On the Duality Between the BSC and Quantum PSC

Narayanan Rengaswamy
Department of Electrical and Computer Engineering
The University of Arizona
Tucson, Arizona 85721, USA
Email: narayananr@arizona.edu

Henry D. Pfister
Department of Electrical and Computer Engineering
Duke University
Durham, North Carolina 27708, USA
Email: henry.pfister@duke.edu

***Abstract*—In 2018, Renes [IEEE Trans. Inf. Theory, vol. 64, no. 1, pp. 577-592 (2018)] developed a general theory of channel duality for classical-input quantum-output channels. His result shows that a number of well-known duality results for linear codes on the binary erasure channel can be extended to general classical channels at the expense of using dual problems which are intrinsically quantum mechanical. One special case of this duality is a connection between coding for error correction on the quantum pure-state channel (PSC) and coding for wire-tap secrecy on the classical binary symmetric channel (BSC). Similarly, coding for error correction on the BSC is related to wire-tap secrecy on the PSC.**

While this result has important implications for classical coding, the machinery behind the general duality result is rather challenging for researchers without a strong background in quantum information theory. In this work, we leverage prior results for linear codes on PSCs to give an alternate derivation of the aforementioned special case by computing closed-form expressions for the performance metrics. The noted prior results include the optimality of square-root measurement for linear codes on the PSC and the Fourier duality of linear codes.

I. Introduction

In the mathematical sciences, duality is a powerful concept that connects two problems so that the solution of one determines the solution of the other. In coding theory, an $[n, k]$ binary linear code $\mathcal{C} \subseteq \mathbb{F}_2^n$ is a k-dimensional subspace of the vector space of length-n binary vectors. In this case, its dual code $\mathcal{C}^{\perp} \subseteq \mathbb{F}_2^n$ is the $(n-k)$-dimensional subspace that is orthogonal to $\mathcal{C}$ under the standard dot product. An early and important implication of this duality is that the weight enumerator (WE) of a linear code can be computed from the WE of the dual code using the MacWilliams identity [1].

For classical channels, the notion of a dual channel did not arise until after the rediscovery of low-density parity-check (LDPC) codes. Even then, it was understood only for the erasure channel. Let BEC(ϵ) denote the binary erasure channel with erasure probability ϵ. It was shown in [2] that the extrinsic information transfer (EXIT) function for a code on the BEC(ϵ) is closely related to the EXIT function for the dual code on the BEC($1-\epsilon$). This and other symmetries in the decoding analysis of LDPC codes led some researchers to treat the BEC($1-\epsilon$) as the dual channel of the BEC(ϵ) [2]–[7].

For more than a decade, it remained an open question whether this notion of channel duality could be extended beyond the erasure case. In 2018, Renes [8] provided such a definition by showing that the dual channel of a binary memoryless channel can be defined in terms of a *classical-input quantum-output*, or simply a *classical-quantum* (CQ), channel [9]. In particular, Renes developed a general theory of CQ channel duality where a channel W and its dual $W^{\perp}$, both with d input symbols, satisfy $H(W) + H^{\perp}(W^{\perp}) = \log d$ for a pair of primal and dual entropies H and $H^{\perp}$ (e.g., see [10]). For example, if we let $W = \text{BEC}(\epsilon)$, then this approach shows that $W^{\perp} = \text{BEC}(1-\epsilon)$ and also recovers some previously known results for the BEC. Additionally, if $W = \text{PSC}(\theta)$ is the CQ binary *pure-state channel (PSC)* with parameter θ [11] (see Section II-C), then $W^{\perp} = \text{BSC}(p)$ is the classical binary symmetric channel (BSC) with parameter $p \triangleq (1-\cos\theta)/2$. Note that any classical channel can be treated as CQ by defining the outputs to be diagonal in the standard basis, i.e., $0 \equiv |0\rangle, 1 \equiv |1\rangle$. Thus, just as the complex numbers play an important role in our understanding of the real numbers, this suggests that CQ channels are also relevant to the theory of classical channels (e.g., see also [12]). Moreover, Renes also extends EXIT function duality for linear codes on the BEC to general CQ channels in [8].

This paper focuses on duality between coding problems for communication and secrecy. For communication on a channel W, a key quantity of interest is the maximum guessing probability, $P(W)$, for the input symbol given the channel output. Of course, this equals one minus the minimum decoder error probability. For secret communication where the channel to the eavesdropper is W, one measure of information leakage is the minimum squared Bhattacharyya coefficient, $Q(W)$, between the prior on the secret message X and the posterior distribution on X at the eavesdropper. Renes showed that for a pair of CQ channels W and $W^{\perp}$, channel duality implies that $P(W) = Q(W^{\perp})$ [8], where both quantities are optimized over all valid quantum measurements on the output system. For CQ channels W and $W^{\perp}$ with input $X \in \mathcal{X}$ and output B, $P(W)$ and $Q(W^{\perp})$ are related to quantum conditional entropies $H_{\min}(W) = H_{\min}(X|B)$ and $H_{\max}(W^{\perp}) = H_{\max}(X|B)$ as $P(W) = 2^{-H_{\min}(W)}$ and $Q(W) = \frac{1}{|\mathcal{X}|} 2^{H_{\max}(W^{\perp})}$. So, Renes' approach is to show that $H_{\min}(W) + H_{\max}(W^{\perp}) = \log |\mathcal{X}|$, which is part of his results

This work was supported in part by the National Science Foundation (NSF) under Grant No. 1908730, 1910571, and 1855879. Any opinions, findings, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these sponsors.

978-1-5386-8209-8/21/$31.00 ©2021 IEEE

for general dual entropies [10], and then the result follows.

Renes also extends this to the case where linear codes are used for both problems. In this paper, we give an alternate derivation of that result for the PSC-BSC pair by directly calculating closed-form expressions for (i) the block error rate and (ii) the Bhattacharyya coefficient between the posterior distribution of the secret message and the uniform distribution. For the secrecy setting, we emphasize here that the channel to the intended recipient is noiseless and the focus is only on the channel to the eavesdropper. While the approach in [8] uses some sophisticated quantum techniques, our exposition relies on direct calculation and targets an audience of classical information and coding theorists.

The square-root measurement (SRM), also called the least-squares measurement (LSM) or the pretty-good measurement (PGM), is an important and useful measurement for many quantum tasks [13]. Here, its optimality (or suboptimality) is discussed for the problems and channels considered. We have also used this SRM analysis for channel coding over the PSC to verify that belief propagation with quantum messages (BPQM) is quantum-optimal with respect to the block success probability for a 5-bit code [14]. Since BPQM produces a structured receiver circuit, unlike the SRM, this connection enables one to design receivers for optical communications over pure-loss bosonic channels. Hence, our work relates to optimal practical receivers for several communication problems.

II. Background and Notation

A. Quantum States and Measurements

In Dirac notation, the standard basis vectors of $\mathbb{C}^2$ are represented as $|0\rangle = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $|1\rangle = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$. For $n \geq 1$ qubits, the standard basis vectors are denoted by kets $|v\rangle = |v_1\rangle \otimes |v_2\rangle \otimes \cdots \otimes |v_n\rangle$, where $v_i \in \mathbb{Z}_2$ and $\otimes$ denotes the Kronecker product. Hence, a general n-qubit *pure state* is represented as

$$|\psi\rangle = \sum_{v \in \mathbb{Z}_2^n} \alpha_v |v\rangle \in \mathbb{C}^{2^n}, \; ; \; |\langle\psi|\psi\rangle|^2 = \sum_{v \in \mathbb{Z}_2^n} |\alpha_v|^2 = 1, \quad (1)$$

where "bra psi", $\langle\psi| \triangleq |\psi\rangle^\dagger$, is the conjugate transpose of $|\psi\rangle$ and the "bra-ket" $\langle\psi|\psi\rangle$ denotes the self inner product of $|\psi\rangle$.

If a quantum system is in one of several possible states $|\psi_m\rangle$ with probability p_m, for $m = 1, 2, \ldots, T$, then a succinct description of the system is given by its *density matrix* $\rho \triangleq \sum_{m=1}^{T} p_m |\psi_m\rangle \langle\psi_m| \in \mathbb{C}^{2^n \times 2^n}$, also called a *mixed state*.

A *projective measurement* is described by a set of complete orthogonal projectors $\{\Pi_i, \; i = 1, \ldots, M\}$ such that $\Pi_i \Pi_j = \delta_{ij} \Pi_i$ and $\sum_{i=1}^{M} \Pi_i = I_{2^n}$, where I_{2^n} is the $2^n \times 2^n$ identity matrix. If a system is in state ρ, then by the Born rule the measurement yields the result i with probability $p_i = \mathrm{Tr}\,[\Pi_i \rho]$, and projects the system to the state $\rho_i = \Pi_i \rho \Pi_i / p_i$ [9].

B. Linear Codes and their Complements

We briefly review a particular perspective on a linear code and its complements as described by Renes in [8]. A binary linear code $\mathcal{C} \colon [n, k]$ and its unique dual code $\mathcal{C}^\perp \colon [n, n-k]$ can be related through their non-unique complementary codes $\mathcal{C}^\top \colon [n, n-k]$ and $\mathcal{C}^{\underline{\top}} \colon [n, k]$, respectively. In particular, a code and one of its complements define disjoint subspaces and, thus, provide a direct-sum decomposition for the entire space. Note that once $\mathcal{C}$ and $\mathcal{C}^\top$ are fixed, $\mathcal{C}^\perp$ and $\mathcal{C}^{\underline{\top}}$ are fixed automatically. This organization of codes plays an important role in our results (see [15] for more details).

As an example, consider the following invertible matrices:

$$A = \left[\begin{array}{cccc} 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right], \; B = \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]. \quad (2)$$

The first row of A can be interpreted as a generator matrix, $G_{\mathcal{C}}$, for the 4-bit repetition code $\mathcal{C}$, which means that the last 3 rows of B form a generator matrix, $G_{\mathcal{C}^\perp}$, for the dual single-parity-check code $\mathcal{C}^\perp$. The remaining rows of A generate a code, $\mathcal{C}^\top$, that complements $\mathcal{C}$ so that A generates the whole vector space $\mathbb{F}_2^4$. Similarly, the first row of B generates a code, $\mathcal{C}^{\underline{\top}}$, that complements $\mathcal{C}^\perp$ and is also dual to $\mathcal{C}^\top$.

C. Pure-State Channel (PSC)

The pure-state CQ channel [11] can be described by the mapping $W^{\mathrm{PSC}(\theta)} \colon x \in \mathbb{Z}_2 \mapsto |(-1)^x\theta\rangle \equiv |(-1)^x\theta\rangle \langle(-1)^x\theta|$, where $|(-1)^x\theta\rangle \triangleq \cos\frac{\theta}{2}|0\rangle + (-1)^x \sin\frac{\theta}{2}|1\rangle$, $\theta \in \left[0, \frac{\pi}{2}\right]$. The overlap between the states is $\langle-\theta|\theta\rangle = \cos\theta$. For any θ, the optimal (Helstrom) measurement [16], [17] is given by the projectors $\{|+\rangle\langle+|, |-\rangle\langle-|\}$, which achieves the probability of error $P_{\mathrm{Hel}}^{\mathrm{PSC}} = \frac{1}{2}\left[1 - \sqrt{1 - |\langle-\theta|\theta\rangle|^2}\right] = (1 - \sin\theta)/2$. Hence, the PSC combined with this Helstrom measurement induces the binary symmetric channel $\mathrm{BSC}(P_{\mathrm{Hel}}^{\mathrm{PSC}})$. The PSC model can be motivated by the pure-loss bosonic channel in optical communications [14], [18].

Observe that the two output states of the PSC satisfy a symmetry: $|-\theta\rangle = Z|\theta\rangle$, where $Z \triangleq \left[\begin{smallmatrix}1 & 0\\0 & -1\end{smallmatrix}\right]$ is the Pauli Z operator. For a binary vector $b = [b_1, b_2, \ldots, b_n]$ we define $Z(b) \triangleq Z(b_1) \otimes Z(b_2) \otimes \cdots \otimes Z(b_n)$, where $Z(b_i) \triangleq Z^{b_i}$. Then, when $b \in \mathbb{Z}_2^n$ is transmitted over n uses of the PSC, the output state is given by $Z(b)|\theta\rangle^{\otimes n}$.

III. Duality of Channel Coding and Secrecy

We begin by discussing this duality for the BEC. Although the notion of a dual BEC has been known for a while, we describe the duality here in such a way that it builds the correct intuition for the PSC-BSC case that we consider subsequently. At the end of the first two subsections, we will relate the conditional entropy for channel coding on BEC(ϵ) and the conditional entropy for secrecy on BEC($1-\epsilon$). Then, in the next subsection, we prove the duality between channel coding on the PSC and secrecy on the BSC via direct computation of the optimal success (guessing) probability for the former. This computation exploits the optimality of the square root measurement (SRM) for this problem, and combines it with the factor graph duality of linear codes. Finally, we use these results to extend the conditional entropy approach for the BEC to the PSC-BSC pair. Due to space constraints, we omit most proofs and also our results for secrecy on the PSC where, among other things, we show that the SRM is suboptimal. For these details, we refer the interested reader to [15].

A. Duality for Channel Coding on the BEC

Let $\mathcal{C}$ be an (n,k) binary linear code with generator matrix G and parity-check matrix H. Assume that a random codeword X is chosen uniformly and transmitted through a BEC with output $Y \in \{0,1,?\}^n$. For an output realization y, let $\mathcal{E} \triangleq \{i \in [n] \,|\, y_i = ?\}$ be the set of indices where an erasure occurs. It turns out that many duality statements are more natural for a deterministic length-n BEC that erases all bits whose indices are in $\mathcal{E}$. We refer to this channel as BEC($\mathcal{E}$) and its dual channel, which correctly transmits only the bits with indices in $\mathcal{E}$, is BEC($\mathcal{E}^c$).

For a set $\mathcal{E} = (e_1, e_2, \ldots, e_{|\mathcal{E}|})$ with $e_1 < e_2 < \cdots < e_{|\mathcal{E}|}$ and an $m \times n$ matrix $A = (a_1, a_2, \ldots, a_n)$ whose i-th column is a, we let $A_{\mathcal{E}} = (a_{e_1}, a_{e_2}, \ldots, a_{e_{|\mathcal{E}|}})$. The same rule applies to row vectors using $m = 1$. Let $V = \{z \in \mathcal{C} \,|\, z_{\mathcal{E}^c} = y_{\mathcal{E}^c}\}$ be the set of codewords that are compatible with the observations. Then, the posterior distribution of X given Y is $P_{X|Y}(x|y) = 1/|V|$ if $x \in V$ and 0 otherwise.

Since $\mathcal{C}$ is linear, the set V is the affine subspace of $x \in \{0,1\}^n$ satisfying $H_{\mathcal{E}} x_{\mathcal{E}}^T = H_{\mathcal{E}^c} x_{\mathcal{E}^c}^T$ because $x_{\mathcal{E}^c}$ is known at the decoder. Thus, dimension of the solution space is given by $|\mathcal{E}| - \operatorname{rank}(H_{\mathcal{E}})$. Similarly, affine subspace of input vectors $u \in \{0,1\}^k$ compatible with y is defined by $uG_{\mathcal{E}^c} = x_{\mathcal{E}^c}$ and dimension of the solution space is $k - \operatorname{rank}(G_{\mathcal{E}^c})$. Of course, the two spaces must have the same dimension and this implies that $k - \operatorname{rank}(G_{\mathcal{E}^c}) = |\mathcal{E}| - \operatorname{rank}(H_{\mathcal{E}})$. This implies that

$$H(X|X_{\mathcal{E}^c}) = |\mathcal{E}| - \operatorname{rank}(H_{\mathcal{E}}) = k - \operatorname{rank}(G_{\mathcal{E}^c}). \quad (3)$$

Let $X' \in \mathcal{C}^\perp$ be a uniform random dual codeword. Then, the first equality in (3) shows that the entropy of X' over the dual channel BEC($\mathcal{E}^c$) is given by

$$H(X'|X'_{\mathcal{E}}) = |\mathcal{E}^c| - \operatorname{rank}(H^\perp_{\mathcal{E}^c}) = |\mathcal{E}^c| - \operatorname{rank}(G_{\mathcal{E}^c})$$

because $H^\perp = G$. Applying the second equality in (3) gives

$$H(X'|X'_{\mathcal{E}}) = H(X|X_{\mathcal{E}^c}) + |\mathcal{E}^c| - k. \quad (4)$$

B. Duality Between Channel Coding and Secrecy on the BEC

In 1975, Wyner introduced the wire-tap channel and proposed encoding secret messages into cosets of a group code [19]. Encoding proceeds by using the secret message to choose a coset and then encoding to a uniform random element from that coset. In this section, we will see that there is a duality between the information loss of channel coding using $\mathcal{C}$ and the information leakage of Wyner's coset coding using $\mathcal{C}^\perp$.

First, we will consider the standard channel coding problem for $\mathcal{C}$ and Wyner's wire-tap coding using cosets of $\mathcal{C}$. The coding problem transmits the codeword $x \in \{0,1\}^n$ as determined by the information $u \in \{0,1\}^k$ and coset selector $s \in \{0,1\}^{n-k}$ using the definitions

$$A = \begin{bmatrix} G \\ F \end{bmatrix}, \quad x = [u \;\, s]\, A = [u \;\, s] \begin{bmatrix} G \\ F \end{bmatrix} = uG + sF. \quad (5)$$

In this setup, rowspace(F) is a linear complement of $\mathcal{C} =$ rowspace(G) and A is full rank.

To decode the codeword, one would assume that the receiver knows the coset vector sF. In contrast, the wire-tap coding problem assumes u is unknown and tries to decode the secret message s. To make this stochastic, we let X be a uniform random vector over $\{0,1\}^n$ and define the random vectors $U \in \{0,1\}^k$ and $S \in \{0,1\}^{n-k}$ via $[U \; S] = XA^{-1}$. We note that choosing the uniform distribution for X is equivalent to using uniform distributions for U and S.

Since X is uniform, we can also interpret it as the codeword for the coding/secrecy problem using the dual code $\mathcal{C}^\perp$ on the dual channel, where the erased positions are in $\mathcal{E}^c$. Note that

$$xA^{-1} = [u \;\, s]AA^{-1} = [u \;\, s] \quad (6)$$

implies that the last k columns of A^{-1} give the transpose of a parity-check matrix H for $\mathcal{C}$. Also, the first $n-k$ columns of A^{-1} give a right inverse for G which we denote by E^T (i.e,. $GE^T = I$). Thus, we can define $B = (A^{-1})^T$,

$$B = \begin{bmatrix} E \\ H \end{bmatrix}, \quad x = [s' \; u']\, B = [s' \; u'] \begin{bmatrix} E \\ H \end{bmatrix} = s'E + u'H, \quad (7)$$

and view any $x \in \{0,1\}^n$ as the sum of the dual codeword $u'H$ and the coset vector $s'E$. We also define the random vectors $U' \in \{0,1\}^{n-k}$ and $S' \in \{0,1\}^k$ via $[S' \; U'] = XB^{-1}$.

The following derivation relates the secrecy and coding problems for the dual code on the dual channel:

$$\begin{aligned} n &= I(U', S'; X) \\ &= I(U', S'; X_{\mathcal{E}}, X_{\mathcal{E}^c}) \\ &= I(U', S'; X_{\mathcal{E}}) + I(U', S'; X_{\mathcal{E}^c}|X_{\mathcal{E}}) \\ &= I(S'; X_{\mathcal{E}}) + I(U'; X_{\mathcal{E}}|S') + I(U', S'; X_{\mathcal{E}^c}|\mathcal{E}, X_{\mathcal{E}}) \\ &= I(S'; X_{\mathcal{E}}) + I(U'; X_{\mathcal{E}}|, S') + (n - |\mathcal{E}|) \\ &= I(S'; X_{\mathcal{E}}) + (n-k) - H(U'|X_{\mathcal{E}}, S') + (n - |\mathcal{E}|). \quad (8) \end{aligned}$$

The final equation relates the number of erasures $|\mathcal{E}^c| = n - |\mathcal{E}|$, the information leakage $I(S'; X_{\mathcal{E}}|\mathcal{E})$ of the dual-code secrecy problem on the dual channel, and the message uncertainty $H(U'|X_{\mathcal{E}}, \mathcal{E}, S')$ for the coding problem using the dual code and dual channel.

Now, we can use (4) to substitute the uncertainty for primal coding on the primal channel, $H(U|X_{\mathcal{E}^c}, \mathcal{E}, S)$, for $H(U'|X_{\mathcal{E}}, \mathcal{E}, S')$ and a few other terms. This gives

$$H(U'|X_{\mathcal{E}}, \mathcal{E}, S') = H(U|X_{\mathcal{E}^c}, \mathcal{E}, S) + n - k - |\mathcal{E}|\,. \quad (9)$$

For the dual code and channel, combining with (8) allows us to write the information leakage of the secrecy problem as

$$\begin{aligned} H(S'|X_{\mathcal{E}}, \mathcal{E}) &= k - I(S'; X_{\mathcal{E}}|\mathcal{E}) \\ &= k - (k + H(U'|X_{\mathcal{E}}, \mathcal{E}, S') - n + |\mathcal{E}|) \\ &= k - H(U|X_{\mathcal{E}^c}, \mathcal{E}, S). \quad (10) \end{aligned}$$

Block Error Rate and Bhattacharyya Distance: If the uncertainty in U has dimension $d = H(U|X_{\mathcal{E}^c}, \mathcal{E}, S)$, then the probability of correctly guessing the codeword is 2^{-d}. If the posterior of S' given $X_{\mathcal{E}}$ and $\mathcal{E}$ is uniform over an affine subspace of dimension $f = k - I(S'; X_{\mathcal{E}}|\mathcal{E}) = k - d$, then the Bhattacharyya coefficient between this posterior distribution and the uniform distribution is

$$\sum_{s\in\{0,1\}^f} \sqrt{p_s 2^{-k}} = 2^f 2^{-f/2} 2^{-k/2} = 2^{(f-k)/2}.$$

By (10), we see that $f = k - d$ and $\left(2^{\frac{f-k}{2}}\right)^2 = 2^{f-k} = 2^{-d}$. Thus, the block guessing probability, 2^{-d}, equals the square of the Bhattacharyya coefficient, 2^{-d}, between this posterior distribution and the uniform distribution.

C. Channel Coding on the PSC and Secrecy on the BSC

Now, we will explicitly compute the block guessing probability for the PSC and show its equality to the squared Bhattacharyya coefficient for secrecy on the BSC.

For a M-ary hypothesis testing problem with candidate states $\{\rho_i\ ;\ i = 1, 2, \ldots, M\}$ and prior probabilities $\{p_i\ ;\ i = 1, 2, \ldots, M\}$, the minimum Bayes cost is given by

$$C^* = \sum_{j=1}^{M} \mathrm{Tr}\left[\hat{\Pi}_j \rho'_j\right], \quad \rho'_j \triangleq \sum_{i=1}^{M} p_i C_{ji} \rho_i, \tag{11}$$

where C_{ji} is the cost associated to deciding ρ_j when the truth is ρ_i, and $\{\hat{\Pi}_j;\ j = 1, 2, \ldots, M\}$ is the optimal measurement. For the transmission of an $[n, k]$ binary linear code $\mathcal{C}$ on PSC(θ), the minimum block error probability P_e equals C^* using the hypothesis testing problem with $C_{ji} = 1 - \delta_{ji}$ and $\rho_i = |\phi_i\rangle\langle\phi_i|$, where $|\phi_i\rangle$ is the result of transmitting the i-th codeword of $\mathcal{C}$ through PSC(θ).

This problem satisfies the *geometrically uniform (GU)* state set criterion of Eldar and Forney [13]. The criterion is that there is a generator state $|\phi\rangle$ and an abelian group $\mathcal{G}$, of size $|\mathcal{C}|$, such that each $|\phi_i\rangle$ can be written as $|\phi_i\rangle = U_i|\phi\rangle$ for some $U_i \in \mathcal{G}$. For this case, by the PSC symmetry mentioned in Section II-C, we have $|\phi\rangle = |\theta\rangle^{\otimes n}$ and $\mathcal{G} = \{Z(c),\ c \in \mathcal{C}\}$. Hence, for equally likely codewords, the SRM is the optimal measurement $\{\hat{\Pi}_j;\ j = 1, 2, \ldots, M\}$ [13]. We see that calculating P_e involves the inner products $|\langle\psi_j|\phi_i\rangle|$:

Definition 1: The elements of the SRM are $\hat{\Pi}_j = |\psi_j\rangle\langle\psi_j|$, where $|\psi_j\rangle$ is the j-th column of the SRM matrix

$$\Psi \triangleq \Phi\left(\left(\Phi^\dagger\Phi\right)^{1/2}\right)^{-1} \in \mathbb{C}^{2^n \times 2^k}, \tag{12}$$

the columns of Φ are $\{|\phi_i\rangle\ ;\ i = 0, 1, \ldots, 2^k - 1\}$, and the Moore-Penrose pseudo-inverse is used if Φ is rank deficient.

For the setting of transmitting binary linear codes over the PSC, the Gram matrix $\Phi^\dagger\Phi$ has full rank because Φ has full column rank. The optimality of SRM for this problem enables us to combine the results in [13] with Fourier duality of linear codes [20] to calculate the optimal block error rate as follows.

Definition 2: Given a linear code $\mathcal{C}$ that is transmitted over PSC(θ), define the overlap function $s(g) \triangleq \langle\theta|^{\otimes n} Z(c_g)|\theta\rangle^{\otimes n} = (\cos\theta)^{w_H(c_g)}$, where $w_H(c_g)$ is the Hamming weight of the codeword $c_g \triangleq gG_{\mathcal{C}} \in \mathcal{C}, g \in \mathbb{Z}_2^k$. Its Fourier transform is given by

$$\hat{s}(h) = \frac{1}{\sqrt{2^k}} \sum_{g\in\mathbb{Z}_2^k} (-1)^{hg^T} (\cos\theta)^{w_H(c_g)}. \tag{13}$$

The Fourier transform matrix is given by $\mathcal{F}_{gh} = \frac{1}{\sqrt{2^k}}(-1)^{gh^T}$, where the rows and columns are indexed by $g, h \in \mathbb{Z}_2^k$.

Now, using the above definitions, we will state a key result that enables us to calculate P_e in closed-form.

Theorem 3 ([13]): Consider the transmission of an $[n, k]$ binary linear code $\mathcal{C}$ over the channel PSC(θ). The codeword matrix Φ and the SRM matrix Ψ satisfy $\Psi^\dagger\Phi = \mathcal{F}\overline{\Sigma}\mathcal{F}^\dagger$, where $\overline{\Sigma}$ is a $2^k \times 2^k$ diagonal matrix with diagonal elements

$$\sigma(h) \triangleq 2^{k/4}\sqrt{\hat{s}(h)}, \quad h \in \mathbb{Z}_2^k. \tag{14}$$

Since $\Psi^\dagger\Phi = \mathcal{F}\overline{\Sigma}\mathcal{F}^\dagger$ is Hermitian, we observe $(\Psi^\dagger\Phi)_{ji} = (\Psi^\dagger\Phi)^*_{ij} \Rightarrow |\langle\psi_j|\phi_i\rangle|^2 = \left|(\Psi^\dagger\Phi)_{ji}\right|^2 = \left|(\Psi^\dagger\Phi)_{ij}\right|^2$. From [13], the columns of Ψ are given by $\{|\psi_g\rangle\ ,\ g \in \mathbb{Z}_2^k\}$, where

$$\begin{aligned} |\psi_g\rangle = \frac{1}{\sqrt{2^k}} \sum_{h\in\mathbb{Z}_2^k} (-1)^{gh^T} \frac{1}{\sigma(h)} \mathbb{I}(\sigma(h) \neq 0) \\ \cdot \frac{1}{\sqrt{2^k}} \sum_{f\in\mathbb{Z}_2^k} (-1)^{hf^T} Z(c_f)\,|\theta\rangle^{\otimes n}. \end{aligned} \tag{15}$$

and $\mathbb{I}(\cdot)$ denotes the indicator function that is 1 iff its argument is true. Hence, to compute the inner products $|\langle\psi_j|\phi_i\rangle|$ in P_e, we need to calculate $\sigma(h)$ or, equivalently, $\hat{s}(h)$ for all $h \in \mathbb{Z}_2^k$.

Factor Graph Duality Enables Calculation of P_e

We will now introduce the indicator function of $\mathcal{C}$ in $\hat{s}(h)$ in order to apply a factor graph duality [20]–[22] that produces the indicator function of $\mathcal{C}^\perp$ and simplifies the calculation of $\hat{s}(h)$. For this, let us embed $s(g)$ in $\mathbb{Z}_2^n$ by setting

$$s'(x) \triangleq \mathbb{I}(x \in \mathcal{C})\,(\cos\theta)^{w_H(x)},\ x \in \mathbb{Z}_2^n. \tag{16}$$

Then, the Fourier transform over $\mathbb{Z}_2^n$ produces

$$\hat{s}'(y) = \frac{1}{\sqrt{2^n}} \sum_{x\in\mathbb{Z}_2^n} (-1)^{yx^T} \mathbb{I}(x \in \mathcal{C})\,(\cos\theta)^{w_H(x)}. \tag{17}$$

Remark 4: We immediately observe that, if we express $y = y_{\mathcal{C}^{\top\mkern-14mu\perp}} + x_{\mathcal{C}^\perp}$ for some unique $y_{\mathcal{C}^{\top\mkern-14mu\perp}} \in \mathcal{C}^{\top\mkern-14mu\perp}$ and $x_{\mathcal{C}^\perp} \in \mathcal{C}^\perp$, then $\hat{s}'(y) = \hat{s}'(y_{\mathcal{C}^{\top\mkern-14mu\perp}})$ since $yx^T = y_{\mathcal{C}^{\top\mkern-14mu\perp}}x^T$ for all $x \in \mathcal{C}$.

Now, we see that the sum effectively happens over $\mathbb{Z}_2^k$ due to the presence of the indicator function, but the factor $(-1)^{yx^T}$ does not exactly map to $(-1)^{hg^T}$ since the latter is only taken over the "message" part of the codewords x, y assuming a systematic encoding. Hence, for a generic code, we need to make these coefficients $\hat{s}'(y)$ usable exactly in place of $\hat{s}(h)$. For the following results, see [15] for proofs.

Lemma 5: Let y_h denote the codeword in $\mathcal{C}^{\top\mkern-14mu\perp}$ corresponding to the message $h \in \mathbb{Z}_2^k$, i.e., $By_h^T = [0^{n-k}, h]^T$. Then, the functions $\hat{s}(h)$ and $\hat{s}'(y_h)$ are related as $\hat{s}(h) = \sqrt{2^{n-k}}\hat{s}'(y_h)$. This further implies that $\sigma(h) = 2^{k/4}\sqrt{\hat{s}(h)} = 2^{n/4}\sqrt{\hat{s}'(y_h)}$.

We use the following lemma to compute $\hat{s}(h)$ from (17).

Lemma 6 (Factor graph duality [22]): For binary vectors $x \in \mathbb{Z}_2^n$, given functions $\mu_j \colon \mathbb{Z}_2 \to \mathbb{R}$ for each index $j \in \{1, 2, \dots, n\}$, and an $[n, k]$ binary linear code $\mathcal{C}$, we have

$$\sum_{x \in \mathbb{Z}_2^n} \mathbb{I}(x \in \mathcal{C}) \prod_{j=1}^{n} \mu_j(x_j) = \sum_{\hat{x} \in \mathbb{Z}_2^n} 2^{k-n/2} \mathbb{I}(\hat{x} \in \mathcal{C}^{\perp}) \prod_{j=1}^{n} \hat{\mu}_j(\hat{x}_j),$$

where $\hat{\mu}_j(\hat{z}) \triangleq \frac{1}{\sqrt{2}} \sum_{z \in \mathbb{Z}_2} (-1)^{\hat{z}z} \mu_j(z)$.

Proof: See [23] for an algebraic proof rather than the graphical approach in [22]. ■

Lemma 7: Given an $[n, k]$ binary linear code $\mathcal{C}$ and the channel PSC(θ), $\hat{s}(h), h \in \mathbb{Z}_2^k$, can be calculated as

$$\frac{1}{2^{k/2}} \hat{s}(h) = \sum_{z \in y_h \oplus \mathcal{C}^{\perp}} p^{w_H(z)} (1-p)^{n - w_H(z)} \; ; \quad \sum_{h \in \mathbb{Z}_2^k} \frac{\hat{s}(h)}{2^{k/2}} = 1.$$

Here, y_h denotes the codeword in $\mathcal{C}^{\top}$ corresponding to the message $h \in \mathbb{Z}_2^k$. The code $\mathcal{C}^{\top}$ is the complement of $\mathcal{C}^{\perp}$ as defined by Renes [8], i.e., the codewords y_h of $\mathcal{C}^{\top}$ form coset leaders for the 2^k cosets of $\mathcal{C}^{\perp}$ in $\mathbb{Z}_2^n$.

Lemma 8: Consider an $[n, k]$ binary linear code $\mathcal{C}$ and the channel PSC(θ). The overlap between the square root measurement (SRM) vectors and the states obtained by transmitting the codewords of $\mathcal{C}$ over PSC(θ) is given by

$$|\langle \psi_g | \phi_t \rangle|^2 = \frac{\hat{\sigma}(g \oplus t)^2}{2^k}. \tag{18}$$

This is equal to the probability of sending a message $t \in \mathbb{Z}_2^k$ and decoding it as $g \in \mathbb{Z}_2^k$ using the SRM.

This was also used recently to verify the optimality of the BPQM algorithm for decoding a 5-bit code over PSC(θ) [14].

Theorem 9: Given an $[n, k]$ binary linear code $\mathcal{C}$, the optimal block error rate for transmission over PSC(θ) is given by

$$P_e = \frac{1}{2^k} \sum_{j \in \{0,1\}^k} \sum_{\substack{i \in \{0,1\}^k \\ i \neq j}} |\langle \psi_j | \phi_i \rangle|^2 = 1 - \mathcal{B}\left(\frac{\hat{s}}{2^{k/2}}, \frac{1}{2^k} \right)^2, \tag{19}$$

where the Bhattacharyya coefficient (or fidelity) between the probability distributions $\{2^{-\frac{k}{2}} \hat{s}(h)\}$ and $\{2^{-k}\}$ is defined as

$$\mathcal{B}\left(\frac{\hat{s}}{2^{k/2}}, \frac{1}{2^k} \right) \triangleq \sum_{h \in \mathbb{Z}_2^k} \sqrt{\frac{\hat{s}(h)}{2^{k/2}}} \sqrt{\frac{1}{2^k}}. \tag{20}$$

Furthermore, $\{2^{-k/2} \hat{s}(h)\}$ forms the posterior when cosets of $\mathcal{C}^{\perp}$ are used to send secure messages over the dual, BSC$\big(p = \frac{1 - \cos\theta}{2}\big)$. Hence, $\mathcal{B}\left(\frac{\hat{s}}{2^{k/2}}, \frac{1}{2^k}\right)^2$ measures the optimal decoupling of the secret from the intercepted information.

Thus, this approach uses standard linear algebra and group theory to establish that the block error rate for channel coding on the PSC(θ) with $\mathcal{C}$ equals the defined Bhattacharyya performance measure for secrecy using $\mathcal{C}^{\perp}$ on the BSC(p).

To understand this result in terms of channels, let $W_{\mathcal{C}}^{PSC(\theta)}$ denote the CQ channel implied by the channel coding problem over PSC(θ) using $\mathcal{C}$ (i.e., the input alphabet is $\{0, 1\}^k$). Similarly, let $W_{\mathcal{C}^{\perp}}^{BSC(p)}$ denote the CQ channel implied by sending the coset selector $s \in \{0, 1\}^k$ for the secrecy problem using $\mathcal{C}^{\perp}$, where the channel to the eavesdropper is BSC(p) and the channel to the intended recipient is noiseless. Rewriting the above result, we see that the optimal success probability $P\big(W_{\mathcal{C}}^{PSC(\theta)}\big)$ for transmitting a code over PSC(θ) is equal to the optimal secrecy metric $Q\big(W_{\mathcal{C}^{\perp}}^{BSC(p)}\big)$. In [8, Corollary 4], Renes proved a more general result for coding over general CQ channels using methods from quantum information theory.

D. Entropic Duality for the PSC-BSC Pair

Here, we extend the entropic result (10) to the PSC-BSC scenario, thereby completing the intuition we developed.

Lemma 10: For the channel coding problem on the PSC, the overlap (or normalized Grammian) matrix $\Gamma = 2^{-k} \Phi^{\dagger} \Phi$ is diagonalized by the Fourier transform $\mathcal{F}$. From this, we find that the set of (non-zero) eigenvalues of both Γ and the density matrix $\rho^{Y,S=0} = 2^{-k} \Phi \Phi^{\dagger}$ equal $\{2^{-k/2} \hat{s}(h) \,|\, h \in \{0, 1\}^k\}$.

Since the Von Neumann entropy $H(C)_{\rho}$ of a quantum system C with density matrix ρ^C equals the Shannon entropy of the eigenvalues of ρ^C [9], it follows that

$$\begin{aligned} H(Y|S=0)_{\rho^{Y,S=0}} &= \sum_{h \in \{0,1\}^k} 2^{-k/2} \hat{s}(h) \log \frac{1}{2^{-k/2} \hat{s}(h)} \\ &= \sum_{h \in \{0,1\}^k} \mathbb{P}\left(h \,|\, \hat{x}\right) \log \frac{1}{\mathbb{P}\left(h \,|\, \hat{x}\right)} \\ &= H(S'|Y'), \end{aligned} \tag{21}$$

where $\mathbb{P}\left(h \,|\, \hat{x}\right)$ is the posterior when cosets of $\mathcal{C}^{\perp}$ are used for secrecy on the dual BSC, and the final quantity is the classical Shannon entropy. Note that the set of values $\{\mathbb{P}\left(h \,|\, \hat{x}\right), h \in \{0, 1\}^k\}$ remain the same irrespective of the intercepted vector $\hat{x}$, and only the mapping $h \mapsto \mathbb{P}\left(h \,|\, \hat{x}\right)$ depends on $\hat{x}$.

Using the same setup as the BEC analysis in Section III-B, we can use the above result to investigate the Von Neumann conditional entropy. Similar to the BEC result in (10), we have

$$\begin{aligned} &H(U|Y, S=0)_{\rho^{UY,S=0}} \\ &= H(U|S=0) + H(Y|U, S=0)_{\rho^{UY,S=0}} - H(Y|S=0)_{\rho^{Y,S=0}} \\ &= k + 0 - H(S'|Y'). \end{aligned} \tag{22}$$

Next, consider the BEC channel coding duality result defined by (4). One can generalize this to the PSC by observing

$$\begin{aligned} &H(U'|Y', S') \\ &= H(U'|S') + H(Y'|S', U') - H(Y'|S') \\ &= n - k + \eta(p)\, n - \big(H(Y') + H(S'|Y') - H(S')\big) \\ &= n - k + \eta(p)\, n - \big(n + \big(k - H(U|Y, S=0)_{\rho^{UY,S=0}}\big) - k\big) \\ &= H(U|Y, S=0)_{\rho^{UY,S=0}} + \eta(p)\, n - k, \end{aligned} \tag{23}$$

where $\eta(p) \triangleq -p \log p - (1-p) \log(1-p)$ is the binary entropy function and the term $\eta(p)\, n$ can be seen as the total entropy produced by the dual channel.

Therefore, the Fourier transform $\hat{s}(h)$ of the overlap function $s(g)$ for the PSC forms the key link that connects channel coding on the PSC and secrecy on the BSC via performance metrices as well as entropies. This might be a general phenomenon that extends beyond the special case of the PSC and the BSC, and we leave this investigation to future work. For our results on secrecy over the PSC, see [15].

References

[1] J. MacWilliams, "A theorem on the distribution of weights in a systematic code," *The Bell Syst. Techn. J.*, vol. 42, no. 1, pp. 79–94, 1963.

[2] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Trans. Inform. Theory*, vol. 50, no. 11, pp. 2657–2674, Nov. 2004.

[3] S.-Y. Chung, "On the construction of some capacity-approaching coding schemes," Ph.D. dissertation, MIT, 2000.

[4] E. Martinian and J. S. Yedidia, "Iterative quantization using codes on graphs," in *Proc. Annual Allerton Conf. on Commun., Control, and Comp.*, 2003.

[5] H. D. Pfister and I. Sason, "Accumulate–repeat–accumulate codes: Capacity-achieving ensembles of systematic codes for the erasure channel with bounded complexity," *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2088–2115, June 2007.

[6] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of ldpc codes to the wiretap channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 8, pp. 2933–2945, 2007.

[7] N. Obata, Y.-Y. Jian, K. Kasai, and H. D. Pfister, "Spatially-coupled multi-edge type LDPC codes with bounded degrees that achieve capacity on the BEC under BP decoding," in *Proc. IEEE Int. Symp. Inform. Theory*, July 2013, pp. 2433–2437.

[8] J. M. Renes, "Duality of channels and codes," *IEEE Trans. Inform. Theory*, vol. 64, no. 1, pp. 577–592, 2018.

[9] M. M. Wilde, *Quantum Information Theory*. Cambridge University Press, 2013.

[10] M. Tomamichel, *Quantum Information Processing with Finite Resources: Mathematical Foundations*. Springer, 2015, vol. 5.

[11] J. M. Renes, "Belief propagation decoding of quantum channels by passing quantum messages," *New Journal of Physics*, vol. 19, no. 7, p. 072001, 2017. [Online]. Available: http://arxiv.org/abs/1607.04833

[12] M. Dalai, "Lower bounds on the probability of error for classical and classical-quantum channels," *IEEE Trans. Inform. Theory*, vol. 59, no. 12, pp. 8027–8056, 2013.

[13] Y. C. Eldar and G. D. Forney, "On quantum detection and the square-root measurement," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 858–872, 2000. [Online]. Available: http://arxiv.org/abs/quant-ph/0005132

[14] N. Rengaswamy, K. P. Seshadreesan, S. Guha, and H. D. Pfister, "Belief propagation with quantum messages for quantum-enhanced classical communications," *To appear in npj Quantum Information*, 2021. [Online]. Available: http://arxiv.org/abs/2003.04356

[15] N. Rengaswamy and H. D. Pfister, "A semiclassical proof of duality between the classical BSC and the quantum PSC," *arXiv preprint arXiv:2103.09225*, 2021. [Online]. Available: http://arxiv.org/abs/2103.09225

[16] C. W. Helstrom, "Quantum detection and estimation theory," *Journal of Statistical Physics*, vol. 1, no. 2, pp. 231–252, 1969.

[17] C. W. Helstrom, J. W. Liu, and J. P. Gordon, "Quantum-mechanical communication theory," *Proc. of the IEEE*, vol. 58, no. 10, pp. 1578–1598, 1970.

[18] S. Guha and M. M. Wilde, "Polar coding to achieve the Holevo capacity of a pure-loss optical channel," in *Proc. IEEE Int. Symp. Inform. Theory*, 2012, pp. 546–550. [Online]. Available: https://arxiv.org/abs/1202.0533

[19] A. D. Wyner, "The wire-tap channel," *The Bell Syst. Techn. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[20] C. R. P. Hartmann and L. D. Rudolph, "An optimum symbol-by-symbol decoding rule for linear codes," *IEEE Trans. Inform. Theory*, vol. 22, no. 5, pp. 514–517, 1976.

[21] G. D. Forney Jr., "Codes on graphs: Normal realizations," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 520–548, 2001.

[22] G. D. Forney Jr. and P. O. Vontobel, "Partition functions of normal factor graphs," in *Proc. Annual Workshop on Inform. Theory and its Appl.*, 2011. [Online]. Available: http://arxiv.org/abs/1102.0316

[23] H. D. Pfister, "Factor graph duality," 2014. [Online]. Available: http://pfister.ee.duke.edu/courses/ece590_gmi/fg_duality.pdf