Entanglement of Formation for Isotropic States

Barbara M. Terhal[1,*] and Karl Gerd H. Vollbrecht[2,†]

[1]*IBM Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598*
[2]*Institut für Mathematische Physik, TU Braunschweig, Mendelssohnstrasse 3, 38106 Braunschweig, Germany*
(Received 24 May 2000)

We give an explicit expression for the entanglement of formation for isotropic density matrices in arbitrary dimensions in terms of the convex hull of a simple function. For two qutrit isotropic states we determine the convex hull and we have strong evidence for its exact form for arbitrary dimension. Unlike for two qubits, the entanglement of formation for two qutrits or more is found to be a nonanalytic function of the maximally entangled fraction in the regime where the density matrix is entangled.

One of the main goals in quantum information theory is to develop a theory of entanglement. A cornerstone of this theory will be a good measure of bipartite entanglement. Such a measure must obey the essential property that the entanglement of a bipartite density matrix ρ which is shared by Alice and Bob cannot increase, on average, under local quantum operations and classical communication ($LO + CC$) between Alice and Bob. In this way, the entanglement captures the truly quantum correlations in a bipartite density matrix. For pure bipartite states a good measure of entanglement has been found, it is the following quantity:

$$E(|\psi\rangle\langle\psi|) = S(\text{Tr}_B(|\psi\rangle\langle\psi|)), \tag{1}$$

where $S(\rho)$ is the von Neumann entropy of ρ; i.e., $S(\rho) = -\text{Tr}\rho \log\rho$ and $\text{Tr}_B(|\psi\rangle\langle\psi|)$ is the reduced density matrix that we obtain by tracing out over Bob's quantum system. This measure E is unique [1,2] if one requires the entanglement to obey a set of natural properties such as convexity, nonincrease under local measurements, asymptotic continuity, partial additivity, and normalization. Moreover, E is a measure of the asymptotic entanglement costs [3] of making the state $|\psi\rangle$ out of a canonical set of states, which we can choose to be EPR singlets $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$, which have $E = 1$. This process is reversible, in the sense that one can concentrate [3] a set of n states $|\psi\rangle$ with entanglement E to a smaller set of $m = En$ EPR singlets.

The situation for mixed states is much more complex. In Ref. [4] a first measure of mixed state entanglement, called the entanglement of formation, was introduced. This measure is a candidate for measuring the asymptotic costs of making the density matrix out of a supply of EPR singlets. There are no mixed density matrices for which this statement has been proved, but neither have counterexamples been found so far. The search for a possible discrepancy between the entanglement of formation and the asymptotic entanglement costs is hampered by the fact that we know the entanglement of formation only for two qubit systems; Wootters [5] found an analytic expression for the entanglement of formation for all two qubit density matrices.

In this Letter we present the first calculation of the entanglement of formation of a class of density matrices in dimensions higher than $\mathbf{C}^2 \otimes \mathbf{C}^2$. We explicitly determine the entanglement of formation for two qutrit density matrices in this class, and we find an expression in arbitrary dimension in terms of the convex hull of a simple function. We conjecture the explicit form of this convex hull, which can be easily verified in a given dimension. Surprisingly, the entanglement of formation is found to be a nonanalytic function of the parameter characterizing the class of states that we consider.

Let us start by recalling the definition of the entanglement of formation. Let $\mathcal{E}_\rho = \{p_i, |\psi_i\rangle\}$ be an ensemble of pure states which form a decomposition of $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. The entanglement of formation for mixed states ρ is defined as

$$E(\rho) = \min_{\mathcal{E}=\{p_i, |\psi_i\rangle\}} \sum_i p_i E(|\psi_i\rangle\langle\psi_i|). \tag{2}$$

In this Letter we will consider the class of density matrices, sometimes called isotropic density matrices, which are convex mixtures of a maximally entangled state and the maximally mixed state:

$$\rho_F = \frac{1 - F}{d^2 - 1}(\mathbf{1} - |\Psi^+\rangle\langle\Psi^+|) + F|\Psi^+\rangle\langle\Psi^+|, \tag{3}$$

for $0 \leq F \leq 1$ and $|\Psi^+\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^d |ii\rangle$. For $F \leq 1/d$ these density matrices are separable [6]. The entanglement of formation for states with $d = 2$ is equal to [4]

$$H_2(\mu), \quad \mu = \frac{1}{2} + \sqrt{F(1 - F)}, \tag{4}$$

where $H_2(.)$ is the binary entropy function. The states ρ_F have the important property [6] that they are invariant under the operation $U \otimes U^*$ for any unitary transformation U. The $LO + CC$ "twirling" superoperator $S^{U\otimes U^*}$ is defined as

$$S^{U\otimes U^*}(\rho) = \frac{1}{\text{Vol}(U)} \int dU\, U \otimes U^* \rho U^\dagger \otimes U^{*\dagger}. \tag{5}$$

In Ref. [7] the Schmidt number of the isotropic states was determined. Instead of making an isotropic state

out of a set of maximally entangled states, we ask how to construct an isotropic state with a given F out of some state characterized by a Schmidt vector $\vec{\mu}$. So, let us

take an arbitrary initial pure state $|\psi\rangle = \sum_{i=1}^{d} \sqrt{\mu_i} |a_i, b_i\rangle$ and consider the effect of twirling. We can write $|\psi\rangle = U_A \otimes U_B \sum_i \sqrt{\mu_i} |i, i\rangle$ and thus

$$S^{U \otimes U^*}\left(\sum_{i,j} \sqrt{\mu_i \mu_j} |a_i, b_i\rangle\langle a_j, b_j|\right) = S^{U \otimes U^*}\left((\mathbf{1} \otimes V) \sum_{i,j} \sqrt{\mu_i \mu_j} |i, i\rangle\langle j, j| (\mathbf{1} \otimes V^\dagger)\right), \quad (6)$$

where $V = U_A^T U_B$. We define $v_{ij} = \langle i|V|j\rangle$. The twirled state becomes

$$S^{U \otimes U^*}(|\psi\rangle\langle\psi|) = \frac{|\sum_i \nu_i|^2}{d} P_+$$
$$+ \frac{1 - |\sum_i \nu_i|^2/d}{d^2 - 1} (\mathbf{1} - P_+), \quad (7)$$

where $\nu_i = \sqrt{\mu_i} v_{ii}$ and $P_+ = |\Psi^+\rangle\langle\Psi^+|$. When we choose $V = \mathbf{1}$ we find the density matrix ρ_F at $F = [\sum_i \sqrt{\mu_i}]^2/d$. For general V one can bound

$$\left|\sum_i \nu_i\right|^2 \le \left[\sum_i |\nu_i|\right]^2 \le \left[\sum_i \sqrt{\mu_i}\right]^2, \quad (8)$$

since $|v_{ii}| \le 1$ for all i. Thus the largest value for F is obtained by choosing the initial state $\sum_{i=1}^{d} \sqrt{\mu_i} |ii\rangle$.

The use of symmetry makes it possible to give a simplified expression for the entanglement of formation for isotropic states:

Lemma 1.—The entanglement of formation for isotropic states in $\mathbf{C}^d \otimes \mathbf{C}^d$ ($d \ge 2$) for $F \in (1/d, 1]$ is given by

$$E(\rho_F) = \mathrm{co}(R(F)), \quad (9)$$

where $\mathrm{co}(g)$ denotes the convex hull of the function g and $R(F)$ is defined as

$$R(F) = \min_{\vec{\mu}}\left\{H(\vec{\mu}) \Big| F = \left[\sum_{i=1}^{d} \sqrt{\mu_i}\right]^2 \Big/ d\right\}, \quad (10)$$

and $\vec{\mu}$ is a Schmidt vector.

Proof.—Assume that there exists an optimal decomposition of ρ_F formed by the ensemble $\{p_i, |\psi_i(\vec{\mu}^i)\rangle\}$, where $\vec{\mu}^i$ denotes the Schmidt vector of the state $|\psi_i\rangle$. By twirling the left-hand side and right-hand side of the equation $\rho_F = \sum p_i |\psi_i\rangle\langle\psi_i|$ we obtain that $\rho_F = \sum_i p_i \rho_{F_i} = \rho_{\sum_i p_i F_i}$ where

$$F_i(V_i, \vec{\mu}^i) = \frac{|\sum_k v_{kk} \sqrt{\mu_k^i}|^2}{d}, \quad (11)$$

as in Eq. (7). Since the decomposition is optimal, each Schmidt vector $\vec{\mu}^i$ has minimal entropy under this constraint. Consider the function

$$R_V(F) = \min_{\vec{\mu}}\left\{H(\vec{\mu}) \Big| F = \left|\sum_{i=1}^{d} v_{ii} \sqrt{\mu_i}\right|^2 \Big/ d\right\}. \quad (12)$$

An optimal decomposition of ρ_F is a convex combination of pure states each of which corresponds to a certain F under twirling. Thus the entanglement of formation $E(\rho_F)$

can be obtained by taking the convex hull of the functions $\mathrm{co}(R_V(F))$. We can make an additional simplification. Equation (8) implies that $R_V(F) = R_{\mathbf{1}}(F') \equiv R(F')$ where $F' \ge F$ for every V. Thus instead of taking the convex hull of all functions $\mathrm{co}(R_V(F))$, we can take the convex hull of the function $R_+(F) = \min_x\{R(x)|x \ge F\}$. In Lemma 2 we will determine $R(F)$ and it is not hard to show that $R(F)$ is a monotonically increasing function of F. It follows then that $R_+(F) = R(F)$ and $\mathrm{co}(R_+(F)) = \mathrm{co}(R(F))$. ∎

We now determine the function $R(F)$ defined in Eq. (10). Since all the equations are symmetric in μ_i, we can restrict ourselves to solutions which satisfy $\mu_1 \ge \mu_2 \ge \cdots \ge \mu_d$. With the method of Lagrange multipliers we get a necessary condition for the minimum

$$-1 - \log\mu_i + \Lambda_1 + \frac{\Lambda_2}{2}\mu_i^{-1/2} = 0, \quad (13)$$

where Λ_1, Λ_2 denote the Lagrange multipliers. For fixed Λ_1, Λ_2 this determines the whole set $\{\mu_i\}$. Setting $\mu_i = \frac{1}{q_i^2}$ we obtain an expression of the form $\log q_i = A q_i + B$, where A, B depend only on Λ_1, Λ_2. Since a convex and a concave function cross each other in at most two points, this equation has maximally two possible nonzero solutions for q_i. Therefore all Schmidt vectors $\vec{\mu}$ that are possible candidates for the minimum have to satisfy the condition $\mu_i \in \{\gamma, \delta, 0\}$. Let n be the number of entries where $\mu_i = \gamma$ and m the number of entries where $\mu_i = \delta$. The minimization problem has been reduced considerably: For fixed n, m, $n + m \le d$, we minimize the function

$$nh(\gamma) + mh(\delta), \quad (14)$$

where $h(x) = -x \log x$, under the constraints

$$\{n\gamma + m\delta = 1, n\sqrt{\gamma} + m\sqrt{\delta} = \sqrt{dF}\}. \quad (15)$$

The constraints give rise to a quadratic equation in $\sqrt{\gamma}$ which provides two possible solutions for γ for every choice of n, m:

$$\gamma_{nm}^{\pm}(F) = \left[\frac{\sqrt{dF}n \pm \sqrt{mn(m + n - dF)}}{n(n + m)}\right]^2. \quad (16)$$

With the first constraint we get the corresponding $\delta_{nm}^{\pm}(F) = [1 - n\gamma_{nm}^{\pm}(F)]/m$. Since $\gamma_{mn}^- = \delta_{nm}^+$, the function in Eq. (14) takes the same value for γ_{nm}^+ and γ_{mn}^-. Therefore we can restrict ourselves to the solutions $\gamma_{nm} := \gamma_{nm}^+$. The pointwise minimum over all possible choices for n, m of

$$R_{nm}(F) = H_2(n\gamma_{nm}) + n\gamma_{nm}\log\frac{n}{m} + \log m, \quad (17)$$

defined on the domain $\frac{n}{d} \leq F \leq \frac{n+m}{d}$, is the required function $R(F)$. The restriction on the domain comes from requiring that γ_{nm} is a proper solution of Eq. (16) which implies that $F \leq \frac{n+m}{d}$. On the other hand we demand that $\delta_{nm} \geq 0$ which implies that $F \geq \frac{n}{d}$. In this regime one can verify that $\gamma_{nm} \geq \delta_{nm}$.

When $m = 0$, γ and F are uniquely determined by the constraints, i.e., $F = \frac{n}{d}$. Since $R_{n0}(\frac{n}{d}) = R_{n'm'}(\frac{n}{d})$ for all $n' + m' = n$, we can neglect these cases.

When $d = 3$, what remains is a minimization over the three functions $R_{12}(F)$, $R_{21}(F)$, and $R_{11}(F)$, which are plotted in Fig. 1. For $d = 3$ we get $R(F) = R_{12}(F)$ [8]. Thus the optimal vector $\vec{\mu}$ is always of the form

$$\vec{\mu} = \{\gamma, \delta, \delta\}, \quad (18)$$

satisfying $\gamma \geq \delta$.

The case $d = 3$ is the important one, since it turns out that we can relate all the higher dimensional minima to $d = 3$ and prove that

Lemma 2.—For $d \geq 3$ the function $R(F) = R_{1,d-1}(F)$.

Proof.—The case $d = 3$ is discussed above. Note that $R_{1,d-1}(1/d) = 0$, which is clearly minimal, so we provide a proof for $F > 1/d$. Let the minimum be attained in $d > 3$ dimensions by a vector $\vec{\mu} = \{\mu_i\}$. Let us select some subset of the entries of $\vec{\mu}$, the set $\{\mu_{i_j}\}_{j=1}^{d'}$, where $\sum_{j=1}^{d'}\mu_{i_j} = k \leq 1$. Since $\vec{\mu}$ is the minimum, it follows that the set $\{\mu_{i_j}\}_{j=1}^{d'}$ is the minimum when we keep the other entries of the vector $\vec{\mu}$ fixed. Let $\mu'_j := \frac{\mu_{i_j}}{k}$. The vector $\vec{\mu'}$ is the solution for the minimization of

$$\sum_{i=1}^{d'} h(\mu'_i k) = k\left[\sum_{i=1}^{d'} h(\mu'_i)\right] + h(k), \quad (19)$$

under the constraints $\sum_{j=1}^{d'}\mu'_j = 1$ and $\sum_{j=1}^{d'}\sqrt{\mu'_j} = C$, where

$$C = \sqrt{\frac{dF}{k}} - \frac{1}{\sqrt{k}}\sum_{i|\forall j, i\neq i_j}\sqrt{\mu_i}. \quad (20)$$

This last equation can always be written as $C = \sqrt{d'F'}$ for some F'. Thus the restricted minimization problem is equivalent to a d'-dimensional version of the original problem, up to the scaling factor k and the additive term $h(k)$. When $F' \leq \frac{1}{d'}$, we know that solution of this minimization problem is given by a Schmidt vector $\vec{\mu'}$ which corresponds to an unentangled state; i.e., it is of the form $\vec{\mu'} = \{1, 0, \ldots, 0\}$. Let us choose three arbitrary μ_i out of the optimal vector $\vec{\mu}$. When the resulting $F' \leq \frac{1}{3}$, it follows that $\vec{\mu'} = \{1, 0, 0\}$. When $F' > \frac{1}{3}$, the three entries of $\vec{\mu'}$ have to satisfy Eq. (18). So in fact, in both cases they satisfy Eq. (18). Suppose now that one entry of $\vec{\mu}$ is equal to zero. Then it follows that $\vec{\mu}$ cannot have two nonzero entries since this would violate condition (18); in other words it must be that $\vec{\mu} = \{1, 0, \ldots, 0\}$. But this is a solution for $F = 1/d$. Therefore we get $n + m = d$ for $F > 1/d$. Suppose that $n \geq 2$. We can choose the vector $\{\gamma, \gamma, \delta\}$ satisfying $\gamma \geq \delta$. Then condition (18) implies that $\gamma = \delta$. This implies that all entries of $\vec{\mu}$ are identical, or $\vec{\mu} = \{\frac{1}{d}, \ldots, \frac{1}{d}\}$. This corresponds to a maximally entangled state, which is the unique solution for $F = 1$. Therefore $n = 1$ and $m = d - 1$. ∎

Lemma 1 and Lemma 2 together result in the following:

Theorem 1.—The entanglement of formation $E(\rho_F)$ for isotropic states in $\mathbf{C}^d \otimes \mathbf{C}^d$ ($d \geq 2$) for $F \in (1/d, 1)$ is given by

$$E(\rho_F) = \text{co}(R_{1,d-1}(F)), \quad (21)$$

where

$$R_{1,d-1}(F) = H_2(\gamma(F)) + [1 - \gamma(F)]\log(d - 1), \quad (22)$$

with

$$\gamma(F) = \frac{1}{d}\left[\sqrt{F} + \sqrt{(d - 1)(1 - F)}\right]^2. \quad (23)$$

For $d = 3$ the first and second derivative of the function $R_{12}(F)$ are plotted in Fig. 2. The figure shows that the
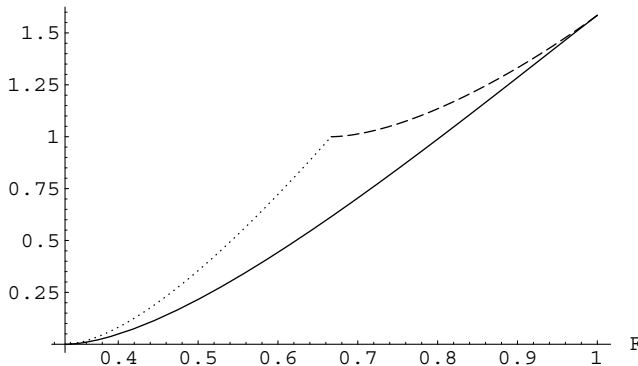

FIG. 1.   The $R_{nm}(F)$ functions for $d = 3$. The solid line is $R_{12}(F)$, which is the minimal one. The dotted and dashed lines are $R_{11}(F)$ and $R_{21}(F)$, respectively.
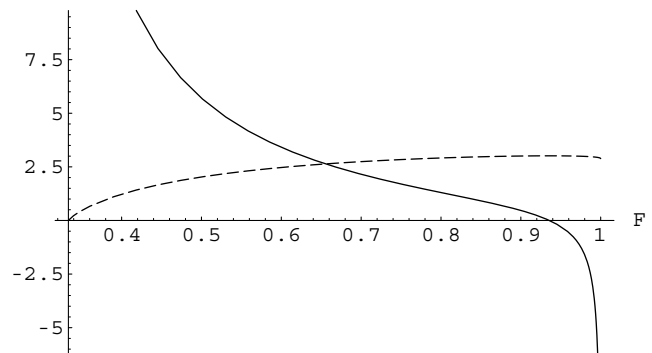

FIG. 2.   The first and second derivative of $R_{1,d-1}(F)$ for $d = 3$. The solid line is the second derivative, which is going to $-\infty$ for $F = 1$. The dashed line is the first derivative.

function $R_{12}(F)$ is not convex near $F = 1$; its second derivative is not positive. In order to determine $\mathrm{co}(R_{12}(F))$ for $d = 3$ we solve the following equations: Let $E_{\text{line}}(F) = aF + \log 3 - a$ be the line crossing through the point $(1, \log 3)$. We solve (i) $E_{\text{line}}(F) = R_{12}(F)$ and (ii) $\frac{dE_{\text{line}}}{dF} = a = \frac{dR_{12}}{dF}$ for a and F. Figure 2 indicates

$$E(\rho_F) = \begin{cases} 0, & F \le \frac{1}{d}, \\ R_{1,d-1}(F), & F \in [\frac{1}{d}, \frac{4(d-1)}{d^2}], \\ \frac{d \log(d-1)}{d-2}(F - 1) + \log d, & F \in [\frac{4(d-1)}{d^2}, 1]. \end{cases} \tag{24}$$

The correctness of this solution can easily be verified for a given d by plotting the function $R_{1,d-1}(F)$ and its second derivative and noting that the convex hull of $R_{1,d-1}(F)$ is obtained by calculating where $R_{1,d-1}(F)$ meets the line going through the point $(F = 1, E = \log d)$ and $R'_{1,d-1}(F)$ equals the slope of this line.

It is surprising to find that $E(\rho_F)$ is nonanalytic in the region where ρ_F is an entangled density matrix. Another feature of our solution is that for, say, $d = 3$ and $F > 8/9$ an optimal decomposition of ρ_F is not one in which every pure state has an equal amount of entanglement. Indeed, the optimal decomposition that we find, is a mixture of the maximally entangled state and the ensemble of states $|\psi\rangle$ obtained by twirling, each of which has entanglement $E = -1/3 + \log 3$. Since every state in the optimal decomposition of ρ_F has, under twirling, a value of entanglement on $R_{1,d-1}(F)$, every optimal decomposition of ρ_F for $d = 3$ in the range $F > 8/9$ will be a mixture of the maximally entangled state and some less entangled states. This is in contrast with optimal decomposition for E for two qubits. For $F > 8/9$ more than $d^2 = 9$ pure states must be used in the optimal decomposition of ρ_F. We make ρ_F from a maximally entangled state and the state $\rho_{F=8/9}$ which has rank 9, and thus needs at least 9 states in its optimal decomposition. In total, this gives 10 states. For $F > 8/9$ there is no optimal decomposition with fewer states: one always has to mix in the maximally entangled state with some probability. The remaining state ρ'_F either has rank 9 (like $\rho_{F=8/9}$) or a lower rank. If it has a lower rank, it must be separable, which would imply that the optimal decomposition is made from mixing a separable state with a maximally entangled state which we know to be false. This is the first example of an entangled state for which it is proved that the number of pure states in the optimal decomposition exceeds the rank of the state (see Ref. [10] for separable states with this property).

Crucial in our method is the invariance of the isotropic states under a symmetry group of local operations. A result similar to Lemma 2 will hold, for example, for the class of Werner states [9] which are invariant under the transformation $U \otimes U$ for all $U \in U(d)$. Let $S^{U \otimes U}$ be defined as in Eq. (5) but with the omission of the complex conjugation. The Werner states ρ_p^W are characterized by a single parameter p. One can prove that $E(\rho_p^W) = \mathrm{co}(R(p))$ where

that $R_{12}(F)$ is monotonically increasing and that there is only one region where $R_{12}(F)$ is not convex, namely near $F = 1$. Therefore the solution to the equations will be unique: we find that $F = 8/9$ and $a = 3$. For higher dimensions, we conjecture, based on examining these two equations, that the entanglement of formation in $\mathbf{C}^d \otimes \mathbf{C}^d$ is given by

$$R(p) = \min\{E(|\psi\rangle\langle\psi|) \mid S^{U \otimes U}(|\psi\rangle\langle\psi|) = \rho_p^W\}. \tag{25}$$

It may thus be possible to carry out a similar analysis as was done here for the Werner states. In a further generalization one could consider the entanglement of formation for $g \otimes g$ or $g \otimes g^*$ invariant states where $g \in G$ and G is a subgroup of $U(d)$.

---

*Email address: terhal@watson.ibm.com
†Email address: k.vollbrecht@tu-bs.de

[1] G. Vidal, J. Mod. Opt. **47**, 355–376 (2000).
[2] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **84**, 2014 (2000).
[3] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[4] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824–3851 (1996).
[5] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).
[6] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206–4216 (1999).
[7] B. M. Terhal and P. Horodecki, Phys. Rev. A **61**, 040301 (2000).
[8] Instead of a proof by inspection of the figure, one can give an entirely rigorous proof of this statement; Karl Vollbrecht (unpublished).
[9] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[10] D. P. DiVincenzo, B. M. Terhal, and A. V. Thapliyal, J. Mod. Opt. **47**, 377–385 (2000).