2014

Bounding Polynomial Entanglement Measures for Mixed States

Peter John Love
Haverford College

Samuel Rodriques '13
Class of 2013, Haverford College

Bounding polynomial entanglement measures for mixed states

Samuel Rodriques,[1,2] Nilanjana Datta,[3,*] and Peter Love[2,†]
[1]*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*
[2]*Department of Physics, Haverford College, Haverford, Pennsylvania 19041, USA*
[3]*Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, United Kingdom*
(Received 25 July 2013; revised manuscript received 15 January 2014; published 30 July 2014)

We generalize the notion of the best separable approximation (BSA) and best W-class approximation (BWA) to arbitrary pure-state entanglement measures, defining the best zero-E approximation (BEA). We show that for any polynomial entanglement measure E, any mixed state ρ admits at least one "S decomposition," i.e., a decomposition in terms of a mixed state on which E is equal to zero, and a single additional pure state with (possibly) nonzero E. We show that the BEA is not, in general, the optimal S decomposition from the point of view of bounding the entanglement of ρ and describe an algorithm to construct the entanglement-minimizing S decomposition for ρ and place an upper bound on $E(\rho)$. When applied to the three-tangle, the cost of the algorithm is linear in the rank d of the density matrix and has an accuracy comparable to a steepest-descent algorithm whose cost scales as $d^8 \log d$. We compare the upper bound to a lower-bound algorithm given by C. Eltschka and J. Siewert [Phys. Rev. Lett. **108**, 020502 (2012)] for the three-tangle and find that on random rank-2 three-qubit density matrices, the difference between the upper and lower bounds is 0.14 on average. We also find that the three-tangle of random full-rank three-qubit density matrices is less than 0.023 on average.

DOI: 10.1103/PhysRevA.90.012340 PACS number(s): 03.67.Mn, 03.65.Ud, 03.67.Bg

I. INTRODUCTION

Nonclassical correlations in quantum states such as entanglement distinguish quantum from classical information theory. The ability to calculate entanglement of mixed quantum states is relevant for the analysis of tomography data for systems of multiple qubits in several implementations [1–3]. Multipartite systems can contain multiple inequivalent types of entanglement that cannot be converted into one another by local operations and classical communication [4].

One approach to characterizing pure-state entanglement in a system of qubits associates a polynomial function that is invariant under determinant-1 local operations with each type of entanglement [5–7]. Examples of such polynomial invariants include the concurrence for two qubits [8] and the three-tangle, which quantifies the amount of entanglement in a three-qubit system that cannot be accounted for by entanglement between pairs of the qubits [9].

A polynomial-invariant E is extended to mixed states by way of the convex roof, given for a rank-d density matrix ρ by

$$E(\rho) = \min_{\mathscr{E} \in \Upsilon_\rho} \sum_i p_i E(\psi_i), \tag{1}$$

where $\mathscr{E} = \{p_i, |\psi_i\rangle\}$ is a pure-state ensemble for ρ and Υ_ρ is the set of all such ensembles. Carathéodory's theorem allows us to restrict the optimization to ensembles containing no more than d^2 elements [10].

An ensemble that minimizes Eq. (1) is said to be *minimal*. We consider the rank d of the density matrix d, rather than the dimension of the Hilbert space on which it acts, because d is the parameter that determines the computational difficulty of the convex roof minimization. A number of special cases of computation of the convex roof have been solved for cases of restricted rank [11–13].

Minimal ensembles have been found analytically for the concurrence of arbitrary two-qubit mixed states [8] and for the three-tangle of rank-2 mixtures of generalized Greenberger-Horne-Zeilinger (GHZ) and generalized W states [12,13], as well as on rank-3 mixtures of a GHZ state, a W state, and a state obtained by flipping all three bits of a W state [14]. When the minimal ensemble is not known analytically, which is the typical case, one may evaluate an upper bound on $E(\rho)$ using, for example, a steepest-descent algorithm [15]. However, the cost of such an upper bound scales like $d^8 \log d$, making calculations infeasible for high rank.

An alternative approach to characterizing the entanglement of a mixed state was given by Lewenstein and Sanpera [16]. Given a two-qubit density matrix ρ, they considered the set S of pure states $\{\psi_i\}$, such that for all $\psi_i \in S$, there exists some $p_i \in (0,1)$ and a separable mixed state π_i such that

$$\rho = p_i \psi_i + (1 - p_i)\pi_i. \tag{2}$$

We refer to any decomposition of a state into a pure state ψ_i and a state π_i such that $E(\pi_i) = 0$, for given polynomial invariant E, as an S decomposition. For the concurrence on two-qubit states, Lewenstein and Sanpera showed that S is nonempty and then considered the S decompositions obtained by finding the element $\psi_e \in S$ that minimizes the corresponding probability p_e [16]. The corresponding separable state π_e is the "best separable approximation" (BSA) of ρ. Their algorithm for finding the BSA of a mixed state ρ determines whether ρ is separable and provides an upper bound on the entanglement of ρ because $E(\rho) \leqslant p_e E(\psi_e)$ for all convex roof entanglement monotones E.

More generally, it has been shown that every bipartite state ρ has a unique convex decomposition of the form $\rho = \lambda\rho_s + (1 - \lambda)\omega$, where ρ_s is a separable state and the parameter $\lambda \in$

*n.datta@statslab.cam.ac.uk
†plove@haverford.edu

$[0,1]$ is maximal [16,17]. State ρ_s is referred to as the BSA of ρ, and λ is referred to as its *separability* [18]. Obviously, for any two-qubit entangled state, $\lambda < 1$. Moreover, it was established that for any two-qubit mixed state ρ, the separability is nonzero (and hence the BSA exists), and ω is a pure state. However, for states of more than two qubits, the separability may be zero, and ω may be mixed. For a two-qubit state, the BSA places an upper bound on the concurrence $C(\rho)$ of ρ because, clearly,

$$C(\rho) \leqslant (1-\lambda)C(\omega). \tag{3}$$

All entanglement measures are zero on separable states. However, there are other interesting classes of states, such as the W class, on which some, but not all, measures are zero. Acin *et al.* extended the above approach to the three-tangle of three-qubit states, defining the "best W approximation" (BWA) [19].

In this paper, we generalize the BSA and BWA to arbitrary polynomial invariants, defining the best zero-E approximation (BEA) of a mixed state ρ. We show that the BEA exists for all ρ (i.e., S is nonempty for all ρ) and that it is unique. For the BSA, BWA, and BEA, the probability of the single entangled state in the pure-state ensemble for ρ is minimized. This does not mean that the BSA, BEA, or BWA gives the best upper bound on the entanglement for an ensemble of this form. One can obtain an improved upper bound on the entanglement of ρ by finding the element ψ_l of S such that $p_l E(\psi_l)$ is minimized, where p_l is the probability with which ψ_l occurs in a convex decomposition of ρ.

We describe an algorithm that finds ψ_l for any mixed state ρ and so places an upper bound on $E(\rho)$. Applied to the three-tangle, the cost of finding this upper bound scales linearly in d and terminates after 10 s on random three-qubit density matrices (using Intel Core 2 CPUs at 2.66 GHz), as opposed to the $d^8 \log d$ scaling and 10-day runtime expected (see Table I) for the steepest-descent algorithm given in [15]. We evaluate the accuracy of this algorithm by comparing this upper bound to the analytical value of the three-tangle for states on which it is known. By comparing the upper-bound and steepest-descent methods on random states we demonstrate that the two algorithms exhibit comparable accuracies on states for which no analytical value is known.

In all that follows, $\mathcal{H}$ denotes a Hilbert space of some number of qubits, $\mathscr{D}(\mathcal{H})$ denotes the set of density matrices (states) acting on $\mathcal{H}$, and $E : \mathscr{D}(\mathcal{H}) \mapsto \mathbb{R}$ is assumed to be the convex roof extension of a polynomial function of pure states that is of homogeneous degree in the expansion coefficients of pure states written relative to the computational basis and that is invariant under determinant-1 local operations. $D(\rho,\pi) = \|\rho - \pi\|_1$ is the trace distance, $\mathrm{supp}(\rho)$ is the support of ρ, and $R(\rho)$ is its range. For any pure state $|\psi\rangle \in \mathcal{H}$, we denote the projector $|\psi\rangle\langle\psi|$ simply as ψ.

TABLE I. Average run times (in seconds) for the algorithm presented in the text (UB) and the steepest-descent algorithm (SD) on 240 uniformly sampled three-qubit density matrices for ranks 2 through 8. The steepest descent has only been calculated up to rank 5 due to the rapid growth in the run time. The average difference in entanglements is shown in the third row (UB-SD).

	Rank						
	2	3	4	5	6	7	8
UB	0.94	1.89	3.07	4.2	5.5	7.0	7.5
SD	0.17	1.39	12.39	125.2			
UB-SD	0.0357	0.0239	0.0164	0.0116			

II. RESULTS

We generalize the BSA and BWA, beginning with the following:

Theorem 1. For any mixed state ρ and polynomial-invariant E there exists a pure-state ensemble containing at most one state with nonzero E.

The proof is given in the Appendix. Theorem 1 leads naturally to an approximation of ρ in terms of a mixed state for which E is equal to zero. By analogy with the BSA and BWA, we define the BEA of ρ as the state $\rho_e := \rho^*/\mathrm{Tr}\rho^*$, where ρ^* is a positive semidefinite operator with $E(\rho^*) = 0$ such that $\rho - \rho^* \geqslant 0$ and $\mathrm{Tr}\rho^* \leqslant 1$ is maximal. [Since $E(\rho^*)$ is a homogeneous polynomial in the expansion coefficients of the pure states in the minimal ensemble, it is well defined even if ρ^* has nonunit trace.] Moreover, we refer to the parameter $\mu := \mathrm{Tr}\rho^* \in [0,1]$ as the *zero-E equivalency* of ρ. Any state ρ has a convex decomposition of the form

$$\rho = \mu\rho_e + (1-\mu)\omega, \tag{4}$$

where ω is a pure state with nonzero E. We refer to (4) as the *optimal zero-E decomposition* of ρ, and ρ_e is the BEA. We now prove the following:

Theorem 2. All mixed states ρ have nonzero zero-E equivalency and have a unique optimal zero-E decomposition with ω being a pure state.

Theorem 2 relies on Lemmas 1, 2, and 3, whose proofs, together with the proofs of Theorems 1 and 2, are given in the Appendix.

Lemma 1. Consider $\rho,\pi \in \mathscr{D}(\mathcal{H})$ and let $E : \mathscr{D}(\mathcal{H}) \mapsto \mathbb{R}$ be a non-negative convex function bounded above by $E_{\max}$. Suppose that there exists some $k > 0$ such that

$$\sigma_\rho = \rho + \frac{k}{D(\rho,\pi)}(\rho - \pi) \tag{5}$$

is a state. Then,

$$E(\rho) - E(\pi) \leqslant \frac{D(\rho,\pi)}{D(\sigma_\rho,\pi)}[E(\sigma_\rho) - E(\pi)]. \tag{6}$$

The question of the existence of states of the form given by Eq. (5) is addressed by Lemma 2.

Lemma 2. For all $\rho,\pi \in \mathcal{D}(\mathcal{H})$ satisfying $\mathrm{supp}(\pi) \subseteq \mathrm{supp}(\rho)$, there exists a positive constant $k > 0$ such that the operator σ_ρ, defined as

$$\sigma_\rho := \rho + \frac{k}{D(\rho,\pi)}(\rho - \pi), \tag{7}$$

is a state and such that $\mathrm{rank}\,\sigma_\rho < \mathrm{rank}\,\rho$.

Equation (6), combined with Lemma 2, provides a nonuniform continuity bound on any non-negative convex function $E : \mathscr{D}(\mathcal{H}) \mapsto \mathbb{R}$. The continuity bound is nontrivial between two density matrices ρ and π as long as ρ and π have equal supports.

We have now generalized the BSA and BWA to arbitrary homogeneous polynomial invariants. However, the BEA for ρ does not, in general, provide the best estimate of $E(\rho)$ over the set of S decompositions. The entanglement of the S decomposition for ρ (note that this is an upper bound on the entanglement of ρ itself) with pure state ψ occurring with probability p is simply $pE(\psi)$. Hence, we define ψ_l to be the state in S such that $p_l E(\psi_l)$ is minimal and note

$$E(\rho) \leqslant p_l E(\psi_l) \leqslant p_e E(\psi_e), \tag{8}$$

where ψ_e is the pure state associated with the BEA for ρ. Because the BEA is unique, the second inequality is only an equality if the BEA minimizes $p_e E(\psi_e)$ as well as p_e, i.e., if $p_e = p_l$ and $\psi_e = \psi_l$.

We now describe an algorithm that may be used to determine ψ_l for any state ρ. We use the fact (from Lemma 3) that every mixed state ρ has in its range at least one pure state on which E is equal to zero.

Lemma 3. For any mixed state ρ, there is a pure state $|\psi\rangle \in R(\rho)$ such that $E(\psi) = 0$.

Given a mixed state $\rho \in \mathcal{D}(\mathcal{H})$ of rank d, we first use a steepest-descent algorithm [20] to identify pure states $\psi_i \in R(\rho)$ which have zero E [21]. For $d > 2$ there is a continuous set of such states, and the steepest-descent algorithm chooses one such state randomly. We repeat this procedure several times to identify a number [22] of such pure states $\{\psi_i\}$.

We then construct the uniform mixture π_1 of the pure states identified by the steepest-descent algorithm. Clearly, $\operatorname{supp}(\pi_1) \subseteq \operatorname{supp}(\rho)$ and $E(\pi_1) = 0$. Then, by Lemma 2, there exists a $k > 0$ such that the operator

$$\rho_1 = \rho + \frac{k}{D(\rho,\pi_1)}(\rho - \pi_1) \tag{9}$$

is a state and such that $\operatorname{rank}\rho_1 < \operatorname{rank}\rho$. We then apply Lemma 1 with $\sigma_\rho \equiv \rho_1$ and $\pi \equiv \pi_1$ to obtain

$$E(\rho) - E(\pi_1) \leqslant \frac{D(\rho,\pi_1)}{D(\rho_1,\pi_1)}[E(\rho_1) - E(\pi_1)]. \tag{10}$$

Hence, because $E(\pi_1) = 0$,

$$E(\rho) \leqslant \frac{D(\rho,\pi_1)}{D(\rho_1,\pi_1)}E(\rho_1). \tag{11}$$

From Eq. (9), ρ may be written as a convex combination of ρ_1 and π_1.

If ρ_1 is a pure state, then ρ may be written as a convex combination of the states ψ (comprising π_1), which have zero E, and the state ρ_1, which may have nonzero E. We have thus identified a pure-state ensemble for ρ containing at most one pure state with nonzero E, and the algorithm terminates since $E(\rho_1)$ can be calculated directly.

If ρ_1 is not pure, the same procedure is applied to ρ_1. We find a density matrix π_2 such that $E(\pi_2) = 0$ and $\operatorname{supp}(\pi_2) \subseteq \operatorname{supp}(\rho_1)$ and construct ρ_2 from it. The state ρ can then be written as a convex combination of the pure states comprising π_1 and π_2, and the (possibly mixed) state ρ_2. Then,

$$E(\rho_1) \leqslant \frac{D(\rho_1,\pi_2)}{D(\rho_2,\pi_2)}E(\rho_2). \tag{12}$$

We can now combine Eq. (12) with Eq. (11) to obtain

$$E(\rho) \leqslant \frac{D(\rho,\pi_1)}{D(\rho_1,\pi_1)}\frac{D(\rho_1,\pi_2)}{D(\rho_2,\pi_2)}E(\rho_2). \tag{13}$$

The procedure is then repeated for ρ_2. The algorithm terminates when one arrives at a state ρ_i which is pure, in which case $E(\rho_i)$ may be calculated directly. Because $\operatorname{rank}\rho_{\rm i} < \operatorname{rank}\rho_{\rm i-1}$ for all i, the algorithm is guaranteed to terminate, and we have

$$E(\rho) \leqslant \frac{D(\rho,\pi_1)}{D(\rho_1,\pi_1)} \cdots \frac{D(\rho_{d-1},\pi_d)}{D(\rho_d,\pi_d)}E(\rho_d), \tag{14}$$

where $\rho_d = \psi_d$ is pure. Note that at the ith step, we only need to find one pure state with $E = 0$ in the range of ρ_i for the algorithm to proceed.

The algorithm described above constructs an ensemble with the property that apart from ψ_d, every pure state in the ensemble has $E = 0$. Thus, $\psi_d \in S$. To find ψ_l, we then use the facts that S is a connected set and that the number of local minima in S is bounded above by the polynomial degree of E. The connectedness of S follows from the connectedness of the set of mixed states with zero E. For many entanglement monotones of interest, the degree of E is only 2 or 4 [5], so it follows that if we perform steepest descent to minimize the function $pE(\psi)$ over the set of pure states $\psi \in S$, where p is the probability of ψ in the S decomposition, starting from the state $\rho_d = \psi_d$, we will converge to the global minimum ψ_l with high probability.

The real dimension of S scales linearly in the rank of the density matrix, so the steepest descent is tractable. To perform this steepest descent in practice, for every state $\psi_i \in S$, we denote by $k_i \in (0,1)$ the smallest value such that for some mixed state π_i with $E(\pi_i) = 0$, $\rho = k_i\psi_i + (1 - k_i)\pi_i$. One can calculate k_i using the algorithm of Lewenstein and Sanpera, suitably adapted to use pure states with zero E, rather than separable states [16]. If no such value of k_i exists, we set $k_i = \infty$. One may then calculate ψ_l by minimizing $k_i E(\psi_i)$ by steepest descent over $\mathcal{H}$.

However, we have found that in many cases one may use Eq. (14) on its own to obtain a tight and computationally tractable numerical upper bound on $E(\rho)$. In this approach, one runs the algorithm up to Eq. (14) many times, getting a different result each time, and then takes the smallest of these results as an upper bound on $E(\rho)$.

To investigate the efficacy of this method we performed calculations of the three-tangle for three-qubit mixed states. The three-tangle is the simplest multipartite entanglement measure, and in this case the BEA is the BWA. The combination of the upper bound obtained in this way from Eq. (14) with the lower bound of [23] provides nontrivial upper and lower bounds on the three-tangle that one may compute rapidly on arbitrary states of three qubits. We evaluate both bounds for mixtures of GHZ-class and W-class states [4] where the three-tangle is known analytically [12,13] and for random rank-d density matrices. We compare the upper bound to analytical values of the three-tangle where available. We also compare the upper bound to the steepest-descent algorithm given in [15] and to the lower-bound algorithm of Eltschka and Siewert [23] for the square root of the three-tangle, whose square gives a lower bound on the three-tangle. Both algorithms are stochastic, so we repeat each one many times on a given state and use the

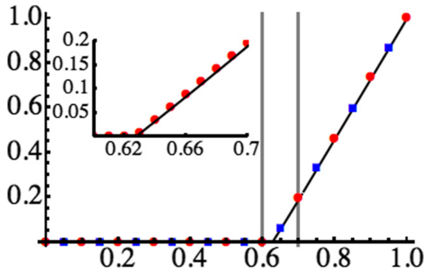

FIG. 1. (Color online) Upper bounds on the three-tangle of the states in Eq. (15) for 11 values of p (red circles) compared to the analytical value (line). For ten values of p (blue squares) the states in Eq. (15) were conjugated by a random element of $\mathrm{SU}(2)^{\otimes 3}$ sampled from the Gaussian unitary ensemble. For $p = 0$ through $p = 0.6$, the algorithm calculated the three-tangle to be zero to numerical precision. The inset shows the upper bounds (red dots) on the three-tangle of the density matrices in Eq. (15) for 11 values of p between 0.6 and 0.7, compared to the analytical value (line). These are results from 400 repetitions of the upper-bound algorithm.

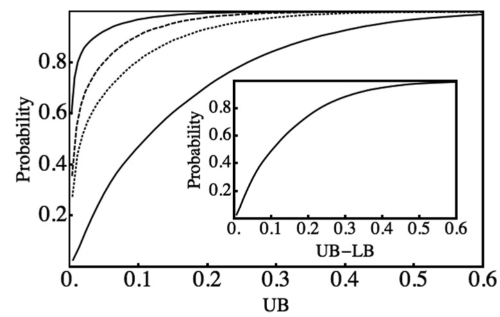

FIG. 2. Cumulative distribution function of the upper bound on the three-tangle of 10 000 randomly sampled three-qubit density matrices of rank 2 (lower solid line), 4 (dotted line), 6 (dashed line), and 8 (upper solid line). The inset shows the cumulative distribution function of the difference between the upper and lower bounds on the three-tangle of 30 000 uniformly sampled three-qubit rank-2 density matrices.

best result. For the calculations described below our upper bound is the minimum value obtained by running our algorithm 200 times on a given density matrix. The lower bound is the maximum value obtained after running the Eltschka and Siewert algorithm [23] 1000 times.

We evaluated the upper bound on mixtures of GHZ and W states,

$$\pi(p) = p|\mathrm{GHZ}\rangle\langle\mathrm{GHZ}| + (1-p)|W\rangle\langle W|, \qquad (15)$$

for which the analytical form is known [13]. Our algorithm was able to provide a tight upper bound for the three-tangle for this mixture (Fig. 1). In addition, whereas the steepest-descent algorithm always yields a nonzero value for the three-tangle, the algorithm presented above can and does identify the three-tangle as exactly zero to within numerical precision if the final pure state ρ_d has zero three-tangle. We also computed the upper bound for the case where the state is given by Eq. (15) convolved with a random element of $\mathrm{SU}(2)^{\otimes 3}$ constructed from three elements of SU(2) sampled independently from the Gaussian unitary ensemble. These states have identical entanglement to the states given by Eq. (15). We performed 400 repetitions of the algorithm and selected the smallest entanglement. Our findings are presented in Fig. 1 and in Table I.

We evaluated the upper bound on random mixtures of generalized GHZ and generalized W states for which the three-tangle is also known analytically [12,13]. On 20 000 rank-2 mixtures the average error of the upper bound was 0.025. On states that had zero three-tangle analytically, our upper bound yielded a value of exactly zero (to within numerical precision) 63% of the time. The lower bound of [23] was nonzero on 72% of density matrices with nonzero three-tangle.

The upper bound obtained from Eq. (14) was compared to the one obtained by the steepest-descent algorithm described in [15] for random density matrices. For ranks 2 through 5, both algorithms calculated upper bounds on the three-tangle of 240 different randomly generated density matrices. For the algorithm presented above, density matrices of ranks 6, 7, and 8 were also tested. The steepest-descent algorithm yielded a lower (better) value on average, but the difference decreased with increasing rank. The steepest descent was considerably slower than our algorithm for evaluating an upper bound, making calculations infeasible for ranks greater than 5. The timings and average differences are shown in Table I.

We computed upper and lower bounds on random three-qubit states with ranks 2 through 8, for which the analytical form is not known. We sampled 30 000 states for rank 2 and 10 000 states for ranks 3 through 8. We generated random rank-d states by sampling a probability distribution uniformly on the $(d-1)$-dimensional probability simplex and sampling pure states uniformly over the Hilbert space. The upper bound tightly constrains the three-tangle in this ensemble of states, as shown in Fig. 2. The median values of the three-tangle for ranks 2, 4, 6, and 8 are 0.11, 0.02, 0.013, and 0.003, respectively. The lower bound was mostly zero on these states: only 2561 of 30 000 states (8.5%) were nonzero for rank 2, 126 of 10 000 for rank 3, 12 of 10 000 for rank 4, 1 of 10 000 for ranks 5 and 6, and none for ranks 7 and 8. Hence, for random states of rank >2 the strategy of bounding the entanglement above and below is ineffective as we do not obtain a nontrivial lower bound from the method of [23] in these cases.

For random rank-2 states the mean upper and lower bounds over 30 000 states are 0.157 and 0.016, respectively, and the upper and lower bounds constrain the three-tangle to lie within a region with an average width of 0.14. If we restrict ourselves to those states on which the lower bound is nonzero so that we are considering states where we have a certificate that there is some entanglement, the mean upper and lower bounds are 0.356 and 0.188, respectively. Hence, states for which the lower bound is nonzero also have significantly larger values of the upper bound, and upper and lower bounds constrain the three-tangle to lie within a region with an average width of 0.167 for these states.

The algorithm will always terminate when applied to a polynomial entanglement monotone. On other convex roof entanglement monotones E, for which it may not be possible to construct ensembles containing at most one state on which E is nonzero, the algorithm should choose π_i to be the pure state in the support of ρ_{i-1} on which E is minimal. Then, Eq. (12) becomes [for the $(i-1)$th step]

$$E(\rho_{i-1}) \leqslant \frac{D(\rho_{i-1},\pi_i)}{D(\rho_i,\pi_i)}[E(\rho_i) - E(\pi_i)] + E(\pi_i). \tag{16}$$

Since π_i is pure, we may evaluate $E(\pi_i)$ analytically.

III. CONCLUSIONS

We generalized the best separable approximation and best W-class approximation to the best zero-E approximation for any polynomial invariant. The BEA (like the BSA and BWA) is defined by minimizing the probability of the single entangled state in the ensemble that defines the BEA. We defined an algorithm that minimizes the entanglement for ensembles that contain a single entangled state. We have presented computations of upper and lower bounds for the three-tangle that are both practical methods for calculations on any three-qubit state. We validated these methods on mixtures of generalized GHZ and W states for which the exact value of the three-tangle is known. The upper and lower bounds are close on a large fraction of random rank-2 states, including the fraction of those states for which the lower bound is nonzero and so for which we can certify that the entanglement is nonzero. Future work on the bounds may further close the gap between them and enable accurate estimation of the three-tangle, even if a closed form for the three-tangle remains out of reach.

ACKNOWLEDGMENTS

S.R. would like to thank D. Turban and J. Sabloff, and P.J.L. would like to thank J. Siewert and D. Meyer for useful discussions. N.D. is grateful to A. Sanpera for helpful discussions. P.J.L. acknowledges support from NSF CCI Center, Quantum Information for Quantum Chemistry (QIQC), Award No. CHE-1037992, by NSF Award No. PHY-0955518, and by AFOSR Award No. FA9550-12-1-0046. S.R. acknowledges support from the Hertz and Beckmann Foundations.

APPENDIX

In this appendix we give the proofs of the lemmas and theorems presented in the paper.

Proof of Lemma 1. Let

$$\omega_p := p\pi + (1-p)\sigma_\rho, \tag{A1}$$

with $p \in (0,1)$. Since π and σ_ρ are states, so is ω_p. Note that for the particular choice of p given by

$$p \equiv p_k = \frac{k}{D(\rho,\pi) + k}, \tag{A2}$$

we have that $\omega_{p_k} = \rho$. Using the expression for the trace distance between any two states ρ_1 and ρ_2,

$$D(\rho_1,\rho_2) = \max_{0 \leqslant P \leqslant I} \mathrm{Tr}[\mathrm{P}(\rho_1 - \rho_2)],$$

it can be readily verified that the following identities hold:

$$D(\pi,\sigma_\rho) = D(\pi,\rho) + D(\rho,\sigma_\rho), \tag{A3}$$

$$D(\rho,\sigma_\rho) = k. \tag{A4}$$

Let $\varepsilon := 1 - p_k$. Then, (A2), (A3), and (A4) imply that

$$p_k = \frac{D(\rho,\sigma_\rho)}{D(\pi,\sigma_\rho)}, \qquad \varepsilon = \frac{D(\rho,\pi)}{D(\pi,\sigma_\rho)}. \tag{A5}$$

Then,

$$\begin{aligned} E(\rho) - E(\pi) &= E(\omega_{p_k}) - E(\omega_{p_k+\varepsilon}) \\ &\leqslant \varepsilon[E(\sigma_\rho) - E(\pi)] \\ &= \frac{D(\rho,\pi)}{D(\pi,\sigma_\rho)}[E(\sigma_\rho) - E(\pi)]. \end{aligned} \tag{A6}$$

The first equality holds since the choice of p_k and ε ensures that $\omega_{p_k} = \rho$ and $\omega_{p_k+\varepsilon} = \omega_1 = \pi$. The inequality in the second line holds since E is a convex function and can be obtained as follows: Since $\varepsilon = 1 - p_k$, we have $\omega_{p_k} = p_k\pi + \varepsilon\sigma_\rho$. Then the convexity of E implies that

$$E(\omega_{p_k}) \leqslant p_k E(\pi) + \varepsilon E(\sigma_\rho),$$

and hence,

$$E(\omega_{p_k}) - E(\pi) \leqslant \varepsilon[E(\sigma_\rho) - E(\pi)].$$

The last equality in (A6) follows from (A5). ■

Proof of Lemma 2. It is clear from (5) that $\mathrm{Tr}\sigma_\rho = 1$ since $\rho,\pi \in \mathcal{D}(\mathcal{H})$. To establish that σ_ρ is a state we only need to show that $\sigma_\rho \geqslant 0$. In the following, for any $|\varphi\rangle \in \mathcal{H}$ and any $\omega \in \mathcal{D}(\mathcal{H})$ let $\omega^\varphi := \langle\varphi|\omega|\varphi\rangle$. Any $|\varphi\rangle \in \mathcal{H}$ can be written as

$$|\varphi\rangle = \Pi_\rho|\varphi\rangle + (I - \Pi_\rho)|\varphi\rangle,$$

where Π_ρ denotes the projection onto the support of ρ. Obviously, $\rho(I - \Pi_\rho)|\varphi\rangle = 0$ and $\pi(I - \Pi_\rho)|\varphi\rangle = 0$ since $\mathrm{supp}\,(\pi) \subseteq \mathrm{supp}\,(\rho)$. These identities imply that $\sigma_\rho(I - \Pi_\rho)|\varphi\rangle = 0$, and hence,

$$\sigma_\rho^\varphi = \langle\varphi|\Pi_\rho\sigma\Pi_\rho|\varphi\rangle.$$

Let us define

$$|\widetilde{\varphi}\rangle := \frac{\Pi_\rho|\varphi\rangle}{\sqrt{\langle\varphi|\Pi_\rho|\varphi\rangle}}. \tag{A7}$$

Then to prove that $\sigma_\rho \geqslant 0$, it suffices to show that $\sigma_\rho^{\widetilde{\varphi}} \geqslant 0$. From (5), it equivalently suffices to prove that

$$\rho^{\widetilde{\varphi}} \geqslant \frac{k\left(\pi^{\widetilde{\varphi}} - \rho^{\widetilde{\varphi}}\right)}{D(\rho,\pi)}. \tag{A8}$$

Note that $D(\rho,\pi) = D(\pi,\rho)$ by symmetry and that

$$\begin{aligned} D(\pi,\rho) &= \max_{0 \leqslant P \leqslant I} \mathrm{Tr}[\mathrm{P}(\pi - \rho)], \\ &\geqslant \mathrm{Tr}[|\widetilde{\varphi}\rangle\langle\widetilde{\varphi}|(\pi - \rho)] \\ &= \pi^{\widetilde{\varphi}} - \rho^{\widetilde{\varphi}}. \end{aligned} \tag{A9}$$

Hence, to prove (A8), it suffices to prove that there exists a positive constant k such that

$$\rho^{\widetilde{\varphi}} \geqslant k. \tag{A10}$$

Let the eigenvalue decomposition of ρ be given by

$$\rho = \sum_{i=1}^{d} \lambda_i |e_i\rangle\langle e_i|, \tag{A11}$$

and let us choose $k = \lambda_{\min}(\rho)$, where $\lambda_{\min}(\rho) := \min_{1\leqslant i\leqslant d}\{\lambda_i : \lambda_i > 0\}$. Obviously, $|\widetilde{\varphi}\rangle \in \operatorname{supp}\rho$, and hence, $|\widetilde{\varphi}\rangle = \sum_{i:\lambda_i>0}\alpha_i|e_i\rangle$, with $\alpha_i = \langle e_i|\widetilde{\varphi}\rangle$ and $\sum_{i:\lambda_i>0}|\alpha_i|^2 = 1$. Hence,

$$\begin{aligned}\rho^{\widetilde{\varphi}} &= \langle\widetilde{\varphi}|\left(\sum_{i:\lambda_i>0}\lambda_i|e_i\rangle\langle e_i|\right)|\widetilde{\varphi}\rangle \\ &\geqslant \lambda_{\min}(\rho)\sum_{i:\lambda_i>0}|\langle\widetilde{\varphi}|e_i\rangle|^2 \\ &= \lambda_{\min}(\rho). \end{aligned}\tag{A12}$$

It follows that the operator σ_ρ defined in Eq. (5) with $k = \lambda_{\min}(\rho)$ is a state. However, if $\rho \neq \pi$, then $\rho - \pi$ has at least one negative eigenvalue. It follows that for $k > 0$ large enough, the operator defined in Eq. (5) is not positive semidefinite and hence is not a state. By continuity, then, there exists a $p > 0$ such the operator σ_ρ defined in Eq. (5) with $k = p$ is a state but such that the operator σ_ρ as defined in Eq. (5) is not a state when $k = p + \epsilon$ for any positive ϵ. Define $\bar{\sigma}_\rho$ to be the operator σ_ρ defined in Eq. (5) with $k = p$.

If the support of π were contained within the support of $\bar{\sigma}_\rho$, then there would exist some $q > 0$ such that

$$\bar{\sigma}_\rho + \frac{q}{D(\bar{\sigma}_\rho,\pi)}(\bar{\sigma}_\rho - \pi) \geqslant 0. \tag{A13}$$

If such a q existed, however, then one could use the fact that from the definitions of σ_ρ and the trace distance,

$$p = D(\rho,\sigma_\rho)$$

and

$$D(\pi,\rho) + D(\rho,\sigma_\rho) = D(\pi,\sigma_\rho)$$

in order to show that

$$\rho + \frac{p+q}{D(\rho,\pi)}(\rho - \pi) \geqslant 0,$$

which is a contradiction by the definition of p. Thus, the support of π is not contained within the support of $\bar{\sigma}_\rho$, so the supports of $\bar{\sigma}_\rho$ and ρ are not equal. However, ρ can be written as a convex combination of $\bar{\sigma}_\rho$ and π, so the support of $\bar{\sigma}_\rho$ must be contained within the support of ρ. It follows that the support of $\bar{\sigma}_\rho$ is strictly smaller than the support of ρ, so the rank of $\bar{\sigma}_\rho$ must be smaller than the rank of ρ. ■

Proof of Lemma 3. Since the range of any mixed state ρ is the set of superpositions of the eigenvectors of ρ, it suffices to show that for any two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$, there exists some θ, ϕ, such that the state

$$|\theta,\phi\rangle = \cos(\theta)|\psi_1\rangle + e^{i\phi}\sin(\theta)|\psi_2\rangle \tag{A14}$$

has zero E.

Let the coefficients of $|\psi_1\rangle$ in the computational basis be written $b_1,b_2,\ldots,b_8$, and let the coefficients of $|\psi_2\rangle$ in the computational basis be written $c_1,c_2,\ldots,c_8$. Then for all k, the coefficients of $|\theta,\phi\rangle$, written in the computational basis, are given by

$$a_k = \cos(\theta)b_k + e^{i\phi}\sin(\theta)c_k. \tag{A15}$$

Since E is a polynomial of homogeneous degree D equal to at most 4 in the coefficients of $|\theta,\phi\rangle$ written in the computational basis, we may write it as

$$E(|\theta,\phi\rangle) = \sum_i^T C_i a_1^{s_{1,i}}\cdots a_N^{s_{N,i}}, \tag{A16}$$

where T is the number of terms in the expression for E, N is the dimension of the Hilbert space, C_i is the coefficient on the ith term, and $\sum_{k=1}^N s_{k,i} = D$ for all i. Substituting the expression for a_k, we have

$$E(|\theta,\phi\rangle) = \sum_i^T C_i\prod_{k=1}^N[\cos(\theta)b_k + e^{i\phi}\sin(\theta)c_k]^{s_{k,i}}. \tag{A17}$$

Factoring out $\cos(\theta)^D$ from every term in the sum, we have

$$E(|\theta,\phi\rangle) = \cos(\theta)^D\sum_i^T C_i\prod_{k=1}^N[b_k + e^{i\phi}\tan(\theta)c_k]^{s_{k,i}}. \tag{A18}$$

We now perform a change of variables, defining $z(\theta,\phi) = e^{i\phi}\tan(\theta)$. We have

$$E(|\theta,\phi\rangle) = \cos(\theta)^D\sum_i^T C_i\prod_{k=1}^N[b_k + z(\theta,\phi)c_k]^{s_{k,i}}. \tag{A19}$$

We now note that the range of θ may be restricted to the interval $[0,\pi/2]$, while the range of ϕ is $[0,2\pi]$. If we assume that $E(|\psi_2\rangle) \neq 0$, then for the purpose of finding the roots of E, the range of θ may be restricted further to $[0,\pi/2)$. On this range, $\cos(\theta)^D$ is nonzero, so the roots of $E(|\theta,\phi\rangle) = 0$ are equivalent to the roots of the polynomial

$$\bar{E} = \sum_i^T C_i\prod_{k=1}^N[b_k + z(\theta,\phi)c_k]^{s_{k,i}}. \tag{A20}$$

The fundamental theorem of algebra guarantees that $\bar{E}$ will have D complex roots, including multiplicities. These roots lie within the range of $z(\theta,\phi)$, which is the entire complex plane. Thus, there exists at least one unique pure state $|\theta,\phi\rangle$ such that $E(|\theta,\phi\rangle) = 0$, completing the proof. ■

Proof of Theorem 2. Every pure state $|\psi\rangle \in R(\omega)$ [where ω is the state appearing in the optimal zero-E decomposition (4)] must have positive E. This is because, if there is a pure state $|\varphi\rangle \in R(\omega)$ with $E(\varphi) = 0$, then we could subtract $\gamma\varphi$ from ω (for some $0 < \gamma < 1$) and add $(1-\mu)\gamma\varphi$ to $\mu\rho_L$ to obtain a decomposition of the form

$$\rho = \tilde{\mu}\tilde{\rho}_L + (1-\tilde{\mu})\tilde{\omega},$$

such that $E(\tilde{\rho}_L) = 0$ and $\tilde{\mu} = \mu + (1-\mu)\lambda$. However, this increases the zero-E equivalency μ and hence leads to a contradiction since μ is maximal by definition. Hence, $E(\psi) > 0\ \forall\,\psi \in R(\omega)$. By Lemma 3, it then follows that ω must be pure. This also implies that if ρ is a mixed state, then $\mu > 0$.

To prove that the optimal zero-E decomposition is unique, we proceed as in [17] and assume that there exist at least two optimal zero-E decompositions $\rho = \lambda\rho_L + (1-\lambda)|\psi\rangle\langle\psi|$ and $\rho = \lambda\rho_L' + (1-\lambda)|\psi'\rangle\langle\psi'|$ with the same maximal λ. Any convex combination of these two decompositions is also an optimal zero-E decomposition, i.e., $\forall\,\varepsilon \in [0,1]$,

$$\begin{aligned}\rho &= \varepsilon[\lambda\rho_L + (1-\lambda)\psi] + (1-\varepsilon)[\lambda\rho_L' + (1-\lambda)\psi'] \\ &= \lambda\tilde{\rho}_L + (1-\lambda)\tilde{\omega}, \end{aligned}\tag{A21}$$

where $\tilde{\omega} := \varepsilon\psi + (1-\varepsilon)\psi'$ and $\tilde{\rho}_L := \varepsilon\rho_L + (1-\varepsilon)\rho'_L$, with $E(\tilde{\rho}_L) = 0$ [since the convex roof extension E is convex and $E(\rho_L) = 0 = E(\rho'_L) = 0$]. Since $\tilde{\omega}$ is a mixed state, by Lemma 3 there must exist a pure state $|\varphi>$ in its range such that (as above) we could subtract $c\varphi$ from $\tilde{\omega}$ [for some $c \in (0,1)$] and add it to $\lambda\tilde{\rho}_L$ to obtain another optimal zero-E decomposition. However, this would increase the zero-E equivalency and thus would result in a contradiction. ∎

Proof of Theorem 1. We prove Theorem 1 by construction, using Lemmas 1, 2, and 3. Since ρ is a mixed state, by Lemma 3 there exists a pure state $|\psi\rangle \in R(\rho)$ such that $E(\psi) = 0$. By Lemma 2 we infer that there exists a positive constant k and a state ρ_1 such that

$$\rho = \lambda\psi + (1-\lambda)\rho_1, \tag{A22}$$

with rank $\rho_1 <$ rank ρ. Here $\lambda \equiv \lambda(k, D(\rho,\psi)) = k/[D(\rho,\psi) + k]$.

If ρ_1 is a pure state, then the claim of Theorem 1 follows since we have constructed an ensemble of two pure states ψ and ρ_1 for ρ, with $E(\psi) = 0$ and $E(\rho_1) \geqslant 0$. If ρ_1 is a mixed state, then we know by Lemma 2 that there exists a pure state $|\psi_1\rangle \in R(\rho_1)$ such that $E(\psi_1) = 0$. Then we repeat the above steps (for ρ_1) to arrive at a state ρ_2 such that

$$\rho = \lambda_1\psi_1 + (1-\lambda_1)\rho_2$$

and $\lambda_1 \in (0,1)$. If ρ_2 is pure, the proof is completed since ρ can be expressed in terms of an ensemble of three pure states, ψ, ψ_1, and ρ_2, with only ρ_2 having possibly nonzero E. If ρ_2 is mixed, we iterate again. We stop after the ith iteration if ρ_i is pure. Since, by Lemma 2, the rank of the state ρ_i obtained after the ith iteration is strictly smaller than the rank of ρ_{i+1}, we definitely arrive at a pure state, and hence, the iteration stops, yielding an ensemble of pure states for ρ with at most one (namely, the one obtained in the last step of the iteration) having nonzero E. ∎

---

[1] J. Volz, M. Weber, D. Schlenk, W. Rosenfeld, J. Vrana, K. Saucke, C. Kurtsiefer, and H. Weinfurter, Phys. Rev. Lett. **96**, 030404 (2006).

[2] K. J. Resch, P. Walther, and A. Zeilinger, Phys. Rev. Lett. **94**, 070402 (2005).

[3] H. Häffner, C. F. Roos, and R. Blatt, Phys. Rep. **469**, 155 (2008).

[4] W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).

[5] C. Eltschka, T. Bastin, A. Osterloh, and J. Siewert, Phys. Rev. A **85**, 022301 (2012).

[6] D. A. Meyer and N. R. Wallach, J. Math. Phys. **43**, 4273 (2002).

[7] G. Vidal, J. Mod. Opt. **47**, 355 (2000).

[8] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[9] V. Coffman, J. Kundu, and W. K. Wootters, Phys. Rev. A **61**, 052306 (2000).

[10] A. Uhlmann, Open Syst. Inf. Dyn. **5**, 209 (1998).

[11] S. Hill and W. K. Wootters, Phys. Rev. Lett. **78**, 5022 (1997).

[12] C. Eltschka, A. Osterloh, J. Siewert, and A. Uhlmann, New J. Phys. **10**, 043014 (2008).

[13] R. Lohmayer, A. Osterloh, J. Siewert, and A. Uhlmann, Phys. Rev. Lett. **97**, 260502 (2006).

[14] E. Jung, M.-R. Hwang, D. K. Park, and J.-W. Son, Phys. Rev. A **79**, 024306 (2009).

[15] J. Zhu, S. Kais, A. Aspuru-Guzik, S. Rodriques, B. Brock, and P. J. Love, J. Chem. Phys. **137**, 074112 (2012).

[16] M. Lewenstein and A. Sanpera, Phys. Rev. Lett. **80**, 2261 (1998).

[17] S. Karnas and M. Lewenstein, J. Phys. A **34**, 6919 (2001).

[18] T. Wellens and M. Kuś, Phys. Rev. A **64**, 052302 (2001).

[19] A. Acin, D. Bruß, M. Lewenstein, and A. Sanpera, Phys. Rev. Lett. **87**, 040401 (2001).

[20] This algorithm, which performs steepest descent to minimize E over $\mathcal{H}$ with a cost independent of d, should not be confused with the steepest-descent convex-roof algorithm, which minimizes the three-tangle over Υ_ρ and which has a cost scaling like $d^8 \log d$.

[21] These pure states could also be found using a root-finding algorithm on the polynomial defined in the proof of Theorem 1.

[22] Carathéodory's theorem implies that pure-state ensembles of size d^2 are required in general in order to minimize the convex roof. For this reason, π_i is constructed as a uniform mixture of $3/2d - i$ distinct pure states selected from the range of ρ_i, each with zero E. Because $\sum_{i=0}^{d-2}(3/2d - i) = d^2 - 1$, this ensures that d^2 pure states are selected in total, including the final pure state.

[23] C. Eltschka and J. Siewert, Phys. Rev. Lett. **108**, 020502 (2012).