Disguising quantum channels by mixing and channel distance trade-off

Chi-Hang Fred Fung and H. F. Chau

Department of Physics and Center of Theoretical and Computational Physics,
University of Hong Kong, Pokfulam Road, Hong Kong

E-mail: chffung@hku.hk

Abstract. We consider the reverse problem to the distinguishability of two quantum channels, which we call the disguising problem. Given two quantum channels, the goal here is to make the two channels identical by mixing with some other channels with minimal mixing probabilities. This quantifies how much one channel can disguise as the other. In addition, the possibility to trade off between the two mixing probabilities allows one channel to be more preserved (less mixed) at the expense of the other. We derive lower- and upper-bounds of the trade-off curve and apply them to a few example channels. Optimal trade-off is obtained in one example. We relate the disguising problem and the distinguishability problem by showing the the former can lower and upper bound the diamond norm. We also show that the disguising problem gives an upper bound on the key generation rate in quantum cryptography.

PACS numbers: 03.67.Hk, 03.67.Lx, 03.67.Dd

1. Introduction

Quantum information processing involves the transformation of quantum states through quantum channels and it is often useful to quantify how far apart quantum states or quantum channels are. Depending on the problem at hand, different ways of measuring the distance may be adopted. Trace distance [1, 2] and fidelity [3, 4, 5] are two widely-used measures for quantum states. Trace distance is particularly interesting because it corresponds to a measurement that distinguishes between two quantum states with the minimum error. Other distances for quantum states have also been studied recently, including the Monge distance [6], the kth operator norm [7], and the partitioned trace distance [8]. For quantum channels, measures [9] have also been proposed based on extending the fidelity measure [10] and the trace distance measure [11] of quantum states. The diamond norm, in particular, is a trace-distance-based measure for quantum channels. It was first introduced in quantum information processing by Kitaev [11] for studying quantum error correction and has a nice operational meaning because it corresponds to minimum-error channel discrimination. As such, the diamond norm has been receiving a lot of attention since its introduction, in both the theoretical aspect [12, 13, 14, 15, 16, 17] and the computational aspect [18, 19].

While distinguishability (of quantum states and channels) is a well studied problem, we consider the reverse problem – the disguising problem for quantum channels. Unlike the distinguishability problem in which the goal is to find a measurement that distinguishes between two (or more) states or channels, the aim in the disguising problem is to find out the minimal mixing needed to make two (or more) quantum channels completely identical. In essence, this quantifies how much the effect of one channel is partially carried out by another channel. In this pilot study, we investigate the disguising problem for two channels.

As we show in this paper, the disguising problem and the distinguishability problem can be considered as dual to each other. We establish this by showing that the solution of the disguising problem can be used to lower and upper bound the diamond norm, which is a measure of distinguishability. This has an interesting implication: the more distinguishable two quantum channels are, the more effort it takes to disguise one as the other. Additionally, the disguising problem can be cast as a semidefinite program and we also show efficient ways to compute lower- and upper-bounds of it. We note that the diamond norm can be computed using semidefinite/convex programming and Monte Carlo methods [20, 18, 21, 19].

The disguising problem can be understood with the following operational interpretation. First note that the operational meaning of the diamond norm is based on the perspective of the receiver who tries to distinguish between two channels. A reverse perspective is to look at the channel intervener who tries to make the channels identical by minimal intervention. The channel intervener possesses the two original channels as black boxes. She is not allowed to open them and is only allowed to occasionally substitute each of them with some other arbitrary channel. We ask what are the minimal

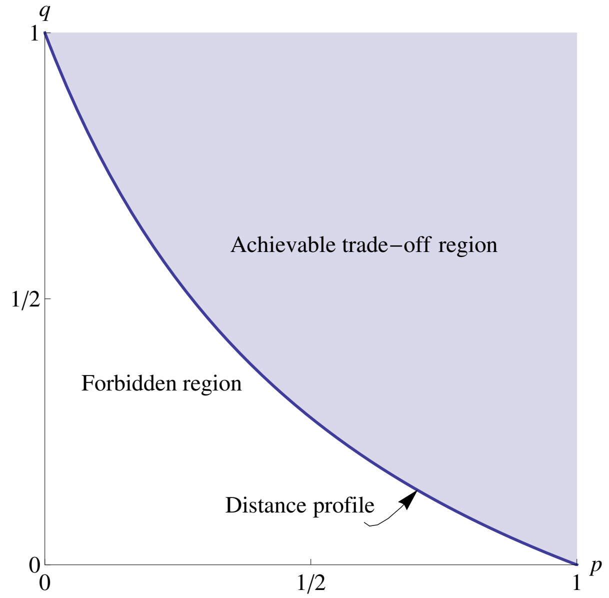


Figure 1. Achievable trade-off region and distance profile of the channel disguising problem. When the mixing probabilities p and q are too small (corresponding to the forbidden region), the channels $\mathcal{E}$ and $\mathcal{F}$ cannot be made equal.

mixing probabilities needed to make the two intervened channels identical?

The precise problem statement is the following. Given two quantum channels $\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger$ and $\mathcal{F}(\rho) = \sum_i F_i \rho F_i^\dagger$ acting on an n-dimensional Hilbert space where E_i and F_i are $n \times n$ complex matrices representing the Kraus operators of the channels, we consider the processing

$$\mathcal{E}(\rho) \to \mathcal{E}'(\rho) \ = (1-p)\mathcal{E}(\rho) + p\mathcal{E}_\Delta(\rho), \tag{1}$$

$$\mathcal{F}(\rho) \to \mathcal{F}'(\rho) = (1-q)\mathcal{F}(\rho) + q\mathcal{F}_\Delta(\rho) \tag{2}$$

such that $\mathcal{E}' = \mathcal{F}'$ with "the smallest" p and q (which will be clarified later as the distance profile). In other words, we are interested in the least amount of substitution of $\mathcal{E}$ and $\mathcal{F}$ needed to make them equal. This is illustrated in figure 2. Operationally, the new channel probabilistically selects between the original channel and some other *harmonizing channel*, $\mathcal{E}_\Delta$ or $\mathcal{F}_\Delta$, which is yet to be determined. The smaller the *mixing probability*, p or q, the closer the new channel is to the original one. Thus, p and q serve as a distance between the two channels. We note that in general the harmonizing channels $\mathcal{E}_\Delta$ and $\mathcal{F}_\Delta$ are not universal and depend on the original channels $\mathcal{E}$ and $\mathcal{F}$. In fact, the problem becomes trivial if we insist $\mathcal{E}_\Delta$ and $\mathcal{F}_\Delta$ to be universal for then $\mathcal{E}_\Delta = \mathcal{F}_\Delta$ (with either $p = q = 1$ when $\mathcal{E} \neq \mathcal{F}$ or $p = q = 0$ when $\mathcal{E} = \mathcal{F}$). Also, note that $\mathcal{E}' = \mathcal{F}'$ is trivially satisfied with $\mathcal{E}_\Delta = \mathcal{F}$, $\mathcal{F}_\Delta = \mathcal{E}$, and $p + q = 1$. This gives a linear trade-off between p and q. However, in general, better sub-linear trade-off can be obtained, as we show later. Note that $\mathcal{E}_\Delta$ and $\mathcal{F}_\Delta$ can be general quantum channels with arbitrary complexity.

The pair of parameters (p, q) represents the trade-off between the two channels'

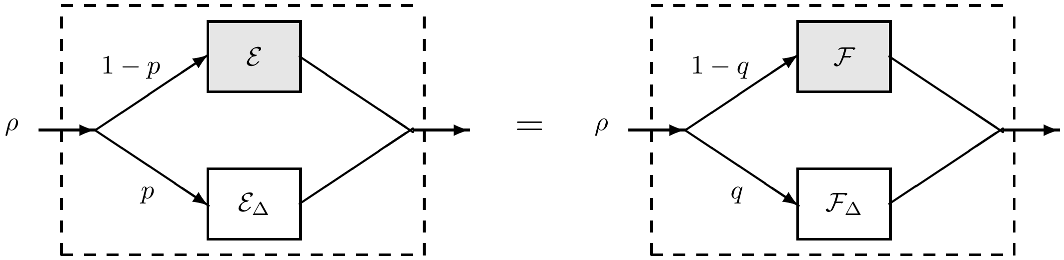


Figure 2. Two quantum channels, $\mathcal{E}$ and $\mathcal{F}$, are made identical by mixing the original channel $\mathcal{E}$ ($\mathcal{F}$) with a harmonizing channel $\mathcal{E}_\Delta$ ($\mathcal{F}_\Delta$) with probability p (q). The designer is free to choose the non-shaded parts $\mathcal{E}_\Delta$ and $\mathcal{F}_\Delta$, but not $\mathcal{E}$ and $\mathcal{F}$.

mixing probabilities. The more mixing is imposed on one channel, the less mixing is required on the other. Note that if (p, q) is a trade-off point, $(p + \delta p, q + \delta q)$ with $\delta p, \delta q \geq 0$ is a also a trade-off point (and we call the former point strictly better than the latter). When given a region of achievable trade-off points, a trade-off curve can be obtained by tracing out the boundary such that no point is strictly better than another. This gives us a distance profile for the two channels (see figure 1). Thus, our measure is unique in that it is represented by a 2-dimensional curve rather than a scalar as in other measures for quantum channels. On the other hand, a scalar distance may be obtained from our measure in several ways, for example, (i) by imposing equal mixing probabilities $p = q$ and regarding the minimum $p = q$ as the distance between the two channels, or (ii) by regarding the minimum $p + q$ as the distance. We will justify that these two are distances by showing that the triangle inequality holds.

The disguising problem admits a geometric interpretation. Given a channel $\mathcal{E}$, we denote the set of all channels achieved by mixing channel $\mathcal{E}$ with arbitrary harmonizing channels and mixing probability p as

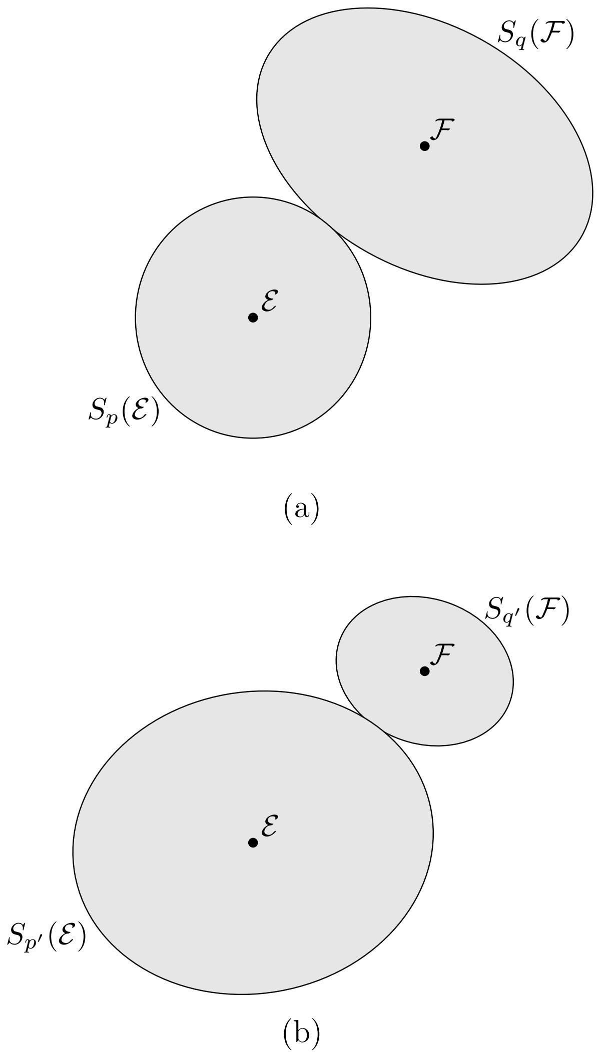
$$S_p(\mathcal{E}) \equiv \{\mathcal{E}' : \mathcal{E}' = (1-p)\mathcal{E} + p\mathcal{E}_\Delta, \text{ for some harmonizing channel } \mathcal{E}_\Delta\}.$$

Note that $S_{p'} \subset S_p$ for $p' < p$ for the following reason. For any $\mathcal{E}' \in S_{p'}$, we have

$$\begin{aligned}\mathcal{E}' &= (1-p')\mathcal{E} + p'\mathcal{E}_\Delta \\ &= (1-p)\mathcal{E} + [p'\mathcal{E}_\Delta + (p-p')\mathcal{E}]\end{aligned}$$

where the term in bracket is a valid quantum channel scaled by p. Thus, $\mathcal{E}'$ can be regarded as having mixing probability p and so $\mathcal{E}' \in S_p$. It can be easily checked that $S_p(\mathcal{E})$ is compact and convex. This enables a geometric interpretation of the disguising problem as a search for p and q so that $S_p(\mathcal{E})$ and $S_q(\mathcal{F})$ just meet (see figure 3).

In this paper, we formulate the disguising problem as an optimization problem and present our main result in section 3. Although solving it turns out to be difficult, we are able to obtain lower-bound and upper-bound on the (p, q) trade-off curve (cf. equation (11)). In section 4, we prove the main result, which is the lower- and upper-bounds. Next, we illustrate the computation of the bounds in a few examples for different quantum channels in section 5. In one special case, the analytical lower- and



(a)

(b)

Figure 3. Geometric interpretation of the disguising problem. The set of channels within mixing probability p (q) from channel $\mathcal{E}$ ($\mathcal{F}$) is denoted as $S_p(\mathcal{E})$ ($S_q(\mathcal{F})$). In (a), we search for p and q so that the two sets meet. In (b), they meet with different parameters p' and q'.

upper-bounds coincide, effectively producing the optimal trade-off curve. For the other cases, the numerically computed lower- and upper-bounds are quite tight, showing the effectiveness of the bounds. In section 6, we show that the disguising problem can lower and upper bound the diamond norm which quantifies the distinguishability of two quantum channels. We also make three remarks about our distance profile for quantum channels in section 7. We discuss one application of the disguising problem in section 8, which is to bound the key generation rate in quantum cryptography. We conclude in section 9.

2. Notation

We denote a matrix A to be positive-semidefinite (PSD) by $A \succeq 0$, the transpose of A by A^t, and the conjugate transpose of A by $A^\dagger$. A is PSD if and only if A is Hermitian and its eigenvalues are non-negative. $\|A\|$ denotes the spectral norm of A which is the largest singular value of A or the largest eigenvalue of A if A is PSD. $\|A\|_1 \triangleq \mathrm{Tr}\sqrt{A^\dagger A}$ denotes the trace norm of A.

$\mathcal{B}(\mathcal{H}_n)$ denotes the set of all bounded linear operators in an n-dimensional Hilbert space $\mathcal{H}_n$. I_p denotes the identity operator in a p-dimensional Hilbert space. A linear map $\mathcal{E} : \mathcal{B}(\mathcal{H}_n) \to \mathcal{B}(\mathcal{H}_n)$ is positive if $\mathcal{E}(A)$ is PSD for all PSD A in $\mathcal{B}(\mathcal{H}_n)$, and $\mathcal{E}$ is completely positive (CP) if $\mathcal{E} \otimes I_p$ is positive for all positive integers p. $\mathcal{E}$ is trace-preserving (TP) if $\mathrm{Tr}(\mathcal{E}(A)) = \mathrm{Tr}(A)$ for all A in $\mathcal{B}(\mathcal{H}_n)$.

A linear map $\mathcal{E} : \mathcal{B}(\mathcal{H}_n) \to \mathcal{B}(\mathcal{H}_n)$ can be represented by a Choi matrix of size $n^2 \times n^2$ [22]:

$$C_{\mathcal{E}} = \begin{bmatrix} \mathcal{E}(|0\rangle\langle 0|) & \mathcal{E}(|0\rangle\langle 1|) & \dots & \mathcal{E}(|0\rangle\langle n-1|) \\ \mathcal{E}(|1\rangle\langle 0|) & \ddots & & \\ \vdots & & & \\ \mathcal{E}(|n-1\rangle\langle 0|) & \dots & & \mathcal{E}(|n-1\rangle\langle n-1|) \end{bmatrix} \tag{3}$$

$$= \sum_{i,j=0}^{n-1} |i\rangle\langle j| \otimes \mathcal{E}(|i\rangle\langle j|). \tag{4}$$

We define a function, which we call the *channel sum* function, of the Choi matrix $C_{\mathcal{E}}$ of a linear map $\mathcal{E}$ as follows:

$$\mathbb{T}(C_{\mathcal{E}}) := \mathrm{Tr}_2^t(C_{\mathcal{E}}) \tag{5}$$

$$= \sum_{i,j=0}^{n-1} |i\rangle\langle j| \cdot \mathrm{Tr}[\mathcal{E}(|j\rangle\langle i|)], \tag{6}$$

where Tr_2 is the partial trace over the second system and t represents transpose. We remark that $\mathbb{T}(C_{\mathcal{E}}) = I$ if and only if $\mathcal{E}$ is trace-preserving (see Lemma 5 in Appendix A).

3. Problem formulation and main result

3.1. Optimization problem formulation

To solve for the optimal distance profile (p, q) of equations (1) and (2) with the condition that the new channels are identical, i.e., $\mathcal{E}' = \mathcal{F}'$, we formulate the problem as (see figure 1):

$$\begin{aligned} &\text{minimize} \quad q \\ &\text{subject to } \mathcal{E}' = \mathcal{F}', \\ &\qquad\qquad C_{\mathcal{E}_\Delta} \succeq 0, \end{aligned} \tag{7}$$

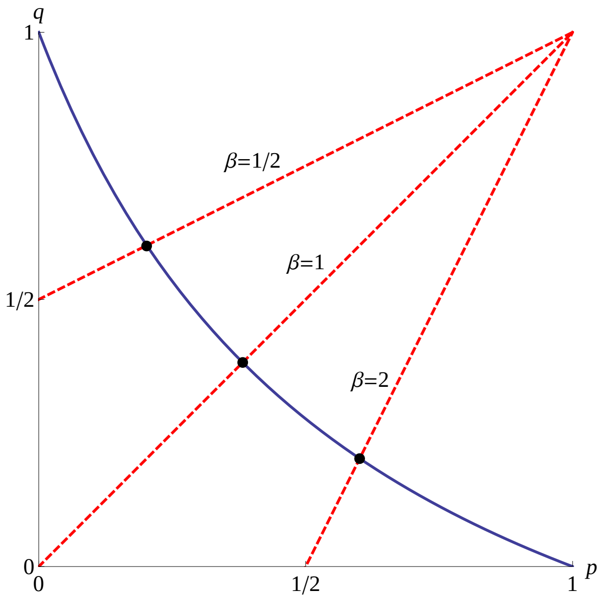


Figure 4. Distance profile obtained by solving problem (8). Among all the points (p, q) along the line corresponding to a fixed value of β, we choose the point with the minimum q subject to the constraints of the problem. For each value of β, we solve the optimization problem and obtain an optimal point (p, q). Repeating this for a range of β produces the solid curve.

$$
\begin{aligned}
&C_{\mathcal{F}_\Delta} \succeq 0, \\
&\mathbb{T}(C_{\mathcal{E}_\Delta}) = I, \\
&\mathbb{T}(C_{\mathcal{F}_\Delta}) = I,
\end{aligned}
$$

where the minimization is over q, $\mathcal{E}_\Delta$, and $\mathcal{F}_\Delta$, for some fixed p. Here, we denote the Choi matrices of $\mathcal{E}_\Delta$ and $\mathcal{F}_\Delta$ by $C_{\mathcal{E}_\Delta}$ and $C_{\mathcal{F}_\Delta}$, respectively (see equation (3)), and $\mathbb{T}$ is the channel sum function defined in equation (5). The last four constraints demand that $\mathcal{E}_\Delta$ and $\mathcal{F}_\Delta$ be quantum channels (TPCP maps) (cf. Theorem 3 and Lemma 5 in Appendix A). Note that the roles of p and q in the formulation of the above optimization problem may be interchanged (i.e., we may have "fix q and minimize p" instead). We remark that this problem can be cast as a semidefinite program with the use of equation (13) and may be solved numerically. However, in this paper, we are interested in the analytical bounds of this problem and the investigation of the trade-off behavior between p and q.

Note that for $p = 1$, the solution of $q = 0$ is trivially obtained to make $\mathcal{E}' = \mathcal{F}'$, since we can choose $\mathcal{E}_\Delta = \mathcal{F}$ in equations (1) and (2). By the same token, $(p, q) = (0, 1)$ is feasible.

The distance profile (such as that in figure 1) should be convex. This is because given two points (p, q) and (p', q') that satisfy $\mathcal{E}' = \mathcal{F}'$ [see equations (1) and (2)], any linear combination of them [i.e., $((1-t)p + tp', (1-t)q + tq')$ for some $0 \le t \le 1$] also satisfies it. It follows that any point on the line $q = 1 - p$ is a feasible solution.

3.2. Main result

Our main result is the lower- and upper-bounds of the distance profile generated by the solutions of problem (7). It turns out that it is easier to analyze and present the main result by expressing the problem as follows:

$$\begin{aligned} &\text{minimize} \quad q \\ &\text{subject to } \mathcal{E}' = \mathcal{F}', \\ &\qquad\qquad \beta = \frac{1-q}{1-p}, \\ &\qquad\qquad C_{\mathcal{E}_\Delta} \succeq 0, \\ &\qquad\qquad C_{\mathcal{F}_\Delta} \succeq 0, \\ &\qquad\qquad \mathbb{T}(C_{\mathcal{E}_\Delta}) = I, \\ &\qquad\qquad \mathbb{T}(C_{\mathcal{F}_\Delta}) = I, \end{aligned} \tag{8}$$

where the minimization is over p, q, $\mathcal{E}_\Delta$, and $\mathcal{F}_\Delta$, for some fixed parameter β. Here, β is a new parameter and the distance profile (p, q) is obtained by solving this optimization problem over a range of β (see figure 4).

Our main result is that the distance profile (p, q) is bounded from below and above as follows:

$$p = 1 - \frac{1}{\alpha + \beta}, \text{ and} \tag{9}$$

$$q = \frac{\alpha}{\alpha + \beta} \tag{10}$$

where β is fixed and

$$n^{-1}\mathrm{Tr}[\mathbb{T}(\Delta_+)] \leq \alpha \leq \min(\|\mathbb{T}(\Delta_+)\|, 1). \tag{11}$$

Here, Δ_+ is a PSD matrix obtained by decomposing $C_\mathcal{E} - \beta C_\mathcal{F}$ into the positive and negative subspaces by eigen-decomposition:

$$C_\mathcal{E} - \beta C_\mathcal{F} = \Delta_+ - \Delta_-, \tag{12}$$

where $\Delta_\pm$ are PSD matrices with support on orthogonal vector spaces (i.e., $\Delta_+\Delta_- = 0$) and $C_\mathcal{E}$ and $C_\mathcal{F}$ are the Choi matrices for the quantum channels $\mathcal{E}$ and $\mathcal{F}$, respectively. Also, n is the dimension of the quantum states on which the channels act.

Note that $dp/d\alpha > 0$ and $dq/d\alpha > 0$ implying that a smaller α gives rise to a "smaller" pair (p, q) in the 2-dimensional space. This means that the lower (upper) bounds of α in equation (11) obtained by varying β correspond to a lower (upper) bound curve in the (p, q) space.

Note that if the lower bound and upper bound of α coincide, the optimal α and thus optimal (p, q) are obtained. This happens if and only if Δ_+ already corresponds to a scaled quantum channel, i.e., $\mathbb{T}(\Delta_+) = \alpha I$, which is not the case in general. Also, note that equation (11) implies that the lower bound of α is always less than or equal to 1, and thus if $\mathbb{T}(\Delta_+) = \alpha I$, $\alpha \leq 1$.

4. Proof of lower- and upper-bounds

As noted earlier, the case of $p = 0$ is trivial and thus we focus on the case $p < 1$ in the following. We now analyze problem (8), and as we will show later, directly solving this problem turns out to be difficult. Let us first focus on the condition that we want: $\mathcal{E}' = \mathcal{F}'$. By Theorem 4 in Appendix A, we convert this condition to the Choi-matrix equivalence $C_{\mathcal{E}'} = C_{\mathcal{F}'}$, which implies that

$$(1-p)C_{\mathcal{E}} + pC_{\mathcal{E}_\Delta} = (1-q)C_{\mathcal{F}} + qC_{\mathcal{F}_\Delta} \tag{13}$$

$$C_{\mathcal{E}} - \beta C_{\mathcal{F}} = \frac{q}{1-p}C_{\mathcal{F}_\Delta} - \frac{p}{1-p}C_{\mathcal{E}_\Delta} \tag{14}$$

where $\beta = \frac{1-q}{1-p}$. We decompose the left-hand side into the positive and negative subspaces by eigen-decomposition:

$$C_{\mathcal{E}} - \beta C_{\mathcal{F}} = \Delta_+ - \Delta_-, \tag{15}$$

where $\Delta_\pm$ are positive semidefinite matrices with support on orthogonal vector spaces (i.e., $\Delta_+\Delta_- = 0$). As such, by Theorem 3 in Appendix A, $\Delta_\pm$ correspond to some CP maps.

Note that $G_{\mathcal{E}}$, $G_{\mathcal{F}}$, $G_{\mathcal{F}_\Delta}$, $G_{\mathcal{E}_\Delta}$, Δ_+, and Δ_- are all Choi matrices.

Comparing equations (14) and (15), since the positive and negative parts on the right-hand sides must match, the Choi matrices of the harmonizing channels must be of the form

$$\frac{q}{1-p}C_{\mathcal{F}_\Delta} = \Delta_+ + X, \tag{16}$$

$$\frac{p}{1-p}C_{\mathcal{E}_\Delta} = \Delta_- + X \tag{17}$$

where X is some Hermitian matrix corresponding to the Choi matrix of some linear map. Note that $\Delta_\pm$ may not correspond to scaled quantum channels because $\mathbb{T}(\Delta_\pm) \neq \alpha I$ for any $\alpha > 0$. The purpose of adding X is to make them scaled quantum channels so that $\mathbb{T}(\Delta_\pm + X) = \alpha_\pm I$.

Lemma 1. $\mathbb{T}(\Delta_+) = \mathbb{T}(\Delta_-) + (1-\beta)I$.

Proof. Rearranging equation (15) and applying Corollary 2, we have

$$\begin{aligned} &\mathbb{T}(C_{\mathcal{E}} + \Delta_-) = \mathbb{T}(\beta C_{\mathcal{F}} + \Delta_+) \\ \Longrightarrow\ &\mathbb{T}(C_{\mathcal{E}}) + \mathbb{T}(\Delta_-) = \mathbb{T}(\beta C_{\mathcal{F}}) + \mathbb{T}(\Delta_+) \\ \Longrightarrow\ &(1-\beta)I + \mathbb{T}(\Delta_-) = \mathbb{T}(\Delta_+) \end{aligned}$$

where we have used the fact that $\mathbb{T}(C_{\mathcal{F}_\Delta}) = \mathbb{T}(C_{\mathcal{E}_\Delta}) = I$ since $C_{\mathcal{F}_\Delta}$ and $C_{\mathcal{E}_\Delta}$ are Choi matrices of quantum channels. □

As a consequence, $\mathbb{T}(\Delta_+ + X) = \alpha I$ if and only if $\mathbb{T}(\Delta_- + X) = (\alpha + \beta - 1)I$. Furthermore, from equation (16), since $\mathbb{T}(C_{\mathcal{F}_\Delta}) = I$, we have

$$\frac{q}{1-p} = \alpha. \tag{18}$$

The same expression is obtained when we consider equation (17) with $\mathbb{T}(C_{\mathcal{E}_\Delta}) = I$. Thus, minimizing q given β fixed is equivalent to minimizing α given β fixed, since

$$\alpha = \frac{q}{1-q}\beta \tag{19}$$

is an increasing function of q. The original problem (8) becomes

$$\begin{aligned}\hat{\alpha} = \text{minimize} \quad & \alpha \\ \text{subject to} \quad & \Delta_+ + X \succeq 0, \\ & \Delta_- + X \succeq 0, \\ & \mathbb{T}(\Delta_+ + X) = \alpha I, \\ & \mathbb{T}(\Delta_- + X) = (\alpha + \beta - 1)I,\end{aligned} \tag{20}$$

where the minimization is over Hermitian matrix X given β fixed, and $\Delta_\pm$ are from equation (15). Note that the fourth constraint is redundant due to Lemma 1 and is shown only for completeness. Once α is found, we can compute p and q from equations (18) and (19).

We investigate the form of X. Since $C_{\mathcal{F}_\Delta}$ and $C_{\mathcal{E}_\Delta}$ represent quantum channels, they are PSD. This means that, according to equations (16) and (17), $\Delta_\pm + X$ are PSD. However, this does not mean that X is also PSD, and this makes finding the optimal X difficult. Nevertheless, we have the following constraint on X which helps us bound $\hat{\alpha}$.

Lemma 2. The constraints of problem (20) implies $\mathrm{Tr}(X) \geq 0$.

Proof. Since Δ_+ and Δ_- are the positive and negative ranges of the matrix in equation (15), we can identify non-overlapping projectors P_+ and P_- onto them respectively. We also define the projector onto the remaining subspace $P_0 = I - P_+ - P_-$. Since $\Delta_\pm + X$ is PSD, we have

$$\begin{aligned}&\mathrm{Tr}[P_-(\Delta_+ + X)] \geq 0, \\ &\mathrm{Tr}[P_+(\Delta_- + X)] \geq 0, \text{ and} \\ &\mathrm{Tr}[P_0(\Delta_\pm + X)] \geq 0,\end{aligned}$$

which implies that

$$\begin{aligned}&\mathrm{Tr}(P_- X) \geq 0, \\ &\mathrm{Tr}(P_+ X) \geq 0, \text{ and} \\ &\mathrm{Tr}(P_0 X) \geq 0.\end{aligned}$$

Summing these terms gives the desired result. □

This lemma implies that the non-zero eigenvalues of X cannot be all negative, but X can have positive and negative eigenvalues.

Theorem 1. The optimal value of problem (20) is upper bounded by $\|\mathbb{T}(\Delta_+)\|$.

Proof. To show an upper bound, we only need to find a feasible X. Choose $M = \|\mathbb{T}(\Delta_+)\|I - \mathbb{T}(\Delta_+)$ as the difference between two channel sums. Certainly, M is PSD and thus can be written as $M = D_0^\dagger D_0$ where D_0 is a square matrix. M represents the channel sum of the channel $\rho \to D_0 \rho D_0^\dagger$. Let $X = |D_0\rangle\langle D_0|$ be the Choi representation of this channel where $|D_0\rangle$ is the vector form of D_0 (cf. equation (A.3)). Since X is PSD, $\Delta_\pm + X$ is PSD and the first two constraints of problem (20) are satisfied. Note that $\mathbb{T}(X) = M$ by construction (cf. Lemma 4 and Definition 1). Therefore, $\mathbb{T}(\Delta_+ + X) = \mathbb{T}(\Delta_+) + \mathbb{T}(X) = \|\mathbb{T}(\Delta_+)\|I$ by Corollary 2.

Note that for this upper bound, we have chosen X to be PSD. □

We computed $\|\mathbb{T}(\Delta_+)\|$ for random quantum channels and found cases with $\|\mathbb{T}(\Delta_+)\| > 1$. Nevertheless, $\hat{\alpha} \le 1$ is also a valid bound.

Lemma 3. The optimal value of problem (20) is upper bounded by unity, i.e., $\hat{\alpha} \le 1$.

Proof. Set the harmonizing channels in equations (1)–(2) to be $\mathcal{E}_\Delta = \mathcal{F}$ and $\mathcal{F}_\Delta = \mathcal{E}$. Then, $\mathcal{E}' = \mathcal{F}'$ is satisfied with $q = 1 - p$, which means that $\alpha = 1$ according to equation (18). Based on equation (15), we set $X = C_\mathcal{E} - \Delta_+ = \beta C_\mathcal{F} - \Delta_-$. Then, we have $\Delta_+ + X = C_\mathcal{E} \succeq 0$ and $\Delta_- + X = \beta C_\mathcal{F} \succeq 0$. As such, the constraints of problem (20) are satisfied with $\alpha = 1$. □

We remark that in the above proofs of the two upper bounds, we have explicitly constructed X. Therefore, problem (20) is always feasible.

Theorem 2. The optimal value of problem (20) is lower bounded by $n^{-1}\mathrm{Tr}[\mathbb{T}(\Delta_+)]$.

Proof. The channel sum is $\mathbb{T}(\Delta_+ + X)$, and the sum of the eigenvalues of the channel sum $\mathbb{T}(\Delta_+ + X)$ is

$$\begin{aligned}\mathrm{Tr}[\mathbb{T}(\Delta_+ + X)] &= \mathrm{Tr}[\mathbb{T}(\Delta_+)] + \mathrm{Tr}[\mathbb{T}(X)] \\ &\ge \mathrm{Tr}[\mathbb{T}(\Delta_+)],\end{aligned}$$

where the first line is due to linearity of $\mathbb{T}$ (cf. Corollary 2) and the second line is due to Corollary 1 and Lemma 2 which imply $\mathrm{Tr}[\mathbb{T}(X)] = \mathrm{Tr}(X) \ge 0$. Finally, since $\mathbb{T}(\Delta_+ + X) = \alpha I$ (cf. problem (20)), we have $\alpha \ge n^{-1}\mathrm{Tr}[\mathbb{T}(\Delta_+)]$.

□

In summary, the solution of problem (20) is bounded as follows:

$$n^{-1}\mathrm{Tr}[\mathbb{T}(\Delta_+)] \le \hat{\alpha} \le \min(\|\mathbb{T}(\Delta_+)\|, 1). \tag{21}$$

If Δ_+ already corresponds to a scaled quantum channel, i.e., $\mathbb{T}(\Delta_+) = \alpha I$ for some α, then the optimal solution can be found: $\hat{\alpha} = n^{-1}\mathrm{Tr}[\mathbb{T}(\Delta_+)] = \|\mathbb{T}(\Delta_+)\| = \alpha$. In this case, $C_{\mathcal{F}_\Delta} = \alpha^{-1}\Delta_+$ and $C_{\mathcal{E}_\Delta} = (\alpha + \beta - 1)^{-1}\Delta_-$ can be found from equations (16) and (17) with $X = 0$ and equation (18).

4.1. Procedure for computing the lower- and upper-bound (p, q) curves

Suppose that we are given two quantum channels $\mathcal{E}$ and $\mathcal{F}$ of dimension n.

(i) Compute the Choi matrices $C_{\mathcal{E}}$ and $C_{\mathcal{F}}$ for the two channels using equation (3).

(ii) Fix β in the range of $(0, \infty)$. [Note that $\beta = 0$ or $\beta = \infty$ corresponds to $q = 1$ or $p = 1$ respectively, and these are trivial cases because either $\mathcal{E}'$ or $\mathcal{F}'$ becomes arbitrary.]

(iii) Eigen-decompose equation (15) to obtain $\Delta_{\pm}$.

(iv) Compute the channel sum $\mathbb{T}(\Delta_+)$ using equation (5).

(v) Compute the lower and upper bounds on $\hat{\alpha}$ using equation (21).

(vi) Given a bound, denoted as α, solve for p and q using equations (9) and (10).

We can repeat this procedure for a range of β to obtain the lower- and upper-bound (p, q) trade-off curves.

5. Examples

5.1. Difference between bit-flip and phase-flip channels

Given the bit-flip and phase-flip channels,

$$\mathcal{E}(\rho) = (1-a)I_2\rho I_2 + aX\rho X \tag{22}$$

$$\mathcal{F}(\rho) = (1-b)I_2\rho I_2 + bZ\rho Z \tag{23}$$

where a and b are the bit-flip and phase-flip probabilities, and

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

we compute the Choi matrices for the two channels and find the difference

$$\begin{aligned} &C_{\mathcal{E}} - \beta C_{\mathcal{F}} \\ =&(1-a-\beta+b\beta)|e_1\rangle\langle e_1| - b\beta|e_2\rangle\langle e_2| + a|e_3\rangle\langle e_3| \end{aligned} \tag{24}$$

where $\langle e_1| = [1,0,0,1]$, $\langle e_2| = [1,0,0,-1]$, and $\langle e_3| = [0,1,1,0]$. Next, we separate this into the positive and negative subspaces as in equation (15). Note that since we consider $a, b, \beta > 0$, the second term of the equation is negative and the third term is positive, while the first term can be non-negative or negative.

Case 1: $1 - a - \beta + b\beta \geq 0$. According to equation (15), we have

$$\begin{aligned} \Delta_+ &= (1-a-\beta+b\beta)|e_1\rangle\langle e_1| + a|e_3\rangle\langle e_3|, \\ \Delta_- &= b\beta|e_2\rangle\langle e_2| \end{aligned}$$

and

$$\begin{aligned} \mathbb{T}(\Delta_+) &= (1-\beta+b\beta)I_2, \\ \mathbb{T}(\Delta_-) &= b\beta I_2. \end{aligned}$$

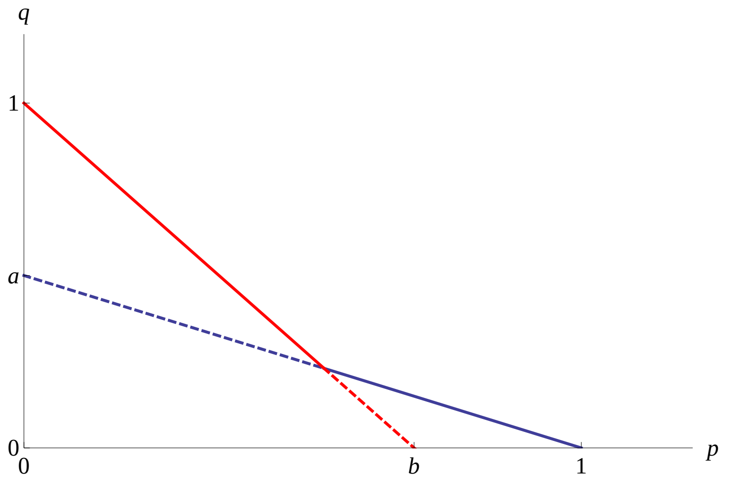


Figure 5. Optimal trade-off curve for the mixing probabilities p and q [defined in equations (1) and (2)] for the bit- and phase-flip channels given in equations (22) and (23). This curve is the solution to problem (8) or problem (20).

Therefore, using the bounds in equation (21), we obtain the optimal solution of problem (20) as $\hat{\alpha} = 1 - \beta + b\beta$.

Case 2: $1 - a - \beta + b\beta < 0$. According to equation (15), we have

$$\begin{aligned}\Delta_+ &= a|e_3\rangle\langle e_3|,\\ \Delta_- &= -(1 - a - \beta + b\beta)|e_1\rangle\langle e_1| + b\beta|e_2\rangle\langle e_2|\end{aligned}$$

and

$$\begin{aligned}\mathbb{T}(\Delta_+) &= aI_2,\\ \mathbb{T}(\Delta_-) &= (-1 + a + \beta)I_2.\end{aligned}$$

Therefore, using the bounds in equation (21), we obtain the optimal solution of problem (20) as $\hat{\alpha} = a$. Note that we are able to obtain the optimal solution in both cases instead of upper and lower bounds.

Finally, with α found for each case, we can compute a relation for p and q using equations (18) and (19):

$$\begin{cases} p = b - bq & , \quad \text{if } 1 - a - \beta + b\beta \geq 0 \\ q = a - ap & , \quad \text{if } 1 - a - \beta + b\beta < 0 \end{cases}. \tag{25}$$

This relation is depicted as the solid curve in figure 5, where the top-left (bottom-right) part corresponds to the first (second) case in equation (25). Essentially, the cusp in the figure is due to the transition from case 1 with 2 positive and 1 negative eigenvalues to case 2 with 1 positive and 2 negative eigenvalues in equation (24).

5.2. A pair of random qubit channels

We randomly generated two qubits channels each having four Kraus operators and they are listed in Appendix B. Using the procedure given in section 4.1, we compute the

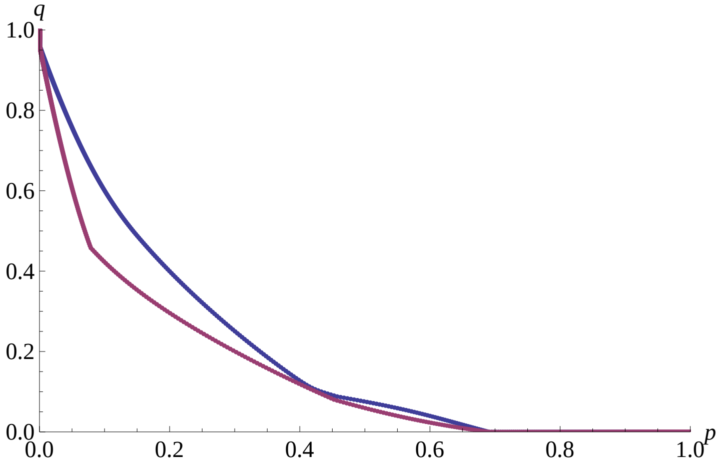


Figure 6. Lower- (purple) and upper-bound (blue) curves for the mixing probabilities p and q [defined in equations (1) and (2)] for two random qubit channels.

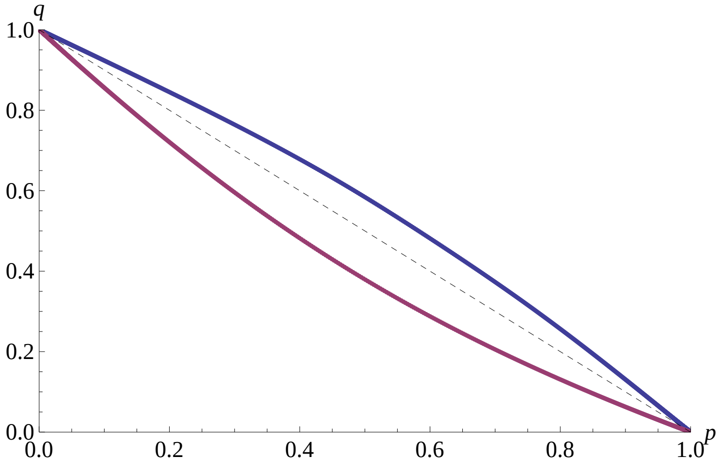


Figure 7. Lower- (purple) and upper-bound (blue) curves for the mixing probabilities p and q [defined in equations (1) and (2)] for two random four-dimensional channels. The solid curves are the bounds computed with equation (21) and the dashed curve is $q = 1 - p$.

lower- and upper-bound curves which are shown in figure 6. We make two observations. First, four cusps are obvious in the lower-bound curve, which are due to the transition of an eigenvalue of equation (15) from positive to negative (or vice versa). Note that at most four cusps can occur since the dimension of the Choi matrices are four. Second, there are regions where $q = 0$ for a range of p and where $p = 0$ for a range of q (the former is much bigger than the latter). These regions correspond to the case that one channel contains another channel and we will clarify this concept in section 7.2 later.

5.3. A pair of random four-dimensional channels

We randomly generated two four-dimensional channels each having four Kraus operators. (We do not list them here as they take up a lot of space.) Figure 7 shows the bounds. We remark that the solid upper-bound curve computed with equation (21) is not useful since it is above the $q = 1-p$ line which is a trivial upper-bound. Nevertheless, the lower-bound curve is useful since it allows only a narrow gap with the upper-bound line.

This example brings up an important point: in general, we should take the convex hull of an upper-bound curve as a refined upper-bound.

We also remark that the maximum number of cusps in the bounding curves is 16 since the dimension of the channels is 4. However, they are not apparent in the figure.

6. Relation with distinguishability

The diamond norm is related to the minimum error in discriminating between two quantum channels and is defined as

$$\|\mathcal{G}\|_\diamond \triangleq \max_\rho \|(I_\mathcal{K} \otimes \mathcal{G})(\rho)\|_1.$$

Here, an ancillary Hilbert space $\mathcal{K}$ is introduced and $I_\mathcal{K}$ is the identity map acting on it. The dimension of $I_\mathcal{K}$ is the same as the dimension of the Hilbert space of $\mathcal{G}$ [11]. The minimum error in distinguishing between $\mathcal{E}$ and $\mathcal{F}$ is given by [23]

$$P(\text{error}) = \frac{1}{2}\left(1 - \frac{1}{2}\|\mathcal{E} - \mathcal{F}\|_\diamond\right).$$

6.1. Upper bound

We can upper bound the diamond norm with the mixing probabilities p and q of our disguising problem as follows:

$$\begin{aligned}\|\mathcal{E} - \mathcal{F}\|_\diamond &= \|(1-p)\mathcal{E} - (1-q)\mathcal{F} + p\mathcal{E} - q\mathcal{F}\|_\diamond \\ &= \|q\mathcal{F}_\Delta - p\mathcal{E}_\Delta + p\mathcal{E} - q\mathcal{F}\|_\diamond \\ &\le q\|\mathcal{F}_\Delta\|_\diamond + p\|\mathcal{E}_\Delta\|_\diamond + p\|\mathcal{E}\|_\diamond + q\|\mathcal{F}\|_\diamond \\ &= 2(p+q), \qquad (26)\end{aligned}$$

where the second line is due to equations (1) and (2) with the disguising condition $\mathcal{E}' = \mathcal{F}'$ satisfied, the third line is due to the triangle inequality of the trace distance, and the fourth line comes from the fact that the trace norm of the channel output (a density matrix) is one. Note that equation (26) holds for any feasible (p, q) satisfying $\mathcal{E}' = \mathcal{F}'$, not just the optimal (p, q) trade-off curve.

When the two channels $\mathcal{E}$ and $\mathcal{F}$ are perfectly distinguishable, $\|\mathcal{E} - \mathcal{F}\|_\diamond = 2$. On the other hand, in our disguising problem, $p+q = 1$ is always achievable in equations (1) and (2) since we can set $p = 1 - q$, $\mathcal{E}_\Delta = \mathcal{F}$, and $\mathcal{F}_\Delta = \mathcal{E}$. Therefore, equation (26) is tight in this case.

6.2. Lower bound

We focus on the case where $p = q$, which means that $\beta = 1$. In this case, from equation (11), the smallest $p = q$ satisfies

$$\frac{p}{1-p} = \frac{q}{1-q} \leq \|\mathbb{T}(\Delta_+)\|, \tag{27}$$

where Δ_+ is the positive subspace of $C_{\mathcal{E}} - C_{\mathcal{F}}$.

We divide Δ_+ (of size $n^2 \times n^2$) into $n \times n$ blocks of equal size and denote the (i, j) block as Δ_{+ij}. Thus, $\Delta_+ = \sum_{i,j=0}^{n-1} |i\rangle\langle j| \otimes \Delta_{+ij}$, and it follows from the definition of $\mathbb{T}$ in equation (5) that

$$\begin{aligned}
\|\mathbb{T}(\Delta_+)\| &= \left\| \sum_{i,j=0}^{n-1} |i\rangle\langle j| \cdot \mathrm{Tr}[\Delta_{+ij}] \right\| \\
&= \max_{|\phi\rangle} \sum_{i,j=0}^{n-1} \langle \phi | i \rangle \langle j | \phi \rangle \cdot \sum_{k=0}^{n-1} \langle z_k | \Delta_{+ij} | z_k \rangle \\
&= \max_{|\phi\rangle} \sum_{k=0}^{n-1} \langle \phi | \otimes \langle z_k | \Delta_+ | \phi \rangle \otimes | z_k \rangle \\
&\leq n \max_{|\phi\rangle, |z\rangle} \langle \phi | \otimes \langle z | \Delta_+ | \phi \rangle \otimes | z \rangle,
\end{aligned} \tag{28}$$

where on the second and third lines $\{|z_k\rangle\}$ is an orthonormal basis, and $\langle \phi | \phi \rangle = \langle z | z \rangle = 1$.

Next, we consider the diamond norm:

$$\begin{aligned}
&\|\mathcal{E} - \mathcal{F}\|_\diamond \\
\geq\ & \max_{|\psi\rangle} \|(I \otimes (\mathcal{E} - \mathcal{F}))(|\psi\rangle\langle\psi|)\| \\
=\ & \max_{|\sigma\rangle, |\psi\rangle} |\langle \sigma | (I \otimes (\mathcal{E} - \mathcal{F}))(|\psi\rangle\langle\psi|) | \sigma \rangle| \triangleq Q,
\end{aligned} \tag{29}$$

where the second line is due to the fact that the trace norm is no less than the spectral norm, and both the auxiliary system and the original system have dimension n. Without loss of generality, using the Schmidt decomposition on the auxiliary and original systems, we can express

$$\begin{aligned}
|\sigma\rangle &= \sum_{i=0}^{n-1} \gamma_i |B_i\rangle |\sigma_i\rangle \\
|\psi\rangle &= \sum_{i=0}^{n-1} \lambda_i |B_i\rangle |\psi_i\rangle,
\end{aligned}$$

where $\{|B_i\rangle\}$ is an orthonormal basis,

$$\begin{aligned}
&\langle \sigma_i | \sigma_i \rangle = \langle \psi_i | \psi_i \rangle = 1 \text{ for all } i = 0, \ldots, n-1, \text{ and} \\
&\sum_{i=0}^{n-1} |\gamma_i|^2 = \sum_{i=0}^{n-1} |\lambda_i|^2 = 1.
\end{aligned} \tag{30}$$

Continuing with equation (29), the term to be maximized is equal to

$$\begin{aligned}&\left|\sum_{i,j=0}^{n-1} \gamma_i^* \lambda_i \langle \sigma_i | (\mathcal{E} - \mathcal{F})(|\psi_i\rangle\langle\psi_j|) |\sigma_j\rangle \lambda_j^* \gamma_j \right| \\ &= |\langle v|G|v\rangle|\end{aligned}$$

where

$$G \triangleq \begin{bmatrix} (\mathcal{E} - \mathcal{F})(|\psi_0\rangle\langle\psi_0|) & (\mathcal{E} - \mathcal{F})(|\psi_0\rangle\langle\psi_1|) & \cdots & \\ (\mathcal{E} - \mathcal{F})(|\psi_1\rangle\langle\psi_0|) & \ddots & & \\ \vdots & & & \\ & & & (\mathcal{E} - \mathcal{F})(|\psi_{n-1}\rangle\langle\psi_{n-1}|) \end{bmatrix}$$

and

$$|v\rangle \triangleq \begin{bmatrix} |\sigma_0\rangle \lambda_0^* \gamma_0 \\ \vdots \\ |\sigma_{n-1}\rangle \lambda_{n-1}^* \gamma_{n-1} \end{bmatrix}.$$

Note that G is not a standard Choi matrix since there is no requirement that $\{|\psi_i\rangle\}$ is an orthonormal basis and also $\langle v|v\rangle = \sum_i |\lambda_i|^2 |\gamma_i|^2$ may not be unity.

Continuing with equation (29), we have

$$Q \triangleq \max_{\{|\sigma_i\rangle\},\{\gamma_i\}} \max_{\{|\psi_i\rangle\},\{\lambda_i\}} |\langle v|G|v\rangle|$$

subject to the constraints in equation (30). Since $|\psi_i\rangle = |i\rangle$ and $\lambda_i = 1/\sqrt{n}$ satisfy the constraints,

$$\begin{aligned} Q &\geq \max_{|\tilde{v}\rangle} \frac{1}{n} |\langle \tilde{v}|(C_{\mathcal{E}} - C_{\mathcal{F}})|\tilde{v}\rangle| \\ &= \frac{1}{n} \|C_{\mathcal{E}} - C_{\mathcal{F}}\| \\ &= \frac{1}{n} \max(\|\Delta_+\|, \|\Delta_-\|), \end{aligned}$$

where the maximization is over any vector $|\tilde{v}\rangle$ with $\langle \tilde{v}|\tilde{v}\rangle = 1$ and G with the substitution $|\psi_i\rangle = |i\rangle$ is equal to $C_{\mathcal{E}} - C_{\mathcal{F}}$. Finally, note that $\|\Delta_+\|$ is larger than or equal to the maximization term in equation (28). Therefore,

$$\|\mathcal{E} - \mathcal{F}\|_\diamond \geq \frac{1}{n^2} \|\mathbb{T}(\Delta_+)\| \tag{31}$$

and combining with equation (27), the smallest $p = q$ must satisfy

$$\frac{p}{n^2(1-p)} \leq \|\mathcal{E} - \mathcal{F}\|_\diamond. \tag{32}$$

6.3. Summary

Using equations (26) and (32), the smallest $p = q$ must satisfy

$$\frac{p}{n^2(1-p)} \leq \|\mathcal{E} - \mathcal{F}\|_\diamond \leq 4p. \tag{33}$$

This shows that the disguising problem and the distinguishability problem are dual: when two channels are easy to distinguish (the diamond norm is large), it requires great effort to disguise one channel as the other ($p = q$ is large); and the reverse also holds. Note that the lower bound of equation (33) is less than or equal to the upper bound since $p = q$ is at most $1/2$ due to the fact that any point on the line $q = 1 - p$ is feasible (cf. section 3.1).

7. Other remarks

7.1. Triangle inequality

We apply the notion of triangle inequality to our mixing probabilities. Suppose $\mathcal{E}$ and $\mathcal{F}$ are compatible with mixing probabilities (p, q) and $\mathcal{G}$ and $\mathcal{F}$ with (p', q'), meaning that

$$\begin{align} (1-p)\mathcal{E}(\rho) + p\mathcal{E}_\Delta(\rho) &= (1-q)\mathcal{F}(\rho) + q\mathcal{F}_\Delta(\rho), \tag{34} \\ (1-p')\mathcal{G}(\rho) + p'\mathcal{G}_\Delta(\rho) &= (1-q')\mathcal{F}(\rho) + q'\mathcal{F}'_\Delta(\rho). \tag{35} \end{align}$$

Note that the harmonizing channels $\mathcal{F}_\Delta$ and $\mathcal{F}'_\Delta$ are different in general. We want to infer the distance profiles (p'', q'') for $\mathcal{E}$ and $\mathcal{G}$ from the distance profiles (p, q) and (p', q'). To do this, we propose the following method: cross-multiply equations (34) and (35) to make the coefficients of $\mathcal{F}$ equal and add additional terms to the two resultant equations to make the overall harmonizing channels on the right-hand sides equal. The result is

$$\begin{aligned} &(1-q')\left[(1-p)\mathcal{E} + p\mathcal{E}_\Delta\right] + (1-q)q'\mathcal{F}'_\Delta \\ &\qquad = (1-q')\left[(1-q)\mathcal{F} + q\mathcal{F}_\Delta\right] + (1-q)q'\mathcal{F}'_\Delta, \text{ and} \\ &(1-q)\left[(1-p')\mathcal{G} + p'\mathcal{G}_\Delta\right] + (1-q')q\mathcal{F}_\Delta \\ &\qquad = (1-q)\left[(1-q')\mathcal{F} + q'\mathcal{F}'_\Delta\right] + (1-q')q\mathcal{F}_\Delta, \end{aligned}$$

where we drop the dependence on ρ for simpler notation and assume not both q and q' equal to 1. Thus, the two left-hand sides are equal, giving

$$\begin{aligned} &(1-q')\left[(1-p)\mathcal{E} + p\mathcal{E}_\Delta\right] + (1-q)q'\mathcal{F}'_\Delta \\ &\qquad = (1-q)\left[(1-p')\mathcal{G} + p'\mathcal{G}_\Delta\right] + (1-q')q\mathcal{F}_\Delta. \end{aligned}$$

This equation is interpreted as $\mathcal{E}$ occurring with probability $(1-q')(1-p)/(1-qq')$ and its harmonizing channel with probability

$$p'' = \frac{p(1-q') + (1-q)q'}{1-qq'}, \tag{36}$$

and $\mathcal{G}$ occurring with probability $(1-q)(1-p')/(1-qq')$ and its harmonizing channel with probability

$$q'' = \frac{p'(1-q) + (1-q')q}{1-qq'}. \tag{37}$$

With equations (36)–(37), given an achievable pair of mixing probabilities (p, q) for $\mathcal{E}$ and $\mathcal{F}$ and another pair (p', q') for $\mathcal{F}$ and $\mathcal{G}$, we can compute an achievable pair (p'', q'')

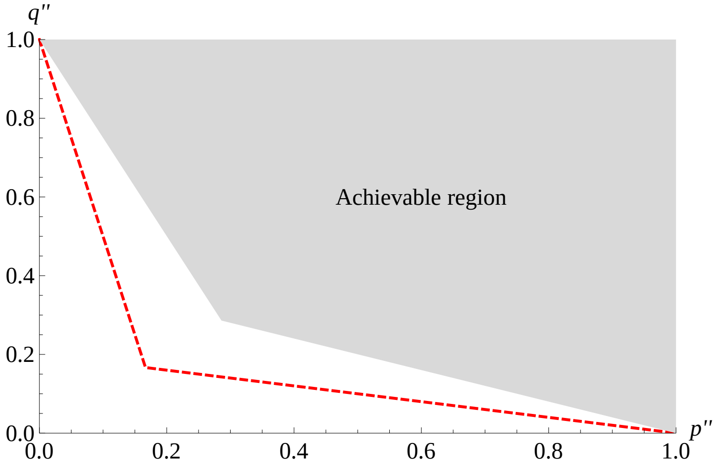


Figure 8. Achievable region for $\mathcal{E}$ and $\mathcal{G}$ obtained by equations (36) and (37) using the mixing probabilities of $\mathcal{E}$ and $\mathcal{F}$ and the mixing probabilities of $\mathcal{F}$ and $\mathcal{G}$. The dashed curve (red) is the optimal trade-off.

for $\mathcal{E}$ and $\mathcal{G}$. Tracing out the entire distance profiles (p, q) and (p', q') produces an achievable region (p'', q''). Note that this region is not in general a curve. Nevertheless, the bounding curve to this region can be regarded as a distance profile for $\mathcal{E}$ and $\mathcal{G}$. This curve is certainly achievable but may not be optimal and so it represents an upper bound to the optimal trade-off distance profile.

Note that even though we have the assumption $qq' < 1$, we can still obtain the end points. In particular, when $(p, q) = (0, 1)$ and $q' < 1$, we get $(p'', q'') = (0, 1)$; when $(p', q') = (0, 1)$ and $q < 1$, we get $(p'', q'') = (1, 0)$.

As an example, consider $\mathcal{E}$ and $\mathcal{F}$ given in equations (22) and (23) and

$$\mathcal{G}(\rho) = (1 - c)I_2\rho I_2 + c(XZ)\rho(ZX)$$

with $a = b = c = 0.2$. Figure 8 shows the achievable region for $\mathcal{E}$ and $\mathcal{G}$ obtained by equations (36) and (37) together with the optimal curve obtained in section 5.1. Note that due to symmetry, any pair of $\mathcal{E}$, $\mathcal{F}$, and $\mathcal{G}$ has the same optimal trade-off curve.

We now consider the triangle inequality for two scalar distances derived from our 2-dimensional measure: (i) the minimum of p subject to $p = q$ and (ii) the minimum of $p + q$. For case (i), suppose that the minimum for $\mathcal{E}$ and $\mathcal{F}$ is p_0 with $q_0 = p_0$ and the minimum for $\mathcal{F}$ and $\mathcal{G}$ is p_0' with $q_0' = p_0'$. Substituting these four parameters into equations (36) and (37) gives

$$p'' = \frac{p_0(1 - p_0') + (1 - p_0)p_0'}{1 - p_0 p_0'} = q''.$$

This means that the $p = q$ condition is preserved. Since $(1 - p_0')(1 - p_0 p_0') \leq 1$ and $(1 - p_0)(1 - p_0 p_0') \leq 1$, we have $p'' \leq p_0 + p_0'$. Finally, since p'' is only an achievable upper bound to the optimal mixing probability for $\mathcal{E}$ and $\mathcal{G}$, the triangle inequality is satisfied.

For case (ii), suppose that the minimum for $\mathcal{E}$ and $\mathcal{F}$ is $p_0 + q_0$ and the minimum for $\mathcal{F}$ and $\mathcal{G}$ is $p_0' + q_0'$. Substituting these four parameters into equations (36) and (37) gives

$$p'' + q'' = \frac{(p_0 + q_0)(1 - q_0') + (p_0' + q_0')(1 - q_0)}{1 - q_0 q_0'}.$$

Using the same argument as in case (i), we have $p'' + q'' \leq (p_0 + q_0) + (p_0' + q_0')$ and the triangle inequality is satisfied.

7.2. Channel containment

Given two quantum channels $\mathcal{E}$ and $\mathcal{F}$, we introduce the notion that $\mathcal{E}$ contains $\mathcal{F}$ if $\mathcal{F}$ is part of a mixture of $\mathcal{E}$:

$$\mathcal{E}(\rho) = (1 - q)\mathcal{F}(\rho) + q\mathcal{F}_\Delta(\rho) \tag{38}$$

where $\mathcal{F}_\Delta$ is some quantum channel and we require that $q < 1$ to make the containment of $\mathcal{F}$ non-trivial. To see if equation (38) holds, we convert it to the Choi representation and proceed as in equation (14) with $p = 0$. The question becomes for what values of q is $C_\mathcal{E} - (1 - q)C_\mathcal{F} = qC_{\mathcal{F}_\Delta}$ PSD. This is the same as asking whether $\Delta_- = 0$ in equation (15). In general, we fix a value of q and perform eigen-decomposition to see if $\Delta_- = 0$. But for the case that $C_\mathcal{F}$ is invertible, we can find the minimum q by equating $1 - q$ with the minimum eigenvalue of $C_\mathcal{F}^{-1/2} C_\mathcal{E} C_\mathcal{F}^{-1/2}$.

Note that $C_{\mathcal{F}_\Delta}$ is automatically trace-preserving, since $q\mathbb{T}(C_{\mathcal{F}_\Delta}) = \mathbb{T}(C_\mathcal{E} - (1 - q)C_\mathcal{F}) = I - (1 - q)I = qI$ as both $\mathcal{E}$ and $\mathcal{F}$ are trace-preserving.

7.3. Composition of quantum channels

Suppose that $\mathcal{E}_i$ and $\mathcal{F}_i$ can be made compatible with p_i and q_i according to the processing in equations (1)–(2), where $i = 1, 2$. This means that $\mathcal{E}_i'(\rho) = \mathcal{F}_i'(\rho)$ or

$$(1 - p_i)\mathcal{E}_i(\rho) + p_i\mathcal{E}_{\Delta i}(\rho) = (1 - q_i)\mathcal{F}_i(\rho) + q_i\mathcal{F}_{\Delta i}(\rho)$$

for all density matrices ρ and $i = 1, 2$. Then the composed quantum channels $\mathcal{E}_2 \circ \mathcal{E}_1$ and $\mathcal{F}_2 \circ \mathcal{F}_1$ can also be made compatible with $p = p_1 + p_2 - p_1p_2$ and $q = q_1 + q_2 - q_1q_2$:

$$(1 - p)\mathcal{E}_2 \circ \mathcal{E}_1(\rho) + p\mathcal{E}_\Delta(\rho) = (1 - q)\mathcal{F}_2 \circ \mathcal{F}_1(\rho) + q\mathcal{F}_\Delta(\rho). \tag{39}$$

To show this, note that since $\mathcal{E}_i(\rho)$ and $\mathcal{F}_i(\rho)$ are density matrices for any ρ, the following holds:

$$\mathcal{E}_2'(\mathcal{E}_1'(\rho)) = \mathcal{F}_2'(\mathcal{F}_1'(\rho)). \tag{40}$$

Expansion of the LHS gives

$$\begin{aligned} \mathcal{E}_2' \circ \mathcal{E}_1' = {} & (1 - p_1)(1 - p_2)\mathcal{E}_2 \circ \mathcal{E}_1 + p_1(1 - p_2)\mathcal{E}_2 \circ \mathcal{E}_{\Delta 1} \\ & + (1 - p_1)p_2\mathcal{E}_{\Delta 2} \circ \mathcal{E}_1 + p_1p_2\mathcal{E}_{\Delta 2} \circ \mathcal{E}_{\Delta 1}. \end{aligned} \tag{41}$$

We can readily see that the first term is the original composed channel with mixing probability $1 - p = (1 - p_1)(1 - p_2)$ and the sum of the last three terms represents the harmonizing channel $\mathcal{E}_\Delta$ with mixing probability $p = p_1 + p_2 - p_1 p_2$. Together with a similar argument for $\mathcal{F}'_2 \circ \mathcal{F}'_1$ proves the claim.

We can also argue that the composed quantum channels $\mathcal{E}_2 \circ \mathcal{E}_1$ and $\mathcal{F}_2 \circ \mathcal{F}_1$ can be made compatible with $p = p_1 + p_2$ and $q = q_1 + q_2$, by breaking up the first term of equation (41) and allocating the portion $p_1 p_2 \mathcal{E}_2 \circ \mathcal{E}_1$ to the harmonizing channel (and similarly for $\mathcal{F}$).

8. Application to quantum cryptography

The disguising condition $\mathcal{E}' = \mathcal{F}'$ can be used to upper bound the key generation rate in quantum cryptography [24, 25]. The intuitive idea is that the more easily the eavesdropper's channel can be disguised as the legitimate user's channel, the smaller is the amount of the generated key. We establish this idea quantitatively relating the mixing probability and the key generation rate.

In quantum cryptography with one-way forward communications, Alice repeatedly sends a quantum state ρ_a to Franky to establish a secret key, where $a = (a_0, a_1)$ and a_0 (a_1) is Alice's raw key basis (value) chosen independently between different transmissions. After the reception of the sequence of states by Franky, Alice sends classical information (including basis information, error correction information, and privacy amplification information) to him in order to correct bit errors and remove any information the eavesdropper Eve may have on the final key. Suppose that Eve launches a collective attack [26] which means that she applies the same unitary transformation U to each state sent by Alice (with sufficient ancillas). Thus, Franky's channel and Eve's channel are given as follows:

$$\mathcal{F}(\rho_a) = \mathrm{Tr}_{\mathtt{E}}(U(\rho_a \otimes |0\rangle\langle 0|)U^\dagger) \tag{42}$$

$$\mathcal{E}(\rho_a) = \mathrm{Tr}_{\mathtt{F}}(U(\rho_a \otimes |0\rangle\langle 0|)U^\dagger) \tag{43}$$

where we assume without loss of generality (w.l.o.g.) that the entire Hilbert space is divided into two systems $\mathtt{E}$ and $\mathtt{F}$. Furthermore, for simplicity and w.l.o.g., we assume that $\mathtt{E}$ and $\mathtt{F}$ have the same dimensions n (which can be assured by padding zeros as needed).

A key rate upper bound is given by the classical secret key capacity formula [27]

$$R = \sup_{\substack{U \leftarrow A \\ V \leftarrow U}} I(U; F|V) - I(U; E|V). \tag{44}$$

Note that the use of the classical formula is valid since when considering the upper bound, we can assume one particular strategy of Eve, which is to measure her quantum states separately. Here, A is Alice's random variable holding the raw key a, F and E are the Franky's and Eve's random variables holding the measurement outcomes f and e respectively.

We consider the upper bound of the key rate for the case where the disguising condition is

$$\mathcal{F}(\rho) = (1-p)\mathcal{E}(\rho) + p\mathcal{E}_\Delta(\rho). \tag{45}$$

That is, $1-p$ fraction of Franky's channel is Eve's channel. Thus, one would expect that only the remaining fraction of p could be used to generate a secret key and the key rate would be on the order of p.

For simplicity of discussion, instead of equation (44), we bound the key rate expression without the processing:

$$R' = I(A;F) - I(A;E). \tag{46}$$

Suppose that Bob's POVM is $\{M_{f,a_0}\}$ where $\sum_f M_{f,a_0} = I$ and it is dependent on the raw key basis a_0. Applying this POVM to equation (45) produces classical probability distributions

$$\begin{aligned} P_{AF}(a,f) &= P_A(a)\mathrm{Tr}(M_{f,a_0}\mathcal{F}(\rho_a)) \\ P_{AE}(a,f) &= P_A(a)\mathrm{Tr}(M_{f,a_0}\mathcal{E}(\rho_a)) \\ P_{AE_\Delta}(a,f) &= P_A(a)\mathrm{Tr}(M_{f,a_0}\mathcal{E}_\Delta(\rho_a)) \end{aligned}$$

which are related by

$$P_{AF}(a,f) = (1-p)P_{AE}(a,f) + pP_{AE_\Delta}(a,f). \tag{47}$$

Note that this relation can be explained by the following hypothetical probability distribution

$$P_{AFZ}(a,f,z) = \begin{cases} (1-p)P_{AE}(a,f), & \text{if } z=0 \\ pP_{AE_\Delta}(a,f), & \text{if } z=1 \end{cases}$$

in that $\sum_{z=0,1} P_{AFZ}(a,f,z)$ is equal to equation (47).

Now, we bound the first term of equation (46) as follows:

$$\begin{aligned} I(A;F) &\le I(A;FZ) \\ &= I(A;F|Z) + I(A;Z) \\ &= (1-p)I(A;F|z=0) + pI(A;F|z=1) \\ &= (1-p)I(A;E) + pI(A;E_\Delta) \end{aligned}$$

where on the second line we have $I(A;Z)=0$ since $P_{AZ}(a,z) = P_A(a)P_Z(z)$. Therefore, using equation (46), the key rate is bounded as

$$\begin{aligned} R' &\le (1-p)I(A;E) + pI(A;E_\Delta) - I(A;E) \\ &= p[I(A;E_\Delta) - I(A;E)] \\ &\le p\log_2 n \end{aligned} \tag{48}$$

where the last inequality is due to Holevo-Schumacher-Westmoreland channel capacity theorem [28, 29] and the fact that the maximum entropy for an n-dimensional state is $\log_2 n$. A similar analysis can be applied to the original key rate expression in equation (44) to obtain the same upper bound in equation (48).

9. Conclusions

The disguising problem tries to make two quantum channels identical, which is the reverse of the distinguishability problem which tries to maximize their difference in the measurement statistics in order to discriminate between them. Indeed, we showed that the two problems are related by proving that a certain combination of the mixing probabilities of the disguising problem upper bounds the diamond norm of the distinguishability problem. We also showed that the triangle equality holds for two scalar distances derived from the mixing probabilities.

Conventional measures on quantum channels are mostly based on trace distance or fidelity which are both concepts derived from measuring the distance between quantum states. In this paper, we propose a new measure genuinely for quantum channels. Note that the application of our measure to quantum states is possible but the result would be rather trivial since there is no more the need of making sure the harmonizing channels satisfy the TP condition for a linear map (which is ensured by the addition of X in equations (16) and (17)). This extra condition makes the calculation of our measure for quantum channels more difficult. Nevertheless, we obtain analytical lower- and upper-bounds which generate curves that are close to each other in many cases.

Our measure is based on the notion of minimizing the probabilities of channel mixing, which can be viewed as the costs for an channel intervener to make two channels the same. We show how these costs are linked to the key generation rate in quantum key distribution. The investigation of how these costs are linked to other quantum information processing tasks is a topic for future research. Also, open problems include efficient/approximate computation of our measure, and the effect of extending the Hilbert space dimensions of the channels by including ancillary systems (i.e., $\mathcal{E}$ becomes $\mathcal{E} \otimes I$) in the calculation of our measure.

Acknowledgments

We thank Masahito Hayashi, Seung-Hyeok Kye, and Jonathan Oppenheim for enlightening discussion. This work is supported in part by RGC under Grant No. 700709P and 700712P of the HKSAR Government.

Appendix A. Useful results related to quantum channels

This section discusses the tools and definitions related to quantum channels. A quantum channel is a linear map that is completely-positive (CP) and trace-preserving (TP).

Theorem 3. (Choi's theorem [22]) Given a linear map $\mathcal{E}$ and its Choi matrix $C_{\mathcal{E}}$, $C_{\mathcal{E}}$ is PSD if and only if $\mathcal{E}$ is a completely-positive map.

When $C_{\mathcal{E}}$ is Hermitian, $\mathcal{E}$ can also be represented in the operator-sum form:

$$\mathcal{E}(\rho) = \sum_i \lambda_i E_i \rho E_i^{\dagger}, \tag{A.1}$$

where $\lambda_i \in \mathbb{R}$ and $E_i \in \mathbb{C}^{n,n}$. This can be seen by taking some decomposition (e.g., eigen-decomposition) of $C_{\mathcal{E}}$ to be

$$C_{\mathcal{E}} = \sum_i \lambda_i |E_i\rangle\langle E_i| \tag{A.2}$$

where $\lambda_i \in \mathbb{R}$ and $|E_i\rangle \in \mathbb{C}^{n^2,1}$, and rearranging the vector $|E_i\rangle$ into the square matrix E_i as follows:

$$|E_i\rangle \equiv \begin{bmatrix} E_i(1,1) \\ E_i(2,1) \\ \vdots \\ E_i(n,1) \\ E_i(1,2) \\ \vdots \\ E_i(n,2) \\ \vdots \\ E_i(n,n) \end{bmatrix} = \mathrm{vec}(E_i) \tag{A.3}$$

where $E_i(k,l)$ is the (k,l) entry of E_i, and the vec operator creates a vector by stacking the columns of its operand (see, e.g., Ref. [30]). The dimension of E_i is $n \times n$ and the dimension of $C_{\mathcal{E}}$ is $n^2 \times n^2$. Note that the operator-sum form is not unique; there can be more than one such form corresponding to the same Choi matrix.

Observation 1. The Choi matrix of any channel given in the operator-sum form of equation (A.1) can be constructed by using equation (3) or equation (A.2). The equivalence of these two ways can be checked easily by direct expansion.

The next Theorem follows directly from the definition of the Choi matrix in equation (3).

Theorem 4. $\mathcal{E}(\rho) = \mathcal{F}(\rho)$ for all density matrices ρ if and only if $C_{\mathcal{E}} = C_{\mathcal{F}}$.

Definition 1. We define the *channel sum* for linear map $\mathcal{E}$ with a Hermitian Choi matrix as

$$T_{\mathcal{E}} \triangleq \sum_i \lambda_i E_i^\dagger E_i, \tag{A.4}$$

which has dimension $n \times n$.

The next lemma shows how to obtain the channel sum from channel outputs directly.

Lemma 4.

$$\begin{aligned} T_{\mathcal{E}} &= \sum_{i,j=0}^{n-1} |i\rangle\langle j| \cdot \mathrm{Tr}[\mathcal{E}(|j\rangle\langle i|)] \\ &= \mathrm{Tr}_2^t(C_{\mathcal{E}}) \end{aligned}$$

where Tr_2 is the partial trace over the second system and t represents transpose. Note that here, it does not matter whether the transpose is taken after or before the partial trace.

Proof. The (i, j) element of $S_{\mathcal{E}}$ defined in equation (A.4) is

$$\begin{aligned}\langle i|T_{\mathcal{E}}|j\rangle &= \langle i|\sum_k \lambda_k E_k^\dagger E_k|j\rangle \\ &= \mathrm{Tr}\left[\sum_k \lambda_k E_k|j\rangle\langle i|E_k^\dagger\right] \\ &= \mathrm{Tr}\left[\mathcal{E}(|j\rangle\langle i|)\right].\end{aligned}$$

□

Therefore, $T_{\mathcal{E}}$ is independent of the operator-sum form of $\mathcal{E}$ and is dependent only on the Choi matrix. This allows us to define the channel sum function

$$\begin{aligned}\mathbb{T}(C_{\mathcal{E}}) &:= \mathrm{Tr}_2^t(C_{\mathcal{E}}) \qquad \text{(A.5)} \\ &= T_{\mathcal{E}}\end{aligned}$$

where we used Lemma 4. The introduction of $\mathbb{T}$ facilitates the discussion of the channel sum with reference to only the Choi matrix.

Lemma 5. A linear map $\mathcal{E}$ is trace-preserving if and only if $\mathbb{T}(C_{\mathcal{E}}) = I$.

Proof. For $\mathcal{E}$ to be trace-preserving, the following must hold for all density matrices ρ:

$$\begin{aligned}\mathrm{Tr}(\rho) &= \mathrm{Tr}[\mathcal{E}(\rho)] \\ &= \mathrm{Tr}(T_{\mathcal{E}}\rho)\end{aligned}$$

where the last equality is due to equation (A.4). Since this holds for all ρ, $T_{\mathcal{E}} = I$. Then, using equation (A.5), $\mathbb{T}(C_{\mathcal{E}}) = I$.

The proof for the other direction is obvious. □

We remark that the eigenvalues of a Choi matrix C and of its channel sum $\mathbb{T}(C)$ are in general not the same. Nevertheless, they have the same trace.

Corollary 1. $\mathrm{Tr}(C) = \mathrm{Tr}(\mathbb{T}(C))$ for any Choi matrix C.

Thus, the trace of the Choi matrix of a quantum channel is its Hilbert space dimension n because the channel sum of a quantum channel is I_n. This corollary will be useful when we consider the lower bound of the channel distance.

Corollary 2. $C_{\mathcal{E}} + C_{\mathcal{F}} = C_{\mathcal{E}+\mathcal{F}}$ and $\mathbb{T}(C_{\mathcal{E}+\mathcal{F}}) = \mathbb{T}(C_{\mathcal{E}}) + \mathbb{T}(C_{\mathcal{F}})$ for linear maps $\mathcal{E}$ and $\mathcal{F}$.

Also, note that it can easily be checked that if C is PSD, $\mathbb{T}(C)$ is also PSD.

Appendix B. Channel specification for section 5.2

The two channels are $\mathcal{E}(\rho) = \sum_{i=1}^{4} E_i \rho E_i^\dagger$ and $\mathcal{F}(\rho) = \sum_{i=1}^{4} F_i \rho F_i^\dagger$, where

$$\begin{aligned}
E_1 &= \begin{bmatrix} -0.504828 & -0.331944 \\ -0.0133105 & 0.295026 \end{bmatrix}, \\
E_2 &= \begin{bmatrix} 0.419485 & 0.158018 \\ 0.330761 & 0.0616354 \end{bmatrix}, \\
E_3 &= \begin{bmatrix} 0.464696 & 0.251826 \\ -0.312786 & 0.165248 \end{bmatrix}, \\
E_4 &= \begin{bmatrix} 0.160149 & -0.346665 \\ -0.346665 & 0.750403 \end{bmatrix}, \\
F_1 &= \begin{bmatrix} -0.20917 & -0.248828 \\ 0.382771 & -0.451866 \end{bmatrix}, \\
F_2 &= \begin{bmatrix} -0.62412 & -0.425856 \\ 0.286902 & -0.0613943 \end{bmatrix}, \\
F_3 &= \begin{bmatrix} 0.216184 & -0.422341 \\ -0.403389 & 0.451605 \end{bmatrix}, \\
F_4 &= \begin{bmatrix} 0.236514 & 0.269256 \\ 0.269256 & 0.306531 \end{bmatrix}.
\end{aligned}$$

References

[1] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge University Press)

[2] Bengtsson I and Życzkowski K 2008 *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press)

[3] Uhlmann A 1976 *Rep. Math. Phys.* **9** 273–279

[4] Alberti P M and Uhlmann A 1983 *Lett. Math. Phys.* **7**(2) 107–112

[5] Jozsa R 1994 *J. Modern Optics* **41** 2315–2323

[6] Życzkowski K and Slomczyński W 2001 *J. Phys. A* **34** 6689

[7] Johnston N and Kribs D W 2010 *J. Math. Phys.* **51** 082202 (pages 16)

[8] Rastegin A 2010 *Quantum Info. Processing* **9**(1) 61–73

[9] Gilchrist A, Langford N K and Nielsen M A 2005 *Phys. Rev. A* **71**(6) 062310

[10] Belavkin V P, DAriano G M and Raginsky M 2005 *J. Math. Phys.* **46** 062106

[11] Kitaev A Y 1997 *Russian Mathematical Surveys* **52** 1191

[12] Aharonov D, Kitaev A and Nisan N 1998 Quantum circuits with mixed states *Proceedings of the thirtieth annual ACM symposium on Theory of computing* STOC '98 (New York, NY, USA: ACM) pp 20–30

[13] Rosgen B and Watrous J 2005 On the hardness of distinguishing mixed-state quantum computations *Proceedings of the 20th Annual IEEE Conference on Computational Complexity* CCC '05 (Washington, DC, USA: IEEE Computer Society) pp 344–354 ISBN 0-7695-2364-1

[14] Sacchi M F 2005 *Phys. Rev. A* **71**(6) 062340

[15] Li L and Qiu D 2008 *J. Phys. A* **41** 335302

[16] Piani M and Watrous J 2009 *Phys. Rev. Lett.* **102**(25) 250501

[17] Yu N, Duan R and Xu Q 2012 ArXiv:1201.1172 [quant-ph] (*Preprint* `e-print arXiv:1201.1172 [quant-ph]`)

[18] Watrous J 2009 *Theory of Computing* **5** 217–238

[19] Benenti G and Strini G 2010 *J. Phys. B* **43** 215508

[20] Johnston N, Kribs D W and Paulsen V I 2009 *Quantum Info. Comput.* **9** 16–35 ISSN 1533-7146

[21] Ben-Aroya A and Ta-Shma A 2010 *Quantum Info. Comput.* **10** 77–86 ISSN 1533-7146

[22] Choi M D 1975 *Linear Algebra and its Applications* **10** 285–290

[23] Helstrom C W 1976 *Quantum detection and estimation theory* (Academic Press)

[24] Bennett C H and Brassard G 1984 Quantum cryptography: Public key distribution and coin tossing *Proc. of IEEE Int. Conference on Computers, Systems, and Signal Processing* (IEEE Press, New York) pp 175–179

[25] Ekert A K 1991 *Phys. Rev. Lett.* **67** 661–663

[26] Biham E and Mor T 1997 *Phys. Rev. Lett.* **78**(11) 2256–2259

[27] Csiszár I and Körner J 1978 *IEEE Trans. Inf. Theory* **24** 339–348

[28] Holevo A 1998 *IEEE Trans. Inf. Theory* **44** 269–273 ISSN 0018-9448

[29] Schumacher B and Westmoreland M D 1997 *Phys. Rev. A* **56**(1) 131–138

[30] Horn R A and Johnson C R 1994 *Topics in Matrix Analysis* (Cambridge University Press)