Two approaches to obtain the strong converse exponent of quantum hypothesis testing for general sequences of quantum states

Milán Mosonyi, Tomohiro Ogawa

Abstract—We present two general approaches to obtain the strong converse exponent of simple quantum hypothesis testing for correlated quantum states. One approach requires that the states satisfy a certain factorization property; typical examples of such states are the temperature states of translation-invariant finite-range interactions on a spin chain. The other approach requires the differentiability of a regularized Rényi α-divergence in the parameter α; typical examples of such states include temperature states of non-interacting fermionic lattice systems, and classical irreducible Markov chains. In all cases, we get that the strong converse exponent is equal to the Hoeffding anti-divergence, which in turn is obtained from the regularized Rényi divergences of the two states.

I. INTRODUCTION

Assume that we have a quantum system with finite-dimensional Hilbert space \( \mathcal{H} \), and we know that the system is either prepared in the state \( \rho_1 \) (null-hypothesis \( H_0 \)), or in the state \( \sigma_1 \) (alternative hypothesis \( H_1 \)). We further assume that we have access to several identical copies of the system, either all prepared in state \( \rho_1 \), or all prepared in state \( \sigma_1 \); for \( n \) copies this means that the state of the system is given by \( \rho_n := \rho_1^\otimes n \) or by \( \sigma_n := \sigma_1^\otimes n \). Our task is to decide which hypothesis is true, by performing measurements on the system. It is easy to see that the most general decision scheme can be described by a binary POVM (positive operator-valued measure), with POVM elements \( T(0) = T \) corresponding to accepting the null-hypothesis, and \( T(1) = I - T \) corresponding to accepting the alternative hypothesis. Here we assume that we are allowed to make collective measurements on all the available copies, i.e., \( T \) can be any positive semidefinite operator on \( \mathcal{H}_n = \mathcal{H}^\otimes n \), satisfying \( T \leq I \). Such an operator is called a test. Obviously, the test \( T \) and the POVM \( \{T, I - T\} \) uniquely determine each other.

There are two possible ways of making an erroneous decision; either by accepting \( H_1 \) when \( H_0 \) is true (type I error), or the other way around (type II error). For a test \( T \), the probabilities of these errors are given by

\[
\alpha_n(T) := \text{Tr} \rho_n (I - T), \quad \text{(type I)} \quad \text{and} \quad \beta_n(T) := \text{Tr} \sigma_n T, \quad \text{(type II)}.
\]

Obviously, there is a trade-off between these two error probabilities, and there are various ways to jointly optimize them. Probably the most studied scenario is where the type I error is required to vanish in the asymptotics or, in a different formulation, to stay below a given threshold for all number of copies. The quantum Stein’s lemma [28], [47] states that in both formulations, the best achievable asymptotics for the type II error is an exponential decay, where the exponent is given by the relative entropy \( D(\rho_1 || \sigma_1) \).

To get a more detailed view of the trade-off between the two error probabilities, one may ask about the asymptotics of the type I error when the type II error is made to decay as \( e^{-r} \), where \( r \) is a fixed rate below or above the optimal rate \( D(\rho_1 || \sigma_1) \). As it turns out, for rates below \( D(\rho_1 || \sigma_1) \), the best achievable asymptotics for the type I error is an exponential decay, with a rate

\[
d(r|\rho||\sigma) := \sup \left\{ -\limsup_{n \to +\infty} \frac{1}{n} \log \alpha_n(T_n) : \limsup_{n \to +\infty} \frac{1}{n} \log \beta_n(T_n) < -r \right\}
\]

(1)

\[
= H_r(\rho_1 || \sigma_1) := \sup_{0 < \alpha < 1} \alpha - 1 \alpha \left[ r - D_\alpha(\rho_1 || \sigma_1) \right].
\]

(2)

Here, \( d(r|\rho||\sigma) \) is the direct exponent of the problem, \( H_r(\rho_1 || \sigma_1) \) is the Hoeffding divergence, and \( D_\alpha(\rho_1 || \sigma_1) := \frac{1}{\alpha - 1} \log \text{Tr} \rho_1^\alpha \sigma_1^{1-\alpha} \) is a quantum version of Rényi’s \( \alpha \)-divergence [48], [50]. On the other hand, for rates above \( D(\rho_1 || \sigma_1) \), we see a strong converse behaviour; namely, the type I error not only does not vanish asymptotically, but it goes to 1 exponentially fast, and the best achievable exponent is

\[
s^c(r|\rho||\sigma) := \inf \left\{ -\liminf_{n \to +\infty} \frac{1}{n} \log (1 - \alpha_n(T_n)) : \limsup_{n \to +\infty} \frac{1}{n} \log \beta_n(T_n) < -r \right\}
\]

(3)

\[
= H^c_r(\rho_1 || \sigma_1) := \sup_{1 < \alpha} \alpha - 1 \alpha \left[ r - D_\alpha^*(\rho_1 || \sigma_1) \right].
\]

(4)
Here, $sc(r|\rho||\sigma)$ is the strong converse exponent of the problem, $H_\alpha^*(\rho_1||\sigma_1)$ is the Hoeffding anti-divergence, and $D_\alpha^*(\rho_1||\sigma_1) := 1/\alpha \log \text{Tr}(\rho_1^{1/\alpha} (1-\sigma_1) / \rho_1^{1/\alpha})$ is an alternative version of the quantum Rényi divergence, recently introduced in [59]. [77].

The expression for the direct exponent in the classical case (corresponding to commuting $\rho_1$ and $\sigma_1$) has been obtained in [7], [32]. The exponential decay of the type I error probabilities in the non-commuting case has been proved in [21], with bounds on $d(r|\rho||\sigma)$ similar to (3) in form but with a different quantum version of the Rényi divergences. The correct form of the direct exponent has been obtained in [23], [41], based on techniques developed for the quantum Chernoff bound in [3], [46]. The strong converse exponent in the classical case has been first determined in [13], expressed as an optimization of relative entropies. The strong converse property of the quantum Stein’s lemma has been first proved in [47], with a suboptimal bound on the strong converse exponent; the proof was later much simplified in [40]. An expression for the strong converse exponent in quantum hypothesis testing was given in [23], with an asymptotic post-measurement version of the $D_\alpha^*$ Rényi divergences in place of $D_\alpha^*(\rho_1||\sigma_1)$. The expression (4) has been obtained recently in [37].

It is worth noting that two different notions of quantum Rényi divergence are needed to completely describe the trade-off curve. Indeed, the direct exponent is expressed in terms of the $D_\alpha^*$ divergences with $\alpha \in (0, 1)$, whereas the strong converse exponent is a function of the $D_\alpha^*$ divergences with $\alpha > 1$. A more explicit operational interpretation of these divergences as generalized cutoff rates has been given in [36]. [37].

These results give a complete description of the trade-off between the exponents of the two error probabilities in the i.i.d. (independent and identically distributed) case described above. In reality, however, there may be correlations between the different copies. That is, while the state of the individual systems are still described by $\rho_1$ if $H_0$ is true, the global state of $n$ copies may not be of the product form $\rho_1^\otimes n$ as assumed above, and similarly, $\sigma_n$ might contain correlations among the different copies. A physically relevant example of such a scenario is where the consecutive copies are parts of a chain of particles governed by the translates of some local Hamiltonian, and $\rho_n$ and $\sigma_n$ are temperature (Gibbs) states of two different Hamiltonians.

Exact trade-off formulas for such states have been obtained in the direct domain in [30], where two general methods have been developed for the hypothesis testing of correlated states. The first method works if the states corresponding to both hypotheses satisfy a certain factorization property, which was shown to be satisfied by temperature states of translation-invariant finite-range Hamiltonians on a spin chain in [29]. The other method requires the existence of the regularized Rényi divergences $\overline{D}_\alpha^*(\rho||\sigma) := \lim_n (1/n) D_\alpha^*(\rho_n||\sigma_n)$ for all $\alpha \in (0, 1)$, and differentiability of the limit in the parameter $\alpha$. Typical examples of such states include classical irreducible Markov chains [13], certain finitely correlated states [15], [30], and temperature states of non-interacting fermions [54] and bosons [65] on a cubic lattice. In both cases, the trade-off formula is a direct generalization of (1)–(2), with $\overline{D}_\alpha^*(\rho||\sigma)$ in place of $D_\alpha^*(\rho_1||\sigma_1)$.

Here we show analogous results for the trade-off in the strong converse region. Namely, we show the following extension of the i.i.d. result (3)–(4):

\[
sc(r|\rho||\sigma) := H_\alpha^*(\rho||\sigma) := \sup_{1<\alpha} \frac{1}{\alpha} \left[ r - \overline{D}_\alpha^*(\rho||\sigma) \right]
\]

with $\overline{D}_\alpha^*(\rho||\sigma) := \lim_n (1/n) D_\alpha^*(\rho_n||\sigma_n)$, $\alpha > 1$, (6)

if one of two conditions is satisfied: (1) the states corresponding to both hypotheses satisfy the factorization property, or (2) the limit $\overline{D}_\alpha^*(\rho||\sigma)$ exists for all $\alpha > 1$, it is a differentiable function of $\alpha$, and it coincides with a variant of this limit, explained later. The main examples satisfying the first condition are again the temperature states of translation-invariant finite-range Hamiltonians on a spin chain. The second condition is satisfied by the same class of classical Markov states and finitely correlated states as in the previous paragraph, and for temperature states of non-interacting fermions on a cubic lattice.

The structure of the paper is as follows. In Section II we list some mathematical preliminaries. In Section III-A we summarize some properties of the quantum Rényi divergences, in Section III-B we introduce their asymptotic versions, which is needed when dealing with correlated states, and in Section III-C we introduce the Hoeffding anti-divergence based on the asymptotic Rényi $\alpha$-divergences.

Section IV is the main contribution of the paper. Here we start with a more general approach than described above, namely, we consider the strong converse exponent of the hypothesis testing problem between two general sequences of states $\{\rho_n\}_{n \in \mathbb{N}}$ and $\{\sigma_n\}_{n \in \mathbb{N}}$, where $\rho_n$ and $\sigma_n$ are states on the same Hilbert space, but states with different indices are not assumed to be related in any particular way. This general approach originates from the information spectrum method [20]. Expressions for the strong converse exponent in the classical [19] and in the quantum case [42] were obtained in the information spectrum framework in terms of the exponents of the Neyman-Pearson tests. Here we impose extra conditions on the two sequences in order to obtain the more explicit expression (5) for the strong converse exponent, which can be further evaluated for various classes of correlated states with physical relevance. We also note that while in the information spectrum method all exponents are evaluated in terms of the asymptotics of the error probabilities along the Neyman-Pearson tests, here we also consider variants of the Neyman-Pearson tests, e.g., in Sections IV-A and IV-B.

In the beginning of Section IV we start with two general observations. First, a straightforward generalization of the results of [40], [47], utilizing the monotonicity of the Rényi divergences under measurements, yields $sc(r|\rho||\sigma) \geq H_\alpha^*(\rho||\sigma)$ (with a slightly more general definition of $\overline{D}_\alpha^*(\rho||\sigma)$); this is the content of Lemma IV.1. Next, we show in Theorems IV.2–IV.3 that if a parametric family of sequences of tests exists with certain properties then the inequality can be reversed, and $sc(r|\rho||\sigma) = H_\alpha^*(\rho||\sigma)$ holds. We then consider two general cases in which the existence of such tests can be verified.
In Section [IV-A] we show that if a certain classicalization of the problem, corresponding to a suitably chosen auxiliary sequence \(\{\sigma_n\}_{n \in \mathbb{N}}\), yields the same asymptotic Rényi divergences as the original problem, which also satisfy some regularity condition (differentiability in \(\alpha\)) then the Neyman-Pearson tests of the classicalized problem can be used to fulfill the conditions of Theorem [IV.3] and obtain [5]. We show in Section [IV-B] that the hypothesis testing problem for gauge-invariant fermionic quasi-free states satisfies these conditions; moreover, the asymptotic Rényi divergences can be explicitly expressed in terms of the symbols of the two states. The standard classical example for which these conditions hold is the hypothesis testing of irreducible Markov chains; an expression for the strong converse exponent of this problem has been determined in [42]. We explain in Appendix D how the results of [44] can be obtained from our general considerations.

In Section [IV-B] we consider a special class of states on an infinite spin chain, satisfying a certain factorization property. We evaluate the exponents of the type I success- and the type II error probabilities corresponding to the standard Neyman-Pearson tests, and show that these tests satisfy the conditions of Theorem [IV.3] from which we can conclude that [5] holds. Once these exponents are available, [5] can also be obtained from the general information spectrum formula for the strong converse exponent, given in [42] Theorem 4, as we explain in Remark IV.14. The factorization property is known to hold for the Gibbs states of finite-range translation-invariant infinite spin chain, satisfying a certain factorization property. We call a map \(H \to B(K)\) positive if and only if \(\Phi (H) = \Phi (H) \geq 0\) for all \(H \geq 0\) on \(H\). The factorization property is known to hold for the Gibbs states of finite-range translation-invariant Hamiltonians [29], as we discuss in Section [IV-A].

We remark that neither the conditions of Section [IV-A], nor the factorization property of Section [IV-B] need to hold in the classical case, i.e., for commuting \(\rho_n\) and \(\sigma_n\). Hence, our results may have non-trivial applications even for classical hypothesis testing.

Background material on classical large deviations, fermionic quasi-free states and Szegö’s theorem is given in the Appendices.

II. Preliminaries

For a finite-dimensional Hilbert space \(\mathcal{H}\), let \(B(\mathcal{H})\) denote the set of linear operators on \(\mathcal{H}\), let \(B(\mathcal{H})_+\) denote the set of non-zero positive semidefinite operators, \(B(\mathcal{H})_{++}\) the set of positive definite operators on \(\mathcal{H}\), and let \(S(\mathcal{H}) := \{\rho \in B(\mathcal{H})_+ : \text{Tr} \rho = 1\}\) be the set of density operators or states.

We call a map \(\Phi : B(\mathcal{H}) \to B(K)\) a positive map if \(\Phi\) is linear, and \(\Phi (B(\mathcal{H})_+) \subseteq B(K)_+\). For every finite-dimensional Hilbert space \(\mathcal{H}\), \((X,Y) \mapsto \text{Tr} X^* Y\) is an inner product on \(B(\mathcal{H})\) (called the Hilbert-Schmidt inner product), and for a linear map \(\Phi : B(\mathcal{H}) \to B(K)\), we denote its adjoint with respect to the Hilbert-Schmidt inner products on \(B(\mathcal{H})\) and \(B(K)\) by \(\Phi^*\). It is easy to see that \(\Phi\) is positive if and only if \(\Phi^*\) is positive, and \(\Phi\) is trace-preserving if and only if \(\Phi^*\) is unital.

For a self-adjoint operator \(X\) on a finite-dimensional Hilbert space, let \(\{X \geq 0\}\) denote the spectral projection of \(X\) corresponding to the non-negative eigenvalues of \(X\). The spectral projections \(\{X > 0\}\), \(\{X \leq 0\}\) and \(\{X < 0\}\) are defined similarly. The positive part \(X_+\) of \(X\) is defined as \(X_+ := X \{X > 0\}\). It is easy to see that

\[
\text{Tr} X_+ = \max \{\text{Tr} XT : 0 \leq T \leq I\},
\]

a fact that we will use without further notice.

Lemma II.1. For any Hermitian operators \(A,B \in B(\mathcal{H})\),

\[
A \geq B \implies \text{Tr} A_+ \geq \text{Tr} B_+.
\]

For any Hermitian operator \(A\) and any positive trace preserving map \(\mathcal{F}\), we have

\[
\text{Tr} A_+ \geq \text{Tr} \mathcal{F}(A)_+.
\]

Proof: The first assertion follows from

\[
\text{Tr} B_+ = \text{Tr} B \{B > 0\} \leq \text{Tr} A \{B > 0\} \leq \text{Tr} A \{A > 0\},
\]

where the first inequality is due to the assumption \(B \leq A\), and the second is due to [4]. The second assertion follows by

\[
\text{Tr} \mathcal{F}(A)_+ = \text{Tr} \mathcal{F}(A) \mathcal{F}(A_+) \geq \text{Tr} A \mathcal{F}^* \mathcal{F}(A_+) \leq \text{Tr} A \{A > 0\},
\]

where \(\mathcal{F}^*\) is the Hilbert-Schmidt adjoint of \(\mathcal{F}\), and we used that \(\mathcal{F}^*\) is positivity preserving and unital.

We will follow the convention that powers of a positive semidefinite operator \(A\) are taken on its support only, and defined to be 0 on the orthocomplement of its support. That is, if \(\lambda_1, \ldots, \lambda_r\) are the strictly positive eigenvalues of \(A\) with corresponding spectral projections \(P_1, \ldots, P_r\), then

\[
A^t := \sum_{i=1}^{r} \lambda_i^t P_i.
\]

In particular, \(A^0\) denotes the projection onto the support of \(A\), and \(A^0 \leq B^0\) is a shorthand for \(\text{supp} A \subseteq \text{supp} B\) when \(A,B \in B(\mathcal{H})_+\). Similarly, we define \(\log A\) to be 0 on the orthocomplement of \(A\).

For an operator \(\sigma \in B(\mathcal{H})\), we denote by \(v(\sigma)\) the number of different eigenvalues of \(\sigma\). If \(\sigma\) is self-adjoint with spectral projections \(P_1, \ldots, P_r\), then the pinching by \(\sigma\) is the map \(\mathcal{E}_\sigma : B(\mathcal{H}) \to B(\mathcal{H})\), defined as

\[
\mathcal{E}_\sigma : X \mapsto \sum_{i=1}^{r} P_i XP_i, \quad X \in B(\mathcal{H}).
\]

The pinching inequality [22], [23] tells us that if \(X\) is positive semidefinite then

\[
X \leq v(\sigma) \mathcal{E}_\sigma (X).
\]

III. Rényi Divergences and Related Quantities

A. Definitions and general properties

For non-zero positive semidefinite operators \(\rho, \sigma\) on a finite-dimensional Hilbert space \(\mathcal{H}\), let

\[
Q_\rho(\rho|\sigma) := \text{Tr} \rho^t \sigma^{1-t}, \quad \psi(t|\rho|\sigma) := \log Q_\rho(\rho|\sigma)
\]

for any \(t \in \mathbb{R}\), and

\[
Q_t(\rho|\sigma) := \text{Tr} \left( \rho^{\frac{t}{2}} \sigma^{\frac{1-t}{2}} \right)^t, \quad \psi(t|\rho|\sigma) := \log Q_t(\rho|\sigma),
\]

for any \(t > 0\). In the following, let \((v)\) denote either * or \(\{\}\), where \(\{\}\) stands for the empty string. That is, \(Q_t^{(v)}(\rho|\sigma)\) with \((v) = \{\}\) is simply \(Q_t(\rho|\sigma)\). When \(\rho\) and \(\sigma\) commute,
the expressions with and without $*$ coincide, and therefore we omit $*$ in the notation.

The Rényi $\alpha$-divergences of $\rho$ w.r.t. $\sigma$ for parameter $\alpha \in [0, +\infty) \setminus \{1\}$ are defined as

$$D_\alpha^{(v)}(\rho \| \sigma) := \lim_{\alpha \to 0} \frac{1}{\alpha - 1} \psi^{(v)}(\alpha |\rho|\sigma + \varepsilon I) - \frac{1}{\alpha - 1} \log \text{Tr} \rho$$

$$= \begin{cases} \frac{1}{\text{Tr} \rho} \left[ \text{Tr} \rho \log \rho - \text{Tr} \rho \log \sigma \right], & \rho^0 \leq \sigma^0, \\ +\infty, & \text{otherwise.} \end{cases}$$

The equality above is straightforward to verify for $(v) = \{ \}$, and it follows from Lemma 12 in [39] for $(v) = *$. For $\alpha = 1$ we define

$$D_1(\rho \| \sigma) := \lim_{\alpha \to 1} D_\alpha^{(v)}(\rho \| \sigma)$$

$$= D(\rho \| \sigma)$$

$$:= \begin{cases} \frac{1}{\text{Tr} \rho} \left[ \text{Tr} \rho \log \rho - \text{Tr} \rho \log \sigma \right], & \rho^0 \leq \sigma^0, \\ +\infty, & \text{otherwise.} \end{cases}$$

(11)

Note that $D(\rho \| \sigma)$ is the relative entropy [48], [53], [54] of $\rho$ w.r.t. $\sigma$. The above limit relation for $D_\alpha$ is straightforward to verify, and for $D_\alpha^*$ it has been shown by different methods in [38], [39], [57]. We also provide a proof for it below. It has been shown in [39] that

$$D_\alpha^*(\rho \| \sigma) := \lim_{\alpha \to +\infty} D_\alpha^{(v)}(\rho \| \sigma)$$

$$= D_{\text{max}}(\rho \| \sigma) := \inf \{ \gamma : \rho \leq e^\gamma \sigma \},$$

where $D_{\text{max}}(\rho \| \sigma)$ is the max-relative entropy of $\rho$ w.r.t. $\sigma$ [12], [51].

Lemma III.1. Let $\rho$ and $\sigma$ be such that $\rho \sigma \neq 0$. Then $\psi^{(v)}(\rho \| \sigma)$ is differentiable on $(0, +\infty)$, and

$$\frac{d}{dt} \psi(t|\rho\|\sigma) = \frac{1}{Q_t(\rho \| \sigma)} \text{Tr} \rho^0 \sigma^0 t^{-1} (\log \rho - \log \sigma),$$

(14)

$$\frac{d}{dt} \psi^*(t|\rho\|\sigma) = \frac{1}{Q_t(\rho \| \sigma)} \left[ \text{Tr} \left( \rho^0 \sigma^0 t^{-1} \rho^0 \right) \right]$$

$$= \frac{1}{1 - \frac{1}{t}} \text{Tr} \left( \frac{\rho^0 \sigma^0}{\rho^0} \right) (\log \rho) \rho^0.$$  

(15)

Moreover, (14) is valid for all $t \in \mathbb{R}$.

Proof: The derivative of $\psi$ is straightforward to compute. To see the derivative of $\psi^*$, first note that for every $t > 0$, there exist $c_t, a_t > 0$ such that $c_t \rho^0 \sigma^0 t \leq \rho^0 \sigma^0 t \leq a_t \rho^0 \sigma^0 t$, and thus $(\rho^0 \sigma^0 t \leq \rho^0 \sigma^0 t) = P$, independently of $t$. Hence, $\rho^0 \sigma^0 t \leq \rho^0 \sigma^0 t$ can be seen as an invertible positive operator on $\mathbb{R}$. Define

$$g : \mathbb{R}_+ \to \mathbb{R}_+ \oplus B(\mathbb{R}^n), \quad g(t) := t \oplus \rho^0 \sigma^0 t \oplus \rho^0 \sigma^0 t,$$

$$f : \mathbb{R}_+ \oplus B(\mathbb{R}^n) \to \mathbb{R}, \quad f(t \oplus X) := \text{Tr} X^t,$$

where $\mathbb{R}_+ = (0, +\infty)$, so that $\text{Tr} \left( \rho^0 \sigma^0 t^{-1} \rho^0 \right) = f(g(t)).$

Then $g$ has derivative

$$\frac{d}{dt} g(t) = 1 \oplus \left( -\frac{1}{t^2} \right) \rho^0 \sigma^0 t^{-1} (\log \rho) \rho^0,$$

and the derivative of $f$ at a point $(t, X)$ is the linear map

$$df(t, X) := (s, Y) \mapsto s \text{Tr} X^t \log X + t \text{Tr} X^{t-1} Y.$$  

(16)

The second term in (16) can be obtained e.g. from Theorem V.3.3 in [3]. Using the chain rule for derivatives, we get that

$$\frac{d}{dt} \text{Tr} \left( \rho^0 \sigma^0 t^{-1} \rho^0 \right) = \frac{d}{dt} f(g(t))$$

$$= \text{Tr} \left( \rho^0 \sigma^0 t^{-1} \rho^0 \right) \frac{d}{dt} \log \left( \rho^0 \sigma^0 t^{-1} \rho^0 \right)$$

$$- \frac{1}{t} \text{Tr} \left( \rho^0 \sigma^0 t^{-1} \rho^0 \right) \left( -1 \rho^0 \sigma^0 t^{-1} (\log \rho) \rho^0 \right),$$

which yields (15).

Corollary III.2. The limit relation in (11) holds.

Proof: Assume first that $\rho^0 \leq \sigma^0$. Then we have $D_\alpha^{(v)}(\rho \| \sigma) = \psi^{(v)}(\alpha |\rho|\sigma) - \psi^{(v)}(1|\rho|\sigma)$, and hence

$$\lim_{\alpha \to 1} D_\alpha^{(v)}(\rho \| \sigma) = \frac{d}{d\alpha} \psi^{(v)}(\alpha |\rho|\sigma) |_{\alpha = 1} = D(\rho \| \sigma),$$

where the last equality is due to lemma III.1. Assume next that $\rho^0 \neq \sigma^0$. Then $D_\alpha^{(v)}(\rho \| \sigma) = +\infty = D(\rho \| \sigma)$ for every $\alpha > 1$. On the other hand, $\psi^{(v)}(1|\rho|\sigma) = \log \text{Tr} \rho \sigma^0 < \log \text{Tr} \rho$, and hence for $\alpha < 1$ we have

$$D_\alpha^{(v)}(\rho \| \sigma) = \frac{\psi^{(v)}(\alpha |\rho|\sigma) - \psi^{(v)}(1|\rho|\sigma)}{\alpha - 1} + \frac{\log \text{Tr} \rho \sigma^0}{\alpha - 1} \log \text{Tr} \rho,$$

The first term has a finite limit, again due to lemma III.1, while the second term goes to $+\infty = D(\rho \| \sigma)$ as $\alpha \to 1$. It is easy to see (by simply computing its second derivative) that for fixed $\rho, \sigma \in B(H)_+$, the function $\alpha \mapsto \psi(\alpha |\rho|\sigma)$ is convex on $\mathbb{R}$. We have the following:

Lemma III.3. Let $\rho, \sigma \in B(H)_+$ be such that $\rho^0 \leq \sigma^0$. For every $\alpha > 1$,

$$\psi^*(\alpha |\rho|\sigma) = \lim_{n \to +\infty} \frac{1}{n} \psi(\alpha |E_{\sigma \otimes n} \rho^0| \sigma^{\otimes n}),$$

(17)

where $E_{\sigma \otimes n}$ is the pinching by $\sigma^{\otimes n}$. In particular, $\alpha \mapsto \psi^*(\alpha |\rho|\sigma)$ is convex.

Proof: The limit relation (17) is due to Theorem III.7 in [57]. By (17), $\psi^*(\rho \| \sigma)$ is the pointwise limit of convex functions, and hence itself is convex on $(1, +\infty)$.

Corollary III.4. Let $\rho, \sigma \in B(H)_+$ be such that $\rho^0 \leq \sigma^0$. Then $\alpha \mapsto D_\alpha^{(v)}(\rho \| \sigma)$ is monotone increasing on $(1, +\infty)$, and

$$D_1(\rho \| \sigma) = \inf_{\alpha > 1} D_\alpha^{(v)}(\rho \| \sigma).$$

Proof: Note that $\rho^0 \leq \sigma^0$ implies that $D_\alpha^{(v)}(\rho \| \sigma) = \psi^{(v)}(\alpha |\rho|\sigma) - \psi^{(v)}(1|\rho|\sigma)$, and hence convexity of $\psi^{(v)}(\alpha |\rho|\sigma)$ in $\alpha$ yields $\alpha \mapsto D_\alpha^{(v)}(\rho \| \sigma)$ is a monotone increasing function of $\alpha$; in particular, (13) holds.

The $D_\alpha$ Rényi divergences are known to be monotone non-increasing under completely positive trace-preserving maps for $\alpha \in [0, 2]$. Monotonicity for the $D_\alpha^*$ Rényi divergences...
has been proved for different ranges of $\alpha$ and with different methods in \cite{6, 16, 31, 37, 39, 57}.

**Lemma III.5.** Let $\rho, \sigma \in B(\mathcal{H})_+$ and $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$ be a linear completely positive trace-preserving map. Then

$$D_\alpha^*(\Phi(\rho)\|\Phi(\sigma)) \leq D_\alpha^*(\rho\|\sigma), \quad \alpha \in [1/2, +\infty].$$

The following lemma is straightforward to verify:

**Lemma III.6.** Let $\rho, \sigma \in B(\mathcal{H})_+$ and $\lambda, \kappa > 0$. For every $\alpha \in (0, +\infty]$,

$$\psi^{(v)}(\alpha|\lambda \rho\|\kappa \sigma) = \alpha \log \lambda + (1 - \alpha) \log \kappa + \psi^{(v)}(\alpha|\rho\|\sigma),$$

$$D_\alpha^{(v)}(\lambda \rho\|\kappa \sigma) = \log \lambda - \log \kappa + D_\alpha^{(v)}(\rho\|\sigma).$$

**B. Asymptotic Rényi quantities**

For every $n \in \mathbb{N}$, let $\mathcal{H}_n$ be a finite-dimensional Hilbert space, let $\rho_n \in B(\mathcal{H}_n)$ be a state, and $\sigma_n \in B(\mathcal{H}_n)_+$ be a positive semidefinite operator. These will play the role of the null- and the alternative hypotheses in the later sections. Note that we don’t require the $\sigma_n$ to be normalized; the reason is that this more general case can be treated the same way as the normalized case, and it turns out to be useful e.g., in state compression (see, e.g., \cite{38}). We will use the notation

$$\rho := \{\rho_n\}_{n \in \mathbb{N}} \quad \text{and} \quad \sigma := \{\sigma_n\}_{n \in \mathbb{N}}.$$  

We will assume throughout that

$$\text{supp} \rho \subseteq \text{supp} \sigma,$$

which we abbreviate as

$$\text{supp} \rho \subseteq \text{supp} \sigma.$$  

We will also consider an additional sequence $\tilde{\sigma} = \{\tilde{\sigma}_n\}_{n \in \mathbb{N}}$ such that

$$\sigma_n \leq \tilde{\sigma}_n, \quad \text{and} \quad (\sigma_n)^0 = (\tilde{\sigma}_n)^0, \quad n \in \mathbb{N}. \quad (21)$$

This sequence will be specified later, depending on the concrete problem. Given the sequence $\tilde{\sigma}$, we introduce

$$\hat{\rho}_n := \mathcal{E}_{\tilde{\sigma}_n}(\rho_n),$$

the pinching of $\rho_n$ by $\tilde{\sigma}$. By the pinching inequality \eqref{eq:pinching}, we have

$$\rho_n \leq \psi(\tilde{\sigma}_n)\hat{\rho}_n, \quad (22)$$

where $\psi(\tilde{\sigma}_n)$ stands for the number of different eigenvalues of $\tilde{\sigma}_n$.

**Remark III.7.** The application of the pinching technique in Quantum Information Theory goes back to \cite{28} and \cite{24}. The pinching of $\rho_n$ with $\tilde{\sigma}_n := \sigma_n$ was the main tool to obtain the first expression for the strong converse exponent of i.i.d. binary quantum state discrimination in \cite{24}, as well as for the expression in terms of the sandwiched Rényi divergences in \cite{24}. The key property used in these applications is the pinching inequality \cite{24}, and that $\lim_{n \to +\infty} \frac{1}{n} \log \psi(\sigma_n) = 0$ in the i.i.d. case, where $\sigma_n = \sigma_n^{\otimes n}$. This latter property, however, need not hold in the non-i.i.d. case, and therefore pinching with $\sigma_n$ may not be a viable way to extend results from the i.i.d. to the non-i.i.d. setting. To circumvent this problem, a clever way of grouping together the eigenvalues of $\sigma_n$ was introduced in \cite{52}, which we review below in Example III.10. This results in a new reference operator $\tilde{\sigma}_n$ satisfying \eqref{eq:pinching}, and with the additional property that $\lim_{n \to +\infty} \frac{1}{n} \log \psi(\tilde{\sigma}_n) = 0$ under much weaker conditions than i.i.d. We will use this trick to obtain the strong converse exponent for gauge-invariant quasi-free states in Section V.-B. We are grateful to an anonymous referee for drawing our attention to this technique.

For $\alpha > 1$, we define the asymptotic Rényi quantities

$$\overline{\psi}(\alpha|\rho\|\sigma) := \limsup_{n \to +\infty} \frac{1}{n} \psi^*(\alpha|\rho_n\|\sigma_n),$$

$$\overline{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \overline{\psi}(\alpha|\rho\|\sigma) = \limsup_{n \to +\infty} \frac{1}{n} D_\alpha^*(\rho_n\|\sigma_n),$$

$$\overline{\psi}(\alpha|\rho\|\sigma) := \limsup_{n \to +\infty} \frac{1}{n} \psi(\tilde{\rho}_n\|\tilde{\sigma}_n),$$

$$\overline{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \overline{\hat{\psi}}(\alpha|\rho\|\sigma) = \limsup_{n \to +\infty} \frac{1}{n} \hat{D}_\alpha(\tilde{\rho}_n\|\tilde{\sigma}_n).$$

**Lemma III.8.** \(\overline{\psi}(\alpha|\rho\|\sigma)\) and \(\overline{\hat{\psi}}(\alpha|\rho\|\sigma)\) are convex in $\alpha$ on $(1, +\infty)$, the functions

$$\alpha \mapsto \overline{D}_\alpha(\rho\|\sigma) \quad \text{and} \quad \alpha \mapsto \overline{\hat{D}}_\alpha(\rho\|\sigma)$$

are monotone increasing, and hence

$$\overline{D}_1(\rho\|\sigma) := \inf_{\alpha > 1} \overline{D}_\alpha(\rho\|\sigma) = \lim_{\alpha \to +\infty} \overline{D}_\alpha(\rho\|\sigma), \quad (26)$$

$$\overline{D}_\infty(\rho\|\sigma) := \sup_{\alpha > 1} \overline{D}_\alpha(\rho\|\sigma) = \lim_{\alpha \to +\infty} \overline{D}_\alpha(\rho\|\sigma), \quad (27)$$

$$\overline{\hat{D}}_1(\rho\|\sigma) := \inf_{\alpha > 1} \overline{\hat{D}}_\alpha(\rho\|\sigma) = \lim_{\alpha \to +\infty} \overline{\hat{D}}_\alpha(\rho\|\sigma), \quad (28)$$

$$\overline{\hat{D}}_\infty(\rho\|\sigma) := \sup_{\alpha > 1} \overline{\hat{D}}_\alpha(\rho\|\sigma) = \lim_{\alpha \to +\infty} \overline{\hat{D}}_\alpha(\rho\|\sigma). \quad (29)$$

**Proof:** Both $\overline{\psi}$ and $\overline{\hat{\psi}}$ are the limsup of convex functions, and hence are convex. Note that $\text{supp} \rho \subseteq \text{supp} \sigma$ implies $\overline{\psi}(1|\rho\|\sigma) = 0$, and hence $\overline{D}_\alpha(\rho\|\sigma) = \frac{1}{\alpha - 1} \overline{\hat{\psi}}(\alpha|\rho\|\sigma) = \frac{1}{\alpha - 1} \overline{\psi}(\alpha|\rho\|\sigma) = \frac{1}{\alpha - 1} \overline{\hat{\psi}}(\alpha|\rho\|\sigma)$, i.e., $\overline{D}_\alpha(\rho\|\sigma)$ and the proof for $\overline{\hat{D}}_\alpha(\rho\|\sigma)$ goes exactly the same way.

We say that $\overline{\psi}(\alpha|\rho\|\sigma)$ (resp., $\overline{\hat{\psi}}(\alpha|\rho\|\sigma)$) exists as a limit, if the corresponding limsup in \eqref{eq:limsup}–\eqref{eq:limsup2} can be replaced with a limit. We have the following:

**Lemma III.9.** For every $\alpha > 1$ and $n \in \mathbb{N}$,

$$\frac{1}{n} \psi^*(\alpha|\rho_n\|\sigma_n) - \frac{\alpha}{n} \log \psi(\tilde{\sigma}_n) + \frac{1 - \alpha}{n} D_{\max}(\tilde{\sigma}_n\|\sigma_n)$$

$$\leq \frac{1}{n} \psi(\tilde{\rho}_n\|\tilde{\sigma}_n) \leq \frac{1}{n} \psi^*(\alpha|\rho_n\|\sigma_n). \quad (30)$$

In particular, if $\lim_{n \to +\infty} \frac{1}{n} \log \psi(\tilde{\sigma}_n) = 0 = \lim_{n \to +\infty} \frac{1}{n} D_{\max}(\tilde{\sigma}_n\|\sigma_n)$ then

$$\overline{\psi}(\alpha|\rho\|\sigma) = \overline{\hat{\psi}}(\alpha|\rho\|\sigma),$$

and $\overline{\psi}(\alpha|\rho\|\sigma)$ exists as a limit if and only if $\overline{\hat{\psi}}(\alpha|\rho\|\sigma)$ exists as a limit.
Proof: By the monotonicity of $D^*_\alpha$ under pinching [39 Proposition 14], we have $\psi(\alpha|\rho_n||\sigma_n) \geq \psi^*(\alpha|\rho_n||\sigma_n)$. For $\alpha > 1$, the function $x \mapsto x^{-\alpha}$ is operator monotone decreasing on $(0, +\infty)$, and $X \mapsto \text{Tr} X^\alpha$ is monotone increasing on positive semidefinite operators (with respect to the positive semidefinite ordering), and hence $\sigma_n \leq \tilde{\sigma}_n$ yields $\psi^*(\alpha|\rho_n||\sigma_n) \leq \psi^*(\alpha|\rho_n||\sigma_n)$. This proves the second inequality in [30].

According to the proof of [37 Theorem 3.7]),

$$\psi(\alpha|\rho_n||\sigma_n) \geq \psi^*(\alpha|\rho_n||\sigma_n) - \alpha \log v(\tilde{\sigma}_n).$$

By [12], $\tilde{\sigma}_n \leq c_n\sigma_n$, where $c_n := e^{D^*_\alpha(\tilde{\sigma}_n||\sigma_n)}$. By the same monotonicity argument as above,

$$\psi^*(\alpha|\rho_n||\sigma_n) \geq \psi^*(\alpha|\rho_n||c_n\sigma_n) = \psi^*(\alpha|\rho_n||\sigma_n) + (1 - \alpha)D^*_\alpha(\tilde{\sigma}_n||\sigma_n),$$

where the last identity is due to [19]. This proves the first inequality in (30). The rest of the Lemma is obvious from (30). ■

The following construction is from the proof of [52 Theorem 14], which we review here in detail for readers convenience:

Example III.10. Let $\lambda_{1,n}, \ldots, \lambda_{r,n}$ be the different non-zero eigenvalues of $\sigma_n$ with corresponding spectral projections $P_{1,n}, \ldots, P_{r,n}$, and let $\lambda_{\max}(\sigma_n) := \lambda_{1,n}$, $\lambda_{\min}(\sigma_n) := \lambda_{r,n}$. Let $q_n := \lambda_{\max}(\sigma_n)/\lambda_{\min}(\sigma_n)$, and $l_n := \log q_n + 1$. Then for every $i$, there exists a unique $k_i \in \{-1, \ldots, l_n - 1\}$ such that $\lambda_{\min}(\sigma_n) q_{n}^{-\frac{k_i}{l_n}} < \lambda_{i,n} \leq \lambda_{\min}(\sigma_n) q_{n}^{-\frac{k+1}{l_n}}$. Define $\hat{\lambda}_{i,n} := \lambda_{\min}(\sigma_n) q_{n}^{-\frac{k_i}{l_n}}$, $i = 1, \ldots, r_n$, and $\tilde{\sigma}_n := \sum_{i=1}^{r_n} \hat{\lambda}_{i,n} P_{i,n}$.

Then $v(\tilde{\sigma}_n) \leq \left| \log \frac{\lambda_{\max}(\sigma_n)}{\lambda_{\min}(\sigma_n)} \right| + 1 - \left| -\log \lambda_{\min}(\sigma_n) \right| + 1$ (31)

and $\sigma_n \leq \tilde{\sigma}_n \leq q_{n}^{-\frac{1}{l_n}} \sigma_n$, (32)

and by the last inequality,

$$D^*_\alpha(\tilde{\sigma}_n||\sigma_n) \leq \frac{1}{l_n} \log q_n \leq 1.$$ (33)

Following [52], we introduce the notation

$$\theta(\sigma_n) := \min \left\{ v(\sigma_n), \left| \log \frac{\lambda_{\max}(\sigma_n)}{\lambda_{\min}(\sigma_n)} \right| + 1 \right\}.$$

Corollary III.11. Consider one of the following scenarios:

(i) $\lim_{n \to +\infty} \frac{1}{l_n} \log v(\sigma_n) = 0$, and we define $\tilde{\sigma}_n := \sigma_n$, $n \in \mathbb{N}$.

(ii) $\lim_{n \to +\infty} \frac{1}{l_n} \log \left( \left| \log \frac{\lambda_{\max}(\sigma_n)}{\lambda_{\min}(\sigma_n)} \right| + 1 \right) = 0$, and we define (30).

(iii) $\lim_{n \to +\infty} \frac{1}{l_n} \log \theta(\sigma_n) = 0$, and for every $n \in \mathbb{N}$, if $\theta(\sigma_n) = v(\sigma_n)$ then let $\tilde{\sigma}_n := \sigma_n$, otherwise let $\tilde{\sigma}_n$ be the state constructed in Example III.10.

Then

$$\tilde{\psi}(\alpha|\rho||\sigma) = \psi(\alpha|\rho||\sigma).$$ (34)

Proof: Immediate from [30-33].

Corollary III.12. Assume that there exist constants $c, d > 0$ and $\nu \in \mathbb{R}$ such that $c^\nu(\sigma_n)^0 \leq \sigma_n \leq d^\nu(\sigma_n)^0$ for all large enough $n$. Then the sequence $\{\sigma_n\}_{n \in \mathbb{N}}$ constructed in Example III.10 satisfies (34).

Proof: Immediate from [31] of Corollary III.11 ■

Remark III.13. A similar condition as in Corollary III.12 was applied to the exponent of secret key generation in [23 Section VI C].

C. Generalized Legendre transforms

In this section we consider the extension of $H^*_\alpha$ in (4) to general correlated states. First we present a more general definition, corresponding to a general convex function $f$, and we will obtain the desired quantity by specializing to $f = -\psi$. For what follows, let $f : [1, +\infty) \to \mathbb{R}$ be a non-negative convex function such that $f(1) = 0$, and let

$$f_0(a) := \sup_{t>1} \{ a(t - 1) - f(t) \}, \quad a \in \mathbb{R},$$ (35)

$$H_{f,r} := \sup_{t>1} \frac{r(t - 1) - f(t)}{t} = \sup_{0<s<1} \left\{ s^r(1-s)f\left(\frac{1}{1-s}\right) \right\}, \quad r \in \mathbb{R}.$$ (36)

Note that $a \mapsto f_0(a) + a$ is the Legendre-Fenchel transform (or polar transform) of $f$ on $(1, +\infty)$, and $r \mapsto H_{f,r}$ is the Legendre-Fenchel transform of $s \mapsto (1-s)f\left(\frac{1}{1-s}\right)$ on $(0,1)$.

Let

$$D_{f,1} := a_{f,\min} := \inf_{1<t<+\infty} \frac{f(t)}{t - 1} = \lim_{t \to 1^+} \frac{f(t)}{t - 1} := \partial^+ f(1),$$ (37)

and

$$D_{f,\infty} := a_{f,\max} := \sup_{1<t<+\infty} \frac{f(t)}{t - 1} = \lim_{t \to +\infty} \frac{f(t)}{t - 1} := \partial^- f(1).$$ (38)

Note that $a_{f,\min}$ is always finite, whereas $a_{f,\max}$ can be $+\infty$, in which case also $\partial^- f = +\infty$.

Lemma III.14. For any $a \in \mathbb{R}$,

$$f_0(a) \geq 0, \quad \text{and} \quad f_0(a) > 0 \iff a > \partial^+ f(1).$$ (41)

For any $r \in [0, +\infty)$, we have

$$0 \leq H_{f,r} = \begin{cases} r - a_r = f_0(a_r), & \quad r < f_0(a_{f,\max}) + a_{f,\max}, \\
 r - a_{f,\max}, & \quad r \geq f_0(a_{f,\max}) + a_{f,\max} \end{cases}$$ (42)

where $a_r$ is the unique solution of $r - a_r = f_0(a_r)$. Moreover,

$$0 < H_{f,r} \iff r > \partial^+ f(1) = a_{f,\min} = \overline{r}_{f,\min}.$$ (43)

Proof: Non-negativity of $f_0(a)$ and $H_{f,r}$ are obvious from their definitions (35) and (36) and the fact that $f(1) = 0$. Due to the convexity of $f$, $t \mapsto \frac{f(t)}{t - 1} = \frac{f(t) - f(1)}{t - 1}$ is monotone increasing, proving the equalities of the limits.
and the infimum/supremum in \((37)\) and \((39)\). There exists a \(t > 1\) such that \(a(t-1) - f(t) > 0\) if and only if \(a > \inf_{t>1} \frac{f(t)}{t-1} = D^+ f(1)\), proving \((41)\). In particular, \(f_\circ(a_{f,\min}) = 0\), and hence \(a_{f,\min} = r_{f,\min}\), proving the last identity in \((43)\). Non-negativity of \(H^*_r\) and the rest of \((43)\) follow the same way as \((41)\). Hence, we have to prove the identities in \((42)\).

First, we consider the case \(0 \leq r < r_{f,\max}\). Note that \(a \mapsto f_\circ(a) + a\) is strictly increasing and continuous on \([0, a_{f,\max}]\), and hence for every \(0 < r < r_{f,\max}\) there exists a unique \(a_r\) such that \(r = f_\circ(a_r) + a_r\). By definition,

\[
f_\circ(a_r) \geq a_r(t-1) - f(t) = (t-1)(r-f_\circ(a_r)) - f(t), \quad t \geq 1,
\]

and equality holds in the above inequality for some \(t_r \in [1, \infty)\). Rearranging, we get

\[
f_\circ(a_r) \geq r(t-1) - f(t), \quad t \geq 1,
\]

with equality for \(t_r\), and hence

\[
f_\circ(a_r) = \max_{t \geq 1} r(t-1) - f(t) = H^*_r.
\]

Next, assume that \(r \geq r_{f,\max}\). Since \(r < +\infty\) by assumption, we only have to consider the case \(r_{f,\max} < +\infty\), which implies \(a_{f,\max} < +\infty\). Note that

\[
\lim_{t \to +\infty} \frac{r(t-1) - f(t)}{t} = r - a_{f,\max}
\]

for every \(t > 1\). Note that \(r \geq r_{f,\max} = f_\circ(a_{f,\max}) + a_{f,\max}\) implies

\[
r - a_{f,\max} \geq f_\circ(a_{f,\max}) \geq a_{f,\max} - f(t)
\]

for every \(t > 1\), from which we obtain

\[
\frac{r + f(t)}{t} \geq a_{f,\max}.
\]

Thus we have

\[
r - a_{f,\max} \geq r - \frac{r + f(t)}{t} = \frac{r(t-1) - f(t)}{t},
\]

and hence \(H^*_r = r - a_{f,\max}\), as required.

We will mainly be interested in the above general Legendre-Fenchel transforms when \(f = \bar{\psi}\) or \(f = \hat{\psi}\). In these special cases we have

\[
\phi(a) := \bar{\psi}_\circ(a) = \sup_{a > 1} \{a(\alpha - 1) - \bar{\psi}(\alpha|\rho|\sigma)\}
\]

\[
\hat{\phi}(a) := \hat{\psi}_\circ(a) = \sup_{a > 1} \{a(\alpha - 1) - \hat{\psi}(\alpha|\rho|\sigma)\}, \quad a \in \mathbb{R},
\]

and for every \(r \geq 0\),

\[
H^*_r(\rho|\sigma) := H^*_{\bar{\psi},r} = \sup_{a > 1} \frac{r(\alpha - 1) - \bar{\psi}(\alpha|\rho|\sigma)}{\alpha}
\]

\[
= \sup_{a > 1} \frac{\alpha - 1}{\alpha} \left[ r - D_\circ(\rho|\sigma) \right],
\]

\[
\hat{H}^*_r(\rho|\sigma) := H^*_{\hat{\psi},r} = \sup_{a > 1} \frac{r(\alpha - 1) - \hat{\psi}(\alpha|\rho|\sigma)}{\alpha}
\]

\[
= \sup_{a > 1} \frac{\alpha - 1}{\alpha} \left[ r - \hat{D}_\circ(\rho|\sigma) \right].
\]

Note that

\[
D_{\bar{\psi},1} = \bar{D}_1(\rho|\sigma), \quad D_{\bar{\psi},\infty} = \bar{D}_\infty(\rho|\sigma),
\]

\[
D_{\hat{\psi},1} = \hat{D}_1(\rho|\sigma), \quad D_{\hat{\psi},\infty} = \hat{D}_\infty(\rho|\sigma).
\]

By lemma \[\ref{lem:anti-divergence} \] we have

\[
0 \leq H^*_r(\rho|\sigma) = \begin{cases} r - a_r = \phi(a_r), & r < \phi(a_{\max}) + a_{\max}, \\ r - a_{\max}, & r \geq \phi(a_{\max}) + a_{\max}, \end{cases}
\]

(51)

where \(a_r\) is the unique solution of \(r - a_r = \phi(a_r)\) and \(a_{\max} = \bar{D}_\infty(\rho|\sigma)\), and

\[
0 < H^*_r(\rho|\sigma) \iff r > \bar{D}_1(\rho|\sigma). \tag{52}
\]

The same relations hold for \(\hat{H}^*_r(\rho|\sigma)\) with \(\hat{\phi}, \hat{D}_1(\rho|\sigma)\) and \(\hat{D}_\infty(\rho|\sigma)\) in place of \(\phi, \bar{D}_1(\rho|\sigma)\) and \(\bar{D}_\infty(\rho|\sigma)\), respectively.

We call \(H^*_r(\rho|\sigma)\) the Hölder anti-divergence of \(\rho\) and \(\sigma\) with parameter \(r\). It differs from the (regularized) Hölder divergence \([4], [24], [30], [41]\) in two ways: first, it is based on the \(D^\circ\) Rényi divergences instead of \(D_\circ\), and second, the optimization is over \(\alpha > 1\) instead of \(\alpha \in (0, 1)\). Due to the latter it is monotone non-decreasing under completely positive trace-preserving maps, which is the reason why we call it an anti-divergence.

We close this section with some observations about a differentiable \(f\). Recall that by \((56)\),

\[
H^*_{f,r} = \sup_{t > 1} \frac{r(t-1) - f(t)}{t} = \sup_{0 < c < 1} \left\{ sr - F(s) \right\}, \tag{53}
\]

with \(F(s) := (1-s)f \left( \frac{1}{1-s} \right)\), where the second equality is due to the change of variables

\[
s := \frac{t-1}{t} \quad \text{so that} \quad t = \frac{1}{1-s}. \tag{54}
\]

For the rest, we will always assume that \(s\) and \(t\) are related as in \((54)\). We start with the following lemma:

\begin{lemma}
Let \(f : (1, +\infty) \to \mathbb{R}\) be a convex function. Then

\[
F : s \mapsto (1-s)f \left( \frac{1}{1-s} \right) \quad \text{is convex on} \quad (0, 1).
\]

\end{lemma}

\begin{proof}
Since \(f\) is convex, it can be written as the supremum of affine functions, i.e., \(f(x) = \sup_{i \in I} \{a_i x + b_i\}\), where \(I\) is some index set, and \(a_i, b_i \in \mathbb{R}\). Hence,

\[
(1-s)f \left( \frac{1}{1-s} \right) = (1-s) \sup_{i \in I} \left\{ \frac{a_i}{1-s} + b_i \right\}
\]

\[
= \sup_{i \in I} \left\{ a_i + b_i(1-s) \right\},
\]

\end{proof}
which, as the supremum of affine functions, is convex in $s$. ■

Assume for the rest that $f$ is differentiable on $(1, +\infty)$, and it is continuous at $1$. Then $F$ is differentiable in $(0, 1)$, and

$$F'(s) = -f \left( \frac{1}{1-s} \right) + \frac{1}{1-s} f' \left( \frac{1}{1-s} \right) = -f(t) + tf'(t).$$

Using the assumption that $\lim_{t \searrow 1} f(t) = f(1) = 0$, we get

$$F'(0^+) := \lim_{s \searrow 0} F'(s) = \lim_{t \searrow 1} (-f(t) + tf'(t)) = \lim_{t \searrow 1} f'(t) = D_{f, 1} = a_f, \min,$n

$$F'(1^-) := \lim_{s \searrow 0} F'(s) = \lim_{t \nearrow \infty} (-f(t) + tf'(t)).$$

Convexity of $F$ guarantees that $F'$ is continuous and monotone increasing, and hence for every $r \in (D_{f, 1}, F'(1^-))$, there exists an $s_r = (t_r - 1)/t_r \in (0, 1)$ such that

$$r = F'(s_r) = -f(t_r) + t_r f'(t_r) = f_o(a_r) + a_r,$n

and hence,

$$H_{f, r} = s_r r - F(s_r) = \frac{t_r - 1}{t_r} (-f(t_r) + t_r f'(t_r)) - \frac{1}{t_r} f(t_r) = -f(t_r) + (t_r - 1) f'(t_r) = f_o(a_r),$$

where

$$a_r := f'(t_r) = r - H_{f, r}.$$

Note that convexity of $f$ implies that

$$D_{f, \infty} = a_{f, \max} = \sup_{1 < t < \infty} \frac{f(t)}{t - 1} = \lim_{t \rightarrow \infty} f'(t) = f'(1^+),$$

and it is easy to see that $a_r \in (D_{f, 1}, D_{f, \infty})$. Note that $a \mapsto f_o(a) + a$ is convex, monotone increasing and lower semicontinuous on $\mathbb{R}$, and hence if $a_{\min} < a_{\max}$, we have

$$r_{f, \max} = f_o(a_{\max}) + a_{\max} = \sup_{a_{\min} < a < a_{\max}} \{ f_o(a) + a \}.$$

Since for every $a \in (a_{\min}, a_{\max})$, there exists an $s = (t - 1)/t \in (0, 1)$ such that $a = f'(t), F'(s) = -f(t) + tf'(t) = f_o(a) + a$, and vice versa, for every $s = (t - 1)/t \in (0, 1)$, we have $a = f'(t) \in (a_{\min}, a_{\max})$ and $F'(s) = f_o(a) + a$, we see that

$$r_{f, \max} = \sup_{a_{\min} < a < a_{\max}} \{ f_o(a) + a \} = \sup_{1 < t < \infty} (-f(t) + tf'(t)) = \sup_{s \in (0, 1)} F'(s) = F'(1^-).$$

IV. THE STRONG CONVERSE EXPONENT IN BINARY HYPOTHESIS TESTING

For every $n \in \mathbb{N}$, let $\mathcal{H}_n, \rho_n$ and $\sigma_n$ be as in Section III-B. As before, we assume that $\sup p \subseteq \supp \sigma$, i.e., $\supp \rho_n \subseteq \supp \sigma_n$ for every $n \in \mathbb{N}$. For every parameter $r > 0$, the lower and upper strong converse exponents $\overline{sc}(r|\rho||\sigma)$ and $\overline{sc}(r|\rho||\sigma)$ of the hypothesis testing problem with null-hypothesis $\rho$ and alternative hypothesis $\sigma$ are defined as

$$\overline{sc}(r|\rho||\sigma) := \inf \left\{ \liminf_{n \rightarrow \infty} -\frac{1}{n} \log Tr \rho_n T_n : \lim sup_{n \rightarrow \infty} -\frac{1}{n} \log Tr \sigma_n T_n \leq -r \right\},$$

$$\overline{sc}(r|\rho||\sigma) := \inf \left\{ \limsup_{n \rightarrow \infty} -\frac{1}{n} \log Tr \rho_n T_n : \lim sup_{n \rightarrow \infty} -\frac{1}{n} \log Tr \sigma_n T_n \leq -r \right\},$$

where the infimum is over all sequences of tests $T_n \in \mathcal{B}(\mathcal{H}_n)_+, T_n \leq I, n \in \mathbb{N}$ (cf. (1)). It is easy to see that $\overline{sc}(r|\rho||\sigma)$ can be alternatively expressed as

$$\overline{sc}(r|\rho||\sigma) = \sup \left\{ R : \forall \{T_n\}_{n=1}^{\infty}, 0 \leq T_n \leq I, \lim sup_{n \rightarrow \infty} -\frac{1}{n} \log Tr \rho_n T_n \leq -r \right\}.$$

and similar expressions hold for $\overline{sc}(r|\rho||\sigma)$ as well.

The following lemma is essentially due to [40] and [23], the only difference is that we use $D_{\alpha}^*$ instead of $D_{\alpha}$.

**Lemma IV.1.** For any $r \geq 0$, we have $\overline{sc}(r|\rho||\sigma) \geq H_{f}^*(\rho|\sigma)$.

Proof: Let $T_n \in \mathcal{B}(\mathcal{H}_n)$ be a test and let $p_n := (Tr \rho_n T_n, Tr \rho_n (I - T_n))$ and $q_n := (Tr \sigma_n T_n, Tr \sigma_n (I - T_n))$ be the post-measurement probability distributions. By the monotonicity of the Rényi divergences under measurements, we have, for any $\alpha > 1$,

$$D_{\alpha}^*(\rho_n || \sigma_n) \geq D_{\alpha}^* (p_n || q_n) \geq \frac{1}{\alpha - 1} \log [(Tr \rho_n T_n)^{\alpha} (Tr \sigma_n T_n)^{1-\alpha}] = \frac{1}{\alpha - 1} \log Tr \rho_n T_n - \log Tr \sigma_n T_n,$$

or equivalently,

$$\frac{1}{\alpha - 1} \log Tr \rho_n T_n \leq \frac{1}{\alpha} \left[ \frac{1}{n} D_{\alpha}^* (\rho_n || \sigma_n) + \frac{1}{n} \log Tr \sigma_n T_n \right].$$

In the limit $\lim sup_{n \rightarrow \infty} \frac{1}{n} \log Tr \sigma_n T_n \leq -r$ then

$$\lim sup_{n \rightarrow \infty} \frac{1}{n} \log Tr \rho_n T_n \leq \frac{1}{\alpha} \left[ D_{\alpha}^* (\rho_n || \sigma) - r \right], \quad \alpha > 1.$$

Taking the infimum in $\alpha > 1$, and multiplying both sides by $-1$, the assertion follows. ■

It is known that the inequality in Lemma [IV.1] holds as an equality in the i.i.d. case [37], and our aim is to extend
We prove the claim by considering three different regions of $a$ to (43). For any $r$ such that $r \in (42)$. For every $n \to \infty$, and (59) and (60) yield
\[
\lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n(a) \geq -f_o(a).
\] Then
\[
\sigma(r) \leq H^*_r, \quad r \geq 0.
\] Proof: Due to the representation (38) of $\sigma(r)$ as an infimum of rates, it is sufficient to show that for any rate $R > H^*_r$, there exists a sequence of tests $\{T_n\}_{n=1}^{\infty}$ satisfying
\[
\limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n \leq -r \quad \text{and} \quad \liminf_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n T_n \geq -R.
\] We prove the claim by considering three different regions of $r$.

(i) In the case $D_{f,1} < r < r_{f,\max}$, there exists a unique $a_r \in (D_{f,1}, D_{f,}\infty)$ satisfying $r - a_r = f_o(a_r)$, and (59) and (60) yield
\[
\limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n(a_r) \leq -(f_o(a_r) + a_r) = -r,
\] and
\[
\liminf_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n T_n(a_r) \geq -f_o(a_r) = H^*_r, r.
\] where the last identity is due to (42).

(ii) In the case $0 \leq r \leq D_{f,1}$, we have $H^*_r = 0$, according to (43). For any $R > 0$, we can find an $a \in (D_{f,1}, D_{f,}\infty)$ such that $0 < f_o(a) < R$. Note that $f_o(a) + a > D_{f,1} \geq r$, and (59) and (60) yield
\[
\limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n(a) \leq -(f_o(a) + a) < -r,
\] and
\[
\liminf_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n T_n(a) \geq -f_o(a) > -R.
\]

(iii) In the case $r \geq r_{f,\max}$, we use a modification of the tests $T_n(a)$, following the method of the proof of Theorem 4 in [42]. For every $a, r \in \mathbb{R}$, let
\[
T_n(r, a) := e^{-\frac{n(r - a - f_o(a))}{T_n(a)}},
\] If $a \in (D_{f,1}, D_{f,}\infty)$ and $r \geq r_{f,\max}$ then $r > f_o(a) + a$, and hence $0 \leq T_n(r, a) \leq 1$, i.e., $T_n(r, a)$ is a test, and
\[
\limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n(r, a) \leq -r + a + f_o(a) - (a + f_o(a)) = -r,
\] and
\[
\liminf_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n T_n(r, a) \leq -r + a + f_o(a) - f_o(a) = -(r - a),
\] by (59) and (60). Now for any $R > H^*_r = r - D_{f,\infty}$, we can find an $a \in (D_{f,1}, D_{f,}\infty)$ such that $r - D_{f,\infty} < r - a < R$, and the assertion follows.

Specializing to $f = \psi$ in the above Theorem yields the following:

**Theorem IV.3.** Assume that for every $a \in (D_1(\rho||\sigma), D_\infty(\rho||\sigma))$ there exists a sequence of tests $0 \leq T_n(a) \leq I_n$, $n \in \mathbb{N}$, such that
\[
\limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n(a) \leq -(\phi(a) + a),
\] and
\[
\liminf_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n T_n(a) \geq -\phi(a),
\] where $\phi$ is given in (49). Then
\[
\sigma(r) \leq H^*_r, \quad r \geq 0.
\] Proof: Immediate from Lemma IV.1 and Theorem IV.2.

**Remark IV.4.** The separate treatment of two different regions of $r$ values for the strong converse exponent, as in (77) and (78) in Theorem IV.2, dates back to (33), where it was noted that randomized tests are necessary for $r$ values above a critical one.

**A. States with differentiable $\psi = \hat{\psi}$**

Now we fix a sequence $\{\hat{\sigma}_n\}_{n \in \mathbb{N}}$ satisfying (21) and, as before, we denote by $\hat{\rho}_n$ the pinching of $\rho_n$ by $\hat{\sigma}_n$. Let
\[
\hat{S}_n(a) := \{\hat{\rho}_n - e^{-a\hat{\sigma}_n} > 0\}
\] be a Neyman-Pearson test for every $a \in \mathbb{R}$ and every $n \in \mathbb{N}$.

**Theorem IV.5.** Assume that for every $a > 1$, $\hat{\psi}(a|\rho||\sigma)$ exists as a limit, and $a \mapsto \hat{\psi}(a|\rho||\sigma)$ is differentiable on $(1, +\infty)$. Then
\[
\limsup_{n \to +\infty} \frac{1}{n} \log \text{Tr} \sigma_n \hat{S}_n(a) \leq \limsup_{n \to +\infty} \frac{1}{n} \log \text{Tr} \hat{\sigma}_n \hat{S}_n(a) = -(\phi(a) + a),
\] and
\[
\limsup_{n \to +\infty} \frac{1}{n} \log \text{Tr} \rho_n \hat{S}_n(a) = -\hat{\phi}(a).
\] for every $a \in (\hat{D}_1(\rho||\sigma), \hat{D}_\infty(\rho||\sigma))$, where $\hat{\phi}$ is given in (50). and
\[
H^*_r(\rho||\sigma) \leq \sigma(r) \leq \hat{\sigma}(r||\sigma) \leq \hat{H}^*_r(\rho||\sigma), \quad r \geq 0.
\] If, moreover, $\hat{\psi}(\alpha|\rho||\sigma) = \hat{\psi}(\alpha|\rho||\sigma)$ for every $\alpha > 1$ then
\[
\sigma(r) \leq \hat{\sigma}(r||\sigma) \leq \hat{H}_r^*(\rho||\sigma), \quad r \geq 0.
\] Proof: First, note that $\hat{S}_n(a) = \mathcal{E}_{\hat{\sigma}_n}(\hat{S}_n(a))$, and hence
\[
\text{Tr} \rho_n \hat{S}_n(a) = \text{Tr} \rho_n \mathcal{E}_{\hat{\sigma}_n}(\hat{S}_n(a)) = \text{Tr} \mathcal{E}_{\hat{\sigma}_n}(\rho_n) \hat{S}_n(a) = \text{Tr} \hat{\rho}_n \hat{S}_n(a).
\] Since $\hat{\rho}_n$ and $\hat{\sigma}_n$ commute, we may consider them as probability mass functions on some finite set $\mathcal{X}_n$, and write
\[
\text{Tr} \sigma_n \hat{S}_n(a) \leq \text{Tr} \hat{\sigma}_n \hat{S}_n(a) = \mathbb{P}_{\hat{\sigma}_n}\{x \in \mathcal{X}_n : Y_n(x) > a\}
\] and
\[
\text{Tr} \rho_n \hat{S}_n(a) = \mathbb{P}_{\hat{\rho}_n}\{x \in \mathcal{X}_n : Y_n(x) > a\}
\]
where $Y_n(x) := \frac{1}{n} \log h_n(x, \sigma_n)$ and $\mu_{n,1}$ and $\mu_{n,2}$ are probability measures on $\mathbb{R}$, defined for any Borel subset $H$ of $\mathbb{R}$ by 
\[ \mu_{n,1}(H) := \mathbb{P}_{\tilde{s}_n} \left( \{ x \in X_n : Y_n(x) \in H \} \right), \]
\[ \mu_{n,2}(H) := \mathbb{P}_{\rho_n} \left( \{ x \in X_n : Y_n(x) \in H \} \right). \]
The first inequality in (71) is due to (21). Let $\Lambda_{n,1}$ and $\Lambda_{n,2}$ be the logarithmic moment generating functions of $\mu_{n,1}$ and $\mu_{n,2}$ respectively (see Appendix A). Then we have 
\[ \Lambda_{n,1}(nt) = \log \mathbb{E}_{\tilde{s}_n} e^{t \log(\tilde{\rho}_n / \tilde{\sigma}_n)} = \log \sum_{x \in X_n} \tilde{\rho}_n(x)^t \tilde{\sigma}_n(x)^{1-t}, \]
\[ = \log \text{Tr} \tilde{\rho}_n^{t} \tilde{\sigma}_n^{1-t} = \psi(t \tilde{\rho}_n / \tilde{\sigma}_n), \]
\[ \Lambda_{n,2}(nt) = \log \mathbb{E}_{\rho_n} e^{t \log(\tilde{\rho}_n / \tilde{\sigma}_n)} = \log \sum_{x \in X_n} \rho_n(x)^t \tilde{\sigma}_n(x)^{1-t}, \]
\[ = \log \text{Tr} \tilde{\rho}_n^{t} \tilde{\sigma}_n^{1-t} = \psi(1 + t \tilde{\rho}_n / \tilde{\sigma}_n). \]

By assumption, 
\[ \mathcal{X}_1(t) := \lim_{n \to +\infty} \Lambda_{n,1}(nt) = \psi(\alpha / \sigma), \quad t > 1, \]
\[ \mathcal{X}_2(t) := \lim_{n \to +\infty} \Lambda_{n,2}(nt) = \psi(1 + t \alpha / \sigma), \quad t > 0. \]

By convexity, it is easy to see that 
\[ \lim_{\alpha \to 1} \mathcal{X}_1(t) = \bar{D}_1(\rho / \sigma), \quad \lim_{t \to +\infty} \mathcal{X}_1(t) \geq \bar{D}_\infty(\rho / \sigma), \]

and, similarly, 
\[ \lim_{\alpha \to 0} \mathcal{X}_2(t) = \bar{D}_1(\rho / \sigma), \quad \lim_{t \to +\infty} \mathcal{X}_2(t) \geq \bar{D}_\infty(\rho / \sigma). \]

Using now Lemmas A.1 and A.2 we get that for every $a \in (\bar{D}_1(\rho / \sigma), \bar{D}_\infty(\rho / \sigma))$,
\[ \lim_{n \to +\infty} \frac{1}{n} \log \text{Tr} \rho_n \tilde{S}_n(a) = \lim_{n \to +\infty} \frac{1}{n} \log \mu_{n,1}((a, +\infty)) = - \sup_{t \geq 1} \{ at - \psi(1 + t \alpha / \sigma) \} = - (\psi(a) + a), \]
\[ \lim_{n \to +\infty} \frac{1}{n} \log \text{Tr} \rho_n \tilde{S}_n(a) = \lim_{n \to +\infty} \frac{1}{n} \log \mu_{n,2}((a, +\infty)) = - \sup_{t > 0} \{ at - \psi(t \tilde{\rho}_n / \tilde{\sigma}_n) \} = - (\psi(a) + a), \]

proving the identities in (67)-(69). The inequality in (67) is obvious from the inequality in (71).

Applying Theorem IV.2 with $f := \tilde{\psi}$ and $T_n(a) := \tilde{S}_n(a)$ yields the last inequality in (69), and the first inequality is immediate from Lemma IV.7. Finally, if $\tilde{\psi}(\alpha / \sigma) = \psi(\alpha / \sigma)$ for every $\alpha > 1$ then $H_{\sigma}(\rho / \sigma) = H_{\sigma}(\rho / \sigma)$ for every $r$, and (69) reduces to (70).

Combining Theorem IV.5 and Corollary III.11 we get immediately the following:

**Corollary IV.6.** Assume that for every $\alpha > 1$, $\tilde{\psi}(\alpha / \sigma)$ exists as a limit, and $\alpha \to \tilde{\psi}(\alpha / \sigma)$ is differentiable on $(1, +\infty)$. Assume also that $\lim_{n \to +\infty} \frac{1}{n} \log \theta(\sigma_n) = 0$. Then 
\[ \underline{s}(r) = \overline{s}(r) = H_{\sigma}(\rho / \sigma), \quad r \geq 0. \]

Moreover, the optimal sequence of tests can be chosen as in (66), with $\sigma_n$ as in (iii) of Corollary III.11.

We say that the hypothesis testing problem is i.i.d. if $\mathbb{H}_n = \mathbb{H}_n^{\sigma_n}$, $\rho_n = \rho_n^{\sigma_n}$ and $\sigma_n = \sigma_n^{\sigma_n}$ for every $n \in \mathbb{N}$. Let $H^*_n(\rho_n / \sigma_n)$ be as given in (4). An expression for the strong converse exponent in the i.i.d. case was first given in (73) using the tests $\tilde{S}_n(a)$ corresponding to the choice $\tilde{\sigma}_n := \sigma_n$.

There it was shown that the inequality 
\[ \underline{s}(r) = \overline{s}(r) \geq H^*_n(\rho_n / \sigma_n) := \sup_{\alpha > 0} \frac{1}{\alpha} \left[ r - \frac{1}{n} \mathbb{E} \left[ \log \mathbb{E}_{\sigma_n} \rho_n / \sigma_n \right] \right] \]

(74)

holds (cf. Lemma IV.1); the converse inequality can be obtained by applying the classical strong converse result of (18) to the commuting states $\tilde{\rho}_n = \mathbb{E}_{\sigma_n} \rho_n$ and $\sigma_n$. It was shown later in (37) that $\mathbb{E}_{\mathcal{E}_{\sigma_n}} \rho_n / \sigma_n = \mathbb{E}_{\mathcal{E}_{\sigma_n}} \rho_n / \sigma_n$; with this addition, (74) yields Lemma IV.1. The strong converse exponent was later shown to be equal to $H^*_n(\rho_n / \sigma_n)$ in (67), by showing that (67)-(69) hold with $\tilde{S}_n(a) := \{ \rho_n - e^{n \alpha} \sigma_n > 0 \}$ in place of $\tilde{S}_n(a)$. Here we give an alternative proof, based on Theorem IV.5.

Note that neither the proof in (37), nor the proof below uses the different expression, based on the Hellinger arc, in (43), where it was also explained how the expression in (18) can obtained from the one in (43). In Appendix [D] we briefly explain how the expressions in (43) can be obtained from Theorem IV.7.

**Theorem IV.7.** In the i.i.d. case, $\underline{s}(r) = \overline{s}(r) = H^*_n(\rho_n / \sigma_n)$ for every $r \geq 0$.

**Proof:** It is easy to see that the i.i.d. assumption implies $\tilde{\psi}(\alpha / \sigma) = \psi(\alpha / \sigma)$, and thus also $H^*_n(\rho_n / \sigma_n) = H^*_n(\rho_n / \sigma_n)$. The choice $\tilde{\sigma}_n := \sigma_n = \sigma_n^{\sigma_n}$ yields $\psi(\sigma_n)$ grows polynomially with $n$, and $\mathbb{D}_{\max}(\sigma_n / \sigma_n) = 0$, and hence, by Lemma III.9 $\psi(\alpha / \sigma) = \psi(\alpha / \sigma)$. Finally, differentiability of $\tilde{\psi}(\alpha / \sigma)$ exists for every $\alpha > 1$ follows from Lemma III.11. Thus, all the conditions of Theorem IV.5 are satisfied, and therefore (70) holds.

An expression for the strong converse exponent in the classical i.i.d. case was first given in (18), followed by a different expression, based on the Hellinger arc, in (43), where it was also explained how the expression in (18) can obtained from the one in (43). In Appendix [D] we briefly explain how the expressions in (43) can be obtained from Theorem IV.7.

In (30) Example B.1] a class of finitely correlated states [15] with commutative auxiliary algebra has been studied, and it has been shown that for these states, $\lim_{n \to +\infty} \frac{1}{n} \psi(\alpha / \sigma_n)$ is differentiable in $\alpha$ for every $\alpha \in \mathbb{R}$. In particular, this class includes all Markovian chains with an irreducible transition matrix. Exactly the same argument as in (30) yields that for this class of states, also $\tilde{\psi}(\alpha / \sigma)$ exists as a limit and is differentiable in $\alpha$ for $\alpha > 1$. It is also easy to verify that if $\sigma$ is in this class then $\lim_{n \to +\infty} \frac{1}{n} \psi(\sigma_n) = 0$, and thus $\tilde{\psi}(\alpha / \sigma)$ exists due to Lemma III.9. In particular, the strong converse exponent can be expressed as (70), due to Theorem IV.5. An alternative expression for the strong converse exponent of classical Markov chains was given before
in \[43\]. In Appendix \[13\] we explain how the exponent of \[44\] can be obtained from Theorem \[IV.5\].

In Section \[IV.B\] we show that Theorem \[IV.5\] can be applied to obtain the strong converse exponent for the hypothesis testing problem of gauge-invariant fermionic quasi-free states.

### B. States with factorization property

Let \( \mathcal{H} \) be a finite-dimensional Hilbert space, and for every \( n \in \mathbb{N} \), let \( \omega_n \) be a state on \( \mathcal{H} \otimes \mathcal{H}^n \). We say that \( \omega := \{ \omega_n \}_{n \in \mathbb{N}} \) satisfies the factorization property if there exists an \( \eta \geq 1 \) such that for every \( k, m, r \in \mathbb{N} \),

\[
\omega_{km+r} \leq \eta^k \omega_k \otimes \omega_r \quad \text{(upper factorization)}, \quad \omega_{kn+m+r} \geq \eta^{-k} \omega_k \otimes \omega_r \quad \text{(lower factorization)}.
\]

We call \( \eta \) a factorization constant for \( \omega \). Note that if \( \rho = \{ \rho_n \}_{n \in \mathbb{N}} \) and \( \sigma = \{ \sigma_n \}_{n \in \mathbb{N}} \) both satisfy the factorization property then we can always choose an \( \eta \) which is a common factorization constant for both \( \rho \) and \( \sigma \).

Obviously, if \( \omega \) is i.i.d., i.e., of the form \( \omega_n = \omega \otimes \mathcal{H}^n \), \( n \in \mathbb{N} \), then it satisfies the factorization property with \( \eta = 1 \). It has been shown in \[29\] that finitely correlated states \[15\] satisfy the factorization property with \( \eta = 1 \).

In Section \[V.A\] we start with showing that under the factorization assumption, \( \psi \) exists as a limit and give bounds on its deviation from the \( \psi \) functions for finite \( n \).

#### Lemma IV.8

Let \( \rho \) and \( \sigma \) satisfy the factorization property, and let \( \eta \) be a common factorization constant. Then \( \overline{T}(\alpha|\rho||\sigma) \) exists as a limit for every \( \alpha > 1 \), and

\[
\overline{T}(\alpha|\rho||\sigma) - \frac{2\alpha - 1}{n} \log \eta \leq \frac{1}{n} \sum_{a}^{\omega^*}(\alpha|\rho_n||\sigma_n) \leq \overline{T}(\alpha|\rho||\sigma) + \frac{2\alpha - 1}{n} \log \eta
\]

for every \( \alpha > 1 \) and every \( n \in \mathbb{N} \).

**Proof:** Given \( m \in \mathbb{N} \), every \( n \in \mathbb{N} \) can be uniquely written in the form \( n = km + r \) with \( k, r \in \mathbb{N} \), \( r \in \{0, \ldots, m-1\} \). Since \( \alpha > 1 \), we have \( -1 < \frac{\alpha}{\alpha} - 1 < 0 \), and hence \( x \mapsto x^{-\alpha} \) is operator monotone decreasing. Thus

\[
\eta^{-k\alpha^{-1}}(\sigma_k \otimes \sigma_r)^{\frac{\alpha}{\alpha}} \leq \frac{1}{n} \sigma_n \leq \eta^{k\alpha^{-1}}(\sigma_k \otimes \sigma_r)^{\frac{\alpha}{\alpha}}.
\]

Taking into account that \( A \mapsto \text{Tr} A^\alpha \) is monotone increasing w.r.t. the positive semidefinite ordering, we obtain

\[
Q^*_\alpha(\rho_n||\sigma_n) = \frac{\text{Tr}(\rho_n^{1/2} \sigma_n^{\frac{\alpha}{\alpha}} \rho_n^{1/2})^\alpha}{\sigma_n^{\frac{\alpha}{\alpha}}} \leq \eta^{k\alpha^{-1}}(\sigma_k \otimes \sigma_r)^{\frac{\alpha}{\alpha}} \leq \frac{1}{n} \sigma_n \leq \eta^{-k\alpha^{-1}}(\sigma_k \otimes \sigma_r)^{\frac{\alpha}{\alpha}}.
\]

Thus

\[
\limsup_{n \to \infty} \frac{1}{n} \log Q^*_\alpha(\rho_n||\sigma_n) \leq \frac{2\alpha - 1}{n} \log \eta + \frac{1}{m} \log Q^*_\alpha(\rho_n||\sigma_m).
\]

Taking now the liminf in \( m \), we get that \( \overline{T}(\alpha|\rho||\sigma) \) exists as a limit, and the first inequality in (76) holds, for every \( \alpha > 1 \).

Using the lower factorization for \( \rho \) and upper factorization for \( \sigma \), an analogous argument to the one above yields the second inequality in (76).

**Corollary IV.9.** For every \( \alpha \in (1, +\infty) \), we have

\[
\overline{T}_\alpha(\rho||\sigma) = \lim_{n \to \infty} \frac{1}{n} D^*_\alpha(\rho_n||\sigma_n).
\]

**Lemma IV.10.** Assume that \( \rho \) and \( \sigma \) satisfy the factorization property. Then

\[
\limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n \sigma_n(a) \leq -\phi(a),
\]

\[
\limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n(\rho_n||\sigma_n(a)) \leq -(\phi(a) + \alpha).
\]

for any \( a \in \mathbb{R} \).

**Proof:** First, we prove that

\[
\text{Tr} \rho_n \sigma_n(a) \leq e^{-\alpha a^{-1}} Q^*_\alpha(\rho_n||\sigma_n)
\]

for every \( n \in \mathbb{N} \), \( \alpha \geq 1 \) and \( n \in \mathbb{N} \). Indeed, this inequality holds trivially if \( S_n(a) = 0 \). Otherwise we can use

\[
\text{Tr} \rho_n \sigma_n(a) \geq e^{\alpha a} \text{Tr} \sigma_n(\rho_n||\sigma_n(a))
\]

to show that for \( \alpha \geq 1 \),

\[
\text{Tr} \rho_n S_n(a) = \{ \text{Tr} \rho_n S_n(a) \}^\alpha \{ \text{Tr} \rho_n S_n(a) \}^{1-\alpha}
\]

\[
\leq e^{-\alpha a^{-1}} \{ \text{Tr} \rho_n S_n(a) \}^{1-\alpha}
\]

\[
\leq e^{-\alpha a^{-1} \{ \text{Tr} \rho_n S_n(a) \}} \{ \text{Tr} \sigma_n(\rho_n||\sigma_n(a)) \}^{1-\alpha}
\]

\[
\leq e^{-\alpha a^{-1} \{ \text{Tr} \sigma_n(a) \} ||\sigma_n)}
\]
where the last inequality is due to Lemma 11.5. From (81) we obtain
\[
\frac{1}{n} \log \Tr \rho_n S_n(a) \leq - \left\{ a(\alpha - 1) - \frac{1}{n} \log Q_\alpha^\ast (\rho_n || \sigma_n) \right\},
\]
and taking first the limsup in \( n \) and then the infimum over \( \alpha > 1 \) yields (79). Finally, combining (82) with (79) yields (80).

**Lemma IV.11.** For any \( A, B \in \mathcal{B}(\mathcal{H})_+ \), and any \( c \in (D_1(A \| B), D^\infty_\alpha(A \| B)) \), we have
\[
\lim_{n \to +\infty} \frac{1}{n} \log \Tr \left( A^{\otimes n} - e^{nc} B^{\otimes n} \right)_+ = (\Tr A)^n \left\{ \log \Tr \left( A^{\otimes n} - e^{nc} B^{\otimes n} \right) + \frac{1}{n} \log \Tr \left( \tilde{A}^{\otimes n} - e^{nc} \tilde{B}^{\otimes n} \right) \right\}.
\]

**Proof:** When \( \Tr A = \Tr B = 1 \), the assertion follows from Theorem IV.4 in (77). In general, let \( \tilde{A} := A/\Tr A, \tilde{B} := B/\Tr B \). Then
\[
\Tr \left( A^{\otimes n} - e^{nc} B^{\otimes n} \right)_+ = (\Tr A)^n \left\{ \log \Tr \left( A^{\otimes n} - e^{nc} \tilde{B}^{\otimes n} \right) + \frac{1}{n} \log \Tr \left( \tilde{A}^{\otimes n} - e^{nc} \tilde{B}^{\otimes n} \right) \right\}.
\]

By (20), we have \( D^\ast_\alpha(A \| B) = \log \Tr A - \log \Tr B + D^\ast_\alpha(A \| B) \), and hence \( c + \log \Tr B - \log \Tr A \in (D_1(A \| B), D^\infty_\alpha(A \| B)) \). Thus, by (83) and Theorem IV.4 in (77), we have
\[
\lim_{n \to +\infty} \frac{1}{n} \log \Tr \left( A^{\otimes n} - e^{nc} B^{\otimes n} \right)_+ = \log \Tr \left( A^{\otimes n} - e^{nc} \tilde{B}^{\otimes n} \right) - \frac{1}{n} \log \Tr \left( \tilde{A}^{\otimes n} - e^{nc} \tilde{B}^{\otimes n} \right),
\]
where the last equality is due to (19).

**Lemma IV.12.** Assume that \( \rho \) and \( \sigma \) satisfy the factorization property. Then
\[
\liminf_{n \to +\infty} \frac{1}{n} \log \rho_n S_n(a) \geq \liminf_{n \to +\infty} \frac{1}{n} \log \Tr (\rho_n - e^{na} \sigma_n)_+ \geq -\phi(a)
\]
for every \( a \in (\mathcal{T}_1(\rho || \sigma), \mathcal{T}_\infty(\rho || \sigma)) \).

**Proof:** We will assume that \( \mathcal{T}_1(\rho || \sigma) \neq \mathcal{T}_\infty(\rho || \sigma) \), since otherwise the statement is empty. Let \( \eta \) denote a common factorization constant for \( \rho \) and \( \sigma \), and let \( b \in \mathbb{R} \) be such that \( \mathcal{T}_1(\rho || \sigma) < a < b < D_\infty(\rho || \sigma) \). Due to (27), there exist \( 1 < \alpha_1 < \alpha_2 < +\infty \) such that \( D_\alpha(\rho || \sigma) < a < b < D_\alpha(\rho || \sigma) \). Note that for every \( \alpha > 1 \),
\[
\frac{1}{m} D^\ast_\alpha (\eta^{-1} \rho_n || \eta \sigma_m) = -\frac{1}{m} \log \eta^2 + \frac{1}{m} D^\ast_\alpha (\rho_n || \sigma_m) \to D_\alpha(\rho || \sigma),
\]
where the limit follows from (78). Thus we see the existence of \( m_b \) such that for all \( m \geq m_b \),
\[
\frac{1}{m} D_1(\eta^{-1} \rho_n || \eta \sigma_m) \leq \frac{1}{m} D^\ast_\alpha (\eta^{-1} \rho_m || \eta \sigma_m) < b < \frac{1}{m} D^\ast_\alpha (\eta^{-1} \rho_m || \eta \sigma_m) \leq \frac{1}{m} D^\ast_\alpha (\eta^{-1} \rho_m || \eta \sigma_m),
\]
where the first and the last inequalities are due to the monotonicity of the Rényi divergences in the parameter \( \alpha \).

For a fixed \( m \geq m_b \), we can write every \( n > m \) uniquely as \( n = km + r \) with \( k \in \mathbb{N} \) and \( r \in \{1, \ldots, m\} \). Then we have
\[
\Tr \rho_n S_n(a) = \Tr (\rho_n - e^{na} \sigma_n) S_n(a) + e^{na} \Tr \sigma_n S_n(a)
\]
\[
\geq \Tr (\rho_n - e^{na} \sigma_n)_+ \geq \Tr (\rho_{km} - e^{na} \sigma_{km})_+ \geq \Tr (\eta^{-1} \rho_m || \eta \sigma_m)_+ + e^{na} \Tr (\eta^{-1} \rho_m || \eta \sigma_m)_+,
\]
where the second inequality follows from the monotonicity of the Rényi divergences applied to the partial trace over subsystems \( km + 1 \) to \( n \), and the last inequality is due to the factorization properties and (8). Note that
\[
na = (km + r)a \leq km + ma = km \left(a + \frac{a}{k}\right) < km b,
\]
whenever \( k > a/(b - a) \), and for any such \( k \) we have
\[
\Tr (\rho_n - e^{na} \sigma_n)_+ \geq \Tr (\eta^{-1} \rho_m || \eta \sigma_m)_+ + e^{na} \Tr (\eta^{-1} \rho_m || \eta \sigma_m)_+ \geq \Tr (\eta^{-1} \rho_m || \eta \sigma_m)_+ + e^{na} \Tr (\eta^{-1} \rho_m || \eta \sigma_m)_+,
\]
due to (86) and (8). By (85), \( mb \in (D_1(\eta^{-1} \rho_m || \eta \sigma_m), D^\infty_\alpha(\eta^{-1} \rho_m || \eta \sigma_m)) \), and hence (87) and Lemma IV.11 yield
\[
\liminf_{n \to +\infty} \frac{1}{n} \log \Tr (\rho_n - e^{na} \sigma_n)_+ \geq \frac{1}{m} \liminf_{k \to +\infty} \frac{1}{k} \log \Tr (\eta^{-1} \rho_m || \eta \sigma_m)_+, = \frac{1}{m} \sup_{\alpha > 1} \left\{ mb(\alpha - 1) - \phi(\alpha || \eta^{-1} \rho_m || \eta \sigma_m) \right\},
\]
and hence (88).

By (19) and (76), we have
\[
\frac{1}{m} \psi(\alpha || \eta^{-1} \rho_m || \eta \sigma_m) = \frac{1}{m} \psi(\alpha || \eta \rho_m || \sigma_m) - \frac{(2a - 1)}{m} \log \eta
\]
\[
\geq \frac{1}{m} \psi(\alpha || \rho \sigma) - \frac{(4a - 2)}{m} \log \eta.
\]
Combining the above inequality and (88), we have
\[
\liminf_{n \to +\infty} \frac{1}{n} \log \Tr (\rho_n - e^{na} \sigma_n)_+ \geq \frac{1}{m} \sup_{\alpha > 1} \left\{ b + \frac{4}{m} \log \eta \right\} (\alpha - 1) - \frac{2}{m} \log \eta
\]
\[
= -\phi \left( b + \frac{4}{m} \log \eta \right) - \frac{2}{m} \log \eta.
\]
Note that \( \phi \) is continuous on \( (\mathcal{T}_1(\rho || \sigma), \mathcal{T}_\infty(\rho || \sigma)) \) and \( b + \frac{4}{m} \log \eta \in (\mathcal{T}_1(\rho || \sigma), \mathcal{T}_\infty(\rho || \sigma)) \) for all sufficiently large \( m \). Hence, by taking the limit \( m \to \infty \), we get
\[
\liminf_{n \to +\infty} \frac{1}{n} \log \Tr (\rho_n - e^{na} \sigma_n)_+ \geq -\phi(b).
\]
Finally, taking the limit $b \searrow a$, we get
\[
\liminf_{n \to \infty} \frac{1}{n} \log \text{Tr}(\rho_n - e^{na} \sigma_n) + \geq -\phi(a). \tag{89}
\]
Now (86) and (89) lead to the assertion. \hfill \blacksquare

**Theorem IV.13.** Assume that $\rho$ and $\sigma$ satisfy the factorization property. Then
\[
\lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n S_n(a) = \lim_{n \to \infty} \frac{1}{n} \log \text{Tr} (\rho_n - e^{na} \sigma_n) +
\]
\[
= -\phi(a), \tag{90}
\]
\[
\lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n S_n(a) = -(\phi(a) + a) \tag{91}
\]
for every $a \in (D_1(\rho||\sigma), D_{\infty}(\rho||\sigma))$. In particular, the conditions of Theorem IV.3 are satisfied with $T_n(a) := S_n(a)$, and hence
\[
\sigma(r||\sigma) = \sigma(r||\sigma) = H_r^s(\rho||\sigma), \quad r \geq 0. \tag{92}
\]

**Proof:** The identities in (90) are immediate from (79) and (83). By (80), $\limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n S_n(a) \leq -(\phi(a) + a)$, and the same argument as in [37] Theorem IV.5 yields $\frac{1}{n} \log \text{Tr} S_n(a) \geq -(\phi(a) + a)$, proving (91). Finally, (92) follows from Theorem IV.3. \hfill \blacksquare

**Remark IV.14.** It has been shown in [42] Theorem 4 that for general sequences $\{\rho_n\}_{n \in \mathbb{N}}, \{\sigma_n\}_{n \in \mathbb{N}}$ as in Section III-B
\[
\sigma(r||\sigma) = \inf_a \max \{r - a, \tau(a)\}, \tag{93}
\]
\[
\tau(a) := \limsup_n \frac{1}{n} \log \text{Tr} \rho_n S_n(a), \tag{94}
\]
where $S_n(a)$ is given in (75). (Note that the roles of $\rho$ and $\sigma$ are reversed here as compared to [42], which is the reason why we have $r - a$ instead of $r + a$ as in [42] Theorem 4.) By (90), if both $\rho$ and $\sigma$ satisfy the factorization property then $\tau(a) = \phi(a)$ for $a \in (D_1(\rho||\sigma), D_{\infty}(\rho||\sigma))$. It is easy to see that both $\tau$ and $\phi$ are non-negative and monotone increasing. Since $\lim_{a \to D_{\infty}(\rho||\sigma)} \phi(a) = 0$, we get $\phi(a) = 0 = \tau(a)$ for every $a \leq D_{\infty}(\rho||\sigma)$. It is also clear from the definitions that $\phi(a) = +\infty = \tau(a)$ for every $a > D_{\infty}(\rho||\sigma)$. Hence, $\tau(a) = \phi(a)$ for all $a \in \mathbb{R} \setminus \{D_{\infty}(\rho||\sigma)\}$, from which we obtain
\[
\sigma(r||\sigma) = \inf_a \max \{r - a, \tau(a)\} = \inf_a \max \{r - a, \phi(a)\}. \tag{95}
\]

(It is easy to see (e.g., by drawing a picture of the graphs of $\tau$, $\phi$ and $a \mapsto r - a$) that the values of these functions at $D_{\infty}(\rho||\sigma)$ do not play a role in the validity of the above identity.) By exactly the same argument as in [37] Lemma IV.16, we have
\[
\inf_a \max \{r - a, \phi(a)\} = H_r^s(\rho||\sigma). \tag{96}
\]

This gives an alternative derivation of (92), based on (92)–(96) and (90)–(91), and without using Theorems IV.2, IV.3.

**V. Examples**

**A. Gibbs states on spin chains**

Let $\mathcal{H}$ be a finite-dimensional Hilbert space. A translation-invariant, finite-range interaction $\Phi$ on $\mathcal{H}$ is specified by a number $r \in \mathbb{N}$, and $\Phi_j \in B(\mathcal{H}^j)$, $j \in [r] := \{1, \ldots, r\}$, where each $\Phi_j$ is self-adjoint. For every $n \in \mathbb{N}$, the local Hamiltonian $H_n$ corresponding to $\Phi$ is defined as
\[
H_n := \sum_{j=1}^r \sum_{k+j \leq n} \Phi_{j,n,k}, \quad \Phi_{j,n,k} := (\delta_{i,j}^{-1} I) \otimes \Phi_j \otimes (\delta_{i,k+j}^{-1} I)
\]
where each $\Phi_j$ is specified by a translation-invariant, finite-range local Gibbs state $\omega_n^{\Phi_j}$ on $n$ sites at inverse temperature $\beta > 0$.

The corresponding local Gibbs state $\omega_n^{\Phi}$ on $n$ sites at inverse temperature $\beta > 0$ is defined as
\[
\omega_n^{\Phi} := e^{-\beta H_n} \text{Tr} e^{-\beta H_n}.
\]

The thermodynamic limit (TDL) Gibbs state $\omega_n^{\Phi}$ on $n$ sites is then given by
\[
\text{Tr} A \omega_n^{\Phi} = \lim_{k \to \infty} \text{Tr} A \omega_n^{\Phi}, \quad A \in B(\mathcal{H}^n).
\]

The existence and uniqueness of the TDL Gibbs state was shown in [11, 22]. The following has been shown in [29] Lemma 4.2:

**Lemma V.1.** Let $\Phi$ be a translation-invariant, finite-range interaction, and $\omega := (\omega_n^{\Phi})_{n \in \mathbb{N}}$ and $\bar{\omega} := (\bar{\omega}_n^{\Phi})_{n \in \mathbb{N}}$. Then both $\omega$ and $\bar{\omega}$ satisfy the factorization property.

**Lemma V.1** and Theorem IV.13 yield the following:

**Theorem V.2.** Let $\Phi^{(1)}$ and $\Phi^{(2)}$ be translation-invariant, finite-range interactions on a finite-dimensional Hilbert space $\mathcal{H}$, and let $\beta_1, \beta_2 > 0$. Let $\rho = (\omega_n^{\Phi^{(1)}(\beta_1)})_{n \in \mathbb{N}}$ or $\rho = (\omega_n^{\Phi^{(2)}(\beta_2)})_{n \in \mathbb{N}}$, and let $\sigma = (\omega_n^{\Phi^{(2)}(\beta_2)})_{n \in \mathbb{N}}$ or $\sigma = (\omega_n^{\Phi^{(2)}(\beta_2)})_{n \in \mathbb{N}}$. Then
\[
\sigma(r||\sigma) = \sigma(r||\sigma) = H_r^s(\rho||\sigma).
\]

**B. Quasi-free states of a fermionic lattice**

In this section we consider the hypothesis testing problem for the case where the null-hypothesis is a temperature state of a non-interacting fermionic lattice system and the alternative hypothesis is a product state. For the basics on fermionic quasi-free states, see Appendix B.

Let $\omega_Q$ and $\omega_R$ be translation-invariant quasi-free states of fermions on the lattice $\mathbb{Z}^d$, with symbols $Q, R \in B(L^2(\mathbb{Z}^d))$. Then $Q$ and $R$ are translation-invariant, and hence there exist measurable functions $q, r : [0, 2\pi)^d \to \mathbb{R}$ such that $Q = F^{-1} M_q F$ and $R = F^{-1} M_r F$, where $M_q$ and $M_r$ denote the corresponding multiplication operators on $L^2([0, 2\pi)^d)$, and $F$ is the Fourier transformation (see Section 4). To avoid technical complications, we assume that there exists a $c \in (0, 1/2)$ such that $c \leq q, r \leq 1 - c$ almost everywhere with respect to the Lebesgue measure, or equivalently, $ci \leq Q, R \leq (1 - c)I$. 
The state of the fermions confined to the hypercube $C_n := \{ k : k_1, \ldots, k_\nu = 0, \ldots, n-1 \}$ is again a quasi-free state, with symbol $Q_n := p_n Q P_n$ or $R_n := p_n R P_n$, where $P_n := \sum_{k_1, \ldots, k_\nu=0}^{n-1} |1(k_1) \cdots |1(k_\nu)\rangle$, and $|1(k)\rangle_{k\in\mathbb{Z}^\nu}$ is the standard basis of $\mathcal{F}(\mathbb{Z}^\nu)$. These states have density operators on the Fock space $\mathcal{H}_n := \mathcal{F} (\text{ran } P_n)$, given by

$$\omega_{Q_n} = \det(I - Q_n) \bigoplus_{k=0}^{n^\nu} \hat{Q}_n$$

$$\omega_{R_n} = \det(I - R_n) \bigoplus_{k=0}^{n^\nu} \hat{R}_n$$

(97)

where $\hat{Q}_n := Q_n/(I - Q_n)$, $\hat{R}_n := Q_n/(I - R_n)$.

With a slight abuse of notation, we identify $\omega_Q$ with $\{\omega_{Q_n}\}_{n\in\mathbb{N}}$ and $\omega_R$ with $\{\omega_{R_n}\}_{n\in\mathbb{N}}$. We consider the hypothesis testing problem with

$$H_0 : \omega_Q \quad \text{vs.} \quad H_1 : \omega_R.$$

Note that the Hilbert space corresponding to one single mode $1_k$ is $\mathcal{F} (\text{ran } |1_k\rangle \langle 1_k|) \cong \mathbb{C}^2$, and

$$\mathcal{F} (\text{ran } P_n) \cong \bigotimes_{k\in C_n} \mathcal{F} (\text{ran } |1_k\rangle \langle 1_k|) \cong (\mathbb{C}^2)^{\otimes n^\nu}.$$

(98)

That is, $\mathcal{H}_n$ is the Hilbert space of $n^\nu$ elementary subsystems. Thus, we replace all the $1/n$ scalings in the previous sections with $1/n^\nu$. For instance, we define the strong converse exponents as

$$\overline{sc}(r|\omega_Q||\omega_R)$$

$$:= \inf \left\{ \lim_{n\to\infty} \inf \frac{1}{n^\nu} \log \text{Tr} \omega_{Q_n} T_n \right\} - r$$

$$:= \inf \left\{ \lim_{n\to\infty} \sup \frac{1}{n^\nu} \log \text{Tr} \omega_{R_n} T_n \right\} - r$$

$$= \inf \left\{ \lim_{n\to\infty} \sup \frac{1}{n^\nu} \log \text{Tr} \omega_{Q_n} T_n \right\} - r$$

$$= \inf \left\{ \lim_{n\to\infty} \sup \frac{1}{n^\nu} \log \text{Tr} \omega_{R_n} T_n \right\} - r$$

It is easy to verify that the results of Section IV-A hold true with appropriately modifying all formulas according to this scaling; e.g., the pinched Neyman-Pearson tests have to be defined as $S_n(\alpha) := \{ \omega_{Q_n} - e^{n^\nu \alpha} \omega_{R_n} > 0 \}$, etc.

We start with showing that $\overline{\psi}(\alpha|\omega_Q||\omega_R)$ exists as a limit, and it is a differentiable function of $\alpha$ for every $\alpha > 0$.

**Theorem V.3.** Let $\omega_Q$ and $\omega_R$ be quasi-free states of a fermion system on the lattice $\mathbb{Z}^\nu$, with symbols $Q = F^{-1} M_q F$, $R = F^{-1} M_r F$, and assume that there exists a constant $c \in (0, 1/2)$ such that $c \leq q, r \leq 1 - c$ almost everywhere with respect to the Lebesgue measure. Then for every $\alpha > 0$ and $(v) = \{ \}$ or $(v) = \ast$, we have

$$\overline{\psi}(\alpha|\omega_Q||\omega_R)$$

$$= \lim_{n \to +\infty} \frac{1}{n^\nu} \psi^{(v)}(\alpha|\omega_{Q_n}||\omega_{R_n})$$

$$= \left( \frac{1}{2\pi} \right)^{\nu} \int_{[0, 2\pi]^\nu} \log \left( q(x)^\alpha r(x)^{1-\alpha} \right) + (1 - q(x))^{\alpha} (1 - r(x))^{1-\alpha} \, dx$$

(99)

Moreover, $\overline{\psi}(\alpha|\omega_Q||\omega_R)$ is differentiable on $(0, +\infty)$, and

$$\frac{d}{d\alpha} |_{\alpha = 1} \overline{\psi}(\alpha|\omega_Q||\omega_R)$$

$$= \frac{1}{n^\nu} \int_{[0, 2\pi]^\nu} \left( q(x) \log \frac{q(x)}{r(x)} \right) + (1 - q(x)) \log \frac{1 - q(x)}{1 - r(x)} \, dx$$

(100)

**Proof:** The identities in (99) and (100) for $(v) = \{ \}$, and the identities in (100) have been shown in Proposition 4.1. Here we use a similar proof for (99) and (100) in the case $(v) = \ast$.

For the rest, we fix an $\alpha > 0$. Let

$$W_{n,\alpha} := \left( \frac{Q_n}{1 - Q_n} \right)^{\frac{1}{2}} \left( \frac{R_n}{1 - R_n} \right)^{\frac{1}{2}} \left( \frac{Q_n}{1 - Q_n} \right)^{\frac{1}{2}},$$

$$w_{n,\alpha} := \left( \frac{q}{1 - q} \right)^{\frac{1}{2}} \left( \frac{r}{1 - r} \right)^{\frac{1}{2}} \left( \frac{q}{1 - q} \right)^{\frac{1}{2}}.$$

Then

$$\frac{1}{n^\nu} \psi^{(v)}(\alpha|\omega_{Q_n}||\omega_{R_n})$$

$$= \frac{1}{n^\nu} \log \text{Tr} \left( \omega_{Q_n}^{\frac{1}{2}} \omega_{R_n}^{\frac{1}{2}} \omega_{Q_n}^{\frac{1}{2}} \right)^\alpha$$

$$= \frac{1}{n^\nu} \log \text{Tr} \left( \left[ \det(I - Q_n) \mathcal{F}(\hat{Q}_n) \right]^{\frac{1}{2}} \left[ \det(I - R_n) \mathcal{F}(\hat{R}_n) \right]^{\frac{1}{2}} \right)^\alpha$$

$$= \frac{1}{n^\nu} \log \text{det}(I - Q_n)^\alpha + \frac{1}{n^\nu} \log \text{det}(I - R_n)^{1-\alpha}$$

$$+ \frac{1}{n^\nu} \log \text{Tr} \mathcal{F}(W_{n,\alpha})$$

$$= \frac{1}{n^\nu} \log \text{Tr} (I - Q_n)^\alpha + \frac{1}{n^\nu} \log \text{Tr} (I - R_n)^{1-\alpha}$$

$$+ \frac{1}{n^\nu} \log \text{det}(I + W_{n,\alpha})$$

$$= \frac{1}{n^\nu} \log \text{Tr} (I - Q_n)^\alpha + \frac{1}{n^\nu} \log \text{Tr} (I - R_n)^{1-\alpha}$$

$$+ \frac{1}{n^\nu} \log (I + W_{n,\alpha})$$

(103)
By lemma \textbf{C.1} we have
\begin{equation}
\lim_{n \to +\infty} \frac{1}{n^\nu} \mathrm{Tr}\log(I - Q_n)^\alpha + \frac{1}{n^\nu} \mathrm{Tr}\log(I - R_n)^{1-\alpha} = \frac{1}{(2\pi)^\nu} \int_{[0,2\pi)^\nu} \log \left[ \left(1 - q(x)\right)^\alpha (1 - r(x))^{1-\alpha} \right] \, dx
\end{equation}
(104)
To evaluate the limit of the last term in (103), we use Corollary \textbf{C.2} with \( a(1) = q, \, a(2) = r, \, f^{(1)}(t) = (t/(1-t))^\frac{1}{\nu}, \, f^{(2)}(t) = (t/(1-t))^\frac{1}{\nu} \) and \( g(x) = \log(1 + x^\alpha) \), and obtain
\begin{equation}
\lim_{n \to +\infty} \frac{1}{n^\nu} \mathrm{Tr}\log(I + W_n^\alpha) = \frac{1}{(2\pi)^\nu} \int_{[0,2\pi)^\nu} \log \left( 1 + \left( \hat{q}(x) \hat{f}(x) \frac{1}{\nu} \hat{q}(x) \right)^\alpha \right),
\end{equation}
(105)
where \( \hat{q} := q/(1-q), \hat{f} := r/(1-r) \). Combining (103), (104), and (105), we get (99)–(100). Differentiability of \( \psi(\alpha | \omega_Q || \omega_R) \) is straightforward to verify.

In particular, Theorem \textbf{V.3} shows that \( D_\alpha \) and \( D_\alpha^* \) give rise to the same asymptotic quantities:

**Corollary V.4.** In the setting of Theorem \textbf{V.3} we have
\begin{equation}
\mathcal{D}_\alpha(r) \omega_Q \omega_R)
\end{equation}
(106)
for every \( \alpha \in (0, +\infty) \setminus \{1\} \) and \( \{v\} \) or \( \{v\} = * \).

Now we can obtain the strong converse exponent for quasi-free states.

**Theorem V.5.** Let \( \omega_Q \) and \( \omega_R \) be quasi-free states of a fermion system on the lattice \( \mathbb{Z}^\nu \), and assume that \( cI \leq Q, R \leq (1-c)I \) for some \( c \in (0, 1/2) \). Then
\begin{equation}
\frac{sc(r | \omega_Q \omega_R)}{H^\alpha(r | \omega_Q \omega_R)} = \sup_{\alpha \in [0,1]} - \frac{1}{\alpha - 1} \left[ r - \mathcal{D}_\alpha(\omega_Q \omega_R) \right],
\end{equation}
(108)
where \( \mathcal{D}_\alpha(\omega_Q \omega_R) \) is given in (106).

**Proof:** By Theorem \textbf{V.3} \( \psi(\alpha | \omega_Q \omega_R) \) exists as a limit for \( \alpha > 1 \), and the limit is differentiable. By (97) and the assumption that \( cI \leq R \leq (1-c)I \),
\begin{align*}
\omega_{R_n} & \geq \det(cI_{n^\nu}) \sum_{k=0}^{n^\nu} \left( \frac{c}{1-c} \right)^k I_{\Lambda^k H_n} \\
& = c^{\nu^\nu} \sum_{k=0}^{n^\nu} \left( \frac{c}{1-c} \right)^k I_{\Lambda^k H_n} \\
& \geq \left( c \min \left\{ 1, \frac{c}{1-c} \right\} \right)^\nu I_{F(\mathrm{ran} P_n)},
\end{align*}
Hence, by Corollary \textbf{III.12} we have \( \overline{\psi}(\alpha | \omega_Q \omega_R) = \psi(\alpha | \omega_Q \omega_R) \) for the states \( \sigma_n = \hat{\omega}_{R_n} \) in Example \textbf{III.10} Combining these two facts, the assertion follows from Theorem \textbf{IV.5}.

**APPENDIX A**

**CLASSICAL LARGE DEVIATIONS**

Let \( \mu_n, n \in \mathbb{N} \), be a sequence of finite positive measures on \( \mathbb{R} \), and let \( c_n, n \in \mathbb{N} \), be a sequence of positive numbers such that \( \lim_n c_n = +\infty \). For each \( n \), define the logarithmic moment generating function \( \Lambda_n \) by
\begin{equation}
\Lambda_n(t) := \log \int_{\mathbb{R}} e^{tx} \, d\mu_n(x).
\end{equation}
(109)
Here we use the convention \( \log +\infty := +\infty \). Define
\begin{equation}
\overline{\Lambda}(t) := \limsup_{n \to +\infty} \frac{1}{c_n} \Lambda_n(c_n t), \quad t \in \mathbb{R}.
\end{equation}
(110)
Hölder’s inequality yields that \( \Lambda_n \) is convex for every \( n \in \mathbb{N} \), and hence \( \overline{\Lambda} \) is convex as well.

The following lemma is a standard generalization of the Markov inequality:

**Lemma A.1.** For every \( x \in \mathbb{R} \),
\begin{align*}
\limsup_{n \to +\infty} \frac{1}{c_n} \log \mu_n ([x, +\infty)) & \leq \sup_{t \geq 0} \left\{ tx - \overline{\Lambda}(t) \right\}, \quad (110) \\
\limsup_{n \to +\infty} \frac{1}{c_n} \log \mu_n ((-\infty, x]) & \leq \sup_{t \leq 0} \left\{ tx - \overline{\Lambda}(t) \right\}. \quad (111)
\end{align*}
**Proof:** For every \( t \geq 0 \),
\begin{align*}
\mu_n ([x, +\infty)) &= \int_x^{+\infty} 1_{[x, +\infty)}(z) \, d\mu_n(z) \\
& \leq \int_x^{+\infty} e^{c_n (z-x)} \, d\mu_n(z) \\
& \leq e^{-c_n t x} \int_{\mathbb{R}} e^{c_n t z} \, d\mu_n(z),
\end{align*}
and hence,
\begin{align*}
\limsup_{n \to +\infty} \frac{1}{c_n} \log \mu_n ([x, +\infty)) & \leq -tx + \limsup_{n \to +\infty} \frac{1}{c_n} \log \int_{\mathbb{R}} e^{c_n t z} \, d\mu_n(z) \\
& = -tx + \overline{\Lambda}(t),
\end{align*}
from which (110) follows. The proof of (111) goes the same way.

The following converse to Lemma \textbf{A.1} was essentially given in (11), under the stronger (for our purposes too strong) condition that \( \overline{\Lambda} \) exists as a limit in a neighbourhood of 0, where it is also differentiable. The more general version below can be easily obtained by following the same line of argument as in (11). Using the same approach, a generalization of the lower bound in (11) has been obtained in (9), that also generalizes Lemma \textbf{A.2} below. For readers’ convenience, we include a detailed proof below, based on the proof of the Gärtner-Ellis theorem in (13, pp. 49–50).
Lemma A.2. Assume that $\overline{X}(t) = \lim_n \frac{1}{n} \Lambda_n(c_n t)$ in some interval $(\alpha, \beta)$, and, moreover, that $\overline{X}$ is a finite-valued differentiable function on $(\alpha, \beta)$. Then, for every $x \in J := \left( \lim_{t \searrow \alpha} \overline{X}(t), \lim_{t \nearrow \beta} \overline{X}(t) \right)$, there exists a $t_x \in (\alpha, \beta)$ such that $\overline{X}'(t_x) = x$, and

$$\overline{X}'(x) := \sup_{t \in \mathbb{R}} \{x - \overline{X}(t)\} = \sup_{t \in \mathbb{I}} \{x - \overline{X}(t)\}$$

$$= x t_x - \overline{X}(t_x) > x - \overline{X}(t), \quad (112)$$

where $\mathbb{I} \subset \mathbb{R}$ is any interval such that $t_x \in \mathbb{I}$, and the last inequality holds for every $t \in \mathbb{R}$ such that $t \neq (\alpha, \beta)$ or $\overline{X}(t) \neq x$. Moreover, for every $x_0, x_1 \in \mathbb{R} \cup \{\pm \infty\}$ such that $x_0 < x < x_1$, we have

$$\liminf_{n} \frac{1}{c_n} \log \mu_n((x_0, x_1)) \geq -\overline{X}'(x), \quad (113)$$

$$\liminf_{n} \frac{1}{c_n} \log \mu_n((x_0, x)) \geq -\overline{X}'(x). \quad (114)$$

Proof: Since $\overline{X}$ is convex, differentiability on $(\alpha, \beta)$ implies that $\overline{X}$ is monotone and continuous on $(\alpha, \beta)$, and hence for every $x \in J$ there exists a $t_x \in (\alpha, \beta)$ such that $\overline{X}(t_x) = x$. The rest of the assertions in (113) are immediate from the concavity of $t \mapsto x - \overline{X}(t)$ on $\mathbb{R}$. Hence, we are left to prove (113) and (114), of which we only prove (113), as the proof of (114) goes exactly the same way.

Let $x \in J$ and $x_1 \in \mathbb{R} \cup \{\pm \infty\}$ be such that $x < x_1$. For every $\delta > 0$ such that $x + \delta < x_1$, $(x, x + \delta) \subset J$, choose a $y \in (x, x + \delta)$. By the above, there is a $t_y \in (\alpha, \beta)$ corresponding to $y$ such that $\overline{X}(t_y) = y$, and hence, $\overline{X}(y) = y t_y - \overline{X}(t_y)$. Since $\overline{X}(t_y) < +\infty$, we have $\Lambda_n(c_n t_y) < +\infty$ for all large enough $n$, and hence we can define the probability measures $\mu_n, \mu_n, \mu_n, \mu_n, \mu_n$ by

$$\mu_n(x, y) := \frac{1}{c_n} \int_{\mathbb{R}} e^{c_n t_y s} d\mu_n(s)$$

$$= \int_{\mathbb{R}} e^{c_n t_y s - \Lambda_n(c_n t_y)} d\mu_n(s),$$

where $B \subset \mathbb{R}$ is any Borel set. Note that

$$\mu_n((x, x + \delta)) = \int_{(x,x+\delta)} e^{c_n t_y s - \Lambda_n(c_n t_y)} d\mu_n(s)$$

$$= e^{c_n t_y x - \Lambda_n(c_n t_y)} \int_{(x,x+\delta)} e^{c_n t_y s - \Lambda_n(c_n t_y)} d\mu_n(s)$$

$$\leq e^{c_n t_y x - \Lambda_n(c_n t_y)} e^{c_n t_y \delta} \int_{(x,x+\delta)} d\mu_n(s)$$

$$= e^{c_n t_y x - \Lambda_n(c_n t_y)} e^{c_n t_y \delta} \mu_n((x, x + \delta)),$$

and therefore,

$$\liminf_{n} \frac{1}{c_n} \log \mu_n((x, x_1)) \geq \liminf_{n} \frac{1}{c_n} \log \mu_n((x, x)) \geq \overline{X}(t_y) - t_y x - |t_y| \delta + \liminf_{n} \frac{1}{c_n} \log \mu_n, \mu_n((x, x + \delta)),$$

where we used that $\overline{X}(t_y)$ exists as a limit. If we can prove that

$$0 > \max \left\{ \limsup_{n \to +\infty} \frac{1}{c_n} \log \mu_n, \mu_n((-\infty, x)) \right\},$$

$$\limsup_{n \to +\infty} \frac{1}{c_n} \log \mu_n, \mu_n([x + \delta, +\infty)) \right\} \quad (116)$$

then we have $\lim_{n \to +\infty} \mu_n, \mu_n((x, x + \delta)) = 1$, and hence, by (115),

$$\liminf_{n \to +\infty} \frac{1}{c_n} \log \mu_n((x, x_1)) \geq \overline{X}(t_y) - t_y x - |t_y| \delta$$

$$> -\overline{X}'(x) - |t_y| \delta,$$

where the second inequality is due to (112). Using that $|t_y| \delta \to 0$ as $\delta \to 0$, (113) follows. Hence, we are left to prove (116).

Let $\Lambda_n, \mu_n$ denote the logarithmic moment generating function of $\mu_n, \mu_n, \mu_n, \mu_n, \mu_n, \mu_n, \mu_n$, i.e., for every $t \in \mathbb{R}$,

$$\Lambda_n, \mu_n : \mathbb{R} \mapsto \log \int_{\mathbb{R}} e^{c_n t s} d\mu_n(s)$$

$$= \log \int_{\mathbb{R}} e^{c_n t s + c_n t y s - \Lambda_n(c_n t y)} d\mu_n(s)$$

$$= \Lambda_n(c_n t + c_n t y) - \Lambda_n(c_n t y),$$

and let $\overline{\Lambda}_y(t) := \limsup_{n \to +\infty} \frac{1}{c_n} \Lambda_n, \mu_n(c_n t)$. By assumption,

$$\overline{\Lambda}_y(t) = \overline{X}(t + y) - \overline{X}(t) \quad (117)$$

for all $t$ such that $t + y \in (\alpha, \beta)$. By lemma A.1,

$$\limsup_{n \to +\infty} \frac{1}{c_n} \log \mu_n, \mu_n((x + \delta, +\infty)) \leq -\limsup_{t \geq 0} \{t(x + \delta) - \overline{\Lambda}_y(t)\},$$

$$\limsup_{n \to +\infty} \frac{1}{c_n} \log \mu_n, \mu_n((-\infty, x)) \leq -\limsup_{t \leq 0} \{t(x - \overline{\Lambda}_y(t))\},$$

and hence (116) will be proved if we can show that

$$0 < \sup_{t \geq 0} \{t(x + \delta) - \overline{\Lambda}_y(t)\} \quad \text{and} \quad 0 < \sup_{t \leq 0} \{t(x - \overline{\Lambda}_y(t))\}. \quad (118)$$

By (117),

$$\sup_{t \geq 0} \{t(x + \delta) - \overline{\Lambda}_y(t)\}$$

$$= \sup_{t \geq 0} \{t(x + \delta) - \overline{X}(t + y) + \overline{X}(t)\}$$

$$= \overline{X}(t_y) - (x + \delta) t_y + \sup_{t \geq 0} \{t + t y)(x + \delta) - \overline{X}(t + y)\}$$

$$= \overline{X}(t_y) - (x + \delta) t_y + \sup_{t \geq t_y} \{t(x + \delta) - \overline{X}(t)\}$$

$$= \overline{X}((x + \delta)) - \{x + \delta\} t_y - \overline{X}(t_y)$$

$$> 0,$$

where the last identity and the inequality follows from (112), since $\overline{X}(t_y) = y < x + \delta$. The other half of (118) follows by the same kind of argument, which we omit.
Thus we get that in this case the eigenvalues of the bounded operators on and interpretation is that anticommutation relations (CAR) is defined as the unique bounded linear extension or fermionic Fock space \( F \), where \( \dim F \) may be countably infinite. Note that for any \( x_1, \ldots, x_k \in H \), their anti-symmetrized vector \( \langle x_1 \wedge \ldots \wedge x_k \rangle = \det(I - Q) \oplus k=0 \frac{Q}{I - Q} = \det(I - Q)F \left( \frac{Q}{I - Q} \right) \), according to [14] Lemma 3).

The dynamics of a system of non-interacting fermions is determined by a single-particle Hamiltonian operator, i.e., a self-adjoint operator \( H \) on \( H \). Assume for the rest that \( H \) is finite-dimensional, and let \( H_{k,i} := \otimes^k H \otimes \otimes^{k-1} F \) be the embedding of \( H \) into the i-th position in \( \otimes^k F \). It is easy to see that \( \sum_{k=1} H_{k,i} \) leaves \( \wedge^k H \) invariant, and we denote the restriction of \( H_{k,i} \) onto \( \wedge^k H \) by \( \Gamma_k(H) \), and define the second-quantized Hamiltonian \( \Gamma(H) := \oplus_{k=1} \Gamma_k(H) \). If the initial state of the system is a pure state given by the vector \( x_1 \wedge \ldots \wedge x_r \), then the state after time \( t \) is the pure state given by the vector \( e^{-itH} x_1 \wedge \ldots \wedge e^{-itH} x_r = e^{-it\Gamma(H)}(x_1 \wedge \ldots \wedge x_r) \). Thus, the dynamics of the many-particle system is governed by the Hamiltonian \( \Gamma(H) \). Hence, the equilibrium state at inverse temperature \( \beta \) (Gibbs state) is \( e^{-\beta \Gamma(H)} / \text{Tr} e^{-\beta \Gamma(H)} \), and a direct computation shows that this is a quasi-free state with

\[
Q = \frac{e^{-\beta H}}{I + e^{-\beta H}}.
\]

Vice versa, any quasi-free state with symbol \( Q \) such that \( 0 < Q < 1 \) is the Gibbs state of non-interacting fermions at some inverse temperature \( \beta \). The infinite-dimensional case is slightly more complicated: it is still true that the equilibrium state of non-interacting fermions at inverse temperature \( \beta \) with one-particle Hamiltonian \( H \) is the quasi-free state with symbol \( e^{-\beta H} / (I + e^{-\beta H}) \), but the equilibrium state in this case is defined through the KMS condition [3] Section 5.2.4]

To describe a system of non-interacting fermions occupying sites of a \( \nu \)-dimensional cubic lattice, we choose \( H := \ell^2(\mathbb{Z}^\nu) \). As we have seen, the equilibrium state of the system at any inverse temperature is a quasi-free state, with some symbol \( Q \in \mathcal{B}(\ell^2(\mathbb{Z}^\nu)) \). Let \( \{1_{(k)} : k \in \mathbb{Z}^\nu \} \) denote the standard basis of \( \ell^2(\mathbb{Z}^\nu) \), and for every \( n \in \mathbb{N} \), let \( P_n := \sum_{k_1, \ldots, k_n = 0} 1_{(k)} 1_{(k)} \). The Hilbert space of the fermions occupying sites of the hypercube \( C_n := \{ k : k_1, \ldots, k_n = 0, \ldots, n - 1 \} \) is then \( F(P_n, H) \), and the state of this subsystem is quasi-free with symbol \( Q_n := P_n Q P_n \).

The translation operators are the unique linear extensions of \( S_j 1_{(k)} \mapsto 1_{(k+j)} \), \( k, j \in \mathbb{Z}^\nu \). The map \( \gamma_j(e(x)) := e(S_j x) \) extends to an automorphism of CAR(\( \ell^2(\mathbb{Z}^\nu) \)) for all \( j \in \mathbb{Z}^\nu \), and \( \gamma_j \), \( j \in \mathbb{Z}^\nu \), is a group of automorphisms, called the group of translation automorphisms. A quasi-free state \( \omega_Q \) is called translation-invariant if \( \omega_Q \circ \gamma_j = \omega_Q \), \( j \in \mathbb{Z}^\nu \), which holds

\[
\omega_Q(I) = 1, \text{ and for any } x_1, \ldots, x_n, y_1, \ldots, y_m, \\
\omega_Q(c(x_1) \ldots c(x_n) c(y_m) \ldots c(y_1)) = \delta_{m,n} \det\{ \langle y_i, Q x_j \rangle \}_{i,j=1}^n.
\]
if and only if its symbol $Q$ is translation-invariant, i.e., it
commutes with all the unitaries $S_j$, $j \in \mathbb{Z}^\nu$.

\section*{Appendix C

Generalizations of Szegő's Theorem}

Translation-invariant operators on $l^2(\mathbb{Z}^\nu)$ commute with each other, and they are simultaneously diagonalized by the Fourier transformation

$$F : l^2(\mathbb{Z}^\nu) \rightarrow L^2([0,2\pi)^\nu)$$

$$F \varphi_k := \varphi_k, \quad \varphi_k(z) := e^{i(k \cdot z)}, \quad k \in [0,2\pi)^\nu, \ k \in \mathbb{Z}^\nu,$$

where $\langle k, z \rangle := \sum_{i=1}^\nu k_i x_i$. That is, every translation-invariant operator $A$ arises in the form $A = F^{-1} M_A F$, where $M_A$ denotes the multiplication operator by a bounded measurable function $a$ on $[0,2\pi)^\nu$. In the rest of this section we use the convention that lower and upper case versions of the same letter refer to a measurable function on $[0,2\pi)^\nu$ and the Fourier transform of the corresponding multiplication operator, respectively. Moreover, we use the notation

$$A_n := P_n A P_n, \quad A \in \mathcal{B}(l^2(\mathbb{Z}^\nu)),$$ where

$$P_n := \sum_{k_1, \ldots, k_n=0}^{\nu} |1(k)| \langle 1(k) \rangle.$$ Let $\Sigma(A)$ denote the convex hull of the spectrum of a self-adjoint operator $A$. The following generalization of Szegő's theorem \cite{[17]} has been shown in \cite{[34]} lemma 3.1:

\begin{lemma}
For all $k = 1, \ldots, r$, let $a^{(k)} : [0,2\pi)^\nu \rightarrow \mathbb{R}$ be a bounded measurable function, and $f^{(k)} : \Sigma(A^{(k)}) \rightarrow \mathbb{R}$ be continuous. Then

$$\lim_{n \rightarrow \infty} \frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(A^{(k_1)}) \cdots f^{(k_n)}(A^{(k_n)}) \right| \leq \frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(a^{(k_1)}) \cdots f^{(k_n)}(a^{(k_n)}) \right| \leq \frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(a^{(k_1)}) \cdots f^{(k_n)}(a^{(k_n)}) \right|$$

\end{lemma}

From this lemma, we can easily get the following:

\begin{corollary}
In the setting of \textit{lemma C.1}, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(A^{(k_1)}) \cdots f^{(k_n)}(A^{(k_n)}) \right| \leq \frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(a^{(k_1)}) \cdots f^{(k_n)}(a^{(k_n)}) \right|$$

for any continuous function $g : [0,\Delta] \rightarrow \mathbb{R}$, where $\Delta := \prod_{k=1}^{\nu} \max_{a \in \Sigma(A^{(k)})} |f^{(k)}(a)|^2$.

\end{corollary}

\begin{proof}
Let $B_n := \prod_{k=1}^{n} f^{(k)}(A^{(k)})$ and $b := \prod_{k=1}^{n} f^{(k)}(a^{(k)})$. Since $g$ is continuous on $[0,\Delta]$, the Stone-Weierstrass theorem tells that for every $\varepsilon > 0$, there exists a polynomial $g_\varepsilon$ such that

$$\|g - g_\varepsilon\|_\infty := \max_{x \in [0,\Delta]} |g(x) - g_\varepsilon(x)| < \varepsilon.$$

Now, for a fixed $\varepsilon > 0$, \textit{lemma C.1} yields that there exists an $N_\varepsilon$ such that for all $n \geq N_\varepsilon$,

$$\frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(A^{(k_1)}) \cdots f^{(k_n)}(A^{(k_n)}) \right| < \varepsilon.$$

Hence, for every $n \geq N_\varepsilon$,

$$\frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(A^{(k_1)}) \cdots f^{(k_n)}(A^{(k_n)}) \right| \leq \frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(a^{(k_1)}) \cdots f^{(k_n)}(a^{(k_n)}) \right| \leq 2 \frac{1}{n^\nu} \sum_{k_1, \ldots, k_n=0} \left| f^{(k_1)}(a^{(k_1)}) \cdots f^{(k_n)}(a^{(k_n)}) \right|$$

\end{proof}

\section*{Appendix D

Classical i.i.d. and Markov chains}

Let $\mathcal{X}$ be a finite set, and $\rho$ and $\sigma$ be probability measures on the sigma-field generated by the cylinder sets of $\mathcal{X}^\infty := \times_{k=1}^\infty \mathcal{X}$. We denote by $\rho_n$ and $\sigma_n$ the restrictions of $\rho$ and $\sigma$, respectively, to $\mathcal{X}^n := \times_{k=1}^n \mathcal{X}$. Then $\rho_n$ and $\sigma_n$ can be identified with their respective probability mass functions, which we also denote by $\rho_n$ and $\sigma_n$. Let $\mathcal{H}_n := l^2(\mathcal{X}^n) = l^2(\mathcal{X})^\otimes n$, where for any finite set $\mathcal{Y}$, $l^2(\mathcal{Y}) = \mathcal{Y}^\otimes$ equipped with the inner product $\langle f, g \rangle := \sum_{y \in \mathcal{Y}} \overline{f(y)} g(y)$. The multiplication operators by $\rho_n$ and $\sigma_n$ on $l^2(\mathcal{X}^n)$ are density operators, which we also denote by $\rho_n$ and $\sigma_n$ if no confusion arises. Moreover, we identify the probability measures $\rho$ and $\sigma$ with $
abla \rho_n \nabla n \in \mathbb{N}$ and $\nabla \sigma_n n \in \mathbb{N}$.

Our aim here is to show how the expressions in \cite{[43], [44]} for the strong converse exponent in the classical i.i.d. and Markov case can be recovered from our Theorems \textit{IV.5} and \textit{IV.7}. Note that in the classical case any choice of the auxiliary sequence $\sigma$ yields $\rho_n = \rho_n$, $n \in \mathbb{N}$, and thus $\psi = \psi$. Hence, in the classical case it is sufficient to verify the differentiability of $\psi$ to apply Theorem \textit{IV.5}.

First, we consider the i.i.d. case, where $\rho_n = \sigma_n^\otimes n$, $\sigma_n = \sigma_n^\otimes n$, $n \in \mathbb{N}$. We assume that $\rho_1 \subseteq \sigma_1$. Then $\psi(t) := \psi(t) |\rho| |\sigma| = \psi(t) |\rho_1| |\sigma_1| = \log Z(t)$, where $Z(t) := \sum_{x \in \mathcal{X}} \rho_1(x)^t |\sigma_1(x)|^{-t}$. Define

$$\omega_1(t)(x) := \rho_1(x)^t |\sigma_1(x)|^{-t}/Z(t), \quad x \in \mathcal{X}, \ t \in \mathbb{R},$$

and let

$$\omega_1(\infty)(x) := \lim_{t \rightarrow \infty} \omega_1(t)(x) = \begin{cases} 0, & x \notin \mathcal{X}^*, \\ \sigma_1(x)/|\sigma_1(x)|, & x \in \mathcal{X}^*, \end{cases}$$

where $x \in \mathcal{X}^*$ if $\rho_1(y)/\sigma_1(y) \leq \rho_1(x)/\sigma_1(x)$ for all $y \in \mathcal{X}$, i.e., if $\log(\rho_1(x)/\sigma_1(x)) = D_{\max}(\rho_1|\sigma_1)$. Then $\omega_1(t)$ is a
probability mass function on $\mathcal{X}$ for every $t \in \mathbb{R} \cup \{+\infty\}$, and a straightforward computation shows

\begin{align}
D(\omega_1(t)\|\sigma_1) &= -\psi(t) + t\psi'(t), \quad (120) \\
D(\omega_1(\infty)\|\sigma_1) &= -\log \sigma(X^*), \quad (121) \\
D(\omega_1(t)\|\rho_1) &= -\psi(t) + (t-1)\psi'(t), \quad (122) \\
D(\omega_1(\infty)\|\rho_1) &= -\log \sigma(X^*) - D_{\max}(\rho_1\|\sigma_1), \quad (123) \\
\tau(t) := D(\omega_1(t)\|\sigma_1) - D(\omega_1(t)\|\rho_1) &= \psi'(t), \quad (124) \\
\tau(\infty) := \lim_{t \to +\infty} \tau(t) &= D_{\max}(\rho_1\|\sigma_1). \quad (125)
\end{align}

Using now the notations and results of Section III-C with $f := \psi$, we get

\begin{align}
D_{\psi,1} &= D(\rho_1\|\sigma_1) = D(\omega_1(t)\|\sigma_1), \\
D_{\psi,\infty} &= D_{\max}(\rho_1\|\sigma_1), \\
r_{\text{max}} &= \sup_{1 < t < +\infty} (-\psi(t) + t\psi'(t)) = \sup_{1 < t < +\infty} D(\omega_1(t)\|\sigma_1) \\
 &= D(\omega_1(\infty)\|\sigma_1),
\end{align}

where in the last line we used (57) and the fact that $t \mapsto -\psi(t) + t\psi'(t) = F(x(t))$ is monotone increasing due to the convexity of $F$. Using now (55)–(56) and (120)–(123), we see that for every $r \in (D(\omega_1(1)\|\sigma_1), D(\omega_1(\infty)\|\sigma_1))$, there exists a $t_r \in (1, +\infty)$ such that

\begin{align}
r &= -\psi(t_r) + t\psi'(t_r) = D(\omega_1(t_r)\|\sigma), \\
H_r^*(\rho_1\|\sigma_1) &= -\psi(t_r) + (t_r - 1)\psi'(t_r) = D(\omega_1(t_r)\|\rho), \quad (126)
\end{align}

and for $r \geq D(\omega_1(\infty)\|\sigma_1)$,

\begin{align}
H_r^*(\rho_1\|\sigma_1) &= r - D_{\max}(\rho_1\|\sigma_1) \\
 &= r - D(\omega_1(\infty)\|\sigma_1) + D(\omega_1(\infty)\|\rho_1), \quad (127)
\end{align}

due to (42), (59) and (120)–(123). Combining now Theorem [IV.7] with (120)–(127), we get Theorems 2 and 3 in [43] with $p_0 = \sigma_1$, $p_1 = \rho_1$.

Next, we consider the case where $\rho$ and $\sigma$ are Markov chains, with transition matrices $R$ and $S$, respectively. That is, for any $x_1, \ldots, x_n \in \mathcal{X}$, we have

\begin{align}
\rho_n(x_1, \ldots, x_n) &= \rho_1(x_1)R_{x_1x_2}R_{x_2x_3}\cdots R_{x_{n-1}x_n}, \\
\sigma_n(x_1, \ldots, x_n) &= \sigma_1(x_1)S_{x_1x_2}S_{x_2x_3}\cdots S_{x_{n-1}x_n}.
\end{align}

We assume that $R$ is irreducible, i.e., there exists an $n \in \mathbb{N}$ such that all the entries of $(I + R)^n$ are strictly positive. We also assume that $\text{supp} \; R \subseteq \text{supp} \; S$, i.e., $R_{xy} > 0 \Rightarrow S_{xy} > 0$ for all $x, y \in \mathcal{X}$; then $S$ is also irreducible. By the Perron-Frobenius theory (see, e.g., [13] Theorem 3.1.1), for any irreducible matrix with non-negative entries, the spectral radius of the matrix is an eigenvalue, and the corresponding left and right eigenvectors can be chosen to have strictly positive entries. Here we don’t assume that the Markov chains are stationary, only that their initial distributions are strictly positive, i.e., $\rho_1(x) > 0$, $\sigma_1(x) > 0$ for all $x \in \mathcal{X}$. For every $t \in \mathbb{R}$, let $T_{xy}(t) := R_{xy}^{\omega_1(t)x_{y-1}}$. Then $T(t)$ is irreducible for every $t \in \mathbb{R}$. Let $\lambda(t)$ be the spectral radius of $T(t)$, and let $v(t)$ be a corresponding right eigenvector of $T(t)$ with strictly positive entries. Noting that $v_x(t)/M(t) \leq 1 \leq v_x(t)/m(t)$ for every $x \in \mathcal{X}$, where $m(t) := \min_x v_x(t)$, $M(t) := \max_x v_x(t)$, we get

\begin{align}
\overline{\psi}(t) := \overline{\psi}(t\|\sigma) &= \lim_{n \to +\infty} \frac{1}{n} \log \sum_{x_1, \ldots, x_n \in \mathcal{X}} \rho_1(x_1)\sigma_1(x_1)^{1-t} \\
 &= \lim_{n \to +\infty} \frac{1}{n} \log(u(t), T(t)^{n-1}1) = \log \lambda(t),
\end{align}

where $u_x(t) := \rho_1(x)^t\sigma_1(x)^{1-t}$, $x \in \mathcal{X}$, and $1$ stands for the constant one vector (see also [13] Theorem 3.1.1). By standard results in perturbation theory [53], $\lambda(t)$ is an analytic function of $t$. In particular, $\overline{\psi}(t\|\sigma)$ is a differentiable function of $t$, and Theorem [IV.5] yields that (70) holds. Our aim now is derive alternative expressions for $H_r^*(\rho\|\sigma)$, from which we can recover the results of [44].

Following [45], [54], we define

\begin{align}
Q_{xy}(t) := \frac{T_{xy}(t)v_y(t)}{\lambda(t)v_x(t)}.
\end{align}

Then $Q(t)$ is irreducible with stationary distribution $q(t)$, and we denote the generated irreducible Markov chain by $\omega(t)$. The following explicit expressions for the asymptotic relative entropies are easy to verify:

\begin{align}
\mathbb{D}(\omega(t)\|\sigma) &= \lim_{n \to +\infty} \frac{1}{n} D(\omega_n(t)\|\sigma_n) \\
 &= \sum_{x,y} q_x(t)Q_{xy}(t) \log \frac{Q_{xy}(t)}{S_{xy}}, \\
\mathbb{D}(\omega(t)\|\rho) &= \lim_{n \to +\infty} \frac{1}{n} D(\omega_n(t)\|\rho_n) \\
 &= \sum_{x,y} q_x(t)Q_{xy}(t) \log \frac{Q_{xy}(t)}{R_{xy}}.
\end{align}

Now we follow a modification of the proof of [54] Lemma 2.3] to connect the above formulas to $\overline{\psi}$. Normalizing $v(t)$ such that $\sum_x q_x(t)v_x(t) = 1$, we have, for every $n \in \mathbb{N}$,

\begin{align}
\overline{\psi}(t\|\sigma) &= \log \lambda(t) \\
 &= \frac{1}{n} \log \sum_{x_0, x_1, \ldots, x_n} q_{x_0}(t) \frac{R_1^{x_{0}x_{1}}} {v_x(t)} \cdots \frac{R_{x_n}^{1-x_{n-1}}}{v_{x_n}(t)}.
\end{align}
A straightforward calculation gives
\[
\overline{\psi}(t) = \frac{1}{t} \log \left( \lambda(t) \right)
+ \frac{1}{t} \frac{1}{n} \sum_{x_0, x_1, \ldots, x_n} q_{x_0}(t) Q_{x_0 x_1} \cdots Q_{x_{n-1} x_n} \log \frac{Q_{x_0 x_1} \cdots Q_{x_{n-1} x_n}}{S_{x_0 x_1} \cdots S_{x_{n-1} x_n}}
\]
\[
+ \frac{1}{t} \frac{1}{n} \sum_{x_0, x_1, \ldots, x_n} q_{x_0}(t) Q_{x_0 x_1} \cdots Q_{x_{n-1} x_n} \left[ v_{x_0}(t) - v_{x_n}(t) \right]
\]
\[
+ \frac{1}{t} \frac{1}{n} \sum_{x_0, x_1, \ldots, x_n} q_{x_0}(t) Q_{x_0 x_1} \cdots Q_{x_{n-1} x_n} \frac{d}{dt} \frac{q_{x_0}(t) v_{x_n}(t)}{v_{x_n}(t)}.
\]
(128)
(129)
(130)
(131)

The term in (130) is equal to zero, and taking the limit \( n \to +\infty \) yields
\[
\overline{\psi}(t) = \frac{1}{t} \log \left( \lambda(t) \right) + \frac{1}{t} \overline{D}(\omega(t) | \sigma).
\]
(132)

It is easy to see that (128)–(131) holds also if (128)–(129) is replaced with
\[
\frac{1}{t - 1} \log \left( \lambda(t) \right) + \frac{1}{t - 1} \frac{1}{n} \sum_{x_0, x_1, \ldots, x_n} q_{x_0}(t) Q_{x_0 x_1} \cdots Q_{x_{n-1} x_n} \log \frac{Q_{x_0 x_1} \cdots Q_{x_{n-1} x_n}}{R_{x_0 x_1} \cdots R_{x_{n-1} x_n}}
\]
and as above, we obtain
\[
\overline{\psi}(t) = \frac{1}{t - 1} \log \left( \lambda(t) \right) + \frac{1}{t - 1} \overline{D}(\omega(t) | \rho).
\]
(133)

Hence,
\[
\overline{D}(\omega(t) | \sigma) = -\overline{\psi}(t) + t \overline{\psi}(t),
\]
(134)
\[
\overline{D}(\omega(t) | \rho) = -\overline{\psi}(t) + (t - 1) \overline{\psi}(t),
\]
(135)
in complete analogy with (120)–(122). Using again the general considerations in Section III-C with \( f = \psi \), and Theorem IV.3 we recover Theorems 2 and 3 from [44].

We refer to [29, 44, 55] for more details on exponentially decaying tail probabilities, hypothesis testing, and the information geometry of classical Markov chains.

ACKNOWLEDGMENTS

MM would like to thank Prof. Fumio Hiai for discussions on the eigenvalues of Gibbs states, and Vincent F. Tan and Marco Tomamichel for pointing out the paper [9]. This work was partially supported by the MEXT Grant-in-Aid (A) No. 20686026 “Project on Multi-user Quantum Network” (TO), and by the European Research Council Advanced Grant “IRQUAT”, the Spanish MINECO Project No. FIS2013-40627-P, the Generalitat de Catalunya CIRIT Project No. 2014 SGR 966, and by the Technische Universität München – Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement no. 291763 (MM). The authors are grateful to two anonymous referees for their comments that helped to improve the paper, and in particular for one referee for suggesting the application of the construction from the proof of [52] Theorem 14.

REFERENCES

[1] H. Araki, Gibbs States of a one dimensional quantum lattice, Commun. Math. Phys. 14, 120–157, (1969).
[2] H. Araki, On uniqueness of KMS states of one-dimensional quantum lattice systems, Commun. Math. Phys. 44, 1–7, (1975).
[3] K.M.R. Audenaert, J. Calsamiglia, Ll. Masanes, R. Munoz-Tapia, A. Acin, E. Bagan, F. Verstraete.: Discriminating states: the quantum Chernoff bound; Phys. Rev. Lett. 98 160501, (2007)
[4] K.M.R. Audenaert, M. Nussbaum, A. Szkoła, F. Verstraete: Asymmetric error rates in quantum hypothesis testing; Commun. Math. Phys. 279, 251–283 (2008).
[5] R. Bhatia: Matrix Analysis; Springer, (1997)
[6] Salman Beigi: Quantum Rényi divergence satisfies data processing inequality; J. Math. Phys., 54, 122202 (2013)
[7] R.E. Blahut: Hypothesis testing and information theory; IEEE Trans. Inform. Theory vol. 20, issue 4, pp. 405–417, (1974)
[8] O. Bratteli, D.W. Robinson: Operator algebras and quantum statistical mechanics II.; Springer, (1981)
[9] Po-Ning Chen: Generalization of Gärtner-Ellis Theorem; IEEE Transactions on Information Theory 46,7, pp. 2752–2760, (2000)
[10] I. Csiszár: Generalized cutoff rates and Rényi’s information measures; IEEE Trans. Inf. Theory 41, 26–34, (1995)
[11] Didier Cucker-CASTelle: Formule de Chernoff pour une suite de variables réelles. In: Grandes Deviations et Applications Statistiques, Astérisque 68, pp. 19–24, (1979)
[12] N. Datta: Min- and Max-Relative Entropies and a New Entanglement Monotone; IEEE Transactions on Information Theory, vol. 55, no. 6, pp. 2816–2826, (2009).
[13] A. Dembo, O. Zeitouni: Large Deviations Techniques and Applications : Second ed., Springer, Application of Mathematics, Vol. 38, (1998)
[14] B. Derrick, M. Fannes, M. Pogorzelska: Fermanic Quasi-Free states and Maps in Information Theory; J. Math. Phys. 49, 032109, (2008)
[15] M. Fannes, B. Nachtergaele, R.F. Werner: Finitely correlated states on quantum spin chains; Commun. Math. Phys. 144, 443–490, (1992)
[16] Rupert L. Frank and Elliott H. Lieb: Monotonicity of a relative Rényi entropy; J. Math. Phys. 54 , 122201, (2013)
[17] U. Grenander, G. Szegő: Toeplitz Forms and Applications; University of California, Berkeley, (1958)
[18] T.S. Han and K. Kobayashi: The strong converse theorem for hypothesis testing; IEEE Trans. Inform. Theory, vol. 35, pp. 178–180, (1989)
[19] T.S. Han: Hypothesis testing with the general source; IEEE Trans. Inf. Theory, vol. 46, pp. 2415–2427, (2000)
[20] T.S. Han: Information-Spectrum Methods in Information Theory; Springer-Verlag, Berlin, Germany, (2003)
[21] M. Hayashi, T. Ogawa: On error exponents in quantum hypothesis testing; IEEE Trans. Inf. Theory, vol. 50, issue 6, pp. 1368–1372, (2004)
[22] M. Hayashi, Optimal sequence of POVM’s in the sense of Stein’s lemma in quantum hypothesis testing; J. Phys. A: Math. Gen. 35, pp. 10759–10773, (2002).
[23] M. Hayashi, Quantum Information Theory: An Introduction; Springer, (2006).
[24] M. Hayashi: Error exponent in asymmetric quantum hypothesis testing and its application to classical–quantum channel coding; Phys. Rev. A 76, 062301, (2007).
[25] Masahito Hayashi: Large deviation analysis for quantum security via smoothing of Rényi entropy of order 2; IEEE Transactions on Information Theory, Volume 60, Issue 10, pp. 6702–6732, (2014)
[26] Masahito Hayashi, Shun Watanabe; Information Geometry Approach to Parameter Estimation in Markov Chains; arXiv:1401.3814, (2014)
[27] Masahito Hayashi, Marco Tomamichel: Correlation Detection and an Operational Interpretation of the Rényi Mutual Information; arXiv:1408.6894
[28] F. Hiai, D. Petz: The proper formula for relative entropy and its asymptotics in quantum probability; Comm. Math. Phys. 143, 99–114 (1991).
[29] F. Hiai, M. Mosonyi, T. Ogawa: Large deviations and Chernoff bound for certain correlated states on a spin chain; J. Math. Phys. 48, (2007)
[30] F. Hiai, M. Mosonyi, T. Ogawa: Error exponents in hypothesis testing for correlated states on a spin chain; J. Math. Phys. 49, 032112, (2008)
Milán Mosonyi, F. Hiai, T. Ogawa, M. Fannes: Asymptotic distinguishability measures for shift-invariant quasifree states of fermionic lattice systems; J. Math. Phys. 49, 072104, (2008)

M. Mosonyi, H. Nagaoka: Hypothesis testing for Gaussian states on bosonic lattices; J. Math. Phys. 50, 032104, (2009)

M. Mosonyi, F. Hiai: On the quantum Rényi relative entropies and related capacity formulas; IEEE Trans. Inf. Theory, 57, pp. 2474–2487, (2011)

Milán Mosonyi, Tomohiro Ogawa: Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies; Communications in Mathematical Physics, Volume 334, Issue 3, pp. 1617–1648, (2015)

Milán Mosonyi, H. Nagaoka: Coding theorems for compound problems via quantum Rényi divergences; arXiv:1310.7525, (2013); to appear in IEEE Transactions on Information Theory

Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, Marco Tomamichel: On quantum Rényi entropies: a new definition and some properties; J. Math. Phys. 54, 122203, (2013)

H. Nagaoka: Strong converse theorems in quantum information theory; in the book “Asymptotic Theory of Quantum Statistical Inference” edited by M. Hayashi, World Scientific, (2005)

H. Nagaoka: The converse part of the theorem for quantum Hoeffding bound; quant-ph/0611289

H. Nagaoka, M. Hayashi: An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses; IEEE Trans. Inform. Theory 53, 534–549, (2007)

Kenji Nakagawa, Fumio Kanaya: On the converse theorem in statistical hypothesis testing; IEEE Transactions on Information Theory, vol. 39, no. 2, pp. 623–628, (1993)

Kenji Nakagawa, Fumio Kanaya: On the converse theorem in statistical hypothesis testing for Markov chains; IEEE Transactions on Information Theory, vol. 39, no. 2, pp. 629–633, (1993)

S. Natarajan: Large deviations, hypothesis testing, and source coding for finite Markov chains; IEEE Transactions on Information Theory 31, pp. 360–365, (1985)

M. Nussbaum, A. Szkoła: A lower bound of Chernoff type for symmetric quantum hypothesis testing; Ann. Statist. 37, 1040–1057, (2009)

T. Ogawa, H. Nagaoka: Strong converse and Stein’s lemma in quantum hypothesis testing; IEEE Trans. Inform. Theory 47, 2428–2433 (2000).

M. Ohya, D. Petz: Quantum Entropy and its Use; Springer, (1993)

D. Petz: Quasi-entropies for finite quantum systems; Rep. Math. Phys. 23, 57–65, (1986)

A. Rényi: On measures of entropy and information; Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I, pp. 547–561, Univ. California Press, Berkeley, California, (1961)

R. Renner: Security of Quantum Key Distribution, PhD dissertation, Swiss Federal Institute of Technology Zurich, Diss. ETH No. 16242, (2005).

Marco Tomamichel, Masahito Hayashi: A Hierarchy of information quantities for finite block length analysis of quantum tasks; IEEE Transactions on Information Theory, Volume: 59 , Issue: 11, pp. 7693–7710, (2013)

H. Umegaki: Conditional expectation in an operator algebra; Kodai Math. Sem. Rep. 14, 59–85, (1962)

Karol Váleg: On the error exponent for ergodic Markov source; Kybernetika, vol. 16, no. 4, pp. 318–329, (1980)

Shun Watanabe, Masahito Hayashi: Finite-length Analysis on Tail probability for Markov Chain and Application to Simple Hypothesis Testing; arXiv:1401.3801, (2014)

Alfred Wehrl: General properties of entropy; Rev. Mod. Phys. 50, pp. 221–260, (1978)

Mark M. Wilde, Andreas Winter, Dong Yang: Strong converse for the classical capacity of entanglement-breaking and Hadamard channels; Communications in Mathematical Physics, 331, pp. 593–622, (2014)