Unital Quantum Channels

by

Christian B. Mendl

Supervisor: Prof. Dr. J. Ignacio Cirac
Closing Date: June 3, 2008
Declaration of Authorship

I hereby declare that the work presented here is original and the result of my own investigations, except as acknowledged.

Garching, June 3, 2008

Christian B. Mendl
Unital Quantum Channels

Christian B. Mendl

June 3, 2008
Contents

1 Introduction 1

2 Fundamental Concepts 3
  2.1 Jamiołkowski isomorphism 5
  2.2 Kraus operators 6
  2.3 Transfer matrix representation 7
  2.4 Duality 8
  2.5 Contractivity and the $p$-Schatten norm 8
  2.6 Stinespring factorization 9
  2.7 Reformulating mixtures of unitaries 10
  2.8 Asymptotic version of Birkhoff's theorem 11
  2.9 Determinants 11

3 General Structure Analysis 13
  3.1 Separation witnesses 13
    3.1.1 Structure inheritance 13
    3.1.2 A special class of separation witnesses 14
  3.2 The affine span of unitary channels 14
  3.3 Extreme points 17

4 Covariant Channels 20
  4.1 Separation witnesses under symmetries 20
  4.2 Orthogonal covariant channels 21

5 Birkhoff's Theorem and Environment Assisted Error Correction 24
  5.1 Definition and correction setup 24
  5.2 Two copies of the same channel 26
  5.3 Adding the completely depolarizing channel 27

6 Negativity 29
  6.1 Base norms and unitaries 29
  6.2 Implementation for the Werner-Holevo channel 30

7 Conclusions 35

A Extreme Point Coefficients 37
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>B Matrix Optimization Problems</td>
<td>39</td>
</tr>
<tr>
<td>B.1 Trace subject to fixed singular values</td>
<td>39</td>
</tr>
<tr>
<td>B.2 Absolute trace under constraints</td>
<td>41</td>
</tr>
<tr>
<td>B.3 Trace, partial transposition and symmetry</td>
<td>42</td>
</tr>
<tr>
<td>B.4 Partial transposition, trace and distinct dimensions</td>
<td>46</td>
</tr>
<tr>
<td>Bibliography</td>
<td>50</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Overview Quantum channels capture in most general terms the transitions of quantum mechanical systems, covering a very broad range of effects. Transmission of photons through a glass fibre, energy dissipation and quantum noise introduced by the coupling to an environment, measurement and time evolution are just a few examples. The subclass of channels which leave the completely mixed state (corresponding to infinite temperature) invariant, so-called unital or doubly-stochastic channels, is particularly interesting. First, they have been identified with contractions (Pérez-García et al. 2006) and exhibit special fixed point properties (Arias et al. 2002) and second, it was recently shown that the additivity conjecture of the minimal output entropy can be reduced to unital channels (Fukuda 2007; Fukuda and Wolf 2007; Rosgen 2008). Interestingly, King (2002) has actually proved this additivity conjecture for unital qubit channels. Finally, Vollbrecht and Werner (2001) investigate highly symmetric channels, termed covariant channels, which are unital by nature, as is the convex hull of the reversible channels.

It turns out that the latter are precisely the unitary conjugations. From an experimentalist’s point of view, given a convex combination of these so-called unitary channels, the only error mechanism involved is caused by classical uncertainty. Moreover, the environment-assisted error correction scheme can completely restore the input fed into a quantum channel if and only if this channel is a convex combination (or mixture) of unitary channels (Gregoratti and Werner 2002).

A classical theorem by Birkhoff (1946) states that every doubly stochastic matrix is a convex combination of the reversible ones, i.e. permutations. However, the quantum analogue holds in 2 dimensions only, that is, for qubits, but fails in general for higher dimensions (Landau and Streater 1993). On the other hand, the results by Smolin et al. (2005) suggest that one may recover this theorem in the asymptotic limit.

Outline As motivated in the last paragraph, we start from unital quantum channels and focus on the relation between covariant channels and the convex structure of unitary channels.

- In chapter 2, the fundamental concepts related to quantum channels are concisely explained, laying the foundation for the forthcoming sections.
We also review some very recent results regarding quantum channels and shed light on the mathematical methods employed.

- In chapter 3 we provide a general (convex) structure analysis, including separation witness based on the Hahn-Banach theorem. We construct a channel which is extremal in the convex set of unital channels, but not extremal within the set of general channels. To the best of our knowledge, no such channel has been provided in the literature so far.

- In chapter 4 we develop mathematical theory for covariant unitary channels and determine the convex hull of $OO$-invariant states by a detailed calculation.

- In chapter 5 we apply the tools developed so far to construct a covariant channel which eludes the environment-assisted error correction scheme but can be brought back by unexpected countermeasures, namely, by using two copies of the same channel in parallel or adding the completely depolarising channel.

- As permitted by the linearity result in chapter 3, we introduce a distance measure in chapter 6 and illustrate it for covariant channels. Furthermore, we explicitly implement it for the Werner-Holevo channel.

- Finally, in chapter 7 we draw conclusions from the obtained results and sketch further lines of investigation.
Chapter 2

Fundamental Concepts

We introduce the basic concepts and ideas related to quantum channels or – interchangeably – quantum operations. Generally speaking, a quantum channel $T$ carries a physical system characterized by the density matrix $\rho$ to a new state $\rho'$,

$$\rho' = T(\rho).$$

The most prominent example is the unitary time evolution of a closed system, namely $\rho' = U \rho U^\dagger$ with $U = e^{-iHt/\hbar}$ and Hamiltonian $H$. Nevertheless, the notion of quantum channels is much broader, in particular, the transformation does not have to be unitary. Consider a system coupled to an environment obeying a master equation\(^2\), and additionally assume that the system is Markovian\(^3\), i.e. the density matrix $\rho(t)$ at time $t$ completely determines $\rho(t')$ for all $t' \geq t$. Then there exists a generator $L$ such that the quantum channel describing the time-evolution satisfies

$$T_t = e^{tL},$$

and $L$ can be written in Lindblad form (Lindblad 1976; Gorini and Kossakowski 1976; Nielsen and Chuang 2000), i.e.

$$\dot{\rho} = L(\rho) = \frac{i}{\hbar} [\rho, H] + \sum_{i,j} g_{ij} \left( F_i \rho F_j^\dagger - \frac{1}{2} \{ F_j^\dagger F_i, \rho \} \right)$$

(2.1)

Here, $H$ describes the Hamiltonian part and $G = (g_{ij})$ together with the $(F_i)$ form the Lindblad operators coupling the system to the environment and accounting for the dissipative part of the process. They are uniquely fixed by $G$ positive semidefinite and $\text{tr} F_i = 0$, $\text{tr} [F_i^\dagger F_j] = \delta_{ij}$ for all $i,j$. The second line in (2.1) basically re-expresses the first, so $\phi$ is a completely positive map.

But not only does the concept of time evolutions fit nicely to quantum channels, also measurements – which are by nature very different from continuous

\(^1\)For any complex matrix $A$ we denote the Hermitian conjugate by $A^\dagger$.

\(^2\)This equation can be derived from the evolution of the total configuration (system and environment) $\rho_{\text{tot}}$, and subsequently tracing out the environment, i.e. the state of the system is $\rho = \text{tr}_{\text{env}} [\rho_{\text{tot}}]$.

\(^3\)That is, the system is memoryless. In mathematical terms, $\rho(t)$ is a completely positive, continuous one-parameter semigroup.
Figure 2.1: Input-output mapping of a closed (left) and open (right) quantum system with a unitary $U$. $\rho$ is the density matrix of the principal system and $\rho_{\text{env}}$ describes the environment. Stinespring’s theorem guarantees that every quantum channel $T$ can be represented as on the right side.

processes – are special classes of quantum operations. Labeling the different eigenvalues of an observable by $m$, the corresponding channel (Nielsen and Chuang 2000, section 2.4.2) reads

$$T_m(\rho) = \frac{M_m \rho M_m^\dagger}{\text{tr} \left[M_m^\dagger M_m \rho \right]}, \quad \sum_m M_m^\dagger M_m = 1.$$

Despite this very broad approach, there are hard – physically motivated – requisites on quantum channels, which will be pinned down in definition 1.

This thesis focuses on the subclass of unital quantum channels, i.e. $T(1) = 1$. This simple to state feature is equivalent to $T$ being a contraction in the sense of the article by Pérez-García, Wolf, Petz, and Ruskai (2006) and is always fulfilled when considering convex combinations of unitary channels, which map $\rho \mapsto U \rho U^\dagger$ for a unitary $U \in \mathcal{B}(\mathcal{H})$. In terms of the Stinespring factorization $T(\rho) = V^\dagger \Phi(\rho) V$ treated in theorem 7 below, the operator $V$ is an isometry for unital channels.

Taking a more abstract viewpoint, we express the required fundamentals in mathematical terms. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be finite-dimensional Hilbert spaces; throughout the thesis, we focus on the case $\mathcal{H}_1 = \mathcal{H}_2 \equiv \mathcal{H}$, but the following definition remains valid in either case.

**Definition 1.** A linear map $T : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$ is called

- **trace preserving (tp) if** $\text{tr} \left[T(\rho)\right] = \text{tr} \left[\rho\right]$ for all $\rho \in \mathcal{B}(\mathcal{H}_1)$.

- **positive** if $T(\rho)$ is positive semidefinite for all positive semidefinite $\rho \in \mathcal{B}(\mathcal{H}_1)$.

- **completely positive (cp)** if $1 \otimes T : \mathcal{K} \otimes \mathcal{H}_1 \rightarrow \mathcal{K} \otimes \mathcal{H}_2$ is positive for any Hilbert space $\mathcal{K}$ of arbitrary dimension.

- **unital if** $T(1) = 1$.

$T$ is a quantum channel if it is completely positive and trace-preserving.

It turns out that $\dim \mathcal{K} = \dim \mathcal{H}$ suffices in the definition of "completely positive", i.e. $T$ is completely positive if and only if $1 \otimes T$ is positive on $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$. In physical terms, any density matrix on the joint system $\mathcal{K} \otimes \mathcal{H}$ must remain positive semidefinite when applying $T$ to the subsystem $\mathcal{H}$ and preserving $\mathcal{K}$.
Notation. Unless stated otherwise, we adopt the following conventions. $A^\dagger$ denotes the Hermitian conjugate of any complex matrix $A \in \mathbb{C}^{d \times d}$ with $d \in \mathbb{N}$, and $\overline{A}$ the component-wise complex conjugate. $U(d)$ is the unitary group of degree $d$. Throughout the thesis, $\mathcal{H} \cong \mathbb{C}^d$ will be a finite $d$-dimensional Hilbert space and $\mathcal{B}(\mathcal{H}) \cong \mathbb{C}^{d \times d}$ the set of bounded linear operators on $\mathcal{H}$. Some numeric quantities will appear in the thesis together with the symbol $\approx$, which means in this context truncation of the decimal representation to numerical accuracy without rounding. So instead of $\frac{1}{\sqrt{2}} = 0.707106\ldots$ or $\frac{1}{\sqrt{2}} \approx 0.707107$ we concisely write $\frac{1}{\sqrt{2}} = 0.707106$.

2.1 Jamiolkowski isomorphism

The maximally entangled state on the joint system $\mathcal{H}_1 \otimes \mathcal{H}_1$ defined by

$$|\Omega\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i, i\rangle$$

with $\{|i\rangle\}_i$ any orthonormal basis set of $\mathcal{H}_1$ – is the key ingredient of the following configuration. $|\Omega\rangle$ generalizes the EPR bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|0\rangle |0\rangle + |1\rangle |1\rangle)$.

The partial traces on $\mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ needed in the following proposition are defined by the linear extension of

$$\text{tr}_1[A \otimes B] := (\text{tr} A) \cdot B \in \mathcal{B}(\mathcal{H}_2),$$

$$\text{tr}_2[A \otimes B] := (\text{tr} B) \cdot A \in \mathcal{B}(\mathcal{H}_1).$$

Proposition 2. (Jamiolkowski isomorphism) The equation

$$\rho_T = (1 \otimes T)(|\Omega\rangle \langle \Omega|)$$

(2.2)

defines a linear isomorphism between completely positive maps $T : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$ and density matrices in $\mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Moreover, $T$ being trace-preserving/unital is equivalent to $\text{tr}_2[\rho_T] = 1/d$ and $\text{tr}_1[\rho_T] = 1/d$, respectively.

Proof. $\rho_T$ is positive semidefinite due to the fact that $T$ is completely positive. To recover $T$ from $\rho_T$, first decompose the density matrix – e.g. via spectral analysis – into

$$\rho_T = \sum_k |e_k\rangle \langle e_k| \quad \text{with} \quad |e_k\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2,$$

where the $|e_k\rangle$ need not be normalized. Next, define matrices $A_k : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ by

$$|e_k\rangle := \frac{1}{\sqrt{d}} \sum_{i} |i\rangle (A_k |i\rangle).$$

(2.3)

Expanding the maximally entangled state in equation (2.2) yields

$$\rho_T = \frac{1}{d} \sum_{i,j} |i\rangle \langle j| \otimes T(|i\rangle \langle j|),$$

(2.4)

\footnote{Compare with Nielsen and Chuang (2000, section 8.2.4).}

\footnote{The arguments that follow do not even require the $|e_k\rangle$ to be perpendicular.}
which allows us to express $T$ in terms of the $A_k$, namely $T(|i\rangle \langle j|) = d \langle i|_1 \rho T |j\rangle_1 = \sum_k A_k |i\rangle \langle j| A_k^\dagger$. Here the subscript 1 refers to the first tensor factor. By linearity,

$$T(\sigma) = \sum_k A_k \sigma A_k^\dagger \quad \text{for all } \sigma \in B(\mathcal{H}_1),$$

(2.5)

and $T$ is indeed completely positive as $(1 \otimes T)(\sigma) = \sum_k (1 \otimes A_k) \sigma (1 \otimes A_k)^\dagger$ for all $\sigma \in B(\mathcal{H}_1 \otimes \mathcal{H}_1)$.

Note that the link between the trace-preserving/unital property and the partial traces $\text{tr}_{1,2}[\rho_T]$ is established by a direct calculation starting from (2.4).

The simple fact that the isomorphism preserves convexity enables us to switch between representations when investigating the convex structure of quantum channels. Moreover, we can immediately deduce that the set of quantum channels must be bounded\(^6\) since the trace-preserving property forces $\|\rho_T\|_{\text{tr}} = 1$ for every quantum channel $T$.

### 2.2 Kraus operators

By the arguments in the last paragraph we have en passant proven the existence of another representation, which leads back to the original paper by Choi (1975) and early work by Kraus (Hellwig and Kraus 1969, 1970; Kraus 1983). The issue of uniqueness will be settled below.

**Proposition 3.** A linear operator $T : B(\mathcal{H}_1) \to B(\mathcal{H}_2)$ is completely positive if and only if there exist $N$ linear maps $A_k : \mathcal{H}_1 \to \mathcal{H}_2$ called **Kraus operators** such that

$$T(\rho) = \sum_{k=1}^N A_k \rho A_k^\dagger \quad \text{for all } \rho \in B(\mathcal{H}_1).$$

(2.6)

The minimal number of terms $N$ is called **Kraus rank** and is equal to the rank of the Jamiolkowski state $\rho_T$. In particular, the matrices $A_k$ can be chosen such that

$$N \leq d_1 d_2, \quad d_i := \dim \mathcal{H}_i \quad \text{and} \quad \text{tr} \left[ A_j^\dagger A_k \right] = 0 \quad \text{for all } j \neq k.$$

$T$ being trace-preserving and unital is equivalent to $\sum_k A_k^\dagger A_k = 1$ and $\sum_k A_k A_k^\dagger = 1$, respectively.

**Proof.** The representation (2.6) follows from the arguments in the last paragraph leading to (2.5). By identifying $A_k$ with $|e_k\rangle$ in (2.3) and noting that

$$\langle e_j | e_k \rangle = \frac{1}{d} \sum_i \langle i| A_j^\dagger A_k |i\rangle = \frac{1}{d} \text{tr} \left[ A_j^\dagger A_k \right],$$

the acclaimed choice is a direct result of the spectral theorem. The trace condition translates to

$$\text{tr} \left[ T(\rho) \right] = \sum_k \text{tr} \left[ A_k^\dagger A_k \rho \right] = \text{tr} \left[ \left( \sum_k A_k^\dagger A_k \right) \rho \right] = \text{tr} [\rho],$$

which can only be fulfilled by all $\rho$ if $\sum_k A_k^\dagger A_k = 1$. \(\square\)

In finite dimensions, all norms are equivalent!\(^6\)
Regarding the uniqueness of the Kraus representation, the following proposition is stated in Nielsen and Chuang (2000, theorem 8.2 in section 8.2.4).

**Proposition 4. (Unitary freedom of the Kraus representation)** Let \( \{ A_1, \ldots, A_N \} \) and \( \{ A'_1, \ldots, A'_{N'} \} \) be two sets of Kraus operators, w.l.o.g. \( N' \leq N \). Then they induce the same quantum channel \( T \) if and only if there exists an isometry \( V = (v_{ij}) \in \mathbb{C}^{N \times N'} \) such that

\[
A_i = \sum_j v_{ij} A'_j,
\]

This freedom is especially useful for quantum error correction since the insight into the correction process depends on the particular representation.

### 2.3 Transfer matrix representation

We endow \( B(\mathcal{H}) \) with the Hilbert-Schmidt inner product \( \langle A | B \rangle := \text{tr} [A^\dagger B] \) and induced Frobenius norm

\[
\| A \| = \sqrt{\text{tr} [A^\dagger A]} = (\text{tr} |A|^p)^{1/p} = \left( \sum_i \sigma_i(A)^p \right)^{1/p}, \quad p := 2,
\]

i.e. the \( p \)-Schatten for \( p = 2 \). In this way, \( B(\mathcal{H}) \) becomes a Hilbert space with orthonormal basis sets, e.g. matrix units \( e_{ij} := |i \rangle \langle j | \) or generalized Gell-Mann matrices, which are themselves generalizations of the Pauli \( \sigma \)-matrices. Thus, eigenvalues and singular values of \( T \) are understood as the respective quantities of the **transfer matrix** \( \hat{T} \) which is a matrix representation of \( T \) with respect to any such orthonormal basis set. For example, consider a Kraus decomposition (2.6) and conveniently choose matrix units \( (e_{ij}) \) as computational basis, then

\[
\langle e_{ij} | T (e_{lm}) \rangle = \sum_k \text{tr} \left[ |j \rangle \langle i| A_k |l \rangle \langle m| A_k^\dagger \right] = \sum_k \langle i| A_k |l \rangle \langle m| A_k^\dagger |j \rangle = \sum_k \langle ij| A_k \otimes A_k^\dagger |lm \rangle,
\]

so lexicographically ordering the matrix units yields

\[
\hat{T} = \sum_k A_k \otimes A_k^\dagger.
\]

With that it becomes easy to compute for example the trace of \( T \), which is of course independent of the specific chosen basis.

\[
\text{tr} T = \sum_k |\text{tr} A_k|^2 = d \langle \Omega | \rho_T | \Omega \rangle,
\]

especially \( \text{tr} T \geq 0 \) and \( \text{tr} T = 0 \) if and only if all Kraus operators are traceless, which is preserved by the unitary freedom introduced in proposition 4.
2.4 Duality

The dual map $T^*$ of $T$ is – by definition – the adjoint operator with respect to the Hilbert-Schmidt inner product, i.e.

$$\langle \sigma | T(\rho) \rangle = \langle T^*(\sigma) | \rho \rangle \quad \text{for all } \rho \in \mathcal{B}(\mathcal{H}_1) \text{ and } \sigma \in \mathcal{B}(\mathcal{H}_2).$$

Proposition 5. Let $T : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ be a linear map. Then $T$ is (completely) positive if and only if $T^*$ is. In this case, assuming $T$ has Kraus representation (2.6),

$$T^*(\rho) = \sum_k A_k^\dagger \rho A_k \quad \text{for all } \rho \in \mathcal{B}(\mathcal{H}_2)$$

is a Kraus representation of $T^*$. Furthermore, $T$ is trace-preserving if and only if $T^*$ is unital and vice versa.

Proof. Refer to section 1.1 in Landau and Streater (1993), or note that the assertions can be proven directly by the Kraus representation and the spectral theorem applied to $\sigma$ and $\tau$ in (2.8).

In particular, $T^*(1) = 1$ for any quantum channel $T$, which means in abstract terms that $T^*$ is a unital mapping from the $C^*$-algebra $\mathcal{A} \cong \mathbb{C}^{d \times d}$ into itself.

Let $\hat{T}(\rho) := \sum_k \hat{A}_k \rho \hat{A}_k^\dagger$ be another completely positive map and consider the composition $T^* \circ \hat{T}$. It follows from the last proposition that, interestingly,

$$\text{tr} \left[ T^* \circ \hat{T} \right] = \sum_{k,k'} \left| \text{tr} \left[ A_k^\dagger \hat{A}_{k'} \right] \right|^2 = \sum_{k,k'} \left| \langle A_k | \hat{A}_{k'} \rangle \right|^2,$$

which can only be zero if all Kraus operators are pairwise perpendicular.

2.5 Contractivity and the $p$-Schatten norm

Although not commonly used in general, the $p$-Schatten norm becomes a fundamental building block regarding contractivity of quantum channels. For all $A \in \mathcal{B}(\mathcal{H})$, set

$$\|A\|_p \equiv \|A\|_{p,\text{Schatten}} = (\text{tr} |A|^p)^{1/p} = \left( \sum \sigma_i(A)^p \right)^{1/p}$$

where $\sigma_i(A)$ are the singular values of $A$. For $p = 2$, this is the Frobenius norm (2.7) as above and $\mathcal{B}(\mathcal{H})$ becomes a Hilbert space by $\langle A | B \rangle := \text{tr} [A^\dagger B]$.

As Pérez-García et al. (2006) we endow $T$ with the operator norm

$$\|T\|_{p-p} := \sup_{\|A\|_p=1} \|T(A)\|_p$$

and call $T$ **contractive under the $p$-norm** if $\|T\|_{p-p} \leq 1$. Unless stated otherwise, we employ $p = 2$ and set

$$\|T\| := \|T\|_{2-2} = \|\hat{T}\|_{\infty,\text{Schatten}}$$

i.e. the greatest singular value of $\hat{T}$ in the transfer matrix representation. For an extensive discussion refer to Pérez-García et al. (2006), where inter alia the following interesting bound is obtained.
Theorem 6. Let $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be a positive and trace-preserving map, then

$$1 \leq \|T\| \leq \sqrt{d}, \quad d := \dim \mathcal{H}. \quad (2.9)$$

Moreover, the following are equivalent:

- $\|T\| = 1$
- $T(1) = 1$, i.e. $T$ is unital
- $T$ is a contraction.

In particular, this statement characterizes unital channels as contractions.

2.6 Stinespring factorization

The following theorem by Stinespring (1955) characterizes completely positive maps as the composition of a *-homomorphism and a conjugation $V^\dagger(\cdot)V$, which are themselves completely positive. It is phrased in an abstract operator-theoretic setting and poses no additional constraints on the underlying Hilbert spaces, in particular, it remains valid for infinite dimensions. A basic familiarity with $C^*$-algebras is helpful – otherwise think of a $C^*$-algebra as the set of bounded operators on a Hilbert space, where * denotes the operator adjoint.

Theorem 7. Let $\mathcal{A}$ be a unital $C^*$-algebra, $\mathcal{H}$ a Hilbert space and $T : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a linear map. Then a necessary and sufficient condition for $T$ being completely positive is that $T$ has the form

$$T(\alpha) = V^\dagger \Phi(\alpha)V \quad \text{for all } \alpha \in \mathcal{A}, \quad (2.10)$$

where $V$ is a bounded linear map from $\mathcal{H}$ to a Hilbert space $\mathcal{K}$ and $\Phi$ is a unital *-homomorphism of $\mathcal{A}$ into $\mathcal{B}(\mathcal{K})$. In this case, $\Phi(1)$ can be chosen 1.

Sketch of the proof. The hard part is the proof of sufficiency. Define a positive Hermitian bilinear form on the vector space $\mathcal{A} \otimes \mathcal{H}$ by the linear extension of

$$(\alpha \otimes x | \beta \otimes y) = \langle x | T(\alpha^* \beta) y \rangle_{\mathcal{H}},$$

and observe that there is a natural mapping $\Phi'$ from $\mathcal{A}$ to the linear maps on $\mathcal{A} \otimes \mathcal{H}$, namely

$$\Phi'(\alpha) \sum_i \beta_i \otimes y_i = \sum_i (\alpha \beta_i) \otimes y.$$

It can be shown that $(\Phi'(\alpha) \zeta | \Phi'(\alpha) \zeta) \leq \|\alpha\|^2 (\zeta | \zeta)$ for all $\zeta \in \mathcal{A} \otimes \mathcal{H}$, so the set $\mathcal{N}$ of all $\zeta \in \mathcal{A} \otimes \mathcal{H}$ with $(\zeta | \zeta) = 0$ is a linear subspace by the Schwarz inequality and invariant under $\Phi'(\mathcal{A})$. Let $\mathcal{K}$ be the completion of the pre-Hilbert space $(\mathcal{A} \otimes \mathcal{H}) / \mathcal{N}$ and $\Phi(\alpha)$ the induced bounded operator on $\mathcal{K}$ for each $\alpha \in \mathcal{A}$. Now define $V$ by

$$Vx = (1 \otimes x) + \mathcal{N} \quad \text{for all } x \in \mathcal{H},$$

which is bounded since $\|Vx\|^2 \leq (x | T(1) x)$. Finally observe that

$$\langle y | V^\dagger \Phi(\alpha)Vx \rangle_{\mathcal{H}} = (Vy | \Phi(\alpha)Vx)_{\mathcal{K}} = (1 \otimes y | \Phi'(\alpha)(1 \otimes x)) = (1 \otimes y | \alpha \otimes x) \overset{\text{def}}{=} \langle y | T(\alpha) x \rangle_{\mathcal{H}}$$

for all $x,y \in \mathcal{H}$, that is, one obtains the representation (2.10). \qed
Remark. Substituting $\alpha = 1$ in the last equation yields the restriction of $\langle \cdot | \cdot \rangle_{K}$ to $\mathcal{H}$ – as embedded by $V$ into $K$ –

$$\langleVy | Vx\rangle_{K} = \langle y | T(1)x\rangle_{\mathcal{H}}.$$ 

Especially for $T$ unital, $V$ is an isometry $H$ may be regarded as a sub-Hilbert space of $K$. In this sense, $T$ is the projection of $\Phi$ onto $\mathcal{H}$.

A Stinespring representation $(\Phi, V, K)$ can be chosen minimal in the sense that $K$ is the closed linear span of $\Phi(A)VH$. It can be shown that this minimal representation is unique up to unitary transformations.

As a corollary, another derivation of the Kraus operator representation can be obtained. For simplicity we switch back to the finite-dimensional setting and assume that $K = A \otimes H$. Identifying $A = B(H_1)$ and $H = H_2$, one gains the key-isomorphism

$$K \cong H_1 \otimes B(H_1, H_2)$$ such that

$$\Phi(\alpha)(x \otimes u) = (\alpha x) \otimes u$$

for all $\alpha \in A$, $x \in H_1$ and $u \in B(H_1, H_2)$. Let $(P_k)_k$ be a set of rank-1 projections on $B(H_1, H_2)$ with $P_k P_l = \delta_{kl} P_k$ and $\sum_k P_k = 1$, and define

$$P_k(x \otimes u) := x \otimes P_k(u).$$

Then it immediately follows that

$$P_k \Phi(\alpha) P_l = \delta_{kl} P_k \Phi(\alpha) P_k$$ for all $\alpha \in A$,

and since $H_1 \cong P_k(K)$ for each $k$, we may identify the last term with $\delta_{kl} \alpha$ and regard $A_k := V^\dagger P_k$ as operator from $H_1$ to $H_2$. Finally plugging everything into (2.10) yields

$$T(\alpha) = \sum_{k,l} (V^\dagger P_k \Phi(\alpha) P_l V) = \sum_k A_k \alpha A_k^\dagger$$

for all $\alpha \in A$, that is, the Kraus representation (2.6).

### 2.7 Reformulating mixtures of unitaries

Audenaert and Scheel (2007) reformulate the problem whether a unital quantum channel $T$ is a convex combination of unitary conjugations. They first note that the set of all possible matrices $Z$ (not necessarily square) with

$$\rho_T = ZZ^\dagger$$

is characterized by

$$Z = \rho_T^{1/2} R, \quad R \text{ right-unitary}.$$ 

Such a $Z$ is called square root of $\rho_T$ and corresponds precisely to a pure state decomposition. Denoting the generalized Gell-Mann matrices by $\tau_i$, the main theorem reads as follows.
Theorem 8. A unital quantum channel $T$ is a convex combination of unitary channels if and only if there exists a right-unitary $d^2 \times K$ matrix $R$ ($K \geq d^2$) such that

$$\operatorname{diag}(R^i G_i R) = 0 \text{ for all } i = 1, \ldots, d^2 - 1,$$

$$G_i := \rho_T^{1/2} (1 \otimes \tau_i) \rho_T^{1/2}.$$  

By Caratheodory’s theorem, $K \leq d^4 + 1$ suffices.

2.8 Asymptotic version of Birkhoff’s theorem

The entropy of entanglement $E$ is a central entanglement measure for bipartite pure states and defined by the von Neumann entropy $S$ of the reduced states, i.e.

$$E(|\psi \rangle \langle \psi|) := S(\operatorname{tr}_1 |\psi \rangle \langle \psi|) = S(\operatorname{tr}_2 |\psi \rangle \langle \psi|), \quad |\psi \rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2.$$  

Bennett et al. (1996) have shown that the asymptotically defined distillation rate of maximally entangled states is precisely $E$.

Considering three parties, Alice, Bob and a helper Charlie, the entanglement of assistance (Cohen 1998; DiVincenzo et al. 1999), which is given by

$$E_{\text{ass}}(\rho) := \max \left\{ \sum p_i E(|\psi_i \rangle \langle \psi_i|) : \rho = \sum p_i |\psi_i \rangle \langle \psi_i| \right\}$$

captures the idea that the helper Charlie can effect any pure state decomposition for Alice and Bob’s state. The asymptotic version of this quantity, the so-called regularization of $E_{\text{ass}}$, is

$$E_{\text{ass}}^\infty(\rho) := \frac{1}{n} E_{\text{ass}}(\rho^\otimes n).$$

Smolin et al. (2005) proved the following theorem.

Theorem 9. Given a pure tripartite state $|\psi \rangle_{ABC}$, then the optimal EPR rate distillable between A and B with the help of C under LOCC is

$$E_{\text{ass}}^\infty(\psi_{ABC}) = \min \{S(A), S(B)\}.$$  

This gives rise to the "asymptotic" version of Birkhoff’s theorem by noting that any Jamiołkowski state $\rho_T$ describing a unital quantum channel $T$ obtains the maximum asymptotic entanglement of assistance since $\operatorname{tr}_{1,2} [\rho_T] \sim 1$, suggesting complete correction by Charlie. Compare this with figure 5.2 in section 5; there, the channels which allow complete correction are precisely the convex combinations of unitaries.

2.9 Determinants

Determinants turn out to be an invaluable tool for understanding divisibility of quantum channels, that is, in the simplest formulation, whether a given channel
$T$ is a non-trivial composition of two channels $T_1$ and $T_2$, i.e. $T = T_1 T_2$. In that case,

$$\det T = \det (T_1 T_2) = (\det T_1)(\det T_2),$$

which frequently provides additional information about the channels $T_1$ and $T_2$, for example, if $|\det T_1| = 1$ or $|\det T_2| = 1$. In what follows, we mainly rely on chapter III in the treatise by Wolf and Cirac (2006).

The next theorem doesn’t even require the channel $T$ to be completely positive. $\lfloor \cdot \rfloor$ denotes the floor function.

**Theorem 10.** Let $T : B(\mathcal{H}) \to B(\mathcal{H})$ be a linear positive and trace-preserving map. Then

- $\det T$ is real and contained in the interval $[-1, 1]$.
- $|\det T| = 1$ if and only if $T$ is either a unitary conjugation $A \mapsto UAU^\dagger$ or unitarily equivalent to a matrix transposition $A \mapsto A^T$; in the former case $\det T = 1$ and in the latter $\det T = (-1)^{\lfloor d^2/2 \rfloor}$.

**Sketch of the proof.** The first assertion follows from the fact that the concatenation of quantum channels is again a quantum channel and thus bounded in norm (2.9), together with Gelfand’s formula for the spectral radius

$$\rho(T) = \lim_{n \to \infty} \|T^n\|^{1/n} \geq 1.$$

Concerning the second assertion, "$\Rightarrow$" is the hard part, so assume $|\det T| = 1$. Wolf and Cirac show that there exists an integer sequence $(n_i)$ such that $T_\infty := \lim_{i \to \infty} T^{n_i} = 1$, so $T^{-1} = T_\infty T^{-1} = \lim_{i \to \infty} T_0^{-1}$ is also a positive and trace-preserving map. From $T$ and $T^{-1}$ both being positive one infers that they map pure states to pure states and are unital. Now by theorem 2.5, both $T$ and $T^{-1}$ are contractions, so $T$ is actually norm-preserving. In particular, applied to pure states, $T$ preserves $|\langle \varphi | \psi \rangle|$ for all $|\varphi \rangle, |\psi \rangle \in \mathcal{H}$. A group theory argument by Wigner (1931) finally asserts that $T$ must be unitary or anti-unitary.

**Corollary 11.** (Monotonicity of the determinant)

- $T, T^{-1}$ are positive, trace-preserving linear endomorphisms on $B(\mathcal{H})$ if and only if $T$ is a unitary conjugation or a matrix transposition.
- $\det T$ is decreasing in magnitude under composition, that is,

$$|\det(TT^\dagger)| \leq |\det T|$$

for all positive trace-preserving $T$ and $T^\dagger$, and equality is equivalent to $T$ being unitary, a matrix transposition or $\det T = 0$.

We finally replicate an elegant formula for the determinant of an Markovian channel (Wolf and Cirac 2006, theorem 6).

**Theorem 12.** (Determinants of Markovian channels) Let $T = e^L$ be a Markovian quantum channel with generator $L$ as in (2.1). Then

$$\det T = \exp \{-d \text{tr} G\}.$$

Consult e.g. Lax (2002, theorem 4 in chapter 17.1).
Chapter 3

General Structure Analysis

3.1 Separation witnesses

The Hahn-Banach theorem is a very powerful tool in functional analysis and guarantees the existence of separation witnesses for convex sets. In section 3.1.1 we formulate this theorem for unital quantum channels, namely, given any nonempty closed convex subset $S$ of the unital quantum channels in the Jamiolkowski representation, we provide sufficient and necessary criteria whether a channel lies within $S$. In section 3.1.2 we solve the minimization problem resulting from a subclass of separation witnesses, applied to mixtures of unitary channels.

3.1.1 Structure inheritance

The general setup of this section is as follows. Let $H$ be a finite $d$-dimensional Hilbert space, $C := \{ \rho_T : \rho_T = (1 \otimes T) (|\Omega\rangle \langle \Omega|), \ T : B(H) \to B(H) \text{ cp, tp} \}$ = $\{ \rho \in B(H \otimes H) : \rho \geq 0, \ tr_1 \rho = \frac{1}{d}, \ tr_2 \rho = \frac{1}{d} \}$ the set of quantum channels in the Jamiolkowski representation and $C_{\text{unital}} := \{ \rho \in B(H \otimes H) : \rho \geq 0, \ tr_1 \rho = \frac{1}{d}, \ tr_2 \rho = \frac{1}{d} \}$ the subset of unital quantum channels.

In the following proposition we explicate the Hahn-Banach theorem for unital quantum channels.

**Proposition 13.** Assume that $S$ is a nonempty closed convex subset of $C_{\text{unital}}$, and let $\rho \in C_{\text{unital}}$ characterize a unital quantum channel. Then $\rho \in S$ if and only if

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

for all Hermitian operators $W \in B(H \otimes H)$ which satisfy

$$tr_1 W = \frac{1}{d}, \ tr_2 W = \frac{1}{d}, \ tr [W \sigma] \geq 0 \ \forall \sigma \in S.$$
Proof. It only remains to be shown that if $\rho \notin \mathcal{S}$, then there exists such a $W$ with $\text{tr}[W\rho] < 0$. First note that

$$\mathcal{X} := \{A \in \mathcal{B}(\mathcal{H} \otimes \mathcal{H}) : A \text{ Hermitian, } \text{tr}A = 0, \text{tr}2A = 0\}$$

is a real linear subspace of all Hermitian operators on $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$ and $C_{\text{unital}} - 1/d^2 \subset \mathcal{X}$. Set $\tilde{\rho} := \rho - 1/d^2 \in \mathcal{X}$. By the Hahn-Banach theorem (Zeidler 1995, theorem 1.C in chapter 1), there is a $\tilde{W} \in \mathcal{X}$ with

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

Now set $W := \tilde{W} + 1/d^2$ and observe that $\text{tr}\tilde{W} = 0$, so

$$\text{tr}[W\rho] = \text{tr}[\tilde{W}\rho] + \frac{1}{d^2}\text{tr}[\rho] = \text{tr}[\tilde{W}\rho] + \frac{1}{d^2} < 0,$$

and similarly $\text{tr}[W\sigma] \geq 0$ for all $\sigma \in \mathcal{S}$. □

### 3.1.2 A special class of separation witnesses

For any $B \in \mathcal{B}(\mathcal{H})$, we investigate $W = (1 \otimes B)F(1 \otimes B^\dagger)$ as separation witness for mixtures of unitary channels and solve the resulting minimization problem analytically in appendix B.1. Our general motivation is deriving nonlinear separation criteria; however, this is postponed for future research and not further developed here.

Plugging in $W$ leads to

$$\min \{\text{tr}[W(1 \otimes U)\Omega(1 \otimes U^\dagger)] : U \in U(d)\}$$

$$= \min \{\text{tr}[A\sigma] : \sigma(A) = \sigma(B)\}$$

which only depends on the singular values of $B$! We solve this problem in corollary 26 in appendix B.1.

### 3.2 The affine span of unitary channels

We prove in this section that the affine linear span of unitary quantum channels\(^1\) covers the whole set of unital quantum channels, that is, every unital quantum channel $T$ can be written as

$$T(\rho) = \sum_i \lambda_i U_i \rho U_i^\dagger, \quad \lambda_i \in \mathbb{R} \text{ for all } i, \quad \sum_i \lambda_i = 1, \quad (3.1)$$

where all $U_i \in \mathcal{B}(\mathcal{H})$ are unitary. Together with the results by Pérez-García et al. (2006) this leads to a novel characterization of unital channels in theorem 18.

As always, assume that $\mathcal{H}$ is a $d$-dimensional Hilbert space. Note that $\sum_i \lambda_i = 1$ in equation (3.1) is simply forced by the condition $T(1) = 1$. As a direct consequence, there is no other linear equation except for $T(1) = 1$ obeyed by the convex hull of unitary channels.

Generally speaking, we hope that the universal decomposition (3.1) may allow the preservation of general structures. For example, the question whether

\(^1\)i.e. the conjugations $\rho \mapsto U\rho U^\dagger$ with $U$ unitary
a given unital quantum channel is a convex combination of unitary channels is reduced to the question whether all \( \lambda_i \) in (3.1) can be chosen non-negative.

Let \( \mathcal{X} \) be the real vector space of all traceless, Hermitian operators on \( \mathcal{H} \),
\[
\mathcal{X} := \{ A \in \mathcal{B}(\mathcal{H}) : A = A^\dagger, \text{tr} A = 0 \} = \{ 1 \}^\perp
\]
equipped with the Hilbert-Schmidt inner product \( \langle A | B \rangle := \text{tr} [A^\dagger B] \), and
\[
\mathcal{V} := \{ A \mapsto UAU^\dagger : U \in \mathcal{B}(\mathcal{H}) \text{ unitary} \} \subset \mathcal{B}(\mathcal{X})
\]
the unitary conjugations on \( \mathcal{X} \). Note that \( \dim \mathcal{X} = d^2 - 1 \), and the operators in \( \mathcal{V} \) are orthogonal w.r.t. the Hilbert-Schmidt product as
\[
\langle A | UBU^\dagger \rangle = \text{tr} [A^\dagger UBU^\dagger] = \text{tr} [U^\dagger AUB] = \langle U^\dagger AU | B \rangle.
\]
That is, \( \mathcal{V} \) is a subgroup of \( O(\mathcal{X}) \), the group of orthogonal operators on \( \mathcal{X} \). Moreover, it is clear from the definitions that \( \text{span}_\mathbb{R} \mathcal{V} \) is an associative algebra over \( \mathbb{R} \) with composition as multiplication.

The main theorem 16 stated below first requires a few lemmas. As the case \( d = 1 \) is trivial, \( d \geq 2 \) is assumed throughout. We first choose a basis of \( \mathcal{X} \) as follows:
\[
\begin{align*}
\sigma^{jk}_x &:= |j \rangle \langle k| + |k \rangle \langle j| \quad \text{for all } j < k \\
\sigma^{jk}_y &:= -i (|j \rangle \langle k| - |k \rangle \langle j|) \quad \text{for all } j < k \\
\sigma^j_z &:= |j \rangle \langle j| - |j + 1 \rangle \langle j + 1| \quad \forall j = 1, \ldots, d - 1.
\end{align*}
\]
Then the set
\[
\{ \sigma^{jk}_x, \sigma^{jk}_y, \sigma^j_z \} \quad (3.2)
\]
is linearly independent and spans \( \mathcal{X} \).

In the following, we denote the subspace of real linear combinations of vectors \( \{ x_1, \ldots, x_n \} \) such that the coefficients sum to zero by \( \text{zerospan}_\mathbb{R} \{ x_1, \ldots, x_n \} \).
Explicitly,
\[
\text{zerospan}_\mathbb{R} \{ x_1, \ldots, x_n \} := \left\{ \sum_i \lambda_i x_i : \lambda_i \in \mathbb{R}, \sum_i \lambda_i = 0 \right\}.
\]

**Lemma 14.** For each base vector \( B \) in (3.2) there exists a \( T \in \text{zerospan}_\mathbb{R} \mathcal{V} \) which maps \( B \) to itself and all other base vectors to zero.

**Proof.** We explicitly construct such a \( T \) w.l.o.g. for \( \sigma^{12}_x \). Set
\[
T_1(\rho) := \frac{1}{2} \left( \rho + U_1 \rho U_1^\dagger \right), \quad U_1 := \begin{pmatrix}
1 & \sqrt{2} \\
1 & -1
\end{pmatrix}.
\]
Then for all \( \alpha \in \mathbb{R}^4 \)
\[
A := \left( \begin{array}{c|c}
\alpha \cdot \sigma_x & B^* \\
\hline
B & C
\end{array} \right) \rightarrow_{T_1} \left( \begin{array}{c|c}
\alpha \cdot \sigma_y & 0 \\
\hline
0 & C
\end{array} \right), \quad \sigma \equiv (\sigma_x, \sigma_y, \sigma_z, 1_2).
\]

15
Compose $T_1$ with

$$T_2(\rho) := \frac{1}{2} \left( \rho - U_2 \rho U_2^\dagger \right) \in \text{zerospan}_\mathbb{R} \mathcal{V}, \quad U_2 := \begin{pmatrix} \sigma_y & 0 \\ \overline{\sigma_x} & 1 \end{pmatrix},$$

which eliminates the $C$ block:

$$\begin{pmatrix} \alpha \cdot \sigma_x & 0 \\ 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} \alpha_1 \sigma_x + \alpha_3 \sigma_z & 0 \\ 0 & 0 \end{pmatrix}.$$

Finally, we map $\alpha_3$ to zero by $T_3$, which is the same as $T_2$, only $\sigma_y$ in $U_2$ replaced by $\sigma_z$. Then

$$T := T_3 \circ T_2 \circ T_1$$

is the desired operator: $T(A) = \alpha_1 \cdot \sigma_x^{12}$, and $T \in \text{zerospan}_\mathbb{R} \mathcal{V}$ as $T_2 \in \text{zerospan}_\mathbb{R} \mathcal{V}$.

**Lemma 15.** For any base vectors $B_1, B_2$ in (3.2), 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 unitary $U_1$ and $U_2$ such that $U_j^\dagger B_j U_j$ ($j = 1, 2$) is diagonal, respectively, so w.l.o.g. $B_1$ and $B_2$ diagonal since $\mathcal{V}$ is a group. Next, note that any two diagonal entries of $B_1$, for example the first two, can be permuted by

$$B_1 \mapsto U B_1 U^\dagger \quad \text{with} \quad U = \begin{pmatrix} \sigma_x & 0 \\ 0 & 1 \end{pmatrix}.$$  

Each basis vector has eigenvalues $(1, -1, 0, \ldots, 0)$, which are exactly these diagonal entries, so a composition of permutations maps $B_1$ to $B_2$. \qed

Now we combine these lemmas to prove the following theorem.

**Theorem 16.** $\mathcal{V}$ zero-spans all linear operators on $\mathcal{X}$, that is,

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

**Proof.** For any two base vectors $B_1, B_2$ in (3.2), 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 base vectors to zero, so a linear combination of these $T$s yields any linear map on $\mathcal{X}$. \qed

Denote the subset of all Hermitian operators in $\mathcal{B}(\mathcal{H})$ by $\mathcal{B}_h(\mathcal{H})$ and the affine linear span of vectors $\{x_1, \ldots, x_n\}$ by $\text{aff} \{x_1, \ldots, x_n\}$.

**Corollary 17.**

$$\text{aff}_\mathbb{R} \{ A \mapsto U A U^\dagger : U \text{ unitary, } A \in \mathcal{B}_h(\mathcal{H}) \} = \{ T : \mathcal{B}_h(\mathcal{H}) \rightarrow \mathcal{B}_h(\mathcal{H}) : T \text{ tp, unital } \}.$$
Proof. Only \( \supset \) needs to be shown. For any \( \text{tp} \), unital \( T \), the restriction of \( T - 1 \) to \( X \) is in \( \mathcal{B}(X) \), so by theorem 16 there are unitary \( U_1, \ldots, U_n \in \mathcal{B}(\mathcal{H}) \) and \( \lambda_1, \ldots, \lambda_n \in \mathbb{R} \) such that
\[
T - 1 \big|_X = \sum_i \lambda_i U_i (\cdot) U_i^\dagger, \quad \sum_i \lambda_i = 0.
\]
It follows that \( T(\rho) = \rho + \sum_i \lambda_i U_i \rho U_i^\dagger \) on \( X \), which evidently holds on \( \mathcal{B}_b(\mathcal{H}) \), too.

Finally, we combine corollary 17 with the results by Pérez-García et al. (2006) and section 2.5.

**Theorem 18. (Characterization of unital channels)**
Let \( T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}) \) be a quantum channel, then the following are equivalent:

1. \( T \) is unital, i.e. \( T(1) = 1 \),
2. \( T \) is contractive with respect to the \( p \)-Schatten norm for every \( p \in [1, \infty] \), that is, \( \| T \|_{p-p} \leq 1 \),
3. \( \| T \|_{p-p} \leq 1 \) for some \( p \in (1, \infty] \),
4. The transfer matrix \( \hat{T} \) is a convex combination of unitary maps on the bipartite system, i.e. \( \hat{T} = \sum p_i U_i \) with \( U_i \in U(d^2) \) and \( p_i \geq 0 \), \( \sum p_i = 1 \),
5. \( T \) is an affine-linear combination of unitary channels, that is, \( T(\rho) = \sum \lambda_i U_i \rho U_i^\dagger \) with \( \lambda_i \in \mathbb{R} \) and \( \sum \lambda_i = 1 \).

**Proof.** The relations 1. \( \Leftrightarrow \) 2. \( \Leftrightarrow \) 3. are established in Pérez-García et al. (2006). It holds true in general that the convex hull of unitary matrices is precisely the set of contractions with respect to the operator norm, which can be deduced from the singular value decomposition. This proves 2. \( \Leftrightarrow \) 4.. Finally, 1. \( \Leftrightarrow \) 5. is a direct consequence of corollary 17.

### 3.3 Extreme points

The compact set of quantum channels\(^2\) \( \mathcal{C}_d \) operating on a finite \( d \)-dimensional Hilbert space \( \mathcal{H} \) is convex, as is the subset of all unital quantum channels \( \mathcal{C}_d^{\text{unital}} \). So quite naturally we are interested in the extreme points of these sets, sketched in figure 3.1, and the question arises whether each extreme point of \( \mathcal{C}_d^{\text{unital}} \) is also an extreme point of \( \mathcal{C}_d \). In the following, we answer this question in the negative by constructing an explicit counterexample.

One building block is the Kraus operator representation of quantum channels\(^3\)
\[
T(\rho) = \sum_{k=1}^N A_k \rho A_k^\dagger, \quad N \leq d^2, \quad A_k \in \mathcal{B}(\mathcal{H}),
\]
where the trace-preserving property is equivalent to \( \sum_k A_k^\dagger A_k = 1 \).

Landau and Streater (1993) state the following theorem (for the duals), which leads back to Choi’s original paper (Choi 1975).

\(^2\)That is, completely positive, trace-preserving endomorphisms \( \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}) \).

\(^3\)Consult section 2.2 and section 8.2.4 in Nielsen and Chuang (2000).
Theorem 19. Let $T \in C_d$ be a quantum channel with Kraus representation (3.3). $T$ is extreme in the set $C_d$ if and only if the set of matrices

$$\left\{ A_k^* A_l \right\}_{k,l=1 \ldots N}$$

(3.4) is linearly independent. Assume further that $T$ is unital. Then $T$ is extreme in $C_{d^\text{unital}}$ if and only if

$$\left\{ A_k^* A_l \oplus A_l A_k^* \right\}_{k,l=1 \ldots N}$$

(3.5) is linearly independent.

We construct a channel $T$ serving as counterexample in the "minimal" configuration $d = 3$ and $N = 4$. Note that (3.4) can never be linearly independent for this setting as $N^2 \not\leq d^2 = \dim \mathcal{B}(H)$. Minimality is understood in the sense that for $d = 2$, the extreme points of $C_{d^\text{unital}}$ are precisely the unitary channels\(^4\) which are also extreme points in $C_d$, and we have found arguments suggesting that also for $d = 3$ and $N \leq 3$, each example is extremal in both sets.

Let $\rho_T$ be the Jamiolkowski representation of $T$,

$$\rho_T = (1 \otimes T)(|\Omega\rangle \langle \Omega|) = \sum_{k=1}^N |e_k \rangle \langle e_k|$$

(3.6) with

$$|e_k \rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i \rangle (A_k |i \rangle );$$

$T$ being trace-preserving and unital is equivalent to

$$\text{tr}_1 \rho_T = \text{tr}_2 \rho_T = 1/d.$$  

(3.7)

\(^4\)Refer e.g. to Landau and Streater (1993, theorem 1).
Now we employ the Ansatz
\[ \rho_T = \sum_{i,j=1}^{6} x_{ij} |\psi_i\rangle \langle \psi_j|, \]
where the \((|\psi_i\rangle)_i\) span the orthogonal complement of \((|kk\rangle)_k\), namely
\[
|\psi_1\rangle = \frac{1}{\sqrt{2}} (|12\rangle + |21\rangle),
|\psi_2\rangle = \frac{1}{\sqrt{2}} (|13\rangle + |31\rangle),
|\psi_3\rangle = \frac{1}{\sqrt{2}} (|23\rangle + |32\rangle),
|\psi_4\rangle = \frac{1}{\sqrt{2}} (|12\rangle - |21\rangle),
|\psi_5\rangle = \frac{1}{\sqrt{2}} (|13\rangle - |31\rangle),
|\psi_6\rangle = \frac{1}{\sqrt{2}} (|23\rangle - |32\rangle),
\]
and the Hermitian matrix \(X \equiv (x_{ij})\) is given by
\[
X := \frac{1}{3} \begin{pmatrix}
\frac{1}{2} & 0 & -i \mu_1 & i \mu_3 & i \mu_4 & 0 \\
0 & \frac{1}{2} & -i \mu_1 & -i \mu_4 & -(2 + i) \mu_3 & 0 \\
-i \mu_3 & i \mu_1 & \frac{1}{2} & 0 & 0 & 2 \mu_2 + i \mu_3 \\
i \mu_4 & i \mu_1 & 0 & \frac{1}{2} & 0 & -i \mu_1 \\
-i (i - 2) \mu_3 & 0 & 0 & \frac{1}{2} & 0 & i \mu_1 \\
0 & 0 & 2 \mu_2 - i \mu_3 & i \mu_1 & -i \mu_1 & \frac{1}{2}
\end{pmatrix}
\]
(3.8)

Stated in matrix notation,
\[ \rho_T = \Psi X \Psi^\dagger \quad \text{with} \quad \Psi = (|\psi_1\rangle \ldots |\psi_6\rangle). \]

This Ansatz guarantees that (3.7) always holds, which can be checked by an explicit calculation.

What remains is carefully choosing algebraic numbers \(\mu_1, \ldots, \mu_4 \in \mathbb{R}\) such that \(X\) is positive semidefinite with rank \(N = 4\), and that at the same time (3.5) is linearly independent when plugging in the corresponding Kraus operators; consult appendix A for the explicit formulas. The first requirement can be verified by Mathematica or a similar computer algebra system, but the second one involves an eigenvector decomposition (3.6) of \(\rho_T\), which is nearly infeasible given the elaborate structure of \(\mu_1, \ldots, \mu_4\). Instead of that, a numeric approach comes down to just calculating the rank of a complex matrix \(R \in \mathbb{C}^{N^2 \times 2d^2} = \mathbb{C}^{16 \times 18}\). We take the correspondig singular values, listed in descending order
\[ \sigma(R) \doteq \{ 0.44072, \ldots, 0.0057623 \} \]
as trustworthy indication that \(R\) has indeed full rank, as required.
Chapter 4

Covariant Channels

4.1 Separation witnesses under symmetries

In this section we relate the concept of separation witnesses to covariant channels. The basic setup and notation adopt (Vollbrecht and Werner 2001, section II.A). Let $\mathcal{H}$ be a $d$-dimensional Hilbert space and $G$ a closed subgroup of

$$\{U_1 \otimes U_2 : U_1, U_2 \in \mathcal{B}(\mathcal{H}) \text{ unitary}\}.$$ 

As $G$ is compact, there exists a unique normalized measure which is invariant under right and left group translation, the Haar measure. Vollbrecht and Werner (2001) define the resulting canonical projection, denoted twirl operation, by

$$P(A) := \int_G UAU^\dagger \, dU \quad \forall A \in \mathcal{B}(\mathcal{H} \otimes \mathcal{H}). \quad (4.1)$$

By definition, each element $A$ of the commutant $G'$ satisfies $[U, A] = 0$ for all $U \in G$, which is equivalent to $PA = A$, as follows directly from the invariance under group translations.

Denote the set of all density operators (positive semidefinite maps on $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$) by $S$ and let

$$S_U := \text{conv} \left\{ (1 \otimes U) |\Omega\rangle \langle \Omega| (1 \otimes U^\dagger) : U \in \mathcal{B}(\mathcal{H}) \text{ unitary} \right\}$$

be the compact subset of convex combinations of unitary channels in the Jamiolkowski representation. $P$ maps density operators to density operators, i.e. $PS \subset S$. Furthermore, $PS_U \subset S_U$, since for all unitaries $U \in \mathcal{B}(\mathcal{H})$

$$P(1 \otimes U) |\Omega\rangle \langle \Omega| (1 \otimes U^\dagger)$$

$$= \int_G (V_1 \otimes V_2) (1 \otimes U) |\Omega\rangle \langle \Omega| (1 \otimes U^\dagger) (V_1 \otimes V_2)^\dagger \, d(V_1 \otimes V_2)$$

$$= \int_G (1 \otimes V_2 U V_1^T) |\Omega\rangle \langle \Omega| \left(1 \otimes (V_2 U V_1^T)^\dagger\right) \, d(V_1 \otimes V_2)$$

is an element of $S_U$ as $S_U$ is convex.

The following definition constitutes the set of covariant separation witnesses, the use of which will be clear from the next proposition.

$$W_U(G) := \{W \in G' : W \text{ Hermitian,}$$

$$\text{tr} \left[(1 \otimes U) |\Omega\rangle \langle \Omega| (1 \otimes U^\dagger) W \right] \geq 0 \, \forall U \in \mathcal{B}(\mathcal{H}) \text{ unitary} \}. $$
The crucial point here is that \( \mathcal{W}_U(G) \subset G' \), which is in general much smaller than the set of all Hermitian operators on \( \mathcal{H} \otimes \mathcal{H} \). This idea correlates directly with the arguments by Vollbrecht and Werner (2001), namely that \( \rho \in P_S \) is already determined by its expectation values \( \text{tr}[\rho A] \) for \( A \in G' \).

**Proposition 20.** Let \( \rho \in P_S \) be a density operator which is invariant under \( G \). Then

\[
\rho \in P_{S_U} \iff \text{tr}[\rho W] \geq 0 \text{ for all } W \in \mathcal{W}_U(G).
\]

*Proof.* Since \( S_U \) is convex, one may equally define \( \mathcal{W}_U(G) \) as

\[
\mathcal{W}_U(G) := \{ W \in G' : W \text{ Hermitian, } \text{tr}[\rho W] \geq 0 \forall \rho \in S_U \}.
\]

Observe that \( \text{tr}[\rho W] = \text{tr}[\rho P(W)] = \text{tr}[P(\rho)W] \), so

\[
\mathcal{W}_U(G) = \{ W \in G' : W \text{ Hermitian, } \text{tr}[\rho W] \geq 0 \forall \rho \in P_{S_U} \}.
\]

The proposition is now an immediate consequence of the Hahn-Banach separation theorem since \( P_{S_U} \subset G' \) is convex and compact.

4.2 Orthogonal covariant channels

In this section, we exemplify the criteria in proposition 20 above for the orthogonal group\(^1\) \( G = O \), i.e. \( G = \{ O \otimes O : O \in \mathcal{B}(\mathcal{H}) \text{ real orthogonal} \} \) (Vollbrecht and Werner 2001, section II.D, example 3). The abelian commutant \( G' \) is spanned by three Hermitian operators,

\[
G' = \langle \{1, F, \hat{F}\} \rangle,
\]

where \( \hat{F} = d |\Omega\rangle \langle \Omega| \) and \( F \) is the flip operator defined in the paper by Vollbrecht and Werner. These operators commute pairwise, i.e. can be simultaneously diagonalized, which enables us to decompose \( G' \) into the so-obtained projections onto the corresponding eigenspaces,

\[
\begin{align*}
P_0 &= \frac{1}{d} \hat{F} = |\Omega\rangle \langle \Omega| \\
P_1 &= \frac{1}{2} (1 - F) = A_2 \\
P_2 &= \frac{1}{2} (1 + F) - \frac{1}{d} \hat{F} = S_2 - |\Omega\rangle \langle \Omega|,
\end{align*}
\]

where \( A_2 \) and \( S_2 \) are the 2-particle (anti-)symmetrizers. Consequently, \( P_S \) is the convex hull of the corresponding normalized density matrices \( \rho_i = [\text{tr} P_i]^{-1} P_i \).

Note that \( \rho_1 \) corresponds exactly to the Werner-Holevo channel. As explained by Vollbrecht and Werner in section II.A, each covariant state \( \rho \in P_S \) is completely characterized by its "coordinates"

\[
\left( \langle F \rangle, \langle \hat{F} \rangle \right)_\rho \equiv \left( \text{tr}[\rho F], \text{tr}[\rho \hat{F}] \right).
\]

Especially for the \( \rho_1 \), we obtain

---

\(^1\)Fix a basis of \( \mathcal{H} \) such that the component-wise complex conjugate \( \overline{A} \) of any \( A \in \mathcal{B}(\mathcal{H}) \) makes sense.
as illustrated in figure 4.1.

Figure 4.1: The set of orthogonal covariant channels in the Jamiolkowski representation $PS$ (green triangle) and the convex subset of unitary channels $PS_U$ (blue area), which is described in analytic terms by proposition 22. Note that the Werner-Holevo channel is "furthest away" from the unitaries. Compare with Vollbrecht and Werner (2001, example 3 and figure 2).

Regarding $PS_U$, proposition 20 guarantees that we may identify

$$PS_U \cong \left\{ \left( \langle F \rangle, \langle \tilde{F} \rangle \right)_\rho : \rho \in S_U \right\}$$

$$= \text{conv} \left\{ \left( \langle F \rangle, \langle \tilde{F} \rangle \right)_U : U \in \mathcal{B}(\mathcal{H}) \text{ unitary} \right\};$$

(4.3)

a short calculation shows that

$$\langle F \rangle_U \equiv \text{tr} \left[ (1 \otimes U) \langle \Omega | \langle \Omega | (1 \otimes U^\dagger) F \right] = \frac{1}{d} \text{tr} \left[ U U^\dagger \right],$$

$$\langle \tilde{F} \rangle_U \equiv \text{tr} \left[ (1 \otimes U) \langle \Omega | \langle \Omega | (1 \otimes U^\dagger) \tilde{F} \right] = \frac{1}{d} |\text{tr} U|^2.$$  

(4.4)
Proposition 21. If \( d \) is even, then \( \mathbf{PSU} = \mathbf{PS} \), in particular, in this case the Werner-Holevo channel \( \rho_{\text{WH}} \sim A_2 \) is a convex combination of unitary channels.

Proof. It suffices to note that the expectation values (4.2) with respect to \( \rho_i \) and \( U_i \) coincide for \( U_0 = 1 \), \( U_1 = \text{diag}(\sigma_y, \ldots, \sigma_y) \) and \( U_2 = \text{diag}(\sigma_z, \ldots, \sigma_z) \), just by plugging in (4.4).

So interesting structures only emerge for \( d \) odd, which we will treat now.

Proposition 22. Let \( d \geq 3 \) be odd. Then the extreme points of (4.3) are

\[
(-1 + 2/d, 0), \quad (1, 0) \quad \text{and} \quad \{(x, m(x)) : x \in [-1 + 2/d, 1]\} \tag{4.5}
\]

with

\[
m(x) := d \left( \frac{1}{2} \left( 1 - \frac{1}{d} \right) \left( 1 - \frac{2}{d} + x \right) \right)^{1/2} + \frac{1}{d} \right)^2.
\]

Proof. "(4.3) \( \subset \operatorname{conv}(4.5) \)"; For all unitary \( U \in U(d) \),

\[
\frac{1}{d} \operatorname{tr} [U\overline{U}] \in [-1 + 2/d, 1],
\]

which follows from the fact that, for any matrix \( A \), the spectrum of \( A\overline{A} \) is symmetric with respect to the real axis, the eigenvalues \( \lambda, \overline{\lambda} \) have the same algebraic multiplicity, and the algebraic multiplicity of all negative eigenvalues of \( A\overline{A} \) (if any) is even (Faßbender and Ikramov 2006). Together with proposition 28 in the appendix we get the required bounds on (4.3).

"(4.5) \( \subset \) (4.3)"; Set

\[
Q_0 := \frac{1}{2} \left( \begin{array}{ccc} 0 & 1 - i & -1 - i \\ -1 + i & -i & 1 \\ 1 + i & 1 & i \end{array} \right)
\]

and \( \varphi := \exp(2\pi i/3) \), then the coordinates (4.5) are obtained by

\( U_0 = \text{diag}(\sigma_y, \ldots, \sigma_y, Q_0) \), \( U_1 = \text{diag}(\sigma_z, \ldots, \sigma_z, \varphi, \varphi^2, 1) \) and the unitary matrices which solve the maximum problem (B.5).
Chapter 5

Birkhoff’s Theorem and Environment Assisted Error Correction

5.1 Definition and correction setup

We prove the somewhat surprising existence of a noisy quantum channel $T$ such that restoring quantum information by a measurement on the environment (Gregoratti and Werner 2002) is impossible, but becomes possible when enlarging the system and applying $T$ simultaneously with another noisy quantum channel $\tilde{T}$, even if intuitively, $\tilde{T}$ does not improve the situation at all. Namely, we set

1. $\tilde{T} = T$, i.e. two copies of the same channel
2. $\tilde{T} : \rho \mapsto \frac{1}{D}$, the completely depolarizing channel.

The basic setup is illustrated in figure 5.1. In the sense of (Gregoratti and Werner 2002) the correction scheme depicted in figure 5.2 restores quantum information if the measurement can be performed in such a way that the overall channel $T_{\text{corr}}$ is the identity.

In the following paragraphs we adopt the notation of Vollbrecht and Werner (2001). The Hilbert spaces $H_1$ and $K_1$ belong to Alice, whereas $H_2$ and $K_2$ belong to Bob. It is throughout assumed that $d := \dim H_1 = \dim H_2$ and $D := \dim K_1 = \dim K_2$, and w.l.o.g. $d$ odd, since otherwise everything is a convex combination of unitary channels. $D$ will of course be equal to $d$ when considering $T \otimes T$.

We introduce a small parameter $\epsilon > 0$ which will be fixed later and define $T$ by its Jamiolkowski representation

$$\rho_T = (1 - \lambda) \rho_{WH} + \lambda \rho_{\text{sym}}, \quad \lambda = \frac{1}{d} - \frac{\epsilon}{2}$$

Here,

$$\rho_{WH} = \frac{1}{d(d - 1)} (1 - F) \quad \text{and}$$

$$\rho_{\text{sym}} = \frac{1}{d(d + 1)} (1 + F) \in B(H_1 \otimes H_2)$$
Figure 5.1: Alice and Bob communicate via two quantum channels $T$ and $\tilde{T}$. Note that in general, we do not require the dimensions of the associated Hilbert spaces to be equal.

![Diagram of quantum channels](image)

Figure 5.2: The correction scheme (Gregoratti and Werner 2002) applied to the simultaneous usage of two noisy channels $T : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ and $\tilde{T} : \mathcal{B}(\mathcal{K}_1) \to \mathcal{B}(\mathcal{K}_2)$. By the Stinespring representation, these channels consist of unitary couplings $U$ and $\tilde{U}$, respectively, between the local system and environment. The result $\alpha$ of the measurement on the global environment performed by Charlie is classically transmitted to Bob, who chooses the recovery quantum channel $R_\alpha$ accordingly.

![Diagram of correction scheme](image)

are the Werner-Holevo channel and the projector onto symmetric states, respectively. Both are Werner states, that is, commute with $U \otimes U$ for all unitary $U : \mathcal{H}_1 \to \mathcal{H}_2$. The channel $T$ is completely characterized by its flip expectation value

$$\text{tr} [\rho_T \mathbb{F}] = -1 + \frac{2}{d} - \epsilon,$$
which is chosen such that $F$ serves as separation witness, i.e. $T$ is not a convex combination of unitary channels. This follows from the fact that for all unitary $U : H_1 \to H_2$,

$$\text{tr} \left[ (1 \otimes U) |\Omega \rangle \langle \Omega| (1 \otimes U^\dagger) F \right] = \frac{1}{d} \text{tr} [U U^\dagger] \geq -1 + \frac{2}{d}$$

by the same arguments as in the proof of proposition 22.¹

### 5.2 Two copies of the same channel

To exploit $\tilde{T} = T$, Vollbrecht and Werner (2001, section V.B) increment the tensor product symmetry group $(U \otimes U) \otimes (V \otimes V)$ by a flip of $H_i$ and $K_i$,

$$F_{H_i \to K_i} \otimes F_{H_2 \to K_2},$$

resulting in a new symmetry group $G$. Then the commutant $G'$ is precisely spanned by $1$ and

$$F := \frac{1}{2} (1 \otimes F_{H_1 \to K_2} + F_{H_1 \to K_2} \otimes 1), \quad F_{12} := F_{H_1 \to K_2} \otimes F_{K_1 \to K_2},$$

i.e. every state $\rho \in G'$ is completely characterized by the expectation values / coordinates

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

Especially for any unitary channel described by $U : H_1 \otimes K_1 \to H_2 \otimes K_2$, we define $U_s := \frac{1}{2} \left( U + U^T \right)$ and carry out a short calculation resulting in

$$\langle F_{12} \rangle_U = \text{tr} \left[ F_{12} (1 \otimes U) |\Omega \rangle \langle \Omega| (1 \otimes U^\dagger) \right] = \frac{1}{d^2} \text{tr} [U U^\dagger],$$

$$\langle F \rangle_U = \text{tr} \left[ F (1 \otimes U) |\Omega \rangle \langle \Omega| (1 \otimes U^\dagger) \right] = \frac{1}{2d^2} \text{tr} \left[ \frac{1}{d} \text{tr} \left[ U (U_s^{T_1} + U_s^{T_2}) \right] \right] = \frac{1}{d^2} \text{tr} \left[ U_s U_s^{T_2} \right].$$

The last equation takes advantage of the fact that $U_s^{T_2}$ is again symmetric. By the explicit example for $d = 3$ in appendix B.3, the coordinates

$$\langle (F), \langle F_{12} \rangle \rangle_{\vartheta} = \frac{1}{9} \left( -\frac{8}{3} (\cos \vartheta + 1)^2 + 3, 16 \cos^2 \vartheta - 7 \right), \quad \vartheta \in [0, \pi/2] \quad (5.2)$$

correspond to convex combinations of unitary channels.

Now matching the coordinates of $T \otimes T$,

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

with (5.2) yields

$$\epsilon = \frac{2}{3} \left( 4 - 3 \sqrt{2} - \sqrt{3} + \sqrt{6} \right) = 0.31653 \ldots$$

¹This is also a special case of section B.1 in the appendix.
as shown in figure 5.3. The blue area equates to convex combinations of unitary channels\(^2\), the orange curve to the coordinates of single-channel tensor products and the red subcurve to those which are convex combinations on the single-channel level. The state at the lower corner is

\[
\rho_m = \frac{1}{2} \left( \rho_{WH} \otimes \rho_{sym} + \rho_{sym} \otimes \rho_{WH} \right).
\]

As can be seen from direct inspection, each point on the curve \((5.2)\) is an extreme point of the blue area. The remaining extreme points \((1,1)\) and \(\left(\frac{1}{3}, -\frac{2}{3}\right)\) are realized by the unitaries \(I\) and \(\left(-1\right)^{\otimes 3} \otimes 1\), respectively.

Figure 5.3: Twofold tensor products of covariant channels for \(d = 3\); compare with figure 9 in Vollbrecht and Werner (2001) and note that \(-1 + \frac{2}{d^2} = -\frac{7}{9}\) is a hard analytic lower bound on \(\langle F_{12} \rangle\) for unitary channels.

### 5.3 Adding the completely depolarizing channel

The Jamiolkowski representation of the completely depolarizing channel \(\tilde{T} : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2), \rho \rightarrow 1/D\) is just

\[
\rho_{\tilde{T}} \equiv \left( 1 \otimes \tilde{T} \right) (\Omega) (\Omega) = 1/D^2.
\]

\(^2\)A prove that \((5.2)\) really solves the minimization problem \((B.6)\) as acclaimed is still outstanding, but strongly supported by numeric tests. In any case, the set of convex combinations of unitary channels is \textit{at least} as big as the blue area, so the arguments regarding \(T\) remain valid.
Let $H$ be the group of all unitaries $V_1 \otimes V_2$ for unitary $V_i \in \mathcal{B}(\mathcal{K}_i)$, then $H' = \text{span}\{1\}$. By the commutation theorem for von Neumann algebras,

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

(Vollbrecht and Werner 2001, example 7 in section II.D), in other words, $\rho_T \otimes \rho_{\tilde{T}} \in (G \otimes H)'$ is completely characterized by the expectation value of the witness $W = \mathbb{F}_{H_1 \rightarrow H_2} \otimes 1_{\mathcal{K}_1 \otimes \mathcal{K}_2}$, yielding

$$\text{tr} \left[ (\rho_T \otimes \rho_{\tilde{T}}) W \right] = \text{tr} [\rho_T \mathbb{F}] = -1 + \frac{2}{d} - \epsilon. \quad (5.3)$$

Since $W$ is Hermitian and satisfies $W^2 = 1$, the only eigenvalues are $-1$ and 1, so every normalized state $\rho$ satisfies $\text{tr}[\rho W] \in [-1, 1]$.

To obtain the subinterval of $[-1, 1]$ covered by convex combinations of unitary channels, calculate $\langle W \rangle_U$ for unitary $U : \mathcal{H}_1 \otimes \mathcal{K}_1 \rightarrow \mathcal{H}_2 \otimes \mathcal{K}_2$,

$$\text{tr} \left[ (1 \otimes U) |\Omega\rangle \langle \Omega| (1 \otimes U^\dagger) W \right] = \frac{1}{dD} \text{tr} \left[ U U^T W \right]; \quad (5.4)$$

the flowchart in figure 5.4 illustrates this calculation. Note that $U = 1$ reaches

![Flowchart](image)

Figure 5.4: Flowchart corresponding to $\langle \Omega \vert (1 \otimes U^\dagger) (\mathbb{F}_{H_1 \rightarrow H_2} \otimes 1) (1 \otimes U) \vert \Omega \rangle$ for a unitary map $U : \mathcal{H}_1 \otimes \mathcal{K}_1 \rightarrow \mathcal{H}_2 \otimes \mathcal{K}_2$. Note that the flip operator acts on the $H_1 \otimes H_2$ subsystem only.

the upper bound 1, so the hard part is the lower bound, which is treated in appendix B.4. For example in case $D = 2$, we conjecture that (5.4) spans the range

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

For $d = 3$ and $d = 5$, we have constructed explicit examples in the appendix showing that (5.4) covers this range at least. In particular, for $\epsilon$ small enough, the expectation value (5.3) lies within this interval and such proves that $\rho_T \otimes \rho_{\tilde{T}}$ is indeed a convex combination of unitary channels.
Chapter 6

Negativity

6.1 Base norms and unitaries

We introduce – in a similar vein as Vidal and Werner (2001) – a base norm $\| \cdot \|_{S_U}$ and associated negativity $N_{S_U}$ based on mixtures of unitaries $S_U$. We put this in abstract terms first.

Definition 23. Let $X$ be any real linear subspace of the Hermitian operators in $B(H \otimes H)$ and $S$ a convex, compact subset of $X$ such that every element of $S$ has unit trace and the real linear span of $S$ equals $X$. Then $S$ defines a so-called base norm and associated negativity on $X$ by

$$\| A \|_{S_U} := \inf \{ \alpha_p + \alpha_n : A = \alpha_p \sigma_p - \alpha_n \sigma_n, \alpha_{p/n} \geq 0, \sigma_{p/n} \in S \},$$

$$N_{S_U}(A) := \inf \{ \alpha_n : A = \alpha_p \sigma_p - \alpha_n \sigma_n, \alpha_{p/n} \geq 0, \sigma_{p/n} \in S \}$$

for all $A \in X$.

Needless to say $\| \cdot \|_{S_U}$ is a norm on $X$. Detailed proofs can be found in Vidal and Werner (2001) and Nagel (1974). If $tr A = 1$, it follows that $\| A \|_{S_U} = 1 + 2N_{S_U}(A)$ since $\alpha_p - \alpha_n = 1$, and $N_{S_U}(A)$ being zero is equivalent to $A \in S$.

More concretely, let

$$C^{\text{unital}} := \{ \rho_T : \rho_T = (1 \otimes T) (|\Omega\rangle \langle \Omega|), T : B(H) \to B(H) \text{ cp, tp, unital} \}$$

$$= \{ \rho \in B(H \otimes H) : \rho \geq 0, tr_1 \rho = 1/d, tr_2 \rho = 1/d \}$$

be the compact set of unital quantum channels in the Jamiolkowski representation and

$$\mathcal{X} := \text{span}_d C^{\text{unital}} = \{ A \in B(H \otimes H) : A^\dagger = A, \text{ tr}_1 A = \text{ tr}_2 A = 1 \}.$$ 

Then the set $S_U$ of mixtures of unitary channels in the Jamiolkowski representation meets all requirements, which follows from the universality result in section 3.2, corollary 17. In mathematical terms,

$$S_U := \text{conv} \{ (1 \otimes U) |\Omega\rangle \langle \Omega| (1 \otimes U^\dagger) : U \in U(d) \}$$

is a convex, closed subset of $C^{\text{unital}}$.

Considering composition of channels, we obtain the following basic result.
Proposition 24. Let $T_1$ and $T_2$ be unital quantum channels with Jamiolkowski representations $\rho_{T_1}$ and $\rho_{T_2}$, respectively. Then

$$\|\rho_{T_1 \circ T_2}\|_{S_U} \leq \|\rho_{T_1}\|_{S_U} \cdot \|\rho_{T_2}\|_{S_U}.$$ 

Proof. This follows from a straightforward computation applied to the decomposition $\rho_T = \alpha_p \sigma_p - \alpha_n \sigma_n$ of a quantum channel $T$ in definition 23 and the observation that

$$(1 \otimes T) [\langle \Omega | (1 \otimes U^\dagger) (1 \otimes T^*) \rangle]$$

$$= (U^T \otimes 1) [(1 \otimes T) \langle \Omega | \Omega \rangle] (U \otimes 1)$$

$$= (U^T \otimes 1) \rho_T (U \otimes 1).$$

for any unitary $U \in \mathcal{B}(\mathcal{H})$.

As we have shown in section 4.1, the set $S_U$ is invariant under $(U_1 \otimes U_2)$-conjugations for any unitary $U_1, U_2 \in \mathcal{B}(\mathcal{H})$, so

$$\|A\|_{S_U} = \|UAU^\dagger\|_{S_U} \quad \text{for all} \quad U = U_1 \otimes U_2, \quad (6.2)$$

which equivalently holds for the negativity $\mathcal{N}_{S_U}$, too. Let $G$ be a closed subgroup of all unitaries of the form $(U_1 \otimes U_2)$, then averaging over the conjugations drawn from $G$ leads to the twist operation\(^1\) introduced in (4.1),

$$P(A) = \int_G UAU^\dagger \, dU.$$ 

Note that this is a LOCC\(^2\) in physical terms due to the tensor product splitting into $U_1 \otimes U_2$. Equation (6.2) and the triangle inequality yield

$$\|P(A)\|_{S_U} \leq \|A\|_{S_U} \quad \text{for all} \quad A \in \mathcal{X}.$$ 

Moreover, for any $G$-invariant $A$, i.e. element of the commutant $G'$, the states $\sigma_{p/n} \in S_U$ in definition 6.1 can also be taken from $G'$ since

$$A = P(A) = \alpha_p P(\sigma_p) - \alpha_n P(\sigma_n)$$

and $P(\sigma_p), P(\sigma_n) \in S_U$. This is exemplified in figure 6.1, which illustrates $\mathcal{N}_{S_U}$ for orthogonal covariant channels.

6.2 Implementation for the Werner-Holevo channel

The results in section 4.2 encourage a concrete realization of the Werner-Holevo channel $T_{WH}$ as a convex combination of unitary channels in case of even dimension $d = \dim \mathcal{H}$ and as an affine combination with optimal negativity in

\(^1\)The integral is taken with respect to the normalized Haar measure corresponding to the compact group of all unitaries on $\mathcal{B}(\mathcal{H})$. This is extensively discussed by Vollbrecht and Werner (2001) and in chapter 4.

\(^2\)Local Operations and Classical Communication
Figure 6.1: Negativity as distance measure, exemplified through orthogonal covariant channels from section 4.2. $N_{SU}$ is constant along each red line, which can be directly seen by geometric interpretation.

case of $d$ odd. That is, in the first case we explicitly provide a finite number of unitaries $U_i$, $i = 1, \ldots, m$ such that

$$\rho_{WH} = \frac{1}{m} \sum_{i=1}^{m} \left( 1 \otimes U_i \right) |\Omega\rangle \langle \Omega| \left( 1 \otimes U_i^\dagger \right), \quad (d \text{ even})$$

and in the second case we first calculate the negativity $N_{SU}(\rho_{WH}) =: \alpha_n$ and afterwards two sets of unitaries $V_i$, $i = 1, \ldots, m$ and $V'_i$, $i = 1, \ldots, m'$ with associated positive weights $q_i, q'_i$ such that $\sum_i q_i = 1$, $\sum_i q'_i = 1$ and

$$\rho_{WH} = (1 + \alpha_n) \sum_{i=1}^{m} q_i \left( 1 \otimes V_i \right) |\Omega\rangle \langle \Omega| \left( 1 \otimes V_i^\dagger \right)$$

$$-\alpha_n \sum_{i=1}^{m'} q'_i \left( 1 \otimes V'_i \right) |\Omega\rangle \langle \Omega| \left( 1 \otimes V'_i^\dagger \right). \quad (d \text{ odd})$$

The Werner-Holevo channel is special since it was the first counterexample to the additivity conjecture of the maximal $p$-norm multiplicativity conjecture (Werner and Holevo 2002; Hayden 2007). Recall that

$$T_{WH} : \rho \mapsto \frac{1}{d - 1} \left( \text{tr}[\rho] \mathbf{1} - \rho^T \right)$$

31
is a unital quantum channel with $UU$-invariant Jamiolkowski representation

$$\rho_{\text{WH}} = (1 \otimes T_{\text{WH}})(\Omega \langle \Omega \rangle)$$

$$= [d(d-1)]^{-1}(1 - F)$$

$$= 2[d(d-1)]^{-1}A_2,$$

where $A_2 = \frac{1}{2}(1 - F)$ is the 2-particle antisymmetrizer.

We need some preliminary definitions for the setup in the next two subsections. First set

$$\mid \phi_{ij}^{a} \rangle := \mid \phi_{ji}^{a} \rangle := |ij\rangle - |ji\rangle \sqrt{2},$$

$$\mid \phi_{ij}^{s} \rangle := \mid \phi_{ji}^{s} \rangle := |ij\rangle + |ji\rangle \sqrt{2}$$

for all $1 \leq i < j \leq d$ such that

$$\{\mid \phi_{ij}^{a} \rangle_{i<j}, \mid \phi_{ij}^{s} \rangle_{i<j}, |ii\rangle\}$$

is an orthonormal basis of $H \otimes H$. Further define

$$\sigma_{y}^{kl} := -i \left( |k\rangle \langle l| - |l\rangle \langle k| \right),$$

$$\sigma_{z}^{kl} := |k\rangle \langle k| - |l\rangle \langle l|$$

for all $k \neq l$.

For all permutations $\sigma$ of $\{1, \ldots, d\}$,

$$P_{\sigma} := \left\{ \{(\sigma_1, \sigma_2), \ldots, (\sigma_{d-1}, \sigma_d)\}, \text{ } d \text{ even} \right\}
\left\{ \{(\sigma_1, \sigma_2), \ldots, (\sigma_{d-2}, \sigma_{d-1}, \sigma_d)\}, \text{ } d \text{ odd} \right\}$$

where $(\sigma_i, \sigma_j)$ is understood as tuple in $\mathbb{N}^2$, and

$$\mathcal{P} := \{P_{\sigma} : \sigma \text{ permutation of } \{1, \ldots, d\}\}.$$}

Note that by a short thought, the cardinality, i.e. the number of elements of $\mathcal{P}$ equals

$$m_d := \begin{cases} \frac{d!}{(d/2)!} \frac{\Gamma(d+1)}{\Gamma(d/2 + 1)}, & d \text{ even}, \\ \frac{d!}{(d-1/2)!} \frac{\Gamma(d+1)}{\Gamma(d/2 + 3/2)}, & d \text{ odd}. \end{cases}$$

Finally, for all $i, j$, set

$$\text{sign}_{ij} := \begin{cases} 1 & i < j, \\ 0 & i = j, \\ -1 & i > j. \end{cases}$$

**The case $d$ even**

Our goal is to find an explicit convex combination of unitary channels which equals $\rho_{\text{WH}}$.

For all $P \in \mathcal{P}$ define the unitary matrix $U_P := \sum_{(k,l) \in P} \sigma_{y}^{kl}$ and

$$\mid u^P \rangle := (1 \otimes U^P) \mid \Omega \rangle = i \sqrt{2^{d-1}} \sum_{(k,l) \in P} \text{sign}_{kl} \mid \phi_{a}^{kl} \rangle.$$

(6.3)
Set $\rho_y := \frac{1}{|P|} \sum_{P \in \mathcal{P}} |u^P\rangle \langle u^P|$. We want to prove that $\rho_y = \rho_{\text{WH}}$. For all $i < j$,

$$
\rho_y |\varphi_{ij}^{(a)}\rangle = \frac{1}{|P|} \sum_{P \in \mathcal{P}} \frac{2}{d} \left( \sum_{(i,j) \in P} \sum_{(k,l) \in P} \text{sgn}_{kl} |\varphi_{ik}^{(a)}\rangle \langle \varphi_{lj}^{(a)}| \right) = 2 \left[ d(d-1) \right]^{-1} |\varphi_{ij}^{(a)}\rangle,
$$

and directly from the definition (6.3) it follows that $\rho_y |\varphi_{ij}^{(a)}\rangle = 0$ for all $i < j$ and $\rho_y |ii\rangle = 0$ for all $i$, so actually

$$
\rho_y = 2 \left[ d(d-1) \right]^{-1} A_2 = \rho_{\text{WH}}.
$$

**The case $d$ odd**

Recall that by definition (6.1), our task consists of finding convex combinations of unitary channels $\sigma_{p/n}$ such that

$$
\rho_{\text{WH}} = (1 + \alpha_n) \sigma_p - \alpha_n \sigma_n, \quad \alpha_n = N_{\mathcal{S}_n}(\rho_{\text{WH}}) \geq 0. \quad (6.4)
$$

The negativity can be directly obtained from the x-axis in figure 6.1, yielding

$$
N_{\mathcal{S}_n}(\rho_{\text{WH}}) = \frac{1}{d-1}.
$$

For all $P \in \mathcal{P}$, we define unitary matrices

$$
U^P := \sum_{(k,l) \in P} \sigma_{kl}^y + \sum_{j \in P} |j\rangle \langle j|,
$$

$$
V^P := \sum_{(k,l) \in P} \sigma_{kl}^z + \sum_{j \in P} |j\rangle \langle j|,
$$

and set

$$
|u^P\rangle := (1 \otimes U^P) |\Omega\rangle = i\sqrt{2d^{-1}} \sum_{(k,l) \in P} \text{sgn}_{kl} |\varphi_{ik}^{(a)}\rangle + d^{-1/2} \sum_{j \in P} |jj\rangle,
$$

$$
|v^P\rangle := (1 \otimes V^P) |\Omega\rangle = d^{-1/2} \sum_{(k,l) \in P} (|kk\rangle - |ll\rangle) + d^{-1/2} \sum_{j \in P} |jj\rangle.
$$

Define corresponding stochastic unitary channels by

$$
\rho_y := \frac{1}{|P|} \sum_{P \in \mathcal{P}} |u^P\rangle \langle u^P|,
$$

$$
\rho_z := \frac{1}{|P|} \sum_{P \in \mathcal{P}} |v^P\rangle \langle v^P|.
$$
Then a calculation similar to the case \( d \) even yields
\[
\rho_y |\varphi_{ij}^y\rangle = 4 m_{d-2} [d \cdot m_d]^{-1} |\varphi_{ij}^a\rangle = 2 d^{-2} |\varphi_{ij}^a\rangle,
\]
\( \rho_y |\varphi_{ij}^y\rangle = 0 \) for all \( i < j \) and
\[
\rho_y |ii\rangle = \frac{1}{|P|} \sum_{P \in P} d^{-1/2} |u_P\rangle = \frac{1}{|P|} \frac{1}{d} \left( \sum_{P \in P} 1 \right) |ii\rangle
\]
\( = m_{d-1} [d \cdot m_d]^{-1} |ii\rangle = d^{-2} |ii\rangle \) for all \( i \).

Moreover, one obtains \( \rho_z |\varphi_{ij}^y\rangle = 0 \), \( \rho_z |\varphi_{ij}^z\rangle = 0 \) for all \( i < j \), and
\[
\rho_z |ii\rangle = \frac{1}{|P|} \frac{1}{\sqrt{d}} \left[ \sum_{j \neq i} \left( \sum_{P \in P (i,j) \in P} |v_P\rangle - \sum_{P \in P (j,i) \in P} |v_P\rangle \right) + \sum_{P \in P i \in P} 1 \right] |ii\rangle
\]
\( = 2 (d-1) m_{d-2} + m_{d-1} |ii\rangle - 2 m_{d-2} [d \cdot m_d]^{-1} \sum_{j \neq i} |jj\rangle
\)
\( = d^{-1} |ii\rangle - d^{-2} \sum_{j \neq i} |jj\rangle \).

Finally, the state \( \rho_1 := |\Omega\rangle \langle \Omega| \) satisfies
\[
\rho_1 |\varphi_{ij}^y\rangle = 0, \quad \rho_1 |\varphi_{ij}^z\rangle = 0,
\]
\( \rho_1 |ii\rangle = d^{-1/2} |\Omega\rangle = d^{-1} |ii\rangle + d^{-1} \sum_{j \neq i} |jj\rangle \).

Plugging in the above values, one verifies that
\[
\rho_{WH} = (1 + \alpha_n) \rho_y - \alpha_n [d(d + 1)^{-1} \rho_z + (d + 1)^{-1} \rho_1], \quad \alpha_n = \frac{1}{d - 1}
\]

since the coefficients are chosen such that the right hand side maps \( |ii\rangle \) to zero for all \( i \). That is, we have found a representation (6.4) with optimal \( \alpha_n \).
Chapter 7

Conclusions

Our first main line of thought in this thesis culminates in corollary 17 stating that every unital quantum channel $T$ can be represented as an affine combination of unitary conjugations,

$$T(\rho) = \sum_i \lambda_i U_i \rho U_i^\dagger, \quad \lambda_i \in \mathbb{R} \text{ for all } i, \quad \sum_i \lambda_i = 1.$$  

We have combined this equation with recent contractivity results (Pérez-García et al. 2006) in theorem 18 to state a novel characterization of unital channels. Moreover, the linear relation permits us to define a base norm $\| \cdot \|_{S_U}$ and associated negativity $N_{S_U}$ as a distance measure based on the convex hull of unitary channels. We have been able to prove a basic result in proposition 24 regarding base norms and the composition of quantum channels, namely, given unital quantum channels $T_1$ and $T_2$,

$$\|\rho_{T_1 \circ T_2}\|_{S_U} \leq \|\rho_{T_1}\|_{S_U} \cdot \|\rho_{T_2}\|_{S_U}.$$  

Thus we hope that $\| \cdot \|_{S_U}$ is a promising candidate for further investigation.

The second line of thought centers around channels with a high degree of symmetry, namely, orthogonal covariant ones. We have been able to completely resolve the structure of unitary channels within this set in figure 4.1, section 4.2. Interestingly, a substructure emerges for odd dimension $d$ only.

In a third line of thought, we have applied the tools developed so far to construct a covariant channel which exhibits counter-intuitive properties regarding environment-assisted error correction, even for very small dimensions. We reason that this correction scheme is far from being understood and may prove fruitful for future explorations.

Intuitively it may seem that asymptotically defined quantities and the "curse of dimensionality" completely exclude a numeric approach in quantum information theory. Nevertheless, in this thesis we have successfully used numeric methods to obtain the extreme point counterexample in section 3.3 and understand the critical parts of the matrix analysis problems in appendix B. (Which have altogether arisen from the application of separation witnesses.) A key ingredient of our proceeding can be described as follows. Given a continuous
target function \( f \) from \( U(d) \) to the real numbers, we want to obtain an analytic solution of
\[
\min \{ f(U) : U \in U(d) \}.
\]
After calculating a numeric minimizer \( \tilde{U} \), the hard part consists of "analytifying" \( \tilde{U} \), that is, replacing the entries in \( \tilde{U} \) by analytic values. We have achieved this by successively confining the search space under the constraint that it still contains the minimizer of \( f \), until the latter becomes unique.

Concluding, it is reasonable that our analytic low-dimensional results can be generalized to arbitrary dimensions. This may open a door for a concrete understanding of e.g. environment-assisted error correction in the asymptotic limit.
Appendix A

Extreme Point Coefficients

The following algebraic values for the coefficients $\mu_1, \ldots, \mu_4$ of $X$ in (3.8) are appropriate; we have obtained them basically by guessing and suppose that at least polynomial degree three is required.

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

The eigenvalues of $X$ as calculated by Mathematica are then

\[
\lambda(X) = \left\{ 0, 0, \frac{1}{3}, \frac{1 \pm \sqrt{3 \cdot 2} \cdot 0.23604}{6} \right\}
\]
\[
\approx \left\{ 0, 0, \frac{1}{3}, \frac{1 \pm 0.097285}{6} \right\}
\]
with
\[
\alpha = \text{Root}_1 \left[ -25957 + 163107x - 78003x^2 + 6561x^3 \right]
\]
\[
= \frac{107}{27} - \frac{104 \cdot 2^{2/3} (1 + i\sqrt{3})}{27 (67 + 23i\sqrt{7})^{1/3}} - \frac{13 (1 - i\sqrt{3}) (67 + 23i\sqrt{7})^{1/3}}{27 \cdot 2^{2/3}}
\]
\[
\geq 0.17329.
\]
In particular, \( X \) is positive semidefinite with rank equal to 4.
Appendix B

Matrix Optimization Problems

B.1 Trace subject to fixed singular values

We solve the minimization problem posed in section 3.1.2 which handles a special class of separation witnesses. Given any matrix \( A \in B(H) \cong \mathbb{C}^{d \times d} \), denote the singular values of \( A \) by \( \sigma_i(A) \) (counting multiplicity) such that

\[
\sigma_1(A) \geq \cdots \geq \sigma_d(A).
\]

Our main result is the following proposition, which is similar to theorem 7.4.10 in Horn and Johnson (1990).

**Proposition 25.** Let \( A \in B(H) \) and \( \sigma_1 \geq \cdots \geq \sigma_d \) denote the singular values of \( A \). If \( A^*A \) is Hermitian, then there exists a permutation \( \tau \) of \( \{1, \ldots, d\} \), an even \( r \leq d \) and a function \( \rho : \{1, \ldots, r/2\} \rightarrow \{0, 1\} \) with

\[
\text{tr}[AA^*] = 2 \sum_{i=1}^{r/2} (-1)^{\rho(i)} \sigma_{\tau(2i-1)} \sigma_{\tau(2i)} + \sum_{i=r+1}^d \sigma_{\tau(i)}^2. 
\]  

(B.1)

Conversely, given any such \( \tau, r, \rho \) and nonnegative numbers \( \sigma_1 \geq \cdots \geq \sigma_d \), there exists a \( A \in B(H) \) such that \( \sigma_i(A) = \sigma_i \forall i \) and (B.1) holds.

**Proof.** We split the "\( \Rightarrow \)" part into the following steps:

1. \( \text{tr}[AA^*] \) and the singular values of \( A \) are invariant under \( A \mapsto UA^TU^T \) for any unitary \( U \). Note that this map sends \( AA^* \mapsto UAA^*U^T \), so by the spectral theorem, w.l.o.g. \( AA^* \) real diagonal.

2. It follows that \( AA^* = \overline{A^*A} = \overline{AA} \), i.e. \( A \) commutes with \( \overline{A} \), and each of the eigenspaces of \( AA^* \) is invariant under \( A \) and \( \overline{A} \). Stated differently, \( A \) is block diagonal, each block corresponding to an eigenspace of \( AA^* \), so w.l.o.g. \( AA^* = \lambda \mathbf{1} \) for a \( \lambda \in \mathbb{R} \).

3. The case \( \lambda \neq 0 \): Applying the singular value decomposition yields unitary matrices \( U \) and \( V \) such that \( A_1 := U^*AV = \text{diag}(\sigma_1, \ldots, \sigma_d) \). Using
\[
\mathcal{A} = \lambda A^{-1}, \text{ we have }

A_2 := V^\dagger \mathcal{A} U = \lambda (U^\dagger A V)^{-1} = \lambda A_1^{-1} = \lambda \text{diag} (\sigma_1^{-1}, \ldots, \sigma_d^{-1}).
\]

\(A_1\) and \(A_2\) sharing the same singular values translates to \(\{\sigma_1, \ldots, \sigma_d\} = |\lambda| \{\sigma_1^{-1}, \ldots, \sigma_d^{-1}\}\), so there is a permutation \(\tau\) of \(\{1, \ldots, d\}\) with
\[
\begin{align*}
\sigma_{\tau(2i-1)} \sigma_{\tau(2i)} &= |\lambda| \quad \text{for } i = 1, \ldots, \frac{d}{2}, \quad \text{d even} \\
\sigma_{\tau(2i-1)} \sigma_{\tau(2i)} &= |\lambda| \quad \text{for } i = 1, \ldots, \frac{d-1}{2}, \quad \text{d odd.}
\end{align*}
\]

Note that \(d\) cannot be odd if \(\lambda < 0\) as the negative eigenvalues of \(AA\) are of even algebraic multiplicity (Faßbender and Ikramov 2006). Concluding, \(tr[AA] = d \lambda\) can always be written in the form (B.1).

4. The case \(\lambda = 0\): Let \(r\) denote the number of nonzero singular values of \(A\). \(A A = 0\) means that \(\text{range}(A) \subseteq \text{kern}(A)\), so
\[
r = \text{rank}(A) = \text{dim range}(A) \leq \text{dim kern}(A) = d - r,
\]
\(\text{i.e.} \ 2r \leq d\), and there is a permutation \(\tau\) such that each summand in the right hand side of (B.1) is zero.

To prove the "\(\Rightarrow\)" part, set \(A := \text{diag} (A_1, \ldots, A_{r/2}, \sigma_{\tau(r+1)}, \ldots, \sigma_{\tau(d)})\) with
\[
A_i := \begin{pmatrix} 0 & (-1)^{r(i)} \sigma_{\tau(2i)} \\ 0 & 0 \end{pmatrix}.
\]

**Corollary 26.** Given any nonnegative numbers \(\sigma_1 \geq \cdots \geq \sigma_d\),
\[
\min \{ tr [A \overline{A}] : A \in B(H), \sigma_i(A) = \sigma_i \forall i \} = \begin{cases} 
-2 \sum_{i=1}^{d/2} \sigma_{2i-1} \sigma_{2i}, & d \text{ even} \\
-2 \sum_{i=1}^{d-1/2} \sigma_{2i-1} \sigma_{2i} + \sigma_d^2, & d \text{ odd.}
\end{cases}
\] (B.2)

**Proof.** What remains to be shown is \(A \overline{A}\) being Hermitian for all solutions \(A\); then proposition 25 guarantees optimality. All relevant terms are invariant under \(A \mapsto VAV^T\) for unitary \(V\); applying this to the singular value decomposition of \(A\), we may w.l.o.g. assume that \(A = UD\) with \(D = \text{diag} (\sigma_1, \ldots, \sigma_d)\) and \(U\) unitary. Now vary \(U\) to minimize \(tr[AA]\); the unitary constraint translates via
\[
U \, dU^\dagger + dU U^\dagger = d \left( U \, U^\dagger \right) = 0
\]
to \(X := i dU U^\dagger\) being Hermitian. Then
\[
0 = \frac{d}{dU} \text{tr} [D U D U^\dagger] = 2 \text{Re} \text{ tr} [D dU D U^\dagger] = 2 \text{Im} \text{ tr} [U D U D X].
\]
This holds true for any Hermitian matrix \(X\). By decomposing \(U D U D = B_1 + iB_2\), \(B_1\) and \(B_2\) Hermitian, it follows that \(B_2 = 0\).
B.2 Absolute trace under constraints

In this section we calculate the analytic solution of $\max |\text{tr} U|$ for fixed $\text{tr} [U U^\dagger]$ over all unitaries $U \in U(d)$. The motivation for this optimization problem comes from equation (4.4) and proposition 22 in section 4.2, which characterizes the convex hull of unitary channels within the set of orthogonal-covariant channels.

We need the following lemma for proving the main result in proposition 28.

**Lemma 27.** Let $U \in U(2)$ be a unitary $2 \times 2$ matrix and $U_s := \frac{1}{2} (U + U^T)$ the symmetric part of $U$. Then the 1-Schatten-norm\(^1\) of $U_s$ equals

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

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

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

Direct calculation shows that

$$\text{tr} [UU^\dagger] = 2 \left( |\alpha|^2 - \Re (\beta^2) \right) = 4 \left( |\alpha|^2 + (\Im \beta)^2 \right) - 2 \quad (B.3)$$

and

$$U_s = \begin{pmatrix} \alpha & i \Im \beta \\ i \Im \beta & -\alpha \end{pmatrix}, \quad U_s U_s^\dagger = \begin{pmatrix} |\alpha|^2 + (\Im \beta)^2 \\ 1 \end{pmatrix}, \quad (B.4)$$

so

$$\sum_{j=1,2} \sigma_j (U_s) = 2 \sqrt{|\alpha|^2 + (\Im \beta)^2} \|UU^\dagger\|_2 = \sqrt{\text{tr} [UU^\dagger]} + 2.$$ \hfill \square

**Proposition 28.** Let $d \geq 1$ be odd. Then for all $x \in \left[-1 + \frac{2}{d}, 1\right]$, there exists a unitary $U \in U(d)$ such that $d^{-1} \text{tr} [UU^\dagger] = x$, and

$$\max \left\{ d^{-1} |\text{tr} U| : U \in U(d), \, d^{-1} \text{tr} [UU^\dagger] = x \right\}$$

$$= \left[ \frac{1}{2} \left( 1 - \frac{1}{d} \right) \left( 1 - \frac{2}{d} + x \right) \right]^{1/2} + \frac{1}{d} \quad (B.5)$$

**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 $\sigma := \begin{pmatrix} \alpha & \beta \\ -\beta & \bar{\alpha} \end{pmatrix}$, then a short calculation shows that $U := \text{diag} (\sigma, \ldots, \sigma, 1)$ satisfies $d^{-1} \text{tr} [UU^\dagger] = x$ with $d^{-1} |\text{tr} U|$ equal to the right hand side of (B.5). 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 $d^{-1} \text{tr} [UU^\dagger] = x$. By the Youla theorem (Faßbender and Ikramov 2006), given any conjugate-normal matrix $A$ (that is, $AA^\dagger = A^\dagger A$), there exists a unitary map $V$ such that $VAV^\dagger$ is a block

\(^1\)The $p$-Schatten norm ($p \in [1, \infty]$) of $A$ is defined by $\|A\|_p := (\text{tr} |A|^p)^{1/p} = (\sum_i \sigma_i^p)^{1/p}$, where $\sigma_i$ are the singular values of $A$. 

41
diagonal matrix with diagonal blocks of order $1 \times 1$ and $2 \times 2$, the $1 \times 1$ blocks corresponding to the real nonnegative eigenvalues of $A\overline{A}$ and the $2 \times 2$ blocks corresponding either to pairs of equal negative eigenvalues of $A\overline{A}$ or to conjugate pairs of non-real eigenvalues of $A\overline{A}$. Applying this to $U$, there is a unitary $V$ with $U =VDV^T$, the block diagonal matrix $D$ as described. Let $r/2$ be the number of $2 \times 2$ blocks (with even $r \leq d$) and denote these blocks by $D_i$. As $U$ and $V$ are unitary, so must be $D_i$, i.e. $D_i \in U(2)$ for all $i$. Moreover, $U\overline{U}$ unitary guarantees $|\lambda| = 1$ for each eigenvalue $\lambda$ of $U\overline{U}$, so each real nonnegative eigenvalue of $U\overline{U}$ must be 1. Altogether one gets $D = \text{diag} (D_1, \ldots, D_2, 1, \ldots, 1)$. Set $c_i := \frac{1}{2} \text{tr} [D_i D_i^T] \in [-1, 1]$ and $D_2 := \frac{1}{2} (D + D^T)$. Using lemma 27 and the fact that $V^T V$ is unitary and symmetric,

$$|\text{tr} U| = |\text{tr} [D V^T V]| = |\text{tr} [D_2 V^T V]| \leq \sum_{j=1}^d \sigma_j (D_2)$$

$$= \sum_{i=1}^{r/2} \sum_{j=1,2} \sigma_j (D_{i,s}) + (d - r) = 2 \cdot \sum_{i=1}^{r/2} \sqrt{c_i + \frac{1}{2} + d - r}.$$ 

Let $c := d \cdot x \in [-d + 2, d]$. A little thought shows that the problem

$$\max \left\{ 2 \cdot \sum_{i=1}^{r/2} \sqrt{c_i + \frac{1}{2} + d - r} \right\}$$

subject to the side conditions

$$2 \cdot \sum_{i=1}^{r/2} c_i + d - r = c$$

$$c_i \in [-1, 1] \text{ for all } i$$

$$r \text{ even, } r \leq d$$

has optimal solution $r = d - 1$, $c_i = \frac{\sqrt{c} - 1}{2}$ $\forall i$ and the obtained maximum is $d \left[ \frac{1}{2} \left( 1 - \frac{1}{3} \right) \left( 1 - \frac{2}{3} + \frac{\sqrt{c}}{3} \right) \right]^{1/2} + 1$. This upper bound on $|\text{tr} U|$ corresponds exactly to the right hand side of (B.5). \hfill \Box

### B.3 Trace, partial transposition and symmetry

Motivated by equation (5.1) in chapter 5, we investigate

$$\min \left\{ \frac{1}{d_1 d_2} \text{tr} \left[ U_s \overline{U}_s^T \right] : U_s = \frac{1}{2} \left( U + U^T \right), \right.$$ 

$$U \in \mathcal{B} (\mathcal{H} \otimes \mathcal{K}) \text{ unitary with } \frac{1}{d_1 d_2} \text{tr} \left[ U\overline{U} \right] = y \right\} \quad (B.6)$$

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

$$(A \otimes B)^{T_1} = A^T \otimes B \quad \text{and}$$

$$(A \otimes B)^{T_2} = A \otimes B^T, \quad A \otimes B \in \mathcal{B} (\mathcal{H} \otimes \mathcal{K}),$$

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

$$(A \otimes B)^{T_1} = A^T \otimes B \quad \text{and}$$

$$(A \otimes B)^{T_2} = A \otimes B^T, \quad A \otimes B \in \mathcal{B} (\mathcal{H} \otimes \mathcal{K}),$$

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.
respectively. Note that for any $A$ and $B$,

$$\text{tr} \left[ AB \right] = \text{tr} \left[ A^T B^\dagger \right],$$

and for any real or complex $A$ with $A^T = A$, the partial transposes are on equal footing, i.e. $A^T = A^T_1 = A^{T_2}$, so (B.6) is inherently symmetric with respect to $d_1 \leftrightarrow d_2$. In what follows, the operators in $B(\mathcal{H} \otimes \mathcal{K})$ are represented with respect to the ordered computational basis $\left\{ |11\rangle, |12\rangle, \ldots |d_2\rangle, \ldots |d_1 d_2\rangle \right\}$.

Immediately by the definitions, all quantities in (B.6), 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 $W_1 \in B(\mathcal{H})$ and $W_2 \in B(\mathcal{K})$.

Since every unitary matrix is also conjugate-normal (that is, $U U^\dagger = U^\dagger U$), the Youla-theorem$^2$ states that there exists a unitary matrix $V \in B(\mathcal{H} \otimes \mathcal{K})$ such that

$$U = V D V^T$$

with $D$ real block-diagonal and blocks of size $1 \times 1$ and $2 \times 2$, the former non-negative and the latter of the form $\left( \begin{smallmatrix} \sigma & -z \\ z & \sigma \end{smallmatrix} \right)$ with $\sigma \geq 0$. Since $D$ must also be unitary, this equals $\left( \begin{smallmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{smallmatrix} \right)$ for a $\vartheta \in \mathbb{R}$, and all $1 \times 1$ blocks are 1. Note that $U = V D_s V^T$, $D_s := \frac{1}{2} (D + D^T) \geq 0$ diagonal, in particular, $D_s$ contains the singular values of $U_s$. Moreover, $\text{tr} [UU^\dagger] \equiv \text{tr} [DD^\dagger]$ is independent of $V$, so $D$ fixes $y$ in (B.6) and one can freely vary $V$. Conversely, Takagi’s theorem (Horn and Johnson 1990) asserts that every complex-symmetric matrix $A \in \mathbb{C}^{n \times n}$ can be decomposed into

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

(B.7)

with unitary $V$ and $\sigma_i \geq 0$ for all $i$, so identifying $A \equiv U_s$ and $\text{diag} (\sigma_1, \ldots, \sigma_n) \equiv D_s$, the minimization problem (B.6) can be reduced to the following problem and a subsequent optimization over $D_s$.

$$\min \left\{ \frac{1}{d_1 d_2} \text{tr} \left[ A A^T \right] : A \in B(\mathcal{H} \otimes \mathcal{K}), A^T = A, \sigma_i(A) = \sigma_i \forall i \right\}. \quad (B.8)$$

This closely resembles (B.2), and we have effectively decoupled the target function from the peculiar unitary constraint in (B.6).

**Proposition 29.** Every (local) minimizer $A$ of (B.8) satisfies $A A^T \text{Hermitian}$.\footnote{Refer to Youla (1961) and Faßbender and Ikramov (2006, theorem 4). The Youla-form corresponds to the Schur-form w.r.t. unitary congruence transformations $A \rightarrow V A V^T$ and is a generalization of Takagi’s factorization (Horn and Johnson 1990).}

**Proof.** Denote the derivative w.r.t. $V$ in (B.7) by $dV$; since $V$ is unitary,

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

43
so \( X := \frac{1}{d} dV V \) must be Hermitian. Plugging
\[
\frac{d}{dV} A = dV \log(\sigma_i) V^T + V \log(\sigma_i) dV^T = i \left( X A + A X^T \right)
\]
into the target function (B.8) yields
\[
\frac{d}{dV} \text{tr} \left[ A \bar{A}^{T_2} \right] = i \text{tr} \left[ (X A + A X^T) \bar{A}^{T_2} \right] - i \text{tr} \left[ A \left( X A + A X^T \right)^{T_2} \right] = 2 i \text{tr} \left[ X A \bar{A}^{T_2} - A \bar{A}^{T_2} X \right] = 2 i \text{tr} \left[ X \left( A \bar{A}^{T_2} - A \bar{A}^{T_2} \text{h.c.} \right) \right] \div 0.
\]
As this must hold for any Hermitian \( X \), the last equation can only be fulfilled if \( A \bar{A}^{T_2} \) is Hermitian. \( \square \)

It is instructive to rewrite the target function as follows, setting \( \sigma := (\sigma_1, \ldots, \sigma_{d_1,d_2}) \). Denote the columns of \( V \) by \( v_i \), i.e. \( V = (v_1| \ldots | v_{d_1,d_2}) \), then \( A = \sum_i \sigma_i v_i v_i^T \) and
\[
\frac{1}{d_1 d_2} \text{tr} \left[ A \bar{A}^{T_2} \right] = \langle \sigma | G \sigma \rangle, \quad G = (g_{ij}) \text{ Hermitian with}
\]
\[
g_{ij} = \frac{1}{d_1 d_2} \text{tr} \left[ v_j v_j^T v_i v_i^T \bar{A}^{T_2} \right]. \tag{B.9}
\]
Writing
\[
v_i = \sum_k |k\rangle \otimes x_{ik}, \quad x_{ik} \in \mathcal{K},
\]
the last expression becomes
\[
g_{ij} = \frac{1}{d_1 d_2} \sum_{k,k',l,l'} \text{tr} \left[ \left( |l'\rangle \langle l| \otimes x_{jl'} x_{jl}^T \right) \left( |k'\rangle \langle k| \otimes (x_{ik'} x_{ik}^T)^\dagger \right) \right]
\]
\[
= \frac{1}{d_1 d_2} \sum_{k,l} \text{tr}_K \left[ x_{jk} x_{jl}^T x_{ik} x_{ik}^T \right]
\]
\[
= \frac{1}{d_1 d_2} \sum_{k,l} \langle x_{ik} | x_{jl} \rangle \langle x_{il} | x_{jk} \rangle
\]
\[
= \frac{1}{d_1 d_2} \text{tr} \left[ s_{ij}^2 \right], \quad s_{ij} := \langle x_{ik} | x_{jl} \rangle_{k,l=1,\ldots, d_1}.
\]
\( V \) being unitary translates to
\[
\text{tr} s_{ij} = \sum_k \langle x_{ik} | x_{jk} \rangle = \langle v_i | v_j \rangle = \delta_{ij},
\]
In the following paragraph we provide an explicit upper bound\(^3\) of (B.6) for \(d_1 = d_2 = d = 3\). Start with the Ansatz that all \(2 \times 2\) blocks in \(D\) belong to the same phase, i.e.

\[
D = \begin{pmatrix}
\cos \vartheta & -\sin \vartheta \\
\sin \vartheta & \cos \vartheta
\end{pmatrix} \otimes 1, \quad \vartheta \in [0, \pi/2], \quad \text{so}
\]

\[
\frac{1}{d^2} \text{tr} [U^T U] = \frac{1}{d^2} \text{tr} [D^T D] = \frac{1}{9} (16 \cos^2 \vartheta - 7) = y,
\]

and set

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

independent of \(\vartheta\)!. Then \(D_8 = \text{diag} (\sigma_1, \ldots, \sigma_9)\) with \(\sigma_1 = \cdots = \sigma_8 = \cos \vartheta\), \(\sigma_9 = 1\), and \(G\) in (B.9) becomes

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

Finally evaluating the target function provides the supposed minimum

\[
\frac{1}{d^2} \text{tr} \left[ U_s U_s^T \right] = (\sigma | G \sigma) = \frac{1}{9} \left( -\frac{8}{3} (\cos \vartheta + 1)^2 + 3 \right)
\]

with \(\vartheta\) defined by (B.10).

Interestingly, the smallest eigenvalue \(-\frac{1}{27}\) of \(G\) is of algebraic multiplicity 1 with corresponding eigenvector \(\sigma(\vartheta)\) evaluated at \(\vartheta = \pi/3\) and coordinates \(-\frac{1}{9}(1, 1)\). Furthermore, (B.11) is minimal w.r.t. \(\vartheta\) exactly for \(\vartheta = 0\), which corresponds to maximal \(y = \frac{1}{d^2} \text{tr} [U^T U] = 1\) and \(\sigma_1 = \cdots = \sigma_8 = 1\). In this case, \(D\) is the identity and \(U = V V^T\) complex symmetric. Comparing with section B.4, notice that we obtain the same minimum value \(-\frac{23}{27}\).

\(^3\)Numeric tests strongly suggest that this is the actual minimum. Most interestingly, the acclaimed minimizer \(V\) does not depend on \(\vartheta\!\!\!\!\!\!\!\!\!\!
B.4 Partial transposition, trace and distinct dimensions

We explore the following minimization problem posed by equation (5.4) in section 5.3.

\[
\min \left\{ \frac{1}{d_1 d_2} \text{tr} \left[ U U^{T_2} \right] : U \in \mathcal{B}(H \otimes K) \text{ unitary} \right\}, \tag{B.12}
\]

where \(H \otimes K\) is the tensor product of two Hilbert spaces with dimensions \(d_1 = \dim H\) and \(d_2 = \dim K\), respectively, \(d_1\) being odd. The partial transposition is introduced in B.3. Note that any transformation

\[
U \rightarrow \left( V^T \otimes W_1^\dagger \right) U \left( V \otimes W_2 \right) \tag{B.13}
\]

for unitary \(V \in \mathcal{B}(H)\) and unitary \(W_1, W_2 \in \mathcal{B}(K)\) leaves the target function invariant. Restricted to tensor products \(U = U_1 \otimes U_2\), the target function would collapse to \(\frac{1}{d_1} \text{tr} \left[ U_1 U_1^T \right] \geq -1 + \frac{2}{d_1}\), which is in general strictly greater than (B.12), see below. It is worth mentioning that (B.12) is inherently asymmetric w.r.t. \(d_1 \leftrightarrow d_2\), as opposed to the previous section B.3.

**Proposition 30.** \(U U^{T_2}\) is Hermitian for every minimizer \(U\) of (B.12).

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

\[
\frac{d}{dU} \text{tr} \left[ U U^{T_2} \right] = i \text{tr} \left[ X U U^{T_2} \right] - i \text{tr} \left[ U^{T_2} (X U)^\dagger \right]
= i \text{tr} \left[ X \left( U U^{T_2} - h.c. \right) \right] \overset{!}{=} 0.
\]

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

**Disassembly and reformulation** Let \(X = \mathcal{B}(K)\) be the Hilbert space equipped with the Hilbert-Schmidt inner product

\[
\langle A | B \rangle = \text{tr} \left[ A^\dagger B \right]
\]

and induced Frobenius norm\(^4\)

\[
\|A\| = \sqrt{\text{tr} [A^\dagger A]} = \left( \sum_i \sigma_i(A)^2 \right)^{1/2}.
\]

By partitioning \(U\) into \(U = \sum_{i,j=1}^{d_1} |i\rangle \langle j| \otimes u_{ij}\) with \(u_{ij} \in X\) we can now reformulate the target function (B.12), namely

\[
\text{tr} \left[ U U^{T_2} \right] = \sum_{i,j} \langle u_{ij} | u_{ji} \rangle.
\]

\(^4\)This is exactly the \(p\)-Schatten norm \(\|A\|_p = \left( \sum_i \sigma_i(A)^p \right)^{1/p}\) for \(p = 2\).

46
The condition $U$ unitary translates to

$$UU^\dagger = 1 \iff \sum_i u_{ij}^\dagger u_{ik} = \delta_{jk}1 \quad \forall j, k = 1, \ldots, d_1$$

$$\iff \sum_i \langle u_{ij} | u_{ik} x \rangle = \delta_{jk} \text{tr} x \quad \forall j, k; \ x \in \mathcal{X}$$

and the condition in proposition 30 to $U U^T_{2}$ Hermitian

$$\iff \sum_i u_{ki} u_{ij}^\dagger = \sum_i u_{ik} u_{ji}^\dagger$$

$$\iff \sum_i \langle u_{ij} | x u_{ki} \rangle = \sum_i \langle u_{ji} | x u_{ik} \rangle \quad \forall j, k; \ x \in \mathcal{X}.$$ 

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 (B.12) is

$$-1 + \frac{2}{d^2}.$$  \hspace{1cm} (B.14)

Interestingly, there emerges a substructure which is best described by quaternions. Recall that quaternions $H = \{ x_0 + x_1 i + x_2 j + x_3 k : x_0, \ldots, x_3 \in \mathbb{R} \}$ 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 $H$, and the quaternion conjugate of $q = x_0 + x_1 i + x_2 j + x_3 k \in H$ is denoted by $q^*$. Furthermore, define $\text{Re} q := x_0$ and $\vec{q} := q - \text{Re} q = x_1 i + x_2 j + x_3 k$.

To bridge the gap between quaternions and operators on Hilbert spaces, employ the $SU(2)$ identification

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

Note that in this representation

$$q \leftrightarrow \hat{q} = \begin{pmatrix} x_0 + i x_1 & x_2 + i x_3 \\ -x_2 + i x_3 & x_0 - i x_1 \end{pmatrix} \in \mathbb{C}^{2 \times 2} \quad (\text{B.15})$$

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

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

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 an $d \times d$ matrix $A \in \mathbb{H}^{d \times d}$. The identification (B.15) provides an algebra isomorphism

---

5This convention is different from Bunse-Gerstner et al. (1989).
between $\mathbb{H}^{d \times d} \cong \mathbb{R}^{d \times d} \otimes \mathbb{H}$ and the complex $2d \times 2d$ matrices consisting of $2 \times 2$ blocks (B.15); to obtain $Ax$, define $u,v \in \mathbb{C}^d$ by $x = u - jv$ and set $\hat{x} = (u_1, v_1, \ldots, u_d, v_d)^T \in \mathbb{C}^{2d}$; 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

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

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

As in the article by Bunse-Gerstner, Byers, and Mehrmann (1989), call $\lambda \in \mathbb{C}$, $\text{Im} \lambda \geq 0$ an eigenvalue of $A$ with 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 31.** Let $A \in \mathbb{H}^{d \times d}$ be Hermitian with eigenvalues $-1 \pm i d$, 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 $\text{Re} A = 0$, then there exists a unitary $U \in \mathbb{B}(\mathcal{H} \otimes \mathcal{K})$ with

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

**Proof.** $A^* = -A$ since $\text{Re} A = 0$, and consequently $A^T = -A$. Set $U = \frac{1}{\sqrt{2}} (1 + A)$, 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 $\lambda(U) = \frac{1}{\sqrt{2}} (1 + (-1 \pm d)) = \pm 1$. Using

$$ U^T U = \frac{1}{2} (1 + A^*) = \frac{1}{2} (1 - A), $$

it follows that

$$ \frac{1}{2 d^2} \text{tr} \left[ U^T \hat{A}^T U \right] = \frac{1}{2 d^2} \text{tr} \left[ U^T U \right] \frac{1}{2} \left[ 1 - A^2 \right] = \frac{1}{4 d^2} \left( 1 - (d^2 - 1) \right) = -1 + \frac{2}{d^2}. $$

The isomorphism (B.15) introduces an additional factor 2 into the trace, which cancels $\frac{1}{2}$ in the last equation. \hfill $\square$

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 $A^2 + 2A - (d^2 - 1)I = 0$. The requirement $\text{Re} A = 0$ implies the respective eigenvalue multiplicities via $\text{tr} A = 0$.

To show that proposition 31 is not void, 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

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

respectively. Note that for these quaternion models, $U$ is Hermitian, as well as $U \hat{U}^T$, in agreement with proposition 30.

48
Higher dimensions  The following table B.1 contains numeric results for different values of \( d_1 \) and \( d_2 \). We have 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\(^6\) -1 can be obtained for \( d_1 = 5 \) and \( d_2 = 4 \), and that the corresponding minimizer can be chosen real orthogonal.

<table>
<thead>
<tr>
<th>( d_1 )</th>
<th>( d_2 = 1 )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>-\frac{1}{3}</td>
<td>-\frac{7}{9}</td>
<td>-\frac{23}{27}</td>
<td>-\frac{4}{5}</td>
<td>-\frac{47}{35}</td>
</tr>
<tr>
<td>5</td>
<td>-\frac{2}{5}</td>
<td>-\frac{23}{25}</td>
<td>-0.929151 \ldots</td>
<td>-1</td>
<td>-0.976326 \ldots</td>
</tr>
<tr>
<td>7</td>
<td>-\frac{5}{7}</td>
<td>-\frac{47}{35}</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

Table B.1: Numeric solutions of (B.12); the columns correspond to different values of \( d_2 \). The case \( d_2 = 1 \) can be treated analytically by an eigenvalue argument and is a special case of B.1. Note that \( d_2 = 2 \) is in agreement with (B.14), and for \( d_1 = d_2 = 3 \) we obtain the same value as the minimum of (B.11) with respect to \( \theta \). In agreement with the nature of problem, the table is asymmetric under \( d_1 \leftrightarrow d_2 \).

\(^6\)The fact that the range of the target function (B.12) is contained in the interval \([-1,1]\) follows from the original equation (5.1) and the spectrum \( \lambda(F) = \{-1,1\} \).
Bibliography


Acknowledgments

First and foremost, I’d like to thank Ignacio Cirac and Michael Wolf for giving me the opportunity of working in the group and constantly supervising my work. Michael has explained to me almost everything I’ve learned about quantum information theory in the last year.

I had a very good time with my fellow students Heike Schwager, Philipp Hauke, Borys Hagemann, Alexander Ewald and Martin Fürst and I’m very grateful for reviewing this thesis.

Without going into details I just want to say thank you to the whole theory group, my parents and my brothers Joachim and Johannes.