Generalizing Multiple Access Wiretap and Wiretap II Channel Models: Achievable Rates and Cost of Strong Secrecy

Mohamed Nafea, *Member, IEEE*, and Aylin Yener, *Fellow, IEEE*

***Abstract*— In this paper, new two-user multiple access wiretap channel models are studied. First, the multiple access wiretap channel II with a discrete memoryless main channel under different wiretapping scenarios is introduced. The wiretapper, as in the classical wiretap channel II model, chooses a fixed-size subset of the channel uses, in which it obtains noise-free observations of one of the codewords: a deterministic function, e.g., superposition, of the two codewords or each of the two codewords. A fourth wiretapping scenario is considered, in which the wiretapper, in each position it chooses, decides to observe either one of the codewords or both codewords, with an overall budget on the number of its noiselessly observed symbols. These, thus, extend the recently examined wiretap channel II with a noisy main channel to a multiple access setting with a variety of attack models for the wiretapper. Next, the proposed multiple access wiretap channel II models are further generalized to the case when the wiretapper observes the outputs of a discrete memoryless channel, instead of erasures, outside the subset of noiseless observations. Achievable strong secrecy rate regions for all the proposed models are derived. Achievability is established by solving dual multi-terminal secret key agreement problems in the source model and converting the solution to the original channel models using probability distribution approximation arguments. The derived achievable rate regions quantify the secrecy cost due to the additional capabilities of the wiretapper with respect to the previous multiple access wiretap models.**

***Index Terms*— Multiple access wiretap channel, wiretap channel II, new wiretap channel models, strategic adversaries, strong secrecy, source-channel duality, random binning, concentration inequalities.**

I. Introduction

THE wiretap channel II in which the legitimate terminals communicate over a noiseless channel and the wiretapper has perfect access to a fixed fraction of its choosing of the transmitted bits, has been introduced in [3]. This model, while similar to a classical wiretap channel [4], [5] with a noiseless main channel and a binary erasure wiretapper channel, models a more powerful wiretapper which is able to select the positions of erasures. Reference [3] has shown that the secrecy capacity of the wiretap channel II model does not increase if the wiretapper is a passive observer with a binary erasure channel whose erasures are randomly chosen by nature.

Considerable amount of research on practical code design for secrecy has been motivated by the coset coding scheme devised in [3], see for example [6]–[10]. However, for several decades, there has been no effort for generalizing the wiretap II model outside the special scenario of the noiseless main channel. Recently, [11] has introduced a discrete memoryless main channel to the wiretap cahnnel II, and derived inner and outer bounds for its capacity-equivocation region. Reference [12] has characterized the secrecy capacity of this model, showing that, once again, the secrecy capacity does not increase when the more powerful wiretapper is replaced with an erasure channel.

More recently, [13] has introduced the generalized wiretap channel model and identified its secrecy capacity. In this model, the main channel is a discrete memoryless channel while the wiretapper, besides noiselessly observing a subset of its choice of the transmitted codeword symbols, observes the remainder through a discrete memoryless channel. This new model subsumes both the classical wiretap channel [5] and the wiretap channel II with a discrete memoryless main channel [11] as its special cases. The secrecy capacity of this generalized model quantifies the secrecy penalty of the additional capability at the wiretapper with respect to the previous wiretap models. In addition, the results in [3], [11]–[13] demonstrate the immunity of wiretap (stochastic) encoding against a more powerful wiretapper which is able to choose where to tap.

The notion of a "strategic adversary", inspired by the wiretap II model, in which the adversary designs a *partial* attack by monitoring, modifying, and/or corrupting a subset of its choice of either the legitimate communication or the information contents at legitimate nodes in the network, has attracted a wide spectrum of research in recent years. This strategic adversary has been considered in several problems such as secure network coding [14], [15], secure distributed storage systems [16], [17], active adversarial attacks [9], and adversarial erasure channels [18].

Manuscript received January 31, 2018; revised January 3, 2019; accepted March 11, 2019. Date of publication April 2, 2019; date of current version July 12, 2019. This work was supported in part by the National Science Foundation under Grant CNS 13-14719. This paper was presented in part at the 2016 IEEE International Symposium on Information Theory [1] and the 2016 IEEE Information Theory Workshop [2].

M. Nafea was with the Department of Electrical Engineering, Pennsylvania State University at University Park, State College, PA 16802 USA. He is now with the Electrical and Computer Engineering Department, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: mnafea3@gatech.edu).

A. Yener is with the Department of Electrical Engineering, Pennsylvania State University at University Park, State College, PA 16802 USA (e-mail: yener@ee.psu.edu).

Communicated by A. Tchamkerten, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2019.2908832

These developments reflect on the wide scope of application for the strategic adversary model in the wiretap channel II. The wiretap channel II provides an intermediate ground for modeling adversarial capabilities between either overly optimistic or overly pessimistic assumptions with regard to what can be known about the adversary. In particular, theoretical frameworks developed so far assume either a *purely passive observer* as in Wyner's classical wiretap channel and its multi-terminal extensions [19]–[25]; or an *over-powerful* arbitrarily varying wiretap model [26]–[29] which captures complete uncertainty about the adversary's jamming and eavesdropping. Being a middle ground between these two extremes makes the wiretap channel II model interesting to consider and motivates further investigation of the model. Generalizing the wiretap channel II outside the point-to-point setting and investigating the impact of *designed adversarial attacks* in multi-terminal networks is the natural next step, which is the focus of this work.

In this paper, we extend the generalized wiretap channel in [13] to the multiple access scenario [21]. In particular, we first consider the special case of the multiple access wiretap channel II with a discrete memoryless main channel, and propose three different attack models for the wiretapper. In each of these models, the wiretapper chooses a fixed-size subset of the channel uses and observes erasures outside this subset. In the first wiretapping model, the wiretapper, in each position of the subset, decides to observe either the first or the second user's symbol. In the second model, the wiretapper observes a deterministic function, e.g., superposition, of the two transmitted symbols in the positions of the subset, while in the third model, the wiretapper observes the transmitted symbols of both users. Further, we consider a fourth wiretapping model in which the wiretapper chooses two subsets of the channel uses whose overall size is fixed, and observes the transmitted symbols of the first (second) user in the positions of the first (second) subset, and erasures otherwise. The fourth wiretapping model thus generalizes the wiretapper's strategy space in the first and third models in which the wiretapper is restricted to choose two non-overlapping and identical subsets, respectively.

The first attack model is a setting in which the wiretapper is able to tap one of the two transmissions but not both. For instance, if two transmitters are distant from each other, the wiretapper may need to get close to one in order to obtain noise-free observations, and thus is able to tap one at a time. The second attack model mimics a medium that, for instance, superposes both transmissions (e.g., wireless), where the attacker is close enough to both transmitters. In the third attack model, the wiretapper is able to tap both codewords individually in the same positions, which can be interpreted as the wiretapper being able to obtain noiseless (partial) side information about both transmitted codewords. The fourth attack model is an intermediate setting between the first and third models, where the wiretapper is able to choose between tapping the two codewords at two identical, overlapping, or distinct sets of positions.

For each of these models, we derive an achievable strong secrecy rate region. Even though the third attack model, in which the wiretapper sees the transmitted symbols of *both* users, is stronger than the first, the ability of the wiretapper in the first model to choose which user's symbol to tap into results in identical achievable strong secrecy rate regions for the two models. That is, each transmitter designs their encoding according to the worst case scenario in which the wiretapper chooses to see its symbols in all positions of the subset. Similarly, in the fourth attack model, the worst case scenario for each transmitter is when the wiretapper spends all of its budget on observing only its symbols, resulting in an achievable rate region that is identical to the first and third models. The achievable secrecy rate region for the second attack model is shown to be larger than the achievable secrecy rate region for the other three models, demonstrating the intrinsic cooperation introduced by the medium.

After obtaining these insights, we generalize these models by replacing the wiretapper's erasures with noisy channel outputs as was done in [13] for the single user channel. In particular, we generalize the multiple access wiretap channel II with a discrete memoryless main channel under the proposed wiretapping scenarios to the case when the wiretapper observes the remainder of the codewords of both users separately through a discrete memoryless channel. The generalized multiple access wiretap channel, under the third wiretapping scenario, also generalizes the multiple access wiretap channel in [20], [21], [30] to the case when the wiretapper is provided with a subset of its choosing of noise-free observations of both users symbols. Achievable strong secrecy rate regions which quantify the secrecy cost, with respect to the multiple access wiretap channel, of the additional capabilities of the wiretapper in these generalized models are derived.

Achievability of the strong secrecy rate regions for all the proposed models is established by muti-terminal extensions of methods in [13], [31], [32]. In particular, for each of the proposed models, a corresponding dual multi-terminal secret key agreement problem in the source model is introduced. In this dual model, two independent sources wish to agree on two indepedent keys with a common decoder in the presence of a *compound* wiretapping source. We solve the problem in the dual source model, and convert the solution to the original channel model by means of deriving the joint distributions of the two problems to become almost identical, in the total variation distance sense. The technical challenge in the present paper lies in generalizing the tool utilized for establishing secrecy of the key in the dual source model from the single source case, [13, Lemma 2], to the case of two *independent* sources. This is done by adapting the lemma in order to establish all the *corner* (extreme) points of the rate region for the two keys, generated at the independent sources, such that the convergence rate for the probability of the two keys being independent from the wiretapper's observation is *doubly-exponential*. Time sharing between the resulting corner points produces the desired rate region. This doubly-exponential convergence rate is needed in order to *exhaust* the exponentially many possible strategies for the wiretapper [12], [13].

Overall, the contributions of this paper are summarized as follows:

1) We introduce the multiple access wiretap channel II with a noisy main channel under different wiretapping scenarios which feature different adversarial capabilities and different transmission media. In these models, the wiretapper chooses subset(s) of noise-free observations, in different forms, and observes erasures outside the subset(s) it chooses.
2) We derive achievable strong secrecy rate regions for all the proposed models, and highlight the insights drawn from the derived regions.
3) We further generalize the proposed multiple access wiretap channel II models to the case when the wiretapper observes noisy outputs, instead of erasures, outside the subset(s) of noiseless observations. We derive achievable strong secrecy rate regions for these generalized models and quantify the secrecy cost of the additional capabilities at the wiretapper.
4) In order to derive the achievable strong secrecy rate regions in this paper, we generalize a one-shot lemma which provides a doubly-exponential convergence rate for the security measure in a single-user source-modeled secret key agreement problem [13, Lemma 2] to the case of two sources. Generalizing this lemma to directly obtain a rate region for the keys generated at the sources is troublesome. Instead, we adapt the lemma to obtain corner points of the rate region and deduce the desired region by time sharing between these points.

The remainder of the paper is organized as follows. Section II provides the notation and definitions. Section III describes the channel models considered in this paper. Section IV presents the main results. The proofs of the results are presented in Sections V and VI. Section VII provides a discussion about the main results and the utilized achievability approach. Section VIII concludes the paper.

II. Notation

We remark the notation we use throughout the paper. Vectors are denoted by bold lower-case super-scripted letters while their components are denoted by lower-case sub-scripted letters. A similar convention but with upper-case letters is used for random vectors and their components. $\mathcal{A}_1 \times \mathcal{A}_2$ denotes the Cartesian product of the sets $\mathcal{A}_1$ and $\mathcal{A}_2$. We use $\mathbb{1}\{\mathcal{A}\}$ to denote the indicator function of the event $\mathcal{A}$. For $a, b \in \mathbb{R}$, $[a : b]$ denotes the set of integers $\{i \in \mathbb{N} : a \leq i \leq b\}$. For a sequence of random variables (vectors) $A_1, \cdots, A_n$, we use $A_{[i:j]}$ to denote the sub-sequence $\{A_i, \cdots, A_j\}$, where $1 \leq i < j \leq n$. We also use $A_S \triangleq \{A_i\}_{i \in S}$ for any $S \subseteq [1 : n]$. For a set S, where $S \subseteq [1 : n]$, we use $|S|$ to denote its cardinality and S^c to denote its complement.

Probability distribution of a random variable X taking values from the countable set $\mathcal{X}$ is denoted by lower-case letters, such as p_X or $\tilde{p}_X$. We use upper-case letters to denote a random probability distribution, e.g., P_X or $\tilde{P}_X$, which is defined as follows:

Definition 1: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where Ω is the sample space, $\mathcal{F}$ is the σ-algebra of events, and $\mathbb{P}$ is the probability measure. For any countable set $\mathcal{X}$, let $\Delta_{\mathcal{X}}$ be the simplex of probability distributions over $\mathcal{X}$. A random probability distribution P_X is a random vector which maps the sample space Ω to the measurable space $\Delta_{\mathcal{X}}$, i.e., $P_X : \omega \in \Omega \mapsto P_X(.; \omega)$. That is, for every $x \in \mathcal{X}$, the mapping $\omega \in \Omega \mapsto P_X(x; \omega)$ is a random variable. A random probability distribution can as well be viewed as an indexed family of distributions with a certain probability distribution over the index set. A random joint probability distribution P_{XY}, over the product space $\mathcal{X} \times \mathcal{Y}$, is defined similarly.

Notice that the law of total probability and the definition of conditional probability continue to hold for random probability distributions. That is, $P_X = \sum_{y \in \mathcal{Y}} P_{XY}(x, y)$ and $P_{X|Y}(x, y) = \frac{P_{XY}(x,y)}{P_Y(y)}$; which means, for all $\omega \in \Omega$, $P_X(x; \omega) = \sum_{y \in \mathcal{Y}} P_{XY}(x, y; \omega)$ and $P_{X|Y}(x|y; \omega) = \frac{P_{XY}(x,y;\omega)}{P_Y(y;\omega)}$. For a countably-generated measurable space $(\mathcal{X}, \mathcal{F})$, the probability distribution $p_X \in \Delta_{\mathcal{X}}$ gives rise to a probability measure over $(\mathcal{X}, \mathcal{F})$, denoted by $\mathbb{P}_{p_X}$; for $\mathcal{A} \in \mathcal{F}$, $\mathbb{P}_{p_X}(\mathcal{A}) = \sum_{x \in \mathcal{A}} p_X(x)$. We use $\mathbb{E}_{p_X}$ to denote the expectation taken with respect to $\mathbb{P}_{p_X}$. Similarly, we use I_{p_X} to denote the mutual information taken with respect to the probability distribution p_X.

We use p_X^U to denote the distribution of a uniform random variable X. The argument of the probability distribution is omitted when it is clear from its subscript. $\mathbb{V}(p_X, q_X)$ and $\mathbb{D}(p_X || q_X)$ denote the total variation distance and the Kullback-Leibler (K-L) divergence between the two probability distributions p_X and q_X.

III. Channel Models

We describe the channel models considered in this paper. In Section III-A, we present the multiple access wiretap channel II with a noisy main channel under the aforementioned attack models for the wiretapper. Section III-B describes the generalized multiple access wiretap channel models.

A. The Multiple Access Wiretap Channel II With a Noisy Main Channel

Consider the channel model in Fig. 1. The main channel $\{\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}, p_{Y|X_1X_2}\}$ is a discrete memoryless channel consisting of two finite input alphabets $\mathcal{X}_1, \mathcal{X}_2$, a finite output alphabet $\mathcal{Y}$, and a transition probability distribution $p_{Y|X_1X_2}$. Each transmitter wishes to reliably communicate an independent message to a common receiver and to keep it secret from the wiretapper. To do so, transmitter j maps its message, W_j, uniformly distributed over $[1 : 2^{nR_j}]$, into the transmitted codeword $\mathbf{X}_j^n = [X_{j,1}, \cdots, X_{j,n}] \in \mathcal{X}_j^n$ using a stochastic encoder, $j = 1, 2$. The receiver observes the sequence $\mathbf{Y}^n = [Y_1, \cdots, Y_n] \in \mathcal{Y}^n$ and outputs the estimates $\hat{W}_1, \hat{W}_2$ of the transmitted messages. As shown in Fig. 1, we consider the following models for the wiretapper channel.

1) Model 1: This model is described in Fig. 1, when the switch is on position 1. The wiretapper chooses the subset $S_p \in \mathcal{S}_p$ and the sequence $\mathbf{u} = [u_1, \cdots, u_\mu] \in \{1, 2\}^\mu$, where $\mathcal{S}_p \triangleq \{S_p \subseteq [1 : n] : \ |S_p| = \mu \leq n\}$. That is, S_p represents the set of positions noiselessly tapped by the wiretapper and $\mathbf{u}$ represents its sequence of decisions to observe *either the first*

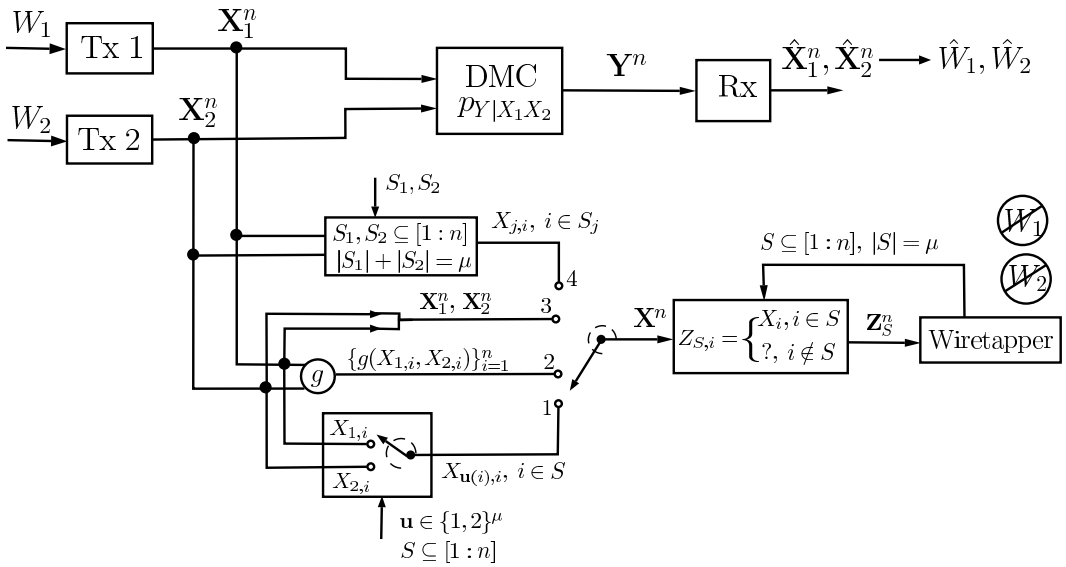

Fig. 1. The two-user multiple access wiretap channel II with a noisy main channel.

or the second user codeword symbols. We define the fraction of the tapped symbols by the wiretapper as

$$\alpha = \frac{\mu}{n}, \qquad 0 \le \alpha \le 1. \tag{1}$$

Let $S_p(k)$ and $\mathbf{u}(k)$ denote the kth elements of the subset S_p and the sequence $\mathbf{u}$, where $k \in [1:\mu]$. The set $\mathcal{S}$ of all possible strategies for the wiretapper is defined as

$$\mathcal{S} \triangleq \left\{(S_p(k), \mathbf{u}(k)) : S_p \in \mathcal{S}_p,\ \mathbf{u} \in \{1,2\}^\mu,\ k \in [1:\mu]\right\}. \tag{2}$$

For $S \in \mathcal{S}$, the wiretapper observes $\mathbf{Z}_S^n = [Z_{S,1}, \cdots, Z_{S,n}] \in \mathcal{Z}^n$, where

$$Z_{S,i} = \begin{cases} X_{j,i}, & (i,j) \in S \\ ?, & (i,j) \notin S, \end{cases} \tag{3}$$

'?' denotes an erasure, and the alphabet is $\mathcal{Z} \triangleq \{\mathcal{X}_1 \cup \mathcal{X}_2\} \cup \{?\}$.

2) Model 2: The model is described in Fig. 1, when the switch is on position 2. The wiretapper chooses the subset $S \in \mathcal{S}$, where $\mathcal{S}$ is redefined as

$$\mathcal{S} \triangleq \{S \subseteq [1:n] :\ |S| = \mu \le n\}. \tag{4}$$

The wiretapper then observes $\mathbf{Z}_S^n = [Z_{S,1}, \cdots, Z_{S,n}] \in \mathcal{Z}^n$, where

$$Z_{S,i} = \begin{cases} g(X_{1,i}, X_{2,i}), & i \in S \\ ?, & i \notin S; \end{cases} \tag{5}$$

$g : \mathcal{X}_1 \times \mathcal{X}_2 \mapsto g(\mathcal{X}_1, \mathcal{X}_2)$ is a fixed deterministic function which is not controlled by the wiretapper; $g(\mathcal{X}_1, \mathcal{X}_2)$ is the codomain of g, and $\mathcal{Z} \triangleq \{g(\mathcal{X}_1, \mathcal{X}_2)\} \cup \{?\}$. Thus, the wiretapper observes the outputs of a deterministic function of the two users codeword symbols in the positions of the subset S, and erasures otherwise. The ratio α is defined as in (1).

An example of the deterministic function g is the noiseless superposition of the two users symbols, i.e., $g(X_{1,i}, X_{2,i}) = X_{1,i} + X_{2,i}$. For this case $\mathcal{Z} \triangleq \{\mathcal{X}_1 + \mathcal{X}_2\} \cup \{?\}$, where $\mathcal{X}_1 + \mathcal{X}_2$ denotes the Minkowski sum of the sets $\mathcal{X}_1$ and $\mathcal{X}_2$, i.e., $\mathcal{X}_1 + \mathcal{X}_2 \triangleq \{x_1 + x_2 :\ x_1 \in \mathcal{X}_1,\ x_2 \in \mathcal{X}_2\}$.

3) Model 3: The model is described in Fig. 1, when the switch is on position 3. The wiretapper chooses the subset $S \in \mathcal{S}$, with $\mathcal{S}$ defined as in (4), and observes $\mathbf{Z}_S^n = [Z_{S,1}, \cdots, Z_{S,n}] \in \mathcal{Z}^n$, where

$$Z_{S,i} = \begin{cases} \{X_{1,i}, X_{2,i}\}, & i \in S \\ ?, & i \notin S, \end{cases} \tag{6}$$

and $\mathcal{Z} \triangleq \{\mathcal{X}_1 \times \mathcal{X}_2\} \cup \{?\}$. That is, the wiretapper observes *the transmitted codeword symbols of both users* in the positions of the subset S, and erasures otherwise.

4) Model 4: The model is described in Fig. 1, when the switch is on position 4. This wiretapping model represents an intermediate setting between Models 1 and 3, where the wiretapper, in every position it chooses, decides to observe either one symbol or both symbols, with an overall budget on the number of the observed symbols. In particular, the wiretapper chooses two (possibly overlapping) subsets $S_1, S_2 \subseteq [1:n]$, such that $|S_1| + |S_2| = \mu$. The set $\mathcal{S}$ of all possible strategies for the wiretapper is defined as

$$\mathcal{S} \triangleq \{(S_1, S_2) : S_1, S_2 \subseteq [1:n],\ |S_1| + |S_2| = \mu \le n\}. \tag{7}$$

For $S \in \mathcal{S}$, the wiretapper observes $\mathbf{Z}_S^n = [Z_{S,1}, \cdots, Z_{S,n}] \in \mathcal{Z}^n$, where

$$Z_{S,i} = \begin{cases} \{X_{1,i}, X_{2,i}\}, & i \in S_1 \cap S_2 \\ X_{j,i}, & i \in S_j \setminus S_k,\ k \ne j \\ ?, & \text{o.w.}, \end{cases} \tag{8}$$

and $\mathcal{Z} \triangleq \{\mathcal{X}_1 \cup \mathcal{X}_2\} \cup \{\mathcal{X}_1 \times \mathcal{X}_2\} \cup \{?\}$. The ratio α is defined as in (1).

Next, we present generalized multiple access wiretap channel models which extend the attack models proposed in this section to the case when the wiretapper sees noisy observations, instead of erasures, outside the subset(s) it chooses.

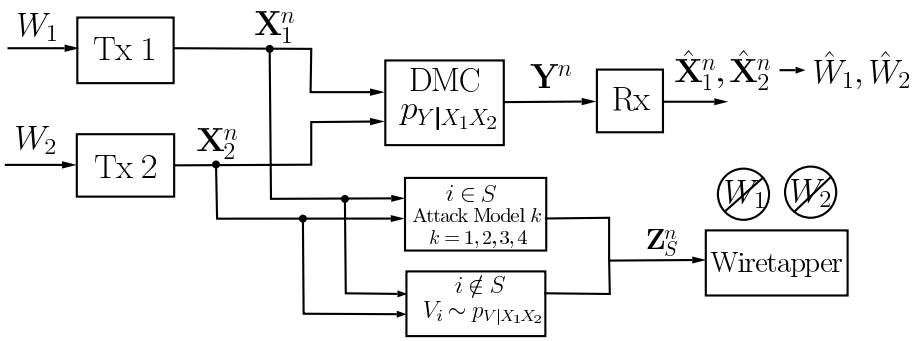

Fig. 2. The generalized two-user multiple access wiretap channel.

B. The Generalized Multiple Access Wiretap Channel

Consider the channel model in Fig. 2. The main channel in this model is identical to the main channel in Section III-A. For the wiretapper channel, we consider the same wiretapping models in Sections III-A.1-III-A.4, while replacing the erasures with the outputs of the discrete memoryless multiple access channel $p_{V|X_1X_2}$. In particular, for the four wiretapping models in Section III-A, $Z_{S,i}$ is defined as in (3), (5), (6), and (8), with replacing '?' by $V_i \in \mathcal{V}$, where $\mathbf{V}^n = [V_1, \cdots, V_n] \in \mathcal{V}^n$ is the n-letter output of $p_{V|X_1X_2}$, and $\mathcal{V}$ is a finite alphabet. The alphabets $\mathcal{Z}$ are thus defined as in Sections III-A.1-III-A.4 with replacing {?} by $\mathcal{V}$.

For the channel models described in Sections III-A and III-B, an $(n, 2^{nR_1}, 2^{nR_2})$ channel code $\mathcal{C}_n \triangleq \{\mathcal{C}_{1,n}, \mathcal{C}_{2,n}\}$ consists of two message sets $\mathcal{W}_1 = [1 : 2^{nR_1}]$, $\mathcal{W}_2 = [1 : 2^{nR_2}]$; two stochastic encoders $P_{\mathbf{X}_1^n|W_1}^{(\mathcal{C}_{1,n})}$, $P_{\mathbf{X}_2^n|W_2}^{(\mathcal{C}_{2,n})}$, and a decoder at the receiver. (R_1, R_2) is an achievable strong secrecy rate pair if there exists a sequence of $(n, 2^{nR_1}, 2^{nR_2})$ codes, $\{\mathcal{C}_n\}_{n\geq 1}$, such that

$$\lim_{n\to\infty} \mathbb{P}^{(\mathcal{C}_n)}\left(\bigcup_{j=1,2} (\hat{W}_j \neq W_j)\right) = 0, \tag{9}$$

$$\text{and} \quad \lim_{n\to\infty} \max_{S\in\mathcal{S}} I^{(\mathcal{C}_n)}(W_1, W_2; \mathbf{Z}_S^n) = 0. \tag{10}$$

$\mathbb{P}^{(\mathcal{C}_n)}$ and $I^{(\mathcal{C}_n)}$ denote the probability measure and the mutual information with respect to the joint probability distribution induced by the code $\mathcal{C}_n$. Strong secrecy capacity region for the channel is the supremum of all achievable strong secrecy rate pairs (R_1, R_2).

IV. Main Results

We first present achievable strong secrecy rate regions for the generalized multiple access wiretap channel in Section III-B, under the four proposed attack models for the wiretapper.

Theorem 1: For $0 \leq \alpha \leq 1$, an achievable strong secrecy rate region for the generalized multiple access wiretap channel under the wiretapper model 1, $\mathcal{R}^{(1)}(\alpha)$, is given by the convex hull of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(U_1; Y|U_2) - \alpha I(U_1; X_1) - (1-\alpha) I(U_1; V) \tag{11}$$

$$R_2 \leq I(U_2; Y|U_1) - \alpha I(U_2; X_2) - (1-\alpha) I(U_2; V) \tag{12}$$

$$R_1 + R_2 \leq I(U_1, U_2; Y) - \alpha I(U_1, U_2; X_1, X_2) - (1-\alpha) I(U_1, U_2; V), \tag{13}$$

for some distribution $p_{U_1X_1} p_{U_2X_2}$ which satisfies the Markov chains $U_1 - X_1 - (Y, V)$ and $U_2 - X_2 - (Y, V)$.

Theorem 2: For $0 \leq \alpha \leq 1$, an achievable strong secrecy rate region for the generalized multiple access wiretap channel under the wiretapper model 2, $\mathcal{R}^{(2)}(\alpha)$, is given by the convex hull of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(U_1; Y|U_2) - \alpha I(U_1; g(X_1, X_2)) - (1-\alpha) I(U_1; V) \tag{14}$$

$$R_2 \leq I(U_2; Y|U_1) - \alpha I(U_2; g(X_1, X_2)) - (1-\alpha) I(U_2; V) \tag{15}$$

$$R_1 + R_2 \leq I(U_1, U_2; Y) - \alpha I(U_1, U_2; g(X_1, X_2)) - (1-\alpha) I(U_1, U_2; V), \tag{16}$$

for some distribution $p_{U_1X_1} p_{U_2X_2}$ which satisfies the Markov chains $U_1 - X_1 - (Y, V)$ and $U_2 - X_2 - (Y, V)$.

Theorem 3: For $0 \leq \alpha \leq 1$, an achievable strong secrecy rate region for the generalized multiple access wiretap channel under the wiretapper model 3, $\mathcal{R}^{(3)}(\alpha)$, is given by the convex hull of all rate pairs (R_1, R_2) satisfying (11)-(13), for some distribution $p_{U_1X_1} p_{U_2X_2}$ which satisfies the Markov chains $U_1 - X_1 - (Y, V)$ and $U_2 - X_2 - (Y, V)$.

Remark 1: The achievable strong secrecy rate region for the generalized multiple access wiretap channel under wiretapper model 1 in Theorem 1 is identical to the achievable strong secrecy rate region for the more powerful wiretapper in Theorem 3, i.e., $\mathcal{R}^{(1)}(\alpha) = \mathcal{R}^{(3)}(\alpha)$. When the wiretapper has the ability of choosing to observe either symbol in every position it chooses, each user designs its transmission according to the worst case scenario in which the wiretapper decides to observe only its symbols in all the positions of the subset. This results in an achievable rate region for the wiretapper model 1 as when the wiretapper observes both users symbols in each position it chooses. However, this argument does not imply the equivalence of the secrecy capacity regions for the two models.

Theorem 4: For $0 \leq \alpha \leq 1$, an achievable strong secrecy rate region for the generalized multiple access wiretap channel under the wiretapper model 4, $\mathcal{R}^{(4)}(\alpha)$, is given by the convex hull of all rate pairs (R_1, R_2) satisfying (11)-(13), for some distribution $p_{U_1X_1} p_{U_2X_2}$ which satisfies the Markov chains $U_1 - X_1 - (Y, V)$ and $U_2 - X_2 - (Y, V)$.

Remark 2: The achievable strong secrecy rate region for the generalized multiple access wiretap channel under wiretapper model 4 in Theorem 4 is identical to the achievable regions for

wiretapper models 1 and 3. The worst case scenario according to which user j, $j = 1, 2$, designs its transmission is when the wiretapper chooses S_1, S_2 such that $|S_j| = \mu$, $|S_k| = 0, k \neq j$. That is, the wiretapper spends all of its budget on observing only user j's symbols, resulting in an achievable rate region as when the wiretapper observes both users symbols in each position it chooses.

Remark 3: The achievable strong secrecy rate region for the generalized multiple access wiretap channel under wiretapper models $1, 3, 4$, is included in the achievable region for the wiretapper model 2, i.e., $\mathcal{R}^{(1)}(\alpha) \subseteq \mathcal{R}^{(2)}(\alpha)$. This follows due to the Markov chains $U_1 - X_1 - g(X_1, X_2)$; $U_2 - X_2 - g(X_1, X_2)$, and $(U_1, U_2) - (X_1, X_2) - g(X_1, X_2)$. By data processing inequality, we have

$$I(U_j; X_j) \geq I(U_j; g(X_1, X_2)), \qquad j = 1, 2, \tag{17}$$

$$I(U_1, U_2; X_1, X_2) \geq I(U_1, U_2; g(X_1, X_2)). \tag{18}$$

Next, we present achievable strong secrecy rate regions for the multiple access wiretap channel II in Section III-A.

Corollary 1: For $0 \leq \alpha \leq 1$, an achievable strong secrecy rate region for the multiple access wiretap channel II under wiretapper models $1, 3, 4$ in Sections III-A.1, III-A.3, III-A.4 is given by the convex hull of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(U_1; Y|U_2) - \alpha I(U_1; X_1), \tag{19}$$

$$R_2 \leq I(U_2; Y|U_1) - \alpha I(U_2; X_2), \tag{20}$$

$$R_1 + R_2 \leq I(U_1, U_2; Y) - \alpha I(U_1, U_2; X_1, X_2), \tag{21}$$

for some distribution $p_{U_1X_1}p_{U_2X_2}$ which satisfies the Markov chains $U_1 - X_1 - Y$ and $U_2 - X_2 - Y$.

Corollary 2: For $0 \leq \alpha \leq 1$, an achievable strong secrecy rate region for the multiple access wiretap channel II under the wiretapper model 2 in Section III-A.2 is given by the convex hull of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(U_1; Y|U_2) - \alpha I(U_1; g(X_1, X_2)), \tag{22}$$

$$R_2 \leq I(U_2; Y|U_1) - \alpha I(U_2; g(X_1, X_2)), \tag{23}$$

$$R_1 + R_2 \leq I(U_1, U_2; Y) - \alpha I(U_1, U_2; g(X_1, X_2)), \tag{24}$$

for some distribution $p_{U_1X_1}p_{U_2X_2}$ which satisfies the Markov chains $U_1 - X_1 - Y$ and $U_2 - X_2 - Y$.

Corollaries 1 and 2 follow directly from Theorems 1-4 by setting $V = \text{const.}$, i.e., the channel $p_{V|X_1X_2}$ is an erasure channel with erasure probability one.

The proofs for Theorems 1-4 are provided in Sections V and VI. A discussion about the main results of this work is provided in Section VII.

V. Proof of Theorem 1

The achievability proof of Theorem 1 follows the same key steps as in [13], with the need of extending the utilized tools to address the multiterminal setting as will be explained shortly. The outline of the proof is as follows:

1) First, we consider the availability of an *additional* shared randomness at the transmitters, receiver, as well as the external wiretapper, in the original channel model. This assumed shared randomness represents the randomness of codebooks generation in the original channel model, whose realizations are available at the legitimate terminals and the wiretapper.
2) We then define a dual *multi-terminal* secret key agreement problem in the source model, which introduces a set of random variables similar to those introduced by the original channel model with the added shared randomness.
3) We provide rate conditions which satisfy certain reliability and secrecy (independence) conditions in the dual source model. We also solve for rate conditions which result in the induced joint distributions from the original and dual models to be almost identical in the total variation distance sense.
4) Next, we use the closeness of the induced joint distributions to show that, under the same rate conditions, the desired reliability and secrecy properties in the original channel model with the added shared randomness are satisfied.
5) Finally, we eliminate the added shared randomness from the original channel model by conditioning on a certain instance of that randomness, while keeping the reliability and security conditions satisfied. This resembles showing the existence of good codebooks which satisfy the desired properties.

The achievability proof is thus threefold: (i) Reliability of the keys in the dual source model, (ii) Security of the keys in the dual source model, and (iii) Closeness of the induced joint distributions. Reliability of the keys follows from Slepian-Wolf source coding theorem for multiple sources [33, Theorem 10.3]. Closeness of joint distributions, and converting the reliability and security conditions from the dual problem to the original problem, are ensured by utilizing an *exponential* convergence rate for the average total variation distance between the two distributions. This is done using a rather straightforward generalization of [13, Lemma 1].

The main challenge in the proof lies in ensuring security for the keys in the dual source model, which requires *doubly-exponential* convergence rate for the probability of the *two keys* being uniform and independent from the wiretapper's observation, in the Kullback-Leibler divergence sense. Generalizing the lemma derived for the single source case, [13, Lemma 2], to the case of two sources in order to directly obtain a rate region for the keys generated at the sources, which satisfies the doubly-exponential convergence, is not easy. Instead, we adapt the lemma so that we derive the *corner points* of this rate region. Time sharing between these corner points results in the desired rate region.

The *exponential* convergence rate for the average total variation distance between the distributions of the dual and original problems is *needed* to prove a *convergence in probability* result, which allows for converting the security condition from the dual problem to the original problem. In addition, note that, for all the proposed models in this paper, the number of possible subsets the wiretappcer can choose is exponential in the block length. The double-exponential convergence rate for the probability that the keys are uniform and independent from the wiretapper's observation, i.e., the security condition

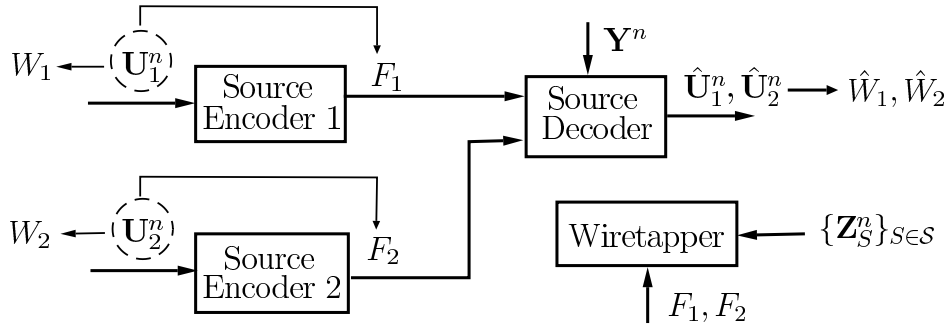

Fig. 3. Problem A: Multi-terminal secret key agreement in the source model.

in the dual source model, is utilized, along with the union bound, in order to ensure security against the exponentially many possible strategies for the wiretapper [13].

Let us first fix the distribution $p_{U_1X_1}p_{U_2X_2} = p_{U_1}p_{U_2}p_{X_1|U_1}p_{X_2|U_2}$. Let $p_{Y|U_1U_2}$ be the distribution resulting from concatenating the discrete memoryless channels $p_{Y|X_1X_2}$ and $p_{X_1X_2|U_1U_2} = p_{X_1|U_1}p_{X_2|U_2}$, where $p_{Y|X_1X_2}$ is the transition probability distribution for the legitimate channel of the channel model in Section III-B. That is,

$$p_{Y|U_1U_2}(y|u_1,u_2) = \sum_{x_1,x_2\in\mathcal{X}_1\times\mathcal{X}_2} p_{X_1|U_1}(x_1|u_1)p_{X_2|U_2}(x_2|u_2)p_{Y|X_1X_2}(y|x_1,x_2). \tag{25}$$

We present the following two problems. Each problem (i) describes a system model, (ii) defines a set of random variables for the model, and (iii) induces a joint distribution over these random variables. We precisely identify the joint distribution induced by each problem.

Problem A: Multi-terminal Secret Key Agreement in the Source Model (Dual Model):

This problem is shown in Fig. 3, and is described as follows. Let $\mathbf{U}_1^n$, $\mathbf{U}_2^n$, and $\mathbf{Y}^n$ be independent and identically distributed (i.i.d.) sequences according to the distribution $p_{U_1}p_{U_2}p_{Y|U_1U_2}$.

Source Encoders: Source encoder j, $j=1,2$, (i) observes the i.i.d. random sequence $\mathbf{U}_j^n$, distributed according to p_{U_j} over the finite alphabet $\mathcal{U}_j$, (ii) assigns two indices $w_j(\mathbf{u}_j^n) \in [1:2^{nR_j}]$ and $f_j(\mathbf{u}_j^n) \in [1:2^{n\tilde{R}_j}]$ to each sequence $\mathbf{u}_j^n \in \mathcal{U}_j^n$, and (iii) sends the index $f_j(\mathbf{u}_j^n)$ to the common decoder over a noiseless public channel, which is perfectly accessed by the wiretapper. The indices w_j, $j=1,2$, represent the confidential keys generated at the source encoders and to be reconstructed at the common decoder.

Random Binning: We consider a random binning argument in which source encoder j, $j=1,2$, randomly and independently assigns two indices $w_j(\mathbf{u}_j^n)$ and $f_j(\mathbf{u}_j^n)$ to each sequence $\mathbf{u}_j^n \in \mathcal{U}_j^n$, according to uniform distributions over $[1:2^{nR_j}]$ and $[1:2^{n\tilde{R}_j}]$, respectively. Let $W_j \triangleq \mathcal{B}_1^{(j)}(\mathbf{U}_j^n)$ and $F_j \triangleq \mathcal{B}_2^{(j)}(\mathbf{U}_j^n)$ denote the random bin indices, where $\mathcal{B}_1^{(j)}$ and $\mathcal{B}_2^{(j)}$ are independent and uniformly distributed over $[1:2^{nR_j}]$ and $[1:2^{n\tilde{R}_j}]$, respectively.

Decoder: The decoder observes the i.i.d. random sequence (side information) $\mathbf{Y}^n$, distributed according to p_Y over the finite alphabet $\mathcal{Y}$, and the public messages F_1 and F_2, and outputs the estimates $\hat{\mathbf{U}}_1^n$ and $\hat{\mathbf{U}}_2^n$ of the source encoders observations, and the estimates $\hat{W}_1$ and $\hat{W}_2$ of their generated keys. In particular, the decoder assigns estimates $(\hat{\mathbf{u}}_1^n,\hat{\mathbf{u}}_2^n) \in \mathcal{U}_1^n\times\mathcal{U}_2^n$ for each sequence $\mathbf{y}^n \in \mathcal{Y}^n$ and index pair $(f_1(\mathbf{u}_1^n)), f_2(\mathbf{u}_2^n))$, which in turn are used to output the estimates $\hat{w}_1(\hat{\mathbf{u}}_1^n)$ and $\hat{w}_2(\hat{\mathbf{u}}_2^n)$.

Wiretapper Model: The wiretapper observes the random sequence $\mathbf{Z}_S^n$, for some $S\in\mathcal{S}$, and the public messages F_1 and F_2; where $\mathcal{S}$ is defined in (2), and $\mathbf{Z}_S^n$, for all $S\in\mathcal{S}$, is defined as in (3) with replacing the erasures '?' by V_i. In particular, the wiretapper chooses the strategy $S\in\mathcal{S}$ whose realization is *unknown* to the legitimate terminals. The cardinality of the set $\mathcal{S}$ of all possible wiretapper's strategies for attack model 1 is upper bounded as

$$|\mathcal{S}| = \binom{n}{\mu}\times 2^{\mu} = \binom{n}{\alpha n}\times 2^{\alpha n} < 2^n\times 2^{\alpha n} = 2^{(1+\alpha)n}. \tag{26}$$

Thus, the distribution of the wiretapper's observation (side information) $\mathbf{Z}_S^n$ is only known to belong to the finite class $\{p_{\mathbf{Z}_S^n}\}_{S\in\mathcal{S}}$ of probability distributions over the finite alphabet $\mathcal{Z}$.

Problem A thus introduces the random variables $W_{[1:2]}, F_{[1:2]}, \mathbf{U}_{[1:2]}^n, \mathbf{Y}^n, \mathbf{Z}_S^n, \hat{\mathbf{U}}_{[1:2]}^n$, and $\hat{W}_{[1:2]}$. The induced joint distribution of Problem A is

$$\begin{aligned}
&\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}_{[1:2]}^n\mathbf{Y}^n\mathbf{Z}_S^n\hat{\mathbf{U}}_{[1:2]}^n} \\
&= p_{\mathbf{U}_{[1:2]}^n\mathbf{Y}^n\mathbf{Z}_S^n}\tilde{P}_{W_{[1:2]}F_{[1:2]}|\mathbf{U}_{[1:2]}^n}\tilde{P}_{\hat{\mathbf{U}}_{[1:2]}^n|\mathbf{Y}^nF_{[1:2]}} && (27)\\
&= p_{\mathbf{U}_{[1:2]}^n\mathbf{Y}^n\mathbf{Z}_S^n}\tilde{P}_{\hat{\mathbf{U}}_{[1:2]}^n|\mathbf{Y}^nF_{[1:2]}} \\
&\quad\times \mathbb{1}\left\{\mathcal{B}_1^{(j)}(\mathbf{U}_j^n)=W_j, \mathcal{B}_2^{(j)}(\mathbf{U}_j^n)=F_j, \forall j=1,2\right\} && (28)\\
&= \tilde{P}_{W_{[1:2]}F_{[1:2]}}\tilde{P}_{\mathbf{U}_{[1:2]}^n|W_{[1:2]}F_{[1:2]}}\, p_{\mathbf{Y}^n\mathbf{Z}_S^n|\mathbf{U}_{[1:2]}^n}\tilde{P}_{\hat{\mathbf{U}}_{[1:2]}^n|\mathbf{Y}^nF_{[1:2]}}, && (29)
\end{aligned}$$

where $\tilde{P}$ denotes the random probability distribution induced by the random binning of $\mathbf{U}_1^n$ and $\mathbf{U}_2^n$, and $p_{\mathbf{U}_{[1:2]}^n\mathbf{Y}^n\mathbf{Z}_S^n}$ denotes the marginal probability distribution of the noisy observations at the source encoders, decoder, and the wiretapper; defined as in the original channel model in Section III-B under attack model 1, and (25). Note that, in (27)-(29), we use $\tilde{P}$ to denote the overall joint distribution of Problem A, as well as the

marginal and conditional probability distributions that depend on the random binning, i.e., the random bins $W_{[1:2]}$ and $F_{[1:2]}$.

Remark 4: Using the random binning argument for the dual problem (Problem A), we show that certain reliability and security conditions, averaged over the random binning, are satisfied. We then show the existence of a binning realization which satisfies the aforementioned conditions. For this fixed binning realization, the decoder's outputs $\hat{W}_1$ and $\hat{W}_2$ are deterministic functions of its estimates of the source encoders observations, i.e., $\hat{\mathbf{U}}_1^n$ and $\hat{\mathbf{U}}_2^n$.

Remark 5: Notice that the joint distribution in (27) does not include the $\hat{W}$ random variables. We will introduce them later as deterministic functions of the $\hat{\mathbf{U}}^n$ random vectors, after fixing the binning realization, i.e., the binning realization which satisfies the desired conditions.

Problem B: Original Channel Model with Additional Shared Randomness:

This problem is described as the original channel model in Section III-B under attack model 1, with the addition of assuming shared randomness F_1 and F_2 that is available at the transmitters, the receiver, as well as the wiretapper. F_1 and F_2 are independent, uniformly distributed over $[1 : 2^{n\tilde{R}_1}]$ and $[1 : 2^{n\tilde{R}_2}]$, and independent from all other random variables.

Notice that the assumed shared randomness F_1, F_2, is available at the wiretapper, and hence not utilized as a secure shared key between the legitimate parties. This shared randomness rather represents the random generation of the codebooks in the original channel model, whose realizations are available both at the legitimate terminals and the wiretapper. That is, adding this shared randomness to the original channel model brings the random generation of the codebooks from the background to the foreground as explicit random variables, F_1 and F_2.

We utilize here the encoders and decoder in (29), i.e., the encoders and decoder from the dual source model in Problem A. That is,

$$P_{\mathbf{U}_{[1:2]}^n|W_{[1:2]}F_{[1:2]}} = \tilde{P}_{\mathbf{U}_{[1:2]}^n|W_{[1:2]}F_{[1:2]}}, \quad \text{and}$$
$$P_{\hat{\mathbf{U}}_{[1:2]}^n|\mathbf{Y}^n F_{[1:2]}} = \tilde{P}_{\hat{\mathbf{U}}_{[1:2]}^n|\mathbf{Y}^n F_{[1:2]}}, \tag{30}$$

where we use P to denote the random joint, marginal, and conditional probability distributions for Problem B. The induced joint distribution for Problem B is given by

$$\begin{aligned} &P_{W_{[1:2]}F_{[1:2]}\mathbf{U}_{[1:2]}^n\mathbf{Y}^n\mathbf{Z}_S^n\hat{\mathbf{U}}_{[1:2]}^n} \\ &= p_{W_{[1:2]}}^U p_{F_{[1:2]}}^U \tilde{P}_{\mathbf{U}_{[1:2]}^n|W_{[1:2]}F_{[1:2]}} p_{\mathbf{Y}^n\mathbf{Z}_S^n|\mathbf{U}_{[1:2]}^n} \tilde{P}_{\hat{\mathbf{U}}_{[1:2]}^n|\mathbf{Y}^n F_{[1:2]}}. \end{aligned} \tag{31}$$

Remark 6: Notice that $P_{\mathbf{U}_{[1:2]}^n|W_{[1:2]}F_{[1:2]}} = \tilde{P}_{\mathbf{U}_{[1:2]}^n|W_{[1:2]}F_{[1:2]}}$ factorizes as $\tilde{P}_{\mathbf{U}_1^n|W_1F_1}\tilde{P}_{\mathbf{U}_2^n|W_2F_2}$. That is, the shared randomness F_i available at the jth transmitter is not utilized to generate its codeword $\mathbf{U}_j^n$; $i, j = 1, 2, i \neq j$. This implies that the transmitted codeword at one transmitter does not depend on the codebook of the other transmitter.

Before continuing with the proof, we state the following lemmas.

A. Useful Lemmas

By comparing the joint distributions for Problems A and B in (29) and (31), we find that they only differ in the distribution for $W_{[1:2]}$ and $F_{[1:2]}$. In particular, $W_{[1:2]}$ and $F_{[1:2]}$ are independent and uniformly distributed in Problem B, while their distribution in Problem A is determined by the random binning of $\mathbf{U}_1^n$ and $\mathbf{U}_2^n$. The following lemma is a one-shot result which provides conditions on the binning rates such that the random binning in Problem A results in a distribution for the bins that is close, in the total variation distance sense, to independent uniform distributions. The convergence rate provided by the lemma, which is *exponential*, is needed for converting the secrecy condition, established for the source model in Problem A, to the original channel model in Problem B [13].

Lemma 1: Let X_1 and X_2 be two independent sources with probability distributions p_{X_1} and p_{X_2} over the alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. For $j = 1, 2$, each $x_j \in \mathcal{X}_j$ is randomly and independently assigned into the two indices $w_j(x_j)$ and $f_j(x_j)$ according to uniform distributions over $[1 : \tilde{W}]$ and $[1 : \tilde{F}]$. Let $W_j = \mathcal{B}_1^{(j)}(X_j)$ and $F_j = \mathcal{B}_2^{(j)}(X_j)$ denote the random bin indices, where $\mathcal{B}_1^{(j)}$ and $\mathcal{B}_2^{(j)}$ are independent and uniformly distributed over $[1 : \tilde{W}_j]$ and $[1 : \tilde{F}_j]$. Let $\mathcal{B} \triangleq \left\{\mathcal{B}_1^{(j)}(x_j), \mathcal{B}_2^{(j)}(x_j) : x_j \in \mathcal{X}_j, j = 1, 2\right\}$, and let $p_{\mathcal{B}}$ denote the joint distribution of all uniform random variables in $\mathcal{B}$. For $\gamma_j > 0$, $j = 1, 2$, define

$$\mathcal{D}_{\gamma_j} \triangleq \left\{x_j \in \mathcal{X}_j : \log \frac{1}{p_{X_j}(x_j)} > \gamma_j\right\}. \tag{32}$$

Then, we have

$$\begin{aligned} &\mathbb{E}_{p_{\mathcal{B}}}\left(\mathbb{V}\left(P_{W_{[1:2]}F_{[1:2]}}, p_{W_{[1:2]}}^U p_{F_{[1:2]}}^U\right)\right) \\ &\quad \leq \sum_{j=1}^{2}\left(\mathbb{P}_{P_{X_j}}\left(X_j \notin \mathcal{D}_{\gamma_j}\right) + \frac{1}{2}\sqrt{\tilde{W}_j\tilde{F}_j 2^{-\gamma_j}}\right), \end{aligned} \tag{33}$$

where P is the induced distribution over $W_{[1:2]}$ and $F_{[1:2]}$.

Proof: Lemma 1 is a generalization of [13, Lemma 1]. In particular, using [13, Lemma 1], we have, for $j = 1, 2$,

$$\begin{aligned} &\mathbb{E}_{p_{\mathcal{B}}}\left(\mathbb{V}\left(P_{W_jF_j}, p_{W_j}^U p_{F_j}^U\right)\right) \\ &\quad \leq \mathbb{P}_{P_{X_j}}\left(X_j \notin \mathcal{D}_{\gamma_j}\right) + \frac{1}{2}\sqrt{\tilde{W}_j\tilde{F}_j 2^{-\gamma_j}}. \end{aligned} \tag{34}$$

Since X_1 and X_2 are independent, so are $\{W_1, F_1\}$ and $\{W_2, F_2\}$. Thus, Lemma 1 follows by using the triangle inequality. ■

Lemma 2 below is again a one-shot result which provides rate conditions for a certain secrecy (independence) condition in the source model. In particular, the lemma provides a *doubly-exponential* convergence rate for the probability of the confidential keys $W_{[1:2]}$ and the public messages $F_{[1:2]}$ being independent, uniformly distributed, and all independent from the wiretapper's observation $\mathbf{Z}_S^n$. This doubly-exponential convergence is utilized, along with the union bound, to guarantee secrecy against the exponentially many choices for the wiretapper.

Lemma 2: Let X_1 and X_2 be two sources with probability distributions p_{X_1} and p_{X_2} over the finite alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$. Both X_1 and X_2 are correlated with the source $\{Z_S\}$ whose distribution is only known to belong to the finite class $\left\{p_{Z_S} : S \in \mathcal{S}\right\}$ of probability distributions over the finite alphabet $\mathcal{Z}$. For $j = 1, 2$, the source X_j is randomly binned into the two indices W_j and F_j as in Lemma 1. For $\gamma_j, \gamma_{ij} > 0, i, j = 1, 2, i \neq j$, and for any $S \in \mathcal{S}$, define

$$\mathcal{D}_j^S \triangleq \Big\{(x_{[1:2]}, z) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Z} : (x_j, z) \in \mathcal{D}_{\gamma_j}^S, \ \ (x_{[1:2]}, z) \in \mathcal{D}_{\gamma_{ij}}^S\Big\}, \tag{35}$$

where

$$\mathcal{D}_{\gamma_j}^S \triangleq \left\{(x_j, z) \in \mathcal{X}_j \times \mathcal{Z} : \log \frac{1}{p_{X_j|Z_S}(x_j|z)} > \gamma_j\right\}, \tag{36}$$

and

$$\mathcal{D}_{\gamma_{ij}}^S \triangleq \Big\{(x_{[1:2]}, z) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Z} : \log \frac{1}{p_{X_i|X_jZ_S}(x_i|x_j, z)} > \gamma_{ij}\Big\}. \tag{37}$$

If there exists a $\delta \in (0, \frac{1}{2})$ such that for $j = 1, 2$, and for all $S \in \mathcal{S}$, we have

$$\mathbb{P}_{p_{X_{[1:2]}Z_S}}\left(\left(X_{[1:2]}, Z_S\right) \in \mathcal{D}_j^S\right) \geq 1 - \delta^2, \tag{38}$$

then, we have, for every $\epsilon \in [0, 1]$, that

$$\begin{aligned}&\mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}} \mathbb{D}(P_{W_{[1:2]}F_{[1:2]}Z_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{Z_S}) \geq 2\tilde{\epsilon}\right)\\ &\leq |\mathcal{S}||\mathcal{Z}| \min_{i,j=1,2,i\neq j}\left\{\exp\left(\left(\frac{-\epsilon^2(1-\delta)2^{\gamma_j}}{3\tilde{W}_j\tilde{F}_j}\right)\right)\right.\\ &\qquad\left. + \exp\left(\left(\frac{-\epsilon^2(1-\delta)2^{\gamma_{ij}}}{3\tilde{W}_i\tilde{F}_i}\right)\right)\right\},\end{aligned} \tag{39}$$

where $p_{\mathcal{B}}$ is defined as in Lemma 1; P is the induced distribution over $W_{[1:2]}$ and $F_{[1:2]}$;

$$\tilde{\epsilon} = \max_{j=1,2}\left\{\epsilon + (\delta + \delta^2)\log(\tilde{W}_j\tilde{F}_j) + H_b(\delta^2)\right\}, \tag{40}$$

and H_b is the binary entropy function.

Proof: See the Appendix. ■

In applying Lemmas 1 and 2 to the dual source model in Problem A, we utilize the following version of Hoeffding's inequality:

Lemma 3: (Hoeffding's Inequality) [34, Theorem 2], [13, Lemma 3]:
Let $X_1, X_2, \cdots, X_n$ be independent random variables with $X_i \in [0, b]$ for all $i \in [1 : n]$, and let $\bar{m} = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(X_i)$. Then, for $\epsilon > 0$, we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i \leq (1-\epsilon)\bar{m}\right) \leq \exp\left(\frac{-2\epsilon^2\bar{m}^2}{b^2}n\right). \tag{41}$$

Remark 7: After showing that the reliability and secrecy properties established for the dual source model hold as well for the original channel model in Problem B, we utilize the selection lemma, [35, Lemma 2.2], in order to prove the existence of a binning realization such that both properties are still satisfied for the channel model with the added shared randomness. It is also utilized to eliminate the shared randomness $F_{[1:2]}$ from the channel model in Problem B.

B. Proof

1) Closeness of Induced Joint Distributions: We first apply Lemma 1 to the dual source model in Problem A to establish the closeness of the induced joint distributions from the two problems. In Lemma 1, set $X_j = \mathbf{U}_j^n$, $\tilde{W}_j = 2^{nR_j}$, and $\tilde{F}_j = 2^{n\tilde{R}_j}$, for $j = 1, 2$; $\mathbf{U}_j^n, \tilde{W}_j, \tilde{F}_j$ are defined as in Problem A. Let $\mathcal{D}_{\gamma_j}$ be defined as in (32) with $X_j = \mathbf{U}_j^n$ for $j = 1, 2$. For $\epsilon_j > 0, j = 1, 2$, choose $\gamma_j = n(1 - \epsilon_j)H(U_j)$. Without loss of generality, assume that for all $\mathbf{u}_j^n$, $j = 1, 2$, $p_{\mathbf{U}_j^n}(\mathbf{u}_j^n) > 0$. Using Hoeffding's inequality in Lemma 3,

$$\mathbb{P}_{p_{\mathbf{U}_j^n}}\left(\mathbf{U}_j^n \notin \mathcal{D}_{\gamma_j}\right) = \mathbb{P}_{p_{\mathbf{U}_j^n}}\left(\log\frac{1}{p_{\mathbf{U}_j^n}(\mathbf{U}_j^n)} \leq \gamma_j\right) \tag{42}$$

$$= \mathbb{P}_{p_{U_j}}\left(\sum_{k=1}^n \log\frac{1}{p_{U_j}(U_{j,k})} \leq n(1-\epsilon_j)H(U_j)\right) \tag{43}$$

$$\leq \exp(-\beta_j n), \tag{44}$$

where $\beta_j > 0$. By substituting the choices for $\tilde{W}_j, \tilde{F}_j, \gamma_j$, and (44) in (33), as long as

$$R_1 + \tilde{R}_1 < (1-\epsilon_1)H(U_1) \tag{45}$$

$$R_2 + \tilde{R}_2 < (1-\epsilon_2)H(U_2), \tag{46}$$

there exists a $\beta > 0$ such that

$$\mathbb{E}_{p_{\mathcal{B}}}\left(\mathbb{V}\left(\tilde{P}_{W_{[1:2]}F_{[1:2]}}, p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}\right)\right) \leq 4\exp(-\beta n), \tag{47}$$

where $p_{\mathcal{B}}$ is defined as in Lemma 1, with replacing $x_j \in \mathcal{X}_j$ by $\mathbf{u}_j^n \in \mathcal{U}_j^n$, for $j = 1, 2$.

Using (29), (31), and (47), we have

$$\begin{aligned}&\mathbb{E}_{p_{\mathcal{B}}}\Big(\mathbb{V}\Big(\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}},\\ &\qquad P_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}}\Big)\Big)\\ &= \mathbb{E}_{p_{\mathcal{B}}}\left(\mathbb{V}\left(\tilde{P}_{W_{[1:2]}F_{[1:2]}}, p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}\right)\right) \leq 4\exp(-\beta n).\end{aligned} \tag{48}$$

2) Reliability of the Dual Source Model (Problem A): Next, we establish a reliability condition for the dual source model in Problem A. We utilize a Slepian-Wolf decoder [36], which implies that [33, Theorem 10.3]

$$\lim_{n\to\infty}\mathbb{E}_{p_{\mathcal{B}}}\left(\mathbb{P}_{\tilde{P}}\left(\hat{\mathbf{U}}^n_{[1:2]} \neq \mathbf{U}^n_{[1:2]}\right)\right) = 0, \tag{49}$$

as long as

$$\tilde{R}_1 \geq H(U_1|U_2, Y), \tag{50}$$

$$\tilde{R}_2 \geq H(U_2|U_1, Y), \tag{51}$$

$$\tilde{R}_1 + \tilde{R}_2 \geq H(U_1, U_2|Y). \tag{52}$$

Using (49) and [32, Lemma 1], which is a variation of the Slepian-Wolf source coding theorem, we have, for all $S \in \mathcal{S}$,

$$\lim_{n\to\infty} \mathbb{E}_{p_{\mathcal{B}}}\Big(\mathbb{V}\Big(\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}}, \tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]} = \mathbf{U}^n_{[1:2]}\}\Big)\Big) = \lim_{n\to\infty}\mathbb{E}_{p_{\mathcal{B}}}\Big(\mathbb{P}_{\tilde{P}}(\hat{\mathbf{U}}^n_{[1:2]} \neq \mathbf{U}^n_{[1:2]})\Big) = 0. \quad (53)$$

3) Secrecy for the Dual Source Model (Problem A): Next, we use Lemma 2 to establish the secrecy condition for the dual source model in Problem A. In Lemma 2, for $j = 1, 2$, set $X_j = \mathbf{U}^n_j$, $\tilde{W}_j = 2^{nR_j}$, $\tilde{F}_j = 2^{n\tilde{R}_j}$, $Z_S = \mathbf{Z}^n_S$, for all $S \in \mathcal{S}$, where $\mathbf{U}^n_j$, $\mathcal{S}$, $\mathbf{Z}^n_S$ are defined as in Problem A. In addition, let $\mathcal{D}^S_j$, $\mathcal{D}^S_{\gamma_j}$, and $\mathcal{D}^S_{\gamma_{ij}}$ be defined as in (35)-(37), with $X_j = \mathbf{U}^n_j$ and $Z_S = \mathbf{Z}^n_S$.

For $S \in \mathcal{S}$, where $\mathcal{S}$ is defined in (2), let $\bar{S}_j \triangleq \{k : (k, j) \in S\}$. That is, $\bar{S}_j$ is the set of positions in which the wiretapper observes the jth transmitter's symbols. Notice that $\bar{S}_1 \cap \bar{S}_2 = \emptyset$. For $j = 1, 2$, let $|\bar{S}_j| = \mu_j$, and thus $\mu_1 + \mu_2 = \mu$. Recall that $S_p = \bar{S}_1 \cup \bar{S}_2$, where $|S_p| = \mu$. The wiretapper's observation can be written as $\mathbf{Z}^n_S = \{\mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}\}$.

The channel $p_{V|U_1U_2}$ is a discrete memoryless channel, since it results from concatenating the discrete memoryless channels $p_{V|X_1X_2}$ and $p_{X_1|U_1}p_{X_2|U_2}$. Thus, we have

$$H(\mathbf{U}^n_1|\mathbf{Z}^n_S) = H(\mathbf{U}^n_1|\mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) \quad (54)$$
$$= H(\mathbf{U}_{1,\bar{S}_1}, \mathbf{U}_{1,\bar{S}_2}, \mathbf{U}_{1,S^c_p}|\mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) \quad (55)$$
$$= H(\mathbf{U}_{1,\bar{S}_1}|\mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) + H(\mathbf{U}_{1,S^c_p}|\mathbf{U}_{1,\bar{S}_1}, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) + H(\mathbf{U}_{1,\bar{S}_2}|\mathbf{U}_{1,\bar{S}^c_2}, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) \quad (56)$$
$$= H(\mathbf{U}_{1,\bar{S}_1}|\mathbf{X}_{1,\bar{S}_1}) + H(\mathbf{U}_{1,S^c_p}|\mathbf{V}_{S^c_p}) + H(\mathbf{U}_{1,\bar{S}_2}) \quad (57)$$
$$= \mu_1 H(U_1|X_1) + (n-\mu)H(U_1|V) + \mu_2 H(U_1), \quad (58)$$

where (56) follows because $\bar{S}^c_2 = S^c_p \cup \bar{S}_1$. (57) follows from the Markov chains $\mathbf{U}_{1,\bar{S}_1} - \mathbf{X}_{1,\bar{S}_1} - (\mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p})$ and $\mathbf{U}_{1,S^c_p} - \mathbf{V}_{S^c_p} - (\mathbf{U}_{1,\bar{S}_1}, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2})$, and since $\mathbf{U}_{1,\bar{S}_2}$ is independent from $\{\mathbf{U}_{1,\bar{S}^c_2}, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}\}$; which hold because (i) $\{\mathbf{U}^n_1, \mathbf{X}^n_1\}$ are independent from $\{\mathbf{U}^n_2, \mathbf{X}^n_2\}$, (ii) $\mathbf{U}^n_1$ and $\mathbf{U}^n_2$ are i.i.d. sequences, and (iii) $p_{X_1|U_1}$, $p_{X_2|U_2}$, and $p_{V|U_1U_2}$ are discrete memoryless channels. Similarly, we have

$$H(\mathbf{U}^n_2|\mathbf{Z}^n_S) = \mu_2 H(U_2|X_2) + (n-\mu)H(U_2|V) + \mu_1 H(U_2) \quad (59)$$

$$H(\mathbf{U}^n_1|\mathbf{U}^n_2, \mathbf{Z}^n_S)$$
$$= H(\mathbf{U}_{1,\bar{S}_1}, \mathbf{U}_{1,\bar{S}_2}, \mathbf{U}_{1,S^c_p}|\mathbf{U}^n_2, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) \quad (60)$$
$$= H(\mathbf{U}_{1,\bar{S}_1}|\mathbf{U}^n_2, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) + H(\mathbf{U}_{1,S^c_p}|\mathbf{U}_{1,\bar{S}_1}, \mathbf{U}_{2,S_p}, \mathbf{U}_{2,S^c_p}, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) + H(\mathbf{U}_{1,\bar{S}_2}|\mathbf{U}_{1,\bar{S}^c_2}, \mathbf{U}^n_2, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p}) \quad (61)$$
$$= H(\mathbf{U}_{1,\bar{S}_1}|\mathbf{X}_{1,\bar{S}_1}) + H(\mathbf{U}_{1,S^c_p}|\mathbf{U}_{2,S^c_p}, \mathbf{V}_{S^c_p}) + H(\mathbf{U}_{1,\bar{S}_2}) \quad (62)$$
$$= \mu_1 H(U_1|X_1) + (n-\mu)H(U_1|U_2, V) + \mu_2 H(U_1), \quad (63)$$
$$H(\mathbf{U}^n_2|\mathbf{U}^n_1, \mathbf{Z}^n_S)$$
$$= \mu_2 H(U_2|X_2) + (n-\mu)H(U_2|U_1, V) + \mu_1 H(U_2), \quad (64)$$

where (62) follows due to the Markov chains $\mathbf{U}_{1,\bar{S}_1} - \mathbf{X}_{1,\bar{S}_1} - (\mathbf{U}^n_2, \mathbf{X}_{2,\bar{S}_2}, \mathbf{V}_{S^c_p})$ and $\mathbf{U}_{1,S^c_p} - (\mathbf{U}_{2,S^c_p}, \mathbf{V}_{S^c_p}) - (\mathbf{U}_{1,\bar{S}_1}, \mathbf{U}_{2,S_p}, \mathbf{X}_{1,\bar{S}_1}, \mathbf{X}_{2,\bar{S}_2})$.

In addition, for the tuples $(\mathbf{x}^n_{[1:2]}, \mathbf{z}^n)$ with $p_{\mathbf{X}^n_j|\mathbf{Z}^n_S}(\mathbf{x}^n_j|\mathbf{z}^n) > 0$ and $p_{\mathbf{X}^n_i|\mathbf{X}^n_j\mathbf{Z}^n_S}(\mathbf{x}^n_i|\mathbf{x}^n_j, \mathbf{z}^n) > 0$, where $i, j = 1, 2, i \neq j$, we have, for all $S \in \mathcal{S}$, that

$$p_{\mathbf{U}^n_j|\mathbf{Z}^n_S}(\mathbf{u}^n_j|\mathbf{z}^n)$$
$$= p(\mathbf{u}_{j,\bar{S}_j}, \mathbf{u}_{j,\bar{S}_i}, \mathbf{u}_{j,S^c_p}|\mathbf{x}_{j,\bar{S}_j}, \mathbf{x}_{i,\bar{S}_i}, \mathbf{v}_{S^c_p}) \quad (65)$$
$$= p(\mathbf{u}_{j,\bar{S}_j}|\mathbf{x}_{j,\bar{S}_j}, \mathbf{x}_{i,\bar{S}_i}, \mathbf{v}_{S^c_p})\, p(\mathbf{u}_{j,S^c_p}|\mathbf{u}_{j,\bar{S}_j}, \mathbf{x}_{j,\bar{S}_j}, \mathbf{x}_{i,\bar{S}_i}, \mathbf{v}_{S^c_p}) \times p(\mathbf{u}_{j,\bar{S}_i}|\mathbf{u}_{j,\bar{S}^c_i}, \mathbf{x}_{j,\bar{S}_j}, \mathbf{x}_{i,\bar{S}_i}, \mathbf{v}_{S^c_p}) \quad (66)$$
$$= p(\mathbf{u}_{j,\bar{S}_j}|\mathbf{x}_{j,\bar{S}_j})\, p(\mathbf{u}_{j,S^c_p}|\mathbf{v}_{S^c_p})\, p(\mathbf{u}_{j,\bar{S}_i})$$
$$= \prod_{k\in\bar{S}_j} p(u_{j,k}|x_{j,k}) \prod_{k\in S^c_p} p(u_{j,k}|v_k) \prod_{k\in\bar{S}_i} p(u_{j,k}), \quad (67)$$

$$p_{\mathbf{U}^n_i|\mathbf{U}^n_j\mathbf{Z}^n_S}(\mathbf{u}^n_i|\mathbf{u}^n_j, \mathbf{z}^n)$$
$$= p(\mathbf{u}_{i,\bar{S}_i}|\mathbf{x}_{i,\bar{S}_i})\, p(\mathbf{u}_{i,S^c_p}|\mathbf{u}_{j,S^c_p}, \mathbf{v}_{S^c_p})\, p(\mathbf{u}_{i,\bar{S}_j}) \quad (68)$$
$$= \prod_{k\in\bar{S}_i} p(u_{i,k}|x_{i,k}) \prod_{k\in S^c_p} p(u_{i,k}|v_k) \prod_{k\in\bar{S}_j} p(u_{i,k}). \quad (69)$$

For $i, j = 1, 2, i \neq j$, and $\tilde{\epsilon}_j > 0$, let

$$\gamma_j = (1-\tilde{\epsilon}_j) \min_{S\in\mathcal{S}} H(\mathbf{U}^n_j|\mathbf{Z}^n_S) \quad (70)$$
$$= (1-\tilde{\epsilon}_j)[\mu H(U_j|X_j) + (n-\mu)H(U_j|V)], \quad (71)$$
$$\gamma_{ij} = (1-\tilde{\epsilon}_j) \min_{S\in\mathcal{S}} H(\mathbf{U}^n_i|\mathbf{U}^n_j, \mathbf{Z}^n_S) \quad (72)$$
$$= (1-\tilde{\epsilon}_j)[\mu H(U_i|X_i) + (n-\mu)H(U_i|U_j, V)], \quad (73)$$

where (71) and (73) follow from (58), (59), (63), (64), and that $\mu_j H(U_j|X_j) + (n-\mu)H(U_j|V) + \mu_i H(U_j)$ is minimized by $\mu_j = \mu$ and $\mu_i = 0$, which occurs when $S = \{(k, j) : k \in \mathcal{S}_p\}$, i.e., when the wiretapper observes the symbols of the jth transmitter in all the positions it chooses. Similarly, $\mu_i H(U_i|X_i) + (n-\mu)H(U_i|U_j, V) + \mu_j H(U_i)$ is minimized by $\mu_i = \mu$ and $\mu_j = 0$.

Using Hoeffding inequality and the definition of $\mathcal{D}^S_{\gamma_j}$ in (36), we have, for all $S \in \mathcal{S}$,

$$\mathbb{P}_{p_{\mathbf{U}^n_j\mathbf{Z}^n_S}}\Big(\Big(\mathbf{U}^n_j, \mathbf{Z}^n_S\Big) \notin \mathcal{D}^S_{\gamma_j}\Big)$$
$$= \mathbb{P}_{p_{\mathbf{U}^n_j\mathbf{Z}^n_S}}\Big(\log\frac{1}{p_{\mathbf{U}^n_j|\mathbf{Z}^n_S}(\mathbf{U}^n_j|\mathbf{Z}^n_S)} \leq \gamma_j\Big) \quad (74)$$
$$= \mathbb{P}_{p_{U_jX_jV}}\Big(\sum_{k\in\bar{S}_j}\log\frac{1}{p(U_{j,k}|X_{j,k})} + \sum_{k\in S^c_p}\log\frac{1}{p(U_{j,k}|V_k)} + \sum_{k\in\bar{S}_i}\log\frac{1}{p(U_{j,k})} \leq (1-\tilde{\epsilon}_j)\big[\mu H(U_j|X_j) + (n-\mu)H(U_j|V)\big]\Big) \quad (75)$$
$$\leq \mathbb{P}_{p_{U_jX_jV}}\Big(\sum_{k\in\bar{S}_j}\log\frac{1}{p(U_{j,k}|X_{j,k})} + \sum_{k\in S^c_p}\log\frac{1}{p(U_{j,k}|V_k)} + \sum_{k\in\bar{S}_i}\log\frac{1}{p(U_{j,k})} \leq (1-\tilde{\epsilon}_j)\big[\mu_j H(U_j|X_j)$$

$$+ (n-\mu)H(U_j|V) + \mu_i H(U_j)]\Big) \tag{76}$$

$$\leq \exp(-\tilde{\beta}_j n), \tag{77}$$

where $\tilde{\beta}_j > 0$ for $j = 1, 2$; (75) follows from (67), and (76) follows because, for $i, j = 1, 2, i \neq j$, and all $S \in \mathcal{S}$,

$$\mu H(U_j|X_j) + (n-\mu)H(U_j|V) \leq \mu_j H(U_j|X_j) + (n-\mu)H(U_j|V) + \mu_i H(U_j). \tag{78}$$

Note that, for any finite γ_j, in order to compute the probability on the left hand side of (74), we only need to consider the tuples $(\mathbf{u}_j^n, \mathbf{z}^n)$ with $p_{\mathbf{U}_j^n|\mathbf{Z}_S^n}(\mathbf{u}_j^n|\mathbf{z}^n) > 0$.

Similarly, for $i, j = 1, 2, i \neq j$ and all $S \in \mathcal{S}$, using Hoeffding's inequality, (69), (73), and the definition for $\mathcal{D}_{\gamma_{ij}}^S$ in (37), we have

$$\mathbb{P}_{p_{\mathbf{U}_{[1:2]}^n \mathbf{Z}_S^n}}\left((\mathbf{U}_{[1:2]}^n, \mathbf{Z}_S^n) \notin \mathcal{D}_{\gamma_{ij}}^S\right) = \mathbb{P}_{p_{\mathbf{U}_{[1:2]}^n \mathbf{Z}_S^n}}\left(\log \frac{1}{p_{\mathbf{U}_i^n|\mathbf{U}_j^n \mathbf{Z}_S^n}(\mathbf{U}_i^n|\mathbf{U}_j^n, \mathbf{Z}_S^n)} \leq \gamma_{ij}\right) \tag{79}$$

$$\leq \exp(-\tilde{\beta}_i n). \tag{80}$$

Taking $\delta^2 = 2\exp(-\tilde{\beta}n)$, where $\tilde{\beta} = \min\{\tilde{\beta}_1, \tilde{\beta}_2\}$, yields

$$\mathbb{P}_{p_{\mathbf{U}_{[1:2]}^n \mathbf{Z}_S^n}}\left((\mathbf{U}_{[1:2]}^n, \mathbf{Z}_S^n) \notin \mathcal{D}_j^S\right) \leq \delta^2, \tag{81}$$

for $j = 1, 2$ and all $S \in \mathcal{S}$. Note that $\lim_{n\to\infty} \delta^2 = 0$. Thus, for n sufficiently large, $\delta^2 \in (0, \frac{1}{4})$. Thus, the conditions for Lemma 2 are satisfied. We also have, for $j = 1, 2$, that

$$\lim_{n\to\infty} (\delta + \delta^2)\log(\tilde{W}_j \tilde{F}_j) = \lim_{n\to\infty} n(R_j + \tilde{R}_j)(2\exp(-\tilde{\beta}n) + \exp(-\frac{1}{2}\tilde{\beta}n)) = 0 \tag{82}$$

$$\lim_{n\to\infty} H_b(\delta^2) = H_b\left(\lim_{n\to\infty} \delta^2\right) = 0, \tag{83}$$

where (83) follows because H_b is a continuous function. Thus, we have

$$\lim_{n\to\infty} \tilde{\epsilon} = \epsilon + \lim_{n\to\infty} (\delta + \delta^2)\log(\tilde{W}_j \tilde{F}_j) + \lim_{n\to\infty} H_b(\delta^2) = \epsilon. \tag{84}$$

By substituting the choices for $\tilde{W}_j, \tilde{F}_j, \gamma_j, \gamma_{ij}$, where $i, j = 1, 2,\ i \neq j$, and

$$|\mathcal{S}||\mathcal{Z}^n| \leq \exp(n[(1+\alpha)\ln 2 + \ln(|\mathcal{X}_1| + |\mathcal{X}_2| + |\mathcal{V}|)]), \tag{85}$$

in (39), and using (84), we have, for every $\epsilon, \epsilon' > 0$, $\tilde{\epsilon} = \epsilon + \epsilon'$, there exist $n^* \in \mathbb{N}$ and $\kappa_\epsilon, \tilde{\kappa} > 0$ such that for all $n \geq n^*$,

$$\mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}} \mathbb{D}\left(\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{Z}_S^n} || p_{W_{[1:2]}}^U p_{F_{[1:2]}}^U p_{\mathbf{Z}_S^n}\right) \geq 2\tilde{\epsilon}\right) \leq \exp\left(-\kappa_\epsilon e^{\tilde{\kappa}n}\right), \tag{86}$$

as long as

$$R_1 + \tilde{R}_1 \leq (1-\tilde{\epsilon}_1)\ [\alpha H(U_1|X_1) + (1-\alpha)H(U_1|V)], \tag{87}$$

$$R_2 + \tilde{R}_2 \leq (1-\tilde{\epsilon}_2)\ [\alpha H(U_2|X_2) + (1-\alpha)H(U_2|V)], \tag{88}$$

$$R_1 + R_2 + \tilde{R}_1 + \tilde{R}_2 \leq (1-\tilde{\epsilon}_1)\left[\alpha H(U_1, U_2|X_1, X_2) + (1-\alpha)H(U_1, U_2|V)\right]. \tag{89}$$

By applying the first Borel-Cantelli Lemma [37, Theorem 4.3] to (86), we get

$$\lim_{n\to\infty} \mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}} \mathbb{D}\left(\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{Z}_S^n} || p_{W_{[1:2]}}^U p_{F_{[1:2]}}^U p_{\mathbf{Z}_S^n}\right) > 0\right) = 0. \tag{90}$$

In addition, using Markov's inequality and (47), we have, for any $r > 0$, that

$$\sum_{n=1}^{\infty} \mathbb{P}_{p_{\mathcal{B}}}\left(\mathbb{V}\left(\tilde{P}_{W_{[1:2]}F_{[1:2]}}, p_{W_{[1:2]}}^U p_{F_{[1:2]}}^U\right) > r\right) \leq \frac{4}{r}\sum_{n=1}^{\infty} \exp(-\beta n) < \infty. \tag{91}$$

Using the first Borel-Cantelli lemma, it follows from (91) that

$$\lim_{n\to\infty} \mathbb{P}_{p_{\mathcal{B}}}\left(\mathbb{V}\left(\tilde{P}_{W_{[1:2]}F_{[1:2]}}, p_{W_{[1:2]}}^U p_{F_{[1:2]}}^U\right) > 0\right) = 0. \tag{92}$$

Remark 8: In the secrecy condition for the source model, (90), we require the independence of the public messages $F_{[1:2]}$ from the confidential keys $W_{[1:2]}$ and the wiretapper's observation $\mathbf{Z}_S^n$. The reason is that, after showing that the secrecy condition in (90) holds as well for the original channel model in Problem B, we need to eliminate the added shared randomness $F_{[1:2]}$ from the channel model by conditioning on a certain instance of it, without disturbing the established independence between the messages $W_{[1:2]}$ and the wiretapper's observation $\mathbf{Z}_S^n$.

Remark 9: By setting $j = 1, i = 2$, instead of the minimum in the right hand side of (39), Lemma 2 results in the maximum binning rate $R_1 + \tilde{R}_1$ of the source $\mathbf{U}_1^n$, and the corresponding maximum conditional binning rate $R_2 + \tilde{R}_2$ for the source $\mathbf{U}_2^n$ given $R_1 + \tilde{R}_1$, such that the probability in the left hand side of (39) is vanishing. In other words, Lemma 2 provides the corner points of the binning rate region such that the probability, over the random binning of the sources, that the bins are independent, uniform, and independent from the wiretapper's observation, is vanishing.

4) Converting Reliability and Secrecy Properties from Problem A to Problem B: Now, we show that the reliability and secrecy conditions in (53) and (90) hold as well for the original channel model in Problem B. First, for the reliability condition,

using (29), (31), and the triangle inequality, we have

$$\begin{aligned}
&\mathbb{V}\Big(P_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}},\\
&\qquad P_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]}=\mathbf{U}^n_{[1:2]}\}\Big)\\
&\le \mathbb{V}\Big(P_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}},\\
&\qquad \tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}}\Big)\\
&\quad+\mathbb{V}\Big(\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}},\\
&\qquad \tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]}=\mathbf{U}^n_{[1:2]}\}\Big)\\
&\quad+\mathbb{V}\Big(\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]}=\mathbf{U}^n_{[1:2]}\},\\
&\qquad P_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]}=\mathbf{U}^n_{[1:2]}\}\Big) \qquad (93)\\
&=\mathbb{V}\Big(\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}},\\
&\qquad \tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]}=\mathbf{U}^n_{[1:2]}\}\Big)\\
&\qquad+2\mathbb{V}\left(P_{W_{[1:2]}F_{[1:2]}},\, p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}\right). \qquad (94)
\end{aligned}$$

Thus, using (47), (53), and (94), we have

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}_{p_{\mathcal{B}}}\Big(\mathbb{V}\Big(P_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}},\\
&\qquad P_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]}=\mathbf{U}^n_{[1:2]}\}\Big)\Big)=0. \qquad (95)
\end{aligned}$$

Second, for the secrecy condition, using the union bound, we have

$$\begin{aligned}
&\mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}}\ \mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}\mathbf{Z}^n_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{\mathbf{Z}^n_S}\right)>0\right)\\
&\le \mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}}\ \mathbb{D}\left(\tilde{P}_{W_{[1:2]}F_{[1:2]}\mathbf{Z}^n_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{\mathbf{Z}^n_S}\right)>0\right)\\
&\qquad+\mathbb{P}_{p_{\mathcal{B}}}\left(\mathbb{V}\left(\tilde{P}_{W_{[1:2]}F_{[1:2]}},\, p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}\right)>0\right). \qquad (96)
\end{aligned}$$

Thus, using (90), (92), and (96), we have

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}}\ \mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}\mathbf{Z}^n_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{\mathbf{Z}^n_S}\right)>0\right)\\
&\qquad=0. \qquad (97)
\end{aligned}$$

By applying the selection lemma, [35, Lemma 2.2], to (95) and (97), there is at least one binning realization $\mathbf{b}^*=\{b_1^{*(j)}, b_2^{*(j)}: j=1,2\}$, with a corresponding joint distribution p^* for Problem B, such that

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{V}\Big(p^*_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}},\\
&\qquad p^*_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]}=\mathbf{U}^n_{[1:2]}\}\Big)=0, \qquad (98)
\end{aligned}$$

and

$$\lim_{n\to\infty}\mathbb{1}\left\{\max_{S\in\mathcal{S}}\ \mathbb{D}\left(p^*_{W_{[1:2]}F_{[1:2]}\mathbf{Z}^n_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{\mathbf{Z}^n_S}\right)>0\right\}=0 \qquad (99)$$

where $W_j=b_1^{*(j)}(\mathbf{U}^n_j)$ and $F_j=b_2^{*(j)}(\mathbf{U}^n_j)$, $j=1,2$.

Next, we introduce the $\hat{W}$ variables to the joint distributions in (98). For $j=1,2$, $\hat{W}_j$ is a deterministic function of the random sequence $\hat{\mathbf{U}}^n_j$. In particular, $p^*_{\hat{W}_j|\hat{\mathbf{U}}^n_j}(\hat{w}_j|\hat{\mathbf{u}}^n_j)=\mathbb{1}\left\{\hat{w}_j=b_1^{*(j)}(\hat{\mathbf{u}}^n_j)\right\}$. Using (98) and a similar analysis as in [13, (58)-(64)], we have

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}_{p^*_{F_{[1:2]}}}\left(\mathbb{P}_{p^*}\left(\hat{W}_{[1:2]}\neq W_{[1:2]}|F_{[1:2]}\right)\right)\\
&=\lim_{n\to\infty}\mathbb{V}\Big(p^*_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S\hat{\mathbf{U}}^n_{[1:2]}},\\
&\qquad p^*_{W_{[1:2]}F_{[1:2]}\mathbf{U}^n_{[1:2]}\mathbf{Y}^n\mathbf{Z}^n_S}\mathbb{1}\{\hat{\mathbf{U}}^n_{[1:2]}=\mathbf{U}^n_{[1:2]}\}\Big)=0, \qquad (100)
\end{aligned}$$

where $p^*_{F_{[1:2]}}=p^U_{F_{[1:2]}}$. Using the union bound, we also have

$$\begin{aligned}
&\mathbb{P}_{p^*_{F_{[1:2]}}}\left(\max_{S\in\mathcal{S}}\ \mathbb{D}\left(p^*_{W_{[1:2]}\mathbf{Z}^n_S|F_{[1:2]}}||p^U_{W_{[1:2]}}p^*_{\mathbf{Z}^n_S|F_{[1:2]}}\right)>0\right)\\
&\le \mathbb{P}_{p^*_{F_{[1:2]}}}\Big(\max_{S\in\mathcal{S}}\mathbb{D}(p^*_{W_{[1:2]}\mathbf{Z}^n_S|F_{[1:2]}}||p^U_{W_{[1:2]}}p^*_{\mathbf{Z}^n_S|F_{[1:2]}})>0,\\
&\qquad \text{and } \forall S,\ p^*_{W_{[1:2]}F_{[1:2]}\mathbf{Z}^n_S}=p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{\mathbf{Z}^n_S}\Big)\\
&\qquad+\mathbb{1}\left\{\max_{S\in\mathcal{S}}\mathbb{D}\left(p^*_{W_{[1:2]}F_{[1:2]}\mathbf{Z}^n_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{\mathbf{Z}^n_S}\right)>0\right\} \qquad (101)\\
&=\mathbb{1}\left\{\max_{S\in\mathcal{S}}\mathbb{D}\left(p^*_{W_{[1:2]}F_{[1:2]}\mathbf{Z}^n_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{\mathbf{Z}^n_S}\right)>0\right\}, \qquad (102)
\end{aligned}$$

where (102) follows since the first term on the right hand side of (101) is equal to zero. Thus, using (99) and (102), we have

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{P}_{p^*_{F_{[1:2]}}}\left(\max_{S\in\mathcal{S}}\ \mathbb{D}\left(p^*_{W_{[1:2]}\mathbf{Z}^n_S|F_{[1:2]}}||p^U_{W_{[1:2]}}p^*_{\mathbf{Z}^n_S|F_{[1:2]}}\right)>0\right)\\
&\qquad=0. \qquad (103)
\end{aligned}$$

5) Eliminating the Added Shared Randomness From Problem B: Once again, applying the selection lemma to (100) and (103), implies that there is at least one realization $f^*_{[1:2]}$ such that

$$\lim_{n\to\infty}\mathbb{P}_{p^*}\left(\hat{W}_{[1:2]}\neq W_{[1:2]}\big|F_{[1:2]}=f^*_{[1:2]}\right)=0, \qquad (104)$$

$$\lim_{n\to\infty}\max_{S\in\mathcal{S}}I_{p^*}\left(W_{[1:2]};\mathbf{Z}^n_S\big|F_{[1:2]}=f^*_{[1:2]}\right)=0. \qquad (105)$$

Let $\tilde{p}^*$ be the induced joint distribution for Problem A which corresponds to the binning realization $\mathbf{b}^*$. We identify $\left\{\tilde{p}^*(\mathbf{u}^n_j|w_j, f^*_j), p(\mathbf{x}^n_j|\mathbf{u}^n_j), j=1,2\right\}$ and $\left\{\tilde{p}^*(\hat{\mathbf{u}}^n_{[1:2]}|\mathbf{y}^n, f^*_{[1:2]}), \{b_1^{*(j)}(\hat{\mathbf{u}}^n_j), j=1,2\}\right\}$ as the encoders and the decoder for the original channel model.

By combining the rate conditions in (45), (46), (50)-(52), (87)-(89), and taking $\tilde{\epsilon}_1, \tilde{\epsilon}_2\to 0$, we obtain the achievable strong secrecy rate region in (11)-(13). The convex hull follows by time sharing independent codes and the fact that maximizing the secrecy constraint over S in the whole block-length is upper bounded by its maximization over the individual segments of the time sharing.

VI. Proofs for Theorems 2, 3, and 4

The proof for Theorem 2 follows similar steps as in the proof for Theorem 1. The difference is that in Problem A, $\mathcal{S}$ is defined as in (4) and $\mathbf{Z}^n_S$, for all $S\in\mathcal{S}$, is defined as

in (5) with replacing the erasures '?' by $V_i \sim p_{V|X_1X_2}$. For $i, j = 1, 2,\ i \neq j$, and all $S \in \mathcal{S}$, we have

$$H(\mathbf{U}_j^n|\mathbf{Z}_S^n) = H\left(\mathbf{U}_{j,S}, \mathbf{U}_{j,S^c}\middle|\{g(X_{1,i}, X_{2,i})\}_{i\in S}, \mathbf{V}_{S^c}\right) \quad (106)$$
$$= H\left(\mathbf{U}_{j,S}\middle|\{g(X_{1,i}, X_{2,i})\}_{i\in S}, \mathbf{V}_{S^c}\right) + H\left(\mathbf{U}_{j,S^c}\middle|\mathbf{U}_{j,S}, \{g(X_{1,i}, X_{2,i})\}_{i\in S}, \mathbf{V}_{S^c}\right) \quad (107)$$
$$= H\left(\mathbf{U}_{j,S}\middle|\{g(X_{1,i}, X_{2,i})\}_{i\in S}\right) + H(\mathbf{U}_{j,S^c}|\mathbf{V}_{S^c}) \quad (108)$$
$$= \mu H(U_j|g(X_1, X_2)) + (n-\mu)H(U_j|V) \quad (109)$$
$$H(\mathbf{U}_i^n|\mathbf{U}_j^n, \mathbf{Z}_S^n)$$
$$= H\left(\mathbf{U}_{i,S}, \mathbf{U}_{i,S^c}\middle|\mathbf{U}_j^n, \{g(X_{1,i}, X_{2,i})\}_{i\in S}, \mathbf{V}_{S^c}\right) \quad (110)$$
$$= H\left(\mathbf{U}_{i,S}\middle|\mathbf{U}_j^n, \{g(X_{1,i}, X_{2,i})\}_{i\in S}, \mathbf{V}_{S^c}\right) + H\left(\mathbf{U}_{i,S^c}\middle|\mathbf{U}_{i,S}, \mathbf{U}_j^n, \{g(X_{1,i}, X_{2,i})\}_{i\in S}, \mathbf{V}_{S^c}\right) \quad (111)$$
$$= H\left(\mathbf{U}_{i,S}\middle|\mathbf{U}_{j,S}, \{g(X_{1,i}, X_{2,i})\}_{i\in S}\right) + H(\mathbf{U}_{i,S^c}|\mathbf{U}_{j,S^c}, \mathbf{V}_{S^c}) \quad (112)$$
$$= \mu H(U_i|U_j, g(X_1, X_2)) + (n-\mu)H(U_i|U_j, V), \quad (113)$$

where (108) follows from the Markov chains $\mathbf{U}_{j,S} - \{g(X_{1,i}, X_{2,i})\}_{i\in S} - \mathbf{V}_{S^c}$ and $\mathbf{U}_{j,S^c} - \mathbf{V}_{S^c} - (\mathbf{U}_{j,S}, \{g(X_{1,i}, X_{2,i})\}_{i\in S})$. Similarly, (112) follows due to the Markov chains $\mathbf{U}_{i,S} - \left(\mathbf{U}_{j,S}, \{g(X_{1,i}, X_{2,i})\}_{i\in S}\right) - (\mathbf{U}_{j,S^c}, \mathbf{V}_{S^c})$ and $\mathbf{U}_{i,S^c} - (\mathbf{U}_{j,S^c}, \mathbf{V}_{S^c}) - (\mathbf{U}_{i,S}, \mathbf{U}_{j,S}, \{g(X_{1,i}, X_{2,i})\}_{i\in S})$.

Thus, in applying Lemma 2 to the dual source model in Problem A, for $i, j = 1, 2,\ i \neq j$, and $\tilde{\epsilon}_j > 0$, we choose

$$\gamma_j = (1-\tilde{\epsilon}_j)\min_{S\in\mathcal{S}} H(\mathbf{U}_j^n|\mathbf{Z}_S^n)$$
$$= (1-\tilde{\epsilon}_j)[\mu H(U_j|g(X_1, X_2)) + (n-\mu)H(U_j|V)] \quad (114)$$
$$\gamma_{ij} = (1-\tilde{\epsilon}_j)\min_{S\in\mathcal{S}} H(\mathbf{U}_i^n|\mathbf{U}_j^n, \mathbf{Z}_S^n)$$
$$= (1-\tilde{\epsilon}_j)[\mu H(U_i|U_j, g(X_1, X_2)) + (n-\mu)H(U_i|U_j, V)] \quad (115)$$

Using Hoeffding inequality, the conditions of the lemma are satisfied, and the rate conditions required for the secrecy property in (90) are

$$R_1 + \tilde{R}_1 \leq \alpha H(U_1|g(X_1, X_2)) + (1-\alpha)H(U_1|V) \quad (116)$$
$$R_2 + \tilde{R}_2 \leq \alpha H(U_2|g(X_1, X_2)) + (1-\alpha)H(U_2|V) \quad (117)$$
$$R_1 + R_2 + \tilde{R}_1 + \tilde{R}_2 \leq \alpha H(U_{[1:2]}|g(X_1, X_2)) + (1-\alpha)H(U_{[1:2]}|V). \quad (118)$$

These conditions, combined with the rate conditions for the Slepian-Wolf decoder in (50)-(52), and using time sharing, establish the achievability for the strong secrecy rate region in Theorem 2.

Similarly, the proof for Theorem 3 follows similar steps as in the proof for Theorem 1. In Problem A, $\mathcal{S}$ and $\mathbf{Z}_S^n$, for all $S \in \mathcal{S}$, are defined as in Section III-A.3, with replacing the erasures '?' by V_i. For $i, j = 1, 2,\ i \neq j$, and all $S \in \mathcal{S}$,

$$H(\mathbf{U}_j^n|\mathbf{Z}_S^n) = H(\mathbf{U}_{j,S}, \mathbf{U}_{j,S^c}|\mathbf{X}_{1,S}, \mathbf{X}_{2,S}, \mathbf{V}_{S^c}) \quad (119)$$
$$= H(\mathbf{U}_{j,S}|\mathbf{X}_{1,S}, \mathbf{X}_{2,S}, \mathbf{V}_{S^c}) + H(\mathbf{U}_{j,S^c}|\mathbf{U}_{j,S}, \mathbf{X}_{1,S}, \mathbf{X}_{2,S}, \mathbf{V}_{S^c}) \quad (120)$$
$$= H(\mathbf{U}_{j,S}|\mathbf{X}_{j,S}) + H(\mathbf{U}_{j,S^c}|\mathbf{V}_{S^c}) \quad (121)$$
$$= \mu H(U_j|X_j) + (n-\mu)H(U_j|V) \quad (122)$$
$$H(\mathbf{U}_i^n|\mathbf{U}_j^n, \mathbf{Z}_S^n) = H(\mathbf{U}_{i,S}, \mathbf{U}_{i,S^c}|\mathbf{U}_j^n, \mathbf{X}_{1,S}, \mathbf{X}_{2,S}, \mathbf{V}_{S^c}) \quad (123)$$
$$= H(\mathbf{U}_{i,S}|\mathbf{U}_j^n, \mathbf{X}_{1,S}, \mathbf{X}_{2,S}, \mathbf{V}_{S^c}) + H(\mathbf{U}_{i,S^c}|\mathbf{U}_{i,S}, \mathbf{U}_{j,S}, \mathbf{U}_{j,S^c}, \mathbf{X}_{1,S}, \mathbf{X}_{2,S}, \mathbf{V}_{S^c}) \quad (124)$$
$$= H(\mathbf{U}_{i,S}|\mathbf{X}_{i,S}) + H(\mathbf{U}_{i,S^c}|\mathbf{U}_{j,S^c}, \mathbf{V}_{S^c}) \quad (125)$$
$$= \mu H(U_i|X_i) + (n-\mu)H(U_i|U_j, V), \quad (126)$$

where (121) follows due to the Markov chains $\mathbf{U}_{j,S} - \mathbf{X}_{j,S} - (\mathbf{X}_{i,S}, \mathbf{V}_{S^c})$ and $(\mathbf{U}_{j,S}, \mathbf{X}_{1,S}, \mathbf{X}_{2,S}) - \mathbf{V}_{S^c} - \mathbf{U}_{j,S^c}$. Equation (125) follows from the Markov chains $\mathbf{U}_{i,S} - \mathbf{X}_{i,S} - (\mathbf{U}_j^n, \mathbf{X}_{j,S}, \mathbf{V}_{S^c})$ and $(\mathbf{U}_{i,S}, \mathbf{U}_{j,S}, \mathbf{X}_{1,S}, \mathbf{X}_{2,S}) - (\mathbf{U}_{j,S^c}, \mathbf{V}_{S^c}) - \mathbf{U}_{i,S^c}$.

Thus, for $i, j = 1, 2,\ i \neq j$, and $\tilde{\epsilon}_j > 0$, by choosing

$$\gamma_j = (1-\tilde{\epsilon}_j)\min_{S\in\mathcal{S}} H(\mathbf{U}_j^n|\mathbf{Z}_S^n)$$
$$= (1-\tilde{\epsilon}_j)[\mu H(U_j|X_j) + (n-\mu)H(U_j|V)] \quad (127)$$
$$\gamma_{ij} = (1-\tilde{\epsilon}_j)\min_{S\in\mathcal{S}} H(\mathbf{U}_i^n|\mathbf{U}_j^n, \mathbf{Z}_S^n)$$
$$= (1-\tilde{\epsilon}_j)[\mu H(U_i|X_i) + (n-\mu)H(U_i|U_j, V)], \quad (128)$$

and using Hoeffding inequality, the conditions of Lemma 2 are satisfied. The rate conditions needed for the secrecy property in (90) are given by (87)-(89). Thus, the achievable strong secrecy rate region in Theorem 3 is identical to the region in Theorem 1.

Finally, the proof for Theorem 4 follows similarly, where in Problem A, $\mathcal{S}$ and $\mathbf{Z}_S^n$, for all $S \in \mathcal{S}$, are defined as in (7) and (8), with replacing the erasures '?' by V_i.

Define $\bar{S} \triangleq S_1 \cup S_2$, and let $|S_j| = \mu_j$, $j = 1, 2$, and $|\bar{S}| = \bar{\mu}$. Notice that $\bar{\mu} \leq \mu_1 + \mu_2 = \mu$. The wiretapper's observation can be written as $\mathbf{Z}_S^n = \{\mathbf{X}_{1,S_1}, \mathbf{X}_{2,S_2}, \mathbf{V}_{\bar{S}^c}\}$. For $i, j = 1, 2,\ i \neq j$, and all $S \in \mathcal{S}$, we have

$$H(\mathbf{U}_j^n|\mathbf{Z}_S^n) = H(\mathbf{U}_{j,\bar{S}}, \mathbf{U}_{j,\bar{S}^c}|\mathbf{X}_{1,S_1}, \mathbf{X}_{2,S_2}, \mathbf{V}_{\bar{S}^c}) \quad (129)$$
$$= H(\mathbf{U}_{j,\bar{S}}|\mathbf{X}_{1,S_1}, \mathbf{X}_{2,S_2}, \mathbf{V}_{\bar{S}^c}) + H(\mathbf{U}_{j,\bar{S}^c}|\mathbf{U}_{j,\bar{S}}, \mathbf{X}_{1,S_1}, \mathbf{X}_{2,S_2}, \mathbf{V}_{\bar{S}^c}) \quad (130)$$
$$= H(\mathbf{U}_{j,\bar{S}}|\mathbf{X}_{j,S_j}) + H(\mathbf{U}_{j,\bar{S}^c}|\mathbf{V}_{\bar{S}^c}) \quad (131)$$
$$= H(\mathbf{U}_{j,S_j}|\mathbf{X}_{j,S_j}) + H(\mathbf{U}_{j,\{\bar{S}\setminus S_j\}}) + H(\mathbf{U}_{j,\bar{S}^c}|\mathbf{V}_{\bar{S}^c}) \quad (132)$$
$$= \mu_j H(U_j|X_j) + (\bar{\mu}-\mu_j)H(U_j) + (n-\bar{\mu})H(U_j|V) \quad (133)$$
$$H(\mathbf{U}_i^n|\mathbf{U}_j^n, \mathbf{Z}_S^n) = H(\mathbf{U}_{i,\bar{S}}, \mathbf{U}_{i,\bar{S}^c}|\mathbf{U}_j^n, \mathbf{X}_{1,S_1}, \mathbf{X}_{2,S_2}, \mathbf{V}_{\bar{S}^c}) \quad (134)$$
$$= H(\mathbf{U}_{i,\bar{S}}|\mathbf{U}_j^n, \mathbf{X}_{1,S_1}, \mathbf{X}_{2,S_2}, \mathbf{V}_{\bar{S}^c}) + H(\mathbf{U}_{i,\bar{S}^c}|\mathbf{U}_{i,\bar{S}}, \mathbf{U}_{j,\bar{S}}, \mathbf{U}_{j,\bar{S}^c}, \mathbf{X}_{1,S_1}, \mathbf{X}_{2,S_2}, \mathbf{V}_{\bar{S}^c}) \quad (135)$$
$$= H(\mathbf{U}_{i,\bar{S}}|\mathbf{X}_{i,S_i}) + H(\mathbf{U}_{i,\bar{S}^c}|\mathbf{U}_{j,\bar{S}^c}, \mathbf{V}_{\bar{S}^c}) \quad (136)$$
$$= H(\mathbf{U}_{i,S_i}|\mathbf{X}_{i,S_i}) + H(\mathbf{U}_{i,\{\bar{S}\setminus S_i\}}) + H(\mathbf{U}_{i,\bar{S}^c}|\mathbf{U}_{j,\bar{S}^c}, \mathbf{V}_{\bar{S}^c}) \quad (137)$$
$$= \mu_i H(U_i|X_i) + (\bar{\mu}-\mu_i)H(U_i) + (n-\bar{\mu})H(U_i|U_j, V). \quad (138)$$

The right hand side of (133) is minimized by $\mu_j = \bar{\mu} = \mu$ and $\mu_i = 0$. Similarly, the right hand side of (138) is

minimized by $\mu_i = \bar{\mu} = \mu$ and $\mu_j = 0$. By choosing

$$\gamma_j = (1-\tilde{\epsilon}_j) \min_{S\in\mathcal{S}} H(\mathbf{U}_j^n|\mathbf{Z}_S^n) = (1-\tilde{\epsilon}_j)[\mu H(U_j|X_j) + (n-\mu)H(U_j|V)] \quad (139)$$

$$\gamma_{ij} = (1-\tilde{\epsilon}_j) \min_{S\in\mathcal{S}} H(\mathbf{U}_i^n|\mathbf{U}_j^n, \mathbf{Z}_S^n) = (1-\tilde{\epsilon}_j)[\mu H(U_i|X_i) + (n-\mu)H(U_i|U_j, V)], \quad (140)$$

and using Hoeffding inequality, the conditions of Lemma 2 are satisfied. Once again, the rate conditions needed for the secrecy property in (90) are the same as (87)-(89). Thus, the achievable strong secrecy rate region in Theorem 4 is identical to the region in Theorem 1.

VII. Discussion

The third wiretapping model in Section III-A.3 is the strongest attack model with regard to the number of noiselessly observed symbols by the wiretapper. In particular, the wiretapper in this model noiselessly observes a total of 2μ symbols; μ symbols from each user. For the first and fourth wiretapping models in Sections III-A.1 and III-A.4, the wiretapper noiselessly observes a total of μ symbols. However, the ability of the wiretapper in these two models to decide on which user's symbol to observe, in each position it chooses, results in achievable strong secrecy rate regions which are identical to the rate region of the third wiretapping model.

In the achievable strong secrecy rate regions for the generalized multiple access wiretap channel under the proposed attack models, in Theorems 1-4, the terms multiplied by α represent the secrecy cost due to the noise-free observations by the wiretapper and the terms multiplied by $(1-\alpha)$ represent the secrecy cost due to its noisy observations. The achievable secrecy rate regions under wiretapper models 1, 3, 4, in Theorems 1, 3, and 4, can be alternatively expressed by the convex hull of all rate pairs (R_1, R_2) satisfying

$$R_1 \le I(U_1; Y|U_2) - I(U_1; V) - \alpha I(U_1; X_1|V), \quad (141)$$

$$R_2 \le I(U_2; Y|U_1) - I(U_2; V) - \alpha I(U_2; X_2|V), \quad (142)$$

$$R_1 + R_2 \le I(U_1, U_2; Y) - I(U_1, U_2; V) - \alpha I(U_1, U_2; X_1, X_2|V), \quad (143)$$

for some distribution $p_{U_1X_1}p_{U_2X_2}$ which satisfies the Markov chains $U_1 - X_1 - (Y, V)$ and $U_2 - X_2 - (Y, V)$. This follows because $I(U_j; X_j) = I(U_j; X_j, V), j = 1, 2$, and $I(U_1, U_2; X_1, X_2) = I(U_1, U_2; X_1, X_2, V)$, due to the Markov chains $U_j - X_j - V$, $j = 1, 2$, and $(U_1, U_2) - (X_1, X_2) - V$.

By setting the size of the subset S (or the overall size of S_1 and S_2 for wiretapper model 4) to zero, i.e., $\alpha = 0$, in Theorems 1-4, we obtain the achievable strong secrecy rate region in [30, Theorem 1] for the two user multiple access wiretap channel. The same region was derived under a weak secrecy criterion in [21], [38]. In addition, by setting $R_2 = 0$ in Theorems 1, 3, 4, we obtain the strong secrecy capacity region for the single user case in [13, Theorem 1]. From the alternative characterization of the rate region for wiretapper models 1, 3, 4, in (141)-(143), the terms multiplied by α quantify the secrecy cost, with respect to the classical multiple access wiretap channel, of the additional capabilities at the wiretapper in these models; cf. [30, Theorem 1].

For the generalized multiple access wiretap channel under wiretapper model 3, the achievable strong secrecy rate region remains the same when the wiretapper observes noisy outputs in all channel uses, i.e., the wiretapper observes the whole sequence $\mathbf{V}^n$. For the wiretapper model 3, in which the wiretapper noiselessly observes both users symbols in the positions of the subset it chooses, observing noisy symbols through the multiple access channel $p_{V|X_1X_2}$ in the same positions does not increase the wiretapper's information about the transmitted messages. The generalized multiple access wiretap channel under the third wiretapper model thus generalizes the multiple access wiretap channel in [21], [30] to the case when the wiretapper is provided with noiseless observations for a subset of its choosing of the transmitted codeword symbols of both users. The terms multiplied by α in (141)-(143) quantify the secrecy cost, with respect to the multiple access wiretap channel, of these additional noise-free observations.

We note that extending the achievability approach utilized in this paper to the case of non-uniform messages, i.e., semantic secrecy [12], does not appear straightforward. In order to handle the case of non-uniform messages, we would need to characterize the distribution of the sources $\mathbf{U}_1^n, \mathbf{U}_2^n$, given the wiretapper's observation $\mathbf{Z}_S^n$, when conditioned on each realization of the keys w_1, w_2, i.e., $p_{\mathbf{U}_{[1:2]}^n|\mathbf{Z}_S^n W_{[1:2]}}(\mathbf{u}_{[1:2]}^n|\mathbf{z}^n, w_{[1:2]})$, for all $\mathbf{u}_{[1:2]}^n \in \mathcal{U}_1^n \times \mathcal{U}_2^n$, $\mathbf{z}^n \in \mathcal{Z}^n$, and $w_{[1:2]} \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$, which is not easy due to the random binning of $\mathbf{U}_1^n$ and $\mathbf{U}_2^n$.

The advantage of the achievability approach we utilize is that it allows for rather straightforward extensions of the proof to the different attack models proposed for the wiretapper, as described in Section VI. In particular, the crucial component of the achievability proof is Lemma 2, which provides a doubly exponential convergence rate for the security measure that exhausts the exponentially many possible strategies for the type-II wiretapper. The different wiretapper models result in different rate conditions required to satisfy the assumptions of the lemma, while the remainder of the proof remains the same.

Finally, we note that, for the wiretapper models considered in this paper, the wiretapper's strategy, i.e., which positions and which user symbols to noiselessly tap, is chosen by the wiretapper before the transmission begins. Investigating the case when the wiretapper is allowed to adaptively update its strategy according to the symbols it has observed so far is of future interest.

VIII. Conclusion

In this paper, we have studied the extension of the wiretap channel II with a noisy main channel in [11] and the generalized wiretap channel model in [13] to the multiple access setting. For the multiple access wiretap channel II with a noisy main channel, we have proposed four different attack models for the wiretapper which feature different adversarial

capabilities, and derived an achievable strong secrecy rate region for each. Further, we have generalized the proposed models to the case when the wiretapper observes the outputs of a noisy multiple access channel instead of erasures outside the subset of noise-free observations it chooses, proposing a *generalized* multiple access wiretap model. We have derived achievable strong secrecy rate regions for this generalized model under the proposed wiretapping scenarios. The tools we have utilized for achievability extend the set of tools utilized for the single-user scenario in [13] to a multi-user setting.

Future work includes other multi-terminal setups with more powerful wiretappers. Additionally, extending the results and the utilized techniques in this work to Gaussian channels is an interesting open problem. Another interesting direction is to investigate whether the strong secrecy capacity in the wiretap channel II setting also implies semantic security, i.e., security over all possible message distributions.

Appendix

First, we rewrite the relative entropy in (39) as follows:

$$\begin{aligned}
&\mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}Z_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{Z_S}\right)\\
&= \sum_{w_{[1:2]},f_{[1:2]},z} P_{W_{[1:2]}F_{[1:2]}Z_S}(w_{[1:2]},f_{[1:2]},z)\\
&\qquad\times \log\frac{P_{W_{[1:2]}F_{[1:2]}Z_S}(w_{[1:2]},f_{[1:2]},z)}{p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{Z_S}(z)} \qquad (144)\\
&= \sum_{w_{[1:2]},f_{[1:2]},z} P_{W_{[1:2]}F_{[1:2]}Z_S}(w_{[1:2]},f_{[1:2]},z)\\
&\qquad\times \log\left(\frac{P_{W_{[1:2]}F_{[1:2]}Z_S}(w_{[1:2]},f_{[1:2]},z)}{P_{W_1F_1Z_S}(w_1,f_1,z)p^U_{W_2}p^U_{F_2}}\right.\\
&\qquad\left.\times\frac{P_{W_1F_1Z_S}(w_1,f_1,z)}{p^U_{W_1}p^U_{F_1}p_{Z_S}(z)}\right) \qquad (145)\\
&= \mathbb{E}_{p_{Z_S}}\left(\mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}|Z_S}||P_{W_1F_1|Z_S}p^U_{W_2}p^U_{F_2}\right)\right)\\
&\qquad+\mathbb{D}\left(P_{W_1F_1Z_S}||p^U_{W_1}p^U_{F_1}p_{Z_S}\right). \qquad (146)
\end{aligned}$$

Thus, the probability in (39) is upper bounded as

$$\begin{aligned}
&\mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}}\mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}Z_S}||p^U_{W_{[1:2]}}p^U_{F_{[1:2]}}p_{Z_S}\right)\geq 2\tilde{\epsilon}\right)\\
&\leq \mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}}\mathbb{D}\left(P_{W_1F_1Z_S}||p^U_{W_1}p^U_{F_1}p_{Z_S}\right)\geq\tilde{\epsilon}\right)\\
&+\mathbb{P}_{p_{\mathcal{B}}}\left(\begin{array}{c}\max_{S\in\mathcal{S}}\mathbb{E}_{p_{Z_S}}\\ \mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}|Z_S}||P_{W_1F_1|Z_S}p^U_{W_2}p^U_{F_2}\right)\geq\tilde{\epsilon}\end{array}\right). \qquad (147)
\end{aligned}$$

We upper bound each term on the right hand side of (147). Using [13, Lemma 2], the first term is upper bounded as

$$\begin{aligned}
&\mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}}\mathbb{D}\left(P_{W_1F_1Z_S}||p^U_{W_1}p^U_{F_1}p_{Z_S}\right)\geq\tilde{\epsilon}\right)\\
&\qquad\leq|\mathcal{S}||\mathcal{Z}|\exp\left(\frac{-\epsilon^2(1-\delta)2^{\gamma_1}}{3\tilde{W}_1\tilde{F}_1}\right). \qquad (148)
\end{aligned}$$

Next, we upper bound the second term in (147). For all $S\in\mathcal{S}$, let us define

$$\mathcal{A}_S\triangleq\left\{z\in\mathcal{Z}:\mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]},Z_S)\in\mathcal{D}^S_1\right)\geq 1-\delta\right\}, \qquad (149)$$

where $\mathcal{D}^S_1$ is defined in (35). Using Markov's inequality and (38), we have

$$\begin{aligned}
\mathbb{P}_{p_{Z_S}}(\mathcal{A}^c_S)&=\mathbb{P}_{p_{Z_S}}\left(\mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]},Z_S)\notin\mathcal{D}^S_1\right)\geq\delta\right) \qquad (150)\\
&\leq\frac{1}{\delta}\,\mathbb{P}_{p_{X_{[1:2]}Z_S}}\left((X_{[1:2]},Z_S)\notin\mathcal{D}^S_1\right)\leq\frac{\delta^2}{\delta}=\delta. \qquad (151)
\end{aligned}$$

For all $w_{[1:2]}, f_{[1:2]}\in[1:\tilde{W}]\times[1:\tilde{F}]$, $z\in\mathcal{Z}$, and $S\in\mathcal{S}$, define

$$\begin{aligned}
&P^S_1(w_{[1:2]},f_{[1:2]}|z)\\
&=\sum_{x_{[1:2]}\in\mathcal{X}_1\times\mathcal{X}_2}p_{X_{[1:2]}|Z_S}(x_{[1:2]}|z)\mathbb{1}\left\{(x_{[1:2]},z)\in\mathcal{D}^S_1\right\}\\
&\qquad\times\mathbb{1}\left\{\mathcal{B}^{(j)}_1(x_j)=w_j,\mathcal{B}^{(j)}_2(x_j)=f_j,\forall j=1,2\right\} \qquad (152)\\
&P^S_2(w_{[1:2]},f_{[1:2]}|z)\\
&=\sum_{x_{[1:2]}\in\mathcal{X}_1\times\mathcal{X}_2}P_{X_{[1:2]}|Z_S}(x_{[1:2]}|z)\mathbb{1}\left\{(x_{[1:2]},z)\notin\mathcal{D}^S_1\right\}\\
&\qquad\times\mathbb{1}\left\{\mathcal{B}^{(j)}_1(x_j)=w_j,\mathcal{B}^{(j)}_2(x_j)=f_j,\forall j=1,2\right\}. \qquad (153)
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&P_{W_{[1:2]}F_{[1:2]}|Z_S}(w_{[1:2]},f_{[1:2]}|z)\\
&\qquad=P^S_1(w_{[1:2]},f_{[1:2]}|z)+P^S_2(w_{[1:2]},f_{[1:2]}|z). \qquad (154)
\end{aligned}$$

Now, for every $x_2\in\mathcal{X}_2$, define

$$\begin{aligned}
U_{x_2}&=\sum_{x_1\in\mathcal{X}_1}p_{X_{[1:2]}|Z_S}(x_{[1:2]}|z)\\
&\times\mathbb{1}\left\{\mathcal{B}^{(2)}_1(x_2)=w_2,\mathcal{B}^{(2)}_2(x_2)=f_2\right\}\mathbb{1}\left\{(x_{[1:2]},z)\in\mathcal{D}^S_1\right\}. \qquad (155)
\end{aligned}$$

The random variables $\left\{U_{x_2}\right\}_{x_2\in\mathcal{X}_2}$ are non-negative and independent since the random variables $\left\{\mathcal{B}^{(2)}_1(x_2),\mathcal{B}^{(2)}_2(x_2)\right\}_{x_2\in\mathcal{X}_2}$ are independent. From the definition of $\mathcal{D}^S_1$ in (35), we have for $(x_{[1:2]},z)\in\mathcal{D}^S_1$ that $(x_{[1:2]},z)\in\mathcal{D}^S_{\gamma_{21}}$. Additionally, from the definition of $\mathcal{D}^S_{\gamma_{21}}$ in (37), we have that $p(x_2|x_1,z)\leq 2^{-\gamma_{21}}$. From (155), we have

$$\begin{aligned}
U_{x_2}&\leq\sum_{x_1}p_{X_1|Z_S}(x_1|z)p_{X_2|X_1,Z_S}(x_2|x_1,z)\\
&\qquad\times\mathbb{1}\left\{(x_{[1:2]},z)\in\mathcal{D}^S_1\right\} \qquad (156)\\
&\leq 2^{-\gamma_{21}}\sum_{x_1}p_{X_1|Z_S}(x_1|z)\mathbb{1}\left\{(x_{[1:2]},z)\in\mathcal{D}^S_1\right\} \qquad (157)\\
&\leq 2^{-\gamma_{21}}. \qquad (158)
\end{aligned}$$

Since for all $x_2\in\mathcal{X}_2$,

$$\mathbb{E}_{p_{\mathcal{B}}}\left(\mathbb{1}\left\{\mathcal{B}^{(2)}_1(x_2)=w_2,\mathcal{B}^{(2)}_2(x_2)=f_2\right\}\right)=\frac{1}{\tilde{W}_2\tilde{F}_2}, \qquad (159)$$

we have,

$$\sum_{x_2\in\mathcal{X}_2}\mathbb{E}_{p_{\mathcal{B}}}(U_{x_2}) = \frac{1}{\tilde{W}_2\tilde{F}_2}\sum_{x_{[1:2]}\in\mathcal{X}_1\times\mathcal{X}_2} p_{X_{[1:2]}|Z_S}(x_{[1:2]}|z)\mathbb{1}\left\{(x_{[1:2]},z)\in\mathcal{D}_1^S\right\} \tag{160}$$

$$= \frac{\mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]},z)\in\mathcal{D}_1^S\right)}{\tilde{W}_2\tilde{F}_2}. \tag{161}$$

In addition, notice that

$$\sum_{w_1,f_1} P_1^S(w_{[1:2]},f_{[1:2]}|z) = \sum_{x_{[1:2]}} p_{X_{[1:2]}|Z_S}(x_{[1:2]}|z)\mathbb{1}\left\{(x_{[1:2]},z)\in\mathcal{D}_1^S\right\} \times \sum_{w_1,f_1}\mathbb{1}\left\{\mathcal{B}_1^{(j)}(x_j)=w_j,\mathcal{B}_2^{(j)}(x_j)=f_j,\forall j=1,2\right\} \tag{162}$$

$$= \sum_{x_2}\sum_{x_1} p_{X_{[1:2]}|Z_S}(x_{[1:2]}|z)\mathbb{1}\left\{(x_{[1:2]},z)\in\mathcal{D}_1^S\right\} \times \mathbb{1}\left\{\mathcal{B}_1^{(2)}(x_2)=w_2,\mathcal{B}_2^{(2)}(x_2)=f_2\right\} \tag{163}$$

$$= \sum_{x_2} U_{x_2}. \tag{164}$$

We now state the following lemma, which is a variation on Chernoff's bound that we need to utilize in the proof.

Lemma 4: (A variation on Chernoff bound [13, Lemma 6]): Let $U_1, U_2, \cdots, U_n$ be a sequence of non-negative independent random variables with respective means $\mathbb{E}(U_i) = \bar{m}_i$. If $U_i \in [0, b]$, for all $i \in [1 : n]$, and $\sum_{i=1}^n \bar{m}_i \le \bar{m}$, then, for every $\epsilon \in [0, 1]$, we have

$$\mathbb{P}\left(\sum_{i=1}^n U_i \ge (1+\epsilon)\bar{m}\right) \le \exp\left(-\epsilon^2\frac{\bar{m}}{3b}\right). \tag{165}$$

The random variables $\{U_{x_2}\}_{x_2\in\mathcal{X}_2}$ are non-negative, independent, and $U_{x_2} \in [0, 2^{-\gamma_{21}}]$ for all $x_2 \in \mathcal{X}_2$. By applying Lemma 4 to the random variables $\{U_{x_2}\}_{x_2\in\mathcal{X}_2}$, we have,

$$\mathbb{P}_{p_{\mathcal{B}}}\left(P_1^S(w_{[1:2]},f_{[1:2]}|z) \ge \frac{1+\epsilon}{\tilde{W}_2\tilde{F}_2}P_{W_1F_1|Z_S}(w_1,f_1|z)\right) \le \mathbb{P}_{p_{\mathcal{B}}}\left(\begin{array}{l}\sum_{w_1,f_1} P_1^S(w_{[1:2]},f_{[1:2]}|z)\\ \ge \frac{1+\epsilon}{\tilde{W}_2\tilde{F}_2}\sum_{w_1,f_1} P_{W_1F_1|Z_S}(w_1,f_1|z)\end{array}\right) \tag{166}$$

$$= \mathbb{P}_{p_{\mathcal{B}}}\left(\sum_{x_2} U_{x_2} \ge \frac{1+\epsilon}{\tilde{W}_2\tilde{F}_2}\right) \tag{167}$$

$$\le \mathbb{P}_{p_{\mathcal{B}}}\left(\sum_{x_2} U_{x_2} \ge \frac{1+\epsilon}{\tilde{W}_2\tilde{F}_2}\mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]},z)\in\mathcal{D}_1^S\right)\right) \tag{168}$$

$$= \mathbb{P}_{p_{\mathcal{B}}}\left(\sum_{x_2} U_{x_2} \ge (1+\epsilon)\sum_{x_2}\mathbb{E}_{p_{\mathcal{B}}}(U_{x_2})\right) \tag{169}$$

$$\le \exp\left(\frac{-\epsilon^2 2^{\gamma_{21}}}{3\tilde{W}_2\tilde{F}_2}\mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]},z)\in\mathcal{D}_1^S\right)\right), \tag{170}$$

where (167) follows from (164), (169) follows from (161), and (170) follows from Lemma 4.

From the definition of $\mathcal{A}_S$ in (149), we have, for all $z \in \mathcal{A}_S$, that $\mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]},z)\in\mathcal{D}_1^S\right) \ge 1-\delta$. Thus, for all $z \in \mathcal{A}_S$,

$$\mathbb{P}_{p_{\mathcal{B}}}\left(P_1^S(w_{[1:2]},f_{[1:2]}|z) \ge \frac{1+\epsilon}{\tilde{W}_2\tilde{F}_2}P_{W_1F_1|Z_S}(w_1,f_1|z)\right) \le \exp\left(\frac{-\epsilon^2(1-\delta)2^{\gamma_{21}}}{3\tilde{W}_2\tilde{F}_2}\right). \tag{171}$$

Note that, for fixed $z \in \mathcal{Z}$ and $S \in \mathcal{S}$, the random variables $\left\{P_1^S(w_{[1:2]},f_{[1:2]}|z)\right\}$ are identically distributed for all $w_{[1:2]}, f_{[1:2]}$ due to the symmetry in the random binning. Let $\mathbf{b} \triangleq \{b_1^{(j)}, b_2^{(j)}, j = 1, 2\}$ be a realization of the random binning $\mathcal{B}$. We define the class $\mathcal{G}$ of binning functions $\mathbf{b}$ as

$$\mathcal{G} \triangleq \left\{\mathbf{b} : P_1^S(w_{[1:2]},f_{[1:2]}|z) < \frac{1+\epsilon}{\tilde{W}_2\tilde{F}_2}P_{W_1F_1|Z_S}(w_1,f_1|z), \text{ for all } S\in\mathcal{S} \text{ and } z\in\mathcal{A}_S\right\}. \tag{172}$$

Using the union bound, we have

$$\mathbb{P}_{p_{\mathcal{B}}}(\mathcal{G}^c) = \mathbb{P}_{p_{\mathcal{B}}}\left(P_1^S(w_{[1:2]},f_{[1:2]}|z) \ge \frac{1+\epsilon}{\tilde{W}_2\tilde{F}_2}P_{W_1F_1|Z_S}(w_1,f_1|z), \text{ for some } S\in\mathcal{S} \text{ or } z\in\mathcal{A}_S\right) \tag{173}$$

$$\le \sum_{S\in\mathcal{S},z\in\mathcal{A}_S}\mathbb{P}_{p_{\mathcal{B}}}\left(P_1^S(w_{[1:2]},f_{[1:2]}|z) \ge \frac{1+\epsilon}{\tilde{W}_2\tilde{F}_2}P_{W_1F_1|Z_S}(w_1,f_1|z)\right) \tag{174}$$

$$\le |\mathcal{S}||\mathcal{Z}|\exp\left(\frac{-\epsilon^2(1-\delta)2^{\gamma_{21}}}{3\tilde{W}_2\tilde{F}_2}\right), \tag{175}$$

where (175) follows from (171).

Take $\mathbf{b}$ such that $\mathbf{b} \in \mathcal{G}$, and set $W_j = b_1^{(j)}(X_j)$ and $F_j = b_2^{(j)}(X_j)$ for $j = 1, 2$. For all $S \in \mathcal{S}$,

$$\mathbb{E}_{p_{Z_S}}\left(\mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}|Z_S}||P_{W_1F_1|Z_S}p_{W_2}^U p_{F_2}^U\right)\right) = \mathbb{E}_{p_{Z_S}}\left(\sum_{w_{[1:2]},f_{[1:2]}} P_{W_{[1:2]}F_{[1:2]}|Z_S}(w_{[1:2]},f_{[1:2]}|Z_S) \times \log\frac{P_{W_{[1:2]}F_{[1:2]}|Z_S}(w_{[1:2]},f_{[1:2]}|Z_S)}{P_{W_1F_1|Z_S}(w_1,f_1|Z_S)p_{W_2}^U p_{F_2}^U}\right) \tag{176}$$

$$= \mathbb{E}_{p_{Z_S}}\left(\sum_{w_{[1:2]},f_{[1:2]}}\sum_{i=1}^2 P_i^S(w_{[1:2]},f_{[1:2]}|Z_S) \times \log\frac{\sum_{i=1}^2 P_i^S(w_{[1:2]},f_{[1:2]}|Z_S)}{\frac{P_{W_1F_1|Z_S}(w_1,f_1|Z_S)}{\tilde{W}_2\tilde{F}_2}\sum_{i=1}^2\sum_{\substack{w_{[1:2]},\\ f_{[1:2]}}} P_i^S(w_{[1:2]},f_{[1:2]}|Z_S)}\right) \tag{177}$$

$$\le \mathbb{E}_{p_{Z_S}}\left(\sum_{i=1}^2\sum_{w_{[1:2]},f_{[1:2]}} P_i^S(w_{[1:2]},f_{[1:2]}|Z_S)\right.$$

$$\times \log \frac{1}{\sum_{w_{[1:2]}, f_{[1:2]}} P_i^S(w_{[1:2]}, f_{[1:2]}|Z_S)}\Bigg)$$
$$+ \mathbb{E}_{p_{Z_S}}\Bigg(\sum_{w_{[1:2]}, f_{[1:2]}} P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S) \times \log \frac{\tilde{W}_2\tilde{F}_2 P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S)}{P_{W_1F_1|Z_S}(w_1, f_1|Z_S)}\Bigg)$$
$$+ \mathbb{E}_{p_{Z_S}}\Bigg(\sum_{w_{[1:2]}, f_{[1:2]}} P_2^S(w_{[1:2]}, f_{[1:2]}|Z_S) \times \log \frac{\tilde{W}_2\tilde{F}_2 P_2^S(w_{[1:2]}, f_{[1:2]}|Z_S)}{P_{W_1F_1|Z_S}(w_1, f_1|Z_S)}\Bigg), \tag{178}$$

where (177) follows since

$$\sum_{w_{[1:2]}, f_{[1:2]}} \sum_{i=1}^{2} P_i^S\left(w_{[1:2]}, f_{[1:2]}|Z_S\right) = 1, \tag{179}$$

and (178) follows from the log-sum inequality.

Now, we upper bound each term in the right hand side of (178) for $\mathbf{b} \in \mathcal{G}$. The second term in the right hand side of (178) is upper bounded as follows:

$$\mathbb{E}_{p_{Z_S}}\Bigg(\sum_{w_{[1:2]}, f_{[1:2]}} P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S) \times \log \frac{\tilde{W}_2\tilde{F}_2\ P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S)}{P_{W_1F_1|Z_S}(w_1, f_1|Z_S)}\Bigg)$$
$$= \mathbb{E}_{p_{Z_S}}\Bigg(\sum_{w_{[1:2]}, f_{[1:2]}} P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S) \times \log \frac{\tilde{W}_2\tilde{F}_2\ P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S)}{P_{W_1F_1|Z_S}(w_1, f_1|Z_S)} \mathbb{1}\{Z_S \notin \mathcal{A}_S\}\Bigg)$$
$$+ \mathbb{E}_{p_{Z_S}}\Bigg(\sum_{w_{[1:2]}, f_{[1:2]}} P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S) \times \log \frac{\tilde{W}_2\tilde{F}_2\ P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S)}{P_{W_1F_1|Z_S}(w_1, f_1|Z_S)} \mathbb{1}\{Z_S \in \mathcal{A}_S\}\Bigg) \tag{180}$$
$$\leq \log(\tilde{W}_2\tilde{F}_2) \sum_{z\in\mathcal{Z}} p_{Z_S}(z)\mathbb{1}\{z \notin \mathcal{A}_S\} \sum_{\substack{w_{[1:2]},\\ f_{[1:2]}}} P_1^S(w_{[1:2]}, f_{[1:2]}|z)$$
$$+ \log(1+\epsilon)\, \mathbb{E}_{p_{Z_S}}\left(\sum_{w_{[1:2]}, f_{[1:2]}} P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S)\right) \tag{181}$$
$$\leq \mathbb{P}_{p_{Z_S}}(Z_S \notin \mathcal{A}_S)\log(\tilde{W}_2\tilde{F}_2) + \log(1+\epsilon) \tag{182}$$
$$\leq \delta \log(\tilde{W}_2\tilde{F}_2) + \epsilon, \tag{183}$$

where (181) follows because (i) for $i = 1, 2$, and all $w_{[1:2]}, f_{[1:2]}$, we have

$$P_i^S(w_{[1:2]}, f_{[1:2]}|Z_S) \leq P_{W_{[1:2]}F_{[1:2]}|Z_S}(w_{[1:2]}, f_{[1:2]}|Z_S) \tag{184}$$
$$= P_{W_1F_1|Z_S}(w_1, f_1|Z_S)P_{W_2F_2|W_1F_1Z_S}(w_2, f_2|w_1, f_1, Z_S) \tag{185}$$
$$\leq P_{W_1F_1|Z_S}(w_1, f_1|Z_S), \tag{186}$$

and (ii) for $\mathbf{b} \in \mathcal{G}$ and $Z_S \in \mathcal{A}_S$, $\frac{\tilde{W}_2\tilde{F}_2\ P_1^S(w_{[1:2]}, f_{[1:2]}|Z_S)}{P_{W_1F_1|Z_S}(w_1, f_1|Z_S)} < (1+\epsilon)$, which follows from (172).

Next, we upper bound the third term in the right hand side of (178). We have that

$$\mathbb{E}_{p_{Z_S}}\left(\sum_{w_{[1:2]}, f_{[1:2]}} P_2^S(w_{[1:2]}, f_{[1:2]}|Z_S)\right)$$
$$= \sum_z p_{Z_S}(z) \sum_{x_{[1:2]}} p_{X_{[1:2]}|Z_S}(x_{[1:2]}, z)\mathbb{1}\left\{(x_{[1:2]}, z) \notin \mathcal{D}_1^S\right\}$$
$$\times \sum_{w_{[1:2]}, f_{[1:2]}} \mathbb{1}\left\{\mathcal{B}_1^{(j)}(x_j) = w_j, \mathcal{B}_2^{(j)}(x_j) = f_j, j = 1, 2\right\} \tag{187}$$
$$= \mathbb{P}_{p_{X_{[1:2]},Z_S}}\left((X_{[1:2]}, Z_S) \notin \mathcal{D}_1^S\right) \leq \delta^2, \tag{188}$$

where the inequality in (188) follows from the assumption of the lemma in (38). Using (186) and (188), we have

$$\mathbb{E}_{p_{Z_S}}\Bigg(\sum_{w_{[1:2]}, f_{[1:2]}} P_2^S(w_{[1:2]}, f_{[1:2]}|Z_S) \times \log \frac{\tilde{W}_2\tilde{F}_2 P_2^S(w_{[1:2]}, f_{[1:2]}|Z_S)}{P_{W_1F_1|Z_S}(w_1, f_1|Z_S)}\Bigg) \leq \delta^2 \log(\tilde{W}_2\tilde{F}_2). \tag{189}$$

Since we have

$$\sum_{w_{[1:2]}, f_{[1:2]}} P_1^S\left(w_{[1:2]}, f_{[1:2]}|Z_S\right) = \mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]}, Z_S) \in \mathcal{D}_1^S\right), \tag{190}$$

and

$$\sum_{w_{[1:2]}, f_{[1:2]}} P_2^S\left(w_{[1:2]}, f_{[1:2]}|Z_S\right) = 1 - \mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]}, Z_S) \in \mathcal{D}_1^S\right), \tag{191}$$

the first term on the right hand side of (178) is upper bounded as follows:

$$\mathbb{E}_{p_{Z_S}}\Bigg(\sum_{i=1}^{2} \sum_{w_{[1:2]}, f_{[1:2]}} P_i^S(w_{[1:2]}, f_{[1:2]}|Z_S) \times \log \frac{1}{\sum_{w_{[1:2]}, f_{[1:2]}} P_i^S(w_{[1:2]}, f_{[1:2]}|Z_S)}\Bigg)$$
$$= \mathbb{E}_{p_{Z_S}}\left(H_b\left(\mathbb{P}_{p_{X_{[1:2]}|Z_S}}\left((X_{[1:2]}, Z_S) \in \mathcal{D}_1^S\right)\right)\right) \tag{192}$$
$$\leq H_b\left(\mathbb{P}_{p_{X_{[1:2]},Z_S}}\left((X_{[1:2]}, Z_S) \in \mathcal{D}_1^S\right)\right) \tag{193}$$
$$\leq H_b(1-\delta^2) = H_b(\delta^2), \tag{194}$$

where (193) follows from Jensen's inequality and the concavity of H_b, and (194) follows from (38) and that $H_b(x)$ is monotonically decreasing in $x \in (\frac{1}{2}, 1)$.

Using (183), (189), and (194), for any $\mathbf{b} \in \mathcal{G}$ and $S \in \mathcal{S}$, the left hand side of (178) is upper bounded as

$$\mathbb{E}_{p_{Z_S}}\left(\mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}|Z_S}||P_{W_1F_1|Z_S}p_{W_2}^U p_{F_2}^U\right)\right)$$
$$\leq \epsilon + (\delta + \delta^2)\log(\tilde{W}_2\tilde{F}_2) + H_b(\delta^2) \leq \tilde{\epsilon}. \tag{195}$$

Thus, the second probability on the right hand side of (147) is upper bounded as

$$\mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}}\ \mathbb{E}_{p_{Z_S}}\mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}|Z_S}||P_{W_1F_1|Z_S}p^U_{W_2}p^U_{F_2}\right)>\tilde{\epsilon}\right)$$
$$=1-$$
$$\mathbb{P}_{p_{\mathcal{B}}}\left(\max_{S\in\mathcal{S}}\ \mathbb{E}_{p_{Z_S}}\mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}|Z_S}||P_{W_1F_1|Z_S}p^U_{W_2}p^U_{F_2}\right)\leq\tilde{\epsilon}\right) \tag{196}$$

$$=1-\mathbb{P}_{p_{\mathcal{B}}}\left(\mathbb{E}_{p_{Z_S}}\mathbb{D}\left(P_{W_{[1:2]}F_{[1:2]}|Z_S}||P_{W_1F_1|Z_S}p^U_{W_2}p^U_{F_2}\right)\leq\tilde{\epsilon}\ \text{for all } S\in\mathcal{S}\right) \tag{197}$$

$$\leq 1-\mathbb{P}_{p_{\mathcal{B}}}(\mathcal{G})=\mathbb{P}_{p_{\mathcal{B}}}(\mathcal{G}^c) \tag{198}$$

$$\leq|\mathcal{S}||\mathcal{Z}|\exp\left(\frac{-\epsilon^2(1-\delta)2^{\gamma_{21}}}{3\tilde{W}_2\tilde{F}_2}\right), \tag{199}$$

where (199) follows from (175).

Finally, by rewriting (146) with switching the roles of (W_1, F_1) and (W_2, F_2) and repeating the whole proof, we obtain the second term in the minimum in (39), which completes the proof for Lemma 2.

References

[1] M. Nafea and A. Yener, "The multiple access wiretap channel II with a noisy main channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2016, pp. 2983–2987.

[2] M. Nafea and A. Yener, "A new multiple access wiretap channel model," in *Proc. IEEE Inf. Theory Workshop*, Sep. 2016, pp. 349–353.

[3] L. H. Ozarow and A. D. Wyner, "Wire-tap channel II," *AT&T Bell Lab. Tech. J.*, vol. 63, no. 10, pp. 2135–2157, 1984.

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

[5] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[6] V. K. Wei, "Generalized Hamming weights for linear codes," *IEEE Trans. Inf. Theory*, vol. 37, no. 5, pp. 1412–1418, Sep. 1991.

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

[8] R. Liu, Y. Liang, H. V. Poor, and P. Spasojevic, "Secure nested codes for type II wiretap channels," in *Proc. IEEE Inf. Theory Workshop*, Sep. 2007, pp. 337–342.

[9] V. Aggarwal, L. Lai, A. R. Calderbank, and H. V. Poor, "Wiretap channel type II with an active eavesdropper," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun./Jul. 2009, pp. 1944–1948.

[10] M. Bloch, M. Hayashi, and A. Thangaraj, "Error-control coding for physical-layer secrecy," *Proc. IEEE*, vol. 103, no. 10, pp. 1725–1746, Oct. 2015.

[11] M. Nafea and A. Yener, "Wiretap channel II with a noisy main channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2015, pp. 1159–1163.

[12] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Semantic-security capacity for wiretap channels of type II," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3863–3879, Jul. 2016.

[13] M. Nafea and A. Yener, "A new wiretap channel model and its strong secrecy capacity," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 2077–2092, Mar. 2018.

[14] N. Cai and R. W. Yeung, "Secure network coding on a wiretap network," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 424–435, Jan. 2011.

[15] S. El Rouayheb, E. Soljanin, and A. Sprintson, "Secure network coding for wiretap networks of type II," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1361–1371, Mar. 2012.

[16] N. B. Shah, K. V. Rashmi, and P. V. Kumar, "Information-theoretically secure regenerating codes for distributed storage," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2011, pp. 1–5.

[17] S. Pawar, S. El Rouayheb, and K. Ramchandran, "Securing dynamic distributed storage systems against eavesdropping and adversarial attacks," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6734–6753, Oct. 2011.

[18] R. Bassily and A. Smith, "Causal erasure channels," in *Proc. ACM-SIAM Symp. Discrete Algorithms*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2014, pp. 1844–1857.

[19] A. Yener and S. Ulukus, "Wireless physical-layer security: Lessons learned from information theory," *Proc. IEEE*, vol. 103, no. 10, pp. 1814–1825, Oct. 2015.

[20] E. Tekin and A. Yener, "The Gaussian multiple access wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5747–5755, Dec. 2008.

[21] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[22] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[23] R. Liu, I. Maric, P. Spasojević, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[24] X. He and A. Yener, "The role of feedback in two-way secure communications," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8115–8130, Dec. 2013.

[25] X. He and A. Yener, "The Gaussian many-to-one interference channel with confidential messages," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2730–2745, May 2011.

[26] X. He and A. Yener, "MIMO wiretap channels with unknown and varying eavesdropper channel states," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6844–6869, Nov. 2014.

[27] E. MolavianJazi, M. Bloch, and J. N. Laneman, "Arbitrary jamming can preclude secure communication," in *Proc. Annu. Allerton Conf. Commun., Control, Comput.*, Sep./Oct. 2009, pp. 1069–1075.

[28] R. Schaefer, H. Boche, and H. Poor, "Secure communication under channel uncertainty and adversarial attacks," *Proc. IEEE*, vol. 103, no. 10, pp. 1796–1813, Oct. 2015.

[29] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Arbitrarily varying wiretap channels with type constrained states," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7216–7244, Dec. 2016.

[30] M. H. Yassaee and M. R. Aref, "Multiple access wiretap channels with strong secrecy," in *Proc. IEEE Inf. Theory Workshop*, Aug./Sep. 2010, pp. 1–5.

[31] R. Ahlswede and I. Csiszàr, "Common randomness in information theory and cryptography—Part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[32] M. H. Yassaee, M. R. Aref, and A. Gohari, "Achievability proof via output statistics of random binning," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6760–6786, Nov. 2014.

[33] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[34] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Stat. Assoc.*, vol. 58, no. 301, pp. 13–30, 1963.

[35] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[36] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[37] B. Patrick, *Probability and Measure*. Hoboken, NJ, USA: Wiley, 1995.

[38] O. Simeone and A. Yener, "The cognitive multiple access wire-tap channel," in *Proc. IEEE Conf. Inf. Sci. Syst.*, Mar. 2009, pp. 158–163.

Mohamed Nafea (S'11–M'19) received the B.Sc. degree in electrical engineering from Alexandria University, Egypt, in 2010; the M.Sc. degree in wireless communications from Wireless Intelligent Networks Center (WINC), Nile University, Egypt, in 2012; the M.A. degree in mathematics from Department of Mathematics, The Pennsylvania State University, University Park, PA, USA, in 2017, and the Ph.D. degree in electrical engineering from Wireless Communications and Networking Laboratory (WCAN Lab), The Pennsylvania State University, University Park, PA, USA, in 2018. He is currently a Postdoctoral Scholar at the Electrical and Computer Engineering Department, Georgia Institute of Technology, Atlanta, GA, USA. His research interests include network information theory, information theoretic security, inference of causal interaction networks, statistical learning, probabilistic graphical models, and algorithmic fairness.

Aylin Yener (S'91–M'01–SM'14–F'15) received the B.Sc. degree in electrical and electronics engineering and the B.Sc. degree in physics from Bogazici University, Istanbul, Turkey, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Wireless Information Network Laboratory (WINLAB), Rutgers University, New Brunswick, NJ, USA. She is a Distinguished Professor of Electrical Engineering at The Pennsylvania State University, University Park, PA, USA, where she joined the faculty as an assistant professor in 2002. Since 2017, she is also a Dean's Fellow in the College of Engineering at The Pennsylvania State University. She was a visiting professor of Electrical Engineering at Stanford University in 2016-2018 and a visiting associate professor in the same department in 2008-2009. Her current research interests are in information security, green communications, caching systems, and more generally in the fields of information theory, communication theory and networked systems. She received the NSF CAREER Award in 2003, the Best Paper Award in Communication Theory from the IEEE International Conference on Communications in 2010, the Penn State Engineering Alumni Society (PSEAS) Outstanding Research Award in 2010, the IEEE Marconi Prize Paper Award in 2014, the PSEAS Premier Research Award in 2014, the Leonard A. Doggett Award for Outstanding Writing in Electrical Engineering at Penn State in 2014, and the IEEE Women in Communications Engineering Outstanding Achievement Award in 2018. She is a distinguished lecturer for the IEEE Information Theory Society (2019-2020), the IEEE Communications Society (2018-2020) and the IEEE Vehicular Technology Society (2017-2019).

Dr. Yener is serving as the vice president of the IEEE Information Theory Society in 2019. Previously she was the second vice president (2018), member of the Board of Governors (2015-2018) and the treasurer (2012-2014) of the IEEE Information Theory Society. She served as the Student Committee Chair for the IEEE Information Theory Society (2007-2011), and was the co-Founder of the Annual School of Information Theory in North America in 2008. She was a Technical (Co)-Chair for various symposia/tracks at the IEEE ICC, PIMRC, VTC, WCNC, and Asilomar in 2005, 2008-2014 and 2018. Previously, she served as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS (2009-2012), an Editor for the IEEE TRANSACTIONS ON MOBILE COMPUTING (2017-2018), and an Editor and an Editorial Advisory Board Member for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2001-2012). She also served a Guest Editor for the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY in 2011, and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS in 2015. Currently, she serves as a Senior Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.