

OPTIMAL DECOMPOSITIONS OF QUANTUM STATES WITH RESPECT TO ENTROPY

F. BENATTI

Dip. Fisica Teorica, Trieste University, Italy (e-mail: Benatti@trieste.infn.it)

H. NARNHOFER

Inst. für Theoretische Physik, Vienna University, Austria (e-mail: Narnh@pap.univie.ac.at)

A. UHLMANN

Inst. für Mathematik, Leipzig University, Germany (e-mail: Uhlmann@qft.physik.uni-leipzig.d400.de)

(Received October 19, 1995)

Any non-pure quantum state admits an infinite number of non-trivial decompositions. A recent proposal how to measure the information content of a quantum state with reference to a given subalgebra of operators, singles out some of them, called optimal decompositions, which depend both on the state and on the subalgebra. In this paper we start exploring their main features.

1. Introduction

Considering a state ω over a large algebra \mathbf{M} , but concentrating on its restriction $\omega|_{\mathbf{A}}$ to a smaller algebra $\mathbf{A} \subseteq \mathbf{M}$, we see that the usual definition of entropy [1, 2] of the restricted state fails to be monotonic for noncommutative algebras. This causes severe problems if one considers this entropy as a starting point for a dynamical entropy of the Kolmogorov–Sinai type [3]. In the framework of relativistic quantum field theories problems arise, too; the local entropy that is needed to do thermodynamics becomes infinite [4].

In [5] an alternative definition is offered, tailor-made to restore monotonicity which turned out to be useful to define both a quantum dynamical entropy [6, 7] and a local entropy in quantum field theory. The definition is based on an appropriate optimization of the decomposition of the state on the large algebra \mathbf{M} with respect to the small algebra $\mathbf{A} \subseteq \mathbf{M}$.

Unfortunately, only under very special circumstances, that is in case a state-preserving conditional expectation exists, we have complete control how these optimal decompositions look like. In addition, for the abelian subalgebra of \mathbf{M}_2 generated by $\mathbb{1}$ and the Pauli matrix σ_z and the state $\omega = (\mathbb{1} + a^2 + 2a\sigma_x)/2(1 + a^2)$, the optimal decomposition was given in [8], though without a detailed argument. Its generalization to arbitrary states can be found in [9], again the surprisingly lengthy and not so much

revealing argument was omitted. In all examples under control all decompositions corresponded to an appropriately chosen maximal abelian subalgebra.

In this paper we want to give a support to the conjecture that all possible decompositions that are needed to define entropy correspond to maximal abelian subalgebras. In order to do so, we will first study the general structure of optimal decompositions and find compatibility relations between the components of such a decomposition.

2. The entropy functional and its structure

DEFINITION 1. Let \mathbf{M} be a von Neumann algebra and ω a state over it. Let \mathbf{A} be a subalgebra of \mathbf{M} . Then, the entropy of \mathbf{A} relative to the state ω is [7]

$$H_{\omega, \mathbf{M}}(\mathbf{A}) := \sup_{\sum_i \lambda_i \omega_i = \omega} \sum_i \lambda_i S(\omega, \omega_i) | \mathbf{A}. \quad (1)$$

In the above definition $\sum_i \lambda_i \omega_i$ is any convex decomposition of ω , as a state on \mathbf{M} , into (normalized) states ω_i (over \mathbf{M}) and $S(\omega, \omega_i) | \mathbf{A}$ is the *relative entropy* of the states ω, ω_i restricted to \mathbf{A} .

Some properties

1. As $S(\omega, \rho) | \mathbf{A}$ is increasing in \mathbf{A} [2, 10, 5], so is $H_{\omega, \mathbf{M}}(\mathbf{A})$.
2. As every decomposition of ω as a state over \mathbf{M} gives a decomposition of the restriction of ω to any subalgebra $\mathbf{N} \subseteq \mathbf{M}$ (the reverse need not be generally true), $H_{\omega, \mathbf{M}}(\mathbf{A})$ is monotonically decreasing in \mathbf{M} .
3. If \mathbf{A} is finite dimensional, the relative entropy can be written

$$S(\omega, \omega_i) | \mathbf{A} = \text{Tr } \omega_i | \mathbf{A} \{ \log \omega_i | \mathbf{A} - \log \omega | \mathbf{A} \}, \quad (2)$$

where

$$S(\omega_i | \mathbf{A}) = -\text{Tr } \omega_i | \mathbf{A} \log \omega_i | \mathbf{A} \quad (3)$$

is the von Neumann entropy of the state ω_i restricted to the (finite dimensional) subalgebra \mathbf{A} .

Accordingly, the entropy functional (1) reads:

$$H_{\omega, \mathbf{M}}(\mathbf{A}) = S(\omega | \mathbf{A}) - \mathcal{R}(\omega, \mathbf{A}), \quad (4)$$

$$\mathcal{R}(\omega, \mathbf{A}) := \inf \left\{ \mathcal{R}(\omega, \{\lambda_j \omega_j\}, \mathbf{A}), \sum_j \lambda_j \omega_j = \omega \right\},$$

$$\mathcal{R}(\omega, \{\lambda_j \omega_j\}, \mathbf{A}) := \sum_j \lambda_j S(\omega_j | \mathbf{A}). \quad (5)$$

4. In (4), it suffices to compute the infimum over decompositions into pure states. In fact, any non-pure state ω_j appearing in $\omega = \sum_j \lambda_j \omega_j$ could be further decomposed into $\omega_j = \sum_k \nu_k^j \omega_k^j$ and due to strict concavity of the von Neumann entropy:

$$S(\omega_j | \mathbf{A}) > \sum_k \nu_k^j S(\omega_k^j | \mathbf{A}), \quad (6)$$

unless $\omega_k^j | \mathbf{A} = \omega_j | \mathbf{A}$, $\forall k$, we could improve on $\mathcal{R}(\omega, \{\lambda_j \omega_j\}, \mathbf{A})$.

THEOREM 1. *Let the set of states $\{\omega(i), i \in I\}$ and corresponding weights $\{\lambda(i)\}$ provide an optimal decomposition for the state ω over \mathbf{M} with respect to \mathbf{A} . that is*

$$\omega = \int_I d\mu(i) \lambda(i) \omega(i), \quad (7)$$

$$\mathcal{R}(\omega, \mathbf{A}) = \mathcal{R}(\omega, \{\lambda(i) \omega(i)\}_{i \in I}, \mathbf{A}), \quad (8)$$

where we have kept the possibility that the set I has finite cardinality or that we can decompose with respect to a continuous measure on it. Consider the state

$$\tilde{\omega} = \int_I d\mu(i) \tilde{\lambda}(i) \omega(i), \quad \int_I d\mu(i) \tilde{\lambda}(i) = 1. \quad (9)$$

with $\tilde{\lambda}(i)$ a positive μ -measurable function on I . Then

$$\mathcal{R}(\tilde{\omega}, \mathbf{A}) = \mathcal{R}(\tilde{\omega}, \{\tilde{\lambda}(i) \omega(i)\}_{i \in I}, \mathbf{A}). \quad (10)$$

Proof: Choose a μ -measurable function on I , $0 \leq g(i) \leq 1$, and set

$$\omega_1 := \int_I d\mu(i) g(i) \lambda(i) \omega(i). \quad (11)$$

$$\omega_2 := \int_I d\mu(i) (1 - g(i)) \lambda(i) \omega(i). \quad (12)$$

by using the $\lambda(i)$ and $\omega(i)$ given in (7).

The states $\omega_{1,2}$ in $\omega = \omega_1 + \omega_2$ are not normalized. Let $\nu_{1,2}$ be the normalizing factors and consider any optimal decomposition $\{\nu_j^1 \tilde{\omega}_j^1\}_{j \in J}$ for $\tilde{\omega}_1 = \nu_1^{-1} \omega_1$, respectively $\{\nu_k^2 \tilde{\omega}_k^2\}_{k \in K}$ for $\tilde{\omega}_2 = \nu_2^{-1} \omega_2$. Together they provide a decomposition $\{\nu_j^1 \tilde{\omega}_j^1\}_{j \in J} \cup \{\nu_k^2 \tilde{\omega}_k^2\}_{k \in K}$ for ω which is not necessarily optimal, thus:

$$\begin{aligned} \mathcal{R}(\omega, \mathbf{A}) &\leq \nu_1 \mathcal{R}(\tilde{\omega}_1, \{\nu_j^1 \tilde{\omega}_j^1\}_{j \in J}, \mathbf{A}) + \nu_2 \mathcal{R}(\tilde{\omega}_2, \{\nu_k^2 \tilde{\omega}_k^2\}_{k \in K}, \mathbf{A}) \\ &= \nu_1 \mathcal{R}(\tilde{\omega}_1, \mathbf{A}) + \nu_2 \mathcal{R}(\tilde{\omega}_2, \mathbf{A}). \end{aligned} \quad (13)$$

On the other hand, through (11) and (12) the optimal decomposition (7) for ω provides two decompositions for $\tilde{\omega}_{1,2}$. The latter need not be optimal for $\tilde{\omega}_1$, respectively $\tilde{\omega}_2$, hence:

$$\begin{aligned} \nu_1 \mathcal{R}(\tilde{\omega}_1, \mathbf{A}) + \nu_2 \mathcal{R}(\tilde{\omega}_2, \mathbf{A}) &\leq \mathcal{R}(\omega_1, \{g(i) \lambda(i) \omega(i)\}_{i \in I}, \mathbf{A}) + \\ &\quad + \mathcal{R}(\omega_2, \{(1 - g(i)) \lambda(i) \omega(i)\}_{i \in I}, \mathbf{A}) \\ &\leq \mathcal{R}(\omega, \mathbf{A}), \end{aligned} \quad (14)$$

so it follows that the two decompositions are indeed optimal and a further normalization of ω_1 and ω_2 concludes the proof. \blacksquare

DEFINITION 2. Any linear convex combination $\omega = \sum_{i \in I} \lambda_i \omega_i$ (I being possibly of infinite cardinality) of states ω_i on \mathbf{M} (possibly not normalized) at which the minimum in (4) is attained we will call *optimal* for $\mathbf{A} \subseteq \mathbf{M}$. That is

$$\mathcal{R}(\omega, \{\lambda_i \omega_i\}_{i \in I}, \mathbf{A}) = \mathcal{R}(\omega, \mathbf{A}). \quad (15)$$

DEFINITION 3. A set of vectors $\{|\phi_i\rangle\}_{i \in I}$ will be called *compatible* (with respect to $\mathbf{A} \subseteq \mathbf{M}$) if every positive, linear functional $\mathbf{M} \ni \hat{x} \mapsto \sum_{i \in I} \lambda_i \langle \phi_i | \hat{x} | \phi_i \rangle$, with $\lambda_i > 0$, $\sum_{i \in I} \lambda_i = 1$ and with mutually different positive linear functionals $\mathbf{M} \ni \hat{x} \mapsto \langle \phi_i | \hat{x} | \phi_i \rangle$ is an optimal decomposition with respect to \mathbf{A} (if I is not of finite cardinality, the corresponding set of vectors is called compatible if every finite subset is compatible).

A set $\{|\phi_i\rangle\}_{i \in I}$ is called *maximally compatible* (with respect to \mathbf{A}) if it is compatible but not a proper subset of any compatible set.

Remarks: 1. In Theorem 1, we have written the decomposition as an integral to cover the possibility of both finite and infinite decompositions. If \mathbf{M} is not of type I , then we know that pure states on \mathbf{M} do not exist, therefore the infimum is not attained. However, the theorem holds if we interpret the result in the sense that $\mathcal{R}(\omega, \{\lambda(i)\omega(i)\}, \mathbf{A})$ can get arbitrarily close to the infimum. To avoid these difficulties, from now on we restrict ourselves to *finite dimensional* algebras \mathbf{M} on which the pure states are well defined and we will take $\mathbf{M} = \mathbf{M}_n(\mathbb{C})$ as a full $(n \times n)$ matrix algebra over a complex Hilbert space $\mathcal{H}_n = \mathbb{C}^n$.

2. The specification ‘‘mutually different’’ in the above definition serves to get rid of the arbitrariness of phase factors. Otherwise, any maximally optimal set might be of infinite cardinality, for it could contain, together with a vector $|\psi\rangle$, the whole ray $\{e^{i\gamma}|\psi\rangle\}_{\gamma \in [0, 2\pi)}$.

LEMMA 1. *If $\mathbf{A} = \mathbf{M}$, all sets $\{|\phi_1\rangle, \dots, |\phi_r\rangle\}$, $r = 1, \dots, +\infty$, are compatible.*

Proof: The states $\phi_j: \hat{x} \mapsto \phi_j(\hat{x}) := \langle \phi_j | \hat{x} | \phi_j \rangle$ are pure on \mathbf{M} and then have von Neumann entropy $S(\phi_j | \mathbf{M}) = 0$. They need not be pure on a strict subalgebra $\mathbf{A} \subset \mathbf{M}$, but in our case $\mathbf{A} = \mathbf{M}$. Then

$$\sum_j \lambda_j S(\phi_j | \mathbf{M}) = 0 = \inf \left\{ \sum_k \nu_k S(\rho_k | \mathbf{M}), \sum_k \nu_k \rho_k = \sum_j \lambda_j \phi_j \right\}. \quad (16)$$

■

LEMMA 2. *Let $\mathbf{M} = \mathbf{M}_n(\mathbb{C})$. Then, there exist compatible sets with cardinality $1, 2, \dots, n$.*

Proof: Compatible single vector sets exist because any pure state $\hat{x} \mapsto \phi(\hat{x}) := \langle \phi | \hat{x} | \phi \rangle$ on \mathbf{M} has $S(\phi | \mathbf{M}) = 0$ and cannot be further decomposed.

Any faithful state ω on \mathbf{M} needs at least n vectors $|\phi_i\rangle$ to be decomposed into pure states $\mathbf{M} \ni \hat{x} \mapsto \langle \phi_i | \hat{x} | \phi_i \rangle$, the latter giving

$$\mathcal{R}(\omega, \{\lambda_i \phi_i\}, \mathbf{M}) = 0 = \mathcal{R}(\omega, \mathbf{M}).$$

Therefore, for faithful states we need at least n pure states to decompose them optimally with respect to \mathbf{M} . But, in principle, we could need more than that.

Cardinalities $2 \leq k \leq n$ can be derived as follows. We consider a sequence of states built from a continuous compatible set $\{\phi(\lambda)\}$ by means of an absolutely continuous measure $\mu(\lambda)$ on some finite measure space Λ :

$$\omega_n := \int_{\Lambda} d\mu(\lambda) g_n(\lambda) \phi(\lambda), \quad (17)$$

$$0 \leq g_n(\lambda), \quad \int_{\Lambda} d\mu(\lambda) g_n(\lambda) = 1 \quad \forall n \in \mathbb{N}, \quad (18)$$

$$\lim_{n \rightarrow +\infty} \int_{\Lambda} d\mu(\lambda) g_n(\lambda) f(\lambda) = \frac{1}{l} \sum_{j=1}^l f(\lambda_j), \quad (19)$$

for any continuous function f on Λ . Then, because of the finite dimensionality of \mathbf{M} , the following limit holds in the norm topology:

$$\omega_n \longrightarrow \frac{1}{l} \sum_{j=1}^l \phi_{\lambda_j} := \omega \quad (20)$$

in the sense that $\lim_{n \rightarrow +\infty} \underbrace{\sup_{\hat{x} \in \mathbf{M}, \|\hat{x}\|=1} |\omega_n(\hat{x}) - \omega(\hat{x})|}_{\|\omega_n - \omega\|} = 0$.

Since $\|\omega_n - \omega\| \rightarrow 0$ implies $\lim_n S(\omega_n|\mathbf{A}) = S(\omega|\mathbf{A})$ [7] for all $\mathbf{A} \subseteq \mathbf{M}$, from Theorem 1 we get

$$\lim_{n \rightarrow +\infty} \mathcal{R}(\omega_n, \mathbf{A}) = \lim_{n \rightarrow +\infty} \mathcal{R}(\omega_n, \{g_n(\lambda)\phi(\lambda)\}_{\lambda \in \Lambda}, \mathbf{A}) = \frac{1}{l} \sum_{j=1}^l S(\phi_{\lambda_j}|\mathbf{A}). \quad (21)$$

On the other hand, $\liminf \leq \inf \lim$ implies

$$\lim_{n \rightarrow +\infty} \mathcal{R}(\omega_n, \mathbf{A}) \leq \mathcal{R}(\omega, \mathbf{A}) \leq \frac{1}{l} \sum_{j=1}^l S(\phi_{\lambda_j}|\mathbf{A}). \quad (22)$$

■

Instead of looking for optimal decompositions of states on $\mathbf{M} = \mathbf{M}_n(\mathbb{C})$, we proceed by investigating the structure of compatible sets.

The von Neumann entropy $S(\omega|\mathbf{A})$ is a smooth function of ω except at the states ω where $\omega|\mathbf{A}$ has eigenvalue 0. Therefore, we can consider the functional $\mathcal{R}(\omega, \{\lambda_j \langle \phi_j | \cdot | \phi_j \rangle\}, \mathbf{A})$ to be differentiable with respect to the various components $\phi_j(\alpha)$ of the vectors $|\phi_j\rangle$ in a suitable orthonormal basis $\{|\alpha\rangle\}_{\alpha=1}^n$ of $\mathcal{H}_n = \mathbb{C}^n$. We then proceed with

DEFINITION 4. Given any subalgebra $\mathbf{A} \subseteq \mathbf{M}_n(\mathbb{C})$, the following vector valued function on \mathcal{H}_n :

$$F(\phi)(\alpha) := \frac{\partial}{\partial \phi^*(\alpha)} \|\phi\|^2 S\left(\frac{|\phi\rangle\langle\phi|}{\|\phi\|^2}|\mathbf{A}\right) \quad (23)$$

is defined everywhere except on a submanifold where it becomes singular. Moreover, it is *homogeneous* in the sense that

$$|F(c\phi)\rangle = c|F(\phi)\rangle \quad \forall c \in \mathbb{C}. \quad (24)$$

Remark 3: If $\mathbf{A} \subset \mathbf{M}_n(\mathbb{C})$ is maximally abelian, then

$$\|\phi\|^2 S\left(\frac{|\phi\rangle\langle\phi|}{\|\phi\|^2}|\mathbf{A}\right) = \langle\phi, F(\phi)\rangle. \quad (25)$$

More explicitly, let $\{|k\rangle\}_{k=1}^n$ be an orthonormal basis in \mathbb{C}^n defined by the minimal projections $|k\rangle\langle k|$ of \mathbf{A} . Then, setting $\phi(k) = \langle k|\phi\rangle$, we get

$$S\left(\frac{|\phi\rangle\langle\phi|}{\|\phi\|^2}|\mathbf{A}\right) = -\sum_{k=1}^n \frac{|\phi(k)|^2}{\|\phi\|^2} \log \frac{|\phi(k)|^2}{\|\phi\|^2}. \quad (26)$$

LEMMA 3. *Let the set $\{|\phi_1\rangle, |\phi_2\rangle\}$ be compatible and assume that $|F(\phi_1)\rangle$ and $|F(\phi_2)\rangle$ exist. Then*

$$\langle\phi_1, F(\phi_2)\rangle = \langle F(\phi_1), \phi_2\rangle. \quad (27)$$

Proof: Consider the state $\omega = \frac{1}{2}\langle\phi_1|\cdot|\phi_1\rangle + \frac{1}{2}\langle\phi_2|\cdot|\phi_2\rangle$, with $\|\phi_1\| = \|\phi_2\| = 1$. This state corresponds to a two-dimensional density matrix over the Hilbert space spanned by the vectors $|\phi_1\rangle$ and $|\phi_2\rangle$. In full generality, representing the state ω by a density matrix $\hat{\rho}$, we can consider its positive square root $\sqrt{\hat{\rho}}$ and decompose the state into pure states according to:

$$\omega(\cdot) = \sum_{j=1}^k \langle\sqrt{\hat{\rho}}\psi_j|\cdot|\sqrt{\hat{\rho}}\psi_j\rangle, \quad \omega(\hat{x}) = \sum_{j=1}^k \text{Tr}\{\sqrt{\hat{\rho}}|\psi_j\rangle\langle\psi_j|\sqrt{\hat{\rho}}\hat{x}\}, \quad (28)$$

where $|\psi_j\rangle$ are vectors in \mathcal{H}_n such that $\sum_{j=1}^k |\psi_j\rangle\langle\psi_j| = \mathbf{1}$. Therefore, only linear combinations $\sqrt{\hat{\rho}}|\psi_j\rangle = \gamma_j|\phi_1\rangle + \delta_j|\phi_2\rangle$ can contribute to the optimal decomposition. Let

1. $\gamma_1 = \sqrt{\frac{1}{2} - \epsilon}$, $\delta_1 = \sqrt{\epsilon}$; $\gamma_2 = \sqrt{\epsilon}$, $\delta_2 = -\sqrt{\frac{1}{2} - \epsilon}$.
2. $\gamma_1 = \sqrt{\frac{1}{2} - \epsilon}$, $\delta_1 = i\sqrt{\epsilon}$; $\gamma_2 = \sqrt{\epsilon}$, $\delta_2 = -i\sqrt{\frac{1}{2} - \epsilon}$.

The stationarity of $\mathcal{R}(\omega, \{\frac{1}{2}\langle\phi_1|\cdot|\phi_1\rangle + \frac{1}{2}\langle\phi_2|\cdot|\phi_2\rangle\}, \mathbf{A})$, which follows from the assumed compatibility of $|\phi_1\rangle$ and $|\phi_2\rangle$, requires that the term of order $\sqrt{\epsilon}$ vanishes, which is exactly relation (27). \blacksquare

Remark 4: The above equality is not sufficient, it might correspond to a maximum or to a saddle point. That the supremum can be found by differentiation is guaranteed because we are considering a continuous functional differentiable everywhere on a manifold without boundary except where the vector components vanish. (Notice that we consider the functional as a functional over normalized vectors and not over density matrices.)

Since, according to Theorem 1, for every compatible set, every subset of it is compatible, too, we can generalize the previous result.

LEMMA 4. *If the set $\{|\phi_1\rangle, \dots, |\phi_k\rangle\}$ is compatible, then*

$$\langle \phi_i, F(\phi_j) \rangle = \langle F(\phi_i), \phi_j \rangle \quad \forall i, j = 1, \dots, k. \quad (29)$$

These conditions can be used in a constructive way to find compatible sets. We consider $|\phi_j\rangle$ as points in a $2n$ -dimensional real manifold. With $|\phi_j\rangle$ also $c|\phi_j\rangle$, $\forall c \in \mathbb{C}$, solves (29). The latter is the only obvious redundancy, and, if we fix the norms and the overall phases such that $\| |\phi_j\rangle \| = 1$ and $\phi_j(1) \in \mathbb{R}$, we remain with $2k(n-1)$ unknown vector components in relations among each other through the $k(k-1)$ real equations (29). In addition, the set $\{|\phi_1\rangle, \dots, |\phi_k\rangle\}$ should be optimal for some density matrix ($\text{Tr } \hat{\rho} = 1$)

$$\hat{\rho} = \sum_{l=1}^k \lambda_l |\phi_l\rangle \langle \phi_l|. \quad (30)$$

Thus, the unknown vector components must satisfy $n^2 - 1$ equations for any fixed $\hat{\rho}$. The total number of constraints to be obeyed by the $2k(n-1) + k - 1$ unknowns is then $k^2 - k + n^2 - 1$. If $k > n \geq 2$, the former exceed the latter. Thus, we expect that any solution to the minimization problem can be found only for k at most equal to n . Of course, the result of Lemma 1 teaches us that great care has to be exercised while using this argument.

LEMMA 5. *Assume the set $\{|\phi_1\rangle, \dots, |\phi_n\rangle\}$ to be linearly independent in \mathcal{H}_n and compatible. Let $\{|\phi_1\rangle, \dots, |\phi_n\rangle, |\phi_{n+1}\rangle\}$ and $\{|\phi_1\rangle, \dots, |\phi_n\rangle, |\phi_{n+2}\rangle\}$ be compatible sets, too. Then,*

$$\langle \phi_{n+1}, F(\phi_{n+2}) \rangle = \langle F(\phi_{n+1}), \phi_{n+2} \rangle. \quad (31)$$

Proof: We write $|\phi_{n+1}\rangle = \sum_{j=1}^n \alpha_j |\phi_j\rangle$, $|\phi_{n+2}\rangle = \sum_{k=1}^n \beta_k |\phi_k\rangle$. From the compatibility of $\{|\phi_i\rangle\}_{i=1}^{n+1}$ it follows

$$\left\langle \phi_l, F\left(\sum_{j=1}^n \alpha_j \phi_j\right) \right\rangle = \sum_{j=1}^n \alpha_j \langle F(\phi_l), \phi_j \rangle = \left\langle \phi_l, \sum_{j=1}^n \alpha_j F(\phi_j) \right\rangle, \quad (32)$$

for all $l = 1, \dots, n$. That is, $F(|\sum_{j=1}^n \alpha_j \phi_j\rangle) = \sum_{j=1}^n \alpha_j F(|\phi_j\rangle)$, because of the assumed linear independence of $\{|\phi_j\rangle\}_{j=1}^n$. Analogously, $F(|\sum_{k=1}^n \beta_k \phi_k\rangle) = \sum_{k=1}^n \beta_k F(|\phi_k\rangle)$. Therefore,

$$\begin{aligned} \langle \phi_{n+2}, F(\phi_{n+1}) \rangle &= \sum_{k,j=1}^n \beta_k^* \alpha_j \langle \phi_k, F(\phi_j) \rangle = \sum_{k,j=1}^n \beta_k^* \alpha_j \langle F(\phi_k), \phi_j \rangle \\ &= \langle F(\phi_{n+2}), \phi_{n+1} \rangle. \end{aligned} \quad (33)$$

■

From the preceding lemma we also deduce that any compatible, linearly independent set $\{|\phi_1\rangle, \dots, |\phi_n\rangle\}$ of vectors of \mathcal{H}_n defines a hermitian $n \times n$ matrix $M \in \mathbf{M}_n(\mathbb{C})$

such that

$$\sum_{\beta=1}^n M_{\alpha\beta} \phi_j(\beta) = F(\phi_j)(\alpha). \quad (34)$$

The latter can be completed to a maximally compatible set by finding out all solutions to (34).

Remark 5: B. Kümmerer and R. Werner arrived at the matrix M by optimizing among all finite decompositions into pure states of a given density matrix. Then, M enters as the matrix of Lagrange multipliers and depends on the matrix $\hat{\rho}$ [11]. The assumption that finite decompositions suffice follows from Choquet theory (every $n \times n$ density matrix is a linear convex combination of n fixed 1-dimensional projections). As a consequence, they could draw a conclusion similar to Theorem 1.

We stick to the other strategy, for not all M can be used, and how to find the subclass of permitted M is still an open problem; solving (34) is by far not a trivial task.

2.1. Continuity considerations

Provided the exactness of considerations in the previous section that “in general” only sets of cardinality at most equal to the cardinality of the underlying Hilbert space \mathcal{H}_n are compatible, we have to examine the consequences of possible exceptional points, namely points where the functional (23) is no longer smooth. In the next section we will see that the few examples where the optimal decompositions can be given explicitly give support to these considerations.

We begin with the following

LEMMA 6. *The cardinality of a minimal optimal set for a fixed $\mathbf{A} \subseteq \mathbf{M}(\mathbb{C})$ is lower semicontinuous as a functional on the states over $\mathbf{M}_n(\mathbb{C})$.*

Proof: Let $\{\omega(\gamma)\}_{\gamma \in \mathbb{N}}$ be a sequence of normalized states such that it converges in norm to a normalized state ω : $\lim_{\gamma \rightarrow +\infty} \|\omega(\gamma) - \omega\| = 0$. Let the set $|\phi_i(\gamma)\rangle$, $i = 1, \dots, k_\gamma$, provide an optimal decomposition for $\omega(\gamma)$: $\omega(\gamma) = \sum_j \lambda_j(\gamma) \phi_j$, $\phi_j(\hat{x}) = \langle \phi_j | \hat{x} | \phi_j \rangle$.

Finally, let $\{|\psi_1\rangle, \dots, |\psi_l\rangle\}$ be an optimal set for the limit state ω , that is $\omega(\hat{x}) = \sum_{j=1}^l \lambda_j \psi_j(\hat{x})$, $\psi_j(\hat{x}) = \lambda_j \langle \psi_j | \hat{x} | \psi_j \rangle$, and assume $l > k^* := \max_{\gamma \in \mathbb{N}} k_\gamma$.

By representing $\omega(\gamma)$ and ω as the density matrices $\hat{\rho}(\gamma)$, respectively $\hat{\rho}$, the states ψ_j can be written as $\sqrt{\hat{\rho}} \hat{x}_j \sqrt{\hat{\rho}}$, where $\hat{x}_j \in \mathbf{M}_n(\mathbb{C})$ fulfil $\mathbf{1} \geq \hat{x}_j > 0$, $\sum_{j=1}^l \hat{x}_j = \mathbf{1}$. The same set of positive operators \hat{x}_j can be used to construct a decomposition (in general not optimal) of $\hat{\rho}(\gamma)$:

$$\hat{\rho}(\gamma) = \sum_{j=1}^l \sqrt{\hat{\rho}(\gamma)} \hat{x}_j \sqrt{\hat{\rho}(\gamma)}. \quad (35)$$

That is, $\omega(\gamma) = \sum_{j=1}^l \psi_j(\gamma)$, where

$$\psi_j(\gamma)(\hat{x}) := \text{Tr} \sqrt{\hat{\rho}(\gamma)} \hat{x}_j \sqrt{\hat{\rho}(\gamma)} \hat{x}. \quad (36)$$

Because of the assumed convergence $\omega(\gamma) \rightarrow \omega$, the same uniformity characterizes the limit $\hat{\rho}(\gamma) \rightarrow \hat{\rho}$ when $\gamma \rightarrow +\infty$. Thus

$$\lim_{\gamma \rightarrow +\infty} \mathcal{R}(\omega, \{\psi_j(\gamma)\}, \mathbf{A}) = \mathcal{R}(\omega, \{\psi_j\}, \mathbf{A}) = \mathcal{R}(\omega, \mathbf{A}). \quad (37)$$

Furthermore, as the ψ_j need not form an optimal decomposition for the $\omega(\gamma)$ we have

$$\mathcal{R}(\omega(\gamma), \{\psi_j(\gamma)\}, \mathbf{A}) \geq \mathcal{R}(\omega(\gamma), \{\lambda_j(\gamma)\phi_j\}, \mathbf{A}) \quad \forall \gamma. \quad (38)$$

Even if $\lim_{\gamma \rightarrow +\infty} \lambda_j(\gamma)\phi_j(\gamma)$ does not make any sense, we can select a norm-converging subsequence $\lambda_j(\gamma_n)\phi_j(\gamma_n)$ to a positive linear functional (not a normalized state) $\lambda_j^*\phi_j^*$, $0 < j \leq k^*$. These limit states provide a decomposition $\omega = \sum_{j=1}^l \lambda_j^*\phi_j^*$ that cannot be optimal, because we assumed the cardinality l of the optimal set for ω to be strictly greater than k^* . But then, by using (37) and (38), we arrive at a contradiction:

$$\begin{aligned} \mathcal{R}(\omega, \mathbf{A}) &< \mathcal{R}(\omega, \{\lambda_j^*\phi_j^*\}, \mathbf{A}) \\ &= \lim_{n \rightarrow +\infty} \mathcal{R}(\omega(\gamma_n), \{\lambda_j(\gamma_n)\phi_j(\gamma_n)\}, \mathbf{A}) \\ &\leq \lim_{n \rightarrow +\infty} \mathcal{R}(\omega(\gamma_n), \{\psi_j(\gamma_n)\}, \mathbf{A}) = \mathcal{R}(\omega, \mathbf{A}). \end{aligned} \quad (39)$$

Hence, $l \leq \lim_{\gamma \rightarrow +\infty} k_\gamma$ which proves the lemma. \blacksquare

LEMMA 7. *Let ω be a state over $\mathbf{A} \subset \mathbf{M}_n(\mathbb{C})$ and assume that some optimal decomposition be given by n pure states. Then,*

$$H_{\omega, \mathbf{M}(\mathbf{A})} = \sup_{\mathbf{B} \subset \mathbf{M}'_n \cap \mathbf{B}'} S(\omega \otimes \omega, \omega) | \mathbf{A} \otimes \mathbf{B}. \quad (40)$$

Remark 6: Given $\mathbf{M}_n(\mathbb{C})$, let $\mathcal{H}_{n^2} = \mathcal{H}_n \otimes \mathcal{H}_n$ and $B(\mathcal{H}_{n^2}) = \mathbf{M}_n(\mathbb{C}) \otimes \mathbf{M}_n(\mathbb{C})$. In this way, ω , which is a density matrix $\hat{\rho}$ with eigenvalues ρ_k and eigenvectors $|r_k\rangle$ on $\mathbf{M}_n(\mathbb{C})$, gets purified as a state on $B(\mathcal{H}_{n^2})$. It corresponds to the vector state $|\sqrt{\hat{\rho}}\rangle = \sum_{k=1}^n \sqrt{\rho_k} |r_k\rangle \otimes |r_k\rangle \in \mathcal{H}_{n^2}$ such that

$$\omega(\hat{a}) = \langle \sqrt{\hat{\rho}} | \hat{a} \otimes \mathbb{1} | \sqrt{\hat{\rho}} \rangle. \quad (41)$$

The commutant $\mathbf{M}_n(\mathbb{C})'$ that appears in (40) is isomorphic to $\mathbb{1} \otimes \mathbf{M}_n(\mathbb{C})$. Therefore the supremum in the same formula has to be taken over all subalgebras $\mathbf{A} \otimes \mathbf{B} \subset B(\mathcal{H}_{n^2})$, where $\mathbf{B} \subseteq \mathbb{1} \otimes \mathbf{M}_n$ being contained in its own commutant must be abelian. Furthermore, via (41), ω also defines a state on $\mathbf{M}_n(\mathbb{C})'$, and in turn, a state on $B(\mathcal{H}_{n^2})$: $\omega \otimes \omega(\hat{a} \otimes \hat{a}') = \omega(\hat{a})\omega(\hat{a}')$.

Therefore, using (2) the argument of the supremum on the right-hand side of the above expression reads

$$\text{Tr} \omega | \mathbf{A} \otimes \mathbf{B} \log \omega | \mathbf{A} \otimes \mathbf{B} + S(\omega | \mathbf{A}) + S(\omega | \mathbf{B}). \quad (42)$$

Proof: We represent ω as a density matrix $\hat{\rho}$ that, by assumption, is optimally decomposed by a set of n pure states. Thus $\hat{\rho} = \sum_{i=1}^n \sqrt{\hat{\rho}} \hat{x}_i \sqrt{\hat{\rho}}$, with the operators \hat{x}_i

being minimal projections \hat{P}_i (in \mathbf{M}_n) of a maximally abelian algebra \mathcal{P} , $\hat{P}_i\hat{P}_j = \delta_{ij}\hat{P}_j$. According to the previous remark, selecting any abelian (not necessarily maximally abelian) subalgebra \mathbf{B} with orthogonal projections (not necessarily minimal in \mathbf{M}_n) \hat{Q}_j , $j = 1, \dots, k \leq n$, such that $\sum_{j=1}^k \hat{Q}_j = \mathbb{1}$, $\omega|_{\mathbf{A} \otimes \mathbf{B}}$ can be represented as the following density matrix on \mathcal{H}_{n^2} :

$$\hat{\rho}|_{\mathbf{A} \otimes \mathbf{B}} = \sum_{j=1}^k \sqrt{\hat{\rho}}\hat{Q}_j\sqrt{\hat{\rho}}|_{\mathbf{A} \otimes \hat{Q}_j}. \quad (43)$$

Then, (42) reads

$$-\sum_{j=1}^k S\left(\sqrt{\hat{\rho}}\hat{Q}_j\sqrt{\hat{\rho}}|_{\mathbf{A}}\right) - \sum_{j=1}^k \text{Tr} \hat{\rho}\hat{Q}_j \log \text{Tr} \hat{\rho}\hat{Q}_j + S(\omega|_{\mathbf{A}}). \quad (44)$$

By going to the normalized states $\omega_j(\cdot) = \text{Tr} \sqrt{\hat{\rho}}\hat{Q}_j\sqrt{\hat{\rho}}(\cdot)/\text{Tr} \hat{\rho}\hat{Q}_j$, we get

$$(42) = S(\omega|_{\mathbf{A}}) - \mathcal{R}(\omega, \{(\text{Tr} \hat{\rho}\hat{Q}_j)\omega_j\}, \mathbf{A}). \quad (45)$$

■

CONJECTURE: We deem it true that

$$H_{\omega, \mathbf{M}}(\mathbf{A}) = \sup_{\mathbf{B} \subset \mathbf{M}' \cap \mathbf{B}'} S(\omega \otimes \omega, \omega)|_{\mathbf{A} \otimes \mathbf{B}} \quad (46)$$

holds in full generality, namely, also for algebras \mathbf{M} that are not finite dimensional as supposed in (40), but that are, instead, the strong operator closure of an increasing sequence of finite dimensional subalgebras $\mathbf{M}_n \subset \mathbf{M}$.

Indeed, for finite dimensional subalgebras we could not find any counterexample. In the infinite dimensional case, we cannot expect that there exists any optimal decomposition into normal states. Therefore, we equally expect not to have a replacement for the compatibility conditions. Nevertheless, in order to pass from formula (40) to formula (46) the continuity considerations are applicable.

LEMMA 8.

$$H_{\omega, \mathbf{A}}(\mathbf{M}) = \inf_n H_{\omega, \mathbf{A}}(\mathbf{M}_n) = \lim_{n \rightarrow \infty} H_{\omega, \mathbf{A}}(\mathbf{M}_n). \quad (47)$$

Proof: Every decomposition corresponds to a choice of positive trace class operators of the form $\sqrt{\hat{\rho}}\hat{x}\sqrt{\hat{\rho}}$ for finite dimensional algebras, otherwise we can write them as

$$\langle \Omega | \hat{x} \Delta_n^{1/2} \hat{a} | \Omega \rangle \quad \forall \hat{x} \in \mathbf{M}_n, \quad \forall \hat{a} \in \mathbf{A} \subset \mathbf{M}_n, \quad (48)$$

where we have considered the GNS representation based on ω and the associated modular operator defined as Friederichs extension of the quadratic form

$$d_n(\phi, \phi) := \inf_{|\phi\rangle = \hat{a}|\Omega\rangle, \hat{a} \in \mathbf{M}_n(\mathbb{C})} \omega(\hat{a}\hat{a}^*). \quad (49)$$

For $n \rightarrow \infty$, Δ_n converges in the strong resolvent sense to Δ [4, 8]. If $\hat{a} \in \mathbf{M}_n$, then $\hat{a}|\Omega\rangle \in \text{dom}(\Delta_n^{1/2})$. Therefore, $st - \lim \Delta_n^{1/2}\hat{a} = \Delta^{1/2}\hat{a}$.

Assume $\|(\Delta_n^{1/2} - \Delta^{1/2})\hat{a}|\Omega\rangle\| \leq \epsilon\|\hat{a}\|$, for all \hat{a} in the finite dimensional subalgebra \mathbf{A} . Let $\hat{x}_i(n)$ give the optimal decomposition for ω relative to \mathbf{M}_n . Then they are nearly optimal relative to \mathbf{M} because the individual states with respect to Δ and Δ_n differ only by ϵ .

In order to conclude the argument for the conjecture, we believe now that the $\hat{x}_i(n)$ form an abelian subalgebra \mathbf{B} of \mathbf{M}_n that can be embedded into \mathbf{M}' . Consequently, the optimum result with respect to \mathbf{M} would be also reachable by restriction to abelian subalgebras as conjectured. ■

3. Examples

We will construct examples of maximally compatible sets for various $\mathbf{A} \subseteq \mathbf{M}_n(\mathbb{C})$. The purpose of these examples is to get a feeling whether the relations (29) are independent constraints so that maximally compatible sets have the cardinality n , that is the dimension of any maximally abelian subalgebra of $\mathbf{M}_n(\mathbb{C})$. If this is not the case, we can still look for other conditions that might allow us to fix the cardinality of compatible sets.

EXAMPLE 1. $\mathbf{A} \subset \mathbf{M}_2(\mathbb{C})$.

We consider an abelian subalgebra \mathbf{A} that we choose to be generated by the identity and the Pauli matrix σ_z (\mathbf{A} is uniquely determined up to unitary equivalences). Its elements are $a = \alpha + \gamma\sigma_z$, whereas those of $\mathbf{M}_2(\mathbb{C})$ will be of the form $m = \beta + \vec{n} \cdot \vec{\sigma}$. We represent vectors (states) and the action of the functional (23) on them with respect to the orthonormal basis associated with \mathbf{A} . Let us take

$$|\phi_1\rangle = \begin{pmatrix} \cos \alpha \\ e^{i\delta} \sin \alpha \end{pmatrix}, \quad |F(\phi_1)\rangle = \begin{pmatrix} \cos \alpha \log \cos^2 \alpha \\ e^{i\delta} \sin \alpha \log \sin^2 \alpha \end{pmatrix}, \quad (50)$$

$$|\phi_2\rangle = \begin{pmatrix} \cos \beta \\ e^{i\gamma} \sin \beta \end{pmatrix}, \quad |F(\phi_2)\rangle = \begin{pmatrix} \cos \beta \log \cos^2 \beta \\ e^{i\gamma} \sin \beta \log \sin^2 \beta \end{pmatrix}. \quad (51)$$

Thus, equation (27) becomes

$$\begin{aligned} & \cos \alpha \cos \beta \log \cos^2 \alpha + e^{i(\delta-\gamma)} \sin \alpha \sin \beta \log \sin^2 \alpha \\ & = \cos \alpha \cos \beta \log \cos^2 \beta + e^{i(\delta-\gamma)} \sin \alpha \sin \beta \log \sin^2 \beta. \end{aligned} \quad (52)$$

We distinguish two possibilities: either $\delta \neq \gamma$, or $\delta = \gamma$.

In the first case, the only solutions to (52) are $\sin \alpha = \pm \sin \beta$. Since $\phi_{1,2}|\mathbf{A} = \phi|\mathbf{A}$, all states $\phi = \lambda_1\langle\phi_1| \cdot |\phi_1\rangle + \lambda_2\langle\phi_2| \cdot |\phi_2\rangle$, $\lambda_{1,2} > 0$, $\lambda_1 + \lambda_2 = 1$, will have $S(\phi|\mathbf{A}) = \lambda_1 S(\phi_1|\mathbf{A}) + \lambda_2 S(\phi_2|\mathbf{A})$. Therefore, $S(\phi, \phi_{1,2})|\mathbf{A} = 0$, which is the minimal value achievable, while we are interested in its maximum.

Thus, we must choose the second alternative in which case by denoting $x = \cos \alpha$ and $y = \sin \alpha$ equation (52) reads $f(x, y, \beta) = 0$, where

$$f(x, y, \beta) := A(\beta)x - B(\beta)x \log x^2 + C(\beta)y - D(\beta)y \log y^2, \quad (53)$$

$$A(\beta) = \cos \beta \log \cos^2 \beta, \quad B(\beta) = \cos \beta, \quad (54)$$

$$C(\beta) = \sin \beta \log \sin^2 \beta, \quad D(\beta) = \sin \beta. \quad (55)$$

Obviously, we seek solutions (x, y) of (53) that satisfy the additional condition $x^2 + y^2 = 1$. It is easily verified that we have the following solutions $(x(\beta), y(\beta))$:

$$(\cos \beta, \sin \beta), \quad (\cos \beta, -\sin \beta), \quad (56)$$

$$(-\cos \beta, \sin \beta), \quad (-\cos \beta, -\sin \beta), \quad (57)$$

$$(\sin \beta, \cos \beta), \quad (-\sin \beta, -\cos \beta). \quad (58)$$

The first four couples provide decompositions that correspond, as before, to the minimum of $S(\phi, \phi_i) | \mathbf{A}$ whereas both the last two give rise to the nontrivial set

$$|\phi_1\rangle = \begin{pmatrix} \sin \beta \\ e^{i\delta} \cos \beta \end{pmatrix}, \quad |\phi_2\rangle = \begin{pmatrix} \cos \beta \\ e^{i\delta} \sin \beta \end{pmatrix}. \quad (59)$$

Notice that the whole of the state space over $\mathbf{M}_2(\mathbb{C})$ is covered by the convex combinations $\lambda \langle \phi_1 | \cdot | \phi_1 \rangle + (1 - \lambda) \langle \phi_2 | \cdot | \phi_2 \rangle$ by varying $0 \leq \lambda \leq 1$ and $\beta, \delta \in [0, 2\pi]$.

In the appendix it is proved that there are no other solutions than those in (56)–(58), whence the second solution will correspond to the optimal one and the cardinality of the maximal optimal sets will be 2.

EXAMPLE 2. $\mathbf{A} \subseteq \mathbf{M}_3(\mathbb{C})$.

We will restrict ourselves to a particular class of states on $\mathbf{M}_3(\mathbb{C})$, those invariant under all possible permutations of the vectors $|k\rangle$, $k = 1, 2, 3$, in the orthonormal basis defined by the maximally abelian subalgebra \mathbf{A} .

As up to a certain point we need not limit ourselves to $n = 3$, let $P_k = |k\rangle\langle k|$, $k = 1, \dots, n$, be the minimal projections of maximally abelian \mathbf{A} in $\mathbf{M}_n(\mathbb{C})$ and let us call “completely symmetric” with respect to the fixed orthonormal basis any density matrix of the form

$$\rho = \begin{pmatrix} 1/n & z & z & \dots & z \\ z & 1/n & z & \dots & z \\ z & z & 1/n & \dots & z \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & z & 1/n \end{pmatrix}, \quad z \in \mathbb{R}. \quad (60)$$

All these are uniquely determined by a real parameter z and are invariant under all possible permutations:

$$|k\rangle \mapsto |\pi(k)\rangle, \quad \pi: (1, 2, \dots, n) \mapsto (\pi(1), \pi(2), \dots, \pi(n)). \quad (61)$$

The range of the possible values of z is fixed by the request that the eigenvalues $\rho_i(z)$ of $\hat{\rho}(z)$ be $0 \leq \rho_i(z) \leq 1$. As the latter are $\rho_1(z) = 1/n + z(n-1)$ and $\rho_2(z) = \dots = \rho_n(z) = 1/n - z$, we get $-1/n(n-1) \leq z \leq 1/n$.

We are now interested in decomposing $\hat{\rho}(z)$ in such a way that $\mathcal{R}(\hat{\rho}(z), \mathbf{A})$ in (5) be attained. Notice, however, that getting control over optimal decompositions of ρ immediately involves controlling optimal decompositions of all $\hat{U}\hat{\rho}\hat{U}^*$, with \hat{U} being unitaries in \mathbf{A} .

According to what was discussed in point 4 after Definition 1, we can limit our considerations to decompositions into pure states:

$$\hat{\rho}(z) = \sum_j \lambda_j |\phi_j\rangle\langle\phi_j|. \quad 0 < \lambda_j < 1. \quad \sum_j \lambda_j = 1. \quad (62)$$

The next lemma makes use of the considerations mentioned above, essentially about the concavity of the von Neumann entropy, and shows that for a restricted class of completely symmetric $\hat{\rho}(z)$ the search for optimal decompositions can be restricted to those given in terms of the real one-dimensional projections.

LEMMA 9. *Given any density matrix $\hat{\rho}(z) \in \mathbf{M}_n(\mathbb{C})$ as in (60), there always exists an optimal decomposition of $\hat{\rho}(z)$ in terms of the real one-dimensional projections $\hat{P}_j = |\phi_j\rangle\langle\phi_j|$ that gives $\mathcal{R}(\hat{\rho}(z), \mathbf{A})$, where $\mathbf{A} \subset \mathbf{M}_n(\mathbb{C})$ is the maximally abelian subalgebra that fixes the representation (60).*

Proof: Let $\hat{\rho}(z) = \sum_j \lambda_j \hat{\sigma}_j$ be a decomposition into non-real projections $\hat{\sigma}_j$. Then, let $\hat{\sigma}_j^* \in \mathbf{M}_n(\mathbb{C})$ be such that in the orthonormal basis defined by \mathbf{A}

$$(\hat{\sigma}_j^*)_{ik} = (\hat{\sigma}_j)_{ik}^*, \quad i \neq k, \quad (\hat{\sigma}_j^*)_{ii} = (\hat{\sigma}_j)_{ii}, \quad i, k = 1, \dots, n. \quad (63)$$

Since $(\hat{\rho}(z))_{ik} = z = \sum_j \lambda_j (\hat{\sigma}_j)_{ik} = \sum_j \lambda_j (\hat{\sigma}_j^*)_{ik}$, we get that $\hat{\rho}(z) = \sum_j \lambda_j \hat{\sigma}_j^*$, too.

Moreover, $S(\hat{\sigma}_j^* | \mathbf{A}) = S(\hat{\sigma}_j | \mathbf{A}) = S(\frac{1}{2}(\hat{\sigma}_j + \hat{\sigma}_j^*) | \mathbf{A})$, hence

$$\mathcal{R}(\hat{\rho}(z), \{\lambda_j \hat{\sigma}_j\}, \mathbf{A}) = \mathcal{R}(\hat{\rho}(z), \{\lambda_j \hat{\rho}_j\}, \mathbf{A}), \quad (64)$$

where $\hat{\rho}_j = \frac{1}{2}(\hat{\sigma}_j + \hat{\sigma}_j^*)$ are not pure states and can thus be further decomposed. For instance, take their spectral decompositions (they are rank two operators) $\hat{\rho}_j = r_{j1} \hat{R}_{j1} + r_{j2} \hat{R}_{j2}$ that must involve real eigenprojectors \hat{R}_{jl} , $l = 1, 2$.

Strict concavity of the von Neumann entropy (inequality (6)) tells us that either an optimal decomposition is already given in terms of real projections, or, if this is not the case and the optimal decomposition involves generic projections, we can always find one with the same optimal contribution $\mathcal{R}(\hat{\rho}(z), \mathbf{A})$, but using real pure states. In fact, from (6)

$$\mathcal{R}(\hat{\rho}(z), \{\lambda_j \hat{\sigma}_j\}, \mathbf{A}) \geq \mathcal{R}(\hat{\rho}(z), \{\lambda_j r_{ji} \hat{R}_{ji}\}, \mathbf{A}). \quad (65)$$

■

We turn now to the case $n = 3$ where we have the following

LEMMA 10. *Given a density matrix $\hat{\rho}(z) \in \mathbf{M}_3(\mathbb{C})$ completely symmetric with respect to a maximally abelian subalgebra $\mathbf{A} \subset \mathbf{M}_3(\mathbb{C})$ as in (60), there exists a one-parameter family of real vectors $|\phi(u, z)\rangle \in \mathbb{C}^3$, $u \in [0, 2\pi]$, such that*

$$\hat{\rho}(z) = \frac{1}{3} \sum_{j=1}^3 \hat{U}_{\pi_j} |\phi(u, z)\rangle \langle \phi(u, z)| \hat{U}_{\pi_j}^*, \quad (66)$$

where the \hat{U}_{π_j} are the unitaries in $\mathbf{M}_3(\mathbb{C})$ that implement the cyclic permutations of the basis vectors defined by \mathbf{A} , that is

$$\hat{U}_{\pi_j} |k\rangle = |\pi_j(k)\rangle \quad \text{and} \quad \begin{cases} \pi_1: (1, 2, 3) \mapsto (1, 2, 3), \\ \pi_2: (1, 2, 3) \mapsto (3, 1, 2), \\ \pi_3: (1, 2, 3) \mapsto (2, 3, 1). \end{cases} \quad (67)$$

Proof: When $n = 3$, the real parameter $z \in [-\frac{1}{6}, \frac{1}{3}]$. Define the quantities

$$a = \sqrt{1 + 6z}, \quad b = 2\sqrt{1 - 3z}, \quad (68)$$

and the vectors, whose components with respect to the basis $|k\rangle$, $k = 1, 2, 3$, defined by \mathbf{A} are:

$$\begin{aligned} \phi_1(u, z) &= \frac{1}{3} [a + b \cos(u - \frac{1}{3}\pi)], \\ \phi_2(u, z) &= \frac{1}{3} [a + b \cos(u + \frac{1}{3}\pi)], \quad u \in [0, 2\pi], \\ \phi_3(u, z) &= \frac{1}{3} [a - b \cos u] \end{aligned} \quad (69)$$

■

Remark 7: Given any vector $\phi \in \mathbb{C}^n$, in order to get out of it a completely symmetric density matrix with respect to a maximally abelian subalgebra $\mathbf{A} \subset \mathbf{M}_n(\mathbb{C})$, one would in principle use all the unitaries $\hat{U}_\pi \in \mathbf{M}_n(\mathbb{C})$ implementing all the possible permutations π of the basis vectors provided by \mathbf{A} and construct

$$\hat{\rho}(z) = \frac{1}{n!} \sum_{\pi} \hat{U}_\pi |\phi\rangle \langle \phi| \hat{U}_\pi^*. \quad (70)$$

In fact, when $n = 3$; not all permutations are necessary, but only the cyclic ones.

From Lemma 9 we know that the class of decompositions of $\hat{\rho}(z)$ consisting of real pure states always contains at least one optimal decomposition with respect to the maximally abelian \mathbf{A} with respect to which $\hat{\rho}(z)$ is completely symmetric. What turns out is that, in the 3-dimensional case, the subclass of the latter containing only three cyclically permuted real projections already provides optimal decompositions of $\hat{\rho}(z)$ with respect to \mathbf{A} .

We start with fixing the notation.

We will denote by $\hat{P}_j(u, z)$ the projections onto the vector states $\hat{U}_{\pi_j}|\phi(u, z)\rangle$ that contribute to (66), and by

$$\begin{aligned} r(u, z) &= \mathcal{R}(\hat{\rho}(z), \{\frac{1}{3}\hat{P}_j(u, z)\}, \mathbf{A}) \\ &= \frac{1}{3} \sum_{j=1}^3 S\left(\hat{P}_j(u, z)|\mathbf{A}\right) = - \sum_{k=1}^3 \phi_k(u, z)^2 \log \phi_k(u, z)^2 \end{aligned} \quad (71)$$

their entropic contribution as a decomposition of $\hat{\rho}(z)$. Notice, in fact, that $\hat{P}_j(u, z)|\mathbf{A} = |\phi(u, z)\rangle\langle\phi(u, z)||\mathbf{A}$, $\forall j = 1, 2, 3$, as \mathbf{A} is mapped into itself by the whole of the \hat{U}_{π} 's.

Finally, we will consider

$$r^*(z) = \inf \{r(u, z), \quad u \in [0, 2\pi]\}, \quad (72)$$

namely the minimum of the entropic contributions coming out of the special subclass of cyclically permuted real decompositions (66) and denote by $u^*(z)$ the point at which it is attained: $r^*(z) = r(u^*(z), z)$.

Notice that we cannot exclude that more than one vector $|\phi(u_l^*(z), z)\rangle$, $l = 1, 2, \dots$, might optimally contribute with $r^l(u_l^*(z), z) = r^*(z)$.

Before coping with the various possibilities we introduce the following

LEMMA 11. *Let $\hat{\rho}(z) \in \mathbf{M}_3(\mathbb{C})$ and $\mathbf{A} \in \mathbf{M}_3(\mathbb{C})$ as in the previous considerations. Assume $r^*(z)$ in (72) to be a convex function of $z \in [-\frac{1}{6}, \frac{1}{3}]$:*

$$r^*\left(\sum_k \lambda_k z_k\right) \leq \sum_k \lambda_k r^*(z_k), \quad (73)$$

for all $0 \leq \lambda_k \leq 1$ with $\sum_k \lambda_k = 1$ and all $z_k \in [-\frac{1}{6}, \frac{1}{3}]$. Then, $\mathcal{R}(\hat{\rho}(z), \mathbf{A}) = r^*(z)$, $\forall z \in [-\frac{1}{6}, \frac{1}{3}]$.

Proof: Let us consider any decomposition $\hat{\rho}(z) = \sum_j \lambda_j \hat{\sigma}_j$ of $\hat{\rho}(z)$ into real one-dimensional projections $\hat{\sigma}_j \in \mathbf{M}_3(\mathbb{C})$. Since $\hat{\rho}(z) = \hat{U}_{\pi} \hat{\rho}(z) \hat{U}_{\pi}^*$, for all permutations π of the basis vectors $|k\rangle$, $k = 1, 2, 3$, provided by \mathbf{A} we can use the cyclic projections and decompose it as follows

$$\hat{\rho}(z) = \sum_j \lambda_j \frac{1}{3} \sum_{k=1}^3 \hat{U}_{\pi_k} \hat{\sigma}_j \hat{U}_{\pi_k}^*. \quad (74)$$

According to Lemma 10, the density matrices $\frac{1}{3} \sum_{k=1}^3 \hat{U}_{\pi_k} \hat{\sigma}_j \hat{U}_{\pi_k}^*$ are completely symmetric with respect to \mathbf{A} and identified by a real parameter z_k , $k = 1, 2, 3$. Let them be denoted by $\hat{\rho}_k(z_k)$.

Let us now assume the initial decomposition $\hat{\rho}(z) = \sum_j \lambda_j \hat{\sigma}_j$ to be optimal with respect to \mathbf{A} , that is

$$\mathcal{R}(\hat{\rho}(z), \{\lambda_j \hat{\sigma}_j\}, \mathbf{A}) = \mathcal{R}(\hat{\rho}(z), \mathbf{A}) = \sum_j \lambda_j S(\hat{\sigma}_j|\mathbf{A}). \quad (75)$$

Because of Theorem 1, it thus turns out that the decompositions

$$\hat{\rho}_j(z_j) = \frac{1}{3} \sum_{k=1}^3 \hat{U}_{\pi_k} \hat{\sigma}_j \hat{U}_{\pi_k}^* \quad (76)$$

must be optimal for $\hat{\rho}_j(z_j)$ and therefore the contributors $\hat{U}_{\pi_k} \hat{\sigma}_j \hat{U}_{\pi_k}^*$, $k = 1, 2, 3$, must coincide with some triple $\hat{P}_k(u_1^*(z), z)$, $k = 1, 2, 3$, that minimizes $r(u, z)$ in (71) giving $r^*(z)$ in (72). That is,

$$\frac{1}{3} \sum_{k=1}^3 S(\hat{U}_{\pi_k} \hat{\sigma}_j \hat{U}_{\pi_k}^* | \mathbf{A}) = S(\hat{\sigma}_j | \mathbf{A}) = r^*(z_j). \quad (77)$$

Using now the assumed convexity of $r^*(z)$ in $[-\frac{1}{6}, \frac{1}{3}]$, and (75)–(77), we get

$$r^*(z) \geq \mathcal{R}(\hat{\rho}(z), \mathbf{A}) = \sum_j \lambda_j r^*(z_j) \geq r^*\left(\sum_j \lambda_j z_j\right) = r^*(z). \quad (78)$$

Indeed, the first inequality follows because the decomposition of $\hat{\rho}(z)$ that minimizes $r(u, z)$ in (71) need not minimize $\mathcal{R}(\hat{\rho}(z), \{\mu_l \hat{\rho}_l\}, \mathbf{A})$. The last equality is a consequence of (74).

Concluding, $r^*(z) = \mathcal{R}(\hat{\rho}(z), \mathbf{A})$ and every decomposition of $\hat{\rho}(z)$ as in (66) that minimizes $r(u, z)$ in (71) is already an optimal decomposition in the larger sense of Definition 2. ■

The convexity of $r^*(z)$ in $[-\frac{1}{6}, \frac{1}{3}]$ is controlled by the numerical means and Fig. 2 supports the conjecture that $r^*(z)$ is indeed convex. As far as the minima of $r(u, z)$ in (71) are concerned, again by the numerical means, the following two possibilities are discovered (see Fig. 1).

There exists a bifurcation value $z^* < 0$:

$$z^* = \frac{1}{3} \frac{(t^*)^2 - 4t^* + 3}{2(t^*)^2 - 4t^* + 3} < 0, \quad (79)$$

where $e^{-t^*} = t^* - 1$, such that, for all $z \geq z^*$, $r(u, z)$ has three absolute minima at $u_1^*(z) = \frac{1}{3}\pi$, $u_2^*(z) = \pi = u_1^*(z) + \frac{2}{3}\pi$ and $u_3^*(z) = \frac{5}{3}\pi = 2\pi - u_1^*(z)$.

On the contrary, for all $z \leq z^*$, there are six z -dependent absolute minima of $r(u, z)$. The first two are attained at $0 < u_1^*(z) = \frac{1}{3}\pi - \alpha(z)$ and $u_2^*(z) = \frac{1}{3}\pi + \alpha(z)$. For $z = -\frac{1}{6}$ we have $u_1^*(z) = \frac{1}{6}\pi$, $u_2^*(z) = \frac{1}{2}\pi$.

The other four are related to these latter as are the second two to the first one for $z > z^*$, namely $u_3^*(z) = u_1^*(z) + \frac{2}{3}\pi$, $u_4^*(z) = u_2^*(z) + \frac{2}{3}\pi$ and, finally, $u_5^*(z) = 2\pi - u_1^*(z)$, $u_6^*(z) = 2\pi - u_2^*(z)$.

Indeed, the functional $r(u, z)$ in (71) is invariant under $u \mapsto u + \frac{2}{3}\pi$ and $u \mapsto 2\pi - u$.

For $z \geq z^*$, the three minima are associated with optimal decompositions that are just the cyclic permutation of the other and thus do not differ; the optimal vectors have

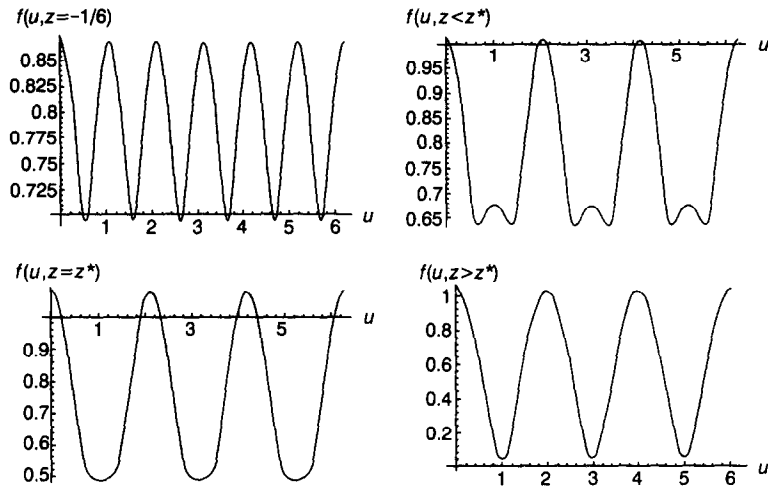


Fig. 1: Various minima of $f(u, z)$ with varying $-1/6 \leq z \leq 1/3$.

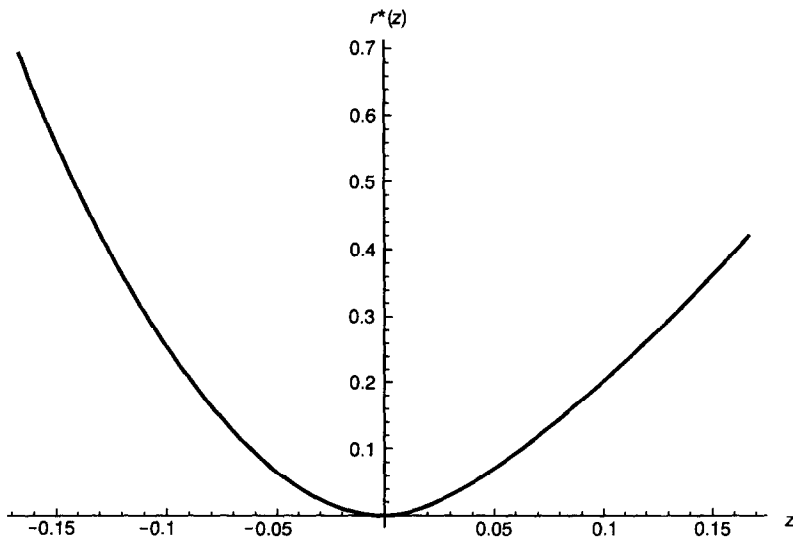


Fig. 2: Convexity of $r^*(z)$.

two equal components. Therefore, our conjecture on the minimal number of optimal decomposers is confirmed.

For $z \leq z^*$, the first two minima correspond to two different optimal decompositions of the kind (66), whereas the other minima contribute with optimal decompositions that come out of the first two by cyclic permutations. It thus turns out that there exist optimal decompositions which contain just three decomposers in agreement with our hypothesis, but there is more of them than one. Together with the two provided by the two minima at $u_1^*(z) = \frac{1}{3}\pi \mp \alpha(z)$, we also have the whole of their convex span.

Appendix

In order to show that the solutions (56)–(58) to equation (52) are the only ones, we follow the idea expounded in [12, Chapter 18]. Precisely, we consider the complex functions

$$F_1(z_1, z_2) := A(\beta)z_1 - B(\beta)z_1 \log z_1^2 - C(\beta)z_2 + D(\beta)z_2 \log z_2^2, \quad (80)$$

$$F_2(z_1, z_2) := z_1^2 + z_2^2 - 1. \quad (81)$$

we choose definite branches of the logarithms in order to make them analytic functions and form the homotopy $H(z_1, z_2, t) := (H_1(z_1, z_2, t), H_2(z_1, z_2, t))$, $t \in [0, 1]$:

$$H_1(z_1, z_2, t) := (1-t)(z_1^2 - 1) + tF_1(z_1, z_2), \quad (82)$$

$$H_2(z_1, z_2, t) := (1-t)(z_2^3 - 1) + tF_2(z_1, z_2). \quad (83)$$

Notice that for $t = 0$, $H(z_1, z_2, 0) = 0$ has 6 solutions and that, for $t = 1$, $H(z_1, z_2, 1) = 0$ is solved by the 6 (real) solutions to (56)–(58).

Via the implicit function theorem, the manifold of solutions to $H(z_1, z_2, t) = 0$ consists of the paths connecting the solutions to $H(z_1, z_2, 0) = 0$ with those to $H(z_1, z_2, 1) = 0$. Now, splitting H_i and z_i in their real and imaginary parts, $\hat{H}_{i1}(z_1, z_2, t)$, $\hat{H}_{i2}(z_1, z_2, t)$ and x_i, y_i , respectively, the determinant of the Jacobian

$$\hat{H}'_w(w, t) := \left[\partial_{w_j} \left(\hat{H}_{i1}(w, t), \hat{H}_{i2}(w, t) \right) \right], \quad w = (x_1, y_1, x_2, y_2), \quad (84)$$

is always ≥ 0 because of the Cauchy–Riemann relations. It happens to be strictly positive at $t = 0$, which means that no more than one path can start from the solutions to $H(z_1, z_2, t) = 0$ at $t = 0$, namely from the 6 solutions to $z_1^2 = 1$ and $z_2^3 = 1$.

Therefore, if there were more than 6 solutions at $t = 1$, then some of the paths issuing from them should diverge somewhere in between 0 and 1. This is excluded. In fact,

$$\lim_{\|z_1\| \rightarrow +\infty} \frac{H_1(z_1, z_2, t)}{z_1^2 - 1} = 1 - t, \quad (85)$$

$$\lim_{\|z_2\| \rightarrow +\infty} \frac{H_2(z_1, z_2, t)}{z_2^3 - 1} = 1 - t. \quad (86)$$

Thus, the diverging solutions to $H(z_1, z_2, t) = 0$ are forbidden for all $0 \leq t < 1$.

REFERENCES

- [1] A. Wehrl: *Rev. Mod. Phys.* **50** (1978), 221.
- [2] M. Ohya and D. Petz: *Quantum Entropy and its Use*, Springer, Berlin 1993.
- [3] I. P. Cornfeld, S. W. Fomin and Ya. G. Sinai: *Ergodic Theory*, Springer, Berlin 1982.
- [4] H. Narnhofer: *Rev. Math. Phys.* **6** (1994), 1127.
- [5] H. Narnhofer and W. Thirring: *Fizika* **17** (1985), 257.
- [6] A. Connes: *C. R. Acad. Sci. Ser. I. Math.* **301** (1985), 1.
- [7] A. Connes, H. Narnhofer and W. Thirring: *Commun. Math. Phys.* **112** (1987), 681.
- [8] A. Connes, H. Narnhofer and W. Thirring: in *Recent Developments in Mathematical Physics*, H. Mitter and L. Pittner (Eds.), Springer, Berlin 1987, p. 102.
- [9] F. Benatti and T. Grava: to appear in the *Proceedings of the Udine Conference in Memory of A. Frigerio*, 1994.
- [10] H. Kosaki: *J. Optim. Theory* **16** (1986), 335.
- [11] B. Kümmerer and R. Werner: private communication.
- [12] C. B. Garcia and W. I. Zangwill: *Pathways to Solutions, Fixed Points, and Equilibria*, Prentice-Hall Series in Computational Mathematics, 1981.