Journal of Physics A: Mathematical and Theoretical

PAPER

Recoverability for optimized quantum *f*-divergences

To cite this article: Li Gao and Mark M Wilde 2021 *J. Phys. A: Math. Theor.* **54** 385302

IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **54** (2021) 385302 (58pp) https://doi.org/10.1088/1751-8121/ac1dc2

Recoverability for optimized quantum *f*-divergences

Li Gao[1,*] **and Mark M Wilde**[2,3]

[1] Department of Mathematics, Texas A & M University, College Station, Texas 77840, United States of America

[2] Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, and Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, United States of America

[3] Stanford Institute for Theoretical Physics, Stanford University, Stanford, California 94305, United States of America

E-mail: ligao@math.tamu.edu and mwilde@lsu.edu

Received 27 November 2020, revised 3 August 2021
Accepted for publication 16 August 2021
Published 29 September 2021

CrossMark

Abstract

The optimized quantum f-divergences form a family of distinguishability measures that includes the quantum relative entropy and the sandwiched Rényi relative quasi-entropy as special cases. In this paper, we establish physically meaningful refinements of the data-processing inequality for the optimized f-divergence. In particular, the refinements state that the absolute difference between the optimized f-divergence and its channel-processed version is an upper bound on how well one can recover a quantum state acted upon by a quantum channel, whenever the recovery channel is taken to be a rotated Petz recovery channel. Not only do these results lead to physically meaningful refinements of the data-processing inequality for the sandwiched Rényi relative entropy, but they also have implications for perfect reversibility (i.e. quantum sufficiency) of the optimized f-divergences. Along the way, we improve upon previous physically meaningful refinements of the data-processing inequality for the standard f-divergence, as established in recent work of Carlen and Vershynina [arXiv:1710.02409, arXiv:1710.08080]. Finally, we extend the definition of the optimized f-divergence, its data-processing inequality, and all of our recoverability results to the general von Neumann algebraic setting, so that all of our results can be employed in physical settings beyond those confined to the most common finite-dimensional setting of interest in quantum information theory.

Keywords: relative entropy, data processing inequality, quantum f-divergence, recoverability, quantum sufficiency, Rényi divergence

*Author to whom any correspondence should be addressed.

1. Introduction

The quantum relative entropy is a fundamental measure in quantum information theory. It was first introduced by Umegaki [Ume62] as a noncommutative generalization of the classical relative entropy (the latter is also called Kullback–Leibler divergence [KL51]). For two quantum states described by density operators ρ and σ, the relative entropy of ρ with respect to σ is defined as

$$D(\rho\|\sigma) := \mathrm{tr}\,(\rho\,\log\,\rho - \rho\,\log\,\sigma)\,,$$

where tr denotes the matrix trace. The relative entropy $D(\rho\|\sigma)$ measures how well the quantum state ρ can be distinguished from σ in an asymptotic setting of quantum hypothesis testing [HP91, ON00]. One of its most important properties is the data-processing inequality [Lin75, Uhl77]: for all quantum channels Φ and states ρ and σ, the following inequality holds

$$D(\rho\|\sigma) \geqslant D(\Phi(\rho)\|\Phi(\sigma)). \tag{1}$$

As the quantum relative entropy is a distinguishability measure, the data-processing inequality asserts that two quantum states cannot become more distinguishable after applying the same quantum channel to them. The data-processing inequality is a key principle underlying the widespread applications of quantum relative entropy in quantum information [Ved02, Wil17].

The wide interest in relative entropy has sparked researchers to study other entropy-type measures that also satisfy the data-processing inequality. Important generalizations in classical information theory are the Rényi relative entropy [Rén61] and the more general notion of f-divergence [AS66, Csi67, Mor63]. For two probability distributions $\{p(x)\}_x$ and $\{q(x)\}_x$ and a convex function f, the classical f-divergence [AS66, Csi67, Mor63] is defined as

$$S_f(p\|q) := \sum_x p(x) f\left(\frac{q(x)}{p(x)}\right),$$

and it satisfies the data-processing inequality for classical channels. In [Pet85, Pet86a], Petz introduced a quantum version of the f-divergence and proved that the quantum f-divergence satisfies the data-processing inequality whenever the underlying function f is *operator convex*. One notable example is the Petz–Rényi relative quasi-entropy [Pet85, Pet86a], which corresponds to $f(t) = t^s$ for $s \in (-1, 0) \cup (0, 1)$, i.e. the power function. From this quantity, the Petz–Rényi relative entropy can be defined, and it has an operational interpretation in quantum hypothesis testing [Hay07, Nag06].

In recent years, the sandwiched Rényi relative entropy [MLDS+13, WWY14] was introduced as another quantum generalization of Rényi relative entropy and has found extensive application in establishing strong converse results for communication tasks [CMW16, DW18, GW15, TWW16, WWY14, WTB17]. It also has a direct operational meaning in quantum hypothesis testing in terms of the strong converse exponent [MO15]. While Petz's definition of quantum f-divergence from [Pet85, Pet86a] is often called the standard f-divergence, it was not clear how to express the sandwiched Rényi relative entropy in terms of a standard f-divergence. This problem was solved in [Wil18a] with the introduction of a different type of quantum f-divergence called the optimized f-divergence. It was also proved in [Wil18a] that the optimized f-divergence satisfies the data-processing inequality for an operator anti-monotone function f.

Over decades, the data-processing inequality of the quantum relative entropy has been refined in various ways. Petz proved that the data-processing inequality in (1) is saturated, i.e. $D(\rho\|\sigma) = D(\Phi(\rho)\|\Phi(\sigma))$, if and only if there exists a quantum recovery channel R satisfying $(R\circ\Phi)(\rho) = \rho$ and $(R\circ\Phi)(\sigma) = \sigma$ [Pet86b, Pet88]. The latter condition is also called 'quantum sufficiency' [Pet86b] because it indicates that the pair $(\Phi(\rho), \Phi(\sigma))$ is just as good as the pair (ρ, σ) in a distinguishability experiment. Moreover, there is a canonical choice of the recovery channel R, now called the Petz recovery map, which is given by

$$R_{\Phi,\sigma}(x) = \sigma^{1/2}\Phi^{\dagger}(\Phi(\sigma)^{-1/2}x\Phi(\sigma)^{-1/2})\sigma^{1/2}, \tag{2}$$

where $\Phi^{\dagger}$ is the adjoint of Φ with respect to the Hilbert–Schmidt inner product.

More recently, much progress has been made on the case of approximate recovery. The idea is that when the data-processing inequality is nearly saturated, then the states (ρ, σ) can be approximately recovered from $(\Phi(\rho), \Phi(\sigma))$ by the action of some quantum channel R. The first precise quantitative result of approximate recovery was obtained in [FR15] for the special case of Φ being a partial trace and σ being a marginal of ρ (this specialized setting is relevant for an information measure called conditional mutual information). The result of [FR15] has been generalized in [JRS+18, SBT17, STH16, Wil15]. In particular, it was proved in [JRS+18] that the following inequality holds for a universal recovery map R:

$$D(\rho\|\sigma) \geqslant D(\Phi(\rho)\|\Phi(\sigma)) - \log F(\rho, (R\circ\Phi)(\rho)), \tag{3}$$

while the equality $(R\circ\Phi)(\sigma) = \sigma$ holds also. In (3) above, F denotes the Uhlmann fidelity [Uhl76] (defined later in (8)) and the recovery map R is explicitly given as follows:

$$R := \int_{\mathbb{R}} R^{\frac{t}{2}}_{\Phi,\sigma}\ \mathrm{d}\beta(t)\ , \qquad R^{t}_{\Phi,\sigma}(x) := \sigma^{-it}R_{\Phi,\sigma}(\Phi(\sigma)^{it}x\Phi(\sigma)^{-it})\sigma^{it}\ , \tag{4}$$

where $R_{\Phi,\sigma}$ is the original Petz map in (2), $R^{t}_{\Phi,\sigma}(x)$ is called a rotated Petz map [Wil15], and R is the expectation of R_t with respect to the following probability density function:

$$\mathrm{d}\beta(t) = \frac{\pi}{2}(\cosh(\pi t) + 1)^{-1}\mathrm{d}t.$$

The recovery map R in (4) is said to be 'universal' because it does not depend on the ρ state; this property is useful in a variety of physical applications such as quantum error correction [JRS+18]. Note that a slightly stronger inequality than the one in (3) is available in [JRS+18].

Most recently, the main result of [JRS+18] has been extended to the von Neumann algebraic setting [FHSW20], and references [CV18, Ver19] established an approximate recovery estimate for the original Petz map. The method of [CV18, Ver19] is based on the integral representation of operator convex functions and further applies to the case of approximate recoverability for standard f-divergences, as well as to the case of Petz–Rényi relative entropies. A similar method has been employed to understand refinements of the data-processing inequality for the maximal f-divergences [BC20].

2. Summary of results

In this paper, we study approximate recoverability for optimized f-divergences and contribute the following findings:

(a) We prove that the difference of optimized f-divergences before and after the action of a quantum channel is an upper bound on the recoverability error for rotated Petz recovery maps (see lemma 4.19). Since the sandwiched Rényi relative quasi-entropy is a special kind of optimized f-divergence [Wil18a], our result gives the first quantitative estimate for approximate recoverability with respect to the sandwiched Rényi relative (quasi-)entropies (see theorem 4.20 and corollary 4.21). The method that we employ here is inspired by [CV18, CV20a, Ver19].

(b) As a corollary, we find the following reversibility result: if the optimized f-divergence is preserved under the action of a quantum channel, then every rotated Petz map is a perfect recovery map (see corollary 4.23). This extends previous reversibility results found for the sandwiched Rényi relative entropy [HM17, Jen17a] (see also [CV20b, LRD17, Zha20] for related conditions regarding the saturation of the data-processing inequality for the sandwiched Rényi relative entropy).

(c) We also improve the results of [CV18, CV20a] for the quantum and Petz–Rényi relative entropies and further generalize these prior results to rotated Petz maps (see theorems 4.5 and 4.7, corollary 4.9, theorems 4.13 and 4.15, and corollary 4.16). One advantage of these new bounds over the previous ones from [CV18, CV20a] is that the remainder term involves the Petz–Rényi relative entropy of order two, rather than the operator norm of the relative modular operator. As such, these bounds are non-trivial for the important class of bosonic Gaussian states [Ser17], whereas the previous bounds from [CV18, CV20a] do not apply for this class of states.

(d) Motivated by the recent works on quantum f-divergences in general von Neumann algebras [Hia18, Hia19], we extend the definition of optimized f-divergence, its data-processing inequality, and our recoverability results to the general context of von Neumann algebras (see definition 5.1, theorems 5.10 and 5.15, and corollary 5.16). Our results also provide a new way for understanding the sandwiched Rényi relative entropy in the von Neumann algebraic setting. Note that the sandwiched Rényi relative entropy was previously defined and analyzed in the von Neumann algebraic setting [BST18, Jen17b, Jen18]. Later on, it was analyzed under a different approach [GYZ19] and studied in the context of conformal field theory [Las19].

The rest of our paper is organized as follows. Section 3 reviews the basic definitions of operator monotone and operator convex functions, quantum (optimized) f-divergences, and (rotated) Petz recovery maps. In section 4, we discuss our main recoverability results in the finite-dimensional setting, while focusing on quantum channels that act as restrictions to a subalgebra. This is the core case, and the argument here avoids technicalities that occur in infinite dimensions. We prove that the recoverability error for a rotated Petz recovery map can be bounded from above by a difference of (optimized) f-divergences. Section 5 is devoted to the optimized f-divergence in general von Neumann algebras. We prove the data-processing inequality and extend our recoverability results to a general quantum channel in this setting.

3. Preliminaries

3.1. Operator convex functions and operator monotone functions

We briefly review the integral representation of operator monotone and operator convex functions. We refer to [Bha13] for more information on this topic.

Let $B(H)$ denote the set of bounded operators acting on a Hilbert space H. An operator $A \in B(H)$ is positive if $\langle v| A |v\rangle \geqslant 0$ for all $|v\rangle \in H$. Let $B(H)^+$ denote the set of positive operators. A function $f : (0, \infty) \to \mathbb{R}$ is operator monotone if the following inequality holds for all invertible positive operators $A, B \in B(H)$ satisfying $A \leqslant B$:

$$f(A) \leqslant f(B) .$$

We say that f is operator convex if the following inequality holds for all invertible positive operators A, B and $\lambda \in [0, 1]$:

$$f(\lambda A + (1-\lambda)B) \leqslant \lambda f(A) + (1-\lambda) f(B)\ .$$

We say that f is operator anti-monotone (resp. concave) if $-f$ is operator monotone (resp. convex). It is known that $f : (0,\infty) \to \mathbb{R}$ is operator concave if it is operator monotone. A function $f : (0,\infty) \to \mathbb{R}$ is operator convex if and only if for all invertible positive operators $A \in B(H)^+$ and Hilbert-space isometries $V : K \to H$, the following inequality holds

$$V^* f(A) V \geqslant f(V^* A V)\ .$$

This inequality is known as the operator Jensen inequality and also extends to positive A (see [Pet85, appendix], as well as [HP03]).

By the Löwner theorem (cf [Bha13, p 144]), an operator monotone function $f : (0,\infty) \to \mathbb{R}$ admits the following integral representation:

$$f(t) = a + bt + \int_0^\infty \left(\frac{\lambda}{\lambda^2+1} - \frac{1}{\lambda+t} \right) \mathrm{d}\nu(\lambda), \tag{5}$$

where $a \in \mathbb{R}$, $b \geqslant 0$, and ν is a positive measure on $[0,\infty)$ such that $\int_0^\infty \frac{\lambda}{\lambda^2+1} \mathrm{d}\nu(\lambda) < \infty$.

Example 3.1. Below we list several important examples of functions $f : (0,\infty) \to \mathbb{R}$ that are either operator monotone or operator anti-monotone.

(a) $f(t) = (\lambda + t)^{-1}$ is operator anti-monotone and operator convex for $\lambda \geqslant 0$. This corresponds to $a = b = 0$ and μ being the point measure at λ.

(b) Let $0 < r < 1$. The power function $t \mapsto t^r$ is operator monotone and operator concave by the following integral representation:

$$t^r = \frac{\sin(r\pi)}{\pi} \int_0^\infty \left(\frac{1}{\lambda} - \frac{1}{\lambda+t} \right) \lambda^r \, \mathrm{d}\lambda\ ,$$

where $\mathrm{d}\lambda$ is the Lebesgue measure on $\mathbb{R}$. On the other hand, $t \mapsto t^{-r}$ is operator anti-monotone because it is a composition of $t \mapsto t^{-1}$ and $t \mapsto t^r$, with the former being operator anti-monotone and the latter operator monotone. It is thus also operator convex. The integral representation of t^{-r} is

$$t^{-r} = \frac{\sin(r\pi)}{\pi} \int_0^\infty \lambda^{-r} \frac{1}{\lambda+t} \, \mathrm{d}\lambda\ .$$

(c) The logarithm function $f(t) = \log t$ is operator monotone and operator concave. These statements are a consequence of the following integral representation of $\log t$:

$$\log\, t = -\int_0^\infty \left(\frac{1}{\lambda+t} - \frac{\lambda}{\lambda^2+1} \right) \mathrm{d}\lambda\ .$$

Following [Ver19], we say that an operator monotone or operator anti-monotone function f is regular if the measure $\nu(\lambda)$ in its integral representation is absolutely continuous with respect to the Lebesgue measure $\mathrm{d}\lambda$ and for all $0 < a < b < \infty$, there is a constant $C_{a,b}$ such that $\mathrm{d}\lambda \leqslant C_{a,b} \mathrm{d}\nu(\lambda)$ on (a, b). All examples (b), (c), and (d) given above are regular operator monotone (or anti-monotone) functions.

3.2. Standard and optimized f-divergences

Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra equipped with a faithful trace τ. For the standard quantum information setup, one can take $\mathcal{M} = B(H)$ for a finite-dimensional Hilbert space H and tr the matrix trace. For $p \in [1, \infty)$, the L_p-norm for $a \in \mathcal{M}$ is defined as $\|a\|_p := \tau(|a|^p)^{1/p}$. We use the same notation and definition for $p \in (0, 1)$, when $\|a\|_p$ is a quasi-norm. We identify $L_1(\mathcal{M}) \cong \mathcal{M}_*$ and $L_\infty(\mathcal{M}) \cong \mathcal{M}$. A state ρ is given by a density operator $\rho \in L_1(\mathcal{M})$ with $\rho \geqslant 0$ and $\tau(\rho) = 1$. We denote the state space of $\mathcal{M}$ by $D(\mathcal{M}) := \{\rho \in L_1(\mathcal{M}) \,|\, \rho \geqslant 0, \tau(\rho) = 1\}$ and the set of invertible density operators by $D_+(\mathcal{M})$. The L_2-space $L_2(\mathcal{M})$ is a Hilbert space with the following trace inner product:

$$\langle x, y\rangle := \tau(x^* y) \ .$$

Let $|x\rangle$ denote the vector in the GNS space $L_2(\mathcal{M})$ corresponding to the element x. The vector $|1\rangle$ corresponding to the identity operator is an analog of the (unnormalized) maximally entangled state. The GNS representation $\pi : \mathcal{M} \to B(L_2(\mathcal{M}))$ is given by

$$\pi(a)\,|x\rangle = |ax\rangle \ .$$

We often omit π and write $a\,|x\rangle := \pi(a)\,|x\rangle$. A state ρ admits a vector representation by $\left|\rho^{1/2}\right\rangle$ (also called a purification) that satisfies

$$\rho(x) = \tau(\rho x) = \left\langle \rho^{1/2} \middle| \, x \, \middle| \rho^{1/2} \right\rangle \ .$$

For simplicity of notation, we sometimes write $|\boldsymbol{\rho}\rangle = \left|\rho^{1/2}\right\rangle$. Let ρ and σ be two states. Let $s(\rho)$ and $s(\sigma)$ denote the support projections onto the supports of ρ and σ, respectively, and let ρ^{-1} denote the inverse of ρ on its support. The relative modular operator is defined as

$$\Delta(\sigma, \rho)\,|x\rangle := \left|\sigma x \rho^{-1}\right\rangle \ ,$$

which for faithful ρ and σ is always a positive and invertible operator on $L_2(\mathcal{M})$.

Let $f : (0, \infty) \to \mathbb{R}$ be an operator anti-monotone function. Given two states ρ and σ with $s(\rho) \leqslant s(\sigma)$, the standard f-divergence is defined as [Pet85, Pet86a] (see also [HMPB11])

$$Q_f(\rho\|\sigma) := \langle \boldsymbol{\rho}|\, f(\Delta(\sigma, \rho))\, |\boldsymbol{\rho}\rangle \ ,$$

where $f(\Delta(\sigma, \rho))$ makes use of the functional calculus, as applied to $\Delta(\sigma, \rho)$. The optimized f-divergence is defined as [Wil18a]

$$\widetilde{Q}_f(\rho\|\sigma) := \sup_{\omega \in D_+(\mathcal{M})} \langle \boldsymbol{\rho}|\, f(\Delta(\sigma, \omega))\, |\boldsymbol{\rho}\rangle \, . \tag{6}$$

We review some important examples of relative entropies defined through standard or optimized f-divergences.

(a) Umegaki relative entropy [Ume62]:

$$D(\rho\|\sigma) := -\langle \boldsymbol{\rho}|\, \log\, \Delta(\sigma, \rho)\, |\boldsymbol{\rho}\rangle = \tau(\rho\, \log\, \rho - \rho\, \log\, \sigma) = Q_{-\log\, x}(\rho\|\sigma). \tag{7}$$

(b) Petz–Rényi relative entropy [Pet85, Pet86a]:

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\,\tau(\rho^\alpha\sigma^{1-\alpha})$$
$$= \frac{1}{\alpha-1}\log\,\langle\boldsymbol{\rho}|\,\Delta(\sigma,\rho)^{1-\alpha}\,|\boldsymbol{\rho}\rangle = \frac{1}{\alpha-1}\log\,Q_{x^{1-\alpha}}(\rho\|\sigma)\;.$$

This quantity is defined for $\alpha\in(0,1)\cup(1,\infty)$. A special case of interest for some of the remainder terms in the entropy inequalities in section 4 occurs when $\alpha=2$ or $\alpha=-1$, for which $Q_{x^{-1}}(\rho\|\sigma)=\tau(\rho^2\sigma^{-1})$ or $Q_{x^2}(\rho\|\sigma)=\tau(\rho^{-1}\sigma^2)$, respectively.

(c) Sandwiched Rényi relative entropy [MLDS+13, WWY14]: for $\alpha>1$,

$$\widetilde{D}_\alpha(\rho\|\sigma) := \frac{\alpha}{\alpha-1}\log\left\|\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right\|_\alpha = \frac{\alpha}{\alpha-1}\log\widetilde{Q}_{x^{\frac{1-\alpha}{\alpha}}}(\rho\|\sigma),$$

where the last equality was identified in [Wil18a]. Also, for $0<\alpha<1$

$$\widetilde{D}_\alpha(\rho\|\sigma) = \frac{\alpha}{\alpha-1}\log\Big(-\widetilde{Q}_{-x^{\frac{1-\alpha}{\alpha}}}(\rho\|\sigma)\Big)\,.$$

The sandwiched Rényi relative entropy is defined for $\alpha\in(0,1)\cup(1,\infty)$.

(d) Holevo fidelity [Hol72]:

$$F_{\mathrm{H}}(\rho,\sigma) := \tau(\rho^{1/2}\sigma^{1/2})^2 := \left\langle\rho^{1/2}\middle|\,\Delta(\sigma,\rho)^{1/2}\middle|\rho^{1/2}\right\rangle^2 = Q_{x^{1/2}}(\rho\|\sigma)^2\;.$$

Observe that $D_{1/2}(\rho\|\sigma) = -\log F_{\mathrm{H}}(\rho,\sigma)$.

(e) Uhlmann fidelity [Uhl76]:

$$F(\rho,\sigma) := \tau(|\sqrt{\rho}\sqrt{\sigma}|)^2 = \left\|\rho^{1/2}\sigma\rho^{1/2}\right\|_{1/2}$$
$$= \inf_{\omega\in\mathcal{D}(\mathcal{M}_+)}\tau(\rho^{1/2}\sigma\rho^{1/2}\omega^{-1}) = \widetilde{Q}_x(\rho\|\sigma)^2. \tag{8}$$

Observe that $\widetilde{D}_{1/2}(\rho\|\sigma) = -\log\,F(\rho,\sigma)$.

3.3. Petz map and rotated Petz map

One of the key properties of a quantum divergence is its monotonicity under quantum channels, which is also called the data-processing inequality. Recall that a quantum channel $\Phi: L_1(\mathcal{M}_1)\to L_1(\mathcal{M}_2)$ is a completely positive, trace-preserving (CPTP) map. The data-processing inequality is as follows [Pet85]:

$$Q_f(\rho\|\sigma) \geqslant Q_f(\Phi(\rho)\|\Phi(\sigma))\,, \tag{9}$$

holding for all quantum channels Φ and states ρ,σ. It was proved (see [Pet86a, Pet88] and [Hia21, theorem 6.19] for the general case) that for a regular operator convex function f, the equality $Q_f(\rho\|\sigma)=Q_f(\Phi(\rho)\|\Phi(\sigma))$ holds if and only if $R_{\Phi,\rho}(\Phi(\sigma))=\sigma$, where $R_{\Phi,\rho}: L_1(\mathcal{M}_2)\to L_1(\mathcal{M}_1)$ is the Petz recovery map, defined as

$$R_{\Phi,\rho}(x) := \rho^{1/2}\Phi^\dagger(\Phi(\rho)^{-1/2}x\Phi(\rho)^{-1/2})\rho^{1/2}\;.$$

The data-processing inequality for the optimized f-divergence in the finite-dimensional setting, i.e.

$$\widetilde{Q}_f(\rho\|\sigma) \geqslant \widetilde{Q}_f(\Phi(\rho)\|\Phi(\sigma)) ,$$

was proved in [Wil18a].

Throughout this paper, we mostly discuss recoverability results for the special case when Φ is the restriction to a subalgebra, as was done in [Pet86a]. For example, a partial-trace map $\mathrm{id}\otimes\mathrm{tr} : B(H_A\otimes H_B)\to B(H_A)$ is a restriction from a tensor-product system AB to the subsystem A. By the Stinespring dilation theorem [Sti55], this is the core step in the data-processing inequality. Indeed, recoverability results for general quantum channels (CPTP maps) follow from the subalgebra case.

Let $\mathcal{N}\subset\mathcal{M}$ be a subalgebra, and let $E:\mathcal{M}\to\mathcal{N}$ be the unital, trace-preserving conditional expectation, defined through

$$\tau(xy) = \tau(xE(y)) , \quad \forall \;\; x\in\mathcal{N}, y\in\mathcal{M}.$$

For $1\leqslant p\leqslant\infty$, the space $L_p(\mathcal{N})$ is a subspace of $L_p(\mathcal{M})$, and the conditional expectation $E: L_p(\mathcal{M})\to L_p(\mathcal{N})$ extends to a projection. The quantum channel corresponding to restriction to a subalgebra is then given by $E: L_1(\mathcal{M})\to L_1(\mathcal{N})$. For an invertible density operator $\rho\in L_1(\mathcal{M})$, we write $\rho_{\mathcal{N}} := E(\rho)$ for the reduced density operator of ρ on $\mathcal{N}$. The ρ-preserving conditional expectation $E_\rho:\mathcal{M}\to\mathcal{N}$ is a unital, completely positive (UCP) map:

$$E_\rho(x) := \rho_{\mathcal{N}}^{-\frac{1}{2}} E\left(\rho^{\frac{1}{2}} x \rho^{\frac{1}{2}}\right) \rho_{\mathcal{N}}^{-\frac{1}{2}}.$$

The Petz recovery map $R_\rho: L_1(\mathcal{N})\to L_1(\mathcal{M})$ is the adjoint of E_ρ and is a completely positive trace-preserving (CPTP) map:

$$R_\rho(x) = \rho^{\frac{1}{2}}\left(\rho_{\mathcal{N}}^{-\frac{1}{2}} x \rho_{\mathcal{N}}^{-\frac{1}{2}}\right)\rho^{\frac{1}{2}}. \tag{10}$$

Let us also define the rotated Petz map R^t_ρ [Wil15] and the universal Petz map R^u_ρ [JRS+18] as follows:

$$R^t_\rho(x) := \rho^{\frac{1}{2}-it}\left(\rho_{\mathcal{N}}^{-\frac{1}{2}+it} x \rho_{\mathcal{N}}^{-\frac{1}{2}-it}\right)\rho^{\frac{1}{2}+it} \quad , \forall\; t\in\mathbb{R}, \tag{11}$$

$$R^u_\rho(x) := \int_{\mathbb{R}} R^{t/2}_\rho(x)\;\mathrm{d}\beta(t), \tag{12}$$

where $\mathrm{d}\beta(t) := \frac{\pi}{2}(\cosh(\pi t)+1)^{-1}\mathrm{d}t$ is a probability measure on $\mathbb{R}$. Both R^t_ρ and R^u_ρ are CPTP maps satisfying $R^t_\rho(\rho_{\mathcal{N}}) = R^u_\rho(\rho_{\mathcal{N}}) = \rho$ because $\rho^{it}\in\mathcal{M}$ (resp. $\rho_{\mathcal{N}}^{-it}\in\mathcal{N}$) is a unitary operator and commutes with ρ (resp. $\rho_{\mathcal{N}}$).

4. Recoverability for *f*-divergences via rotated Petz maps

4.1. Recoverability via a rotated Petz map R^t_ρ

In this section, we discuss recoverability for the f-divergence in the finite-dimensional setting. We start with an improvement of the argument in [CV20a].

Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra, and let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. Let $\rho, \sigma \in D_+(\mathcal{M})$ be two faithful states for $\mathcal{M}$, and let $\rho_\mathcal{N}, \sigma_\mathcal{N}$ be their restrictions on $\mathcal{N}$, respectively. We use the following shorthand notation for the relative modular operators:

$$\Delta_\mathcal{M} \equiv \Delta(\sigma, \rho) \in B(L_2(\mathcal{M})) , \qquad \Delta_\mathcal{N} \equiv \Delta(\sigma_\mathcal{N}, \rho_\mathcal{N}) \ \in B(L_2(\mathcal{N})).$$

Let $f : (0, \infty) \to \mathbb{R}$ be an operator anti-monotone function with the following integral representation:

$$f(x) = a + bx + \int_0^\infty \left(\frac{1}{\lambda + x} - \frac{\lambda}{\lambda^2 + 1} \right) \mathrm{d}\nu(\lambda), \tag{13}$$

where $\mathrm{d}\nu(\lambda)$ is the corresponding measure on $\mathbb{R}$. For $\lambda \geqslant 0$, let

$$Q_\lambda(\rho\|\sigma) := \langle \rho^{1/2} | (\lambda + \Delta(\sigma, \rho))^{-1} | \rho^{1/2} \rangle \tag{14}$$

denote the standard f-divergence for $f(x) = (\lambda + x)^{-1}$. It follows from the integral representation that

$$Q_f(\rho\|\sigma) - Q_f(\rho_\mathcal{N}\|\sigma_\mathcal{N}) = \int_0^\infty \left[Q_\lambda(\rho\|\sigma) - Q_\lambda(\rho_\mathcal{N}\|\sigma_\mathcal{N}) \right] \ \mathrm{d}\nu(\lambda) \ . \tag{15}$$

By the data-processing inequality and inspection of (14), the following function

$$F(\lambda) := Q_\lambda(\rho\|\sigma) - Q_\lambda(\rho_\mathcal{N}\|\sigma_\mathcal{N}) \ , \quad \lambda \in [0, \infty)$$

is a continuous non-negative function of all faithful states ρ and σ.

We now recall [CV20a, lemma 2.1]:

Lemma 4.1 (Lemma 2.1 of [CV20a]). *Let $U : K \to H$ be a Hilbert-space isometry, and let A be an invertible positive operator on H. Then for all $h \in K$, the following identity holds*

$$\langle h | U^* A^{-1} U | h \rangle - \langle h | (U^* A U)^{-1} | h \rangle = \langle v | A | v \rangle \geqslant 0,$$

where $|v\rangle := A^{-1} U |h\rangle - U(U^* A U)^{-1} |h\rangle$.

Define the isometry $V_\rho : L_2(\mathcal{N}) \to L_2(\mathcal{M})$ as

$$V_\rho \, |x\rangle = \left| x \rho_\mathcal{N}^{-\frac{1}{2}} \rho^{\frac{1}{2}} \right\rangle \ , \quad \forall \ x \in \mathcal{N} \ .$$

The adjoint is $V_\rho^*(x) = E(x\rho^{1/2})\rho_\mathcal{N}^{-1/2}$. Since $V_\rho^* \Delta_\mathcal{M} V_\rho = \Delta_\mathcal{N}$ (we require the assumption of faithfulness of ρ here) and by lemma 4.1 above, we find that

$$\begin{aligned}
F(\lambda) &= Q_\lambda(\rho\|\sigma) - Q_\lambda(\rho_\mathcal{N}\|\sigma_\mathcal{N}) && (16)\\
&= \langle \rho^{1/2} | (\Delta_\mathcal{M} + \lambda)^{-1} | \rho^{1/2} \rangle - \langle \rho_\mathcal{N}^{1/2} | (\Delta_\mathcal{N} + \lambda)^{-1} | \rho_\mathcal{N}^{1/2} \rangle && (17)\\
&= \langle \rho_\mathcal{N}^{1/2} | V_\rho^* (\Delta_\mathcal{M} + \lambda)^{-1} V_\rho | \rho_\mathcal{N}^{1/2} \rangle - \langle \rho_\mathcal{N}^{1/2} | (V_\rho^* (\Delta_\mathcal{M} + \lambda) V_\rho)^{-1} | \rho_\mathcal{N}^{1/2} \rangle && (18)\\
&= \langle w_\lambda | (\Delta_\mathcal{M} + \lambda) | w_\lambda \rangle && (19)\\
&= \left\| \Delta_\mathcal{M}^{1/2} | w_\lambda \rangle \right\|_2^2 + \lambda \| | w_\lambda \rangle \|_2^2, && (20)
\end{aligned}$$

where

$$|w_\lambda\rangle := (\Delta_{\mathcal{M}} + \lambda)^{-1} \left|\rho^{1/2}\right\rangle - V_\rho(\Delta_{\mathcal{N}} + \lambda)^{-1} \left|\rho_{\mathcal{N}}^{\frac{1}{2}}\right\rangle \tag{21}$$

is a vector in $L_2(\mathcal{M})$.

Lemma 4.2. *Let $t \in \mathbb{R}$ and set*

$$|w_t\rangle := \Delta_{\mathcal{M}}^{\frac{1}{2}+it} \left|\rho^{1/2}\right\rangle - V_\rho \Delta_{\mathcal{N}}^{\frac{1}{2}+it} \left|\rho_{\mathcal{N}}^{1/2}\right\rangle = \left|\sigma^{\frac{1}{2}+it}\rho^{-it}\right\rangle - \left|\sigma_{\mathcal{N}}^{\frac{1}{2}+it}\rho_{\mathcal{N}}^{-\frac{1}{2}-it}\rho^{\frac{1}{2}}\right\rangle. \tag{22}$$

Then the following equality holds

$$|w_t\rangle = -\frac{\cosh(\pi t)}{\pi}\left(\int_0^\infty \lambda^{\frac{1}{2}+it}|w_\lambda\rangle\ \mathrm{d}\lambda\right) \tag{23}$$

and the following inequality holds

$$\left\|\sigma - R_\rho^t(\sigma)\right\|_1 \leqslant 2\||w_t\rangle\|_2. \tag{24}$$

Proof. Recall the operator integral from [Kom66], which states that the following integral formula holds for the imaginary power of a positive operator A:

$$\begin{aligned} A^{\frac{1}{2}+it} &= \frac{\sin\left(\pi\left(\frac{1}{2}+it\right)\right)}{\pi}\int_0^\infty \lambda^{\frac{1}{2}+it}\left(\frac{1}{\lambda} - \frac{1}{\lambda + A}\right)\mathrm{d}\lambda \\ &= \frac{\cosh(\pi t)}{\pi}\int_0^\infty \lambda^{\frac{1}{2}+it}\left(\frac{1}{\lambda} - \frac{1}{\lambda + A}\right)\mathrm{d}\lambda. \end{aligned}$$

Then

$$\begin{aligned} |w_t\rangle &= \Delta_{\mathcal{M}}^{\frac{1}{2}+it}\left|\rho^{1/2}\right\rangle - V_\rho\Delta_{\mathcal{N}}^{\frac{1}{2}+it}\left|\rho_{\mathcal{N}}^{1/2}\right\rangle \\ &= \frac{\cosh(\pi t)}{\pi}\left(\int_0^\infty \lambda^{1/2+it}\left(\frac{1}{\lambda} - \frac{1}{\Delta_{\mathcal{M}}+\lambda}\right)\left|\rho^{1/2}\right\rangle \mathrm{d}\lambda\right. \\ &\quad \left. - V_\rho\int_0^\infty \lambda^{1/2+it}\left(\frac{1}{\lambda} - \frac{1}{\Delta_{\mathcal{N}}+\lambda}\right)\left|\rho_{\mathcal{N}}^{1/2}\right\rangle \mathrm{d}\lambda\right) \\ &= -\frac{\cosh(\pi t)}{\pi}\left(\int_0^\infty \lambda^{1/2+it}\left(\frac{1}{\Delta_{\mathcal{M}}+\lambda}\left|\rho^{1/2}\right\rangle - V_\rho\frac{1}{\Delta_{\mathcal{N}}+\lambda}\left|\rho_{\mathcal{N}}^{1/2}\right\rangle\right)\mathrm{d}\lambda\right) \\ &= -\frac{\cosh(\pi t)}{\pi}\left(\int_0^\infty \lambda^{1/2+it}|w_\lambda\rangle\ \mathrm{d}\lambda\right). \end{aligned}$$

This establishes (23).

We now prove (24). Note that $(\sigma^{\frac{1}{2}+it})^*\sigma^{\frac{1}{2}+it} = \sigma$, and

$$\begin{aligned} &\left(\sigma_{\mathcal{N}}^{\frac{1}{2}+it}\rho_{\mathcal{N}}^{-\frac{1}{2}-it}\rho^{\frac{1}{2}+it}\right)^*\left(\sigma_{\mathcal{N}}^{\frac{1}{2}+it}\rho_{\mathcal{N}}^{-1/2-it}\rho^{\frac{1}{2}+it}\right) \\ &\quad = \rho^{\frac{1}{2}-it}\rho_{\mathcal{N}}^{-\frac{1}{2}+it}\sigma_{\mathcal{N}}\rho_{\mathcal{N}}^{-\frac{1}{2}-it}\rho^{\frac{1}{2}+it} = R_\rho^t(\sigma), \end{aligned}$$

where $R^t_\rho(\sigma)$ is defined in (11). Recall the following inequality from [CV20a, lemma 2.2]:

$$\|x^*x - y^*y\|_1 \leqslant 2\|x-y\|_2\,, \tag{25}$$

which holds for x and y satisfying $\|x\|_2 = \|y\|_2 = 1$. Then (24) follows because

$$\begin{aligned}\left\|\sigma - R^t_\rho(\sigma)\right\|_1 &\leqslant 2\left\|\sigma^{1/2+it} - \sigma_{\mathcal{N}}^{1/2+it}\rho_{\mathcal{N}}^{-1/2-it}\rho^{\frac{1}{2}+it}\right\|_2 \\ &= 2\left\|\sigma^{1/2+it}\rho^{-it} - \sigma_{\mathcal{N}}^{1/2+it}\rho_{\mathcal{N}}^{-1/2-it}\rho^{\frac{1}{2}}\right\|_2 = 2\||w_t\rangle\|_2.\end{aligned}$$

□

For a regular operator anti-monotone function f, we have the following estimate of $\||w_t\rangle\|_2$:

Lemma 4.3. *Let $f:(0,\infty)\to\mathbb{R}$ be a regular operator anti-monotone function, and let $\mathrm{d}\nu$ be the measure in its integral representation. Suppose $Q_{x^2}(\rho\|\sigma) = \left\langle\rho^{1/2}\right|\Delta^2_{\mathcal{M}}\left|\rho^{1/2}\right\rangle < \infty$. Suppose for some S and T, satisfying $0\leqslant S < T\leqslant\infty$, that there exists $c(S,T) > 0$ such that on the interval (S,T),*

$$\mathrm{d}\lambda \leqslant c(S,T)\,\mathrm{d}\nu(\lambda).$$

Then, for $|w_t\rangle$ as defined in (22), the following inequality holds

$$\begin{aligned}\||w_t\rangle\|_2 \leqslant \frac{\cosh(\pi t)}{\pi}\Big(4S^{1/2} + [c(S,T)(T-S)]^{1/2}\big(Q_f(\rho\|\sigma) \\ - Q_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})\big)^{1/2} + 4T^{-1/2}Q_{x^2}(\rho\|\sigma)^{1/2}\Big)\,.\end{aligned}$$

Proof. To estimate the norm of $|w_t\rangle$, we break $|w_t\rangle$ into three separate terms after applying (23):

$$\begin{aligned}-|w_t\rangle &= \frac{\cosh(\pi t)}{\pi}\left(\int_0^\infty \lambda^{1/2+it}\,|w_\lambda\rangle\,\mathrm{d}\lambda\right) \\ &= \frac{\cosh(\pi t)}{\pi}\left(\int_0^S \lambda^{1/2+it}\,|w_\lambda\rangle\,\mathrm{d}\lambda + \int_T^S \lambda^{1/2+it}\,|w_\lambda\rangle\,\mathrm{d}\lambda + \int_T^\infty \lambda^{1/2+it}\,|w_\lambda\rangle\,\mathrm{d}\lambda\right) \\ &=: \frac{\cosh(\pi t)}{\pi}\,(\mathrm{I}+\mathrm{II}+\mathrm{III})\,,\end{aligned}$$

where $|w_\lambda\rangle$ is defined in (21). For the first term I, we define the function

$$h^t_S(x) := \int_0^S \lambda^{\frac{1}{2}+it}\left(\frac{1}{\lambda} - \frac{1}{\lambda+x}\right)\mathrm{d}\lambda.$$

Thus

$$\int_0^S \lambda^{\frac{1}{2}+it}\,|w_\lambda\rangle\,\mathrm{d}\lambda = V_\rho h^t_S(\Delta_{\mathcal{N}})\left|\rho_{\mathcal{N}}^{1/2}\right\rangle - h^t_S(\Delta_{\mathcal{M}})\left|\rho^{1/2}\right\rangle.$$

For $x \geqslant 0$,

$$|h_S^t(x)| \leqslant \int_0^S \left|\lambda^{\frac{1}{2}+it}\left(\frac{1}{\lambda} - \frac{1}{\lambda+x}\right)\right| \mathrm{d}\lambda \leqslant \int_0^S \lambda^{\frac{1}{2}}\left(\frac{1}{\lambda} - \frac{1}{\lambda+x}\right) \mathrm{d}\lambda$$
$$= 2x^{1/2}\arctan\left(\frac{\sqrt{S}}{\sqrt{x}}\right) = h_S^0(x).$$

Note that h_S^0 is the function h_S^t with $t = 0$. So we conclude that

$$\left\| h_S^t(\Delta_{\mathcal{M}}) \left|\rho^{1/2}\right\rangle \right\|_2^2 = \left\langle \rho^{1/2} \right| h_S^t(\Delta_{\mathcal{M}})^* h_S^t(\Delta_{\mathcal{M}}) \left|\rho^{1/2}\right\rangle$$
$$\leqslant \left\langle \rho^{1/2} \right| 4\Delta_{\mathcal{M}} \arctan^2\left(\frac{\sqrt{S}}{\sqrt{\Delta_{\mathcal{M}}}}\right) \left|\rho^{1/2}\right\rangle$$
$$= \int_0^\infty 4s \arctan^2\left(\frac{\sqrt{S}}{\sqrt{s}}\right) \mathrm{d}\mu(s)$$
$$\leqslant \int_0^\infty 4s\left(\frac{S}{s}\right) \mathrm{d}\mu(s) \leqslant 4S.$$

Here $\mathrm{d}\mu(s)$ is the probability measure $\mathrm{d}\mu(s) = \mathrm{d}\left\langle \rho^{1/2} \right| E_{[0,s]}(\Delta_{\mathcal{M}}) \left|\rho^{1/2}\right\rangle$ from the spectral decomposition, and the second inequality uses the fact that $\arctan(x) \leqslant x$ for $x \geqslant 0$. Similarly,

$$\left\| V_\rho h_S^t(\Delta_{\mathcal{N}}) \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \right\|_2^2 \leqslant \left\langle \rho_{\mathcal{N}}^{1/2} \right| 4\Delta_{\mathcal{N}} \arctan^2\left(\frac{\sqrt{S}}{\sqrt{\Delta_{\mathcal{N}}}}\right) \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \leqslant 4S\ .$$

Thus we have

$$\|\mathrm{I}\|_2 \leqslant \left\| h_S^t(\Delta_{\mathcal{M}}) \left|\rho^{1/2}\right\rangle \right\|_2 + \left\| V_\rho h_S^t(\Delta_{\mathcal{N}}) \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \right\|_2 \leqslant 2S^{1/2} + 2S^{1/2} = 4S^{1/2}. \tag{26}$$

For the second term, consider that

$$\|\mathrm{II}\|_2 = \left\| \int_S^T \lambda^{1/2+it} \left|w_\lambda\right\rangle \mathrm{d}\lambda \right\|_2 \tag{27}$$
$$\leqslant \int_S^T \lambda^{1/2} \| \left|w_\lambda\right\rangle \|_2 \, \mathrm{d}\lambda \tag{28}$$
$$\leqslant \left(\int_S^T 1\,\mathrm{d}\lambda\right)^{1/2} \left(\int_S^T \lambda \| \left|w_\lambda\right\rangle \|_2^2 \, \mathrm{d}\lambda\right)^{1/2} \tag{29}$$
$$\leqslant (T-S)^{1/2} \left(\int_S^T F(\lambda)\,\mathrm{d}\lambda\right)^{1/2} \tag{30}$$
$$\leqslant (T-S)^{1/2} \left(c(S,T) \int_S^T F(\lambda)\ \mathrm{d}\nu(\lambda)\right)^{1/2} \tag{31}$$
$$\leqslant \left[c(S,T)(T-S)(Q_f(\rho\|\sigma) - Q_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}))\right]^{1/2}. \tag{32}$$

The first inequality follows from the triangle inequality, the second from Cauchy–Schwarz, the third from (20), the fourth from the assumption of a regular operator anti-monotone function f, and the fifth from (15).

For the third term, consider that

$$\begin{aligned}\mathrm{III} &= \int_T^\infty \lambda^{1/2+it} \left|w_\lambda\right\rangle \, \mathrm{d}\lambda \\ &= -\int_T^\infty \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda + \Delta_\mathcal{M}}\right) \left|\rho^{1/2}\right\rangle \mathrm{d}\lambda + V_\rho \int_T^\infty \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda + \Delta_\mathcal{N}}\right) \left|\rho_\mathcal{N}^{1/2}\right\rangle \mathrm{d}\lambda.\end{aligned}$$

Let us consider the integral

$$\int_T^\infty \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda + x}\right) \mathrm{d}\lambda = x^{1/2+it} \int_{\frac{T}{x}}^\infty \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda + 1}\right) \mathrm{d}\lambda.$$

Note that the function $\lambda \mapsto \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda+1}\right)$ is bounded and integrable on $(0, \infty)$. We define the following continuous function:

$$g_T^t(x) := \int_T^\infty \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda + x}\right) \mathrm{d}\lambda.$$

Then

$$\int_T^\infty \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda + \Delta_\mathcal{M}}\right) \mathrm{d}\lambda = g_T^t(\Delta_\mathcal{M}) \ .$$

For $x \geqslant 0$,

$$|g_T^t(x)|^2 \leqslant x \left(\int_{\frac{T}{x}}^\infty \lambda^{1/2} \left(\frac{1}{\lambda} - \frac{1}{\lambda + 1}\right) \mathrm{d}\lambda\right)^2 = 4x \arctan^2\left(\frac{\sqrt{x}}{\sqrt{T}}\right) \ . \tag{33}$$

Therefore

$$\begin{aligned}\left\| \int_T^\infty \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda + \Delta_\mathcal{M}}\right) \mathrm{d}\lambda \left|\rho^{1/2}\right\rangle \right\|_2^2 &= \left\langle \rho^{1/2} \right| |g_T^t(\Delta_\mathcal{M})|^2 \left|\rho^{1/2}\right\rangle \\ &\leqslant \left\langle \rho^{1/2} \right| 4\Delta_\mathcal{M} \arctan^2\left(\frac{\sqrt{\Delta_\mathcal{M}}}{\sqrt{T}}\right) \left|\rho^{1/2}\right\rangle \leqslant \frac{4}{T} \left\langle \rho^{1/2} \right| \Delta_\mathcal{M}^2 \left|\rho^{1/2}\right\rangle \\ &= \frac{4}{T} Q_{x^2}(\rho\|\sigma),\end{aligned}$$

where we used (33) and the bound $\arctan(x) \leqslant x$, holding for $x \geqslant 0$. Similarly,

$$\begin{aligned}\left\| V_\rho \int_T^\infty \lambda^{1/2+it} \left(\frac{1}{\lambda} - \frac{1}{\lambda + \Delta_\mathcal{N}}\right) \left|\rho_\mathcal{N}^{1/2}\right\rangle \mathrm{d}\lambda \right\|_2^2 &\leqslant \frac{4}{T} \left\langle \rho_\mathcal{N}^{1/2} \right| \Delta_\mathcal{N}^2 \left|\rho_\mathcal{N}^{1/2}\right\rangle \\ &= \frac{4}{T} Q_{x^2}(\rho_\mathcal{N}\|\sigma_\mathcal{N}) \leqslant \frac{4}{T} Q_{x^2}(\rho\|\sigma).\end{aligned}$$

The final inequality follows from the data-processing inequality for the Petz–Rényi relative quasi-entropy Q_{x^2}. So we conclude from the analysis above and the triangle inequality that

$$\|\mathrm{III}\|_2 \leqslant 4T^{-1/2} Q_{x^2}(\rho\|\sigma)^{1/2}. \tag{34}$$

Putting the estimates from (26), (32) and (34) together, we find that

$$\begin{aligned}\| |w_t\rangle \|_2 &\leqslant \frac{\cosh(\pi t)}{\pi}\left(\|\mathrm{I}\|_2 + \|\mathrm{II}\|_2 + \|\mathrm{III}\|_2\right) \\ &\leqslant \frac{\cosh(\pi t)}{\pi}\Big(4S^{1/2} + c(S,T)^{1/2}(T-S)^{1/2}(Q_f(\rho\|\sigma) - Q_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}))^{1/2} \\ &\quad + 4T^{-1/2}Q_{x^2}(\rho\|\sigma)^{1/2}\Big)\,.\end{aligned}$$

This completes the proof. □

A direct consequence of lemmas 4.2 and 4.3 is the following general bound on the recoverability error in terms of a standard f-divergence:

Corollary 4.4. *Considering the same hypotheses of lemma 4.3, the following inequality holds*

$$\begin{aligned}\left\|\sigma - R^t_\rho(\sigma_{\mathcal{N}})\right\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi}\Big(4S^{1/2} + [c(S,T)(T-S)]^{1/2}\big(Q_f(\rho\|\sigma) \\ - Q_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})\big)^{1/2} + 4T^{-1/2}Q_{x^2}(\rho\|\sigma)^{1/2}\Big)\,.\end{aligned}$$

For a particular choice of f, the estimate from corollary 4.4 simplifies, depending on the measure ν. In the next section, we consider some important examples.

4.1.1. Recoverability for quantum relative entropy. We begin with the quantum relative entropy $D(\rho\|\sigma)$, as defined in (7).

Theorem 4.5. *Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra, and let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. Then for all faithful states ρ and σ, the following inequalities hold*

$$D(\rho\|\sigma) - D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) \geqslant \left(\frac{\pi}{8}\right)^4 Q_{x^2}(\rho\|\sigma)^{-1}\|\sigma - R_\rho(\sigma_{\mathcal{N}})\|_1^4, \tag{35}$$

$$D(\rho\|\sigma) - D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) \geqslant \left(\frac{\pi}{8\cosh(\pi t)}\right)^4 Q_{x^2}(\rho\|\sigma)^{-1}\left\|\sigma - R^t_\rho(\sigma_{\mathcal{N}})\right\|_1^4, \tag{36}$$

$$D(\rho\|\sigma) - D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) \geqslant \frac{1}{256}\, Q_{x^2}(\rho\|\sigma)^{-1}\left\|\sigma - R^u_\rho(\sigma_{\mathcal{N}})\right\|_1^4. \tag{37}$$

Here $Q_{x^2}(\rho\|\sigma) = \langle\rho^{1/2}|\Delta(\sigma,\rho)^2|\rho^{1/2}\rangle = \tau(\rho^{-1}\sigma^2)$.

Proof. Consider from (7) that $f(x) = -\log x$, for which we have the following integral representation:

$$-\log\, x = \int_0^\infty \left(\frac{1}{\lambda + x} - \frac{\lambda}{1+\lambda^2}\right) \mathrm{d}\lambda\,,$$

where $\mathrm{d}\lambda$ is the Lebesgue measure. Thus $c(S,T) = 1$ for $0 \leqslant S \leqslant T \leqslant \infty$. Choose $S = 0$ and

$$T = 4\,Q_{x^2}(\rho\|\sigma)^{1/2}(D(\rho\|\sigma) - D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}))^{-1/2}\,.$$

Applying corollary 4.4, we obtain the following:

$$\left\|\sigma - R_\rho^t(\sigma_{\mathcal{N}})\right\|_1 \leqslant \frac{2\,\cosh(\pi t)}{\pi}\left[T^{1/2}(D(\rho\|\sigma) - D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}))^{1/2} + 4T^{-1/2}Q_{x^2}(\rho\|\sigma)^{1/2}\right]$$
$$= \frac{8\,\cosh(\pi t)}{\pi}(D(\rho\|\sigma) - D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}))^{1/4}Q_{x^2}(\rho\|\sigma)^{1/4}. \tag{38}$$

Equation (36) follows from rewriting (38), and (37) follows from integrating (38) with respect to the measure $\mathrm{d}\beta(t) = \frac{\pi}{2}(\cosh(\pi t) + 1)^{-1}\mathrm{d}t$, from convexity of the trace norm, the fact that $\int_{\mathbb{R}} \frac{\cosh(\pi t/2)}{(\cosh(\pi t)+1)}\mathrm{d}t = 1$, and by the integral expression $R_\rho^u = \int_{\mathbb{R}} R_\rho^{\frac{t}{2}}\,\mathrm{d}\beta(t)$. Equation (35) is a special case of (36) at $t = 0$. □

Remark 4.6. As mentioned in section 2, an advantage of the bounds in theorem 4.5 over previous bounds from [CV18, CV20a] is that the remainder term features the quantity $Q_{x^2}(\rho\|\sigma)$ rather than the operator norm of the relative modular operator. As such, these bounds are non-trivial for the important class of bosonic Gaussian states [Ser17], whereas the previous bounds from [CV18, CV20a] are trivial for this class of states. Moreover, explicit formulas are available for evaluating the Petz– and sandwiched-Rényi relative entropies of bosonic Gaussian states (see [SLW18] and references therein). This remark applies not only to theorem 4.5, but also to theorem 4.7, corollary 4.9, theorems 4.13 and 4.15, and corollary 4.16.

4.1.2. Recoverability for Petz–Rényi relative (quasi-)entropy. For $\alpha \in (0, 1) \cup (1, 2)$, the Petz–Rényi relative entropy is given by

$$D_\alpha(\rho\|\sigma) = \frac{1}{\alpha - 1}\,\log\,\tau(\rho^\alpha\sigma^{1-\alpha}) = \frac{1}{\alpha - 1}\,\log\,Q_{x^{1-\alpha}}(\rho\|\sigma)\ . \tag{39}$$

In some of the statements that follow, we also adopt a different parameterization by setting

$$s := 1 - \alpha \in (-1, 0) \cup (0, 1), \tag{40}$$

so that

$$D_\alpha(\rho\|\sigma) = D_{1-s}(\rho\|\sigma) = -\frac{1}{s}\,\log\,\tau(\rho^{1-s}\sigma^s) = -\frac{1}{s}\,\log\,Q_{x^s}(\rho\|\sigma)\ , \tag{41}$$

and we also adopt the abbreviation

$$Q_s(\rho\|\sigma) \equiv Q_{x^s}(\rho\|\sigma) = \tau(\rho^{1-s}\sigma^s). \tag{42}$$

We begin by focusing on the Petz–Rényi relative quasi-entropy in theorem 4.7, and then we extend the result to the Petz–Rényi relative entropy in corollary 4.9.

Theorem 4.7. *Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra, and let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. Then for all faithful states ρ and σ, the following inequalities hold for the Petz–Rényi relative quasi-entropy $Q_s(\rho\|\sigma) = \tau(\rho^{1-s}\sigma^s)$ for $s \in (-1, 0) \cup (0, 1)$:*

$$|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})| \geqslant K(s, Q_{x^2}(\rho\|\sigma))\left(\frac{\pi}{4 + 2|s|}\|\sigma - R_\rho(\sigma_{\mathcal{N}})\|_1\right)^{4+2|s|}, \tag{43}$$

$$|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})| \geqslant K(s, Q_{x^2}(\rho\|\sigma)) \times \left(\frac{\pi}{(4+2|s|)\cosh \pi t}\left\|\sigma - R^t_\rho(\sigma_{\mathcal{N}})\right\|_1\right)^{4+2|s|}, \tag{44}$$

$$|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})| \geqslant K(s, Q_{x^2}(\rho\|\sigma))\left(\frac{1}{(2+|s|)}\left\|\sigma - R^u_\rho(\sigma_{\mathcal{N}})\right\|_1\right)^{4+2|s|}, \tag{45}$$

where $Q_{x^2}(\rho\|\sigma) = \tau(\rho^{-1}\sigma^2)$ *and*

$$K(s, Q_{x^2}(\rho\|\sigma)) := \begin{cases} \dfrac{\sin(\pi|s|)}{\pi}\left(\dfrac{(|s|+1)^2}{16Q_{x^2}(\rho\|\sigma)}\right)^{|s|+1} & \textit{for } -1 < s < 0 \\ \dfrac{\sin(\pi s)}{\pi Q_{x^2}(\rho\|\sigma)}\dfrac{s^{2s}}{16^{s+1}} & \textit{for } 0 < s < 1 \end{cases}. \tag{46}$$

Proof. For $0 < s < 1$, the function $f(x) = x^s$ is operator monotone and operator concave. An integral representation for it is

$$x^s = \frac{\sin(\pi s)}{\pi}\int_0^\infty \lambda^s\left(\frac{1}{\lambda} - \frac{1}{\lambda + x}\right)\mathrm{d}\lambda\ .$$

So for $0 < S < T < 0$, the constant $c(S,T) \leqslant \frac{\pi}{\sin(\pi s)}S^{-s}$. Then by applying corollary 4.4, we find that

$$\left\|\sigma - R^t_\rho(\sigma_{\mathcal{N}})\right\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi}\left(4S^{1/2} + \left(\frac{\pi}{\sin(\pi s)}\right)^{1/2} S^{-s/2}(T-S)^{1/2}|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})|^{1/2} + 4T^{-1/2}Q_{x^2}(\rho\|\sigma)^{1/2}\right).$$

Define the following constants:

$$a = 4\ , \qquad b = \left(\frac{\pi}{\sin(\pi s)}\right)^{\frac{1}{2}}|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})|^{1/2}\ , \qquad c = 4Q_{x^2}(\rho\|\sigma)^{1/2}.$$

We then have

$$\begin{aligned}\left\|\sigma - R^t_\rho(\sigma_{\mathcal{N}})\right\|_1 &\leqslant \frac{2\cosh(\pi t)}{\pi}\left(aS^{1/2} + bS^{-s/2}(T-S)^{\frac{1}{2}} + cT^{-1/2}\right) \\ &\leqslant \frac{2\cosh(\pi t)}{\pi}\left(aS^{1/2} + bS^{-s/2}(T-S)^{\frac{1}{2}} + c(T-S)^{-\frac{1}{2}}\right).\end{aligned}$$

Minimizing the following function over $0 < S < T < \infty$:

$$F(S,T) = aS^{1/2} + bS^{-s/2}(T-S)^{\frac{1}{2}} + c(T-S)^{-\frac{1}{2}},$$

we obtain the following inequality at the choices $S = \left(\frac{cbs^2}{a^2}\right)^{\frac{2}{s+2}}$ and $T - S = \left(\frac{cbs^2}{a^2}\right)^{\frac{s}{s+2}}\left(\frac{c}{b}\right)$:

$$\begin{aligned}\left\|\sigma - R^t_\rho(\sigma_\mathcal{N})\right\|_1 &\leqslant \frac{2\cosh(\pi t)}{\pi}(s+2)(a^s s^{-s}\, bc)^{\frac{1}{s+2}}\\ &= \frac{2\cosh(\pi t)}{\pi}(s+2)\left(16^{s+1}s^{-2s}\frac{\pi}{\sin(\pi s)}|Q_s(\rho\|\sigma)\right.\\ &\qquad \left. - Q_s(\rho_\mathcal{N}\|\sigma_\mathcal{N})|Q_{x^2}(\rho\|\sigma)\right)^{\frac{1}{2s+4}}. \end{aligned} \tag{47}$$

For $-1 < s < 0$, the function $f(x) = x^s$ is operator anti-monotone and operator convex. An integral representation of x^s for $s \in (-1, 0)$ is

$$x^s = \frac{\sin(\pi|s|)}{\pi}\int_0^\infty \frac{\lambda^s}{\lambda + x}\,\mathrm{d}\lambda\ .$$

Then we can choose $S = 0$ and the constant $c(0, T) \leqslant \frac{\pi}{\sin(\pi|s|)}T^{|s|}$. By corollary 4.4, we find that

$$\begin{aligned}\left\|\sigma - R^t_\rho(\sigma_\mathcal{N})\right\|_1 &\leqslant \frac{2\cosh(\pi t)}{\pi}\left(\left(\frac{\pi}{\sin(\pi|s|)}T^{|s|}\right)^{1/2}T^{1/2}\left(Q_s(\rho\|\sigma)\right.\right.\\ &\qquad \left.\left. - Q_s(\rho_\mathcal{N}\|\sigma_\mathcal{N})\right)^{1/2} + 4T^{-1/2}Q_{x^2}(\rho\|\sigma)^{1/2}\right). \end{aligned}$$

Define the following constants:

$$b = \left(\frac{\pi}{\sin(\pi|s|)}\right)^{\frac{1}{2}}(Q_s(\rho\|\sigma) - Q_s(\rho_\mathcal{N}\|\sigma_\mathcal{N}))^{1/2}\ , \qquad c = 4Q_{x^2}(\rho\|\sigma)^{1/2}.$$

We want to minimize the following function over $0 < T < \infty$:

$$G(T) = bT^{\frac{|s|+1}{2}} + cT^{-1/2}\ .$$

Choosing $T = \left(\frac{c}{b(|s|+1)}\right)^{\frac{2}{|s|+2}}$, we find that

$$\begin{aligned}\left\|\sigma - R^t_\rho(\sigma_\mathcal{N})\right\|_1 &\leqslant \frac{2\cosh(\pi t)}{\pi}(|s|+2)(|s|+1)^{-\frac{|s|+1}{|s|+2}}c^{\frac{|s|+1}{|s|+2}}b^{\frac{1}{|s|+2}}\\ &= \frac{2\cosh(\pi t)}{\pi}(|s|+2)(|s|+1)^{-\frac{|s|+1}{|s|+2}}\\ &\quad \times\left(\left(\frac{\pi}{\sin(\pi|s|)}\right)(Q_s(\rho\|\sigma) - Q_s(\rho_\mathcal{N}\|\sigma_\mathcal{N}))\right)^{\frac{1}{2(|s|+2)}}\left(4\sqrt{Q_{x^2}(\rho\|\sigma)}\right)^{\frac{|s|+1}{|s|+2}}. \end{aligned} \tag{48}$$

Putting together the conclusions in (47) and (48), we arrive at (44). Equation (43) is a special case of (44). The proof of (45) is similar to the proof in theorem 4.5. □

Note that the estimate above fails for $s = -1$ because the measure in the integral representation of x^{-1} is a point mass at $\lambda = 0$.

Example 4.8. For $s = 1/2$, we have the Holevo fidelity

$$F_{\mathrm{H}}(\rho,\sigma) = Q_{x^{1/2}}(\rho\|\sigma)^2 = \tau(\rho^{1/2}\sigma^{1/2})^2 .$$

Then

$$\sqrt{F_{\mathrm{H}}(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})} - \sqrt{F_{\mathrm{H}}}(\rho,\sigma) \geqslant \frac{1}{128\pi Q_{x^2}(\rho\|\sigma)}\left(\frac{\pi}{5}\|\sigma - R_\rho(\sigma_{\mathcal{N}})\|_1\right)^5. \tag{49}$$

The inequality in (49) can be compared with the main result of [Wil18b]. The prefactor $[Q_{x^2}(\rho\|\sigma)]^{-1}$ is an improvement, but the fifth power on the trace distance is not.

The estimate in theorem 4.7 leads to a strengthened data-processing inequality for the Petz–Rényi relative entropy, as defined in (39), by following the same argument used to arrive at [CV18, theorem 6.1], along with an additional argument:

Corollary 4.9. *Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra, and let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. Let ρ and σ be two faithful states. For $\alpha \in (0,1)$ and $t \in \mathbb{R}$,*

$$\begin{aligned} D_\alpha(\rho\|\sigma) - D_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) \geqslant \frac{1}{1-\alpha} \log\Big(1 + K(1-\alpha, Q_{x^2}(\rho\|\sigma)) \\ \times \left(\frac{\pi}{2(3-\alpha)\cosh \pi t}\left\|\sigma - R_\rho^t(\sigma_{\mathcal{N}})\right\|_1\right)^{2(3-\alpha)}\Big), \end{aligned}$$

and for $\alpha \in (1,2)$ and $t \in \mathbb{R}$,

$$\begin{aligned} D_\alpha(\rho\|\sigma) - D_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) \geqslant \frac{1}{\alpha-1} \log\Bigg(1 + \frac{K(1-\alpha, Q_{x^2}(\rho\|\sigma))}{Q_{x^{-1}}(\rho\|\sigma)^{\alpha-1}} \\ \times \left(\frac{\pi}{2(\alpha+1)\cosh \pi t}\left\|\sigma - R_\rho^t(\sigma_{\mathcal{N}})\right\|_1\right)^{2(\alpha+1)}\Bigg), \end{aligned}$$

where $Q_{x^2}(\rho\|\sigma)) = \tau(\rho^{-1}\sigma^2)$, $Q_{x^{-1}}(\rho\|\sigma)) = \tau(\rho^2\sigma^{-1})$, and the constant $K(1-\alpha, Q_{x^2}(\rho\|\sigma))$ is given by (46).

Proof. For $0 < \alpha < 1$, we find that

$$\begin{aligned} D_\alpha(\rho\|\sigma) - D_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) &= \frac{1}{1-\alpha} \log \frac{Q_{x^{1-\alpha}}(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})}{Q_{x^{1-\alpha}}(\rho\|\sigma)} \\ &= \frac{1}{1-\alpha} \log\left(1 + \frac{Q_{x^{1-\alpha}}(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) - Q_{x^{1-\alpha}}(\rho\|\sigma)}{Q_{x^{1-\alpha}}(\rho\|\sigma)}\right) \\ &\geqslant \frac{1}{1-\alpha} \log\left(1 + \frac{K(1-\alpha, Q_{x^2}(\rho\|\sigma))}{Q_{x^{1-\alpha}}(\rho\|\sigma)}\right. \end{aligned}$$

$$
\begin{aligned}
&\times \left(\frac{\pi}{2(3-\alpha)\cosh \pi t} \left\| \sigma - R_\rho^t(\sigma_\mathcal{N}) \right\|_1 \right)^{2(3-\alpha)} \Bigg), \\
&\geqslant \frac{1}{1-\alpha} \log \Bigg(1 + K(1-\alpha, Q_{x^2}(\rho\|\sigma)) \\
&\quad \times \left(\frac{\pi}{2(3-\alpha)\cosh \pi t} \left\| \sigma - R_\rho^t(\sigma_\mathcal{N}) \right\|_1 \right)^{2(3-\alpha)} \Bigg).
\end{aligned}
$$

The first inequality follows from (44), and the second follows because $Q_{x^{1-\alpha}}(\rho\|\sigma) \leqslant 1$ for $\alpha \in (0, 1)$.

For $\alpha \in (1, 2)$, consider that

$$
\begin{aligned}
D_\alpha(\rho\|\sigma) - D_\alpha(\rho_\mathcal{N}\|\sigma_\mathcal{N}) &= \frac{1}{\alpha-1} \log \frac{Q_{x^{1-\alpha}}(\rho\|\sigma)}{Q_{x^{1-\alpha}}(\rho_\mathcal{N}\|\sigma_\mathcal{N})} \\
&= \frac{1}{\alpha-1} \log \left(1 + \frac{Q_{x^{1-\alpha}}(\rho\|\sigma) - Q_{x^{1-\alpha}}(\rho_\mathcal{N}\|\sigma_\mathcal{N})}{Q_{x^{1-\alpha}}(\rho_\mathcal{N}\|\sigma_\mathcal{N})} \right) \\
&\geqslant \frac{1}{\alpha-1} \log \Bigg(1 + \frac{K(1-\alpha, Q_{x^2}(\rho\|\sigma))}{Q_{x^{1-\alpha}}(\rho_\mathcal{N}\|\sigma_\mathcal{N})} \\
&\quad \times \left(\frac{\pi}{2(\alpha+1)\cosh \pi t} \left\| \sigma - R_\rho^t(\sigma_\mathcal{N}) \right\|_1 \right)^{2(\alpha+1)} \Bigg) \\
&\geqslant \frac{1}{\alpha-1} \log \Bigg(1 + \frac{K(1-\alpha, Q_{x^2}(\rho\|\sigma))}{Q_{x^{-1}}(\rho\|\sigma)^{\alpha-1}} \\
&\quad \times \left(\frac{\pi}{2(\alpha+1)\cosh \pi t} \left\| \sigma - R_\rho^t(\sigma_\mathcal{N}) \right\|_1 \right)^{2(\alpha+1)} \Bigg).
\end{aligned}
$$

The first inequality follows from (44), and the second follows because $Q_{x^{1-\alpha}}(\rho_\mathcal{N}\|\sigma_\mathcal{N}) \leqslant Q_{x^{1-\alpha}}(\rho\|\sigma) \leqslant Q_{x^{-1}}(\rho\|\sigma)^{\alpha-1}$, the latter following from data processing and the fact that the Petz–Rényi relative entropies are monotone non-decreasing with respect to α. □

4.2. *Recoverability via another rotated Petz map R_σ^t*

We now modify the argument from the previous section to obtain a recoverability statement involving the other rotated Petz map R_σ^t. This time we use the following integral representation, holding for $t \in \mathbb{R}$:

$$
x^{-\frac{1}{2}-it} = -\frac{\sin\left(\pi\left(-\frac{1}{2}-it\right)\right)}{\pi} \int_0^\infty \lambda^{-\frac{1}{2}-it} (\lambda + x)^{-1} \mathrm{d}\lambda \tag{50}
$$

$$
= \frac{\cosh(\pi t)}{\pi} \int_0^\infty \lambda^{-\frac{1}{2}-it} (\lambda + x)^{-1} \mathrm{d}\lambda\ . \tag{51}
$$

Lemma 4.10. *Let $t \in \mathbb{R}$ and*

$$|v_t\rangle := \left|\rho^{1/2}\right\rangle - \Delta_{\mathcal{M}}^{\frac{1}{2}+it} V_\rho \Delta_{\mathcal{N}}^{-\frac{1}{2}-it} \left|\rho_{\mathcal{N}}^{1/2}\right\rangle = \left|\rho^{1/2}\right\rangle - \left|\sigma^{\frac{1}{2}+it} \sigma_{\mathcal{N}}^{-\frac{1}{2}-it} \rho_{\mathcal{N}}^{it+\frac{1}{2}} \rho^{-it}\right\rangle.$$

Then the following equality holds

$$|v_t\rangle = \frac{\cosh(\pi t)}{\pi} \Delta_{\mathcal{M}}^{\frac{1}{2}+it} \int_0^\infty \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda, \tag{52}$$

where $|w_\lambda\rangle$ is defined in (21), and the following inequality holds

$$\left\|\rho - R_\sigma^{-t}(\rho_{\mathcal{N}})\right\|_1 \leqslant 2\||v_t\rangle\|_2. \tag{53}$$

Proof. Using the integral representation in (51) for $\Delta_{\mathcal{M}}$ and $\Delta_{\mathcal{N}}$, we find that

$$\Delta_{\mathcal{M}}^{-\frac{1}{2}-it} |v_t\rangle = \Delta_{\mathcal{M}}^{-\frac{1}{2}-it} \left|\rho^{1/2}\right\rangle - V_\rho \Delta_{\mathcal{N}}^{-\frac{1}{2}-it} \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \tag{54}$$

$$= \frac{\cosh(\pi t)}{\pi} \left(\int_0^\infty \lambda^{-\frac{1}{2}-it} (\Delta_{\mathcal{M}} + \lambda)^{-1} \, \mathrm{d}\lambda \left|\rho^{1/2}\right\rangle \right.$$

$$\left. - V_\rho \left(\int_0^\infty \lambda^{-\frac{1}{2}-it} (\Delta_{\mathcal{N}} + \lambda)^{-1} \, \mathrm{d}\lambda \right) \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \right) \tag{55}$$

$$= \frac{\cosh(\pi t)}{\pi} \int_0^\infty \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda, \tag{56}$$

where $|w_\lambda\rangle$ is defined in (21). Applying $\Delta_{\mathcal{M}}^{1/2+it}$ leads to (52):

$$\Delta_{\mathcal{M}}^{1/2+it} \left(\Delta_{\mathcal{M}}^{-\frac{1}{2}-it} \left|\rho^{1/2}\right\rangle - V_\rho \Delta_{\mathcal{N}}^{-\frac{1}{2}-it} \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \right)$$

$$= \left|\rho^{1/2}\right\rangle - \Delta_{\mathcal{M}}^{\frac{1}{2}+it} V_\rho \Delta_{\mathcal{N}}^{-\frac{1}{2}-it} \left|\rho_{\mathcal{N}}^{1/2}\right\rangle$$

$$= \left|\rho^{1/2}\right\rangle - \left|\sigma^{\frac{1}{2}+it} \sigma_{\mathcal{N}}^{-\frac{1}{2}-it} \rho_{\mathcal{N}}^{\frac{1}{2}+it} \rho^{-it}\right\rangle.$$

The inequality in (53) follows from (25) and the following identity:

$$\sigma^{\frac{1}{2}+it} \sigma_{\mathcal{N}}^{-\frac{1}{2}-it} \rho_{\mathcal{N}}^{\frac{1}{2}+it} \rho^{-it} \left(\sigma^{\frac{1}{2}+it} \sigma_{\mathcal{N}}^{-\frac{1}{2}-it} \rho_{\mathcal{N}}^{\frac{1}{2}+it} \rho^{-it} \right)^* = \sigma^{\frac{1}{2}+it} \sigma_{\mathcal{N}}^{-\frac{1}{2}-it} \rho_{\mathcal{N}} \sigma_{\mathcal{N}}^{-\frac{1}{2}+it} \sigma^{\frac{1}{2}-it}$$

$$= R_\sigma^{-t}(\rho_{\mathcal{N}}),$$

where $R_\sigma^{-t}(\rho_{\mathcal{N}})$ is defined through (11). □

We have the following estimate of $\||v_t\rangle\|_2$:

Lemma 4.11. *Let $f : (0, \infty) \to \mathbb{R}$ be a regular operator anti-monotone function, and let $\mathrm{d}\nu$ be the measure in its integral representation. Let $Q_{x^{-1}}(\rho\|\sigma) = \left\langle \rho^{1/2} \right| \Delta_{\mathcal{M}}^{-1} \left|\rho^{1/2}\right\rangle = \tau(\rho^2\sigma^{-1})$. Suppose for $0 < S < T < \infty$ that there exists $c(S, T) > 0$ such that on the interval (S, T)*

$$\mathrm{d}\lambda \leqslant c(S, T) \, \mathrm{d}\nu(\lambda).$$

Then

$$\| |v_t\rangle \|_2 \leqslant \frac{\cosh(\pi t)}{\pi} \Big(4(Q_{x^{-1}}(\rho\|\sigma)S)^{1/2} + [c(S,T)\ln(T/S)]^{1/2} \\ \times (Q_f(\rho\|\sigma) - Q_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}))^{1/2} + 4T^{-1/2} \Big).$$

Proof. By applying (52), consider that

$$\begin{aligned}
\| |v_t\rangle \|_2 &= \frac{\cosh(\pi t)}{\pi} \left\| \Delta_{\mathcal{M}}^{1/2} \int_0^\infty \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda \right\|_2 \\
&\leqslant \frac{\cosh(\pi t)}{\pi} \left(\left\| \Delta_{\mathcal{M}}^{1/2} \int_0^S \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda \right\|_2 + \left\| \Delta_{\mathcal{M}}^{1/2} \int_S^T \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda \right\|_2 \right. \\
&\quad \left. + \left\| \Delta_{\mathcal{M}}^{1/2} \int_T^\infty \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda \right\|_2 \right) \\
&=: \frac{\cosh(\pi t)}{\pi} (\|\mathrm{I}\|_2 + \|\mathrm{II}\|_2 + \|\mathrm{III}\|_2).
\end{aligned}$$

For the terms above, we show the following estimates:

$$\begin{aligned}
\|\mathrm{I}\|_2 &= \left\| \Delta_{\mathcal{M}}^{\frac{1}{2}} \left(\int_0^S \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda \right) \right\|_2 \leqslant 4S^{1/2}[Q_{x^{-1}}(\rho\|\sigma)]^{1/2}, \\
\|\mathrm{II}\|_2 &= \left\| \Delta_{\mathcal{M}}^{\frac{1}{2}} \left(\int_S^T \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda \right) \right\|_2 \leqslant \left(c(S,T) \ln\left(\frac{T}{S}\right) (Q_f(\rho\|\sigma) - Q_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})) \right)^{1/2}, \\
\|\mathrm{III}\|_2 &= \left\| \Delta_{\mathcal{M}}^{\frac{1}{2}} \left(\int_T^\infty \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda \right) \right\|_2 \leqslant 4T^{-1/2}.
\end{aligned}$$

For the first term, we define the following function:

$$h_S^t(x) := \int_0^S \lambda^{-\frac{1}{2}-it} \frac{1}{\lambda + x} \, \mathrm{d}\lambda \,.$$

Thus

$$\int_0^S \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda = h_S^t(\Delta_{\mathcal{M}}) \left| \rho^{1/2} \right\rangle - V_\rho h_S^t(\Delta_{\mathcal{N}}) \left| \rho_{\mathcal{N}}^{1/2} \right\rangle,$$

leading to

$$\Delta_{\mathcal{M}}^{\frac{1}{2}} \left(\int_0^S \lambda^{-\frac{1}{2}-it} |w_\lambda\rangle \, \mathrm{d}\lambda \right) = \Delta_{\mathcal{M}}^{\frac{1}{2}} h_S^t(\Delta_{\mathcal{M}}) \left| \rho^{1/2} \right\rangle - \Delta_{\mathcal{M}}^{\frac{1}{2}} V_\rho h_S^t(\Delta_{\mathcal{N}}) \left| \rho_{\mathcal{N}}^{1/2} \right\rangle.$$

Note that for $x \geqslant 0$,

$$|h_S^t(x)| \leqslant \int_0^S \lambda^{-\frac{1}{2}} \frac{1}{\lambda + x} \, \mathrm{d}\lambda = 2x^{-1/2} \arctan\left(\frac{\sqrt{S}}{\sqrt{x}}\right) =: h_S(x),$$

so that

$$|h_S^t(x)|^2 \leqslant 4x^{-1} \arctan^2 \left(\frac{\sqrt{S}}{\sqrt{x}}\right).$$

Here h_S is the function h_S^t with $t = 0$. Using the spectral theorem for the probability measure $\mathrm{d}\mu(s) = \mathrm{d}\left\langle \rho^{1/2}\right| E_{[0,s]}(\Delta_{\mathcal{M}}) \left|\rho^{1/2}\right\rangle$, we find that

$$\begin{aligned}
\left\| \Delta_{\mathcal{M}}^{\frac{1}{2}} h_S^t(\Delta_{\mathcal{M}}) \left|\rho^{1/2}\right\rangle \right\|_2^2 &= \left\langle \rho^{1/2} \right| \Delta_{\mathcal{M}} h_S^t(\Delta_{\mathcal{M}})^* h_S^t(\Delta_{\mathcal{M}}) \left|\rho^{1/2}\right\rangle \\
&\leqslant \left\langle \rho^{1/2} \right| \Delta_{\mathcal{M}} h_S^0(\Delta_{\mathcal{M}})^* h_S^0(\Delta_{\mathcal{M}}) \left|\rho^{1/2}\right\rangle = \int_0^\infty s\, h_S^2(s)\, \mathrm{d}\mu(s) \\
&= \int_0^\infty 4 \arctan^2\left(\frac{\sqrt{S}}{\sqrt{s}}\right) \mathrm{d}\mu(s) \\
&\leqslant 4S \int_0^\infty \frac{1}{s}\, \mathrm{d}\mu(s) = 4S \left\langle \rho^{1/2} \right| \Delta_{\mathcal{M}}^{-1} \left|\rho^{1/2}\right\rangle \\
&= 4S\, Q_{x^{-1}}(\rho\|\sigma),
\end{aligned}$$

where we use again the inequality $\arctan(x) \leqslant x$, holding for $x \geqslant 0$. Similarly,

$$\begin{aligned}
\left\| \Delta_{\mathcal{M}}^{\frac{1}{2}} V_\rho h_S^t(\Delta_{\mathcal{N}}) \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \right\|_2^2 &= \left\langle \rho_{\mathcal{N}}^{1/2} \right| h_S^t(\Delta_{\mathcal{N}})^* V_\rho^* \Delta_{\mathcal{M}} V_\rho h_S^t(\Delta_{\mathcal{N}}) \left|\rho^{1/2}\right\rangle \\
&\leqslant \left\langle \rho_{\mathcal{N}}^{1/2} \right| h_S^t(\Delta_{\mathcal{N}})^* \Delta_{\mathcal{N}} h_S^t(\Delta_{\mathcal{N}}) \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \\
&= \left\langle \rho_{\mathcal{N}}^{1/2} \right| \Delta_{\mathcal{N}} h_S(\Delta_{\mathcal{N}})^2 \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \\
&\leqslant 4S \left\langle \rho_{\mathcal{N}}^{1/2} \right| \Delta_{\mathcal{N}}^{-1} \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \\
&\leqslant 4S \left\langle \rho^{1/2} \right| \Delta_{\mathcal{M}}^{-1} \left|\rho^{1/2}\right\rangle = 4S\, Q_{x^{-1}}(\rho\|\sigma),
\end{aligned}$$

where we used $V_\rho^* \Delta_{\mathcal{M}} V_\rho = \Delta_{\mathcal{N}}$ and the fact that $x \mapsto x^{-1}$ is an operator anti-monotone and operator convex function. Therefore

$$\begin{aligned}
\|\mathrm{I}\|_2 &\leqslant \left\| \Delta_{\mathcal{M}}^{\frac{1}{2}} \int_0^S \lambda^{-\frac{1}{2}-it} (\Delta_{\mathcal{M}} + \lambda) \left|\rho^{1/2}\right\rangle \right\|_2 + \left\| \Delta_{\mathcal{M}}^{\frac{1}{2}} \int_0^S \lambda^{-\frac{1}{2}-it} V_\rho (\Delta_{\mathcal{N}} + \lambda) \left|\rho_{\mathcal{N}}^{1/2}\right\rangle \right\|_2 \\
&\leqslant 2S^{1/2} [Q_{x^{-1}}(\rho\|\sigma)]^{1/2} + 2S^{1/2} [Q_{x^{-1}}(\rho\|\sigma)]^{1/2} \leqslant 4S^{1/2} [Q_{x^{-1}}(\rho\|\sigma)]^{1/2}.
\end{aligned}$$

For the second term, consider that

$$\begin{aligned}
\|\mathrm{II}\|_2 &= \left\| \Delta_{\mathcal{M}}^{1/2} \int_S^T \lambda^{-1/2-it} |w_\lambda\rangle\, \mathrm{d}\lambda \right\|_2 \leqslant \int_S^T \lambda^{-1/2} \left\| \Delta_{\mathcal{M}}^{1/2} |w_\lambda\rangle \right\|_2 \mathrm{d}\lambda \\
&\leqslant \left(\int_S^T \lambda^{-1} \mathrm{d}\lambda \right)^{1/2} \left(\int_S^T \left\| \Delta_{\mathcal{M}}^{1/2} |w_\lambda\rangle \right\|_2^2 \mathrm{d}\lambda \right)^{1/2} \leqslant (\ln(T/S))^{1/2} \left(\int_S^T F(\lambda)\, \mathrm{d}\lambda \right)^{1/2} \\
&\leqslant (\ln(T/S))^{1/2} \left(c(S,T) \int_S^T F(\lambda)\, \mathrm{d}\nu(\lambda) \right)^{1/2} \\
&\leqslant \left(c(S,T) \ln(T/S) (Q_f(\rho\|\sigma) - Q_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})) \right)^{1/2}.
\end{aligned}$$

For the third term, consider that

$$\Delta_{\mathcal{M}}^{1/2}\int_T^\infty \lambda^{-1/2-it}\,|w_\lambda\rangle\,\mathrm{d}\lambda = \Delta_{\mathcal{M}}^{1/2}\int_T^\infty \lambda^{-1/2-it}\frac{1}{\lambda+\Delta_{\mathcal{M}}}\,\mathrm{d}\lambda\left|\rho^{1/2}\right\rangle - \Delta_{\mathcal{M}}^{1/2}V_\rho\int_T^\infty \lambda^{-1/2-it}\frac{1}{\lambda+\Delta_{\mathcal{N}}}\,\mathrm{d}\lambda\left|\rho_{\mathcal{N}}^{1/2}\right\rangle.$$

Let us consider the integral

$$\int_T^\infty \lambda^{-1/2-it}\frac{1}{\lambda+x}\,\mathrm{d}\lambda = x^{-1/2-it}\int_{\frac{T}{x}}^\infty \lambda^{-1/2-it}\frac{1}{\lambda+1}\,\mathrm{d}\lambda.$$

Note that the function $\lambda\mapsto\lambda^{-1/2-it}\frac{1}{\lambda+1}$ is bounded and integrable on $(0,\infty)$. We define the continuous function

$$g_T^t(x) := \int_T^\infty \lambda^{-1/2-it}\frac{1}{\lambda+x}\,\mathrm{d}\lambda.$$

Then

$$\int_T^\infty \lambda^{-1/2-it}\frac{1}{\lambda+\Delta_{\mathcal{M}}}\,\mathrm{d}\lambda = g_T^t(\Delta_{\mathcal{M}})\,.$$

For $x\geqslant 0$,

$$|g_T^t(x)|^2 \leqslant x^{-1}\left(\int_{\frac{T}{x}}^\infty \lambda^{-1/2}\frac{1}{\lambda+1}\,\mathrm{d}\lambda\right)^2 = 4x^{-1}\arctan^2\left(\frac{\sqrt{x}}{\sqrt{T}}\right)\,.$$

Therefore

$$\begin{aligned}\left\|\Delta_{\mathcal{M}}^{1/2}\int_T^\infty \lambda^{-1/2-it}\left(\frac{1}{\lambda+\Delta_{\mathcal{M}}}\right)\mathrm{d}\lambda\left|\rho^{1/2}\right\rangle\right\|_2^2 &= \left\langle\rho^{1/2}\right|g_T^t(\Delta_{\mathcal{M}})^*\Delta_{\mathcal{M}}g_T^t(\Delta_{\mathcal{M}})\left|\rho^{1/2}\right\rangle\\ &\leqslant \left\langle\rho^{1/2}\right|4\arctan^2\left(\frac{\sqrt{\Delta_{\mathcal{M}}}}{\sqrt{T}}\right)\left|\rho^{1/2}\right\rangle\\ &\leqslant \frac{4}{T}\left\langle\rho^{1/2}\right|\Delta_{\mathcal{M}}\left|\rho^{1/2}\right\rangle\\ &= \frac{4}{T},\end{aligned}$$

where we used again the inequality $\arctan(x)\leqslant x$, holding for $x\geqslant 0$. Similarly,

$$\left\|\Delta_{\mathcal{M}}^{1/2}V_\rho\int_T^\infty \lambda^{-1/2-it}\frac{1}{\lambda+\Delta_{\mathcal{N}}}\mathrm{d}\lambda\left|\rho_{\mathcal{N}}^{1/2}\right\rangle\right\|_2^2 \leqslant \frac{4}{T}\left\langle\rho_{\mathcal{N}}^{1/2}\right|\Delta_{\mathcal{N}}\left|\rho_{\mathcal{N}}^{1/2}\right\rangle = \frac{4}{T}.$$

Putting the estimates together, we conclude the proof. □

A direct consequence of lemmas 4.10 and 4.11, as well as the symmetry of the function $\cosh(\pi t)$ about $t = 0$, is the following general bound on the recoverability error in terms of a standard f-divergence:

Corollary 4.12. *Considering the same hypotheses of lemma 4.11, the following inequality holds*

$$\|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi}\Big(4S^{1/2}Q_{x^{-1}}(\rho\|\sigma)^{1/2} + [c(S,T)\ln(T/S)]^{1/2} \times (Q_f(\rho\|\sigma) - Q_f(\rho_\mathcal{N}\|\sigma_\mathcal{N}))^{1/2} + 4T^{-1/2}\Big).$$

4.2.1. Recoverability for quantum relative entropy. We have the following estimate for the quantum relative entropy $D(\rho\|\sigma)$, as defined in (7):

Theorem 4.13. *Let $\mathcal{M}$ be a finite von Neumann algebra, and let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. Let ρ and σ be faithful density operators of $\mathcal{M}$, and let $\rho_\mathcal{N}$ and $\sigma_\mathcal{N}$ be the respective reduced density operators on $\mathcal{N}$. Denote $Q_{x^{-1}}(\rho\|\sigma) = \tau(\rho^2\sigma^{-1})$. Then for all $\varepsilon \in (0, 1/2)$ and $t \in \mathbb{R}$,*

$$D(\rho\|\sigma) - D(\rho_\mathcal{N}\|\sigma_\mathcal{N}) \geqslant \left(K(Q_{x^{-1}}(\rho\|\sigma), \varepsilon)\frac{\pi}{2}\|\rho - R_\sigma(\rho_\mathcal{N})\|_1\right)^{\frac{1}{1/2-\varepsilon}}, \tag{57}$$

$$D(\rho\|\sigma) - D(\rho_\mathcal{N}\|\sigma_\mathcal{N}) \geqslant \left(K(Q_{x^{-1}}(\rho\|\sigma), \varepsilon)\frac{\pi}{2\cosh(\pi t)}\|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1\right)^{\frac{1}{1/2-\varepsilon}}, \tag{58}$$

$$D(\rho\|\sigma) - D(\rho_\mathcal{N}\|\sigma_\mathcal{N}) \geqslant \left(K(Q_{x^{-1}}(\rho\|\sigma), \varepsilon)\|\rho - R^u_\sigma(\rho_\mathcal{N})\|_1\right)^{\frac{1}{1/2-\varepsilon}}, \tag{59}$$

where the constant

$$K(Q_{x^{-1}}(\rho\|\sigma), \varepsilon) := \left(4\sqrt{Q_{x^{-1}}(\rho\|\sigma)} + 4 + (\varepsilon e)^{-1/2}\right)^{-1}. \tag{60}$$

Proof. Consider from (7) that $f(x) = -\log x$, for which we have the following integral representation:

$$-\log x = \int_0^\infty \left(\frac{1}{\lambda + x} - \frac{\lambda}{\lambda^2 + 1}\right) \mathrm{d}\lambda,$$

where $\mathrm{d}\lambda$ is the Lebesgue measure. Thus $c(S, T) = 1$ for all $0 \leqslant S \leqslant T \leqslant \infty$. Then, by applying corollary 4.12, we find that

$$\|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi}\Big(4Q_{x^{-1}}(\rho\|\sigma)^{1/2}S^{1/2} + \left(\ln(T/S)(D(\rho\|\sigma) - D(\rho_\mathcal{N}\|\sigma_\mathcal{N}))\right)^{1/2} + 4T^{-1/2}\Big). \tag{61}$$

Define the following constants:

$$a := 4Q_{x^{-1}}(\rho\|\sigma)^{1/2}, \qquad b := (D(\rho\|\sigma) - D(\rho_\mathcal{N}\|\sigma_\mathcal{N}))^{1/2}, \qquad c := 4.$$

We want to minimize the following function over $0 < S < T < \infty$,

$$F(S,T) = aS^{1/2} + b\sqrt{\ln(T/S)} + cT^{-1/2}\ .$$

Set $\delta := \min\{D(\rho\|\sigma) - D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}), 1\}$. Then a rough choice is $S = T^{-1} = \delta$, and we find that

$$\begin{aligned}\|\rho - R^t_\sigma(\rho_{\mathcal{N}})\|_1 &\leqslant \frac{2\ \cosh(\pi t)}{\pi}\left(4Q_{x^{-1}}(\rho\|\sigma)^{1/2}\delta^{1/2} + (2\delta|\ln\ \delta|)^{1/2} + 4\delta^{1/2}\right)\\ &= \frac{2\ \cosh(\pi t)}{\pi}\left(4\sqrt{Q_{x^{-1}}(\rho\|\sigma)} + 4 + \sqrt{2|\ln\ \delta|}\right)\delta^{1/2}\\ &\leqslant \frac{2\ \cosh(\pi t)}{\pi}\left(4\sqrt{Q_{x^{-1}}(\rho\|\sigma)} + 4 + (\varepsilon e)^{-1/2}\right)\delta^{1/2-\varepsilon}.\end{aligned}$$

In the case that $\delta = 1$, the first inequality is trivial, following because $\|\rho - R^t_\sigma(\rho_{\mathcal{N}})\|_1 \leqslant 2$, $4Q_{x^{-1}}(\rho\|\sigma)^{1/2}\delta^{1/2} + (2\delta|\ln\ \delta|)^{1/2} \geqslant 0$, and $\frac{2\ \cosh(\pi t)}{\pi}4\delta^{1/2} \geqslant 2$ for all $t \in \mathbb{R}$ in this case. Otherwise, the first inequality is a consequence of (61). The last inequality is a consequence of the inequalities $\delta^\varepsilon < 1$ and $\delta^\varepsilon\sqrt{2|\ln\ \delta|} \leqslant (\varepsilon e)^{-1/2}$, holding for $0 < \delta < 1$ and $\varepsilon > 0$.

The rest of the proof is similar to the proof of theorem 4.5. □

Remark 4.14. It is a consequence of the result in [JRS+18] that the following inequality holds for the universal recovery map R^u_σ:

$$D(\rho\|\sigma) - D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) \geqslant 4\|\rho - R^u_\sigma(\rho_{\mathcal{N}})\|_1^2\ . \tag{62}$$

For R^u_σ, the inequality in (62) is stronger than our estimate in (59). Nevertheless, theorem 4.13 above provides an error estimate with the rotated Petz map R^t_σ for each t.

4.2.2. Recoverability for Petz–Rényi relative (quasi-)entropy. We now turn to the Petz–Rényi relative quasi-entropy, as defined in (39)–(42).

Theorem 4.15. *Let $s \in (-1,0)\cup(0,1)$. Denote $Q_{x^{-1}}(\rho\|\sigma) = \tau(\rho^2\sigma^{-1})$. Then the following inequalities hold for all $t \in \mathbb{R}$ and $\varepsilon \in (0,(1-|s|)/2)$:*

$$|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})| \geqslant \left(K(s, Q_{x^{-1}}(\rho\|\sigma), \varepsilon)\frac{\pi}{2}\|\rho - R_\sigma(\rho_{\mathcal{N}})\|_1\right)^{\frac{1}{\frac{1-|s|}{2}-\varepsilon}}, \tag{63}$$

$$|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})| \geqslant \left(K(s, Q_{x^{-1}}(\rho\|\sigma), \varepsilon)\frac{\pi}{2\ \cosh\ \pi t}\|\rho - R^t_\sigma(\rho_{\mathcal{N}})\|_1\right)^{\frac{1}{\frac{1-|s|}{2}-\varepsilon}}, \tag{64}$$

$$|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})| \geqslant \left(K(s, Q_{x^{-1}}(\rho\|\sigma), \varepsilon)\|\rho - R^u_\sigma(\rho_{\mathcal{N}})\|_1\right)^{\frac{1}{\frac{1-|s|}{2}-\varepsilon}}, \tag{65}$$

where the constant

$$K(s, Q_{x^{-1}}(\rho\|\sigma), \varepsilon) := \left(4Q_{x^{-1}}(\rho\|\sigma)^{1/2} + \left(\frac{\pi}{e\varepsilon\ \sin(\pi|s|)}\right)^{1/2} + 4\right)^{-1}. \tag{66}$$

Proof. For $0 < s < 1$, the function $f(x) = x^s$ is operator monotone and operator concave. The integral representation of x^s is

$$x^s = \frac{\sin(\pi s)}{\pi}\int_0^\infty \lambda^s\left(\frac{1}{\lambda} - \frac{1}{\lambda + x}\right)\mathrm{d}\lambda\ .$$

Corollary 4.12 holds for $c(S,T) \leqslant \frac{\pi}{\sin(\pi s)} S^{-s}$, and we find that

$$\|\rho - R_\sigma^t(\rho_{\mathcal{N}})\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi}\Big(4S^{1/2} Q_{x^{-1}}(\rho\|\sigma)^{1/2} + \sqrt{S^{-s}\frac{\pi}{\sin(\pi s)}\ln(T/S)} |Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})|^{1/2} + 4T^{-1/2}\Big). \quad (67)$$

Define the following constants:

$$a := 4Q_{x^{-1}}(\rho\|\sigma)^{1/2}, \qquad b := \left(\frac{\pi}{\sin(\pi s)}|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})|\right)^{1/2}, \qquad c := 4,$$

and the function

$$F(S,T) = aS^{1/2} + b\sqrt{S^{-s}\ln(T/S)} + cT^{-1/2}.$$

Setting $\delta := \min\{|Q_s(\rho\|\sigma) - Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})|, 1\}$ and $S = T^{-1} = \delta$, we find that

$$\begin{aligned}
&\|\rho - R_\sigma^t(\rho_{\mathcal{N}})\|_1 \\
&\leqslant \frac{2\cosh(\pi t)}{\pi}\left(4\delta^{1/2} Q_{x^{-1}}(\rho\|\sigma)^{1/2} + \delta^{(1-s)/2}\sqrt{\frac{2\pi}{\sin(\pi s)}|\ln\delta|} + 4\delta^{1/2}\right) && (68)\\
&= \frac{2\cosh(\pi t)}{\pi}\left(4Q_{x^{-1}}(\rho\|\sigma)^{1/2}\delta^{s/2} + \sqrt{\frac{2\pi}{\sin(\pi s)}|\ln\delta|} + 4\delta^{s/2}\right)\delta^{(1-s)/2} && (69)\\
&\leqslant \frac{2\cosh(\pi t)}{\pi}\left(4Q_{x^{-1}}(\rho\|\sigma)^{1/2} + \left(\frac{\pi}{e\varepsilon\,\sin(\pi s)}\right)^{1/2} + 4\right)\delta^{(1-s)/2-\varepsilon}. && (70)
\end{aligned}$$

In the case that $\delta = 1$, the first inequality is trivial, following from the facts that $\|\rho - R_\sigma^t(\rho_{\mathcal{N}})\|_1 \leqslant 2$,

$$4\delta^{1/2} Q_{x^{-1}}(\rho\|\sigma)^{1/2} + \delta^{(1-s)/2}\sqrt{\frac{2\pi}{\sin(\pi s)}|\ln\delta|} \geqslant 0,$$

and $\frac{2\cosh(\pi t)}{\pi} 4\delta^{1/2} \geqslant 2$ for all $t \in \mathbb{R}$ in this case. Otherwise, the first inequality is a consequence of (67). The last inequality is a consequence of the inequalities $\delta^{s/2+\varepsilon} \leqslant 1$ and $\delta^{\varepsilon}\sqrt{2|\ln\delta|} \leqslant (\varepsilon e)^{-1/2}$, holding for $0 < \delta < 1$.

For $-1 < s < 0$, the function $f(x) = x^s$ is operator anti-monotone and operator convex. The integral representation of x^s in this case is

$$x^s = \frac{\sin(\pi|s|)}{\pi}\int_0^\infty \frac{\lambda^s}{\lambda + x}\,d\lambda.$$

Then the constant $c(S,T) \leqslant \frac{\pi}{\sin(\pi|s|)} T^{|s|}$. The following inequality holds as a consequence of corollary 4.12:

$$\|\rho - R_\sigma^t(\rho_{\mathcal{N}})\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi}\Big(4S^{1/2} Q_{x^{-1}}(\rho\|\sigma)^{1/2}$$

$$+\sqrt{\frac{\pi}{\sin(\pi|s|)}T^{|s|}\ \ln(T/S)}\left(Q_s(\rho\|\sigma)\right)\ -Q_s(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})\big)^{1/2}+4T^{-1/2}\bigg).$$

The rest of the analysis is the same as that given for the case $0<s<1$, by taking $S=T^{-1}=\delta$.
□

Following the same method of proof given for corollary 4.9, we arrive at the following corollary:

Corollary 4.16. *Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra, and let $\mathcal{N}\subset\mathcal{M}$ be a subalgebra. Let ρ and σ be two faithful states. For $\alpha\in(0,1)$, $\varepsilon\in(0,\alpha/2)$, and $t\in\mathbb{R}$,*

$$\begin{aligned}
&D_\alpha(\rho\|\sigma)-D_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})\\
&\qquad\geqslant\frac{1}{1-\alpha}\ \log\left(1+\left(K(1-\alpha,Q_{x^{-1}}(\rho\|\sigma),\varepsilon)\frac{\pi}{2\cosh\pi t}\left\|\rho-R_\sigma^t(\rho_{\mathcal{N}})\right\|_1\right)^{\frac{1}{\alpha/2-\varepsilon}}\right),
\end{aligned}$$

and for $\alpha\in(1,2)$, $\varepsilon\in(0,(2-\alpha)/2)$, and $t\in\mathbb{R}$,

$$\begin{aligned}
D_\alpha(\rho\|\sigma)-D_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})\geqslant\frac{1}{\alpha-1}\log\Bigg(1+&\frac{1}{Q_{x^{-1}}(\rho\|\sigma)^{\alpha-1}}\\
&\times\left(\frac{K(1-\alpha,Q_{x^{-1}}(\rho\|\sigma),\varepsilon)\pi}{2(\alpha+1)\cosh\pi t}\left\|\rho-R_\sigma^t(\rho_{\mathcal{N}})\right\|_1\right)^{\frac{1}{(2-\alpha)/2-\varepsilon}}\Bigg),
\end{aligned}$$

where the constant $K(1-\alpha,Q_{x^{-1}}(\rho\|\sigma),\varepsilon)$ is given by (66).

4.3. Recoverability for optimized f-divergence

We now discuss recoverability for the optimized *f*-divergence. Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra with trace τ, and let $\mathcal{N}\subset\mathcal{M}$ be a subalgebra. Let $\rho,\sigma\in\mathcal{M}$ be two faithful states, and let $E(\rho)=\rho_{\mathcal{N}}$ and $E(\sigma)=\sigma_{\mathcal{N}}$ be the respective reduced density operators on $\mathcal{N}$. Let f be an operator anti-monotone function. Recall from (6) that the optimized *f*-divergences are defined as follows:

$$\begin{aligned}
\widetilde{Q}_f(\rho\|\sigma)&=\sup_{\omega\in D_+(\mathcal{M})}\left\langle\rho^{1/2}\right|f(\Delta_{\mathcal{M}}(\sigma,\omega))\left|\rho^{1/2}\right\rangle,\\
\widetilde{Q}_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})&=\sup_{\omega_{\mathcal{N}}\in D_+(\mathcal{N})}\left\langle\rho_{\mathcal{N}}^{1/2}\right|f(\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}}))\left|\rho_{\mathcal{N}}^{1/2}\right\rangle\ ,
\end{aligned}$$

where the supremum is with respect to all invertible density operators $\omega\in D_+(\mathcal{M})$ (resp. $\omega_{\mathcal{N}}\in D_+(\mathcal{N})$). Let $V_\rho:L_2(\mathcal{N})\to L_2(\mathcal{M})$ denote the following isometry:

$$V_\rho(a\left|\rho_{\mathcal{N}}^{1/2}\right\rangle)=a\left|\rho^{1/2}\right\rangle\ ,\quad\forall\ a\in\mathcal{N},$$

with a similar definition for V_σ.

Lemma 4.17. *Let $\rho,\sigma,\omega\in D_+(\mathcal{M})$. The following equality holds*

$$V_\rho^*\Delta_{\mathcal{M}}(\sigma,R_\rho(\omega_{\mathcal{N}}))V_\rho=\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}}),$$

where R_ρ is the Petz recovery map from (10). As a consequence, the following inequality holds for all operator anti-monotone functions $f:(0,\infty)\to\mathbb{R}$:

$$\left\langle \rho_{\mathcal{N}}^{1/2}\right| f(\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}}))\left|\rho_{\mathcal{N}}^{1/2}\right\rangle \leqslant \left\langle \rho^{1/2}\right| f(\Delta_{\mathcal{M}}(\sigma,R_\rho(\omega_{\mathcal{N}})))\left|\rho^{1/2}\right\rangle .$$

Proof. Recall the recovery map $R_\rho(\omega_{\mathcal{N}})=\rho^{1/2}\rho_{\mathcal{N}}^{-1/2}\omega_{\mathcal{N}}\rho_{\mathcal{N}}^{-1/2}\rho^{1/2}$. By definition of Δ and V_ρ, we have for any $a\in\mathcal{N}$,

$$\begin{aligned}
V_\rho^*\Delta(\sigma,R_\rho(\omega_{\mathcal{N}}))V_\rho\left|a\rho_{\mathcal{N}}^{1/2}\right\rangle &= V_\rho^*\Delta(\sigma,R_\rho(\omega_{\mathcal{N}}))\left|a\rho^{1/2}\right\rangle\\
&= V_\rho^*\left|\sigma a\rho^{1/2}R_\rho(\omega_{\mathcal{N}})^{-1}\right\rangle\\
&= \left|E(\sigma a\rho_{\mathcal{N}}^{1/2}\omega_{\mathcal{N}}^{-1}\rho_{\mathcal{N}}^{1/2})\rho_{\mathcal{N}}^{-1/2}\right\rangle\\
&= \left|\sigma_{\mathcal{N}}a\rho_{\mathcal{N}}^{1/2}\omega_{\mathcal{N}}^{-1}\right\rangle\\
&= \Delta(\sigma_{\mathcal{N}},\omega_{\mathcal{N}})\left|a\rho_{\mathcal{N}}^{1/2}\right\rangle,
\end{aligned}$$

where in the second last equality we used the fact that $\rho_{\mathcal{N}}^{1/2}\omega_{\mathcal{N}}^{-1}\rho_{\mathcal{N}}^{1/2}\in\mathcal{N}$. This verifies the claimed equality.

Now using operator convexity and operator anti-monotonicity of f, we find that

$$\begin{aligned}
\left\langle \rho_{\mathcal{N}}^{1/2}\right| f(\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}}))\left|\rho_{\mathcal{N}}^{1/2}\right\rangle &= \left\langle \rho_{\mathcal{N}}^{1/2}\right| f(V_\rho^*\Delta_{\mathcal{M}}(\sigma,R_\rho(\omega_{\mathcal{N}}))V_\rho)\left|\rho_{\mathcal{N}}^{1/2}\right\rangle\\
&\leqslant \left\langle \rho_{\mathcal{N}}^{1/2}\right| V_\rho^* f(\Delta_{\mathcal{M}}(\sigma,R_\rho(\omega_{\mathcal{N}})))V_\rho\left|\rho_{\mathcal{N}}^{1/2}\right\rangle\\
&= \left\langle \rho^{1/2}\right| f(\Delta_{\mathcal{M}}(\sigma,R_\rho(\omega_{\mathcal{N}})))\left|\rho^{1/2}\right\rangle .
\end{aligned}$$ □

For all $\varepsilon>0$, we can choose $\omega_{\mathcal{N}}\in D_+(\mathcal{N})$ such that

$$\left\langle \rho_{\mathcal{N}}^{1/2}\right| f(\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}}))\left|\rho_{\mathcal{N}}^{1/2}\right\rangle \geqslant \widetilde{Q}_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})-\varepsilon.$$

Note that by lemma 4.17,

$$\begin{aligned}
\left\langle \rho^{1/2}\right| \Delta_{\mathcal{M}}(\sigma,R_\rho(\omega_{\mathcal{N}}))\left|\rho^{1/2}\right\rangle &= \left\langle \rho_{\mathcal{N}}^{1/2}\right| V_\rho^*\Delta_{\mathcal{M}}(\sigma,R_\rho(\omega_{\mathcal{N}}))V_\rho\left|\rho_{\mathcal{N}}^{1/2}\right\rangle\\
&= \left\langle \rho_{\mathcal{N}}^{1/2}\right| \Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}})\left|\rho_{\mathcal{N}}^{1/2}\right\rangle .
\end{aligned}$$

Then

$$\begin{aligned}
\widetilde{Q}_f(\rho\|\sigma)-\widetilde{Q}_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) &= \sup_{\omega\in D_+(\mathcal{M})}\left\langle \rho^{1/2}\right| f(\Delta_{\mathcal{M}}(\sigma,\omega))\left|\rho^{1/2}\right\rangle\\
&\quad - \sup_{\omega\in D_+(\mathcal{N})}\left\langle \rho_{\mathcal{N}}^{1/2}\right| f(\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega))\left|\rho_{\mathcal{N}}^{1/2}\right\rangle\\
&\geqslant \sup_{\omega\in D_+(\mathcal{M})}\left\langle \rho^{1/2}\right| f(\Delta_{\mathcal{M}}(\sigma,\omega))\left|\rho^{1/2}\right\rangle\\
&\quad - \left\langle \rho_{\mathcal{N}}^{1/2}\right| f(\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}}))\left|\rho_{\mathcal{N}}^{1/2}\right\rangle - \varepsilon
\end{aligned}$$

$$
\begin{aligned}
&\geqslant \left\langle \rho^{1/2} \middle| f(\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}}))) \middle| \rho^{1/2} \right\rangle \\
&- \left\langle \rho_{\mathcal{N}}^{1/2} \middle| f(\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}}, \omega_{\mathcal{N}})) \middle| \rho_{\mathcal{N}}^{1/2} \right\rangle - \varepsilon \\
&= b \left\langle \rho^{1/2} \middle| \Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) \middle| \rho^{1/2} \right\rangle - b \left\langle \rho_{\mathcal{N}}^{1/2} \middle| \Delta_{\mathcal{N}}(\sigma_{\mathcal{N}}, \omega_{\mathcal{N}}) \middle| \rho_{\mathcal{N}}^{1/2} \right\rangle \\
&+ \int_0^\infty \left(\left\langle \rho^{1/2} \middle| (\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) + \lambda)^{-1} \middle| \rho^{1/2} \right\rangle \right. \\
&\left. - \left\langle \rho_{\mathcal{N}}^{1/2} \middle| (\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}}, \omega_{\mathcal{N}}) + \lambda)^{-1} \middle| \rho_{\mathcal{N}}^{1/2} \right\rangle \right) \mathrm{d}\nu(\lambda) - \varepsilon \ . \\
&= \int_0^\infty \left(\left\langle \rho^{1/2} \middle| (\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) + \lambda)^{-1} \middle| \rho^{1/2} \right\rangle \right. \\
&\left. - \left\langle \rho_{\mathcal{N}}^{1/2} \middle| (\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}}, \omega_{\mathcal{N}}) + \lambda)^{-1} \middle| \rho_{\mathcal{N}}^{1/2} \right\rangle \right) \mathrm{d}\nu(\lambda) - \varepsilon \ ,
\end{aligned}
$$

where b is the parameter and $\mathrm{d}\nu$ is the measure in the integral representation (5) of f. Denote

$$
\begin{aligned}
F(\lambda) &:= \left\langle \rho^{1/2} \middle| (\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) + \lambda)^{-1} \middle| \rho^{1/2} \right\rangle - \left\langle \rho_{\mathcal{N}}^{1/2} \middle| (\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}}, \omega_{\mathcal{N}}) + \lambda)^{-1} \middle| \rho_{\mathcal{N}}^{1/2} \right\rangle , \\
&= \langle \rho_{\mathcal{N}}^{1/2} | V_\rho^* (\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) + \lambda)^{-1} V_\rho | \rho_{\mathcal{N}}^{1/2} \rangle - \langle \rho_{\mathcal{N}}^{1/2} | (V_\rho^* (\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) + \lambda) V_\rho)^{-1} | \rho_{\mathcal{N}}^{1/2} \rangle \\
&= \langle u_\lambda | \, \Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) + \lambda \, | u_\lambda \rangle \geqslant 0 \, ,
\end{aligned}
$$

where

$$
|u_\lambda\rangle := (\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) + \lambda)^{-1} \left| \rho^{1/2} \right\rangle - V_\rho (\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}}, \omega_{\mathcal{N}}) + \lambda)^{-1} \left| \rho_{\mathcal{N}}^{1/2} \right\rangle \ ,
$$

and the last line follows from lemma 4.1. Thus, we find that

$$
\widetilde{Q}_f(\rho \| \sigma) - \widetilde{Q}_f(\rho_{\mathcal{N}} \| \sigma_{\mathcal{N}}) \geqslant \int_0^\infty F(\lambda) \, \mathrm{d}\nu(\lambda) - \varepsilon. \tag{71}
$$

Lemma 4.18. *Let $t \in \mathbb{R}$ and*

$$
|u_t\rangle := \frac{\cosh(\pi t)}{\pi} \Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}}))^{1/2+it} \int_0^\infty \lambda^{-1/2-it} \, |u_\lambda\rangle \ \mathrm{d}\lambda \ .
$$

Then the following inequality holds

$$
\left\| \rho - R_\sigma^{-t}(\rho) \right\|_1 \leqslant 2 \| |u_t\rangle \|_2 .
$$

Proof. Using the integral representation in (51), i.e.

$$
x^{-1/2-it} = \frac{\cosh(\pi t)}{\pi} \int_0^\infty \lambda^{-1/2-it} (\lambda + x)^{-1} \mathrm{d}\lambda,
$$

we find, by a similar argument to that given for (54)–(56), that

$$\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}}))^{-1/2-it}\left|\rho^{1/2}\right\rangle - V_\rho\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}})^{-1/2-it}\left|\rho_{\mathcal{N}}^{1/2}\right\rangle$$
$$= \frac{\cosh(\pi t)}{\pi}\int_0^\infty \lambda^{-1/2-it}\left|u_\lambda\right\rangle \mathrm{d}\lambda\ .$$

Applying $\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}}))^{1/2+it}$, we find that

$$|u_t\rangle = \left|\rho^{1/2}\right\rangle - \Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}}))^{1/2+it}V_\rho\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}})^{-1/2-it}\left|\rho_{\mathcal{N}}^{1/2}\right\rangle$$
$$= \left|\rho^{1/2}\right\rangle - \left|\sigma^{1/2+it}\sigma_{\mathcal{N}}^{-1/2-it}\rho_{\mathcal{N}}^{1/2}\omega_{\mathcal{N}}^{1/2+it}\rho_{\mathcal{N}}^{-1/2}\rho^{1/2}R_\rho(\omega_{\mathcal{N}})^{-1/2-it}\right\rangle.$$

For the second term above, we have the following collapse:

$$\sigma^{\frac{1}{2}+it}\sigma_{\mathcal{N}}^{-\frac{1}{2}-it}\rho_{\mathcal{N}}^{\frac{1}{2}}\omega_{\mathcal{N}}^{\frac{1}{2}+it}\rho_{\mathcal{N}}^{-\frac{1}{2}}\rho^{\frac{1}{2}}R_\rho(\omega_{\mathcal{N}})^{-\frac{1}{2}-it}\left(\sigma^{\frac{1}{2}+it}\sigma_{\mathcal{N}}^{-\frac{1}{2}-it}\rho_{\mathcal{N}}^{\frac{1}{2}}\omega_{\mathcal{N}}^{\frac{1}{2}+it}\rho_{\mathcal{N}}^{-\frac{1}{2}}\rho^{\frac{1}{2}}R_\rho(\omega_{\mathcal{N}})^{-\frac{1}{2}-it}\right)^*$$
$$= \sigma^{\frac{1}{2}+it}\sigma_{\mathcal{N}}^{-\frac{1}{2}-it}\rho_{\mathcal{N}}^{\frac{1}{2}}\omega_{\mathcal{N}}^{\frac{1}{2}+it}\rho_{\mathcal{N}}^{-\frac{1}{2}}\rho^{\frac{1}{2}}R_\rho(\omega_{\mathcal{N}})^{-\frac{1}{2}-it}R_\rho(\omega_{\mathcal{N}})^{-\frac{1}{2}+it}\rho^{\frac{1}{2}}\rho_{\mathcal{N}}^{-\frac{1}{2}}\omega_{\mathcal{N}}^{\frac{1}{2}-it}\rho_{\mathcal{N}}^{\frac{1}{2}}\sigma_{\mathcal{N}}^{-\frac{1}{2}+it}\sigma^{\frac{1}{2}-it}$$
$$= \sigma^{\frac{1}{2}+it}\sigma_{\mathcal{N}}^{-\frac{1}{2}-it}\rho_{\mathcal{N}}^{\frac{1}{2}}\omega_{\mathcal{N}}^{\frac{1}{2}+it}\rho_{\mathcal{N}}^{-\frac{1}{2}}\rho^{\frac{1}{2}}R_\rho(\omega_{\mathcal{N}})^{-1}\rho^{\frac{1}{2}}\rho_{\mathcal{N}}^{-\frac{1}{2}}\omega_{\mathcal{N}}^{\frac{1}{2}-it}\rho_{\mathcal{N}}^{\frac{1}{2}}\sigma_{\mathcal{N}}^{-\frac{1}{2}+it}\sigma^{\frac{1}{2}-it}$$
$$= \sigma^{\frac{1}{2}+it}\sigma_{\mathcal{N}}^{-\frac{1}{2}-it}\rho_{\mathcal{N}}^{\frac{1}{2}}\omega_{\mathcal{N}}^{\frac{1}{2}+it}\omega_{\mathcal{N}}^{-1}\omega_{\mathcal{N}}^{\frac{1}{2}-it}\rho_{\mathcal{N}}^{\frac{1}{2}}\sigma_{\mathcal{N}}^{-\frac{1}{2}+it}\sigma^{\frac{1}{2}-it}$$
$$= \sigma^{\frac{1}{2}+it}\sigma_{\mathcal{N}}^{-\frac{1}{2}-it}\rho_{\mathcal{N}}\sigma_{\mathcal{N}}^{-\frac{1}{2}+it}\sigma^{\frac{1}{2}-it}$$
$$= R_\sigma^{-t}(\rho),$$

where R_σ^{-t} is defined through (11). For the third equality above, we used the following:

$$R_\rho(\omega_{\mathcal{N}}) = \rho^{1/2}\rho_{\mathcal{N}}^{-1/2}\omega_{\mathcal{N}}\rho_{\mathcal{N}}^{-1/2}\rho^{1/2} \quad\iff\quad \rho_{\mathcal{N}}^{1/2}\rho^{-1/2}R_\rho(\omega_{\mathcal{N}})\rho^{-1/2}\rho_{\mathcal{N}}^{1/2} = \omega_{\mathcal{N}}.$$

After applying (25), we find that

$$\left\|\rho - R_\sigma^{-t}(\rho)\right\|_1 \leqslant 2\left\|\left|\rho^{1/2}\right\rangle - \Delta_{\mathcal{M}}^{1/2+it}(\sigma, R_\rho(\omega_{\mathcal{N}}))^{1/2}V_\rho\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}},\omega_{\mathcal{N}})^{-1/2-it}\left|\rho_{\mathcal{N}}^{1/2}\right\rangle\right\|_2$$
$$= \frac{2\,\cosh(\pi t)}{\pi}\left\|\Delta_{\mathcal{M}}^{1/2}(\sigma, R_\rho(\omega_{\mathcal{N}}))\int_0^\infty \lambda^{-1/2-it}\left|u_\lambda\right\rangle\right\|_2 = 2\||u_t\rangle\|_2. \qquad \square$$

Lemma 4.19. *Let $f:(0,\infty)\to\mathbb{R}$ be a regular operator anti-monotone function, and let $\mathrm{d}\nu$ be the measure in its integral representation. Suppose for some S and T, satisfying $0 < S < T < \infty$, that $\mathrm{d}\lambda \leqslant c(S,T)\,\mathrm{d}\nu(\lambda)$ for $c(S,T) > 0$. Then*

$$\|\rho - R_\sigma^t(\rho_{\mathcal{N}})\|_1 \leqslant \frac{2\,\cosh(\pi t)}{\pi}\Big(4S^{1/2}\widetilde{Q}_{x^{-1}}(\rho\|\sigma)^{1/2}$$
$$+\left(c(S,T)\ln\left(T/S\right)\right)^{1/2}(\widetilde{Q}_f(\rho\|\sigma) - \widetilde{Q}_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}))^{1/2} + 4T^{-1/2}\Big),$$

where $\widetilde{Q}_{x^{-1}}(\rho\,\|\sigma) = \left\|\rho^{1/2}\sigma^{-1}\rho^{1/2}\right\|_\infty = \inf\{\lambda > 0 \mid \rho \leqslant \lambda\sigma\}$.

Proof. The following argument is similar to the case of the non-optimized Q_f, as presented in the proof of lemma 4.11. We employ the shorthand $\Delta_{\mathcal{M}} \equiv \Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}}))$. Applying lemma 4.18 and the triangle inequality, we find that

$$\begin{aligned}
\left\|\rho - R_\sigma^{-t}(\rho)\right\|_1 &\leqslant \frac{2\cosh(\pi t)}{\pi}\left(\left\|\Delta_{\mathcal{M}}^{1/2+it}\int_0^S \lambda^{-1/2-it}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2\right.\\
&\quad \left.+\left\|\Delta_{\mathcal{M}}^{1/2+it}\int_S^T \lambda^{-1/2-it}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2+\left\|\Delta_{\mathcal{M}}^{1/2+it}\int_T^\infty \lambda^{-1/2-it}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2\right)\\
&= \frac{2\cosh(\pi t)}{\pi}\left(\left\|\Delta_{\mathcal{M}}^{1/2}\int_0^S \lambda^{-1/2-it}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2+\left\|\Delta_{\mathcal{M}}^{1/2}\int_S^T \lambda^{-1/2-it}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2\right.\\
&\quad \left.+\left\|\Delta_{\mathcal{M}}^{1/2}\int_T^\infty \lambda^{-1/2-it}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2\right)\\
&= \frac{2\cosh(\pi t)}{\pi}\left(\|I\|_2+\|II\|_2+\|III\|_2\right).
\end{aligned}$$

For each term, we argue similarly as in the proof of lemma 4.11, but implicitly using lemma 4.18 and (71):

$$\begin{aligned}
\|\mathrm{I}\|_2 &= \left\|\Delta_{\mathcal{M}}^{1/2}\int_0^S \lambda^{-1/2-it}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2 \leqslant 4S^{1/2}\left\langle \rho^{1/2}\right|\Delta_{\mathcal{M}}^{-1}\left|\rho^{1/2}\right\rangle^{1/2} \leqslant 4S^{1/2}\widetilde{Q}_{x^{-1}}(\rho\|\sigma)^{1/2},\\
\|\mathrm{II}\|_2 &= \left\|\int_S^T \lambda^{-1/2-it}\Delta_{\mathcal{M}}^{1/2}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2 \leqslant \left(\int_S^T \lambda^{-1}\mathrm{d}\lambda\right)^{1/2}\left(\int_S^T \left\|\Delta_{\mathcal{M}}^{1/2}\,|u_\lambda\rangle\right\|_2^2\mathrm{d}\lambda\right)^{1/2}\\
&\leqslant \sqrt{c(S,T)\ln(T/S)(\widetilde{Q}_f(\rho\|\sigma)-\widetilde{Q}_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})+\varepsilon)}\,,\\
\|\mathrm{III}\|_2 &= \left\|\int_T^\infty \lambda^{-1/2-it}\Delta_{\mathcal{M}}^{1/2}\,|u_\lambda\rangle\,\mathrm{d}\lambda\right\|_2 \leqslant 4T^{-1/2}.
\end{aligned}$$

Note that here

$$\begin{aligned}
\widetilde{Q}_{x^{-1}}(\rho\|\sigma) &= \sup_{\omega\in D_+(\mathcal{M})}\left\langle \rho^{1/2}\right|\Delta(\sigma,\omega)^{-1}\left|\rho^{1/2}\right\rangle\\
&= \sup_{\omega\in D_+(\mathcal{M})}\tau(\rho^{1/2}\sigma^{-1}\rho^{1/2}\,\omega) = \left\|\rho^{1/2}\sigma^{-1}\rho^{1/2}\right\|_\infty
\end{aligned}$$

is related to the max-relative entropy $D_\infty(\rho\|\sigma) = \log\inf\{\lambda|\rho\leqslant\lambda\sigma\}$ [Dat09]. Since ε is arbitrary and the upper bound is symmetric in t, we arrive at the statement of the lemma. □

4.3.1. Recoverability for sandwiched Rényi relative (quasi-)entropy. We now turn to the sandwiched Rényi relative (quasi-)entropy and identify physically meaningful refinements of its data-processing inequality. Let $\alpha\in[1/2,1)\cup(1,\infty]$ and set $\alpha' := \alpha/(\alpha-1)$, so that $\frac{1}{\alpha}+\frac{1}{\alpha'}=1$. The sandwiched Rényi relative entropy is given by

$$\widetilde{D}_\alpha(\rho\,\|\sigma) = \alpha'\,\log\left\|\rho^{1/2}\sigma^{-\frac{1}{\alpha'}}\rho^{1/2}\right\|_\alpha .$$

Theorem 4.20. *Let* $\alpha \in (1/2, 1) \cup (1, \infty)$, $\alpha' = \frac{\alpha}{\alpha-1}$, *and* $\varepsilon \in \left(0, \frac{1-1/|\alpha'|}{2}\right)$. *Let* $\widetilde{Q}_\alpha(\rho\|\sigma) := \left\|\rho^{1/2}\sigma^{-\frac{1}{\alpha'}}\rho^{1/2}\right\|_\alpha$ *denote the sandwiched* α*-Rényi relative quasi-entropy. Let* $\widetilde{Q}_\infty(\rho\|\sigma) := \left\|\rho^{1/2}\sigma^{-1}\rho^{1/2}\right\|_\infty$. *Then*

$$|\widetilde{Q}_\alpha(\rho\,\|\sigma) - \widetilde{Q}_\alpha(\rho_\mathcal{N}\|\sigma_\mathcal{N})| \geqslant \left(K(\alpha, \widetilde{Q}_\infty(\rho\|\sigma), \varepsilon)\frac{\pi}{2}\|\rho - R_\sigma(\rho_\mathcal{N})\|_1\right)^{\frac{1}{\frac{1-1/|\alpha'|}{2}-\varepsilon}}, \tag{72}$$

$$|\widetilde{Q}_\alpha(\rho\|\sigma) - \widetilde{Q}_\alpha(\rho_\mathcal{N}\|\sigma_\mathcal{N})| \geqslant \left(K(\alpha, \widetilde{Q}_\infty(\rho\|\sigma), \varepsilon)\frac{\pi}{2\cosh \pi t}\|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1\right)^{\frac{1}{\frac{1-1/|\alpha'|}{2}-\varepsilon}}, \tag{73}$$

$$|\widetilde{Q}_\alpha(\rho\|\sigma) - \widetilde{Q}_\alpha(\rho_\mathcal{N}\|\sigma_\mathcal{N})| \geqslant \left(K(\alpha, \widetilde{Q}_\infty(\rho\|\sigma), \varepsilon)\,\frac{1}{2}\|\rho - R^u_\sigma(\rho_\mathcal{N})\|_1\right)^{\frac{1}{\frac{1-1/|\alpha'|}{2}-\varepsilon}}, \tag{74}$$

where the constant

$$K(\widetilde{Q}_\infty(\rho\,\|\sigma), \alpha, \varepsilon) := \left(4\widetilde{Q}_\infty(\rho\|\sigma)^{1/2} + \left(\frac{\pi}{e\varepsilon\,\sin\left(\pi|\frac{1}{\alpha'}|\right)}\right)^{1/2} + 4\right)^{-1}. \tag{75}$$

Proof. For $1 < \alpha, \alpha' < \infty$, the function $x^{-\frac{1}{\alpha'}}$ is operator convex and operator anti-monotone. We have

$$\begin{aligned}\widetilde{Q}_{x^{-\frac{1}{\alpha'}}}(\rho\|\sigma) &= \sup_{\omega\in D_+(\mathcal{M})} \left\langle \rho^{1/2}\middle|\,\Delta(\sigma,\omega)^{-1/\alpha'}\,\middle|\rho^{1/2}\right\rangle \\ &= \sup_{\omega\in D_+(\mathcal{M})} \tau(\rho^{1/2}\sigma^{-1/\alpha'}\rho^{1/2}\omega^{1/\alpha'}) = \left\|\rho^{1/2}\sigma^{-\frac{1}{\alpha'}}\rho^{1/2}\right\|_\alpha.\end{aligned}$$

Thus, for $1 < \alpha \leqslant \infty$,

$$\widetilde{D}_\alpha(\rho\,\|\sigma) = \alpha'\,\log \widetilde{Q}_{x^{-1/\alpha'}}(\rho\|\sigma)\,.$$

Writing $0 < \beta := 1/\alpha' < 1$, the integral representation of $x^{-\beta}$ is as follows:

$$x^{-\beta} = \frac{\sin(\pi\beta)}{\pi}\int_0^\infty \lambda^{-\beta}\frac{1}{\lambda + x}\,\mathrm{d}\lambda\,.$$

The constant $c(S,T) \leqslant \frac{\pi}{\sin(\pi\beta)}T^\beta$. Then by lemma 4.19, we have

$$\begin{aligned}\|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi}&\left(4S^{1/2}\widetilde{Q}_{x^{-1}}(\rho\|\sigma)^{1/2} + \sqrt{\frac{\pi}{\sin(\pi\beta)}T^\beta\,\ln(T/S)}\right.\\ &\left.\times(\widetilde{Q}_{x^{-\beta}}(\rho\|\sigma) - \widetilde{Q}_{x^{-\beta}}(\rho_\mathcal{N}\|\sigma_\mathcal{N}))^{1/2} + 4T^{-1/2}\right)\,.\end{aligned}$$

Choose $S = T^{-1} = \delta$ and $\delta := \min\{|\widetilde{Q}_{x^{-\beta}}(\rho\|\sigma) - \widetilde{Q}_{x^{-\beta}}(\rho_\mathcal{N}\|\sigma_\mathcal{N})|, 1\}$. Thus

$$\begin{aligned}\|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1 &\leqslant \frac{2\cosh(\pi t)}{\pi}\left(4\delta^{1/2}\widetilde{Q}_{x^{-1}}(\rho\|\sigma)^{1/2} + \sqrt{\frac{2\pi}{\sin(\pi\beta)}|\ln\,\delta|}\delta^{\frac{1-\beta}{2}} + 4\delta^{1/2}\right) \\ &\leqslant \frac{2\cosh(\pi t)}{\pi}\left(4\widetilde{Q}_{x^{-1}}(\rho\|\sigma)^{1/2} + \sqrt{\frac{\pi}{e\varepsilon\,\sin(\pi\beta)}} + 4\right)\delta^{\frac{1-\beta}{2}-\varepsilon}.\end{aligned}$$

The reasoning for these steps is similar to that given for (68)–(70).

For $1/2 \leqslant \alpha < 1$, which implies that $\alpha' \leqslant -1$, the function $x^{-\frac{1}{\alpha'}}$ is operator monotone and operator concave because $-\frac{1}{\alpha'} \in (0, 1)$. We have

$$\widetilde{Q}_{-x^{-\frac{1}{\alpha'}}}(\rho\|\sigma) = \sup_{\omega \in D_+(\mathcal{M})} -\left\langle \rho^{1/2} \middle| \Delta(\sigma, \omega)^{-1/\alpha'} \middle| \rho^{1/2} \right\rangle$$
$$= -\inf_{\omega \in D_+(\mathcal{M})} \tau(\rho^{1/2}\sigma^{-1/\alpha'}\rho^{1/2}\omega^{1/\alpha'}) = -\left\| \rho^{1/2}\sigma^{-1/\alpha'}\rho^{1/2} \right\|_\alpha .$$

Then for $1/2 \leqslant \alpha < 1$,

$$\widetilde{D}_\alpha(\rho\, \|\sigma) = \alpha' \log\left(-\widetilde{Q}_{-x^{-1/\alpha'}}(\rho\|\sigma) \right) .$$

Let $\gamma := -1/\alpha'$. For $0 < \gamma < 1$, the integral representation is

$$x^\gamma = \frac{\sin(\pi\gamma)}{\pi} \int_0^\infty \lambda^\gamma \left(\frac{1}{\lambda} - \frac{1}{\lambda + x} \right) \mathrm{d}\lambda .$$

Then the constant $c(S, T) \leqslant \frac{\pi}{\sin(\pi\gamma)} S^{-\gamma}$. By lemma 4.19, we have

$$\|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi} \left(4S^{1/2}\widetilde{Q}_{x^{-1}}(\rho\|\sigma)^{1/2} + \sqrt{\frac{\pi}{\sin(\pi\gamma)} S^{-\gamma} \ln(T/S)} |\widetilde{Q}_{-x^\gamma}(\rho\|\sigma) \right.$$
$$\left. - \widetilde{Q}_{-x^\gamma}(\rho_\mathcal{N}\|\sigma_\mathcal{N})|^{1/2} + 4T^{-1/2} \right) .$$

Set $S = T^{-1} = \delta$ and $\delta := \min\{|\widetilde{Q}_{-x^\gamma}(\rho\|\sigma) - \widetilde{Q}_{-x^\gamma}(\rho_\mathcal{N}\|\sigma_\mathcal{N})|, 1\}$. Then

$$\|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1 \leqslant \frac{2\cosh(\pi t)}{\pi} \left(4\delta^{1/2}\widetilde{Q}_{x^{-1}}(\rho\|\sigma)^{1/2} + \sqrt{\frac{2\pi}{\sin(\pi\gamma)} |\ln\delta|}\, \delta^{\frac{1-\gamma}{2}} + 4\delta^{1/2} \right)$$
$$\leqslant \frac{2\cosh(\pi t)}{\pi} \left(4\widetilde{Q}_{x^{-1}}(\rho\|\sigma)^{1/2} + \sqrt{\frac{\pi}{e\varepsilon\, \sin(\pi\gamma)}} + 4 \right) \delta^{\frac{1-\gamma}{2} - \varepsilon}.$$

The reasoning for these steps is similar to that given for (68)–(70). □

We then find the following for the sandwiched Rényi relative entropy:

Corollary 4.21. *Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra, and let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. Let ρ and σ be two faithful states. Let $\alpha \in (1/2, 1) \cup (1, \infty)$ and $\alpha' = \alpha/(\alpha - 1)$, so that $1/\alpha + 1/\alpha' = 1$. Set $t \in \mathbb{R}$ and $\varepsilon \in (0, \frac{1-1/|\alpha'|}{2})$. For $\alpha \in (1/2, 1)$, the following inequality holds*

$$\widetilde{D}_\alpha(\rho\|\sigma) - \widetilde{D}_\alpha(\rho_\mathcal{N}\|\sigma_\mathcal{N})$$
$$\geqslant |\alpha'| \log\left(1 + \left(K(\alpha, \varepsilon, \widetilde{Q}_\infty(\rho\|\sigma)) \frac{\pi}{2\cosh \pi t} \|\rho - R^t_\sigma(\rho_\mathcal{N})\|_1 \right)^{\frac{1}{\frac{1-1/|\alpha'|}{2} - \varepsilon}} \right),$$

and for $\alpha > 1$, *the following inequality holds*

$$\widetilde{D}_\alpha(\rho\,\|\sigma) - \widetilde{D}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) \geqslant \alpha' \log\left(1 + \frac{1}{\widetilde{Q}_\infty(\rho\|\sigma)^{\frac{1}{\alpha'}}}\left(K(\alpha,\varepsilon,\widetilde{Q}_\infty(\rho\|\sigma))\right.\right.$$
$$\left.\left.\times \frac{\pi}{2\cosh \pi t}\|\rho - R^t_\sigma(\rho_{\mathcal{N}})\|_1\right)^{\frac{\frac{1}{1-1/|\alpha'|}}{2}-\varepsilon}\right),$$

where the constant $K(\alpha,\varepsilon,\widetilde{Q}_\infty(\rho\|\sigma))$ *is given in* (75).

Proof. For $1/2 < \alpha < 1$ and $\alpha' \leqslant -1$, we find that

$$\begin{aligned}
\widetilde{D}_\alpha(\rho\|\sigma) - \widetilde{D}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) &= |\alpha'| \log \frac{\widetilde{Q}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})}{\widetilde{Q}_\alpha(\rho\|\sigma)} \\
&= |\alpha'| \log\left(1 + \frac{\widetilde{Q}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) - \widetilde{Q}_\alpha(\rho\|\sigma)}{\widetilde{Q}_\alpha(\rho\|\sigma)}\right) \\
&\geqslant |\alpha'| \log\left(1 + \frac{1}{\widetilde{Q}_\alpha(\rho\|\sigma)}\left(K(\alpha,\varepsilon,\widetilde{Q}_\infty(\rho\|\sigma))\right.\right. \\
&\quad \left.\left.\times \frac{\pi}{2\cosh \pi t}\|\rho - R^t_\sigma(\rho_{\mathcal{N}})\|_1\right)^{\frac{\frac{1}{1-1/|\alpha'|}}{2}-\varepsilon}\right), \\
&\geqslant |\alpha'| \log\left(1 + \left(K(\alpha,\varepsilon,\widetilde{Q}_\infty(\rho\|\sigma))\right.\right. \\
&\quad \left.\left.\times \frac{\pi}{2\cosh \pi t}\|\rho - R^t_\sigma(\rho_{\mathcal{N}})\|_1\right)^{\frac{\frac{1}{1-1/|\alpha'|}}{2}-\varepsilon}\right).
\end{aligned}$$

The first inequality follows from (73), and the second follows because $\widetilde{Q}_\alpha(\rho\|\sigma) \leqslant 1$ for $\alpha \in (1/2, 1)$.

For $\alpha > 1$, consider that

$$\begin{aligned}
\widetilde{D}_\alpha(\rho\|\sigma) - \widetilde{D}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) &= \alpha' \log \frac{\widetilde{Q}_\alpha(\rho\|\sigma)}{\widetilde{Q}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})} \\
&= \alpha' \log\left(1 + \frac{\widetilde{Q}_\alpha(\rho\|\sigma) - \widetilde{Q}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})}{\widetilde{Q}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})}\right) \\
&\geqslant \alpha' \log\left(1 + \frac{1}{\widetilde{Q}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})}\left(K(\alpha,\varepsilon,\widetilde{Q}_\infty(\rho\|\sigma))\right.\right. \\
&\quad \left.\left.\times \frac{\pi}{2\cosh \pi t}\|\rho - R^t_\sigma(\rho_{\mathcal{N}})\|_1\right)^{\frac{\frac{1}{1-1/|\alpha'|}}{2}-\varepsilon}\right) \\
&\geqslant \alpha' \log\left(1 + \frac{1}{\widetilde{Q}_\infty(\rho\|\sigma)^{\frac{1}{\alpha'}}}\left(K(\alpha,\varepsilon,\widetilde{Q}_\infty(\rho\|\sigma))\right.\right.
\end{aligned}$$

$$\times \frac{\pi}{2\cosh \pi t}\|\rho - R_\sigma^t(\rho_{\mathcal{N}})\|_1\Big)^{\frac{\frac{1}{1-1/|\alpha'|}}{2}-\varepsilon}\Bigg).$$

The first inequality follows from (73). The second inequality follows from the inequalities $\widetilde{Q}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}}) \leqslant \widetilde{Q}_\alpha(\rho\|\sigma) \leqslant \widetilde{Q}_\infty(\rho\|\sigma)^{\frac{1}{\alpha'}}$. The first is a consequence of the data-processing inequality and the second a consequence of the monotonicity of the sandwiched Rényi relative entropies with respect to α (for the latter, see [MLDS+13, theorem 7] and [BST18, lemma 8]). □

Remark 4.22. For $\alpha = 1$, $\widetilde{D}_\alpha$ coincides with the standard relative entropy D, for which results are given in theorem 4.13. For the two boundary cases $\alpha = 1/2$ and $\alpha = \infty$, the recoverability result in corollary 4.21 does not hold. The $\alpha = 1/2$ case corresponds to the root fidelity

$$\sqrt{F}(\rho,\sigma) = \left\|\rho^{1/2}\sigma\rho^{1/2}\right\|_{1/2} = -\widetilde{Q}_{-x}(\rho\|\sigma),$$

and $\alpha = \infty$ to

$$\widetilde{Q}_\infty(\rho\ \|\sigma) = \widetilde{Q}_{x^{-1}}(\rho\|\sigma) = \inf\{\lambda \mid \rho \leqslant \lambda\sigma\}\ .$$

Our method fails for these two cases because both operator anti-monotone functions $f(x) = -x$ and $g(x) = x^{-1}$ have trivial measure $d\nu$ in their integral representations. Indeed, for both cases, it was already observed in [HM17, remarks 5.15 and 5.16] that there are examples for which the data-processing inequality for fidelity is saturated, i.e. $F(\rho,\sigma) = F(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})$ (resp. $\widetilde{Q}_\infty(\rho\|\sigma) = \widetilde{Q}_\infty(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})$), but it is not for the relative entropy $D(\rho\|\sigma) > D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})$, which implies that the existence of any exact recovery map is impossible. This extends the results in [Jen17a].

The following are reversibility results that are consequences of recoverability estimates. Note that the faithfulness assumption can be weakened to $s(\rho) \leqslant s(\sigma)$, as in corollary 5.16 for the von Neumann algebra case.

Corollary 4.23. *Let ρ and σ be faithful quantum states. The following are equivalent:*

(*a*) $D(\rho\|\sigma) = D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})$.
(*b*) $\widetilde{D}_\alpha(\rho\|\sigma) = \widetilde{D}_\alpha(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})$ *for some* $\alpha \in (1/2, 1) \cup (1,\infty)$ *where* $\widetilde{D}_\alpha$ *is the* α*-sandwiched Rényi relative entropy.*
(*c*) $\widetilde{Q}_f(\rho\|\sigma) = \widetilde{Q}_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})$ *for some regular operator anti-monotone function* f.
(*d*) $\widetilde{Q}_f(\rho\|\sigma) = \widetilde{Q}_f(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})$ *for all operator anti-monotone functions* f.
(*e*) $R^t_\rho(\sigma_{\mathcal{N}}) = \sigma$ *for all* $t \in \mathbb{R}$.
(*f*) $R^t_\sigma(\rho_{\mathcal{N}}) = \rho$ *for all* $t \in \mathbb{R}$.
(*g*) *There exists some CPTP map* $\Phi : L_1(\mathcal{N}) \to L_1(\mathcal{M})$ *such that* $\Phi(\rho_{\mathcal{N}}) = \rho$ *and* $\Phi(\sigma_{\mathcal{N}}) = \sigma$.

Proof. The implications (e) ⇒ (g) and (f) ⇒ (g) are trivial. (g) ⇒ (a)–(d) follows from the data-processing inequality. Note that for faithful ρ and σ, $Q_{x^2}(\rho\|\sigma), Q_{x^{-1}}(\rho\|\sigma), \widetilde{Q}_\infty(\rho\|\sigma) < \infty$ are finite.

(a) ⇒ (e) follows from theorem 4.5. (a) ⇒ (f) uses theorem 4.13. (b) ⇒ (f) uses corollary 4.21. (c) ⇒ (f) follows from theorem 4.20. (d) ⇒ (c) is trivial. □

Remark 4.24. It follows from [Pet86a] and [JRS+18] that the same equivalences hold for the standard f-divergence Q_f and the Petz–Rényi relative entropy D_α. Corollary 4.23 above shows that the preservation of a 'regular' optimized f-divergence is also equivalent to the existence of a recovery map.

5. Optimized *f*-divergence in von Neumann algebras

5.1. Definition of optimized *f*-divergence

In this section, we define the optimized f-divergence for states of a general von Neumann algebra. We also prove the data-processing inequality for the optimized f-divergence. We refer to appendix A for a review of the basics of von Neumann algebras and the notations used in this section. We first define the optimized f-divergence for two states ρ and σ satisfying the support assumption $s(\rho) \leqslant s(\sigma)$.

Definition 5.1. Let $\mathcal{M} \subset B(H)$ be a von Neumann algebra. Let ρ, σ be two normal states such that $s(\rho) \leqslant s(\sigma)$, and let $\boldsymbol{\rho}, \boldsymbol{\sigma} \in H$ be their corresponding vector representations. For an operator anti-monotone function $f : (0, \infty) \to \mathbb{R}$, we define the optimized f-divergence $\widetilde{Q}_f(\rho\|\sigma)$ as follows:

$$\widetilde{Q}_f(\rho\|\sigma) = \sup_{\boldsymbol{\omega}\,:\,\|\boldsymbol{\omega}\|_2=1,\ \boldsymbol{\rho}\in[\mathcal{M}\boldsymbol{\omega}]} \langle\boldsymbol{\rho}|\, f(\Delta(\boldsymbol{\sigma},\boldsymbol{\omega}))\,|\boldsymbol{\rho}\rangle\,, \tag{76}$$

where the supremum runs over all unit vectors $\boldsymbol{\omega} \in H$ such that $\boldsymbol{\rho} \in [\mathcal{M}\boldsymbol{\omega}]$ and $\Delta(\boldsymbol{\sigma}, \boldsymbol{\omega})$ is the relative modular operator. This definition of $\widetilde{Q}_f$ only depends on the states ρ and σ, and is independent of the choice of the algebra representation $\mathcal{M} \subset B(H)$ and the vector representations $|\boldsymbol{\rho}\rangle$ and $|\boldsymbol{\sigma}\rangle$ for ρ and σ respectively.

If f is a continuous function on $[0, \infty)$, we do not need the restriction $\boldsymbol{\rho} \in [\mathcal{M}\boldsymbol{\omega}]$ and can take the supremum over all $\boldsymbol{\omega}$ satisfying $\|\boldsymbol{\omega}\|_2 = 1$. Otherwise, we have to require $\boldsymbol{\rho} \in [\mathcal{M}'\boldsymbol{\sigma}]$ and $\boldsymbol{\rho} \in [\mathcal{M}\boldsymbol{\omega}]$, since $\Delta(\boldsymbol{\sigma}, \boldsymbol{\omega})$ is supported on $s_{\mathcal{M}}(\sigma)s_{\mathcal{M}'}(\omega')$, where ω' is the state of the commutant $\mathcal{M}'$ implemented by the vector $\boldsymbol{\omega}$ and $s_{\mathcal{M}}(\sigma)$ (resp. $s_{\mathcal{M}'}(\omega')$) is the support projection of σ (resp. ω') on $\mathcal{M}$ (resp. $\mathcal{M}'$). The relative modular operator connects to the spatial derivative as follows:

$$\Delta(\boldsymbol{\sigma}, \boldsymbol{\omega}) = \Delta(\sigma/\boldsymbol{\omega}) = \Delta(\sigma/\omega'),$$

where $\omega' \in \mathcal{M}'_*$ is the state on $\mathcal{M}'$ implemented by the vector $\boldsymbol{\omega} \in H$. Note that $\Delta(\sigma/\omega')$ and $\Delta(\omega'/\sigma)$ have the same support and

$$\Delta(\sigma/\omega') = \Delta(\omega'/\sigma)^{-1}$$

on their support. Then we have the following equivalent definition for the optimized f-divergence:

$$\widetilde{Q}_f(\rho\|\sigma) = \sup_{\boldsymbol{\omega}\,:\,\|\boldsymbol{\omega}\|_2=1,\boldsymbol{\rho}\in[\mathcal{M}\boldsymbol{\omega}]} \langle\boldsymbol{\rho}|\, \widetilde{f}(\Delta(\omega'/\sigma))\,|\boldsymbol{\rho}\rangle\,, \tag{77}$$

where $\widetilde{f}(x) = f(x^{-1})$ is operator monotone. This latter definition via the spatial derivative is closer to the definition of the sandwiched Rényi relative entropy from [BST18], which used Araki–Masuda L_p spaces [AM82].

We now verify that the definition of $\widetilde{Q}_f$ in (76) is independent of vector representations. Note that the representation π in the following need not be faithful.

Proposition 5.2. *Let $\pi : \mathcal{M} \to B(H_1)$ be a $*$-representation, and let $\boldsymbol{\rho}_1, \boldsymbol{\sigma}_1 \in H_1$ be the unit vectors implementing ρ and σ, respectively, via π. Then*

$$\widetilde{Q}_f(\rho\|\sigma) = \sup_{\boldsymbol{\omega}_1 \,:\, \|\boldsymbol{\omega}_1\|_2 = 1, \boldsymbol{\rho}_1 \in [\pi(\mathcal{M})\boldsymbol{\omega}_1]} \langle \boldsymbol{\rho}_1 |\, f(\Delta(\boldsymbol{\sigma}_1, \boldsymbol{\omega}_1))\, | \boldsymbol{\rho}_1 \rangle \,.$$

Proof. We follow the idea of [BST18, lemma 3] and use the equivalent definition from (77) with the spatial derivative. Consider that

$$\widetilde{Q}_f(\rho\|\sigma) = \sup_{\boldsymbol{\omega} \,:\, \|\boldsymbol{\omega}\|_2 = 1, \boldsymbol{\rho} \in [\mathcal{M}\boldsymbol{\omega}]} \langle \boldsymbol{\rho} |\, \widetilde{f}(\Delta(\boldsymbol{\omega}/\sigma))\, | \boldsymbol{\rho} \rangle \,.$$

Define $V_\rho : H \to H_1$ as the partial isometry such that, for $\boldsymbol{\eta} \in [\mathcal{M}\rho]^\perp$,

$$V(a\boldsymbol{\rho} + \boldsymbol{\eta}) = \pi(a)\boldsymbol{\rho}_1 \,, \qquad a \in \mathcal{M}.$$

Since $\pi(a)V_\rho = V_\rho a$, we have $R_\sigma(V_\rho\boldsymbol{\rho}) = V_\rho R_\sigma(\boldsymbol{\rho})$ (see (91) for the definition of operator $R_\sigma(\boldsymbol{\rho})$). Let $V \equiv V_\rho$. Then for all $\boldsymbol{\xi} \in [\mathcal{M}\boldsymbol{\omega}]s(\sigma)H$ and $\boldsymbol{\omega}_1 \in H_1$, we find that

$$\begin{aligned}
\langle \boldsymbol{\xi} |\, V^*\Delta(\boldsymbol{\omega}_1/\sigma)V\, | \boldsymbol{\xi} \rangle &= \langle \boldsymbol{\omega}_1 |\, R_\sigma(V\boldsymbol{\xi})R_\sigma(V\boldsymbol{\xi})^*\, | \boldsymbol{\omega}_1 \rangle \\
&= \langle \boldsymbol{\omega}_1 |\, VR_\sigma(\boldsymbol{\xi})R_\sigma(\boldsymbol{\xi})^*V^*\, | \boldsymbol{\omega}_1 \rangle \\
&= \langle V^*\boldsymbol{\omega}_1 |\, R_\sigma(\boldsymbol{\xi})R_\sigma(\boldsymbol{\xi})^*\, | V^*\boldsymbol{\omega}_1 \rangle \\
&= \langle \boldsymbol{\xi} |\, \Delta(V^*\boldsymbol{\omega}_1/\sigma)\, | \boldsymbol{\xi} \rangle \,.
\end{aligned}$$

Moreover $s'(V^*\boldsymbol{\omega}_1) = [\mathcal{M}V^*\boldsymbol{\omega}_1] = [V^*\pi(\mathcal{M})\boldsymbol{\omega}_1] = V^*s'(\boldsymbol{\omega}_1)V$ and hence

$$V^*\Delta(\boldsymbol{\omega}_1/\sigma)V = \Delta(V^*\boldsymbol{\omega}_1/\sigma),$$

with the same support for all $\boldsymbol{\omega}_1 \in H_1$ with $\rho \in [\pi(\mathcal{M})\boldsymbol{\omega}_1]$. Since $\widetilde{f}$ is operator concave and operator monotone

$$\begin{aligned}
\langle \boldsymbol{\rho}_1 |\, \widetilde{f}(\Delta(\boldsymbol{\omega}_1/\sigma))\, | \boldsymbol{\rho}_1 \rangle &= \langle \boldsymbol{\rho} |\, V^*\widetilde{f}(\Delta(\boldsymbol{\omega}_1/\sigma))V\, | \boldsymbol{\rho} \rangle \\
&\leqslant \langle \boldsymbol{\rho} |\, \widetilde{f}(\Delta(V^*\boldsymbol{\omega}_1/\sigma))\, | \boldsymbol{\rho} \rangle \leqslant \langle \boldsymbol{\rho} |\, \widetilde{f}(\Delta(\overline{V^*\boldsymbol{\omega}_1}/\sigma))\, | \boldsymbol{\rho} \rangle \;,
\end{aligned}$$

where $\overline{V^*\boldsymbol{\omega}_1}$ is the normalization of $V^*\boldsymbol{\omega}_1$. Here we view V as an isometry by restricting on the support $V^*V = [\pi_1(\mathcal{M})\boldsymbol{\rho}_1]$. Therefore

$$\sup_{\boldsymbol{\omega}_1 \,:\, \|\boldsymbol{\omega}_1\|_2 = 1, \boldsymbol{\rho}_1 \in [\pi(\mathcal{M})\boldsymbol{\omega}_1]} \langle \boldsymbol{\rho}_1 |\, \widetilde{f}(\Delta(\boldsymbol{\omega}_1/\sigma))\, | \boldsymbol{\rho}_1 \rangle \leqslant \sup_{\boldsymbol{\omega} \,:\, \|\boldsymbol{\omega}\|_2 = 1, \boldsymbol{\rho} \in [\mathcal{M}\boldsymbol{\omega}]} \langle \boldsymbol{\rho} |\, \widetilde{f}(\Delta(\boldsymbol{\omega}/\sigma))\, | \boldsymbol{\rho} \rangle \;.$$

The converse direction follows by the symmetric role of the representations $\pi_1(\mathcal{M}) \subset B(H_1)$ and $\mathcal{M} \subset B(H)$. □

By the independence above, we can then carry the definition to the standard form $(\mathcal{M}, L_2(\mathcal{M}), J, L_2(\mathcal{M})^+)$ using Haagerup L_2-spaces. Let $h_\rho \in L_1(\mathcal{M})$ be the density operator corresponding to ρ. Then

$$\widetilde{Q}_f(\rho\|\sigma) = \sup_{\omega} \left\langle h_\rho^{1/2} \middle|\, f\,(\Delta(\sigma, \omega)\, \middle| h_\rho^{1/2} \right\rangle,$$

where the supremum runs over all states ω such that $s(\omega) \geqslant s(\rho)$. The next proposition shows that the definition above coincides with the finite-dimensional definition in [Wil18a], and one can further restrict to $\omega \gg \rho$; i.e. there exists $\lambda > 0$ such that $\rho \leqslant \lambda\omega$.

Proposition 5.3. *Let $f : (0, \infty) \to \mathbb{R}$ be an operator anti-monotone and ν be the measure in the integral representation of f as in (5). Suppose ν does not contain a point mass at $\lambda = 0$. Then*

$$\widetilde{Q}_f(\rho\|\sigma) = \sup_{\omega\in D(\mathcal{M})} \lim_{\varepsilon\to 0^+} \left\langle h_\rho^{1/2} \right| f(\Delta(\sigma, \omega + \varepsilon\phi)) \left| h_\rho^{1/2} \right\rangle \tag{78}$$

$$= \sup_{\omega\in D(\mathcal{M}),\, \omega\gg\rho} \left\langle h_\rho^{1/2} \right| f(\Delta(\sigma, \omega)) \left| h_\rho^{1/2} \right\rangle, \tag{79}$$

where in (78), ϕ can be any normal state with $s(\rho) \leqslant s(\phi)$.

Proof. For the first expression, we note that $\Delta(\sigma, \omega + \varepsilon\phi)^{1/2} \to \Delta(\sigma, \omega)^{1/2}$ strongly in the resolvent sense by [OP04, proposition 4.9]. This implies (by the integral representation of f) that

$$\lim_{\varepsilon\to 0^+} \langle h_\rho^{1/2} | f(\Delta(\sigma, \omega + \varepsilon\phi)) | h_\rho^{1/2} \rangle = \langle h_\rho^{1/2} | f(\Delta(\sigma, \omega)) | h_\rho^{1/2} \rangle \ .$$

For the second expression, we can choose $\omega_\varepsilon = \varepsilon\rho + (1 - \varepsilon)\omega$. By the same reasoning,

$$\lim_{\varepsilon\to 0^+} \langle h_\rho^{1/2} | f(\Delta(\sigma, \omega_\varepsilon)) | h_\rho^{1/2} \rangle = \langle h_\rho^{1/2} | f(\Delta(\sigma, \omega)) | h_\rho^{1/2} \rangle \ .$$

Note that $\rho \leqslant \varepsilon^{-1}\omega_\varepsilon$. Then we have

$$\sup_{\omega\in D(\mathcal{M}),\omega\gg\rho} \left\langle h_\rho^{1/2} \right| f(\Delta(\sigma, \omega)) \left| h_\rho^{1/2} \right\rangle \geqslant \widetilde{Q}_f(\rho\|\sigma) \ .$$

The inverse inequality is obvious. □

Following the same idea above, the optimized divergence for general two states ρ and σ can be defined as follows

$$\widetilde{Q}_f(\rho\|\sigma) := \lim_{\varepsilon\to 0^+} \widetilde{Q}_f(\rho\|\sigma + \varepsilon\rho) = \sup_{\varepsilon>0} \widetilde{Q}_f(\rho\|\sigma + \varepsilon\rho) \ . \tag{80}$$

The above limit is increasing as $\varepsilon \to 0^+$ for all ω because $\Delta(\sigma + \varepsilon\rho, \omega) = \Delta(\sigma, \omega) + \varepsilon\Delta(\rho, \omega)$ and f is operator anti-monotone. For ρ and σ with $s(\rho) \leqslant s(\sigma)$, this recovers the definition 5.1

$$\begin{aligned}\lim_{\varepsilon\to 0^+} \widetilde{Q}_f(\rho\|\sigma + \varepsilon\rho) &= \sup_{\varepsilon>0} \sup_{\omega\in D(\mathcal{M}),\, \omega\gg\rho} \left\langle h_\rho^{1/2} \right| f(\Delta(\sigma + \varepsilon\rho, \omega)) \left| h_\rho^{1/2} \right\rangle \\ &= \sup_{\omega\in D(\mathcal{M}),\, \omega\gg\rho} \sup_{\varepsilon>0} \left\langle h_\rho^{1/2} \right| f(\Delta(\sigma + \varepsilon\rho, \omega)) \left| h_\rho^{1/2} \right\rangle \\ &= \widetilde{Q}_f(\rho\|\sigma).\end{aligned} \tag{81}$$

For the last step above, by [OP04, proposition 4.9] we have that for each $\omega \gg \rho$ and $t > 0$,

$$\left\langle h_\rho^{1/2} \right| (\Delta(\sigma + \varepsilon\rho, \omega) + t)^{-1} \left| h_\rho^{1/2} \right\rangle \to \left\langle h_\rho^{1/2} \right| (\Delta(\sigma, \omega) + t)^{-1} \left| h_\rho^{1/2} \right\rangle .$$

Using integral representation (5) of f and monotone convergence theorem over $\varepsilon \to 0^+$, we have

$$\left\langle h_\rho^{1/2}\right| f(\Delta(\sigma+\varepsilon\rho,\omega)) \left|h_\rho^{1/2}\right\rangle \to \left\langle h_\rho^{1/2}\right| f(\Delta(\sigma,\omega)) \left|h_\rho^{1/2}\right\rangle,$$

which verifies (81).

As the optimized f-divergence for general ρ and σ is defined through approximation, for most of the following discussion it suffices to consider $\widetilde{Q}_f(\rho\|\sigma)$ with support assumption.

5.2. Comparison to standard f-divergence

In this section, we first review the definition of f-divergence introduced by Petz in [Pet85, Pet86a], which we call the standard f-divergence. Let $\mathcal{M} \subset B(H)$ be a von Neumann algebra, and let ρ, σ be two normal states implemented by $\boldsymbol{\rho}, \boldsymbol{\sigma} \in H$, respectively. For an operator convex function $f : (0, \infty) \to \mathbb{R}$, the standard f-divergence is defined as follows:

$$Q_f(\rho\|\sigma) := \langle \boldsymbol{\rho}| \, f(\Delta(\boldsymbol{\sigma}, \boldsymbol{\rho})) \, |\boldsymbol{\rho}\rangle \ , \qquad \text{if} \ \ s(\rho) \leqslant s(\sigma)$$

which is also independent of the particular vector representation, as in proposition 5.2. Because the standard f-divergence Q_f for general ρ and σ also admits approximation as in (80) (see [Hia18]), it is clear from definitions that

$$\widetilde{Q}_f(\rho\|\sigma) \geqslant Q_f(\rho\|\sigma) \ .$$

Example 5.4. The sandwiched Rényi relative entropy was defined in [BST18] as $\widetilde{D}_\alpha(\rho\ \|\sigma) := \alpha' \log \widetilde{Q}_\alpha(\rho\|\sigma)$, where $\alpha' := \alpha/(\alpha-1)$ and

$$\widetilde{Q}_\alpha(\rho\|\sigma) := \begin{cases} \sup\limits_{\omega\,:\,\|\omega\|=1} \left\| \Delta(\boldsymbol{\omega}/\sigma)^{\frac{1}{2\alpha'}} \, |\boldsymbol{\rho}\rangle \right\|_H^2 & \text{if } 1 < \alpha \leqslant \infty \\ \inf\limits_{\omega\,:\,\|\omega\|=1\ ,\boldsymbol{\rho}\in[\mathcal{M}\boldsymbol{\rho}]} \left\| \Delta(\boldsymbol{\omega}/\sigma)^{\frac{1}{2\alpha'}} \, |\boldsymbol{\rho}\rangle \right\|_H^2 & \text{if } \dfrac{1}{2} \leqslant \alpha < 1. \end{cases}$$

Note that

$$\left\| \Delta(\boldsymbol{\omega}/\sigma)^{\frac{1}{2\alpha'}} \, |\boldsymbol{\rho}\rangle \right\|_H^2 = \langle \boldsymbol{\rho}|\Delta(\boldsymbol{\omega}/\sigma)^{\frac{1}{\alpha'}}|\boldsymbol{\rho}\rangle = \langle \boldsymbol{\rho}|\Delta(\boldsymbol{\sigma},\boldsymbol{\omega})^{-\frac{1}{\alpha'}}|\boldsymbol{\rho}\rangle \ .$$

Thus we have

$$\widetilde{Q}_\alpha(\rho\|\sigma) = \begin{cases} \widetilde{Q}_{x^{-\frac{1}{\alpha'}}}(\rho\|\sigma) & \text{if } 1 < \alpha \leqslant \infty \\ -\widetilde{Q}_{-x^{-\frac{1}{\alpha'}}}(\rho\|\sigma) & \text{if } \dfrac{1}{2} \leqslant \alpha < 1 \end{cases} .$$

Example 5.5. For $f(x) = -\log x$, it was shown in [Wil18a], by invoking the Klein inequality, that for $\mathcal{M} = B(H)$, the following equality holds

$$\widetilde{Q}_{-\log\ x}(\rho\|\sigma) = D(\rho\|\sigma) \ .$$

For the general case, we immediately have that

$$\widetilde{Q}_{-\log\ x}(\rho\ \|\sigma) \geqslant Q_{-\log\ x}(\rho\|\sigma) = D(\rho\|\sigma).$$

On the other hand, since $t \mapsto \alpha' \log t$ is concave for $\alpha > 1$ (and hence $\alpha' > 1$), we find that

$$\widetilde{Q}_{-\log x}(\rho\|\sigma) = \sup_{\omega} \langle \boldsymbol{\rho}| -\log \Delta(\sigma,\omega) |\boldsymbol{\rho}\rangle = \sup_{\omega} \langle \boldsymbol{\rho}| \alpha' \log \Delta(\sigma,\omega)^{-\frac{1}{\alpha'}} |\boldsymbol{\rho}\rangle$$
$$\leqslant \alpha' \log \sup_{\omega} \langle \boldsymbol{\rho}| \Delta(\sigma,\omega)^{-\frac{1}{\alpha'}} |\boldsymbol{\rho}\rangle \leqslant \widetilde{D}_{\alpha}(\rho\|\sigma) .$$

Moreover, it was proved in [BST18, theorem 13] that if $\rho \leqslant c\sigma$ for some $c > 0$, then

$$\lim_{\alpha\to 1^+} \widetilde{D}_{\alpha}(\rho\|\sigma) = D(\rho\|\sigma) .$$

For general case, we have

$$\widetilde{Q}_{-\log x}(\rho\ \|\sigma) = \lim_{\varepsilon\to 0^+} \widetilde{Q}_{-\log x}(\rho\|\sigma+\varepsilon\rho) = \lim_{\varepsilon\to 0^+} D(\rho\|\sigma+\varepsilon\rho) = D(\rho\|\sigma) .$$

Here the second limit follows from the fact that $D(\rho\|\sigma+\varepsilon\rho)$ is monotone non-decreasing and lower semi-continuity of D.

Recall that we denote by D_α the Petz–Rényi relative entropy. For $\alpha = 1$, we write $\widetilde{D}_1(\rho\|\sigma) = D_1(\rho\|\sigma) := D(\rho\|\sigma)$ as the standard relative entropy. The following lemma enables us to approximate relative entropy $D_\alpha(\rho\|\sigma)$ and $\widetilde{D}_\alpha(\rho\|\sigma)$ by ρ,σ with $s(\rho) = s(\sigma)$.

Lemma 5.6. *Let $\rho,\sigma \in D(\mathcal{M})$ be two normal states with $s(\rho) \leqslant s(\sigma)$. For $0 < \varepsilon < 1$, denote $\rho_\varepsilon = (1-\varepsilon)\rho + \varepsilon\sigma$. Then*

(*a*) *For all $0 < \alpha < 2$, $\lim_{\varepsilon\to 0} D_\alpha(\rho_\varepsilon\|\sigma) = D_\alpha(\rho\|\sigma)$;*
(*b*) *For all $1/2 \leqslant \alpha \leqslant \infty$, $\lim_{\varepsilon\to 0} \widetilde{D}_\alpha(\rho_\varepsilon\|\sigma) = \widetilde{D}_\alpha(\rho\|\sigma)$.*

Proof. Write $\mathrm{id}^n_{\infty,2} : \mathcal{M} \to \mathcal{M}$ as the identity map and define the normal UCP map

$$\Psi_\sigma : \mathcal{M} \to \mathcal{M}, \Psi_\sigma(x) = \sigma(x)1 .$$

It is clear that the adjoint $\Psi^\dagger_\sigma(\rho) = \sigma$ for any state $\rho \in D(\mathcal{M})$. Take the normal UCP map $\Psi_\varepsilon = (1-\varepsilon)\mathrm{id} + \varepsilon\Psi$. Then $\Psi^\dagger_\varepsilon(\rho) = \rho \circ \Psi_\varepsilon = \rho_\varepsilon$ and $\Psi^\dagger_\varepsilon(\sigma) = \sigma$. For (a), using data processing inequality of D_α

$$\limsup_{\varepsilon\to 0} D_\alpha(\rho_\varepsilon\|\sigma) = \limsup_{\varepsilon\to 0} D_\alpha(\Psi^\dagger_\varepsilon(\rho)\|\Psi^\dagger_\varepsilon(\sigma)) \leqslant \limsup_{\varepsilon\to 0} D_\alpha(\rho_\varepsilon\|\sigma)$$
$$\leqslant D_\alpha(\rho\|\sigma) \leqslant \liminf_{\varepsilon\to 0} D_\alpha(\rho_\varepsilon\|\sigma) ,$$

where the last inequality uses the lower semi-continuity [Hia18, theorem 4.1]. The argument for (b) and $\alpha > 1$ is similar by using the data processing inequality and the lower semi-continuity of $\widetilde{D}_\alpha$ [Jen18, proposition 3.7 and theorem 3.11]. For $1/2 \leqslant \alpha < 1$, the lower semi-continuity can be replaced by

$$\liminf_{\varepsilon\to\infty} \widetilde{D}_\alpha(\rho_\varepsilon\|\sigma) \geqslant \liminf_{\varepsilon\to\infty} \widetilde{D}_\alpha((1-\varepsilon)\rho\|\sigma)$$
$$= \liminf_{\varepsilon\to\infty} \widetilde{D}_\alpha(\rho\|\sigma) + \alpha' \log(1-\varepsilon) = \widetilde{D}_\alpha(\rho\|\sigma) . \qquad \square$$

5.3. Data-processing inequality for optimized f-divergence

We now establish the data-processing inequality for the optimized f-divergence $\widetilde{Q}_f$. We start with the key case of restricting to a subalgebra.

Lemma 5.7. *Let $\mathcal{M} \subset B(H)$ be a von Neumann algebra, and let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. Let $\boldsymbol{\rho}, \boldsymbol{\sigma} \in H$ be two unit vectors, and let $\rho_{\mathcal{M}}, \sigma_{\mathcal{M}}$ (resp. $\rho_{\mathcal{N}}, \sigma_{\mathcal{N}}$) be the corresponding normal states on $\mathcal{M}$ (resp. $\mathcal{N}$). Then for an operator anti-monotone function $f : (0, \infty) \to \mathbb{R}$, the following inequality holds*

$$\widetilde{Q}_f(\rho_{\mathcal{M}} \| \sigma_{\mathcal{M}}) \geqslant \widetilde{Q}_f(\rho_{\mathcal{N}} \| \sigma_{\mathcal{N}}). \tag{82}$$

Proof. For two vectors $\boldsymbol{\sigma}, \boldsymbol{\omega} \in H$, we write $\Delta^{\mathcal{M}}(\boldsymbol{\sigma}, \boldsymbol{\omega})$ (resp. $\Delta^{\mathcal{N}}(\boldsymbol{\sigma}, \boldsymbol{\omega})$) as the relative modular operator with respect to $\mathcal{M}$ (resp. $\mathcal{N}$). Let $S^{\mathcal{M}}_{\boldsymbol{\sigma},\boldsymbol{\omega}}$ and $S^{\mathcal{N}}_{\boldsymbol{\sigma},\boldsymbol{\omega}}$ be the corresponding anti-linear operators such that

$$(S^{\mathcal{M}}_{\boldsymbol{\sigma},\boldsymbol{\omega}})^* \bar{S}^{\mathcal{M}}_{\boldsymbol{\sigma},\boldsymbol{\omega}} = \Delta^{\mathcal{M}}(\boldsymbol{\sigma}, \boldsymbol{\omega}) , \qquad (S^{\mathcal{N}}_{\boldsymbol{\sigma},\boldsymbol{\omega}})^* \bar{S}^{\mathcal{N}}_{\boldsymbol{\sigma},\boldsymbol{\omega}} = \Delta^{\mathcal{N}}(\boldsymbol{\sigma}, \boldsymbol{\omega}) .$$

Recall the support projections are given by

$$s_{\mathcal{M}}(\boldsymbol{\omega}) = [\mathcal{M}' \boldsymbol{\omega}] , \qquad s_{\mathcal{N}'}(\boldsymbol{\omega}) = [\mathcal{N} \boldsymbol{\omega}] .$$

By the definition of the S operators, we find that

$$S^{\mathcal{M}}_{\boldsymbol{\sigma},\boldsymbol{\omega}} s_{\mathcal{N}'}(\boldsymbol{\omega}) = s_{\mathcal{M}}(\boldsymbol{\omega}) S^{\mathcal{N}}_{\boldsymbol{\sigma},\boldsymbol{\omega}} , \qquad \Delta^{\mathcal{N}}(\boldsymbol{\sigma}, \boldsymbol{\omega}) \geqslant s_{\mathcal{N}'}(\boldsymbol{\omega}) \Delta^{\mathcal{M}}(\boldsymbol{\sigma}, \boldsymbol{\omega}) s_{\mathcal{N}'}(\boldsymbol{\omega}).$$

Then for all $\boldsymbol{\omega}$ such that $\|\boldsymbol{\omega}\|_2 = 1$ and $\boldsymbol{\rho} \in [\mathcal{N}\boldsymbol{\omega}] = s_{\mathcal{N}'}(\boldsymbol{\omega})$, we find that

$$\begin{aligned} \langle \boldsymbol{\rho} | \, f(\Delta^{\mathcal{N}}(\boldsymbol{\sigma}, \boldsymbol{\omega})) \, | \boldsymbol{\rho} \rangle &\leqslant \langle \boldsymbol{\rho} | \, f(s_{\mathcal{N}'}(\boldsymbol{\omega}) \Delta^{\mathcal{M}}(\boldsymbol{\sigma}, \boldsymbol{\omega}) s_{\mathcal{N}'}(\boldsymbol{\omega})) \, | \boldsymbol{\rho} \rangle \\ &\leqslant \langle \boldsymbol{\rho} | \, s_{\mathcal{N}'}(\boldsymbol{\omega}) f(\Delta^{\mathcal{M}}(\boldsymbol{\sigma}, \boldsymbol{\omega})) s_{\mathcal{N}'}(\boldsymbol{\omega}) \, | \boldsymbol{\rho} \rangle \\ &= \langle \boldsymbol{\rho} | \, f(\Delta^{\mathcal{M}}(\boldsymbol{\sigma}, \boldsymbol{\omega})) \, | \boldsymbol{\rho} \rangle . \end{aligned}$$

Here we view the projection $s_{\mathcal{N}'}(\boldsymbol{\omega})$ as an isometry on its support. Noting that $\boldsymbol{\rho} \in [\mathcal{N}\boldsymbol{\omega}] \subset [\mathcal{M}\boldsymbol{\omega}]$, then

$$\begin{aligned} \widetilde{Q}_f(\rho_{\mathcal{N}} \| \sigma_{\mathcal{N}}) &= \sup_{\boldsymbol{\omega} \,:\, \|\boldsymbol{\omega}\|_2 = 1, \boldsymbol{\rho} \in [\mathcal{N}\boldsymbol{\omega}]} \langle \boldsymbol{\rho} | f(\Delta^{\mathcal{N}}(\sigma, \boldsymbol{\omega})) | \boldsymbol{\rho} \rangle \\ &\leqslant \sup_{\boldsymbol{\omega} \,:\, \|\boldsymbol{\omega}\|_2 = 1, \boldsymbol{\rho} \in [\mathcal{N}\boldsymbol{\omega}]} \langle \boldsymbol{\rho} | f(\Delta^{\mathcal{M}}(\sigma, \boldsymbol{\omega})) | \boldsymbol{\rho} \rangle \\ &\leqslant \sup_{\boldsymbol{\omega} \,:\, \|\boldsymbol{\omega}\|_2 = 1, \boldsymbol{\rho} \in [\mathcal{M}\boldsymbol{\omega}]} \langle \boldsymbol{\rho} | f(\Delta^{\mathcal{M}}(\sigma, \boldsymbol{\omega})) | \boldsymbol{\rho} \rangle = \widetilde{Q}_f(\rho_{\mathcal{M}} \| \sigma_{\mathcal{M}}). \end{aligned}$$

This concludes the proof. □

Lemma 5.8. *Let $\mathcal{M}$ be a von Neumann algebra, and let $e \in \mathcal{M}$ be a projection. Let $\rho, \sigma \in D(\mathcal{M})$ be two normal states with support $s(\rho) \leqslant s(\sigma) \leqslant e$. Let σ_e, ρ_e denote the corresponding normal states on $e\mathcal{M}e$. Then for all operator anti-monotone functions $f : (0, \infty) \to \mathbb{R}$, the following equality holds*

$$\widetilde{Q}_f(\rho \| \sigma) = \widetilde{Q}_f(\rho_e \| \sigma_e). \tag{83}$$

Proof. We use the standard form $(\mathcal{M}, L_2(\mathcal{M}), J, L_2(\mathcal{M})_+)$ from appendix A.3. The standard form of $e\mathcal{M}e$ is $(e\mathcal{M}e, eL_2(\mathcal{M})e, J, eL_2(\mathcal{M})_+e)$. Let $V : eL_2(\mathcal{M})e \hookrightarrow L_2(\mathcal{M})$ be the isometry that is the adjoint of the projection $P : L_2(\mathcal{M}) \to eL_2(\mathcal{M})e$ with $P(x) = exe$. Let $h_\rho^{1/2}$ and $h_\sigma^{1/2}$ be the vectors in $L_2(\mathcal{M})_+$ corresponding to ρ and σ, respectively. Since $s(\rho) \leqslant s(\sigma) \leqslant e$, we have that $h_\rho^{1/2} = eh_\rho^{1/2} = eh_\rho^{1/2}e$ and similarly for $h_\sigma^{1/2}$. Let $\omega \in D(\mathcal{M})$ be a normal state, and let $h_\omega^{1/2} \in L_2(\mathcal{M})_+$ be the corresponding unit vector. Let $\omega_e \in (e\mathcal{M}e)_+$ be the restriction of ω on $e\mathcal{M}e$. Note that ω_e is a sub-state corresponding to $eh_\omega e \in eL_1(\mathcal{M})e \cong L_1(e\mathcal{M}e)$. By proposition 5.3, it suffices to consider ω such that $\omega_e \neq 0$. Otherwise we can always replace ω by $\omega_\varepsilon = (1-\varepsilon)\omega + \varepsilon\rho$.

Recall that $\Delta_{\mathcal{M}}(\sigma,\omega)^{-1} = J\Delta_{\mathcal{M}}(\omega,\sigma)J$ and for $x \in \mathcal{M}$, $\Delta(\omega,\sigma)^{1/2}JP\left|h_\sigma^{1/2}x\right\rangle = \left|h_\omega^{1/2}exe\right\rangle$. Then we find that

$$\begin{aligned}
&\left\langle h_\sigma^{1/2}x\right| P\Delta_{\mathcal{M}}(\sigma,\omega)^{-1}P\left|h_\sigma^{1/2}x\right\rangle \\
&\quad= \left\langle h_\omega^{1/2}exe\,\middle|\,h_\omega^{1/2}exe\right\rangle \\
&\quad= \mathrm{tr}\,(ex^*eh_\omega exe) = \left\langle h_\sigma^{1/2}exe\right|\Delta_{e\mathcal{M}e}(\sigma,\omega_e)^{-1}\left|h_\sigma^{1/2}exe\right\rangle.
\end{aligned}$$

This implies that

$$P\Delta_{\mathcal{M}}(\sigma,\omega)^{-1}P = \Delta_{e\mathcal{M}e}(\sigma,\omega_e)^{-1}.$$

For $f : (0,\infty) \to \mathbb{R}$ operator anti-monotone, $\widetilde{f}(x) = f(x^{-1})$ is operator monotone and operator concave. Since $h_\rho^{1/2} \in eL_2(\mathcal{M})e = PL_2(\mathcal{M})$,

$$\begin{aligned}
\left\langle h_\rho^{1/2}\right| f(\Delta_{\mathcal{M}}(\sigma,\omega))\left|h_\rho^{1/2}\right\rangle &= \left\langle h_\rho^{1/2}\right| P\widetilde{f}(\Delta_{\mathcal{M}}(\sigma,\omega)^{-1})P\left|h_\rho^{1/2}\right\rangle \\
&\leqslant \left\langle h_\rho^{1/2}\right| \widetilde{f}(P\Delta_{\mathcal{M}}(\sigma,\omega)^{-1}P)\left|h_\rho^{1/2}\right\rangle \\
&= \left\langle h_\rho^{1/2}\right| \widetilde{f}(\Delta_{e\mathcal{M}e}(\sigma,\omega_e)^{-1})\left|h_\rho^{1/2}\right\rangle \\
&\leqslant \left\langle h_\rho^{1/2}\right| \widetilde{f}(\Delta_{e\mathcal{M}e}(\sigma,\overline{\omega}_e)^{-1})\left|h_\rho^{1/2}\right\rangle \\
&\leqslant \widetilde{Q}_f(\rho_e\|\sigma_e),
\end{aligned} \tag{84}$$

where $\overline{\omega}_e = \frac{\omega_e}{\omega_e(1)}$ is the normalized state of ω_e. By taking all $\omega \gg \rho$,

$$\widetilde{Q}_f(\rho\|\sigma) \leqslant \widetilde{Q}_f(\rho_e\|\sigma_e)\ .$$

The reverse inequality follows from lemma 5.7 because $e\mathcal{M}e \subset \mathcal{M}$ as a (non-unital) subalgebra. □

Remark 5.9. The lemma above is an extension of isometric invariance [Wil18a, proposition 4] in the finite-dimensional case. It implies that it suffices to consider the optimized f-divergence on σ-finite von Neumann algebras. Indeed, we can always restrict to $e\mathcal{M}e$

for $e = s(\rho + \sigma)$ because $\widetilde{Q}_f(\rho\|\sigma) = \widetilde{Q}_f(\rho_e\|\sigma_e)$. Based on that, one can further deduce the following variant of proposition 5.3:

$$\widetilde{Q}_f(\rho\|\sigma) = \sup_{\omega \in D_+(\mathcal{M})} \left\langle h_\rho^{1/2} \middle| f(\Delta(\sigma,\omega)) \middle| h_\rho^{1/2} \right\rangle,$$

where $D_+(\mathcal{M})$ is the set of all faithful normal states.

Theorem 5.10 (Data-processing inequality). *Let $\Phi : \mathcal{N} \to \mathcal{M}$ be a normal completely positive unital map, and let $\rho, \sigma \in D(\mathcal{M})$ be two normal states. For $f : (0,\infty) \to \mathbb{R}$ operator anti-monotone, the following data-processing inequality holds*

$$\widetilde{Q}_f(\rho\|\sigma) \geqslant \widetilde{Q}_f(\rho \circ \Phi\|\sigma \circ \Phi).$$

Proof. Let $\mathcal{M} \subset B(H)$, and let $\boldsymbol{\rho}, \boldsymbol{\sigma} \in H$ be the vectors implementing ρ, σ, respectively. Let $\Phi(\cdot) = V^*\pi(\cdot)V$ be the Stinespring dilation of Φ, where $\pi : \mathcal{N} \to B(K)$ is a normal $*$-homomorphism and $V : H \to K$ is an isometry [Sti55]. Let $\rho_1 = \rho \circ \Phi$ and $\sigma_1 = \sigma \circ \Phi$ denote states on $\mathcal{N}$. Then $\boldsymbol{\rho}_1 = V\boldsymbol{\rho}$ and $\boldsymbol{\sigma}_1 = V\boldsymbol{\sigma}$ are vector representations of ρ_1 and σ_1, respectively, via π because

$$\rho \circ \Phi(x) = \rho(V^*\pi(x)V) = \langle \boldsymbol{\rho}| V^*\pi(x)V |\boldsymbol{\rho}\rangle \ ,$$
$$\sigma \circ \Phi(x) = \sigma(V^*\pi(x)V) = \langle \boldsymbol{\sigma}| V^*\pi(x)V |\boldsymbol{\sigma}\rangle \ .$$

Take the projection $e = VV^* \in B(H)$. Let $\mathcal{L} \subset B(K)$ denote the von Neumann subalgebra in $B(K)$ generated by $V\mathcal{M}V^*$ and $\pi(\mathcal{N})$. Note that $V : H \to eK$ is a surjective isometry and define the map $T : B(eK) \to B(H)$ as

$$x \mapsto V^*xV \ .$$

The map T is a $*$-isomorphism that sends $e\mathcal{L}e$ to $\mathcal{M}$. Thus we have the following factorization of Φ:

$$\mathcal{N} \xrightarrow{\pi} \pi(\mathcal{N}) \hookrightarrow \mathcal{L} \to e\mathcal{L}e \xrightarrow{T} \mathcal{M} \ . \tag{85}$$

Let us introduce the shorthand $\widetilde{Q}_f^{\mathcal{M}}(\boldsymbol{\rho}\|\boldsymbol{\sigma}) \equiv \widetilde{Q}_f(\rho_{\mathcal{M}}\|\sigma_{\mathcal{M}})$, where $\rho_{\mathcal{M}}, \sigma_{\mathcal{M}}$ are the states on $\mathcal{M}$ implemented by the vectors $\boldsymbol{\rho}, \boldsymbol{\sigma}$. Using this notation, we have

$$\begin{aligned}\widetilde{Q}_f(\rho \circ \Phi\|\sigma \circ \Phi) &= \widetilde{Q}_f^{\pi(\mathcal{N})}(\boldsymbol{\rho}_1\|\boldsymbol{\sigma}_1) \leqslant \widetilde{Q}_f^{\mathcal{L}}(\boldsymbol{\rho}_1\|\boldsymbol{\sigma}_1) = \widetilde{Q}_f^{e\mathcal{L}e}(\boldsymbol{\rho}_1\|\boldsymbol{\sigma}_1)\\ &= \widetilde{Q}_f^{\mathcal{M}}(\boldsymbol{\rho}\|\boldsymbol{\sigma}) = \widetilde{Q}_f(\rho\|\sigma).\end{aligned}$$

Here the first equality follows from the independence in lemma 5.2. The inequality follows from the inclusion $\pi(\mathcal{N}) \subset \mathcal{L}$ and lemma 5.7. The second equality follows because $\boldsymbol{\rho}_1, \boldsymbol{\sigma}_1 \in eK$ and by applying lemma 5.8. The last step is a $*$-isomorphism. □

It is clear from the argument above that the actual inequality in data processing is the inclusion $\pi(\mathcal{N}) \subset \mathcal{L}$.

5.4. Recoverability results

In this section, we discuss recoverability results in the setting of general von Neumann algebras. We first review the generalized conditional expectation introduced in [AC82], which is the (dual of) Petz map of the inclusion $\mathcal{N} \subset \mathcal{M}$ in the Heisenberg picture.

Let $\mathcal{M}$ be a von Neumann algebra, and let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. We denote by $(\mathcal{M}, L_2(\mathcal{M}), J, L_2(\mathcal{M})^+)$ (resp. $(\mathcal{N}, L_2(\mathcal{N}), J_0, L_2(\mathcal{N})^+)$) the standard form of $\mathcal{M}$ (resp. $\mathcal{N}$) using Haagerup L_2-spaces. Given a normal state $\rho \in D(\mathcal{M})$ and its restriction $\rho_{\mathcal{N}}$ in $D(\mathcal{N})$, we denote by h_ρ (resp. $h_{\rho_{\mathcal{N}}}$) the density operator of ρ (resp. $\rho_{\mathcal{N}}$) in $L_1(\mathcal{M})$ (resp. $L_1(\mathcal{N})$). Thus $h_\rho^{1/2} \in L_2(\mathcal{M})$ (resp. $h_{\rho_{\mathcal{N}}}^{1/2} \in L_2(\mathcal{N})$) is a vector representation of ρ (resp. $\rho_{\mathcal{N}}$). Define the partial isometry $V_\rho : L_2(\mathcal{N}) \to L_2(\mathcal{M})$ as

$$V_\rho(ah_{\rho_{\mathcal{N}}}^{1/2} + \xi) = ah_\rho^{1/2}\ , \quad \forall\ a \in \mathcal{N}, \xi \in [\mathcal{N}h_{\rho_{\mathcal{N}}}^{1/2}]^\perp\ .$$

Indeed,

$$\left\| V_\rho(ah_{\rho_{\mathcal{N}}}^{1/2}) \right\|_{L_2(\mathcal{M})}^2 = \left\| ah_\rho^{1/2} \right\|_2^2 = \operatorname{tr}(a^*ah_\rho) = \rho(a^*a) = \left\| ah_{\rho_{\mathcal{N}}}^{1/2} \right\|_{L_2(\mathcal{M})}^2.$$

The ρ-preserving generalized conditional expectation $E_\rho : \mathcal{M} \to \mathcal{N}$ is defined as follows:

$$E_\rho(x) := J_0 V_\rho J x J V_\rho J_0\ .$$

Observe that $E_\rho : \mathcal{M} \to \mathcal{N}$ is a normal completely positive sub-unital map. Moreover $E_\rho(s(\rho)) = s_{\mathcal{N}}(\rho)$ and $E_\rho(1 - s(\rho)) = 0$ where $s(\rho)$ (resp. $s_{\mathcal{N}}(\rho)$) is the support of ρ (resp. $\rho_{\mathcal{N}}$). It was proved by Petz [Pet88] that if $D(\rho\|\sigma) < \infty$, then the equality $D(\rho\|\sigma) = D(\rho_{\mathcal{N}}\|\sigma_{\mathcal{N}})$ is equivalent to the following conditions:

(a) $E_\rho = E_\sigma$;
(b) $\rho_{\mathcal{N}} \circ E_\sigma = \rho$;
(c) $\sigma_{\mathcal{N}} \circ E_\rho = \sigma$.

In this sense E_ρ (or equivalently E_σ) is a recovery for the inclusion $\mathcal{N} \subset \mathcal{M}$.

In general, consider a normal completely positive unital map $\Phi : \mathcal{N} \to \mathcal{M}$. Let $\rho \in D(\mathcal{M})$ be a state, and set $\rho_0 = \rho \circ \Phi \in D(\mathcal{N})$. The Petz map $R := R_{\Phi,\rho} : \mathcal{M} \to \mathcal{N}$ is the unique normal completely positive sub-unital map such that

$$R(s(\rho)) = s(\rho_0)\ , \qquad R(1 - s(\rho)) = 0\ ,$$

and $\forall\ \ x \in \mathcal{N},\ y \in \mathcal{M}$,

$$\langle Jyh_\rho^{1/2}, J_0\Phi(x)h_{\rho_0}^{1/2} \rangle = \langle JR(y)h_\rho^{1/2}, J_0xh_{\rho_0}^{1/2} \rangle\ . \tag{86}$$

In particular, if $\rho_0 = \rho \circ \Phi$ is faithful, then R is unital.

Recall that the modular automorphism group $\alpha_t^\rho : \mathcal{M} \to \mathcal{M}$ for a state ρ is given by

$$\alpha_t^\rho(x) = \Delta(\rho,\rho)^{-it} x \Delta(\rho,\rho)^{it}\ .$$

The rotated Petz map is defined as follows:

$$E_\rho^t(x) := \alpha_t^{\rho_{\mathcal{N}}} \circ E_\rho \circ \alpha_{-t}^\rho\ , \qquad R_{\Phi,\rho}^t(x) = \alpha_t^{\rho_0} \circ R_{\Phi,\rho} \circ \alpha_{-t}^\rho\ . \tag{87}$$

Recall that in the Stinespring dilation $\Phi(\cdot) = V^*\pi(\cdot)V$, π can be faithful (cf [Pis20, theorem 1.41]). By the same argument in the proof of theorem 5.10, it suffices to consider two cases:

(a) For an inclusion $\iota : \mathcal{N} \to \mathcal{M}$, $R_{\iota,\rho} = E_\rho$ is the generalized conditional expectation.

(b) Consider the projection map

$$P : \mathcal{M} \to e\mathcal{M}e\ ,\ P(x) = exe$$

for a projection $e \in \mathcal{M}$ and let ρ be a state with $s(\rho) \leqslant e$. The recovery map $R_{P,\rho} = \iota_\rho : s(\rho)\mathcal{M}s(\rho) \to \mathcal{M}$ is the embedding and so is the rotated Petz map $R^t_{P,\rho} = \alpha^\rho_t \circ \iota_\rho \circ \alpha^\rho_{-t} = \iota_\rho$.

Let $\Phi : \mathcal{N} \to \mathcal{M}$ be a general normal UCP map given by the composition $\Phi = P \circ \iota$. Note that by the symmetric role of Φ and $R_{\Phi,\rho}$ in (86), the Petz map $R_{\Phi,\rho} = R_{\iota,\rho} \circ R_{P,\rho} = E_\rho \circ \iota_\rho$ is a composition of the Petz map of the above two cases. Similarly for a rotated Petz map,

$$R^t_{\Phi,\rho} = \alpha^{\rho_0}_t \circ R_{\Phi,\rho} \circ \alpha^\rho_{-t} = (\alpha^{\rho_0}_t \circ E_\rho \circ \alpha^\rho_{-t}) \circ (\alpha^\rho_t \circ \iota_\rho \circ \alpha^\rho_{-t}) = E^t_\rho \circ \iota_\rho\ .$$

Since the embedding $\iota_\rho : s(\rho)\mathcal{M}s(\rho) \hookrightarrow \mathcal{M}$ always preserves the L_1-norm and (optimized) f-divergence on its support (lemma 5.8), it suffices to consider the recovery result for E^t_ρ.

We now extend the recovery results in section 4 to the general setting. For simplicity, we will mainly focus on faithful cases. The main steps that need adaptation are lemmas 4.2, 4.10, and 4.18, which we reproduce here using standard form on Haagerup L_2-spaces.

Lemma 5.11. *Let ρ, σ, and ω be normal states, and let $|\boldsymbol{\rho}\rangle = h^{1/2}_\rho \in L_2(\mathcal{M})$ be the vector representation of ρ. Suppose $|\boldsymbol{\rho}\rangle \in supp(\Delta(\sigma,\omega)) = s(\sigma)s(\omega')$. Then for all $t \in \mathbb{R}$,*

$$\langle \boldsymbol{\rho}|\, \Delta(\sigma,\omega)^{-it} x \Delta(\sigma,\omega)^{it}\, |\boldsymbol{\rho}\rangle = \rho \circ \alpha^\sigma_t(x).$$

Thus $\Delta(\sigma,\omega)^{-it} x \Delta(\sigma,\omega)^{it} = \alpha^\sigma_t(x)$.

Proof. Let h_ρ, h_σ, and h_ω be the density operators of ρ, σ, and ω, respectively. We have

$$|\boldsymbol{\rho}\rangle = \left|h^{1/2}_\rho\right\rangle\ , \qquad \Delta(\sigma,\omega)^{it}\left|h^{1/2}_\rho\right\rangle = \left|h^{it}_\sigma h^{1/2}_\rho h^{-it}_\omega\right\rangle\ .$$

Then for $x \in \mathcal{M}$,

$$\begin{aligned}\langle \boldsymbol{\rho}|\, \Delta(\sigma,\omega)^{-it} x \Delta(\sigma,\omega)^{it}\, |\boldsymbol{\rho}\rangle &= \mathrm{tr}\,((h^{it}_\sigma h^{1/2}_\rho h^{-it}_\omega)^* x h^{it}_\sigma h^{1/2}_\rho h^{-it}_\omega) \\ &= \mathrm{tr}\,(h^{it}_\omega h^{1/2}_\rho h^{-it}_\sigma x h^{it}_\sigma h^{1/2}_\rho h^{-it}_\omega) \\ &= \mathrm{tr}\,(h_\rho h^{-it}_\sigma x h^{it}_\sigma) \\ &= \mathrm{tr}\,(h_\rho \alpha^t_\sigma(x)) \\ &= \rho \circ \alpha^t_\sigma(x).\end{aligned}$$

□

Lemma 5.12. *Let $\rho \in D_+(\mathcal{M})$ and $\omega_\mathcal{N} \in D_+(\mathcal{N})$ be faithful. Then*

$$V^*_\rho \Delta_\mathcal{M}(\sigma, E^\dagger_\rho(\omega_\mathcal{N}))V_\rho = \Delta_\mathcal{N}(\sigma_\mathcal{N}, \omega_\mathcal{N})\ .$$

As a consequence, for all operator anti-monotone functions $f : (0,\infty) \to \mathbb{R}$,

$$\left\langle h^{1/2}_{\rho_\mathcal{N}}\right| f(\Delta_\mathcal{N}(\sigma_\mathcal{N}, \omega_\mathcal{N}))\left|h^{1/2}_{\rho_\mathcal{N}}\right\rangle \leqslant \left\langle h^{1/2}_\rho\right| f(\Delta_\mathcal{M}(\sigma, R_\rho(\omega_\mathcal{N})))\left|h^{1/2}_\rho\right\rangle\ .$$

Proof. Let $h_{\rho_\mathcal{N}}, h_{\omega_\mathcal{N}}, h_{\sigma_\mathcal{N}}$ and $h_\rho, h_\omega, h_\sigma$ be the corresponding density operators. Let $S_{\sigma_\mathcal{N},\omega_\mathcal{N}} : L_2(\mathcal{N}) \to L_2(\mathcal{N})$ (resp. $S_{\sigma,\omega}$) be the anti-linear operator for the standard form of $\mathcal{N}$ (resp. $\mathcal{M}$). We have for $a \in \mathcal{N}$,

$$V_\sigma S_{\sigma_\mathcal{N},\omega_\mathcal{N}}(ah_{\omega_\mathcal{N}}^{1/2}) = V_\sigma(a^* h_{\sigma_\mathcal{N}}^{1/2}) = a^* h_\sigma^{1/2} \ .$$

On the other hand, for any $a, b \in \mathcal{N}$,

$$V_\rho(abh_\rho^{1/2}) = abh_\rho^{1/2} = aV_\rho(bh_\rho^{1/2}) \ .$$

By the density of $\mathcal{N}h_{\rho_\mathcal{N}}^{1/2}$ in $L_2(\mathcal{N})$, this implies $aV_\rho = V_\rho a$. Then if we choose the L_2 vector $\boldsymbol{\omega} = V_\rho h_{\omega_\mathcal{N}}^{1/2}$,

$$S_{\sigma,\boldsymbol{\omega}} V_\rho(ah_{\omega_\mathcal{N}}^{1/2}) = S_{\sigma,\boldsymbol{\omega}}(a\boldsymbol{\omega}) = a^* h_\sigma^{1/2} \ .$$

Note that $\Delta_\mathcal{M}(\sigma, \boldsymbol{\omega})$ only depends on $\omega' \in \mathcal{M}'$ induced by $\boldsymbol{\omega}$. Indeed, for $x \in \mathcal{M}$,

$$\begin{aligned}
\langle \boldsymbol{\omega} | JxJ | \boldsymbol{\omega} \rangle &= \left\langle h_{\omega_\mathcal{N}}^{1/2} \right| V_\rho^* JxJV_\rho \left| h_{\omega_\mathcal{N}}^{1/2} \right\rangle \\
&= \left\langle h_{\omega_\mathcal{N}}^{1/2} \right| J_0 V_\rho^* JxJV_\rho J_0 \left| h_{\omega_\mathcal{N}}^{1/2} \right\rangle \\
&= \omega_\mathcal{N} \circ E_\rho(x) \\
&= \mathrm{tr}\,(xh_{\omega_\mathcal{N} \circ E_\rho}).
\end{aligned}$$

Thus $S_{\sigma,\boldsymbol{\omega}}^* \bar{S}_{\sigma,\boldsymbol{\omega}} = \Delta_\mathcal{M}(\sigma, \omega)$ for $\omega = \omega_\mathcal{N} \circ E_\rho$. Thus for this choice $\boldsymbol{\omega} = V_\rho h_{\omega_\mathcal{N}}^{1/2}$,

$$S_{\sigma,\boldsymbol{\omega}} V_\rho = V_\sigma S_{\sigma_\mathcal{N},\omega_\mathcal{N}} \ , \ V_\rho^* \Delta_\mathcal{M}(\sigma, \omega) V_\rho = \Delta_\mathcal{N}(\sigma_\mathcal{N}, \omega_\mathcal{N}) \ .$$

The other assertion follows from operator convexity and operator monotonicity of f. □

Lemma 5.13. *Let $\rho, \sigma \in D_+(\mathcal{M})$ and $\omega_\mathcal{N} \in D_+(\mathcal{N})$ be faithful. Define the vectors*

$$\begin{aligned}
|a_t\rangle &:= J\Delta(\sigma, \rho)^{-it} V_\rho \Delta(\sigma_\mathcal{N}, \rho_\mathcal{N})^{\frac{1}{2}+it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle, \\
|b_t\rangle &:= \Delta(\sigma, \rho)^{\frac{1}{2}+it} V_\rho \Delta(\sigma_\mathcal{N}, \rho_\mathcal{N})^{-\frac{1}{2}-it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle, \\
|c_t\rangle &:= \Delta(\sigma, E_\rho(\omega_\mathcal{N}))^{1/2+it} V_\rho \Delta(\sigma_\mathcal{N}, \omega_\mathcal{N})^{-1/2-it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle.
\end{aligned}$$

The following equalities hold for $x \in \mathcal{M}$:

$$\langle a_t | x | a_t \rangle = \sigma_\mathcal{N} \circ E_\rho^t(x) \ , \quad \langle b_t | x | b_t \rangle = \rho_\mathcal{N} \circ E_\sigma^{-t}(x) \ , \quad \langle c_t | x | c_t \rangle = \rho_\mathcal{N} \circ E_\sigma^{-t}(x) \ .$$

Proof. For the first one,

$$\begin{aligned}
|a_t\rangle &= J\Delta(\sigma, \rho)^{-it} V_\rho \Delta(\sigma_\mathcal{N}, \rho_\mathcal{N})^{\frac{1}{2}+it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle \\
&= J\Delta(\sigma, \rho)^{-it} JJV_\rho J_0 J_0 \Delta(\sigma_\mathcal{N}, \rho_\mathcal{N})^{\frac{1}{2}+it} J_0 \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle
\end{aligned}$$

$$= \Delta(\rho,\sigma)^{-it} J V_\rho J_0 \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{it} \left| h_{\sigma_{\mathcal{N}}}^{1/2} \right\rangle .$$

Then

$$\begin{aligned}
&\langle a_t | \, x \, | a_t \rangle \\
&= \left\langle h_{\sigma_{\mathcal{N}}}^{1/2} \right| \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{-it} J_0 V_\rho^* J \Delta(\rho,\sigma)^{it} x \Delta(\rho,\sigma)^{-it} J V_\rho J_0 \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{it} \left| h_{\sigma_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \left\langle h_{\sigma_{\mathcal{N}}}^{1/2} \right| \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{-it} J_0 V_\rho^* J \alpha_{-t}^{\rho}(x) J V_\rho J_0 \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{it} \left| h_{\sigma_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \left\langle h_{\sigma_{\mathcal{N}}}^{1/2} \right| \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{-it} E_\rho \circ \alpha_{-t}^{\rho}(x) \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{it} \left| h_{\sigma_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \left\langle h_{\sigma_{\mathcal{N}}}^{1/2} \right| (\alpha_t^{\rho} \circ E_\rho \circ \alpha_{-t}^{\rho_{\mathcal{N}}})(x) \left| h_{\sigma_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \left\langle h_{\sigma_{\mathcal{N}}}^{1/2} \right| E_\rho^t(x) \left| h_{\sigma_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \sigma_{\mathcal{N}} \circ E_\rho^t(x).
\end{aligned}$$

For the second one, we first show that

$$\Delta(\sigma,\rho)^{\frac{1}{2}} V_\rho \Delta(\sigma_{\mathcal{N}},\rho_{\mathcal{N}})^{-\frac{1}{2}} = J V_\sigma J_0.$$

Indeed, for $a \in \mathcal{M}$

$$\begin{aligned}
J V_\sigma J_0 \left| h_{\sigma_{\mathcal{N}}}^{1/2} a \right\rangle &= J V_\sigma \left| a^* h_{\sigma_{\mathcal{N}}}^{1/2} \right\rangle = J \left| a^* h_{\sigma}^{1/2} \right\rangle = \left| h_{\sigma}^{1/2} a \right\rangle \\
\Delta(\sigma,\rho)^{\frac{1}{2}} V_\rho \Delta(\sigma_{\mathcal{N}},\rho_{\mathcal{N}})^{-\frac{1}{2}} \left| h_{\sigma_{\mathcal{N}}}^{1/2} a \right\rangle &= \Delta(\sigma,\rho)^{\frac{1}{2}} V_\rho J \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{\frac{1}{2}} J \left| h_{\sigma_{\mathcal{N}}}^{1/2} a \right\rangle \\
&= \Delta(\sigma,\rho)^{\frac{1}{2}} V_\rho J \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})^{\frac{1}{2}} \left| a^* h_{\sigma_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \Delta(\sigma,\rho)^{\frac{1}{2}} V_\rho J \left| h_{\rho_{\mathcal{N}}}^{1/2} a^* \right\rangle \\
&= \Delta(\sigma,\rho)^{\frac{1}{2}} V_\rho \left| a h_{\rho_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \Delta(\sigma,\rho)^{\frac{1}{2}} \left| a h_{\rho}^{1/2} \right\rangle \\
&= \left| h_{\sigma}^{1/2} a \right\rangle .
\end{aligned}$$

Then for $x \in \mathcal{M}$,

$$\begin{aligned}
\langle b_t | \, x \, | b_t \rangle &= \left\langle h_{\rho_{\mathcal{N}}}^{1/2} \right| \Delta(\sigma_{\mathcal{N}},\rho_{\mathcal{N}})^{-\frac{1}{2}+it} V_\rho^* \Delta(\sigma,\rho)^{\frac{1}{2}-it} x \Delta(\sigma,\rho)^{\frac{1}{2}+it} \\
&\quad \times V_\rho \Delta(\sigma_{\mathcal{N}},\rho_{\mathcal{N}})^{-\frac{1}{2}-it} \left| h_{\rho_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \left\langle h_{\rho_{\mathcal{N}}}^{1/2} \right| \Delta(\sigma_{\mathcal{N}},\rho_{\mathcal{N}})^{it} J_0 V_\sigma^* J \Delta(\sigma,\rho)^{-it} x \Delta(\sigma,\rho)^{it} \\
&\quad \times J V_\sigma J_0 \Delta(\sigma_{\mathcal{N}},\rho_{\mathcal{N}})^{-it} \left| h_{\rho_{\mathcal{N}}}^{1/2} \right\rangle \\
&= \left\langle h_{\rho_{\mathcal{N}}}^{1/2} \right| \alpha_{-t}^{\sigma_{\mathcal{N}}} \circ E_\sigma \circ \alpha_t^{\sigma}(x) \left| h_{\rho_{\mathcal{N}}}^{1/2} \right\rangle
\end{aligned}$$

$$= \left\langle h_{\rho_\mathcal{N}}^{1/2} \right| E_\sigma^{-t}(x) \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle$$
$$= \rho_\mathcal{N} \circ E_\sigma^{-t}(x).$$

For the third assertion, note that we have shown in lemma 5.12 that

$$\left\langle h_{\omega_\mathcal{N}}^{1/2} \right| V_\rho^* J x J V_\rho \left| h_{\omega_\mathcal{N}}^{1/2} \right\rangle = \mathrm{tr}\,(x h_{\omega_\mathcal{N}\circ E_\rho}) = \left\langle h_{\omega_\mathcal{N}\circ E_\rho}^{1/2} \right| JxJ \left| h_{\omega_\mathcal{N}\circ E_\rho}^{1/2} \right\rangle \ .$$

Then we have

$$V_\rho h_{\omega_\mathcal{N}}^{1/2} = u h_{\omega_\mathcal{N}\circ E_\rho}^{1/2} \ .$$

for some unitary u in $\mathcal{M}$. For ease of notation, we write $\Delta_\mathcal{N} = \Delta(\sigma_\mathcal{N}, \omega_\mathcal{N})$ and $\Delta_\mathcal{M} = \Delta(\sigma, \omega_\mathcal{N} \circ E_\rho)$. Then for $a \in \mathcal{M}$

$$\begin{aligned}
\Delta_\mathcal{M}^{\frac{1}{2}} V_\rho \Delta_\mathcal{N}^{-\frac{1}{2}} \left| h_{\sigma_\mathcal{N}}^{1/2} a \right\rangle &= \Delta_\mathcal{M}^{\frac{1}{2}} V_\rho J \Delta_\mathcal{N}^{\frac{1}{2}} \left| a^* h_{\sigma_\mathcal{N}}^{1/2} \right\rangle = \Delta_\mathcal{M}^{\frac{1}{2}} V_\rho \left| a h_{\omega_\mathcal{N}}^{1/2} \right\rangle \\
&= \Delta_\mathcal{M}^{\frac{1}{2}} a V_\rho \left| h_{\omega_\mathcal{N}}^{1/2} \right\rangle \\
&= \Delta_\mathcal{M}^{\frac{1}{2}} \left| a u h_{\omega_\mathcal{N}\circ E_\rho}^{1/2} \right\rangle = \left| h_\sigma^{1/2} a u \right\rangle = J u^* V_\sigma J_0 \left| h_{\sigma_\mathcal{N}}^{1/2} a \right\rangle .
\end{aligned}$$

Thus we have shown that

$$J u^* V_\sigma J_0 = \Delta_\mathcal{M}^{\frac{1}{2}} V_\rho \Delta_\mathcal{N}^{-\frac{1}{2}},$$

where u is the unitary from the polar decomposition of $V_\rho h_{\omega_\mathcal{N}}^{1/2}$. Then

$$\begin{aligned}
\langle c_t | \, x \, | c_t \rangle &= \left\langle h_{\rho_\mathcal{N}}^{1/2} \right| \Delta_\mathcal{N}^{-1/2+it} V_\rho^* \Delta_\mathcal{M}^{1/2-it} x \Delta_\mathcal{M}^{1/2+it} V_\rho \Delta_\mathcal{N}^{-1/2-it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle \\
&= \left\langle h_{\rho_\mathcal{N}}^{1/2} \right| \Delta_\mathcal{N}^{it} J_0 V_\sigma^* J J u J \Delta_\mathcal{M}^{-it} x \Delta_\mathcal{M}^{+it} J u^* J J V_\sigma J_0 \Delta_\mathcal{N}^{-it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle \\
&= \left\langle h_{\rho_\mathcal{N}}^{1/2} \right| \Delta_\mathcal{N}^{it} J_0 V_\sigma^* J J u J \alpha_t^\sigma(x) J u^* J J V_\sigma J_0 \Delta_\mathcal{N}^{-it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle \\
&= \left\langle h_{\rho_\mathcal{N}}^{1/2} \right| \Delta_\mathcal{N}^{it} J_0 V_\sigma^* J \alpha_t^\sigma(x) J V_\sigma J_0 \Delta_\mathcal{N}^{-it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle \\
&= \left\langle h_{\rho_\mathcal{N}}^{1/2} \right| \Delta_\mathcal{N}^{it} E_\sigma \circ \alpha_t^\sigma(x) \Delta_\mathcal{N}^{-it} \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle \\
&= \left\langle h_{\rho_\mathcal{N}}^{1/2} \right| \alpha_{-t}^{\sigma_\mathcal{N}} \circ E_\sigma \circ \alpha_t^\sigma(x) \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle \\
&= \left\langle h_{\rho_\mathcal{N}}^{1/2} \right| E_{-\sigma}^t(x) \left| h_{\rho_\mathcal{N}}^{1/2} \right\rangle \\
&= \rho_\mathcal{N} \circ E_\sigma^{-t}(x).
\end{aligned}$$

□

Now we can recover the estimate in lemmas 4.2, 4.10, and 4.18.

Lemma 5.14. *Define*

$$|w_\lambda\rangle := (\Delta(\sigma, \rho) + \lambda)^{-1} \left| h_\rho^{1/2} \right\rangle - V_\rho (\Delta(\sigma_\mathcal{N}, \rho_\mathcal{N}) + \lambda)^{-1} \left| h_{\rho_\mathcal{N}}^{\frac{1}{2}} \right\rangle \ ,$$

$$|u_\lambda\rangle := (\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}})) + \lambda)^{-1} \left|h_\rho^{1/2}\right\rangle - V_\rho(\Delta_{\mathcal{N}}(\sigma_{\mathcal{N}}, \omega_{\mathcal{N}}) + \lambda)^{-1} \left|h_{\rho_{\mathcal{N}}}^{1/2}\right\rangle \ ,$$

and

$$\begin{aligned}
|w_t\rangle &:= -\frac{\cosh(\pi t)}{\pi}\left(\int_0^\infty \lambda^{1/2+it}\,|w_\lambda\rangle\ \mathrm{d}\lambda\right),\\
|v_t\rangle &:= \frac{\cosh(\pi t)}{\pi}\Delta_{\mathcal{M}}^{\frac{1}{2}+it}\int_0^\infty \lambda^{-\frac{1}{2}-it}\,|w_\lambda\rangle\ \mathrm{d}\lambda,\\
|u_t\rangle &:= \frac{\cosh(\pi t)}{\pi}\Delta_{\mathcal{M}}(\sigma, R_\rho(\omega_{\mathcal{N}}))^{1/2+it}\int_0^\infty \lambda^{-1/2-it}\,|u_\lambda\rangle\ \mathrm{d}\lambda.
\end{aligned}$$

Then the following inequalities hold

$$\begin{aligned}
\left\|\sigma - \sigma_{\mathcal{N}}\circ E_\rho^t\right\|_1 &\leqslant 2\||w_t\rangle\|_2,\\
\left\|\rho - \rho_{\mathcal{N}}\circ E_\sigma^{-t}\right\|_1 &\leqslant 2\||v_t\rangle\|_2,\\
\left\|\rho - \rho_{\mathcal{N}}\circ E_\sigma^{-t}\right\|_1 &\leqslant 2\||u_t\rangle\|_2.
\end{aligned}$$

Proof. For ease of notation, we write $\Delta_{\mathcal{M}} := \Delta(\rho,\sigma)$ and $\Delta_{\mathcal{N}} := \Delta(\rho_{\mathcal{N}},\sigma_{\mathcal{N}})$. As in the finite-dimensional case,

$$\begin{aligned}
|w_t\rangle &= \Delta_{\mathcal{M}}^{1/2+it}\left|h_\rho^{1/2}\right\rangle - V_\rho\Delta_{\mathcal{N}}^{1/2+it}\left|h_{\rho_{\mathcal{N}}}^{1/2}\right\rangle\\
&= \Delta_{\mathcal{M}}^{it}\left(\Delta_{\mathcal{M}}^{1/2}\left|h_\rho^{1/2}\right\rangle - \Delta_{\mathcal{M}}^{-it}V_\rho\Delta_{\mathcal{N}}^{1/2+it}\left|h_{\rho_{\mathcal{N}}}^{1/2}\right\rangle\right)\\
&= \Delta_{\mathcal{M}}^{it}\left(\left|h_\sigma^{1/2}\right\rangle - J\,|a_t\rangle\right),\\
|v_t\rangle &= \left|h_\rho^{1/2}\right\rangle - \Delta_{\mathcal{M}}^{\frac{1}{2}+it}V_\rho\Delta_{\mathcal{N}}^{-1/2-it}\left|h_{\rho_{\mathcal{N}}}^{1/2}\right\rangle = \left|h_\rho^{1/2}\right\rangle - |b_t\rangle\,,\\
|u_t\rangle &= \left|h_\rho^{1/2}\right\rangle - \Delta(\sigma,\omega_{\mathcal{N}}\circ E_\rho)^{\frac{1}{2}+it}V_\rho\Delta(\sigma_{\mathcal{N}},\omega_{\mathcal{N}})^{-\frac{1}{2}-it}\left|h_{\rho_{\mathcal{N}}}^{1/2}\right\rangle = \left|h_\rho^{1/2}\right\rangle - |c_t\rangle\,.
\end{aligned}$$

For the first one, by $J\left|h_\sigma^{1/2}\right\rangle = \left|h_\sigma^{1/2}\right\rangle$ and lemma 5.13

$$2\||w_t\rangle\|_2 = 2\left\|J\left|h_\sigma^{1/2}\right\rangle - J\,|a_t\rangle\right\|_2 = 2\left\|\left|h_\sigma^{1/2}\right\rangle - |a_t\rangle\right\|_2 \geqslant \left\|\sigma - \sigma_{\mathcal{N}}\circ E_\rho^t\right\|_1\,,$$

where we have used the inequality in (25). This inequality remains valid in Haagerup L_p-spaces since its proof in [CV20a, lemma 2.2] only uses Hölder's inequality. The other two assertions follow similarly. □

Based on the lemma above, the rest of the argument is identical to that given for lemmas 4.3, 4.11, and 4.19, which estimate the Hilbert-space norm of $|v_t\rangle\,, |w_t\rangle$, and $|u_t\rangle$, respectively. In particular, the argument of lemmas 4.3, 4.11, and 4.19 implies the integral expression of $|v_t\rangle\,, |w_t\rangle$, and $|u_t\rangle$ converges absolutely if $Q_f(\rho\|\sigma)$ and $\widetilde{Q}_f(\rho\|\sigma)$ are finite for some regular f.

We now state our recovery results for quantum channels on general von Neumann algebras. Recall that we denote D as the standard relative entropy and $\widetilde{D}_\alpha$ as the α-sandwiched Rényi relative entropy. The maps $R^t_{\Phi,\sigma}$ and $R^t_{\Phi,\rho}$ are the rotated Petz maps defined in (87).

Theorem 5.15. *Let $\Phi : \mathcal{N} \to \mathcal{M}$ be a normal unital completely positive map. Let $\rho, \sigma \in D(\mathcal{M})$ be two states and denote $\rho_0 = \rho \circ \Phi, \sigma_0 = \sigma \circ \Phi$. Suppose $s(\rho) \leqslant s(\sigma)$. For $t \in \mathbb{R}$,*

(a) *If $s(\rho) = s(\sigma)$ and $Q_{x^2}(\rho\|\sigma) < \infty$,*

$$D(\rho \,\|\sigma) - D(\rho_0\|\sigma_0) \geqslant \left(\frac{\pi}{8\,\cosh(\pi t)}\right)^4 Q_{x^2}(\rho\|\sigma)^{-1}\left\|\sigma - \sigma_0 \circ R^t_{\Phi,\rho}\right\|_1^4 .$$

(b) *If $Q_{x^{-1}}(\rho\|\sigma) < \infty$, then for all $\varepsilon \in (0, 1/2)$,*

$$D(\rho \,\|\sigma) - D(\rho_0\|\sigma_0) \geqslant \left(K(Q_{x^{-1}}(\rho\|\sigma), \varepsilon)\frac{\pi}{2\,\cosh(\pi t)}\left\|\rho - \rho_0 \circ R^t_{\Phi,\sigma}\right\|_1\right)^{\frac{1}{1/2-\varepsilon}} .$$

(c) *If $\widetilde{Q}_\infty(\rho \,\|\sigma) = \inf\{\lambda|\rho \leqslant \lambda\sigma\} < \infty$, then for all $\alpha \in (1/2, 1)$, $\alpha' = \alpha/(\alpha-1)$, and $\varepsilon \in (0, (1 - 1/|\alpha'|)/2)$,*

$$\widetilde{D}_\alpha(\rho \,\|\sigma) - \widetilde{D}_\alpha(\rho_0\|\sigma_0) \geqslant |\alpha'| \log\left(1 + \left(K(\alpha, \varepsilon, \widetilde{Q}_\infty(\rho\|\sigma))\right.\right.$$
$$\left.\left.\times \frac{\pi}{2\,\cosh\,\pi t}\left\|\rho - \rho_0 \circ R^t_{\Phi,\sigma}\right\|_1\right)^{\frac{1}{\frac{1-1/|\alpha'|}{2}} - \varepsilon}\right) .$$

For all $\alpha > 1$, $\alpha' = \alpha/(\alpha-1)$, and $\varepsilon \in (0, (1 - 1/|\alpha'|)/2)$,

$$\widetilde{D}_\alpha(\rho \,\|\sigma) - \widetilde{D}_\alpha(\rho_0\|\sigma_0) \geqslant \alpha' \,\log\left(1 + \frac{1}{\widetilde{Q}_\infty(\rho\|\sigma)^{\frac{1}{\alpha'}}}\left(K(\alpha, \varepsilon, \widetilde{Q}_\infty(\rho\|\sigma))\right.\right.$$
$$\left.\left.\times \frac{\pi}{2\,\cosh\,\pi t}\left\|\rho - \rho_0 \circ R^t_{\Phi,\sigma}\right\|_1\right)^{\frac{1}{\frac{1-1/|\alpha'|}{2}} - \varepsilon}\right) .$$

In the inequalities above, $K(Q_{x^{-1}}(\rho\|\sigma), \varepsilon)$ and $K(\alpha, \varepsilon, \widetilde{Q}_\infty(\rho\|\sigma)))$ are constants defined as in (60) and (75), respectively.

Proof. Note that the assumption $s(\rho) = s(\sigma)$ is equivalent to ρ and σ being faithful because we can always restrict our considerations to $s(\sigma)\mathcal{M}s(\sigma)$, as mentioned in remark 5.9. The faithfulness is needed for lemma 5.12 where we used the identity $V^*_\rho\Delta(\sigma, \rho)V_\rho = \Delta(\sigma_\mathcal{N}, \rho_\mathcal{N})$ for the estimates in (a). For (b) and (c), we first obtain the faithful cases by the lemma 5.12 and 5.13. For the general case of $s(\rho) \leqslant s(\sigma)$, we use the approximation in lemma 5.6. Indeed, take $\rho_\varepsilon = (1 - \varepsilon)\rho + \varepsilon\sigma$ and $\rho_{0,\varepsilon} = (1 - \varepsilon)\rho_0 + \varepsilon\sigma_0$. Then $s(\rho_\varepsilon) = s(\sigma)$, $\rho_{0,\varepsilon} = \rho_\varepsilon \circ \Phi$ and moreover

$$\lim_{\varepsilon\to 0} D(\rho_\varepsilon\|\sigma) = D(\rho\|\sigma) \text{ , } \lim_{\varepsilon\to 0} D(\rho_{0,\varepsilon}\|\sigma) = D(\rho_0\|\sigma_0) \text{ .}$$

Then the estimate follows the faithful cases and $\left\|\rho - \rho_0 \circ R^t_{\Phi,\sigma}\right\|_1 = \lim_{\varepsilon\to 0} \left\|\rho_\varepsilon - \rho_{0,\varepsilon} \circ R^t_{\Phi,\sigma}\right\|_1$. The argument for (c) is similar. □

We have the following corollary regarding reversibility:

Corollary 5.16. *Let $\Phi : \mathcal{N} \to \mathcal{M}$ be a normal unital completely positive map. Let $\rho, \sigma \in D(\mathcal{M})$ be two states, and let $\rho_0 := \rho \circ \Phi$ and $\sigma_0 := \sigma \circ \Phi$. Suppose $s(\rho) \leqslant s(\sigma)$ and $D(\rho\|\sigma) < \infty$. The following are equivalent:*

(*a*) $D(\rho\|\sigma) = D(\rho_0\|\sigma_0) < \infty$.
(*b*) $\widetilde{D}_\alpha(\rho\|\sigma) = \widetilde{D}_\alpha(\rho_0\|\sigma_0) < \infty$ *for some* $\alpha \in (1/2, 1) \cup (1, \infty)$.
(*c*) $\widetilde{Q}_f(\rho\|\sigma) = \widetilde{Q}_f(\rho_0\|\sigma_0)$ *for some regular operator anti-monotone function* f.
(*d*) $\widetilde{Q}_f(\rho\|\sigma) = \widetilde{Q}_f(\rho_0\|\sigma_0)$ *for all operator anti-monotone functions* f.
(*e*) *There exists a normal UCP map* $\Phi : \mathcal{M} \to \mathcal{N}$ *such that* $\rho_0 \circ \Phi = \rho$ *and* $\sigma_0 \circ \Phi = \sigma$.
(*f*) $\rho_0 \circ R^t_{\Phi,\sigma} = \rho$ *for all* t.
(*g*) $\sigma_0 \circ R^t_{\Phi,\rho} = \sigma$ *for all* t.

Proof. We argue for the subalgebra case $\Phi = \iota : \mathcal{N} \hookrightarrow \mathcal{M}$. (f), (g) $\Rightarrow$ (e) is trivial. Note that by the monotonicity $\alpha \mapsto \widetilde{D}_\alpha$, $D(\rho\|\sigma) < \infty$ implies $\widetilde{D}_\alpha(\rho\|\sigma) < \infty$. Then (e) $\Rightarrow$ (a)–(d) by data processing inequality. Also (d) $\Rightarrow$ (c) is trivial and (c) $\Rightarrow$ (a) by example 5.5. (a) implies $|w_\lambda\rangle = 0$ for $\lambda > 0$, which by lemma 5.14 further implies (f) and (g). (b) $\Rightarrow$ (e) is proved in [Jen17b, theorem 5.1]. □

6. Conclusion

In summary, we have established physically meaningful remainder terms for the data-processing inequality for the optimized f-divergence, and we have improved upon prior results like this for the standard f-divergence. As a consequence, we have established the first physically meaningful remainder terms for the data-processing inequality for the sandwiched Rényi relative entropy. Finally, we generalized all of our results to the von Neumann algebraic setting of the optimized f-divergence, by suitably generalizing its definition, its data-processing inequality, and refinements to this setting.

Going foward from here, we consider it to be a great challenge to establish universal remainder terms for the data-processing inequalities of the standard and optimized f-divergences, in the sense of [JRS+18]. Such results would significantly extend the domain of applicability of these refined data-processing inequalities.

Acknowledgments

We thank Anna Vershynina for feedback on our paper. This project was initiated during the workshop ‘Algebraic and Statistical Ways into Quantum Resource Theories’, which took place during the summer of 2019 at the Banff International Research Station (BIRS). The authors thank the Banff International Research Station for Mathematical Innovation and Discovery for the kind hospitality during the workshop. MMW acknowledges support from the National Science Foundation under Grant No. 1714215, from Stanford QFARM, and from AFOSR under Grant No. FA9550-19-1-03.

Data availability statement

No new data were created or analysed in this study.

Appendix A. Preliminaries on von Neumann algebras

In this appendix, we briefly review some of the von Neumann algebra theory used in section 5. We refer to the classic texts [Tak79, Tak03] for more information on von Neumann algebras and to [OP04] for a similar introduction related to quantum divergences.

A.1. Spatial derivative and relative modular operator

Let $\mathcal{M} \subset B(H)$ be a von Neumann algebra. A linear functional $\phi : \mathcal{M} \to \mathbb{C}$ is

(a) **Normal** if it is weak*-continuous;
(b) **Positive** if $\phi(x^*x) \geqslant 0$, $\forall\ x \in \mathcal{M}$;
(c) **Unital** if $\phi(1) = 1$;
(d) A **state** if ϕ is positive and unital.

The predual $\mathcal{M}_*$ of $\mathcal{M}$ is the space of all normal linear functionals. We denote by $\mathcal{M}_*^+$ the set of all normal positive linear functionals and by $D(\mathcal{M})$ the set of all normal states. A positive normal linear functional ϕ is **faithful** if $\phi(x^*x) = 0$ implies $x = 0$. A von Neumann algebra is σ-finite if it admits a normal faithful state. For $\phi \in \mathcal{M}_*^+$, its support $s(\phi)$ is the smallest projection $e \in \mathcal{M}$ such that $\phi(e) = \phi(1)$. We say that $\pi : \mathcal{M} \to B(H)$ is a $*$-representation if π is a normal $*$-homomorphism (not necessarily unital). We say that the vector $\boldsymbol{\phi} \in H$ implements $\phi \in \mathcal{M}_*^+$ via π if for all $x \in \mathcal{M}$,

$$\phi(x) = \langle \boldsymbol{\phi}|\, \pi(x)\, |\boldsymbol{\phi}\rangle \ .$$

We typically use Greek letters ρ, σ, ϕ, ψ to denote states and linear functionals, and boldface letters $\boldsymbol{\rho}, \boldsymbol{\sigma}, \boldsymbol{\phi}, \boldsymbol{\psi}$ to denote vectors implementing the corresponding states. Let G_ϕ be the Hilbert space completion of $\mathcal{M}$ with respect to the ϕ-inner product:

$$\langle x, y\rangle_\phi = \phi(x^*y) \ .$$

Let $\boldsymbol{\eta}_\phi(x)$ (resp. $\boldsymbol{\eta}_\phi$) be the vector corresponding to $x \in \mathcal{M}$ (resp. identity 1). The GNS representation $\pi_\phi : \mathcal{M} \to B(G_\phi)$ is the normal $*$-homomorphism given by

$$\pi_\sigma(a)\boldsymbol{\eta}_\phi(x) = \boldsymbol{\eta}_\phi(ax) \ .$$

In particular, $\boldsymbol{\eta}_\phi$ implements ϕ via π_ϕ. Letting $\boldsymbol{\phi} \in H$ be a vector implementing ϕ via $\pi : \mathcal{M} \to B(H)$, we can define the isometry $V : G_\phi \to H$ as follows:

$$V(\pi_\phi(x)\boldsymbol{\eta}_\phi) = \pi(x)\boldsymbol{\phi} \ .$$

We denote $[\pi(\mathcal{M})\boldsymbol{\phi}]$ as the closure of $\pi(\mathcal{M})\boldsymbol{\phi} \subset H$ as a subspace, and with slight abuse of notation, also identify it as the projection onto $[\pi(\mathcal{M})\boldsymbol{\phi}]$. Thus $G_\phi \cong [\pi(\mathcal{M})\boldsymbol{\phi}]$ for all $\boldsymbol{\phi}$ implementing ϕ.

Let $\mathcal{M} \subset B(H)$ be a von Neumann algebra acting on H, and let

$$\mathcal{M}' := \{x \in B(H) \mid xa = ax \quad \forall\ a \in \mathcal{M}\}$$

be its commutant. For a vector $\boldsymbol{\phi} \in H$, we denote by $\phi \in \mathcal{M}_*$ and $\phi' \in (\mathcal{M}')_*$ the corresponding states implemented on $\mathcal{M}$ and $\mathcal{M}'$. The support projections are given by

$$s_{\mathcal{M}}(\boldsymbol{\phi}) := s(\phi) = [\mathcal{M}'\boldsymbol{\phi}] \in \mathcal{M}\ , \qquad s_{\mathcal{M}'}(\boldsymbol{\phi}) := s(\phi') = [\mathcal{M}\boldsymbol{\phi}] \in \mathcal{M}'.$$

Given two vectors $\boldsymbol{\phi}, \boldsymbol{\psi} \in H$, we define the anti-linear operator $S_{\psi,\phi}$ as follows:

$$S_{\psi,\phi}(a\boldsymbol{\phi} + \boldsymbol{\eta}) = s(\boldsymbol{\phi})a^*\boldsymbol{\psi}\ . \quad a \in \mathcal{M}\ , \tag{88}$$

where $a\boldsymbol{\phi} \in [\mathcal{M}\boldsymbol{\phi}], \boldsymbol{\eta} \in [\mathcal{M}\boldsymbol{\phi}]^\perp$. Then $S_{\psi,\phi}$ is a closable operator, and the relative modular operator is the positive self-adjoint operator defined as

$$\Delta(\boldsymbol{\psi}, \boldsymbol{\phi}) := (S_{\psi,\phi})^* \bar{S}_{\psi,\phi}\ , \tag{89}$$

where $\bar{S}_{\psi,\phi}$ is the closure of $S_{\psi,\phi}$. For $a\boldsymbol{\phi} \in \mathcal{M}\boldsymbol{\phi}$,

$$\langle a\boldsymbol{\phi}|\, \Delta(\boldsymbol{\psi}, \boldsymbol{\phi})\, |a\boldsymbol{\phi}\rangle = \langle \boldsymbol{\psi}|\, as(\boldsymbol{\phi})a^*\, |\boldsymbol{\psi}\rangle\ . \tag{90}$$

and the support $\mathrm{supp}(\Delta(\boldsymbol{\psi}, \boldsymbol{\phi})) = s(\psi)s(\phi')$.

We also recall the spatial derivative. Given $\phi \in \mathcal{M}_*^+$, define the lineal of ϕ as the subspace

$$H_\phi = \{\boldsymbol{\xi} \in H \mid \|a\boldsymbol{\xi}\|_H^2 \leqslant C\phi(a^*a)\ \forall\ \ a \in \mathcal{M}\ , \quad \text{for some } C \geqslant 0\}\ .$$

The closure $\overline{H_\phi} = s(\phi)H$. For $\boldsymbol{\xi} \in H_\phi$, we define the bounded operator $R_\phi(\boldsymbol{\xi}) : G_\phi \to H$ as follows:

$$R_\phi(\boldsymbol{\xi})\boldsymbol{\eta}_\phi(x) = x\boldsymbol{\xi}\ . \tag{91}$$

Then $R_\phi(\boldsymbol{\xi})\pi_\phi(a) = aR_\phi(\boldsymbol{\xi})$, which implies $R_\phi(\boldsymbol{\xi})R_\phi(\boldsymbol{\xi})^* \in \mathcal{M}'$. For a vector $\boldsymbol{\psi} \in H$, the spatial derivative $\Delta(\boldsymbol{\psi}/\phi)$ is the positive self-adjoint operator on H_ϕ defined by

$$\langle \boldsymbol{\xi}|\, \Delta(\boldsymbol{\psi}/\phi)\, |\boldsymbol{\xi}\rangle := \langle \boldsymbol{\psi}|\, R_\phi(\boldsymbol{\xi})R_\phi(\boldsymbol{\xi})^*\, |\boldsymbol{\psi}\rangle\ .$$

We can write $\Delta(\boldsymbol{\psi}/\phi) = \Delta(\psi'/\phi)$ because it only depends on $\psi' \in (\mathcal{M}')_*^+$ implemented by $\boldsymbol{\psi}$. The connection to the relative modular operator is given by

$$\Delta(\boldsymbol{\psi}, \boldsymbol{\phi}) = \Delta(\psi/\phi'),$$

where $\psi \in (\mathcal{M}_*)^+$ is implemented by $\boldsymbol{\psi}$ and $\phi' \in (\mathcal{M}'_*)^+$ implemented by $\boldsymbol{\phi}$. Indeed, $R_{\phi'}(a\boldsymbol{\phi}) = aR_{\phi'}(\boldsymbol{\phi})$ for $a \in \mathcal{M}$ and $R_{\phi'}(\boldsymbol{\phi})R_{\phi'}(\boldsymbol{\phi})^* = [\mathcal{M}'\boldsymbol{\phi}] = s(\phi) \in \mathcal{M}$. Then for $a\boldsymbol{\phi} \in \mathcal{M}\boldsymbol{\phi}$,

$$\langle a\boldsymbol{\phi}|\, \Delta(\psi/\phi')\, |a\boldsymbol{\phi}\rangle = \psi(R_{\phi'}(a\boldsymbol{\phi})R_{\phi'}(a\boldsymbol{\phi})^*) = \psi(as(\boldsymbol{\phi})a^*)\ ,$$

which coincides with (90). Thus we verify that $\Delta(\boldsymbol{\psi}, \boldsymbol{\phi}) = \Delta(\psi/\phi')$ for all $\boldsymbol{\psi}, \boldsymbol{\phi} \in H$.

The relative modular operator $\Delta(\boldsymbol{\psi}, \boldsymbol{\phi})$ is independent of vector representations up to isometry. Let ϕ and ψ be two normal states of $\mathcal{M}$. Let $\pi_1 : \mathcal{M} \to B(H_1)$ (resp. $\pi_2 : \mathcal{M} \to B(H_2)$) be a representation, and suppose that $\boldsymbol{\phi}_1, \boldsymbol{\psi}_1 \in H_1$ (resp. $\boldsymbol{\phi}_2, \boldsymbol{\psi}_2 \in H_2$) implement ϕ and ψ via π_1 (resp. π_2). Define the partial isometries $V_\phi : H_1 \to H_2$ and $V_\psi : H_1 \to H_2$ as follows:

$$\begin{aligned} &V_\phi(\pi_1(a)\boldsymbol{\phi}_1 + \boldsymbol{\eta}) = \pi_2(a)\boldsymbol{\phi}_2\ , \\ &V_\psi(\pi_1(a)\boldsymbol{\psi}_1 + \boldsymbol{\zeta}) = \pi_2(a)\boldsymbol{\psi}_2\ ,\ \forall\ \ a \in \mathcal{M}, \end{aligned} \tag{92}$$

where $\eta \in [\pi_1(\mathcal{M})\phi_1]^\perp$ and $\zeta \in [\pi_1(\mathcal{M})\psi_1]^\perp$. Let S_{ψ_1,ϕ_1} and $\Delta(\psi_1,\phi_1)$ (resp. S_{ψ_2,ϕ_2} and $\Delta(\psi_2,\phi_2)$) be the operators defined in (88) and (89) for $\pi_1(\mathcal{M})$ (resp. $\pi_2(\mathcal{M})$). Note that $\pi_1(s(\phi)) = s(\phi_1), \pi_2(s(\phi)) = s(\phi_2)$ and $V_\psi^* V_\psi = s(\psi_1) \supset \mathrm{Ran}(S_{\psi_1,\phi_1})$. We have

$$S_{\psi_2,\phi_2} V_\phi = V_\psi S_{\psi_1,\phi_1}, \qquad \Delta(\psi_1,\phi_1) = V_\phi^* \Delta(\psi_2,\phi_2) V_\phi.$$

A.2. Standard form of von Neumann algebras

The theory of the standard form of von Neumann algebras was developed by Araki [Ara74], Connes [Con76], and Haagerup [Haa76]. Recall that the standard form $(\mathcal{M}, H, J, P)$ of a von Neumann algebra $\mathcal{M}$ is given by an injective $*$-homomorphism $\pi : \mathcal{M} \to B(H)$, an anti-linear isometry J on H, and a self-dual cone P such that

(a) $J^2 = 1, J\mathcal{M}J = \mathcal{M}'$,
(b) $JaJ = a^*$ for $a \in \mathcal{M} \cap \mathcal{M}'$,
(c) $J\xi = \xi$ for $\xi \in P$,
(d) $aJaJP = P$ for $a \in \mathcal{M}$.

The standard form is unique up to unitary equivalence. For each normal state $\phi \in \mathcal{M}_*^+$, there exists a unique unit vector $\xi_\phi \in P$ implementing ϕ. We write the standard form of the relative modular operator as

$$\Delta(\phi,\psi) := \Delta(\xi_\phi, \xi_\psi) \ .$$

By the symmetric role of $\mathcal{M}$ and $\mathcal{M}'$, we have

$$\Delta(\phi,\psi) = J\Delta(\psi,\phi)^{-1}J. \tag{93}$$

In particular, the modular operator of ϕ is $\Delta(\phi,\phi)$ and the modular automorphism group $\alpha_t^\phi : \mathcal{M} \to \mathcal{M}$ is as follows:

$$\alpha_t^\phi(x) = \Delta(\phi,\phi)^{-it} x \Delta(\phi,\phi)^{it}.$$

When $\mathcal{M}$ is semifinite equipped with a normal faithful semi-finite trace τ, the standard form is basically given by the GNS construction. Define the τ-inner product and L_2-norm respectively as

$$\langle a, b\rangle = \tau(a^* b) \ , \qquad \|a\|_2^2 = \langle a, a\rangle \ .$$

The L_2-space $L_2(\mathcal{M})$ is a Hilbert space as the norm completion of $\{a \in \mathcal{M} \mid \tau(s(|a|)) < \infty\}$, where $s(|a|)$ is the support of $|a|$. The GNS representation $\pi : \mathcal{M} \to L_2(\mathcal{M})$ has the following action for all $x \in \mathcal{M}$:

$$\pi(x)a = xa.$$

This gives a standard form $(\mathcal{M}, L_2(\mathcal{M}), J, L_2(\mathcal{M})^+)$, where the anti-linear isometry is $Ja = a^*$ and $L_2(\mathcal{M})^+$ is the positive cone in L_2.

A.3. Haagerup L_p-spaces

In this part, we briefly review Haagerup's L_p-space [Haa79] as our tool to section 5.4. We refer to [Ter82] and [Jen18, appendix] for more details on this topic.

Let $\mathcal{M} \subset B(H)$ be a von Neumann algebra acting on a Hilbert space H. Given a distinguished normal faithful state $\omega \in D(\mathcal{M})$, we denote by $\alpha_t := \alpha_t^\omega : \mathcal{M} \to \mathcal{M}, t \in \mathbb{R}$ the one parameter modular automorphism group. The crossed product

$$\mathcal{R} = \mathcal{M} \rtimes_\alpha \mathbb{R}$$

is the von Neumann algebra acting on $L_2(\mathbb{R}, H)$, generated by the operator $\pi(x), x \in \mathcal{M}$, and the operator $\lambda(s), s \in \mathbb{R}$, defined as follows: for all $\xi \in L_2(\mathbb{R}, H)$ and $t \in \mathbb{R}$

$$\pi(x)(\xi)(t) := \alpha_{-t}(x)\xi(t) , \qquad \lambda(s)(\xi)(t) = \xi(t-s) .$$

Note that π is a normal faithful representation of $\mathcal{M}$ on $H \otimes_2 L_2(\mathbb{R}) \cong L_2(\mathbb{R}, H)$ and $(\pi, \lambda(s))$ gives a covariant representation such that $\alpha_t(x) = \lambda(t)x\lambda(t)^*, x \in \mathcal{M}, t \in \mathbb{R}$. The dual action $\hat{\alpha}_t$ of $\mathbb{R}$ on $\mathcal{R}$ is a one-parameter automorpshim group of $\mathbb{R}$ on $\mathcal{R}$, implemented by the unitary representation $\{W(t)\}_{t\in\mathbb{R}}$ on $L_2(\mathbb{R}, H)$,

$$\hat{\alpha}_t(x) = W(t)xW(t)^*,$$

where

$$W(t)(\xi)(s) = e^{-its}\xi(s), \quad \xi \in L_2(\mathbb{R}, H),\ t, s \in \mathbb{R} .$$

The dual action $\hat{\alpha}$ satisfies (and is uniquely determined by)

$$\hat{\alpha}_t(x) = x, \qquad \hat{\alpha}_t(\lambda(s)) = e^{-ist}\lambda(s) , \quad x \in \mathcal{M},\ s, t \in \mathbb{R} ,$$

and $\mathcal{M} = \{x \in \mathcal{R} \mid \hat{\alpha}_t(x) = x , \forall\ t \in \mathbb{R}\}$. This cross product algebra $\mathcal{R}$ admits a normal faithful semi-finite trace τ satisfying

$$\tau \circ \hat{\alpha}_t = e^{-t}\tau, \quad \forall\ t \in \mathbb{R}.$$

For $0 < p \leqslant \infty$, the Haagerup noncommutative L_p-space is then defined as

$$L_p(\mathcal{M}, \omega) = \{x \in L_0(\mathcal{R}, \tau)\ : \hat{\alpha}_t = e^{-t/p}x, \quad \forall\ t \in \mathbb{R}\} .$$

We will suppress 'ω' in the notation $L_p(\mathcal{M})$ since the L_p-spaces constructed for different states are isomorphic. The positive part is $L_p(\mathcal{M})_+ = L_p(\mathcal{M}) \cap L_0(\mathcal{R})_+$. For all $\phi \in \mathcal{M}_*^+$, there exists a Radon–Nikodym derivative $h_\phi \in L_1(\mathcal{R}, \tau)$ with respect to τ such that

$$\widetilde{\phi}(x) = \tau(h_\phi x), \quad x \in \mathcal{R}_+ , \qquad \hat{\alpha}_t(h_\phi) = e^{-t}h_\phi ,$$

where $\widetilde{\phi}$ is the dual weight of ϕ on $\mathcal{R}$. This gives a linear bijection

$$\phi \in \mathcal{M}_*^+ \Longleftrightarrow h_\phi \in L_1(\mathcal{M})^+.$$

This bijection further extends an identification $\phi \in \mathcal{M}_* \leftrightarrow h_\phi \in L_1(\mathcal{M})$ with the property $h_{x\phi y} = xh_\phi y, x, y \in \mathcal{M}$. Moreover, if $\phi = u|\phi|$ is the polar decomposition, $h_\phi = uh_{|\phi|}$. Using this linear bijection, the trace and L_1-norm on $L_1(\mathcal{M})$ is defined as

$$\mathrm{tr}\,(h_\phi) := \phi(1) ,\ \|h_\phi\|_1 := \mathrm{tr}\,(|h_\phi|) = \mathrm{tr}\,(h_{|\phi|}) = |\phi|(1) = \|\phi\|_{\mathcal{M}_*}.$$

For $a \in \mathcal{R}$, we have the polar decomposition $a = u|a|$ and for $p \in [1, \infty)$

$$a \in L_p(\mathcal{M}) \Longleftrightarrow |a| \in L_p(\mathcal{M}) \Longleftrightarrow |a|^p \in L_1(\mathcal{M}) .$$

which leads to the L_p-norm, defined as

$$\|a\|_{L_p(\mathcal{M})} = \mathrm{tr}\,(|a|^p)^{1/p}\,, \qquad \|a\|_\infty = \|a\|_{\mathcal{M}}.$$

For $a \in L_p(\mathcal{M}), b \in L_q(\mathcal{M})$ with $1/p + 1/q = 1$, $ab, ba \in L_1(\mathcal{M})$ and the trace 'tr' has the following tracial property:

$$\mathrm{tr}\,(ab) = \mathrm{tr}\,(ba)\,.$$

In particular, the L_2-space $L_2(\mathcal{M})$ is a Hilbert space with inner product $\langle a, b\rangle = \mathrm{tr}(a^*b)$. Define the left regular representation

$$\pi : \mathcal{M} \to B(L_2(\mathcal{M})), \qquad \pi(x)a = xa$$

and the anti-linear isometry

$$J : L_2(\mathcal{M}) \to L_2(\mathcal{M})\,, \quad Ja = a^*\,.$$

Identifying $\pi(\mathcal{M}) \cong \mathcal{M}$, the quadruple $(\mathcal{M}, L_2(\mathcal{M}), J, L_2(\mathcal{M})^+)$ is a standard form of $\mathcal{M}$. In particular, JMJ acts on $L_2(\mathcal{M})$ as the right multiplication $JxJa = ax^*$.

ORCID iDs

Li Gao https://orcid.org/0000-0002-9220-0119
Mark M Wilde https://orcid.org/0000-0002-3916-4462

References

[AC82] Accardi L and Cecchini C 1982 Conditional expectations in von Neumann algebras and a theorem of Takesaki *J. Funct. Anal.* **45** 245–73

[AM82] Araki H and Masuda T 1982 Positive cones and L_p-spaces for von Neumann algebras *Publ. Res. Inst. Math. Sci.* **18** 339–411

[Ara74] Araki H 1974 Some properties of modular conjugation operator of von Neumann algebras and a non-commutative Radon–Nikodym theorem with a chain rule *Pacific J. Math.* **50** 309–54

[AS66] Ali S M and Silvey S D 1966 A general class of coefficients of divergence of one distribution from another *J. Royal Stat. Soc.* B **28** 131–42

[BC20] Bluhm A and Capel Á 2020 A strengthened data processing inequality for the Belavkin–Staszewski relative entropy *Rev. Math. Phys.* **32** 2050005

[Bha13] Bhatia R 2013 *Matrix Analysis* vol 169 (Berlin: Springer)

[BST18] Berta M, Scholz V B and Tomamichel M 2018 Rényi divergences as weighted non-commutative vector-valued L_p-spaces *Ann. Henri Poincare* **19** 1843–67

[CMW16] Cooney T, Mosonyi M and Wilde M M 2016 Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication *Commun. Math. Phys.* **344** 797–829

[Con76] Connes A 1976 Classification of injective factors cases II_1, II_∞, III_λ, $\lambda \neq 1$ *Ann. Math.* **104** 73–115

[Csi67] Csiszár I 1967 Information-type measures of difference of probability distributions and indirect observation *Stud. Sci. Math. Hung.* **2** 229–318

[CV18] Carlen E A and Vershynina A 2018 Recovery and the data processing inequality for quasi-entropies *IEEE Trans. Inf. Theory* **64** 6929–38

[CV20a] Carlen E A and Vershynina A 2020 Recovery map stability for the data processing inequality *J. Phys. A: Math. Theor.* **53** 035204

[CV20b] Chehade S and Vershynina A 2020 Saturating the data processing inequality for α-z Rényi relative entropy (arXiv:2006.07726)

[Dat09] Datta N 2009 Min- and max-relative entropies and a new entanglement monotone *IEEE Trans. Inf. Theory* **55** 2816–26

[DW18] Ding D and Wilde M M 2018 Strong converse for the feedback-assisted classical capacity of entanglement-breaking channels *Probl. Inf. Transm.* **54** 1–19

[FHSW20] Faulkner T, Hollands S, Swingle B and Wang Y 2020 Approximate recovery and relative entropy I. General von Neumann subalgebras (arXiv:2006.08002)

[FR15] Fawzi O and Renner R 2015 Quantum conditional mutual information and approximate Markov chains *Commun. Math. Phys.* **340** 575–611

[GW15] Gupta M K and Wilde M M 2015 Multiplicativity of completely bounded p-norms implies a strong converse for entanglement-assisted capacity *Commun. Math. Phys.* **334** 867–87

[GYZ19] Gu J, Yin Z and Zhang H 2019 Interpolation of quasi noncommutative L_p-spaces (arXiv:1905.08491)

[Haa76] Haagerup U 1976 The standard form of von Neumann algebras *Math. Scand.* **37** 271–83

[Haa79] Haagerup U 1979 L_p-spaces associated with an arbitrary von Neumann algebra *Algebres d'Opérateurs et leurs Applications en Physique Mathématiqueournal of Functional Analysis* (Marseille: Proc. Colloq.) vol 32 pp 175–206 1977

[Hay07] Hayashi M 2007 Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding *Phys. Rev.* A **76** 062301

[Hia18] Hiai F 2018 Quantum f-divergences in von Neumann algebras. I. Standard f-divergences *J. Math. Phys.* **59** 102202

[Hia19] Hiai F 2019 Quantum f-divergences in von Neumann algebras. II. Maximal f-divergences *J. Math. Phys.* **60** 012203

[Hia21] Hiai F 2021 *Quantum F-divergences in Von Neumann Algebras: Reversibility of Quantum Operations* (Berlin: Springer)

[HM17] Hiai F and Mosonyi M 2017 Different quantum f-divergences and the reversibility of quantum operations *Rev. Math. Phys.* **29** 1750023

[HMPB11] Hiai F, Mosonyi M, Petz D and Bény C 2011 Quantum f-divergences and error correction *Rev. Math. Phys.* **23** 691–747

[Hol72] Kholevo A S 1972 On quasiequivalence of locally normal states *Theor. Math. Phys.* **13** 1071–82

[HP91] Hiai F and Petz D 1991 The proper formula for relative entropy and its asymptotics in quantum probability *Commun. Math. Phys.* **143** 99–114

[HP03] Hansen F and Pedersen G K 2003 Jensen's operator inequality *Bull. Math. Soc.* **35** 553–64

[Jen17a] Jenčová A 2017 Preservation of a quantum Rényi relative entropy implies existence of a recovery map *J. Phys. A: Math. Theor.* **50** 085303

[Jen17b] Jenčová A 2017 Rényi relative entropies and noncommutative L_p-spaces II (arXiv:1707.00047)

[Jen18] Jenčová A 2018 Rényi relative entropies and noncommutative L_p-spaces *Ann. Henri Poincare* **19** 2513–42

[JRS+18] Junge M, Renner R, Sutter D, Wilde M M and Winter A 2018 Universal recovery maps and approximate sufficiency of quantum relative entropy *Ann. Henri Poincare* **19** 2955–78

[KL51] Kullback S and Leibler R A 1951 On information and sufficiency *Ann. Math. Stat.* **22** 79–86

[Kom66] Komatsu H 1966 Fractional powers of operators *Pacific J. Math.* **19** 285–346

[Las19] Lashkari N 2019 Constraining quantum fields using modular theory *J. High Energy Phys.* JHEP01(2019)059

[Lin75] Lindblad G 1975 Completely positive maps and entropy inequalities *Commun. Math. Phys.* **40** 147–51

[LRD17] Leditzky F, Rouzé C and Datta N 2017 Data processing for the sandwiched Rényi divergence: a condition for equality *Lett. Math. Phys.* **107** 61–80

[MLDS+13] Müller-Lennert M, Dupuis F, Szehr O, Fehr S and Tomamichel M 2013 On quantum Rényi entropies: a new generalization and some properties *J. Math. Phys.* **54** 122203

[MO15] Mosonyi M and Ogawa T 2015 Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies *Commun. Math. Phys.* **334** 1617–48

[Mor63] Morimoto T 1963 Markov processes and the H-theorem *J. Phys. Soc. Japan* **18** 328–31

[Nag06] Nagaoka H 2006 The converse part of the theorem for quantum Hoeffding bound (arXiv:quant-ph/0611289)

[ON00] Ogawa T and Nagaoka H 2000 Strong converse and Stein's lemma in quantum hypothesis testing *IEEE Trans. Inf. Theory* **46** 2428–33

[OP04] Ohya M and Petz D 2004 *Quantum Entropy and Its Use* (Berlin: Springer)

[Pet85] Petz D 1985 Quasi-entropies for states of a von Neumann algebra *Publ. Res. Inst. Math. Sci.* **21** 787–800

[Pet86a] Petz D 1986 Quasi-entropies for finite quantum systems *Rep. Math. Phys.* **23** 57–65

[Pet86b] Petz D 1986 Sufficient subalgebras and the relative entropy of states of a von Neumann algebra *Commun. Math. Phys.* **105** 123–31

[Pet88] Petz D 1988 Sufficiency of channels over von Neumann algebras *Q. J. Math.* **39** 97–108

[Pis20] Pisier G 2020 *Tensor Products of C^*-Algebras and Operator Spaces: The Connes–Kirchberg Problem* vol 96 (Cambridge: Cambridge University Press)

[Rén61] Rényi A 1961 On measures of entropy and information *Proc. 4th Berkeley Symp. Mathematical Statistics and probability, Volume 1: Contributions to the Theory of Statistics* (The Regents of the University of California) pp 547–61

[SBT17] Sutter D, Berta M and Tomamichel M 2017 Multivariate trace inequalities *Commun. Math. Phys.* **352** 37–58

[Ser17] Serafini A 2017 *Quantum Continuous Variables* (Boca Raton, FL: CRC Press)

[SLW18] Seshadreesan K P, Lami L and Wilde M M 2018 Rényi relative entropies of quantum Gaussian states *J. Math. Phys.* **59** 072204

[STH16] Sutter D, Tomamichel M and Harrow A W 2016 Strengthened monotonicity of relative entropy via pinched Petz recovery map *IEEE Trans. Inf. Theory* **62** 2907–13

[Sti55] Stinespring W F 1955 Positive functions on C^*-algebras *Proc. Am. Math. Soc.* **6** 211–6

[Tak79] Takesaki M 1979 *Theory of Operator Algebras I* vol 124 (Berlin: Springer)

[Tak03] Takesaki M 2003 *Theory of Operator Algebras II* vol 125 (Berlin: Springer)

[Ter82] Terp M 1982 Interpolation spaces between a von Neumann algebra and its predual *J. Operator Theory* 327–60

[TWW16] Tomamichel M, Wilde M M and Winter A 2016 Strong converse rates for quantum communication *IEEE Trans. Inf. Theory* **63** 715–27

[Uhl76] Uhlmann A 1976 The transition probability in the state space of a^*-algebra *Rep. Math. Phys.* **9** 273–9

[Uhl77] Uhlmann A 1977 Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory *Commun. Math. Phys.* **54** 21–32

[Ume62] Umegaki H 1962 Conditional expectation in an operator algebra. IV. Entropy and information *Kodai Mathematical Seminar Reports* vol 14 59–85

[Ved02] Vedral V 2002 The role of relative entropy in quantum information theory *Rev. Mod. Phys.* **74** 197

[Ver19] Vershynina A 2019 On quantum quasi-relative entropy *Rev. Math. Phys.* **31** 1950022

[Wil15] Wilde M M 2015 Recoverability in quantum information theory *Proc. R. Soc.* A **471** 20150338

[Wil17] Wilde M M 2017 *Quantum Information Theory* 2nd edn (Cambridge: Cambridge University Press)

[Wil18a] Wilde M M 2018 Optimized quantum f-divergences and data processing *J. Phys. A: Math. Theor.* **51** 374002

[Wil18b] Wilde M M 2018 Recoverability for Holevo's just-as-good fidelity *Proc. 2018 IEEE Int. Symp. Information Theory* (Piscataway, NJ: IEEE) pp 2331–5

[WTB17] Wilde M M, Tomamichel M and Berta M 2017 Converse bounds for private communication over quantum channels *IEEE Trans. Inf. Theory* **63** 1792–817

[WWY14] Wilde M M, Winter A and Yang D 2014 Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy *Commun. Math. Phys.* **331** 593–622

[Zha20] Zhang H 2020 Equality conditions of data processing inequality for α-z Rényi relative entropies (arXiv:2007.06644)