Multiple Quantum Hypothesis Testing Expressions and Classical-Quantum Channel Converse Bounds

Gonzalo Vazquez-Vilar
Universidad Carlos III de Madrid, Spain
Email: gvazquez@ieee.org

Abstract—Alternative exact expressions are derived for the minimum error probability of a hypothesis test discriminating among \( M \) quantum states. The first expression corresponds to the error probability of a binary hypothesis test with certain parameters; the second involves the optimization of a given information-spectrum measure. Particularized in the classical-quantum channel coding setting, this characterization implies the tightness of two existing converse bounds; one derived by Matthews and Wehner using hypothesis-testing, and one obtained by Hayashi and Nagaoka via an information-spectrum approach.

I. INTRODUCTION

Optimal discrimination among multiple quantum states—quantum hypothesis testing—is at the core of several information processing tasks involving quantum-mechanical systems. When the number of hypotheses is two, quantum hypothesis testing allows a simple formulation in terms of two kinds of pairwise errors. The quantum version of the Neyman-Pearson lemma establishes the optimum binary test in this setting. This problem was first studied by Helstrom in [11] (see also [2], [3]). When the number of hypotheses is larger than two, a (classical) prior distribution is usually placed over the hypotheses. While there exists no closed form for the optimal test in general, optimality conditions can be obtained [4], [5].

For historical notes on the subject see [6, Ch. IV].

In the context of reliable communication, hypothesis testing has been instrumental in the derivation of converse bounds to the error probability both in the classical and quantum settings (see, e.g., [7], [8]). Recently, hypothesis testing gained interest as a very general approach to obtain converse bounds in the finite block-length regime. In classical channel coding, Polyanisky, Poor and Verdú derived the meta-converse bound based on an instance of binary hypothesis testing [9]. A similar approach was used by Wang and Renner to derive a finite block-length converse bound for classical-quantum channels [10], and by Matthews and Wehner to obtain a family of converse bounds for general quantum channels [11]. The results by Matthews and Wehner are general enough to recover the meta-converse bound in the classical setting and Wang-Renner converse bound in the classical-quantum setting.

The information-spectrum method studies the asymptotics of a certain random variable, often referred to as information density or information random variable. Using a quantum analogue of this quantity, Hayashi and Nagaoka studied quantum hypothesis testing [12], and classical-quantum channel coding [13], obtaining general bounds for both problems.

In this paper, we derive two alternative exact expressions for the minimum error probability of multiple quantum hypothesis testing when a (classical) prior distribution is placed over the hypotheses. The expressions obtained illustrate connections among hypothesis testing, information-spectrum measures and converse bounds in classical-quantum channel coding. An application to classical-quantum channel coding shows that Matthews-Wehner converse bound [11, Th. 19] and Hayashi-Nagaoka lemma [13, Lemma 4] with certain parameters yield the exact error probability. This work thus generalizes several results derived in [14] in the classical setting.

II. BACKGROUND

A. Notation

In the general case, a quantum state is described by a density operator \( \rho \) acting on some finite dimensional complex Hilbert space \( \mathcal{H} \). Density operators are self-adjoint, positive semidefinite, and have unit trace. A measurement on a quantum system is a mapping from the state of the system \( \rho \) to a classical outcome \( m \in \{1, \ldots, M\} \). A measurement is represented by a collection of positive self-adjoint operators \( \{\Pi_1, \ldots, \Pi_M\} \) such that \( \sum \Pi_m = \mathbb{1} \), where \( \mathbb{1} \) is the identity operator. These operators form a POVM (positive operator-valued measure). A measurement \( \{\Pi_1, \ldots, \Pi_M\} \) applied to \( \rho \) has outcome \( m \) with probability \( \text{Tr}(\rho \Pi_m) \).

For two self-adjoint operators \( A, B \), the notation \( A \geq B \) means that \( A - B \) is positive semidefinite. Similarly \( A \leq B \), \( A > B \), and \( A < B \) means that \( A - B \) is negative semidefinite, positive definite and negative definite, respectively. For a self-adjoint operator \( A \) with spectral decomposition \( A = \sum_i \lambda_i E_i \), where \( \{\lambda_i\} \) are the eigenvalues and \( \{E_i\} \) are the orthogonal projections onto the corresponding eigenspaces, we define

\[
\{A > 0\} \triangleq \sum_{i: \lambda_i > 0} E_i.
\]

This corresponds to the projector associated to the positive eigenspace of \( A \). We shall also use \( \{A \geq 0\} \triangleq \sum_{i: \lambda_i \geq 0} E_i \), \( \{A < 0\} \triangleq \sum_{i: \lambda_i < 0} E_i \) and \( \{A \leq 0\} \triangleq \sum_{i: \lambda_i \leq 0} E_i \).
B. Binary Hypothesis Testing

Let us consider a binary hypothesis test (with simple hypotheses) discriminating between the density operators $\rho_0$ and $\rho_1$ acting on $\mathcal{H}$. In order to distinguish between the two hypotheses we perform a measurement. We define a test measurement $\{T, T^\dagger\}$, such that $T$ and $T^\dagger \equiv \mathbb{1} - T$ are positive semidefinite. The test decides $\rho_0$ (resp. $\rho_1$) when the measurement outcome corresponding to $T$ (resp. $T^\dagger$) occurs.

Let $\epsilon_{ij}$ denote the probability of deciding $\rho_j$ when $\rho_i$ is the true hypothesis, $i, j = 0, 1, i \neq j$. More precisely, we define

$$\epsilon_{1|0}(T) \equiv 1 - \text{Tr}\left(\rho_0 T\right) = \text{Tr}\left(\rho_0 T^\dagger\right),$$

(2)

$$\epsilon_{0|1}(T) \equiv \text{Tr}\left(\rho_1 T\right).$$

(3)

Let $\alpha_\beta(\rho_0, \rho_1)$ denote the minimum error probability $\epsilon_{1|0}$ among all tests with $\epsilon_{0|1}$ at most $\beta$, that is,

$$\alpha_\beta(\rho_0, \rho_1) \equiv \inf_{T} \epsilon_{1|0}(T),$$

(4)

subject to $\epsilon_{0|1}(T) \leq \beta$.

The function $\alpha_\beta(\cdot, \cdot)$ is the inverse of the function $\beta_\alpha(\cdot, \cdot)$ appearing in $\text{[5, Th. I]}$ which is itself related to the hypothesis-testing relative entropy $\text{D}_\text{rel}(\rho_0(\cdot) | | \rho_1(\cdot)) = -\log \beta_\alpha(\rho_0(\cdot) | | \rho_1(\cdot))$.

When $\rho_0$ and $\rho_1$ commute, the test $T$ in (4) can be restricted to be diagonal in the (common) eigenbasis of $\rho_0$ and $\rho_1$, then (4) reduces to the classical case [13].

The quantum version of the Neyman-Pearson lemma characterizes the form of the test minimizing (4). Let $t \geq 0$ and let $P_+^t, P_-^t, P_0^t$ denote the projectors spanning the positive, negative and null eigenspaces of the matrix $\rho_0 - t \rho_1$, respectively, i.e.,

$$P_+^t \equiv \left\{ \rho_0 - t \rho_1 > 0 \right\},$$

(5)

$$P_-^t \equiv \left\{ \rho_0 - t \rho_1 < 0 \right\},$$

(6)

$$P_0^t \equiv \mathbb{1} - P_+^t - P_-^t.$$  

(7)

Lemma 1 (Neyman-Pearson lemma): The operator $T_{NP}$ is an optimal test between $\rho_0$ and $\rho_1$ if and only if

$$T_{NP} = P_+^t + P_0^t,$$

(8)

where $0 \leq p_0^t \leq P_0^t$.

Proof: A slightly different formulation of this result is usually given in the literature. The statement included here can be found in, e.g., [13, LRM. 3].

Therefore, for any $t \geq 0$ and $0 \leq p_0^t \leq P_0^t$ such that $\text{Tr}\{\rho_1 T_{NP}\} = \beta$, the resulting test $T_{NP}$ minimizes (4). Moreover, the following lower bound holds.

Lemma 2: For any test discriminating between $\rho_0$ and $\rho_1$, and for any $t' \geq 0$, it holds that

$$\alpha_\beta(\rho_0, \rho_1) \geq \text{Tr}\left(\rho_0 (P_{+}^{t'} + P_{0}^{t'})\right) - t' \beta.$$  

(9)

Proof: For any operator $A \geq 0$ and $0 \leq t \leq \mathbb{1}$, it holds that $\text{Tr}\{A(A > 0)\} \geq \text{Tr}\{AT\}$ [12, Eq. 8]. For $A = \rho_0 - t' \rho_1$ and $T = T_{NP}$, this inequality becomes

$$\text{Tr}\{(\rho_0 - t' \rho_1) P_{+}^{t'}\} \geq \text{Tr}\{(\rho_0 - t' \rho_1) T_{NP}\},$$

(10)

which after some algebra yields

$$-\text{Tr}\{\rho_0 T_{NP}\} \geq -\text{Tr}\{\rho_0 P_{+}^{t'}\} + t' \text{Tr}\{\rho_1 (P_{+}^{t'} - T_{NP})\}.$$  

(11)

Summing one to both sides of (11) and noting that $\alpha_\beta(\rho_0, \rho_1) = 1 - \text{Tr}\{\rho_0 T_{NP}\}$ and $\beta = \text{Tr}\{\rho_1 T_{NP}\}$, we obtain

$$\alpha_\beta(\rho_0, \rho_1) \geq \text{Tr}\{\rho_0 (P_{+}^{t'} + P_{0}^{t'})\} + t' \text{Tr}\{\rho_1 (P_{+}^{t'} - t' \beta).$$

(12)

The result thus follows by lower-bounding $\text{Tr}\{\rho_1 P_{+}^{t'}\} \geq 0$.

III. MULTIPLE QUANTUM HYPOTHESIS TESTING

We consider a hypothesis testing problem discriminating among $M$ possible states acting on $\mathcal{H}$, where $M$ is assumed to be finite. The $M$ alternative hypotheses $\tau_1, \ldots, \tau_M$ are assumed to occur with (classical) probabilities $p_1, \ldots, p_M$, respectively.

A $M$-ary hypothesis test is a POVM $\mathcal{P} \equiv \{\Pi_1, \Pi_2, \ldots, \Pi_M\}$, where $\sum_i \Pi_i = \mathbb{1}$. The test decides the alternative $\tau_i$ when the measurement with respect to $\mathcal{P}$ has outcome $i$. The probability that the test $\mathcal{P}$ decides $\tau_i$ when $\tau_i$ is the true underlying state is thus $\text{Tr}\{\Pi_i \tau_i\}$ and the average error probability is

$$\epsilon(\mathcal{P}) \equiv 1 - \sum_{i=1}^{M} p_i \text{Tr}\{\tau_i \Pi_i\}.$$  

We define the minimum average error probability as

$$\epsilon \equiv \min_{\mathcal{P}} \epsilon(\mathcal{P}).$$

(13)

The test $\mathcal{P}$ minimizing (13) has no simple form in general.

Lemma 3 (Holevo-Yuen-Kennedy-Lax Conditions): A test $\mathcal{P}^* = \{\Pi_1^*, \ldots, \Pi_M^*\}$ minimizes (13) if and only if, for each $m = 1, \ldots, M$,

$$\Lambda(\mathcal{P}^*) - p_m \tau_m \Pi_m^* = \Pi_m^* (\Lambda(\mathcal{P}^*) - p_m \tau_m) = 0,$$

(15)

$$\Lambda(\mathcal{P}^*) - p_m \tau_m \geq 0.$$  

(16)

where

$$\Lambda(\mathcal{P}^*) \equiv \sum_{i=1}^{M} p_i \tau_i \Pi_i^* = \sum_{i=1}^{M} p_i \Pi_i^* \tau_i.$$  

(17)

is required to be self-adjoint [1].

Proof: The theorem follows from [3, Th. 4.1, Eq. (4.8)] or [5, Th. 1] after simplifying the resulting optimality conditions.

We next show an alternative characterization of the minimum error probability $\epsilon$ as a function of a binary hypothesis test with certain parameters.

Let $\text{diag}(\rho_1, \ldots, \rho_M)$ denote the block-diagonal matrix with diagonal blocks $\rho_1, \ldots, \rho_M$. We define

$$\mathcal{T} \equiv \text{diag}(\rho_1 \tau_1, \ldots, \rho_M \tau_M),$$

(18)

$$\mathcal{D}(\mu_0) \equiv \text{diag}(\mu_0, \ldots, \mu_0),$$

(19)

where $\mu_0$ is an arbitrary density operator acting on $\mathcal{H}$. Note that both $\mathcal{T}$ and $\mathcal{D}(\mu_0)$ are density operators themselves, since they are self-adjoint, positive semidefinite and have unit trace.

1 The operator $\Lambda(\mathcal{P})$ takes a role of the Lagrange multiplier associated to the constraint $\sum_i \Pi_i = \mathbb{1}$, which involving self-adjoint operators requires $\Lambda$ to be self-adjoint.
Theorem 1: The minimum error probability of an $M$-ary test discriminating among states $\{\tau_1, \ldots, \tau_M\}$ with prior classical probabilities $\{p_1, \ldots, p_M\}$ satisfies
\[
\epsilon = \max_{\mu_0} \alpha \left( \mathcal{T} \| D(\mu_0) \right),
\] (20)
where $\mathcal{T}$ and $D(\cdot)$ are given in (18) and (19), respectively, and where the optimization is carried out over (unit-trace non-negative) density operators $\mu_0$.

Proof: For any $\mathcal{P} = \{\Pi_1, \Pi_2, \ldots, \Pi_M\}$ let us define the binary test $T' \triangleq \text{diag} (\Pi_1, \ldots, \Pi_M)$. For this test we obtain
\[
\epsilon_{1|0}(T') = 1 - \sum_{i=1}^{M} p_i \text{Tr} (\tau_i \Pi_i) = \epsilon(\mathcal{P}),
\] (21)
\[
\epsilon_{0|1}(T') = \frac{1}{M} \sum_{i=1}^{M} \text{Tr} (\mu_0 \Pi_i)
\] (22)
\[
= \frac{1}{M} \text{Tr} \left( \mu_0 \left( \sum_{i=1}^{M} \Pi_i \right) \right)
\] (23)
\[
= \frac{1}{M} \text{Tr} (\mu_0) = \frac{1}{M}
\] (24)
The (possibly suboptimal) test $T'$ has thus $\epsilon_{1|0}(T') = \epsilon(\mathcal{P})$ and $\epsilon_{0|1}(T') = \frac{1}{M}$. Therefore, using (4) and maximizing the resulting expression over $\mu_0$, we obtain
\[
\epsilon(\mathcal{P}) \geq \max_{\mu_0} \alpha \left( \mathcal{T} \| D(\mu_0) \right).
\] (25)

It remains to show that, for $\mathcal{P} = \mathcal{P}^*$ defined in Lemma 3, the lower bound (25) holds with equality. To this end, we next demonstrate that the optimality conditions for $T_{NP}$ in Lemma 1 and for $\mathcal{P}^* = \{\Pi_1^*, \ldots, \Pi_M^*\}$ in Lemma 4 are equivalent for a specific choice of $\mu_0$.

Let $\mathcal{P}^* = \{\Pi_1^*, \ldots, \Pi_M^*\}$ satisfy (15)–(16) and define
\[
\mu_0^* \triangleq \frac{1}{c_0^*} \sum_{i=1}^{M} p_i \tau_i \Pi_i^* = \frac{1}{c_0^*} \Lambda(\mathcal{P}^*),
\] (26)
where $c_0^*$ is a normalizing constant such that $\mu_0^*$ is unit trace.

Lemma 1 shows that the test $T_{NP}$ achieving (25) is associated to the non-negative eigenspace of the matrix $\mathcal{T} - tD(\mu_0)$. Given the block-diagonal structure of the matrix $\mathcal{T} - tD(\mu_0)$, it is enough to consider binary tests $T_{NP}$ with block-diagonal structure. Then, we write $T_{NP} = \text{diag} (T_{NP}^1, \ldots, T_{NP}^M)$.

For the choice $\mu_0 = \mu_0^*$, and $t = M c_0^*$, the $m$-th block-diagonal term in $\mathcal{T} - tD(\mu_0)$ is given by
\[
p_m \tau_m - \frac{1}{M} \mu_0 = p_m \tau_m - \Lambda(\mathcal{P})
\] (27)
The $m$-th block of the Neyman-Pearson test $T_{NP}^m$ must lie in the non-negative eigenspace of the matrix (27). However, since (16) implies that (27) is negative semidefinite, each block $T_{NP}^m$ can only lie in the null eigenspace of (27), $m = 1, \ldots, M$. According to (15), the operator $\Pi_m^*$ belongs to the null eigenspace of (27), $m = 1, \ldots, M$. As a result, the choice
\[
T_{NP} = \text{diag} (\Pi_1^*, \ldots, \Pi_M^*)
\] (28)
satisfies the optimality conditions in Lemma 1. Moreover, since $\epsilon_{1|0}(T_{NP}) = \epsilon(\mathcal{P}^*) = \epsilon$ and $\epsilon_{0|1}(T_{NP}) = \frac{1}{M}$, Lemma 1 implies that (20) holds with equality for $\mu_0 = \mu_0^*$. Given the bound in (25), other choices of $\mu_0$ cannot improve the result, and Theorem 1 thus follows.

Combining Theorem 1 and Lemma 2 we obtain a characterization for $\epsilon$ based on information-spectrum measures.

Theorem 2: The minimum error probability of an $M$-ary test discriminating among states $\{\tau_1, \ldots, \tau_M\}$ with prior classical probabilities $\{p_1, \ldots, p_M\}$ satisfies
\[
\epsilon = \max_{\mu_0, t \geq 0} \left\{ \sum_{i=1}^{M} p_i \text{Tr} (\tau_i \{\tau_i - t \mu_0 \leq 0\}) - t \right\},
\] (29)
where the optimization is carried out over (unit-trace non-negative) density operators $\mu_0$ acting on $\mathcal{H}$, and over the scalar threshold $t \geq 0$.

Proof: Applying Lemma 2 to (20), and using the definitions of $\mathcal{T}$ in (18) and $D(\cdot)$ in (19), yields, for any $\mu_0, t' \geq 0$,
\[
\epsilon \geq \sum_{i=1}^{M} p_i \text{Tr} (\tau_i \{\tau_i - t' \mu_0 \leq 0\}) - t'
\] (30)
It remains to show that there exist $\mu_0$ and $t' \geq 0$ such that (30) holds with equality. In particular, let us choose $\mu_0 = \mu_0^*$ defined in (26), and $t' = M c_0^*$ where $c_0^* = \sum_{i=1}^{M} p_i \text{Tr}(\tau_i \Pi_i^*)$ is the normalizing constant from (26).

For this choice of $\mu_0$ and $t'$, the projector spanning the negative semidefinite eigenspace of the operator $p_i \tau_i - \frac{t'}{M} \mu_0$ can be rewritten as
\[
\{p_i \tau_i - \frac{t'}{M} \mu_0 \leq 0\} = \{p_i \tau_i - \Lambda(\mathcal{P}^*) \leq 0\}
\] (31)
\[
= \mathbb{1},
\] (32)
where the last identity follows from (16). The right-hand side of (30) thus becomes
\[
\sum_{i=1}^{M} p_i \text{Tr}(\tau_i) - \frac{t'}{M} = 1 - \frac{t'}{M}
\] (33)
The result follows since $\frac{t'}{M} = c_0^* = \sum_{i=1}^{M} p_i \text{Tr}(\tau_i \Pi_i^*) = 1 - \epsilon$.

The alternative expressions derived in Theorems 1 and 2 are not easier to compute than the original optimization in (14), all of them requiring to solve a semidefinite program. We recall from the proofs of the theorems that a density operator $\mu_0$ maximizing (20) and (29) is
\[
\mu_0^* \triangleq \frac{1}{c_0^*} \sum_{i=1}^{M} p_i \tau_i \Pi_i^*,
\] (34)
for some $\mathcal{P}^* = \{\Pi_1^*, \ldots, \Pi_M^*\}$ satisfying the conditions in Lemma 3 and where $c_0^*$ is a normalizing constant. Hence, the optimal $M$-ary hypothesis test $\mathcal{P}^*$ characterizes the optimal $\mu_0$. Conversely, the optimal $\mu_0$ is precisely the Lagrange multiplier associated to the minimization in (14), after an appropriate re-scaling.

The expressions derived in Theorems 1 and 2 can be used to determine the tightness of several converse bounds from the literature, as we show in the next section.
IV. APPLICATION TO CLASSICAL-QUANTUM CHANNELS

We consider the channel coding problem of transmitting $M$ equiprobable messages over a one-shot classical-quantum channel $x \rightarrow W_x$, with $x \in \mathcal{X}$ and $W_x \in \mathcal{H}$.

A channel code is defined as a mapping from the message set $\{1, \ldots, M\}$ into a set of $M$ codewords $\mathcal{C} = \{x_1, \ldots, x_M\}$. For a source message $m$, the decoder receives the associated density operator $W_{x_m}$ and must decide on the transmitted message. The minimum error probability for a code $\mathcal{C}$ is

$$P_e(\mathcal{C}) \triangleq \min_{(\Pi_1, \ldots, \Pi_M)} \left\{ 1 - \frac{1}{M} \sum_{m=1}^{M} \text{Tr}(W_{x_m} \Pi_m) \right\}. \quad (35)$$

This problem corresponds precisely to the $M$-ary quantum hypothesis testing problem described in Section III. Then, direct application of Theorems 1 and 2 yields two alternative expressions for $P_e(\mathcal{C})$.

Let $A$ and $B$ denote the input and output of the system, respectively. The joint state induced by a codebook $\mathcal{C}$ is

$$\rho_{AB}^0 = \frac{1}{M} \sum_{x \in \mathcal{C}} |x\rangle \langle x|^A \otimes W_x^B, \quad (36)$$

and $\rho_{C}^0 = \frac{1}{M} \sum_{x \in \mathcal{C}} |x\rangle \langle x|^A$ its input marginal. According to (20) in Theorem 1 we obtain

$$P_e(\mathcal{C}) = \max_{\mu_0} \alpha_{\mathcal{C}} \left( \rho_{AB}^0 \parallel \rho_C^0 \otimes \rho_C^0 \right). \quad (37)$$

The expression (37) is precisely the finite block-length converse bound by Matthews and Wehner [11, Eq. (45)], particularized for a classical-quantum channel with an input state induced by the codebook $\mathcal{C}$. Therefore, Theorem 1 implies that the quantum generalization of the meta-converse bound proposed by Matthews and Wehner is tight for a fixed codebook $\mathcal{C}$.

Minimizing the right-hand side of (37) over all distributions $P_X$ defined over the input alphabet $\mathcal{X}$, not necessarily induced by a codebook, yields a lower bound on $P_e(\mathcal{C})$ for any codebook $\mathcal{C}$. By fixing $\mu_0$ to be the state induced at the system output, this lower bound recovers the converse bound by Wang and Renner [10, Th. 1]. This bound is not tight in general since (i) the minimizing $P_X$ does not need to coincide with the input state induced by the best codebook, and (ii) the choice of $\mu_0$ in [10, Th. 1] does not maximize the resulting bound in general.

Using the characterization in Theorem 2 the error probability $P_e(\mathcal{C})$ can be equivalently written as

$$P_e(\mathcal{C}) = \max_{\mu_0, t' \geq 0} \left\{ \frac{1}{M} \sum_{x \in \mathcal{C}} \text{Tr}\left(W_x \{W_x - t' \mu_0 \leq 0\}\right) - t' / M \right\}. \quad (38)$$

The objective of the maximization in (38) coincides with the information-spectrum bound [13, Lemma 4]. Then, (38) shows that the Hayashi-Nagaoka lemma yields the exact error probability for a fixed code, after optimization over the free parameters $\mu_0$, $t' \geq 0$.

V. CONCLUDING REMARKS

In Theorem 1 the minimum error probability of an $M$-ary quantum hypothesis test is expressed as an instance of a binary quantum hypothesis test with certain parameters. This expression implies the tightness of the converse bound [11, Th. 19] by Matthews and Wehner, and identifies the weakness of [10, Th. 1] by Wang and Renner in classical-quantum channel coding. For more general channels and entanglement-assisted codes, it is not clear whether the bounds in [11, Th. 18 and Th. 19] coincide with the exact error probability. To study this, a generalization of Theorem 1 imposing less structure over the test alternatives is needed. Theorem 2 shows that the minimum error probability can be written as an optimization problem involving information-spectrum measures. In particular, this expression shows that the Hayashi-Nagaoka lemma [13, Lemma 4] yields the exact error probability after optimization over its free parameters.

ACKNOWLEDGMENT

The problem studied here was suggested to the author by Alfonso Martinez. The author thanks him, Albert Guillén i Fábregas and William Matthews for stimulating discussions related to this work.

REFERENCES

[1] C. W. Helstrom, “Detection theory and quantum mechanics,” Inf. and Control, vol. 10, no. 3, pp. 254–291, 1967.
[2] P. A. Bakut and S. S. Shchurov, “Optimal detection of a quantum signal,” Problemy Peredachi Informatsii, vol. 4, no. 1, pp. 77–82, 1968, (in Russian, English translation: Probl. Inf. Transm., vol. 4, pp. 61-65, 1968).
[3] A. S. Holevo, “An analog of the theory of statistical decisions in noncommutative theory of probability,” Trudy Moskov. Mat. Obšč., vol. 26, pp. 133–149, 1972, (in Russian).
[4] ——, “Statistical decision theory for quantum systems,” J. Multivariate Anal., vol. 3, no. 4, pp. 337–394, 1973.
[5] H. P. Yuen, R. S. Kennedy, and M. Lax, “Optimum testing of multiple hypotheses in quantum detection theory,” IEEE Trans. Inf. Theory, vol. 21, no. 2, pp. 125–134, Mar 1975.
[6] C. W. Helstrom, Quantum Detection and Estimation Theory. NY: Academic Press, 1976.
[7] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, “Lower bounds to error probability for coding on discrete memoryless channels,” I. Inf. Contr., vol. 10, no. 1, pp. 65–103, 1967.
[8] H. Nagaoka, “Strong converse theorems in quantum information theory,” in Proc. ERATO Conf. Quantum Inf. Science, Tokyo, Japan, 2001, p. 33.
[9] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” IEEE Trans. Inf. Theory, vol. 56, no. 5, pp. 2307–2359, 2010.
[10] L. Wang and R. Renner, “One-shot classical-quantum capacity and hypothesis testing,” Phys. Rev. Lett., vol. 108, no. 20, p. 200501, 2012.
[11] W. Matthews and S. Wehner, “Finite blocklength converse bounds for quantum channels,” IEEE Trans. Inf. Theory, vol. 60, no. 11, pp. 7317–7329, 2014.
[12] H. Nagaoka and M. Hayashi, “An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses,” IEEE Trans. Inf. Theory, vol. 53, no. 2, pp. 534–549, 2007.
[13] M. Hayashi and H. Nagaoka, “General formulas for capacity of classical-quantum channels,” IEEE Trans. Inf. Theory, vol. 49, no. 7, pp. 1755–1768, 2003.
[14] G. Vazquez-Vilar, A. Tauste Campo, A. Guillén i Fábregas, and A. Martínez, “Bayesian $M$-ary hypothesis testing: The meta-converse and Verdú-Han bounds are tight,” IEEE Trans. Inf. Theory, 2016, to appear. Preprint available at arXiv:1411.3292.
[15] A. Jenčová, “Quantum hypothesis testing and sufficient subalgebras,” Lett. Math. Phys., vol. 93, no. 1, pp. 15–27, 2010.