary qubit prepared in the state $|+\rangle$ and apply the controlled cyclic shift gate $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \mathbb{W}_\pi$ on the entire system then yields

$$\begin{aligned} & \frac{|0\rangle\langle 0|}{2} \otimes \bigotimes_{n=1}^N \rho_n + \frac{|1\rangle\langle 1|}{2} \otimes \mathbb{W}_\pi \bigotimes_{n=1}^N \rho_n \mathbb{W}_\pi^{-1} \\ &+ \frac{|0\rangle\langle 1|}{2} \otimes \bigotimes_{n=1}^N \rho_n \mathbb{W}_\pi^{-1} + \frac{|1\rangle\langle 0|}{2} \otimes \mathbb{W}_\pi \bigotimes_{n=1}^N \rho_n. \end{aligned} \quad (98)$$

The reduced state on the ancilla qubit reads

$$\frac{I}{2} + \frac{1}{2} \text{Tr}(\rho_1 \rho_2 \cdots \rho_N) X. \quad (99)$$

Measure the ancilla qubit in the Pauli X basis, then the probabilities of obtaining outcomes $\pm$ are given by

$$P_{\pm} = \frac{1}{2} \pm \frac{1}{2} \text{Tr}(\rho_1 \rho_2 \cdots \rho_N), \quad (100)$$

which means

$$P_+ - P_- = \text{Tr}(\rho_1 \rho_2 \cdots \rho_N). \quad (101)$$

The full protocol is termed as the *permutation test*, which leads to a simple recipe for estimating the N th moment $\text{Tr}(\rho^N)$ of a general quantum state $\rho \in \mathcal{S}(\mathcal{H}_{AB})$, and the sample complexity is independent of the system size. This method can readily be applied to estimating all the relevant parameters featuring in the moment-based k -reduction criteria.

In practice, the controlled cyclic shift gate can be implemented as follows. A controlled-SWAP gate (also known as the Fredkin gate) can be constructed using Toffoli and CNOT gates (see Ref. [51] for a detailed decomposition). Furthermore, any full permutation of N elements can be expressed as a sequence of adjacent transpositions. Denote the controlled-SWAP between the i th and $(i+1)$ th qubits

by $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \mathbb{W}_{(i,i+1)}$, then a controlled cyclic shift gate acting on N qubits is given by the product

$$\prod_{i=1}^{N-1} [|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \mathbb{W}_{(i,i+1)}]. \quad (102)$$

Incidentally, a work about the quantum advantage enabled by permutation tests can be found in Ref. [52].

VII. CONCLUSION

By combining the k -reduction map and the moment method, we proposed a simple yet powerful approach for certifying the Schmidt number via k -reduction moments. These moments can be efficiently estimated by virtue of randomized measurements based on unitary 3-designs, such as the Clifford group, which are significantly more practical than previous approaches based on 4-designs. Our approach does not require any prior assumption on the target state, is applicable to both k -faithful and k -unfaithful states, and thus has a much wider scope of applications than fidelity-based methods. The efficacy of the proposed method is substantiated by rigorous theoretical guarantees and comprehensive numerical benchmarks across a range of practically relevant settings.

When k -reduction moments with sufficiently high orders are accessible, our criteria become equivalent to the k -reduction criterion itself, enabling accurate certification of the entanglement dimensionality of all pure states. Moreover, only the first three k -reduction moments are required for certifying the entanglement dimensionality of isotropic states. In the course of study, we clarified the fundamental properties of the spectrum of the k -reduced operator and showed that the k -reduction negativity is monotonic under LOCC for pure states.

Our findings open several avenues for further exploration. First, is the k -reduction negativity monotonic under LOCC for general mixed states? Second, are all k -faithful states certifiable by the k -reduction map? Intuitively, the k -reduction criterion is more powerful than fidelity-based methods, but a strict proof is still absent. Third, how robust are the k -reduction moment criteria to statistical fluctuation? To clarify the third question would require knowledge about the perturbation theory of Hankel matrices, which we leave for future study. Finally, can we construct better criteria for certifying the entanglement dimensionality by virtue of other k -positive maps in conjunction with the moment method? We anticipate that these questions can stimulate further progress on the study of high-dimensional entanglement.

Note added. After posting this paper, a recent paper working on a similar topic [53] was posted, which considered other k -positive maps and the application of Schmidt-number certification on the task of quantum channel discrimination.

ACKNOWLEDGMENTS

We thank Xiaodong Yu, Zihao Li, Datong Chen, Zhenhuan Liu, and Shuheng Liu for helpful discussions. This work is supported by Shanghai Science and Technology Innovation Action Plan (Grant No. 24LZ1400200), the National Key Research and Development Program of China (Grant No. 2022YFA1404204), Shanghai Municipal Science and Technology Major Project (Grant No. 2019SHZDZX01), and National Natural Science Foundation of China (Grant No. 92165109).

DATA AVAILABILITY

The data that support the findings of this article are openly available [54].

APPENDIX A: PROOFS OF THEOREMS 1, 2 AND PROPOSITION 3

In this appendix and the following appendices we prove Theorems 1, 2, 4–7 and Propositions 3, 7–9 presented in the main text.

1. Auxiliary lemmas

Here we introduce two auxiliary lemmas to clarify the basic properties of the operator $\Omega_k(\lambda)$ defined in Eq. (27) and the function $\theta_k(\lambda)$ defined Eq. (28). Recall that Δ_d is the $(d-1)$ -dimensional probability simplex defined in Eq. (3).

Lemma A1. Suppose k is a positive integer and $\lambda \in \Delta_d$. Then $\Omega_k(\lambda)$ has at most one negative eigenvalue, and the function $\theta_k(\lambda)$ is Schur concave in λ and nonincreasing in k . Let r be the number of positive entries in λ , then

$$\text{rank}(\Omega_k(\lambda)) = \begin{cases} r-1, & k=r, \\ r, & k \neq r. \end{cases} \quad (A1)$$

If $k \geq r$, then $\Omega_k(\lambda) \geq 0$ and $\theta_k(\lambda) = 0$. If $k < r$, then $-\theta_k(\lambda)$ is the unique negative eigenvalue of $\Omega_k(\lambda)$, and $\theta_k(\lambda)$ satisfies the following equation:

$$\sum_{i=0}^{d-1} \frac{\lambda_i}{k\lambda_i + \theta_k(\lambda)} = 1. \quad (A2)$$

Proof. Let $\tilde{\lambda}$ be the vector obtained from λ by deleting the zero entries. Then $\Omega_k(\lambda)$ and $\Omega_k(\tilde{\lambda})$ have the same nonzero eigenvalues, including the multiplicities, which means $\theta_k(\lambda) = \theta_k(\tilde{\lambda})$. In addition, $\Omega_k(\lambda)$ and $\theta_k(\lambda)$ are continuous in λ . Therefore, it suffices to prove Lemma A1 under the assumption $\lambda > 0$, which means $\lambda_0, \lambda_1, \dots, \lambda_{d-1} > 0$ and $r = d$.

Let $\Lambda = \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_{d-1})$, then $\Omega_k(\lambda) = \Lambda^{1/2} \Omega' \Lambda^{1/2}$, where Ω' is the $d \times d$ matrix with entries

$\Omega'_{ij} = k\delta_{ij} - 1$. Since Λ is an invertible real diagonal matrix, Ω and Ω' have the same rank and the same number of positive (negative) eigenvalues. It is straightforward to verify that Ω' has $r - 1$ eigenvalues equal to k and one eigenvalue equal to $k - r$, which implies Eq. (A1). If $k \geq r = d$, then $\Omega_k(\lambda) \geq 0$ and $\theta_k(\lambda) = 0$. If $k < r$, then $\Omega_k(\lambda)$ has exactly one negative eigenvalue, which is equal to the smallest eigenvalue $-\theta_k$.

Suppose $k < r$ and $\mathbf{y} = (y_0, y_1, \dots, y_{d-1})^\top$ is an eigenvector of $\Omega_k(\lambda)$ associated with the eigenvalue $-\theta_k$. Then

$$k\lambda_i y_i - \sum_j \sqrt{\lambda_i \lambda_j} y_j = -\theta_k y_i \quad \forall i, \quad (\text{A3})$$

which implies that

$$y_i = \frac{\sqrt{\lambda_i}}{k\lambda_i + \theta_k} \left(\sum_{j=0}^{d-1} \sqrt{\lambda_j} y_j \right) \quad \forall i. \quad (\text{A4})$$

Note that $\sum_{j=0}^{d-1} \sqrt{\lambda_j} y_j \neq 0$; otherwise, $\mathbf{y}$ is a zero vector. Now, we can deduce Eq. (A2) by multiplying both sides with $\sqrt{\lambda_i}$ and summing over i .

Next, we prove that $\theta_k(\lambda)$ is nonincreasing in k and Schur concave in λ . If $k' \geq k$, then $\Omega_{k'}(\lambda) \geq \Omega_k(\lambda)$, which implies that $\theta_{k'}(\lambda) \leq \theta_k(\lambda)$, so $\theta_k(\lambda)$ is nonincreasing in k .

If $k \geq d$, then $\theta_k(\lambda) = 0$, so $\theta_k(\lambda)$ is Schur concave in λ . To address the case with $k < d$, we need to introduce an auxiliary function,

$$G(y, k, \lambda) \equiv \sum_{i=0}^{d-1} \frac{\lambda_i}{k\lambda_i + y}, \quad k, y > 0, \quad \lambda > 0. \quad (\text{A5})$$

Then $\theta_k(\lambda)$ is the unique solution of $y > 0$ to the equation $G(y, k, \lambda) = 1$. Note that $G(y, k, \lambda)$ is nonincreasing in y , convex in each entry of λ , and invariant under any permutation of these entries. Therefore, $G(y, k, \lambda)$ is Schur concave in λ . Suppose $\lambda' \in \Delta_d$ and $\lambda' \prec \lambda$; let $y = \theta_k(\lambda)$, then $G(y, k, \lambda') \geq G(y, k, \lambda) = 1$, which implies that $\theta_k(\lambda') \geq y = \theta_k(\lambda)$. Therefore, $\theta_k(\lambda)$ is Schur concave in λ , which completes the proof of Lemma A1. ■

Lemma A2. Suppose k is a positive integer, $\lambda \in \Delta_d$ has r nonzero entries with $\tilde{r}$ distinct nonzero values labeled as $\{\ell_j\}_{j=0}^{\tilde{r}-1}$, where ℓ_j has multiplicity m_j . Let $\tilde{\Omega}_k(\lambda)$ be the $\tilde{r} \times \tilde{r}$ matrix with entries $\tilde{\Omega}_k(\lambda)_{ij} = k\ell_i \delta_{ij} - \sqrt{m_i m_j \ell_i \ell_j}$. Then the eigenvalues $\{\omega_j\}_{j=0}^{\tilde{r}-1}$ of $\tilde{\Omega}_k(\lambda)$ are nondegenerate and satisfy

$$-1 < \omega_{\tilde{r}-1} < k\ell_{\tilde{r}-1} < \omega_{\tilde{r}-2} < k\ell_{\tilde{r}-2} < \dots < \omega_0 < k\ell_0. \quad (\text{A6})$$

In addition,

$$\text{rank}(\tilde{\Omega}_k(\lambda)) = \begin{cases} \tilde{r} - 1, & k = r, \\ \tilde{r}, & k \neq r, \end{cases} \quad (\text{A7})$$

$$\sigma(\Omega_k(\lambda)) = \bigcup_{j=0}^{\tilde{r}-1} \{k\ell_j\}^{\times(m_j-1)} \cup \sigma(\tilde{\Omega}_k(\lambda)) \cup \{0\}^{\times(d-r)}. \quad (\text{A8})$$

Proof of Lemma A2. According to the Cauchy Interlacing Theorem, the eigenvalues $\{\omega_j\}_{j=0}^{\tilde{r}-1}$ of $\tilde{\Omega}_k(\lambda)$ satisfy

$$-1 < \omega_{\tilde{r}-1} \leq k\ell_{\tilde{r}-1} \leq \omega_{\tilde{r}-2} \leq k\ell_{\tilde{r}-2} \leq \dots \leq \omega_0 \leq k\ell_0. \quad (\text{A9})$$

Here the first inequality holds because $\ell_j > 0$ and $\sum_{j=0}^{\tilde{r}-1} m_j \ell_j = 1$. In addition, using proof by contradiction, it is straightforward to verify that $k\ell_j$ is not an eigenvalue of $\tilde{\Omega}_k(\lambda)$ for $j = 0, 1, \dots, \tilde{r} - 1$. So all the inequalities in the above equation are necessarily strict, which confirms Eq. (A6) and shows that the eigenvalues $\{\omega_j\}_{j=0}^{\tilde{r}-1}$ of $\tilde{\Omega}_k(\lambda)$ are nondegenerate.

Next, we prove Eq. (A8) before Eq. (A7). To this end, we need to introduce the following index sets:

$$\begin{aligned} \mathcal{I}_j &\equiv \{i : \lambda_i = \ell_j\}, \quad j = 0, 1, \dots, \tilde{r} - 1; \\ \mathcal{I}_{\tilde{r}} &\equiv \{i : \lambda_i = 0\}. \end{aligned} \quad (\text{A10})$$

On this basis we can define the following subspaces of $\mathbb{R}^d$:

$$\begin{aligned} \mathcal{V}_j &= \left\{ v \in \mathbb{R}^d : v_i = 0 \quad \forall i \notin \mathcal{I}_j, \quad \sum_i v_i = 0 \right\}, \\ j &= 0, 1, \dots, \tilde{r} - 1; \quad \mathcal{V}_{\tilde{r}} = \{v \in \mathbb{R}^d : v_i = 0 \quad \forall i \notin \mathcal{I}_{\tilde{r}}\}. \end{aligned} \quad (\text{A11})$$

Then every nonzero vector in $\mathcal{V}_j$ is an eigenvector of $\Omega_k(\lambda)$ with eigenvalue $k\ell_j$, where $\ell_{\tilde{r}} = 0$. In addition,

$$\dim \mathcal{V}_j = |\mathcal{I}_j| = m_j - 1, \quad j = 0, 1, \dots, \tilde{r} - 1; \quad (\text{A12})$$

$$\dim \mathcal{V}_{\tilde{r}} = |\mathcal{I}_{\tilde{r}}| = d - r, \quad (\text{A13})$$

Define $\mathcal{W} \equiv (\mathcal{V}_1 + \mathcal{V}_2 + \dots + \mathcal{V}_{\tilde{r}})^\perp$. Then $\mathcal{W}$ is an invariant subspace of $\Omega_k(\lambda)$ that has dimension $\tilde{r}$ and is spanned

by the following $\tilde{r}$ vectors in $\mathbb{R}^d$:

$$(v_j)_i = \begin{cases} \frac{1}{\sqrt{m_j}}, & i \in \mathcal{I}_j \\ 0, & i \notin \mathcal{I}_j \end{cases} \quad j = 0, 1, \dots, \tilde{r} - 1. \quad (\text{A14})$$

Now, consider the following map from $\mathbb{R}^{\tilde{r}}$ to $\mathbb{R}^d$:

$$x \mapsto v_x \equiv \sum_{j=0}^{\tilde{r}-1} x_j v_j. \quad (\text{A15})$$

Then v_x is an eigenvector of $\Omega_k(\lambda)$ iff x is an eigenvector of $\tilde{\Omega}_k(\lambda)$ with the same eigenvalue. This observation implies Eq. (A8).

Equation (A7) is a simple corollary of Eqs. (A1) and (A8) given that $\sum_{j=0}^{\tilde{r}-1} (m_j - 1) = r - \tilde{r}$, which completes the proof of Lemma A2. ■

2. Proof of Theorem 1

Proof. Without loss of generality we can assume that $|\psi\rangle$ has the following Schmidt decomposition: $|\psi\rangle = \sum_{i=0}^{d-1} \sqrt{\lambda_i} |ii\rangle$, where $d = \min\{d_A, d_B\} = d_A$. Define $\mathcal{W}_1 \equiv \text{span}(\{|ii\rangle\}_{i=0}^{d-1})$ and $\mathcal{W}_2 = \mathcal{W}_1^\perp$. Then $\mathcal{W}_1$ and $\mathcal{W}_2$ have dimensions d and $d(d_B - 1)$, respectively. The k -reduced operator $\mathcal{R}_k(\rho)$ can be expressed as follows:

$$\begin{aligned} \mathcal{R}_k(\rho) &= k \sum_{i=0}^{d-1} \lambda_i |i\rangle \langle i| \otimes I_B - \sum_{i,j=0}^{d-1} \sqrt{\lambda_i \lambda_j} |ii\rangle \langle jj| \\ &= R_1 + R_2, \end{aligned} \quad (\text{A16})$$

where

$$\begin{aligned} R_1 &\equiv k \sum_{i=0}^{d-1} \lambda_i |ii\rangle \langle ii| - \sum_{i,j=0}^{d-1} \sqrt{\lambda_i \lambda_j} |ii\rangle \langle jj|, \\ R_2 &\equiv k \sum_{i=0}^{d-1} \sum_{j=0, j \neq i}^{d_B-1} \lambda_i |ij\rangle \langle ij| \end{aligned} \quad (\text{A17})$$

are supported in $\mathcal{W}_1, \mathcal{W}_2$, respectively. Note that $\sigma(R_1) = \sigma(\Omega_k(\lambda))$ and $\sigma(R_2) = [k\sigma(\psi)]^{\times(d_B-1)}$. Therefore,

$$\begin{aligned} \sigma(\mathcal{R}_k(\rho)) &= \sigma(R_1) \cup \sigma(R_2) \\ &= \sigma(\Omega_k(\lambda)) \cup \{k\sigma(\psi)\}^{\times(d_B-1)}, \end{aligned} \quad (\text{A18})$$

which confirms Eq. (29). Since all eigenvalues of R_2 are non-negative, it follows that $\mathcal{R}_k(\rho)$ and $\Omega_k(\lambda)$ have the same negative spectrum, including multiplicities. According to Lemma A1, if $k \geq r$, then $\Omega_k(\lambda) \geq 0$ and $\mathcal{R}_k(\rho) \geq 0$. If instead $1 \leq k < r$, then both $\Omega_k(\lambda)$ and $\mathcal{R}_k(\rho)$ have exactly one negative eigenvalue.

Let $\{\ell_j\}_{j=0}^{\tilde{r}-1}$ be the set of distinct nonzero Schmidt coefficients of $|\psi\rangle$, where ℓ_j has multiplicity m_j . In conjunction with Lemma A2 we can further deduce that

$$\sigma(\mathcal{R}_k(\rho)) = \bigcup_{j=0}^{\tilde{r}-1} \{k\ell_j\}^{\times(d_B m_j - 1)} \cup \sigma(\tilde{\Omega}_k(\lambda)) \cup \{0\}^{\times d_B(d-r)}, \quad (\text{A19})$$

where $\tilde{\Omega}_k(\lambda)$ is the $\tilde{r} \times \tilde{r}$ matrix defined in Lemma A2, and its eigenvalues satisfy Eq. (A6). Moreover, $\tilde{\Omega}_k(\lambda)$ is non-singular when $k \neq r$ and has exactly one eigenvalue equal to 0 when $k = r$. Therefore, $\mathcal{R}_k(\rho)$ has $2\tilde{r}$ distinct nonzero eigenvalues when $k \neq r$ and $2\tilde{r} - 1$ distinct nonzero eigenvalues when $k = r$. In any case, $\mathcal{R}_k(\rho)$ has at most $2\tilde{r}$ distinct nonzero eigenvalues. This observation completes the proof of Theorem 1. ■

3. Proof of Theorem 2

Proof. According to Theorem 1 and Lemma A1, $\mathcal{N}_k(\psi) = \theta_k(\lambda)$ is Schur concave in λ and nonincreasing in k . Next, suppose $\text{SN}(|\psi\rangle) = r$ and let $\lambda' \equiv (r^{-1}, \dots, r^{-1}, 0, \dots, 0) \in \Delta_d$, where the value r^{-1} is repeated r times. Then $\lambda \succ \lambda'$, which implies that $\mathcal{N}_k(\lambda) = \theta_k(\lambda) \leq \theta_k(\lambda') = \max\{0, 1 - k/r\}$. ■

4. Proof of Proposition 3

Proof. Define $\mathcal{W}_0 \equiv \text{span}(\{|i\rangle_A \otimes |j\rangle_B\}_{i,j=0}^{r-1})$. Let $I_{\mathcal{W}_0}$ be the projector onto this subspace and $I_{\mathcal{W}_0^\perp} = I - I_{\mathcal{W}_0}$, then

$$\begin{aligned} \rho_{\varepsilon,r} &= (1 - \varepsilon)|+_r\rangle \langle +_r| + \varepsilon \frac{I_{\mathcal{W}_0} + I_{\mathcal{W}_0^\perp}}{D} \\ &= (1 - \varepsilon')\rho_{\varepsilon,r}^{\text{ef}} + \varepsilon' \frac{I_{\mathcal{W}_0^\perp}}{D - r^2}, \end{aligned} \quad (\text{A20})$$

where $D \equiv d_A d_B$ and

$$\rho_{\varepsilon,r}^{\text{ef}} \equiv \frac{1 - \varepsilon}{1 - \varepsilon'} |+_r\rangle \langle +_r| + \frac{\varepsilon}{1 - \varepsilon'} \frac{I_{\mathcal{W}_0}}{D}, \quad \varepsilon' \equiv \left(1 - \frac{r^2}{D}\right) \varepsilon. \quad (\text{A21})$$

Every pure state decomposition of $\rho_{\varepsilon,r}^{\text{ef}}$ corresponds to a pure state decomposition of $\rho_{\varepsilon,r}$ with the same maximal Schmidt number. Thus, $\text{SN}(\rho_{\varepsilon,r}) \leq \text{SN}(\rho_{\varepsilon,r}^{\text{ef}})$. Note that $\rho_{\varepsilon,r}^{\text{ef}}$ is an isotropic state of a bipartite system with local dimensions both equal to r . It is known that the Schmidt

number of

$$\frac{r^2 F - 1}{r^2 - 1} |+_r\rangle\langle+_r| + \frac{1 - F}{r^2 - 1} I_{\mathcal{W}_0} \quad (\text{A22})$$

equals $\lceil rF \rceil$ [15]. Hence, if F satisfies

$$\frac{r^2 F - 1}{r^2 - 1} = \frac{1 - \varepsilon}{1 - \varepsilon'}, \quad (\text{A23})$$

then we have

$$\begin{aligned} \text{SN}(\rho_{\varepsilon,r}) &\leq \text{SN}(\rho_{\varepsilon,r}^{\text{ef}}) = \lceil rF \rceil \\ &= \left\lceil \frac{r^2 - 1}{r} \cdot \frac{D(1 - \varepsilon)}{D(1 - \varepsilon) + \varepsilon r^2} + \frac{1}{r} \right\rceil \\ &= \left\lceil \frac{(1 + u)r}{1 + r^2 u} \right\rceil, \end{aligned} \quad (\text{A24})$$

which confirms the upper bound in Eq. (39).

On the other hand, if

$$k \leq \left\lceil \frac{(1 + u)r}{1 + d_B r u} \right\rceil - 1, \quad (\text{A25})$$

then $\mathcal{N}_k(\rho_{\varepsilon,r}) > 0$ by Proposition 2, which implies that

$$\text{SN}(\rho_{\varepsilon,r}) \geq \max \left\{ \frac{(1 + u)r}{1 + d_B r u}, \frac{(1 + u)r}{1 + d_A r u} \right\} = \frac{(1 + u)r}{1 + d r u}. \quad (\text{A26})$$

This confirms the lower bound in Eq. (39).

Finally, if $r \leq \sqrt{d}$ and $\varepsilon < 1/2$, then we have $u < (d_A d_B)^{-1}$ and

$$\frac{(1 + u)r}{1 + d r u} > r - 1, \quad (\text{A27})$$

which implies that $\text{SN}(\rho_{\varepsilon,r}) = r$. This observation completes the proof of Proposition 3. ■

APPENDIX B: PROOFS OF THEOREMS 4 AND 5

1. An auxiliary lemma

Given $x \in \mathbb{R}$ and $l \in \mathbb{N}_0$, define

$$e(x, l) \equiv (1, x, x^2, \dots, x^l)^\top. \quad (\text{B1})$$

Consider the moment sequence $S = (s_n)_{n \in \mathbb{N}_0}$ with $s_n = x^n$ and let $S_{k,2l+k} \equiv (s_k, s_{k+1}, \dots, s_{2l+k})$ be a truncated subsequence. The Hankel matrix $H(S_{k,2l+k})$ has entries

$$\begin{aligned} [H(S_{k,2l+k})]_{m,n} &= s_{m+n+k} = x^k e(x, l)_m e(x, l)_n, \\ m, n &= 0, 1, \dots, l. \end{aligned} \quad (\text{B2})$$

So the Hankel matrix itself can be expressed as follows:

$$H(S_{k,2l+k}) = x^k e(x, l) e(x, l)^\top. \quad (\text{B3})$$

Next, as a generalization, consider the moment sequence $S' = (s'_n)_{n \in \mathbb{N}_0}$ with $s'_n = \sum_{i=0}^{\chi-1} m_i x_i^n$, where $x_i \in \mathbb{R}$ and $m_i \in \mathbb{N}$. By linearity, we have

$$H(S'_{k,2l+k}) = \sum_{i=0}^{\chi-1} m_i x_i^k e(x_i, l) e(x_i, l)^\top. \quad (\text{B4})$$

The following lemma is instructive for understanding the properties of the Hankel matrix $H(S'_{k,2l+k})$.

Lemma B1. Suppose $\{x_i\}_{i=0}^{\chi-1}$ is a set of distinct real numbers and $l \in \mathbb{N}_0$. Then

$$\text{rank}([e(x_0, l), e(x_1, l), \dots, e(x_{\chi-1}, l)]) = \min\{l + 1, \chi\}. \quad (\text{B5})$$

If $l + 1 \geq \chi$, then the set $\{e(x_i, l)\}_{i=0}^{\chi-1}$ is linearly independent.

Proof. Let V be the Vandermonde matrix associated with the set $\{x_i\}_{i=0}^{\chi-1}$, that is,

$$V_{ij} \equiv x_i^j \quad i, j = 0, 1, \dots, \min\{l, \chi - 1\}. \quad (\text{B6})$$

Thanks to the Vandermonde determinant formula, we have

$$\det(V) = \prod_{0 \leq i < j \leq \dim(V)-1} (x_i - x_j) \neq 0. \quad (\text{B7})$$

Therefore,

$$\begin{aligned} \min\{l + 1, \chi\} \\ &= \text{rank}(V) \leq \text{rank}([e(x_0, l), e(x_1, l), \dots, e(x_{\chi-1}, l)]) \\ &\leq \min\{l + 1, \chi\}, \end{aligned} \quad (\text{B8})$$

which implies Eq. (B5).

If $l + 1 \geq \chi$, then $\text{rank}([e(x_0, l), e(x_1, l), \dots, e(x_{\chi-1}, l)]) = \chi$ by Eq. (B5), so the set $\{e(x_i, l)\}_{i=0}^{\chi-1}$ is linearly independent, which completes the proof of Lemma B1. ■

2. Proofs of Theorems 4 and 5

In preparation for proving Theorems 4 and 5, suppose $\rho \in \mathcal{S}(\mathcal{H}_{\text{AB}})$ and $\mathcal{R}_k(\rho)$ has χ distinct nonzero eigenvalues, namely, $\{x_i\}_{i=0}^{\chi-1}$, where x_i has multiplicity m_i . According to the definition in Eq. (65) and the formula in Eq. (B4),

$B_N[\rho, k]$ can be expressed as follows:

$$B_N[\rho, k] = \sum_{i=0}^{\chi-1} m_i p_{N,k}(x_i) b_N(x_i) b_N(x_i)^\top, \quad (\text{B9})$$

where

$$b_N(x) \equiv e(x, \lfloor (N-1)/2 \rfloor) = (1, x, x^2, \dots, x^{\lfloor \frac{N-1}{2} \rfloor})^\top, \quad (\text{B10})$$

$$p_{N,k}(x) \equiv \begin{cases} x, & N \text{ odd}, \\ x(k-x), & N \text{ even}. \end{cases} \quad (\text{B11})$$

Proof of Theorem 4. Suppose $\mathcal{R}_k(\rho)$ has χ different nonzero eigenvalues $\{x_i\}_{i=0}^{\chi-1}$, where x_i has multiplicity $m_i \geq 1$. Then B_N has the decomposition in Eq. (B9) with $x_i \in \mathbb{R}$ and $m_i \in \mathbb{N}$. Since $\text{SN}(\rho) \leq k$ by assumption, we have $0 \leq x_i \leq k$ for all i thanks to Theorem 1 and Corollary 1. Therefore, $m_i p_{N,k}(x_i) \geq 0$ for all i , which means $B_N \geq 0$ and completes the proof of Theorem 4. ■

Proof of Theorem 5. Let $\{x_i\}_{i=0}^{\chi-1}$ be the set of distinct nonzero eigenvalues of $\mathcal{R}_k(\rho)$ and denote by m_i the multiplicity of x_i . Then B_N has the decomposition in Eq. (B9) with $m_i \in \mathbb{N}$ and $x_i \in [-1 + k/d, k]$ by Corollary 1. In addition, $\{b_N(x_i)\}_{i=0}^{\chi-1}$ are linearly independent according to Lemma B1, given that $N \geq 2\chi - 1$ by assumption. Hence, $B_N \geq 0$ iff $m_i p_{N,k}(x_i) \geq 0$ for all i . Meanwhile, $m_i p_{N,k}(x_i) \geq 0$ iff $x_i \geq 0$. Therefore, $B_N \geq 0$ iff $x_i \geq 0$ for all i , that is, $\mathcal{R}_k(\rho) \geq 0$, which completes the proof of Theorem 5. ■

APPENDIX C: PROOFS OF THEOREM 6 AND PROPOSITION 7

1. Proof of Theorem 6

Proof of Theorem 6. Let $k = 1, 2, \dots, r-1$; then $\mathcal{R}_k(\psi) \not\geq 0$ and $\mathcal{R}_k(\psi)$ has at most $2\tilde{r}$ distinct nonzero eigenvalues according to Theorem 1. Suppose $N \geq 4\tilde{r} - 1$; then $\mathcal{R}_k(\psi) \not\geq 0$ iff $B_N[\psi, k] \not\geq 0$ thanks to Theorem 5. Therefore, $B_N[\psi, k] \not\geq 0$, which confirms Eq. (74).

Let $\{\ell_j\}_{j=0}^{\tilde{r}-1}$ be the set of distinct nonzero Schmidt coefficients of $|\psi\rangle$, where ℓ_j has multiplicity m_j . According to Lemma A2, the spectrum of $\mathcal{R}_k(\psi)$ can be expressed as follows:

$$\sigma(\mathcal{R}_k(\psi)) = \bigcup_{j=0}^{\tilde{r}-1} \{k\ell_j\}^{\times(d_B m_j - 1)} \cup \{\omega_j\}_{j=0}^{\tilde{r}-1} \cup \{0\}^{\times d_B(d-r)}, \quad (\text{C1})$$

where ω_j for $j = 0, 1, \dots, \tilde{r} - 1$ are the eigenvalues of the matrix $\tilde{\Omega}(\lambda)$ defined in Lemma A2. Note that $\sum_j m_j = r$

and

$$\begin{aligned} 0 < k\ell_j \leq k, \quad -1 < \omega_j < k \quad \forall j = 0, 1, \dots, \tilde{r} - 1; \\ 0 < \omega_j < k \quad \forall j = 0, 1, \dots, \tilde{r} - 2. \end{aligned} \quad (\text{C2})$$

In conjunction with Eqs. (65) and (B9), the Hankel matrix $B_N[\psi, k]$ can be expressed as follows:

$$\begin{aligned} B_N &= B_N[\psi, k] = A_1 + A_2, \\ A_1 &\equiv \sum_{j=0}^{\tilde{r}-1} p_{N,k}(\omega_j) b_N(\omega_j) b_N(\omega_j)^\top, \\ A_2 &\equiv \sum_{j=0}^{\tilde{r}-1} (d_B m_j - 1) p_{N,k}(k\ell_j) b_N(k\ell_j) b_N(k\ell_j)^\top, \end{aligned} \quad (\text{C3})$$

where b_N and $p_{N,k}(x)$ are defined in Eqs. (B10) and (B11), respectively. Note that $p_{N,k}(k\ell_j) \geq 0$, while $p_{N,k}(\omega_j)$ has the same sign as ω_j for $j = 0, 1, \dots, \tilde{r} - 1$.

Now suppose $N \leq 2\tilde{r}$ and $k = r - 1$. Let $\ell_{\min} \equiv \min_j \{\ell_j\}$ and $\omega_{\min} \equiv \min_j \{\omega_j\} = -\mathcal{N}_{r-1}(\psi)$; then $\omega_{\min} \in (-1, 0)$. In addition, A_1 , A_2 , and B_N are square matrices with at most $\tilde{r}$ rows (columns), and $1 \leq \|b_N(\omega_{\min})\|_2^2 \leq \tilde{r}$. In conjunction with Lemma B1, we can deduce that A_2 has full rank, which means $\text{Im}(A_1) \subset \text{Im}(A_2)$.

By definition, in Eq. (75), we have

$$A_0 = \sum_{j=0}^{\tilde{r}-1} p_{N,r-1}[(r-1)\ell_j] b_N[(r-1)\ell_j] b_N[(r-1)\ell_j]^\top, \quad (\text{C4})$$

which means $\sigma_{\min}(A_2) \geq (d_B - 1)\sigma_{\min}(A_0)$ given that $m_j \geq 1$ and $p_{N,r-1}[(r-1)\ell_j] \geq 0$ for all j . If in addition N is odd, then

$$\sigma_{\min}(A_1) \geq p_{N,r-1}(\omega_{\min}) \|b_N(\omega_{\min})\|_2^2 \geq \tilde{r}\omega_{\min}, \quad (\text{C5})$$

$$\begin{aligned} \sigma_{\min}(B_N) &\geq \sigma_{\min}(A_1) + \sigma_{\min}(A_2) \\ &\geq \tilde{r}\omega_{\min} + (d_B - 1)\sigma_{\min}(A_0). \end{aligned} \quad (\text{C6})$$

If instead N is even, then

$$\sigma_{\min}(A_1) \geq p_{N,r-1}(\omega_{\min}) \|b_N(\omega_{\min})\|_2^2 \geq r\tilde{r}\omega_{\min}, \quad (\text{C7})$$

$$\sigma_{\min}(B_N) \geq r\tilde{r}\omega_{\min} + (d_B - 1)\sigma_{\min}(A_0). \quad (\text{C8})$$

Note that $\omega_{\min} = -\mathcal{N}_{r-1}(\psi)$ and that Eq. (C8) holds in both cases. If the condition in Eq. (76) is satisfied, that is, $d_B - 1 \geq r\tilde{r}\mathcal{N}_{r-1}(\psi)/\sigma_{\min}(A_0)$, then $\sigma_{\min}(B_N) > 0$, which confirms Eq. (77) and completes the proof of Theorem 6. ■

2. Proof of Proposition 7

Proof. The k -reduced operator $\mathcal{R}_k(\rho_F)$ has spectrum

$$\sigma(\mathcal{R}_k(\rho_F)) = \{\lambda_1\}^{\times(d^2-1)} \sqcup \{\lambda_0\},$$

$$\lambda_1 = \frac{F-1}{d^2-1} + \frac{k}{d} > 0, \quad \lambda_0 = \frac{k}{d} - F. \quad (\text{C9})$$

Note that $\mathcal{R}_k(\rho_F)$ has at most two distinct nonzero eigenvalues. Given $N = 3$, then $B_N[\rho_F, k] \geq 0$ iff $\mathcal{R}_k(\rho_F) \geq 0$ according to Theorem 5. If in addition $k \leq \lceil dF \rceil - 1$, then $\lambda_0 < 0$ and $\mathcal{R}_k(+_r) \not\geq 0$, which further implies that $B_N[+_r, k] \not\geq 0$. ■

APPENDIX D: PROOF OF PROPOSITION 8

Proof. By assumption the local Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ have the same dimension d and thus can be identified as two copies of a d -dimensional Hilbert space $\mathcal{H}$. Let $\{P_j\}_{j=1}^{d^2-1}$ be an orthogonal basis on $\mathcal{L}_0^H(\mathcal{H})$ that satisfies the normalization condition $\text{tr}(P_j P_k) = d\delta_{j,k}$. Then $\{P_j^*\}_{j=1}^{d^2-1}$ is also an orthogonal basis on this space. In conjunction with the definition of ρ_F in Eq. (78), the entries of its CM can be calculated as follows:

$$\begin{aligned} T_{jk} &= \frac{d^2 F - 1}{d(d^2 - 1)} \langle +_d | P_j \otimes P_k^* | +_d \rangle \\ &= \frac{d^2 F - 1}{d^2(d^2 - 1)} \sum_{m,n=0}^{d-1} \langle mm | P_j \otimes P_k^* | nm \rangle \\ &= \frac{d^2 F - 1}{d^2(d^2 - 1)} \sum_{m,n=0}^{d-1} \langle m | P_j | n \rangle \langle n | P_k^* | m \rangle \\ &= \frac{d^2 F - 1}{d^2(d^2 - 1)} \text{Tr}(P_j P_k^*) = \frac{d^2 F - 1}{d(d^2 - 1)} \delta_{j,k}, \end{aligned} \quad (\text{D1})$$

where $j, k = 1, 2, \dots, d^2 - 1$. Therefore, all singular values of T are equal to $(d^2 F - 1)/d(d^2 - 1)$, which has multiplicity $d^2 - 1$. As a simple corollary, we have

$$\|T\|_1 = (d^2 - 1) \times \frac{d^2 F - 1}{d(d^2 - 1)} = \frac{d^2 F - 1}{d}, \quad (\text{D2})$$

$$\|T\|_2^2 = (d^2 - 1) \times \left[\frac{d^2 F - 1}{d(d^2 - 1)} \right]^2 = \frac{(d^2 F - 1)^2}{d^2(d^2 - 1)}, \quad (\text{D3})$$

$$\|T\|_4^4 = (d^2 - 1) \times \left[\frac{d^2 F - 1}{d(d^2 - 1)} \right]^4 = \frac{(d^2 F - 1)^4}{d^4(d^2 - 1)^3}, \quad (\text{D4})$$

which completes the proof of Proposition 8. ■

APPENDIX E: PROOF OF PROPOSITION 9

Proof. To prove that ρ is k -unfaithful, we need to find one solution to Eq. (50). Let $\mu_A = \mu_B = 1/2$, then E_A, E_B

need to satisfy

$$E_A \otimes I_B + I_A \otimes E_B \geq \rho, \quad E_A \leq \frac{I_A}{2}, \quad E_B \leq \frac{I_B}{2}, \quad (\text{E1})$$

$$\text{Tr}(E_A) = \text{Tr}(E_B) = \frac{k-1}{2}.$$

Because $\pi(j) \neq \tau(j)$ for all $j = 0, \dots, k-1$, we have $\langle +_k | \Phi \rangle = 0$. Then the condition $E_A \otimes I_B + I_A \otimes E_B \geq \rho$ leads to

$$\frac{1}{k} \sum_{i=0}^{k-1} \langle i | E_A | i \rangle + \frac{1}{k} \sum_{i=0}^{k-1} \langle i | E_B | i \rangle \geq \langle +_k | \rho | +_k \rangle = \frac{1}{2}, \quad (\text{E2})$$

$$\begin{aligned} &\frac{1}{\chi} \sum_{j=0}^{\chi-1} \phi_j \langle \pi(j) | E_A | \pi(j) \rangle + \frac{1}{\chi} \sum_{j=0}^{\chi-1} \phi_j \langle \tau(j) | E_B | \tau(j) \rangle \\ &\geq \langle \Phi | \rho | \Phi \rangle = \frac{1}{2}. \end{aligned} \quad (\text{E3})$$

We assume that E_A, E_B are diagonal in the computational basis with diagonal entries equal to either $1/4$ or 0 . Then both operators have $2k-2$ nonzero entries because $\text{Tr}(E_A) = \text{Tr}(E_B) = (k-1)/2$. The conditions in Eqs. (E2) and (E3) can be fulfilled by setting

$$\langle i | E_A | i \rangle = \langle i | E_B | i \rangle = 1/4, \quad i = 0, 1, \dots, k-1, \quad (\text{E4})$$

$$\begin{aligned} &\langle \pi(j) | E_A | \pi(j) \rangle = \langle \tau(j) | E_B | \tau(j) \rangle = 1/4, \\ &j = 0, 1, \dots, \chi-1, \end{aligned} \quad (\text{E5})$$

which is equivalent to

$$\langle i | E_A | i \rangle = \frac{1}{4} \quad \forall i \in \mathcal{I}_L, \quad \langle i | E_B | i \rangle = \frac{1}{4} \quad \forall i \in \mathcal{I}_R. \quad (\text{E6})$$

Then, as long as $|\mathcal{I}_L|, |\mathcal{I}_R| \leq 2k-2$, the operators E_A, E_B can be constructed as

$$\begin{aligned} E_A &= \frac{1}{4} \sum_{i \in \mathcal{I}_L} |i\rangle \langle i| + \frac{1}{4} \sum_{i \in \mathcal{I}'} |i\rangle \langle i|, \\ E_B &= \frac{1}{4} \sum_{i \in \mathcal{I}_R} |i\rangle \langle i| + \frac{1}{4} \sum_{i \in \mathcal{I}''} |i\rangle \langle i|, \end{aligned} \quad (\text{E7})$$

where $\mathcal{I}'$ can be any index set that has no overlap with $\mathcal{I}_L$ and satisfies $|\mathcal{I}_L| + |\mathcal{I}'| = 2k-2$. The other set $\mathcal{I}''$ can be constructed in the same way. Clearly, $E_A \leq I_A/2$ and $E_B \leq I_B/2$, hence all requirements in Eq. (E1) are satisfied. This proves that ρ is k -unfaithful.

Now we prove that $\mathcal{R}_{k-1}(\rho) \not\geq 0$ with the extra condition in Eq. (85). The reduced density matrix reads

$$\begin{aligned} \rho_A &= \frac{1}{2k} \sum_{i=0}^{k-1} |i\rangle\langle i| + \frac{1}{2} \sum_{j=0}^{k-1} \phi_j |\pi(j)\rangle\langle\pi(j)| \\ &= \sum_{i=0}^{k-1} \frac{1 + k\phi_{\pi^{-1}(i)}}{2k} |i\rangle\langle i| + \frac{1}{2} \sum_{j:\pi(j)\geq k} \phi_j |\pi(j)\rangle\langle\pi(j)|. \end{aligned} \quad (\text{E8})$$

Let $\Pi = \sum_{i=0}^{k-1} |ii\rangle\langle ii|$. Then

$$\begin{aligned} \Pi \mathcal{R}_{k-1}(\rho) \Pi &= (k-1) \sum_{i=0}^{k-1} \langle i|\rho_A|i\rangle |ii\rangle\langle ii| - \frac{1}{2} |+_k\rangle\langle+_k| \\ &\quad - \frac{1}{2} \Pi |\Phi\rangle\langle\Phi| \Pi, \end{aligned} \quad (\text{E9})$$

where $\Pi |\Phi\rangle\langle\Phi| \Pi = 0$ because $\pi(j) \neq \tau(j)$ for all $j = 0, \dots, k-1$. Clearly, if $\Pi \mathcal{R}_{k-1}(\rho) \Pi \not\geq 0$, then $\mathcal{R}_{k-1}(\rho) \not\geq 0$. Suppose $\mathbf{y} = (y_j)_{j=0}^{k-1}$ is an eigenvector of $\Pi \mathcal{R}_{k-1}(\rho) \Pi$ with eigenvalue $-\theta$. Then

$$\begin{aligned} (k-1) \langle i|\rho_A|i\rangle y_i - \frac{1}{2k} \left(\sum_{i=0}^{k-1} y_i \right) &= -\theta y_i, \\ i &= 0, \dots, k-1. \end{aligned} \quad (\text{E10})$$

Using the same technique in the proof of Lemma A1, we get

$$\sum_{i=0}^{k-1} \frac{1}{(k-1) \langle i|\rho_A|i\rangle + \theta} = 2k. \quad (\text{E11})$$

Hence, $\Pi \mathcal{R}_k(\rho) \Pi$ has a negative eigenvalue iff

$$\sum_{i=0}^{k-1} \frac{1}{(k-1) \langle i|\rho_A|i\rangle} > 2k, \quad (\text{E12})$$

which is equivalent to Eq. (85). This observation completes the proof of Proposition 9. $\blacksquare$

APPENDIX F: PROOF OF THEOREM 7

In this appendix, we will show that, given access to a unitary 3-design, the sample complexity of estimating the functions $p_2 \equiv \text{Tr}(\rho^2)$, $p_3 \equiv \text{Tr}(\rho^3)$, and $t_2 \equiv \text{Tr}[\rho_A \text{Tr}_B(\rho^2)]$ using the classical shadow method is $\mathcal{O}(D)$, where D is the dimension of the underlying Hilbert space $\mathcal{H}_{AB}$. Note that the protocol for estimating a_2, a_3 is identical to the counterpart for estimating p_2, p_3 , but the underlying Hilbert space $\mathcal{H}_A$ has dimension $\sqrt{D}$ instead. Hence, the sample complexity of estimating a_2, a_3 is smaller than that of p_2, p_3 , and the sample complexity of the third order moment-based k -reduction criterion is $\mathcal{O}(D)$.

1. An auxiliary lemma

We first recall the necessary background on unitary designs. Denote by $\Phi^{(k)}$ the k th twirling superoperator defined as follows:

$$\Phi^{(k)}(O) \equiv \int d\mu(U) U^{\otimes k} O (U^\dagger)^{\otimes k}, \quad O \in \mathcal{L}(\mathcal{H}_{AB}^{\otimes k}). \quad (\text{F1})$$

Here $\mu(U)$ denotes the Haar measure on the group $U(\mathcal{H}_{AB})$ of unitary operators on $\mathcal{H}_{AB}$. A unitary ensemble $\mathcal{U}$ on $U(\mathcal{H}_{AB})$ is called a unitary k -design if

$$\mathbb{E}_{U \sim \mathcal{U}} [U^{\otimes k} \cdot (U^\dagger)^{\otimes k}] = \Phi^{(k)}(\cdot). \quad (\text{F2})$$

When $O = (|\psi\rangle\langle\psi|)^{\otimes k}$ is the k th tensor power of a pure state, by virtue of Schur's lemma we can deduce that

$$\begin{aligned} \Phi^{(k)}[(|\psi\rangle\langle\psi|)^{\otimes k}] &= \int_{\text{Haar}} dv (|v\rangle\langle v|)^{\otimes k} \\ &= \frac{(D-1)!}{(D+k-1)!} \sum_{\pi \in \text{Sym}_k} \mathbb{W}_\pi, \end{aligned} \quad (\text{F3})$$

where D is the dimension of the Hilbert space $\mathcal{H}_{AB}$, Sym_k is the symmetric group on k elements, and $\mathbb{W}_\pi$ denotes the permutation operator on $\mathcal{H}_{AB}^{\otimes k}$ corresponding to $\pi \in \text{Sym}_k$. The above equation will be useful in the following derivations.

In the classical shadow protocol, we first sample a random unitary U from a given unitary ensemble $\mathcal{U}$, then measure the rotated state $U\rho U^\dagger$ in the computational basis and record the measurement outcome b . When $\mathcal{U}$ forms a unitary 2-design, an unbiased estimator for the state ρ can be constructed as follows:

$$\hat{\rho} \equiv (D+1)U^\dagger |b\rangle\langle b| U - I. \quad (\text{F4})$$

On this basis, we can construct unbiased estimators for p_2 , p_3 , and t_2 . To analyze the variances of these estimators, we need to introduce some notation and an auxiliary lemma. Define

$$R_\rho \equiv I \otimes I + I \otimes \rho + \rho \otimes I, \quad \mathbb{W} \equiv \mathbb{W}_{(12)}, \quad (\text{F5})$$

where (12) denotes the transposition swapping the first and second elements. The following lemma is quite helpful for computing the variances of the estimators p_2, p_3 , and t_2 . A similar result can be found in Lemma 14 of Ref. [55].

Lemma F1. Suppose $\mathcal{U}$ is a unitary 3-design, and $\hat{\rho}, \hat{\rho}'$ are two independent estimators for ρ . Then $\mathbb{E}[\hat{\rho}] = \rho$ and

$$\mathbb{E}[\hat{\rho} \otimes \hat{\rho}] = \frac{D+1}{D+2} R_{\rho} \mathbb{W} - \frac{1}{D+2} R_{\rho}, \quad (\text{F6})$$

$$\mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^2 = \frac{(D+1)^2 + 1}{(D+2)^2} R_{\rho}^2 - \frac{2(D+1)}{(D+2)} R_{\rho}^2 \mathbb{W}, \quad (\text{F7})$$

$$\begin{aligned} \mathbb{E}[(\hat{\rho} \otimes \hat{\rho}')(\hat{\rho}' \otimes \hat{\rho})] &= \frac{(D+1)^3}{(D+2)^2} R_{\rho} \mathbb{W} - \frac{2(D+1)}{(D+2)^2} R_{\rho}^2 \mathbb{W} + \frac{1}{(D+2)^2} R_{\rho}^2 \\ &\quad + \frac{D(D+1)^2}{(D+2)^2} (\rho \otimes \rho) \mathbb{W} + \frac{(D+1)^2}{(D+2)^2} (1 + p_2) \mathbb{W}. \end{aligned} \quad (\text{F8})$$

Proof. The equality $\mathbb{E}[\hat{\rho}] = \rho$ follows from Eq. (F4) and the equation below:

$$\begin{aligned} \mathbb{E}[U^{\dagger}|b\rangle\langle b|U] &= \mathbb{E}_{U \sim \mathcal{U}} \left[\sum_b \langle b|U\rho U^{\dagger}|b\rangle U^{\dagger}|b\rangle\langle b|U \right] = \sum_b \text{Tr}_2 \left\{ \mathbb{E}_{U \sim \mathcal{U}} [(U^{\dagger})^{\otimes 2} (|b\rangle\langle b|)^{\otimes 2} U^{\otimes 2}] (I \otimes \rho) \right\} \\ &= \sum_b \text{Tr}_2 \left\{ \Phi^{(2)} [(|b\rangle\langle b|)^{\otimes 2}] (I \otimes \rho) \right\} \quad \text{since } \mathcal{U} \text{ is a unitary 2-design} \\ &= \sum_b \text{Tr}_2 \left[\frac{1}{D(D+1)} (I + \mathbb{W})(I \otimes \rho) \right] = \frac{\rho + I}{D+1}, \end{aligned} \quad (\text{F9})$$

where the fourth equality follows from Eq. (F3) with $k = 2$ and in deriving the last equality we have taken into account a factor of D from the summation over b .

To prove Eq. (F6), we first evaluate the following expectation value:

$$\begin{aligned} \mathbb{E}[(U^{\dagger}|b\rangle\langle b|U)^{\otimes 2}] &= \mathbb{E}_{U \sim \mathcal{U}} \left[\sum_b \langle b|U\rho U^{\dagger}|b\rangle (U^{\dagger}|b\rangle\langle b|U)^{\otimes 2} \right] \\ &= \sum_b \text{Tr}_3 \left\{ \mathbb{E}_{U \sim \mathcal{U}} [(U^{\dagger})^{\otimes 3} (|b\rangle\langle b|)^{\otimes 3} U^{\otimes 3}] (I \otimes I \otimes \rho) \right\} \\ &= \sum_b \text{Tr}_3 \left\{ \Phi^{(3)} [(|b\rangle\langle b|)^{\otimes 3}] (I \otimes I \otimes \rho) \right\} \quad \text{since } \mathcal{U} \text{ is a unitary 3-design} \\ &= \frac{1}{(D+1)(D+2)} (I \otimes I + \rho \otimes I + I \otimes \rho)(I + \mathbb{W}), \end{aligned} \quad (\text{F10})$$

where the last equality follows from Eq. (F3) with $k = 3$. Now, Eq. (F6) can be proved as follows:

$$\begin{aligned} \mathbb{E}[\hat{\rho} \otimes \hat{\rho}] &= (D+1)^2 \mathbb{E}[(U^{\dagger}|b\rangle\langle b|U)^{\otimes 2}] - (D+1) \left\{ \mathbb{E}[U^{\dagger}|b\rangle\langle b|U] \otimes I + I \otimes \mathbb{E}[U^{\dagger}|b\rangle\langle b|U] \right\} + I \otimes I \\ &= \frac{D+1}{D+2} (I \otimes I + I \otimes \rho + \rho \otimes I) \mathbb{W} - \frac{1}{D+2} (I \otimes I + \rho \otimes I + I \otimes \rho). \end{aligned} \quad (\text{F11})$$

Equation (F7) can be proved using a similar reasoning employed above.

To prove Eq. (F8), we first rewrite the LHS as follows:

$$\begin{aligned}
\mathbb{E}[(\hat{\rho} \otimes \hat{\rho}')(\hat{\rho}' \otimes \hat{\rho})] &= \mathbb{E} \left\{ \text{Tr}_{34} \left[(\hat{\rho} \otimes \hat{\rho}' \otimes \hat{\rho}' \otimes \hat{\rho}) \cdot \mathbb{W}_{(13)(24)} \right] \right\} \\
&= \mathbb{E} \left\{ \text{Tr}_{34} \left[\mathbb{W}_{(234)} (\hat{\rho} \otimes \hat{\rho} \otimes \hat{\rho}' \otimes \hat{\rho}') \mathbb{W}_{(234)}^{-1} \cdot \mathbb{W}_{(13)(24)} \right] \right\} \\
&= \text{Tr}_{34} \left\{ \mathbb{W}_{(234)} (\mathbb{E}[\hat{\rho} \otimes \hat{\rho}] \otimes \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]) \mathbb{W}_{(134)} \right\} \\
&= \frac{(D+1)^2}{(D+2)^2} \text{Tr}_{34} [\mathbb{W}_{(234)} (R_\rho \mathbb{W} \otimes R_\rho \mathbb{W}) \mathbb{W}_{(134)}] - \frac{(D+1)}{(D+2)^2} \text{Tr}_{34} [\mathbb{W}_{(234)} (R_\rho \mathbb{W} \otimes R_\rho) \mathbb{W}_{(134)}] \\
&\quad - \frac{(D+1)}{(D+2)^2} \text{Tr}_{34} [\mathbb{W}_{(234)} (R_\rho \otimes R_\rho \mathbb{W}) \mathbb{W}_{(134)}] + \frac{1}{(D+2)^2} \text{Tr}_{34} [\mathbb{W}_{(234)} (R_\rho \otimes R_\rho) \mathbb{W}_{(134)}].
\end{aligned} \tag{F12}$$

In addition, direct calculation yields

$$\text{Tr}_{34} [\mathbb{W}_{(234)} (R_\rho \mathbb{W} \otimes R_\rho \mathbb{W}) \mathbb{W}_{(134)}] = (D+1)R_\rho \mathbb{W} + (1+p_2)\mathbb{W} + D(\rho \otimes \rho)\mathbb{W}, \tag{F13}$$

$$\text{Tr}_{34} [\mathbb{W}_{(234)} (R_\rho \mathbb{W} \otimes R_\rho) \mathbb{W}_{(134)}] = (1+2\rho \otimes I)R_\rho \mathbb{W}, \tag{F14}$$

$$\text{Tr}_{34} [\mathbb{W}_{(234)} (R_\rho \otimes R_\rho \mathbb{W}) \mathbb{W}_{(134)}] = (1+2I \otimes \rho)R_\rho \mathbb{W}, \tag{F15}$$

$$\text{Tr}_{34} [\mathbb{W}_{(234)} (R_\rho \otimes R_\rho) \mathbb{W}_{(134)}] = R_\rho^2. \tag{F16}$$

In conjunction with Eq. (F12), we can derive Eq. (F8) and complete the proof of Lemma F1. $\blacksquare$

2. Variance of $\text{Tr}(\rho^2)$ by classical shadow

The same computation was done in Ref. [9]. Here we write down the computation process with our notation for completeness. Label the state estimators for the M runs as $\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_M$ separately, where M is the total number of samples. Then the estimator for the purity is (recall that $p_n \equiv \text{Tr}(\rho^n)$)

$$\hat{p}_2 = \binom{M}{2}^{-1} \sum_{j < k} \text{Tr}(\hat{\rho}_j \hat{\rho}_k). \tag{F17}$$

Its variance $\text{Var}[\hat{p}_2]$ can be computed by

$$\begin{aligned}
\text{Var}[\hat{p}_2] &= \mathbb{E}[\hat{p}_2^2] - p_2^2 = \binom{M}{2}^{-2} \sum_{j_1 < k_1, j_2 < k_2} \mathbb{E}[\text{Tr}(\hat{\rho}_{j_1} \hat{\rho}_{k_1}) \text{Tr}(\hat{\rho}_{j_2} \hat{\rho}_{k_2}) - p_2^2] \\
&= \binom{M}{2}^{-1} \{ 2(M-2) [\text{Tr}\{(\rho \otimes \rho) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]\} - p_2^2] + [\text{Tr}(\mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^2) - p_2^2] \} \\
&\leq \frac{4(M-2)}{M(M-1)} \text{Tr}\{(\rho \otimes \rho) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]\} + \frac{2}{M(M-1)} \text{Tr}(\mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^2).
\end{aligned} \tag{F18}$$

Note that

$$\text{Tr}(R_\rho \mathbb{W} R_\rho \mathbb{W}) = \text{Tr}(R_\rho^2) = D^2 + 4D + 2Dp_2 + 2, \quad \text{Tr}(R_\rho^2 \mathbb{W}) = D + 4 + 4p_2. \tag{F19}$$

In conjunction with Eqs. (F6) and (F7) in Lemma F1, we obtain

$$\text{Tr}\{(\rho \otimes \rho) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]\} = \frac{D+1}{D+2} (p_2 + 2p_3) - \frac{1}{D+2} (1 + 2p_2) \leq p_2 + 2p_3, \tag{F20}$$

$$\text{Tr}(\mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^2) = (D+1)^2 - \frac{D^2 + 2(D+1)(D+4)}{(D+2)^2} + 2 \frac{D^2 - 2}{D+2} p_2 \leq (D+1)^2 + 2Dp_2. \tag{F21}$$

Hence,

$$\text{Var}[\hat{p}_2] \leq \frac{4(M-2)}{M(M-1)}(p_2 + 2p_3) + \frac{2}{M(M-1)}[(D+1)^2 + 2Dp_2]. \quad (\text{F22})$$

To guarantee that $\text{Var}[\hat{p}_2] = \mathcal{O}(1)$, we need $M = \mathcal{O}(D)$.

3. Variance of $\text{Tr}(\rho^3)$ by classical shadow

The third moment p_3 has estimator

$$\hat{p}_3 = \frac{1}{2} \binom{M}{3}^{-1} \sum_{i < j < k} \text{Tr}(\hat{\rho}_i \hat{\rho}_j \hat{\rho}_k) + \frac{1}{2} \binom{M}{3}^{-1} \sum_{i < j < k} \text{Tr}(\hat{\rho}_j \hat{\rho}_i \hat{\rho}_k). \quad (\text{F23})$$

Its variance is

$$\begin{aligned} \text{Var}[\hat{p}_3] &= \frac{1}{4} \binom{M}{3}^{-2} \sum_{i_1 < j_1 < k_1, i_2 < j_2 < k_2} \mathbb{E}[\text{Tr}(\hat{\rho}_{i_1} \hat{\rho}_{j_1} \hat{\rho}_{k_1}) \text{Tr}(\hat{\rho}_{i_2} \hat{\rho}_{j_2} \hat{\rho}_{k_2}) - p_3^2] \\ &\quad + \frac{1}{4} \binom{M}{3}^{-2} \sum_{i_1 < j_1 < k_1, i_2 < j_2 < k_2} \mathbb{E}[\text{Tr}(\hat{\rho}_{j_1} \hat{\rho}_{i_1} \hat{\rho}_{k_1}) \text{Tr}(\hat{\rho}_{j_2} \hat{\rho}_{i_2} \hat{\rho}_{k_2}) - p_3^2] \\ &\quad + \frac{1}{2} \binom{M}{3}^{-2} \sum_{i_1 < j_1 < k_1, i_2 < j_2 < k_2} \mathbb{E}[\text{Tr}(\hat{\rho}_{i_1} \hat{\rho}_{j_1} \hat{\rho}_{k_1}) \text{Tr}(\hat{\rho}_{j_2} \hat{\rho}_{i_2} \hat{\rho}_{k_2}) - p_3^2]. \end{aligned} \quad (\text{F24})$$

Now we analyze the expectation value of each summand with different indices. Based on how many pairs of $\hat{\rho}$ sharing the same index, we can classify the summands into four categories.

(1) All indices are distinct, then the summand equals

$$\mathbb{E}[\text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3) \text{Tr}(\hat{\rho}_4 \hat{\rho}_5 \hat{\rho}_6)] - p_3^2 = 0. \quad (\text{F25})$$

(2) There exists exactly a pair of identical indices, then in conjunction with Eq. (F6), the summand equals

$$\begin{aligned} \mathbb{E}[\text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3) \text{Tr}(\hat{\rho}_1 \hat{\rho}_4 \hat{\rho}_5)] &= \text{Tr}\{\mathbb{E}[\hat{\rho} \otimes \hat{\rho}](\rho^2 \otimes \rho^2)\} \\ &= \frac{D+1}{D+2} \text{Tr}[R_\rho \mathbb{W}(\rho^2 \otimes \rho^2)] - \frac{1}{D+2} \text{Tr}[R_\rho(\rho^2 \otimes \rho^2)] \\ &= \frac{D+1}{D+2} (p_4 + 2p_5) - \frac{1}{D+2} (p_3^2 + 2p_2 p_3) = \mathcal{O}(1). \end{aligned} \quad (\text{F26})$$

(3) There exist exactly two pairs of identical indices, then the summand has the following three types. Using Eq. (F7), we have

$$\begin{aligned} \mathbb{E}[\text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3) \text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_4)] &= \text{Tr}\{\mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^2(\rho \otimes \rho)\} \\ &= \frac{(D+1)^2}{(D+2)^2} \text{Tr}[R_\rho \mathbb{W} R_\rho \mathbb{W}(\rho \otimes \rho)] + \frac{1}{(D+2)^2} \text{Tr}[R_\rho^2(\rho \otimes \rho)] \\ &\quad - \frac{D+1}{(D+2)^2} \text{Tr}[R_\rho \mathbb{W} R_\rho(\rho \otimes \rho)] - \frac{D+1}{(D+2)^2} \text{Tr}[R_\rho^2 \mathbb{W}(\rho \otimes \rho)] \\ &= \frac{(D+1)^2}{(D+2)^2} (1 + 4p_2 + 2p_2^2 + 2p_3) + \frac{(1+2p_2)^2}{(D+2)^2} - \frac{2D+2}{(D+2)^2} (p_2 + 4p_3 + 4p_4) = \mathcal{O}(1). \end{aligned} \quad (\text{F27})$$

Using Eq. (F6), we have

$$\begin{aligned}
\mathbb{E}[\text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3) \text{Tr}(\hat{\rho}_1 \hat{\rho}_4 \hat{\rho}_2)] &= \text{Tr}\{\mathbb{E}[\hat{\rho} \otimes \hat{\rho}](I \otimes \rho) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}](\rho \otimes I)\} \\
&= \frac{(D+1)^2}{(D+2)^2} \text{Tr}[R_\rho \mathbb{W}(I \otimes \rho) R_\rho \mathbb{W}(\rho \otimes I)] + \frac{1}{(D+2)^2} \text{Tr}[R_\rho(I \otimes \rho) R_\rho(\rho \otimes I)] \\
&\quad - \frac{D+1}{(D+2)^2} \text{Tr}[R_\rho \mathbb{W}(I \otimes \rho) R_\rho(\rho \otimes I)] - \frac{D+1}{(D+2)^2} \text{Tr}[R_\rho(I \otimes \rho) R_\rho \mathbb{W}(\rho \otimes I)] \\
&= \frac{(D+1)^2}{(D+2)^2} [D(p_2 + 2p_3 + p_4) + 2p_3 + 2p_2 + p_2^2] + \frac{(1+2p_2)^2}{(D+2)^2} - \frac{2D+2}{(D+2)^2} (p_2 + 4p_3 + 4p_4) = \mathcal{O}(D).
\end{aligned} \tag{F28}$$

Using Eq. (F8), we have

$$\begin{aligned}
\mathbb{E}[\text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3) \text{Tr}(\hat{\rho}_2 \hat{\rho}_1 \hat{\rho}_4)] &= \text{Tr}\{\mathbb{E}[(\hat{\rho} \otimes \hat{\rho}')(\hat{\rho}' \otimes \hat{\rho})](\rho \otimes \rho)\} \\
&= \frac{(D+1)^3}{(D+2)^2} \text{Tr}[R_\rho \mathbb{W}(\rho \otimes \rho)] + \frac{1}{(D+2)^2} \text{Tr}[R_\rho \mathbb{W} R_\rho \mathbb{W}(\rho \otimes \rho)] - \frac{2(D+1)}{(D+2)^2} \text{Tr}[R_\rho^2 \mathbb{W}(\rho \otimes \rho)] \\
&\quad + \frac{D(D+1)^2}{(D+2)^2} \text{Tr}[(\rho \otimes \rho) \mathbb{W}(\rho \otimes \rho)] + \frac{(D+1)^2}{(D+2)^2} (1+p_2) \text{Tr}[\mathbb{W}(\rho \otimes \rho)] \\
&= \frac{(D+1)^3}{(D+2)^2} (p_2 + 4p_3 + 4p_4) + \frac{D(D+1)^2}{(D+2)^2} p_4 + \frac{(D+1)^2}{(D+2)^2} (1+p_2) p_2 \\
&\quad + \frac{1}{(D+2)^2} (1 + 4p_2 + 2p_2^2 + 2p_3) - \frac{2(D+1)}{(D+2)^2} (p_2 + 4p_3 + 4p_4) = \mathcal{O}(D).
\end{aligned} \tag{F29}$$

(4) There exist three pairs of identical indices, then the summand has the following two types.

Using Eq. (F6), we have

$$\begin{aligned}
\mathbb{E}[\text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3) \text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3)] &= \text{Tr}(\mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^3) \\
&= \frac{(D+1)^3}{(D+2)^3} \text{Tr}(R_\rho \mathbb{W} R_\rho \mathbb{W} R_\rho \mathbb{W}) - \frac{3(D+1)^2}{(D+2)^3} \text{Tr}(R_\rho^2 \mathbb{W} R_\rho \mathbb{W}) + \frac{3(D+1)}{(D+2)^3} \text{Tr}(R_\rho^3 \mathbb{W}) - \frac{1}{(D+2)^3} \text{Tr}(R_\rho^3) \\
&= \frac{(D+1)^3 + 3(D+1)}{(D+2)^3} (D + 6 + 12p_2 + 8p_3) \\
&\quad - \frac{1 + 3(D+1)^2}{(D+2)^3} [D^2 + 2D(3 + 3p_2 + p_3) + 6(1 + p_2)] = \mathcal{O}(D).
\end{aligned} \tag{F30}$$

Using Eqs. (F6) and (F8), we have

$$\begin{aligned}
\mathbb{E}[\text{Tr}(\hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3) \text{Tr}(\hat{\rho}_2 \hat{\rho}_1 \hat{\rho}_3)] &= \text{Tr}(\mathbb{E}[(\hat{\rho} \otimes \hat{\rho}')(\hat{\rho}' \otimes \hat{\rho})] \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]) \\
&= \frac{(D+1)^4}{(D+2)^3} \text{Tr}(R_\rho^2) + \frac{D(D+1)^3}{(D+2)^2} \text{Tr}[(\rho \otimes \rho) R_\rho] + \frac{(D+1)^3}{(D+2)^2} (1+p_2) \text{Tr}(R_\rho) \\
&\quad + \frac{3(D+1)}{(D+2)^3} \text{Tr}(R_\rho^3 \mathbb{W}) - \frac{2(D+1)^2 + 1}{(D+2)^3} \text{Tr}(R_\rho^3) - \frac{(D+1)^3}{(D+2)^3} \text{Tr}(R_\rho^2 \mathbb{W}) \\
&\quad - \frac{D(D+1)^2}{(D+2)^3} \text{Tr}[R_\rho(\rho \otimes \rho) \mathbb{W}] - \frac{(D+1)^2}{(D+2)^3} (1+p_2) \text{Tr}(R_\rho \mathbb{W})
\end{aligned}$$

$$\begin{aligned}
 &= \frac{(D+1)^4}{(D+2)^3} (D^2 + 4D + 2Dp_2 + 2) + \frac{D(D+1)^3}{(D+2)^3} (1 + 2p_2) \\
 &\quad + \frac{(D+1)^3}{(D+2)^2} (1 + p_2)(D^2 + 2Dp_2) + \frac{3(D+1)}{(D+2)^3} (D + 6 + 12p_2 + 8p_3) \\
 &\quad - \frac{2(D+1)^2 + 1}{(D+2)^3} [D^2 + 2D(3 + 3p_2 + p_3) + 6(1 + p_2)] - \frac{(D+1)^2}{(D+2)^3} (D + 4 + 4p_2) \\
 &\quad - \frac{D(D+1)^2}{(D+2)^3} (p_2 + 2p_3) - \frac{(D+1)^2}{(D+2)^3} (1 + p_2)(D + 2) = \mathcal{O}(D^3).
 \end{aligned} \tag{F31}$$

One can verify that in Eq. (F24), there are $\mathcal{O}(M^6), \mathcal{O}(M^5), \mathcal{O}(M^4), \mathcal{O}(M^3)$ summands of types 1, 2, 3, 4, respectively. Hence,

$$\text{Var}[\hat{p}_3] = \mathcal{O}(M^{-1}) + \mathcal{O}(M^{-2}D) + \mathcal{O}(M^{-3}D^3). \tag{F32}$$

To guarantee $\text{Var}[\hat{p}_3] = \mathcal{O}(1)$, we need $M = \mathcal{O}(D)$.

4. Variance of $\text{Tr}[\rho_A \text{Tr}_B(\rho^2)]$ by classical shadow

We estimate $t_2 \equiv \text{Tr}[\rho_A \text{Tr}_B(\rho^2)]$ using the classical shadow method in the following steps. Use M samples to generate M state estimators for ρ labeled as $\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_M$. Then use L samples to generate L state estimators for ρ_A labeled as $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_L$. The total sample cost is $L + M$.

Let $\mathcal{H}$ denote a local Hilbert space with $\dim(\mathcal{H}) = \sqrt{D}$ such that $\mathcal{H}_A, \mathcal{H}_B \cong \mathcal{H}$. Introduce a permutation $\pi^* \equiv (135)(24)$, and let $\mathbb{W}_{\pi^*}$ be the corresponding permutation operator on $\mathcal{H}^{\otimes 5}$. Then we have

$$\text{Tr}[(\rho \otimes \rho \otimes \rho_A) \mathbb{W}_{\pi^*}] = \text{Tr}[\rho_A \text{Tr}_B(\rho^2)] = t_2. \tag{F33}$$

Therefore, an estimator for t_2 can be constructed as follows:

$$\hat{t}_2 = \frac{1}{LM(M-1)} \sum_{\ell=0}^{L-1} \sum_{i=0, j \neq i}^{M-1} \text{Tr}[(\hat{\rho}_i \otimes \hat{\rho}_j \otimes \hat{a}_\ell) \mathbb{W}_{\pi^*}] = \frac{1}{LM(M-1)} \sum_{\ell=0}^{L-1} \sum_{i=0, j \neq i}^{M-1} \text{Tr}[(\hat{a}_\ell \otimes I_B) \hat{\rho}_i \hat{\rho}_j]. \tag{F34}$$

The variance of $\hat{t}_2$ is

$$\text{Var}[\hat{t}_2] = \frac{1}{L^2 M^2 (M-1)^2} \sum_{\ell_1, \ell_2}^{L-1} \sum_{i_1 \neq j_1, i_2 \neq j_2}^{M-1} \mathbb{E} \{ \text{Tr}[(\hat{a}_{\ell_1} \otimes I) \hat{\rho}_{i_1} \hat{\rho}_{j_1}] \text{Tr}[(\hat{a}_{\ell_2} \otimes I_B) \hat{\rho}_{i_2} \hat{\rho}_{j_2}] - t_2^2 \}. \tag{F35}$$

Note that

$$\sum_{\ell_1, \ell_2} \mathbb{E}[\hat{a}_{\ell_1} \otimes I_B \otimes \hat{a}_{\ell_2} \otimes I_B] = L(L-1) \rho_A \otimes I_B \otimes \rho_A \otimes I_B + L \mathbb{W}_{(23)} (\mathbb{E}[\hat{a} \otimes \hat{a}] \otimes I_B \otimes I_B) \mathbb{W}_{(23)}, \tag{F36}$$

$$\begin{aligned}
 \sum_{i_1 \neq j_1, i_2 \neq j_2} \mathbb{E}[\hat{\rho}_{i_1} \hat{\rho}_{j_1} \otimes \hat{\rho}_{i_2} \hat{\rho}_{j_2}] &= \mathcal{O}(M^4) \rho^2 \otimes \rho^2 + \mathcal{O}(M^2) \{ \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^2 + \mathbb{E}[(\hat{\rho} \otimes \hat{\rho}')(\hat{\rho}' \otimes \hat{\rho})] \} \\
 &\quad + \mathcal{O}(M^3) \{ \mathbb{E}[\hat{\rho} \otimes \hat{\rho}](\rho \otimes \rho) + (\rho \otimes \rho) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}] \} \\
 &\quad + \mathcal{O}(M^3) \{ (I \otimes \rho) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}](\rho \otimes I) + (\rho \otimes I) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}](I \otimes \rho) \}.
 \end{aligned} \tag{F37}$$

In the expansion of Eq. (F34), we have the following terms (recall that $a_n \equiv \text{Tr}(\rho_A^n)$).

Using Eq. (F7), we have

$$\begin{aligned}
& \text{Tr}\{(\rho_A \otimes I_B \otimes \rho_A \otimes I_B) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^2\} \\
&= \frac{(D+1)^2 + 1}{(D+2)^2} \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B) R_\rho^2] - \frac{2(D+1)}{(D+2)^2} \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B) R_\rho^2 \mathbb{W}_{(13)(24)}] \\
&= \frac{(D+1)^2 + 1}{(D+2)^2} \left[\left(\sqrt{D} + 2a_2 + t_2 \right)^2 + t_2^2 - 2(a_2 + t_2)^2 \right] \\
&\quad - \frac{2(D+1)}{(D+2)^2} \left\{ \sqrt{D}a_2 + 4a_3 + 2\text{Tr}[\rho_A^2 \text{Tr}_B(\rho^2)] + 2\text{Tr}[(\rho_A \otimes I_B)\rho(\rho_A \otimes I_B)\rho] \right\} = \mathcal{O}(D). \tag{F38}
\end{aligned}$$

Using Eq. (F6), we have

$$\begin{aligned}
& \text{Tr}\{(\rho_A \otimes I_B \otimes \rho_A \otimes I_B)(\rho \otimes \rho) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]\} \\
&= \frac{D+1}{D+2} \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B)(\rho \otimes \rho) R_\rho \mathbb{W}_{(13)(24)}] - \frac{1}{D+2} \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B)(\rho \otimes \rho) R_\rho] \\
&= \frac{D+1}{D+2} \left\{ \text{Tr}[(\rho_A \otimes I_B)\rho(\rho_A \otimes I_B)\rho] + 2\text{Tr}[(\rho_A \otimes I_B)\rho(\rho_A \otimes I_B)\rho^2] \right\} - \frac{1}{D+2} (a_2^2 + 2a_2t_2) = \mathcal{O}(1). \tag{F39}
\end{aligned}$$

Using Eq. (F8), we have

$$\begin{aligned}
& \text{Tr}\{(\rho_A \otimes I_B \otimes \rho_A \otimes I_B) \mathbb{E}[(\hat{\rho} \otimes \hat{\rho}')(\hat{\rho}' \otimes \hat{\rho})]\} \\
&= \frac{(D+1)^3}{(D+2)^2} \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B) R_\rho \mathbb{W}_{(13)(24)}] + \frac{D(D+1)^2}{(D+2)^2} \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B)(\rho \otimes \rho) \mathbb{W}_{(13)(24)}] \\
&\quad + \frac{(D+1)^2}{(D+2)^2} (1 + p_2) \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B) \mathbb{W}_{(13)(24)}] \\
&\quad + \frac{1}{(D+2)^2} \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B) R_\rho^2] - \frac{2(D+1)}{(D+2)^2} \text{Tr}[(\rho_A \otimes I_B \otimes \rho_A \otimes I_B) R_\rho^2 \mathbb{W}_{(13)(24)}] \\
&= \frac{(D+1)^3}{(D+2)^2} (\sqrt{D}a_2 + 2a_3) + \frac{D(D+1)^2}{(D+2)^2} \text{Tr}[(\rho_A \otimes I_B)\rho]^2 + \frac{(D+1)^2}{(D+2)^2} (1 + p_2) \sqrt{D}a_2 \\
&\quad + \frac{1}{(D+2)^2} \left[\left(\sqrt{D} + 2a_2 + t_2 \right)^2 + t_2^2 - 2(a_2 + t_2)^2 \right] \\
&\quad - \frac{2(D+1)}{(D+2)^2} \left\{ \sqrt{D}a_2 + 4a_3 + 2\text{Tr}[\rho_A^2 \text{Tr}_B(\rho^2)] + 2\text{Tr}[(\rho_A \otimes I_B)\rho(\rho_A \otimes I_B)\rho] \right\} = \mathcal{O}(D^{\frac{3}{2}}). \tag{F40}
\end{aligned}$$

For notational simplicity, we introduce the operator $R_A^{(12)} \equiv R_{\rho_A} \otimes I_B \otimes I_B$. Note that $\mathbb{E}[\hat{a} \otimes \hat{a}]$ takes a form very similar to that of $\mathbb{E}[\hat{\rho} \otimes \hat{\rho}]$, with the only difference being the Hilbert-space dimensions involved.

Using Eq. (F6), we obtain

$$\begin{aligned}
& \text{Tr}\{\mathbb{W}_{(23)}(\mathbb{E}[\hat{a} \otimes \hat{a}] \otimes I_B \otimes I_B) \mathbb{W}_{(23)}(\rho^2 \otimes \rho^2)\} \\
&= \frac{\sqrt{D}+1}{\sqrt{D}+2} \text{Tr}[\mathbb{W}_{(23)}(R_{\rho_A} \mathbb{W}_{(12)} \otimes I \otimes I) \mathbb{W}_{(23)}(\rho^2 \otimes \rho^2)] - \frac{1}{\sqrt{D}+2} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)}(\rho^2 \otimes \rho^2)] \\
&= \frac{\sqrt{D}+1}{\sqrt{D}+2} \left\{ \text{Tr}_A[\text{Tr}_B(\rho^2) \text{Tr}_B(\rho^2)] + 2\text{Tr}_A[\rho_A \text{Tr}_B(\rho^2) \text{Tr}_B(\rho^2)] \right\} - \frac{1}{\sqrt{D}+2} (p_2^2 + 2p_2t_2) = \mathcal{O}(1). \tag{F41}
\end{aligned}$$

Using Eqs. (F6) and (F7), we have

$$\begin{aligned}
& \text{Tr} \{ \mathbb{W}_{(23)}(\mathbb{E}[\hat{a} \otimes \hat{a}] \otimes I_B \otimes I_B) \mathbb{W}_{(23)} \mathbb{E}[\hat{\rho} \otimes \hat{\rho}]^2 \} \\
&= \frac{(D+1)^2 + 1}{(D+2)^2} \text{Tr}[\mathbb{W}_{(23)}(\mathbb{E}[\hat{a} \otimes \hat{a}] \otimes I_B \otimes I_B) \mathbb{W}_{(23)} R_\rho^2] - \frac{2(D+1)}{(D+2)^2} \text{Tr}[\mathbb{W}_{(23)}(\mathbb{E}[\hat{a} \otimes \hat{a}] \otimes I_B \otimes I_B) \mathbb{W}_{(23)} R_\rho^2 \mathbb{W}_{(13)(24)}] \\
&= \frac{(D+1)^2 + 1}{(D+2)^2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)} R_\rho^2] + \frac{2(D+1)}{(D+2)^2} \frac{1}{\sqrt{D} + 2} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)} R_\rho^2 \mathbb{W}_{(13)(24)}] \\
&\quad - \frac{(D+1)^2 + 1}{(D+2)^2} \frac{1}{\sqrt{D} + 1} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)} R_\rho^2] - \frac{2(D+1)}{(D+2)^2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)} R_\rho^2 \mathbb{W}_{(13)(24)}] \\
&= \frac{(D+1)^2 + 1}{(D+2)^2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \left(D^{\frac{3}{2}} + 2D + 2\sqrt{D} + 4\sqrt{D}a_2 \right) + \frac{2(D+1)}{(D+2)^2} \frac{1}{\sqrt{D} + 2} \left(D + 2\sqrt{D} + 4 + 4p_2 + 8a_2 + 8t_2 \right) \\
&\quad - \frac{(D+1)^2 + 1}{(D+2)^2} \frac{1}{\sqrt{D} + 1} \left(D^2 + 4D + 4\sqrt{D}a_2 + 2Dt_2 + 4a_2 + 2\sqrt{D}p_2 \right) \\
&\quad - \frac{2(D+1)}{(D+2)^2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \left[D^{\frac{3}{2}} + 2D(2 + a_2 + t_2) + 2\sqrt{D}(2 + p_2) + 4a_2 + 2p_2 \right] = \mathcal{O}\left(D^{\frac{3}{2}}\right). \tag{F42}
\end{aligned}$$

Using Eq. (F6), we have

$$\begin{aligned}
& \text{Tr} \{ \mathbb{W}_{(23)}(\mathbb{E}[\hat{a} \otimes \hat{a}] \otimes I_B \otimes I_B) \mathbb{W}_{(23)}(\rho \otimes \rho) \mathbb{E}[\hat{\rho} \otimes \hat{\rho}] \} \\
&= \frac{D+1}{D+2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)}(\rho \otimes \rho) R_\rho \mathbb{W}_{(13)(24)}] + \frac{1}{D+2} \frac{1}{\sqrt{D} + 2} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)}(\rho \otimes \rho) R_\rho] \\
&\quad - \frac{D+1}{D+2} \frac{1}{\sqrt{D} + 2} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)}(\rho \otimes \rho) R_\rho \mathbb{W}_{(13)(24)}] - \frac{1}{D+2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \text{Tr}[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)}(\rho \otimes \rho) R_\rho] \\
&= \frac{D+1}{D+2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \{ a_2 + 2\text{Tr}[(\rho_A \otimes \rho_B)\rho] + 2t_2 + 2\text{Tr}[(\rho_A \otimes \text{Tr}_B(\rho^2))\rho] + 2\text{Tr}[(\rho_A \otimes \rho_B)\rho^2] \} \\
&\quad - \frac{D+1}{D+2} \frac{1}{\sqrt{D} + 2} \{ a_2 + 2\text{Tr}[(\rho_A \otimes \rho_B)\rho] + 2t_2 + 2\text{Tr}[(\rho_A \otimes \text{Tr}_A(\rho^2) + \rho_A \otimes \rho_B)\rho] \} \\
&\quad + \frac{1}{D+2} \frac{1}{\sqrt{D} + 2} (1 + 4p_2 + 2t_2 + 2p_2a_2) - \frac{1}{D+2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \{ a_2 + 2a_3 + 2t_2 + 4\text{Tr}[\rho_A^2 \text{Tr}_B(\rho^2)] \} = \mathcal{O}(1). \tag{F43}
\end{aligned}$$

Using Eqs. (F6) and (F8), we have

$$\begin{aligned}
& \text{Tr} \{ \mathbb{W}_{(23)}(\mathbb{E}[\hat{a} \otimes \hat{a}] \otimes I_B \otimes I_B) \mathbb{W}_{(23)} \mathbb{E}[(\hat{\rho} \otimes \hat{\rho}')(\hat{\rho}' \otimes \hat{\rho})] \} \\
&= \frac{(D+1)^3}{(D+2)^2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)} R_\rho \mathbb{W}_{(13)(24)} \right] \\
&\quad + \frac{D(D+1)^2}{(D+2)^2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)}(\rho \otimes \rho) \mathbb{W}_{(13)(24)} \right] \\
&\quad + \frac{(D+1)^2}{(D+2)^2} (1 + p_2) \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)} \mathbb{W}_{(13)(24)} \right] \\
&\quad + \frac{1}{(D+2)^2} \frac{\sqrt{D} + 1}{\sqrt{D} + 2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)} R_\rho^2 \right] + \frac{2(D+1)}{(D+2)^2} \frac{1}{\sqrt{D} + 2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)} R_\rho^2 \mathbb{W}_{(13)(24)} \right]
\end{aligned}$$

$$\begin{aligned}
& - \frac{2(D+1)}{(D+2)^2} \frac{\sqrt{D}+1}{\sqrt{D}+2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(12)} \mathbb{W}_{(23)} R_{\rho}^2 \mathbb{W}_{(13)(24)} \right] \\
& - \frac{(D+1)^3}{(D+2)^2} \frac{1}{\sqrt{D}+2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)} R_{\rho} \mathbb{W}_{(13)(24)} \right] \\
& - \frac{D(D+1)^2}{(D+2)^2} \frac{1}{\sqrt{D}+2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)} (\rho \otimes \rho) \mathbb{W}_{(13)(24)} \right] \\
& - \frac{(D+1)^2}{(D+2)^2} (1+p_2) \frac{1}{\sqrt{D}+2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)} \mathbb{W}_{(13)(24)} \right] - \frac{1}{(D+2)^2} \frac{1}{\sqrt{D}+2} \text{Tr} \left[\mathbb{W}_{(23)} R_{\rho_A}^{(12)} \mathbb{W}_{(23)} R_{\rho}^2 \right]. \quad (\text{F44})
\end{aligned}$$

After computation, we prove that it equals

$$\begin{aligned}
& \frac{(D+1)^3}{(D+2)^2} \frac{\sqrt{D}+1}{\sqrt{D}+2} \left(D^{\frac{3}{2}} + 2D + 2\sqrt{D} + 2\sqrt{D}a_2 + 2 \right) \\
& + \frac{D(D+1)^2}{(D+2)^2} \frac{\sqrt{D}+1}{\sqrt{D}+2} \left\{ a_2 + \text{Tr}[(\rho_A \otimes \rho_B)\rho] + \text{Tr}(\rho_A \rho_B^2) \right\} \\
& + \frac{(D+1)^2}{(D+2)^2} (1+p_2) \frac{\sqrt{D}+1}{\sqrt{D}+2} \left(D^{\frac{3}{2}} + 2D \right) + \frac{1}{(D+2)^2} \frac{\sqrt{D}+1}{\sqrt{D}+2} \left(D^{\frac{3}{2}} + 2D + 2\sqrt{D} + 4\sqrt{D}a_2 \right) \\
& + \frac{2(D+1)}{(D+2)^2} \frac{1}{\sqrt{D}+2} \left(D + 2\sqrt{D} + 4 + 4p_2 + 8a_2 + 8t_2 \right) \\
& - \frac{2(D+1)}{(D+2)^2} \frac{\sqrt{D}+1}{\sqrt{D}+2} \left[D^{\frac{3}{2}} + 2D(2+a_2+t_2) + 2\sqrt{D}(2+p_2) + 4a_2 + 2p_2 \right] \\
& - \frac{(D+1)^3}{(D+2)^2} \frac{1}{\sqrt{D}+2} \left[D^2 + 2D^{\frac{3}{2}} + D(4+2p_2+4a_2+2t_2) \right] \\
& - \frac{(D+1)^3}{(D+2)^2} \frac{1}{\sqrt{D}+2} \left\{ \text{Tr}[(\rho \otimes \rho) \mathbb{W}_{(1243)}] + \text{Tr}[(\rho_A \otimes I_B \otimes I)(\rho \otimes \rho)(\mathbb{W}_{(1243)} + \mathbb{W}_{(1342)})] \right\} \\
& - \frac{(D+1)^2}{(D+2)^2} (1+p_2) \frac{1}{\sqrt{D}+2} [D + 2\sqrt{D}] - \frac{(D+1)^3}{(D+2)^2} \frac{1}{\sqrt{D}+2} \left[\sqrt{D}(4+2p_2) + 2 + 4a_2 \right] \\
& - \frac{1}{(D+2)^2} \frac{1}{\sqrt{D}+2} \left(D^2 + 4D + 4\sqrt{D}a_2 + 2Dt_2 + 4a_2 + 2\sqrt{D}p_2 \right) = \mathcal{O} \left(D^{\frac{5}{2}} \right). \quad (\text{F45})
\end{aligned}$$

Therefore,

$$\begin{aligned}
\text{Var}[\hat{t}_2] &= \frac{1}{L^2 M^4} \left[\mathcal{O}(L^2 M^2 D) + \mathcal{O}(L^2 M^3) + \mathcal{O}(L^2 M^2 D^{\frac{3}{2}}) + \mathcal{O}(LM^4) + \mathcal{O}(LM^2 D^{\frac{5}{2}}) \right] \\
&= \mathcal{O}(M^{-2} D) + \mathcal{O}(M^{-1}) + \mathcal{O}(M^{-2} D^{\frac{3}{2}}) + \mathcal{O}(L^{-1}) + \mathcal{O}(L^{-1} M^{-2} D^{\frac{5}{2}}). \quad (\text{F46})
\end{aligned}$$

To guarantee that $\text{Var}[\hat{t}_2] = \mathcal{O}(1)$, we need

$$M = \mathcal{O} \left(D^{\frac{3}{4}} \right), \quad LM^2 = \mathcal{O} \left(D^{\frac{5}{2}} \right). \quad (\text{F47})$$

If we choose $L = \mathcal{O}(D)$, then the total sample complexity is $L + M = \mathcal{O}(D)$.

- [1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81**, 865 (2009).
- [2] O. Gühne and G. Tóth, Entanglement detection, *Phys. Rep.* **474**, 1 (2009).
- [3] N. Friis, G. Vitagliano, M. Malik, and M. Huber, Entanglement certification from theory to experiment, *Nat. Rev. Phys.* **1**, 72 (2019).
- [4] A. Peres, Separability criterion for density matrices, *Phys. Rev. Lett.* **77**, 1413 (1996).
- [5] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: Necessary and sufficient conditions in terms of linear maps, *Phys. Lett. A* **283**, 1 (2001).
- [6] G. Aubrun, S. J. Szarek, and D. Ye, Phase transitions for random states and a semicircle law for the partial transpose, *Phys. Rev. A* **85**, 030302(R) (2012).
- [7] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, The randomized measurement toolbox, *Nat. Rev. Phys.* **5**, 9 (2023).
- [8] P. Cieřliński, S. Imai, J. Dziewior, O. Gühne, L. Knips, W. Laskowski, J. Meinecke, T. Paterek, and T. Vértesi, Analysing quantum systems with randomised measurements, *Phys. Rep.* **1095**, 1 (2024).
- [9] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, *et al.*, Mixed-state entanglement from local randomized measurements, *Phys. Rev. Lett.* **125**, 200501 (2020).
- [10] A. Neven, J. Carrasco, V. Vitale, C. Kokail, A. Elben, M. Dalmonte, P. Calabrese, P. Zoller, B. Vermersch, R. Kueng, *et al.*, Symmetry-resolved entanglement detection using partial transpose moments, *npj Quantum Inf.* **7**, 152 (2021).
- [11] X.-D. Yu, S. Imai, and O. Gühne, Optimal entanglement certification from moments of the partial transpose, *Phys. Rev. Lett.* **127**, 060504 (2021).
- [12] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Probing Rényi entanglement entropy via randomized measurements, *Science* **364**, 260 (2019).
- [13] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, *Nat. Phys.* **16**, 1050 (2020).
- [14] Y. Zhou, P. Zeng, and Z. Liu, Single-copies estimation of entanglement negativity, *Phys. Rev. Lett.* **125**, 200502 (2020).
- [15] B. M. Terhal and P. Horodecki, Schmidt number for density matrices, *Phys. Rev. A* **61**, 040301(R) (2000).
- [16] Z. Huang, L. Maccone, A. Karim, C. Macchiavello, R. J. Chapman, and A. Peruzzo, High-dimensional entanglement certification, *Sci. Rep.* **6**, 27637 (2016).
- [17] J. Bavaresco, N. Herrera Valencia, C. Klöckl, M. Pivoluska, P. Erker, N. Friis, M. Malik, and M. Huber, Measurements in two bases are sufficient for certifying high-dimensional entanglement, *Nat. Phys.* **14**, 1032 (2018).
- [18] M. Weilenmann, B. Dive, D. Trillo, E. A. Aguilar, and M. Navascués, Entanglement detection beyond measuring fidelities, *Phys. Rev. Lett.* **124**, 200502 (2020).
- [19] S. Morelli, M. Huber, and A. Tavakoli, Resource-efficient high-dimensional entanglement detection via symmetric projections, *Phys. Rev. Lett.* **131**, 170201 (2023).
- [20] N. K. H. Li, M. Huber, and N. Friis, High-dimensional entanglement witnessed by correlations in arbitrary bases, *npj Quantum Inf.* **11**, 50 (2025).
- [21] S. Imai, N. Wyderka, A. Ketterer, and O. Gühne, Bound entanglement from randomized measurements, *Phys. Rev. Lett.* **126**, 150501 (2021).
- [22] S. Liu, Q. He, M. Huber, O. Gühne, and G. Vitagliano, Characterizing entanglement dimensionality from randomized measurements, *PRX Quantum* **4**, 020324 (2023).
- [23] N. Wyderka and A. Ketterer, Probing the geometry of correlation matrices with randomized measurements, *PRX Quantum* **4**, 020325 (2023).
- [24] S. Liu, M. Fadel, Q. He, M. Huber, and G. Vitagliano, Bounding entanglement dimensionality from the covariance matrix, *Quantum* **8**, 1236 (2024).
- [25] D. Gross, K. Audenaert, and J. Eisert, Evenly distributed unitaries: On the structure of unitary designs, *J. Math. Phys.* **48**, 052104 (2007).
- [26] H. Zhu, R. Kueng, M. Grassl, and D. Gross, The Clifford group fails gracefully to be a unitary 4-design, *arXiv:1609.08172*.
- [27] E. Bannai, G. Navarro, N. Rizo, and P. H. Tiep, Unitary t -groups, *J. Math. Soc. Japan* **72**, 909 (2020).
- [28] O. Lib, S. Liu, R. Shekel, Q. He, M. Huber, Y. Bromberg, and G. Vitagliano, Experimental certification of high-dimensional entanglement with randomized measurements, *Phys. Rev. Lett.* **134**, 210202 (2025).
- [29] A. Sanpera, D. Bruß, and M. Lewenstein, Schmidt-number witnesses and bound entanglement, *Phys. Rev. A* **63**, 050301(R) (2001).
- [30] K. Schmüdgen, *The Moment Problem* (Springer, Cham, Switzerland, 2017).
- [31] K. Schmüdgen, Ten lectures on the moment problem, *arXiv:2008.12698*.
- [32] Z. Webb, The Clifford group forms a unitary 3-design, *Quantum Inf. Comput.* **16**, 1379 (2016).
- [33] H. Zhu, Multiqubit Clifford groups are unitary 3-designs, *Phys. Rev. A* **96**, 062336 (2017).
- [34] R. Kueng, H. Rauhut, and U. Terstiege, Low rank matrix recovery from rank one measurements, *Appl. Comput. Harmon. Anal.* **42**, 88 (2017).
- [35] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University, Cambridge, UK, 2010), 2nd ed.
- [36] M.-D. Choi, Positive linear maps on C^* -algebras, *Can. J. Math.* **24**, 520 (1972).
- [37] T. Takasaki and J. Tomiyama, On the geometry of positive maps in matrix algebras, *Math. Z.* **184**, 101 (1983).
- [38] J. Tomiyama, On the geometry of positive maps in matrix algebras. II, *Linear Algebra Appl.* **69**, 169 (1985).
- [39] D. Chruściński and G. Sarbicki, Entanglement witnesses: Construction, analysis and classification, *J. Phys. A: Math. Theor.* **47**, 483001 (2014).
- [40] G. Vidal and R. F. Werner, Computable measure of entanglement, *Phys. Rev. A* **65**, 032314 (2002).
- [41] A. W. Roberts, in *Handbook of Convex Geometry* (Elsevier, Amsterdam, Netherlands, 1993), p. 1081.
- [42] M. A. Nielsen, Conditions for a class of entanglement transformations, *Phys. Rev. Lett.* **83**, 436 (1999).

- [43] K. Życzkowski and H.-J. Sommers, Induced measures in the space of mixed quantum states, *J. Phys. A: Math. Gen.* **34**, 7111 (2001).
- [44] O. Gühne, Y. Mao, and X.-D. Yu, Geometry of faithful entanglement, *Phys. Rev. Lett.* **126**, 140503 (2021).
- [45] F. Bohnet-Waldraff, D. Braun, and O. Giraud, Entanglement and the truncated moment problem, *Phys. Rev. A* **96**, 032312 (2017).
- [46] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Quantum state tomography via compressed sensing, *Phys. Rev. Lett.* **105**, 150401 (2010).
- [47] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, Sample-optimal tomography of quantum states, in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, NY, 2016), pp. 913–925.
- [48] R. O’Donnell and J. Wright, Efficient quantum tomography, in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, NY, 2016), pp. 899–912.
- [49] S. Aaronson, Shadow tomography of quantum states, in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing* (Association for Computing Machinery, New York, NY, 2018), pp. 325–338.
- [50] S. Aaronson and G. N. Rothblum, Gentle measurement of quantum states and differential privacy, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (Association for Computing Machinery, New York, NY, 2019), pp. 322–333.
- [51] P. M. Q. Cruz and B. Murta, Shallow unitary decompositions of quantum Fredkin and Toffoli gates for connectivity-aware equivalent circuit averaging, *APL Quantum* **1**, 016105 (2024).
- [52] Z. Liu and F. Wei, Separation between entanglement criteria and entanglement detection protocols, *Phys. Rev. Res.* **7**, 033121 (2025).
- [53] B. Mallick, A. G. Maity, N. Ganguly, and A. S. Majumdar, Higher-dimensional-entanglement detection and quantum-channel characterization using moments of generalized positive maps, *Phys. Rev. A* **112**, 012416 (2025).
- [54] https://github.com/CYI1995/PRXQ_SN.
- [55] D. Grier, H. Pashayan, and L. Schaeffer, Sample-optimal classical shadows for pure states, *Quantum* **8**, 1373 (2024).