Uhlmann's Theorem for Measured Divergences

Kun Fang, *Member, IEEE*, Hamza Fawzi, and Omar Fawzi

***Abstract*—Uhlmann's theorem is a cornerstone of quantum information theory, stating that for any quantum state ρ_{AB} and any state σ_A, there exists an extension σ_{AB} of σ_A such that the fidelity between ρ_{AB} and σ_{AB} equals the fidelity between their marginals ρ_A and σ_A. This property underpins many results and applications in quantum information science. In this work, we generalize Uhlmann's theorem to a broad class of measured f-divergences, including the measured α-Rényi divergences for all $\alpha \geq 0$. The well-known Uhlmann's theorem for the fidelity corresponds to the special case $\alpha = 1/2$. Since most commonly used quantum Rényi divergences, including the Petz and sandwiched Rényi divergences, cannot satisfy this property (except for degenerate cases), this fundamentally distinguishes measured f-divergences from other quantum divergences and highlights their unique mathematical structure.**

***Index Terms*—Uhlmann's theorem, measured f-divergences, quantum information theory, Rényi divergences.**

I. Introduction

FOR a convex function $f : \mathbb{R}_{>0} \to \mathbb{R}$, the f-divergence between two discrete probability vectors P and Q is defined as[1]

$$S_f(P\|Q) := \sum_x Q(x) f\left(\frac{P(x)}{Q(x)}\right). \tag{1}$$

Received 8 March 2025; revised 26 November 2025; accepted 20 December 2025. Date of publication 29 December 2025; date of current version 19 February 2026. The work of Kun Fang was supported in part by the National Natural Science Foundation of China under Grant 12404569 and Grant 92470113, in part by Shenzhen Science and Technology Program under Grant JCYJ20240813113519025, in part by Shenzhen Fundamental Research Program under Grant JCYJ20241202124023031, in part by the General R&D Projects of 1+1+1 CUHK-CUHK(SZ)-GDST Joint Collaboration Fund under Grant GRDP2025-022, and in part by the University Development Fund under Grant UDF01003565. The work of Hamza Fawzi was supported in part by U.K. Research and Innovation (UKRI) under U.K. Government's Horizon Europe Funding Guarantee under Grant EP/X032051/1. The work of Omar Fawzi was supported in part by European Research Council (ERC) Grant AlgoQIP under Agreement 851716, in part by European Union's Horizon 2020 Research and Innovation Programme under Grant 101017733 (VERIqTAS), and in part by the Agence Nationale de la Recherche through the Plan France 2030 under Grant ANR-22-PETQ-0009. *(Corresponding author: Kun Fang.)*

Kun Fang is with the School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China (e-mail: kunfang@cuhk.edu.cn).

Hamza Fawzi is with the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, CB3 0WA Cambridge, U.K. (e-mail: h.fawzi@damtp.cam.ac.uk).

Omar Fawzi is with Inria, ENS de Lyon, UCBL, LIP, 69342 Lyon Cedex 07, France (e-mail: omar.fawzi@ens-lyon.fr).

Communicated by E. Chitambar, Associate Editor for Quantum.

Digital Object Identifier 10.1109/TIT.2025.3649040

[1]We make the following standard conventions to define this expression: $f(0) = \lim_{t\downarrow 0} f(t), 0f\left(\frac{0}{0}\right) = 0$ and $0f\left(\frac{a}{0}\right) = \lim_{t\downarrow 0} tf\left(\frac{a}{t}\right)$ for $a > 0$. Note that the limits may be $+\infty$.

This family already appears in Rényi's work [1] and was further studied by Csiszár [2] and Ali and Silvey [3]. Important examples include the Kullback-Leibler divergence (also called relative entropy) $D(\cdot\|\cdot)$ corresponding to $f_1(t) = t\log t$ with $D(P\|Q) = S_{f_1}(P\|Q)$ and the α-Rényi divergences $D_\alpha(\cdot\|\cdot)$ corresponding to $f_\alpha(t) = \operatorname{sign}(\alpha-1)t^\alpha$ with $D_\alpha(P\|Q) = \frac{1}{\alpha-1}\log\left(\operatorname{sign}(\alpha-1)S_{f_\alpha}(P\|Q)\right)$. Such divergences play an important role in various areas of statistics and information theory, we refer to [4, Chapter 7] for an overview.

For quantum states modeled by positive semidefinite operators ρ and σ, there are multiple ways of extending (1); see [5] for a systematic exposition. The main focus in this paper is on the *measured* f-divergence [6], [7] which is obtained by performing a measurement M on the states ρ and σ and computing the f-divergence between the resulting probability distributions $P_{\rho,M}$ and $P_{\sigma,M}$. We then take the supremum over all possible measurements or all projective measurements:

$$S_{\mathrm{M},f}(\rho\|\sigma) := \sup_{M \text{POVM}} S_f(P_{\rho,M}\|P_{\sigma,M}), \tag{2}$$

$$S^{\mathrm{P}}_{\mathrm{M},f}(\rho\|\sigma) := \sup_{M \text{PVM}} S_f(P_{\rho,M}\|P_{\sigma,M}), \tag{3}$$

where in (2) the supremum is over positive operator-valued measures (POVM) M described by a discrete set $\mathcal{X}$ and positive semidefinite operators $\{M_x\}_{x\in\mathcal{X}}$ satisfying $\sum_{x\in\mathcal{X}} M_x = I$, the supremum in (3) is over projection-valued measures (PVM) M, i.e., POVMs satisfying in addition that M_x is an orthogonal projector for all $x \in \mathcal{X}$ and $P_{\rho,M}(x) = \operatorname{Tr}[\rho M_x]$.

A. Main Results

We consider the Uhlmann property that is well known for the fidelity [8]: given a quantum state ρ_A on A and a state σ_{AR} on $A \otimes R$, is there an extension ρ_{AR} of ρ_A such that $S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) = S_{\mathrm{M},f}(\rho_A\|\sigma_A)$? We show in Theorem 2 that if f^* is operator convex and operator monotone and the domain of f^* is unbounded from below, then Uhlmann's property holds. Similarly, if $(f^*)^{-1}$ is operator concave and operator monotone and the range of f^* is unbounded from above, then Uhlmann's property holds with the roles of ρ and σ exchanged. In the special case of Rényi divergences, this shows the Uhlmann property for all $\alpha \geq 0$ (Corollary 1). This was previously known for $\alpha = \frac{1}{2}$ [8] and in the limit $\alpha \to \infty$ [9]. Our result can also be used to recover the regularized Uhlmann's theorem of [9]. We conclude this paper in Section IV with an intriguing "duality" property (Proposition 2) relating the Umegaki relative entropy $D(\rho\|\mathscr{C})$ between ρ and a compact convex set $\mathscr{C}$ and the measured relative entropy $D_{\mathrm{M}}(\rho\|\mathscr{C}^\circ_{++})$ between ρ and the corresponding polar set $\mathscr{C}^\circ_{++}$. A similar duality holds between $D(\rho\|\mathscr{C}^\circ_{++})$ and

$D_{\mathrm{M}}(\rho\|\mathscr{C})$. These duality relations provide a clearer understanding of why the measured relative entropy to a set of quantum states becomes superadditive when the polar sets are closed under tensor products. This superadditivity property is fundamental for several recent applications, including [10], [11], [12], [13]. The significance and some potential applications of our results are discussed in Section V.

B. Notation

For a finite-dimensional Hilbert space A, we let $\mathscr{H}(A)$ denote the space of Hermitian operators on A, $\mathscr{H}_+(A) = \{\omega \in \mathscr{H}(A) : \omega \geq 0\}$ the set of positive semidefinite operators and $\mathscr{H}_{++}(A) = \{\omega \in \mathscr{H}(A) : \omega > 0\}$ the set of positive definite operators. The set of density operators is denoted by $\mathscr{D}(A) = \{\rho \in \mathscr{H}_+(A) : \mathrm{Tr}[\rho] = 1\}$. We drop the Hilbert space from the notation when it is clear from the context. For $\omega \in \mathscr{H}$, $\mathrm{spec}(\omega)$ denotes the spectrum of ω. The measured Rényi divergences for $\alpha \geq 0$ are defined as

$$D_{\mathrm{M},\alpha}(\rho\|\sigma) = \frac{1}{\alpha-1}\log\left(\mathrm{sign}(\alpha-1)S_{\mathrm{M},f_\alpha}(\rho\|\sigma)\right), \quad (4)$$

for $\alpha \neq 1$, and

$$D_{\mathrm{M}}(\rho\|\sigma) = S_{\mathrm{M},f_1}(\rho\|\sigma), \quad (5)$$

for $\alpha = 1$. In the limit $\alpha \to \infty$, it is known that $D_{\mathrm{M},\alpha} \to D_{\max}$, see e.g., [14, Appendix A].

1) Independent Work: We note that in independent and concurrent work, Mazzola et al. [15] established a regularized Uhlmann's theorem for the sandwiched Rényi divergence for $\alpha \geq \frac{1}{2}$ as well as an Uhlmann inequality for the measured Rényi divergences i.e., they show the existence of an extension satisfying $D_{\mathrm{M},\alpha}(\rho_{AR}\|\sigma_{AR})$ is upper bounded by the corresponding sandwiched Rényi divergence.

II. Variational Expression for Measured Divergences

In the classical setting, variational expressions for f-divergences play an important role (see [4, Chapter 7]), but they are also fundamental in the quantum case in particular for the measured divergences as they are at the core of many applications of such divergences; see e.g., [10], [16], [17]. In this section, we discuss the variational expression for measured f-divergences, which forms an important starting point for finding the Uhlmann's theorem for measured divergences in the next section. These variational expressions have been studied in the von Neumann algebra setting in [18, Chapter 5]. Here, we present the results and their proofs in the finite-dimensional setting for completeness and for accessiblility to general readers.

In order to state the variational expression, we need to introduce the Fenchel conjugate of a function. For a function $f : \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$, the Fenchel conjugate (also called Legendre transform) is defined as [19]

$$f^*(z) := \sup_{t\in\mathrm{dom}(f)} \{tz - f(t)\}, \quad (6)$$

where $\mathrm{dom}(f) = \{t \in \mathbb{R} : f(t) < +\infty\}$. The domain of f^* is defined in the same way as

$$\mathrm{dom}(f^*) = \{z \in \mathbb{R} : f^*(z) < +\infty\}. \quad (7)$$

As f should define a divergence, we assume throughout this paper that $(0,\infty) \subset \mathrm{dom}(f)$ and that f is convex. We assume in addition that f is lower semicontinuous. Note that the f-divergence only depends on the values f takes on $(0,\infty)$, but the Fenchel conjugate (and hence the variational expressions) depends on the values of f on $\mathrm{dom}(f)$.

We present the following general variational formula for the projective measured f-divergence under the sole condition that f is convex and lower semicontinuous. This result corresponds to [18, Theorem 5.7]. It can also be seen as a quantum analog of [4, Theorem 7.26].

Proposition 1: Let $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a convex and lower semicontinuous function and $(0,\infty) \subset \mathrm{dom}(f)$. For any $\rho, \sigma \in \mathscr{H}_+$,

$$S^{\mathrm{P}}_{\mathrm{M},f}(\rho\|\sigma) = \sup_{\omega\in\mathscr{H},\ \mathrm{spec}(\omega)\subset\mathrm{dom}(f^*)} \mathrm{Tr}[\rho\omega] - \mathrm{Tr}[\sigma f^*(\omega)], \quad (8)$$

where f^* is the Fenchel conjugate of f as defined in (6).

Proof: Since f is convex and lower semicontinuous, we know that $f = (f^*)^*$ [20, Theorem 12.2], i.e., for any $t \in \mathbb{R}$

$$f(t) = \sup_{z\in\mathrm{dom}(f^*)} \{tz - f^*(z)\}. \quad (9)$$

Hence for any $r, s > 0$ we get

$$sf(r/s) = \sup_{z\in\mathrm{dom}(f^*)} \{rz - sf^*(z)\}. \quad (10)$$

This expression also matches the conventions we took when $r = 0$ or $s = 0$. In fact, if $r = 0, s = 1$, we use (9) to get $f(0) = \sup_{z\in\mathrm{dom}(f^*)}\{-f^*(z)\}$. If $r = s = 0$, both sides are equal to 0. For $r > 0$ and $s = 0$, we need to show that $\lim_{s\downarrow 0} sf(r/s) = \sup_{z\in\mathrm{dom}(f^*)}\{rz\}$. By performing the change of variable $x = r/s$, this is equivalent to showing that $\lim_{x\to\infty}\frac{f(x)}{x} = \sup\mathrm{dom}(f^*)$. To establish this equality, we prove the two inequalities separately. We start with ($\geq$): we have $\frac{f(x)}{x} = \sup_{z\in\mathrm{dom}(f^*)}\left\{z - \frac{f^*(z)}{x}\right\} \geq z - \frac{f^*(z)}{x}$ for any $z \in \mathrm{dom}(f^*)$. Thus, $\lim_{x\to\infty}\frac{f(x)}{x} \geq \sup\mathrm{dom}(f^*)$. Now for ($\leq$), let $w < \lim_{x\to\infty}\frac{f(x)}{x}$ and let $x_n \geq n$ be such that $f(x) \geq wx$ for all $x \geq x_n$ (n is an integer $n \geq 2$). Let y_n be any subderivative of f at x_n. By definition, for any $x \in \mathbb{R}$, $f(x) \geq f(x_n)+y_n(x-x_n)$. This implies that $y_n x - f(x) \leq y_n x_n - f(x_n)$ and thus $f^*(y_n) \leq y_n x_n - f(x_n) < +\infty$ so $y_n \in \mathrm{dom}(f^*)$. In addition, we have $y_n \geq \frac{f(x_n)-f(1)}{x_n-1} \geq w\frac{x_n}{x_n-1} - \frac{f(1)}{x_n-1}$ and $1 \in \mathrm{dom}(f)$ by assumption. As a result, for any $\epsilon > 0$, there exists an n such that we obtain $y_n \geq w-\epsilon$. Thus $\sup\mathrm{dom}(f^*) \geq w$ which concludes the second inequality.[2] Plugging into (3), we get

$$\begin{aligned} &S^{\mathrm{P}}_{\mathrm{M},f}(\rho\|\sigma) \\ &= \sup_{\{M_x\}\mathrm{PVM}} \sum_x \mathrm{Tr}[M_x\sigma] f\left(\frac{\mathrm{Tr}[M_x\rho]}{\mathrm{Tr}[M_x\sigma]}\right) \quad (11)\\ &= \sup_{\{M_x\}\mathrm{PVM}} \sum_x \sup_{z_x\in\mathrm{dom}(f^*)} \mathrm{Tr}[M_x\rho]z_x - \mathrm{Tr}[M_x\sigma]f^*(z_x) \quad (12)\\ &= \sup_{\{M_x\}\mathrm{PVM}} \sup_{\{z_x\}} \sum_x \mathrm{Tr}[z_x M_x\rho] - \mathrm{Tr}[f^*(z_x)M_x\sigma] \quad (13) \end{aligned}$$

[2]We thank an anonymous reviewer for suggesting this argument.

TABLE I
CONVEX FUNCTIONS f AND THEIR FENCHEL CONJUGATES

$f(t)$	$\mathrm{dom}(f)$	$f^*(z)$	$\mathrm{dom}(f^*)$	Props. of f^* on $\mathrm{dom}(f^*)$	Props. of $(f^*)^{-1}$ on $\mathrm{range}(f^*)$
$-t^\alpha$ $(\alpha \in [0,1))$	$\mathbb{R}_{>0}$	$(1-\alpha)\left(\frac{-z}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}$	$\mathbb{R}_{<0}$	$\alpha \in [0,\frac{1}{2}]$: op. convex and monotone	$\alpha \in [\frac{1}{2},1)$: op. concave and monotone
$\begin{cases} t^\alpha & t \geq 0 \\ 0 & t < 0 \end{cases}$ $(\alpha > 1)$	$\mathbb{R}$	$(\alpha-1)\left(\frac{z}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}$	$\mathbb{R}_{>0}$	$\alpha \geq 2$: op. convex	$\alpha > 1$: op. concave and monotone
$t\log(t)$	$\mathbb{R}_{>0}$	e^{z-1}	$\mathbb{R}$	Not applicable	Op. concave and monotone
$\frac{1}{2}\lvert t-1\rvert$	$\mathbb{R}$	z	$[-\frac{1}{2},\frac{1}{2}]$	Linear monotone	Linear monotone

$$= \sup_{\{M_x\}\mathrm{PVM}} \sup_{\{z_x\}} \mathrm{Tr}\left[\sum_x z_x M_x \rho\right] - \mathrm{Tr}\left[f^*\left(\sum_x z_x M_x\right)\sigma\right] \tag{14}$$

$$= \sup_{\omega\in\mathscr{H},\mathrm{spec}(\omega)\subset\mathrm{dom}(f^*)} \mathrm{Tr}[\rho\omega] - \mathrm{Tr}[\sigma f^*(\omega)], \tag{15}$$

where everywhere $\{M_x\}$ is a projective measurement and $z_x \in \mathrm{dom}(f^*)$. Note that in the third line we used the fact that M_x are orthogonal projectors. 0 ■

Table I shows the Fenchel conjugate of some common functions.

Remark 1: We can recover Rényi divergences for $\alpha \in [0,1)$ by choosing

$$f_\alpha(t) = -t^\alpha, \qquad \mathrm{dom}(f_\alpha) = \mathbb{R}_{>0} \tag{16}$$

whose Fenchel conjugate is given in Table I. In this case, the variational expression (8) becomes:

$$S^{\mathrm{P}}_{\mathrm{M},f_\alpha}(\rho\|\sigma) = \sup_{\omega<0} \mathrm{Tr}[\rho\omega] - (1-\alpha)\mathrm{Tr}\left[\sigma\left(-\frac{\omega}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}\right], \tag{17}$$

which becomes after the change of variable $\gamma = -\frac{\omega}{\alpha}$:

$$S^{\mathrm{P}}_{\mathrm{M},f_\alpha}(\rho\|\sigma) = \sup_{\gamma>0} -\alpha\mathrm{Tr}[\rho\gamma] - (1-\alpha)\mathrm{Tr}\left[\sigma\gamma^{-\frac{\alpha}{1-\alpha}}\right]. \tag{18}$$

Then, since $D_{\mathrm{M},\alpha}(\rho\|\sigma) = \frac{1}{\alpha-1}\log\left(-S^{\mathrm{P}}_{\mathrm{M},f_\alpha}(\rho\|\sigma)\right)$, we get the variational expression established in [16] for $\alpha \in (0,1)$ (up to a simple change of variable that is described in Remark 3). Note that for $\alpha = 0$, we interpret γ^0 as the projector on the support of γ.

When $\alpha > 1$, we consider the function[3]

$$f_\alpha(t) = \begin{cases} t^\alpha & t \geq 0 \\ 0 & t < 0, \end{cases} \qquad \mathrm{dom}(f_\alpha) = \mathbb{R} \tag{19}$$

whose Fenchel conjugate is given in Table I. The expression (8) gives

$$S^{\mathrm{P}}_{\mathrm{M},f_\alpha}(\rho\|\sigma) = \sup_{\omega>0} \alpha\mathrm{Tr}[\rho\omega] + (1-\alpha)\mathrm{Tr}\left[\sigma\omega^{\frac{\alpha}{1-\alpha}}\right], \tag{20}$$

which also matches with the expression found in [16].

[3]The reason we consider the function (19) defined on the whole of $\mathbb{R}$ instead of just $\mathbb{R}_{>0}$ is that the resulting Fenchel conjugate is defined on $\mathbb{R}_{>0}$ instead of the whole of $\mathbb{R}$ which is useful in the application of Theorem 1.

For the measured relative entropy, we choose $f_1(t) = t\log t$ and $\mathrm{dom}(f_1) = \mathbb{R}_{>0}$. This gives

$$S^{\mathrm{P}}_{\mathrm{M},f_1}(\rho\|\sigma) = \sup_{\omega\in\mathscr{H}} \mathrm{Tr}[\rho\omega] - \mathrm{Tr}\left[\sigma e^{\omega-1}\right], \tag{21}$$

which again is equivalent to the expressions obtained in [16] and [21].

The variational expression (8) holds for all convex lower semicontinuous functions f. However, the optimization program as written is in general not a convex one unless additional conditions are imposed, such as operator convexity of f^*. The theorem below gives a general sufficient condition under which $S^{\mathrm{P}}_{\mathrm{M},f}(\rho\|\sigma)$ can be expressed as a convex program and which guarantees that the measured divergence and the one restricted to projective measurements match. This is similar to the result of [18, Theorem 5.8].

Theorem 1: Let $f : \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ be a convex and lower semicontinuous function such that $(0,\infty) \subset \mathrm{dom}(f)$, and let f^* be its Fenchel conjugate with domain $\mathrm{dom}(f^*)$.

Assume $\psi : J \to \mathrm{dom}(f^*)$ is a one-to-one map from an interval $J \subset \mathbb{R}$ to $\mathrm{dom}(f^*)$ such that

- ψ is operator concave on J,
- $f^* \circ \psi$ is operator convex on J.

Then we have, for any $\rho,\sigma \in \mathscr{H}_+$

$$S_{\mathrm{M},f}(\rho\|\sigma) = S^{\mathrm{P}}_{\mathrm{M},f}(\rho\|\sigma)$$
$$= \sup_{\gamma\in\mathscr{H},\,\mathrm{spec}(\gamma)\subset J} \{\mathrm{Tr}[\rho\psi(\gamma)] - \mathrm{Tr}[\sigma(f^*\circ\psi)(\gamma)]\}. \tag{22}$$

Remark 2: Combining Theorem 1 with [5, Corollary 4.20], it follows that under the conditions of the theorem, the data-processing inequality of $S_{\mathrm{M},f}$ holds under any positive trace-preserving maps.

Remark 3: For all the examples mentioned in Remark 1, such a ψ can be found. In fact, for $\alpha \in \left(0,\frac{1}{2}\right]$, $f^*_\alpha(z) = (1-\alpha)\left(\frac{-z}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}$ is already operator convex so $\psi(z) = z$ works. For $\alpha \in \left[\frac{1}{2},1\right)$, we can choose $J = \mathbb{R}_{>0}$ and $\psi(\lambda) = (f^*_\alpha)^{-1}(\lambda) = -\alpha\left(\frac{\lambda}{1-\alpha}\right)^{-\frac{1-\alpha}{\alpha}}$. For $\alpha > 1$, we can also choose $J = \mathbb{R}_{>0}$ and $\psi(\lambda) = (f^*_\alpha)^{-1}(\lambda) = \alpha\left(\frac{\lambda}{\alpha-1}\right)^{\frac{\alpha-1}{\alpha}}$. Similarly, for $f(t) = t\log t$, we choose $J = \mathbb{R}_{>0}$ and $\psi(\lambda) = (f^*)^{-1}(\lambda) = 1+\log\lambda$. For all these examples, the result (22) was already established in [16].

Another example is the measured total variation distance which is obtained by taking $f_{TV}(t) = \frac{1}{2}|t-1|$ and $\mathrm{dom}(f) = \mathbb{R}$. With this choice $f^*_{TV}(z) = z$ for $z \in \mathrm{dom}(f^*_{TV}) = \left[-\frac{1}{2},\frac{1}{2}\right]$.

Expression (22) then gives the well-known variational formulation of the trace distance.

Proof of Theorem 1: The second equality in (22) is a simple change of variables to (8), namely $\omega = \psi(\gamma)$. The inequality $S_{\mathrm{M},f}(\rho\|\sigma) \geq S^{\mathrm{P}}_{\mathrm{M},f}(\rho\|\sigma)$ is clear and so it remains to show the reverse inequality. Fix $\{M_x\}$ any POVM and note that we can write, using the same argument as in (12)

$$\sum_x \mathrm{Tr}[M_x\sigma] f\left(\frac{\mathrm{Tr}[M_x\rho]}{\mathrm{Tr}[M_x\sigma]}\right) \tag{23}$$
$$= \sup_{\{z_x\}\subset \mathrm{dom}(f^*)} \mathrm{Tr}\left[\sum_x z_x M_x \rho\right] - \mathrm{Tr}\left[\sum_x f^*(z_x) M_x \sigma\right].$$

Since ψ is a one-to-one map from J to $\mathrm{dom}(f^*)$, we can do the change of variables $z_x = \psi(\lambda_x)$ where $\lambda_x \in J$, and so we get

$$\sum_x \mathrm{Tr}[M_x\sigma] f\left(\frac{\mathrm{Tr}[M_x\rho]}{\mathrm{Tr}[M_x\sigma]}\right) \tag{24}$$
$$= \sup_{\{\lambda_x\}\subset J} \mathrm{Tr}\left[\sum_x \psi(\lambda_x) M_x \rho\right] - \mathrm{Tr}\left[\sum_x f^*(\psi(\lambda_x)) M_x \sigma\right].$$

By assumption, ψ is operator concave, and so using the operator Jensen inequality [22]

$$\psi\left(\sum_x \lambda_x M_x\right) = \psi\left(\sum_x \sqrt{M_x}(\lambda_x I)\sqrt{M_x}\right) \tag{25}$$
$$\geq \sum_x \psi(\lambda_x) M_x, \tag{26}$$

and similarly, since $f^* \circ \psi$ is operator convex we have

$$\sum_x f^*(\psi(\lambda_x)) M_x \geq (f^* \circ \psi)\left(\sum_x \lambda_x M_x\right). \tag{27}$$

Combining this relation with (24) gives

$$\sum_x \mathrm{Tr}[M_x\sigma] f\left(\frac{\mathrm{Tr}[M_x\rho]}{\mathrm{Tr}[M_x\sigma]}\right)$$
$$\leq \sup_{\{\lambda_x\}\subset J} \mathrm{Tr}\left[\psi\left(\sum_x \lambda_x M_x\right)\rho\right] - \mathrm{Tr}\left[f^*\circ\psi\left(\sum_x \lambda_x M_x\right)\sigma\right] \tag{28}$$
$$\leq \sup_{\gamma\in\mathscr{H},\ \mathrm{spec}(\gamma)\subset J} \mathrm{Tr}[\psi(\gamma)\rho] - \mathrm{Tr}[(f^*\circ\psi)(\gamma)\sigma] \tag{29}$$
$$= S^{\mathrm{P}}_{\mathrm{M},f}(\rho\|\sigma) \tag{30}$$

where in the second inequality above we used $\gamma = \sum_x \lambda_x M_x$ whose spectrum is in J since $M_x \geq 0$ and $\sum_x M_x = I$. 0 ■

III. Uhlmann's Theorem for Measured Divergences

Uhlmann's theorem is a fundamental result in quantum information theory that describes the relationship between the fidelity of mixed quantum states and their extensions [8]. More specifically, it states that for any two quantum states ρ_A and σ_{AR}, there exists an extension ρ_{AR} of ρ_A such that the fidelity $F(\rho_{AR}, \sigma_{RA})$ equals the fidelity $F(\rho_A, \sigma_A)$. The following result generalizes the renowned Uhlmann's theorem to measured f-divergences in general.

Theorem 2 (Uhlmann's theorem for measured f-divergences): Let f be a convex lower semicontinuous function such that $(0,\infty) \subset \mathrm{dom}(f)$. If its Fenchel conjugate f^* is operator convex and operator monotone on $\mathrm{dom}(f^*)$ and $\inf\{z : z \in \mathrm{dom}(f^*)\} = -\infty$, then for any $\rho_A \in \mathscr{D}(A)$ and $\sigma_{AR} \in \mathscr{H}_+(AR)$, we have

$$\inf_{\substack{\rho_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\rho_{AR}=\rho_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) = S_{\mathrm{M},f}(\rho_A\|\sigma_A). \tag{31}$$

Similarly, if $f^* : \mathrm{dom}(f^*) \to \mathrm{range}(f^*)$ is one-to-one and $(f^*)^{-1}$ is operator concave and operator monotone on $\mathrm{range}(f^*)$ and $\sup\{\lambda : \lambda \in \mathrm{range}(f^*)\} = +\infty$, then for any $\rho_{AR} \in \mathscr{D}(AR)$ and $\sigma_A \in \mathscr{H}_+(A)$, we have

$$\inf_{\substack{\sigma_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\sigma_{AR}=\sigma_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) = S_{\mathrm{M},f}(\rho_A\|\sigma_A). \tag{32}$$

In addition, the infimum in both equations is attained.

Remark 4: We leave it as an open question whether such a result can be extended to other choices of f. However, only assuming that f^* is operator convex and monotone on $\mathrm{dom}(f^*)$ is not sufficient. For example, the trace distance does not satisfy the Uhlmann property, but as mentioned in Remark 3, we have $f^*_{TV}(z) = z$ for $z \in \mathrm{dom}(f^*) = \left[-\frac{1}{2}, +\frac{1}{2}\right]$ which is clearly operator convex and monotone.

Remark 5: Let ϕ_{RA} be a purification of ρ_A. Then the set of all extensions of ρ_A can be obtained by ranging over all completely positive trace-preserving (CPTP) maps $\mathcal{E}_{R\to R}$ acting on the purifying system R, i.e.,

$$\{\rho_{AR} \in \mathscr{H}_+(RA) : \mathrm{Tr}_R\rho_{AR} = \rho_A\}$$
$$= \{(\mathrm{id}_A \otimes \mathcal{E}_{R\to R})(\phi_{AR}) : \mathcal{E}_{R\to R} \in \mathrm{CPTP}\}. \tag{33}$$

This can be similarly proved as [23, Lemma 10]. More explicitly, it is clear that "⊃" holds. Now for any extension ρ_{AR} of ρ_A. We have $\bar\rho_{AR} = \rho_A^{-1/2}\rho_{AR}\rho_A^{-1/2}$. Then we have $\bar\rho_{AR} \geq 0$ and $\mathrm{Tr}_R\bar\rho_{AR} = I_A$. From the Choi-Jamiolkowski isomorphism, we know that there exists a CPTP map $\mathcal{N}_{R\to R}$ such that $\bar\rho_{AR} = \mathrm{id}_A \otimes \mathcal{N}_{R\to R}(\Phi_{RA})$, where Φ_{RA} denotes the unnormalized maximally entangled state. Thus, we get $\rho_{AR} = \mathrm{id}_A \otimes \mathcal{N}_{R\to R}(\rho_A^{1/2}\Phi_{RA}\rho_A^{1/2})$. Denote $\psi_{RA} := \rho_A^{1/2}\Phi_{RA}\rho_A^{1/2}$. We know that ψ_{RA} is a purification of ρ_A. Due to the isometric equivalence between purifications, there exists an isometry $\mathcal{U}_{R\to R}$ on the system R such that $\psi_{RA} = \mathrm{id}_A \otimes \mathcal{U}_{R\to R}(\phi_{RA})$. This gives $\rho_{AR} = \mathrm{id}_A \otimes \mathcal{N}_{R\to R} \circ \mathcal{U}_{R\to R}(\phi_{RA})$. So we have the "⊂" direction. As a result, Eq. (31) can be equivalently reformulated in terms of purifications and channels. Let ϕ_{AR} be any purification of ρ_A. Then we have

$$\inf_{\mathcal{E}\in\mathrm{CPTP}} S_{\mathrm{M},f}((\mathrm{id}_A \otimes \mathcal{E}_{R\to R})(\phi_{AR})\|\sigma_{AR}) = S_{\mathrm{M},f}(\rho_A\|\sigma_A),$$

where the infimum is over all CPTP maps $\mathcal{E}_{R\to R}$ acting on the purifying system R. When σ_{AR} is a pure state, this formulation closely parallels the original Uhlmann's theorem for the fidelity [8], which involves optimization over isometries on the purifying system. Similarly, Eq. (32) can be reformulated as: for any purification ψ_{AR} of σ_A, we have

$$\inf_{\mathcal{E}\in\mathrm{CPTP}} S_{\mathrm{M},f}(\rho_{AR}\|(\mathrm{id}_A \otimes \mathcal{E}_{R\to R})(\psi_{AR})) = S_{\mathrm{M},f}(\rho_A\|\sigma_A).$$

Proof: We establish the result for Eq. (31) first. Recall the variational expression in Eq. (22) with $\psi = \mathrm{id}$:

$$S_{\mathrm{M},f}(\rho\|\sigma) = \sup_{\substack{\gamma\in\mathscr{H}\\ \mathrm{spec}(\gamma)\subset\mathrm{dom}(f^*)}} \{\mathrm{Tr}[\rho\gamma] - \mathrm{Tr}[\sigma f^*(\gamma)]\}. \quad (34)$$

We now use Sion's minimax theorem [24, Corollary 3.3] to get

$$\inf_{\substack{\rho_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\rho_{AR}=\rho_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) = \sup_{\substack{\gamma_{AR}\in\mathscr{H}(AR)\\ \mathrm{spec}(\gamma_{AR})\subset\mathrm{dom}(f^*)}} \inf_{\substack{\rho_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\rho_{AR}=\rho_A}} \{\mathrm{Tr}[\rho_{AR}\gamma_{AR}] - \mathrm{Tr}[\sigma_{AR} f^*(\gamma_{AR})]\}. \quad (35)$$

In fact, the set $\{\rho_{AR} \in \mathscr{H}_+(AR) : \mathrm{Tr}_R\rho_{AR} = \rho_A\}$ is compact and convex and $\{\gamma_{AR} \in \mathscr{H}(AR) : \mathrm{spec}(\gamma_{AR}) \subset \mathrm{dom}(f^*)\}$ is convex. In addition the expression $\mathrm{Tr}[\rho_{AR}\gamma_{AR}]-\mathrm{Tr}[\sigma_{AR}f^*(\gamma_{AR})]$ is concave in γ_{AR} (because f^* is operator convex) and linear in ρ_{AR}. Now, by semidefinite programming duality (e.g. [25, Section 1.2.3]), we have for a fixed $\gamma_{AR} \in \mathscr{H}(AR)$

$$\inf_{\substack{\rho_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\rho_{AR}=\rho_A}} \mathrm{Tr}[\rho_{AR}\gamma_{AR}] = \sup_{\substack{\Lambda_A\in\mathscr{H}(A)\\ \Lambda_A\otimes I_R\le\gamma_{AR}}} \mathrm{Tr}[\rho_A\Lambda_A]. \quad (36)$$

In fact it is clear that the infimum is finite as $0 \le \rho_{AR} \le \mathrm{Tr}[\rho_A]I_{AR}$ and $\Lambda_A = -(\|\gamma_{AR}\|_\infty + 1)I_A$ is strictly feasible for the dual. As a result,

$$\inf_{\substack{\rho_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\rho_{AR}=\rho_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) = \sup_{\substack{\gamma_{AR}\in\mathscr{H}(AR)\\ \mathrm{spec}(\gamma_{AR})\subset\mathrm{dom}(f^*)}} \sup_{\substack{\Lambda_A\in\mathscr{H}(A)\\ \Lambda_A\otimes I_R\le\gamma_{AR}}} \{\mathrm{Tr}[\rho_A\Lambda_A] - \mathrm{Tr}[\sigma_{AR} f^*(\gamma_{AR})]\}. \quad (37)$$

Note that we can obtain a lower bound on the supremum (the easy direction) by restricting ourselves to $\gamma_{AR} = \Lambda_A \otimes I_R$ and observing that $f^*(\Lambda_A \otimes I_R) = f^*(\Lambda_A) \otimes I_R$:

$$\inf_{\substack{\rho_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\rho_{AR}=\rho_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) \ge \sup_{\substack{\Lambda_A\in\mathscr{H}(A)\\ \mathrm{spec}(\Lambda_A)\subset\mathrm{dom}(f^*)}} \{\mathrm{Tr}[\rho_A\Lambda_A] - \mathrm{Tr}[\sigma_A f^*(\Lambda_A)]\} \quad (38)$$

$$= S_{\mathrm{M},f}(\rho_A\|\sigma_A). \quad (39)$$

For the other direction, as f^* is operator monotone we have for any feasible (γ_{AR}, Λ_A), it holds that $f^*(\gamma_{AR}) \ge f^*(\Lambda_A)\otimes I_R$. This then implies

$$\inf_{\substack{\rho_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\rho_{AR}=\rho_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) \le \sup_{\substack{\gamma_{AR}\in\mathscr{H}(AR)\\ \mathrm{spec}(\gamma_{AR})\subset\mathrm{dom}(f^*)}} \sup_{\substack{\Lambda_A\in\mathscr{H}(A)\\ \Lambda_A\otimes I_R\le\gamma_{AR}}} \{\mathrm{Tr}[\rho_A\Lambda_A] - \mathrm{Tr}[\sigma_A f^*(\Lambda_A)]\} \quad (40)$$

$$\le \sup_{\substack{\Lambda_A\in\mathscr{H}(A)\\ \mathrm{spec}(\Lambda_A)\subset\mathrm{dom}(f^*)}} \{\mathrm{Tr}[\rho_A\Lambda_A] - \mathrm{Tr}[\sigma_A f^*(\Lambda_A)]\} \quad (41)$$

$$= S_{\mathrm{M},f}(\rho_A\|\sigma_A). \quad (42)$$

In the second inequality, we used the condition $\inf\mathrm{dom}(f^*) = -\infty$ to get that $\Lambda_A\otimes I_R \le \gamma_{AR}$ and $\mathrm{spec}(\gamma_{AR}) \subset \mathrm{dom}(f^*)$ implies $\mathrm{spec}(\Lambda_A) \subset \mathrm{dom}(f^*)$. This establishes (31).

To prove that the infimum is attained, it is well-known that a lower semicontinuous function attains its infimum on a compact set [26, Theorem 7.3.1]. This can be more explicitly argued as follows. For any $n \ge 1$, we can take $\rho_{AR}^{(n)} \in \mathscr{H}_+(AR)$ satisfying $\mathrm{Tr}_R\rho_{AR}^{(n)} = \rho_A$ such that $0 \le S_{\mathrm{M},f}(\rho_{AR}^{(n)}\|\sigma_{AR}) - S_{\mathrm{M},f}(\rho_A\|\sigma_A) \le \frac{1}{n}$. As the set $\{\rho_{AR} \in \mathscr{H}_+(AR) : \mathrm{Tr}_R\rho_{AR} = \rho_A\}$ is compact, we can find a converging subsequence $\lim_{i\to\infty}\rho_{AR}^{(n_i)} = \rho_{AR}^\infty$. By construction, $\mathrm{Tr}_R\rho_{AR}^\infty = \rho_A$ so the data processing inequality (see Remark 2) gives $S_{\mathrm{M},f}(\rho_{AR}^\infty\|\sigma_{AR}) \ge S_{\mathrm{M},f}(\rho_A\|\sigma_A)$. As $S_{\mathrm{M},f}$ can be written as the supremum of lower semicontinuous functions, it is also lower semi-continuous. This implies that

$$S_{\mathrm{M},f}(\rho_A\|\sigma_A) = \lim_{i\to\infty} S_{\mathrm{M},f}(\rho_{AR}^{(n_i)}\|\sigma_{AR}) \ge S_{\mathrm{M},f}(\rho_{AR}^\infty\|\sigma_{AR}). \quad (43)$$

So the infimum is attained at ρ_{AR}^∞.

The proof of Eq. (32) is similar. Recall the variational expression in Eq. (22) with $\psi = (f^*)^{-1}$:

$$S_{\mathrm{M},f}(\rho\|\sigma) = \sup_{\substack{\gamma\in\mathscr{H}\\ \mathrm{spec}(\gamma)\subset\mathrm{range}(f^*)}} \left\{\mathrm{Tr}[\rho(f^*)^{-1}(\gamma)] - \mathrm{Tr}[\sigma\gamma]\right\}. \quad (44)$$

We now use Sion's minimax theorem [24, Corollary 3.3] to get

$$\inf_{\substack{\sigma_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\sigma_{AR}=\sigma_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) = \sup_{\substack{\gamma_{AR}\in\mathscr{H}(AR)\\ \mathrm{spec}(\gamma_{AR})\subset\mathrm{range}(f^*)}} \inf_{\substack{\sigma_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\sigma_{AR}=\sigma_A}} \left\{\mathrm{Tr}[\rho_{AR}(f^*)^{-1}(\gamma_{AR})] - \mathrm{Tr}[\sigma_{AR}\gamma_{AR}]\right\}. \quad (45)$$

In fact, the set $\{\sigma_{AR} \in \mathscr{H}_+(AR) : \mathrm{Tr}_R\sigma_{AR} = \sigma_A\}$ is compact and convex and $\{\gamma_{AR} \in \mathscr{H}(AR) : \mathrm{spec}(\gamma_{AR}) \subset \mathrm{range}(f^*)\}$ is convex. In addition the expression $\mathrm{Tr}[\rho_{AR}(f^*)^{-1}(\gamma_{AR})] - \mathrm{Tr}[\sigma_{AR}\gamma_{AR}]$ is concave in γ_{AR} and linear in σ_{AR}. Now, by semidefinite programming duality, we have

$$\sup_{\substack{\sigma_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\sigma_{AR}=\sigma_A}} \mathrm{Tr}[\sigma_{AR}\gamma_{AR}] = \inf_{\substack{\Lambda_A\in\mathscr{H}(A)\\ \Lambda_A\otimes I_R\ge\gamma_{AR}}} \mathrm{Tr}[\sigma_A\Lambda_A]. \quad (46)$$

In fact it is clear that the supremum is finite as $0 \le \sigma_{AR} \le \mathrm{Tr}[\sigma_A]I_{AR}$ and $\Lambda_A = (\|\gamma_{AR}\|_\infty + 1)I_A$ is strictly feasible for the dual. As a result,

$$\inf_{\substack{\sigma_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\sigma_{AR}=\sigma_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) = \sup_{\substack{\gamma_{AR}\in\mathscr{H}(AR)\\ \mathrm{spec}(\gamma_{AR})\subset\mathrm{range}(f^*)}} \sup_{\substack{\Lambda_A\in\mathscr{H}(A)\\ \Lambda_A\otimes I_R\ge\gamma_{AR}}} \left\{\mathrm{Tr}[\rho_{AR}(f^*)^{-1}(\gamma_{AR})] - \mathrm{Tr}(\sigma_A\Lambda_A)]\right\}. \quad (47)$$

Note that we can in particular choose $\gamma_{AR} = \Lambda_A \otimes I_R$ and get the easy direction:

$$\inf_{\substack{\sigma_{AR}\in\mathscr{H}_+(AR)\\ \mathrm{Tr}_R\sigma_{AR}=\sigma_A}} S_{\mathrm{M},f}(\rho_{AR}\|\sigma_{AR}) \ge \sup_{\substack{\Lambda_A\in\mathscr{H}(A)\\ \mathrm{spec}(\Lambda_A)\subset\mathrm{range}(f^*)}} \left\{\mathrm{Tr}[\rho_A(f^*)^{-1}(\Lambda_A)] - \mathrm{Tr}[\sigma_A\Lambda_A]\right\} \quad (48)$$

$$= S_{\mathrm{M},f}(\rho_A\|\sigma_A). \quad (49)$$

For the other direction, as $(f^*)^{-1}$ is operator monotone we have for any feasible (γ, Λ), it holds that $(f^*)^{-1}(\gamma_{AR}) \le$

$(f^*)^{-1}(\Lambda_A) \otimes I_R$. In addition, as $\sup \operatorname{range}(f^*) = +\infty$ and $\operatorname{spec}(\gamma_{AR}) \subset \operatorname{range}(f^*)$, we have that $\operatorname{spec}(\Lambda_A) \subset \operatorname{range}(f^*)$. In fact, f^* is continuous on its domain since it is a univariate lower semicontinuous convex function, and thus $\operatorname{range}(f^*)$ is an interval, which by our additional assumption, is of the form $[a, +\infty)$ or $(a, +\infty)$. This then implies

$$\inf_{\substack{\sigma_{AR} \in \mathscr{H}_+(AR) \\ \operatorname{Tr}_R \sigma_{AR} = \sigma_A}} S_{\mathrm{M},f}(\rho_{AR} \| \sigma_{AR})$$
$$\leq \sup_{\substack{\Lambda_A \in \mathscr{H}(A) \\ \operatorname{spec}(\Lambda_A) \subset \operatorname{range}(f^*)}} \left\{ \operatorname{Tr}[\rho_A (f^*)^{-1}(\Lambda_A)] - \operatorname{Tr}[\sigma_A \Lambda_A] \right\} \tag{50}$$
$$= S_{\mathrm{M},f}(\rho_A \| \sigma_A). \tag{51}$$

This establishes (32). The attainment of the infimum follows the same proof as before. 0 ■

As an immediate corollary, we obtain the following result of Rényi divergences.

Corollary 1 (Uhlmann's theorem for measured Rényi divergences): Let $\rho_A \in \mathscr{D}(A)$, $\sigma_{AR} \in \mathscr{H}_+(AR)$, and $\alpha \in \left[0, \frac{1}{2}\right]$, then

$$\inf_{\substack{\rho_{AR} \in \mathscr{H}_+(AR) \\ \operatorname{Tr}_R \rho_{AR} = \rho_A}} D_{\mathrm{M},\alpha}(\rho_{AR} \| \sigma_{AR}) = D_{\mathrm{M},\alpha}(\rho_A \| \sigma_A). \tag{52}$$

Similarly, let $\rho_{AR} \in \mathscr{D}(AR), \sigma_A \in \mathscr{H}_+(A)$ and $\alpha \in \left[\frac{1}{2}, +\infty\right]$ 1 [4], then

$$\inf_{\substack{\sigma_{AR} \in \mathscr{H}_+(AR) \\ \operatorname{Tr}_R \sigma_{AR} = \sigma_A}} D_{\mathrm{M},\alpha}(\rho_{AR} \| \sigma_{AR}) = D_{\mathrm{M},\alpha}(\rho_A \| \sigma_A). \tag{53}$$

In addition, the infimum in both equations is attained.

Remark 6: Note that for $\alpha \geq 1$, it is easy to see that Uhlmann's property in the first argument (i.e., (52)) cannot hold in general. In fact, we can choose $\rho_A = |0\rangle\langle 0|$ and the Bell state $\sigma_{AR} = |\Phi\rangle\langle\Phi|$ where $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, then $D_{\mathrm{M},\alpha}(\rho_A \| \sigma_A) = \log 2$ but $D_{\mathrm{M},\alpha}(\rho_{AR} \| \sigma_{AR}) = +\infty$ for any extension ρ_{AR} of ρ_A.

Proof: For $\alpha \in \left[0, \frac{1}{2}\right]$, as previously mentioned, we have $f_\alpha^*(z) = (1-\alpha)\left(\frac{-z}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}$ and $\operatorname{dom}(f^*) = \mathbb{R}_{<0}$. In this case, f_α^* is operator convex, operator monotone and the domain is unbounded from below.

For $\alpha \in \left[\frac{1}{2}, 1\right)$, we also have $f_\alpha^*(z) = (1-\alpha)\left(\frac{-z}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}$ and $\operatorname{dom}(f_\alpha^*) = \mathbb{R}_{<0}$. But in this case, we check the conditions for (32): we have f_α^* is one-to-one with $\operatorname{range}(f_\alpha^*) = \mathbb{R}_{>0}$ and $(f_\alpha^*)^{-1}(\lambda) = -\alpha\left(\frac{\lambda}{1-\alpha}\right)^{-\frac{1-\alpha}{\alpha}}$ which is operator concave, operator monotone and we have $\sup \operatorname{range}(f_\alpha^*) = +\infty$.

For $\alpha > 1$, we have $f_\alpha^*(z) = (\alpha - 1)\left(\frac{z}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}$ with $\operatorname{dom}(f^*) = \operatorname{range}(f^*) = \mathbb{R}_{>0}$. Thus, $(f_\alpha^*)^{-1}(\lambda) = \alpha\left(\frac{\lambda}{\alpha-1}\right)^{\frac{\alpha-1}{\alpha}}$ is operator concave, operator monotone and $\sup \operatorname{range}(f_\alpha^*) = +\infty$.

For $\alpha = 1$, recall that $f_1(t) = t \log t$, and thus we have $f_1^*(z) = e^{z-1}$ with $\operatorname{dom}(f_1^*) = \mathbb{R}$ and $\operatorname{range}(f_1^*) = \mathbb{R}_{>0}$. Thus $(f_1^*)^{-1}(\lambda) = 1 + \log \lambda$ is operator concave, operator monotone and $\sup \operatorname{range}(f_1^*) = +\infty$.

For $\alpha = +\infty$, we take supremum over $\alpha \in [1, +\infty)$. Note that $D_{\mathrm{M},\alpha}(\rho \| \sigma)$ is lower semicontinuous in (ρ, σ) for every $\alpha \in [1, +\infty)$ [27, Proposition 18] and monotonic increasing in α for every (ρ, σ) [27, Proposition 22]. So we can use the minimax theorem in [28, Corollary A.2] to exchange the supremum and infimum. Finally, noting that $\sup_{\alpha \geq 1} D_{\mathrm{M},\alpha}(\rho \| \sigma) = D_{\max}(\rho \| \sigma)$, we have the desired result. As $D_{\max}(\rho \| \sigma)$ is lower semicontinous in (ρ, σ) [27, Proposition 18], the infimum is attained on a compact set [26, Theorem 7.3.1]. ■

The case $\alpha = \frac{1}{2}$ corresponds to the fidelity as $D_{\mathrm{M},\frac{1}{2}} = -2 \log F$, where F is the fidelity; see [14], [29]. Thus, Uhlmann's theorem is the special case $\alpha = \frac{1}{2}$ of this theorem. The case $\alpha = +\infty$ was shown in [9]. Moreover, the regularized Uhlmann's theorem from [9] follows easily from this theorem. Note that the result in [9] is shown for $\alpha > 1$, whereas the following result applies more generally.

Let $D_{\mathrm{S},\alpha}(\rho \| \sigma) := \frac{1}{\alpha - 1} \log \operatorname{Tr}\left[\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right]^\alpha$ be the sandwiched Rényi divergence [30], [31].

Corollary 2: Let $\rho_{AR} \in \mathscr{D}(AR), \sigma_A \in \mathscr{H}_+(A)$ and $\alpha \in \left[\frac{1}{2}, +\infty\right]$, then

$$\lim_{n \to \infty} \frac{1}{n} \inf_{\substack{\sigma_{A^n R^n} \in \mathscr{H}_+(A^n R^n) \\ \operatorname{Tr}_{R^n} \sigma_{A^n R^n} = \sigma_A^{\otimes n}}} D_{\mathrm{S},\alpha}(\rho_{AR}^{\otimes n} \| \sigma_{A^n R^n}) = D_{\mathrm{S},\alpha}(\rho_A \| \sigma_A). \tag{54}$$

Proof: Applying Corollary 1 to the states $\rho_{AR}^{\otimes n}$ and $\sigma_A^{\otimes n}$, we get

$$\inf_{\substack{\sigma_{A^n R^n} \in \mathscr{H}_+(A^n R^n) \\ \operatorname{Tr}_{R^n} \sigma_{A^n R^n} = \sigma_A^{\otimes n}}} D_{\mathrm{M},\alpha}(\rho_{AR}^{\otimes n} \| \sigma_{A^n R^n}) = D_{\mathrm{M},\alpha}(\rho_A^{\otimes n} \| \sigma_A^{\otimes n}). \tag{55}$$

By the asymptotic equivalence of the measured Rényi divergence and the sandwiched Rényi divergence [10, Lemma 28], applied with $\mathscr{B}_n = \{\sigma_{A^n R^n} \in \mathscr{H}_+(A^n R^n) : \operatorname{Tr}_{R^n} \sigma_{A^n R^n} = \sigma_A^{\otimes n}\}$ (which is convex, compact, and permutation invariant), we obtain

$$\lim_{n \to \infty} \frac{1}{n} \inf_{\sigma_{A^n R^n} \in \mathscr{B}_n} D_{\mathrm{S},\alpha}(\rho_{AR}^{\otimes n} \| \sigma_{A^n R^n})$$
$$= \lim_{n \to \infty} \frac{1}{n} \inf_{\sigma_{A^n R^n} \in \mathscr{B}_n} D_{\mathrm{M},\alpha}(\rho_{AR}^{\otimes n} \| \sigma_{A^n R^n}). \tag{56}$$

Combining Eq. (55), Eq. (56) and the fact that $\lim_{n \to \infty} \frac{1}{n} D_{\mathrm{M},\alpha}(\rho_A^{\otimes n} \| \sigma_A^{\otimes n}) = D_{\mathrm{S},\alpha}(\rho_A \| \sigma_A)$ [14, Corollary 3.8], we have the asserted result. 0 ■

Remark 7: Note that the set $\mathscr{B}_n$ defined in the proof above satisfies all assumptions (A.1)-(A.4) in [10, Assumption 25]. In fact, (A.1) and (A.2) are clear and $\operatorname{Tr}_{R_1 R_2}(\sigma_{A_1 R_1}^{(1)} \otimes \sigma_{A_2 R_2}^{(2)}) = \sigma_{A_1}^{(1)} \otimes \sigma_{A_2}^{(2)}$ so (A.3) is satisfied. The condition (A.4) states that the support function is submultiplicative, i.e., $h_{\mathscr{B}_{m+k}}(X_{A^m R^m} \otimes X_{A^k R^k}) \leq h_{\mathscr{B}_m}(X_{A^m R^m}) h_{\mathscr{B}_k}(X_{A^k R^k})$. For that, recall that we have obtained a dual formulation for the support function of the sets $\mathscr{B}_n$ in (46). We use this same dual formulation to establish the desired submultiplicativity:

$$h_{\mathscr{B}_{m+k}}(X_{A^m R^m} \otimes X_{A^k R^k})$$
$$= \sup_{\substack{\sigma_{A^{m+k} R^{m+k}} \in \mathscr{H}_+(A^{m+k} R^{m+k}) \\ \operatorname{Tr}_{R^{m+k}} \sigma_{A^{m+k} R^{m+k}} = \sigma_A^{\otimes m+k}}} \operatorname{Tr}(\sigma_{A^{m+k} R^{m+k}} X_{A^m R^m} \otimes X_{A^k R^k}) \tag{57}$$
$$= \inf_{\substack{\Lambda_{A^{m+k}} \in \mathscr{H}(A^{m+k}) \\ \Lambda_{A^{m+k}} \otimes I_{R^{m+k}} \geq X_{A^m R^m} \otimes X_{A^m R^m}}} \operatorname{Tr}(\sigma_A^{\otimes m+k} \Lambda_{A^{m+k}}) \tag{58}$$
$$\leq \inf_{\substack{\Lambda_{A^m} \in \mathscr{H}(A^m), \Lambda_{A^k} \in \mathscr{H}(A^k) \\ \Lambda_{A^m} \otimes \Lambda_{A^k} \otimes I_{R^{m+k}} \geq X_{A^m R^m} \otimes X_{A^k R^k}}} \operatorname{Tr}(\sigma_A^{\otimes m+k} \Lambda_{A^m} \otimes \Lambda_{A^k}) \tag{59}$$

[4] When $\alpha = 1$, $D_{\mathrm{M},\alpha}$ is interpreted as D_{M}. When $\alpha = +\infty$, $D_{\mathrm{M},\alpha}$ is interpreted as $D_{\max}$.

$$\leq \inf_{\substack{\Lambda_{A^m}\in\mathscr{H}(A^m),\Lambda_{A^k}\in\mathscr{H}(A^k)\\ \Lambda_{A^m}\otimes I_{R^m}\geq X_{A^mR^m}\\ \Lambda_{A^k}\otimes I_{R^k}\geq X_{A^kR^k}}} \mathrm{Tr}(\sigma_A^{\otimes m}\Lambda_{A^m})\mathrm{Tr}(\sigma_A^{\otimes k}\Lambda_{A^k}) \quad (60)$$

$$= h_{\mathscr{B}_m}(X_{A^mR^m})h_{\mathscr{B}_k}(X_{A^kR^k}). \quad (61)$$

Thus, the results of [10] characterize the asymmetric hypothesis testing problem between $\{\rho_{AR}^{\otimes n}\}$ and $\mathscr{B}_n$. A related hypothesis testing scenario was considered in [32]. The authors of [32] studied hypothesis testing between $\{\rho_{AR}^{\otimes n}\}$ and $\{\rho_A^{\otimes n}\otimes\sigma_{R^n} : \sigma\in\mathscr{D}(R^n)\}$ and demonstrated that the Rényi mutual information is the strong converse exponent in this context. Similarly, we expect Corollary 2 can be used to establish the strong converse exponent for hypothesis testing between $\rho_{AR}^{\otimes n}$ and $\mathscr{B}_n$.

IV. Duality With the Quantum Relative Entropy

This section is specific to the measured relative entropy (i.e., $f(t) = t\log t$): we utilize the variational expression for D_{M} to derive a duality relation with the Umegaki quantum relative entropy which is defined by $D(\rho\|\sigma) = \mathrm{Tr}(\rho\log\rho) - \mathrm{Tr}(\rho\log\sigma)$ when the support of ρ is included in the support of σ and $+\infty$ otherwise. For a quantum divergence $\mathbb{D}$ and a set $\mathscr{C}$ of positive operators, we write $\mathbb{D}(\rho\|\mathscr{C}) := \inf_{\sigma\in\mathscr{C}}\mathbb{D}(\rho\|\sigma)$.

Proposition 2: Let $\rho\in\mathscr{D}$ and $\mathscr{C}\subset\mathscr{H}_+$ be an non-empty compact convex set, such that $\mathrm{supp}(\rho)\subset\mathrm{supp}(\omega)$ for some $\omega\in\mathscr{C}$. Then it holds that

$$D(\rho\|\mathscr{C}) + D_{\mathrm{M}}(\rho\|\mathscr{C}^\circ_{++}) = D(\rho\|I), \quad (62)$$

where $\mathscr{C}^\circ := \{X : \mathrm{Tr}[XY]\leq 1, \forall\, Y\in\mathscr{C}\}$ is the polar set and $\mathscr{C}^\circ_{++} := \mathscr{C}^\circ\cap\mathscr{H}_{++}$.

Proof: By the variational expression for the measured relative entropy [16], [21] also presented in (21), we have

$$\begin{aligned} &D_{\mathrm{M}}(\rho\|\sigma)\\ &= \sup_{\omega\in\mathscr{H}_{++}} \mathrm{Tr}[\rho\log\omega] + 1 - \mathrm{Tr}[\sigma\omega] && (63)\\ &= \sup_{\omega\in\mathscr{H}_{++}} -D(\rho\|\omega) - \mathrm{Tr}[\sigma\omega] + 1 + D(\rho\|I) && (64)\\ &= g_\rho^*(-\sigma) + 1 + D(\rho\|I), && (65)\end{aligned}$$

where the second line follows from the definition of Umegaki relative entropy and the third line follows by denoting g_ρ^* as the Fenchel conjugate of $g_\rho : \omega\mapsto D(\rho\|\omega)$, with $\mathrm{dom}(g_\rho) = \{\omega\geq 0 : \rho\ll\omega\}$, i.e., for $\tau\in\mathscr{H}$, $g_\rho^*(\tau) = \sup_{\omega>0}\mathrm{Tr}[\omega\tau] - D(\rho\|\omega)$.[5] Let us show that $\mathrm{dom}(g_\rho^*) = \{\tau\leq 0 : \mathrm{supp}(\rho)\subset\mathrm{supp}(\tau)\}$. In fact, we can write

$$\begin{aligned} &g_\rho^*(\tau)\\ &= \sup_{\{\lambda_i,\Pi_i\}_i}\sum_i \lambda_i\mathrm{Tr}[\Pi_i\tau] + \log\lambda_i\mathrm{Tr}[\rho\Pi_i] - \mathrm{Tr}[\rho\log\rho] && (66)\\ &= -\mathrm{Tr}[\rho\log\rho] + \sup_{\{\Pi_i\}_i}\sum_i\sup_{\{\lambda_i\}_i}\lambda_i\mathrm{Tr}[\Pi_i\tau] + \log\lambda_i\mathrm{Tr}[\rho\Pi_i]. && (67)\end{aligned}$$

where the supremum is over orthogonal projectors Π_i with $\sum_i\Pi_i = I$ and $\lambda_i > 0$. This quantity is finite if and only if for all i, we have $\mathrm{Tr}[\Pi_i\tau] < 0$ or if $\mathrm{Tr}[\Pi_i\tau] = 0$, then $\mathrm{Tr}[\rho\Pi_i] = 0$. As g_ρ is lower semicontinuous, the biduality relation of Fenchel conjugate [20, Theorem 12.2] gives

$$\begin{aligned} D(\rho\|\sigma) &= g_\rho(\sigma) && (68)\\ &= \sup_{\omega\in\mathrm{dom}(g_\rho^*)}\mathrm{Tr}[\sigma\omega] - g_\rho^*(\omega) && (69)\\ &= \sup_{\omega<0}\mathrm{Tr}[\sigma\omega] - g_\rho^*(\omega) && (70)\\ &= \sup_{\omega<0}\mathrm{Tr}[\sigma\omega] - D_{\mathrm{M}}(\rho\|-\omega) + D(\rho\|I) + 1 && (71)\\ &= \sup_{\omega>0} -\mathrm{Tr}[\sigma\omega] - D_{\mathrm{M}}(\rho\|\omega) + D(\rho\|I) + 1. && (72)\end{aligned}$$

In the second line we used the fact that for any $\omega\in\mathrm{dom}(g_\rho^*)$ and any $\epsilon>0$, we have $(1-\epsilon)\omega + \epsilon(-I) < 0$ and

$$g_\rho^*(\omega) = \lim_{\epsilon\downarrow 0} g_\rho^*((1-\epsilon)\omega + \epsilon(-I)) \quad (73)$$

which follows by applying [20, Theorem 7.5]. Note that the theorem applies because $-I$ is in the interior of $\mathrm{dom}(g_\rho^*)$ and g_ρ^* is lower semicontinuous. The third line (71) follows from Eq. (65), and the last line follows by replacing ω with $-\omega$. Optimizing over $\sigma\in\mathscr{C}$, we have

$$\begin{aligned} D(\rho\|\mathscr{C}) &= \inf_{\sigma\in\mathscr{C}} D(\rho\|\sigma) && (74)\\ &= \inf_{\sigma\in\mathscr{C}}\sup_{\omega\in\mathscr{H}_{++}} -\mathrm{Tr}[\sigma\omega] - D_{\mathrm{M}}(\rho\|\omega) + 1 + D(\rho\|I). && (75)\end{aligned}$$

Since $D_{\mathrm{M}}(\rho\|\omega)$ is convex in ω, the above objective function is concave in ω and convex (actually linear) in σ. Moreover, since $\mathscr{C}$ is a compact convex set by assumption, we can use Sion's minimax theorem [24, Corollary 3.3] to exchange the infimum and supremum. This gives

$$D(\rho\|\mathscr{C}) = \sup_{\omega\in\mathscr{H}_{++}}\inf_{\sigma\in\mathscr{C}} -\mathrm{Tr}[\sigma\omega] - D_{\mathrm{M}}(\rho\|\omega) + 1 + D(\rho\|I). \quad (76)$$

Let $h_{\mathscr{C}}(\omega) := \sup_{\sigma\in\mathscr{C}}\mathrm{Tr}[\sigma\omega]$ be the support function of $\mathscr{C}$. We get

$$\begin{aligned} D(\rho\|\mathscr{C}) &= \sup_{\omega\in\mathscr{H}_{++}} -h_{\mathscr{C}}(\omega) - D_{\mathrm{M}}(\rho\|\omega) + 1 + D(\rho\|I) && (77)\\ &\overset{(a)}{=} \sup_{\substack{\omega\in\mathscr{H}_{++}\\ h_{\mathscr{C}}(\omega)=1}} -D_{\mathrm{M}}(\rho\|\omega) + D(\rho\|I) && (78)\\ &\overset{(b)}{=} \sup_{\substack{\omega\in\mathscr{H}_{++}\\ h_{\mathscr{C}}(\omega)\leq 1}} -D_{\mathrm{M}}(\rho\|\omega) + D(\rho\|I) && (79)\\ &\overset{(c)}{=} -D_{\mathrm{M}}(\rho\|\mathscr{C}^\circ_{++}) + D(\rho\|I), && (80)\end{aligned}$$

To see why (a) holds, let ω be any feasible solution in the first line and define $\widetilde{\omega} = \omega/h_{\mathscr{C}}(\omega)$. Then $h_{\mathscr{C}}(\widetilde{\omega}) = 1$ and we will see that $\widetilde{\omega}$ achieves an objective value no smaller than ω as

$$\begin{aligned} -D_{\mathrm{M}}(\rho\|\widetilde{\omega}) &= -D_{\mathrm{M}}(\rho\|\omega) - \log h_{\mathscr{C}}(\omega) && (81)\\ &\geq -D_{\mathrm{M}}(\rho\|\omega) - h_{\mathscr{C}}(\omega) + 1 && (82)\end{aligned}$$

where the inequality follows from the fact that $\log x\leq x-1$. The step (b) follows by the same argument. Suppose ω is a solution in the third line, let $\widetilde{\omega} = \omega/h_{\mathscr{C}}(\omega)$ and we can check that $\widetilde{\omega}$ gives an objective value no smaller than the second line. Finally, we have the step (c) by the fact that $h_{\mathscr{C}}(\omega)\leq 1$ if and only if $\omega\in\mathscr{C}^\circ$ and the notation that $\mathscr{C}^\circ_{++} = \mathscr{C}^\circ\cap\mathscr{H}_{++}$. This completes the proof. 0 ■

[5]Note that $g_\rho^*(\tau) := \sup_{\omega\in\mathrm{dom}(g_\rho)}\mathrm{Tr}[\omega\tau] - D(\rho\|\omega) = \sup_{\omega>0}\mathrm{Tr}[\omega\tau] - D(\rho\|\omega)$ since for any $\omega\geq 0$, $D(\rho\|\omega) = \lim_{\epsilon\downarrow 0} D(\rho\|\omega+\epsilon I)$.

Remark 8: By the bipolar theorem in optimization theory $(\mathscr{C}^\circ)^\circ = \overline{\mathrm{conv}(\mathscr{C} \cup \{0\})}$ [33, Exercise 1.15] together with the dominance property $D_{\mathrm{M}}(\rho\|\lambda\sigma) \geq D_{\mathrm{M}}(\rho\|\sigma)$ for $\lambda \in (0,1]$, one obtains a similar duality $D(\rho\|\mathscr{C}^\circ_{++}) + D_{\mathrm{M}}(\rho\|\mathscr{C}) = D(\rho\|I)$. Such duality relations allow one to transfer additivity properties between the two terms. For example, if the polar set $(\mathscr{C}_n)^\circ$ is closed under tensor products—a technical assumption in many recent applications (e.g., [10], [11], [12], [13])—then $D(\rho^{\otimes n}\|(\mathscr{C}_n)^\circ_{++})$ is clearly subadditive by definition, which in turn implies superadditivity of $D_{\mathrm{M}}(\rho^{\otimes n}\|\mathscr{C}_n)$. This superadditivity underpins several recent results, as cited above. The duality relation thus provides a more intuitive explanation for why measured relative entropy exhibits superadditivity when the polar sets satisfy this property. The same reasoning applies when swapping the roles of D and D_{M}. These insights are in the same spirit as recent progress on the Frenkel representation of quantum relative entropy [34], which offers a direct route to proving the data-processing inequality—often challenging from the original definition, but straightforward from this new perspective.

Since the measured relative entropy coincides with the quantum relative entropy in the classical case, we can replace D_{M} with D, obtaining an equality expressed solely in terms of the relative entropy. In the quantum case, if we consider the regularized divergence $\mathbb{D}^\infty(\rho\|\mathscr{C}) := \lim_{n\to\infty} \frac{1}{n}\mathbb{D}(\rho^{\otimes n}\|\mathscr{C}_n)$, we get a self-duality for the quantum relative entropy as follows.

Corollary 3: Let $\rho \in \mathscr{D}$. Let $\{\mathscr{C}_n\}_{n\in\mathbb{N}}$ be a sequences of sets satisfying $\mathscr{C}_n \subset \mathscr{H}_+(\mathcal{H}^{\otimes n})$ and $D_{\max}(\rho^{\otimes n}\|\mathscr{C}_n) \leq cn$, for all $n \in \mathbb{N}$ and a constant $c \in \mathbb{R}_+$. In addition, each $\mathscr{C}_n$ is convex, compact and permutation-invariant, and $\mathscr{C}_m \otimes \mathscr{C}_k \subset \mathscr{C}_{m+k}$ for any $m, k \in \mathbb{N}$. Then it holds that

$$D^\infty(\rho\|\mathscr{C}) + D^\infty(\rho\|\mathscr{C}^\circ_{++}) = D(\rho\|I). \tag{83}$$

Proof: By Proposition 2, we know that

$$\frac{1}{n}D(\rho^{\otimes n}\|\mathscr{C}_n) + \frac{1}{n}D_{\mathrm{M}}(\rho^{\otimes n}\|(\mathscr{C}_n)^\circ_{++}) = D(\rho\|I). \tag{84}$$

Since the first term on the left-hand side is subadditive (see [10, Lemma 26]), its limit exists when n goes to infinity by Fekete's lemma. By the equality relation, the limit for the second term on the left-hand side also exists. Therefore, we have

$$\left[\lim_{n\to\infty}\frac{1}{n}D(\rho^{\otimes n}\|\mathscr{C}_n)\right] + \left[\lim_{n\to\infty}\frac{1}{n}D_{\mathrm{M}}(\rho^{\otimes n}\|(\mathscr{C}_n)^\circ_{++})\right] = D(\rho\|I). \tag{85}$$

Finally, we complete the proof by noting that the second term on the left-hand side is equal to the one with quantum relative entropy (see [10, Lemma 28]). ■

V. Discussions

In this work, we have established Uhlmann's theorem for general measured f-divergences, which includes the measured Rényi divergences as special cases and generalizes the well-known Uhlmann's theorem for the fidelity. The significance and potential applications are as follows:

- (**Fundamental importance:**) Uhlmann's theorem is a cornerstone in quantum information theory. It has applications in various subareas, ranging from quantum Shannon theory [25] and quantum computing [35] to quantum gravity [36]. A concrete example for the latter is the construction of the holographic dual of the bulk symplectic form in an entanglement wedge. The construction relies on evaluating the fidelity via a particular replica trick and uses Uhlmann's theorem as a key ingredient. This naturally prompts the question (see Footnote 4 of [36]) of whether the construction extends to Rényi divergences and, if so, whether the resulting conclusions remain consistent. Moreover, having an Uhlmann's theorem for a quantum divergence is of fundamental interest, as it serves as a criterion for classifying divergences into different types. Note that Uhlmann's theorem cannot hold for any "sufficient" quantum divergence (in the sense that saturation of the data processing inequality $\mathbb{D}(\Phi(\rho)\|\Phi(\sigma)) = \mathbb{D}(\rho\|\sigma)$ implies the reversibility of Φ on $\{\rho,\sigma\}$). Since most commonly used quantum Rényi divergences, including the Petz and sandwiched Rényi divergences, are sufficient in this sense, they cannot satisfy Uhlmann's theorem (except for degenerate cases). This fundamentally distinguishes measured f-divergences from other quantum divergences and highlights their unique mathematical structure.
- (**Existing application:**) In the case of $\alpha > 1$, the Uhlmann's theorem for regularized sandwiched Rényi divergence proved in [37] already plays a crucial role in deriving the generalized entropy accumulation theorem, which is then used in quantum cryptographic tasks such as blind randomness expansion. Our work establishes Uhlmann's theorem directly for the measured Rényi divergence in the single-letter case, which provides a conceptually clearer and technically simpler proof of the chain rule in [37]. Our Uhlmann theorem for measured divergences can potentially be used to obtain improved chain rules, e.g. improving the measured divergence term in [38, Corollary 4] by taking partial traces when the channel $\mathcal{F}$ satisfies a non-signalling property. We leave this for future work.
- (**Potentially new application:**) The Uhlmann's theorem for measured divergence provides a potential approach to computing the amortized relative entropy for quantum channels. Let $\mathcal{N}, \mathcal{M}$ be two quantum channels; their amortized relative entropy is defined as

$$D^A(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{RA},\sigma_{RA}} D(\mathcal{N}_{A\to B}(\rho_{RA})\|\mathcal{M}_{A\to B}(\sigma_{RA})) - D(\rho_{RA}\|\sigma_{RA}), \tag{86}$$

where the supremum is over all bipartite states ρ_{RA}, σ_{RA} with R being a reference system of arbitrary dimension. This quantity serves as the ultimate limit for the asymptotic quantum channel discrimination problem, but its computability remains open [39], [40]. The main challenge for this is the potentially unbounded dimension of the reference system R in the optimization. However, our Uhlmann's theorem implies that the dimension in an analogous formula for measured relative entropy can be restricted to the same size as system A. Specifically, if we define the amortized measured relative entropy for channels by

$$D^A_{\mathrm{M}}(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{RA},\sigma_{RA}} D_{\mathrm{M}}(\mathcal{N}_{A\to B}(\rho_{RA})\|\mathcal{M}_{A\to B}(\sigma_{RA})) - D_{\mathrm{M}}(\rho_{RA}\|\sigma_{RA}), \tag{87}$$

and consider a purification of ρ_{RA} as $|\phi\rangle_{ERA}$, then by Uhlmann's theorem, there exists an extension σ_{ERA} for σ_{RA} such that $D_{\mathrm{M}}(\phi_{ERA}\|\sigma_{ERA}) = D_{\mathrm{M}}(\rho_{RA}\|\sigma_{RA})$. Moreover, by the data-processing inequality, we have

$$D_{\mathrm{M}}(\mathcal{N}_{A\to B}(\rho_{RA})\|\mathcal{M}_{A\to B}(\sigma_{RA})) \le D_{\mathrm{M}}(\mathcal{N}_{A\to B}(\phi_{ERA})\|\mathcal{M}_{A\to B}(\sigma_{ERA})). \tag{88}$$

This implies that

$$D^A_{\mathrm{M}}(\mathcal{N}\|\mathcal{M}) = \sup_{\phi_{ERA},\sigma_{ERA}} D_{\mathrm{M}}(\mathcal{N}_{A\to B}(\phi_{ERA})\|\mathcal{M}_{A\to B}(\sigma_{ERA})) - D_{\mathrm{M}}(\phi_{ERA}\|\sigma_{ERA}). \tag{89}$$

That is, the optimization can be restricted to pure states ϕ_{ERA}. Rewriting the systems ER as R, we get

$$D^A_{\mathrm{M}}(\mathcal{N}\|\mathcal{M}) = \sup_{\phi_{RA},\sigma_{RA}} D_{\mathrm{M}}(\mathcal{N}_{A\to B}(\phi_{RA})\|\mathcal{M}_{A\to B}(\sigma_{RA})) - D_{\mathrm{M}}(\phi_{RA}\|\sigma_{RA}). \tag{90}$$

Suppose $|R| \ge |A|$ and let A' be isomorphic to A. For any two purifications ϕ_{RA} and $\phi_{A'A}$ of ϕ_A, there exists an isometry $V_{R\to A'}$ such that $|\phi\rangle_{A'A} = V_{R\to A'}|\phi\rangle_{RA}$. By isometry invariance of the measured relative entropy, we get

$$D^A_{\mathrm{M}}(\mathcal{N}\|\mathcal{M}) = \sup_{\phi_{A'A},\sigma_{A'A}} D_{\mathrm{M}}(\mathcal{N}_{A\to B}(\phi_{A'A})\|\mathcal{M}_{A\to B}(\sigma_{A'A})) - D_{\mathrm{M}}(\phi_{A'A}\|\sigma_{A'A}), \tag{91}$$

with $|A'| = |A|$. Let $D(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{RA}} D(\mathcal{N}_{A\to B}(\rho_{RA})\|\mathcal{M}_{A\to B}(\rho_{RA}))$ and its regularization $D^\infty(\mathcal{N}\|\mathcal{M}) := \lim_{n\to\infty}\frac{1}{n}D(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$. We can have similar definitions for D_{M}. Then we have the relations that

$$\frac{1}{n}D_{\mathrm{M}}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) \le D^\infty(\mathcal{N}\|\mathcal{M}) \tag{92}$$

$$= D^\infty_{\mathrm{M}}(\mathcal{N}\|\mathcal{M}) \tag{93}$$

$$\le \frac{1}{n}D^A_{\mathrm{M}}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}), \tag{94}$$

where the first inequality follows as $D_{\mathrm{M}}(\cdot\|\cdot) \le D(\cdot\|\cdot)$, the equality follows by the asymptotic equivalence of measured relative entropy and quantum relative entropy for permutation-invariant states (e.g. [10, Lemma 16, 17]), and the last inequality follows from the fact that $D_{\mathrm{M}}(\mathcal{N}\|\mathcal{M}) \le D^A_{\mathrm{M}}(\mathcal{N}\|\mathcal{M})$ and the subadditivity of $D^A_{\mathrm{M}}(\mathcal{N}\|\mathcal{M})$ by definition. Note that $D^\infty(\mathcal{N}\|\mathcal{M}) = D^A(\mathcal{N}\|\mathcal{M})$ as proved in [40]. Therefore, we have

$$\frac{1}{n}D_{\mathrm{M}}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}) \le D^A(\mathcal{N}\|\mathcal{M}) \tag{95}$$

$$\le \frac{1}{n}D^A_{\mathrm{M}}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n}). \tag{96}$$

It is clear that the lower bound $\frac{1}{n}D_{\mathrm{M}}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$ converges to $D^\infty(\mathcal{N}\|\mathcal{M})$ and therefore $D^A(\mathcal{N}\|\mathcal{M})$. If we can show that the upper bound converges to $D^A(\mathcal{N}\|\mathcal{M})$, then we would have established a sandwiching of $D^A(\mathcal{N}\|\mathcal{M})$ between two sequences converging to it, and both sequences can be computed in finite time. This would imply the computability of $D^A(\mathcal{N}\|\mathcal{M})$; that is, there exists an algorithm that terminates in finite time within additive error.

ACKNOWLEDGMENT

The authors would like to thank David Sutter for a comment that helped them clarify Corollary 2 and Remark 7. They also thank anonymous reviewers for their valuable comments and for bringing Hiai's book to their attention.

REFERENCES

[1] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkeley Symp. Math. Statist. Probab., Contrib. Theory Statist.*, vol. 4. Berkeley, CA, USA: Univ. California Press, 1961, pp. 547–562.
[2] I. Csiszár, "On information-type measure of difference of probability distributions and indirect observations," *Studia Sci. Math. Hungar.*, vol. 2, pp. 299–318, Jan. 1967.
[3] S. M. Ali and S. D. Silvey, "A general class of coefficients of divergence of one distribution from another," *J. Roy. Stat. Soc. Ser. B, Stat. Methodol.*, vol. 28, no. 1, pp. 131–142, Jan. 1966.
[4] Y. Polyanskiy and Y. Wu, *Information Theory: From Coding to Learning*. Cambridge, U.K.: Cambridge Univ. Press, 2025.
[5] F. Hiai and M. Mosonyi, "Different quantum f-divergences and the reversibility of quantum operations," *Rev. Math. Phys.*, vol. 29, no. 7, 2017, Art. no. 1750023.
[6] M. J. Donald, "On the relative entropy," *Commun. Math. Phys.*, vol. 105, no. 1, pp. 13–34, 1986.
[7] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Commun. Math. Phys.*, vol. 143, no. 1, pp. 99–114, Dec. 1991.
[8] A. Uhlmann, "The 'transition probability' in the state space of a*-algebra," *Rep. Math. Phys.*, vol. 9, no. 2, pp. 273–279, 1976.
[9] T. Metger, O. Fawzi, D. Sutter, and R. Renner, "Generalised entropy accumulation," *Commun. Math. Phys.*, vol. 405, no. 11, p. 261, Nov. 2024.
[10] K. Fang, H. Fawzi, and O. Fawzi, "Generalized quantum asymptotic equipartition," 2024, *arXiv:2411.04035*.
[11] K. Fang, H. Fawzi, and O. Fawzi, "Efficient approximation of regularized relative entropies and applications," 2025, *arXiv:2502.15659*.
[12] K. Fang, H. Fawzi, and O. Fawzi, "Adversarial quantum channel discrimination," *Phys. Rev. Lett.*, vol. 135, no. 26, Dec. 2025, Art. no. 260201, doi: 10.1103/mtbj-grbk.
[13] A. Arqand and E. Y.-Z. Tan, "Marginal-constrained entropy accumulation theorem," 2025, *arXiv:2502.02563*.
[14] M. Mosonyi and T. Ogawa, "Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies," *Commun. Math. Phys.*, vol. 334, no. 3, pp. 1617–1648, Mar. 2015.
[15] G. Mazzola, D. Sutter, and R. Renner, "Uhlmann's theorem for relative entropies," *IEEE Trans. Inf. Theory*, vol. 71, no. 9, pp. 7039–7051, Feb. 2025.
[16] M. Berta, O. Fawzi, and M. Tomamichel, "On variational expressions for quantum relative entropies," *Lett. Math. Phys.*, vol. 107, no. 12, pp. 2239–2265, Dec. 2017.
[17] D. Sutter, M. Berta, and M. Tomamichel, "Multivariate trace inequalities," *Commun. Math. Phys.*, vol. 352, no. 1, pp. 37–58, May 2017.

[18] F. Hiai, "Quantum f-divergences in von Neumann algebras," in *Mathematical Physics Studies*. Singapore: Springer, 2021.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[20] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1970, doi: 10.1515/9781400873173.

[21] D. Petz, *Quantum Information Theory and Quantum Statistics*. Cham, Switzerland: Springer, 2008.

[22] F. Hansen and G. K. Pedersen, "Jensen's operator inequality," *Bull. London Math. Soc.*, vol. 35, no. 4, pp. 553–564, 2003.

[23] K. Fang, X. Wang, M. Tomamichel, and M. Berta, "Quantum channel simulation and the channel's smooth max-information," *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 2129–2140, Apr. 2020.

[24] M. Sion, "On general minimax theorems," *Pacific J. Math.*, vol. 8, no. 1, pp. 171–176, Mar. 1958. [Online]. Available: https://msp.org/pjm/1958/8-1/p14.xhtml

[25] J. Watrous, *The Theory of Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2018.

[26] A. Kurdila and M. Zabarankin, *Convex Functional Analysis* (Systems & Control: Foundations & Applications). Basel, Switzerland: Birkhäuser, 2005. [Online]. Available: https://books.google.com.hk/books?id=OoMHPdsafUYC

[27] M. Mosonyi and F. Hiai, "Some continuity properties of quantum Rényi divergences," *IEEE Trans. Inf. Theory*, vol. 70, no. 4, pp. 2674–2700, Apr. 2024.

[28] M. Mosonyi and F. Hiai, "On the quantum Rényi relative entropies and related capacity formulas," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2474–2487, Apr. 2011.

[29] C. A. Fuchs, "Distinguishability and accessible information in quantum theory," 1996, *arXiv:quant-ph/9601020*.

[30] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," *J. Math. Phys.*, vol. 54, no. 12, Dec. 2013, Art. no. 122203.

[31] M. M. Wilde, A. Winter, and D. Yang, "Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy," *Commun. Math. Phys.*, vol. 331, no. 2, pp. 593–622, Oct. 2014.

[32] M. Hayashi and M. Tomamichel, "Correlation detection and an operational interpretation of the Rényi mutual information," *J. Math. Phys.*, vol. 57, no. 10, 2016, Art. no. 102201.

[33] G. Aubrun and S. J. Szarek, *Alice and Bob Meet Banach*, vol. 223. Providence, RI, USA: American Mathematical Society, 2017.

[34] P. E. Frenkel, "Integral formula for quantum relative entropy implies data processing inequality," *Quantum*, vol. 7, p. 1102, Sep. 2023.

[35] J. Bostanci, Y. Efron, T. Metger, A. Poremba, L. Qian, and H. Yuen, "Unitary complexity and the uhlmann transformation problem," 2023, *arXiv:2306.13073*.

[36] J. Kirklin, "The holographic dual of the entanglement wedge symplectic form," *J. High Energy Phys.*, vol. 2020, no. 1, p. 71, Jan. 2020, doi: 10.1007/jhep01(2020)071.

[37] T. Metger, O. Fawzi, D. Sutter, and R. Renner, "Generalised entropy accumulation," in *Proc. IEEE 63rd Annu. Symp. Found. Comput. Sci. (FOCS)*, Oct. 2022, pp. 844–850.

[38] M. Berta and M. Tomamichel, "Chain rules for quantum channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2022, pp. 2427–2432.

[39] X. Wang and M. M. Wilde, "Resource theory of asymmetric distinguishability for quantum channels," *Phys. Rev. Res.*, vol. 1, no. 3, Dec. 2019, Art. no. 033169.

[40] K. Fang, O. Fawzi, R. Renner, and D. Sutter, "Chain rule for the quantum relative entropy," *Phys. Rev. Lett.*, vol. 124, no. 10, Mar. 2020, Art. no. 100501.

Kun Fang (Member, IEEE) received the B.S. degree in mathematics from Wuhan University in 2015 and the Ph.D. degree in quantum information from the University of Technology Sydney in 2018. After that, he was a Post-Doctoral Researcher with the University of Cambridge and the University of Waterloo from 2018 to 2020. He was a Senior Researcher and the Tech Lead with the Institute for Quantum Computing, Baidu, from 2020 to 2023. He is currently an Assistant Professor with the School of Data Science, The Chinese University of Hong Kong, Shenzhen. His research interests include understanding the capabilities and limitations of quantum resources, as well as their usage in quantum computation and quantum communication.

Hamza Fawzi received the Ph.D. degree from MIT in 2016. Since then he has been with the University of Cambridge, where is currently a Professor of applied mathematics. His research interests include convex optimization and its applications to quantum information and many-body theory.

Omar Fawzi received the Ph.D. degree in computer science from McGill University in 2012. From 2012 to 2014, he was a Post-Doctoral Researcher with ETH Zürich. From 2014 to 2020, he was an Associate Professor of computer science with ENS de Lyon, where he is currently the Research Director with Inria. His research interests include classical and quantum information, with an emphasis on algorithmic questions.