

Unital Quantum Channels – Convex Structure and Revivals of Birkhoff’s Theorem

Christian B. Mendl¹, Michael M. Wolf^{1,2}

¹ Max-Planck-Institute for Quantum Optics, Garching, Germany.

E-mail: christian_mendl@hotmail.com; wolf@nbi.dk

² Niels Bohr Institute, Copenhagen, Denmark

Received: 13 July 2008 / Accepted: 5 March 2009

Published online: 26 May 2009 – © Springer-Verlag 2009

Abstract: The set of doubly-stochastic quantum channels and its subset of mixtures of unitaries are investigated. We provide a detailed analysis of their structure together with computable criteria for the separation of the two sets. When applied to $O(d)$ -covariant channels this leads to a complete characterization and reveals a remarkable feature: instances of channels which are not in the convex hull of unitaries can become elements of this set by either taking two copies of them or supplementing with a completely depolarizing channel. These scenarios imply that a channel whose noise initially resists any environment-assisted attempt of correction can become perfectly correctable.

Contents

I.	Introduction	1058
II.	Unital Quantum Channels	1059
	A Preliminaries	1059
	B Representations	1060
	C Extreme points	1061
III.	Mixtures of Unitary Channels	1063
	A Separation witnesses	1063
	B A negativity measure	1064
IV.	Covariant Channels	1065
	A $O(d)$ covariance	1066
	B A complete picture	1067
V.	Restoring Birkhoff’s Theorem	1070
	A Two copies of a channel	1071
	B Help from a noisy friend	1072
	C Environment-assisted error correction	1074
VI.	Discussion	1074
VII.	Appendix A–Matrix Optimization Problems	1075

A Minimizing $\text{tr}[A\bar{A}]$ subject to fixed singular values 1075

B Maximizing $|\text{tr}[U]|$ subject to fixed $\text{tr}[U\bar{U}]$ 1076

C Minimizing $\text{tr} \begin{bmatrix} U_s & \bar{U}_s^{-T_2} \end{bmatrix}$ subject to fixed $\text{tr}[U\bar{U}]$ 1078

D Minimizing $\text{tr} \begin{bmatrix} U & \bar{U}^{-T_2} \end{bmatrix}$ 1081

VIII. Appendix B—A Special Extremal Channel 1084

References 1086

I. Introduction

Quantum channels are the most general input-output relations which the framework of quantum mechanics allows for arbitrary inputs. Physically, they describe any transmission in space, e.g., through optical fibres, and/or evolution in time, as in quantum memories, from a general open-systems point of view. Mathematically, they are characterized by linear, completely positive maps acting (in the Schrödinger picture) on density operators in a trace-preserving manner.

The present work investigates the particular class of quantum channels which leaves the maximally mixed (chaotic or infinite-temperature) state invariant. These channels are called *unital* or *doubly-stochastic* (referring to unital *and* trace-preserving) and they appear naturally in contexts with an irreducible symmetry. Apart from their practical relevance the interest in these channels has various origins: (i) they exhibit many special properties, e.g., regarding contractivity [1] or fixed points [2] — often allowing for a more geometric intuition, (ii) for small dimensions their additional constraint is strong enough to considerably simplify problems [3], and (iii) for sufficiently large dimensions problems on general channels can often be reduced to their unital counterparts [4–6].

The line of interest taken up by this article concerns the convex structure of the set of unital channels and, in particular, its relation to the subset of mixtures of unitary channels. This question was addressed and touched upon in [7–9] where a crucial difference between the classical and the quantum case was realized: whereas, by Garrett Birkhoff’s theorem [10], every doubly stochastic matrix (describing a classical channel) is a convex combination of reversible ones (i.e., permutations), not every doubly-stochastic quantum channel has to be a mixture of unitaries. The latter phenomenon became more significant when it was realized in Ref.[11] that a quantum channel allows for perfect environment-assisted error correction if and only if it is a mixture of unitaries. Another remarkable step was made in Ref.[12] where evidence has been provided that asymptotically many copies of a unital channel might always be well approximated by a mixture of unitaries—a conjectured restoration of Birkhoff’s theorem in the asymptotic limit.

We will (in Sec.V) provide a proof of this conjecture for special instances. This will result as a simple corollary from a thorough investigation of the convex structure of the sets of unital channels and of mixtures of unitary conjugations—in particular under symmetries. Along the way we obtain various other results on the mentioned convex structures which may be interesting in their own right. An outline of the paper and a summary of its results follows:

- In Sec.II we provide two characterizations of *unital channels*: (i) as channels which are convex combinations of unitaries acting on Hilbert-Schmidt space, and (ii) as channels which are affine combinations of unitary channels. Moreover, we show that extreme points of the set of unital channels need not be extremal within the set of all channels.

- In Sec.III computable criteria for the separation of unital channels from the set of mixtures of unitaries are provided and a respective negativity measure is introduced.
- In Sec.IV we focus on covariant channels (in particular w.r.t. $O(d)$) and show how symmetry enables us to explicitly determine the above sets and to compute the negativity measure.
- In Sec.V we apply the acquired tools in order show that families of covariant channels outside the convex hull of unitary channels can come into this set by either taking several copies of them or supplementing with a completely depolarizing channel.

II. Unital Quantum Channels

A. *Preliminaries.* We begin with introducing some notation and basic concepts. Throughout we will work in the Schrödinger picture and consider quantum channels T with finite and equal input and output dimensions, i.e., $T : \mathcal{M}_d \rightarrow \mathcal{M}_d$ is a linear map on $d \times d$ (density)matrices. Complete positivity enables a *Kraus decomposition*

$$T(\rho) = \sum_i A_i \rho A_i^\dagger, \quad \sum_i A_i^\dagger A_i = \mathbb{1}, \tag{1}$$

where the second relation expresses the trace preserving property. A channel is called *unital* if $T(\mathbb{1}) = \mathbb{1}$ and as we include the trace preserving property in the definition of a *channel*, a unital channel is a *doubly-stochastic* completely positive map.

It is often convenient to regard \mathcal{M}_d as a vector space which, when equipped with the inner product $\langle A, B \rangle := \text{tr}[A^\dagger B]$, forms the *Hilbert-Schmidt* Hilbert space \mathfrak{H}_d . Every channel is thus a linear map on this space and has as such a respective matrix representation $\hat{T} \in \mathcal{M}_{d^2} \simeq \mathcal{B}(\mathfrak{H})$.¹ We will occasionally use a (non-orthogonal) basis for \mathfrak{H} which is obtained from embedded Pauli-matrices in the form

$$\begin{aligned} \sigma_x^{jk} &:= |j\rangle \langle k| + |k\rangle \langle j| \quad \text{for all } j < k, \\ \sigma_y^{jk} &:= -i(|j\rangle \langle k| - |k\rangle \langle j|) \quad \text{for all } j < k, \\ \sigma_z^j &:= |j\rangle \langle j| - |j+1\rangle \langle j+1| \quad \forall j = 1, \dots, d-1, \end{aligned} \tag{2}$$

together with the identity matrix.

Another useful concept is the *state-channel duality* introduced by Jamiolkowski [13] which assigns a density operator $\rho_T \in \mathcal{M}_{d^2}$ to every channel T via

$$\rho_T = (\text{id} \otimes T)(|\Omega\rangle \langle \Omega|), \quad |\Omega\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^d |j, j\rangle,$$

where Ω is a *maximally entangled state*. The states ρ_T corresponding to unital channels are exactly those with reduced density matrices

$$\text{tr}_1 [\rho_T] = \text{tr}_2 [\rho_T] = \mathbb{1}/d. \tag{3}$$

Note that due to the linearity of the correspondence the convex structure of channels is entirely reflected by the convex structure of the set of their dual states. Depending on what is more convenient we will switch back and forth between T and ρ_T .

¹ $\mathcal{B}(X)$ denotes the space of *bounded linear operators* on X .

B. Representations. In the remainder of this subsection we will prove the following characterization of unital channels:

Theorem 1 (Characterization of unital channels). *Let $T : \mathcal{M}_d \rightarrow \mathcal{M}_d$ be a quantum channel. Then the following are equivalent:*

1. T is unital (i.e., doubly-stochastic),
2. T is a convex combination of unitaries on \mathfrak{H}_d , i.e., $\hat{T} = \sum_{\alpha} p_{\alpha} W_{\alpha}$ with the p 's being probabilities and each $W_{\alpha} \in \mathcal{M}_{d^2}$ unitary,
3. $T(\cdot) = \sum_i \lambda_i U_i \cdot U_i^{\dagger}$ is an affine combination of unitary channels, i.e., the λ 's are real and sum up to one and each $U_i \in \mathcal{M}_d$ is unitary.

In order to see $1 \Leftrightarrow 2$ we use a result from [1]: for any $p > 1$ a positive trace-preserving map T is a contraction in the sense² of $\|T\|_{p \rightarrow p} \leq 1$ iff³ T is unital. In addition we have $\|T\|_{2 \rightarrow 2} = \|\hat{T}\|_{\infty}$ so that T is unital iff \hat{T} is a contraction with respect to the operator norm. The set of these contractions in turn is the convex hull of unitaries (as can be seen from the singular value decomposition) which completes $1 \Leftrightarrow 2$.

As $3 \Rightarrow 1$ is obvious it remains to show $1 \Rightarrow 3$. To this end we introduce

$$\mathcal{X} := \{A \in \mathcal{M}_d : A = A^{\dagger}, \text{tr } A = 0\},$$

$$\mathcal{V} := \{A \mapsto UAU^{\dagger} : U \in \mathcal{M}_d \text{ unitary}\} \subset \mathcal{B}(\mathcal{X}).$$

That is, \mathcal{X} is a real linear subspace of \mathfrak{H} containing all Hermitian operators orthogonal to $\mathbb{1}$ and \mathcal{V} are the unitary conjugations on \mathcal{X} . Note that the real linear span of \mathcal{V} is invariant under composition and that the set in Eq. (2) (without the identity) forms a basis of \mathcal{X} . The idea is now to show first how $\mathcal{B}(\mathcal{X})$ can be obtained from \mathcal{V} and then to extend this to the claimed implication $1 \Rightarrow 3$ in Thm. 1.

Denote the subspace of real linear combinations of vectors $\{x_1, \dots, x_n\}$ such that the coefficients sum to zero by

$$\text{zerospan}_{\mathbb{R}} \{x_1, \dots, x_n\} := \left\{ \sum_i \lambda_i x_i : \lambda_i \in \mathbb{R}, \sum_i \lambda_i = 0 \right\}.$$

Lemma 2. *For each basis vector $B \in \mathcal{X}$ in (2) there exists a $T \in \text{zerospan}_{\mathbb{R}} \mathcal{V}$ which maps B to itself and all other basis vectors to zero.*

Proof. We explicitly construct such a T w.l.o.g. for σ_x ¹². Set

$$T_1(\rho) := \frac{1}{2} \left(\rho + U_1 \rho U_1^{\dagger} \right), \quad U_1 := \begin{pmatrix} \boxed{\mathbb{1}_2} & \\ & \boxed{-\mathbb{1}} \end{pmatrix},$$

$$T_2(\rho) := \frac{1}{2} \left(\rho - U_2 \rho U_2^{\dagger} \right) \in \text{zerospan}_{\mathbb{R}} \mathcal{V}, \quad U_2 := \begin{pmatrix} \boxed{\sigma_y} & \\ & \boxed{\mathbb{1}} \end{pmatrix}.$$

Then for all $\alpha \in \mathbb{R}^4$ and $\sigma \equiv (\sigma_x, \sigma_y, \sigma_z, \mathbb{1}_2)$,

$$A := \left(\begin{array}{c|c} \alpha \cdot \sigma & B^* \\ \hline & C \end{array} \right) \xrightarrow{T_1} \left(\begin{array}{c|c} \alpha \cdot \sigma & 0 \\ \hline 0 & C \end{array} \right) \xrightarrow{T_2} \left(\begin{array}{c|c} \alpha_1 \sigma_x + \alpha_3 \sigma_z & 0 \\ \hline 0 & 0 \end{array} \right).$$

² The norms are defined as $\|T\|_{p \rightarrow q} = \sup_A \|T(A)\|_q / \|A\|_p$ with $\|A\|_p = (\text{tr}[(A^{\dagger}A)^{p/2}])^{1/p}$.

³ As usual ‘iff’ should be read ‘if and only if’.

In a similar vein we can finally map α_3 to zero by a T_3 , defined as T_2 only with σ_y in U_2 replaced by σ_z . Then $T := T_3 \circ T_2 \circ T_1$ is the desired operator which satisfies $T(A) = \alpha_1 \sigma_x^{12}$, and $T \in \text{zerospan}_{\mathbb{R}} \mathcal{V}$ as $T_2 \in \text{zerospan}_{\mathbb{R}} \mathcal{V}$. Clearly, the same type of construction works for all basis vectors in (2). \square

Lemma 3. *For every pair of basis vectors $B_1, B_2 \in \mathcal{X}$ in (2), there is a $T \in \mathcal{V}$ such that $T(B_1) = B_2$.*

Proof. As B_1 and B_2 are Hermitian, there are unitaries U_1 and U_2 such that $U_j^\dagger B_j U_j$ ($j = 1, 2$) are both diagonal. These can in turn be mapped onto each other by a permutation in \mathcal{V} since they both have eigenvalues $(1, -1, 0, \dots, 0)$. Exploiting that \mathcal{V} forms a group we can compose these steps to obtain $T(B_1) = B_2$. \square

Proposition 4. \mathcal{V} zero-spans all linear operators on \mathcal{X} , that is,

$$\text{zerospan}_{\mathbb{R}} \mathcal{V} = \mathcal{B}(\mathcal{X}).$$

Proof. For any two basis vectors B_1, B_2 in (2), by the above lemmas there is a $T \in \text{zerospan}_{\mathbb{R}} \mathcal{V}$ which maps B_1 to B_2 and all other basis vectors to zero, so that a linear combination of these T ’s generates any linear map on \mathcal{X} . \square

This immediately implies $1 \Rightarrow 3$ in Thm.1 as for every unital quantum channel T we have that $T - \text{id} \in \mathcal{B}(\mathcal{X})$ so that we can write $T(\rho) = \rho + \sum_i \lambda_i U_i \rho U_i^\dagger$ with the λ ’s summing up to zero.

Note that Thm.1 implies that mixtures of unitaries form a set of non-zero measure within the set of doubly stochastic channels. Conversely, assuming that they have non-zero measure implies the equivalence of 1 and 3.⁴

C. Extreme points. The set of all unital quantum channels on \mathcal{M}_d is convex and compact. That is, every unital channel T can be decomposed as

$$T = \sum_i p_i T_i, \tag{4}$$

where the p ’s are probabilities and the T_i ’s are *extremal* unital channels, i.e., those which cannot be further decomposed in a non-trivial way. Despite considerable effort [7, 9, 15–17] not much is known about the explicit structure of these extreme points beyond $d = 2$ (in which case they are all unitary conjugations [7]). The small contribution of this subsection is to review the existing results and to apply them in order to show that channels which are extremal within the set of unital channels are not necessarily extremal within the convex set of all channels. To the best of our knowledge all known examples so far were extremal within both sets—although the numerical results stated in [17] already indicate that this might not be generally true. The main ingredient is the following theorem which is stated in [9] and based on [18].

⁴ The fact that they have non-zero measure was proven independently (after the present paper has been made public) in [14].

Theorem 5. (Extremal channels). *Consider a quantum channel with Kraus operators $\{A_i\}_{i=1,\dots,N}$. It is an extreme point within the convex set of quantum channels iff the set of matrices*

$$\left\{ A_k^\dagger A_l \right\}_{k,l=1\dots N} \tag{5}$$

is linearly independent. Assume further that the channel is unital. Then it is extremal within the convex set of unital channels iff

$$\left\{ A_k^\dagger A_l \oplus A_l A_k^\dagger \right\}_{k,l=1\dots N} \tag{6}$$

is linearly independent.

We will exploit the fact that (5) allows less linearly independent operators than (6): while (5) gives the simple bound $N \leq d$, the set (6) yields $N \leq \sqrt{2}d$.⁵ For our example we choose dimension $d = 3$ and $N = 4$ linearly independent Kraus operators. The former ensures that there are non-trivial extreme points, and the latter already implies that (5) can never be linearly independent as $N \not\leq d$. We start with an Ansatz for the Jamiolkowski state of the sought channel of the form

$$\rho_T = (\text{id} \otimes T)(|\Omega\rangle\langle\Omega|) = \sum_{i,j=1}^6 x_{ij} |\psi_i\rangle\langle\psi_j|,$$

where the $(|\psi_i\rangle)_i$ span the orthogonal complement of $(|kk\rangle)_k$, namely

$$\begin{aligned} |\psi_1\rangle &= \frac{1}{\sqrt{2}} (|12\rangle + |21\rangle), & |\psi_2\rangle &= \frac{1}{\sqrt{2}} (|13\rangle + |31\rangle), & |\psi_3\rangle &= \frac{1}{\sqrt{2}} (|23\rangle + |32\rangle), \\ |\psi_4\rangle &= \frac{1}{\sqrt{2}i} (|12\rangle - |21\rangle), & |\psi_5\rangle &= \frac{1}{\sqrt{2}i} (|13\rangle - |31\rangle), & |\psi_6\rangle &= \frac{1}{\sqrt{2}i} (|23\rangle - |32\rangle), \end{aligned}$$

and the Hermitian matrix $X \equiv (x_{ij})$ is given by

$$X := \frac{1}{3} \begin{pmatrix} \frac{1}{2} & 0 & -i\mu_1 & i\mu_3 & i\mu_4 & 0 \\ 0 & \frac{1}{2} & -i\mu_1 & -i\mu_4 & -(2+i)\mu_3 & 0 \\ i\mu_1 & i\mu_1 & \frac{1}{2} & 0 & 0 & 2\mu_2 + i\mu_3 \\ -i\mu_3 & i\mu_4 & 0 & \frac{1}{2} & 0 & -i\mu_1 \\ -i\mu_4 & (i-2)\mu_3 & 0 & 0 & \frac{1}{2} & i\mu_1 \\ 0 & 0 & 2\mu_2 - i\mu_3 & i\mu_1 & -i\mu_1 & \frac{1}{2} \end{pmatrix}. \tag{7}$$

The latter is chosen such that ρ_T satisfies the conditions (3) corresponding to a unital and trace-preserving map. It remains to choose algebraic numbers $\mu_1, \dots, \mu_4 \in \mathbb{R}$ such that X is positive semidefinite with rank $N = 4$, and that at the same time (6) is linearly independent when plugging in the corresponding Kraus operators. A possible choice for such a set of parameters is provided in Appendix VIII.

⁵ In fact, in [16] it was shown that $N \leq \sqrt{2d^2 - 1}$ which is, however, practically the same as $N \leq \sqrt{2}d$ when applied to integer N .

III. Mixtures of Unitary Channels

This section deals with the class of unital channels which can be represented as

$$T(\rho) = \sum_{i=1}^N p_i U_i \rho U_i^\dagger, \quad U_i U_i^\dagger = \mathbb{1}, \quad p_i > 0 \quad \forall i. \tag{8}$$

The Jamiolkowski states corresponding to these mixtures of unitary conjugations are exactly the states which are convex combinations of maximally entangled states. The rank of the Jamiolkowski state ρ_T gives a simple bound [19] for the minimal N as there exists always a decomposition with $N \leq (\text{rank } \rho_T)^2$. For $d = 2$ we can achieve equality in the general lower bound $N \geq \text{rank } \rho_T$ and, as mentioned before, every unital channel on \mathcal{M}_2 is a mixture of unitaries [7]. For $d \geq 3$ the question whether a given unital channel allows for such a representation was investigated and reformulated in [17] but a general operational way of deciding it remains to be found. The approach in the following subsection provides a class of easily computable necessary conditions which, when applied to covariant channels, will later be extended to necessary and sufficient criteria.

A. Separation witnesses. Since the set (8) of mixtures of unitary channels is convex and compact, every unital channel which lies outside this set can be separated from it by a hyperplane — a *witness*. As this can most easily be expressed on the level of Jamiolkowski states we introduce the corresponding sets

$$\begin{aligned} \mathcal{S} &:= \{ \rho_T : \rho_T = (\text{id} \otimes T)(|\Omega\rangle\langle\Omega|), T : \mathcal{M}_d \rightarrow \mathcal{M}_d \text{ cp, tp, unital} \} \\ &= \{ \rho \in \mathcal{M}_{d^2} : \rho \geq 0, \text{tr}_1 \rho = \text{tr}_2 \rho = \mathbb{1}/d \}, \\ \mathcal{U} &:= \text{conv} \left\{ (\mathbb{1} \otimes U) |\Omega\rangle\langle\Omega| (\mathbb{1} \otimes U^\dagger) : U U^\dagger = \mathbb{1} \right\}, \end{aligned}$$

which we will, with some abuse of notation, occasionally also use for channels, i.e., we will write ‘ $T \in \mathcal{S}$ ’ meaning $\rho_T \in \mathcal{S}$. The following shows that we may impose some structure on the witnesses — they can be taken from the affine span of \mathcal{U} .

Proposition 6. (Separation witnesses). *Let $\rho \in \mathcal{S}$ characterize a unital quantum channel. Then $\rho \in \mathcal{U}$, i.e., it is a mixture of maximally entangled states, iff*

$$\text{tr}[W\rho] \geq 0$$

for all Hermitian operators $W \in \mathcal{M}_{d^2}$ which satisfy

$$\text{tr}_1 W = \text{tr}_2 W = \mathbb{1}/d, \quad \text{tr}[W\sigma] \geq 0 \quad \forall \sigma \in \mathcal{U}. \tag{9}$$

Proof. We have to show that if $\rho \notin \mathcal{U}$, then there exists such a W with $\text{tr}[W\rho] < 0$. First note that

$$\mathcal{X} := \left\{ A \in \mathcal{M}_{d^2} : A = A^\dagger, \text{tr}_1 A = \text{tr}_2 A = 0 \right\}$$

is a real linear space and $\mathcal{S} - \mathbb{1}/d^2 \subset \mathcal{X}$. Set $\tilde{\rho} := \rho - \mathbb{1}/d^2 \in \mathcal{X}$. Using the Hahn-Banach separation theorem [20, Theorem 1.C in Chap. 1] we find a $\tilde{W} \in \mathcal{X}$ with

$$\text{tr}[\tilde{W}\tilde{\rho}] < -1/d^2, \quad \text{but} \quad \text{tr}[\tilde{W}\tilde{\sigma}] \geq -1/d^2 \quad \forall \tilde{\sigma} \in (\mathcal{U} - \mathbb{1}/d^2).$$

Setting $W := \tilde{W} + \mathbb{1}/d^2$ yields the sought witness. \square

To simplify matters we will in the following also consider Hermitian witnesses which do not fulfill the l.h.s. of (9) as long as the r.h.s. is satisfied. A class of this kind which turns out to be particularly useful are operators constructed from the *flip* operator $\mathbb{F} : |k, l\rangle \mapsto |l, k\rangle$ in the form

$$W = (\mathbb{1} \otimes B) \mathbb{F} (\mathbb{1} \otimes B^\dagger) + w(B) \mathbb{1}, \quad B \in \mathcal{M}_d, \tag{10}$$

where $w(B) \in \mathbb{R}$ is a constant depending on B such that W fulfills the r.h.s. in (9). Before we determine this dependence let us note that replacing $(\mathbb{1} \otimes B)$ by $(A \otimes B)$ in Eq. (10) won't lead to a more general class of witnesses since $(A \otimes B) \mathbb{F} (A^\dagger \otimes B^\dagger) = (\mathbb{1} \otimes BA^\dagger) \mathbb{F} (\mathbb{1} \otimes AB^\dagger)$.

The sharpest constant $w(B)$ for which (10) fulfills the witness condition $\text{tr}[W\rho] \geq 0$ for all $\rho \in \mathcal{U}$ is obtained from

$$\begin{aligned} w(B) &= -\min_U \text{tr}[(\mathbb{1} \otimes B) \mathbb{F} (\mathbb{1} \otimes B^\dagger) (\mathbb{1} \otimes U) |\Omega\rangle \langle \Omega| (\mathbb{1} \otimes U^\dagger)] \\ &= -\frac{1}{d} \min_U \text{tr} [B^\dagger U B^T \bar{U}] \\ &= -\frac{1}{d} \min_A \{ \text{tr} [A \bar{A}] : \sigma(A) = \sigma(B) \}, \end{aligned}$$

where U is unitary, $A \in \mathcal{M}_d$ and $\sigma(A)$ denotes the singular values of A . We solve this matrix optimization problem in Appendix VII A arriving at the following result.

Theorem 7. (Tight witnesses). *For any $B \in \mathcal{M}_d$ with singular values $\sigma_1 \geq \dots \geq \sigma_d$ the operator in Eq. (10) is a separation witness iff*

$$w(B) \geq \frac{1}{d} \begin{cases} 2 \sum_{i=1}^{d/2} \sigma_{2i-1} \sigma_{2i}, & d \text{ even} \\ 2 \left(\sum_{i=1}^{(d-1)/2} \sigma_{2i-1} \sigma_{2i} \right) - \sigma_d^2, & d \text{ odd.} \end{cases} \tag{11}$$

Note in particular that for $B = \mathbb{1}$ and d odd we get $w \geq 1 - 2/d$ while for d even $w \geq 1$. Hence, for even d no channel is separated from \mathcal{U} by such a witness (since $\mathbb{F} + \mathbb{1} \geq 0$). However, we will see in Sec. IV B that for d odd it becomes a powerful tool.

B. A negativity measure. There are several possible ways of quantifying the deviation of a channel $T \in \mathcal{S} \setminus \mathcal{U}$ from being a mixture of unitary channels: one may for instance follow [17], use the *entanglement of assistance* [12,21] or the minimal distance to the set \mathcal{U} w.r.t. some distance measure. The representation Thm.1 enables a very natural alternative approach—a *base norm* (inspired by [22]). That is, the deviation is quantified by the smallest negative contribution when representing T as an affine combination of terms in \mathcal{U} . More formally:

Definition 8. (Negativity). *For all $\rho \in \mathcal{S}$ the base norm associated with \mathcal{U} is*

$$\|\rho\|_{\mathcal{U}} := \inf \{ \alpha_p + \alpha_n : \rho = \alpha_p \sigma_p - \alpha_n \sigma_n, \alpha_{p,n} \geq 0, \sigma_{p,n} \in \mathcal{U} \},$$

and the corresponding negativity is given by

$$\mathcal{N}_{\mathcal{U}}(\rho) := \inf \{ \alpha_n : \rho = \alpha_p \sigma_p - \alpha_n \sigma_n, \alpha_{p,n} \geq 0, \sigma_{p,n} \in \mathcal{U} \}. \tag{12}$$

For $\text{tr}[\rho] = 1$ the two are related via $\|\rho\|_{\mathcal{U}} = 1 + 2\mathcal{N}_{\mathcal{U}}(\rho)$ and obviously $\mathcal{N}_{\mathcal{U}}(\rho) = 0$ iff $\rho \in \mathcal{U}$. The base norm behaves nicely under concatenation and convex combination. Writing $\|T\|_{\mathcal{U}} := \|\rho_T\|_{\mathcal{U}}$ we get

Proposition 9. *Let $T_i \in \mathcal{S}$ be a set of quantum channels and $p_i \geq 0$ probabilities, then*

$$\left\| \prod T_i \right\|_{\mathcal{U}} \leq \prod \|T_i\|_{\mathcal{U}}, \quad \text{and} \quad \left\| \sum p_i T_i \right\|_{\mathcal{U}} \leq \sum p_i \|T_i\|_{\mathcal{U}}.$$

Both can easily be proven from the definition. The latter can be interpreted as coming from triangle inequality and homogeneity of the norm. Note also that the above norm is unitarily invariant in the sense of $\|TV\|_{\mathcal{U}} = \|VT\|_{\mathcal{U}} = \|T\|_{\mathcal{U}}$ for every unitary conjugation V .

As always measures are easy to define but hard to compute. For covariant channels we will show the calculation in Sec. IV B.

IV. Covariant Channels

In order to arrive at more explicit results we need some help—coming in the form of symmetries imposed on the channels. Consider any subgroup $\mathcal{G} \subset U(d)$ with elements $g \in \mathcal{G}$ and two unitary representations V_g, \tilde{V}_g on \mathbb{C}^d . We say that a channel $T : \mathcal{M}_d \rightarrow \mathcal{M}_d$ is \mathcal{G} -covariant w.r.t. these representations if for all $g \in \mathcal{G}$:

$$T \left(V_g \cdot V_g^\dagger \right) = \tilde{V}_g T(\cdot) \tilde{V}_g^\dagger. \tag{13}$$

In this sense the action of the channel ‘commutes with the symmetry’. If \tilde{V} is an irreducible representation then T is unital as $T(\mathbb{1}) = \int dg T(V_g V_g^\dagger) = \int dg \tilde{V}_g T(\mathbb{1}) \tilde{V}_g^\dagger = \mathbb{1}$ by invoking Schur’s Lemma (where dg is the Haar measure). In order to express Eq. (13) in terms of the Jamiołkowski state ρ_T we introduce $G = \{ \tilde{V}_g \otimes V_g \}_{g \in \mathcal{G}}$ and its commutant $G' = \{ X \in \mathcal{M}_{d^2} : \forall U_g \in G : [X, U_g] = 0 \}$. Covariance of the channel translates then simply to

$$\rho_T \in G'.$$

As we will see below most of the analysis can w.l.o.g. be restricted to this commutant which considerably simplifies matters as $\dim G'$ is for a sufficiently large symmetry group much smaller than d^4 , the dimensionality we would have to deal with otherwise. The map

$$P(A) := \int dg U_g A U_g^\dagger$$

defines a projection in $\mathcal{B}(\mathcal{M}_{d^2})$, often called *twirl*, which maps every matrix A into G' and acts as the identity on G' . Moreover, since G' is an algebra it is spanned by a set of minimal projections $\{P_i\}$. These are orthogonal if G' is abelian (which happens for large enough symmetry groups) so that every $X \in G'$ can be written as $X = \sum_i x_i P_i$ with $x_i = \text{tr}[X P_i] / \text{tr}[P_i]$. In this case we can easily determine⁶

$$P(A) = \sum_i \frac{\text{tr}[A P_i]}{\text{tr}[P_i]} P_i. \tag{14}$$

⁶ Here we have used that $\text{tr}[P(A) P_i] = \text{tr}[A P(P_i)] = \text{tr}[A P_i]$.

If G' fails to be abelian a similar reasoning still applies—for a detailed exposition of these matters we refer to [22]. In order to see how covariance helps for our purposes let us denote the set of witnesses by $\mathcal{W} := \{W = W^\dagger : \forall \sigma \in \mathcal{U} : \text{tr}[W\sigma] \geq 0\}$.

Proposition 10 (Reduction to the commutant). *Let $\rho \in \mathcal{S} \cap G'$ be the Jamiołkowski state corresponding to a covariant unital channel. Then $\rho \in \mathcal{U}$ iff $\text{tr}[W\rho] \geq 0$ for all $W \in \mathcal{W} \cap G'$. Moreover,*

$$\|\rho\|_{\mathcal{U}} = \inf \{ \alpha_p + \alpha_n : \rho = \alpha_p \sigma_p - \alpha_n \sigma_n, \alpha_{p,n} \geq 0, \sigma_{p,n} \in \mathcal{U} \cap G' \},$$

which equivalently holds for the negativity $\mathcal{N}_{\mathcal{U}}$.

Proof. The crucial point for both parts is that $\sigma \in \mathcal{U}$ implies $\mathbf{P}(\sigma) \in \mathcal{U}$ which in turn means that $\mathbf{P}(W) \in \mathcal{W}$ for every $W \in \mathcal{W}$. Therefore due to $\text{tr}[\rho \mathbf{P}(W)] = \text{tr}[\mathbf{P}(\rho)W] = \text{tr}[\rho W]$ the set \mathcal{W} can w.l.o.g. be restricted to G' . Regarding the base norm we arrive at the stated result when starting with any optimal decomposition $\rho = \alpha_p \sigma_p - \alpha_n \sigma_n$ and applying the twirl to both sides of the equation. \square

This suggests the program for the next subsections: fix a symmetry group, identify the commutant G' and determine \mathcal{U} , $\|\cdot\|_{\mathcal{U}}$ and $\mathcal{N}_{\mathcal{U}}$ by exploiting the reduction to G' .

A. $O(d)$ covariance. The symmetry we will consider is the one of the real orthogonal group, i.e., $G = \{O \otimes O : O \in \mathcal{M}_d \text{ real orthogonal}\}$. The most prominent non-trivial example of a channel having this symmetry is

$$T(\rho) = (\text{tr}[\rho] \mathbb{1} - \rho^T)/(d - 1), \tag{15}$$

which (for $d = 3$) gained some popularity as a steady source of counterexamples: for the multiplicativity of the output p -norm [23], the additivity of the relative entropy of entanglement [24] and, most relevant in our context, the quantum analogue of Birkhoff’s theorem [9]. On the level of Jamiołkowski states we can make use of the analysis in [24] where the commutant G' was shown to be abelian and spanned by

$$G' = \text{span} \{ \mathbb{1}, \mathbb{F}, \widehat{\mathbb{F}} \},$$

where $\widehat{\mathbb{F}} := d |\Omega\rangle \langle \Omega|$. From there the minimal projections can be identified as

$$\begin{aligned} P_0 &= \frac{1}{d} \widehat{\mathbb{F}} = |\Omega\rangle \langle \Omega|, \\ P_1 &= \frac{1}{2} (\mathbb{1} - \mathbb{F}), \\ P_2 &= \frac{1}{2} (\mathbb{1} + \mathbb{F}) - \frac{1}{d} \widehat{\mathbb{F}}, \end{aligned}$$

where $(\mathbb{1} \pm \mathbb{F})/2$ are the projections onto the symmetric and anti-symmetric subspace, respectively. Consequently, every density operator in G' is in the convex hull of the corresponding normalized density matrices $\rho_i = P_i/\text{tr}[P_i]$ of which ρ_1 corresponds to the *Werner-Holevo channel* in (15), ρ_0 is the ideal channel and $\sum_i P_i/d^2$ corresponds to the completely depolarizing channel $T(\rho) = \text{tr}[\rho] \mathbb{1}/d$. Clearly, all of them are unital, i.e., elements of \mathcal{S} . Every state $\rho \in G'$ is completely characterized by its “coordinates”

$$((\mathbb{F}), (\widehat{\mathbb{F}}))_\rho \equiv (\text{tr}[\rho \mathbb{F}], \text{tr}[\rho \widehat{\mathbb{F}}]).$$

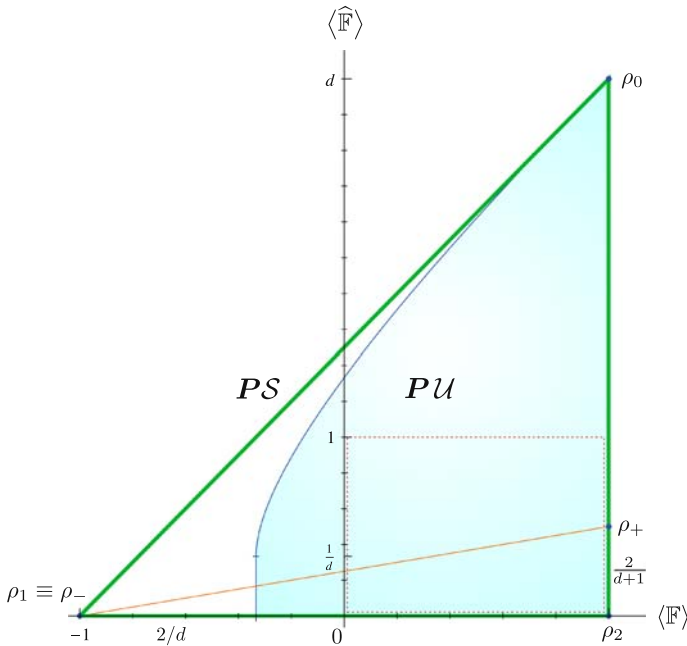


Fig. 1. The set of orthogonal covariant channels PS in the Jamiolkowski representation (outer/green triangle) and the convex hull of unitary channels PU (blue/shaded area), which is described in analytic terms by Proposition 12. Note that the Werner-Holevo channel ρ_- is “furthest away” from the unitaries. The orange line (from ρ_- to ρ_+) depicts the $U(d)$ covariant channels, and the dotted unit-square corresponds to entanglement-breaking channels. Compare with Ref. [24, Fig. 2]

Especially for the extreme points ρ_i we obtain (see Fig. 1)

state	ρ_0	ρ_1	ρ_2	(16)
coords	$(1, d)$	$(-1, 0)$	$(1, 0)$	

B. A complete picture. We will now determine the subset \mathcal{U} of mixtures of unitary channels within the set of $O(d)$ -covariant channels. Following the above considerations this amounts to identifying the corresponding region in the two-dimensional parameter space

$$\begin{aligned}
 \mathcal{U} \cap G' &\cong \left\{ (\langle \mathbb{F} \rangle, \langle \widehat{\mathbb{F}} \rangle)_\rho : \rho \in \mathcal{U} \right\} \\
 &= \text{conv} \left\{ (\langle \mathbb{F} \rangle, \langle \widehat{\mathbb{F}} \rangle)_U : U \text{ unitary} \right\},
 \end{aligned}
 \tag{17}$$

where the index U stands for the expectation value w.r.t. $(\mathbb{1} \otimes U) |\Omega\rangle$ which parameterizes an extreme point within \mathcal{U} . A short calculation reveals that

$$\begin{aligned}
 \langle \mathbb{F} \rangle_U &\equiv \text{tr} \left[(\mathbb{1} \otimes U) |\Omega\rangle \langle \Omega| (\mathbb{1} \otimes U^\dagger) \mathbb{F} \right] = \frac{1}{d} \text{tr} [U \overline{U}], \\
 \langle \widehat{\mathbb{F}} \rangle_U &\equiv \text{tr} \left[(\mathbb{1} \otimes U) |\Omega\rangle \langle \Omega| (\mathbb{1} \otimes U^\dagger) \widehat{\mathbb{F}} \right] = \frac{1}{d} |\text{tr } U|^2.
 \end{aligned}
 \tag{18}$$

The picture depends crucially on whether d is even or odd.

Theorem 11 (Even dimension). *If d is even then $\mathcal{U} \cap G' = \mathcal{S} \cap G'$, i.e., every $O(d)$ -covariant channel is a mixture of unitary channels.*

Proof. It suffices to note that the expectation values (16,18) with respect to ρ_i and U_i coincide for $U_0 = \mathbb{1}$, $U_1 = \text{diag}(\sigma_y, \dots, \sigma_y)$ and $U_2 = \text{diag}(\sigma_z, \dots, \sigma_z)$, just by plugging in (18). \square

So the interesting structure only emerges for d odd (see Fig. 1 for $d = 3$), for which we need the following result proven in Appendix VII B:

Proposition 12. *Let $d \geq 1$ be odd. Then for all $x \in [-1 + \frac{2}{d}, 1]$ there exists a unitary $U \in U(d)$ such that $\text{tr}[U\bar{U}]/d = x$, and*

$$\begin{aligned} & \max \{ |\text{tr} U|/d : U \in U(d), \text{tr}[U\bar{U}]/d = x \} \\ & = \left[\frac{1}{2} \left(1 - \frac{1}{d} \right) \left(1 - \frac{2}{d} + x \right) \right]^{1/2} + \frac{1}{d} =: m(x). \end{aligned} \tag{19}$$

Theorem 13 (Odd dimension). *Let $d \geq 3$ be odd. Then the extreme points of the set (17) corresponding to mixtures of unitary channels are*

$$(-1 + 2/d, 0), \quad (1, 0) \quad \text{and} \quad \left\{ (x, d(m(x))^2) : x \in [-1 + 2/d, 1] \right\}. \tag{20}$$

Proof. “(17) \subset conv(20)”: For all unitary $U \in U(d)$,

$$\frac{1}{d} \text{tr}[U\bar{U}] \in [-1 + 2/d, 1],$$

which follows from the fact that for any matrix A the spectrum of $A\bar{A}$ is symmetric with respect to the real axis, the eigenvalues $\lambda, \bar{\lambda}$ have the same algebraic multiplicity, and the algebraic multiplicity of all negative eigenvalues of $A\bar{A}$ (if any) is even, see [25]. Together with Prop. 12 we obtain the stated bounds on (17).

“(20) \subset (17)”: Set

$$Q_0 := \frac{1}{2} \begin{pmatrix} 0 & 1-i & -1-i \\ -1+i & -i & 1 \\ 1+i & 1 & i \end{pmatrix}$$

and $\varphi := \exp(2\pi i/3)$, then the coordinates (20) are obtained by

$U_0 = \text{diag}(\sigma_y, \dots, \sigma_y, Q_0)$, $U_1 = \text{diag}(\sigma_z, \dots, \sigma_z, \varphi, \varphi^2, 1)$ and the unitary matrices which solve the maximization problem (19) (explicitly given in Appendix VII B). \square

The fact that according to Prop. 10, we can restrict to decompositions within the two-dimensional parameter space, together with the explicit characterization of the set $\mathcal{U} \cap G'$ enables us now to compute the negativity $\mathcal{N}_{\mathcal{U}}$, as follows.

We show first that in Eq. (12), $\sigma_n = \rho_2$ always obtains the infimum, as illustrated in Fig. 2. Since $\mathcal{U} \cap G'$ is convex and closed, the optimal $\sigma_{p,n}$ in (12) are on the boundary of $\mathcal{U} \cap G'$, and σ_n lies either on the segment joining ρ_2 and ρ_0 or the one joining ρ_2 and the covariant state with coordinates $(-1 + 2/d, 0)$. We may w.l.o.g. assume the former, i.e., $\sigma_n = \lambda \rho_0 + (1 - \lambda)\rho_2$ for a $\lambda \in [0, 1]$. Considering decompositions $\rho = \alpha_p \sigma_p - \alpha_n \sigma_n$

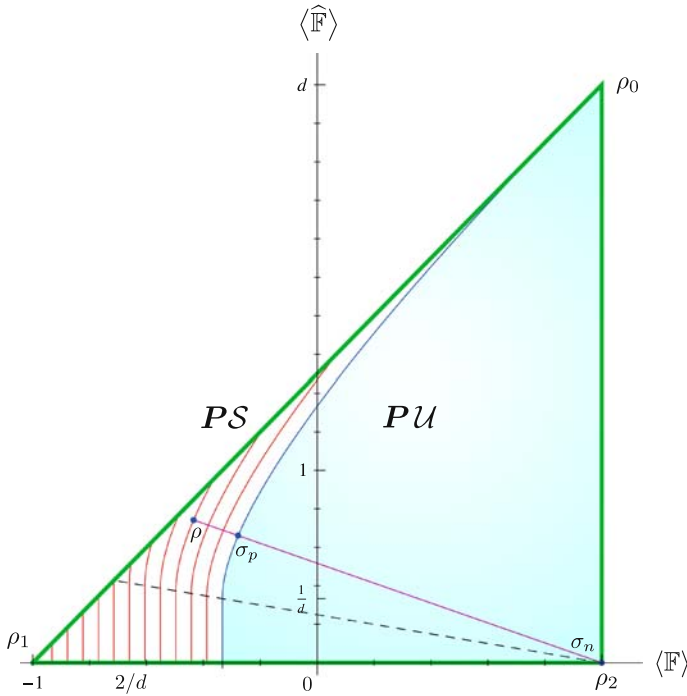


Fig. 2. Negativity as distance measure, exemplified by orthogonal covariant channels. $\mathcal{N}_{\mathcal{U}}$ is constant along red lines, which are obtained – in geometric terms – by a uniform scaling of the unitary channel boundary about ρ_2 as origin. $\sigma_{p,n} \in PU$ are the optimal states in (12) given $\rho \in PS \setminus PU$

with optimal σ_p (depending on λ) given ρ and σ_n via λ , both $\alpha_{p,n}$ are already determined by the x -coordinates of ρ and $\sigma_{p,n}$ due to $\alpha_p + \alpha_n = 1$. Note that the x -coordinates of ρ and σ_n remain fixed for all values of λ whereas the x -coordinate of σ_p is non-increasing as λ decreases, and so is α_n . That is, $\lambda = 0$ or equivalently $\sigma_n = \rho_2$ minimizes α_n .

It follows that a uniform scaling of the boundary of $\mathcal{U} \cap G'$ by a factor $(1 + \alpha_n)$ starting from ρ_2 as origin yields precisely the set of points with negativity α_n .

We may write each $\rho \in \mathcal{S}\mathcal{U} \cap G'$ in terms of a convex combination of the ρ_i listed in Table 16, that is, $\rho = \sum_i q_i \rho_i$ with $q_i \geq 0$ and $\sum_i q_i = 1$. From Fig. 2 it is evident that $q_1 > 0$ for all $\rho \notin \mathcal{U}$. Set $q := q_0/q_1$ and distinguish the following two cases due to the particular shape of $\mathcal{U} \cap G'$.

- $q > \frac{1}{d(d-1)}$. This corresponds exactly to the area above the dashed line in Fig. 2. Applying the scaling $(1 + \alpha_n)$ to the curve in Theorem 13, an explicit calculation shows that

$$\mathcal{N}_{\mathcal{U}}(\rho) = \frac{1}{d-2} \left(d-1 + \left(d + \frac{2}{d-2} \right) q - 2 \frac{d-1}{d-2} \sqrt{q^2 + \frac{d-2}{d-1} q} \right) q_1 - 1,$$

$$\rho = \sum_{i=1}^3 q_i \rho_i \in \mathcal{S}\mathcal{U} \text{ with } q_i \geq 0, \sum_i q_i = 1, \quad q := q_0/q_1.$$

- $q \leq \frac{1}{d(d-1)}$. In this case the negativity does not depend on q , and we get

$$\mathcal{N}_{\mathcal{U}}(\rho) = \frac{d}{d-1} q_1 - 1.$$

In particular, $\mathcal{N}_{\mathcal{U}}(\rho)$ is maximal exactly for the Werner-Holevo channel ρ_- , namely $\mathcal{N}_{\mathcal{U}}(\rho_-) = 1/(d-1)$.

V. Restoring Birkhoff’s Theorem

Measures quantifying the deviation of a unital channel from being a mixture of unitary channels are known to be not additive (or multiplicative). That is, a naive extrapolation from the ‘distance’ between a given $T \in \mathcal{S} \setminus \mathcal{U}$ and \mathcal{U} typically leads to an overestimation of the respective quantity for $T^{\otimes n}$, i.e., several copies of the channel. This effect was studied in detail in the context of the *entanglement of assistance* [12,21]

$$E_a(\rho) := \sup \left\{ \sum_i p_i S(\text{tr}_1 \Psi_i) : \rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \right\},$$

where $S(\rho) = -\text{tr}[\rho \log \rho]$ is the von Neumann entropy, and the supremum has to be taken over all convex decompositions of the given state $\rho \in \mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$ into pure ones. As $S(\text{tr}_1 \Psi_i) \leq \log d$ with equality iff Ψ_i is a maximally entangled state we have that $E_a(\rho) \leq \log d$ with equality iff $\rho \in \mathcal{U}$. It was shown in [12] that

$$\forall \rho \in \mathcal{S} : \lim_{n \rightarrow \infty} \frac{1}{n} E_a(\rho^{\otimes n}) = \log d,$$

which suggests that the approximation of $\rho^{\otimes n}$ by an element of \mathcal{U} improves as n increases. This would mean a restoration of Birkhoff’s theorem in the asymptotic limit. Whether this statement is valid in general when formulated in terms of norm distances (either for channels or, supposedly weaker, for states) remains an open problem [26].

In the following subsections we will prove it in the strongest possible sense for a class of $O(d)$ -covariant channels. We will see that at least for these cases neither the asymptotic limit nor an approximation is required—a remarkable effect from the perspective of environment-assisted error correction (Sec. V C).

More specifically we will show that for a $T \notin \mathcal{U}$ we find $T \otimes \tilde{T} \in \mathcal{U}$ when choosing

$$T : \mathcal{M}_d \rightarrow \mathcal{M}_d, \quad T(\rho) = \frac{1+\delta}{d} \text{tr}[\rho] \mathbb{1} - \delta \rho^T, \quad d \text{ odd}, \quad (21)$$

with appropriate δ and either $\tilde{T} = T$ (Sec. V A) or $\tilde{T} : \rho \in \mathcal{M}_D \mapsto \text{tr}[\rho] \mathbb{1} \in \mathcal{M}_D$ completely depolarizing (Sec. V B). The symmetry of the channels will help us in two stages: (i) we can use Thm.13 which tells us that $T \notin \mathcal{U}$ for $\delta > 1/(d+1)$, and (ii) it circumvents having to find an explicit decomposition in terms of unitary channels for $T \otimes \tilde{T}$: if the convex hull of the relevant G' expectation values of any set of unitary channels contains the ones of $T \otimes \tilde{T}$, then the twirling projection \mathbf{P} does the rest of the job.

A. *Two copies of a channel.* As usual we switch to the Jamiolkowski representation, where the family of channels in Eq. (21) becomes

$$\rho_T = \left(1 - \frac{1}{d} + \frac{\epsilon}{2}\right) \rho_- + \left(\frac{1}{d} - \frac{\epsilon}{2}\right) \rho_+, \quad d \text{ odd,}$$

where ρ_{\pm} are the normalized projections onto the symmetric and anti-symmetric subspace, respectively. The parametrization is chosen such that $\rho_T \in \mathcal{S} \setminus \mathcal{U}$ for $\epsilon \in (0, 2/d)$ since

$$\text{tr} [\rho_T \mathbb{F}] = -1 + \frac{2}{d} - \epsilon,$$

and $\mathbb{F} + (1 - 2/d) \mathbb{1}$ is a tight separation witness according to Thms.7,13. To exploit the full symmetry coming from $\tilde{T} = T$ we follow Sect. V.B of [24] and increment the tensor product symmetry group⁷ $\{(U \otimes U) \otimes (V \otimes V)\}$ by a flip operator which interchanges the tensor factors in the product $T \otimes \tilde{T}$. This results in a larger symmetry group G , thus yielding a smaller commutant $G' \subset \mathcal{M}_{d^4}$ which is spanned by $\mathbb{1}$ and

$$F := \frac{1}{2} (\mathbb{1} \otimes \mathbb{F} + \mathbb{F} \otimes \mathbb{1}), \quad F_{12} := \mathbb{F} \otimes \mathbb{F}.$$

That is, every state $\rho \in G'$ is now completely characterized by the expectation values/coordinates

$$(\langle F \rangle, \langle F_{12} \rangle)_\rho \equiv (\text{tr} [\rho F], \text{tr} [\rho F_{12}]).$$

Especially for any unitary channel described by $U \in \mathcal{M}_{d^2}$, setting $U_5 := \frac{1}{2} (U + U^T)$ and denoting partial transposes by T_i gives

$$\begin{aligned} \langle F_{12} \rangle_U &= \text{tr} \left[F_{12} (\mathbb{1} \otimes U) |\Omega\rangle \langle \Omega| (\mathbb{1} \otimes U^\dagger) \right] = \frac{1}{d^2} \text{tr} [U \bar{U}], \\ \langle F \rangle_U &= \text{tr} \left[F (\mathbb{1} \otimes U) |\Omega\rangle \langle \Omega| (\mathbb{1} \otimes U^\dagger) \right] \\ &= \frac{1}{2d^2} \text{tr} \left[U (\bar{U}^{T_1} + \bar{U}^{T_2}) \right] \\ &= \frac{1}{d^2} \text{tr} \left[U \bar{U}_5^{T_2} \right] = \frac{1}{d^2} \text{tr} \left[U_5 \bar{U}_5^{T_2} \right]. \end{aligned} \tag{22}$$

The last equation uses the fact that $\bar{U}_5^{T_2}$ is again symmetric. The ranges of the expectation values in Eq. (22) are studied in Appendix VII C. In particular for $d = 3$ we provide an explicit construction for the coordinates

$$(\langle F \rangle, \langle F_{12} \rangle)_\vartheta = \frac{1}{9} \left(-\frac{8}{3} (\cos \vartheta + 1)^2 + 3, 16 \cos^2 \vartheta - 7 \right), \quad \vartheta \in [0, \pi/2] \tag{23}$$

corresponding to convex combinations of unitary channels. Now matching the coordinates of $T \otimes T$,

$$(\langle F \rangle, \langle F_{12} \rangle)_T = \left(\text{tr} [\rho_T \mathbb{F}], \text{tr} [\rho_T \mathbb{F}]^2 \right) = (x, x^2), \quad x := -1 + \frac{2}{d} - \epsilon$$

⁷ The aim is to use the full symmetry group. That is, we use $V \otimes V$, $V \in U(d)$ for ρ_T which already allows us to discard $\hat{\mathbb{F}}$ from the commutant. We use further that the commutant of a tensor product is the tensor product of the commutants and that Prop. 10 remains true when adding the additional flip operator.

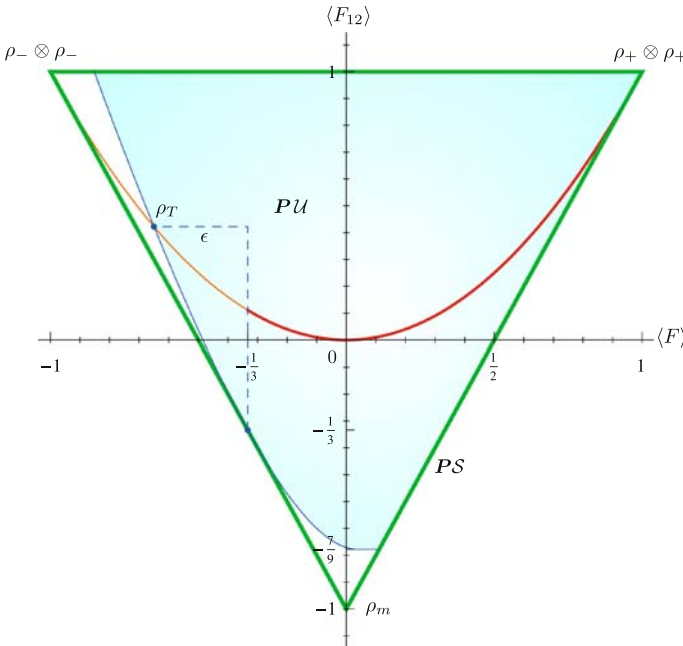


Fig. 3. Twofold tensor products of covariant channels for $d = 3$ (outer/green triangle). The mixtures of unitary channels correspond to the blue/shaded region, the orange parabola to tensor products of a channel with itself and the red part of that curve to the elements of \mathcal{U} on the single-channel level. Compare with Fig. 9 in [24]

with (23) yields

$$\epsilon = \frac{2}{3} \left(4 - 3\sqrt{2} - \sqrt{3} + \sqrt{6} \right) \approx \frac{1}{\sqrt{10}}$$

as shown in Fig. 3. The blue area corresponds to convex combinations of unitary channels,⁸ i.e., elements of \mathcal{U} , the orange curve to coordinates of single-channel tensor products and the red part of this curve to the elements of \mathcal{U} on the single-channel level. The state at the lower corner is

$$\rho_m = \frac{1}{2} (\rho_- \otimes \rho_+ + \rho_+ \otimes \rho_-).$$

As can be seen from direct inspection, each point on the curve (23) is an extreme point of the blue area. The remaining extreme points $(1, 1)$ and $(\frac{1}{9}, -\frac{7}{9})$ are realized by the unitaries $\mathbb{1}$ and $(\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix})^{\otimes 4} \oplus 1$, respectively.

B. Help from a noisy friend. Instead of adding a second copy of the channel, we will now supplement it by a completely depolarizing \tilde{T} . The Jamiolkowski representation of

⁸ A proof that (23) really solves the minimization problem (32) as acclaimed is still outstanding, but strongly supported by numerical tests. In any case, the set of convex combinations of unitary channels is *at least* as big as the blue area.

Table 1. Numerical minimization of (25). Note that the table is asymmetric w.r.t $d \leftrightarrow D$, and that the lower bound -1 can apparently be obtained for $d = 5$ and $D = 4$, that is, each channel-tensor product $T \otimes \tilde{T}$ (with \tilde{T} the completely depolarizing channel) becomes then a convex combination of unitaries

d	$D = 1$	2	3	4	5
1	1	1	1	1	1
3	$-\frac{1}{3}$	$-\frac{7}{9}$	$-\frac{23}{27}$	$-\frac{4}{5}$	$-\frac{37}{45}$
5	$-\frac{3}{5}$	$-\frac{23}{25}$	$-0.929\dots$	-1	$-0.976\dots$
7	$-\frac{5}{7}$	$-\frac{47}{49}$			

the completely depolarizing channel $\tilde{T} : \mathcal{B}(\mathcal{K}_1) \rightarrow \mathcal{B}(\mathcal{K}_2)$, $\mathcal{K}_i = \mathbb{C}^D$ is

$$\rho_{\tilde{T}} \equiv \left(\text{id} \otimes \tilde{T} \right) (|\Omega\rangle \langle \Omega|) = \mathbb{1}/D^2.$$

Let H be the corresponding symmetry group of all local unitaries $V_1 \otimes V_2$ with $V_i \in \mathcal{B}(\mathcal{K}_i)$, then $H' = \text{span}\{\mathbb{1}\}$. Using again that

$$(G \otimes H)' = G' \otimes H'$$

(see Example 7 in Sect. II.D of [24]) we get that every element of the commutant is completely characterized by the expectation value of $Y = \mathbb{F}_{\mathcal{H}_1 \otimes \mathcal{H}_2} \otimes \mathbb{1}_{\mathcal{K}_1 \otimes \mathcal{K}_2}$, yielding

$$\text{tr} [(\rho_T \otimes \rho_{\tilde{T}}) Y] = \text{tr} [\rho_T \mathbb{F}] = -1 + \frac{2}{d} - \epsilon. \tag{24}$$

Since $-\mathbb{1} \leq Y \leq \mathbb{1}$ every normalized state ρ satisfies $\text{tr} [\rho Y] \in [-1, 1]$. In order to obtain the subinterval of $[-1, 1]$ covered by convex combinations of unitary channels, we have to calculate $\langle Y \rangle_U$ for unitary $U : \mathcal{H}_1 \otimes \mathcal{K}_1 \rightarrow \mathcal{H}_2 \otimes \mathcal{K}_2$,

$$\text{tr} [(\mathbb{1} \otimes U) |\Omega\rangle \langle \Omega| (\mathbb{1} \otimes U^\dagger) Y] = \frac{1}{dD} \text{tr} [U \bar{U}^{T_2}]. \tag{25}$$

As $U = \mathbb{1}$ reaches the upper bound 1, the hard part is the lower bound which is treated in Appendix VII D. The results suggest that for $D = 2$, Eq. (25) gives

$$\left[-1 + \frac{2}{d^2}, 1 \right] \tag{26}$$

for the range in which $T \otimes \tilde{T} \in \mathcal{U}$ while, recall, $T \in \mathcal{U}$ only within $[-1 + 2/d, 1]$. The interval (26) can be related to the conjectured existence of a certain quaternion matrix, which we construct explicitly for $d = 3$ and $d = 5$. This means that in this case (25) covers this range *at least*. In particular, for $\epsilon \leq 2(d - 1)/d^2$, the expectation value (24) lies within this interval such that $\rho_T \otimes \rho_{\tilde{T}}$ becomes then indeed a convex combination of maximally entangled states.

For higher values of D , we reproduce Table 2 from Appendix VII D.

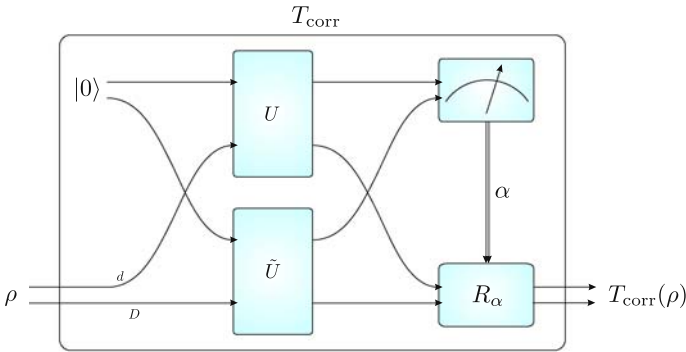


Fig. 4. The correction scheme as in [11], applied to the simultaneous usage of two noisy channels $T : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$ and $\tilde{T} : \mathcal{B}(\mathcal{K}_1) \rightarrow \mathcal{B}(\mathcal{K}_2)$. The channels are represented by unitary couplings U and \tilde{U} to an environment which is initially in a pure product state. The classical result α of the measurement on the global environment is used by the receiver who chooses the recovery operation R_α (again a quantum channel) accordingly. As discussed in the text, $T \otimes \tilde{T}$ can become perfectly correctable (i.e., $T_{\text{corr}} = \text{id}$) although neither T nor \tilde{T} is so

C. Environment-assisted error correction. The above results become especially remarkable from the point of view of environment-assisted error correction—a concept introduced in [11]. There it was studied which channels allow complete correction, given a suitable feedback of classical information from the environment (see Fig. 4). The class of perfectly correctable channels was identified with the set of convex combinations of unitary channels. In this way the above observations yield examples of channels which are not perfectly correctable on their own but become so when either taking several copies, or supplementing with a completely depolarizing channel.

VI. Discussion

The presented investigation of the set of unital quantum channels is to a large extent based on and inspired by methods and ideas from entanglement theory. The tools acquired in this context could be directly applied to the Jamiolkowski representation of the channel. This approach as such leads to questions about further analogies between the two fields. It would in particular be interesting whether a useful counterpart to *positive maps*, i.e., powerful non-linear criteria can be found for the separation of the set of mixtures of unitary channels.⁹

Clearly, the asymptotic Birkhoff conjecture[26] remains an important open problem for which the present work might be regarded as supporting evidence as it provides the first class of examples for which there is a rigorous proof. In this context it might be interesting to investigate $T \otimes \tilde{T}_D$ with \tilde{T}_D a D -dimensional maximally depolarizing channel, as studied in Sec. V B. Is there a dense subset of unital channels such that a finite D makes $T \otimes \tilde{T}_D$ a mixture of unitary channels?

Acknowledgements. We thank R. F. Werner for valuable discussions and the F, and J.I. Cirac for many useful comments and the steady support along the way. M.M.W. acknowledges financial support by QUANTOP and the Danish Natural Science Research Council (FNU).

⁹ Here it might be helpful to replace positivity by some form of contractivity.

VII. Appendix A–Matrix Optimization Problems

A. Minimizing $\text{tr}[A\bar{A}]$ subject to fixed singular values. We solve the minimization problem posed in Sect. III A which deals with a special class of separation witnesses. Considering any complex $d \times d$ matrix A , we denote its singular values by $\sigma_i(A)$ (counting multiplicity) such that $\sigma_1(A) \geq \dots \geq \sigma_d(A)$. Our main result is the following proposition, which is similar to Theorem 7.4.10 in [27].

Proposition 14. *Let $A \in \mathcal{M}_d$ and $\sigma_1 \geq \dots \geq \sigma_d$ denote the singular values of A . If $A\bar{A}$ is Hermitian, then there exists a permutation τ of $\{1, \dots, d\}$, an even $r \leq d$ and a function $\rho : \{1, \dots, r/2\} \rightarrow \{0, 1\}$ with*

$$\text{tr}[A\bar{A}] = 2 \sum_{i=1}^{r/2} (-1)^{\rho(i)} \sigma_{\tau(2i-1)} \sigma_{\tau(2i)} + \sum_{i=r+1}^d \sigma_{\tau(i)}^2. \tag{27}$$

Conversely, given any such τ, r, ρ and nonnegative numbers $\sigma_1 \geq \dots \geq \sigma_d$, there exists an $A \in \mathcal{M}_d$ such that $\sigma_i(A) = \sigma_i$ for all i and (27) holds.

Proof. We split the “ \Rightarrow ” part into the following steps:

1. $\text{tr}[A\bar{A}]$ and the singular values of A are invariant under $A \mapsto UAU^T$ for any unitary U . Note that this map sends $A\bar{A} \mapsto UA\bar{A}U^\dagger$, so by the spectral theorem, w.l.o.g. $A\bar{A}$ real diagonal.
2. It follows that $A\bar{A} = \overline{A\bar{A}} = \bar{A}A$, i.e. A commutes with \bar{A} , and each of the eigenspaces of $A\bar{A}$ is invariant under A and \bar{A} . Stated differently, A is block diagonal, each block corresponding to an eigenspace of $A\bar{A}$, so w.l.o.g. $A\bar{A} = \lambda \mathbb{1}$ for a $\lambda \in \mathbb{R}$.
3. The case $\lambda \neq 0$: Applying the singular value decomposition yields unitary matrices U, V such that $A_1 := U^\dagger AV = \text{diag}(\sigma_1, \dots, \sigma_d)$. Using $\bar{A} = \lambda A^{-1}$, we have

$$A_2 := V^\dagger \bar{A}U = \lambda \left(U^\dagger AV \right)^{-1} = \lambda A_1^{-1} = \lambda \text{diag} \left(\sigma_1^{-1}, \dots, \sigma_d^{-1} \right).$$

A_1 and A_2 sharing the same singular values translates to $\{\sigma_1, \dots, \sigma_d\} = |\lambda| \left\{ \sigma_1^{-1}, \dots, \sigma_d^{-1} \right\}$, so there is a permutation τ of $\{1, \dots, d\}$ with

$$\begin{cases} \sigma_{\tau(2i-1)} \sigma_{\tau(2i)} = |\lambda| & \text{for } i = 1, \dots, \frac{d}{2}, & d \text{ even} \\ \sigma_{\tau(2i-1)} \sigma_{\tau(2i)} = |\lambda| & \text{for } i = 1, \dots, \frac{d-1}{2}, \sigma_{\tau(d)}^2 = |\lambda|, & d \text{ odd.} \end{cases}$$

Note that d cannot be odd if $\lambda < 0$ as the negative eigenvalues of $A\bar{A}$ are of even algebraic multiplicity (see e.g. [27], pp. 252, 253). Concluding, $\text{tr}[A\bar{A}]$ can always be written in the form (27).

4. The case $\lambda = 0$: Let r denote the number of nonzero singular values of A . $A\bar{A} = 0$ means that $\text{range}(\bar{A}) \subseteq \text{kern}(A)$, so

$$r = \text{rank}(A) = \dim \text{range}(\bar{A}) \leq \dim \text{kern}(A) = d - r,$$

i.e. $2r \leq d$, and there is a permutation τ such that each summand in the right-hand side of (27) is zero.

To prove the “ \Rightarrow ” part, set $A := \text{diag} (A_1, \dots, A_{r/2}, \sigma_{\tau(r+1)}, \dots, \sigma_{\tau(d)})$ with

$$A_i := \begin{pmatrix} 0 & (-1)^{\rho(i)} \sigma_{\tau(2i)} \\ \sigma_{\tau(2i-1)} & 0 \end{pmatrix}.$$

□

Corollary 15. *Given any nonnegative numbers $\sigma_1 \geq \dots \geq \sigma_d$,*

$$\begin{aligned} & \min \{ \text{tr} [A\bar{A}] : A \in \mathcal{B}(\mathcal{H}), \sigma_i(A) = \sigma_i \forall i \} \\ & = \begin{cases} -2 \sum_{i=1}^{d/2} \sigma_{2i-1} \sigma_{2i}, & d \text{ even} \\ -2 \sum_{i=1}^{d-1/2} \sigma_{2i-1} \sigma_{2i} + \sigma_d^2, & d \text{ odd.} \end{cases} \end{aligned} \tag{28}$$

Proof. What remains to be shown is $A\bar{A}$ being Hermitian for optimal A ; then Proposition 14 guarantees optimality. Exploiting invariance under $A \mapsto VAV^T$ for unitary V we can w.l.o.g. assume that $A = UD$ with $D = \text{diag} (\sigma_1, \dots, \sigma_d)$ and U unitary. Now vary U to minimize $\text{tr} [A\bar{A}]$; the unitary constraint translates via

$$U \, dU^\dagger + dU \, U^\dagger = d (U \, U^\dagger) = 0$$

to $X := i \, dU \, U^\dagger$ being Hermitian. We have

$$\begin{aligned} 0 & \stackrel{!}{=} \frac{d}{dU} \text{tr} [DU\bar{D}U] \\ & = 2 \text{Re} \text{tr} [D \, dU\bar{D}U] \\ & = 2 \text{Im} \text{tr} [U\bar{D}U\bar{D}X]. \end{aligned}$$

This has to hold true for any Hermitian matrix X . Decomposing $U\bar{D}U\bar{D} = B_1 + iB_2$, B_1 and B_2 Hermitian, it follows that $B_2 = 0$. □

B. Maximizing $|\text{tr}[U]|$ subject to fixed $\text{tr}[U\bar{U}]$. In this subsection we prove Proposition 12 from Sect. IV B, i.e. we calculate the analytic solution of $\max |\text{tr} U|$ for fixed $\text{tr} [U\bar{U}]$ over all unitaries $U \in U(d)$. The motivation for this optimization problem comes from Eq. (18), which characterizes the convex hull of unitary channels within the set of orthogonal-covariant channels.

We need the following lemma first.

Lemma 16. *Let $U \in U(2)$ be a unitary 2×2 matrix and $U_s := \frac{1}{2} (U + U^T)$ be the symmetric part of U . Then the trace-norm of U_s equals*

$$\|U_s\|_1 \equiv \sum_{j=1,2} \sigma_j (U_s) = \sqrt{\text{tr} [U\bar{U}] + 2}.$$

Proof. There are $\alpha, \beta \in \mathbb{C}$ such that up to an unimportant phase factor

$$U = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \quad \text{with} \quad |\alpha|^2 + |\beta|^2 = 1.$$

Direct calculation shows that

$$\text{tr} [U\bar{U}] = 2 \left(|\alpha|^2 - \text{Re} (\beta^2) \right) = 4 \left(|\alpha|^2 + (\text{Im} \beta)^2 \right) - 2, \tag{29}$$

$$U_{\mathfrak{s}} = \begin{pmatrix} \alpha & i \text{Im} \beta \\ i \text{Im} \beta & \bar{\alpha} \end{pmatrix}, \quad U_{\mathfrak{s}} U_{\mathfrak{s}}^{\dagger} = \left(|\alpha|^2 + (\text{Im} \beta)^2 \right) \mathbb{1}, \tag{30}$$

$$\sum_{j=1,2} \sigma_j (U_{\mathfrak{s}}) \stackrel{(30)}{=} 2\sqrt{|\alpha|^2 + (\text{Im} \beta)^2} \stackrel{(29)}{=} \sqrt{\text{tr} [U\bar{U}] + 2}.$$

□

Proposition 17. *Let $d \geq 1$ be odd. Then for all $x \in [-1 + \frac{2}{d}, 1]$, there exists a unitary $U \in U(d)$ such that $d^{-1} \text{tr} [U\bar{U}] = x$, and*

$$\begin{aligned} & \max \left\{ d^{-1} |\text{tr} U| : U \in U(d), d^{-1} \text{tr} [U\bar{U}] = x \right\} \\ &= \left[\frac{1}{2} \left(1 - \frac{1}{d} \right) \left(1 - \frac{2}{d} + x \right) \right]^{1/2} + \frac{1}{d}. \end{aligned} \tag{31}$$

Proof. Set $\alpha := \frac{d}{d-1} \left[\frac{1}{2} \left(1 - \frac{1}{d} \right) \left(1 - \frac{2}{d} + x \right) \right]^{1/2} \in [0, 1]$, $\beta := \sqrt{1 - \alpha^2}$ and $\sigma := \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$, then a short calculation shows that $U := \text{diag} (\sigma, \dots, \sigma, 1)$ satisfies $d^{-1} \text{tr} [U\bar{U}] = x$ with $d^{-1} |\text{tr} U|$ equal to the right hand side of (31). So what remains to be shown is an upper bound on $d^{-1} |\text{tr} U|$.

Let $U \in U(d)$ be a unitary matrix with $d^{-1} \text{tr} [U\bar{U}] = x$. By the Youla theorem [25], given any *conjugate-normal* matrix A (that is, $AA^{\dagger} = \overline{A^{\dagger}A}$), there exists a unitary V such that VAV^T is a block diagonal matrix with diagonal blocks of order 1×1 and 2×2 , the 1×1 blocks corresponding to the real nonnegative eigenvalues of $A\bar{A}$ and the 2×2 blocks corresponding either to pairs of equal negative eigenvalues of $A\bar{A}$ or to conjugate pairs of non-real eigenvalues of $A\bar{A}$. Applying this to U , there is a unitary V with $U = VDV^T$, the block diagonal matrix D as described. Let $\frac{r}{2}$ be the number of 2×2 blocks (with even $r \leq d$) and denote these blocks by D_i . As U and V are unitary, so must be D , i.e. $D_i \in U(2)$ for all i . $U\bar{U}$ unitary guarantees $|\lambda| = 1$ for each eigenvalue λ of $U\bar{U}$, so each real nonnegative eigenvalue of $U\bar{U}$ must be 1. Altogether we have $D = \text{diag} (D_1, \dots, D_{\frac{r}{2}}, 1, \dots, 1)$. Set $c_i := \frac{1}{2} \text{tr} [D_i \bar{D}_i] \in [-1, 1]$ and $D_{\mathfrak{s}} := \frac{1}{2} (D + D^T)$. Using Lemma 16 and the fact that $V^T V$ is unitary and symmetric,

$$\begin{aligned} |\text{tr} U| &= \left| \text{tr} [D V^T V] \right| = \left| \text{tr} [D_{\mathfrak{s}} V^T V] \right| \leq \sum_{j=1}^d \sigma_j (D_{\mathfrak{s}}) \\ &= \sum_{i=1}^{r/2} \sum_{j=1,2} \sigma_j (D_{i,\mathfrak{s}}) + (d - r) = 2 \cdot \sum_{i=1}^{r/2} \sqrt{\frac{c_i + 1}{2}} + d - r. \end{aligned}$$

Let $c := d \cdot x \in [-d + 2, d]$. Some elementary analysis shows that the problem

$$\begin{aligned} & \max \left\{ 2 \cdot \sum_{i=1}^{r/2} \sqrt{\frac{c_i + 1}{2}} + d - r \right\} \\ \text{subject to } & 2 \cdot \sum_{i=1}^{r/2} c_i + d - r = c, \\ & c_i \in [-1, 1] \text{ for all } i, \\ & r \text{ even, } r \leq d \end{aligned}$$

has optimal solution $r = d - 1, c_i = \frac{c-1}{d-1} \forall i$ and the obtained maximum is $d \left[\frac{1}{2} \left(1 - \frac{1}{d} \right) \left(1 - \frac{2}{d} + \frac{c}{d} \right) \right]^{1/2} + 1$. This upper bound on $|\text{tr } U|$ corresponds exactly to the right-hand side of (31). \square

C. *Minimizing* $\text{tr} \left[U_{\mathfrak{s}} \overline{U_{\mathfrak{s}}}^{T_2} \right]$ *subject to fixed* $\text{tr}[U\overline{U}]$. Motivated by Eq. (22) in Sect. V A, we investigate

$$\min \left\{ \frac{1}{d_1 d_2} \text{tr} \left[U_{\mathfrak{s}} \overline{U_{\mathfrak{s}}}^{T_2} \right] : U_{\mathfrak{s}} = \frac{1}{2} (U + U^T), \right. \\ \left. U \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K}) \text{ unitary with } \frac{1}{d_1 d_2} \text{tr} [U\overline{U}] = y \right\} \tag{32}$$

for Hilbert spaces \mathcal{H} and \mathcal{K} with dimensions d_1 and d_2 , respectively, and provided $y \in [-1 + 2/d_1 d_2, 1]$, where $d_1 \geq 3$ is odd.

The partial transposes T_1 and T_2 are defined w.r.t. a fixed product basis by the linear extension of

$$\begin{aligned} (A \otimes B)^{T_1} &= A^T \otimes B \quad \text{and} \\ (A \otimes B)^{T_2} &= A \otimes B^T, \quad A \otimes B \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K}), \end{aligned}$$

respectively. Note that for any A and B ,

$$\text{tr} \left[A \overline{B}^{T_2} \right] = \text{tr} \left[A^{T_1} B^\dagger \right],$$

and for any real or complex A with $A^T = A$, the partial transposes are on equal footing, i.e. $A^{T_1} = A^{T_2}$, so (32) is inherently *symmetric* with respect to $d_1 \leftrightarrow d_2$. We identify $\mathcal{B}(\mathcal{H} \otimes \mathcal{K}) \cong \mathbb{C}^{d_1 d_2 \times d_1 d_2}$ by means of the ordered computational basis $(|11\rangle, |12\rangle, \dots, |1d_2\rangle, \dots, |d_1 d_2\rangle)$.

All quantities in (32), especially the minimizers, will stay invariant if we send

$$U \rightarrow (W_1 \otimes W_2) U (W_1 \otimes W_2)^T$$

with arbitrary unitaries $W_1 \in \mathcal{B}(\mathcal{H})$ and $W_2 \in \mathcal{B}(\mathcal{K})$.

Since every unitary matrix is also conjugate-normal (that is, $U U^\dagger = \overline{U^\dagger U}$), the Youla-theorem¹⁰ states that there exists a unitary matrix $V \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$ such that

$$U = V D V^T$$

with D real block-diagonal and blocks of size 1×1 and 2×2 , the former non-negative and the latter of the form $\begin{pmatrix} \sigma & -z \\ z & \sigma \end{pmatrix}$ with $\sigma \geq 0$. Since D must also be unitary, this equals $\begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}$ for a $\vartheta \in \mathbb{R}$, and all 1×1 blocks are 1. Note that

$$U_{\mathfrak{s}} = V D_{\mathfrak{s}} V^T, \quad D_{\mathfrak{s}} := \frac{1}{2} (D + D^T) \geq 0 \text{ diagonal,}$$

in particular, $D_{\mathfrak{s}}$ contains the singular values of $U_{\mathfrak{s}}$. Moreover, $\text{tr} [U \overline{U}] \equiv \text{tr} [D \overline{D}]$ is independent of V , so D fixes y in (32) and we may freely vary V . Conversely, Takagi’s theorem [27] asserts that every complex-symmetric matrix $A \in \mathbb{C}^{n \times n}$ can be decomposed into

$$A = V \text{diag} (\sigma_1, \dots, \sigma_n) V^T \tag{33}$$

with unitary V and $\sigma_i \geq 0$ for all i , so identifying $A \equiv U_{\mathfrak{s}}$ and $\text{diag} (\sigma_1, \dots, \sigma_n) \equiv D_{\mathfrak{s}}$, the minimization problem (32) can be reduced to the following problem and a subsequent optimization over $D_{\mathfrak{s}}$:

$$\min \left\{ \frac{1}{d_1 d_2} \text{tr} [A \overline{A}^{T_2}] : A \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K}), A^T = A, \sigma_i(A) = \sigma_i \forall i \right\}. \tag{34}$$

This closely resembles (28), and we have effectively decoupled the target function from the peculiar unitary constraint in (32).

Proposition 18. *Every (local) minimizer A of (34) satisfies $A \overline{A}^{T_2}$ Hermitian.*

Proof. Denote the derivative w.r.t. V in (33) by dV ; since V is unitary,

$$V dV^\dagger + dV V^\dagger = d(V V^\dagger) = 0$$

so $X := \frac{1}{i} dV V^\dagger$ must be Hermitian. Plugging

$$\frac{d}{dV} A = dV \text{diag} (\sigma_i) V^T + V \text{diag} (\sigma_i) dV^T = i (XA + AX^T)$$

into the target function (34) yields

$$\begin{aligned} \frac{d}{dV} \text{tr} [A \overline{A}^{T_2}] &= i \text{tr} \left[(XA + AX^T) \overline{A}^{T_2} \right] \\ &\quad - i \text{tr} \left[A \overline{(XA + AX^T)^{T_2}} \right] \\ &= 2i \text{tr} \left[X A \overline{A}^{T_2} \right] - i \text{tr} \left[A^{T_2} (XA + AX^T)^\dagger \right] \\ &= 2i \text{tr} \left[X \left(A \overline{A}^{T_2} - A^{T_2} \overline{A} \right) \right] \\ &= 2i \text{tr} \left[X \left(A \overline{A}^{T_2} - \text{h.c.} \right) \right] \stackrel{!}{=} 0. \end{aligned}$$

¹⁰ Refer in particular to Thm. 4 in [25]. The Youla-form corresponds to the Schur-form w.r.t. unitary congruence transformations $A \mapsto VAV^T$ and is a generalization of Takagi’s factorization [27].

As this must hold for any Hermitian X , the last equation can only be fulfilled if $A \overline{A}^{T_2}$ is Hermitian. \square

It is instructive to rewrite the target function as follows, setting $\sigma := (\sigma_1, \dots, \sigma_{d_1 d_2})$. Denote the columns of V by v_i , i.e. $V = (v_1 | \dots | v_{d_1 d_2})$, then $A = \sum_i \sigma_i v_i v_i^T$ and

$$\begin{aligned} \frac{1}{d_1 d_2} \text{tr} \left[A \overline{A}^{T_2} \right] &= \langle \sigma | G \sigma \rangle, \\ G &= (g_{ij}) \text{ Hermitian with} \\ g_{ij} &= \frac{1}{d_1 d_2} \text{tr} \left[v_j v_j^T \overline{v_i v_i^T}^{T_2} \right]. \end{aligned} \tag{35}$$

Writing $v_i =: \sum_k |k\rangle \otimes x_{ik}$, $x_{ik} \in \mathcal{K}$ the last expression becomes

$$\begin{aligned} g_{ij} &= \frac{1}{d_1 d_2} \sum_{k,k',l,l'} \text{tr} \left[(|l'\rangle \langle l| \otimes x_{j l'} x_{j l}^T) \left(|k'\rangle \langle k| \otimes (x_{i k'} x_{i k}^T)^\dagger \right) \right] \\ &= \frac{1}{d_1 d_2} \sum_{k,l} \text{tr}_{\mathcal{K}} \left[x_{j k} x_{j l}^T \overline{x_{i k}} x_{i l}^\dagger \right] \\ &= \frac{1}{d_1 d_2} \sum_{k,l} \langle x_{i k} | x_{j l} \rangle \langle x_{i l} | x_{j k} \rangle \\ &= \frac{1}{d_1 d_2} \text{tr} \left[s_{ij}^2 \right], \quad s_{ij} := (\langle x_{i k} | x_{j l} \rangle)_{k,l=1,\dots,d_1}. \end{aligned}$$

V being unitary translates to $\text{tr} s_{ij} = \sum_k \langle x_{i k} | x_{j k} \rangle = \langle v_i | v_j \rangle \stackrel{!}{=} \delta_{ij}$.

In what follows, we provide an explicit upper bound¹¹ of (32) for $d_1 = d_2 =: d$ and $d = 3$. Start with the *Ansatz* that all 2×2 blocks in D belong to the same phase, i.e.

$$\begin{aligned} D &= \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}^{\otimes 4} \oplus 1, \quad \vartheta \in [0, \pi/2], \quad \text{so} \\ \frac{1}{d^2} \text{tr} [U \overline{U}] &= \frac{1}{d^2} \text{tr} [D \overline{D}] = \frac{1}{9} (16 \cos^2 \vartheta - 7) \stackrel{!}{=} y, \end{aligned} \tag{36}$$

and set

$$V = \begin{pmatrix} \frac{i}{\sqrt{2}} & -\frac{i}{2\sqrt{6}} & & -\frac{i}{2\sqrt{2}} & & \frac{1}{\sqrt{3}} \\ \frac{i}{\sqrt{2}} & -\sqrt{\frac{3}{8}} & & \frac{1}{2\sqrt{2}} & & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \sqrt{\frac{3}{8}} & & -\frac{1}{2\sqrt{2}} & & \frac{1}{\sqrt{3}} \\ & \frac{i}{\sqrt{6}} & & \frac{1}{\sqrt{2}} & & \frac{i}{\sqrt{2}} \\ & & & -\frac{1}{\sqrt{2}} & & \frac{i}{\sqrt{2}} \\ & & & -\frac{1}{\sqrt{2}} & & \frac{i}{\sqrt{2}} \\ & -\frac{i}{\sqrt{2}} & -\frac{i}{2\sqrt{6}} & -\frac{i}{2\sqrt{2}} & & \frac{1}{\sqrt{3}} \end{pmatrix}$$

¹¹ Numeric tests strongly suggest that this is the actual minimum. Most interestingly, the acclaimed minimizer V does not depend on ϑ !

independent of ϑ ! Then $D_s = \text{diag}(\sigma_1, \dots, \sigma_9)$ with $\sigma_1 = \dots = \sigma_8 = \cos \vartheta$, $\sigma_9 = 1$, and G in (35) becomes

$$G = \frac{1}{27} \begin{pmatrix} \frac{3}{2} & 0 & -\frac{11}{8} & 0 & -\frac{9}{8} & 0 & 0 & 0 & -1 \\ 0 & \frac{3}{2} & \frac{1}{8} & 0 & \frac{3}{8} & -\frac{3}{2} & 0 & -\frac{3}{2} & -1 \\ -\frac{11}{8} & \frac{1}{8} & \frac{3}{2} & -\frac{1}{4} & -1 & \frac{1}{8} & -\frac{1}{4} & \frac{1}{8} & -1 \\ \hline 0 & 0 & -\frac{1}{4} & \frac{3}{2} & -\frac{3}{4} & 0 & -\frac{3}{2} & 0 & -1 \\ -\frac{9}{8} & \frac{3}{8} & -1 & -\frac{3}{4} & \frac{3}{2} & \frac{3}{8} & -\frac{3}{4} & \frac{3}{8} & -1 \\ 0 & -\frac{3}{2} & \frac{1}{8} & 0 & \frac{3}{8} & \frac{3}{2} & 0 & -\frac{3}{2} & -1 \\ \hline 0 & 0 & -\frac{1}{4} & -\frac{3}{2} & -\frac{3}{4} & 0 & \frac{3}{2} & 0 & -1 \\ 0 & -\frac{3}{2} & \frac{1}{8} & 0 & \frac{3}{8} & -\frac{3}{2} & 0 & \frac{3}{2} & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 \end{pmatrix} \in \mathbb{Q}^{9 \times 9}.$$

Finally evaluating the target function provides the supposed minimum

$$\frac{1}{d^2} \text{tr} \left[U_s \overline{U_s}^{T_2} \right] = \langle \sigma \mid G \sigma \rangle = \frac{1}{9} \left(-\frac{8}{3} (\cos \vartheta + 1)^2 + 3 \right) \tag{37}$$

with ϑ defined by (36).

Interestingly, the smallest eigenvalue $-\frac{1}{d^2}$ of G is of algebraic multiplicity 1 with corresponding eigenvector $\sigma(\vartheta)$ evaluated at $\vartheta = \pi/3$ and coordinates $-\frac{1}{3}(1, 1)$. Furthermore, (37) is minimal w.r.t. ϑ exactly for $\vartheta = 0$, which corresponds to maximal $y = \frac{1}{d^2} \text{tr} [U \overline{U}] = 1$ and $\sigma_1 = \dots = \sigma_{d^2} = 1$. In this case, D is the identity and $U = V V^T$ complex symmetric. Comparing with Appendix VII D, notice that we obtain the same minimum value $-\frac{23}{27}$.

D. Minimizing $\text{tr} [U \overline{U}^{T_2}]$. We explore the following minimization problem posed by Eq. (25) in Sect. V B.

$$\min \left\{ \frac{1}{d_1 d_2} \text{tr} [U \overline{U}^{T_2}] : U \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K}) \text{ unitary} \right\}, \tag{38}$$

where $\mathcal{H} \otimes \mathcal{K}$ is the tensor product of two Hilbert spaces with dimensions $d_1 = \dim \mathcal{H}$ and $d_2 = \dim \mathcal{K}$, respectively, d_1 being odd. The partial transposition is introduced in VII C. Note that any transformation

$$U \rightarrow (V^T \otimes W_1^\dagger) U (V \otimes W_2)$$

for unitary $V \in \mathcal{B}(\mathcal{H})$ and unitary $W_1, W_2 \in \mathcal{B}(\mathcal{K})$ leaves the target function invariant. If we allowed tensor products only, i.e. $U = U_1 \otimes U_2$, the target function would collapse to $\frac{1}{d_1} \text{tr} [U_1 \overline{U_1}] \geq -1 + \frac{2}{d_1}$, which is in general strictly greater than (38), see below. It is worth mentioning that (38) is inherently asymmetric w.r.t. $d_1 \leftrightarrow d_2$, as opposed to the previous Sect. VII C.

Proposition 19. $U \overline{U}^{T_2}$ is Hermitian for every minimizer U of (38).

Proof. As in previous sections, we differentiate the target function with respect to U . As U is unitary, $X := \frac{1}{i} dU U^\dagger$ must be Hermitian, and we get

$$\begin{aligned} \frac{d}{dU} \operatorname{tr} [U \bar{U}^{T_2}] &= i \operatorname{tr} [X U \bar{U}^{T_2}] - i \operatorname{tr} [U^{T_1} (XU)^\dagger] \\ &= i \operatorname{tr} [X (U \bar{U}^{T_2} - \text{h.c.})] \stackrel{!}{=} 0. \end{aligned}$$

This holds for any Hermitian X , so $U \bar{U}^{T_2}$ must be Hermitian, too. \square

Disassembly and reformulation. Let $\mathcal{X} = \mathcal{B}(\mathcal{K})$ be the Hilbert space equipped with the Hilbert-Schmidt inner product and induced Frobenius norm. By partitioning U as $U = \sum_{i,j=1}^{d_1} |i\rangle \langle j| \otimes u_{ij}$ with $u_{ij} \in \mathcal{X}$, we can now reformulate the target function (38) as

$$\operatorname{tr} [U \bar{U}^{T_2}] = \sum_{i,j} \langle u_{ij} | u_{ji} \rangle.$$

The condition U unitary translates to

$$\begin{aligned} U U^\dagger = \mathbb{1} &\Leftrightarrow \sum_i u_{ij}^\dagger u_{ik} = \delta_{jk} \mathbb{1} \quad \forall j, k = 1, \dots, d_1 \\ &\Leftrightarrow \sum_i \langle u_{ij} | u_{ik} x \rangle = \delta_{jk} (\mathbb{1} | x) \equiv \delta_{jk} \operatorname{tr} x \quad \forall j, k; x \in \mathcal{X} \end{aligned}$$

and the condition in Proposition 19 to

$$\begin{aligned} U \bar{U}^{T_2} \text{ Hermitian} &\Leftrightarrow \sum_i u_{ki} u_{ij}^\dagger = \sum_i u_{ik} u_{ji}^\dagger \\ &\Leftrightarrow \sum_i \langle u_{ij} | x u_{ki} \rangle = \sum_i \langle u_{ji} | x u_{ik} \rangle \quad \forall j, k; x \in \mathcal{X}. \end{aligned}$$

Note that these equations can be rewritten in terms of the Hilbert-Schmidt inner product as shown.

Quaternion structure. In this paragraph we assume $d_2 = 2$ and set $d = d_1$. Numeric tests suggest that in this case, the minimum value (38) is

$$-1 + \frac{2}{d^2}. \tag{39}$$

Interestingly, there emerges a substructure which is best described by quaternions. Recall that quaternions

$$\mathbb{H} = \{x_0 + x_1 i + x_2 j + x_3 k : x_0, \dots, x_3 \in \mathbb{R}\}$$

are a non-abelian division ring and form a 4-dimensional normed division algebra over the real numbers. We regard \mathbb{R} and \mathbb{C} as subalgebras of \mathbb{H} and denote the quaternion-conjugate of $q = x_0 + x_1 i + x_2 j + x_3 k \in \mathbb{H}$ by q^* . Furthermore, define $\operatorname{Re} q := x_0$ and $\vec{q} := q - \operatorname{Re} q = x_1 i + x_2 j + x_3 k$.

To bridge the gap between quaternions and operators on Hilbert spaces, employ the identification¹²

$$i \leftrightarrow i \sigma_z = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad j \leftrightarrow i \sigma_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad k \leftrightarrow i \sigma_x = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Note that in this representation

$$q \leftrightarrow \hat{q} = \begin{pmatrix} x_0 + i x_1 & x_2 + i x_3 \\ -x_2 + i x_3 & x_0 - i x_1 \end{pmatrix} \in \mathbb{C}^{2 \times 2}, \tag{40}$$

the quaternion conjugate is the Hermitian conjugate of the corresponding matrix, and the quaternion norm is the square root of the determinant,

$$q^* \leftrightarrow \hat{q}^\dagger, \quad \|q\| = \sqrt{\det(\hat{q})}.$$

Let \mathbb{H}^d be the d -dimensional “vector space” over \mathbb{H} with multiplication from the right, then each linear transformation on \mathbb{H}^d can be represented by a $d \times d$ matrix $A \in \mathbb{H}^{d \times d}$. The identification (40) provides an algebra isomorphism between $\mathbb{H}^{d \times d} \cong \mathbb{R}^{d \times d} \otimes \mathbb{H}$ and the complex $2d \times 2d$ matrices consisting of 2×2 blocks (40); to obtain Ax , define $u, v \in \mathbb{C}^d$ by $x =: u - jv$ and set $\hat{x} = (u_1, v_1, \dots, u_d, v_d)^T \in \mathbb{C}^{2d}$; then Ax corresponds exactly to $\hat{A}\hat{x}$. In the following it will be clear from context which representation is employed.

For all $A \in \mathbb{H}^{d \times d}$, the component-wise quaternion conjugate A^* and the quaternion conjugate-transpose A^\dagger are intuitively translated to

$$A^* \leftrightarrow \overline{\hat{A}}^{T_2}, \quad A^\dagger \leftrightarrow \hat{A}^\dagger.$$

Consequently, we say that A is Hermitian if $A^\dagger = A$.

As in [28], call $\lambda \in \mathbb{C}$, $\text{Im } \lambda \geq 0$ an eigenvalue of A with the corresponding eigenvector $x \in \mathbb{H}^d$ if $Ax = x\lambda$. These are exactly the eigenvalues of \hat{A} which have nonnegative imaginary part. Note that most of the well-known linear algebra theorems can be generalized straightforward to quaternions.

Proposition 20. *Let $A \in \mathbb{H}^{d \times d}$ be Hermitian with eigenvalues $-1 \pm d$, the algebraic multiplicity of $(-1 + d)$ being $\frac{d+1}{2}$ and of $(-1 - d)$ being $\frac{d-1}{2}$, respectively. Suppose A can be chosen such that $\text{Re } A = \frac{0}{2}$, then there exists a unitary $U \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$ with*

$$\frac{1}{2d} \text{tr} \left[U \overline{U}^{T_2} \right] = -1 + \frac{2}{d^2}.$$

Proof. Since $\text{Re } A = 0$, we have $A^* = -A$, and consequently $A^T = -A$. Set $U = \frac{1}{d}(\mathbb{1} + A)$, embedding $\mathbb{H}^{d \times d}$ into $\mathbb{C}^{d \times d} \otimes \mathbb{C}^{2 \times 2}$ as described above, then U will be Hermitian and unitary since the eigenvalues satisfy $\lambda(U) = \frac{1}{d}(1 + (-1 \pm d)) = \pm 1$. Using $\overline{U}^{T_2} \equiv \frac{1}{d}(\mathbb{1} + A^*) = \frac{1}{d}(\mathbb{1} - A)$, we get

$$\frac{1}{2d} \text{tr} \left[U \overline{U}^{T_2} \right] = \frac{1}{2d} \frac{1}{d^2} \text{tr} \left[\mathbb{1} - A^2 \right] = \frac{1}{d^2} \left(1 - (d^2 - 1) \right) = -1 + \frac{2}{d^2}.$$

The isomorphism (40) introduces an additional factor 2 into the trace, which cancels $\frac{1}{2}$ in the last equation. \square

¹² We adhere here to a different convention than e.g. [28].

Table 2. Numeric solutions of (38); the columns correspond to different values of d_2 . The case $d_2 = 1$ is treated analytically and is a special case of Sec. VII A. Note that $d_2 = 2$ is in agreement with (39), and for $d_1 = d_2 = 3$ we obtain the same value as the minimum of (37) with respect to ϑ

d_1	$d_2 = 1$	2	3	4	5
1	1	1	1	1	1
3	$-\frac{1}{3}$	$-\frac{7}{9}$	$-\frac{23}{27}$	$-\frac{4}{5}$	$-\frac{37}{45}$
5	$-\frac{3}{5}$	$-\frac{23}{25}$	$-0.92915 \dots$	-1	$-0.97632 \dots$
7	$-\frac{5}{7}$	$-\frac{47}{49}$			

Note that the conditions on A can be rephrased as follows. $A \in \mathbb{H}^{d \times d}$ is Hermitian with $\text{Re } A = 0$ such that $A^2 + 2A - (d^2 - 1)\mathbb{1} = 0$. The requirement $\text{Re } A = 0$ implies the respective eigenvalue multiplicities via $\text{tr } A = 0$.

To ensoul Proposition 20, we provide explicit examples of A meeting all requirements for $d = 3$ and $d = 5$, namely

$$A = 2 \begin{pmatrix} 0 & -i & j \\ i & 0 & -k \\ -j & k & 0 \end{pmatrix}$$

and

$$A = \begin{pmatrix} 0 & -2i & -\sqrt{12}j & 2k & -2j \\ 2i & 0 & 0 & -2j & 4k \\ \sqrt{12}j & 0 & 0 & -\sqrt{12}i & 0 \\ -2k & 2j & \sqrt{12}i & 0 & -2i \\ 2j & -4k & 0 & 2i & 0 \end{pmatrix}$$

respectively. Note that for these quaternion models, U is Hermitian, as well as $U \bar{U}^{T_2}$, in agreement with Proposition 19.

Higher dimensions. Table 2 contains numerical results for different values of d_1 and d_2 . We have simply employed $U = \exp[iX]$ with Hermitian X to represent unitary matrices. The local convergence error is about 10^{-6} , but it is still difficult to find the global minimizers. Quite remarkably, it seems that the lower bound -1 can be obtained for $d_1 = 5$ and $d_2 = 4$, even if we restrict to real orthogonal matrices.

VIII. Appendix B–A Special Extremal Channel

The following algebraic values for the coefficients μ_1, \dots, μ_4 of X in (7) are appropriate; we have obtained them basically by guessing and suppose that at least polynomial

degree 3 is required.

$$\begin{aligned} \mu_1 &= \frac{1}{6} \text{Root}_1 \left[-356 + 312x - 66x^2 + 3x^3 \right]^{1/2} \\ &= \frac{1}{6} \left(\frac{22}{3} - \frac{43 \cdot 2^{2/3}}{3 (977 + 213i\sqrt{7})^{1/3}} (1 + i\sqrt{3}) \right. \\ &\quad \left. - \frac{(977 + 213i\sqrt{7})^{1/3}}{3 \cdot 2^{2/3}} (1 - i\sqrt{3}) \right)^{1/2} \\ &\doteq 0.21821, \\ \mu_2 &= \text{Root}_1 \left[-1 + 432x^2 + 2592x^3 \right] \\ &= -\frac{1}{18} - \frac{1 + i\sqrt{3}}{18 (1 + 3i\sqrt{7})^{1/3}} - \frac{1}{72} (1 - i\sqrt{3}) (1 + 3i\sqrt{7})^{1/3} \\ &\doteq -0.14937, \\ \mu_3 &= \frac{1}{6}, \\ \mu_4 &= \text{Root}_2 \left[1 - 6x + 18x^3 \right] \\ &= -\frac{(1 + i\sqrt{3}) (-3 + i\sqrt{7})^{1/3}}{6 \cdot 2^{2/3}} - \frac{1 - i\sqrt{3}}{3 (2 (-3 + i\sqrt{7}))^{1/3}} \\ &\doteq 0.18595. \end{aligned}$$

The eigenvalues of X as calculated by a computer algebra program are then

$$\begin{aligned} \lambda(X) &= \left\{ 0, 0, \frac{1}{3}, \frac{1}{3}, \frac{1 \pm \sqrt{\alpha}}{6} \right\} \\ &\doteq \left\{ 0, 0, \frac{1}{3}, \frac{1}{3}, 0.23604, 0.097285 \right\} \end{aligned}$$

with

$$\begin{aligned} \alpha &= \text{Root}_1 \left[-25957 + 163107x - 78003x^2 + 6561x^3 \right] \\ &= \frac{107}{27} - \frac{104 \cdot 2^{2/3} (1 + i\sqrt{3})}{27 (67 + 23i\sqrt{7})^{1/3}} - \frac{13 (1 - i\sqrt{3}) (67 + 23i\sqrt{7})^{1/3}}{27 \cdot 2^{2/3}} \\ &\doteq 0.17329. \end{aligned}$$

In particular, X is positive semidefinite with rank equal to 4. Similarly, the linear independence of (6) can be verified explicitly.

References

1. Pérez-García, D., Wolf, M.M., Petz, D., Ruskai, M.B.: Contractivity of positive and trace-preserving maps under L_p norms. *J. Math. Phys.* **47**(8), 083506 (2006)
2. Arias, A., Gheondea, A., Gudder, S.: Fixed points of quantum operations. *J. Math. Phys.* **43**(12), 5872–5881 (2002)
3. King, C.: Additivity for unital qubit channels. *J. Math. Phys.* **43**, 4641–4653 (2002)
4. Fukuda, M.: Simplification of additivity conjecture in quantum information theory. *Quant. Inf. Comp.* **6**, 179–186 (2007)
5. Fukuda, M., Wolf, M.M.: Simplifying additivity problems using direct sum constructions. *J. Math. Phys.* **48**, 2101 (2007)
6. Rosgen, B.: Additivity and distinguishability of random unitary channels. *J. Math. Phys.* **49**, 102107 (2008)
7. Tregub, S.L.: Bistochastic operators on finite-dimensional von-Neumann algebras. *Soviet Math.* **30**, 105 (1986)
8. Kümmerer, B., Maassen, H.: The essentially commutative dilations of dynamical semigroups on M_n . *Commun. Math. Phys.* **109**, 1–22 (1987)
9. Landau, L.J., Streater, R.F.: On Birkhoff's theorem for doubly stochastic completely positive maps of matrix algebras. *Lin. Alg. Appl.* **193**, 107–127 (1993)
10. Birkhoff, G.: Three observations on linear algebra. *Univ. Nac. Tucuan. Revista A* **5**, 147–151 (1946)
11. Gregoratti, M., Werner, R.F.: Quantum lost and found. *J. Mod. Opt.* **50**, 915–933 (2003)
12. Smolin, J.A., Verstraete, F., Winter, A.: Entanglement of assistance and multipartite state distillation. *Phys. Rev. A* **72**, 052317 (2005)
13. Jamiolkowski, A.: Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.* **3**, 275–278 (1972)
14. Watrous, J.: *Mixing doubly stochastic quantum channels with the completely depolarizing channel*. [http://arXiv.org/abs/0807.2668v1\[quant-ph\]](http://arXiv.org/abs/0807.2668v1[quant-ph]), 2008
15. Rudolph, O.: On extremal quantum states of composite systems with fixed marginals. *J. Math. Phys.* **45**(11), 4035–4041 (2004)
16. Parthasarath, K.R.: Extremal quantum states in coupled systems. *Ann. l'Inst. H. Poincaré (B) Prob. Stat.* **41**, 257–268 (2005)
17. Audenaert, K.M.R., Scheel, S.: On random unitary channels. *New J. Phys.* **10**, 023011 (2008)
18. Choi, M.-D.: Completely positive linear maps on complex matrices. *Lin. Alg. Appl.* **10**, 285–290 (1975)
19. Buscemi, F.: On the minimal number of unitaries needed to describe a random-unitary channel. *Phys. Lett. A* **360**, 256–258 (2006)
20. Zeidler, E.: *Applied Functional Analysis*. Main Principles and Their Applications. New York: Springer, 1995
21. DiVincenzo, D.P., Fuchs, C.A., Mabuchi, H., Smolin, J.A., Thapliyal, A.V., Uhlmann, A.: Entanglement of assistance. In: *Proc. Quantum Computing and Quantum Communications*, First NASA Intl. Conf., Palm Springs, Berlin-Heidelberg-New York: Springer, 1999, pp. 247–257
22. Vidal, G., Werner, R.F.: Computable measure of entanglement. *Phys. Rev. A* **65**, 032314 (2002)
23. Werner, R.F., Holevo, A.S.: Counterexample to an additivity conjecture for output purity of quantum channels. *J. Math. Phys.* **43**, 4353–4357 (2002)
24. Vollbrecht, K.G.H., Werner, R.F.: Entanglement measures under symmetry. *Phys. Rev. A* **64**, 062307 (2001)
25. Faßbender, H., Ikramov, Kh.D.: Some observations on the Youla form and conjugate-normal matrices. *Lin. Alg. Appl.* **422**, 29–38 (2006)
26. Open problems in QIT. <http://www.imaph.tu-bs.de/qi/problems/>
27. Horn, R.A., Johnson, C.R.: *Matrix Analysis*. Cambridge, Cambridge University Press, 1990
28. Bunse-Gerstner, A., Byers, R., Mehrmann, V.: A quaternion qr algorithm. *Numer. Math.* **55**, 83–95 (1989)

Communicated by M.B. Ruskai