Multivariate Fidelities

Theshani Nuradha* Hemant K. Mishra* Felix Leditzky†
Mark M. Wilde*

Abstract

The bivariate classical fidelity is a widely used measure of the similarity of two probability distributions. There exist a few extensions of the notion of the bivariate classical fidelity to more than two probability distributions; herein we call these extensions multivariate classical fidelities, with some examples being the Matusita multivariate fidelity and the average pairwise fidelity. Hitherto, quantum generalizations of multivariate classical fidelities have not been systematically explored, even though there are several well known generalizations of the bivariate classical fidelity to quantum states, such as the Uhlmann and Holevo fidelities. The main contribution of our paper is to introduce a number of multivariate quantum fidelities and show that they satisfy several desirable properties that are natural extensions of those of the Uhlmann and Holevo fidelities. We propose three variants that reduce to the average pairwise fidelity for commuting states: average pairwise z-fidelities, the multivariate semi-definite programming (SDP) fidelity, and a multivariate fidelity inspired by an existing secrecy measure. The second one is obtained by extending the SDP formulation of the Uhlmann fidelity to more than two states. All three of these variants satisfy the following properties: (i) reduction to multivariate classical fidelities for commuting states, (ii) the data-processing inequality, (iii) invariance under permutations of the states, (iv) its values are in the interval $[0,1]$; they are faithful, that is, their values are equal to one if and only if all the states are equal, and they satisfy orthogonality, that is their values are equal to zero if and only if the states are mutually orthogonal to each other, (v) direct-sum property, (vi) joint concavity, and (vii) uniform continuity bounds under certain conditions. Furthermore, we establish inequalities relating these different variants, indeed clarifying that all these definitions coincide with the average pairwise fidelity for commuting states. Lastly, we introduce another multivariate fidelity called multivariate log-Euclidean fidelity, which is a quantum generalization of the Matusita multivariate fidelity. We also show that it satisfies most of the desirable properties listed above, it is a function of a multivariate log-Euclidean divergence, and has an operational interpretation in terms of quantum hypothesis testing with an arbitrarily varying null hypothesis.

*School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14850, USA (pt388@cornell.edu, hemant.mishra@cornell.edu, wilde@cornell.edu).

†Department of Mathematics & Illinois Quantum Information Science and Technology (IQUIST) Center, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA (leditzky@illinois.edu)

Contents

1 Introduction

Distinguishability and similarity are essential concepts that hold significance across all scientific disciplines. The fundamental toolkit for understanding these concepts revolves around measures of distinguishability and similarity. This has led to several measures of distinguishability between probability distributions in classical information theory, such as the Kullback–Leibler divergence [KL51] and Rényi relative entropy [Rén61], and similarity measures such as the Bhattacharyya overlap [Bha46]. In quantum information theory, we have distinguishability measures between quantum states such as quantum relative entropy [Ume62], Petz–Rényi relative entropy [Pet86], and sandwiched Rényi relative entropy [MLDS+13, WWY14], and information quantities derived from these measures, such as quantum mutual information [Str65]. Moreover, the Uhlmann [Uhl76] and Holevo [Hol72] fidelities are similarity measures between two quantum states. All of the aforementioned distinguishability and similarity measures take into account two probability distributions in the classical case and two quantum states in the quantum case.

There exist a few extensions of the notion of the Bhattacharyya overlap or bivariate classical fidelity to more than two probability distributions. Examples include the Matusita affinity [Mat67] and the average pairwise classical fidelity; herein we call all such extensions multivariate classical fidelities. Hitherto, quantum generalizations of multivariate classical fidelities have not been systematically explored. Existing multivariate measures that do not reduce to classical multivariate fidelities are as follows: Holevo information of a tuple of states with uniform prior characterizes the correlations between the classical system and the quantum system (see (5.105) for a precise definition). To this end, Holevo information being approximately zero implies that the states in this tuple are similar. Multivariate Chernoff divergence is defined in [MNW23, Definition 3] for commuting states while discussing generalizations to non-commuting states. Moreover, multivariate generalizations for Rényi divergences have been explored in [MBV22] for general states and in [FFHT23] for commuting states.

The main contribution of our paper is to introduce quantum generalizations of the Matusita affinity and average pairwise classical fidelity; here we call them multivariate quantum fidelities. We show that these multivariate fidelities satisfy several desirable properties that are natural generalizations of those of bivariate fidelities. Consequently, the multivariate quantum fidelities are bona fide measures of similarity between multiple states. See Section 1.2 for a more detailed discussion of our contributions.

1.1 Classical and quantum bivariate fidelities

We first provide a brief overview of classical and quantum fidelities.

The bivariate classical fidelity quantifies the similarity of two discrete probability distributions p and q on a finite set $\mathcal{X}$; it is defined as

$$F(p,q) \coloneqq \sum_{x\in\mathcal{X}} \sqrt{p(x)q(x)}. \tag{1.1}$$

It satisfies $0 \le F(p,q) \le 1$ for all distributions p and q, it is equal to one if and only if $p = q$, and it is equal to zero if and only if the supports of p and q are disjoint. The above formulation can be naturally extended to commuting states as follows: for commuting states ρ and σ, define

$$F(\rho,\sigma) \coloneqq \sum_{x} \sqrt{\rho(x)\sigma(x)}, \tag{1.2}$$

where $(\rho(x))_x$ and $(\sigma(x))_x$ are the tuples of eigenvalues of ρ and σ, respectively, in their common eigenbasis.

In the same spirit of distinguishability measures, there is an infinite number of similarity measures that generalize the bivariate classical fidelity to quantum states. Here we call them z-fidelities and define them for $z \ge 1/2$ and quantum states ρ and σ as

$$F_z(\rho,\sigma) \coloneqq \mathrm{Tr}\left[\left(\sigma^{\frac{1}{4z}}\rho^{\frac{1}{2z}}\sigma^{\frac{1}{4z}}\right)^z\right] = \left\|\rho^{\frac{1}{4z}}\sigma^{\frac{1}{4z}}\right\|_{2z}^{2z}. \tag{1.3}$$

For every positive integer z, cyclicity of trace leads to the following expression:

$$F_z(\rho,\sigma) = \mathrm{Tr}\left[\left(\rho^{1/2z}\sigma^{1/2z}\right)^z\right]. \tag{1.4}$$

The z-fidelities are obtained by fixing $\alpha = 1/2$ in the α-z Rényi relative entropies [AD15]. They reduce to the classical fidelity for commuting states, and each of them satisfies the data-processing inequality [Zha20, Theorem 1.2]; i.e., for all $z \ge 1/2$ and for every quantum channel $\mathcal{N}$ and pair of states ρ and σ, the following inequality holds:

$$F_z(\rho,\sigma) \le F_z(\mathcal{N}(\rho),\mathcal{N}(\sigma)). \tag{1.5}$$

The z-fidelities are monotonically decreasing and continuous in z [LT15, Proposition 6]. Notably, for $z = 1/2$, the z-fidelity reduces to the Uhlmann fidelity [Uhl76]:

$$F_{1/2}(\rho,\sigma) = F(\rho,\sigma) \coloneqq \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1, \tag{1.6}$$

which was expounded upon in [Joz94]. For $z = 1$, the z-fidelity reduces to the Holevo fidelity [Hol72]:

$$F_1(\rho,\sigma) = F_H(\rho,\sigma) \coloneqq \mathrm{Tr}[\sqrt{\rho}\sqrt{\sigma}]. \tag{1.7}$$

The log-Euclidean fidelity is defined as the $z \to \infty$ limit of (1.3), for which it is possible to obtain a closed-form expression by an application of the Lie–Trotter product formula. Indeed, for $\varepsilon > 0$, define $\rho(\varepsilon) \coloneqq \rho + \varepsilon I$ and $\sigma(\varepsilon) \coloneqq \sigma + \varepsilon I$. Then

$$\begin{aligned} F_\flat(\rho, \sigma) \coloneqq \lim_{z\to\infty} F_z(\rho, \sigma) &= \inf_{\varepsilon>0} \operatorname{Tr}\left[\exp\left(\frac{1}{2}(\ln \rho(\varepsilon) + \ln \sigma(\varepsilon)\right)\right] \\ &= \lim_{\varepsilon\to 0^+} \operatorname{Tr}\left[\exp\left(\frac{1}{2}(\ln \rho(\varepsilon) + \ln \sigma(\varepsilon)\right)\right]. \end{aligned} \tag{1.8}$$

See Appendix A.1 for a short proof of (1.8), and see [AD15, Section 4] and [MO17, Eq. (17)] for a quantity that generalizes the log-Euclidean fidelity.

For two pure states $|\psi\rangle\langle\psi|$ and $|\phi\rangle\langle\phi|$, the z-fidelity reduces to

$$F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\langle\psi|\phi\rangle|^{2z}. \tag{1.9}$$

Operationally, this is the zth power of the probability for the state $|\psi\rangle\langle\psi|$ to pass a test for being the state $|\phi\rangle\langle\phi|$.

The bivariate setting focuses on the similarity of two states. Going forward, the main goal of our paper is to identify measures for quantifying the similarity of a given *tuple of states*. We focus on two main approaches to arrive at multivariate quantum fidelities, as shown in Fig. 1. The first approach is to generalize multivariate classical fidelities, which reduce to bivariate classical fidelity for two commuting states, to a tuple of non-commuting states. The second is to generalize bivariate quantum fidelities to multivariate quantum fidelities that reduce to the bivariate setting when two states are considered. We also note that some generalizations can be obtained by following either of the aforementioned approaches.

1.2 Contributions

In this paper, we comprehensively study multivariate generalizations of bivariate fidelity, while establishing connections between different formulations.

First, we recall some multivariate classical fidelities for commuting states, namely, the Matusita multivariate fidelity and average pairwise fidelity, and then we establish an inequality between them (Proposition 4.6). Thereafter we generalize these quantities to average k-wise fidelities (Definition 4.7) and prove that they are ordered (Proposition 4.8).

Next, we introduce quantum generalizations, mainly focusing on three variants that reduce to average pairwise fidelity for commuting states: average pairwise fidelity using z-fidelities for $z \geq 1/2$ (Definition 5.2); multivariate semi-definite programming (SDP) fidelity (Definition 5.8); and secrecy-based multivariate fidelity (Definition 5.22). The multivariate SDP fidelity is obtained by generalizing the SDP of Uhlmann fidelity from [Wat09] to multiple states. The secrecy-based multivariate fidelity is inspired by an existing secrecy measure from [KRS09, Eq. (19)]—the average fidelity between each state of the tuple and a fixed state, where it is maximized over all such fixed states (c.f., (5.48) for the definition). With these definitions, we show that both SDP fidelity and secrecy-based multivariate fidelity are sandwiched between the average pairwise Holevo fidelity ($z = 1$) and the average pairwise

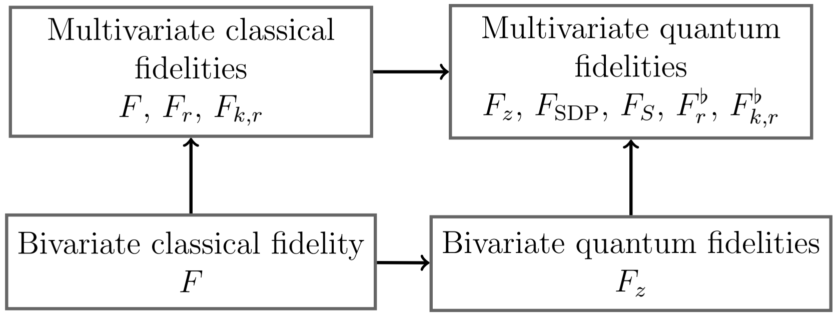

Figure 1: Two main approaches employed in this work to arrive at multivariate quantum fidelities are as follows. Our first approach is to generalize multivariate classical fidelities to a tuple of non-commuting states. These include the average pairwise fidelity F in (4.14), the Matusita multivariate fidelity F_r in (4.2), and the average k-wise fidelities in (4.21), all of which reduce to the bivariate classical fidelity in (1.1) for two commuting states. Our second approach is to generalize bivariate quantum fidelities (z-fidelity F_z in (1.3)) to multivariate quantum fidelities, each of which reduces to a bivariate quantum fidelity when two general states are considered. With these two approaches, we present several multivariate quantum fidelities: average pairwise z-fidelity F_z (Section 5.1); multivariate SDP fidelity F_{SDP} (Section 5.2); secrecy-based multivariate fidelity F_S (Section 5.3); multivariate log-Euclidean fidelity $F_r^\flat$ (Section 5.6); and average k-wise log-Euclidean fidelity $F_{k,r}^\flat$ (Section 5.7).

Uhlmann fidelity ($z = 1/2$), implying that all of these variants converge to average pairwise fidelities in the commuting case (see Theorem 5.28).

Theorem 5.5, Theorem 5.15, and Theorem 5.25 assert that all three of the aforementioned quantum variants satisfy a number of properties desired for a multivariate fidelity: (i) reduction to multivariate classical fidelities for commuting states; (ii) data processing; (iii) permutation invariance; (iv) faithfulness (i.e., it is equal to one if and only if the states are the same); (v) orthogonality (i.e., it is equal to zero if and only if the tuple of states forms an orthogonal set); (vi) direct-sum property; and (vii) joint concavity. In addition, the average pairwise fidelity satisfies super-multiplicativity, whereas the SDP fidelity and secrecy-based multivariate fidelity do not satisfy this property in general. Corollary 5.14 establishes that the multivariate SDP fidelity satisfies a uniform continuity bound, which follows by means of an alternative formulation for the SDP fidelity in Theorem 5.13. Proposition 5.4 and Proposition 5.24 state that the average pairwise Uhlmann and Holevo fidelities and secrecy-based multivariate fidelity also satisfy uniform continuity bounds, respectively. In addition, we define maximal and minimal extensions of multivariate classical fidelities (Definition 5.30 and Definition 5.31) and analyse their properties.

Lastly, we explore a quantum generalization of the Matusita multivariate fidelity, and we show that it is a special case of a multivariate log-Euclidean divergence. We call this variant the multivariate log-Euclidean fidelity. We show that this satisfies most of the desirable properties of a multivariate fidelity in Theorem 5.45. We give an operational interpretation of all multivariate log-Euclidean divergences in terms of quantum hypothesis testing with an arbitrarily varying null hypothesis (Corollary 5.48), by making use of the main result of [Nö14]. In addition, we define the *oveloh information* in (5.109) and show that it is related to the multivariate log-Euclidean fidelity in (5.110). In passing, we also derive a connection between the Holevo information and oveloh information in Corollary 5.47. We also define average k-wise log-Euclidean fidelities (Definition 5.49) and show that they are ordered (Proposition 5.50), as a generalization of the aforementioned average k-wise classical fidelities.

1.3 Paper organization

The rest of our paper is organized as follows. In Section 2, we introduce notation and background needed to understand the rest of the paper. Section 3 reviews and presents different formulations of the bivariate Uhlmann fidelity. Multivariate classical fidelities are defined and studied in Section 4. In Section 5, we introduce proposals for multivariate quantum fidelity, focusing on four main variants, their properties, relationships between different formulations, and operational interpretations of some of them. In the appendices, we provide mathematical proofs of results presented in the paper. Finally, Section 6 provides concluding remarks and future directions.

2 Notations and background

We begin by reviewing basic concepts from quantum information theory and refer the reader to [KW20] for more details. A quantum system R is identified with a finite-dimensional complex Hilbert space $\mathcal{H}_R$ with inner product denoted in bra-ket notation as $\langle\psi|\phi\rangle$ for $|\psi\rangle, |\phi\rangle \in \mathcal{H}_R$. We denote the set of linear operators acting on $\mathcal{H}_R$ by $\mathscr{L}_R$. The support of a linear operator $X \in \mathscr{L}_R$ is defined to be the orthogonal complement of its kernel, and we denote it by $\operatorname{supp}(X)$. We denote the *Hermitian conjugate* or *adjoint* of X by $X^\dagger$, which is the unique linear operator acting on $\mathcal{H}_R$ that satisfies $\langle\psi|(X^\dagger|\phi\rangle) = (X|\psi\rangle)^\dagger|\phi\rangle$ for all $|\psi\rangle, |\phi\rangle \in \mathcal{H}_R$. Here, $|\psi\rangle^\dagger \equiv \langle\psi|$ is the linear functional $\mathcal{H}_R \to \mathbb{C}$ given by $|\phi\rangle \mapsto \langle\psi|\phi\rangle$. The set $\mathscr{L}_R$ is a linear space and has a Hilbert-space structure given by the Hilbert–Schmidt inner product, defined as $\langle X, Y\rangle \coloneqq \operatorname{Tr}[X^\dagger Y]$ for all $X, Y \in \mathscr{L}_R$. Given two quantum systems A and B, with respective Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the Hilbert space of the composite system AB is given by $\mathcal{H}_{AB} \coloneqq \mathcal{H}_A \otimes \mathcal{H}_B$. For $K_{AB} \in \mathscr{L}_{AB}$, let $\operatorname{Tr}[K_{AB}]$ denote the trace of K_{AB}, and let $\operatorname{Tr}_A[K_{AB}]$ denote the partial trace of K over the system A. We use the standard notation $K_A \equiv \operatorname{Tr}_B[K_{AB}]$ and $K_B \equiv \operatorname{Tr}_A[K_{AB}]$ to denote the marginals of K_{AB}. The trace norm of an operator K is defined as $\|K\|_1 \coloneqq \operatorname{Tr}[\sqrt{K^\dagger K}]$. For Hermitian operators K and L, the notation $K \geq L$ indicates that $K - L$ is a positive semi-definite (PSD) operator, while $K > L$ indicates that $K - L$ is a positive definite operator. Given a complex number z, we denote the real part of z by $\mathfrak{R}[z]$.

A quantum state of a system R is identified with a density operator $\rho_R \in \mathscr{L}_R$, which is a PSD operator of unit trace. We denote the set of all density operators acting on $\mathcal{H}_R$ by $\mathscr{D}_R$. We shall also use the notations $\mathscr{D}$ and $\mathscr{L}$ to denote the sets of density operators and linear operators, respectively, when there is no ambiguity regarding the underlying quantum system. A quantum state ρ_R is called a pure state if its rank is equal to one, and in this case, there exists a state vector $|\psi\rangle_R \in \mathcal{H}_R$ such that $\rho_R = |\psi\rangle\!\langle\psi|_R$. Otherwise, ρ_R is called a mixed state. By the spectral decomposition theorem, every state can be written as a convex combination of pure and mutually orthogonal states. A quantum channel from system A to system B is a linear, completely positive and trace-preserving (CPTP) map from $\mathscr{L}_A$ to $\mathscr{L}_B$. A measurement of a quantum system R is described by a positive operator-valued measure (POVM), which is defined as a tuple of PSD operators $(M_y)_{y\in\mathcal{Y}}$ satisfying $\sum_{y\in\mathcal{Y}} M_y = I_R$, where I_R is the identity operator acting on $\mathcal{H}_R$ and $\mathcal{Y}$ is a finite alphabet. The Born rule asserts that, when applying the above POVM to a state ρ, the probability of observing the outcome y is given by $\mathrm{Tr}[M_y\rho]$ [Bor26]. Associated with any POVM $(M_y)_{y\in\mathcal{Y}}$ is a measurement or quantum-to-classical channel $\mathcal{M}$ described as follows. Let $\mathcal{K}$ be a complex Hilbert space of dimension $|\mathcal{Y}|$, and let $\{|y\rangle : y \in \mathcal{Y}\}$ be a fixed orthonormal basis of $\mathcal{K}$. Given an input state ω, the output of the measurement channel is given by

$$\mathcal{M}(\omega) \coloneqq \sum_{y\in\mathcal{Y}} \mathrm{Tr}[M_y\omega]\, |y\rangle\!\langle y|. \tag{2.1}$$

Finally, throughout our paper, we adopt the shorthand $[r] \coloneqq \{1,\ldots,r\}$, and we let S_r denote the permutation group on $[r]$.

3 Many forms of Uhlmann fidelity and its properties

In this section, we discuss several equivalent formulations for the Uhlmann fidelity, along with its various properties. This section also serves as a foundation for subsequent sections, in which we extend the Uhlmann fidelity to multivariate fidelities.

We adopt the following notations here: $\mathscr{U}_R$ denotes the set of unitary operators acting on $\mathcal{H}_R$; POVM_A denotes the set of POVMs acting on system A; and $\mathscr{L}_A^+ \subset \mathscr{L}_A$ denotes the set of positive semi-definite operators.

Let us begin by recalling that the Uhlmann fidelity has several equivalent formulations, as stated in the following proposition. This is one of the reasons that it is the mostly widely used bivariate quantum fidelity.

Proposition 3.1 (Equivalent formulations of Uhlmann fidelity)**.** *Let ρ and σ be quantum states of a system A. Let R be a reference system of the same dimension as A, and let $|\psi^\rho\rangle$ and $|\psi^\sigma\rangle$ be purifications of ρ and σ, respectively. The Uhlmann fidelity $F(\rho,\sigma)$ between ρ and σ is equal to any one of the following expressions:*

$$\begin{aligned} F(\rho,\sigma) &\coloneqq \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1 \\ &= \sup_{U_R\in\mathscr{U}_R} |\langle\psi^\rho|U_R\otimes I_A|\psi^\sigma\rangle| \qquad (3.1)\\ &= \inf_{(\Lambda_x)_x\in\,\mathrm{POVM}_A} \sum_x \sqrt{\mathrm{Tr}[\Lambda_x\rho]\mathrm{Tr}[\Lambda_x\sigma]} \qquad (3.2) \end{aligned}$$

$$= \frac{1}{2} \inf_{Y_1, Y_2 \geq 0} \left\{ \operatorname{Tr}[Y_1 \rho] + \operatorname{Tr}[Y_2 \sigma] : \begin{bmatrix} Y_1 & -I \\ -I & Y_2 \end{bmatrix} \geq 0 \right\} \tag{3.3}$$

$$= \sup_{X \in \mathscr{L}_A} \left\{ \mathfrak{R}[\operatorname{Tr}[X]] : \begin{bmatrix} \rho & X \\ X^\dagger & \sigma \end{bmatrix} \geq 0 \right\}. \tag{3.4}$$

Uhlmann's theorem (3.1) was established in [Uhl76]. The fact that the fidelity is achieved by a quantum measurement (3.2) was realized in [FC95, Eq. (7)]. The SDPs (3.3) and (3.4) for bivariate fidelity were established in [Wat09]. We refer to [KW20] for proofs of the expressions above (i.e., Theorem 6.8, Theorem 6.12, and Proposition 6.6 therein). In addition to these expressions, a corollary of Theorem 5.13 in Section 5 is a new formulation for bivariate fidelity of two states ρ and σ, which we state in Corollary 5.20.

In the following proposition, we list several properties of the Uhlmann fidelity that we use as a guide for defining multivariate quantum fidelities.

Proposition 3.2 (Properties of Uhlmann fidelity)**.** *The Uhlmann fidelity satisfies the following properties for states ρ and σ:*

(i) Reduction to classical fidelity: If the states ρ and σ commute, then

$$F(\rho, \sigma) = \sum_x \sqrt{\rho(x)\sigma(x)}, \tag{3.5}$$

where $(\rho(x))_x$ and $(\sigma(x))_x$ are the spectra of ρ and σ, respectively, in their common eigenbasis.

(ii) Data-processing inequality: For every quantum channel $\mathcal{N}$, we have

$$F(\rho, \sigma) \leq F(\mathcal{N}(\rho), \mathcal{N}(\sigma)). \tag{3.6}$$

(iii) Symmetry: Uhlmann fidelity is independent of the order of the quantum states, i.e.,

$$F(\rho, \sigma) = F(\sigma, \rho). \tag{3.7}$$

(iv) Faithfulness and orthogonality: Uhlmann fidelity satisfies the inequalities $0 \leq F(\rho, \sigma) \leq 1$. Furthermore, $F(\rho, \sigma) = 1$ if and only if $\rho = \sigma$, and $F(\rho, \sigma) = 0$ if and only if ρ and σ are orthogonal to each other, i.e., $\rho\sigma = 0$.

(v) Direct-sum property: Let $(\rho_x)_{x \in \mathcal{X}}$, and $(\sigma_x)_{x \in \mathcal{X}}$ be tuples of quantum states, where $\mathcal{X}$ is a finite alphabet. For the classical-quantum states formed using those tuples, the following equality holds:

$$F\left(\sum_{x \in \mathcal{X}} p(x) |x\rangle\!\langle x| \otimes \rho_x, \sum_{x \in \mathcal{X}} p(x) |x\rangle\!\langle x| \otimes \sigma_x \right) = \sum_{x \in \mathcal{X}} p(x) F(\rho_x, \sigma_x), \tag{3.8}$$

where p is an arbitrary probability distribution on $\mathcal{X}$.

(vi) Joint concavity: Let $(\rho_x)_{x\in\mathcal{X}}$, and $(\sigma_x)_{x\in\mathcal{X}}$ be tuples of quantum states for a finite alphabet $\mathcal{X}$, and let $(p(x))_{x\in\mathcal{X}}$ be a probability distribution. Then,

$$F\left(\sum_{x\in\mathcal{X}} p(x)\rho_x, \sum_{x\in\mathcal{X}} p(x)\sigma_x\right) \geq \sum_{x\in\mathcal{X}} p(x) F(\rho_x, \sigma_x). \tag{3.9}$$

We note that the joint concavity of Uhlmann fidelity is a consequence of its properties (b) data-processing inequality under partial trace channel and (e) direct-sum property. We refer the reader to [KW20] for proofs of these properties.[1]

4 Multivariate classical fidelities

The basic requirement that we set for a multivariate classical fidelity is that it should reduce to the bivariate classical fidelity in (1.1), when only two probability distributions are being considered. Further requirements include satisfying the properties outlined in Proposition 3.2, such as the data-processing inequality, symmetry under exchange of the probability distributions, faithfulness, orthogonality, and the direct-sum property. As such, there are many possible functions that satisfy these requirements, and the aims of this section are to outline several variants of multivariate classical fidelity and to establish relationships between them. In subsequent sections, we show how quantum generalizations reduce to these definitions for commuting states (states represented using classical distributions), which is one of the desired properties for a quantum generalization of a classical measure.

4.1 Matusita multivariate fidelity

Matusita introduced a generalization of the bivariate classical fidelity of two probability distributions to several probability distributions [Mat67], and therein it was called the affinity of several distributions. Here we call it the Matusita multivariate fidelity, and we recall its definition now.

Definition 4.1 (Matusita multivariate fidelity)**.** *For $r \in \mathbb{N}$, let $p_1, \ldots, p_r$ be probability distributions on a finite set $\mathcal{X}$. The Matusita multivariate fidelity is defined as*

$$F_r(p_1, \ldots, p_r) := \sum_{x\in\mathcal{X}} \left(p_1(x) \cdots p_r(x)\right)^{\frac{1}{r}}. \tag{4.1}$$

Furthermore, for a tuple of commuting states $(\rho_i)_{i=1}^r$, the Matusita multivariate fidelity is defined as

$$F_r(\rho_1, \ldots, \rho_r) := \sum_{x} \left(\rho_1(x) \cdots \rho_r(x)\right)^{\frac{1}{r}} \tag{4.2}$$

where, for $i \in [r]$, $(\rho_i(x))_x$ is the spectrum of the state ρ_i in the common eigenbasis of the states.

[1] In particular, we refer to Theorem 6.9 for (ii), Theorem 6.7 for (iv), and Theorem 6.11 for (vi) in [KW20], while (i) and (iii) follow by definition of fidelity.

The Matusita multivariate fidelity satisfies the data-processing inequality, symmetry under exchange of the probability distributions, and the direct-sum property. It also satisfies $0 \leq F_r(p_1, \ldots, p_r) \leq 1$, and the equality $F_r(p_1, \ldots, p_r) = 1$ holds if and only if $p_1 = \cdots = p_r$. See Theorem 1 and Corollary 3 of [Mat67]. Also, if there exists at least one pair of distributions p_i and p_j with $i \neq j \in [r]$ that have disjoint support, then $F_r(p_1, \ldots, p_r) = 0$. However, $F_r(p_1, \ldots, p_r) = 0$ does not imply that at least one pair of distributions is disjoint. For example, the probability vectors

$$p_1 = (0, 1/2, 1/2), \qquad p_2 = (1/2, 0, 1/2), \qquad p_3 = (1/2, 1/2, 0) \tag{4.3}$$

satisfy $F_r(p_1, p_2, p_3) = 0$, but no two distributions have disjoint support.

The Matusita multivariate fidelity can be written in terms of the classical relative entropy

$$D(p\|q) \coloneqq \begin{cases} \sum_x p(x) \ln\left(\frac{p(x)}{q(x)}\right) & \text{if } \operatorname{supp}(p) \subseteq \operatorname{supp}(q) \\ +\infty & \text{else} \end{cases} \tag{4.4}$$

as follows:

$$F_r(p_1, \ldots, p_r) = \exp\left(-\inf_w \frac{1}{r} \sum_x D(w\|p_i)\right), \tag{4.5}$$

where the infimum is over every probability distribution w with support contained in the intersection of the supports of $p_1, \ldots, p_r$ (and if no such distribution exists, then the relative entropy is equal to $+\infty$). We review this representation of the Matusita fidelity in the more general quantum case in Section 5.6, where we also provide an operational interpretation of it. Since the relative entropy is well known to obey the data-processing inequality, the formulation of the Matusita fidelity in (4.5) makes it easier to see that it obeys the data-processing inequality.

Remark 4.2 (Reduction to Rényi relative entropy & Hellinger transform)**.**
The Rényi relative entropy between two probability distributions p and q is defined for $\alpha \in (0,1) \cup (1, \infty)$ as [Rén61]

$$D_\alpha(p\|q) \coloneqq \begin{cases} \frac{1}{\alpha-1} \ln(\sum_x p(x)^\alpha q(x)^{1-\alpha}) & \text{if } \alpha \in (0,1) \vee (\alpha > 1 \wedge \operatorname{supp}(p) \subseteq \operatorname{supp}(q)) \\ +\infty & \text{else} \end{cases}. \tag{4.6}$$

Let $\alpha \in (0,1)$ be rational such that $\alpha = t/r$ with $t < r$. Then, for such α we have that

$$D_\alpha(p\|q) = \frac{1}{\alpha - 1} \ln(F_r(p, \ldots, p, q, \ldots, q)), \tag{4.7}$$

where, on the right-hand side, p occurs t times and q occurs $r - t$ times. Thus, the Rényi relative entropy of rational order is a special case of the Matusita fidelity.

For a tuple of probability distributions $p_1, \ldots, p_r$ and a probability vector $s = (s_1, \ldots, s_r)$, the Hellinger transform is defined as [Hel09] (see also [Tou74, Eq. (34)])

$$H_s(p_1, ..., p_r) \coloneqq \sum_x \prod_{i=1}^r p_i(x)^{s_i}. \tag{4.8}$$

If every s_i is rational such that $s_i = t_i/t$, then we have

$$H_s(p_1, ..., p_r) = F_t(p_1, \ldots, p_1, p_2, \ldots, p_2, \ldots, p_r, \ldots, p_r), \tag{4.9}$$

where, on the right-hand side, p_i appears t_i times, for all $i \in [r]$. As such, the Hellinger transform for rational s is a special case of the Matusita fidelity.

The Hellinger distance between two probability distributions p and q is defined as

$$d_H(p, q) \coloneqq \|\sqrt{p} - \sqrt{q}\|_2 = \left(\sum_x \left(\sqrt{p(x)} - \sqrt{q(x)} \right)^2 \right)^{\frac{1}{2}}. \tag{4.10}$$

Next, we state that the Matusita multivariate fidelity obeys a uniform continuity bound, which quantifies its deviation in terms of the deviation of corresponding probability distributions in two different tuples of probability distributions.

Proposition 4.3 (Uniform continuity of Matusita multivariate fidelity)**.** *Let $(p_1, \ldots, p_r)$ and $(q_1, \ldots, q_r)$ be two tuples of probability distributions, and let $\varepsilon > 0$ be such that $\frac{1}{r} \left(\sum_{i=1}^r [d_H(p_i, q_i)]^2 \right) \le \varepsilon$, where $d_H(\cdot, \cdot)$ is defined in (4.10). Then,*

$$|F_r(p_1, \ldots, p_r) - F_r(q_1, \ldots, q_r)| \le r\, (\varepsilon)^{\frac{1}{r}}. \tag{4.11}$$

Proof. See Appendix A.2. □

4.2 Average pairwise fidelity

The average pairwise Bhattacharyya overlap is defined in [ŞTA09, Eq. (7)], which serves as an upper bound on the error probability of an uncoded transmission of classical data through polarized channels. We also call this quantity the average pairwise fidelity, and note here that it has been considered previously in a similar context for the quantum case [NR18, Section II].

Here we recall the definition of the average pairwise fidelity for classical probability distributions and for commuting states more generally.

Definition 4.4 (Average pairwise fidelity)**.** *Let $p_1, \ldots, p_r$ be probability distributions on a finite set $\mathcal{X}$. The average pairwise fidelity is defined as*

$$F(p_1, \ldots, p_r) \coloneqq \frac{2}{r(r-1)} \sum_{i<j} F(p_i, p_j) \tag{4.12}$$

$$= \frac{2}{r(r-1)} \sum_{i<j} \sum_{x \in \mathcal{X}} \sqrt{p_i(x) p_j(x)}. \tag{4.13}$$

Furthermore, for commuting states $\rho_1, \ldots, \rho_r$, the average pairwise fidelity is defined as

$$F(\rho_1, \ldots, \rho_r) \coloneqq \frac{2}{r(r-1)} \sum_{i<j} F(\rho_i, \rho_j) \tag{4.14}$$

$$= \frac{2}{r(r-1)} \sum_{i<j} \sum_{x} \sqrt{\rho_i(x)\rho_j(x)}, \tag{4.15}$$

where $(\rho_i(x))_x$ is the tuple of eigenvalues of ρ_i.

Remark 4.5 (Uniform continuity of average pairwise fidelity)**.** *The average pairwise fidelity of commuting states obeys a uniform continuity bound, which follows as an immediate consequence of Proposition 5.4 below, the latter holding for general quantum states.*

Proposition 4.6 (Inequalities relating multivariate classical fidelities)**.** *For a tuple of commuting states $(\rho_i)_{i=1}^r$, the following inequality holds, relating the average pairwise fidelity $F(\rho_1, \ldots, \rho_r)$ and the Matusita fidelity $F_r(\rho_1, \ldots, \rho_r)$:*

$$F(\rho_1, \ldots, \rho_r) \geq F_r(\rho_1, \ldots, \rho_r). \tag{4.16}$$

Proof. By Definition 4.4, we have

$$F(\rho_1, \ldots, \rho_r) = \frac{2}{r(r-1)} \sum_{i<j} \sum_{x} \sqrt{\rho_i(x)\rho_j(x)} \tag{4.17}$$

$$= \sum_{x} \frac{2}{r(r-1)} \sum_{i<j} \sqrt{\rho_i(x)\rho_j(x)}, \tag{4.18}$$

where $(\rho_i(x))_x$ is the tuple of eigenvalues of ρ_i, for $i \in [r]$. For each x, the right-hand side of the above expression is an average of $r(r-1)/2$ terms. Applying the inequality of arithmetic and geometric means then gives

$$\frac{2}{r(r-1)} \sum_{i<j} \sqrt{\rho_i(x)\rho_j(x)} \geq \prod_{i<j} \left(\sqrt{\rho_i(x)\rho_j(x)} \right)^{\frac{2}{r(r-1)}} \tag{4.19}$$

$$= \prod_{i<j} \left(\rho_i(x)\rho_j(x)\right)^{\frac{1}{r(r-1)}}. \tag{4.20}$$

The desired inequality (4.16) then follows from the above inequality because each $\rho_i(x)$ appears exactly $r-1$ times in the product in (4.20). ☐

Later on in Proposition 5.46, we generalize the inequality in (4.16) to the log-Euclidean class of quantum fidelities.

4.3 Average k-wise fidelities

As a generalization of the average pairwise fidelity and to interpolate between this quantity and the Matusita fidelity of order r, we define the average k-wise fidelities of a tuple of r commuting states as follows:

Definition 4.7 (Average k-wise fidelities)**.** *For $r \in \{3, 4, \ldots\}$, let $\rho_1, \ldots, \rho_r$ be a tuple of commuting states, where $(\rho_i(x))_x$ is the tuple of eigenvalues of ρ_i. For $k \in \{2, \ldots, r\}$, we define the average k-wise fidelity of $\rho_1, \ldots, \rho_r$ as*

$$F_{k,r}(\rho_1, \ldots, \rho_r) := \frac{1}{\binom{r}{k}} \sum_{i_1 < i_2 < \cdots < i_k} F_k(\rho_{i_1}, \rho_{i_2}, \ldots, \rho_{i_k}), \tag{4.21}$$

where $F_k(\rho_{i_1}, \rho_{i_2}, \ldots, \rho_{i_k}) = \sum_x (\rho_{i_1}(x) \cdots \rho_{i_k}(x))^{\frac{1}{k}}$ is the Matusita fidelity of order k.

As a consequence of the properties of Matusita fidelity, all of these quantities obey the data-processing inequality, symmetry under exchange of the states, faithfulness, equal to zero for orthogonal states, and the direct-sum property. They are all also jointly concave, as a consequence of the data-processing inequality and the direct-sum property.

Observe that $F(\rho_1, \ldots, \rho_r) = F_{2,r}(\rho_1, \ldots, \rho_r)$ and $F_r(\rho_1, \ldots, \rho_r) = F_{r,r}(\rho_1, \ldots, \rho_r)$, so that both the average pairwise fidelity and the Matusita fidelity of order r are special cases of the average k-wise fidelities.

As a refinement of Proposition 4.6, the average k-wise fidelities are sorted in a descending order, which again follows from an application of the inequality of arithmetic and geometric means, as well as basic combinatorial reasoning:

Proposition 4.8 (Inequalities relating average k-wise fidelities)**.** *For $r \in \{3, 4, \ldots\}$, let $\rho_1, \ldots, \rho_r$ be a tuple of commuting states. Then*

$$F_{2,r} \geq F_{3,r} \geq \cdots \geq F_{r-1,r} \geq F_{r,r}, \tag{4.22}$$

where, for brevity, we have suppressed the dependence of each quantity $F_{k,r}$ on $\rho_1, \ldots, \rho_r$.

Proof. See Appendix A.3. □

5 Proposed quantum generalizations

In this section, we first introduce several desirable properties of multivariate quantum fidelity, which are generalizations of the properties satisfied by the Uhlmann fidelity in Proposition 3.2. Then we propose four main generalizations of the bivariate quantum fidelities from Section 1.1:

1. The first generalization is the average pairwise z-fidelity, which is the simplest generalization of the classical average pairwise fidelity in Definition 4.4.

2. The next generalization, called the multivariate SDP fidelity, uses the SDP formulation of the Uhlmann fidelity presented in Proposition 3.1.

3. The third generalization, called the secrecy-based multivariate fidelity, is inspired by an existing secrecy measure from [KRS09].

4. The fourth generalization, called the multivariate log-Euclidean fidelity, is defined through the log-Euclidean divergences given in Definition 5.42.

We also show that the second and the third generalizations are quantum generalizations of the classical average pairwise fidelity introduced in Definition 4.4. Furthermore, we show that the last one is a quantum generalization of the Matusita multivariate fidelity given in Definition 4.1. In addition, we define maximal and minimal extensions of multivariate classical fidelities (see Definition 5.30 and Definition 5.31). We also show that the proposed multivariate quantum fidelities satisfy several desired properties presented in Definition 5.1.

Definition 5.1 (Desired properties of multivariate fidelity)**.**
For quantum states $\rho_1, \ldots, \rho_r$, we consider the following properties of a multivariate fidelity quantity $\mathbf{F}(\rho_1, \ldots, \rho_r)$:

(i) Reduction to multivariate classical fidelity: For commuting states $\rho_1, \ldots, \rho_r$,

$$\begin{aligned}\mathbf{F}(\rho_1, \ldots, \rho_r) &= \frac{2}{r(r-1)} \sum_{i<j} F(\rho_i, \rho_j) &(5.1)\\ &= \frac{2}{r(r-1)} \sum_{i<j} \sum_x \sqrt{\rho_i(x)\rho_j(x)}, &(5.2)\end{aligned}$$

or

$$\begin{aligned}\mathbf{F}(\rho_1, \ldots, \rho_r) &= F_r(\rho_1, \ldots, \rho_r) &(5.3)\\ &= \sum_x \left(\rho_1(x) \cdots \rho_r(x)\right)^{\frac{1}{r}}, &(5.4)\end{aligned}$$

or, for some $k \in \{3, \ldots, r-1\}$,

$$\begin{aligned}\mathbf{F}(\rho_1, \ldots, \rho_r) &= F_{k,r}(\rho_1, \ldots, \rho_r) &(5.5)\\ &= \frac{1}{\binom{r}{k}} \sum_{i_1 < \cdots < i_k} \sum_x \left(\rho_{i_1}(x) \cdots \rho_{i_k}(x)\right)^{\frac{1}{k}}, &(5.6)\end{aligned}$$

where $(\rho_i(x))_x$ is the spectrum of the state ρ_i in the common eigenbasis of all the states.

(ii) Data processing: For a quantum channel $\mathcal{N}$,

$$\mathbf{F}(\rho_1, \ldots, \rho_r) \leq \mathbf{F}(\mathcal{N}(\rho_1), \ldots, \mathcal{N}(\rho_r)). \tag{5.7}$$

(iii) Symmetry: For every permutation π of $[r]$,

$$\mathbf{F}(\rho_1, \ldots, \rho_r) = \mathbf{F}(\rho_{\pi(1)}, \ldots, \rho_{\pi(r)}). \tag{5.8}$$

(iv) Faithfulness: $\mathbf{F}(\rho_1, \ldots, \rho_r) = 1$ if and only if all the states are the same, i.e., $\rho_i = \rho_j$ for $i, j \in [r]$.

(v) Orthogonality: $\mathbf{F}(\rho_1, \ldots, \rho_r) = 0$ *if and only if all states are orthogonal to each other, i.e.,* $\rho_i \rho_j = 0$ *for* $i \neq j,\ i, j \in [r]$.[2]

(vi) Direct-sum property: Let $(\rho_i^x)_x$ *be tuples of quantum states for all* $i \in [r]$. *For a probability distribution* $(p(x))_{x\in\mathcal{X}}$ *where* $\mathcal{X}$ *is a finite alphabet and classical–quantum states formed using those tuples, the following holds:*

$$\mathbf{F}\left(\sum_{x\in\mathcal{X}} p(x)|x\rangle\!\langle x| \otimes \rho_1^x, \ldots, \sum_{x\in\mathcal{X}} p(x)|x\rangle\!\langle x| \otimes \rho_r^x\right) = \sum_{x\in\mathcal{X}} p(x)\mathbf{F}(\rho_1^x, \ldots, \rho_r^x). \tag{5.9}$$

Another desirable property for a multivariate fidelity is joint concavity, stated as follows. Let $(\rho_i^x)_{x\in\mathcal{X}}$ be tuples of quantum states for all $i \in [r]$, and let $(p(x))_{x\in\mathcal{X}}$ be a probability distribution. Then

$$\mathbf{F}\left(\sum_{x\in\mathcal{X}} p(x)\rho_1^x, \ldots, \sum_{x\in\mathcal{X}} p(x)\rho_r^x\right) \geq \sum_{x\in\mathcal{X}} p(x)\mathbf{F}(\rho_1^x, \ldots, \rho_r^x). \tag{5.10}$$

This property is an immediate consequence of data processing and the direct-sum property. Indeed, if these latter two properties hold, then one obtains joint concavity by applying (5.9) and then (5.7) with the channel $\mathcal{N}$ as the partial trace over the classical register.

5.1 Average pairwise fidelities

In this subsection, we generalize the average pairwise classical fidelity in the simplest way possible by using the z-fidelity defined in (1.3).

Definition 5.2 (Average pairwise z-fidelity). *For* $z \geq 1/2$ *and quantum states* $\rho_1, \ldots, \rho_r$, *the average pairwise* z*-fidelity is defined as*

$$F_z(\rho_1, \ldots, \rho_r) \coloneqq \frac{2}{r(r-1)} \sum_{i<j} F_z(\rho_i, \rho_j). \tag{5.11}$$

Remark 5.3 (Average pairwise Holevo and Uhlmann fidelities). *By fixing* $z = 1/2$, *we obtain the average pairwise Holevo fidelity, denoted by*

$$F_H(\rho_1, \ldots, \rho_r) \coloneqq \frac{2}{r\,(r-1)} \sum_{i<j} F_H(\rho_i, \rho_j). \tag{5.12}$$

For $z = 1$, *we obtain the average pairwise Uhlmann fidelity, denoted by*

$$F_U(\rho_1, \ldots, \rho_r) \coloneqq \frac{2}{r\,(r-1)} \sum_{i<j} F(\rho_i, \rho_j). \tag{5.13}$$

[2] Note that it is only possible for Matusita fidelity to satisfy $\mathbf{F}(\rho_1, \ldots, \rho_r) = 0$ if all states are orthogonal to each other. The example presented in (4.3) excludes the other implication from holding for Matusita fidelity. The same statements apply to the average k-wise fidelities for $r \geq 3$ and $3 \leq k \leq r$.

As mentioned previously, the average pairwise Uhlmann fidelity has appeared in [NR18] as a measure of the reliability of a classical-quantum channel. Note that in the classical (viz., commuting states) case, the average pairwise z-fidelities are equal for all $z \geq 1/2$, and they reduce to the multivariate classical fidelity given in Definition 4.4.

5.1.1 Uniform continuity bound for average pairwise Uhlmann and Holevo fidelities

Here we prove uniform continuity bounds for the Uhlmann and Holevo average pairwise fidelities. Before doing so, let us define the Bures [Hel67, Bur69] and Hellinger distances [Jen04, LZ04] and recall that they obey the triangle inequality:

$$d_B(\rho, \sigma) := \sqrt{2\left[1 - F(\rho, \sigma)\right]}, \tag{5.14}$$

$$d_H(\rho, \sigma) := \sqrt{2\left[1 - F_H(\rho, \sigma)\right]}. \tag{5.15}$$

Proposition 5.4 (Uniform continuity of average pairwise fidelities). *Let $\rho_1, \ldots, \rho_r, \sigma_1, \ldots, \sigma_r$ be quantum states, and let $\varepsilon > 0$ be such that $\frac{1}{r}\sum_{i=1}^{r} d_B(\rho_i, \sigma_i) \leq \varepsilon$. Then*

$$\left|F_U(\rho_1, \ldots, \rho_r) - F_U(\sigma_1, \ldots, \sigma_r)\right| \leq 2\sqrt{2}\varepsilon. \tag{5.16}$$

Similarly, if $\varepsilon > 0$ is such that $\frac{1}{r}\sum_{i=1}^{r} d_H(\rho_i, \sigma_i) \leq \varepsilon$, then

$$\left|F_H(\rho_1, \ldots, \rho_r) - F_H(\sigma_1, \ldots, \sigma_r)\right| \leq 2\sqrt{2}\varepsilon. \tag{5.17}$$

Proof. See Appendix A.4. □

5.1.2 Properties of average pairwise z-fidelities

Here we show that the average pairwise z-fidelities satisfy all desirable properties from Definition 5.1. In addition, we show that it satisfies super-multiplicativity and coarse-graining.

Theorem 5.5 (Properties of average pairwise z-fidelity). *For $z \in [1/2, +\infty)$, the average pairwise z-fidelity satisfies all desired properties of a multivariate fidelity, as listed in Definition 5.1.*

Proof. Proofs of these properties follow because the z-fidelity (for $r = 2$) satisfies all the properties listed in Proposition 3.2. In particular, reduction to the classical setting follows by plugging commuting states into (1.3) and simplifying. Faithfulness follows because, if the states are the same, then $F_z(\rho, \rho) = \left\|\rho^{\frac{1}{4z}}\rho^{\frac{1}{4z}}\right\|_{2z}^{2z} = \left\|\rho^{\frac{1}{2z}}\right\|_{2z}^{2z} = 1$. If $F_z(\rho, \sigma) = 1$, then due to the fact that the z-fidelities are monotonically decreasing and continuous in z [LT15, Proposition 6], we conclude that $F(\rho, \sigma) = 1$, from which we conclude that $\rho = \sigma$. Orthogonality follows because, if $\rho\sigma = 0$, then $\rho^{\frac{1}{4z}}\sigma^{\frac{1}{4z}} = 0$, from which we conclude that $F_z(\rho, \sigma) = 0$. If $F_z(\rho, \sigma) = 0$, then $\rho^{\frac{1}{4z}}\sigma^{\frac{1}{4z}} = 0$ (from positive definiteness of the norm expression in (1.3)), which implies that $\rho\sigma = 0$. Moreover, data-processing follows from [Zha20, Theorem 1.2] and the direct-sum property because

$$\operatorname{Tr}\left[\left(\left(\sum_x p(x)|x\rangle\langle x|\otimes\rho_x\right)^{1/4z}\left(\sum_x p(x)|x\rangle\langle x|\otimes\sigma_x\right)^{1/2z}\left(\sum_x p(x)|x\rangle\langle x|\otimes\rho_x\right)^{1/4z}\right)^z\right]$$
$$= \sum_x p(x)\operatorname{Tr}\left[\left(\rho_x^{1/4z}\sigma_x^{1/2z}\rho_x^{1/4z}\right)^z\right], \quad (5.18)$$

thus concluding the proof. □

Proposition 5.6 (Super-multiplicativity of average pairwise fidelities)**.** *For states $\rho_1, \ldots, \rho_r$ and $z \geq 1/2$, the average pairwise z-fidelity is supermultiplicative:*

$$(F_z(\rho_1, \ldots, \rho_r))^n \leq F_z\left(\rho_1^{\otimes n}, \ldots, \rho_r^{\otimes n}\right). \quad (5.19)$$

Proof. Using Definition 5.2 for the average pairwise z-fidelity, consider the following steps:

$$\begin{aligned}
F_z\left(\rho_1^{\otimes n}, \ldots, \rho_r^{\otimes n}\right) &= \frac{2}{r(r-1)}\sum_{i<j} F_z(\rho_i^{\otimes n}, \rho_j^{\otimes n}) && (5.20)\\
&= \frac{2}{r(r-1)}\sum_{i<j}(F_z(\rho_i, \rho_j))^n && (5.21)\\
&\geq \left(\frac{2}{r(r-1)}\sum_{i<j} F_z(\rho_i, \rho_j)\right)^n && (5.22)\\
&= (F_z(\rho_1, \ldots, \rho_r))^n, && (5.23)
\end{aligned}$$

where the second equality follows from the multiplicativity of z-fidelity. For the inequality, we use convexity of the function x^n for $n > 1$ and apply Jensen's inequality as follows: Let Y be a random variable with a uniform probability distribution over the values y_{ij}, where $y_{ij} := F_z(\rho_i, \rho_j)$ for all $i < j$ with $i, j \in [r]$. Then Jensen's inequality gives $\mathbb{E}[Y^n] \geq (\mathbb{E}[Y])^n$, where $\mathbb{E}$ denotes the expectation. □

Proposition 5.7 (Coarse-graining property of average pairwise fidelities)**.**
For states $\rho_1, \ldots, \rho_r, \ldots, \rho_{r+m}$, the average pairwise z-fidelity satisfies the following inequality for all $z \geq 1/2$:

$$F_z(\rho_1, \ldots, \rho_r) \leq \frac{(r+m)(r+m-1)}{r(r-1)} F_z(\rho_1, \ldots, \rho_r, \ldots, \rho_{r+m}). \quad (5.24)$$

Proof. Consider that

$$\sum_{i,j=1:i\neq j}^{r} F_z(\rho_i, \rho_j) \leq \sum_{i,j=1:i\neq j}^{r+m} F_z(\rho_i, \rho_j) \quad (5.25)$$

due to $0 \leq F_z(\rho, \sigma)$. Then, by rewriting the above terms by using Definition 5.2, we arrive at the desired inequality. □

5.2 Multivariate semi-definite programming fidelity

Here we propose a multivariate generalization of Uhlmann fidelity by generalizing its SDP formulation given in (3.3):

Definition 5.8 (Multivariate SDP fidelity)**.** *Given quantum states $\rho_1, \ldots, \rho_r$, we define their multivariate SDP fidelity as*

$$F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) \coloneqq \frac{1}{r(r-1)} \inf_{Y_1,\ldots,Y_r \geq 0} \left\{ \sum_{i=1}^{r} \operatorname{Tr}[Y_i \rho_i] : \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes Y_i \geq \sum_{i \neq j} |i\rangle\!\langle j| \otimes I \right\}. \tag{5.26}$$

Remark 5.9 (Multivariate SDP fidelity for PSD operators)**.** *Definition 5.8 extends to general positive semi-definite (PSD) operators, simply by replacing the tuple of states $(\rho_i)_{i=1}^{r}$ with the tuple of PSD operators $(A_i)_{i=1}^{r}$.*

In view of the other SDP formulations of the Uhlmann fidelity in Proposition 3.1, it is natural to ask if analogous formulations can be constructed for the multivariate SDP fidelity. It is indeed the case, as stated in Proposition 5.10 and Theorem 5.13.

Proposition 5.10 (Dual of multivariate SDP fidelity)**.** *The multivariate SDP fidelity in Definition 5.8 has the following dual formulation:*

$$F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) = \frac{2}{r(r-1)} \sup_{\substack{(X_{ij})_{i\neq j} \text{ s.t.} \\ X_{ji} = X_{ij}^{\dagger}}} \left\{ \sum_{i<j} \mathfrak{R}[\operatorname{Tr}[X_{ij}]] : \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes \rho_i + \sum_{i\neq j} |i\rangle\!\langle j| \otimes X_{ij} \geq 0 \right\}. \tag{5.27}$$

Proof. See Appendix A.5. □

5.2.1 Uniform continuity bound for multivariate SDP fidelity

The goal of this subsection is to prove a uniform continuity bound for the multivariate SDP fidelity. One important step in doing so is to introduce the following quantity, which we show later to be equal to the multivariate SDP fidelity.

Definition 5.11 ($K^{\star}$-representation)**.** *For a tuple $(\rho_i)_{i=1}^{r}$ of states, define*

$$F_{K^{\star}}(\rho_1, \ldots, \rho_r) \coloneqq \frac{1}{r-1} \sup_{K \geq 0} \left\{ \langle \psi | K \otimes I_d | \psi \rangle - 1 : K = I_r \otimes I_d + \sum_{i \neq j} |i\rangle\!\langle j| \otimes K_{ij} \right\}, \tag{5.28}$$

where

$$|\psi\rangle \coloneqq \frac{1}{\sqrt{r}} \sum_{i=1}^{r} |i\rangle |\phi^{\rho_i}\rangle, \tag{5.29}$$

with $|\phi^{\rho_i}\rangle$ being a purification of ρ_i for all $i \in [r]$ and d the dimension of the Hilbert space of the states.

We call this the $K^\star$-representation, and later on, Theorem 5.13 states that it is equal to the multivariate SDP fidelity:

$$F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) = F_{K^\star}(\rho_1, \ldots, \rho_r) \tag{5.30}$$

Let us note that the $K^\star$-representation can also be expressed as

$$F_{K^\star}(\rho_1, \ldots, \rho_r) = \frac{1}{r-1} \sup_{K' \in \mathrm{Herm}} \left\{ \langle\psi|K' \otimes I_d|\psi\rangle : K' \coloneqq \sum_{i \neq j} |i\rangle\!\langle j| \otimes K_{ij} \leq I_r \otimes I_d \right\}, \tag{5.31}$$

where the optimization is over every Hermitian matrix K'.

Next, we show that the $K^\star$-representation satisfies a uniform continuity bound. This property finds use in the proof of Theorem 5.13.

Theorem 5.12 (Uniform continuity of $K^\star$-representation)**.** *Let $\rho_1, \ldots, \rho_r, \sigma_1, \ldots, \sigma_r$ be quantum states, and let $\varepsilon \in [0, 1]$ be such that*

$$\frac{1}{r} \sum_{i=1}^{r} F(\rho_i, \sigma_i) \geq 1 - \varepsilon. \tag{5.32}$$

Then,

$$|F_{K^\star}(\rho_1, \ldots, \rho_r) - F_{K^\star}(\sigma_1, \ldots, \sigma_r)| \leq \frac{r}{r-1} \sqrt{\varepsilon\,(2 - \varepsilon)}. \tag{5.33}$$

Proof. See Appendix A.6. □

Theorem 5.13 below provides another equivalent formulation for the multivariate SDP fidelity, stating that it is equivalent to $K^\star$-representation. This formulation will be useful in proving the lower bound stated in Theorem 5.28 and some consequential properties in Theorem 5.15.

Theorem 5.13 (Multivariate SDP fidelity and $K^\star$-representation)**.** *For quantum states $\rho_1, \ldots, \rho_r$, the multivariate SDP fidelity (5.26) is equal to the $K^\star$-representation:*

$$F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) = F_{K^\star}(\rho_1, \ldots, \rho_r). \tag{5.34}$$

Proof. See Appendix A.8. □

By using Theorem 5.12 and Theorem 5.13, we arrive at the following uniform continuity bound for the multivariate SDP fidelity.

Corollary 5.14 (Uniform continuity of multivariate SDP fidelity)**.** *Let $\rho_1, \ldots, \rho_r, \sigma_1, \ldots, \sigma_r$ be quantum states, and let $\varepsilon \in [0, 1]$ be such that*

$$\frac{1}{r} \sum_{i=1}^{r} F(\rho_i . \sigma_i) \geq 1 - \varepsilon, \tag{5.35}$$

Then,

$$|F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) - F_{\mathrm{SDP}}(\sigma_1, \ldots, \sigma_r)| \leq \frac{r}{r-1} \sqrt{\varepsilon\,(2 - \varepsilon)}. \tag{5.36}$$

5.2.2 Properties of multivariate SDP fidelity

Here we present several properties satisfied by multivariate SDP fidelity.

Theorem 5.15 (Properties of multivariate SDP fidelity)**.** *Multivariate SDP fidelity satisfies all the desired properties of multivariate fidelity listed in Definition 5.1. For commuting states, it reduces to the classical average pairwise fidelity.*

Proof. See Appendix A.9. □

Remark 5.16 (Properties for a set of PSD operators)**.** *For a tuple of PSD operators $(A_i)_{i=1}^r$, the following properties hold. For $c > 0$, by applying Definition 5.8 directly, we have the following scaling property:*

$$F_{\mathrm{SDP}}(cA_1, \ldots cA_r) = cF_{\mathrm{SDP}}(A_1, \ldots, A_r). \tag{5.37}$$

Let $(A_i^x)_x$ be tuples of PSD operators for all $i \in [r]$. Then

$$F_{\mathrm{SDP}}\left(\sum_{x\in\mathcal{X}} |x\rangle\!\langle x| \otimes A_1^x, \ldots, \sum_{x\in\mathcal{X}} |x\rangle\!\langle x| \otimes A_r^x\right) = \sum_{x\in\mathcal{X}} F_{\mathrm{SDP}}(A_1^x, \ldots, A_r^x). \tag{5.38}$$

A proof of (5.38) follows similarly to the proof of direct-sum property of multivariate SDP fidelity in Theorem 5.15. To this end, note that property (vi) in Theorem 5.15 can be obtained as a special case of (5.38) by using (5.37) and (5.38). The other properties (i)-(iii) in Theorem 5.15 also hold by following the same proof arguments, due to the fact that Theorem 5.28 holds even for PSD operators, among other things.

Proposition 5.17 (Coarse-graining property)**.** *Let $r, m \in \mathbb{N}$, and let $\rho_1, \ldots, \rho_{r+m}$ be quantum states. We have the following inequality:*

$$F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) \leq \frac{(r+m)(r+m-1)}{r(r-1)} F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r, \ldots, \rho_{r+m}). \tag{5.39}$$

Proof. See Appendix A.10. □

Remark 5.18 (Super-multiplicativity of multivariate SDP fidelity)**.** *Through numerical calculations, we found that there exists a tuple of states for which super-multiplicativity does not hold for the SDP fidelity, i.e., $(F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r))^n > F_{\mathrm{SDP}}\big(\rho_1^{\otimes n}, \ldots, \rho_r^{\otimes n}\big)$ for some r and $(\rho_i)_{i=1}^r$. An example of a tuple of pure states (i.e., $\rho_i \coloneqq |\psi_i\rangle\!\langle\psi_i|$ for $i \in \{1, 2, 3\}$) for $r = 3$, $n = 2$, and $d = 3$ is as follows (up to numerical approximations of MATLAB):*

$$|\psi_1\rangle = \begin{pmatrix} -0.8954 + 0.2791i \\ 0.2061 - 0.0805i \\ 0.2418 + 0.1135i \end{pmatrix}, \tag{5.40}$$

$$|\psi_2\rangle = \begin{pmatrix} -0.2422 + 0.2315i \\ 0.4386 - 0.4318i \\ -0.6928 - 0.1704i \end{pmatrix}, \tag{5.41}$$

$$|\psi_3\rangle = \begin{pmatrix} 0.2560 + 0.5837i \\ 0.4811 - 0.3057i \\ -0.5175 - 0.314i \end{pmatrix}. \tag{5.42}$$

For the above example,

$$0.4075 = (F_{\text{SDP}}(\rho_1, \rho_2, \rho_3))^2 > F_{\text{SDP}}\left(\rho_1^{\otimes 2}, \rho_2^{\otimes 2}, \rho_3^{\otimes 2}\right) = 0.3820, \tag{5.43}$$

which shows a violation of super-multiplicativity.[3]

5.2.3 Other reductions of multivariate SDP fidelity

Remark 5.19 (Multivariate SDP fidelity as a pairwise quantity)**.** *The development in the proof of Theorem 5.13 gives insight into how we can view the SDP fidelity as being somewhat like an average pairwise quantity (see (A.92)). Indeed, we see that*

$$F_{\text{SDP}}(\rho_1, \ldots, \rho_r) = \frac{2}{r\,(r-1)} \sup_{K \geq 0} \left\{ \begin{array}{c} \sum_{i<j} \mathfrak{R}\left[\langle \psi^{\rho_i} | K_{i,j} \otimes I | \psi^{\rho_j} \rangle\right] : \\ K = \sum_{i,j=1}^{r} |i\rangle\!\langle j| \otimes K_{i,j}, \;\; K_{i,i} = I_d \;\forall i \in [r] \end{array} \right\}. \tag{5.44}$$

By choosing $r = 2$ in Theorem 5.13 and applying (5.31), we obtain yet another formulation for Uhlmann fidelity.

Corollary 5.20 (Another formulation for Uhlmann fidelity)**.** *For all quantum states ρ and σ, the Uhlmann fidelity can be expressed as*

$$F(\rho, \sigma) = \sup_{K \in \text{Herm}} \left\{ \langle \phi | K \otimes I_d | \phi \rangle : I_2 \otimes I_d \geq \begin{bmatrix} 0 & X \\ X^\dagger & 0 \end{bmatrix} =: K \right\}, \tag{5.45}$$

where

$$|\phi\rangle \coloneqq \frac{1}{\sqrt{2}} \left(|0\rangle |\psi^\rho\rangle + |1\rangle |\psi^\sigma\rangle \right) \tag{5.46}$$

with $|\psi^\rho\rangle$ and $|\psi^\sigma\rangle$ being purifications of ρ and σ, respectively.

We obtain a formulation for multivariate SDP fidelity of pure states by simplifying the $K^\star$-representation and applying Theorem 5.13.

Corollary 5.21 (Multivariate SDP fidelity for pure states)**.** *For a tuple of pure states, $(|\psi_1\rangle\!\langle\psi_1|, \ldots, |\psi_r\rangle\!\langle\psi_r|)$,*

$$F_{\text{SDP}}(|\psi_1\rangle\!\langle\psi_1|, \ldots, |\psi_r\rangle\!\langle\psi_r|) = \frac{2}{r\,(r-1)} \sup_{k_{i,j} \in \mathbb{C}} \left\{ \sum_{i<j} \mathfrak{R}[k_{i,j} \langle \psi_i | \psi_j \rangle] : \sum_{i,j=1}^{r} k_{i,j} |i\rangle\!\langle j| \geq 0, \;\; k_{i,i} = 1 \;\forall i \in [r] \right\}. \tag{5.47}$$

Proof. See Appendix A.11. □

[3] The MATLAB code used to find the given counterexample is available as arXiv ancillary files along with the arXiv posting of this paper.

5.3 Secrecy-based multivariate fidelity

In this subsection, we introduce another multivariate quantum fidelity inspired by an existing secrecy measure in (5.48) and analyse its properties.

Recall the following secrecy measure from [KRS09], defined in terms of the average Uhlmann fidelity of a tuple of states to a reference state:

$$S(\rho_1, \ldots, \rho_r) \coloneqq \sup_{\sigma \in \mathscr{D}} \frac{1}{r} \sum_{i=1}^{r} F(\rho_i, \sigma) = \sup_{\sigma \in \mathscr{D}} F(\rho_{XA}, \rho_X \otimes \sigma), \tag{5.48}$$

where $\rho_{XA} \coloneqq \sum_{i=1}^{r} \frac{1}{r} |i\rangle\langle i|_X \otimes \rho_i$. Observe that $\rho_X = I/r$. The square of this secrecy measure was given a direct operational interpretation in [RASW23, Theorem 6] as the maximum success probability, in a quantum interactive proof, that a prover could pass a test for the states being similar. See [AKF22] for a comprehensive study on the optimization problem involved with this secrecy measure. In particular, by putting together Eqs. (22)–(23), Definition 3, Lemma 4, and Theorems 5 and 6 of [AKF22] and performing some other simplifications, we arrive at the following primal and dual SDP formulations of $S(\rho_1, \ldots, \rho_r)$:

$$\begin{aligned} & S(\rho_1, \ldots, \rho_r) \\ & = \frac{1}{r} \sup_{\substack{X_{i,j} = X_{j,i}^\dagger \in \mathscr{L}, \\ \forall i,j \in \{0,1,\ldots,r\}, \\ \sigma \geq 0}} \left\{ \begin{array}{c} \sum_{i=1}^{r} \Re[\operatorname{Tr}[X_{i,0}]] : \operatorname{Tr}[\sigma] = 1, \\ |0\rangle\langle 0| \otimes \sigma + \sum_{i=1}^{r} |i\rangle\langle i| \otimes \rho_i + \sum_{i,j=0: i \neq j}^{r} |i\rangle\langle j| \otimes X_{i,j} \geq 0 \end{array} \right\} \qquad (5.49) \\ & = \frac{1}{2r} \inf_{\substack{Y_1, \ldots, Y_r \geq 0, \\ \mu \geq 0}} \left\{ \begin{array}{c} \mu + \sum_{i=1}^{r} \operatorname{Tr}[Y_i \rho_i] : \\ |0\rangle\langle 0| \otimes \mu I_d + \displaystyle\sum_{i=1}^{r} |i\rangle\langle i| \otimes Y_i \geq \sum_{i=1}^{r} (|i\rangle\langle 0| + |0\rangle\langle i|) \otimes I_d \end{array} \right\}. \qquad (5.50) \end{aligned}$$

Using the secrecy measure, we define another multivariate fidelity as follows.

Definition 5.22 (Secrecy-based multivariate fidelity)**.** *Let $\rho_1, \ldots, \rho_r$ be quantum states. We define the multivariate secrecy fidelity as*

$$F_S(\rho_1, \ldots, \rho_r) \coloneqq \frac{1}{r-1} \left(r \left[S(\rho_1, \ldots, \rho_r) \right]^2 - 1 \right), \tag{5.51}$$

where $S(\rho_1, \ldots, \rho_r)$ is defined in (5.48).

Remark 5.23 (Equivalent formulation)**.** *By applying [RASW23, Eqs. (A70)–(A76)], the following equality holds:*

$$\left[S(\rho_1, \ldots, \rho_r) \right]^2 = \sup_{P_{T'RF \to T''F'} \in \mathscr{U}} \frac{1}{r^2} \sum_{x,y=1}^{r} \langle \phi^{\rho_x} |_{RA} \left(P^x_{R \to F'} \right)^\dagger P^y_{R \to F'} | \phi^{\rho_y} \rangle_{RA}, \tag{5.52}$$

where

$$P^x_{R \to F'} \coloneqq \langle x |_{T''} P_{T'RF \to T''F'} | x \rangle_{T'} | 0 \rangle_F, \tag{5.53}$$

with $P_{T'RF\to T''F'}$ being a unitary, and $|\phi^{\rho_x}\rangle$ is a purification of state ρ_x for all $x \in [r]$. Using this expression, the secrecy-based multivariate fidelity can be written somewhat similar to an average pairwise fidelity as follows:

$$\begin{aligned} F_S(\rho_1,\ldots,\rho_r) &= \sup_{P_{T'RF\to T''F'}\in\mathscr{U}} \left\{ \frac{2}{r(r-1)} \sum_{x<y} \mathfrak{R}\Big[\langle\phi^{\rho_x}|_{RA} \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'} |\phi^{\rho_y}\rangle_{RA}\Big] \right. \\ &\qquad \left. - \frac{1}{r(r-1)} \sum_{x=1}^{r} \langle\phi^{\rho_x}|_{RA} \left(I_R - \left(P^x_{R\to F'}\right)^\dagger P^x_{R\to F'}\right) |\phi^{\rho_x}\rangle_{RA} \right\} \qquad (5.54) \\ &\le \sup_{P_{T'RF\to T''F'}\in\mathscr{U}} \frac{2}{r(r-1)} \sum_{x<y} \mathfrak{R}\Big[\langle\phi^{\rho_x}|_{RA} \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'} |\phi^{\rho_y}\rangle_{RA}\Big], \qquad (5.55) \end{aligned}$$

where the inequality follows because, for all x, $P^x_{R\to F'}$ is a contraction.

5.3.1 Uniform continuity bound for secrecy-based multivariate fidelity

Proposition 5.24 (Uniform continuity)**.** *Let $\rho_1, \ldots, \rho_r, \sigma_1, \ldots, \sigma_r$ be quantum states, and let $\varepsilon \in [0,1]$ be such that*

$$\frac{1}{r}\sum_{i=1}^{r} d_B(\rho_i,\sigma_i) \le \varepsilon, \qquad (5.56)$$

where the Bures distance d_B is defined in (5.14). Then,

$$|F_S(\rho_1,\ldots,\rho_r) - F_S(\sigma_1,\ldots,\sigma_2)| \le 2\sqrt{2}\left(\frac{r}{r-1}\right)\varepsilon. \qquad (5.57)$$

Proof. See Appendix A.12. □

5.3.2 Properties of secrecy-based multivariate fidelity

Next we prove that secrecy-based multivariate fidelity also satisfies the desirable properties of a multivariate fidelity from Definition 5.1.

Theorem 5.25 (Properties of secrecy-based multivariate fidelity)**.** *The secrecy-based multivariate fidelity satisfies all the desired properties of a multivariate fidelity listed in Definition 5.1. For commuting states, it reduces to the classical average pairwise fidelity.*

Proof. See Appendix A.13. □

Similar to the average pairwise fidelity and multivariate SDP fidelity, F_S also satisfies the coarse-graining property with an extra additive factor.

Proposition 5.26 (Coarse-graining property)**.** *Let $r, m \in \mathbb{N}$, and let $\rho_1, \ldots, \rho_{r+m}$ be quantum states. We have the following inequality:*

$$F_S(\rho_1,\ldots,\rho_r) \le \frac{(r+m)(r+m-1)}{r(r-1)} F_S(\rho_1,\ldots,\rho_r,\ldots,\rho_{r+m}) + \frac{m}{r(r-1)}. \qquad (5.58)$$

Proof. See Appendix A.14. □

Remark 5.27 (Super-multiplicativity)**.** *Through numerical calculations, we found that there exists a tuple of states for which super-multiplicativity does not hold for the SDP fidelity, i.e.,* $(F_S(\rho_1, \ldots, \rho_r))^n > F_S\big(\rho_1^{\otimes n}, \ldots, \rho_r^{\otimes n}\big)$ *for some* r *and* $(\rho_i)_{i=1}^r$. *The same example that violates super-multiplicativity of SDP multivariate fidelity in Remark 5.18, also violates the super-multiplicativity of secrecy-based multivariate fidelity.*

5.4 Relations between proposed generalizations

In this section, we present some inequalities relating the average pairwise Holevo fidelity, the secrecy-based multivariate fidelity, the multivariate SDP fidelity, and the average pairwise Uhlmann fidelity. In particular, we show that the multivariate SDP fidelity and secrecy-based multivariate fidelity are sandwiched between the average pairwise Holevo fidelity and the average pairwise Uhlmann fidelity, from below, and above, respectively.

Theorem 5.28 (Inequalities relating multivariate fidelities)**.** *For a tuple* $(\rho_i)_{i=1}^r$ *of states, the multivariate SDP fidelity and secrecy-based multivariate fidelity lie between the average pairwise Holevo fidelity and the average pairwise Uhlmann fidelity:*

$$F_H(\rho_1, \ldots, \rho_r) \le F_S(\rho_1, \ldots, \rho_r) \le F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) \le F_U(\rho_1, \ldots, \rho_r) \le \sqrt{F_H(\rho_1, \ldots, \rho_r)}. \tag{5.59}$$

Proof. See Appendix A.15. □

In Theorem 5.28 we showed that multivariate SDP fidelity is bounded from above by the average pairwise Uhlmann fidelity. Next, we show that it is also bounded from below by a function of pairwise Uhlmann fidelities (note that it is not an average).

Proposition 5.29 (Lower bound)**.** *The multivariate SDP fidelity is bounded from below by a multiple of the maximum of sums of pairwise fidelities of the given states. More precisely, given quantum states* $\rho_1, \ldots, \rho_r$*, we have*

$$\frac{2}{r(r-1)} \sup_{\pi \in S_r} \sum_{i=1}^{\lfloor \frac{r}{2} \rfloor} F(\rho_{\pi(i)}, \rho_{\pi(i+\lfloor r/2 \rfloor)}) \le F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r). \tag{5.60}$$

Proof. See Appendix A.16. □

5.5 Minimal and maximal extensions of multivariate classical fidelities

One way to construct a multivariate quantum fidelity is by optimizing classical multivariate fidelities, and there are at least two ways of doing so. The first way is to perform a measurement of the r quantum states, resulting in r probability distributions, and then to minimize a classical multivariate fidelity of the r distributions over all possible measurements. The measurement is also called a quantum-to-classical transformation. The second way is to begin with r probability distributions and act on them with a preparation channel that realizes the r quantum states, and then to maximize a classical multivariate fidelity over all possible preparation channels. A preparation channel is also known as a classical-to-quantum transformation. The former approach leads to a quantity that we call a maximal extension of multivariate fidelity (or measured multivariate fidelity), and the latter leads to a minimal extension of multivariate fidelity (or preparation multivariate fidelity).

These concepts have been introduced and studied extensively in prior work. The measured relative entropy was defined in [Don86,HP91,Pia09] and its extension to Petz-Rényi divergences in [Fuc96, Eqs. (3.116)–(3.117)] (see also [BFT15,MH22]). The preparation relative entropy was defined in [Mat10a] and shown therein to be equal to the Belavkin–Staszewski relative entropy from [BS82]. The preparation fidelity and more general f-divergences were defined in [Mat10b, Mat18] and studied further in [HM17, Hia19]. See also [GT20] for maximal and minimal extensions of resource measures, including generalized fidelity for sub-normalized states. Most recently, maximal and minimal extensions of multivariate Chernoff divergence were defined in [MNW23, Section V-A].

We define the maximal and minimal extensions of classical multivariate fidelity as follows.

Definition 5.30 (Maximal extension of multivariate fidelity)**.** *We define a maximal extension of multivariate fidelity (also called a measured multivariate fidelity) as*

$$\overline{\mathbf{F}}(\rho_1, \ldots, \rho_r) := \inf_{(\Lambda_x)_x} \mathbf{F}(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)) \tag{5.61}$$

where $\mathbf{F}$ *is one of the multivariate fidelities listed in (i) of Definition 5.1 and* $\Lambda(\rho_i) := \sum_x \operatorname{Tr}[\Lambda_x \rho] |x\rangle\langle x|$ *is the measurement channel associated with the POVM* $(\Lambda_x)_x$ *(see also (2.1)).*

Definition 5.31 (Minimal extension of multivariate fidelity)**.** *We define a minimal extension of multivariate fidelity (also called a preparation multivariate fidelity) as*

$$\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) := \sup_{\mathcal{P} \in \mathrm{CPTP},\, \omega_i \in \mathscr{D}_c \forall i \in [r]} \left\{ \mathbf{F}(\omega_1, \ldots, \omega_r) : \mathcal{P}(\omega_i) = \rho_i \, \forall i \in [r] \right\}, \tag{5.62}$$

where $\mathbf{F}$ *is one of the multivariate fidelities listed in (i) of Definition 5.1 and* $\mathcal{P}$ *is a classical-quantum channel satisfying* $\mathcal{P}(\omega_i) = \rho_i$ *for all* $i \in [r]$ *and* $\mathscr{D}_c$ *denotes the set of all states diagonal in the standard basis (so that* $\mathscr{D}_c$ *is a set of commuting states).*

Note that, by the above definitions and data processing of multivariate fidelity (Definition 5.1), we have the following inequalities:

$$\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) \leq \mathbf{F}(\rho_1, \ldots, \rho_r) \leq \overline{\mathbf{F}}(\rho_1, \ldots, \rho_r). \tag{5.63}$$

Proofs for these inequalities follow the same approach used in [MNW23, Appendices E & F]. Both the maximal and minimal extensions also satisfy the data-processing inequality.

Remark 5.32 (Bivariate setting)**.** *In the bivariate setting, the maximal extension of the classical fidelity is equal to the Uhlmann fidelity, and the minimal extension is equal to the geometric fidelity, as defined in* (5.65)*: By [Mat10b, Theorem 1] we have*

$$\underline{\mathbf{F}}(\rho,\sigma) \le \mathbf{F}(\rho,\sigma) \le \overline{\mathbf{F}}(\rho,\sigma) = F(\rho,\sigma), \tag{5.64}$$

where the equality follows from (3.2)*. Furthermore, [Mat10b, Remark 3] shows that*

$$\underline{\mathbf{F}}(\rho,\sigma) = \inf_{\varepsilon>0} \operatorname{Tr}[\rho(\varepsilon)\#\sigma(\varepsilon)], \tag{5.65}$$

where $A\#B$ denotes the geometric mean of the positive definite operators A and B:

$$A\#B \coloneqq A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2}, \tag{5.66}$$

$\rho(\varepsilon) \coloneqq \rho + \varepsilon I$, and $\sigma(\varepsilon) \coloneqq \sigma + \varepsilon I$. Due to this formulation, the minimal extension of the classical bivariate fidelity is also called the geometric fidelity. In [CS20], the geometric fidelity was called the Matsumoto fidelity, and several of its properties were established by means of semi-definite programming.

5.5.1 Measured multivariate fidelity

In this section, we consider various properties satisfied by the measured multivariate fidelity defined in Definition 5.30.

Proposition 5.33 (Properties of measured multivariate fidelity)**.** *If the underlying classical multivariate fidelity satisfies data processing, symmetry, faithfulness, orthogonality, and the direct-sum property, then the maximal extension satisfies data processing, symmetry, faithfulness, orthogonality, and the direct-sum property.*

Proof. See Appendix A.17. □

By choosing the underlying classical multivariate fidelity to be the average pairwise fidelity, here we define measured average pairwise fidelity.

Definition 5.34 (Measured average pairwise fidelity)**.** *For states $\rho_1, \ldots, \rho_r$, the measured average pairwise fidelity is defined as*

$$F_{\mathrm{M}}(\rho_1, \ldots, \rho_r) \coloneqq \inf_{(\Lambda_x)_x} \frac{2}{r(r-1)} \sum_{i<j} F\big(\Lambda(\rho_i), \Lambda(\rho_j)\big) \tag{5.67}$$

where $\Lambda(\rho_i) \coloneqq \sum_x \operatorname{Tr}[\Lambda_x \rho] |x\rangle\!\langle x|$ is the measurement channel induced by applying the POVM $(\Lambda_x)_x$ on the state ρ_i (see also (2.1)*).*

Note that Definition 5.34 is the largest quantum generalization of average pairwise classical fidelity that satisfies data processing.

When $\mathbf{F}$ is chosen to be a multivariate fidelity that generalizes average pairwise classical fidelity, we have that

$$\overline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = F_{\mathrm{M}}(\rho_1, \ldots, \rho_r). \tag{5.68}$$

Similarly, when $\mathbf{F}$ is chosen to be a multivariate fidelity that generalizes Matusita multivariate fidelity, we have that

$$\overline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = \inf_{(\Lambda_x)_x} F_r(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)). \tag{5.69}$$

The next proposition generalizes Proposition 4.6 to the case of measured multivariate fidelities.

Proposition 5.35 (Measured Matusita fidelity and average pairwise fidelity)**.**
The following inequality holds:

$$F_{\mathrm{M}}(\rho_1, \ldots, \rho_r) \geq \inf_{(\Lambda_x)_x} F_r(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)), \tag{5.70}$$

where the Matusita multivariate fidelity F_r is defined in (4.2).

Proof. The proof of the first statement follows from applying Proposition 4.6 for each measurement channel Λ and then taking the infimum over all measurement channels. □

Corollary 5.36 (Uniform continuity of measured average pairwise fidelity)**.**
Let $\rho_1, \ldots, \rho_r, \sigma_1, \ldots, \sigma_r$ be quantum states, and let $\varepsilon > 0$ be such that $\frac{1}{r}\sum_{i=1}^{r} d_B(\rho_i, \sigma_i) \leq \varepsilon$. Then,

$$|F_{\mathrm{M}}(\rho_1, \ldots, \rho_r) - F_{\mathrm{M}}(\sigma_1, \ldots, \sigma_r)| \leq 2\sqrt{2}\varepsilon. \tag{5.71}$$

Furthermore, if $\varepsilon > 0$ is such that $\frac{1}{r}\sum_{i=1}^{r} F(\rho_i, \sigma_i) \geq 1 - \varepsilon$, then

$$|F_{\mathrm{M}}(\rho_1, \ldots, \rho_r) - F_{\mathrm{M}}(\sigma_1, \ldots, \sigma_r)| \leq \frac{r}{r-1}\sqrt{\varepsilon\,(2-\varepsilon)}. \tag{5.72}$$

Proof. For $F_{\mathrm{M}}(\rho_1, \ldots, \rho_r)$, the underlying classical multivariate fidelity is the average pairwise fidelity. Considering the reduction to the classical case for the quantum generalizations we have proposed, the following equalities hold for every measurement channel Λ:

$$F(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)) = F_U(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)) = F_{\mathrm{SDP}}(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)). \tag{5.73}$$

Since classical multivariate fidelity satisfies data processing we also have the following implications:

$$\frac{1}{r}\sum_{i=1}^{r} d_B(\rho_i, \sigma_i) \leq \varepsilon \implies \frac{1}{r}\sum_{i=1}^{r} d_B(\Lambda(\rho_i), \Lambda(\sigma_i)) \leq \varepsilon, \tag{5.74}$$

$$\frac{1}{r}\sum_{i=1}^{r} F(\rho_i,\sigma_i) \geq 1-\varepsilon \implies \frac{1}{r}\sum_{i=1}^{r} F(\Lambda(\rho_i),\Lambda(\sigma_i)) \geq 1-\varepsilon. \tag{5.75}$$

Now consider that

$$F_{\mathrm{M}}(\rho_1,\ldots,\rho_r) - F_{\mathrm{M}}(\sigma_1,\ldots,\sigma_r)$$
$$= \inf_{(\Lambda_x)_x} F(\Lambda(\rho_1),\ldots,\Lambda(\rho_r)) - \inf_{(\Lambda'_x)_x} F(\Lambda'(\sigma_1),\ldots,\Lambda'(\sigma_r)) \tag{5.76}$$
$$= \sup_{(\Lambda'_x)_x} \left(\inf_{(\Lambda_x)_x} F(\Lambda(\rho_1),\ldots,\Lambda(\rho_r)) - F(\Lambda'(\sigma_1),\ldots,\Lambda'(\sigma_r)) \right) \tag{5.77}$$
$$\leq \sup_{(\Lambda'_x)_x} \left(F(\Lambda'(\rho_1),\ldots,\Lambda'(\rho_r)) - F(\Lambda'(\sigma_1),\ldots,\Lambda'(\sigma_r)) \right). \tag{5.78}$$

At this point, we conclude the two inequalities by applying Proposition 5.4 and Corollary 5.14 along with (5.73). ☐

Remark 5.37 (Uniform continuity of measured Matusita multivariate fidelity)**.** *Using a similar proof technique as given in the proof of Corollary 5.36 and using uniform continuity of the Matusita multivariate fidelity in Proposition 4.3, it is possible to obtain uniform continuity bound for measured Matusita multivariate fidelity as well under the assumption that* $\frac{1}{r}\left(\sum_{i=1}^{r}[d_B(\rho_i,\sigma_i)]^2\right) \leq \varepsilon$.

The following proposition establishes inequalities between the measured average pairwise fidelity and average pairwise Uhlmann fidelity.

Proposition 5.38 (Multivariate fidelities and their measured variant)**.** *The following inequality holds for every tuple* $(\rho_1,\ldots,\rho_r)$ *of states:*

$$F_{\mathrm{M}}(\rho_1,\ldots,\rho_r) \geq F_U(\rho_1,\ldots,\rho_r). \tag{5.79}$$

Proof. Consider the following:

$$F_{\mathrm{M}}(\rho_1,\ldots,\rho_r) = \inf_{(\Lambda_x)_x} \frac{2}{r(r-1)} \sum_{i<j} F\big(\Lambda(\rho_i),\Lambda(\rho_j)\big) \tag{5.80}$$
$$\geq \frac{2}{r(r-1)} \sum_{i<j} \inf_{(\Lambda_x)_x} F\big(\Lambda(\rho_i),\Lambda(\rho_j)\big) \tag{5.81}$$
$$= \frac{2}{r(r-1)} \sum_{i<j} F(\rho_i,\rho_j), \tag{5.82}$$

where the first inequality follows from placing the infimum inside the sum, the second equality follows from the Fuchs–Caves characterization of Uhlmann fidelity in terms of the measured bivariate fidelity (see (3.2)). ☐

Remark 5.39 (Multivariate SDP fidelity and the measured version)**.** *When $r = 2$, the Uhlmann fidelity is equal to the measured fidelity, as stated in (3.2). However, when $r > 2$, through numerical evaluations, it follows that there exists a tuple of states satisfying $F_{\mathrm{M}}(\rho_1, \ldots, \rho_r) > F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r)$. This is numerically validated by showing that $F_U \geq F_{\mathrm{SDP}}$ in (5.59) is a strict inequality for some tuples of states and then combining with (5.79). One such tuple of states (up to numerical approximations of MATLAB) is the following:*

$$\rho_1 = \begin{bmatrix} 0.3465 + 0.0000i & -0.2036 + 0.2643i \\ -0.2036 - 0.2643i & 0.6535 + 0.0000i \end{bmatrix}, \tag{5.83}$$

$$\rho_2 = \begin{bmatrix} 0.6546 + 0.0000i & -0.3308 - 0.2297i \\ -0.3308 + 0.2297i & 0.3454 + 0.0000i \end{bmatrix}, \tag{5.84}$$

$$\rho_3 = \begin{bmatrix} 0.6169 + 0.0000i & 0.0327 - 0.0321i \\ 0.0327 + 0.0321i & 0.3831 + 0.0000i \end{bmatrix}. \tag{5.85}$$

The MATLAB code used to find the given counterexample is available as arXiv ancillary files along with the arXiv posting of this paper.

Furthermore, it is interesting to determine whether there exists a tuple of states for which $F_M(\rho_1, \ldots, \rho_r) > F_U(\rho_1, \ldots, \rho_r)$—we leave this as an open question.

5.5.2 Preparation multivariate fidelity

The preparation multivariate fidelity, as outlined in Definition 5.31, is also known as the minimal extension of multivariate fidelity. Next, we discuss various properties satisfied by the preparation multivariate fidelity.

Proposition 5.40 (Properties of preparation multivariate fidelity)**.**
The preparation multivariate fidelity in Definition 5.31 satisfies data processing, symmetry, faithfulness, and the direct-sum property, as listed in Definition 5.1.

Furthermore, when the states are orthogonal to each other, we have $\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = 0$ (weak orthogonality).

Proof. See Appendix A.18. ☐

Remark 5.41 (Lack of uniform continuity)**.** *The minimal extension of the multivariate fidelity does not satisfy uniform continuity in general. To see that consider the following counterexample: Fix $r = 2$ and choose*

$$\rho_1 = \rho_2 = \sigma_1 = |0\rangle\!\langle 0| \quad \text{and} \quad \sigma_2 = |\psi\rangle\!\langle\psi|, \tag{5.86}$$

where $|\psi\rangle := \sqrt{1-\varepsilon}|0\rangle + \sqrt{\varepsilon}|1\rangle$ for $\varepsilon > 0$. Then, we obtain the following by using the properties of Uhlmann fidelity from Proposition 3.2:

$$F(\rho_1, \sigma_1) = 1, \tag{5.87}$$

$$F(\rho_2, \sigma_2) = \sqrt{1-\varepsilon}, \tag{5.88}$$

$$\underline{\mathbf{F}}(\rho_1, \rho_2) = 1, \tag{5.89}$$

$$\underline{\mathbf{F}}(\sigma_1, \sigma_2) = 0, \tag{5.90}$$

where the third equality follows by the faithfulness of preparation fidelity and the last equality by the fact that the minimal extension of the fidelity between two pure states evaluates to zero [Mat10b, Eq. (23)] (see also [CS20, Lemma 3.17 and Eq. (3.59)]). With the above evaluations we see that $\frac{1}{2}\left(F(\rho_1, \sigma_1) + F(\rho_2, \sigma_2)\right) \geq 1 - \varepsilon$ *and yet*

$$|\underline{\mathbf{F}}(\rho_1, \rho_2) - \underline{\mathbf{F}}(\sigma_1, \sigma_2)| = 1, \tag{5.91}$$

implying that the preparation multivariate fidelity does not satisfy uniform continuity in general.

5.6 Multivariate log-Euclidean fidelity

In this section, we recall the definition of multivariate log-Euclidean divergences [MBV22, MNW23], we show how one possible quantum Matusita fidelity is a special case, and we give operational interpretations of them in terms of quantum hypothesis testing with an arbitrarily varying null hypothesis, specifically making use of the main result of [Nö14].

Definition 5.42 (Multivariate log-Euclidean divergence)**.** *Let* $\rho_1, \ldots, \rho_r$ *be positive definite states and let* $s = (s_1, \ldots, s_r)$ *be a probability distribution. The multivariate log-Euclidean divergence is defined as*

$$H_s(\rho_1, \ldots, \rho_r) := -\ln \operatorname{Tr}\left[\exp\left(\sum_{i=1}^{r} s_i \ln \rho_i\right)\right]. \tag{5.92}$$

For general states, the multivariate log-Euclidean divergence is defined as

$$H_s(\rho_1, \ldots, \rho_r) := \inf_{\varepsilon > 0} H_s(\rho_1(\varepsilon), \ldots, \rho_r(\varepsilon)) = \lim_{\varepsilon \to 0^+} H_s(\rho_1(\varepsilon), \ldots, \rho_r(\varepsilon)), \tag{5.93}$$

where $\rho_i(\varepsilon) := \rho_i + \varepsilon I$ *for all* $i \in [r]$ *and the rightmost equality follows from operator monotonicity of the logarithm and monotonicity of* $(\cdot) \to \operatorname{Tr}[\exp(\cdot)]$*. The argument of the trace in* (5.92) *is known as the log-Euclidean Frechet mean [AFPA07, Section 3.4].*

For simplicity, in most of what follows we focus on positive definite states and note that all of the statements can be extended as in (5.93).

Now, we define the multivariate log-Euclidean fidelity that extends (4.2) for commuting states to the general quantum setting.

Definition 5.43 (Multivariate log-Euclidean fidelity)**.** *Let* $(\rho_i)_{i=1}^r$ *be a tuple of states. We define the multivariate log-Euclidean fidelity for positive definite states as*

$$F_r^\flat(\rho_1, \ldots, \rho_r) := \operatorname{Tr}\left[\exp\left(\frac{1}{r}\sum_{i=1}^{r} \ln \rho_i\right)\right], \tag{5.94}$$

and for the general case as

$$F_r^\flat(\rho_1,\ldots,\rho_r) \coloneqq \inf_{\varepsilon>0} F_r^\flat(\rho_1(\varepsilon),\ldots,\rho_r(\varepsilon)) = \lim_{\varepsilon\to 0^+} F_r^\flat(\rho_1(\varepsilon),\ldots,\rho_r(\varepsilon)), \tag{5.95}$$

where $\rho_i(\varepsilon) \coloneqq \rho_i + \varepsilon I$.

Note that Definition 5.43 reduces to (1.8) for the bivariate setting.

Remark 5.44. *When s is uniform (i.e., $s_i = 1/r$ for all $i \in [r]$), the multivariate log-Euclidean divergence is equal to the negative logarithm of the multivariate log-Euclidean fidelity in Definition 5.43:*

$$-\ln\left(F_r^\flat(\rho_1,\ldots,\rho_r)\right) = -\ln \operatorname{Tr}\left[\exp\left(\sum_{i=1}^{r} \frac{1}{r}\ln\rho_i\right)\right]. \tag{5.96}$$

For the special case of commuting states, the multivariate log-Euclidean divergence reduces to the negative logarithm of the Hellinger transform from (4.8):

$$H_s(\rho_1,\ldots,\rho_r) = -\ln \operatorname{Tr}\left[\prod_{i=1}^{r} \rho_i^{s_i}\right]. \tag{5.97}$$

Theorem 5.45 (Properties of multivariate log-Euclidean fidelity)**.** *The multivariate log-Euclidean fidelity in Definition 5.43 satisfies several desirable properties of a multivariate fidelity listed in Definition 5.1, including reduction to classical Matusita multivariate fidelity in (4.2), data processing, symmetry, faithfulness, direct-sum property, and uniform continuity.*

If all states are orthogonal to each other, i.e., $\rho_i\rho_j = 0$ for $i \neq j$, $i,j \in [r]$, then $F_r^\flat(\rho_1,\ldots,\rho_r) = 0$ (weak orthogonality). However, $F_r^\flat(\rho_1,\ldots,\rho_r) = 0$ does not imply that the states are orthogonal to each other.

Proof. See Appendix A.19 for a proof of the properties. As a counterexample for one direction of the orthogonality property mentioned above, consider the diagonal states formed by the probability vectors given in (4.3). □

Next, we establish a quantum generalization of the inequality from Proposition 4.6.

Proposition 5.46 (Multivariate log-Euclidean fidelity upper bound)**.** *Let $(\rho_i)_{i=1}^r$ be a tuple of states. Then, we have*

$$F_r^\flat(\rho_1,\ldots,\rho_r) \leq \frac{2}{r(r-1)} \sum_{i<j} F_\flat(\rho_i,\rho_j). \tag{5.98}$$

Proof. For positive definite states $\rho_1, \ldots, \rho_r$, consider that

$$\frac{2}{r\,(r-1)} \sum_{i<j} F_\flat(\rho_i,\rho_j)$$

$$
= \frac{2}{r\,(r-1)} \sum_{i<j} \exp\!\left(- \inf_{\omega \in \mathscr{D}} \left[\frac{1}{2} D(\omega\|\rho_i) + \frac{1}{2} D(\omega\|\rho_j) \right] \right) \tag{5.99}
$$

$$
\geq \exp\!\left(- \frac{2}{r\,(r-1)} \sum_{i<j} \inf_{\omega \in \mathscr{D}} \left[\frac{1}{2} D(\omega\|\rho_i) + \frac{1}{2} D(\omega\|\rho_j) \right] \right) \tag{5.100}
$$

$$
\geq \exp\!\left(- \inf_{\omega \in \mathscr{D}} \frac{2}{r\,(r-1)} \sum_{i<j} \left[\frac{1}{2} D(\omega\|\rho_i) + \frac{1}{2} D(\omega\|\rho_j) \right] \right) \tag{5.101}
$$

$$
= \exp\!\left(- \inf_{\omega \in \mathscr{D}} \frac{1}{r} \sum_{i=1}^{r} D(\omega\|\rho_i) \right) \tag{5.102}
$$

$$
= F_r^{\flat}(\rho_1, \ldots, \rho_r), \tag{5.103}
$$

where the first equality follows from (5.110); the first inequality from convexity of the exponential function; the second inequality follows because

$$
\inf_{\omega \in \mathscr{D}} \frac{2}{r\,(r-1)} \sum_{i<j} \left[\frac{1}{2} D(\omega\|\rho_i) + \frac{1}{2} D(\omega\|\rho_j) \right] \geq \frac{2}{r\,(r-1)} \sum_{i<j} \inf_{\omega \in \mathscr{D}} \left[\frac{1}{2} D(\omega\|\rho_i) + \frac{1}{2} D(\omega\|\rho_j) \right]; \tag{5.104}
$$

the penultimate equality is from a counting argument, and the final equality is again from (5.110). The general case follows by a limiting argument. □

5.6.1 Connection to oveloh information

The generalized Holevo information of a tuple of states $(\rho_i)_i$ with uniform prior distribution is defined as

$$
\inf_{\sigma_A \in \mathscr{D}} \boldsymbol{D}(\rho_{XA} \| \rho_X \otimes \sigma_A), \tag{5.105}
$$

where $\boldsymbol{D}$ is a generalized divergence that satisfies data processing under quantum channels [SW12, LKDW18] and

$$
\rho_{XA} \coloneqq \frac{1}{r} \sum_{i=1}^{r} |i\rangle\!\langle i|_X \otimes \rho_i. \tag{5.106}
$$

Fix $\alpha \in (0,1) \cup (1,\infty)$. The Petz–Rényi relative entropy of a state ρ and a PSD operator σ is defined as [Pet85, Pet86]

$$
D_\alpha(\rho\|\sigma) \coloneqq \begin{cases} \frac{1}{\alpha-1} \ln \operatorname{Tr}[\rho^\alpha \sigma^{1-\alpha}], & \text{if } \alpha \in (0,1) \vee (\alpha > 1 \wedge \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)) \\ +\infty, & \text{otherwise} \end{cases} \tag{5.107}
$$

It is a generalized divergence for $\alpha \in [0,1) \cup (1,2]$ [Pet86]. Taking the limit $\alpha \to 1$ gives the quantum relative entropy defined as [Ume62]

$$
D(\rho\|\sigma) \coloneqq \begin{cases} \operatorname{Tr}[\rho(\log\rho - \log\sigma)], & \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma), \\ +\infty, & \text{otherwise.} \end{cases} \tag{5.108}
$$

In the spirit of the definition of *lautum information* (reverse of mutual information) [PV08], we define the *oveloh information* (reverse of Holevo information):

$$\mathscr{O}(X;A)_\rho := \inf_{\sigma\in\mathscr{D}} D(\rho_X\otimes\sigma\|\rho_{XA}) = \inf_{\sigma\in\mathscr{D}} \frac{1}{r}\sum_{i=1}^{r} D(\sigma\|\rho_i), \tag{5.109}$$

where ρ_{XA} is given in (5.106). Also note that oveloh information is a special case of log-Euclidean divergence with the choice of s being a uniform distribution (see (5.118)). To this end, $\exp\left[-\mathscr{O}(X;A)_\rho\right]$ is precisely the multivariate log-Euclidean fidelity in Definition 5.43:

$$F_r^\flat(\rho_1,\ldots,\rho_r) = \exp\left[-\mathscr{O}(X;A)_\rho\right] = \exp\left(-\inf_{\sigma\in\mathscr{D}} \frac{1}{r}\sum_{i=1}^{r} D(\sigma\|\rho_i)\right). \tag{5.110}$$

With this connection, we establish a relationship between Holevo and oveloh information as given below.

Corollary 5.47 (Holevo and oveloh information)**.** *Let $(\rho_i)_i$ be a tuple of states. Then the following inequality holds:*

$$\exp\left[-\inf_{\sigma\in\mathscr{D}} D_{\frac{1}{2}}(\rho_{XA}\|\rho_X\otimes\sigma)\right] \geq \frac{1}{r} + \frac{r-1}{r}\exp[-\mathscr{O}(X;A)_\rho]\,, \tag{5.111}$$

where ρ_{XB} is given in (5.106) and $D_{1/2}(\cdot\|\cdot)$ is the Petz–Rényi relative entropy of order $1/2$ in (5.107).

Proof. By using (5.110) together with (5.109), we have

$$\begin{aligned}\exp\left[-\inf_{\sigma\in\mathscr{D}} D(\rho_X\otimes\sigma\|\rho_{XA})\right] &= F_r^\flat(\rho_1,\ldots,\rho_r) && (5.112)\\ &\leq \frac{2}{r(r-1)}\sum_{i<j} F_\flat(\rho_i,\rho_j), && (5.113)\end{aligned}$$

where the inequality follows from Proposition 5.46. Then, by recalling that z-fidelities are antimonotone for $z \geq 1/2$ together with (1.8), we have $F_\flat(\rho,\sigma) \leq F_H(\rho,\sigma)$ leading to

$$\exp\left[-\inf_{\sigma\in\mathscr{D}} D(\rho_X\otimes\sigma\|\rho_{XA})\right] \leq \frac{2}{r(r-1)}\sum_{i<j} F_H(\rho_i,\rho_j). \tag{5.114}$$

Also, using (A.208), we obtain the following:

$$\exp\left[-\inf_{\sigma\in\mathscr{D}} D_{\frac{1}{2}}(\rho_{XA}\|\rho_X\otimes\sigma)\right] = \frac{1}{r} + \frac{r-1}{r}\left[\frac{2}{r(r-1)}\sum_{i<j} F_H(\rho_i,\rho_j)\right]. \tag{5.115}$$

Combining (5.114) and the previous equality completes the proof. ∎

Note that the inequality stated above is tighter than that which one would obtain by using the facts that

$$\inf_{\sigma\in\mathscr{D}} D_{\frac{1}{2}}(\rho_{XA}\|\rho_X\otimes\sigma) = \inf_{\sigma\in\mathscr{D}} D_{\frac{1}{2}}(\rho_X\otimes\sigma\|\rho_{XA}), \tag{5.116}$$

$$\inf_{\sigma\in\mathscr{D}} D_{\frac{1}{2}}(\rho_X\otimes\sigma\|\rho_{XA}) \leq \inf_{\sigma\in\mathscr{D}} D(\rho_X\otimes\sigma\|\rho_{XA}). \tag{5.117}$$

5.6.2 Operational interpretation of multivariate log-Euclidean divergences

The following formula is known (see [MBV22], and [MNW23, Appendix J], as well as [BGJ19, Eq. (47)] for a related observation):

$$H_s(\rho_1,\ldots,\rho_r) = \inf_{\omega\in\mathscr{D}} \sum_{i=1}^{r} s_i D(\omega\|\rho_i) = \inf_{\omega\in\mathscr{D}} D(\pi(s)\otimes\omega\|\rho(s)), \tag{5.118}$$

where $\mathscr{D}$ is the set of density matrices and the quantum relative entropy is given in (5.108). The second equality follows from the direct-sum property of quantum relative entropy [KW20, Eq. (7.2.27)], along with the following definitions:

$$\pi(s) \coloneqq \sum_{i=1}^{r} s_i|i\rangle\!\langle i|, \qquad \rho(s) \coloneqq \sum_{i=1}^{r} s_i|i\rangle\!\langle i|\otimes\rho_i. \tag{5.119}$$

Now we present the hypothesis testing setting and result that provides an operational interpretation of the multivariate log-Euclidean divergence. The setting was considered in a general way in [Nö14] and is known as quantum hypothesis testing with arbitrarily varying null hypothesis and simple alternative hypothesis. The null hypothesis for our case is that a state is chosen adversarially from the following set:

$$\Omega_s^{(n)} \coloneqq \left\{ \bigotimes_{i=1}^{n} (\pi(s)\otimes\omega_i) : \omega_i\in\mathscr{D}\ \forall i\in[n] \right\}. \tag{5.120}$$

This set is an example of what is called an arbitrarily varying source in [Nö14]. The alternative hypothesis is that the state $\rho(s)^{\otimes n}$ is chosen. The receiver of the unknown state is allowed to perform a measurement $\{M^{(n)}, I^{\otimes n} - M^{(n)}\}$ acting on all n systems in order to figure out which is the case, such that $0 \leq M^{(n)} \leq I^{\otimes n}$. The outcome $M^{(n)}$ is associated with choosing the null hypothesis, and the outcome $I^{\otimes n} - M^{(n)}$ is associated with choosing the alternative hypothesis. Formally, the type I and type II error probabilities are respectively defined in this setting as

$$\alpha(M^{(n)}) \coloneqq \sup_{\sigma^{(n)}\in\Omega_s^{(n)}} \operatorname{Tr}\!\left[\left(I^{\otimes n} - M^{(n)}\right)\sigma^{(n)}\right], \tag{5.121}$$

$$\beta(M^{(n)}) \coloneqq \operatorname{Tr}[M^{(n)}\rho(s)^{\otimes n}]. \tag{5.122}$$

The fact that the type I error probability is maximized over arbitrary $\sigma^{(n)}\in\Omega_s^{(n)}$ is what makes the hypothesis testing scenario adversarial, as the state $\sigma^{(n)}$ is selected from $\Omega_s^{(n)}$ in such a way, for a given measurement operator $M^{(n)}$, to make the type I error probability as large as possible. For $\varepsilon \in [0,1]$ and $n \in \mathbb{N}$, the optimal non-asymptotic type II error exponent is defined as

$$\frac{1}{n} D_H^{\varepsilon}(\Omega_s^{(n)} \| \rho(s)^{\otimes n}) := -\frac{1}{n} \ln\left(\inf_{M^{(n)} \in \mathcal{M}^{(n)}} \left\{ \beta(M^{(n)}) : \alpha(M^{(n)}) \leq \varepsilon \right\} \right), \tag{5.123}$$

where $\mathcal{M}^{(n)}$ denotes the set of all measurement operators:

$$\mathcal{M}^{(n)} := \left\{ M^{(n)} : 0 \leq M^{(n)} \leq I^{\otimes n} \right\}. \tag{5.124}$$

Corollary 5.48. *Let s be a probability distribution on $[r]$, and let $\rho_1, \ldots, \rho_r$ be states. For all $\varepsilon \in (0,1)$, the following equality holds:*

$$\lim_{n \to \infty} \frac{1}{n} D_H^{\varepsilon}(\Omega_s^{(n)} \| \rho(s)^{\otimes n}) = H_s(\rho_1, \ldots, \rho_r), \tag{5.125}$$

where $\frac{1}{n} D_H^{\varepsilon}(\Omega_s^{(n)} \| \rho(s)^{\otimes n})$ is defined in (5.123) and the multivariate log-Euclidean divergence in Definition 5.42.

Proof. The proof follows by applying [Nö14, Theorem 1] and the identity in (5.118). □

Thus, we obtain a fundamental operational meaning for $H_s(\rho_1, \ldots, \rho_r)$ in this hypothesis testing scenario, as the optimal asymptotic type II error exponent. By picking s to be uniform, we obtain an operational meaning for the negative logarithm of the multivariate log-Euclidean fidelity in (5.96).

5.7 Average k-wise log-Euclidean fidelities

As a generalization of the ideas presented in Section 4.3 for the classical case, we define the average k-wise log-Euclidean fidelities as follows:

Definition 5.49 (Average k-wise log-Euclidean fidelity)**.** *For $r \in \{3, 4, \ldots\}$, let $\rho_1, \ldots, \rho_r$ be quantum states. For $k \in \{2, 3, \ldots, r\}$, define the average k-wise log-Euclidean fidelity as follows:*

$$F_{k,r}^{\flat}(\rho_1, \ldots, \rho_r) := \frac{1}{\binom{r}{k}} \sum_{i_1 < \cdots < i_k} F_k^{\flat}(\rho_{i_1}, \ldots, \rho_{i_k}). \tag{5.126}$$

We can establish an ordering relationship between them, generalizing that found in Proposition 4.8 for the classical case. The proof below is similar to previous ones, combining elements of the proofs of Proposition 4.8 and Proposition 5.46:

Proposition 5.50 (Ordering of average k-wise log-Euclidean fidelities)**.** *For $r \in \{3, 4, \ldots\}$, let $\rho_1, \ldots, \rho_r$ be quantum states. Then the following inequalities hold:*

$$F_{2,r}^{\flat} \geq F_{3,r}^{\flat} \geq \cdots \geq F_{r-1,r}^{\flat} \geq F_{r,r}^{\flat}, \tag{5.127}$$

where, for brevity, we have suppressed the dependence of $F_{k,r}^{\flat}$ on $\rho_1, \ldots, \rho_r$.

Proof. See Appendix A.20. □

6 Concluding remarks and future directions

In this work, we extended bivariate fidelities to multivariate fidelities, which are measures characterizing the similarity between multiple quantum states. In the classical case, we analysed the Matusita multivariate fidelity [Mat67], average pairwise fidelity, and average k-wise fidelities. We proposed three variants for the quantum case that reduce to the average pairwise fidelity for commuting states, namely, average pairwise z-fidelities for $z \geq 1/2$, the multivariate SDP fidelity, inspired by the SDP of the Uhlmann fidelity, and the secrecy-based multivariate fidelity inspired by the secrecy measure proposed in [KRS09]. We analysed their mathematical properties, such as data processing, faithfulness, uniform continuity bounds, and connections between these three variants. To this end, we proved that multivariate SDP fidelity and secrecy-based multivariate fidelity are sandwiched between the average pairwise z-fidelities at $z = 1/2$ and $z = 1$. We also explored a quantum extension of Matusita multivariate fidelity, which we called as multivariate log-Euclidean fidelity while analysing its properties and providing an operational interpretation to this in terms of quantum hypothesis testing with an arbitrarily varying null hypothesis.

Open questions for future work include the following (in no particular order):

1. What are similarity measures for a tuple of channels, which generalize multivariate fidelity of states? One approach towards this is to extend the SDP of channel fidelity [KW21, Proposition 55] from two channels to multiple channels, by using the same approach we followed here when defining multivariate SDP fidelity (Definition 5.8). In particular, we can define multivariate channel SDP fidelity for a tuple of channels $(\mathcal{N}_i)_{i=1}^r$ (each one of those is a CPTP map from $\mathscr{L}_A$ to $\mathscr{L}_B$) as follows:

$$\begin{aligned} &F_{\mathrm{SDP}}(\mathcal{N}_1, \ldots, \mathcal{N}_r) \\ &:= \inf_{\rho_{RA} \in \mathscr{D}_{RA}} F_{\mathrm{SDP}}(\mathcal{N}_1(\rho_{RA}), \ldots, \mathcal{N}_r(\rho_{RA})) \qquad (6.1) \\ &= \frac{1}{r(r-1)} \inf_{\substack{Y^i_{RB} \geq 0 \ \forall i \in [r], \\ \rho_R \in \mathscr{D}}} \left\{ \begin{array}{c} \sum_{i=1}^r \operatorname{Tr}\!\left[Y^i_{RB} \Gamma^{\mathcal{N}_i}_{RB}\right] : \\ \sum_{i=1}^r |i\rangle\!\langle i| \otimes Y^i_{RB} \geq \sum_{i \neq j} |i\rangle\!\langle j| \otimes \rho_R \otimes I_B \end{array} \right\}, \qquad (6.2) \end{aligned}$$

   where $\Gamma^{\mathcal{N}_i}_{RB}$ is the Choi operator of channel $\mathcal{N}_i$ for all $i \in [r]$ (i.e., $\Gamma^{\mathcal{N}_i}_{RB} := \sum_{k,\ell} |k\rangle\!\langle \ell| \otimes \mathcal{N}_i(|k\rangle\!\langle \ell|)$). The proof of the equality in the last line follows from the same reasoning used to prove [KW21, Proposition 55].

   Furthermore, similar to average pairwise z-fidelity for states in Definition 5.2, we can define average pairwise channel z-fidelity as follows for $z \geq 1/2$:

$$\begin{aligned} F_z(\mathcal{N}_1, \ldots, \mathcal{N}_r) &:= \inf_{\rho_{RA} \in \mathscr{D}_{RA}} F_z(\mathcal{N}_1(\rho_{RA}), \ldots, \mathcal{N}_r(\rho_{RA})) \qquad (6.3) \\ &= \inf_{\rho_{RA} \in \mathscr{D}_{RA}} \frac{2}{r(r-1)} \sum_{i<j} F_z(\mathcal{N}_i(\rho_{RA}), \mathcal{N}_j(\rho_{RA})). \qquad (6.4) \end{aligned}$$

2. Another interesting future direction is to investigate multivariate geometric fidelities based on average pairwise fidelities built from (5.65) or by generalizing the following

SDP formulation of geometric fidelity from [Mat14, Section 6] and [CS20, Eq. (1.11)]:

$$\underline{\mathbf{F}}(\rho,\sigma) = \sup_{W \in \mathrm{Herm}} \left\{ \mathrm{Tr}[W] : \begin{bmatrix} \rho & W \\ W & \sigma \end{bmatrix} \geq 0 \right\}. \tag{6.5}$$

Indeed a multivariate generalization of the SDP above is as follows:

$$\frac{2}{r(r-1)} \sup_{\substack{(X_{ij})_{i\neq j} \text{ s.t.} \\ X_{ji}=X_{ij}\in\mathrm{Herm}}} \left\{ \sum_{i<j} \mathrm{Tr}[X_{ij}] : \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes \rho_i + \sum_{i\neq j} |i\rangle\!\langle j| \otimes X_{ij} \geq 0 \right\}, \tag{6.6}$$

and one can verify that it satisfies several basic requirements for a multivariate fidelity listed in Definition 5.1, by following our proof approaches for Theorem 5.15.

3. Are there information-theoretic tasks that provide operational interpretations to quantum generalizations of multivariate fidelity, including average pairwise z-fidelities and multivariate SDP fidelity?

4. Is $\sqrt{2(1-F_z(\rho,\sigma))}$ a distance measure for $z \in (1/2,1)\cup(1,\infty)$? Knowing this in turn would provide uniform continuity bounds for average pairwise z-fidelities, as Proposition 5.4 does for the special cases $z \in \{1/2, 1\}$.

5. Another question that we pose is related to multivariate generalizations of the Holevo fidelity. It was shown in [Mat14, page 14] that the Holevo fidelity has the following SDP characterization:

$$F_H(\rho,\sigma) = \sup_{C\geq 0} \left\{ \langle\Gamma|C|\Gamma\rangle : \begin{bmatrix} \rho\otimes I & C \\ C & I\otimes\sigma^T \end{bmatrix} \geq 0 \right\}, \tag{6.7}$$

where $|\Gamma\rangle \coloneqq \sum_i |i\rangle|i\rangle$. This characterization follows because an optimal choice for C is the matrix geometric mean of the commuting operators $\rho\otimes I$ and $I\otimes\sigma^T$ (see, e.g., [Bha07, Theorem 4.1.3]), which is simply $(\rho\otimes\sigma^T)^{1/2}$, and because $\langle\Gamma|(\rho\otimes\sigma^T)^{1/2}|\Gamma\rangle = \mathrm{Tr}[\rho^{1/2}\sigma^{1/2}]$. The dual of the SDP in (6.7) is given by

$$\inf_{\substack{Y,Z\geq 0,\\ W\in\mathscr{L}}} \left\{ \mathrm{Tr}[Y(\rho\otimes I)] + \mathrm{Tr}[Z(I\otimes\sigma^T)] : W + W^\dagger \geq |\Gamma\rangle\!\langle\Gamma|, \begin{bmatrix} Y & W \\ W^\dagger & Z \end{bmatrix} \geq 0 \right\}. \tag{6.8}$$

However, it is unclear how to generalize the constructions in (6.7) and (6.8) to multiple states, given that it makes use of properties specific to bipartite systems. Interestingly, the constructions in (5.27) and (6.6) do not encounter this difficulty.

6. Is there a possibility of finding analytical or alternative variational expressions for minimal and maximal extensions of multivariate classical fidelity in Definition 5.31 and Definition 5.30, respectively?

Acknowledgements

We thank Komal Malik for helpful discussions. TN acknowledges support from the NSF under grant no. 2329662. HKM acknowledges support from the NSF under grant no. 2304816 and AFRL under agreement no. FA8750-23-2-0031. FL acknowledges support from the Department of Mathematics at UIUC for a research visit to Cornell University, as well as the hospitality of Mark Wilde's research group during that stay. MMW acknowledges support from the NSF under grants 2329662, 2315398, 1907615, 2304816.

This material is based on research sponsored by Air Force Research Laboratory under agreement number FA8750-23-2-0031. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory or the U.S. Government.

References

[AD15] Koenraad M. R. Audenaert and Nilanjana Datta. α-z-Rényi relative entropies. *Journal of Mathematical Physics*, 56(2):022202, 2015. arXiv:1310.7178 [quant-ph].

[AFPA07] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2007.

[AKF22] A. Afham, Richard Kueng, and Chris Ferrie. Quantum mean states are nicer than you think: fast algorithms to compute states maximizing average fidelity, 2022. arXiv:2206.08183 [quant-ph].

[ANSV08] Koenraad M. R. Audenaert, Michael Nussbaum, Arleta Szkoła, and Frank Verstraete. Asymptotic error rates in quantum hypothesis testing. *Communications in Mathematical Physics*, 279:251–283, 2008. arXiv:0708.4282 [quant-ph].

[BFT15] Mario Berta, Omar Fawzi, and Marco Tomamichel. On variational expressions for quantum relative entropies. *Letters in Mathematical Physics*, 107(12):2239–2265, December 2015. arXiv:1512.02615 [quant-ph].

[BGJ19] Rajendra Bhatia, Stephane Gaubert, and Tanvi Jain. Matrix versions of the Hellinger distance. *Letters in Mathematical Physics*, 109:1777–1804, 2019. arXiv:1901.01378 [math-ph].

[Bha46] A. Bhattacharyya. On a measure of divergence between two multinomial populations. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 7(4):401–406, July 1946.

[Bha07] Rajendra Bhatia. *Positive Definite Matrices*. Princeton University Press, 2007.

[Bor26] Max Born. Quantum mechanics of collision processes. *Uspekhi Fizich*, June 1926.

[BS82] V. P. Belavkin and P. Staszewski. C*-algebraic generalization of relative entropy and entropy. *Annales de l'I.H.P. Physique théorique*, 37:51–58, 1982.

[Bur69] Donald Bures. An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite ω^*-algebras. *Transactions of the American Mathematical Society*, 135(0):199–212, 1969.

[CS20] Sam Cree and Jamie Sikora. A fidelity measure for quantum states based on the matrix geometric mean, 2020. arXiv:2006.06918 [quant-ph].

[DL14] Nilanjana Datta and Felix Leditzky. A limit of the quantum Rényi divergence. *Journal of Physics A*, 47(4):045304, 2014. arXiv:1308.5961 [quant-ph].

[Don86] Matthew J. Donald. On the relative entropy. *Communications in Mathematical Physics*, 105:13–34, 1986.

[FC95] Christopher A. Fuchs and Carlton M. Caves. Mathematical techniques for quantum communication theory. *Open Systems & Information Dynamics*, 3(3):345–356, 1995. arXiv:quant-ph/9604001.

[FFHT23] Muhammad Usman Farooq, Tobias Fritz, Erkka Haapasalo, and Marco Tomamichel. Asymptotic and catalytic matrix majorization, 2023. arXiv:2301.07353 [math.ST].

[Fuc96] Christopher Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. PhD thesis, University of New Mexico, December 1996. arXiv:quant-ph/9601020.

[GT20] Gilad Gour and Marco Tomamichel. Optimal extensions of resource measures and their applications. *Physical Review A*, 102(6):062401, December 2020. arXiv:2006.12408 [quant-ph].

[GW15] Manish K. Gupta and Mark M. Wilde. Multiplicativity of completely bounded p-norms implies a strong converse for entanglement-assisted capacity. *Communications in Mathematical Physics*, 334:867–887, 2015. arXiv:1310.7028 [quant-ph].

[Hel09] Ernst Hellinger. Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen. *Journal für die reine und angewandte Mathematik*, 1909(136):210–271, July 1909.

[Hel67] Carl W. Helstrom. Minimum mean-squared error of estimates in quantum statistics. *Physics Letters A*, 25(2):101–102, July 1967.

[Hia19] Fumio Hiai. Quantum f-divergences in von Neumann algebras. II. Maximal f-divergences. *Journal of Mathematical Physics*, 60(1):012203, January 2019. arXiv:1807.03118 [math-ph].

[HM17] Fumio Hiai and Milán Mosonyi. Different quantum f-divergences and the reversibility of quantum operations. *Reviews in Mathematical Physics*, 29(07):1750023, 2017. arXiv:1604.03089 [math-ph].

[Hol72] Alexander S. Holevo. On quasiequivalence of locally normal states. *Theoretical and Mathematical Physics*, 13(2):1071–1082, November 1972.

[HP91] Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143:99–114, 1991.

[Jen04] Anna Jencova. Geodesic distances on density matrices. *Journal of Mathematical Physics*, 45(5):1787–1794, May 2004. arXiv:math-ph/0312044.

[Joz94] Richard Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41(12):2315–2323, 1994.

[KL51] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, March 1951.

[KRS09] Robert König, Renato Renner, and Christian Schaffner. The operational meaning of min-and max-entropy. *IEEE Transactions on Information Theory*, 55(9):4337–4347, 2009. arXiv:0807.1338 [quant-ph].

[KW20] Sumeet Khatri and Mark M. Wilde. Principles of quantum communication theory: A modern approach, 2020. arXiv:2011.04672v2 [quant-ph].

[KW21] Vishal Katariya and Mark M. Wilde. Geometric distinguishability measures limit quantum channel estimation and discrimination. *Quantum Information Processing*, 20(2):78, 2021. arXiv:2004.10708 [quant-ph].

[LKDW18] Felix Leditzky, Eneet Kaur, Nilanjana Datta, and Mark M. Wilde. Approaches for approximate additivity of the Holevo information of quantum channels. *Physical Review A*, 97(1):012332, January 2018. arXiv:1709.01111 [quant-ph].

[LT15] Simon M. Lin and Marco Tomamichel. Investigating properties of a family of quantum Rényi divergences. *Quantum Information Processing*, 14(4):1501–1512, 2015. arXiv:1408.6897 [quant-ph].

[LZ04] Shunlong Luo and Qiang Zhang. Informational distance on quantum-state space. *Physical Review A*, 69(3):032106, March 2004.

[Mat67] Kameo Matusita. On the notion of affinity of several distributions and some of its applications. *Annals of the Institute of Statistical Mathematics*, 19:181–192, December 1967.

[Mat10a] Keiji Matsumoto. Reverse test and characterization of quantum relative entropy, 2010. arXiv:1010.1030 [quant-ph].

[Mat10b] Keiji Matsumoto. Reverse test and quantum analogue of classical fidelity and generalized fidelity, 2010. arXiv:1006.0302 [quant-ph].

[Mat14] Keiji Matsumoto. Quantum fidelities, their duals, and convex analysis, 2014. arXiv:1408.3462 [quant-ph].

[Mat18] Keiji Matsumoto. A new quantum version of f-divergence. In *Reality and Measurement in Algebraic Quantum Theory*, volume 261, pages 229–273. Springer Singapore, 2018.

[MBV22] Milán Mosonyi, Gergely Bunth, and Péter Vrana. Geometric relative entropies and barycentric Rényi divergences, 2022. arXiv:2207.14282 [quant-ph].

[MH22] Milán Mosonyi and Fumio Hiai. Test-measured Rényi divergences. *IEEE Transactions on Information Theory*, 69(2):1074–1092, 2022. arXiv:2201.05477 [quant-ph].

[MLDS+13] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, 2013. arXiv:1306.3142 [quant-ph].

[MNW23] Hemant K. Mishra, Michael Nussbaum, and Mark M. Wilde. On the optimal error exponents for classical and quantum antidistinguishability, 2023. arXiv:2309.03723v2 [quant-ph].

[MO17] Milán Mosonyi and Tomohiro Ogawa. Strong converse exponent for classical-quantum channel coding. *Communications in Mathematical Physics*, 355:373–426, 2017. arXiv:1409.3562 [quant-ph].

[NR18] Rajai Nasser and Joseph M. Renes. Polar codes for arbitrary classical-quantum channels and arbitrary cq-MACs. *IEEE Transactions on Information Theory*, 64(11):7424–7442, 2018. arXiv:1701.03397 [cs.IT].

[Nö14] Janis Nötzel. Hypothesis testing on invariant subspaces of the symmetric group: Part I. Quantum Sanov's theorem and arbitrarily varying sources. *Journal of Physics A: Mathematical and Theoretical*, 47(23):235303, May 2014. arXiv:1310.5553 [quant-ph].

[Pet85] Dénes Petz. Quasi-entropies for states of a von Neumann algebra. *Publications of the Research Institute for Mathematical Sciences*, 21(4):787–800, 1985.

[Pet86] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports on Mathematical Physics*, 23:57–65, 1986.

[Pia09] Marco Piani. Relative entropy of entanglement and restricted measurements. *Physical Review Letters*, 103(16):160504, October 2009. arXiv:0904.2705 [quant-ph].

[PV08] Daniel P. Palomar and Sergio Verdu. Lautum information. *IEEE Transactions on Information Theory*, 54(3):964–975, 2008.

[RASW23] Soorya Rethinasamy, Rochisha Agarwal, Kunal Sharma, and Mark M. Wilde. Estimating distinguishability measures on quantum computers. *Physical Review A*, 108(1):012409, July 2023. arXiv:2108.08406 [quant-ph].

[Rén61] Alfréd Rényi. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, volume 4, pages 547–562. University of California Press, January 1961.

[ŞTA09] Eren Şaşoğlu, Emre Telatar, and Erdal Arikan. Polarization for arbitrary discrete memoryless channels. In *IEEE Information Theory Workshop*, pages 144–148, 2009. arXiv:0908.0302 [cs.IT].

[Str65] R. L. Stratonovich. Information capacity of a quantum communications channel. I. *Soviet Radiophysics*, 8(1):82–91, January 1965.

[SW12] Naresh Sharma and Naqueeb A. Warsi. On the strong converses for the quantum channel capacity theorems, May 2012. arXiv:1205.1712 [quant-ph].

[Tou74] Godfried T. Toussaint. Some properties of Matusita's measure of affinity of several distributions. *Annals of the Institute of Statistical Mathematics*, 26(3):389–394, 1974.

[Uhl76] Armin Uhlmann. The "transition probability" in the state space of a $*$-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[Ume62] Hisaharu Umegaki. Conditional expectation in an operator algebra, iv (entropy and information). In *Kodai Mathematical Seminar Reports*, volume 14, pages 59–85. Department of Mathematics, Tokyo Institute of Technology, 1962.

[Wat09] John Watrous. Semidefinite programs for completely bounded norms. *Theory of Computing*, 5(11):217–238, 2009. arXiv:0901.4709 [quant-ph].

[WWY14] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331:593–622, 2014. arXiv:1306.1586 [quant-ph].

[Zha20] Haonan Zhang. From Wigner-Yanase-Dyson conjecture to Carlen-Frank-Lieb conjecture. *Advances in Mathematics*, 365:107053, 2020. arXiv:1811.01205 [math.FA].

A Proofs

In the following appendices, we present technical proofs of the results that are stated in the main text of the paper.

A.1 Proof of Equation (1.8) (Log-Euclidean fidelity as limit of z-fidelity)

For $\varepsilon > 0$, define $\rho(\varepsilon) \coloneqq \rho + \varepsilon I$ and $\sigma(\varepsilon) \coloneqq \sigma + \varepsilon I$. Observe that

$$F_z(\rho, \sigma) = \lim_{\varepsilon \to 0^+} F_z(\rho(\varepsilon), \sigma(\varepsilon)) = \inf_{\varepsilon > 0} F_z(\rho(\varepsilon), \sigma(\varepsilon)), \tag{A.1}$$

which follows from the operator monotonicity of the function x^p for $p \in (0, 1]$ and monotonicity of $(\cdot) \to \operatorname{Tr}[(\cdot)^z]$ for all $z \geq 1/2$. Now recalling that the z-fidelities are antimonotone for $z \geq 1/2$ [LT15, Proposition 6] and continuous in z, we conclude that

$$\begin{aligned}
\lim_{z \to \infty} F_z(\rho, \sigma) &= \inf_{z \geq \frac{1}{2}} F_z(\rho, \sigma) & \text{(A.2)}\\
&= \inf_{z \geq \frac{1}{2}} \inf_{\varepsilon > 0} F_z(\rho(\varepsilon), \sigma(\varepsilon)) & \text{(A.3)}\\
&= \inf_{\varepsilon > 0} \inf_{z \geq \frac{1}{2}} F_z(\rho(\varepsilon), \sigma(\varepsilon)) & \text{(A.4)}\\
&= \inf_{\varepsilon > 0} \lim_{z \to \infty} F_z(\rho(\varepsilon), \sigma(\varepsilon)) & \text{(A.5)}\\
&= \inf_{\varepsilon > 0} \operatorname{Tr}\left[\exp\left(\frac{1}{2}(\ln \rho(\varepsilon) + \ln \sigma(\varepsilon))\right)\right]. & \text{(A.6)}
\end{aligned}$$

The last equality follows from an application of the Lie–Trotter product formula. Similar observations were made in [AD15, Section 4] using the fact that for operators A and B, $\lim_{z \to \infty} \exp(A/z) \exp(B/z) = \exp(A + B)$ with the choice $A = \frac{1}{2} \ln \rho(\varepsilon)$ and $B = \frac{1}{2} \ln \sigma(\varepsilon)$.

The last equality in (1.8) follows from operator monotonicity of the logarithm function and monotonicity of $(\cdot) \to \operatorname{Tr}[\exp(\cdot)]$.

A.2 Proof of Proposition 4.3 (Uniform continuity of Matusita fidelity)

Consider that

$$
\begin{aligned}
&|F_r(p_1,\ldots,p_r)-F_r(q_1,\ldots,q_r)| \\
&=\left|\sum_x (p_1(x)\cdots p_r(x))^{\frac{1}{r}}-\sum_x (q_1(x)\cdots q_r(x))^{\frac{1}{r}}\right| && \text{(A.7)}\\
&=\left|\sum_x \left[(p_1(x)\cdots p_r(x))^{\frac{1}{r}}-(q_1(x)\cdots q_r(x))^{\frac{1}{r}}\right]\right| && \text{(A.8)}\\
&=\left|\sum_{i=1}^r\sum_x \left[(p_1(x)\cdots p_{i-1}(x))^{\frac{1}{r}}\left(p_i(x)^{\frac{1}{r}}-q_i(x)^{\frac{1}{r}}\right)(q_{i+1}(x)\cdots q_r(x))^{\frac{1}{r}}\right]\right| && \text{(A.9)}\\
&\leq\sum_{i=1}^r\left|\sum_x \left[(p_1(x)\cdots p_{i-1}(x))^{\frac{1}{r}}\left(p_i(x)^{\frac{1}{r}}-q_i(x)^{\frac{1}{r}}\right)(q_{i+1}(x)\cdots q_r(x))^{\frac{1}{r}}\right]\right| && \text{(A.10)}\\
&\leq\sum_{i=1}^r\left\|(p_1\cdots p_{i-1}q_{i+1}\cdots q_r)^{\frac{1}{r}}\right\|_{\frac{r}{r-1}}\left\|p_i^{\frac{1}{r}}-q_i^{\frac{1}{r}}\right\|_r && \text{(A.11)}\\
&=\sum_{i=1}^r F_{r-1}(p_1,\ldots,p_{i-1},q_{i+1},\ldots,q_r)^{\frac{r-1}{r}}\left\|p_i^{\frac{1}{r}}-q_i^{\frac{1}{r}}\right\|_r && \text{(A.12)}\\
&\leq\sum_{i=1}^r\left\|p_i^{\frac{1}{r}}-q_i^{\frac{1}{r}}\right\|_r && \text{(A.13)}\\
&=r\left(\frac{1}{r}\sum_{i=1}^r\left(\left\|p_i^{\frac{1}{r}}-q_i^{\frac{1}{r}}\right\|_r^r\right)^{\frac{1}{r}}\right) && \text{(A.14)}\\
&\leq r\left(\frac{1}{r}\sum_{i=1}^r\left\|p_i^{\frac{1}{r}}-q_i^{\frac{1}{r}}\right\|_r^r\right)^{\frac{1}{r}} && \text{(A.15)}\\
&\leq r\left(\frac{1}{r}\sum_{i=1}^r\|\sqrt{p_i}-\sqrt{q_i}\|_2^2\right)^{\frac{1}{r}} && \text{(A.16)}\\
&\leq r\,(\varepsilon)^{\frac{1}{r}}. && \text{(A.17)}
\end{aligned}
$$

The third equality follows from rewriting as a telescoping sum. The first inequality follows from the triangle inequality. The second inequality follows from Hölder's inequality with parameters $r,s\geq 1$ satisfying $\frac{1}{r}+\frac{1}{s}=1$, so that $s=\frac{r}{r-1}$. The fourth equality follows from applying definitions, and the third inequality follows because $F_r\leq 1$ for all probability distributions and for every integer $r\geq 2$. The fourth inequality follows from the concavity of function $x^{1/r}$ for $r\geq 2$ for all $x\geq 0$. To see the penultimate inequality, consider that the following inequality holds for all distributions p and q by using [Mat67, Theorem 5]:

$$\left\| p^{\frac{1}{r}} - q^{\frac{1}{r}} \right\|_r^r = \sum_x \left| p(x)^{\frac{1}{r}} - q(x)^{\frac{1}{r}} \right|^r \leq \sum_x \left(\sqrt{p(x)} - \sqrt{q(x)} \right)^2 = \|\sqrt{p} - \sqrt{q}\|_2^2 . \tag{A.18}$$

Finally, the last inequality follows from the the assumption $\frac{1}{r}\left(\sum_{i=1}^r \left[d_H(p_i, q_i)\right]^2\right) \leq \varepsilon$, along with (4.10).

A.3 Proof of Proposition 4.8 (Inequalities relating classical average k-wise fidelities)

It suffices to prove the following inequality:

$$F_{k-1,r}(\rho_1, \ldots, \rho_r) \geq F_{k,r}(\rho_1, \ldots, \rho_r), \tag{A.19}$$

for all $k \in \{3, \ldots, r\}$. So we prove this one, using a generalization of the approach used in the proof of Proposition 4.6. Let $S_{k-1}(i_1, \ldots, i_k)$ denote all size-$(k-1)$ subsets of $i_1, \ldots, i_k$, of which there are k of them. For example, if $r = 7$, $k = 4$, and $i_1 = 2$, $i_2 = 4$, $i_3 = 5$, $i_4 = 7$, then

$$S_{k-1}(i_1, i_2, i_3, i_4) = S_3(2, 4, 5, 7) = \{\{2, 4, 5\}, \{2, 4, 7\}, \{2, 5, 7\}, \{4, 5, 7\}\} . \tag{A.20}$$

Note that each symbol i_j appears $k-1$ times in each subset of $S_{k-1}(i_1, \ldots, i_k)$. Then consider that

$$\begin{aligned}
F_k(\rho_{i_1}, \ldots, \rho_{i_k}) &= \sum_x \left(\rho_{i_1}(x) \cdots \rho_{i_k}(x)\right)^{\frac{1}{k}} &\text{(A.21)}\\
&= \sum_x \left(\prod_{t \in S_{k-1}(i_1, \ldots, i_k)} \left[\rho_{t(1)}(x) \cdots \rho_{t(k-1)}(x)\right]^{\frac{1}{k-1}} \right)^{\frac{1}{k}} &\text{(A.22)}\\
&\leq \sum_x \frac{1}{k} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} \left[\rho_{t(1)}(x) \cdots \rho_{t(k-1)}(x)\right]^{\frac{1}{k-1}} &\text{(A.23)}\\
&= \frac{1}{k} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} \sum_x \left[\rho_{t(1)}(x) \cdots \rho_{t(k-1)}(x)\right]^{\frac{1}{k-1}} &\text{(A.24)}\\
&= \frac{1}{k} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} F_{k-1}(\rho_{t(1)}, \ldots, \rho_{t(k-1)}), &\text{(A.25)}
\end{aligned}$$

where the inequality follows as a consequence of the inequality of arithmetic and geometric means. Now consider that

$$F_{k,r}(\rho_1, \ldots, \rho_r) = \frac{1}{\binom{r}{k}} \sum_{i_1 < \cdots < i_k} F_k(\rho_{i_1}, \ldots, \rho_{i_k})$$

$$\leq \frac{1}{\binom{r}{k}} \sum_{i_1<\cdots<i_k} \frac{1}{k} \sum_{t \in S_{k-1}(i_1,\ldots,i_k)} F_{k-1}(\rho_{t(1)}, \ldots, \rho_{t(k-1)}) \tag{A.26}$$

$$= \frac{k-1!r-k!}{r!} \sum_{i_1<\cdots<i_k} \sum_{t \in S_{k-1}(i_1,\ldots,i_k)} F_{k-1}(\rho_{t(1)}, \ldots, \rho_{t(k-1)}) \tag{A.27}$$

$$= \frac{k-1!r-k!}{r!} (r-k+1) \sum_{i_1<\cdots<i_{k-1}} F_{k-1}(\rho_{i_1}, \ldots, \rho_{i_{k-1}}) \tag{A.28}$$

$$= \frac{1}{\binom{r}{k-1}} \sum_{i_1<\cdots<i_{k-1}} F_{k-1}(\rho_{i_1}, \ldots, \rho_{i_{k-1}}) \tag{A.29}$$

$$= F_{k-1,r}(\rho_1, \ldots, \rho_r). \tag{A.30}$$

The second equality follows because the sum in the third line includes each distinct term $r-k+1$ times. Indeed, the sum in the third line arises by picking size-k subsets from $[r]$, and from each of these size-k subsets, picking size-$(k-1)$ subsets, thus leading to a total number of subsets given by $\binom{r}{k}\binom{k}{k-1}$. However, picking subsets in this way leads to $\binom{r}{k-1}$ distinct subsets, each occurring with multiplicity $\frac{\binom{r}{k}\binom{k}{k-1}}{\binom{r}{k-1}} = r-k+1$.

A.4 Proof of Proposition 5.4 (Uniform continuity of average pairwise fidelities)

Define the shorthand $\mathcal{S}_\rho \coloneqq (\rho_1, \ldots, \rho_r)$ and $\mathcal{S}_\sigma \coloneqq (\sigma_1, \ldots, \sigma_r)$. Consider that

$$F_U(\mathcal{S}_\rho) - F_U(\mathcal{S}_\sigma) = \frac{2}{r(r-1)} \sum_{i<j} F(\rho_i, \rho_j) - \frac{2}{r(r-1)} \sum_{i<j} F(\sigma_i, \sigma_j) \tag{A.31}$$

$$= \frac{2}{r(r-1)} \sum_{i<j} \left[F(\rho_i, \rho_j) - F(\sigma_i, \sigma_j)\right]. \tag{A.32}$$

We have

$$\begin{aligned}
&2\left[F(\rho_i, \rho_j) - F(\sigma_i, \sigma_j)\right] \\
&= 2\left[1 - F(\sigma_i, \sigma_j) - (1 - F(\rho_i, \rho_j))\right] && \text{(A.33)} \\
&= 2\left[\sqrt{1 - F(\sigma_i, \sigma_j)}\sqrt{1 - F(\sigma_i, \sigma_j)} - \sqrt{1 - F(\rho_i, \rho_j)}\sqrt{1 - F(\rho_i, \rho_j)}\right] && \text{(A.34)} \\
&= d_B(\sigma_i, \sigma_j) d_B(\sigma_i, \sigma_j) - d_B(\rho_i, \rho_j) d_B(\rho_i, \rho_j) && \text{(A.35)} \\
&= d_B(\sigma_i, \sigma_j) d_B(\sigma_i, \sigma_j) - d_B(\sigma_i, \sigma_j) d_B(\rho_i, \rho_j) \\
&\qquad + d_B(\sigma_i, \sigma_j) d_B(\rho_i, \rho_j) - d_B(\rho_i, \rho_j) d_B(\rho_i, \rho_j) && \text{(A.36)} \\
&= \left[d_B(\sigma_i, \sigma_j) + d_B(\rho_i, \rho_j)\right]\left[d_B(\sigma_i, \sigma_j) - d_B(\rho_i, \rho_j)\right] && \text{(A.37)} \\
&\leq \left[d_B(\sigma_i, \sigma_j) + d_B(\rho_i, \rho_j)\right] \times \\
&\qquad \left[d_B(\sigma_i, \rho_i) + d_B(\rho_i, \rho_j) + d_B(\rho_j, \sigma_j) - d_B(\rho_i, \rho_j)\right] && \text{(A.38)}
\end{aligned}$$

$$= [d_B(\sigma_i,\sigma_j) + d_B(\rho_i,\rho_j)]\,[d_B(\sigma_i,\rho_i) + d_B(\rho_j,\sigma_j)] \tag{A.39}$$

$$\leq 2\sqrt{2}\,[d_B(\sigma_i,\rho_i) + d_B(\rho_j,\sigma_j)] \tag{A.40}$$

$$= 2\sqrt{2}\,[d_B(\rho_i,\sigma_i) + d_B(\rho_j,\sigma_j)]\,, \tag{A.41}$$

where the first inequality follows from the triangular inequality of the Bures distance and the second inequality because $d_B(\omega,\tau) \leq \sqrt{2}$ for all quantum states ω and τ. This implies that

$$F(\rho_i,\rho_j) - F(\sigma_i,\sigma_j) \leq \sqrt{2}\,[d_B(\rho_i,\sigma_i) + d_B(\rho_j,\sigma_j)]\,. \tag{A.42}$$

Substituting (A.42) into (A.32) and using the fact that

$$\sum_{i<j} [d_B(\rho_i,\sigma_i) + d_B(\rho_j,\sigma_j)] = (r-1)\sum_{i=1}^{r} d_B(\rho_i,\sigma_i) \tag{A.43}$$

gives

$$F_U(\mathcal{S}_\rho) - F_U(\mathcal{S}_\sigma) \leq \sqrt{2}\left(\frac{2}{r\,(r-1)}\right)\sum_{i<j}[d_B(\rho_i,\sigma_i) + d_B(\rho_j,\sigma_j)] \tag{A.44}$$

$$= 2\sqrt{2}\left(\frac{1}{r}\sum_{i=1}^{r} d_B(\rho_i,\sigma_i)\right) \tag{A.45}$$

$$\leq 2\sqrt{2}\varepsilon, \tag{A.46}$$

where the last inequality follows from the assumption that $\frac{1}{r}\sum_{i=1}^{r} d_B(\rho_i,\sigma_i) \leq \varepsilon$. This concludes the proof of the inequality in (5.16).

A similar proof works for the inequality in (5.17) by replacing all instances of Uhlmann fidelity and Bures distance with Holevo fidelity and quantum Hellinger distance, respectively.

A.5 Proof of Proposition 5.10 (Dual of multivariate SDP fidelity)

Recall that a standard primal SDP and its dual are given as follows (cf. [KW20, Definition 2.20]):

$$\mathtt{PRIMAL} \coloneqq \inf_{Y\geq 0}\left\{\mathrm{Tr}[BY] : \Phi^\dagger(Y) \geq A\right\}, \tag{A.47}$$

$$\mathtt{DUAL} \coloneqq \sup_{X\geq 0}\left\{\mathrm{Tr}[AX] : \Phi(X) \leq B\right\}, \tag{A.48}$$

where A and B are Hermitian matrices and Φ is a Hermiticity-preserving superoperator. It is always true that $\mathtt{PRIMAL} \geq \mathtt{DUAL}$, and strong duality is said to hold in the case of equality. A sufficient condition for strong duality to hold is that there exists a feasible point for the $\mathtt{DUAL}$ and a *strictly feasible* point for the $\mathtt{PRIMAL}$; i.e., the latter meaning that there exists a Y such that $Y > 0$ and $\Phi^\dagger(Y) > A$. This is known as Slater's condition. In what follows, we identify the A, B, and $\Phi^\dagger$ from the multivariate SDP formulation (5.26) to derive its dual, and we then argue that strong duality holds using Slater's theorem.

Comparing the primal SDP (A.47) to the multivariate SDP fidelity (5.26) without considering the normalization term $r(r-1)$, we make the following choices of A, B, and $\Phi^\dagger$: given $Y = \sum_{i,j=1}^{r} |i\rangle\!\langle j| \otimes Y_{ij}$, define

$$A \coloneqq \sum_{i,j\in[r]:i\neq j} |i\rangle\!\langle j| \otimes I, \tag{A.49}$$

$$B \coloneqq \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes \rho_i, \tag{A.50}$$

$$\Phi^\dagger(Y) \coloneqq \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes Y_{ii}. \tag{A.51}$$

Thus, given any Hermitian matrix $X = \sum_{i,j=1}^{r} |i\rangle\!\langle j| \otimes X_{ij}$ and using the relation $\mathrm{Tr}[Y\Phi(X)] = \mathrm{Tr}\big[\Phi^\dagger(Y)X\big]$, we obtain

$$\Phi(X) = \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes X_{ii}. \tag{A.52}$$

Also, we have

$$\mathrm{Tr}[AX] = \sum_{i\neq j} \mathrm{Tr}[X_{ij}] = 2\sum_{i<j} \Re\left[\mathrm{Tr}[X_{ij}]\right]. \tag{A.53}$$

The `DUAL` condition $\Phi(X) \leq B$ is given by

$$\sum_{i=1}^{r} |i\rangle\!\langle i| \otimes X_{ii} \leq \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes \rho_i, \tag{A.54}$$

which, by adding $\sum_{i\neq j} |i\rangle\!\langle i| \otimes X_{ij}$ to both sides, evaluates to

$$0 \leq X \leq \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes \rho_i + \sum_{i\neq j} |i\rangle\!\langle i| \otimes X_{ij}. \tag{A.55}$$

Using the relations (A.53) and (A.55) in the `DUAL` form and the fact that X is Hermitian gives the desired dual formulation (5.27).

We use Slater's condition to show that strong duality holds, which requires finding a strictly feasible point in the SDP formulation (5.26) and a feasible point in the dual formulation (5.27). The positive definite matrix $Y = \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes rI$ is a strictly feasible point because $\Phi^\dagger(Y) - A > 0$. To see this, consider that

$$\Phi^\dagger(Y) - A = \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes rI - \sum_{i\neq j} |i\rangle\!\langle j| \otimes I \tag{A.56}$$

$$= \sum_{i=1}^{r} (r+1)|i\rangle\!\langle i| \otimes I - \sum_{i,j=1}^{r} |i\rangle\!\langle j| \otimes I \tag{A.57}$$

$$= \left((r+1)I - uu^T\right) \otimes I, \tag{A.58}$$

where u is the r-column vector of all ones. The matrix $(r+1)I - uu^T$ is positive definite because its (distinct) eigenvalues are 1 and $r+1$. We thus conclude that $\Phi^\dagger(Y) - A > 0$. A feasible point for (5.27) consists of simply picking $X_{ij} = 0$ for all i, j.

A.6 Proof of Theorem 5.12 (Uniform continuity of $K^\star$-representation)

Recall by definition that

$$F_{K^\star}(\rho_1, \ldots, \rho_r) = \frac{1}{r-1} \sup_{K \geq 0} \left\{ \begin{array}{c} \langle \psi^\rho | K \otimes I | \psi^\rho \rangle - 1 : \\ K = I_r \otimes I_d + \sum_{i \neq j} |i\rangle\langle j| \otimes K_{ij} \geq 0, \end{array} \right\}, \tag{A.59}$$

where $|\psi^\rho\rangle = \frac{1}{\sqrt{r}} \sum_{i=1}^r |i\rangle|\phi^{\rho_i}\rangle$ and $|\phi^{\rho_i}\rangle$ is an arbitrary purification of ρ_i. Fix

$$K = I_r \otimes I_d + \sum_{i \neq j} |i\rangle\langle j| \otimes K_{ij} \geq 0. \tag{A.60}$$

Also fix

$$|\psi^\rho\rangle = \frac{1}{\sqrt{r}} \sum_{i=1}^r |i\rangle|\phi^{\rho_i}\rangle, \tag{A.61}$$

$$|\psi^\sigma\rangle = \frac{1}{\sqrt{r}} \sum_{i=1}^r |i\rangle|\phi^{\sigma_i}\rangle, \tag{A.62}$$

where these purifications are those that achieve the (root) fidelity, so that $\langle \phi^{\sigma_i} | \phi^{\rho_i} \rangle = F(\rho_i, \sigma_i)$. Hölder's inequality implies that

$$\begin{aligned}
&|\langle \psi^\rho | K \otimes I | \psi^\rho \rangle - \langle \psi^\sigma | K \otimes I | \psi^\sigma \rangle| \\
&= |\mathrm{Tr}[(|\psi^\rho\rangle\langle\psi^\rho| - |\psi^\sigma\rangle\langle\psi^\sigma|)\, K \otimes I]| && \text{(A.63)} \\
&\leq \||\psi^\rho\rangle\langle\psi^\rho| - |\psi^\sigma\rangle\langle\psi^\sigma|\|_1 \, \|K \otimes I\|_\infty && \text{(A.64)} \\
&= \||\psi^\rho\rangle\langle\psi^\rho| - |\psi^\sigma\rangle\langle\psi^\sigma|\|_1 \, \|K\|_\infty && \text{(A.65)} \\
&\leq \||\psi^\rho\rangle\langle\psi^\rho| - |\psi^\sigma\rangle\langle\psi^\sigma|\|_1 \, \|r\,(I_r \otimes I_d)\|_\infty && \text{(A.66)} \\
&= r\, \||\psi^\rho\rangle\langle\psi^\rho| - |\psi^\sigma\rangle\langle\psi^\sigma|\|_1 && \text{(A.67)} \\
&= r\sqrt{1 - |\langle \psi^\sigma | \psi^\rho \rangle|^2} && \text{(A.68)} \\
&= r\sqrt{1 - \left| \frac{1}{r} \sum_{i=1}^r \langle \phi^{\sigma_i} | \phi^{\rho_i} \rangle \right|^2} && \text{(A.69)} \\
&= r\sqrt{1 - \left(\frac{1}{r} \sum_{i=1}^r F(\rho_i, \sigma_i) \right)^2} && \text{(A.70)}
\end{aligned}$$

$$\leq r\sqrt{1-(1-\varepsilon)^2} \tag{A.71}$$
$$= r\sqrt{\varepsilon\,(2-\varepsilon)}. \tag{A.72}$$

The second to last inequality follows because

$$(I_r \otimes I_d) = \sum_x \left(|x\rangle\!\langle x| \otimes I\right) K \left(|x\rangle\!\langle x| \otimes I\right) \tag{A.73}$$
$$= \frac{1}{r}\sum_{k=0}^{r-1} \left(Z(k) \otimes I\right) K \left(Z(k) \otimes I\right)^\dagger \tag{A.74}$$
$$\geq \frac{1}{r}K, \tag{A.75}$$

where $Z(k) := \sum_{\ell=0}^{r-1} e^{2\pi i \ell k/r}|\ell\rangle\!\langle\ell|$ is the Heisenberg phase shift operator. The last inequality follows from antimonotonicity of the function $\sqrt{1-x^2}$ on the interval $x \in [0,1]$ and the fact that $\frac{1}{r}\sum_{i=1}^{r} F(\rho_i,\sigma_i) \geq 1-\varepsilon$ by assumption (see (5.32)). From (A.72) we thus get

$$\frac{1}{r-1}\left(\langle\psi^\rho|K \otimes I|\psi^\rho\rangle - 1\right)$$
$$\leq \frac{1}{r-1}\left(\langle\psi^\sigma|K \otimes I|\psi^\sigma\rangle - 1\right) + \frac{r}{r-1}\sqrt{\varepsilon\,(2-\varepsilon)} \tag{A.76}$$
$$\leq F_{K^\star}(\sigma_1,\ldots,\sigma_r) + \frac{r}{r-1}\sqrt{\varepsilon\,(2-\varepsilon)}, \tag{A.77}$$

where the last inequality follows from the expression in (A.59). Since the inequality holds for all K satisfying the condition, we conclude that

$$F_{K^\star}(\rho_1,\ldots,\rho_r) \leq F_{K^\star}(\sigma_1,\ldots,\sigma_r) + \frac{r}{r-1}\sqrt{\varepsilon\,(2-\varepsilon)}, \tag{A.78}$$

after applying (A.59) once again. The other inequality follows from the same reasoning, but instead swapping $\rho_1,\ldots,\rho_r$ and $\sigma_1,\ldots,\sigma_r$.

A.7 Monotonicity of multivariate SDP fidelity

For every tuple of states $\rho_1,\ldots,\rho_r$ (not necessarily invertible) and for all $\varepsilon > 0$, all states of the following form are invertible:

$$\rho_i^{(\varepsilon)} := \frac{1}{1+\varepsilon}\left(\rho_i + \frac{\varepsilon}{d}I\right). \tag{A.79}$$

As stated below, the multivariate SDP fidelity $F_{\mathrm{SDP}}(\rho_1^{(\varepsilon)},\ldots,\rho_r^{(\varepsilon)})$ is monotonic as a function of ε, up to a multiplicative factor.

Lemma 1 (Monotonicity)**.** *Let $\rho_1,\ldots,\rho_r$ be quantum states, and fix ε_1 and ε_2 such that $0 \leq \varepsilon_1 \leq \varepsilon_2$. Then*

$$(1+\varepsilon_1)F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon_1)},\ldots,\rho_r^{(\varepsilon_1)}\right) \leq (1+\varepsilon_2)F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon_2)},\ldots,\rho_r^{(\varepsilon_2)}\right). \tag{A.80}$$

Proof. Let $(Y_i)_{i=1}^r$ be a candidate tuple for the SDP for $F_{\text{SDP}}\left(\rho_1^{(\varepsilon_2)}, \ldots, \rho_r^{(\varepsilon_2)}\right)$. Then

$$\frac{1}{r(r-1)} \sum_{i=1}^{n} \operatorname{Tr}\left[Y_i \rho_i^{(\varepsilon_2)}\right] = \frac{1}{r(r-1)(1+\varepsilon_2)} \sum_{i=1}^{n} \operatorname{Tr}[Y_i(\rho_i + \varepsilon_2 I/d)] \tag{A.81}$$

$$\geq \frac{1}{r(r-1)(1+\varepsilon_2)} \sum_{i=1}^{n} \operatorname{Tr}[Y_i(\rho_i + \varepsilon_1 I/d)] \tag{A.82}$$

$$= \frac{(1+\varepsilon_1)}{r(r-1)(1+\varepsilon_2)} \sum_{i=1}^{n} \operatorname{Tr}\left[Y_i \rho_i^{(\varepsilon_1)}\right] \tag{A.83}$$

$$\geq \frac{(1+\varepsilon_1)}{(1+\varepsilon_2)} F\left(\rho_1^{(\varepsilon_1)}, \ldots, \rho_r^{(\varepsilon_1)}\right). \tag{A.84}$$

The first inequality holds because $\varepsilon_1 \leq \varepsilon_2$, and the last inequality follows from Definition 5.8. By optimizing the left-hand side of the above over all $(Y_i)_{i=1}^r$ satisfying the constraints in the SDP for $F_{\text{SDP}}\left(\rho_1^{(\varepsilon_2)}, \ldots, \rho_r^{(\varepsilon_2)}\right)$, and then rearranging the terms, we arrive at the desired inequality (A.80). □

A.8 Proof of Proposition 5.13 (Alternative formulation of multivariate SDP fidelity)

The proof of Theorem 5.13 is structured in four steps as follows: In the first step we show an equivalent formulation for the $K^\star$-representation in (5.28). This helps us in the second step to prove that $F_{K^\star}(\rho_1, \ldots, \rho_r) \leq F_{\text{SDP}}(\rho_1, \ldots, \rho_r)$ for general states. In the third step, we show that $F_{K^\star}(\rho_1, \ldots, \rho_r) = F_{\text{SDP}}(\rho_1, \ldots, \rho_r)$ for invertible states. Finally, in the fourth step, with the assistance of Lemma 1 and Theorem 5.12 among other things, we establish the equivalence (5.34) for the general setting.

Step 1: Equivalent formulation for $K^\star$-representation

Define

$$F_{\text{Int}}(\rho_1, \ldots, \rho_r) \coloneqq \frac{2}{r(r-1)} \sup_{K_{ji} = K_{ij}^\dagger} \left\{ \sum_{i<j} \Re\left[\operatorname{Tr}\left[\rho_i^{1/2} K_{ij} \rho_j^{1/2}\right]\right] : I_r \otimes I_d + \sum_{i \neq j} |i\rangle\!\langle j| \otimes K_{ij} \geq 0 \right\}. \tag{A.85}$$

We first show that $F_{\text{Int}}(\rho_1, \ldots, \rho_r) = F_{K^\star}(\rho_1, \ldots, \rho_r)$. For all $i \in [r]$, define the canonical purification of ρ_i as

$$|\phi^{\rho_i}\rangle = (\rho_i^{1/2} \otimes I_d)|\Gamma\rangle, \tag{A.86}$$

where $|\Gamma\rangle \coloneqq \sum_{i=1}^d |i\rangle \otimes |i\rangle$ is the unnormalized maximally entangled vector and d denotes the dimension of the Hilbert space on which each ρ_i is defined. With this, we can write each term in the objective function as

$$\Re\left[\operatorname{Tr}\left[\rho_i^{1/2} K_{ij} \rho_j^{1/2}\right]\right] = \Re[\langle \phi^{\rho_i}| K_{ij} \otimes I_d |\phi^{\rho_j}\rangle]. \tag{A.87}$$

Recalling the definition of K from (A.85)

$$K \coloneqq I_r \otimes I_d + \sum_{i \neq j} |i\rangle\!\langle j| \otimes K_{ij}, \tag{A.88}$$

and considering that $K \geq 0$, we also obtain that

$$\frac{1}{\sqrt{r}} \left(\sum_{k=1}^{r} \langle k| \langle \phi^{\rho_k}| \right) \left(I_r \otimes I_d \otimes I_d + \sum_{i \neq j=1}^{r} |i\rangle\!\langle j| \otimes K_{ij} \otimes I_d \right) \frac{1}{\sqrt{r}} \left(\sum_{\ell=1}^{r} |\ell\rangle |\phi^{\rho_\ell}\rangle \right) \geq 0 \tag{A.89}$$

and

$$\begin{aligned} &\frac{1}{\sqrt{r}} \left(\sum_{k=1}^{r} \langle k| \langle \phi^{\rho_k}| \right) \left(I_r \otimes I_d \otimes I_d + \sum_{i \neq j=1}^{r} |i\rangle\!\langle j| \otimes K_{ij} \otimes I_d \right) \frac{1}{\sqrt{r}} \left(\sum_{\ell=1}^{r} |\ell\rangle |\phi^{\rho_\ell}\rangle \right) \\ &= 1 + \frac{1}{r} \left(\sum_{i \neq j=1}^{r} \langle \phi^{\rho_i}| K_{ij} \otimes I_d |\phi^{\rho_j}\rangle \right) && \text{(A.90)} \\ &= 1 + \frac{2}{r} \left(\sum_{i<j} \Re \left[\langle \phi^{\rho_i}| K_{ij} \otimes I_d |\phi^{\rho_j}\rangle \right] \right). && \text{(A.91)} \end{aligned}$$

The above evaluates to the following equality:

$$\sum_{i<j} \Re \left[\langle \phi^{\rho_i}| K_{ij} \otimes I_d |\phi^{\rho_j}\rangle \right] = \frac{r}{2} \left(\langle \psi| K \otimes I_d |\psi\rangle - 1 \right), \tag{A.92}$$

where

$$|\psi\rangle = \frac{1}{\sqrt{r}} \sum_{i=1}^{r} |i\rangle |\phi^{\rho_i}\rangle. \tag{A.93}$$

Then, for the canonical purification of each ρ_i, we deduce that the objective function of the optimizations in both $F_{\text{Int}}(\rho_1, \ldots, \rho_r)$ and $F_{K^\star}(\rho_1, \ldots, \rho_r)$ are equivalent.

For any other purification, recall that all purifications are related by an isometry. Let

$$|\Phi^{\rho_i}\rangle = (U_i \otimes I_d) |\phi^{\rho_i}\rangle, \tag{A.94}$$

where U_i is an isometry for all $i \in [r]$. Then, the objective function inside the SDP (i.e., without considering the constants) evaluates to

$$\sum_{i<j} \Re \left[\langle \Phi^{\rho_i}| K_{ij} \otimes I_d |\Phi^{\rho_j}\rangle \right] = \sum_{i<j} \Re \left[\langle \phi^{\rho_i}| U_i^\dagger K_{ij} U_j \otimes I_d |\phi^{\rho_j}\rangle \right]. \tag{A.95}$$

Also note that the matrix formed by replacing K_{ij} in (A.88) with $U_i^\dagger K_{ij} U_j$ also satisfies the constraint due to:

$$\sum_{i=1}^{r} |i\rangle\!\langle i| \otimes U_i^\dagger \left(I_r \otimes I_d + \sum_{i \neq j} |i\rangle\!\langle j| \otimes K_{ij} \right) \sum_{k=1}^{r} |k\rangle\!\langle k| \otimes U_k \geq 0, \tag{A.96}$$

which follows from the fact that $U_i^\dagger U_i = I$. Then, the previous arguments in the canonical purification follow for all purifications of the considered states. We then conclude that

$$F_{K^\star}(\rho_1, \ldots, \rho_r) = F_{\mathrm{Int}}(\rho_1, \ldots, \rho_r) \tag{A.97}$$

$$= \frac{2}{r(r-1)} \sup_{K_{ji}=K_{ij}^\dagger} \left\{ \sum_{i<j} \mathfrak{R}\left[\mathrm{Tr}\left[\rho_i^{1/2} K_{ij} \rho_j^{1/2}\right]\right] : I_r \otimes I_d + \sum_{i \neq j} |i\rangle\!\langle j| \otimes K_{ij} \geq 0 \right\}. \tag{A.98}$$

Step 2: $F_{K^\star}(\rho_1, \ldots, \rho_r) \leq F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r)$ for general states

Consider the constraint: $I_r \otimes I_d + \sum_{i \neq j} |i\rangle\!\langle j| \otimes K_{ij} \geq 0$ in (A.98). By left and right multiplying the above inequality by $\sum_{i=1}^r |i\rangle\!\langle i| \otimes \rho_i^{1/2}$, we obtain

$$\sum_{i=1}^r |i\rangle\!\langle i| \otimes \rho_i + \sum_{i \neq j} |i\rangle\!\langle j| \otimes \rho_i^{1/2} K_{ij} \rho_j^{1/2} \geq 0. \tag{A.99}$$

Then referring to the dual SDP of multivariate SDP fidelity in Proposition 5.10, we see that a feasible point in the optimization (5.27) is given by $X_{ij} = \rho_i^{1/2} K_{ij} \rho_j^{1/2}$ for $i \neq j$. With that we have:

$$\frac{2}{r(r-1)} \sum_{i<j} \mathfrak{R}\left[\mathrm{Tr}\left[\rho_i^{1/2} K_{ij} \rho_j^{1/2}\right]\right] \leq F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r). \tag{A.100}$$

Then, optimizing over all K satisfying the required constraint, we arrive at the claim:

$$F_{K^\star}(\rho_1, \ldots, \rho_r) \leq F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r). \tag{A.101}$$

Step 3: $F_{K^\star}(\rho_1, \ldots, \rho_r) = F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r)$ for invertible states

Consider an arbitrary feasible point in the dual SDP (5.27) of multivariate SDP fidelity:

$$\sum_{i=1}^r |i\rangle\!\langle i| \otimes \rho_i + \sum_{i \neq j} |i\rangle\!\langle j| \otimes X_{ij} \geq 0. \tag{A.102}$$

Using the assumption that ρ_i is invertible for all $i \in [r]$, by left and right multiplying the above inequality by the positive definite matrix $\sum_{i=1}^r |i\rangle\!\langle i| \otimes \rho_i^{-1/2}$, we obtain

$$\sum_{i=1}^r |i\rangle\!\langle i| \otimes I_d + \sum_{i \neq j} |i\rangle\!\langle j| \otimes \rho_i^{-1/2} X_{ij} \rho_j^{-1/2} \geq 0. \tag{A.103}$$

This implies that the matrices $K_{ij} = \rho_i^{-1/2} X_{ij} \rho_j^{-1/2}$ for $i \neq j$ form a feasible point for the optimization in (A.98). With this choice of a feasible point, we get

$$\frac{2}{r(r-1)} \sum_{i<j} \mathfrak{R}[\mathrm{Tr}[X_{ij}]] \leq F_{K^\star}(\rho_1, \ldots, \rho_r). \tag{A.104}$$

Since the above inequality holds for an arbitrary feasible point in (5.27), it follows by the definition that

$$F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) \leq F_{K^\star}(\rho_1, \ldots, \rho_r). \tag{A.105}$$

Combining (A.105) with (A.101), we get

$$F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) = F_{K^\star}(\rho_1, \ldots, \rho_r) \tag{A.106}$$

for invertibles states.

Step 4: Extending to non-invertible states

For all $i \in [r]$ and $\varepsilon > 0$,

$$\rho_i^{(\varepsilon)} \coloneqq \frac{1}{1+\varepsilon}\left(\rho_i + \frac{\varepsilon}{d} I\right) \tag{A.107}$$

is an invertible state. Then, we have that

$$\frac{1}{2}\left\|\rho_i - \rho_i^{(\varepsilon)}\right\|_1 = \frac{\varepsilon}{2(1+\varepsilon)}\left\|\rho_i - \frac{I}{d}\right\|_1 \leq \frac{\varepsilon}{1+\varepsilon} =: \varepsilon', \tag{A.108}$$

leading to $F\left(\rho_i, \rho_i^{(\varepsilon)}\right) \geq 1 - \varepsilon'$. Then by Theorem 5.12, we get

$$\left|F_{K^\star}(\rho_1, \ldots, \rho_r) - F_{K^\star}\left(\rho_1^{(\varepsilon)}, \ldots, \rho_r^{(\varepsilon)}\right)\right| \leq \frac{r}{r-1}\sqrt{\varepsilon'(2-\varepsilon')}. \tag{A.109}$$

Using Lemma 1 with the choice $\varepsilon_1 = 0$ and $\varepsilon_2 = \varepsilon > 0$, we have that

$$\begin{aligned} F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) &\leq (1+\varepsilon) F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon)}, \ldots, \rho_r^{(\varepsilon)}\right) && \text{(A.110)} \\ &\leq F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon)}, \ldots, \rho_r^{(\varepsilon)}\right) + \varepsilon, && \text{(A.111)} \end{aligned}$$

where the last inequality holds because $F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon)}, \ldots, \rho_r^{(\varepsilon)}\right) \leq 1$. Then consider that

$$\begin{aligned} F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon)}, \ldots, \rho_r^{(\varepsilon)}\right) &= F_{K^\star}\left(\rho_1^{(\varepsilon)}, \ldots, \rho_r^{(\varepsilon)}\right) && \text{(A.112)} \\ &\leq F_{K^\star}(\rho_1, \ldots, \rho_r) + \frac{r}{r-1}\sqrt{\varepsilon'(2-\varepsilon')} && \text{(A.113)} \\ &\leq F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) + \frac{r}{r-1}\sqrt{\varepsilon'(2-\varepsilon')}, && \text{(A.114)} \end{aligned}$$

where the first equality follows from (A.106); the second inequality from (A.109); and finally the third inequality from (A.101) for general states.

Combining (A.111) and (A.114), we arrive at

$$\left|F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) - F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon)}, \ldots, \rho_r^{(\varepsilon)}\right)\right| \leq \max\left\{\varepsilon, \frac{r}{r-1}\sqrt{\varepsilon'(2-\varepsilon')}\right\} \tag{A.115}$$

$$\leq 2\sqrt{\varepsilon'(2-\varepsilon')} \tag{A.116}$$

$$= 2\sqrt{\frac{\varepsilon}{1+\varepsilon}\left(2-\frac{\varepsilon}{1+\varepsilon}\right)}. \tag{A.117}$$

With the established machinery, consider

$$\begin{aligned}
&|F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r) - F_{K^\star}(\rho_1,\ldots,\rho_r)| \\
&\leq \left|F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r) - F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon)},\ldots,\rho_r^{(\varepsilon)}\right)\right| \\
&\qquad + \left|F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon)},\ldots,\rho_r^{(\varepsilon)}\right) - F_{K^\star}\left(\rho_1^{(\varepsilon)},\ldots,\rho_r^{(\varepsilon)}\right)\right| \\
&\qquad + \left|F_{K^\star}(\rho_1,\ldots,\rho_r) - F_{K^\star}\left(\rho_1^{(\varepsilon)},\ldots,\rho_r^{(\varepsilon)}\right)\right| \qquad \text{(A.118)} \\
&\leq 2\sqrt{\frac{\varepsilon}{1+\varepsilon}\left(2-\frac{\varepsilon}{1+\varepsilon}\right)} + 0 + \frac{r}{r-1}\sqrt{\frac{\varepsilon}{1+\varepsilon}\left(2-\frac{\varepsilon}{1+\varepsilon}\right)}, \qquad \text{(A.119)}
\end{aligned}$$

where the first inequality follows from triangular inequality; the second inequality from using (A.109) and (A.117) along with the fact that $F_{\mathrm{SDP}}\left(\rho_1^{(\varepsilon)},\ldots,\rho_r^{(\varepsilon)}\right) = F_{K^\star}\left(\rho_1^{(\varepsilon)},\ldots,\rho_r^{(\varepsilon)}\right)$ due to (A.106).

Finally, by taking limit $\varepsilon \to 0$ in the above inequality, we conclude that

$$|F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r) - F_{K^\star}(\rho_1,\ldots,\rho_r)| = 0, \tag{A.120}$$

which completes the proof for the general case.

A.9 Proof of Theorem 5.15 (Properties of multivariate SDP fidelity)

Note that multivariate SDP fidelity satisfies reduction to classical average pairwise fidelity, faithfulness, and orthogonality by Theorem 5.28 and the fact that average pairwise fidelities satisfy these properties as stated in Theorem 5.5.

Data processing: Let $\mathcal{N} : \mathscr{L}(\mathcal{H}_A) \to \mathscr{L}(\mathcal{H}_B)$ be a channel, where the dimension of $\mathcal{H}_A$ and $\mathcal{H}_B$ be d_A and d_B, respectively. For $i \in [r]$, let Y_i be such that $\sum_{i=1}^r |i\rangle\!\langle i| \otimes Y_i \geq \sum_{i\neq j} |i\rangle\!\langle j| \otimes I_{d_B}$. Consider that

$$\frac{1}{r(r-1)}\sum_{i=1}^{r} \mathrm{Tr}[Y_i\mathcal{N}(\rho_i)] = \frac{1}{r(r-1)}\sum_{i=1}^{r} \mathrm{Tr}\left[\mathcal{N}^\dagger(Y_i)\rho_i\right]. \tag{A.121}$$

Then, due to $\mathcal{N}$ being completely positive, along with $\mathcal{N}$ being trace-preserving (so that $\mathcal{N}^\dagger(I_{d_B}) = I_{d_A}$), we have that

$$0 \leq \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes \mathcal{N}^\dagger(Y_i) - \sum_{i\neq j} |i\rangle\!\langle j| \otimes \mathcal{N}^\dagger(I_{d_B}) = \sum_{i=1}^{r} |i\rangle\!\langle i| \otimes \mathcal{N}^\dagger(Y_i) - \sum_{i\neq j} |i\rangle\!\langle j| \otimes I_{d_A}. \tag{A.122}$$

With that, the tuple $(\mathcal{N}^\dagger(Y_i))_{i=1}^r$ is a candidate for $F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r)$. Thus,

$$F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) \leq \frac{1}{r(r-1)} \sum_{i=1}^r \mathrm{Tr}[Y_i \mathcal{N}(\rho_i)] \,. \tag{A.123}$$

This holds for all $(Y_i)_{i=1}^r$ satisfying $\sum_{i=1}^r |i\rangle\!\langle i| \otimes Y_i \geq \sum_{i \neq j} |i\rangle\!\langle j| \otimes I_{d_A}$. Finally, taking the infimum over all such $(Y_i)_{i=1}^r$ concludes the proof.

Remark A.1. *We note that the proof for data processing given above holds not just for completely positive, trace-preserving maps, but more generally for r-positive, trace-preserving maps.*

Symmetry: This follows directly from the definition of the SDP fidelity.

Direct-sum property: Let $(Y_i^x)_{i=1}^r$ for all $x \in \mathcal{X}$ satisfy $\sum_{i=1}^r |i\rangle\!\langle i| \otimes Y_i^x - \sum_{i \neq j} |i\rangle\!\langle j| \otimes I \geq 0$. Consider that

$$\begin{aligned}
&\frac{1}{r(r-1)} \sum_{x \in \mathcal{X}} p(x) \sum_{i=1}^r \mathrm{Tr}[Y_i^x \rho_i^x] \\
&= \frac{1}{r(r-1)} \sum_{i=1}^r \mathrm{Tr}\left[\sum_{x \in \mathcal{X}} p(x) Y_i^x \rho_i^x\right] && \text{(A.124)}\\
&= \frac{1}{r(r-1)} \sum_{i=1}^r \mathrm{Tr}\left[\left(\sum_{x' \in \mathcal{X}} |x'\rangle\!\langle x'| \otimes Y_i^{x'}\right) \sum_{x \in \mathcal{X}} p(x) |x\rangle\!\langle x| \otimes \rho_i^x\right]. && \text{(A.125)}
\end{aligned}$$

From the conditions satisfied by $(Y_i^x)_{i=1}^r$, we also get

$$\sum_{i=1}^r |i\rangle\!\langle i| \otimes |x\rangle\!\langle x| \otimes Y_i^x - \sum_{i \neq j} |i\rangle\!\langle j| \otimes |x\rangle\!\langle x| \otimes I_d \geq 0. \tag{A.126}$$

By summing over $x \in \mathcal{X}$, we get

$$\sum_{i=1}^r |i\rangle\!\langle i| \otimes \sum_{x \in \mathcal{X}} |x\rangle\!\langle x| \otimes Y_i^x - \sum_{i \neq j} |i\rangle\!\langle j| \otimes I_{|\mathcal{X}|} \otimes I_d \geq 0. \tag{A.127}$$

Then, $\left(\sum_{x \in \mathcal{X}} |x\rangle\!\langle x| \otimes Y_i^x\right)_{i=1}^r$ forms a feasible point for the optimization in $F_{\mathrm{SDP}}\left(\sum_{x \in \mathcal{X}} p(x) |x\rangle\!\langle x| \otimes \rho_1^x, \ldots, \sum_{x \in \mathcal{X}} p(x) |x\rangle\!\langle x| \otimes \rho_r^x\right)$. Using the relation (A.125), this leads to

$$F_{\mathrm{SDP}}\left(\sum_{x \in \mathcal{X}} p(x) |x\rangle\!\langle x| \otimes \rho_1^x, \ldots, \sum_{x \in \mathcal{X}} p(x) |x\rangle\!\langle x| \otimes \rho_r^x\right) \leq \frac{1}{r(r-1)} \sum_{x \in \mathcal{X}} p(x) \sum_{i=1}^r \mathrm{Tr}[Y_i^x \rho_i^x]. \tag{A.128}$$

Since the last inequality holds for all $(Y_i^x)_{i=1}^r$ and for all $x \in \mathcal{X}$, this also holds for the optimal ones achieving $F(\rho_1^x, \ldots, \rho_r^x)$. We thus get

$$F_{\mathrm{SDP}}\left(\sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_1^x, \ldots, \sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_r^x\right) \leq \sum_{x\in\mathcal{X}} p(x) F_{\mathrm{SDP}}(\rho_1^x, \ldots, \rho_r^x). \quad \text{(A.129)}$$

To prove the reverse direction, let $(Z_i)_{i=1}^r$ satisfy $\sum_{i=1}^r |i\rangle\langle i| \otimes Z_i - \sum_{i\neq j} |i\rangle\langle j| \otimes I_{|\mathcal{X}|} \otimes I_d \geq 0$. Then

$$\frac{1}{r(r-1)} \sum_{i=1}^r \mathrm{Tr}\left[Z_i \sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_i^x\right] = \sum_{x\in\mathcal{X}} p(x) \frac{1}{r(r-1)} \sum_{i=1}^r \mathrm{Tr}[(\langle x| \otimes I_d) Z_i (|x\rangle \otimes I_d) \rho_i^x]\,. \quad \text{(A.130)}$$

From the conditions satisfied by $(Z_i)_{i=1}^r$, we also get for all $x \in \mathcal{X}$ that

$$\sum_{i=1}^r |i\rangle\langle i| \otimes (\langle x| \otimes I_d) Z_i (|x\rangle \otimes I_d) - \sum_{i\neq j} |i\rangle\langle j| \otimes \langle x|x\rangle \otimes I_d \geq 0$$
$$\implies \quad \sum_{i=1}^r |i\rangle\langle i| \otimes (\langle x| \otimes I_d) Z_i (|x\rangle \otimes I_d) - \sum_{i\neq j} |i\rangle\langle j| \otimes I_d \geq 0. \quad \text{(A.131)}$$

Then $((\langle x| \otimes I_d) Z_i (|x\rangle \otimes I_d))_{i=1}^r$ is a candidate for $F_{\mathrm{SDP}}(\rho_1^x, \ldots, \rho_r^x)$, leading to

$$F_{\mathrm{SDP}}(\rho_1^x, \ldots, \rho_r^x) \leq \frac{1}{r(r-1)} \sum_{i=1}^r \mathrm{Tr}[(\langle x| \otimes I_d) Z_i (|x\rangle \otimes I_d) \rho_i^x]\,. \quad \text{(A.132)}$$

Summing over all $x \in \mathcal{X}$ together with (A.130) leads to

$$\sum_{x\in\mathcal{X}} p(x) F_{\mathrm{SDP}}(\rho_1^x, \ldots, \rho_r^x) \leq \frac{1}{r(r-1)} \sum_{i=1}^r \mathrm{Tr}\left[Z_i \sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_i^x\right]. \quad \text{(A.133)}$$

Since the above inequality also holds for an optimal choice of $(Z_i)_{i=1}^r$, we arrive at

$$\sum_{x\in\mathcal{X}} p(x) F_{\mathrm{SDP}}(\rho_1^x, \ldots, \rho_r^x) \leq F_{\mathrm{SDP}}\left(\sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_1^x, \ldots, \sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_r^x\right). \quad \text{(A.134)}$$

Lastly, by combining (A.129) and (A.134), we conclude the proof of (3.8).

Joint concavity: Using the direct-sum property (property 6 in Theorem 5.15), we have

$$F_{\mathrm{SDP}}\left(\sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_1^x, \ldots, \sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_r^x\right) = \sum_{x\in\mathcal{X}} p(x) F_{\mathrm{SDP}}(\rho_1^x, \ldots, \rho_r^x). \quad \text{(A.135)}$$

Then, using the above and applying data processing with respect to the partial trace channel (i.e., taking a partial trace over the classical system) to the left-hand side concludes the proof.

A.10 Proof of Proposition 5.17 (Coarse-graining property of SDP fidelity)

Let $(Y_i)_{i=1}^{r+m}$ be such that

$$\sum_{i=1}^{r+m} |i\rangle\langle i| \otimes Y_i \geq \sum_{i \neq j} |i\rangle\langle j| \otimes I. \tag{A.136}$$

Then that choice is a possible candidates for the SDP of $F_{\text{SDP}}(\rho_1, \ldots, \rho_r, \ldots, \rho_{r+m})$ with the extension of (5.26) to $r+m$ states. Consider

$$\begin{aligned} \sum_{i=1}^{r+m} \operatorname{Tr}[Y_i \rho_i] &\geq \sum_{i=1}^{r} \operatorname{Tr}[Y_i \rho_i] & \text{(A.137)} \\ &\geq r(r-1) F_{\text{SDP}}(\rho_1, \ldots, \rho_r) & \text{(A.138)} \end{aligned}$$

where the first inequality follows from $Y_i \geq 0$ for all $i \in [r+m]$, and the last inequality from the fact that $(Y_i)_{i=1}^{r}$ (r elements from the tuple $(Y_i)_{i=1}^{r+m}$) forms a possible candidate for the SDP of $F_{\text{SDP}}(\rho_1, \ldots, \rho_r)$ in (5.26) due to

$$\sum_{i=1}^{r} |i\rangle\langle i| \otimes Y_i \geq \sum_{i \neq j} |i\rangle\langle j| \otimes I. \tag{A.139}$$

We conclude the proof by infimizing the LHS above over all possible candidates $(Y_i)_{i=1}^{r+m}$ satisfying the constraints, which renders

$$(r+m)(r+m-1) F_{\text{SDP}}(\rho_1, \ldots, \rho_r, \ldots, \rho_{r+m}) \geq r(r-1) F_{\text{SDP}}(\rho_1, \ldots, \rho_r). \tag{A.140}$$

A.11 Proof of Proposition 5.21 (Multivariate SDP fidelity for pure states)

Consider the general formula for arbitrary states $(\rho_1, \ldots, \rho_r)$ from Theorem 5.13:

$$F_{\text{SDP}}(\rho_1, \ldots, \rho_r) = \frac{1}{r-1} \sup_{K \geq 0} \left\{ \langle \psi | K \otimes I_d | \psi \rangle - 1 : K = \sum_{i,j=1}^{r} |i\rangle\langle j| \otimes K_{i,j},\ K_{i,i} = I_d \ \forall i \in [r] \right\}, \tag{A.141}$$

where $|\psi\rangle = \frac{1}{\sqrt{r}} \sum_{i=1}^{r} |i\rangle |\psi^{\rho_i}\rangle$ and $|\psi^{\rho_i}\rangle$ is a purification of ρ_i. In the case that each ρ_i is pure, the purification system for each state can be taken to be a trivial one-dimensional system. So then each matrix $K_{i,j}$ reduces to a scalar $k_{i,j} \in \mathbb{C}$, and the constraints in (A.141) reduce to those in (5.47). Then

$$\begin{aligned} &\langle \psi | K \otimes I_d | \psi \rangle \\ &= \left(\frac{1}{\sqrt{r}} \sum_{i''=1}^{r} \langle i'' | \langle \psi_{i''} | \right) \left(\sum_{i,j=1}^{r} k_{i,j} |i\rangle\langle j| \otimes I_d \right) \left(\frac{1}{\sqrt{r}} \sum_{i'=1}^{r} |i'\rangle |\psi_{i'}\rangle \right) & \text{(A.142)} \end{aligned}$$

$$= \frac{1}{r} \sum_{i'',i,j,i'=1}^{r} k_{i,j} \langle i''|i\rangle\langle j|i'\rangle \; \langle\psi_{i''}|\psi_{i'}\rangle \tag{A.143}$$

$$= \frac{1}{r} \sum_{i,j=1}^{r} k_{i,j} \langle\psi_i|\psi_j\rangle \tag{A.144}$$

$$= \frac{1}{r} \sum_{i=1}^{r} k_{i,i} \langle\psi_i|\psi_i\rangle + \frac{1}{r} \sum_{i,j=1:i\neq j}^{r} k_{i,j} \langle\psi_i|\psi_j\rangle \tag{A.145}$$

$$= 1 + \frac{2}{r} \sum_{i<j} \mathfrak{R}[k_{i,j} \langle\psi_i|\psi_j\rangle]. \tag{A.146}$$

Substituting (A.146) into the objective function in (A.141) and simplifying then gives (5.47).

A.12 Proof of Proposition 5.24 (Uniform continuity of secrecy-based multivariate fidelity)

By Definition 5.22,

$$\begin{aligned}
&|F_S(\rho_1, \ldots, \rho_r) - F_S(\sigma_1, \ldots, \sigma_2)| \\
&= \frac{r}{r-1} \left| [S(\rho_1, \ldots, \rho_r)]^2 - [S(\sigma_1, \ldots, \sigma_r)]^2 \right| && \text{(A.147)} \\
&= \frac{r}{r-1} \left[S(\rho_1, \ldots, \rho_r) + S(\rho_1, \ldots, \rho_r) \right] \left| S(\rho_1, \ldots, \rho_r) - S(\rho_1, \ldots, \rho_r) \right| && \text{(A.148)} \\
&\leq \frac{2r}{r-1} \left| S(\rho_1, \ldots, \rho_r) - S(\rho_1, \ldots, \rho_r) \right|, && \text{(A.149)}
\end{aligned}$$

where the last inequality follows because $S(\rho_1, \ldots, \rho_r) \leq 1$ for every tuple $(\rho_i)_{i=1}^r$ of states.

By using (5.48), we have that

$$\begin{aligned}
|S(\rho_1, \ldots, \rho_r) - S(\sigma_1, \ldots, \sigma_r)| &\leq \sup_{\omega \in \mathscr{D}} \left| \frac{1}{r} \sum_{i=1}^{r} (F(\rho_i, \omega) - F(\sigma_i, \omega)) \right| && \text{(A.150)} \\
&\leq \sup_{\omega \in \mathscr{D}} \frac{1}{r} \sum_{i=1}^{r} |F(\rho_i, \omega) - F(\sigma_i, \omega)| && \text{(A.151)}
\end{aligned}$$

For an arbitrary $\omega \in \mathscr{D}$, consider that

$$\begin{aligned}
|F(\rho_i, \omega) - F(\sigma_i, \omega)| &= |[1 - F(\sigma_i, \omega)] - [1 - F(\rho_i, \omega)]| && \text{(A.152)} \\
&= \frac{1}{2} \left| [d_B(\sigma_i, \omega)]^2 - [d_B(\rho_i, \omega)]^2 \right| && \text{(A.153)} \\
&= \frac{1}{2} \left[d_B(\sigma_i, \omega) + d_B(\rho_i, \omega) \right] \left| d_B(\sigma_i, \omega) - d_B(\rho_i, \omega) \right| && \text{(A.154)} \\
&\leq \frac{1}{2} \left[d_B(\sigma_i, \omega) + d_B(\rho_i, \omega) \right] d_B(\sigma_i, \rho_i) && \text{(A.155)}
\end{aligned}$$

$$\leq \sqrt{2} d_B(\sigma_i, \rho_i), \tag{A.156}$$

where the penultimate inequality follows from triangular inequality of the Bures distance and the last inequality because the Bures distance is bounded by $\sqrt{2}$. Together with that, we have

$$|S(\rho_1, \ldots, \rho_r) - S(\sigma_1, \ldots, \sigma_r)| \leq \frac{\sqrt{2}}{r} \sum_{i=1}^{r} d_B(\sigma_i, \rho_i) \leq \sqrt{2}\varepsilon, \tag{A.157}$$

where the last inequality follows from the assumption $\frac{1}{r}\sum_{i=1}^{r} d_B(\rho_i, \sigma_i) \leq \varepsilon$. Finally combining (A.147)–(A.149) and (A.157), we conclude that

$$|F_S(\rho_1, \ldots, \rho_r) - F_S(\sigma_1, \ldots, \sigma_2)| \leq \frac{r}{r-1} 2\sqrt{2}\varepsilon, \tag{A.158}$$

completing the proof.

A.13 Proof of Theorem 5.25 (Properties of secrecy-based multivariate fidelity)

Note that the secrecy-based multivariate fidelity satisfies reduction to classical average pairwise fidelity, faithfulness, and orthogonality by Theorem 5.28 and the fact that average pairwise fidelities satisfy these properties as stated in Theorem 5.5.

Symmetry: This follows by the definition of secrecy measure in (5.48).

Data processing: Consider

$$\sup_{\sigma \in \mathscr{D}} \frac{1}{r} \sum_{i=1}^{r} F(\rho_i, \sigma) \leq \sup_{\sigma \in \mathscr{D}} \frac{1}{r} \sum_{i=1}^{r} F(\mathcal{N}(\rho_i), \mathcal{N}(\sigma)) \tag{A.159}$$

$$\leq \sup_{\sigma' \in \mathscr{D}} \frac{1}{r} \sum_{i=1}^{r} F(\mathcal{N}(\rho_i), \sigma'), \tag{A.160}$$

where the first inequality follows from data processing of the Uhlmann fidelity (Proposition 3.2, Property 2); and the last inequality by supremizing over a larger set. With that we arrive at

$$S(\rho_1, \ldots, \rho_r) \leq S(\mathcal{N}(\rho_1), \ldots, \mathcal{N}(\rho_r)). \tag{A.161}$$

By recalling Definition 5.22, we conclude that

$$F_S(\rho_1, \ldots, \rho_r) \leq F_S(\mathcal{N}(\rho_1), \ldots, \mathcal{N}(\rho_r)). \tag{A.162}$$

Direct-sum property: A purification of the state $\sum_{x \in \mathcal{X}} p(x) |x\rangle\langle x| \otimes \rho_i^x$ is

$$|\Psi_i\rangle \coloneqq \sum_{x} \sqrt{p(x)} |x\rangle \otimes |\Phi^{\rho_i^x}\rangle, \tag{A.163}$$

where $|\Phi^{\rho_i^x}\rangle$ is a purification of state ρ_i^x. Let an optimal unitary of the prover achieving the optimum value of $[S(\rho_1^x,\ldots,\rho_r^x)]^2$ related to the tuple of states $(\rho_i^x)_i$ in (A.211) be $Z^x_{T'RF\to T''F'}$ for all $x\in\mathcal{X}$. i.e.,

$$[S(\rho_1^x,\ldots,\rho_r^x)]^2 = \frac{1}{r^2}\sum_{i,j=1}^{r}\left\langle\Phi^{\rho_i^x}\right|_{RXA}\left(Z^{x,i}_{R\to F'}\right)^\dagger Z^{y,j}_{R\to F'}\left|\Phi^{\rho_j^x}\right\rangle_{RXA}, \tag{A.164}$$

where

$$Z^{x,y}_{R\to F'} \coloneqq \langle y|_{T''}Z^x_{T'RF\to T''F'}|y\rangle_{T'}|0\rangle_F \tag{A.165}$$

Then, choose the controlled unitary

$$Z \coloneqq \sum_x |x\rangle\!\langle x|\otimes Z^x. \tag{A.166}$$

and the resulting contraction

$$Z^y_{R\to XF'} \coloneqq \sum_x |x\rangle\!\langle x|\otimes Z^{x,y}_{R\to F'} \tag{A.167}$$

So, now with the notations developed, consider

$$\begin{aligned}
&= \frac{1}{r^2}\sum_{i,j=1}^{r}\left\langle\Psi^i\right|_{RXA}\left(Z^i_{R\to XF'}\right)^\dagger Z^j_{R\to XF'}\left|\Psi^j\right\rangle_{RXA}\\
&= \frac{1}{r^2}\sum_{i,j=1}^{r}\left(\sum_x\sqrt{p(x)}|x\rangle\otimes\langle\Phi^{\rho_i^x}|\right)\left(\sum_x|x\rangle\!\langle x|\otimes\left(Z^{x,i}_{R\to F'}\right)^\dagger Z^{y,j}_{R\to F'}\right)\times\\
&\qquad\qquad\left(\sum_x\sqrt{p(x)}|x\rangle\otimes|\Phi^{\rho_i^x}\rangle\right) &&\text{(A.168)}\\
&= \frac{1}{r^2}\sum_{i,j=1}^{r}\sum_x p(x)\left\langle\Phi^{\rho_i^x}\right|_{RXA}\left(Z^{x,i}_{R\to F'}\right)^\dagger Z^{y,j}_{R\to F'}\left|\Phi^{\rho_j^x}\right\rangle_{RXA} &&\text{(A.169)}\\
&= \sum_x p(x)\frac{1}{r^2}\sum_{i,j=1}^{r}\left\langle\Phi^{\rho_i^x}\right|_{RXA}\left(Z^{x,i}_{R\to F'}\right)^\dagger Z^{y,j}_{R\to F'}\left|\Phi^{\rho_j^x}\right\rangle_{RXA} &&\text{(A.170)}\\
&= \sum_x p(x)\left[S(\rho_1^x,\ldots,\rho_r^x)\right]^2, &&\text{(A.171)}
\end{aligned}$$

where the first inequality follows by substituting the definitions in (A.163) and (A.167), and the last equality from (A.164).

This shows that controlled unitary in (A.166) is a possible candidate for the unitary in the optimization of $\left[S\left(\sum_{x\in\mathcal{X}}p(x)|x\rangle\!\langle x|\otimes\rho_1^x,\ldots,\sum_{x\in\mathcal{X}}p(x)|x\rangle\!\langle x|\otimes\rho_r^x\right)\right]^2$. With that, we conclude the following inequality:

$$\left[S\left(\sum_{x\in\mathcal{X}}p(x)|x\rangle\!\langle x|\otimes\rho_1^x,\ldots,\sum_{x\in\mathcal{X}}p(x)|x\rangle\!\langle x|\otimes\rho_r^x\right)\right]^2 \geq \sum_{x\in\mathcal{X}}p(x)\left[S(\rho_1^x,\ldots,\rho_r^x)\right]^2. \tag{A.172}$$

Now we prove the opposite inequality. Let $\Delta(\cdot) \coloneqq \sum_x |x\rangle\!\langle x|(\cdot)|x\rangle\!\langle x|$ denote the dephasing channel. Recall that

$$S^2\left(\sum_x p(x)|x\rangle\!\langle x| \otimes \rho_1^x, \ldots, \sum_x p(x)|x\rangle\!\langle x| \otimes \rho_r^x\right) \\ = \left[\sup_{\sigma_{XA}} \frac{1}{r}\sum_{i=1}^{r} F\left(\sum_x p(x)|x\rangle\!\langle x| \otimes \rho_i^x, \sigma_{XA}\right)\right]^2. \quad \text{(A.173)}$$

Let σ_{XA} be an arbitrary state, and define the probability distribution q and states $(\sigma^x)_x$ in terms of

$$(\Delta \otimes \mathrm{id})\,(\sigma_{XA}) = \sum_x q(x)|x\rangle\!\langle x| \otimes \sigma^x, \quad \text{(A.174)}$$

where id is the Identity channel. Then

$$\frac{1}{r}\sum_{i=1}^{r} F\left(\sum_x p(x)|x\rangle\!\langle x| \otimes \rho_i^x, \sigma_{XA}\right)$$

$$\leq \frac{1}{r}\sum_{i=1}^{r} F\left((\Delta \otimes \mathrm{id})\left(\sum_x p(x)|x\rangle\!\langle x| \otimes \rho_i^x\right), (\Delta \otimes \mathrm{id})\,(\sigma_{XA})\right) \quad \text{(A.175)}$$

$$= \frac{1}{r}\sum_{i=1}^{r} F\left(\sum_x p(x)|x\rangle\!\langle x| \otimes \rho_i^x, \sum_x q(x)|x\rangle\!\langle x| \otimes \sigma^x\right) \quad \text{(A.176)}$$

$$= \frac{1}{r}\sum_{i=1}^{r}\sum_x \sqrt{p(x)q(x)}F(\rho_i^x, \sigma^x) \quad \text{(A.177)}$$

$$= \sum_x \sqrt{p(x)q(x)}\frac{1}{r}\sum_{i=1}^{r} F(\rho_i^x, \sigma^x)\,. \quad \text{(A.178)}$$

Now applying the Cauchy–Schwarz inequality, we conclude that

$$\sum_x \sqrt{p(x)q(x)}\frac{1}{r}\sum_{i=1}^{r} F(\rho_i^x, \sigma^x)$$

$$\leq \sqrt{\sum_x q(x)}\sqrt{\sum_x p(x)\left[\frac{1}{r}\sum_{i=1}^{r} F(\rho_i^x, \sigma^x)\right]^2} \quad \text{(A.179)}$$

$$= \sqrt{\sum_x p(x)\left[\frac{1}{r}\sum_{i=1}^{r} F(\rho_i^x, \sigma^x)\right]^2} \quad \text{(A.180)}$$

$$\leq \sqrt{\sum_x p(x)S^2(\rho_1^x, \ldots, \rho_r^x)}. \quad \text{(A.181)}$$

We have thus shown that the following inequality holds for every state σ_{XA}:

$$\left[\frac{1}{r}\sum_{i=1}^{r} F\left(\sum_x p(x)|x\rangle\!\langle x|\otimes\rho_i^x, \sigma_{XA}\right)\right]^2 \leq \sum_x p(x) S^2(\rho_1^x,\ldots,\rho_r^x). \tag{A.182}$$

By taking the optimization over every state σ_{XA}, we conclude that

$$S^2\left(\sum_x p(x)|x\rangle\!\langle x|\otimes\rho_1^x,\ldots,\sum_x p(x)|x\rangle\!\langle x|\otimes\rho_r^x\right) \leq \sum_x p(x) S^2(\rho_1^x,\ldots,\rho_r^x). \tag{A.183}$$

Note that equality is achieved by picking $\sigma_{XA} = \sum_x p(x)|x\rangle\!\langle x|\otimes\sigma^x$, where σ^x is an optimal choice for $S^2(\rho_1^x,\ldots,\rho_r^x)$. This completes the proof.

A.14 Proof of Proposition 5.26 (Coarse-graining property of secrecy-based fidelity)

Let $\mu \geq 0$ and $(Y_i)_{i=1}^{r+m}$ be a tuple of PSD matrices that satisfy $|0\rangle\!\langle 0|\otimes\mu I_d + \sum_{i=1}^{r+m}|i\rangle\!\langle i|\otimes Y_i \geq \sum_{i=1}^{r+m}(|i\rangle\!\langle 0| + |0\rangle\!\langle i|)\otimes I_d$ and consider the SDP formulation of the secrecy measure in (5.50). Given that Y_i for all $i\in[r+m]$ is PSD, we get

$$\mu + \sum_{i=1}^{r+m}\mathrm{Tr}[Y_i\rho_i] \geq \mu + \sum_{i=1}^{r}\mathrm{Tr}[Y_i\rho_i]. \tag{A.184}$$

Also note that by choosing the first r elements of the tuple $(Y_i)_{i=1}^{r+m}$, $(Y_i)_{i=1}^{r}$, those also satisfy

$$|0\rangle\!\langle 0|\otimes\mu I_d + \sum_{i=1}^{r}|i\rangle\!\langle i|\otimes Y_i \geq \sum_{i=1}^{r}\left(|i\rangle\!\langle 0| + |0\rangle\!\langle i|\right). \tag{A.185}$$

With that, μ and $(Y_i)_{i=1}^{r}$ are possible candidates to the SDP of $S(\rho_1,\ldots,\rho_r)$. Then, we arrive at

$$\frac{1}{2r}\left(\mu + \sum_{i=1}^{r+m}\mathrm{Tr}[Y_i\rho_i]\right) \geq S(\rho_1,\ldots,\rho_r). \tag{A.186}$$

By optimizing over all $\mu\geq 0$ and $(Y_i)_{i=1}^{r+m}$ that satisfy the required constraints, we arrive at

$$\frac{(r+m)}{r}S(\rho_1,\ldots,\rho_{r+m}) \geq S(\rho_1,\ldots,\rho_r). \tag{A.187}$$

Note that by rearranging the terms in Definition 5.22, we have the equality

$$(r-1)F_S(\rho_1,\ldots,\rho_r) + 1 = r\left[S(\rho_1,\ldots,\rho_r)\right]^2. \tag{A.188}$$

Then, by combining above-mentioned relations, we get

$$\frac{r+m}{r}\left((r+m-1)F_S(\rho_1,\ldots,\rho_{r+m}) + 1\right) \geq (r-1)F_S(\rho_1,\ldots,\rho_r) + 1. \tag{A.189}$$

Finally, we arrive at the desired result by algebraic simplifications.

A.15 Proof of Theorem 5.28 (Relations between multivariate fidelities)

We prove the chain of inequalities in the theorem statement from left to right.

$F_H(\rho_1, \ldots, \rho_r) \leq F_S(\rho_1, \ldots, \rho_r)$:

Recall the Petz [Pet86] and sandwiched [MLDS+13, WWY14] Rényi relative entropies, defined for $\alpha \in (0,1) \cup (1,\infty)$ respectively as follows:

$$D_\alpha(\rho\|\sigma) \coloneqq \frac{1}{\alpha-1} \ln \mathrm{Tr}\left[\rho^\alpha \sigma^{1-\alpha}\right], \tag{A.190}$$

$$\widetilde{D}_\alpha(\rho\|\sigma) \coloneqq \frac{1}{\alpha-1} \ln \mathrm{Tr}\left[\left(\sigma^{(1-\alpha)/2\alpha} \rho \sigma^{(1-\alpha)/2\alpha}\right)^\alpha\right]. \tag{A.191}$$

Also note that $\widetilde{D}_{\frac{1}{2}}(\rho\|\sigma) = -2\ln F(\rho,\sigma)$ and $D_{\frac{1}{2}}(\rho\|\sigma) = -2\ln F_H(\rho,\sigma)$. The following inequality is known for $\alpha \in (0,1)$ from [DL14, Lemma 3]:

$$\widetilde{D}_\alpha(\rho\|\sigma) \leq D_\alpha(\rho\|\sigma). \tag{A.192}$$

Let

$$\rho_{XA} \coloneqq \frac{1}{r}\sum_{i=1}^{r} |i\rangle\!\langle i|_X \otimes \rho_i. \tag{A.193}$$

Plugging into (A.192), taking infima with respect to a state σ_A, and setting $\alpha = 1/2$ then gives

$$\inf_{\sigma_A} \widetilde{D}_{\frac{1}{2}}(\rho_{XA}\|\rho_X \otimes \sigma_A) \leq \inf_{\sigma_A} D_{\frac{1}{2}}(\rho_{XA}\|\rho_X \otimes \sigma_A). \tag{A.194}$$

Consider that

$$\inf_{\sigma_A} \widetilde{D}_{\frac{1}{2}}(\rho_{XA}\|\rho_X \otimes \sigma_A) = \inf_{\sigma_A}\left[-2\ln F(\rho_{XA}, \rho_X \otimes \sigma_A)\right] \tag{A.195}$$

$$= -2\ln \sup_{\sigma_A} F(\rho_{XA}, \rho_X \otimes \sigma_A) \tag{A.196}$$

$$= -2\ln \sup_{\sigma} \frac{1}{r}\sum_{i=1}^{r} F(\rho_i, \sigma) \tag{A.197}$$

$$= -2\ln S(\rho_1, \ldots, \rho_r). \tag{A.198}$$

For the penultimate line, we used the direct-sum property of the fidelity in (3.8).

Now applying [GW15, Corollary 8] (which holds for $\alpha \in (0,1)$), consider that

$$\inf_{\sigma_A} D_{\frac{1}{2}}(\rho_{XA}\|\rho_X \otimes \sigma_A)$$

$$= -\ln \mathrm{Tr}\left[\left(\mathrm{Tr}_X[\rho_X^{1/2}\rho_{XA}^{1/2}]\right)^2\right] \tag{A.199}$$

$$= -\ln \operatorname{Tr}\left[\left(\operatorname{Tr}_X\left[\left(\sum_i \frac{1}{r}|i\rangle\langle i| \otimes I\right)^{1/2}\left(\sum_j \frac{1}{r}|j\rangle\langle j| \otimes \rho_j\right)^{1/2}\right]\right)^2\right] \tag{A.200}$$

$$= -\ln \operatorname{Tr}\left[\left(\operatorname{Tr}_X\left[\sum_i \frac{1}{r}|i\rangle\langle i| \otimes \rho_i^{1/2}\right]\right)^2\right] \tag{A.201}$$

$$= -\ln \operatorname{Tr}\left[\left(\sum_i \frac{1}{r}\rho_i^{1/2}\right)^2\right] \tag{A.202}$$

$$= -\ln \operatorname{Tr}\left[\left(\sum_i \frac{1}{r}\rho_i^{1/2}\right)\left(\sum_j \frac{1}{r}\rho_j^{1/2}\right)\right] \tag{A.203}$$

$$= -\ln \operatorname{Tr}\left[\frac{1}{r^2}\sum_i \rho_i + \frac{1}{r^2}\sum_{i\neq j}\rho_i^{1/2}\rho_j^{1/2}\right] \tag{A.204}$$

$$= -\ln\left(\frac{1}{r} + \frac{1}{r^2}\sum_{i\neq j}\operatorname{Tr}\left[\rho_i^{1/2}\rho_j^{1/2}\right]\right) \tag{A.205}$$

$$= -\ln\left(\frac{1}{r} + \frac{2}{r^2}\sum_{i<j}F_H(\rho_i,\rho_j)\right) \tag{A.206}$$

$$= -\ln\left(\frac{1}{r} + \frac{r-1}{r}\left[\frac{2}{r\,(r-1)}\sum_{i<j}F_H(\rho_i,\rho_j)\right]\right) \tag{A.207}$$

$$= -\ln\left(\frac{1}{r} + \frac{r-1}{r}F_H(\rho_1,\ldots,\rho_r)\right). \tag{A.208}$$

Combining (A.194), (A.198), and (A.208), we get

$$-2\ln S(\rho_1,\ldots,\rho_r) \leq -\ln\left(\frac{1}{r} + \frac{r-1}{r}F_H(\rho_1,\ldots,\rho_r)\right), \tag{A.209}$$

which is the same as

$$[S(\rho_1,\ldots,\rho_r)]^2 \geq \frac{1}{r} + \frac{r-1}{r}F_H(\rho_1,\ldots,\rho_r). \tag{A.210}$$

Then, by rearranging the terms in the above inequality, we arrive at the required bound.

$\underline{F_S(\rho_1,\ldots,\rho_r) \leq F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r)}$:

By using (5.52), we have

$$[S(\rho_1,\ldots,\rho_r)]^2 = \sup_{P_{T'RF\to T''F'}\in\mathscr{U}} \frac{1}{r^2}\sum_{x,y=1}^{r} \langle\phi^{\rho_x}|_{RA}\,(P^x_{R\to F'})^\dagger\, P^y_{R\to F'}\,|\phi^{\rho_y}\rangle_{RA} \tag{A.211}$$

$$
= \sup_{P_{T'RF\to T''F'}\in\mathscr{U}} \frac{1}{r^2}\sum_{x=1}^{r} \langle\phi^{\rho_x}|_{RA} \left(P^x_{R\to F'}\right)^\dagger P^x_{R\to F'} |\phi^{\rho_x}\rangle_{RA} + \\
\frac{2}{r^2}\sum_{x<y=1}^{r} \mathfrak{R}\left[\langle\phi^{\rho_x}|_{RA} \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'} |\phi^{\rho_y}\rangle_{RA}\right], \tag{A.212}
$$

where

$$
P^x_{R\to F'} \coloneqq \langle x|_{T''} P_{T'RF\to T''F'} |x\rangle_{T'} |0\rangle_F \tag{A.213}
$$

with $P_{T'RF\to T''F'}$ being a unitary, and $|\phi^{\rho_x}\rangle$ is a purification of state ρ_x for all $x\in[r]$. With $P^x_{R\to F'}$ being a contraction, we have $\left(P^x_{R\to F'}\right)^\dagger P^x_{R\to F'} \leq I_d$ (also see the proof of [RASW23, Theorem 6]). Applying that, we have

$$
[S(\rho_1,\ldots,\rho_r)]^2 \leq \frac{1}{r} + \sup_{P_{T'RF\to T''F'}\in\mathscr{U}} \frac{2}{r^2}\sum_{x<y=1}^{r} \mathfrak{R}\left[\langle\phi^{\rho_x}|_{RA} \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'} |\phi^{\rho_y}\rangle_{RA}\right]. \tag{A.214}
$$

Now, observe that

$$
\sum_{x,y} |x\rangle\!\langle y| \otimes \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'} = \left(\sum_x |x\rangle \otimes \left(P^x_{R\to F'}\right)^\dagger\right)\left(\sum_y \langle y| \otimes P^y_{R\to F'}\right) \geq 0. \tag{A.215}
$$

Again applying $\left(P^x_{R\to F'}\right)^\dagger P^x_{R\to F'} \leq I_d$, we see that

$$
I_r\otimes I_d + \sum_{x\neq y} |x\rangle\!\langle y| \otimes \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'} \geq \sum_{x,y} |x\rangle\!\langle y| \otimes \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'} \geq 0. \tag{A.216}
$$

Then the choice $K_{xy} = \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'}$ for $x\neq y\in[r]$ and $K_{xx}=I_d$ for $x\in[r]$ is a possible candidate to the $F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r)$ in (5.44). To this end, we have

$$
\sum_{x<y=1}^{r} \mathfrak{R}\left[\langle\phi^{\rho_x}|_{RA} \left(P^x_{R\to F'}\right)^\dagger P^y_{R\to F'} |\phi^{\rho_y}\rangle_{RA}\right] \leq \frac{r(r-1)}{2} F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r). \tag{A.217}
$$

Supremizing over $P_{T'RF\to T''F'}\in\mathscr{U}$ and incorporating (A.214) , we have

$$
[S(\rho_1,\ldots,\rho_r)]^2 \leq \frac{1}{r} + \frac{(r-1)}{r} F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r). \tag{A.218}
$$

Rearranging the terms and recalling Definition 5.22, we conclude by arriving at

$$
F_S(\rho_1,\ldots,\rho_r) \leq F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r). \tag{A.219}
$$

$\underline{F_{\mathrm{SDP}}(\rho_1,\ldots,\rho_r) \leq F_U(\rho_1,\ldots,\rho_r)}$:

Let X_{ij} satisfy $X_{ji} = X_{ij}^\dagger$ along with the following relation:

$$
\sum_{i=1}^{r} |i\rangle\!\langle i| \otimes \rho_i + \sum_{i\neq j} |i\rangle\!\langle j| \otimes X_{ij} \geq 0. \tag{A.220}
$$

Then for each pair of states ρ_i, ρ_j with $i < j$, we also have

$$\begin{bmatrix} \rho_i & X_{ij} \\ X_{ij}^\dagger & \rho_j \end{bmatrix} \geq 0. \tag{A.221}$$

Using this together with the dual SDP of Uhlmann fidelity given in (3.4), we get

$$2 \sum_{i<j} \Re\left[\mathrm{Tr}[X_{ij}]\right] \leq 2 \sum_{i<j} F(\rho_i, \rho_j). \tag{A.222}$$

Taking the supremum over all $(X_{ij})_{ij}$ satisfying the relation $\sum_{i=1}^r |i\rangle\!\langle i| \otimes \rho_i + \sum_{i \neq j} |i\rangle\!\langle j| \otimes X_{ij} \geq 0$, we obtain

$$r(r-1) F_{\mathrm{SDP}}(\rho_1, \ldots, \rho_r) \leq 2 \sum_{i<j}^{r} F(\rho_i, \rho_j), \tag{A.223}$$

which implies the desired inequality after rearrangement of terms.

$\underline{F_U(\rho_1, \ldots, \rho_r) \leq \sqrt{F_H(\rho_1, \ldots, \rho_r)}:}$

Notice that for states ρ and σ

$$F(\rho, \sigma) \leq \sqrt{F_H(\rho, \sigma)}, \tag{A.224}$$

which follows by adapting [ANSV08, Theorem 6] with the choice $s = 1/2$, $A = \rho$, and $B = \sigma$. This gives

$$\begin{aligned} F_U(\rho_1, \ldots, \rho_r) &= \frac{2}{r\,(r-1)} \sum_{i<j} F(\rho_i, \rho_j) && \text{(A.225)} \\ &\leq \frac{2}{r\,(r-1)} \sum_{i<j} \sqrt{F_H(\rho_i, \rho_j)} && \text{(A.226)} \\ &\leq \sqrt{\frac{2}{r\,(r-1)} \sum_{i<j} F_H(\rho_i, \rho_j)} && \text{(A.227)} \\ &= \sqrt{F_H(\rho_1, \ldots, \rho_r)}, && \text{(A.228)} \end{aligned}$$

where the last inequality holds due to the concavity of the square root function.

A.16 Proof of Proposition 5.29 (Lower bound on multivariate SDP fidelity)

Fix an arbitrary permutation $\pi \in S_r$. For all $1 \leq i \leq \lfloor r/2 \rfloor$, we have from (5.27) that

$$F(\rho_{\pi(i)}, \rho_{\pi(i+\lfloor r/2 \rfloor)}) = \sup_{Z_i} \left\{ \Re[\mathrm{Tr}[Z_i]] : \begin{bmatrix} \rho_{\pi(i)} & Z_i \\ Z_i^\dagger & \rho_{\pi(i+\lfloor r/2 \rfloor)} \end{bmatrix} \geq 0 \right\}. \tag{A.229}$$

Consider arbitrary operators Z_i satisfying the constraint of (A.229):

$$\begin{bmatrix} \rho_{\pi(i)} & Z_i \\ Z_i^\dagger & \rho_{\pi(i+\lfloor r/2 \rfloor)} \end{bmatrix} \geq 0, \qquad 1 \leq i \leq \lfloor r/2 \rfloor. \tag{A.230}$$

Define operators X_{ij} for $1 \leq i < j \leq r$ as

$$X_{ij} = \begin{cases} Z_i & \text{if } 1 \leq i \leq \lfloor r/2 \rfloor \text{ and } j = i + \lfloor r/2 \rfloor, \\ 0 & \text{otherwise,} \end{cases} \tag{A.231}$$

where 0 denotes the zero operator. The positive semi-definiteness of the matrices in (A.230) implies that

$$\sum_{i=1}^{r} |i\rangle\langle i| \otimes \rho_{\pi(i)} + \sum_{i \neq j} |i\rangle\langle j| \otimes X_{ij} \geq 0. \tag{A.232}$$

Hence, the SDP dual (5.27) and the permutation symmetry of the multivariable SDP fidelity yield

$$F_{\text{SDP}}(\rho_1, \ldots, \rho_r) \geq \frac{2}{r(r-1)} \sum_{1 \leq i < j \leq r} \Re[\text{Tr}[X_{ij}]] \tag{A.233}$$

$$= \frac{2}{r(r-1)} \sum_{i=1}^{\lfloor \frac{r}{2} \rfloor} \Re[\text{Tr}[Z_i]] \,. \tag{A.234}$$

Combining (A.229) and (A.234), we get

$$F_{\text{SDP}}(\rho_1, \ldots, \rho_r) \geq \frac{2}{r(r-1)} \sum_{i=1}^{\lfloor \frac{r}{2} \rfloor} F(\rho_{\pi(i)}, \rho_{\pi(i)+\lfloor \frac{r}{2} \rfloor}). \tag{A.235}$$

Since the above inequality holds for arbitrary $\pi \in S_r$, the desired inequality (5.60) follows.

A.17 Proof of Proposition 5.33 (Properties of measured multivariate fidelity)

Symmetry follows by the symmetry followed by the underlying classical multivariate fidelity.

Data processing: Let Λ' be a measurement channel and $\mathcal{N}$ be a channel. Then consider

$$\mathbf{F}\big(\Lambda'\big(\mathcal{N}(\rho_1)\big), \ldots, \Lambda'\big(\mathcal{N}(\rho_r)\big)\big) \geq \inf_{(\Lambda_x)_x} \mathbf{F}(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)) \tag{A.236}$$

due to the composed measurement channel $\Lambda' \circ \mathcal{N}$ being one of the choices of measurement channels. Infimizing over all measurement channels Λ', we arrive at

$$\inf_{(\Lambda'_x)_x} \mathbf{F}\big(\Lambda'\big(\mathcal{N}(\rho_1)\big), \ldots, \Lambda'\big(\mathcal{N}(\rho_r)\big)\big) \geq \inf_{(\Lambda_x)_x} \mathbf{F}(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)) \,, \tag{A.237}$$

proving the data processing property.

Faithfulness: When $\rho_i = \rho_j$, by the faithfulness of multivariate fidelity we have

$$\overline{\mathbf{F}}(\rho_1, \ldots, \rho_r) \geq \mathbf{F}(\rho_1, \ldots, \rho_r) = 1, \tag{A.238}$$

Since the underlying classical multivariate fidelity is less than or equal to one, we also have $\overline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = 1$.

To prove the reverse implication, assume $\overline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = 1$. Again due to the fact that the underlying classical multivariate fidelity is less than or equal to one, under all measurement channels we have $\mathbf{F}(\Lambda(\rho_1), \ldots, \Lambda(\rho_r)) = 1$. This includes tomographically complete measurements, which lead to distributions that are in one-to-one correspondence with the underlying density operators (i.e., there is a linear invertible map relating the resulting distributions and the original density operators). From the faithfulness of the underlying classical multivariate fidelities, the distributions are the same, and due to the aforementioned bijection, the states are also the same. i.e., $\{M_x\}_{x=1}^r$ with $M_x > 0$, $\sum_{x=1}^r M_x = I$ and $\operatorname{Tr}[M_x(\rho_i - \rho_j)]$ implies that $\rho_i = \rho_j$.

Orthogonality: Let $\overline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = 0$. Then we have

$$0 = \overline{\mathbf{F}}(\rho_1, \ldots, \rho_r) \geq \mathbf{F}(\rho_1, \ldots, \rho_r). \tag{A.239}$$

Then, by orthogonality of multivariate fidelity that generalizes the underlying classical multivariate fidelity (e.g., average pairwise z-fidelity that generalizes average pairwise fidelities), we have the condition that for $\rho_i\rho_j = 0$ for all $i \neq j \in [r]$.

To prove the reverse, assume that $\rho_i\rho_j = 0$ for all $i \neq j \in [r]$. Then there exists a measurement to distinguish them perfectly, leading to orthogonal distributions. Then the underlying multivariate classical fidelity is equal to zero.

Direct-sum property: For all $i \in [r]$, define

$$\rho^i_{XA} \coloneqq \sum_x p(x)|x\rangle\!\langle x| \otimes \rho_i^x, \tag{A.240}$$

where $(p(x))_x$ is a probability distribution, $\{|x\rangle\}_x$ is an orthonormal basis, and each ρ_i^x is a state.

Defining the measurement channels

$$\Lambda(\omega_{XA}) \coloneqq \sum_y \operatorname{Tr}[\Lambda_y \omega_{XA}]|y\rangle\!\langle y|, \tag{A.241}$$

$$\Lambda^x(\tau_A) \coloneqq \sum_y \operatorname{Tr}[\Lambda_y^x \tau_A]|y\rangle\!\langle y|, \tag{A.242}$$

$$\widetilde{\Lambda}(\omega_{XA}) \coloneqq \sum_{x,y} \operatorname{Tr}\!\left[\left(|x\rangle\!\langle x| \otimes \Lambda_y^x\right) \omega_{XA}\right] |x\rangle\!\langle x| \otimes |y\rangle\!\langle y|, \tag{A.243}$$

consider that

$$
\begin{aligned}
&\overline{\mathbf{F}}(\rho_{XA}^1, \ldots, \rho_{XA}^r) \\
&= \inf_{\Lambda} \mathbf{F}(\Lambda(\rho_{XA}^1), \ldots, \Lambda(\rho_{XA}^r)) && \text{(A.244)} \\
&\leq \inf_{\widetilde{\Lambda}} \mathbf{F}(\widetilde{\Lambda}(\rho_{XA}^1), \ldots, \widetilde{\Lambda}(\rho_{XA}^r)) && \text{(A.245)} \\
&= \inf_{\widetilde{\Lambda}} \mathbf{F}\left(\sum_x p(x)|x\rangle\!\langle x| \otimes \Lambda^x(\rho_1^x), \ldots, \sum_x p(x)|x\rangle\!\langle x| \otimes \Lambda^x(\rho_r^x)\right) && \text{(A.246)} \\
&= \inf_{\widetilde{\Lambda}} \sum_x p(x)\mathbf{F}(\Lambda^x(\rho_1^x), \ldots, \Lambda^x(\rho_r^x)) && \text{(A.247)} \\
&= \sum_x p(x) \inf_{\Lambda^x} \mathbf{F}(\Lambda^x(\rho_1^x), \ldots, \Lambda^x(\rho_r^x)) && \text{(A.248)} \\
&= \sum_x p(x)\overline{\mathbf{F}}(\rho_1^x, \ldots, \rho_r^x). && \text{(A.249)}
\end{aligned}
$$

The first inequality follows because $\widetilde{\Lambda}$ is a special kind of measurement. The third equality follows from the direct-sum property of the underlying classical multivariate fidelity.

For the opposite inequality, consider that

$$
\begin{aligned}
&\overline{\mathbf{F}}(\rho_{XA}^1, \ldots, \rho_{XA}^r) \\
&= \inf_{\Lambda} \mathbf{F}(\Lambda(\rho_{XA}^1), \ldots, \Lambda(\rho_{XA}^r)) && \text{(A.250)} \\
&= \inf_{\Lambda} \mathbf{F}\left(\sum_x p(x)\Lambda\left(|x\rangle\!\langle x| \otimes \rho_1^x\right), \ldots, \sum_x p(x)\Lambda\left(|x\rangle\!\langle x| \otimes \rho_r^x\right)\right) && \text{(A.251)} \\
&\geq \inf_{\Lambda} \sum_x p(x)\mathbf{F}(\Lambda\left(|x\rangle\!\langle x| \otimes \rho_1^x\right), \ldots, \Lambda\left(|x\rangle\!\langle x| \otimes \rho_r^x\right)) && \text{(A.252)} \\
&\geq \sum_x p(x)\overline{\mathbf{F}}(\rho_1^x, \ldots, \rho_r^x). && \text{(A.253)}
\end{aligned}
$$

The first equality follows from linearity of the measurement channel Λ. The first inequality follows because the underlying classical fidelities are jointly concave (as a consequence of the data-processing inequality and the direct-sum property). The final inequality follows because tensoring in the state $|x\rangle\!\langle x|$ and performing the measurement Λ is a particular kind of measurement channel for the tuple $(\rho_1^x, \ldots, \rho_r^x)$, and so the resulting fidelity cannot be smaller than $\overline{\mathbf{F}}(\rho_1^x, \ldots, \rho_r^x)$, which is defined as the infimum over all measurement channels.

By combining (A.249) and (A.253), we conclude the proof.

A.18 Proof of Proposition 5.40 (Properties of minimal extension of multivariate fidelity)

Recall

$$
\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) \coloneqq \sup_{\mathcal{P} \in \mathrm{CPTP}, \, \omega_i \in \mathscr{D}_c \forall i \in [r]} \left\{ \mathbf{F}(\omega_1, \ldots, \omega_r) : \mathcal{P}(\omega_i) = \rho_i \; \forall i \in [r] \right\}, \tag{A.254}
$$

where $\mathcal{P}$ is a classical-quantum channel with $\mathcal{P}(\omega_i) = \rho_i$.

Symmetry follows by the symmetry of the underlying classical multivariate fidelity.

Data processing: Let $\mathcal{N}$ be a quantum channel. Let $\omega'_i \in \mathscr{D}_c$ such that $\mathcal{P}'(\omega'_i) = \rho_i$. This also leads to $\mathcal{N}\big(\mathcal{P}'(\omega'_i)\big) = \mathcal{N}(\rho_i)$, providing a feasible preparation channel for $\mathcal{N}(\rho_i)$ resulted from $\mathcal{N} \circ \mathcal{P}'$. So we have

$$\mathbf{F}(\omega'_i, \ldots, \omega'_r) \leq \underline{\mathbf{F}}(\mathcal{N}(\rho_1), \ldots, \mathcal{N}(\rho_r)) \tag{A.255}$$

Then, supremizing over all $\mathcal{P}$ and ω_i, we obtain that

$$\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) \leq \underline{\mathbf{F}}(\mathcal{N}(\rho_1), \ldots, \mathcal{N}(\rho_r))\,. \tag{A.256}$$

Faithfulness: Let $\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = 1$. Then, we also have that $\mathbf{F}(\rho_1, \ldots, \rho_r) = 1$. With the faithfulness property of multivariate fidelity (e.g., average z-pairwise fidelity), we find that $\rho_i = \rho_j$ for all $i, j \in [r]$.

To prove the reverse implication, assume that $\rho_i = \rho_j = \rho$ for all $i, j \in [r]$. Then a simple preparation channel is trace and replace the deterministic state $\omega_i := |0\rangle\!\langle 0|$ with ρ. The fidelity of the common state is equal to one, due to the faithfulness of the classical multivariate fidelity. Since the prep fidelity involves a supremum over all preparations and the underlying fidelity cannot exceed one, this proves that $\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = 1$.

Direct-sum property: For all $i \in [r]$, define

$$\rho^i_{XA} := \sum_x p(x) |x\rangle\!\langle x| \otimes \rho^x_i, \tag{A.257}$$

where $(p(x))_x$ is a probability distribution, $\{|x\rangle\}_x$ is an orthonormal basis, and each ρ^x_i is a state.

Define the preparation channel $\mathcal{P}^x$ and the state $\omega^x_i \in \mathscr{D}_c$, such that

$$\mathcal{P}^x(\omega^x_i) = \rho^x_i \;\; \forall i \in [r]\,. \tag{A.258}$$

Defining

$$\mathcal{P}'(\tau_{XS}) := \sum_x |x\rangle\!\langle x| \otimes \mathcal{P}^x(\langle x|_X \tau_{XS} |x\rangle_X), \tag{A.259}$$

$$\omega'_i := \sum_x p(x) |x\rangle\!\langle x| \otimes \omega^x_i \tag{A.260}$$

we then find that, for all $i \in [r]$,

$$\mathcal{P}'\left(\omega'_i\right) = \sum_x p(x) |x\rangle\!\langle x| \otimes \mathcal{P}^x(\omega^x_i) = \rho^i_{XA}. \tag{A.261}$$

Then consider that

$$\underline{\mathbf{F}}(\rho^1_{XA}, \ldots, \rho^r_{XA})$$

$$= \sup_{\substack{\mathcal{P}, \\ \omega_1,\ldots,\omega_r \in \mathscr{D}_c}} \left\{ \mathbf{F}(\omega_1,\ldots,\omega_r) : \mathcal{P}(\omega_i) = \rho^i_{XA}\ \forall i \in [r] \right\} \tag{A.262}$$

$$\geq \sup_{\substack{\mathcal{P}', \\ \omega'_1,\ldots,\omega'_r \in \mathscr{D}_c}} \left\{ \mathbf{F}(\omega'_1,\ldots,\omega'_r) : \mathcal{P}'(\omega'_i) = \rho^i_{XA}\ \forall i \in [r] \right\} \tag{A.263}$$

$$= \sup_{\substack{\mathcal{P}', \\ \omega'_1,\ldots,\omega'_r \in \mathscr{D}_c}} \left\{ \begin{array}{c} \mathbf{F}(\sum_x p(x)|x\rangle\langle x| \otimes \omega^x_i, \ldots, \sum_x p(x)|x\rangle\langle x| \otimes \omega^x_i) : \\ \mathcal{P}'(\omega'_i) = \rho^i_{XA}\ \forall i \in [r] \end{array} \right\} \tag{A.264}$$

$$= \sup_{\substack{\mathcal{P}', \\ \omega'_1,\ldots,\omega'_r \in \mathscr{D}_c}} \left\{ \sum_x p(x) \mathbf{F}(\omega^x_i,\ldots,\omega^x_i) : \mathcal{P}'(\omega'_i) = \rho^i_{XA}\ \forall i \in [r] \right\} \tag{A.265}$$

$$= \sum_x p(x) \sup_{\substack{\mathcal{P}^x, \\ \omega^x_1,\ldots,\omega^x_r \in \mathscr{D}_c}} \left\{ \mathbf{F}(\omega^x_i,\ldots,\omega^x_i) : \mathcal{P}^x(\omega^x_i) = \rho^x_i\ \forall i \in [r] \right\} \tag{A.266}$$

$$= \sum_x p(x) \underline{\mathbf{F}}(\rho^x_1,\ldots,\rho^x_r). \tag{A.267}$$

Now we prove the following inequality

$$\underline{\mathbf{F}}(\rho^1_{XA},\ldots,\rho^r_{XA}) \leq \sum_x p(x) \underline{\mathbf{F}}(\rho^x_1,\ldots,\rho^x_r). \tag{A.268}$$

For a probability distribution $(q_i(y))_y$, define the state

$$\omega_i \coloneqq \sum_y q_i(y)|y\rangle\langle y|, \tag{A.269}$$

for all $i \in [r]$. Then consider that the action of a general preparation channel $\mathcal{P}$ is as follows

$$\mathcal{P}(\omega_i) = \rho^i_{XA} \quad \forall i \in [r]\,. \tag{A.270}$$

Now define

$$\mathcal{P}(|y\rangle\langle y|) = \sigma^y_{XA}, \tag{A.271}$$

so that

$$\sum_y q_i(y)\sigma^y_{XA} = \rho^i_{XA} \quad \forall i \in [r]\,. \tag{A.272}$$

Then from the action of the completely dephasing channel Δ, define the probability distribution $(r(x|y))_x$ and tuple $(\sigma^{x,y}_A)_x$ of states as

$$(\Delta \otimes \mathrm{id})\left(\sigma^y_{XA}\right) = \sum_x r(x|y)|x\rangle\langle x| \otimes \sigma^{x,y}_A, \tag{A.273}$$

and consider that

$$\rho^i_{XA} = (\Delta \otimes \mathrm{id})\left(\rho^i_{XA}\right) \tag{A.274}$$

$$= \sum_y q_i(y) \left(\Delta \otimes \mathrm{id}\right) \left(\sigma^y_{XA}\right) \tag{A.275}$$

$$= \sum_{x,y} q_i(y) r(x|y) |x\rangle\!\langle x| \otimes \sigma^{x,y}_A. \tag{A.276}$$

Now observe that

$$\sum_x p(x) |x\rangle\!\langle x| = \mathrm{Tr}_A[\rho^i_{XA}] = \sum_x \left(\sum_y q_i(y) r(x|y) \right) |x\rangle\!\langle x|, \tag{A.277}$$

proving that the marginal over y of the joint distribution $q_i(y)r(x|y)$ is $p(x)$. Now let us define the conditional distribution $s_i(y|x)$ by

$$s_i(y|x)p(x) = q_i(y)r(x|y), \tag{A.278}$$

and we note that

$$\rho^i_{XA} = \sum_{x,y} s_i(y|x)p(x)|x\rangle\!\langle x| \otimes \sigma^{x,y}_A \tag{A.279}$$

$$= \sum_x p(x)|x\rangle\!\langle x| \otimes \sum_y s_i(y|x)\sigma^{x,y}_A \tag{A.280}$$

$$= \sum_x p(x)|x\rangle\!\langle x| \otimes \rho^x_i. \tag{A.281}$$

Thus, conditioned on x, generating y at random according to $s_i(y|x)$ and preparing the state $\sigma^{x,y}_A$ is a particular way of preparing ρ^x_i. That is, we can define the preparation channel $\mathcal{P}^x$ as

$$\mathcal{P}^x(|y\rangle\!\langle y|) = \sigma^{x,y}_A, \tag{A.282}$$

so that, for all $i \in [r]$,

$$\mathcal{P}^x \left(\sum_y s_i(y|x)|y\rangle\!\langle y| \right) = \sum_y s_i(y|x)\mathcal{P}^x\left(|y\rangle\!\langle y|\right) = \sum_y s_i(y|x)\sigma^{x,y}_A = \rho^x_i. \tag{A.283}$$

Putting everything together, let $\mathcal{P}$ and $\omega_1, \ldots, \omega_r$ be a particular way of preparing $\rho^1_{XA}, \ldots, \rho^r_{XA}$, respectively. Then

$$\mathbf{F}(\omega_1, \ldots, \omega_r)$$

$$= \mathbf{F}\left(\sum_y q_1(y)|y\rangle\!\langle y|, \ldots, \sum_y q_r(y)|y\rangle\!\langle y| \right) \tag{A.284}$$

$$\leq \mathbf{F}\left(\sum_{x,y} q_1(y)r(x|y)|x\rangle\!\langle x| \otimes |y\rangle\!\langle y|, \ldots, \sum_{x,y} q_r(y)r(x|y)|x\rangle\!\langle x| \otimes |y\rangle\!\langle y| \right) \tag{A.285}$$

$$= \mathbf{F}\left(\sum_{x,y} s_1(y|x)p(x)|x\rangle\langle x| \otimes |y\rangle\langle y|, \ldots, \sum_{x,y} s_1(y|x)p(x)|x\rangle\langle x| \otimes |y\rangle\langle y|\right) \tag{A.286}$$

$$= \sum_x p(x)\mathbf{F}\left(\sum_y s_1(y|x)|y\rangle\langle y|, \ldots, \sum_{x,y} s_1(y|x)|y\rangle\langle y|\right) \tag{A.287}$$

$$\leq \sum_x p(x) \sup_{\substack{\mathcal{P}^x, \\ \omega_1^x,\ldots,\omega_r^x \in \mathscr{D}_c}} \{\mathbf{F}(\omega_i^x, \ldots, \omega_i^x) : \mathcal{P}^x(\omega_i^x) = \rho_i^x \ \forall i \in [r]\} \tag{A.288}$$

$$= \sum_x p(x)\underline{\mathbf{F}}(\rho_1^x, \ldots, \rho_r^x). \tag{A.289}$$

The first inequality follows from data processing. The second equality follows from the identity in (A.278). The third equality follows from the direct-sum property for the underlying classical multivariate fidelity. The last inequality follows because, from $\mathcal{P}$ and $\omega_1, \ldots, \omega_r$, we have constructed a particular method of preparing $\rho_1^x, \ldots, \rho_r^x$ from the channel $\mathcal{P}^x$ and the commuting states $\sum_y s_1(y|x)|y\rangle\langle y|, \ldots, \sum_y s_r(y|x)|y\rangle\langle y|$, as given in (A.283). The proof of the inequality in (A.268) is complete after noticing that we have proven the inequality $\mathbf{F}(\omega_1, \ldots, \omega_r) \leq \sum_x p(x)\underline{\mathbf{F}}(\rho_1^x, \ldots, \rho_r^x)$ for all possible preparations of the states $\rho_{XA}^1, \ldots, \rho_{XA}^r$, and so we can take a supremum over all such preparations.

Weak orthogonality: Suppose that $\rho_i\rho_j = 0$ for all $i \neq j \in [r]$. Then, by the orthogonality of multivariate fidelity that generalizes the underlying classical multivariate fidelity (e.g., average pairwise z-fidelity that generalizes average pairwise fidelity and multivariate log-Euclidean fidelity that generalizes Matusita multivariate fidelity), we have

$$\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) \leq \mathbf{F}(\rho_1, \ldots, \rho_r) = 0, \tag{A.290}$$

which implies that $\underline{\mathbf{F}}(\rho_1, \ldots, \rho_r) = 0$ concluding the proof.

A.19 Proof of Theorem 5.45 (Properties of multivariate log-Euclidean fidelity)

Symmetry follows by the definition of oveloh information in (5.109).

Reduction to classical Matusita multivariate fidelity: By (5.110), we have

$$\inf_{\sigma \in \mathscr{D}} \frac{1}{r}\sum_{i=1}^r D(\sigma\|\rho_i) = -\ln \operatorname{Tr}\left[\exp\left(\sum_{i=1}^r \frac{1}{r}\ln \rho_i\right)\right]. \tag{A.291}$$

Then, when the states are commuting, we also have the following identity:

$$\exp\left(\sum_{i=1}^r \frac{1}{r}\ln \rho_i\right) = \prod_{i=1}^r \rho_i^{\frac{1}{r}}. \tag{A.292}$$

Combining the above two identitites, we arrive at

$$\exp\left[-\inf_{\sigma\in\mathscr{D}} D(\rho_X\otimes\sigma\|\rho_{XA})\right] = \exp\left[-\inf_{\sigma\in\mathscr{D}} \frac{1}{r}\sum_{i=1}^{r} D(\sigma\|\rho_i)\right] \tag{A.293}$$

$$= \operatorname{Tr}\left[\prod_{i=1}^{r} \rho_i^{\frac{1}{r}}\right]. \tag{A.294}$$

Observing that the last expression is precisely the Matusita multivariate fidelity for commuting states then establishes the claim.

Data processing: Consider an arbitrary $\sigma\in\mathscr{D}$, due to the data processing of quantum relative entropy, we have

$$D(\rho_X\otimes\sigma\|\rho_{XA}) \geq D(\rho_X\otimes\mathcal{N}(\sigma)\|\operatorname{id}\otimes\mathcal{N}(\rho_{XA}))\,, \tag{A.295}$$

where id is the identity channel and $\mathcal{N}:\mathscr{L}(\mathcal{H}_A)\to\mathscr{L}(\mathcal{H}_B)$ is a quantum channel. Note that

$$(\operatorname{id}\otimes\mathcal{N})(\rho_{XA}) = \frac{1}{r}\sum_{i=1}^{r} |i\rangle\!\langle i|_X\otimes\mathcal{N}(\rho_i). \tag{A.296}$$

So, we arrive at

$$\inf_{\sigma\in\mathscr{D}}\frac{1}{r}\sum_{i=1}^{r} D(\sigma\|\rho_i) \geq \inf_{\sigma\in\mathscr{D}}\frac{1}{r}\sum_{i=1}^{r} D(\mathcal{N}(\sigma)\|\mathcal{N}(\rho_i))$$

$$\geq \inf_{\sigma'\in\mathscr{D}}\frac{1}{r}\sum_{i=1}^{r} D(\sigma'\|\mathcal{N}(\rho_i))\,, \tag{A.297}$$

where the last inequality follows from optimization over a larger set of states. This then leads to

$$\exp\left(-\inf_{\sigma\in\mathscr{D}}\frac{1}{r}\sum_{i=1}^{r} D(\sigma\|\rho_i)\right) \leq \exp\left(-\inf_{\sigma'\in\mathscr{D}}\frac{1}{r}\sum_{i=1}^{r} D(\sigma'\|\mathcal{N}(\rho_i))\right), \tag{A.298}$$

due to the monotonicity of the exponential function.

Faithfulness: If $\rho_i=\rho$ for all $i\in[r]$, then for arbitrary $\sigma\in\mathscr{D}$

$$D(\rho_X\otimes\sigma\|\rho_{XA}) = D(\rho_X\otimes\sigma\|\rho_X\otimes\rho) \geq 0. \tag{A.299}$$

Then $D(\rho_X\otimes\sigma\|\rho_{XA})=0$ by the choice $\sigma=\rho$. With that choice, infimum is achieved leading to $\exp\left[-\mathscr{O}(X;A)_\rho\right]=\exp(0)=1$.

To prove the reverse direction, assume that $\exp\left[-\mathscr{O}(X;A)_\rho\right]=1$, so that we have

$$\inf_{\sigma\in\mathscr{D}}\frac{1}{r}\sum_{i=1}^{r} D(\sigma\|\rho_i) = 0 \tag{A.300}$$

Since quantum relative entropy is non-negative, there exists $\sigma \in \mathscr{D}$ such that $D(\sigma\|\rho_i) = 0$ for all $i \in [r]$. Then, faithfulness of quantum relative entropy [KW20, Proposition 7.3] implies that $\sigma = \rho_i$ for all i, which is the desired implication.

Direct-sum property: Let

$$\rho_i \coloneqq \sum_x p(x)|x\rangle\!\langle x|_X \otimes \rho_i^x \tag{A.301}$$

for all $i \in [r]$. Then, by the spectral theorem, we have

$$\ln \rho_i = \sum_x |x\rangle\!\langle x|_X \otimes \ln(p(x)\rho_i^x) = \sum_x |x\rangle\!\langle x|_X \otimes \ln(p(x))\, I + \sum_x |x\rangle\!\langle x|_X \otimes \ln(\rho_i^x)\,. \tag{A.302}$$

Summing over all $i \in [r]$ and averaging, we get

$$\sum_{i=1}^r \frac{1}{r} \ln \rho_i = \sum_x |x\rangle\!\langle x|_X \otimes \ln(p(x))\, I + \sum_x |x\rangle\!\langle x|_X \otimes \sum_{i=1}^r \frac{1}{r} \ln(\rho_i^x)\,. \tag{A.303}$$

Using the fact that the two summands on the right above commute, that $\mathrm{Tr}[\exp(A+B)] = \mathrm{Tr}[\exp(A)\exp(B)]$ for commuting operators A and B, and again applying the spectral theorem, we arrive at

$$\begin{aligned}
&\mathrm{Tr}\left[\exp\left(\sum_{i=1}^r \frac{1}{r}\ln \rho_i\right)\right] \\
&= \mathrm{Tr}\left[\exp\left(\sum_x |x\rangle\!\langle x|_X \otimes \ln(p(x)) I\right) \exp\left(\sum_{x'} |x'\rangle\!\langle x'|_X \otimes \sum_{i=1}^r \frac{1}{r}\ln\left(\rho_i^{x'}\right)\right)\right] && \text{(A.304)}\\
&= \mathrm{Tr}\left[\left(\sum_x |x\rangle\!\langle x|_X \otimes p(x) I\right)\left(\sum_{x'} |x'\rangle\!\langle x'|_X \otimes \exp\left(\sum_{i=1}^r \frac{1}{r}\ln\left(\rho_i^{x'}\right)\right)\right)\right] && \text{(A.305)}\\
&= \mathrm{Tr}\left[\sum_x |x\rangle\!\langle x|_X \otimes p(x)\exp\left(\sum_{i=1}^r \frac{1}{r}\ln(\rho_i^x)\right)\right] && \text{(A.306)}\\
&= \sum_x p(x)\mathrm{Tr}\left[\exp\left(\sum_{i=1}^r \frac{1}{r}\ln(\rho_i^x)\right)\right]. && \text{(A.307)}
\end{aligned}$$

Recalling (5.110) together with Definition 5.43, we conclude the proof.

Weak orthogonality: Let $\rho_i\rho_j = 0$ for all $i \neq j \in [r]$. Then, the channel mapping $\rho_i \to |i\rangle\!\langle i|$ for all $i \in [r]$ is a reversible mapping. Then, due to the data processing property of log-Euclidean multivariate fidelity, we have

$$F_r^\flat(\rho_1, \ldots, \rho_r) = F_r^\flat(|1\rangle\!\langle 1|, \ldots, |r\rangle\!\langle r|). \tag{A.308}$$

For $\inf_\sigma \frac{1}{r}\sum_{i=1}^r D(\sigma\||i\rangle\!\langle i|) < \infty$, there should be $\sigma \in \mathscr{D}$ such that $D(\sigma\||i\rangle\!\langle i|) < \infty$ for all $i \in [r]$. This indeed happens only if $\mathrm{supp}(\sigma) \subseteq \mathrm{supp}(|i\rangle\!\langle i|)$ for all i. Since, $\{|i\rangle\!\langle i|\}$ is a collection of orthonormal projections, we cannot find σ such that it satisfies the finiteness condition for all $i \in [r]$. This implies that

$$\inf_{\sigma} \frac{1}{r} \sum_{i=1}^{r} D(\sigma \| |i\rangle\langle i|) = \infty. \tag{A.309}$$

Then, by recalling Eq. (5.110), we conclude that $F_r^{\flat}(\rho_1, \ldots, \rho_r) = 0$.

Continuity: The multivariate log-Euclidean fidelity $F_\flat$ given in Definition 5.42 is continuous at tuples $(\rho_1, \ldots, \rho_r)$ of invertible states. This follows from the fact that the matrix exponential and logarithmic functions are continuous at positive definite operators, which implies that the multivariate log-Euclidean fidelity is a composition of continuous functions.

A.20 Proof of Proposition 5.50 (Inequalities relating average k-wise log-Euclidean fidelities)

Let $S_{k-1}(i_1, \ldots, i_k)$ denote all size-$(k-1)$ subsets of $i_1, \ldots, i_k$, of which there are k of them. Note that each symbol i_j appears $k-1$ times in each subset of $S_{k-1}(i_1, \ldots, i_k)$. Then consider that

$$\begin{aligned}
&\frac{1}{k} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} F_{k-1}^{\flat}(\rho_{t(1)}, \ldots, \rho_{t(k-1)}) \\
&= \frac{1}{k} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} \exp\left(- \inf_{\omega \in \mathscr{D}} \frac{1}{k-1} \sum_{j \in [k-1]} D(\omega \| \rho_{t(j)}) \right) && \text{(A.310)} \\
&\geq \exp\left(-\frac{1}{k} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} \inf_{\omega \in \mathscr{D}} \frac{1}{k-1} \sum_{j \in [k-1]} D(\omega \| \rho_{t(j)}) \right) && \text{(A.311)} \\
&\geq \exp\left(- \inf_{\omega \in \mathscr{D}} \frac{1}{k(k-1)} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} \sum_{j \in [k-1]} D(\omega \| \rho_{t(j)}) \right) && \text{(A.312)} \\
&= \exp\left(- \inf_{\omega \in \mathscr{D}} \frac{1}{k} \sum_{\ell \in \{i_1, \ldots, i_k\}} D(\omega \| \rho_\ell) \right) && \text{(A.313)} \\
&= F_k^{\flat}(\rho_{i_1}, \ldots, \rho_{i_k}), && \text{(A.314)}
\end{aligned}$$

where the first inequality follows from the convexity of the exponential function and the second from the fact that

$$\frac{1}{k} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} \inf_{\omega \in \mathscr{D}} \frac{1}{k-1} \sum_{j \in [k-1]} D(\omega \| \rho_{t(j)}) \leq \inf_{\omega \in \mathscr{D}} \frac{1}{k(k-1)} \sum_{t \in S_{k-1}(i_1, \ldots, i_k)} \sum_{j \in [k-1]} D(\omega \| \rho_{t(j)}). \tag{A.315}$$

The penultimate equality follows because

$$\frac{1}{k\,(k-1)} \sum_{t\in S_{k-1}(i_1,\ldots,i_k)} \sum_{j\in[k-1]} D(\omega\|\rho_{t(j)}) = \frac{1}{k} \sum_{\ell\in\{i_1,\ldots,i_k\}} D(\omega\|\rho_\ell), \tag{A.316}$$

as a generalization of the justification behind (A.22). Now consider that

$$\frac{1}{\binom{r}{k}} \sum_{i_1<\cdots<i_k} F_k^\flat(\rho_{i_1},\ldots,\rho_{i_k})$$

$$\leq \frac{1}{\binom{r}{k}} \sum_{i_1<\cdots<i_k} \frac{1}{k} \sum_{t\in S_{k-1}(i_1,\ldots,i_k)} F_{k-1}^\flat(\rho_{t(1)},\ldots,\rho_{t(k-1)}) \tag{A.317}$$

$$= \frac{k-1!r-k!}{r!} \sum_{i_1<\cdots<i_k} \sum_{t\in S_{k-1}(i_1,\ldots,i_k)} F_{k-1}^\flat(\rho_{t(1)},\ldots,\rho_{t(k-1)}) \tag{A.318}$$

$$= \frac{k-1!r-k!}{r!}\,(r-k+1) \sum_{i_1<\cdots<i_{k-1}} F_{k-1}^\flat(\rho_{i_1},\ldots,\rho_{i_{k-1}}) \tag{A.319}$$

$$= \frac{1}{\binom{r}{k-1}} \sum_{i_1<\cdots<i_{k-1}} F_{k-1}^\flat(\rho_{i_1},\ldots,\rho_{i_{k-1}}). \tag{A.320}$$

The reasoning for these steps is precisely the same as that given for (A.26)–(A.30).