On a Block Matrix Inequality quantifying the Monogamy of the Negativity of Entanglement

Koenraad M.R. Audenaert

Department of Mathematics, Royal Holloway University of London, Egham TW20 0EX, United Kingdom

Department of Physics and Astronomy, Ghent University, S9, Krijgslaan 281, B-9000 Ghent, Belgium

Abstract

We convert a conjectured inequality from quantum information theory, due to He and Vidal, into a block matrix inequality and prove a very special case. Given n matrices A_i, $i = 1, \ldots, n$, of the same size, let Z_1 and Z_2 be the block matrices $Z_1 := (A_j A_i^*)_{i,j=1}^n$ and $Z_2 := (A_j^* A_i)_{i,j=1}^n$. Then the conjectured inequality is

$$(||Z_1||_1 - \operatorname{Tr} Z_1)^2 + (||Z_2||_1 - \operatorname{Tr} Z_2)^2 \leq \left(\sum_{i \neq j} ||A_i||_2 ||A_j||_2 \right)^2,$$

where $|| \cdot ||_1$ and $|| \cdot ||_2$ denote the trace norm and the Hilbert-Schmidt norm, respectively. We prove this inequality for the already challenging case $n = 2$ with $A_1 = \mathbb{I}$.

1 Introduction

Quantum Information Theory (QIT), a recent physical theory combining concepts of information theory with quantum mechanics, has proven to be a rich source of challenging matrix analysis problems [1,2]. In this paper one such problem is presented and some progress towards its resolution is reported.

Consider a set of n given general $n_1 \times n_2$ matrices A_i, and with them form the two $n \times n$ block matrices

$$Z_1 := (A_j A_i^*)_{i,j=1}^n = \begin{pmatrix} A_1 A_1^* & A_2 A_1^* & \ldots \\ A_1 A_2^* & A_2 A_2^* & \ldots \\ \vdots & \vdots & \end{pmatrix}$$

and

$$Z_2 := (A_j^* A_i)_{i,j=1}^n = \begin{pmatrix} A_1^*A_1 & A_2^*A_1 & \dots \\ A_1^*A_2 & A_2^*A_2 & \dots \\ \vdots & \vdots & \end{pmatrix}.$$

These two matrices are Hermitian, but not in general positive semidefinite. We wish to investigate whether the following inequality holds:

$$(||Z_1||_1 - \operatorname{Tr} Z_1)^2 + (||Z_2||_1 - \operatorname{Tr} Z_2)^2 \leq \left(\sum_{i\neq j} ||A_i||_2 ||A_j||_2\right)^2. \tag{1}$$

Here $||\cdot||_1$ and $||\cdot||_2$ denote the trace norm (Schatten 1-norm) and Frobenius norm (Schatten 2-norm), respectively.

This inequality is the block matrix formulation of an equivalent inequality in QIT, conjectured recently by He and Vidal in [6], regarding the so-called 'monogamy of the negativity of entanglement'. We present this conjecture and prove its equivalence to (1) in Section 2. In our opinion, proving this inequality is a very hard problem, and we have only succeeded in proving a very special case. Namely, in Section 3 we prove the special case $n = 2$, where there are only two matrices A_1 and A_2, and where in addition we also require A_1 to be the identity matrix.

Let us recall the notations we will use. The modulus of a matrix X will be denoted as $|X|$, and is given by $(X^*X)^{1/2}$. Any Hermitian matrix can be decomposed as a difference of its positive and negative part: $X = X_+ - X_-$, with $X_\pm := (|X| \pm X)/2$. This is the so-called Jordan decomposition. The Schatten q-norm of a matrix, for $q \geq 1$ is denoted as $||X||_q$ and is defined as $||X||_q := (\operatorname{Tr}|X|^q)^{1/q}$. The trace norm is just the Schatten 1-norm, $||X||_1 = \operatorname{Tr} X$, and the Frobenius norm is the Schatten 2-norm. We will also need the quantity $||X||_q$ for $0 < q < 1$, which is no longer a norm but a quasi-norm. Finally, we denote the eigenvalues of a Hermitian matrix, sorted either in non-increasing or non-decreasing order as $\lambda_j^\downarrow$ and $\lambda_j^\uparrow$, respectively.

2 The He-Vidal Conjecture

Let us begin with introducing the He-Vidal Conjecture, along with the operations of partial transpose and partial trace, on which the conjecture is based. We will be very brief and refer to the quantum information literature (e.g. [1]) for a more in-depth discussion of the basic concepts.

For convenience we will use Dirac notation for vectors. A general vector of a Hilbert space $\mathcal{H}$ will be denoted by $|\psi\rangle$, where the symbol ψ is merely a label. The Hermitian conjugate of the vector $|\psi\rangle$ is denoted by $\langle\psi|$. The tensor product of two vectors $|\phi\rangle$ and $|\psi\rangle$ is denoted $|\phi\rangle \otimes |\psi\rangle$ or $|\phi\rangle|\psi\rangle$ for short. The elements of an orthonormal basis for a finite-dimensional Hilbert space $\mathcal{H}$ (with dimension d) will be denoted by $|i\rangle$, with $i = 1, \ldots, d$.

The He-Vidal conjecture is concerned with a Hilbert space $\mathcal{H}$ built up as a tensor product $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ of three Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_C$, of dimension d_A, d_B and d_C, respectively. Let us denote the orthonormal bases of these Hilbert spaces by $\{|i\rangle\}_{i=1}^{d_A}$ (for $\mathcal{H}_A$), $\{|j\rangle\}_{j=1}^{d_B}$ (for $\mathcal{H}_B$) and $\{|k\rangle\}_{k=1}^{d_C}$ (for $\mathcal{H}_C$). We choose as basis of $\mathcal{H}$ the tensor product of these three bases: $\{|i\rangle|j\rangle|k\rangle\}_{i,j,k}$, or in short $\{|ijk\rangle\}_{i,j,k}$.

Next, we need the *partial traces* with respect to $\mathcal{H}_B$ and $\mathcal{H}_C$, denoted as Tr_B and Tr_C, respectively, and the *partial transpose* with respect to $\mathcal{H}_A$, denoted by the superscript Γ. Let ρ_A, ρ_B and ρ_C be positive semidefinite matrices acting on $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_C$. Then the three mentioned operations are defined by their action on tensor products as

$$\begin{aligned}
\mathrm{Tr}_B(\rho_A \otimes \rho_B \otimes \rho_C) &= \mathrm{Tr}(\rho_B)\ \rho_A \otimes \rho_C \\
\mathrm{Tr}_C(\rho_A \otimes \rho_B \otimes \rho_C) &= \mathrm{Tr}(\rho_C)\ \rho_A \otimes \rho_B \\
(\rho_A \otimes \rho_B \otimes \rho_C)^\Gamma &= \rho_A^T \otimes \rho_B \otimes \rho_C,
\end{aligned}$$

with their action on general matrices (acting on the whole space $\mathcal{H}$) defined by linear extension.

Let us finally define $N(X) := ||X||_1 - \mathrm{Tr}\, X$, the negativity function of a Hermitian matrix X [8], and define $N_{A|BC}(\rho) = N(\rho^\Gamma)$, $N_{A|B}(\rho) = N(\mathrm{Tr}_C\, \rho^\Gamma)$ and $N_{A|C}(\rho) = N(\mathrm{Tr}_B\, \rho^\Gamma)$. Then the He-Vidal conjecture can be stated as follows:

Conjecture 1 (He-Vidal) *For any normalised complex vector $|\psi\rangle$ in the tensor product Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, the following inequality holds:*

$$N_{A|B}^2(|\psi\rangle\langle\psi|) + N_{A|C}^2(|\psi\rangle\langle\psi|) \leq N_{A|BC}^2(|\psi\rangle\langle\psi|). \tag{2}$$

In the following paragraphs, we rephrase this problem as an inequality for certain block matrices that can be readily understood without requiring any background knowledge in quantum information. In the remainder of the paper we then prove the conjecture in an important special case, using some well-established techniques of matrix analysis.

Given any set of orthonormal bases $\{|i\rangle\}$, $\{|j\rangle\}$ and $\{|k\rangle\}$ for the spaces $\mathcal{H}_A$,

$\mathcal{H}_B$ and $\mathcal{H}_C$, respectively, we can write the vector $|\psi\rangle$ as

$$|\psi\rangle = \sum_{i=1}^{d_A}\sum_{j=1}^{d_B}\sum_{k=1}^{d_C} c_{ijk}|ijk\rangle.$$

The coefficients c_{ijk} can be rearranged into d_A matrices A_i with elements $(A_i)_{jk} = c_{ijk}$. We write $|A_i\rangle$ for the reshaping of A_i as a vector: $|A_i\rangle = \sum_{jk} c_{ijk}|jk\rangle$. Then $|\psi\rangle$ can be written in terms of the $|A_i\rangle$ as

$$|\psi\rangle = \sum_{ijk} c_{ijk}|i\rangle|jk\rangle = \sum_i |i\rangle \otimes \sum_{jk} c_{ijk}|jk\rangle = \sum_i |i\rangle \otimes |A_i\rangle.$$

The normalisation of $|\psi\rangle$ yields a condition on the A_i:

$$1 = \langle\psi|\psi\rangle = \sum_i \langle A_i|A_i\rangle = \sum_i \mathrm{Tr}\, A_i^* A_i. \tag{3}$$

The negativities can now be rewritten in terms of these matrices A_i. For $N_{A|BC}$ we need ρ^Γ:

$$\rho^\Gamma = (|\psi\rangle\langle\psi|)^\Gamma = \sum_{i,i'} (|i\rangle\langle i'| \otimes |A_i\rangle\langle A_{i'}|)^\Gamma = \sum_{i,i'} |i'\rangle\langle i| \otimes |A_i\rangle\langle A_{i'}|.$$

Introducing the matrix $\mathcal{A} := \sum_i |A_i\rangle\langle i|$, which is a reshape of the vector of coefficients c_{ijk}, it is easy to check that $\left(\rho^\Gamma\right)^2 = \mathcal{A}^*\mathcal{A} \otimes \mathcal{A}\mathcal{A}^*$. Hence, $|\rho^\Gamma| = |\mathcal{A}| \otimes |\mathcal{A}^*|$, and for the negativity we get:

$$N_{A|BC} = ||\rho^\Gamma||_1 - 1 = ||\mathcal{A}||_1 \;\; ||\mathcal{A}^*||_1 - 1 = ||\mathcal{A}||_1^2 - 1 = ||\mathcal{A}^*\mathcal{A}||_{1/2} - 1.$$

To find the other two negativities we need the partial traces of ρ^Γ:

$$\mathrm{Tr}_B\, \rho^\Gamma = \sum_{i,i'} |i'\rangle\langle i| \otimes \mathrm{Tr}_B\, |A_i\rangle\langle A_{i'}| = \sum_{i,i'} |i'\rangle\langle i| \otimes \overline{A_i^* A_{i'}}$$
$$\mathrm{Tr}_C\, \rho^\Gamma = \sum_{i,i'} |i'\rangle\langle i| \otimes \mathrm{Tr}_C\, |A_i\rangle\langle A_{i'}| = \sum_{i,i'} |i'\rangle\langle i| \otimes A_i A_{i'}^*.$$

These partial traces can be expressed as $d_A \times d_A$ block matrices:

$$\mathrm{Tr}_C\, \rho^\Gamma = Z_1 := \begin{pmatrix} A_1A_1^* & A_2A_1^* & \dots \\ A_1A_2^* & A_2A_2^* & \dots \\ \vdots & \vdots & \end{pmatrix}, \quad \mathrm{Tr}_B\, \rho^\Gamma = Z_2 := \begin{pmatrix} A_1^*A_1 & A_2^*A_1 & \dots \\ A_1^*A_2 & A_2^*A_2 & \dots \\ \vdots & \vdots & \end{pmatrix}. \tag{4}$$

By the normalisation condition (3), we have $\mathrm{Tr}\, Z_1 = \mathrm{Tr}\, Z_2 = 1$. The corresponding negativities are

$$N_{A|B} = ||Z_1||_1 - 1 \text{ and } N_{A|C} = ||Z_2||_1 - 1.$$

We can now reformulate the conjecture in terms of block matrices: any d_A block matrices A_i satisfying the normalisation condition $\sum_i \mathrm{Tr}\, A_i^* A_i = 1$ also satisfy the inequality

$$(||Z_1||_1 - 1)^2 + (||Z_2||_1 - 1)^2 \leq \left(||\mathcal{A}^*\mathcal{A}||_{1/2} - 1\right)^2. \tag{5}$$

It is possible to simplify the right-hand side of (5) by choosing a particular orthonormal basis $\{|i\rangle\}$ for $\mathcal{H}_A$. Let the singular value decomposition of $\mathcal{A}^*$ be given as $\mathcal{A}^* = U\Sigma V^*$, where U and V are unitary matrices of dimension d_A and $d_B d_C$, respectively, and Σ is essentially diagonal. If we choose $|i\rangle$ to be the i-th column of U, for all i, then $\langle A_i|$ is the i-th row of ΣV^*. By this choice the vectors $|A_i\rangle$ become mutually orthogonal, and $\mathcal{A}^*\mathcal{A}$ becomes diagonal. The right-hand side of (5) then simplifies by the identity

$$||\mathcal{A}^*\mathcal{A}||_{1/2} = \left(\sum_i \sqrt{\langle A_i|A_i\rangle}\right)^2 = \left(\sum_i ||A_i||_2\right)^2.$$

Inequality (5) would therefore follow from the somewhat simpler inequality

$$(||Z_1||_1 - 1)^2 + (||Z_2||_1 - 1)^2 \leq \left(\left(\sum_i ||A_i||_2\right)^2 - 1\right)^2, \tag{6}$$

Remarkably, however, inequalities (6) and (5) are actually equivalent. This can be seen from the fact that replacing the positive semidefinite matrix $\mathcal{A}^*\mathcal{A}$ by its diagonal can not decrease the $||\cdot||_{1/2}$ quasinorm. By a further rescaling we can drop the normalisation condition $\sum_i \mathrm{Tr}\, A_i^* A_i = 1$, upon which (6) turns into the inequality (1) mentioned in the Introduction.

Note that, from a mathematical viewpoint, it would already be interesting to show that the following holds:

$$(||Z_1||_1 - \mathrm{Tr}\, Z_1)^2 + (||Z_2||_1 - \mathrm{Tr}\, Z_2)^2 \leq c \left(\sum_{i \neq j} ||A_i||_2 ||A_j||_2\right)^2,$$

with $c < 2$. However, to be of interest for the QIT community, it is essential that $c = 1$.

3 Proof of a special case

The task of proving inequality (1) is a hard one because of the inequality's tightness. It is easy to see that every term of the left-hand side of (1) is itself bounded above by the right-hand side. In entanglement theory, this corresponds to the fact that the negativity is a so-called entanglement monotone, which among other things means that it can not increase under taking partial traces [8]. A matrix analytical proof proceeds by first exploiting the triangle inequality to show that $||Z_1||_1 \leq \sum_{i,j} ||A_j A_i^*||_1$, and then the Cauchy-Schwartz inequality to bound $\sum_{i,j} ||A_j A_i^*||_1$ by $\sum_{i,j} ||A_j||_2 ||A_i^*||_2 = (\sum_i ||A_i||_2)^2$.

To prove (1), however, we must show that the sum of $(||Z_1||_1 - \operatorname{Tr} Z_1)^2$ and $(||Z_2||_1 - \operatorname{Tr} Z_2)^2$ is bounded above by the exact same expression that bounds each of the terms separately. Finding the proof of that statement is an extremely delicate process, where picking up proportionality constants has to be avoided at all costs. Any such constant larger than 1 (no matter how close to 1) would ruin the tightness and render the result irrelevant. For example, it is clear from the above that (1) certainly holds with an extra factor of 2 in the right-hand side (just add the inequalities for each term separately) but this is a trivial result and says absolutely nothing about monogamy of negativity.

In what follows we will restrict to the case $d_A = 2$ (i.e. system A is a qubit); even this simple case already turned out to be a major undertaking. To simplify notations, we will replace A_1 and A_2 by A and B. Then Z_1 and Z_2 are given by the 2×2 block matrices

$$Z_1 = \begin{pmatrix} AA^* & BA^* \\ AB^* & BB^* \end{pmatrix} \text{ and } Z_2 = \begin{pmatrix} A^*A & B^*A \\ A^*B & B^*B \end{pmatrix}.$$

Furthermore, we were obliged to restrict to the case $A = \mathbb{I}$. This requires taking $d_B = d_C$. We will henceforth write d for $d_B = d_C$.

In this case both terms of the left-hand side of (1) turn out to be less than one half the right-hand side (but note that numerical experiments reveal that this is not true in general). Adding up then proves (1). The goal is therefore to show, for all $d \times d$ matrices B,

$$||Z_1||_1 - \operatorname{Tr} Z_1 \leq \frac{1}{\sqrt{2}}\, 2\, ||\mathbb{I}||_2 ||B||_2 = 2\sqrt{d/2}\, ||B||_2.$$

Replacing B by B^* yields the corresponding inequality for Z_2. Henceforth, we will write Z for Z_1, and we have

$$Z = \begin{pmatrix} \mathbb{I} & B \\ B^* & BB^* \end{pmatrix}.$$

Noting that $||X||_1 - \operatorname{Tr} X = 2 \operatorname{Tr} X_-$ for any Hermitian X, we can rewrite the inequality as

$$\operatorname{Tr} Z_- \leq \sqrt{d/2}\, ||B||_2. \tag{7}$$

Our proof proceeds by splitting this inequality into two inequalities. First we show

$$\operatorname{Tr} Z_- \leq \operatorname{Tr} \sqrt{(BB^* - B^*B)_-} \tag{8}$$

and then we show

$$\operatorname{Tr} \sqrt{(BB^* - B^*B)_-} \leq \sqrt{d/2}\, ||B||_2. \tag{9}$$

3.1 Proof of inequality (8)

It is well-known that any given square matrix B is weakly unitarily equivalent to its Hermitian conjugate B^*. Indeed, let $B = U|B|$ be the polar decomposition of B, then $B^* = |B|U^*$, so that $B^* = U^*BU^*$. So, by multiplying B on the left and on the right by some unitary matrix, we obtain B^*. However, there is another way to relate B and B^* requiring only a left multiplication, by extending both matrices.

Let $\Delta = BB^* - B^*B$ and let its Jordan decomposition be $\Delta = \Delta_+ - \Delta_-$, where $\Delta_\pm \geq 0$. Then $BB^* + \Delta_- = B^*B + \Delta_+$. By positive semidefiniteness of all four terms we can write this as

$$\begin{pmatrix} B & \sqrt{\Delta_-} \end{pmatrix} \begin{pmatrix} B^* \\ \sqrt{\Delta_-} \end{pmatrix} = \begin{pmatrix} B^* & \sqrt{\Delta_+} \end{pmatrix} \begin{pmatrix} B \\ \sqrt{\Delta_+} \end{pmatrix}.$$

This immediately implies that there must exist a unitary matrix U such that

$$\begin{pmatrix} B^* \\ \sqrt{\Delta_-} \end{pmatrix} = U \begin{pmatrix} B \\ \sqrt{\Delta_+} \end{pmatrix}. \tag{10}$$

These two block matrices are the abovementioned extensions of B and B^*, respectively. If B is not square, it can be made so by zero-padding and the same statement therefore holds for general matrices B.

According to Cauchy's interlacing theorem, the eigenvalues of an $m \times m$ principal submatrix A' of an $n \times n$ Hermitian matrix A satisfy the relation $\lambda_j^{\uparrow}(A) \leq \lambda_j^{\uparrow}(A')$ for $j = 1, \ldots, m$ (there is also an upper bound, but we will not need it). In particular, as Z is a submatrix of the matrix

$$Z_1 := \begin{pmatrix} \mathbb{I} & 0 & B \\ 0 & \mathbb{I} & \sqrt{\Delta_+} \\ B^* & \sqrt{\Delta_+} & BB^* \end{pmatrix}$$

we have $\lambda_j^{\uparrow}(Z) \geq \lambda_j^{\uparrow}(Z_1)$ for $j = 1, \ldots, 2d$.

By (10), and the fact that for unitary U two block matrices of the form

$$\begin{pmatrix} \mathbb{I} & UX \\ X^*U^* & Y \end{pmatrix} \text{ and } \begin{pmatrix} \mathbb{I} & X \\ X^* & Y \end{pmatrix},$$

are equal up to a unitary conjugation and therefore have the same spectrum, Z_1 has the same spectrum as

$$Z_2 := \begin{pmatrix} \mathbb{I} & 0 & B^* \\ 0 & \mathbb{I} & \sqrt{\Delta_-} \\ B & \sqrt{\Delta_-} & BB^* \end{pmatrix}.$$

Now, Z_2 can be split as a sum of two matrices, the first one being positive semidefinite:

$$Z_2 = Z_3 + Z_4, \quad Z_3 := \begin{pmatrix} \mathbb{I} & 0 & B^* \\ 0 & \mathbb{I} & 0 \\ B & 0 & BB^* \end{pmatrix}, \quad Z_4 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \sqrt{\Delta_-} \\ 0 & \sqrt{\Delta_-} & 0 \end{pmatrix}.$$

By Weyl's monotonicity theorem, we therefore have $\lambda_j^{\uparrow}(Z_2) \geq \lambda_j^{\uparrow}(Z_4)$. The d smallest eigenvalues of Z_4 are non-positive and given by $-\sqrt{\mu_j}$, where μ_j are the eigenvalues of Δ_-. Thus, $\lambda_j^{\uparrow}(Z) \geq -\sqrt{\mu_j^{\downarrow}}$, for $j = 1, \ldots, d$. Furthermore,

Z_4 has at most d negative eigenvalues. Tracing back through the previous argument then reveals that this is also true for Z_2, Z_1 and finally Z itself.

So we have that the number of negative eigenvalues n_- of Z is at most d, and they are larger than $-\sqrt{\mu_j}$. Hence,

$$\operatorname{Tr} Z_- = \sum_{j=1}^{n_-}(-\lambda_j^\uparrow(Z)) \le \sum_{j=1}^{n_-}\sqrt{\mu_j^\downarrow} \le \sum_{j=1}^{d}\sqrt{\mu_j} = \operatorname{Tr}\sqrt{\Delta_-},$$

which is inequality (8).

3.2 Proof of inequality (9)

In this section we prove that the inequality (9) is valid for any $d \times d$ matrix B. For convenience we will actually prove the equivalent statement that

$$\operatorname{Tr}\sqrt{(BB^* - B^*B)_+} \le \sqrt{d/2}\,||B||_2;$$

the latter turns into the former by replacing B with B^*.

Note that BB^* and B^*B have the same eigenvalues, hence they are unitarily equivalent. Another way to phrase the inequality is that

$$\operatorname{Tr}\sqrt{(L - ULU^*)_+} \le \sqrt{d/2}\,||L||_2,$$

for any unitary matrix U and any non-negative diagonal matrix L. One way to attack this problem is to first try and prove it for U that are permutation matrices, so that both L and ULU^* are diagonal, and then extend this result from the commutative case to the general case. It turns out that this extension can indeed be done thanks to a theorem by Drury.

In [5], Drury stated the following theorem (without explicit proof, but with the remark that it can be proven easily using the method he has developed in a preceding publication, [4]):

Theorem 1 (Drury) *Let X and Y be $d \times d$ Hermitian matrices with given eigenvalues $x_1 \ge x_2 \ge \cdots \ge x_d$ and $y_1 \ge y_2 \ge \cdots \ge y_d$, respectively. Let $I = [x_n + y_n, x_1 + y_1]$. Let the function $\phi : I \to \mathbb{R}$ be isoclinally metaconvex on I. Then*

$$\operatorname{Tr}\phi(X+Y) \le \max_{\pi \in S_d}\sum_{j=1}^{d}\phi(x_j + y_{\pi(j)}).$$

The class of isoclinally metaconvex (IM) functions has been introduced by Drury in [5] exactly for this purpose:

Definition 1 *Let I be an interval in $\mathbb{R}$. An infinitely differentiable function $\phi : I \to \mathbb{R}$ is said to be IM on I if whenever $t_1, t_2 \in I$ with $t_1 \neq t_2$ and $\phi'(t_1) = \phi'(t_2)$, then $\phi''(t_1) + \phi''(t_2) > 0$.*

For example, strictly concave and strictly convex functions are both IM. It is possible for other functions to be in this class as well, provided that for every point where the curvature is negative there is another point with the same gradient and with positive curvature greater in absolute value.

This theorem would allow us to reduce the problem of proving inequality (9) to the commutative case if only the function $x \mapsto f(x) = \sqrt{x_+}$ were IM. Clearly it is not, as it is not even differentiable. However, $f(x)$ can be approximated arbitrarily well by a sequence of IM functions, as shown in Section 3.3, and this is all what is needed. Hence, Drury's result when applied to the matrices $X = BB^*$ and $Y = -B^*B$ implies that inequality (9) is valid if we can show that the inequality

$$\left(\sum_{i=1}^{d} \sqrt{(\mu_i - \mu_{\pi(i)})_+}\right)^2 \leq \frac{d}{2} \sum_{i=1}^{d} \mu_i$$

holds for any permutation $\pi \in S_d$, and for any set μ_i of non-negative numbers (the eigenvalues of BB^*). Without loss of generality we can assume that $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_d$ and $\sum_i \mu_i = 1$.

The key to the proof is to decompose a given permutation $\pi \in S_d$ in what we will call here *maximal ascending chains* (MA chains). Let an ascending chain be a sequence of increasing integers from $\{1, 2, \ldots, d\}$ such that the image under π of each integer in the chain is given by the next integer in the chain (if any). That is, it is a sequence $I := (i_1, i_2, \ldots, i_r)$ such that $i_{j+1} > i_j$ and $i_{j+1} = \pi(i_j)$, for $j = 1, 2, \ldots, r-1$. An MA chain is one that is as long as possible. For a general permutation, more than one such chain may exist. Clearly, chains are disjoint.

For example, the permutation $\begin{pmatrix} 1234 \\ 2341 \end{pmatrix}$ has one MA chain, namely $I = (1, 2, 3)$. The element 4 is not included because its image is 1, which is less than 4. The permutation $\begin{pmatrix} 1234 \\ 3421 \end{pmatrix}$ has two such chains, namely $I_1 = (1, 3)$ and $I_2 = (2, 4)$.

To proceed with the proof, we split the sum $\sum_i \sqrt{(\mu_i - \mu_{\pi(i)})_+}$ into several components, one per MA chain of the permutation π. Let the lengths of the various MA chains I_1, I_2, ..., I_K of a permutation be $r_1, r_2, \ldots, r_K$, respectively. Clearly, as MA chains are disjoint, the r_k sum up to at most d. Then we split the sum as follows:

$$\sum_{i=1}^{d} \sqrt{(\mu_i - \mu_{\pi(i)})_+} = \sum_{k=1}^{K} \sum_{i \in I_k} \sqrt{(\mu_i - \mu_{\pi(i)})_+}.$$

We can do this because the i-th term has a nonzero contribution to the sum unless i appears in some MA chain. Indeed, if i does not appear in any of the MA chains, this means that $\pi(i) < i$, whence, by the ordering of the μ_i, we have $\mu_{\pi(i)} > \mu_i$, so that $(\mu_i - \mu_{\pi(i)})_+ = 0$.

Let us now consider one such component, for a chain $I = (i_1, i_2, \ldots, i_r)$ of length r, namely $\sum_{j=1}^{r-1} \sqrt{(\mu_{i_j} - \mu_{\pi(i_j)})_+}$. Because the i_j form an MA chain, we have $\mu_{i_j} > \mu_{\pi(i_j)}$ and $\pi(i_j) = i_{j+1}$. We can therefore simplify this sum as $\sum_{j=1}^{r-1} \sqrt{\mu_{i_j} - \mu_{i_{j+1}}}$. We now claim that this sum is bounded above as

$$\sum_{j=1}^{r-1} \sqrt{\mu_{i_j} - \mu_{i_{j+1}}} \le \left(\frac{r}{2} \sum_{j=1}^{r} \mu_{i_j} \right)^{1/2}$$

For $r = 2$ this is trivially true, as the sum has only one term:

$$\left(\sum_{j=1}^{r-1} \sqrt{\mu_{i_j} - \mu_{i_{j+1}}} \right)^2 = \mu_{i_1} - \mu_{i_2} \le \mu_{i_1} + \mu_{i_2}.$$

For $r > 2$ we can exploit the following inequality, which can be seen as a Hölder-type inequality for the $l_{1/2}$-(quasi)-norm: for any vector x with non-negative real elements x_j, and any probability vector p (that is, $p_j \ge 0$ and $\sum_j p_j = 1$),

$$\left(\sum_{j=1}^{d} \sqrt{x_j} \right)^2 \le \sum_{j=1}^{d} \frac{x_j}{p_j}. \tag{11}$$

We will apply this in the following instance: $d = r - 1$, $x_j = \mu_{i_j} - \mu_{i_{j+1}}$ and $p_1 = 2/r$ and $p_2 = \cdots = p_{r-1} = 1/r$, to obtain, as required,

$$\left(\sum_{j=1}^{r-1} \sqrt{\mu_{i_j} - \mu_{i_{j+1}}} \right)^2 \le \frac{r}{2}(\mu_{i_1} - \mu_{i_2}) + r(\mu_{i_2} - \mu_{i_3}) + \cdots + r(\mu_{i_{r-1}} - \mu_{i_r})$$

$$= \frac{r}{2}(\mu_{i_1} + \mu_{i_2}) - r\mu_{i_r} \leq \frac{r}{2}\sum_{j=1}^{r} \mu_{i_j}.$$

Having one such bound per MA-chain component, we can now easily get a bound on the entire sum. By the previous result we have

$$\sum_{i=1}^{d} \sqrt{(\mu_i - \mu_{\pi(i)})_+} \leq \sum_{k=1}^{K} \sqrt{\frac{r_k}{2} \sum_{i \in I_k} \mu_i}.$$

We can now simply exploit the Cauchy-Schwarz inequality and find the upper bound

$$\sum_{k=1}^{K} \sqrt{\frac{r_k}{2} \sum_{i \in I_k} \mu_i} \leq \left(\sum_{k=1}^{K} \frac{r_k}{2}\right)^{1/2} \left(\sum_{k=1}^{K} \sum_{i \in I_k} \mu_i\right)^{1/2} \leq \sqrt{\frac{d}{2}} \sqrt{\sum_{i=1}^{d} \mu_i},$$

which ends the proof.

3.3 The function $\sqrt{x_+}$ can be approximated by IM functions

Here we prove the statement used in Section 3.2 that the function $f(x) = \sqrt{x_+}$ can be approximated arbitrarily well by IM functions. More precisely, we show that there exists a sequence of IM functions that converges uniformly to $f(x)$. Many functions do so, but we construct this sequence in such a way that its metaconvexity is easy to prove.

We start by defining a particular function $h(x)$ and then show two things: first, that $h(x)$ is IM and second, that $|h(x) - f(x)|$ is bounded by a finite constant $c > 0$. Using such a function $h(x)$ we can easily construct a sequence of IM functions converging uniformly to $f(x)$: we just have to consider the functions $h_s(x) := h(sx)/\sqrt{s}$. These functions inherit the property of being IM from $h(x)$, and $|h_s(x) - f(x)| = |h(sx) - f(sx)|/\sqrt{s} < c/\sqrt{s}$, which tends to 0 as s tends to $+\infty$, proving their uniform convergence.

To construct $h(x)$ consider the functions $w(x) = (1/2)(x^2+1)^{-1/4}$ and $\alpha(x) = 1 + \exp(-x)$, and let $g(x) = w(\alpha(x)x)$. The function $\alpha(x)$ satisfies $\alpha(x) \geq 1$, is monotonically decreasing, and tends to 1 as x tends to $+\infty$. Our function $h(x)$ of choice is the integral of $g(x)$, namely $h(x) = \int_{-\infty}^{x} dy\, g(y)$. Note that for x tending to $+\infty$, $w(x)$ tends to $1/(2\sqrt{x})$, so that in that regime $h(x)$ tends to $\sqrt{x}$ plus some finite constant arising from the integration over all smaller values of x.

We first show that $h(x)$ is IM. This involves the first and second derivatives of h, which are given by:

$$
\begin{aligned}
h'(x) &= g(x) = w(\alpha(x)x) \\
h''(x) &= g'(x) = w'(\alpha(x)x)\ (\alpha'(x)x + \alpha(x)).
\end{aligned}
$$

We therefore need to show that distinct x_1 and x_2 with the same value of $g(x)$ must satisfy $g'(x_1) + g'(x_2) > 0$. It is essential that $w(x)$ is an even function that is monotonically increasing for $x < 0$, and monotonically decreasing for $x > 0$, so that any pair of distinct x having the same $g(x)$ must have opposite sign. Let $x_1 < 0$ and $x_2 > 0$ be such points. By the evenness of $w(x)$, this is so if and only if $-\alpha(x_1)x_1 = \alpha(x_2)x_2$. For such points, the factor $w'(\alpha(x)x)$ in $g'(x)$ has the same absolute value (again by virtue of w being even), and is positive for x_1 and negative for x_2. The condition $g'(x_1) + g'(x_2) > 0$ is therefore equivalent to

$$
(\alpha'(x_1)x_1 + \alpha(x_1)) - (\alpha'(x_2)x_2 + \alpha(x_2)) > 0.
$$

This condition is easily seen to be satisfied as $\alpha'(x)x+\alpha(x) = -\theta x \exp(-\theta x) + 1 + \exp(-\theta x)$ is always larger than 2 for $x < 0$ and less than 2 for $x > 0$. This proves that $h(x)$ is IM.

Secondly, we have to show that $h(x)$ is an approximation of $f(x) = \sqrt{x_+}$, in the sense that $|h(x) - f(x)|$ is bounded by a finite constant. For $x < 0$ we have $f(x) = 0$ and $h(x) > 0$. To show that $h(x)$ is bounded above for $x < 0$ we only have to show that $h(0)$ is finite, since $h(x)$ is an increasing function (as $w(x) > 0$). Since $w(x) < 1/(2\sqrt{-x})$ for $x < 0$ and $\alpha(x) > \exp(-\theta x)$, we get, indeed,

$$
h(0) = \int_{-\infty}^{0} dy\ w(\alpha(y)y) < \int_{-\infty}^{0} dy\ \frac{1}{2\sqrt{\exp(-\theta y)y}} = \sqrt{\frac{\pi}{2\theta}}.
$$

For $x > 0$, $f(x) = \sqrt{x}$. As $\alpha(x) > 1$, we have that $h'(x) = w(\alpha(x)x) < w(x)$. For $x > 0$, we also have $w(x) < 1/(2\sqrt{x}) = f'(x)$, so that $h'(x) < f'(x)$. Integrating over x yields $h(x) - h(0) \leq f(x) - f(0)$ from which we obtain the upper bound $h(x) - f(x) < h(0) - f(0) = h(0)$, which is finite.

To obtain a lower bound we can exploit the two inequalities

$$
w(x) = \frac{1}{2(1+x^2)^{1/4}} > \frac{1}{2\sqrt{x}} - \frac{1}{8x^{5/2}} \text{ and } \frac{1}{\sqrt{1+\exp(-x)}} > 1 - \frac{1}{2}\exp(-x).
$$

This yields

$$h'(x) = w(\alpha(x)x) > \frac{1}{2\sqrt{x}\sqrt{1+\exp(-x)}} - \frac{1}{8(1+\exp(-x))^{5/2}}$$
$$> \frac{1}{2\sqrt{x}}\left(1 - \frac{1}{2}\exp(-x)\right) - \frac{1}{8x^{5/2}}$$

so that

$$h'(x) - f'(x) > -\frac{\exp(-x)}{4\sqrt{x}} - \frac{1}{8x^{5/2}}.$$

Integrating from 1 to x yields, for $x > 1$,

$$h(x) - f(x) > h(1) - \left(\frac{1}{4}\int_1^x dx\ \frac{\exp(-x)}{\sqrt{x}} + \frac{1}{8}\int_1^x dx\ \frac{1}{x^{5/2}}\right).$$

The first integral is bounded above by $\int_0^\infty dx\ \exp(-x)/\sqrt{x} = \sqrt{\pi}$ and the second integral is equal to $(2/3)(1-x^{-3/2})$, which is bounded above by 2/3. Thus, for $x > 1$, $h(x) - f(x)$ is bounded below by a finite constant. It is clear that, for $0 < x < 1$, $h(x) - f(x)$ is bounded below as well since $h(x) > 0$ and $f(x) < 1$. We conclude that $|h(x) - f(x)|$ is bounded everywhere by a finite constant.

Acknowledgments

This work is supported by an Odysseus grant from the Flemish FWO. Thanks to Frank Verstraete for discussions and rekindling my interest in entanglement theory. The helpful comments of an anonymous referee are also greatly appreciated.

References

[1] K.M.R. Audenaert, "Mathematical aspects of quantum information theory", in: Physics and Theoretical Computer Science, J.-P. Gazeau et al. (Eds.), 3–24, IOS Press (2007).

[2] K.M.R. Audenaert and F. Kittaneh, "Problems and Conjectures in Matrix and Operator Inequalities", in: Operator Theory, J. Zemanek ed., Banach Center Publications Series. (In Press). See also eprint arXiv:1201.5232.

[3] V. Coffman, J. Kundu and W. Wootters, "Distributed entanglement", Phys. Rev. A **61**, 052306 (2000).

[4] S.W. Drury, "On symmetric functions of the eigenvalues of the sum of two Hermitian matrices", Linear Algebra Appl. **176**, 211-222 (1992).

[5] S.W. Drury, "Maximizing traces of matrix functions", Linear Algebra Appl. **387**, 221–234 (2004).

[6] H. He and G. Vidal, "Disentangling theorem and monogamy for entanglement negativity", eprint arXiv:1401.5843 (2014).

[7] T.J. Osborne and F. Verstraete, "General monogamy inequality for bipartite qubit entanglement", Phys. Rev. Lett. **96**, 220503 (2006).

[8] G. Vidal and R.F. Werner, "A computable measure of entanglement", Phys. Rev. A **65**, 032314 (2002).