

Quantum Fisher information matrices from Rényi relative entropies

Mark M. Wilde

*School of Electrical and Computer Engineering,
Cornell University, Ithaca, New York 14850, USA*

October 7, 2025

Abstract

Quantum generalizations of the Fisher information are important in quantum information science, with applications in high energy and condensed matter physics and in quantum estimation theory, machine learning, and optimization. One can derive a quantum generalization of the Fisher information matrix in a natural way as the Hessian matrix arising in a Taylor expansion of a smooth divergence. Such an approach is appealing and intuitive for quantum information theorists, given the ubiquity of divergences in quantum information theory. In contrast to the classical case, there is not a unique quantum generalization of the Fisher information matrix, similar to how there is not a unique quantum generalization of the relative entropy or the Rényi relative entropy. In this paper, I derive information matrices arising from the log-Euclidean, α - z , and geometric Rényi relative entropies, with the main technical tool for doing so being the method of divided differences for calculating matrix derivatives. Interestingly, for all non-negative values of the Rényi parameter α , the log-Euclidean Rényi relative entropy leads to the Kubo–Mori information matrix, and the geometric Rényi relative entropy leads to the right-logarithmic derivative Fisher information matrix. Thus, the resulting information matrices obey the data-processing inequality for all non-negative values of the Rényi parameter α even though the original quantities do not. Additionally, I derive and establish basic properties of α - z information matrices resulting from the α - z Rényi relative entropies, while leaving it open to determine all values of α and z for which the α - z information matrices obey the data-processing inequality. For parameterized thermal states and time-evolved states, I establish formulas for their α - z information matrices and hybrid quantum–classical algorithms for estimating them, with applications in quantum Boltzmann machine learning.

This paper is dedicated to Professor Fumio Hiai on the occasion of his forthcoming 80th birthday. His seminal insights in matrix analysis [Hia10, HP14], mathematical physics [Hia18, Hia19, Hia21], and quantum information theory [HP91, HMPB11, MH11] have left indelible impacts on these fields and have had a profound influence on younger generations of quantum information scientists.

Contents

Contents	2
1 Introduction	3
1.1 Background	3
1.2 Summary of results	4
1.3 Paper organization	7
2 Preliminaries	7
2.1 Classical Fisher information	8
2.2 Quantum generalizations of Fisher information	9
3 Kubo–Mori information matrix from log-Euclidean Rényi relative entropies	11
4 RLD Fisher information matrix from geometric relative entropies	16
4.1 RLD Fisher information matrix from geometric Rényi relative entropies . . .	16
4.2 RLD Fisher information matrix from Belavkin–Staszewski relative entropy .	22
5 α-z Information matrices from α-z Rényi relative entropies	23
6 Special cases of α-z information matrices	34
6.1 Kubo–Mori information matrix	34
6.2 Petz–Rényi information matrices	36
6.3 Sandwiched Rényi information matrices	39
7 Orderings and relations between α-z information matrices	41
7.1 Ordering of general information matrices	42
7.2 Ordering of Petz–Rényi information matrices	45
7.3 Ordering of sandwiched Rényi information matrices	47
8 α-z Information matrices of parameterized thermal states	49
8.1 General formula for information matrices of parameterized thermal states . .	51
8.2 α - z High-peak tent probability densities	57
9 α-z Information matrices of time-evolved states	69
9.1 General formula for information matrices of time-evolved states	71
10 Information matrices of classical–quantum states	75
11 Conclusion	78
References	78

A Proofs of Equations (2.13) and (2.14)	86
B Review: Matrix derivatives from divided differences	89
B.1 First derivative	89
B.2 Examples	95
C Proof of Equation (5.55)	101
D Proof of Equation (5.6)	103
E Proof of Equation (5.83)	104

1 Introduction

1.1 Background

Fisher information is a core concept in information science [CT05, Section 11.10]. It arises in a number of contexts, perhaps most prominently in estimation theory [KK11] but also in machine learning and optimization in the form of the natural gradient descent method [Ama98]. Additionally, it is the centerpiece of the field of information geometry [Ama16, Nie22], because the Fisher information matrix defines a Riemannian metric on the parameter space underlying a parameterized family of probability distributions. In fact, up to a multiplicative constant, it is the unique Riemannian metric for which the data-processing inequality holds [Čen82]. The fact that it goes by the moniker “information” suggests that every information theorist should know something about it.

The Fisher information concept has been generalized to the quantum case [Hel67, Hel69, BC94, Hol11] with the original goal of developing a theory of estimation relevant for quantum information science. Since then, it has been broadly applied in a variety of contexts, including high energy physics [MNS⁺15, LVR16, BES18, MH18], condensed matter [CVZ07, ZGC07, Gu10, HHTZ16, CVS20], quantum machine learning and optimization [SIKC20, SMP⁺22, KB22, SKP24, PW24, MPW25a, MPW25b], and quantum resource theories [TNR21, Mar22, YMST25]. Signifying its fundamental role in quantum information science, several reviews of quantum Fisher information have appeared to date [BZ06, Hay17, LYLW19, SK20, JK20, KW21, Mey21, SMP⁺22, SASS24].

One of the simplest ways for an information theorist to connect with the Fisher information concept is through smooth divergences that measure the distinguishability of nearby probability distributions. More precisely, the Fisher information matrix is equal to the Hessian matrix in a Taylor expansion of a smooth divergence evaluated at nearby probability distributions. Remarkably, the Čencov (Chentsov) theorem states that, up to a constant prefactor, for every smooth divergence, the Fisher information matrix is the unique matrix that arises in this way [Čen82]. Thus, in the classical case, we can speak of a single notion of Fisher information matrix.

Similar to the classical case, a quantum generalization of the Fisher information matrix arises through smooth divergences that measure the distinguishability of nearby quantum states. As before, a quantum generalization of the Fisher information matrix is equal to the Hessian matrix in a Taylor expansion of a smooth divergence evaluated at nearby quantum states. However, in distinction to the classical case, there is not a unique quantum Fisher information matrix, and instead there are an infinite number of possibilities, similar to how there is not a unique quantum generalization of the relative entropy or the Rényi relative entropy. There is a characterization of information matrices that arise in the aforementioned way [MC91, Pet96] (see also [BZ06, Theorem 14.1]), which is recalled as Theorem 1 below.

1.2 Summary of results

In this paper, I explore quantum generalizations of the Fisher information matrix that arise from the log-Euclidean, geometric, and α - z Rényi relative entropies (defined in (3.1), (4.1), and (5.1), respectively). These are called the log-Euclidean information matrix, the geometric information matrix, and the α - z information matrix, and are defined in (3.3), (4.3), and (5.4), respectively.

The main technical results of this paper are as follows:

- For all $\alpha \in (0, 1) \cup (1, \infty)$, the log-Euclidean information matrix is equal to the Kubo–Mori information matrix (Theorem 3), the latter defined in (3.4), being the information matrix that arises from the standard (Umegaki) relative entropy.
- For all $\alpha \in (0, 1) \cup (1, \infty)$, the geometric information matrix is equal to the right logarithmic derivative (RLD) information matrix (Theorem 5), the latter defined in (4.4) and well studied in the literature on quantum estimation theory (see, e.g., [SK20, Section V-E]). I also prove that the same is true when considering the information matrix that arises from the Belavkin–Staszewski relative entropy (Theorem 7). These derivations generalize those from the single-parameter case [Mat10, Mat13, Mat18, KW21].
- For all $\alpha \in (0, 1) \cup (1, \infty)$ and $z > 0$, I derive a formula for the α - z information matrix (Theorem 10). This formula, applicable for the multiparameter case, generalizes the formula reported in [MH18, Eq. (2.12)] for the single-parameter case and that in [CDL⁺18, Eqs. (132)–(133)] for the Bloch-sphere qubit case. Additionally, the formula demonstrates that, for all $z > 0$, the α - z information matrix converges to the Kubo–Mori information matrix in the limit as $\alpha \rightarrow 1$. It also does so for all $\alpha \in (0, 1) \cup (1, \infty)$ in the limit as $z \rightarrow \infty$.
- The Petz– and sandwiched Rényi information matrices arise as special cases of the α - z information matrix when $z = 1$ and $z = \alpha$ (Corollary 17 and Corollary 21), respectively, which correspond to the cases in which the underlying Rényi relative entropy is set to be the Petz– and sandwiched Rényi relative entropies, respectively. I establish integral representations for these information matrices when $\alpha \in (0, 1)$ (Proposition 18 and Proposition 22) and derive simple formulas for them for $\alpha = 2$.

- I establish ordering relations for the Petz– and sandwiched Rényi information matrices. Namely, as a function of the Rényi parameter α , the Petz–Rényi information matrix is monotone decreasing on $\alpha \in (0, \frac{1}{2}]$ and monotone increasing on $\alpha \in [\frac{1}{2}, \infty)$ (Theorem 28), and the sandwiched Rényi information matrix is monotone increasing on $\alpha \in (0, \infty)$ (Theorem 30).
- For all $\alpha \in (0, 1)$ and $z > 0$, I derive a formula for the α - z information matrix when the underlying family of states consists of parameterized thermal states (Theorem 32). Specifically, for $\rho(\theta) := \frac{e^{-H(\theta)}}{\text{Tr}[e^{-H(\theta)}]}$ and $H(\theta) := \sum_{j=1}^L \theta_j H_j$, with $\theta_j \in \mathbb{R}$ and H_j Hermitian for all $j \in \{1, \dots, L\}$, the following formula holds for the α - z information matrix:

$$[I_{\alpha,z}(\theta)]_{i,j} = \frac{1}{2} \langle \{ \Phi_{q_{\alpha,z},\theta}(H_i), H_j \} \rangle_{\rho(\theta)} - \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)}, \quad (1.1)$$

where the quantum channel $\Phi_{q_{\alpha,z},\theta}$ is given by

$$\Phi_{q_{\alpha,z},\theta}(X) := \int_{-\infty}^{\infty} dt q_{\alpha,z}(t) e^{-itH(\theta)} X e^{itH(\theta)}, \quad (1.2)$$

and $q_{\alpha,z} := p * p_{\alpha,z}$ is a probability density function equal to the convolution of the probability density functions p and $p_{\alpha,z}$, each defined on $t \in \mathbb{R}$ as

$$p(t) := \frac{2}{\pi} \ln \left| \coth \left(\frac{\pi t}{2} \right) \right|, \quad p_{\alpha,z}(t) := \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi z t)} \right)^2 \right). \quad (1.3)$$

- For all $\alpha \in (0, 1)$ and $z > 0$, I derive a formula for the α - z information matrix when the underlying parameterized family consists of time-evolved states (Theorem 35). Specifically, for $\sigma(\phi) := e^{-iH(\phi)} \rho e^{iH(\phi)}$, where $H(\phi) := \sum_{j=1}^L \phi_j H_j$, with $\phi_j \in \mathbb{R}$ and H_j Hermitian for all $j \in \{1, \dots, L\}$, and $\rho := \frac{e^{-G}}{\text{Tr}[e^{-G}]}$ with G Hermitian, the following formula holds for the α - z information matrix:

$$[I_{\alpha,z}(\phi)]_{i,j} = \langle [\Phi_{p_{\alpha,z}}(\Psi_\phi(H_i)), [G, \Psi_\phi(H_j)]] \rangle_\rho, \quad (1.4)$$

where the quantum channels Ψ_ϕ and $\Phi_{p_{\alpha,z}}$ are defined as

$$\Psi_\phi(X) := \int_0^1 dt e^{iH(\phi)t} X e^{-iH(\phi)t}, \quad (1.5)$$

$$\Phi_{p_{\alpha,z}}(X) := \int_{-\infty}^{\infty} dt p_{\alpha,z}(t) e^{-iGt} X e^{iGt}, \quad (1.6)$$

and $p_{\alpha,z}$ is the probability density function defined in (1.3).

- I evaluate the α - z , Kubo–Mori, and RLD information matrices for parameterized families of classical-quantum states, showing that they decompose as a sum of a classical part and a quantum part (Theorem 40), thus generalizing the previous finding from [KW21, Proposition 7] for the single-parameter case. A direct consequence of this and the data-processing inequality is that these information matrices are convex (Corollary 41).

The first two results mentioned above are reminiscent of the classical Čencov theorem. That is, for all $\alpha \in (0, 1) \cup (1, \infty)$, the log-Euclidean information matrices collapse to the same information matrix, namely, the Kubo–Mori one, and for all $\alpha \in (0, 1) \cup (1, \infty)$, the geometric information matrices collapse to the same information matrix, namely, the RLD Fisher information matrix. Thus, the underlying Rényi relative entropies belong to two distinct families in this sense. Additionally, in spite of the fact that the log-Euclidean Rényi relative entropy does not obey the data-processing inequality for $\alpha > 1$, the resulting log-Euclidean information matrix does. Similarly, even though the geometric Rényi relative entropy does not obey the data-processing inequality for $\alpha > 2$, the resulting RLD Fisher information matrix does.

The result in (1.1) is a broad generalization of the findings reported in [PW24, Theorems 1 and 2] and [MPW25a, Theorem 14], such that these earlier findings are now special cases of Theorem 32. Specifically,

- [PW24, Theorem 1] follows from Theorem 32 with $\alpha = z = \frac{1}{2}$,
- [PW24, Theorem 2] follows from Theorem 32 with $\alpha \rightarrow 1$ and $z > 0$ (or $\alpha > 0$ and $z \rightarrow \infty$),
- [MPW25a, Theorem 14] follows from Theorem 32 with $\alpha = \frac{1}{2}$ and $z = 1$, and I note here that Theorem 32 provides a simpler formula than that given in [MPW25a, Theorem 14].

Given that parameterized thermal states are also known as quantum Boltzmann machines [AAR⁺18, KW17, BRGBPO17], the result stated in (1.1) is applicable to quantum Boltzmann machine learning. In particular, if one wishes to perform quantum optimization using natural gradient descent with quantum Boltzmann machines and the geometry induced by the α - z Rényi relative entropy, then the formula in (1.1) is applicable. To estimate the elements of the α - z information matrix by means of a hybrid quantum-classical algorithm, one can employ a procedure similar to that outlined in [MPW25a, Figure 4(a)], with the underlying probability density replaced by $q_{\alpha,z}$, as defined in (8.72). Given the connections between the α - z information matrix and high energy physics, as considered in [MH18], the formula in (1.1) and the corresponding hybrid quantum-classical algorithm could find applications there as well.

Similar to what was stated above for parameterized thermal states, the result in (1.4), regarding time-evolved states, is also a broad generalization of the findings reported in [MPW25a, Theorems 11, 15, and 19], such that these earlier findings are now special cases of Theorem 35. Specifically,

- [MPW25a, Theorem 11] follows from Theorem 35 with $\alpha = z = \frac{1}{2}$.
- [MPW25a, Theorem 15] follows from Theorem 35 with $\alpha = \frac{1}{2}$ and $z = 1$, and I note here that Theorem 35 provides a simpler formula than that given in [MPW25a, Theorem 15].

- [MPW25a, Theorem 19] follows from Theorem 35 with $\alpha \rightarrow 1$ and $z > 0$ (or $\alpha > 0$ and $z \rightarrow \infty$).

Given that time-evolved states are also known as quantum evolution machines [MPW25a], the result stated in (1.4) is also applicable to quantum machine learning. In particular, if one wishes to perform quantum optimization using natural gradient descent with quantum evolution machines and the geometry induced by the α - z Rényi relative entropy, then the formula in (1.4) is applicable. To estimate the elements of the α - z information matrix by means of a hybrid quantum-classical algorithm, one can employ a procedure similar to that outlined in [MPW25a, Figure 2(b)], with the underlying probability density p therein replaced by $p_{\alpha,z}$, as defined in (1.3).

The results in (1.1) and (1.4) are also reminiscent of the classical Čencov theorem. Indeed, when considering the α - z information matrices of parameterized thermal states and time-evolved states for $\alpha \in (0, 1)$ and $z > 0$, the only α - z dependence that these expressions have is with respect to the underlying probability density, which is $q_{\alpha,z}$ in (1.1) and $p_{\alpha,z}$ in (1.4). Thus, these results represent a near unification of the α - z information matrices for these classes of states.

1.3 Paper organization

This paper is structured as follows. Section 2 reviews preliminary material needed to understand the rest of the paper, including classical Fisher information and quantum generalizations thereof. Section 3 presents the first result stated above, that the log-Euclidean information matrix is equal to the Kubo–Mori information matrix (Theorem 3). Section 4 presents the second result stated above, that the geometric information matrix is equal to the RLD Fisher information matrix (Theorem 5). Section 5 presents the formula for the α - z information matrix (Theorem 10), Section 6 presents special cases of the α - z information matrices, and Section 7 presents the ordering relations for the Petz– and sandwiched Rényi information matrices. Section 8 presents the formula for the α - z information matrix of parameterized thermal states (Theorem 32), and Section 9 presents the formula for the α - z information matrix of time-evolved states (Theorem 35). Section 10 provides the decomposition formula for information matrices evaluated on classical-quantum states. Finally, Section 11 concludes with a summary of the results and points to directions for future work. Appendix B provides a detailed review of matrix derivatives.

2 Preliminaries

Throughout the paper, I use the notation

$$\mathbb{R}_+ := (0, \infty). \quad (2.1)$$

The rest of this section provides an overview of classical Fisher information (Section 2.1) and quantum generalizations of Fisher information (Section 2.2), from the perspective of smooth divergences.

2.1 Classical Fisher information

One of the simplest ways for an information theorist to connect with the Fisher information concept is through smooth divergences that measure the distinguishability of nearby probability distributions. Examples of smooth divergences include the relative entropy and the Rényi relative entropies, respectively defined for $\alpha \in (0, 1) \cup (1, \infty)$ and probability distributions p and q over a finite alphabet $\mathcal{X}$ as follows:

$$D(p\|q) := \sum_{x \in \mathcal{X}} p(x) \ln \left(\frac{p(x)}{q(x)} \right), \quad (2.2)$$

$$D_\alpha(p\|q) := \frac{1}{\alpha - 1} \ln \sum_{x \in \mathcal{X}} p(x)^\alpha q(x)^{1-\alpha}. \quad (2.3)$$

To see this in more detail, suppose that $(p_\theta)_{\theta \in \Theta}$ is a second-order differentiable parameterized family of probability distributions over a finite alphabet $\mathcal{X}$, where $\Theta \subseteq \mathbb{R}^L$ is open and $L \in \mathbb{N}$, and suppose that $\mathbf{D}$ is a smooth divergence. By the latter, I mean that

1. $\mathbf{D}$ is second-order differentiable, so that $\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(p_\theta \| p_{\theta+\varepsilon})$ exists for all $\theta, \varepsilon \in \Theta$.
2. $\mathbf{D}$ obeys the data-processing inequality, so that the following inequality holds for every classical channel $N \equiv (N(y|x))_{x \in \mathcal{X}, y \in \mathcal{Y}}$ and every pair (p, q) of probability distributions:

$$\mathbf{D}(p\|q) \geq \mathbf{D}(N(p)\|N(q)), \quad (2.4)$$

where $N(p)$ denotes the probability distribution $(\sum_{x \in \mathcal{X}} N(y|x)p(x))_{y \in \mathcal{Y}}$, with a similar meaning for $N(q)$.

3. $\mathbf{D}$ is faithful, so that the following holds for probability distributions p and q :

$$\mathbf{D}(p\|q) = 0 \iff p = q. \quad (2.5)$$

Taken together, data processing and faithfulness imply that

$$\mathbf{D}(p\|q) \geq 0 \quad (2.6)$$

for every pair (p, q) of probability distributions, so that the minimum value of $\mathbf{D}$ is equal to zero. Under all three assumptions, the following Taylor expansion holds for $\varepsilon \in \Theta$ such that $\|\varepsilon\| \ll 1$:

$$\mathbf{D}(p_\theta \| p_{\theta+\varepsilon}) = \mathbf{D}(p_\theta \| p_\theta) + \varepsilon^T (\nabla \mathbf{D})(\theta) + \frac{1}{2} \varepsilon^T I_{\mathbf{D}}(\theta) \varepsilon + o(\|\varepsilon\|^2) \quad (2.7)$$

$$= \frac{1}{2} \varepsilon^T I_{\mathbf{D}}(\theta) \varepsilon + o(\|\varepsilon\|^2), \quad (2.8)$$

where $(\nabla \mathbf{D})(\theta)$ denotes the gradient vector and $I_{\mathbf{D}}(\theta)$ the Hessian matrix, defined in terms of their elements as follows:

$$[(\nabla \mathbf{D})(\theta)]_i := \left. \frac{\partial}{\partial \varepsilon_i} \mathbf{D}(p_\theta \| p_{\theta+\varepsilon}) \right|_{\varepsilon=0}, \quad (2.9)$$

$$[I_{\mathbf{D}}(\theta)]_{i,j} := \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(p_\theta \| p_{\theta+\varepsilon}) \right|_{\varepsilon=0}. \quad (2.10)$$

The equality in (2.7) follows from the smoothness assumption, and the equality in (2.8) follows from all three assumptions. Indeed, $\mathbf{D}(p_\theta \| p_\theta) = 0$ by faithfulness and $(\nabla \mathbf{D})(\theta) = 0$ because we have expanded the function $\varepsilon \mapsto \mathbf{D}(p_\theta \| p_{\theta+\varepsilon})$ about a critical point $\varepsilon = 0$ (as previously stated, $\mathbf{D}(p_\theta \| p_{\theta+\varepsilon})$ takes its minimum value at $\varepsilon = 0$). The Hessian matrix $I_{\mathbf{D}}(\theta)$ is also known as “information matrix” in this context. From the equality in (2.8) and assumed properties of the smooth divergence $\mathbf{D}$, one can deduce that $I_{\mathbf{D}}(\theta)$ is positive semi-definite and obeys the data-processing inequality (as a matrix inequality) for all $\theta \in \Theta$ (see, e.g., [MPW25a, Proposition 4]).

Remarkably, the Čencov theorem states that, up to a constant prefactor, the Fisher information matrix is the unique matrix such that (2.8) holds [Čen82]. That is, for every smooth divergence $\mathbf{D}$, there exists a constant $\kappa > 0$ such that the following equality holds:

$$\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(p_\theta \| p_{\theta+\varepsilon}) \right|_{\varepsilon=0} = \kappa [I_F(\theta)]_{i,j}, \quad (2.11)$$

where the Fisher information matrix $I_F(\theta)$ is defined in terms of its elements as

$$[I_F(\theta)]_{i,j} := \sum_{x \in \mathcal{X}} \frac{1}{p(x)} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right). \quad (2.12)$$

As a special case, it is known that the following equality holds for all $\alpha \in (0, 1) \cup (1, \infty)$ when choosing the smooth divergence $\mathbf{D}$ to be the Rényi relative entropy D_α [vEH14, Eq. (50)]:

$$\frac{1}{\alpha} \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha(p_\theta \| p_{\theta+\varepsilon}) \right|_{\varepsilon=0} = [I_F(\theta)]_{i,j}. \quad (2.13)$$

Additionally, the following equality holds [vEH14, Eq. (49)]:

$$\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D(p_\theta \| p_{\theta+\varepsilon}) \right|_{\varepsilon=0} = [I_F(\theta)]_{i,j}, \quad (2.14)$$

which is consistent with the well known fact that $\lim_{\alpha \rightarrow 1} D_\alpha(p \| q) = D(p \| q)$ for every pair (p, q) of probability distributions. See Appendix A for brief proofs of (2.13) and (2.14), which are helpful guidelines for how analogous proofs will proceed for the more complicated case of quantum generalizations of Fisher information.

2.2 Quantum generalizations of Fisher information

Several of the developments above can be generalized to the quantum case, with the main exception being that there is no longer a unique quantum generalization of Fisher information, similar to how there are not unique generalizations of relative entropy and Rényi relative

entropy in quantum information theory, and instead there are several sensible generalizations (see, e.g., [KW20, Chapter 7]).

Let us briefly review quantum generalizations of the concepts from Section 2.1. Now suppose that $(\rho(\theta))_{\theta \in \Theta}$ is a second-order differentiable parameterized family of density operators acting on finite-dimensional Hilbert space, where $\Theta \subseteq \mathbb{R}^L$ is open and $L \in \mathbb{N}$, and suppose that $\mathbf{D}$ is a smooth quantum divergence. By the latter, I mean that

1. $\mathbf{D}$ is second-order differentiable, so that $\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(\rho(\theta) \| \rho(\theta + \varepsilon))$ exists for all $\theta, \varepsilon \in \Theta$.
2. $\mathbf{D}$ obeys the data-processing inequality, so that the following inequality holds for every quantum channel $\mathcal{N}$ (a completely positive, trace-preserving superoperator) and every pair (ρ, σ) of states:

$$\mathbf{D}(\rho \| \sigma) \geq \mathbf{D}(\mathcal{N}(\rho) \| \mathcal{N}(\sigma)). \quad (2.15)$$

3. $\mathbf{D}$ is faithful, so that the following holds for states ρ and σ :

$$\mathbf{D}(\rho \| \sigma) = 0 \iff \rho = \sigma. \quad (2.16)$$

As in the classical case, under all three assumptions, the following Taylor expansion holds for $\varepsilon \in \Theta$ such that $\|\varepsilon\| \ll 1$:

$$\mathbf{D}(\rho(\theta) \| \rho(\theta + \varepsilon)) = \frac{1}{2} \varepsilon^T I_{\mathbf{D}}(\theta) \varepsilon + o(\|\varepsilon\|^2) \quad (2.17)$$

where $I_{\mathbf{D}}(\theta)$ denotes the Hessian matrix, defined in terms of its elements as follows:

$$[I_{\mathbf{D}}(\theta)]_{i,j} := \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(\rho(\theta) \| \rho(\theta + \varepsilon)) \right|_{\varepsilon=0}. \quad (2.18)$$

The reasoning for (2.17) is the same as that given for (2.8) (see, e.g., [MPW25a, Eq. (36)]). Furthermore, the information matrix $I_{\mathbf{D}}(\theta)$ is positive semi-definite and obeys the data-processing inequality (see, e.g., [MPW25a, Proposition 4]). Additionally, additivity of the information matrix $I_{\mathbf{D}}(\theta)$ with respect to tensor-product families of states follows if the underlying smooth divergence is additive. Indeed, for a tensor-product family $(\rho(\theta) \otimes \sigma(\theta))_{\theta \in \Theta}$,

$$[I_{\mathbf{D}}(\theta; (\rho(\theta) \otimes \sigma(\theta))_{\theta \in \Theta})]_{i,j} := \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(\rho(\theta) \otimes \sigma(\theta) \| \rho(\theta + \varepsilon) \otimes \sigma(\theta + \varepsilon)) \right|_{\varepsilon=0} \quad (2.19)$$

$$= \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} [\mathbf{D}(\rho(\theta) \| \rho(\theta + \varepsilon)) + \mathbf{D}(\sigma(\theta) \| \sigma(\theta + \varepsilon))] \right|_{\varepsilon=0} \quad (2.20)$$

$$= \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(\rho(\theta) \| \rho(\theta + \varepsilon)) \right|_{\varepsilon=0} + \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(\sigma(\theta) \| \sigma(\theta + \varepsilon)) \right|_{\varepsilon=0} \quad (2.21)$$

$$= [I_{\mathbf{D}}(\theta; (\rho(\theta))_{\theta \in \Theta})]_{i,j} + [I_{\mathbf{D}}(\theta; (\sigma(\theta))_{\theta \in \Theta})]_{i,j}, \quad (2.22)$$

where the second equality follows from the assumption that $\mathbf{D}$ is additive on tensor-product states.

As stated previously, the main distinction between the classical and quantum cases is that there is no longer a unique quantum generalization of Fisher information. One of the main reasons for this is that a state $\rho(\theta)$ and each derivative $\frac{\partial}{\partial \theta_i} \rho(\theta)$ need not commute. Regardless, there is still a remarkable characterization due to [MC91, Pet96] (see also [BZ06, Theorem 14.1]), which I now recall.

Theorem 1 ([MC91, Pet96]). *Let $(\rho(\theta))_{\theta \in \Theta}$ be a second-order differentiable parameterized family of positive definite states, where $\Theta \subseteq \mathbb{R}^L$ is open and $L \in \mathbb{N}$. For every smooth quantum divergence $\mathbf{D}$, there exists a function $\zeta: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ satisfying properties 1–4 below, such that the following equality holds:*

$$\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \mathbf{D}(\rho(\theta) \| \rho(\theta + \varepsilon)) \right|_{\varepsilon=0} = [I_\zeta(\theta)]_{i,j}, \quad (2.23)$$

where the quantum information matrix $I_\zeta(\theta)$ is defined in terms of its elements as

$$[I_\zeta(\theta)]_{i,j} := \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr}[\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))], \quad (2.24)$$

$\partial_i \equiv \frac{\partial}{\partial \theta_i}$, $\rho(\theta)$ has a spectral decomposition as $\rho(\theta) = \sum_k \lambda_k \Pi_k$, (which suppresses the dependence of each eigenvalue λ_k and each eigenprojection Π_k on θ), and the function ζ satisfies the following:

1. There exists a constant $\kappa > 0$ such that $\zeta(x, x) = \frac{\kappa}{x}$ for all $x > 0$.
2. ζ is symmetric, i.e., $\zeta(x, y) = \zeta(y, x)$ for all $x, y > 0$.
3. ζ satisfies $\zeta(sx, sy) = s^{-1} \zeta(x, y)$ for all $s, x, y > 0$.
4. The function $t \mapsto f(t) := \frac{1}{\zeta(t, 1)}$ is operator monotone on $t \in (0, \infty)$.

3 Kubo–Mori information matrix from log-Euclidean Rényi relative entropies

The main result of this section is Theorem 3, which asserts that the information matrix resulting from the log-Euclidean Rényi relative entropy is equal to the Kubo–Mori information matrix, for all values of the Rényi parameter α .

The log-Euclidean Rényi relative entropy of positive definite states ρ and σ is defined for all $\alpha \in (0, 1) \cup (1, \infty)$ as

$$D_\alpha^b(\rho \| \sigma) := \frac{1}{\alpha - 1} \ln \text{Tr}[\exp(\alpha \ln \rho + (1 - \alpha) \ln \sigma)]. \quad (3.1)$$

See [AD15, Section 4] and [MO17, Eq. (17)]. It is equal to the standard (Umegaki) quantum relative entropy [Ume62] in the limit $\alpha \rightarrow 1$:

$$\lim_{\alpha \rightarrow 1} D_\alpha^b(\rho \parallel \sigma) = D(\rho \parallel \sigma) := \text{Tr}[\rho (\ln \rho - \ln \sigma)]. \quad (3.2)$$

The data-processing inequality holds for D_α^b if $\alpha \in (0, 1)$, and it does not hold if $\alpha > 1$, as a special case of Fact 9 with $z \rightarrow \infty$.

Let $(\rho(\theta))_{\theta \in \mathbb{R}^L}$ be a parameterized family of positive definite states, where θ is an L -dimensional real parameter vector. The log-Euclidean information matrix is defined in terms of (3.1) and for all $\alpha \in (0, 1) \cup (1, \infty)$ as follows:

$$[I_\alpha^b(\theta)]_{i,j} := \frac{1}{\alpha} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha^b(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \right). \quad (3.3)$$

Definition 2 (Kubo–Mori information matrix). The Kubo–Mori information matrix $I^{\text{KM}}(\theta)$ is defined in terms of its matrix elements as

$$[I_{\text{KM}}(\theta)]_{i,j} := \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \right) \quad (3.4)$$

and it has the following alternative expressions:

$$[I_{\text{KM}}(\theta)]_{i,j} = \text{Tr}[(\partial_i \rho(\theta)) (\partial_j \ln \rho(\theta))], \quad (3.5)$$

$$= \int_0^\infty dt \text{Tr}[(\partial_i \rho(\theta)) (\rho(\theta) + tI)^{-1} (\partial_j \rho(\theta)) (\rho(\theta) + tI)^{-1}] \quad (3.6)$$

$$= \sum_{k,\ell} f_{\ln}^{[1]}(\lambda_k, \lambda_\ell) \text{Tr}[(\partial_i \rho(\theta)) \Pi_k (\partial_j \rho(\theta)) \Pi_\ell], \quad (3.7)$$

where $\partial_i \equiv \frac{\partial}{\partial \theta_i}$, a spectral decomposition of $\rho(\theta)$ is given by $\rho(\theta) = \sum_k \lambda_k \Pi_k$, and the first divided difference $f_{\ln}^{[1]}(x, y)$ of the logarithm function $x \mapsto \ln x$ is defined for all $x, y > 0$ as

$$f_{\ln}^{[1]}(x, y) := \begin{cases} \frac{\ln x - \ln y}{x - y} & : x \neq y \\ \frac{1}{x} & : x = y \end{cases}. \quad (3.8)$$

The equalities in (3.6) and (3.7) follow from, e.g., [SMP⁺22, Eqs. (B9), (B12), (B22), (B23)], and the first equality follows from applying Proposition 48 to (3.6). See also Theorem 42 for concluding (3.7) from (3.5).

Theorem 3. *The log-Euclidean information matrix in (3.3) is equal to the Kubo–Mori information matrix for all $\alpha \in (0, 1) \cup (1, \infty)$:*

$$I_\alpha^b(\theta) = I_{\text{KM}}(\theta). \quad (3.9)$$

Proof. Defining

$$\omega(\alpha, \theta, \varepsilon) := \exp(\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta + \varepsilon)), \quad (3.10)$$

consider that

$$\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha^b(\rho(\theta) \parallel \rho(\theta + \varepsilon)) = \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \left(\frac{1}{\alpha - 1} \ln \text{Tr}[\omega(\alpha, \theta, \varepsilon)] \right) \quad (3.11)$$

$$= \frac{1}{\alpha - 1} \frac{\partial}{\partial \varepsilon_i} \left(\frac{\partial}{\partial \varepsilon_j} \ln \text{Tr}[\omega(\alpha, \theta, \varepsilon)] \right) \quad (3.12)$$

$$= \frac{1}{\alpha - 1} \frac{\partial}{\partial \varepsilon_i} \left(\frac{\frac{\partial}{\partial \varepsilon_j} \text{Tr}[\omega(\alpha, \theta, \varepsilon)]}{\text{Tr}[\omega(\alpha, \theta, \varepsilon)]} \right). \quad (3.13)$$

Now consider that

$$\frac{\partial}{\partial \varepsilon_j} \text{Tr}[\omega(\alpha, \theta, \varepsilon)] = \frac{\partial}{\partial \varepsilon_j} \text{Tr}[\exp(\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta + \varepsilon))] \quad (3.14)$$

$$= \text{Tr} \left[\frac{\exp(\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta + \varepsilon)) \times \frac{\partial}{\partial \varepsilon_j} (\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta + \varepsilon))}{\exp(\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta + \varepsilon))} \right] \quad (3.15)$$

$$= (1 - \alpha) \text{Tr} \left[\omega(\alpha, \theta, \varepsilon) \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \right], \quad (3.16)$$

where the penultimate equality follows from Corollary 43. Then, plugging (3.16) into (3.13), we find that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha^b(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \\ &= - \frac{\partial}{\partial \varepsilon_i} \left(\frac{\text{Tr} \left[\omega(\alpha, \theta, \varepsilon) \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \right]}{\text{Tr}[\omega(\alpha, \theta, \varepsilon)]} \right) \end{aligned} \quad (3.17)$$

$$\begin{aligned} &= \frac{\left(\frac{\partial}{\partial \varepsilon_i} \text{Tr}[\omega(\alpha, \theta, \varepsilon)] \right) \text{Tr} \left[\omega(\alpha, \theta, \varepsilon) \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \right]}{(\text{Tr}[\omega(\alpha, \theta, \varepsilon)])^2} \\ &\quad - \frac{\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\omega(\alpha, \theta, \varepsilon) \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \right]}{\text{Tr}[\omega(\alpha, \theta, \varepsilon)]} \end{aligned} \quad (3.18)$$

$$\begin{aligned} &= (1 - \alpha) \frac{\text{Tr} \left[\omega(\alpha, \theta, \varepsilon) \left(\frac{\partial}{\partial \varepsilon_i} \ln \rho(\theta + \varepsilon) \right) \right] \text{Tr} \left[\omega(\alpha, \theta, \varepsilon) \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \right]}{(\text{Tr}[\omega(\alpha, \theta, \varepsilon)])^2} \\ &\quad - \frac{\text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_i} \omega(\alpha, \theta, \varepsilon) \right) \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \right]}{\text{Tr}[\omega(\alpha, \theta, \varepsilon)]} \\ &\quad - \frac{\text{Tr} \left[\omega(\alpha, \theta, \varepsilon) \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \right]}{\text{Tr}[\omega(\alpha, \theta, \varepsilon)]}. \end{aligned} \quad (3.19)$$

The last equality follows from (3.14)–(3.16) and the product rule for derivatives. By observing that

$$\omega(\alpha, \theta, 0) := \exp(\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta)) \quad (3.20)$$

$$= \rho(\theta), \quad (3.21)$$

$$\left. \frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right|_{\varepsilon=0} = \frac{\partial}{\partial \theta_j} \ln \rho(\theta), \quad (3.22)$$

$$\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right|_{\varepsilon=0} = \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln \rho(\theta), \quad (3.23)$$

it follows that

$$\begin{aligned} & \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha^b(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \right|_{\varepsilon=0} \\ &= (1 - \alpha) \frac{\text{Tr} \left[\omega(\alpha, \theta, 0) \left(\left. \frac{\partial}{\partial \varepsilon_i} \ln \rho(\theta + \varepsilon) \right|_{\varepsilon=0} \right) \right] \text{Tr} \left[\omega(\alpha, \theta, 0) \left(\left. \frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right|_{\varepsilon=0} \right) \right]}{(\text{Tr}[\omega(\alpha, \theta, 0)])^2} \\ & \quad - \frac{\text{Tr} \left[\left(\left. \frac{\partial}{\partial \varepsilon_i} \omega(\alpha, \theta, \varepsilon) \right|_{\varepsilon=0} \right) \left(\left. \frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right|_{\varepsilon=0} \right) \right]}{\text{Tr}[\omega(\alpha, \theta, 0)]} \\ & \quad - \frac{\text{Tr} \left[\omega(\alpha, \theta, 0) \left(\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right|_{\varepsilon=0} \right) \right]}{\text{Tr}[\omega(\alpha, \theta, 0)]} \end{aligned} \quad (3.24)$$

$$\begin{aligned} &= (1 - \alpha) \text{Tr} \left[\rho(\theta) \left(\frac{\partial}{\partial \theta_i} \ln \rho(\theta) \right) \right] \text{Tr} \left[\rho(\theta) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] \\ & \quad - \text{Tr} \left[\left(\left. \frac{\partial}{\partial \varepsilon_i} \omega(\alpha, \theta, \varepsilon) \right|_{\varepsilon=0} \right) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] - \text{Tr} \left[\rho(\theta) \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln \rho(\theta) \right) \right] \end{aligned} \quad (3.25)$$

$$= - \text{Tr} \left[\left(\left. \frac{\partial}{\partial \varepsilon_i} \omega(\alpha, \theta, \varepsilon) \right|_{\varepsilon=0} \right) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] - \text{Tr} \left[\rho(\theta) \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln \rho(\theta) \right) \right], \quad (3.26)$$

where the final equality follows because

$$\text{Tr} \left[\rho(\theta) \left(\frac{\partial}{\partial \theta_i} \ln \rho(\theta) \right) \right] = \text{Tr} \left[\rho(\theta) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] = 0, \quad (3.27)$$

due to [MPW25a, Eqs. (E78)–(E84)]. Now consider, from Proposition 47, that

$$\begin{aligned} & \frac{\partial}{\partial \varepsilon_i} \omega(\alpha, \theta, \varepsilon) \\ &= \frac{\partial}{\partial \varepsilon_i} \exp(\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta + \varepsilon)) \\ &= \int_0^1 dt \, e^{t[\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta + \varepsilon)]} \left[\frac{\partial}{\partial \varepsilon_i} (\alpha \ln \rho(\theta) + (1 - \alpha) \ln \rho(\theta + \varepsilon)) \right] \times \end{aligned} \quad (3.28)$$

$$e^{(1-t)[\alpha \ln \rho(\theta) + (1-\alpha) \ln \rho(\theta + \varepsilon)]} \quad (3.29)$$

$$= (1-\alpha) \int_0^1 dt [\omega(\alpha, \theta, \varepsilon)]^t \left[\frac{\partial}{\partial \varepsilon_i} \ln \rho(\theta + \varepsilon) \right] [\omega(\alpha, \theta, \varepsilon)]^{1-t}. \quad (3.30)$$

Then we find that

$$\left. \frac{\partial}{\partial \varepsilon_i} \omega(\alpha, \theta, \varepsilon) \right|_{\varepsilon=0} = (1-\alpha) \int_0^1 dt [\omega(\alpha, \theta, 0)]^t \left[\frac{\partial}{\partial \varepsilon_i} \ln \rho(\theta + \varepsilon) \right]_{\varepsilon=0} [\omega(\alpha, \theta, 0)]^{1-t} \quad (3.31)$$

$$= (1-\alpha) \int_0^1 dt \rho(\theta)^t \left[\frac{\partial}{\partial \theta_i} \ln \rho(\theta) \right] \rho(\theta)^{1-t} \quad (3.32)$$

$$= (1-\alpha) \frac{\partial}{\partial \theta_i} \rho(\theta). \quad (3.33)$$

where the last equality follows from Proposition 47 because

$$\frac{\partial}{\partial \theta_i} \rho(\theta) = \frac{\partial}{\partial \theta_i} (e^{\ln \rho(\theta)}) \quad (3.34)$$

$$= \int_0^1 dt \rho(\theta)^t \left[\frac{\partial}{\partial \theta_i} \ln \rho(\theta) \right] \rho(\theta)^{1-t}. \quad (3.35)$$

Substituting (3.33) into (3.26), we conclude that

$$\begin{aligned} & \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha^\flat(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \right|_{\varepsilon=0} \\ &= (\alpha - 1) \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] - \text{Tr} \left[\rho(\theta) \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln \rho(\theta) \right) \right] \end{aligned} \quad (3.36)$$

$$= (\alpha - 1) \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] + \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] \quad (3.37)$$

$$= \alpha \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right], \quad (3.38)$$

where the second equality follows from [MPW25a, Eq. (E87)]. Applying Proposition 48 and Theorem 42, we conclude that

$$\begin{aligned} & \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] \\ &= \int_0^\infty dt \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) (\rho(\theta) + tI)^{-1} \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) (\rho(\theta) + tI)^{-1} \right] \end{aligned} \quad (3.39)$$

$$= \sum_{k, \ell} f_{\ln}^{[1]}(\lambda_k, \lambda_\ell) \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_k \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \Pi_\ell \right], \quad (3.40)$$

thus completing the proof after dividing (3.38) by α . $\square$

Remark 4. As stated previously, the log-Euclidean Rényi relative entropy D_α^b obeys the data-processing inequality for all $\alpha \in (0, 1)$, and it does not when $\alpha > 1$. However, as indicated by Theorem 3, the log-Euclidean information matrix is equal to the Kubo–Mori information matrix for all $\alpha \in (0, 1) \cup (1, \infty)$. As such, the log-Euclidean information matrix obeys the data-processing inequality for all $\alpha \in (0, 1) \cup (1, \infty)$ even though the log-Euclidean Rényi relative entropy obeys it only for $\alpha \in (0, 1)$.

4 RLD Fisher information matrix from geometric relative entropies

The main results of this section are Theorem 5, which asserts that the information matrix resulting from the geometric Rényi relative entropy is equal to the RLD Fisher information matrix, for all values of the Rényi parameter α , and Theorem 7, which asserts that the information matrix resulting from the Belavkin–Staszewski relative entropy is also equal to the RLD Fisher information matrix.

4.1 RLD Fisher information matrix from geometric Rényi relative entropies

The geometric Rényi relative entropy is defined for all $\alpha \in (0, 1) \cup (1, \infty)$ and positive definite states ρ and σ as follows [Mat13, Mat18]:

$$\widehat{D}_\alpha(\rho \parallel \sigma) := \frac{1}{\alpha - 1} \ln \operatorname{Tr} \left[\sigma \left(\sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}} \right)^\alpha \right] \quad (4.1)$$

$$= \frac{1}{\alpha - 1} \ln \operatorname{Tr} \left[\rho \left(\rho^{-\frac{1}{2}} \sigma \rho^{-\frac{1}{2}} \right)^{1-\alpha} \right], \quad (4.2)$$

where the second equality follows from [KW20, Proposition 7.39]. The data-processing inequality holds for $\widehat{D}_\alpha$ if $\alpha \in (0, 1) \cup (1, 2]$, and it does not for $\alpha > 2$ [MBV24, Example 3.36].

The elements of the geometric information matrix induced by this relative entropy are defined for a second-order differentiable parameterized family $(\rho(\theta))_{\theta \in \mathbb{R}^L}$ of positive definite states as

$$\left[\widehat{I}_\alpha(\theta) \right]_{i,j} := \frac{1}{\alpha} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \widehat{D}_\alpha(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \right). \quad (4.3)$$

The elements of the right-logarithmic derivative (RLD) information matrix are given by

$$[I_{\text{RLD}}(\theta)]_{i,j} := \Re \left[\operatorname{Tr} \left[(\partial_i \rho(\theta)) \rho(\theta)^{-1} (\partial_j \rho(\theta)) \right] \right] \quad (4.4)$$

$$= \frac{1}{2} \operatorname{Tr} \left[\{ \partial_i \rho(\theta), \partial_j \rho(\theta) \} \rho(\theta)^{-1} \right], \quad (4.5)$$

where $\partial_i \equiv \frac{\partial}{\partial \theta_i}$. This is actually the real part of what is commonly called the RLD information matrix (see, e.g., [SK20, Eq. (176)]), but for brevity and simplicity, I refer to it here and throughout as the RLD information matrix.

Theorem 5. Let $(\rho(\theta))_{\theta \in \mathbb{R}^L}$ be a second-order differentiable parameterized family of positive definite states. The geometric information matrix $\hat{I}_\alpha(\theta)$ in (4.3) is equal to the RLD information matrix for all $\alpha \in (0, 1) \cup (1, \infty)$:

$$\hat{I}_\alpha(\theta) = I_{\text{RLD}}(\theta). \quad (4.6)$$

Proof. Consider that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \hat{D}_\alpha(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \\ &= \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \frac{1}{\alpha - 1} \ln \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \end{aligned} \quad (4.7)$$

$$= \frac{1}{\alpha - 1} \frac{\partial}{\partial \varepsilon_i} \frac{\partial}{\partial \varepsilon_j} \ln \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \quad (4.8)$$

$$= \frac{1}{\alpha - 1} \frac{\partial}{\partial \varepsilon_i} \left(\frac{\frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right]}{\text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right]} \right) \quad (4.9)$$

$$\begin{aligned} &= -\frac{1}{\alpha - 1} \left(\frac{\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \times \frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right]}{\left(\text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \right)^2} \right) \\ &+ \frac{1}{\alpha - 1} \left(\frac{\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right]}{\text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right]} \right). \end{aligned} \quad (4.10)$$

Then we find that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \hat{D}_\alpha(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \\ &= -\frac{1}{\alpha - 1} \left(\frac{\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0} \times \frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0}}{\left(\text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \right)^2} \right) \end{aligned}$$

$$+ \frac{1}{\alpha - 1} \left(\frac{\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0}}{\text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right]} \right) \quad (4.11)$$

$$= -\frac{1}{\alpha - 1} \left(\frac{\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0}}{\frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0}} \times \right. \\ \left. + \frac{1}{\alpha - 1} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0} \right) \right). \quad (4.12)$$

Let us now prove that

$$\begin{aligned} \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0} \\ = \frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0} = 0. \end{aligned} \quad (4.13)$$

To this end, define

$$\Delta_\varepsilon \equiv \rho(\theta + \varepsilon) - \rho(\theta). \quad (4.14)$$

By the assumption that the positive operator-valued function $\theta \mapsto \rho(\theta)$ is second-order differentiable, it follows that it is continuous. In particular, for all $\theta \in \mathbb{R}^L$ and $\delta > 0$, there exists $\gamma > 0$ such that for all $\varepsilon \in \mathbb{R}^L$ satisfying $\|\varepsilon\| \leq \gamma$, the inequality $\|\rho(\theta + \varepsilon) - \rho(\theta)\| \leq \delta$ holds. Let us choose $\delta \in (0, \lambda_{\min}(\rho(\theta)))$, where $\lambda_{\min}(\rho(\theta))$ denotes the minimum eigenvalue of $\rho(\theta)$. Then by continuity, there exists $\gamma > 0$, such that for all $\varepsilon \in \mathbb{R}^L$ satisfying $\|\varepsilon\| \leq \gamma$, the inequality $\|\rho(\theta + \varepsilon) - \rho(\theta)\| \leq \delta$ holds. Proceeding with this choice of γ and restricting to $\varepsilon \in \mathbb{R}^L$ satisfying $\|\varepsilon\| \leq \gamma$, consider that

$$\begin{aligned} \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \\ = \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} (\Delta_\varepsilon + \rho(\theta)) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \end{aligned} \quad (4.15)$$

$$= \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\rho(\theta) \left(I + \rho(\theta)^{-\frac{1}{2}} \Delta_\varepsilon \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \quad (4.16)$$

$$= \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\rho(\theta) \left[\sum_{n=0}^{\infty} \binom{1-\alpha}{n} \left(\rho(\theta)^{-\frac{1}{2}} \Delta_\varepsilon \rho(\theta)^{-\frac{1}{2}} \right)^n \right] \right] \quad (4.17)$$

$$= \frac{\partial}{\partial \varepsilon_i} \left(\sum_{n=0}^{\infty} \binom{1-\alpha}{n} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \Delta_\varepsilon \rho(\theta)^{-\frac{1}{2}} \right)^n \right] \right) \quad (4.18)$$

$$= \frac{\partial}{\partial \varepsilon_i} \left(\begin{aligned} &1 + (1-\alpha) \text{Tr}[\Delta_\varepsilon] + \frac{(1-\alpha)(-\alpha)}{2} \text{Tr}[\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon] \\ &+ \sum_{n=3}^{\infty} \binom{1-\alpha}{n} \text{Tr}[(\Delta_\varepsilon \rho(\theta)^{-1})^{n-1} \Delta_\varepsilon] \end{aligned} \right) \quad (4.19)$$

$$= \frac{\partial}{\partial \varepsilon_i} \left(\frac{(1-\alpha)(-\alpha)}{2} \text{Tr}[\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon] + \sum_{n=3}^{\infty} \binom{1-\alpha}{n} \text{Tr}[(\Delta_\varepsilon \rho(\theta)^{-1})^{n-1} \Delta_\varepsilon] \right). \quad (4.20)$$

The third equality follows from the binomial series expansion

$$(1+x)^{1-\alpha} = \sum_{n=0}^{\infty} \binom{1-\alpha}{n} x^n, \quad (4.21)$$

which converges for all $x \in \mathbb{R}$ satisfying $|x| < 1$. To see that the series

$$\sum_{n=0}^{\infty} \binom{1-\alpha}{n} \left(\rho(\theta)^{-\frac{1}{2}} \Delta_\varepsilon \rho(\theta)^{-\frac{1}{2}} \right)^n \quad (4.22)$$

indeed converges to

$$\left(I + \rho(\theta)^{-\frac{1}{2}} \Delta_\varepsilon \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha}, \quad (4.23)$$

consider that, by the choice of γ mentioned above, the following inequality holds for all $\varepsilon \in \mathbb{R}^L$ satisfying $\|\varepsilon\| \leq \gamma$:

$$\left\| \rho(\theta)^{-\frac{1}{2}} \Delta_\varepsilon \rho(\theta)^{-\frac{1}{2}} \right\| = \left\| \rho(\theta)^{-\frac{1}{2}} (\rho(\theta + \varepsilon) - \rho(\theta)) \rho(\theta)^{-\frac{1}{2}} \right\| \quad (4.24)$$

$$\leq \left\| \rho(\theta)^{-\frac{1}{2}} \right\| \left\| \rho(\theta + \varepsilon) - \rho(\theta) \right\| \left\| \rho(\theta)^{-\frac{1}{2}} \right\| \quad (4.25)$$

$$= \left\| \rho(\theta)^{-1} \right\| \left\| \rho(\theta + \varepsilon) - \rho(\theta) \right\| \quad (4.26)$$

$$= \frac{\left\| \rho(\theta + \varepsilon) - \rho(\theta) \right\|}{\lambda_{\min}(\rho(\theta))} \quad (4.27)$$

$$< 1, \quad (4.28)$$

where the first inequality follows from submultiplicativity of the spectral norm. The last equality in (4.20) follows because $\text{Tr}[\Delta_\varepsilon] = 0$ and $\frac{\partial}{\partial \varepsilon_i}(1) = 0$. Now consider that

$$\frac{\partial}{\partial \varepsilon_i} \left(\frac{(1-\alpha)(-\alpha)}{2} \text{Tr}[\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon] + \sum_{n=3}^{\infty} \binom{1-\alpha}{n} \text{Tr}[(\Delta_\varepsilon \rho(\theta)^{-1})^{n-1} \Delta_\varepsilon] \right) \Big|_{\varepsilon=0} = 0. \quad (4.29)$$

To see this, observe that applying the product rule to the first term and evaluating it at $\varepsilon = 0$ gives

$$\begin{aligned} \left(\frac{\partial}{\partial \varepsilon_i} \text{Tr}[\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon] \right) \Big|_{\varepsilon=0} &= \\ \left(\text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \right] + \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \right] \right) \Big|_{\varepsilon=0} &= 0, \end{aligned} \quad (4.30)$$

which follows because $\lim_{\varepsilon \rightarrow 0} \Delta_\varepsilon = 0$. Similar reasoning implies that

$$\frac{\partial}{\partial \varepsilon_i} \text{Tr}[(\Delta_\varepsilon \rho(\theta)^{-1})^{n-1} \Delta_\varepsilon] \Big|_{\varepsilon=0} = 0 \quad (4.31)$$

for all $n \geq 3$. We thus conclude (4.13).

Then the expression in (4.12) simplifies as follows:

$$\begin{aligned} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \widehat{D}_\alpha(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} &= \\ \frac{1}{\alpha - 1} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0} \right). \end{aligned} \quad (4.32)$$

To evaluate this last term, we can employ the same binomial series expansion from (4.20) to conclude that

$$\begin{aligned} & \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{1-\alpha} \right] \Big|_{\varepsilon=0} \right) \\ &= \left(\frac{(1-\alpha)(-\alpha)}{2} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} [\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon] \right. \\ & \quad \left. + \sum_{n=3}^{\infty} \binom{1-\alpha}{n} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} [(\Delta_\varepsilon \rho(\theta)^{-1})^{n-1} \Delta_\varepsilon] \right) \Big|_{\varepsilon=0} \end{aligned} \quad (4.33)$$

$$= \frac{\alpha(\alpha-1)}{2} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} [\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon] \Big|_{\varepsilon=0}. \quad (4.34)$$

The last equality follows because, for all terms for which $n \geq 3$, applying the product rule to $\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} [(\Delta_\varepsilon \rho(\theta)^{-1})^{n-1} \Delta_\varepsilon]$ leaves an expression featuring a sum of terms, each of which contains at least one Δ_ε remaining, which in turn evaluates to zero after taking the limit $\varepsilon \rightarrow 0$. Thus, for all $n \geq 3$,

$$\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} [(\Delta_\varepsilon \rho(\theta)^{-1})^{n-1} \Delta_\varepsilon] \Big|_{\varepsilon=0} = 0. \quad (4.35)$$

For example, for $n = 3$, we find that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} [(\Delta_\varepsilon \rho(\theta)^{-1})^2 \Delta_\varepsilon] \\ &= \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} [\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon] \end{aligned} \quad (4.36)$$

$$= \frac{\partial}{\partial \varepsilon_i} \left(\frac{\partial}{\partial \varepsilon_j} \text{Tr} [\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon] \right) \quad (4.37)$$

$$= \frac{\partial}{\partial \varepsilon_i} \left(\text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \right] + \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \right] \right. \\ \left. + \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \right] \right) \quad (4.38)$$

$$\begin{aligned} &= \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \right] + \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \right] \\ & \quad + \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \right] \end{aligned} \quad (4.39)$$

$$\begin{aligned}
&= \text{Tr} \left[\left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \right] + \text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \right] \\
&\quad + \text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \right] + \text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \right] \\
&\quad + \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \right] + \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \right] \\
&\quad + \text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \right] + \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \right] \\
&\quad + \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \Delta_\varepsilon \right) \right]. \tag{4.40}
\end{aligned}$$

Thus, in the final expression in (4.40), each term contains Δ_ε , and after taking the $\varepsilon \rightarrow 0$ limit, each term evaluates to zero. This analysis in fact implies, by means of the Hölder and triangle inequalities, that

$$\left| \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[(\Delta_\varepsilon \rho(\theta)^{-1})^2 \Delta_\varepsilon \right] \right| \leq 3d \|\rho(\theta)^{-1}\|^2 \|\Delta_\varepsilon\| \left(\|\Delta_\varepsilon\| \left\| \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \Delta_\varepsilon \right\| + 2 \left\| \frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right\| \left\| \frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right\| \right), \tag{4.41}$$

where d is the dimension of the underlying Hilbert space. As such, taking the limit $\varepsilon \rightarrow 0$ implies that $\lim_{\varepsilon \rightarrow 0} \|\Delta_\varepsilon\| = 0$, so that the term on the left-hand side vanishes. For general $n \geq 3$, this upper bound becomes

$$\left| \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[(\Delta_\varepsilon \rho(\theta)^{-1})^n \Delta_\varepsilon \right] \right| \leq nd \|\rho(\theta)^{-1}\|^n \|\Delta_\varepsilon\|^{n-2} \left(\|\Delta_\varepsilon\| \left\| \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \Delta_\varepsilon \right\| + (n-1) \left\| \frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right\| \left\| \frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right\| \right), \tag{4.42}$$

so that each of the terms indeed vanish in the limit $\varepsilon \rightarrow 0$ as claimed.

It thus remains to evaluate (4.34). To this end, consider that

$$\begin{aligned}
&\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \Delta_\varepsilon \right] \right|_{\varepsilon=0} \\
&= \left(\begin{aligned} &\text{Tr} \left[\left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \Delta_\varepsilon \right] + \text{Tr} \left[\Delta_\varepsilon \rho(\theta)^{-1} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \Delta_\varepsilon \right) \right] \\ &+ \text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \right] + \text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_j} \Delta_\varepsilon \right) \rho(\theta)^{-1} \left(\frac{\partial}{\partial \varepsilon_i} \Delta_\varepsilon \right) \right] \end{aligned} \right) \Big|_{\varepsilon=0} \tag{4.43}
\end{aligned}$$

$$= \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \rho(\theta)^{-1} \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right] + \text{Tr} \left[\left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \rho(\theta)^{-1} \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \right] \tag{4.44}$$

$$= \text{Tr} \left[\left\{ \frac{\partial}{\partial \theta_j} \rho(\theta), \frac{\partial}{\partial \theta_i} \rho(\theta) \right\} \rho(\theta)^{-1} \right]. \tag{4.45}$$

Finally, putting together the expressions from (4.3), (4.32), (4.34), and (4.45), we conclude the desired equality in (4.6). $\square$

Remark 6. As stated previously, the geometric Rényi relative entropy $\widehat{D}_\alpha$ obeys the data-processing inequality for all $\alpha \in (0, 1) \cup (1, 2]$, and it does not when $\alpha > 2$. However, as indicated by Theorem 5, the geometric information matrix is equal to the RLD Fisher information matrix for all $\alpha \in (0, 1) \cup (1, \infty)$. As such, the geometric information matrix obeys the data-processing inequality for all $\alpha \in (0, 1) \cup (1, \infty)$ even though the geometric Rényi relative entropy obeys it only for $\alpha \in (0, 1) \cup (1, 2]$.

4.2 RLD Fisher information matrix from Belavkin–Staszewski relative entropy

The Belavkin–Staszewski relative entropy is defined for positive definite states ρ and σ as follows [BS82]:

$$\widehat{D}(\rho\|\sigma) := \text{Tr} \left[\rho \ln \left(\rho^{\frac{1}{2}} \sigma^{-1} \rho^{\frac{1}{2}} \right) \right]. \quad (4.46)$$

The data-processing inequality holds for $\widehat{D}$ [HP91], and the Belavkin–Staszewski relative entropy is known to be the following limit of the geometric Rényi relative entropy [Mat13, Mat18] (see also [KW21, Proposition 79]):

$$\lim_{\alpha \rightarrow 1} \widehat{D}_\alpha(\rho\|\sigma) = \widehat{D}(\rho\|\sigma). \quad (4.47)$$

Let us define the Belavkin–Staszewski information matrix to be the information matrix induced by this relative entropy. That is, for a second-order differentiable parameterized family $(\rho(\theta))_{\theta \in \mathbb{R}^L}$ of positive definite states, it is defined as

$$\left[\widehat{I}(\theta) \right]_{i,j} := \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \widehat{D}(\rho(\theta)\|\rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \right). \quad (4.48)$$

Theorem 7. *Let $(\rho(\theta))_{\theta \in \mathbb{R}^L}$ be a second-order differentiable parameterized family of positive definite states. The Belavkin–Staszewski information matrix $\widehat{I}(\theta)$, defined in (4.48), is equal to the RLD information matrix:*

$$\widehat{I}(\theta) = I_{\text{RLD}}(\theta). \quad (4.49)$$

Proof. Employing the definition $\Delta_\varepsilon \equiv \rho(\theta + \varepsilon) - \rho(\theta)$ as in (4.14), consider that

$$\begin{aligned} & \widehat{D}(\rho(\theta)\|\rho(\theta + \varepsilon)) \\ &= \text{Tr} \left[\rho(\theta) \ln \left(\rho(\theta)^{\frac{1}{2}} \rho(\theta + \varepsilon)^{-1} \rho(\theta)^{\frac{1}{2}} \right) \right] \end{aligned} \quad (4.50)$$

$$= \text{Tr} \left[\rho(\theta) \ln \left(\left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right)^{-1} \right) \right] \quad (4.51)$$

$$= - \text{Tr} \left[\rho(\theta) \ln \left(\rho(\theta)^{-\frac{1}{2}} \rho(\theta + \varepsilon) \rho(\theta)^{-\frac{1}{2}} \right) \right] \quad (4.52)$$

$$= - \text{Tr} \left[\rho(\theta) \ln \left(\rho(\theta)^{-\frac{1}{2}} [\Delta_\varepsilon + \rho(\theta)] \rho(\theta)^{-\frac{1}{2}} \right) \right] \quad (4.53)$$

$$= - \text{Tr} \left[\rho(\theta) \ln \left(I + \rho(\theta)^{-\frac{1}{2}} \Delta_\varepsilon \rho(\theta)^{-\frac{1}{2}} \right) \right] \quad (4.54)$$

$$= -\operatorname{Tr} \left[\rho(\theta) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \left(\rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right)^n \right] \quad (4.55)$$

$$= \operatorname{Tr} \left[\rho(\theta) \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \left(\rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right)^n \right] \quad (4.56)$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \operatorname{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right)^n \right] \quad (4.57)$$

$$= -\operatorname{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right) \right] + \frac{1}{2} \operatorname{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right)^2 \right] \\ + \sum_{n=3}^{\infty} \frac{(-1)^n}{n} \operatorname{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right)^n \right] \quad (4.58)$$

$$= \frac{1}{2} \operatorname{Tr} [\Delta_{\varepsilon} \rho(\theta)^{-1} \Delta_{\varepsilon}] + \sum_{n=3}^{\infty} \frac{(-1)^n}{n} \operatorname{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right)^n \right]. \quad (4.59)$$

The equality in (4.55) follows from the series expansion $\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$, which converges for $|x| < 1$. The fact that $\left\| \rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right\| < 1$ holds is justified by the same arguments given in (4.24)–(4.28). So then

$$\left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \widehat{D}(\rho(\theta) \| \rho(\theta + \varepsilon)) \right) \Big|_{\varepsilon=0} \\ = \frac{1}{2} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \operatorname{Tr} [\Delta_{\varepsilon} \rho(\theta)^{-1} \Delta_{\varepsilon}] \right) \Big|_{\varepsilon=0} \\ + \sum_{n=3}^{\infty} \frac{(-1)^n}{n} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \operatorname{Tr} \left[\rho(\theta) \left(\rho(\theta)^{-\frac{1}{2}} \Delta_{\varepsilon} \rho(\theta)^{-\frac{1}{2}} \right)^n \right] \right) \Big|_{\varepsilon=0} \quad (4.60)$$

$$= \operatorname{Tr} \left[\left\{ \frac{\partial}{\partial \theta_j} \rho(\theta), \frac{\partial}{\partial \theta_i} \rho(\theta) \right\} \rho(\theta)^{-1} \right]. \quad (4.61)$$

The second term in (4.60) vanishes for the same reasons given in (4.35)–(4.42). The final equality follows from (4.43)–(4.45). $\square$

Remark 8. In the special case of a single parameter, Theorem 7 was established by [Mat13, Mat18, Section 6.4] (see also [KW21, Proposition 53]).

5 α - z Information matrices from α - z Rényi relative entropies

The main result of this section is Theorem 10, which provides a formula for the information matrix resulting from the α - z Rényi relative entropies.

Recall that the α - z Rényi relative entropies are defined for all $\alpha \in (0, 1) \cup (1, \infty)$ and $z > 0$ and all positive definite states ρ and σ as follows [AD15]:

$$D_{\alpha,z}(\rho\|\sigma) := \frac{1}{\alpha-1} \ln \text{Tr} \left[\left(\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^z \right] \quad (5.1)$$

$$= \frac{1}{\alpha-1} \ln \text{Tr} \left[\left(\rho^{\frac{\alpha}{2z}} \sigma^{\frac{1-\alpha}{z}} \rho^{\frac{\alpha}{2z}} \right)^z \right]. \quad (5.2)$$

The data-processing inequality is the statement that the following inequality holds for all states ρ and σ and every channel $\mathcal{N}$:

$$D_{\alpha,z}(\rho\|\sigma) \geq D_{\alpha,z}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)). \quad (5.3)$$

The full range of $\alpha, z > 0$ for which the data-processing inequality holds was established in [Zha20, Theorem 1.2] and is recalled as Fact 9 below:

Fact 9. *The data-processing inequality holds for $D_{\alpha,z}$ if and only if one of the following conditions holds [Zha20, Theorem 1.2]:*

- $0 < \alpha < 1$ and $z \geq \max\{\alpha, 1 - \alpha\}$,
- $\alpha > 1$ and $\alpha - 1 \leq z \leq \alpha \leq 2z$.

It should be stated that the data-processing inequality for $0 < \alpha < 1$ and $z \geq \max\{\alpha, 1 - \alpha\}$ follows from [Hia13, Theorem 2.1], as observed in [AD15, Theorem 1].

The α - z information matrix is defined for a parameterized family $(\rho(\theta))_{\theta \in \mathbb{R}^L}$ of positive definite states as follows:

$$[I_{\alpha,z}(\theta)]_{i,j} := \frac{1}{\alpha} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha,z}(\rho(\theta)\|\rho(\theta + \varepsilon)) \Big|_{\varepsilon=0}. \quad (5.4)$$

Theorem 10. *For a parameterized family $(\rho(\theta))_{\theta \in \mathbb{R}^L}$ of second-order differentiable, positive definite states and for all $\alpha \in (0, 1) \cup (1, \infty)$ and $z > 0$, the following equality holds:*

$$[I_{\alpha,z}(\theta)]_{i,j} = \sum_{k,\ell} \zeta_{\alpha,z}(\lambda_k, \lambda_\ell) \text{Tr}[\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))], \quad (5.5)$$

where $\partial_i \equiv \frac{\partial}{\partial \theta_i}$, the spectral decomposition of $\rho(\theta)$ is given by $\rho(\theta) = \sum_k \lambda_k \Pi_k$, and for all $x, y > 0$,

$$\zeta_{\alpha,z}(x, y) := \begin{cases} \frac{z}{\alpha(1-\alpha)} \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x-y} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) & : x \neq y \\ \frac{1}{x} & : x = y \end{cases}. \quad (5.6)$$

Proof. Our first reduction is to prove that

$$\left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha,z}(\rho(\theta)\|\rho(\theta + \varepsilon)) \right) \Big|_{\varepsilon=0} =$$

$$\frac{1}{\alpha - 1} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right) \Big|_{\varepsilon=0}. \quad (5.7)$$

To this end, consider that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha, z}(\rho(\theta) \| \rho(\theta + \varepsilon)) \\ &= \frac{1}{\alpha - 1} \frac{\partial}{\partial \varepsilon_i} \left[\frac{\partial}{\partial \varepsilon_j} \ln \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right] \end{aligned} \quad (5.8)$$

$$= \frac{1}{\alpha - 1} \frac{\partial}{\partial \varepsilon_i} \left[\frac{\frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right]}{\text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right]} \right] \quad (5.9)$$

$$\begin{aligned}
&= -\frac{1}{\alpha-1} \frac{\left[\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \times \right.}{\left(\text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right)^2} \\
&\quad \left. + \frac{1}{\alpha-1} \left[\frac{\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right]}{\text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right]} \right] \right]. \quad (5.10)
\end{aligned}$$

Then it follows that

$$\begin{aligned}
& \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha, z}(\rho(\theta) \| \rho(\theta + \varepsilon)) \right|_{\varepsilon=0} \\
&= -\frac{1}{\alpha - 1} \frac{\left[\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right]_{\varepsilon=0} \times \left[\frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right]_{\varepsilon=0}}{\left(\text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right)^2} \\
& \quad + \frac{1}{\alpha - 1} \left[\frac{\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right]_{\varepsilon=0}}{\text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right]} \right] \quad (5.11)
\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{\alpha-1} \left[\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right]_{\varepsilon=0} \times \\
&\quad \left[\frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right]_{\varepsilon=0} \Bigg] \\
&\quad + \frac{1}{\alpha-1} \left[\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right]_{\varepsilon=0} \Bigg]. \quad (5.12)
\end{aligned}$$

Lemma 11 implies that

$$\left. \frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right|_{\varepsilon=0} = 0 \quad (5.13)$$

for all i , so that (5.7) follows.

Now proceeding to analyze (5.7), consider that

$$\begin{aligned} & \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha, z}(\rho(\theta) \|\rho(\theta + \varepsilon)) \right|_{\varepsilon=0} \\ &= \frac{1}{\alpha - 1} \left[\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \right]_{\varepsilon=0} \end{aligned} \quad (5.14)$$

$$= \frac{z}{\alpha - 1} \left(\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\begin{array}{c} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \times \\ \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \varepsilon_j} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \end{array} \right] \right) \right|_{\varepsilon=0} \quad (5.15)$$

$$\begin{aligned} &= \frac{z}{\alpha - 1} \text{Tr} \left[\begin{array}{c} \left(\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \right) \times \\ \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \varepsilon_j} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \end{array} \right] \bigg|_{\varepsilon=0} \\ &\quad + \frac{z}{\alpha - 1} \text{Tr} \left[\begin{array}{c} \left(\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \right) \times \\ \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \end{array} \right] \bigg|_{\varepsilon=0} \end{aligned} \quad (5.16)$$

$$\begin{aligned} &= \frac{z}{\alpha - 1} \text{Tr} \left[\begin{array}{c} \left(\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \right) \times \\ \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \end{array} \right] \bigg|_{\varepsilon=0} \\ &\quad + \frac{z}{\alpha - 1} \text{Tr} \left[\begin{array}{c} \left(\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \right) \times \\ \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \end{array} \right] \end{aligned} \quad (5.17)$$

$$\begin{aligned} &= \frac{z}{\alpha - 1} \text{Tr} \left[\begin{array}{c} \left(\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \right) \times \\ \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \end{array} \right] \bigg|_{\varepsilon=0} \\ &\quad + \frac{z}{\alpha - 1} \text{Tr} \left[\rho(\theta)^{\frac{z-1+\alpha}{z}} \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right]. \end{aligned} \quad (5.18)$$

For the penultimate equality, we applied the following:

$$\left. \frac{\partial}{\partial \varepsilon_j} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right|_{\varepsilon=0} = \frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}}, \quad (5.19)$$

$$\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right|_{\varepsilon=0} = \frac{\partial^2}{\partial \theta_i \partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}}. \quad (5.20)$$

Let us handle each of the terms in (5.18) individually. Beginning with the first term in (5.18) and applying Proposition 49 with the substitutions $r \rightarrow z - 1$, $\frac{\partial}{\partial x} \rightarrow \frac{\partial}{\partial \varepsilon_i}$, and

$A(x) \rightarrow \rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}}$, consider that

$$\begin{aligned} & \left. \frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \right|_{\varepsilon=0} \\ &= (z-1) \int_0^1 dt \int_0^\infty ds \left[\frac{\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{(z-1)t}}{\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} + sI} \times \right. \\ & \quad \left. \left(\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right) \right) \times \right. \\ & \quad \left. \frac{\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{(z-1)(1-t)}}{\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} + sI} \right] \Big|_{\varepsilon=0} \end{aligned} \quad (5.21)$$

$$\begin{aligned} &= (z-1) \int_0^1 dt \int_0^\infty ds \left[\frac{\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{(z-1)t}}{\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} + sI} \times \right. \\ & \quad \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right) \right) \Big|_{\varepsilon=0} \rho(\theta)^{\frac{\alpha}{2z}} \times \\ & \quad \left. \frac{\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{(z-1)(1-t)}}{\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} + sI} \right] \end{aligned} \quad (5.22)$$

$$\begin{aligned} &= (z-1) \int_0^1 dt \int_0^\infty ds \left[\frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)t} \rho(\theta)^{\frac{\alpha}{2z}}}{\rho(\theta)^{\frac{1}{z}} + sI} \times \right. \\ & \quad \left. \left(\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right) \right) \Big|_{\varepsilon=0} \times \right. \\ & \quad \left. \rho(\theta)^{\frac{\alpha}{2z}} \frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)(1-t)}}{\rho(\theta)^{\frac{1}{z}} + sI} \right] \end{aligned} \quad (5.23)$$

$$\begin{aligned} &= (z-1) \int_0^1 dt \int_0^\infty ds \left[\frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)t} \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right)}{\rho(\theta)^{\frac{1}{z}} + sI} \times \right. \\ & \quad \left. \rho(\theta)^{\frac{\alpha}{2z}} \frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)(1-t)}}{\rho(\theta)^{\frac{1}{z}} + sI} \right]. \end{aligned} \quad (5.24)$$

In the last line, we applied the identity

$$\left(\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right) \right) \Big|_{\varepsilon=0} = \frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}}. \quad (5.25)$$

Substituting (5.24) into the first term of (5.18), we find that

$$\begin{aligned} & \text{Tr} \left[\frac{\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1}}{\rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}}} \Big|_{\varepsilon=0} \times \right. \\ & \quad \left. \right] \\ &= (z-1) \int_0^1 dt \int_0^\infty ds \text{Tr} \left[\frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)t} \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right)}{\rho(\theta)^{\frac{1}{z}} + sI} \times \right. \\ & \quad \rho(\theta)^{\frac{\alpha}{2z}} \frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)(1-t)}}{\rho(\theta)^{\frac{1}{z}} + sI} \times \\ & \quad \left. \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \right] \end{aligned} \quad (5.26)$$

$$= (z-1) \int_0^1 dt \int_0^\infty ds \operatorname{Tr} \left[\frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)t + \frac{\alpha}{z}}}{\rho(\theta)^{\frac{1}{z}} + sI} \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \times \right. \\ \left. \frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)(1-t) + \frac{\alpha}{z}}}{\rho(\theta)^{\frac{1}{z}} + sI} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right]. \quad (5.27)$$

Now substituting the spectral decomposition $\rho(\theta) = \sum_k \lambda_k \Pi_k$, consider that

$$(z-1) \int_0^1 dt \int_0^\infty ds \operatorname{Tr} \left[\frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)t + \frac{\alpha}{z}}}{\rho(\theta)^{\frac{1}{z}} + sI} \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \times \right. \\ \left. \frac{\rho(\theta)^{\left(\frac{z-1}{z}\right)(1-t) + \frac{\alpha}{z}}}{\rho(\theta)^{\frac{1}{z}} + sI} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \\ = (z-1) \int_0^1 dt \int_0^\infty ds \operatorname{Tr} \left[\frac{\sum_k \frac{\lambda_k^{\left(\frac{z-1}{z}\right)t + \frac{\alpha}{z}}}{\lambda_k^{\frac{1}{z}} + s} \Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \times}{\sum_\ell \frac{\lambda_\ell^{\left(\frac{z-1}{z}\right)(1-t) + \frac{\alpha}{z}}}{\lambda_\ell^{\frac{1}{z}} + s} \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right)} \right] \quad (5.28)$$

$$= (z-1) \sum_{k,\ell} \int_0^1 dt \int_0^\infty ds \left(\frac{\lambda_k^{\left(\frac{z-1}{z}\right)t + \frac{\alpha}{z}}}{\lambda_k^{\frac{1}{z}} + s} \right) \left(\frac{\lambda_\ell^{\left(\frac{z-1}{z}\right)(1-t) + \frac{\alpha}{z}}}{\lambda_\ell^{\frac{1}{z}} + s} \right) \times \\ \operatorname{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \quad (5.29)$$

$$= (z-1) \sum_{k,\ell} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \int_0^1 dt \lambda_k^{\left(\frac{z-1}{z}\right)t} \lambda_\ell^{\left(\frac{z-1}{z}\right)(1-t)} \int_0^\infty ds \left(\frac{1}{\lambda_k^{\frac{1}{z}} + s} \right) \left(\frac{1}{\lambda_\ell^{\frac{1}{z}} + s} \right) \times \\ \operatorname{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \quad (5.30)$$

$$= (z-1) \sum_{k,\ell} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \frac{1}{\frac{z-1}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\ln \lambda_k - \ln \lambda_\ell} \right) \left(\frac{\ln \lambda_k^{\frac{1}{z}} - \ln \lambda_\ell^{\frac{1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \times \\ \operatorname{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \quad (5.31)$$

$$= \sum_{k,\ell} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \operatorname{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right]. \quad (5.32)$$

For the penultimate equality, we applied the following integrals, which hold for $x, y > 0$:

$$\int_0^1 dt x^{\left(\frac{z-1}{z}\right)t} y^{\left(\frac{z-1}{z}\right)(1-t)} = \frac{1}{\frac{z-1}{z}} \left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{\ln x - \ln y} \right), \quad (5.33)$$

$$\int_0^\infty ds \left(\frac{1}{x^{\frac{1}{z}} + s} \right) \left(\frac{1}{y^{\frac{1}{z}} + s} \right) = \frac{\ln x^{\frac{1}{z}} - \ln y^{\frac{1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}}, \quad (5.34)$$

and we have left it implicit above that one evaluates these expressions in the limit $x \rightarrow y$ when $x = y$. Now applying Theorem 42 to evaluate $\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}}$ and $\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}}$, consider that

$$\begin{aligned} & \sum_{k,\ell} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \\ &= \sum_{k,\ell} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \times \\ & \quad \text{Tr} \left[\begin{array}{c} \Pi_k \left(\sum_{m,n} \left(\frac{\lambda_m^{\frac{1-\alpha}{z}} - \lambda_n^{\frac{1-\alpha}{z}}}{\lambda_m - \lambda_n} \right) \Pi_m \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_n \right) \times \\ \Pi_\ell \left(\sum_{p,r} \left(\frac{\lambda_p^{\frac{1-\alpha}{z}} - \lambda_r^{\frac{1-\alpha}{z}}}{\lambda_p - \lambda_r} \right) \Pi_p \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \Pi_r \right) \end{array} \right] \end{aligned} \quad (5.35)$$

$$\begin{aligned} &= \sum_{k,\ell,m,n,p,r} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \left(\frac{\lambda_m^{\frac{1-\alpha}{z}} - \lambda_n^{\frac{1-\alpha}{z}}}{\lambda_m - \lambda_n} \right) \times \\ & \quad \left(\frac{\lambda_p^{\frac{1-\alpha}{z}} - \lambda_r^{\frac{1-\alpha}{z}}}{\lambda_p - \lambda_r} \right) \text{Tr} \left[\Pi_k \Pi_m \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_n \Pi_\ell \Pi_p \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \Pi_r \right] \end{aligned} \quad (5.36)$$

$$\begin{aligned} &= \sum_{k,\ell} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \left(\frac{\lambda_\ell^{\frac{1-\alpha}{z}} - \lambda_k^{\frac{1-\alpha}{z}}}{\lambda_\ell - \lambda_k} \right) \times \\ & \quad \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right] \end{aligned} \quad (5.37)$$

$$\begin{aligned} &= \sum_{k,\ell} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right)^2 \times \\ & \quad \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right]. \end{aligned} \quad (5.38)$$

Let us now simplify the second term in (5.18). Before doing so, consider that

$$\begin{aligned} \frac{\partial}{\partial \theta_i} \text{Tr} \left[\rho(\theta)^{\frac{z-1+\alpha}{z}} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] &= \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{z-1+\alpha}{z}} \right) \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \\ & \quad + \text{Tr} \left[\rho(\theta)^{\frac{z-1+\alpha}{z}} \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right]. \end{aligned} \quad (5.39)$$

Now observe that

$$\begin{aligned} & \text{Tr} \left[\rho(\theta)^{\frac{z-1+\alpha}{z}} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \\ &= \text{Tr} \left[\sum_k \lambda_k^{\frac{z-1+\alpha}{z}} \Pi_k \left(\sum_{m,\ell} \left(\frac{\lambda_m^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_m - \lambda_\ell} \right) \Pi_m \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \Pi_\ell \right) \right] \end{aligned} \quad (5.40)$$

$$= \sum_{k,m,\ell} \lambda_k^{\frac{z-1+\alpha}{z}} \left(\frac{\lambda_m^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_m - \lambda_\ell} \right) \text{Tr} \left[\Pi_k \Pi_m \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \Pi_\ell \right] \quad (5.41)$$

$$= \sum_k \lambda_k^{\frac{z-1+\alpha}{z}} \left(\frac{1-\alpha}{z} \right) \lambda_k^{\frac{1-\alpha}{z}-1} \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right] \quad (5.42)$$

$$= \left(\frac{1-\alpha}{z} \right) \sum_k \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right] \quad (5.43)$$

$$= \left(\frac{1-\alpha}{z} \right) \text{Tr} \left[\frac{\partial}{\partial \theta_j} \rho(\theta) \right] \quad (5.44)$$

$$= \left(\frac{1-\alpha}{z} \right) \frac{\partial}{\partial \theta_j} \text{Tr}[\rho(\theta)] \quad (5.45)$$

$$= 0. \quad (5.46)$$

By combining (5.39) and (5.40)–(5.46), we thus conclude that

$$\text{Tr} \left[\rho(\theta)^{\frac{z-1+\alpha}{z}} \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] = - \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{z-1+\alpha}{z}} \right) \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right]. \quad (5.47)$$

Then

$$\begin{aligned} & \text{Tr} \left[\rho(\theta)^{\frac{z-1+\alpha}{z}} \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \\ &= - \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \rho(\theta)^{\frac{z-1+\alpha}{z}} \right) \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \end{aligned} \quad (5.48)$$

$$= - \text{Tr} \left[\left(\sum_{k,\ell} \left(\frac{\lambda_k^{\frac{z-1+\alpha}{z}} - \lambda_\ell^{\frac{z-1+\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \right) \times \right. \\ \left. \left(\sum_{m,n} \left(\frac{\lambda_m^{\frac{1-\alpha}{z}} - \lambda_n^{\frac{1-\alpha}{z}}}{\lambda_m - \lambda_n} \right) \Pi_m \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \Pi_n \right) \right] \quad (5.49)$$

$$= - \sum_{k,\ell,m,n} \left(\frac{\lambda_k^{\frac{z-1+\alpha}{z}} - \lambda_\ell^{\frac{z-1+\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \left(\frac{\lambda_m^{\frac{1-\alpha}{z}} - \lambda_n^{\frac{1-\alpha}{z}}}{\lambda_m - \lambda_n} \right) \times \\ \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \Pi_m \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \Pi_n \right] \quad (5.50)$$

$$= - \sum_{k,\ell} \left(\frac{\lambda_k^{\frac{z-1+\alpha}{z}} - \lambda_\ell^{\frac{z-1+\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \left(\frac{\lambda_\ell^{\frac{1-\alpha}{z}} - \lambda_k^{\frac{1-\alpha}{z}}}{\lambda_\ell - \lambda_k} \right) \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right] \quad (5.51)$$

$$= - \sum_{k,\ell} \left(\frac{\lambda_k^{\frac{z-1+\alpha}{z}} - \lambda_\ell^{\frac{z-1+\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right]. \quad (5.52)$$

Combining both terms in (5.18), using (5.38) and (5.52) (while omitting the prefactor $\frac{z}{\alpha-1}$

for now), we conclude that

$$\begin{aligned}
& \text{Tr} \left[\left(\frac{\partial}{\partial \varepsilon_i} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \right) \times \right. \\
& \quad \left. \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \right] + \text{Tr} \left[\rho(\theta)^{\frac{z-1+\alpha}{z}} \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \rho(\theta)^{\frac{1-\alpha}{z}} \right) \right] \\
&= \sum_{k,\ell} (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right)^2 \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right] \\
& \quad - \sum_{k,\ell} \left(\frac{\lambda_k^{\frac{z-1+\alpha}{z}} - \lambda_\ell^{\frac{z-1+\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right] \quad (5.53) \\
&= \sum_{k,\ell} \left[(\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right)^2 \right. \\
& \quad \left. - \left(\frac{\lambda_k^{\frac{z-1+\alpha}{z}} - \lambda_\ell^{\frac{z-1+\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \right] \text{Tr} \left[\Pi_k \left(\frac{\partial}{\partial \theta_i} \rho(\theta) \right) \Pi_\ell \left(\frac{\partial}{\partial \theta_j} \rho(\theta) \right) \right]. \quad (5.54)
\end{aligned}$$

Appendix C provides a long sequence of algebraic steps proving that

$$\begin{aligned}
& (\lambda_k \lambda_\ell)^{\frac{\alpha}{z}} \left(\frac{\lambda_k^{\frac{z-1}{z}} - \lambda_\ell^{\frac{z-1}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right)^2 - \left(\frac{\lambda_k^{\frac{z-1+\alpha}{z}} - \lambda_\ell^{\frac{z-1+\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \\
& \quad = - \left(\frac{\lambda_k^{\frac{1-\alpha}{z}} - \lambda_\ell^{\frac{1-\alpha}{z}}}{\lambda_k - \lambda_\ell} \right) \left(\frac{\lambda_k^{\frac{\alpha}{z}} - \lambda_\ell^{\frac{\alpha}{z}}}{\lambda_k^{\frac{1}{z}} - \lambda_\ell^{\frac{1}{z}}} \right). \quad (5.55)
\end{aligned}$$

After combining with (5.54) and incorporating the prefactor $\frac{z}{\alpha-1}$, the proof of (5.5) is concluded. For a proof of the case $x = y$ in (5.6), see Appendix D. $\square$

Lemma 11. *The following equality holds for all $\alpha, z > 0$:*

$$\frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \Big|_{\varepsilon=0} = 0. \quad (5.56)$$

Proof. If $1 - \alpha = 0$, then

$$\frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \Big|_{\varepsilon=0} = \frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \Big|_{\varepsilon=0} \quad (5.57)$$

$$= 0. \quad (5.58)$$

So suppose that $1 - \alpha \neq 0$, and consider that

$$\frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right]$$

$$= z \operatorname{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \frac{\partial}{\partial \varepsilon_j} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right) \right] \quad (5.59)$$

$$= z \operatorname{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \varepsilon_j} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right) \rho(\theta)^{\frac{\alpha}{2z}} \right] \quad (5.60)$$

$$= z \operatorname{Tr} \left[\rho(\theta)^{\frac{\alpha}{2z}} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \varepsilon_j} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \right) \right], \quad (5.61)$$

where the first equality follows from Corollary 43. Now applying Proposition 47, consider that

$$\begin{aligned} & \frac{\partial}{\partial \varepsilon_j} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \\ &= \frac{\partial}{\partial \varepsilon_j} e^{\left(\frac{1-\alpha}{z}\right) \ln \rho(\theta + \varepsilon)} \end{aligned} \quad (5.62)$$

$$= \int_0^1 dt e^{\left(\frac{1-\alpha}{z}\right)t \ln \rho(\theta + \varepsilon)} \left(\frac{\partial}{\partial \varepsilon_j} \left(\left(\frac{1-\alpha}{z} \right) \ln \rho(\theta + \varepsilon) \right) \right) e^{\left(\frac{1-\alpha}{z}\right)(1-t) \ln \rho(\theta + \varepsilon)} \quad (5.63)$$

$$= \frac{1-\alpha}{z} \int_0^1 dt \rho(\theta + \varepsilon)^{\left(\frac{1-\alpha}{z}\right)t} \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \rho(\theta + \varepsilon)^{\left(\frac{1-\alpha}{z}\right)(1-t)}. \quad (5.64)$$

It then follows that

$$\begin{aligned} & \frac{\partial}{\partial \varepsilon_j} \operatorname{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \\ &= (1-\alpha) \int_0^1 dt \operatorname{Tr} \left[\frac{\rho(\theta)^{\frac{\alpha}{2z}} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \rho(\theta)^{\frac{\alpha}{2z}} \times}{\rho(\theta + \varepsilon)^{\left(\frac{1-\alpha}{z}\right)t} \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \rho(\theta + \varepsilon)^{\left(\frac{1-\alpha}{z}\right)(1-t)}} \right] \end{aligned} \quad (5.65)$$

We then find that

$$\begin{aligned} & \frac{1}{1-\alpha} \frac{\partial}{\partial \varepsilon_i} \operatorname{Tr} \left[\left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \Big|_{\varepsilon=0} \\ &= \int_0^1 dt \operatorname{Tr} \left[\frac{\rho(\theta)^{\frac{\alpha}{2z}} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \rho(\theta)^{\frac{\alpha}{2z}} \times}{\rho(\theta + \varepsilon)^{\left(\frac{1-\alpha}{z}\right)t} \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \rho(\theta + \varepsilon)^{\left(\frac{1-\alpha}{z}\right)(1-t)}} \right] \Big|_{\varepsilon=0} \end{aligned} \quad (5.66)$$

$$= \int_0^1 dt \operatorname{Tr} \left[\frac{\rho(\theta)^{\frac{\alpha}{2z}} \left(\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \right)^{z-1} \rho(\theta)^{\frac{\alpha}{2z}} \times}{\rho(\theta)^{\left(\frac{1-\alpha}{z}\right)t} \left(\frac{\partial}{\partial \varepsilon_j} \ln \rho(\theta + \varepsilon) \right) \Big|_{\varepsilon=0} \rho(\theta)^{\left(\frac{1-\alpha}{z}\right)(1-t)}} \right] \quad (5.67)$$

$$= \int_0^1 dt \operatorname{Tr} \left[\frac{\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}(1-t)} \rho(\theta)^{\frac{z-1}{z}} \rho(\theta)^{\frac{1-\alpha}{z}t} \rho(\theta)^{\frac{\alpha}{2z}} \times}{\left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right)} \right] \quad (5.68)$$

$$= \operatorname{Tr} \left[\rho(\theta)^{\frac{\alpha}{2z}} \rho(\theta)^{\frac{1-\alpha}{z}} \rho(\theta)^{\frac{z-1}{z}} \rho(\theta)^{\frac{\alpha}{2z}} \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] \quad (5.69)$$

$$= \text{Tr} \left[\rho(\theta) \left(\frac{\partial}{\partial \theta_j} \ln \rho(\theta) \right) \right] \quad (5.70)$$

$$= 0. \quad (5.71)$$

The last equality follows from (E78)–(E84) of [MPW25a]. $\square$

Remark 12. For the special case of a single parameter, Theorem 10 was stated in [MH18]. Therein, a proof was not given and instead a brief suggestion for establishing the proof was provided in [MH18, Appendix A], with the authors indicating that second derivatives of matrix powers are needed to carry out the calculation. In contrast, in the approach given above, there is no need to compute a second derivative, due to the equality established in (5.47). As such, the proof given above is presumably simpler than the approach described in [MH18, Appendix A].

Remark 13. For two pure states $\psi \equiv |\psi\rangle\langle\psi|$ and $\phi \equiv |\phi\rangle\langle\phi|$ and $\alpha \in (0, 1)$, the α - z Rényi relative entropy reduces as follows:

$$D_{\alpha,z}(\psi\|\phi) := \frac{1}{\alpha-1} \ln \text{Tr} \left[\left(\phi^{\frac{1-\alpha}{2z}} \psi^{\frac{\alpha}{z}} \phi^{\frac{1-\alpha}{2z}} \right)^z \right] \quad (5.72)$$

$$= \frac{1}{\alpha-1} \ln \text{Tr} [(|\phi\rangle\langle\phi|\psi\rangle\langle\psi|\phi\rangle\langle\phi|)^z] \quad (5.73)$$

$$= \frac{1}{\alpha-1} \ln \text{Tr} [(|\phi\rangle\langle\phi|\psi\rangle\langle\psi|^2|\phi\rangle\langle\phi|)^z] \quad (5.74)$$

$$= \frac{1}{\alpha-1} \ln |\langle\phi|\psi\rangle|^{2z} \text{Tr} [(|\phi\rangle\langle\phi|)^z] \quad (5.75)$$

$$= \frac{1}{\alpha-1} \ln |\langle\phi|\psi\rangle|^{2z} \text{Tr} [|\phi\rangle\langle\phi|] \quad (5.76)$$

$$= \frac{1}{\alpha-1} \ln |\langle\phi|\psi\rangle|^{2z} \quad (5.77)$$

$$= \frac{z}{1-\alpha} (-\ln |\langle\phi|\psi\rangle|^2) \quad (5.78)$$

$$= \frac{z}{1-\alpha} (-\ln F(\psi, \phi)), \quad (5.79)$$

where $F(\psi, \phi) := |\langle\phi|\psi\rangle|^2$ is the fidelity. As expected, for $\psi \neq \phi$, the limit $D_{\alpha,z}(\psi\|\phi) \rightarrow \infty$ holds as $\alpha \rightarrow 1$ for fixed $z > 0$ or $D_{\alpha,z}(\psi\|\phi) \rightarrow \infty$ as $z \rightarrow \infty$ for fixed $\alpha \in (0, 1)$, consistent with the fact that $D_{\alpha,z}$ converges to the standard (Umegaki) relative entropy in these limits and this latter quantity being equal to ∞ for two distinct pure states. Thus, for a parameterized family $(\psi(\theta))_{\theta \in \mathbb{R}^L}$ of pure states, the α - z information matrix reduces to

$$[I_{\alpha,z}(\theta)]_{i,j} := \frac{1}{\alpha} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha,z}(\psi(\theta)\|\psi(\theta + \varepsilon)) \Big|_{\varepsilon=0}. \quad (5.80)$$

$$= \frac{z}{\alpha(1-\alpha)} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (-\ln F(\psi(\theta), \psi(\theta + \varepsilon))) \Big|_{\varepsilon=0} \quad (5.81)$$

$$= \frac{2z}{\alpha(1-\alpha)} \Re [\langle \partial_i \psi(\theta) | (I - |\psi(\theta)\rangle\langle\psi(\theta)|) | \partial_j \psi(\theta) \rangle], \quad (5.82)$$

consistent with the known result [LYLW19, Theorem 2.5] that

$$\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (-2 \ln F(\psi(\theta), \psi(\theta + \varepsilon))) \right|_{\varepsilon=0} = 4 \Re [\langle \partial_i \psi(\theta) | (I - |\psi(\theta)\rangle\langle\psi(\theta)|) | \partial_j \psi(\theta) \rangle]. \quad (5.83)$$

Appendix E reviews the proof of (5.83).

Corollary 14. *As a consequence of Theorem 1 and Theorem 10, the following function is operator monotone on $x \in (0, \infty)$ for the values of α and z stated in Fact 9:*

$$x \mapsto \frac{\alpha(1-\alpha)}{z} \frac{(x-1) \left(x^{\frac{1}{z}} - 1 \right)}{\left(x^{\frac{1-\alpha}{z}} - 1 \right) \left(x^{\frac{\alpha}{z}} - 1 \right)}. \quad (5.84)$$

6 Special cases of α - z information matrices

In this section, I consider special cases of the α - z information matrices, as previously done as well in [MH18, CDL⁺18], which amounts to evaluating them for particular values or in various limits. In particular, the Kubo–Mori information matrix arises for all $\alpha \in (0, 1) \cup (1, \infty)$ in the limit $z \rightarrow \infty$ or for all $z > 0$ in the limit $\alpha \rightarrow 1$ (Section 6.1). The sandwiched Rényi information matrix arises when $z = \alpha$ (Section 6.3), and the Petz–Rényi information matrix arises when $z = 1$ (Section 6.2). All information matrices considered in this section are evaluated for a second-order differentiable family of positive definite states, $(\rho(\theta))_{\theta \in \mathbb{R}^L}$.

6.1 Kubo–Mori information matrix

Proposition 15. *For all $x, y > 0$ such that $x \neq y$ and for all $\alpha \in (0, 1) \cup (1, \infty)$,*

$$\lim_{z \rightarrow \infty} \zeta_{\alpha, z}(x, y) = \frac{\ln x - \ln y}{x - y}. \quad (6.1)$$

Thus, for all $\alpha \in (0, 1) \cup (1, \infty)$,

$$\lim_{z \rightarrow \infty} I_{\alpha, z}(\theta) = I_{\text{KM}}(\theta). \quad (6.2)$$

Proof. Consider that

$$\lim_{z \rightarrow \infty} \zeta_{\alpha, z}(x, y) = \lim_{z \rightarrow \infty} \frac{z}{\alpha(1-\alpha)} \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x - y} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \quad (6.3)$$

$$= \lim_{z \rightarrow \infty} \frac{1}{\alpha(1-\alpha)} \left(\frac{1}{x - y} \right) \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{\frac{1}{z}} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \quad (6.4)$$

$$= \frac{1}{\alpha(1-\alpha)} \left(\frac{1}{x-y} \right) \left[\lim_{z \rightarrow \infty} \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{\frac{1}{z}} \right) \right] \left[\lim_{z \rightarrow \infty} \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \right] \quad (6.5)$$

$$= \frac{1}{\alpha(1-\alpha)} \left(\frac{1}{x-y} \right) \left[\lim_{h \rightarrow 0} \left(\frac{x^{(1-\alpha)h} - y^{(1-\alpha)h}}{h} \right) \right] \left[\lim_{h \rightarrow 0} \left(\frac{x^{\alpha h} - y^{\alpha h}}{x^h - y^h} \right) \right]. \quad (6.6)$$

Now consider that

$$\lim_{h \rightarrow 0} \frac{x^{(1-\alpha)h} - y^{(1-\alpha)h}}{h} = \lim_{h \rightarrow 0} \frac{x^{(1-\alpha)h}}{h} - \lim_{h \rightarrow 0} \frac{y^{(1-\alpha)h}}{h} \quad (6.7)$$

$$= (1-\alpha) \ln x - (1-\alpha) \ln y \quad (6.8)$$

$$= (1-\alpha) (\ln x - \ln y) \quad (6.9)$$

and

$$\lim_{h \rightarrow 0} \frac{x^{\alpha h} - y^{\alpha h}}{x^h - y^h} = \lim_{h \rightarrow 0} \frac{x^{\alpha h} \alpha \ln x - y^{\alpha h} \alpha \ln y}{x^h \ln x - y^h \ln y} \quad (6.10)$$

$$= \alpha \lim_{h \rightarrow 0} \frac{x^{\alpha h} \ln x - y^{\alpha h} \ln y}{x^h \ln x - y^h \ln y} \quad (6.11)$$

$$= \alpha \left(\frac{\ln x - \ln y}{\ln x - \ln y} \right) \quad (6.12)$$

$$= \alpha. \quad (6.13)$$

Then putting together (6.6), (6.9), and (6.13), we conclude (6.1). The equality in (6.2) follows from (6.1) and (3.7). $\square$

Consistent with [LT15, Proposition 3], the following holds:

Proposition 16. *For all $x, y, z > 0$ such that $x \neq y$,*

$$\lim_{\alpha \rightarrow 1} \zeta_{\alpha, z}(x, y) = \frac{\ln x - \ln y}{x - y}. \quad (6.14)$$

Thus, for all $z > 0$,

$$\lim_{\alpha \rightarrow 1} I_{\alpha, z}(\theta) = I_{\text{KM}}(\theta). \quad (6.15)$$

Proof. Consider that

$$\lim_{\alpha \rightarrow 1} \zeta_{\alpha, z}(x, y) = \lim_{\alpha \rightarrow 1} \frac{z}{\alpha(1-\alpha)} \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x - y} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \quad (6.16)$$

$$= \left(\frac{z}{x - y} \right) \lim_{\alpha \rightarrow 1} \frac{1}{\alpha} \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{1 - \alpha} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \quad (6.17)$$

$$= \left(\frac{z}{x - y} \right) \left(\lim_{\alpha \rightarrow 1} \frac{1}{\alpha} \right) \left(\lim_{\alpha \rightarrow 1} \frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{1 - \alpha} \right) \left(\lim_{\alpha \rightarrow 1} \frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \quad (6.18)$$

$$= \left(\frac{z}{x-y} \right) \left(\lim_{\alpha \rightarrow 1} \frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{1-\alpha} \right) \left(\frac{x^{\frac{1}{z}} - y^{\frac{1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \quad (6.19)$$

$$= \left(\frac{z}{x-y} \right) \left(\lim_{\alpha \rightarrow 1} \frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{1-\alpha} \right) \quad (6.20)$$

$$= \left(\frac{z}{x-y} \right) \left(\lim_{\alpha \rightarrow 1} \frac{x^{\frac{1-\alpha}{z}} \left(-\frac{1}{z} \ln x \right) - y^{\frac{1-\alpha}{z}} \left(-\frac{1}{z} \ln y \right)}{-1} \right) \quad (6.21)$$

$$= \left(\frac{z}{x-y} \right) \frac{1}{z} (\ln x - \ln y) \quad (6.22)$$

$$= \frac{\ln x - \ln y}{x-y}. \quad (6.23)$$

The equality in (6.15) follows from (6.14) and (3.7). $\square$

6.2 Petz–Rényi information matrices

The Petz–Rényi relative entropy is defined for positive definite states ρ and σ and $\alpha \in (0, 1) \cup (1, \infty)$ as [Pet85, Pet86]

$$\overline{D}_\alpha(\rho \parallel \sigma) := \frac{1}{\alpha - 1} \ln \text{Tr} [\rho^\alpha \sigma^{1-\alpha}]. \quad (6.24)$$

It obeys the data-processing inequality for $\alpha \in (0, 1) \cup (1, 2]$ [Pet85, Pet86]. It is a special case of the α - z Rényi relative entropy when $z = 1$.

The elements of the Petz–Rényi information matrix, considered previously in [Has93], are defined for all $\alpha \in (0, 1) \cup (1, \infty)$ as

$$[\overline{I}_\alpha(\theta)]_{i,j} := \frac{1}{\alpha} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \overline{D}_\alpha(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \quad (6.25)$$

$$= \frac{1}{\alpha} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha,1}(\rho(\theta) \parallel \rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \quad (6.26)$$

Given that, for all $x, y > 0$ such that $x \neq y$,

$$\lim_{z \rightarrow 1} \zeta_{\alpha,z}(x, y) = \lim_{z \rightarrow 1} \frac{z}{\alpha(1-\alpha)} \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x-y} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \quad (6.27)$$

$$= \frac{1}{\alpha(1-\alpha)} \left(\frac{x^{1-\alpha} - y^{1-\alpha}}{x-y} \right) \left(\frac{x^\alpha - y^\alpha}{x-y} \right) \quad (6.28)$$

$$= \frac{1}{\alpha(1-\alpha)} \frac{(x^\alpha - y^\alpha)(x^{1-\alpha} - y^{1-\alpha})}{(x-y)^2}, \quad (6.29)$$

we can conclude the following corollary of Theorem 10:

Corollary 17. For all $\alpha \in (0, 1) \cup (1, \infty)$, the elements of the Petz–Rényi information matrix are as follows:

$$[\bar{I}_\alpha(\theta)]_{i,j} = \sum_{k,\ell} \bar{\zeta}_\alpha(\lambda_k, \lambda_\ell) \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))], \quad (6.30)$$

where $\partial_i \equiv \frac{\partial}{\partial \theta_i}$, the spectral decomposition of $\rho(\theta)$ is given by $\rho(\theta) = \sum_k \lambda_k \Pi_k$, and for all $x, y > 0$,

$$\bar{\zeta}_\alpha(x, y) := \begin{cases} \frac{1}{\alpha(1-\alpha)} \frac{(x^\alpha - y^\alpha)(x^{1-\alpha} - y^{1-\alpha})}{(x-y)^2} & : x \neq y \\ \frac{1}{x} & : x = y \end{cases}. \quad (6.31)$$

Interestingly, the elements of the Petz–Rényi information matrix $\bar{I}_\alpha(\theta)$ are α -dependent, in contrast to the log-Euclidean and geometric information matrices (recall Theorem 3 and Theorem 5). I explore this point further in Section 7.

For $\alpha \in (0, 1)$, Proposition 18 gives an integral representation for the elements of the Petz–Rényi information matrix, which is basis independent, and for $\alpha = 2$, Corollary 19 states that the Petz–Rényi information matrix is equal to the RLD information matrix.

Proposition 18. For all $\alpha \in (0, 1)$, the following integral representation holds for the Petz–Rényi information matrix:

$$[\bar{I}_\alpha(\theta)]_{i,j} = \frac{\sin^2(\alpha\pi)}{\alpha(1-\alpha)\pi^2} \int_0^\infty \int_0^\infty ds \, dt \, s^\alpha t^{1-\alpha} \text{Tr} \left[\frac{(\rho(\theta) + sI)^{-1}(\rho(\theta) + tI)^{-1}(\partial_i \rho(\theta))}{(\rho(\theta) + sI)^{-1}(\rho(\theta) + tI)^{-1}(\partial_j \rho(\theta))} \times \right]. \quad (6.32)$$

Proof. Using the integral representation from (B.102), it follows that, for $x, y > 0$ such that $x \neq y$,

$$\bar{\zeta}_\alpha(x, y) = \frac{1}{\alpha(1-\alpha)} \frac{(x^\alpha - y^\alpha)(x^{1-\alpha} - y^{1-\alpha})}{(x-y)^2} \quad (6.33)$$

$$= \frac{\sin(\alpha\pi)}{\alpha\pi} \frac{\sin((1-\alpha)\pi)}{(1-\alpha)\pi} \int_0^\infty ds \, \frac{s^\alpha}{(x+s)(y+s)} \int_0^\infty dt \, \frac{t^{1-\alpha}}{(x+t)(y+t)} \quad (6.34)$$

$$= \frac{\sin^2(\alpha\pi)}{\alpha(1-\alpha)\pi^2} \int_0^\infty \int_0^\infty ds \, dt \, \frac{s^\alpha t^{1-\alpha}}{(x+s)(x+t)(y+s)(y+t)}, \quad (6.35)$$

which implies, after defining $g(\alpha) \equiv \frac{\sin^2(\alpha\pi)}{\alpha(1-\alpha)\pi^2}$, that

$$[\bar{I}_\alpha(\theta)]_{i,j} = \sum_{k,\ell} \bar{\zeta}_\alpha(\lambda_k, \lambda_\ell) \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))] \quad (6.36)$$

$$= \sum_{k,\ell} g(\alpha) \int_0^\infty \int_0^\infty ds \, dt \, \frac{s^\alpha t^{1-\alpha}}{(\lambda_k + s)(\lambda_k + t)(\lambda_\ell + s)(\lambda_\ell + t)} \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))] \quad (6.37)$$

$$= g(\alpha) \int_0^\infty \int_0^\infty ds dt s^\alpha t^{1-\alpha} \sum_{k,\ell} \frac{\text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))]}{(\lambda_k + s)(\lambda_k + t)(\lambda_\ell + s)(\lambda_\ell + t)} \quad (6.38)$$

$$= g(\alpha) \int_0^\infty \int_0^\infty ds dt s^\alpha t^{1-\alpha} \text{Tr} \left[\frac{\sum_k \left(\frac{1}{(\lambda_k + s)(\lambda_k + t)} \right) \Pi_k(\partial_i \rho(\theta)) \times}{\sum_\ell \left(\frac{1}{(\lambda_\ell + s)(\lambda_\ell + t)} \right) \Pi_\ell(\partial_j \rho(\theta))} \right] \quad (6.39)$$

$$= g(\alpha) \int_0^\infty \int_0^\infty ds dt s^\alpha t^{1-\alpha} \text{Tr} \left[\frac{(\rho(\theta) + sI)^{-1} (\rho(\theta) + tI)^{-1} (\partial_i \rho(\theta)) \times}{(\rho(\theta) + sI)^{-1} (\rho(\theta) + tI)^{-1} (\partial_j \rho(\theta))} \right], \quad (6.40)$$

thus concluding the proof. $\square$

Corollary 19. For $\alpha = 2$, the following equality holds:

$$\bar{I}_2(\theta) = I_{\text{RLD}}(\theta). \quad (6.41)$$

Proof. Considering that, for all $x, y > 0$ such that $x \neq y$,

$$\lim_{\alpha \rightarrow 2} \bar{\zeta}_\alpha(x, y) = \lim_{\alpha \rightarrow 2} \frac{1}{\alpha(1-\alpha)} \frac{(x^\alpha - y^\alpha)(x^{1-\alpha} - y^{1-\alpha})}{(x-y)^2} \quad (6.42)$$

$$= -\frac{1}{2} \frac{(x^2 - y^2)(x^{-1} - y^{-1})}{(x-y)^2} \quad (6.43)$$

$$= -\frac{1}{2} \frac{(x+y)(x-y)x^{-1}y^{-1}(y-x)}{(x-y)^2} \quad (6.44)$$

$$= \frac{1}{2} (x+y)x^{-1}y^{-1} \quad (6.45)$$

$$= \frac{1}{2} \left(\frac{1}{x} + \frac{1}{y} \right), \quad (6.46)$$

it follows that

$$[\bar{I}_2(\theta)]_{i,j} = \sum_{k,\ell} \frac{1}{2} \left(\frac{1}{\lambda_k} + \frac{1}{\lambda_\ell} \right) \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))] \quad (6.47)$$

$$\begin{aligned} &= \frac{1}{2} \sum_{k,\ell} \frac{1}{\lambda_k} \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))] \\ &\quad + \frac{1}{2} \sum_{k,\ell} \frac{1}{\lambda_\ell} \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))] \end{aligned} \quad (6.48)$$

$$\begin{aligned} &= \frac{1}{2} \text{Tr} \left[\left(\sum_k \frac{1}{\lambda_k} \Pi_k \right) (\partial_i \rho(\theta)) \left(\sum_\ell \Pi_\ell \right) (\partial_j \rho(\theta)) \right] \\ &\quad + \frac{1}{2} \text{Tr} \left[\left(\sum_k \Pi_k \right) (\partial_i \rho(\theta)) \left(\sum_\ell \frac{1}{\lambda_\ell} \Pi_\ell \right) (\partial_j \rho(\theta)) \right] \end{aligned} \quad (6.49)$$

$$= \frac{1}{2} \text{Tr}[\rho(\theta)^{-1} (\partial_i \rho(\theta)) (\partial_j \rho(\theta))] + \frac{1}{2} \text{Tr}[(\partial_i \rho(\theta)) \rho(\theta)^{-1} (\partial_j \rho(\theta))] \quad (6.50)$$

$$= \frac{1}{2} \text{Tr}[\{\partial_i \rho(\theta), \partial_j \rho(\theta)\} \rho(\theta)^{-1}] \quad (6.51)$$

$$= [I_{\text{RLD}}(\theta)]_{i,j}, \quad (6.52)$$

thus concluding the proof. $\square$

Remark 20. Corollary 19 is consistent with the fact that $\overline{D}_2(\rho\|\sigma) = \widehat{D}_2(\rho\|\sigma)$ and Theorem 5.

6.3 Sandwiched Rényi information matrices

The sandwiched Rényi relative entropy is defined for positive definite states ρ and σ and $\alpha \in (0, 1) \cup (1, \infty)$ as [MLDS⁺13, WWY14]

$$\widetilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \ln \text{Tr} \left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right]. \quad (6.53)$$

It obeys the data-processing inequality for $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$ [FL13] (see also [Wil18]). It is a special case of the α - z Rényi relative entropy when $z = \alpha$.

The elements of the sandwiched Rényi information matrix, considered previously in [TF17], are defined for all $\alpha \in (0, 1) \cup (1, \infty)$ as

$$[\widetilde{I}_\alpha(\theta)]_{i,j} := \frac{1}{\alpha} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \widetilde{D}_\alpha(\rho(\theta)\|\rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \quad (6.54)$$

$$= \frac{1}{\alpha} \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha,\alpha}(\rho(\theta)\|\rho(\theta + \varepsilon)) \Big|_{\varepsilon=0} \quad (6.55)$$

Given that, for all $x, y > 0$ such that $x \neq y$,

$$\lim_{z \rightarrow \alpha} \zeta_{\alpha,z}(x, y) = \lim_{z \rightarrow \alpha} \frac{z}{\alpha(1-\alpha)} \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x - y} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \quad (6.56)$$

$$= \frac{1}{1-\alpha} \left(\frac{x^{\frac{1-\alpha}{\alpha}} - y^{\frac{1-\alpha}{\alpha}}}{x - y} \right) \left(\frac{x^{\frac{\alpha}{\alpha}} - y^{\frac{\alpha}{\alpha}}}{x^{\frac{1}{\alpha}} - y^{\frac{1}{\alpha}}} \right) \quad (6.57)$$

$$= \frac{1}{1-\alpha} \left(\frac{x^{\frac{1-\alpha}{\alpha}} - y^{\frac{1-\alpha}{\alpha}}}{x^{\frac{1}{\alpha}} - y^{\frac{1}{\alpha}}} \right), \quad (6.58)$$

we conclude the following corollary of Theorem 10, related to what was previously reported as [TF17, Lemma 4]:

Corollary 21. *For all $\alpha \in (0, 1) \cup (1, \infty)$, the elements of the sandwiched Rényi information matrix are as follows:*

$$[\widetilde{I}_\alpha(\theta)]_{i,j} = \sum_{k,\ell} \widetilde{\zeta}_\alpha(\lambda_k, \lambda_\ell) \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))], \quad (6.59)$$

where $\partial_i \equiv \frac{\partial}{\partial \theta_i}$, the spectral decomposition of $\rho(\theta)$ is given by $\rho(\theta) = \sum_k \lambda_k \Pi_k$, and for all $x, y > 0$,

$$\tilde{\zeta}_\alpha(x, y) := \begin{cases} \frac{1}{1-\alpha} \left(\frac{x^{\frac{1-\alpha}{\alpha}} - y^{\frac{1-\alpha}{\alpha}}}{x^{\frac{1}{\alpha}} - y^{\frac{1}{\alpha}}} \right) & : x \neq y \\ \frac{1}{x} & : x = y \end{cases}. \quad (6.60)$$

Similar to the Petz–Rényi information matrix, the elements of the sandwiched Rényi information matrix $\tilde{I}_\alpha(\theta)$ are α -dependent, again in contrast to the log-Euclidean and geometric information matrices (recall Theorem 3 and Theorem 5). I explore this point further in Section 7.

For $\alpha \in (0, 1)$, Proposition 22 gives an integral representation for the elements of the sandwiched Rényi information matrix, and for $\alpha = 2$, Corollary 23 gives a simple form for them. Both of the expressions given are basis independent.

Proposition 22. *For all $\alpha \in (0, 1)$, the following integral representation holds for the sandwiched Rényi information matrix:*

$$\begin{aligned} [\tilde{I}_\alpha(\theta)]_{i,j} &= \\ \frac{\sin((1-\alpha)\pi)}{(1-\alpha)\pi} \int_0^\infty dt \, t^{1-\alpha} \operatorname{Tr} \left[\left(\rho(\theta)^{\frac{1}{\alpha}} + tI \right)^{-1} (\partial_i \rho(\theta)) \left(\rho(\theta)^{\frac{1}{\alpha}} + tI \right)^{-1} (\partial_j \rho(\theta)) \right]. \end{aligned} \quad (6.61)$$

Proof. Using the integral representation from (B.102), it follows that, for $x, y > 0$ such that $x \neq y$,

$$\tilde{\zeta}_\alpha(x, y) = \frac{1}{1-\alpha} \left(\frac{x^{\frac{1-\alpha}{\alpha}} - y^{\frac{1-\alpha}{\alpha}}}{x^{\frac{1}{\alpha}} - y^{\frac{1}{\alpha}}} \right) \quad (6.62)$$

$$= \frac{\sin((1-\alpha)\pi)}{(1-\alpha)\pi} \int_0^\infty dt \, \frac{t^{1-\alpha}}{\left(x^{\frac{1}{\alpha}} + t\right) \left(y^{\frac{1}{\alpha}} + t\right)}, \quad (6.63)$$

which implies that

$$\begin{aligned} [\tilde{I}_\alpha(\theta)]_{i,j} &= \\ &= \sum_{k,\ell} \tilde{\zeta}_\alpha(\lambda_k, \lambda_\ell) \operatorname{Tr}[\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))] \end{aligned} \quad (6.64)$$

$$= \sum_{k,\ell} \frac{\sin((1-\alpha)\pi)}{(1-\alpha)\pi} \int_0^\infty dt \, \frac{t^{1-\alpha}}{\left(\lambda_k^{\frac{1}{\alpha}} + t\right) \left(\lambda_\ell^{\frac{1}{\alpha}} + t\right)} \operatorname{Tr}[\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))] \quad (6.65)$$

$$= \frac{\sin((1-\alpha)\pi)}{(1-\alpha)\pi} \int_0^\infty dt \, t^{1-\alpha} \sum_{k,\ell} \frac{1}{\left(\lambda_k^{\frac{1}{\alpha}} + t\right) \left(\lambda_\ell^{\frac{1}{\alpha}} + t\right)} \operatorname{Tr}[\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))] \quad (6.66)$$

$$= \frac{\sin((1-\alpha)\pi)}{(1-\alpha)\pi} \int_0^\infty dt \, t^{1-\alpha} \operatorname{Tr} \left[\sum_k \left(\frac{1}{\lambda_k^{\frac{1}{\alpha}} + t} \right) \Pi_k (\partial_i \rho(\theta)) \sum_\ell \left(\frac{1}{\lambda_\ell^{\frac{1}{\alpha}} + t} \right) \Pi_\ell (\partial_j \rho(\theta)) \right] \quad (6.67)$$

$$= \frac{\sin((1-\alpha)\pi)}{(1-\alpha)\pi} \int_0^\infty dt \, t^{1-\alpha} \text{Tr} \left[\left(\rho(\theta)^{\frac{1}{\alpha}} + tI \right)^{-1} (\partial_i \rho(\theta)) \left(\rho(\theta)^{\frac{1}{\alpha}} + tI \right)^{-1} (\partial_j \rho(\theta)) \right], \quad (6.68)$$

thus concluding the proof. $\square$

Corollary 23. *For $\alpha = 2$, the following equality holds:*

$$\left[\tilde{I}_2(\theta) \right]_{i,j} = \text{Tr} \left[\rho(\theta)^{-\frac{1}{2}} (\partial_i \rho(\theta)) \rho(\theta)^{-\frac{1}{2}} (\partial_j \rho(\theta)) \right]. \quad (6.69)$$

Proof. Consider that

$$\lim_{\alpha \rightarrow 2} \tilde{\zeta}_\alpha(x, y) = \lim_{\alpha \rightarrow 2} \frac{1}{1-\alpha} \left(\frac{x^{\frac{1-\alpha}{\alpha}} - y^{\frac{1-\alpha}{\alpha}}}{x^{\frac{1}{\alpha}} - y^{\frac{1}{\alpha}}} \right) \quad (6.70)$$

$$= - \left(\frac{x^{-\frac{1}{2}} - y^{-\frac{1}{2}}}{x^{\frac{1}{2}} - y^{\frac{1}{2}}} \right) \quad (6.71)$$

$$= -x^{-\frac{1}{2}} y^{-\frac{1}{2}} \left(\frac{y^{\frac{1}{2}} - x^{\frac{1}{2}}}{x^{\frac{1}{2}} - y^{\frac{1}{2}}} \right) \quad (6.72)$$

$$= \frac{1}{\sqrt{xy}}. \quad (6.73)$$

Then consider that, from Corollary 21,

$$\left[\tilde{I}_2(\theta) \right]_{i,j} = \sum_{k,\ell} \tilde{\zeta}_2(\lambda_k, \lambda_\ell) \text{Tr} [\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))] \quad (6.74)$$

$$= \sum_{k,\ell} \frac{1}{\sqrt{\lambda_k \lambda_\ell}} \text{Tr} [\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))] \quad (6.75)$$

$$= \text{Tr} \left[\left(\sum_k \frac{1}{\sqrt{\lambda_k}} \Pi_k \right) (\partial_i \rho(\theta)) \left(\sum_\ell \frac{1}{\sqrt{\lambda_\ell}} \Pi_\ell \right) (\partial_j \rho(\theta)) \right] \quad (6.76)$$

$$= \text{Tr} \left[\rho(\theta)^{-\frac{1}{2}} (\partial_i \rho(\theta)) \rho(\theta)^{-\frac{1}{2}} (\partial_j \rho(\theta)) \right], \quad (6.77)$$

thus concluding the proof. $\square$

7 Orderings and relations between α -z information matrices

In this section, I establish ordering relations for the Petz- and sandwiched Rényi information matrices. I begin with some orderings that are direct consequences of previous results. After that, Section 7.1 provides a lemma for orderings of general information matrices, which was mentioned in passing in [JK20, Eq. (60)]. Sections 7.2 and 7.3 then provide orderings for the Petz- and sandwiched Rényi information matrices, respectively.

Let us begin by recalling that orderings of smooth divergences imply orderings of information matrices, as reviewed in [MPW25a, Proposition 4]. As such, we can conclude some orderings as direct consequences of orderings of Rényi relative entropies that have been previously established in the literature. Given that, for states ρ and σ ,

$$\tilde{D}_\alpha(\rho\|\sigma) \leq \overline{D}_\alpha(\rho\|\sigma) \quad (7.1)$$

for all $\alpha > 0$ [WWY14, DL14] and that

$$\alpha \overline{D}_\alpha(\rho\|\sigma) \leq \tilde{D}_\alpha(\rho\|\sigma) \quad (7.2)$$

for all $\alpha \in (0, 1)$ [IRS17, Corollary 2.3], we can divide by α in both of the inequalities and apply [MPW25a, Proposition 4] to conclude the following matrix inequalities for the Petz- and sandwiched Rényi information matrices for all $\theta \in \mathbb{R}^L$:

$$\tilde{I}_\alpha(\theta) \leq \overline{I}_\alpha(\theta) \quad (7.3)$$

for all $\alpha > 0$ and that

$$\alpha \overline{I}_\alpha(\theta) \leq \tilde{I}_\alpha(\theta) \quad (7.4)$$

for all $\alpha \in (0, 1)$. Additionally, it is known that, for fixed states ρ and σ and $\alpha > 0$, the function $z \mapsto D_{\alpha,z}(\rho\|\sigma)$ is monotone increasing on $\mathbb{R}_+$ if $\alpha \in (0, 1)$ and it is monotone decreasing on $\mathbb{R}_+$ if $\alpha > 1$ [LT15, Proposition 1]. A direct consequence is the following:

Corollary 24. *For all $\theta \in \mathbb{R}^L$, the matrix-valued function $z \mapsto I_{\alpha,z}(\theta)$ is monotone increasing on $\mathbb{R}_+$ if $\alpha \in (0, 1)$ and it is monotone decreasing on $\mathbb{R}_+$ if $\alpha > 1$, with respect to the Loewner order.*

7.1 Ordering of general information matrices

Suppose that $\zeta: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a function satisfying items 1–3 of Theorem 1, and define $f(t) := \frac{1}{\zeta(t,1)}$ for all $t > 0$. Recall that a general information matrix has the following form:

$$[I_\zeta(\theta)]_{i,j} = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \operatorname{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))], \quad (7.5)$$

$$= \sum_{k,\ell} \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)} \operatorname{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))], \quad (7.6)$$

where the second equality follows because

$$\zeta(\lambda_k, \lambda_\ell) = \frac{1}{\lambda_\ell} \zeta\left(\frac{\lambda_k}{\lambda_\ell}, 1\right) = \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)}. \quad (7.7)$$

In this section, I drop the requirement for f to be operator monotone on $(0, \infty)$, in which case $I_\zeta(\theta)$ need not satisfy the data-processing inequality. However, the information matrix

$I_\zeta(\theta)$ is still positive semi-definite, as seen in Corollary 27 below, and we can still consider ordering relations between information matrices.

Let us begin with a general statement regarding ordering of information matrices built from functions f_1 and f_2 that satisfy an ordering relation on $\mathbb{R}_+$:

Lemma 25. *Let $\zeta_1, \zeta_2: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be functions that satisfy items 1–3 of Theorem 1, and define $f_i(t) := \frac{1}{\zeta_i(t,1)}$ for all $t > 0$ and $i \in \{1, 2\}$. Suppose that $f_1(t) \geq f_2(t)$ for all $t > 0$. Then the following matrix inequality holds for all $\theta \in \mathbb{R}^L$:*

$$I_{\zeta_1}(\theta) \leq I_{\zeta_2}(\theta). \quad (7.8)$$

Proof. To see this, consider that the desired matrix inequality is equivalent to

$$v^T I_{\zeta_1}(\theta) v \leq v^T I_{\zeta_2}(\theta) v \quad (7.9)$$

holding for all $v \in \mathbb{R}^L$. Defining

$$W := \sum_i v_i \partial_i \rho(\theta), \quad (7.10)$$

$$|W\rangle := (W \otimes I) |\Gamma\rangle, \quad (7.11)$$

$$|\Gamma\rangle := \sum_i |i\rangle \otimes |i\rangle, \quad (7.12)$$

consider that, for a general ζ and f , defined in the same way as ζ_i and f_i ,

$$v^T I_\zeta(\theta) v = \sum_{i,j} v_i \left(\sum_{k,\ell} \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)} \text{Tr}[\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))] \right) v_j \quad (7.13)$$

$$= \sum_{k,\ell} \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)} \text{Tr} \left[\Pi_k \left(\sum_i v_i \partial_i \rho(\theta) \right) \Pi_\ell \left(\sum_j v_j \partial_j \rho(\theta) \right) \right] \quad (7.14)$$

$$= \sum_{k,\ell} \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)} \text{Tr}[\Pi_k W \Pi_\ell W] \quad (7.15)$$

$$= \sum_{k,\ell} \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)} \langle W | (\Pi_k \otimes \Pi_\ell^T) | W \rangle \quad (7.16)$$

$$= \langle W | \left[\sum_{k,\ell} \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)} (\Pi_k \otimes \Pi_\ell^T) \right] | W \rangle, \quad (7.17)$$

where we used that

$$\text{Tr}[\Pi_k W \Pi_\ell W] = \text{Tr}[W \Pi_k W \Pi_\ell] \quad (7.18)$$

$$= \langle \Gamma | (W \Pi_k W \Pi_\ell \otimes I) | \Gamma \rangle \quad (7.19)$$

$$= \langle \Gamma | (W \Pi_k W \otimes \Pi_\ell^T) | \Gamma \rangle \quad (7.20)$$

$$= \langle \Gamma | (W \otimes I) (\Pi_k \otimes \Pi_\ell^T) (W \otimes I) | \Gamma \rangle \quad (7.21)$$

$$= \langle W | (\Pi_k \otimes \Pi_\ell^T) | W \rangle. \quad (7.22)$$

Then the desired inequality in (7.9) follows because

$$v^T I_{\zeta_1}(\theta) v = \langle W | \left[\sum_{k,\ell} \frac{1}{\lambda_\ell f_1\left(\frac{\lambda_k}{\lambda_\ell}\right)} (\Pi_k \otimes \Pi_\ell^T) \right] | W \rangle \quad (7.23)$$

$$\leq \langle W | \left[\sum_{k,\ell} \frac{1}{\lambda_\ell f_2\left(\frac{\lambda_k}{\lambda_\ell}\right)} (\Pi_k \otimes \Pi_\ell^T) \right] | W \rangle \quad (7.24)$$

$$= v^T I_{\zeta_2}(\theta) v, \quad (7.25)$$

where we used the matrix inequality

$$\sum_{k,\ell} \frac{1}{\lambda_\ell f_1\left(\frac{\lambda_k}{\lambda_\ell}\right)} (\Pi_k \otimes \Pi_\ell^T) \leq \sum_{k,\ell} \frac{1}{\lambda_\ell f_2\left(\frac{\lambda_k}{\lambda_\ell}\right)} (\Pi_k \otimes \Pi_\ell^T), \quad (7.26)$$

which follows from the assumption that $f_1(t) \geq f_2(t)$ for all $t > 0$. $\square$

Remark 26. The expression in (7.17) is a rewriting of [Pet96, Eqs. (8) & (11)] that does not make use of the left and right multiplication superoperators.

Corollary 27. *Let $\zeta: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a function that satisfies items 1–3 of Theorem 1, and define $f(t) := \frac{1}{\zeta(t,1)}$ for all $t > 0$. Then the following matrix inequality holds for all $\theta \in \mathbb{R}^L$:*

$$I_\zeta(\theta) \geq 0. \quad (7.27)$$

Proof. By the same reasoning as in the proof of Lemma 25, it follows that, for all $v \in \mathbb{R}^L$,

$$v^T I_{\zeta_1}(\theta) v = \langle W | \left[\sum_{k,\ell} \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)} (\Pi_k \otimes \Pi_\ell^T) \right] | W \rangle. \quad (7.28)$$

The matrix $\sum_{k,\ell} \left[\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right) \right]^{-1} (\Pi_k \otimes \Pi_\ell^T)$ is positive semi-definite because $\lambda_\ell > 0$, $f\left(\frac{\lambda_k}{\lambda_\ell}\right) > 0$ given that $\lambda_k, \lambda_\ell > 0$, and $\Pi_k \otimes \Pi_\ell^T \geq 0$ for all k and ℓ . Then the inequality

$$\langle W | \left[\sum_{k,\ell} \frac{1}{\lambda_\ell f\left(\frac{\lambda_k}{\lambda_\ell}\right)} (\Pi_k \otimes \Pi_\ell^T) \right] | W \rangle \geq 0 \quad (7.29)$$

holds for all $|W\rangle$, thus concluding the proof. $\square$

7.2 Ordering of Petz–Rényi information matrices

Recall that, for all $\alpha \in (0, 1) \cup (1, \infty)$, the Petz–Rényi information matrix has the form given in Corollary 17, which we can rewrite as follows by employing (7.6):

$$[\bar{I}_\alpha(\theta)]_{i,j} = \sum_{k,\ell} \frac{1}{\lambda_\ell \bar{f}\left(\frac{\lambda_k}{\lambda_\ell}, \alpha\right)} \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))], \quad (7.30)$$

where

$$\bar{f}(x, \alpha) := \alpha(1 - \alpha) \left(\frac{(x - 1)^2}{(x^\alpha - 1)(x^{1-\alpha} - 1)} \right). \quad (7.31)$$

As a consequence of Lemma 25 and Lemma 29, the following theorem holds:

Theorem 28. *For all $\theta \in \mathbb{R}^L$, the matrix-valued function $\alpha \mapsto \bar{I}_\alpha(\theta)$ is monotone decreasing on $\alpha \in (0, \frac{1}{2}]$ and monotone increasing on $\alpha \in [\frac{1}{2}, \infty)$, with respect to the Loewner order.*

Lemma 29. *For all $x > 0$, the function*

$$\alpha \mapsto \bar{f}(x, \alpha) := \alpha(1 - \alpha) \left(\frac{(x - 1)^2}{(x^\alpha - 1)(x^{1-\alpha} - 1)} \right) \quad (7.32)$$

is monotone increasing on $\alpha \in (0, \frac{1}{2}]$ and monotone decreasing on $\alpha \in [\frac{1}{2}, \infty)$.

Proof. The plan is to calculate the derivative of the function in (7.32) and show that it is non-negative for all $\alpha \in (0, \frac{1}{2}]$ and non-positive for all $\alpha \geq \frac{1}{2}$. To begin with, consider that $\bar{f}(x, \alpha) > 0$ for all $\alpha > 0$ and $x > 0$ (one can verify this by considering various cases $\alpha \in (0, 1)$, $\alpha > 1$, $x \in (0, 1)$, and $x > 1$). The derivative of $\bar{f}(x, \alpha)$ with respect to α is given by

$$\begin{aligned} & \frac{\partial}{\partial \alpha} \bar{f}(x, \alpha) \\ &= \frac{\partial}{\partial \alpha} \left(\frac{\alpha(1 - \alpha)(x - 1)^2}{(x^\alpha - 1)(x^{1-\alpha} - 1)} \right) \end{aligned} \quad (7.33)$$

$$= \frac{\frac{\partial}{\partial \alpha} [\alpha(1 - \alpha)] (x - 1)^2}{(x^\alpha - 1)(x^{1-\alpha} - 1)} - \left(\frac{\alpha(1 - \alpha)(x - 1)^2}{[(x^\alpha - 1)(x^{1-\alpha} - 1)]^2} \frac{\partial}{\partial \alpha} [(x^\alpha - 1)(x^{1-\alpha} - 1)] \right) \quad (7.34)$$

$$= \frac{(1 - 2\alpha)(x - 1)^2}{(x^\alpha - 1)(x^{1-\alpha} - 1)} - \alpha(1 - \alpha)(x - 1)^2 \frac{-x^\alpha \ln x + x^{1-\alpha} \ln x}{[(x^\alpha - 1)(x^{1-\alpha} - 1)]^2} \quad (7.35)$$

$$= \frac{(x - 1)^2}{(x^\alpha - 1)(x^{1-\alpha} - 1)} \left[(1 - 2\alpha) - \alpha(1 - \alpha) \frac{x^{1-\alpha} - x^\alpha}{(x^\alpha - 1)(x^{1-\alpha} - 1)} \ln x \right]. \quad (7.36)$$

Let us analyze the term on the right in square brackets:

$$(1 - 2\alpha) - \alpha(1 - \alpha) \frac{x^{1-\alpha} - x^\alpha}{(x^\alpha - 1)(x^{1-\alpha} - 1)} \ln x$$

$$= (1 - 2\alpha) - \alpha(1 - \alpha) \frac{x^{1-\alpha} - 1 - (x^\alpha - 1)}{(x^\alpha - 1)(x^{1-\alpha} - 1)} \ln x \quad (7.37)$$

$$= (1 - 2\alpha) - \alpha(1 - \alpha) \left(\frac{1}{x^\alpha - 1} - \frac{1}{x^{1-\alpha} - 1} \right) \ln x \quad (7.38)$$

$$= 1 - \alpha - \alpha - (1 - \alpha) \frac{\alpha \ln x}{x^\alpha - 1} - \alpha \frac{(1 - \alpha) \ln x}{x^{1-\alpha} - 1} \quad (7.39)$$

$$= (1 - \alpha) \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right) - \alpha \left(1 - \frac{(1 - \alpha) \ln x}{x^{1-\alpha} - 1} \right) \quad (7.40)$$

$$= \alpha(1 - \alpha) \left[\frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right) - \frac{1}{1 - \alpha} \left(1 - \frac{(1 - \alpha) \ln x}{x^{1-\alpha} - 1} \right) \right]. \quad (7.41)$$

Then it follows that

$$\frac{\partial}{\partial \alpha} \bar{f}(x, \alpha) = \bar{f}(x, \alpha) \left[\frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right) - \frac{1}{1 - \alpha} \left(1 - \frac{(1 - \alpha) \ln x}{x^{1-\alpha} - 1} \right) \right]. \quad (7.42)$$

Since $\bar{f}(x, \alpha) > 0$ for all $x > 0$ and $\alpha > 0$, it suffices to prove that

$$\frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right) \geq \frac{1}{1 - \alpha} \left(1 - \frac{(1 - \alpha) \ln x}{x^{1-\alpha} - 1} \right) \quad (7.43)$$

for all $\alpha \in (0, \frac{1}{2}]$ and that

$$\frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right) \leq \frac{1}{1 - \alpha} \left(1 - \frac{(1 - \alpha) \ln x}{x^{1-\alpha} - 1} \right) \quad (7.44)$$

for all $\alpha \geq \frac{1}{2}$. Since the function $\alpha \mapsto \frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right)$ is symmetric about $\alpha = \frac{1}{2}$ on $\alpha \in (0, 1)$, it suffices to prove the inequality in (7.44) for all $\alpha \geq \frac{1}{2}$. The inequality in (7.44) will follow if we prove that the function $\alpha \mapsto \frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right)$ is decreasing on $\alpha \in (-\infty, \infty)$. To do so, let us evaluate the derivative of this function as follows and prove that it is non-positive for all $x > 0$:

$$\frac{\partial}{\partial \alpha} \left(\frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right) \right) = \frac{\partial}{\partial \alpha} \left(\frac{1}{\alpha} - \frac{\ln x}{x^\alpha - 1} \right) \quad (7.45)$$

$$= -\frac{1}{\alpha^2} + \frac{\ln x}{(x^\alpha - 1)^2} x^\alpha \ln x \quad (7.46)$$

$$= -\frac{1}{\alpha^2} + \left(\frac{\ln x}{x^\alpha - 1} \right)^2 x^\alpha \quad (7.47)$$

The desired inequality $\frac{\partial}{\partial \alpha} \left(\frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right) \right) \leq 0$ is then equivalent to

$$-\frac{1}{\alpha^2} + \left(\frac{\ln x}{x^\alpha - 1} \right)^2 x^\alpha \leq 0 \quad (7.48)$$

$$\iff \left(\frac{\ln x}{x^\alpha - 1} \right)^2 x^\alpha \leq \frac{1}{\alpha^2} \quad (7.49)$$

$$\iff (\ln x)^2 \alpha^2 \leq \frac{(x^\alpha - 1)^2}{x^\alpha} \quad (7.50)$$

$$= x^\alpha - 2 + x^{-\alpha} \quad (7.51)$$

Now set $x = e^\lambda$ for $\lambda \in \mathbb{R}$ and observe that

$$(\ln x)^2 \alpha^2 \leq x^\alpha - 2 + x^{-\alpha} \quad (7.52)$$

$$\iff 2 + (\lambda \alpha)^2 \leq e^{\lambda \alpha} + e^{-\lambda \alpha} \quad (7.53)$$

$$\iff 1 + \frac{1}{2} (\lambda \alpha)^2 \leq \cosh(\lambda \alpha) \quad (7.54)$$

$$= 1 + \frac{1}{2} (\lambda \alpha)^2 + \sum_{k=2}^{\infty} \frac{(\lambda \alpha)^{2k}}{(2k)!} \quad (7.55)$$

$$\iff 0 \leq \sum_{k=2}^{\infty} \frac{(\lambda \alpha)^{2k}}{(2k)!}. \quad (7.56)$$

Thus, the function $\alpha \mapsto \frac{1}{\alpha} \left(1 - \frac{\alpha \ln x}{x^\alpha - 1} \right)$ is indeed decreasing on $\alpha \in (-\infty, \infty)$ for all $x > 0$, concluding the proof. $\square$

7.3 Ordering of sandwiched Rényi information matrices

Recall that, for all $\alpha \in (0, 1) \cup (1, \infty)$, the sandwiched Rényi information matrix has the form given in Corollary 21, which we can rewrite as follows by employing (7.6):

$$\left[\tilde{I}_\alpha(\theta) \right]_{i,j} = \sum_{k,\ell} \frac{1}{\lambda_\ell \tilde{f}\left(\frac{\lambda_k}{\lambda_\ell}, \alpha\right)} \text{Tr}[\Pi_k(\partial_i \rho(\theta)) \Pi_\ell(\partial_j \rho(\theta))], \quad (7.57)$$

where

$$\tilde{f}(x, \alpha) := (1 - \alpha) \left(\frac{x^{\frac{1}{\alpha}} - 1}{x^{\frac{1-\alpha}{\alpha}} - 1} \right). \quad (7.58)$$

As a consequence of Lemma 25 and Lemma 29, the following theorem holds:

Theorem 30. *For all $\theta \in \mathbb{R}^L$, the matrix-valued function $\alpha \mapsto \tilde{I}_\alpha(\theta)$ is monotone increasing on $\alpha \in (0, \infty)$, with respect to the Loewner order.*

Lemma 31. *For all $x > 0$, the function*

$$\alpha \mapsto \tilde{f}(x, \alpha) := (1 - \alpha) \left(\frac{x^{\frac{1}{\alpha}} - 1}{x^{\frac{1-\alpha}{\alpha}} - 1} \right) \quad (7.59)$$

is monotone decreasing on $\alpha \in (0, \infty)$.

Proof. The plan is to calculate the derivative of the function in (7.59) and show that it is non-negative for all $\alpha \in (0, \infty)$. To begin with, consider that the inequality $\tilde{f}(x, \alpha) > 0$ holds for all $\alpha > 0$ and $x > 0$ (one can verify this by considering various cases $\alpha \in (0, 1)$, $\alpha > 1$, $x \in (0, 1)$, and $x > 1$). The derivative of $\tilde{f}(x, \alpha)$ with respect to α is given by

$$\begin{aligned} & \frac{\partial}{\partial \alpha} \tilde{f}(x, \alpha) \\ &= \frac{\partial}{\partial \alpha} \left((1 - \alpha) \left(\frac{x^{\frac{1}{\alpha}} - 1}{x^{\frac{1-\alpha}{\alpha}} - 1} \right) \right) \end{aligned} \quad (7.60)$$

$$= - \left(\frac{x^{\frac{1}{\alpha}} - 1}{x^{\frac{1-\alpha}{\alpha}} - 1} \right) + (1 - \alpha) \left(\frac{-\frac{1}{\alpha^2} x^{\frac{1}{\alpha}} \ln x}{x^{\frac{1-\alpha}{\alpha}} - 1} \right) + (1 - \alpha) \left(\frac{x^{\frac{1}{\alpha}} - 1}{\left(x^{\frac{1-\alpha}{\alpha}} - 1 \right)^2} \right) \frac{1}{\alpha^2} x^{\frac{1-\alpha}{\alpha}} \ln x \quad (7.61)$$

$$= - \left(\frac{x^{\frac{1}{\alpha}} - 1}{x^{\frac{1-\alpha}{\alpha}} - 1} \right) - (1 - \alpha) \left(\frac{x^{\frac{1}{\alpha}} \ln x}{\alpha^2 \left(x^{\frac{1-\alpha}{\alpha}} - 1 \right)} \right) + (1 - \alpha) \left(\frac{x^{\frac{1}{\alpha}} - 1}{x^{\frac{1-\alpha}{\alpha}} - 1} \right) \frac{x^{\frac{1-\alpha}{\alpha}} \ln x}{\alpha^2 \left(x^{\frac{1-\alpha}{\alpha}} - 1 \right)} \quad (7.62)$$

$$= \frac{(1 - \alpha)}{\alpha^2} \left(\frac{x^{\frac{1}{\alpha}} - 1}{x^{\frac{1-\alpha}{\alpha}} - 1} \right) \left[-\frac{\alpha^2}{1 - \alpha} - \frac{x^{\frac{1}{\alpha}} \ln x}{x^{\frac{1}{\alpha}} - 1} + \frac{x^{\frac{1-\alpha}{\alpha}} \ln x}{x^{\frac{1-\alpha}{\alpha}} - 1} \right] \quad (7.63)$$

$$= \frac{\tilde{f}(x, \alpha)}{\alpha^2} \left[\frac{1}{\frac{1}{\alpha}} - \frac{1}{\frac{1-\alpha}{\alpha}} - \frac{\ln x}{1 - x^{-\frac{1}{\alpha}}} + \frac{\ln x}{1 - x^{-(\frac{1-\alpha}{\alpha})}} \right] \quad (7.64)$$

$$= \frac{\tilde{f}(x, \alpha)}{\alpha^2} \left[- \left(\frac{\ln x}{1 - x^{-\frac{1}{\alpha}}} - \frac{1}{\frac{1}{\alpha}} \right) + \frac{\ln x}{1 - x^{-(\frac{1-\alpha}{\alpha})}} - \frac{1}{\frac{1-\alpha}{\alpha}} \right]. \quad (7.65)$$

Since $\frac{\tilde{f}(x, \alpha)}{\alpha^2} > 0$ for all $\alpha > 0$ and $x > 0$, the desired inequality $\frac{\partial}{\partial \alpha} \tilde{f}(x, \alpha) \leq 0$ will follow if we prove that

$$\frac{\ln x}{1 - x^{-(\frac{1-\alpha}{\alpha})}} - \frac{1}{\frac{1-\alpha}{\alpha}} \leq \frac{\ln x}{1 - x^{-\frac{1}{\alpha}}} - \frac{1}{\frac{1}{\alpha}} \quad (7.66)$$

for all $\alpha \in (0, \infty)$. Defining $\gamma := \frac{1-\alpha}{\alpha}$, it follows that $\frac{1}{\alpha} = \gamma + 1$, so that the inequality in (7.66) is equivalent to

$$\frac{\ln x}{1 - x^{-\gamma}} - \frac{1}{\gamma} \leq \frac{\ln x}{1 - x^{-(\gamma+1)}} - \frac{1}{\gamma+1} \quad (7.67)$$

for all $\gamma \in (-\infty, \infty)$. The inequality in (7.67) will follow if we prove that the function

$$\gamma \mapsto \frac{\ln x}{1 - x^{-\gamma}} - \frac{1}{\gamma} \quad (7.68)$$

is increasing on $\gamma \in (-\infty, \infty)$. Note that $\lim_{\gamma \rightarrow 0} \left(\frac{\ln x}{1 - x^{-\gamma}} - \frac{1}{\gamma} \right) = \frac{\ln x}{2}$. Taking the derivative

of $\frac{\ln x}{1-x^{-\gamma}} - \frac{1}{\gamma}$ with respect to γ , we find that

$$\frac{\partial}{\partial \gamma} \left(\frac{\ln x}{1-x^{-\gamma}} - \frac{1}{\gamma} \right) = \frac{\ln x}{(1-x^{-\gamma})^2} (-x^{-\gamma} \ln x) + \frac{1}{\gamma^2} \quad (7.69)$$

$$= -\frac{x^{-\gamma} (\ln x)^2}{(1-x^{-\gamma})^2} + \frac{1}{\gamma^2}. \quad (7.70)$$

Note that $\lim_{\gamma \rightarrow 0} \left(-\frac{x^{-\gamma} (\ln x)^2}{(1-x^{-\gamma})^2} + \frac{1}{\gamma^2} \right) = \frac{(\ln x)^2}{12}$, so that $\frac{\partial}{\partial \gamma} \left(\frac{\ln x}{1-x^{-\gamma}} - \frac{1}{\gamma} \right) > 0$ at $\gamma = 0$ for all $x > 0$. Setting $x = e^\lambda$ for $\lambda \in \mathbb{R}$, the inequality $\frac{\partial}{\partial \gamma} \left(\frac{\ln x}{1-x^{-\gamma}} - \frac{1}{\gamma} \right) \geq 0$ for all $\gamma \in (-\infty, 0) \cup (0, \infty)$ is equivalent to

$$-\frac{x^{-\gamma} (\ln x)^2}{(1-x^{-\gamma})^2} + \frac{1}{\gamma^2} \geq 0 \quad (7.71)$$

$$\iff \frac{1}{\gamma^2} \geq \frac{e^{-\lambda\gamma} (\lambda)^2}{(1-e^{-\lambda\gamma})^2} \quad (7.72)$$

$$\iff \frac{(1-e^{-\lambda\gamma})^2}{e^{-\lambda\gamma}} \geq (\gamma\lambda)^2 \quad (7.73)$$

$$\iff e^{\lambda\gamma} (1 - 2e^{-\lambda\gamma} + e^{-2\lambda\gamma}) \geq (\gamma\lambda)^2 \quad (7.74)$$

$$\iff e^{\lambda\gamma} + e^{-\lambda\gamma} \geq 2 + (\gamma\lambda)^2 \quad (7.75)$$

$$\iff \cosh(\lambda\gamma) \geq 1 + \frac{1}{2} (\gamma\lambda)^2 \quad (7.76)$$

$$\iff 1 + \frac{1}{2} (\lambda\gamma)^2 + \sum_{k=2}^{\infty} \frac{(\lambda\gamma)^{2k}}{(2k)!} \geq 1 + \frac{1}{2} (\gamma\lambda)^2 \quad (7.77)$$

$$\iff \sum_{k=2}^{\infty} \frac{(\lambda\gamma)^{2k}}{(2k)!} \geq 0, \quad (7.78)$$

thus concluding the proof. $\square$

8 α - z Information matrices of parameterized thermal states

In this section, I establish a formula for the α - z information matrix of parameterized thermal states, for all $\alpha \in (0, 1)$ and $z > 0$ (Theorem 32). Like the former results of [PW24, MPW25a], this formula leads to a hybrid quantum-classical algorithm for estimating the elements of the α - z information matrix of parameterized thermal states, assuming that one has the ability to prepare thermal states on a quantum computer. As such, this formula has applications in quantum Boltzmann machine learning, namely, in natural gradient descent algorithms for performing optimization using quantum Boltzmann machines, as put forward in [PW24, MPW25a].

Let $(\rho(\theta))_{\theta \in \mathbb{R}^L}$ be a parameterized family of positive definite states. A quantum generalization of the Fisher information matrix is denoted by $I_\zeta(\theta)$, and it has the form stated in Theorem 1, wherein the properties of the function $\zeta(x, y)$ are stated. For convenience, let us restate the form of $I_\zeta(\theta)$ here:

$$[I_\zeta(\theta)]_{i,j} = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr}[\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))], \quad (8.1)$$

where $\rho(\theta) = \sum_k \lambda_k \Pi_k$ is a spectral decomposition of $\rho(\theta)$ and $\zeta(x, y)$ is a function of $x, y > 0$ that satisfies the following properties for all $x, y, s > 0$:

$$\zeta(x, y) = \zeta(y, x), \quad (8.2)$$

$$\zeta(sx, sy) = \frac{1}{s} \zeta(x, y), \quad (8.3)$$

$$\zeta(x, x) = \frac{\kappa}{x}. \quad (8.4)$$

with $\kappa > 0$ a constant. Also, the function $t \mapsto \frac{1}{\zeta(t, 1)}$ is operator monotone on $t \in (0, \infty)$.

Let us now evaluate the expression in (8.1) for parameterized thermal states. In this case, let

$$H(\theta) := \sum_j \theta_j H_j, \quad (8.5)$$

$$\rho(\theta) := \frac{e^{-H(\theta)}}{Z(\theta)}, \quad (8.6)$$

$$Z(\theta) := \text{Tr}[e^{-H(\theta)}], \quad (8.7)$$

where $\theta_j \in \mathbb{R}$ and H_j is Hermitian for all $j \in \{1, \dots, L\}$. As a consequence of Lemma 33 and Lemma 34, the following theorem holds for the α - z information matrices of parameterized thermal states:

Theorem 32. *For parameterized thermal states of the form in (8.6), the following equality holds for all $\alpha \in (0, 1)$ and $z > 0$:*

$$[I_{\alpha,z}(\theta)]_{i,j} = \frac{1}{2} \langle \{ \Phi_{q_{\alpha,z},\theta}(H_i), H_j \} \rangle_{\rho(\theta)} - \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)}, \quad (8.8)$$

where the channel $\Phi_{q_{\alpha,z},\theta}$ is given by

$$\Phi_{q_{\alpha,z},\theta}(X) := \int_{-\infty}^{\infty} dt q_{\alpha,z}(t) e^{-itH(\theta)} X e^{itH(\theta)}, \quad (8.9)$$

and $q_{\alpha,z}$ is the probability density function defined in (8.72).

As stated above, Theorem 32 follows in part from Lemma 33, which establishes a formula for the information matrix $I_\zeta(\theta)$ of parameterized thermal states, under the assumption that a certain Fourier transform exists. Additionally, Theorem 32 follows from Lemma 34, which precisely determines the needed Fourier transform for all $\alpha \in (0, 1)$ and $z > 0$ when the function $\zeta(x, y)$ is set to $\zeta_{\alpha,z}(x, y)$, as defined in (5.6).

8.1 General formula for information matrices of parameterized thermal states

Lemma 33. *Let $\zeta(x, y)$ be a function defined for $x, y > 0$ satisfying the properties stated in (8.2)–(8.4), and suppose that the Fourier transform of the function*

$$\omega \mapsto \zeta(e^{-\omega}, 1) \frac{(e^{-\omega} - 1)^2}{\omega^2 (e^{-\omega} + 1)} \quad (8.10)$$

exists, where $\omega \in \mathbb{R}$. Then, for parameterized thermal states of the form in (8.6), the following equality holds:

$$[I_\zeta(\theta)]_{i,j} = \frac{1}{2} \langle \{\Phi_{f,\theta}(H_i), H_j\} \rangle_{\rho(\theta)} - \kappa \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)}, \quad (8.11)$$

where $I_\zeta(\theta)$ is defined in (8.1),

$$\Phi_{f,\theta}(X) := \int_{-\infty}^{\infty} dt f(t) e^{-itH(\theta)} X e^{itH(\theta)}, \quad (8.12)$$

and $f(t)$ is a real-valued function satisfying the following for all $\omega \in \mathbb{R}$:

$$\int_{-\infty}^{\infty} dt f(t) e^{it\omega} = 2 \zeta(e^{-\omega}, 1) \frac{(e^{-\omega} - 1)^2}{\omega^2 (e^{-\omega} + 1)}, \quad (8.13)$$

$$\int_{-\infty}^{\infty} dt f(t) = \kappa. \quad (8.14)$$

Proof. Consider that

$$\frac{\partial}{\partial \theta_i} \rho(\theta) = \frac{\partial}{\partial \theta_i} \left(\frac{e^{-H(\theta)}}{Z(\theta)} \right) \quad (8.15)$$

$$= \frac{\frac{\partial}{\partial \theta_i} e^{-H(\theta)}}{Z(\theta)} - \frac{e^{-H(\theta)}}{Z(\theta)^2} \frac{\partial}{\partial \theta_i} Z(\theta) \quad (8.16)$$

$$= \frac{1}{Z(\theta)} \left(\frac{\partial}{\partial \theta_i} e^{-H(\theta)} - \rho(\theta) \frac{\partial}{\partial \theta_i} \text{Tr}[e^{-H(\theta)}] \right) \quad (8.17)$$

$$= \frac{1}{Z(\theta)} \left(\frac{\partial}{\partial \theta_i} e^{-H(\theta)} + \rho(\theta) \text{Tr} \left[e^{-H(\theta)} \frac{\partial}{\partial \theta_i} H(\theta) \right] \right) \quad (8.18)$$

$$= \frac{1}{Z(\theta)} \left(\frac{\partial}{\partial \theta_i} e^{-H(\theta)} + \rho(\theta) \text{Tr}[e^{-H(\theta)} H_i] \right) \quad (8.19)$$

$$= \frac{1}{Z(\theta)} \left(\frac{\partial}{\partial \theta_i} e^{-H(\theta)} \right) + \rho(\theta) \text{Tr}[H_i \rho(\theta)] \quad (8.20)$$

$$= \frac{1}{Z(\theta)} \left(\frac{\partial}{\partial \theta_i} e^{-H(\theta)} \right) + \rho(\theta) \langle H_i \rangle_{\rho(\theta)}, \quad (8.21)$$

where the fourth equality follows from Corollary 43. Now applying Proposition 47, we find that

$$\frac{\partial}{\partial \theta_i} e^{-H(\theta)} = - \int_0^1 dt e^{-tH(\theta)} \left(\frac{\partial}{\partial \theta_i} H(\theta) \right) e^{-(1-t)H(\theta)} \quad (8.22)$$

$$= - \int_0^1 dt e^{-tH(\theta)} H_i e^{-(1-t)H(\theta)}, \quad (8.23)$$

which implies from (8.21) that

$$\frac{\partial}{\partial \theta_i} \rho(\theta) = - \frac{1}{Z(\theta)} \int_0^1 dt e^{-tH(\theta)} H_i e^{-(1-t)H(\theta)} + \rho(\theta) \langle H_i \rangle_{\rho(\theta)} \quad (8.24)$$

$$= - \int_0^1 dt \left(\frac{e^{-H(\theta)}}{Z(\theta)} \right)^t H_i \left(\frac{e^{-H(\theta)}}{Z(\theta)} \right)^{1-t} + \rho(\theta) \langle H_i \rangle_{\rho(\theta)} \quad (8.25)$$

$$= - \int_0^1 dt \rho(\theta)^t H_i \rho(\theta)^{1-t} + \rho(\theta) \langle H_i \rangle_{\rho(\theta)}. \quad (8.26)$$

Then it follows that

$$[I_\zeta(\theta)]_{i,j} = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr}[\Pi_k (\partial_i \rho(\theta)) \Pi_\ell (\partial_j \rho(\theta))] \quad (8.27)$$

$$\begin{aligned} &= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr} \left[\Pi_k \left(- \int_0^1 dt \rho(\theta)^t H_i \rho(\theta)^{1-t} \right) \Pi_\ell \left(- \int_0^1 ds \rho(\theta)^s H_j \rho(\theta)^{1-s} \right) \right] \\ &\quad + \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr} \left[\Pi_k \left(- \int_0^1 dt \rho(\theta)^t H_i \rho(\theta)^{1-t} \right) \Pi_\ell \left(\rho(\theta) \langle H_j \rangle_{\rho(\theta)} \right) \right] \\ &\quad + \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr} \left[\Pi_k \left(\rho(\theta) \langle H_i \rangle_{\rho(\theta)} \right) \Pi_\ell \left(- \int_0^1 ds \rho(\theta)^s H_j \rho(\theta)^{1-s} \right) \right] \\ &\quad + \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr} \left[\Pi_k \left(\rho(\theta) \langle H_i \rangle_{\rho(\theta)} \right) \Pi_\ell \left(\rho(\theta) \langle H_j \rangle_{\rho(\theta)} \right) \right] \end{aligned} \quad (8.28)$$

$$\begin{aligned} &= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \int_0^1 \int_0^1 dt ds \text{Tr} [\Pi_k \rho(\theta)^t H_i \rho(\theta)^{1-t} \Pi_\ell \rho(\theta)^s H_j \rho(\theta)^{1-s}] \\ &\quad - \langle H_j \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \int_0^1 dt \text{Tr} [\Pi_k \rho(\theta)^t H_i \rho(\theta)^{1-t} \Pi_\ell \rho(\theta)] \\ &\quad - \langle H_i \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \int_0^1 ds \text{Tr} [\Pi_k \rho(\theta) \Pi_\ell \rho(\theta)^s H_j \rho(\theta)^{1-s}] \\ &\quad + \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr} [\Pi_k \rho(\theta) \Pi_\ell \rho(\theta)] \end{aligned} \quad (8.29)$$

$$= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \int_0^1 \int_0^1 dt ds \text{Tr} [\Pi_k \lambda_k^{1-s+t} H_i \lambda_\ell^{1-t+s} \Pi_\ell H_j]$$

$$\begin{aligned}
& - \langle H_j \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \int_0^1 dt \operatorname{Tr} [\Pi_k \lambda_k^t H_i \Pi_\ell \lambda_\ell^{2-t}] \\
& - \langle H_i \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \int_0^1 ds \operatorname{Tr} [\Pi_k \lambda_k^{2-s} \Pi_\ell \lambda_\ell^s H_j] \\
& + \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \operatorname{Tr} [\Pi_k \Pi_\ell \lambda_\ell^2]
\end{aligned} \tag{8.30}$$

$$\begin{aligned}
& = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \lambda_k \lambda_\ell \int_0^1 \int_0^1 dt ds \left(\frac{\lambda_k}{\lambda_\ell} \right)^{t-s} \operatorname{Tr} [\Pi_k H_i \Pi_\ell H_j] \\
& - \langle H_j \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \int_0^1 dt \lambda_k^t \lambda_\ell^{2-t} \operatorname{Tr} [\Pi_\ell \Pi_k H_i] \\
& - \langle H_i \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \int_0^1 ds \lambda_k^{2-s} \lambda_\ell^s \operatorname{Tr} [\Pi_k \Pi_\ell H_j] \\
& + \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)} \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \lambda_\ell^2 \operatorname{Tr} [\Pi_k \Pi_\ell]
\end{aligned} \tag{8.31}$$

$$\begin{aligned}
& = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \lambda_k \lambda_\ell \int_0^1 \int_0^1 dt ds \left(\frac{\lambda_k}{\lambda_\ell} \right)^{t-s} \operatorname{Tr} [\Pi_k H_i \Pi_\ell H_j] \\
& - \langle H_j \rangle_{\rho(\theta)} \sum_k \zeta(\lambda_k, \lambda_k) \int_0^1 dt \lambda_k^t \lambda_k^{2-t} \operatorname{Tr} [\Pi_k H_i] \\
& - \langle H_i \rangle_{\rho(\theta)} \sum_k \zeta(\lambda_k, \lambda_k) \int_0^1 ds \lambda_k^{2-s} \lambda_k^s \operatorname{Tr} [\Pi_k H_j] \\
& + \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)} \sum_k \zeta(\lambda_k, \lambda_k) \lambda_k^2 \operatorname{Tr} [\Pi_k]
\end{aligned} \tag{8.32}$$

$$\begin{aligned}
& = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \lambda_k \lambda_\ell \int_0^1 \int_0^1 dt ds \left(\frac{\lambda_k}{\lambda_\ell} \right)^{t-s} \operatorname{Tr} [\Pi_k H_i \Pi_\ell H_j] \\
& - \langle H_j \rangle_{\rho(\theta)} \sum_k \frac{\kappa}{\lambda_k} \lambda_k^2 \operatorname{Tr} [\Pi_k H_i] \\
& - \langle H_i \rangle_{\rho(\theta)} \sum_k \frac{\kappa}{\lambda_k} \lambda_k^2 \operatorname{Tr} [\Pi_k H_j] \\
& + \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)} \sum_k \frac{\kappa}{\lambda_k} \lambda_k^2 \operatorname{Tr} [\Pi_k]
\end{aligned} \tag{8.33}$$

$$\begin{aligned}
& = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \lambda_k \lambda_\ell \int_0^1 \int_0^1 dt ds \left(\frac{\lambda_k}{\lambda_\ell} \right)^{t-s} \operatorname{Tr} [\Pi_k H_i \Pi_\ell H_j] \\
& - \kappa \langle H_j \rangle_{\rho(\theta)} \operatorname{Tr} \left[\sum_k \lambda_k \Pi_k H_i \right] - \kappa \langle H_i \rangle_{\rho(\theta)} \operatorname{Tr} \left[\sum_k \lambda_k \Pi_k H_j \right]
\end{aligned}$$

$$+ \kappa \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)} \text{Tr} \left[\sum_k \lambda_k \Pi_k \right] \quad (8.34)$$

$$= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \lambda_k \lambda_\ell \int_0^1 \int_0^1 dt ds \left(\frac{\lambda_k}{\lambda_\ell} \right)^{t-s} \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \\ - \kappa \langle H_j \rangle_{\rho(\theta)} \langle H_i \rangle_{\rho(\theta)} - \kappa \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)} + \kappa \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)} \quad (8.35)$$

$$= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \lambda_k \lambda_\ell \int_0^1 \int_0^1 dt ds \left(\frac{\lambda_k}{\lambda_\ell} \right)^{t-s} \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \\ - \kappa \langle H_i \rangle_{\rho(\theta)} \langle H_j \rangle_{\rho(\theta)}. \quad (8.36)$$

Thus, it remains to evaluate the first term in the last expression above. To this end, considering that $\rho(\theta)$ and $H(\theta)$ commute, they have the same eigenprojections, so that the spectral decomposition of $H(\theta)$ is given by $\sum_k \mu_k \Pi_k$ and that of $\rho(\theta)$ is given by $\rho(\theta) = \sum_k \lambda_k \Pi_k = \sum_k \frac{e^{-\mu_k}}{Z} \Pi_k$, where $Z \equiv Z(\theta)$. Then consider that

$$\sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \lambda_k \lambda_\ell \int_0^1 \int_0^1 dt ds \left(\frac{\lambda_k}{\lambda_\ell} \right)^{t-s} \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \\ = \sum_{k,\ell} \zeta\left(\frac{e^{-\mu_k}}{Z}, \frac{e^{-\mu_\ell}}{Z}\right) \frac{e^{-\mu_k}}{Z} \frac{e^{-\mu_\ell}}{Z} \int_0^1 \int_0^1 dt ds \left(e^{-(\mu_k - \mu_\ell)} \right)^{t-s} \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \quad (8.37)$$

$$= \sum_{k,\ell} \zeta(e^{-\mu_k}, e^{-\mu_\ell}) e^{-\mu_k} \frac{e^{-\mu_\ell}}{Z} \int_0^1 \int_0^1 dt ds e^{-t(\mu_k - \mu_\ell)} e^{s(\mu_k - \mu_\ell)} \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \quad (8.38)$$

$$= \sum_{k,\ell} \zeta(e^{-(\mu_k - \mu_\ell)}, 1) \frac{e^{-\mu_k}}{Z} \left(\int_0^1 dt e^{-t(\mu_k - \mu_\ell)} \right) \left(\int_0^1 ds e^{s(\mu_k - \mu_\ell)} \right) \text{Tr}[\Pi_k H_i \Pi_\ell H_j]. \quad (8.39)$$

Observing that

$$\int_0^1 dt e^{-t(\mu_k - \mu_\ell)} = \frac{e^{-t(\mu_k - \mu_\ell)}}{-(\mu_k - \mu_\ell)} \Big|_0^1 \quad (8.40)$$

$$= \frac{e^{-(\mu_k - \mu_\ell)}}{-(\mu_k - \mu_\ell)} - \frac{1}{-(\mu_k - \mu_\ell)} \quad (8.41)$$

$$= \frac{1 - e^{-(\mu_k - \mu_\ell)}}{\mu_k - \mu_\ell}, \quad (8.42)$$

$$\int_0^1 ds e^{s(\mu_k - \mu_\ell)} = \frac{e^{s(\mu_k - \mu_\ell)}}{\mu_k - \mu_\ell} \Big|_0^1 \quad (8.43)$$

$$= \frac{e^{\mu_k - \mu_\ell}}{\mu_k - \mu_\ell} - \frac{1}{\mu_k - \mu_\ell} \quad (8.44)$$

$$= \frac{e^{\mu_k - \mu_\ell} - 1}{\mu_k - \mu_\ell}, \quad (8.45)$$

we conclude that

$$\begin{aligned} & \sum_{k,\ell} \zeta(e^{-(\mu_k-\mu_\ell)}, 1) \frac{e^{-\mu_k}}{Z} \left(\int_0^1 dt e^{-t(\mu_k-\mu_\ell)} \right) \left(\int_0^1 ds e^{s(\mu_k-\mu_\ell)} \right) \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \\ &= \sum_{k,\ell} \zeta(e^{-(\mu_k-\mu_\ell)}, 1) \frac{e^{-\mu_k}}{Z} \left(\frac{1 - e^{-(\mu_k-\mu_\ell)}}{\mu_k - \mu_\ell} \right) \left(\frac{e^{\mu_k-\mu_\ell} - 1}{\mu_k - \mu_\ell} \right) \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \end{aligned} \quad (8.46)$$

$$= \sum_{k,\ell} \zeta(e^{-(\mu_k-\mu_\ell)}, 1) \frac{e^{-\mu_k}}{Z} e^{\mu_k-\mu_\ell} \left(\frac{1 - e^{-(\mu_k-\mu_\ell)}}{\mu_k - \mu_\ell} \right) \left(\frac{1 - e^{-(\mu_k-\mu_\ell)}}{\mu_k - \mu_\ell} \right) \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \quad (8.47)$$

$$= \sum_{k,\ell} \zeta(e^{-(\mu_k-\mu_\ell)}, 1) \frac{e^{-\mu_\ell}}{Z} \left(\frac{1 - e^{-(\mu_k-\mu_\ell)}}{\mu_k - \mu_\ell} \right)^2 \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \quad (8.48)$$

$$= \sum_{k,\ell} \zeta(e^{-(\mu_k-\mu_\ell)}, 1) \frac{e^{-\mu_\ell}}{Z} \left(\frac{e^{-(\mu_k-\mu_\ell)} - 1}{\mu_k - \mu_\ell} \right)^2 \left(\frac{e^{-(\mu_k-\mu_\ell)} + 1}{e^{-(\mu_k-\mu_\ell)} + 1} \right) \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \quad (8.49)$$

$$= \frac{1}{2} \sum_{k,\ell} 2 \zeta(e^{-(\mu_k-\mu_\ell)}, 1) \left(\frac{e^{-(\mu_k-\mu_\ell)} - 1}{\mu_k - \mu_\ell} \right)^2 \left(\frac{1}{e^{-(\mu_k-\mu_\ell)} + 1} \right) \left(\frac{e^{-\mu_k}}{Z} + \frac{e^{-\mu_\ell}}{Z} \right) \text{Tr}[\Pi_k H_i \Pi_\ell H_j]. \quad (8.50)$$

Now suppose that $f(t)$ is a function satisfying the following Fourier transform relation:

$$\int_{-\infty}^{\infty} dt f(t) e^{it\omega} = 2 \zeta(e^{-\omega}, 1) \frac{(e^{-\omega} - 1)^2}{\omega^2 (e^{-\omega} + 1)}. \quad (8.51)$$

To establish that $f(t)$ is real-valued, it suffices to prove that $\omega \mapsto 2 \zeta(e^{-\omega}, 1) \frac{(e^{-\omega}-1)^2}{\omega^2(e^{-\omega}+1)}$ is an even function. To this end, consider that

$$-\omega \mapsto 2 \zeta(e^{-(-\omega)}, 1) \frac{(e^{-(-\omega)} - 1)^2}{(-\omega)^2 (e^{-(-\omega)} + 1)} \quad (8.52)$$

$$= 2 \zeta(e^\omega, 1) \frac{(e^\omega - 1)^2}{\omega^2 (e^\omega + 1)} \quad (8.53)$$

$$= 2 \zeta(1, e^\omega) \frac{(e^\omega - 1)^2}{\omega^2 (e^\omega + 1)} \quad (8.54)$$

$$= 2 e^{-\omega} \zeta(e^{-\omega}, 1) \frac{(e^\omega - 1)^2}{\omega^2 (e^\omega + 1)} \quad (8.55)$$

$$= 2 \zeta(e^{-\omega}, 1) \frac{e^{-\omega} (e^\omega - 1) e^{-\omega} (e^\omega - 1)}{e^{-\omega} \omega^2 (e^\omega + 1)} \quad (8.56)$$

$$= 2 \zeta(e^{-\omega}, 1) \frac{(1 - e^{-\omega}) (1 - e^{-\omega})}{\omega^2 (1 + e^{-\omega})} \quad (8.57)$$

$$= 2 \zeta(e^{-\omega}, 1) \frac{(1 - e^{-\omega})^2}{\omega^2 (e^{-\omega} + 1)} \quad (8.58)$$

$$= 2\zeta(e^{-\omega}, 1) \frac{(e^{-\omega} - 1)^2}{\omega^2 (e^{-\omega} + 1)}, \quad (8.59)$$

thus establishing the claim. Then it follows that

$$\begin{aligned} & \frac{1}{2} \sum_{k,\ell} 2\zeta(e^{-(\mu_k - \mu_\ell)}, 1) \left(\frac{e^{-(\mu_k - \mu_\ell)} - 1}{\mu_k - \mu_\ell} \right)^2 \left(\frac{1}{e^{-(\mu_k - \mu_\ell)} + 1} \right) \left(\frac{e^{-\mu_k}}{Z} + \frac{e^{-\mu_\ell}}{Z} \right) \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \\ &= \frac{1}{2} \sum_{k,\ell} \left(\int_{-\infty}^{\infty} dt f(t) e^{it(\mu_k - \mu_\ell)} \right) \left(\frac{e^{-\mu_k}}{Z} + \frac{e^{-\mu_\ell}}{Z} \right) \text{Tr}[\Pi_k H_i \Pi_\ell H_j] \end{aligned} \quad (8.60)$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} dt f(t) \sum_{k,\ell} \left(\frac{e^{-\mu_k}}{Z} + \frac{e^{-\mu_\ell}}{Z} \right) \text{Tr}[\Pi_k e^{it\mu_k} H_i \Pi_\ell e^{-it\mu_\ell} H_j] \quad (8.61)$$

$$\begin{aligned} &= \frac{1}{2} \int_{-\infty}^{\infty} dt f(t) \sum_{k,\ell} \text{Tr} \left[\Pi_k \frac{e^{-\mu_k}}{Z} e^{it\mu_k} H_i \Pi_\ell e^{-it\mu_\ell} H_j \right] \\ &\quad + \frac{1}{2} \int_{-\infty}^{\infty} dt f(t) \sum_{k,\ell} \text{Tr} \left[\Pi_k e^{it\mu_k} H_i \Pi_\ell \frac{e^{-\mu_\ell}}{Z} e^{-it\mu_\ell} H_j \right] \end{aligned} \quad (8.62)$$

$$\begin{aligned} &= \frac{1}{2} \int_{-\infty}^{\infty} dt f(t) \text{Tr} \left[\left(\sum_k \Pi_k \frac{e^{-\mu_k}}{Z} e^{it\mu_k} \right) H_i \left(\sum_\ell \Pi_\ell e^{-it\mu_\ell} \right) H_j \right] \\ &\quad + \frac{1}{2} \int_{-\infty}^{\infty} dt f(t) \text{Tr} \left[\left(\sum_k \Pi_k e^{it\mu_k} \right) H_i \left(\sum_\ell \Pi_\ell \frac{e^{-\mu_\ell}}{Z} e^{-it\mu_\ell} \right) H_j \right] \end{aligned} \quad (8.63)$$

$$\begin{aligned} &= \frac{1}{2} \int_{-\infty}^{\infty} dt f(t) \text{Tr}[\rho(\theta) e^{itH(\theta)} H_i e^{-itH(\theta)} H_j] \\ &\quad + \frac{1}{2} \int_{-\infty}^{\infty} dt f(t) \text{Tr}[e^{itH(\theta)} H_i e^{-itH(\theta)} \rho(\theta) H_j] \end{aligned} \quad (8.64)$$

$$= \frac{1}{2} \text{Tr}[\Phi_{f,\theta}(H_i) H_j \rho(\theta)] + \text{Tr}[H_j \Phi_{f,\theta}(H_i) \rho(\theta)] \quad (8.65)$$

$$= \frac{1}{2} \langle \{ \Phi_{f,\theta}(H_i), H_j \} \rangle_{\rho(\theta)}. \quad (8.66)$$

Finally, the claim in (8.14) follows because

$$\int_{-\infty}^{\infty} dt f(t) = \lim_{\omega \rightarrow 0} 2\zeta(e^{-\omega}, 1) \frac{(e^{-\omega} - 1)^2}{\omega^2 (e^{-\omega} + 1)} \quad (8.67)$$

$$= 2\zeta(1, 1) \left(\lim_{\omega \rightarrow 0} \frac{e^{-\omega} - 1}{\omega} \right)^2 \frac{1}{2} \quad (8.68)$$

$$= \kappa \left(\frac{d}{d\omega} e^{-\omega} \Big|_{\omega=0} \right)^2 \quad (8.69)$$

$$= \kappa, \quad (8.70)$$

where the third equality follows from (8.4). $\square$

8.2 α - z High-peak tent probability densities

In this subsection, I prove Lemma 34, which is the second ingredient needed for the proof of Theorem 32. The probability density in (8.73) is known as the high-peak tent, a name given in [PKPW24], due to its form when plotted (see [PKPW24, Figure 3]). It was shown in [EO19, Appendix C], [AAKS21, Section 5.1, Supplementary Information], and [PKPW24, Lemma 12] that the function $p(t)$ is the Fourier transform of $\tanh(\frac{\omega}{2}) / (\frac{\omega}{2})$, the latter function having been considered in the context of quantum belief propagation [Has07, Kim12, KB19].

The family of probability densities in (8.74)–(8.75), parameterized by $\alpha \in (0, 1)$ and $z > 0$, have a similar form to $p(t)$ when plotted, and so let us refer to them as the α - z high-peak tent probability densities.

Lemma 34. *For all $\alpha \in (0, 1)$ and $z > 0$, the Fourier transform of the function*

$$\omega \mapsto 2 \zeta_{\alpha,z}(e^{-\omega}, 1) \frac{(e^{-\omega} - 1)^2}{\omega^2 (e^{-\omega} + 1)}, \quad (8.71)$$

where $\zeta_{\alpha,z}(e^{-\omega}, 1)$ is defined in (5.6), is equal to the following probability density function:

$$q_{\alpha,z}(t) := (p * p_{\alpha,z})(t) = \int_{-\infty}^{\infty} d\tau p(\tau) p_{\alpha,z}(t - \tau), \quad (8.72)$$

where the probability density functions p and $p_{\alpha,z}$ are defined on $t \in \mathbb{R}$ as

$$p(t) := \frac{2}{\pi} \ln \left| \coth \left(\frac{\pi t}{2} \right) \right|, \quad (8.73)$$

$$p_{\alpha,z}(t) := \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi z t)} \right)^2 \right) \quad (8.74)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\coth^2(\pi z t) - \left(\frac{\cos(\pi\alpha)}{\sinh(\pi z t)} \right)^2 \right). \quad (8.75)$$

Additionally,

1. The identity $p_{\alpha=\frac{1}{2}, z=\frac{1}{2}}(t) = p(t)$ holds for all $t \in \mathbb{R}$.
2. For all $z > 0$, the characteristic function of $p_{\alpha,z}(t)$ converges pointwise to one everywhere in the limit $\alpha \rightarrow 1$, implying that, for all $z > 0$, the random variable $X_{\alpha,z} \sim q_{\alpha,z}$ converges in distribution to the random variable $X \sim p$ in the limit $\alpha \rightarrow 1$.
3. Similarly, for all $\alpha \in (0, 1)$, the characteristic function of $p_{\alpha,z}(t)$ converges pointwise to one everywhere in the limit $z \rightarrow \infty$, implying that, for all $\alpha \in (0, 1)$, the random variable $X_{\alpha,z} \sim q_{\alpha,z}$ converges in distribution to the random variable $X \sim p$ in the limit $z \rightarrow \infty$.

Proof. Recall from (5.6) that

$$\zeta_{\alpha,z}(e^{-\omega}, 1) = \frac{z}{\alpha(1-\alpha)} \left(\frac{e^{-\left(\frac{1-\alpha}{z}\right)\omega} - 1}{e^{-\omega} - 1} \right) \left(\frac{e^{-\left(\frac{\alpha}{z}\right)\omega} - 1}{e^{-\left(\frac{1}{z}\right)\omega} - 1} \right). \quad (8.76)$$

After defining

$$f_{\alpha,z}(\omega) := \frac{z \left(1 - e^{-\left(\frac{1-\alpha}{z}\right)\omega} \right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)\omega} \right)}{\alpha(1-\alpha)\omega \left(1 - e^{-\left(\frac{1}{z}\right)\omega} \right)}, \quad (8.77)$$

it follows that

$$2\zeta_{\alpha,z}(e^{-\omega}, 1) \frac{(e^{-\omega} - 1)^2}{\omega^2(e^{-\omega} + 1)} = \frac{2z \left(e^{-\left(\frac{1-\alpha}{z}\right)\omega} - 1 \right) \left(e^{-\left(\frac{\alpha}{z}\right)\omega} - 1 \right) (e^{-\omega} - 1)}{\alpha(1-\alpha) \left(e^{-\left(\frac{1}{z}\right)\omega} - 1 \right) \omega^2 (e^{-\omega} + 1)} \quad (8.78)$$

$$= \frac{z \left(1 - e^{-\left(\frac{1-\alpha}{z}\right)\omega} \right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)\omega} \right)}{\alpha(1-\alpha)\omega \left(1 - e^{-\left(\frac{1}{z}\right)\omega} \right)} \frac{(1 - e^{-\omega})}{\frac{\omega}{2}(1 + e^{-\omega})} \quad (8.79)$$

$$= f_{\alpha,z}(\omega) \left(\frac{1 - e^{-\omega}}{\frac{\omega}{2}(1 + e^{-\omega})} \right) \quad (8.80)$$

$$= f_{\alpha,z}(\omega) \left(\frac{e^{\frac{\omega}{2}} - e^{-\frac{\omega}{2}}}{\frac{\omega}{2}(e^{\frac{\omega}{2}} + e^{-\frac{\omega}{2}})} \right) \quad (8.81)$$

$$= f_{\alpha,z}(\omega) \frac{\tanh\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}}. \quad (8.82)$$

Our goal is to evaluate the following expression for all $t \in \mathbb{R}$:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega f_{\alpha,z}(\omega) \frac{\tanh\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}} e^{i\omega t}, \quad (8.83)$$

and by the convolution theorem [Wei], this is equal to the convolution of the following two Fourier transforms:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega f_{\alpha,z}(\omega) e^{i\omega t}, \quad (8.84)$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \frac{\tanh\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}} e^{i\omega t} = \frac{2}{\pi} \ln \left| \coth \left(\frac{\pi t}{2} \right) \right|, \quad (8.85)$$

where (8.85) follows from [PKPW24, Lemma 12] and the function $t \mapsto \frac{2}{\pi} \ln \left| \coth \left(\frac{\pi t}{2} \right) \right|$ is a probability density function on $\mathbb{R}$ known as the high-peak tent [PKPW24]. The equality in (8.85) was calculated in [EO19, Appendix C] and [AAKS21, Section 5.1, Supplementary

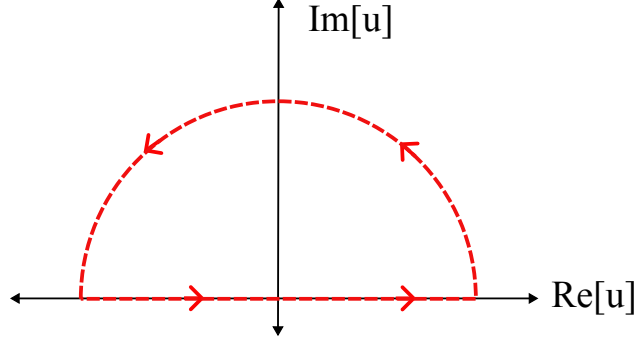


Figure 1: Contour γ_R^+ for $t > 0$.

Information] by means of contour integration. By setting $\alpha = z = \frac{1}{2}$, the following equality holds:

$$f_{\alpha=\frac{1}{2}, z=\frac{1}{2}}(\omega) = \frac{\tanh\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}}, \quad (8.86)$$

which implies that the Fourier transform in (8.85) is a special case of that in (8.84). Indeed, consider that

$$f_{\alpha=\frac{1}{2}, z=\frac{1}{2}}(\omega) = \lim_{\alpha, z \rightarrow \frac{1}{2}} \frac{z \left(1 - e^{-\left(\frac{1-\alpha}{z}\right)\omega}\right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)\omega}\right)}{\alpha (1 - \alpha) \omega \left(1 - e^{-\left(\frac{1}{z}\right)\omega}\right)} \quad (8.87)$$

$$= \frac{(1 - e^{-\omega})(1 - e^{-\omega})}{\frac{\omega}{2} (1 - e^{-2\omega})} \quad (8.88)$$

$$= \frac{(1 - e^{-\omega})(1 - e^{-\omega})}{\frac{\omega}{2} (1 - e^{-\omega})(1 + e^{-\omega})} \quad (8.89)$$

$$= \frac{1 - e^{-\omega}}{\frac{\omega}{2} (1 + e^{-\omega})} \quad (8.90)$$

$$= \frac{\tanh\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}}. \quad (8.91)$$

In what follows, I prove that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega f_{\alpha, z}(\omega) e^{i\omega t} = \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi z t)} \right)^2 \right), \quad (8.92)$$

for all $\alpha \in (0, 1)$ and $z > 0$, and that

$$\lim_{\alpha, z \rightarrow \frac{1}{2}} \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi z t)} \right)^2 \right) = \frac{2}{\pi} \ln \left| \coth \left(\frac{\pi t}{2} \right) \right|. \quad (8.93)$$

Let us now evaluate the Fourier transform in (8.84) by employing contour integration and the Cauchy residue theorem. We begin by considering the case when $t > 0$, and consider the

positively oriented contour γ_R^+ depicted in Figure 1, where $R > 0$ denotes the radius of the semicircle depicted there. Consider that

$$\lim_{R \rightarrow \infty} \oint_{\gamma_R^+} du f_{\alpha,z}(u) e^{iut} = \lim_{R \rightarrow \infty} \left[\int_{-R}^R d\omega f_{\alpha,z}(\omega) e^{i\omega t} + \int_0^\pi d\theta f_{\alpha,z}(Re^{i\theta}) e^{iRe^{i\theta}t} \right], \quad (8.94)$$

such that the contour integral is broken up into the line integral along the real axis and the line integral around the arc of the semicircle. Let us now prove that the line integral around the arc of the semicircle evaluates to zero in the limit $R \rightarrow \infty$. To this end, consider that

$$\begin{aligned} & \left| \int_0^\pi d\theta f_{\alpha,z}(Re^{i\theta}) e^{iRe^{i\theta}t} \right| \\ &= \left| \int_0^\pi d\theta f_{\alpha,z}(Re^{i\theta}) e^{iR \cos(\theta)t} e^{-R \sin(\theta)t} \right| \end{aligned} \quad (8.95)$$

$$\leq \int_0^\pi d\theta |f_{\alpha,z}(Re^{i\theta}) e^{iR \cos(\theta)t} e^{-R \sin(\theta)t}| \quad (8.96)$$

$$= \int_0^\pi d\theta |f_{\alpha,z}(Re^{i\theta})| e^{-R \sin(\theta)t} \quad (8.97)$$

$$= \int_0^\pi d\theta \left| \frac{z \left(1 - e^{-\left(\frac{1-\alpha}{z}\right)Re^{i\theta}}\right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)Re^{i\theta}}\right)}{\alpha(1-\alpha) Re^{i\theta} \left(1 - e^{-\left(\frac{1}{z}\right)Re^{i\theta}}\right)} \right| e^{-R \sin(\theta)t} \quad (8.98)$$

$$\leq \frac{z}{\alpha|1-\alpha|R} \int_0^\pi d\theta \left| \frac{\left(1 - e^{-\left(\frac{1-\alpha}{z}\right)Re^{i\theta}}\right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)Re^{i\theta}}\right)}{\left(1 - e^{-\left(\frac{1}{z}\right)Re^{i\theta}}\right)} \right| \quad (8.99)$$

$$\leq \frac{z}{\alpha|1-\alpha|R} \int_0^\pi d\theta 2 \quad (8.100)$$

$$= \frac{2\pi z}{\alpha|1-\alpha|R} \quad (8.101)$$

Thus, we conclude that

$$\lim_{R \rightarrow \infty} \left| \int_0^\pi d\theta f_{\alpha,z}(Re^{i\theta}) e^{iRe^{i\theta}t} \right| \leq \lim_{R \rightarrow \infty} \frac{\pi z}{\alpha|1-\alpha|R}. \quad (8.102)$$

$$= 0. \quad (8.103)$$

To see (8.100), setting $u = x + iy$, consider that

$$\begin{aligned} & \left| \frac{\left(1 - e^{-\left(\frac{1-\alpha}{z}\right)u}\right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)u}\right)}{\left(1 - e^{-\left(\frac{1}{z}\right)u}\right)} \right| \\ &= \left| \frac{\left(1 - e^{-\left(\frac{1-\alpha}{z}\right)u}\right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)u}\right)}{\left(1 - e^{-\left(\frac{1}{z}\right)u}\right)} \right| \end{aligned} \quad (8.104)$$

$$= \left| \frac{e^{-\left(\frac{1-\alpha}{2z}\right)u} \left(e^{\left(\frac{1-\alpha}{2z}\right)u} - e^{-\left(\frac{1-\alpha}{2z}\right)u} \right) e^{-\left(\frac{\alpha}{2z}\right)u} \left(e^{\left(\frac{\alpha}{2z}\right)u} - e^{-\left(\frac{\alpha}{2z}\right)u} \right)}{e^{-\left(\frac{1}{2z}\right)u} \left(e^{\left(\frac{1}{2z}\right)u} - e^{-\left(\frac{1}{2z}\right)u} \right)} \right| \quad (8.105)$$

$$= \left| \frac{\left(e^{\left(\frac{1-\alpha}{2z}\right)u} - e^{-\left(\frac{1-\alpha}{2z}\right)u} \right) \left(e^{\left(\frac{\alpha}{2z}\right)u} - e^{-\left(\frac{\alpha}{2z}\right)u} \right)}{\left(e^{\left(\frac{1}{2z}\right)u} - e^{-\left(\frac{1}{2z}\right)u} \right)} \right| \quad (8.106)$$

$$= 2 \left| \frac{\sinh\left(\left(\frac{1-\alpha}{2z}\right)u\right) \sinh\left(\left(\frac{\alpha}{2z}\right)u\right)}{\sinh\left(\left(\frac{1}{2z}\right)u\right)} \right| \quad (8.107)$$

$$= 2 \frac{|\sinh\left(\left(\frac{1-\alpha}{2z}\right)u\right)| |\sinh\left(\left(\frac{\alpha}{2z}\right)u\right)|}{|\sinh\left(\left(\frac{1}{2z}\right)u\right)|} \quad (8.108)$$

$$= 2 \frac{|\sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right)| |\sinh\left(\left(\frac{\alpha}{2z}\right)x\right)| \sqrt{1 + \sin^2\left(\left(\frac{1-\alpha}{2z}\right)y\right)} \sqrt{1 + \sin^2\left(\left(\frac{\alpha}{2z}\right)y\right)}}{|\sinh\left(\left(\frac{1}{2z}\right)x\right)| \sqrt{1 + \sin^2\left(\left(\frac{1}{2z}\right)y\right)}} \leq 2, \quad (8.109)$$

where we used the following identities:

$$|\sinh(x + iy)| = |\sinh(x) \cos(y) + i \cosh(x) \sin(y)| \quad (8.110)$$

$$= \sqrt{\sinh^2(x) \cos^2(y) + \cosh^2(x) \sin^2(y)} \quad (8.111)$$

$$= \sqrt{\sinh^2(x) \cos^2(y) + (\sinh^2(x) + 1) \sin^2(y)} \quad (8.112)$$

$$= \sqrt{\sinh^2(x) (\cos^2(y) + \sin^2(y)) + \sinh^2(x) \sin^2(y)} \quad (8.113)$$

$$= \sqrt{\sinh^2(x) (1 + \sin^2(y))} \quad (8.114)$$

$$= |\sinh(x)| \sqrt{1 + \sin^2(y)}, \quad (8.115)$$

and the facts that, for all $x, y \in \mathbb{R}$, $\alpha \in (0, 1)$, and $z > 0$,

$$\frac{\sqrt{1 + \sin^2\left(\left(\frac{1-\alpha}{2z}\right)y\right)} \sqrt{1 + \sin^2\left(\left(\frac{\alpha}{2z}\right)y\right)}}{\sqrt{1 + \sin^2\left(\left(\frac{1}{2z}\right)y\right)}} \leq 2, \quad (8.116)$$

$$\frac{|\sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right)| |\sinh\left(\left(\frac{\alpha}{2z}\right)x\right)|}{|\sinh\left(\left(\frac{1}{2z}\right)x\right)|} \leq \frac{1}{2}. \quad (8.117)$$

The inequality in (8.117) follows because, for all $\alpha \in (0, 1)$ and $z > 0$,

$$\lim_{x \rightarrow \infty} \frac{\sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right) \sinh\left(\left(\frac{\alpha}{2z}\right)x\right)}{\sinh\left(\left(\frac{1}{2z}\right)x\right)} = \lim_{x \rightarrow \infty} \frac{\frac{e^{\left(\frac{1-\alpha}{2z}\right)x}}{2} \frac{e^{\left(\frac{\alpha}{2z}\right)x}}{2}}{\frac{e^{\left(\frac{1}{2z}\right)x}}{2}} = \frac{1}{2}, \quad (8.118)$$

$$\lim_{x \rightarrow -\infty} \frac{\sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right) \sinh\left(\left(\frac{\alpha}{2z}\right)x\right)}{\sinh\left(\left(\frac{1}{2z}\right)x\right)} = \lim_{x \rightarrow -\infty} \frac{\frac{-e^{-\left(\frac{1-\alpha}{2z}\right)x}}{2} \frac{e^{-\left(\frac{\alpha}{2z}\right)x}}{2}}{\frac{-e^{-\left(\frac{1}{2z}\right)x}}{2}} = -\frac{1}{2}, \quad (8.119)$$

$$\begin{aligned} & \lim_{x \rightarrow 0} \frac{\sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right) \sinh\left(\left(\frac{\alpha}{2z}\right)x\right)}{\sinh\left(\left(\frac{1}{2z}\right)x\right)} \\ &= \lim_{x \rightarrow 0} \frac{\left(\frac{1-\alpha}{2z}\right) \cosh\left(\left(\frac{1-\alpha}{2z}\right)x\right) \sinh\left(\left(\frac{\alpha}{2z}\right)x\right) + \sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right) \left(\frac{\alpha}{2z}\right) \cosh\left(\left(\frac{\alpha}{2z}\right)x\right)}{\left(\frac{1}{2z}\right) \cosh\left(\left(\frac{1}{2z}\right)x\right)} \end{aligned} \quad (8.120)$$

$$= 0, \quad (8.121)$$

and the fact that the function $x \mapsto \sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right) \sinh\left(\left(\frac{\alpha}{2z}\right)x\right) / \sinh\left(\left(\frac{1}{2z}\right)x\right)$ is increasing on $x \in \mathbb{R}$. Indeed, since this latter function is odd, to see that it is increasing on $x \in \mathbb{R}$, it suffices to prove that it is increasing on $(0, \infty)$. To see this, let us calculate the derivative of this function with respect to x and prove that it is non-negative on $(0, \infty)$. Consider that

$$\begin{aligned} & 2 \frac{\partial}{\partial x} \left(\frac{\sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right) \sinh\left(\left(\frac{\alpha}{2z}\right)x\right)}{\sinh\left(\left(\frac{1}{2z}\right)x\right)} \right) \\ &= \frac{\partial}{\partial x} \left(\frac{\left(e^{\left(\frac{1-\alpha}{2z}\right)x} - e^{-\left(\frac{1-\alpha}{2z}\right)x}\right) \left(e^{\frac{\alpha}{2z}x} - e^{-\frac{\alpha}{2z}x}\right)}{e^{\frac{1}{2z}x} - e^{-\frac{1}{2z}x}} \right) \end{aligned} \quad (8.122)$$

$$= \frac{\partial}{\partial x} \left(\frac{\left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right)}{e^{\frac{1}{z}x} - 1} \right) \quad (8.123)$$

$$= \frac{\frac{1-\alpha}{z} e^{\left(\frac{1-\alpha}{z}\right)x} \left(e^{\frac{\alpha}{z}x} - 1\right) + \frac{\alpha}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) e^{\frac{\alpha}{z}x}}{e^{\frac{1}{z}x} - 1} - \frac{\frac{1}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right) e^{\frac{1}{z}x}}{\left(e^{\frac{1}{z}x} - 1\right)^2} \quad (8.124)$$

$$\begin{aligned} &= \frac{\frac{1-\alpha}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1 + 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right) + \frac{\alpha}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1 + 1\right)}{e^{\frac{1}{z}x} - 1} \\ &\quad - \frac{\frac{1}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right) e^{\frac{1}{z}x}}{\left(e^{\frac{1}{z}x} - 1\right)^2} \end{aligned} \quad (8.125)$$

$$\begin{aligned} &= \frac{\left[\frac{1-\alpha}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right) + \frac{\alpha}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right) \right]}{e^{\frac{1}{z}x} - 1} \\ &\quad - \frac{\frac{1}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right) e^{\frac{1}{z}x}}{\left(e^{\frac{1}{z}x} - 1\right)^2} \end{aligned} \quad (8.126)$$

$$\begin{aligned}
&= \frac{\left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right)}{\left(e^{\frac{1}{z}x} - 1\right)} \left(\frac{1-\alpha}{z} + \frac{\alpha}{z} - \frac{e^{\frac{1}{z}x}}{z\left(e^{\frac{1}{z}x} - 1\right)}\right) \\
&\quad + \frac{\frac{1-\alpha}{z} \left(e^{\frac{\alpha}{z}x} - 1\right) + \frac{\alpha}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right)}{e^{\frac{1}{z}x} - 1} \tag{8.127}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right)}{\left(e^{\frac{1}{z}x} - 1\right)} \frac{1}{z} \left(1 - \frac{e^{\frac{1}{z}x}}{e^{\frac{1}{z}x} - 1}\right) + \frac{\frac{1-\alpha}{z} \left(e^{\frac{\alpha}{z}x} - 1\right) + \frac{\alpha}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right)}{e^{\frac{1}{z}x} - 1} \tag{8.128}
\end{aligned}$$

$$\begin{aligned}
&= -\frac{\left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right)}{\left(e^{\frac{1}{z}x} - 1\right)} \left(\frac{1}{z\left(e^{\frac{1}{z}x} - 1\right)}\right) + \frac{\frac{1-\alpha}{z} \left(e^{\frac{\alpha}{z}x} - 1\right)}{e^{\frac{1}{z}x} - 1} + \frac{\frac{\alpha}{z} \left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right)}{e^{\frac{1}{z}x} - 1} \tag{8.129}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right)}{z\left(e^{\frac{1}{z}x} - 1\right)} \left(\frac{1-\alpha}{e^{\left(\frac{1-\alpha}{z}\right)x} - 1} + \frac{\alpha}{e^{\frac{\alpha}{z}x} - 1} - \frac{1}{e^{\frac{1}{z}x} - 1}\right). \tag{8.130}
\end{aligned}$$

Since $\frac{\left(e^{\left(\frac{1-\alpha}{z}\right)x} - 1\right) \left(e^{\frac{\alpha}{z}x} - 1\right)}{z\left(e^{\frac{1}{z}x} - 1\right)} > 0$ for all $x > 0$, the desired statement will follow if the inequality

$$\frac{1-\alpha}{e^{\left(\frac{1-\alpha}{z}\right)x} - 1} + \frac{\alpha}{e^{\frac{\alpha}{z}x} - 1} \geq \frac{1}{e^{\frac{1}{z}x} - 1} \tag{8.131}$$

holds for all $\alpha \in (0, 1)$ and $x > 0$. Thus, the statement follows if the function $x \mapsto \frac{1}{e^{\frac{1}{z}x} - 1}$ is convex and decreasing on $x \in (0, \infty)$. Indeed, if these properties hold, then

$$\frac{1-\alpha}{e^{\left(\frac{1-\alpha}{z}\right)x} - 1} + \frac{\alpha}{e^{\frac{\alpha}{z}x} - 1} \geq \frac{1}{e^{\frac{1}{z}((1-\alpha)^2 + \alpha^2)x} - 1} \geq \frac{1}{e^{\frac{1}{z}x} - 1}, \tag{8.132}$$

where the first inequality follows from convexity and the second from the decreasing property because $(1-\alpha)^2 + \alpha^2 < 1$ for $\alpha \in (0, 1)$. To prove these properties, let us calculate the first and second derivatives of the function $x \mapsto \frac{1}{e^{\frac{1}{z}x} - 1}$:

$$\frac{\partial}{\partial x} \left(\frac{1}{e^{\frac{1}{z}x} - 1} \right) = -\frac{e^{\frac{x}{z}}}{z \left(e^{\frac{1}{z}x} - 1 \right)^2}, \tag{8.133}$$

$$\frac{\partial^2}{\partial x^2} \left(\frac{1}{e^{\frac{1}{z}x} - 1} \right) = \frac{e^{\frac{x}{z}} \left(e^{\frac{x}{z}} + 1 \right)}{z^2 \left(e^{\frac{1}{z}x} - 1 \right)^3}. \tag{8.134}$$

Since $\frac{\partial}{\partial x} \left(\frac{1}{e^{\frac{1}{z}x} - 1} \right) \leq 0$ and $\frac{\partial^2}{\partial x^2} \left(\frac{1}{e^{\frac{1}{z}x} - 1} \right) \geq 0$ on $x \in (0, \infty)$, the claim follows. Thus, it follows that

$$\lim_{x \rightarrow \pm\infty} \frac{|\sinh\left(\left(\frac{1-\alpha}{2z}\right)x\right)| |\sinh\left(\left(\frac{\alpha}{2z}\right)x\right)|}{\left|\sinh\left(\left(\frac{1}{2z}\right)x\right)\right|} = \frac{1}{2}, \tag{8.135}$$

$$\lim_{x \rightarrow 0} \frac{|\sinh((\frac{1-\alpha}{2z})x)| |\sinh((\frac{\alpha}{2z})x)|}{|\sinh((\frac{1}{2z})x)|} = 0, \quad (8.136)$$

and that the function $x \mapsto \frac{|\sinh((\frac{1-\alpha}{2z})x)| |\sinh((\frac{\alpha}{2z})x)|}{|\sinh((\frac{1}{2z})x)|}$ is decreasing on $(-\infty, 0]$ and increasing on $[0, \infty)$. We then conclude that

$$\frac{1}{2\pi} \lim_{R \rightarrow \infty} \oint_{\gamma_R^+} du f_{\alpha,z}(u) e^{iut} = \frac{1}{2\pi} \lim_{R \rightarrow \infty} \int_{-R}^R d\omega f_{\alpha,z}(\omega) e^{i\omega t} \quad (8.137)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega f_{\alpha,z}(\omega) e^{i\omega t}. \quad (8.138)$$

We can evaluate the expression on the left-hand side of (8.137) by means of the Cauchy residue theorem. For $u \in \mathbb{C}$ and $\alpha \in (0, 1)$, the singularities of the function

$$f_{\alpha,z}(u) e^{iut} = \frac{z \left(1 - e^{-(\frac{1-\alpha}{z})u}\right) \left(1 - e^{-(\frac{\alpha}{z})u}\right)}{\alpha(1-\alpha)u \left(1 - e^{-(\frac{1}{z})u}\right)} e^{iut} \quad (8.139)$$

in the region enclosed by the contour γ_R^+ , as $R \rightarrow \infty$, occur at $u \in \{2\pi i z m : m \in \mathbb{N}\}$. Indeed, there is not a singularity of $f_{\alpha,z}(u) e^{iut}$ at $u = 0$ because

$$\begin{aligned} & \lim_{u \rightarrow 0} f_{\alpha,z}(u) e^{iut} \\ &= \lim_{u \rightarrow 0} \frac{z \left(1 - e^{-(\frac{1-\alpha}{z})u}\right) \left(1 - e^{-(\frac{\alpha}{z})u}\right)}{\alpha(1-\alpha)u \left(1 - e^{-(\frac{1}{z})u}\right)} e^{iut} \end{aligned} \quad (8.140)$$

$$= \frac{z}{\alpha(1-\alpha)} \lim_{u \rightarrow 0} \frac{\left(1 - e^{-(\frac{1-\alpha}{z})u}\right) \left(1 - e^{-(\frac{\alpha}{z})u}\right)}{u \left(1 - e^{-(\frac{1}{z})u}\right)} \lim_{u \rightarrow 0} e^{iut} \quad (8.141)$$

$$= \frac{z}{\alpha(1-\alpha)} \lim_{u \rightarrow 0} \frac{\left(\frac{1-\alpha}{z}\right) e^{-(\frac{1-\alpha}{z})u} \left(1 - e^{-(\frac{\alpha}{z})u}\right) + \left(1 - e^{-(\frac{1-\alpha}{z})u}\right) \left(\frac{\alpha}{z}\right) e^{-(\frac{\alpha}{z})u}}{1 - e^{-(\frac{1}{z})u} + u \left(\frac{1}{z}\right) e^{-(\frac{1}{z})u}} \quad (8.142)$$

$$= \frac{z}{\alpha(1-\alpha)} \lim_{u \rightarrow 0} \frac{\left(\frac{1-\alpha}{z}\right) \left(e^{-(\frac{1-\alpha}{z})u} - e^{-(\frac{1}{z})u}\right) + \left(\frac{\alpha}{z}\right) \left(e^{-(\frac{\alpha}{z})u} - e^{-(\frac{1}{z})u}\right)}{1 - e^{-(\frac{1}{z})u} + u \left(\frac{1}{z}\right) e^{-(\frac{1}{z})u}} \quad (8.143)$$

$$= \frac{z}{\alpha(1-\alpha)} \lim_{u \rightarrow 0} \frac{\left(\frac{1-\alpha}{z}\right) e^{-(\frac{1-\alpha}{z})u} + \left(\frac{\alpha}{z}\right) e^{-(\frac{\alpha}{z})u} - \left(\frac{1}{z}\right) e^{-(\frac{1}{z})u}}{1 - e^{-(\frac{1}{z})u} + u \left(\frac{1}{z}\right) e^{-(\frac{1}{z})u}} \quad (8.144)$$

$$= \frac{z}{\alpha(1-\alpha)} \lim_{u \rightarrow 0} \frac{-\left(\frac{1-\alpha}{z}\right)^2 e^{-(\frac{1-\alpha}{z})u} - \left(\frac{\alpha}{z}\right)^2 e^{-(\frac{\alpha}{z})u} + \left(\frac{1}{z}\right)^2 e^{-(\frac{1}{z})u}}{\left(\frac{1}{z}\right) e^{-(\frac{1}{z})u} + \left(\frac{1}{z}\right) e^{-(\frac{1}{z})u} - u \left(\frac{1}{z}\right)^2 e^{-(\frac{1}{z})u}} \quad (8.145)$$

$$= \frac{z}{\alpha(1-\alpha)} \frac{-\left(\frac{1-\alpha}{z}\right)^2 - \left(\frac{\alpha}{z}\right)^2 + \left(\frac{1}{z}\right)^2}{\frac{1}{z} + \frac{1}{z}} \quad (8.146)$$

$$= \frac{-(1-\alpha)^2 - \alpha^2 + 1}{2\alpha(1-\alpha)} \quad (8.147)$$

$$= \frac{-(1-2\alpha+\alpha^2) - \alpha^2 + 1}{2\alpha(1-\alpha)} \quad (8.148)$$

$$= \frac{2\alpha - 2\alpha^2}{2\alpha(1-\alpha)} \quad (8.149)$$

$$= 1. \quad (8.150)$$

For all $m \in \mathbb{N}$, we then need to calculate the residues. Since the poles of the function $u \mapsto f_{\alpha,z}(u)e^{iut}$ at $2\pi izm$ are simple poles, we can use the formula $\text{Res}_{u=c} \left[\frac{P(z)}{Q(z)} \right] = \frac{P(c)}{Q'(c)}$ to conclude that

$$\text{Res}_{u=2\pi imz} [f_{\alpha,z}(u)e^{iut}] = \frac{z}{\alpha(1-\alpha)} \frac{\left(1 - e^{-\left(\frac{1-\alpha}{z}\right)u}\right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)u}\right) e^{iut}}{\frac{\partial}{\partial u} \left[u \left(1 - e^{-\left(\frac{1}{z}\right)u}\right) \right]} \Big|_{u=2\pi imz} \quad (8.151)$$

$$= \frac{z \left(1 - e^{-\left(\frac{1-\alpha}{z}\right)2\pi imz}\right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)2\pi imz}\right)}{\alpha(1-\alpha) \left(1 - e^{-\left(\frac{1}{z}\right)u} + \frac{u}{z} e^{-\left(\frac{1}{z}\right)u}\right) \Big|_{u=2\pi imz}} e^{i(2\pi imz)t} \quad (8.152)$$

$$= \frac{z \left(1 - e^{-(1-\alpha)2\pi im}\right) \left(1 - e^{-\alpha 2\pi im}\right)}{\alpha(1-\alpha) 2\pi im} e^{-2\pi mzt} \quad (8.153)$$

$$= \frac{z}{\alpha(1-\alpha)} \left(\frac{2 - e^{\alpha 2\pi im} - e^{-\alpha 2\pi im}}{2\pi im} \right) e^{-2\pi mzt}. \quad (8.154)$$

Thus, by applying the Cauchy residue theorem, it follows that

$$\begin{aligned} & \frac{1}{2\pi} \lim_{R \rightarrow \infty} \oint_{\gamma_R^+} du f_{\alpha,z}(u) e^{iut} \\ &= i \sum_{m=1}^{\infty} \text{Res}_{u=i2\pi mz} [f_{\alpha,z}(u) e^{iut}] \end{aligned} \quad (8.155)$$

$$= i \sum_{m=1}^{\infty} \frac{z}{\alpha(1-\alpha)} \left(\frac{2 - e^{\alpha 2\pi im} - e^{-\alpha 2\pi im}}{2\pi im} \right) e^{-2\pi mzt} \quad (8.156)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \sum_{m=1}^{\infty} \frac{2e^{-2\pi mzt}}{m} - \frac{(e^{2\pi(i\alpha-zt)})^m}{m} - \frac{(e^{2\pi(-i\alpha-zt)})^m}{m} \quad (8.157)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \left(-2 \ln(1 - e^{-2\pi zt}) + \ln(1 - e^{2\pi(i\alpha-zt)}) + \ln(1 - e^{2\pi(-i\alpha-zt)}) \right) \quad (8.158)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\frac{(1 - e^{2\pi(i\alpha-zt)}) (1 - e^{2\pi(-i\alpha-zt)})}{(1 - e^{-2\pi zt})^2} \right) \quad (8.159)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\frac{1 - e^{2\pi(i\alpha-zt)} - e^{2\pi(-i\alpha-zt)} + e^{-4\pi zt}}{(1 - e^{-2\pi zt})^2} \right) \quad (8.160)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\frac{1 + e^{-4\pi zt} - 2e^{-2\pi zt} \cos(2\pi\alpha)}{(1 - e^{-2\pi zt})^2} \right) \quad (8.161)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\frac{(1 - e^{-2\pi zt})^2 + 2e^{-2\pi zt} (1 - \cos(2\pi\alpha))}{(1 - e^{-2\pi zt})^2} \right) \quad (8.162)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \frac{2e^{-2\pi zt} (1 - \cos(2\pi\alpha))}{(1 - e^{-2\pi zt})^2} \right) \quad (8.163)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \frac{2(1 - \cos(2\pi\alpha))}{(e^{\pi zt} - e^{-\pi zt})^2} \right) \quad (8.164)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi zt)} \right)^2 \right). \quad (8.165)$$

where the fourth equality follows from the Taylor expansion $-\ln(1-u) = \sum_{m=1}^{\infty} \frac{u^m}{m}$, which holds for all $u \in \mathbb{C}$ such that $|u| < 1$. Thus, for $t > 0$, we conclude that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega f_{\alpha,z}(\omega) e^{i\omega t} = \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi zt)} \right)^2 \right). \quad (8.166)$$

We can handle the case when $t < 0$ by a similar, yet symmetric argument, instead considering the negatively oriented contour γ_R^- depicted in Figure 2. To summarize the argument succinctly, we again use the fact that

$$\lim_{R \rightarrow \infty} \oint_{\gamma_R^-} du f_{\alpha,z}(u) e^{iut} = \lim_{R \rightarrow \infty} \left[\int_{-R}^R d\omega f_{\alpha,z}(\omega) e^{i\omega t} + \int_0^{-\pi} d\theta f_{\alpha,z}(Re^{i\theta}) e^{iRe^{i\theta}t} \right], \quad (8.167)$$

and apply a similar argument as in (8.95)–(8.117) to conclude that

$$\lim_{R \rightarrow \infty} \left| \int_0^{-\pi} d\theta f_{\alpha,z}(Re^{i\theta}) e^{iRe^{i\theta}t} \right| = 0, \quad (8.168)$$

while keeping in mind that $\sin(\theta) \leq 0$ for all $\theta \in [-\pi, 0]$. Then, similar to (8.139)–(8.166), we apply the Cauchy residue theorem to conclude that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega f_{\alpha,z}(\omega) e^{i\omega t} = \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi zt)} \right)^2 \right) \quad (8.169)$$

for all $t < 0$.

The function $p_{\alpha,z}(t)$ is a probability density for all $\alpha \in (0, 1)$ and $z > 0$ because

$$p_{\alpha,z}(t) = \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi zt)} \right)^2 \right) \quad (8.170)$$

$$\geq \frac{z}{2\pi\alpha(1-\alpha)} \ln(1) \quad (8.171)$$

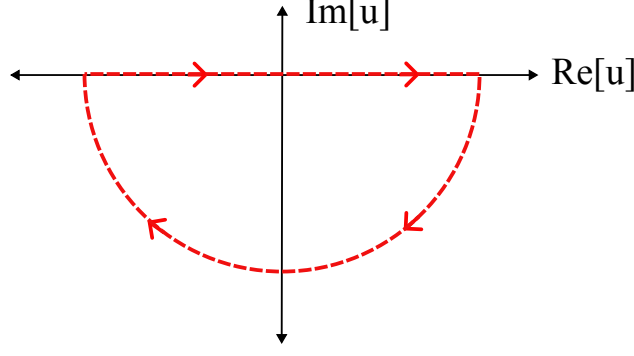


Figure 2: Contour γ_R^- for $t < 0$.

$$= 0, \quad (8.172)$$

where we used the fact that $\left(\frac{\sin(\pi\alpha)}{\sinh(\pi z t)}\right)^2 \geq 0$ for all $t, z > 0$ and $\alpha \in (0, 1)$. Additionally,

$$\int_{-\infty}^{\infty} dt \, p_{\alpha,z}(t) = \lim_{\omega \rightarrow 0} f_{\alpha,z}(\omega) = 1, \quad (8.173)$$

the latter equality already having been argued in (8.141)–(8.150).

The expression in (8.75) follows because

$$\begin{aligned} & \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi z t)} \right)^2 \right) \\ &= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\frac{\sinh^2(\pi z t) + \sin^2(\pi\alpha)}{\sinh^2(\pi z t)} \right) \end{aligned} \quad (8.174)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\frac{\sinh^2(\pi z t) + 1 - \cos^2(\pi\alpha)}{\sinh^2(\pi z t)} \right) \quad (8.175)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\frac{\cosh^2(\pi z t) - \cos^2(\pi\alpha)}{\sinh^2(\pi z t)} \right) \quad (8.176)$$

$$= \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(\coth^2(\pi z t) - \left(\frac{\cos(\pi\alpha)}{\sinh(\pi z t)} \right)^2 \right). \quad (8.177)$$

The equality $p_{\alpha=\frac{1}{2}, z=\frac{1}{2}}(t) = p(t)$ holds because

$$p_{\alpha=\frac{1}{2}, z=\frac{1}{2}}(t) = \frac{\frac{1}{2}}{2\pi\frac{1}{2}(1-\frac{1}{2})} \ln \left(\coth^2 \left(\pi \left(\frac{1}{2} \right) t \right) - \left(\frac{\cos(\pi(\frac{1}{2}))}{\sinh(\pi(\frac{1}{2})t)} \right)^2 \right) \quad (8.178)$$

$$= \frac{1}{\pi} \ln \left(\coth^2 \left(\frac{\pi t}{2} \right) \right) \quad (8.179)$$

$$= \frac{2}{\pi} \ln \left| \coth \left(\frac{\pi t}{2} \right) \right|. \quad (8.180)$$

The final statement about the characteristic function $f_{\alpha,z}(\omega)$ of $p_{\alpha,z}(t)$ holds because, for all $z > 0$ and $\omega \in \mathbb{R} \setminus \{0\}$,

$$\begin{aligned} & \lim_{\alpha \rightarrow 1} f_{\alpha,z}(\omega) \\ &= \lim_{\alpha \rightarrow 1} \frac{z \left(1 - e^{-\left(\frac{1-\alpha}{z}\right)\omega} \right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)\omega} \right)}{\alpha(1-\alpha)\omega \left(1 - e^{-\left(\frac{1}{z}\right)\omega} \right)} \end{aligned} \quad (8.181)$$

$$= \frac{z}{\omega \left(1 - e^{-\left(\frac{1}{z}\right)\omega} \right)} \lim_{\alpha \rightarrow 1} \frac{\left(1 - e^{-\left(\frac{1-\alpha}{z}\right)\omega} \right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)\omega} \right)}{\alpha(1-\alpha)} \quad (8.182)$$

$$= \frac{z}{\omega \left(1 - e^{-\left(\frac{1}{z}\right)\omega} \right)} \lim_{\alpha \rightarrow 1} \frac{\left(-\frac{\omega}{z} e^{-\left(\frac{1-\alpha}{z}\right)\omega} \right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)\omega} \right) + \left(1 - e^{-\left(\frac{1-\alpha}{z}\right)\omega} \right) \left(\frac{\omega}{z} e^{-\left(\frac{\alpha}{z}\right)\omega} \right)}{1 - 2\alpha} \quad (8.183)$$

$$= \frac{z}{\omega \left(1 - e^{-\left(\frac{1}{z}\right)\omega} \right)} \left(-\left(-\frac{\omega}{z} \right) \left(1 - e^{-\left(\frac{1}{z}\right)\omega} \right) \right) \quad (8.184)$$

$$= 1. \quad (8.185)$$

In the case that $\omega = 0$, we already showed in (8.141)–(8.150) that $\lim_{\omega \rightarrow 0} f_{\alpha,z}(\omega) = 1$ for all $\alpha \in (0, 1) \cup (1, \infty)$ and $z > 0$, so that $\lim_{\alpha \rightarrow 1} f_{\alpha,z}(0) = 1$. It follows from Levy's continuity theorem that, for all $z > 0$, the random variable $X_{\alpha,z} \sim q_{\alpha,z}$ converges in distribution to the random variable $X \sim p$ in the limit $\alpha \rightarrow 1$.

Similarly, for all $\alpha \in (0, 1)$ and $\omega \in \mathbb{R} \setminus \{0\}$,

$$\lim_{z \rightarrow \infty} f_{\alpha,z}(\omega) \quad (8.186)$$

$$= \lim_{z \rightarrow \infty} \frac{\left(1 - e^{-\left(\frac{1-\alpha}{z}\right)\omega} \right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)\omega} \right)}{\alpha(1-\alpha)\omega \left(\frac{1}{z} \right) \left(1 - e^{-\left(\frac{1}{z}\right)\omega} \right)} \quad (8.187)$$

$$= \frac{1}{\alpha(1-\alpha)\omega} \lim_{h \rightarrow 0} \frac{\left(1 - e^{-h(1-\alpha)\omega} \right) \left(1 - e^{-h\alpha\omega} \right)}{h \left(1 - e^{-h\omega} \right)} \quad (8.188)$$

$$= \frac{1}{\alpha(1-\alpha)\omega} \lim_{h \rightarrow 0} \frac{(1-\alpha)\omega e^{-h(1-\alpha)\omega} (1 - e^{-h\alpha\omega}) + (1 - e^{-h(1-\alpha)\omega}) \alpha\omega e^{-h\alpha\omega}}{1 - e^{-h\omega} + h\omega e^{-h\omega}} \quad (8.189)$$

$$= \frac{1}{\alpha(1-\alpha)\omega} \lim_{h \rightarrow 0} \frac{\left[-[(1-\alpha)\omega]^2 e^{-h(1-\alpha)\omega} (1 - e^{-h\alpha\omega}) + (1-\alpha)\omega e^{-h(1-\alpha)\omega} \alpha\omega e^{-h\alpha\omega} \right] + (1-\alpha)\omega e^{-h(1-\alpha)\omega} \alpha\omega e^{-h\alpha\omega} - (1 - e^{-h(1-\alpha)\omega}) (\alpha\omega)^2 e^{-h\alpha\omega}}{\omega e^{-h\omega} + \omega e^{-h\omega} - h\omega^2 e^{-h\omega}} \quad (8.190)$$

$$= \frac{1}{\alpha(1-\alpha)\omega} \left(\frac{2(1-\alpha)\alpha\omega^2}{2\omega} \right) \quad (8.191)$$

$$= 1. \tag{8.192}$$

In the case that $\omega = 0$, we already showed in (8.141)–(8.150) that $\lim_{\omega \rightarrow 0} f_{\alpha,z}(\omega) = 1$ for all $\alpha \in (0, 1) \cup (1, \infty)$ and $z > 0$, so that $\lim_{z \rightarrow \infty} f_{\alpha,z}(0) = 1$. It follows from Levy’s continuity theorem that, for all $\alpha \in (0, 1)$, the random variable $X_{\alpha,z} \sim q_{\alpha,z}$ converges in distribution to the random variable $X \sim p$ in the limit $z \rightarrow \infty$. $\square$

9 α - z Information matrices of time-evolved states

In this section, I establish a formula for the α - z information matrix of time-evolved states, for all $\alpha \in (0, 1)$ and $z > 0$ (Theorem 35). By appealing to the former results of [MPW25a], this formula leads to a hybrid quantum–classical algorithm for estimating the elements of the α - z information matrix of time-evolved states. As such, this formula has applications in quantum machine learning, namely, in natural gradient descent algorithms for performing optimization using quantum evolution machines, as put forward in [MPW25a].

Let us now evaluate the expression in (8.1) for time-evolved states. Let

$$H(\phi) := \sum_{j=1}^L \phi_j H_j, \tag{9.1}$$

$$\rho := \frac{e^{-G}}{Z}, \tag{9.2}$$

$$\sigma(\phi) := e^{-iH(\phi)} \rho e^{iH(\phi)}, \tag{9.3}$$

where $\phi_j \in \mathbb{R}$ and H_j is Hermitian for all $j \in \{1, \dots, L\}$, G is Hermitian, and $Z := \text{Tr}[e^{-G}]$. The state $\sigma(\phi)$ is what we refer to as a “time-evolved state.” For analyzing these states, the following quantum channel Ψ_ϕ and its corresponding adjoint channel $\Psi_\phi^\dagger$ play a role:

$$\Psi_\phi(X) := \int_0^1 dt e^{iH(\phi)t} X e^{-iH(\phi)t} = e^{iH(\phi)} \Psi_\phi^\dagger(H_i) e^{-iH(\phi)}, \tag{9.4}$$

$$\Psi_\phi^\dagger(X) := \int_0^1 dt e^{-iH(\phi)t} X e^{iH(\phi)t}, \tag{9.5}$$

As a consequence of Lemma 39 and Corollary 36, the following theorem holds for the α - z information matrices of time-evolved states:

Theorem 35. *For time-evolved states of the form in (9.3), the following equality holds for all $\alpha \in (0, 1)$ and $z > 0$:*

$$[I_{\alpha,z}(\phi)]_{i,j} = \langle [\Phi_{p_{\alpha,z}}(\Psi_\phi(H_i)), [G, \Psi_\phi(H_j)]] \rangle_\rho, \tag{9.6}$$

where the channel Ψ_ϕ is defined in (9.4), the channel $\Phi_{p_{\alpha,z}}$ is defined as

$$\Phi_{p_{\alpha,z}}(X) := \int_{-\infty}^{\infty} dt p_{\alpha,z}(t) e^{-iGt} X e^{iGt}, \tag{9.7}$$

and $p_{\alpha,z}$ is the probability density function defined in (8.74).

As stated above, Theorem 35 follows in part from Lemma 39, which establishes a formula for a general information matrix $I_\zeta(\phi)$ of time-evolved states, under the assumption that a certain Fourier transform exists. Additionally, Theorem 35 follows from Corollary 36, which precisely determines the needed Fourier transform for all $\alpha \in (0, 1)$ and $z > 0$ when the function $\zeta(x, y)$ is set to $\zeta_{\alpha, z}(x, y)$, as defined in (5.6).

The following corollary was established already in (8.92) (therein, recall the definition of $f_{\alpha, z}(\omega)$ in (8.77)):

Corollary 36. *For all $\alpha \in (0, 1)$ and $z > 0$, the Fourier transform of the function*

$$\omega \mapsto \frac{\zeta_{\alpha, z}(e^{-\omega}, 1) (1 - e^{-\omega})}{\omega}, \quad (9.8)$$

where $\zeta_{\alpha, z}(e^{-\omega}, 1)$ is defined in (5.6), is equal to the probability density function $p_{\alpha, z}(t)$ defined in (8.74). That is, for all $\alpha \in (0, 1)$ and $z > 0$,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \frac{z \left(1 - e^{-\left(\frac{1-\alpha}{z}\right)\omega}\right) \left(1 - e^{-\left(\frac{\alpha}{z}\right)\omega}\right)}{\alpha(1-\alpha)\omega \left(1 - e^{-\left(\frac{1}{z}\right)\omega}\right)} e^{i\omega t} \\ = \frac{z}{2\pi\alpha(1-\alpha)} \ln \left(1 + \left(\frac{\sin(\pi\alpha)}{\sinh(\pi z t)}\right)^2\right) =: p_{\alpha, z}(t). \end{aligned} \quad (9.9)$$

Remark 37. We recover the result from [MPW25a, Theorem 19] as a special case of Theorem 35 in the limit $\alpha \rightarrow 1$ for fixed $z > 0$ or in the limit $z \rightarrow \infty$ for fixed $\alpha \in (0, 1)$. Indeed, Lemma 34 states that the characteristic function of $p_{\alpha, z}(t)$ converges to one everywhere in these limits, so that the expression

$$\frac{\zeta_{\alpha, z}(e^{-(\mu_k - \mu_\ell)}, 1) (1 - e^{-(\mu_k - \mu_\ell)})}{\mu_k - \mu_\ell} \quad (9.10)$$

in (9.43) is equal to one for all $\mu_k, \mu_\ell \in \mathbb{R}$ in either of these limits. Then proceeding through the rest of the analysis after (9.43), one finds that

$$[I_{\text{KM}}(\phi)]_{i,j} = \lim_{\alpha \rightarrow 1} [I_{\alpha, z}(\phi)]_{i,j} = \lim_{z \rightarrow \infty} [I_{\alpha, z}(\phi)]_{i,j} = \left\langle \left[\left[\Psi_\phi^\dagger(H_j), G \right], \Psi_\phi^\dagger(H_i) \right] \right\rangle_\rho, \quad (9.11)$$

consistent with the expression reported in [MPW25a, Theorem 19].

Remark 38. By using the same methods use to prove Theorem 35, the results from [MPW25b], regarding information matrices for parameterized quantum circuits, also generalize to α - z information matrices. Indeed, by adopting the same notation used in [MPW25b], we arrive at the following statement: Let ρ be an initial thermal state of the form in (9.2). For a general layered parameterized circuit of the form in [MPW25b, Eq. (1)], the following equality holds for all $\alpha \in (0, 1)$ and $z > 0$:

$$[I_{\alpha, z}(\phi)]_{i,j} = \left\langle \left[\Phi_{p_{\alpha, z}} \left(\mathcal{U}_{R_i}^\dagger(H_i) \right), \left[G, \mathcal{U}_{R_j}^\dagger(H_j) \right] \right] \right\rangle_\rho, \quad (9.12)$$

where $\mathcal{U}_{R_i}^\dagger$ denotes the unitary channel defined in [MPW25b, Eq. (8)] and $\Phi_{p_{\alpha, z}}$ is the channel defined in (9.7).

9.1 General formula for information matrices of time-evolved states

Lemma 39. Let $\zeta(x, y)$ be a function defined for $x, y > 0$ satisfying the properties stated in (8.2)–(8.4), and suppose that the Fourier transform of the function

$$\omega \mapsto \frac{\zeta(e^{-\omega}, 1)(1 - e^{-\omega})}{\omega} \quad (9.13)$$

exists, where $\omega \in \mathbb{R}$. Then, for time-evolved states of the form in (9.3), the following equality holds:

$$[I_\zeta(\phi)]_{i,j} = \langle [\Phi_g(\Psi_\phi(H_i)), [G, \Psi_\phi(H_j)]] \rangle_\rho, \quad (9.14)$$

where $I_\zeta(\phi)$ is defined in (8.1), the channel Ψ_ϕ is defined in (9.4),

$$\Phi_g(X) := \int_{-\infty}^{\infty} dt g(t) e^{-iGt} X e^{iGt}, \quad (9.15)$$

and $g(t)$ is a real-valued function satisfying the following for all $\omega \in \mathbb{R}$:

$$\int_{-\infty}^{\infty} dt g(t) e^{it\omega} = \frac{\zeta(e^{-\omega}, 1)(1 - e^{-\omega})}{\omega}, \quad (9.16)$$

$$\int_{-\infty}^{\infty} dt g(t) = \kappa. \quad (9.17)$$

Proof. As a consequence of Proposition 47, consider that

$$\begin{aligned} & \frac{\partial}{\partial \phi} \sigma(\phi) \\ &= \frac{\partial}{\partial \phi} (e^{-iH(\phi)} \rho e^{iH(\phi)}) \end{aligned} \quad (9.18)$$

$$= \left(\frac{\partial}{\partial \phi} e^{-iH(\phi)} \right) \rho e^{iH(\phi)} + e^{-iH(\phi)} \rho \left(\frac{\partial}{\partial \phi} e^{iH(\phi)} \right) \quad (9.19)$$

$$= -i \int_0^1 dt e^{-iH(\phi)t} H_i e^{-iH(\phi)(1-t)} \rho e^{iH(\phi)} + i e^{-iH(\phi)} \rho \int_0^1 dt e^{iH(\phi)t} H_i e^{iH(\phi)(1-t)} \quad (9.20)$$

$$= -i \int_0^1 dt e^{-iH(\phi)t} H_i e^{iH(\phi)t} e^{-iH(\phi)} \rho e^{iH(\phi)} + i e^{-iH(\phi)} \rho \int_0^1 dt e^{iH(\phi)(1-t)} H_i e^{iH(\phi)t} \quad (9.21)$$

$$= -i \Psi_\phi^\dagger(H_i) \sigma(\phi) + i \sigma(\phi) \Psi_\phi^\dagger(H_i) \quad (9.22)$$

$$= i \left[\sigma(\phi), \Psi_\phi^\dagger(H_i) \right]. \quad (9.23)$$

Let $G = \sum_k \mu_k \tilde{\Pi}_k$ be a spectral decomposition of G , where μ_k is an eigenvalue and $\tilde{\Pi}_k$ is an eigenprojection. Then $\rho = \sum_k \frac{e^{-\mu_k}}{Z} \tilde{\Pi}_k$ is a spectral decomposition of ρ , and set $\lambda_k := \frac{e^{-\mu_k}}{Z}$. Defining $\Pi_k := e^{-iH(\phi)} \tilde{\Pi}_k e^{iH(\phi)}$, it follows that $\sigma(\phi) = \sum_k \lambda_k \Pi_k$ is a spectral decomposition of $\sigma(\phi)$. Applying (9.23), we conclude that

$$\Pi_k (\partial_i \sigma(\phi)) \Pi_\ell = \Pi_k i \left[\sigma(\phi), \Psi_\phi^\dagger(H_i) \right] \Pi_\ell \quad (9.24)$$

$$= i\Pi_k\sigma(\phi)\Psi_\phi^\dagger(H_i)\Pi_\ell - i\Pi_k\Psi_\phi^\dagger(H_i)\sigma(\phi)\Pi_\ell \quad (9.25)$$

$$= i\lambda_k\Pi_k\Psi_\phi^\dagger(H_i)\Pi_\ell - i\lambda_\ell\Pi_k\Psi_\phi^\dagger(H_i)\Pi_\ell \quad (9.26)$$

$$= i(\lambda_k - \lambda_\ell)\Pi_k\Psi_\phi^\dagger(H_i)\Pi_\ell. \quad (9.27)$$

Then, considering (8.1), we find that

$$[I_\zeta(\phi)]_{i,j} = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) \text{Tr}[\Pi_k(\partial_i\sigma(\phi))\Pi_\ell(\partial_j\sigma(\phi))] \quad (9.28)$$

$$= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) i(\lambda_k - \lambda_\ell) \text{Tr}[\Pi_k\Psi_\phi^\dagger(H_i)\Pi_\ell(\partial_j\sigma(\phi))] \quad (9.29)$$

$$= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) i(\lambda_k - \lambda_\ell) \text{Tr}[\Psi_\phi(H_i)\Pi_\ell(\partial_j\sigma(\phi))\Pi_k] \quad (9.30)$$

$$= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) i(\lambda_k - \lambda_\ell) i(\lambda_\ell - \lambda_k) \text{Tr}[\Psi_\phi^\dagger(H_i)\Pi_\ell\Psi_\phi^\dagger(H_j)\Pi_k] \quad (9.31)$$

$$= \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) (\lambda_k - \lambda_\ell)^2 \text{Tr}[\Pi_k\Psi_\phi^\dagger(H_i)\Pi_\ell\Psi_\phi^\dagger(H_j)] \quad (9.32)$$

Note that

$$\text{Tr}[\Pi_k\Psi_\phi^\dagger(H_i)\Pi_\ell\Psi_\phi^\dagger(H_j)] = \text{Tr}[e^{-iH(\phi)}\tilde{\Pi}_k e^{iH(\phi)}\Psi_\phi^\dagger(H_i)e^{-iH(\phi)}\tilde{\Pi}_\ell e^{iH(\phi)}\Psi_\phi^\dagger(H_j)] \quad (9.33)$$

$$= \text{Tr}[\tilde{\Pi}_k e^{iH(\phi)}\Psi_\phi^\dagger(H_i)e^{-iH(\phi)}\tilde{\Pi}_\ell e^{iH(\phi)}\Psi_\phi^\dagger(H_j)e^{-iH(\phi)}] \quad (9.34)$$

$$= \text{Tr}[\tilde{\Pi}_k\Psi_\phi(H_i)\tilde{\Pi}_\ell\Psi_\phi(H_j)], \quad (9.35)$$

where we applied (9.4). Then, continuing from (9.32),

$$[I_\zeta(\phi)]_{i,j} = \sum_{k,\ell} \zeta(\lambda_k, \lambda_\ell) (\lambda_k - \lambda_\ell)^2 \text{Tr}[\tilde{\Pi}_k\Psi_\phi(H_i)\tilde{\Pi}_\ell\Psi_\phi(H_j)] \quad (9.36)$$

$$= \sum_{k,\ell} \zeta\left(\frac{e^{-\mu_k}}{Z}, \frac{e^{-\mu_\ell}}{Z}\right) \left(\frac{e^{-\mu_k}}{Z} - \frac{e^{-\mu_\ell}}{Z}\right)^2 \text{Tr}[\tilde{\Pi}_k\Psi_\phi(H_i)\tilde{\Pi}_\ell\Psi_\phi(H_j)] \quad (9.37)$$

$$= \sum_{k,\ell} Z\zeta(e^{-\mu_k}, e^{-\mu_\ell}) \frac{1}{Z^2} (e^{-\mu_k} - e^{-\mu_\ell})^2 \text{Tr}[\tilde{\Pi}_k\Psi_\phi(H_i)\tilde{\Pi}_\ell\Psi_\phi(H_j)] \quad (9.38)$$

$$= \sum_{k,\ell} \zeta(e^{-\mu_k}, e^{-\mu_\ell}) \frac{1}{Z} (e^{-\mu_k} - e^{-\mu_\ell})^2 \text{Tr}[\tilde{\Pi}_k\Psi_\phi(H_i)\tilde{\Pi}_\ell\Psi_\phi(H_j)] \quad (9.39)$$

$$= \sum_{k,\ell} e^{\mu_\ell} \zeta(e^{-(\mu_k - \mu_\ell)}, 1) \frac{e^{-2\mu_\ell}}{Z} (e^{-(\mu_k - \mu_\ell)} - 1)^2 \text{Tr}[\tilde{\Pi}_k\Psi_\phi(H_i)\tilde{\Pi}_\ell\Psi_\phi(H_j)] \quad (9.40)$$

$$= \sum_{k,\ell} \zeta(e^{-(\mu_k - \mu_\ell)}, 1) \frac{e^{-\mu_\ell}}{Z} (1 - e^{-(\mu_k - \mu_\ell)})^2 \text{Tr}[\tilde{\Pi}_k\Psi_\phi(H_i)\tilde{\Pi}_\ell\Psi_\phi(H_j)] \quad (9.41)$$

$$= \sum_{k,\ell} \frac{\zeta(e^{-(\mu_k-\mu_\ell)}, 1) (1 - e^{-(\mu_k-\mu_\ell)})}{\mu_k - \mu_\ell} \frac{e^{-\mu_\ell}}{Z} (\mu_k - \mu_\ell) (1 - e^{-(\mu_k-\mu_\ell)}) \text{Tr} \left[\tilde{\Pi}_k \Psi_\phi(H_i) \tilde{\Pi}_\ell \Psi_\phi(H_j) \right] \quad (9.42)$$

$$= \sum_{k,\ell} \frac{\zeta(e^{-(\mu_k-\mu_\ell)}, 1) (1 - e^{-(\mu_k-\mu_\ell)})}{\mu_k - \mu_\ell} (\mu_k - \mu_\ell) \left(\frac{e^{-\mu_\ell}}{Z} - \frac{e^{-\mu_k}}{Z} \right) \text{Tr} \left[\tilde{\Pi}_k \Psi_\phi(H_i) \tilde{\Pi}_\ell \Psi_\phi(H_j) \right] \quad (9.43)$$

Now suppose that $g(t)$ is a function satisfying

$$\int_{-\infty}^{\infty} dt g(t) e^{it\omega} = \frac{\zeta(e^{-\omega}, 1) (1 - e^{-\omega})}{\omega}. \quad (9.44)$$

To establish that $g(t)$ is real-valued, it suffices to prove that its Fourier transform is even. To this end, consider that

$$-\omega \mapsto \frac{\zeta(e^{-(-\omega)}, 1) (1 - e^{-(-\omega)})}{-\omega} \quad (9.45)$$

$$= \frac{\zeta(e^\omega, 1) (1 - e^\omega)}{-\omega} \quad (9.46)$$

$$= \frac{\zeta(1, e^\omega) (1 - e^\omega)}{-\omega} \quad (9.47)$$

$$= \frac{e^{-\omega} \zeta(e^{-\omega}, 1) (1 - e^\omega)}{-\omega} \quad (9.48)$$

$$= \frac{\zeta(e^{-\omega}, 1) (e^{-\omega} - 1)}{-\omega} \quad (9.49)$$

$$= \frac{\zeta(e^{-\omega}, 1) (1 - e^{-\omega})}{\omega}. \quad (9.50)$$

Then, continuing from (9.43),

$$[I_\zeta(\phi)]_{i,j} = \sum_{k,\ell} \int_{-\infty}^{\infty} dt g(t) e^{it(\mu_k-\mu_\ell)} (\mu_k - \mu_\ell) \left(\frac{e^{-\mu_\ell}}{Z} - \frac{e^{-\mu_k}}{Z} \right) \text{Tr} \left[\tilde{\Pi}_k \Psi_\phi(H_i) \tilde{\Pi}_\ell \Psi_\phi(H_j) \right] \quad (9.51)$$

$$= \int_{-\infty}^{\infty} dt g(t) \sum_{k,\ell} (\mu_k - \mu_\ell) \left(\frac{e^{-\mu_\ell}}{Z} - \frac{e^{-\mu_k}}{Z} \right) \text{Tr} \left[e^{it\mu_k} \tilde{\Pi}_k \Psi_\phi(H_i) e^{-it\mu_\ell} \tilde{\Pi}_\ell \Psi_\phi(H_j) \right] \quad (9.52)$$

$$\begin{aligned} &= - \int_{-\infty}^{\infty} dt g(t) \sum_{k,\ell} \mu_k \frac{e^{-\mu_k}}{Z} \text{Tr} \left[e^{it\mu_k} \tilde{\Pi}_k \Psi_\phi(H_i) e^{-it\mu_\ell} \tilde{\Pi}_\ell \Psi_\phi(H_j) \right] \\ &\quad + \int_{-\infty}^{\infty} dt g(t) \sum_{k,\ell} \mu_k \frac{e^{-\mu_\ell}}{Z} \text{Tr} \left[e^{it\mu_k} \tilde{\Pi}_k \Psi_\phi(H_i) e^{-it\mu_\ell} \tilde{\Pi}_\ell \Psi_\phi(H_j) \right] \\ &\quad + \int_{-\infty}^{\infty} dt g(t) \sum_{k,\ell} \mu_\ell \frac{e^{-\mu_k}}{Z} \text{Tr} \left[e^{it\mu_k} \tilde{\Pi}_k \Psi_\phi(H_i) e^{-it\mu_\ell} \tilde{\Pi}_\ell \Psi_\phi(H_j) \right] \end{aligned}$$

$$- \int_{-\infty}^{\infty} dt g(t) \sum_{k,\ell} \mu_\ell \frac{e^{-\mu_\ell}}{Z} \text{Tr} \left[e^{it\mu_k} \tilde{\Pi}_k \Psi_\phi(H_i) e^{-it\mu_\ell} \tilde{\Pi}_\ell \Psi_\phi(H_j) \right] \quad (9.53)$$

$$\begin{aligned} &= - \int_{-\infty}^{\infty} dt g(t) \text{Tr} \left[\left(\sum_k \mu_k \frac{e^{-\mu_k}}{Z} e^{it\mu_k} \right) \tilde{\Pi}_k \Psi_\phi(H_i) \left(\sum_\ell e^{-it\mu_\ell} \tilde{\Pi}_\ell \right) \Psi_\phi(H_j) \right] \\ &\quad + \int_{-\infty}^{\infty} dt g(t) \text{Tr} \left[\left(\sum_k \mu_k e^{it\mu_k} \tilde{\Pi}_k \right) \Psi_\phi(H_i) \left(\sum_\ell e^{-it\mu_\ell} \frac{e^{-\mu_\ell}}{Z} \tilde{\Pi}_\ell \right) \Psi_\phi(H_j) \right] \\ &\quad + \int_{-\infty}^{\infty} dt g(t) \text{Tr} \left[\left(\sum_k \frac{e^{-\mu_k}}{Z} e^{it\mu_k} \tilde{\Pi}_k \right) \Psi_\phi(H_i) \left(\sum_\ell e^{-it\mu_\ell} \mu_\ell \tilde{\Pi}_\ell \right) \Psi_\phi(H_j) \right] \\ &\quad - \int_{-\infty}^{\infty} dt g(t) \text{Tr} \left[\left(\sum_k e^{it\mu_k} \tilde{\Pi}_k \right) \Psi_\phi(H_i) \left(\sum_\ell e^{-it\mu_\ell} \mu_\ell \frac{e^{-\mu_\ell}}{Z} \tilde{\Pi}_\ell \right) \Psi_\phi(H_j) \right] \end{aligned} \quad (9.54)$$

$$\begin{aligned} &= - \int_{-\infty}^{\infty} dt g(t) \text{Tr} [G \rho e^{itG} \Psi_\phi(H_i) e^{-itG} \Psi_\phi(H_j)] \\ &\quad + \int_{-\infty}^{\infty} dt g(t) \text{Tr} [G e^{itG} \Psi_\phi(H_i) e^{-itG} \rho \Psi_\phi(H_j)] \\ &\quad + \int_{-\infty}^{\infty} dt g(t) \text{Tr} [\rho e^{itG} \Psi_\phi(H_i) e^{-itG} G \Psi_\phi(H_j)] \\ &\quad - \int_{-\infty}^{\infty} dt g(t) \text{Tr} [e^{itG} \Psi_\phi(H_i) e^{-itG} G \rho \Psi_\phi(H_j)] \end{aligned} \quad (9.55)$$

$$\begin{aligned} &= - \text{Tr} [G \rho \Phi_g(\Psi_\phi(H_i)) \Psi_\phi(H_j)] + \text{Tr} [G \Phi_g(\Psi_\phi(H_i)) \rho \Psi_\phi(H_j)] \\ &\quad + \text{Tr} [\rho \Phi_g(\Psi_\phi(H_i)) G \Psi_\phi(H_j)] - \text{Tr} [\Phi_g(\Psi_\phi(H_i)) G \rho \Psi_\phi(H_j)] \end{aligned} \quad (9.56)$$

$$\begin{aligned} &= - \text{Tr} [\Phi_g(\Psi_\phi(H_i)) \Psi_\phi(H_j) G \rho] + \text{Tr} [\Psi_\phi(H_j) G \Phi_g(\Psi_\phi(H_i)) \rho] \\ &\quad + \text{Tr} [\Phi_g(\Psi_\phi(H_i)) G \Psi_\phi(H_j) \rho] - \text{Tr} [G \Psi_\phi(H_j) \Phi_g(\Psi_\phi(H_i)) \rho] \end{aligned} \quad (9.57)$$

$$= \text{Tr} [\Phi_g(\Psi_\phi(H_i)) [G, \Psi_\phi(H_j)] \rho] - \text{Tr} [[G, \Psi_\phi(H_j)] \Phi_g(\Psi_\phi(H_i)) \rho] \quad (9.58)$$

$$= \text{Tr} [[\Phi_g(\Psi_\phi(H_i)), [G, \Psi_\phi(H_j)]] \rho] \quad (9.59)$$

$$= \langle [\Phi_g(\Psi_\phi(H_i)), [G, \Psi_\phi(H_j)]] \rangle_\rho. \quad (9.60)$$

For the equality in (9.57), we used that $[\rho, G] = 0$. Finally, the claim in (9.17) follows because

$$\int_{-\infty}^{\infty} dt g(t) = \lim_{\omega \rightarrow 0} \frac{\zeta(e^{-\omega}, 1) (1 - e^{-\omega})}{\omega} \quad (9.61)$$

$$= \zeta(1, 1) \left(\lim_{\omega \rightarrow 0} \frac{1 - e^{-\omega}}{\omega} \right) \quad (9.62)$$

$$= \kappa \left(- \frac{d}{d\omega} e^{-\omega} \Big|_{\omega=0} \right) \quad (9.63)$$

$$= \kappa, \quad (9.64)$$

thus concluding the proof. $\square$

10 Information matrices of classical–quantum states

In this section, I show how the various information matrices considered in this paper decompose whenever the underlying parameterized family of states have a classical–quantum structure, generalizing the previous finding reported for the single-parameter case in [KW21, Proposition 7]. In particular, the various information matrices decompose as a sum of a classical part and a quantum part. For the purposes of this section, I adopt a revised notation in which the family on which the information matrix is being evaluated is explicitly written. That is, rather than the abbreviated notation used in (2.18), I write the following instead:

$$[I_D(\theta; (\rho(\theta))_{\theta \in \Theta})]_{i,j} := \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D(\rho(\theta) \| \rho(\theta + \varepsilon)) \Big|_{\varepsilon=0}, \quad (10.1)$$

where $\Theta \subseteq \mathbb{R}^L$ is open.

A parameterized family $(\rho_{XA}(\theta))_{\theta \in \Theta}$ of classical–quantum states has the following form:

$$\rho_{XA}(\theta) := \sum_{x \in \mathcal{X}} p_\theta(x) |x\rangle\langle x| \otimes \rho_x(\theta), \quad (10.2)$$

where $(p_\theta)_{\theta \in \Theta}$ is a parameterized family of probability distributions and, for all $x \in \mathcal{X}$, $(\rho_x(\theta))_{\theta \in \Theta}$ is a parameterized family of quantum states.

Theorem 40. *Let $(\rho_{XA}(\theta))_{\theta \in \Theta}$ be a parameterized family of second-order differentiable, positive definite, classical–quantum states, of the form in (10.2). For all $\alpha \in (0, 1) \cup (1, \infty)$ and $z > 0$, the following equality holds:*

$$I_{\alpha,z}(\theta; (\rho_{XA}(\theta))_{\theta \in \Theta}) = I_F(\theta; (p_\theta)_{\theta \in \Theta}) + \sum_{x \in \mathcal{X}} p_\theta(x) I_{\alpha,z}(\theta; (\rho_x(\theta))_{\theta \in \Theta}), \quad (10.3)$$

where the classical Fisher information matrix $I_F(\theta; (p_\theta)_{\theta \in \Theta})$ is defined in (2.12). Also, the following equalities hold:

$$I_{\text{KM}}(\theta; (\rho_{XA}(\theta))_{\theta \in \Theta}) = I_F(\theta; (p_\theta)_{\theta \in \Theta}) + \sum_{x \in \mathcal{X}} p_\theta(x) I_{\text{KM}}(\theta; (\rho_x(\theta))_{\theta \in \Theta}), \quad (10.4)$$

$$I_{\text{RLD}}(\theta; (\rho_{XA}(\theta))_{\theta \in \Theta}) = I_F(\theta; (p_\theta)_{\theta \in \Theta}) + \sum_{x \in \mathcal{X}} p_\theta(x) I_{\text{RLD}}(\theta; (\rho_x(\theta))_{\theta \in \Theta}). \quad (10.5)$$

Proof. By applying (5.7), it follows that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_{\alpha,z}(\rho_{XA}(\theta) \| \rho_{XA}(\theta + \varepsilon)) \Big|_{\varepsilon=0} \\ &= \frac{1}{\alpha - 1} \left[\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left(\rho_{XA}(\theta)^{\frac{\alpha}{2z}} \rho_{XA}(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho_{XA}(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \Big|_{\varepsilon=0} \right]. \end{aligned} \quad (10.6)$$

Now consider that

$$\text{Tr} \left[\left(\rho_{XA}(\theta)^{\frac{\alpha}{2z}} \rho_{XA}(\theta + \varepsilon)^{\frac{1-\alpha}{z}} \rho_{XA}(\theta)^{\frac{\alpha}{2z}} \right)^z \right]$$

$$= \text{Tr} \left[\left(\frac{(\sum_x |x\rangle\langle x| \otimes p_\theta(x) \rho_x(\theta))^{\frac{\alpha}{2z}} (\sum_{x'} |x'\rangle\langle x'| \otimes p_{\theta+\varepsilon}(x') \rho_{x'}(\theta+\varepsilon))^{\frac{1-\alpha}{z}} \times}{(\sum_{x''} |x''\rangle\langle x''| \otimes p_\theta(x'') \rho_{x''}(\theta))^{\frac{\alpha}{2z}}} \right)^z \right] \quad (10.7)$$

$$= \text{Tr} \left[\left(\frac{\left(\sum_x |x\rangle\langle x| \otimes [p_\theta(x) \rho_x(\theta)]^{\frac{\alpha}{2z}} \right) \left(\sum_{x'} |x'\rangle\langle x'| \otimes [p_{\theta+\varepsilon}(x') \rho_{x'}(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} \right) \times}{\left(\sum_{x''} |x''\rangle\langle x''| \otimes [p_\theta(x'') \rho_{x''}(\theta)]^{\frac{\alpha}{2z}} \right)} \right)^z \right] \quad (10.8)$$

$$= \text{Tr} \left[\left(\sum_x |x\rangle\langle x| \otimes \left([p_\theta(x) \rho_x(\theta)]^{\frac{\alpha}{2z}} [p_{\theta+\varepsilon}(x) \rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [p_\theta(x) \rho_x(\theta)]^{\frac{\alpha}{2z}} \right) \right)^z \right] \quad (10.9)$$

$$= \text{Tr} \left[\sum_x |x\rangle\langle x| \otimes \left([p_\theta(x) \rho_x(\theta)]^{\frac{\alpha}{2z}} [p_{\theta+\varepsilon}(x) \rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [p_\theta(x) \rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \quad (10.10)$$

$$= \sum_x \text{Tr} \left[\left([p_\theta(x) \rho_x(\theta)]^{\frac{\alpha}{2z}} [p_{\theta+\varepsilon}(x) \rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [p_\theta(x) \rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \quad (10.11)$$

$$= \sum_x p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right].$$

Then we find that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left(\rho_{XA}(\theta)^{\frac{\alpha}{2z}} \rho_{XA}(\theta+\varepsilon)^{\frac{1-\alpha}{z}} \rho_{XA}(\theta)^{\frac{\alpha}{2z}} \right)^z \right] \\ &= \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \left(\sum_x p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \right) \end{aligned} \quad (10.12)$$

$$= \frac{\partial}{\partial \varepsilon_i} \left(\sum_x p_\theta(x)^\alpha \frac{\partial}{\partial \varepsilon_j} p_{\theta+\varepsilon}(x)^{1-\alpha} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \right. \\ \left. + \sum_x p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha} \frac{\partial}{\partial \varepsilon_j} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \right) \quad (10.13)$$

$$\begin{aligned} &= \sum_x p_\theta(x)^\alpha \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} p_{\theta+\varepsilon}(x)^{1-\alpha} \right) \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \\ &+ \sum_x p_\theta(x)^\alpha \left(\frac{\partial}{\partial \varepsilon_j} p_{\theta+\varepsilon}(x)^{1-\alpha} \right) \left(\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \right) \\ &+ \sum_x p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \right). \end{aligned} \quad (10.14)$$

Now evaluating this last expression at $\varepsilon = 0$ gives

$$\begin{aligned} & \left(\sum_x p_\theta(x)^\alpha \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} p_{\theta+\varepsilon}(x)^{1-\alpha} \right) \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \right) \Big|_{\varepsilon=0} \\ &+ \left(\sum_x p_\theta(x)^\alpha \left(\frac{\partial}{\partial \varepsilon_j} p_{\theta+\varepsilon}(x)^{1-\alpha} \right) \left(\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta+\varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \right) \right) \Big|_{\varepsilon=0} \end{aligned}$$

$$+ \left(\sum_x p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta + \varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \right) \right) \Big|_{\varepsilon=0} \quad (10.15)$$

$$\begin{aligned} &= \sum_x p_\theta(x)^\alpha \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} p_{\theta+\varepsilon}(x)^{1-\alpha} \Big|_{\varepsilon=0} \right) \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \\ &\quad + \sum_x p_\theta(x)^\alpha \left(\frac{\partial}{\partial \varepsilon_j} p_{\theta+\varepsilon}(x)^{1-\alpha} \Big|_{\varepsilon=0} \right) \left(\frac{\partial}{\partial \varepsilon_i} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta + \varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \Big|_{\varepsilon=0} \right) \\ &\quad + \sum_x p_\theta(x)^\alpha p_\theta(x)^{1-\alpha} \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \text{Tr} \left[\left([\rho_x(\theta)]^{\frac{\alpha}{2z}} [\rho_x(\theta + \varepsilon)]^{\frac{1-\alpha}{z}} [\rho_x(\theta)]^{\frac{\alpha}{2z}} \right)^z \right] \Big|_{\varepsilon=0} \right) \quad (10.16) \end{aligned}$$

$$= [I(\theta; (p_\theta)_{\theta \in \Theta})]_{i,j} + \sum_x p_\theta(x) [I_{\alpha,z}(\theta; (\rho_x(\theta))_{\theta \in \Theta})]_{i,j}. \quad (10.17)$$

Proofs of (10.4) and (10.5) are similar. $\square$

A simple consequence of Theorem 40 is that information matrices that satisfy (10.3)–(10.5) are convex whenever the data-processing inequality holds:

Corollary 41 (Convexity). *For all $x \in \mathcal{X}$, let $(\rho_x(\theta))_{\theta \in \Theta}$ be a parameterized family of second-order differentiable, positive-definite states, let $(p(x))_{x \in \mathcal{X}}$ be a probability distribution, and let $\bar{\rho}(\theta) := \sum_{x \in \mathcal{X}} p(x) \rho_x(\theta)$. Then, for all $\alpha, z > 0$ satisfying the conditions stated in Fact 9, the following convexity inequality holds:*

$$\sum_{x \in \mathcal{X}} p(x) I_{\alpha,z}(\theta; (\rho_x(\theta))_{\theta \in \Theta}) \geq I_{\alpha,z}(\theta; (\bar{\rho}(\theta))_{\theta \in \Theta}). \quad (10.18)$$

Additionally,

$$\sum_{x \in \mathcal{X}} p(x) I_{\text{KM}}(\theta; (\rho_x(\theta))_{\theta \in \Theta}) \geq I_{\text{KM}}(\theta; (\bar{\rho}(\theta))_{\theta \in \Theta}), \quad (10.19)$$

$$\sum_{x \in \mathcal{X}} p(x) I_{\text{RLD}}(\theta; (\rho_x(\theta))_{\theta \in \Theta}) \geq I_{\text{RLD}}(\theta; (\bar{\rho}(\theta))_{\theta \in \Theta}). \quad (10.20)$$

Proof. This is a consequence of Theorem 40 and the data-processing inequality. Indeed, we can apply Theorem 40 to the following classical-quantum state family:

$$\sigma_{XA}(\theta) := \sum_{x \in \mathcal{X}} p(x) |x\rangle\langle x| \otimes \rho_x(\theta), \quad (10.21)$$

to conclude that

$$I_{\alpha,z}(\theta; (\sigma_{XA}(\theta))_{\theta \in \Theta}) = I_F(\theta; (p)_{\theta \in \Theta}) + \sum_{x \in \mathcal{X}} p(x) I_{\alpha,z}(\theta; (\rho_x(\theta))_{\theta \in \Theta}) \quad (10.22)$$

$$= \sum_{x \in \mathcal{X}} p(x) I_{\alpha,z}(\theta; (\rho_x(\theta))_{\theta \in \Theta}), \quad (10.23)$$

while noting that $I_F(\theta; (p)_{\theta \in \Theta}) = 0$ because the family $(p)_{\theta \in \Theta}$ has no dependence on the parameter vector θ . Applying the data-processing inequality for $I_{\alpha,z}$, we then find that

$$I_{\alpha,z}(\theta; (\sigma_{XA}(\theta))_{\theta \in \Theta}) \geq I_{\alpha,z}(\theta; (\text{Tr}_X[\sigma_{XA}(\theta)])_{\theta \in \Theta}) \quad (10.24)$$

$$= I_{\alpha,z}(\theta; (\bar{\rho}(\theta))_{\theta \in \Theta}), \quad (10.25)$$

because $\text{Tr}_X[\sigma_{XA}(\theta)] = \bar{\rho}(\theta)$. Proofs for (10.19) and (10.20) follow similarly. $\square$

11 Conclusion

In summary, this paper explored quantum generalizations of the Fisher information matrix that are derived from quantum Rényi relative entropies, including the log-Euclidean, geometric, and α - z Rényi relative entropies. Among the various results of this paper, key contributions include a detailed proof of the formula in Theorem 10 for the α - z information matrix, the formula in Theorem 32 for the α - z information matrix of parameterized thermal states, the formula in Theorem 35 for the α - z information matrix of time-evolved states, and orderings for the Petz- and sandwiched Rényi information matrices (Theorem 28 and Theorem 30).

Going forward from here, it is an interesting open question to determine the full range of $\alpha, z > 0$ for which the data processing inequality holds for the α - z information matrix. Fact 9 provides sufficient conditions on $\alpha, z > 0$ for which it holds, but Theorem 3 implies that these conditions are not necessary. By [Pet96, Theorems 3 and 5], determining the full range of $\alpha, z > 0$ for which data processing holds is equivalent to determining the full range of $\alpha, z > 0$ for which the function in Corollary 14 is operator monotone. It is also an intriguing open question to determine if there is an appealing expression for the α - z information matrices of parameterized thermal states and time-evolved states, like those in Theorem 32 and Theorem 35, for $\alpha > 1$.

Acknowledgements—I am grateful to Milan Mosonyi for helpful comments on the manuscript, and I thank Hami Mehrabi for a helpful discussion. I am also grateful to Michele Minervini and Dhrumil Patel for our collaborations on [PKPW24, PW24, MPW25a, MPW25b], which were foundational for the present paper. I acknowledge support from the National Science Foundation under grant no. 2329662 and from the Cornell School of Electrical and Computer Engineering.

References

- [AAKS21] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. Sample-efficient learning of interacting quantum systems. *Nature Physics*, 17(8):931–935, 2021. [doi:10.1038/s41567-021-01232-0](https://doi.org/10.1038/s41567-021-01232-0).
- [AAR⁺18] Mohammad H. Amin, Evgeny Andriyash, Jason Rolfe, Bohdan Kulchyt-sky, and Roger Melko. Quantum Boltzmann machine. *Physical Re-*

- view X, 8:021050, May 2018. URL: <https://link.aps.org/doi/10.1103/PhysRevX.8.021050>, doi:10.1103/PhysRevX.8.021050.
- [AD15] Koenraad M. R. Audenaert and Nilanjana Datta. α - z -Rényi relative entropies. *Journal of Mathematical Physics*, 56(2):022202, February 2015. doi:10.1063/1.4906367.
- [Ama98] Shun-Ichi Amari. Natural gradient works efficiently in learning. *Neural Computation*, 10(2):251–276, 1998. doi:10.1162/089976698300017746.
- [Ama16] Shun-ichi Amari. *Information Geometry and Its Applications*, volume 194 of *Applied Mathematical Sciences*. Springer Tokyo, 2016. doi:10.1007/978-4-431-55978-8.
- [BC94] Samuel L. Braunstein and Carlton M. Caves. Statistical distance and the geometry of quantum states. *Physical Review Letters*, 72:3439–3443, May 1994. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.72.3439>, doi:10.1103/PhysRevLett.72.3439.
- [BES18] Souvik Banerjee, Johanna Erdmenger, and Debajyoti Sarkar. Connecting Fisher information to bulk entanglement in holography. *Journal of High Energy Physics*, 2018(8):1, 2018. doi:10.1007/JHEP08(2018)001.
- [Bha97] Rajendra Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics. Springer New York, NY, 1997. doi:10.1007/978-1-4612-0653-8.
- [BRGBPO17] Marcello Benedetti, John Realpe-Gómez, Rupak Biswas, and Alejandro Perdomo-Ortiz. Quantum-assisted learning of hardware-embedded probabilistic graphical models. *Physical Review X*, 7(4):041052, November 2017. URL: <http://dx.doi.org/10.1103/PhysRevX.7.041052>, doi:10.1103/physrevx.7.041052.
- [BS82] V. P. Belavkin and P. Staszewski. C^* -algebraic generalization of relative entropy and entropy. *Annales de l’I.H.P. Physique théorique*, 37(1):51–58, 1982. URL: <http://eudml.org/doc/76163>.
- [BZ06] Ingemar Bengtsson and Karol Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2006. doi:10.1017/CB09780511535048.
- [CDL⁺18] F. M. Ciaglia, F. Di Cosmo, M. Laudato, G. Marmo, F. M. Mele, F. Ventriglia, and P. Vitale. A pedagogical intrinsic approach to relative entropies as potential functions of quantum metrics: The q - z family. *Annals of Physics*, 395:238–274, 2018. URL: <https://www.sciencedirect.com/science/article/pii/S0003491618301519>, doi:10.1016/j.aop.2018.05.015.

- [Čen82] N. N. Čencov. *Statistical Decision Rules and Optimal Inference*. Translations of Mathematical Monographs 53. American Mathematical Society, Providence, 1982. Originally Published in Russian. Nauka, Moscow (1972).
- [CT05] Thomas M. Cover and Joy A. Thomas. *Information Theory and Statistics*, chapter 11, pages 347–408. John Wiley & Sons, Ltd., 2005. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/047174882X.ch11>, doi:10.1002/047174882X.ch11.
- [CVS20] Angelo Carollo, Davide Valenti, and Bernardo Spagnolo. Geometry of quantum phase transitions. *Physics Reports*, 838:1–72, 2020. URL: <https://www.sciencedirect.com/science/article/pii/S0370157319303655>, doi:10.1016/j.physrep.2019.11.002.
- [CVZ07] Lorenzo Campos Venuti and Paolo Zanardi. Quantum critical scaling of the geometric tensors. *Physical Review Letters*, 99:095701, August 2007. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.99.095701>, doi:10.1103/PhysRevLett.99.095701.
- [DL14] Nilanjana Datta and Felix Leditzky. A limit of the quantum Rényi divergence. *Journal of Physics A: Mathematical and Theoretical*, 47(4):045304, January 2014. doi:10.1088/1751-8113/47/4/045304.
- [EO19] S. Ejima and Y. Ogata. Perturbation theory of KMS states. *Annales Henri Poincaré*, 20(9):2971–2986, 2019. doi:10.1007/s00023-019-00824-0.
- [FL13] Rupert L. Frank and Elliott H. Lieb. Monotonicity of a relative Rényi entropy. *Journal of Mathematical Physics*, 54:122201, December 2013. URL: <https://doi.org/10.1063/1.4838835>, arXiv:1306.5358.
- [Gu10] Shi-Jian Gu. Fidelity approach to quantum phase transitions. *International Journal of Modern Physics B*, 24(23):4371–4458, 2010. arXiv:<https://doi.org/10.1142/S0217979210056335>, doi:10.1142/S0217979210056335.
- [Has93] Hiroshi Hasegawa. α -divergence of the non-commutative information geometry. *Reports on Mathematical Physics*, 33(1):87–93, 1993. URL: <https://www.sciencedirect.com/science/article/pii/003448779390043E>, doi:10.1016/0034-4877(93)90043-E.
- [Has07] Matthew B. Hastings. Quantum belief propagation: An algorithm for thermal quantum systems. *Physical Review B*, 76:201102, November 2007. URL: <https://link.aps.org/doi/10.1103/PhysRevB.76.201102>, doi:10.1103/PhysRevB.76.201102.
- [Hay17] Masahito Hayashi. *Quantum Information Geometry and Quantum Estimation*, pages 253–322. Springer Berlin Heidelberg, Berlin, Heidelberg, 2017. doi:10.1007/978-3-662-49725-8_6.

- [Hel67] Carl W. Helstrom. Minimum mean-squared error of estimates in quantum statistics. *Physics Letters A*, 25(2):101–102, 1967. URL: <https://www.sciencedirect.com/science/article/pii/0375960167903660>, doi: [10.1016/0375-9601\(67\)90366-0](https://doi.org/10.1016/0375-9601(67)90366-0).
- [Hel69] Carl W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1:231–252, 1969. doi:[10.1007/BF01007479](https://doi.org/10.1007/BF01007479).
- [HHTZ16] Philipp Hauke, Markus Heyl, Luca Tagliacozzo, and Peter Zoller. Measuring multipartite entanglement through dynamic susceptibilities. *Nature Physics*, 12(8):778–782, 2016. doi:[10.1038/nphys3700](https://doi.org/10.1038/nphys3700).
- [Hia10] Fumio Hiai. Matrix analysis: Matrix monotone functions, matrix means, and majorization. *Interdisciplinary Information Sciences*, 16(2):139–248, 2010. doi:[10.4036/iis.2010.139](https://doi.org/10.4036/iis.2010.139).
- [Hia13] Fumio Hiai. Concavity of certain matrix trace and norm functions. *Linear Algebra and its Applications*, 439(5):1568–1589, 2013. URL: <https://www.sciencedirect.com/science/article/pii/S0024379513002826>, doi:[10.1016/j.laa.2013.04.020](https://doi.org/10.1016/j.laa.2013.04.020).
- [Hia18] Fumio Hiai. Quantum f -divergences in von Neumann algebras. I. standard f -divergences. *Journal of Mathematical Physics*, 59(10):102202, September 2018. doi:[10.1063/1.5039973](https://doi.org/10.1063/1.5039973).
- [Hia19] Fumio Hiai. Quantum f -divergences in von Neumann algebras. II. maximal f -divergences. *Journal of Mathematical Physics*, 60(1):012203, January 2019. doi:[10.1063/1.5051427](https://doi.org/10.1063/1.5051427).
- [Hia21] Fumio Hiai. *Quantum f -divergences in von Neumann Algebras: Reversibility of Quantum Operations*. Mathematical Physics Studies. Springer Singapore, 2021. doi:[10.1007/978-981-33-4199-9](https://doi.org/10.1007/978-981-33-4199-9).
- [HMPB11] Fumio Hiai, Milán Mosonyi, Dénes Petz, and Cédric Bény. Quantum f -divergences and error correction. *Reviews in Mathematical Physics*, 23(07):691–747, 2011. doi:[10.1142/S0129055X11004412](https://doi.org/10.1142/S0129055X11004412).
- [Hol11] Alexander S. Holevo. *Probabilistic and statistical aspects of quantum theory*. Publications of the Scuola Normale Superiore. Edizioni della Normale Pisa, 2011. doi:[10.1007/978-88-7642-378-9](https://doi.org/10.1007/978-88-7642-378-9).
- [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. doi:[10.1007/BF02100287](https://doi.org/10.1007/BF02100287).
- [HP14] Fumio Hiai and Dénes Petz. *Introduction to Matrix Analysis and Applications*. Springer Cham, 2014. doi:[10.1007/978-3-319-04150-6](https://doi.org/10.1007/978-3-319-04150-6).

- [IRS17] Raban Iten, Joseph M. Renes, and David Sutter. Pretty good measures in quantum information theory. *IEEE Transactions on Information Theory*, 63(2):1270–1279, 2017. doi:[10.1109/TIT.2016.2639521](https://doi.org/10.1109/TIT.2016.2639521).
- [JK20] Marcin Jarzyna and Jan Kolodynski. Geometric approach to quantum statistical inference. *IEEE Journal on Selected Areas in Information Theory*, 1(2):367–386, August 2020. doi:[10.1109/JSAIT.2020.3017469](https://doi.org/10.1109/JSAIT.2020.3017469).
- [KB19] Kohtaro Kato and Fernando G. S. L. Brandão. Quantum approximate Markov chains are thermal. *Communications in Mathematical Physics*, 370(1):117–149, 2019. doi:[10.1007/s00220-019-03485-6](https://doi.org/10.1007/s00220-019-03485-6).
- [KB22] Bálint Koczor and Simon C. Benjamin. Quantum natural gradient generalized to noisy and nonunitary circuits. *Physical Review A*, 106:062416, December 2022. URL: <https://link.aps.org/doi/10.1103/PhysRevA.106.062416>, doi:[10.1103/PhysRevA.106.062416](https://doi.org/10.1103/PhysRevA.106.062416).
- [Kim12] Isaac H. Kim. Perturbative analysis of topological entanglement entropy from conditional independence. *Physical Review B*, 86:245116, December 2012. URL: <https://link.aps.org/doi/10.1103/PhysRevB.86.245116>, doi:[10.1103/PhysRevB.86.245116](https://doi.org/10.1103/PhysRevB.86.245116).
- [KK11] Aleksandr Petrovich Korostelev and Olga Korosteleva. *Mathematical Statistics: Asymptotic Minimax Theory*, volume 119. American Mathematical Society, 2011. doi:[10.1090/gsm/119](https://doi.org/10.1090/gsm/119).
- [KW17] Maria Kieferova and Nathan Wiebe. Tomography and generative data modeling via quantum Boltzmann training. *Physical Review A*, 96(6):062327, December 2017. arXiv:[1612.05204](https://arxiv.org/abs/1612.05204), doi:[10.1103/PhysRevA.96.062327](https://doi.org/10.1103/PhysRevA.96.062327).
- [KW20] Sumeet Khatri and Mark M. Wilde. Principles of quantum communication theory: A modern approach, 2020. URL: <https://arxiv.org/abs/2011.04672v2>, arXiv:[2011.04672v2](https://arxiv.org/abs/2011.04672v2).
- [KW21] Vishal Katariya and Mark M. Wilde. Geometric distinguishability measures limit quantum channel estimation and discrimination. *Quantum Information Processing*, 20(2):78, 2021. doi:[10.1007/s11128-021-02992-7](https://doi.org/10.1007/s11128-021-02992-7).
- [LT15] Simon M. Lin and Marco Tomamichel. Investigating properties of a family of quantum Rényi divergences. *Quantum Information Processing*, 14(4):1501–1512, 2015. doi:[10.1007/s11128-015-0935-y](https://doi.org/10.1007/s11128-015-0935-y).
- [LVR16] Nima Lashkari and Mark Van Raamsdonk. Canonical energy is quantum Fisher information. *Journal of High Energy Physics*, 2016(4):153, 2016. doi:[10.1007/JHEP04\(2016\)153](https://doi.org/10.1007/JHEP04(2016)153).

- [LYLW19] Jing Liu, Haidong Yuan, Xiao-Ming Lu, and Xiaoguang Wang. Quantum Fisher information matrix and multiparameter estimation. *Journal of Physics A: Mathematical and Theoretical*, 53(2):023001, December 2019. URL: <https://dx.doi.org/10.1088/1751-8121/ab5d4d>, doi:10.1088/1751-8121/ab5d4d.
- [Mar22] Iman Marvian. Operational interpretation of quantum Fisher information in quantum thermodynamics. *Physical Review Letters*, 129:190502, October 2022. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.129.190502>, doi:10.1103/PhysRevLett.129.190502.
- [Mat10] Keiji Matsumoto. Reverse test and quantum analogue of classical fidelity and generalized fidelity, 2010. URL: <https://arxiv.org/abs/1006.0302>, arXiv:1006.0302.
- [Mat13] Keiji Matsumoto. A new quantum version of f -divergence, 2013. URL: <https://arxiv.org/abs/1311.4722>, arXiv:1311.4722.
- [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*, volume 261, pages 229–273, Singapore, 2018. Springer Singapore. Series Title: Springer Proceedings in Mathematics & Statistics. URL: http://link.springer.com/10.1007/978-981-13-2487-1_10, doi:10.1007/978-981-13-2487-1_10.
- [MBV24] Milán Mosonyi, Gergely Bunn, and Péter Vrana. Geometric relative entropies and barycentric Rényi divergences. *Linear Algebra and its Applications*, 699:159–276, 2024. URL: <https://www.sciencedirect.com/science/article/pii/S0024379524002490>, doi:10.1016/j.laa.2024.06.005.
- [MC91] E. A. Morozova and N. N. Chentsov. Markov invariant geometry on manifolds of states. *Journal of Soviet Mathematics*, 56(5):2648–2669, 1991. doi:10.1007/BF01095975.
- [Mey21] Johannes Jakob Meyer. Fisher information in noisy intermediate-scale quantum applications. *Quantum*, 5:539, September 2021. doi:10.22331/q-2021-09-09-539.
- [MH11] Milán Mosonyi and Fumio Hiai. On the quantum Rényi relative entropies and related capacity formulas. *IEEE Transactions on Information Theory*, 57(4):2474–2487, 2011. doi:10.1109/TIT.2011.2110050.
- [MH18] Alex May and Eliot Hijano. The holographic entropy zoo. *Journal of High Energy Physics*, 2018(10):36, 2018. doi:10.1007/JHEP10(2018)036.

- [MLDS⁺13] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: a new generalization and some properties. *Journal of Mathematical Physics*, 54:122203, December 2013. URL: <https://doi.org/10.1063/1.4838856>, [arXiv:1306.3142](#).
- [MNS⁺15] Masamichi Miyaji, Tokiro Numasawa, Noburo Shiba, Tadashi Takayanagi, and Kento Watanabe. Distance between quantum states and gauge-gravity duality. *Physical Review Letters*, 115:261602, December 2015. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.115.261602>, [doi:10.1103/PhysRevLett.115.261602](#).
- [MO17] Milán Mosonyi and Tomohiro Ogawa. Strong converse exponent for classical-quantum channel coding. *Communications in Mathematical Physics*, 355:373–426, 2017. [arXiv:1409.3562](#), [doi:10.1007/s00220-017-2928-4](#).
- [MPW25a] Michele Minervini, Dhruvil Patel, and Mark M. Wilde. Evolved quantum Boltzmann machines, 2025. URL: <https://arxiv.org/abs/2501.03367v2>, [arXiv:2501.03367v2](#).
- [MPW25b] Michele Minervini, Dhruvil Patel, and Mark M. Wilde. Quantum natural gradient with thermal-state initialization. *Physical Review A*, 112:022424, August 2025. URL: <https://link.aps.org/doi/10.1103/49b1-fbc3>, [doi:10.1103/49b1-fbc3](#).
- [Nie22] Frank Nielsen. The many faces of information geometry. *Notices of the American Mathematical Society*, 69(1):36–45, 2022. [doi:10.1090/noti2403](#).
- [Pet85] Dénes Petz. Quasi-entropies for states of a von Neumann algebra. *Publications of the Research Institute for Mathematical Sciences*, 21:787–800, 1985. [doi:10.2977/prims/1195178929](#).
- [Pet86] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports in Mathematical Physics*, 23:57–65, 1986. URL: [https://doi.org/10.1016/0034-4877\(86\)90067-4](https://doi.org/10.1016/0034-4877(86)90067-4).
- [Pet96] Dénes Petz. Monotone metrics on matrix spaces. *Linear Algebra and its Applications*, 244:81–96, 1996. URL: <https://www.sciencedirect.com/science/article/pii/0024379594002118>, [doi:10.1016/0024-3795\(94\)00211-8](#).
- [PKPW24] Dhruvil Patel, Daniel Koch, Saahil Patel, and Mark M. Wilde. Quantum Boltzmann machine learning of ground-state energies, 2024. URL: <https://arxiv.org/abs/2410.12935v2>, [arXiv:2410.12935v2](#).
- [PW24] Dhruvil Patel and Mark M. Wilde. Natural gradient and parameter estimation for quantum Boltzmann machines, 2024. URL: <https://arxiv.org/abs/2410.24058>, [arXiv:2410.24058](#).

- [SASS24] Matteo Scandi, Paolo Abiuso, Jacopo Surace, and Dario De Santis. Quantum Fisher information and its dynamical nature, 2024. URL: <https://arxiv.org/abs/2304.14984>, [arXiv:2304.14984](https://arxiv.org/abs/2304.14984).
- [SIKC20] James Stokes, Josh Izaac, Nathan Killoran, and Giuseppe Carleo. Quantum natural gradient. *Quantum*, 4:269, May 2020. doi:[10.22331/q-2020-05-25-269](https://doi.org/10.22331/q-2020-05-25-269).
- [SK20] Jasmininder S. Sidhu and Pieter Kok. Geometric perspective on quantum parameter estimation. *AVS Quantum Science*, 2(1):014701, February 2020. doi:[10.1116/1.5119961](https://doi.org/10.1116/1.5119961).
- [SKP24] Mohammad Aamir Sohail, Mohsen Heidari Khoozani, and S. Sandeep Pradhan. Quantum natural stochastic pairwise coordinate descent, 2024. URL: <https://arxiv.org/abs/2407.13858>, [arXiv:2407.13858](https://arxiv.org/abs/2407.13858).
- [SMP⁺22] Faris M. Sbahi, Antonio J. Martinez, Sahil Patel, Dmitri Saberi, Jae Hyeon Yoo, Geoffrey Roeder, and Guillaume Verdon. Provably efficient variational generative modeling of quantum many-body systems via quantum-probabilistic information geometry, 2022. URL: <https://arxiv.org/abs/2206.04663v1>, [arXiv:2206.04663v1](https://arxiv.org/abs/2206.04663v1).
- [TF17] Kaito Takahashi and Akio Fujiwara. Information geometry of sandwiched Rényi α -divergence. *Journal of Physics A: Mathematical and Theoretical*, 50(16):165301, mar 2017. doi:[10.1088/1751-8121/aa6326](https://doi.org/10.1088/1751-8121/aa6326).
- [TNR21] Kok Chuan Tan, Varun Narasimhachar, and Bartosz Regula. Fisher information universally identifies quantum resources. *Physical Review Letters*, 127:200402, November 2021. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.127.200402>, doi:[10.1103/PhysRevLett.127.200402](https://doi.org/10.1103/PhysRevLett.127.200402).
- [Ume62] Hisaharu Umegaki. Conditional expectation in an operator algebra, IV (entropy and information). *Kodai Mathematical Journal*, 14:59–85, January 1962. doi:[10.2996/kmj/1138844604](https://doi.org/10.2996/kmj/1138844604).
- [vEH14] Tim van Erven and Peter Harremoës. Rényi divergence and Kullback-Leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014. doi:[10.1109/TIT.2014.2320500](https://doi.org/10.1109/TIT.2014.2320500).
- [Wei] Eric W. Weisstein. Convolution theorem. From MathWorld—A Wolfram Resource. URL: <https://mathworld.wolfram.com/ConvolutionTheorem.html>.
- [Wil18] Mark M. Wilde. Optimized quantum f -divergences and data processing. *Journal of Physics A*, 51:374002, September 2018. URL: <https://doi.org/10.1088/1751-8121/aad5a1>, [arXiv:1710.10252](https://arxiv.org/abs/1710.10252).

- [WWY14] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331:593–622, October 2014. URL: <https://doi.org/10.1007/s00220-014-2122-x>, [arXiv:1306.1586](https://arxiv.org/abs/1306.1586).
- [YMST25] Koji Yamaguchi, Yosuke Mitsuhashi, Tomohiro Shitara, and Hiroyasu Tajima. Quantum geometric tensor determines the i.i.d. conversion rate in the resource theory of asymmetry for any compact Lie group, 2025. URL: <https://arxiv.org/abs/2411.04766>, [arXiv:2411.04766](https://arxiv.org/abs/2411.04766).
- [ZGC07] Paolo Zanardi, Paolo Giorda, and Marco Cozzini. Information-theoretic differential geometry of quantum phase transitions. *Physical Review Letters*, 99:100603, September 2007. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.99.100603>, [doi:10.1103/PhysRevLett.99.100603](https://doi.org/10.1103/PhysRevLett.99.100603).
- [Zha20] Haonan Zhang. From Wigner-Yanase-Dyson conjecture to Carlen-Frank-Lieb conjecture. *Advances in Mathematics*, 365:107053, 2020. URL: <https://www.sciencedirect.com/science/article/pii/S0001870820300785>, [arXiv:1811.01205](https://arxiv.org/abs/1811.01205), [doi:10.1016/j.aim.2020.107053](https://doi.org/10.1016/j.aim.2020.107053).

A Proofs of Equations (2.13) and (2.14)

This appendix provides short proofs of (2.13) and (2.14), when the parameterized family $(p_\theta)_{\theta \in \Theta}$ of probability distributions are in the interior of the simplex (i.e., $p_\theta(x) > 0$ for all $x \in \mathcal{X}$).

Beginning with (2.13), consider that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha(p_\theta \| p_{\theta+\varepsilon}) \\ &= \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \frac{1}{\alpha - 1} \ln \left(\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha} \right) \end{aligned} \quad (\text{A.1})$$

$$= \frac{1}{\alpha - 1} \frac{\partial}{\partial \varepsilon_i} \left(\frac{\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \varepsilon_j} (p_{\theta+\varepsilon}(x)^{1-\alpha})}{\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha}} \right) \quad (\text{A.2})$$

$$\begin{aligned} &= \frac{\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (p_{\theta+\varepsilon}(x)^{1-\alpha})}{(\alpha - 1) \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha}} \\ &\quad - \frac{\left[\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \varepsilon_i} (p_{\theta+\varepsilon}(x)^{1-\alpha}) \right] \left[\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \varepsilon_j} (p_{\theta+\varepsilon}(x)^{1-\alpha}) \right]}{(\alpha - 1) \left(\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha} \right)^2}. \end{aligned} \quad (\text{A.3})$$

It then follows that

$$\begin{aligned}
& \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha(p_\theta \| p_{\theta+\varepsilon}) \right|_{\varepsilon=0} \\
&= \left. \frac{\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (p_{\theta+\varepsilon}(x)^{1-\alpha})}{(\alpha-1) \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha}} \right|_{\varepsilon=0} \\
&\quad - \left. \frac{\left[\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \varepsilon_i} (p_{\theta+\varepsilon}(x)^{1-\alpha}) \right] \left[\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \varepsilon_j} (p_{\theta+\varepsilon}(x)^{1-\alpha}) \right]}{(\alpha-1) \left(\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_{\theta+\varepsilon}(x)^{1-\alpha} \right)^2}} \right|_{\varepsilon=0} \tag{A.4}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (p_{\theta+\varepsilon}(x)^{1-\alpha}) \Big|_{\varepsilon=0}}{(\alpha-1) \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_\theta(x)^{1-\alpha}} \\
&\quad - \frac{\left[\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \varepsilon_i} (p_{\theta+\varepsilon}(x)^{1-\alpha}) \Big|_{\varepsilon=0} \right] \left[\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \varepsilon_j} (p_{\theta+\varepsilon}(x)^{1-\alpha}) \Big|_{\varepsilon=0} \right]}{(\alpha-1) \left(\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_\theta(x)^{1-\alpha} \right)^2} \tag{A.5}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\alpha-1} \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial^2}{\partial \theta_i \partial \theta_j} (p_\theta(x)^{1-\alpha}) \\
&\quad - \frac{1}{\alpha-1} \left[\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \theta_i} (p_\theta(x)^{1-\alpha}) \right] \left[\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \theta_j} (p_\theta(x)^{1-\alpha}) \right]. \tag{A.6}
\end{aligned}$$

Now consider that

$$\sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \theta_i} (p_\theta(x)^{1-\alpha}) = (1-\alpha) \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_\theta(x)^{-\alpha} \frac{\partial}{\partial \theta_i} (p_\theta(x)) \tag{A.7}$$

$$= (1-\alpha) \sum_{x \in \mathcal{X}} \frac{\partial}{\partial \theta_i} p_\theta(x) \tag{A.8}$$

$$= (1-\alpha) \frac{\partial}{\partial \theta_i} \sum_{x \in \mathcal{X}} p_\theta(x) \tag{A.9}$$

$$= (1-\alpha) \frac{\partial}{\partial \theta_i} (1) \tag{A.10}$$

$$= 0. \tag{A.11}$$

Furthermore,

$$\begin{aligned}
& \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial^2}{\partial \theta_i \partial \theta_j} (p_\theta(x)^{1-\alpha}) \\
&= (1-\alpha) \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \frac{\partial}{\partial \theta_i} \left(p_\theta(x)^{-\alpha} \frac{\partial}{\partial \theta_j} p_\theta(x) \right) \tag{A.12}
\end{aligned}$$

$$= (1-\alpha) \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha \left(\frac{\partial}{\partial \theta_i} p_\theta(x)^{-\alpha} \frac{\partial}{\partial \theta_j} p_\theta(x) + p_\theta(x)^{-\alpha} \frac{\partial^2}{\partial \theta_i \partial \theta_j} p_\theta(x) \right) \tag{A.13}$$

$$\begin{aligned}
&= \alpha(\alpha-1) \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_\theta(x)^{-\alpha-1} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right) \\
&\quad + (1-\alpha) \sum_{x \in \mathcal{X}} p_\theta(x)^\alpha p_\theta(x)^{-\alpha} \frac{\partial^2}{\partial \theta_i \partial \theta_j} p_\theta(x)
\end{aligned} \tag{A.14}$$

$$= \alpha(\alpha-1) \sum_{x \in \mathcal{X}} p_\theta(x)^{-1} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right) + (1-\alpha) \sum_{x \in \mathcal{X}} \frac{\partial^2}{\partial \theta_i \partial \theta_j} p_\theta(x) \tag{A.15}$$

$$= \alpha(\alpha-1) \sum_{x \in \mathcal{X}} p_\theta(x)^{-1} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right) + (1-\alpha) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \sum_{x \in \mathcal{X}} p_\theta(x) \tag{A.16}$$

$$= \alpha(\alpha-1) \sum_{x \in \mathcal{X}} p_\theta(x)^{-1} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right). \tag{A.17}$$

Thus, we conclude that

$$\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D_\alpha(p_\theta \| p_{\theta+\varepsilon}) \right|_{\varepsilon=0} = \alpha \sum_{x \in \mathcal{X}} p_\theta(x)^{-1} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right), \tag{A.18}$$

which is equivalent to the desired equality in (2.13).

Now let us prove (2.14). Consider that

$$\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D(p_\theta \| p_{\theta+\varepsilon}) = \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \sum_{x \in \mathcal{X}} p_\theta(x) \ln \left(\frac{p_\theta(x)}{p_{\theta+\varepsilon}(x)} \right) \tag{A.19}$$

$$= \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \left(\sum_{x \in \mathcal{X}} p_\theta(x) \ln(p_\theta(x)) - \sum_{x \in \mathcal{X}} p_\theta(x) \ln(p_{\theta+\varepsilon}(x)) \right) \tag{A.20}$$

$$= - \sum_{x \in \mathcal{X}} p_\theta(x) \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \ln(p_{\theta+\varepsilon}(x)) \tag{A.21}$$

$$= - \sum_{x \in \mathcal{X}} p_\theta(x) \frac{\partial}{\partial \varepsilon_i} \left(\frac{\frac{\partial}{\partial \varepsilon_j} p_{\theta+\varepsilon}(x)}{p_{\theta+\varepsilon}(x)} \right) \tag{A.22}$$

$$= \sum_{x \in \mathcal{X}} p_\theta(x) \left(- \frac{\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} p_{\theta+\varepsilon}(x)}{p_{\theta+\varepsilon}(x)} + \frac{\left(\frac{\partial}{\partial \varepsilon_i} p_{\theta+\varepsilon}(x) \right) \left(\frac{\partial}{\partial \varepsilon_j} p_{\theta+\varepsilon}(x) \right)}{(p_{\theta+\varepsilon}(x))^2} \right). \tag{A.23}$$

Then it follows that

$$\left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} D(p_\theta \| p_{\theta+\varepsilon}) \right|_{\varepsilon=0} \tag{A.24}$$

$$= \sum_{x \in \mathcal{X}} p_\theta(x) \left(- \frac{\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} p_{\theta+\varepsilon}(x)}{p_{\theta+\varepsilon}(x)} + \frac{\left(\frac{\partial}{\partial \varepsilon_i} p_{\theta+\varepsilon}(x) \right) \left(\frac{\partial}{\partial \varepsilon_j} p_{\theta+\varepsilon}(x) \right)}{(p_{\theta+\varepsilon}(x))^2} \right) \Big|_{\varepsilon=0} \tag{A.25}$$

$$= \sum_{x \in \mathcal{X}} p_\theta(x) \left(-\frac{\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} p_{\theta+\varepsilon}(x) \Big|_{\varepsilon=0}}{p_\theta(x)} + \frac{\left(\frac{\partial}{\partial \varepsilon_i} p_{\theta+\varepsilon}(x) \Big|_{\varepsilon=0} \right) \left(\frac{\partial}{\partial \varepsilon_j} p_{\theta+\varepsilon}(x) \Big|_{\varepsilon=0} \right)}{(p_\theta(x))^2} \right) \quad (\text{A.26})$$

$$= \sum_{x \in \mathcal{X}} p_\theta(x) \left(-\frac{\frac{\partial^2}{\partial \theta_i \partial \theta_j} p_\theta(x)}{p_\theta(x)} + \frac{\left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right)}{(p_\theta(x))^2} \right) \quad (\text{A.27})$$

$$= -\sum_{x \in \mathcal{X}} \frac{\partial^2}{\partial \theta_i \partial \theta_j} p_\theta(x) + \sum_{x \in \mathcal{X}} p_\theta(x)^{-1} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right) \quad (\text{A.28})$$

$$= -\frac{\partial^2}{\partial \theta_i \partial \theta_j} \sum_{x \in \mathcal{X}} p_\theta(x) + \sum_{x \in \mathcal{X}} p_\theta(x)^{-1} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right) \quad (\text{A.29})$$

$$= \sum_{x \in \mathcal{X}} p_\theta(x)^{-1} \left(\frac{\partial}{\partial \theta_i} p_\theta(x) \right) \left(\frac{\partial}{\partial \theta_j} p_\theta(x) \right), \quad (\text{A.30})$$

thus completing the proof of (2.14).

B Review: Matrix derivatives from divided differences

This appendix provides a review of the method of divided differences for calculating matrix derivatives. See also [Bha97, Section V.3] and [Hia10, Sections 2.2 and 2.3].

For a general d -dimensional matrix $A(x)$, the matrix derivative $\frac{\partial}{\partial x} A(x)$ is defined as

$$\frac{\partial}{\partial x} A(x) := \begin{bmatrix} \frac{\partial}{\partial x} a_{1,1}(x) & \frac{\partial}{\partial x} a_{1,2}(x) & \cdots & \frac{\partial}{\partial x} a_{1,d}(x) \\ \frac{\partial}{\partial x} a_{2,1}(x) & \frac{\partial}{\partial x} a_{2,2}(x) & & \frac{\partial}{\partial x} a_{2,d}(x) \\ \vdots & & \ddots & \vdots \\ \frac{\partial}{\partial x} a_{d,1}(x) & \frac{\partial}{\partial x} a_{d,2}(x) & \cdots & \frac{\partial}{\partial x} a_{d,d}(x) \end{bmatrix}. \quad (\text{B.1})$$

where the matrix elements of $A(x)$ are denoted by $a_{i,j}(x)$. From this, one can deduce that the product rule holds for matrices $A(x)$ and $B(x)$:

$$\frac{\partial}{\partial x} [A(x)B(x)] = \left[\frac{\partial}{\partial x} A(x) \right] B(x) + A(x) \left[\frac{\partial}{\partial x} B(x) \right]. \quad (\text{B.2})$$

B.1 First derivative

Let $A(x)$ be a Hermitian matrix parameterized by $x \in \mathbb{R}$. Suppose that $A(x)$ has a spectral decomposition as follows:

$$A(x) = \sum_{\ell} \lambda_{\ell} \Pi_{\ell}, \quad (\text{B.3})$$

where we have suppressed the dependence of each eigenvalue λ_{ℓ} and eigenprojection Π_{ℓ} on the parameter x .

Theorem 42. Let $x \mapsto A(x)$ be a Hermitian operator-valued function, and let f be an analytic function that has a power series expansion convergent on an open interval $I \subseteq \mathbb{R}$, such that, for all x , all of the eigenvalues of $A(x)$ are contained in I . Then the following equality holds:

$$\frac{\partial}{\partial x} f(A(x)) = \sum_{\ell, m} f^{[1]}(\lambda_\ell, \lambda_m) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m, \quad (\text{B.4})$$

where the first divided difference function $f^{[1]}(y_1, y_2)$ is defined for $y_1, y_2 \in I$ as

$$f^{[1]}(y_1, y_2) := \begin{cases} f'(y_1) & : y_1 = y_2 \\ \frac{f(y_1) - f(y_2)}{y_1 - y_2} & : y_1 \neq y_2 \end{cases}. \quad (\text{B.5})$$

Proof. By applying the product rule for differentiation, the following chain of equalities holds for all $n \in \mathbb{N}$:

$$\frac{\partial}{\partial x} (A(x)^n) = \sum_{k=0}^{n-1} A(x)^k \left(\frac{\partial}{\partial x} A(x) \right) A(x)^{n-k-1} \quad (\text{B.6})$$

$$= \sum_{k=0}^{n-1} \left(\sum_{\ell} \lambda_\ell^k \Pi_\ell \right)^k \left(\frac{\partial}{\partial x} A(x) \right) \left(\sum_m \lambda_m^{n-k-1} \Pi_m \right) \quad (\text{B.7})$$

$$= \sum_{k=0}^{n-1} \left(\sum_{\ell} \lambda_\ell^k \Pi_\ell \right) \left(\frac{\partial}{\partial x} A(x) \right) \left(\sum_m \lambda_m^{n-k-1} \Pi_m \right) \quad (\text{B.8})$$

$$= \sum_{\ell, m} \left(\sum_{k=0}^{n-1} \lambda_\ell^k \lambda_m^{n-k-1} \right) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.9})$$

$$= \sum_{\ell, m} f_{x^n}^{[1]}(\lambda_\ell, \lambda_m) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.10})$$

where the last equality follows because

$$\sum_{k=0}^{n-1} \lambda_\ell^k \lambda_m^{n-k-1} = f_{x^n}^{[1]}(\lambda_\ell, \lambda_m) := \begin{cases} n \lambda_\ell^{n-1} & : \lambda_\ell = \lambda_m \\ \frac{\lambda_\ell^n - \lambda_m^n}{\lambda_\ell - \lambda_m} & : \lambda_\ell \neq \lambda_m \end{cases}. \quad (\text{B.11})$$

The notation $f_{x^n}^{[1]}(\lambda_\ell, \lambda_m)$ means that this function is the first divided difference for $x \mapsto x^n$. To see this, if $\lambda_\ell = \lambda_m$, then

$$\sum_{k=0}^{n-1} \lambda_\ell^k \lambda_m^{n-k-1} = \sum_{k=0}^{n-1} \lambda_\ell^{n-1} = n \lambda_\ell^{n-1}, \quad (\text{B.12})$$

and if $\lambda_\ell \neq \lambda_m$, then

$$\sum_{k=0}^{n-1} \lambda_\ell^k \lambda_m^{n-k-1} = \lambda_m^{n-1} \left[\sum_{k=0}^{n-1} \left(\frac{\lambda_\ell}{\lambda_m} \right)^k \right] \quad (\text{B.13})$$

$$= \lambda_m^{n-1} \left(\frac{1 - \left(\frac{\lambda_\ell}{\lambda_m} \right)^n}{1 - \frac{\lambda_\ell}{\lambda_m}} \right) \quad (\text{B.14})$$

$$= \frac{\lambda_m^n - \lambda_\ell^n}{\lambda_m - \lambda_\ell}. \quad (\text{B.15})$$

Thus, $\sum_{k=0}^{n-1} \lambda_\ell^k \lambda_m^{n-k-1}$ is indeed the first divided difference for the function $x \mapsto x^n$.

We now extend this development to an arbitrary analytic function that has a power series expansion convergent on an open interval $I \subseteq \mathbb{R}$. Toward this, let $y \in I$, and let f be a function that has the following power series expansion centered at $c \in I$ and convergent on I :

$$f(y) = \sum_{n=0}^{\infty} a_n (y - c)^n. \quad (\text{B.16})$$

Then

$$f(A(x)) = \sum_{n=0}^{\infty} a_n (A(x) - cI)^n. \quad (\text{B.17})$$

By applying (B.10) and noting that $\frac{\partial}{\partial x} (A(x) - cI) = \frac{\partial}{\partial x} A(x)$ and $A(x) - cI = \sum_\ell (\lambda_\ell - c) \Pi_\ell$, we conclude that

$$\frac{\partial}{\partial x} f(A(x)) = \frac{\partial}{\partial x} \sum_{n=0}^{\infty} a_n (A(x) - cI)^n \quad (\text{B.18})$$

$$= \sum_{n=0}^{\infty} a_n \frac{\partial}{\partial x} ((A(x) - cI)^n) \quad (\text{B.19})$$

$$= \sum_{n=0}^{\infty} a_n \sum_{\ell, m} f_{x^n}^{[1]}(\lambda_\ell - c, \lambda_m - c) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.20})$$

$$= \sum_{\ell, m} \left(\sum_{n=0}^{\infty} a_n f_{x^n}^{[1]}(\lambda_\ell - c, \lambda_m - c) \right) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m. \quad (\text{B.21})$$

The last equality follows because, if $\lambda_\ell = \lambda_m$, then

$$\sum_{n=0}^{\infty} a_n f_{x^n}^{[1]}(\lambda_\ell - c, \lambda_m - c) = \sum_{n=0}^{\infty} a_n n (\lambda_\ell - c)^{n-1} \quad (\text{B.22})$$

$$= \sum_{n=0}^{\infty} a_n \frac{\partial}{\partial \lambda_\ell} (\lambda_\ell - c)^n \quad (\text{B.23})$$

$$= \frac{\partial}{\partial \lambda_\ell} \sum_{n=0}^{\infty} a_n (\lambda_\ell - c)^n \quad (\text{B.24})$$

$$= \frac{\partial}{\partial \lambda_\ell} f(\lambda_\ell) \quad (\text{B.25})$$

$$= f'(\lambda_\ell), \quad (\text{B.26})$$

and if $\lambda_\ell \neq \lambda_m$, then

$$\begin{aligned} & \sum_{n=0}^{\infty} a_n f_{x^n}^{[1]}(\lambda_\ell - c, \lambda_m - c) \\ &= \sum_{n=0}^{\infty} a_n \left(\frac{(\lambda_m - c)^n - (\lambda_\ell - c)^n}{\lambda_m - \lambda_\ell} \right) \end{aligned} \quad (\text{B.27})$$

$$= \frac{(\sum_{n=0}^{\infty} a_n (\lambda_m - c)^n) - (\sum_{n=0}^{\infty} a_n (\lambda_\ell - c)^n)}{\lambda_m - \lambda_\ell} \quad (\text{B.28})$$

$$= \frac{f(\lambda_m) - f(\lambda_\ell)}{\lambda_m - \lambda_\ell} \quad (\text{B.29})$$

$$= f^{[1]}(\lambda_\ell, \lambda_m). \quad (\text{B.30})$$

This concludes the proof. $\square$

Corollary 43. *Let $x \mapsto A(x)$ be a Hermitian operator-valued function, and let f be an analytic function that has a power series expansion convergent on an open interval $I \subseteq \mathbb{R}$, such that, for all x , all of the eigenvalues of $A(x)$ are contained in I . Then the following equality holds:*

$$\text{Tr} \left[\frac{\partial}{\partial x} f(A(x)) \right] = \text{Tr} \left[f'(A(x)) \frac{\partial}{\partial x} A(x) \right]. \quad (\text{B.31})$$

Proof. Applying Theorem 42, consider that

$$\text{Tr} \left[\frac{\partial}{\partial x} f(A(x)) \right] = \text{Tr} \left[\sum_{\ell, m} f^{[1]}(\lambda_\ell, \lambda_m) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \right], \quad (\text{B.32})$$

$$= \sum_{\ell, m} f^{[1]}(\lambda_\ell, \lambda_m) \text{Tr} \left[\Pi_m \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \right] \quad (\text{B.33})$$

$$= \sum_{\ell} f^{[1]}(\lambda_\ell, \lambda_\ell) \text{Tr} \left[\Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \right] \quad (\text{B.34})$$

$$= \sum_{\ell} f'(\lambda_\ell) \text{Tr} \left[\Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \right] \quad (\text{B.35})$$

$$= \text{Tr} \left[\sum_{\ell} f'(\lambda_\ell) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \right] \quad (\text{B.36})$$

$$= \text{Tr} \left[f'(A(x)) \frac{\partial}{\partial x} A(x) \right], \quad (\text{B.37})$$

thus concluding the proof. $\square$

Proposition 44. *The following expression holds for a holomorphic function f and for a contour γ containing all the eigenvalues of the Hermitian operator-valued function $x \mapsto A(x)$:*

$$\frac{\partial}{\partial x} f(A(x)) = \frac{1}{2\pi i} \oint_{\gamma} dz f(z) (zI - A(x))^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (zI - A(x))^{-1}. \quad (\text{B.38})$$

Proof. The Cauchy integral formula implies that the following equality holds for a holomorphic function f and a contour γ containing $a \in \mathbb{C}$:

$$f(a) = \frac{1}{2\pi i} \oint_{\gamma} dz \frac{f(z)}{z - a}. \quad (\text{B.39})$$

We can then apply this to establish the following formula for the first divided difference at $a, b \in \mathbb{C}$ and now with γ a contour containing a and b :

$$\frac{f(a) - f(b)}{a - b} = \left(\frac{1}{a - b} \right) \left(\frac{1}{2\pi i} \oint_{\gamma} dz \frac{f(z)}{z - a} - \frac{1}{2\pi i} \oint_{\gamma} dz \frac{f(z)}{z - b} \right) \quad (\text{B.40})$$

$$= \left(\frac{1}{a - b} \right) \frac{1}{2\pi i} \oint_{\gamma} dz f(z) \left(\frac{1}{z - a} - \frac{1}{z - b} \right) \quad (\text{B.41})$$

$$= \left(\frac{1}{a - b} \right) \frac{1}{2\pi i} \oint_{\gamma} dz f(z) \left(\frac{z - b - (z - a)}{(z - a)(z - b)} \right) \quad (\text{B.42})$$

$$= \frac{1}{2\pi i} \oint_{\gamma} dz \frac{f(z)}{(z - a)(z - b)}. \quad (\text{B.43})$$

Applying this and Theorem 42, we find that, when γ is a contour containing all the eigenvalues of $A(x)$,

$$\begin{aligned} & \frac{\partial}{\partial x} f(A(x)) \\ &= \sum_{\ell, m} f^{[1]}(\lambda_{\ell}, \lambda_m) \Pi_{\ell} \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \end{aligned} \quad (\text{B.44})$$

$$= \sum_{\ell, m} \left(\frac{1}{2\pi i} \oint_{\gamma} dz \frac{f(z)}{(z - \lambda_{\ell})(z - \lambda_m)} \right) \Pi_{\ell} \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.45})$$

$$= \frac{1}{2\pi i} \oint_{\gamma} dz f(z) \left(\sum_{\ell} \frac{1}{(z - \lambda_{\ell})} \Pi_{\ell} \right) \left(\frac{\partial}{\partial x} A(x) \right) \left(\sum_m \frac{1}{(z - \lambda_m)} \Pi_m \right) \quad (\text{B.46})$$

$$= \frac{1}{2\pi i} \oint_{\gamma} dz f(z) (zI - A(x))^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (zI - A(x))^{-1}. \quad (\text{B.47})$$

This concludes the proof. $\square$

We can also consider the case of operator monotone functions. Recall from [Bha97, pp. 144–145] the following representation theorem regarding operator monotone functions:

Theorem 45. A function $f: (0, \infty) \rightarrow \mathbb{R}$ is operator monotone on $(0, \infty)$ if and only if it has the representation

$$f(t) = a + bt + \int_0^\infty d\mu(v) \frac{tv}{t+v}, \quad (\text{B.48})$$

where $a \in \mathbb{R}$, $b \geq 0$, and μ is a positive measure on $(0, \infty)$ such that

$$\int_0^\infty d\mu(v) \frac{v}{1+v} < +\infty. \quad (\text{B.49})$$

If f is operator monotone on $[0, \infty)$, then $a = f(0)$.

We can use this theorem to establish the following representation for the derivative of operator monotone functions.

Proposition 46. Let $x \mapsto A(x)$ be a positive operator-valued function. The following expression holds for the derivative of a function f that is operator monotone on $(0, \infty)$:

$$\frac{\partial}{\partial x} f(A(x)) = b \frac{\partial}{\partial x} A(x) + \int_0^\infty d\mu(v) v^2 (A(x) + vI)^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (A(x) + vI)^{-1}. \quad (\text{B.50})$$

Proof. For $x, y > 0$ such that $x \neq y$, consider that

$$f(x) - f(y) = \left(a + bx + \int_0^\infty d\mu(v) \frac{xv}{x+v} \right) \quad (\text{B.51})$$

$$- \left(a + by + \int_0^\infty d\mu(v) \frac{yv}{y+v} \right) \quad (\text{B.52})$$

$$= b(x - y) + \int_0^\infty d\mu(v) \left(\frac{xv}{x+v} - \frac{yv}{y+v} \right) \quad (\text{B.53})$$

$$= b(x - y) + \int_0^\infty d\mu(v) \frac{xv(y+v) - yv(x+v)}{(x+v)(y+v)} \quad (\text{B.54})$$

$$= b(x - y) + (x - y) \int_0^\infty d\mu(v) \frac{v^2}{(x+v)(y+v)} \quad (\text{B.55})$$

$$= (x - y) \left(b + \int_0^\infty d\mu(v) \frac{v^2}{(x+v)(y+v)} \right). \quad (\text{B.56})$$

Then it follows that

$$\frac{f(x) - f(y)}{x - y} = b + \int_0^\infty d\mu(v) \frac{v^2}{(x+v)(y+v)}, \quad (\text{B.57})$$

and furthermore that

$$f'(x) = b + \int_0^\infty d\mu(v) \frac{v^2}{(x+v)^2}. \quad (\text{B.58})$$

By applying Theorem 42, consider that

$$\begin{aligned} & \frac{\partial}{\partial x} f(A(x)) \\ &= \sum_{\ell, m} f^{[1]}(\lambda_\ell, \lambda_m) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \end{aligned} \quad (\text{B.59})$$

$$= \sum_{\ell, m} \left(b + \int_0^\infty d\mu(v) \frac{v^2}{(\lambda_\ell + v)(\lambda_m + v)} \right) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.60})$$

$$\begin{aligned} &= b \sum_{\ell, m} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \\ &\quad \int_0^\infty d\mu(v) v^2 \left(\sum_\ell \frac{1}{(\lambda_\ell + v)} \Pi_\ell \right) \left(\frac{\partial}{\partial x} A(x) \right) \left(\sum_m \frac{1}{(\lambda_m + v)} \Pi_m \right) \end{aligned} \quad (\text{B.61})$$

$$= b \frac{\partial}{\partial x} A(x) + \int_0^\infty d\mu(v) v^2 (A(x) + vI)^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (A(x) + vI)^{-1}, \quad (\text{B.62})$$

thus concluding the proof. $\square$

B.2 Examples

We can verify the consistency of the first divided difference expression for the derivative with other known expressions for derivatives. As a first example, let us verify Duhamel's formula:

Proposition 47 (Exponential function). *For a Hermitian operator-valued function $x \mapsto A(x)$, the following identity holds:*

$$\frac{\partial}{\partial x} e^{A(x)} = \int_0^1 dt e^{tA(x)} \left(\frac{\partial}{\partial x} A(x) \right) e^{(1-t)A(x)}. \quad (\text{B.63})$$

Proof. Suppose that $A(x)$ has a spectral decomposition as in (B.3). Then

$$\begin{aligned} & \int_0^1 dt e^{tA(x)} \left(\frac{\partial}{\partial x} A(x) \right) e^{(1-t)A(x)} \\ &= \int_0^1 dt \left(\sum_\ell e^{t\lambda_\ell} \Pi_\ell \right) \left(\frac{\partial}{\partial x} A(x) \right) \left(\sum_m e^{(1-t)\lambda_m} \Pi_m \right) \end{aligned} \quad (\text{B.64})$$

$$= \sum_{\ell, m} \int_0^1 dt e^{t\lambda_\ell} e^{(1-t)\lambda_m} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.65})$$

$$= \sum_{\ell, m} e^{\lambda_m} \int_0^1 dt e^{t(\lambda_\ell - \lambda_m)} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.66})$$

$$= \sum_\ell e^{\lambda_m} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_\ell$$

$$+ \sum_{\ell \neq m} e^{\lambda_m} \int_0^1 dt e^{t(\lambda_\ell - \lambda_m)} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m. \quad (\text{B.67})$$

Now consider that

$$\begin{aligned} & \sum_{\ell \neq m} e^{\lambda_m} \int_0^1 dt e^{t(\lambda_\ell - \lambda_m)} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \\ &= \sum_{\ell \neq m} e^{\lambda_m} \left[\frac{e^{t(\lambda_\ell - \lambda_m)}}{\lambda_\ell - \lambda_m} \Big|_0^1 \right] \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \end{aligned} \quad (\text{B.68})$$

$$= \sum_{\ell \neq m} e^{\lambda_m} \left[\frac{e^{(\lambda_\ell - \lambda_m)}}{\lambda_\ell - \lambda_m} - \frac{1}{\lambda_\ell - \lambda_m} \right] \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.69})$$

$$= \sum_{\ell \neq m} \frac{e^{\lambda_\ell} - e^{\lambda_m}}{\lambda_\ell - \lambda_m} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m. \quad (\text{B.70})$$

Substituting back into (B.67), we conclude that

$$\begin{aligned} & \int_0^1 dt e^{tA(x)} \left(\frac{\partial}{\partial x} A(x) \right) e^{(1-t)A(x)} \\ &= \sum_\ell e^{\lambda_m} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_\ell + \sum_{\ell \neq m} \frac{e^{\lambda_\ell} - e^{\lambda_m}}{\lambda_\ell - \lambda_m} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \end{aligned} \quad (\text{B.71})$$

$$= \sum_{\ell, m} f_{e^x}^{[1]}(\lambda_\ell, \lambda_m) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.72})$$

$$= \frac{\partial}{\partial x} e^{A(x)}, \quad (\text{B.73})$$

where $f_{e^x}^{[1]}$ denotes the first divided difference for the function $x \mapsto e^x$ and the last equality follows from Theorem 42 and the fact that the Taylor series expansion $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges for all $x \in \mathbb{R}$. $\square$

We can also verify a well known expression for the derivative of the logarithm:

Proposition 48 (Logarithmic function). *For a positive operator-valued function $x \mapsto A(x)$, the following identity holds:*

$$\frac{\partial}{\partial x} \ln A(x) = \int_0^\infty ds (A(x) + sI)^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (A(x) + sI)^{-1}. \quad (\text{B.74})$$

Proof. Suppose that $A(x)$ has a spectral decomposition as in (B.3). Then

$$\int_0^\infty ds (A(x) + sI)^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (A(x) + sI)^{-1}$$

$$= \int_0^\infty ds \left(\sum_\ell \left(\frac{1}{\lambda_\ell + s} \right) \Pi_\ell \right) \left(\frac{\partial}{\partial x} A(x) \right) \left(\sum_m \left(\frac{1}{\lambda_m + s} \right) \Pi_m \right) \quad (\text{B.75})$$

$$= \sum_{\ell, m} \left[\int_0^\infty ds \left(\frac{1}{\lambda_\ell + s} \right) \left(\frac{1}{\lambda_m + s} \right) \right] \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.76})$$

$$= \sum_\ell \left[\int_0^\infty ds \frac{1}{(\lambda_\ell + s)^2} \right] \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_\ell + \sum_{\ell \neq m} \left[\int_0^\infty ds \left(\frac{1}{\lambda_\ell + s} \right) \left(\frac{1}{\lambda_m + s} \right) \right] \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.77})$$

$$= \sum_\ell \frac{1}{\lambda_\ell} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_\ell + \sum_{\ell \neq m} \frac{\ln \lambda_\ell - \ln \lambda_m}{\lambda_\ell - \lambda_m} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.78})$$

$$= \sum_{\ell, m} f_{\ln x}^{[1]}(\lambda_\ell, \lambda_m) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.79})$$

$$= \frac{\partial}{\partial x} \ln A(x), \quad (\text{B.80})$$

where $f_{\ln x}^{[1]}$ denotes the first divided difference for the function $x \mapsto \ln x$. The penultimate equality is a consequence of the following integral formulas, holding for all $x, y > 0$:

$$\int_0^\infty ds \frac{1}{(x+s)^2} = \frac{1}{x}, \quad (\text{B.81})$$

$$\int_0^\infty ds \left(\frac{1}{x+s} \right) \left(\frac{1}{y+s} \right) = \frac{\ln x - \ln y}{x - y}, \quad (\text{B.82})$$

and the last equality follows from Theorem 42 and the fact that the logarithm has the following power series expansion that converges for all $y \in (0, 2c)$, where $c > 0$:

$$\ln y = \ln c + \sum_{n=1}^\infty \frac{(-1)^{n-1}}{n} \left(\frac{y-c}{c} \right)^n. \quad (\text{B.83})$$

To see this, we can write

$$\ln y = \ln \left(c \left(1 + \frac{y-c}{c} \right) \right) \quad (\text{B.84})$$

$$= \ln c + \ln \left(1 + \frac{y-c}{c} \right) \quad (\text{B.85})$$

and then apply the standard Taylor expansion for $\ln(1+y)$ that converges for all $|y| < 1$. Substituting $y \rightarrow \frac{y-c}{c}$, the convergence condition becomes $|\frac{y-c}{c}| < 1$, which is equivalent to $y \in (0, 2c)$. In order to apply Theorem 42, we take $c > 0$ to be larger than the largest eigenvalue of $A(x)$. $\square$

We now use Propositions 47 and 48 to find an integral formula for the derivative of all power functions.

Proposition 49 (Power function). *Let $x \mapsto A(x)$ be a positive operator-valued function. For all $r \in \mathbb{R}$, the following equality holds:*

$$\frac{\partial}{\partial x} (A(x)^r) = r \int_0^1 dt \int_0^\infty ds \frac{A(x)^{rt}}{A(x) + sI} \left(\frac{\partial}{\partial x} A(x) \right) \frac{A(x)^{r(1-t)}}{A(x) + sI}. \quad (\text{B.86})$$

Proof. Applying Propositions 47 and 48, consider that

$$\begin{aligned} & \frac{\partial}{\partial x} (A(x)^r) \\ &= \frac{\partial}{\partial x} e^{r \ln A(x)} \end{aligned} \quad (\text{B.87})$$

$$= \int_0^1 dt e^{tr \ln A(x)} \left(\frac{\partial}{\partial x} r \ln A(x) \right) e^{(1-t)r \ln A(x)} \quad (\text{B.88})$$

$$= r \int_0^1 dt A(x)^{rt} \left(\frac{\partial}{\partial x} \ln A(x) \right) A(x)^{r(1-t)} \quad (\text{B.89})$$

$$= r \int_0^1 dt A(x)^{rt} \left(\frac{\partial}{\partial x} \ln A(x) \right) A(x)^{r(1-t)} \quad (\text{B.90})$$

$$= r \int_0^1 dt A(x)^{rt} \int_0^\infty ds (A(x) + sI)^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (A(x) + sI)^{-1} A(x)^{r(1-t)} \quad (\text{B.91})$$

$$= r \int_0^1 dt \int_0^\infty ds \frac{A(x)^{rt}}{A(x) + sI} \left(\frac{\partial}{\partial x} A(x) \right) \frac{A(x)^{r(1-t)}}{A(x) + sI}, \quad (\text{B.92})$$

thus concluding the proof. $\square$

By taking a spectral decomposition of $A(x)$ as in (B.3), we can further manipulate the expression from Proposition 49 to see that it is consistent with the first divided difference expression from Theorem 42:

$$\begin{aligned} & \frac{\partial}{\partial x} (A(x)^r) \\ &= r \int_0^1 dt \int_0^\infty ds \frac{A(x)^{rt}}{A(x) + sI} \left(\frac{\partial}{\partial x} A(x) \right) \frac{A(x)^{r(1-t)}}{A(x) + sI} \end{aligned} \quad (\text{B.93})$$

$$= r \int_0^1 dt \int_0^\infty ds \left(\sum_\ell \frac{\lambda_\ell^{rt}}{\lambda_\ell + sI} \Pi_\ell \right) \left(\frac{\partial}{\partial x} A(x) \right) \left(\sum_m \frac{\lambda_m^{r(1-t)}}{\lambda_m + sI} \Pi_m \right) \quad (\text{B.94})$$

$$= r \sum_{\ell, m} \int_0^1 dt \lambda_\ell^{rt} \lambda_m^{r(1-t)} \int_0^\infty ds \left(\frac{1}{\lambda_\ell + sI} \right) \left(\frac{1}{\lambda_m + sI} \right) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.95})$$

$$= r \sum_\ell \lambda_\ell^r \int_0^\infty ds \frac{1}{(\lambda_\ell + sI)^2} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_\ell$$

$$+ r \sum_{\ell \neq m} \int_0^1 dt \lambda_\ell^{rt} \lambda_m^{r(1-t)} \int_0^\infty ds \left(\frac{1}{\lambda_\ell + sI} \right) \left(\frac{1}{\lambda_m + sI} \right) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.96})$$

$$= \sum_\ell r \lambda_\ell^{r-1} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_\ell + r \sum_{\ell \neq m} \left(\frac{\lambda_\ell^r - \lambda_m^r}{\ln \lambda_\ell - \ln \lambda_m} \right) \left(\frac{\ln \lambda_\ell - \ln \lambda_m}{\lambda_\ell - \lambda_m} \right) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.97})$$

$$= \sum_\ell r \lambda_\ell^{r-1} \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_\ell + \sum_{\ell \neq m} \left(\frac{\lambda_\ell^r - \lambda_m^r}{\lambda_\ell - \lambda_m} \right) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.98})$$

$$= \sum_{\ell, m} f_{x^r}^{[1]}(\lambda_\ell, \lambda_m) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m, \quad (\text{B.99})$$

where $f_{x^r}^{[1]}$ denotes the first divided difference for the function $x \mapsto x^r$.

There is an alternative simpler expression for the derivative of the power function when the power $r \in (0, 1)$:

Proposition 50 (Power function). *For a positive operator-valued function $x \mapsto A(x)$, the following equality holds for all $r \in (-1, 0) \cup (0, 1)$:*

$$\frac{\partial}{\partial x} (A(x)^r) = \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^r (A(x) + tI)^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (A(x) + tI)^{-1}. \quad (\text{B.100})$$

Proof. The following integral representation holds for $r \in (0, 1)$ and $x > 0$ [Bha97, Exercise V.4.20]:

$$x^r = \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^{r-1} \left(\frac{x}{x+t} \right). \quad (\text{B.101})$$

From this, we can conclude the following integral representation for $r \in (0, 1)$ and $x, y > 0$ such that $x \neq y$:

$$\frac{x^r - y^r}{x - y} = \frac{\sin(r\pi)}{\pi} \int_0^\infty dt \frac{t^r}{(x+t)(y+t)}. \quad (\text{B.102})$$

Indeed, consider that

$$\begin{aligned} & \frac{x^r - y^r}{x - y} \\ &= \left(\frac{1}{x - y} \right) \left(\frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^{r-1} \left(\frac{x}{x+t} \right) - \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^{r-1} \left(\frac{y}{y+t} \right) \right) \end{aligned} \quad (\text{B.103})$$

$$= \left(\frac{1}{x - y} \right) \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^{r-1} \left(\frac{x}{x+t} - \frac{y}{y+t} \right) \quad (\text{B.104})$$

$$= \left(\frac{1}{x-y} \right) \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^{r-1} \left(\frac{x(y+t) - y(x+t)}{(x+t)(y+t)} \right) \quad (\text{B.105})$$

$$= \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^{r-1} \left(\frac{t}{(x+t)(y+t)} \right) \quad (\text{B.106})$$

$$= \frac{\sin(r\pi)}{\pi} \int_0^\infty dt \frac{t^r}{(x+t)(y+t)}. \quad (\text{B.107})$$

This allows us to write, for $r \in (0, 1)$,

$$\begin{aligned} & \frac{\partial}{\partial x} (A(x)^r) \\ &= \sum_{\ell, m} f_{x^r}^{[1]}(\lambda_\ell, \lambda_m) \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \end{aligned} \quad (\text{B.108})$$

$$= \sum_{\ell, m} \left[\frac{\sin(r\pi)}{\pi} \int_0^\infty dt \frac{t^r}{(\lambda_\ell + t)(\lambda_m + t)} \right] \Pi_\ell \left(\frac{\partial}{\partial x} A(x) \right) \Pi_m \quad (\text{B.109})$$

$$= \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^r \left(\sum_\ell \frac{1}{(\lambda_\ell + t)} \Pi_\ell \right) \left(\frac{\partial}{\partial x} A(x) \right) \left(\sum_m \frac{1}{(\lambda_m + t)} \Pi_m \right) \quad (\text{B.110})$$

$$= \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^r (A(x) + tI)^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (A(x) + tI)^{-1}. \quad (\text{B.111})$$

The first equality follows from applying Theorem 42 and the fact that the power function has the following power series expansion that converges for all $y \in (0, 2c)$, where $c > 0$:

$$y^r = c^r \sum_{n=0}^\infty \binom{r}{n} \left(\frac{y-c}{c} \right)^n. \quad (\text{B.112})$$

To see this, we can write

$$y^r = \left(c \left(1 + \frac{y-c}{c} \right) \right)^r \quad (\text{B.113})$$

$$= c^r \left(1 + \frac{y-c}{c} \right)^r \quad (\text{B.114})$$

and then apply the standard binomial series for $(1+y)^r$ that converges for all $|y| < 1$. Substituting $y \rightarrow \frac{y-c}{c}$, the convergence condition becomes $\left| \frac{y-c}{c} \right| < 1$, which is equivalent to $y \in (0, 2c)$. In order to apply Theorem 42, we take $c > 0$ to be larger than the largest eigenvalue of $A(x)$. This concludes the proof of (B.100) for $r \in (0, 1)$.

Dividing the integral representation in (B.101) by $x > 0$ gives

$$x^{r-1} = \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^{r-1} \left(\frac{1}{x+t} \right). \quad (\text{B.115})$$

This is then equivalent to the following for $r \in (-1, 0)$ and $x > 0$:

$$x^r = \frac{\sin((r+1)\pi)}{\pi} \int_0^\infty dt t^r \left(\frac{1}{x+t} \right). \quad (\text{B.116})$$

$$= -\frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^r \left(\frac{1}{x+t} \right) \quad (\text{B.117})$$

Then it follows that, for $r \in (-1, 0)$ and $x, y > 0$ such that $x \neq y$:

$$\frac{x^r - y^r}{x - y} = \left(\frac{-\frac{\sin(r\pi)}{\pi}}{x - y} \right) \left(\int_0^\infty dt t^r \left(\frac{1}{x+t} \right) - \int_0^\infty dt t^r \left(\frac{1}{y+t} \right) \right) \quad (\text{B.118})$$

$$= \left(\frac{\frac{\sin(r\pi)}{\pi}}{y - x} \right) \int_0^\infty dt t^r \left(\frac{1}{x+t} - \frac{1}{y+t} \right) \quad (\text{B.119})$$

$$= \left(\frac{\frac{\sin(r\pi)}{\pi}}{y - x} \right) \int_0^\infty dt t^r \left(\frac{y+t - (x+t)}{(x+t)(y+t)} \right) \quad (\text{B.120})$$

$$= \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^r \frac{1}{(x+t)(y+t)}. \quad (\text{B.121})$$

Now performing the same manipulations as in (B.108)–(B.111) and similar justification, we conclude the following for $r \in (-1, 0)$:

$$\frac{\partial}{\partial x} (A(x)^r) = \frac{\sin(r\pi)}{\pi} \int_0^\infty dt t^r (A(x) + tI)^{-1} \left(\frac{\partial}{\partial x} A(x) \right) (A(x) + tI)^{-1}, \quad (\text{B.122})$$

thus concluding the proof of (B.100) for $r \in (-1, 0)$. $\square$

C Proof of Equation (5.55)

Consider that for $x, y > 0$, such that $x \neq y$,

$$\begin{aligned} (xy)^{\frac{\alpha}{z}} & \left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x - y} \right)^2 - \left(\frac{x^{\frac{z-1+\alpha}{z}} - y^{\frac{z-1+\alpha}{z}}}{x - y} \right) \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x - y} \right) \\ &= \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{(x - y)^2} \right) \left[(xy)^{\frac{\alpha}{z}} \left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \left(x^{\frac{z-1+\alpha}{z}} - y^{\frac{z-1+\alpha}{z}} \right) \right] \end{aligned} \quad (\text{C.1})$$

$$= \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x - y} \right) \left(\frac{1}{x - y} \right) \left[(xy)^{\frac{\alpha}{z}} \left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \left(x^{\frac{z-1+\alpha}{z}} - y^{\frac{z-1+\alpha}{z}} \right) \right]. \quad (\text{C.2})$$

Now consider that

$$\left(\frac{1}{x - y} \right) \left[(xy)^{\frac{\alpha}{z}} \left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \left(x^{\frac{z-1+\alpha}{z}} - y^{\frac{z-1+\alpha}{z}} \right) \right]$$

$$= \left(\frac{(xy)^{\frac{\alpha}{z}}}{x-y} \right) \left[\left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \left(x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} - y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} \right) \right] \quad (\text{C.3})$$

$$= \left(\frac{(xy)^{\frac{\alpha}{z}}}{x-y} \right) \left[\left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \frac{\left(x^{\frac{1}{z}} - y^{\frac{1}{z}} \right) \left(x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} - y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} \right)}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right] \quad (\text{C.4})$$

$$= \left(\frac{(xy)^{\frac{\alpha}{z}}}{(x-y) \left(x^{\frac{1}{z}} - y^{\frac{1}{z}} \right)} \right) \times \left[\left(x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \left(x^{\frac{1}{z}} - y^{\frac{1}{z}} \right) \left(x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} - y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} \right) \right] \quad (\text{C.5})$$

Focusing on the term in the bottom line of (C.5), observe that

$$\begin{aligned} & \left(x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \left(x^{\frac{1}{z}} - y^{\frac{1}{z}} \right) \left(x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} - y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} \right) \\ &= \left(x^{\frac{z-1}{z}} \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - y^{\frac{z-1}{z}} \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) \right) \\ & \quad - \left(x^{\frac{1}{z}} \left(x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} - y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} \right) - y^{\frac{1}{z}} \left(x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} - y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} \right) \right) \end{aligned} \quad (\text{C.6})$$

$$\begin{aligned} &= \left(x^{\frac{z-1}{z}} x^{\frac{1-\alpha}{z}} - x^{\frac{z-1}{z}} y^{\frac{1-\alpha}{z}} - y^{\frac{z-1}{z}} x^{\frac{1-\alpha}{z}} + y^{\frac{z-1}{z}} y^{\frac{1-\alpha}{z}} \right) \\ & \quad - \left(x^{\frac{1}{z}} x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} - x^{\frac{1}{z}} y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} - y^{\frac{1}{z}} x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} + y^{\frac{1}{z}} y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} \right) \end{aligned} \quad (\text{C.7})$$

$$\begin{aligned} &= \left(x^{\frac{z-\alpha}{z}} - x^{\frac{z-1}{z}} y^{\frac{1-\alpha}{z}} - y^{\frac{z-1}{z}} x^{\frac{1-\alpha}{z}} + y^{\frac{z-\alpha}{z}} \right) \\ & \quad - \left(xy^{-\frac{\alpha}{z}} - y^{\frac{z-1}{z}} x^{\frac{1-\alpha}{z}} - x^{\frac{z-1}{z}} y^{\frac{1-\alpha}{z}} + yx^{-\frac{\alpha}{z}} \right) \end{aligned} \quad (\text{C.8})$$

$$= (xx^{-\frac{\alpha}{z}} + yy^{-\frac{\alpha}{z}}) - (xy^{-\frac{\alpha}{z}} + yx^{-\frac{\alpha}{z}}) \quad (\text{C.9})$$

$$= (x-y) \left(x^{-\frac{\alpha}{z}} - y^{-\frac{\alpha}{z}} \right). \quad (\text{C.10})$$

Plugging back into (C.5), we conclude that

$$\left(\frac{(xy)^{\frac{\alpha}{z}}}{(x-y) \left(x^{\frac{1}{z}} - y^{\frac{1}{z}} \right)} \right) \left[\left(x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \left(x^{\frac{1}{z}} - y^{\frac{1}{z}} \right) \left(x^{\frac{z-1}{z}} y^{-\frac{\alpha}{z}} - y^{\frac{z-1}{z}} x^{-\frac{\alpha}{z}} \right) \right]$$

$$= \left(\frac{(xy)^{\frac{\alpha}{z}}}{(x-y) \left(x^{\frac{1}{z}} - y^{\frac{1}{z}} \right)} \right) (x-y) \left(x^{-\frac{\alpha}{z}} - y^{-\frac{\alpha}{z}} \right) \quad (\text{C.11})$$

$$= \frac{(xy)^{\frac{\alpha}{z}} \left(x^{-\frac{\alpha}{z}} - y^{-\frac{\alpha}{z}} \right)}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \quad (\text{C.12})$$

$$= - \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right). \quad (\text{C.13})$$

We have thus now proved that

$$\begin{aligned} \left(\frac{1}{x-y}\right) \left[(xy)^{\frac{\alpha}{z}} \left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \left(x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}} \right) - \left(x^{\frac{z-1+\alpha}{z}} - y^{\frac{z-1+\alpha}{z}} \right) \right] \\ = - \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right). \quad (\text{C.14}) \end{aligned}$$

Finally, plugging back into (C.2), we conclude that

$$\begin{aligned} (xy)^{\frac{\alpha}{z}} \left(\frac{x^{\frac{z-1}{z}} - y^{\frac{z-1}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x-y} \right)^2 - \left(\frac{x^{\frac{z-1+\alpha}{z}} - y^{\frac{z-1+\alpha}{z}}}{x-y} \right) \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x-y} \right) \\ = - \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x-y} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right), \quad (\text{C.15}) \end{aligned}$$

thus completing the proof of (5.55).

D Proof of Equation (5.6)

Consider that

$$\left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x-y} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) = \left(\frac{y^{\frac{1-\alpha}{z}}}{y} \right) \left(\frac{\left(\frac{x}{y}\right)^{\frac{1-\alpha}{z}} - 1}{\frac{x}{y} - 1} \right) \left(\frac{y^{\frac{\alpha}{z}}}{y^{\frac{1}{z}}} \right) \left(\frac{\left(\frac{x}{y}\right)^{\frac{\alpha}{z}} - 1}{\left(\frac{x}{y}\right)^{\frac{1}{z}} - 1} \right) \quad (\text{D.1})$$

$$= \left(\frac{y^{\frac{1-\alpha}{z}}}{y} \right) \left(\frac{y^{\frac{\alpha}{z}}}{y^{\frac{1}{z}}} \right) \left(\frac{\left(\frac{x}{y}\right)^{\frac{1-\alpha}{z}} - 1}{\frac{x}{y} - 1} \right) \left(\frac{\left(\frac{x}{y}\right)^{\frac{\alpha}{z}} - 1}{\left(\frac{x}{y}\right)^{\frac{1}{z}} - 1} \right) \quad (\text{D.2})$$

$$= \left(\frac{1}{y} \right) \left(\frac{\left(\frac{x}{y}\right)^{\frac{1-\alpha}{z}} - 1}{\frac{x}{y} - 1} \right) \left(\frac{\left(\frac{x}{y}\right)^{\frac{\alpha}{z}} - 1}{\left(\frac{x}{y}\right)^{\frac{1}{z}} - 1} \right). \quad (\text{D.3})$$

Now consider that

$$\lim_{x \rightarrow y} \left(\frac{\left(\frac{x}{y}\right)^{\frac{1-\alpha}{z}} - 1}{\frac{x}{y} - 1} \right) \left(\frac{\left(\frac{x}{y}\right)^{\frac{\alpha}{z}} - 1}{\left(\frac{x}{y}\right)^{\frac{1}{z}} - 1} \right) = \lim_{x \rightarrow 1} \left(\frac{x^{\frac{1-\alpha}{z}} - 1}{x - 1} \right) \left(\frac{x^{\frac{\alpha}{z}} - 1}{x^{\frac{1}{z}} - 1} \right) \quad (\text{D.4})$$

$$= \left(\lim_{x \rightarrow 1} \frac{x^{\frac{1-\alpha}{z}} - 1}{x - 1} \right) \left(\lim_{x \rightarrow 1} \frac{x^{\frac{\alpha}{z}} - 1}{x^{\frac{1}{z}} - 1} \right) \quad (\text{D.5})$$

$$= \left(\lim_{x \rightarrow 1} \frac{\frac{1-\alpha}{z} x^{\frac{1-\alpha}{z}-1}}{1} \right) \left(\lim_{x \rightarrow 1} \frac{\frac{\alpha}{z} x^{\frac{\alpha}{z}-1}}{\frac{1}{z} x^{\frac{1}{z}-1}} \right) \quad (\text{D.6})$$

$$= \frac{\left(\frac{1-\alpha}{z}\right) \left(\frac{\alpha}{z}\right)}{\frac{1}{z}} \lim_{x \rightarrow 1} x^{\frac{1-\alpha}{z}-1+\frac{\alpha}{z}-1-(\frac{1}{z}-1)} \quad (\text{D.7})$$

$$= \frac{\alpha(1-\alpha)}{z} \lim_{x \rightarrow 1} x^{-1} \quad (\text{D.8})$$

$$= \frac{\alpha(1-\alpha)}{z}. \quad (\text{D.9})$$

Thus,

$$\lim_{x \rightarrow y} \left(\frac{x^{\frac{1-\alpha}{z}} - y^{\frac{1-\alpha}{z}}}{x - y} \right) \left(\frac{x^{\frac{\alpha}{z}} - y^{\frac{\alpha}{z}}}{x^{\frac{1}{z}} - y^{\frac{1}{z}}} \right) = \alpha \left(\frac{1-\alpha}{z} \right) \left(\frac{1}{y} \right). \quad (\text{D.10})$$

E Proof of Equation (5.83)

Consider that

$$\begin{aligned} & \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (-2 \ln F(\psi(\theta), \psi(\theta + \varepsilon))) \\ &= \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (-2 \ln \text{Tr}[\psi(\theta)\psi(\theta + \varepsilon)]) \end{aligned} \quad (\text{E.1})$$

$$= -2 \frac{\partial}{\partial \varepsilon_i} \left(\frac{\text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \varepsilon_j} \psi(\theta + \varepsilon) \right) \right]}{\text{Tr}[\psi(\theta)\psi(\theta + \varepsilon)]} \right) \quad (\text{E.2})$$

$$= -2 \left(\frac{\text{Tr} \left[\psi(\theta) \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \psi(\theta + \varepsilon) \right) \right]}{\text{Tr}[\psi(\theta)\psi(\theta + \varepsilon)]} - \frac{\text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \varepsilon_i} \psi(\theta + \varepsilon) \right) \right] \text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \varepsilon_j} \psi(\theta + \varepsilon) \right) \right]}{(\text{Tr}[\psi(\theta)\psi(\theta + \varepsilon)])^2} \right). \quad (\text{E.3})$$

Then it follows that

$$\begin{aligned} & \left. \frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (-2 \ln F(\psi(\theta), \psi(\theta + \varepsilon))) \right|_{\varepsilon=0} \\ &= -2 \left. \frac{\text{Tr} \left[\psi(\theta) \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \psi(\theta + \varepsilon) \right) \right]}{\text{Tr}[\psi(\theta)\psi(\theta + \varepsilon)]} \right|_{\varepsilon=0} \\ & \quad + 2 \left. \frac{\text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \varepsilon_i} \psi(\theta + \varepsilon) \right) \right] \text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \varepsilon_j} \psi(\theta + \varepsilon) \right) \right]}{(\text{Tr}[\psi(\theta)\psi(\theta + \varepsilon)])^2} \right|_{\varepsilon=0} \\ &= -2 \frac{\text{Tr} \left[\psi(\theta) \left(\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} \psi(\theta + \varepsilon) \right) \right]_{\varepsilon=0}}{\text{Tr}[\psi(\theta)\psi(\theta)]} \end{aligned} \quad (\text{E.4})$$

$$+ 2 \frac{\text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \varepsilon_i} \psi(\theta + \varepsilon) \Big|_{\varepsilon=0} \right) \right] \text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \varepsilon_j} \psi(\theta + \varepsilon) \Big|_{\varepsilon=0} \right) \right]}{(\text{Tr} [\psi(\theta) \psi(\theta)])^2} \quad (\text{E.5})$$

$$= -2 \left(\text{Tr} \left[\psi(\theta) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \psi(\theta) \right] - \text{Tr} \left[\psi(\theta) \frac{\partial}{\partial \theta_i} \psi(\theta) \right] \text{Tr} \left[\psi(\theta) \frac{\partial}{\partial \theta_j} \psi(\theta) \right] \right) \quad (\text{E.6})$$

Now, given that $\text{Tr}[\psi(\theta)\psi(\theta)] = 1$, consider that

$$0 = \frac{\partial}{\partial \theta_i} \text{Tr}[\psi(\theta)\psi(\theta)] \quad (\text{E.7})$$

$$= \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \psi(\theta) \right) \psi(\theta) \right] + \text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \theta_i} \psi(\theta) \right) \right] \quad (\text{E.8})$$

$$= 2 \text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \theta_i} \psi(\theta) \right) \right]. \quad (\text{E.9})$$

Thus,

$$\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j} (-2 \ln F(\psi(\theta), \psi(\theta + \varepsilon))) \Big|_{\varepsilon=0} = -2 \text{Tr} \left[\psi(\theta) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \psi(\theta) \right]. \quad (\text{E.10})$$

$$= 2 \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \psi(\theta) \right) \left(\frac{\partial}{\partial \theta_j} \psi(\theta) \right) \right], \quad (\text{E.11})$$

where the equality follows because

$$0 = \frac{\partial}{\partial \theta_i} \text{Tr} \left[\psi(\theta) \left(\frac{\partial}{\partial \theta_j} \psi(\theta) \right) \right] \quad (\text{E.12})$$

$$= \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \psi(\theta) \right) \left(\frac{\partial}{\partial \theta_j} \psi(\theta) \right) \right] + \text{Tr} \left[\psi(\theta) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \psi(\theta) \right]. \quad (\text{E.13})$$

Finally, while using $\partial_i \equiv \frac{\partial}{\partial \theta_i}$, observe that

$$\text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} \psi(\theta) \right) \left(\frac{\partial}{\partial \theta_j} \psi(\theta) \right) \right] \quad (\text{E.14})$$

$$= \text{Tr} \left[\left(\frac{\partial}{\partial \theta_i} |\psi(\theta)\rangle \langle \psi(\theta)| \right) \left(\frac{\partial}{\partial \theta_j} |\psi(\theta)\rangle \langle \psi(\theta)| \right) \right] \quad (\text{E.15})$$

$$= \text{Tr} [(|\partial_i \psi(\theta)\rangle \langle \psi(\theta)| + |\psi(\theta)\rangle \langle \partial_i \psi(\theta)|) (|\partial_j \psi(\theta)\rangle \langle \psi(\theta)| + |\psi(\theta)\rangle \langle \partial_j \psi(\theta)|)] \quad (\text{E.16})$$

$$= \text{Tr} [|\partial_i \psi(\theta)\rangle \langle \psi(\theta)| \partial_j \psi(\theta)\rangle \langle \psi(\theta)|] + \text{Tr} [|\partial_i \psi(\theta)\rangle \langle \psi(\theta)| \psi(\theta)\rangle \langle \partial_j \psi(\theta)|] \\ + \text{Tr} [|\psi(\theta)\rangle \langle \partial_i \psi(\theta)| \partial_j \psi(\theta)\rangle \langle \psi(\theta)|] + \text{Tr} [|\psi(\theta)\rangle \langle \partial_i \psi(\theta)| \psi(\theta)\rangle \langle \partial_j \psi(\theta)|] \quad (\text{E.17})$$

$$= \langle \psi(\theta) | \partial_i \psi(\theta) \rangle \langle \psi(\theta) | \partial_j \psi(\theta) \rangle + \langle \partial_j \psi(\theta) | \partial_i \psi(\theta) \rangle \\ + \langle \partial_i \psi(\theta) | \partial_j \psi(\theta) \rangle + \langle \partial_j \psi(\theta) | \psi(\theta) \rangle \langle \partial_i \psi(\theta) | \psi(\theta) \rangle \quad (\text{E.18})$$

$$= 2\Re [\langle \partial_i \psi(\theta) | \partial_j \psi(\theta) \rangle + \langle \psi(\theta) | \partial_i \psi(\theta) \rangle \langle \psi(\theta) | \partial_j \psi(\theta) \rangle] \quad (\text{E.19})$$

$$= 2\Re [\langle \partial_i \psi(\theta) | \partial_j \psi(\theta) \rangle - \langle \partial_i \psi(\theta) | \psi(\theta) \rangle \langle \psi(\theta) | \partial_j \psi(\theta) \rangle] \quad (\text{E.20})$$

$$= 2\Re [\langle \partial_i \psi(\theta) | (I - |\psi(\theta)\rangle\langle\psi(\theta)|) \partial_j \psi(\theta) \rangle], \quad (\text{E.21})$$

where the penultimate equality follows because

$$0 = \partial_i (\langle \psi(\theta) | \psi(\theta) \rangle) = \langle \partial_i \psi(\theta) | \psi(\theta) \rangle + \langle \psi(\theta) | \partial_i \psi(\theta) \rangle. \quad (\text{E.22})$$