

Classification of four-qubit entangled states via machine learning

S. V. Vintskevich¹, N. Bao,² A. Nomerotski,² P. Stankus,² and D. A. Grigoriev³

¹*Technology Innovation Institute, Abu Dhabi, Masdar City 9639, United Arab Emirates*

²*Brookhaven National Laboratory, Upton New York 11973, USA*

³*LLP Eqvium, Almaty 050009, Kazakhstan*



(Received 6 July 2022; accepted 2 March 2023; published 23 March 2023)

We apply the support vector machine (SVM) algorithm to derive a set of entanglement witnesses (EW) to identify entanglement patterns in families of four-qubit states. The effectiveness of SVM for practical EW implementations stems from the coarse-grained description of families of equivalent entangled quantum states. The equivalence criteria in our work is based on the stochastic local operations and classical communication classification and the description of the four-qubit entangled Werner states. We numerically verify that the SVM approach provides an effective tool to address the entanglement witness problem when the coarse-grained description of a given family state is available. We also discuss and demonstrate the efficiency of nonlinear kernel SVM methods as applied to four-qubit entangled state classification.

DOI: [10.1103/PhysRevA.107.032421](https://doi.org/10.1103/PhysRevA.107.032421)

I. INTRODUCTION

The ability to quantify, detect and analyze the structure of quantum entanglement [1] is essential for quantum computation [2–4], quantum communication [5–7], quantum networks [8–10], and quantum metrology [11,12]. Moreover, an improper “amount” of entanglement, incorrect structure or pattern of multipartite entangled state, or action of quantum noise could all severely affect the overall efficiency of given quantum computation tasks [13] or performance of given quantum protocols, such as that of entanglement purification [14]. Thus, it is crucial to detect and describe the structure of entanglement in a set of states to maximize the efficiency of a given protocol.

As a specific case of this, one can attempt to detect the presence of entanglement in particular families of states relevant in specific quantum protocols; note that the success probability of the protocol may vary when one uses different families of entangled states [15–17]. A characteristic example of such a difference is that between the multipartite GHZ- and W-type states. The GHZ-type states are fragile against losses and are more utilized in quantum information sharing protocols, whereas W-type states are robust against the noise and are used in multi-party quantum network protocols [18]. For instance, in the case of tripartite W-state it is easy to verify that each pair of qubits in this state is in an entangled state in contrast to the tripartite GHZ-type state. This particular example emphasizes robustness of the W state, e.g., we can assume that the third qubit can be traced out to emulate losses.

The entanglement witness (EW) technique is one of the most common, effective, and practical methods to detect the presence of entanglement for a given *multipartite* quantum state, see, e.g., Refs. [1,19], and references therein. In the present work, we analyze this technique applied to arbitrary four-qubit systems. The essence of the EW technique is briefly summarized below. By finding a specific Hermitian operator, called the entanglement witness \hat{W} , one calculates a linear functional (EW functional), which maps a given multipartite

quantum state described by density operator $\hat{\rho}$ into a real number.¹ An EW functional has a non-negative value for all separable states $\hat{\rho}_{\text{sep}}$ and that there exists a particular set of entangled states for which the EW functional has negative values [1,20,21].

More specifically, let us formally denote a set of arbitrary multiqubit states as $S_{n_q} := \{\hat{\rho}_{n_q} \in \mathcal{T}(\mathcal{H}_{n_q}) \mid \text{tr}(\hat{\rho}_{n_q}) = 1, \hat{\rho}_{n_q} \geq O\}$. We denoted $\mathcal{T}(\mathcal{H}_{n_q})$ as a linear space of trace class operators acting on Hilbert space \mathcal{H}_{n_q} . The condition $\hat{\rho}_{n_q} \geq O$ states that operator $\hat{\rho}$ is a positive operator [22]. Mathematically, the EW functional will have the following property:

$$\text{tr}(\hat{\rho}_{\text{sep}} \hat{W}) \geq 0, \forall \hat{\rho}_{\text{sep}} \in S; \text{tr}(\hat{\rho}_{\text{ent}} \hat{W}) < 0, \quad (1)$$

where we denote S as a subset of all separable states. By definition [1,22] a state $\hat{\rho}_{\text{sep}}$ is separable if and only if it can be represented as a convex combination of factorized states: $\hat{\rho} = \sum_i^k p_i \hat{\rho}_1^i \otimes \dots \otimes \hat{\rho}_N^i$; where each $\hat{\rho}^i$ is a state of the i -th qubit subsystem, a density operator acting on a subspace \mathcal{H}_i of a N -partite quantum system with joint space $\mathcal{H} = \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_N$; $\sum_{i=1}^k p_i = 1$, $k \leq \dim \mathcal{H}^2$. Thus, the set of all separable states is a convex subset of all states S_{n_q} defined above with respect to the trace norm [22], while, by definition, the entangled states are those states that are not separable.

Note that the EW operator \hat{W} , which can detect the entanglement for a given state $\hat{\rho}_{\text{ent}}$ is not universal. There are always *other* entangled states $\hat{\sigma}_{\text{ent}}$ such that $\text{tr}(\hat{W} \hat{\sigma}_{\text{ent}}) > 0$, but at the same time $\text{tr}(\hat{W} \hat{\rho}_{\text{ent}}) < 0$. Unfortunately, one cannot determine the entanglement witness operator for all possible entangled states for an arbitrary multipartite quantum system [1,22]. It was shown that the problem of a general description of all entangled states, pure and mixed, for multipartite quantum systems does not have a solution, see e.g., [23–26]. However, in some cases, the aforementioned disadvantage

¹The usage of density operators allows the accommodation of both mixed and pure states.

might be partially overcome. The general idea is to split the quantum states into specific families - sets of states that may share certain symmetries or specific structure. Another possibility is if the quantum states can be created via a specific protocol that maps a certain set of states to another set that can be mathematically described via a specific parametrization.

The idea of splitting sets of entangled quantum states (generally, an infinite set of states) into families with some inner mathematical structure can be termed a *coarse-grained classification*. For instance one can specify equivalence classes, families of the pure entangled states. Each state within a given family can be transformed into another state of the same family with nonzero probability through local operations and classical communications (LOCC) [27,28]. If one can describe or specify a structure of entangled states one may expect that the complexity of finding entanglement witnesses for such set of states will be significantly reduced. The mentioned above GHZ and W states are members of two distinguishable families of states that cannot be converted into one another by any (SLOCC) as was emphasized in Ref. [28]. Other possible approaches include inductive entanglement classification [29], entanglement classification with matrix product states [30], and coarse graining of entanglement classes in $2 \times m \times n$ systems [31].

Here we aim to employ the coarse-grained classifications of four-qubit quantum states, following the results of Ref. [28], to construct a set of entanglement witness operators. We avoid difficulties of analytical derivation of the entanglement witness operators by approaching the problem numerically. The core of our numerical analysis is based on the well-known support vector machine (SVM) method in machine-learning (ML) [32,33]. The SVM-based algorithm is designed to detect the presence of entanglement not only in arbitrary four-qubit states but also to assign this state to a particular family of states. To train our SVM model we sample a data set of quantum states for each class of entangled and separable states based on this coarse-grained classification. We are sampling 20 000 states for each class of entangled states, including separable states.

It is worth emphasizing that in recent years, ML-based methods have demonstrated remarkable efficiency in application to various areas of quantum physics [34–37]. For instance, others have used the neural networks [38,39] and also the SVM [40] to find EW operators, which efficiently distinguish between separable and entangled states of a particular type.

The simplest linear SVM approach can classify quantum states if they belong to one of two classes (entangled and probably separable) by computing the *decision function* and constructing a *decision boundary*. Thus, the SVM approach directly corresponds to the EW problem. Note that SVM fits our problem naturally, as quantum states can be equivalently considered as vectors in $\mathcal{T}(\mathcal{H}_{n_q})$ space. Consequently, the SVM approach allows one to find a hyper-plane in the $\mathcal{T}(\mathcal{H}_{n_q})$, which corresponds to the entanglement witness. We describe our approach and results of its application in Sec. II. The Appendix provides a detailed description of the SLOCC classification of four-qubit states used in this work.

II. ENTANGLEMENT WITNESSES FOR FOUR-QUBIT STATES AND APPLICATION OF SVM

In this section we construct a set of EW operators by considering a set of four qubit states ($n_q = 4$). As the first step, let us provide a general mathematical description for the linear space of trace class operators $\mathcal{T}(\mathcal{H}_{n_q})$ in the case of an arbitrary number of qubits n_q . The space $\mathcal{T}(\mathcal{H}_{n_q})$ is endowed with Hilbert-Schmidt inner product for any two operators \hat{R}_1, \hat{R}_2 : $\langle \hat{R}_1, \hat{R}_2 \rangle_{\text{HS}} = \text{tr}(\hat{R}_1^\dagger \hat{R}_2)$, which induces the norm $\|\hat{R}\|_{\text{HS}} = \sqrt{\langle \hat{R}^\dagger \hat{R} \rangle_{\text{HS}}}$. Let us choose the standard multi-qubit Pauli basis $\{\hat{\mathbf{I}}/\sqrt{2}, \hat{\sigma}_x/\sqrt{2}, \hat{\sigma}_y/\sqrt{2}, \hat{\sigma}_z/\sqrt{2}\}^{\otimes n_q}$ as a self-adjoint orthonormal basis for $\mathcal{T}(\mathcal{H}_{n_q})$. For simplicity we denoted a particular basis operator as \hat{E}_i , $i = 0, 4^{n_q} - 1$ assuming $\hat{E}_0 = \hat{\mathbf{I}}^{\otimes n_q}$, $\text{tr}(\hat{E}_i) = 0$, $i \neq 0$ and $\langle \hat{E}_i \hat{E}_j \rangle_{\text{HS}} = \delta_{ij}$. Thus, an arbitrary operator \hat{R} can be represented as a vector \vec{r} :

$$\hat{R} = \sum_j r_j \hat{E}_j, \quad r_k = \text{tr}(\hat{E}_k \hat{R}), \quad \vec{r} = (r_0, \dots, r_{4^{n_q}-1}). \quad (2)$$

If operator \hat{R} is Hermitian all elements of a corresponding vector \vec{r} are real numbers and $\|\hat{R}\|_{\text{HS}} = \|\vec{r}\|_e$, the norm $\|\cdot\|_e$ is the standard Euclidean norm of a vector. Consequently, the Hilbert-Schmidt inner product of two Hermitian operators $\langle \hat{R}_1, \hat{R}_2 \rangle_{\text{HS}}$ corresponds to the standard ‘‘Euclidean’’ inner product (\vec{r}_1, \vec{r}_2) . The vector representation emphasizes the direct correspondence between EW problems and the linear SVM method, and, indeed, the EW problems can be viewed as problems of classification. We consider two classes: the set of all separable states and a subset of entangled states. In the vector representation an EW operator and an arbitrary state can be denoted $\hat{W} \leftrightarrow \vec{w}$ and $\hat{\rho} \leftrightarrow \vec{\rho}$ respectively, in accordance with Eq. (2). Thus, the decision function used in SVM can be written as follows:

$$t_{\hat{W}}(\hat{\rho}) = \begin{cases} -1, & \text{if } (\vec{w}, \vec{\rho}) < 0 \\ 1, & \text{if } (\vec{w}, \vec{\rho}) > 0, \end{cases} \quad (3)$$

where we explicitly write the linear form of EW functional $f_{\hat{W}}(\rho) = (\vec{w}, \vec{\rho}) \equiv \text{tr}(\hat{W} \hat{\rho})$. Accordingly, a training data set $DS = \{\hat{\rho}_j\}_{j=1}^{N_{\text{data}}}$ consists of separable and entangled states in vector representation labeled with $t_{\hat{W}}(\hat{\rho}_{\text{sep.}}) = 1$ and $t_{\hat{W}}(\hat{\rho}_{\text{ent.}}) = -1$ respectively, whereas a decision hyperplane is defined by $f_{\hat{W}}(\hat{\rho}) = 0$.

For a given sampled state of a train data set $\hat{\rho}_j$ we will simplify the notation: $f_{\hat{W}}(\hat{\rho}_j) \equiv f_j$ and $t_{\hat{W}}(\hat{\rho}_j) \equiv t_j$. The SVM training objective is to find a decision boundary by maximizing the margin m , which is the smallest distance between the decision hyperplane (i.e., the boundary) and the closest quantum states, also named *support vectors*, from the training data set. We illustrate the SVM approach in Fig. 1.

To find the optimal margin m_* one can consider *minimization* of the following objective function $\text{argmin}_{\vec{w}, DS} L(\vec{w}, DS)$ named the *hinge error function*. The reader may find a comprehensive description in Refs. [33,41]:

$$L(\vec{w}, DS) = \sum_j^{N_{\text{data}}} \max(0, 1 - f_j t_j) + \lambda \|\vec{w}\|_e^2. \quad (4)$$

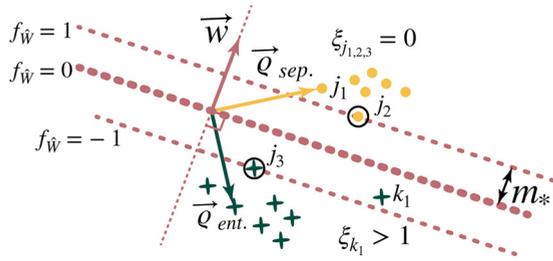


FIG. 1. Illustration of SVM approach in two dimensions. The distance from a given sample state $\hat{\rho}_j \leftrightarrow \vec{\rho}_j$ to a decision boundary $f_{\hat{W}} = 0$ is given by $m_j = t_j f_j / \|\vec{w}\|$. Mathematically, the value of an optimal margin value m_* can be found by minimizing the hinge error function (4). For simplicity we denote $\xi_j = \max(0, 1 - f_j t_j)$. The location of the decision boundary is determined by a subset of the data points, known as *support vectors*, which is located on the margin boundary ($f_{\hat{W}} = 1$ and $f_{\hat{W}} = -1$, respectively). The support vectors are marked by the circles. Thus, for the states that are on the correct side of the decision boundary $\xi_j = 0$ and for the outliers $\xi_j = 1 - t_j f_j$. Note that adding new samples that lie outside of the optimal margin region will not affect the decision boundary.

The parameter λ plays the role of regularization parameter to control the model's accuracy and generalization abilities.

Due to the given vector representation (2), the linear structure of entanglement witness functional (1) and relatively small dimensionality of the linear space of operators $\mathcal{T}(\mathcal{H}_{n_q})$ [in our case $n_q = 4$, $n_f = \dim(\mathcal{T}(\mathcal{H}_{n_q})) = 256$], the SVM approach is computationally feasible. We denoted the dimension of a feature space as n_f , which in our case coincides with the dimension of space $\mathcal{T}(\mathcal{H}_{n_q})$. For instance, a rough estimate of the SVM computational complexity is $O(N_{\text{data}} n_f)$, based on the implementation of SCIKIT LEARN python package [41]. Thus, the SVM computational complexity can be estimated by $O(N_{\text{data}} 4^{n_q})$. Therefore, the SVM is a good match for analyzes of complex but relatively small data sets.

In addition, the case of four qubits has another advantage, the SLOCC classification of entangled states. This coarse-grained classification of four-qubit entangled states further reduces the complexity of the EW problem and allows efficient application of SVM. In Sec. II A we focus on this classification following the results of [28] and applying it to the construction of EW operators. Section II B provides results of the SVM approach for construction of the EW operators and entanglement detection of arbitrary four-qubit states.

A. Classification of four-qubit states

It was shown in Ref. [28] how one can classify all pure states of four qubits into distinguishable classes of entangled *pure* states. The authors specified nine equivalent classes of this form. The equivalence criteria were determined with respect to the SLOCC applied to a particular quantum state. More specifically, states in a set can be treated as equivalent if any state for this set can be transformed into *any other state* with nonzero probability by means of LOCC. Note that in the present analysis we restrict ourselves to only local operations that are unitary transformations. A local unitary transformation of arbitrary four-qubit state can be written as

$$\hat{\rho} = \hat{U}_1 \otimes \hat{U}_2 \otimes \hat{U}_3 \otimes \hat{U}_4 \hat{\rho} \hat{U}_1^\dagger \otimes \hat{U}_2^\dagger \otimes \hat{U}_3^\dagger \otimes \hat{U}_4^\dagger, \quad (5)$$

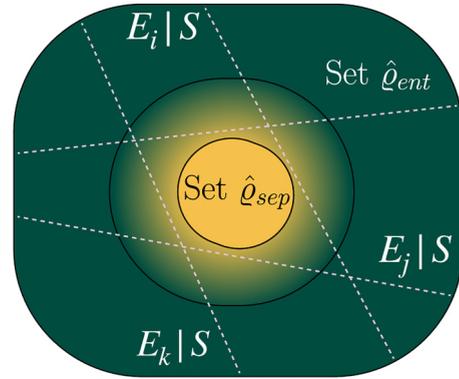


FIG. 2. Schematic representation of results obtained with the SVM approach. The decision boundaries between the set of separable states S and each class of entangled states E_i are represented by dashed lines. Each line corresponds to a particular class and is specified by a vector representation \vec{w}_i of the entanglement witness operator \hat{W}_i in accordance with Eqs. (1) and Eqs. (3).

where \hat{U} are arbitrary unitary operators acting on Hilbert space of a single qubit \mathcal{H}_2 . Operationally, the states belonging to the same class can be used in a given quantum protocol but they will have a different protocol efficiency. We list the explicit state classification in the Appendix based on results of Ref. [28]. We denote these *pure states* via ket vectors $|G\rangle, |E\rangle_i, i = 1, \dots, 8, |F\rangle_0$, where $|G\rangle$ represents a generic state, as $|E\rangle_i$ we denoted “specific” classes of entangled states, and $|F\rangle_0$ is a factorized state, so is a class member of separable four-qubit states.

Some of the states, like the class representatives in Eqs. (A1)–(A9), are parametrized with complex numbers, e.g., the generic states such as $|G\rangle$ in Eq. (A1) and states $|E\rangle_i, i = 1, 5$ in Eqs. (A2)–(A6). However, some classes are represented just by a single pure state. To form a data set needed to train our SVM model we sample the complex numbers for each class $\{E_i, i = 1, 5\}$. In addition, we sample the random unitary matrices \hat{U} to increase the diversity of a particular data set, in accordance with Eq. (5).

The sampling algorithm of random unitary matrices is implemented via the QUTIP PYTHON PACKAGE [42,43] with a slight modification as explained below. We took the standard QUTIP function `rand_unitary` and added an additional parameter ε to control the value of U_ε . For this we assumed that

$$U_\varepsilon = \exp(i\varepsilon H), \quad (6)$$

where H is a random 2×2 full rank Hermitian matrix. The sampling of H is done with another QUTIP function, `rand_herm`.

Summarizing, for each SLOCC class of four-qubit states we create $N_{\text{data}} = 20000$ samples of pure states, including factorized states, utilizing the LOCC transformations for each state (5) with random unitary operators in accordance with Eq. (6). We normalize each sample by replacing $\hat{\rho}_{E_i} \rightarrow \hat{\rho}_{E_i} / \text{tr}(\hat{\rho}_{E_i})$. As the next step, we find the EW operators via a SVM algorithm, which can distinguish separable four-qubit states \hat{Q}_{sep} and each class of entangled states represented in families $E_i, i = 1, \dots, 8$. Figure 2 illustrates the expected result.

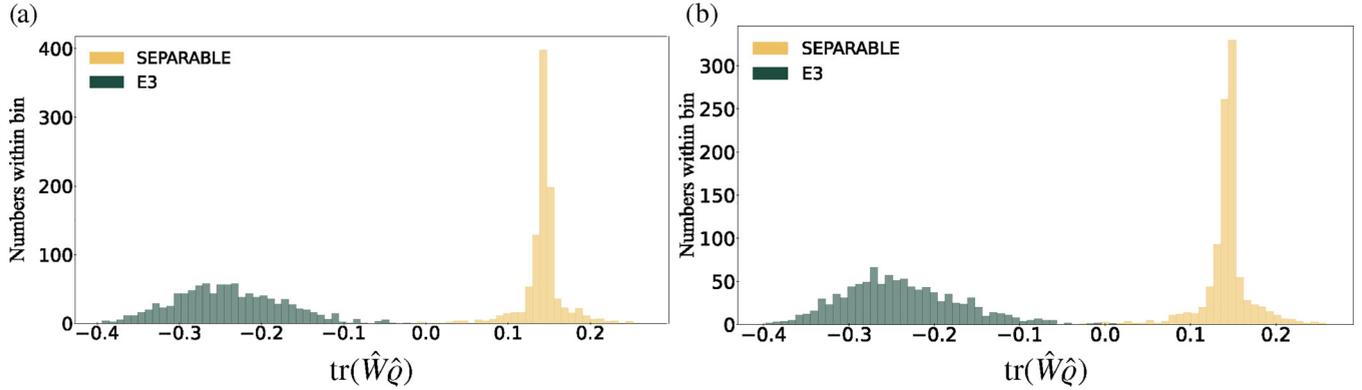


FIG. 3. Distribution of $\text{tr}(\hat{\rho}\hat{W})$, mean value of the entanglement witness for (a) validation set and (b) test set of the trained EW model via linear SVM algorithm in the case of E_3 family of states. Both test set and validation set consist of 2000 samples: 1000 separable states and 1000 entangled states of E_3 family. For both validation and test sets there were only few (≤ 5) miss-classifications of entangled states and zero miss-classifications for separable states. Note that the training data set included mixed Werner states to achieve better generalization.

The set of separable states S is a closed convex set and any separable state can be represented as a convex hull of factorized states. We directly employ this property to create a data set of separable states to train and validate our results achieving high accuracy, see Sec. II B. To create a single separable state we sample $4 * 4^4$ random factorized unitary operators (matrices): $\hat{U}_f^{(4)} \equiv \hat{U}_1 \otimes \hat{U}_2 \otimes \hat{U}_3 \otimes \hat{U}_4$. Each 2×2 single qubit unitary operator \hat{U} is distributed according to the Haar measure [44]. As a result one can get a random pure factorized state $|F\rangle' = \hat{U}_f^{(4)} |F\rangle_0$. The need for $4 * 4^4$ unitary operators is driven by the requirement of Carathéodory's theorem for the convex sets and dimension of state space \mathcal{H}_{n_q} , see [22] for a detailed discussion. To sample an arbitrary separable four-qubit state we construct a convex hull: $\hat{\rho}_S = \sum_{i=1}^{4^4} p_i |F\rangle'_i \langle F|'_i$, where the probability distribution $\{p_i\}$ is generated from the Dirichlet distribution [33]. Note that it is possible to train SVM and find the EW operator by using only pure factorized states. It follows directly from the definition of separable states and properties of entanglement witnesses. We will return to this topic in the next subsection.

Summarizing, our central goal is to demonstrate that the SVM approach allows one to classify and detect the entanglement for an *arbitrary* four-qubit state, not only for pure states. In the next subsection we confirm that the presented above SVM approach can achieve this goal with high accuracy.

B. Constructing entanglement witness operators via linear support vector machines

To implement the linear SVM algorithm, we prepared a data set comprised of pairs of entangled and separable states. The data set had 20 000 samples of each family of states $\{G, E_i, i = 1, \dots, 8\}$ and the same number of samples of separable states S . Specifically, in the case of $G, E_1 - E_5$ we randomly sampled complex numbers a, b, c, d describing parametrization of a particular family (A1)–(A9). All states in the data set are normalized and converted to the density operator form as we mentioned above.

It is important to emphasize that, in principle, an increase of the training data set size will improve the overall performance of any algorithm. However, the computational

resources required for such an improvement were not available for the full scope of this work. The doubling of the training data set demonstrates at maximum $\approx 0.1\%$ improvements of overall accuracy but roughly doubling computational time, plus one requires additional time to tune the algorithm's parameters. Based on performance results we have chosen 20 000 as a reasonable number of samples. As a result, for each member of $\{E_i, i = 1, \dots, 8\}$, we constructed a corresponding set of normalized EW operators such that $|\text{tr}(\hat{\rho}\hat{W})| \leq 1$.

To train an SVM model for a given family, we split the corresponding data subset (for each pair of states) into three parts: 18 000 samples were used for the training set, 1000 samples for a *validation set* to tune hyperparameters and 1000 samples for a *test set* to check the generalization capabilities of an algorithm by evaluating the accuracy on another data set. We presented the results of the trained model performance via SVM in Figs. 3 and 4 for the cases of families E_3 and G . Histograms of E_3 and G consist of 75 bins versus $\text{tr}(\hat{\rho}\hat{W})$ covering the range of $|\text{tr}(\hat{\rho}\hat{W}_G)| \leq 1$. Note that for the both presented cases we have chosen the parameter $\varepsilon = 0.5$ in (6) to assemble data sets of states. We set the regularization parameter in (4), $\lambda = 0.5 * 10^{-4}$ to yield an acceptable generalization, based on the performance on the validation data set.

We used the Adam optimizer implemented in TENSORFLOW library [45] with the following parameters: learning rate = 0.005, epsilon = 10^{-6} . The total number of training steps, or epochs, was chosen to be equal to 20 000 steps, but an acceptable convergence of SVM algorithm with the batch size 50 was achieved after approximately 5000–7000 training steps. Based on the performance analysis on a validation set we utilized the following regularization strategy. Each batch consisting of 50 samples was randomly sampled from the whole data set for each epoch. This additional randomization strategy had demonstrated significant improvements on a test data set. It is clear from the performance presented in Fig. 3 that the resulting witness can reliably distinguish separable and entangled states. Let us further analyze the performance of the linear SVM algorithm. It is worth mentioning that, in accordance with the properties of multipartite EW (1) $f_W(\rho) \geq 0$ for all separable states but $f_W(\hat{\rho}_{\text{sep.}}) < 0$ for *at*

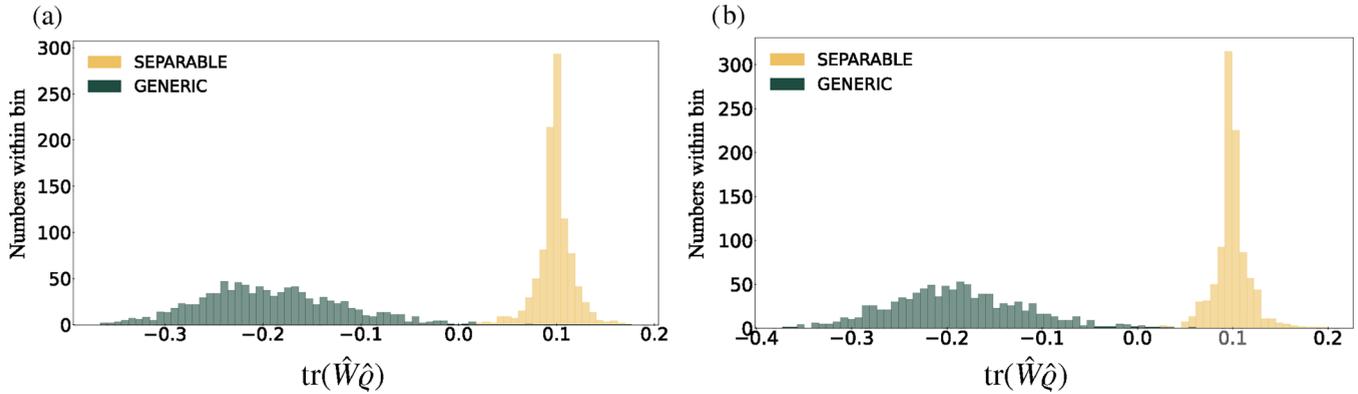


FIG. 4. Distribution of $\text{tr}(\hat{Q}\hat{W})$, mean value of entanglement witness, for (a) validation set and (b) test set of trained EW model via Linear SVM algorithm in the case of generic G family of states. Both validation and test set consist of 2000 samples: 1000 separable states and 1000 entangled states of E_3 family. For both validation and test sets there were only few (≤ 5) misclassifications of entangled states and zero misclassifications for separable states. Note that the training data set included mixed Werner states to achieve better generalization.

least one entangled state \hat{Q}_{ent} ; there does not exist an EW operator that can detect all entangled states. We aim to find the best possible EW operator that can detect the presence of entanglement in a maximum number of states of a given family of states.

Usually, the optimal EW implies an operator \hat{W}_{opt} that can detect the maximum number of entangled states [22]. In geometrical terms it means that the decision hyperplane is tangent to the set of all separable states S . Note that for a single entangled state it is guaranteed that this entangled state and all separable states can be linearly separated. On the other hand, in terms of linear SVM we tend to find the maximal distance to the closest sample of a given set of states from the decision boundary. It is known [33] that the linear SVM works perfectly for linearly separable data, but even for nonlinearly separable data it might perform quite well [41]. Below we will demonstrate that analysis of the SVM performance can help to determine to what extent the entangled states from a given family can be linearly separated from a set of all separable states.

The analysis presented here directly tests the ability to distinguish entangled and separable states for each SLOCC family individually. The SVM generalization performance is considered for the newly resampled entangled and separable states. In what follows, we analyze a variety of states correctly detected by EW without limiting them to only pure entangled states from families $\{G, E_i, i = 1, \dots, 8\}$, but also considering a specific family of mixed entangled states, Werner states. Note that doing this we also included mixed Werner states and pure factorized states in the construction of a training data set.

Let us analyze the performance of a linear SVM classifier to train the entanglement witnesses corresponding to each family of quantum states. We consider three different cases correspond to three parameters $\varepsilon_1, \varepsilon_2, \varepsilon_3$. In these cases we operate with different data sets of entangled states. The diversity of each sample is varying by parameter ε in Eq. (6), in accordance with the SLOCC classification criteria (5). We tested the procedure for the following parameters: $\varepsilon_1 = 0.5, \varepsilon_2 = 0.75, \varepsilon_3 = 1$. The results are presented in Fig. 5. It is clear from the figure that the ability to detect entanglement with an already trained EW operator for the corresponding

family members drops dramatically for $\varepsilon_3 = 1$: the score is equal to about 50% on average for all classes compared to almost 100% for $\varepsilon_1 = 0.5$. In these cases, the value of ε has a possible interpretation of ability to distinguish entangled states from separable ones employing linear SVM and, therefore, the EW theory approach.

The approach works similarly for all families $\{E_i\}, i = 1 \dots 8$. In other words, one can roughly estimate the vicinity for any state if it is a member of a given family of entangled states, for which a linear model can detect the presence of entanglement. We also analyzed the SVM generalization performance in the worst case scenario. In this case, each state is modified via Eq. (5) but each unitary operator $\hat{U}_f^{(4)} \equiv \hat{U}_1 \otimes \hat{U}_2 \otimes \hat{U}_3 \otimes \hat{U}_4$ is distributed according to the Haar measure. As expected, in this case we have a much inferior performance: scores $\leq 1\%$ for all SLOCC families. However, the correct detection scores for separable and factorized states have not changed and are still high, about 99.5%. Such behavior agrees with the EW properties.

A pure factorized state always has an infinitely large number of entangled states within ε , the vicinity, calculated, for example, with respect to the trace norm or Hilbert-Schmidt norm. This may lead to inferior performance of the SVM algorithm in its attempt to derive EWs capable of detecting the *mixed entangled states*. Consequently, to improve the algorithm further we have included mixed entangled states in our training model. For instance, let us consider another important family of multipartite quantum states \hat{Q}_W , so called Werner states [46,47], which involve entanglement of mixed states.

It is known [1] that, in general, the detection of entanglement presence in Werner states utilizing entanglement witnesses is not an easy task. Werner states are invariant under the diagonal action of the unitary group $U(2) \otimes U(2)$. Thus, in the case of four-qubit quantum states the Werner states are defined as follows:

$$\hat{Q}_W = [\hat{U} \otimes \hat{U}] \hat{Q}_W [\hat{U}^\dagger \otimes \hat{U}^\dagger], \quad (7)$$

where $\hat{U} = \hat{U}_1 \otimes \hat{U}_2, U_1, U_2 \in U(2)$.

To construct the sample data sets of both entangled and separable Werner states we followed derivations presented in

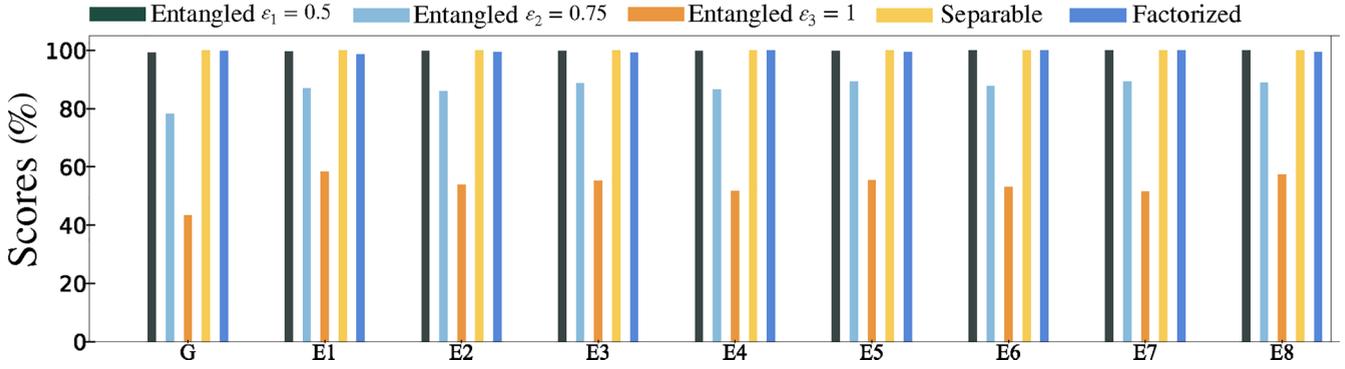


FIG. 5. General analysis of correct detection of entangled states based on trained set of EW operators. For each family G, E_1, \dots, E_8 we used 20000 new samples of states. Each state is modified according to Eq. (5). The random unitary operators are controlled by parameters ϵ according to Eq. (6). The depicted bars represent the probability (or score) to correctly detect entangled states of a given family for various values of ϵ : 0.5, 0.75, 1, correspondingly for the bars 1, 2, 3 starting from the left. For each family index $G, E_1 - E_8$ the fourth bar 4 corresponds to mixed separable states, the fifth bar 5 corresponds to pure separable (factorized) states. The set of EW is obtained employing SVM with parameters described in the caption of Figs. 3 and 4.

Ref. [47]. An arbitrary Werner state can be obtained by action of a *twirl channel*: $\hat{\rho}_W = \tau[\hat{\rho}_W] = \int [\hat{U} \otimes \hat{U}] \hat{\rho} [\hat{U} \otimes \hat{U}] d_{\text{Haar}} \mathbf{U}$ where $d_{\text{Haar}} \mathbf{U} = dU_1 dU_2$, invariant Haar measure on group $U(2) \otimes U(2)$. This operation can be considered as a projection on $\hat{U} \otimes \hat{U}$, an invariant subspace.

On the other hand, any Werner state can be represented as a decomposition $\hat{\rho}_W = \sum_{\alpha} q_{\alpha} \hat{Q}_{\alpha}$, where $\{\hat{Q}_{\alpha}\}$ is a set of four-partite orthogonal projectors that characterize $\hat{U} \otimes \hat{U}$ invariant subspace, see Ref. [47] where an explicit form of the operators is presented). It is evident that action of the twirling channel on a separable state results in a separable state. Thus, as it was shown in Ref. [47], the Werner state is separable if and only when the following conditions hold: $q_1 \leq 1, q_2, q_3 \leq 1/2, q_4 \leq 1/4; q_4 \leq q_2, q_3 \leq 1$ for mentioned above decomposition of $\hat{\rho}_W$. In our analysis we applied this criteria to sample both separable and entangled states by considering the following procedures. To sample the separable Werner states one can sample a simple four-qubit separable state and then apply a twirling channel.

The action of a twirling channel is equivalent to projection of a quantum state $\hat{\rho}$ by using operators $\{\hat{Q}_{\alpha}\}$ with coefficients $q_{\alpha}(\hat{\rho}) = \text{tr}(\hat{Q}_{\alpha} \hat{\rho}), \alpha = 1, \dots, 4$. to obtain corresponding $\hat{\rho}_W(\hat{\rho}) = \sum_{\alpha} q_{\alpha}(\hat{\rho}) \hat{Q}_{\alpha}$. Such construction of Werner states works because the following properties of twirling channel hold. First, the dual channel to the twirling channel is again a twirling channel. Second, any \hat{Q}_{α} is invariant under action of a twirling channel. Entangled Werner states are obtained by sampling a random four-qubit state and collecting states for which the separability criteria had been violated.

Utilizing already trained witness operators for each SLOCC family we found that an entanglement detection for the random entangled Werner states yielded poor accuracy. We obtained correct entanglement detection scores of just $\approx 1\%$ based on a trained witness operator that corresponded to the family G, for all other families of states we got scores $\approx 0.1\%$. On the other hand, the separable states are again classified with almost 100% accuracy. Summarizing, almost all sampled entangled Werner states were classified as separable states. Thus, it was essential to train the SVM linear modes for the Werner states and consider such states as a separate family.

The performance results of trained EW in the case of Werner states is presented in Fig. 6. It is not surprising that the linear SVM performance was poor because the Werner states are not linearly separable in the space of trace class operators. Nevertheless, it is clear from the separability criteria above (see also details in work [47]) that the separable and entangled Werner states Ref. can be distinguished in invariant subspace $\hat{U} \otimes \hat{U}$. Furthermore, one may project subsets of states parameterized with $\vec{\beta}$ from subsets $(S_{\vec{\beta}})$ of a given SLOCC family $\{\hat{\rho}(E_i, \vec{\beta}), \vec{\beta} \in S_{\vec{\beta}} \subseteq \mathcal{R}^m\}$ onto invariant subspace $\hat{U} \otimes \hat{U}$. In a such case one may assume that even a linear SVM model can yield good performance. Indeed, one may expect that knowing a trained entanglement witness $\hat{W}(\vec{\beta}) : \text{tr}(\hat{W}(\vec{\beta}) \hat{\rho}(\vec{\beta})) < 0, \forall \vec{\beta} \in S_{\vec{\beta}}$ one may construct an EW operator for the corresponding subset of Werner states obtained via a twirling channel $\hat{\rho}_W(\vec{\beta}) = \tau[\hat{\rho}(\vec{\beta})]$. Mathematically, it is evident from the following example. By choosing a family $\hat{\rho}(\vec{\beta}) \in G, E_i, i = 1, \dots, 8, \forall \vec{\beta} \in S_{\vec{\beta}}$, and assuming that there exist an EW operator $\hat{W}'(\vec{\beta})$ such that $\text{sign}[\text{tr}(\tau[\hat{\rho}(\vec{\beta})] \hat{W}'(\vec{\beta}))] = \text{sign}[\text{tr}(\hat{\rho}(\vec{\beta}) \tau[\hat{W}'(\vec{\beta})])] < 0 \forall \vec{\beta} \in S_{\vec{\beta}}$. In the above equality we utilized the self-duality property of a twirling channel by swapping action of $\tau[\hat{\rho}(\vec{\beta})] \rightarrow \tau[\hat{W}'(\vec{\beta})]$ under the trace operation. Thus, one may imply that $\tau[\hat{W}'(\vec{\beta})] \subset \text{span}(\{\hat{P}_{\lambda_j}(\hat{W}(\vec{\beta}))\})$. We denoted "span(\cdot)" as linear span for a set of operators.

In our case these sets of operators $\{\hat{P}_{\lambda_j}(\hat{W}(\vec{\beta}))\}$ are orthogonal projectors on an eigenspace that corresponds to an eigenvalue $\lambda_j \in \text{spec}(\hat{W}(\vec{\beta}))$. It is evident that if the set $S_{\vec{\beta}}$ is simple enough the correspondence can be specified straightforwardly. As we can see from Fig. 6 the direct numerical evaluation of linear SVM algorithm support the discussed assumption. Note, an approach discussed above might be generalized further to construct more sophisticated algorithm for entanglement detection of multipartite quantum states.

In this subsection, we considered several families of entangled states, including entangled Werner states. Our analysis clearly indicates that a linear SVM algorithm can be efficiently employed to construct a set of EW operators. However, it was also demonstrated that the approach has fundamental

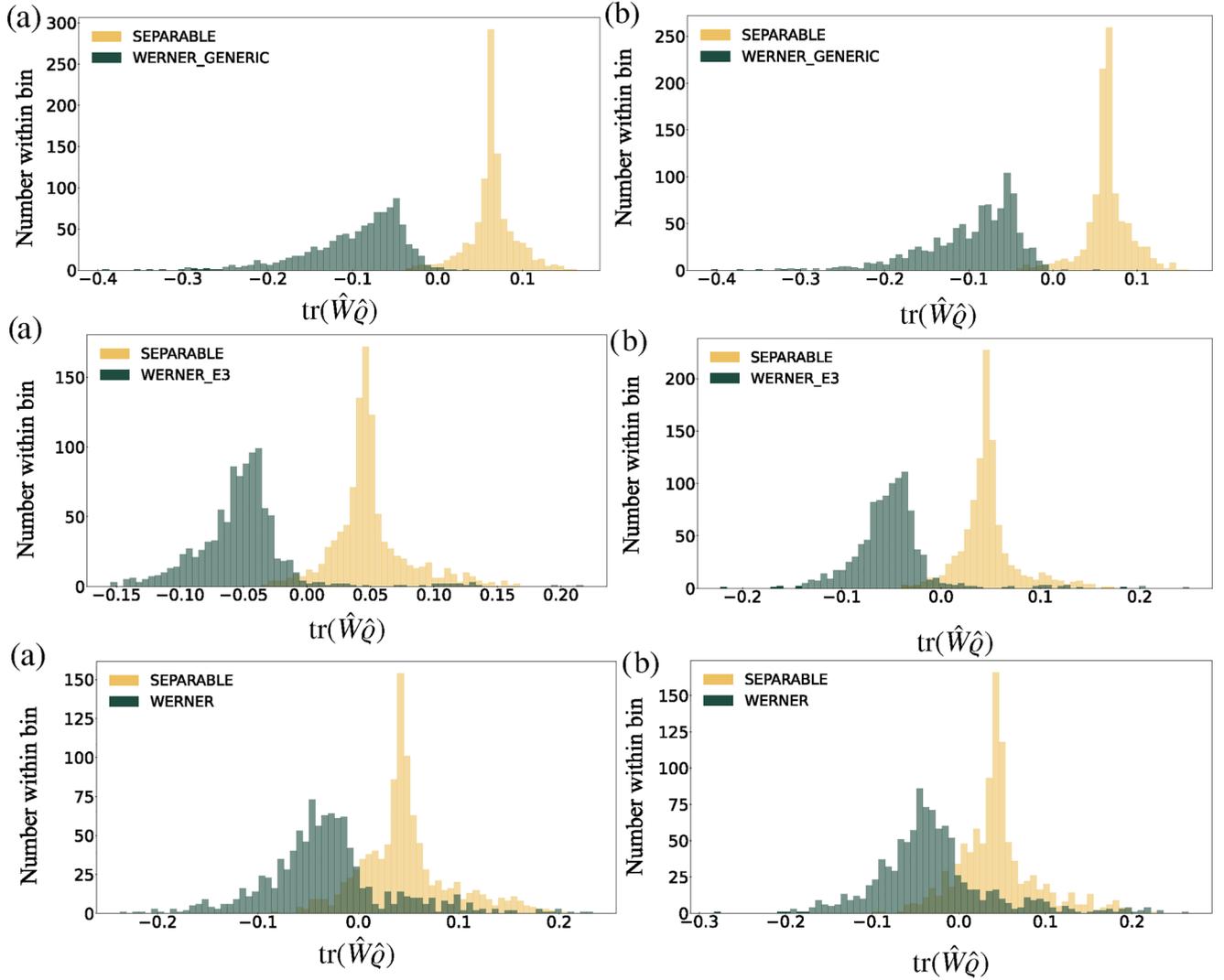


FIG. 6. Distribution of $\text{tr}(\hat{Q}\hat{W})$, mean value of entanglement witness, for (a) validation set and (b) test set of trained EW model via the linear SVM algorithm in the case of Werner family of entangled states. All parameters are the same as was described in Figs. 3 and 4. For the first and second rows in a plot the corresponding Werner states are generated from Generic-G and E3 family of states by applying twirling channel and separability criteria, in a third row provides performance in the case of general sampling strategy described in a main text.

limitations. It is evident because each family of states is quite complex and cannot be separated by applying just a linear SVM algorithm. In the following, we discuss possible non-linear extensions of the SVM algorithms applicable to the problem of entanglement detection.

C. Constructing entanglement witness operators via support vector machines with nonlinear kernels

Let us consider possible improvements of the discussed above approach. The SVM method is a quadratic constrained programming problem, which can be seen from Eq. (4). Importantly, the maximum margin problem has a dual representation operating with kernel functions [33,41]. We can upgrade our simple linear model (3) with the following expression:

$$f_{\vec{\phi}}(\vec{\varrho}) = (\vec{w}(\vec{\varrho}), \vec{\phi}(\vec{\varrho})), \quad (8)$$

where a map $\vec{\phi}(\cdot)$ transforms the initial feature space, but the overall model is still linear. One can write the error function

(loss) in the following form:

$$\tilde{L}(\vec{a}) = \sum_i a_i - \frac{1}{2} \sum_{n,m} a_n a_m t_n t_m k(\vec{\varrho}_n, \vec{\varrho}_m), \quad (9)$$

where we introduced a new vector $\vec{a} = (a_1, \dots, a_N)$ related to $\vec{w}(\vec{\varrho})$ as follows: $\vec{w}(\vec{\varrho}) = \sum_i a_i t_i \vec{\phi}(\vec{\varrho}_i)$. The symbol $k(\vec{\varrho}_n, \vec{\varrho}_m)$ represents a *kernel function*, which can be efficiently used to evaluate nonlinear transformations. The optimization (training by finding \vec{a}) of quadratic function (9) yields the following predictive model in terms of kernel function:

$$f_{\vec{\phi}}(\vec{\varrho}) \longrightarrow f_{k(\vec{\varrho}_n, \cdot)}(\vec{\varrho}) = \sum_i a_i t_i k(\vec{\varrho}_n, \vec{\varrho}). \quad (10)$$

Note that working with kernel functions allows one to avoid explicit operations with the featured space, though it could produce, in principle, more successful and complex models. For instance, it could happen that one finds a specific kernel that can perfectly distinguish separable states

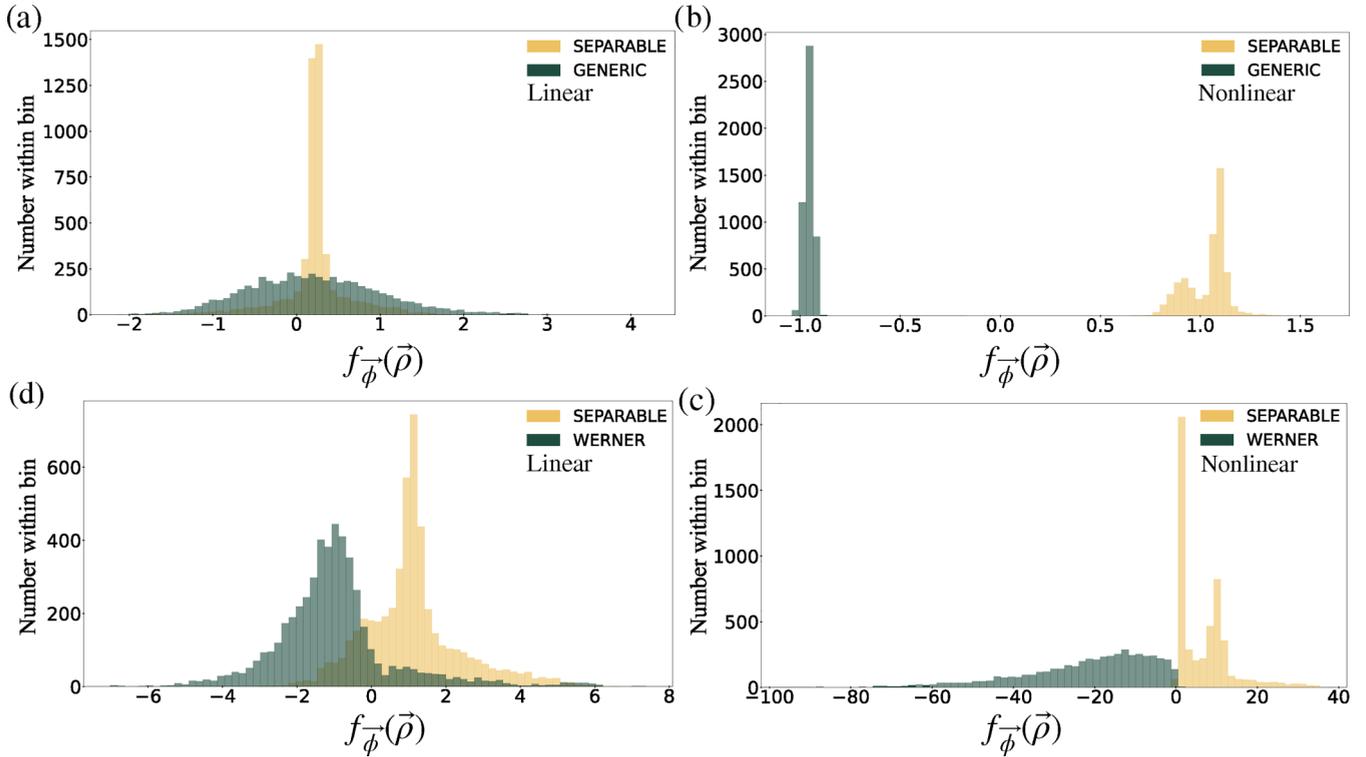


FIG. 7. Distribution of $f_{\phi}(\vec{\rho})$ described in Eqs. (8) and (10), and comparison of trained linear and nonlinear SVM algorithm with RBF kernel using SciKit Learn Python package. Cases (a) and (d) correspond to the standard SVM classification with linear kernel on a newly generated test set, whereas (b) and (c) correspond to the generic and entangled Werner states classified with a nonlinear SVM with RBF kernel. We verify results by sampling 5000 states of each family and for each case. Note that the assembled data set of Generic states is obtained by modifying each sample by unitary operators distributed according to the Haar measure and Eq. (5). It is clear that the performance of nonlinear SVM is very good, 100% and 99.8% in the cases of Generic and Werner states, respectively.

and the *whole* family of entangled states based on the SLOCC classification. Indeed, we verified this assumption directly applying a nonlinear SVM using SCIKIT LEARN package. Figure 7 provides a comparison of models with linear and nonlinear kernels to classify the entangled Werner, generic, and separable states. Furthermore, we modify the generic states according to the Haar measure to assemble a new data set of 1000 samples. We observe 100% accuracy on a test set in the case of nonlinear SVM with radial basis function (RBF) kernel: $k(\vec{\rho}, \vec{\sigma}) = \exp(-\gamma \|\vec{\rho} - \vec{\sigma}\|_2^2)$, compared to the linear model.

Unfortunately, there are constraints among the feature values that restrict the dimension of feature space so evaluating kernel functions for all samples might be computationally demanding. To reduce the problem complexity, one needs to find certain heuristics or another advanced coarse-grained classification. In the next section we summarize the obtained results and discuss possible applications and further steps.

III. DISCUSSION

Let us summarize the salient points regarding the problem of entanglement detection in multipartite states using the support vector machine algorithms as we considered in our work.

The entanglement detection problem is rather complex. Ideally, one would aim to find tangent hyperplanes at each point of a set of separable states. Several methods were

proposed to find an approximate distance between a given witness and a set of separable states, which help to detect the entanglement of a particular state [25]. In this context the considered SVM approach has a similar objective to the one mentioned above. The SVM aims to maximize the shortest distance between the decision boundary and “support” state (vector). However, as it was shown in Refs. [23–25] the characterization of all EWs for a given multipartite system has non-deterministic polynomial time hardness (NP-hard) problem as the dimensionality increases. The geometrical interpretation of this problem is also a highly nontrivial task [26].

Nevertheless, the EW problem complexity can be reduced by employing a coarse-grained classification such as the considered four-qubit SLOCC classification or description in terms of Werner states. We demonstrated successful applications of the linear SVM algorithm for both of these cases. It is important to mention that our analysis is akin to that proposed in Ref. [40], where the authors also emphasized the direct relationship of the entanglement witness functional and linear SVM. In our work here, we focus on the successful applicability of the SVM technique specifically for the coarse-grained classification of entangled states joined in to certain families. However, the success of linear models comes with a price owing to the above mentioned NP-hardness of the EW problem. In general, the linear SVM approach cannot provide high accuracy for all states in a particular family. As we

showed, one can detect the entanglement with high accuracy for all family members only within a vicinity of a fiducial state of the family. Thus, one can interpret the accuracy of a trained EW model based on linear SVM as a measure of complexity for a given family or parametrization of entangled states.

One of the main advantages of the discussed linear SVM approach is its universality in terms of implementation. Note that the modern Python packages such as QUTIP, TENSORFLOW, SCIKIT LEARN [42,43,48,49] provide all necessary tools for implementation of both linear and nonlinear SVM to find EW. The source code used for this work is also provided in a GitHub repository [50]. In addition, a nonlinear version of SVM can be a valuable tool for theoretical analysis, especially in the case of coarse-grained classifications. We demonstrated that nonlinear SVM allows us to detect the entanglement with almost perfect accuracy for any arbitrary state from a given family, while the linear SVM has almost zero accuracy. It is not surprising because an arbitrary family of states based on a given classification or parametrization is highly likely to be not linearly separable from a convex set of separable states. We also showed that introduction of the kernel function could be useful for specific cases. At the same time it may require considerable computing resources making the nonlinear kernels less attractive in general.

It is important to mention that local unitary transformations limit our analysis. These limitations are dictated by the simplicity of analysis and implementation of the SVM algorithm. Nevertheless, we believe that our results are sufficient to demonstrate the universality of the proposed approach. It is also worth noting that the previously described problematic Werner states are not members of the SLOCC family, as their definition is based on the unitary group structure described above. In addition, we tailored our work to specific practical problems, such as quantum astrometry [51–53], where the local unitary operations are reasonable models for the information encoding present in those environments. On the other hand, it is also crucial to emphasize that the general analysis of the symmetry of quantum states and operations, for example, by applying group theory, plays a prominent role in the construction of the coarse-grained description.

Application of the group theory and symmetry analysis are some of the most powerful tools in theoretical physics. In particular, these approaches led to a significant progress in developing novel theoretical and practical tools for quantum technologies and their applications, such as entanglement detection, foundations of quantum theory [54–57], randomized benchmarking [58–60], etc. The aim for the next stage of research is to elaborate on the developments of SVM, such as algorithms to train more sophisticated and, if possible, universal frameworks of classifiers to combine specific sets of multipartite quantum states. It is our hope that the synthesis of machine learning-based group theory methods and symmetry analysis of states and quantum channels can be valuable tools to accelerate further developments of quantum technology.

Another advantage of linear models is that they allow for a physical interpretation of EW operators, as observables that can be measured directly in an experiment. An interesting goal would be to analyze further the connection between kernel methods and optimal collective measurements within the concept of collective entanglement witnesses [61].

The presented SVM-based analysis of EW can be useful for practical applications. For instance, the SLOCC operations described above directly relate to the field of quantum metrology. In particular, optical interferometers were proposed where an entangled ancilla was shared between two or more stations to improve the accuracy of astrometrical measurements [51–53,62,63]. Each ancilla's subsystem interacts locally with a fiducial state in the station that carries valuable information to be extracted. For practical implementation of such schemes, one needs to understand how the noise affects the ancilla's state and also specifics of local interactions to optimize the measurement protocol and to quantify the structure of the entangled states. Simple and effective tools such as the proposed SVM approach can help to solve the problem of detecting and classifying the structure of entangled states used in those quantum astrometry schemes.

Finally, another promising application of this approach is its employment in quantum-enhanced sensor networks [64] and quantum reservoir computing [65,66]. It was already demonstrated that both SVM and EW problems could be mapped and processed by such multipartite quantum systems working as quantum-enhanced processors [64]. These promising developments open new directions for applications of multipartite quantum systems to address the cross-cutting interdisciplinary problems of machine learning and quantum metrology.

ACKNOWLEDGMENTS

We are grateful to M. Graydon, Z. Chen, A. Slosar, K. Katamadze, and R. Reimann for useful discussions and support. This work was supported in part by the U.S. Department of Energy QuantISED award and BNL LDRD Grant No. 19-30.

APPENDIX: SLOCC CLASSIFICATION OF FOUR-QUBIT PURE STATES

All trained SVM models discussed above were derived by combining all states of one family based on a given property such as the SLOCC equivalence or another particular state property (parametrization) intrinsic to the Werner states. This Appendix provides an explicit description of the SLOCC classification used in this manuscript. This classification was investigated in detail in Ref. [28], where the authors derived nine families of states corresponding to nine different ways of entangling four qubits.

The main idea of derivation follows from the equivalence of groups $SU(2) \otimes SU(2)$ and $SO(4)$ in the Lie-group theory, where $SU(2)$ is a special group of 2×2 unitary matrices and $SO(4)$ is a group of orthogonal matrices with unit determinant. Mathematically, this equivalence is represented by $\forall U_1, U_2 \in SU(2) \exists T : O = T(U_1 \otimes U_2)T^\dagger \in SO(4)$. Additionally, each four-qubit pure state can be represented as a 4×4 complex matrix $R = T\psi_{(i_1 i_2)(i_3 i_4)}T^\dagger$, where $\psi_{(i_1 i_2)(i_3 i_4)}$ is a reshaped matrix element of four-qubit state representation in the computational basis. Thus, the aforementioned equivalence and state representation allows one to represent the SLOCC transformation of four-qubit state in (5) as an action of orthogonal matrices: $R' = O_1 R O_2$. On the other hand it was proven that a given 4×4

complex matrix R can always be transformed via O_1 and O_2 to the Jordan block normal form. This means that the normal form encodes the genuine nonlocal properties of the four-qubit state and allows classification of the states. Entangled

states from the same family can perform the same quantum protocol but with a different probability. The following distinct classes of (unnormalized) pure states were specified in the computational basis:

$$|G_{abcd}\rangle = \frac{a+d}{2}(|0000\rangle + |1111\rangle) + \frac{a-d}{2}(|0011\rangle + |1100\rangle) + \frac{b+c}{2}(|0101\rangle + |1010\rangle) + \frac{b-c}{2}(|0110\rangle + |1001\rangle), \quad (\text{A1})$$

$$|E_1\rangle = \frac{a+b}{2}(|0000\rangle + |1111\rangle) + \frac{a-b}{2}(|0011\rangle + |1100\rangle) + c(|0101\rangle + |1010\rangle) + |0110\rangle, \quad (\text{A2})$$

$$|E_2\rangle = a(|0000\rangle + |1111\rangle) + b(|0101\rangle + |1010\rangle) + |0110\rangle + |0011\rangle, \quad (\text{A3})$$

$$|E_3\rangle = a(|0000\rangle + |1111\rangle) + \frac{a+b}{2}(|0101\rangle + |1010\rangle) + \frac{a-b}{2}(|0110\rangle + |1001\rangle) + \frac{i}{\sqrt{2}}(|0001\rangle + |0010\rangle + |0111\rangle + |1011\rangle), \quad (\text{A4})$$

$$|E_4\rangle = a(|0000\rangle + |0101\rangle + |1010\rangle + |1111\rangle) + (i|0001\rangle + |0110\rangle - i|1011\rangle), \quad (\text{A5})$$

$$|E_5\rangle = a(|0000\rangle + |1111\rangle) + (|0011\rangle + |0101\rangle + |0110\rangle), \quad (\text{A6})$$

$$|E_6\rangle = |0000\rangle + |0101\rangle + |1000\rangle + |1110\rangle, \quad (\text{A7})$$

$$|E_7\rangle = |0000\rangle + |1011\rangle + |1101\rangle + |1110\rangle, \quad (\text{A8})$$

$$|E_8\rangle = |0000\rangle + |0111\rangle, \quad (\text{A9})$$

$$|S_0\rangle = |0110\rangle \equiv |E\rangle_1 : a = b = c = 0. \quad (\text{A10})$$

The state $|S_0\rangle$ represents a subclass of the pure factorized state. All other states (classes) possess unique properties and entanglement structure. For instance, the state $|G\rangle_{abcd}$ represents a class of generic pure states. It is claimed that this is a class of states with maximal four-partite entanglement on the orbit generated by SLOCC measured in accordance with majorization criteria [28].

-
- [1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).
- [2] R. Jozsa and N. Linden, *Proc. R. Soc. London A* **459**, 2011 (2003).
- [3] H. J. Briegel and R. Raussendorf, *Phys. Rev. Lett.* **86**, 910 (2001).
- [4] A.-N. Zhang, C.-Y. Lu, X.-Q. Zhou, Y.-A. Chen, Z. Zhao, T. Yang, and J.-W. Pan, *Phys. Rev. A* **73**, 022330 (2006).
- [5] S.-H. Wei, B. Jing, X.-Y. Zhang, J.-Y. Liao, C.-Z. Yuan, B.-Y. Fan, C. Lyu, D.-L. Zhou, Y. Wang, G.-W. Deng *et al.*, *Laser Photonics Rev.* **16**, 2100219 (2022).
- [6] X.-M. Hu, C.-X. Huang, Y.-B. Sheng, L. Zhou, B.-H. Liu, Y. Guo, C. Zhang, W.-B. Xing, Y.-F. Huang, C.-F. Li *et al.*, *Phys. Rev. Lett.* **126**, 010503 (2021).
- [7] S. Lloyd, *Phys. Rev. Lett.* **90**, 167902 (2003).
- [8] W. McCutcheon, A. Pappa, B. A. Bell, A. McMillan, A. Chailloux, T. Lawson, M. Mafu, D. Markham, E. Diamanti, I. Kerenidis *et al.*, *Nat. Commun.* **7**, 13251(2016).
- [9] J. Illiano, M. Caleffi, A. Manzalini, and A. S. Cacciapuoti, *Computer Networks* **213**, 109092 (2022).
- [10] P. Drews, G. Williams, B. Goldfain, E. A. Theodorou, and J. M. Rehg, *IEEE Robotics and Automation Letters* **4**, 1564 (2019).
- [11] G. Tóth, *Phys. Rev. A* **85**, 022322 (2012).
- [12] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi, *Phys. Rev. A* **85**, 022321 (2012).
- [13] D. Gross, S. T. Flammia, and J. Eisert, *Phys. Rev. Lett.* **102**, 190501 (2009).
- [14] W. Dür and H. J. Briegel, *Rep. Prog. Phys.* **70**, 1381 (2007).
- [15] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, *Nat. Rev. Phys.* **2**, 382 (2020).
- [16] M. Kliesch and I. Roth, *PRX Quantum* **2**, 010201 (2021).
- [17] O. Gühne, Y. Mao, and X.-D. Yu, *Phys. Rev. Lett.* **126**, 140503 (2021).
- [18] D. Cruz, R. Fournier, F. Gremion, A. Jeannerot, K. Komagata, T. Tomic, J. Thiesbrummel, C. L. Chan, N. Macris, M.-A. Dupertuis *et al.*, *Adv. Quantum Technol.* **2**, 1900015 (2019).
- [19] O. Gühne and G. Tóth, *Phys. Rep.* **474**, 1 (2009).
- [20] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223**, 1 (1996).
- [21] B. M. Terhal, *Phys. Lett. A* **271**, 319 (2000).
- [22] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory* (Cambridge University Press, Cambridge, 2009).

- [23] L. M. Ioannou, B. C. Travaglione, D. C. Cheung, and A. K. Ekert, *Phys. Rev. A* **70**, 060303(R) (2004).
- [24] L. M. Ioannou and B. C. Travaglione, *Phys. Rev. A* **73**, 052314 (2006).
- [25] L. M. Ioannou, *Quantum Inf. Comput.* **7**, 335 (2007).
- [26] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2017).
- [27] W. Dür, G. Vidal, and J. I. Cirac, *Phys. Rev. A* **62**, 062314 (2000).
- [28] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, *Phys. Rev. A* **65**, 052112 (2002).
- [29] L. Lamata, J. León, D. Salgado, and E. Solano, *Phys. Rev. A* **75**, 022318 (2007).
- [30] M. Sanz, I. L. Egusquiza, R. D. Candia, H. Saberi, L. Lamata, and E. Solano, *Sci. Rep.* **6**, 30188 (2016).
- [31] M. Hebenstreit, M. Gachechiladze, O. Gühne, and B. Kraus, *Phys. Rev. A* **97**, 032330 (2018).
- [32] C. Cortes and V. Vapnik, *Mach. Learn.* **20**, 273 (1995).
- [33] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)* (Springer, New York, 2007), 1st ed., ISBN 0387310738.
- [34] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [35] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, *Nat. Phys.* **14**, 447 (2018).
- [36] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [37] I. A. Luchnikov, S. V. Vintskevich, D. A. Grigoriev, and S. N. Filippov, *Phys. Rev. Lett.* **124**, 140502 (2020).
- [38] Y.-C. Ma and M.-H. Yung, *npj Quantum Inf.* **4**, 34(2018)
- [39] S. Lu, S. Huang, K. Li, J. Li, J. Chen, D. Lu, Z. Ji, Y. Shen, D. Zhou, and B. Zeng, *Phys. Rev. A* **98**, 012315 (2018).
- [40] E. Y. Zhu, L. T. Wu, O. Levi, and L. Qian, [arXiv:2107.02301](https://arxiv.org/abs/2107.02301).
- [41] A. Géron, *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (O'Reilly Media, Sebastopol, CA, 2019).
- [42] J. Johansson, P. Nation, and F. Nori, *Comput. Phys. Commun.* **183**, 1760 (2012).
- [43] J. Johansson, P. Nation, and F. Nori, *Comput. Phys. Commun.* **184**, 1234 (2013).
- [44] M. Fasi and L. Robol, *Linear Algebra Appl.* **620**, 297 (2021).
- [45] TensorFlow Developers, TensorFlow, <https://zenodo.org/record/4724125>.
- [46] F. Huber, I. Klep, V. Magron, and J. Volčič, *Commun. Math. Phys.* **396**, 1051 (2022).
- [47] D. Chruściński and A. Kossakowski, *Phys. Rev. A* **73**, 062314 (2006).
- [48] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems (2015), software available from tensorflow.org, <https://www.tensorflow.org/>.
- [49] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, *J. Mach. Learn. Research* **12**, 2825 (2011).
- [50] GitHub, Classification of four-qubit entangled states via machine learning: Source code, <https://github.com/StephenVintskevich/qCSC>.
- [51] P. Stankus, A. Nomerotski, A. Slosar, and S. Vintskevich, *Open J. Astrophys.* **5** (2022).
- [52] E. T. Khabiboulline, J. Borregaard, K. De Greve, and M. D. Lukin, *Phys. Rev. Lett.* **123**, 070504 (2019).
- [53] E. T. Khabiboulline, J. Borregaard, K. De Greve, and M. D. Lukin, *Phys. Rev. A* **100**, 022316 (2019).
- [54] M. A. Graydon and D. M. Appleby, 33LT02 (2016).
- [55] M. A. Graydon, [arXiv:1703.06800](https://arxiv.org/abs/1703.06800).
- [56] M. A. Graydon, J. Skanes-Norman, and J. J. Wallman, Clifford groups are not always 2-designs, [arXiv:2108.04200](https://arxiv.org/abs/2108.04200).
- [57] M. A. Graydon, J. Skanes-Norman, and J. J. Wallman, Designing stochastic channels, [arXiv:2201.07156](https://arxiv.org/abs/2201.07156).
- [58] A. Erhard, J. J. Wallman, L. Postler, M. Meth, R. Stricker, E. A. Martinez, P. Schindler, T. Monz, J. Emerson, and R. Blatt, *Nat. Commun.* **10**, 5347 (2019).
- [59] R. Harper, S. T. Flammia, and J. J. Wallman, *Nat. Phys.* **16**, 1184 (2020).
- [60] A. Winick, J. J. Wallman, and J. Emerson, *Phys. Rev. Lett.* **126**, 230502 (2021).
- [61] M. Horodecki, K. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), and U. Sen, *Phys. Rev. Lett.* **90**, 100402 (2003).
- [62] D. Gottesman, T. Jennewein, and S. Croke, *Phys. Rev. Lett.* **109**, 070503 (2012).
- [63] A. Nomerotski, P. Stankus, A. Slosar, S. Vintskevich, S. Andrews, G. Carini, D. Dolzhenko, D. England, E. Figueroa, S. Gera *et al.*, in *Optical and Infrared Interferometry and Imaging VII*, (SPIE, 2020), Vol. 11446, pp. 290–306, <https://doi.org/10.1117/12.2560272>.
- [64] Q. Zhuang and Z. Zhang, *Phys. Rev. X* **9**, 041023 (2019).
- [65] S. Ghosh, A. Opala, M. Matuszewski, T. Paterek, and T. C. H. Liew, *npj Quantum Inf.* **5**, 35 (2019).
- [66] S. Vintskevich and D. Grigoriev, [arXiv:2201.07969](https://arxiv.org/abs/2201.07969).