A Variational Approach to the Quantum Separability Problem

Mirko Consiglio,^{1,*} Tony J. G. Apollaro,¹ and Marcin Wieśniak^{2,3}

¹Department of Physics, University of Malta, Msida MSD 2080, Malta

²Institute of Theoretical Physics and Astrophysics, Faculty of Mathematics,
Physics, and Informatics, University of Gdańsk, 80-308 Gdańsk, Poland

³International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-308 Gdańsk, Poland

(Dated: November 18, 2022)

We present the variational separability verifier (VSV), which is a novel variational quantum algorithm (VQA) that determines the closest separable state (CSS) of an arbitrary quantum state with respect to the Hilbert–Schmidt distance (HSD). We first assess the performance of the VSV by investigating the convergence of the optimization procedure for Greenberger–Horne–Zeilinger (GHZ) states of up to seven qubits, using both statevector and shot-based simulations. We also numerically determine the CSS of maximally-entangled mixed X-states (X-MEMS), and subsequently use the results of the algorithm to surmise the analytical form of the aforementioned CSS. Our results indicate that current noisy intermediate-scale quantum (NISQ) devices may be useful in addressing the NP-hard full separability problem using the VSV, due to the shallow quantum circuit imposed by employing the destructive SWAP test to evaluate the HSD. The VSV may also possibly lead to the characterization of multipartite quantum states, once the algorithm is adapted and improved to obtain the closest k-separable state (k-CSS) of a multipartite entangled state.

I. INTRODUCTION

Entanglement is the principal defining feature of quantum mechanics [1], and a quantitative description of the phenomenon started with Bell's inequalities [2]. Different measures of entanglement quantify distinct resources, although most require the fulfillment of specific conditions to be considered "good" entanglement measures. These conditions can be defined as a set of axioms, such as the nullification of the measure for separable states, invariance under local unitary operations, and monotonicity under local operations and classical communication (LOCC) [3]. To this end, many entanglement measures have been defined, such as distillable entanglement, entanglement cost, and the entanglement of formation [4, 5]. The difficulty in calculating these measures for generic states lies in their non-closed form, given that they generally include the computation of a minimum over a large Hilbert space, which is typically intractable as the dimension of the system scales. On the other hand, the problem of measuring entanglement is more approachable for pure states, given that even for different entanglement classes, that is under stochastic local operations and classical communication (SLOCC) [5], many different measures of entanglement can be computed.

The simplest system that has been completely characterized is a two-qubit one. An interesting property for two qubits is that only two SLOCC classes exist, one which contains the only type of entanglement possible, i.e. bipartite entanglement, and the other consists of fully separable states, i.e. non-entangled states. The relevant literature for three- and four-qubit pure states, as well as the general multipartite case can be found in Refs. [6–9], [10–14], and [8, 15–18] respectively. Only highly

specific mixed states are known to have closed forms of multipartite entanglement measures, such as combinations of Greenberger–Horne–Zeilinger (GHZ) and W states [19], GHZ-symmetric states [20], and maximally-entangled mixed X-states (X-MEMS) [21], among a few others. It is also known that any arbitrary teleportation protocol [22], using an arbitrary multipartite entanglement channel requires an entanglement monotone termed localizable concurrence [23], as a quantum resource [24].

We can also relax our specificity of requiring an exact quantifier of entanglement, and focus on whether a particular state is certified to be entangled. This is known as the quantum separability problem, and is in fact deemed to be NP-hard, even for the bipartite case [25–27]. Consequently, verifying a state is fully separable requires that the state is separable with respect to all bipartitions, which implies that the full separability problem is at least as hard as the bipartite separability problem. As a result, we look towards variational quantum algorithms (VQAs), which incorporate hybrid quantum—classical computation aimed at harnessing the power of noisy intermediate-scale quantum (NISQ) computers [28], to solve challenging computational problems.

To this end, we describe a novel VQA, the variational separability verifier (VSV), capable of determining the closest separable state (CSS) of an arbitrary quantum state (assuming it can be prepared on a quantum device), with respect to the Hilbert–Schmidt distance (HSD) [29]. In addition, the HSD induces an entanglement measure, denoted as the Hilbert–Schmidt entanglement (HSE) [29, 30]. The HSE is still up for debate whether it is an entanglement monotone [31], however, it could in some cases: quantify the amount of entanglement present in specific states; behave as a separability witness; or else provide useful constructions of entanglement witnesses [32].

We will also briefly mention the prospect of extending

^{*} mirko.consiglio@um.edu.mt

the VSV to find the closest k-separable state (k-CSS), while for the time being, the algorithm is designed to find the closest fully separable state, i.e. the 1-CSS, which for the sake of brevity we refer to as the CSS. Although we investigate only qubit systems, the VSV can be adapted to find the CSS for qudit systems, however, this requires an encoding of the d-level system onto a 2-level system, as well as adjustments to the gates needed to generate (k-)separable qudit states.

The paper is organized as follows: Sec. II consists of definitions for the HSD, the induced HSE measure, and the CSS, along with a discussion on why we use the HSD as our metric for the VSV. In Sec. III, we introduce the framework of the VSV, specifically how minimizing the HSD over the set of fully separable states, with respect to a state ρ , leads to the CSS of ρ and its HSE. Since the computation of the HSD requires the calculation of state overlaps, we discuss how this is tackled on a quantum device using the destructive SWAP test [33, 34]. We then introduce the variational form of the algorithm and explain the corresponding bilevel optimization [35] approach. Lastly, we discuss some of the computational complexity aspects of the VSV, such as the scaling of the number of parameters with respect to the dimension of the system. In Sec. IV we present the results of the VSV applied to n-qubit GHZ states, ranging from two to seven qubits with statevector optimization, and from two to five qubits with shot-based optimization. We also apply the VSV to two- and three-qubit X-MEMS. Since the analytical form of the CSS for n-qubit X-MEMS is not known, we employed a technique similar to the one in Ref. [11] to surmise the analytical form of the CSS in Appendix A. Furthermore, we show that the genuine multipartite entanglement (GME) concurrence [5] and HSE are not related in the case of three-qubit X-States [21]. Finally, in Sec. V, we summarize the results of the paper, and acknowledge the inspiration of the VSV: the quantum Gilbert algorithm (QGA) [32, 36–38], which is further discussed in Appendix B.

II. THE HILBERT-SCHMIDT DISTANCE AS A MEASURE OF ENTANGLEMENT

The HSD between two quantum states $D_{\text{HS}}(\rho, \sigma)$, is defined as

$$D_{\mathrm{HS}}(\rho, \sigma) \equiv \mathrm{Tr}\left\{ (\rho - \sigma)^2 \right\},$$
 (1)

which is a non-monotonic Riemannian metric [29]. Using this definition, an entanglement measure can be induced by the HSD [29, 30], denoted as the HSE,

$$E_{\mathrm{HS}}\left(\rho\right) \equiv \min_{\sigma \in \mathcal{C}} D_{\mathrm{HS}}\left(\rho, \sigma\right),$$
 (2)

with the CSS defined as

$$\rho_{\text{CSS}} \equiv \underset{\sigma \in \mathcal{C}}{\operatorname{arg \, min}} \ D_{\text{HS}} \left(\rho, \sigma \right). \tag{3}$$

where \mathcal{C} is the convex set of states for some degree k of separability. To define it concretely, a state ρ is said to be k-separable if it can be written as a convex sum of k-separable states, that is

$$\rho = \sum_{i} p_{i} \left| \Psi_{i}^{k} \middle\langle \Psi_{i}^{k} \right|, \tag{4}$$

where a k-separable pure state can be written as a tensor product of k-local states

$$\left|\Psi_{i}^{k}\right\rangle = \bigotimes_{j=1}^{k} \left|\psi_{j}\right\rangle,$$
 (5)

such that $|\psi_j\rangle$ are k-separable states on subsets of the n-qubit parties [37]. More specifically, states are called biseparable for k=2, triseparable for k=3, up to fully separable for k=n. Although the k-CSS (3) (and corresponding HSE (2)) can be distinctly defined for all convex sets of k-separable states, as stated in Sec. I, we designed the VSV to find the CSS (and corresponding HSE) of entangled states with respect to the set of fully separable states.

It was proven in Ref. [31] that the HSD is not non-increasing under completely-positive trace-preserving (CPTP) maps, posing a principal question of whether the HSE is a good entanglement measure. Nevertheless, the HSD, and the corresponding HSE, are still useful quantities to investigate, since they are utilized in generalized Bell inequalities [39, 40], while also providing insight into the geometry of entangled states [29, 41]. The HSD is also relatively straightforward to evaluate on a quantum device, since it can be decomposed as

$$\operatorname{Tr}\left\{ (\rho - \sigma)^{2} \right\} = \operatorname{Tr}\left\{ \rho^{2} \right\} + \operatorname{Tr}\left\{ \sigma^{2} \right\} - 2\operatorname{Tr}\left\{ \rho\sigma \right\}, \quad (6)$$

coupled with using quantum primitives such as the destructive SWAP test to measure the overlap and purities of ρ and σ .

On the other hand, while it was shown in Ref. [41] that distance measures, such as the Bures measure of entanglement [29], are directly related to the fidelity [42] of a CSS, the fidelity is a harder quantity to evaluate on both a classical and quantum device [30, 33, 34, 43–45] when compared with the simplicity of the HSD.

It is also interesting to mention that the HSD is being used for tackling the quantum low-rank approximation problem [46], with applications in principal component analysis and preparing arbitrary mixed states on quantum computers [47].

III. FRAMEWORK OF THE ALGORITHM

In our procedure, we designate ρ as being the test state, and σ as our trial state(s), which are the states that iteratively approach the CSS. As a result, we choose to

prepare separable states on a quantum computer in the following way:

$$\sigma(\boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{\phi}) = \sum_{i=1}^{s} p_i |\psi(\boldsymbol{\theta}_i, \boldsymbol{\phi}_i)\rangle \langle \psi(\boldsymbol{\theta}_i, \boldsymbol{\phi}_i)|, \qquad (7)$$

where

$$\boldsymbol{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_s \end{pmatrix}, \quad \sum_{i=1}^s p_i = 1, \quad p_i \in [0, 1],$$

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \\ \vdots \\ \boldsymbol{\theta}_s \end{pmatrix} = \begin{pmatrix} \theta_{11} & \theta_{12} & \cdots & \theta_{1n} \\ \theta_{21} & \theta_{22} & \cdots & \theta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{s1} & \theta_{s2} & \cdots & \theta_{sn} \end{pmatrix}, \quad \theta_{ij} \in [0, 2\pi),$$

$$\boldsymbol{\phi} = \begin{pmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \\ \vdots \\ \boldsymbol{\phi}_s \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{s1} & \phi_{s2} & \cdots & \phi_{sn} \end{pmatrix}, \quad \phi_{ij} \in [0, 2\pi),$$

totaling s(2n+1) parameters, where n is the number of qubits and s is the number of separable pure states needed to generate the trial state. Due to Carathéodory's theorem, we require $s \leq d^2$ separable pure states [41, 48, 49], where $d=2^n$ is the dimension of the Hilbert space. However, in our simulations we found that s=d suffices to find the CSS of all our test states. $\{p_i, |\psi(\theta_i, \phi_i)\rangle\}$ represent the ensemble of separable pure states, which can be decomposed as

$$|\psi(\boldsymbol{\theta}_{i}, \boldsymbol{\phi}_{i})\rangle = \bigotimes_{j=1}^{n} \left(\cos\left(\theta_{ij}\right)|0_{j}\rangle + e^{\imath\phi_{ij}}\sin\left(\theta_{ij}\right)|1_{j}\rangle\right).$$
(8)

A separable pure state can thus be generated on a quantum computer by applying a set of one-qubit gates to the all-zero state:

$$|\psi(\boldsymbol{\theta}_i, \boldsymbol{\phi}_i)\rangle = \bigotimes_{j=1}^n U(\theta_{ij}, \phi_{ij}) |0\rangle_j,$$
 (9)

where

$$U(\theta_{ij}, \phi_{ij}) = \begin{pmatrix} \cos(\theta_{ij}) & -\sin(\theta_{ij}) \\ e^{i\phi_{ij}}\sin(\theta_{ij}) & e^{i\phi_{ij}}\cos(\theta_{ij}) \end{pmatrix}, \quad (10)$$

which is equivalent to applying an $R_y(\theta_{ij})$ gate followed by an $R_z(\phi_{ij})$ gate.

A. Measuring the Hilbert-Schmidt Distance

Preparing a mixed state on a quantum computer, requires, in general, twice as many qubits when compared

to the size of the state [50]. This, coupled with the issue of preparing an arbitrary separable state, leads us to find an alternative method for computing the overlap of the trial states with the test state. Our solution is to individually supply separable (non-orthogonal) pure states of our separable mixed state on the quantum computer, evaluate the individual overlaps using the destructive SWAP test, and then classically mix them during the optimization procedure. The benefit of this method is that we do not require the direct preparation of the CSS on the quantum computer, as we are only interested in the computation of the HSD.

Suppose, we can decompose the trial state σ as in Eq. (7), then we can evaluate the overlap between ρ and σ as

$$\operatorname{Tr}\{\rho\sigma\} = \operatorname{Tr}\left\{\rho \sum_{i=1}^{s} p_{i} |\psi_{i}\rangle\langle\psi_{i}|\right\}$$

$$= \sum_{i=1}^{s} p_{i} \operatorname{Tr}\{\rho |\psi_{i}\rangle\langle\psi_{i}|\}$$

$$= \sum_{i=1}^{s} p_{i} \langle\psi_{i}|\rho|\psi_{i}\rangle. \tag{11}$$

If we want to compute the purity of σ , then

$$\operatorname{Tr}\left\{\sigma^{2}\right\} = \operatorname{Tr}\left\{\sum_{i=1}^{s} p_{i} |\psi_{i}\rangle\langle\psi_{i}| \sum_{j=1}^{s} p_{j} |\psi_{j}\rangle\langle\psi_{j}|\right\}$$

$$= \sum_{i,j=1}^{s} p_{i}p_{j} \operatorname{Tr}\left\{|\psi_{i}\rangle\langle\psi_{i}|\psi_{j}\rangle\langle\psi_{j}|\right\}$$

$$= \sum_{i,j=1}^{s} p_{i}p_{j} |\langle\psi_{i}|\psi_{j}\rangle|^{2}$$

$$= \sum_{i=1}^{s} p_{i}^{2} + \sum_{i\neq j}^{s} p_{i}p_{j} |\langle\psi_{i}|\psi_{j}\rangle|^{2}$$

$$= \sum_{i=1}^{s} p_{i}^{2} + 2\sum_{i\neq j}^{s} p_{i}p_{j} |\langle\psi_{i}|\psi_{j}\rangle|^{2}. \tag{12}$$

The purity of ρ is trivially obtained assuming we can prepare it directly on a quantum computer; given that it is our test state.

B. Measuring Overlaps on a Quantum Computer

Given Eqs (11) and (12), we require subroutines in the VSV capable of measuring the purity and overlap of quantum states. Specifically, one can utilize the destructive SWAP test [33, 34] to calculate these quantities. The concept of this procedure stems from the fact that measuring in the Bell basis determines the amount of correlations present between two systems, and it can be shown that it is equivalent to the non-destructive SWAP

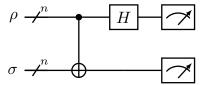


FIG. 1. Destructive SWAP test for determining the overlap between two quantum states. The bundled CNOT gate here actually represents CNOT gates between each pair of qubits of ρ and σ , with the bundled Hadamard gate equating to performing a Hadamard gate on each qubit of ρ . This effectively reduces to a Bell basis transformation for each pair of qubits of ρ and σ .

test [51]. Fig. 1 shows the quantum circuit for performing the destructive SWAP test.

By placing a CNOT gate followed by a Hadamard gate on the control qubit, each pair of qubits of ρ and σ are measured in the Bell basis. Following this, each pair of measurement results are post-processed using the vector $\vec{c} = (1, 1, 1, -1)$, which corresponds to summing the probabilities of getting 00, 01 and 10, and subtracting the probability of getting 11. This is equivalent to measuring the expectation value of a CZ operator, resulting in obtaining the overlap $Tr\{\rho\sigma\}$, as long as the qubits are rearranged as $R_1S_1R_2S_2...R_nS_n$, where R_i and S_i denote the subsystems of ρ and σ , respectively. This results in a linear scaling in post-processing, given that we do not directly compute $\vec{c} \cdot \vec{p}$, where \vec{p} is the probability vector, but rather binning the paired measurement outcomes into a +1 and -1 bin, and then averaging the results [34]. The purity of a state can be similarly obtained by supplying two copies of the state as inputs to the test. The one-qubit gates necessary to generate the separable pure states of Eq. (9), coupled with the two-depth circuit needed for the destructive SWAP test, results in a noticeably shallow circuit for the VSV.

C. Variational Optimization

Eqs. (11) and (12) give us the means to compute the HSD as in Eq. (6), and by combining these with Eqs. (2) and (3), we can devise a VQA that is able to compute the CSS and the corresponding HSE.

The optimizer in this scenario is tasked with providing angles $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ to prepare the separable pure states, and probabilities \boldsymbol{p} to classically mix them during post-processing, to generate a separable state $\sigma(\boldsymbol{p},\boldsymbol{\theta},\boldsymbol{\phi})$. One can notice that the quantum computer is only tasked with computing $\text{Tr}\{\rho^2\}$, $\langle \psi(\boldsymbol{\theta}_i,\phi_i)|\rho|\psi(\boldsymbol{\theta}_i,\phi_i)\rangle$, and $|\langle \psi(\boldsymbol{\theta}_i,\phi_i)|\psi(\boldsymbol{\theta}_j,\phi_j)\rangle|^2 \,\,\forall\,\, i,j\in[s]$. The probabilities \boldsymbol{p} are only incorporated classically when computing the final results for calculating the purity $\text{Tr}\{\sigma(\boldsymbol{p},\boldsymbol{\theta},\boldsymbol{\phi})^2\}$, and the overlap $\text{Tr}\{\rho\sigma(\boldsymbol{p},\boldsymbol{\theta},\boldsymbol{\phi})\}$. As a result, we split our optimization routine into a bilevel system, which is essentially a nested optimization routine [35].

The structure of the bilevel system is as follows: at the beginning of every iteration, an upper level optimizer selects the parameters θ and ϕ , and proceeds to call the quantum computer to compute the overlaps in Eqs. (11) and (12). The upper level optimizer then launches a lower level optimizer to obtain the parameters p to minimize Eq. (6). We thus obtain (near-)optimal parameters of the minimization of the cost function (1),

$$\{p^*, \theta^*, \phi^*\} \approx \underset{p,\theta,\phi}{\operatorname{arg min}} D_{\mathrm{HS}}(\rho, \sigma(p, \theta, \phi)),$$
 (13)

such that the CSS of ρ would then be

$$\rho_{\rm CSS} \approx \sigma(\boldsymbol{p}^*, \boldsymbol{\theta}^*, \boldsymbol{\phi}^*),$$
(14)

from which we can calculate the HSE of ρ as

$$E_{\rm HS}(\rho) \approx D_{\rm HS}(\rho, \sigma(\boldsymbol{p}^*, \boldsymbol{\theta}^*, \boldsymbol{\phi}^*)).$$
 (15)

A diagrammatic representation of the VSV algorithm is shown in Fig. 2.

D. Complexity Analysis

The scalability of the VSV, similar to other hybrid quantum-classical algorithms, is characterized from both the classical and quantum point of view. The algorithm prepares the test state ρ (which may be unknown, but reproducible) and trial states σ on a quantum computer, which provides an exponential memory reduction when compared with purely classical algorithms. We also assume ρ can be efficiently prepared on a quantum device: either via direct input, for example, from another quantum system; or else using quantum state preparation methods [52]. Aside from preparing ρ , the algorithm requires n CNOT gates and n Hadamard gates for carrying out the destructive SWAP test, which scales linearly with the number of qubits, and uses a constant circuit depth of two, assuming the quantum computer has a ladder-like connectivity.

In the current implementation of the VSV, the number of parameters scales as s(2n+1) for determining the closest fully separable state of ρ , where s is the number of (non-orthogonal) components representing the trial state σ . Unless ρ is separable, we found that the CSS of ρ being full rank. Thus, we require at least $s \geq d$, although we found that in all of our simulations s = dsuffices to find the CSS of an entangled state. This implies that we have an exponential scaling in the number of parameters needed to determine the CSS via this method. This may be alleviated if one finds a way to directly prepare the CSS as a separable mixed state on the quantum device, which will (generally) require doubling the number of qubits [50]. On the other hand, the number of parameters necessary to define the CSS may be reduced such that it scales sub-exponentially with the number of qubits. It should also be noted that finding the k-CSS would require more complex ansätze, which are more likely to require a larger set of parameters to represent k-separable states.

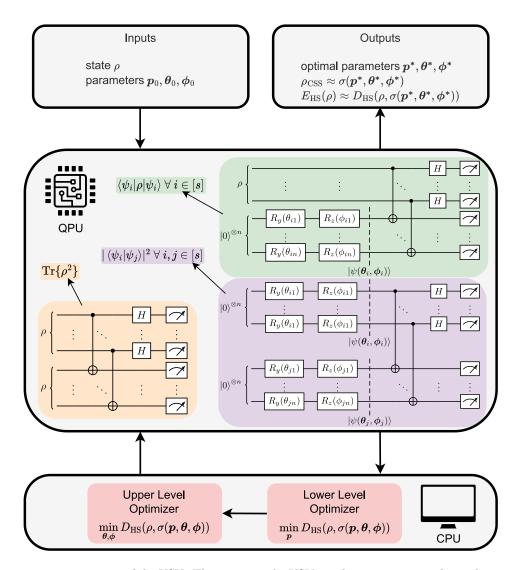


FIG. 2. Diagrammatic representation of the VSV. The inputs to the VSV are the test state ρ , and initial parameters ρ_0 , θ_0 , ϕ_0 . The VSV then computes the orange (bottom-left QPU) circuit once, to compute $\text{Tr}\{\rho^2\}$, with the green (top-right QPU) and purple (bottom-right QPU) circuits evaluated to compute overlaps $\langle \psi_i | \rho | \psi_i \rangle$ and $\langle \psi_i | \psi_j \rangle \forall i, j \in [s]$, which are used in Eqs. (11) and (12), respectively. The red (bottom CPU) section is the classical feedback loop computed by the CPU, consisting of first minimizing the parameters \boldsymbol{p} , followed by proposing new parameters $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ to evaluate on the QPU. After a sufficient number of iterations, the VSV outputs optimal parameters $\boldsymbol{\rho}^*$, $\boldsymbol{\phi}^*$, $\boldsymbol{\phi}^*$, along with the computed CSS and corresponding HSE.

IV. RESULTS

The results involving the VSV are presented in this section for both state vector and shot-based simulations, with all initial points being generated randomly. State vector simulations refer to instances where we have perfect information about the quantum states (that is access to the vector of the state), which is used to benchmark and test the algorithm. On the other hand, shot-based simulations refer to the cases where we sample the state obtained at the end of each circuit to obtain N bit strings, where N is denoted as the number of shots.

A plot of the HSD convergence using state vector optimization for GHZ states, ranging from two to seven qubits, is shown in Fig. 3 — with an inset similarly showing two to five qubits but using shot-based optimization with 8192 shots. The optimizer used in this instance is the generalized simulated annealing (GSA) algorithm [53, 54] from the SciPy library [55], for upper parameter (θ, ϕ) optimization, with the sequential least squares programming (SLSQP) algorithm [56] for the lower parameter (p) optimizer. The GHZ states are specifically chosen so as to demonstrate the performance of the VSV, and also since the CSS is analytically known, as given in Ref. [32].

GHZ states are the basis of numerous quantum communication protocols, such as quantum secret sharing [57], and quantum communication complexity reduction [58], beside being the experimental workhorse for more refined tests of quantum non-locality with respect

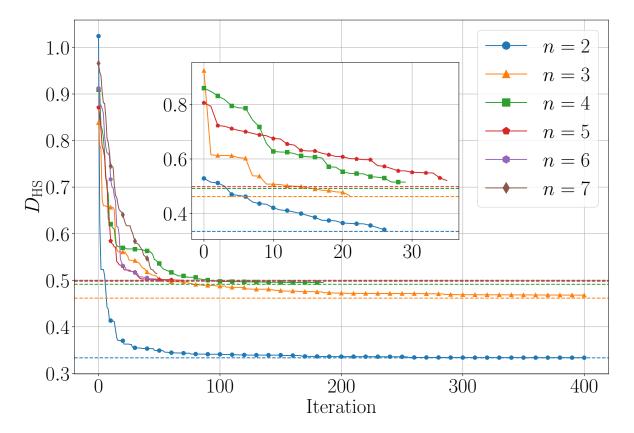


FIG. 3. Plot of the convergence of HSD with respect to the number of iterations for the statevector optimization for GHZ states. The solid lines represent the convergence for two to seven qubits (bottom-up). The dashed lines represent the analytical HSE of the corresponding number of qubits, that is $E_{\rm HS}$ ($|{\rm GHZ}\rangle_n$). The inset figure represents the shot-based optimization convergence with 8192 shots for two to five qubits (bottom-up). Note that there are less iterations for the shot-based cases (although same number of function evaluations), since it is harder for the GSA optimizer to find a strictly lower HSE at each function evaluation.

to two-qubit Bell states [59]. GHZ-diagonal states may inherit their non-classical features from the former [60], since they are typically a consequence of local damping channels acting on individual qubits. Thus, investigating the CSS may provide us with a better understanding of the optimal strategies in the presence of noise, as well as insight into the amount of error that is admissible in these protocols, in order to retain a quantum advantage.

The following application of the VSV is on two- and three-qubit X-MEMS, which are X-states that have the maximal amount of multipartite entanglement for a given linear entropy [21]. These states are specifically chosen since the GME concurrence [5], is known exactly for X-states. The GME concurrence is an extension of the bipartite concurrence, which is a measure related to the entanglement of formation [61, 62], and in fact the GME concurrence reduces to the bipartite concurrence for two-qubit states. The intention is to relate the HSE with the GME concurrence, possibly enabling the HSE to behave as an entanglement monotone for X-MEMS, albeit it does not satisfy the conditions for one in general [29]. Figs. 4 and 5 show the analytical HSE along with the values determined by the VSV, for two- and three-qubit

X-MEMS. In the case of the inset figures, the Nakanishi–Fujii–Todo (NFT) optimizer [63] was used as an upper level optimizer, since it offered significant improvement over GSA when evaluating the overlaps using shots rather than statevector calculations.

By inspection of the CSS determined by the VSV, we were able to surmise the form of the CSS for n-qubit X-MEMS, which is

$$\begin{pmatrix}
\frac{a}{2} & & & & & & d \\
& \frac{b}{N-1} & & & & 0 & \\
& & \ddots & & & \ddots & \\
& & \frac{b}{N-1} & 0 & & & \\
& & 0 & \frac{1-a-b}{N-1} & & & \\
& & \ddots & & & \ddots & \\
& 0 & & & \frac{1-a-b}{N-1} & \\
d^* & & & & \frac{a}{2}
\end{pmatrix}, (16)$$

where $N=2^{n-1},\,a,b\in[0,1],\,a+b\leq 1,\,$ and $|d|\in\left[0,\frac{a}{2}\right],$ to ensure a valid density matrix. The parameters a,b and d can be directly determined by using the Mathematica code on GitHub [64]. We found that for $n\geq 3,\,a$ is

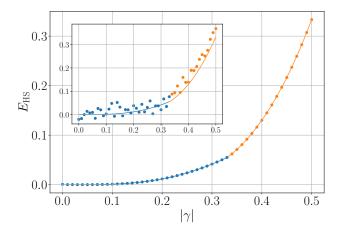


FIG. 4. Plot of the results of the statevector optimization applied to two-qubit X-MEMS. The solid lines represent the analytical HSE as a function of $|\gamma|$. The left blue segment represents the range between $0 \le |\gamma| \le \frac{1}{3}$, while the right orange segment represents the range between $\frac{1}{3} \le |\gamma| \le \frac{1}{2}$. The inset figure represents the data acquired with shot-based optimization using 8192 shots.

a root of a quartic equation including γ , with b and d depending on both a and γ . More details can be found in Appendix A.

Similar to GHZ states and their CSSs (which are also X-states), the X-state itself is robust against local damping channels [65, 66]. This is due to the fact that decoherence effects that introduce additional non-X elements, to a density matrix, cannot decrease the degree of entanglement [21]. As a consequence, the X part of a density matrix provides a lower bound on the GME present in the state [67]. Furthermore, the possibility of an arbitrary initial pure state decohering into an X-state was studied in Ref. [68].

In the case of two-qubit states, the HSE and the concurrence is known to be equal for a restricted set of states [69, 70]. With our VSV, we have numerical evidence supporting the conjecture that the HSE of two distinct, arbitrary, two-qubit states, ρ and σ , is equal if, and only if, the concurrence of ρ and σ is also equal, that is

$$C(\rho) = C(\sigma) \iff E_{\text{HS}}(\rho) = E_{\text{HS}}(\sigma), \quad (17)$$

and similarly, that the HSE of ρ is greater than that of σ if, and only if, the concurrence of ρ is also greater than that of σ , that is

$$C(\rho) > C(\sigma) \iff E_{\rm HS}(\rho) > E_{\rm HS}(\sigma)$$
. (18)

Following this, we tested whether similar statements hold when comparing GME concurrence and HSE for $(n \ge 3)$ -qubit X-states. All of the simulations we carried out pointed towards the null hypothesis in both cases, meaning that equality does not hold:

$$C_{\text{GME}}(\rho) = C_{\text{GME}}(\sigma) \iff E_{\text{HS}}(\rho) = E_{\text{HS}}(\sigma), \quad (19)$$

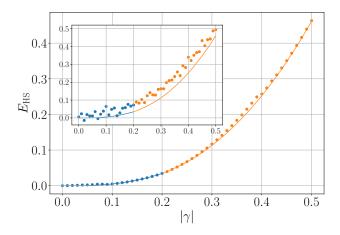


FIG. 5. Plot of the results of the statevector optimization applied to three-qubit X-MEMS. The solid lines represent the analytical HSE as a function of $|\gamma|$. The left blue segment represents the range between $0 \le |\gamma| \le \frac{1}{5}$, while the right orange segment represents the range between $\frac{1}{5} \le |\gamma| \le \frac{1}{2}$. The inset figure represents the data acquired with shot-based optimization using 8192 shots.

nor does ordered inequality:

$$C_{\text{GME}}(\rho) > C_{\text{GME}}(\sigma) \iff E_{\text{HS}}(\rho) > E_{\text{HS}}(\sigma).$$
 (20)

This is not surprising, since the GME concurrence characterizes n-qubit entanglement. On the other hand, the HSE seems to describe the departure of a state from a fully separable one, where in the case of three-qubits, may also consist of biseparable states which are not detected by GME concurrence.

V. CONCLUSION

The variational separability verifier (VSV) is a novel variational quantum algorithm (VQA) that finds the closest separable state (CSS) of arbitrary quantum states, with respect to the Hilbert-Schmidt distance (HSD), and obtain the Hilbert-Schmidt entanglement (HSE). The VSV has been applied to n-qubit Greenberger-Horne–Zeilinger (GHZ) states, showing the convergence of the optimizer. It has also been applied to investigate maximally-entangled mixed X-states (X-MEMS), producing a relation between the genuine multipartite entanglement (GME) concurrence and HSE, as well as helping deduce the analytical form of the CSS for n-qubit X-MEMS. In this respect, the VSV can be useful for shedding light on possible analytical forms of CSSs for arbitrary states. Further simulations carried out on twoqubit states demonstrated that while equality (or ordering) of the concurrence of two states is present if, and only if, equality (or ordering) of the HSE is also present, this does not hold in general for the GME concurrence and HSE in the case of X-states with three or more qubits.

If the calculated HSE of an entangled state is equal

to zero, then we trivially acquire $\rho_{\text{CSS}} = \rho$, which immediately implies that our test state ρ is fully separable, and hence, is not entangled. In either case, if one knows the CSS, then an entanglement witness [71–73] can be defined as $\mathcal{W} \equiv \rho - \rho_{\text{CSS}}$. In general however, one does not exactly acquire the CSS, but rather a close approximation to it, say, σ . In this case, the optimal entanglement witness would be of the form $\mathcal{W} \equiv \rho - \sigma - \max_{|\psi\rangle \in \mathcal{C}} \langle \psi | (\rho - \sigma) | \psi \rangle$ [32].

The strength of the VSV as a potential noisy intermediate-scale quantum (NISQ) algorithm stems from its simplicity. While the algorithm requires in general many calls to the quantum computer for evaluating overlaps, the destructive SWAP test, coupled with the ease of preparing separable states through one-qubit gates, results in a three-depth circuit which is remarkably tractable on a NISQ device. It should be noted that the entire trial state(s) can be directly prepared on the quantum computer by utilizing (up to n) ancillary qubits, which are then traced out after performing the relevant unitary gates [50], leading to less overlap computations for the cost of more qubits, as discussed in Sec. III. It is also noteworthy to mention that the evaluation of state overlaps could be improved by looking towards alternative methods, such as in Refs. [74–78].

Future work would involve enhancing the VSV to find the closest k-separable state (k-CSS) of an arbitrary quantum state, which is not straightforward, since even the closest biseparable state for three-qubit states is significantly challenging to implement. The issue lies in properly expressing a biseparable state in a variational form, since the number of pure state bipartitions scales as $2^{n-1}-1$ for an n-qubit state. The set of all biseparable states is generated by the convex hull of all pure state bipartitions. This, coupled with the fact that an n-qubit state can have rank up to 2^n makes the extension of the VSV to determine the k-CSS extremely non-trivial. The question of finding the k-CSS, along with improvements to the performance of the VSV, are left as open problems.

Apart from the VSV, there are alternative approaches that aim to find the CSS of arbitrary states using other methods, such as neural networks [79] and an adaptive polytope approximation [80]. However, work that inspired the variational approach to the problem at hand is the so-called quantum Gilbert algorithm (QGA) [32, 36–38]. The implementation of the QGA on a quantum device and related discussion can be found in Appendix B.

The Python code for running the VSV simulations, using the package Qulacs [81], and the Mathematica code for determining the CSS of X-MEMS, can be found on GitHub [64].

ACKNOWLEDGMENTS

MC acknowledges funding by the Tertiary Education Scholarships Scheme and the MCST Research Excellence Programme 2022 on the QVAQT project at the Univer-

sity of Malta. TJGA acknowledges funding by the SEA-EU Alliance, European University of the Seas, Research Seed Fund "STATED". MW acknowledges partial support by NCN Grant No. 2017/26/E/ST2/01008 and the Foundation for Polish Science (IRAP project, ICTQT, Contract No. 2018/MAB/5, co-financed by EU within Smart Growth Operational Programme).

Appendix A: Deriving the CSS for X-MEMS

The general form of X-MEMS is given by

$$\tilde{X} = \begin{pmatrix} f(\gamma) & & & & \gamma \\ g(\gamma) & & & & 0 \\ & \ddots & & \ddots & \\ & & g(\gamma) & 0 & \\ & & 0 & 0 & \\ & & \ddots & & \ddots & \\ 0 & & & & 0 \\ \gamma^* & & & & g(\gamma) \end{pmatrix}, \quad (A1)$$

where

$$f(\gamma) = \begin{cases} \frac{1}{N+1} & 0 \le |\gamma| \le \frac{1}{N+1}, \\ |\gamma| & \frac{1}{N+1} \le |\gamma| \le \frac{1}{2}, \end{cases}$$
 (A2a)

$$g(\gamma) = \begin{cases} \frac{1}{N+1} & 0 \le |\gamma| \le \frac{1}{N+1}, \\ \frac{1-2|\gamma|}{N-1} & \frac{1}{N+1} \le |\gamma| \le \frac{1}{2}, \end{cases}$$
 (A2b)

and $N=2^{n-1}$. The GME concurrence was shown to be equal to $2|\gamma|$ for X-MEMS [21].

1. Two-qubit X-MEMS

The analytical form of the CSS for two-qubit X-MEMS was derived using analytical optimization with the Karush–Kuhn–Tucker (KKT) conditions [82], since the numerical results by the VSV hinted towards a CSS of the form

$$\begin{pmatrix}
\frac{a}{2} & 0 & 0 & d \\
0 & b & 0 & 0 \\
0 & 0 & 1 - a - b & 0 \\
d^* & 0 & 0 & \frac{a}{2}
\end{pmatrix},$$
(A3)

where $a, b \in [0, 1]$, $a + b \le 1$, and $|d| \in [0, \frac{a}{2}]$, to ensure a valid density matrix. For this state to be separable, then the concurrence must be equal to zero, meaning that

$$|d|^2 < b(1 - a - b). \tag{A4}$$

Optimizing this problem results in parameters

$$a = \frac{1}{9} \left(7 - \sqrt{1 + 36|\gamma|^2} \right),$$
 (A5a)

$$b = \frac{1 + 12|\gamma|^2 + \sqrt{1 + 36|\gamma|^2}}{6\sqrt{1 + 36|\gamma|^2}},$$
 (A5b)

$$d = \frac{\gamma}{3} \left(1 + \frac{2}{\sqrt{1 + 36|\gamma|^2}} \right),$$
 (A5c)

with an HSE of

$$E_{\rm HS}\left(\tilde{X}\right) = \frac{2}{27} \left(1 + 18|\gamma|^2 - \sqrt{1 + 36|\gamma|^2}\right), \quad (A6)$$

for $0 \le |\gamma| \le \frac{1}{3}$, and

$$a = \frac{1}{3} \left(1 + 4|\gamma| - \sqrt{1 - 4|\gamma| + 8|\gamma|^2} \right),$$
 (A7a)

$$b = \frac{1}{6} \left(3 - 6|\gamma| + \frac{3 - 12|\gamma| + 16|\gamma|^2}{\sqrt{1 - 4|\gamma| + 8|\gamma|^2}} \right), \quad (A7b)$$

$$d = \frac{\gamma \left(2 - 4|\gamma| + \sqrt{1 - 4|\gamma| + 8|\gamma|^2}\right)}{3\sqrt{1 - 4|\gamma| + 8|\gamma|^2}},$$
 (A7c)

with an HSE of

$$E_{\rm HS}\left(\tilde{X}\right) = \frac{2}{3} \left(1 - 4|\gamma| + 6|\gamma|^2 + (2|\gamma| - 1)\sqrt{1 - 4|\gamma| + 8|\gamma|^2}\right), \quad (A7d)$$

for $\frac{1}{3} \le |\gamma| \le \frac{1}{2}$.

2. Three-qubit X-MEMS

The analytical form of the CSS for three-qubit X-MEMS was derived using analytical optimization with the KKT conditions [82], since the numerical results by the VSV hinted towards a CSS of the form

$$\begin{pmatrix}
\frac{a}{2} & 0 & 0 & 0 & 0 & 0 & 0 & d \\
0 & \frac{b}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{b}{3} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{b}{3} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1-a-b}{3} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1-a-b}{3} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1-a-b}{3} & 0 \\
d^* & 0 & 0 & 0 & 0 & 0 & 0 & \frac{a}{2}
\end{pmatrix},$$
(A8)

where similarly, $a, b \in [0, 1]$, $a+b \le 1$, and $|d| \in \left[0, \frac{a}{2}\right]$, to ensure a valid density matrix. Since we require the state to be separable, we first need to certify that the GME concurrence is zero, which results in the same condition as in (A4). However, the GME concurrence only assures us that the state does not have tripartite entanglement in this scenario, yet we still need to ensure that the state has no bipartite entanglement. It can be immediately

seen from the density matrix (A8) that the concurrence between pairs of individual qubits is zero — since the partial trace with respect to each qubit results in a diagonal matrix.

What is left to check is the entanglement between each qubit and the rest, and so, we look to the negativity [83], although the negativity does not capture bound entangled states [71, 84, 85]. However, we conjecture that the state in (A8) is not bound entangled, since the numerical HSE, determined by the VSV, coincides with the HSE of (A8), as shown in Fig. 5. Due to the uniqueness of the CSS [32], this implies that the fully separable state determined by the VSV, is exactly the same state as (A8), meaning that it is fully separable.

The partial transpose of (A8) with respect to each qubit is equivalent, and the negativity is non-zero if the partial transpose has negative eigenvalues. Now, the partial transpose of (A8) only has one eigenvalue which can be negative, for some parameters a, b and d, which is

$$\frac{1}{6} \left(1 - a - \sqrt{(1 - a - 2b)^2 + 36|d|^2} \right), \quad (A9)$$

meaning we need to ascertain that the above condition is non-negative, for (A8) to be separable. This condition can instead be rewritten and shown that it is equivalent to

$$|d|^2 \le \frac{b(1-a-b)}{9},$$
 (A10)

which interestingly encompasses condition (A4) as well, resulting in an analytical optimization problem using the KKT conditions with only constraint (A10) to insure separability. For the sake of brevity, we shall not provide the parameters a, b and d, as well as the HSE here, since they are in terms of a root of a quartic equation and so the explicit equations would be too cumbersome to present. Nevertheless, they can be reproduced using the Mathematica code available on GitHub [64].

3. n-qubit X-MEMS

While we have not explicitly derived the CSS of n-qubit X-MEMS in general, we conjecture that it is of the form

$$\begin{pmatrix} \frac{a}{2} & & & & & & d \\ & \frac{b}{N-1} & & & & 0 & & \\ & & \ddots & & & \ddots & & \\ & & \frac{b}{N-1} & 0 & & & \\ & & 0 & \frac{1-a-b}{N-1} & & & \\ & & \ddots & & & \ddots & \\ & 0 & & & \frac{1-a-b}{N-1} & \frac{a}{2} \end{pmatrix}, (A11)$$

where $N=2^{n-1}$, $a,b\in[0,1]$, $a+b\leq 1$, and $|d|\in[0,\frac{a}{2}]$, to ensure a valid density matrix. This form is also

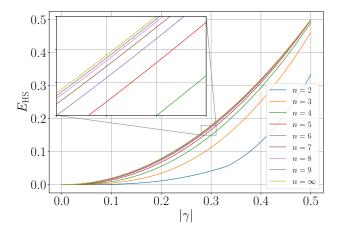


FIG. 6. Analytical HSE of two- to nine-qubit X-MEMS as a function of $|\gamma|$ (bottom-up). The infinite-qubit (top-most line) case represents the analytical upper bound of $2|\gamma|^2$ for the HSE of X-MEMS.

invariant under the symmetries imposed by X-MEMS, which is a necessary condition for the CSS as shown in Ref. [32]. The symmetries of the CSS consist of any qubit permutations from the set $\{2,\ldots,n\}$, and the condition that the all-zero element must be equal to the all-one element. The parameter a in the CSS seems to in general be a root of a quartic equation, with the parameters b and d depending on both a and γ . The condition for zero negativity also encompasses the cases for $(k \geq 3)$ -separability, which is equivalent to

$$|d|^2 \le \frac{b(1-a-b)}{(N-1)^2},$$
 (A12)

presumably related to the hierarchy of mixed state entanglement [5].

Fig. 6 shows the analytical HSE of two- to nine-qubit X-MEMS as a function of $|\gamma|$. Each function for the HSE is piece-wise continuous, with the domains divided at the point $|\gamma| = \frac{1}{N+1}$, as is similarly given by the definition of the X-MEMS in Eqs. (A2). The CSSs possess an HSE of zero at $|\gamma| = 0$ and an HSE of

$$E_{\text{HS}}\left(|\text{GHZ}\rangle_{n}\right) = \frac{2^{n} - 2}{2^{n+1} + 2^{3-n} - 4},$$
 (A13)

at $|\gamma|=\frac{1}{2}$, since the *n*-qubit *X*-MEMS correspond to the *n*-qubit GHZ states at that point [32]. It is also interesting to note the limit of the functions in Fig. 6, corresponding to the infinite-qubit *X*-MEMS, which is equal to

$$\lim_{n \to \infty} E_{\rm HS} \left(\tilde{X}_n \right) = 2|\gamma|^2, \tag{A14}$$

meaning that the HSE for X-MEMS is bounded above.

Appendix B: Comparison with the Quantum Gilbert Algorithm

The QGA is the quantum analogue of Gilbert's algorithm [86], which is utilized to determine the distance between a given point and a convex set, according to the Hilbert-Schmidt norm. The QGA has insofar only been implemented on classical computers to test for membership of states in different stochastic local operations and classical communication (SLOCC) classes [32, 36–38].

It is only natural to attempt to extend the QGA to operate on a quantum computer, owing to many reasons. The first of which is the storage of quantum mixed state data, which is known to increase exponentially as 4^n , with respect to the number of qubits n. The second is the input of arbitrary mixed states into a quantum computer, created via open quantum systems as an example, where these reasons also apply to the VQA counterpart. The issue in simply extending the QGA to be utilized on a quantum computer, in order to improve the performance over its classical counterpart, is due to the nature of the algorithm itself, requiring possibly, an infinite number of state mixtures to the trial state. The QGA relies on computing overlaps of different trial states, similar to the VSV, however, the main difference is that the VSV keeps a constant number of states in memory. On the other hand, the QGA, at each successful iteration of the algorithm adds one state to memory, increasing the number of future overlap calculations. This means that more calls to the quantum computer are required for subsequent iterations of the algorithm. In general if we suppose we require one million trial states for a three-qubit state to acquire a reasonable CSS, which is exactly the amount of states taken by the QGA, as given in Ref. [32], and we assume that the number of successes, c_s , is proportional to the number of trial states, c_t , by $c_s \propto c_t^{0.44127}$, then the number of overlaps is equal to

$$\sum_{i=1}^{10^6} \sum_{j=1}^{c_s^{(i)}} j = \sum_{i=1}^{10^6} \frac{c_s^{(i)} \left(c_s^{(i)} + 1\right)}{2},$$
 (B1)

where $c_s^{(i)}$ is the number of successes at iteration i, which starts at 1 for i=1, and increases by 1 every $\frac{c_t}{c_s} \approx \frac{c_t}{c_t^{0.44127}} = c_t^{0.55873}$ times, which for one million trial states is approximately every 2251 trial states. However, in the beginning, almost every trial state will correspond to a success, due to the initial state on average lying farther away from the actual CSS. As the number of successes increases, and the state moves closer to the border of the set of separable states, nearer to the CSS, it will be harder to find a closer separable state. But, even if we take it to be split evenly, then this results to be around 3.3×10^{10} calls to the quantum computer. On the other hand, the VSV calls the quantum computer 5×10^3 times to obtain a CSS comparable to the QGA, which would equate to more than 6×10^6 calls for the QGA for each call of the VSV.

In either case, we shall describe the direct implementation of the QGA on a quantum device. Given an initial guess in the form of a pure product state $\sigma_0 = |\psi_0\rangle\langle\psi_0|$, after the n^{th} success, the new CSS ρ_n to the test state ρ is given iteratively as

$$\rho_n = p_n \rho_{n-1} + (1 - p_n) \sigma_n, \tag{B2}$$

which equates to

$$\rho_n = \sum_{i=0}^n \left(\prod_{j=i+1}^n p_j \right) (1 - p_i) \, \sigma_i,$$
(B3)

where $\sigma_i = |\psi_i\rangle\langle\psi_i|$ is a pure product state and $p_i \in (0,1)$, such that

$$\sum_{i=0}^{n} \left(\prod_{j=i+1}^{n} p_j \right) (1 - p_i) = 1.$$
 (B4)

Note that this definition entails that $\rho_0 \equiv \sigma_0$. The implementation of a QGA can be thus be described as follows [32]:

- (a) Input data: the state to be tested ρ and any pure product state σ_0 .
- (b) Output data: the closest state found after n successes ρ_n , and lists of values of $D_{\rm HS}(\rho, \rho_i) = \text{Tr}\{(\rho \rho_i)^2\}$, p_i and σ_i .
- 0. Calculate the value of $D_{\rm HS}(\rho, \sigma_0) \equiv D_{\rm HS}(\rho, \rho_0)$ and add it to the list.
- 1. Increase the counter of trials c_t by 1. Draw a random pure product state σ_1 , hereafter called the first trial state.
- 2. Run a preselection for the trial state by checking a value of a linear functional. If it fails, go back to step 1.
- 3. In the case of successful preselection, find the minimum of $\text{Tr}\{(\rho-\rho_n)^2\}$ with respect to p_n , which in the case of the first trial state is $\text{Tr}\{(\rho-p_1\rho_0-(1-p_1)\sigma_1)^2\}$ with respect to p_1 .
- 4. If the minimum occurs for $p_1 \in [0,1]$, then $\rho_1 = p_1\rho_0 + (1-p_1)\sigma_1$ is saved to the list as the new CSS, along with the value of $D_{\text{HS}}(\rho, \rho_1)$, p_1 and σ_1 , increase the success c_s counter value by 1.
- 5. Go to step 1., now drawing a new random pure product state σ_2 , such that after a successful preselection, one determines a new p_2 , and consequently ρ_2 , and then saves $D_{\text{HS}}(\rho, \rho_2)$, p_2 and σ_2 . Continue in this fashion until a specified HALT criterion is met.

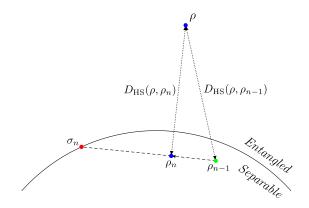


FIG. 7. Visualization of the $n^{\rm th}$ iteration of the QGA. A random pure product state σ_n is generated that satisfies the preselection criterion. The next step is to find $\rho_n = p_n \rho_{n-1} + (1-p_n)\sigma_n$, $p_n \in (0,1)$, such that $D_{\rm HS}\left(\rho,\rho_n\right) < D_{\rm HS}\left(\rho,\rho_{n-1}\right)$. Figure adapted from Ref. [32].

Fig. 7 provides a visual representation of the n^{th} iteration of the algorithm presented above, while the preselection criterion for the n^{th} trial state is given by

$$Tr\{(\sigma_n - \rho_{n-1})(\rho - \rho_{n-1})\} > 0,$$
 (B5)

with the geometrical interpretation being that the angle between the vectors $\sigma_n - \rho_{n-1}$ and $\rho - \rho_{n-1}$ is not larger than $\frac{\pi}{2}$ [32], ensuring that a state between ρ_{n-1} and σ_n exists such that it is closer to ρ . Eq. (B5) can rewritten and computed as

$$\operatorname{Tr}\{\rho_{n-1}^2\} + \operatorname{Tr}\{\rho\sigma_n\} > \operatorname{Tr}\{\rho\rho_{n-1}\} + \operatorname{Tr}\{\rho_{n-1}\sigma_n\}.$$
 (B6)

The purity of each new CSS when calculated iteratively (starting at 1 due to the initial trial state σ_0 being a pure state), is given by

$$\operatorname{Tr}\{\rho_n^2\} = \operatorname{Tr}\{(p_n \rho_{n-1} + (1 - p_n)\sigma_n)^2\}$$

$$= p_n^2 \operatorname{Tr}\{\rho_{n-1}^2\} + (1 - p_n)^2$$

$$+ 2p(1 - p_n) \operatorname{Tr}\{\rho_{n-1}\sigma_n\},$$
 (B7)

while the overlap is calculated as

$$\operatorname{Tr}\{\rho\rho_n\} = p_n \operatorname{Tr}\{\rho\rho_{n-1}\} + (1 - p_n) \operatorname{Tr}\{\rho\sigma_n\}.$$
 (B8)

Thus, the cost function for the $n^{\rm th}$ iteration is defined as follows:

$$D_{\text{HS}}(\rho, \rho_n) = \text{Tr}\{(\rho - p_n \rho_{n-1} - (1 - p_n)\sigma_n)^2\}$$

$$= \text{Tr}\{\rho^2\} + p_n^2 \text{Tr}\{\rho_{n-1}^2\} + (1 - p_n)^2$$

$$+ 2p_n(1 - p_n) \text{Tr}\{\rho_{n-1}\sigma_n\}$$

$$- 2p_n \text{Tr}\{\rho\rho_{n-1}\} - 2(1 - p_n) \text{Tr}\{\rho\sigma_n\},$$
(B9)

which is explicitly written and computed as

$$D_{\text{HS}}(\rho, \rho_n) = D_{\text{HS}}\left(\rho, \sum_{i=0}^n \left(\prod_{j=i+1}^n p_j\right) (1 - p_i) \,\sigma_i\right)$$

$$= \text{Tr}\left\{\left(\rho - \sum_{i=0}^n \left(\prod_{j=i+1}^n p_j\right) (1 - p_i) \,\sigma_i\right)^2\right\}$$

$$= \text{Tr}\left\{\rho^2\right\} + \text{Tr}\left\{\left(\sum_{i=0}^n \left(\prod_{j=i+1}^n p_j\right) (1 - p_i) \,\sigma_i\right)^2\right\} - 2 \,\text{Tr}\left\{\rho \sum_{i=0}^n \left(\prod_{j=i+1}^n p_j\right) (1 - p_i) \,\sigma_i\right\}$$

$$= \text{Tr}\left\{\rho^2\right\} + 2 \sum_{i < k}^n \left(\prod_{j=i+1}^n p_j\right) \left(\prod_{l=k+1}^n p_l\right) (1 - p_i) (1 - p_k) \,\text{Tr}\left\{\sigma_i \sigma_k\right\} +$$

$$\sum_{i=0}^n \left(\prod_{j=i+1}^n p_j^2\right) (1 - p_i)^2 - 2 \sum_{i=0}^n \left(\prod_{j=i+1}^n p_j\right) (1 - p_i) \,\text{Tr}\left\{\rho \sigma_i\right\}. \tag{B10}$$

This entails that the overlap between every new trial state σ_n and every previous trial state $\sigma_0, \ldots, \sigma_{n-1}$, as well as with the test state ρ , must be calculated at every iteration. This implies that the parameters of each

separable pure trial state must be saved at each success, resulting in an ever-increasing memory and quantum device utilization.

- [1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. 47, 777 (1935).
- [2] J. S. Bell, Physics Physique 1, 195 (1964).
- [3] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. 78, 2275 (1997).
- [4] M. Plenio and S. Virmani, Proc. Spie. 7, 1 (2007).
- [5] C. Eltschka and J. Siewert, J. Phys. A: Math. Theor. 47, 424005 (2014).
- [6] V. Coffman, J. Kundu, and W. K. Wootters, Phys. Rev. A 61, 052306 (2000).
- [7] M. M. Cunha, A. Fonseca, and E. O. Silva, Universe 5, 209 (2019).
- [8] O. Gühne and M. Seevinck, New J. Phys. 12, 053002 (2010).
- [9] W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A 62, 062314 (2000).
- [10] B. Regula, S. Di Martino, S. Lee, and G. Adesso, Phys. Rev. Lett. 113, 1 (2014).
- [11] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, Phys. Rev. A 65, 052112 (2002).
- [12] G. Gour and N. R. Wallach, J. Math. Phys. 51, 112201 (2010).
- [13] M. G. Ghahi and S. J. Akhtarshenas, Eur. Phys. J. D 70, 1 (2016).
- [14] A. Osterloh, Phys. Rev. A 94, 1 (2016).
- [15] Y. Guo and L. Zhang, Phys. Rev. A 101, 032301 (2020).
- [16] J. Eisert and H. J. Briegel, Phys. Rev. A **64**, 022306 (2001)
- [17] P. J. Love, A. M. v. d. Brink, A. Y. Smirnov, M. H. S. Amin, M. Grajcar, E. Il'ichev, A. Izmalkov, and A. M.

- Zagoskin, A characterization of global entanglement (2006)
- [18] M. Walter, D. Gross, and J. Eisert, Multipartite Entanglement (2016).
- [19] C. Eltschka, A. Osterloh, J. Siewert, and A. Uhlmann, New J. Phys. 10, 043014 (2008).
- [20] C. Eltschka and J. Siewert, Phys. Rev. Lett. 108, 020502 (2012).
- [21] S. Agarwal and S. M. Hashemi Rafsanjani, Int. J. Quantum Inf. 11, 1350043 (2013).
- [22] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. 70, 1895 (1993).
- [23] M. Popp, F. Verstraete, M. A. Martín-Delgado, and J. I. Cirac, Phys. Rev. A 71, 042306 (2005).
- [24] M. Consiglio, L. Z. Mangion, and T. J. G. Apollaro, Int. J. Quantum Inf. 19, 10.1142/s0219749921500246 (2021).
- [25] L. Gurvits, J. Comput. Syst. Sci. 69, 448 (2004), special Issue on STOC 2003.
- [26] L. Ioannou, Proc. Spie. 7, 336 (2007), arXiv:0603199 [quant-ph].
- [27] S. Gharibian, Proc. Spie. 10, 343 (2010), arXiv:0810.4507.
- [28] J. Preskill, Quantum 2, 79 (2018).
- [29] I. Bengtsson and K. Zyczkowski, Geometry of Quantum States (Cambridge University Press, 2006).
- [30] C. Witte and M. Trucks, Phys. Lett. A 257, 14 (1999).
- [31] M. Ozawa, Phys. Lett. A 268, 158 (2000).
- [32] P. Pandya, O. Sakarya, and M. Wieśniak, Phys. Rev. A 102, 012409 (2020).

- [33] J. C. Garcia-Escartin and P. Chamorro-Posada, Phys. Rev. A 87, 052330 (2013).
- [34] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, New J. Phys. 20, 113022 (2018).
- [35] S. Dempe, Foundations of bilevel programming, Nonconvex optimization and its applications, Vol. 61 (Kluwer Academic Publishers, Dordrecht, 2002).
- [36] S. Brierley, M. Navascues, and T. Vertesi, Convex separation from convex optimization for large-scale problems (2016).
- [37] J. Shang and O. Gühne, Phys. Rev. Lett. 120, 050506 (2018).
- [38] M. Wieśniak, P. Pandya, O. Sakarya, and B. Woloncewicz, Quantum Reports 2, 49 (2020).
- [39] R. A. Bertlmann, H. Narnhofer, and W. Thirring, Phys. Rev. A 66, 032319 (2002).
- [40] S. L. L. Silva and D. H. T. Franco, Quantum Studies: Mathematics and Foundations 9, 219 (2022).
- [41] A. Streltsov, H. Kampermann, and D. Bruß, New J. Phys. 12, 123004 (2010).
- [42] A. Uhlmann, Rep. Math. Phys. 9, 273 (1976).
- [43] T. A. Brun, Measuring polynomial functions of states (2004).
- [44] K. Bartkiewicz, K. Lemr, and A. Miranowicz, Phys. Rev. A 88, 052104 (2013).
- [45] M. Cerezo, A. Poremba, L. Cincio, and P. J. Coles, Quantum 4, 248 (2020).
- [46] N. Ezzell, Z. Holmes, and P. J. Coles, The quantum low-rank approximation problem (2022).
- [47] N. Ezzell, E. M. Ball, A. U. Siddiqui, M. M. Wilde, A. T. Sornborger, P. J. Coles, and Z. Holmes, Quantum Mixed State Compiling (2022).
- [48] P. Horodecki, Phys. Lett. A 232, 333 (1997).
- [49] V. Vedral and M. B. Plenio, Phys. Rev. A 57, 1619 (1998).
- [50] G. Benenti and G. Strini, Phys. Rev. A 79, 052301 (2009).
- [51] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. 87, 167902 (2001).
- [52] I. F. Araujo, D. K. Park, F. Petruccione, and A. J. da Silva, Sci. Rep. 11, 1 (2021), arXiv:2008.01511.
- [53] C. Tsallis and D. A. Stariolo, Physica A 233, 395 (1996).
- [54] Y. Xiang, D. Sun, W. Fan, and X. Gong, Phys. Lett. A 233, 216 (1997).
- [55] SciPy, Dual Annealing, https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.dual_annealing.html (2022).
- [56] D. Kraft, A Software Package for Sequential Quadratic Programming, Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt Köln: Forschungsbericht (Wiss. Berichtswesen d. DFVLR, 1988).
- [57] M. Hillery, V. Bužek, and A. Berthiaume, Phys. Rev. A 59, 1829 (1999).
- [58] J. Ho, G. Moreno, S. Brito, F. Graffitti, C. L. Morrison, R. Nery, A. Pickston, M. Proietti, R. Rabelo, A. Fedrizzi, and R. Chaves, npj Quantum Inf. 8, 13 (2022).
- [59] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter,

- and A. Zeilinger, Nature 403, 515 (2000).
- [60] X.-y. Chen, L.-z. Jiang, P. Yu, and M. Tian, Entanglement and genuine entanglement of three qubit GHZ diagonal states (2012).
- [61] S. Hill and W. K. Wootters, Phys. Rev. Lett. 78, 5022 (1997).
- [62] W. K. Wootters, Phys. Rev. Lett. 80, 2245 (1998).
- [63] K. M. Nakanishi, K. Fujii, and S. Todo, Phys. Rev. Research 2, 043158 (2020).
- [64] M. Consiglio, https://github.com/mirkoconsiglio/ VSV (2022).
- [65] T. Yu and J. Eberly, Proc. Spie. 7, 459 (2007).
- [66] S. M. Hashemi Rafsanjani, M. Huber, C. J. Broadbent, and J. H. Eberly, Phys. Rev. A 86, 062303 (2012).
- [67] Z.-H. Ma, Z.-H. Chen, J.-L. Chen, C. Spengler, A. Gabriel, and M. Huber, Phys. Rev. A 83, 062325 (2011).
- [68] N. Quesada, A. Al-Qasimi, and D. F. James, J. Mod. Optic. 59, 1322 (2012), https://doi.org/10.1080/09500340.2012.713130.
- [69] D. Mundarain and J. Stephany, Concurrence and negativity as distances (2007).
- [70] R. Ganardi, M. Miller, T. Paterek, and M. Żukowski, Quantum 6, 654 (2022).
- [71] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A 223, 1 (1996).
- [72] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. A 69, 022308 (2004).
- [73] O. Gühne and G. Tóth, Phys. Rep. 474, 1 (2009).
- [74] S. T. Flammia and Y.-K. Liu, Phys. Rev. Lett. 106, 230501 (2011).
- [75] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller, Phys. Rev. A 99, 052323 (2019).
- [76] A. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller, Phys. Rev. Lett. 124, 010504 (2020).
- [77] M. Fanizza, M. Rosati, M. Skotiniotis, J. Calsamiglia, and V. Giovannetti, Phys. Rev. Lett. 124, 060503 (2020).
- [78] W. Liu and H.-W. Yin, Physica A 606, 128117 (2022).
- [79] A. Girardin, N. Brunner, and T. Kriváchy, Phys. Rev. Research 4, 023238 (2022).
- [80] T.-A. Ohst, X.-D. Yu, O. Gühne, and H. C. Nguyen, Certifying Quantum Separability with Adaptive Polytopes (2022).
- [81] Y. Suzuki, Y. Kawase, Y. Masumura, Y. Hiraga, M. Nakadai, J. Chen, K. M. Nakanishi, K. Mitarai, R. Imai, S. Tamiya, T. Yamamoto, T. Yan, T. Kawakubo, Y. O. Nakagawa, Y. Ibe, Y. Zhang, H. Yamashita, H. Yoshimura, A. Hayashi, and K. Fujii, Quantum 5, 559 (2021).
- [82] D. Tabak and B. C. Kuo (1971).
- [83] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A 58, 883 (1998).
- [84] A. Peres, Phys. Rev. Lett. 77, 1413 (1996).
- [85] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. 80, 5239 (1998).
- [86] E. G. Gilbert, Siam J. Control 4, 61 (1966).