Correlation Detection and an Operational Interpretation of the Rényi Mutual Information

Masahito Hayashi\textsuperscript{1,2} and Marco Tomamichel\textsuperscript{3,2}

\textsuperscript{1}Graduate School of Mathematics, Nagoya University, Furocho, Chikusaku, Nagoya, 464-860, Japan
\textsuperscript{2}Centre for Quantum Technologies, National University of Singapore, Singapore 117543, Singapore
\textsuperscript{3}School of Physics, The University of Sydney, Sydney 2006, Australia

Recently, a variety of new measures of quantum Rényi mutual information and quantum Rényi conditional entropy have been proposed, and some of their mathematical properties explored. Here, we show that the Rényi mutual information attains operational meaning in the context of composite hypothesis testing, when the null hypothesis is a fixed bipartite state and the alternate hypothesis consists of all product states that share one marginal with the null hypothesis. This hypothesis testing problem occurs naturally in channel coding, where it corresponds to testing whether a state is the output of a given quantum channel or of a “useless” channel whose output is decoupled from the environment. Similarly, we establish an operational interpretation of Rényi conditional entropy by choosing an alternative hypothesis that consists of product states that are maximally mixed on one system. Specialized to classical probability distributions, our results establish an operational interpretation of Rényi mutual information and Rényi conditional entropy.

Full version available at arxiv: 1408.6894.

Introduction. In order to distill useful measures of Rényi mutual information and Rényi conditional entropy from a plethora of possible definitions, it is important to find out which definitions correspond to relevant operational quantities. For this purpose, let us consider how efficiently an arbitrary bipartite correlated state $\rho_{AB}$ on systems $A$ and $B$ can be distinguished from product states when the marginal of $\rho_{AB}$ on $A$ is known to be $\rho_A$. This problem can be regarded as the problem of detecting correlations in the state $\rho_{AB}$. Formally, we consider the following binary composite hypothesis testing problem for $n$ copies of such a state\textsuperscript{1}:

Null Hypothesis: The state is $\rho_{AB}^\otimes n$.

Alternate Hypothesis: The state is of the form $\rho_A^\otimes n \otimes \sigma_B^n$ with $\sigma_B^n$ any state on $n$ copies of $B$.

This problem figures prominently when analyzing the converse to various channel coding questions in classical as well as quantum information processing.\textsuperscript{2} There, the problem is specified by a description of a channel $\mathcal{E}_{A' \rightarrow B}$ and a bipartite state $\rho_{AA'}$ where the system $A$ constitutes an environment of the channel, $A'$ is the channel input, and $B$ its output. We are given an unknown state on $n$ copies of $A$ and $B$ and consider the following two hypotheses.

Null Hypothesis: The state is the output of $n$ uses of the channel $\mathcal{E}_{A' \rightarrow B}$, namely the state is exactly $\rho_{AB}^\otimes n$ where $\rho_{AB} := \mathcal{E}_{A' \rightarrow B}[\rho_{AA'}]$.

Alternate Hypothesis: The state is the output of a “useless” channel and decoupled from the environment, namely it is of the form $\rho_A^\otimes n \otimes \sigma_B^n$ with $\sigma_B^n$ any state on $n$ copies of $B$.

\begin{itemize}
  \item[1] We want to consider the speed with which the probability that we erroneously support the state $\rho_{AB}^\otimes n$ when the actual state is a product state of the form $\rho_A^\otimes n \otimes \sigma_B^n$ under a constraint for the opposite error. As is explained later, this problem can be discussed as the Hoeffding bound and Stein’s lemma under this formulation.
  \item[2] There exists an intimate connection between quantum channel coding and binary hypothesis testing (see, e.g., [8]). This connection is particularly important when analyzing how much information can be transmitted with a single use of a quantum channel [11, 23] or when approximating how much information can be transmitted with finitely many uses of the channel [3, 22]. (See also [7, 19] for the classical case. In particular, Polyanskiy [18, Sec. II] discusses the classical special case of this hypothesis testing problem.)
\end{itemize}
A hypothesis test for this problem is a binary positive operator-valued measure \( \{Q_{A^n B^n}, 1_{A^n B^n} - Q_{A^n B^n}\} \) on the \( n \) copies of the systems \( A \) and \( B \), determined by an operator \( 0 \leq Q_{A^n B^n} \leq 1_{A^n B^n} \). If the operator \( Q_{A^n B^n} \) “clicks” on our state, we conclude that the null hypothesis is correct, whereas otherwise we conclude that the alternate hypothesis is correct. The error of the first kind, \( \alpha_n(Q_{A^n B^n}) \), is defined as the probability with which we wrongly conclude that the alternate hypothesis is correct even if the state is \( \rho_{AB}^n \), given by

\[
\alpha_n(Q_{A^n B^n}) = \text{tr}[\rho^n_{AB}(1_{A^n B^n} - Q_{A^n B^n})].
\]

(1)

Conversely, the error of the second kind, \( \beta_n(Q_{A^n B^n}) \), is defined as the probability with which we wrongly conclude that the null hypothesis is correct even if the state is of the form \( \rho^n_A \otimes \sigma_{B^n} \) for some \( \sigma_{B^n} \), given by

\[
\beta_n(Q_{A^n B^n}) = \max_{\sigma_{B^n}} \text{tr}[\rho^n_A \otimes \sigma_{B^n} Q_{A^n B^n}],
\]

(2)

where the maximum is taken over all states \( \sigma_{B^n} \) on \( n \) copies of \( B \).

**Main Results.** The main contribution of this paper is an asymptotic analysis of the fundamental trade-off between these two errors as \( n \) goes to infinity. To investigate this trade-off, we ask the following questions: let us assume that our test is such that \( \hat{\alpha}_n(R) \) must approach one exponentially fast in \( n \), and investigate the exact exponents with which this error vanishes as \( n \) goes to infinity, yielding a quantum Hoeffding bound [6, 15] for our composite hypothesis testing problem. We find that the exponents are determined by the Rényi mutual information, defined as

\[
I_\alpha(A:B)_\rho = \min_{\sigma_B} D_\alpha(\rho_{AB}\|\rho_A \otimes \sigma_B), \quad \text{for} \quad \alpha \in (0, 1),
\]

(4)

where \( D_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \log \text{tr} \left[ \sigma^{1-\alpha} \rho^\alpha \sigma^{1-\alpha} \right] \) is the Rényi relative entropy first investigated by Petz (see, e.g. [17]) and the minimization is over all states \( \sigma_B \) on \( B \). We obtain

\[
\lim_{n \to \infty} \left\{ -\frac{1}{n} \log \hat{\alpha}_n(nR) \right\} = \sup_{s \in (0, 1)} \left\{ \frac{1}{s} (I_\alpha(A:B)_\rho - R) \right\}. \tag{5}
\]

On the other hand, if \( R > I(A:B)_\rho \), we show that \( \hat{\alpha}_n(R) \) must approach one exponentially fast in \( n \). This implies the strong converse for quantum Stein’s lemma [16] for our problem. We then find the exact exponents (also called strong converse exponents, see [5, Ch. 3] and [12, 16]) with which the error of the first kind goes to one as \( n \) goes to infinity and we find that in our case the exponent is determined by the sandwiched Rényi mutual information [1, 4], given as

\[
\tilde{I}_\alpha(A:B)_\rho = \min_{\sigma_B} \tilde{D}_\alpha(\rho_{AB}\|\rho_A \otimes \sigma_B), \quad \text{for} \quad \alpha > 1,
\]

(6)

where \( \tilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \log \text{tr} \left[ (\sigma^{1-\alpha} \rho^\alpha \sigma^{1-\alpha})^\alpha \right] \) is the (sandwiched) Rényi divergence [14, 24]. We obtain

\[
\lim_{n \to \infty} \left\{ -\frac{1}{n} \log (1 - \hat{\alpha}_n(nR)) \right\} = \sup_{s > 1} \left\{ \frac{s}{s - 1} (R - \tilde{I}_s(A:B)_\rho) \right\}. \tag{7}
\]
Hence, we show that the above composite hypothesis testing problem yields an operational interpretation for different definitions of the Rényi mutual information for the two ranges of $\alpha$, paralleling the observation in [12].

Finally, we also perform a second-order analysis for quantum Stein’s lemma [10, 21] and show that the minimal error of the first kind converges to a constant if $\beta_n(Q_{\rho}^A \otimes B^n) \leq \exp(-nI(A:B)_\rho - \sqrt{n}r)$ for some $r \in \mathbb{R}$. Then, for any $r \in \mathbb{R}$, we have

$$\lim_{n \to \infty} \left\{ \hat{\alpha}_n(nI(A:B)_\rho + \sqrt{n}\,r) \right\} = \Phi\left( \frac{r}{\sqrt{V(A:B)_\rho}} \right),$$

where $\Phi$ is the cumulative standard normal (Gaussian) distribution and

$$V(A:B)_\rho := \text{tr}\left[ \rho_{AB} \left( \log \rho_{AB} - \log \rho_A \otimes \rho_B - I(A:B)_\rho \right)^2 \right].$$

(9)

is the mutual information variance.

**Conditional Entropy.** Analogously, an operational interpretation for conditional Rényi entropies is established by considering the following binary hypotheses testing problem, which is motivated by the task of decoupling of quantum states. The problem is specified by a description of a state $\rho_{AB}$. Given an unknown state on $A$ and $B$, consider the following two hypotheses:

**Null Hypothesis:** The state is the $n$-fold product of $\rho_{AB}$, namely $\rho_{AB}^\otimes n$.

**Alternate Hypothesis:** The state is uniform on $A^n$ and decoupled form $B^n$, i.e. it is of the form $\rho_{AB}^\otimes n \otimes \sigma_{B^n}$, where $\sigma_A$ is the fully mixed state on $A$.

The same analysis as above applied to this problem reveals that the exponents in the quantum Hoeffding bound are determined by the Rényi conditional entropies defined as [20]

$$H^\alpha_{\rho}(A|B) = -\min_{\sigma_B} D_\alpha(\rho_{AB}||1_A \otimes \sigma_B), \quad \text{for} \quad \alpha \in (0, 1),$$

and the strong converse exponents are determined by the sandwiched conditional Rényi entropies [14]

$$\tilde{H}^\alpha_{\rho}(A|B) = -\min_{\sigma_B} \tilde{D}_\alpha(\rho_{AB}||1_A \otimes \sigma_B), \quad \text{for} \quad \alpha > 1.$$  

(11)

**Main Proof Ideas.** The main ideas are quickly summarized as follows:

1. We show that there exists a test (a different one in the direct and converse regime, respectively) which works uniformly for all choices of $\sigma_{B^n}$ by using some elementary tools from group representation theory. Various tools, including a new minimax theorem are derived in order to prove the quantum Hoeffding bound.

2. To investigate the strong converse exponents, we show that the sandwiched Rényi mutual information can be achieved by pinching and show various other properties, including that it is differentiable in $\alpha$ for all $\alpha \geq \frac{1}{2}$. Then, we use the Gärtner-Ellis theorem for the large deviation analysis of the correlated distributions that result from pinching.

**Related Work.** Complementary and concurrent to this work, Cooney et al. [2] investigated the strong converse exponents for a similar hypothesis testing problem when adaptive strategies are allowed — however, they did not treat the case of a composite alternate hypothesis and they also did not analyze the error exponents in the quantum Hoeffding bound.

Our proof of the strong converse exponents parallels the development in a very recent preprint by Mosonyi and Ogawa [13]. There, the authors consider correlated states and use the Gärtner-Ellis theorem of classical large deviation theory in order to investigate the asymptotic error exponents in the presence of correlations. Here, we are not interested in correlated states per se, but our proof technique based on pinching naturally leads us to a classical hypothesis testing problem with correlated distributions, for which the Gärtner-Ellis theorem again provides the right solution.