# On quantum Rényi divergences

**Gergely Bunth**

Doctoral thesis

Supervisor: Dr. Milán Mosonyi

MŰEGYETEM 1782

Budapest University of Technology and Economics
Department of Analysis and Operations Research

2023

# Contents

## Acknowledgements

I am grateful to my supervisor, Milán Mosonyi, for his supervision of my Ph.D. studies. I was lucky to work in a team during my studies, a team Milán brought to life. I would like to express my heartfelt gratitude to my consultant and colleague Péter Vrana, for his unwavering support and guidance. His patience never ran out.

This work was partially funded by the National Research, Development and Innovation Office of Hungary via the research grants K124152 and KH129601, by the Ministry of Culture and Innovation and the National Research, Development and Innovation Office within the Quantum Information National Laboratory of Hungary (Grant No. 2022-2.1.1-NL-2022-00004), by the ÚNKP-20-5 New National Excellence Program of the Ministry for Innovation and Technology and by the Momentum Program of the Hungarian Academy of Sciences (grant no. LP2021-15/2021).

# 1 Introduction

In 1961 Alfréd Rényi gave an axiomatic description of the quantities that are today known as *classical Rényi $\alpha$-divergences* [Rén61]. These functions are evaluated on pairs of vectors of equal length $(\rho, \sigma)$ consisting of positive numbers, both indexed by a finite index set $\mathcal{I}$, as

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\underbrace{\sum_{i\in\mathcal{I}}\rho(i)^\alpha\sigma(i)^{1-\alpha}}_{=:Q_\alpha(\rho\|\sigma)} - \frac{1}{\alpha-1}\log\sum_{i\in\mathcal{I}}\rho(i), \tag{1.1}$$

where $\alpha \in (0, +\infty) \setminus \{1\}$ is a parameter (for the general definitions see Subsection 2.2.1). For a probability vector $\rho$ (i.e., in addition, elements of $\rho$ sum up to 1), these functions give the so-called *Rényi entropies*, the *Kullback-Leibler divergence* (or *classical relative entropy*) and the *Shannon entropy* as special cases:

$$\begin{aligned}
H_\alpha(\rho) &:= -D_\alpha(\rho\|\mathbf{1}) = \frac{1}{1-\alpha}\log\sum_{i\in\mathcal{I}}\rho(i)^\alpha, \\
D(\rho\|\sigma) = D_1(\rho\|\sigma) &:= \lim_{\alpha\to1} D_\alpha(\rho\|\sigma) = \sum_{i\in\mathcal{I}}\rho(i)(\log\rho(i) - \log\sigma(i)), \\
H(\rho) = H_1(\rho) &:= \lim_{\alpha\to1} H_\alpha(\rho) = -D(\rho\|\mathbf{1}) = -\sum_{i\in\mathcal{I}}\rho(i)\log\rho(i),
\end{aligned}$$

where $\mathbf{1}$ is the vector with $\mathbf{1}(i) = 1\ \forall i \in \mathcal{I}$. For the study and applications of the (classical) Rényi divergences, the relevant quantity is actually $Q_\alpha = \sum_{i\in\mathcal{I}}\rho(i)^\alpha\sigma(i)^{1-\alpha}$. Interestingly, the relative entropy in itself also determines the whole one-parameter family of Rényi divergences, as for every $\alpha \in (0,1) \cup (1,+\infty)$,

$$-\log Q_\alpha(\rho\|\sigma) = \min_{\omega\in\mathcal{P}(\mathcal{I})}\{\alpha D(\omega\|\rho) + (1-\alpha)D(\omega\|\sigma)\},$$

where the optimization is over all finitely supported probability distributions $\omega$ on $\mathcal{I}$ (see [CM03] and [MO21]).

The Shannon entropy and the classical relative entropy by 1961 had already been playing a distinguished role in statistical physics, information theory and economics. The Rényi $\alpha$-divergences and Rényi $\alpha$-entropies then turned out to be the unified picture of central information quantities in these fields. See, for instance, [Csi93] for the role of the Rényi divergences and derived information measures (entropy, divergence radius, channel capacity) in classical state discrimination, as well as source- and channel coding. Coming decades have seen a surge first in digital technology and information theory, then in quantum technology with the ever-increasing feasibility of nanomanipulation of systems on the molecular or atomic level. From the viewpoint of applications this surge made it necessary to develop on the one hand quantum information theory dealing with coding and decoding, hypothesis testing, compression and decompression and more, all in the quantum framework. On the other hand further development of governing quantities of quantum physics of bound systems on the atomic level was needed. From the viewpoint of mathematics, this was the surge of externally motivated and mathematically interesting dissimilarity measures in matrix analysis. Brand new possibilities opened in choosing and axiomatizing the fundamental information quantities and new applications of notions of matrix analysis meant new approaches in studying them e.g., different relaxations of majorization or a whole spectrum of different matrix means.

From both the theoretical and practical point of view in quantum information theory the most important quantum extensions of (1.1) are the *Petz-type* [Pet86b] and the *sandwiched Rényi divergences* [MDS+13, WWY14], given by

$$D_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1}\log \underbrace{\operatorname{Tr}\rho^\alpha\sigma^{1-\alpha}}_{=:Q_\alpha(\rho\|\sigma)} - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho, \tag{1.2}$$

$$D^*_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1}\log \underbrace{\operatorname{Tr}\left(\rho^{\frac{1}{2}}\sigma^{\frac{1-\alpha}{\alpha}}\rho^{\frac{1}{2}}\right)^\alpha}_{=:Q^*_\alpha(\rho\|\sigma)} - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho, \tag{1.3}$$

for $\alpha \in (0,\infty)\setminus\{1\}$ and for positive definite operators on the same finite dimensional Hilbert space (for general definitions see Section 2.2.2). For a quantum state $\rho$ of full support (i.e. in addition to being positive definite, $\rho$ satisfies the trace-condition $\operatorname{Tr}\rho = 1$), the *quantum relative entropy* or *Umegaki relative entropy* is a quantum extension of the classical relative entropy and can be given as a limit of either (1.2) or (1.3):

$$D^{\mathrm{Um}}(\rho\|\sigma) := \lim_{\alpha\to 1} D_\alpha(\rho\|\sigma) = \lim_{\alpha\to 1} D^*_\alpha(\rho\|\sigma) = \operatorname{Tr}\rho(\log\rho - \log\sigma).$$

Among other usages [HKM+02, JV18, LY22b, LY22c, LY22a, LY23, LYH23, MO21], the Petz-type Rényi divergences found operational application in determining the direct part of the error exponents in i.i.d. hypothesis testing (for $\alpha \in (0,1)$) [ANSV08, Nag06], the sandwiched Rényi divergences found operational usage in determining the so-called strong-converse exponents (for $\alpha \in (1,\infty)$) [MO17] and their meeting point, the Umegaki relative entropy found application in the meeting point of the two problems, the so-called Stein lemma [ON05, Hay07, ANSV08].

In quantum information theory, the quantum framework is a generalization of the classical framework, or in other words, the classical framework is embedded in the quantum framework. To recover the classical framework one has to fix a basis and require all operators in the quantum framework to be diagonalizable in that basis. Or equivalently to require all operators to be commuting with each other. A quantum extension of the classical Rényi divergences (1.1) is then a function on operators that reduce to the classical Rényi divergences, when all operators commute. (1.2) or (1.3) are obviously such extensions but because of noncommutativity there are infinitely many other quantum extensions.

Interestingly, if we pose a further requirement on quantum Rényi divergences, the data processing inequality (DPI), that is, monotonicity under completely positive trace-preserving maps (CPTP maps), then the extensions $D_\alpha^{\mathrm{meas}}$, $D_\alpha^{\max}$ can be given that are minimal and maximal [Mat18]. This is to say, that if $D_\alpha^q$ is a quantum extension of $D_\alpha$ that satisfies the DPI, then

$$D_\alpha^{\mathrm{meas}}(\rho\|\sigma) \le D_\alpha^q(\rho\|\sigma) \le D_\alpha^{\max}(\rho\|\sigma), \tag{1.4}$$

for all $\rho, \sigma$ states over the same Hilbert space and for all $\alpha \in (0,\infty) \setminus \{1\}$ (for the definitions see Subsection 2.2.2). Another quantum extension of the classical relative entropy that will be of use to us is the so-called Belavkin-Staszewski entropy $D^{\max}$ [BS82]. For a quantum state $\rho$ of full support, $D^{\max}$ can be given as a limit of the maximal Rényi divergences $D_\alpha^{\max}$:

$$D^{\max}(\rho\|\sigma) := \operatorname{Tr} \rho \log\left(\rho^{\frac{1}{2}} \sigma^{-1} \rho^{\frac{1}{2}}\right) = \lim_{\alpha \to 1} D_\alpha^{\max}(\rho\|\sigma).$$

If we require the DPI and additivity on tensor products, then it can be shown that the Umegaki relative entropy is the minimal and the Belavkin-Staszewski entropy is the maximal extension of the classical relative entropy that satisfy these two properties similar to the above sense of (1.4) [Mat18].

The Rényi $(\alpha, z)$-divergences [AD15] unify the Petz-type and sandwiched divergences in a single family, at least by giving a single formula:

$$D_{\alpha,z}(\rho\|\sigma) := \frac{1}{\alpha-1} \log \underbrace{\operatorname{Tr}\left(\rho^{\frac{\alpha}{2z}} \sigma^{\frac{1-\alpha}{z}} \rho^{\frac{\alpha}{2z}}\right)^z}_{=:Q_{\alpha,z}(\rho\|\sigma)} - \log \operatorname{Tr} \rho,$$

with

$$D_\alpha^*(\rho\|\sigma) = D_{\alpha,\alpha}(\rho\|\sigma), \quad \text{and} \quad D_\alpha(\rho\|\sigma) = D_{\alpha,1}(\rho\|\sigma).$$

The somewhat philosophical question of the basis on which a set of quantum extensions can be called a family is not closed. Indeed on one hand one could plot any $z(\alpha)$ function on the $\alpha - z$ quarter plane of the $(\alpha, z)$-divergences and get a quantum extension of the classical divergences (and there are even more extensions, that are not $(\alpha, z)$-divergences). It has been already suggested that one could unify the operational parts (see above) of the Petz-type and sandwiched Rényi divergences with the Umegaki relative entropy to get "family". This is further backed by the fact that such a "family" would consist only of quantum divergences that satisfy the DPI, whereas the sandwiched Rényi divergences do not satisfy the DPI for $\alpha < \frac{1}{2}$ and the Petz-type Rényi divergences do not satisfy the DPI for $\alpha > 2$. For this "family" a unified, albeit regularized integral representation has been given recently in [HT23]. Axiomatic study and research of quantum extensions could therefore help to unify the perspective on the ever-increasing set of quantum Rényi divergences, as was done for the classical information measures in [Rén61].

There are many other quantum extensions of the classical Rényi divergences than the ones already mentioned. Although their operational interpretations are not as explicit as that of the Petz-type or sandwiched Rényi divergences, it is often useful to consider and study other extensions as well. Indeed, apart from their study being interesting from the purely mathematical point of view of matrix analysis, some of these quantities serve as useful tools in proofs to arrive at the operationally relevant Rényi information quantities in various problems; see, e.g., the role played by the so-called log-Euclidean Rényi divergences $D_{\alpha,+\infty}$ in determining the strong converse exponent in various problems [LY22b, LY22a, LY23, LYH23, MO17, MO21], or the family of Rényi divergences $D_\alpha^{\#}$ introduced in [FF20], where it was used to determine the strong converse exponent of binary channel discrimination.

From our mathematical point of view, only quantum extensions with good mathematical properties are interesting. The most important one is the DPI: $D_\alpha(\Phi(\rho)\|\Phi(\sigma)) \leq D_\alpha(\rho\|\sigma)$ for all $\Phi$ CPTP maps. The DPI, in turn, is strongly connected to joint convexity. Another desirable property is additivity on tensor products $D_\alpha(\rho_1 \otimes \rho_2\|\sigma_1 \otimes \sigma_2) = D_\alpha(\rho_1\|\sigma_1) + D_\alpha(\rho_2\|\sigma_2)$. However, quantum Rényi divergences without these properties might still be useful; for example $D_{\alpha,+\infty}$ is additive, but not monotone for $\alpha > 1$ (the range of $\alpha$ values for which it was used in [MO17, MO21]), and $D_\alpha^{\#}$ is monotone, but not additive. The study of the mathematical properties of various quantum Rényi divergences and related information quantities and trace functionals has been the subject of intensive research in matrix analysis, functional analysis, and operator algebras in the past several decades; see e.g. [And79, Bei13, CFL16, CL08, FL13, Hia01, Hia13, Hia21, HM23, Jen17, JP06, Kos82, LR73, OP93, Pet86b, Zha20] and references therein. These quantities have also been extended to the most general von Neumann algebra setting in [BST18, Jen18, Jen21, Pet85, Pet86a, Hia18, Hia19], however, in this work we only concern ourselves with the Rényi divergences of operators on finite dimensional Hilbert spaces.

Certain quantum Rényi divergences also found usage in characterizing state transformability questions. The general question in such a problem is giving necessary and/or sufficient conditions whether a set of states $\{\rho\}_{\mathcal{X}}$ can be mapped to another set of states $\{\rho'\}_{\mathcal{X}}$ point-by-point. This can be relaxed to e.g., catalytic or asymptotic transformations and one can also study approximate versions of these, in hope to arrive more easily to a simpler set of conditions. Quantum Rényi divergences can be used to give such transformability conditions on pairs of states, e.g., the maximal Rényi $\alpha$-divergences [Mat18] with $\alpha \in [0, 2]$, and the Rényi $(\alpha, z)$-divergences [AD15] for certain values of $\alpha$ and $z$ [Zha20] or commuting states, as was done in [BHN+15, Jen19, Kli07, MPST21, Tur07]. The sufficient conditions for pairs of commuting states have been extended very recently in [FFHT23] to a complete characterization of asymptotic as well as approximate catalytic convertibility between finite sets of commuting states in terms of the monotonicity of the multivariable Rényi quantities

$$Q_{\underline{\alpha}}(\rho_1, \ldots, \rho_r) := \mathrm{Tr}(\rho_1^{\alpha_1} \cdot \ldots \cdot \rho_r^{\alpha_r}),$$

where $\alpha_1 + \ldots + \alpha_r = 1$ and either all of them are nonnegative or exactly one of them is positive. No sufficient conditions, however, are known in the general noncommutative case.

Rényi divergences have other connections with more distant fields of mathematics as well. Many problems of quantum physics and quantum information theory, e.g., hypothesis testing or the state-transformability questions above, can be formulated in a *resource theory*. In a resource theory a set of possible states of the setting are given, and the question is the set of achievable states from a given one, if there are free states and operations, that can be used infinitely, and some or no costly sates and operations. Some resource theories in turn have a natural connection with semiring theory. The states of such a setting is usually described by positivity (no chance or ratio can be negative), and in the classical as well as the quantum framework composition of independent scenarios are usually described by a multiplication (usually the tensor-product), and composition of mutually exclusive scenarios are usually described by a summation (usually the direct sum). The success chance or distillation ratio in such a setting then is given by the product or the sum of chances or ratios. This, together with the achievability question of resource theories which gives rise to a preorder between the possible states, suggests a generalization of resource theories on a semiring level where the resource theory is a preordered semiring and the success chance or distillation ratio is a monotone homomorphism from the semiring into the nonnegative real numbers. This has been recently exploited by using for example [Fri23] on several accounts.

For example, one could loosen the requirements on state convertibility in a way that the target operators are only bounds. Such a relation makes it possible to describe resource theories, that otherwise can not be tackled with strict state convertibility, like hypothesis testing problems (see e.g., [BV21]). In this setting, it is inevitable to work with unnormalized positive operators instead of states. Further relaxing physical channels to completely positive trace-nonincreasing maps gives rise to a generalized preorder, the so-called *relative submajorization*. In the classical setting a pair of positive vectors $(p,q) \in \mathbb{R}^d_{\geq 0} \times \mathbb{R}^d_{\geq 0}$ is said to relatively submajorize another pair $(p',q')$ if there exists a substochastic map $T$ such that $T(p) \geq p'$ and $T(q) \leq q'$ componentwise [Ren16]. This relation can be used to characterize probabilistic and work-assisted thermal operations between incoherent states, as well as error probabilities in hypothesis testing.

Quantum majorization is a relation between bipartite quantum states sharing a marginal. A state $\rho_{AB}$ quantum majorizes $\rho'_{AB'}$ if there is a quantum channel $T : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_{B'})$ such that $\rho'_{AB'} = (\mathrm{id}_A \otimes T)(\rho_{AB})$. In [GJB+18] it was shown that this relation, as well as a $G$-equivariant version (for some compact group $G$) can be characterized using an infinite family of monotones defined in terms of the conditional min-entropy. For specific classical-quantum states (with system $A$ as a classical bit), quantum majorization with covariance encodes time-translation symmetric Gibbs-preserving transformations. These transformations like thermal operations, put constraints on the evolution of states with coherence between energy eigenstates.

In this work, I will also study transformations between pairs of positive operators by equivariant maps in a sense similar to relative submajorization: given representations $\pi : G \to U(\mathcal{H})$ and $\pi' : G \to U(\mathcal{H}')$, and pairs of positive operators $(\rho,\sigma)$ on $\mathcal{H}$ and $(\rho',\sigma')$ on $\mathcal{H}'$, I say that $(\pi,\rho,\sigma)$ equivariantly relatively submajorizes $(\pi',\rho',\sigma')$ if there is a completely positive trace-nonincreasing map $T$ that is equivariant, i.e., satisfies $T(\pi(g)A\pi(g)^*) = \pi'(g)T(A)\pi'(g)^*$ for all $g \in G$ and operator $A$, in addition to the inequalities $T(\rho) \geq \rho'$ and $T(\sigma) \leq \sigma'$.

An averaging argument shows that this relation can equivalently be understood as transformations between families of positive operators parametrized by two copies of $G$. Somewhat more generally, we will consider pairs of continuous families of positive operators $\rho : X \to \mathcal{B}(\mathcal{H})_{>0}$, $\sigma : Y \to \mathcal{B}(\mathcal{H})_{>0}$, where $X$ and $Y$ are fixed nonempty compact topological spaces (when studying $G$-equivariant transformations for a compact group $G$, one would use $X = Y = G$). In this case we say that $(\rho,\sigma)$ relatively submajorizes $(\rho',\sigma')$ (notation: $(\rho,\sigma) \succcurlyeq (\rho',\sigma')$) if there is a completely positive trace-nonincreasing map $T$ such that $T(\rho(x)) \geq \rho'(x)$ and $T(\sigma(y)) \leq \sigma'(y)$ for all $x \in X$ and $y \in Y$.

The structure of this work is as follows. In Chapter 2, I discuss the necessary background from the literature on divergences as well as some recent results from semiring theory. In Chapter 3 I introduce a way to derive multivariable quantum divergences from previously known relative entropies by computing the relative entropy radius of multiple states. This is one of the two methods that we introduced in [MBV22]. I discuss some properties of these quantities in Section 3.2 and in Section 3.3. I argument that the radius computed from the Belavkin-Staszewski entropy is strictly smaller than the maximal Rényi $\alpha$-divergence $D_\alpha^{\max}$ (in the sense of (1.4)) and in fact defines a new family of Rényi $\alpha$-divergences in Section 3.4. A similar, albeit simpler argument is also given in Section 3.4 showing that such radii computed from quantum relative entropies can never be equal to the minimal Rényi $\alpha$-divergence $D_\alpha^{\mathrm{meas}}$ on all pairs of states.

In Chapter 4, I derive axiomatically the quantities governing some further relaxations of relative submajorization based on the results in [BV23]. This, in turn, also yields the axiomatic definition of a new 2-parameter family of quantum Rényi divergences that are extensions of the sandwiched Rényi divergences for $\alpha > 1$ (much like the $(\alpha, z)$-divergences [AD15]). I note that this new family does not have any obvious relation to the $(\alpha, z)$-divergences and that a specialization of them is the first axiomatic description of the sandwiched Rényi $\alpha$ divergences for $\alpha > 1$.

The main result in Chapter 4 is a characterization of an asymptotic relaxation of this relative submajorization and a sufficient condition for the possibility of a catalytic transformation. We say that $(\rho, \sigma)$ asymptotically relatively submajorizes $(\rho', \sigma')$ if $(2^{o(n)}\rho^{\otimes n}, \sigma^{\otimes n}) \succcurlyeq ({\rho'}^{\otimes n}, {\sigma'}^{\otimes n})$. Assuming that the image of $\sigma$ and $\sigma'$ consist of commuting operators, the characterization is in terms of explicitly given monotones: $(\rho, \sigma)$ asymptotically relatively submajorizes $(\rho', \sigma')$ if and only if the inequalities

$$Q_\alpha^*\left(\rho(x)\middle\|\exp\int_Y \log\sigma \,\mathrm{d}\gamma\right) \geq Q_\alpha^*\left(\rho'(x)\middle\|\exp\int_Y \log\sigma' \,\mathrm{d}\gamma\right) \tag{1.5}$$

hold for every $\alpha > 1$, $x \in X$ and probability measure $\gamma$ on (the Borel $\sigma$-algebra of) $Y$. Note that, the expression $\exp\int_Y \log\sigma \,\mathrm{d}\gamma$ in (1.5) can be viewed as some commutative geometric mean of the operators in the family $\sigma$, that needs to plugged in the formula of the sandwiched Rényi divergences $Q_\alpha^*$ (see (1.3)). If the inequalities are strict, then relative submajorization holds after tensoring both pairs with a suitable catalyst. Without the commutativity assumption, I give generalizations of the conditions (1.5) that are necessary for asymptotic or catalytic relative submajorization. In these, the second argument is replaced with a suitable noncommutative geometric mean. For example,

$$Q_\alpha^*(\rho(x)\|\sigma(y_1)\#\sigma(y_2))$$

is one of these monotones, where $x \in X$, $y_1, y_2 \in Y$, and $\sigma(y_1)\#\sigma(y_2)$ is the matrix geometric mean [PW75].

To prove my results in Chapter 4, I use recent results from the theory of preordered semirings to find conditions in terms of monotone quantities that are additive under direct sums and multiplicative under the tensor product, following some of the ideas of [PVW22, BV21]. While in general these monotones are defined only implicitly, under the additional assumption that the image of $\sigma$ consists of commuting operators, we obtain a complete classification, identifying them as exponentiated sandwiched Rényi divergences between one of the $\rho$ operators and a weighted geometric mean of the $\sigma$ operators. Finding all the relevant quantum extensions appears to be a difficult problem, although our heuristic approach reveals a way to systematically construct some of them. Interestingly, these also give new monotones for pair transformations by specialization: for example, it follows that

$$(\rho, \sigma) \mapsto D_\alpha^*(\rho \| \rho \# \sigma)$$

is a quantity that satisfies the data processing inequality (although it is not a monotone under relative submajorization). This is a special case of the 2-parameter family mentioned above given by the sandwiched Rényi $\alpha$-divergence and the matrix geometric mean. These results are also the precursor and motivation for the second systematic way of deriving Rényi $\alpha$-divergences presented in our work in [MBV22] using matrix means. I do not discuss this method in this work in greater depth.

At the end of Chapter 4 I give enticing applications of the axiomatic approach given in the chapter in terms of composite hypothesis testing, hypothesis testing with group symmetry, asymptotic transformations by thermal processes and approximate joint transformations. I note here that the application on the strong converse regime of composite hypothesis testing given in Subsection 4.4.1 is a generalization of our results from [BV21] which considered a composite null hypothesis consisting of finitely many noncommuting positive operators and a simple alternative hypothesis. This result was already a direct and axiomatic derivation of the strong-converse exponents for composite null hypothesis, that was first derived for simple null and alternative hypotheses in [MO15]. The application given in Subsection 4.4.1 considers composite null and alternative hypotheses consisting of infinitely many positive operators and requires commutativity only between the operators of the alternative hypothesis.

# 2 Preliminaries

## 2.1 Quantum framework

For a finite-dimensional Hilbert space $\mathcal{H}$, let $\mathcal{B}(\mathcal{H})$ denote the set of all linear operators on $\mathcal{H}$, and let $\mathcal{B}(\mathcal{H})_{\mathrm{sa}}$, $\mathcal{B}(\mathcal{H})_{\geq 0}$, $\mathcal{B}(\mathcal{H})_{\gneq 0}$, and $\mathcal{B}(\mathcal{H})_{>0}$ denote the set of self-adjoint, positive semi-definite (PSD), nonzero positive semi-definite, and positive definite operators, respectively. For an interval $J \subseteq \mathbb{R}$, let $\mathcal{B}(\mathcal{H})_{\mathrm{sa},J} := \{A \in \mathcal{B}(\mathcal{H})_{\mathrm{sa}} : \mathrm{spec}(A) \subseteq J\}$, i.e., the set of self-adjoint operators on $\mathcal{H}$ with all their eigenvalues in $J$. Let $\mathcal{S}(\mathcal{H}) := \left\{\rho \in \mathcal{B}(\mathcal{H})_{\geq 0} : \operatorname{Tr} \rho = 1\right\}$ denote the set of *density operators*, or *states*. For an operator $X \in \mathcal{B}(\mathcal{H})$,

$$\|X\|_\infty := \max\{\|X\psi\| : \psi \in \mathcal{H}, \|\psi\| = 1\}$$

denotes the *operator norm* of $X$ (i.e., its largest singular value).

Similarly, for a finite set $\mathcal{I}$, we will use the notation $\mathcal{F}(\mathcal{I}) := \mathbb{C}^{\mathcal{I}}$ for the set of complex-valued functions on $\mathcal{I}$, and $\mathcal{F}(\mathcal{I})_{\geq 0}$, $\mathcal{F}(\mathcal{I})_{\gneq 0}$, $\mathcal{F}(\mathcal{I})_{>0}$ for the set of nonnegative, nonnegative and not constant zero, and strictly positive functions on $\mathcal{I}$. The set of probability density functions on $\mathcal{I}$ will be denoted by $\mathcal{P}(\mathcal{I})$. When equipped with the maximum norm, $\mathcal{F}(\mathcal{I})$ becomes a commutative $C^*$-algebra, which we denote by $\ell^\infty(\mathcal{I})$. In the more general case when $\mathcal{I}$ is an arbitrary nonempty set, we will also use the notations $\mathcal{P}_f(\mathcal{I})$ for the set of finitely supported probability measures, and $\mathcal{P}_f^\pm(\mathcal{I})$ for the set of finitely supported signed probability measures on $\mathcal{I}$, i.e.,

$$\mathcal{P}_f^\pm(\mathcal{I}) := \left\{P \in \mathbb{R}^{\mathcal{I}} : |\operatorname{supp} P| < +\infty, \sum_{i\in\mathcal{I}} P(i) = 1\right\}, \quad \operatorname{supp} P := \{i \in \mathcal{I} : P(i) \neq 0\}.$$

We also introduce the following subset of signed probability measures:

$$\mathcal{P}_{f,1}^\pm(\mathcal{I}) := \left\{P \in \mathcal{P}_f^\pm(\mathcal{I}) : \exists i_+ \in \mathcal{I} \text{ s.t. } P(i_+) > 0 \text{ and } P(i) \leq 0, i \in \mathcal{I} \setminus \{i_+\}\right\},$$

which plays an important role in the definition of multivariable Rényi divergences.

In Chapter 4, we will make use of some facts on positive functionals on $C(X)$ for compact Hausdorff $X$ (see e.g., [Fol99, Chapter 7]). On such a space, a Radon measure is a finite regular Borel measure.

**Theorem 2.1.1** (Riesz representation theorem, [Fol99, 7.2 Theorem])**.** *Let $X$ be a compact Hausdorff topological space and $L : C(X) \to \mathbb{R}$ a positive linear functional. Then there exists a unique Radon* *measure $\mu$ on $X$ such that for all $\xi \in C(X)$ the equality*

$$L(\xi) = \int_X \xi(x)\,\mathrm{d}\mu(x) \tag{2.1}$$

*holds. Conversely, every Radon measure gives rise to a positive linear functional via* (2.1).

Examples of Radon measures include positive linear combinations of Dirac measures and Haar measures of locally compact topological groups. On a compact space $X$, every Radon measure $\mu$ is in the closure (with respect to the vague topology) of the set of positive linear combinations of Dirac measures with total mass $\mu(X)$ [Bou04, III, §2, No. 4, Cor. 3.]. We will use an abuse of notation and just say that $\gamma$ is a measure on the Hausdorff topological space $Y$, it is always implied in such a case that $\gamma$ is a Radon measure, in particular it is a measure on the Borel $\sigma$-algebra of $Y$.

For any nonempty set $\mathcal{X}$, let

$$\mathcal{B}(\mathcal{X},\mathcal{H}),\ \ \mathcal{B}(\mathcal{X},\mathcal{H})_{\geq 0},\ \ \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0},\ \ \mathcal{B}(\mathcal{X},\mathcal{H})_{>0},\ \ \mathcal{S}(\mathcal{X},\mathcal{H}),$$

denote the set of functions mapping from $\mathcal{X}$ into $\mathcal{B}(\mathcal{H})$, $\mathcal{B}(\mathcal{H})_{\geq 0}$, $\mathcal{B}(\mathcal{H})_{\gneq 0}$, $\mathcal{B}(\mathcal{H})_{>0}$, and $\mathcal{S}(\mathcal{H})$, respectively. Elements of $\mathcal{S}(\mathcal{X},\mathcal{H})$ are called *classical-quantum channels*, or *cq channels*, and we will use the terminology *generalized classical-quantum channels*, or *gcq channels*, for the elements of $\mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$. We will normally use the notation $W = (W_x)_{x\in\mathcal{X}}$ to denote elements of $\mathcal{B}(\mathcal{X},\mathcal{H})_{\geq 0}$. We say that $W \in \mathcal{B}(\mathcal{X},\mathcal{H})$ is *classical* if there exists an orthonormal basis $(e_i)_{i\in\mathcal{I}}$ in $\mathcal{H}$ such that $W_x = \sum_{i\in\mathcal{I}}\langle e_i, W_x e_i\rangle |e_i\rangle\langle e_i|$, $x \in \mathcal{X}$; we call any such orthonormal basis a *$W$-basis*. Equivalently, we may identify $W$ with the collection of functions $((\widetilde{W}_x(i) := \langle e_i, W_x e_i\rangle)_{i\in\mathcal{I}})_{x\in\mathcal{X}} \in \mathcal{F}(\mathcal{X},\mathcal{I})$, where we use the notations

$$\mathcal{F}(\mathcal{X},\mathcal{I}),\ \ \mathcal{F}(\mathcal{X},\mathcal{I})_{\geq 0},\ \ \mathcal{F}(\mathcal{X},\mathcal{I})_{\gneq 0},\ \ \mathcal{F}(\mathcal{X},\mathcal{I})_{>0},\ \ \mathcal{P}(\mathcal{X},\mathcal{I}),$$

for the sets of functions mapping elements of $\mathcal{X}$ into functions $f_x \in \mathcal{F}(\mathcal{I})$, $x \in \mathcal{X}$, on the finite set $\mathcal{I}$, such that the $f_x$ are arbitrary/nonnegative/nonnegative and not constant zero/strictly positive/probability density functions on $\mathcal{I}$.

Operations on elements of $\mathcal{B}(\mathcal{X},\mathcal{H})$ are always meant pointwise; e.g., for any $W, W^{(1)}, W^{(2)} \in \mathcal{B}(\mathcal{X},\mathcal{H})$, $V \in \mathcal{B}(\mathcal{H},\mathcal{K})$, and $\sigma \in \mathcal{B}(\mathcal{K})$, where $\mathcal{K}$ is an arbitrary finite-dimensional Hilbert space

$$\begin{aligned}
&VWV^* := (VW_xV^*)_{x\in\mathcal{X}},\\
&W^{(1)} \otimes W^{(2)} := \left(W_x^{(1)} \otimes W_x^{(2)}\right)_{x\in\mathcal{X}},\\
&W \otimes \sigma := (W_x \otimes \sigma)_{x\in\mathcal{X}}.\\
&W^{(1)} \oplus W^{(2)} := \left(W_x^{(1)} \oplus W_x^{(2)}\right)_{x\in\mathcal{X}},\\
&W \oplus \sigma := (W_x \oplus \sigma)_{x\in\mathcal{X}}.
\end{aligned} \tag{2.2}$$

Note that here we only consider the (pointwise) tensor product of functions defined on the same set, and that this notion of tensor product is different from the one used to describe the parallel action of two cq channels, given by

$$W^{(1)} \otimes W^{(2)} := \left(W^{(1)}_{x_1} \otimes W^{(2)}_{x_2}\right)_{(x_1,x_2)\in\mathcal{X}_1\times\mathcal{X}_2},$$

where $W^{(i)} \in \mathcal{B}(\mathcal{X}^{(i)}, \mathcal{H}^{(i)})$, $i = 1, 2$, and possibly $\mathcal{X}^{(1)} \neq \mathcal{X}^{(2)}$, $\mathcal{H}^{(1)} \neq \mathcal{H}^{(2)}$. The tensor product in (2.2) can be interpreted either in this setting, with $\mathcal{X}^{(1)} = \mathcal{X}$, $W^{(1)} = W$ and $\mathcal{X}^{(2)} = \{0\}$, $W^{(2)}_0 = \sigma$, or as the pointwise tensor product between $W^{(1)} = W \in \mathcal{B}(\mathcal{X}, \mathcal{H})$ and the constant function $W^{(2)} \in \mathcal{B}(\mathcal{X}, \mathcal{H})$, $W^{(2)}_x = \sigma$, $x \in \mathcal{X}$.

The set of orthogonal projections on $\mathcal{H}$ is denoted by $\mathbb{P}(\mathcal{H}) := \left\{P \in \mathcal{B}(\mathcal{H}) : P^2 = P = P^*\right\}$. For $P, Q \in \mathbb{P}(\mathcal{H})$, the projection onto $(\mathrm{ran}P) \cap (\mathrm{ran}Q)$ is denoted by $P \wedge Q$. For a sequence of projections $P_1, \ldots, P_r \in \mathcal{B}(\mathcal{H})$ summing to $I$, the corresponding *pinching* operation is

$$\mathcal{B}(\mathcal{H}) \ni X \mapsto \sum_{i=1}^{r} P_i X P_i.$$

For a self-adjoint operator $A$, let $P^A_a := \mathbf{1}_{\{a\}}(A)$ denote the spectral projection of $A$ corresponding to the singleton $\{a\} \subset \mathbb{R}$. (Here and henceforth $\mathbf{1}_H$ stands for the characteristic (or indicator) function of a set $H$.) The projection onto the support of $A$ is $\sum_{a\neq 0} P^A_a$; in particular, if $A$ is positive semi-definite, it is equal to $\lim_{\alpha\searrow 0} A^\alpha =: A^0$. In general, we follow the convention that real powers of a positive semi-definite operator $A$ are taken only on its support, i.e., for any $x \in \mathbb{R}$, $A^x := \sum_{a>0} a^x P^A_a$. In particular, $A^{-1} := \sum_{a>0} a^{-1} P^A_a$ stands for the generalized inverse of $A$, and $A^{-1}A = AA^{-1} = A^0$. For $A \in \mathcal{B}(\mathcal{H})_{\geq 0}$ and a projection $P$ on $\mathcal{H}$ we write $A \in \mathcal{B}(P\mathcal{H})_{\geq 0}$ if $A^0 \leq P$.

For two PSD operators $\rho$, $\sigma$, we write $\rho \perp \sigma$ if $\mathrm{ran}\rho \perp \mathrm{ran}\sigma$, which is equivalent to $\rho\sigma = 0$, and further to $\langle \rho, \sigma\rangle_{HS} := \mathrm{Tr}\,\rho\sigma = 0$, and to $\rho^0\sigma^0 = 0$. In particular, it implies $\rho^0 \wedge \sigma^0 = 0$, but not the other way around.

For two finite-dimensional Hilbert spaces $\mathcal{H}, \mathcal{K}$, we will use the notations $\mathrm{PTP}(\mathcal{H}, \mathcal{K})$ and $\mathrm{CPTP}(\mathcal{H}, \mathcal{K})$ for the set of positive trace-preserving linear maps and the set of completely positive trace-preserving linear maps, respectively, from $\mathcal{B}(\mathcal{H})$ to $\mathcal{B}(\mathcal{K})$. We will also use the notation $\mathrm{P}^+(\mathcal{H}, \mathcal{K})$ for the set of (positive) linear maps from $\mathcal{B}(\mathcal{H})$ to $\mathcal{B}(\mathcal{K})$ such that $\Phi(\rho) \in \mathcal{B}(\mathcal{K})_{\gneq 0}$ for all $\rho \in \mathcal{B}(\mathcal{H})_{\gneq 0}$. We will also consider (completely) positive maps of the form $\Phi : \mathcal{B}(\mathcal{H}) \to \ell^\infty(\mathcal{I})$ and $\Phi : \ell^\infty(\mathcal{I}) \to \mathcal{B}(\mathcal{H})$.

For a finite-dimensional Hilbert space $\mathcal{H}$ and a positive integer $n$, we denote by

$$\mathrm{POVM}(\mathcal{H}, [n]) := \left\{ M = (M_i)_{i=0}^{n-1} \in \mathcal{B}(\mathcal{H})^{[n]}_{\geq 0} : \sum_{i=0}^{n-1} M_i = I \right\}$$

the set of *n-outcome positive operator valued measures* (POVMs) on $\mathcal{H}$, where

$$[n] := \{0, \ldots, n-1\}.$$

Any $M \in \mathrm{POVM}(\mathcal{H}, [n])$ determines a CPTP map $M : \mathcal{B}(\mathcal{H}) \to \ell^\infty([n])$ by

$$M(\cdot) := \sum_{i=0}^{n-1} (\mathrm{Tr}\, M_i(\cdot))\mathbf{1}_{\{i\}}.$$

By log we denote the natural logarithm, and we use two different extensions of it to $[0,+\infty]$, defined as

$$\log x := \begin{cases} -\infty, & x=0,\\ \log x, & x\in(0,+\infty),\\ +\infty, & x=+\infty, \end{cases} \qquad \widehat{\log}\, x := \begin{cases} 0, & x=0,\\ \log x, & x\in(0,+\infty),\\ +\infty, & x=+\infty. \end{cases}$$

Throughout this work I use the convention

$$0\cdot(\pm\infty) := 0.$$

For a function $f:(0,+\infty)\to\mathbb{R}$, the corresponding *operator perspective function* [ENEG11, Eff09, HM17] $\mathcal{P}_f$ is defined on pairs of positive definite operators $\rho,\sigma\in\mathcal{B}(\mathcal{H})_{>0}$ as

$$\mathcal{P}_f(\rho,\sigma) := \sigma^{1/2} f\left(\sigma^{-1/2}\rho\sigma^{-1/2}\right)\sigma^{1/2},$$

and it is extended to pairs of positive semi-definite operators $\rho,\sigma$ as $\mathcal{P}_f(\rho,\sigma) := \lim_{\varepsilon\searrow 0}\mathcal{P}_f(\rho+\varepsilon I,\sigma+\varepsilon I)$, whenever the limit exists. It is easy to see that for the *transpose function* $\tilde{f}(x) := xf(1/x)$, $x>0$, we have

$$\mathcal{P}_f(\rho,\sigma) = \mathcal{P}_{\tilde{f}}(\sigma,\rho),$$

whenever both sides are well-defined. For any $\gamma\in(0,1)$, the choice $f_\gamma := \mathrm{id}_{[0,+\infty)}^{\gamma}$ gives the *Kubo-Ando $\gamma$-weighted geometric mean*, denoted by $\mathcal{P}_{f_\gamma}(\rho,\sigma) =: \sigma\#_\gamma\rho$, $\rho,\sigma\in\mathcal{B}(\mathcal{H})\geq 0$.

### 2.2 Rényi divergences

In the following two subsections I give a review of 2-variable classical Rényi divergences and the 2-variable quantum Rényi divergences that are most commonly used in the literature. Then in subsection 2.2.3, I give a possible way to define multivariable Rényi divergences. Finally, in Subsection 2.2.4, I discuss the most important properties of Rényi divergences, in the most general multivariable quantum setting. Some of these properties, however, are already touched in Subsection 2.2.2, when discussing 2-variable quantum Rényi divergences. Although these properties are somewhat frequent in the literature, the Reader might find it more useful to first skip the more substantial parts of Subsection 2.2.2 and revisit it after Subsection 2.2.4. I structured the contents of this Section so that it advances from the more widely known ideas to the newer ones.

#### 2.2.1 Classical Rényi divergences

**Theorem 2.2.1** ( [Rén61, Theorem 3.]). *Let $\Delta(\cdot\|\cdot):\mathcal{F}([n])_{>0}\times\mathcal{F}([n])_{>0}\to\mathbb{R}$ be a bivariable, real-valued function, so that the following five properties hold:*

*(i) $\Delta$ is invariant under permutations, i.e., $\Delta(\mathcal{P}(p)\|\mathcal{P}(q)) = \Delta(p\|q)$ for all permutations $\mathcal{P}$ and $p,q\in\mathcal{F}([n])_{>0}$;*

*(ii)* *If $p, q \in \mathcal{F}([n])_{>0}$ are such that $p_k \geq q_k$ $\forall k \in [n]$, then $\Delta(p\|q) \geq 0$, and if $p_k \leq q_k$ $\forall k \in [n]$, then $\Delta(p\|q) \leq 0$;*

*(iii)* $\Delta\big(1\|\frac{1}{2}\big) = 2$*;*

*(iv)* *For all $p, q \in \mathcal{F}([n])_{>0}$ and $p', q' \in \mathcal{F}([n'])_{>0}$, $\Delta(p \otimes p'\|q \otimes q') = \Delta(p\|q) + \Delta(p'\|q')$;*

*(v)* *If $p, q \in \mathcal{F}([n])_{>0}$ and $p', q' \in \mathcal{F}([n'])_{>0}$ are such that $\sum_{k=1}^{n} p_k + \sum_{k=1}^{n'} p'_k \leq 1$ and $\sum_{k=1}^{n} q_k + \sum_{k=1}^{n'} q'_k \leq 1$, then there exists a continuous and strictly increasing function $g : \mathbb{R} \to \mathbb{R}$ so that the following hold:*

$$\Delta(p \oplus p'\|q \oplus q') = g^{-1}\left[\frac{\sum_{k=1}^{n} p_k g(\Delta(p\|q)) + \sum_{k=1}^{n'} p'_k g(\Delta(p'\|q'))}{\sum_{k=1}^{n} p_k + \sum_{k=1}^{n'} p'_k}\right].$$

*Then for the function $\Delta$ satisfying (i)-(v) above, the function $g$ is either a linear or an exponential function. If $g$ is linear, then the function $\Delta$ satisfying the above properties must have the form $\Delta(p\|q) = D_1(p\|q)$ for all $p, q \in \mathcal{F}([n])_{>0}$, where*

$$D_1(p\|q) = \frac{\sum_{k=1}^{n} p_k \log_2 \frac{p_k}{q_k}}{\sum_{k=1}^{n} p_k}, \tag{2.3}$$

*and $D_1$ does satisfy the above properties, whereas if $g$ is exponential, then the function $\Delta$ satisfying the above properties must have the form $\Delta(p\|q) = D_\alpha(p\|q)$ with some $\alpha \neq 1$, for all $p, q \in \mathcal{F}([n])_{>0}$, where*

$$D_\alpha(p\|q) = \frac{1}{\alpha - 1} \log_2 \frac{\sum_{k=1}^{n} \frac{p_k^\alpha}{q_k^{\alpha-1}}}{\sum_{k=1}^{n} p_k}, \tag{2.4}$$

*and $D_\alpha$ does satisfy the above properties.*

**Remark 2.2.2.** Requiring the additional property in Theorem 2.2.1 that $\lim_{\varepsilon \searrow 0} \Delta(p \oplus \varepsilon\|p \oplus p) = 0$ for some $p$ with $0 < p < \frac{1}{2}$ excludes $\alpha \leq 0$ from the solutions making all solutions ($\alpha > 0$) continuous in all $p_k, q_k$ $k \in [n]$.

**Remark 2.2.3.** It is easy to check that for $p, q \in \mathcal{F}([n])_{>0}$ $\lim_{\alpha \to 1} D_\alpha(p, q) = D_1(p, q)$ confirming the validity of the subscript of $D_1$.

The quantities in (2.3) and (2.4) for $\alpha > 0$ are called classical Rényi $\alpha$-divergences and can be generalized three-fold: in terms of support, from the classical setting to quantum setting, and to multivariable Rényi $\alpha$-divergences. We will first deal with the generalization with regard the support.

**Definition 2.2.4** (Kullback-Leibler divergence)**.** For a finite set $\mathcal{I}$, and $\rho, \sigma \in \mathcal{F}(\mathcal{I})_{\gneq 0}$ the relative entropy or Kullback-Leibler divergence of $\rho$ and $\sigma$ is defined as

$$D(\rho\|\sigma) := \begin{cases} \sum_{i \in \mathcal{I}} \Big[\rho(i) \widehat{\log} \rho(i) - \rho(i) \widehat{\log} \sigma(i)\Big], & \operatorname{supp} \rho \subseteq \operatorname{supp} \sigma, \\ +\infty, & \text{otherwise.} \end{cases}$$

**Definition 2.2.5** (Rényi $\alpha$-divergences)**.** The classical Rényi $\alpha$-divergences [Rén61] are defined for $\rho, \sigma \in \mathcal{F}(\mathcal{I})_{\gneq 0}$ and $\alpha \in (0,1) \cup (1, +\infty)$ as

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \underbrace{\log Q_\alpha(\rho\|\sigma)}_{=:\psi_\alpha(\rho\|\sigma)} - \frac{1}{\alpha - 1} \log \sum_{i \in \mathcal{I}} \rho(i), \tag{2.5}$$

$$\begin{aligned} Q_\alpha(\rho\|\sigma) &:= \lim_{\varepsilon \searrow 0} \sum_{i \in \mathcal{I}} (\rho(i) + \varepsilon)^\alpha (\sigma(i) + \varepsilon)^{1-\alpha} \\ &= \begin{cases} \sum_{i \in \mathcal{I}} \rho(i)^\alpha \sigma(i)^{1-\alpha} & \alpha \in (0,1) \text{ or } \operatorname{supp} \rho \subseteq \operatorname{supp} \sigma, \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.6}$$

For $\alpha \in \{0, 1, +\infty\}$, the Rényi $\alpha$-divergence is defined by the corresponding limit.

It is easy to see that

$$D_0(\rho\|\sigma) := \lim_{\alpha \searrow 0} D_\alpha(\rho\|\sigma) = -\log \sum_{i \in \operatorname{supp} \rho} \sigma(i) + \log \sum_{i \in \mathcal{I}} \rho(i), \tag{2.7}$$

$$D_1(\rho\|\sigma) := \lim_{\alpha \to 1} D_\alpha(\rho\|\sigma) = \frac{1}{\sum_{i \in \mathcal{I}} \rho(i)} D(\rho\|\sigma),$$

$$D_{+\infty}(\rho\|\sigma) := \lim_{\alpha \to +\infty} D_\alpha(\rho\|\sigma) = \log \inf\{\lambda > 0 : \rho \leq \lambda\sigma\}. \tag{2.8}$$

In particular, the Rényi 1-divergence is the same as the relative entropy up to normalization.

We extend the definitions of the Rényi divergences to the case when the second argument is zero as

$$D_\alpha(\rho\|0) := +\infty, \quad \rho \gneq 0,\ \alpha \in [0, +\infty), \tag{2.9}$$

and the definition of the relative entropy to the case when one or both arguments are zero as

$$D(0\|\sigma) := 0 \ \ \sigma \geq 0, \qquad D(\rho\|0) := +\infty \ \ \rho \gneq 0. \tag{2.10}$$

For any fixed $\rho, \sigma \in \mathcal{F}(\mathcal{I})_{\gneq 0}$, the Rényi divergence is nondecreasing as a function of $\alpha$, and it is continuous on the set of $\alpha$ for which it is finite [vH14].

For the study and applications of the (classical) Rényi divergences, the relevant quantity is actually $Q_\alpha$ (equivalently, $\psi_\alpha$); the normalizations in (2.5) are somewhat arbitrary, and are mainly relevant only for the limits in (2.7)–(2.8). The Rényi $\alpha$-divergences with $\alpha \in (0,1) \cup (1, +\infty)$ can be recovered from the relative entropy as

$$-\log Q_\alpha(\rho\|\sigma) = \inf_\omega \{\alpha D(\omega\|\rho) + (1 - \alpha) D(\omega\|\sigma)\}, \tag{2.11}$$

where the infimum is taken over all $\omega \in \mathcal{P}(\mathcal{I})$ with $\operatorname{supp} \omega \subseteq \operatorname{supp} \rho$, and it is uniquely attained at

$$\omega_\alpha(\rho\|\sigma) := \sum_{i \in S} \frac{\rho(i)^\alpha \sigma(i)^{1-\alpha}}{\sum_{j \in S} \rho(j)^\alpha \sigma(j)^{1-\alpha}} \mathbf{1}_{\{i\}}, \tag{2.12}$$

where $S := \operatorname{supp}\rho \cap \operatorname{supp}\sigma$, provided that $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$, or $\operatorname{supp}\rho \cap \operatorname{supp}\sigma \neq \emptyset$ and $\alpha \in (0,1)$. The case $\alpha \in (0,1)$ was discussed in [CM03] in the more general setting where $\mathcal{I}$ is not finite, while the case $\alpha > 1$ was discussed in the finite-dimensional quantum case in [MO21]; see also Chapter 3 below.

### 2.2.2 Quantum Rényi divergences

In this section I give a brief review of the (2-variable) quantum Rényi divergences most commonly used in the literature, which will also play an important role in the rest. We will discuss various ways to define multivariable Rényi divergences in Chapters 3 and 4.

**Definition 2.2.6.** Let $D_\alpha$ be a classical Rényi $\alpha$-divergence for some $\alpha \in [0, \infty]$. We say that $D_\alpha^q$ is a quantum extension of $D_\alpha$ and a quantum Rényi $\alpha$-divergence if for any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ that are jointly diagonalizable, and any orthonormal basis $(e_i)_{i\in\mathcal{I}}$ jointly diagonalizing $\rho$ and $\sigma$, the following holds

$$D_\alpha^q(\rho\|\sigma) = D_\alpha\big((\langle e_i, \rho e_i\rangle)_{i\in\mathcal{I}} \| (\langle e_i, \sigma e_i\rangle)_{i\in\mathcal{I}}\big).$$

**Remark 2.2.7.** We understand the extension analogously in terms of the $Q_\alpha$ quantities in $D_\alpha$ and note that for any $\alpha \in [0,1) \cup (1, +\infty)$, there is an obvious bijection between quantum extensions of $Q_\alpha$ and quantum extensions of $D_\alpha$.

**Remark 2.2.8.** Since 0 commutes with any other operator, any quantum Rényi $\alpha$-divergence $D_\alpha^q$ must satisfy

$$D_\alpha^q(\rho\|0) = +\infty, \quad \rho \gneq 0,$$

according to (2.9), and any quantum relative entropy $D^q$ must satisfy

$$D^q(0\|\sigma) = 0 \;\; \sigma \geq 0, \qquad D^q(\rho\|0) = +\infty \;\; \rho \gneq 0, \tag{2.13}$$

according to (2.10).

Since these values are fixed by definition, in the discussion of different quantum Rényi divergences and relative entropies below, it is sufficient to consider nonzero arguments most of the time.

**Remark 2.2.9.** Note that there is a bijection between quantum extensions of the Rényi 1-divergence and quantum extensions of the relative entropy, given in one direction by $D^q(\rho\|\sigma) := (\operatorname{Tr}\rho) D_1^q(\rho\|\sigma)$, and in the other direction by $D_1^q(\rho\|\sigma) := D^q(\rho\|\sigma)/\operatorname{Tr}\rho$, for any nonzero $\rho$.

The following examples of quantum Rényi $\alpha$-divergences are well studied in the literature. We review them in some detail for later use.

**Example 2.2.10.** For any $\alpha \in [0,1) \cup (1,+\infty)$ and $z \in (0,+\infty)$, the Rényi $(\alpha,z)$-divergence of $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ is defined as [AD15]

$$D_{\alpha,z}(\rho,\sigma) := \frac{1}{\alpha-1}\log Q_{\alpha,z}(\rho,\sigma) - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho,$$
$$Q_{\alpha,z}(\rho,\sigma) := \begin{cases} \operatorname{Tr}\Big(\rho^{\frac{\alpha}{2z}}\sigma^{\frac{1-\alpha}{z}}\rho^{\frac{\alpha}{2z}}\Big)^z, & \alpha \in [0,1) \text{ or } \rho^0 \leq \sigma^0, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is easy to see that the Rényi $(\alpha, z)$-divergence defines a quantum Rényi $\alpha$-divergence in the sense of Definition 2.2.6. $D_{\alpha,1}(\rho\|\sigma)$ is called the Petz-type (or standard) Rényi $\alpha$-divergence [Pet86b] of $\rho$ and $\sigma$, and $D^*_\alpha(\rho\|\sigma) := D_{\alpha,\alpha}(\rho\|\sigma)$ their sandwiched Rényi $\alpha$-divergence [MDS+13, WWY14]. The limit

$$\begin{aligned} D_{\alpha,+\infty}(\rho\|\sigma) &:= \lim_{z=\to+\infty} D_{\alpha,z}(\rho\|\sigma) \\ &= \begin{cases} \frac{1}{\alpha-1}\log \underbrace{\operatorname{Tr} P e^{\alpha P(\widehat{\log}\,\rho)P + P(\widehat{\log}\,\sigma)P}}_{=:Q_{\alpha,+\infty}(\rho\|\sigma))} - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho, & \alpha \in (0,1) \text{ or } \rho^0 \le \sigma^0, \\ +\infty, & \text{otherwise}, \end{cases} \end{aligned}$$

where $P := \rho^0 \wedge \sigma^0$, is also a quantum Rényi $\alpha$-divergence, often referred to as the log-Euclidean Rényi $\alpha$-divergence [AD15, HP93, MO15]. It is known [LT15, MH23a] that for any function $z : (1-\delta, 1+\delta) \to (0, +\infty]$ such that $\liminf_{\alpha\to 1} z(\alpha) > 0$, and for any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$,

$$\lim_{\alpha\to 1} D_{\alpha,z(\alpha)}(\rho\|\sigma) = \frac{1}{\operatorname{Tr}\rho} D^{\mathrm{Um}}(\rho\|\sigma) =: D^{\mathrm{Um}}_1(\rho\|\sigma),$$

where the Umegaki relative entropy $D^{\mathrm{Um}}(\rho\|\sigma)$ is defined as

$$D^{\mathrm{Um}}(\rho\|\sigma) := \begin{cases} \operatorname{Tr}\Big(\rho\,\widehat{\log}\,\rho - \rho\,\widehat{\log}\,\sigma\Big), & \rho^0 \le \sigma^0, \\ +\infty, & \text{otherwise}. \end{cases} \tag{2.14}$$

In particular, for any $z \in (0, +\infty]$, we define $D_{1,z}(\rho\|\sigma) := D^{\mathrm{Um}}_1(\rho\|\sigma)$.

For every $\alpha \in (0, +\infty)$ and $z \in (0, +\infty]$, the Rényi $(\alpha, z)$-divergence is strictly positive [Mos23, Corollary III.28]. The range of $(\alpha, z)$-values for which $D_{\alpha,z}$ is monotone under CPTP maps was studied in a series of works [Bei13, FL13, Hia13, Pet86b], and was finally characterized completely in [Zha20]. It is clear from their definitions that for every $\alpha \in (0, +\infty)$ and $z \in (0, +\infty]$, the Rényi $(\alpha, z)$-divergence is additive on tensor products.

**Example 2.2.11.** For any quantum Rényi $\alpha$-divergence $D^q_\alpha$, its regularization on a pair $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ is defined as

$$\overline{D}^q_\alpha(\rho\|\sigma) := \lim_{n\to+\infty} \frac{1}{n} D^q_\alpha(\rho^{\otimes n}\|\sigma^{\otimes n}),$$

whenever the limit exists. If the limit exists for all $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, then $\overline{D}^q_\alpha$ is a quantum Rényi $\alpha$-divergence that is weakly additive, and if $D^q_\alpha$ is monotone under CPTP maps then so is $\overline{D}^q_\alpha$.

The classical Rényi $\alpha$-divergences admit two canonical quantum extensions, the minimal and the maximal ones:

**Example 2.2.12.** For any $\alpha \in [0, +\infty]$ and $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ their measured Rényi $\alpha$-divergence is defined as

$$D^{\mathrm{meas}}_\alpha(\rho\|\sigma) := \sup\big\{D(M(\rho)\|M(\sigma)) : M \in \mathrm{POVM}(\mathbb{C}^d, [n]),\, n \in \mathbb{N}\big\}.$$

It is easy to see that $D^{\mathrm{meas}}_\alpha$ is a quantum Rényi $\alpha$-divergence, called the measured, or minimal extension.

As introduced in [Mat18], a reverse test for $\rho, \omega \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ is a pair of vectors $p, q \in \mathcal{F}(\mathcal{X})$ together with $\Gamma : \ell^\infty(\mathcal{I}) \to \mathcal{B}(\mathcal{H})$ a (completely) positive trace-preserving map such that $\Gamma(p) = \rho$, $\Gamma(q) = \sigma$. For a classical Rényi $\alpha$-divergence $D$, let

$$D_\alpha^{\max}(\rho\|\sigma) := \inf\{D(p\|q) : (p, q, \Gamma) \text{ is a reverse test for } (\rho, \sigma)\}.$$

It is easy to see that $D_\alpha^{\max}$ is a quantum Rényi $\alpha$-divergence, called the maximal extension.

It is straightforward to verify from their definitions that both $D_\alpha^{\mathrm{meas}}$ and $D_\alpha^{\max}$ are monotone under PTP maps, and for any quantum extension $D_\alpha^q$ of $D_\alpha$ that is monotone under CPTP maps,

$$D_\alpha^{\mathrm{meas}}(\rho\|\sigma) \leq D_\alpha^q(\rho\|\sigma) \leq D_\alpha^{\max}(\rho\|\sigma) \tag{2.15}$$

holds.

It is also clear that since $D_\alpha$ is additive that $D_\alpha^{\max}$ is subadditive and $D_\alpha^{\mathrm{meas}}$ is superadditive, and the regularized measured and the regularized maximal $D$-divergences

$$\overline{D}_\alpha^{\mathrm{meas}}(\rho\|\sigma) := \sup_{n\in\mathbb{N}} \frac{1}{n} D_\alpha^{\mathrm{meas}}(\rho^{\otimes n}\|\sigma^{\otimes n}) = \lim_{n\to+\infty} \frac{1}{n} D_\alpha^{\mathrm{meas}}(\rho^{\otimes n}\|\sigma^{\otimes n})$$
$$\overline{D}_\alpha^{\max}(\rho\|\sigma) := \inf_{n\in\mathbb{N}} \frac{1}{n} D_\alpha^{\max}(\rho^{\otimes n}\|\sigma^{\otimes n}) = \lim_{n\to+\infty} \frac{1}{n} D_\alpha^{\max}(\rho^{\otimes n}\|\sigma^{\otimes n}),$$

are quantum extensions of $D_\alpha$ that are weakly additive. Obviously, $\overline{D}_\alpha^{\mathrm{meas}}$ and $\overline{D}_\alpha^{\max}$ are monotone under CPTP maps, and for any quantum extension $D_\alpha^q$ of $D_\alpha$ that is monotone under CPTP maps, and any $\rho, \omega \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, $n \in \mathbb{N}$, we have

$$\exists\, \overline{D}_\alpha^q(\rho\|\sigma) := \lim_{n\to+\infty} \frac{1}{n} D_\alpha^q(\rho^{\otimes n}\|\sigma^{\otimes n}) \implies \overline{D}_\alpha^{\mathrm{meas}}(\rho\|\sigma) \leq \overline{D}_\alpha^q(\rho\|\sigma) \leq \overline{D}_\alpha^{\max}(\rho\|\sigma).$$

In particular, if $D_\alpha^q$ is additive then

$$\overline{D}_\alpha^{\mathrm{meas}}(\rho\|\sigma) \leq D_\alpha^q(\rho\|\sigma) \leq \overline{D}_\alpha^{\max}(\rho\|\sigma)$$

for any $(\rho\|\sigma) \in \mathcal{B}(\mathcal{H})_{\gneq 0}$.

For any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, their measured relative entropy is $D^{\mathrm{meas}}(\rho\|\sigma) = (\operatorname{Tr}\rho) D_1^{\mathrm{meas}}(\rho\|\sigma)$. We have $D_0^{\mathrm{meas}} = D_{0,1}$, $D_{1/2}^{\mathrm{meas}} = D_{1/2,1/2}$ (see [NC10, Chapter 9]) and for any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$,

$$D_{+\infty}^{\mathrm{meas}}(\rho\|\sigma) = D_{+\infty}^*(\rho\|\sigma) := D_{+\infty,+\infty}(\rho\|\sigma) := \lim_{\alpha\to+\infty} D_{\alpha,\alpha}(\rho\|\sigma) = \log\inf\{\lambda \geq 0 : \rho \leq \lambda\sigma\}. \tag{2.16}$$

We will call the last quantity the max divergence, it was introduced in [Dat09] under the name max-relative entropy, and its equality to the limit above has been shown in [MDS+13, Theorem 5]. No explicit expression is known for $D_\alpha^{\mathrm{meas}}$ for other $\alpha$ values.

Surprisingly, $\overline{D}_\alpha^{\mathrm{meas}}(\rho\|\sigma)$ has a closed formula for every $\alpha\in[0,+\infty]$, given by

$$\overline{D}_\alpha^{\mathrm{meas}}(\rho\|\sigma)=\begin{cases}D_{\alpha,\alpha}(\rho\|\sigma), & \alpha\in[1/2,+\infty],\\ \frac{\alpha}{1-\alpha}D_{1-\alpha,1-\alpha}(\sigma\|\rho)+\frac{1}{\alpha-1}\log\frac{\operatorname{Tr}\rho}{\operatorname{Tr}\sigma}=D_{\alpha,1-\alpha}(\rho\|\sigma), & \alpha\in(0,1/2),\\ D_{\alpha,1-\alpha}(\rho\|\sigma), & \alpha=0;\end{cases}\tag{2.17}$$

see [HP91] for $\alpha=1$, [MO15] for $\alpha\in(1,+\infty)$, and [HT16] for $\alpha=(1/2,1)$; the last expression for $\alpha\in(0,\frac{1}{2})$ above was first observed by Péter Vrana in August 2022, to the best of my knowledge. The case $\alpha=0$ follows from $D_0^{\mathrm{meas}}=D_{0,1}$ and the additivity of the latter.

For every $\alpha\in(0,1)$, strict positivity of $D_\alpha^{\mathrm{meas}}$ is immediate from the strict positivity of the classical Rényi $\alpha$-divergence, which is a straightforward corollary of Hölder's inequality, and strict positivity of $D_\alpha^{\mathrm{meas}}$ for $\alpha\in[1,+\infty]$ follows from this and the easily verifiable fact that $\alpha\mapsto D_\alpha^{\mathrm{meas}}$ is monotone increasing. Strict positivity of $\overline{D}_\alpha^{\mathrm{meas}}$ follows from $D_\alpha^{\mathrm{meas}}\le\overline{D}_\alpha^{\mathrm{meas}}$.

For any $\alpha\in[0,+\infty]$, the measured Rényi $\alpha$-divergence is superadditive on tensor products, but not additive unless $\alpha\in\{0,1/2,+\infty\}$; see, e.g., [HM17, Remark 4.27] and [MH23b, Proposition III.13] for the latter. On the other hand, for every $\alpha\in[0,+\infty]$, the regularized measured Rényi $\alpha$-divergence is not only weakly additive but even additive on tensor products, according to (2.17) and Example 2.2.10.

$\overline{D}_\alpha^{\mathrm{meas}}$, $\alpha\in[0,+\infty]$, are monotone under PTP maps, according to [Bei13, Jen21, MR17] and (2.17). In particular, the Umegaki relative entropy $D^{\mathrm{Um}}$ is monotone under PTP maps [MR17].

For any $\rho,\sigma\in\mathcal{B}(\mathcal{H})_{\gneq 0}$, their maximal relative entropy is $D^{\max}(\rho\|\sigma)=(\operatorname{Tr}\rho)D_1^{\max}(\rho\|\sigma)$.

Let $\rho_{\sigma,\mathrm{ac}}:=\max\{0\le C\le\rho: C^0\le\sigma^0\}=P\rho P-P\rho(P^\perp\rho P^\perp)^{-1}\rho P$ be the absolutely continuous part of $\rho$ w.r.t. $\sigma$ [AT75], where $P:=\sigma^0$, and let $\lambda_i$, $i\in[r]$, be the different eigenvalues of $\sigma^{-1/2}\rho_{\sigma,\mathrm{ac}}\sigma^{-1/2}$ with corresponding spectral projections $P_i$. Let $\mathcal{I}:=[r]\cup\{r+1\}$, let $\tau_0\in\mathcal{S}(\mathcal{H})$ be arbitrary, and

$$\tau_1:=\begin{cases}\frac{\rho-\rho_{\sigma,\mathrm{ac}}}{\operatorname{Tr}(\rho-\rho_{\sigma,\mathrm{ac}})}, & \rho^0\not\le\sigma^0,\\ \tau_0, & \text{otherwise.}\end{cases}$$

According to [Mat18],

$$\hat{p}(i):=\begin{cases}\lambda_i\operatorname{Tr}\sigma P_i, & i\in[r],\\ \operatorname{Tr}(\rho-\rho_{\sigma,\mathrm{ac}}), & i=r+1,\end{cases}\qquad \hat{q}(i):=\begin{cases}\operatorname{Tr}\sigma P_i, & i\in[r],\\ 0, & i=r+1,\end{cases}\tag{2.18}$$

$$\hat{\Gamma}(\mathbf{1}_{\{i\}}):=\begin{cases}\frac{\sigma^{1/2}P_i\sigma^{1/2}}{\operatorname{Tr}\sigma P_i}, & i\in[r],\ \operatorname{Tr}\sigma P_i\ne 0,\\ \tau_0, & i\in[r],\ \operatorname{Tr}\sigma P_i=0,\\ \tau_1, & i=r+1,\end{cases}$$

is a reverse test for $(\rho, \sigma)$ that is optimal for every $D_\alpha^{\max}(\rho\|\sigma)$, $\alpha \in [0,2] \cup \{+\infty\}$, and

$$Q_\alpha^{\max}(\rho\|\sigma) = \widehat{Q}_\alpha(\rho\|\sigma) := Q_\alpha(\hat{p}\|\hat{q}) = \operatorname{Tr} \mathcal{P}_{f_\alpha}(\rho\|\sigma) \tag{2.19}$$

$$= \begin{cases} \operatorname{Tr} \sigma\big(\sigma^{-1/2}\rho_{\sigma,\mathrm{ac}}\sigma^{-1/2}\big)^\alpha, & \alpha \in [0,1), \\ \operatorname{Tr} \sigma\big(\sigma^{-1/2}\rho\sigma^{-1/2}\big)^\alpha, & \alpha \in (1,2],\ \rho^0 \le \sigma^0, \\ +\infty, & \alpha \in (1,2],\ \rho^0 \not\le \sigma^0, \end{cases} \tag{2.20}$$

$$D^{\max}(\rho\|\sigma) = (\operatorname{Tr}\rho) D_1^{\max}(\rho\|\sigma) = D(\hat{p}\|\hat{q}) = \operatorname{Tr} \mathcal{P}_\eta(\rho,\sigma) \tag{2.21}$$

$$= \begin{cases} \operatorname{Tr} \sigma^{1/2}\rho\sigma^{-1/2}\, \widehat{\log}\big(\sigma^{-1/2}\rho\sigma^{-1/2}\big) = \operatorname{Tr} \rho\, \widehat{\log}\big(\rho^{1/2}\sigma^{-1}\rho^{1/2}\big), & \rho^0 \le \sigma^0, \\ +\infty, & \text{otherwise}, \end{cases} \tag{2.22}$$

$$D_{+\infty}^{\max}(\rho\|\sigma) = D_{+\infty}(\hat{p}\|\hat{q}) = D_{+\infty}^*(\rho\|\sigma),$$

where $f_\alpha := \mathrm{id}_{[0,+\infty)}^\alpha$, $\eta(x) := x\log x$, $x \ge 0$. (For the expressions in terms of the perspective functions, see also [HM17, Hia19], and for the realation between the Kubo-Ando means and the expressions in (2.20) see [MBV22].) The expression in (2.22) is called the Belavkin-Staszewski relative entropy [BS82]. Note that the optimal reverse test above is independent of $\alpha$. No explicit expression is known for $D_\alpha^{\max}$ when $\alpha \in (2,+\infty)$, in which case the above reverse test is known not to be optimal.

Strict positivity of $D_\alpha^{\max}$ for all $\alpha \in (0,+\infty]$ follows from that of $D_\alpha^{\mathrm{meas}}$ and the inequality $D_\alpha^{\mathrm{meas}} \le D_\alpha^{\max}$, which is due to the monotonicity of the classical Rényi divergences under stochastic maps.

For any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ and $\alpha \in (0,2]$, we have

$$D_\alpha^{\max}(\rho\|\sigma) = \lim_{\varepsilon \searrow 0} D_\alpha^{\max}(\rho + \varepsilon I\|\sigma + \varepsilon I); \tag{2.23}$$

see, e.g., [Hia19, HM17, Mat18].

It is immediate from their definition that $D_\alpha^{\max}$, $\alpha \in [0,+\infty]$, are subadditive on tensor products. For $\alpha \in [0,2] \cup \{+\infty\}$, $D_\alpha^{\max}$ is even additive, as one can easily verify from the representation $Q_\alpha^{\max}(\rho\|\sigma) = \operatorname{Tr} \mathcal{P}_{f_\alpha}(\rho\|\sigma)$, $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ in (2.19). However, additivity of $D_\alpha^{\max}$ is not known for $\alpha \in (2,+\infty)$. In particular, we have

$$\overline{D}_\alpha^{\max} \begin{cases} = D_\alpha^{\max}, & \alpha \in [0,2] \cup \{+\infty\}, \\ \le D_\alpha^{\max}, & \alpha \in (2,+\infty). \end{cases}$$

**Remark 2.2.13.** Note that, with the above notations,

$$\lim_{\alpha \searrow 0} Q_\alpha^{\max}(\rho\|\sigma) = \operatorname{Tr} \sigma(\sigma^{-1/2}\rho_{\sigma,\mathrm{ac}}\sigma^{-1/2})^0 = \operatorname{Tr} \sigma \sum_{i:\lambda_i > 0} P_i, \tag{2.24}$$

while

$$Q_0^{\max}(\rho\|\sigma) = Q_0(\hat{p}\|\hat{q}) = \sum_i \hat{p}(i)^0 \hat{q}(i) = \sum_{i:\lambda_i\ \operatorname{Tr} \sigma P_i > 0} \operatorname{Tr} \sigma P_i = \operatorname{Tr} \sigma \sum_{i:\lambda_i\ \operatorname{Tr} \sigma P_i > 0} P_i. \tag{2.25}$$

Since

$$P_i \sigma^{-1/2} \rho_{\sigma,\mathrm{ac}} \sigma^{-1/2} P_i = \lambda_i P_i,$$

we see that $\lambda_i > 0 \implies P_i \not\perp \sigma^0 \iff \operatorname{Tr} \sigma P_i > 0$, and hence (2.24) and (2.25) are equal to each other, i.e.,

$$\lim_{\alpha \searrow 0} D_\alpha^{\max}(\rho\|\sigma) = D_0^{\max}(\rho\|\sigma), \quad \rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}.$$

**Remark 2.2.14.** Note that if $D_1^q$ is an additive quantum Rényi 1-divergence then the corresponding quantum relative entropy $D^q$ is not additive; instead, it satisfies

$$D^q(\rho_1 \otimes \rho_2 \| \sigma_1 \otimes \sigma_2) = (\operatorname{Tr} \rho_2) D^q(\rho_1\|\sigma_1) + (\operatorname{Tr} \rho_1) D^q(\rho_2\|\sigma_2)$$

for any $\rho_k, \sigma_k \in \mathcal{B}(\mathcal{H}_k)_{\gneq 0}$, $k = 1, 2$. Thus, the natural notion of regularization for a quantum relative entropy $D^q$ on a pair $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ is

$$\overline{D}(\rho\|\sigma) := (\operatorname{Tr} \rho) \overline{D}_1^q(\rho\|\sigma),$$

which is well-defined whenever $\overline{D}_1^q(\rho\|\sigma)$ is. Clearly, if $\overline{D}(\rho\|\sigma)$ is well-defined for all $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ then it gives a quantum relative entropy that is weakly additive, and if $D^q$ is monotone under CPTP maps then so is $\overline{D}^q$.

**Remark 2.2.15.** According to Remark 2.2.12, for any given $\alpha \in [0, +\infty]$, and any quantum Rényi $\alpha$-divergence $D_\alpha^q$ that is monotone under CPTP maps,

$$D_\alpha^{\mathrm{meas}} \le D_\alpha^q \le D_\alpha^{\max}.$$

If the regularization of $D_\alpha^q$ is well-defined then we further have

$$\overline{D}_\alpha^{\mathrm{meas}} \le \overline{D}_\alpha^q \le \overline{D}_\alpha^{\max};$$

in particular, this is the case if $D_\alpha^q$ is additive, when we also have $\overline{D}_\alpha^q = D_\alpha^q$.

Likewise, for any quantum relative entropy $D^q$ that is monotone under CPTP maps,

$$D^{\mathrm{meas}} \le D^q \le D^{\max}, \tag{2.26}$$

and if the regularization of $D^q$ is well-defined then we further have

$$D^{\mathrm{Um}} = \overline{D}^{\mathrm{meas}} \le \overline{D}^q \le \overline{D}^{\max} = D^{\max}.$$

It is also known that

$$D^{\mathrm{meas}} < D^{\mathrm{Um}} < D^{\max};$$

see [HM17, Theorem 4.18] for the first inequality (also [BFT17] for a slightly weaker statement), and [HM17, Theorem 4.3] for the second inequality.

**Remark 2.2.16.** Note that $D_{+\infty}^{\mathrm{meas}} = \overline{D}_{+\infty}^{\mathrm{meas}} = \overline{D}_{+\infty}^{\mathrm{max}} = D_{+\infty}^{\mathrm{max}}$ is the unique quantum extension of $D_{+\infty}$ that is monotone under (completely) positive trace-preserving maps, as it was observed in [Tom16], and this unique extension also happens to be additive. On the other hand, for any other $\alpha \in [0, +\infty)$, there are infinitely many different monotone and additive quantum Rényi $\alpha$-divergences; see, e.g., Example 2.2.10.

**Remark 2.2.17.** According to Remark 2.2.33, any additive quantum Rényi $\alpha$-divergence $D_\alpha^q$ satisfies the scaling law

$$D_\alpha^q(t\rho\|s\sigma) = D_\alpha^q(\rho\|\sigma) + D_\alpha(t\|s) = D_\alpha^q(\rho\|\sigma) + \log t - \log s. \tag{2.27}$$

In particular, this holds for $D_{\alpha,z}$, $\alpha \in [0, +\infty)$, $z \in (0, +\infty]$, and $D_\alpha^{\mathrm{max}}$, $\alpha \in [0, 2] \cup \{+\infty\}$. It is easy to verify that $D_\alpha^{\mathrm{max}}$ also satisfies (2.27) for every $\alpha \in (2, +\infty)$, where additivity is not known, and $D_\alpha^{\mathrm{meas}}$ also satisfies (2.27) for every $\alpha \in [0, +\infty]$, even though they are not additive unless $\alpha \in \{0, 1/2, +\infty\}$.

Note that a quantum Rényi 1-divergence $D_1^q$ satisfies the scaling law (2.27) if and only if the corresponding quantum relative entropy $D^q$ satisfies the scaling law

$$D^q(t\rho\|s\sigma) = tD^q(\rho\|\sigma) + (\operatorname{Tr}\rho)D(t\|s), \tag{2.28}$$

which in turn equivalent to

$$D^q(t\rho\|\sigma) = (t\log t)\operatorname{Tr}\rho + tD^q(\rho\|\sigma), \tag{2.29}$$

$$D^q(\rho\|s\sigma) = D^q(\rho\|\sigma) - (\log s)\operatorname{Tr}\rho. \tag{2.30}$$

**Remark 2.2.18.** By definition, a quantum Rényi $\alpha$-divergence $D_\alpha^q$ is trace-monotone, if

$$D_\alpha^q(\rho\|\sigma) \ge D_\alpha(\operatorname{Tr}\rho\|\operatorname{Tr}\sigma) \; (= \log\operatorname{Tr}\rho - \log\operatorname{Tr}\sigma) \tag{2.31}$$

for any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, and it is strictly trace-monotone if equality holds in (2.31) if and only if $\rho = \sigma$. Likewise, a quantum relative entropy $D^q$ is trace-monotone, if

$$D^q(\rho\|\sigma) \ge D(\operatorname{Tr}\rho\|\operatorname{Tr}\sigma) \; (= (\operatorname{Tr}\rho)\log\operatorname{Tr}\rho - (\operatorname{Tr}\rho)\log\operatorname{Tr}\sigma). \tag{2.32}$$

for any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, and it is strictly trace-monotone if equality holds in (2.32) if and only if $\rho = \sigma$. Obviously, any trace-monotone Rényi $\alpha$-divergence or relative entropy is nonnegative. Moreover, is is easy to see that if a quantum Rényi $\alpha$-divergence $D_\alpha^q$ satisfies the scaling law (2.27) then it is nonnegative (strictly positive) if and only if it is (strictly) trace-monotone, and similarly, if a quantum relative entropy $D^q$ satisfies the scaling law (2.28) then it is nonnegative (strictly positive) if and only if it is (strictly) trace-monotone.

**Remark 2.2.19.** If a quantum relative entropy $D^q$ satisfies the trace monotonicity (2.32) then for any $\tau, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$,

$$D^q(\tau\|\sigma) \geq -(\operatorname{Tr}\tau)\log\frac{\operatorname{Tr}\sigma}{\operatorname{Tr}\tau} \geq \operatorname{Tr}\tau\left(1 - \frac{\operatorname{Tr}\sigma}{\operatorname{Tr}\tau}\right) = \operatorname{Tr}\tau - \operatorname{Tr}\sigma, \tag{2.33}$$

and equality holds everywhere when $\tau = \sigma$. As an immediate consequence of this, for any $\sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$,

$$\operatorname{Tr}\sigma = \max_{\tau\in\mathcal{B}(\mathcal{H})_{\geq 0}}\{\operatorname{Tr}\tau - D^q(\tau\|\sigma)\} = \max_{\tau\in\mathcal{B}(\sigma^0\mathcal{H})_{\geq 0}}\{\operatorname{Tr}\tau - D^q(\tau\|\sigma)\}, \tag{2.34}$$

$$\log\operatorname{Tr}\sigma = \max_{\tau\in\mathcal{B}(\mathcal{H})_{\gneq 0}}\left\{\log\operatorname{Tr}\tau - \frac{1}{\operatorname{Tr}\tau}D^q(\tau\|\sigma)\right\} = \max_{\tau\in\mathcal{B}(\sigma^0\mathcal{H})_{\gneq 0}}\left\{\log\operatorname{Tr}\tau - \frac{1}{\operatorname{Tr}\tau}D^q(\tau\|\sigma)\right\}. \tag{2.35}$$

Note that $\tau$ is a maximizer for (2.34) if and only if $\operatorname{Tr}\tau = \operatorname{Tr}\sigma$ and $D^q(\tau\|\sigma) = 0$ (since the second inequality in (2.33) holds as an equality if and only if $\operatorname{Tr}\tau = \operatorname{Tr}\sigma$), and if $D^q$ also satisfies the scaling property (2.28) then $\tau$ is a maximizer for (2.35) if and only if $D\big(\frac{\tau}{\operatorname{Tr}\tau}\big\|\frac{\sigma}{\operatorname{Tr}\sigma}\big) = 0$. If $D^q$ is strictly trace monotone then $\tau = \sigma$ is the unique maximizer for all the expressions in (2.34)–(2.35).

The variational formula (2.34) has already been noted in [Tro12, Lemma 6] in the case $D^q = D^{\mathrm{Um}}$.

**Remark 2.2.20.** It is easy to see from their definitions that $D^{\mathrm{meas}}$, $D^{\mathrm{Um}}$, and $D^{\max}$ are all regular and anti-monotone in their second argument (due to the operator monotonicity of log and operator anti-monotonicity of the inverse [Bha97]), i.e.,

$$D^q(\rho\|\sigma + \varepsilon I) \nearrow D^q(\rho\|\sigma) \ \text{ as } \ \varepsilon \searrow 0. \tag{2.36}$$

By Remark 2.2.35, they are also strongly regular. It is clear from (2.14) and (2.22) that for any fixed $\varepsilon > 0$, $\mathcal{B}(\mathcal{H})^2_{\gneq 0} \ni (\rho, \sigma) \mapsto D^q(\rho\|\sigma + \varepsilon I)$ is continuous when $q = \mathrm{Um}$ or $q = \max$. Hence, by (2.36), $D^{\mathrm{Um}}$ and $D^{\max}$ are both jointly lower semi-continuous in their arguments. In particular, the classical relative entropy is jointly lower semi-continuous, whence $D^{\mathrm{meas}}$, as the supremum of lower semi-continuous functions, is also jointly lower semi-continuous.

**Remark 2.2.21.** It is clear from (2.14) and (2.22) that $D^{\mathrm{Um}}$ and $D^{\max}$ are block additive. For $D^{\mathrm{meas}}$ block sub-additivity can be proven [MBV22].

### 2.2.3 Multivariable Rényi divergences

It is natural to ask whether the concept of Rényi divergences can be generalized to more than two variables. Formulas (2.6) and (2.11) offer two different approaches to do that. In a very general setting, one may consider a set $\mathcal{X}$ equipped with a $\sigma$-algebra $\mathcal{A}$. Then for any measurable $w \in \mathcal{F}(\mathcal{X}, \mathcal{I})_{\gneq 0}$ and signed measure $P$ on $\mathcal{A}$ with $P(\mathcal{X}) = 1$, one may consider

$$\hat{Q}_P(w) := \lim_{\varepsilon\searrow 0}\sum_{i\in\mathcal{I}}\exp\left(\int_{\mathcal{X}}\log(w_x(i) + \varepsilon)\,dP(x)\right), \tag{2.37}$$

or

$$\hat{Q}_P(w) := \lim_{\varepsilon\searrow 0}\sum_{i\in\mathcal{I}}\exp\left(\int_{\mathcal{X}}\log((1-\varepsilon)w_x(i) + \varepsilon/|\mathcal{I}|)\,dP(x)\right), \tag{2.38}$$

where the latter is somewhat more natural when the $w_x$ are probability density functions on $\mathcal{I}$. In the most general case, various issues regarding the existence of the integrals and the limits arise, which are important from a mathematical, but not particularly relevant from a conceptual point, and hence for the rest we will restrict our attention to the case where $P$ is finitely supported. In that case the integrals always exist, and the $\varepsilon \searrow 0$ limit can be easily determined as

$$\hat{Q}_P(w) = \sum_{i\in\mathcal{I}} \left( \left( \prod_{x:\, w_x(i)>0} w_x(i)^{P(x)} \right) \cdot \left\{ \begin{array}{ll} 0, & \text{if } \sum_{x:\, w_x(i)=0} P(x) > 0, \\ 1, & \text{if } \sum_{x:\, w_x(i)=0} P(x) = 0, \\ +\infty, & \text{if } \sum_{x:\, w_x(i)=0} P(x) < 0, \end{array} \right\} \right) \tag{2.39}$$

independently of whether (2.37) or (2.38) is used.

Alternatively, one may define

$$\tilde{Q}_P^{\mathrm{b}}(w) := \sup_{\tau\in[0,+\infty)^{\mathcal{I}}} \left\{ \sum_{i\in\mathcal{I}} \tau(i) - \int_{\mathcal{X}} D(\tau\|w_x)\, dP(x) \right\},$$

which is well-defined at least when $P$ is a probability measure, all the $w_x$ are probability density functions, and $\mathcal{X} \ni x \mapsto D(\tau\|w_x)$ is measurable. Again, we restrict to the case when $P$ is finitely supported, but allow it to be a signed probability measure, in which case we use a slight modification of the above to define

$$Q_P^{\mathrm{b}}(w) := \sup_{\substack{\tau\in[0,+\infty)^{\mathcal{I}} \\ \operatorname{supp}\tau\subseteq\bigcap_{x:\, P(x)>0}\operatorname{supp} w_x}} \left\{ \sum_{i\in\mathcal{I}} \tau(i) - \sum_{x\in\mathcal{X}} P(x) D(\tau\|w_x) \right\}. \tag{2.40}$$

We will show in Section 3.1 that this is equivalent to (2.11) when $\mathcal{X} = \{0,1\}$, $P(0) = \alpha \in (0,1) \cup (1,+\infty)$. It is not too difficult to see that with the definition in (2.40), we have

$$Q_P^{\mathrm{b}}(w) = +\infty \iff \bigcap_{x:\, P(x)>0} \operatorname{supp} w_x \not\subseteq \bigcap_{x:\, P(x)<0} \operatorname{supp} w_x.$$

Thus, while (2.39) and (2.40) coincide when $P$ is a probability measure, they may differ when $P$ can take negative values. The following is easy to verify from (2.39) and (2.40):

**Lemma 2.2.22.** *Let $P \in \mathcal{P}_f^{\pm}(\mathcal{X})$ and $w \in \mathcal{F}(\mathcal{X},\mathcal{I})_{\gneq 0}$, and assume that at least one of the following holds true:*

- *(i)* $\operatorname{supp} w_x = \operatorname{supp} w_{x'}$, $x, x' \in \operatorname{supp} P$;
- *(ii)* $P \in \mathcal{P}_f(\mathcal{X}) \cup \mathcal{P}_{f,1}^{\pm}(\mathcal{X})$, *i.e., either* $P(x) \geq 0$, $x \in \mathcal{X}$, *or there exists a unique* $x_+$ *with* $P(x_+) > 0$ *and* $P(x) \leq 0$, $x \in \mathcal{X} \setminus \{x_+\}$.

*Then* $\hat{Q}_P(w) = Q_P^{\mathrm{b}}(w)$.

**Definition 2.2.23.** For any $P \in \mathcal{P}_f^{\pm}(\mathcal{X})$ and any $w \in \mathcal{F}(\mathcal{X},\mathcal{I})_{\gneq 0}$ such that $\hat{Q}_P(w) = Q_P^{\mathrm{b}}(w)$, we call this common value the *multivariable Rényi $Q_P$ of $w$*, and denote it by $Q_P(w)$.

For any $P \in \left(\mathcal{P}_f(\mathcal{X}) \cup \mathcal{P}^{\pm}_{f,1}(\mathcal{X})\right) \setminus \{\mathbf{1}_{\{x\}} : x \in \mathcal{X}\}$, we define the *(symmetrically normalized) classical $P$-weighted Rényi-divergence* of $w \in \mathcal{F}(\mathcal{X}, \mathcal{I})_{\gneq 0}$ as

$$\begin{aligned} D_P(w) &:= \frac{1}{\prod_{x\in\mathcal{X}}(1-P(x))}\left(-\log Q_P(w) + \sum_{x\in\mathcal{X}} P(x)\log\sum_{i\in\mathcal{I}} w_x(i)\right) \\ &= \frac{1}{\prod_{x\in\mathcal{X}}(1-P(x))}\left(-\log Q_P\left(\left(\frac{w_x}{\sum_{i\in\mathcal{I}} w_x(i)}\right)_{x\in\mathcal{X}}\right)\right). \end{aligned}$$

In this case we also define

$$\widetilde{Q}_P(w) := s(P) Q_P(w),$$

where

$$s(P) := \begin{cases} -1, & P \in \mathcal{P}_f(\mathcal{X}), \\ 1, & P \in \mathcal{P}^{\pm}_{f,1}(\mathcal{X}). \end{cases}$$

Lemma 2.2.22 and (2.39) yield immediately the following:

**Corollary 2.2.24.** *Let $P \in \mathcal{P}_f(\mathcal{X}) \cup \mathcal{P}^{\pm}_{f,1}(\mathcal{X})$ and $w \in \mathcal{F}(\mathcal{X}, \mathcal{I})_{\gneq 0}$. Then*

$$Q_P(w) = \lim_{\varepsilon\searrow 0} Q_P(w+\varepsilon) = \lim_{\varepsilon\searrow 0} Q_P((1-\varepsilon)w + \varepsilon/|\mathcal{I}|). \tag{2.41}$$

**Remark 2.2.25.** In the case when $P$ is a probability measure, $Q_P(w)$ was introduced in classical decision theory, and called the *Hellinger transform* of $P$; see [Str11]. The case where $P(x) > 0$ for exactly one $x$ was very recently considered in [MPST21] in the context of (classical) Blackwell dominance of experiments, and in [FFHT23] in the case where all $w$ are strictly positive, in the context of classical state convertibility.

Note that in the case when $\mathcal{X} = \{0,1\}$ and $\alpha := P(0) \in (0,1) \cup (1,+\infty)$, condition (ii) in Lemma 2.2.22 is always satisfied, and we have

$$Q_P(w) = Q_\alpha(w_0 \| w_1).$$

That is, the multivariable $Q_P$ give multivariable extensions of the $Q_\alpha$ quantities.

**Remark 2.2.26.** Note that when $\mathcal{X} = \{0,1\}$ and $\alpha := P(0) = 0$, neither $\hat{Q}_P(w)$ nor $Q^{\mathrm{b}}_P(w)$ coincides with $Q_0(w_0\|w_1)$ in general. The reason for this in the case of $\hat{Q}_P(w)$ is that the limits $\varepsilon \searrow 0$ and $\alpha \searrow 0$ are not interchangeable, while in the case of $Q^{\mathrm{b}}_P(w)$, it is clear that it only depends on $(w_x)_{x\in\operatorname{supp} P}$, while $Q_0$ depends on $w_0$ (or at least its support) even though $0 \notin \operatorname{supp} P = \operatorname{supp}(0,1) = \{1\}$.

Recall that classical divergences can be identified with quantum divergences defined on commuting operators; in particular, monotonicity under (completely) positive trace-preserving maps makes sense for the former. For the purposes of applications, it is monotone divergences that are relevant, and monotonicity is closely related to joint convexity. The following is easy to verify; see, e.g., [MPST21, Lemma 8].

**Lemma 2.2.27.** *Let $P$ be a finitely supported signed probability measure on $\mathcal{X}$. The following are equivalent:*

*(i)* $\widetilde{Q}_P$ *is jointly convex on* $\mathcal{F}(\mathcal{X},\mathcal{I})_{\gneq 0}$ *for any/some finite non-empty* $\mathcal{I}$*;*

*(ii)* $\widetilde{Q}_P$ *is jointly convex on* $\mathcal{F}(\mathcal{X},\mathcal{I})_{>0}$ *for any/some finite non-empty* $\mathcal{I}$*;*

*(iii)* $P \in \mathcal{P}_f(\mathcal{X}) \cup \mathcal{P}^{\pm}_{f,1}(\mathcal{X})$.

*Proof.* The equivalence of (ii) and (iii) is due to [MPST21, Lemma 8], and the equivalence of (ii) and (i) is immediate from (2.41). □

**Corollary 2.2.28.** *For any $P \in \mathcal{P}_f(\mathcal{X}) \cup \mathcal{P}^{\pm}_{f,1}(\mathcal{X})$, $\widetilde{Q}_P$ is jointly convex and monotone under positive trace-preserving maps, and $D_P$ is monotone under positive trace-preserving maps whenever $P \in \Big(\mathcal{P}_f(\mathcal{X}) \cup \mathcal{P}^{\pm}_{f,1}(\mathcal{X})\Big) \setminus \{\mathbf{1}_{\{x\}} : x \in \mathcal{X}\}$.*

*Proof.* It is straightforward to verify that $\widetilde{Q}_P$ is homogeneous, block additive, and stable under tensoring with an arbitrary state, and hence the assertion follows from Lemma 2.2.27 and Lemma 2.2.34. □

**Definition 2.2.29.** As a generalization of Definition 2.2.6, for any $P \in \Big(\mathcal{P}_f(\mathcal{X}) \cup \mathcal{P}^{\pm}_{f,1}(\mathcal{X})\Big) \setminus \{\mathbf{1}_{\{x\}} : x \in \mathcal{X}\}$ let $D_P$ be a $\mathcal{X}$-variable classical Rényi divergence. We say that $D_P^q$ is a quantum extension of $D_P$ and a $\mathcal{X}$-variable quantum Rényi divergence if for all $W \in \Big(\mathcal{B}(\mathbb{C}^d)_{\gneq 0}\Big)^{\mathcal{X}}$ that is jointly diagonalizable, and any orthonormal basis $(e_i)_{i\in\mathcal{I}}$ jointly diagonalizing all $W_x$, the following holds

$$D_P^q(W) = D_P\Big(\big([\langle e_i, W_x e_i\rangle]_{i\in\mathcal{I}}\big)_{x\in\mathcal{X}}\Big).$$

### 2.2.4 Properties of multivariable Rényi divergences

**Definition 2.2.30.** For two $\mathcal{X}$-variable quantum Rényi divergences $D_P$ and $D'_P$, we write

$$D_P \le D'_P \quad \text{if} \quad D_P(W) \le D'_P(W), \ \ \forall W \in \Big(\mathcal{B}(\mathbb{C}^d)_{\gneq 0}\Big)^{\mathcal{X}}.$$

For bivariable quantum Rényi $\alpha$-divergences on pairs of nonzero PSD operators we also introduce the following strict ordering that will be useful when comparing quantum relative entropies and Rényi $\alpha$-divergences:

**Definition 2.2.31.** Let $D_\alpha, D'_\alpha$ be binary quantum Rényi $\alpha$-divergences. We write

$$D_\alpha < D'_\alpha \ \text{ if } \ \rho^0 \le \sigma^0,\ \rho\sigma \ne \sigma\rho \ \Rightarrow\ D_\alpha(\rho\|\sigma) < D'_\alpha(\rho\|\sigma).$$

**Definition 2.2.32.** Let $D_P$ be a $\mathcal{X}$-variable quantum Rényi divergence. We say that $D_P$ is

- nonnegative if $D_P(W) \ge 0$ for all collections of density operators $W$, and it is strictly positive if it is nonnegative and $D_P(W) = 0$ if and only if $W_x = W_y$, $\forall x, y \in \mathcal{X}$, again for density operators;

- monotone under a given map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ for some finite-dimensional Hilbert spaces $\mathcal{H}, \mathcal{K}$, if

$$\forall W : \quad D_P(\Phi(W)) \le D_P(W),$$

where $\Phi(W) := (\Phi(W_x))_{x\in\mathcal{X}}$; in particular, it is monotone under CPTP maps/PTP maps/ pinchings if monotonicity holds for any map in the given class for any two finite-dimensional Hilbert spaces $\mathcal{H}, \mathcal{K}$, and it is trace-monotone, if

$$\forall W : \quad D_P(\operatorname{Tr} W) \le D_P(W), \tag{2.42}$$

with $\operatorname{Tr} W := (\operatorname{Tr} W_x)_{x\in\mathcal{X}}$. We say that $D_P$ is strictly trace-monotone, if equality in (2.42) implies the existence of a state $\omega$ and numbers $\lambda_x$, $x \in \mathcal{X}$ such that $W_x = \lambda_x \omega$, $\forall x \in \mathcal{X}$.

- jointly convex if for all $W^{(k)}$, $k \in [r]$, and probability distribution $\{p_k\}_{k\in[r]}$,

$$D_P\left(\sum_{k\in[r]} p_k W^{(k)}\right) \le \sum_{k\in[r]} p_k D_P\Big(W^{(k)}\Big),$$

and it is jointly concave if $-D_P$ is jointly convex;

- Additive, if for all $\big\{W^{(1)}, W^{(2)}\big\}$,

$$D_P\Big(W^{(1)} \otimes W^{(2)}\Big) = D_P\Big(W^{(1)}\Big) + D_P\Big(W^{(2)}\Big),$$

and subadditive (superadditive) if LHS $\le$ RHS (LHS $\ge$ RHS) holds above;

- weakly additive, if for all $W$,

$$D_P\big(W^{\otimes n}\big) = n D_P(W), \quad \forall n \in \mathbb{N},$$

and weakly subadditive (superadditive) if LHS $\le$ RHS (LHS $\ge$ RHS) holds above;

- block subadditive, if for any $W$, and any sequence of orthogonal projections $P_0, \dots, P_{r-1} \in \mathcal{B}(\mathcal{H})$ summing to $I$,

$$D_P\left(\sum_{i=0}^{r-1} P_i W P_i\right) \le \sum_{i=0}^{r-1} D_P(P_i W P_i).$$

Conversely, if the inequality in the above always holds in the opposite direction then $D_P$ is called block superadditive, and if it is always an equality, then $D_P$ is block additive.

- (positive) homogeneous, if for every $W$ and $t \in (0, +\infty)$,

$$D_P((tW_x)_{x\in\mathcal{X}}) = t D_P((W_x)_{x\in\mathcal{X}}).$$

**Remark 2.2.33.** Note that $z \otimes X \mapsto zX$ gives a canonical identification between $\mathbb{C} \otimes \mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{H})$. In particular, any additive $\mathcal{X}$-variable quantum Rényi divergence $D_P$ satisfies the scaling law:

$$D_P((t_x W_x)_{x\in\mathcal{X}}) = D_P(t) + D_P(W),$$

where $t = (t_x)_{x\in\mathcal{X}}$.

The following is a lemma to relate the DPI and joint convexity properties of general, multivariable Rényi divergences using the well-known idea of relating trace-monotonicity via the Stinespring representation to the joint convexity property in the case of bivariable Rényi divergences.

**Lemma 2.2.34.** *Assume that a $\mathcal{X}$-variable quantum Rényi divergence $D_P$ is block superadditive, homogeneous, monotone nondecreasing under partial traces. Then $D_P$ is jointly concave and jointly superadditive. Vice versa, if $D_P$ is jointly concave and it is stable under tensoring with the maximally mixed state, i.e., for any $W$ and any $\mathcal{K}$, $D_P(W \otimes (I_{\mathcal{K}}/\dim\mathcal{K})) = D_P(W)$, then for any $W$ and any CPTP map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ we have $D_P(\Phi(W)) \geq D_P(W)$.*

*Proof.* Suppose that $D_P$ is block superadditive and homogeneous. Let $(t_i)_{i\in[r]}$ be a probability distribution, and let $(e_i)_{i\in[r]}$ be an orthonormal system in some Hilbert space $\mathcal{K}$. Then

$$\begin{aligned}
D_P\left(\left(\sum_{i=0}^{r-1} t_i W_x^{(i)}\right)_{x\in\mathcal{X}}\right) &= D_P\left(\left(\mathrm{Tr}_{\mathcal{K}} \sum_{i=0}^{r-1} t_i W_x^{(i)} \otimes |e_i\rangle\langle e_i|_{\mathcal{K}}\right)_{x\in\mathcal{X}}\right) \\
&\geq D_P\left(\left(\sum_{i=0}^{r-1} t_i W_x^{(i)} \otimes |e_i\rangle\langle e_i|_{\mathcal{K}}\right)_{x\in\mathcal{X}}\right) \\
&\geq \sum_{i=0}^{r-1} D_P\left(\left(t_i W_x^{(i)} \otimes |e_i\rangle\langle e_i|_{\mathcal{K}}\right)_{x\in\mathcal{X}}\right) \\
&= \sum_{i=0}^{r-1} t_i D_P\left(\left(W_x^{(i)} \otimes |e_i\rangle\langle e_i|_{\mathcal{K}}\right)_{x\in\mathcal{X}}\right) \\
&= \sum_{i=0}^{r-1} t_i D_P\left(\left(W_x^{(i)}\right)_{x\in\mathcal{X}}\right),
\end{aligned}$$

where the first equality is obvious, the first inequality is by the assumption that $D_P$ is monotone nondecreasing under partial traces, the second inequality is due to the block superadditivity of $D_P$, the second equality follows from homogeneity, and the last equality is due to the isometric invariance of $D_P$. This proves joint concavity, and joint superadditivity follows from it immediately due to homogeneity.

Assume now that $D_P$ is jointly concave and it is stable under tensoring with the maximally mixed state. Let $W \in \mathcal{D}_{\mathcal{H}}(D_P)$ and $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be a CPTP map such that $\Phi(W) \in \mathcal{D}_{\mathcal{K}}(D_P)$. Let $\Phi(.) = \mathrm{Tr}_E\, V(.)V^*$ be a Stinespring representation of $\Phi$, where $V : \mathcal{H} \to \mathcal{H}_E \otimes \mathcal{K}$ is an isometry. Let $(U_{ab})_{a,b=0}^{d_E-1}$ be the discrete Weyl unitaries in some ONB of $\mathcal{H}_E$, so that $(1/d_E^2)\sum_{a,b=0}^{d_E-1} U_{a,b}(.)U_{a,b}^* =$

$(1/d_E)I_E\,\mathrm{Tr}(.)$ Then

$$\begin{aligned} D_P(\Phi(W)) &= D_P(\mathrm{Tr}_E\, VWV^*) \\ &= D_P((1/d_E)I_E \otimes \mathrm{Tr}_E\, VWV^*) \\ &= D_P\left(\frac{1}{d_E^2}\sum_{a,b=0}^{d_E-1}(U_{a,b}\otimes I_{\mathcal{K}})W(U_{a,b}\otimes I_{\mathcal{K}})^*\right) \\ &\geq \frac{1}{d_E^2}\sum_{a,b=0}^{d_E-1} D_P((U_{a,b}\otimes I_{\mathcal{K}})W(U_{a,b}\otimes I_{\mathcal{K}})^*) \\ &= \frac{1}{d_E^2}\sum_{a,b=0}^{d_E-1} D_P(W) = D_P(W), \end{aligned}$$

where the second equality is due to stability, the inequality is due to the joint concavity of $D_P$, and the fourth equality is due to isomeric invariance. □

We will furthermore consider properties that only concern one variable of a divergence. We formulate these only for the case when this is the second variable of a 2-variable divergence; the definitions in the general case can be obtained by straightforward modifications. In particular, we say that a 2-variable quantum Rényi $\alpha$-divergence $D_\alpha$ is

- anti-monotone in the second argument (AM), if for all $\rho, \sigma_1, \sigma_2 \in \mathcal{B}(\mathcal{H})$,

$$\sigma_1 \leq \sigma_2 \implies D_\alpha(\rho\|\sigma_1) \geq D_\alpha(\rho\|\sigma_2);$$

- weakly anti-monotone in the second argument, if for any $(\rho, \sigma)$, and for some $\kappa_{\rho,\sigma} > 0$,

$$[0, \kappa_{\rho,\sigma}) \ni \varepsilon \mapsto D_\alpha(\rho\|\sigma + \varepsilon I) \text{ is decreasing;}$$

- regular, if for any $(\rho, \sigma)$, and for some $\kappa_{\rho,\sigma} > 0$,

$$D_\alpha(\rho\|\sigma) = \lim_{\varepsilon\searrow 0} D_\alpha(\rho\|\sigma + \varepsilon I);$$

- strongly regular, if for any $(\rho, \sigma)$, and any sequence of operators $(\sigma_n)_{n\in\mathbb{N}}$ converging decreasingly to $\sigma$ such that we have

$$D_\alpha(\rho\|\sigma) = \lim_{n\to+\infty} D_\alpha(\rho\|\sigma_n);$$

**Remark 2.2.35.** Note that

$$\text{AM} + \text{regularity} \implies \text{strong regularity.}$$

Indeed, assume that $D_\alpha$ is regular and anti-monotone in its second argument. Let $(\sigma_n)_{n\in\mathbb{N}}$ be a sequence of operators converging decreasingly to $\sigma$. Then for every $n \in \mathbb{N}$,

$$\sigma \leq \sigma_n = \sigma + \sigma_n - \sigma \leq \sigma + \underbrace{\|\sigma_n - \sigma\|_\infty}_{=:c_n} I = \sigma + c_n I,$$

hence

$$D_\alpha(\rho\|\sigma) \geq D_\alpha(\rho\|\sigma_n) \geq D_\alpha(\rho\|\sigma + c_n I), \quad n \in \mathbb{N}.$$

By the regularity assumption, the RHS above tends to $D_\alpha(\rho\|\sigma)$ as $n \to +\infty$, whence also $\lim_{n\to+\infty} D_\alpha(\rho\|\sigma_n) = D_\alpha(\rho\|\sigma)$. Thus, $D_\alpha$ is strongly regular.

### 2.3 On preordered semirings

**Definition 2.3.1.** A *preordered semiring* $(S, +, \cdot, 0, 1, \preccurlyeq)$ consists of a set $S$, two commutative and associative binary operations $+, \cdot : S \times S \to S$ that satisfy $(x + y) \cdot z = x \cdot z + y \cdot z$ for all $x, y, z \in S$, a zero element and a unit element $0, 1 \in S$ (i.e., $0 + x = x$ and $1 \cdot x = x$ for all $x$), and a transitive and reflexive relation (preorder) $\preccurlyeq \subseteq S \times S$. For every $x, y, z \in S$ the $x \preccurlyeq y$ preorder is required to satisfy both $x + z \preccurlyeq y + z$ and $x \cdot z \preccurlyeq y \cdot z$.

**Remark 2.3.2.** Hereinafter I use the same $+, \cdot, 0, 1, \preccurlyeq$ symbols for the notation of binary operations, neutral elements (with the multiplication sign often omitted as usual) and the preorder in the different cases of different semirings. To emphasize the relationship of the preordered semirings to natural numbers or nonnegative reals, I use the symbol $\preccurlyeq$ instead of $\leq$ in Chapter 4 and the rest of this Section to denote the preorder in these semirings as well. In accordance with the above, I simply denote the preordered semiring with the symbol of the underlying set and if otherwise not noted the operations and the preorder is then clear from the context.

**Definition 2.3.3.** For all nonzero $x \in S$ we define $x^k$ to be the $k$-term product $x \cdot x \cdot \ldots \cdot x$ and $x^0 := 1$.

**Proposition 2.3.4.** *Let $S$ be a preordered semiring and let $x, y, s, t \in S$, then $x \preccurlyeq y$ and $s \preccurlyeq t$ imply both $xs \preccurlyeq yt$ and $x + s \preccurlyeq y + t$.*

*Proof.* The relation $x \preccurlyeq y$ implies $xs \preccurlyeq ys$, whereas $s \preccurlyeq t$ implies $ys \preccurlyeq yt$. Then $xs \preccurlyeq yt$ follows as claimed. The proof is analogous for $+$. □

**Remark 2.3.5.** Two preordered semirings will play a distinguished role: the first is the set $\mathbb{R}_{\geq 0}$ of *nonnegative real numbers* with its usual addition, multiplication and total order; the second is the *tropical semiring*. In the multiplicative picture, as a set, the tropical real semiring is the set of nonnegative real numbers $\mathbb{TR} = \mathbb{R}_{\geq 0}$, with the sum of $x$ and $y$ being defined as $\max\{x, y\}$, while $\cdot$ is the usual multiplication. Equipped with the usual total order of the real numbers, this set is a preordered semiring.

Semirings considered in this work have the natural numbers embedded (with the sole exception of the tropical numbers above). More precisely the canonical map which sends $n \in \mathbb{N}$ to the $n$-term sum $1 + 1 + \cdots + 1$ should be an order embedding (i.e., injective and $m \preccurlyeq n$ as natural numbers if and only if their images, also denoted by $m$ and $n$, satisfy $m \preccurlyeq n$ in the semiring considered). All semirings considered are of *polynomial growth* [Fri23].

**Definition 2.3.6.** A semiring is of polynomial growth if there exist a $u \in S$ *power universal* element such that $u \succcurlyeq 1$ and for every nonzero $x \in S$ there is a nonnegative $k$ such that $x \preccurlyeq u^k$ and $1 \preccurlyeq u^k x$.

**Proposition 2.3.7.** *Let $S$ be a semiring of polynomial growth and let $u$ be a power universal of $S$. Then for all nonzero $x, y \in S$ there is a nonnegative $k$ such that $u^k x \succcurlyeq y$.*

*Proof.* Follows directly from Definition 2.3.6 and Proposition 2.3.4. □

The power universal element is usually not unique, but it can be shown that the subsequent definition of the asymptotic preorder does not depend on a particular choice [Vra22, Lemma 1].

**Definition 2.3.8.** Let $S$ be a preordered semiring of polynomial growth and $u \in S$ a power universal element. The asymptotic preorder is defined by $x \succsim y$ if there is a sequence $\{k_n\}_{n=1}^{\infty}$ of nonnegative integers such that $\lim_{n\to\infty} k_n/n = 0$ (i.e., the sequence is sublinear) and for all $n \in \mathbb{N}$ the inequality $u^{k_n} x^n \succcurlyeq y^n$ holds.

**Definition 2.3.9.** Let $S$ be a preordered semiring of polynomial growth and $u \in S$ a power universal element. Let $K_u$ be a map from $(S \setminus \{0\}) \times (S \setminus \{0\})$ to the nonnegative integers defined as $K_u(x, y) := \min\{k \in \mathbb{N} : u^k x \succcurlyeq y\}$. By Proposition 2.3.7, $K_u$ is well-defined and finite for all nonzero $x, y \in S$.

**Proposition 2.3.10.** *Let $S$ be a preordered semiring of polynomial growth and $u \in S$ a power universal element. For all nonzero $x, y \in S$ the sequence $\{K_u(x^n, y^n)\}_{n=1}^{\infty}$ has the following properties:*

*(i) if $x \succsim y$, then the condition of $x \succsim y$ in Definition 2.3.8 is also satisfied by $k_n = K_u(x^n, y^n)$,*

*(ii) it enjoys subadditivity,*

*(iii) $\overline{K}_u(x, y) := \lim_{n\to+\infty} \frac{1}{n} K_u(x^n, y^n)$ exists and $\overline{K}_u(x, y) = \inf \frac{1}{n} K_u(x^n, y^n) \geq 0$.*

*Proof.* $(i)$ follows from the definitions, $ii)$ is immediate from Definition 2.3.9 and Proposition 2.3.4, $(iii)$ then follows from Fekete's lemma. □

**Proposition 2.3.11.** *Let $S$ be a preordered semiring of polynomial growth and $u \in S$ a power universal element. For all nonzero $x, y \in S$ the following are equivalent:*

*(i) there is a sublinear sequence of nonnegative integers $\{k_n\}_{n=1}^{\infty}$ such that $u^{k_n} x^n \succcurlyeq y^n$ holds for infinitely many $n \in \mathbb{N}$;*

*(ii) $\overline{K}_u(x, y) = 0$;*

*(iii) $x \succsim y$.*

*Proof.* $(i) \Rightarrow (ii)$ : by the definition of $K_u$, $K_u(x^n, y^n) \leq k_n$ holds infinitely many times, thus $\overline{K}_u(x, y) \leq \lim_{n\to+\infty} \frac{1}{n} k_n = 0$. $(ii) \Rightarrow (iii)$ : by Proposition 2.3.10, the condition of $x \succsim y$ in Definition 2.3.8 is satisfied by $k_n = K_u(x^n, y^n)$. $(iii) \Rightarrow (i)$ : follows from definition. □

**Corollary 2.3.12.** *Let $S$ be a preordered semiring of polynomial growth and $u \in S$ a power universal element. For all nonzero $x, y \in S$ the following are equivalent:*

*(i) for all sublinear sequences of nonnegative integers $\{k_n\}_{n=1}^{\infty}$, the inequality $u^{k_n}x^n \succcurlyeq y^n$ holds for finitely many $n \in \mathbb{N}$;*

*(ii) $\overline{K}_u(x,y) > 0$;*

*(iii) $x \not\gtrsim y$.*

*Proof.* Immediate from Proposition 2.3.11. □

**Definition 2.3.13.** Let $S$ be a preordered semiring and let $x, y \in S$. If $\exists a \in S \setminus \{0\})$ such that $ax \succcurlyeq ay$ then $x$ is catalytically larger than $y$, in notation $x \succcurlyeq_{\mathrm{c}} y$. If such an $a$ exists, it is in turn called a catalyst.

**Proposition 2.3.14.** *Let $S$ be a preordered semiring and let $a, x, y \in S$, then $ax \succcurlyeq ay$ implies $ax^n \succcurlyeq ay^n$, for all $n \in \mathbb{N}$.*

*Proof.* Suppose that $ax \succcurlyeq ay$ and $ax^k \succcurlyeq ay^k$ hold for a natural number $k$, then $ax^{k+1} \succcurlyeq axy^k \succcurlyeq ay^{k+1}$, where the first inequality follows from multiplying $ax^k \succcurlyeq ay^k$ by $x$ and the second inequality follows from multiplying $ax \succcurlyeq ay$ by $y^k$. The proposition then follows from induction. □

**Remark 2.3.15.** The above proposition underlines the reason to call the element realizing $\succcurlyeq_{\mathrm{c}}$ a catalyst in accordance with the notion of the same name in chemistry.

**Proposition 2.3.16.** *$x \succcurlyeq y \implies x \succcurlyeq_{\mathrm{c}} y \implies x \gtrsim y$.*

*Proof.* The first implication is obvious. For the second implication consider $ax \succcurlyeq ay$. Then by Proposition 2.3.14, $ax^n \succcurlyeq ay^n$ for all $n \in \mathbb{N}$. There exist nonnegative integers $k_1, k_2$ such that $u^{k_1} \succcurlyeq a$ and $u^{k_2}a \succcurlyeq 1$. Thus, for all $n \in \mathbb{N}$, we have $u^{k_1+k_2}x^n \succcurlyeq u^{k_2}ax^n \succcurlyeq u^{k_2}ay^n \succcurlyeq y^n$. In fact there is a constant power realizing the asymptotic ordering. □

**Remark 2.3.17.** Note that while in the case of some simple semirings (e.g., the nonnegative reals or the tropical reals) all the above preorders are equivalent, generally the catalytic preorder can be strictly weaker than the preorder and the asymptotic preorder can be strictly weaker than the catalytic preorder. The catalytic and the asymptotic preorders are called relaxations of the preorder of the semiring.

**Definition 2.3.18.** A map $\varphi : S_1 \to S_2$ is a *homomorphism* between the semirings $(S_1, \preccurlyeq_1)$ and $(S_2, \preccurlyeq_2)$ if $\varphi(0) = 0$, $\varphi(1) = 1$, $\varphi(x+y) = \varphi(x) + \varphi(y)$, $\varphi(xy) = \varphi(x)\varphi(y)$ and $x \preccurlyeq_1 y \implies \varphi(x) \preccurlyeq_2 \varphi(y)$ $\forall x, y \in S_1$. In notation $\varphi \in \mathrm{Hom}(S_1, S_2)$.

**Remark 2.3.19.** A homomorphism is a structure-preserving map between two algebraic structures of the same type. Since in this work we are interested in homomorphisms from various preordered semirings into the real and tropical semirings and their real orderings, the definition above includes the preservation of the preorder in the notion of homomorphisms. The Reader may, however, come across in the literature the term 'monotone semiring homomorphisms' denoting the same notion as the definition above.

**Proposition 2.3.20** (Positive definiteness of homomorphisms)**.** *Let $S$ be a preordered semiring of polynomial growth and $u \in S$ a power universal element. Then for every homomorphism $f$ from $S$ into either $\mathbb{R}$ or $\mathbb{TR}$ and for every nonzero $x \in S$ we have that $f(x) > 0$.*

*Proof.* Since $u$ is power universal, $u \succcurlyeq 1$ and thus $f(u) \succcurlyeq f(1) = 1 > 0$. Then for every nonzero $x \in S$ there is a nonnegative integer $k$ such that $1 \preccurlyeq u^k x$. This yields $f(u)^k f(x) \succcurlyeq f(1) = 1$ and $f(x) \succcurlyeq f(1) f(u)^{-k} > 0$. □

**Proposition 2.3.21** (The trivial homomorphism)**.** *Let $S$ be a preordered semiring of polynomial growth and $u \in S$ a power universal element. If $f \in \mathrm{Hom}(S, \mathbb{TR}) \cup \mathrm{Hom}(S, \mathbb{R}_{\geq 0})$ and $f(u) = 1$ then $f \in \mathrm{Hom}(S, \mathbb{TR})$ and $f(x) \equiv 1$ for $x \in S \setminus \{0\}$.*

*Proof.* Let $f \in \mathrm{Hom}(S, \mathbb{TR}) \cup \mathrm{Hom}(S, \mathbb{R}_{\geq 0})$ and suppose that $f(u) = 1$. Let $x \in S \setminus \{0\}$ then there is a nonnegative integer $k$ such that $1 \preccurlyeq u^k x$ and $u^k \succcurlyeq x$. Applying $f$ and rearranging yields

$$1 = f(1) \preccurlyeq f(u)^k f(x) = f(x) \preccurlyeq f(u)^k = 1,$$

showing that $f(x) = 1$ for any $x \in S \setminus \{0\}$. Now

$$1 = f(1+1) = f(1) + f(1) = 1 + 1$$

yields a contradiction if $f \in \mathrm{Hom}(S, \mathbb{R}_{\geq 0})$. □

The tool to explore the relaxed preorders of a semiring by exploiting its real and tropical homomorphisms is the following result from [Fri23].

**Theorem 2.3.22** ( [Fri23, second part of 7.15. Theorem])**.** *Let $S$ be a preordered semiring of polynomial growth with a power universal element $u$ and with $0 \preccurlyeq 1$. Suppose that $x, y \in S \setminus \{0\}$ such that for all $f \in \mathrm{Hom}(S, \mathbb{TR})$ with $f(u) > 1$ and for all $f \in \mathrm{Hom}(S, \mathbb{R}_{\geq 0})$ the strict inequality $f(x) > f(y)$ holds. Then also the following hold:*

*(i) there is a nonnegative integer $k$ such that $u^k x^n \succcurlyeq u^k y^n$ for every sufficiently large $n$,*

*(ii) if in addition $x$ is power universal then $x^n \succcurlyeq y^n$ for every sufficiently large $n$,*

*(iii) there is a nonzero $a \in S$ such that $ax \succcurlyeq ay$.*

*Moreover, there is a nonnegative integer $k$ such that the catalyst can be chosen as $a := u^k \sum_{j=0}^{n} x^j y^{n-j}$ for any sufficiently large $n$.*

**Corollary 2.3.23.** *Let $S$ be a preordered semiring of polynomial growth with a power universal $u \in S$ and with $0 \preccurlyeq 1$. Then $x \succsim y$ if and only if for all $f \in \mathrm{Hom}(S, \mathbb{TR})$ with $f(u) > 1$ and for all $f \in \mathrm{Hom}(S, \mathbb{R}_{\geq 0})$ the non-strict inequality $f(x) \succcurlyeq f(y)$ holds.*

*Proof.* The only if direction is clear: $u^{k_n} x^n \succcurlyeq y^n$ implies $f(u)^{k_n/n} f(x) \succcurlyeq f(y)$, and by taking the limit as $n \to \infty$, also $f(x) \succcurlyeq f(y)$. For the if direction, assuming $f(x) \succcurlyeq f(y)$ for all such $f$ implies that for all $n \in \mathbb{N}$ and such $f$ we have the strict inequalities $f(ux^n) > f(y^n)$, since by Proposition 2.3.21 $f(u) > 1$. By Theorem 2.3.22, there exists nonzero $a_n \in S$, such that $a_n u x^n \succcurlyeq a_n y^n$, i.e., $ux^n \succcurlyeq_{\mathrm{c}} y^n$.

By Proposition 2.3.16 this implies $ux^n \gtrsim y^n$ for all $n$, which in turn by [Vra22, Lemma 3] implies $x \gtrsim y$. $\square$

**Remark 2.3.24.** The conditions in Theorem 2.3.22 and Corollary 2.3.23 are redundant. Indeed, since exponentiation by any positive real acts as an automorphism on $\mathbb{TR}$ and since for a power universal $u$ one has $u \succcurlyeq 1 \Rightarrow f(u) \succcurlyeq 1$ under any homomorphism $f$, it is sufficient to take into account only a normalized part $\{f \in \mathrm{Hom}(S, \mathbb{TR}) : f(u) = 2\}$ of the non-trivial tropical homomorphisms $\{f \in \mathrm{Hom}(S, \mathbb{TR}) : f(u) > 1\}$ in Theorem 2.3.22 and Corollary 2.3.23 (see also [Fri23, Section 7.]).

**Definition 2.3.25.** We will call $\mathrm{TSper}_1(S, \preccurlyeq) := \mathrm{Hom}(S, \mathbb{R}_{\geq 0}) \cup \{f \in \mathrm{Hom}(S, \mathbb{TR}) : f(u) = 2\}$ the 1-test spectrum [Fri23] of $S$.

Now we are ready to reformulate Theorem 2.3.22 and Corollary 2.3.23.

**Theorem 2.3.26** (Theorem 2.3.22 and Corollary 2.3.23)**.** *Let $S$ be a preordered semiring of polynomial growth with a power universal $u$ and with $0 \preccurlyeq 1$. Suppose that $x, y \in S \setminus \{0\}$. Then we have (i) $\Rightarrow$ all of the conditions in (ii), any one of the conditions in (ii) $\Rightarrow$ (iii) and (iii) $\Leftrightarrow$ (iv) in the following:*

*(i) For all $f \in \mathrm{TSper}_1(S, \preccurlyeq)$ : $f(x) > f(y)$.*

*(ii) (a) There is a nonnegative integer $k$ such that $u^k x^n \succcurlyeq u^k y^n$ for every sufficiently large $n$.*

*(b) If in addition $x$ is power universal then $x^n \succcurlyeq y^n$ for every sufficiently large $n$.*

*(c) There is a nonzero $a \in S$ such that $ax \succcurlyeq ay$ (i.e., $ax \succcurlyeq_{\mathrm{c}} ay$).*

*Moreover, there is a nonnegative integer $k$ such that the catalyst can be chosen as $a := u^k \sum_{j=0}^{n} x^j y^{n-j}$ for any sufficiently large $n$.*

*(iii) For all $f \in \mathrm{TSper}_1(S, \preccurlyeq)$ : $f(x) \succcurlyeq f(y)$.*

*(iv) $x \gtrsim y$.*

**Remark 2.3.27.** If in addition to the requirements in Theorem 2.3.26 the power universal $u$ is invertible, then the conditions $(ii)/(a)$ and $(ii)/(b)$ can be switched with the simpler condition of $x^n \succcurlyeq y^n$ for every sufficiently large $n$.

# 3 Barycentric Rényi divergences

In this chapter we use the term "quantum relative entropy" in a more restrictive (though still very general) sense than in the previous chapter. Namely, a quantum divergence $D^q$ will be called a quantum relative entropy if, on top of being a quantum extension of the classical relative entropy, it is also nonnegative, it satisfies the scaling law (2.28), and the following support condition:

$$D^q(\rho\|\sigma) < +\infty \iff \rho^0 \le \sigma^0.$$

Note that by Remark 2.2.18, any quantum relative entropy in the above sense is also trace-monotone. In particular, no quantum relative entropy can take the value $-\infty$.

**Example 3.0.1.** It is easy to verify that that $D^{\mathrm{Um}}$, $D^{\mathrm{meas}}$ and $D^{\max}$ are all quantum relative entropies in the above more restrictive sense.

## 3.1 Definitions

**Definition 3.1.1.** Let $W \in \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$ be a gcq channel, let $P \in \mathcal{P}_f^{\pm}(\mathcal{X})$ and

$$S_+ := \bigwedge_{x:P(x)>0} W_x^0, \quad S_- := \bigwedge_{x:P(x)<0} W_x^0,$$

and for every $x \in \mathcal{X}$, let $D^{q_x}$ be a quantum relative entropy. We define

$$Q_P^{\mathrm{b},\mathbf{q}}(W) := \sup_{\tau\in\mathcal{B}(S_+\mathcal{H})_{\geq 0}} \left\{ \operatorname{Tr}\tau - \sum_{x\in\mathcal{X}} P(x) D^{q_x}(\tau\|W_x) \right\}, \tag{3.1}$$

$$\psi_P^{\mathrm{b},\mathbf{q}}(W) := \log Q_P^{\mathrm{b},\mathbf{q}}(W), \tag{3.2}$$

$$R_{D^{\mathbf{q}},\mathrm{left}}(W,P) := \inf_{\omega\in\mathcal{S}(S_+\mathcal{H})} \sum_{x\in\mathcal{X}} P(x) D^{q_x}(\omega\|W_x). \tag{3.3}$$

Here, $\mathbf{q} := (q_x)_{x\in\mathcal{X}}$, $D^{\mathbf{q}} := (D^{q_x})_{x\in\mathcal{X}}$, and $R_{D^{\mathbf{q}},\mathrm{left}}(W,P)$ is the $P$-weighted left $D^{\mathbf{q}}$-radius of $W$. We call any $\omega$ attaining the infimum in (3.3) a $P$-weighted left $D^{\mathbf{q}}$-center for $W$.

When $P \notin \{\mathbf{1}_{\{x\}} : x \in \mathcal{X}\}$, we also define the $P$-weighted barycentric Rényi-divergence of $W$ corresponding to $D^{\mathbf{q}}$ as

$$D_P^{\mathrm{b},\mathbf{q}}(W) := \frac{1}{\prod_{x\in\mathcal{X}}(1-P(x))} \left( -\log Q_P^{\mathrm{b},\mathbf{q}}\left( \left( \frac{W_x}{\operatorname{Tr} W_x} \right)_{x\in\mathcal{X}} \right) \right).$$

**Remark 3.1.2.** Since I almost exclusively consider only left divergence radii and left divergence centers in this work, I will normally omit "left" from the terminology.

**Remark 3.1.3.** Note that by definition,

$$P(x) \geq 0,\, x \in \mathcal{X} \implies S_- = I, \quad P(x) \leq 0,\, x \in \mathcal{X} \implies S_+ = I.$$

**Definition 3.1.4.** Let $D^{\mathbf{q}} = (D^{q_0}, D^{q_1})$ be quantum relative entropies. For any two nonzero PSD operators $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, and any $\alpha \in [0, +\infty)$, let

$$Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) := \sup_{\tau \in \mathcal{B}(\rho^0\mathcal{H})_{\geq 0}} \{\operatorname{Tr}\tau - \alpha D^{q_0}(\tau\|\rho) - (1-\alpha)D^{q_1}(\tau\|\sigma)\}, \tag{3.4}$$

$$\psi_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) := \log Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma), \tag{3.5}$$

$$D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) := \frac{\psi_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) - \psi_1^{\mathrm{b},\mathbf{q}}(\rho\|\sigma)}{\alpha - 1}, \tag{3.6}$$

where we define the last quantity only for $\alpha \in [0,1) \cup (1, +\infty)$. $D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma)$ is called the barycentric Rényi $\alpha$-divergence of $\rho$ and $\sigma$ corresponding to $D^{\mathbf{q}}$.

**Remark 3.1.5.** When $D^{q_0} = D^{q_1} = D^q$, we will use the simpler notation $D_\alpha^{\mathrm{b},q}$ instead of $D_\alpha^{\mathrm{b},(q_0,q_1)}$.

**Remark 3.1.6.** Note that with the choice $\mathcal{X} = \{0,1\}$, $W_0 = \rho$, $W_1 = \sigma$, and $P(0) = \alpha$, (3.4) and (3.5) are special cases of (3.1) and (3.2), respectively, when $\alpha \in (0, +\infty)$, and we will show in Lemma 3.1.9 that also (3.6) is a special case of (3.3) in this case. When $\alpha = 0$, the restriction $\tau^0 \leq S_+$ in (3.1) would give $\tau^0 \leq \sigma^0$, while we use $\tau^0 \leq \rho^0$ in (3.4). The reason for this is to guarantee the continuity of $D_\alpha^{\mathrm{b},q}$ at 0; see [MBV22, Proposition V.35.].

**Remark 3.1.7.** It is easy to see that when $P$ is a probability measure, the supremum in (3.1) and the infimum in (3.3) can be equivalently taken over $\mathcal{B}(\mathcal{H})_{\geq 0}$ and $\mathcal{S}(\mathcal{H})$, respectively, i.e.,

$$Q_P^{\mathrm{b},\mathbf{q}}(W) = \sup_{\tau \in \mathcal{B}(\mathcal{H})_{\geq 0}} \left\{\operatorname{Tr}\tau - \sum_{x \in \mathcal{X}} P(x) D^{q_x}(\tau\|W_x)\right\}, \tag{3.7}$$
$$R_{D^{\mathbf{q}},\mathrm{left}}(W, P) = \inf_{\omega \in \mathcal{S}(\mathcal{H})} \sum_{x \in \mathcal{X}} P(x) D^{q_x}(\tau\|W_x),$$

and in the 2-variable case we have

$$\begin{aligned} Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) &= \sup_{\tau \in \mathcal{B}(\mathcal{H})_{\geq 0}} \{\operatorname{Tr}\tau - \alpha D^{q_0}(\tau\|\rho) - (1-\alpha)D^{q_1}(\tau\|\sigma)\} \\ &= \sup_{\tau \in \mathcal{B}((\rho^0 \wedge \sigma^0)\mathcal{H})_{\geq 0}} \{\operatorname{Tr}\tau - \alpha D^{q_0}(\tau\|\rho) - (1-\alpha)D^{q_1}(\tau\|\sigma)\}, \quad \alpha \in (0,1). \end{aligned} \tag{3.8}$$

In the general case, the restriction $\tau^0 \leq S_+$ is introduced to avoid the appearance of infinities of opposite signs in $\sum_{x \in \mathcal{X}} P(x) D^{q_x}(\tau\|W_x)$. In the 2-variable case (3.4), the restriction $\tau^0 \leq \rho^0$ also serves to guarantee that $Q_\alpha^{\mathrm{b},\mathbf{q}}$ is a quantum extension of $Q_\alpha$ for $\alpha > 1$, which would not be true, for instance, if it was replaced by $\tau^0 \leq \rho^0 \wedge \sigma^0$; see, e.g., [MBV22, Corollary 5.40.].

**Remark 3.1.8.** Note that (3.4) can be seen as a 2-variable extension of the variational formula (2.34). In particular, we have

$$Q_1^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \max_{\tau\in\mathcal{B}(\rho^0\mathcal{H})}\{\operatorname{Tr}\tau - D^{q_0}(\tau\|\rho)\} = \operatorname{Tr}\rho, \tag{3.9}$$

where the first equality is by definition (3.4), and the second equality is due to (2.34). Thus,

$$\psi_1^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \log\operatorname{Tr}\rho, \tag{3.10}$$

and for every $\alpha\in[0,1)\cup(1,+\infty)$,

$$D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \frac{1}{\alpha-1}\log Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho. \tag{3.11}$$

By Remark 2.2.19, the maximum in (3.9) is attained at $\tau$ if and only if $\operatorname{Tr}\tau = \operatorname{Tr}\rho$ and $D^{q_0}(\tau\|\rho) = 0$; in particular, $\tau=\rho$ is the unique maximizer in (3.9) when $D^{q_0}$ is strictly positive.

At $\alpha=0$, (3.4) and (2.34) give

$$\sigma^0\le\rho^0 \implies Q_0^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \operatorname{Tr}\sigma \implies D_0^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \log\operatorname{Tr}\rho - \log\operatorname{Tr}\sigma.$$

**Lemma 3.1.9.** *(i) In the setting of Definition 3.1.1,*

$$-\log Q_P^{\mathrm{b},\mathbf{q}}(W) = R_{D^{\mathbf{q}},\mathrm{left}}(W,P) \tag{3.12}$$

*Moreover, if $S_+\le S_-$ then a $\tau\in\mathcal{B}(S_+\mathcal{H})_{\ge 0}$ is optimal in (3.1) if and only if*

$$Q_P^{\mathrm{b},\mathbf{q}}(W) = \operatorname{Tr}\tau \quad\textit{and}\quad \sum_{x\in\mathcal{X}} P(x)D^{q_x}(\tau\|W_x) = 0, \tag{3.13}$$

*and if $\tau\ne 0$ is optimal in (3.1) then $\omega := \tau/\operatorname{Tr}\tau$ is optimal in (3.12). Conversely, for any $\omega$ that is optimal in (3.12) $\tau := e^{-R_{D^{\mathbf{q}},\mathrm{left}}(W,P)}\omega$ is optimal in (3.1).*

*(ii) In the setting of Definition 3.1.4,*

$$-\log Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \inf_{\omega\in\mathcal{S}(\rho^0\mathcal{H})}\{\alpha D^{q_0}(\omega\|\rho) + (1-\alpha)D^{q_1}(\omega\|\sigma)\},\quad \alpha\in[0,+\infty). \tag{3.14}$$

*Assume for the rest that $\alpha\in[0,1]$ or $\rho^0\le\sigma^0$. Then $\tau$ is optimal in (3.4) if and only if*

$$Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \operatorname{Tr}\tau \quad\textit{and}\quad \alpha D^{q_0}(\tau\|\rho) + (1-\alpha)D^{q_1}(\tau\|\sigma) = 0, \tag{3.15}$$

*and if $\tau\ne 0$ is optimal in (3.4) then $\omega := \tau/\operatorname{Tr}\tau$ is optimal in (3.14). Conversely, for any $\omega$ that is optimal in (3.14), $\tau := e^{\psi_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma)}\omega$ optimal in (3.4).*

*Proof.* (i) Assume first that $S_+=0$. Then the only admissible $\tau\in\mathcal{B}(\mathcal{S}_+\mathcal{H})_{\ge 0}$ in (3.1) is $\tau=0$, whence $Q_P^{\mathrm{b},\mathbf{q}}(W)=0$, according to (2.13), and thus $\psi_P^{\mathrm{b},\mathbf{q}}(W)=-\infty$. On the other hand, the infimum in (3.3) is taken over the empty set, and hence it is equal to $+\infty$. Thus, (3.12) and (3.13) hold.

Assume next that $S_+\ne 0$. If there exists an $x\in\mathcal{X}$ such that $P(x)<0$ and $S_+\not\le W_x^0$ then taking

$\tau := \omega := S_+/\operatorname{Tr} S_+$ yields

$$Q_P^{\mathrm{b},\mathbf{q}}(W) \geq \operatorname{Tr}\tau - \underbrace{\sum_{x:P(x)>0} P(x)D^{q_x}(\tau\|W_x)}_{\in\mathbb{R}} - \underbrace{\sum_{x:P(x)<0} P(x)D^{q_x}(\tau\|W_x)}_{=-\infty} = +\infty,$$

and

$$R_{D^{\mathbf{q}},\mathrm{left}}(W,P) \leq \underbrace{\sum_{x:P(x)>0} P(x)D^{q_x}(\omega\|W_x)}_{\in\mathbb{R}} + \underbrace{\sum_{x:P(x)<0} P(x)D^{q_x}(\omega\|W_x)}_{=-\infty} = -\infty,$$

(where we used that $D^{q_x}$ does not take the value $-\infty$), whence (3.12) holds.

Finally, if $0 \neq S_+ \leq S_-$ then the proof follows easily from representing a positive semi-definite operator $\tau \in \mathcal{B}(S_+\mathcal{H})_{\geq 0}$ as a pair $(\omega,t) \in \mathcal{S}(S_+\mathcal{H}) \times [0,+\infty)$. Indeed, we have

$$\begin{aligned} Q_P^{\mathrm{b},\mathbf{q}}(W) &= \sup_{\omega\in\mathcal{S}(S_+\mathcal{H})}\sup_{t\in[0,+\infty)} \left\{ \operatorname{Tr} t\omega - \sum_{x\in\mathcal{X}} P(x)D^{q_x}(t\omega\|W_x) \right\} \\ &= \sup_{\omega\in\mathcal{S}(S_+\mathcal{H})}\sup_{t\in[0,+\infty)} \left\{ t - t\log t - t\underbrace{\sum_{x\in\mathcal{X}} P(x)D^{q_x}(\omega\|W_x)}_{=:c(\omega)} \right\}, \end{aligned} \tag{3.16}$$

where the first equality is by definition, and the second equality follows from the scaling property (2.29). Note that $c(\omega) \neq \pm\infty$ by assumption, and the inner supremum in (3.16) is equal to $e^{-c(\omega)}$, attained at $t = e^{-c(\omega)}$, as for the function $[0,+\infty) \ni t \mapsto t - t\log t - tc =: f(t)$ we have $f'(t) = -\log t - c = 0 \iff t = e^{-c}$, $f''(t) = -1/t < 0$, $t \in (0,+\infty)$. From these, all the remaining assertions in (i) follow immediately.

The assertions in (ii) are special cases of the corresponding ones in (i) when $\alpha \in (0,+\infty)$ (also taking into account (3.8) when $\alpha \in (0,1)$). The case $\alpha = 0$ can be verified analogously to the above; we omit the easy details. $\square$

**Remark 3.1.10.** Clearly, when $\alpha > 1$ and $\rho^0 \not\leq \sigma^0$ then the set of optimal $\tau$ operators in (3.4) is exactly $\mathcal{B}(\rho^0\mathcal{H})_{\gneq 0}$, and the set of optimal $\omega$ states in (3.14) is exactly $\mathcal{S}(\rho^0\mathcal{H})$.

**Corollary 3.1.11.** *Assume that $S_+ \leq S_-$. Then*

$$Q_P^{\mathrm{b},\mathbf{q}}(W) = \max\left\{ \operatorname{Tr}\tau : \tau \in \mathcal{B}(S_+\mathcal{H})_{\geq 0},\ \sum_{x\in\mathcal{X}} P(x)D^{q_x}(\tau\|W_x) = 0 \right\}.$$

*Likewise, if $\alpha \in [0,1]$ or $\rho^0 \leq \sigma^0$, then*

$$Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \max\left\{ \operatorname{Tr}\tau : \tau \in \mathcal{B}(\rho^0\mathcal{H})_{\geq 0},\ \alpha D^{q_0}(\tau\|\rho) + (1-\alpha)D^{q_1}(\tau\|\sigma) = 0 \right\}.$$

*Proof.* Immediate from the characterizations of the optimal $\tau$ in (3.13) and (3.15). $\square$

**Remark 3.1.12.** Note that in the case $S_+ \leq S_-$, the condition $\sum_{x\in\mathcal{X}} P(x)D^{q_x}(\tau\|W_x) = 0$ is necessary for the optimality of $\tau$, but not sufficient. Indeed, it is easy to see from the scaling property (2.29) that

$$\begin{aligned}&\left\{\tau \in \mathcal{B}(S_+\mathcal{H}) : \sum_{x\in\mathcal{X}} P(x)D^{q_x}(\tau\|W_x) = 0\right\}\\&\quad= \left\{\exp\left(-\sum_{x\in\mathcal{X}} P(x)D^{q_x}(\tau\|W_x)\right)\omega : \omega \in \mathcal{S}(S_+\mathcal{H})\right\} \cup \{0\}.\end{aligned}$$

On the other hand, each $\tau \in \mathcal{B}(S_+\mathcal{H}) \setminus \{0\}$ with $\sum_{x\in\mathcal{X}} P(x)D^{q_x}(\tau\|W_x) = 0$ has the extremality property

$$\begin{aligned}\operatorname{Tr}(\lambda\tau) - \sum_{x\in\mathcal{X}} P(x)D^{q_x}(\lambda\tau\|W_x) &= (\lambda - \lambda\log\lambda)\operatorname{Tr}\tau\\&< \operatorname{Tr}\tau = \operatorname{Tr}(\tau) - \sum_{x\in\mathcal{X}} P(x)D^{q_x}(\tau\|W_x)\end{aligned}$$

for every $\lambda \in (0,1)\cup(1,+\infty)$, where the first equality is again due to the scaling property (2.29).

**Remark 3.1.13.** Under the conditions given in Lemma 3.1.9, for the supremum in (3.1) to be a maximum, it is sufficient if the infimum in (3.3) is a minimum. For the latter, a natural sufficient condition is that each $D^{q_x}$ with $x \in \operatorname{supp} P$ is lower semi-continuous in its first argument (when $P$ is a probability measure), or continuous in its first argument with its support dominated by the support of a fixed second argument (when $P$ can take negative values), since the domain of optimization, namely, $\mathcal{S}(\rho^0\mathcal{H})$, is a compact set.

Examples of quantum relative entropies that are lower semi-continuous in their first argument (in fact, in both of their arguments), include $D^{\mathrm{meas}}$, $D^{\mathrm{Um}}$, and their $\gamma$-weighted versions, as well $D^{\max}$, and obviously, all possible convex combinations of these. $D^{\mathrm{Um}}$ and $D^{\max}$ are also clearly continuous in their first argument when its support is dominated by the support of a fixed second argument.

**Remark 3.1.14.** Using (3.10) and the scaling law (2.30), (3.6) can be rewritten as

$$D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = \frac{1}{\alpha-1}\sup_{\tau\in\mathcal{B}(\rho^0\mathcal{H})_{\geq 0}}\left\{\operatorname{Tr}\tau - \alpha D^{q_0}\left(\tau\middle\|\frac{\rho}{\operatorname{Tr}\rho}\right) - (1-\alpha)D^{q_1}\left(\tau\middle\|\frac{\sigma}{\operatorname{Tr}\sigma}\right)\right\} + \log\operatorname{Tr}\rho - \log\operatorname{Tr}\sigma.$$

On the other hand, using Lemma 3.1.9 we get that for every $\alpha \in [0,1)\cup(1,+\infty)$,

$$\begin{aligned}D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) &= \frac{1}{\alpha-1}\log Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho\\&= \frac{1}{1-\alpha}\inf_{\omega\in\mathcal{S}(\rho^0\mathcal{H})}\{\alpha D^{q_0}(\omega\|\rho) + (1-\alpha)D^{q_1}(\omega\|\sigma)\} - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho \qquad (3.17)\\&= \frac{1}{1-\alpha}\inf_{\omega\in\mathcal{S}(\rho^0\mathcal{H})}\left\{\alpha D^{q_0}\left(\omega\middle\|\frac{\rho}{\operatorname{Tr}\rho}\right) + (1-\alpha)D^{q_1}(\omega\|\sigma)\right\} + \log\operatorname{Tr}\rho\\&= \frac{1}{1-\alpha}\inf_{\omega\in\mathcal{S}(\rho^0\mathcal{H})}\left\{\alpha D^{q_0}\left(\omega\middle\|\frac{\rho}{\operatorname{Tr}\rho}\right) + (1-\alpha)D^{q_1}\left(\omega\middle\|\frac{\sigma}{\operatorname{Tr}\sigma}\right)\right\} + \log\operatorname{Tr}\rho - \log\operatorname{Tr}\sigma, \qquad (3.18)\end{aligned}$$

where the first equality is by (3.11), the second equality follows from (3.14), and the third and the fourth equalities from the scaling laws (2.29)–(2.30). Moreover, for $\alpha \in (0,1)$, the infimum can be taken over $\mathcal{S}(\mathcal{H})$, i.e.,

$$
\begin{aligned}
D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) &= \frac{1}{1-\alpha}\inf_{\omega\in\mathcal{S}(\mathcal{H})}\{\alpha D^{q_0}(\omega\|\rho)+(1-\alpha)D^{q_1}(\omega\|\sigma)\} - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho \\
&= \frac{1}{1-\alpha}\inf_{\omega\in\mathcal{S}(\mathcal{H})}\left\{\alpha D^{q_0}\left(\omega\middle\|\frac{\rho}{\operatorname{Tr}\rho}\right)+(1-\alpha)D^{q_1}\left(\omega\middle\|\frac{\sigma}{\operatorname{Tr}\sigma}\right)\right\} + \log\operatorname{Tr}\rho - \log\operatorname{Tr}\sigma,
\end{aligned}
$$

because if $\omega^0 \not\leq \rho^0$ then $D^{q_0}(\omega\|\rho) = D^{q_0}\left(\omega\middle\|\frac{\rho}{\operatorname{Tr}\rho}\right) = +\infty$. The above formulas explain the term "barycentric Rényi divergence".

**Definition 3.1.15.** For $\alpha \in (0,1)$, any $\omega$ attaining the infimum in (3.17) will be called an $\alpha$-weighted (left) $D^{\mathbf{q}}$-center for $(\rho,\sigma)$.

## 3.2 Barycentric Rényi divergences are quantum Rényi divergences

In this section we show that the barycentric Rényi $\alpha$-divergences are quantum Rényi divergences for every $\alpha \in (0,1)$, provided that the defining quantum relative entropies are monotone under pinchings. This latter condition does not pose a serious restriction; indeed, all the concrete quantum relative entropies that we consider in this work (e.g., measured, Umegaki, maximal, and the $\gamma$-weighted versions of these) are monotone under PTP maps, and hence also under pinchings.

Isometric invariance holds even without this mild restriction, and also for $\alpha > 1$:

**Lemma 3.2.1.** *All the quantities in* (3.1)–(3.6) *are invariant under isometries, and hence they are all quantum divergences.*

*Proof.* We prove the statement only for $Q_P^{\mathrm{b},\mathbf{q}}$, as for the other quantities it either follows from that, or the proof goes the same way. Let $W \in \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$ be a gcq channel, $P \in \mathcal{P}_f^{\pm}(\mathcal{X})$, and $V : \mathcal{H} \to \mathcal{K}$ be an isometry. Obviously, $\tilde{S}_+ := \bigwedge_{x:P(x)>0}(VW_xV^*)^0 = V(\bigwedge_{x:P(x)>0} W_x^0)V^* = VS_+V^*$, and for any $\tau \in \mathcal{B}(\tilde{S}_+\mathcal{K})_{\geq 0}$ there exists a unique $\hat{\tau} \in \mathcal{B}(S_+\mathcal{H})_{\geq 0}$ such that $\tau = V\hat{\tau}V^*$. Thus,

$$
\begin{aligned}
Q_P^{\mathrm{b},\mathbf{q}}(VWV^*) &= \sup_{\tau\in\mathcal{B}(\tilde{S}_+\mathcal{K})_{\geq 0}}\left\{\operatorname{Tr}\tau - \sum_{x\in\mathcal{X}}P(x)D^{q_x}(\tau\|VW_xV^*)\right\} \\
&= \sup_{\hat{\tau}\in\mathcal{B}(S_+\mathcal{H})_{\geq 0}}\left\{\operatorname{Tr}V\hat{\tau}V^* - \sum_{x\in\mathcal{X}}P(x)D^{q_x}(V\hat{\tau}V^*\|VW_xV^*)\right\} \\
&= \sup_{\hat{\tau}\in\mathcal{B}(S_+\mathcal{H})_{\geq 0}}\left\{\operatorname{Tr}\hat{\tau} - \sum_{x\in\mathcal{X}}P(x)D^{q_x}(\hat{\tau}\|W_x)\right\} \\
&= Q_P^{\mathrm{b},\mathbf{q}}(W),
\end{aligned}
$$

where the third equality follows by the isometric invariance of the relative entropies. □

Recall that $D^q$ is said to be monotone under pinchings if

$$D^q\left(\sum_{i=1}^r P_i\rho P_i\Big\|\sum_{i=1}^r P_i\sigma P_i\right)\le D^q(\rho\|\sigma)$$

for any $\rho,\sigma\in\mathcal{B}(\mathcal{H})_{\gneq 0}$ and $P_1,\dots,P_r\in\mathbb{P}(\mathcal{H})$ such that $\sum_{i=1}^r P_i=I$.

**Lemma 3.2.2.** *Let $W\in\mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$ be a gcq channel that is classical on the support of some $P\in\mathcal{P}_f(\mathcal{X})$, i.e., there exists an ONB $(e_i)_{i=0}^{d-1}$ in $\mathcal{H}$ such that $W_x=\sum_{i=0}^{d-1}\widetilde{W}_x(i)|e_i\rangle\langle e_i|$, where $\widetilde{W}_x(i):=\langle e_i,W_xe_i\rangle$, $i\in[d]$, $x\in\operatorname{supp}P$. If all $D^{q_x}$, $x\in\operatorname{supp}P$, are monotone under pinchings then*

$$Q_P^{\mathrm{b},\mathbf{q}}(W)=\sum_{i\in\tilde{S}}\prod_{x\in\operatorname{supp}P}\widetilde{W}_x(i)^{P(x)},\tag{3.19}$$

*where $\tilde{S}:=\cap_{x\in\operatorname{supp}P}\operatorname{supp}\widetilde{W}_x$ and $\operatorname{supp}\widetilde{W}_x=\left\{i\in[d]:\widetilde{W}_x(i)>0\right\}$; moreover, there exists a unique optimal $\tau$ in* (3.1)*, given by*

$$\tau_P^{\mathbf{q}}(W):=\tau_P(\widetilde{W}):=\sum_{i\in S}|e_i\rangle\langle e_i|\prod_{x\in\operatorname{supp}P}\widetilde{W}_x(i)^{P(x)}\ .\tag{3.20}$$

*Proof.* If $S_+=0$ then $Q_P^{\mathrm{b},\mathbf{q}}(W)=0$, and the RHS of (3.19) is an empty sum, whence the equality in (3.19) holds trivially.

Thus, for the rest we assume that $S_+\neq 0$. Let $\mathcal{E}(\cdot):=\sum_{i=0}^{d-1}|e_i\rangle\langle e_i|(\cdot)|e_i\rangle\langle e_i|$ be the pinching corresponding to the joint eigenbasis of the $W_x$, $x\in\operatorname{supp}P$, guaranteed by the classicality assumption. For any $\tau\in\mathcal{B}(S_+\mathcal{H})_{\geq 0}$,

$$\begin{aligned}\operatorname{Tr}\tau-\sum_{x\in\mathcal{X}}P(x)\underbrace{D^{q_x}(\tau\|W_x)}_{\geq D^{q_x}(\mathcal{E}(\tau)\|\mathcal{E}(W_x))}&\leq\underbrace{\operatorname{Tr}\tau}_{=\operatorname{Tr}\mathcal{E}(\tau)}-\sum_{x\in\mathcal{X}}P(x)D^{q_x}(\mathcal{E}(\tau)\|\underbrace{\mathcal{E}(W_x)}_{=W_x})\\&=\operatorname{Tr}\mathcal{E}(\tau)-\sum_{x\in\mathcal{X}}P(x)D^{q_x}(\mathcal{E}(\tau)\|W_x)),\end{aligned}$$

where the inequality follows from the monotonicity of the $D^{q_x}$ under pinchings. Thus, the supremum in (3.1) can be restricted to $\tau$ operators that can be written as $\tau=\sum_{i=1}^d\tilde{\tau}(i)|e_i\rangle\langle e_i|$ with some $\tilde{\tau}(i)\in[0,+\infty)$, $i\in[d]$. Clearly, $\tau^0\leq S_+^0$ is equivalent to $\operatorname{supp}\tilde{\tau}\subseteq\tilde{S}$. For any such $\tau$,

$$\begin{aligned}\operatorname{Tr}\tau-\sum_{x\in\mathcal{X}}P(x)D^{q_x}(\tau\|W_x)&=\operatorname{Tr}\tau-\sum_{x\in\mathcal{X}}P(x)\sum_{i\in\tilde{S}}[\tilde{\tau}(i)\log\tilde{\tau}(i)-\tilde{\tau}(i)\log W_x(i)]\\&=\sum_{i\in\tilde{S}}\Big[\tilde{\tau}(i)-\tilde{\tau}(i)\log\tilde{\tau}(i)+\tilde{\tau}(i)\sum_{x\in\operatorname{supp}P}P(x)\log\widetilde{W}_x(i)\Big].\end{aligned}$$

The supremum of this over all such $\tau$ is

$$\sum_{i\in\tilde{S}}e^{\sum_{x\in\operatorname{supp}P}P(x)\log\widetilde{W}_x(i)}=\sum_{i\in\tilde{S}}\prod_{x\in\operatorname{supp}P}\widetilde{W}_x(i)^{P(x)},$$

which is uniquely attained at the $\tau=\tau_P^{\mathbf{q}}(W)$ given in (3.20), as for the function $[0,+\infty)\ni t\mapsto$

$t - t\log t - tc =: f(t)$ we have $f'(t) = -\log t - c = 0 \iff t = e^{-c}$, $f''(t) = -1/t < 0$, $t \in (0, +\infty)$. This proves (3.19). $\square$

**Corollary 3.2.3.** *In the setting of Lemma 3.2.2, the $P$-weighted left $D^{\mathbf{q}}$-radius of $W$ can be given explicitly as*

$$R_{D^{\mathbf{q}},\text{left}}(W, P) = -\log \sum_{i\in\tilde{S}} \prod_{x\in\operatorname{supp} P} \widetilde{W}_x(i)^{P(x)},$$

*and if $\tilde{S} \neq \emptyset$ then there is a unique $P$-weighted left $D^{\mathbf{q}}$-center for $W$, given by*

$$\omega_P^{\mathbf{q}}(W) := \frac{\tau^{\mathbf{q}}(W, P)}{\operatorname{Tr}\tau^{\mathbf{q}}(W, P)} = \sum_{i\in\tilde{S}} |e_i\rangle\langle e_i| \frac{\prod_{x\in\operatorname{supp} P} \widetilde{W}_x(i)^{P(x)}}{\sum_{j\in\tilde{S}} \prod_{x\in\operatorname{supp} P} \widetilde{W}_x(j)^{P(x)}} =: \omega_P(\widetilde{W}). \tag{3.21}$$

*Proof.* Immediate from Lemmas 3.1.9 and 3.2.2. $\square$

Lemma 3.2.2 yields immediately the following:

**Corollary 3.2.4.** *Assume that $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ commute, and hence can be written as $\rho = \sum_{i=1}^d \tilde{\rho}(i)|e_i\rangle\langle e_i|$, $\sigma = \sum_{i=1}^d \tilde{\sigma}(i)|e_i\rangle\langle e_i|$, in some ONB $(e_i)_{i=1}^d$. If $D^{q_0}$ and $D^{q_1}$ are monotone under pinchings then*

$$Q_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = Q_\alpha(\tilde{\rho}\|\tilde{\sigma}) = \sum_{i=1}^d \tilde{\rho}(i)^\alpha \tilde{\sigma}(i)^{1-\alpha}, \quad \alpha \in (0,1),$$

*and there exists a unique optimal $\tau$ in (3.4), given by*

$$\tau_\alpha^{\mathbf{q}}(\rho\|\sigma) := \tau_\alpha(\tilde{\rho}\|\tilde{\sigma}) := \sum_{i=1}^d |e_i\rangle\langle e_i| \tilde{\rho}(i)^\alpha \tilde{\sigma}(i)^{1-\alpha}. \tag{3.22}$$

As a special case of Corollary 3.2.3, we get the following:

**Corollary 3.2.5.** *In the setting of Corollary 3.2.4, if $\rho^0 \wedge \sigma^0 \neq 0$ then for every $\alpha \in (0,1)$ there exists a unique $\alpha$-weighted $D^{\mathbf{q}}$-center for $(\rho, \sigma)$, given by*

$$\omega_\alpha^{\mathbf{q}}(\rho\|\sigma) := \frac{\tau_\alpha^{\mathbf{q}}(\rho\|\sigma)}{\operatorname{Tr}\tau_\alpha^{\mathbf{q}}(\rho\|\sigma)} = \sum_{i=1}^d |e_i\rangle\langle e_i| \frac{\tilde{\rho}(i)^\alpha \tilde{\sigma}(i)^{1-\alpha}}{\sum_{j=1}^d \tilde{\rho}(j)^\alpha \tilde{\sigma}(j)^{1-\alpha}} =: \omega_\alpha(\tilde{\rho}\|\tilde{\sigma}). \tag{3.23}$$

*Proof.* Immediate from Corollary 3.2.4 and Lemma 3.1.9. $\square$

**Remark 3.2.6.** Note that $\tau_P^{\mathbf{q}}(W)$ in (3.20) and $\omega_P^{\mathbf{q}}(W)$ in (3.21) are independent of $D^{\mathbf{q}}$, as long as all $D^{q_x}$, $x \in \operatorname{supp} P$, are monotone under pinchings. Likewise, $\tau_\alpha^{\mathbf{q}}(\rho\|\sigma)$ in (3.22) and $\omega_\alpha^{\mathbf{q}}(\rho\|\sigma)$ in (3.23) are independent of $D^{q_0}$ and $D^{q_1}$, as long as both of them are monotone under pinchings.

Lemma 3.2.1 and Corollary 3.2.4 together give the following:

**Proposition 3.2.7.** *If $D^{q_0}$ and $D^{q_1}$ are two quantum relative entropies that are monotone under pinchings then for every $\alpha \in (0,1)$ the corresponding barycentric Rényi $\alpha$-divergence $D_\alpha^{\mathrm{b};\mathbf{q}}$ is a quantum Rényi $\alpha$-divergence in the sense of Definition 2.2.6.*

**Remark 3.2.8.** Note that in the classical case the barycentric Rényi $\alpha$-divergence is equal to the unique classical Rényi $\alpha$-divergence also for $\alpha > 1$; see (2.11). On the other hand, if $D^{q_0} \neq D^{q_1}$ then it may happen that $D_\alpha^{\mathrm{b};\mathbf{q}}$ is not a quantum Rényi $\alpha$-divergence for some $\alpha > 1$.

Note that for a fixed $i \in \cap_{x\in\operatorname{supp} P} \operatorname{supp} \widetilde{W}_x$, the expression $\prod_{x\in\operatorname{supp} P} \widetilde{W}_x(i)^{P(x)}$ in (3.19) is the weighted geometric mean of $(\widetilde{W}_x(i))_{x\in\operatorname{supp} P}$ with weights $(P(x))_{x\in\operatorname{supp} P}$. This motivates the following:

**Definition 3.2.9.** If $D^{\mathbf{q}}$, $W \in \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$, and $P \in \mathcal{P}_f^{\pm}(\mathcal{X})$ are such that there exists a unique optimizer $\tau =: \tau_P^{\mathbf{q}}(W)$ in (3.1) then this $\tau$ is called the $P$-weighted $D^{\mathbf{q}}$-geometric mean of $W$, and is also denoted by $G_P^{D^{\mathbf{q}}}(W) := \tau_P^{\mathbf{q}}(W)$.

Similarly, if there exists a unique optimizer $\tau =: \tau_\alpha^{\mathbf{q}}(\rho\|\sigma)$ in (3.4) then it is called the $\alpha$-weighted $D^{\mathbf{q}}$-geometric mean of $\rho$ and $\sigma$, and it is also denoted by $G_\alpha^{D^{\mathbf{q}}}(\rho\|\sigma) := \tau_\alpha^{\mathbf{q}}(\rho\|\sigma)$.

**Remark 3.2.10.** Note that if $G_P^{D^{\mathbf{q}}}(W)$ exists then by definition,

$$Q_P^{\mathrm{b};\mathbf{q}} = \operatorname{Tr} G_P^{D^{\mathbf{q}}}(W) = Q_P^{G_P^{D^{\mathbf{q}}}}(W)$$

(in particular, if $G_\alpha^{D^{\mathbf{q}}}(\rho\|\sigma)$ exists then $Q_\alpha^{\mathrm{b};\mathbf{q}}(\rho\|\sigma) = \operatorname{Tr} G_\alpha^{D^{\mathbf{q}}}(\rho\|\sigma)$).

In classical statistics, the family of states $(\omega_\alpha(\tilde\rho\|\tilde\sigma))_{\alpha\in(0,1)}$ given in (3.23) is called the Hellinger arc. (Note that if $\tilde\rho$ and $\tilde\sigma$ are probability distributions with equal supports then the Hellinger arc connects them in the sense that $\lim_{\alpha\searrow 0} \omega_\alpha(\tilde\rho\|\tilde\sigma) = \tilde\rho$, $\lim_{\alpha\nearrow 1} \omega_\alpha(\tilde\rho\|\tilde\sigma) = \tilde\sigma$.) This motivates the following:

**Definition 3.2.11.** Assume that $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ and $D^{q_0}$, $D^{q_1}$ are such that for every $\alpha \in (0,1)$ there exists a unique $\alpha$-weighted $D^{\mathbf{q}}$-center $\omega_\alpha^{\mathbf{q}}(\rho\|\sigma)$ for $(\rho,\sigma)$. Then $(\omega_\alpha^{\mathbf{q}}(\rho\|\sigma))_{\alpha\in(0,1)}$ is called the $D^{\mathbf{q}}$-Hellinger arc for $\rho$ and $\sigma$.

More generally, if $W$ and $D^{\mathbf{q}}$ are such that for every $P \in \mathcal{P}_f(\mathcal{X})$ there exists a unique $P$-weighted $D^{\mathbf{q}}$-center $\omega_P^{\mathbf{q}}(W)$ for $W$ then we call $(\omega_P^{\mathbf{q}}(W))_{P\in\mathcal{P}_f(\mathcal{X})}$ the $D^{\mathbf{q}}$-Hellinger body for $W$.

**Remark 3.2.12.** Note that by Lemma 3.1.9, for given $P \in \mathcal{P}_f^{\pm}(\mathcal{X})$, $W \in \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$, and $D^{\mathbf{q}}$, there exists a unique nonzero $P$-weighted $D^{\mathbf{q}}$-geometric mean $G_P^{D^{\mathbf{q}}}(W) = \tau_P^{\mathbf{q}}(W)$ if and only if there exists a unique $P$-weighted $D^{\mathbf{q}}$-center $\omega_P^{\mathbf{q}}(W)$ for $W$, and in this case we have

$$\begin{aligned}
\omega_P^{\mathbf{q}}(W) &= \frac{G_P^{D^{\mathbf{q}}}(W)}{\operatorname{Tr} G_P^{D^{\mathbf{q}}}(W)},\\
Q_P^{\mathbf{q}}(W) &= \operatorname{Tr} G_P^{D^{\mathbf{q}}}(W),\\
0 &= \sum_{x\in\mathcal{X}} P(x) D^{q_x}(G_P^{D^{\mathbf{q}}}(W)\|W_x),\\
-\log Q_P^{\mathbf{q}}(W) &= \sum_{x\in\mathcal{X}} P(x) D^{q_x}(\omega_P^{\mathbf{q}}(W)\|W_x).
\end{aligned}$$

The following is a multivariable generalization of [MO17, Theorem 3.6] :

**Proposition 3.2.13.** *Let $P \in \mathcal{P}_f^{\pm}(\mathcal{X})$, $W \in \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$, and for every $x \in \operatorname{supp} P$, let $D^{q_x} = D^{\mathrm{Um}}$. Assume that $0 \neq S_+ \leq S_-$. Then there exist a unique $P$-weighted $D^{\mathrm{Um}}$-geometric mean $G_P^{D^{\mathrm{Um}}}(W) := \tau_P^{\mathrm{Um}}(W)$ of $W$ and a unique $P$-weighted $D^{\mathrm{Um}}$-center $\omega_P^{\mathrm{Um}}$ for $W$, given by*

$$G_P^{D^{\mathrm{Um}}}(W) = \tau_P^{\mathrm{Um}}(W) = S_+ e^{\sum_{x\in\operatorname{supp} P} P(x) S_+ (\widehat{\log} W_x) S_+},$$

$$\omega_P^{\mathrm{Um}}(W) = \frac{G_P^{D^{\mathrm{Um}}}(W)}{\operatorname{Tr} G_P^{D^{\mathrm{Um}}}(W)} = \frac{S_+ e^{\sum_{x\in\operatorname{supp} P} P(x) S_+ (\widehat{\log} W_x) S_+}}{\operatorname{Tr} S_+ e^{\sum_{x\in\operatorname{supp} P} P(x) S_+ (\widehat{\log} W_x) S_+}},$$

*respectively, and*

$$-\log Q_P(W) = -\log \operatorname{Tr} G_P^{D^{\mathrm{Um}}}(W) = \sum_{x\in\mathcal{X}} D^{\mathrm{Um}}(\omega_P^{\mathrm{Um}}(W) \| W_x).$$

*Proof.* Note that for $\sigma := S_+ e^{\sum_{x\in\operatorname{supp} P} P(x) S_+ (\log W_x) S_+}$ and any $\tau \in \mathcal{B}(S_+\mathcal{H})_{\geq 0}$, we have

$$\begin{aligned} D^{\mathrm{Um}}(\tau\|\sigma) &= \operatorname{Tr}\tau\log\tau - \operatorname{Tr}\tau\log\left(S_+ e^{\sum_{x\in\operatorname{supp} P} P(x) S_+ (\widehat{\log} W_x) S_+}\right) \\ &= \operatorname{Tr}\tau\log\tau - \operatorname{Tr}\tau \sum_{x\in\operatorname{supp} P} P(x) \log W_x \\ &= \sum_{x\in\operatorname{supp} P} P(x) \underbrace{\left(\operatorname{Tr}\tau\log\tau - \operatorname{Tr}\tau\,\widehat{\log}\, W_x\right)}_{=D^{\mathrm{Um}}(\tau\|W_x)} \end{aligned} \tag{3.24}$$

Thus,

$$\begin{aligned} Q_P^{\mathrm{b},\mathbf{q}}(W) &= \sup_{\tau\in\mathcal{B}(S_+\mathcal{H})_{\geq 0}} \left\{ \operatorname{Tr}\tau - \sum_{x\in\operatorname{supp} P} P(x) D^{\mathrm{Um}}(\tau\|W_x) \right\} \\ &= \max_{\tau\in\mathcal{B}(S_+\mathcal{H})_{\geq 0}} \left\{ \operatorname{Tr}\tau - D^{\mathrm{Um}}(\tau\|\sigma) \right\} \\ &= \operatorname{Tr}\sigma, \end{aligned} \tag{3.25}$$

where the first equality is by definition, the second equality is by (3.24), and last equality is due to (2.34). Moreover, since $D^{\mathrm{Um}}$ is strictly trace monotone (see, e.g., [HMPB11, Proposition A.4]), Remark 2.2.19 yields that $\tau = \sigma$ is the unique state attaining the maximum in (3.25). This proves the assertion about the $P$-weighted $D^{\mathrm{Um}}$-geometric mean, and the rest of the assertions follow from this according to Remark 3.2.12. $\square$

**Corollary 3.2.14.** *Let $D^{q_x} = D^{\mathrm{Um}}$, $x \in \mathcal{X}$, and let $W \in \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$ be such that $\wedge_{x\in\mathcal{X}_0} W_x^0 \neq 0$ for any finite subset $W_0 \subseteq W$. Then the $D^{\mathrm{Um}}$-Hellinger body for $W$ exists.*

**Remark 3.2.15.** Note that for $P \in \mathcal{P}_f(\mathcal{X})$,

$$G_P^{D^{\mathrm{Um}}}(W) = \widehat{G}_{P,+\infty}(W) := W_P^0 e^{\sum_x P(x) \widehat{\log} W_x}.$$

### 3.3 Properties

One of the most important properties of quantum divergences is monotonicity under quantum operations (i.e., CPTP maps). Many of the important quantum divergences are monotone under more general trace-preserving maps, e.g., dual Schwarz maps in the case of Petz-type Rényi divergences for $\alpha \in [0,2]$ [Pet86b], or PTP maps in the case of the sandwiched Rényi divergences for $\alpha \geq 1/2$ [Bei13, Jen21, MR17], and the measured as well as the maximal Rényi divergences for $\alpha \in [0,+\infty]$, by definition. It is easy to see that for $\alpha \in [0,1]$, the barycentric Rényi $\alpha$-divergences are monotone under the same class of PTP maps as their generating quantum relative entropies. More generally, we have the following:

**Proposition 3.3.1.** *If all $D^{q_x}$, $x \in \mathcal{X}$, are monotone nonincreasing under a trace nondecreasing positive map $\Phi \in \mathrm{P}^+(\mathcal{H},\mathcal{K})$ then $Q^{\mathrm{b},\mathbf{q}}$ is monotone nondecreasing, and $R_{D^{\mathbf{q}},\mathrm{left}}$ is monotone nonincreasing under $\Phi$, i.e., for every gcq channel $W \in \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$ and every $P \in \mathcal{P}_f(\mathcal{X})$,*

$$Q_P^{\mathrm{b},\mathbf{q}}(\Phi(W)) \geq Q_P^{\mathrm{b},\mathbf{q}}(W), \tag{3.26}$$

$$R_{D^{\mathbf{q}},\mathrm{left}}(\Phi(W),P) \leq R_{D^{\mathbf{q}},\mathrm{left}}(W,P). \tag{3.27}$$

*Proof.* We have

$$\begin{aligned}
Q_P^{\mathrm{b},\mathbf{q}}(\Phi(W)) &= \sup_{\tau\in\mathcal{B}(\mathcal{K})_{\geq 0}} \left\{ \operatorname{Tr}\tau - \sum_{x\in\mathcal{X}} P(x) D^{q_x}(\tau\|\Phi(W_x)) \right\} \\
&\geq \sup_{\tilde\tau\in\mathcal{B}(\mathcal{H})_{\geq 0}} \left\{ \underbrace{\operatorname{Tr}\Phi(\tilde\tau)}_{\geq \operatorname{Tr}\tilde\tau} - \sum_{x\in\mathcal{X}} P(x) \underbrace{D^{q_x}(\Phi(\tilde\tau)\|\Phi(W_x))}_{\leq D^{q_x}(\tilde\tau\|W_x)} \right\} \\
&\geq \sup_{\tilde\tau\in\mathcal{B}(\mathcal{H})_{\geq 0}} \left\{ \operatorname{Tr}\tilde\tau - \sum_{x\in\mathcal{X}} P(x) D^{q_x}(\tilde\tau\|W_x) \right\} \\
&= Q_P^{\mathrm{b},\mathbf{q}}(W),
\end{aligned}$$

where the equalities are by definition (3.1) and by (3.7), the first inequality is obvious, and the second one follows from the assumptions. This proves (3.26), and (3.27) follows immediately by Lemma 3.1.9. □

Note that the normalized relative entropies $D_1^{q_0}$ and $D_1^{q_1}$ satisfy the scaling property (2.27) by assumption. This property is inherited by all the corresponding barycentric Rényi divergences $D_\alpha^{\mathbf{q}}$. More generally, we have the following:

**Lemma 3.3.2.** *For any $P \in \mathcal{P}_f^{\pm}(\mathcal{X})$, any gcq channel $W \in \mathcal{B}(\mathcal{X},\mathcal{H})_{\gneq 0}$ and any $t \in (0,+\infty)^{\mathcal{X}}$,*

$$Q_P^{\mathrm{b},\mathbf{q}}((t_x W_x)_{x\in\mathcal{X}}) = \left( \prod_{x\in\operatorname{supp} P} t_x^{P(x)} \right) Q_P^{\mathrm{b},\mathbf{q}}(W), \tag{3.28}$$

$$-\log Q_P^{\mathrm{b},\mathbf{q}}((t_x W_x)_{x\in\mathcal{X}}) = -\log Q_P^{\mathrm{b},\mathbf{q}}(W) - \sum_x P(x) \log t_x. \tag{3.29}$$

*In particular, $Q_\alpha^{\mathrm{b},\mathbf{q}}$ is homogeneous.*

*Proof.* (3.29) is straightforward to verify from the definition (3.2), and the scaling law (2.28), and (3.28) follows immediately from it. $\square$

**Corollary 3.3.3.** *The barycentric Rényi divergences satisfy the scaling law* (2.27)*, i.e.,*

$$D_\alpha^{\mathrm{b},\mathbf{q}}(t\rho\|s\sigma) = D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) + \log t - \log s,$$

*for every $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, $t, s \in (0, +\infty)$, $\alpha \in [0, +\infty]$.*

*Proof.* Immediate from Lemma 3.3.2, or alternatively, from (3.18). $\square$

We make use of the following proposition from [MBV22]

**Proposition 3.3.4.** *[Corollary V.40] [MBV22] We have*

$$D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = +\infty \begin{cases} \iff \rho^0 \wedge \sigma^0 = 0, & \text{when } \alpha \in [0,1), \\ \Longleftarrow \rho^0 \not\leq \sigma^0, & \text{when } \alpha > 1. \end{cases}$$

*If $D^{q_1}$ is anti-monotone in its second argument then the one-sided implication above is also an equivalence.*

This characterization of the finiteness of the 2-variable barycentric Rényi divergences gives an easily verifiable condition for a quantum Rényi $\alpha$-divergence not being a barycentric Rényi $\alpha$-divergence, as follows:

**Proposition 3.3.5.** *Let $D_\alpha^q$ be a quantum Rényi $\alpha$-divergence for some $\alpha \in (0,1)$ with the property that $D_\alpha^q(\rho\|\sigma) = +\infty \iff \rho \perp \sigma$. Then there exist no quantum relative entropies $D^{q_0}$ and $D^{q_1}$ with which $D_\alpha^q = D_\alpha^{\mathrm{b},\mathbf{q}}$.*

*Proof.* Let $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ be such that $\rho^0 \wedge \sigma^0 = 0$ and $\rho \not\perp \sigma$. Then

$$D_\alpha^{\mathrm{b},\mathbf{q}}(\rho\|\sigma) = +\infty > D_\alpha^q(\rho\|\sigma)$$

for any quantum relative entropies $D^{q_0}$ and $D^{q_1}$, according to Proposition 3.3.4. Since such pairs exist in any dimension larger than 1, we get that $D_\alpha^{\mathrm{b},\mathbf{q}} \neq D_\alpha^q$. $\square$

**Corollary 3.3.6.** *$D_{\alpha,z}$ is not a barycentric Rényi $\alpha$-divergence for any $\alpha \in (0,1)$ and $z \in (0,+\infty)$.*

*Proof.* It is obvious by definition that for any $\alpha \in (0,1)$ and $z \in (0,+\infty)$, and any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, $D_{\alpha,z}(\rho\|\sigma) = +\infty \iff \rho \perp \sigma$, and hence the assertion follows immediately from Proposition 3.3.5. $\square$

**Corollary 3.3.7.** *The measured Rényi $\alpha$-divergence $D_\alpha^{\mathrm{meas}}$ is not a barycentric Rényi $\alpha$-divergence for any $\alpha \in (0,1)$.*

*Proof.* According to Proposition 3.3.5 we only need to prove that for any $\alpha \in (0,1)$ and any $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$, $D_\alpha^{\mathrm{meas}}(\rho\|\sigma) = +\infty \iff \rho \perp \sigma$. This is well known and easy to verify, but we give the details for the readers' convenience. If $\rho \perp \sigma$ then the measurement $M_0 := \rho^0$, $M_1 := I - \rho^0$ gives $D_\alpha^{\mathrm{meas}}(\rho\|\sigma) \geq D_\alpha(M(\rho)\|M(\sigma)) = +\infty$. If $\rho \not\perp \sigma$ then we have $D_\alpha^{\mathrm{meas}}(\rho\|\sigma) \leq D_{\alpha,1}(\rho\|\sigma) < +\infty$, where the first inequality is due to the monotonicity of the Petz-type Rényi $\alpha$-divergence under measurements [Pet86b]. $\square$

### 3.4 Maximal Rényi divergences vs. the barycentric maximal Rényi divergences

The following lemma is trivial from the definitions.

**Lemma 3.4.1.** *If $P \in \mathcal{P}_f(\mathcal{X})$ and $D^{q_x} \leq D^{\tilde{q}_x}$, $x \in \operatorname{supp} P$, then*

$$-\log Q_P^{\mathrm{b},\mathbf{q}} \leq -\log Q_P^{\mathrm{b},\tilde{\mathbf{q}}}. \tag{3.30}$$

*In particular, if $\mathcal{X} = \{0,1\}$ then*

$$D^{q_0} \leq D^{\tilde{q}_0},\, D^{q_1} \leq D^{\tilde{q}_1} \implies D_\alpha^{\mathrm{b},\mathbf{q}} \leq D_\alpha^{\mathrm{b},\tilde{\mathbf{q}}}, \quad \alpha \in [0,1].$$

**Proposition 3.4.2.** *Let $P \in \mathcal{P}_f(\mathcal{X})$, and let $D^{q_x}$, $x \in \operatorname{supp} P$, be monotone under CPTP maps. Then*

$$-\log Q_P^{\mathrm{meas}} \leq -\log Q_P^{\mathrm{b,meas}} \leq -\log Q_P^{\mathrm{b},\mathbf{q}} \leq -\log Q_P^{\mathrm{b,max}} \leq -\log Q_P^{\mathrm{max}}. \tag{3.31}$$

*In particular, if $D^{q_0}$ and $D^{q_1}$ are quantum relative entropies that are monotone under CPTP maps then*

$$D_\alpha^{\mathrm{meas}} \leq D_\alpha^{\mathrm{b,meas}} \leq D_\alpha^{\mathrm{b},\mathbf{q}} \leq D_\alpha^{\mathrm{b,max}} \leq D_\alpha^{\mathrm{max}}, \quad \alpha \in [0,1]. \tag{3.32}$$

*Proof.* The second and the third inequalities in (3.31) are immediate from (3.30) and (2.26). Since $D^{\mathrm{meas}}$ and $D^{\mathrm{max}}$ are monotone under CPTP maps, so are $-\log Q_P^{\mathrm{b,meas}}$ and $-\log Q_P^{\mathrm{b,max}}$ as well, according to Proposition 3.3.1. Hence, the first and the last inequalities in (3.31) follow immediately from (2.15). The inequalities in (3.32) follow the same way. $\square$

By Proposition 3.4.2, for every $\alpha \in (0,1)$, $D_\alpha^{\mathrm{meas}} \leq D_\alpha^{\mathrm{b,meas}}$, but we have already seen in Corollary 3.3.7 that in fact $D_\alpha^{\mathrm{meas}}$ cannot be a barycentric Rényi $\alpha$-divergence, or in other words that all barycentric Rényi $\alpha$-divergences are larger than the minimal Rényi $\alpha$-divergence for a given $\alpha \in (0,1)$.

Also by Proposition 3.4.2, for every $\alpha \in (0,1)$, $D_\alpha^{\mathrm{b,max}} \leq D_\alpha^{\mathrm{max}}$. Our aim in this section is to show that equality does not hold. In fact, we conjecture the stronger relation that for noncommuting invertible PSD operators $\rho, \sigma$, $D_\alpha^{\mathrm{b,max}}(\rho\|\sigma) < D_\alpha^{\mathrm{max}}(\rho\|\sigma)$, $\alpha \in (0,1)$, which is supported by numerical examples. We will prove this below in the special case where the inputs are 2-dimensional. Of course, this already gives at least that

$$D_\alpha^{\mathrm{b,max}} \lneq D_\alpha^{\mathrm{max}}, \quad \alpha \in (0,1).$$

Let $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\gneq 0}$ be such that $\rho^0 \wedge \sigma^0 \neq 0$. Recall the definition of the maximal Rényi $\alpha$-divergence and the optimal reverse test $(\hat{p}, \hat{q}, \hat{\Gamma})$ from Example 2.2.12. Let $\omega_\alpha := \omega_\alpha(\hat{p}\|\hat{q}) =$

$\sum_{x\in\operatorname{supp}\hat q}\hat p(x)^\alpha \hat q(x)^{1-\alpha}\mathbf{1}_{\{x\}}$ be as in (3.23). Then for any $\alpha\in(0,1)$,

$$\begin{aligned}\frac{1}{\alpha-1}\log\operatorname{Tr}\rho + D_\alpha^{\max}(\rho\|\sigma) &= \frac{1}{\alpha-1}\log\operatorname{Tr}\hat p + D_\alpha(\hat p\|\hat q)\\ &= \frac{\alpha}{1-\alpha}D(\omega_\alpha\|\hat p) + D(\omega_\alpha\|\hat q)\\ &\geq \frac{\alpha}{1-\alpha}D^{\max}(\hat\Gamma(\omega_\alpha)\|\underbrace{\hat\Gamma(\hat p)}_{=\rho}) + D^{\max}(\hat\Gamma(\omega_\alpha)\|\underbrace{\hat\Gamma(\hat q)}_{=\sigma}) \qquad (3.33)\\ &= \frac{\alpha}{1-\alpha}D^{\max}(\hat\Gamma(\omega_\alpha)\|\rho) + D^{\max}(\hat\Gamma(\omega_\alpha)\|\sigma),\end{aligned}$$

where the first two equalities follow from Example 2.2.12 and (2.11)–(2.12), the inequality is due to the monotonicity of $D^{\max}$ under positive trace-preserving maps, and the third equality is by the definition of $\hat\Gamma$. By (2.12) and (2.18)–(2.20),

$$\begin{aligned}&\omega_\alpha = \sum_{i=1}^r \frac{\lambda_i^\alpha \operatorname{Tr}\sigma P_i}{Q_\alpha^{\max}(\rho\|\sigma)}1_{\{i\}},\\ &Q_\alpha^{\max}(\rho\|\sigma) = Q_\alpha(\hat p\|\hat q) = \sum_{i=1}^r \lambda_i^\alpha\operatorname{Tr}\sigma P_i = \operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha = \operatorname{Tr}\sigma\#_\alpha\rho,\\ &\hat\Gamma(\omega_\alpha) = \frac{1}{Q_\alpha^{\max}(\rho\|\sigma)}\sum_i \lambda_i^\alpha\sigma^{1/2}P_i\sigma^{1/2} = \frac{1}{Q_\alpha^{\max}(\rho\|\sigma)}\underbrace{\sigma^{1/2}(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha\sigma^{1/2}}_{=\sigma\#_\alpha\rho} =: \widehat{\sigma\#_\alpha\rho}, \qquad (3.34)\end{aligned}$$

where $\sigma\#_\alpha\rho$ is the $\alpha$-weighted Kubo-Ando geometric mean of $\rho$ and $\sigma$ (see Example 4.3.9 and the discussion before it).

The inequality in (3.33) is in fact an equality, according to the following:

**Lemma 3.4.3.** *Let $\rho,\sigma\in\mathcal{B}(\mathcal{H})_{\gneq 0}$ be such that $\rho^0\wedge\sigma^0\neq 0$, and assume that $\alpha\in(0,1)$, or that $\alpha\in(1,2]$ and $\rho^0\leq\sigma^0$. Then*

$$D_\alpha^{\max}(\rho\|\sigma) = \frac{\alpha}{1-\alpha}D^{\max}\left(\frac{\sigma\#_\alpha\rho}{Q_\alpha^{\max}(\rho\|\sigma)}\middle\|\rho\right) + D^{\max}\left(\frac{\sigma\#_\alpha\rho}{Q_\alpha^{\max}(\rho\|\sigma)}\middle\|\sigma\right) - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho. \qquad (3.35)$$

*Proof.* Let $Q_\alpha^{\max} := Q_\alpha^{\max}(\rho\|\sigma)$. Assume first that $\rho$ and $\sigma$ are invertible, and recall that in this case,

$$\sigma\#_\alpha\rho = \rho\#_{1-\alpha}\sigma = \rho^{1/2}(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha}\rho^{1/2}. \qquad (3.36)$$

Thus, by (2.21),

$$\begin{aligned}D^{\max}\left(\frac{\sigma\#_\alpha\rho}{Q_\alpha^{\max}(\rho\|\sigma)}\middle\|\rho\right) &= \frac{1}{Q_\alpha^{\max}}\operatorname{Tr}\rho(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha}\log\frac{(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha}}{Q_\alpha^{\max}}\\ &= -\frac{\log Q_\alpha^{\max}}{Q_\alpha^{\max}}\underbrace{\operatorname{Tr}\rho(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha}}_{=\operatorname{Tr}\sigma\#_\alpha\rho = Q_\alpha^{\max}}\\ &\quad+\frac{1-\alpha}{Q_\alpha^{\max}}\operatorname{Tr}\rho(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha}\log\left(\rho^{-1/2}\sigma\rho^{-1/2}\right)\\ &= -\log Q_\alpha^{\max} - \frac{1-\alpha}{Q_\alpha^{\max}}\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha\log\left(\sigma^{-1/2}\rho\sigma^{-1/2}\right), \qquad (3.37)\end{aligned}$$

where in the last equality we used that the transpose function of $f(\cdot) := (\cdot)^{1-\alpha}\log(\cdot)$ is $\tilde{f}(\cdot) := -(\cdot)^{\alpha}\log(\cdot)$, whence

$$\operatorname{Tr}\rho(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha}\log\Big(\rho^{-1/2}\sigma\rho^{-1/2}\Big) = -\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^{\alpha}\log\Big(\sigma^{-1/2}\rho\sigma^{-1/2}\Big), \tag{3.38}$$

according to the connection between the transpose function and the operator perspective function see ( [MBV22, Eq. (II.13)] and the discussion before it for example). Similarly,

$$\begin{aligned} D^{\max}\left(\frac{\sigma\#_\alpha\rho}{Q_\alpha^{\max}(\rho\|\sigma)}\middle\|\sigma\right) &= \frac{1}{Q_\alpha^{\max}}\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^{\alpha}\log\frac{(\sigma^{-1/2}\rho\sigma^{-1/2})^{\alpha}}{Q_\alpha^{\max}} \\ &= -\frac{\log Q_\alpha^{\max}}{Q_\alpha^{\max}}\underbrace{\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^{\alpha}}_{=Q_\alpha^{\max}} \\ &\quad + \frac{\alpha}{Q_\alpha^{\max}}\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^{\alpha}\log\Big(\sigma^{-1/2}\rho\sigma^{-1/2}\Big). \end{aligned} \tag{3.39}$$

From (3.37) and (3.39) we obtain that the RHS of (3.35) is

$$\frac{1}{\alpha-1}\log Q_\alpha^{\max} - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho = D_\alpha^{\max}(\rho\|\sigma),$$

where the equality is by definition. This proves (3.35) for invertible $\rho$ and $\sigma$.

Assume now that $\alpha \in (0,1)$, or that $\alpha \in (1,2]$ and $\rho^0 \le \sigma^0$. By the above, for any $\varepsilon > 0$, (3.35) holds with $\rho_\varepsilon := \rho + \varepsilon I$ and $\sigma_\varepsilon := \sigma + \varepsilon I$ in place of $\rho$ and $\sigma$, respectively, and taking the limit $\varepsilon \searrow 0$ yields (3.35), according to the properties of the Kubo-Ando mean (see [MBV22, Eq. IV.107-108.] for example), and (2.23). □

**Remark 3.4.4.** Using (3.37)–(3.39), a straightforward computation gives

$$\begin{aligned} D^{\max}(\hat{\Gamma}(\omega_\alpha)\|\rho) &= D(\omega_\alpha\|\hat{p}) = -\log Q_\alpha^{\max}(\rho\|\sigma) + (1-\alpha)\sum_i \omega_\alpha(i)\log\frac{\hat{q}(i)}{\hat{p}(i)}, \\ D^{\max}(\hat{\Gamma}(\omega_\alpha)\|\sigma) &= D(\omega_\alpha\|\hat{q}) = -\log Q_\alpha^{\max}(\rho\|\sigma) - \alpha\sum_i \omega_\alpha(i)\log\frac{\hat{q}(i)}{\hat{p}(i)}. \end{aligned}$$

From these, the equality in (3.33) can also be verified directly.

Our aim now is to prove that $\widehat{\sigma\#_\alpha\rho}$ is not an optimal $\omega$ in the variational formula (3.17) for $D_\alpha^{\mathrm{b},\max}$ when $\alpha \in (0,1)$. We prove this (at least in the 2-dimensional case) by showing that any state $\omega$ on the line segment connecting $\widehat{\sigma\#_\alpha\rho}$ and the maximally mixed state $\pi_{\mathcal{H}} := I/d$, $d := \dim\mathcal{H}$, that is close enough to $\widehat{\sigma\#_\alpha\rho}$ but is not equal to it, gives a strictly lower value than the RHS of (3.35) when substituted into $\frac{\alpha}{1-\alpha}D^{\max}(\cdot\|\rho) + D^{\max}(\cdot\|\sigma) - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho$.

**Lemma 3.4.5.** *Let $\rho,\sigma \in \mathcal{B}(\mathcal{H})_{>0}$, and let $P_1,\dots,P_r \in \mathbb{P}(\mathcal{H})$ and $\lambda_1,\dots,\lambda_r \in \mathbb{R}$, be such that $\sum_{i=1}^r P_i = I$, and*

$$\sigma^{-1/2}\rho\sigma^{-1/2} = \sum_{i=1}^r \lambda_i P_i.$$

*Then*

$$\partial_{\pi_{\mathcal{H}}} := \frac{d}{dt}\Big[\alpha D^{\max}\Big((1-t)\widehat{\sigma\#_\alpha\rho} + t\pi_{\mathcal{H}}\Big\|\rho\Big) + (1-\alpha)D^{\max}\Big((1-t)\widehat{\sigma\#_\alpha\rho} + t\pi_{\mathcal{H}}\Big\|\sigma\Big)\Big]\Bigg|_{t=0}$$

$$= -1 + \frac{1}{d}\sum_{i,j}\operatorname{Tr} P_i\sigma P_j\sigma^{-1}\underbrace{\Big[\alpha\log^{[1]}\big(\lambda_i^{\alpha-1},\lambda_j^{\alpha-1}\big)\lambda_i^{\alpha-1} + (1-\alpha)\log^{[1]}\big(\lambda_i^{\alpha},\lambda_j^{\alpha}\big)\lambda_i^{\alpha}\Big]}_{=:\Lambda_{\alpha,i,j}}, \quad (3.40)$$

*where*

$$\Lambda_{\alpha,i,j} = \begin{cases}\alpha(1-\alpha)(\log\lambda_i - \log\lambda_j)\frac{(\lambda_i-\lambda_j)}{(\lambda_i^\alpha-\lambda_j^\alpha)(\lambda_i^{1-\alpha}-\lambda_j^{1-\alpha})}, & \lambda_i \neq \lambda_j,\\ 1, & \lambda_i = \lambda_j.\end{cases} \quad (3.41)$$

*Proof.* Let us introduce the notation

$$(\rho/\sigma) := \sigma^{-1/2}\rho\sigma^{-1/2} = \sum_i \lambda_i P_i.$$

Let

$$X := \rho^{1/2}\sigma^{-1/2}, \text{ and } X = U|X| = U\sum_i \lambda_i^{1/2}P_i \quad (3.42)$$

be its polar decomposition. Then

$$\rho^{-1/2}\sigma\rho^{-1/2} = (X^{-1})^*(X^{-1}) = U\left(\sum_i \lambda_i^{-1}P_i\right)U^* = \sum_i \lambda_i^{-1}\underbrace{UP_iU^*}_{=:R_i} \quad (3.43)$$

is a spectral decomposition of $\rho^{-1/2}\sigma\rho^{-1/2}$. Recall from (3.34) that

$$\widehat{\sigma\#_\alpha\rho} := \frac{1}{Q_\alpha^{\max}}\sigma\#_\alpha\rho = \frac{1}{Q_\alpha^{\max}}\rho\#_{1-\alpha}\sigma,$$

where $Q_\alpha^{\max} := Q_\alpha^{\max}(\rho\|\sigma)$, and note the following identities:

$$\sigma^{-1/2}\widehat{\sigma\#_\alpha\rho}\,\sigma^{-1/2} = \frac{1}{Q_\alpha^{\max}}(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha = \sum_i \frac{\lambda_i^\alpha}{Q_\alpha^{\max}}P_i,$$

$$\rho^{-1/2}\widehat{\sigma\#_\alpha\rho}\,\rho^{-1/2} = \frac{1}{Q_\alpha^{\max}}(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha} = \sum_i \frac{\lambda_i^{\alpha-1}}{Q_\alpha^{\max}}R_i,$$

where in the last line we used (3.36).

Recall that $\pi_{\mathcal{H}} = I/d$ denotes the maximally mixed state on $\mathcal{H}$. We have

$$
\begin{aligned}
&\frac{d}{dt} D^{\max}\Big((1-t)\widehat{\sigma\#_\alpha\rho} + t\pi_{\mathcal{H}} \big\| \sigma\Big)\Big|_{t=0} \\
&= \operatorname{Tr} \underbrace{\sigma^{1/2}\Big(\pi_{\mathcal{H}} - \widehat{\sigma\#_\alpha\rho}\Big)\sigma^{-1/2}}_{=I/d - \frac{1}{Q_\alpha^{\max}}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha} \log \underbrace{\sigma^{-1/2}\widehat{\sigma\#_\alpha\rho}\,\sigma^{-1/2}}_{=\frac{1}{Q_\alpha^{\max}}(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha} \\
&+ \operatorname{Tr} \underbrace{\sigma^{1/2}\widehat{\sigma\#_\alpha\rho}\,\sigma^{-1/2}}_{=\frac{1}{Q_\alpha^{\max}}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha} \sum_{i,j} \log^{[1]}\left(\frac{\lambda_i^\alpha}{Q_\alpha^{\max}}, \frac{\lambda_j^\alpha}{Q_\alpha^{\max}}\right) \underbrace{P_i\sigma^{-1/2}\Big(\pi_{\mathcal{H}} - \widehat{\sigma\#_\alpha\rho}\Big)\sigma^{-1/2}P_j}_{=d^{-1}P_i\sigma^{-1}P_j - \frac{1}{Q_\alpha^{\max}}\delta_{i,j}\lambda_i^\alpha P_i} \\
&= -\log Q_\alpha^{\max} + \frac{\alpha}{d}\operatorname{Tr}\log\Big(\sigma^{-1/2}\rho\sigma^{-1/2}\Big) \\
&+ \frac{1}{Q_\alpha^{\max}}(\log Q_\alpha^{\max}) \underbrace{\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha}_{=Q_\alpha^{\max}} - \frac{\alpha}{Q_\alpha^{\max}}\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha \log\Big(\sigma^{-1/2}\rho\sigma^{-1/2}\Big) \\
&+ \frac{1}{dQ_\alpha^{\max}} \sum_{i,j} \log^{[1]}\left(\frac{\lambda_i^\alpha}{Q_\alpha^{\max}}, \frac{\lambda_j^\alpha}{Q_\alpha^{\max}}\right) \operatorname{Tr}\sigma \underbrace{(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha P_i\sigma^{-1}P_j}_{=\lambda_i^\alpha P_i\sigma^{-1}P_j} \\
&- \frac{1}{Q_\alpha^{\max}}\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha \sum_i \underbrace{\log^{[1]}\left(\frac{\lambda_i^\alpha}{Q_\alpha^{\max}}, \frac{\lambda_i^\alpha}{Q_\alpha^{\max}}\right)\frac{\lambda_i^\alpha}{Q_\alpha^{\max}}}_{=1} P_i \\
&= \frac{\alpha}{d}\operatorname{Tr}\log\Big(\sigma^{-1/2}\rho\sigma^{-1/2}\Big) - \frac{\alpha}{Q_\alpha^{\max}}\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha\log\Big(\sigma^{-1/2}\rho\sigma^{-1/2}\Big) \\
&+ \frac{1}{dQ_\alpha^{\max}}\sum_{i,j}\log^{[1]}\left(\frac{\lambda_i^\alpha}{Q_\alpha^{\max}}, \frac{\lambda_j^\alpha}{Q_\alpha^{\max}}\right)\lambda_i^\alpha \operatorname{Tr}\sigma P_i\sigma^{-1}P_j - \frac{1}{Q_\alpha^{\max}} \underbrace{\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^\alpha}_{=Q_\alpha^{\max}}
\end{aligned}
$$

An exactly analogous calculation yields

$$
\begin{aligned}
&\frac{d}{dt} D^{\max}\Big((1-t)\widehat{\sigma\#_\alpha\rho} + t\pi_{\mathcal{H}} \big\| \rho\Big)\Big|_{t=0} \\
&= \frac{1-\alpha}{d}\operatorname{Tr}\log\Big(\rho^{-1/2}\sigma\rho^{-1/2}\Big) - \frac{1-\alpha}{Q_\alpha^{\max}}\operatorname{Tr}\rho(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha}\log\Big(\rho^{-1/2}\sigma\rho^{-1/2}\Big) \\
&+ \frac{1}{dQ_\alpha^{\max}}\sum_{i,j}\log^{[1]}\left(\frac{\lambda_i^{\alpha-1}}{Q_\alpha^{\max}}, \frac{\lambda_j^{\alpha-1}}{Q_\alpha^{\max}}\right)\lambda_i^{\alpha-1}\operatorname{Tr}\rho R_i\rho^{-1}R_j - 1.
\end{aligned}
$$

Thus,

$$\begin{aligned}
\partial_{\pi_{\mathcal{H}}} &= \frac{d}{dt}\Big[\alpha D^{\max}\Big((1-t)\widehat{\sigma\#_\alpha\rho}+t\pi_{\mathcal{H}}\big\|\rho\Big)+(1-\alpha)D^{\max}\Big((1-t)\widehat{\sigma\#_\alpha\rho}+t\pi_{\mathcal{H}}\big\|\sigma\Big)\Big]\Big|_{t=0}\\
&= \frac{\alpha(1-\alpha)}{d}\underbrace{\Big[\underbrace{\operatorname{Tr}\log\Big(\rho^{-1/2}\sigma\rho^{-1/2}\Big)}_{=\sum_i\log\lambda_i^{-1}}+\underbrace{\operatorname{Tr}\log\Big(\sigma^{-1/2}\rho\sigma^{-1/2}\Big)}_{=\sum_i\log\lambda_i}\Big]}_{=0}\\
&\quad-\frac{\alpha(1-\alpha)}{Q_\alpha^{\max}}\underbrace{\Big[\operatorname{Tr}\rho(\rho^{-1/2}\sigma\rho^{-1/2})^{1-\alpha}\log\Big(\rho^{-1/2}\sigma\rho^{-1/2}\Big)+\operatorname{Tr}\sigma(\sigma^{-1/2}\rho\sigma^{-1/2})^{\alpha}\log\Big(\sigma^{-1/2}\rho\sigma^{-1/2}\Big)\Big]}_{=0}\\
&\quad+\frac{\alpha}{d}\sum_{i,j}\log^{[1]}\big(\lambda_i^{\alpha-1},\lambda_j^{\alpha-1}\big)\lambda_i^{\alpha-1}\operatorname{Tr}\rho R_i\rho^{-1}R_j+\frac{1-\alpha}{d}\sum_{i,j}\log^{[1]}\big(\lambda_i^{\alpha},\lambda_j^{\alpha}\big)\lambda_i^{\alpha}\operatorname{Tr}\sigma P_i\sigma^{-1}P_j\ -1\\
&= \frac{\alpha}{d}\sum_{i,j}\log^{[1]}\big(\lambda_i^{\alpha-1},\lambda_j^{\alpha-1}\big)\lambda_i^{\alpha-1}\operatorname{Tr}\rho R_i\rho^{-1}R_j+\frac{1-\alpha}{d}\sum_{i,j}\log^{[1]}\big(\lambda_i^{\alpha},\lambda_j^{\alpha}\big)\lambda_i^{\alpha}\operatorname{Tr}\sigma P_i\sigma^{-1}P_j\ -1,
\end{aligned}\tag{3.44}$$

where the first expression above is equal to 0 due to (3.43), and the second expression is equal to 0 according to (3.38).

Note that by (3.42),

$$U=X|X|^{-1}=\rho^{1/2}\sigma^{-1/2}(\rho/\sigma)^{-1/2},$$

hence

$$\begin{array}{ccc}
U^* & = & (\rho/\sigma)^{-1/2}\sigma^{-1/2}\rho^{1/2}\\
\| & & \\
U^{-1} & = & (\rho/\sigma)^{1/2}\sigma^{1/2}\rho^{-1/2},
\end{array}$$

which in turn yields

$$U=(U^{-1})^*=\rho^{-1/2}\sigma^{1/2}(\rho/\sigma)^{1/2}.$$

Thus,

$$\begin{aligned}
&\operatorname{Tr}\rho R_i\rho^{-1}R_j\\
&=\operatorname{Tr}\rho(UP_iU^*)\rho^{-1}(UP_jU^*)\\
&=\operatorname{Tr}\rho\,\underbrace{\rho^{-1/2}\sigma^{1/2}(\rho/\sigma)^{1/2}}_{=U}P_i\underbrace{(\rho/\sigma)^{-1/2}\sigma^{-1/2}\rho^{1/2}}_{=U^*}\rho^{-1}\underbrace{\rho^{1/2}\sigma^{-1/2}(\rho/\sigma)^{-1/2}}_{=U}P_j\underbrace{(\rho/\sigma)^{1/2}\sigma^{1/2}\rho^{-1/2}}_{=U^*}\\
&=\operatorname{Tr}\sigma^{1/2}\underbrace{(\rho/\sigma)^{1/2}P_i(\rho/\sigma)^{-1/2}}_{=P_i}\sigma^{-1}\underbrace{(\rho/\sigma)^{-1/2}P_j(\rho/\sigma)^{1/2}}_{=P_j}\sigma^{1/2}\\
&=\operatorname{Tr}\sigma P_i\sigma^{-1}P_j.
\end{aligned}$$

Writing this back into (3.44), we get

$$\partial_{\pi_{\mathcal{H}}} = -1 + \frac{1}{d}\sum_{i,j} \operatorname{Tr} \sigma P_i \sigma^{-1} P_j \underbrace{\left[\alpha \log^{[1]}\left(\lambda_i^{\alpha-1}, \lambda_j^{\alpha-1}\right)\lambda_i^{\alpha-1} + (1-\alpha)\log^{[1]}\left(\lambda_i^{\alpha}, \lambda_j^{\alpha}\right)\lambda_i^{\alpha}\right]}_{=:\Lambda_{\alpha,i,j}}. \quad (3.45)$$

It follows by a straightforward computation that $\Lambda_{\alpha,i,j}$ can be written as in (3.41). Note that $\Lambda_\alpha$ is symmetric, i.e., $\Lambda_{\alpha,i,j} = \Lambda_{\alpha,j,i}$. Exchanging the indices $i$ and $j$ in (3.45) yields (3.40). □

Our aim is therefore to prove that $\partial_{\pi_{\mathcal{H}}} < 0$. For this, we will need the following:

**Lemma 3.4.6.** *The following equivalent inequalities are true: for every $\alpha \in (0,1)$,*

$$\frac{\log\lambda - \log\eta}{\lambda - \eta} > \frac{1}{\alpha}\frac{\lambda^\alpha - \eta^\alpha}{\lambda - \eta} \cdot \frac{1}{1-\alpha}\frac{\lambda^{1-\alpha} - \eta^{1-\alpha}}{\lambda - \eta}, \qquad \lambda, \eta \in (0,+\infty),\ \lambda \neq \eta, \quad (3.46)$$

$$\frac{\log x}{x-1} > \frac{1}{\alpha}\frac{x^\alpha - 1}{x-1}\frac{1}{1-\alpha}\frac{x^{1-\alpha}-1}{x-1}, \qquad x \in (0,+\infty)\setminus\{1\},$$

$$\int_0^1 \frac{1}{tx+1-t}\,dt > \int_0^1 \frac{1}{(tx+1-t)^\alpha}\,dt \int_0^1 \frac{1}{(tx+1-t)^{1-\alpha}}\,dt, \qquad x \in (0,+\infty)\setminus\{1\}. \quad (3.47)$$

*Proof.* It is straightforward to verify that the above inequalities are equivalent to each other. The inequality in (3.47) follows from the strict concavity of the power functions, as

$$\int_0^1 \frac{1}{(tx+1-t)^\gamma}\,dt < \left(\int_0^1 \frac{1}{tx+1-t}\,dt\right)^\gamma, \quad \gamma \in (0,1).$$

□

**Corollary 3.4.7.** *In the setting of Lemma 3.4.5,*

$$\Lambda_{\alpha,i,j} > 1, \quad i \neq j.$$

*Proof.* Immediate from (3.46). □

Note that we may take the $P_i$ in Lemma 3.4.5 to be rank 1, i.e., $P_i = |e_i\rangle\langle e_i|$, $i = 1,\ldots,d$, for some orthonormal eigenbasis of $\sigma^{-1/2}\rho\sigma^{-1/2}$. Then (3.40) can be rewritten as

$$\begin{aligned}\partial_{\pi_{\mathcal{H}}} &= -1 + \frac{1}{d}\sum_{i,j} \underbrace{\langle e_i, \sigma e_j\rangle}_{=:S_{i,j}} \underbrace{\langle e_j, \sigma^{-1} e_i\rangle}_{=(S^{-1})^{\mathrm{T}}_{i,j}} \cdot \Lambda_{\alpha,i,j} \\ &= -1 + \langle u, (S \star (S^{-1})^{\mathrm{T}} \star \Lambda_\alpha) u\rangle,\end{aligned} \quad (3.48)$$

where $u = \frac{1}{\sqrt{d}}(1,1,\ldots,1)$ and $A \star B$ denotes the component-wise (also called Hadamard, or Schur) product of two matrices $A$ and $B$.

Next, note that

$$(S^{-1})_{j,i} = (-1)^{i+j}\frac{\det([S]_{i,j})}{\det S},$$

where $[S]_{i,j}$ is the matrix that we get by omitting the $i$-th row and $j$-th column of $S$. Thus, (3.48) can be rewritten as

$$\partial_{\pi_{\mathcal{H}}} = -1 + \frac{1}{d}\sum_{i=1}^{d}\frac{1}{\det S}\sum_{j=1}^{d}(-1)^{i+j}S_{i,j}\det([S]_{i,j})\Lambda_{\alpha,i,j}.$$

Note that for every $i$,

$$\frac{1}{\det S}\sum_{j=1}^{d}(-1)^{i+j}S_{i,j}\det([S]_{i,j}) = (SS^{-1})_{i,i} = 1.$$

**Theorem 3.4.8.** *Let $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{>0}$, where $\dim\mathcal{H} = 2$, and assume that $\rho\sigma \neq \sigma\rho$. Then*

$$D_\alpha^{\mathrm{b,max}}(\rho\|\sigma) < D_\alpha^{\mathrm{max}}(\rho\|\sigma), \quad \alpha \in (0,1).$$

*Proof.* By Corollary 3.3.3, we may assume that $\operatorname{Tr}\rho = \operatorname{Tr}\sigma = 1$. By the above, it is sufficient to prove that $\partial_{\pi_{\mathcal{H}}} < 0$. Let $(e_1, e_2)$ be an orthonormal eigenbasis of $\sigma^{-1/2}\rho\sigma^{-1/2}$ with corresponding eigenvalues $\lambda_1, \lambda_2$. By assumption, $\rho\sigma \neq \sigma\rho$, which implies that $\lambda_1 \neq \lambda_2$. (In fact, $\lambda_1 = \lambda_2 \iff \rho = c\sigma$ for some $c > 0$, in which case $c = \lambda_1 = \lambda_2$.) Writing out everything in the ONB $(e_1, e_2)$, we have

$$S = \frac{1}{2}\begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix}$$

with some $r := (x,y,z) \in \mathbb{R}^3$ such that $\|r\|^2 = x^2+y^2+z^2 < 1$, and

$$(S^{-1})^{\mathrm{T}} = \frac{4}{1-\|r\|^2}\cdot\frac{1}{2}\begin{bmatrix} 1-z & -x-iy \\ -x+iy & 1+z \end{bmatrix},$$

whence

$$S \star (S^{-1})^{\mathrm{T}} = \frac{1}{1-\|r\|^2}\begin{bmatrix} 1-z^2 & -(x^2+y^2) \\ -(x^2+y^2) & 1-z^2 \end{bmatrix}.$$

Hence, by (3.48) and the symmetry $\Lambda_{\alpha,1,2} = \Lambda_{\alpha,2,1}$,

$$\partial_{\pi_{\mathcal{H}}} = -1 + \frac{1}{1-\|r\|^2}\Big[1 - z^2 - (x^2+y^2)\Lambda_{\alpha,1,2}\Big].$$

Since $\sigma$ is not diagonal in the given ONB (otherwise it would commute with $\rho$), we have $(x^2+y^2) > 0$. Combining this with Corollary 3.4.7, we get $\partial_{\pi_{\mathcal{H}}} < 0$, as required. $\square$

# 4 An axiomatic derivation of quantum Rényi divergences

In this chapter we turn to the study of relative submajorization of pairs of families $(\rho, \sigma)$. We rely on Section 2.3 and in particular on Theorem 2.3.26 to characterize relaxations of relative submajorization by monotone functions. To use Theorem 2.3.26, we first define the precise setup and show that the preorder of relative submajorization on the semiring of pairs of families satisfies the required technical conditions outlined in Section 2.3. Then we provide a classification of the relevant monotones when restricted to classical families, to then extend the classification to arbitrary $\rho$ and commuting $\sigma$, and construct some monotones for general pairs.

In the last section of this chapter we discuss some application of the results of this chapter. Relative submajorization between $(\rho, \sigma)$ and constant functions $(1-\alpha, \beta)$ is equivalent to $(\alpha, \beta)$ being achievable as a pair of type I and type II errors in hypothesis testing with $\rho$ and $\sigma$ as composite hypotheses of states of full support (following the idea of [Ren16]). This in turn means that asymptotic relative submajorization is an adequate way to characterize achievable exponents in the strong converse regime. In Subsection 4.4.1 we give necessary and sufficient conditions for pairs of error exponents being achievable for the strong converse regime of hypothesis testing with composite hypotheses as above, with the additional requirement, that the operators in $\sigma$ commute with each other. These conditions are given in terms of pair-wise sandwiched Rényi divergences between states from $\rho$ and naive geometric means of the states from $\sigma$.

In the rest of the last section of this chapter we describe in more detail how equivariant relative submajorization (including non-compact groups) can be encoded as the relative submajorization of certain families, with applications to timetranslation symmetric Gibbs-preserving maps and to groupsymmetric hypothesis testing. We relate a type of approximate asymptotic transformation to asymptotic relative submajorization. Then using the monotones in the fully quantum case, we find a new family of monotone quantum Renyi divergences.

## 4.1 The preordered semiring of pairs of families

Let $X, Y$ be nonempty compact Hausdorff topological spaces. Let $(\rho, \sigma)$ be a pair of continuous maps, where $\rho : X \to \mathcal{B}(\mathcal{H})_{>0}$ and $\sigma : Y \to \mathcal{B}(\mathcal{H})_{>0}$ for some finite dimensional Hilbert space $\mathcal{H}$. Two pairs $(\rho, \sigma)$ and $(\rho', \sigma')$ are equivalent if there is a unitary $U : \mathcal{H} \to \mathcal{H}'$ such that $\forall x \in X : U\rho(x)U^* = \rho'(x)$ and $\forall y \in Y : U\sigma(y)U^* = \sigma'(y)$. Let $S_{X,Y,\mathcal{H}}$ denote the set of equivalence classes of pairs of such families for a given finite dimensional Hilbert space $\mathcal{H}$. Let $S_{X,Y} = \cup_{\mathcal{H}} S_{X,Y,\mathcal{H}}$.

**Definition 4.1.1** (Pairs of families)**.** We will call elements of $S_{X,Y}$ pairs of families (of positive operators).

The pointwise direct sum and tensor product operations are well-defined on equivalence classes, and turn $S_{X,Y}$ into a commutative semiring. The zero element is represented by the unique pair over a zero dimensional Hilbert space, while 1 is represented by the pair consisting of constant functions with value $I$ over $\mathbb{C}$.

$X, Y$ will be the index sets for the families. I adopt the convention that if any map $\rho : X \to \mathcal{B}(\mathcal{H})_{>0}$ and $\sigma : Y \to \mathcal{B}(\mathcal{H})_{>0}$ appear outside of brackets, then any operations or relations they appear in, are to be understood pointwise, including sum, product, direct sum, tensor product or image under a linear super operator, usually notated by $T$. More precisely, given any map $T : \mathcal{B}(\mathcal{H})_{>0} \to \mathcal{B}(\mathcal{H}')_{>0}$ the composition is to be understood as $T(\rho) : x \mapsto T(\rho(x))$ and $T(\sigma) : y \mapsto T(\sigma(y))$.

**Definition 4.1.2.** $(\rho, \sigma)$ relatively submajorizes $(\rho', \sigma')$, in notation $(\rho, \sigma) \succcurlyeq (\rho', \sigma')$, if there exists a completely positive trace-nonincreasing map $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ such that $T(\rho) \geq \rho'$ and $T(\sigma) \leq \sigma'$.

**Remark 4.1.3.** For the relative submajorization or any relaxations of it, I will note the completely positive trace-nonincreasing map explicitly if it needs to be emphasized: $(\rho, \sigma) \succcurlyeq^T (\rho', \sigma')$.

**Example 4.1.4.** Let $X$ and $Y$ be one-point spaces and let $\rho, \sigma$ be states of full support over some Hilbert space $\mathcal{H}$ and let $\rho', \sigma'$ be states of full support over some Hilbert space $\mathcal{K}$. Then $(\rho, \sigma) \succcurlyeq (\rho', \sigma')$ translates to relative majorization between the pairs of states.

(i) If we further specify $\mathcal{H} = \mathcal{H}'$ and $\sigma = \sigma' = \frac{I_{\mathcal{H}}}{\dim \mathcal{H}}$, then $(\rho, \sigma) \succcurlyeq (\rho', \sigma')$ translates to the so-called majorization preorder between $\rho$ and $\rho'$, meaning there is a unital channel mapping $\rho$ to $\rho'$.

(ii) If we specify instead $\sigma = \sigma' = \frac{e^{-\beta H}}{\operatorname{Tr} e^{-\beta H}}$ to be Gibbs states for the Hamiltonian $H$, then $(\rho, \sigma) \succcurlyeq (\rho', \sigma')$ translates to the so-called thermomajorization preorder [HO13] between $\rho$ and $\rho'$, meaning there is a Gibbs preserving channel mapping $\rho$ to $\rho'$.

(iii) We will see in Proposition 4.4.4 that for the constant functions $(1 - \alpha)_X, \beta_Y$ on $X$ and $Y$, respectively, $(\rho, \sigma) \succcurlyeq ((1 - \alpha)_X, \beta_Y)$ will translate to $(\alpha, \beta)$ being achievable as a pair of tpye I and type II error in a composite hypothesis testing problem.

**Proposition 4.1.5.** *$S_{X,Y}$ is a preordered semiring with relative submajorization.*

*Proof.* Identity and the composition of two completely positive trace-nonincreasing maps are once again completely positive trace-nonincreasing maps for which the conditions in Definition 4.1.2 are met and thus $\succcurlyeq$ is a reflexive and transitive relation, i.e., a preorder. We need to verify that the preorder is compatible with the semiring operations. Suppose that $(\rho, \sigma) \succcurlyeq (\rho', \sigma')$ and let $T$ be a completely positive trace nonincreasing map as in Definition 4.1.2. Let $(\omega, \tau) \in S_{X,Y}$ be a pair of families on $\mathcal{K}$, i.e., $\omega : X \to \mathcal{B}(\mathcal{K})_{>0}$ and $\tau : Y \to \mathcal{B}(\mathcal{K})_{>0}$. Then

$$
\begin{aligned}
(T \otimes \mathrm{id}_{\mathcal{B}(\mathcal{K})})(\rho \otimes \omega) &= T(\rho) \otimes \omega \geq \rho' \otimes \omega \\
(T \otimes \mathrm{id}_{\mathcal{B}(\mathcal{K})})(\sigma \otimes \tau) &= T(\sigma) \otimes \tau \leq \sigma' \otimes \tau,
\end{aligned}
$$

therefore $(\rho, \sigma)(\omega, \tau) \succcurlyeq (\rho', \sigma')(\omega, \tau)$.

The map $\tilde{T} : \mathcal{B}(\mathcal{H} \oplus \mathcal{K}) \to \mathcal{B}(\mathcal{H}' \oplus \mathcal{K})$ defined as

$$
\tilde{T}\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) = \begin{bmatrix} T(A) & 0 \\ 0 & D \end{bmatrix}
$$

is also completely positive and trace nonincreasing, and satisfies

$$\tilde{T}(\rho \oplus \omega) = T(\rho) \oplus \omega \geq \rho' \oplus \omega$$
$$\tilde{T}(\sigma \oplus \tau) = T(\sigma) \oplus \tau \leq \sigma' \oplus \tau,$$

therefore $(\rho, \sigma) + (\omega, \tau) \succcurlyeq (\rho', \sigma') + (\omega, \tau)$. $\square$

**Proposition 4.1.6.** *$S_{X,Y}$, is of polynomial growth and $u = (b_X, 1_Y)$ is power universal for any constant $b > 1$.*

*Proof.* Recall the definition of the power universal element from Definition 2.3.6. For the pair of families $(\rho, \sigma)$ let us choose the substochastic map $T(.) := c_1 \operatorname{Tr}(.)$, with $c_1 = \min\{1, [\max_{y \in Y} \operatorname{Tr} \sigma(y)]^{-1}\}$. Then we have $c_1 \operatorname{Tr} \sigma \leq 1$ and $T(u^k(\rho, \sigma)) = (T(b^k \rho), T(\sigma)) = (c_1 b^k \operatorname{Tr} \rho, c_1 \operatorname{Tr} \sigma)$. Choosing a large enough $k$ will ensure $c_1 b^k \operatorname{Tr} \rho \geq 1$, since $\operatorname{Tr} \rho$ is bounded away from 0 on $X$ and so $\exists k \in \mathbb{N} : u^k(\rho, \sigma) \succcurlyeq (1, 1)$. For $d := \dim \mathcal{H}$, let us choose now $T(.) := c_2(.) \otimes \frac{1}{d} I_{\mathbb{C}^d}$, where we set $c_2 = \min\{1, d[\min_{y \in Y} \min(\operatorname{spec}(\sigma(y))]\}$. Then we have $\sigma \geq \frac{c_2}{d} I_{\mathbb{C}^d}$ and $T(u^k) = (T(b^k), T(1)) = (\frac{b^k c_2}{d} I_{\mathbb{C}^d}, \frac{c_2}{d} I_{\mathbb{C}^d})$. Choosing a large enough $k$ will ensure $\frac{b^k c_2}{d} I_{\mathbb{C}^d} \geq \rho$, since $\max \operatorname{spec}(\rho)$ is bounded on $X$ and so $\exists k \in \mathbb{N} : u^k \succcurlyeq (\rho, \sigma)$. $\square$

$S_{X,Y}$ is then a semiring of polynomial growth and in $S_{X,Y}$ we have $0 \preccurlyeq 1$ and thus Theorem 2.3.22 and Corollary 2.3.23 are applicable.

**Remark 4.1.7.** Recall from Definition 2.3.6 and the subsequent discussion that, even if the power universal element is not unique the definition of the asymptotic preorder, given in Definition 2.3.8 is independent of a particular choice. With a few, explicitly stated exceptions we will use $u = (2_X, 1_Y)$.

It will be useful to consider to subsemiring of $S_{X,Y}$, which we will call classical and semi-classical subsemiring.

**Definition 4.1.8.** The subsemiring generated by the pairs of families of one-dimensional positive operators is called the subsemiring of classical families, in notation $S^{\mathrm{c}}_{X,Y}$. It is easy to see that $(\rho, \sigma) \in S^{\mathrm{c}}_{X,Y}$ if and only if $[\rho(x), \rho(x')] = [\rho(x), \sigma(y)] = [\sigma(y), \sigma(y')] = 0,\ \forall x, x' \in X,\ \forall y, y' \in Y$.

**Definition 4.1.9.** The elements $(\rho, \sigma) \in S_{X,Y}$ satisfying $[\sigma(y), \sigma(y')] = 0\ \forall y, y' \in Y$ form a subsemiring of $S_{X,Y}$, which I will call the semi-classical subsemiring and denote with $S^{\mathrm{sc}}_{X,Y}$.

**Remark 4.1.10.** It is easy to check that $S^{\mathrm{c}}_{X,Y}$ and $S^{\mathrm{sc}}_{X,Y}$ equipped with the relative submajorization are indeed preordered semirings, they are subsemirings of $S_{X,Y}$ and since $u = (2_X, 1_Y) \in S^{\mathrm{c}}_{X,Y}$ it is a power universal element of both subsemirings as well.

### 4.2 Classical families

We turn to the classification of the 1-test spectrum (see Definition 2.3.25), which is practically the real and tropical real valued homomorphisms, on the subsemiring of classical families. By definition every element in $S^{\mathrm{c}}_{X,Y}$ is a sum of one-dimensional elements. A one-dimensional element of the semiring on the other hand can be identified by a pair of strictly positive continuous functions on $X$ and $Y$.

Suppose that $f$ is a multiplicative map from the one-dimensional pairs into either the real or the tropical numbers. Then the extension of $f$ to multidimensional pairs via additivity also enjoys multiplicativity. Since $S^{\mathrm{c}}_{X,Y}$ is generated by the one-dimensional pairs, the value of every $f \in \mathrm{TSper}_1(S, \preccurlyeq)$ is determined by its behavior on one-dimensional pairs.

**Proposition 4.2.1.** *If $f \in \mathrm{TSper}_1(S^{\mathrm{c}}_{X,Y})$ then there exists unique, nonnegative Radon measures $\mu$ and $\nu$ on $X$ and $Y$ such that for every multidimensional classical pair $(\bigoplus_{i=1}^d p_i, \bigoplus_{i=1}^d q_i)$ $(p_i \in C(X), q_i \in C(Y))$, if $f$ is real valued, it admits the form*

$$\sum_{i=1}^{d} \exp\biggl(\int_X \log p_i \,\mathrm{d}\mu - \int_Y \log q_i \,\mathrm{d}\nu\biggr), \qquad \mu(X) - \nu(Y) = 1, \tag{4.1}$$

*while if $f$ is tropical valued it admits the form*

$$\max_{i\in[d]} \exp\biggl(\int_X \log p_i \,\mathrm{d}\mu - \int_Y \log q_i \,\mathrm{d}\nu\biggr), \qquad \mu(X) - \nu(Y) = 0. \tag{4.2}$$

*Proof.* Let $f \in \mathrm{TSper}_1(S^{\mathrm{c}}_{X,Y})$ be an element of the spectrum. For every $\xi, \eta > 0$ one has $(e^{\xi}, 1_Y) \succcurlyeq (1_X, 1_Y)$ and $(1_X, e^{-\eta}) \succcurlyeq (1_X, 1_Y)$, thus the maps $\xi \mapsto \log f(e^{\xi}, 1_Y)$, from $C(X)$ to $\mathbb{R}$ and $\eta \mapsto \log f(1_X, e^{-\eta})$, from $C(Y)$ to $\mathbb{R}$ are well defined positive linear functionals on $C(X)$ and $C(Y)$. Thus by Theorem 2.1.1, $\log f(e^{\xi}, 1_Y) = \int_X \xi(x) \,\mathrm{d}\mu(x)$ and $\log f(1_Y, e^{-\eta}) = \int_Y \eta(y) \,\mathrm{d}\nu(y)$ for some unique $\mu, \nu$ Radon measures on $X$ and $Y$. Since $f$ is multiplicative $\log f(e^{\xi}, e^{-\eta}) = \log f(e^{\xi}, 1_Y) + \log f(1_X, e^{-\eta}) = \int_X \xi \,\mathrm{d}\mu + \int_Y \eta \,\mathrm{d}\nu$. From this $f(p,q) = \exp\bigl(\int_X \log p \,\mathrm{d}\mu - \int_Y \log q \,\mathrm{d}\nu\bigr)$. We used only the multiplicative property of $f$ but not the additive property, thus this part of the proof works for either real or tropical valued elements of the spectrum. Consider now $f(1_X + 1_X, 1_Y + 1_Y) = f(1_X, 1_Y) + f(1_X, 1_Y)$. In the real case this translates to $f(1_X{+}1_X, 1_Y{+}1_Y) = 2$, in the tropical case to $f(1_X{+}1_X, 1_Y{+}1_Y) = 1$. Then $t \mapsto \log f(e^t 1_X, e^t 1_Y)$ is additive, normalized and monotone, therefore satisfies Cauchy's functional equation and admits the form $\log f(e^t 1_X, e^t 1_Y) = t$ in the real case and the form $\log f(e^t 1_X, e^t 1_Y) = 0$ in the tropical case. This leads to $1 = \log f(e1_X, e1_Y) = \mu(X) - \nu(Y)$ in the real case and $0 = \log f(e1_X, e1_Y) = \mu(X) - \nu(Y)$ in the tropical case.

Now from additivity any real or tropical valued element of the spectrum admits the forms (4.1) and (4.2). $\square$

**Remark 4.2.2.** Functions of the form (4.1) are homogeneous of degree 1 and functions of the form (4.2) are homogeneous of degree 0. From the previous proposition it follows that in particular this is true for elements of the spectrum going to the real and tropical numbers respectively.

**Proposition 4.2.3.** *Functions of the form* (4.1) *and* (4.2) *are elements of the spectrum if and only if they satisfy the data-processing inequality.*

*Proof.* Elements of the spectrum need to be monotone under relative submajorization, so in particular relative majorization. This shows that satisfying the data-processing inequality is a necessary condition. Let us assume that $f(p,q)$ is of the form of (4.1) or (4.2) and $f$ is monotone under relative majorization. Let us assume that $(p', q') \in S^{\mathrm{c}}_{X,Y}$ and $(p', q') \preccurlyeq^{T} (p, q)$. We need to show that $f(p', q') \leq f(p, q)$. Now if $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$, then let us complete $T$ into a completely positive trace-preserving map $\tilde{T}:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H}'')$ by adding an extra dimension to $\mathcal{H}'$. We arrive to

$$(p,q)\succcurlyeq^{\tilde{T}}(\tilde{T}(p),\tilde{T}(q))\succcurlyeq(T(p),T(q))\succcurlyeq^{\mathrm{id}}(p',q'),$$

where the first ordering is by definition, the second ordering is realized by the map that erases the extra dimension (this is a completely positive trace-nonincreasing map) and the last ordering is by assumption. Let us apply f,

$$f(p,q)\geq f(\tilde{T}(p),\tilde{T}(q))\geq f(T(p),T(q))\geq f(p',q'),$$

where the first inequality is by assumption of $f$ satisfying the data-processing inequality, the second and last inequality follows from the forms (4.1) and (4.2), since these functions are monotone decreasing under deletion of an index and increase of any of the $q_i$ or decrease of any of the $p_i$. □

We will use the following lemma to relate the DPI to the joint convexity property of elements of the spectrum. This lemma is stated in a more specialized way than the quite generally stated Lemma 2.2.34, however it also treats functions going to the tropical numbers, an aspect that we need.

**Lemma 4.2.4.** *Let $f$ be an additive function from $S_{X,Y}$ into either the real or tropical numbers. Then*

*(i) if $f$ is homogeneous of degree 1 and $f$ goes into the real numbers, then it satisfies the data-processing inequality if and only if it is jointly convex;*

*(ii) if $f$ is homogeneous of degree 0 and $f$ goes into the tropical numbers, then it satisfies the data-processing inequality if and only if it is jointly quasi-convex.*

*Proof.* Let $f$ be an additive function from $S_{X,Y}$ into either the real or tropical numbers and let $f$ be homogeneous of degree $k$. Whenever the $\sum$ symbol is outside of $f$ let it stand as summing with respect to the semiring: usual summing in the real case and maximum in the tropical case. Suppose $f$ is monotone under quantum channels. Applying monotonicity to $\hat{\rho}:=\sum_i p_i|i\rangle\langle i|_E\otimes\rho_i$ and $\hat{\sigma}:=\sum_i p_i|i\rangle\langle i|_E\otimes\sigma_i$ under the partial trace $\mathrm{Tr}_E$, where $(|i\rangle)_{i=1}^r$ is an ONS in $\mathcal{H}_E$ yields

$$\sum_i p_i^k f(\rho_i,\sigma_i)=f\left(\sum_i p_i|i\rangle\langle i|_E\otimes\rho_i,\sum_i p_i|i\rangle\langle i|_E\otimes\sigma_i\right)\geq f\left(\sum_i p_i\rho_i,\sum_i p_i\sigma_i\right),$$

which translates to joint convexity when $k=1$ and $f$ goes into the real numbers and joint quasi-convexity, when $k=0$ and $f$ goes into the tropical numbers. Suppose now that $f$ is homogeneous of degree $k$ and it is jointly convex in the real case or jointly quasi-convex in the tropical case. Using Stinespring dilation $\Phi(.)=\mathrm{Tr}_E\,V(.)V^*$ with an isometry $V:\mathcal{H}\to\mathcal{K}\otimes\mathcal{H}_E$, and writing the partial trace multiplied by the maximally mixed state as a convex combination of unitary conjugations (e.g., by the discrete Weyl unitaries):

$$
\begin{aligned}
f(\Phi(\rho),\Phi(\sigma)) &= \sum_{i=1}^{d_E} \frac{1}{d_E^k} f(\Phi(\rho),\Phi(\sigma)) \\
&= f\left(\frac{1}{d_E} I_E \otimes \operatorname{Tr}_E V(\rho)V^*, \frac{1}{d_E} I_E \otimes \operatorname{Tr}_E V(\sigma)V^*\right) \\
&= f\left(\sum_i \frac{1}{n} U_i V \rho V^* U_i^*, \sum_i \frac{1}{n} U_i V \sigma V^* U_i^*\right) \\
&\leq \frac{1}{n^k} \sum_{i=1}^{n} f(\rho,\sigma) = f(\rho,\sigma).
\end{aligned}
$$

□

**Proposition 4.2.5.** *If functions of the form* (4.1) *and* (4.2) *satisfy the data processing inequality then the measure $\mu$ is concentrated on one point.*

*Proof.* By Lemma 4.2.4 we require functions of the form (4.1) and (4.2) to be jointly convex and jointly quasi-convex. In particular both family of functions needs to be jointly quasiconvex in the one dimensional special case. These functions are totally differentiable and if we restrict $f$ to a line segment then having a zero directional derivative and negative second derivative would mean strict local maximum and would contradict quasiconvexity. Consider the general directional derivative of $f$ at 1. The forms of $f$ in (4.1) and (4.2) are differentiable and the derivative of the integrands are continuous on $X$ and $Y$ and thus bounded. Then by [Fol99, Theorem 2.27] the differentiation and the integration commute.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} f(1_X + s\xi, 1_Y)\Big|_{s=0} &= \frac{\mathrm{d}}{\mathrm{d}s}\left[\exp \int_X \log(1_X + s\xi)\,\mathrm{d}\mu\right]\Big|_{s=0} \\
&= f(1_X + s\xi, 1_Y)\left[\int_X \frac{\xi}{1_X + s\xi}\,\mathrm{d}\mu\right]\Big|_{s=0} = \int_X \xi\,\mathrm{d}\mu
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}s^2} f(1_X + s\xi, 1_Y)\Big|_{s=0} &= \frac{\mathrm{d}^2}{\mathrm{d}s^2}\left[\exp \int_X \log(1_X + s\xi)\,\mathrm{d}\mu\right]\Big|_{s=0} \\
&= f(1_X + s\xi, 1_Y)\left[\left(\int_X \frac{\xi}{1_X + s\xi}\,\mathrm{d}\mu\right)^2 - \int_X \left(\frac{\xi}{1_X + s\xi}\right)^2 \mathrm{d}\mu\right]\Big|_{s=0} \\
&= \left(\int_X \xi\,\mathrm{d}\mu\right)^2 - \int_X \xi^2\,\mathrm{d}\mu.
\end{aligned}
$$

To get a contradiction with quasiconvexity, we need to find a continuous function $\xi$ whose integral is zero and such that $\xi^2$ has nonzero integral, supposing that $\mu$ is not concentrated on a point. To this end let $x_1, x_2$ be distinct points in the support of $\mu$. Choose disjoint closed neighborhoods $A_1$ and $A_2$ (possible since $X$ is a compact Hausdorff space). By Urysohn's lemma, there exist continuous functions $\xi_1, \xi_2 : X \to [0,1]$ such that $\xi_1$ is 1 on $A_1$ and 0 on $A_2$, and $\xi_2$ is 0 on $A_1$ and 1 on $A_2$. Let

$$\xi = \left(\int_X \xi_2 \,\mathrm{d}\mu\right)\xi_1 - \left(\int_X \xi_1 \,\mathrm{d}\mu\right)\xi_2.$$

By construction, the integral of $\xi$ vanishes, while

$$\int_X \xi^2 \,\mathrm{d}\mu \geq \mu(A_1)\left(\int_X \xi_2 \,\mathrm{d}\mu\right)^2 > 0.$$

□

**Corollary 4.2.6.** *Let $\alpha := \mu(X)$ and $\gamma := \frac{\nu}{\nu(Y)}$. Taking Propositions 4.2.1 and 4.2.5 into account an element of the 1-test spectrum must have one of the following forms.*

$$f\left(\bigoplus_{i=1}^{d} p_i, \bigoplus_{i=1}^{d} q_i\right) = \sum_{i=1}^{d} p_i(x)^{\alpha} \exp\left[(1-\alpha)\int_Y \log q_i \,\mathrm{d}\gamma\right] \tag{4.3}$$

*if $f$ goes into the reals and*

$$f\left(\bigoplus_{i=1}^{d} p_i, \bigoplus_{i=1}^{d} q_i\right) = \max_{i\in[d]} p_i(x) \exp\left[-\int_Y \log q_i \,\mathrm{d}\gamma\right], \tag{4.4}$$

*if $f$ goes into the tropicals, where we further took into account that $f(u) = 2$ is required for elements of the spectrum going into the tropical numbers.* (4.3) *is characterized by the point $x$, the weight $\alpha \geq 1$ and the probability measure $\gamma$, while* (4.4) *is characterized by the point $x$ and the probability measure $\gamma$.*

**Proposition 4.2.7.** (4.3) *and* (4.4) *satisfy the data processing inequality.*

*Proof.* By Lemma 4.2.4 the functions (4.3) are monotone under channels if and only if they are jointly convex. Since the sum of such functions also enjoy joint convexity, it is sufficient to show joint convexity in the 1-dimensional case. Joint convexity then equivalently translates to the second directional derivative being nonnegative at any point in any direction. Let us compute the general directional second derivative relying on the commutativity of the differentiation and integration by [Fol99, Theorem 2.27].

$$f(p + s\xi, q + s\eta) = f(p,q) f(1_X + s\tilde{\xi}, 1_Y + s\tilde{\eta}),$$

with $\tilde{\xi} := \frac{\xi}{p}$ and $\tilde{\eta} := \frac{\eta}{q}$. Then the second directional derivative of (4.3) is

$$\begin{aligned}
&\frac{\mathrm{d}^2}{\mathrm{d}s^2} f(p+s\xi, q+s\eta)\bigg|_{s=0} \\
=& f(p,q)\frac{\mathrm{d}}{\mathrm{d}s} f(1_X + s\tilde{\xi}, 1_Y + s\tilde{\eta})\left[\alpha\frac{\tilde{\xi}(x)}{1_X + s\tilde{\xi}(x)} + (1-\alpha)\int_Y \frac{\tilde{\eta}}{1_Y + s\tilde{\eta}}\,\mathrm{d}\gamma\right]\bigg|_{s=0} \\
=& f(p,q)\left[\left(\alpha\tilde{\xi}(x) + (1-\alpha)\int_Y \tilde{\eta}\,\mathrm{d}\gamma\right)^2 - \alpha\tilde{\xi}(x)^2 - (1-\alpha)\int_Y \tilde{\eta}^2\,\mathrm{d}\gamma\right] \\
=& f(p,q)(\alpha-1)\left[\alpha\tilde{\xi}(x)^2 - 2\alpha\tilde{\xi}(x)\int_Y \tilde{\eta}\,\mathrm{d}\gamma + \int_Y \tilde{\eta}^2\,\mathrm{d}\gamma + (\alpha-1)\left(\int_Y \tilde{\eta}\,\mathrm{d}\gamma\right)^2\right] \\
\geq& f(p,q)(\alpha-1)\alpha\left[\tilde{\xi}(x)^2 - 2\tilde{\xi}(x)\int_Y \tilde{\eta}\,\mathrm{d}\gamma + \left(\int_Y \tilde{\eta}\,\mathrm{d}\gamma\right)^2\right] \\
=& f(p,q)(\alpha-1)\alpha\left(\tilde{\xi}(x) - \int_Y \tilde{\eta}\,\mathrm{d}\gamma\right)^2 \geq 0,
\end{aligned}$$

where we used that the second moment of $\tilde{\eta}$ is greater than the square of the first moment. We conclude that the functions (4.3) are jointly convex and thus satisfy the data processing inequality, they are monotone decreasing under stochastic maps. Now for a function $f$ of the form (4.3) consider

$$\begin{aligned}
g_\alpha\left(\bigoplus_{i=1}^d p_i, \bigoplus_{i=1}^d q_i\right) :=& \left(\sum_{i=1}^d f(p_i, q_i)\right)^{\frac{1}{\alpha}} \\
=& \left(\sum_{i=1}^d p_i^\alpha(x)\exp\left[(1-\alpha)\int_Y \log q_i\,\mathrm{d}\gamma\right]\right)^{\frac{1}{\alpha}}.
\end{aligned} \tag{4.5}$$

$g_\alpha$ then satisfies the data processing inequality and preserves this property in the $\alpha \to \infty$ limit. However

$$\lim_{\alpha\to\infty} g_\alpha\left(\bigoplus_{i=1}^d p_i, \bigoplus_{i=1}^d q_i\right) = \max_{i\in[d]} p_i(x)\exp\left[-\int_Y \log q_i\,\mathrm{d}\gamma\right],$$

showing that functions of the form (4.4) satisfy the data processing inequality too. □

Note that functions of the form (4.3) can be viewed as the $\alpha$-Rényi quasidivergences of a positive vector $p$ and some pointwise geometric mean of positive vectors $q_i$. What we used in the last proof is that the max divergence, i.e., functions in (4.4), can be given as a $\alpha \to \infty$ limit of Rényi divergences.

**Remark 4.2.8.** In [Fri23, 7.3 Definition.] the so-called *logarithmic evaluation map* is introduced and defined by

$$\mathrm{lev}_x : \mathrm{TSper}_1(S, \preccurlyeq) \to [0,\infty), \qquad \phi \mapsto \frac{\log\phi(x)}{\log\phi(u)}.$$

It is continuous for all nonzero $x \in S$. We note that with our notation above in (4.5), $\log g_\alpha(p,q) = \mathrm{lev}_{(p,q)}(f)$.

**Theorem 4.2.9.** $\mathrm{TSper}_1(S^{\mathrm{c}}_{X,Y})$ *consists of the functions*

$$f_{\alpha,x,\gamma}\left(\bigoplus_{i=1}^{d} p_i, \bigoplus_{i=1}^{d} q_i\right) = \sum_{i=1}^{d} p_i(x)^{\alpha} \exp\left[(1-\alpha)\int_Y \log q_i \,\mathrm{d}\gamma\right], \tag{4.6}$$

*where $\alpha \geq 1$, $x \in X$ and $\gamma$ is a probability measure on $Y$, if $f$ is real-valued and*

$$f_{x,\gamma}\left(\bigoplus_{i=1}^{d} p_i, \bigoplus_{i=1}^{d} q_i\right) = \max_{i\in[d]} p_i(x) \exp\left[-\int_Y \log q_i \,\mathrm{d}\gamma\right], \tag{4.7}$$

*where, $x \in X$ and $\gamma$ is a probability measure on $Y$, if $f$ is tropical real-valued.*

*Proof.* Follows from Propositions 4.2.1, 4.2.5 and 4.2.7. □

We will be referring to elements of $\mathrm{TSper}_1(S^{\mathrm{c}}_{X,Y})$ as $f_{(\alpha),x,\gamma}$ or $g_{(\alpha),x,\gamma}$, if we want distinct multiple elements of the spectrum, signifying the characterizing quantities $x, \gamma$ and possibly $\alpha$ as well, the same time.

If $X$ and $Y$ are one-point spaces, then the elements of the spectrum are exponentiated Rényi divergences. With normalized arguments, these are bounded from below by 1, with equality if and only if the two arguments are the same. We conclude this section with an analogous statement for the spectral points on the classical semiring for general spaces.

**Proposition 4.2.10.** *Let $\alpha > 1$, $x \in X$ and $\gamma$ a probability measure on $Y$. Suppose that $(p,q) \in S^{\mathrm{c}}_{X,Y}$ is normalized, i.e., $\sum_{i=1}^{d} p_i(x') = 1$ and $\sum_{i=1}^{d} q_i(y') = 1$ for all $x' \in X$ and $y' \in Y$. Then $f_{\alpha,x,\gamma}(p,q) \geq 1$ with equality if and only if $p(x) = q(y)$ for all $y \in \operatorname{supp}\gamma$. Similarly, $f_{x,\gamma}(p,q) \geq 1$ with equality if and only if $p(x) = q(y)$ for all $y \in \operatorname{supp}\gamma$.*

*Proof.* The inequality follows from monotonicity under the stochastic map $p \mapsto \sum_{i=1}^{d} p_i$ and the normalization $f(1_X, 1_Y) = 1$. Let

$$\bar{q}_i = \exp \int_Y \log q_i \,\mathrm{d}\gamma,$$

so that

$$f_{\alpha,x,\gamma}(p,q) = \sum_{i=1}^{d} p_i(x)^{\alpha} \bar{q}_i^{1-\alpha} = 2^{(\alpha-1)D_{\alpha}(p\|\bar{q})}.$$

By the Jensen inequality,

$$\sum_{i=1} \bar{q}_i \leq \sum_{i=1} \int_Y q_i \,\mathrm{d}\gamma = \int_Y \underbrace{\sum_{i=1} q_i}_{1_Y} \mathrm{d}\gamma = 1.$$

Since the Rényi divergence is anti-monotone in the second argument and strictly positive when the arguments are distinct probability distributions, the equality $f_{\alpha,x,\gamma}(p,q) = 1$ is equivalent to $p(x) = \bar{q}$. This means that the Jensen inequality holds with equality which, by strict concavity of the logarithm, implies $q(y) = \bar{q} = p$ for all $y \in \operatorname{supp}\gamma$.

Similarly, if $f_{x,\gamma}(p,q) = 1$ then for all $i$ we have $p_i(x) \leq \bar{q}_i$ but $\sum_{i=1}^{d} p_i(x) = 1 \geq \sum_{i=1} \bar{q}_i$, which implies $p(x) = \bar{q}$. □

### 4.3 Quantum extensions

In Section 4.1 we defined the preordered semiring $S_{X,Y}$, the set of pairs of families of positive operators equipped with the realative submajorization as a preorder. In Section 4.2 we introduced the classical subsemiring $S^{\mathrm{c}}_{X,Y}$ of $S_{X,Y}$ and we succeeded in characterizing the whole set of the spectrum $\mathrm{TSper}_1(\mathrm{S}^{\mathrm{c}}_{\mathrm{X,Y}}, \succcurlyeq)$, that is the homomorphisms from $S^{\mathrm{c}}_{X,Y}$ into the real numbers or the normalized homomorphisms from $S^{\mathrm{c}}_{X,Y}$ to the tropical numbers. Our ultimate goal is to use Theorem 2.3.26 in the most general case possible of our setting.

By Theorem 4.2.9 we could already apply Theorem 2.3.26 in its full force to $S^{\mathrm{c}}_{X,Y}$, this would allow us to fully characterize asymptotic relative submajorization and give sufficient conditions for catalytic relative submajorization between pairs of families, assuming that all the operators commute with each other. This would already make it available, for example, to fully characterize the so called strong converse regime of hypothesis testing, with composite hypotheses in the form of continuous families of states. This characterization would be in terms of the classical Rényi divergences derived in Theorem 4.2.9. In Subsection 4.4.1, for example, we will discuss this problem, however with less restriction. We will only need to assume that the operators in the second family $\sigma$ commute with each other and we will be able to assume general noncommutativity between any other pairs of operators. For this purpose and to be able to use Theorem 2.3.26 in the most general case possible we need to try and characterize further elements of $\mathrm{TSper}_1(\mathrm{S}_{\mathrm{X,Y}}, \succcurlyeq)$ in this section.

**Proposition 4.3.1.** *Suppose that $\tilde{f}$ is an element of the spectrum. Let $f_{(\alpha),x,\gamma}$ be $\tilde{f}$ constrained on the classical semiring according to* (4.6) *or* (4.7) *in Theorem 4.2.9. Then for any $\rho, \rho' : X \to \mathcal{B}(\mathcal{H})_{>0}$ such that $\rho(x) = \rho'(x)$ and for any $\sigma, \sigma' : Y \to \mathcal{B}(\mathcal{H})_{>0}$ such that $\sigma(y') = \sigma'(y')$ for all $y' \in \operatorname{supp}\gamma$, it follows that $\tilde{f}(\rho, \sigma) = \tilde{f}(\rho', \sigma')$.*

*Proof.* Let

$$\begin{aligned} c_1(x') &= \min\{t : \rho'(x') \leq t\rho(x')\} \\ d_1(y') &= \max\{t : \sigma'(y') \geq t\sigma(y')\} \\ c_2(x') &= \min\{t : \rho(x') \leq t\rho'(x')\} \\ d_2(y') &= \max\{t : \sigma(y') \geq t\sigma'(y')\}. \end{aligned}$$

Then $c_1, c_2, d_1, d_2$ are strictly positive continuous functions on $X$ and $Y$, respectively, and

$$(\rho', \sigma') \preccurlyeq (c_1, d_1)(\rho, \sigma) \preccurlyeq (c_1, d_1)(c_2, d_2)(\rho', \sigma').$$

$c_1(x) = c_2(x) = 1$ and $d_1(y') = d_2(y') = 1$ for all $y' \in \operatorname{supp}\gamma$, $(c_1, d_1)$ and $(c_2, d_2)$ are classical pairs and thus $\tilde{f}(c_1, d_1) = f_{(\alpha),x,\gamma}(c_1, d_1) = 1$ and $\tilde{f}(c_1, d_1) = f_{(\alpha),x,\gamma}(c_2, d_2) = 1$. Applying now $\tilde{f}$ to all three parts of the above inequality yields

$$\tilde{f}(\rho', \sigma') \leq \tilde{f}(\rho, \sigma) \leq \tilde{f}(\rho', \sigma').$$

□

The following proposition classifies the test-spectrum of $S^{\mathrm{sc}}_{X,Y}$.

**Proposition 4.3.2.** *Let $(\rho,\sigma)\in S_{X,Y}$ and let $\tilde{f}$ be a real element of the spectrum and let $f_{\alpha,x,\gamma}$ be its restriction onto the classical subsemiring. Let $\tilde{g}$ be a tropical element of the spectrum and let $g_{x,\gamma}$ be its restriction onto the classical subsemiring. If $[\sigma(y),\sigma(y')]=0\ \forall y,y'\in Y$ then*

$$\tilde{f}(\rho,\sigma)=Q^*_\alpha\left(\rho(x)\middle\|\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)$$

*and*

$$\tilde{g}(\rho,\sigma)=\left\|\rho^{\frac{1}{2}}(x)\left(\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)^{-1}\rho^{\frac{1}{2}}(x)\right\|_\infty.$$

*Proof.* There is a positive definite operator $\tilde{\sigma}$ such that the eigenbasis of $\tilde{\sigma}$ simultaneously diagonalizes all $\sigma(y)$. Let $\mathcal{P}_{\tilde{\sigma}_n}$ denote the pinching by $\tilde{\sigma}^{\otimes n}$, then $\mathcal{P}_{\tilde{\sigma}_n}$ leaves $\sigma^{\otimes n}(y)$ invariant for all $y\in Y$. It follows that

$$\begin{aligned}(\rho,\sigma)^n &\succcurlyeq \left(\mathcal{P}_{\tilde{\sigma}_n}\left(\rho^{\otimes n}\right),\mathcal{P}_{\tilde{\sigma}_n}\left(\sigma^{\otimes n}\right)\right)\\ &=\left(\mathcal{P}_{\tilde{\sigma}_n}\left(\rho^{\otimes n}\right),\sigma^{\otimes n}\right)\\ &\succcurlyeq\left(\frac{1}{\mathrm{poly}(n)},1\right)\left(\rho^{\otimes n},\sigma^{\otimes n}\right),\end{aligned}$$

where $\mathrm{poly}(n)$ is a polynomial of $n$ and we used that any pinching is a completely positive trace preserving map and the pinching inequality:

$$\rho^{\otimes n}\le|\operatorname{spec}(\tilde{\sigma}^{\otimes n})|\mathcal{P}_{\tilde{\sigma}_n}\left(\rho^{\otimes n}\right)=\mathrm{poly}(n)\mathcal{P}_{\tilde{\sigma}_n}\left(\rho^{\otimes n}\right).$$

We have shown in Proposition 4.3.1 that $\tilde{f}$ and $\tilde{g}$ only depend on one point of $\rho$ apart from $\sigma$, but after the pinching all these operators commute and thus we are evaluating $\tilde{f}$ and $\tilde{g}$ on the classical subsemiring, where $\tilde{f}$ and $\tilde{g}$ are determined by Theorem 4.2.9. Applying $f$ to all three parts and taking the $n$-th root yields

$$\begin{aligned}\tilde{f}(\rho,\sigma)&\ge\sqrt[n]{f_{\alpha,x,\gamma}(\mathcal{P}_{\tilde{\sigma}_n}(\rho^{\otimes n}(x)),\sigma^{\otimes n})}\\ &=\sqrt[n]{\operatorname{Tr}(\mathcal{P}_{\tilde{\sigma}_n}(\rho^{\otimes n}))^\alpha\left(\exp\int_Y\log\sigma^{\otimes n}\,\mathrm{d}\gamma\right)^{1-\alpha}}\\ &=\sqrt[n]{\operatorname{Tr}(\mathcal{P}_{\tilde{\sigma}_n}(\rho^{\otimes n}(x)))^\alpha\left(\left(\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)^{\otimes n}\right)^{1-\alpha}}\\ &\ge\sqrt[n]{\frac{1}{(\mathrm{poly}(n))^\alpha}}\tilde{f}(\rho,\sigma).\end{aligned}$$

Taking the limit $n\to+\infty$ gives us

$$\tilde{f}(\rho(x),\sigma)\ge Q^*_\alpha\left(\rho\middle\|\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)\ge\tilde{f}(\rho,\sigma),\tag{4.8}$$

where we refer to [Tom16, Proposition 4.12.] (see also [PVW22, Theorem 4.4.]) in taking the limit of the middle term.

Now applying $\tilde{g}$ to all three parts yields

$$\begin{aligned}\tilde{g}(\rho,\sigma) &\geq \sqrt[n]{g_{x,\gamma}(\mathcal{P}_{\tilde{\sigma}_n}(\rho^{\otimes n}(x)),\sigma^{\otimes n})} \\ &= \sqrt[n]{\left\|(\mathcal{P}_{\tilde{\sigma}_n}(\rho^{\otimes n}(x)))\left(\exp\int_Y \log\sigma^{\otimes n}\,\mathrm{d}\gamma\right)\right\|_\infty} \\ &= \sqrt[n]{\left\|(\mathcal{P}_{\tilde{\sigma}_n}(\rho^{\otimes n}(x)))\left(\left(\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)^{\otimes n}\right)\right\|_\infty} \\ &\geq \sqrt[n]{\frac{1}{\mathrm{poly}(n)}}\tilde{g}(\rho,\sigma).\end{aligned}$$

Taking the limit $n\to+\infty$ gives us

$$\tilde{g}(\rho,\sigma)\geq\left\|\rho^{\frac{1}{2}}(x)\left(\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)^{-1}\rho^{\frac{1}{2}}(x)\right\|_\infty\geq\tilde{g}(\rho,\sigma), \tag{4.9}$$

where we refer to [Dat09] and [Tom16, Section 4.2.4] in taking the limit of the middle term. ☐

**Corollary 4.3.3.** *Suppose that $\tilde{f}$ is an element of the spectrum and let $f_{\alpha,x,\gamma}$ be its restriction to the classical semiring. If $\sigma(y)=\sigma_0$ for all $y\in\operatorname{supp}\gamma$, then $\tilde{f}(\rho,\sigma)=Q^*_\alpha(\rho(x)\|\sigma_0)$.*

*Proof.* By Proposition 4.3.1, $\tilde{f}(\rho,\sigma)$ is equal to $\tilde{f}(\rho,\bar{\sigma})$ where $\bar{\sigma}(y):=\sigma_0$ for all $y\in Y$. $\bar{\sigma}(y)$ is constant, hence commuting. The statement then follows from Proposition 4.3.2. ☐

We prove an analogue of Proposition 4.2.10 for the quantum extensions:

**Proposition 4.3.4.** *Suppose that $\tilde{f}$ is an element of the spectrum and let $f_{\alpha,x,\gamma}$ be its restriction to the classical semiring. Suppose that $(\rho,\sigma)$ is normalized, i.e., $\operatorname{Tr}\rho(x')=1$ and $\operatorname{Tr}\sigma(y')$ for all $x'\in X$ and $y'\in Y$. Then $\tilde{f}(\rho,\sigma)\geq 1$ with equality if and only if $\rho(x)=\sigma(y')$ for all $y'\in\operatorname{supp}\gamma$. Similarly, if $\tilde{f}$ is tropical and its restriction is $f_{x,\gamma}$, then $\tilde{f}(\rho,\sigma)\geq 1$ with equality if and only if $\rho(x)=\sigma(y')$ for all $y'\in\operatorname{supp}\gamma$.*

*Proof.* As in Proposition 4.2.10, the inequality follows by applying monotonicity under the trace map. The "if" part is a consequence of Corollary 4.3.3. For the "only if" direction, suppose that $\rho(x)\neq\sigma(y)$ for some $y\in\operatorname{supp}\gamma$. Let $F$ be a measurement channel such that $F(\rho(x))\neq F(\sigma(y))$. Then $(\rho,\sigma)\not\succcurlyeq(F(\rho),F(\sigma))$, so by monotonicity and Proposition 4.2.10 we have the strict inequality. ☐

The expression $\exp\int_Y\log\sigma\,\mathrm{d}\gamma$ in (4.8) and (4.9) can be viewed as a continuous analogue of a weighted geometric mean of positive numbers. The form of the spectrum elements in the case of commuting $\sigma$ suggests looking for fully quantum generalizations of the form $f(\rho,\sigma)=Q^*_\alpha(\rho(x)\|M(\sigma))$, where $\alpha\geq 1$, $x\in X$ and $M$ is some noncommutative version of the weighted geometric mean. We note that multivariable Rényi divergences constructed using Kubo–Ando means have been studied in [FLO23, MBV22].

In the following definition we make the requirements more precise and also more flexible by allowing the result to be also a continuous family of positive operators. The advantages of this formulation are that a simple composition property conveniently allows for the construction of many examples, and that these objects also give rise to homomorphisms between different semirings $S_{X,Y} \to S_{X',Y'}$. We equip $C(Y, \mathcal{B}(\mathcal{H})_{>0})$ with the pointwise semidefinite partial order.

**Definition 4.3.5.** Let $Y$ and $Y'$ be nonempty compact Hausdorff topological spaces. A family of continuous geometric means indexed by $Y'$ is a collection of maps $M : \bigtimes_{\mathcal{H}} C(Y, \mathcal{B}(\mathcal{H})_{>0}) \to \bigtimes_{\mathcal{H}} C(Y', \mathcal{B}(\mathcal{H})_{>0})$ satisfying the following properties for all $\sigma \in C(Y, \mathcal{B}(\mathcal{H})_{>0})$, $\sigma' \in C(Y, \mathcal{B}(\mathcal{H}')_{>0})$ and $\lambda \in \mathbb{R}_{>0}$:

(i) $M(U\sigma U^*) = UM(\sigma)U^*$, for all unitary $U$

(ii) $M(\sigma \oplus \sigma') = M(\sigma) \oplus M(\sigma')$,

(iii) $M(\sigma \otimes \sigma') = M(\sigma) \otimes M(\sigma')$,

(iv) $M(\lambda\sigma) = \lambda M(\sigma)$,

(v) if $\sigma \leq \sigma'$, then $M(\sigma) \leq M(\sigma')$,

(vi) $M$ is concave.

The set of families of geometric means is denoted by $\mathcal{G}(Y, Y')$. When $Y'$ is a one-point space, we identify $C(Y', \mathcal{B}(\mathcal{H})_{>0})$ with $\mathcal{B}(\mathcal{H})_{>0}$ and write $\mathcal{G}(Y)$ instead of $\mathcal{G}(Y, Y')$. We will see in Proposition 4.3.14, that in the commutative case, elements of $\mathcal{G}(Y)$ reduce to $\exp \int_Y \log \sigma \, \mathrm{d}\gamma$ as in (4.8) and (4.9), and thus we call elements of $\mathcal{G}(Y, Y')$ geometric means.

**Remark 4.3.6.** Because of unitary equivariance, it is sufficient to specify a family of means for families of operators on $\mathbb{C}^d$ for all $d$. Note also that the properties of families of geometric means that we consider imply that they can be extended to positive semidefinite operators by $\lim_{\epsilon \to 0} M(\sigma + \epsilon (I_{\mathcal{H}})_Y)$. This extension will automatically satisfy (ii)-(vi) in Definition 4.3.5. If $\sigma \in C(Y, \mathcal{B}(\mathcal{H})_{\geq 0})$ is such that the operators $\sigma(y)$ share the same support for all $y \in Y$, then this extension can be given by mapping the positive definite part of $\sigma$ by $M$ in the restrictive sense of Definition 4.3.5 and then adding the zero operator on the missing subspace $\ker \sigma$ by the direct sum. In particular, in this case, $M(\sigma(y'))$ will share the same support with $\sigma(y)$ for all $y \in Y$ and $y' \in Y'$. Thus it is clear that if $\sigma \in C(Y, \mathcal{B}(\mathcal{H})_{\geq 0})$ is such that the operators $\sigma(y)$ share the same support for all $y \in Y$, then this extension is invariant under isometries. Such an extension, however, may map a $\sigma \in C(Y, \mathcal{B}(\mathcal{H})_{\geq 0})$ into a discontinuous function (see e.g., the Kubo-Ando means below).

**Example 4.3.7.** An example of an element of $\mathcal{G}(\{1, 2\})$ is given by

$$\sigma(1) \# \sigma(2) = \sigma(1)^{1/2} (\sigma(1)^{-1/2} \sigma(2) \sigma(1)^{-1/2})^{1/2} \sigma(1)^{1/2},$$

the (unweighted) geometric mean of a pair of matrices, introduced in [PW75] and put in a general context by Kubo and Ando [KA80]. Extensions to several variables have been constructed by building on the bivariable geometric mean or generalizing characterizations thereof (see e.g. [Moa05, PT05, BH06,LL11,LP12]), and also studied from an axiomatic point of view [ALM04]. Examples of elements of $\mathcal{G}\{1,2,\ldots,n\}$ are the Karcher means [LL11, LP12].

**Remark 4.3.8.** Since the axioms in Definition 4.3.5 are directly motivated by the use of means in constructing homomorphisms, the axioms given here differ from the ones considered in the literature on geometric means of matrices, in particular in emphasis on relating the means of matrices of different sizes (by the tensor product or the direct sum). In addition, we need to consider every possible weighting of the arguments, therefore symmetry is not a relevant property in our problem.

**Example 4.3.9.** Further examples of elements of $\mathcal{G}(\{1,2\})$ are given by the weighted geometric mean of pairs of matrices (see for example [LL11]):

$$\sigma(1)\#_t\sigma(2) = \sigma(1)^{1/2}(\sigma(1)^{-1/2}\sigma(2)\sigma(1)^{-1/2})^t\sigma(1)^{1/2}.$$

**Proposition 4.3.10.** *If the family of positive operators $C_Y$ is a constant function, then for any $M \in \mathcal{G}(Y,Y')$ one has $M(C_Y) = C_{Y'}$.*

*Proof.* First consider the constant function $(I_{\mathbb{C}^1})_Y$ mapping each point of $Y$ to $I_{\mathbb{C}^1}$. For this function

$$M((I_{\mathbb{C}^1})_Y) = M((I_{\mathbb{C}^1})_Y \otimes (I_{\mathbb{C}^1})_Y) = M((I_{\mathbb{C}^1})_Y) \otimes M((I_{\mathbb{C}^1})_Y).$$

The only positive definite operator satisfying this is $I_{\mathbb{C}^1}$ itself yielding $M((I_{\mathbb{C}^1})_Y) = (I_{\mathbb{C}^1})_{Y'}$. Then from the eigendecomposition of $C = U\bigoplus_{i=1}^n \lambda_i I_{\mathbb{C}^1} U^*$,

$$M(C_Y) = M\left(U\bigoplus_{i=1}^n \lambda_i (I_{\mathbb{C}^1})_Y U^*\right) = U\bigoplus_{i=1}^n \lambda_i M((I_{\mathbb{C}^1})_Y)U^* = U\bigoplus_{i=1}^n \lambda_i (I_{\mathbb{C}^1})_{Y'} U^* = C_{Y'}.$$

□

**Remark 4.3.11.** I will use an abuse of notation and write simply $M(C) = C$ omitting the topological spaces from the indices, the underlying topological space will be clear from context.

The following proposition is a specialization of the main idea behind Lemma 2.2.34 and Lemma 4.2.4 to geometric means.

**Proposition 4.3.12.** *Families of geometric means are increasing under completely positive trace-preserving maps in the sense that if $M \in \mathcal{G}(Y,Y')$, $\sigma \in C(Y,\mathcal{B}(\mathcal{H})_{>0})$ and $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ is a completely positive trace-preserving map, then $M(T(\sigma)) \geq T(M(\sigma))$.*

*Proof.* Consider the Stinespring dilation of $T(.) = \operatorname{Tr}_E V(.)V^*$, with an isometry $V : \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}_E$, and writing the partial trace multiplied by the maximally mixed state as a convex combination of unitary conjugations (e.g., by the discrete Weyl unitaries):

$$\frac{1}{d_E} I_E \otimes \operatorname{Tr}_E V(.)V^* = \sum_i \frac{1}{n} U_i V(.)V^* U_i^*,$$

□

then

$$
\begin{aligned}
&\frac{1}{d_E} I_E \otimes M(T(.)) = M\left(\frac{1}{d_E} I_E \otimes \operatorname{Tr}_E V(.)V^*\right) \\
=&M\left(\sum_i \frac{1}{n} U_i V(.)V^* U_i^*\right) \leq \sum_i \frac{1}{n} M(U_i V(.)V^* U_i^*) \\
=&\sum_i \frac{1}{n} U_i V M(.) V^* U_i^* = \frac{1}{d_E} I_E \otimes \operatorname{Tr}_E V(M(.))V^* = \frac{1}{d_E} I_E \otimes T(M(.)),
\end{aligned}
$$

where we used (iii) and (iv) from Definition 4.3.5, the extended version of $M$ in the sense of Remark 4.3.6 in all the formulas, the equivariance of the extended geometric means under isometries, (vi) from Definition 4.3.5 i.e., the concavity of the extended geometric means to obtain the inequality and Proposition 4.3.10 on the constant function $I_E$. The partial trace is a positive map and so tracing the environment out finishes the proof.

**Lemma 4.3.13.** *For any $a \in \mathbb{R}_{>0}$, any $M \in \mathcal{G}(Y)$ and any $\sigma \in C(Y, \mathcal{B}(\mathbb{C}^1)_{>0})$ one has $M(\sigma^a) = M(\sigma)^a$.*

*Proof.* Let $b$ and $c$ be positive integer numbers, then by the multiplicative property (iii) in Definition 4.3.5 we have

$$
M\left(\sigma^{\frac{b}{c}}\right) = \left(M\left(\left(\sigma^{\frac{b}{c}}\right)^c\right)\right)^{\frac{1}{c}} = \left(M(\sigma^b)\right)^{\frac{1}{c}} = (M(\sigma))^{\frac{b}{c}}.
$$

Now let $\{b_n, c_n, b'_n, c'_n\}_{n=1}^{\infty}$ be sequences of positive integers such that $\frac{b_n}{c_n} \nearrow a \swarrow \frac{b'_n}{c'_n}$, then from the monotonicity of $c^{(\cdot)}$ on real numbers and the monotonicity property (v) in Definition 4.3.5 we have

$$
M(\sigma)^{\frac{b_n}{c_n}} = M\left(\sigma^{\frac{b_n}{c_n}}\right) \leq M(\sigma^a) \leq M\left(\sigma^{\frac{b'_n}{c'_n}}\right) = M(\sigma)^{\frac{b'_n}{c'_n}}.
$$

However, $\lim_{n\to\infty} \left\| M(\sigma)^a - M(\sigma)^{\frac{b_n}{c_n}} \right\|_\infty = \lim_{n\to\infty} \left\| M(\sigma)^a - M(\sigma)^{\frac{b'_n}{c'_n}} \right\|_\infty = 0$, since $M(\sigma)$ is a bounded function. □

The following Proposition underlines the validity of calling elements of $\mathcal{G}(Y, Y')$ geometric means.

**Proposition 4.3.14.** *If $\sigma \in C(Y, \mathcal{B}(\mathbb{C}^1)_{>0})$, then for all $M \in \mathcal{G}(Y)$ $M(\sigma) = \exp \int_Y \log \sigma \,\mathrm{d}\gamma$ for some probability measure $\gamma$.*

*Proof.* $\log M\left(e^{(\cdot)}\right)$ is a positive linear functional on $C(Y, \mathcal{B}(\mathbb{C}^1)_{>0})$ by Lemma 4.3.13 and the multiplicative property (iii) in Definition 4.3.5. Thus by Theorem 2.1.1 there exists some unique $\gamma$ Radon measure such that

$$
\log M\left(e^{(\xi)}\right) = \int_Y \xi(y) \,\mathrm{d}\gamma(y),
$$

then

$$
M(\sigma) = \exp \int_Y \log \sigma \,\mathrm{d}\gamma(y).
$$

From Proposition 4.3.10 we have

$$M(e^1) = \exp \int_Y \mathrm{d}\gamma(y) = e^1$$

yielding $\int_Y \mathrm{d}\gamma(y) = 1$. $\square$

**Remark 4.3.15.** From Proposition 4.3.14 and Theorem 4.2.9 it follows that for all $M \in \mathcal{G}(Y)$

$$f\left(\bigoplus_{i=1}^{d} p_i, \bigoplus_{i=1}^{d} q_i\right) := \sum_{i=1}^{d} p_i(x)^{\alpha} M(q_i)^{1-\alpha}$$

is an element of the classical spectrum underlying the connection between the elements of the spectrum and the geometric means.

The following proposition lists basic constructions that allow one to exhibit many elements of $\mathcal{G}(Y)$. Geometric means that can be obtained in this way include the Ando–Li–Mathias mean [ALM04] and the Bini–Meini–Polini means [BMP10].

**Proposition 4.3.16.** *Let $Y, Y', Y''$ be nonempty compact spaces.*

*(i) If $M \in \mathcal{G}(Y, Y')$ and $N \in \mathcal{G}(Y', Y'')$, then $N \circ M \in \mathcal{G}(Y, Y'')$.*

*(ii) If $f : Y' \to Y$ is a continuous map, then $M(\sigma) = \sigma \circ f$ defines an element of $\mathcal{G}(Y, Y')$.*

*(iii) $M(\sigma) = \sigma'$ with*

$$\sigma'(y_1, y_2, \gamma) := \sigma(y_1)\#_{\gamma}\sigma(y_2) = \sigma(y_1)^{1/2}\left(\sigma(y_1)^{-1/2}\sigma(y_2)\sigma(y_1)^{-1/2}\right)^{\gamma}\sigma(y_1)^{1/2}$$

*defines an element of $\mathcal{G}(Y, Y \times Y \times [0,1])$.*

*(iv) $\mathcal{G}(Y)$ is compact with respect to the pointwise convergence (i.e., convergence of $i \mapsto M_i(\sigma)$ for all $\sigma$).*

*Proof.* (i): The composition is clearly additive, multiplicative, homogeneous and monotone. For concavity, we apply $N$ to the inequality $M(\lambda\sigma + (1-\lambda)\sigma') \geq \lambda M(\sigma) + (1-\lambda)M(\sigma')$, which expresses the concavity of $M$, using that first $N$ is monotone and then that it is concave as well:

$$\begin{aligned}(N \circ M)(\lambda\sigma + (1-\lambda)\sigma') &= N(M(\lambda\sigma + (1-\lambda)\sigma')) \\ &\geq N(\lambda M(\sigma) + (1-\lambda)M(\sigma')) \\ &\geq \lambda N(M(\sigma)) + (1-\lambda)N(M(\sigma')).\end{aligned}$$

(ii): $M$ is clearly additive, multiplicative, homogeneous, monotone and affine (hence concave).

(iii): $\sigma'$ is clearly continuous for every continuous $\sigma$. The geometric mean is clearly additive, multiplicative and homogeneous. For concavity and monotonicity see [KA80] and [Sim19, Theorem 37.1].

(iv): If $\sigma \in C(Y, \mathcal{B}(\mathcal{H})_{>0})$, then there exist constants $c_1, c_2 > 0$ such that $c_1 I \leq \sigma \leq c_2 I$. By the monotonicity property (v) in Definition 4.3.5, it follows that $c_1 I \leq M(\sigma) \leq c_2 I$ for every $M \in$

$\mathcal{G}(Y)$. The interval $[c_1 I, c_2 I] = \left\{A \in \mathcal{B}(\mathbb{C}^d)_{>0} : c_1 I \le A \le c_2 I\right\}$ is compact for every $d$, therefore the evaluations embed $\mathcal{G}(Y)$ into the compact space $\prod_{d\in\mathbb{N}} \prod_{\sigma \in C(Y,\mathcal{B}(\mathbb{C}^d)_{>0})} [c_1(\sigma)I, c_2(\sigma)I]$. The conditions defining $\mathcal{G}$ are closed (equalities and non-strict inequalities with respect to the semidefinite partial order), therefore the image under the embedding is closed. $\square$

**Remark 4.3.17.** We state without proof that $\mathcal{G}\{1,2\}$ consists only of the Kubo-Ando means given in Example 4.3.9.

So far we studied the properties of the geometric means we introduced. We argued that that geometric means in the above sense could be a noncommutative generalization of $\sigma \mapsto \exp \int_Y \log \sigma \, \mathrm{d}\gamma$. We also showed in Proposition 4.3.14 that, indeed, geometric means on real-valued functions (and thus on commuting operators) will take this form. Our goal is to find more general means using the above and plug them in the elements of the spectrum of $S^{\mathrm{sc}}_{X,Y}$ for some $X, Y$ to arrive to more general elements of the spectrum, than the ones already characterized in Proposition 4.3.2 and Corollary 4.3.3. To this end we will use the following proposition.

**Proposition 4.3.18.** *Let $X, Y, X', Y'$ be nonempty compact spaces, $M \in \mathcal{G}(Y, Y')$ and $f : X' \to X$ continuous. Then the map $S_{X,Y} \to S_{X',Y'} : (\rho,\sigma) \mapsto (\rho \circ f, M(\sigma))$ is a semiring homomorphism.*

*Proof.* This map is by definition additive and multiplicative. We have yet to show monotonicity. Suppose that the completely positive trace-nonincreasing map $T$ realizes $(\rho,\sigma) \succcurlyeq (\rho',\sigma')$. Then $T(\rho) \ge \rho'$ and $T(\sigma) \le \sigma'$. From monotonicity of $M$ in its variables under completely positive trace-nonincreasing maps:

$$T(M(\sigma)) \le M(T(\sigma)) \le M(\sigma').$$

This yields $(\rho, M(\sigma)) \succcurlyeq (\rho', M(\sigma'))$ by the same map $T$. $\square$

**Theorem 4.3.19.** *Let $X, Y$ be nonempty compact spaces. For all $\alpha \ge 1$, $x \in X$ and $M \in \mathcal{G}(Y)$ the functional*

$$f(\rho,\sigma) = Q^*_\alpha(\rho(x) \| M(\sigma)) \tag{4.10}$$

*is an element of the real part of the spectrum, and*

$$f(\rho,\sigma) = \left\| M(\sigma)^{-1/2} \rho(x) M(\sigma)^{-1/2} \right\|_\infty \tag{4.11}$$

*is an element of the tropical part.*

*Proof.* By Proposition 4.3.18, the map $(\rho,\sigma) \mapsto (\rho(x), M(\sigma))$ determines a semiring homomorphism from $S_{X,Y}$ to $S_{1,1}$, where 1 is a one-point space. On $S_{1,1}$ the functionals $f_\alpha(\rho,\sigma) = Q^*_\alpha(\rho\|\sigma)$ are in the real spectrum and $(\rho,\sigma) \mapsto \left\|\sigma^{-1/2}\rho\sigma^{-1/2}\right\|_\infty$ is in the tropical spectrum, as follows from Proposition 4.3.2 (see also [PVW22, BV21]). Therefore (4.10) and (4.11) are compositions of semiring homomorphisms, which implies that they are points in the real (respectively tropical) part of the spectrum. $\square$

## 4.4 Applications

The largest subsemiring of $S_{X,Y}$ for which we classified the whole test-spectrum is the semi-classical subsemiring $S^{\text{sc}}_{X,Y}$. Theorem 2.3.26, Remark 2.3.27, Theorem 4.2.9 and Proposition 4.3.2 yield the following:

**Theorem 4.4.1.** *Let $(\rho,\sigma),(\rho',\sigma')\in S^{\text{sc}}_{X,Y}\setminus\{0\}$. Then we have (i) $\Rightarrow$ all of the conditions in (ii), any one of the conditions in (ii) $\Rightarrow$ (iii) and (iii) $\Leftrightarrow$ (iv) in the following:*

*(i) for all $x\in X$, $\alpha>1$ and $\gamma$ probability distribution on $Y$,*

$$Q^*_\alpha\left(\rho(x)\middle\|\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)>Q^*_\alpha\left(\rho'(x)\middle\|\exp\int_Y\log\sigma'\,\mathrm{d}\gamma\right);$$

*(ii) (a) $(\rho,\sigma)^n\succcurlyeq(\rho',\sigma')^n$ for every sufficiently large $n$;*

*(b) $(\rho,\sigma)\succcurlyeq_{\mathrm{c}}(\rho',\sigma')$,*

*moreover, there is a nonnegative integer $k$ such that the catalyst can be chosen as $\bigoplus_{j=0}^{n}\left(2^k\rho^{\otimes j}\otimes\rho'^{\otimes n-j},\sigma^{\otimes j}\otimes\sigma'^{\otimes n-j}\right)$ for any sufficiently large $n$;*

*(iii) For all $x\in X$, $\alpha>1$ and $\gamma$ probability distribution on $Y$,*

$$Q^*_\alpha\left(\rho(x)\middle\|\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)\geq Q^*_\alpha\left(\rho'(x)\middle\|\exp\int_Y\log\sigma'\,\mathrm{d}\gamma\right);$$

*(iv) $(\rho,\sigma)\gtrsim(\rho',\sigma')$.*

*Proof.* Theorem 2.3.26 modified by Remark 2.3.27 needs to be applied to the elements of the spectrum derived in Theorem 4.2.9 and Proposition 4.3.2. The quantities $Q^*_\alpha$ on positive definite operators are continuous in $\alpha$. Thus it follows that from the ordering of the quantities of $Q^*_\alpha$ for $\alpha>1$ directly follows the ordering of

- $\lim_{\alpha\to1}Q^*_\alpha=Q^*_1$ and
- $\lim_{\alpha\to\infty}\frac{1}{\alpha-1}\log Q^*_\alpha=D^*_\infty$, where for states $\omega,\tau$ of full support $D^*_\infty(\omega\|\tau)=\log\left\|\omega^{\frac12}\tau^{-1}\omega^{\frac12}\right\|_\infty$ is the max divergence of $\omega$ and $\tau$ (see (2.16)).

This shows that the orderings

$$Q^*_1\left(\rho(x)\middle\|\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)\quad\text{vs.}\quad Q^*_1\left(\rho'(x)\middle\|\exp\int_Y\log\sigma'\,\mathrm{d}\gamma\right)$$

and

$$\left\|\rho^{\frac12}(x)\left(\exp\int_Y\log\sigma\,\mathrm{d}\gamma\right)^{-1}\rho^{\frac12}(x)\right\|_\infty\quad\text{vs.}\quad\left\|\rho'^{\frac12}(x)\left(\exp\int_Y\log\sigma'\,\mathrm{d}\gamma\right)^{-1}\rho'^{\frac12}(x)\right\|_\infty$$

can be omitted from the conditions in Theorem 4.4.1 for simplicity. $\square$

In the general case of $S_{X,Y}$, the whole spectrum is not classified, and so generally we can only state necessary conditions. Theorem 2.3.26, Remark 2.3.27 and Theorem 4.3.19 yield the following:

**Corollary 4.4.2.** *Let $(\rho,\sigma),(\rho',\sigma')\in S_{X,Y}\setminus\{0\}$. Then we have (i) $\Rightarrow$ (ii), (ii) $\Rightarrow$ all of the conditions in (iii) in the following:*

*(i) there exists a $x\in X$, $\alpha>1$ and $M\in\mathcal{G}(Y)$ such that*

$$Q^*_\alpha(\rho(x)\|M(\sigma))<Q^*_\alpha(\rho'(x)\|M(\sigma'));$$

*(ii) $(\rho,\sigma)\not\gtrsim(\rho',\sigma')$;*

*(iii) (a) $\nexists n\in\mathbb{N}$ such that $(\rho,\sigma)^n\succcurlyeq(\rho',\sigma')^n$;*

*(b) $(\rho,\sigma)\not\succcurlyeq_{\mathrm{c}}(\rho',\sigma')$.*

*Proof.* Completely analogous to the above proof of Theorem 4.4.1 but now using the characterized part of the general spectrum of $S_{X,Y}$ given in Theorem 4.3.19. □

**Remark 4.4.3.** Neither Theorem 4.4.1 or Corollary 4.4.2 uses the tropical part of the characterized part of the spectrum (fully characterized for $S^{\mathrm{sc}}_{X,Y}$ and partially for $S_{X,Y}$) in the characterization of the catalytic or the asymptotic relative submajorization. However the reason behind it is that the sandwiched Rényi divergence $Q^*_\alpha$ giving the elements of the known part of the spectrum is continuous in $\alpha$ on positive definite operators. We know no general argument showing that the general spectrum of $S_{X,Y}$ could not contain tropical functions that are independent from the real part.

### 4.4.1 Composite hypothesis testing

One interpretation of a (normalized) pair of families is that the states in $\rho(x),x\in X$ form a composite null hypothesis which is to be tested against the composite alternative hypothesis $\sigma(y),y\in Y$. In this hypothesis testing problem one considers a two-outcome POVM $(\Pi,I-\Pi)$, or *test*, and the decision is based on the measurement result, rejecting the null hypothesis if the second outcome is observed. Such a test is uniquely specified by an operator $\Pi$ such that $0\le\Pi\le I$ and every such operator gives rise to a valid POVM.

A type I error occurs when the null hypothesis is falsely rejected. For every member in the family $\rho$ we define a probability of type I error,

$$\alpha_x(\Pi):=\operatorname{Tr}\rho(x)(I-\Pi)=1-\operatorname{Tr}\rho(x)\Pi,$$

and the maximum

$$\alpha(\Pi):=\max_{x\in X}\alpha_x(\Pi)$$

is the *significance level* of the test.

In contrast, a type II error means that the correct state was from $\sigma$ but the null hypothesis does not get rejected. The probability of a type II error is

$$\beta_y(\Pi):=\operatorname{Tr}\sigma(y)\Pi,$$

and their maximum is

$$\beta(\Pi) := \max_{y \in Y} \beta_y(\Pi).$$

In general it is not possible to have arbitrary low probability for both types of errors but there is a trade-off between the two quantities. Following the idea of [Ren16]), the possible values are exactly characterized by the preordered semiring $S_{X,Y}$:

**Proposition 4.4.4.** *Let $(\rho, \sigma) \in S_{X,Y}$ be a pair of families of states of full support and let $\alpha : X \to \mathbb{R}$ and $\beta : Y \to \mathbb{R}$ be continuous functions, such that $\alpha < 1$ and $\beta > 0$. Then $(1 - \alpha, \beta)$ is an element of $S_{X,Y}$ and the following are equivalent:*

*(i) there exists a test $\Pi$ with $\alpha_x(\Pi) \leq \alpha(x)$ and $\beta_y(\Pi) \leq \beta(y)$ for all $x \in X$ and $y \in Y$,*

*(ii) $(\rho, \sigma) \succcurlyeq (1 - \alpha, \beta)$.*

*Proof.* Let $(\rho, \sigma) \in S_{X,Y}$ be a pair of families of states on $\mathcal{H}$ and suppose that a test exists with the properties above. Consider the map $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathbb{C})$ given by $T(X) = \mathrm{Tr}(X\Pi)$. $T$ is completely positive because $\Pi \geq 0$ and trace nonincreasing because $\Pi \leq I$. We apply $T$ to the families:

$$\begin{aligned} T(\rho(x)) &= \mathrm{Tr}(\rho(x)\Pi) = 1 - \alpha_x(\Pi) \geq 1 - \alpha(x) \\ T(\sigma(y)) &= \mathrm{Tr}(\sigma(y)\Pi) = \beta_y(\Pi) \leq \beta(y), \end{aligned}$$

for all $x \in X$ and $y \in Y$, therefore $(\rho, \sigma) \succcurlyeq^T (1 - \alpha, \beta)$.

Conversely, suppose that $(\rho, \sigma) \succcurlyeq (1 - \alpha, \beta)$. This means that there exists a completely positive trace nonincreasing map $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathbb{C})$ such that $T(\rho) \geq (1 - \alpha)$ and $T(\sigma) \leq \beta$. Pick such a map and let $\Pi = T^*(I_{\mathbb{C}^1})$. Then $0 \leq \Pi \leq I$ and

$$\begin{aligned} \alpha_x(\Pi) &= \mathrm{Tr}\, \rho(x)(I - \Pi) = 1 - \mathrm{Tr}\, \rho(x) T^*(I_{\mathbb{C}^1}) = 1 - T(\rho(x)) \leq \alpha(x) \\ \beta_y(\Pi) &= \mathrm{Tr}\, \sigma(y)\Pi = \mathrm{Tr}\, \sigma(y) T^*(I_{\mathbb{C}^1}) = T(\sigma(y)) \leq \beta(y), \end{aligned}$$

for all $x \in X$ and $y \in Y$.

□

**Remark 4.4.5.** Note that Proposition 4.4.4

- is trivially fulfilled if $\exists x \in X$ such that $\alpha(x) < 0$, as neither of the two conditions can be satisfied in this case;
- is trivially fulfilled if $\alpha \geq 0$ and $\beta \geq 1$, since both conditions can be satisfied with the trivial test $\Pi = I$ and the trace-functional $\mathrm{Tr}(\cdot)$;
- is trivially fulfilled if $\exists x \in X$ and $\exists y \in Y$ such that $\alpha(x) = 0$ and $\beta(y) < 1$, since from $\alpha(x) = 0$ we have that the test and the map must be as in the previous point: the trivial test and the trace-functional, and these do not satisfy either condition;
- gives a nontrivial characterization in the case when $0 < \alpha < 1$ and $\exists y \in Y$ such that $\beta(y) \in (0, 1)$. However, we can assume that $0 < \beta \leq 1$ by restricting $Y$ to $\beta^{-1}([0, 1])$, as neither $(i)$ or $(ii)$ in Proposition 4.4.4 results in any constraints for $y \in \beta^{-1}((1, +\infty))$.

Suppose that we have access to $n$ copies of such identically prepared pairs of families. The element of the semiring describing this situation is the power $(\rho, \sigma)^n = (\rho^{\otimes n}, \sigma^{\otimes n})$. If we are allowed to perform a joint measurement, i.e., $\Pi_n \in \mathcal{B}(\mathcal{H}^{\otimes n})_{\geq 0}$ with $\Pi_n \leq I_{\mathcal{H}^{\otimes n}}$, then we expect to be able to achieve lower probabilities of both types of errors than with a single copy. In particular, an extension of the quantum Stein's lemma says that when the alternative hypothesis is simple, as the number of copies $n \to \infty$ and the maximum of the probability of the type I error is required to go to 0, it is possible to achieve an exponential decay of the type II error, where the exponent is given by the minimum of the relative entropies $D^{\mathrm{Um}}(\rho(x)\|\sigma)$ [BDK+05, BBH21, Mos15].

**Remark 4.4.6.** Although in most asymptotic settings, the errors given in terms of the number of copies we have access to are exponential, there are notable exceptions to this rule of thumb. See for example our work in [BMMZ22]. In our case here however, the asymptotic errors are exponential in the number of copies.

The asymptotic preorder $\gtrsim$ is able to capture the exponential decay of the type II error and the exponential convergence of the type I error to one, called the *strong converse regime*. We have the following characterization.

**Proposition 4.4.7.** *For compact Hausdorff topological spaces $X, Y$ and composite null and alternative hypotheses in the form of continuous functions $\rho : X \to S(\mathcal{H})$, $\sigma : Y \to S(\mathcal{H})$ consisting of states of full support and for all $R : X \to \mathbb{R}$ and $r : Y \to \mathbb{R}$ continuous functions (rate functions), the following are equivalent:*

*(i) there is a sequence of tests $\{\Pi_n\}_{n=1}^{\infty}$ and a sublinear sequence of nonnegative integers $\{k_n\}_{n=1}^{\infty}$ such that $\alpha_x(\Pi_n) \leq 1 - e^{-R(x)n-k_n}$ and $\beta_y(\Pi_n) \leq e^{-r(y)n}$ for all $x \in X$, $y \in Y$ and $n \in \mathbb{N}$,*

*(ii) $(\rho, \sigma) \gtrsim (e^{-R}, e^{-r})$.*

*Proof.* We have that $(e^{-R}, e^{-r})$ is an element of $S_{X,Y}$ for all $R : X \to \mathbb{R}$ and $r : Y \to \mathbb{R}$ continuous functions. Recall from Remark 4.1.7 that in the definition of the asymptotic preorder, we can choose $u = (e_X, 1_Y)$, and since $u = (e_X, 1_Y)$ is invertible, the condition appearing in the definition of the asymptotic preorder may be written as

$$
\begin{aligned}
(\rho^{\otimes n}, \sigma^{\otimes n}) &\succcurlyeq u^{-k_n}(e^{-R}, e^{-r})^n \\
&= (e^{-Rn-k_n}, e^{-rn}),
\end{aligned}
$$

for all $n \in \mathbb{N}$ and some $\{k_n\}_{n=1}^{\infty}$ sublinear sequence of nonnegative integers. According to Proposition 4.4.4, the above is equivalent to the existence of a sequence of tests $\{\Pi_n\}_{n=1}^{\infty}$ such that $\alpha_x(\Pi_n) \leq 1 - e^{-R(x)n-k_n}$ and $\beta_y(\Pi_n) \leq e^{-r(y)n}$ for all $x \in X$, $y \in Y$ and $n \in \mathbb{N}$. $\square$

**Remark 4.4.8.** Note that, following directly from Remark 4.4.5, Proposition 4.4.4

- is trivially fulfilled if $\exists x \in X$ such that $R(x) < 0$, as neither of the two conditions can be satisfied in this case;
- is trivially fulfilled if $R \geq 0$ and $r \leq 0$, since both conditions can be satisfied with the sequence of tests $\Pi_n = I_{\mathcal{H}^{\otimes n}}$ and the $\mathrm{Tr}(\cdot)$ functional;

- gives a nontrivial characterization, when $R \geq 0$ and $\exists y \in Y$ such that $r(y) > 0$. However, similar to the case in Remark 4.4.5, we can assume that $r \geq 0$ without losing generality.

**Proposition 4.4.9.** *For the one-point space $X = \{x\}$ and any compact Hausdorff topological space $Y$ and simple null hypothesis $\rho > 0$ and composite alternative hypothesis in the form of continuous function $\sigma : Y \to S(\mathcal{H})$ consisting of states of full support and for all continuous rate function $r : Y \to \mathbb{R}$ and error rate $R \in \mathbb{R}$, the following are equivalent:*

*(i) for all sequences of tests $\{\Pi_n\}_{n=1}^{\infty}$ such that $\beta_y(\Pi_n) \leq e^{-r(y)n}$ for all $y \in Y$ and $n \in \mathbb{N}$, and for all sublinear sequences of nonnegative integers $\{k_n\}_{n=1}^{\infty}$, we have $\alpha_x(\Pi_n) > 1 - e^{-Rn-k_n}$ for all but finitely many exceptions in $n \in \mathbb{N}$,*

*(ii) $(\rho, \sigma) \not\gtrsim (e^{-R}, e^{-r})$.*

*Proof.* $(ii) \Rightarrow (i)$ : By Corollary 2.3.12 $(\rho, \sigma) \not\gtrsim (e^{-R}, e^{-r}) \iff$ for all sublinear sequences of positive integers $\{k_n\}_{n=1}^{\infty}$ we have that $(\rho^{\otimes n}, \sigma^{\otimes n}) \not\succcurlyeq (e^{-Rn-k_n}, e^{-rn})$ for finitely many $n \in \mathbb{N}$. Then by Proposition 4.4.4 it follows that for all sequences of tests $\{\Pi_n\}_{n=1}^{\infty}$ and all sublinear sequences of positive integers $\{k_n\}_{n=1}^{\infty}$ we have $\big(\alpha_x(\Pi_n) > 1 - e^{-Rn-k_n}$ or $\beta_y(\Pi_n) > e^{-r(y)n}$ for some $y \in Y\big)$ for all but finitely many $n \in \mathbb{N}$. $(i) \Rightarrow (ii)$ is immediate from Proposition 4.4.7. □

**Remark 4.4.10.** Similarly as in Remark 4.4.8 we can assume in Proposition 4.4.9 that $R, r \geq 0$, without losing generality.

**Remark 4.4.11.** Note that, although, Theorem 4.4.1 only allows us characterization of the asymptotic preorder in the case of $(\rho, \sigma) \in S^{\mathrm{sc}}_{X,Y}$, i.e., both $\rho$ and $\sigma$ are positive definite operator valued continuous functions on compact topological spaces and the operators in $\sigma$ are required to commute, we have not yet used these assumptions in the above propositions, in Proposition 4.4.4, Proposition 4.4.7 and Proposition 4.4.9. The reason to still state these propositions in the above forms is that we could not apply Theorem 4.4.1 to a more general setting. Precise construction of the more general semiring $S^{\geq 0}_{X,Y}$ consisting of families of positive semi-definite operators as continous functions over topological spaces is straightforward. Characterization of the elements of the spectrum, however, is a much harder task, if positive semi-definite operators are allowed in the families of the semiring. The requirement of compactness however is unavoidable for many arguments in the previous sections, even though one could define type I and II errors with suprema instead of maxima.

Thus, to achieve asymptotically the smallest type II error probabilities for a family of given exponents $r$, we need to find the smallest $R(x)$ values satisfying the equivalent conditions in terms of the spectral points for the asymptotic preorder. Proposition 4.4.7 and Theorem 4.4.1 imply the following theorem:

**Theorem 4.4.12.** *For compact Hausdorff topological spaces $X, Y$ and composite null and alternative hypotheses in the form of continuous functions $\rho : X \to S(\mathcal{H})$, $\sigma : Y \to S(\mathcal{H})$ consisting of states of full support, where in addition $[\sigma(y), \sigma(y')] = 0$ for all $y, y' \in Y$, and for all $R : X \to \mathbb{R}$ and $r : Y \to \mathbb{R}$ continuous rate functions, the following are equivalent:*

*(i) there is a sequence of tests $\{\Pi_n\}_{n=1}^{\infty}$ and a sublinear sequence of positive integers $\{k_n\}_{n=1}^{\infty}$ for which $\alpha_x(\Pi_n) \leq 1 - e^{-R(x)n-k_n}$ and $\beta_y(\Pi_n) \leq e^{-r(y)n}$ for all $x \in X$, $y \in Y$ and $n \in \mathbb{N}$;*

*(ii)* $(\rho, \sigma) \succsim (e^{-R}, e^{-r})$;

*(iii)*

$$\begin{aligned} R(x) &\geq \sup_{\alpha>1} \sup_{\gamma} \frac{\alpha-1}{\alpha} \left[ \int_Y r \,\mathrm{d}\gamma - D^*_\alpha\left( \rho(x) \middle\| \exp \int_Y \log \sigma \,\mathrm{d}\gamma \right) \right] \\ &= \sup_{\alpha>1} \frac{\alpha-1}{\alpha} \max_{\gamma} \left[ \int_Y r \,\mathrm{d}\gamma - D^*_\alpha\left( \rho(x) \middle\| \exp \int_Y \log \sigma \,\mathrm{d}\gamma \right) \right], \end{aligned} \tag{4.12}$$

*for all* $x \in X$.

*In addition, for the bounds we have*

$$\sup_{\alpha>1} \frac{\alpha-1}{\alpha} \max_{\gamma} \left[ \int_Y r \,\mathrm{d}\gamma - D^*_\alpha\left( \rho(x) \middle\| \exp \int_Y \log \sigma \,\mathrm{d}\gamma \right) \right] \geq 0$$

*and*

$$\sup_{\alpha>1} \frac{\alpha-1}{\alpha} \max_{\gamma} \left[ \int_Y r \,\mathrm{d}\gamma - D^*_\alpha\left( \rho(x) \middle\| \exp \int_Y \log \sigma \,\mathrm{d}\gamma \right) \right] > 0 \iff \max_y \left[ r(y) - D^{\mathrm{Um}}(\rho(x) \| \sigma(y)) \right] > 0.$$

*Proof.* According to Proposition 4.4.7, *(i)* $\iff$ *(ii)*. According to Theorem 4.4.1, $(\rho, \sigma) \succsim (e^{-R}, e^{-r})$ if and only if for all $x \in X$, $\alpha > 1$ and $\gamma$ probability distribution on $Y$,

$$\begin{aligned} Q^*_\alpha\left( \rho(x) \middle\| \exp \int_Y \log \sigma \,\mathrm{d}\gamma \right) &\geq Q^*_\alpha\left( e^{-R(x)} \middle\| \exp \int_Y \log e^{-r} \,\mathrm{d}\gamma \right) \\ &= Q^*_\alpha\left( e^{-R(x)} \middle\| e^{-\int_Y r \,\mathrm{d}\gamma} \right) \\ &= e^{-R(x)\alpha - \int_Y r \,\mathrm{d}\gamma (1-\alpha)}. \end{aligned}$$

Applying $\frac{1}{\alpha-1} \log(\cdot)$ to both sides yield

$$D^*_\alpha\left( \rho(x) \middle\| \exp \int_Y \log \sigma \,\mathrm{d}\gamma \right) \geq -R(x) \frac{\alpha}{\alpha-1} + \int_Y r \,\mathrm{d}\gamma,$$

(recall the definition of $D^*_\alpha$ in Example 2.2.10 and note that $\operatorname{Tr} \rho = 1_X$, since we assumed $\rho$ is a family of states). We rearrange for $R(x)$:

$$R(x) \geq \frac{\alpha-1}{\alpha} \left[ \int_Y r \,\mathrm{d}\gamma - D^*_\alpha\left( \rho(x) \middle\| \exp \int_Y \log \sigma \,\mathrm{d}\gamma \right) \right],$$

for all $x \in X$, $\alpha > 1$ and $\gamma$ probability distribution on Y. We also used the lower semicontinuity of $D^*_\alpha$ in its arguments (see [MO17, Corollary 3.27]) and thus changed the supremum in $\gamma$ to maximum.

In addition for the characterizing bounds we have

$$\lim_{\alpha \searrow 1} \frac{\alpha-1}{\alpha} \max_{\gamma} \left[ \int_Y r \,\mathrm{d}\gamma - D^*_\alpha\left( \rho(x) \middle\| \exp \int_Y \log \sigma \,\mathrm{d}\gamma \right) \right] = 0,$$

since the term in the brackets is bounded for a fixed $r$ and

$$\begin{aligned}\sup_{\alpha>1}\frac{\alpha-1}{\alpha}\max_{\gamma}\left[\int_Y r\,\mathrm{d}\gamma - D^*_\alpha\left(\rho(x)\middle\|\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)\right] > 0 \iff \\ \exists\alpha>1 : \max_{\gamma}\left[\int_Y r\,\mathrm{d}\gamma - D^*_\alpha\left(\rho(x)\middle\|\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)\right] > 0 \iff \\ \max_{\gamma}\left[\int_Y r\,\mathrm{d}\gamma - D^{\mathrm{Um}}\left(\rho(x)\middle\|\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)\right] > 0,\end{aligned}$$

where in the first equivalence we dropped the factor $\frac{\alpha-1}{\alpha}$ as it is positive for all $\alpha > 1$, and in the second equivalence we used the monotonicity and continuity of $D^*_\alpha$ in $\alpha$ on positive definite operators (see [MDS+13]) and the fact that on states $\lim_{\alpha\to 1} D^*_\alpha$ evaluates as $D^{\mathrm{Um}}$ (see Example 2.2.10). However for the derived bound we have

$$\begin{aligned}&\max_{\gamma}\left[\int_Y r\,\mathrm{d}\gamma - D^{\mathrm{Um}}\left(\rho(x)\middle\|\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)\right] \\ &= \max_{\gamma}\left[\int_Y r\,\mathrm{d}\gamma - \operatorname{Tr}\rho(x)\left(\log\rho(x) - \log\left(\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)\right)\right] \\ &= \max_{\gamma}\left[\int_Y r\,\mathrm{d}\gamma - \operatorname{Tr}\rho(x)\left(\log\rho(x) - \int_Y \log\sigma\,\mathrm{d}\gamma\right)\right] \\ &= \max_{\gamma}\int_Y [r - \operatorname{Tr}\rho(x)(\log\rho(x) - \log\sigma)]\,\mathrm{d}\gamma \\ &= \max_{y}[r - \operatorname{Tr}\rho(x)(\log\rho(x) - \log\sigma)] \\ &= \max_{y}\big[r - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big].\end{aligned}$$

□

**Remark 4.4.13.** Note that similarly as in the case of Proposition 4.4.7 and Remark 4.4.8, Theorem 4.4.12 gives nontrivial characterization for rate functions $R, r \geq 0$.

**Theorem 4.4.14.** *For compact Hausdorff topological spaces $X, Y$ and composite null and alternative hypotheses in the form of continuous functions $\rho : X \to S(\mathcal{H})$, $\sigma : Y \to S(\mathcal{H})$ consisting of states of full support, where in addition $[\sigma(y), \sigma(y')] = 0$ for all $y, y' \in Y$, and for all $r : Y \to \mathbb{R}$ continuous rate functions and all $x \in X$, the following are equivalent:*

*(i)* *$\exists R > 0$ such that for all sequences of tests $\{\Pi_n\}_{n=1}^\infty$ such that $\beta_y(\Pi_n) \leq e^{-r(y)n}$ for all $y \in Y$ and $n \in \mathbb{N}$, and for all sequences of nonnegative integers $\{k_n\}_{n=1}^\infty$, we have $\alpha_x(\Pi_n) > 1 - e^{-Rn-k_n}$ for all but finitely many $n \in \mathbb{N}$;*

*(ii)* *the interval $\big(0, \max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big]\big)$ is nonempty and for all $R > 0$, we have $R < \max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big]$ if any only if for all sequences of tests $\{\Pi_n\}_{n=1}^\infty$ such that $\beta_y(\Pi_n) \leq e^{-r(y)n}$ for all $y \in Y$ and $n \in \mathbb{N}$, and for all sequences of nonnegative integers $\{k_n\}_{n=1}^\infty$, we have $\alpha_x(\Pi_n) > 1 - e^{-Rn-k_n}$ for all but finitely many $n \in \mathbb{N}$;*

*(iii)* *$\exists R > 0$ such that $(\rho(x), \sigma) \not\gtrsim (e^{-R}, e^{-r})$ in $S^{\mathrm{sc}}_{\{x\},Y}$;*

*(iv) the interval* $\big(0, \max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big]\big)$ *is nonempty and for all* $R > 0$ *we have* $R < \max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big]$ *if and only if* $(\rho(x), \sigma) \not\gtrsim (e^{-R}, e^{-r})$ *in* $S^{\mathrm{sc}}_{\{x\},Y}$*;*

*(v)* $\max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big] > 0$*;*

*Proof.* The equivalence relations $(i) \iff (iii)$ and $(ii) \iff (iv)$ follow directly from Proposition 4.4.9. $(v) \Rightarrow (iv)$ : Let $x \in X$ be such that $\max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big] > 0$. Choose $\{x\}$ and $Y$ as indexing topological spaces in Theorem 4.4.12. It follows that for a $R > 0$ in $S^{\mathrm{sc}}_{\{x\},Y}$, we have $(\rho(x), \sigma) \not\gtrsim (e^{-R}, e^{-r})$ if and only if $R \in \big(0, \max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big]\big)$. $(iv) \Rightarrow (iii)$ is trivial. $(iii) \Rightarrow (v)$ : Follows directly from Theorem 4.4.12 by choosing $\{x\}$ and $Y$ as indexing topological spaces. □

In particular we have the following immediate consequence of Theorem 4.4.12 and Theorem 4.4.14:

**Corollary 4.4.15.** *For compact Hausdorff topological spaces* $X, Y$ *and composite null and alternative hypotheses in the form of continuous functions* $\rho : X \to S(\mathcal{H})$, $\sigma : Y \to S(\mathcal{H})$ *consisting of states of full support, where in addition* $[\sigma(y), \sigma(y')] = 0$ *for all* $y, y' \in Y$*, for all* $r : Y \to \mathbb{R}$ *continuous rate functions, and for all* $x \in X$*, we have the following*

$$\begin{aligned}\inf \liminf_{n\to+\infty}\left\{-\frac{1}{n}\log(1-\alpha_x(\Pi_n))\right\} &= \inf \limsup_{n\to+\infty}\left\{-\frac{1}{n}\log(1-\alpha_x(\Pi_n))\right\}\\ &= \max\left\{0, \max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big]\right\},\end{aligned}$$

*where the infimum is taken over all sequences of tests* $\{\Pi_n\}_{n=1}^{\infty}$*, such that* $\beta_y(\Pi_n) \leq e^{-r(y)n}$ *for all* $y \in Y$ *and* $n \in \mathbb{N}$*.*

*Proof.* By Theorem 4.4.12 for $R = \max\big\{0, \max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big]\big\}$ there is a sequence of tests $\{\Pi_n^{\mathrm{opt}}\}_{n=1}^{\infty}$ and a sublinear sequence of positive integers $\{k_n\}_{n=1}^{\infty}$ for which $\alpha_x(\Pi_n^{\mathrm{opt}}) \leq 1 - e^{-R(x)n-k_n}$ and $\beta_y(\Pi_n^{\mathrm{opt}}) \leq e^{-r(y)n}$ for all $x \in X$, $y \in Y$ and $n \in \mathbb{N}$. Whereas by Theorem 4.4.14, for all $0 < R < \max\big\{0, \max_y\big[r(y) - D^{\mathrm{Um}}(\rho(x)\|\sigma(y))\big]\big\}$ and for all sequences of tests $\{\Pi_n\}_{n=1}^{\infty}$ such that $\beta_y(\Pi_n) \leq e^{-r(y)n}$ for all $y \in Y$ and $n \in \mathbb{N}$, and for all sequences of nonnegative integers $\{k_n\}_{n=1}^{\infty}$, we have $\alpha_x(\Pi_n) > 1 - e^{-Rn-k_n}$ for all but finitely many $n \in \mathbb{N}$. □

In Theorem 4.4.12 we succeeded in controlling the error rate functions individually for all operators in the families $\rho$ and $\sigma$. One can simplify this setting by requiring bounds on the maximuma of the errors, this simplification results in the so-called worst case scenario, where we want to control the quantities $\alpha(\Pi_n) = \max_{x\in X}\alpha_x(\Pi_n)$ and $\beta(\Pi_n) = \max_{y\in Y}\beta_y(\Pi_n)$. We have the following immediate consequence of Theorem 4.4.12.

**Corollary 4.4.16.** *For compact Hausdorff topological spaces* $X, Y$ *and composite null and alternative hypotheses in the form of continuous functions* $\rho : X \to S(\mathcal{H})$, $\sigma : Y \to S(\mathcal{H})$ *consisting of states of full support, where in addition* $[\sigma(y), \sigma(y')] = 0$ *for all* $y, y' \in Y$*, and for all* $R, r \in \mathbb{R}$*, the following are equivalent:*

*(i) there is a sequence of tests* $\Pi_n$ *on* $\mathcal{H}^{\otimes n}$ *and a sublinear sequence of positive integers* $\{k_n\}_{n=1}^{\infty}$ *for which* $\alpha(\Pi_n) \leq 1 - e^{-Rn-k_n}$ *and* $\beta(\Pi_n) \leq e^{-rn}$ *for all* $n \in \mathbb{N}$*,*

*(ii)*

$$R \geq \sup_{\alpha>1} \frac{\alpha-1}{\alpha}\left[r - \min_{\gamma}\min_{x} D^*_\alpha\left(\rho(x)\middle\|\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)\right].$$

*In addition, for the bounds we have*

$$\sup_{\alpha>1} \frac{\alpha-1}{\alpha}\left[r - \min_{\gamma}\min_{x} D^*_\alpha\left(\rho(x)\middle\|\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)\right] \geq 0$$

*and*

$$\sup_{\alpha>1} \frac{\alpha-1}{\alpha}\left[r - \min_{\gamma}\min_{x} D^*_\alpha\left(\rho(x)\middle\|\exp\int_Y \log\sigma\,\mathrm{d}\gamma\right)\right] > 0 \iff \min_x\min_y D^{\mathrm{Um}}(\rho(x)\|\sigma(y)) < r.$$

*Proof.* Choose the constant functions $R_X$ and $r_y$ as error rate functions in Theorem 4.4.12 and use the lower semicontinuity of $D^*_\alpha$ in its first argument [MO17, Corollary 3.27]. □

In particular for a simple alternative hypothesis, the exponent is given by the minimum of the pairwise exponents. This was first derived for simple hypotheses in [MO15]. The above result for simple alternative hypothesis was first derived axiomatically in [BV21]. Corollary 4.4.16 is a generalization of both results. Note that the result for a simple alternative hypothesis, but a finitely composite null hypothesis in [BV21] can also be obtained from the simple null hypothesis case by an averaging argument (see [BV21, Proposition 9]). Theorem 4.4.12 is generalized even further, it allows individual control of all the type I and type II errors.

### 4.4.2 Equivariant relative submajorization

In this section we consider pairs of operators on a representation space of some fixed group, and a variant of relative submajorization that takes into account the group actions. Let $G$ be a topological group. Let $\pi$ and $\pi'$ be finite dimensional unitary representations of $G$ on $\mathcal{H}$ and $\mathcal{H}'$, respectively. Suppose that $(\rho_0,\sigma_0)\in\mathcal{B}(\mathcal{H})^2_{>0}$ and $(\rho'_0,\sigma'_0)\in\mathcal{B}(\mathcal{H}')^2_{>0}$. We say that $(\pi,\rho_0,\sigma_0)$ equivariantly relatively submajorizes $(\pi',\rho'_0,\sigma'_0)$ if there exists a completely positive trace-nonincreasing map $T:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H}')$ such that

$$\begin{aligned} T(\rho_0) &\geq \rho'_0 \\ T(\sigma_0) &\leq \sigma'_0 \\ \forall g\in G,\forall A\in\mathcal{B}(\mathcal{H}) : T(\pi(g)A\pi(g)^*) &= \pi'(g)T(A)\pi'(g)^* \end{aligned}$$

On these triples the direct sum and tensor product (of representations and of operators) give binary operations that are compatible with equivariant relative submajorization.

It will be convenient to restrict to compact groups $G$, and it can be done without loss of generality for the following reason. Consider the closure $K$ of $\{(\pi(g),\pi'(g)) : g\in G\}\subseteq U(\mathcal{H})\times U(\mathcal{H}')$. This is a compact group (in fact, a Lie group), the map $g\mapsto(\pi(g),\pi'(g))$ is a homomorphism and the representations $\pi, \pi'$ of $G$ extend to representations of $K$ (namely the first and second projections provide the required homomorphisms). By continuity, a map $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ is $G$-equivariant if and only if it is $K$-equivariant. Therefore the condition for $(\pi, \rho_0, \sigma_0) \succcurlyeq (\pi', \rho_0', \sigma_0')$ can be formulated in terms of the compact group $K$ instead of $G$. Note that in this case the compact group in general depends on the specific pair of triples to be compared (through the representations), which may not always be desirable. Alternatively, one may construct $K$ in a universal way, by taking the Bohr compactification of $G$. Recall that the Bohr compactification of a topological group $G$ is a compact Hausdorff topological group $\mathrm{b}(G)$ together with a continuous homomorphism $b : G \to \mathrm{b}(G)$ that is universal in the sense that every continuous homomorphism from $G$ into a compact group factors through $b$ in a unique way. Every topological group has an essentially unique Bohr compactification. We can apply the universal property to the homomorphisms $\pi : G \to U(\mathcal{H})$ to get a representation $\mathrm{b}(\pi) : \mathrm{b}(G) \to U(\mathcal{H})$. Thus, instead of each triple $(\pi, \rho_0, \sigma_0)$ we may consider the modified triple $(\mathrm{b}(\pi), \rho_0, \sigma_0)$. For notational simplicity, from now on we will assume that $G$ itself is a compact Hausdorff group.

We now show how to map the triples $(\pi, \rho_0, \sigma_0)$ to pairs of families in such a way that the operations are preserved and equivariant relative submajorization translates to the relative submajorization of the families. In this way a triple $(\pi, \rho_0, \sigma_0)$ gives rise to the following pair of families, parametrized by $G$:

$$\begin{aligned} \rho(g) &= \pi(g)\rho_0\pi(g)^* \\ \sigma(g) &= \pi(g)\sigma_0\pi(g)^*. \end{aligned}$$

$(\rho, \sigma)$ determines an element of $S_{G,G}$, and this element remains the same if we replace the triple $(\pi, \rho_0, \sigma_0)$ by a unitary equivalent one. This map clearly respects the sum and product operations.

**Example 4.4.17.**

(i) Let $G = \{0, 1\}$ with addition modulo 2 as the group operation. A representation $\pi$ of $G$ on $\mathcal{H}$ is determined by the image of 1, which is a unitary $Z$ satisfying $Z^2 = I$. The triple $(\pi, \rho, \sigma)$ then gives rise to the element $((\rho, Z\rho Z), (\sigma, Z\sigma Z))$.

(ii) Let $G = U(1)$. A representation $\pi$ of $G$ is given by the infinitesimal generator $A \in \mathcal{B}(\mathcal{H})_{\mathrm{sa}}$ as $e^{it} \mapsto e^{itA}$, which is well-defined if $A$ has integer spectrum. The element corresponding to $(\pi, \rho, \sigma)$ may be identified with $((e^{itA}\rho e^{-itA})_{t\in[0,2\pi]}, (e^{itA}\sigma e^{-itA})_{t\in[0,2\pi]})$ (where 0 and $2\pi$ both represent the identity element of $U(1)$).

Suppose that $(\pi, \rho_0, \sigma_0)$ equivariantly relatively submajorizes $(\pi', \rho_0', \sigma_0')$, and let $T$ be an equivariant completely positive trace-nonincreasing map satisfying $T(\rho_0) \geq \rho_0'$ and $T(\sigma_0) \leq \sigma_0'$. Consider the corresponding elements $(\sigma, \rho)$ and $(\sigma', \rho')$ of $S_{G,G}$. Then for all $g \in G$ the inequality

$$T(\rho(g)) = T(\pi(g)\rho_0\pi(g)^*) = \pi'(g)T(\rho_0)\pi'(g)^* \geq \pi'(g)\rho_0'\pi'(g)^* = \rho'(g)$$

holds and similarly $T(\sigma(g)) \leq \sigma'(g)$. This means that $(\rho, \sigma) \succcurlyeq (\rho', \sigma')$ holds.

Conversely, suppose that $(\rho, \sigma) \succcurlyeq (\rho', \sigma')$ is true in $S_{G,G}$ for the families defined above. This means that there exists a (not necessarily equivariant) completely positive trace-nonincreasing map $T_0$ such that for all $g \in G$ the inequalities $T_0(\rho(g)) \geq \rho'(g)$ and $T_0(\sigma(g)) \leq \sigma'(g)$ hold. We construct an equivariant map $T$ by averaging:

$$T(X) = \int_G \pi'(g)^* T_0(\pi(g) X \pi(g)^*) \pi'(g) \, \mathrm{d}\mu(g),$$

where $\mu$ is the Haar probability measure on $G$. Then $T$ is $G$-equivariant and in addition

$$\begin{aligned} T(\rho_0) &= \int_G \pi'(g)^* T_0(\pi(g)\rho_0\pi(g)^*)\pi'(g) \, \mathrm{d}\mu(g) \\ &= \int_G \pi'(g)^* T_0(\rho(g))\pi'(g) \, \mathrm{d}\mu(g) \\ &\geq \int_G \pi'(g)^* \rho'(g) \pi'(g) \, \mathrm{d}\mu(g) = \rho_0', \end{aligned}$$

and similarly $T(\sigma_0) \leq \sigma_0'$.

Note that even though the map $(\pi, \rho_0, \sigma_0) \mapsto (\rho, \sigma)$ is order-preserving and order-reflecting, it is in general not injective (on equivalence classes). Now we can apply our result on general pairs of families to the question of asymptotic equivariant relative submajorization.

**Theorem 4.4.18.** *Let $G$ be a topological group and consider the triples $(\pi, \rho_0, \sigma_0)$ and $(\pi', \rho_0', \sigma_0')$, where $\pi$ is a unitary representation of $G$ on $\mathcal{H}$, $\rho_0, \sigma_0$ are positive definite operators on $\mathcal{H}$ and similarly for $\pi', \rho_0', \sigma_0'$ on $\mathcal{H}'$. The following are equivalent:*

*(i) there exists a sequence of $G$-equivariant completely positive trace-nonincreasing maps $T_n : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{H}'^{\otimes n})$ such that for all $x \in X$ the inequalities*

$$\begin{aligned} T_n(\rho_0^{\otimes n}) &\geq 2^{-o(n)} {\rho_0'}^{\otimes n} \\ T_n(\sigma_0^{\otimes n}) &\leq {\sigma_0'}^{\otimes n} \end{aligned}$$

*hold, with the $o(n)$ uniform in $x$,*

*(ii) $f((\pi(g)\rho_0\pi(g)^*)_{g\in G}, (\pi(g)\sigma_0\pi(g)^*)_{g\in G}) \geq f((\pi(g)\rho_0'\pi(g)^*)_{g\in G}, (\pi(g)\sigma_0'\pi(g)^*)_{g\in G})$ for all $f \in \mathrm{TSper}_1(S_{G,G})$.*

**Remark 4.4.19.** Compact groups include many familiar groups, in particular finite groups and compact Lie groups such as $U(1)$ or $SU(2)$. Theorem 4.4.18 can be applied to any of these, but due to our incomplete knowledge of the test-spectrum, in practice it only gives necessary conditions for catalytic or asymptotic transformations (it would become sufficient if we could evaluate *all* elements of the real and tropical spectrum). Under the additional assumption that the orbit of $\sigma$ consists of commuting operators, Corollary 4.4.21 below gives an explicit necessary and sufficient condition.

We note that in general many elements of $\mathrm{TSper}_1(S_{G,G})$ collapse to the same function when restricted to pairs of the form $((\pi(g)\rho_0\pi(g)^*)_{g\in G}, (\pi(g)\sigma_0\pi(g)^*)_{g\in G})$. The reason is that left translations of $G$ give rise to automorphisms of $S_{G,G}$ of the form $(\rho, \sigma) \mapsto (\rho \circ L_h, \sigma \circ L_h)$ (where $h \in G$ and $L_h : G \to G$ is the map $L_h(g) = hg$), which in turn induce nontrivial automorphisms of $\mathrm{TSper}_1(S_{G,G})$, while the equivalence class of $((\pi(g)\rho_0\pi(g)^*)_{g\in G}, (\pi(g)\sigma_0\pi(g)^*)_{g\in G})$ remains unchanged by these transformations. This can be seen explicitly on the subsemiring of pairs with commuting $\sigma$, where the precise form of spectral points is known: if $\rho(g) = \pi(g)\rho_0\pi(g)^*$ and $\sigma(g) = \pi(g)\sigma_0\pi(g)^*$ such that $\sigma(g)\sigma(e) = \sigma(e)\sigma(g)$ for all $g \in G$, then

$$\begin{aligned}
f_{\alpha,x,\gamma}(\rho\circ L_h,\sigma\circ L_h) &= Q^*_\alpha\left(\rho(hx)\middle\|\exp\int_G \log\sigma(hg)\,\mathrm{d}\gamma(g)\right)\\
&= Q^*_\alpha\left(\pi(h)\rho(x)\pi(h)^*\middle\|\pi(h)\exp\int_G \log\sigma(g)\,\mathrm{d}\gamma(g)\pi(h)^*\right)\\
&= Q^*_\alpha\left(\rho(x)\middle\|\exp\int_G \log\sigma(g)\,\mathrm{d}\gamma(g)\right)\\
&= f_{\alpha,x,\gamma}(\rho,\sigma).
\end{aligned}$$

The first line of this calculation also shows that

$$f_{\alpha,x,\gamma}(\rho\circ L_h,\sigma\circ L_h) = f_{\alpha,hx,(L_h)_*(\gamma)}(\rho,\sigma).$$

In particular, $f_{\alpha,h,\gamma}$ and $f_{\alpha,e,(L_{h^{-1}})_*(\gamma)}$ coincide on these elements.

**Example 4.4.20.**

(i) Let $G = \{0,1\}$ with addition modulo 2. $\mathrm{TSper}_1(S_{G,G})$ contains (at least) the maps $((\rho_0,\rho_1),(\sigma_0,\sigma_1)) \mapsto Q^*_\alpha(\rho_0\|\sigma_0\#_\gamma\sigma_1)$ and $((\rho_0,\rho_1),(\sigma_0,\sigma_1)) \mapsto Q^*_\alpha(\rho_1\|\sigma_0\#_\gamma\sigma_1)$ for all $\alpha \geq 1$ and $\gamma \in [0,1]$. On an element of the form $((\rho, Z\rho Z),(\sigma, Z\sigma Z))$ with $Z^2 = I$, $Z$ unitary, the two are related as

$$\begin{aligned}
Q^*_\alpha(\rho\|\sigma\#_\gamma(Z\sigma Z)) &= \mathrm{Tr}\left(\rho^{1/2}(\sigma\#_\gamma(Z\sigma Z))^{\frac{1-\alpha}{\alpha}}\rho^{1/2}\right)^\alpha\\
&= \mathrm{Tr}\left(Z\rho^{1/2}Z((Z\sigma Z)\#_\gamma\sigma)^{\frac{1-\alpha}{\alpha}}Z\rho^{1/2}Z\right)^\alpha\\
&= \mathrm{Tr}\left((Z\rho Z)^{1/2}(\sigma\#_{1-\gamma}(Z\sigma Z))^{\frac{1-\alpha}{\alpha}}(Z\rho Z)^{1/2}\right)^\alpha\\
&= Q^*_\alpha(Z\rho Z\|\sigma\#_{1-\gamma}(Z\sigma Z)).
\end{aligned}$$

(ii) Let $G = U(1)$, and consider the representation $\pi(e^{it}) = e^{itA}$. An example of a geometric mean in $\mathcal{G}(U(1))$ is the map $M_{t_1,t_2,t_3}(\sigma) = (\sigma(t_1)\#\sigma(t_2))\#\sigma(t_3)$ for some $t_1,t_2,t_3 \in [0,2\pi]$ (for notational simplicity, we identify $[0,2\pi)$ with $U(1)$, see 4.4.17). For any $t$ the maps $(\rho,\sigma) \mapsto Q^*_\alpha(\rho(t)\|M_{t_1,t_2,t_3}(\sigma))$ are in $\mathrm{TSper}_1(S_{U(1),U(1)})$. When evaluated on an element $(\rho,\sigma)$ of the form $\rho(t) = e^{itA}\rho(0)e^{-itA}$, $\sigma(t) = e^{itA}\sigma(0)e^{-itA}$, it gives

$$\begin{aligned}
Q^*_\alpha(\rho(t)\|M_{t_1,t_2,t_3}(\sigma)) &= \mathrm{Tr}\left(\rho(t)^{1/2}M_{t_1,t_2,t_3}(\sigma)^{\frac{1-\alpha}{\alpha}}\rho(t)^{1/2}\right)^\alpha\\
&= \mathrm{Tr}\left(e^{itA}\rho(0)^{1/2}e^{-itA}M_{t_1,t_2,t_3}(\sigma)^{\frac{1-\alpha}{\alpha}}e^{itA}\rho_0^{1/2}e^{-itA}\right)^\alpha\\
&= \mathrm{Tr}\left(\rho(0)^{1/2}M_{t_1,t_2,t_3}(e^{-itA}\sigma e^{itA})^{\frac{1-\alpha}{\alpha}}\rho(0)^{1/2}\right)^\alpha\\
&= \mathrm{Tr}\left(\rho(0)^{1/2}M_{t_1-t,t_2-t,t_3-t}(\sigma)^{\frac{1-\alpha}{\alpha}}\rho(0)^{1/2}\right)^\alpha\\
&= Q^*_\alpha(\rho(0)\|M_{t_1-t,t_2-t,t_3-t}(\sigma)),
\end{aligned}$$

which only depends on the differences $t_1 - t$, $t_2 - t$, and $t_3 - t$.

**Corollary 4.4.21.** *Under the conditions of Theorem 4.4.18, suppose that* $[\sigma_0, \pi(g)\sigma_0\pi(g)^*] = 0$ *and* $[\sigma'_0, \pi(g)\sigma'_0\pi(g)^*] = 0$ *for all* $g \in G$*. Then* $(\pi, \rho_0, \sigma_0) \succsim (\pi', \rho'_0, \sigma'_0)$ *(in the sense of asymptotic equivariant relative submajorization) if and only if for all* $\alpha \geq 0$ *and Radon probability measure* $\gamma$ *on* $G$ *the inequality*

$$D^*_\alpha\left(\rho_0 \middle\| \exp \int_G \log \pi(g)\sigma_0\pi(g)^* \,\mathrm{d}\gamma(g)\right) \geq D^*_\alpha\left(\rho'_0 \middle\| \exp \int_G \log \pi(g)\sigma'_0\pi(g)^* \,\mathrm{d}\gamma(g)\right)$$

*holds.*

**Remark 4.4.22.** The assumption that the orbit of $\sigma$ consists of commuting operators is a strong one due to the following rigidity property: if the orbit of $\sigma$ under the action of a connected group contains only operators that commute with $\sigma$, then $\sigma$ is a fixed point of the action. To see this, note that $\pi(G)$ is a connected Lie subgroup of $U(\mathcal{H})$, therefore the exponential map is surjective. If $iA$ is an element of the Lie algebra, then $[e^{itA}\sigma e^{-itA}, \sigma] = 0$ for all $t \in \mathbb{R}$, which implies by differentiation that $[[A, \sigma], \sigma] = 0$. Since $\sigma$ is diagonalizable, so is the map $X \mapsto [X, \sigma]$, therefore its kernel is equal to the kernel of its square. It follows that $[A, \sigma] = 0$, i.e., $e^{itA}\sigma e^{-itA} = \sigma$ for all $t$.

#### Asymptotic transformations by thermal processes

Thermal operations are central to the resource theoretic approach to quantum thermodynamics. This is the class of quantum channels that can be obtained by preparing Gibbs states at a fixed temperature $T$, performing energy-preserving unitaries and tracing out subsystems [JWZ+00,BHO+13,HO13]. This characterization does not suggest a simple way to decide whether a given channel is a thermal operation or whether a transformation between given states is feasible by a thermal operation, which motivates the study of channels and transformations admitting a simpler description at the cost of satisfying only some of the constraints governing thermal operations.

In the absence of coherence between energy eigenspaces, Gibbs-preserving maps provide an especially useful relaxation, which turns out to allow the same transitions as thermal operations. This is no longer true if coherence is present [FOR15], and in addition to being Gibbs-preserving, the condition of time-translation symmetry has been identified as another key property of thermal operations [LJR15]. Adding this requirement leads to the notion of thermal processes [GJB+18].

Transformations by such processes are an instance of equivariant relative majorization: the group is that of time-translations, isomorphic to $\mathbb{R}$, and to a system with Hilbert space $\mathcal{H}$, Hamiltonian $H \in \mathcal{B}(\mathcal{H})$ and state $\rho$ we associate the triple $(\pi_H, \rho, e^{-\beta H})$, where $\pi_H : \mathbb{R} \to U(\mathcal{H})$ is the representation $t \mapsto e^{-itH}$ and $\beta$ is the inverse temperature. By definition, $(\pi_H, \rho, e^{-\beta H}) \succcurlyeq (\pi_H, \sigma, e^{-\beta H})$ if there is a completely positive trace-preserving map $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ such that $T(\rho) = \sigma$, $T(e^{-\beta H}) = e^{-\beta H}$ (by linearity, this amounts to preserving the Gibbs state $e^{-\beta H}/\operatorname{Tr} e^{-\beta H}$), and $T(e^{-itH}\omega e^{itH}) = e^{-itH}T(\omega)e^{itH}$ for all states $\omega$ and $t \in \mathbb{R}$.

If we relax these transformations to equivariant relative *sub*majorization and consider the asymptotic limit, then Theorem 4.4.18 provides a characterization of the resulting preorder in terms of the spectrum $\operatorname{TSper}_1(S_{\mathrm{b}(\mathbb{R}),\mathrm{b}(\mathbb{R})})$. Moreover, since $e^{-itH}e^{-\beta H}e^{itH} = e^{-\beta H}$, the orbit of $e^{-\beta H}$ has only one element, the simpler characterization of Corollary 4.4.21 can be applied.

**Proposition 4.4.23.** *Let $H \in \mathcal{B}(\mathcal{H})$ be a Hamiltonian on a Hilbert space $\mathcal{H}$, and $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Then the following are equivalent:*

*(i) there exists a sequence of trace-nonincreasing thermal processes $T_n : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{H}^{\otimes n})$ such that $T_n(\rho^{\otimes n}) \geq 2^{-o(n)}\sigma^{\otimes n}$,*

*(ii) for all $\alpha \geq 1$ the inequality $D^*_\alpha\left(\rho \middle\| e^{-\beta H}\right) \geq D^*_\alpha\left(\sigma \middle\| e^{-\beta H}\right)$ holds.*

*Proof.* As explained in Section 4.4.2, we replace the noncompact group $\mathbb{R}$ with its Bohr compactification $\mathrm{b}(\mathbb{R})$ and consider the unique representation $\pi : \mathrm{b}(\mathbb{R}) \to U(\mathcal{H})$ such that $\pi(b(t)) = e^{-itH}$, where $b : \mathbb{R} \to \mathrm{b}(\mathbb{R})$ is the universal map. We claim that the image of $\pi$ is the closure of $\left\{e^{-itH} : t \in \mathbb{R}\right\}$, which we denote by $K$. Since $\pi(\mathrm{b}(\mathbb{R}))$ is compact (hence closed) and contains $e^{-itH}$ for all $t$, it must contain $K$ as well. $K$ is a closed subgroup of $U(\mathcal{H})$, therefore compact. By the universal property, there is map $\varphi : \mathrm{b}(\mathbb{R}) \to K$ such that $\varphi(b(t)) = e^{-itH}$ for all $t \in \mathbb{R}$. Composing with the inclusion $\iota : K \to U(\mathcal{H})$ we get $\pi(b(t))$ for all $t \in \mathbb{R}$, so by uniqueness we have $\iota \circ \varphi = \pi$.

By continuity, we have $\pi(g)e^{-\beta H}\pi(g)^* = e^{-\beta H}$ for every $g \in \mathrm{b}(\mathbb{R})$, therefore

$$\exp \int_{\mathrm{b}(\mathbb{R})} \log \pi(g) e^{-\beta H} \pi(g)^* \,\mathrm{d}\gamma(g) = e^{-\beta H}$$

for every probability measure $\gamma$ on $\mathrm{b}(\mathbb{R})$. This means that the condition $D^*_\alpha\left(\rho \middle\| e^{-\beta H}\right) \geq D^*_\alpha\left(\sigma \middle\| e^{-\beta H}\right)$ for all $\alpha \geq 1$ is equivalent to the condition in Corollary 4.4.21.

To see the equivalence as stated, note that for any trace-nonincreasing equivariant Gibbs subpreserving completely positive map $T_n$ there exists a trace-preserving thermal process $T'_n$ such that $T_n \leq T'_n$ (in the completely positive partial order), for example the map

$$T'_n(X) = T_n(X) + (\operatorname{Tr} X - \operatorname{Tr} T_n(X)) \frac{(e^{-\beta H})^{\otimes n} - T_n((e^{-\beta H})^{\otimes n})}{\operatorname{Tr}(e^{-\beta H})^{\otimes n} - \operatorname{Tr} T_n((e^{-\beta H})^{\otimes n})}.$$

$T'_n$ preserves the Gibbs state $(e^{-\beta H})^{\otimes n} / \operatorname{Tr}\left(e^{-\beta H}\right)^{\otimes n}$ by construction, it is the sum of $T_n$ and a completely positive equivariant map (since $(e^{-\beta H})^{\otimes n}$ is invariant and therefore its image under the equivariant map $T_n$ is invariant as well). Finally, we have $T'_n(\rho^{\otimes n}) \geq T_n(\rho^{\otimes n}) \geq 2^{-o(n)}\sigma^{\otimes n}$. □

We note that this characterization is exactly the same as the one obtained in [PVW22] for Gibbs-preserving maps without the equivariance condition. This means that in this asymptotic limit, Gibbs-preserving maps are no more powerful than thermal processes.

### Hypothesis testing with group symmetry

In the group-symmetric variant of binary state discrimination problems, the measurements are restricted to be invariant with respect to a group representation $\pi : G \to U(\mathcal{H})$. This problem was considered in [HMH09] in three asymptotic regimes: when both error probabilities decay exponentially with the same exponent (Chernoff); when the exponential decays are different (Hoeffding); and when the type I error probability approaches zero arbitrarily, and the type II error decays exponentially (Stein).

We now focus on the strong converse domain. In the unrestricted (i.e., without group symmetry) case, this is characterized in Corollary 4.4.16, providing an operational interpretation of the sandwiched Rényi divergences with orders $\alpha > 1$. Following the strategy of [PVW22], we show that the strong converse error exponent can be characterized in terms of the asymptotic preorder associated with equivariant relative submajorization.

We begin with single-copy measurements. Given a representation $\pi : G \to U(\mathcal{H})$, by a group-symmetric POVM (or invariant measurement) we mean a POVM $(F_i)_{i=1}^r$ with measurement operators $F_i$ satisfying $\pi(g)F_i\pi(g)^* = F_i$ for all $g \in G$. Since we are interested in discriminating between two hypotheses, the measurement will be a test, i.e., a two-outcome POVM with measurement operators $(\Pi, I - \Pi)$. We interpret the outcome associated with $\Pi$ as accepting (not rejecting) the null hypothesis $\rho$, and the other outcome corresponds to accepting $\sigma$. The probability of a type I error is therefore $\operatorname{Tr}(I - \Pi)\rho = 1 - \operatorname{Tr}\Pi\rho$, while that of the type II error is $\operatorname{Tr}\Pi\sigma$.

**Lemma 4.4.24.** *Let $\pi : G \to U(\mathcal{H})$ be a representation and $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Then the following are equivalent for $a, b \in \mathbb{R}_{>0}$:*

*(i) there exists a two-outcome group-symmetric POVM $(\Pi, I - \Pi)$ on $\mathcal{H}$ such that the type I and type II errors satisfy*

$$\begin{aligned} \operatorname{Tr}(I - \Pi)\rho &\leq 1 - a, \\ \operatorname{Tr}\Pi\sigma &\leq b. \end{aligned}$$

*(ii) $(\pi, \rho, \sigma) \succcurlyeq (1, a, b)$, where $1$ denotes the trivial representation of $G$ on $\mathbb{C}$, and we identify $\mathcal{B}(\mathbb{C}) \simeq \mathbb{C}$.*

*Proof.* An invariant measurement on $\mathcal{H}$ can be identified with an equivariant (completely) positive trace-nonincreasing map $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathbb{C}) \simeq \mathbb{C}$, where $\mathbb{C}$ carries the trivial representation $1$: the POVM $(\Pi, I - \Pi)$ corresponds to the map $T(X) = \operatorname{Tr}(X\Pi)$. This is indeed equivariant (which in this special case means invariant), since

$$\begin{aligned} T(\pi(g)X\pi(g)^*) &= \operatorname{Tr}(\pi(g)X\pi(g)^*\Pi) \\ &= \operatorname{Tr}\big(X\pi(g^{-1})\Pi\pi(g^{-1})^*\big) \\ &= \operatorname{Tr}(X\Pi) = T(X). \end{aligned}$$

Conversely, any equivariant linear map $T : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ is of the form $T(X) = \operatorname{Tr}(X\Pi)$ for a unique $\Pi$ that is necessarily invariant, and $T$ is completely positive and trace-nonincreasing if and only if $0 \leq \Pi \leq I$. It follows that $(\pi, \rho, \sigma) \succcurlyeq (1, a, b)$ for some $a, b > 0$ if and only if there is a POVM $(\Pi, I - \Pi)$ such that the type I error satisfies $1 - \operatorname{Tr}\Pi\rho \leq 1 - a$ and the type II error satisfies $\operatorname{Tr}\Pi\sigma \leq b$. $\square$

The following is an immediate consequence of Lemma 4.4.24, the definition of the asymptotic preorder, Theorem 4.4.18, and Corollary 4.4.21:

**Corollary 4.4.25.** *Let $G$ be a compact group and $\pi : G \to U(\mathcal{H})$ a representation and $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Then the following are equivalent for $r, R \geq 0$:*

*(i) there exists a sequence of $\pi^{\otimes n}$-invariant measurements $(\Pi_n, I-\Pi_n)$ with $\Pi_n \in \mathcal{B}(\mathcal{H}^{\otimes n})$ such that the type I error is at most $1-2^{-Rn+o(n)}$ and the type II error is at most $2^{-rn+o(n)}$;*

*(ii) $(\pi,\rho,\sigma) \gtrsim (1, 2^{-R}, 2^{-r})$;*

*(iii) $\forall f \in \mathrm{TSper}_1(S_{G,G}) : f((\pi(g)\rho\pi(g)^*)_{g\in G}, (\pi(g)\sigma\pi(g)^*)_{g\in G}) \geq f(2^{-R}, 2^{-r})$.*

*Moreover, if the orbit of $\sigma$ consists of operators that commute with $\sigma$, then the condition is equivalent to*

$$R \geq \sup_{\alpha>1}\max_{\gamma} \frac{\alpha-1}{\alpha}\left[r - D^*_\alpha\left(\rho\middle\|\exp\int_G \log\pi(g)\sigma\pi(g)^*\,\mathrm{d}\gamma(g)\right)\right], \tag{4.13}$$

*where the maximum is over Radon probability measures $\gamma$ on $G$.*

In the following, $R^*(r)$ denotes the smallest possible error exponent for a given $r$.

**Example 4.4.26.** Consider the states and representation as in [HMH09, Example 6.1]: The group is $G=\{0,1\}$ with addition modulo 2, $\mathcal{H}=\mathbb{C}^2$, and the nontrivial element acts as the Pauli $Z$ matrix. Let the states be

$$\begin{aligned}\rho &= \lambda\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} + (1-\lambda)\frac{1}{2}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\\ \sigma &= \mu\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} + (1-\mu)\frac{1}{2}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\end{aligned}$$

for some $\lambda,\mu\in(0,1)$. Then $Z\sigma Z$ is also of the same form with $\mu$ replaced with $1-\mu$, and these states commute with each other. From (4.13) we obtain

$$\begin{aligned}R^*(r) = \sup_{\alpha>1}\max_{\gamma\in[0,1]}\frac{\alpha-1}{\alpha}\Big[r - \frac{1}{\alpha-1}\log\Big(&\lambda^\alpha\mu^{(1-\alpha)\gamma}(1-\mu)^{(1-\alpha)(1-\gamma)}\\ &+(1-\lambda)^\alpha(1-\mu)^{(1-\alpha)\gamma}\mu^{(1-\alpha)(1-\gamma)}\Big)\Big].\end{aligned}$$

This can be strictly larger than the maximum of the strong converse exponents for testing $\rho$ against $\sigma$ or $Z\sigma Z$, which correspond to $\gamma=0$ and $\gamma=1$. Note that a similar strict inequality was shown for unrestricted hypothesis testing with a composite alternative hypothesis in [MSW22].

**Example 4.4.27.** Let $G=\{0,1\}$, and $\pi : G\to U(\mathcal{H})$, $\rho,\sigma\in\mathcal{S}(\mathcal{H})$ arbitrary full-rank, and let $Z=\pi(1)$. Without the commuting orbit assumption, we can only give a lower bound on the strong converse exponent, corresponding to the inner approximation of the test-spectrum given by Theorem 4.3.19 and Lemma 4.3.16:

$$R^*(r) \geq \sup_{\alpha>1}\max_{\gamma\in[0,1]}\frac{\alpha-1}{\alpha}[r - D^*_\alpha(\rho\|\sigma\#_\gamma(Z\sigma Z))].$$

We do not know whether this inequality can be strict for some $Z,\rho,\sigma$.

**Example 4.4.28.** Let $G$ be an arbitrary compact group, $\pi : G\to U(\mathcal{H})$, $\rho\in\mathcal{S}(\mathcal{H})$ arbitrary an suppose that $\sigma\in\mathcal{S}(\mathcal{H})$ is $G$-invariant, i.e., $\pi(g)\sigma\pi(g)^*=\sigma$ for all $g\in G$ (as always, both $\rho$ and $\sigma$ are assumed to be of full rank). Then the orbit of $\sigma$ is a single point, therefore (4.13) may be used and simplified because the exponentiated integral is equal to $\sigma$ independently of the measure $\gamma$. Therefore the bound reduces to

$$R^*(r) = \sup_{\alpha>1} \frac{\alpha-1}{\alpha}[r - D^*_\alpha(\rho\|\sigma)],$$

which is the same as with unrestricted measurements [MO15]. This phenomenon is similar to the equality of the Stein exponents in the group-symmetric and unrestricted problems, and is in contrast with the Chernoff and Hoeffding exponents, which can be strictly worse with group-symmetric tests, even if $\sigma$ is $G$-invariant [HMH09, Example 6.2].

**Reference frames in hypothesis testing**

When the dynamics of a system is constrained by symmetries, an additional supply of asymmetric states (imperfect reference frames) becomes a resource, which allows to partially overcome the limitations of symmetric evolutions [BRS07]. Suppose that $\pi_{\mathrm{ref}} : G \to U(\mathcal{K})$ is a representation and $\Omega \in \mathcal{S}(\mathcal{K})$ is a state with full support and trivial stabilizer. In the setting of group-symmetric hypothesis testing as modeled above in terms of equivariant relative submajorization, the reference frame corresponds to the triple $(\pi_{\mathrm{ref}}, \Omega, \Omega)$. The two states in this triple are equal, therefore they cannot be distinguished even with unrestricted measurements. The mathematical property that makes it a useful resource in group-symmetric hypothesis testing is the following:

**Proposition 4.4.29.** *Let $\pi_{\mathrm{ref}} : G \to U(\mathcal{K})$ be a representation and $\Omega \in \mathcal{S}(\mathcal{K})$ a state with full support and trivial stabilizer. Let $f \in \mathrm{TSper}_1(S_{G,G})$ and let its restriction to $S^{\mathrm{c}}_{G,G}$ be characterized by $\alpha > 1$ (for real points), $x \in G$, and the probability measure $\gamma$. Then there are two possibilities: either $\gamma$ is the Dirac measure at $x$, or*

$$f((\pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^*)_{g\in G}, (\pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^*)_{g\in G}) > 1.$$

*Proof.* Since the stabilizer of $\Omega$ is trivial, for all $g, g' \in G$, $g \neq g'$ we have $\pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^* \neq \pi_{\mathrm{ref}}(g')\Omega\pi_{\mathrm{ref}}(g')^*$. By assumption, $\mathrm{Tr}\,\pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^* = \mathrm{Tr}\,\Omega = 1$, so by Proposition 4.3.4, the condition for $f((\pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^*)_{g\in G}, (\pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^*)_{g\in G}) = 1$ is that $\gamma$ is concentrated at $x$. □

**Example 4.4.30.**

(i) Let $G = \{0,1\}$, and let the nontrivial element act on $\mathbb{C}^2$ by the Pauli $X$ operator. A biased coin $\Omega = q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|$ with $q \in (0, 1/2)$ is an example of a reference frame. Since $[\Omega, X\Omega X] = 0$, we can use the explicit form of the spectrum of the classical semiring $S^{\mathrm{c}}_{G,G}$. Identifying the probability measure $\gamma$ with the value $\gamma(\{0\}) \in [0,1]$, for $\alpha > 1$ we have

$$\begin{aligned}
&f_{\alpha,0,\gamma}((\Omega, X\Omega X), (\Omega, X\Omega X)) \\
&\qquad = q^\alpha q^{(1-\alpha)\gamma}(1-q)^{(1-\alpha)(1-\gamma)} + (1-q)^\alpha (1-q)^{(1-\alpha)\gamma} q^{(1-\alpha)(1-\gamma)} \\
&\qquad\qquad = q^{1-(1-\alpha)(1-\gamma)}(1-q)^{(1-\alpha)(1-\gamma)} + (1-q)^{1-(1-\alpha)(1-\gamma)} q^{(1-\alpha)(1-\gamma)},
\end{aligned}$$

which is equal to 1 if and only if $\gamma = 1$.

(ii) Let $G = U(1)$, which we can think of as the group of complex numbers with modulus 1. To find a reference frame, we consider the representation

$$\pi(z) = \begin{bmatrix} 1 & 0 \\ 0 & z \end{bmatrix},$$

and let $\Omega$ be any full-rank state with nonzero off-diagonal entries, such as $\Omega = (1-\epsilon)|+\rangle\langle+| + \epsilon|-\rangle\langle-|$, where $\epsilon \in (0, 1/2)$ and $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Conjugation by $\pi(z)$ means that the off-diagonal elements pick up the phase factor $z$ and $\overline{z}$, therefore the stabilizer is trivial.

We can find a lower bound on the values of the spectral points by applying a suitable measurement. Let us measure in the basis $(|+\rangle, |-\rangle)$. The probability of obtaining the plus outcome in the state $\pi(z)\Omega\pi(z)^*$ is

$$\frac{1}{4}\operatorname{Tr}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & (1-2\epsilon)\overline{z} \\ (1-2\epsilon)z & 1 \end{bmatrix} = \frac{1}{2} + \frac{1-2\epsilon}{2}\operatorname{Re} z,$$

which has a strict maximum at $z = 1$, therefore this measurement gives a nontrivial bound for any probability measure $\gamma \neq \delta_1$. More precisely, if $f \in \operatorname{TSper}_1(S_{U(1),U(1)})$, and its restriction to the classical subsemiring is characterized by $\alpha > 1$, $x = 1 \in U(1)$ and the probability measure $\gamma$, then

$$\begin{aligned}
&f((\pi(z)\Omega\pi(z)^*)_{z\in U(1)}, (\pi(z)\Omega\pi(z)^*)_{z\in U(1)}) \\
\geq &f_{\alpha,1,\gamma}((\frac{1}{2} + \frac{1-2\epsilon}{2}\operatorname{Re} z)_{z\in U(1)}, (\frac{1}{2} + \frac{1-2\epsilon}{2}\operatorname{Re} z)_{z\in U(1)}) \\
= &(1-\epsilon)^\alpha \exp\left[(1-\alpha)\int_0^{2\pi} \log\left(\frac{1}{2} + \frac{1-2\epsilon}{2}\cos\varphi\right) \mathrm{d}\gamma(\varphi)\right] \\
&\quad + \epsilon^\alpha \exp\left[(1-\alpha)\int_0^{2\pi} \log\left(\frac{1}{2} - \frac{1-2\epsilon}{2}\cos\varphi\right) \mathrm{d}\gamma(\varphi)\right].
\end{aligned}$$

Testing $(\pi, \rho, \sigma)$ aided by the reference frame corresponds to comparing $(\pi \otimes \pi_{\mathrm{ref}}, \rho \otimes \Omega, \sigma \otimes \Omega)$ with a one-dimensional triple with trivial representation. In an asymptotic setting, more generally, we may use $\kappa$ copies of the reference frame per sample of the state to be discriminated. In this case the exponent pair $(R, r)$ is achievable if and only if

$$\begin{aligned}
\forall f \in \operatorname{TSper}_1(S_{G,G}) : &f((\pi(g)\rho_0\pi(g)^*)_{g\in G}, (\pi(g)\sigma_0\pi(g)^*)_{g\in G}) \\
&\cdot f((\pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^*)_{g\in G}, (\pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^*)_{g\in G})^\kappa \geq f(2^{-R}, 2^{-r}),
\end{aligned}$$

therefore

$$\begin{aligned}
R^*(r,\kappa) = \sup_{\alpha>1} \sup_{\substack{f\in\operatorname{TSper}_1(S_{G,G}) \\ \log f(2,1)=\alpha}} \frac{\alpha-1}{\alpha}\Bigg[r - \frac{1}{\alpha-1}\log f((\pi(g)\rho_0\pi(g)^*)_{g\in G}, (\pi(g)\sigma_0\pi(g)^*)_{g\in G}) \\
- \kappa\frac{1}{\alpha-1}\log f((\pi(g)\Omega\pi(g)^*)_{g\in G}, (\pi(g)\Omega\pi(g)^*)_{g\in G})\Bigg],
\end{aligned}$$

where we may as well restrict to those elements of the spectrum that depend on the first family through its value at the identity. As $\kappa \to \infty$ (i.e., in the limit of unlimited supply of the reference frame), the supremum is achieved for the $\gamma$ (which, as before, is determined by the restriction of $f$ to the classical subsemiring) that is concentrated at the identity by Proposition 4.4.29, since this is the only point where the last term vanishes. This means that we recover the unrestricted strong converse exponent [MO15], which is potentially much smaller than the group-symmetric one. In an extreme example, $\rho_0$ and $\sigma_0$ might be in the same $G$-orbit, in which case $R^*(r, 0) = r$, i.e., a group-symmetric measurement cannot offer any advantage over guessing.

When the orbits of $\sigma_0$ and $\Omega$ consist of commuting operators, we can use Corollary 4.4.21 to obtain an explicit form of the smallest type I strong converse exponent $R^*$ for a given decay rate $r$ of the type II error:

$$R^*(r,\kappa) = \sup_{\alpha>1} \max_{\gamma} \frac{\alpha-1}{\alpha} \left[ r - D^*_\alpha\left(\rho_0 \middle\| \exp \int_G \log \pi(g)\sigma_0\pi(g)^* \, \mathrm{d}\gamma(g)\right) \right.$$
$$\left. - \kappa D^*_\alpha\left(\Omega \middle\| \exp \int_G \log \pi_{\mathrm{ref}}(g)\Omega\pi_{\mathrm{ref}}(g)^* \, \mathrm{d}\gamma(g)\right)\right].$$

### 4.4.3 Approximate joint transformations

In this subsection we specialize our results and derive a characterization of approximate joint transformations with respect to the symmetrized max divergence, in the asymptotic limit. Recall that the max divergence between a pair of states $\rho, \sigma$ is $D_{\max}(\rho\|\sigma) = \log \left\|\sigma^{-1/2}\rho\sigma^{-1/2}\right\|_\infty = \min\left\{\lambda \in \mathbb{R} : 2^\lambda \sigma \geq \rho\right\}$. We will use the max divergence as a measure of dissimilarity between states. It vanishes if and only if the two states are equal, but for subnormalized states this is no longer true, and it is not symmetric. However, the closely related quantity $d_{\mathrm{T}}(\rho, \sigma) := \max\{D_{\max}(\rho\|\sigma), D_{\max}(\sigma\|\rho)\}$ is a metric on the set of positive definite operators (the Thompson metric [Tho63] associated with the semidefinite cone). This metric is unbounded even on a fixed Hilbert space and satisfies $d_{\mathrm{T}}(\rho^{\otimes n}, \sigma^{\otimes n}) = n d_{\mathrm{T}}(\rho, \sigma)$. The notion of approximate transformations that we consider will be that the distance increases sublinearly as the number of copies grow.

**Proposition 4.4.31.** *Let $X$ be a compact space, $\rho \in C(X, \mathcal{B}(\mathcal{H})_{>0})$, and $\rho' \in C(X, \mathcal{B}(\mathcal{H}')_{>0})$. The following are equivalent:*

*(i) there is a sequence of completely positive trace-nonincreasing maps $T_n : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{H}'^{\otimes n})$ such that for all $n$*
$$2^{-o(n)}\rho'^{\otimes n} \leq T_n(\rho^{\otimes n}) \leq \rho'^{\otimes n},$$
*i.e., for all $x \in X$ we have*
$$\lim_{n\to\infty} \frac{1}{n} d_{\mathrm{T}}(T_n(\rho(x)^{\otimes n}), \rho'(x)^{\otimes n}) = 0,$$
*uniformly in $x$;*

*(ii) $(\rho, \rho) \gtrsim (\rho', \rho')$ in the preordered semiring $S_{X,X}$;*

*(iii) for all $f \in \mathrm{TSper}_1(S_{X,X})$ the inequality $f(\rho, \rho) \geq f(\rho', \rho')$ holds.*

*Proof.* The equivalence of (ii) and (iii) is a special case of Theorem 2.3.26. Using the definition of the asymptotic preorder with the power universal element $u = (2 \cdot 1_X, 1_X)$, the condition for $(\rho, \rho) \gtrsim (\rho', \rho')$ is that there is a sequence of completely positive trace-nonincreasing maps $T_n : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{H}'^{\otimes n})$ and a sequence $k_n$ of natural numbers such that $k_n/n \to 0$ and for all $n \in \mathbb{N}$ and $x \in X$ the inequalities $T_n(2^{k_n}\rho(x)) \geq \rho'(x)$ and $T_n(\rho(x)) \leq \rho'(x)$ hold, which is the same as (i). $\square$

**Remark 4.4.32.** The (non-asymptotic) relative submajorization preorder between the pairs $(\rho, \rho)$ and $(\rho', \rho')$ means that $T(\rho(x)) \leq \rho'(x)$ and $T(\rho(x)) \geq \rho'(x)$ for some subchannel $T$ and all $x$, i.e., $T(\rho(x)) = \rho'(x)$.

Specializing to classical families and using the explicit form of the 1-test spectrum of the classical semiring (Theorem 4.2.9), we have the following characterization of asymptotic joint transformations in the above sense.

**Theorem 4.4.33.** *Let $p : X \to \mathcal{P}([d])$, $p' : X \to \mathcal{P}([d'])$ where $X$ is a compact Hausdorff space and $d, d' \in \mathbb{N}_{>0}$ are finite sets. The following are equivalent:*

*(i) there exists a sequence of substochastic maps $T_n$ from $(\mathbb{R}^d)^{\otimes n}$ to $(\mathbb{R}^{d'})^{\otimes n}$ such that for all $x \in X$*

$$\lim_{n\to\infty} \frac{1}{n} d_{\mathrm{T}}(T_n(p(x)^{\otimes n}), p'(x)^{\otimes n}) = 0,$$

*uniformly in $x$;*

*(ii) for all $x \in X$, $\alpha \geq 1$ and probability measure $\gamma$ on $X$ the inequality*

$$\sum_{i=1}^{d} p_i(x)^\alpha \exp(1-\alpha) \int_X \log p_i \,\mathrm{d}\gamma \geq \sum_{i=1}^{d'} p'_i(x)^\alpha \exp(1-\alpha) \int_X \log p'_i \,\mathrm{d}\gamma$$

*holds.*

The ideas in this section can be combined with our considerations on equivariant transformations:

**Proposition 4.4.34.** *Let $X_0$ be a compact space, $G$ a compact group, $\pi : G \to U(\mathcal{H})$, $\pi' : G \to U(\mathcal{H}')$ unitary representations, $\rho_0 \in C(X_0, \mathcal{B}(\mathcal{H})_{>0})$ and $\rho'_0 \in C(X_0, \mathcal{B}(\mathcal{H}')_{>0})$. Let $X = X_0 \times G$ and consider $\rho \in C(X, \mathcal{B}(\mathcal{H})_{>0})$ defined as $\rho(x, g) = \pi(g)\rho_0(x)\pi(g)^*$, and similarly $\rho'$, which determine the elements $(\rho, \rho)$ and $(\rho', \rho')$ in $S_{X,X}$. Then the following are equivalent:*

*(i) there exists an equivariant completely positive trace-nonincreasing map $T$ such that $T(\rho_0(x)) = \rho'_0(x)$ for all $x \in X_0$;*

*(ii) $(\rho, \rho) \succcurlyeq (\rho', \rho')$ in $S_{X,X}$.*

*Proof.* (i) $\Longrightarrow$ (ii): Let $T$ be an equivariant subchannel such that $T(\rho_0(x)) = \rho'_0(x)$ for all $x \in X_0$. Then for all $x \in X_0$ and $g \in G$ we have

$$T(\rho(x,g)) = T(\pi(g)\rho_0(x)\pi(g)^*) = \pi(g)T(\rho_0(x))\pi(g)^* = \pi(g)\rho'_0(x)\pi(g)^* = \rho'(x,g).$$

Therefore $(\rho, \rho) \succcurlyeq (\rho', \rho')$ (see Remark 4.4.32).

(ii) $\implies$ (i): Suppose that $(\rho, \rho) \succcurlyeq (\rho', \rho')$ in $S_{X,X}$. Then there exists a subchannel $T_0$ such that $T_0(\rho(x,g)) = \rho'(x,g)$ for all $x \in X_0$ and $g \in G$. Let us define the completely positive trace-nonincreasing map $T$ as

$$T(\sigma) = \int_G \pi'(g)^* T_0(\pi(g)\sigma\pi(g)^*)\pi'(g)\,\mathrm{d}\mu(g),$$

where $\mu$ is the Haar probability measure on $G$. This is equivariant by construction, and satisfies

$$\begin{aligned} T(\rho_0(x)) &= \int_G \pi'(g)^* T_0(\pi(g)\rho_0(x)\pi(g)^*)\pi'(g)\,\mathrm{d}\mu(g) \\ &= \int_G \pi'(g)^* T_0(\rho(x,g))\pi'(g)\,\mathrm{d}\mu(g) \\ &= \int_G \pi'(g)^* \rho'(x,g)\pi'(g)\,\mathrm{d}\mu(g) \\ &= \int_G \rho'_0\,\mathrm{d}\mu(g) = \rho'_0(x). \end{aligned}$$

□

While a complete classification of $\mathrm{TSper}_1(S_{X,X})$ is required to obtain an explicit characterization of the asymptotic preorder, any subset of the spectrum gives necessary conditions for asymptotic, single or multiple copy, and catalytic transformations. In particular, in the setting of Section 4.4.2, the spectrum gives rise to many "second laws" of thermodynamics in the sense of [BHN+15]. For example, if a transformation is possible by a thermal process under the Hamiltonian $H$ and at temperature $\beta^{-1}$, then the value of

$$D^*_\alpha\big(\rho\big\|(e^{-itH}\rho e^{itH})\#_\gamma e^{-\beta H}\big) \tag{4.14}$$

cannot increase under the process, for every $\alpha \geq 1$, $\gamma \in [0,1]$ and $t \in \mathbb{R}$. Here the second argument may be replaced with any weighted geometric mean of the Gibbs state and arbitrary time-translated versions of $\rho$, and in addition the first argument may be replaced with the Gibbs state. To ensure that the quantity is finite, one generally needs to restrict to full-rank states $\rho$.

**Example 4.4.35.** As a concrete example, consider a transition studied in [FOR15], perturbed slightly to get full-rank states. With the Hamiltonian $H = |1\rangle\langle 1|$ on $\mathbb{C}^2$, the transition $|1\rangle\langle 1| \to |+\rangle\langle +|$ was shown to be possible by Gibbs-preserving maps, but not possible with thermal processes. Let $\tau = e^{-\beta H}/\operatorname{Tr} e^{-\beta H}$ be the Gibbs state at temperature $\beta^{-1}$. Then the transition $(1-\epsilon)|1\rangle\langle 1| + \epsilon\tau \to (1-\epsilon)|+\rangle\langle +| + \epsilon\tau$ is still possible by a Gibbs-preserving map. However, (4.14) with $\alpha = 2$, $\gamma = 1/2$ and $t = \pi$ evaluates to $\frac{1}{2}\log\big(1+e^{\beta}\big) + O(\epsilon)$ on the initial state while it diverges logarithmically on the target state as $\epsilon \to 0$. This implies that, for sufficiently small $\epsilon$, the transition is not possible under a thermal process, even in the presence of a catalyst or assuming multicopy transformations. In fact, a numerical comparison suggests that this holds for all $\epsilon \in (0,1)$.

### 4.4.4 A two-parameter family of quantum Rényi divergences

A defining property of the monotone quantities in $\mathrm{TSper}_1(S_{X,Y})$ is that they are increasing in the first argument and decreasing in the second one. From the point of view of relative majorization, this is a severe and unnecessary restriction, and it is reasonable to expect that by dropping this requirement one gets more constraints on joint transformations that are violated by relative submajorization.

In this section we point out that it is possible to derive some of these additional constraints by specialization, thanks to the possibility of relative submajorization to express relative majorization as a special case. We also used this in Section 4.4.3 for classical families, but now, with a different viewpoint, we consider quantum pairs instead, and introduce a two-parameter family of monotone quantum Rényi divergences. We note that $\alpha$-$z$-divergences, another two-parameter quantum extension of the Rényi divergences defined and summarized in [AD15], do not seem have any obvious relation to ours.

**Definition 4.4.36.** For $\gamma \in [0,1)$ and $\alpha > 1$, let us define the $\gamma$-weighted geometric sandwiched Rényi divergence for positive definite arguments $\rho, \sigma$ as

$$\begin{aligned} D_\alpha^{*,\gamma}(\rho\|\sigma) &:= \frac{1}{1-\gamma} D^*_{\frac{\alpha-\gamma}{1-\gamma}}(\rho\|\sigma\#_\gamma\rho) \\ &= \frac{1}{\alpha-1}\log\operatorname{Tr}\left(\sqrt{\rho}\left(\sqrt{\sigma}\left(\sigma^{-1/2}\rho\sigma^{-1/2}\right)^\gamma\sqrt{\sigma}\right)^{\frac{1-\alpha}{\alpha-\gamma}}\sqrt{\rho}\right)^{\frac{\alpha-\gamma}{1-\gamma}}. \end{aligned} \tag{4.15}$$

The $\alpha \to 1$ limit is studied in a more general context in [MBV22]. We summarize some properties in the following proposition:

**Proposition 4.4.37.** *Let $\gamma \in [0,1)$ and $\alpha > 1$. Then for all $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{>0}$ and $\rho', \sigma' \in \mathcal{B}(\mathcal{H}')_{>0}$ we have*

*(i) if $[\rho, \sigma] = 0$, then $D_\alpha^{*,\gamma}(\rho\|\sigma) = \frac{1}{\alpha-1}\operatorname{Tr}\rho^\alpha\sigma^{1-\alpha}$ (extension of classical Rényi divergence);*

*(ii) $D_\alpha^{*,\gamma}(\rho\otimes\rho'\|\sigma\otimes\sigma') = D_\alpha^{*,\gamma}(\rho\|\sigma) + D_\alpha^{*,\gamma}(\rho'\|\sigma')$ (additivity);*

*(iii) $2^{(\alpha-1)D_\alpha^{*,\gamma}(\rho\oplus\rho'\|\sigma\oplus\sigma')} = 2^{(\alpha-1)D_\alpha^{*,\gamma}(\rho\|\sigma)} + 2^{(\alpha-1)D_\alpha^{*,\gamma}(\rho'\|\sigma')}$ (block additivity);*

*(iv) if there is a channel $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ such that $T(\rho) = \rho'$ and $T(\sigma) = \sigma'$, then $D_\alpha^{*,\gamma}(\rho\|\sigma) \geq D_\alpha^{*,\gamma}(\rho'\|\sigma')$ (data processing inequality);*

*(v) if $\operatorname{Tr}\rho = \operatorname{Tr}\sigma = 1$, then $D_\alpha^{*,\gamma}(\rho\|\sigma) \geq 0$ with equality if and only if $\rho = \sigma$ (positive definiteness);*

*(vi) if $\mathcal{H} = \mathcal{H}'$ and $\sigma \leq \sigma'$ in the semidefinite partial order, then $D_\alpha^{*,\gamma}(\rho\|\sigma) \geq D_\alpha^{*,\gamma}(\rho\|\sigma')$ (anti-monotonicity in the second argument).*

*Proof.* Although the properties can be proved directly, in order to illustrate the general idea, we show them by a reduction to Theorem 4.3.19 and Proposition 4.3.16 that give us a partial description of the test-spectrum $\operatorname{TSper}_1(S_{X,Y})$ for any spaces $X$ and $Y$. Let $X = \{1\}$ and $Y = \{1, 2\}$, so that the elements of $S_{X,Y}$ may be written as $(\rho, (\sigma_1, \sigma_2))$. Then for $\alpha > 1$, $\alpha_0 := \frac{\alpha-\gamma}{1-\gamma}$ and $\gamma \in [0, 1]$ the functional

$$f(\rho, (\sigma_1, \sigma_2)) = Q^*_{\alpha_0}(\rho\|\sigma_2\#_\gamma\sigma_1)$$

belongs to $\operatorname{TSper}_1(S_{\{1\},\{1,2\}})$ by Theorem 4.3.19 and Proposition 4.3.16. It is related to the $\gamma$-weighted geometric sandwiched Rényi divergence as

$$D_\alpha^{*,\gamma}(\rho\|\sigma) = \frac{1}{\alpha-1}\log f(\rho, (\rho, \sigma)).$$

Note that the map $\varphi : (\rho, \sigma) \mapsto (\rho, (\rho, \sigma))$ is clearly a semiring-homomorphism, which implies properties (ii) and (iii). If $T$ is a channel such that $T(\rho) = \rho'$ and $T(\sigma) \leq \sigma'$, then $(\rho, (\rho, \sigma)) \succcurlyeq (\rho', (\rho', \sigma'))$, which implies (iv) and (vi) (the equality $T(\rho) = \rho'$ is important: $\varphi$ is *not* monotone with respect to relative submajorization). Property (i) can be seen by a direct calculation, and (v) is a consequence of (i) and that the classical Rényi divergence is positive definite. □

**Remark 4.4.38.** For completeness, we give a direct proof of properties (iv) and (vi) (the remaining ones either follow easily from the definition or by a reduction to the classical case as in the proof of Proposition 4.4.37).

If $T$ is a completely positive trace-preserving map, then

$$\begin{aligned}
D_{\alpha}^{*,\gamma}(T(\rho)\|T(\sigma)) &= \frac{1}{1-\gamma} D^*_{\frac{\alpha-\gamma}{1-\gamma}}(T(\rho)\|T(\sigma)\#_\gamma T(\rho)) \\
&\geq \frac{1}{1-\gamma} D^*_{\frac{\alpha-\gamma}{1-\gamma}}(T(\rho)\|T(\sigma\#_\gamma \rho)) \\
&\geq \frac{1}{1-\gamma} D^*_{\frac{\alpha-\gamma}{1-\gamma}}(\rho\|\sigma\#_\gamma \rho) \\
&= D_{\alpha}^{*,\gamma}(\rho\|\sigma),
\end{aligned}$$

where the first inequality uses that $T(\sigma)\#_\gamma T(\rho) \geq T(\sigma\#_\gamma\rho)$ and that the sandwiched Rényi divergence is anti-monotone in its second argument, and the second inequality uses the data processing inequality for the sandwiched Rényi divergence. This proves (iv). Property (vi) is true since the matrix geometric mean is monotone in both arguments and the sandwiched Rényi divergence is anti-monotone in its second argument.

Finally, we note that when $\gamma = 0$, (4.15) agrees with the minimal Rényi divergence, and when $\alpha = 1$, the limit $\gamma \to 1$ is the Belavkin–Staszewski relative entropy [MBV22], but for $\alpha > 1$ we do not know what the limit $\gamma \to 1$ is. We leave the detailed study of these divergences for future work.

# 5 References


[AD15] Koenraad M. R. Audenaert and Nilanjana Datta. Alpha-z-relative Renyi entropies. *Journal of Mathematical Physics*, 56(2):022202, February 2015.

[ALM04] T. Ando, Chi-Kwong Li, and Roy Mathias. Geometric means. *Linear Algebra and its Applications*, 385:305–334, July 2004.

[And79] T. Ando. Concavity of certain maps on positive definite matrices and applications to Hadamard products. *Linear Algebra and its Applications*, 26:203–241, August 1979.

[ANSV08] K. M. R. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete. Asymptotic Error Rates in Quantum Hypothesis Testing. *Communications in Mathematical Physics*, 279(1):251–283, April 2008.

[AT75] W. N. Anderson and G. E. Trapp. Shorted Operators. II. *SIAM Journal on Applied Mathematics*, 28(1):60–71, 1975.

[BBH21] Mario Berta, Fernando G. S. L. Brandão, and Christoph Hirche. On Composite Quantum Hypothesis Testing. *Communications in Mathematical Physics*, 385(1):55–77, July 2021.

[BDK+05] Igor Bjelaković, Jean-Dominique Deuschel, Tyll Krüger, Ruedi Seiler, Rainer Siegmund-Schultze, and Arleta Szkoła. A Quantum Version of Sanov's Theorem. *Communications in Mathematical Physics*, 260(3):659–671, December 2005.

[Bei13] Salman Beigi. Sandwiched Rényi Divergence Satisfies Data Processing Inequality. *Journal of Mathematical Physics*, 54(12):122202, December 2013.

[BFT17] Mario Berta, Omar Fawzi, and Marco Tomamichel. On Variational Expressions for Quantum Relative Entropies. *Letters in Mathematical Physics*, 107(12):2239–2265, December 2017.

[BH06] Rajendra Bhatia and John Holbrook. Riemannian geometry and matrix geometric means. *Linear Algebra and its Applications*, 413(2):594–618, March 2006.

[Bha97] Rajendra Bhatia. *Matrix Analysis.* Number 169 in Graduate Texts in Mathematics. Springer, New York, 1997.

[BHN+15] Fernando Brandão, Michał Horodecki, Nelly Ng, Jonathan Oppenheim, and Stephanie Wehner. The second laws of quantum thermodynamics. *Proceedings of the National Academy of Sciences*, 112(11):3275–3279, March 2015.

[BHO+13] Fernando G. S. L. Brandão, Michał Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens. Resource Theory of Quantum States Out of Thermal Equilibrium. *Physical Review Letters*, 111(25):250404, December 2013.

[BMMZ22] Gergely Bunth, Gábor Maróti, Milán Mosonyi, and Zoltán Zimborás. Super-exponential distinguishability of correlated quantum states. *arXiv:2203.16511 [math-ph, physics:quant-ph]*, March 2022.

[BMP10] Dario Bini, Beatrice Meini, and Federico Poloni. An effective matrix geometric mean satisfying the Ando–Li–Mathias properties. *Mathematics of Computation*, 79(269):437–452, January 2010.

[Bou04] Nicolas Bourbaki. *Elements of Mathematics: Integration I.* Springer, Berlin, Heidelberg, 2004.

[BRS07] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Reference frames, superselection rules, and quantum information. *Reviews of Modern Physics*, 79(2):555–609, April 2007.

[BS82] V. P. Belavkin and P. Staszewski. $C^*$-algebraic generalization of relative entropy and entropy. *Annales de l'institut Henri Poincaré. Section A, Physique Théorique*, 37(1):51–58, 1982.

[BST18] Mario Berta, Volkher B. Scholz, and Marco Tomamichel. Rényi Divergences as Weighted Non-commutative Vector-Valued $L_p$-Spaces. *Annales Henri Poincaré*, 19(6):1843–1867, June 2018.

[BV21] Gergely Bunth and Péter Vrana. Asymptotic relative submajorization of multiple-state boxes. *Letters in Mathematical Physics*, 111(4):94, July 2021.

[BV23] Gergely Bunth and Péter Vrana. Equivariant Relative Submajorization. *IEEE Transactions on Information Theory*, 69(2):1057–1073, February 2023.

[CFL16] Eric A. Carlen, Rupert L. Frank, and Elliott H. Lieb. Some operator and trace function convexity theorems. *Linear Algebra and its Applications*, 490:174–185, February 2016.

[CL08] Eric A. Carlen and Elliott H. Lieb. A Minkowski Type Trace Inequality and Strong Subadditivity of Quantum Entropy II: Convexity and Concavity. *Letters in Mathematical Physics*, 83(2):107–126, February 2008.

[CM03] I. Csiszar and F. Matus. Information projections revisited. *IEEE Transactions on Information Theory*, 49(6):1474–1490, June 2003.

[Csi93] Imre Csiszár. Generalized cutoff rates and Renyi's information measures. *IEEE Transactions on Information Theory*, January 1993.

[Dat09] Nilanjana Datta. Min- and Max-Relative Entropies and a New Entanglement Monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826, June 2009.

[Eff09] Edward G. Effros. A matrix convexity approach to some celebrated quantum inequalities. *Proceedings of the National Academy of Sciences*, 106(4):1006–1008, January 2009.

[ENEG11] Ali Ebadian, Ismail Nikoufar, and Madjid Eshaghi Gordji. Perspectives of matrix convex functions. *Proceedings of the National Academy of Sciences*, 108(18):7313–7314, May 2011.

[FF20] Hamza Fawzi and Omar Fawzi. Defining quantum divergences via convex optimization. *arXiv:2007.12576 [quant-ph]*, July 2020.

[FFHT23] Muhammad Usman Farooq, Tobias Fritz, Erkka Haapasalo, and Marco Tomamichel. Asymptotic and catalytic matrix majorization, January 2023.

[FL13] Rupert L. Frank and Elliott H. Lieb. Monotonicity of a relative Rényi entropy. *Journal of Mathematical Physics*, 54(12):122201, December 2013.

[FLO23] Keiichiro Furuya, Nima Lashkari, and Shoy Ouseph. Monotonic multi-state quantum f-divergences. *Journal of Mathematical Physics*, 64(4):042203, April 2023.

[Fol99] G. B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics. Wiley, New York, 2nd ed edition, 1999.

[FOR15] Philippe Faist, Jonathan Oppenheim, and Renato Renner. Gibbs-Preserving Maps outperform Thermal Operations in the quantum regime. *New Journal of Physics*, 17(4):043003, April 2015.

[Fri23] Tobias Fritz. Abstract Vergleichsstellensätze for Preordered Semifields and Semirings I. *SIAM Journal on Applied Algebra and Geometry*, 7(2):505–547, June 2023.

[GJB+18] Gilad Gour, David Jennings, Francesco Buscemi, Runyao Duan, and Iman Marvian. Quantum majorization and a complete set of entropic conditions for quantum thermodynamics. *Nature Communications*, 9(1):5352, December 2018.

[Hay07] Masahito Hayashi. Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding. *Physical Review A*, 76(6):062301, December 2007.

[Hia01] Fumio Hiai. Concavity of Certain Matrix Trace Functions. *Taiwanese Journal of Mathematics*, 5(3):535–554, 2001.

[Hia13] Fumio Hiai. Concavity of certain matrix trace and norm functions. *Linear Algebra and its Applications*, 439(5):1568–1589, September 2013.

[Hia18] Fumio Hiai. Quantum f-divergences in von Neumann algebras. I. Standard f-divergences. *Journal of Mathematical Physics*, 59(10):102202, September 2018.

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

[Hia21] Fumio Hiai. *Quantum F-Divergences in von Neumann Algebras: Reversibility of Quantum Operations*. Mathematical Physics Studies. Springer, Singapore, 2021.

[HKM$^+$02] Masahito Hayashi, Masato Koashi, Keiji Matsumoto, Fumiaki Morikoshi, and Andreas Winter. Error exponents for entanglement concentration. *Journal of Physics A: Mathematical and General*, 36(2):527, December 2002.

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

[HM23] Fumio Hiai and Milán Mosonyi. Quantum Rényi Divergences and the Strong Converse Exponent of State Discrimination in Operator Algebras. *Annales Henri Poincaré*, 24(5):1681–1724, May 2023.

[HMH09] Fumio Hiai, Milán Mosonyi, and Masahito Hayashi. Quantum hypothesis testing with group symmetry. *Journal of Mathematical Physics*, 50(10):103304, October 2009.

[HMPB11] Fumio Hiai, Milán Mosonyi, Dénes Petz, and Cédric Bény. Quantum f-divergences and error correction. *Reviews in Mathematical Physics*, 23(07):691–747, August 2011.

[HO13] Michał Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature Communications*, 4(1):2059, June 2013.

[HP91] Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143(1):99–114, 1991.

[HP93] Fumio Hiai and Dénes Petz. The Golden-Thompson trace inequality is complemented. *Linear Algebra and its Applications*, 181:153–185, March 1993.

[HT16] Masahito Hayashi and Marco Tomamichel. Correlation detection and an operational interpretation of the Rényi mutual information. *Journal of Mathematical Physics*, 57(10):102201, October 2016.

[HT23] Christoph Hirche and Marco Tomamichel. Quantum Rényi and f-divergences from integral representations, August 2023.

[Jen17] Anna Jenčová. Preservation of a quantum Rényi relative entropy implies existence of a recovery map. *Journal of Physics A: Mathematical and Theoretical*, 50(8):085303, January 2017.

[Jen18] Anna Jenčová. Rényi Relative Entropies and Noncommutative $\mathrm{L_p}$-Spaces. *Annales Henri Poincaré*, 19(8):2513–2542, August 2018.

[Jen19] Asger Kjærulff Jensen. Asymptotic majorization of finite probability distributions. *IEEE Transactions on Information Theory*, pages 1–1, 2019.

[Jen21] Anna Jenčová. Rényi Relative Entropies and Noncommutative $\mathrm{L_p}$-Spaces II. *Annales Henri Poincaré*, 22(10):3235–3254, October 2021.

[JP06] Anna Jenčová and Dénes Petz. Sufficiency in Quantum Statistical Inference. *Communications in Mathematical Physics*, 263(1):259–276, April 2006.

[JV18] Asger Kjærulff Jensen and Péter Vrana. The asymptotic spectrum of LOCC transformations. *arXiv:1807.05130 [math-ph, physics:quant-ph]*, July 2018.

[JWZ+00] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th. Beth. Thermodynamic Cost of Reliability and Low Temperatures: Tightening Landauer's Principle and the Second Law. *International Journal of Theoretical Physics*, 39(12):2717–2753, December 2000.

[KA80] Fumio Kubo and Tsuyoshi Ando. Means of positive linear operators. *Mathematische Annalen*, 246(3):205–224, October 1980.

[Kli07] Matthew Klimesh. Inequalities that Collectively Completely Characterize the Catalytic Majorization Relation. *arXiv:0709.3680 [quant-ph]*, September 2007.

[Kos82] Hideki Kosaki. Interpolation theory and the Wigner-Yanase-Dyson-Lieb concavity. *Communications in Mathematical Physics*, 87(3):315–329, December 1982.

[LJR15] Matteo Lostaglio, David Jennings, and Terry Rudolph. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. *Nature Communications*, 6(1):6383, March 2015.

[LL11] Jimmie Lawson and Yongdo Lim. Monotonic properties of the least squares mean. *Mathematische Annalen*, 351(2):267–279, October 2011.

[LP12] Yongdo Lim and Miklós Pálfia. Matrix power means and the Karcher mean. *Journal of Functional Analysis*, 262(4):1498–1514, February 2012.

[LR73] Elliott H. Lieb and Mary Beth Ruskai. A Fundamental Property of Quantum-Mechanical Entropy. *Physical Review Letters*, 30(10):434–436, March 1973.

[LT15] Mingyan Simon Lin and Marco Tomamichel. Investigating Properties of a Family of Quantum Renyi Divergences. *Quantum Information Processing*, 14(4):1501–1512, April 2015.

[LY22a] Ke Li and Yongsheng Yao. Operational Interpretation of the Sandwiched Rényi Divergence of Order 1/2 to 1 as Strong Converse Exponents, December 2022.

[LY22b] Ke Li and Yongsheng Yao. Reliability Function of Quantum Information Decoupling via the Sandwiched Rényi Divergence, April 2022.

[LY22c] Ke Li and Yongsheng Yao. Reliable Simulation of Quantum Channels, November 2022.

[LY23] Ke Li and Yongsheng Yao. Strong Converse Exponent for Entanglement-Assisted Communication, June 2023.

[LYH23] Ke Li, Yongsheng Yao, and Masahito Hayashi. Tight Exponential Analysis for Smoothing the Max-Relative Entropy and for Quantum Privacy Amplification. *IEEE Transactions on Information Theory*, 69(3):1680–1694, March 2023.

[Mat18] Keiji Matsumoto. A New Quantum Version of f-Divergence. In Masanao Ozawa, Jeremy Butterfield, Hans Halvorson, Miklós Rédei, Yuichiro Kitajima, and Francesco Buscemi, editors, *Reality and Measurement in Algebraic Quantum Theory*, Springer Proceedings in Mathematics & Statistics, pages 229–273, Singapore, 2018. Springer.

[MBV22] Milán Mosonyi, Gergely Bunth, and Péter Vrana. Geometric relative entropies and barycentric Rényi divergences, July 2022.

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

[MH23a] Milán Mosonyi and Fumio Hiai. Some continuity properties of quantum Rényi divergences, June 2023.

[MH23b] Milán Mosonyi and Fumio Hiai. Test-Measured Rényi Divergences. *IEEE Transactions on Information Theory*, 69(2):1074–1092, February 2023.

[MO15] Milán Mosonyi and Tomohiro Ogawa. Quantum Hypothesis Testing and the Operational Interpretation of the Quantum Rényi Relative Entropies. *Communications in Mathematical Physics*, 334(3):1617–1648, March 2015.

[MO17] Milán Mosonyi and Tomohiro Ogawa. Strong Converse Exponent for Classical-Quantum Channel Coding. *Communications in Mathematical Physics*, 355(1):373–426, October 2017.

[MO21] Milán Mosonyi and Tomohiro Ogawa. Divergence Radii and the Strong Converse Exponent of Classical-Quantum Channel Coding With Constant Compositions. *IEEE Transactions on Information Theory*, 67(3):1668–1698, March 2021.

[Moa05] Maher Moakher. A Differential Geometric Approach to the Geometric Mean of Symmetric Positive-Definite Matrices. *SIAM Journal on Matrix Analysis and Applications*, 26(3):735–747, January 2005.

[Mos15] Milan Mosonyi. Coding Theorems for Compound Problems via Quantum Rényi Divergences. *IEEE Transactions on Information Theory*, 61(6):2997–3012, June 2015.

[Mos23] Milán Mosonyi. The Strong Converse Exponent of Discriminating Infinite-Dimensional Quantum States. *Communications in Mathematical Physics*, 400(1):83–132, May 2023.

[MPST21] Xiaosheng Mu, Luciano Pomatto, Philipp Strack, and Omer Tamuz. From Blackwell Dominance in Large Samples to Rényi Divergences and Back Again. *Econometrica*, 89(1):475–506, 2021.

[MR17] Alexander Müller-Hermes and David Reeb. Monotonicity of the Quantum Relative Entropy Under Positive Maps. *Annales Henri Poincaré*, 18(5):1777–1788, May 2017.

[MSW22] Milán Mosonyi, Zsombor Szilágyi, and Mihály Weiner. On the Error Exponents of Binary State Discrimination With Composite Hypotheses. *IEEE Transactions on Information Theory*, 68(2):1032–1067, February 2022.

[Nag06] Hiroshi Nagaoka. The Converse Part of The Theorem for Quantum Hoeffding Bound, November 2006.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge ; New York, 10th anniversary ed edition, 2010.

[ON05] Tomohiro Ogawa and Hiroshi Nagaoka. Strong Converse and Stein's Lemma in the Quantum Hypothesis Testing. *arXiv:quant-ph/9906090*, pages 28–42, February 2005.

[OP93] M. Ohya and Denes Petz. *Quantum Entropy and Its Use*. Theoretical and Mathematical Physics. Springer-Verlag, Berlin Heidelberg, 1993.

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

[Pet86a] Dénes Petz. Properties of the relative entropy of states of von Neumann algebras. *Acta Mathematica Hungarica*, 47(1-2), 1986.

[Pet86b] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports on Mathematical Physics*, 23(1):57–65, February 1986.

[PT05] Dénes Petz and Róbert Temesi. Means of Positive Numbers and Matrices. *SIAM Journal on Matrix Analysis and Applications*, 27(3):712–720, January 2005.

[PVW22] Christopher Perry, Péter Vrana, and Albert H. Werner. The Semiring of Dichotomies and Asymptotic Relative Submajorization. *IEEE Transactions on Information Theory*, 68(1):311–321, January 2022.

[PW75] W. Pusz and S. L. Woronowicz. Functional calculus for sesquilinear forms and the purification map. *Reports on Mathematical Physics*, 8(2):159–170, October 1975.

[Rén61] Alfréd Rényi. On Measures of Entropy and Information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California, 1961.

[Ren16] Joseph M. Renes. Relative submajorization and its use in quantum resource theories. *Journal of Mathematical Physics*, 57(12):122202, December 2016.

[Sim19] Barry Simon. *Loewner's Theorem on Monotone Matrix Functions*. Grundlehren Der Mathematischen Wissenschaften. Springer International Publishing, 2019.

[Str11] Helmut Strasser. Mathematical Theory of Statistics: Statistical Experiments and Asymptotic Decision Theory. In *Mathematical Theory of Statistics*. De Gruyter, April 2011.

[Tho63] A. C. Thompson. On Certain Contraction Mappings in a Partially Ordered Vector Space. *Proceedings of the American Mathematical Society*, 14(3):438–443, 1963.

[Tom16] Marco Tomamichel. *Quantum Information Processing with Finite Resources*, volume 5 of *SpringerBriefs in Mathematical Physics*. Springer International Publishing, Cham, 2016.

[Tro12] Joel Tropp. From joint convexity of quantum relative entropy to a concavity theorem of Lieb. *Proceedings of the American Mathematical Society*, 140(5):1757–1760, May 2012.

[Tur07] S Turgut. Catalytic transformations for bipartite pure states. *Journal of Physics A: Mathematical and Theoretical*, 40(40):12185–12212, October 2007.

[vH14] Tim van Erven and Peter Harremos. Rényi Divergence and Kullback-Leibler Divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, July 2014.

[Vra22] Péter Vrana. A Generalization of Strassen's Theorem on Preordered Semirings. *Order*, 39(2):209–228, July 2022.

[WWY14] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy. *Communications in Mathematical Physics*, 331(2):593–622, October 2014.

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