

Nonconvex optimization strategy for computing convex-roof entanglement

Jimmie Adriazola¹ and Katarzyna Roszak²

¹*Department of Mathematics, Southern Methodist University, Dallas, Texas, USA*

²*FZU - Institute of Physics of the Czech Academy of Sciences, 182 00 Prague, Czech Republic*



(Received 21 December 2024; accepted 14 March 2025; published 28 March 2025)

We develop a numerical methodology for the computation of entanglement measures for mixed quantum states. Using the well-known Schrödinger-Hughston-Jozsa-Wooters theorem, the computation of convex-roof entanglement measures is reframed as a search for unitary matrices: a nonconvex optimization problem. To address this nonconvexity, we modify a genetic algorithm, known in the literature as differential evolution, constraining the search space to unitary matrices by using a QR factorization. We then refine results using a quasi-Newton method. We benchmark our method on simple test problems and, as an application, compute entanglement between a system and its environment over time for pure dephasing evolutions. We also study the temperature dependence of Gibbs state entanglement for a class of block-diagonal Hamiltonians to provide a complementary test scenario with a set of entangled states that are qualitatively different. We find that the method works well enough to reliably reproduce entanglement curves, even for comparatively large systems. The modified genetic algorithm is a derivative-free, nonconvex computational method that broadly applies to the computation of convex-roof entanglement measures.

DOI: [10.1103/PhysRevA.111.032435](https://doi.org/10.1103/PhysRevA.111.032435)

I. INTRODUCTION

Quantification of entanglement for *pure* quantum states is relatively straightforward because classical correlations cannot be present in such states. Thus, any measured correlation is quantum, implying that entanglement between two parts of a pure quantum system is inversely proportional to the purity of the state of one (either) part of the system that is obtained after tracing out the other. Measures such as entanglement entropy [1] operate in this fashion—that is, they quantify entanglement by quantifying the purity of these subsystems (via von Neumann entropy here, but any measure of purity, such as linear entropy [2], etc., is usable).

Unfortunately, realistic quantum systems are rarely pure, owing to outside influences such as interactions with uncontrollable (often quantum) environments [3,4], classical noise sources [5], and difficulties related to reliable repetition of experiments. The presence of noise necessitates density matrices instead of quantum state vectors, since density matrices are capable of capturing statistical properties of quantum ensembles. Such states can contain classical correlations (and quantum-classical [6]), making the detection of entanglement through the purity of part of the system alone impossible.

A straightforward solution to this is to quantify the entanglement of the mixed state by calculating the average of the entanglement of the pure state of the eigendecomposition of the density matrix. This fails because of the way the set of *separable* states is defined: Any state is separable if and only if it is possible to decompose it into a mixture of pure separable states. The definition has a deeper physical meaning that entanglement cannot be obtained via local operations and classical communication (LOCC) [7,8].

A trivial example would be an equal mixture of two maximally entangled Bell states,

$$\hat{\rho}_{\text{sep1}} = \frac{1}{2}|\Psi_+\rangle\langle\Psi_+| + \frac{1}{2}|\Psi_-\rangle\langle\Psi_-|, \quad (1)$$

with $|\Psi_{\pm}\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$. Since each Bell state $|\Psi_{\pm}\rangle$ is maximally entangled, the state $\hat{\rho}$ should also be maximally entangled, but it is in fact separable. This is evident once it is written in the separable basis $\{|ij\rangle\}$, with $i, j = 0, 1$, where we get

$$\hat{\rho}_{\text{sep1}} = \frac{1}{2}|01\rangle\langle 01| + \frac{1}{2}|10\rangle\langle 10|. \quad (2)$$

The state is correlated, but only classically. In less trivial cases, the decomposition that fulfills the definition and demonstrates separability typically cannot be obtained via a change of (orthonormal) basis (since it is not an eigendecomposition) and is thus much harder to find. Such mixed states also include other mixtures of Bell states, such as separable X states [9], e.g., $\hat{\rho}_{\text{sep2}} = 1/4|00\rangle\langle 00| + 1/4|11\rangle\langle 11| + 1/4|\Psi_+\rangle\langle\Psi_+|$ or separable Werner states [10]. Prescriptions to find explicit separable decompositions for two-qubit states can be found in Ref. [11], while Werner states are discussed in Ref. [12].

To overcome this, the notion of convex-roof entanglement measures was developed. Convex-roof entanglement measures are based on the idea that the actual amount of entanglement can be quantified as an average of pure-state entanglement of some pure-state decomposition of a state, but the pure-state decomposition has to be the one that minimizes entanglement. If the von Neumann entropy is used as the purity measure in the quantification of pure-state entanglement then the convex-roof entanglement measure is called entanglement of formation (EoF) [13,14]. The minimization problem is complex because it requires the use of an in principle

infinite set of unitary matrices and the number of minimization parameters grows with system size N as $N^2 - 1$.

A side effect of the complexity of minimization is the relatively small number of studies of mixed-state entanglement for larger systems. For two-qubit entanglement, there exists a direct method for the quantification of the EoF [11] and consequently there is an immense number of papers that discuss two-qubit correlations, providing a detailed study of its, often curious, properties. For larger systems, the dominant tool is negativity [15], which does not relate to any measure of pure-state entanglement and cannot quantify certain types of entanglement [16], but is much more accessible numerically.

Beyond that, studies of the EoF or other convex-roof entanglement measures are restricted to scenarios where certain limitations on the symmetry of the density matrices are imposed [17–20], which limit the scope of the minimization. This does not mean that larger entangled states are just a generalization of two-qubit states. On the contrary, there exist properties characteristic for qudits that cannot be reproduced or explained by qubit systems [16,21] and not all of them can be captured by negativity [16]. However, the numerical resources required for the study of entanglement between larger systems are prohibitive for systematic studies to be performed.

An excellent methodology, owing to R  thlisberger *et al.* [22] already exists for the efficient calculation of entanglement measures. This study even led to the development of a software library titled libCreme [23]. In short, R  thlisberger, and others, adapt gradient-based line searches on matrix manifolds [24] to compute convex-roof entanglement measures. Although this study successfully paved the way for quantifying entanglement using powerful numerical optimization methods, the methodology has two drawbacks. First, deterministic line search methods are inefficient for identifying globally optimal points owing to the curse of nonconvexity. Indeed, despite its name, the computation of convex-roof entanglement measures is not formulated as a convex optimization problem. Second, the computation of gradients on matrix manifolds is not automatic and impedes the adoptability of these gradient-based methods to a wide class of problems encountered in practice.

In this paper, we introduce an algorithm for the efficient computation of convex-roof entanglement measures. In principle, it can be treated as a black box, and thus can be broadly applied for the computation of entanglement. Moreover, it is based on a genetic algorithm, a heuristic search method that addresses the nonconvexity of the optimization problem. The method is then accelerated toward the nearest local minimum via a black box implementation of a quasi-Newton method. Our intention with this hybrid, nonconvex-line search approach is to present an efficient algorithm while maintaining

a high level of usability for communities interested in quantifying entanglement in applications and less interested in the mathematical or computational details under the hood.

We test the algorithm on a series of examples for which the amount of entanglement is known, or can be at least estimated. We choose examples that involve a strong dependence on a parameter, such as time or temperature, so that the level of success of the algorithm can also be evaluated on its capability to describe the trends in entanglement behavior. This is particularly important in the study of entanglement in realistic scenarios and in order to identify the properties of entanglement between qudits that are not manifested in qubit-qubit entanglement.

We use two-qubit examples to tune the parameters of the procedure for optimal operation and to demonstrate that the algorithm is capable of demonstrating sudden death of entanglement. We then use it to quantify the evolution of entanglement between a qubit and an environment of many qubits. We obtain smooth, reliable curves, which are expected in a unitary evolution as long as the overall purity of the system is not too small. The last example pertains to the amount of entanglement found in a bipartite temperature-equilibrium state. We find that at low temperatures, when the overall purity of the state is still large, the algorithm performs well, and the obtained curves are smooth, but at larger temperatures the noisiness of the results becomes non-negligible, although the overall trend for the temperature dependence of entanglement is clearly visible.

The paper is organized as follows. In Sec. II we specify the optimization problem under study. The modified genetic algorithm/quasi-Newton method used to tackle the optimization is discussed in Sec. III. Section IV contains a number of examples of the operation of the algorithm on cases with known or partially quantifiable entanglement. Finally, Sec. V concludes the paper.

II. THE OPTIMIZATION PROBLEM

The computation of entanglement measures we pursue uses the well-known convex-roof construction [13,25]. To begin with its description, we define the notion of a pure-state decomposition. Given a density matrix ρ of rank r acting on a Hilbert space \mathcal{H} of finite dimension d , a pure-state decomposition is defined by a set of $s \geq r$ probabilities p_i and vectors $|\psi_i\rangle \in \mathcal{H}$ such that

$$\rho = \sum_{i=1}^s p_i |\psi_i\rangle \langle \psi_i|. \quad (3)$$

The set of all pure-state decompositions is then simply given by

$$\mathfrak{D}(\rho) = \left\{ \{p_i, |\psi_i\rangle\}_{i=1}^s, \{|\psi_i\rangle\}_{i=1}^s \subset \mathcal{H}, p_i \geq 0, \sum_{i=1}^s p_i = 1 \mid \rho = \sum_{i=1}^s p_i |\psi_i\rangle \langle \psi_i| \right\}.$$

The construction of convex roofs further requires a monotone entanglement $m : \mathcal{H} \rightarrow \mathbb{R}^+$ that quantifies the amount of

entanglement in the system [26]. The function m is monotonically decreasing, under local operations and classical

communication [27], and evaluates to zero if the density matrix ρ is separable, i.e., ρ can be written as a convex combination of product states.

As a relevant example of an entanglement monotone, consider the entropy of entanglement. For a pure bipartite quantum state of a pure composite system ρ_{AB} , the entropy of entanglement is given by the von Neumann entropy of the reduced density matrix,

$$m_{\text{vN}}(\rho_{AB}) = -\text{Tr}[\rho_A \ln \rho_A] = -\text{Tr}[\rho_B \ln \rho_B], \quad (4)$$

where $\rho_A = \text{Tr}_B(\rho_{AB})$ and $\rho_B = \text{Tr}_A(\rho_{AB})$ are the reduced density matrices given as partial traces for each partition.

The convex-roof entanglement measure $M : \mathcal{H} \rightarrow \mathbb{R}^+$ is thus defined by the following optimization problem

$$M(\rho) = \inf_{\{p_i, |\psi_i\rangle\} \in \mathcal{D}(\rho)} \sum_i p_i m(|\psi_i\rangle). \quad (5)$$

In words, this computation constitutes a search for the smallest mean evaluation of a given entanglement monotone m over the space \mathcal{D} of all pure-state decompositions corresponding to a given density matrix ρ . As written, this search is intractable.

Fortunately, a theorem owing to Schrödinger [28] and Hughston *et al.* [29], colloquially known as the Schrödinger-HJW theorem, furnishes a unitary parametrization of the space \mathcal{D} . More precisely, let $\mathcal{U}(k, r)$ denote the set of all $k \times r$ matrices $U \in \mathbb{C}^{k \times r}$ with the property that $U^\dagger U$ equals the $r \times r$ identity matrix denoted by $\mathbb{I}_{r \times r}$. The theorem states that every $U \in \mathcal{U}(k, r)$ produces a pure-state decomposition $\{p_i, |\psi_i\rangle\}_{i=1}^k \in \mathcal{D}(\rho)$ of the density matrix ρ .

To see this, let λ_i and $|v_i\rangle$ denote the eigenvalues and corresponding normalized eigenvectors of ρ in the representation

$$\rho = \sum_{j=1}^r \lambda_j |v_j\rangle \langle v_j|. \quad (6)$$

Note that since ρ is a density matrix, it is by definition a positive definite matrix with a well-defined spectral decomposition. Now, given a matrix $U \in \mathcal{U}(k, r)$, define the auxiliary state

$$|\varphi_i\rangle = \sum_{j=1}^r U_{ij} \sqrt{\lambda_j} |v_j\rangle, \quad i = 1, \dots, k,$$

corresponding to the spectral content of ρ in Eq. (6). Observe that the auxiliary states $|\varphi_i\rangle$ yield a pure-state decomposition through Eq. (3) with the probabilities and state vectors given by

$$p_i = \langle \varphi_i | \varphi_i \rangle, \quad |\psi_i\rangle = \frac{1}{\sqrt{p_i}} |\varphi_i\rangle.$$

Moreover, and this is key, the theorem states that every pure-state decomposition of ρ can be realized from a matrix $U \in \mathcal{U}(k, r)$. Thus, the optimization problem can be reformulated as the matrix optimization problem

$$\min_{k \geq r} \inf_{U \in \mathcal{U}(k, r)} J(U) = \min_{k \geq r} \inf_{U \in \mathcal{U}(k, r)} \sum_{i=1}^k p_i(U) m(|\psi_i(U)\rangle), \quad (7)$$

ALGORITHM 1. Differential mutation.

Result: A vector z mutated from agents in a given generation as required by the DE Algorithm (2).

Input: 4 distinct members a, b, c, d from the current generation of agents each with N components, the crossover ratio $C_R \in (0, 1)$, and weight $F \in (0, 2)$.

```

for  $j=1:N$  do
    Compute a random variable rand;
    if rand <  $C_R$  then
        |  $z[j] \leftarrow a[j] + F * (b[j] - c[j])$ 
    else
        |  $z[j] \leftarrow d[j]$ 
    end
end
end

```

where it is implied that p_i and ψ_i are computed from the Schrödinger-HJW method, for some $U \in \mathcal{U}(k, r)$, applied to a given density matrix ρ . Although the Schrödinger-HJW theorem allows us to reformulate this problem as a more tractable search over unitary matrices, the ambiguity in the parameter k forces us to perform a brute force search over the size of the unitary matrices furnishing the pure-state decompositions. We address all computational issues concerning this optimization problem in the following section.

III. THE OPTIMIZATION METHOD

To search for the optimal unitary matrix in problem (7), we modify a preexisting optimization method owing to Storn and Price [30,31]. The algorithm is called differential evolution (DE) and is a stochastic optimization method used to search for candidate solutions to generally nonconvex optimization problems. The idea behind DE is inspired by evolutionary genetics and is thus part of a class of so-called genetic algorithms. Genetic algorithms are used in a wide context of optimization and design problems where derivative information of the objective function is inaccessible or difficult to compute.

DE searches the space of candidate solutions by initializing a population set of vectors, known as agents, within some region of the search space. These agents are then mutated (see Algorithm 1) into a new population set, or generation. The mutation operates via two mechanisms: a weighted combination and a random “crossover.”

At each generation, Algorithm 1 generates a candidate z to replace each agent y . In the mutation step, it chooses at random three agents a, b , and c to create a new trial agent \tilde{z} through the linear combination

$$\tilde{z} = a + F \cdot (b - c),$$

where $F \in [0, 2]$. In the crossover step, a new candidate vector z is constructed by randomly choosing some components \tilde{z} and others from an additional randomly chosen agent d . If we evaluate an objective functional $J : \mathbb{R}^N \rightarrow \mathbb{R}$ and find that $J(z) < J(y)$, then z replaces y in the next generation.

DE ensures that the objective functional J of the optimization problem decreases monotonically with (the optimal

member of) each generation. As each iteration “evolves” into the next, inferior agents “inherit” optimal traits from superior agents via mutation, or else they are discarded. After a sufficient number of iterations, the best vector in the final generation is chosen as the candidate global optimizer. A more detailed discussion about further implementation and benchmarking details can be found in the book by Price *et al.* [31].

As discussed, the classical DE algorithm is broken into two parts: a mutation step and an evolution step. The mutation step remains unchanged. However, the evolution step is modified so that candidate optimizers remain in the space \mathcal{U} . To enforce the unitary constraint, we use the QR decomposition [32], which we now briefly review.

It is a fact that any real square matrix A may be decomposed as

$$A = QR,$$

where Q is an orthogonal matrix and R is an upper triangular matrix. If instead A is a complex square matrix, then Q is unitary. More generally, we can factor a complex $k \times r$ matrix A , with $k \geq r$, as the product of a $k \times k$ unitary matrix Q and a $k \times r$ upper triangular matrix R .

The key steps to adapting the DE method to the calculation of convex-roof entanglement measures are as follows. To make DE mutation compatible with an optimization over unitary matrices, we must first reshape candidate optimizing matrices from $k \times r$ to $k * r \times 1$ since mutation is performed at the vector level. We then perform mutation via Algorithm 1. The resulting mutated vector is then reshaped back into a $k \times r$ matrix \tilde{Q} that is most likely no longer unitary because mutation does not preserve unitarity.

In this step, we use the QR decomposition to ensure that a function evaluation of J appearing in Eq. (7) is done so with a unitary matrix. That is, we apply a QR factorization in \tilde{Q} to obtain a unitary matrix Q of size $k \times k$ and delete its last $k - r$ columns so that it is a unitary matrix U of size $k \times r$. The candidate matrix U now belongs to the feasible optimization space $\mathcal{U}(k, r)$ as desired. This evolutionary process is summarized entirely by Algorithm 2.

Typically, DE calculates the convex-roof entanglement measure correctly to one or two digits of accuracy. To refine the results and accelerate convergence, we use the DE output and feed it into a BFGS method [33], as implemented by MATLAB’s `fminunc` function. Of course, inside of the evaluation of the entanglement measure, we perform a QR decomposition to ensure that the BFGS method remains constrained to searches over unitary spaces. The calculation of the QR factorization is handled via MATLAB’s built-in `qr` function.

One final comment we make is about the use of the QR decomposition to constrain our search space to unitary matrices. Perhaps, a more natural way to project onto this space is to find the closest unitary matrix \hat{U} to A , in the sense of the operator norm. This would be facilitated by finding \hat{U} such that

$$\hat{U} = \arg \min_{\hat{U}} \|A - \hat{U}\|.$$

The solution to this problem is well known [34]. The unitary matrix \hat{U} that minimizes $\|A - \hat{U}\|$ can be calculated via the

ALGORITHM 2. Unitary differential evolution.

Result: A vector likely to be globally optimal with respect to an objective J .

Input: A maximum number of iterations N_{\max} , crossover ratio $C_R \in (0, 1)$, weight $F \in (0, 2)$, and matrix of size $k \times r$

Generate a random population of $k \times r$ unitary matrices

Reshape into a population `pop` of N_{pop} vectors of dimension $k * r$.

while `counter` < N_{\max} **do**

for $i = 1 : N_{\text{pop}}$ **do**

`CurrentMember` \leftarrow `Popi`;

 Choose three distinct vectors a_i, b_i, c_i different from the vector `Popi`;

 Mutate a_i, b_i, c_i , and the `CurrentMember` into the mutated vector z using the mutation parameters C_R, F and Algorithm 1;

 Reshape z and `CurrentMember` into $k \times r$ matrices \tilde{Q} and \tilde{U} ;

 Perform a QR decomposition on the matrix \tilde{Q} and assign the resulting $k \times k$ unitary matrix to Q ;

 Delete the last $k - r$ columns of Q and assign this to U ;

if $J(U) < J(\tilde{U})$ **then**

 Reshape U into a $k * r$ dimensional vector \tilde{z} ;

`TemporaryPopi` \leftarrow \tilde{z} ;

end

end

`Pop` \leftarrow `TemporaryPop`;

`counter` \leftarrow `counter` + 1;

end

singular value decomposition (SVD),

$$A = \hat{V} \hat{\Sigma} \hat{W}^\dagger, \quad \hat{U} = \hat{V} \hat{W}^\dagger.$$

In all computational examples, we consistently found that our results were more optimal when we adopted a QR approach over the SVD approach. We discuss further implementation details in the following section alongside our computational examples.

IV. COMPUTATIONAL EXAMPLES

We test our method on four examples of density matrices where the computation of the objective function J in Eq. (7) is done with the entropy of entanglement given by Eq. (4); in this case, J is called the EoF. Although we only report the EoF in our numerical results, it is clear that our methodology can be used for any other pure-state entanglement measure. The first two examples we consider two qubits, and EoF can be calculated using Wootters formula [11]. This helps with benchmarking the accuracy of our method. The third example considers the qubit-environment scenario studied in Ref. [35], but with the qubit initially in a mixed states, so the general trends of entanglement evolution are known, but not the specifics of EoF in particular. The fourth example involves the temperature dependence of a qubit-environment entanglement

at equilibrium for a special class of Hamiltonians for which EoF can be found owing to symmetry as shown in Ref. [20].

Also note that DE requires a choice of parameters F , C_R , and N_{pop} . For this reason, we use the first example as a way to also tune the parameters in the DE algorithm.

A. First example: Decohered Bell-like state

A decohered two-qubit Bell-like state, which is confined to the $\{|01\rangle, |10\rangle\}$ subspace of the full two-qubit Hilbert space $\{|ij\rangle\}$, is the simplest example of a mixed state that can be entangled. It is given by the density matrix

$$\rho_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & b & x & 0 \\ 0 & x^* & 1-b & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (8)$$

for which the convex-roof entanglement measure coincides with its von Neumann entropy. Indeed, the von Neumann entropy can be computed from ρ_1 's eigenvalues

$$\begin{aligned} \lambda_1 &= \frac{1}{2}(1 + \sqrt{4b^2 - 4b + 4|x|^2 + 1}), \\ \lambda_2 &= \frac{1}{2}(1 - \sqrt{4b^2 - 4b + 4|x|^2 + 1}), \\ \lambda_3 &= \lambda_4 = 0. \end{aligned}$$

When $b = x = 1/3$, the von Neumann entropy is thus calculated to be $-\lambda_1 \ln \lambda_1 - \lambda_2 \ln \lambda_2 = 0.381264053728103$ to 15 digits of precision.

Running Algorithm 2 to solve optimization problem (7) for the density matrix ρ_1 , we find that $M(U_*)$, with the underscore $*$ signifying numerical optimality, evaluates correctly to 14 digits. We find this result after 2^{13} iterations with DE parameters $F = 0.1$, $C_R = 0.9$, and $N_{\text{pop}} = 30$. On an average workstation, this computation takes about 100 seconds. With a more manageable 2^{10} iterations, the optimization finds the correct answer to 6 digits in about 12 seconds.

Choosing the DE parameters is done empirically. Our empirical finding is that $F = 0.1$ and $C_R = 0.9$ consistently work well together. In Fig. 1, we show how the optimization converges as the number of iterations increases for different choices of F and C_R . Although our numerical exploration with this hyperparameter study could be done more thoroughly, we find that our result is sufficient in practice. In the future, we commit to using these parameters in numerical computations.

Two more comments are in order about our results. The first is that the convergence in Fig. 1 is not necessarily monotonic with an increasing number of iterations. The reason is that the optimization method is inherently stochastic. The realizations that we show are fairly representative of what to expect when using our method. The second is about the choice of $k \geq r$ defining the feasible space $\mathcal{U}(k, r)$. We empirically find that $k = r$ works sufficiently well here and in most of the remaining examples. This is in opposition to what is reported in [22], where the choice $k = r + 4$ is made. We find that our results are reliably more optimal for this study when $k = r = 2$ than any value of k between r and $r + 8$. As it turns out, this simplifies the methodology slightly as a byproduct, since we now work with square unitary matrices as opposed to rectangular ones.

B. Second example: Sudden death and rebirth of entanglement

In this example, we still restrict the study to the case of two qubits, so the exact value of the EoF can be found without minimization using the Wootters formula [11]. We study the time evolution of the entanglement of an initial state that is separable and mixed, while the entanglement is driven by an interaction between the $|1\rangle$ states of each qubit. The time-dependent density matrix is given by

$$\rho_2(t) = \frac{1}{4} \begin{pmatrix} 1 & c & c & c^2 e^{i\omega t} \\ c & 1 & c^2 & c e^{i\omega t} \\ c & c^2 & 1 & c e^{i\omega t} \\ c^2 e^{-i\omega t} & c e^{-i\omega t} & c e^{-i\omega t} & 1 \end{pmatrix}, \quad (9)$$

where $\omega \in \mathbb{R}$ and $t \in [0, T]$, $T > 0$.

The parameter $c \in [0, 1]$ corresponds to pure single-qubit dephasing channels that act on both qubits [36] on the two-qubit pure state $|\psi\rangle = |++\rangle$, with $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. The evolution is driven by an entangling interaction Hamiltonian $H_{\text{int}} = \hbar\omega|11\rangle\langle 11|$. The evolution of entanglement for such states is known to display sudden death and sudden birth of entanglement for $c < 1$ [37].

In Fig. 2, we show the evolution of the EoF obtained using our minimization method for different values of the dephasing parameter c . We find that not only the continuous, smooth evolution of entanglement is reproduced, but the life-death cycles for the mixed initial state are very clearly visible (both in agreement with the Wootters formula). Thus, the method is capable not only of quantitative description of entanglement present in a system, but can reliably reproduce evolution, including nontrivial features characteristic for entanglement, such as sudden death.

C. Third example: Qubit-environment entanglement evolution

Since the algorithm parameters are set and we have demonstrated that it is capable of reproducing continuous, smooth evolution, as well as sudden death of entanglement, in agreement with the two-qubit formula, which does not require minimization, we now turn to the study of larger systems. In this case, there is no formula for EoF that does not require minimization, but there exist examples of larger systems, for which the calculation of some entanglement measure is relatively straightforward.

One such example involves a qubit interacting with an environment of arbitrary size via an interaction that can only lead to pure dephasing of the qubit. We study entanglement between a qubit and its environment composed of N_e environmental qubits, following the example of Ref. [35]. Contrarily to the results presented in that paper, we will keep the environment pure, while making the initial qubit state mixed, producing a different set of curves.

The qubit is initially in a product state between the qubit and each environmental qubit. The qubit is in a mixed state given by

$$\rho_q(0) = \frac{1}{2} \begin{pmatrix} 1 & d \\ d & 1 \end{pmatrix},$$

which corresponds to an equal superposition state that already underwent some pure dephasing, quantified by the real parameter d . Each environmental qubit (labelled by k) is initially in

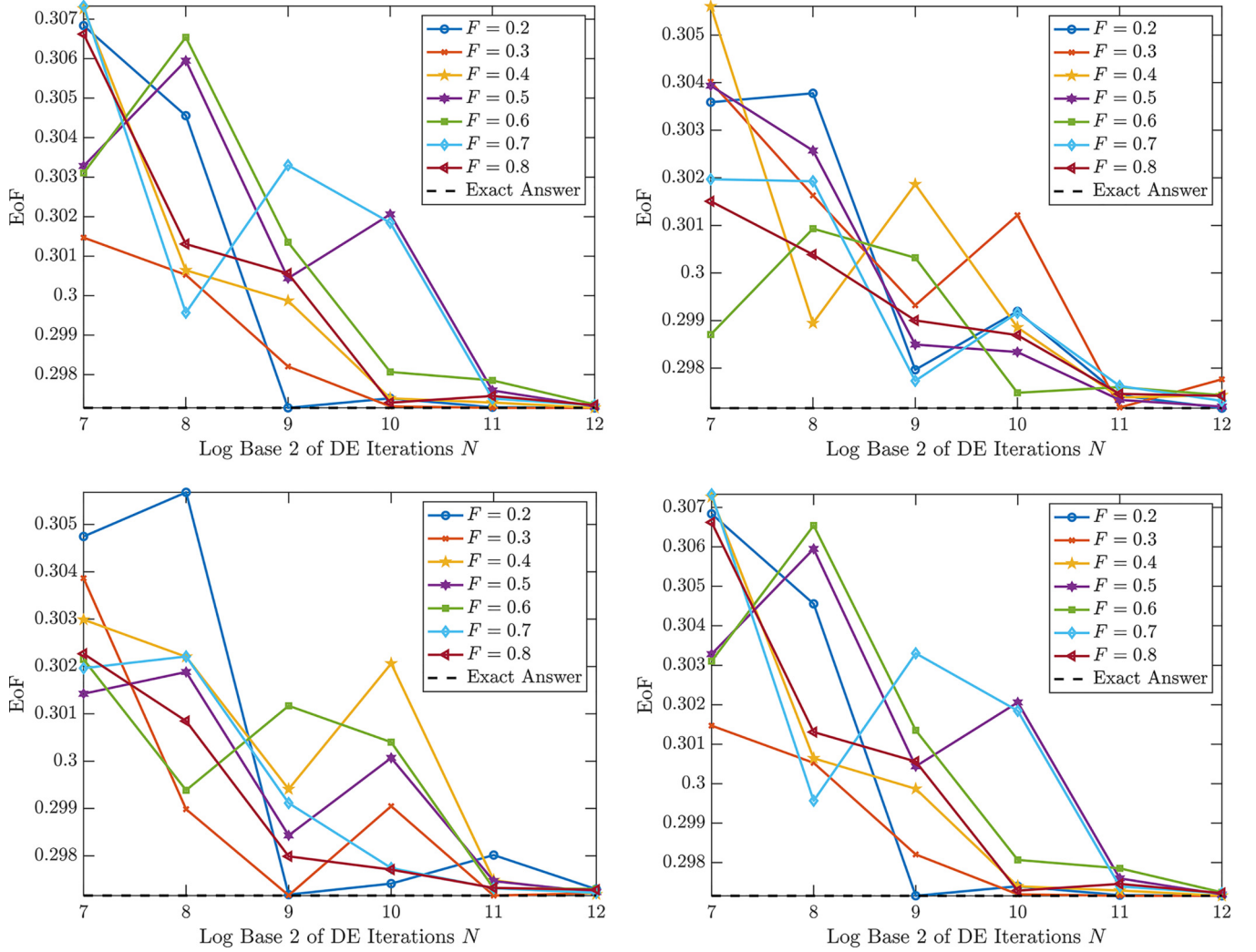


FIG. 1. The impact of varying DE parameters with N_{pop} fixed to 30. In each panel, we show the value of the objective function J after an optimization has been performed on the density matrix ρ_1 from Eq. (8). The top left panel has $C_R = 0.3$, the top right has $C_R = 0.5$, the bottom left has $C_R = 0.7$, and the bottom right has $C_R = 0.9$. We observe that over the many possible choices in the weighting F and the crossover C_R , a small weight and high crossover perform well. The optimal choice in this hyperparameter study, $F = 0.1$ and $C_R = 0.9$, is shown in the bottom right panel.

its ground state $|0\rangle_k$. Thus, the density matrix of the whole environment is

$$\rho_e(0) = \bigotimes_{k=1}^{N_e} |0\rangle_{kk} \langle 0|. \quad (10)$$

The total density matrix of the system is given by the product state

$$\rho_s(0) = \rho_q(0) \otimes \rho_e(0). \quad (11)$$

For pure-dephasing Hamiltonians, it is always possible to write the qubit-environment Hamiltonian in the form [21]

$$H = \sum_{i=0,1} |i\rangle \langle i| \otimes V_i, \quad (12)$$

where the operators V_i act on the space of the environment. Then the evolution operator for the whole system can be

written as

$$U(t) = \sum_{i=0,1} |i\rangle \langle i| \otimes w_i(t), \quad (13)$$

with $w_i(t) = \exp(-iV_i t/\hbar)$. The interpretation of such a process is that the evolution of the environment is conditional on the qubit pointer state, and thus the source of the decoherence is the distinguishability of the qubit state via the state of the environment.

In the following, we omit the free qubit evolution (which has no bearing on entanglement) and the free evolution of the environment, while assuming that no correlations between different environmental qubits is formed during the evolution. For simplicity, we also assume that the interaction is fully asymmetric, which means that $w_0(t) = \mathbb{I}$, which does not have a meaningful bearing on the evolution of the entanglement [38]. Hence, the relevant environmental evolution

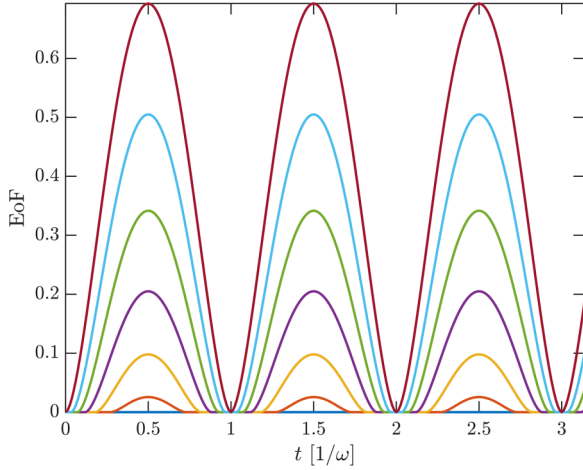


FIG. 2. Evolution of two-qubit entanglement in state (9), for different values of the coherence parameter c . The top curve corresponds to pure-state evolution $c = 1$, which cannot and does not display life/death cycles and reaches the maximum value of unnormalized EoF, $2 \ln 2$. Other curves correspond to mixed states and life/death cycles are observed. The coherence parameter changes from top to bottom as follows: $c = 1$ (red), 0.9 (light blue), 0.8 (green), 0.7 (purple), 0.6 (yellow), 0.5 (orange), and 0.4 (blue).

operator is given by

$$\hat{w}_1(t) = \bigotimes_{k=1}^{N_e} \hat{w}^k(t). \quad (14)$$

Following Ref. [35], we take the evolution operators corresponding to each environmental qubit to be of the form,

$$\hat{w}^k(t) = e^{i\omega_k t} |+\rangle\langle +| + e^{-i\omega_k t} |-\rangle\langle -|, \quad (15)$$

with the states $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and differing phases $\omega_k = \frac{2\pi}{k}$, $k = 1, \dots, N_e$.

The time dependence of entanglement of the qubit-environment density matrix

$$\rho_s(t) = \hat{U}(t)\rho_s(0)\hat{U}^\dagger(t). \quad (16)$$

is shown in Fig. 3. The plotted evolution curves are for environments of two, four, eight, and sixteen qubits, the last of which corresponds to a Hilbert space of dimension 2^{17} for the whole system; thus, the minimization is over an extremely large number of parameters. There is no other way to obtain the EoF curves other than minimization (the presented evolution of entanglement for $d < 1$ is new), but the results for the evolution of the entanglement presented in Ref. [35] allow a qualitative assessment of their validity. The EoF curves we find are smooth and qualitatively in good agreement with the results of [35] where applicable. Moreover, the computation time for a 16 qubit environment is still a manageable 20 minutes for each moment along the time axis. We used parallel computing (24 cores) to maintain the computation of each entanglement curve over 96 moments in time to just under an hour and a half.

D. Fourth example: Temperature dependence of qubit-environment entanglement at equilibrium

Another example of a larger system for which entanglement can be efficiently computed is a density matrix that is block diagonal in disjoint subspaces [20]. Following the example of Ref. [20], we use our algorithm to find the EoF of a thermal equilibrium state corresponding to a Hamiltonian that has this special block form. Thus, the density matrices under study are obtained from the Gibbs state,

$$\rho_{\text{Gibbs}} = \frac{e^{-H/k_B T}}{\text{Tr } e^{-H/k_B T}}. \quad (17)$$

The Hamiltonian that enters Eq. (17) is a qubit-environment Hamiltonian and is given by $H = \sum_m H_m$, where m numbers the blocks. Each block is given by [20]

$$H_m = \begin{pmatrix} E_{m_1} & 0 & 0 & M_m \\ 0 & E_m & 0 & 0 \\ 0 & 0 & E_m & 0 \\ M_m^* & 0 & 0 & E_{m_2} \end{pmatrix}. \quad (18)$$

The critical difference between the blocks is that the subspace of each block is disjoint. By this we mean that in the effectively two-qubit subspace of each block, the two states corresponding to the qubit remain unchanged, but the states corresponding to the environment are different in each block (form a separate subspace). Thus, the matrix for a single block of the Hamiltonian (18) is written in the bases' subspace $\{|0m\rangle, |0m_\perp\rangle, |1m\rangle, |1m_\perp\rangle\}$, where $\langle m|m_\perp\rangle = 0$, but also all states that occupy different subspaces are orthogonal to each other. This requirement is stronger than for the Hamiltonian to simply have block-diagonal form, but it allows for convex-roof entanglement measures to be averaged over different blocks without the necessity of minimization beyond each block [20]. This gives us the opportunity to compare the results obtained by full minimization performed by our algorithm with the results of [20], which took advantage of the special feature.

The parameters within each block of the Hamiltonian (18) are given by

$$E_{m_1} = \alpha \left(m + \frac{\Omega}{2} \right), \quad (19a)$$

$$E_{m_2} = -\alpha \left(m + 1 + \frac{\Omega}{2} \right), \quad (19b)$$

$$E_m = \frac{1}{2}(E_{m_1} + E_{m_2}), \quad (19c)$$

$$M_m = \alpha \sqrt{K(K+1) - m(m+1)}, \quad (19d)$$

where $m = 0, \pm 1, \pm 2, \dots, \pm K$. The parameter α (which we set to 1) dictates the coupling strength, while Ω is responsible for the energy splittings within the system.

The temperature dependence of the equilibrium entanglement between a qubit and an environment with $K = 1$, which corresponds to a Hamiltonian with three blocks, is shown in Fig. 4. Note that a three-block environment translates to six environmental states, so we use the algorithm to quantify entanglement within a twelve-dimensional Hilbert space.

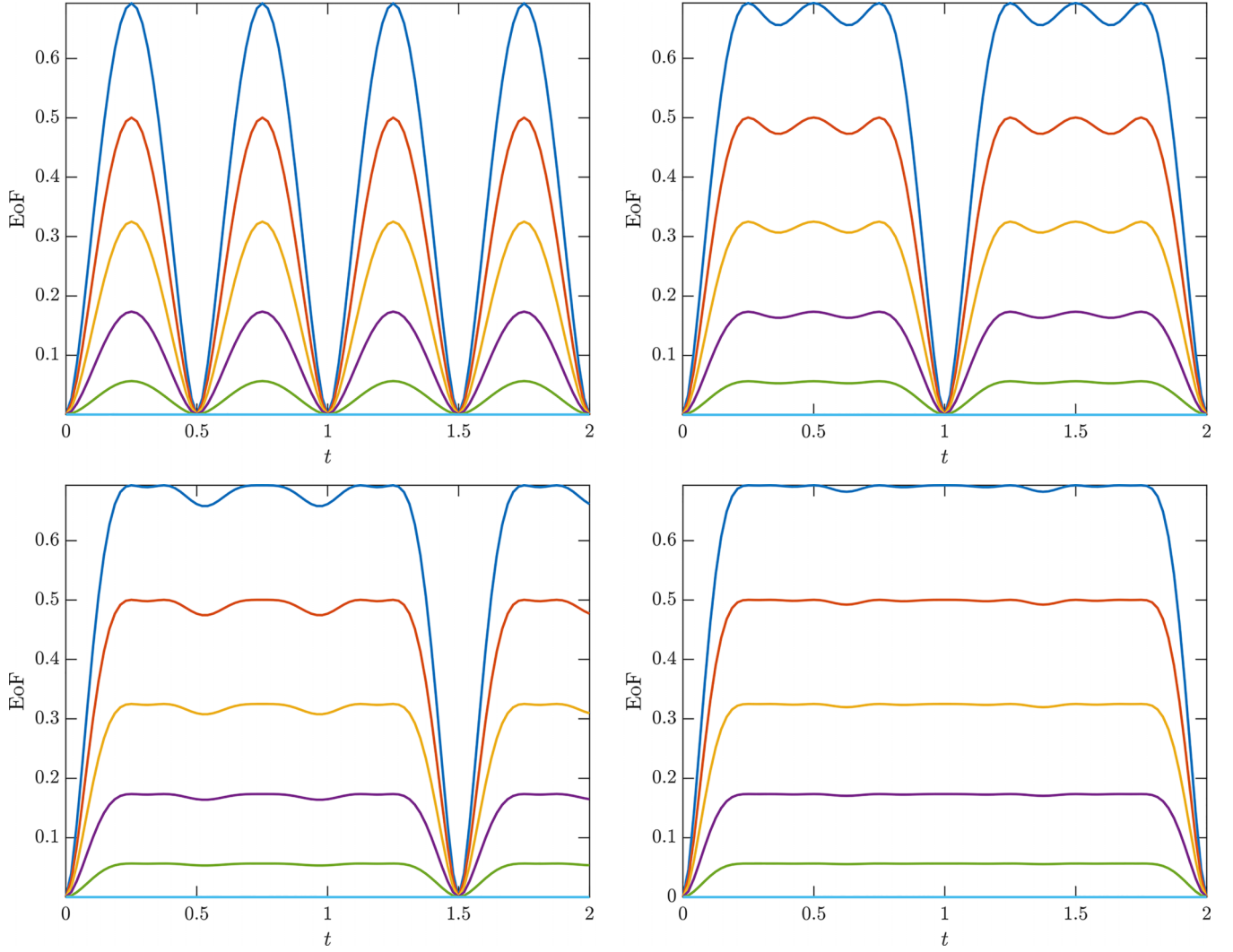


FIG. 3. Evolution of entanglement for density matrix (16) for an environment of (top left) 2, (top right) 4, (bottom left) 8, and (bottom right) 16 qubits for different purity of the initial qubit state. The qubit coherence parameter changes from top to bottom with $d = 1$ (blue), 0.8 (orange), 0.6 (yellow), 0.4 (purple), 0.2 (green), 0 (light blue).

For larger values of Ω (top panel), which roughly correspond to large magnetic fields in spin systems, the resulting curves are smooth and in agreement with corresponding results of Ref. ([20]) up to reasonably high temperatures (but not for the whole scope of non-negligible entanglement). The EoF as a function of the temperature of $\Omega = 1$, is plotted for a wider range of temperatures in the bottom panel. At low temperatures, the curve obtained is smooth and does not differ qualitatively from the higher Ω curves. At larger temperatures, which exceed the temperature at which maximum entanglement is reached, convergence cannot be reached in a computation time of a whole day. Despite this, the general trend is still visible and it agrees very well with the results of Ref. [20].

V. CONCLUSIONS

We developed a method for the efficient computation of mixed-state convex-roof entanglement measures. Such

measures require minimization over all pure-state decompositions of the density matrix. Operationally, the task requires minimization over a set of unitary matrices of the same dimension as the Hilbert space of the problem under study, N , and the number of minimization parameters scales as $N^2 - 1$. Our method thoroughly searches this space using a nonconvex method that remains constrained to the manifold of unitary matrices via a QR factorization. The convergence to the nearest local minimum is then accelerated by using a BFGS method.

We used the developed algorithm on a number of examples. The examples were chosen in such a way that at least the trends of entanglement behavior would be known. Furthermore, the studied scenarios involved smooth or partially smooth parameter dependence (time or temperature) of entanglement in order to provide a better test of the reproducibility and reliability of the outcomes. We have found that the algorithm is extremely reliable (and fast) in identifying the expected trends as long as the overall purity of the state

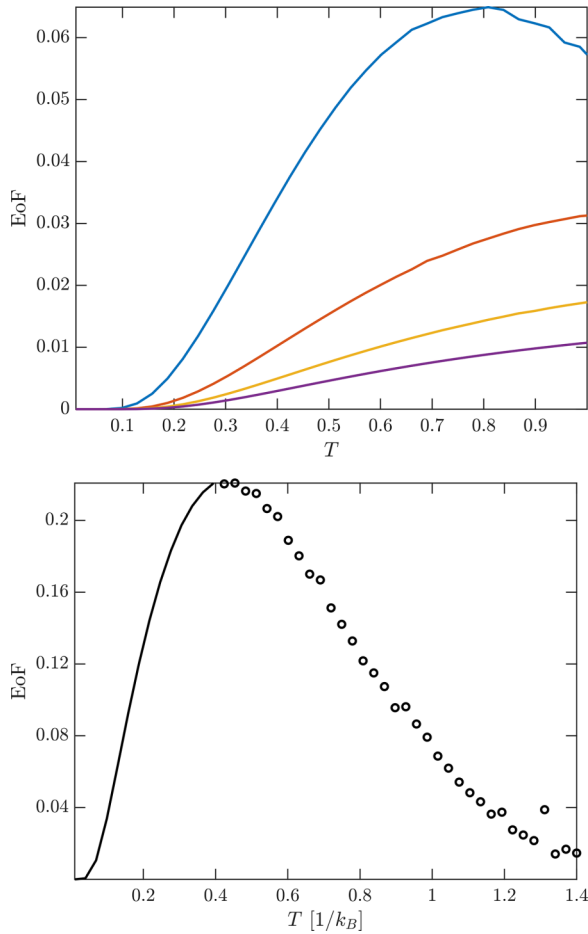


FIG. 4. Temperature dependence of qubit-environment entanglement at equilibrium for a system with three blocks, $K = 1$. The top panel shows entanglement curves for different Ω (5, 10, 15, and 20 from bottom to top), which corresponds to different applied magnetic fields in a spin system. The bottom panel shows a computation for a wider range of temperatures for $\Omega = 1$. We use points at higher temperatures, when the algorithm no longer produces a smooth temperature dependence.

is fairly high. This holds true for very large systems (in the third example we produce entanglement evolution curves for a Hilbert state of dimension 2^{17}).

The algorithm is unable to quantify the EoF with high precision at low purity, as seen at high temperatures in the fourth example. This offers two important opportunities for future work. The first is mathematical and is to explore why the algorithm is unable to thoroughly search the nonconvex landscape at low purities. We conjecture that, loosely speaking, this is caused by an increasing nonconvexity as the purity decreases and thus makes the identification of the global optimum unlikely.

The second direction is computational. As can be gleaned from the presentation of our methodology, any black box numerical optimization method can be used, in principle, if the QR factorization is deployed within the objective function evaluation. This offers an opportunity to explore the numerical optimization methods typically used by the machine learning community, such as adaptive moment estimation [39], that address high-dimensional nonconvex optimization problems.

ACKNOWLEDGMENTS

We appreciate the help of Alejandro Aceves in preparing this manuscript for publication. J.A. also acknowledges support from NSF Award No. 2316622. K.R. is funded within the QuantERA II Programme that has received funding from the EU H2020 research and innovation programme under Grant Agreement No. 101017733, and with funding organization MEYS.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available because they are owned by a third party and the terms of use prevent public distribution. The data are available from the authors upon reasonable request.

- [1] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, *Phys. Rev. A* **53**, 2046 (1996).
- [2] E. Santos and M. Ferrero, Linear entropy and Bell inequalities, *Phys. Rev. A* **62**, 024101 (2000).
- [3] S. Touzard, A. Kou, N. E. Frattini, V. V. Sivak, S. Puri, A. Grimm, L. Frunzio, S. Shankar, and M. H. Devoret, Gated conditional displacement readout of superconducting qubits, *Phys. Rev. Lett.* **122**, 080502 (2019).
- [4] O. Katz and C. Monroe, Programmable quantum simulations of bosonic systems with trapped ions, *Phys. Rev. Lett.* **131**, 033604 (2023).
- [5] E. J. Connors, J. J. Nelson, H. Qiao, L. F. Edge, and J. M. Nichol, Low-frequency charge noise in Si/SiGe quantum dots, *Phys. Rev. B* **100**, 165305 (2019).
- [6] K. Modi, A pedagogical overview of quantum discord, *Open Syst. Inf. Dyn.* **21**, 1440006 (2014).
- [7] M. A. Nielsen, Conditions for a class of entanglement transformations, *Phys. Rev. Lett.* **83**, 436 (1999).
- [8] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Quantum nonlocality without entanglement, *Phys. Rev. A* **59**, 1070 (1999).
- [9] P. Mazurek, K. Roszak, R. W. Chhajlany, and P. Horodecki, Sensitivity of entanglement decay of quantum-dot spin qubits to the external magnetic field, *Phys. Rev. A* **89**, 062318 (2014).
- [10] T. Hiroshima and S. Ishizaka, Local and nonlocal properties of Werner states, *Phys. Rev. A* **62**, 044302 (2000).
- [11] W. K. Wootters, Entanglement of formation of an arbitrary state of two qubits, *Phys. Rev. Lett.* **80**, 2245 (1998).
- [12] H. Azuma and M. Ban, Another convex combination of product states for the separable Werner state, *Phys. Rev. A* **73**, 032315 (2006).

- [13] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824 (1996).
- [14] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of noisy entanglement and faithful teleportation via noisy channels, *Phys. Rev. Lett.* **76**, 722 (1996).
- [15] M. B. Plenio, Logarithmic negativity: A full entanglement monotone that is not convex, *Phys. Rev. Lett.* **95**, 090503 (2005).
- [16] M. Horodecki, P. Horodecki, and R. Horodecki, Mixed-state entanglement and distillation: Is there a “bound” entanglement in nature? *Phys. Rev. Lett.* **80**, 5239 (1998).
- [17] K. Audenaert, F. Verstraete, and B. De Moor, Variational characterizations of separability and entanglement of formation, *Phys. Rev. A* **64**, 052304 (2001).
- [18] M. M. Wolf, G. Giedke, O. Krüger, R. F. Werner, and J. I. Cirac, Gaussian entanglement of formation, *Phys. Rev. A* **69**, 052320 (2004).
- [19] K. Shiokawa, Non-Markovian dynamics, nonlocality, and entanglement in quantum Brownian motion, *Phys. Rev. A* **79**, 012308 (2009).
- [20] M. Jędrzejewski, K. Kinastowski, and K. Roszak, Convex-roof entanglement measures of density matrices block diagonal in disjoint subspaces for the study of thermal states, *Phys. Rev. A* **105**, 062419 (2022).
- [21] K. Roszak, Criteria for system-environment entanglement generation for systems of any size in pure-dephasing evolutions, *Phys. Rev. A* **98**, 052344 (2018).
- [22] B. Röthlisberger, J. Lehmann, and D. Loss, Numerical evaluation of convex-roof entanglement measures with applications to spin rings, *Phys. Rev. A* **80**, 042301 (2009).
- [23] B. Röthlisberger, J. Lehmann, and D. Loss, libCreme: An optimization library for evaluating convex-roof entanglement measures, *Comput. Phys. Commun.* **183**, 155 (2012).
- [24] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds* (Princeton University Press, Princeton, NJ, 2008).
- [25] L. Aolita, A. Buchleitner, and F. Mintert, Scalable method to estimate experimentally the entanglement of multipartite systems, *Phys. Rev. A* **78**, 022308 (2008).
- [26] G. Vidal, Entanglement monotones, *J. Mod. Opt.* **47**, 355 (2000).
- [27] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, Everything you always wanted to know about LOCC (but were afraid to ask), *Commun. Math. Phys.* **328**, 303 (2014).
- [28] E. Schrödinger, Discussion of probability relations between separated systems, in *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 31 (Cambridge University Press, Cambridge, 1935), pp. 555–563.
- [29] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, *Phys. Lett. A* **183**, 14 (1993).
- [30] R. Storn and K. Price, Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces, *J. Global Optim.* **11**, 341 (1997).
- [31] K. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization* (Springer-Verlag, Berlin, 2018).
- [32] L. N. Trefethen and D. Bau, *Numerical Linear Algebra* (SIAM, Philadelphia, 2022).
- [33] J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, New York, 1999).
- [34] K. Fan and A. J. Hoffman, Some metric inequalities in the space of matrices, *Proc. Am. Math. Soc.* **6**, 111 (1955).
- [35] K. Roszak, Measure of qubit-environment entanglement for pure dephasing evolutions, *Phys. Rev. Res.* **2**, 043062 (2020).
- [36] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [37] K. Roszak and P. Machnikowski, Complete disentanglement by partial pure dephasing, *Phys. Rev. A* **73**, 022313 (2006).
- [38] K. Roszak and L. Cywiński, Characterization and measurement of qubit-environment-entanglement generation during pure dephasing, *Phys. Rev. A* **92**, 032310 (2015).
- [39] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).