Generalized Perfect Codes for Symmetric Classical-Quantum Channels

Andreu Blasco Coll, *Student Member, IEEE,* Gonzalo Vazquez-Vilar, *Member, IEEE,*
Javier Rodríguez Fonollosa, *Senior Member, IEEE*

***Abstract*—We define a new family of codes for symmetric classical-quantum channels and establish their optimality. To this end, we extend the classical notion of generalized perfect and quasi-perfect codes to channels defined over some finite dimensional complex Hilbert output space. The resulting optimality conditions depend on the channel considered and on an auxiliary state defined on the output space of the channel. For certain N-qubit classical-quantum channels, we show that codes based on a generalization of Bell states are quasi-perfect and, therefore, they feature the smallest error probability among all codes of the same blocklength and cardinality.**

Andreu Blasco Coll and Javier Rodríguez Fonollosa are with the Department de Teoria del Senyal i Comunicacions, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain (e-mail: andreu.blasco@upc.edu, javier.fonollosa@upc.edu).

G. Vazquez-Vilar is with the Signal Theory and Communications Department, Universidad Carlos III de Madrid, 28911 Leganés, Spain, and with the Gregorio Marañón Health Research Institute, 28007 Madrid, Spain (e-mail: gonzalo.vazquez@uc3m.es).

This work was supported in part by the European Research Council (ERC) under grant 714161, by the Agencia Estatal de Investigacin, Ministerio de Ciencia e Innovación, of the Spanish Government, under grants RED2018-102668-T, PID2019-104958RB-C41 and PID2020-116683GB-C21, and the Catalan Government, 2017 SGR 578 AGAUR and 001-P-001644 QuantumCAT within the ERDF Program of Catalunya.

This work was presented in part at the 2016 IEEE International Symposium on Information Theory, Barcelona, Spain.

***Index Terms*—Classical-quantum channel, finite blocklength analysis, quantum meta-converse, perfect code, quasi-perfect code, quantum hypothesis testing.**

I. Introduction

One of the goals of information theory is to characterize the best achievable performance of any transmission scheme and to establish the structure of codes and decoders attaining this limit. In certain regimes the best performance of a communication system is accurately characterized and there exist practical codes attaining it. In his landmark paper [1], Shannon demonstrated that for every communication channel there exists a fundamental limit, named channel capacity, that determines the highest rate at which a sender can transmit data to a receiver with vanishing decoding error probability. Nowadays, several code constructions achieve the channel capacity or perform very close to it. Specific examples, for sufficiently large codelengths, are low-density parity check (LDPC) codes [2], turbo codes [3], or polar codes [4].

If the length of the code is limited –e.g., due to delay constraints or due to the nature of the channel– the channel capacity is not a good benchmark anymore. To characterize the achievable performance in this regime, we shall use non-asymptotic bounds on the error probability of the best coding scheme. Two early instances of these limits are the sphere-packing bound [5, Eq.

(5.8.19)] on the error probability of the binary symmetric channel (BSC) and Shannon's non-asymptotic results for the additive white Gaussian noise (AWGN) channel [6]. Recent advances in information theory have lead to several (upper and lower) bounds in the finite block-length regime [7]–[9]. Code designers have optimized the finite-length performance of certain codes and now they perform close to those non-asymptotic limits (see, e.g., [10], [11] and references therein). Moreover, certain codes can even attain these limits with equality, thus proving their non-asymptotic optimality. For example, perfect and quasi-perfect binary codes attain the sphere-packing bound [5, Eq. (5.8.19)] for the BSC, and they were generalized beyond binary alphabets in [12].

The results presented above consider transmission channels –or random transformations– which are modeled by a transition probability distribution. Certain physical systems, however, can only be described using the laws of quantum mechanics. For these systems, the classical channel capacity and the corresponding non-asymptotic results have to be extended to encompass the quantum properties of the system. Holevo, Schumacher and Westmoreland studied the task of sending classical data over a channel with classical inputs and quantum outputs [13], [14]; this setting is usually referred to as *classical-quantum channel coding*. Their coding theorem guarantees the existence of reliable codes if their rate is below a fundamental limit, known as *Holevo capacity*, provided that the codelength is sufficiently long. While the proof of this result does not provide an explicit code construction, it guided the design of practical coding schemes. For example, quantum polar codes are practical constructions shown to attain this limit [15], the codes proposed in [16], [17] feature the superadditivity of mutual information, and other coding schemes exploiting the quantum properties of optical channels were proposed in [18], [19]. Holevo capacity is an asymptotic quantity that, in general, can only be attained by a large number of channel uses via a joint measurement on the combined channel outputs. For a finite number of channel uses –which is relevant for practical quantum systems– non-asymptotic performance limits need to be used. Converse non-asymptotic bounds were studied in [20], [21, Sec. 4.6] and [22], among other works. However, to the best of our knowledge, these works have not been applied in the design and/or benchmark of practical code constructions.

The derivation of converse bounds for classical and quantum systems is often based on hypothesis testing. An early application of this technique was used by Shannon, Gallager and Berlekamp to obtain the sphere-packing exponent [23] (see also [24] for a more explicit derivation for symmetric channels). For classical-quantum channels, Nagaoka applied binary hypothesis testing in the derivation of strong-converse bounds [25], and Hayashi used this technique to stablish the converse part of the channel-coding theorem [21, Sec. 4.6]. Later, Polyanskiy, Poor and Verdú formulated a hypothesis testing finite-length bound for classical channels [7, Th. 27], and Matthews and Wehner obtained finite-length bounds for general quantum channels in [22].

In this work, we extend the notion of generalized perfect and quasi-perfect codes [12] to symmetric classical-quantum channels. The results presented here are based on quantum hypothesis testing. In particular, we derive two alternative expressions for the error probability of quantum multiple hypothesis testing, which are then used to determine the exact error probability for a fixed classical-quantum channel code. A weakening of this result yields the non-asymptotic converse bound [22, Eq. (45)] (see also [21, Sec. 4.6]) and coincides with the error probability of the generalized perfect and quasi-perfect codes, whenever they exist. Therefore, these codes yield the best performance among any code with the same rate and block-length. While these codes are possibly rare, we characterize a family of codes based on a generalization of Bell states which are quasi-perfect for certain non-asymptotic 2-qubit classical-quantum channels and their N-qubit extension.

A separate line of research in the literature considers the evolution of quantum information in a noisy environment. In [26], Shor showed that quantum errors can be controlled by encoding the state of the system in a quantum code and periodically measuring the redundant parts of the code. This observation opened the field of *quantum error correction*. While it is possible to encapsulate classical information over quantum channels using quantum error correction codes, they are highly inefficient for this task and they will not be treated here.

The organization of this article is as follows. In Section II we formalize the problems of binary and multiple hypothesis testing and establish an unexplored connection between them. Section III presents the classical-quantum channel model and establishes the accuracy of different converse bounds in the literature. Section IV defines perfect and quasi-perfect codes for classical-quantum symmetric channels and proves their optimality whenever they exist. In Section V we study a family of codes which are quasi-perfect for 2-qubit classical-quantum channels affected by quantum depolarization. Section VI concludes the article with some final remarks.

A. Notation

Let $\mathcal{D}(\mathcal{H})$ denote the space of density operators acting on some finite dimensional complex Hilbert space $\mathcal{H}$. A quantum state is described by a density operator $\rho \in \mathcal{D}(\mathcal{H})$. 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 ρ to a classical outcome $m \in \{1, \dots, M\}$. A measurement is represented by a collection of positive semidefinite self-adjoint operators $\{\Pi_1, \dots, \Pi_M\}$ such that $\sum \Pi_m = \mathbb{1}$, where $\mathbb{1}$ is the identity operator. These operators form a *positive operator-valued measure* (POVM). A POVM measurement $\{\Pi_1, \dots, \Pi_M\}$ applied to ρ has outcome m with probability $\mathrm{Tr}(\rho\Pi_m)$ where Tr is the trace operator.

For 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 real eigenvalues and $\{E_i\}$ are the orthogonal projections onto the corresponding eigenspaces, we define

$$\{A > 0\} \triangleq \sum_{i:\lambda_i > 0} E_i. \tag{1}$$

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 $\{A \leq 0\} \triangleq \sum_{i:\lambda_i \leq 0} E_i$.

II. Quantum Hypothesis Testing

A. Binary Hypothesis Testing

Let us consider a binary hypothesis test (with simple hypotheses) discriminating between the density operators ρ_0 and ρ_1, where $\rho_0, \rho_1 \in \mathcal{D}(\mathcal{H})$. In order to distinguish between the two hypotheses we perform a measurement. We define a test measurement $\{T, \bar{T}\}$, such that T and $\bar{T} \triangleq \mathbb{1} - T$ are positive semidefinite, self-adjoint operators. The test decides ρ_0 (resp. ρ_1) when the measurement outcome corresponding to T (resp. $\bar{T}$) occurs.

Let $\epsilon_{j|i}$ denote the probability of deciding ρ_j when ρ_i is the true hypothesis, $i, j = 0, 1$, $i \neq j$. More precisely,

$$\epsilon_{1|0}(T) \triangleq 1 - \operatorname{Tr}(\rho_0 T) = \operatorname{Tr}\left(\rho_0 \bar{T}\right), \tag{2}$$

$$\epsilon_{0|1}(T) \triangleq \operatorname{Tr}(\rho_1 T). \tag{3}$$

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

$$\alpha_\beta(\rho_0 \| \rho_1) \triangleq \inf_{T : \epsilon_{0|1}(T) \leq \beta} \epsilon_{1|0}(T). \tag{4}$$

The function $\alpha_\beta(\cdot \| \cdot)$ is the inverse of the function $\beta_\alpha(\cdot \| \cdot)$ appearing in [22], which is itself related to the hypothesis-testing relative entropy as $D^{\alpha}_{\mathrm{H}}(\rho_0 \| \rho_1) = -\log \beta_\alpha(\rho_0 \| \rho_1)$ [27].

When ρ_0 and ρ_1 commute, the test T in (4) can be restricted to be diagonal in the (common) eigenbasis of ρ_0 and ρ_1, then (4) reduces to the classical case [28].

The form of the test minimizing (4) is given by the quantum Neyman-Pearson lemma, presented next.

Lemma 1 (Neyman-Pearson lemma): The best trade-off between type-I and type-II error probabilities is attained by tests of the form

$$T_{\mathrm{NP}} = \bigl\{\rho_0 - t\rho_1 > 0\bigr\} + \theta^0_t, \tag{5}$$

for some t and θ^0_t, and where $0 \leq \theta^0_t \leq \bigl\{\rho_0 - t\rho_1 = 0\bigr\}$.

Proof: A slightly different formulation of this result is usually given in the literature (see, e.g., [29, Ch. IV, Eq. (2.18)]). The precise statement included here can be found, e.g., in [30, Lem. 3]. ■

Then, for any choice of t and θ^0 such that $\operatorname{Tr}(\rho_1 T_{\mathrm{NP}}) = \beta$, the resulting test T_{NP} in (5) minimizes (4). The following result is a corollary to the Neyman-Pearson lemma that will be useful in the sequel.

Lemma 2: For any binary hypothesis test discriminating between the quantum states ρ_0 and ρ_1, it follows that

$$\begin{aligned}
\alpha_\beta(\rho_0 \| \rho_1) = \sup_{t \geq 0} \Bigl\{ & \operatorname{Tr}\bigl(\rho_0 \bigl\{\rho_0 - t\rho_1 \leq 0\bigr\}\bigr) \\
& + t\Bigl(\operatorname{Tr}\bigl(\rho_1 \bigl\{\rho_0 - t\rho_1 > 0\bigr\}\bigr) - \beta\Bigr)\Bigr\} \quad (6) \\
\geq \operatorname{Tr}\bigl(\rho_0 & \bigl\{\rho_0 - t'\rho_1 \leq 0\bigr\}\bigr) - t'\beta, \quad (7)
\end{aligned}$$

for any $t' \geq 0$.

Proof: The identity (6) is the quantum analogue of [12, Lem. 1] and the relaxation (7) coincides with [31, Lem. 2]. For completeness, we include next the proof of (6)-(7).

For any operator $A \geq 0$ and $0 \leq T \leq \mathbb{1}$, it holds that $\operatorname{Tr}\bigl(A\{A > 0\}\bigr) \geq \operatorname{Tr}\bigl(AT\bigr)$ [32, Eq. 8]. For $A = \rho_0 - t'\rho_1$ and $T = T_{\mathrm{NP}}$ defined in (5), this inequality becomes

$$\operatorname{Tr}\bigl((\rho_0 - t'\rho_1) P^{+}_{t'}\bigr) \geq \operatorname{Tr}\bigl((\rho_0 - t'\rho_1) T_{\mathrm{NP}}\bigr), \tag{8}$$

where we defined $P^{+}_{t'} \triangleq \{\rho_0 - t'\rho_1 > 0\}$. Indeed, (8) holds with equality for the value $t' = t$ appearing in (5), as $\operatorname{Tr}\bigl((\rho_0 - t\rho_1)\theta^0_t\bigr) = 0$ for any $0 \leq \theta^0_t \leq \bigl\{\rho_0 - t\rho_1 = 0\bigr\}$, tantamount to θ^0_t being in the null-space of $\rho_0 - t\rho_1$.

After some algebra, (8) yields

$$-\operatorname{Tr}\bigl(\rho_0 T_{\mathrm{NP}}\bigr) \geq -\operatorname{Tr}\bigl(\rho_0 P^{+}_{t'}\bigr) + t' \operatorname{Tr}\bigl(\rho_1 (P^{+}_{t'} - T_{\mathrm{NP}})\bigr). \tag{9}$$

Summing one to both sides of (9) and noting that $\alpha_\beta(\rho_0 \| \rho_1) = 1 - \operatorname{Tr}\bigl(\rho_0 T_{\mathrm{NP}}\bigr)$ and $\beta = \operatorname{Tr}\bigl(\rho_1 T_{\mathrm{NP}}\bigr)$, we obtain

$$\begin{aligned}
&\alpha_\beta(\rho_0 \| \rho_1) \\
&\geq \operatorname{Tr}\bigl(\rho_0 \{\rho_0 - t'\rho_1 \leq 0\}\bigr) + t' \operatorname{Tr}\bigl(\rho_1 P^{+}_{t'}\bigr) - t'\beta. \quad (10)
\end{aligned}$$

As (8) holds with equality for the value $t' = t$ appearing in (5), so it does (10) after optimization over the parameter $t' \geq 0$. Then, (6) follows. To obtain the lower bound (7), we fix $t' \geq 0$ and use that $\operatorname{Tr}\bigl(\rho_1 \bigl\{\rho_0 - t'\rho_1 > 0\bigr\}\bigr) \geq 0$. ■

B. Bayesian Multiple Hypothesis Testing

We consider now a hypothesis testing problem discriminating among M possible density operators acting on $\mathcal{H}$, where M is assumed to be finite. We consider the Bayesian setting, where the M alternatives $\tau_1, \dots, \tau_M$ occur with (classical) probabilities $p_1, \dots, p_M$, respectively.

A M-ary hypothesis test is a POVM $\mathcal{T} \triangleq \{\Pi_1, \Pi_2, \dots, \Pi_M\}$, $\sum \Pi_i = \mathbb{1}$. The test decides the alternative τ_i when the measurement with respect to $\mathcal{T}$ has outcome i. The probability that the test $\mathcal{T}$ decides τ_j when τ_i is the true underlying state is thus $\operatorname{Tr}\big(\tau_i \Pi_j\big)$ and the average error probability is

$$\epsilon(\mathcal{T}) \triangleq 1 - \sum_{i=1}^{M} p_i \operatorname{Tr}\left(\tau_i \Pi_i\right). \tag{11}$$

We define the minimum average error probability as

$$\epsilon^\star \triangleq \min_{\mathcal{T}} \epsilon(\mathcal{T}). \tag{12}$$

The test $\mathcal{T}$ minimizing (12) has no simple form in general.

Lemma 3 (Holevo-Yuen-Kennedy-Lax conditions): A test $\mathcal{T}^\star = \{\Pi_1^\star, \dots, \Pi_M^\star\}$ minimizes (12) if and only if, for each $m = 1, \dots, M$,

$$\big(\Lambda(\mathcal{T}^\star) - p_m\tau_m\big)\Pi_m^\star = \Pi_m^\star\big(\Lambda(\mathcal{T}^\star) - p_m\tau_m\big) = 0, \tag{13}$$

$$\Lambda(\mathcal{T}^\star) - p_m\tau_m \geq 0, \tag{14}$$

where

$$\Lambda(\mathcal{T}^\star) \triangleq \sum_{i=1}^{M} p_i\tau_i\Pi_i^\star = \sum_{i=1}^{M} p_i\Pi_i^\star\tau_i \tag{15}$$

is required to be self-adjoint[1].

Proof: This result follows from [33, Th. 4.1, Eq. (4.8)] or [34, Th. I] after simplifying the resulting optimality conditions. ■

We next show an alternative characterization of the minimum error probability $\epsilon^\star$ as a function of a binary hypothesis test with certain parameters. Let $\operatorname{diag}(\rho_1, \dots, \rho_M)$ denote the block-diagonal matrix with diagonal blocks $\rho_1, \dots, \rho_M$ and define

$$\boldsymbol{P} \triangleq \operatorname{diag}\big(p_1\tau_1, \dots, p_M\tau_M\big), \tag{16}$$

$$\boldsymbol{D}(\mu_0) \triangleq \operatorname{diag}\big(\tfrac{1}{M}\mu_0, \dots, \tfrac{1}{M}\mu_0\big), \tag{17}$$

where μ_0 is an arbitrary density operator acting on $\mathcal{H}$. Note that both $\boldsymbol{P}$ and $\boldsymbol{D}(\mu_0)$ are density operators themselves, since they are self-adjoint, positive semidefinite and have unit trace.

Theorem 1: The minimum error probability of a Bayesian M-ary test discriminating among states $\{\tau_1, \dots, \tau_M\}$ with prior probabilities $\{p_1, \dots, p_M\}$ satisfies

$$\epsilon^\star = \max_{\mu_0} \alpha_{\frac{1}{M}}\big(\boldsymbol{P} \,\|\, \boldsymbol{D}(\mu_0)\big), \tag{18}$$

where $\boldsymbol{P}$ and $\boldsymbol{D}(\cdot)$ are given in (16) and (17), respectively, and where the optimization is carried out over (unit-trace non-negative) density operators $\mu_0 \in \mathcal{D}(\mathcal{H})$.

Proof: This result can be proven extending the convex analysis approach from [28, Sec. III.B] to the quantum setting considered here. Nevertheless, in the following we provide an alternative proof based on a constructive approach that explicitly shows the connection between the optimal measurements for both the binary and M-ary discrimination problems.

For any $\mathcal{T} = \{\Pi_1, \Pi_2, \dots, \Pi_M\}$ let us define the binary test $T' \triangleq \operatorname{diag}\left(\Pi_1, \dots, \Pi_M\right)$. The error probabilities $\epsilon_{1|0}$ and $\epsilon_{0|1}$ of the test T' are given by

$$\epsilon_{1|0}(T') = 1 - \sum_{i=1}^{M} p_i \operatorname{Tr}\left(\tau_i \Pi_i\right) = \epsilon(\mathcal{T}), \tag{19}$$

[1] The operator $\Lambda(\mathcal{T})$ takes a role of the Lagrange multiplier associated to the constraint $\sum \Pi_i = \mathbb{1}$, which, involving self-adjoint operators requires Λ to be self-adjoint.

and

$$\epsilon_{0|1}(T') = \frac{1}{M}\sum_{i=1}^{M} \mathrm{Tr}\left(\mu_0 \Pi_i\right) \tag{20}$$

$$= \frac{1}{M}\,\mathrm{Tr}\left(\mu_0 \left(\textstyle\sum_{i=1}^{M} \Pi_i\right)\right) \tag{21}$$

$$= \frac{1}{M}\,\mathrm{Tr}\left(\mu_0\right) = \frac{1}{M}. \tag{22}$$

The (possibly suboptimal) test T' has thus $\epsilon_{1|0}(T') = \epsilon(\mathcal{T})$ where $\epsilon(\mathcal{T})$ is defined in (11) and $\epsilon_{0|1}(T') = \frac{1}{M}$. Therefore, using (4) and maximizing the resulting expression over μ_0, we obtain

$$\epsilon(\mathcal{T}) \geq \max_{\mu_0} \alpha_{\frac{1}{M}}\big(\boldsymbol{P} \,\|\, \boldsymbol{D}(\mu_0)\big). \tag{23}$$

To prove the theorem, it remains to show that the lower bound (23) holds with equality for $\mathcal{T} = \mathcal{T}^\star$ defined in Lemma 3. To this end, we next demonstrate that the optimality conditions in Lemma 1 and in Lemma 3 are equivalent for a certain choice of μ_0.

Let $\mathcal{T}^\star = \{\Pi_1^\star, \dots, \Pi_M^\star\}$ satisfy (13)-(14), i.e., $\mathcal{T}^\star$ corresponds to the optimal M-ary hypothesis test discriminating among states $\{\tau_1, \dots, \tau_M\}$ with prior probabilities $\{p_1, \dots, p_M\}$. We define

$$\mu_0^\star \triangleq \frac{1}{c_0^\star}\sum_{i=1}^{M} p_i \tau_i \Pi_i^\star = \frac{1}{c_0^\star}\Lambda(\mathcal{T}^\star), \tag{24}$$

where $c_0^\star$ is a normalizing constant such that $\mu_0^\star$ is unit trace. Lemma 1 shows that the test T_{NP} achieving (23) is associated to the non-negative eigenspace of the matrix

$$\boldsymbol{A} \triangleq \boldsymbol{P} - t\boldsymbol{D}(\mu_0), \tag{25}$$

which features a block-diagonal structure. According to (16)-(17), for the choice $\mu_0 = \mu_0^\star$, and $t = Mc_0^\star$, the m-th block-diagonal term in $\boldsymbol{A}$ is given by

$$A_m \triangleq p_m\tau_m - \tfrac{t}{M}\mu_0 = p_m\tau_m - \Lambda(\mathcal{T}^\star). \tag{26}$$

Given the block-diagonal structure of the matrix $\boldsymbol{A}$, it is enough to consider a test T_{NP} with block-diagonal structure. Then, we write $T_{\text{NP}} = \mathrm{diag}\left(T_1^{\text{NP}}, \dots, T_M^{\text{NP}}\right)$ and recall that the m-th block T_m^{NP} must lie in the non-negative eigenspace of the matrix A_m.

Using the optimality condition (14), it follows that the matrices A_m, $m = 1, \dots, M$, are negative semidefinite. Therefore, each block T_m^{NP} can only lie in the null eigenspace of A_m, for $m = 1, \dots, M$. Also, acording to the optimality condition (13), the operators $\Pi_m^\star$ precisely lie in the null eigenspace of A_m. As a result, the choice

$$T_{\text{NP}} = \mathrm{diag}\left(\Pi_1^\star, \dots, \Pi_M^\star\right) \tag{27}$$

satisfies the optimality conditions in Lemma 1. Moreover, since $\epsilon_{1|0}(T_{\text{NP}}) = \epsilon\big(\mathcal{T}^\star\big) = \epsilon^\star$ and $\epsilon_{0|1}(T_{\text{NP}}) = \frac{1}{M}$, Lemma 1 implies that (18) holds with equality for $\mu_0 = \mu_0^\star$. Given the bound in (23), other choices of μ_0 cannot improve the result, and Theorem 1 thus follows. ■

Example 1: Consider a hypothesis testing problem between the $M = 4$ (non-equiprobable) alternatives

$$\tau_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \tau_2 = \tfrac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$
$$\tau_3 = \tfrac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad \tau_4 = \tfrac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{28}$$

with prior probabilities $p_1 = \frac{2}{5}$ and $p_2 = p_3 = p_4 = \frac{1}{5}$. By solving (12), we obtain $\epsilon^\star = 7/15 = 0.4\overline{6}$ which is attained by the measurement $\mathcal{T}^\star = \{\Pi_1^\star, \dots, \Pi_4^\star\}$ with $\Pi_1^\star = \begin{bmatrix} 8/9 & 0 \\ 0 & 0 \end{bmatrix}$, $\Pi_2^\star = \begin{bmatrix} 1/18 & 1/6 \\ 1/6 & 1/2 \end{bmatrix}$, $\Pi_3^\star = \begin{bmatrix} 1/18 & -1/6 \\ -1/6 & 1/2 \end{bmatrix}$ and $\Pi_4^\star = 0$. Note that even when the dimension of the Hilbert space is 2, there are 3 active measurement operators. Since they are positive semidefinite and $\sum_{i=1}^4 \Pi_i^\star = \mathbb{1}$, the POVM is well defined. The POVM $\mathcal{T}^\star$ satisfies the optimality conditions from Lemma 3 and therefore $\epsilon^\star = 0.4\overline{6}$ is the lowest average error probability for this testing problem.

According to (24), the auxiliary state

$$\mu_0^\star = \frac{1}{c_0^\star}\sum_{i=1}^{4} p_i \tau_i \Pi_i^\star = \tfrac{1}{4}\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, \tag{29}$$

is optimal in Theorem 1. Indeed, it follows that[2]

$$\alpha_{\frac{1}{4}}\big(\boldsymbol{P} \,\|\, \boldsymbol{D}(\mu_0^\star)\big) = 0.4\overline{6} = \epsilon^\star. \tag{30}$$

[2] This computation can be done, e.g., by using (6) from Lemma 2 or by solving a semidefinite program.

Other choices of μ_0 yield a lower bound on the average error probability $\epsilon^\star$. For example, considering μ_0 the average state for this testing problem,

$$\mu_0 = \sum_{m=1}^{4} p_m \tau_m = \left[\begin{smallmatrix} 0.7 & 0 \\ 0 & 0.3 \end{smallmatrix}\right], \tag{31}$$

yields

$$\alpha_{\frac{1}{4}}\big(\boldsymbol{P} \,\|\, \boldsymbol{D}(\mu_0)\big) \approx 0.4571 < 0.4\overline{6} = \epsilon^\star. \tag{32}$$

Theorem 1 thus provides an alternative expression for the error probability $\epsilon^\star$ for the optimal choice of the auxiliary state, and a lower bound for other choices of μ_0. Combining Theorem 1 and Lemma 2, we obtain an alternative characterization for $\epsilon^\star$ based on information-spectrum measures.

Corollary 1: The minimum error probability of an M-ary test discriminating among states $\{\tau_1, \dots, \tau_M\}$ with prior classical probabilities $\{p_1, \dots, p_M\}$ satisfies

$$\epsilon^\star = \max_{\mu_0, t \geq 0} \left\{ \sum_{i=1}^{M} p_i \operatorname{Tr}\Big(\tau_i \big\{p_i\tau_i - t\mu_0 \leq 0\big\}\Big) - t \right\}. \tag{33}$$

where the optimization is carried out over (unit-trace non-negative) density operators μ_0 acting on $\mathcal{H}$, and over the scalar threshold $t \geq 0$.

Proof: Applying the lower bound (7) from Lemma 2 to the identity (18), and using the definitions of $\boldsymbol{P}$ in (16) and $\boldsymbol{D}(\cdot)$ in (17), it yields, for any μ_0, $t' \geq 0$,

$$\epsilon^\star \geq \sum_{i=1}^{M} p_i \operatorname{Tr}\Big(\tau_i \big\{p_i\tau_i - \tfrac{t'}{M}\mu_0 \leq 0\big\}\Big) - \tfrac{t'}{M}. \tag{34}$$

It remains to show that there exist μ_0 and $t' \geq 0$ such that (34) holds with equality. In particular, let us choose $\mu_0 = \mu_0^\star$ defined in (24), and $t' = Mc_0^\star$ where $c_0^\star = \sum_{i=1}^{M} p_i \operatorname{Tr}(\tau_i \Pi_i^\star)$ is the normalizing constant from (24).

For this choice of μ_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

$$\Big\{p_i\tau_i - \tfrac{t'}{M}\mu_0 \leq 0\Big\} = \big\{p_i\tau_i - \Lambda(\mathcal{T}^\star) \leq 0\big\} = \mathbb{1}, \tag{35}$$

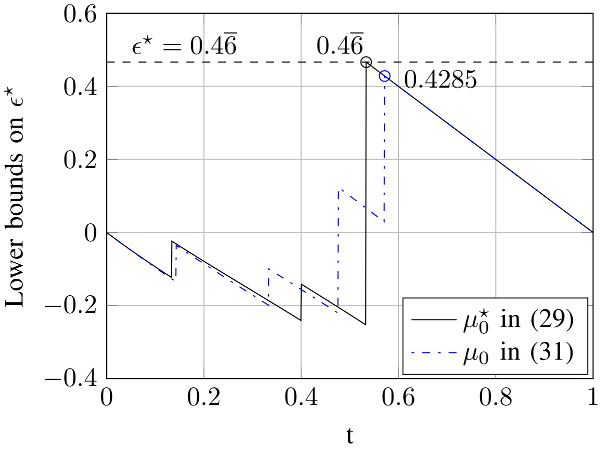

Fig. 1. Minimum error probability $\epsilon^\star$ (horizontal dashed line) for the hypothesis testing problem described in (28), compared with the lower bound that follows from (33) for fixed values of t and μ_0.

where the last identity follows from (14). The right-hand side of (34) thus becomes

$$\sum_{i=1}^{M} p_i \operatorname{Tr}(\tau_i) - \frac{t'}{M} = 1 - \frac{t'}{M}. \tag{36}$$

Using that $\frac{t'}{M} = c_0^\star = \sum_i p_i \operatorname{Tr}(\tau_i \Pi_i^\star) = 1 - \epsilon^\star$, the result follows. ■

For illustration, let us consider again the testing problem from Example 1, c.f. (28). Figure 1 shows the objective of (33) as a function of t for the auxiliary state $\mu_0 = \mu_0^\star$ in (29) and for the value of μ_0 given in (31). We can see that considering $\mu_0 = \mu_0^\star$, after maximization over t, yields the exact error probability $\epsilon^\star = 0.4\overline{6}$. In contrast, considering the value of μ_0 in (31), it yields a strict lower bound with a largest value of 0.4285, approximately. Comparing this value with (32), we conclude that by fixing a suboptimal auxiliary state μ_0, the right-hand side of (18) from Theorem 1 yields tighter bounds than (33) from Corollary 1. This could be expected as the expression in Corollary 1 follows from a weakening of (18).

We recall from the proofs of both Theorem 1 and Corollary 1 that a density operator μ_0 maximizing (18) and (33) is

$$\mu_0^\star = \frac{1}{c_0^\star} \sum_{i=1}^{M} p_i \tau_i \Pi_i^\star, \tag{37}$$

for some $\mathcal{T}^\star = \{\Pi_1^\star, \ldots, \Pi_M^\star\}$ satisfying the conditions in Lemma 3 and where $c_0^\star$ is a normalizing constant. Hence, the optimal M-ary hypothesis test $\mathcal{T}^\star$ characterizes the optimal μ_0. Conversely, the optimal μ_0 is precisely the Lagrange multiplier associated to the minimization in (12), after an appropriate re-scaling.

While the expressions in Theorem 1 and Corollary 1 are not easier to compute than the exact error probability, we show in the next section that they can be used to determine the tightness of several converse bounds in the context of reliable communication over classical-quantum channels.

III. Classical-Quantum Channels

We consider the channel coding problem of transmitting M equiprobable messages[3] over a one-shot classical-quantum channel $x \to W_x$, with $x \in \mathcal{X}$ and $W_x \in \mathcal{D}(\mathcal{H})$. 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.

With some abuse of notation, for a fix code, sometimes we shall write $W_m \triangleq W_{x_m}$. The minimum error probability for a code $\mathcal{C}$ is then given by

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) \triangleq \min_{\{\Pi_1, \ldots, \Pi_M\}} \left\{ 1 - \frac{1}{M} \sum_{m=1}^{M} \mathrm{Tr}\big(W_m \Pi_m\big) \right\}. \tag{38}$$

This problem corresponds precisely to the M-ary quantum hypothesis testing problem described in Section II-B. In contrast to the classical setting, in which (38) is minimized by the maximum likelihood decoder, the minimizer of (38) corresponds to any POVM satisfying the optimality conditions from Lemma 3.

A direct application of Theorem 1 yields an alternative expression for $\mathsf{P}_{\mathrm{e}}(\mathcal{C})$. Let P denote a (classical) distribution over the input alphabet $\mathcal{X}$ and define

$$PW \triangleq \sum\nolimits_{x \in \mathcal{X}} P(x) \Big(|x\rangle\langle x| \otimes W_x \Big), \tag{39}$$

$$P \otimes \mu \triangleq \Big(\sum\nolimits_{x \in \mathcal{X}} P(x) \, |x\rangle\langle x| \Big) \otimes \mu. \tag{40}$$

We denote by $P_{\mathcal{C}}$, the input distribution induced by the codebook $\mathcal{C}$, hence $P_{\mathcal{C}} W = \frac{1}{M} \sum_{x \in \mathcal{C}} \big(|x\rangle\langle x| \otimes W_x\big)$ and $P_{\mathcal{C}} \otimes \mu = \big(\frac{1}{M} \sum_{x \in \mathcal{C}} |x\rangle\langle x|\big) \otimes \mu$. Using the alternative expression introduced in Theorem 1 we obtain the following result.

Theorem 2 (Classical-quantum meta-converse bound): Let $\mathcal{C}$ be any codebook of cardinality M for a channel $x \to W_x$, with $x \in \mathcal{X}$ and $W_x \in \mathcal{D}(\mathcal{H})$. Then,

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = \sup_{\mu} \Big\{ \alpha_{\frac{1}{M}} \big(P_{\mathcal{C}} W \,\|\, P_{\mathcal{C}} \otimes \mu \big) \Big\} \tag{41}$$

$$\geq \inf_{P} \sup_{\mu} \Big\{ \alpha_{\frac{1}{M}} \big(PW \,\|\, P \otimes \mu \big) \Big\}. \tag{42}$$

where the maximization is over auxiliary states $\mu \in \mathcal{D}(\mathcal{H})$, and the minimization is over (classical) input distributions P.

Proof: The identity (41) is a direct application of (18) in Theorem 1. The relaxation (42) follows by minimizing (41) over all input distributions, not necessarily induced by a codebook. ■

The right-hand-side of (41) coincides with the finite block-length converse bound by Matthews and Wehner [22, Eq. (45)], particularized for a classical-quantum channel with an input state induced by the codebook $\mathcal{C}$. The lower bound (42) corresponds to [22, Eq. (46)] specialized to the classical-quantum setting (see also [21, Sec. 4.6] for a direct derivation for classical quantum channels). The classical analogous of (42) is usually referred to as meta-converse bound, since several converse

[3] While the results from Section II-B were derived for discrimination among non-equiprobable alternatives, in the remainder of this paper we consider the channel coding problem with equiprobable messages for clarity of exposition.

bounds in the literature can be derived from it. As it is the case in the classical-quantum setting, in the following we shall refer to this result as *meta-converse*.

Theorem 2 implies that the quantum generalization of the meta-converse bound proposed by Matthews and Wehner in [22, Eq. (45)] is tight for a fixed codebook $\mathcal{C}$. By fixing μ to be the state induced at the system output, the lower bound (42) recovers the converse bound [27, Th. 1], which is a rederivation of previous results in [20] (see [20, Remarks 10 and 15]). This bound is not tight in general since (i) the minimizing P does not need to coincide with the input state induced by the best codebook, and (ii) the choice of μ_0 in [27, Th. 1] does not maximize the resulting bound in general.

Using the characterization from Corollary 1, the error probability $\mathsf{P}_{\mathrm{e}}(\mathcal{C})$ can be equivalently written as

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = \max_{\mu_0, t\geq 0} \left\{ \frac{1}{M} \sum_{x\in\mathcal{C}} \mathrm{Tr}\Big(W_x \{ W_x - t\mu_0 \leq 0 \} \Big) - \frac{t}{M} \right\}. \tag{43}$$

The objective of the maximization in (43) coincides with the information-spectrum bound [20, Lemma 4]. Then, (43) shows that the Hayashi-Nagaoka lemma yields the exact error probability for a fixed code, after optimization over the free parameters μ_0, $t \geq 0$.

IV. Quasi-Perfect Codes

While the alternative expressions (41) and (43) derived in the previous section yield the exact error probability, they still depend on the codebook $\mathcal{C}$. To obtain a practical converse bound, these expressions need to be minimized over a family of codes or input distributions. One practical converse bound is the relaxation given in (42) which can be evaluated in several cases of interest. Since the converse bound (42) is a weakening of (41), it does not coincide with the exact error probability in general. Nevertheless, we next show that this is still the case for certain symmetric channels and the family of codes defined in this section.

A. Symmetric channels

Definition 1: We say that a classical-quantum channel $x \to W_x$, with $x \in \mathcal{X}$ and $W_x \in \mathcal{D}(\mathcal{H})$, is *symmetric* if

$$W_x = U_x \bar{W} U_x^\dagger \tag{44}$$

for every $x \in \mathcal{X}$, where $\bar{W} \in \mathcal{D}(\mathcal{H})$ does not depend on x and U_x is a unitary linear operator acting on $\mathcal{H}$ and parametrized by x.

This definition is related to that of covariant quantum channels (see, e.g., [35] or [36, Sec. 9.7]). Let us consider a quantum channel $\mathcal{N} : \mathcal{D}(\mathcal{H}_A) \to \mathcal{D}(\mathcal{H}_B)$, denoted by $\mathcal{N}_{A\to B}$, where $\mathcal{H}_A$ and $\mathcal{H}_B$ are the input and output Hilbert spaces of the channel, respectively. Let G be a compact group and $g \to V_g$, $g \to U_g$, $g \in G$, be continuous (projective) unitary representations of G in the input and output Hilbert spaces of the channel, respectively. The quantum channel $\mathcal{N}_{A\to B}$ is called covariant with respect to the representation if $\mathcal{N}_{A\to B}(V_g S V_g^\dagger) = U_g \mathcal{N}_{A\to B}(S) U_g^\dagger$ for every input state $S \in \mathcal{D}(\mathcal{H}_A)$. For a classical-quantum channel $x \to W_x$, $x \in \mathcal{X}$, $W_x \in \mathcal{D}(\mathcal{H}_B)$ denoted by $\mathcal{N}_{X\to B}$, we can define the orthogonal basis $|x\rangle\langle x|$. Then, letting V_x be such that $|x\rangle\langle x| = V_x |0\rangle\langle 0| V_x^\dagger$, it follows that $W_x = U_x \mathcal{N}_{X\to B}(|0\rangle\langle 0|) U_x^\dagger = U_x W_0 U_x^\dagger$. We conclude that any covariant quantum channel with classical (orthogonal) inputs also satisfies (44).

Definition 1 was also considered in [37], [38] with the additional assumption $U_x = U^x$, $U^{|\mathcal{X}|} = \mathbb{1}$, for some unitary U. For this family of symmetric channels, H.-C. Cheng *et al.* derived a sphere-packing lower bound for the optimal error probability in finite blocklengths. The proposed bound features the sphere-packing error exponent and the symmetry of the channel allows to tighten the prefactor order to polynomial.

In Definition 1, however, we do not impose further structure on the unitary representations U_x. The results in this section hold for this family of symmetric channels provided that there exist codes satisfying certain properties, as it will be shown next.

B. Quasi-perfect codes

For any operator $\mu \in \mathcal{D}(\mathcal{H})$ and parameter $t \in \mathbb{R}$, we define

$$\mathcal{E}_x(t,\mu) \triangleq \bigl\{W_x - t\mu \geq 0\bigr\}, \tag{45}$$

For a symmetric channel $x \to W_x = U_x \bar{W} U_x^\dagger$, $x \in \mathcal{X}$, we consider the set of auxiliary operators $\mu \in \mathcal{D}(\mathcal{H})$ such that they commute with the unitary transformations U_x, $x \in \mathcal{X}$. More precisely, for a symmetric channel $x \to W_x$, we define

$$\mathcal{U}_W \triangleq \bigl\{\mu \in \mathcal{D}(\mathcal{H}) \mid U_x\mu = \mu U_x\bigr\}. \tag{46}$$

Then, for any symmetric channel $x \to W_x$, $x \in \mathcal{X}$, $W_x \in \mathcal{D}(\mathcal{H})$, and $\mu \in \mathcal{U}_W$, it follows that

$$\begin{aligned}\mathcal{E}_x(t,\mu) &= \bigl\{U_x\bar{W}U_x^\dagger - t\mu \geq 0\bigr\} &(47)\\ &= U_x\bigl\{\bar{W} - tU_x^\dagger\mu U_x \geq 0\bigr\}U_x^\dagger &(48)\\ &= U_x\bar{\mathcal{E}}(t,\mu)U_x^\dagger, &(49)\end{aligned}$$

where in the last step we used the fact that $\mu U_x = U_x\mu$ and defined $\bar{\mathcal{E}}(t,\mu) \triangleq \bigl\{\bar{W} - t\mu \geq 0\bigr\}$, which does not depend on $x \in \mathcal{X}$.

Similarly to (45), we define

$$\mathcal{E}_x^\circ(t,\mu) \triangleq \bigl\{W_x - t\mu = 0\bigr\}. \tag{50}$$

$$\mathcal{E}_x^\bullet(t,\mu) \triangleq \bigl\{W_x - t\mu > 0\bigr\}, \tag{51}$$

and,

$$F_x^\bullet(t,\mu) \triangleq \operatorname{Tr}\bigl(W_x\mathcal{E}_x^\bullet(t,\mu)\bigr), \tag{52}$$

$$G_x^\bullet(t,\mu) \triangleq \operatorname{Tr}\bigl(\mu\,\mathcal{E}_x^\bullet(t,\mu)\bigr), \tag{53}$$

where, $F_\bullet(\cdot) \triangleq F_x^\bullet(\cdot)$, $G_\bullet(\cdot) \triangleq G_x^\bullet(\cdot)$, independent of $x \in \mathcal{X}$ for symmetric channels.

Definition 2: A code $\mathcal{C}$ is *perfect* for a classical-quantum channel $x \to W_x$, if there exist a scalar t and a state $\mu \in \mathcal{D}(\mathcal{H})$ such that the projectors $\bigl\{\mathcal{E}_x(t,\mu)\bigr\}_{x\in\mathcal{C}}$ are orthogonal to each other and $\sum_{x\in\mathcal{C}}\mathcal{E}_x(t,\mu) = \mathbb{1}$. More generally, a code is *quasi-perfect* if there exist t and $\mu \in \mathcal{D}(\mathcal{H})$ such that the projectors $\bigl\{\mathcal{E}_x^\bullet(t,\mu)\bigr\}_{x\in\mathcal{C}}$ are orthogonal to each other, and for $I_\bullet \triangleq \sum_{x\in\mathcal{C}}\mathcal{E}_x^\bullet(t,\mu)$, $I_\circ \triangleq \mathbb{1} - I_\bullet$, it holds that $\sum_{x\in\mathcal{C}}\mathcal{E}_x^\circ(t,\mu) = cI_\circ$ where $c \in \mathbb{R}$, $c > 0$ is a normalizing constant that depends on the code $\mathcal{C}$.

Example 2: Let us consider the pure-state channel $x \to W_x = |\varphi_x\rangle\langle\varphi_x|$ acting on a n-dimensional Hilbert space $\mathcal{H}$. This channel is symmetric, as any pure-state W_x can be constructed via unitary transformations from an arbitrary pure-state $\bar{W} = |\psi\rangle\langle\psi|$. If we do not impose further restrictions on the output of the system, i.e., it can be an arbitrary pure state $W_x = U_x\bar{W}U_x^\dagger$, then the only state μ which commutes with all unitary linear operator U_x, $x \in \mathcal{X}$, is the maximally mixed state $\mu = \frac{1}{n}\mathbb{1}$. According to Definition 2, a code $\mathcal{C}$ with $M = n$ orthogonal pure states is perfect for this channel with parameters $t = n$ and $\mu = \frac{1}{n}\mathbb{1}$, since the projectors $\mathcal{E}_x\bigl(n, \frac{1}{n}\mathbb{1}\bigr) = \bigl\{|\varphi_x\rangle\langle\varphi_x| - \mathbb{1} \geq 0\bigr\} = |\varphi_x\rangle\langle\varphi_x|$ are orthogonal for $x \in \mathcal{C}$, and they form a basis for $\mathcal{H}$. Note that this particular case can be reduced to a classical problem, since the channel outputs commute with each other. Similarly, a code with $M \geq n$ is quasi-perfect for this channel with parameters $t = n$ and $\mu = \frac{1}{n}\mathbb{1}$ provided that $\sum_{x\in\mathcal{C}}|\varphi_x\rangle\langle\varphi_x| = c\mathbb{1}$ with $c = \frac{M}{n}$. A family of codes fulfilling this properties will be studied in detail in Section V. For these quasi-perfect codes, the interiors $\mathcal{E}_x^\bullet\bigl(n, \frac{1}{n}\mathbb{1}\bigr) = \bigl\{|\varphi_x\rangle\langle\varphi_x| - \mathbb{1} > 0\bigr\} = 0$, hence they are orthogonal to each other, and the channel outputs don't commute with each other. For $M < n$, the codes for this channel and the auxiliary state $\mu = \frac{1}{n}\mathbb{1}$ are neither perfect nor quasi-perfect.

To avoid ambiguities, we shall denote by $\bar{t}$ the smallest value of t such that the projectors $\{\mathcal{E}_x^\bullet(t,\mu)\}_{x\in\mathcal{C}}$ are orthogonal to each other for a certain code $\mathcal{C}$. We shall refer to $\bar{t}$ as the *packing radius* of the code $\mathcal{C}$ with respect to state μ.

The next result provides an alternative expression for the error probability of perfect and quasi-perfect codes.

Theorem 3 (Error probability of quasi-perfect codes): Let the channel $x \to W_x$, $x \in \mathcal{X}$, $W_x \in \mathcal{D}(\mathcal{H})$, and $\mu \in \mathcal{U}_W$ be symmetric, and let $\mathcal{C}$ be perfect or quasi-perfect with parameters t and μ. Then,

$$\mathsf{P}_\text{e}(\mathcal{C}) = 1 - F_\bullet(t,\mu) + t\big(G_\bullet(t,\mu) - |\mathcal{C}|^{-1}\big), \tag{54}$$

where $|\mathcal{C}|$ denotes the cardinality of the codebook $\mathcal{C}$.

Proof: Let $\mathcal{C} = \{x_1,\dots,x_M\}$ be an arbitrary code for the (symmetric) channel $x \to W_x$. Let $\bar{t}$ be the packing radius of $\mathcal{C}$ with respect to the auxiliary state μ.

We define the orthogonal basis $\{\bar{E}(i)\}$ associated to the eigenspace of $\{\bar{W} - \bar{t}\mu \geq 0\}$ such that

$$\bar{\mathcal{E}}^\bullet(\bar{t},\mu) = \sum_{i\in\mathcal{I}^\bullet} \bar{E}(i), \tag{55}$$

$$\mathcal{E}_x^\bullet(\bar{t},\mu) = U_x\bar{\mathcal{E}}^\bullet(\bar{t},\mu)U_x^\dagger \tag{56}$$

$$= \sum_{i\in\mathcal{I}^\bullet} U_x\bar{E}(i)U_x^\dagger = \sum_{i\in\mathcal{I}^\bullet} E_x(i), \tag{57}$$

where we let $E_x(i) \triangleq U_x\bar{E}(i)U_x^\dagger$. Here, $\mathcal{I}^\bullet$ denotes the set of basis indexes associated to the strictly positive eigenvalues. Note that the projectors $E_x(i)$ are orthogonal to $E_{x'}(i)$ for $x \neq x'$, $i \in \mathcal{I}^\bullet$ since the projectors $\{\mathcal{E}_x^\bullet(\bar{t},\mu)\}$ for $x \in \mathcal{C}$ are orthogonal to each other. Similarly, we also write

$$\bar{\mathcal{E}}^\circ(\bar{t},\mu) = \sum_{i\in\mathcal{I}^\circ} \bar{E}(i), \tag{58}$$

$$\mathcal{E}_x^\circ(\bar{t},\mu) = U_x\bar{\mathcal{E}}^\circ(\bar{t},\mu)U_x^\dagger \tag{59}$$

$$= \sum_{i\in\mathcal{I}^\circ} U_x\bar{E}(i)U_x^\dagger = \sum_{i\in\mathcal{I}^\circ} E_x(i). \tag{60}$$

where $\mathcal{I}^\circ$ denotes the set of basis indexes associated to the zero eigenvalues. In this later case however, there is no orthogonality condition between the projectors $E_x(i)$ for $i \in \mathcal{I}^\circ$ for different codewords $x \in \mathcal{C}$. Now define $d_\bullet \triangleq M|\mathcal{I}^\bullet|$ and $d_\circ \triangleq n - d_\bullet$, where $n = \dim(\mathcal{H})$. The code specific constant associated with a quasi-perfect code $\mathcal{C}$ is $c \triangleq \frac{M|\mathcal{I}^\circ|}{d_\circ}$.

We consider the decoder $\mathcal{T} = \{\Pi_1,\dots,\Pi_M\}$ with projectors

$$\Pi_m = \mathcal{E}_{x_m}^\bullet(\bar{t},\mu) + \frac{1}{c}\mathcal{E}_{x_m}^\circ(\bar{t},\mu) \tag{61}$$

$$= U_{x_m}\bar{\mathcal{E}}^\bullet(\bar{t},\mu)U_{x_m}^\dagger + \frac{1}{c}U_{x_m}\bar{\mathcal{E}}^\circ(\bar{t},\mu)U_{x_m}^\dagger \tag{62}$$

$$= U_{x_m}\bar{\Pi}U_{x_m}^\dagger, \quad m = 1,\dots,M, \tag{63}$$

where

$$\bar{\Pi} = \bar{\mathcal{E}}^\bullet(\bar{t},\mu) + \frac{1}{c}\bar{\mathcal{E}}^\circ(\bar{t},\mu) \tag{64}$$

Note that this definition implies

$$\sum_{m=1}^M \Pi_m = I_\bullet + I_\circ = \mathbb{1}, \tag{65}$$

as required.

We next show that this decoder satisfies the Holevo-Yuen-Kennedy-Lax conditions from Lemma 3 and it thus minimizes (38). According to (15), let

$$\Lambda(\mathcal{T}) = \frac{1}{M}\sum_{\ell=1}^M W_\ell\Pi_\ell. \tag{66}$$

To verify the condition (13), we write

$$\begin{aligned}&\left(\Lambda(\mathcal{T}) - \frac{1}{M}W_m\right)\Pi_m \\ &= \left(\frac{1}{M}\sum_{\ell\neq m} W_\ell\Pi_\ell\right)\Pi_m + \frac{1}{M}W_m\Pi_m(\Pi_m - I).\end{aligned} \tag{67}$$

Let us consider the first term in (67) only. Using (61), we decompose this term as

$$\begin{aligned}&\left(\frac{1}{M}\sum_{\ell\neq m} W_\ell\Pi_\ell\right)\Pi_m \\ &= \frac{1}{M}\left(\sum_{\ell\neq m} W_\ell\mathcal{E}_\ell^\bullet(\bar{t},\mu) + \frac{1}{c}\sum_{\ell\neq m} W_\ell\mathcal{E}_\ell^\circ(\bar{t},\mu)\right) \\ &\qquad \times \left(\mathcal{E}_m^\bullet(\bar{t},\mu) + \frac{1}{c}\mathcal{E}_m^\circ(\bar{t},\mu)\right).\end{aligned} \tag{68}$$

We note that $\mathcal{E}_\ell^\bullet(\bar{t},\mu)$ and $\mathcal{E}_m^\bullet(\bar{t},\mu)$ are mutually orthogonal for $\ell \neq m$; hence $\mathcal{E}_\ell^\bullet(\bar{t},\mu)\mathcal{E}_m^\bullet(\bar{t},\mu) = 0$. Using that the subspaces $I_\circ$ and $I_\bullet$ are mutually orthogonal, we conclude that $\mathcal{E}_\ell^\circ(\bar{t},\mu)\mathcal{E}_m^\bullet(\bar{t},\mu) = 0$, and $\mathcal{E}_\ell^\bullet(\bar{t},\mu)\mathcal{E}_m^\circ(\bar{t},\mu) = 0$. Then, (68) becomes

$$\begin{aligned}&\left(\frac{1}{M}\sum_{\ell\neq m} W_\ell \Pi_\ell\right)\Pi_m \\ &\quad = \frac{1}{Mc^2}\left(\sum_{\ell\neq m} W_\ell \mathcal{E}_\ell^\circ(\bar{t},\mu)\right)\mathcal{E}_m^\circ(\bar{t},\mu), \qquad (69)\end{aligned}$$

where in the last step we used that $\mathcal{E}_m^\bullet(\bar{t},\mu)$ and $\mathcal{E}_\ell^\bullet(\bar{t},\mu)$ are mutually orthogonal for $\ell \neq m$; and so they are $\mathcal{E}_m^\bullet(\bar{t},\mu)$ and $\mathcal{E}_\ell^\circ(\bar{t},\mu)$.

Since $\mathcal{E}_\ell^\circ(\bar{t},\mu)$ is the projector associated to the nullspace of $W_\ell - \bar{t}\mu$, it holds that

$$W_\ell \mathcal{E}_\ell^\circ(\bar{t},\mu) = \bar{t}\mu \mathcal{E}_\ell^\circ(\bar{t},\mu). \qquad (70)$$

Then, proceeding from (69), we obtain

$$\begin{aligned}&\left(\frac{1}{M}\sum_{\ell\neq m} W_\ell \Pi_\ell\right)\Pi_m \\ &\quad = \frac{1}{Mc^2}\left(\sum_{\ell\neq m} \bar{t}\mu \mathcal{E}_\ell^\circ(\bar{t},\mu)\right)\mathcal{E}_m^\circ(\bar{t},\mu) \qquad (71)\\ &\quad = \frac{1}{Mc^2}\bar{t}\mu\left(\sum_{\ell\neq m} \mathcal{E}_\ell^\circ(\bar{t},\mu)\right)\mathcal{E}_m^\circ(\bar{t},\mu) \qquad (72)\\ &\quad = \frac{1}{Mc^2}\bar{t}\mu\big(cI_\circ - \mathcal{E}_m^\circ(\bar{t},\mu)\big)\mathcal{E}_m^\circ(\bar{t},\mu) \qquad (73)\\ &\quad = \frac{c-1}{Mc^2}\bar{t}\mu\mathcal{E}_m^\circ(\bar{t},\mu), \qquad (74)\end{aligned}$$

where we used first that $\sum_{x\in\mathcal{C}} \mathcal{E}_x^\circ(t,\mu) = cI_\circ$, and then $I_\circ \mathcal{E}_m^\circ(\bar{t},\mu) = \mathcal{E}_m^\circ(\bar{t},\mu)\mathcal{E}_m^\circ(\bar{t},\mu) = \mathcal{E}_m^\circ(\bar{t},\mu)$.

For the second term in (67), we can show that

$$\begin{aligned}\frac{1}{M}W_m\Pi_m(\Pi_m - I) &= \frac{1-c}{Mc^2}W_m\mathcal{E}_m^\circ(\bar{t},\mu) \qquad (75)\\ &= \frac{1-c}{Mc^2}\bar{t}\mu\mathcal{E}_m^\circ(\bar{t},\mu) \qquad (76)\end{aligned}$$

Combining (74) and (76) with (67), we prove that $\left(\Lambda(\mathcal{T}) - \frac{1}{M}W_m\right)\Pi_m = 0$. Following analogous steps, we show that $\Pi_m\left(\Lambda(\mathcal{T}) - \frac{1}{M}W_m\right) = 0$ and hence the decoder satisfies the optimality condition (13).

Using similar arguments to those in the previous derivation, we write

$$\begin{aligned}\Lambda(\mathcal{T}) - \frac{1}{M}W_m &= \frac{1}{M}\sum_{\ell=1}^{M} W_\ell\Pi_\ell - \frac{1}{M}W_m \qquad (77)\\ &= \frac{1}{M}\sum_{\ell\neq m} W_\ell\Pi_\ell + \frac{1}{M}W_m(\Pi_m - I). \qquad (78)\end{aligned}$$

For the first term in (78), we have that

$$\begin{aligned}&\frac{1}{M}\sum_{\ell\neq m} W_\ell\Pi_\ell \\ &\quad = \frac{1}{M}\sum_{\ell\neq m} W_\ell\mathcal{E}_\ell^\bullet(\bar{t},\mu) + \frac{1}{Mc}\bar{t}\mu\sum_{\ell\neq m}\mathcal{E}_\ell^\circ(\bar{t},\mu). \qquad (79)\end{aligned}$$

For the second term, using (61), we obtain

$$\begin{aligned}&\frac{1}{M}W_m(\Pi_m - I) \\ &\quad = \frac{1}{M}W_m\mathcal{E}_m^\bullet(\bar{t},\mu) + \frac{1}{Mc}W_m\mathcal{E}_m^\circ(\bar{t},\mu) - \frac{1}{M}W_m \qquad (80)\\ &\quad = \frac{1}{Mc}\bar{t}\mu\mathcal{E}_m^\circ(\bar{t},\mu) - \frac{1}{M}W_m I_\circ - \frac{1}{M}W_m\sum_{\ell\neq m}\mathcal{E}_\ell^\bullet(\bar{t},\mu), \qquad (81)\end{aligned}$$

where in (81) we used the identities $I_\bullet = \sum_\ell \mathcal{E}_\ell^\bullet(\bar{t},\mu)$ and $W_m = W_m I_\bullet + W_m I_\circ$.

We now use that according to the definition of quasi-perfect codes, $I_\circ = \frac{1}{c}\sum_\ell \mathcal{E}_\ell^\circ(\bar{t},\mu)$. Then, substituting (79) and (81) in (78), grouping terms, we obtain

$$\begin{aligned}&\Lambda(\mathcal{T}) - \frac{1}{M}W_m \\ &\quad = \frac{1}{M}\big(\bar{t}\mu - W_m\big)I_\circ + \frac{1}{M}\sum_{\ell\neq m}\big(W_\ell - W_m\big)\mathcal{E}_\ell^\bullet(\bar{t},\mu). \qquad (82)\end{aligned}$$

The eigenvectors of $W_m - \bar{t}\mu$ corresponding to positive eigenvalues belong to the subspace spanned by $I_\bullet$. Therefore $\frac{1}{M}\big(\bar{t}\mu - W_m\big)I_\circ \geq 0$.

On the other hand, we have that

$$\frac{1}{M}\sum_{\ell\neq m}(W_\ell - W_m)\mathcal{E}_\ell^\bullet(\bar{t},\mu) \tag{83}$$

$$\geq \frac{1}{M}\sum_{\ell\neq m}(\bar{t}\mu - W_m)\mathcal{E}_\ell^\bullet(\bar{t},\mu) \tag{84}$$

$$\geq \frac{1}{M}\sum_{\ell\neq m}(\bar{t}\mu - \bar{t}\mu)\mathcal{E}_\ell^\bullet(\bar{t},\mu) = 0 \tag{85}$$

where in (84) we used $W_\ell\mathcal{E}_\ell^\bullet(\bar{t},\mu) \geq \bar{t}\mu\mathcal{E}_\ell^\bullet(\bar{t},\mu)$, and (85) follows since $W_m\mathcal{E}_{\ell\neq m}^\bullet(\bar{t},\mu) \leq \bar{t}\mu\mathcal{E}_{\ell\neq m}^\bullet(\bar{t},\mu)$ which holds since $\mathcal{E}_{\ell\neq m}^\bullet(\bar{t},\mu)$ and $\mathcal{E}_m^\bullet(\bar{t},\mu)$ being orthogonal implies that $\mathcal{E}_{\ell\neq m}^\bullet(\bar{t},\mu)$ must belong to the negative eigenspace of $W_m - \bar{t}\mu$. We conclude that $\Lambda(\mathcal{T}) - \frac{1}{M}W_m \geq 0$.

As the decoder $\mathcal{T} = \{\Pi_1,\ldots,\Pi_M\}$ satisfies the optimality conditions from Lemma 3, it minimizes (38). Then, combining (15) and (38), we obtain that the error probability of this code is

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = 1 - \operatorname{Tr}\big(\Lambda(\mathcal{T}^\star)\big) \tag{86}$$

$$= 1 - \operatorname{Tr}\Bigg(\frac{1}{M}\sum_{m=1}^{M} W_m\mathcal{E}_m^\bullet(\bar{t},\mu) + \frac{1}{Mc}\sum_{m=1}^{M} W_m\mathcal{E}_m^\circ(\bar{t},\mu)\Bigg). \tag{87}$$

Using (52), the error probability $\mathsf{P}_{\mathrm{e}}(\mathcal{C})$ becomes

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = 1 - \frac{1}{M}\sum_{m=1}^{M} F_{x_m}^\bullet(\bar{t},\mu) - \frac{1}{Mc}\operatorname{Tr}\left(\sum_{m=1}^{M}\bar{t}\mu\mathcal{E}_m^\circ(\bar{t},\mu)\right) \tag{88}$$

$$= 1 - \frac{1}{M}\sum_{m=1}^{M} F_{x_m}^\bullet(\bar{t},\mu) - \frac{\bar{t}}{M}\operatorname{Tr}\big(\mu I_\circ\big) \tag{89}$$

Now, noting that $\mu = \mu(I_\circ + I_\bullet)$, we write

$$\operatorname{Tr}\big(\mu I_\circ\big) = 1 - \operatorname{Tr}\left(\mu\sum_{m=1}^{M}\mathcal{E}_m^\bullet(\bar{t},\mu)\right) \tag{90}$$

$$= 1 - \sum_{m=1}^{M} G_{x_m}^\bullet(\bar{t},\mu) \tag{91}$$

where the second equality follows from (53). Then, substituting (91) in (89), using the fact that, for symmetric channels, $F_x^\bullet(\bar{t},\mu) = F_\bullet(\bar{t},\mu)$ and $G_x^\bullet(\bar{t},\mu) = G_\bullet(\bar{t},\mu)$, and noting that $M = |\mathcal{C}|$, we obtain

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = 1 - F_\bullet(\bar{t},\mu) + \bar{t}(G_\bullet(\bar{t},\mu) - |\mathcal{C}|^{-1}) \tag{92}$$

■

Example 3: For the pure-state channel $W_x = |\varphi_x\rangle\langle\varphi_x|$ introduced in Example 2 above, let $t = n$ be the number of dimensions of Hilber space $\mathcal{H}$ and $\mu = \frac{1}{n}\mathbb{1}$ be the maximally mixed state. Then, $F_\bullet(t,\mu) = G_\bullet(t,\mu) = 0$ and using (54) we obtain that for any perfect or quasi-perfect code $\mathcal{C}$ with cardinality $|\mathcal{C}| = M$, the error probability is given by

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = 1 - \frac{n}{M}. \tag{93}$$

Note that $\mathsf{P}_{\mathrm{e}}(\mathcal{C})$ is the average error probability of the code and that it does not describe how the errors are distributed among the different messages. It could happen that some of the projectors are inactive and the corresponding messages always yield an error, and that some messages may be decoded with no error.

We next show that perfect and quasi-perfect codes attain the converse bound (42) with equality. This result is based on the following auxiliary lemma.

Lemma 4: Let $\rho_0 = PW$ and $\rho_1 = P\otimes\mu$ be defined in (39) and (40), respectively. Then, the optimal trade-off (4) for a hypothesis test between ρ_0 and ρ_1 satisfies

$$\alpha_\beta\big(PW \,\|\, P\otimes\mu\big) = \inf_{\substack{\{\beta'_x\}:\\ \beta=\sum_x P(x)\beta'_x}} \sum_{x\in\mathcal{X}} P(x)\alpha_{\beta'_x}\big(W_x \,\|\, \mu\big). \tag{94}$$

Proof: We consider Lemma 2 with $\rho_0 \leftarrow PW$ and $\rho_1 \leftarrow P\otimes\mu$. Then, using the block-diagonal structure of PW and $P\otimes\mu$, the identity (6) yields

$$\begin{aligned}&\alpha_{\beta}\bigl(PW \,\|\, P\otimes\mu\bigr)\\&=\sup_{t\geq 0}\Biggl\{\sum_{x\in\mathcal{X}} P(x)\operatorname{Tr}\bigl(W_x\{W_x-t\mu\leq 0\}\bigr)\\&\qquad+t\Biggl(\sum_{x\in\mathcal{X}} P(x)\operatorname{Tr}\bigl(\mu\{W_x-t\mu>0\}\bigr)-\beta\Biggr)\Biggr\}\end{aligned}\tag{95}$$

$$\begin{aligned}&=\sup_{t\geq 0}\Biggl\{\sum_{x\in\mathcal{X}} P(x)\Bigl(\operatorname{Tr}\bigl(W_x\{W_x-t\mu\leq 0\}\bigr)\\&\qquad+t\Bigl(\operatorname{Tr}\bigl(\mu\{W_x-t\mu>0\}\bigr)-\beta'_x\Bigr)\Bigr)\Biggr\}\end{aligned}\tag{96}$$

for any $\{\beta'_x\}$, $x\in\mathcal{X}$, such that $\sum_x P(x)\beta'_x=\beta$.

We relax the optimization (96) by letting the parameter t be different for each x. Then, we obtain the following upper bound on $\alpha_{\beta}\bigl(PW \,\|\, P\otimes\mu\bigr)$,

$$\begin{aligned}&\alpha_{\beta}\bigl(PW \,\|\, P\otimes\mu\bigr)\\&\leq\sum_{x\in\mathcal{X}} P(x)\sup_{t_x\geq 0}\Bigl\{\operatorname{Tr}\bigl(W_x\{W_x-t_x\mu\leq 0\}\bigr)\\&\qquad+t_x\Bigl(\operatorname{Tr}\bigl(\mu\{W_x-t_x\mu>0\}\bigr)-\beta'_x\Bigr)\Bigr\}\end{aligned}\tag{97}$$

$$=\sum_{x\in\mathcal{X}} P(x)\alpha_{\beta'_x}\bigl(W_x \,\|\, \mu\bigr),\tag{98}$$

where in the last step we applied the identity (6) from Lemma 2 with $\rho_0\leftarrow W_x$ and $\rho_1\leftarrow\mu$. The bound (97)-(98) holds for any $\{\beta'_x\}$, $x\in\mathcal{X}$, such that $\sum_x P(x)\beta'_x=\beta$. Then, to prove (94) it suffices to show that there exist $\{\beta'_x\}$ satisfying $\sum_x P(x)\beta'_x=\beta$ and such that (97) holds with equality.

Indeed, the value of t maximizing (96) induces the Neyman-Pearson test (5), which due to the block-diagonal structure of the problem, can be decomposed into the sub-tests

$$T'_x=\bigl\{W_x-t\mu>0\bigr\}+\theta^0_x.\tag{99}$$

Each of these subtests induces a type-I error probability α'_x and type-II error probability β'_x, which, according to the NP lemma, satisfy $\sum_x P(x)\alpha'_x=\alpha_{\beta}\bigl(PW \,\|\, P\otimes\mu\bigr)$ and $\sum_x P(x)\beta'_x=\beta$. It follows that, for this choice of $\{\beta'_x\}$, the optimization in (97) yields $t_x=t$ (as the t parameter in the NP subtests is unique), and therefore (97) holds with equality. The result thus follows. ■

Lemma 4 asserts that, for a binary hypothesis test between classical-quantum distributions, it is possible to express the optimal type-I error probability as a convex combination of that of disjoint sub-tests provided that the type-II error is optimally distributed among them. The next result follows from combining Theorem 3 and Lemmas 2 and 4.

Theorem 4 (Quasi-perfect codes attain the meta-converse): Let the channel $x\to W_x$ be symmetric and let $\mathcal{C}$ be perfect or quasi-perfect with parameters t and $\mu\in\mathcal{U}_W$. Then, for $M=|\mathcal{C}|$,

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C})=\inf_{P}\sup_{\mu'}\alpha_{\frac{1}{M}}\bigl(PW \,\|\, P\otimes\mu'\bigr).\tag{100}$$

Proof: According to (42) in Theorem 2, the right-hand side of (100) is a lower bound to the error probability of any code. Then, to prove (100), it suffices to show that the error probability of $\mathcal{C}$ coincides with this lower bound. Using Lemma 4, fixing the auxiliary state μ to that from Definition 2, we obtain

$$\begin{aligned}&\inf_{P}\sup_{\mu'}\alpha_{\frac{1}{M}}\bigl(PW \,\|\, P\otimes\mu'\bigr)\\&\qquad\geq\inf_{\substack{\{P(x),\beta_x\}:\\ \sum_x P(x)\beta_x=\frac{1}{M}}}\sum_{x\in\mathcal{X}} P(x)\alpha_{\beta_x}\bigl(W_x \,\|\, \mu\bigr).\end{aligned}\tag{101}$$

Now, using (6) from Lemma 2, letting $t'=t$, and using the definitions of $F^{\bullet}_x(t,\mu)$ and $G^{\bullet}_x(t,\mu)$, it follows that

$$\begin{aligned}&\alpha_{\beta_x}\bigl(W_x \,\|\, \mu\bigr)\\&\qquad\geq 1-F^{\bullet}_x(t,\mu)+t\bigl(G^{\bullet}_x(t,\mu)-\beta_x\bigr)\end{aligned}\tag{102}$$

$$\qquad=1-F_{\bullet}(t,\mu)+t\bigl(G_{\bullet}(t,\mu)-\beta_x\bigr),\tag{103}$$

where in the last step we used that for symmetric channels, $F_{\bullet}(t,\mu)=F^{\bullet}_x(t,\mu)$ and $G_{\bullet}(t,\mu)=G^{\bullet}_x(t,\mu)$.

Then, using (103) in (101), we obtain

$$\begin{aligned}\inf_{P}\sup_{\mu'}\alpha_{\frac{1}{M}}\big(PW \,\|\, P\otimes\mu'\big)&\\ \geq \inf_{\substack{\{P(x),\beta_x\}:\\ \sum_x P(x)\beta_x=\frac{1}{M}}} &\bigg(1-F_{\bullet}(t,\mu)\\ &+t\Big(G_{\bullet}(t,\mu)-\textstyle\sum_x P(x)\beta_x\Big)\bigg) \quad (104)\\ =1-F_{\bullet}(t,\mu)+t\bigg(G_{\bullet}(t,\mu)-\frac{1}{M}\bigg)& \quad (105)\end{aligned}$$

where in the second step we used that the constraint $\sum_x P(x)\beta_x = \frac{1}{M}$ implies that the objective does not depend on the optimization variables.

The right-hand side of (105) coincides with the error probability of the quasi-perfect codes given in (54). Then, using this observation and (42) we conclude that, whenever $\mathcal{C}$ is perfect or quasi-perfect,

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) \;\leq\; \inf_{P}\sup_{\mu'}\alpha_{\frac{1}{M}}\big(PW \,\|\, P\otimes\mu'\big) \;\leq\; \mathsf{P}_{\mathrm{e}}(\mathcal{C}), \quad (106)$$

and the converse bound (100) must hold with equality. ■

The computation of the right-hand side of (100) can be simplified by exploiting the symmetry of the channel. Indeed, for a symmetric channel $x \to W_x$ and $\mu \in \mathcal{D}(\mathcal{H})$ being invariant under this symmetry, namely $\mu \in \mathcal{U}_W$, it follows that (see [22, Sec. V.E])

$$\inf_{P}\sup_{\mu'}\alpha_{\frac{1}{M}}\big(PW \,\|\, P\otimes\mu'\big) = \alpha_{\frac{1}{M}}\big(W_x \,\|\, \mu\big). \quad (107)$$

Therefore, we obtain the following corollary to Theorem 4; when the channel $x \to W_x$ is symmetric and $\mathcal{C}$ is perfect or quasi-perfect with parameters t and $\mu \in \mathcal{U}_W$, the error probability of $\mathcal{C}$ is

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = \alpha_{\frac{1}{M}}\big(W_x \,\|\, \mu\big). \quad (108)$$

Remark 1: Our definition of quasi-perfect codes is restricted to channels satisfying certain symmetry conditions. In general, we could define a quasi-perfect code C of cardinality M to be quasi-perfect whenever it attains the meta-converse bound with equality, i.e., if it satisfies (100). While this operational definition would apply for symmetric and non-symmetric channels, the condition itself is difficult to verify and it does not yield any simple expression for the error probability of these codes.

Theorem 4 shows that, whenever they exist, quasi-perfect codes attain the meta-converse bound (42) with equality. Particularizing this result in the classical case, for codes satisfying the technical conditions in Definition 2, we obtain [12, Th. 1]. This occurs for example for quasi-perfect binary codes for the BSC. Definition 2 extends the notion of generalized perfect and quasi-perfect codes to classical-quantum symmetric channels and Theorem 4 shows their optimality.

In the classical setting the codes belonging to this class are rare and only exist for short blocklengths. Then, one may wonder if they exist at all for classical-quantum channels of interest. In the next section we show that this is the case for a family of 2-qubit classical-quantum channels and certain code parameters.

V. 2-Qubit Classical-Quantum Channels and Bell Codes

A. *Pure 2-qubit classical-quantum channel*

We consider a 2-qubit pure-state channel

$$x \to W_x = |\varphi_x\rangle\langle\varphi_x|. \quad (109)$$

We define the codebook $\mathcal{C} = \{x_1, \ldots, x_M\}$, with even cardinality $M = 2K \geq 4$, such that the channel output of the m-th codeword is $W_m = |\varphi_{x_m}\rangle\langle\varphi_{x_m}|$ with

$$|\varphi_{x_m}\rangle = \begin{cases}\frac{1}{\sqrt{2}}\big(|00\rangle + e^{j\phi_k}\,|11\rangle\big), & m = 1+2k,\\ \frac{1}{\sqrt{2}}\big(|01\rangle + e^{j\phi_k}\,|10\rangle\big), & m = 2+2k,\end{cases} \quad (110)$$

where $\phi_k = 2\pi k/K$, for $k = 0 \ldots K-1$.

For $M = 4$, the channel outputs $|\varphi_{x_m}\rangle$ correspond precisely to the Bell states [39]. For $M \geq 4$, we refer to this family of codes as *Bell codes*, since they follow from a generalization of the Bell states.

Since $\sum_{m=1}^{M} |\varphi_{x_m}\rangle \langle\varphi_{x_m}| = \frac{M}{4}\mathbb{1}$ for $M \geq 4$, these codes are either perfect (when $M = 4$) or quasi-perfect (when $M > 4$) for the 2-qubit pure-state channel.

Proposition 1: The 2-qubit classical-quantum channel $W_x = |\varphi_x\rangle \langle\varphi_x|$ is symmetric with respect to $\mu_0 = \frac{1}{4}\mathbb{1}_4$ and the Bell code $\mathcal{C}$ is quasi-perfect for this channel. Moreover,

$$\mathsf{P}_\mathrm{e}(\mathcal{C}) = \alpha_{\frac{1}{M}}\bigl(W_x \,\|\, \mu_0\bigr) = 1 - \frac{4}{M}. \tag{111}$$

Proof: See Examples 2 and 3 in Section IV, with error probability given in (93). ■

When $M = 4$, the code corresponds precisely to the (orthogonal) Bell states and the transmitted message can be determined without errors. For $M = 2K > 4$, the codewords are no longer orthogonal to each other and therefore they incur in measurement errors even for the ideal pure 2-qubit classical-quantum channel. Nevertheless, as shown in Theorem 2 and in Proposition 1, there exist no other packing of pure states with lower error probability.

B. Classical-quantum depolarizing channel

This family of codes is not only optimal for the ideal pure-state channel but also when the transmission is affected by certain errors, as we will see next.

Consider the 2-qubit classical-quantum channel (109) where W_x is observed through a quantum depolarizing channel, defined as

$$\mathcal{N}^D_{A\to B}(\rho_A) = p\frac{1}{4}\mathbb{1}_4 + (1-p)\rho_A, \tag{112}$$

where $0 \leq p \leq 1$ is the depolarization parameter.

The combined classical-quantum channel is thus

$$x \to W_x = \mathcal{N}^D_{A\to B}\bigl(|\varphi_x\rangle \langle\varphi_x|\bigr). \tag{113}$$

Using the Bell code defined in (110), the channel output is $x_m \to W_m = \mathcal{N}^D_{A\to B}\bigl(|\varphi_{x_m}\rangle \langle\varphi_{x_m}|\bigr)$, $m = 1,\dots,M$.

Proposition 2: Let $\mu_0 = \frac{1}{4}\mathbb{1}_4$. Then, the 2-qubit classical-quantum depolarizing channel is symmetric with respect to μ_0 and the Bell code $\mathcal{C}$ is quasi-perfect for this channel. Moreover,

$$\mathsf{P}_\mathrm{e}(\mathcal{C}) = \alpha_{\frac{1}{M}}\bigl(W_x \,\|\, \mu_0\bigr) = 1 - \frac{1}{M}(4-3p), \tag{114}$$

which is obtained using decoder $\mathcal{T} = \{\Pi_1,\dots,\Pi_M\}$ with Π_i given by

$$\Pi_i = \frac{4}{M} |\varphi_{x_i}\rangle \langle\varphi_{x_i}|. \tag{115}$$

Proof: Consider the decoder $\mathcal{T} = \{\Pi_1,\dots,\Pi_M\}$ with Π_i defined in (115).

1) Decoder optimality: One can check that $\Pi_i \geq 0$ and $\sum_{i=1}^{M} \Pi_i = \mathbb{1}_4$. For this decoder,

$$\begin{aligned}\Lambda(\mathcal{T}) &\triangleq \frac{1}{M}\sum_{i=1}^{M} W_i \Pi_i &(116)\\ &= \frac{4}{M^2}\sum_{i=1}^{M} W_i |\varphi_{x_i}\rangle \langle\varphi_{x_i}| &(117)\\ &= \frac{1}{4M}(4-3p)\mathbb{1}_4. &(118)\end{aligned}$$

Then, it follows that

$$\Lambda(\mathcal{T})\Pi_i = \frac{1}{M} W_i \Pi_i, \tag{119}$$

which implies (13). Equation (14) is satisfied since, for arbitrary unit norm vector $|\psi\rangle$,

$$\begin{aligned}\frac{\langle\psi| \Lambda(\mathcal{T}) |\psi\rangle}{\frac{1}{M}\langle\psi| W_i |\psi\rangle} &= \frac{\frac{1}{4M}(4-3p)}{\frac{1}{4M}(p + 4(1-p)|\langle\psi|\varphi_{x_i}\rangle|^2)} &(120)\\ &\geq \frac{4-3p}{p+4(1-p)} = 1 &(121)\end{aligned}$$

So $\mathcal{T} = \{\Pi_1,\dots,\Pi_M\}$ minimizes the error probability for the Bell code $\mathcal{C}$.

2) Symmetry of the channel with respect to μ_0: We will prove next that

$$\mathcal{E}_x(t,\mu_0) = \begin{cases} \mathbb{1}_4, & t < 0, \\ |v\rangle \langle v|, & 0 \leq t \leq t_0, \\ 0, & t > t_0, \end{cases} \tag{122}$$

for $|v\rangle = |\varphi_x\rangle$ and $t_0 = 4 - 3p$ independent of x. Then, using (122) in $F_x(t, \mu_0) = \mathrm{Tr}\big(W_x \mathcal{E}_x(t, \mu_0)\big)$, it yields

$$F_x(t, \mu_0) = \begin{cases} 1, & t < 0, \\ 1 - \frac{3}{4}p, & 0 \leq t \leq t_0, \\ 0, & t > t_0, \end{cases} \tag{123}$$

and $\mathrm{Tr}\big(W_x \,|v\rangle \langle v|\big)$ is independent of φ_x, so the channel is symmetric with respect to μ_0.

It remains to show that (122) holds. The identity for $t < 0$ follows trivially. We consider an arbitrary unit-norm vector $|v\rangle$. Then, the largest eigenvalue of $W_x - t\mu_0$ is given by

$$\max_v \langle v| \left(W_x - t\mu_0\right) |v\rangle \tag{124}$$

$$= \max_v \left\{ \frac{p}{4} + (1-p) |\langle v|\varphi_x\rangle|^2 - \frac{t}{4} \right\} \tag{125}$$

$$= 1 - \frac{3}{4}p - \frac{t}{4}. \tag{126}$$

The eigenvalue (126) is negative for $t > 4 - 3p$ and non-negative otherwise. Then, we obtain that $F_x(t, \mu_0) = 0$, for $t > 4 - 3p$. For $0 \leq t \leq 4 - 3p$, (126) is the only non-negative eigenvalue with associated eigenvector $|v\rangle = |\varphi_x\rangle$. Therefore, considering the three regions, we obtain (122).

3) $\mathcal{C}$ is quasi-perfect with respect to μ_0: Comparing (118) with the auxiliary state μ_0 considered in the statement of Proposition 2, we observe that

$$\mu_0 = \frac{1}{c_0} \Lambda(\mathcal{T}) = \frac{1}{M c_0} \sum_{m=1}^{M} W_m \Pi_m, \tag{127}$$

where $c_0 = \frac{4-3p}{M}$ is a normalizing constant and where $\mathcal{T}$ satisfies the optimality conditions.

Take $t = M c_0 = 4 - 3p$, then $\frac{1}{M} W_m - \Lambda(\mathcal{T})$ is negative semidefine and $\mathcal{E}^{\bullet}_{x_m}(t, \mu_0) = 0$. As a result, $\big\{\mathcal{E}^{\bullet}_{x_m}(t, \mu_0)\big\}_{x \in \mathcal{C}}$ are orthogonal to each other. Similarly, for this choice of t and μ_0, it follows that $\mathcal{E}^{\circ}_{x_m}(t, \mu_0) = |\varphi_{x_m}\rangle \langle \varphi_{x_m}|$. Therefore $\sum_{x \in \mathcal{C}} \mathcal{E}^{\circ}_x(t, \mu) = \frac{M}{4} \mathbb{1}_4$ and the code is quasi-perfect.

4) Error probability: Using Theorem 4, it follows that $\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = \alpha_{\frac{1}{M}}\big(W_x \,\|\, \mu_0\big)$. Moreover, using the optimal decoder $\mathcal{T}$, we obtain

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = 1 - \frac{1}{M} \sum_{i=1}^{M} \mathrm{Tr}\big(W_i \Pi_i\big) \tag{128}$$

$$= 1 - \mathrm{Tr}\big(\Lambda(\mathcal{T})\big) \tag{129}$$

$$= 1 - \frac{4 - 3p}{M}, \tag{130}$$

where in the last step we used (118). ■

C. Extension to N-qubit classical-quantum channels

Consider now an arbitrary N-qubit classical-quantum channel with pure outputs given by

$$|\varphi\rangle = \sum_{l=0}^{2^N - 1} \alpha_l \,|l\rangle \tag{131}$$

$$= \sum_{l=0}^{2^N - 1} \alpha_l \,|l_{N-1} \dots l_0\rangle \tag{132}$$

$$= \alpha_0 \,|0 \dots 00\rangle + \alpha_1 \,|0 \dots 01\rangle + \alpha_{2^N - 1} \,|1 \dots 11\rangle \tag{133}$$

for $\sum_{l=0}^{2^N-1} |\alpha_l|^2 = 1$ and where $l_{N-1} \dots l_0$ are the digits of the binary representation of l. The channel is then $x \to W_x = |\varphi_x\rangle \langle \varphi_x|$. For $M = 2^{N-1} K \geq 2^N$, we define the N-qubit Bell codebook of cardinality M given by $\mathcal{C} = \big\{x_1, \dots, x_M\big\}$ with channel outputs

$$|\varphi_{x_m}\rangle = \begin{cases} \frac{1}{\sqrt{2}}\big(|00\rangle + e^{j\phi_k} |11\rangle\big) \otimes |l_{N-3} \dots l_0\rangle\,, \\ \qquad m = 1 + 2k + 2Kl, \\ \\ \frac{1}{\sqrt{2}}\big(|01\rangle + e^{j\phi_k} |10\rangle\big) \otimes |l_{N-3} \dots l_0\rangle\,, \\ \qquad m = 2 + 2k + 2Kl, \end{cases} \tag{134}$$

where $\phi_k = 2\pi k / K$, $k = 0, \dots, K - 1$, and where $l = 0, \dots, 2^{N-2} - 1$.

The channel ouput for codeword x_m is thus given by the pure state $W_m = |\varphi_{x_m}\rangle \langle \varphi_{x_m}|$.

Proposition 3: Let $\mu_0 = \frac{1}{2^N}\mathbb{1}_{2^N}$. The N-qubit classical-quantum channel is symmetric with respect to μ_0 and the N-qubit Bell code $\mathcal{C}$ is quasi-perfect for this channel. Moreover,

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = \alpha_{\frac{1}{M}}\big(W_x \,\|\, \mu_0\big) = 1 - \frac{2^N}{M}. \tag{135}$$

which is obtained using $\mathcal{T} = \{\Pi_1, \ldots, \Pi_M\}$ with

$$\Pi_i = \frac{2^N}{M} W_i. \tag{136}$$

Proof: The proof follows analogous steps to that of Proposition 1 and it is omitted here. ■

The N-qubit Bell code is also quasi-perfect for channels affected by depolarization as stated by the following result which is a generalization of Proposition 2.

Consider the N-qubit classical-quantum channel in (131) observed after a quantum depolarizing channel:

$$\mathcal{N}^D_{A\to B}(\rho_A) = p\frac{1}{2^N}\mathbb{1}_{2^N} + (1-p)\rho_A, \tag{137}$$

The combined classical-quantum channel is $x \to W_x = \mathcal{N}^D_{A\to B}\big(|\varphi_x\rangle\langle\varphi_x|_A\big)$. Using the Bell code defined in (134), the channel output to codeword x_m is given by $W_m = \mathcal{N}^D_{A\to B}\big(|\varphi_{x_m}\rangle\langle\varphi_{x_m}|_A\big)$, $m = 1, \ldots, M$.

Proposition 4: Let $\mu_0 = \frac{1}{2^N}\mathbb{1}_{2^N}$. Then, the N-qubit classical-quantum depolarizing channel is symmetric with respect to μ_0 and the N-qubit Bell code $\mathcal{C}$ is quasi-perfect for this channel. Moreover,

$$\mathsf{P}_{\mathrm{e}}(\mathcal{C}) = \alpha_{\frac{1}{M}}\big(W_x \,\|\, \mu_0\big) = 1 - \frac{1}{M}\big(2^N(1-p) + p\big). \tag{138}$$

which is obtained using decoder $\mathcal{T} = \{\Pi_1, \ldots, \Pi_M\}$ with

$$\Pi_i = \frac{2^N}{M}|\varphi_{x_i}\rangle\langle\varphi_{x_i}|. \tag{139}$$

Proof: The proof follows analogous steps to that of Proposition 2 and it is omitted here. ■

VI. Discussion

In this work we explored the connections between hypothesis testing and classical-quantum channel coding. First, we obtained 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 expression in Theorem 1 illustrates connections among the different settings of hypothesis testing and Corollary 1 provides an alternative formulation based on information-spectrum measures. A direct application of these results to a classical-quantum channel coding problem shows that Matthews-Wehner converse bound [22, Th. 19] and Hayashi-Nagaoka lemma [20, Lemma 4] with certain parameters yield the exact error probability in this setting.

While these results are of theoretical interest, the resulting expressions still depend on the the codebook and their application as performance benchmarks for classical-quantum channels is limited. We studied different relaxations and connections with practical converse bounds in the literature, thus characterizing the weaknesses of these bounds and the gap to the exact channel-coding error probability. Of special interest for this work is the so-called meta-converse bound [22, Eq. (46)], presented here in Theorem 2, which corresponds to the error probability of a binary hypothesis test with certain parameters.

In the second part of this work, we introduced the notion of perfect and quasi-perfect codes for symmetric classical-quantum channels. It is interesting to note that this notion is channel dependent –since a code being perfect for a channel it is not necessarily perfect for another one– and that it encompasses classical perfect and quasi-perfect codes as a special case [12, Sec. IV] provided some technical conditions hold. Theorem 3 provides an expression of the error probability of perfect and quasi-perfect codes for symmetric classical-quantum channels, which is then used in Theorem 4 to prove that these codes attain the meta-converse bound with equality. These codes, whenever they exist, are thus optimal in the sense that they achieve the smallest error probability among all codes of the same blocklength and cardinality.

Establishing the existence of generalized perfect and quasi-perfect codes for a given set of system parameters is a difficult problem, even for simple classical channels. For instance, [40] studies their existence for the BSC channel and [41] shows that MDS codes, which are generalized quasi-perfect for the q-ary erasure channel, only exist for blocklengths $n \leq q+1$. In this work, we consider a family of 2-qubit classical-quantum channels affected by depolarization. Using the framework presented, we established that a generalization of Bell states, that we name Bell codes, are quasi-perfect for these channels when the code cardinality is $M \geq 4$. For these channels and code parameters, we have thus established the error probability and structure of the best coding scheme. Proving the existence of perfect and quasi-perfect codes for other classical-quantum channels of practical interest is an unexplored line of research.

References

[1] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, July and Oct. 1948.

[2] R. Urbanke and T. Richardson, *Modern Coding Theory*. Cambridge University Press, 2008.

[3] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. Comm.*, Geneva, Switzerland, May 23–16 1993.

[4] E. Arikan, "Channel polarization: A method constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, Inc., 1968.

[6] C. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell System Tech. J.*, vol. 38, pp. 611–656, 1959.

[7] 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.

[8] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Feedback in the non-asymptotic regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4903–4925, 2011.

[9] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, 2014.

[10] M. C. Coşkun, G. Durisi, T. Jerkovits, G. Liva, W. Ryan, B. Stein, and F. Steiner, "Efficient error-correcting codes in the short blocklength regime," *Physical Comm.*, vol. 34, pp. 66 – 79, 2019.

[11] I. E. Bocharova, B. D. Kudryashov, E. P. Ovsyannikov, V. Skachek, and T. Uustalu, "Design and analysis of NB QC-LDPC codes over small alphabets," *IEEE Transactions on Communications*, pp. 1–1, 2022, in press.

[12] G. Vazquez-Vilar, A. Guillén i Fàbregas, and S. Verdú, "The error probability of generalized perfect codes via the meta-converse," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5705–5717, Sep. 2019.

[13] A. S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 269–273, Jan. 1998.

[14] M. D. W. B. Schumacher, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol. 56, no. 1, pp. 131–138, 1997.

[15] M. M. Wilde and S. Guha, "Polar codes for classical-quantum channels," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 1175–1187, 2013.

[16] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, "A demonstration of superadditivity in the classical capacity of a quantum channel," *Phys. Lett. A*, 1997.

[17] ——, "Quantum channels showing superadditivity in classical capacity," *Phys. Rev. A*, vol. 58, pp. 146–158, Jul 1998.

[18] S. Guha, "Structured optical receivers to attain superadditive capacity and the holevo limit," *Physical review letters*, vol. 106, no. 24, p. 240502, June 2011.

[19] M. T. DiMario, L. Kunz, K. Banaszek, and F. E. Becerra, "Optimized communication strategies with binary coherent states over phase noise channels," *npj Quantum Information*, vol. 5, no. 65, July 2019.

[20] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1753–1768, 2003.

[21] M. Hayashi, *Quantum Information: An Introduction*. Springer, 2006.

[22] W. Matthews and S. Wehner, "Finite blocklength converse bounds for quantum channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7317–7329, 2014.

[23] C. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Information and Control*, vol. 10, no. 1, pp. 65 – 103, 1967.

[24] R. E. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 4, pp. 405–417, 1974.

[25] H. Nagaoka, "Strong converse theorems in quantum information theory," in *ERATO Workshop on Quantum Information Science*, Tokyo, Japan, Sep. 2001, p. 68.

[26] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Phys. Rev. A*, vol. 52, no. 4, 1995.

[27] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," *Phys. Rev. Lett.*, vol. 108, no. 20, p. 200501, 2012.

[28] G. Vazquez-Vilar, A. Tauste Campo, A. Guillén i Fàbregas, and A. Martinez, "Bayesian M-ary hypothesis testing: The meta-converse and Verdú-Han bounds are tight," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2324–2333, 2016.

[29] C. W. Helstrom, *Quantum Detection and Estimation Theory*. NY: Academic Press, 1976.

[30] A. Jenčová, "Quantum hypothesis testing and sufficient subalgebras," *Lett. Math. Phys.*, vol. 93, no. 1, pp. 15–27, 2010.

[31] G. Vazquez-Vilar, "Multiple quantum hypothesis testing expressions and classical-quantum channel converse bounds," in *2016 IEEE Int. Symp. on Inf. Theory*, Barcelona, Spain, July 2016.

[32] 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.

[33] A. S. Holevo, "Statistical decision theory for quantum systems," *J. Multivariate Anal. 3*, vol. 3, no. 4, pp. 337–394, 1973.

[34] 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.

[35] A. S. Holevo, "A note on covariant dynamical semigroups," *Reports on Mathematical Physics*, vol. 32, no. 2, pp. 211–216, Apr. 1993.

[36] M. Hayashi, *Quantum Information Theory: Mathematical Foundation*. Springer, 2017, graduate Texts in Physics.

[37] H.-C. Cheng, M.-H. Hsieh, and M. Tomamichel, "Sphere-packing bound for symmetric classical-quantum channels," in *IEEE Symposium on Information Theory (ISIT)*, 2017, pp. 286–290.

[38] ——, "Quantum sphere-packing bounds with polynomial prefactors," *IEEE Trans. Inf. Theory*, vol. 65, no. 5, p. 28722898, May 2019.

[39] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[40] T. Etzion and B. Mounits, "Quasi-perfect codes with small distance," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3938–3946, Nov 2005.

[41] G. Seroussi and R. M. Roth, "On MDS extensions of generalized Reed-Solomon codes," *IEEE Trans. Inf. Theory*, vol. 32, no. 3, pp. 349–354, May 1986.

Andreu Blasco Coll received the Telecommunication Systems Engineering degree from the Universitat Politècnica de Catalunya (UPC), in 2015, the Master's degree in Telecommunications Engineering from the Universitat Politècnica de Catalunya (UPC), in 2017. In 2018, he joined the Department of Signal Theory and Communications of Universitat Politècnica de Catalunya (UPC) as a Ph.D. student.

Gonzalo Vazquez-Vilar (S'08–M'12) received the Telecommunication Engineering degree from the University of Vigo, Spain, in 2004, the Master of Science degree from Stanford University, U.S., in 2008 and the Ph.D. in Communication Systems from the University of Vigo, Spain, in 2011.

In 2011-2014 he was a post-doctoral fellow in the Department of Information and Communication Technologies, Universitat Pompeu Fabra, Spain and since 2014 he has been with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Spain. He has held appointments as visiting researcher at Stanford University, U.S., University of Cambridge, U.K., and Princeton University, U.S. His research interests lie in the field of Shannon theory, with emphasis on finite-length information theory and communications.

Javier Rodríguez Fonollosa (S'90, M'92, SM'98) received the Ph.D. degree in electrical and computer engineering at Northeastern University, Boston, MA, in 1992. In 1993 he joined the Department of Signal Theory and Communications of Universitat Politècnica de Catalunya (UPC) where he became Associate Professor in 1996, Professor in 2003 and Department Head from 2006 until 2010.

In 1995 he led UPC's participation in the European Commission funded ACTS Mobile projects TSUNAMI (II) and SUNBEAM that included the analysis of adaptive antennas in 2nd and 3rd generation cellular mobile communication systems. Since January 2000 until 2003 he was technical and project coordinator of the IST projects METRA and I-METRA dedicated to the introduction of multi-antenna terminals in UMTS and Systems beyond 3G. Since January 2006 to December 2008 he coordinated the Sixth Framework Programme IST project SURFACE which evaluated the performance of a generalized air interface with self-configuration capabilities.

In November 2006 he initiated the coordination of the 5 year Type C-Consolider project Fundamental bounds in Network Information Theory of the National Research Plan of Spain. Since December 2008 to December 2014 he was the Coordinator of the CONSOLIDER-INGENIO 2010 Foundations and Methodologies for Future Communication and Sensor Networks (COMONSENS), a 6 year 3.5 Million effort of 135 researchers belonging to 10 universities and research centers in Spain. The project continued under his coordination as a research network Red COMONSENS since December 2015 for seven more years.

He was department head of the Signal Theory and Communications department of UPC from October 2006 to January 2010. In 2009 he coordinated the ERASMUS MUNDUS 2009-2013 Master of Science in Research on Information and Communication Technologies (MERIT) EMMC Joint Master Program.

Since May 2005 he is member of the Editorial Board of the EURASIP Signal Processing Journal. In June 1995 and September 2001 he was co-chairman and organizer of the IEEE Signal Processing/ATHOS Workshop on Higher-Order Statistics held in Begur, Girona, Spain and of the IST Mobile Communications Summit 2001 held in Sitges, Barcelona, Spain. He was elected IEEE Senior Member and member of the Signal Processing for Communications (SPCOM) Technical Committee of the IEEE Signal Processing Society in February 1998 and January 1999 respectively.

Since February 2010 until July 2014 he was manager of the Communications and Electronic Technologies (TEC) area of the National Research Plan of Spain.