The Multiple Access Wiretap Channel II with a Noisy Main Channel

Mohamed Nafea and Aylin Yener

Wireless Communications and Networking Laboratory (WCAN)
Electrical Engineering Department
The Pennsylvania State University, University Park, PA 16802.
mnafEA@psu.edu yener@engr.psu.edu

Abstract—A two transmitter multiple access wiretap channel II (MAC-WT-II) with a discrete memoryless (DM) main channel is investigated. Two models for the wiretapper, who chooses a fixed-length subset of the channel uses and observes erasures outside this subset, are proposed. In the first model, in each position of the subset, the wiretapper noiselessly observes either the first or the second user’s symbol, while in the second model, the wiretapper observes a noiseless superposition of the two symbols. Achievable strong secrecy rate regions for the two models are derived. The secrecy of the keys in the dual problem in the source model is established by deriving a lemma which provides a doubly exponential convergence rate for the probability of the keys being uniform and independent from the wiretapper’s observation. The results extend the recently examined WTC-II with a DM main channel to a multiple access setting.

I. INTRODUCTION

The wiretap channel II (WTC-II), in which the legitimate terminals communicate over a noiseless channel while the wiretapper has perfect access to a fixed fraction of her choosing of the transmitted symbols, was introduced in [1]. This model while being similar to the discrete memoryless (DM) WTC with a noiseless main channel and binary erasure wiretapper channel, models a more capable wiretapper since the wiretapper noiselessly observes either user in each tapped position, while in second model, the wiretapper is proposed. The wiretapper in the first model observes either user in each tapped position, while in second model, she observes a noiseless superposition of the two users. Achievable strong secrecy rate regions for both models are derived by adopting the output statistics of random binning framework in [8], where an appropriate dual source coding problem is solved and the solution is converted to the original model using probability distribution approximation arguments.

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

Notation: For $S, X_S = \{X_i\}_{i\in S}$, $p_{X}^{S}$ denotes a uniform distribution over $X$. $\forall (p_X, q_X)$, $D(p_X||q_X)$ denote the variational distance and K-L divergence between $p_X$ and $q_X$. $\text{Conv}(R)$ denotes the convex hull of region $R$.

II. CHANNEL MODEL

Consider the model in Fig. 1. The main channel consists of two finite input alphabets $X_1, X_2$, a finite output alphabet $Y$, and transition probability $p_{Y|X_1, X_2}$. Each transmitter wishes to reliably communicate an independent message to a common receiver and to keep it secret from the wiretapper. To do so, the wiretapper maps its message $M_j$, uniformly distributed over $[1, 2^{nR_j}]$, into the codeword $X_j^n = [X_{j,1}, \ldots, X_{j,n}] \in X_j^n$ using a stochastic encoder, $j = 1, 2$. The receiver observes $Y^n \in Y^n$ and outputs the estimates $\hat{M}_j, j = 1, 2$. We consider the following two models for the wiretapper channel.

Model I: The wiretapper chooses the subset $S_p \in S_p$ and the sequence $S_u \in \{1, 2\}^n$, where $S_p = \{S_p : S_p \subseteq [1, n], |S_p| = \mu, \mu \leq n, \alpha = \frac{\mu}{n} \in [0, 1]\}$. That is, $S_p$ represents the subset of positions tapped by the wiretapper and $S_u$ represents her sequence of decisions to observe either the first or the second user symbols. Let $S_p(k), S_u(k)$ denote the $k$th elements of $S_p$ and $S_u$, and let $S$ be a set of pairs which represents the wiretapper strategy, and defined as

$$S \triangleq \{(S_p(k), S_u(k)) : k = 1, 2, \ldots, \mu\} \subseteq S,$$  

where $S$ is the set of all possible strategies for the wiretapper. The wiretapper observes $Z_{S}^{i} = [Z_1^{S}, \ldots, Z_n^{S}] \in Z^n$, where

$$Z_{S}^{i} = \begin{cases} X_{j,i}, & (i, j) \in S \\ ?, & \text{otherwise}. \end{cases}$$

Model II: The wiretapper chooses the subset $S \subseteq [1, n], |S| = \mu \leq n, \alpha = \frac{\mu}{n} \in [0, 1]$,

$$S \triangleq \{S : S \subseteq [1, n], |S| = \mu \leq n, \alpha = \frac{\mu}{n} \in [0, 1]\}.$$
and observes \( Z^n_S = [Z^n_1, \ldots, Z^n_S] \in \mathbb{Z}^n \), where
\[
Z^n_i = \begin{cases} X_{1,i} + X_{2,i} & i \in S \\ \hat{?} & \text{otherwise.} \end{cases}
\]

That is, the wiretapper observes a noiseless superposition of the two symbols in the subset \( S \), and erasures otherwise.

Remark 2

The individual rates in Theorem 1 represent the worst-case scenarios in which the wiretapper chooses to observe only one user’s symbols in the worst-case among all the possible sets.

Remark 2

\( \mathcal{R}^n_{\alpha} \subset \mathcal{R}^n \), as every \( (R_1, R_2) \in \mathcal{R}^n_{\alpha} \) also belongs to \( \mathcal{R}^n \). The wiretapper chooses to observe only one user’s symbols in the worst-case scenarios among all the possible sets.

III. MAIN RESULTS

Theorem 1 For \( \alpha \in [0,1] \), an achievable secret rate region for the MAC-WT-II with the wiretapper Model I is
\[
\mathcal{R}^I_{\alpha} = \text{Conv} \bigcup_{p_{U_1}, p_{U_2}} \left\{ (R_1, R_2) : R_1 \leq I(U_1; Y|U_2) - \alpha I(U_1; X_1) , R_2 \leq I(U_2; Y|U_1) - \alpha I(U_2; X_2) , R_1 + R_2 \leq I(U_1, U_2; Y) - \alpha I(U_1, U_2; X_1, X_2) \right\},
\]
where the union is over all distributions \( p_{U_1, U_2} \) which satisfy the Markov chains \( U_1 - X_1 - Y - U_2 - X_2 \).

Theorem 2 For \( \alpha \in [0,1] \), an achievable secret rate region for the MAC-WT-II with the wiretapper Model II is
\[
\mathcal{R}^\Pi_{\alpha} = \text{Conv} \bigcup_{p_{U_1}, p_{U_2}} \left\{ (R_1, R_2) : R_1 \leq I(U_1; Y|U_2) - \alpha I(U_1; X_1 + X_2) , R_2 \leq I(U_2; Y|U_1) - \alpha I(U_2; X_1 + X_2) , R_1 + R_2 \leq I(U_1, U_2; Y) - \alpha I(U_1, U_2; X_1 + X_2) \right\},
\]
where the union is over all distributions \( p_{U_1, U_2} \) which satisfy the Markov chains \( U_1 - X_1 - Y - U_2 - X_2 \).

The proofs for Theorems 1, 2 are provided in Section IV.

Remark 1 The individual rates in Theorem 1 represent the worst-case scenarios in which the wiretapper chooses to observe only one user’s symbols in the worst-case among all the possible sets.

Remark 2 \( \mathcal{R}^n_{\alpha} \subset \mathcal{R}^n \), as every \( (R_1, R_2) \in \mathcal{R}^n_{\alpha} \) also belongs to \( \mathcal{R}^n \). The wiretapper chooses to observe only one user’s symbols in the worst-case scenarios among all the possible sets.

IV. PROOFS FOR THEOREM 1 AND THEOREM 2

We first prove Theorem 1. Let us first consider \( U_1 = X_1 \) and \( U_2 = X_2 \). We fix \( p_{X_1,X_2} = p_{X_1}p_{X_2} \) and describe two protocols, where each protocol defines a set of random variables and induces a joint distribution over them. Throughout the paper, we use the convention \( A_{[i,j]} = (A_i, A_j) \) for random variables (vectors) and their realizations.
in the variational distance sense to (10), and that protocol A is reliable and secure, (ii) we utilize the closeness of the two induced distributions to show that, under the same rate conditions, protocol B is reliable and secure as well, and finally (iii) we eliminate the assumed common randomness $C_1,C_2$ from protocol B by conditioning on certain instances of them.

We now state the following two lemmas by which we derive conditions on the rates $R_j,R_j,j = 1,2$, required for the closeness of the two induced distributions and the security of protocol A. In particular, Lemma 1 provides an exponential decay of the average variational distance between the two induced distributions, which is used to show a convergence in probability result needed in the proof. Lemma 2 provides a doubly-exponential decay of the probability of not achieving secrecy for protocol A, which is needed, with the union bound, to guarantee secrecy for the exponentially many choices of $S$.

**Lemma 1** Let $X_j \triangleq \{X_j,p_{X_j}\}$, $X_2 \triangleq \{X_2,p_{X_2}\}$ be two independent sources. The source $X_j$ is randomly binned into $M_j = B_{1/2}^{(j)}(X_j)$, $C_j = B_{1/2}^{(j)}(X_j)$, where $B_{1/2}^{(j)}(X_j)$ are independent and uniform over $\{0,1\}$, $\{1,1\}$. Let $B \triangleq \{B_{1/2}^{(j)}(x_j)\}_{i=1,2,\infty} \in \mathbb{R}$, and for $\gamma_j > 0$, $j = 1,2$, define $D_{\gamma_j} \triangleq \{x_j \in \mathcal{X}_j : \log \frac{1}{p_{X_j}(x_j)} > \gamma_j\}$. Then, we have

$$\mathbb{E}(P_{\gamma_1}^{(j)}(S_{\gamma_j})) \leq \sum_{i=1}^{\gamma_j} \mathbb{P}(X_j \notin D_{\gamma_j}) + \frac{1}{2}(M_jC_j)^2 - \gamma_j^2). \quad (11)$$

**Proof:** Using the triangle inequality, we obtain

$$\mathbb{V}(P_{\gamma_1}^{(j)}(S_{\gamma_j})) \leq \sum_{i=1}^{\gamma_j} \mathbb{V}(P_{\gamma_1}(S_{\gamma_j})) \leq \sum_{j=1}^{\gamma_j} \mathbb{V}(X_j) \notin D_{\gamma_j} + \frac{1}{2}(M_jC_j)^2 - \gamma_j^2).$$

Let $X_1 \triangleq \{X_1,p_{X_1}\}$, $X_2 \triangleq \{X_2,p_{X_2}\}$ be two independent sources, both correlated with the compound source $\{Z_S\} \triangleq \{Z,p_{Z_S}\}, S \in \mathcal{S}$, where $|X_1|,|X_2|,|Z|,|S| < \infty$. The source $X_1$ is randomly binned into $M_1,C_1,j$ as in Lemma 1. For $\gamma_j > 0$, $j = 1,2$, i.e. $j \neq j$, and any $S \in \mathcal{S}$, define

$$D_j \triangleq \{(x_{i,j},z) \in \mathcal{X}_i \times \mathcal{X}_j : \gamma_j \gamma_j \in \gamma_j, (x_{i,j},z) \in D_j \gamma_j \},$$

where $D_j \gamma_j \triangleq \{(x_{i,j},z) : -\log p_{X_i}(x_{i,j}) > \gamma_j\}, j \neq j$, and $D_j \gamma_j \triangleq \{(x_{i,j},z) : -\log p_{X_i}(x_{i,j}) > \gamma_j\}.$

If $\exists j \in [0,\frac{1}{4}]\ s.t. \forall S \in \mathcal{S}$, $\operatorname{min}_{j=1,2} \mathbb{P}_{p_{X_i},p_{Z_i}}(X_{i,j},Z_S) \in D_j \gamma_j > 1 - \delta$, then we have, for every $\epsilon \in [0,1]$, that

$$\mathbb{P}\left(\max_{S \in \mathcal{S}} \mathbb{D}(P_{\gamma_1}^{(j)}(S_{\gamma_j})) \leq 2\epsilon \right) \leq \epsilon^{2\gamma_j} \prod_{i,j=1,2,\neq j} \left(1 - \epsilon + H_2(\gamma_j^2)\right), \quad (12)$$

where $\epsilon = \max_{j=1,2} \epsilon + (\delta + \delta^3) \log(2M_1C_j) + H_2(\delta^2), H_2$ is the binary entropy function, and $P$ is the induced distribution.

**Proof:** See the Appendix.

We now use Lemma 1 to establish the closeness of the induced distributions. In Lemma 1, set $X_j = X_j, M_j = 2^{nR_j}, C_j = 2^{nR_j}, \gamma_j = (1 - \epsilon)H(X_j), j = 1,2$ (where $X_j$ is defined as in protocol A). Note that for $\gamma_j < \infty$, any $x_j$ with $p(x_j) > 0$ belongs to $D_\gamma_j, j = 1,2$, by definition. Thus, in order to calculate $P(D_\gamma_j)$, we only consider $x_j$s with $p(x_j) > 0$. Without loss of generality, let $p(x_j) > 0, \forall x_j \in X_j$. Let $p_{\gamma_j}^\min = \min_{x_j} p(x_j)$. The random variables $\frac{1}{p(x_j)}, i \in [1,n], j$ are i.i.d. and each is bounded by the interval $[0,\log \frac{1}{p_{\gamma_j}^\min}]$. Using Hoeffding inequality [10], for any $\epsilon > 0$, $\exists \beta_j > 0$ s.t.

$$P(D_\gamma_j) = \mathbb{P}\left(\sum_{k=1}^{n} \frac{1}{p(x_j)} \leq (1 - \epsilon)nH(X_j)\right) \leq e^{-\beta_j n}.$$
there exists $\tilde{\beta}_{ij} > 0$ s.t. $\Pr(\{X_{1,2}, Z_S\} \notin D^S_{\tilde{\beta}_{ij}}) \leq \exp(-\tilde{\beta}_{ij} n)$. Taking $\delta^2 = 2 \exp(-\tilde{\beta}_{ij})$, where $\tilde{\beta} = \min\{\tilde{\beta}_{ij}, \tilde{\beta}_{ij}, \tilde{\beta}_{ij}\}$, gives $\Pr(\{X_{1,2}, Z_S\} \notin D^S_{\tilde{\beta}_{ij}}) \leq \delta^2$, for all $S \in S$ and $j = 1, 2$. Note that $\lim_{n \to \infty} \delta^2 = 0$, and hence, for $n$ sufficiently large, $\delta^2 \leq 0, 1/4$. Using (12), we have, for every $\epsilon, \epsilon_1 > 0, \epsilon = \epsilon + \epsilon_1$, there exists $n^* \in \mathbb{N}$ and $\kappa_n, \tilde{\kappa} > 0$ s.t. for all $n \geq n^*$,

$$\Pr(\text{max}_{S \in S} \{\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq \epsilon) \leq \epsilon, \epsilon \in \mathbb{N}^*$$

when $R_j + \tilde{R}_j < (1 - \epsilon)(1 - \alpha) H(X_j)$, $\forall j = 1, 2$.

(15) since $|S| \geq |\mathbb{N}| \leq \exp[n(1 + \alpha) + 2 n \ln |X_1| + |X_2| + 1)]$. Let $D_n \triangleq \max_{S \in S} \{\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq n \exp[n(1 + \alpha) + 2 n \ln |X_1| + |X_2| + 1)]$. Let $\kappa_n \triangleq \{D_n \geq r\}, r > 0$. By (15), $\sum_{n=1}^{\infty} \Pr(\kappa_n) < \infty$. Thus, $\Pr(\kappa_n$ infinitely often (i.o.)$) = 0$ by the Borel-Cantelli lemma. This implies that $\lim_{n \to \infty} \gamma_n = 0$, and $\Pr(\{D_n \geq r\}$ i.o. $) = 1$, i.e., $D_n$ converges to 0 almost surely. Thus, as $n \to \infty$, we have

$$\Pr(\text{max}_{S \in S} \{\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq \epsilon) \to 0. \quad (16)$$

Now, we show that protocol B is also reliable and secure with the rate conditions above. (13) and (14) imply that

$$\lim_{n \to \infty} \Pr(V(\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq \epsilon) \to 0. \quad (17)$$

Similar to the derivation of (16), Markov inequality and (13) imply that $\lim_{n \to \infty} \Pr(\{V(\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq \epsilon) \to 0$. Thus, by the union bound and (16), we have

$$\lim_{n \to \infty} \Pr(V(\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq \epsilon) \to 0. \quad (18)$$

The selection lemma [12, Lemma 2.2] when applied to (17), (18), implies that there is at least one binning realization $\tilde{b}^*$, with a corresponding joint distribution $p_{S_{ij}, S_{ij}}^* \} \text{for protocol B, s.t.}$,

$$\lim_{n \to \infty} \Pr(\{max_{S \in S} \{\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq \epsilon) \to 0, \quad (19)$$

with $M_j = b_{ij}^{(1)}(X_j), C_j = b_{ij}^{(2)}(X_j), j = 1, 2$. We introduce $p_{M_{ij}, X_{ij}}^* = \{M_j = b_{ij}^{(1)}(X_j), \forall j = 1, 2\}$ to (19). Then,

$$\Pr(\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq n \to \infty 0. \quad (20)$$

follows from (19). Using the union bound, we also have

$$\Pr(\tilde{\beta}_{ij}^* S_{ij}, S_{ij}^* \} \leq n \to \infty 0. \quad (21)$$

as the second term in the RHS of (21) is equal to zero.

Applying the selection lemma to (20), (21), implies that there is at least one $c_{ij}^* \} \text{s.t. both } P(M_{ij} \neq M_{ij}^* S_{ij}^* \} \leq n \to \infty 0. \quad (21)$$

and $\max_{S \in S} I(M_{ij} I_X; Z_S ) \leq c_{ij}^* \} \text{ converge to zero as } n \to \infty$. Let $p^*$ be the induced distribution for protocol A corresponding to $b^*$. We use $p^*(x_{ij}^* | M_{ij}^* S_{ij}^* \} \leq n \to \infty 0. \quad (21)$ as the encoder and $(p^*(x_{ij}^* | y_{ij}^*, c_{ij}^* \} \leq n \to \infty 0. \quad (21)$ as the decoder for the original model.

Combining the conditions $R_j + \tilde{R}_j < (1 - \epsilon)(1 - \alpha) H(X_j)$, $j = 1, 2$; $R_1 \geq H(X_1 X_2 Y), R_2 \geq H(X_2 X_1 Y), R_1 + \tilde{R}_j < H(X_1 X_2 Y)$, and taking $\epsilon \to 0$, establish the achievability of the union over all $p_{X_1 Y_1}^* S_1$, of the region of all pairs $(R_1, R_2)$ satisfying $R_1 \leq I(X_1 Y_1 X_2 Y) - \alpha H(X_1), R_2 \leq I(X_2 Y_1 X_2 Y) - \alpha H(X_2 X_1 Y)$, and $R_1 + \tilde{R}_j < I(X_2 Y_1 X_2 Y) - \alpha H(X_1 X_2 Y)$, by prefixing two independent channels, $p_{X_1 Y_1}^* S_1$, at the transmitters of the original model, we obtain the achievability of the union of the region in (7). The convex hull of the union follows by time sharing independent codewords and that 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.

The proof for Theorem 2 is similar to the proof of (7). The difference is that $S, Z_S, VS \in S$, in protocol A are as in (3), (4). Applying Lemma 2 to protocol A, after prefixing the channels $p_{X_1 Y_1} S_1, p_{X_2 Y_2} S_2$, gives, VS s, $i, j = 1, 2, i \neq j$;

$$H(U_j | Z_S) = n [[(1 - \alpha) H(U_j) + \alpha H(U_j X_1 X_2)] - \alpha H(U_j X_1 X_2)]$$

which we use, along with Hoeffding inequality, to satisfy the conditions of the lemma and derive the rate conditions,

$$R_j + \tilde{R}_j < (1 - \alpha) H(U_j) + \alpha H(U_j X_1 X_2), j = 1, 2,$$

$$\sum_{j=1}^{2} R_j + \tilde{R}_j < (1 - \alpha) H(U_j) + \alpha H(U_j X_1 X_2), j = 1, 2,$$

needed for secrecy. These conditions, combined with $\tilde{R}_j \geq H(U_1 X_2 Y), \tilde{R}_2 \geq H(U_2 X_1 Y), \tilde{R}_1 + \tilde{R}_2 \geq H(U_1, U_2) Y$ for the Slepian-Wolf decoder, and using time sharing (8).

**Remark 3** By setting $j = 1, i = 2$, instead of the minimum, in the RHS of (12), Lemma 2 results in the maximum rate $R_1 + \tilde{R}_1$, and the corresponding rate $R_2 + \tilde{R}_2$ (according to the maximum sum rate) such that the probability in the LHS of (12) is vanishing. By switching $i$ and $j$, the Lemma gives the maximum rate $R_2 + \tilde{R}_2$, and the corresponding rate $R_1 + \tilde{R}_1$ according to the maximum sum rate. Using this, one can deduce the maximum rate region, i.e., the maximum individual and sum rates, required for a vanishing probability.

**V. Conclusion**

In this paper, we have extended the WTC-II with a DM main channel [4] to a multiple access setting. We have proposed two models for the wiretapper and derived a strong secrecy achievable rate region for each. The achievable rate region for the model, where the wiretapper observes a noiseless
superposition of the two signals in the positions of the subset she selects, is larger than the achievable region for the more powerful wiretapper who decides to perfectly access either the first or the second signal at each position. The tools we have used for achievability extend a set of tools, utilized for a single-user scenario in a recent work [9], to a multi-user setting. Future work includes upper bounds for these models and other multi-terminal setups with more capable wiretappers.

APPENDIX

For all $S$, $D(P_{M|Z};C_{1|Z}||P_{M|Z}^{L}P_{C_{1|Z}}P_{Z})$ is equal to

$$E_{p_{Z}}\left(D(P_{M|Z};C_{1|Z}||P_{M|Z}^{L}P_{C_{1|Z}}P_{Z})\right) = D(P_{M|Z}^{L};P_{C_{1|Z}}P_{Z}).$$

Thus, the probability in the LHS of (12) is upper bounded by

$$P_{B}(\max_{S \in S} E_{p_{Z}}\left(D(P_{M|Z};C_{1|Z}||P_{M|Z}^{L}P_{C_{1|Z}}P_{Z}) > \epsilon\right) + \epsilon P_{B}(\max_{S \in S} E_{p_{Z}}\left(D(P_{M|Z}^{L};P_{C_{1|Z}}P_{Z}) > \epsilon\right).$$

We upper bound each term in (23). For all $S \in \mathcal{S}$, define

$$\mathcal{A}_{S} \triangleq \{z \in Z : P_{p_{X},|z}x \geq \mathcal{D}_{S} \geq 1 - \epsilon\}.$$ 

Using Markov inequality, we have, for all $S \in \mathcal{S}$

$$P_{p_{Z}}(\mathcal{A}_{S}^{c}) \leq \frac{1}{\delta} P_{p_{X},|z}x \geq \mathcal{D}_{S} \leq \delta.\quad (24)$$

Let $\mathbf{I}_{(x,m,c)} \triangleq \{B_{j}^{(j)}(x_{j}) = m_{j}, B_{j}^{(j)}(x_{j}) = c_{j}, \forall j \in J\}$, where $J \subseteq \{1, 2\}$, and let $\mathbf{I}_{(x,m,c)}^{(j)} = \{B_{j}^{(j)}(x_{j}), z \in \mathcal{D}_{j}\}$ and $\mathbf{I}_{(x,m,c)}^{(j)} = \{B_{j}^{(j)}(x_{j}), z \in \mathcal{D}_{j}\}$. For any $m_{1}, c_{1}, 1, z \in Z$, and $S \in \mathcal{S}$, define

$$P_{1}^{S}(m_{1}, c_{1}|z) = \sum_{x_{1}} p(x_{1}|z) \mathbf{I}_{(x,m,c)}^{(1)},$$

$$P_{2}^{S}(m_{2}, c_{2}|z) = \sum_{x_{2}} p(x_{2}|z) \mathbf{I}_{(x,m,c)}^{(2)},$$

hence $P_{M|Z}^{L};C_{1|Z}|z = P_{S}^{L} + P_{S}^{C}$. For every $x_{2} \in \mathcal{X}_{2}$, define

$$U_{x_{2}} = \sum_{x_{1} \in \mathcal{X}_{1}} p(x_{1}|z), \mathbf{I}_{(x,m,c)}^{(1))} \mathbf{D}_{j}.$$ 

Thus, $|U_{x_{2}}| \leq \sum_{x_{1}} p(x_{1}|z) |p(x_{2}|x_{1}, z) \mathbf{I}_{(x,m,c)}^{(2)}, z \in \mathcal{D}_{j} \mathbf{D}_{j} \leq 2^{-\gamma_{2}}$. Therefore, $\mathcal{A}_{S} \leq \sum_{x_{2}} p(x_{2}|z) |p(x_{2}|z) \mathbf{I}_{(x,m,c)}^{(2)}, z \in \mathcal{D}_{j} \mathbf{D}_{j} \leq 2^{-\gamma_{2}}$.

Since $E_{p_{Z}} \mathbf{I}_{(x,m,c)}^{(1)} = \frac{1}{M_{2}C_{2}}$, we have

$$\tilde{m} = \sum_{x_{2}} E_{p_{Z}}(U_{x_{2}}) = \frac{1}{M_{2}C_{2}} P_{p_{X},|z}x \geq \mathcal{D}_{S} \leq \frac{\tilde{m}}{M_{2}C_{2}}.$$ 

Also, notice that $\sum_{m_{1}c_{1}} P_{S}^{L}(m_{1}, c_{1}|z) = \sum_{x_{2}} U_{x_{2}}$ since $\sum_{m_{1}c_{1}} \mathbf{I}_{(x,m,c)}^{(1)} = \mathbf{I}_{(x,m,c)}^{(2)}$. Using a variation of Chernoff bound [9, Lemma 3], we have, $\forall \epsilon \in [0, 1], z \in \mathcal{A}_{S},$

$$P_{B}(P_{S}^{L}(m_{1}, c_{1}|z) \geq \frac{1 + \epsilon}{M_{2}C_{2}} P_{M,C_{1}|z}(m_{1}, c_{1}|z)).$$

(25)

Therefore, the probability in the LHS of (12) is upper bounded by

$$P_{B}(\max_{S \in S} E_{p_{Z}}\left(D(P_{M|Z}^{L};P_{C_{1|Z}}P_{Z}) > \epsilon\right) + \epsilon P_{B}(\max_{S \in S} E_{p_{Z}}\left(D(P_{M|Z}^{L};P_{C_{1|Z}}P_{Z}) > \epsilon\right).$$

Using the union bound and (25), we have

$$P_{B}(\mathcal{G}^{c}) \leq |S| E_{p_{Z}}\left(D(P_{M|Z}^{L};P_{C_{1|Z}}P_{Z}) \leq \epsilon + (\delta + \epsilon) log(M_{2}C_{2}) + H_{B}\left(\delta^{2}\right) \leq \epsilon.$$ 

(27)

Then, the first probability in (23) is upper bounded by $P_{B}(\mathcal{G}^{c})$ in (27). Using similar arguments, we show that the second term in (23) is upper bounded by $|S| E_{p_{Z}}\left(D(P_{M|Z}^{L};P_{C_{1|Z}}P_{Z}) \leq \epsilon + (\delta + \epsilon) log(M_{2}C_{2}) + H_{B}\left(\delta^{2}\right) \leq \epsilon.$$ 

Finally, by rewriting (22) with switching the roles of $(M_{1}, C_{1})$ and $(M_{2}, C_{2})$ and repeating the proof, we obtain the second term in the minimum in (12), which completes the proof.

REFERENCES


