On the Additivity of the Entanglement of Formation

Fabio Benatti\(^1\), Heide Narnhofer\(^2\)

\(^1\) Dep. Theor. Phys., University. di Trieste, Strada Costiera 11, I-34100, Trieste, Italy
\(^2\) Inst. Theor. Phys., University of Vienna, Boltzmanngasse 5, A-1090, Wien, Austria

We study whether the entanglement of formation is additive over tensor products and derive a necessary and sufficient condition for optimality of vector states that enables us to show additivity in two special cases.

PACS: 03.67.-a

Entanglement plays a crucial role in teleportation and quantum cryptography and is currently the focus of investigations in the developing field of quantum information [1]–[3]. Quantifying entanglement in a satisfactory way is a major issue in quantum information; there is a list of minimal desiderata and the available proposals have been proved to comply with all of them but for additivity for which only numerical support exists [3].

In this letter we are concerned with the entanglement of formation, defined by [4],

\[
E_f(p; M_1) = \min \left\{ \sum_{i \in I} p_i S(\sigma_i|_{M_1}) \right\},
\]

where \(p \) is a generic mixed state on the tensor product of two full matrix algebras \(M_1 \otimes M_2\), the minimum is computed over all decompositions \(p = \sum_{i \in I} p_i \sigma_i\) of \(p\) into pure states \(\sigma_i\) on \(M_1 \otimes M_2\) and \(S(\sigma_i|_{M_1})\) is the von Neumann entropy of the restriction of \(\sigma_i\) onto \(M_1\).

The question we want to address is whether

\[
E_f(p \otimes p; M_1 \otimes M_1) = 2 E_f(p; M_1),
\]

where \(M_1 \otimes M_1\) is short for \(M_1 \otimes 1_2 \otimes 1_2 \otimes M_1\).

Additivity or its failure will have a quantum information theoretic counterpart; there is indeed a connection between [1] and the maximal accessible information \(I(p)\) of a quantum source described by a mixed state \(p\) on a matrix algebra \(M\) [4,6]. If \(p = \sum_{\ell \in L} q_\ell p_\ell\), then

\[
I(p) = \sup_{B, p_\ell} I_B(p),
\]

where

\[
I_B(p) = - \sum_{\ell \in L} (\text{Tr}(p_\ell b_\ell)) \log(\text{Tr}(p_\ell b_\ell))
\]

\[
+ \sum_{\ell \in L} q_\ell \sum_{i \in I} (\text{Tr}(p_\ell b_i)) \log(\text{Tr}(p_\ell b_i)),
\]

the maximum being computed over all choices \(B = \{b_\ell\}\) of positive \(b_\ell \in M\) such that \(\sum_{i \in I} b_i = 1_M\). Optimal choices correspond to optimal detection of the classical information carried by the quantum states \(p_\ell\).

Since \(q_\ell p_\ell \leq p\) there exists a unique choice of operators \(0 < a_\ell \in M, \ell \in L\), with \(\sum_{\ell \in L} a_\ell = 1\), such that

\[
q_\ell p_\ell = \sqrt{p_\ell} a_\ell \sqrt{p}, \quad q_\ell = \text{Tr}(p_\ell a_\ell).
\]

Let \(\mathcal{A}\) be a commutative \(L\) dimensional algebra with identity \(1_{\mathcal{A}}\) and orthogonal projectors \(A_\ell\) with \(\sum_\ell A_\ell = 1_{\mathcal{A}}\). The map \(\gamma_A : \mathcal{A} \rightarrow M\) obtained by linear extension of \(A_\ell \mapsto \gamma_A(A_\ell) = a_\ell\) is positive and \(\gamma_A(1_\mathcal{A}) = 1_M\). Therefore, given any state \(\sigma\) on \(M\), the linear functional \(\sigma \circ \gamma_A : \mathcal{A} \rightarrow \mathbb{C}, \sigma \circ \gamma_A(A_\ell) = \text{Tr}(\sigma a_\ell)\) defines a state on \(\mathcal{A}\). Using (4) and the cyclicity of the trace,

\[
\text{Tr}(\rho_\ell b_i) = \frac{\text{Tr}(\rho b_i)}{\text{Tr}(p a_\ell)} \cdot \frac{\text{Tr}(\sigma a_\ell) \cdot \gamma(\alpha)}{\sqrt{\gamma(p a_\ell)}}.
\]

Setting \(p_\ell^B = \text{Tr}(\rho_\ell b_i), (6)\) becomes

\[
I_B(p) = S(\rho \circ \gamma_A) - \sum_{i \in I} p_i^B S(\sigma_i, \circ \gamma_A).
\]

Therefore, \(I(p)\) is the maximum of (6) over all possible decompositions of \(p\) into pure states:

\[
I(p) = S(\rho \circ \gamma_A) - \min_{p_\ell = \sum_{i \in I} p_i \sigma_i} \sum_{i \in I} p_i S(\sigma_i, \circ \gamma_A).
\]

If \(N\) is a subalgebra of \(M\), substituting the restrictions \(\rho|_N, \sigma_i|_N\) for \(\rho \circ \gamma_A, \sigma_i \circ \gamma_A\), respectively, \(\gamma_A\), we obtain the so-called entropy of a subalgebra

\[
H_p(N) := S(p|_N) - \min_{p_\ell = \sum_{i \in I} p_i \sigma_i} \sum_{i \in I} p_i S(\sigma_i|_N).
\]

The latter quantity is the building block of an extension of the Kolmogorov-Sinai dynamical entropy (or entropy per unit time) to the quantum realm. According to the above \(E_f(p; M_1) = S(p|_{M_1}) - H_p(M_1)\).

As the von Neumann entropy is additive over matrix algebras, if additivity fails for the entanglement of formation, it also fails for the entropy of a subalgebra. Then, from an information-theoretic point of view, we would deem possible to extract more information about the tensor product of two states over two independent subalgebras than that obtainable from the two of them independently.

In the following we try to use some of the properties of \(H_p(N)\) to investigate the general question whether

\[
E_f(p \otimes \sigma; M_1 \otimes M_3) = E_f(p; M_1) + E_f(\sigma; M_3),
\]

where \(p\) and \(\sigma\) are states on the (finite dimensional) algebras \(M_1 \otimes M_2\), respectively \(M_3 \otimes M_4\).

If \(E_f(p; M_1)\) and \(E_f(\sigma; M_3)\) are achieved at optimal decompositions \(p = \sum_\ell p_\ell p_\ell\) and \(\sigma = \sum_j q_\ell \sigma_j\), the factorized decomposition \(\rho \otimes \sigma = \sum_{j, \ell} q_{j, \ell} p_\ell \rho_\ell \otimes \sigma_j\) contribute...
to $E_f(\rho \otimes \sigma; M_1 \otimes M_3)$ with $E_f(\rho_1; M_1) + E_f(\sigma; M_3)$. However, the latter need not be optimal and the strict inequality $E_f(\rho \otimes \sigma; M_1 \otimes M_3) < E_f(\rho_1; M_1) + E_f(\sigma; M_3)$ is not excluded. In fact, a decomposition

$$\rho \otimes \sigma = \sum_i \alpha_i |\psi_i\rangle \langle \psi_i|, \quad \alpha_i > 0, \quad \sum_i \alpha_i = 1,$$

might be optimal with the $\psi_i$ entangled states over $M_1 \otimes M_3$. Let us consider the Schmidt decomposition

$$|\psi_i\rangle = \sum_j \beta_{ij} |\phi_{ij}^{12}\rangle \otimes |\phi_{ij}^{34}\rangle, \quad \|\psi_i\| = 1, \quad \beta_{ij} > 0,$$

where, for fixed $i$, the $|\phi_{ij}^{12}\rangle$’s and $|\phi_{ij}^{34}\rangle$’s form orthonormal bases over $M_1 \otimes M_2$, respectively $M_3 \otimes M_4$. If it held that

$$S(|\psi_i\rangle \langle |\psi_i|_{M_1 \otimes M_3}) \geq \sum_j \beta_{ij}^2 \left( S(|\phi_{ij}^{12}\rangle \langle \phi_{ij}^{12}|_{M_1}) + S(|\phi_{ij}^{34}\rangle \langle \phi_{ij}^{34}|_{M_3}) \right)$$

additivity would follow because tensor-product states would then never be worse than correlated ones.

Proving the sufficient condition (12) has so far escaped us; there are however particular cases where one can show additivity by using two results obtained for the entropy of a subalgebra (8). Both results concern general properties of optimal decompositions for $H_p(N)$ that we adapt to the entanglement of formation (12).

**Proposition 1.** If $\rho$ is a state on $M_1 \otimes M_2$, $E_f(\rho; M_1)$ is achieved at $\rho = \sum_i p_i \rho_i \otimes U$ is a unitary operator on $M_1 \otimes M_2$, then $E_f(U \rho U; M_2)$ is achieved at the optimal decomposition $U \rho U = \sum_i p_i \rho_i U U$$

**Proposition 2.** Let $\rho$ be a state on $M_1 \otimes M_2$ and $E_f(\rho; M_1)$ be achieved at $\rho = \sum_i \rho_i \otimes \rho_i$, that is, $E_f(\rho; M_1) = \sum_i q_i S(\rho_i |\rho_i\rangle)$, then

$$E_f(\sigma; M_1) = \sum_j q_j S(\rho_j |\rho_j\rangle),$$

where $\sigma = \sum_j q_j \rho_j$ is any linear convex combination of optimal states of $\rho$.

To the above, we add a new property. With some abuse of notation, we denote by $E_f(\rho; N)$ the minimum in (8), even if there is no tensor product structure in $N$.

**Proposition 3.** Let $|\psi_i\rangle \langle \psi_i|, i = 1, 2$, contribute to $E_f(\rho; N)$ and denote

$$\sigma_i := |\psi_i\rangle \langle \psi_i|_N, \quad i = 1, 2; \quad \sigma_{12} := |\psi_1\rangle \langle \psi_2|_N$$

$$\sigma_{ov}(\gamma) := \gamma \sigma_{21} + \gamma^* \sigma_{12}$$

$$\sigma(\gamma) := |\gamma|^{2} \sigma_1 + \sigma_2 - \sigma_{ov}(\gamma), \quad \hat{\sigma}(\gamma) := \frac{\sigma(\gamma)}{\text{Tr}(\sigma(\gamma))},$$

Then, for all complex $\gamma$, $|\gamma|^2 S(\sigma_1) + S(\sigma_2) + \text{Tr}\left(\sigma_{ov}(\gamma) \log \sigma_1 \right)$

$$\text{Tr}(\hat{\sigma}(\gamma)) \leq S(\hat{\sigma}(\gamma)).$$

Vice versa, if inequality (14) holds for all complex $\gamma$, then for all $\rho_3 = \lambda |\psi_1\rangle \langle \psi_1| + (1 - \lambda)|\psi_2\rangle \langle \psi_2|, 1 \geq \lambda \geq 0$, one gets $E_f(\rho_3; N) = \lambda S(\sigma_1) + (1 - \lambda) S(\sigma_2)$.

**Proof of Necessity:** Let $\varepsilon > 0$ and set

$$\rho_{\varepsilon, \gamma} = (1 + \varepsilon |\gamma|^2) |\psi_1\rangle \langle \psi_1| + \varepsilon + (1 + \varepsilon |\gamma|^2)|\psi_2\rangle \langle \psi_2|$$

be a not normalized state on $M$. As $\psi_i, i = 1, 2$ are optimal, Proposition 2 yields

$$E_f(\rho_{\varepsilon, \gamma}; N) = (1 + \varepsilon |\gamma|^2) S(\sigma_1) + \varepsilon + (1 + \varepsilon |\gamma|^2) S(\sigma_2).$$

Indeed, in taking the minimum in (12) normalization is not necessary. With $|\phi_1\rangle := |\psi_1\rangle + \varepsilon |\gamma\rangle |\sigma_2\rangle$ and $|\phi_2\rangle := |\psi_1\rangle - \gamma^{-1} |\psi_2\rangle$, we construct a new decomposition $\rho_{\varepsilon, \gamma} = |\phi_1\rangle \langle \phi_1| + \gamma^2 |\phi_2\rangle \langle \phi_2|$

The latter cannot contribute more than $E_f(\rho_{\varepsilon, \gamma}; N)$; therefore, $E_f(\rho_{\varepsilon, \gamma}; N) \leq f(\varepsilon)$, where

$$f(\varepsilon) := \|\phi_1\|^2 S \left( |\phi_1\rangle \langle \phi_1|_N \right) + \varepsilon |\gamma|^2 \|\phi_2\|^2 S \left( |\phi_2\rangle \langle \phi_2|_N \right).$$

Inequality (14) must then hold at first order in $\varepsilon$.

**Proof of Sufficiency:** By assumption, inequality (14) holds for all $\gamma$’s. Thus, choosing $\alpha_i \geq 0$ and $\gamma_i$ such that $\sum_i \alpha_i |\gamma_i|^2 = \lambda$, $\sum_i \alpha_i = 1 - \lambda$ and $\sum_i \alpha_i \gamma_i = 0$, we get

$$\lambda S(\sigma_1) + (1 - \lambda) S(\sigma_2) \leq \sum_i \alpha_i \left( \text{Tr}(\sigma(\gamma_i)) \right) S(\hat{\sigma}(\gamma_i)) \right).$$

In the above, the left hand side is the contribution to $E_f(\rho_3; N)$ of $\rho_\lambda = \lambda |\psi_1\rangle \langle \psi_1| + (1 - \lambda)|\psi_2\rangle \langle \psi_2|$, whereas the right hand side is the contribution of

$$\rho_3 = \sum_i \alpha_i |\psi_1 + \gamma_i |\psi_2\rangle \langle \psi_1 + \gamma_i |\psi_2|.$$
Further, let $|\chi^T_j\rangle$ be eigenvectors of $\sigma_3$ and consider the unitary operator ($\in M_3$)

$$\hat{U}_j = 1 - (1 - i)P_t,$$  
$$P_t := |\chi^T_j\rangle\langle\chi^T_j|.$$  
From Proposition 1 it follows that the vectors $U_j|\psi_i\rangle$, $U_j = 1_2 \otimes U_j \otimes 1_4$, also give $E_j(\rho \otimes \sigma; M_1 \otimes M_3)$. Let us concentrate on $|\psi_i\rangle$ together with $U_j|\psi_i\rangle$, they have to satisfy (13) for all $\gamma$. Then, according to the notation of Proposition 3,

$$(17) \quad \sigma_1 = \sum_j c^2_j |\phi^T_{j1}\rangle\langle\phi^T_{j1}|_{M_1 \otimes M_5}$$

$$(18) \quad \sigma_2 = U_j \sigma_1 U_j^\dagger, \quad \sigma_{\alpha \nu}(\gamma) = \gamma U_j \sigma_1 + \gamma^* \sigma_1 U_j^\dagger.$$  
Taking $\gamma = 1$, it follows that $\hat{\sigma}(1) = \frac{P_t \sigma_1 P_t}{Tr(P_t \sigma_1)}$ and $\sigma_{\alpha \nu}(1) = 2\sigma_1 - (1 - i)P_t \sigma_1 - (1 + i)\sigma_1 P_t$. Inequality (14) thus becomes

$$(19) \quad -2Tr(P_t \sigma_1 \log \sigma_1) \leq Tr\left(P_t \sigma_1 \right) S\left(\hat{\sigma}(1)\right).$$  
We develop $|\phi^T_{13}\rangle = \sum_p \beta_{j\ell}^2 |\phi^T_{j\ell}\rangle \otimes |\chi^T_{j\ell}\rangle$, along an orthonormal basis for the factor $M_1$, then, by means of the spectral decomposition (12), setting $\Delta_{j\ell} := \langle \phi^T_{j\ell}|P_t|\phi^T_{j\ell}\rangle = \sum_p |\beta_{j\ell}^2|^2$, we get $P_t \sigma_1 P_t = \left(\sum_j c^2_j \Delta_{j\ell} Q_{j\ell}\right) \otimes P$ where

$$Q_{j\ell} := |\chi^T_{j\ell}\rangle\langle\chi^T_{j\ell}|$$

$$|\chi^T_{j\ell}\rangle = \frac{\sum_j \beta_{j\ell}^2}{\Delta_{j\ell}} |\phi^T_{j\ell}\rangle.$$  
Insertion in (19) leads to

$$0 \geq \sum_j c^2_j \Delta_{j\ell} \log \frac{\Delta_{j\ell}}{c^2_j \text{Tr}(P_t \sigma_1)} \geq \sum_j c^2_j \left(\Delta_{j\ell} - c^2_j \text{Tr}(P_t \sigma_1)\right),$$

the latter inequality coming from $x \log x/y \geq x - y$ and holding for all orthogonal projectors $P_t$. Since $\sum_j c^2_j = 1$ and $\sum_j \Delta_{j\ell} = 1$, summing over $\ell$, we get that $c_{j1} = 1$ for one $j$ and $c_{1k} = 0$ if $k \neq j$. Thus, the supposed optimal vectors $|\psi_i\rangle$ must be of the form $|\psi_i\rangle = |\phi^T_{13}\rangle \otimes |\phi^T_{24}\rangle$ and the supposed optimal decomposition (10) must reduce to

$$\rho \otimes \sigma_3 \otimes \sigma_4 = \sum_i \sigma^4_{i\ell} |\phi^T_{i1}\rangle\langle\phi^T_{i1}| \otimes |\phi^T_{j\ell}\rangle\langle\phi^T_{j\ell}|.$$  

Tracing over $M_2 \otimes M_3$ with respect to the Schmidt decompositions $|\phi^T_{13}\rangle = \sum_j \delta_{j1}^2 |\phi^T_{j1}\rangle \otimes |\phi^T_{i\ell}\rangle$ and $|\phi^T_{24}\rangle = \sum_j \delta_{j2}^2 |\phi^T_{j2}\rangle \otimes |\phi^T_{i\ell}\rangle$, orthogonality yields

$$\rho = \sum_i \sigma^4_{i\ell} \left(\delta_{j1}^2)^2 |\phi^T_{i1}\rangle\langle\phi^T_{i1}| \otimes |\phi^T_{j1}\rangle\langle\phi^T_{j1}| + \right. $$

$$\left. \left(\delta_{j2}^2)^2 |\phi^T_{i2}\rangle\langle\phi^T_{i2}| \otimes |\phi^T_{j2}\rangle\langle\phi^T_{j2}| \right).$$  
We thus conclude that a decomposition of $\rho \otimes \sigma_3 \otimes \sigma_4$ as in (14) can be optimal with respect to $M_1 \otimes M_3$ only if $\rho$ is not entangled over $M_1 \otimes M_2$, in which case $E_j(\rho \otimes \sigma_3 \otimes \sigma_4; M_1 \otimes M_2) = 0$ is obviously additive. If $\rho$ is entangled over $M_1 \otimes M_2$, the contradiction is avoided only if the optimal decompositions have the form

$$\rho \otimes \sigma_3 \otimes \sigma_4 = \sum_i \sigma_i |\phi^T_{i1}\rangle\langle\phi^T_{i1}| \otimes |\phi^T_{j1}\rangle\langle\phi^T_{j1}|.$$  
Thus, the optimal states cannot carry any entanglement over $M_1 \otimes M_3$ and additivity follows.

The second case we want to discuss is somewhat the opposite of the previous one where we proved that optimal projections for the tensor products are products of optimal projectors for the factors. In the second case, we want to show that putting together couples of optimal projectors for the factors we get optimal decompositions.

**Case 2:** we consider the state

$$(23) \quad \rho_{\lambda} = \rho \lambda + (1 - \lambda) \dot{\rho}$$
on $M_1 \otimes M_2 \otimes M_3 \otimes M_4$, where $\rho := |\phi^{12}\rangle \langle\phi^{12}| \otimes |\phi^{34}\rangle \langle\phi^{34}|$ and $\dot{\rho} := |\hat{\phi}^{12}\rangle \langle\hat{\phi}^{12}| \otimes |\hat{\phi}^{34}\rangle \langle\hat{\phi}^{34}|$.

Let $|\phi^{12}\rangle$ and $|\phi^{12}\rangle$ be optimal vectors for some state $\rho$ on $M_1 \otimes M_2$ relative to $M_1$ and $|\phi^{34}\rangle = |\phi^{34}\rangle \otimes |\phi^{34}\rangle$, $|\hat{\phi}^{34}\rangle = |\hat{\phi}^{34}\rangle \otimes |\hat{\phi}^{34}\rangle$ on $M_3 \otimes M_4$ so that $E\left(\rho_{\lambda}; M_1 \otimes M_3\right) = 0$.

The contribution to $E\left(\rho_{\lambda}; M_1 \otimes M_3\right)$ of the decomposition (24) is thus

$$(24) \quad E_{\lambda} := \lambda S\left(|\phi^{12}\rangle \langle\phi^{12}|_{M_1}\right) + (1 - \lambda) S\left(|\phi^{34}\rangle \langle\phi^{34}|_{M_3}\right)$$
and we want to prove that this is the best we can have.

We proceed as follows: as for (13), a general decomposition of $\rho_{\lambda}$ is of the form $\rho_{\lambda} = \sum_i a_i |\psi_i\rangle\langle\psi_i|$, where $|\psi_i\rangle = |\phi^{12} \otimes \phi^{34}\rangle + \gamma_i |\hat{\phi}^{12} \otimes \hat{\phi}^{34}\rangle$, with $a_i > 0$ and

$$\sum_i a_i = \lambda, \quad \sum_i a_i |\gamma_i|^2 = 1 - \lambda, \quad \sum_i a_i |\gamma_i|^2 = 0.$$  
We now set $b := |\phi^4\rangle \langle\phi^4|$, $a := \sqrt{1 - b^2}$ and construct the normalized vector state $|\psi^4\rangle := \frac{a |\phi^4\rangle - b |\phi^4\rangle}{a}$ such that $<\psi^4|\phi^4> = 0$. We can thus rewrite

$$(26) \quad |\psi_i\rangle = a_i |\phi^T_{i12}\rangle \otimes |\phi^4\rangle + a_i \gamma_i |\phi^T_{i12}\rangle \otimes \hat{\phi}^3 \otimes |\phi^4\rangle,$$  
where

$$|\phi^T_{i12}\rangle := |\phi^{12} \otimes \phi^3\rangle + b \gamma_i |\hat{\phi}^{12} \otimes \hat{\phi}^3 \otimes \hat{\phi}^4\rangle$$

$$a_i^2 := 1 + |b|^2 |\gamma_i|^2 + 2 Re(\bar{\gamma}_i \langle \phi^{12} | \phi^{12} | \phi^3 \rangle) \langle \phi^4 | \phi^4 \rangle.$$  
With $|\gamma_i\rangle := \frac{|\psi_i\rangle}{\sqrt{a_i}}$, $\gamma_i := a_i^2 + a_i^2 |\gamma_i|^2$, the decomposition (23) reads $\rho_{\lambda} = \sum_i a_i |\psi_i\rangle\langle\psi_i|$.
The contribution of the latter to the entanglement of formation \( E_f(\rho_A; M_1 \otimes M_3) \) is

\[
E := \sum_i \alpha_i \delta_i S \left( \hat{\psi_i} \langle \hat{\psi_i} \mid_{M_1 \otimes M_3} \right) .
\]  

(27)

From the orthogonality of \( \psi^4 \) and \( \phi^4 \) it follows that

\[
|\hat{\psi_i} \rangle \langle \hat{\psi_i} |_{M_1 \otimes M_3} = \frac{a_i^2}{\delta_i} \sigma_1^{123} + \frac{a_i^2 |\gamma|^2}{\delta_i} \sigma_1^{123} ,
\]

where

\[
\begin{align*}
\sigma_1^{123} & := |\phi_1^{123}\rangle \langle \phi_1^{123}|_{M_1 \otimes M_3} \\
\sigma_1^{123} & := |\phi_1^{123}\rangle \langle \phi_1^{123}|_{M_1 \otimes M_3} .
\end{align*}
\]

(28)

Concavity of the von Neumann entropy yields

\[
E \geq \sum_i \alpha_i \left\{ a_i^2 S \left( \sigma_1^{123} \right)_{M_1 \otimes M_3} \right\} .
\]

(29)

As done before, we construct the normalized vector

\[
|\psi^3\rangle := \frac{c}{\sqrt{1 - |d|^2}} |\phi^3\rangle ,
\]

where \( d := \langle \phi^3 | \phi^3 \rangle \), \( c := \sqrt{1 - |d|^2} \), and

\[
|\phi_1^{123}\rangle := \frac{b_1 \langle \psi_1^{12} | \phi^3 \rangle + b c \gamma_i |\phi_1^{12} \rangle}{a_1} ,
\]

(30)

Introducing the Schmidt decompositions over \( M_1 \otimes M_2 \):

\[
\psi_1^{12} = \sum_j c_{ij} |\phi_1^{12}\rangle \otimes |\phi_2^j\rangle ,
\]

setting \( \rho_1 := |\psi_1^{12}\rangle \langle \psi_1^{12}|_{M_1} \), \( \rho_2 := |\phi_1^{12}\rangle \langle \phi_1^{12}|_{M_2} \), because of the orthogonality of \( \phi^3 \) and \( \psi^3 \), the state \( \sigma_1^{123} \) restricted to \( M_1 \otimes M_3 \) can be represented as

\[
\begin{align*}
\sigma_1^{123} & = \frac{1}{a_i^2} \left( \begin{array}{cc}
\frac{b_i^2 \rho_1}{\sqrt{\rho_1} V \sqrt{\rho_2}} & \frac{b_i b c \rho_1 V \sqrt{\rho_2}}{\sqrt{\rho_1} V \sqrt{\rho_2}} \\
\frac{b_i b c \rho_1 V \sqrt{\rho_2}}{\sqrt{\rho_1} V \sqrt{\rho_2}} & \frac{b_i^2 c \gamma_i \rho_1 V \sqrt{\rho_2}}{\sqrt{\rho_1} V \sqrt{\rho_2}} \\
\end{array} \right) \\
& = \frac{1}{a_i^2} \left( \begin{array}{cc}
\frac{b_i \rho_1 V \sqrt{\rho_2}}{\sqrt{\rho_1}} & 0 \\
0 & \frac{b_i V \sqrt{\rho_2}}{\sqrt{\rho_1}} \\
\end{array} \right) ,
\end{align*}
\]

where \( V := \sum_{i,j} \langle \phi_1^{12} | \phi_1^{12} \rangle |\phi_1^{12}\rangle |\phi_1^{12}\rangle \) is a unitary operator and \( V \sqrt{\rho_2} = |\psi_1^{12}\rangle \langle \psi_1^{12}|_{M_1} \).

Since \( \sigma_1^{123} = A^t A \) has the same entropy as \( A A^t = \frac{1}{a_i^2} \left( \begin{array}{cc}
\frac{b_i^2 \rho_1 V \sqrt{\rho_2}}{\sqrt{\rho_1}} & 0 \\
0 & \frac{b_i V \sqrt{\rho_2}}{\sqrt{\rho_1}} \\
\end{array} \right) \), concavity and invariance under unitary transformations of the von Neumann entropy yield

\[
S \left( \sigma_1^{123} \right) = S \left( A A^t \right) \geq \frac{b_i^2}{a_i^2} S \left( \rho_1 \right) + \frac{c^2 |b|^2 |\gamma|^2}{a_i^2} S \left( \rho_2 \right) ,
\]

whence \( [9] \) becomes

\[
E \geq \sum_i \alpha_i \left\{ b_i^2 S \left( \rho_1 \right) + |\gamma|^2 (a^2 + b^2 |\gamma|^2) S \left( \rho_2 \right) \right\} .
\]

(31)

Since we assumed the states \( |\phi^{12}\rangle \) and \( |\phi^{12}\rangle \) in \( [10] \) to be optimal for some state on \( M_1 \otimes M_2 \) when restricted to \( M_1 \), we can use the necessary condition \( [1] \). According to the notation of Proposition 3, we have \( \sigma_1 = \rho_2 \), \( \sigma_2 = |\phi^{12}\rangle \langle \phi^{12}|_{M_1} \), \( \gamma = \gamma_1 b^* d^* \), \( \delta (\gamma) = \rho_1 \) and

\[
\sigma_{ov} (\gamma) = -b^* d^* \gamma_1 \sqrt{\rho_2} V^{\dagger} \sqrt{\rho_1} - b d \gamma_1 \sqrt{\rho_1} V \sqrt{\rho_2} .
\]

From \( [1] \) and the conditions \( [23] \) it follows that

\[
E \geq \sum_i \alpha_i \left\{ S \left( \sigma_1 \mid \phi^{12} \rangle \langle \phi^{12} \mid_{M_1} \right) + |\gamma|^2 S \left( \sigma_2 \mid \phi^{12} \rangle \langle \phi^{12} \mid_{M_1} \right) \right\}
\]

\[
- \text{Tr} \left( \sigma_{ov} (b^* d^* \gamma_1) \log \rho_2 \right) \right\} = E_{\lambda} ,
\]

(32)

where \( E_{\lambda} \) is the contribution \( [22] \) to the entanglement of formation \( E \left( \rho_{\lambda}; M_1 \otimes M_3 \right) \) of the decomposition \( [23] \), which turns out then to be already optimal.

In this letter we have derived a necessary and sufficient condition for the optimality of two vector states and showed its usefulness by proving additivity in two cases. While in case 1. additivity was rather expected because of the tensor-product state \( \rho \otimes \sigma \), it was less so in case 2. We requested, however, additional properties on the state structure over \( M_1 \otimes M_2 \); in case 1. factorization of \( \sigma = \sigma_3 \otimes \sigma_4 \) and in case 2. factorization into pure states of the optimal decomposers. In both cases the state was thus separable with respect to \( M_3 \otimes M_4 \).

[1] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, and W.K. Wootters, Phys. Rev. A 54, 3824 (1996)
[2] S. Hill, and W.K. Wootters, Phys. Rev. Lett. 78, 5022 (1997)
[3] W.K. Wootters, Phys. Rev. Lett. 80, 2245 (1998)
[4] V. Vedral, and M.B. Plenio, Pys. Rev. A 57, 1619 (1998)
[5] M. Horodecki, P. Horodecki, R. Horodecki, lanl e-print quant-ph/9908066
[6] G.G. Amosov, A.S. Hleovo, and R.F. Werner, lanl e-print math-ph/0003002
[7] A.S. Hleovo, Prob. Inf. Transmission USSR 9 31 (1993)
[8] F. Benatti, J. Math. Phys. 37 5244 (1996)
[9] A. Connes, H. Narnhofer, and W. Thirring, Comm. Math. Phys. 112, 691 (1987); H. Narnhofer, and W. Thirring: Fizika 17, 257 (1985)
[10] A. Peres, and W.K. Wootters, Phys. Rev. Lett. 66, 1119 (1991)
[11] N. Gisin and S. Popescu, lanl e-print quant-ph/9901075
[12] F. Benatti, H. Narnhofer, and A. Uhlmann, Rep. Math. Phys. 38, 123 (1996)
[13] F. Benatti, H. Narnhofer, and A. Uhlmann, Lett. Math. Phys. 47, 237 (1999)