




Quantifying subspace entanglement with geometric measures

Xuanran Zhu ^{*}, Chao Zhang ^{*}, and Bei Zeng [†]

Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China



(Received 27 January 2024; revised 21 June 2024; accepted 25 June 2024; published 22 July 2024)

Determining whether a subspace spanned by certain quantum states is entangled and its entanglement dimensionality remains a fundamental challenge in quantum information science. This paper introduces a geometric measure of r -bounded rank, $E_r(\mathcal{S})$, for a given subspace \mathcal{S} . Derived from the established geometric measure of entanglement, this measure is specifically designed to assess the entanglement within \mathcal{S} . It not only serves as a tool for determining the entanglement dimensionality but also illuminates the subspace's capacity to preserve such entanglement. By employing developed nonconvex optimization techniques utilized in machine learning area, we can accurately calculate $E_r(\mathcal{S})$ within the manifold optimization framework. Our approach demonstrates notable advantages over existing hierarchical methods, PPT relaxation techniques, and the seesaw strategy, particularly by combining computational efficiency with broad applicability. More importantly, it paves the way for high-dimensional entanglement certification, which is crucial for numerous quantum information tasks. We showcase its effectiveness in validating high-dimensional entangled subspaces in bipartite systems, determining the border rank of multipartite pure states, and identifying genuinely or completely entangled subspaces.

DOI: [10.1103/PhysRevA.110.012452](https://doi.org/10.1103/PhysRevA.110.012452)

I. INTRODUCTION

Quantum entanglement, a foundational pillar of quantum physics, facilitates the nonlocal exchange of information amidst entangled particles, regardless of the spatial distance separating them. This fundamental phenomenon underpins transformative applications in realms such as quantum computing [1,2], cryptography [3], and teleportation [4]. A critical query within this domain pertains to the determination of whether a given subspace is entangled, and if so, discerning its entanglement dimensionality [5–15].

In the bipartite settings, entanglement dimensionality for a pure state can be characterized by Schmidt rank, which can be effectively determined via the Schmidt decomposition [16]. However, ascertaining the entanglement within a subspace presents a more intricate challenge [17]. This process of entanglement certification plays a pivotal role in numerous applications, such as the creation of entangled mixed states [18], construction of entanglement witnesses [19,20], quantum error correction [21,22], and the validation of protocols such as superdense coding [23].

The complexity of these calculations intensifies in a multipartite context [24], even when examining a single multipartite pure state. In contrast to bipartite pure states, which can be represented as matrices, multipartite pure states are depicted by tensors [25]. While Schmidt decomposition offers a way to determine Schmidt rank, the determination of tensor rank presents a far greater challenge. Specifically, tensor rank, defined as the minimum number of product terms necessary

for decomposition, is notoriously difficult to calculate. Its complexity is believed to be at least NP complete [26,27].

Meanwhile, the notion of border rank arises [28,29]. This concept acknowledges that certain tensors can be approximated asymptotically to arbitrary precision using other tensors with smaller tensor ranks—a phenomenon not possible with matrices. A prominent example is the W state in a three-qubit system, which has a tensor rank of 3 but can be approximated by states with a minimum tensor rank of 2, to any desired degree of precision [9]. In such instances, we denote the border rank of the W state as 2. From a pragmatic perspective, border rank may be more suitable for characterizing entanglement dimensionality in the multipartite scenario. Experimentally, a multipartite pure state with border rank r can always be simulated to any desired precision using states with tensor rank r , even if it has a higher tensor rank.

Remarkably, research into tensor or border ranks holds significance not only in the realm of quantum information due to the demand for high-dimensional entanglement in certain applications [30–33] but also in other disciplines, such as algebraic complexity theory [34–36]. Regarding multipartite subspaces, the principal focus rests on completely or genuinely entangled subspaces. The former is characterized as a subspace devoid of any fully product states [37], while the latter contains no product states across any bipartition [38,39]. The value of completely entangled subspaces stems from their capacity to discriminate pure quantum states locally [6,40]. Conversely, genuinely entangled subspaces have proven their worth in the sphere of quantum cryptography [41].

Significant progress has been made in the quest to certify the presence of entanglement within a subspace. One approach involves deriving lower bounds on the subspace's entanglement mathematically and then investigating under which conditions these bounds produce a positive value. This

^{*}These authors contributed equally to this work.

[†]Contact author: zengb@ust.hk

line of research has a rich history [42–48], with origins in Ref. [49]. An alternative strategy involves bounding the minimum entanglement using semidefinite programming (SDP), such as the positive partial transpose (PPT) relaxation [12,50] and symmetric extension [51]. However, the former mathematical approach may require complex or counterintuitive proofs and may not be broadly applicable. The latter is limited by its computational complexity and is incapable of handling certain scenarios, such as bound entanglement [52] and high-dimensional entanglement [53]. Recently, a new method was introduced that employs a hierarchy of linear systems [11]. Although this technique can certify high-dimensional entanglement within the given subspace, there remains room for improvement in its computational efficiency. Additionally, it cannot provide any desired entangled state.

In this work, entanglement dimensionality of a given subspace is characterized by a quantity called minimal rank. In the bipartite settings, it is defined as the minimum Schmidt rank of quantum states within the given subspace. And we generalize this definition to the multipartite scenario by replacing Schmidt rank with border rank, based on the consideration mentioned earlier. Here, we also present an intuitive, yet nontrivial approach to determine the minimal rank of a given subspace.

Building upon the geometric measure of entanglement (GME) [8,12,13,54], which quantifies the geometric distance between a given pure state $|\psi\rangle$ and fully product states, as follows:

$$E_G(|\psi\rangle) = 1 - \max_{|\varphi_{\text{prod}}\rangle} |\langle\varphi_{\text{prod}}|\psi\rangle|^2, \quad (1)$$

we generalize this concept to the geometric measure of r -bounded rank for a given subspace \mathcal{S} , denoted by $E_r(\mathcal{S})$, as the geometric distance between \mathcal{S} and σ_{r-1} , where σ_{r-1} represents the states with at most tensor rank $r-1$. The minimal rank $r(\mathcal{S})$ of a subspace \mathcal{S} can now be determined by finding the largest value of r such that $E_r(\mathcal{S}) > 0$ [if $E_2 = 0$ then $r(\mathcal{S}) = 1$]. Through further analysis, we find that this kind of geometric measure not only provides a criterion for certifying the entanglement dimensionality within a given subspace (from zero or nonzero), but also furnishes insights into the ability of the subspace to maintain such entanglement (indicated by the magnitude).

By employing the gradient back-propagation technique [55] from the field of machine learning, we are able to calculate this geometric measure through gradient descent with notable efficiency, which traditional convex optimization methods like SDP cannot handle. The core concept involves transforming the optimization problem defined on the complicated set σ_r into an optimization over a product manifold composed of several simpler manifolds. We then apply the concept of trivialization [56], established in the domain of manifold optimization [57,58], to address the original problem within an unconstrained Euclidean space.

We demonstrate its effectiveness in validating high-dimensional entangled subspaces in bipartite systems, determining the border rank of multipartite states, and identifying genuinely or completely entangled subspaces. The results reveal that this method has some attractive advantages compared to other known approaches, especially as it combines

TABLE I. Comparison between different methods for detecting entanglement: Hier (hierarchical method [11]), PPT (positive partial transpose relaxation [12,50]), and the method we propose based on GD (gradient descent approach). The presence of a check mark (✓) indicates the method’s capability to address the specified scenario, while a cross mark (×) implies its limitation. Entanglement is denoted as ENT for simplicity.

Comparison criterion		Hier	PPT	GD
Bipartite	low-dimensional ENT	✓	✓	✓
	bound ENT	✓	×	✓
	high-dimensional ENT	✓	×	✓
Multipartite	border rank of pure state	×	×	✓
	complete ENT	✓	✓	✓
	genuine ENT	✓	✓	✓
Quantify the robustness?		×	✓	✓
Time efficiency		✓	✓✓	✓✓✓

computational efficiency and broad applicability. More importantly, it paves the way for high-dimensional entanglement certification, which is crucial for numerous quantum information tasks. A preview of the main results is presented in Table I.

The structure of this paper is as follows. In Sec. II, we introduce some fundamental definitions related to bipartite and multipartite entanglement. In Sec. III, we propose the geometric measure of r -bounded rank and shed light on the connections among pure states, mixed states, and subspaces. We also briefly explain the key idea about the numerical calculation proposed in the manifold optimization framework, with the developed nonconvex optimization techniques. In Sec. IV, we illustrate the application of this type of geometric measure in various situations and also provide comparisons with other methods. Conclusions and outlooks are summarized and discussed in Sec. V.

II. PRELIMINARIES

Notation. In the paper, $|\psi\rangle$, $|\phi\rangle$, state, etc., all refer to normalized quantum states unless otherwise specified.

A. Bipartite entanglement

Schmidt decomposition. Let us begin with the simplest bipartite case and consider a Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. A state $|\psi\rangle \in \mathcal{H}$ is separable iff $|\psi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle$ for some pure states $|\phi_i\rangle \in \mathcal{H}_i$; otherwise it is called entangled. A useful tool for the characterization of bipartite entanglement is the Schmidt decomposition: any $|\psi\rangle \in \mathcal{H}$ can be written as

$$|\psi\rangle = \sum_{i=1}^r \lambda_i |e_i\rangle |f_i\rangle, \quad (2)$$

where the coefficients λ_i are positive numbers that can be ordered as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$, whose squares sum up to one, $\{|e_i\rangle\}$ and $\{|f_i\rangle\}$ are orthonormal bases of the local Hilbert spaces and $r \leq \min(d_1, d_2)$ is called the Schmidt rank of $|\psi\rangle$ representing the minimum number of terms for decomposing a state. It is equivalent to the singular value decomposition

(SVD) of the corresponding coefficient matrix. Therefore, in this paper, we will use matrix rank and Schmidt rank interchangeably, as they refer to the same concept.

Minimal rank of bipartite subspace. Given a subspace \mathcal{S} spanned by a set of bipartite states $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_m\rangle\}$, minimal rank of \mathcal{S} is defined as the following:

$$r(\mathcal{S}) = \min_{|\psi\rangle \in \mathcal{S}} R_S(|\psi\rangle),$$

where R_S represents the Schmidt rank. $|\psi\rangle \in \mathcal{S}$ means it can be derived by a combination of the states within \mathcal{S} , i.e., $|\psi\rangle = \sum_{i=1}^m c_i |\psi_i\rangle$, where c_i 's are some complex coefficients constrained to the normalization condition. If the minimal rank of a given subspace is larger than one, we say that the subspace is entangled.

B. Multipartite entanglement

Tensor decomposition. For a general multipartite pure state $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_n$, tensor rank is defined as the minimum number r such that there exist states $|\phi_i^{(k)}\rangle \in \mathcal{H}_k (1 \leq i \leq r, 1 \leq k \leq n)$ satisfying

$$|\psi\rangle = \sum_{i=1}^r \lambda_i |\phi_i^{(1)}\rangle \otimes |\phi_i^{(2)}\rangle \otimes \dots \otimes |\phi_i^{(n)}\rangle, \quad (3)$$

where the coefficients λ_i are positive but the different product terms may not be orthonormal to each other.

Tensor rank is a legitimate entanglement measure generalized from the Schmidt rank [10,59] since it is a strictly nonincreasing quantity under local operations (even stochastic) [60]. It provides a method of detecting multipartite entanglement: a multipartite pure state is entangled if and only if its tensor rank is larger than one.

Border rank is defined as the minimum number r of product terms that are sufficient to approximate the given tensor asymptotically with arbitrarily small error (in this paper, we choose the geometric measure).

A famous example of the gap between tensor rank and border rank is that the tensor rank of the W state in three-qubit is 3 while its border rank is 2 since

$$|W\rangle = \frac{1}{\sqrt{3}} \left[\lim_{t \rightarrow 0} \frac{1}{t} ((|0\rangle + t|1\rangle)^{\otimes 3} - |0\rangle^{\otimes 3}) \right]. \quad (4)$$

For any $t \neq 0$, the tensor on the right-hand side has rank 2 and in the limit $t \rightarrow 0$, it becomes W state. For brevity, in the following, we will denote the tensor and border rank as R and \underline{R} separately.

The gap between these two different ranks appears due to the nonexistence of a best rank- r approximation for an arbitrary tensor A [61], implying $R(A) > \underline{R}(A)$. However, there exist two notable exceptions where $R = \underline{R}$: the cases $n = 2$ (the given tensors are matrices) and $r = 1$ (approximation by rank-1 tensors). The former implies $R(A) = \underline{R}(A)$ for a matrix A while the latter means that $\underline{R}(A) = 1$ is equivalent to with $R(A) = 1$.

Minimal rank of multipartite subspace. Given a subspace \mathcal{S} spanned by a set of multipartite states $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_m\rangle\}$, minimal rank of \mathcal{S} is defined as the following:

$$r(\mathcal{S}) = \min_{|\psi\rangle \in \mathcal{S}} \underline{R}(|\psi\rangle).$$

In this work, we choose the border rank as the definition of minimal rank in the multipartite settings because, as mentioned before, experimentally, a multipartite pure state of border rank r can always be simulated to any desired precision using those states of tensor rank r even if it has higher tensor rank. This definition is also consistent across both bipartite and multipartite cases, as the Schmidt rank is equivalent to the border rank for bipartite states.

Completely entangled subspace. Consider a subspace $\mathcal{S} \subset \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_n$, we call \mathcal{S} is completely entangled iff it contains no fully product state. Due to the existence of best rank-1 approximation as mentioned above, we know it is equivalent to the detection of subspaces with $r(\mathcal{S}) > 1$.

Genuinely entangled subspace. A subspace $\mathcal{S} \subset \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_n$ is genuinely entangled iff it contains no product state for any bipartition $K|K^c$, where K is a subset of $\{1, 2, \dots, n\}$ and K^c denotes the complementary set. For example, for a tripartite state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, there are three possible bipartition $AB|C, AC|B$, and $BC|A$. The genuine entanglement of a subspace is equivalent to $r(\mathcal{S}) > 1$ for any bipartition $K|K^c$.

III. METHODOLOGY

A. Geometric measure for entangled subspace

Geometric measure for a pure state $|\psi\rangle$ is defined through the following general formula:

$$E(|\psi\rangle) = 1 - \max_{|\varphi\rangle \in V} |\langle \varphi | \psi \rangle|^2,$$

where V is chosen according to the specific purpose.

In order to quantify entanglement dimensionality via minimal rank, we generalize the GME defined in Eq. (1) to the geometric measure of r -bounded rank

$$E_r(|\psi\rangle) = 1 - \max_{|\varphi\rangle \in \sigma_{r-1}} |\langle \varphi | \psi \rangle|^2. \quad (5)$$

Here, σ_{r-1} is the states with at most tensor rank $r - 1$. Obviously, it can be reduced into the form of geometric measure of entanglement when $r = 2$, i.e., $E_G(|\psi\rangle) = E_2(|\psi\rangle)$.

For a bipartite state $|\psi\rangle$ it has been shown that [8]

$$E_r(|\psi\rangle) = 1 - (\lambda_1^2 + \lambda_2^2 + \dots + \lambda_{r-1}^2),$$

where λ_i is the coefficients after the Schmidt decomposition [see Eq. (2)]. As for a multipartite state $|\psi\rangle$ of border rank r , r is the largest value that makes $E_r(|\psi\rangle) > 0$. Thus, this measure can serve as a tool for certifying the border rank of a given state.

We can generalize this geometric measure, defined for a pure state, to a subspace \mathcal{S} as the following:

$$\begin{aligned} E_r(\mathcal{S}) &= \min_{|\psi\rangle \in \mathcal{S}} E_r(|\psi\rangle) \\ &= \min_{|\psi\rangle \in \mathcal{S}} (1 - \max_{|\varphi\rangle \in \sigma_{r-1}} |\langle \varphi | \psi \rangle|^2) \\ &= 1 - \max_{|\varphi\rangle \in \sigma_{r-1}} \max_{|\psi\rangle \in \mathcal{S}} |\langle \varphi | \psi \rangle|^2 \\ &= 1 - \max_{|\varphi\rangle \in \sigma_{r-1}} \langle \varphi | \mathcal{P}_{\mathcal{S}} | \varphi \rangle \\ &= \min_{|\varphi\rangle \in \sigma_{r-1}} \langle \varphi | \mathcal{P}_{\mathcal{S}}^\perp | \varphi \rangle, \end{aligned} \quad (6)$$

where \mathcal{P}_S projects onto S and \mathcal{P}_S^\perp onto the orthogonal complementary space S^\perp . The most crucial step is the transition from the third line to the fourth line follows from the fact that for a given $|\varphi\rangle$, the vector from S maximizing the quantity will be the projection of $|\varphi\rangle$ onto S . Similarly, for a subspace S of minimal rank r , r is the largest value that makes $E_r(S) > 0$.

Besides the certification of entanglement dimensionality, E_r also characterizes the robustness of entangled subspaces due to the following theorem (some similar consideration appears in Ref. [42]):

Theorem 1 (Robustness of entangled subspace). Given a subspace $S \subset \mathcal{H}_1 \otimes \mathcal{H}_2$ with $r(S) \geq r$, for any perturbation $U = e^{-iH}$ with the trace norm of H less than $E_r(S)^{\frac{1}{2}}$, the subspace after the perturbation S' will maintain the minimal rank $r(S') \geq r$.

Proof. Suppose S is spanned by $\{\phi_1, \phi_2, \dots, \phi_m\}$, the subspace after some perturbation U is defined as S' , spanned by $\{U\phi_1, U\phi_2, \dots, U\phi_m\}$. It is easy to prove that

$$\mathcal{P}_{S'}^\perp = U\mathcal{P}_S^\perp U^\dagger.$$

For convenience, we denote the square root of $\langle x, \mathcal{P}_S^\perp x \rangle$ as $F_S(x)$. Obviously, there exist the following inequalities:

$$F_S(x) + F_S(y) \geq F_S(x+y) \geq F_S(x) - F_S(y),$$

$$F_S\left(\sum_i a_i x_i\right) \leq \sum_i |a_i| \cdot F_S(x_i),$$

where x, y, x_i can be unnormalized states and a_i 's are complex numbers. Thus, the square root of the geometric measure of r -bounded rank for S'

$$\begin{aligned} E_r(S')^{\frac{1}{2}} &= \min_{|\varphi\rangle \in \sigma_{r-1}} (\langle \varphi | \mathcal{P}_{S'}^\perp | \varphi \rangle)^{\frac{1}{2}} \\ &= \min_{|\varphi\rangle \in \sigma_{r-1}} (\langle \varphi | U \mathcal{P}_S^\perp U^\dagger | \varphi \rangle)^{\frac{1}{2}} \\ &\geq \min_{|\varphi\rangle \in \sigma_{r-1}} [F_S(|\varphi\rangle) - F_S((I - U^\dagger)|\varphi\rangle)] \\ &= E_r(S)^{\frac{1}{2}} - \max_{|\varphi\rangle \in \sigma_{r-1}} F_S((I - U^\dagger)|\varphi\rangle). \end{aligned}$$

Here, $I - U^\dagger$ can be diagonalized into $\sum_i (1 - e^{ih_i})|u_i\rangle\langle u_i|$, where h_i is the eigenvalue of Hamiltonian H . Then

$$\begin{aligned} &\max_{|\varphi\rangle \in \sigma_{r-1}} F_S((I - U^\dagger)|\varphi\rangle) \\ &= \max_{|\varphi\rangle \in \sigma_{r-1}} F_S\left(\sum_i (1 - e^{ih_i})|u_i\rangle\langle u_i|\varphi\rangle\right) \\ &\leq \max_{|\varphi\rangle \in \sigma_{r-1}} \sum_i |1 - e^{ih_i}| \cdot F_S(|u_i\rangle\langle u_i|\varphi\rangle) \\ &\leq \sum_i |1 - e^{ih_i}| \leq \sum_i |h_i| = \|H\|_{tr}. \end{aligned}$$

Therefore, if the trace norm of Hamiltonian H is less than $E_r(S)^{\frac{1}{2}}$, $E_r(S')$ will be nonzero, i.e., $r(S') \geq r$. ■

A simple geometric interpretation can be made for this theorem, as shown in Fig. 1: notice that $E_r(S)^{\frac{1}{2}}$ characterizes the sine value of the smallest angle θ^* between the states in σ_{r-1} and states in the given subspace S , see Eq. (5). A perturbation $U = e^{-iH}$ can be considered as a rotation operator in high-dimensional space, the trace norm of H indicates its rotation

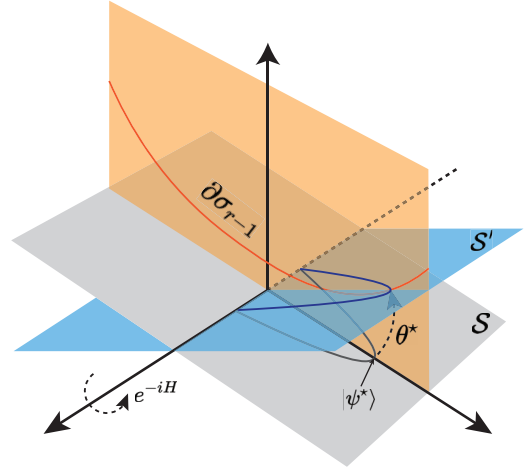


FIG. 1. Geometric interpretation of the robustness of entangled subspace. The gray line represents the quantum states within the subspace S , and $|\psi^*\rangle$ is the closest state to the σ_{r-1} under the geometric measure. The blue line represents the quantum states within the subspace S' after the perturbation. The orange line denotes the boundary of σ_{r-1} .

angle. In order to maintain its entanglement dimensionality, we need to limit the rotation angle $\theta \leq \theta^*$.

The definition can also be extended into mixed states as follows,

$$E_r(\rho) = \min_{\{p_i, |\psi_i\rangle\}} \sum_i p_i E_r(|\psi_i\rangle),$$

where the minimum is computed over all pure state ensembles, i.e., $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Obviously, ρ is separable iff $E_2(\rho) = 0$, i.e., it can be decomposed into a product state ensemble. It is known that $E_2(\rho) \geq E_2[\text{supp}(\rho)]$, where $\text{supp}(\rho)$ is the support space of the density matrix ρ [62]. This result can be generalized to the high-dimensional entangled cases easily, i.e., $E_r(\rho) \geq E_r[\text{supp}(\rho)]$ for $r > 2$.

Based on Theorem 1 and the above definition for mixed states, we can further conclude that quantum states supported on the robust entangled subspaces are also robust:

Theorem 2 (Formation of robust entangled states). Given an entangled state $\rho \in \mathcal{H}_1 \otimes \mathcal{H}_2$ with the support space $\text{supp}(\rho)$, $E_2[\text{supp}(\rho)] > 0$. For any perturbation $U = e^{-iH}$ with the trace norm of H less than $E_2[\text{supp}(\rho)]^{\frac{1}{2}}$, the mixed state after the perturbation $\rho' = U\rho U^\dagger$ will still be entangled.

Proof. Suppose $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, the corresponding support space $\text{supp}(\rho)$ is spanned by $\{|\psi_i\rangle\}$. After the perturbation, it becomes $\rho' = \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger$, the support space becomes $\text{supp}(\rho')$ spanned by $\{U|\psi_i\rangle\}$. According to Theorem 1, if $\|H\|_{tr} < E_2[\text{supp}(\rho)]^{\frac{1}{2}}$, then $E_2(\rho') \geq E_2[\text{supp}(\rho')] > 0$. ■

This theorem also offers a feasible way to construct robust entangled states from the robust entangled subspace, which can also be generalized to high-dimensional entanglement with $r > 2$.

In conclusion, this kind of geometric measure serves a dual purpose: it provides a criterion for certifying the entanglement dimensionality within a given subspace, quantified by the

minimal rank, and offers insights into the robustness of the subspace and mixed states supported on it.

B. Parametrization and optimization

Owing to the challenges in characterizing σ_r (states with at most tensor rank r) and the nonconvex nature of optimization, previous work sidestepped a direct computation of E_r . Instead, it focused on exploring the lower bound of E_r through mathematical derivation, or by relaxing it in a manner that allows for solution via the semidefinite programming (SDP) method, such as the positive partial transpose (PPT) relaxation:

$$\begin{aligned} E_2(\mathcal{S}) &= \min_{|\psi_{\text{prod}}\rangle} \langle \psi_{\text{prod}} | \mathcal{P}_{\mathcal{S}}^\perp | \psi_{\text{prod}} \rangle \\ &= \min_{|\psi_{\text{prod}}\rangle} \text{tr}[\mathcal{P}_{\mathcal{S}}^\perp |\psi_{\text{prod}}\rangle \langle \psi_{\text{prod}}|] \\ &\geq \min_{\substack{\rho \geq 0 \\ \forall_i \rho^T_i \geq 0}} \text{tr}[\mathcal{P}_{\mathcal{S}}^\perp \rho], \end{aligned} \quad (7)$$

where $|\psi_{\text{prod}}\rangle$ denotes the fully product state in σ_1 and ρ^T_i denotes the partial transpose of ρ with respect to the i th local Hilbert space \mathcal{H}_i .

In the following, we will introduce a parametrization strategy to directly characterize the set σ_r in the manifold optimization framework [57,58]. First, we can consider the set of unnormalized states with tensor rank r , denoted by $\tilde{\sigma}_r$, for every $|\tilde{\varphi}_r\rangle \in \tilde{\sigma}_r$, it can be written as

$$|\tilde{\varphi}_r\rangle = \sum_{i=1}^r \tilde{\lambda}_i |\phi_i^{(1)}\rangle \otimes |\phi_i^{(2)}\rangle \otimes \cdots \otimes |\phi_i^{(n)}\rangle.$$

For simplicity, we denote $\tilde{\lambda} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_r) \in \mathbb{R}_+^r$ and $|\phi_i\rangle = \otimes_{k=1}^n |\phi_i^{(k)}\rangle \in \mathcal{B} = \mathbb{S}^{2d_1-1} \times \mathbb{S}^{2d_2-1} \times \cdots \times \mathbb{S}^{2d_n-1}$, where \mathbb{R}_+^r represents the r -dimensional manifold of positive numbers and \mathcal{B} denotes the product manifold composed of a series of $(2d_k - 1)$ -dimensional sphere manifold \mathbb{S}^{2d_k-1} (d_k is the local dimension of k th Hilbert space). Then the original problem [Eq. (6)] can be formulated numerically as an optimization over the manifold \mathcal{M}_{r-1}

$$E_r(\mathcal{S}) = \min_{x \in \mathcal{M}_{r-1}} \langle \varphi_{r-1}(x) | \mathcal{P}_{\mathcal{S}}^\perp | \varphi_{r-1}(x) \rangle = \min_{x \in \mathcal{M}_{r-1}} f_{\mathcal{S}}(x).$$

Here, $|\varphi_{r-1}(x)\rangle$ is the normalized state from $|\tilde{\varphi}_{r-1}\rangle$ via adjusting $\tilde{\lambda}$, where $x = (\tilde{\lambda}, \phi_1, \phi_2, \dots, \phi_{r-1}) \in \mathcal{M}_{r-1} = \mathbb{R}_+^{r-1} \times \mathcal{B}^{r-1}$. And here, $f_{\mathcal{S}}$ is a simplified notation for the related function determined by \mathcal{S} .

Trivialization [56] provides a fresh and efficient perspective for manifold optimization, which requires a mapping g from the free Euclidean space to the manifold \mathcal{M}_{r-1} , such that

$$E_r(\mathcal{S}) = \min_{x \in \mathcal{M}_{r-1}} f_{\mathcal{S}}(x) = \min_{\tilde{\theta} \in \mathbb{R}^D} f_{\mathcal{S}}[g(\tilde{\theta})].$$

The problem is now transferred into an unconstrained optimization problem over the D -dimensional Euclidean space. Thus, the remaining task is to find such a satisfying mapping g . Since the manifold \mathcal{M}_r is a product manifold composed of \mathbb{R}_+^r and \mathcal{B}^r , we can consider them separately. For the positive numbers $\tilde{\lambda} \in \mathbb{R}_+^r$, we can use the SOFTPLUS mapping [63] to

ensure the positivity

$$\tilde{\lambda} = \log(1 + e^{\tilde{y}}),$$

where \tilde{y} is a r -dimensional real vector. As for each normalized states $|\phi_i^{(k)}\rangle \in \mathbb{S}^{2d_k-1}$, we can obtain them by mapping some d_k -dimensional real vectors $\vec{\alpha}_i, \vec{\beta}_i$ onto the sphere manifold \mathbb{S}^{2d_k-1} as follows:

$$|\phi_i^{(k)}\rangle = \frac{\vec{\alpha}_i + i\vec{\beta}_i}{\|\vec{\alpha}_i + i\vec{\beta}_i\|_2}.$$

Given these two fundamental mappings, finally we are able to map a free Euclidean space \mathbb{R}^D onto the manifold \mathcal{M}_{r-1} with the dimension

$$D = \left(2 \sum_k d_k + 1\right)(r - 1).$$

As for the projector $\mathcal{P}_{\mathcal{S}}^\perp$, it can be constructed as

$$\mathcal{P}_{\mathcal{S}}^\perp = I - \mathcal{P}_{\mathcal{S}} = I - \sum_{i=1}^{d_{\mathcal{S}}} |e_i\rangle \langle e_i|,$$

where $\{|e_i\rangle\}$ is the orthonormal basis of the given subspace \mathcal{S} that can be obtained through Gram-Schmidt process. Then we can define a loss function $\mathcal{L}_r(\tilde{\theta}; \mathcal{S}) = \langle \varphi_r(\tilde{\theta}) | \mathcal{P}_{\mathcal{S}}^\perp | \varphi_r(\tilde{\theta}) \rangle$ to compute $E_r(\mathcal{S})$. The relevant algorithm is summarized in Algorithm 1.

ALGORITHM 1. Geometric measure of r -bounded rank E_r .

Input: A subspace $\mathcal{S} \subset \otimes_{k=1}^n \mathcal{H}_k$ with $\dim(\mathcal{H}_k) = d_k$, spanned by $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_m\rangle\}$, the level of entanglement r , converge tolerance ϵ , the number of trials N

Output: Geometric measure of r -bounded rank $E_r(\mathcal{S})$

Orthogonalize the set $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_m\rangle\}$ through the Gram-Schmidt process, then obtain the projector of the subspace $\mathcal{P}_{\mathcal{S}}$

$\mathcal{P}_{\mathcal{S}}^\perp = I - \mathcal{P}_{\mathcal{S}}$

$t \leftarrow 1$

while $t \leq N$ **do**

Initialize a random real vector $\vec{\theta} \in \mathbb{R}^D$,

$D = (2 \sum_k d_k + 1)(r - 1)$

while E_r has not converged under tolerance ϵ **do**

Compute

$\mathcal{L}_{r-1}(\vec{\theta}; \mathcal{S}) = \langle \varphi_{r-1}(\vec{\theta}) | \mathcal{P}_{\mathcal{S}}^\perp | \varphi_{r-1}(\vec{\theta}) \rangle$

Update $\vec{\theta}$ based on the gradient descent

end

$E_r^{(t)} = \mathcal{L}_{r-1}(\vec{\theta}; \mathcal{S})$

$t \leftarrow t + 1$

end

$E_r(\mathcal{S}) = \min\{E_r^{(1)}, \dots, E_r^{(N)}\}$

To minimize the loss function, we adopt the gradient-based optimization method with the gradients obtained from gradient back-propagation provided in PYTORCH deep learning

framework [55]. Different from the common neural networks, there is no randomness in our loss function, so the classical optimization method, L-BFGS-B implemented in SCIPY package [64], should converge faster than those commonly used in training neural networks (e.g., SGD, Adam). However, gradient-based methods do not guarantee convergence to the global minimum, which provide upper bounds in principle. In order to obtain accurate values, we employ a probabilistic approach: minimizing the loss with randomly initialized parameters across N trials. Here, N can be considered as a hyperparameter in the optimization process.

IV. RESULTS

In this section, we demonstrate the application of the proposed method across various scenarios, including both bipartite and multipartite cases. We primarily compare our results with other established methods, such as the PPT relaxation [12,50] described in Eq. (7) and the hierarchical method [11]. Additionally, we compare the seesaw strategy [65] with our approach, which also provides upper bounds for GME. For each optimization, we fix the trial number $N = 3$ and the tolerance $\epsilon = 10^{-10}$ in the optimization. Numerical results show that this choice is effective enough but may not be optimal for certain cases.

A. Bipartite cases

1. Subspace with analytically known E_r

Example 1. The subspace $\mathcal{S}_{2 \times d}^\theta \subset \mathcal{H}_A \otimes \mathcal{H}_B$ with $d_A = 2$ and $d_B = d$, is given by the span of the following vectors [12]:

$$|\psi_i\rangle_{AB} = a|0\rangle_A|i\rangle_B + b|1\rangle_A|i+1\rangle_B, \quad (8)$$

where $i = 0, 1, \dots, d-2$, with $a = \cos(\theta/2)$ and $b = \exp(i\xi) \sin(\theta/2)$, $\theta \in (0, \pi)$, $\xi \in [0, 2\pi)$.

Clearly, the dimension of $\mathcal{S}_{2 \times d}^\theta$ is $d-1$, which is also the maximal available dimension of an entangled subspace in this scenario. Furthermore, it has been proved that

$$E_2(\mathcal{S}_{2 \times d}^\theta) = \frac{1}{2} \left[1 - \sqrt{1 - \sin^2 \theta \sin^2 \left(\frac{\pi}{d} \right)} \right].$$

In Fig. 2, we plot $E_2(\mathcal{S}_{2 \times d}^\theta)$ as a function of θ for different dimensions and compare the analytical results with those obtained via the gradient descent and PPT relaxation methods. For convenience, we fix $\xi = 0$. As observed, for this kind of low-dimensional entanglement, both methods provide accurate results.

The robustness of different entangled subspaces can also be explored, as depicted in Fig. 3. Three different subspaces ($\theta = \frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{6}$) with $d = 3$ are selected for comparison. We randomly generate 1000 unitary perturbations $U = e^{-iH}$ for each different trace norm of H . Utilizing gradient descent strategy, we compute the minimum geometric measure E_2 after the perturbations with different trace norms. A nonzero value of E_2 indicates that the subspace remains entangled after perturbations. As Theorem 1 proves, entangled subspaces with larger E_2 values are more robust against perturbations.

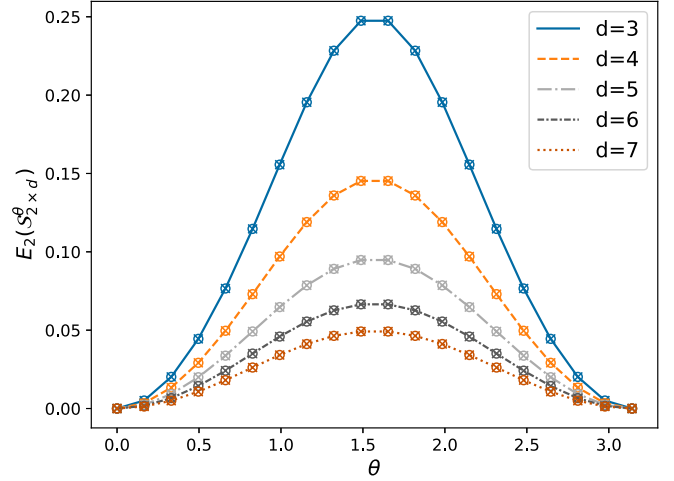


FIG. 2. $E_2(\mathcal{S}_{2 \times d}^\theta)$ as a function of θ . Different lines denote the analytical results of different dimensions, circles denote the results obtained by gradient descent, and crosses represent the results from the PPT relaxation. Numerical results from both two different methods match the analytical results.

2. Bound and high-dimensional entanglement

Although the PPT relaxation can give accurate E_2 for most subspaces in practice, it fails when faced with bound entanglement and cannot estimate the geometric measure of higher bounded rank, i.e., $E_r(r > 2)$.

Example 2. One famous approach to construct bound entangled states is using the unextendible product basis (UPB) [37], as follows:

$$\rho = \frac{1}{d_A d_B - d_S} \mathcal{P}_S^\perp,$$

where d_A, d_B are the local dimensions for bipartite states, \mathcal{S} represents the subspace spanned by the given UPB and d_S is the dimension of that subspace. For example, we can consider

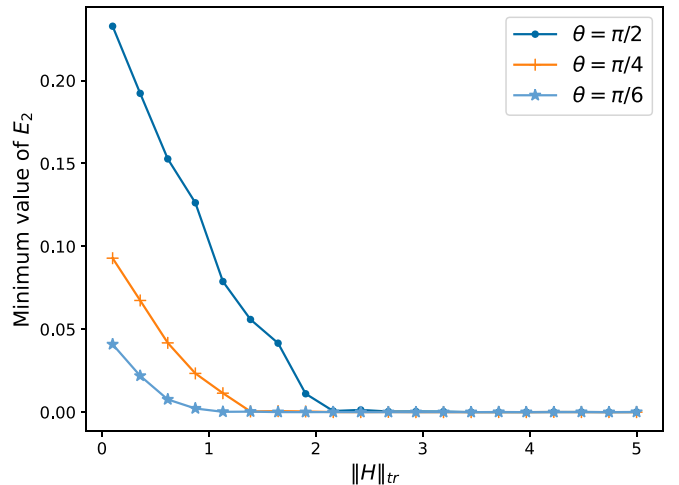


FIG. 3. Minimum values of E_2 after the random perturbations $U = e^{-iH}$ versus different trace norm of H . A nonzero value of E_2 means the perturbed subspace maintains its entanglement. Larger $E_2(\mathcal{S})$ brings more robustness as expected.

the five-state tiles UPB [66] (here we omit normalization for brevity),

$$\begin{aligned} \mathcal{S}_{\text{tiles}} = \text{span}\{ & |0\rangle \otimes (|0\rangle - |1\rangle), |2\rangle \otimes (|1\rangle - |2\rangle), \\ & (|0\rangle - |1\rangle) \otimes |2\rangle, (|1\rangle - |2\rangle) \otimes |0\rangle, (|0\rangle + |1\rangle + |2\rangle) \\ & \otimes (|0\rangle + |1\rangle + |2\rangle)\} \subset \mathcal{H}_A \otimes \mathcal{H}_B. \end{aligned}$$

As we have mentioned before, the geometric measure of the support space can give a lower bound for the geometric measure of the state, i.e., $E_r(\rho) \geq E_r[\text{supp}(\rho)]$. For the state ρ_{tiles} constructed by the tiles UPB, we can estimate $E_2[\text{supp}(\rho_{\text{tiles}})]$ by the PPT relaxation, which gives around 10^{-12} close to 0 while the gradient descent gives around 0.0284. That means PPT cannot detect bound entanglement while the gradient descent does.

Example 3. Let $d_A = d_B = 4$ and consider the following mixed state from [11]:

$$\rho = \frac{1}{3} \sum_{i=1}^3 |\psi_i\rangle\langle\psi_i| \in \mathcal{H}_A \otimes \mathcal{H}_B,$$

where (we omit normalization for brevity)

$$\begin{aligned} |\psi_1\rangle &= |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle + |2\rangle \otimes |2\rangle + |3\rangle \otimes |3\rangle, \\ |\psi_2\rangle &= |0\rangle \otimes |1\rangle + |1\rangle \otimes |2\rangle + |2\rangle \otimes |3\rangle + |3\rangle \otimes |0\rangle, \\ |\psi_3\rangle &= |0\rangle \otimes |2\rangle + |1\rangle \otimes |3\rangle + |2\rangle \otimes |0\rangle - |3\rangle \otimes |1\rangle. \end{aligned}$$

By using the PPT relaxation, we figure out that $E_2(\rho) \geq 0.382367$, indicates the presence of entanglement. Further, by computing the geometric measure of higher bounded rank via gradient descent, we can conclude that ρ is high-dimensional entangled since $E_3(\rho) \geq 0.06558$, i.e., it can only be an ensemble contains at least rank-3 entangled states. This information is not available for the PPT relaxation.

B. Multipartite cases

1. Border rank of pure states

The hierarchical method can also detect bound or high-dimensional entanglement in the bipartite scenario. However, both the PPT relaxation and hierarchical method fall short when certifying high-dimensional entanglement in the multipartite settings even for a single multipartite pure state $|\psi\rangle$. Conversely, through gradient descent, we can efficiently obtain the value of $E_r(|\psi\rangle)$, which provides information about the high-dimensional entanglement, i.e., border rank of the given state.

Example 4. Permutation symmetric states in n -qubit system are given by the Dicke states, which are mathematically expressed as the sum of all permutations of computational basis states with $n - k$ qubits being $|0\rangle$ and k being $|1\rangle$:

$$|D_n^k\rangle = \binom{n}{k}^{-1/2} \sum_{\text{perm}} \underbrace{|0\rangle|0\rangle \cdots |0\rangle}_{n-k} \underbrace{|1\rangle|1\rangle \cdots |1\rangle}_k,$$

with $0 \leq k \leq n$. It is proved that the closest product state in σ_1 for computing E_2 [13,67] has the form

$$|\varphi\rangle = \left(\sqrt{\frac{n-k}{n}}|0\rangle + \sqrt{\frac{k}{n}}|1\rangle \right)^{\otimes n}, \quad (9)$$

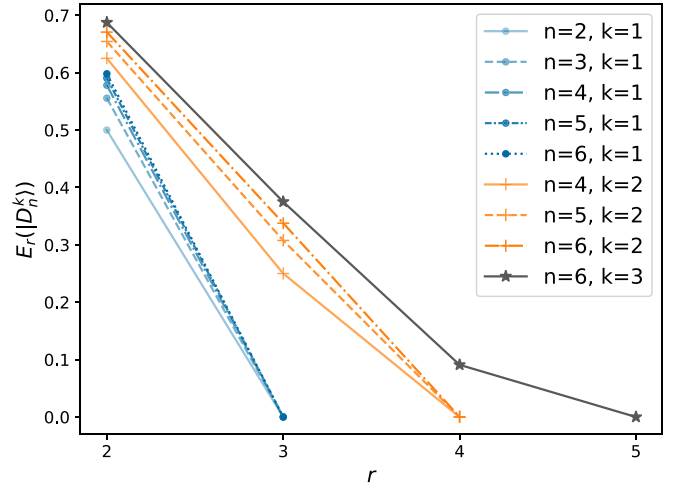


FIG. 4. $E_r(|D_n^k\rangle)$ as a function of r for different n and k . The transition of E_r between zero and nonzero at $k + 1$ and $k + 2$ implies the border rank of $|D_n^k\rangle$ is $k + 1$ as predicted.

i.e., a tensor product of n identical single qubit states. From that, the geometric measure is found to be

$$E_2(|D_n^k\rangle) = 1 - \binom{n}{k} \left(\frac{k}{n}\right)^k \left(\frac{n-k}{n}\right)^{n-k}.$$

Due to the symmetry of Dicke states, we can just discuss the cases where $k \leq \lfloor \frac{n}{2} \rfloor$. It has been proved that for a Dicke state $|D_n^k\rangle$, its border rank $R = k + 1$ [68], which means there should be a transition between zero and nonzero for E_{k+2} and E_{k+1} .

We compute the geometric measure $E_r(|D_n^k\rangle)$ for different n and k , as shown in Fig. 4. For $r = 2$, we compare the numerical results with analytical results, which implies the error is close to the machine precision. The analytical form in Eq. (9) can also be found in the optimization process. Generally, uncovering the closest state proves challenging unless the given state exhibits certain symmetries [69]. Our methodology furnishes a potent instrument for the exploration of these closest states even in $\sigma_r (r > 2)$. Moreover, we observe distinct transitions from nonzero to zero for different Dicke states, indicating the border rank of the given $|D_n^k\rangle$ as postulated. In Appendix, we compare the results obtained using the seesaw strategy and the PPT relaxation for calculating $E_2(|D_n^k\rangle)$. We find that the time efficiency of gradient descent lies between these two methods. However, as we observe, neither the seesaw strategy nor the PPT relaxation can detect higher-dimensional entanglement, which highlights the major advantage of the gradient descent approach we propose.

Border rank and tensor rank are also closely related to algebraic complexity theory. One of the most important open questions in computer science is to understand the computational complexity of matrix multiplication. One would like to know how many multiplication operations are required in order to multiply $n \times n$ matrices. The naive approach uses n^3 multiplications while Strassen found an algorithm only needs seven multiplications for 2×2 matrices [34].

The precise number of least multiplications for the matrix multiplication is called the rank of matrix multiplication while the concept of border rank arises when some matrix multiplications can be approximated with arbitrary precision by less complicated multiplications, which have lower ranks [35,36]. These approximations can give faster algorithms for matrix multiplication in practice.

Example 5. It turns out that the rank of $n \times n$ matrix multiplication is equivalent to the rank of the following tripartite state (unnormalized) [70,71]:

$$\begin{aligned} |\Phi_n\rangle_{ABC} &= |\Phi\rangle_{AB}|\Phi\rangle_{AC}|\Phi\rangle_{BC} \\ &= \sum_{i,j,k=0}^{n-1} |ij\rangle_A |ik\rangle_B |jk\rangle_C, \end{aligned}$$

where any two parties share an EPR pair $|\Phi\rangle$ of dimension n . By computing the values of $E_r(|\Phi_2\rangle_{ABC})$ for different r , we find that E_7 is very close to $\frac{1}{8}$ while E_8 is around 10^{-14} , which implies that $\underline{R}(|\Phi_2\rangle_{ABC}) = 7$. This numerical result matches the previous finding [72], validating the practicability of our approach for studying border rank in many relevant fields, particularly in mathematics and computer science.

2. Completely entangled subspaces

Example 6. Consider the following four three-qubit states [62]:

$$\begin{aligned} |\psi_0\rangle &= |0\rangle_A |0\rangle_B |0\rangle_C, \\ |\psi_1\rangle &= |1\rangle_A |+\rangle_B |-\rangle_C, \\ |\psi_2\rangle &= |-\rangle_A |1\rangle_B |+\rangle_C, \\ |\psi_3\rangle &= |+\rangle_A |-\rangle_B |1\rangle_C. \end{aligned}$$

They form a UPB, which means no product state can be found in the orthogonal space of the subspace \mathcal{S} spanned by them. In other words, \mathcal{S}^\perp is a completely entangled subspace. However, similar to the bipartite UPB case, PPT relaxation cannot detect its entanglement since the lower bound of E_2 obtained close to 10^{-14} while the gradient descent gives the value around 0.08144, close to the analytical value $1 - \frac{3\sqrt{6}}{8}$.

Example 7. The largest possible dimension of a completely entangled subspace (CES) of $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ is $d_1 d_2 d_3 - d_1 - d_2 - d_3 + 2$ and one particular example of such a subspace is [73]

$$\begin{aligned} \mathcal{S} &= \text{span}\{|i_1\rangle \otimes |i_2\rangle \otimes |i_3\rangle - |j_1\rangle \otimes |j_2\rangle \otimes |j_3\rangle\} : \\ &\sum_{r=1}^3 i_r = \sum_{r=1}^3 j_r, 0 \leq i_r \leq d_r - 1, 1 \leq r \leq 3. \end{aligned}$$

Our method is able to certify these maximal-dimension CES for different values of d_1, d_2 , and d_3 quite efficiently. We calculate the computational time for certifying these CESs via different methods in Table II.

As we can see, the hierarchical method is most easily limited by the dimension sizes. Although the PPT method can detect most of them successfully and give the

TABLE II. Computational time for certifying CESs via different methods, including Hier (hierarchical method), PPT (PPT relaxation), and GD (gradient descent). The values of geometric measure E_2 obtained by PPT and GD are also presented. “-” means it is not available in acceptable time.

(d_1, d_2, d_3)	Time (s)			Geometric measure	
	Hier	PPT	GD	E_2^{PPT}	E_2^{GD}
(2,2,2)	0.01	0.09	0.07	2.00×10^{-1}	2.50×10^{-1}
(2,2,4)	0.07	0.16	0.08	4.49×10^{-2}	4.50×10^{-2}
(2,2,6)	0.30	0.33	0.10	1.23×10^{-2}	1.23×10^{-2}
(2,3,4)	180.42	0.32	0.08	1.41×10^{-2}	1.41×10^{-2}
(2,3,6)	-	1.07	0.15	2.86×10^{-3}	2.86×10^{-3}
(2,3,8)	-	2.10	0.21	7.62×10^{-4}	7.62×10^{-4}
(3,3,6)	-	2.71	0.18	7.20×10^{-4}	7.20×10^{-4}
(3,3,8)	-	8.55	0.32	1.57×10^{-4}	1.57×10^{-4}
(3,4,7)	-	18.27	0.40	7.98×10^{-5}	7.98×10^{-5}
(4,4,7)	-	189.31	0.55	2.02×10^{-5}	2.02×10^{-5}
(4,5,10)	-	-	2.92	-	3.54×10^{-7}

accurate lower bounds of E_2 for relatively small dimensions, its computational time will increase significantly since it requires a quantum state ρ with the dimension $d_A d_B d_C \times d_A d_B d_C$. On the contrary, the gradient descent approach can deal with much larger dimensions due to the need for only $2(d_A + d_B + d_C) + 1$ real parameters. This example shows the practicability and efficiency of our approach clearly.

3. Genuinely entangled subspaces

Example 8. The entangled subspace spanned by the vectors in Eq. (8) can be generalized to genuinely entangled subspace (GES) $\mathcal{S}_{2 \times d^2}^\theta$ in $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, spanned by

$$|\psi_{i_1 i_2}\rangle_{ABC} = a|0\rangle_A |i_1\rangle_B |i_2\rangle_C + b|1\rangle_A |i_1 + 1\rangle_B |i_2 + 1\rangle_C,$$

where $d_A = 2, d_B = d_C = d, i_1, i_2 = 0, 1, \dots, d - 2$ and $a = \cos(\theta/2), b = e^\xi \sin(\theta/2), \theta \in (0, \pi), \xi \in [0, 2\pi)$.

For this particular genuinely entangled subspace, it has been proved that [12], for any possible bipartition $K|K^c$ in $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$:

$$E_2^{K|K^c}(\mathcal{S}_{2 \times d^2}^\theta) = \frac{1}{2} \left[1 - \sqrt{1 - \sin^2 \theta \sin^2 \left(\frac{\pi}{d} \right)} \right].$$

We compare this analytical result with ones obtained by gradient descent, which implies the errors are close to the machine precision. That means our approach can successfully detect their genuine entanglement and give accurate values for geometric measure over bipartitions.

V. CONCLUSIONS AND OUTLOOKS

In this work, we introduce a potent tool for examining the entanglement within a given subspace. Through several examples, the universality, practicability, and efficiency of our approach become evident. The core idea of our

methodology involves finding a suitable way to map a free Euclidean space onto the complicated manifold in the given quantum scenario. Rather than augmenting the dimension of quantum problems to supplement convexity, we employ the highly developed nonconvex optimization techniques from machine learning to expedite the process. Consequently, we can address various issues related to quantum subspace, ranging from the certification of bipartite entangled subspaces, to the computation of the border rank of multipartite states, and the verification of completely and genuinely entangled subspaces. All of these results can be reproduced in our public repository [74].

Convex optimization techniques, particularly semidefinite programming (SDP) method, are often used for quantifying entanglement. Although SDP can guarantee a global minimum, mathematically, it still only provides lower bounds. This means that when it returns zero, we cannot exclude the presence of entanglement, such as bound entanglement. Moreover, as highlighted by the results compared in the paper, we can see that its computational complexity increases significantly with dimension, implying that it may only provide entanglement information for relatively small quantum systems.

In contrast, algorithms based on gradient descent strategies provide upper bounds. These also have their drawbacks and advantages. For example, when they return a nonzero value, we theoretically cannot exclude the possibility of a local minimum. However, as we have observed, they offer a more universal and versatile approach, enabling application to a broader range of scenarios, such as the detection of high-dimensional entanglement. Furthermore, thanks to the rapid advancement of nonconvex optimization techniques in recent years, especially those applied in machine learning, their computational efficiency significantly surpasses that of SDP, making it feasible to address quantum problems in more complex systems.

In practice, these two approaches do not conflict. To calculate actual values, we always need both upper and lower bounds to better estimate them. We hope that our research can inspire more applications of nonconvex optimization in quantum information.

ACKNOWLEDGMENTS

We gratefully acknowledge N. Cao, S. Zhou, J. Hu, K. She, and Y. T. Poon for insightful discussions and the assistance of ChatGPT in facilitating the writing process. X.Z., C.Z., and B.Z. are supported by General Research Fund (GRF) Grant No. 16305121.

TABLE III. Comparison of gradient descent (GD), seesaw strategy and the PPT relaxation method in calculating the geometric measure of entanglement for various Dicke states $|D_n^k\rangle$. Time is measured in seconds. The error and time columns denote the absolute error (in unit $\times 10^{-12}$) compared to the analytical value E_G and computational time for the corresponding method. NA means not available within acceptable time.

(n, k)	E_G	GD		Seesaw		PPT
		Error	Time	Error	Time	Time
(5,1)	0.5904	0.21	0.0101	0.23	0.103	1.73
(5,2)	0.6544	0.03	0.087	0.78	0.0091	1.64
(5,3)	0.6544	0.13	0.084	0.70	0.0091	1.59
(6,1)	0.5981	0.21	0.129	0.27	0.017	25.1
(6,2)	0.6708	0.02	0.122	0.44	0.018	26.5
(6,3)	0.6875	0.12	0.166	0.65	0.018	25.7
(7,1)	0.6034	0.01	0.173	1.08	0.019	NA
(7,2)	0.6813	0.22	0.163	0.24	0.019	NA
(7,3)	0.7062	0.04	0.150	0.16	0.020	NA

APPENDIX: COMPARISON OF TIME AND ACCURACY FOR COMPUTING THE GEOMETRIC MEASURE OF ENTANGLEMENT FOR DICKE STATES

Here, we introduce the seesaw strategy proposed in Ref. [65] for the calculation of geometric measure of entanglement E_G [see Eq. (1)]. Given a pure state $|\psi\rangle$ in an n -partite Hilbert space $\bigotimes_{i=1}^n \mathcal{H}_i$, one straightforward strategy to find the closest fully product state $|\varphi\rangle$ under the geometric measure is as follows.

Start with a random fully product state $|\varphi_0\rangle = \bigotimes_{i=1}^n |\phi_0^{(i)}\rangle$. Then, consider the unnormalized state $|\tilde{\psi}\rangle = (\bigotimes_{i=2}^n \langle \phi_0^{(i)} |) |\psi\rangle$. To minimize E_G , i.e., to maximize the overlap $|\langle \varphi | \psi \rangle|$ for fixed states $\bigotimes_{i=2}^n |\phi_0^{(i)}\rangle$, replace $|\phi_0^{(1)}\rangle$ with the state $|\phi_1^{(1)}\rangle = \frac{|\tilde{\psi}\rangle}{\sqrt{\langle \tilde{\psi} | \tilde{\psi} \rangle}}$.

This process is then repeated for each of the other parties. This iterative method, known as the seesaw strategy, is continued until convergence. The convergence criterion is set such that the fidelity of every local state before and after the update is greater than $1 - 10^{-10}$. This strategy can also provide an upper bound for the geometric measure of entanglement.

In Table III, we compare the computational time and accuracy for calculating the geometric measure of entanglement for Dicke states $|D_n^k\rangle$ using the seesaw strategy, the PPT relaxation, and gradient descent (GD). Here, for each state, we repeat the optimization three times for both seesaw and gradient descent strategy.

[1] A. Steane, Quantum computing, *Rep. Prog. Phys.* **61**, 117 (1998).
 [2] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
 [3] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, *Rev. Mod. Phys.* **74**, 145 (2002).

[4] S. Pirandola, J. Eisert, C. Weedbrook, A. Furusawa, and S. L. Braunstein, Advances in quantum teleportation, *Nat. Photon.* **9**, 641 (2015).
 [5] K. R. Parthasarathy, On the maximal dimension of a completely entangled subspace for finite level quantum systems, *Proc. Math. Sci.* **114**, 365 (2004).

- [6] J. Walgate and A. J. Scott, Generic local distinguishability and completely entangled subspaces, *J. Phys. A: Math. Theor.* **41**, 375305 (2008).
- [7] M. Demianowicz, Universal construction of genuinely entangled subspaces of any size, *Quantum* **6**, 854 (2022).
- [8] M. Demianowicz, G. Rajchel-Mieldzióć, and R. Augusiak, Simple sufficient condition for subspace to be completely or genuinely entangled, *New J. Phys.* **23**, 103016 (2021).
- [9] W. Bruzda, S. Friedland, and K. Życzkowski, Rank of a tensor and quantum entanglement, *Linear Multilinear Algebra*, **1** (2023).
- [10] J. Eisert and H. J. Briegel, Schmidt measure as a tool for quantifying multiparticle entanglement, *Phys. Rev. A* **64**, 022306 (2001).
- [11] N. Johnston, B. Lovitz, and A. Vijayaraghavan, Complete hierarchy of linear systems for certifying quantum entanglement of subspaces, *Phys. Rev. A* **106**, 062443 (2022).
- [12] M. Demianowicz and R. Augusiak, Entanglement of genuinely entangled subspaces and states: Exact, approximate, and numerical results, *Phys. Rev. A* **100**, 062318 (2019).
- [13] T.-C. Wei and P. M. Goldbart, Geometric measure of entanglement and applications to bipartite and multipartite quantum states, *Phys. Rev. A* **68**, 042307 (2003).
- [14] S. Liu, Q. He, M. Huber, O. Gühne, and G. Vitagliano, Characterizing entanglement dimensionality from randomized measurements, *PRX Quantum* **4**, 020324 (2023).
- [15] I. Nape, V. Rodríguez-Fajardo, F. Zhu, H.-C. Huang, J. Leach, and A. Forbes, Measuring dimensionality and purity of high-dimensional entangled states, *Nat. Commun.* **12**, 5159 (2021).
- [16] A. Ekert and P. L. Knight, Entangled quantum systems and the Schmidt decomposition, *Am. J. Phys.* **63**, 415 (1995).
- [17] J. F. Buss, G. S. Frandsen, and J. O. Shallit, The computational complexity of some problems of linear algebra, *J. Comput. Syst. Sci.* **58**, 572 (1999).
- [18] P. Horodecki, Separability criterion and inseparable mixed states with positive partial transposition, *Phys. Lett. A* **232**, 333 (1997).
- [19] R. Augusiak, J. Tura, and M. Lewenstein, A note on the optimality of decomposable entanglement witnesses and completely entangled subspaces, *J. Phys. A: Math. Theor.* **44**, 212001 (2011).
- [20] D. Chruściński and G. Sarbicki, Entanglement witnesses: construction, analysis and classification, *J. Phys. A: Math. Theor.* **47**, 483001 (2014).
- [21] G. Gour and N. R. Wallach, Entanglement of subspaces and error-correcting codes, *Phys. Rev. A* **76**, 042309 (2007).
- [22] F. Huber and M. Grassl, Quantum codes of maximal distance and highly entangled subspaces, *Quantum* **4**, 284 (2020).
- [23] M. Horodecki and M. Piani, On quantum advantage in dense coding, *J. Phys. A: Math. Theor.* **45**, 105306 (2012).
- [24] M. Walter, D. Gross, and J. Eisert, Multipartite entanglement, *Quantum Inf.*, 293 (2016).
- [25] P. A. M. Dirac, A new notation for quantum mechanics, *Math. Proc. Cambridge Philos. Soc.* **35**, 416 (1939).
- [26] J. Håstad, Tensor rank is NP-complete, *J. Algo.* **11**, 644 (1990).
- [27] C. J. Hillar and L.-H. Lim, Most tensor problems are NP-hard, *J. ACM* **60**, 1 (2013).
- [28] T. Lickteig, A note on border rank, *Inf. Process. Lett.* **18**, 173 (1984).
- [29] J. M. Landsberg and Z. Teitler, On the ranks and border ranks of symmetric tensors, *Found. Comput. Math.* **10**, 339 (2010).
- [30] M. Erhard, M. Krenn, and A. Zeilinger, Advances in high-dimensional quantum entanglement, *Nat. Rev. Phys.* **2**, 365 (2020).
- [31] L.-J. Kong, Y. Sun, F. Zhang, J. Zhang, and X. Zhang, High-dimensional entanglement-enabled holography, *Phys. Rev. Lett.* **130**, 053602 (2023).
- [32] D. Cozzolino, B. Da Lio, D. Bacco, and L. K. Oxenløwe, High-dimensional quantum communication: benefits, progress, and future challenges, *Adv. Quantum Technol.* **2**, 1900038 (2019).
- [33] S. Etcheverry, G. Cañas, E. Gómez, W. Nogueira, C. Saavedra, G. Xavier, and G. Lima, Quantum key distribution session with 16-dimensional photonic states, *Sci. Rep.* **3**, 2316 (2013).
- [34] V. Strassen *et al.*, Gaussian elimination is not optimal, *Numer. Math.* **13**, 354 (1969).
- [35] D. Bini, M. Capovani, F. Romani, and G. Lotti, $O(n^{2.7799})$ complexity for $n \times n$ approximate matrix multiplication, *Inf. Process. Lett.* **8**, 234 (1979).
- [36] D. Bini, G. Lotti, and F. Romani, Approximate solutions for the bilinear form computational problem, *SIAM J. Comput.* **9**, 692 (1980).
- [37] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, Unextendible product bases and bound entanglement, *Phys. Rev. Lett.* **82**, 5385 (1999).
- [38] M. Demianowicz and R. Augusiak, From unextendible product bases to genuinely entangled subspaces, *Phys. Rev. A* **98**, 012313 (2018).
- [39] S. Agrawal, S. Halder, and M. Banik, Genuinely entangled subspace with all-encompassing distillable entanglement across every bipartition, *Phys. Rev. A* **99**, 032335 (2019).
- [40] B. Lovitz and N. Johnston, Entangled subspaces and generic local state discrimination with pre-shared entanglement, *Quantum* **6**, 760 (2022).
- [41] A. H. Shenoy and R. Srikanth, Maximally nonlocal subspaces, *J. Phys. A: Math. Theor.* **52**, 095302 (2019).
- [42] D. Cavalcanti, Connecting the generalized robustness and the geometric measure of entanglement, *Phys. Rev. A* **73**, 044302 (2006).
- [43] Y.-C. Ou and H. Fan, Bounds on negativity of superpositions, *Phys. Rev. A* **76**, 022320 (2007).
- [44] J. Niset and N. J. Cerf, Tight bounds on the concurrence of quantum superpositions, *Phys. Rev. A* **76**, 042328 (2007).
- [45] W. Song, N.-L. Liu, and Z.-B. Chen, Bounds on the multipartite entanglement of superpositions, *Phys. Rev. A* **76**, 054303 (2007).
- [46] Y. Xiang, S.-J. Xiong, and F.-Y. Hong, The bound of entanglement of superpositions with more than two components, *Eur. Phys. J. D* **47**, 257 (2008).
- [47] S. J. Akhtarshenas, Concurrence of superpositions of many states, *Phys. Rev. A* **83**, 042306 (2011).
- [48] Z. Ma, Z. Chen, S. Han, S. Fei, and S. Severini, Improved bounds on negativity of superpositions, *Quantum Inf. Comput.* **12**, 983 (2012).
- [49] N. Linden, S. Popescu, and J. A. Smolin, Entanglement of superpositions, *Phys. Rev. Lett.* **97**, 100502 (2006).
- [50] Z. Zhang, Y. Dai, Y.-L. Dong, and C. Zhang, Numerical and analytical results for geometric measure of coherence and geometric measure of entanglement, *Sci. Rep.* **10**, 12122 (2020).

- [51] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Complete family of separability criteria, *Phys. Rev. A* **69**, 022308 (2004).
- [52] M. Horodecki, P. Horodecki, and R. Horodecki, Mixed-state entanglement and distillation: Is there a “bound” entanglement in nature? *Phys. Rev. Lett.* **80**, 5239 (1998).
- [53] T. Cubitt, A. Montanaro, and A. Winter, On the dimension of subspaces with bounded Schmidt rank, *J. Math. Phys.* **49**, 022107 (2008).
- [54] M. Blazone, F. Dell’Anno, S. De Siena, and F. Illuminati, Hierarchies of geometric entanglement, *Phys. Rev. A* **77**, 062304 (2008).
- [55] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, PyTorch: An Imperative Style, High-Performance Deep Learning Library, in *Advances in Neural Information Processing Systems 32*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché Buc, E. Fox, and R. Garnett (Curran Associates, New York, 2019), pp. 8024–8035.
- [56] M. L. Casado, Trivializations for gradient-based optimization on manifolds, [arXiv:1909.09501](https://arxiv.org/abs/1909.09501).
- [57] J. Hu, X. Liu, Z.-W. Wen, and Y.-X. Yuan, A brief introduction to manifold optimization, *J. Oper. Res. Soc. China* **8**, 199 (2020).
- [58] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds* (Princeton University Press, Princeton, 2008).
- [59] H. A. Carteret, A. Higuchi, and A. Sudbery, Multipartite generalization of the Schmidt decomposition, *J. Math. Phys.* **41**, 7932 (2000).
- [60] L. Chen, E. Chitambar, R. Duan, Z. Ji, and A. Winter, Tensor rank and stochastic entanglement catalysis for multipartite pure states, *Phys. Rev. Lett.* **105**, 200501 (2010).
- [61] V. De Silva and L.-H. Lim, Tensor rank and the ill-posedness of the best low-rank approximation problem, *SIAM J. Matrix Anal. Appl.* **30**, 1084 (2008).
- [62] C. Branciard, H. Zhu, L. Chen, and V. Scarani, Evaluation of two different entanglement measures on a bound entangled state, *Phys. Rev. A* **82**, 012327 (2010).
- [63] H. Zheng, Z. Yang, W. Liu, J. Liang, and Y. Li, Improving deep neural networks using softplus units, *2015 International Joint Conference on Neural Networks (IJCNN)* (IEEE, Killarney, 2015), pp. 1–4.
- [64] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, SCIPY 1.0: fundamental algorithms for scientific computing in python, *Nat. Methods* **17**, 261 (2020).
- [65] A. Streltsov, H. Kampermann, and D. Bruss, Simple algorithm for computing the geometric measure of entanglement, *Phys. Rev. A* **84**, 022323 (2011).
- [66] D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, Unextendible product bases, uncompletable product bases and bound entanglement, *Commun. Math. Phys.* **238**, 379 (2003).
- [67] T.-C. Wei and S. Severini, Matrix permanent and quantum entanglement of permutation invariant states, *J. Math. Phys.* **51**, 092203 (2010).
- [68] M. Gharahi and S. Mancini, Algebraic-geometric characterization of tripartite entanglement, *Phys. Rev. A* **104**, 042402 (2021).
- [69] R. Hübener, M. Kleinmann, T.-C. Wei, C. González-Guillén, and O. Gühne, Geometric measure of entanglement for symmetric states, *Phys. Rev. A* **80**, 032324 (2009).
- [70] E. Chitambar, R. Duan, and Y. Shi, Tripartite entanglement transformations and tensor rank, *Phys. Rev. Lett.* **101**, 140502 (2008).
- [71] M. Christandl, V. Lyikov, V. Steffan, A. H. Werner, and F. Witteveen, [arXiv:2307.07394](https://arxiv.org/abs/2307.07394).
- [72] J. Landsberg, The border rank of the multiplication of 2×2 matrices is seven, *J. Am. Math. Soc.* **19**, 447 (2006).
- [73] B. R. Bhat, A completely entangled subspace of maximal dimension, *Int. J. Quantum Inf.* **04**, 325 (2006).
- [74] C. Zhang and X. Zhu, numqi/entangled-subspace, <https://doi.org/10.5281/zenodo.12166896> (2024).