Rate-Equivocation Region of Cyclic Shift Symmetric Wiretap Channels

Omur Ozel Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
omur@umd.edu *ulukus@umd.edu*

***Abstract*—In this paper, we study cyclic shift symmetric wiretap channels in which the channels between Alice and Bob and Alice and Eve are both cyclic shift symmetric. We characterize the rate-equivocation region by determining the optimal selection of rate splitting U and channel prefixing V for these channels. We show that optimal U and V are determined via cyclic shifts of the solution of an auxiliary optimization problem that involves only one auxiliary random variable. We find the cardinality bound on the necessary auxiliary variable and formulate the problem as a constrained optimization problem. We determine the optimality conditions for the binary-input cyclic shift symmetric wiretap channels. We find the optimum by inspecting each point of the $I(X;Y) - I(X;Z)$ function and ruling out the sub-optimal candidates that satisfy the optimality conditions. In particular, we address BSC-BEC and BEC-BSC wiretap channels. By using the optimality conditions, we determine the optimal selections of U and V for the rate-equivocation regions of these channels.**

I. Introduction

We consider the discrete memoryless wiretap channel shown in Fig. 1 where the main channel and the eavesdropper's channel are both cyclic shift symmetric. Cyclic shift symmetric channels are an important class that includes binary symmetric, binary erasure, and type-writer channels. The capacity region of a general wiretap channel is characterized by the rate, R, between the legitimate users Alice and Bob, and the equivocation, R_e, at the eavesdropper Eve. Wyner [1] characterized the rate-equivocation region when the received signal at Eve is a degraded version of the signal received at Bob. Csiszár and Körner [2] characterized the rate-equivocation region for general, not necessarily degraded, wiretap channels.

Csiszár and Körner's characterization involves two auxiliary random variables: U, for rate splitting, and V, for channel prefixing. In this paper, we explore the implications of cyclic shift symmetry on the structure of the optimal selections of U and V in a wiretap channel. Nair [3] characterized broadcast channel capacity region of a specific class of two-user cyclic shift symmetric broadcast channels. We studied similar classes in the wiretap channel context in [4] where we showed that optimal U and V have simplified structures for specific wiretap channels. In particular, we proved that rate splitting U and/or channel prefixing V is not necessary (i.e., optimal selection is $U = \phi$ and/or $V = X$) for particular classes of cyclic shift symmetric wiretap channels.

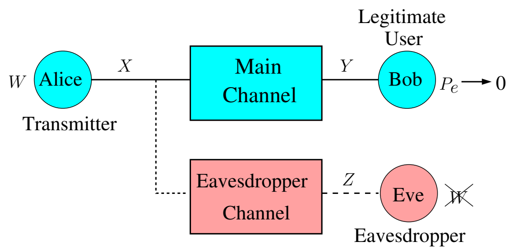

Fig. 1. The wiretap channel.

In this paper, we characterize the rate-equivocation region for the entire class of cyclic shift symmetric wiretap channels. We find the optimal selection of U and V for these channels in terms of the solution of an auxiliary optimization problem that involves only one auxiliary random variable, which is a significant reduction for the amount of calculations needed for the characterization of the rate-equivocation region. We provide a cardinality bound on the auxiliary random variable for the calculation of the optimal value. Then, we formulate the problem as a constrained optimization problem. Finally, we apply our results to the binary-input cyclic shift symmetric wiretap channels. In this case, each point on the boundary of the rate-equivocation region is calculated by a three-variable constrained optimization problem. We show that the problem gains sufficient structure under the cyclic shift symmetry assumption and the sub-optimal candidates for the optimal solution can be ruled out by inspecting the $I(X;Y) - I(X;Z)$ function. We investigate two specific examples: BSC-BEC and BEC-BSC wiretap channels. We provide full characterizations for the rate-equivocation regions of the BSC-BEC and BEC-BSC wiretap channels. In particular, we find that rate-splitting is never necessary for the BSC-BEC wiretap channel.

II. Model and Background

As in Fig. 1, Alice communicates with Bob in the presence of an eavesdropper, Eve. The input and output alphabets, $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$, are finite. The main channel is characterized by $p(y|x)$ and has capacity $C_B = \max_{P_x} I(X;Y)$. Similarly the wiretapper channel is characterized by $p(z|x)$ and has capacity $C_E = \max_{P_x} I(X;Z)$. Both $p(y|x)$ and $p(z|x)$ are cyclic shift symmetric. A channel $p(y|x)$ is cyclic shift symmetric if $I(X;Y)$ is invariant under cyclic shifts of the input distribution [5]. A key property of cyclic shift symmetric

This work was supported by NSF Grants CCF 07-29127, CNS 09-64632, CCF 09-64645 and CCF 10-18185.

channels is that the input distribution that maximizes the mutual information is the uniform distribution [5, Theorem 2]. Hence, $C_B = I_{\mathbf{u}}(X;Y)$ and $C_E = I_{\mathbf{u}}(X;Z)$ where $\mathbf{u}$ is the $|\mathcal{X}|$ dimensional uniform distribution.

W represents the message to be sent to Bob and kept secret from Eve with $W \in \mathcal{W} = \{1, \dots, 2^{nR}\}$. Alice uses an encoder $\varphi : \mathcal{W} \rightarrow \mathcal{X}^n$ to map each message to a channel input of length n. Bob uses a decoder $\psi : \mathcal{Y}^n \rightarrow \mathcal{W}$. The probability of error is: $P_e = \Pr[\psi(Y^n) \neq W]$. The rate R is achievable with equivocation R_e, if $P_e \rightarrow 0$ as $n \rightarrow \infty$, and

$$R_e = \lim_{n\rightarrow\infty} \frac{1}{n} H(W|Z^n) \tag{1}$$

Perfect secrecy[1] is achieved if $\frac{1}{n}I(W;Z^n) \rightarrow 0$ and the secrecy capacity C_s is the highest achievable perfectly secure rate R. The maximum possible equivocation is also C_s.

Throughout the paper, $f_\mu(.)$ denotes the following function of the input distribution P_x

$$f_\mu(P_x) = (\mu + 1)I(X;Y) - I(X;Z) \tag{2}$$

where $\mu \geq 0$ is an arbitrary parameter. We denote $f_0(.)$ simply as $f(.)$. Note that $f_\mu(.)$ is continuous and differentiable for all $\mu \geq 0$.

Csiszár and Körner [2] characterized the entire rate-equivocation region as stated in the following theorem.

Theorem 1 ([2, Corollary 2]) *(R, R_e) pair is in the rate-equivocation region if and only if there exist $U \rightarrow V \rightarrow X \rightarrow Y, Z$ such that $I(U;Y) \leq I(U;Z)$, and*

$$0 \leq R_e \leq I(V;Y|U) - I(V;Z|U) \tag{3}$$
$$R_e \leq R \leq I(V;Y) \tag{4}$$

Further, the secrecy capacity is

$$C_s = \max_{V\rightarrow X\rightarrow Y,Z} I(V;Y) - I(V;Z) \tag{5}$$

Finally, the cardinality bounds on the alphabets of the auxiliary random variables are

$$|\mathcal{U}| \leq |\mathcal{X}| + 3 \tag{6}$$
$$|\mathcal{V}| \leq |\mathcal{X}|^2 + 4|\mathcal{X}| + 3 \tag{7}$$

The rate-equivocation region of a wiretap channel is a convex region. Therefore, the upper right boundary is traced by solving the following optimization problem for all $\mu \geq 0$ as in Fig. 2:

$$\max_{U,V,X} \quad \mu I(V;Y) + I(V;Y|U) - I(V;Z|U) \tag{8}$$

Note that this optimization problem is computable due to the bounds on the sizes of U and V in (6) and (7) in Theorem 1. In the sequel, we refer to the solution of the optimization problem in (8) as the optimal selections U^*, V^* and X^*. These optimal selections depend implicitly on the value of μ. The optimal value of the objective function in (8) at $\mu = 0$ is the secrecy capacity C_s. In this case, U is unnecessary, and in fact, we get (5) [2]. Note that the bounds on the cardinalities of U and V in (6)-(7) in Theorem 1 are valid in general. However, the specific cardinality bound on V for the optimization problem in (8) when $\mu = 0$, or equivalently the problem in (5), is

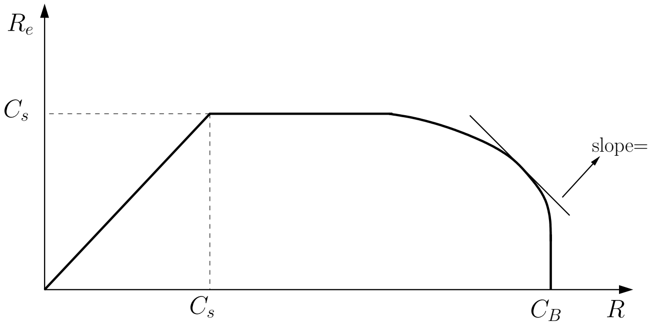

Fig. 2. Characterization of the upper right boundary of the rate-equivocation region.

$$|\mathcal{V}| \leq |\mathcal{X}| \tag{9}$$

In order to prove the cardinality bound in (9), given $V \rightarrow X \rightarrow Y, Z$ with PMFs $p(v)$ and $p(x|v)$, we fix the following $|\mathcal{X}|$ continuous functions of $p(x|v)$:

$$g_j(p_{X|V}(x|v)) = \begin{cases} p_{X|V}(j|v), & j = 1, \dots, |\mathcal{X}| - 1 \\ f(p(x|v)), & j = |\mathcal{X}| \end{cases} \tag{10}$$

From Lemma 3 in [8] and the strengthened Caretheodory theorem of Fenchel-Eggleston in [9], we can find another random variable $V^{'}$ with cardinality at most $|\mathcal{X}|$ such that $V^{'} \rightarrow X \rightarrow Y, Z$ and

$$I(X;Y|V) - I(X;Z|V) = I(X;Y|V^{'}) - I(X;Z|V^{'}) \tag{11}$$

as well as

$$p(x) = \int p_{X|V}(x|v)dF(v) = \sum_{v^{'}} p_{X|V^{'}}(x|v^{'})p(v^{'}) \tag{12}$$

for $x = 1, \dots, |\mathcal{X}| - 1$; see also Appendix C in [10]. Since $I(X;Y) - I(X;Z)$ remains unchanged, $I(V;Y) - I(V;Z) = I(V^{'};Y) - I(V^{'};Z)$. Therefore, $|\mathcal{V}| \leq |\mathcal{X}|$ cardinality is sufficient to solve (5).

III. Characterization of the Rate-Equivocation Region

In the following theorem, we determine the structure of the optimal auxiliary random variables U^* and V^* as well as the channel input X^* for cyclic shift symmetric wiretap channels. Remarkably, the optimizing rate splitting U^* and channel prefixing V^* parameters can be determined by solving an auxiliary optimization problem over only one auxiliary random variable. In addition, the cardinality bounds on U^* and V^* are reduced to $|\mathcal{X}|$ and $|\mathcal{X}|^2$, respectively, compared to the general case in (6) and (7). We provide the proof of this theorem in the Appendix.

[1] We use the weak secrecy notion. However, for discrete wiretap channels weak and strong secrecy are equivalent [6], [7].

Theorem 2 *In a cyclic shift symmetric wiretap channel, the optimal selection of the auxiliary random variables U^* and V^* in (8) have the cardinalities $|\mathcal{U}| \leq |\mathcal{X}|$ and $|\mathcal{V}| \leq |\mathcal{X}|^2$, respectively, with the following structure:*

$$p(U^* = u) = \frac{1}{|\mathcal{X}|}, \qquad u \in \{1, \ldots, |\mathcal{X}|\} \tag{13}$$

$$p(V^* = (u-1)|\mathcal{X}| + v | U^* = u) = p(\hat{V} = v), \quad u, v \in \{1, \ldots, |\mathcal{X}|\} \tag{14}$$

$$p(V^* = v | U^* = u) = 0, \quad u \in \{1, \ldots, |\mathcal{X}|\}, v \notin \{(u-1)|\mathcal{X}| + 1, \ldots, u|\mathcal{X}|\} \tag{15}$$

$$p(x | V^* = v + (u-1)|\mathcal{X}|) = p(x|\hat{V} = v)(u-1), \quad u, v, x \in \{1, \ldots, |\mathcal{X}|\} \tag{16}$$

where $p(X = x|\hat{V} = v)(u-1)$ denotes the $u-1$st cyclic shift of the distribution $p(x|\hat{V} = v)$. Moreover, the distributions $p(\hat{V} = v)$ and $p(X = x|\hat{V} = v)$ with $|\hat{\mathcal{V}}| \leq |\mathcal{X}|$ are the optimizers of the following auxiliary problem:

$$\max_{p(\hat{V}), p(X|\hat{V})} \left(f(P_x) - \mathbb{E}_{\hat{V}} \left[f_\mu \left(p(X|\hat{V}) \right) \right] \right)^+ \tag{17}$$

where $(x)^+ = \max\{0, x\}$.

We illustrate the specific structure of the optimal auxiliary random variables and the channel input in Fig. 3. In particular, each element of U^* generates the optimizing PMF $p(\hat{V})$ over $|\mathcal{X}|$ elements of V^*. The first $|\mathcal{X}|$ elements of V^* generate the optimizing conditional PMF $p(X|\hat{V} = v)$ over X. The remaining elements of V^* generate cyclic shifts of $p(X|\hat{V} = v)$ over X. An equivalent representation for the optimal selections can be obtained by letting $V^* = (V_1^*, V_2^*)$ with $|\mathcal{V}_1^*| = |\mathcal{V}_2^*| = |\mathcal{X}|$:

$$p(U^* = u) = \frac{1}{|\mathcal{X}|}, \quad u \in \{1, \ldots, |\mathcal{X}|\} \tag{18}$$

$$p(V^* = (v_1, v_2)|U^* = u) = p(\hat{V} = v_1)\delta(v_2 - u) \quad u, v_1, v_2 \in \{1, \ldots, |\mathcal{X}|\} \tag{19}$$

$$p(x|V^* = (v_1, v_2)) = p(x|\hat{V} = v_1)(v_2 - 1) \quad v_1, v_2 \in \{1, \ldots, |\mathcal{X}|\} \tag{20}$$

Note that U^* is a deterministic function of V^* as stated in [2, Theorem 1]. This is verified easily from the equivalent representation in (18)-(20). Given $V^* = (V_1^* = v_1, V_2^* = v_2)$, $U^* = v_1$ with probability 1. However, V^* is a stochastic function of U^*. These can also be verified from Fig. 3.

The optimization problem in (17) is a constrained optimization problem over $|\mathcal{X}|^2 - 1$ variables: $|\mathcal{X}|$ probability distributions on X, $p(X = x|\hat{V} = v_i)$. Each probability distribution accounts for $|\mathcal{X}| - 1$ variables for $i = 1, \ldots, |\mathcal{X}|$. In addition, the distribution for $\hat{V}$ accounts for $|\mathcal{X}| - 1$ variables. Let us define $\lambda_i \triangleq p(V = v_i)$ and $\left[p_1^{(i)} p_2^{(i)} \ldots p_{|\mathcal{X}|}^{(i)}\right] \triangleq p(X = x|V = v_i)$. We have $\lambda_i \geq 0$, $p_j^{(i)} \geq 0$ and $\sum_i \lambda_i = 1$, $\sum_j p_j^{(i)} = 1$. The following is a restatement of the constrained optimization problem in (17):

$$\max_{\{\lambda_i\}, \{p_j^{(i)}\}} \quad f\left(\sum_i \lambda_i p_j^{(i)}\right) - \sum_i \lambda_i f_\mu(p_j^{(i)})$$
$$\text{s.t.} \quad \sum_i \lambda_i = 1, \ \sum_j p_j^{(i)} = 1, \ \lambda_i \geq 0, \ p_j^{(i)} \geq 0 \tag{21}$$

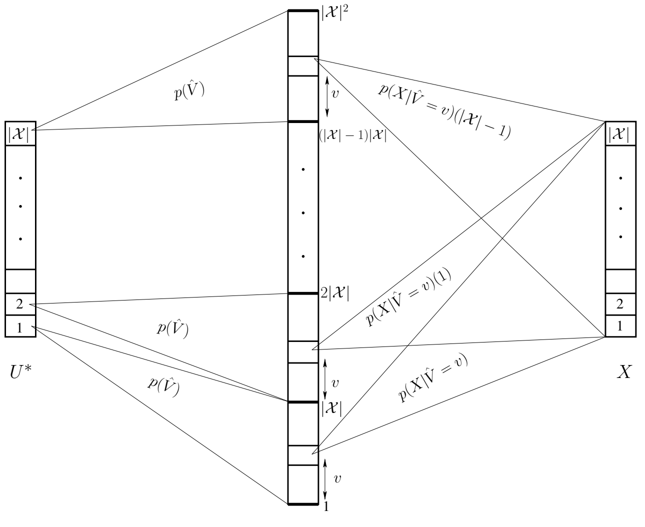

Fig. 3. The structure of the optimal $U^* \rightarrow V^* \rightarrow X$ for cyclic shift symmetric wiretap channels. $p(\hat{V})$ and $p(X|\hat{V} = v)$, $v \in \{1, \ldots, |\mathcal{X}|\}$ are the solutions of the auxiliary optimization problem in (17).

Note that the cyclic shift symmetry assumption on Bob's and Eve's channels yields a significant reduction in the cardinalities of the auxiliary random variables. In particular, the bound on the rate splitting variable reduces from $|\mathcal{X}| + 3$ to $|\mathcal{X}|$ and the bound on the channel prefixing variable reduces from $|\mathcal{X}|^2 + 4|\mathcal{X}| + 3$ to $|\mathcal{X}|^2$. In fact, the problem in (17) for $\mu = 0$ is equivalent to finding the secrecy capacity C_s. Thus, in cyclic shift symmetric wiretap channels, solving a problem of the same number of variables as finding the secrecy capacity is sufficient to characterize the optimal selections of U and V for any point on the boundary of the rate-equivocation region. Another remark is that the constrained optimization problem in (21) for $\mu = 0$ is equivalent to finding the secrecy capacity for general wiretap channels not necessarily cyclic shift symmetric.

The structure of the optimal auxiliary selections U^* and V^* for cyclic shift symmetric wiretap channels in Theorem 2 indicates a sufficient condition for $U = \phi$ to be an optimal selection: If the optimizing $p(\hat{V})$ is uniform and $p(X = x|\hat{V} = v_k)$, $k = 1, \ldots, |\mathcal{X}|$ in (17) is such that

$$p(X = x|\hat{V} = v_k)(k-1) = p(X = x|\hat{V} = v_1), \quad \forall k \tag{22}$$

that is, if the prefix channel $\hat{V} \rightarrow X$ has symmetric transition probabilities and the optimum $p(\hat{V})$ is uniform, then rate splitting is not necessary. In this case, $p(\hat{V})$ is uniform and as $\hat{V}$ has cardinality $|\mathcal{X}|$, it generates a uniform distribution over X. Therefore, selecting $\hat{V}$ uniform with the prefix channel

$\hat{V} \to X$ maximizes $I(X;Y)$ and hence the objective function in (8) (c.f. the proof of Theorem 2 in Appendix). In other words, if (22) is satisfied, then U^* and V^* as selected in (13)-(16) yield a uniform PMF for $p(X|U^* = u)$ for all $u \in \{1, \ldots, |\mathcal{X}|\}$, i.e., U^* is independent of X. Therefore, if (22) is satisfied, $U = \phi$ can be selected without losing optimality, i.e., U is not necessary.

Next, we consider a sub-class of cyclic shift symmetric channels, namely dominantly cyclic shift symmetric channels (c.f. [3, Definition 5]).

Definition 1 *A cyclic shift symmetric wiretap channel is dominantly cyclic shift symmetric if $f(\mathbf{u}) \geq f(P_x)$, $\forall P_x \in \Delta$, where $\mathbf{u}$ is the $|\mathcal{X}|$ dimensional uniform distribution.*

Note that from [11, Theorem 3] and the fact that the uniform distribution is capacity achieving for cyclic shift symmetric channels, a less noisy cyclic shift symmetric wiretap channel is also dominantly cyclic shift symmetric (see also [12]). Based on the analysis in [4], we observe that the solution of (17) satisfies the property in (22) if the wiretap channel is dominantly cyclic shift symmetric.

Lemma 1 *For dominantly cyclic shift symmetric wiretap channels, a solution of the problem in (17) is a selection $p(X = x|\hat{V} = v)$ that satisfies the condition in (22).*

By Lemma 1 and the structure in Theorem 2, $U^* = \phi$ for dominantly cyclic shift symmetric channels and the entire rate-equivocation region can be attained by channel prefixing alone.

We remark here that if the wiretap channel is dominantly cyclic symmetric, then known inner and outer bounds on the corresponding broadcast channel capacity region are shown to coincide in [3]. Therefore, the broadcast channel capacity region, which is in general an open problem, can be fully characterized for dominantly cyclic shift symmetric channels. We observe here that dominant cyclic symmetry yields a similar simplification for the wiretap channel, rendering rate splitting variable U unnecessary. However, note that the class of cyclic shift symmetric wiretap channels for which rate splitting is unnecessary is strictly larger than the class of dominantly cyclic shift symmetric channels. In fact, for all cyclic shift symmetric channels which satisfy (22), $U = \phi$ is optimal and dominant cyclic shift symmetry is just a sufficient but not necessary condition for the property (22). In Section IV, we provide examples for binary-input cyclic shift symmetric wiretap channels that are not dominantly cyclic shift symmetric but for which rate splitting is still unnecessary.

We note that the results obtained for discrete alphabet cyclic shift symmetric channels naturally extend if the alphabets are bounded continuous intervals. In particular, the definition of cyclic shift symmetry extends naturally for $\mathcal{X} = [0, b)$: If $I(X;Y)$ is invariant under any modular shift in the input PDF, the channel is cyclic shift symmetric. Typical examples of continuous alphabet cyclic shift symmetric channels are modulo additive noise channels [13]. If cyclic shift symmetry holds, the channel capacity is achieved at uniform distribution over $\mathcal{X}$. Hence, if both the main and eavesdropping channels are cyclic shift symmetric, then the optimal selections U^* and V^* have the same structure as in Theorem 2. The definition of dominant cyclic shift symmetry also extends similarly for continuous alphabets and rate splitting is not necessary for continuous alphabet dominantly cyclic shift symmetric wiretap channels.

The result does not directly extend for unbounded input alphabets, i.e., for $b = \infty$, with an average power constraint. Even if the cyclic shift symmetry holds, it may not be possible to generate Bob's capacity achieving input PDF by shifting the solution of the auxiliary optimization problem and therefore the proof method in Theorem 2 is not directly applicable.

IV. Binary-Input Cyclic Shift Symmetric Wiretap Channels

In this section, we consider cyclic shift symmetric wiretap channels with binary input: $|\mathcal{X}| = 2$. Note that the cardinality requirement on V to solve the problem in (17) is $|\mathcal{V}| = 2$ for binary input wiretap channels. Let $p(v_1) = \lambda$, $p(x|v_1) = [p_1, 1 - p_1]$ and $p(x|v_2) = [p_2, 1 - p_2]$. Let the resulting input distribution be $P_x = [p_x, 1 - p_x]$. The optimization problem in (17) and (21) for the binary-input case reduces to:

$$\begin{aligned} \max_{\lambda, p_1, p_2} \quad & f(\lambda p_1 + (1-\lambda)p_2) - \lambda f_\mu(p_1) - (1-\lambda) f_\mu(p_2) \\ \text{s.t.} \quad & 0 \leq \lambda, p_1, p_2 \leq 1 \end{aligned} \tag{23}$$

A geometrical visualization for the problem in (23) is provided in Fig. 4. Two points are picked from the x-axis and their image on f_μ are combined to form a line. λ determines the point of operation and the value of the objective function is the difference between f and the formed line segment at that particular point.

The necessary optimality conditions for the problem in (23) are found by taking the derivative of the objective function with respect to p_1, p_2 and λ, respectively. If p_1^*, p_2^*, λ^* are strictly interior to the $[0, 1]$ interval, i.e., not equal to 0 or 1, then

$$\lambda^* \left(f'(\lambda^* p_1^* + (1-\lambda^*)p_2^*) - f'_\mu(p_1^*) \right) = 0 \tag{24}$$

$$(1-\lambda^*) \left(f'(\lambda^* p_1^* + (1-\lambda^*)p_2^*) - f'_\mu(p_2^*) \right) = 0 \tag{25}$$

$$(p_1^* - p_2^*) f'(\lambda^* p_1^* + (1-\lambda^*)p_2^*) - (f_\mu(p_1^*) - f_\mu(p_2^*)) = 0 \tag{26}$$

Note that $\lambda^* \neq 0, 1$ as the objective function in (23) takes the value zero for $\lambda = 0, 1$. Hence, the optimality condition in (26) always holds and we get:

$$f'(\lambda^* p_1^* + (1-\lambda^*)p_2^*) = \frac{f_\mu(p_1^*) - f_\mu(p_2^*)}{p_1^* - p_2^*} \tag{27}$$

If, in addition, $p_1^*, p_2^* \neq 0, 1$,

$$f'(\lambda^* p_1^* + (1-\lambda^*)p_2^*) = f'_\mu(p_1^*) = f'_\mu(p_2^*) \tag{28}$$

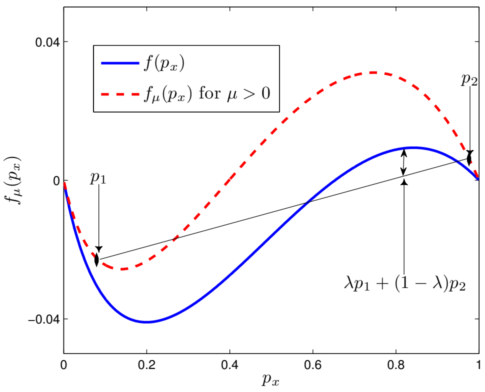

Fig. 4. Visualization of the optimization problem in terms of p_1, p_2 and λ.

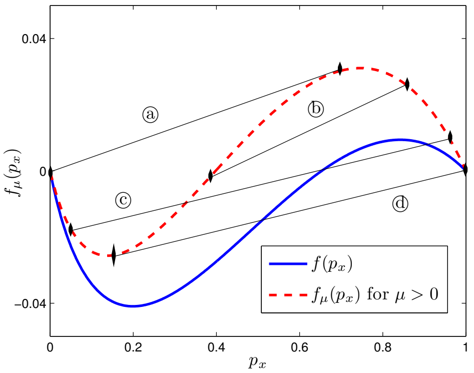

Fig. 5. Optimality conditions for p_1, p_2 and λ. ⓐ and ⓓ satisfy the optimality conditions while ⓑ and ⓒ do not satisfy the optimality condition.

If $p_1^* = 0, 1$ and $p_2^* \neq 0, 1$, then

$$f'(\lambda^* p_1^* + (1-\lambda^*)p_2^*) = f'_\mu(p_2^*) \tag{29}$$

and similarly if $p_2^* = 0, 1$ and $p_1^* \neq 0, 1$, then

$$f'(\lambda^* p_1^* + (1-\lambda^*)p_2^*) = f'_\mu(p_1^*) \tag{30}$$

The conditions in (27)-(30) have the following geometric interpretation: In Fig. 4, the line drawn from $(p_1^*, f_\mu(p_1^*))$ and $(p_2^*, f_\mu(p_2^*))$ must be tangent to the f_μ curve at both points. If p_1^* or p_2^* are 0 or 1, then this tangency does not have to hold at that point. We illustrate these conditions in Fig. 5. We observe that for the selections of p_1 and p_2 in the configurations ⓐ and ⓓ, the line is tangent to $f_\mu(p_x)$ at only one point and the other point is either 0 or 1. However, ⓑ and ⓒ do not satisfy the optimality condition as the p_1 and p_2 points lie interior to $[0, 1]$ but the line is not tangent to f_μ. In fact, we observe by inspection that ⓐ and ⓓ are the only possible configurations in the particular values chosen in Fig. 5 that satisfy the optimality conditions in (27)-(30).

Note that the geometric interpretation of the optimality conditions provide a simple check if a point $p \in (0, 1)$ is one of the optimal selections p_1^*, p_2^*: Draw the tangent line for f_μ at p. If this tangent line does not intersect f_μ other than p or if it intersects at a point $p' \in (0, 1)$ but it is not tangent at p', then p cannot be an optimal selection. Also note that optimality conditions do not rule out the trivial selection $p_1 = 0$ and $p_2 = 1$. Hence, this selection is always a candidate to be an optimal selection. This selection is indeed optimal if $f(p_x) \geq 0$ for all $p_x \in [0, 1]$, i.e., when the wiretap channel is more capable [4]. Geometrically, when $f(p_x) \geq 0$ the points $(p_i, f_\mu(p_i))$ have nonnegative y-axis as $f_\mu(p_x) \geq f(p_x) \geq 0$. The level of the line segment is the smallest when $p_1 = 0$, $p_2 = 1$. Moreover, the points $\lambda p_1 + (1-\lambda)p_2$ span the space of two-dimensional PMFs when $p_1 = 0$ and $p_2 = 1$. Hence, the difference of the function and the line segment has the highest value when $p_1 = 0$ and $p_2 = 1$, that is, the optimal selection is $p_1 = 0$ and $p_2 = 1$.

A. The BSC-BEC Wiretap Channel

Let the main channel be BSC(ϵ) and the eavesdropper's channel be BEC(α). Note that both BSC and BEC are cyclic shift symmetric. $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ and $\mathcal{Z} = \{0, e, 1\}$. It can be shown [3] that when $p(y|x)$ is BSC(ϵ) and $p(z|x)$ is BEC(α):

1) If $\alpha < 4\epsilon(1-\epsilon)$, then $f(p_x)$ is convex.
2) If $4\epsilon(1-\epsilon) \leq \alpha \leq h(\epsilon)$, $f(p_x)$ is non-convex and $f(p_x) \leq 0$.
3) If $h(\epsilon) < \alpha$, then $f(p_x)$ is maximized at $p_x = 0.5$.

Rate splitting is not necessary [4] for $\alpha > h(\epsilon)$ as $f(p_x)$ is maximized at $p_x = 0.5$ when $\alpha > h(\epsilon)$. Moreover the required channel prefixing has $|\mathcal{V}^*| = 2$ with $p(v_1) = p(v_2) = 1/2$ and $p(x|v_1) = [a, 1-a]$, $p(x|v_2) = [1-a, a]$ where $[a, 1-a]$ is an input distribution that maximizes $[(\mu+1)I(X;Y) - I(X;Z)]$ [4]. For $\alpha < 4\epsilon(1-\epsilon)$, Eve is less noisy than Bob, and the secrecy capacity is $C_s = 0$. We investigate the remaining case, which is $4\epsilon(1-\epsilon) \leq \alpha \leq h(\epsilon)$, in the next subsection.

1) The Case of $4\epsilon(1-\epsilon) \leq \alpha \leq h(\epsilon)$: When $4\epsilon(1-\epsilon) \leq \alpha \leq h(\epsilon)$ in the BSC-BEC channel, neither Eve is less noisy nor the dominant cyclic shift symmetry holds. Secrecy capacity is still non-zero in this case and the rate-equivocation region has a non-empty interior. One can verify easily by tracing all points in $[0, 1]$ that there are only 5 configurations that satisfy the necessary optimality conditions as well as the trivial selection $V = X$, i.e., $p_1 = 0$ and $p_2 = 1$. However, the trivial selection is immediately eliminated as $f_\mu(p_x) \leq 0$ for some p_x in this case, and hence $V = X$ is strictly suboptimal [4].

The other 5 configurations are shown in Fig. 6: In configurations ⓐ, ⓑ, ⓒ and ⓓ, either p_1 or p_2 is on the boundary and in configuration ⓔ, both p_1 and p_2 are in the interior with the property $p_1 \in \arg\min_{p_x \in [0,1]} f_\mu(p_x)$ and $p_2 = 1 - p_1$. By comparing these three configurations, we observe that the optimum selection is always configuration ⓔ. In other words, we have for $\mu \geq 0$ and for all $0 \leq \lambda, p_1, p_2 \leq 1$,

$$\begin{aligned} & f(0.5) - \min_{p_x} f_\mu(p_x) \\ & \geq f(\lambda p_1 + (1-\lambda)p_2) - \lambda f_\mu(p_1) - (1-\lambda) f_\mu(p_2) \end{aligned} \tag{31}$$

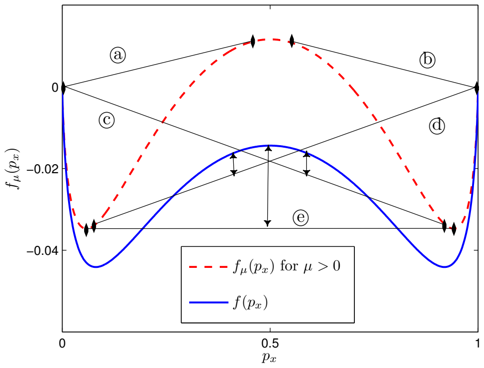

Fig. 6. Five configurations that satisfy optimality conditions. ⓔ is always the optimal.

Note that ⓔ has the desirable property that cyclic shift of $[p_1, 1-p_1]$ is $[p_2, 1-p_2]$, i.e., $p_2 = 1-p_1$. This property is equivalent to the one in (22) in the binary-input case. Therefore, $U = \phi$ is optimal, and the upper right boundary of the rate-equivocation region can be traced by V only. However, unlike the case of $h(\epsilon) \leq \alpha$, if $4\epsilon(1-\epsilon) \leq \alpha \leq h(\epsilon)$, there exists $\mu \geq 0$ such that $f(0.5) < \min_{p_x} f_\mu(p_x)$. We define

$$\mu^* = \min\{\mu : f(0.5) \leq \min_{p_x} f_\mu(p_x)\} \tag{32}$$

For $\mu > \mu^*$, V defined as above cannot improve the objective function. Thus, trivial V is the optimal selection for $\mu > \mu^*$. However, the highest achievable equivocation with a trivial V selection is zero as Eve's channel is more capable with respect to Bob's channel in this case. Hence, for $\mu > \mu^*$, the only possible achievable point is $(C_B, 0)$. The general form of the rate-equivocation region is given in Fig. 7. The upper right boundary includes the line segment that combines the point for which the supporting line slope is μ^* and the $(C_B, 0)$ point. This line segment has the slope μ^*.

In conclusion, rate splitting U is not necessary for determining the rate-equivocation region of the BSC-BEC wiretap channel and in particular the secrecy capacity is

$$C_s = f(0.5) - \min_{p_x} f(p_x) \tag{33}$$

B. The BEC-BSC Wiretap Channel

Now, let the main channel be BEC(α) and the eavesdropper's channel be BSC(ϵ). $\mathcal{X} = \mathcal{Z} = \{0, 1\}$ and $\mathcal{Y} = \{0, e, 1\}$. We have the following facts [3]:

1) If $\alpha < 4\epsilon(1-\epsilon)$, then Bob is less noisy than Eve.
2) If $4\epsilon(1-\epsilon) \leq \alpha \leq h(\epsilon)$, then Bob is more capable but not less noisy than Eve.

and the rate-equivocation region is characterized for the above cases in [4]. We investigate the remaining case, which is $\alpha \geq h(\epsilon)$ in the next subsection.

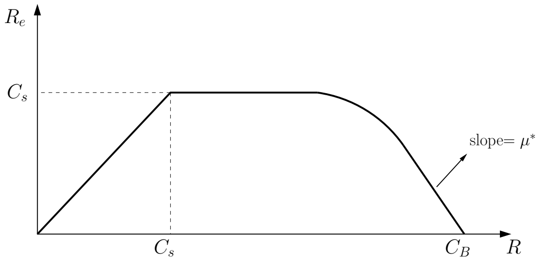

Fig. 7. The general form of the rate-equivocation region of the BSC-BEC wiretap channel for $4\epsilon(1-\epsilon) \leq \alpha \leq h(\epsilon)$.

1) The Case of $\alpha \geq h(\epsilon)$: In the BEC-BSC wiretap channel, if $\alpha \geq h(\epsilon)$, neither less noisy nor more capable condition holds. We first solve the optimization problem in (23) by inspecting the tangent lines drawn at interior points $p \in (0, 1)$. One can easily verify that, as in the BSC-BEC channel, there are only 5 possible configurations that satisfy the necessary optimality conditions in addition to the trivial selection $p_1 = 0$ and $p_2 = 1$. As Bob's channel is not more capable with respect to Eve in this case and $f(p_x)$ is maximized at an interior point and the trivial selection $V = X$ is strictly suboptimal in this case [4]. In particular, the trivial selection is not optimal for all μ such that $f_\mu(p_1) < 0$ for some $p_1 \in [0, 1]$. We show the other 5 configurations in Fig. 8. In configuration ⓔ, $p_1, p_2 \in (0, 1)$ with $f_\mu(p_1) = f_\mu(p_2)$ and $f'_\mu(p_1) = f'_\mu(p_2) = 0$. Hence, it satisfies the optimality condition for $\lambda = 0.5$. However, the objective function $f(0.5) - 0.5\,(f_\mu(p_1) + f_\mu(p_2)) < 0$; therefore, this configuration cannot be optimal. The other configurations ⓐ, ⓑ, ⓒ and ⓓ have p_1 or p_2 on the boundary of $[0, 1]$ interval as shown in Fig. 8. We observe that ⓐ and ⓑ achieve the same value of the objective function and it is always higher compared to that achieved by ⓒ and ⓓ. Therefore, ⓐ and ⓑ are optimal selections for the problem in (23). Note that ⓐ is obtained by cyclic shifts of ⓑ. The configuration in ⓐ is also represented as $p(X = 0|V = v_1) = 0$ and $p(X = 0|V = v_2) = p_1$ where the line segment that combines $(0, 0)$ and $(p_1, f_\mu(p_1))$ is tangent to the curve $(p_x, f_\mu(p_x))$. Similarly, ⓑ is equivalent to $p(X = 0|V = v_1) = 1$ and $p(X = 0|V = v_2) = 1 - p_1$. The rate equivocation region is traced by varying μ and finding p_1 that satisfies the tangency and λ^* that yields the optimal value of the objective function given p_1. In particular, we define

$$\mu^* = \min\{\mu \geq 0 | \min_{p_x} f_\mu(p_x) \geq 0\} \tag{34}$$

For $\mu \leq \mu^*$, we use the following U and V:

$$p(U = u_1) = p(U = u_2) = 0.5 \tag{35}$$

$$p(V = v_1|U = u_1) = \lambda^*, \quad p(V = v_2|U = u_1) = 1 - \lambda^*, \tag{36}$$

$$p(V = v_3|U = u_2) = \lambda^*, \quad p(V = v_4|U = u_2) = 1 - \lambda^*, \tag{37}$$

$$p(X = 0|V = v_1) = 0, \quad p(X = 0|V = v_2) = p_1, \tag{38}$$

$$p(X = 0|V = v_3) = 1, \quad p(X = |V = v_4) = 1 - p_1 \tag{39}$$

For $\mu > \mu^*$, V is not necessary as $f_\mu(p_x) \geq 0$ in this case. We obtain a case similar to the more capable condition and one can easily show that a non-trivial V does not improve the objective function. As V is not used for $\mu > \mu^*$, the achieved rate $I(V; Y) = I(X; Y)$ and optimal selection of U

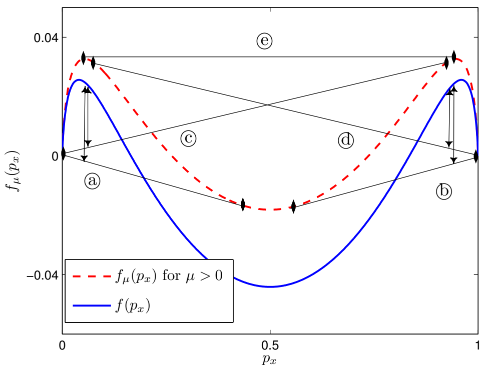

Fig. 8. Five configurations that satisfy the necessary optimality conditions for the BEC-BSC wiretap channel. ⓐ and ⓑ are both optimal.

as in Theorem 2 generates uniform distribution on the channel input X, which is capacity achieving for Bob's channel. Hence, for $\mu > \mu^*$, C_B is achieved. The general form of the rate-equivocation region is depicted in Fig. 9. Note that the supporting line with slope μ^* is on the boundary of the rate-equivocation region.

V. Conclusions

In this paper, we studied cyclic shift symmetric wiretap channels. We explicitly determined the optimal rate splitting and channel prefixing variables that trace the rate-equivocation region. We showed that in cyclic shift symmetric wiretap channels, it suffices to solve an optimization problem only over one auxiliary random variable, which significantly reduces the amount of calculations needed and the required cardinalities of the auxiliary random variables. We formulated the problem as a constrained optimization problem and we applied our results to binary-input cyclic shift symmetric wiretap channels for which we solve a three-variable constrained optimization problem. We characterized the rate-equivocation regions of the BSC-BEC and BEC-BSC wiretap channels.

Appendix

For given $\mu \geq 0$, the optimal selections U^* and V^* are the solutions of the following optimization problem:

$$\max_{U \to V \to X \to Y,Z} \mu I(V;Y) + I(V;Y|U) - I(V;Z|U) \qquad (40)$$

By using the fact that $I(V;Y) = I(X;Y) - I(X;Y|V)$, we obtain an equivalent statement for (40) as:

$$\max_{U \to V \to X \to Y,Z} \mu I(X;Y) + I(X;Y|U) - I(X;Z|U) \\ - [(\mu+1)I(X;Y|V) - I(X;Z|V)] \qquad (41)$$

We have the following for the objective function in (41):

$$\mu I(X;Y) + I(X;Y|U) - I(X;Z|U) \\ -[(\mu+1)I(X;Y|V) - I(X;Z|V)]$$

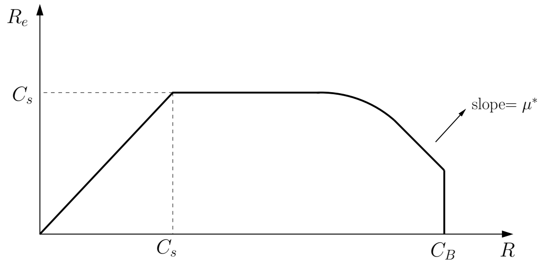

Fig. 9. General form of the rate-equivocation region of the BEC-BSC wiretap channel when $\alpha \geq h(\epsilon)$.

$$\leq \max_{P_x} \mu I(X;Y) \\ + \max_{U \to V \to X \to Y,Z} I(X;Y|U) - I(X;Z|U) \\ - [(\mu+1)I(X;Y|V) - I(X;Z|V)] \qquad (42)$$

$$\leq \mu I_{\mathbf{u}}(X;Y) + \max_{\hat{V} \to X \to Y,Z} I(X;Y) - I(X;Z) \\ - [(\mu+1)I(X;Y|\hat{V}) - I(X;Z|\hat{V})] \qquad (43)$$

where $\mathbf{u}$ denotes the $|\mathcal{X}|$ dimensional discrete uniform random variable, and $I_{\mathbf{u}}(X;Y)$ denotes the mutual information obtained by choosing the PMF of X as $\mathbf{u}$. In (43), we used the fact that $\max_{P_x} I(X;Y) = I_{\mathbf{u}}(X;Y)$ as Bob's channel is cyclic shift symmetric. Moreover, we used the fact that U is not needed, i.e., $U = \phi$, for maximizing $I(X;Y|U) - I(X;Z|U) - [(\mu+1)I(X;Y|V) - I(X;Z|V)]$. Because, for given $U \to V \to X \to Y, Z$, we can always pick $u_i \in \mathcal{U}$ that maximizes

$$\max_{u_i \in \mathcal{U}} I(X;Y|U=u_i) - I(X;Z|U=u_i) \\ - [(\mu+1)I(X;Y|V,U=u_i) - I(X;Z|V,U=u_i)] \qquad (44)$$

and therefore, choose a deterministic U with $U = u^*$, where u^* is the argument of the maximization in (44). Note that the last maximization in (43) is the claimed auxiliary optimization problem in the statement of the theorem. We use $\hat{V}$ notation to emphasize that the channel prefixing auxiliary random variables in (42) and (43) are different.

Next, we will show that the bound in (43) is satisfied with equality for any cyclic shift symmetric wiretap channel. Let $\hat{V}$ with $p(\hat{V} = v)$ and $p(X|\hat{V} = v)$, $v \in \mathcal{V}$, be the solution of the auxiliary problem in (42). First, we note that it suffices to consider $\hat{V}$ such that $|\hat{\mathcal{V}}| \leq |\mathcal{X}|$. Given $\hat{V} \to X \to Y, Z$, we fix $|\mathcal{X}| - 1$ components of $P_{X|\hat{V}}(x|\hat{v})$, $j = 1, \ldots, |\mathcal{X}| - 1$, together with $(\mu+1)I(X;Y|\hat{V}) - I(X;Z|\hat{V})$. By Lemma 3 in [8] and the strengthened Caretheodory theorem of Fenchel-Eggleston in [9], the problem in (42) can be solved with the cardinality bound $|\hat{\mathcal{V}}| \leq |\mathcal{X}|$. Note the equivalence of the operations performed for proving the bound in this problem and those in (5).

Next, we will show that the bound in (43) is satisfied with equality for any cyclic shift symmetric wiretap channel by a specific structure of U^*, V^* in terms of the optimal $\hat{V}$ for the auxiliary problem in (42) as in the statement of the theorem. In particular, we select U^* and V^* as in the statement of Theorem 2 in (13)-(16) and as depicted in Fig. 3 with cardinalities $|\mathcal{U}^*| = |\mathcal{X}|$ and $|\mathcal{V}^*| = |\mathcal{X}|^2$. Each element of U^* generates the optimizing selection $p(\hat{V})$ for the problem in (42) over disjoint $|\mathcal{X}|$ elements of V^*. Each $|\mathcal{X}|$-element block of V^* generates cyclic shifts of the optimizing selection $p(X|\hat{V})$ for the input X. In (13)-(16), we denote the kth cyclic shift of the conditional PMF for the channel input X, $p(x|\hat{V}=v)$, as $p(x|\hat{V}=v)(k)$. Note that the cardinality of $\hat{V}$ is $|\mathcal{X}|$ while that of the optimum V^* is $|\mathcal{X}|^2$ and $|\mathcal{X}|^2$ conditional input PMFs, $p(x|V^*=v)$, are obtained by cyclic shifts of $|\mathcal{X}|$ conditional input PMFs, $p(x|\hat{V}=v)$.

We first observe that $p(x|U^*=i)$ are cyclic shifts of a fixed PMF over X for different i. In particular, in the construction in (13)-(16), we selected $p(x|V^*=v)$ as cyclic shifts of $p(x|\hat{V}=v)$ while we kept $p(V^*=v)$ the same as $p(\hat{V}=v)$. Hence, we have

$$p(x|U^*=i) = p_f(x)(i-1) \tag{45}$$

where $p_f(x) = \sum_{v=1}^{|\mathcal{X}|} p(x|\hat{V}=v)p(\hat{V}=v)$. Note that $p_f(x)$ is the maximizing input PMF for the auxiliary problem. Therefore, U^* and V^* generate a uniform PMF for X:

$$p(x) = \sum_{i=1}^{|\mathcal{X}|} p(U^*=i)p(x|U^*=i) = \sum_{i=1}^{|\mathcal{X}|} \frac{1}{|\mathcal{X}|} p_f(x)(i-1) \tag{46}$$

$$= \frac{1}{|\mathcal{X}|} \tag{47}$$

Moreover, by construction of U^* and V^* and the cyclic shift symmetry of the channels, we observe that, for any given i,

$$\begin{aligned}&\sum_{v=1}^{|\mathcal{X}|}\Big[(\mu+1)I(X;Y|V^*=v+(i-1)|\mathcal{X}|) \\ &\qquad - I(X;Z|V^*=v+(i-1)|\mathcal{X}|)\Big] \\ &\qquad p(V^*=v+(i-1)|\mathcal{X}||U^*=i) \\ &= \sum_{v=1}^{|\mathcal{X}|}\Big[(\mu+1)I(X;Y|\hat{V}=v) - I(X;Z|\hat{V}=v)\Big]p(\hat{V}=v) \end{aligned} \tag{48}$$

$$= (\mu{+}1)I(X;Y|\hat{V}) - I(X;Z|\hat{V}) \tag{49}$$

Therefore, we have

$$\begin{aligned}&(\mu+1)I(X;Y|V^*) - I(X;Z|V^*) \\ &\quad = \sum_{v=1}^{|\mathcal{X}|^2}\Big((\mu+1)I(X;Y|V^*=v) \\ &\qquad - I(X;Z|V^*=v)\Big)p(V^*=v) \end{aligned} \tag{50}$$

$$\begin{aligned}&\quad = \sum_{i=1}^{|\mathcal{X}|}\sum_{v=1}^{|\mathcal{X}|^2}\Big((\mu+1)I(X;Y|V^*=v) \\ &\qquad - I(X;Z|V^*=v)\Big)p(v|U^*=i)p(U^*=i) \end{aligned} \tag{51}$$

$$\begin{aligned}&= \frac{1}{|\mathcal{X}|}\sum_{i=1}^{|\mathcal{X}|}\sum_{v=1}^{|\mathcal{X}|}\Big((\mu+1)I(X;Y|V^*=v+(i-1)|\mathcal{X}|) \\ &\qquad - I(X;Z|V^*=v+(i-1)|\mathcal{X}|)\Big)p(\hat{V}=v) \end{aligned} \tag{52}$$

$$= (\mu+1)I(X;Y|\hat{V}) - I(X;Z|\hat{V}) \tag{53}$$

where (53) is obtained by using (49) and the fact that $p(v|U^*=i)$ is non-zero only for $(i-1)|\mathcal{X}|+1 \leq v \leq i|\mathcal{X}|$. Note that

$$I(X;Y|U^*=i) - I(X;Z|U^*=i) = f(p_f(x)(i-1)) \tag{54}$$

$$= f(p_f(x)), \quad \forall i \tag{55}$$

Hence, given $U^*=i$, we have

$$\begin{aligned}&I(X;Y|U^*=i) - I(X;Z|U^*=i) \\ &\qquad - \Big[(\mu+1)I(X;Y|V^*) - I(X;Z|V^*)\Big] \\ &= f(p_f(x)) - \Big[(\mu+1)I(X;Y|\hat{V}) - I(X;Z|\hat{V})\Big] \end{aligned} \tag{56}$$

As $p_f(x)$ is the maximizing input PMF for the auxiliary problem, we have

$$\begin{aligned}&I(X;Y|U^*) - I(X;Z|U^*) \\ &\qquad - [(\mu+1)I(X;Y|V^*) - I(X;Z|V^*)] \\ &= \max_{\hat{V}\to X} I(X;Y) - I(X;Z) \end{aligned} \tag{57}$$

$$-[(\mu+1)I(X;Y|\hat{V}) - I(X;Z|\hat{V})] \tag{58}$$

Since U^* and V^* generate a uniform PMF for X by (47), $I(X;Y)$ achieves its maximum, as well. Combining this with (58), we conclude that the constructed U^* and V^* achieve the upper bound in (43) and hence are optimal.

References

[1] A. Wyner. The wire-tap channel. *Bell Sys. Tech. Journal*, 54(8):1355–1387, October 1975.

[2] I. Csiszár and J. Körner. Broadcast channels with confidential messages. *IEEE Trans. on Inform. Theory*, 24(3):339–348, May 1978.

[3] C. Nair. Capacity regions of two new classes of two-receiver broadcast channels. *IEEE Trans. on Inform. Theory*, 56(9):4207–4214, September 2010.

[4] O. Ozel and S. Ulukus. Wiretap channels: Roles of rate splitting and channel prefixing. In *IEEE ISIT*, July 2011.

[5] B. Xie and R. Wesel. A mutual information invariance approach to symmetry in discrete memoryless channels. In *UCSD ITA*, Feb. 2008.

[6] I. Csiszár. Almost independence and secrecy capacity. *Prob. of Inform. Trans.*, 32(1):48–57, 1996.

[7] U. M. Maurer and S. Wolf. Information-theoretic key agreement: from weak to strong secrecy for free. In *EUROCRYPT*, 2000.

[8] R. Ahlswede and J. Körner. Source coding with side information and a converse for degraded broadcast channels. *IEEE Trans. on Inform. Theory*, IT-21(6):629–638, November 1975.

[9] M. Salehi. Cardinality bounds on auxiliary variables in multiple user theory via the method of Ahlswede and Körner. *Technical Report No. 33*, August 1978, Stanford Univ., Stanford, CA.

[10] A. El Gamal and Y.-H. Kim. Lecture notes on network information theory. Available at [ArXiv:1001.3404].

[11] M. van Dijk. On a special class of broadcast channels with confidential messages. *IEEE Trans. on Inform. Theory*, 43(2):712–714, March 1997.

[12] S. K. Leung-Yan-Cheung. On a special class of wiretap channels. *IEEE Trans. on Inform. Theory*, 23(5):625–627, September 1977.

[13] U. Erez and R. Zamir. Noise prediction for channels with side information at the transmitter. *IEEE Trans. on Inform. Theory*, 46(4):1610–1617, July 2000.