Contraction coefficients for noisy quantum channels*

Fumio Hiai
Tohoku University (Emeritus)
Hakusan 3-8-16-303, Abiko 270-1154, Japan
hiai.fumio@gmail.com

Mary Beth Ruskai
Institute for Quantum Computing, University of Waterloo
Waterloo, Ontario, Canada
ruskai@member.ams.org

November 1, 2015

Abstract

Generalized relative entropy, monotone Riemannian metrics, geodesic distance, and trace distance are all known to decrease under the action of quantum channels. We give some new bounds on, and relationships between, the maximal contraction for these quantities.

2010 Mathematics Subject Classification: 46L60, 46L87, 15A63, 81P45, 53B50

Key Words: quantum channel; contraction coefficient; relative entropy; quantum divergence; monotone Riemannian metric; geodesic distance; Bures distance

Contents

*Dedicated to the memory of Professor Uffe Haagerup

1 Introduction

It is well-known that many quantities of interest in quantum information theory contract under the action of completely positive and trace-preserving (CPT) maps, which represent quantum channels, including the relative entropy $H(P,Q) \equiv \operatorname{Tr} P(\log P - \log Q)$ of two positive definite operators with $\operatorname{Tr} P = \operatorname{Tr} Q$. When Φ is a quantum channel, we can define

$$\eta^{\mathrm{RelEnt}}(\Phi) \equiv \sup \left\{ \frac{H(\Phi(P), \Phi(Q))}{H(P,Q)} : P, Q > 0,\ P \neq Q,\ \operatorname{Tr} P = \operatorname{Tr} Q \right\}, \tag{1}$$

which describes the maximal contraction under Φ. Another contraction coefficient can be defined with respect to the trace distance as $\eta^{\mathrm{Tr}}(\Phi) \equiv \sup \|\Phi(P-Q)\|_1 / \|P-Q\|_1$, where the supremum is taken over P, Q as above. This can be regarded as the quantum version of the Dobrushin coefficient of ergodicity [12].

The concept of contraction coefficient was defined in the classical case [10] and generalized to the quantum setting in [38]. Similar definitions can be given to describe the contraction of many other quantities. We consider here primarily contraction with respect to quantum divergences (a special case of quasi-entropies), monotone Riemannian metrics and geodesic distances arising from them. There are many relations between the contraction coefficients of these quantities, which are our main concern in this paper. We also study the dependence of these contraction coefficients on the particular operator convex functions used to define them. Both classical and quantum contraction coefficients have important applications to the problem of mixing time bounds of (quantum) Markov processes and in particular, (quantum) Markov chains, as demonstrated in, e.g., [10, 11, 32, 50].

We first recall what is known in the classical setting. A classical channel from $\mathbf{C}^d$ to $\mathbf{C}^{d'}$ can be represented by a $d' \times d$ column-stochastic matrix Λ. The trace-norm (or Dobrushin) contraction coefficient is $\eta^{\mathrm{Tr}}(\Lambda) \equiv \sup \|\Lambda x\|_1/\|x\|_1$, where the supremum is taken over non-zero $x \in \mathbf{R}^d$ with $\sum_i x_i = 0$. On the Riemannian manifold $\mathcal{P}_d$ of probability vectors $p = (p_1, \ldots, p_d)$, $p_i > 0$, $\sum_{i=1}^d p_i = 1$, the so-called Fisher-Rao metric is a unique classical monotone Riemannian metric, for which we have the Riemannian contraction coefficient $\eta^{\mathrm{Riem}}(\Lambda)$. When g is a strictly convex function on $(0, \infty)$ with $g(1) = 0$, the classical g-divergence extending the classical relative entropy is defined as $H_g(p, q) \equiv \sum_{i=1}^d g(p_i/q_i) q_i$ for $p, q \in \mathcal{P}_d$, for which we have the contraction coefficient $\eta_g^{\mathrm{RelEnt}}(\Lambda)$. The following relations between these contraction coefficients were proved in [9, 10]:

$$\eta_g^{\mathrm{RelEnt}}(\Lambda) = \eta^{\mathrm{Riem}}(\Lambda) \le \eta^{\mathrm{Tr}}(\Lambda) \le \sqrt{\eta^{\mathrm{Riem}}(\Lambda)} \tag{2}$$

whenever g is operator convex on $(0, \infty)$.

In the quantum setting, the study of monotone Riemannian metrics on the manifold $\mathcal{D}_d$ of $d \times d$ positive definite density matrices was begun by Morozova and Čencov [42]. Petz [46] then showed that there were infinitely many such metrics, corresponding to positive operator monotone functions on $(0, \infty)$. Following [22, 38] we use the set of operator convex functions $\kappa > 0$ on $(0, \infty)$ with $\kappa(1) = 1$ and $x\kappa(x) = \kappa(x^{-1})$ to parametrize symmetric monotone metrics on $\mathcal{D}_d$, $d \in \mathbb{N}$. Such κ functions correspond one-to-one, by $\kappa = 1/f$, to operator monotone functions $f > 0$ on $(0, \infty)$ with $f(1) = 1$ and $f(x) = xf(x^{-1})$ giving the same family of such metrics as in [46]. Thus, for each κ function we can define the Riemannian contraction coefficient $\eta_\kappa^{\mathrm{Riem}}(\Phi)$ of a channel Φ from the $d \times d$ matrix algebra $\mathbb{M}_d$ to $\mathbb{M}_{d'}$, and do so explicitly in (19) of Section 2.4. On the other hand, for each operator convex function g on $(0, \infty)$ with $g(1) = 0$ and $g''(1) > 0$ we have the quantum g-divergence $H_g(\rho, \gamma)$ for $\rho, \gamma \in \mathcal{D}_d$ and the corresponding contraction coefficient $\eta_g^{\mathrm{RelEnt}}(\Phi)$, as defined in (12) and (15) of Section 2.3. As shown in [38] and developed further here, the relation between $\eta_\kappa^{\mathrm{Riem}}(\Phi)$ and $\eta_g^{\mathrm{RelEnt}}(\Phi)$ and their dependence on the κ and g functions in the quantum setting are not as simple as in the classical setting.

This paper is organized as follows. In Section 2 precise definitions of quantum g-divergences and monotone metrics parametrized by the κ functions are given, for which we introduce the contraction coefficients $\eta_g^{\mathrm{RelEnt}}(\Phi)$ and $\eta_\kappa^{\mathrm{Riem}}(\Phi)$. Section 3 provides familiar examples of g-divergences and monotone metrics such as the BKM, the Wigner-Yanase, and the Bures metrics. In Section 4 a description of $\eta_\kappa^{\mathrm{Riem}}(\Phi)$ in terms of a certain eigenvalue problem developed in [38] is recalled, which establishes the relation $\eta_\kappa^{\mathrm{Riem}}(\Phi) \le \eta^{\mathrm{Tr}}(\Phi)$ when $\kappa(x) = x^{-1/2}$.

The main results in Section 5 are the general relations

$$\eta_\kappa^{\mathrm{geod}}(\Phi) = \eta_\kappa^{\mathrm{Riem}}(\Phi) \le \eta_g^{\mathrm{RelEnt}}(\Phi), \qquad \eta^{\mathrm{Tr}}(\Phi) \le \sqrt{\eta_\kappa^{\mathrm{Riem}}(\Phi)}, \tag{3}$$

when g is related to κ by $g(x) = (x-1)^2\kappa(x)$. Here $\eta_\kappa^{\mathrm{geod}}(\Phi)$ is the contraction with respect to the geodesic distance induced by the monotone metric for κ as defined in (21). A lemma slightly modified from [25] is given in Appendix A to prove the equality in (3). The first inequality in (3) was given in [38]; the second is proved in Section 5.1 strengthening results from [50, 48]. In Section 5.2, a partial ordering of contraction coefficients is shown to hold when the domain or range of the channel is a commutative subalgebra, i.e., classical. Some remarks on extensions to weak Schwarz maps are given in Section 5.3.

In Section 6 we treat qubit channels using the Bloch sphere representation. The section includes proofs of statements announced in [38], as well as additional results. In particular when Φ_T is a unital qubit channel described by a 3×3 matrix T as in [34] we prove that

$$\eta_\kappa^{\text{Riem}}(\Phi_T) = \eta_\kappa^{\text{geod}}(\Phi_T) = \eta_g^{\text{RelEnt}}(\Phi_T) = \|T\|_\infty^2 \tag{4}$$

for every κ and g so that in this case these contraction coefficients are independent of the defining functions κ and g. Next, for Φ in the simplest possible family of non-unital qubit channels (which collapse the Bloch sphere to a line), we estimate $\eta_\kappa^{\text{Riem}}(\Phi)$ for several particular cases of κ as well as $\eta_g^{\text{RelEnt}}(\Phi)$ for special g corresponding to the extreme κ functions. These estimates suffice to show that the equality conditions in (4) above do not extend to non-unital channels; that the contraction coefficients depend on the functions κ and g; and that several natural conjectures are false. Complete proofs of the results in Section 6, which are elementary but somewhat lengthy, are given in Appendix B

Finally in Section 7 we present further results on contraction coefficients for some special examples of Section 3. A remarkable result here is that the equality $\eta_{\text{BKM}}^{\text{Riem}}(\Phi) = \eta_{\text{BKM}}^{\text{RelEnt}}(\Phi)$ holds for every channel Φ, where $\eta_{\text{BKM}}^{\text{Riem}}$ denotes the BKM metric contraction and $\eta_{\text{BKM}}^{\text{RelEnt}}$ the contraction with respect to the *symmetrized* relative entropy $H(P,Q) + H(Q,P)$. But the equality between $\eta_{\text{BKM}}^{\text{Riem}}(\Phi)$ and $\eta^{\text{RelEnt}}(\Phi)$ in (1) is left open.

2 Notation and Definitions

2.1 Basic notation

For each $d \in \mathbb{N}$ we write $\mathbb{M}_d$, $\mathbb{H}_d$, $\mathbb{P}_d$, and $\overline{\mathbb{P}}_d$ for the sets of $d \times d$ complex, Hermitian, positive definite, and positive semi-definite matrices, respectively. We also denote by $\mathcal{D}_d$ the set of $d \times d$ positive definite density matrices and $\overline{\mathcal{D}}_d$ the set of all $d \times d$ density matrices, i.e., $\mathcal{D}_d = \{\rho \in \mathbb{P}_d : \operatorname{Tr} \rho = 1\}$ and $\overline{\mathcal{D}}_d = \{\rho \in \overline{\mathbb{P}}_d : \operatorname{Tr} \rho = 1\}$, where Tr is the usual trace functional on $\mathbb{M}_d$. The trace-norm of $X \in \mathbb{M}_d$ is $\|X\|_1 \equiv \operatorname{Tr} |X|$. Recall that $\mathbb{M}_d$ identified with $\mathcal{B}(\mathbf{C}^d)$ becomes a Hilbert space when equipped with the Hilbert-Schmidt inner product

$$\langle X, Y \rangle \equiv \operatorname{Tr} X^* Y, \qquad X, Y \in \mathbb{M}_d,$$

together with the Hilbert-Schmidt norm $\|X\|_2 \equiv (\operatorname{Tr} X^* X)^{1/2}$. A real subspace $\mathbb{H}_d$ of $\mathbb{M}_d$ is identified with the Euclidean space of dimension d^2, and $\mathcal{D}_d$ is a smooth Riemannian manifold whose tangent space at any foot point is identified with $\mathbb{H}_d^0 \equiv \{A \in \mathbb{H}_d : \operatorname{Tr} A = 0\}$. Functions $f(A)$ of matrices $A \in \mathbb{H}_d$ are defined via the usual functional calculus.

We will use linear maps $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$, and denote by $\widehat{\Phi}$ the adjoint of Φ with respect to the Hilbert-Schmidt inner product, i.e., $\langle \Phi(X), Y \rangle = \langle X, \widehat{\Phi}(Y) \rangle$ for all $X \in \mathbb{M}_d$ and $Y \in \mathbb{M}_{d'}$. As usual in quantum information, we call Φ a (quantum) *channel* if Φ is CPT (i.e., completely positive and trace-preserving) map. Most of the maps we consider will be constructed from the *left* and *right multiplication operators*, respectively, i.e., $L_A X \equiv AX$ and $R_B X \equiv XB$ for $A, B, X \in \mathbb{M}_d$. For each $A, B \in \mathbb{P}_d$, L_A and R_B are commuting positive invertible operators on the Hilbert space $\mathbb{M}_d$ (however, they are not positive in the sense of mapping $\mathbb{P}_d$ into $\overline{\mathbb{P}}_d$). More generally, for functions $f : (0, \infty) \to \mathbf{R}$ we have $L_{f(A)} = f(L_A)$ and $R_{f(B)} = f(R_B)$.

2.2 Operator convex functions

A real function f on $(0,\infty)$ is said to be *operator monotone* (or operator monotone increasing) if $A \ge B$ implies $f(A) \ge f(B)$ for every $A, B \in \mathbb{P}_d$ with any $d \in \mathbb{N}$, and *operator monotone decreasing* if $-f$ is operator monotone. A real function g on $(0,\infty)$ is said to be *operator convex* if

$$g(\lambda A + (1-\lambda)B) \le \lambda g(A) + (1-\lambda)g(B)$$

for all $A, B \in \mathbb{P}_d$ with any $d \in \mathbb{N}$ and all $\lambda \in (0,1)$, and *operator concave* if $-g$ is operator convex. The theory of operator monotone and operator convex functions was initiated by Löwner [41] and Kraus [37], respectively. For details, see, e.g., [5, Section V.4], also [1, 13, 21].

In this work the following classes of operator convex functions play a special role:

$$\begin{aligned}
&\mathcal{G} \equiv \{g : (0,\infty) \to \mathbf{R}, \text{ operator convex}, \ g(1) = 0, \ g''(1) > 0\}, \\
&\mathcal{G}_{\mathrm{sym}} \equiv \{g : (0,\infty) \to [0,\infty), \text{ operator convex}, \ g(x) = xg(x^{-1}) \text{ for } x > 0, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad g(1) = g'(1) = 0, \ g''(1) = 2\}, \\
&\mathcal{K} \equiv \{\kappa : (0,\infty) \to (0,\infty), \text{ operator convex}, \ x\kappa(x) = \kappa(x^{-1}) \text{ for } x > 0, \ \kappa(1) = 1\}.
\end{aligned}$$

By Proposition 2.2 below there is a one-to-one correspondence $\kappa \in \mathcal{K} \leftrightarrow g \in \mathcal{G}_{\mathrm{sym}}$ determined by

$$g(x) = (x-1)^2 \kappa(x). \tag{5}$$

It is easy to see that if $g \in \mathcal{G}$ then $\widetilde{g}(x) \equiv xg(x^{-1})$ is also in $\mathcal{G}$. Indeed, if $g \in \mathcal{G}$, then $g(x)/(x-1) = (g(x) - g(1))/(x-1)$ is operator monotone on $(0,\infty)$ by Kraus' theorem, and hence

$$\frac{\widetilde{g}(x) - \widetilde{g}(1)}{x-1} = \frac{xg(x^{-1})}{x-1} = -\frac{g(x^{-1})}{x^{-1}-1}$$

is also operator monotone so that $\widetilde{g}$ is operator convex on $(0,\infty)$. Moreover, noting that $g''(1) = \widetilde{g}''(1)$, we define the *symmetrization* of g by

$$g_{\mathrm{sym}} \equiv \frac{g + \widetilde{g}}{g''(1)} \in \mathcal{G}_{\mathrm{sym}}. \tag{6}$$

Proposition 2.1. (i) *If $g : (0,\infty) \to \mathbf{R}$ is an operator convex function, then there exist a unique constant $c \ge 0$ and a unique positive measure μ on $[0,\infty)$ with $\int_{[0,\infty)} (1+s)^{-1}\, d\mu(s) < +\infty$ such that*

$$g(x) = g(1) + g'(1)(x-1) + c(x-1)^2 + \int_{[0,\infty)} \frac{(x-1)^2}{x+s}\, d\mu(s), \quad x \in (0,\infty). \tag{7}$$

(ii) *If $\kappa : (0,\infty) \to \mathbf{R}$ is an operator convex function and it satisfies the normalization $\kappa(1) = 1$ and the symmetry condition $x\kappa(x) = \kappa(x^{-1})$ for all $x > 0$, then $\kappa(x) > 0$ for all $x > 0$ (hence $\kappa \in \mathcal{K}$) and there exists a unique probability measure m on $[0,1]$ such that*

$$\begin{aligned}
\kappa(x) &= \int_{[0,1]} \frac{1+x}{(x+s)(1+sx)} \cdot \frac{(1+s)^2}{2}\, dm(s) \\
&= \int_{[0,1]} \left(\frac{1}{x+s} + \frac{1}{sx+1}\right) \frac{1+s}{2}\, dm(s), \qquad x \in (0,\infty).
\end{aligned} \tag{8}$$

The integral expression (7) was given in [38] and is a special case of [14, (5.2)]. Since

$$g''(1) = 2\left(c + \int_{[0,\infty)} \frac{1}{1+s}\, d\mu(s)\right), \tag{9}$$

we note that $g''(1) > 0$ if and only if $c + \mu([0,\infty)) > 0$, or equivalently, g is not a linear function. For the proof of (8), see [22, Appendix A.2]. It is also obvious that $\kappa(x) > 0$ for all $x > 0$ whenever κ is a convex function with $\kappa(1) > 0$ and $x\kappa(x) = \kappa(x^{-1})$ for all $x > 0$.

By Proposition 2.1, $\mathcal{K}$ is a Bauer simplex (in a locally convex topological vector space consisting of real functions on $(0,\infty)$ in the pointwise convergence topology), whose extreme points are

$$\kappa_s(x) \equiv \frac{(1+s)^2}{2} \cdot \frac{1+x}{(x+s)(1+sx)} = \frac{1+s}{2}\left(\frac{1}{x+s} + \frac{1}{1+sx}\right), \qquad 0 \le s \le 1. \tag{10}$$

It is well-known (and immediately seen from the integral expression (8)) that $\kappa_1(x) = 2/(1+x)$ is the smallest element of $\mathcal{K}$ and $\kappa_0(x) = (1+x)/2x$ is the largest in $\mathcal{K}$ so that

$$\frac{2}{1+x} \le \kappa(x) \le \frac{1+x}{2x}, \qquad \kappa \in \mathcal{K}. \tag{11}$$

In the sequel we use the more explicit notations $\kappa_{\min}$ for κ_1 and $\kappa_{\max}$ for κ_0.

Proposition 2.2. *For a function $\kappa : (0,\infty) \to (0,\infty)$ consider the following conditions:*

(a) *κ is operator convex,*

(b) *κ is operator monotone decreasing,*

(c) *$g(x) \equiv (x-1)^2\kappa(x)$ is operator convex.*

Then (a) $\Leftarrow$ (b) $\Leftrightarrow$ (c). *Moreover, if $x\kappa(x) = \kappa(x^{-1})$ for all $x > 0$ or equivalently $g(x) = xg(x^{-1})$ for all $x > 0$, then the above conditions* (a)–(c) *are all equivalent.*

Proof. For (b) $\Rightarrow$ (a), see [22, Theorem 2.4]. As for (c), Kraus' theorem (see, e.g., [21, Corollary 2.7.8]) implies that g is operator convex if and only if

$$h(x) \equiv \frac{g(x) - g(1)}{x-1} = (x-1)\kappa(x), \qquad x > 0,$$

is operator monotone. The latter is also equivalent to the condition that $\kappa(x) = (h(x) - h(1))/(x-1)$ is operator monotone decreasing. Indeed, this is seen from the facts that h is operator monotone on $(0,\infty)$ if and only if it has an integral expression

$$h(x) = h(1) + \gamma(x-1) + \int_{[0,\infty)} \frac{x-1}{x+s}\, d\mu(s), \qquad x \in (0,\infty),$$

where $\gamma \ge 0$ and μ is a positive measure on $[0,\infty)$ with $\int_{[0,\infty)} (1+s)^{-1}\, d\mu(s) < +\infty$ (see [14, Theorem 1.9]), and that κ is operator monotone decreasing on $(0,\infty)$ if and only if it has an integral expression

$$\kappa(x) = \gamma + \int_{[0,\infty)} \frac{1}{x+s}\, d\mu(s), \qquad x \in (0,\infty),$$

where γ and μ are same as above (see [19], also [3, Theorem 3.1]).

Next it is immediate to check that the conditions $x\kappa(x) = \kappa(x^{-1})$ and $g(x) = xg(x^{-1})$ for all $x > 0$ are equivalent. Under this symmetry condition, the implication (a) $\Rightarrow$ (b) follows immediately from the integral expression (8), as shown in [22, Theorem 2.4]. □

2.3 Relative entropy or g-divergence

For every $g \in \mathcal{G}$ and every $A, B \in \mathbb{P}_d$ the (quantum) *g-divergence* of A relative to B is defined by

$$H_g(A, B) \equiv \big\langle B^{1/2}, g(L_A R_B^{-1})(B^{1/2})\big\rangle, \tag{12}$$

which is a generalization of the relative entropy and is a special case of *quasi-entropies* [36, 44, 45]. The most important property of $H_g(A, B)$ is the *monotonicity*

$$H_g(\Phi(A), \Phi(B)) \le H_g(A, B)$$

for every $A, B \in \mathbb{P}_d$ and every CPT map $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$. This was first proved by Petz [44, 45] under slightly more restricted situations, and the above extension is in [38], see also [51, 23].

From the integral expression (7) it is known [38, Theorem II.5] that for every $\rho, \gamma \in \mathcal{D}_d$,

$$H_g(\rho, \gamma) = c \operatorname{Tr}(\rho - \gamma)^2 \gamma^{-1} + \int_{[0,\infty)} \operatorname{Tr}(\rho - \gamma) \frac{1}{L_\rho + sR_\gamma}(\rho - \gamma)\, d\mu(s). \tag{13}$$

Since $(L_\rho + sR_\gamma)^{-1}$ is a positive invertible operator on $\mathbb{M}_d$, the above expression together with (9) implies that $H_g(\rho, \gamma) \ge 0$ and that $H_g(\rho, \gamma) = 0$ if and only if $\rho = \gamma$.

In particular, when $g \in \mathcal{G}_{\mathrm{sym}}$ and $\kappa \in \mathcal{K}$ are given with (5), we note (see [38, Theorem II.5]) that for every $A, B \in \mathbb{P}_d$,

$$\begin{aligned} H_g(A, B) &= \big\langle A - B, R_B^{-1}\kappa(L_A R_B^{-1})(A - B)\big\rangle \\ &= \big\langle A - B, L_A^{-1}\kappa(R_B L_A^{-1})(A - B)\big\rangle = H_g(B, A). \end{aligned}$$

For every $g \in \mathcal{G}$ with symmetrization g_{sym} in (6) we have

$$H_{\tilde{g}}(A, B) = H_g(B, A), \qquad H_{g_{\mathrm{sym}}}(A, B) = \frac{H_g(A, B) + H_g(B, A)}{g''(1)}. \tag{14}$$

For each CPT map $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$ and each $g \in \mathcal{G}$ we introduce the *contraction coefficient* of Φ with respect to the g-divergence H_g by

$$\eta_g^{\mathrm{RelEnt}}(\Phi) \equiv \sup_{\rho, \gamma \in \mathcal{D}_d,\, \rho \ne \gamma} \frac{H_g(\Phi(\rho), \Phi(\gamma))}{H_g(\rho, \gamma)} \quad (\le 1). \tag{15}$$

From (14) we easily see that

$$\eta_{\tilde{g}}^{\mathrm{RelEnt}}(\Phi) = \eta_g^{\mathrm{RelEnt}}(\Phi), \qquad \eta_{g_{\mathrm{sym}}}^{\mathrm{RelEnt}}(\Phi) \le \eta_g^{\mathrm{RelEnt}}(\Phi). \tag{16}$$

2.4 Riemannian metrics and geodesic distance

Given a function $\kappa \in \mathcal{K}$ we define, for any $A \in \mathbb{P}_d$, a linear map $\Omega_A^\kappa : \mathbb{M}_d \to \mathbb{M}_d$ by

$$\Omega_A^\kappa(X) \equiv R_A^{-1}\kappa\big(L_A R_A^{-1}\big)(X) = L_A^{-1}\kappa\big(R_A L_A^{-1}\big)(X), \qquad X \in \mathbb{M}_d,$$

where the equality of the two expressions follows from $x\kappa(x) = \kappa(x^{-1})$. The positivity condition in the sense that $\Omega_A^\kappa(\mathbb{P}_d) \subset \overline{\mathbb{P}}_d$ (equivalent to complete positivity) for the map Ω_A^κ has thoroughly been investigated in [22] with a lot of sample discussions.

Associated with $\kappa \in \mathcal{K}$ a Riemannian metric M^κ on the Riemannian manifold $\mathcal{D}_d$ is defined by

$$M_\rho^\kappa(A,B) \equiv \big\langle A, \Omega_\rho^\kappa(B)\big\rangle, \qquad A,B \in \mathbb{H}_d^0,\ \rho \in \mathcal{D}_d. \tag{17}$$

This family of Riemannian metrics on $\mathcal{D}_d$ ($d \in \mathbb{N}$) induced by $\kappa \in \mathcal{K}$ is called *monotone metrics* since the class was characterized by Petz [46] with the monotonicity property:

$$M_{\Phi(\rho)}(\Phi(A),\Phi(A)) \le M_\rho(A,A), \qquad A \in \mathbb{H}_d^0,\ \rho \in \mathcal{D}_d, \tag{18}$$

for every CPT map $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$ with any d, d'. Here, note that although $\Phi(\rho)$ in $\mathbb{M}_{d'}$ is not necessarily positive definite, the left-hand side of (18) is well defined by regarding $\Phi(\rho)$ and $\Phi(A)$ as matrices in $\Pi\mathbb{M}_{d'}\Pi \cong \mathbb{M}_k$ where $\Pi \equiv \operatorname{supp}\Phi(I_d)$, the support projection, and $k \equiv \dim \Pi$. More recent results on monotone metrics are found in [25, 26]. Also, note that if ρ and A commute, then $\langle A, \Omega_\rho^\kappa(A)\rangle = \operatorname{Tr}\rho^{-1}A^2$ independently of the choice of $\kappa \in \mathcal{K}$. This fact is essentially same as the classical result that there is only one monotone metric in the classical setting, known as the *Fisher-Rao metric*.

For each $\kappa \in \mathcal{K}$ the *contraction coefficient* of a CPT map Φ with respect to the monotone metric M^κ induced by κ is defined by

$$\eta_\kappa^{\mathrm{Riem}}(\Phi) \equiv \sup_{\rho\in\mathcal{D}_d} \sup_{A\in\mathbb{H}_d^0,\, A\ne 0} \frac{\big\langle \Phi(A), \Omega_{\Phi(\rho)}^\kappa(\Phi(A))\big\rangle}{\langle A, \Omega_\rho^\kappa(A)\rangle} \quad (\le 1). \tag{19}$$

Remark 2.3. The expression (17) makes sense for general $A, B \in \mathbb{M}_d$ and $\rho \in \mathbb{P}_d$, which defines quadratic forms on $\mathbb{M}_d \times \mathbb{M}_d$. Then the monotonicity property (18) holds in this general setting for $A \in \mathbb{M}_d$ and $\rho \in \mathbb{P}_d$, as explicitly stated in [38, Theorems II.10, II.14] (also implicitly in [46]). The space $\mathbb{P}_d$ ($\supset \mathcal{D}_d$) is a smooth Riemannian manifold with the tangent space $\mathbb{H}_d$ ($\supset \mathbb{H}_d^0$). For each $\kappa \in \mathcal{K}$ a Riemannian metric M^κ on $\mathbb{P}_d$ can be defined by (17) extended to $A, B \in \mathbb{H}_d$ and $\rho \in \mathbb{P}_d$. Observe that for any CPT map Φ, if $A = \rho \in \mathcal{D}_d$ then we have

$$\frac{\big\langle \Phi(A), \Omega_{\Phi(\rho)}^\kappa(\Phi(A))\big\rangle}{\langle A, \Omega_\rho^\kappa(A)\rangle} = \frac{\operatorname{Tr}\Phi(\rho)}{\operatorname{Tr}\rho} = 1.$$

Therefore, the restriction to the tangent space of trace zero matrices as in (19) is essential in our definition of a meaningful contraction coefficient for M^κ.

We write $\eta_{\max}^{\mathrm{Riem}}(\Phi)$ and $\eta_{\min}^{\mathrm{Riem}}(\Phi)$ for the contraction coefficients associated with the largest $\kappa_{\max}(x) = (1+x)/2x$ and smallest $\kappa_{\min}(x) = 2/(1+x)$ functions in $\mathcal{K}$, while these need not be the largest and smallest contraction coefficients for a given Φ. (Indeed, Proposition 5.5 below suggests that this is not true in general.)

We write $D_\kappa(\rho,\gamma)$ for the *geodesic distance* with respect to M^κ; namely,

$$D_\kappa(\rho,\gamma) \equiv \inf_\xi \int_0^1 \sqrt{\big\langle \xi'(t), \Omega_{\xi(t)}^\kappa(\xi'(t))\big\rangle}\, dt, \tag{20}$$

where the infimum is taken over all (piecewise) smooth curves joining ρ, γ in $\mathcal{D}_d$. The monotonicity of M^κ obviously implies that

$$D_\kappa(\Phi(\rho),\Phi(\gamma)) \le D_\kappa(\rho,\gamma)$$

for every CPT map $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$. The *contraction coefficient* of Φ with respect to the geodesic distance D_κ is then defined by

$$\eta_\kappa^{\mathrm{geod}}(\Phi) \equiv \left[\sup_{\rho,\gamma\in\mathcal{D}_d,\, \rho\neq\gamma} \frac{D_\kappa(\Phi(\rho),\Phi(\gamma))}{D_\kappa(\rho,\gamma)} \right]^2 \quad (\leq 1). \tag{21}$$

We will prove the equality $\eta_\kappa^{\mathrm{Riem}}(\Phi) = \eta_\kappa^{\mathrm{geod}}(\Phi)$ in Section 5.1 while the inequality $\eta_\kappa^{\mathrm{Riem}}(\Phi) \geq \eta_\kappa^{\mathrm{geod}}(\Phi)$ was shown in [38, Theorem IV.2].

Remark 2.4. For any pair $\rho,\gamma \in \mathcal{D}_d$, in addition to (20) one can consider the geodesic distance $\widetilde{D}_\kappa(\rho,\gamma)$ with the same expression now taken over all smooth curves joining ρ,γ in $\mathbb{P}_d$ without confining them to $\mathcal{D}_d$. The difference between the manifolds $\mathcal{D}_d$ and $\mathbb{P}_d$ implies that $\widetilde{D}_\kappa(\rho,\gamma) \leq D_\kappa(\rho,\gamma)$. Moreover, for every $\kappa \in \mathcal{K}$ this inequality must be strict at least in the case when $\rho\gamma = \gamma\rho$ and $\rho \neq \gamma$. Indeed, one can apply the monotonicity of M^κ to the trace-preserving conditional expectation onto a commutative subalgebra $\mathcal{A}$ containing ρ,σ to see that curves ξ can be confined, in both definitions of $D_\kappa(\rho,\gamma)$ and $\widetilde{D}_\kappa(\rho,\gamma)$, to those inside $\mathcal{A}$. (See also the proof of Lemma 5.4.) Therefore, in this situation both $D_\kappa(\rho,\gamma)$ and $\widetilde{D}_\kappa(\rho,\gamma)$ are independent of the choice of $\kappa \in \mathcal{K}$ so that formulas (27) and (28) of the next section hold for any $\kappa \in \mathcal{K}$. Then inequality (29) implies that $\widetilde{D}_\kappa(\rho,\gamma) < D_\kappa(\rho,\gamma)$.

3 Examples

Basic examples of g-divergences and monotone metrics are in order here. Further discussions and results for these cases will be later given in Section 7.

Example 1. (Relative entropy and BKM metric) The function $g(x) = x\log x \in \mathcal{G}$ gives the (usual) *relative entropy*, i.e.,

$$H_{x\log x}(\rho,\gamma) = H(\rho,\gamma) \equiv \mathrm{Tr}\,\rho(\log\rho - \log\gamma).$$

Moreover, $\widetilde{g}(x) = -\log x$ and $H_{-\log x}(\rho,\gamma) = H(\gamma,\rho)$. The symmetrization

$$g_{\mathrm{BKM}}(x) \equiv x\log x - \log x \in \mathcal{G}_{\mathrm{sym}}$$

corresponds to the function

$$\kappa_{\mathrm{BKM}}(x) \equiv \frac{\log x}{x-1} \in \mathcal{K},$$

which gives

$$\Omega_\rho^{\mathrm{BKM}}(X) = \frac{\log L_\rho - \log R_\rho}{L_\rho - R_\rho}(X) = \int_0^\infty \frac{1}{\rho + tI}\, X\, \frac{1}{\rho + tI}\, dt. \tag{22}$$

The corresponding monotone metric is the so-called *Bogolieubov* (or *Kubo-Mori*) *metric*. We write $\eta_{\mathrm{BKM}}^{\mathrm{Riem}}(\Phi)$ and $\eta_{\mathrm{BKM}}^{\mathrm{RelEnt}}(\Phi)$ for the contraction coefficients associated with κ_{BKM} and g_{BKM}. Rather surprisingly, it will be shown in Theorem 7.1 that $\eta_{\mathrm{BKM}}^{\mathrm{Riem}}(\Phi) = \eta_{\mathrm{BKM}}^{\mathrm{RelEnt}}(\Phi)$ holds for every CPT map Φ. However, we know from Theorem 6.6 that this property does not hold in general.

Example 2. (Maximal metric) The function $g(x) = (x-1)^2 \in \mathcal{G}$ yields the quadratic relative entropy

$$H_{(x-1)^2}(\rho,\gamma) = \mathrm{Tr}\,(\rho-\gamma)^2\gamma^{-1} = \mathrm{Tr}\,\rho^2\gamma^{-1} - 1. \tag{23}$$

The function $g_{\max}(x) \equiv (x-1)^2(1+x)/2x$ in $\mathcal{G}_{\mathrm{sym}}$ is the symmetrization of g, which corresponds to the largest function $\kappa_{\max}$ in $\mathcal{K}$. The function $\kappa_{\max}$ defines the largest monotone metric with

$$\Omega_\rho^{\max}(X) = \frac{L_\rho^{-1} + R_\rho^{-1}}{2}(X) = \frac{\rho^{-1}X + X\rho^{-1}}{2}.$$

Note that for the choice $X = \rho - \gamma$,

$$\big\langle \rho - \sigma, \Omega_\rho^{\max}(\rho - \sigma)\big\rangle = H_{(x-1)^2}(\gamma, \rho). \tag{24}$$

Example 3. (Central power metric) The function $x^{-1/2} \in \mathcal{K}$ gives

$$\Omega_\rho^{x^{-1/2}}(X) = \rho^{-1/2} X \rho^{-1/2},$$

which may be considered as the center of $\mathcal{K}$ from some aspects. For instance, it is known [22, Theorem 3.5] that $x^{-1/2}$ is the only function $\kappa \in \mathcal{K}$ such that both Ω_ρ^κ and $(\Omega_\rho^\kappa)^{-1}$ are CP for all $\rho \in \mathcal{D}_d$. The function in $\mathcal{G}_{\mathrm{sym}}$ corresponding to $x^{-1/2}$ is $(x-1)^2 x^{-1/2}$ and the corresponding divergence is

$$H_{(x-1)^2 x^{-1/2}}(\rho, \gamma) = \mathrm{Tr}\,(\rho - \gamma)\rho^{-1/2}(\rho - \gamma)\gamma^{-1/2}.$$

Example 4. (Wigner-Yanase-Dyson metric) For any $t \in (0, 1) \cup (1, 2]$ define

$$g^{(t)}(x) \equiv \frac{x - x^t}{t(1-t)} \in \mathcal{G}, \tag{25}$$

whose symmetrized function corresponds to

$$\kappa_t^{\mathrm{WYD}}(x) \equiv \frac{1}{t(1-t)} \cdot \frac{(1-x^t)(1-x^{1-t})}{(1-x)^2} \in \mathcal{K}. \tag{26}$$

Note that the functions κ_t^{WYD} extend to the parameter $t \in [-1, 2]$ with $\kappa_t^{\mathrm{WYD}} = \kappa_{\mathrm{BKM}}$ for $t = 0, 1$ by taking the limit as $t \to 0, 1$ and with $\kappa_t^{\mathrm{WYD}} = \kappa_{1-t}^{\mathrm{WYD}}$ as symmetric around $t = 1/2$. For the particular case $t = 1/2$ the monotone metric associated with $\kappa_{\mathrm{WY}} \equiv \kappa_{1/2}^{\mathrm{WYD}}$ is called the *Wigner-Yanase metric* with

$$\Omega_\rho^{\mathrm{WY}} = \frac{4}{\big(\sqrt{L_\rho} + \sqrt{R_\rho}\big)^2},$$

and the divergence for $g^{(1/2)}(x) = 4(x - \sqrt{x})$ is

$$H_{4(x-\sqrt{x})}(\rho, \gamma) = 4(1 - \mathrm{Tr}\,\rho^{1/2}\gamma^{1/2}).$$

It seems that Hasegawa [18] was the first to realize the WYD metric as well as the WYD divergences could be extended to the full parameter range $[-1, 2]$. See also [29] where equality conditions were given for the convexity of g-divergences for the WYD functions.

It is known [17, Theorem 5.4] that for every $\rho, \gamma \in \mathcal{D}_d$ the geodesic distance $D_{\mathrm{WY}}(\rho, \gamma)$ with respect to the metric for κ_{WY} is given as

$$D_{\mathrm{WY}}(\rho, \gamma) = \arccos \mathrm{Tr}\,\rho^{1/2}\gamma^{1/2}. \tag{27}$$

On the other hand, the geodesic distance $\widetilde{D}_{\mathrm{WY}}(\rho,\gamma)$ taken over curves in $\mathbb{P}_d$ (see Remark 2.4) is included in [24, Theorem 2.1] and we have

$$\widetilde{D}_{\mathrm{WY}}(\rho,\gamma)=\left\|\rho^{1/2}-\gamma^{1/2}\right\|_2=\sqrt{2-2\mathrm{Tr}\,\rho^{1/2}\gamma^{1/2}}. \tag{28}$$

Since $\sqrt{2-2t}<\arccos t$ for $0\le t<1$, we see that

$$\widetilde{D}_{\mathrm{WY}}(\rho,\gamma)<D_{\mathrm{WY}}(\rho,\gamma) \tag{29}$$

unless $\rho=\gamma$.

Example 5. (Minimal or Bures metric) The smallest function $\kappa_{\min}(x)=2/(1+x)$ in $\mathcal{K}$ defines the smallest monotone metric with

$$\Omega_\rho^{\min}(X)=\frac{2}{L_\rho+R_\rho}(X)=2\int_0^\infty e^{-t\rho}Xe^{-t\rho}\,dt, \tag{30}$$

which is often called the *SLD metric* (symmetric logarithmic derivative). This is considered as the infinitesimal form of the *Bures distance* introduced in [7] and was intensively studied by Uhlmann, e.g., [52], so it is also called the *Bures* or *Bures-Uhlmann metric*. The corresponding function in $\mathcal{G}_{\mathrm{sym}}$ is $g_{\min}(x)\equiv 2(x-1)^2/(x+1)$ and the corresponding divergence is

$$H_{\min}(\rho,\gamma)=\left\langle\rho-\gamma,\frac{2}{L_\rho+R_\gamma}(\rho-\gamma)\right\rangle.$$

Recall that the Bures distance [7] between $\rho,\gamma\in\overline{\mathcal{D}}_d$ is

$$d_{\mathrm{Bures}}(\rho,\gamma)\equiv\sqrt{2-2F(\rho,\gamma)}, \tag{31}$$

where $F(\rho,\gamma)\equiv\mathrm{Tr}\,(\rho^{1/2}\gamma\rho^{1/2})^{1/2}$ is the *fidelity* of ρ,γ. It is known (see [53] and [4, (9.32)]) that the geodesic distance between $\rho,\gamma\in\mathcal{D}_d$ with respect to the Bures metric is given as

$$D_{\min}(\rho,\gamma)=\arccos F(\rho,\gamma). \tag{32}$$

Since $\mathrm{Tr}\,\rho^{1/2}\gamma^{1/2}<F(\rho,\gamma)$ unless $\rho\gamma=\gamma\rho$ (see [15, Corollary 3.4] and [20, Theorem 2.1]), by comparing (27) and (32) one can see that $D_{\mathrm{WY}}(\rho,\gamma)>D_{\min}(\rho,\gamma)$ whenever $\rho\gamma\ne\gamma\rho$.

4 Trace Distance and Eigenvalue Formulation

4.1 Trace distance

The most widely used distance for density matrices is the trace-norm distance $\|\rho-\gamma\|_1=\mathrm{Tr}\,|\rho-\gamma|$. The next trace-norm monotonicity property is a slight extension of [48, Theorem 1]. We give a proof for completeness.

Proposition 4.1. *Let $\Phi:\mathbb{M}_d\to\mathbb{M}_{d'}$ be a positive (not necessarily CP) trace-preserving map. Then*

$$\|\Phi(X)\|_1\le\|X\|_1,\qquad X\in\mathbb{M}_d. \tag{33}$$

Proof. Note that the adjoint map $\widehat{\Phi}:\mathbb{M}_{d'}\to\mathbb{M}_d$ is positive and unital. Due to the Russo-Dye theorem (see, e.g., [6, Theorem 2.3.7]) we have

$$\|\widehat{\Phi}(Z)\|_\infty \le \|Z\|_\infty, \qquad Z\in\mathbb{M}_{d'},$$

where $\|\cdot\|_\infty$ denotes the operator norm. Since $\|\cdot\|_1$ is the dual norm of $\|\cdot\|_\infty$, we have the asserted inequality as follows:

$$\begin{aligned}\|\Phi(X)\|_1 &= \sup\bigl\{|\langle X,\widehat{\Phi}(Z)\rangle| : Z\in\mathbb{M}_{d'},\ \|Z\|_\infty\le 1\bigr\}\\ &\le \sup\bigl\{|\langle X,Y\rangle| : Y\in\mathbb{M}_d,\ \|Y\|_\infty\le 1\bigr\} = \|X\|_1\end{aligned}$$

for every $X\in\mathbb{M}_d$. □

The *contraction coefficient* of a positive trace-preserving map Φ with respect to the trace-norm distance is defined by

$$\eta^{\mathrm{Dobrushin}}(\Phi)=\eta^{\mathrm{Tr}}(\Phi)\equiv \sup_{\rho,\gamma\in\mathcal{D}_d,\,\rho\ne\gamma}\frac{\|\Phi(\rho)-\Phi(\gamma)\|_1}{\|\rho-\gamma\|_1}=\sup_{A\in\mathbb{H}_d^0,\,A\ne 0}\frac{\|\Phi(A)\|_1}{\|A\|_1}\quad(\le 1), \tag{34}$$

which is the quantum generalization of the classical *Dobrushin coefficient of ergodicity.* It was shown in [48, Theorem 2] that

$$\eta^{\mathrm{Tr}}(\Phi)=\frac{1}{2}\sup\bigl\{\|\Phi(E-F)\|_1 : E,F\in\mathbb{M}_d,\ \text{rank 1 projections},\ E\perp F\bigr\}. \tag{35}$$

Next we extend the notion of *scrambling* column-stochastic matrices in [10] to the matrix algebra setting.

Proposition 4.2. *Let $\Phi:\mathbb{M}_d\to\mathbb{M}_{d'}$ be a positive (not necessarily CP) trace-preserving map. Then the following conditions are equivalent:*

(i) $\eta^{\mathrm{Tr}}(\Phi)<1$*;*

(ii) *for every rank 1 projections $E,F\in\mathbb{M}_d$ with $E\perp F$, $\mathrm{Tr}\,\Phi(E)\Phi(F)>0$;*

(iii) *for every non-zero $A,B\in\overline{\mathbb{P}}_d$, $\mathrm{Tr}\,\Phi(A)\Phi(B)>0$.*

Proof. We denote by Σ the set of pairs (E,F) of rank 1 projections in $\mathbb{M}_d$ with $E\perp F$. Since Σ is compact in $\mathbb{M}_d\times\mathbb{M}_d$, there is an $(E_0,F_0)\in\Sigma$ such that $\|\Phi(E_0)-\Phi(F_0)\|_1=\sup_{(E,F)\in\Sigma}\|\Phi(E)-\Phi(F)\|_1$. Here, note that, for any $\rho,\gamma\in\overline{\mathcal{D}}_d$, $\|\rho-\gamma\|_1=2$ if and only if the support projections of ρ,γ are orthogonal, if and only if $\mathrm{Tr}\,\rho\gamma=0$. Hence (i) ⇔ (ii) is immediate from (35). That (iii) ⇒ (ii) is trivial. To prove that (i) ⇒ (iii), assume (i) and let $A,B\in\overline{\mathbb{P}}_d$ be non-zero. Set $\rho\equiv A/\mathrm{Tr}\,A$ and $\gamma\equiv B/\mathrm{Tr}\,B$; then $\rho,\gamma\in\overline{\mathcal{D}}_d$ and so $\Phi(\rho),\Phi(\gamma)\in\overline{\mathcal{D}}_{d'}$. If $\rho=\gamma$, then $\mathrm{Tr}\,\Phi(A)\Phi(B)>0$ is clear. If $\rho\ne\gamma$, then assumption (i) implies that $\|\Phi(\rho)-\Phi(\gamma)\|_1<\|\rho-\gamma\|_1\le 2$. Hence we have $\mathrm{Tr}\,\Phi(\rho)\Phi(\gamma)>0$, i.e., $\mathrm{Tr}\,\Phi(A)\Phi(B)>0$. □

4.2 Eigenvalue formulation

We here summarize, for the convenience of the reader, an observation in [38] to link the Riemannian metric coefficient with an eigenvalue problem, which is of some interest in its own right. Moreover, it allows us to connect $\eta^{\mathrm{Tr}}(\Phi)$ with $\eta_g^{\mathrm{Riem}}(\Phi)$ in some very special cases, as in Theorem 4.4 below.

Let $\rho \in \mathcal{D}_d$ and $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$ be a CPT map. Define a linear map $\Psi : \mathbb{M}_d \to \mathbb{M}_{d'}$ by

$$\Psi = \Psi^{\kappa}_{\Phi,\rho} \equiv (\Omega^{\kappa}_{\Phi(\rho)})^{1/2} \circ \Phi \circ (\Omega^{\kappa}_{\rho})^{-1/2}.$$

Note that Ψ is a contraction since

$$\begin{aligned} \langle \Psi(X), \Psi(X)\rangle &= \big\langle \Phi[(\Omega^{\kappa}_{\rho})^{-1/2}(X)], \Omega^{\kappa}_{\Phi(\rho)}\Phi[(\Omega^{\kappa}_{\rho})^{-1/2}(X)]\big\rangle \\ &\le \big\langle (\Omega^{\kappa}_{\rho})^{-1/2}(X), \Omega^{\kappa}_{\rho}(\Omega^{\kappa}_{\rho})^{-1/2}(X)\big\rangle = \langle X, X\rangle, \qquad X \in \mathbb{M}_d, \end{aligned}$$

where we have used (18) for general $A \in \mathbb{M}_d$ (see Remark 2.3). Since the adjoint of a trace-preserving map is unital, we also find

$$\begin{aligned} \widehat{\Psi} \circ \Psi\big[(\Omega^{\kappa}_{\rho})^{-1/2}(I_d)\big] &= (\Omega^{\kappa}_{\rho})^{-1/2} \circ \widehat{\Phi} \circ \Omega^{\kappa}_{\Phi(\rho)}[\Phi(\rho)] \\ &= (\Omega^{\kappa}_{\rho})^{-1/2}(I_d), \end{aligned}$$

where we have used the facts $(\Omega^{\kappa}_{\rho})^{-1}(I_d) = \rho$ and $\Omega^{\kappa}_{\Phi(\rho)}[\Phi(\rho)] = I_d$. Therefore, $(\Omega^{\kappa}_{\rho})^{-1/2}(I_d)$ is an eigenvector of $\widehat{\Psi}\Psi$ corresponding to the largest eigenvalue 1. Let $\lambda_2^{\kappa}(\Phi, \rho)$ denote the second largest eigenvalue (with multiplicities counted) of $\widehat{\Psi}\Psi$. Then $\lambda_2^{\kappa}(\Phi, \rho)$ is represented as

$$\lambda_2^{\kappa}(\Phi, \rho) = \sup_{X \in \mathbb{M}_d^0,\, X \neq 0} \frac{\big\langle \Phi(X), \Omega^{\kappa}_{\Phi(\rho)}(\Phi(X))\big\rangle}{\big\langle X, \Omega^{\kappa}_{\rho}(X)\big\rangle}, \tag{36}$$

where $\mathbb{M}_d^0 \equiv \{X \in \mathbb{M}_d : \mathrm{Tr}\, X = 0\}$. From the fact that $\Omega^{\kappa}_{\rho}(A) \in \mathbb{H}_d$ and $\langle A, \Omega^{\kappa}_{\rho}(B)\rangle = \langle B, \Omega^{\kappa}_{\rho}(A)\rangle$ for all $A, B \in \mathbb{H}_d$, one can easily see that the right-hand side of (36) coincides with $\eta_{\kappa}^{\mathrm{Riem}}(\Phi)$. Therefore, we have

Theorem 4.3. [38, Theorem IV.4] *For every $\kappa \in \mathcal{K}$ and every CPT map Φ,*

$$\eta_{\kappa}^{\mathrm{Riem}}(\Phi) = \sup_{\rho \in \mathcal{D}_d} \lambda_2^{\kappa}(\Phi, \rho).$$

Now assume that an eigenvector of $\widehat{\Psi}\Psi$ corresponding to $\lambda_2 \equiv \lambda_2^{\kappa}(\Phi, \rho)$ is given by $(\Omega^{\kappa}_{\rho})^{1/2}(X)$, orthogonal to $(\Omega^{\kappa}_{\rho})^{-1/2}(I_d)$. Then $\mathrm{Tr}\, X = 0$ and $\widehat{\Psi}\Psi(\Omega^{\kappa}_{\rho})^{1/2}(X) = \lambda_2(\Omega^{\kappa}_{\rho})^{1/2}(X)$, which is equivalently written as

$$(\Omega^{\kappa}_{\rho})^{-1}\widehat{\Phi}\Omega^{\kappa}_{\Phi(\rho)}(\Phi(X)) = \lambda_2 X. \tag{37}$$

Thus, finding $\lambda_2^{\kappa}(\Phi, \rho)$ is equivalent to solving the eigenvalue problem (37) under the constraint $\mathrm{Tr}\, X = 0$. In connection with (37) we define a linear map $\Upsilon : \mathbb{M}_{d'} \to \mathbb{M}_d$ by

$$\Upsilon = \Upsilon^{\kappa}_{\Phi,\rho} \equiv (\Omega^{\kappa}_{\rho})^{-1}\widehat{\Phi}\Omega^{\kappa}_{\Phi(\rho)}.$$

Since $\widehat{\Upsilon}(I_d) = I_{d'}$, Υ is trace-preserving. Here, assume that Υ is positive in the sense that $\Upsilon(\mathbb{P}_{d'}) \in \overline{\mathbb{P}}_d$. Then, thanks to (33), $\|\Upsilon(Z)\|_1 \le \|Z\|_1$ for all $Z \in \mathbb{M}_{d'}$. Therefore, if $X \in \mathbb{M}_d^0$ is a solution of the eigenvalue equation (37), then we have $\lambda_2\|X\|_1 = \|\Upsilon(\Phi(X))\|_1 \le \|\Phi(X)\|_1$ so that $\lambda_2 \le \|\Phi(X)\|_1/\|X\|_1$. Since all linear maps involving in (37) are self-adjoint, note that a solution X of (37) with $\operatorname{Tr} X = 0$ can always be taken in $\mathbb{H}_d^0$.

From the above argument, if Υ is positive for every $\rho \in \mathcal{D}_d$, then we would have $\eta_\kappa^{\mathrm{Riem}}(\Phi) \le \eta^{\mathrm{Tr}}(\Phi)$. This situation indeed occurs, in particular, when both Ω_ρ^κ and $(\Omega_\rho^\kappa)^{-1}$ are positive (equivalently, CP) for every $\rho \in \mathcal{D}_d$ and every $d \in \mathbb{N}$. But it is known [22, Proposition 3.5] that this latter condition holds only when $\kappa(x) = x^{-1/2}$. So we have

Theorem 4.4. [50, Theorem 14] *For every CPT map Φ,*

$$\eta_{x^{-1/2}}^{\mathrm{Riem}}(\Phi) \le \eta^{\mathrm{Tr}}(\Phi).$$

5 General Contraction Results

5.1 Results for arbitrary channels

In this subsection we present a few general relations between the contraction coefficients defined in Section 2. The next theorem says the general equality between the Riemannian metric contraction coefficient and the geodesic contraction coefficient. The proof is based on a limit formula in [24] for the geodesic distance, whose proof is presented in Appendix A for completeness.

11 February 2025 16:32:16

Theorem 5.1. *For every $\kappa \in \mathcal{K}$ and every CPT map $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$,*

$$\eta_\kappa^{\mathrm{Riem}}(\Phi) = \eta_\kappa^{\mathrm{geod}}(\Phi).$$

Proof. The inequality $\eta_\kappa^{\mathrm{geod}}(\Phi) \le \eta_\kappa^{\mathrm{Riem}}(\Phi)$ was shown in [38, Theorem IV.2]. To prove the reverse inequality, we use Lemma A.1 in the appendix. For every $\rho \in \mathcal{D}_d$ and every $A \in \mathbb{H}_d^0$ with $A \ne 0$, by the lemma we have

$$\frac{\langle \Phi(A), \Omega_{\Phi(\rho)}^\kappa(\Phi(A)) \rangle}{\langle A, \Omega_\rho^\kappa(A) \rangle} = \lim_{\varepsilon \searrow 0} \left[\frac{D_\kappa(\Phi(\rho), \Phi(\rho + \varepsilon A))}{D_\kappa(\rho, \rho + \varepsilon A)} \right]^2 \le \eta_\kappa^{\mathrm{geod}}(\Phi),$$

which implies that $\eta_\kappa^{\mathrm{Riem}}(\Phi) \le \eta_\kappa^{\mathrm{geod}}(\Phi)$. □

For completeness we state the following theorem. The first inequality was proved in [38], and the rest is a straightforward consequence of (14) as discussed earlier.

Theorem 5.2. [38, Theorem IV.2] *For every $g \in \mathcal{G}$ let g_{sym} be the symmetrization of g as in* (6) *and $\kappa(x) \equiv g_{\mathrm{sym}}(x)/(x-1)^2 \in \mathcal{K}$ as in* (5)*. Then, for every CPT map $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$,*

$$\eta_\kappa^{\mathrm{Riem}}(\Phi) \le \eta_{g_{\mathrm{sym}}}^{\mathrm{RelEnt}}(\Phi) \le \eta_g^{\mathrm{RelEnt}}(\Phi) = \eta_{\tilde{g}}^{\mathrm{RelEnt}}(\Phi).$$

In the next theorem we give a general inequality between the contraction coefficients for Riemannian metrics and the trace-norm, generalizing [48, Theorem 3] and [50, Theorem 13].

Theorem 5.3. *For every $\kappa \in \mathcal{K}$ and every CPT map $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$,*

$$\eta^{\mathrm{Tr}}(\Phi) \le \sqrt{\eta_\kappa^{\mathrm{Riem}}(\Phi)}.$$

To prove this result, we first give a lemma generalizing [50, Lemma 5].

Lemma 5.4. *For every $\kappa \in \mathcal{K}$ and every $\rho \in \mathcal{D}_d$,*

$$\|A\|_1^2 \le \langle A, \Omega_\rho^\kappa(A)\rangle, \qquad A \in \mathbb{H}_d^0.$$

Proof. Let $\mathcal{E}$ be the trace-preserving conditional expectation from $\mathbb{M}_d$ onto the subalgebra generated by A. The monotonicity of M^κ implies that

$$\langle A, \Omega_\rho^\kappa(A)\rangle \ge \big\langle A, \Omega_{\mathcal{E}(\rho)}^\kappa(A)\big\rangle = \mathrm{Tr}\,\mathcal{E}(\rho)^{-1}A^2,$$

where the latter equality follows since $\mathcal{E}(\rho)$ and A commute. By the Schwarz inequality we have

$$\begin{aligned}\|A\|_1^2 = (\mathrm{Tr}\,|A|)^2 &= \big(\mathrm{Tr}\,\mathcal{E}(\rho)^{1/2}\cdot\mathcal{E}(\rho)^{-1/2}|A|\big)^2 \\ &\le \mathrm{Tr}\,\mathcal{E}(\rho)\cdot\mathrm{Tr}\,\mathcal{E}(\rho)^{-1}A^2 = \mathrm{Tr}\,\mathcal{E}(\rho)^{-1}A^2.\end{aligned}$$

Therefore, $\|A\|_1^2 \le \langle A, \Omega_\rho^\kappa(A)\rangle$. □

Proof of Theorem 5.3. Let $A \in \mathbb{H}_d^0$ with $A \neq 0$ and assume further that A is invertible. Set $\rho \equiv |A|/\|A\|_1 \in \mathcal{D}_d$. By the above lemma we have

$$\|\Phi(A)\|_1^2 \le \big\langle \Phi(A), \Omega_{\Phi(\rho)}^\kappa(\Phi(A))\big\rangle.$$

On the other hand, since ρ and A commute, we have

$$\langle A, \Omega_\rho^\kappa(A)\rangle = \mathrm{Tr}\,\rho^{-1}A^2 = \|A\|_1\mathrm{Tr}\,|A| = \|A\|_1^2.$$

Therefore,

$$\frac{\|\Phi(A)\|_1^2}{\|A\|_1^2} \le \frac{\big\langle \Phi(A), \Omega_{\Phi(\rho)}^\kappa(\Phi(A))\big\rangle}{\langle A, \Omega_\rho^\kappa(A)\rangle} \le \eta_\kappa^{\mathrm{Riem}}(\Phi).$$

By continuity we have $\|\Phi(A)\|_1^2/\|A\|_1^2 \le \eta_\kappa^{\mathrm{Riem}}(\Phi)$ for all $A \in \mathbb{H}_d^0$ with $A \neq 0$, proving the desired inequality. □

5.2 QC and CQ channels

Let $\Phi : \mathbb{M}_d \to \mathbb{M}_{d'}$ be a positive trace-preserving map. We call Φ a *quantum-classical* (QC) channel if the range of Φ is included in a commutative subalgebra of $\mathbb{M}_{d'}$, and a *classical-quantum* (CQ) channel if the range of $\widehat{\Phi}$ is in a commutative subalgebra of $\mathbb{M}_d$. Note that if Φ is QC or CQ, then positivity is the same as CP, so it is indeed a channel. The following facts are easy to see: Φ is a QC channel if and only if there are an orthonormal basis $\{\psi_k\}_{k=1}^{d'}$ of $\mathbf{C}^{d'}$ and a POVM $\{F_k\}_{k=1}^{d'}$ in $\mathbb{M}_d$ such that

$$\Phi(\rho) = \sum_k (\mathrm{Tr}\,F_k\rho)|\psi_k\rangle\langle\psi_k|.$$

Also, Φ is a CQ channel if and only if there are an orthonormal basis $\{\phi_k\}_{k=1}^{d}$ of $\mathbf{C}^d$ and density matrices $\gamma_k \in \mathcal{D}_{d'}$, $1 \le k \le d$, such that

$$\Phi(\rho) = \sum_k \langle\phi_k, \rho\phi_k\rangle\gamma_k.$$

Thus, our notions of QC and CQ channels coincide with those introduced in [27]. We note that both QC and CQ channels have the form $\Phi(\rho)=\sum_k(\operatorname{Tr}F_k\rho)\gamma_k$ with a POVM $\{F_k\}$ in $\mathbb{M}_d$ and $\gamma_k\in\mathcal{D}_{d'}$, introduced in [27] and shown in [28] to be *entanglement breaking*. Several equivalent characterizations of this class were also given in [28].

We remark that when Φ is purely classical, i.e., the ranges of Φ and $\widehat{\Phi}$ are in commutative subalgebras of $\mathbb{M}_{d'}$ and $\mathbb{M}_d$, respectively, then Φ can be represented by a $d'\times d$ column-stochastic matrix in some orthonormal bases of $\mathbf{C}^d$ and $\mathbf{C}^{d'}$. Thus [9, Theorem 1] implies that

$$\eta_\kappa^{\mathrm{Riem}}(\Phi)=\eta_g^{\mathrm{RelEnt}}(\Phi)$$

for any independent choices of $\kappa\in\mathcal{K}$ and $g\in\mathcal{G}$.

Proposition 5.5. *Let $\kappa_1,\kappa_2\in\mathcal{K}$ and assume that $\kappa_1(x)\le\kappa_2(x)$ for all $x>0$. Then $\eta_{\kappa_1}^{\mathrm{Riem}}(\Phi)\ge\eta_{\kappa_2}^{\mathrm{Riem}}(\Phi)$ for every QC channel Φ, and $\eta_{\kappa_1}^{\mathrm{Riem}}(\Phi)\le\eta_{\kappa_2}^{\mathrm{Riem}}(\Phi)$ for every CQ channel Φ.*

Proof. The assumption $\kappa_1\le\kappa_2$ implies that $\Omega_\rho^{\kappa_1}\le\Omega_\rho^{\kappa_2}$ as operators on the Hilbert space $\mathbb{M}_d$ for every $\rho\in\mathcal{D}_d$. When Φ is QC, we have

$$\eta_\kappa^{\mathrm{Riem}}(\Phi)=\sup_{\rho\in\mathcal{D}_d}\ \sup_{A\in\mathbb{H}_d^0,\,A\ne0}\frac{\operatorname{Tr}\Phi(\rho)^{-1}\Phi(A)^2}{\langle A,\Omega_\rho^\kappa(A)\rangle}$$

for every $\kappa\in\mathcal{K}$. Hence $\eta_{\kappa_1}^{\mathrm{Riem}}(\Phi)\ge\eta_{\kappa_2}^{\mathrm{Riem}}(\Phi)$. When Φ is CQ, choose a subalgebra $\mathcal{A}$ of $\mathbb{M}_d$ including the range of $\widehat{\Phi}$, and let $\mathcal{E}:\mathbb{M}_d\to\mathcal{A}$ be the trace-preserving conditional expectation. Since $\widehat{\mathcal{E}}$ is nothing but the inclusion $\mathcal{A}\hookrightarrow\mathbb{M}_d$, we have $\widehat{\Phi}=\widehat{\mathcal{E}}\widehat{\Phi}$ so that $\Phi=\Phi\mathcal{E}$. Therefore, we have

$$\eta_\kappa^{\mathrm{Riem}}(\Phi)=\sup_{\rho\in\mathcal{D}_d\cap\mathcal{A}}\ \sup_{A\in\mathbb{H}_d^0\cap\mathcal{A},\,A\ne0}\frac{\langle\Phi(A),\Omega_{\Phi(\rho)}^\kappa(\Phi(A))\rangle}{\operatorname{Tr}\rho^{-1}A^2}$$

for every $\kappa\in\mathcal{K}$. Hence $\eta_{\kappa_1}^{\mathrm{Riem}}(\Phi)\le\eta_{\kappa_2}^{\mathrm{Riem}}(\Phi)$. □

Thanks to (11), by Proposition 5.5 and Theorem 4.4 we have

Corollary 5.6. *If Φ is a QC channel and $\kappa\in\mathcal{K}$ satisfies $\kappa(x)\ge x^{-1/2}$ for all $x>0$, then*

$$\eta_{\max}^{\mathrm{Riem}}(\Phi)\le\eta_\kappa^{\mathrm{Riem}}(\Phi)\le\eta_{x^{-1/2}}^{\mathrm{Riem}}(\Phi)\le\eta^{\mathrm{Tr}}(\Phi).$$

If Φ is a CQ channel and $\kappa\in\mathcal{K}$ satisfies $\kappa(x)\le x^{-1/2}$ for all $x>0$, then

$$\eta_{\min}^{\mathrm{Riem}}(\Phi)\le\eta_\kappa^{\mathrm{Riem}}(\Phi)\le\eta_{x^{-1/2}}^{\mathrm{Riem}}(\Phi)\le\eta^{\mathrm{Tr}}(\Phi).$$

The next theorem shows a relation of $\eta_g^{\mathrm{RelEnt}}(\Phi)$ for any $g\in\mathcal{G}_{\mathrm{sym}}$ with the coefficient $\eta_{\min}^{\mathrm{Riem}}(\Phi)$ with respect to the minimal metric when Φ is a QC channel. The proof is based on the integral decomposition (8).

Theorem 5.7. *Assume that $\Phi:\mathbb{M}_d\to\mathbb{M}_{d'}$ is a QC channel. Then for every $\kappa\in\mathcal{K}$ and the corresponding $g\in\mathcal{G}_{\mathrm{sym}}$ in (5),*

$$\eta_\kappa^{\mathrm{Riem}}(\Phi)\le\eta_g^{\mathrm{RelEnt}}(\Phi)\le\eta_{\min}^{\mathrm{Riem}}(\Phi),$$

and in particular,

$$\eta_{\min}^{\mathrm{Riem}}(\Phi)=\eta_{g_{\min}}^{\mathrm{RelEnt}}(\Phi)\tag{38}$$

for $g_{\min}\in\mathcal{G}_{\mathrm{sym}}$ given in Example 5.

Proof. The first inequality holds for a general CPT map Φ due to Theorem 5.2. Now we assume that Φ is a QC channel and let $\kappa \in \mathcal{K}$ be arbitrary. Since κ admits an integral expression (or the extremal decomposition) in (8) so that we write

$$\kappa(x) = \int_{[0,1]} \kappa_s(x)\, dm(s), \qquad x \in (0,\infty)$$

with $\kappa_s \in \mathcal{K}$, $0 \le s \le 1$, given in (10). Moreover, for $0 \le s \le 1$ let

$$g_s(x) \equiv \frac{(x-1)^2}{x+s} \in \mathcal{G}, \tag{39}$$

whose symmetrization $(g_s)_{\rm sym} \in \mathcal{G}_{\rm sym}$ corresponds in (5) to κ_s, i.e., $(g_s)_{\rm sym}(x) = (x-1)^2\kappa_s(x)$. For every $\rho, \gamma \in \mathcal{D}_d$ we then have

$$H_g(\rho,\gamma) = \int_{[0,1]} H_{(g_s)_{\rm sym}}(\rho,\gamma)\, dm(s).$$

So it suffices to prove that

$$H_{(g_s)_{\rm sym}}(\Phi(\rho),\Phi(\gamma)) \le \eta_{\min}^{\rm Riem}(\Phi) H_{(g_s)_{\rm sym}}(\rho,\gamma), \qquad 0 \le s \le 1. \tag{40}$$

Since

$$\begin{aligned} R_\gamma^{-1}\kappa_s(L_\rho R_\gamma^{-1}) &= \frac{1+s}{2}\, R_\gamma^{-1}\biggl(\frac{1}{L_\rho R_\gamma^{-1}+s} + \frac{1}{sL_\rho R_\gamma^{-1}+1}\biggr) \\ &= \frac{1+s}{2}\biggl(\frac{1}{L_\rho + sR_\gamma} + \frac{1}{sL_\rho + R_\gamma}\biggr), \end{aligned} \tag{41}$$

we have

$$\begin{aligned} &\frac{R_\gamma^{-1}\kappa_s(L_\rho R_\gamma^{-1}) + R_\rho^{-1}\kappa_s(L_\gamma R_\rho^{-1})}{2} \\ &\quad = \frac{1+s}{2}\biggl\{\frac{(L_\rho+sR_\gamma)^{-1} + (sL_\gamma+R_\rho)^{-1}}{2} + \frac{(sL_\rho+R_\gamma)^{-1} + (L_\gamma+sR_\rho)^{-1}}{2}\biggr\} \\ &\quad \ge \frac{1+s}{2}\biggl\{\biggl(\frac{L_\rho+sR_\gamma+sL_\gamma+R_\rho}{2}\biggr)^{-1} + \biggl(\frac{sL_\rho+R_\gamma+L_\gamma+sR_\rho}{2}\biggr)^{-1}\biggr\} \\ &\quad = \frac{1}{2}\biggl\{\biggl(\frac{L_{\frac{\rho+s\gamma}{1+s}} + R_{\frac{\rho+s\gamma}{1+s}}}{2}\biggr)^{-1} + \biggl(\frac{L_{\frac{\gamma+s\rho}{1+s}} + R_{\frac{\gamma+s\rho}{1+s}}}{2}\biggr)^{-1}\biggr\}. \end{aligned}$$

In the above the operator convexity of x^{-1} on $(0,\infty)$ has been used. We then have

$$\begin{aligned} H_{(g_s)_{\rm sym}}(\rho,\gamma) &= \frac{H_{(g_s)_{\rm sym}}(\rho,\gamma) + H_{(g_s)_{\rm sym}}(\gamma,\rho)}{2} \\ &= \biggl\langle \rho-\gamma, \biggl\{\frac{R_\gamma^{-1}\kappa_s(L_\rho R_\gamma^{-1}) + R_\rho^{-1}\kappa_s(L_\gamma R_\rho^{-1})}{2}\biggr\}(\rho-\gamma)\biggr\rangle \\ &\ge \frac{1}{2}\biggl\langle (\rho-\gamma, \biggl\{\biggl(\frac{L_{\frac{\rho+s\gamma}{1+s}} + R_{\frac{\rho+s\gamma}{1+s}}}{2}\biggr)^{-1} + \biggl(\frac{L_{\frac{\gamma+s\rho}{1+s}} + R_{\frac{\gamma+s\rho}{1+s}}}{2}\biggr)^{-1}\biggr\}(\rho-\gamma)\biggr\rangle \\ &= \frac{1}{2}\biggl\langle \rho-\gamma, \biggl\{\Omega^{\min}_{\frac{\rho+s\gamma}{1+s}} + \Omega^{\min}_{\frac{\gamma+s\rho}{1+s}}\biggr\}(\rho-\gamma)\biggr\rangle \end{aligned} \tag{42}$$

thanks to (30).

Now we consider the case where ρ and γ commute. Since ρ and γ commute with $\rho-\gamma$, it is easy to see from (41) that

$$\begin{aligned} H_{(g_s)_{\mathrm{sym}}}(\rho,\gamma) &= \langle \rho-\gamma, R_\gamma^{-1}\kappa_s(L_\rho R_\gamma^{-1})(\rho-\gamma)\rangle \\ &= \frac{1+s}{2}\langle \rho-\gamma, \{(\rho+s\gamma)^{-1}+(\gamma+s\rho)^{-1}\}(\rho-\gamma)\rangle \\ &= \frac{1}{2}\left\langle \rho-\gamma, \left\{\Omega^{\min}_{\frac{\rho+s\gamma}{1+s}}+\Omega^{\min}_{\frac{\gamma+s\rho}{1+s}}\right\}(\rho-\gamma)\right\rangle. \end{aligned}$$

Since Φ has the commutative range, we can apply the above to $\Phi(\rho)$ and $\Phi(\gamma)$ to obtain

$$H_{(g_s)_{\mathrm{sym}}}(\Phi(\rho),\Phi(\gamma)) = \frac{1}{2}\left\langle \Phi(\rho-\gamma), \left\{\Omega^{\min}_{\Phi\left(\frac{\rho+s\gamma}{1+s}\right)}+\Omega^{\min}_{\Phi\left(\frac{\gamma+s\rho}{1+s}\right)}\right\}\Phi(\rho-\gamma)\right\rangle. \tag{43}$$

Hence (40) follows from (42) and (43). □

5.3 Weak Schwarz maps

Although we have restricted our consideration to quantum channels in the usual sense of CPT maps, the monotonicity property of g-divergences and monotone metrics holds more generally under a positive trace-preserving map $\Phi:\mathbb{M}_d\to\mathbb{M}_{d'}$ whose adjoint $\widehat{\Phi}$ is a *weak* Schwarz map in the sense that $\widehat{\Phi}(Y^*)\widehat{\Phi}(Y)\le\widehat{\Phi}(Y^*Y)$ for all $Y\in\mathbb{M}_{d'}$. The proofs of monotonicity of g-divergences in [44, 45] as well as the argument in [46] for monotone metrics requires only this weaker condition, as discussed further in [23]. Thus, we can define the contraction coefficients for such maps Φ rather than CPT maps, and all the results in the paper extend to this slightly more general situation. In Section 4.1 we consider the weaker condition of positive trace-preserving maps and in Theorem 6.1 positive maps which are unital and trace-preserving.

In the above comment we have introduced the term "weak Schwarz map" for the following reason. For a positive linear functional ϕ on an operator algebra, the Schwarz inequality can be written as $|\phi(A^*B)|^2\le\phi(A^*A)\phi(B^*B)$. The analogous result for a linear map Φ on an operator algebra is the *operator* inequality

$$\Phi(A^*B)[\Phi(B^*B)]^{-1}\Phi(B^*A)\le\Phi(A^*A) \tag{44}$$

first proved for CP maps by Lieb and Ruskai [40] in 1974. (In finite dimensions, an equivalent inequality was proved much earlier by Kiefer [33] in 1959.) In 1980 Choi [8] showed that (44) holds if and only if Φ is 2-positive. Thus, it would be natural to consider the Schwarz maps as precisely the class of 2-positive maps. However, earlier in 1952 Kadison [30] proved a special case of (44) with $A=A^*$ and $B=I$, and it was later found that the condition $A=A^*$ could be dropped in many situations. Thus, the terms "Schwarz inequality" and "Schwarz map" were associated with the weaker inequality $\Phi(A^*)\Phi(A)\le\Phi(A^*A)$. However, this inequality does not hold for arbitrary positive linear maps and we know of no characterization of the subclass for which it holds other than the inequality itself.

6 Qubit Channels

We now consider some special qubit channels. Some of the results here were stated without proof at the end of [38]. Others are new and resolve conjectures discussed elsewhere in the paper [38].

We first recall the description of $\mathcal{D}_2$ as the Bloch ball briefly, see, e.g., [43, 47, 49] for more details. Pauli matrices $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ with $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ form an orthogonal basis of the qubit Hilbert space $\mathbb{M}_2$. Any $\rho \in \mathcal{D}_2$ is represented as $\rho = \frac{1}{2}(I + \mathbf{w}\cdot\sigma)$ by a unique $\mathbf{w} = (w_1, w_2, w_3)^t \in \mathbf{R}^3$ with $|\mathbf{w}| \equiv \sqrt{w_1^2 + w_2^2 + w_3^2} \le 1$, where $\mathbf{w}\cdot\sigma \equiv w_1\sigma_1 + w_2\sigma_2 + w_3\sigma_3$. Here, ρ is pure if and only if $\mathbf{w}$ is on the unit sphere, i.e., $|\mathbf{w}| = 1$. A trace-preserving linear map $\Phi : \mathbb{M}_2 \to \mathbb{M}_2$ is represented as

$$\Phi(w_0 I + \mathbf{w}\cdot\sigma) = w_0 I + (w_0\mathbf{t} + T\mathbf{w})\cdot\sigma, \qquad w_0 \in \mathbf{R},\ \mathbf{w} \in \mathbf{R}^3,$$

by a vector $\mathbf{t} \in \mathbf{R}^3$ and a 3×3 real matrix $T = (t_{ij})_{i,j=1}^3$, which, as observed in [34], can be assumed to be diagonal without loss of generality. Clearly, Φ is positive if and only if $|\mathbf{t} + T\mathbf{w}| \le 1$ for all $\mathbf{w} \in \mathbf{R}^3$ with $|\mathbf{w}| \le 1$, and Φ is unital and positive if and only if $\mathbf{t} = 0$ and $\|T\|_\infty \le 1$, where $\|T\|_\infty$ is the operator norm of T. Necessary and sufficient conditions for complete positivity were given in [49]. In the special case when only $t_3 \ne 0$, it was shown earlier by Fujiwara and Algoet [16] that a map of this form is CPT if and only if

$$(\lambda_1 \pm \lambda_2)^2 \le (1 \pm \lambda_3)^2 - t_3^2$$

when λ_k are the diagonal elements of T.

Theorem 6.1. *For any unital map $\Phi_T : I + \mathbf{w}\cdot\sigma \mapsto I + (T\mathbf{w})\cdot\sigma$ where T is a real matrix with $\|T\|_\infty \le 1$,*

$$\eta_\kappa^{\mathrm{Riem}}(\Phi_T) = \eta_\kappa^{\mathrm{geod}}(\Phi_T) = \eta_g^{\mathrm{RelEnt}}(\Phi_T) = \|T\|_\infty^2$$

for every $\kappa \in \mathcal{K}$ and every $g \in \mathcal{G}$. Furthermore, $\eta^{\mathrm{Tr}}(\Phi_T) = \|T\|_\infty$.

This result does not require that the map Φ_T be CP. Note that $\eta^{\mathrm{Tr}}(\Phi_T) = \|T\|_\infty = \sqrt{\eta_\kappa^{\mathrm{Riem}}(\Phi_T)}$, which is consistent with Theorem 5.3.

The next theorem treats a family of trace-preserving maps $\Phi_{\alpha,\tau} : \mathbb{M}_2 \to \mathbb{M}_2$ with two real parameters α, τ determined by $\mathbf{t} = (0, 0, \tau)^t$ and $T = \mathrm{diag}(\alpha, 0, 0)$; more explicitly,

$$\Phi_{\alpha,\tau}(w_0 I + \mathbf{w}\cdot\sigma) = w_0 I + \alpha w_1\sigma_1 + \tau w_0\sigma_3.$$

In this case, the condition $\alpha^2 + \tau^2 \le 1$ is necessary and sufficient for both positivity and complete positivity as shown in [16, 49]. It is also easy to check that the range of the adjoint map $\widehat{\Phi}$ is included in the commutative subalgebra generated by $\{I, \sigma_1\}$. Thus, when $\tau \ne 0$, $\Phi_{\alpha,\tau}$ is a non-unital CQ channel. Below we assume that $\alpha \ge 0$ and $\alpha^2 + \tau^2 \le 1$.

Theorem 6.2. *Let $\Phi \equiv \Phi_{\alpha,\tau}$ be a non-unital CQ channel with α, τ specified above. Then*

$$\eta^{\mathrm{Tr}}(\Phi) = \alpha, \tag{45a}$$

$$\eta_{\max}^{\mathrm{Riem}}(\Phi) = \frac{\alpha^2}{1 - \tau^2}, \tag{45b}$$

$$\eta^{\text{Riem}}_{\widehat{\text{WY}}}(\Phi) \equiv \eta^{\text{Riem}}_{(1+\sqrt{x})^2/4x}(\Phi) \geq \alpha^2 \frac{1+\sqrt{1-\tau^2}}{2(1-\tau^2)}, \tag{45c}$$

$$\eta^{\text{Riem}}_{x^{-1/2}}(\Phi) \geq \frac{\alpha^2}{\sqrt{1-\tau^2}}, \tag{45d}$$

$$\eta^{\text{Riem}}_{\text{BKM}}(\Phi) \equiv \eta^{\text{Riem}}_{(\log x)/(x-1)}(\Phi) \geq \frac{\alpha^2}{2\tau} \log \tfrac{1+\tau}{1-\tau}, \tag{45e}$$

$$\eta^{\text{Riem}}_{\text{WY}}(\Phi) \equiv \eta^{\text{Riem}}_{4/(1+\sqrt{x})^2}(\Phi) = \frac{2\alpha^2}{1+\sqrt{1-\tau^2}}, \tag{45f}$$

$$\eta^{\text{Riem}}_{\min}(\Phi) = \alpha^2. \tag{45g}$$

Moreover, for the extreme points κ_s of $\mathcal{K}$ given in (10),

$$\eta^{\text{Riem}}_{\kappa_s}(\Phi) = \frac{\alpha^2}{1 - \left(\frac{1-s}{1+s}\right)^2 \tau^2}, \qquad 0 \leq s \leq 1. \tag{46}$$

In the above, the function $(1+\sqrt{x})^2/4x \in \mathcal{K}$ is the dual of κ_{WY} in Example 4, i.e., $1/\kappa_{\text{WY}}(x^{-1}) = (1+\sqrt{x})^2/4x$. The identities (45b) and (45g) are of course contained in (46). By Proposition 5.5 (for CQ channels) together with (45b) and (45g) we observe that

$$\alpha^2 \leq \eta^{\text{Riem}}_{\kappa}(\Phi_{\alpha,\tau}) \leq \frac{\alpha^2}{1-\tau^2} \tag{47}$$

for every $\kappa \in \mathcal{K}$.

Although the bounds in the above theorem are sufficient to disprove two conjectures as remarked below, we believe that they are optimal, i.e.,

Conjecture 6.3. *Equality holds in* (45c) *through* (45e) *above.*

In those cases in which we can compute $\eta^{\text{Riem}}_{\kappa}(\Phi_{\alpha,\tau})$ exactly, the supremum is attained when $\rho = \frac{1}{2}I$ or, equivalently, $\mathbf{w} = (0,0,0)^t$ and $A = \mathbf{y}{\cdot}\sigma$ with $\mathbf{y} = (y_1, 0, 0)^t$. Since the output of Φ does not involve w_2, w_3 it is reasonable that there is no loss of generality in choosing $w_2 = w_3 = 0$. And since the channel is symmetric around $w_1 = 0$ or $P = \frac{1}{2}I$, this choice is also reasonable. But a proof for arbitrary choices of κ does not seem easy.

If the above conjecture is true, then for these examples $\kappa_1 \leq \kappa_2$ implies $\eta^{\text{Riem}}_{\kappa_1}(\Phi_{\alpha,\tau}) \leq \eta^{\text{Riem}}_{\kappa_2}(\Phi_{\alpha,\tau})$, which is consistent with Proposition 5.5.

Remark 6.4. It follows from (45a) and (45b) that the conjecture [38] that $\eta^{\text{Riem}}_{g}(\Phi) \leq \eta^{\text{Tr}}(\Phi)$ is false. Indeed, whenever $\alpha > 1-\tau^2$, it follows that

$$\eta^{\text{Riem}}_{\max}(\Phi) = \frac{\alpha^2}{1-\tau^2} > \alpha = \eta^{\text{Tr}}(\Phi). \tag{48}$$

Since $\alpha > \alpha^2$, parameters can be found that are consistent with the CP condition but satisfy (48). In fact, $\alpha = \tau = 1/\sqrt{2}$ will do.

Remark 6.5. Although we do not know when equality holds in the bounds above, we have sufficient information to conclude that the largest contraction coefficient is not necessarily given by $\kappa(x) = x^{-1/2}$ as conjectured in [31]. In particular, when $\alpha^2 + \tau^2 = 1$, the bound $\eta^{\text{Riem}}_{x^{-1/2}}(\Phi) \leq \eta^{\text{Tr}}(\Phi)$ (Theorem 4.4) implies that equality holds in (45d), so $\eta^{\text{Riem}}_{x^{-1/2}}(\Phi) = \alpha <$

$1 = \eta^{\mathrm{Riem}}_{\max}(\Phi)$ in this case. This is foreshadowed by the fact that Proposition 5.5 says that $\eta^{\mathrm{Riem}}_{\kappa}(\Phi)$ is monotone increasing in $\kappa \in \mathcal{K}$ for a CQ channel.

We also note that the inequality $\eta^{\mathrm{Riem}}_{x^{-1/2}} \le \eta^{\mathrm{Tr}}$ implies that $\frac{\alpha^2}{\sqrt{1-\tau^2}} \le \alpha$ or $\frac{\alpha}{\sqrt{1-\tau^2}} \le 1$, which is equivalent to the CP condition $\alpha^2 + \tau^2 \le 1$.

The next theorem disproves the conjecture [38] that $\eta^{\mathrm{Riem}}_{\kappa}(\Phi) = \eta^{\mathrm{RelEnt}}_{g}(\Phi)$ for every $\kappa \in \mathcal{K}$ and the corresponding $g \in \mathcal{G}_{\mathrm{sym}}$. For $0 \le s \le 1$ let $g_s \in \mathcal{G}$ be given as in (39), so $(g_s)_{\mathrm{sym}} \in \mathcal{G}_{\mathrm{sym}}$ corresponds to to the extreme $\kappa_s \in \mathcal{K}$.

Theorem 6.6. *Let $\Phi_{\alpha,\tau}$ be as above. If $4\tau^2 > (1-\alpha^2)(4-\alpha^2)$ (this is the case when $\alpha^2+\tau^2 = 1$ with $\alpha > 0$), then*

$$\eta^{\mathrm{RelEnt}}_{g_s}(\Phi_{\alpha,\tau}) \ge \eta^{\mathrm{RelEnt}}_{(g_s)_{\mathrm{sym}}}(\Phi_{\alpha,\tau}) > \eta^{\mathrm{Riem}}_{\kappa_s}(\Phi_{\alpha,\tau}) \tag{49}$$

for any $s \le 1$ sufficiently near 1 (depending on α, τ). Moreover, if $s > \sqrt{\frac{5-\sqrt{21}}{2}} \approx 0.457$, then $\eta^{\mathrm{RelEnt}}_{(g_s)_{\mathrm{sym}}}(\Phi_{\alpha,\tau}) > \eta^{\mathrm{Riem}}_{\kappa_s}(\Phi_{\alpha,\tau})$ when $\alpha^2 = 1-\tau^2$ is sufficiently small.

Note that $(g_0)_{\mathrm{sym}}$ is $g_{\max}$ given in Example 2 and the three coefficients in (49) are equal in particular when $s = 0$, as will be shown in Proposition 7.2 for general CPT maps.

Theorems 6.1, 6.2 and 6.6 will be proved in the whole Appendix B.

7 Results in Special Cases

7.1 BKM metric

Theorem 7.1. *For every CPT map Φ, $\eta^{\mathrm{Riem}}_{\mathrm{BKM}}(\Phi) = \eta^{\mathrm{RelEnt}}_{\mathrm{BKM}}(\Phi)$.*

Proof. By Theorem 5.2 it suffices to prove that

$$\eta^{\mathrm{RelEnt}}_{\mathrm{BKM}}(\Phi) \le \eta^{\mathrm{Riem}}_{\mathrm{BKM}}(\Phi) \tag{50}$$

for every CPT map Φ. To do this, consider the line segment $\xi(t) \equiv (1-t)\rho + t\gamma = \rho + t(\gamma - \rho)$, $0 \le t \le 1$, joining $\rho, \gamma \in \mathcal{D}_d$. By using Daleckii and Krein's differential formula (see, e.g., [21, Section 2.3]) we compute the derivative

$$\begin{aligned}\frac{d}{dt} H(\rho, \xi(t)) &= \frac{d}{dt} \operatorname{Tr} \rho(\log \rho - \log \xi(t)) = -\operatorname{Tr} \rho \biggl(\frac{d}{dt} \log \xi(t)\biggr) \\ &= -\operatorname{Tr} \rho \log^{[1]}(L_{\xi(t)}, R_{\xi(t)})(\gamma - \rho),\end{aligned}$$

where $\log^{[1]}(x,y) \equiv (\log x - \log y)/(x-y)$, the divided difference of $\log x$. We hence have

$$\frac{d}{dt} H(\rho, \xi(t)) = -\operatorname{Tr} \rho \, \frac{\log L_{\xi(t)} - \log R_{\xi(t)}}{L_{\xi(t)} - R_{\xi(t)}} \, (\gamma - \rho) = -\bigl\langle \rho, \Omega^{\mathrm{BKM}}_{\xi(t)}(\gamma - \rho) \bigr\rangle \tag{51}$$

thanks to (22), and similarly

$$\frac{d}{dt} H(\gamma, \xi(t)) = -\bigl\langle \gamma, \Omega^{\mathrm{BKM}}_{\xi(t)}(\gamma - \rho) \bigr\rangle. \tag{52}$$

Therefore,

$$\frac{d}{dt}\{H(\rho,\xi(t))-H(\gamma,\xi(t))\}=\big\langle\gamma-\rho,\Omega_{\xi(t)}^{\mathrm{BKM}}(\gamma-\rho)\big\rangle$$

so that

$$\begin{aligned}H_{\mathrm{BKM}}(\rho,\gamma)&\equiv H(\rho,\gamma)+H(\gamma,\rho)\\&=\{H(\rho,\xi(1))-H(\gamma,\xi(1))\}-\{H(\rho,\xi(0))-H(\gamma,\xi(0))\}\\&=\int_0^1\frac{d}{dt}\{H(\rho,\xi(t))-H(\gamma,\xi(t))\}\,dt\\&=\int_0^1\big\langle\gamma-\rho,\Omega_{\xi(t)}^{\mathrm{BKM}}(\gamma-\rho)\big\rangle\,dt.\end{aligned}$$

By replacing ρ,γ with $\Phi(\rho),\Phi(\gamma)$ we also have

$$H_{\mathrm{BKM}}(\Phi(\rho),\Phi(\gamma))=\int_0^1\big\langle\Phi(\gamma-\rho),\Omega_{\Phi(\xi(t))}^{\mathrm{BKM}}(\Phi(\gamma-\rho))\big\rangle\,dt.$$

Since

$$\big\langle\Phi(\gamma-\rho),\Omega_{\Phi(\xi(t))}^{\mathrm{BKM}}(\Phi(\gamma-\rho))\big\rangle\le\eta_{\mathrm{BKM}}^{\mathrm{Riem}}(\Phi)\big\langle\gamma-\rho,\Omega_{\xi(t)}^{\mathrm{BKM}}(\gamma-\rho)\big\rangle,\qquad 0\le t\le 1,$$

the desired inequality (50) follows. □

The differential expressions in (51) and (52) are quite special, so it seems that we cannot apply the differential method as above for other $\kappa\in\mathcal{K}$. One may also consider the contraction coefficient defined in (1)

$$\eta_{x\log x}^{\mathrm{RelEnt}}(\Phi)\equiv\sup_{\rho,\gamma\in\mathcal{D}_d,\,\rho\neq\gamma}\frac{H(\Phi(\rho),\Phi(\gamma))}{H(\rho,\gamma)}$$

with respect to the standard (non-symmetrized) relative entropy. One has $\eta_{\mathrm{BKM}}^{\mathrm{RelEnt}}(\Phi)\le\eta_{x\log x}^{\mathrm{RelEnt}}(\Phi)$ by the inequality in (16), but it is unknown whether both contraction coefficients coincide or not.

7.2 Other special results

We collect here some additional special relations that may be of interest.

Since the maximal metric has the special property (24) that every metric $\langle A,\Omega_\rho^{\max}(A)\rangle$ can be realized as a quadratic relative entropy, Theorem 5.2 immediately implies the following:

Proposition 7.2. *For every CPT map Φ,*

$$\eta_{\max}^{\mathrm{Riem}}(\Phi)=\eta_{g_{\max}}^{\mathrm{RelEnt}}(\Phi)=\eta_{(x-1)^2}^{\mathrm{RelEnt}}(\Phi),\tag{53}$$

where $g_{\max}\in\mathcal{G}_{\mathrm{sym}}$ is given in Example 2.

The identities (53) and (38) show some asymmetry between the contraction properties of $\kappa_{\max}$ and $\kappa_{\min}$; (53) holds for all quantum channels while (38) does for only QC channels, see a counter-example ($s=1$) in Theorem 6.6 for a CQ channel.

The functions κ_t^{WYD} given in (26) showed up through the representation of the Wigner-Yanase-Dyson skew information in terms of monotone metrics, as described in [22, Section 2.4, Example 4.8]. Furthermore, as studied in [29], the trace functional of WYD concavity/convexity [39, 2] is recovered by the quasi-entropy for $g^{(t)}$ given in (25) as follows:

$$J_t(K,A,B) \equiv \langle KB^{1/2}, g^{(t)}(L_A R_B^{-1})(KB^{1/2})\rangle = \frac{1}{t(1-t)}\left(\mathrm{Tr}\,K^*AK - \mathrm{Tr}\,K^*A^tKB^{1-t}\right)$$

for $A,B \in \mathbb{P}_d$ with a linear term. The $g^{(t)}$-divergence is

$$H_t(\rho,\gamma) \equiv J_t(I,\rho,\gamma) = \frac{1-\mathrm{Tr}\,\rho^t\gamma^{1-t}}{t(1-t)}.$$

Note that $H(\rho,\gamma) = \lim_{t\to 1} H_t(\rho,\gamma)$ and $H(\gamma,\rho) = \lim_{t\to 0} H_t(\rho,\gamma)$ so that $H_t(\rho,\gamma)$ forms a one-parameter extension of the relative entropy. By Theorem 5.2 we have

Proposition 7.3. *For every CPT map* Φ *and every* $t \in (0,1)$*,*

$$\eta^{\mathrm{Riem}}_{\kappa_t^{\mathrm{WYD}}}(\Phi) \le \eta^{\mathrm{RelEnt}}_{g^{(t)}}(\Phi) = \sup_{\rho,\gamma\in\mathcal{D}_d,\,\rho\ne\gamma} \frac{1-\mathrm{Tr}\,\Phi(\rho)^t\Phi(\gamma)^{1-t}}{1-\mathrm{Tr}\,\rho^t\gamma^{1-t}}.$$

For $\kappa_{\mathrm{WY}}(x) = 4/(1+\sqrt{x})^2$ and $\kappa_{\min}(x) = 2/(1+x)$, Theorem 5.1 together with (27) and (32) yields

$$\eta^{\mathrm{Riem}}_{\mathrm{WY}}(\Phi) = \sup_{\rho,\gamma\in\mathcal{D}_d,\,\rho\ne\gamma}\left[\frac{\arccos \mathrm{Tr}\,\Phi(\rho)^{1/2}\Phi(\gamma)^{1/2}}{\arccos \mathrm{Tr}\,\rho^{1/2}\gamma^{1/2}}\right]^2,$$
$$\eta^{\mathrm{Riem}}_{\min}(\Phi) = \sup_{\rho,\gamma\in\mathcal{D}_d,\,\rho\ne\gamma}\left[\frac{\arccos F(\Phi(\rho),\Phi(\gamma))}{\arccos F(\rho,\gamma)}\right]^2. \tag{54}$$

Since

$$\left(\frac{\arccos t}{\arccos s}\right)^2 < \frac{1-t}{1-s} \quad \text{for } 0<s<t<1,$$

from (54) and (31) we also have

Proposition 7.4. *For every CPT map* Φ*,*

$$\eta^{\mathrm{Riem}}_{\min}(\Phi) \le \sup_{\rho,\gamma\in\mathcal{D}_d,\,\rho\ne\gamma} \frac{1-F(\Phi(\rho),\Phi(\gamma))}{1-F(\rho,\gamma)} = \sup_{\rho,\gamma\in\mathcal{D}_d,\,\rho\ne\gamma}\left[\frac{d_{\mathrm{Bures}}(\Phi(\rho),\Phi(\gamma))}{d_{\mathrm{Bures}}(\rho,\gamma)}\right]^2.$$

Acknowledgments

Part of this work was done when both authors participated in a workshop in 2013 at Centro de Ciencias de Benasque Pedro Pascual in Benasque, Spain and during the ICM Satellite Conference in 2014 on Operator Algebras and Applications in Cheongpung, Korea. The work of FH was supported in part by Grant-in-Aid for Scientific Research (C)21540208. The work of MBR was partially supported by NSF grant CCF 1018401 which was administered by Tufts University. The authors are grateful to an anonymous referee for careful reading of the manuscript.

A Hiai-Petz Lemma

The next lemma was proved in [24] in a slightly different setting of the Riemannian manifold $\mathbb{P}_d$ and its tangent space $\mathbb{H}_d$ instead of $\mathcal{D}_d$ and $\mathbb{H}_d^0$ here, whose proof can work in the present setting as well. The proof is provided below for completeness.

Lemma A.1. [24, Lemma 4.2] *For every $\kappa \in \mathcal{K}$, $\rho \in \mathcal{D}_d$ and $A \in \mathbb{H}_d^0$,*

$$\lim_{\varepsilon \searrow 0} \frac{D_\kappa(\rho, \rho + \varepsilon A)}{\varepsilon} = \sqrt{\langle A, \Omega_\rho^\kappa(A) \rangle}.$$

Proof. First, recall that if $\mathbb{T}$ is a linear operator on the Hilbert space $(\mathbb{M}_d, \langle \cdot, \cdot \rangle)$ represented as the Schur multiplication by a matrix $(t_{ij}) \in \mathbb{H}_d$, then $\mathbb{T} \geq 0$ if and only if $t_{ij} \geq 0$ for all $i, j = 1, \ldots, n$. We denote by $\mathbb{I}$ the identity operator on $(\mathbb{M}_d, \langle \cdot, \cdot \rangle)$, which is represented as the Schur multiplication by the matrix of all entries equal to 1. To prove the lemma, we may assume that $\rho = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$. For brevity let $a_{ij} \equiv \lambda_j^{-1} \kappa(\lambda_i \lambda_j^{-1})$ for $i, j = 1, \ldots, d$ and $\alpha \equiv \min_{i,j} a_{ij} > 0$. Since

$$\Omega_\rho^\kappa(X) = (a_{ij})_{ij} \circ X, \qquad X \in \mathbb{M}_d$$

(see, e.g., [22, Section 2.2]), it follows that $\Omega_\rho^\kappa \geq \alpha \mathbb{I}$ as operators on $(\mathbb{M}_d, \langle \cdot, \cdot \rangle)$. For each $\delta \in (0, \alpha)$, since $\gamma \in \mathcal{D}_d \mapsto \Omega_\gamma^\kappa$ is continuous, there exists an $r_1 > 0$ such that if $\gamma \in \mathcal{D}_d$ and $\|\gamma - \rho\|_2 < r_1$ then $\|\Omega_\gamma^\kappa - \Omega_\rho^\kappa\|_\infty < \delta$, where $\| \cdot \|_\infty$ denotes the operator norm for operators on $(\mathbb{M}_d, \langle \cdot, \cdot \rangle)$. Furthermore, since D_κ and $\| \cdot \|_2$ define the same topology on $\mathcal{D}_d$ (see [35, Chapter IV, Proposition 3.5]), there exists an $r_0 > 0$ such that if $\gamma \in \mathcal{D}_d$ and $D_\kappa(\gamma, \rho) < r_0$ then $\|\gamma - \rho\|_2 < r_1$.

Now let $A \in \mathbb{H}_d^0$ and choose a sufficiently small $\varepsilon > 0$ so that $D_\kappa(\rho, \rho + \varepsilon A) < r_0$ and $\varepsilon \|A\|_2 < r_1$. Let $\xi : [0, 1] \to \mathcal{D}_d$ be any smooth curve from ρ to $\rho + \varepsilon A$ such that $L_\kappa(\xi) < r_0$, where $L_\kappa(\xi)$ is the length of ξ with respect to the monotone metric induced by κ. Since $D_\kappa(\xi(t), \rho) < r_0$ and so $\|\xi(t) - \rho\|_2 < r_1$ for all $0 \leq t \leq 1$, we have

$$\begin{aligned} L_\kappa(\xi) &= \int_0^1 \sqrt{\langle \xi'(t), \Omega_{\xi(t)}^\kappa(\xi'(t)) \rangle}\, dt \\ &\geq \int_0^1 \sqrt{\langle \xi'(t), (\Omega_\rho^\kappa - \delta \mathbb{I})(\xi'(t)) \rangle}\, dt \\ &= \int_0^1 \|(\Omega_\rho^\kappa - \delta \mathbb{I})^{1/2}(\xi'(t))\|\, dt \\ &\geq \|(\Omega_\rho^\kappa - \delta \mathbb{I})^{1/2}(\varepsilon A)\|_2 \\ &= \varepsilon \left\| \left((a_{ij} - \delta)^{1/2} \right)_{ij} \circ A \right\|_2. \end{aligned}$$

In the above, note that $\Omega_\rho^\kappa - \delta \mathbb{I} \geq 0$ on $(\mathbb{M}_d, \langle \cdot, \cdot \rangle)$ since $\delta < \alpha$. Also, the second inequality above follows since $\int_0^1 \|(\Omega_\rho^\kappa - \delta I)^{1/2}(\xi'(t))\|\, dt$ is the length of the curve $(\Omega_\rho^\kappa - \delta \mathbb{I})^{1/2}(\xi(t))$, $0 \leq t \leq 1$, from $(\Omega_\rho^\kappa - \delta \mathbb{I})^{1/2}(\rho)$ to $(\Omega_\rho^\kappa - \delta \mathbb{I})^{1/2}(\rho + \varepsilon A)$ in the Euclidean space $(\mathbb{H}_d, \| \cdot \|_2)$ and it is shortest if ξ is the segment between ρ and $\rho + \varepsilon A$. Taking the infimum of $L_\kappa(\xi)$ yields

$$D_\kappa(\rho, \rho + \varepsilon A) \geq \varepsilon \left\| \left((a_{ij} - \delta)^{1/2} \right)_{ij} \circ A \right\|_2.$$

On the other hand, let $\xi_0(t) \equiv \rho + t\varepsilon A$. Since $\|\xi_0(t) - \rho\|_2 \le \varepsilon \|A\|_2 < r_1$ for $0 \le t \le 1$, we have

$$\begin{aligned} D_\kappa(\rho, \rho + \varepsilon A) \le L_\kappa(\xi_0) &= \int_0^1 \sqrt{\langle \xi_0'(t), \Omega^\kappa_{\xi_0(t)}(\xi_0'(t)) \rangle}\, dt \\ &\le \int_0^1 \sqrt{\langle \xi_0'(t), (\Omega^\kappa_\rho + \delta \mathbb{I})(\xi_0'(t)) \rangle}\, dt \\ &= \|(\Omega^\kappa_\rho + \delta\mathbb{I})^{1/2}(\varepsilon A)\|_2 \\ &= \varepsilon \left\| \left((a_{ij} + \delta)^{1/2} \right)_{ij} \circ A \right\|_2 . \end{aligned}$$

Since δ is arbitrary,

$$\lim_{\varepsilon \searrow 0} \frac{D_\kappa(\rho, \rho + \varepsilon A)}{\varepsilon} = \left\| \left(a_{ij}^{1/2} \right)_{ij} \circ A \right\|_2 = \|(\Omega^\kappa_\rho)^{1/2}(A)\|_2 = \sqrt{\langle A, \Omega^\kappa_\rho(A) \rangle},$$

as desired. □

B Qubit Proofs

B.1 Useful results

B.1.1 Basic formulas

We observe that any Hermitian matrix with $\operatorname{Tr} A = 0$ can be written as $A = \mathbf{y} \cdot \sigma$ with $\mathbf{y} \in \mathbf{R}^3$, and that $\operatorname{Tr}(aI + \mathbf{w} \cdot \sigma) = 2a$.

The following formulas will be useful:

$$(aI + \mathbf{w}\cdot\sigma)(bI + \mathbf{y}\cdot\sigma) = (ab + \mathbf{w}\cdot\mathbf{y})I + (a\mathbf{y} + b\mathbf{w} + i\mathbf{w}\times\mathbf{y})\cdot\sigma, \tag{55}$$

$$(aI + \mathbf{w}\cdot\sigma)^{-1} = \frac{1}{a^2 - |\mathbf{w}|^2}(aI - \mathbf{w}\cdot\sigma), \tag{56}$$

$$(bI + \mathbf{w}\cdot\sigma)^{1/2} = \sqrt{\frac{\zeta(b,\mathbf{w})}{2}}\left[I + \frac{\mathbf{w}\cdot\sigma}{\zeta(b,\mathbf{w})}\right], \tag{57}$$

where $\zeta(b, \mathbf{w}) \equiv b + \sqrt{b^2 - |\mathbf{w}|^2}$.

It will be convenient to use the physicists bra and ket notation for vectors in $\mathbf{R}^3$ as well as $\mathbf{C}^d$ in which $|\mathbf{x}\rangle\langle\mathbf{x}|$ denotes $|\mathbf{x}|^2$ times the projection onto $\mathbf{x}$, more generally $|\mathbf{w}\rangle\langle\mathbf{x}| : \mathbf{y} \mapsto (\mathbf{x}\cdot\mathbf{y})\mathbf{w}$. In that notation, if $a \ne 0$ and $a \ne b|\mathbf{w}|^2$ then

$$\big(aI - b|\mathbf{w}\rangle\langle\mathbf{w}|\big)^{-1} = a^{-1}\left[I + \frac{b}{a - b|\mathbf{w}|^2}\,|\mathbf{w}\rangle\langle\mathbf{w}|\right]. \tag{58}$$

The following lemmas will be useful in Sections B.2–B.4 to prove the theorems of Section 6.

Lemma B.1. *Let A, B be positive invertible linear operators on $\mathbf{R}^3$. Then*

$$\sup_{\mathbf{y}\ne 0} \frac{\langle T\mathbf{y}, A^{-1}T\mathbf{y}\rangle}{\langle \mathbf{y}, B^{-1}\mathbf{y}\rangle} = \sup_{\mathbf{y}\ne 0} \frac{\langle T^*\mathbf{y}, BT^*\mathbf{y}\rangle}{\langle \mathbf{y}, A\mathbf{y}\rangle}.$$

Proof. Writing $\mathbf{z} = B^{-1/2}\mathbf{y}$ one finds

$$\begin{aligned}
\sup_{\mathbf{y}\neq 0} \frac{\langle T\mathbf{y}, A^{-1}T\mathbf{y}\rangle}{\langle \mathbf{y}, B^{-1}\mathbf{y}\rangle} &= \sup_{\mathbf{z}\neq 0} \frac{\langle B^{1/2}\mathbf{z}, T^* A^{-1} T B^{1/2}\mathbf{z}\rangle}{\langle \mathbf{z}, \mathbf{z}\rangle} \\
&= \|B^{1/2} T^* A^{-1} T B^{1/2}\| \\
&= \|A^{-1/2} T B T^* A^{-1/2}\| \\
&= \sup_{\mathbf{z}\neq 0} \frac{\langle A^{-1/2}\mathbf{z}, T B T^* A^{-1/2}\mathbf{z}\rangle}{\langle \mathbf{z}, \mathbf{z}\rangle},
\end{aligned}$$

where we have used the fact that $\|\Upsilon^*\Upsilon\| = \|\Upsilon\Upsilon^*\|$ with $\Upsilon = A^{-1/2} T B^{1/2}$. Then redefining $\mathbf{y} = A^{-1/2}\mathbf{z}$ gives the desired result. □

Lemma B.2. *For any fixed $\mu > 0$, $\nu \geq 0$ and $\mathbf{w}$, the minimum of*

$$F(y_2, y_3) \equiv \mu(1 + y_2^2 + y_3^2) + \nu(w_1 + w_2 y_2 + w_3 y_3)^2$$

is

$$\min_{y_2, y_2} F(y_2, y_3) = \frac{\mu(\mu + \nu\,|\mathbf{w}|^2)}{\mu + \nu(|\mathbf{w}|^2 - w_1^2)}.$$

Proof. The condition that $\nabla F = 0$ yields two linear equations which can be written in the form

$$\begin{pmatrix} \mu + \nu\, w_2^2 & \nu\, w_2 w_3 \\ \nu\, w_2 w_3 & \mu + \nu\, w_3^2 \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix} = -\nu\, w_1 \begin{pmatrix} w_2 \\ w_3 \end{pmatrix}.$$

This has the solution

$$\begin{aligned}
\begin{pmatrix} y_2 \\ y_3 \end{pmatrix} &= \frac{-\nu\, w_1}{\det} \begin{pmatrix} \mu + \nu\, w_3^2 & -\nu\, w_2 w_3 \\ -\nu\, w_2 w_3 & \mu + \nu\, w_2^2 \end{pmatrix} \begin{pmatrix} w_2 \\ w_3 \end{pmatrix} \\
&= \frac{-\mu\nu\, w_1}{\det} \begin{pmatrix} w_2 \\ w_3 \end{pmatrix} = \frac{-\nu\, w_1}{\mu + \nu\,(w_2^2 + w_3^2)} \begin{pmatrix} w_2 \\ w_3 \end{pmatrix}
\end{aligned}$$

since $\det = \mu^2 + \mu\nu\,(w_2^2 + w_3^2)$. It is now easy to compute the value of F at this solution as

$$\begin{aligned}
&\frac{\mu\big[\mu + \nu(w_2^2 + w_3^2)\big]^2 + \mu\nu^2 w_1^2 (w_2^2 + w_3^2) + \mu^2 \nu w_1^2}{\big[\mu + \nu(w_2^2 + w_3^2)\big]^2} \\
&= \frac{\mu\big[\mu + \nu(w_2^2 + w_3^2)\big] + \mu\nu w_1^2}{\mu + \nu(w_2^2 + w_3^2)} = \frac{\mu(\mu + \nu\,|\mathbf{w}|^2)}{\mu + \nu(|\mathbf{w}|^2 - w_1^2)}. \quad \square
\end{aligned}$$

Lemma B.3. *Let $P = I + \mathbf{w}\cdot\sigma$ with $|\mathbf{w}| < 1$. Then for every $\mathbf{y} \in \mathbf{R}^3$,*

$$\frac{2}{L_P + R_P}(\mathbf{y}\cdot\sigma) = -\frac{\mathbf{w}\cdot\mathbf{y}}{1 - |\mathbf{w}|^2} I + \left[\mathbf{y} + \frac{(\mathbf{w}\cdot\mathbf{y})\,\mathbf{w}}{1 - |\mathbf{w}|^2}\right]\cdot\sigma \tag{59}$$

Proof. Write $\beta I + \mathbf{z}\cdot\sigma = 2(L_P + R_P)^{-1}(\mathbf{y}\cdot\sigma)$. Then (55) and $\mathbf{z}\times\mathbf{w} = -\mathbf{w}\times\mathbf{z}$ imply that

$$\begin{aligned}
\mathbf{y}\cdot\sigma &= \tfrac{1}{2}(I + \mathbf{w}\cdot\sigma)(\beta I + \mathbf{z}\cdot\sigma) + \tfrac{1}{2}(\beta I + \mathbf{z}\cdot\sigma)(I + \mathbf{w}\cdot\sigma) \\
&= (\beta + \mathbf{z}\cdot\mathbf{w}) I + (\mathbf{z} + \beta\mathbf{w})\cdot\sigma.
\end{aligned}$$

Since I and the Pauli matrices form a basis for $\mathbb{M}_2$, this implies $\beta = -\mathbf{z}\cdot\mathbf{w}$ and $\mathbf{y} = \mathbf{z} + \beta\mathbf{w}$ so that

$$\mathbf{y} = \mathbf{z} - (\mathbf{z}\cdot\mathbf{w})\mathbf{w} = (I - |\mathbf{w}\rangle\langle\mathbf{w}|)\mathbf{z}.$$

Then using (58) with $a = b = 1$ we find that

$$\mathbf{z} = \left[I + \frac{|\mathbf{w}\rangle\langle\mathbf{w}|}{1-|\mathbf{w}|^2}\right]\mathbf{y} = \mathbf{y} + \frac{(\mathbf{w}\cdot\mathbf{y})\mathbf{w}}{1-|\mathbf{w}|^2}.$$

Inserting this into $\beta = -\mathbf{z}\cdot\mathbf{w}$ yields (59). □

B.1.2 Lemmas for extreme points

Lemma B.4. *Let $s \geq 0$, $P = I + \mathbf{w}\cdot\sigma$ and $Q = I + \mathbf{x}\cdot\sigma$ with $|\mathbf{w}|, |\mathbf{x}| < 1$. Let $\mathbf{u} = \mathbf{w} - s\mathbf{x}$ and $\mathbf{v} = \mathbf{w} + s\mathbf{x}$. Then for every $\mathbf{y} \in \mathbf{R}^3$,*

$$\begin{aligned}&\mathrm{Tr}\,(\mathbf{y}\cdot\sigma)\frac{1+s}{L_P + sR_Q}(\mathbf{y}\cdot\sigma)\\ &\quad = 2(1+s)^2\left\langle \mathbf{y}, \left[\left\{(1+s)^2 - |\mathbf{u}|^2\right\}I + |\mathbf{u}\rangle\langle\mathbf{u}| - |\mathbf{v}\rangle\langle\mathbf{v}| - \frac{|\mathbf{u}\times\mathbf{v}\rangle\langle\mathbf{u}\times\mathbf{v}|}{(1+s)^2 - |\mathbf{v}|^2}\right]^{-1}\mathbf{y}\right\rangle, \qquad (60)\end{aligned}$$

where the operator inside $[\quad]^{-1}$ of (60) is positive and invertible.

Proof. As above, let $\beta I + \mathbf{z}\cdot\sigma = \big(L_P + sR_Q\big)^{-1}(\mathbf{y}\cdot\sigma)$ so that

$$\begin{aligned}\mathbf{y}\cdot\sigma &= \big(L_P + sR_Q\big)(\beta I + \mathbf{z}\cdot\sigma)\\ &= [(1+s)\beta + \mathbf{z}\cdot(\mathbf{w} + s\mathbf{x})]I + [(1+s)\mathbf{z} + \beta(\mathbf{w}+s\mathbf{x}) - i\mathbf{z}\times(\mathbf{w} - s\mathbf{x})]\cdot\sigma\\ &= [(1+s)\beta + \mathbf{z}\cdot\mathbf{v}]I + [(1+s)\mathbf{z} + \beta\mathbf{v} - i\mathbf{z}\times\mathbf{u}]\cdot\sigma. \qquad (61)\end{aligned}$$

Since the coefficient of I on the right side of (61) must be 0, we find

$$\beta = -\frac{\mathbf{z}\cdot\mathbf{v}}{1+s}.$$

Inserting this into (61) and equating real and imaginary parts yield

$$\mathbf{y} = \left[(1+s)I - \frac{|\mathbf{v}\rangle\langle\mathbf{v}|}{1+s}\right]\mathbf{z}_1 + \mathbf{z}_2\times\mathbf{u}, \qquad (62a)$$

$$0 = \left[(1+s)I - \frac{|\mathbf{v}\rangle\langle\mathbf{v}|}{1+s}\right]\mathbf{z}_2 - \mathbf{z}_1\times\mathbf{u}, \qquad (62b)$$

where we have written $\mathbf{z} = \mathbf{z}_1 + i\mathbf{z}_2$ and $(\mathbf{z}\cdot\mathbf{v})\mathbf{v} = |\mathbf{v}\rangle\langle\mathbf{v}|\mathbf{z}$. Solving (62b) for $\mathbf{z}_2$ with use of (58) yields

$$\mathbf{z}_2 = \frac{1}{1+s}\left[I + \frac{|\mathbf{v}\rangle\langle\mathbf{v}|}{(1+s)^2 - |\mathbf{v}|^2}\right](\mathbf{z}_1\times\mathbf{u}).$$

Inserting this into (62a) gives

$$\mathbf{y} = \frac{1}{1+s}\left[\left\{(1+s)^2 - |\mathbf{u}|^2\right\}I + |\mathbf{u}\rangle\langle\mathbf{u}| - |\mathbf{v}\rangle\langle\mathbf{v}| - \frac{|\mathbf{u}\times\mathbf{v}\rangle\langle\mathbf{u}\times\mathbf{v}|}{(1+s)^2 - |\mathbf{v}|^2}\right]\mathbf{z}_1, \qquad (63)$$

where we have used

$$(\mathbf{z}_1 \times \mathbf{u}) \times \mathbf{u} = -(\mathbf{u}\cdot\mathbf{u})\mathbf{z}_1 + (\mathbf{u}\cdot\mathbf{z}_1)\mathbf{u},$$
$$[|\mathbf{v}\rangle\langle\mathbf{v}|(\mathbf{z}_1 \times \mathbf{u})] \times \mathbf{u} = [\mathbf{v}\cdot(\mathbf{z}_1 \times \mathbf{u})]\,\mathbf{v}\times\mathbf{u} = -|\mathbf{v}\times\mathbf{u}\rangle\langle\mathbf{v}\times\mathbf{u}|\,\mathbf{z}_1$$

for the first and the second terms of $\mathbf{z}_2$, respectively. Thus we have proved that for every $\mathbf{y} \in \mathbf{R}^3$ there exists a $\mathbf{z}_1 \in \mathbf{R}^3$ satisfying (63). This implies that the operator inside $[\quad]$ of (63) is surjective and hence invertible. Since

$$\mathrm{Tr}\,(\mathbf{y}\cdot\sigma)\frac{1}{L_P + sR_Q}(\mathbf{y}\cdot\sigma) = \mathrm{Tr}\,(\mathbf{y}\cdot\sigma)(\beta I + \mathbf{z}\cdot\sigma) = 2\mathbf{y}\cdot\mathbf{z}_1,$$

we obtain (60) by solving for $\mathbf{z}_1$ in (63). Moreover, since the LHS of (60) is ≥ 0, the operator inside $[\quad]^{-1}$ is indeed positive. □

In the case where $\mathbf{y} = \mathbf{w} - \mathbf{x} \in \mathrm{span}\{\mathbf{u}, \mathbf{v}\}$ and so $\mathbf{y}$ is orthogonal to $\mathbf{u}\times\mathbf{v}$, we can simplify the expression above.

Lemma B.5. *In the notation of Lemma B.4, when* $\mathbf{y}$ *is orthogonal to* $\mathbf{u}\times\mathbf{v}$ *equation* (60) *becomes*

$$2(1+s)^2\left\langle \mathbf{y}, \left[\left\{(1+s)^2 - |\mathbf{u}|^2\right\}I + |\mathbf{u}\rangle\langle\mathbf{u}| - |\mathbf{v}\rangle\langle\mathbf{v}|\right]^{-1}\mathbf{y}\right\rangle, \tag{64}$$

where the operator inside $[\quad]^{-1}$ *of* (64) *is positive and invertible.*

Proof. First note that when $X \geq 0$ and $I - W - X \geq 0$ is invertible, then $I - W \geq 0$ is also invertible. Then it suffices to observe that when $WX = 0$ and $X\mathbf{y} = 0$,

$$(I - W - X)^{-1}\mathbf{y} = \sum_{k=0}^{\infty}(W + X)^k\mathbf{y} = \sum_{k=0}^{\infty} W^k\mathbf{y} = (I - W)^{-1}\mathbf{y},$$

and apply this with

$$W = \frac{|\mathbf{v}\rangle\langle\mathbf{v}| - |\mathbf{u}\rangle\langle\mathbf{u}|}{(1+s)^2 - |\mathbf{u}|^2}, \qquad X = \frac{|\mathbf{u}\times\mathbf{v}\rangle\langle\mathbf{u}\times\mathbf{v}|}{\left[(1+s)^2 - |\mathbf{u}|^2\right]\left[(1+s)^2 - |\mathbf{v}|^2\right]}. \quad \square$$

Note that when $P = Q$ so that $\mathbf{u} = (1-s)\mathbf{w}$ and $\mathbf{v} = (1+s)\mathbf{w}$, the s-dependence of all terms in (64) has the form $(1 \pm s)^2/(1+s)^2 = (1 \pm s^{-1})^2/(1+s^{-1})^2$, which implies that for $A = \mathbf{y}\cdot\sigma$

$$\langle A, \Omega_P^{\kappa_s}(A)\rangle = \left\langle A, \frac{1+s}{L_P + sR_P}(A)\right\rangle = \left\langle A, \frac{1+s}{R_P + sL_P}(A)\right\rangle \tag{65}$$

as

$$\frac{1+s^{-1}}{L_P + s^{-1}R_P} = \frac{1+s}{R_P + sL_P}.$$

In fact, we have a general fact that for any function $h : (0,\infty) \to \mathbf{R}$ and $A \in \mathbb{H}_d$,

$$\langle A, R_P^{-1}h\big(L_P R_P^{-1}\big)(A)\rangle = \langle A, R_P^{-1}\widetilde{h}\big(L_P R_P^{-1}\big)(A)\rangle,$$

where $\widetilde{h}(x) \equiv x^{-1}h(x^{-1})$. (65) is the case when $h(x) = 1/(x+s)$ and $\widetilde{h}(x) = 1/(1+sx)$.

It is not obvious that the operator in Lemma B.5 is positive and invertible. When $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, one finds

$$|\mathbf{u}|^2 + |\mathbf{v}|^2 = |\mathbf{u} \pm \mathbf{v}|^2 = 4|\mathbf{w}|^2 = 4s^2|\mathbf{x}|^2 < (1+s)^2.$$

Then one can readily verify that the expectation of the operator inside $[\quad]^{-1}$ in Lemma B.4 is positive for the choices $\mathbf{y} = \mathbf{v}$ and $\mathbf{y} = \mathbf{u} \times \mathbf{v}$. Since $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u} \times \mathbf{v}$ are orthogonal, this implies that the expectation is positive for arbitrary $\mathbf{y}$ with the additional assumption $\mathbf{u} \cdot \mathbf{v} = 0$.

For the general case, observe that since $|\mathbf{u} \times \mathbf{v}\rangle\langle \mathbf{u} \times \mathbf{v}|$ is orthogonal to $|\mathbf{u}\rangle\langle \mathbf{u}|$ and $|\mathbf{v}\rangle\langle \mathbf{v}|$, it suffices to prove that the two operators

$$\big\{(1+s)^2 - |\mathbf{u}|^2\big\}I + |\mathbf{u}\rangle\langle \mathbf{u}| - |\mathbf{v}\rangle\langle \mathbf{v}|, \qquad \big\{(1+s)^2 - |\mathbf{u}|^2\big\}I - \frac{|\mathbf{u}\times\mathbf{v}\rangle\langle \mathbf{u}\times\mathbf{v}|}{(1+s)^2 - |\mathbf{v}|^2}$$

are positive and invertible. Since $\mathbf{u} = \mathbf{w} - s\mathbf{x}$ and $\mathbf{v} = \mathbf{w} + s\mathbf{x}$, one can show that the problem reduces to showing that for every $|\mathbf{w}|, |\mathbf{x}| < 1$ and $|\mathbf{y}| = 1$,

$$(1+s)^2 - |\mathbf{w}|^2 - s^2|\mathbf{x}|^2 + 2s\langle \mathbf{w}, \mathbf{x}\rangle - 4s\langle \mathbf{y}, \mathbf{w}\rangle\langle \mathbf{y}, \mathbf{x}\rangle > 0,$$

$$\big[(1+s)^2 - \big(|\mathbf{w}|^2 + s^2|\mathbf{x}|^2\big)\big]^2 - 4s^2|\mathbf{w}|^2|\mathbf{x}|^2 > 0.$$

The first inequality above can be proved by using the cosine rule for spherical triangles, and the second is straightforward.

B.2 Unital qubit channels: Proof of Theorem 6.1

Writing rank 1 projections as $E = \frac{1}{2}(I + \mathbf{w}\cdot\sigma)$ and $F = \frac{1}{2}(I + \mathbf{x}\cdot\sigma)$ with $|\mathbf{w}| = |\mathbf{x}| = 1$, we find $EF = 0 \Leftrightarrow \mathbf{w} = -\mathbf{x}$, in which case $E - F = \mathbf{w}\cdot\sigma$ so that by (35),

$$\eta^{\mathrm{Tr}}(\Phi_T) = \tfrac{1}{2} \sup_{|\mathbf{w}|=1} \mathrm{Tr}\, |(T\mathbf{w})\cdot\sigma| = \sup_{|\mathbf{w}|=1} |T\mathbf{w}| = \|T\|_\infty. \tag{66}$$

Note that this implies that Φ_T is "non-scrambling" if and only if $\|T\|_\infty = 1$.

It then follows from Theorems 5.1 and 5.3 that

$$\eta^{\mathrm{geod}}_\kappa(\Phi_T) = \eta^{\mathrm{Riem}}_\kappa(\Phi_T) \geq \|T\|^2_\infty$$

for all $\kappa \in \mathcal{K}$. Thus it suffices to show that $\eta^{\mathrm{RelEnt}}_g(\Phi_T) \leq \|T\|^2_\infty$ for all $g \in \mathcal{G}$. In fact, from the integral expression (13) it suffices to do this for $g_s \equiv (x-1)^2/(x+s)$, $s \geq 0$, as in (39), for which we have

$$H_{g_s}(\rho, \gamma) = \mathrm{Tr}\,(\rho - \gamma)\frac{1}{L_\rho + sR_\gamma}(\rho - \gamma). \tag{67}$$

Let $\rho = \frac{1}{2}(I + \mathbf{w}\cdot\sigma)$ and $\gamma = \frac{1}{2}(I + \mathbf{x}\cdot\sigma)$ with $|\mathbf{w}|, |\mathbf{x}| < 1$ and $\mathbf{y} \equiv \mathbf{w} - \mathbf{x} \neq 0$ (which guarantee $\rho, \gamma \in \mathcal{D}_2$ and $\rho \neq \gamma$). For $s = 0$ it follows immediately from the elementary formulas (55) and (56) in Section B.1.1 that

$$H_{g_0}(\rho, \gamma) = 2\mathrm{Tr}\,(\mathbf{y}\cdot\sigma)^2(I + \mathbf{w}\cdot\sigma)^{-1} = \frac{4|\mathbf{y}|^2}{1 - |\mathbf{w}|^2}.$$

Since $\Phi_T(\rho) = \frac{1}{2}[I + (T\mathbf{w})\cdot\sigma]$ and $\Phi_T(\mathbf{y}\cdot\sigma) = (T\mathbf{y})\cdot\sigma$, we have

$$\frac{H_{g_0}(\Phi_T(\rho), \Phi_T(\gamma))}{H_{g_0}(\rho,\gamma)} = \frac{1-|\mathbf{w}|^2}{4|\mathbf{y}|^2}\cdot\frac{4|T\mathbf{y}|^2}{1-|T\mathbf{w}|^2} = \frac{1-|\mathbf{w}|^2}{1-|T\mathbf{w}|^2}\cdot\frac{|T\mathbf{y}|^2}{|\mathbf{y}|^2}$$

so that

$$\eta_{g_0}^{\mathrm{RelEnt}}(\Phi_T) = \sup_{|\mathbf{w}|<1}\frac{1-|\mathbf{w}|^2}{1-|T\mathbf{w}|^2}\cdot\sup_{\mathbf{y}\neq 0}\frac{|T\mathbf{y}|^2}{|\mathbf{y}|^2} = \|T\|_\infty^2.$$

Next, for $s>0$ define $\mathbf{u} = \mathbf{w} - s\mathbf{x}$ and $\mathbf{v} = \mathbf{w} + s\mathbf{x}$. Then observe that $\mathbf{y} = \mathbf{w} - \mathbf{x} = \frac{1}{2s}[(1+s)\mathbf{u} + (1-s)\mathbf{v}]$, is in the span of $\mathbf{u}$ and $\mathbf{v}$ so that $\mathbf{y}$ is orthogonal to $\mathbf{u}\times\mathbf{v}$. Similarly $T\mathbf{y}$ is orthogonal to $T\mathbf{u}\times T\mathbf{v}$. Therefore we can use Lemma B.5 to conclude

$$\eta_{g_s}^{\mathrm{RelEnt}}(\Phi_T) = \sup_{\mathcal{M}_s}\frac{\left\langle T\mathbf{y}, \left(\left[(1+s)^2 - |T\mathbf{u}|^2\right]I + |T\mathbf{u}\rangle\langle T\mathbf{u}| - |T\mathbf{v}\rangle\langle T\mathbf{v}|\right)^{-1} T\mathbf{y}\right\rangle}{\left\langle \mathbf{y}, \left(\left[(1+s)^2 - |\mathbf{u}|^2\right]I + |\mathbf{u}\rangle\langle \mathbf{u}| - |\mathbf{v}\rangle\langle \mathbf{v}|\right)^{-1} \mathbf{y}\right\rangle},$$

where we have suppressed the dependence of $\mathbf{y} = \frac{1}{2s}[(1+s)\mathbf{v} - (1-s)\mathbf{u}]$ on $\mathbf{u}$ and $\mathbf{v}$ and

$$\begin{aligned}\mathcal{M}_s &\equiv \{(\mathbf{u},\mathbf{v}) : \mathbf{u} = \mathbf{w} - s\mathbf{x},\ \mathbf{v} = \mathbf{w} + s\mathbf{x},\ |\mathbf{w}|<1,\ |\mathbf{x}|<1,\ \mathbf{w}\neq\mathbf{x}\}\\ &= \left\{(\mathbf{u},\mathbf{v}) : |\mathbf{u}+\mathbf{v}|<2,\ |\mathbf{u}-\mathbf{v}|<2s,\ \mathbf{v}\neq\frac{1-s}{1+s}\mathbf{u}\right\}.\end{aligned}$$

Since relaxing the constraints on $\mathbf{y}$ can only increase the supremum, we find using Lemma B.1 that

$$\begin{aligned}\eta_{g_s}^{\mathrm{RelEnt}}(\Phi_T) &\leq \sup_{\mathcal{M}_s}\sup_{\mathbf{y}\neq 0}\frac{\left\langle T^*\mathbf{y}, \left(\left[(1+s)^2 - |\mathbf{u}|^2\right]I + |\mathbf{u}\rangle\langle\mathbf{u}| - |\mathbf{v}\rangle\langle\mathbf{v}|\right)T^*\mathbf{y}\right\rangle}{\left\langle \mathbf{y}, \left(\left[(1+s)^2 - |T\mathbf{u}|^2\right]I + |T\mathbf{u}\rangle\langle T\mathbf{u}| - |T\mathbf{v}\rangle\langle T\mathbf{v}|\right)\mathbf{y}\right\rangle}\\ &= \sup_{\mathcal{M}_s}\sup_{\mathbf{y}\neq 0}\frac{|T^*\mathbf{y}|^2[(1+s)^2 - |\mathbf{u}|^2] + |\langle\mathbf{y}, T\mathbf{u}\rangle|^2 - |\langle\mathbf{y}, T\mathbf{v}\rangle|^2}{|\mathbf{y}|^2[(1+s)^2 - |T\mathbf{u}|^2] + |\langle\mathbf{y}, T\mathbf{u}\rangle|^2 - |\langle\mathbf{y}, T\mathbf{v}\rangle|^2}.\end{aligned}\tag{68}$$

It follows from Lemma B.5 that both the numerator and denominator in (68) are expectations of positive invertible operators and are, hence, positive. If we let $a = |T^*\mathbf{y}|^2(1+s)^2 - |\mathbf{u}|^2|T^*\mathbf{y}|^2 + |\mathbf{y}|^2|T\mathbf{u}|^2$, $b = |\mathbf{y}|^2(1+s)^2$, and $c = |\mathbf{y}|^2|T\mathbf{u}|^2 - |\langle\mathbf{y}, T\mathbf{u}\rangle|^2 + |\langle\mathbf{y}, T\mathbf{v}\rangle|^2$, (68) can be written as $(a-c)/(b-c)$ with $a\geq c\geq 0$ and $b>c\geq 0$. We estimate

$$\begin{aligned}\frac{a}{b} &= \frac{|T^*\mathbf{y}|^2(1+s)^2 - |\mathbf{u}|^2|T^*\mathbf{y}|^2 + |\mathbf{y}|^2|T\mathbf{u}|^2}{|\mathbf{y}|^2(1+s)^2}\\ &= \frac{|T^*\mathbf{y}|^2}{|\mathbf{y}|^2} + \frac{|\mathbf{u}|^2}{(1+s)^2}\left[\frac{|T\mathbf{u}|^2}{|\mathbf{u}|^2} - \frac{|T^*\mathbf{y}|^2}{|\mathbf{y}|^2}\right]\\ &\leq \begin{cases}\dfrac{|T^*\mathbf{y}|^2}{|\mathbf{y}|^2} & \text{if } |\mathbf{y}|^2|T\mathbf{u}|^2 \leq |\mathbf{u}|^2|T^*\mathbf{y}|^2\\ \dfrac{|T\mathbf{u}|^2}{|\mathbf{u}|^2} & \text{if } |\mathbf{y}|^2|T\mathbf{u}|^2 > |\mathbf{u}|^2|T^*\mathbf{y}|^2\end{cases} &(69)\\ &\leq \|T\|_\infty^2. &(70)\end{aligned}$$

In the second case above we have used $|\mathbf{u}|\leq 1+s$. Since $a\geq c\geq 0$ and $a/b\leq\|T\|_\infty^2\leq 1$ by the estimate above, for every $\mathbf{y}\neq 0$ and $(\mathbf{u},\mathbf{v})\in\mathcal{M}_s$ we have

$$\frac{a-c}{b-c}\leq\frac{a}{b}\leq\|T\|_\infty^2,$$

which implies that $\eta_{g_s}^{\mathrm{RelEnt}}(\Phi_T)\leq\|T\|_\infty^2$. □

B.3 Non-unital CQ qubit channels: Proof of Theorem 6.2

Before proving Theorem 6.2 observe that the effect of $\Phi \equiv \Phi_{\alpha,\tau}$ on $P = \frac{1}{2}(I + \mathbf{w}\cdot\sigma)$ is to map $\mathbf{w} \mapsto (\alpha w_1, 0, \tau)^t$ and on $A = \mathbf{y}\cdot\sigma$ it is $\mathbf{y} \mapsto (\alpha y_1, 0, 0)^t$. There is no loss of generality in simply using $P = I + \mathbf{w}\cdot\sigma$ whenever the factors of $\frac{1}{2}$ will cancel.

B.3.1 Dobrushin coefficient

As in the proof of (66),

$$\eta^{\mathrm{Tr}}(\Phi_{\alpha,\tau}) = \tfrac{1}{2} \sup_{|\mathbf{w}|=1} \mathrm{Tr}\, |\alpha w_1 \sigma_1| = \alpha$$

with the supremum attained at $\mathbf{w} = (1, 0, 0)^t$.

B.3.2 Maximal metric

It follows from (56) and (55) that for $P = I + \mathbf{w}\cdot\sigma$,

$$\big\langle \mathbf{y}\cdot\sigma, \Omega_P^{\max}(\mathbf{y}\cdot\sigma)\big\rangle = \mathrm{Tr}\,(\mathbf{y}\cdot\sigma)^2 P^{-1} = \frac{2|\mathbf{y}|^2}{1-|\mathbf{w}|^2}$$

so that

$$\begin{aligned}
\eta^{\mathrm{Riem}}_{\max}(\Phi_{\alpha,\tau}) &= \alpha^2 \sup_{|\mathbf{w}|<1} \sup_{\mathbf{y}\neq 0} \frac{y_1^2}{|\mathbf{y}|^2} \frac{1-|\mathbf{w}|^2}{1-\alpha^2 w_1^2 - \tau^2} \\
&= \alpha^2 \sup_{|w_1|<1} \frac{1-w_1^2}{1-\alpha^2 w_1^2-\tau^2} \qquad (71)\\
&= \frac{\alpha^2}{1-\tau^2} \;\leq\; 1
\end{aligned}$$

thanks to $\alpha^2 + \tau^2 \leq 1$.

B.3.3 $\kappa(x) = x^{-1/2}$

For $\kappa(x) = x^{-1/2}$, first observe that it follows from (56) and (57) that

$$(\mathbf{y}\cdot\sigma)(I+\mathbf{w}\cdot\sigma)^{-1/2} = \sqrt{\tfrac{\zeta}{2(1-|\mathbf{w}|^2)}}\Big[\tfrac{-1}{\zeta}(\mathbf{w}\cdot\mathbf{y})I + \big[\mathbf{y} + \tfrac{i}{\zeta}(\mathbf{w}\times\mathbf{y})\big]\cdot\sigma\Big] \qquad (72)$$

with $\zeta = \zeta(|\mathbf{w}|) \equiv 1 + \sqrt{1-|\mathbf{w}|^2}$ so that for $P = I + \mathbf{w}\cdot\sigma$

$$\begin{aligned}
\mathrm{Tr}\,(\mathbf{y}\cdot\sigma)\Omega_P^{x^{-1/2}}(\mathbf{y}\cdot\sigma) &= \mathrm{Tr}\Big[(\mathbf{y}\cdot\sigma)(I+\mathbf{w}\cdot\sigma)^{-1/2})\Big]^2 \\
&= \frac{\zeta}{1-|\mathbf{w}|^2}\Big[\tfrac{1}{\zeta^2}(\mathbf{w}\cdot\mathbf{y})^2 + |\mathbf{y}|^2 - \tfrac{1}{\zeta^2}|\mathbf{w}\times\mathbf{y}|^2\Big] \\
&= \frac{\zeta}{1-|\mathbf{w}|^2}\Big[\tfrac{1}{\zeta^2}|\mathbf{w}|^2|\mathbf{y}|^2(\cos^2\theta - \sin^2\theta) + |\mathbf{y}|^2\Big] \\
&= \frac{|\mathbf{y}|^2}{(1-|\mathbf{w}|^2)\zeta}\Big(|\mathbf{w}|^2\cos 2\theta + \zeta^2\Big), \qquad (73)
\end{aligned}$$

where θ denotes the angle between $\mathbf{w}$ and $\mathbf{y}$. Thus

$$\begin{aligned} &\eta^{\text{Riem}}_{x^{-1/2}}(\Phi_{\alpha,\tau}) \qquad (74)\\ &= \sup_{\mathbf{w},\mathbf{y}} \frac{\alpha^2 y_1^2}{|\mathbf{y}|^2} \frac{1-|\mathbf{w}|^2}{1-\alpha^2 w_1^2 - \tau^2} \frac{\zeta(|\mathbf{w}|)}{\zeta(|(\alpha w_1, 0, \tau)|)} \frac{(\alpha^2 w_1^2 + \tau^2)\cos 2\widetilde{\theta} + \zeta^2(|(\alpha w_1, 0, \tau)|)}{|\mathbf{w}|^2 \cos 2\theta + \zeta^2(|\mathbf{w}|)} \end{aligned}$$

with $\widetilde{\theta}$ the angle between $(\alpha y_1, 0, 0)$ and $(\alpha w_1, 0, \tau)$. The first ratio in this product is largest when $\mathbf{y} = (y_1, 0, 0)$ and $\mathbf{y}$ enters only implicitly in the last one in $\theta, \widetilde{\theta}$. Assuming $\mathbf{y} = (y_1, 0, 0)$ and $w_1 = 0$ gives $\widetilde{\theta} = \theta = \pi/2$ and $\cos 2\widetilde{\theta} = \cos 2\theta = -1$. Then the identity $\zeta^2(x) - x^2 = 2\sqrt{1-x^2}\,\zeta(x)$ can be used to simplify the RHS of (74) to give

$$\begin{aligned} \eta^{\text{Riem}}_{x^{-1/2}}(\Phi_{\alpha,\tau}) &\geq \sup_{\mathbf{w}=(0,w_2,w_3)} \alpha^2 \frac{1-|\mathbf{w}|^2}{1-\tau^2} \frac{\zeta(|\mathbf{w}|)}{\zeta(|\tau|)} \frac{2\sqrt{1-\tau^2}\,\zeta(|\tau|)}{2\sqrt{1-|\mathbf{w}|^2}\zeta(|\mathbf{w}|)} \\ &= \sup_{\mathbf{w}=(0,w_2,w_3)} \alpha^2 \frac{\sqrt{1-|\mathbf{w}|^2}}{\sqrt{1-\tau^2}} = \frac{\alpha^2}{\sqrt{1-\tau^2}}. \end{aligned}$$

We could have obtained this bound more easily by considering the special case $P = I$. However, since the methods used to obtain (73) are used again later, there is some merit in presenting the details in this relatively simple setting. For the special case, observe that $\Phi(I) = I + \tau\sigma_3$ so that $[\Phi(I)]^{-1/2} = \begin{pmatrix} \frac{1}{\sqrt{1+\tau}} & 0 \\ 0 & \frac{1}{\sqrt{1-\tau}} \end{pmatrix}$ and

$$\Phi(\mathbf{y}\cdot\sigma)[\Phi(I)]^{-1/2}\Phi(\mathbf{y}\cdot\sigma)[\Phi(I)]^{-1/2} = \alpha^2 y_1^2 \frac{1}{\sqrt{1-\tau^2}} I,$$

from which it follows that

$$\eta^{\text{Riem}}_{x^{-1/2}}(\Phi_{\alpha,\tau}) \geq \sup_{\mathbf{y}\neq 0} \alpha^2 \frac{y_1^2}{|\mathbf{y}|^2} \frac{1}{\sqrt{1-\tau^2}} = \frac{\alpha^2}{\sqrt{1-\tau^2}}.$$

B.3.4 BKM metric

For $\kappa_{\text{BKM}}(x) = (\log x)/(x-1)$ we can repeat some of the strategy above to get a lower bound. It follows from (22) that

$$\Omega^{\text{BKM}}_P(A) = \int_0^\infty \frac{1}{P+uI} A \frac{1}{P+uI}\, du.$$

Then using (56) with $a = 1 + u$ and (55), we find for $A = \mathbf{y}\cdot\sigma$ and $P = I + \mathbf{w}\cdot\sigma$

$$(\mathbf{y}\cdot\sigma)\frac{1}{(1+u)I + \mathbf{w}\cdot\sigma} = \frac{-\mathbf{w}\cdot\mathbf{y}I + [(1+u)\mathbf{y} + i\mathbf{w}\times\mathbf{y}]\cdot\sigma}{(1+u)^2 - |\mathbf{w}|^2},$$

from which it follows that

$$\begin{aligned} \text{Tr}\,(\mathbf{y}\cdot\sigma)\frac{1}{(1+u)I + \mathbf{w}\cdot\sigma}(\mathbf{y}\cdot\sigma)\frac{1}{(1+u)I + \mathbf{w}\cdot\sigma} &= \text{Tr}\left[(\mathbf{y}\cdot\sigma)\frac{1}{(1+u)I + \mathbf{w}\cdot\sigma}\right]^2 \\ &= \frac{2}{\left[(1+u)^2 - |\mathbf{w}|^2\right]^2}\left[|\mathbf{w}\cdot\mathbf{y}|^2 + (1+u)^2|\mathbf{y}|^2 - |\mathbf{w}\times\mathbf{y}|^2\right] \end{aligned}$$

$$= \frac{2\,|\mathbf{y}|^2}{\big[(1+u)^2 - |\mathbf{w}|^2\big]^2}\big[(1+u)^2 + |\mathbf{w}|^2\cos 2\theta\big], \tag{75}$$

where for the second equality above we have used $\mathbf{y}\cdot(\mathbf{w}\times\mathbf{y}) = 0$ and θ is the angle between $\mathbf{w}$ and $\mathbf{y}$. For $\theta = \pi/2$, $\cos 2\theta = -1$ so that (75) becomes $2|\mathbf{y}|^2/[(1+u)^2 - |\mathbf{w}|^2]$. Then since the integral

$$\int_0^\infty \frac{1}{(1+u)^2 - |\mathbf{w}|^2}\,du = \frac{1}{2|\mathbf{w}|}\log\tfrac{1+|\mathbf{w}|}{1-|\mathbf{w}|}$$

is elementary, we can conclude

$$\begin{aligned}
\eta^{\rm Riem}_{\rm BKM}(\Phi) &\geq \sup_{w_1=0,\,\mathbf{y}\cdot\mathbf{w}=0} \alpha^2 \frac{y_1^2}{|\mathbf{y}|^2}\frac{|\mathbf{w}|}{\tau}\frac{\log\frac{1+\tau}{1-\tau}}{\log\frac{1+|\mathbf{w}|}{1-|\mathbf{w}|}} \\
&= \sup_{\mathbf{w}=(0,w_2,w_3)} \frac{\alpha^2}{\tau}\log\tfrac{1+\tau}{1-\tau}\frac{|\mathbf{w}|}{\log\frac{1+|\mathbf{w}|}{1-|\mathbf{w}|}} \\
&= \frac{\alpha^2}{2\tau}\log\tfrac{1+\tau}{1-\tau},
\end{aligned}$$

since the inequality $\log\frac{1+x}{1-x} \geq 2x$ implies that

$$\sup_{|\mathbf{w}|<1}\frac{|\mathbf{w}|}{\log\frac{1+|\mathbf{w}|}{1-|\mathbf{w}|}} = \lim_{|\mathbf{w}|\to 0}\frac{|\mathbf{w}|}{\log\frac{1+|\mathbf{w}|}{1-|\mathbf{w}|}} = \frac{1}{2}.$$

This appears to be a reasonable bound when τ is small. Although it might appear to blow up when $\tau \to 1$, the CP condition $\alpha^2 + \tau^2 = 1$ implies that if $\tau \to 1$ then $\alpha \to 0$. Since $1 < \frac{1}{2\tau}\log\frac{1+\tau}{1-\tau} < \frac{1}{1-\tau^2}$ for all $\tau \in (0,1)$, we can conclude that $\alpha^2 < \eta^{\rm Riem}_{\rm BKM}(\Phi) < \frac{\alpha^2}{1-\tau^2}$ consistent with (47).

B.3.5 Dual of WY metric

For the WY function $\kappa_{\rm WY}(x) = 4/(1+\sqrt{x})^2$, the dual function $1/\kappa_{\rm WY}(x^{-1}) = (1+\sqrt{x})^2/4x$ gives the operator

$$\Omega_P^{\widehat{\rm WY}} = \tfrac{1}{4}\big(L_P^{-1/2} + R_P^{-1/2}\big)^2.$$

To proceed as above, for $P = I + \mathbf{w}\cdot\sigma$ we first find

$$\begin{aligned}
\big(L_P^{-1/2} + R_P^{-1/2}\big)(\mathbf{y}\cdot\sigma) &= (I+\mathbf{w}\cdot\sigma)^{-1/2}(\mathbf{y}\cdot\sigma) + (\mathbf{y}\cdot\sigma)(I+\mathbf{w}\cdot\sigma)^{-1/2} \\
&= \sqrt{\frac{2\zeta(|\mathbf{w}|)}{1-|\mathbf{w}|^2}}\Big[\tfrac{-1}{\zeta(|\mathbf{w}|)}(\mathbf{y}\cdot\mathbf{w})I + \mathbf{y}\cdot\sigma\Big],
\end{aligned} \tag{76}$$

where we have used (72) and the fact that the $\mathbf{w}\times\mathbf{y}$ terms have the opposite sign and cancel. As in Sections B.3.3 and B.3.4 the resulting expressions are difficult to deal with unless we make the simplifying assumption $\mathbf{y} = (y_1, 0, 0)^t$ and $w_1 = 0$ which yields the lower bound

$$\begin{aligned}
\eta^{\rm Riem}_{\widehat{WY}}(\Phi) &\geq \alpha^2\frac{1+\sqrt{1-\tau^2}}{1-\tau^2}\sup_{\mathbf{w}=(0,w_2,w_3)}\frac{1-|\mathbf{w}|^2}{1+\sqrt{1-|\mathbf{w}|^2}} \\
&= \alpha^2\frac{1+\sqrt{1-\tau^2}}{2(1-\tau^2)}.
\end{aligned}$$

B.3.6 Minimal or Bures metric

For the smallest function $\kappa_{\min}(x) = 2/(1+x)$ in $\mathcal{K}$ we begin by using (59) of Lemma B.3 to conclude that for $P = I + \mathbf{w}\cdot\sigma$

$$\mathrm{Tr}\,(\mathbf{y}\cdot\sigma)\frac{2}{L_P + R_P}(\mathbf{y}\cdot\sigma) = \mathbf{y}\cdot\left[\mathbf{y} + \frac{(\mathbf{w}\cdot\mathbf{y})\,\mathbf{w}}{1-|\mathbf{w}|^2}\right] = \frac{|\mathbf{y}|^2(1-|\mathbf{w}|^2) + (\mathbf{w}\cdot\mathbf{y})^2}{1-|\mathbf{w}|^2}.$$

Thus

$$\eta^{\mathrm{Riem}}_{\min}(\Phi) = \sup_{|\mathbf{w}|<1}\sup_{\mathbf{y}\neq 0}\frac{\alpha^2 y_1^2(1-|\mathbf{w}|^2)}{|\mathbf{y}|^2(1-|\mathbf{w}|^2)+(\mathbf{w}\cdot\mathbf{y})^2}\,\frac{1-\tau^2}{1-\tau^2-\alpha^2 w_1^2}.$$

For fixed $\mathbf{w}$ we can optimize the first term, which depends only on the ratios y_2/y_1 and y_3/y_1 (so we may assume $y_1 = 1$), directly or use Lemma B.2 to conclude that the maximum is achieved when $\mathbf{y} = \big(1, -w_1w_2/(1-w_1^2), -w_1w_3/(1-w_1^2)\big)^t$. In this case,

$$\frac{y_1^2(1-|\mathbf{w}|^2)}{|\mathbf{y}|^2(1-|\mathbf{w}|^2)+(\mathbf{w}\cdot\mathbf{y})^2} = 1 - w_1^2.$$

Thus

$$\eta^{\mathrm{Riem}}_{\min}(\Phi) = \alpha^2(1-\tau^2)\sup_{|w_1|<1}\frac{1-w_1^2}{1-\tau^2-\alpha^2 w_1^2}.$$

When the CP condition $\alpha^2+\tau^2 \leq 1$ holds, it is elementary to verify that

$$\frac{1-w_1^2}{1-\tau^2-\alpha^2 w_1^2} \leq \frac{1}{1-\tau^2} \tag{77}$$

so that the supremum is achieved when $w_1 = 0$, and

$$\eta^{\mathrm{Riem}}_{\min}(\Phi) = \alpha^2. \tag{78}$$

B.3.7 Extreme points

The extreme functions have the form

$$\kappa_s(x) = \frac{1+s}{2}\left(\frac{1}{x+s} + \frac{1}{1+sx}\right),$$

for which the corresponding operator is

$$\Omega^{\kappa_s}_P = \frac{1+s}{2}\left(\frac{1}{L_P + sR_P} + \frac{1}{R_P + sL_P}\right).$$

We let $P = I+\mathbf{w}\cdot\sigma$, $A = \mathbf{y}\cdot\sigma$ and apply Lemma B.5 and (65) with $\mathbf{u} = (1-s)\mathbf{w}$, $\mathbf{v} = (1+s)\mathbf{w}$ to obtain

$$\begin{aligned}\langle A, \Omega^{\kappa_s}_P(A)\rangle &= 2(1+s)^2\Big\langle \mathbf{y}, \big[\xi_s(|\mathbf{w}|^2)I - 4s|\mathbf{w}\rangle\langle\mathbf{w}|\big]^{-1}\mathbf{y}\Big\rangle \\ &= \frac{2(1+s)^2}{\xi_s(|\mathbf{w}|^2)}\left[|\mathbf{y}|^2 + \frac{4s(\mathbf{w}\cdot\mathbf{y})^2}{(1+s)^2(1-|\mathbf{w}|^2)}\right],\end{aligned}$$

where we have used (58) and

$$\xi_s(x) \equiv (1+s)^2 - (1-s)^2 x = (1+s)^2(1-x) + 4sx. \tag{79}$$

Thus

$$\frac{\langle \Phi(A), \Omega^{\kappa_s}_{\Phi(P)} \Phi(A)\rangle}{\langle A, \Omega^s_P A\rangle}$$
$$= \frac{(1-|\mathbf{w}|^2)\xi_s(|\mathbf{w}|^2)}{|\mathbf{y}|^2(1+s)^2(1-|\mathbf{w}|^2)+4s(\mathbf{w}\cdot\mathbf{y})^2}\ \frac{\alpha^2 y_1^2\big[(1+s)^2(1-\alpha^2 w_1^2-\tau^2)+4s\alpha^2 w_1^2\big]}{(1-\alpha^2 w_1^2-\tau^2)\,\xi_s(\alpha^2 w_1^2+\tau^2)}$$

so that

$$\eta^{\mathrm{Riem}}_{\kappa_s}(\Phi) \tag{80}$$
$$= \alpha^2 \sup_{|\mathbf{w}|<1}\left[\frac{(1-|\mathbf{w}|^2)\xi_s(|\mathbf{w}|^2)}{(1-\alpha^2 w_1^2-\tau^2)\,\xi_s(\alpha^2 w_1^2+\tau^2)}\ \sup_{\mathbf{y}\neq 0}\frac{y_1^2\big[(1+s)^2(1-\alpha^2 w_1^2-\tau^2)+4s\alpha^2 w_1^2\big]}{|\mathbf{y}|^2(1+s)^2(1-|\mathbf{w}|^2)+4s(\mathbf{w}\cdot\mathbf{y})^2}\right].$$

We first consider $\sup_{\mathbf{y}}$ with $\mathbf{w}$ fixed. This term depends only on the ratios y_2/y_1, y_3/y_1 so that there is no loss of generality in assuming $y_1 = 1$, in which case only the denominator depends on $\mathbf{y}$ and we consider instead its minimum, i.e., we seek

$$\min_{y_2,\,y_3}\big[(1+y_2^2+y_3^2)(1+s)^2(1-|\mathbf{w}|^2)+4s(w_1+w_2y_2+y_3y_3)^2\big]$$

which is found in Lemma B.2 with $\mu = (1+s)^2(1-|\mathbf{w}|^2)$ and $\nu = 4s$. The minimum is

$$\frac{(1+s)^2(1-|\mathbf{w}|^2)\xi_s(|\mathbf{w}|^2)}{\xi_s(|\mathbf{w}|^2)-4sw_1^2}. \tag{81}$$

Inserting (81) into (80) yields

$$\eta^{\mathrm{Riem}}_{\kappa_s}(\Phi) = \alpha^2 \sup_{|\mathbf{w}|<1}\frac{\big[(1+s)^2(1-\alpha^2 w_1^2-\tau^2)+4s\alpha^2 w_1^2\big]\,\big[\xi_s(|\mathbf{w}|^2)-4sw_1^2\big]}{(1-\alpha^2 w_1^2-\tau^2)\,\xi_s(\alpha^2 w_1^2+\tau^2)\,(1+s)^2}.$$

The only term in the expression above which involves $\mathbf{w}$ rather than w_1 is $\xi_s(|\mathbf{w}|^2) = (1+s)^2-(1-s)^2|\mathbf{w}|^2$, which is largest when $|\mathbf{w}|$ is smallest, i.e., $|\mathbf{w}| = |w_1|$ or, equivalently, $\mathbf{w} = (w_1, 0, 0)$. Thus we find

$$\eta^{\mathrm{Riem}}_{\kappa_s}(\Phi) = \alpha^2 \sup_{|w_1|<1}\frac{(1+s)^2(1-\alpha^2 w_1^2-\tau^2)+4s\alpha^2 w_1^2}{(1+s)^2(1-\alpha^2 w_1^2-\tau^2)+4s(\alpha^2 w_1^2+\tau^2)}\ \frac{1-w_1^2}{1-\alpha^2 w_1^2-\tau^2}. \tag{82}$$

For $s = 0, 1$ this reduces to the expression for the maximal and minimal functions, (71) and (77), respectively. As in (77) the second factor is largest when $w_1 = 0$. The first factor can be written as

$$\frac{R}{R+4s\tau^2} = 1 - \frac{4s\tau^2}{R+4s\tau^2},$$

which is largest when

$$R \equiv (1+s)^2(1-\alpha^2 w_1^2-\tau^2)+4s\alpha^2 w_1^2 = (1+s)^2(1-\tau^2)-(1-s)^2\alpha^2 w_1^2$$

is largest, which also occurs when $w_1^2 = 0$. Using these observations in (82) we can conclude that

$$\eta^{\mathrm{Riem}}_{\kappa_s}(\Phi) = \frac{(1+s)^2\alpha^2}{(1+s)^2-(1-s)^2\tau^2} = \frac{\alpha^2}{1-\left(\frac{1-s}{1+s}\right)^2\tau^2}. \tag{83}$$

When $s = 0, 1$ we recover the expressions (71) and (78).

B.3.8 Wigner-Yanase metric

Although we have had to make simplifying assumptions to obtain lower bounds for all but the extremal κ, it is quite remarkable that we can obtain an exact expression in the case of the Wigner-Yanase function $\kappa_{\text{WY}}(x) = 4/(1+\sqrt{x})^2$. Then

$$\Omega_P^{\text{WY}} = \frac{4}{\left(\sqrt{L_P}+\sqrt{R_P}\right)^2}.$$

For $P = I + \mathbf{w}\cdot\sigma$, using (57) and Lemma B.3 we can write

$$\frac{2}{\sqrt{L_P}+\sqrt{R_P}}(\mathbf{y}\cdot\sigma) = \sqrt{\frac{2}{\zeta}}\left[-\frac{\zeta\mathbf{w}\cdot\mathbf{y}}{\zeta^2-|\mathbf{w}|^2}I + \left(\mathbf{y} + \frac{(\mathbf{w}\cdot\mathbf{y})\mathbf{w}}{\zeta^2-|\mathbf{w}|^2}\right)\cdot\sigma\right]$$

with $\zeta = \zeta(|\mathbf{w}|) = 1 + \sqrt{1-|\mathbf{w}|^2}$. Therefore,

$$\begin{aligned}
\operatorname{Tr}(\mathbf{y}\cdot\sigma)\frac{4}{\left(\sqrt{L_P}+\sqrt{R_P}\right)^2}(\mathbf{y}\cdot\sigma) &= \operatorname{Tr}\left[\frac{2}{\sqrt{L_P}+\sqrt{R_P}}(\mathbf{y}\cdot\sigma)\right]^2 \\
&= \frac{4}{\zeta}\left[\frac{\zeta^2(\mathbf{w}\cdot\mathbf{y})^2}{(\zeta^2-|\mathbf{w}|^2)^2} + |\mathbf{y}|^2 + 2\,\frac{(\mathbf{w}\cdot\mathbf{y})^2}{\zeta^2-|\mathbf{w}|^2} + \frac{(\mathbf{w}\cdot\mathbf{y})^2|\mathbf{w}|^2}{(\zeta^2-|\mathbf{w}|^2)^2}\right] \\
&= \frac{4}{\zeta}\left[|\mathbf{y}|^2 + (\mathbf{w}\cdot\mathbf{y})^2\,\frac{3\zeta^2-|\mathbf{w}|^2}{(\zeta^2-|\mathbf{w}|^2)^2}\right] \\
&= 4\left\langle\mathbf{y},\left[\zeta I - (2-\zeta^{-1})|\mathbf{w}\rangle\langle\mathbf{w}|\right]^{-1}\mathbf{y}\right\rangle,
\end{aligned}$$

where the last equality is the key to our ability to evaluate $\eta^{\text{Riem}}_{\text{WY}}(\Phi)$ exactly. To obtain this, one can apply $|\mathbf{w}|^2 = \zeta(2-\zeta)$ and (58) to see that

$$\begin{aligned}
\frac{1}{\zeta}\left[I + \frac{3\zeta^2-|\mathbf{w}|^2}{(\zeta^2-|\mathbf{w}|^2)^2}\,|\mathbf{w}\rangle\langle\mathbf{w}|\right] &= \frac{1}{\zeta}\left[I + \frac{2\zeta-1}{2\zeta(\zeta-1)^2}\,|\mathbf{w}\rangle\langle\mathbf{w}|\right] \\
&= \left[\zeta I - (2-\zeta^{-1})|\mathbf{w}\rangle\langle\mathbf{w}|\right]^{-1}.
\end{aligned}$$

Then with $T \equiv \operatorname{diag}(\alpha,0,0)$, $\widetilde{\mathbf{w}} \equiv (\alpha w_1, 0, \tau)^t$ and $\widetilde{\zeta} \equiv \zeta(|\widetilde{\mathbf{w}}|)$, we use Lemma B.1 to obtain

$$\begin{aligned}
\eta^{\text{Riem}}_{\text{WY}}(\Phi) &= \sup_{|\mathbf{w}|<1}\sup_{\mathbf{y}\neq 0}\frac{\left\langle T\mathbf{y},\left[\widetilde{\zeta}I-(2-\widetilde{\zeta}^{-1})|\widetilde{\mathbf{w}}\rangle\langle\widetilde{\mathbf{w}}|\right]^{-1}T\mathbf{y}\right\rangle}{\left\langle\mathbf{y},\left[\zeta I-(2-\zeta^{-1})|\mathbf{w}\rangle\langle\mathbf{w}|\right]^{-1}\mathbf{y}\right\rangle} \\
&= \sup_{|\mathbf{w}|<1}\sup_{\mathbf{y}\neq 0}\frac{\left\langle T\mathbf{y},\left[\zeta I-(2-\zeta^{-1})|\mathbf{w}\rangle\langle\mathbf{w}|\right]T\mathbf{y}\right\rangle}{\left\langle\mathbf{y},\left[\widetilde{\zeta}I-(2-\widetilde{\zeta}^{-1})|\widetilde{\mathbf{w}}\rangle\langle\widetilde{\mathbf{w}}|\right]\mathbf{y}\right\rangle} \\
&= \sup_{|\mathbf{w}|<1}\sup_{\mathbf{y}\neq 0}\frac{\zeta\,|T\mathbf{y}|^2-(2-\zeta^{-1})|\langle\mathbf{w},T\mathbf{y}\rangle|^2}{\widetilde{\zeta}\,|\mathbf{y}|^2-(2-\widetilde{\zeta}^{-1})|\langle\widetilde{\mathbf{w}},\mathbf{y}\rangle|^2} \\
&= \sup_{|\mathbf{w}|<1}\sup_{(y_1,y_3)\neq(0,0)}\frac{\alpha^2\left[\zeta-w_1^2(2-\zeta^{-1})\right]y_1^2}{\widetilde{\zeta}(y_1^2+y_3^2)-(2-\widetilde{\zeta}^{-1})(\alpha w_1y_1+\tau y_3)^2}.
\end{aligned}$$

Here we may assume that $\alpha > 0$ and so $\tau < 1$. Then it is easily verified that $\widetilde{\zeta}-\tau^2(2-\widetilde{\zeta}^{-1}) > 0$. Since the denominator of the last ratio is

$$\left[\widetilde{\zeta}-\tau^2(2-\widetilde{\zeta}^{-1})\right]\left(y_3-\frac{\alpha\tau w_1(2-\widetilde{\zeta}^{-1})y_1}{\widetilde{\zeta}-\tau^2(2-\widetilde{\zeta}^{-1})}\right)^2 + \frac{\widetilde{\zeta}^2-(\alpha^2w_1^2+\tau^2)(2\widetilde{\zeta}-1)}{\widetilde{\zeta}-\tau^2(2-\widetilde{\zeta}^{-1})}\,y_1^2,$$

it follows that

$$\eta_{\rm WY}^{\rm Riem}(\Phi) = \sup_{|\mathbf{w}|<1} \frac{\alpha^2\big[\zeta - w_1^2(2-\zeta^{-1})\big]\big[\widetilde{\zeta} - \tau^2(2-\widetilde{\zeta}^{-1})\big]}{\widetilde{\zeta}^2 - (\alpha^2 w_1^2 + \tau^2)(2\widetilde{\zeta}-1)}.$$

Then $\widetilde{\zeta} = 1 + \sqrt{1-\tau^2-\alpha^2 w_1^2}$ depends on w_1 only and

$$\begin{aligned}\zeta - w_1^2(2-\zeta^{-1}) &= 1 + \sqrt{1-|\mathbf{w}|^2} - w_1^2\left(2 - \frac{1-\sqrt{1-|\mathbf{w}|^2}}{|\mathbf{w}|^2}\right) \\ &= 2(1-w_1^2) - \frac{(w_2^2+w_3^2)\big(1-\sqrt{1-|\mathbf{w}|^2}\big)}{|\mathbf{w}|^2},\end{aligned}$$

which obviously takes the maximum $2(1-w_1^2)$ when $w_2 = w_3 = 0$. Therefore,

$$\begin{aligned}\eta_{\rm WY}^{\rm Riem}(\Phi_{\alpha,\tau}) &= \sup_{|w_1|<1} \frac{2\alpha^2(1-w_1^2)\Big[1+\sqrt{1-\tau^2-\alpha^2w_1^2} - \tau^2\Big(2 - \frac{1}{1+\sqrt{1-\tau^2-\alpha^2w_1^2}}\Big)\Big]}{\big(1+\sqrt{1-\tau^2-\alpha^2w_1^2}\big)^2 - (\alpha^2w_1^2+\tau^2)\big(1+2\sqrt{1-\tau^2-\alpha^2w_1^2}\big)} \\ &= \sup_{|w_1|<1} \frac{\alpha^2(1-w_1^2)\big[2(1-\tau^2) - \alpha^2w_1^2 + 2(1-\tau^2)\sqrt{1-\tau^2-\alpha^2w_1^2}\big]}{(1-\tau^2-\alpha^2w_1^2)\big(1+\sqrt{1-\tau^2-\alpha^2w_1^2}\big)^2} \\ &\le \alpha^2\left[\sup_{|w_1|<1} \frac{1-w_1^2}{1-\tau^2-\alpha^2w_1^2}\right]\left[\sup_{|w_1|<1} \frac{2(1-\tau^2)\big(1+\sqrt{1-\tau^2-\alpha^2w_1^2}\big) - \alpha^2w_1^2}{\big(1+\sqrt{1-\tau^2-\alpha^2w_1^2}\big)^2}\right].\end{aligned}$$

As in (77) the first supremum is attained when $w_1 = 0$. For the second, let $\rho \equiv 1-\tau^2$ and $x \equiv \alpha^2 w_1^2 \in [0,\alpha^2)$ where $\alpha^2 \le \rho < 1$, and observe that the ratio can be written as

$$\frac{2\rho(1+\sqrt{\rho-x}) - x}{(1+\sqrt{\rho-x})^2} = \rho + (1-\rho)\frac{\rho-x}{(1+\sqrt{\rho-x})^2},$$

which is maximized when $x = 0$ (i.e., $w_1 = 0$). Thus we conclude

$$\sup_{|w_1|<1} \frac{2(1-\tau^2)\big(1+\sqrt{1-\tau^2-\alpha^2w_1^2}\big) - \alpha^2w_1^2}{\big(1+\sqrt{1-\tau^2-\alpha^2w_1^2}\big)^2} = \rho + (1-\rho)\frac{\rho}{(1+\sqrt{\rho})^2} = \frac{2(1-\tau^2)}{1+\sqrt{1-\tau^2}}.$$

Combining the two suprema yields

$$\eta_{\rm WY}^{\rm Riem}(\Phi_{\alpha,\tau}) \le \frac{2\alpha^2}{1+\sqrt{1-\tau^2}},$$

It is straightforward to see that the special case $\mathbf{y} = (y_1,0,0)^t$ and $\mathbf{w} = 0$ yields the reverse inequality.

The proof of Theorem 6.2 is now complete. □

B.4 Proof of Theorem 6.6

For $P = I + \mathbf{w}\cdot\sigma$, $Q = I + \mathbf{x}\cdot\sigma$ and $0 \le s \le 1$, by (67) and Lemma B.5 we have

$$\frac{H_{g_s}(Q,P)}{2(1+s)} = \frac{1}{2(1+s)}\,\mathrm{Tr}\,(\mathbf{y}\cdot\sigma)\frac{1}{L_P + sR_Q}(\mathbf{y}\cdot\sigma)$$

$$= \Big\langle \mathbf{y}, \big[\{(1+s)^2 - |\mathbf{u}|^2\}I + |\mathbf{u}\rangle\langle\mathbf{u}| - |\mathbf{v}\rangle\langle\mathbf{v}|\big]^{-1}\mathbf{y}\Big\rangle,$$

where $\mathbf{y} = \mathbf{w} - \mathbf{x}$, $\mathbf{u} = \mathbf{w} - s\mathbf{x}$ and $\mathbf{v} = \mathbf{w} + s\mathbf{x}$, and we note that $\mathbf{y}$ is orthogonal to $\mathbf{u} \times \mathbf{v}$. The formula for $H_{g_s}(P,Q)$ is similar with $\mathbf{w}$, $\mathbf{x}$ interchanged. Since the first inequality of (49) holds in general by (16), we will estimate $\eta^{\mathrm{RelEnt}}_{(g_s)_{\mathrm{sym}}}(\Phi)$ for $\Phi = \Phi_{\alpha,\tau}$. For this we take $\mathbf{w} = (w_1, 0, 0)^t$ and $\mathbf{x} = 0$ for simplicity, for which we have

$$\frac{H_{g_s}(Q,P)}{2(1+s)} = \frac{w_1^2}{(1+s)^2 - w_1^2}, \qquad \frac{H_{g_s}(P,Q)}{2(1+s)} = \frac{w_1^2}{(1+s)^2 - s^2 w_1^2}. \tag{84}$$

Since $\Phi(P) = I + \widetilde{\mathbf{w}}\cdot\sigma$ and $\Phi(Q) = I + \widetilde{\mathbf{x}}\cdot\sigma$ where $\widetilde{\mathbf{w}} = (\alpha w_1, 0, \tau)^t$ and $\widetilde{\mathbf{x}} = (0,0,\tau)^t$, we have the expression

$$\frac{H_{g_s}(\Phi(Q),\Phi(P))}{2(1+s)} = \Big\langle \widetilde{\mathbf{y}}, \big[\{(1+s)^2 - |\widetilde{\mathbf{u}}|^2\}I + |\widetilde{\mathbf{u}}\rangle\langle\widetilde{\mathbf{u}}| - |\widetilde{\mathbf{v}}\rangle\langle\widetilde{\mathbf{v}}|\big]^{-1}\widetilde{\mathbf{y}}\Big\rangle, \tag{85}$$

where $\widetilde{\mathbf{y}} = \widetilde{\mathbf{w}} - \widetilde{\mathbf{x}} = (\alpha w_1, 0, 0)^t$, $\widetilde{\mathbf{u}} = \widetilde{\mathbf{w}} - s\widetilde{\mathbf{x}} = (\alpha w_1, 0, (1-s)\tau)^t$ and $\widetilde{\mathbf{v}} = \widetilde{\mathbf{w}} + s\widetilde{\mathbf{x}} = (\alpha w_1, 0, (1+s)\tau)^t$. The matrix form of the operator inside $[\quad]^{-1}$ of (85) is

$$\begin{pmatrix} \xi_s(\tau^2) - \alpha^2 w_1^2 & 0 & -2s\alpha\tau w_1 \\ 0 & \xi_s(\tau^2) - \alpha^2 w_1^2 & 0 \\ -2s\alpha\tau w_1 & 0 & \xi_s(\tau^2) - 4s\tau^2 - \alpha^2 w_1^2 \end{pmatrix},$$

where $\xi_s(\cdot)$ is in (79). The $(1,1)$-entry of the inverse of this matrix is $\big[\xi_s(\tau^2) - 4s\tau^2 - \alpha^2 w_1^2\big]/\det$ where det is the determinant of the 2×2 matrix of the first and the third rows and columns. Therefore, the exact form of (85) is

$$\frac{H_{g_s}(\Phi(Q),\Phi(P))}{2(1+s)} = \frac{\alpha^2 w_1^2\big[\xi_s(\tau^2) - \alpha^2 w_1^2 - 4s\tau^2\big]}{\big[\xi_s(\tau^2) - \alpha^2 w_1^2\big]\big[\xi_s(\tau^2) - \alpha^2 w_1^2 - 4s\tau^2\big] - 4s^2\alpha^2\tau^2 w_1^2}. \tag{86}$$

A similar computation with $\widetilde{\mathbf{w}}, \widetilde{\mathbf{x}}$ interchanged yields

$$\frac{H_{g_s}(\Phi(P),\Phi(Q))}{2(1+s)} = \frac{\alpha^2 w_1^2\big[\xi_s(\tau^2) - s^2\alpha^2 w_1^2 - 4s\tau^2\big]}{\big[\xi_s(\tau^2) - s^2\alpha^2 w_1^2\big]\big[\xi_s(\tau^2) - s^2\alpha^2 w_1^2 - 4s\tau^2\big] - 4s^2\alpha^2\tau^2 w_1^2}. \tag{87}$$

We define

$$H(s) \equiv \lim_{|w_1|\nearrow 1} \frac{H_{g_s}(Q,P) + H_{g_s}(P,Q)}{2(1+s)},$$
$$\widetilde{H}(s) \equiv \lim_{|w_1|\nearrow 1} \frac{H_{g_s}(\Phi(Q),\Phi(P)) + H_{g_s}(\Phi(P),\Phi(Q))}{2(1+s)}.$$

Since $\eta^{\mathrm{RelEnt}}_{(g_s)_{\mathrm{sym}}}(\Phi) \ge \widetilde{H}(s)/H(s)$ for every $s \in [0,1]$, we may compare $\widetilde{H}(s)/H(s)$ with $\eta^{\mathrm{Riem}}_{\kappa_s}(\Phi)$ for s near 1. For this we observe by (84), (86) and (87) that $H(1) = 2/3$ and

$$\widetilde{H}(1) = \frac{2\alpha^2(4 - \alpha^2 - 4\tau^2)}{(4-\alpha^2)(4-\alpha^2-4\tau^2) - 4\alpha^2\tau^2} = \frac{2\alpha^2(4-\alpha^2-4\tau^2)}{(4-\alpha^2)^2 - 16\tau^2}.$$

Now assume that $4\tau^2 > (1-\alpha^2)(4-\alpha^2)$ (in particular, $\alpha^2 > 0$). Since $(4-\alpha^2)^2 - 16\tau^2 > 0$ (thanks to $\alpha^2 + \tau^2 \le 1$ and $\tau^2 < 1$) and

$$3(4-\alpha^2-4\tau^2) - \big[(4-\alpha^2)^2 - 16\tau^2\big] = 4\tau^2 - (1-\alpha^2)(4-\alpha^2) > 0,$$

it follows that $\widetilde{H}(1)/H(1) > \alpha^2$. From the continuity of the s-dependence of $\widetilde{H}(s)/H(s)$ and $\eta_{\kappa_s}^{\mathrm{Riem}}(\Phi)$ in (83), we arrive at the first assertion stated in the theorem.

For the second assertion, when $\alpha^2 + \tau^2 = 1$, a tedious computation gives

$$\frac{\widetilde{H}(s)}{H(s)} = \alpha^2 \frac{s(s+2)(2s+1)\bigl[12s(s+1)^2 + (2s^4 + s^3 + s + 2)\alpha^2\bigr]}{(s^2+4s+1)\bigl[4s(s+1) + s^3\alpha^2\bigr]\bigl[4s(s+1) + \alpha^2\bigr]}.$$

The limit of $\bigl[\widetilde{H}(s)/H(s)\bigr]/\eta_{\kappa_s}^{\mathrm{Riem}}(\Phi)$ as $\alpha^2 = 1 - \tau^2$ tends to 0 is

$$\frac{3s(s+2)(2s+1)}{(s^2+4s+1)(s+1)^2}.$$

The numerator minus the denominator of the above ratio is $-(s^4 - 5s^2 + 1)$, which is positive when $s^2 > \frac{5-\sqrt{21}}{2}$. This yields the second assertion of the theorem. □

References

[1] T. Ando, *Topics on Operator Inequalities*, Lecture notes (mimeographed), Hokkaido Univ., Sapporo, 1978.

[2] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.* **26** (1979) 203–241.

[3] T. Ando and F. Hiai, Operator log-convex functions and operator means, *Math. Ann.* **350** (2011), 611–630.

[4] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States. An Introduction to Quantum Entanglement*, Cambridge Univ. Press, Cambridge, 2006.

[5] R. Bhatia, *Matrix Analysis*, Springer, New York, 1996.

[6] R. Bhatia, *Positive Definite Matrices*, Princeton Univ. Press, Princeton, 2007.

[7] D. Bures, An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w^*-algebras, *Trans. Amer. Math. Soc.* **135**, (1969), 199–212.

[8] M.-D. Choi, Some assorted inequalities for positive linear maps on C^*-algebras, *J. Operator Theory* **4** (1980), 271–285.

[9] M.-D. Choi, M. B. Ruskai and E. Seneta, Equivalence of certain entropy contraction coefficients, *Linear Algebra Appl.* **208/209** (1994), 29–36.

[10] J. E. Cohen, Y. Iwasa, Gh. Rautu, M. B. Ruskai, E. Seneta and Gh. Zbaganu, Relative entropy under mappings by stochastic matrices, *Linear Algebra Appl.* **179** (1993), 211–235.

[11] J. E. Cohen, J. H. B Kemperman and Gh. Zbăganu, *Comparisons of Stochastic Matrices with Applications in Information Theory, Statistics, Economics and Population Sciences*, Brikhäuser, Boston, 1998.

[12] R. L. Dobrushin, Central limit theorem for nonstationary Markov chains. I, II, *Theory Probab. Appl.* **1** (1956), 65–80; 329–383.

[13] W. F. Donoghue, Jr., *Monotone Matrix Functions and Analytic Continuation*, Springer, Berlin-Heidelberg-New York, 1974.

[14] U. Franz, F. Hiai and É. Ricard, Higher order extension of Löwner's theory: Operator k-tone functions, *Trans. Amer. Math. Soc.* **366** (2014), 3043–3074.

[15] S. Friedland and W. So, On the product of matrix exponentials, *Linear Algebra Appl.* **196** (1994), 193–205.

[16] A. Fujiwara and P. Algoet, One-to-one parametrization of quantum channels, *Phys. Rev. A* **59** (1990), 3290–3294.

[17] P. Gibilisco and T. Isola, Wigner-Yanase information on quantum state space: the geometric approach, *J. Math. Phys.* **44** (2003), 3752–3762.

[18] H. Hasegawa, α-divergence of the non-commutative information geometry, *Rep. Math. Phys.* **33** (1993) 87–93.

[19] F. Hansen, Trace functions as Laplace transforms, *J. Math. Phys.* **47** (2006), 043504, 11 pp.

[20] F. Hiai, Equality cases in matrix norm inequalities of Golden-Thompson type, *Linear Multilinear Algebra* **36** (1994), 239–249.

[21] F. Hiai, Matrix Analysis: Matrix Monotone Functions, Matrix Means, and Majorization, *Interdisciplinary Information Sciences* **16** (2010), 139–248.

[22] F. Hiai, H. Kosaki, D. Petz and M. B. Ruskai, Families of completely positive maps associated with monotone metrics, *Linear Algebra Appl.* **439** (2013), 1749–1791.

[23] F. Hiai, M. Mosonyi, D. Petz and C. Bény, Quantum f-divergences and error correction, *Rev. Math. Phys.* **23** (2011), 691–747.

[24] F. Hiai and D. Petz, Riemannian metrics on positive definite matrices related to means, *Linear Algebra Appl.* **430** (2009), 3105–3130.

[25] F. Hiai and D. Petz, From quasi-entropy to various quantum information quantities, *Publ. Res. Inst. Math. Sci.* **48** (2012), 525–542.

[26] F. Hiai and D. Petz, Convexity of quasi-entropy type functions: Lieb's and Ando's convexity theorems revisited, *J. Math. Phys.* **54** (2013), 062201.

[27] A. S. Holevo, Coding theorems for quantum channels, *Russian Math. Surveys* **53** (1999), 1295–1331; arXiv:quant-ph/9809023.

[28] M. Horodecki, P. W. Shor and M. B. Ruskai, Entanglement breaking channels, *Rev. Math. Phys.* **15** (2003), 629–641.

[29] A. Jenčová and M. B. Ruskai, A unified treatment of convexity of relative entropy and related trace functions, with conditions for equality, *Rev. Math. Phys.* **22** (2010), 1099–1121.

[30] R. Kadison, A generalized Schwarz inequality and algebraic invariants for operator algebras, *Ann. of Math.* **56** (1952), 494–503.

[31] M. J. Kastoryano and K. Temme, private communication.

[32] M. J. Kastoryano and K. Temme, Quantum logarithmic Sobolev inequalities and rapid mixing, *J. Math. Phys.* **54** (2013), 052202.

[33] J. Kiefer, Optimum experimental designs, *J. R. Stat. Soc. Ser. B (Methodol.)* **21** (1959), 272–310.

[34] C. King and M. B. Ruskai, Minimal entropy of states emerging from noisy quantum channels *IEEE Trans. Inform. Theory* **47** (2001), 192–209.

[35] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. 1, Wiley Interscience, New York-London, 1963.

[36] H. Kosaki, Interpolation theory and the Wigner-Yanase-Dyson-Lieb concavity, *Comm. Math. Phys.* **87** (1982), 315–329.

[37] F. Kraus, Über konvexe Matrixfunktionen, *Math. Z.* **41** (1936), 18–42.

[38] A. Lesniewski and M. B. Ruskai, Monotone Riemannian metrics and relative entropy on noncommutative probability spaces, *J. Math. Phys.* **40** (1999), 5702–5724.

[39] E. H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Adv. Math.* **11** (1973), 267–288.

[40] E. H. Lieb and M. B. Ruskai, Some operator inequalities of the Schwarz type, *Adv. Math.* **12** (1974), 269–273.

[41] K. Löwner, Über monotone Matrix Funktionen, *Math. Z.* **38** (1934), 177–216.

[42] E. A. Morozova and N. N. Chentsov, Markov invariant geometry on state manifolds (Russian), *Itogi Nauki i Tekhniki* **36**, 69–102 (1989); Translated in *J. Soviet Math.* **56** (1991), 2648–2669.

[43] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

[44] D. Petz, Quasi-entropies for states of a von Neumann algebra, Publ. RIMS. Kyoto Univ. **21** (1985), 781–800.

[45] D. Petz, Quasi-entropies for finite quantum systems, *Rep. Math. Phys.* **23** (1986), 57–65.

[46] D. Petz, Monotone metrics on matrix spaces, *Linear Algebra Appl.* **244** (1996), 81–96.

[47] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer, Berlin-Heidelberg, 2008.

[48] M. B. Ruskai, Beyond strong subadditivity? Improved bounds on the contraction of generalized relative entropy, *Rev. Math. Phys.* **6** (1994), 1147–1161.

[49] M. B. Ruskai, S. Szarek and E. Werner, An analysis of completely positive trace-preserving maps on $\mathcal{M}_2$, *Linear Algebra Appl.* **347** (2002), 159–187.

[50] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf and F. Verstraete, The χ^2-divergence and mixing times of quantum Markov Processes, *J. Math. Phys.* **51** (2010), 122201.

[51] M. Tomamichel, R. Colbeck and R. Renner, A fully quantum asymptotic equipartition property, *IEEE Trans. Inform Theory* **55** (2009) 5840–5847.

[52] A. Uhlmann, Density operators an an arena for differential geometry, *Rep. Math. Phys.* **33** (1993), 253–263.

[53] A. Uhlmann, Geometric phases and related structures, *Rep. Math. Phys.* **36** (1995), 461–481.