Broadcast Channels With Privacy Leakage Constraints

Ziv Goldfeld, *Student Member, IEEE*, Gerhard Kramer, *Fellow, IEEE*, and Haim H. Permuter, *Senior Member, IEEE*

***Abstract*— The broadcast channel (BC) with one common and two private messages with leakage constraints is studied, where leakage rate refers to the normalized mutual information between a message and a channel symbol string. Each private message is destined for a different user and the leakage rate to the other receiver must satisfy a constraint. This model captures several scenarios concerning secrecy, i.e., when both, either or neither of the private messages are secret. Inner and outer bounds on the leakage-capacity region are derived when the eavesdropper knows the codebook. The inner bound relies on a Marton-like code construction and the likelihood encoder. A uniform approximation lemma is established that states that the marginal distribution induced by the encoder on each of the bins in the Marton codebook is approximately uniform. Without leakage constraints the inner bound recovers Marton's region and the outer bound reduces to the UVW-outer bound. The bounds match for semi-deterministic (SD) and physically degraded (PD) BCs, as well as for BCs with a degraded message set. The leakage-capacity regions of the SD-BC and the BC with a degraded message set recover past results for different secrecy scenarios. A Blackwell BC example illustrates the results and shows how its leakage-capacity region changes from the capacity region without secrecy to the secrecy-capacity regions for different secrecy scenarios.**

***Index Terms*— Broadcast channel, Marton's inner bound, privacy leakage, secrecy, physical-layer security.**

I. Introduction

PUBLIC and confidential messages are often transmitted over the same channel. However, the underlying principles for constructing codes without and with secrecy are different. Without secrecy constraints, codes should use all available channel resources to reliably convey information to the destinations. Confidential messages, on the other hand, require that some channel resources are allocated to preserve security. We study relationships between the coding strategies and the fundamental limits of communication with and without secrecy. To this end we simultaneously account for secret and non-secret transmissions over a two-user broadcast channel (BC) by means of privacy leakage constraints (Fig. 1).

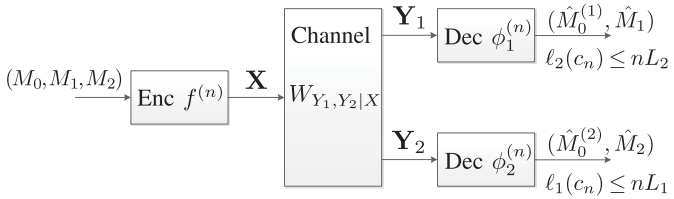

Fig. 1. A BC with a common message and privacy leakage constraints $\ell_1(c_n) \triangleq I_{c_n}(M_1; \mathbf{Y}_2) \le nL_1$ and $\ell_2(c_n) \triangleq I_{c_n}(M_2; \mathbf{Y}_1) \le nL_2$, where I_{c_n} denotes that the mutual information terms are taken with respect to the distribution induced by the code $c_n = \left(f^{(n)}, \phi_1^{(n)}, \phi_2^{(n)}\right)$.

A. Past Work

Information theoretic secrecy was introduced by Shannon [1] who studied communication between a source and a receiver in the presence of an eavesdropper. Wyner modeled secret communication over noisy channels (also known as physical layer security) when he introduced the degraded wiretap channel (WTC) and derived its secrecy capacity [2]. Csiszár and Körner [3] extended Wyner's result to a general BC where the source also transmits a common message to both users. The development of wireless communication, whose inherent open nature makes it vulnerable to security attacks, has inspired a growing interest in the fundamental limits of secure communication.

Multiuser settings with secrecy were extensively treated in the literature. Broadcast and interference channels with two confidential messages were studied in [4], where inner and outer bounds on the secrecy-capacity region of both problems were derived. The secrecy-capacity region for the semi-deterministic (SD) BC was established in [5]. The capacity region of a SD-BC where only the message of the stochastic user is kept secret from the deterministic user was derived in [6]. The opposite case, i.e., when the message of the deterministic user is confidential, was solved in [7]. Secret cooperative communication was considered in [8], where the authors derive inner and outer bounds on the rate-equivocation region of the relay-BC (RBC) with one or two confidential messages. Gaussian multiple-input multiple-output (MIMO)

Manuscript received April 26, 2015; revised April 19, 2016; accepted August 31, 2016. Date of publication May 25, 2017; date of current version July 12, 2017. Z. Goldfeld and H. H. Permuter were supported in part by the European Research Council through the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC under Grant 337752, in part by the Israel Science Foundation, and in part by the Cyber Security Research Center within the Ben-Gurion University of the Negev. G. Kramer was supported by an Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research.

Z. Goldfeld and H. H. Permuter are with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel (e-mail: gziv@post.bgu.ac.il; haimp@bgu.ac.il).

G. Kramer is with the Institute for Communications Engineering, Technical University of Munich, D-80333 Munich, Germany (e-mail: gerhard.kramer@tum.de).

Communicated by Y. Liang, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2017.2708086

BCs and WTCs were studied in [9]–[14], while [15]–[17] focused on BCs with an eavesdropper as an external entity from which all messages are kept secret.

Many of the aforementioned achievability results were derived by combining Marton's coding for BCs [18], [19] and Wyner's wiretap coding [2], [3]. Marton coding usually uses a joint typicality encoder (JTE) whose success is guaranteed by invoking the Mutual Covering Lemma (MCL) [20, Lemma 8.1]. However, the JTE and the MCL have a cumbersome security analysis. Several past works avoid the complications by performing the security analysis without conditioning on the random codebook. This significantly simplifies the derivations, but one would like to have security even if the codebooks are known by the eavesdropper.

B. Model

We study a two-user BC over which a common message for both users and a pair of private messages, each destined for a different user, are transmitted. A limited amount of rate of each private message may be leaked to the opposite receiver. The leaked rate is quantified as the normalized mutual information between the message of interest and the channel output sequence at the opposite user. Setting either leakage to zero or infinity reduces the problem to the case where the associated message is confidential or non-confidential, respectively. Thus, our problem setting specializes to all four scenarios concerning secrecy: when both, either or neither of the private messages are secret. We derive inner and outer bounds on the leakage-capacity region of the BC. The inner bound relies on a leakage-adaptive coding scheme that accounts for the codebook being known to the eavesdropper.

The derived bounds are tight for SD-BCs, physically degraded (PD) BCs, and BCs with a degraded message set, thus characterizing their leakage-capacity regions. Furthermore, we derive a condition for identifying the privacy leakage threshold above which the inner bound saturates. Various past results are captured as special cases. By taking the leakage thresholds to infinity, our inner bound recovers Marton's inner bound with a common message [21], which is tight for every BC with a known capacity region. Making the leakage constraint inactive in our outer bound recovers the UVW-outer bound [22] or the New-Jersey outer bound [23]. These bounds are at least as good as previously known bounds (see [24]–[26]). The leakage-capacity region of the SD-BC reduces to each of the regions in [5], [6], [21], and [27] by discarding the common message and choosing the leakage constraints appropriately. The capacity result also recovers the optimal regions for the BC with confidential messages [3] and the BC with a degraded message set (without secrecy) [28]. Finally, a Blackwell BC (BW-BC) [29], [30] illustrates the results and visualizes the transition of the leakage-capacity region from the capacity region without secrecy to the secrecy-capacity regions for different secrecy scenarios.

C. Organization

This paper is organized as follows. Section II establishes notation and preliminary definitions. In Section III we discuss the need for replacing the JTE with the likelihood encoder and state a Uniform Approximation Lemma. Section IV describes the BC with privacy leakage constraints, states inner and outer bounds on the leakage-capacity region and characterizes the optimal regions of several special cases. In Section V we discuss past results that are captured within our framework and Section VI visualizes the results by means of a BW-BC example. Finally, Section VIII summarizes the main achievements and insights of this work.

II. Notations and Preliminary Definitions

A. Notations

We use the following notations. As customary $\mathbb{N}$ is the set of natural numbers (which does not include 0), while $\mathbb{R}$ denotes the reals. We further define $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$ and $\mathbb{R}_{++} = \mathbb{R} \setminus \{0\}$. Given two real numbers a, b, we denote by $[a : b]$ the set of integers $\{n \in \mathbb{N} | a \leq n \leq b\}$. Calligraphic letters such are $\mathcal{X}$ denote sets, the complement of $\mathcal{X}$ is denoted by $\mathcal{X}^c$, while $|\mathcal{X}|$ stands for its cardinality. $\mathcal{X}^n$ denoted the n-fold Cartesian product of $\mathcal{X}$. An element of $\mathcal{X}^n$ is denoted by $x^n = (x_1, x_2, \ldots, x_n)$; whenever the dimension n is clear from the context, vectors (or sequences) are denoted by boldface letters, e.g., $\mathbf{x}$. A substring of $\mathbf{x} \in \mathcal{X}^n$ is denoted by $x_i^j = (x_i, x_{i+1}, \ldots, x_j)$, for $1 \leq i \leq j \leq n$; when $i = 1$, the subscript is omitted. We also define $x^{n \setminus i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

Let $(\mathcal{X}, \mathcal{F}, \mathbb{P})$ be a probability space, where $\mathcal{X}$ is the sample space, $\mathcal{F}$ is the σ-algebra and $\mathbb{P}$ is the probability measure. Random variables over $(\mathcal{X}, \mathcal{F}, \mathbb{P})$ are denoted by uppercase letters, e.g., X, with conventions for random vectors similar to those for deterministic vectors. The probability of an event $\mathcal{A} \in \mathcal{F}$ is denoted by $\mathbb{P}(\mathcal{A})$, while $\mathbb{P}(\mathcal{A}|\mathcal{B})$ denotes the conditional probability of $\mathcal{A}$ given $\mathcal{B}$. We use $\mathbb{1}_{\mathcal{A}}$ to denote the indicator function of $\mathcal{A}$. The set of all probability mass functions (PMFs) on a finite set $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{X})$, i.e.,

$$\mathcal{P}(\mathcal{X}) = \left\{ P : \mathcal{X} \to [0, 1] \middle| \sum_{x \in \mathcal{X}} P(x) = 1 \right\}. \tag{1}$$

PMFs are denoted by uppercase letters such as P or Q, often with a subscript that identifies the random variable and its possible conditioning. For example, for a discrete probability space $(\mathcal{X}, \mathcal{F}, \mathbb{P})$ and two correlated random variables X and Y over that space, we use P_X, $P_{X,Y}$ and $P_{X|Y}$ to denote, respectively, the marginal PMF of X, the joint PMF of (X, Y) and the conditional PMF of X given Y. In particular, $P_{X|Y}$ represents the stochastic matrix whose elements are given by $P_{X|Y}(x|y) = \mathbb{P}(X = x | Y = y)$. Expressions such as $P_{X,Y} = P_X P_{Y|X}$ are to be understood as $P_{X,Y}(x, y) = P_X(x) P_{Y|X}(y|x)$, for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Accordingly, when three random variables X, Y and Z satisfy $P_{X|Y,Z} = P_{X|Y}$, they form a Markov chain, which we denote by $X - Y - Z$. We omit subscripts if the arguments of a PMF are lowercase versions of the random variables. The support of a PMF P and the expectation of a random variable $X \sim P$ are denoted by $\mathrm{supp}(P)$ and $\mathbb{E}_P[X]$, respectively; when the distribution of X is clear from the context we write its expectation simply as $\mathbb{E}[X]$. Similarly, H_P and I_P denote entropy and mutual

information that are calculated with respect to an underlying PMF P.

For a discrete measurable space $(\mathcal{X}, \mathcal{F})$, a PMF $Q \in \mathcal{P}(\mathcal{X})$ gives rise to a probability measure on $(\mathcal{X}, \mathcal{F})$, which we denote by $\mathbb{P}_Q$; accordingly, $\mathbb{P}_Q(\mathcal{A}) = \sum_{x\in\mathcal{A}} Q(x)$, for every $\mathcal{A} \in \mathcal{F}$. For a random vector X^n, if the entries of X^n are drawn in an independent and identically distributed (i.i.d.) manner according to P_X, then for every $\mathbf{x} \in \mathcal{X}^n$ we have $P_{X^n}(\mathbf{x}) = \prod_{i=1}^n P_X(x_i)$ and we write $P_{X^n}(\mathbf{x}) = P_X^n(\mathbf{x})$. Similarly, if for every $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ we have $P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n P_{Y|X}(y_i|x_i)$, then we write $P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) = P_{Y|X}^n(\mathbf{y}|\mathbf{x})$. The conditional product PMF $P_{Y|X}^n$ given a specific sequence $\mathbf{x} \in \mathcal{X}^n$ is denoted by $P_{Y|X=\mathbf{x}}^n$.

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets. The empirical PMF $\nu_{\mathbf{x}}$ of a sequence $\mathbf{x} \in \mathcal{X}^n$ is

$$\nu_{\mathbf{x}}(x) \triangleq \frac{N(x|\mathbf{x})}{n} \tag{2}$$

where $N(x|\mathbf{x}) = \sum_{i=1}^n \mathbb{1}_{\{x_i=x\}}$. We use $\mathcal{T}_\delta^n(P_X)$ to denote the set of letter-typical sequences of length n with respect to the PMF $P_X \in \mathcal{P}(\mathcal{X})$ and the positive number δ [31, Ch. 3], i.e., we have

$$\mathcal{T}_\delta^n(P_X) = \Big\{\mathbf{x} \in \mathcal{X}^n \Big| \ \big|\nu_{\mathbf{x}}(x) - P_X(x)\big| \le \delta P_X(x), \ \forall x \in \mathcal{X}\Big\}. \tag{3}$$

Furthermore, for a joint PMF $P_{X,Y} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$, a $\delta > 0$, and a fixed sequence $\mathbf{y} \in \mathcal{Y}^n$, we define

$$\mathcal{T}_\delta^n(P_{X,Y}|\mathbf{y}) = \Big\{\mathbf{x} \in \mathcal{X}^n \Big| (\mathbf{x}, \mathbf{y}) \in \mathcal{T}_\delta^n(P_{X,Y})\Big\}. \tag{4}$$

Another notion used throughout this work is information density. Let $(\mathcal{X}\times\mathcal{Y}, \mathcal{F}, P_{X,Y})$ be a probability space, where $\mathcal{X}$ and $\mathcal{Y}$ are arbitrary sets. The *information density* $i_P : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}_+$ of $P_{X,Y}$ is given by

$$i_P(x; y) = \log \frac{\mathsf{d}P_{X,Y}}{\mathsf{d}P_X P_Y}(x, y) \tag{5a}$$

where $\frac{\mathsf{d}P}{\mathsf{d}Q}$ is the Radon-Nikodym derivative of P with respect to Q and P_X and P_Y are the marginal probability measures induced by $P_{X,Y}$ on $\mathcal{X}$ and $\mathcal{Y}$, respectively. If $\mathcal{X}$ and $\mathcal{Y}$ are discrete and $P_{X,Y} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$, then (5a) simplifies as

$$i_P(x; y) = \log \frac{P_{X,Y}(x, y)}{P_X(x)P_Y(y)}. \tag{5b}$$

Whenever the underlying distribution is clear from the context, we drop the subscript P from i_P.

B. Measures of Distribution Proximity

We measure the proximity between two distributions by using total variation (TV).

Definition 1 (Total Variation): Let $(\mathcal{X}, \mathcal{F})$ be a measurable space and P and Q be two probability measures on $\mathcal{F}$. The total variation between P and Q is

$$||P - Q||_{\mathsf{TV}} = \sup_{\mathcal{A}\in\mathcal{F}} \big|P(\mathcal{A}) - Q(\mathcal{A})\big|. \tag{6a}$$

If the sample space $\mathcal{X}$ is countable and $P, Q \in \mathcal{P}(\mathcal{X})$, then (6a) reduces to

$$||P - Q||_{\mathsf{TV}} = \frac{1}{2}\sum_{x\in\mathcal{X}} \big|P(x) - Q(x)\big|. \tag{6b}$$

We also consider the fidelity between two distributions.

Definition 2 (Fidelity): Let $(\mathcal{X}, \mathcal{F})$ be a measurable space and P and Q be two probability measures on $\mathcal{F}$, such that $P \ll Q$, i.e., P is absolutely continuous with respect to Q. The fidelity between P and Q is

$$\mathsf{F}(P, Q) = \mathbb{E}_Q\sqrt{\frac{\mathsf{d}P}{\mathsf{d}Q}}. \tag{7a}$$

If the sample space $\mathcal{X}$ is countable and $P, Q \in \mathcal{P}(\mathcal{X})$, then (7a) reduces to

$$\mathsf{F}(P, Q) = \sum_{x\in\mathcal{X}} \sqrt{P(x)Q(x)}. \tag{7b}$$

Fidelity satisfies $\mathsf{F}(P, Q) \in [0, 1]$ and is related to the TV as follows [32, Lemma 1].

Lemma 1 (Fidelity and Total Variation): For any two probability measures P and Q over the same measurable space $(\mathcal{X}, \mathcal{F})$, we have

$$1 - \mathsf{F}(P, Q) \le ||P - Q||_{\mathsf{TV}} \le \sqrt{1 - \mathsf{F}^2(P, Q)}. \tag{8}$$

Via Jensen's inequality, the right-most inequality in (8) extends to the expected values of the fidelity and the TV between two conditional distributions as follows [32, Lemma 2], [33, Lemma 7].

Lemma 2 (Extension to Expected Values): Let $(\mathbf{\Omega}, \mathcal{G}, \mu)$ be a probability space, $(\mathcal{X}, \mathcal{F})$ be a measurable space and P and Q be two transition probability kernels from $(\mathbf{\Omega}, \mathcal{G})$ to $(\mathcal{X}, \mathcal{F})$.[1] We have

$$\mathbb{E}_\mu\big|\big|P - Q\big|\big|_{\mathsf{TV}} \le \sqrt{1 - \Big(\mathbb{E}_\mu \mathsf{F}\big(P, Q\big)\Big)^2}. \tag{9}$$

By virtue of Lemma 2, if $\big\{P_n\big\}_{n\in\mathbb{N}}$ and $\big\{Q_n\big\}_{n\in\mathbb{N}}$ are two sequences of transition probability kernels[2] then

$$\mathbb{E}_{\mu_n}\mathsf{F}\big(P_n, Q_n\big) \xrightarrow[n\to\infty]{} 1 \tag{10a}$$

implies

$$\mathbb{E}_{\mu_n}\big|\big|P_n - Q_n\big|\big|_{\mathsf{TV}} \xrightarrow[n\to\infty]{} 0. \tag{10b}$$

III. Uniform Distribution Approximation Lemma

A. Marton Coding

A Marton code involves two independent codebooks from which a pair of codewords is usually selected by means of a JTE [19]. A standard tool for the encoding error probability analysis is the MCL [20, Lemma 8.1]. While the JTE and the MCL are convenient for analysing reliability, security (equivocation or leakage) analysis seems cumbersome.

[1] A transition probability kernel between two measurable spaces $(\mathbf{\Omega}, \mathcal{G})$ and $(\mathcal{X}, \mathcal{F})$ is a mapping $\kappa : \mathbf{\Omega} \times \mathcal{F} \to [0, 1]$ such that: (i) $\omega \mapsto \kappa(\omega, \mathcal{A})$ is a $\mathcal{G}$-measurable function for every $\mathcal{A} \in \mathcal{F}$; (ii) $\mathcal{A} \mapsto \kappa(\omega, \mathcal{A})$ is a probability measure on $(\mathcal{X}, \mathcal{F})$ for every $\omega \in \mathbf{\Omega}$.

[2] The formal definition is in accordance with Lemma 2 where we replace $(\mathbf{\Omega}, \mathcal{G}, \mu)$, $(\mathcal{X}, \mathcal{F})$, P and Q with the sequences $\big\{(\mathbf{\Omega}_n, \mathcal{G}_n, \mu_n)\big\}_n$, $\big\{(\mathcal{X}_n, \mathcal{F}_n)\big\}_n$, $\{P_n\}_n$ and $\{Q_n\}_n$, respectively.

Several past works employ Marton coding without conditioning the security analysis on the random codebook. Attempting to repeat the steps from these derivations while conditioning the equivocation on the codebook turns out to be problematic. The principal difficulty is showing that the marginal distribution of an index chosen by the JTE is approximately uniform.[3] More precisely, let the output index pair of the encoder be (I, J); the corresponding alphabets are $\mathcal{I}_n$ and $\mathcal{J}_n$. Several existing proofs rely on the following relations holding true:

$$H(I|\mathsf{C}_n) \geq \log|\mathcal{I}_n| - n\delta_n; \quad H(J|\mathsf{C}_n) \geq \log|\mathcal{J}_n| - n\delta_n, \tag{11}$$

where C_n is the random codebook[4] and $\lim_{n\to\infty}\delta_n = 0$. Proving these inequalities while using the JTE is cumbersome. A potential proof would rely on analysing the output distribution of the JTE. However, the complex structure of this distribution seems to makes the analysis intractable.

B. Likelihood Encoder

Our coding scheme also uses a Marton code. We circumvent the problems with the JTE by replacing it with a likelihood encoder for Marton codebooks [7], [34], [35]. A similar encoding rule was used in [36] and [32] under the name stochastic mutual information encoder. This encoder induces a probability distribution over the possible pairs of indices (or, equivalently, codewords). Given two independently generated bins, the probability of each codeword pair is proportional to the ratio of their joint probability (under the coding distribution) to the product of the marginal distributions. Namely, if $\mathbf{u}_i$ and $\mathbf{v}_j$ are the i-th and j-th codewords for each bin, respectively, and $Q_{U,V}$ is the coding distribution (the codebooks are generated by i.i.d. samples of the marginals Q_U and Q_V), then the encoder chooses (i, j) with probability proportional to

$$\frac{Q^n_{U,V}(\mathbf{u}_i, \mathbf{v}_j)}{Q^n_U(\mathbf{u}_i)Q^n_V(\mathbf{v}_j)}. \tag{12}$$

Thus, the further the joint distribution is from the product of the marginals the more favorable the corresponding pair of codewords is.

Replacing the JTE with the likelihood encoder comes at no cost in reliability. This is because, like the JTE, if the sum of the bin rates is greater than $I(U; V)$, then the likelihood encoder chooses jointly typical codeword pairs with high probability [32, Th. 3]. The leakage analysis, on the other hand, tremendously simplifies. This allows to derive the achievability result for the BC with privacy leakage constraints. Key to the leakage analysis is that the marginal distribution of the indices at the encoder's output is indeed approximately uniform. This relation is formulated in the next subsection and the proof is provided in Section VII-A.

[3]Without the conditioning, uniformity follows by symmetry.

[4]The conditioning on C_n is not present in many existing works. Instead, the relations (11) were replaced with their unconditioned versions $H(I) = \log|\mathcal{I}|$ and $H(J) = \log|\mathcal{J}|$. Although these relations are true, an unconditioned analysis does not imply achievability when the codebook is known to the eavesdropper.

C. Setup and Statement of the Lemma

For notational convenience we formulate the setup and state the result in terms of random variables with finite alphabets. Nonetheless, as can be seen in the proof of Lemma 3 (Section VII-A), the derivation is valid for random variables with general alphabets.

Fix $Q_{W,U,V} \in \mathcal{P}(\mathcal{W}\times\mathcal{U}\times\mathcal{V})$ and for every $n \in \mathbb{N}$ define $\mathcal{I}_n \triangleq \left[1 : 2^{nS_1}\right]$, $\mathcal{J}_n \triangleq \left[1 : 2^{nS_2}\right]$ and $\mathcal{K}_n = \left[1 : 2^{nT}\right]$, where $S_1, S_2, T \in \mathbb{R}_+$. Let $\mathbf{W} \sim Q^n_W$ and fix $\mathbf{w} \in \mathcal{W}^n$ with $Q^n_W(\mathbf{w}) > 0$. Let $\mathsf{B}^{(n)}_U(\mathbf{w}) \triangleq \left\{\mathbf{U}_i(\mathbf{w})\right\}_{i\in\mathcal{I}_n}$ be a random codebook that comprises $|\mathcal{I}_n|$ vectors of length n that are i.i.d. according to $Q^n_{U|W=\mathbf{w}}$. Furthermore, for every $k \in \mathcal{K}_n$ let $\mathsf{B}^{(n)}_V(k, \mathbf{w}) \triangleq \left\{\mathbf{V}_{j,k}(\mathbf{w})\right\}_{j\in\mathcal{J}_n}$ be a random codebook with i.i.d. codewords according to $Q^n_{V|W=\mathbf{w}}$. The codebooks in the set $\mathsf{B}^{(n)}_V(\mathbf{w}) \triangleq \left\{\mathsf{B}^{(n)}_V(k, \mathbf{w})\right\}_{k\in\mathcal{K}_n}$ are conditionally independent of one another given $\mathbf{W} = \mathbf{w}$. For any $\mathbf{w} \in \mathcal{W}^n$ with $Q^n_W(\mathbf{w}) > 0$ we also define $\mathsf{B}_n(\mathbf{w}) \triangleq \left\{\mathsf{B}^{(n)}_U(\mathbf{w}), \mathsf{B}^{(n)}_V(\mathbf{w})\right\}$ and finally we set $\mathsf{B}_n \triangleq \left\{\mathbf{W}, \mathsf{B}_n(\mathbf{W})\right\}$.

A realization of $\mathsf{B}^{(n)}_U(\mathbf{w})$ or $\mathsf{B}^{(n)}_V(k, \mathbf{w})$, $k \in \mathcal{K}_n$, is denoted by $\mathcal{B}^{(n)}_U(\mathbf{w}) \triangleq \left\{\mathbf{u}_i(\mathbf{w})\right\}_{i\in\mathcal{I}_n}$ and $\mathcal{B}^{(n)}_V(k, \mathbf{w}) \triangleq \left\{\mathbf{v}_{j,k}(\mathbf{w})\right\}_{j\in\mathcal{J}_n}$, respectively. In accordance to the above, we also set $\mathcal{B}^{(n)}_V(\mathbf{w}) \triangleq \left\{\mathcal{B}^{(n)}_V(k, \mathbf{w})\right\}_{k\in\mathcal{K}_n} = \left\{\mathbf{v}_{j,k}(\mathbf{w})\right\}_{(j,k)\in\mathcal{J}_n\times\mathcal{K}_n}$, $\mathcal{B}_n(\mathbf{w}) \triangleq \left\{\mathcal{B}^{(n)}_U(\mathbf{w}), \mathcal{B}^{(n)}_V(\mathbf{w})\right\}$ and $\mathcal{B}_n \triangleq \left\{\mathbf{w}, \mathcal{B}_n(\mathbf{w})\right\}$. Letting $\mathfrak{B}_n$ denote the collection of all possible realizations of B_n, the above construction induces a PMF $\lambda \in \mathcal{P}(\mathfrak{B}_n)$ on $\mathfrak{B}_n$ that is given by

$$\begin{aligned}\lambda(\mathcal{B}_n) = {} & Q^n_W(\mathbf{w})\prod_{i\in\mathcal{I}_n} Q^n_{U|W}\left(\mathbf{u}_i(\mathbf{w})\middle|\mathbf{w}\right) \\ & \times \prod_{(j,k)\in\mathcal{J}_n\times\mathcal{K}_n} Q^n_{V|W}\left(\mathbf{v}_{j,k}(\mathbf{w})\middle|\mathbf{w}\right).\end{aligned} \tag{13}$$

Now, let K be a random variable independent of B_n and uniformly distributed over $\mathcal{K}_n$. For each $\mathcal{B}_n \in \mathfrak{B}_n$ and $k \in \mathcal{K}_n$, the index pair $(i, j) \in \mathcal{I}_n \times \mathcal{J}_n$ is drawn according to

$$P^{(\mathcal{B}_n,k)}_{I,J}(i, j) = \frac{2^{i_{Q^n}\left(\mathbf{u}_i(\mathbf{w});\mathbf{v}_{j,k}(\mathbf{w})\middle|\mathbf{w}\right)}}{\sum\limits_{(\bar{\ell},\bar{j})\in\mathcal{I}_n\times\mathcal{J}_n} 2^{i_{Q^n}\left(\mathbf{u}_{\bar{\ell}}(\mathbf{w});\mathbf{v}_{\bar{j},k}(\mathbf{w})\middle|\mathbf{w}\right)}}, \tag{14}$$

where

$$i_{Q^n}(\mathbf{u}; \mathbf{v}|\mathbf{w}) = \log\frac{Q^n_{U,V|W}(\mathbf{u}, \mathbf{v}|\mathbf{w})}{Q^n_{U|W}(\mathbf{u}|\mathbf{w})Q^n_{V|W}(\mathbf{v}|\mathbf{w})}. \tag{15}$$

$P^{(\mathcal{B}_n,k)}_{I,J}$ describes our likelihood encoder. Finally, on account of (13)-(14) we set

$$P_{\mathsf{B}_n,K,I,J}(\mathcal{B}_n, k, i, j) \triangleq \lambda(\mathcal{B}_n)\frac{1}{|\mathcal{K}_n|}P^{(\mathcal{B}_n,k)}_{I,J}(i, j), \tag{16}$$

which induces a probability measure $\mathbb{P}_P$.

The following lemma specifies sufficient conditions on the sizes of the index sets for approximating the induced marginal distribution of I with a uniform distribution over $\mathcal{I}_n$. To state

the result let $p_{\mathcal{I}_n}^{(U)}$ be the uniform distribution over $\mathcal{I}_n$ and note that for every $\mathcal{B}_n \in \mathfrak{B}_n$, we have

$$P_{I|\mathsf{B}_n}(i|\mathcal{B}_n) = \frac{1}{|\mathcal{K}_n|} \sum_{\substack{(j,k) \\ \in \mathcal{J}_n \times \mathcal{K}_n}} \frac{2^{iQ^n}\left(\mathbf{u}_i(\mathbf{w}); \mathbf{v}_{j,k}(\mathbf{w}) \middle| \mathbf{w}\right)}{\sum\limits_{\substack{(\tilde{\ell},\tilde{j}) \\ \in \mathcal{I}_n \times \mathcal{J}_n}} 2^{iQ^n}\left(\mathbf{u}_{\tilde{\ell}}(\mathbf{w}); \mathbf{v}_{\tilde{j},k}(\mathbf{w}) \middle| \mathbf{w}\right)}. \quad (17)$$

Lemma 3 (Uniform Approximation Lemma): For any $Q_{W,U,V} \in \mathcal{P}(\mathcal{W} \times \mathcal{U} \times \mathcal{V})$ if

$$S_2 + \min\left\{S_1, T\right\} > I_{Q_{W,U,V}}(U; V|W) \quad (18a)$$

then

$$\mathbb{E}_{\mathsf{B}_n}\left|\left|P_{I|\mathsf{B}_n} - p_{\mathcal{I}_n}^{(U)}\right|\right|_{\mathsf{TV}} \xrightarrow[n\to\infty]{} 0. \quad (18b)$$

The Lemma is proven in Section VII-A via an analysis of the expected fidelity between the induced marginal distribution of I and the uniform distribution. Inspired by ideas from [32], we employ the Cauchy-Schwarz inequality and Jensen's inequality to show that the expected fidelity converges to 1 with the blocklengh. The result of the lemma then follows by (10).

IV. Broadcast Channels With Privacy Leakage Constraints

A. Problem Setting

The $\left(\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, W_{Y_1,Y_2|X} : \mathcal{X} \to \mathcal{P}(\mathcal{Y}_1 \times \mathcal{Y}_2)\right)$ BC with privacy leakage constraints is illustrated in Fig. 1. The channel has one sender and two receivers. The sender randomly chooses a triple (m_0, m_1, m_2) of indices uniformly and independently from the set $\left[1 : 2^{nR_0}\right] \times \left[1 : 2^{nR_1}\right] \times \left[1 : 2^{nR_2}\right]$ and maps them to a sequence $\mathbf{x} \in \mathcal{X}^n$, which is the channel input (the mapping may be random). The sequence $\mathbf{x}$ is transmitted over a BC with transition probability $W_{Y_1,Y_2|X}$. The output sequence $\mathbf{y}_j \in \mathcal{Y}_j^n$, where $j = 1, 2$, is received by decoder j. Decoder j produces a pair of estimates $\left(\hat{m}_0^{(j)}, \hat{m}_j\right)$ of (m_0, m_j).

Remark 1 (Specific Classes of BCs): We sometimes specialize to the following classes of BCs:

- *Semi-Deterministic BCs:* A BC is SD if its channel transition matrix factors as $W_{Y_1,Y_2|X} = \mathbb{1}_{\{Y_1 = y_1(X)\}} W_{Y_2|X}$, where $y_1 : \mathcal{X} \to \mathcal{Y}_1$ and $W_{Y_2|X} : \mathcal{X} \to \mathcal{P}(\mathcal{Y}_2)$.
- *Physically-Degraded BCs:* A BC is PD if its channel transition matrix factors as $W_{Y_1,Y_2|X} = W_{Y_1|X} W_{Y_2|Y_1}$, where $W_{Y_1|X} : \mathcal{X} \to \mathcal{P}(\mathcal{Y}_1)$ and $W_{Y_2|Y_1} : \mathcal{Y}_1 \to \mathcal{P}(\mathcal{Y}_2)$.
- *Deterministic BCs:* A BC is deterministic if its channel transition matrix factors as $W_{Y_1,Y_2|X} = \mathbb{1}_{\{Y_1 = y_1(X)\} \cap \{Y_2 = y_2(X)\}}$, where $y_j : \mathcal{X} \to \mathcal{Y}_j$, for $j = 1, 2$.

Definition 3 (Code): An (n, R_0, R_1, R_2) code c_n for the BC with leakage constraints has:

1) Three message sets $\mathcal{M}_j^{(n)} \triangleq \left[1 : 2^{nR_j}\right]$, $j = 0, 1, 2$.
2) A stochastic encoder $f^{(n)} : \mathcal{M}_0^{(n)} \times \mathcal{M}_1^{(n)} \times \mathcal{M}_2^{(n)} \to \mathcal{P}(\mathcal{X}^n)$.
3) Two decoding functions, $\phi_j^{(n)} : \mathcal{Y}_j^n \to \hat{\mathcal{M}}_{0j}^{(n)}$, where $\hat{\mathcal{M}}_{0j}^{(n)} \triangleq \mathcal{M}_0^{(n)} \times \mathcal{M}_j^{(n)}$, for $j = 1, 2$.

A code $c_n = \left(f^{(n)}, \phi_1^{(n)}, \phi_2^{(n)}\right)$ for the $W_{Y_1,Y_2|X}$ BC with privacy leakage constraints induces a PMF $P^{(c_n)}$ on $\mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{X}^n \times \mathcal{Y}_1^n \times \mathcal{Y}_2^n \times \hat{\mathcal{M}}_{01} \times \hat{\mathcal{M}}_{02}$, that is given by

$$\begin{aligned} &P^{(c_n)}\left(m_0, m_1, m_2, \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2, \left(\hat{m}_0^{(1)}, \hat{m}_1\right), \left(\hat{m}_0^{(2)}, \hat{m}_2\right)\right) \\ &= \prod_{j=0,1,2} \frac{1}{\left|\mathcal{M}_j^{(n)}\right|} f^{(n)}(\mathbf{x}|m_0, m_1, m_2) W_{Y_1,Y_2|X}^n(\mathbf{y}_1, \mathbf{y}_2|\mathbf{x}) \\ &\quad \times \mathbb{1}_{\bigcap_{j=1,2}\left\{(\hat{m}_0^{(j)}, m_j) = \phi_j^{(n)}(\mathbf{y}_j)\right\}}. \end{aligned} \quad (19)$$

The induced PMF gives rise to the probability measure $\mathbb{P}_{P^{(c_n)}}$, which we abbreviate by $\mathbb{P}_{c_n}$. Similarly, we use the shorthand I_{c_n} instead of $I_{P^{(c_n)}}$ to denote a mutual information expression taken with respect to $P^{(c_n)}$.

Definition 4 (Average Error Probability): The average error probability for an (n, R_0, R_1, R_2) code c_n is

$$P_{\mathsf{e}}(c_n) \triangleq \mathbb{P}_{c_n}\left(\bigcup_{j=1,2}\left\{\left(\hat{M}_0^{(j)}, \hat{M}_j\right) \neq (M_0, M_j)\right\}\right) \quad (20)$$

where $\left(\hat{M}_0^{(j)}, \hat{M}_j\right) = \phi_j^{(n)}(\mathbf{Y}_j)$, for $j = 1, 2$.

Definition 5 (Information Leakage Rate): The information leakage rate of M_1 to receiver 2 under an (n, R_0, R_1, R_2) code c_n is

$$\ell_1(c_n) \triangleq \frac{1}{n} I_{c_n}(M_1; \mathbf{Y}_2). \quad (21a)$$

Similarly, the information leakage rate of M_2 to receiver 1 under c_n is

$$\ell_2(c_n) \triangleq \frac{1}{n} I_{c_n}(M_2; \mathbf{Y}_1). \quad (21b)$$

Definition 6 (Achievable Rates): Let $(L_1, L_2) \in \mathbb{R}_+^2$. A rate triple $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ is (L_1, L_2)-*achievable* if, for any $\epsilon > 0$, there exists a sufficiently large $n \in \mathbb{N}$ and an (n, R_0, R_1, R_2) code c_n such that

$$P_{\mathsf{e}}(c_n) \leq \epsilon \quad (22a)$$

$$\ell_1(c_n) \leq L_1 + \epsilon \quad (22b)$$

$$\ell_2(c_n) \leq L_2 + \epsilon. \quad (22c)$$

Definition 7 (Leakage-Capacity Region): The (L_1, L_2)-*leakage-capacity region* $\mathcal{C}(L_1, L_2)$ is the closure of the set of the (L_1, L_2)-achievable rates.

Remark 2 (Inactive Leakage Constraints): Setting $L_j = R_j$, for $j = 1, 2$, makes (22b)-(22c) inactive and reduces the BC with privacy leakage constraints to the classic BC with a common message. This is a simple consequence of the non-negativity of entropy, which implies that $I_{c_n}(M_1; \mathbf{Y}_2) \leq nR_1$ and $I_{c_n}(M_2; \mathbf{Y}_1) \leq nR_2$ always hold. To simplify notation we write $L_j \to \infty$, $j = 1, 2$ to refer to leakage threshold values under which (22b)-(22c) are satisfied by default.

B. Leakage-Capacity Results

This section states inner and outer bounds on the (L_1, L_2)-leakage-capacity region $\mathcal{C}(L_1, L_2)$ of a BC. The bounds match for SD-BCs, BCs with a degraded message set and PD-BCs, which characterizes the leakage-capacity regions for these three cases. We start with the inner bound.

In the following, the transition probability $W_{Y_1,Y_2|X}$ describing the BC stays fixed unless stated otherwise. When specifying to particular instances of BCs (see Remark 1), we explicitly mention the corresponding structure of $W_{Y_1,Y_2|X}$.

Theorem 1 (Inner Bound): Let $\mathcal{R}_{\mathsf{I}}(L_1, L_2)$ be the closure of the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ satisfying:

$$R_0 \leq \min\Big\{I(U_0;Y_1), I(U_0;Y_2)\Big\} \tag{23a}$$

$$R_1 \leq I(U_1;Y_1|U_0) - I(U_1;U_2,Y_2|U_0) + L_1 \tag{23b}$$

$$R_0+R_1 \leq I(U_1;Y_1|U_0) + \min\Big\{I(U_0;Y_1), I(U_0;Y_2)\Big\} \tag{23c}$$

$$R_2 \leq I(U_2;Y_2|U_0) - I(U_2;U_1,Y_1|U_0) + L_2 \tag{23d}$$

$$R_0+R_2 \leq I(U_2;Y_2|U_0) + \min\Big\{I(U_0;Y_1), I(U_0;Y_2)\Big\} \tag{23e}$$

$$\sum_{j=0,1,2} R_j \leq I(U_1;Y_1|U_0) + I(U_2;Y_2|U_0) - I(U_1;U_2|U_0) + \min\Big\{I(U_0;Y_1), I(U_0;Y_2)\Big\} \tag{23f}$$

where the union is over all PMFs $Q_{U_0,U_1,U_2,X} \in \mathcal{P}(\mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{X})$, each inducing a joint distribution $Q_{U_0,U_1,U_2,X}W_{Y_1,Y_2|X}$. The following inclusion holds:

$$\mathcal{R}_{\mathsf{I}}(L_1, L_2) \subseteq \mathcal{C}(L_1, L_2). \tag{24}$$

The proof of Theorem 1 is given in Section VII-B and uses a leakage-adaptive Marton-like code construction. Rate-splitting is first used to decompose each private message M_j, $j = 1, 2$, into a public part M_{0j} and a private part M_{jj}. A Marton codebook with an extra layer of bins is then constructed while treating (M_0, M_{10}, M_{20}) as a public message and M_{jj}, for $j = 1, 2$, as private message j. The double-binning of the private messages permits joint encoding (outer layer) and controlling the total rate leakage to the other user (inner layer). In contrast to the classic Marton coding scheme [19] that employes a JTE, we execute joint encoding by means of the likelihood encoder from (14). Doing so doesn't affect the reliability analysis (as the likelihood encoder chooses jointly typical pairs of codewords with high probability), but it is of consequence for analysing the leakage rate.

The leakage analysis takes into account the rate leaked due to the decoding of the public message by both users. Also, additional leakage occurs due to the joint encoding process, which introduces correlation between the private message codewords. We account for the latter by relating the bin sizes in the inner and outer coding layers to the rate of the public parts M_{10} and M_{20}. The leakage analysis relies heavily on the structure of the likelihood encoder that lets us establish several crucial properties of our random coding experiment. The main challenge is showing that the induced marginal distribution describing the choice of the private message codewords is approximately uniform. This follows by virtue of the Uniform Approximation Lemma (Lemma 3).

Remark 3 (Relation to Marton's Region): Gelfand and Pinsker [21, Th. 1] generalized Marton's inner bound [18] to include a common message. An alternative form of Gelfand and Pinsker's inner bound was given in [37, Th. 5] (see also [38]). This region is the best known inner bound on the capacity region of the BC with a common message. $\mathcal{R}_{\mathsf{I}}(\infty, \infty)$ recovers the Gelfand-Pinsker region since (23b) and (23d) are redundant. A full discussion of the special cases of $\mathcal{R}_{\mathsf{I}}(L_1, L_2)$ is given in Section V-D.

The following corollary states a sufficient condition on the leakage thresholds L_1 and L_2 to become inactive in the bounds from (23) when $R_0 = 0$ (i.e., no common message is present). To state the result, let $\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2, Q_{U_0,U_1,U_2,X})$ denote the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying (23) with $R_0 = 0$ when the mutual information terms are calculated with respect to $Q_{U_0,U_1,U_2,X}W_{Y_1,Y_2|X}$. Accordingly,

$$\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2) \triangleq \bigcup_{Q_{U_0,U_1,U_2,X}} \tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2, Q_{U_0,U_1,U_2,X}) \tag{25}$$

is the region obtained by setting $R_0 = 0$ in $\mathcal{R}_{\mathsf{I}}(L_1, L_2)$.

Corollary 2 (Inactive Leakage Constraints): Let $Q_{U_0,U_1,U_2,X} \in \mathcal{P}(\mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{X})$. For $j = 1, 2$ define

$$L_j^\star(Q_{U_0,U_1,U_2,X}) = \min\Big\{I(U_0;Y_1), I(U_0;Y_2)\Big\} + I(U_j;U_{\bar{j}},Y_{\bar{j}}|U_0), \tag{26}$$

where $\bar{j} = j + (-1)^{j+1}$. The following implications hold:

1) If $L_1 \geq L_1^\star(Q_{U_0,U_1,U_2,X})$ then $\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2, Q_{U_0,U_1,U_2,X}) = \tilde{\mathcal{R}}_{\mathsf{I}}(\infty, L_2, Q_{U_0,U_1,U_2,X})$.
2) If $L_2 \geq L_2^\star(Q_{U_0,U_1,U_2,X})$ then $\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2, Q_{U_0,U_1,U_2,X}) = \tilde{\mathcal{R}}_{\mathsf{I}}(L_1, \infty, Q_{U_0,U_1,U_2,X})$.
3) If $L_j \geq L_j^\star(Q_{U_0,U_1,U_2,X})$, for $j = 1, 2$, then $\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2, Q_{U_0,U_1,U_2,X}) = \tilde{\mathcal{R}}_{\mathsf{I}}(\infty, \infty, Q_{U_0,U_1,U_2,X})$.

For the proof of Corollary 2 see Section VII-C. According to the above, if any of the leakage thresholds L_j, $j = 1, 2$ surpasses the critical value from (26), then the corresponding inner bound remains unchanged if L_j is further increased, and is therefore equivalent to the region where $L_j \to \infty$.

Remark 4 (Application of Corollary 2): Corollary 2 specifies a condition for L_1 and/or L_2 being inactive for each input probability. Getting a condition for the inactivity of the thresholds with respect to the entire region $\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2)$ from (25) is a more challenging task. Identifying such a condition involves identifying which input distributions achieve the boundary of $\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2)$. In some communication scenarios this is possible, e.g., for the MIMO Gaussian BC with or without secrecy requirements the boundary achieving distributions are Gaussian vectors [11], [39]–[42]. However, the structure of the optimizing distribution is unknown in general.

The merit of Corollary 2 becomes clear when explicitly calculating $\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2)$. One can then identify the optimizing distribution, e.g., by means of an analytical characterization or via an exhaustive search. In turn, one can calculate the maximum of $L_j^\star(Q_{U_0,U_1,U_2,X})$ over those distributions. Denoting by $L_j^\star$ this maximal value, if $L_j < L_j^\star$ then increasing L_j will further enlarge the region. If, on the other hand, $L_j \geq L_j^\star$, then the region remains unchanged even if L_j grows. This idea is demonstrated in Section VI where we calculate the (L_1, L_2)-leakage-capacity region of the Blackwell BC.

Next, we state an outer bound on $\mathcal{C}(L_1, L_2)$. A proof of Theorem 3 is given in Section VII-D.

Theorem 3 (Outer Bound): Let $\mathcal{R}_{\mathsf{O}}(L_1, L_2)$ be the closure of the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ satisfying:

$$R_0 \le \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{27a}$$
$$R_1 \le I(U;Y_1|W,V) - I(U;Y_2|W,V) + L_1 \tag{27b}$$
$$R_1 \le I(U;Y_1|W) - I(U;Y_2|W) + L_1 \tag{27c}$$
$$R_0 + R_1 \le I(U;Y_1|W) + \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{27d}$$
$$R_2 \le I(V;Y_2|W,U) - I(V;Y_1|W,U) + L_2 \tag{27e}$$
$$R_2 \le I(V;Y_2|W) - I(V;Y_1|W) + L_2 \tag{27f}$$
$$R_0 + R_2 \le I(V;Y_2|W) + \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{27g}$$
$$\sum_{j=0,1,2} R_j \le I(U;Y_1|W,V) + I(V;Y_2|W) + \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{27h}$$
$$\sum_{j=0,1,2} R_j \le I(U;Y_1|W) + I(V;Y_2|W,U) + \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{27i}$$

where the union is over all PMFs $Q_{W,U,V}Q_{X|U,V} \in \mathcal{P}(\mathcal{W}\times\mathcal{U}\times\mathcal{V}\times\mathcal{X})$, each inducing a joint distribution $Q_{W,U,V}Q_{X|U,V}W_{Y_1,Y_2|X}$. $\mathcal{R}_{\mathsf{O}}(L_1, L_2)$ is convex and the following inclusion holds:

$$\mathcal{C}(L_1, L_2) \subseteq \mathcal{R}_{\mathsf{O}}(L_1, L_2). \tag{28}$$

Remark 5 (Relation to UVW-Outer Bound): The best known outer bounds on the capacity region of a BC with a common message are the UVW-outer bound [22, Bound 2] and the New-Jersey outer bound [23] which are equivalent. The region $\mathcal{R}_{\mathsf{O}}(\infty,\infty)$ recovers the UVW-outer bound since (27b)-(27c) and (27e)-(27f) are redundant.

The inner and outer bounds in Theorems 1 and 3 are tight for SD-BCs and give rise to the following theorem.

Theorem 4 (Leakage-Capacity - SD-BC): The (L_1, L_2)-leakage-capacity region $\mathcal{C}_{\mathsf{SD}}(L_1, L_2)$ of a SD-BC $\mathbb{1}_{\{Y_1=y_1(X)\}}W_{Y_2|X}$ is the closure of the union of rate triples $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ satisfying:

$$R_0 \le \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{29a}$$
$$R_1 \le H(Y_1|W,V,Y_2) + L_1 \tag{29b}$$
$$R_0 + R_1 \le H(Y_1|W) + \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{29c}$$
$$R_2 \le I(V;Y_2|W) - I(V;Y_1|W) + L_2 \tag{29d}$$
$$R_0 + R_2 \le I(V;Y_2|W) + \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{29e}$$
$$\sum_{j=0,1,2} \le H(Y_1|W,V) + I(V;Y_2|W) + \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{29f}$$

where the union is over all PMFs $Q_{W,V,X} \in \mathcal{P}(\mathcal{W}\times\mathcal{V}\times\mathcal{X})$, each inducing a joint distribution $Q_{W,V,X}\mathbb{1}_{\{Y_1=y_1(X)\}}W_{Y_2|X}$. Furthermore, $\mathcal{C}_{\mathsf{SD}}(L_1, L_2)$ is convex.

The direct part of Theorem 4 follows from Theorem 1 by taking $U_0 = W$, $U_1 = Y_1$ and $U_2 = V$, while Theorem 3 is used for the converse. See Section VII-E for the details.

Remark 6 (SD-BC Result - Special Cases): All four cases of the SD-BC concerning secrecy (i.e., when neither, either or both messages are secret) are solved and their solutions are retrieved from $\mathcal{C}_{\mathsf{SD}}(L_1, L_2)$ by inserting the appropriate values of L_j, $j = 1, 2$. This property of $\mathcal{C}_{\mathsf{SD}}(L_1, L_2)$ is discussed in Section V-D.

The inner and outer bounds in Theorems 1 and 3 also match when the message set is degraded, i.e., when $M_2 = 0$ and there is only one private message.

Theorem 5 (Leakage-Capacity - Degraded Message Set): The L_1-leakage-capacity region $\mathcal{C}_{\mathsf{DM}}(L_1)$ of a BC with a degraded message set ($M_2 = 0$) and a privacy leakage constraint is the closure of the union of rate pairs $(R_0, R_1) \in \mathbb{R}_+^2$ satisfying:

$$R_0 \le \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{30a}$$
$$R_1 \le I(U;Y_1|W) - I(U;Y_2|W) + L_1 \tag{30b}$$
$$R_0 + R_1 \le I(U;Y_1|W) + \min\Big\{I(W;Y_1), I(W;Y_2)\Big\} \tag{30c}$$

where the union is over all PMFs $Q_{W,U}Q_{X|U} \in \mathcal{P}(\mathcal{W}\times\mathcal{U}\times\mathcal{X})$, each inducing a joint distribution $Q_{W,U}Q_{X|U}W_{Y_1,Y_2|X}$. Furthermore, $\mathcal{C}_{\mathsf{DM}}(L_1)$ is convex.

Proof: The direct part follows by setting $R_2 = 0$, $U_0 = W$, $U_1 = U$ and $U_2 = 0$ in Theorem 1. For the converse we show that $\mathcal{R}_{\mathsf{O}}(L_1, L_2) \subseteq \mathcal{C}_{\mathsf{DM}}(L_1)$. Clearly, (30a), (30b) and (30c) coincide with (27a), (27c) and (27d), respectively. Dropping the rest of the inequalities from (27) completes the proof. ■

Remark 7 (Degraded Message Set Result - Special Cases): The BC with a degraded message set and a privacy leakage constraint captures the BC with confidential messages [3] and the BC with a degraded message set [28]. The former is obtained by taking $L_1 = 0$, while $L_1 \to \infty$ recovers the latter. Setting $L_1 = 0$ or $L_1 \to \infty$ into $\mathcal{C}_{\mathsf{DM}}(L_1)$ recovers the capacity regions of these special cases (see Section V-E for more details).

We next characterize the leakage-capacity region of a PD-BC $W_{Y_1|X}W_{Y_2|Y_1}$ with privacy leakage constraints and without a common message ($M_0 = 0$). Since $X - Y_1 - Y_2$ forms a Markov chain, it is impossible to achieve non-trivial leakage constraints on the message M_2. Accordingly, the leakage-capacity region of the PD-BC (where $X - Y_1 - Y_2$) is defined only through L_1.

Corollary 6 (Leakage-Capacity - PD-BC): The L_1-leakage-capacity region $\mathcal{C}_{\mathsf{PD}}(L_1)$ of a PD-BC $W_{Y_1|X}W_{Y_2|Y_1}$ without a common message is the closure of the union over the same domain as $\mathcal{C}_{\mathsf{DM}}(L_1)$ of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying (30), while recasting R_0 as R_2 and noting that $\min\big\{I(W;Y_1), I(W;Y_2)\big\} = I(W;Y_2)$.

The proof of Corollary 6 is similar to that of Theorem 5 and is omitted.

Remark 8 (Cardinality Bounds): Cardinality bounds for the auxiliary random variables in Theorems 1, 3, 4 and 5 can be derived using the perturbation method [20, Appendix C] or techniques such as in [22] and [43]. The computability of the derived regions is not in the scope of this work.

V. Special Cases

A. The Gelfand-Pinsker Inner Bound

Theorem 1 generalizes the Gelfand-Pinsker region for the BC with a common message [21, Th. 1] to the case with

privacy leakage constraints. In other words, $\mathcal{R}_{\mathsf{I}}(\infty,\infty)$ recovers the result from [21], which is tight for every BC (without secrecy) whose capacity region is known.

B. UVW-Outer Bound

The New-Jersey outer bound was derived in [23] and shown to be at least as good as the previously known bounds. A simpler version of this outer bound was established in [22] and was named the UVW-outer bound. The UVW-outer bound is given by $\mathcal{R}_{\mathsf{O}}(\infty,\infty)$.

C. Liu-Marić-Spasojević-Yates Inner Bound

In [4] an inner bound on the secrecy-capacity region of a BC $W_{Y_1,Y_2|X}$ with two confidential messages (each destined for one of the receivers and kept secret from the other) was characterized as the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \leq I(U_1; Y_1|U_0) - I(U_1; U_2, Y_2|U_0) \tag{31a}$$
$$R_2 \leq I(U_2; Y_2|U_0) - I(U_2; U_1, Y_1|U_0) \tag{31b}$$

where the union is over all PMFs $Q_{U_0,U_1,U_2}Q_{X|U_1,U_2} \in \mathcal{P}(\mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{X})$, each inducing a joint distribution $Q_{U_0,U_1,U_2}Q_{X|U_1,U_2}W_{Y_1,Y_2|X}$. This inner bound is tight for SD-BCs [5] and MIMO Gaussian BCs [11]. Setting $R_0 = 0$ in $\mathcal{R}_{\mathsf{I}}(0,0)$ recovers (31).

D. SD-BCs With and Without Secrecy

The SD-BC $\mathbb{1}_{\{Y_1=y_1(X)\}}W_{Y_2|X}$ without a common message, i.e., when $R_0 = 0$, is solved when both, either or neither private messages are secret (see [5], [6], [27], [21], respectively). Setting $L_j = 0$, for $j = 1, 2$, reduces the SD-BC with privacy leakage constraints to the problem where M_j is secret. Taking $L_j \to \infty$ results in a SD-BC without a leakage constraint on M_j. We use Theorem 4 to obtain the leakage-capacity region of the SD-BC without a common message.

Corollary 7 (Leakage-Capacity - SD-BC Without M_0): The (L_1, L_2)-leakage-capacity region $\mathcal{C}_{\mathsf{SD}}^0(L_1, L_2)$ of a SD-BC $\mathbb{1}_{\{Y_1=y_1(X)\}}W_{Y_2|X}$ without a common message is the closure of the union over the domain stated in Theorem 4 of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \leq H(Y_1|W, V, Y_2) + L_1 \tag{32a}$$
$$R_1 \leq H(Y_1|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \tag{32b}$$
$$R_2 \leq I(V; Y_2|W) - I(V; Y_1|W) + L_2 \tag{32c}$$
$$R_2 \leq I(V; Y_2|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \tag{32d}$$
$$R_1 + R_2 \leq H(Y_1|W, V) + I(V; Y_2|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\}. \tag{32e}$$

1) Neither Message Is Secret: If $L_1, L_2 \to \infty$, the SD-BC with privacy leakage constraints reduces to the classic case without secrecy [21]. We recover the capacity region by choosing $W = 0$ so that (32) becomes

$$R_1 \leq H(Y_1) \tag{33a}$$
$$R_2 \leq I(V; Y_2) \tag{33b}$$
$$R_1 + R_2 \leq H(Y_1|V) + I(V; Y_2) \tag{33c}$$

This agrees with the discussion in Section V-A since Marton's inner bound is tight for SD-BCs.

2) Only M_1 is Secret: The SD-BC where M_1 is a secret is obtained by taking $L_1 = 0$ and $L_2 \to \infty$. The secrecy-capacity region was derived in [27, Corollary 4] and is the closure of the union over the same domain as (33) of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \leq H(Y_1|V, Y_2) \tag{34a}$$
$$R_2 \leq I(V; Y_2). \tag{34b}$$

To see that $\mathcal{C}_{\mathsf{SD}}^0(0,\infty)$ and (34) match, first note that when $L_1 = 0$, (32b) is redundant due to (32a). The sum rate bound (32e) also becomes inactive as it is implied by adding (32a) and (32d). Setting $W = 0$ in $\mathcal{C}_{\mathsf{SD}}^0(0,\infty)$ now recovers (34).

Remark 9 (Relation to Optimal Coding Scheme): The optimal code for the SD-BC with a secret message M_1 employs no public message and relies on double-binning the codebook of M_1, while M_2 is transmitted at maximal rate and no binning of its codebook is performed. The optimality of $W = 0$ in $\mathcal{C}_{\mathsf{SD}}^0(0,\infty)$ corresponds to the absence of the public messages. Furthermore, referring to the bounds in Section VII-B, inserting $L_1 = 0$ and $L_2 \to \infty$ into our code construction results in (68a) and (87b) becoming inactive since (86b) is the dominant constraint. Consequently, the redundancy used for correlating the transmission and ensuring security (i.e., the double-binning) is present only in the M_1 codebook.

3) Only M_2 is Secret: The SD-BC where M_2 is secret is obtained by taking $L_1 \to \infty$ and $L_2 = 0$. The secrecy-capacity region is the closure of the union of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \leq H(Y_1) \tag{35a}$$
$$R_1 \leq H(Y_1|W) + I(W; Y_2) \tag{35b}$$
$$R_2 \leq I(V; Y_2|W) - I(V; Y_1|W) \tag{35c}$$

where the union is over all PMFs $Q_{W,V,X} \in \mathcal{P}(\mathcal{W} \times \mathcal{V} \times \mathcal{X})$, each inducing a joint distribution $Q_{W,V,X}\mathbb{1}_{\{Y_1=y_1(X)\}}W_{Y_2|X}$ [6, Th. 1]. Using Corollary 7, the bounds (32) become

$$R_1 \leq H(Y_1|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \tag{36a}$$
$$R_2 \leq I(V; Y_2|W) - I(V; Y_1|W) \tag{36b}$$
$$R_1 + R_2 \leq H(Y_1|W, V) + I(V; Y_2|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\}. \tag{36c}$$

and (36c) is redundant by adding (36a) and (36b). The regions from (35) and (36) thus coincide.

The effect of $L_1 \to \infty$ and $L_2 = 0$ on the bins in our coding scheme (Section VII-B) is analogous to the one described in Remark 9. In contrast to Section V-D2, however, here the achievability of (36) requires a common message. Since $L_2 = 0$, (60c) implies that the public message is a portion of M_1 *only*. Keeping in mind that the public message is decoded by both receivers, unless $R_{20} = 0$ (i.e., unless the public message contains no information about M_2) the secrecy constraint will be violated.

4) Both Messages are Secret: Taking $L_1 = L_2 = 0$ recovers the SD-BC where both messages are secret. The secrecy-capacity region for this case was found in [5, Th. 1] and is the closure of the union of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \le H(Y_1|W, V, Y_2) \tag{37a}$$
$$R_2 \le I(V; Y_2|W) - I(V; Y_1|W) \tag{37b}$$

where the union is over all PMFs $Q_{W,V}Q_{X|V} \in \mathcal{P}(\mathcal{W} \times \mathcal{V} \times \mathcal{X})$, each inducing a joint distribution $Q_{W,V}Q_{X|V}\mathbb{1}_{\{Y_1=y_1(X)\}}W_{Y_2|X}$. The region (37) coincides with $\mathcal{C}_{\mathsf{SD}}^0(0,0)$. Restricting the union in $\mathcal{C}_{\mathsf{SD}}^0(0,0)$ to encompass only PMFs that satisfy the Markov relation $W - V - X$ does not shrink the region. This is since in the proof of Theorem 3 we define $V_q \triangleq (M_2, W_q)$, and therefore, $X_q - V_q - W_q$ forms a Markov chain for every $q \in [1:n]$.

Remark 10 (Relation to Optimal Coding Scheme): The coding scheme that achieves (37) uses double-binning for the codebooks of both private messages. To ensure confidentiality, the rate bounds of each message include the penalty term $I(U_1; U_2|U_0)$. Note that without the confidentiality constraints, Marton's coding scheme [18] requires only that the sum-rate has that penalty term. This is evident from our scheme by setting $L_1 = L_2 = 0$ in (60c), (86b) and (87b), which makes (68a) redundant.

E. BCs With One Private Message

Consider the BC with leakage constraints in which $M_2 = 0$; its leakage-capacity region $\mathcal{C}_{\mathsf{DM}}(L_1)$ is stated in Theorem 5. We show that $\mathcal{C}_{\mathsf{DM}}(L_1)$ recovers the secrecy-capacity region of the BC with confidential messages [3] and the capacity region of the BC with a degraded message set (without secrecy) [28].

1) BCs With Confidential Messages: The secrecy-capacity region of the BC with confidential messages was derived in [3] and is the union over the same domain as in Theorem 5 of rate pairs $(R_0, R_1) \in \mathbb{R}_+^2$ satisfying:

$$R_0 \le \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \tag{38a}$$
$$R_1 \le I(U; Y_1|W) - I(U; Y_2|W). \tag{38b}$$

Inserting $L_1 = 0$ into the result of Theorem 5 produces (38).

Our code construction (Section VII-B) with $L_1 = 0$ and $U_2 = 0$ reduces to a superposition code for which the outer codebook (that is associated with the confidential message) is binned. This is a secrecy-capacity achieving coding scheme for the BC with confidential messages.

Remark 11 (Wiretap Channel): The BC with confidential messages captures the WTC by setting $M_0 = 0$. Thus, the WTC is also a special case of the BC with privacy leakage constraints.

2) BCs With a Degraded Message Set: If $L_1 \to \infty$, we get the BC with a degraded message set [28]. Inserting $L_1 \to \infty$ into $\mathcal{C}_{\mathsf{DM}}(L_1)$ and setting $U = X$ we recover the union of rate pairs $(R_0, R_1) \in \mathbb{R}_+$ satisfying:

$$R_0 \le \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \tag{39a}$$
$$R_0 + R_1 \le I(X; Y_1|W) + I(W; Y_2) \tag{39b}$$
$$R_0 + R_1 \le I(X; Y_1) \tag{39c}$$

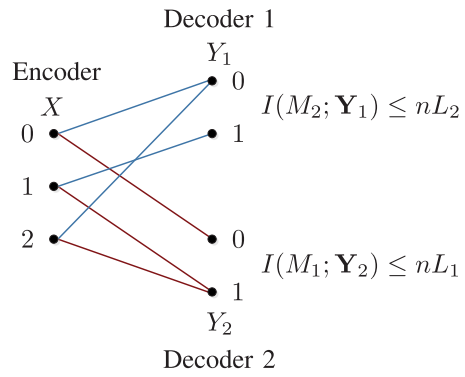

Fig. 2. Blackwell BC with privacy leakage constraints.

where the union is over all PMFs $Q_{W,X} \in \mathcal{P}(\mathcal{V} \times \mathcal{X})$, each induces a joint distribution $Q_{W,X}W_{Y_1,Y_2|X}$.

The region from (39) matches [37, Th. 7] which establishes the union over all PMFs $Q_{T,U,X} \in \mathcal{P}(\mathcal{T} \times \mathcal{U} \times \mathcal{X})$ of rate pairs $(R_0, R_1) \in \mathbb{R}_+$ with

$$R_0 \le \min\Big\{I(T; Y_1), I(T; Y_2)\Big\} \tag{40a}$$
$$R_0 + R_1 \le I(X; Y_1|T, U) + I(T, U; Y_2) \tag{40b}$$
$$R_0 + R_1 \le I(X; Y_1) \tag{40c}$$

as an outer bound on the capacity region of interest. The RHS of (40a) can be bounded as

$$\min\Big\{I(T; Y_1), I(T; Y_2)\Big\} \le \min\Big\{I(T, U; Y_1), I(T, U; Y_2)\Big\} \tag{41}$$

and relabeling $W = (T, U)$ matches (39).

VI. Example

Suppose the channel from the transmitter to Receivers 1 and 2 is the BW-BC without a common message as illustrated in Fig. 2 [29], [30]. Using Corollary 7, the (L_1, L_2)-leakage-capacity region of a deterministic BC (DBC) is the following.

Corollary 8 (Leakage-Capacity - Deterministic BC): The (L_1, L_2)-leakage-capacity region $\mathcal{C}_{\mathsf{D}}(L_1, L_2)$ of the DBC $\mathbb{1}_{\{Y_1=y_1(X)\}\cap\{Y_2=y_2(X)\}}$ without a common message is the union of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \le \min\big\{H(Y_1)\,,\ H(Y_1|Y_2) + L_1\big\} \tag{42a}$$
$$R_2 \le \min\big\{H(Y_2)\,,\ H(Y_2|Y_1) + L_2\big\} \tag{42b}$$
$$R_1 + R_2 \le H(Y_1, Y_2) \tag{42c}$$

where the union is over all input PMFs $Q_X \in \mathcal{P}(\mathcal{X})$.

The proof of Corollary 8 is relegated to Appendix A. For the BW-BC, we parametrize the input PMF $Q_X \in \mathcal{P}\big(\{0, 1, 2\}\big)$ in Corollary 8 as

$$Q_X(0) = \alpha, \quad Q_X(1) = \beta, \quad Q_X(2) = 1 - \alpha - \beta, \tag{43}$$

where $\alpha, \beta \in \mathbb{R}_+$ and $\alpha + \beta \le 1$. Using (43), the (L_1, L_2)-leakage-capacity region $\mathcal{C}_{\mathsf{BW}}(L_1, L_2)$ of the BW-BC is

descried as the union of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying:

$$R_1 \leq \min\left\{H_b(\beta), (1-\alpha)H_b\left(\frac{\beta}{1-\alpha}\right) + L_1\right\} \quad (44a)$$

$$R_2 \leq \min\left\{H_b(\alpha), (1-\beta)H_b\left(\frac{\alpha}{1-\beta}\right) + L_2\right\} \quad (44b)$$

$$R_1 + R_2 \leq H_b(\alpha) + (1-\alpha)H_b\left(\frac{\beta}{1-\alpha}\right) \quad (44c)$$

where the union is over all $\alpha, \beta \in \mathbb{R}_+$ with $\alpha + \beta \leq 1$.

Fig. 3 illustrates $\mathcal{C}_{\mathsf{BW}}(L_1, L_2)$ for three cases. In Fig. 3(a) $L_2 \to \infty$ while $L_1 \in \{0, 0.05, 0.1, 0.4\}$. The blue (inner) line corresponds to $L_1 = 0$ and is the secrecy-capacity region of a BW-BC where M_1 is secret [27, Fig. 5]. The red (outer) line corresponds to $L_1 = 0.4$ (which is large enough to be thought of as $L_1 \to \infty$) and depicts the capacity region of the classic BW-BC. As L_1 grows, the inner (blue) region converges to coincide with the outer (red) region. Fig. 3(b) considers the opposite case, i.e., where $L_1 \to \infty$ and $L_2 \in \{0, 0.05, 0.1, 0.4\}$, and is analogous to Fig. 3(a). In Fig. 3(c) we choose $L_1 = L_2 = L$, where $L \in \{0, 0.05, 0.1, 0.4\}$, and we demonstrate the impact of two leakage constraints on the region. When $L = 0$, one obtains the secrecy-capacity region of the BW-BC when both messages are confidential [5]. In each case, the capacity region grows with L and saturates at the red (outer) region, for which neither message is secret. Focusing on the symmetric case in Fig. 3(c), we note that the saturation of the region at $L = 0.4$ is implied by Corollary 2. For the Blackwell BC with $L_1 = L_2 = L$, and some α, $\beta \in \mathbb{R}_+$ with $\alpha + \beta \leq 1$, we denote by $L^\star(\alpha, \beta)$ the threshold from (26), which reduces to

$$L^\star(\alpha, \beta) = I(Y_1; Y_2) = H_b(\beta) - (1-\alpha)H_b\left(\frac{\beta}{1-\alpha}\right). \quad (45)$$

As explained in Remark 4, for each leakage value L, Corollary 2 (along with some numerical calculations) can be used to tell whether a further increase of L will induce a larger region or not. Accordingly, for each $L \in \{0, 0.05, 0.1, 0.4\}$, we have calculated the maximum of $L^\star(\alpha, \beta)$ over the distributions that achieve the boundary points of the capacity region $\mathcal{C}_{\mathsf{BW}}(L, L)$. Denoting the value of the maximal $L^\star$ that corresponds to the allowed leakage $L \in \{0, 0.05, 0.1, 0.4\}$ by $L^\star(L)$, we have

$$\begin{aligned} L^\star(0) &= L^\star(0.05) = 0.15897 \\ L^\star(0.1) &= 0.20101 \\ L^\star(0.4) &= 0.38317. \end{aligned} \quad (46)$$

Observing that $L^\star(0.4) \leq L$, Corollary 2 and Remark 4 imply that increasing L beyond 0.4 will not change the leakage-capacity region. Evidently, $\mathcal{C}_{\mathsf{BW}}(L, L)$ saturates at $L = 0.4$. For $L \in \{0, 0.05, 0.1\}$, however, $L^\star(L) > L$ and consequently $\mathcal{C}_{\mathsf{BW}}(L', L') \subsetneq \mathcal{C}_{\mathsf{BW}}(L, L)$, for $L, L' \in \{0, 0.05, 0.1\}$ with $L' < L$.

The variation of the sum of rates $R_1 + R_2$ as a function of L is shown by the blue curve in Fig. 4; the red dashed vertical lines correspond to the values of L considered in Fig. 3. Note that for $0 \leq L \leq 0.09818$, (44c) is inactive, and therefore, $R_1 + R_2$ is bounded by the summation of (44a) and (44b).

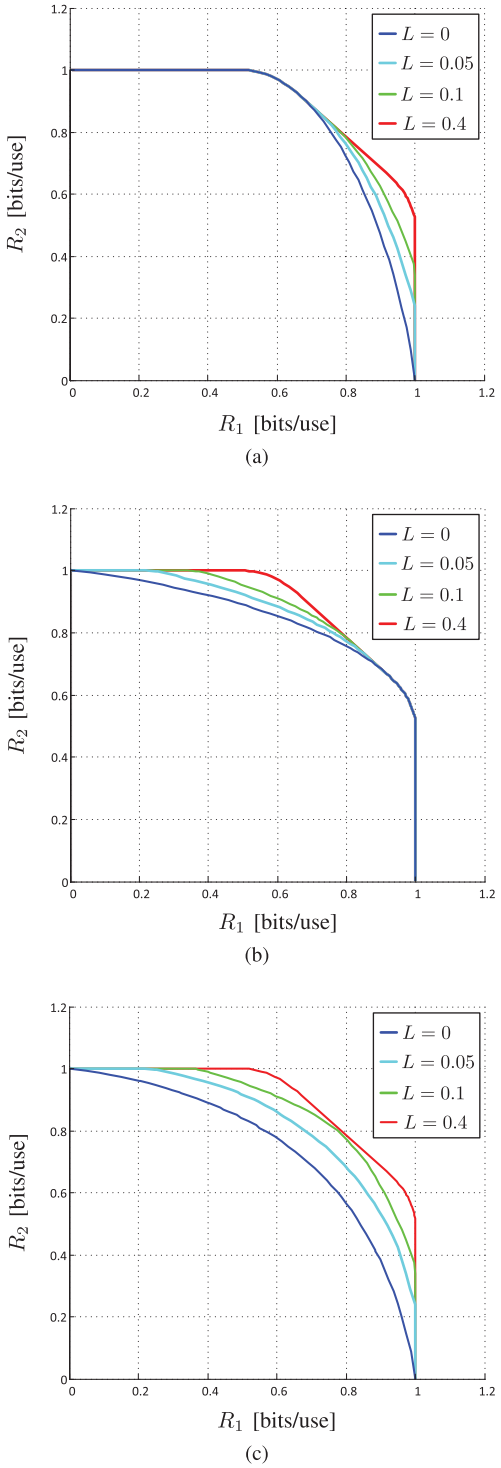

Fig. 3. (L_1, L_2)-leakage-capacity region of the BW-BC for three cases: (a) $L_1 = L$ and $L_2 \to \infty$; (b) $L_1 \to \infty$ and $L_2 = L$; (c) $L_1 = L_2 = L$.

Thus, for $0 \leq L \leq 0.09818$, the sum $R_1 + R_2$ increases linearly with L. For $L > 0.09818$, the bound in (44c) is no longer redundant, and because it is independent of L, the sum rate saturates.

The regions in Fig. 3 are a union of rectangles or pentagons, each corresponds to a different input PMF $Q_X \in \mathcal{P}(\{0, 1, 2\})$. In Fig. 5 we illustrate a typical structure of these rectangles and

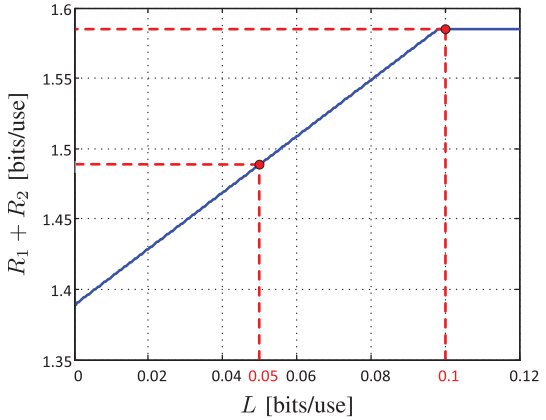

Fig. 4. The sum-rate capacity versus the allowed leakage for $L_1 = L_2 = L$.

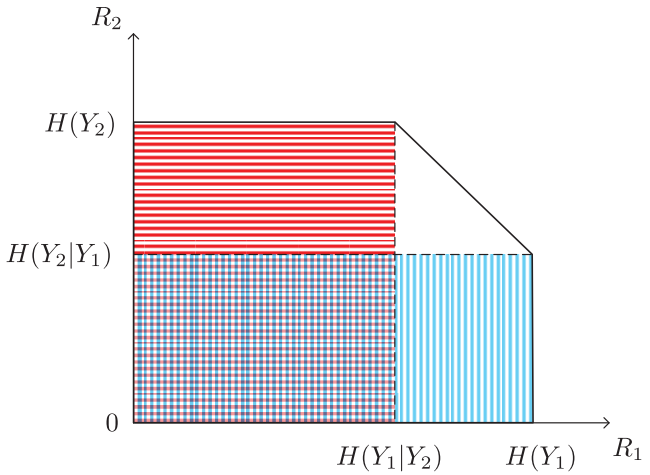

Fig. 5. The pentagons/rectangles whose union produces the capacity region of a BW-BC for different secrecy cases: The outer pentagon corresponds to the case without secrecy; the red and blue rectangles correspond to $L_1 = 0$ and $L_2 = 0$, respectively; the inner rectangle corresponds to $L_1 = L_2 = 0$.

pentagons for a fixed Q_X at the extreme values of L_1 and L_2. When both L_1 and L_2 are sufficiently large, the leakage constraints degenerate and the classic BW-BC is obtained. Its capacity region (the red (outer) line in, e.g., Fig. 3(c)) is a union of the pentagons depicted in Fig. 5. The secrecy-capacity region for $L_1 = 0$ and $L_2 \to \infty$ (depicted by the blue line in Fig. 3(a)) is a union of the red rectangles in Fig. 5. Similarly, when $L_2 = 0$ and $L_1 \to \infty$ the secrecy-capacity region is a union of the blue rectangles in Fig. 5. Finally, if $L_1 = L_2 = 0$ and both messages are secret, the secrecy-capacity region of the BW-BC is the union of the dark rectangles in Fig. 5, i.e., the intersection of the blue and the red regions. Fig. 5 highlights that as L_1 and/or L_2 decrease, the underlying pentagons/rectangles (the union of which produces the admissible rate region) shrink, which results in a smaller region.

VII. Proofs

A. Proof of Lemma 3

We derive sufficient conditions for

$$\mathbb{E}_{\mathsf{B}_n} \mathsf{F}\left(P_{I|\mathsf{B}_n}, p^{(U)}_{\mathcal{I}_n}\right) \xrightarrow[n\to\infty]{} 1 \tag{47}$$

which implies Lemma 3 using (10).

First, for each $\mathcal{B}_n \in \mathfrak{B}_n$ the fidelity between the induced and the desired (uniform) distribution is

$$\mathsf{F}\left(P_{I|\mathsf{B}_n=\mathcal{B}_n}, p^{(U)}_{\mathcal{I}_n}\right) = \sum_{\ell}\left(\frac{1}{|\mathcal{I}_n||\mathcal{K}_n|}\sum_{(j,k)}\frac{2^{i\left(\mathbf{u}_\ell(\mathbf{w});\mathbf{v}_{j,k}(\mathbf{w})\middle|\mathbf{w}\right)}}{\sum\limits_{(\bar\ell,\bar j)} 2^{i\left(\mathbf{u}_{\bar\ell}(\mathbf{w});\mathbf{v}_{\bar j,k}(\mathbf{w})\middle|\mathbf{w}\right)}}\right)^{\frac{1}{2}} \tag{48}$$

where, as in (17), the information density is taken with respect to $Q^n_{U,V}$. Now, by the Cauchy-Schwarz inequality we have the following bound for $\{a_{j,k}\}, \{b_{j,k}\} \subset \mathbb{R}_+$:

$$\sum_{j,k}\sqrt{a_{j,k}b_{j,k}} \le \left(\sum_{j,k} a_{j,k}\right)^{\frac{1}{2}}\left(\sum_{j,k} b_{j,k}\right)^{\frac{1}{2}}. \tag{49}$$

Using this bound on each of the summands from the right-hand side (RHS) of (48) with

$$a_{j,k} = \frac{1}{|\mathcal{I}_n||\mathcal{K}_n|}\frac{2^{i\left(\mathbf{u}_\ell(\mathbf{w});\mathbf{v}_{j,k}(\mathbf{w})\middle|\mathbf{w}\right)}}{\sum\limits_{(\bar\ell,\bar j)} 2^{i\left(\mathbf{u}_{\bar\ell}(\mathbf{w});\mathbf{v}_{\bar j,k}(\mathbf{w})\middle|\mathbf{w}\right)}} \tag{50a}$$

and

$$b_{j,k} = \frac{2^{i\left(\mathbf{u}_\ell(\mathbf{w});\mathbf{v}_{j,k}(\mathbf{w})\middle|\mathbf{w}\right)}}{\sum\limits_{(\bar\ell,\bar k)} 2^{i\left(\mathbf{u}_{\bar\ell}(\mathbf{w});\mathbf{v}_{\bar j,\bar k}(\mathbf{w})\middle|\mathbf{w}\right)}} \tag{50b}$$

we obtain

$$F\left(P_{I|\mathsf{B}_n=\mathcal{B}_n}, p^{(U)}_{\mathcal{I}_n}\right) \ge \sum_{(\ell,j,k)}\frac{2^{i\left(\mathbf{u}_\ell(\mathbf{w});\mathbf{v}_{j,k}(\mathbf{w})\middle|\mathbf{w}\right)}}{\left(|\mathcal{I}_n||\mathcal{K}_n|\sum\limits_{(\bar\ell,\bar j)} 2^{i\left(\mathbf{u}_{\bar\ell}(\mathbf{w});\mathbf{v}_{\bar j,k}(\mathbf{w})\middle|\mathbf{w}\right)}\right)^{\frac{1}{2}}} \times \frac{1}{\left(\sum\limits_{(\bar j,\bar k)} 2^{i\left(\mathbf{u}_{\bar\ell}(\mathbf{w});\mathbf{v}_{\bar j,\bar k}(\mathbf{w})\middle|\mathbf{w}\right)}\right)^{\frac{1}{2}}}. \tag{51}$$

For any $\mathbf{w} \in \mathcal{W}^n$ with $Q^n_W(\mathbf{w}) > 0$, we evaluate the conditional expectation of the fidelity given $\mathbf{W} = \mathbf{w}$ as given in (52) at the top of the next page. First, note that with respect to the notation from Section III, we have

$$\mathbb{E}_{\mathsf{B}_n|\mathbf{W}=\mathbf{w}}\mathsf{F}\left(P_{I|\mathsf{B}_n}, p^{(U)}_{\mathcal{I}_n}\right) = \mathbb{E}_{\mathsf{B}_n(\mathbf{w})}\mathsf{F}\left(P_{I|\mathsf{B}_n}, p^{(U)}_{\mathcal{I}_n}\right). \tag{53}$$

Now consider the following justifications for the steps of (52):
(a) uses (51) and the symmetry of the random codebook;
(b) is the law of total expectation;
(c) uses Jensen's inequality for the two-valued convex function $f : (x, y) \mapsto (xy)^{-\frac{1}{2}}$ and the relation

$$\mathbb{E}_{\mathsf{B}_n(\mathbf{w})|\mathbf{U}_1(\mathbf{w}),\mathbf{V}_{1,1}(\mathbf{w})} 2^{i\left(\mathbf{U}_{\bar\ell}(\mathbf{w});\mathbf{V}_{\bar j,\bar k}(\mathbf{w})\middle|\mathbf{w}\right)} = 1 \tag{54}$$

which holds for any $(\bar\ell, \bar j, \bar k) \neq (1, 1, 1)$ (see [32], [44] for a similar derivation);

$$
\begin{aligned}
&\mathbb{E}_{\mathsf{B}_n(\mathbf{w})}\mathsf{F}\left(P_{I|\mathsf{B}_n}, p_{\mathcal{I}_n}^{(U)}\right)\\
&\overset{(a)}{\geq} |\mathcal{I}_n|^{\frac{1}{2}}|\mathcal{J}_n||\mathcal{K}_n|^{\frac{1}{2}}\mathbb{E}_{\mathsf{B}_n(\mathbf{w})}\left[2^{i\left(\mathbf{U}_1(\mathbf{w});\mathbf{V}_{1,1}(\mathbf{w})\middle|\mathbf{w}\right)}\left(\sum_{(\bar{\ell},\bar{j})}2^{i\left(\mathbf{U}_{\bar{\ell}}(\mathbf{w});\mathbf{V}_{\bar{j},1}\middle|\mathbf{w}\right)}\right)^{-\frac{1}{2}}\left(\sum_{(\tilde{j},\tilde{k})}2^{i\left(\mathbf{U}_1(\mathbf{w});\mathbf{V}_{\tilde{j},\tilde{k}}(\mathbf{w})\right)}\right)^{-\frac{1}{2}}\right]\\
&\overset{(b)}{=} |\mathcal{I}_n|^{\frac{1}{2}}|\mathcal{J}_n||\mathcal{K}_n|^{\frac{1}{2}}\mathbb{E}_{\mathbf{U}_1(\mathbf{w}),\mathbf{V}_{1,1}(\mathbf{w})}\Bigg[2^{i\left(\mathbf{U}_1(\mathbf{w}),\mathbf{V}_{1,1}(\mathbf{w})\middle|\mathbf{w}\right)}\\
&\qquad\times\mathbb{E}_{\mathsf{B}_n(\mathbf{w})|\mathbf{U}_1(\mathbf{w}),\mathbf{V}_{1,1}(\mathbf{w})}\left\{\left(\sum_{(\bar{\ell},\bar{j})}2^{i\left(\mathbf{U}_{\bar{\ell}}(\mathbf{w});\mathbf{V}_{\bar{j},1}(\mathbf{w})\middle|\mathbf{w}\right)}\right)^{-\frac{1}{2}}\left(\sum_{(\tilde{j},\tilde{k})}2^{i\left(\mathbf{U}_1(\mathbf{w});\mathbf{V}_{\tilde{j},\tilde{k}}(\mathbf{w})\middle|\mathbf{w}\right)}\right)^{-\frac{1}{2}}\right\}\Bigg]\\
&\overset{(c)}{\geq} |\mathcal{I}_n|^{\frac{1}{2}}|\mathcal{J}_n||\mathcal{K}_n|^{\frac{1}{2}}\mathbb{E}_{\mathbf{U}_1(\mathbf{w}),\mathbf{V}_{1,1}(\mathbf{w})}\Bigg[2^{i\left(\mathbf{U}_1(\mathbf{w});\mathbf{V}_{1,1}(\mathbf{w})\middle|\mathbf{w}\right)}\left(2^{i\left(\mathbf{U}_1(\mathbf{w});\mathbf{V}_{1,1}(\mathbf{w})\middle|\mathbf{w}\right)}+|\mathcal{I}_n||\mathcal{J}_n|-1\right)^{-\frac{1}{2}}\\
&\qquad\times\left(2^{i\left(\mathbf{U}_1(\mathbf{w});\mathbf{V}_{1,1}(\mathbf{w})\middle|\mathbf{w}\right)}+|\mathcal{J}_n||\mathcal{K}_n|-1\right)^{-\frac{1}{2}}\Bigg]\\
&\overset{(d)}{>} \mathbb{E}_{Q_{U|W=\mathbf{w}}^n Q_{V|W=\mathbf{w}}^n}\left[2^{i(\mathbf{U};\mathbf{V}|\mathbf{w})}\left(1+\left(|\mathcal{I}_n||\mathcal{J}_n|\right)^{-1}2^{i(\mathbf{U};\mathbf{V}|\mathbf{w})}\right)^{-\frac{1}{2}}\left(1+\left(|\mathcal{J}_n||\mathcal{K}_n|\right)^{-1}2^{i(\mathbf{U};\mathbf{V}|\mathbf{w})}\right)^{-\frac{1}{2}}\right]\\
&\overset{(e)}{=} \mathbb{E}_{Q_{U,V|W=\mathbf{w}}^n}\left[\left(1+\left(|\mathcal{I}_n||\mathcal{J}_n|\right)^{-1}2^{i(\mathbf{U};\mathbf{V}|\mathbf{w})}\right)^{-\frac{1}{2}}\left(1+\left(|\mathcal{J}_n||\mathcal{K}_n|\right)^{-1}2^{i(\mathbf{U};\mathbf{V}|\mathbf{w})}\right)^{-\frac{1}{2}}\right] \qquad (52)
\end{aligned}
$$

(d) is by increasing each term in the parenthesis by 1;
(e) is because $\left(\mathbf{U}_1(\mathbf{w}), \mathbf{V}_{1,1}(\mathbf{w})\right) \sim Q_{U|W=\mathbf{w}}^n Q_{V|W=\mathbf{w}}^n$.

Taking an expectation over $\mathbf{W}$ of both sides of (52), while making use of the law of total expectation and of the monotonicity of expectation, gives

$$
\begin{aligned}
&\mathbb{E}_{\mathsf{B}_n}\mathsf{F}\left(P_{I|\mathsf{B}_n}, p_{\mathcal{I}_n}^{(U)}\right)\\
&= \mathbb{E}_{\mathbf{W}}\mathbb{E}_{\mathsf{B}_n|\mathbf{W}}\mathsf{F}\left(P_{I|\mathsf{B}_n}, p_{\mathcal{I}_n}^{(U)}\right)\\
&\geq \mathbb{E}_{Q_{W,U,V}^n}\Bigg[\left(1+\left(|\mathcal{I}_n||\mathcal{J}_n|\right)^{-1}2^{i(\mathbf{U};\mathbf{V}|\mathbf{W})}\right)^{-\frac{1}{2}}\\
&\qquad\times\left(1+\left(|\mathcal{J}_n||\mathcal{K}_n|\right)^{-1}2^{i(\mathbf{U};\mathbf{V}|\mathbf{W})}\right)^{-\frac{1}{2}}\Bigg]. \qquad (55)
\end{aligned}
$$

Finally, note that

$$
\frac{1}{n}\mathbb{E}_{Q_{W,U,V}^n}2^{i_{Q_{W,U,V}}(\mathbf{U};\mathbf{V}|\mathbf{W})} = I_{Q_{W,U,V}}(U;V|W). \qquad (56)
$$

Therefore, by the weak law of large numbers for any $\zeta > 0$ there exists a sequence $\{\delta_n\}_{n\in\mathbb{N}}$ with $\lim_{n\to\infty}\delta_n = 0$, such that

$$
\mathbb{P}_{Q_{W,U,V}^n}\left(\left|\frac{1}{n}i_{Q_{W,U,V}^n}(\mathbf{U};\mathbf{V}|\mathbf{W}) - I_{Q_{W,U,V}}(U;V|W)\right| > \zeta\right) \leq \delta_n \qquad (57)
$$

for all $n \in \mathbb{N}$. Combining (55) and (57) we see that as long as

$$
S_2 + \min\left\{S_1, T\right\} > I_{Q_{W,U,V}}(U;V|W) + \zeta \qquad (58)
$$

then

$$
\mathbb{E}_{\mathsf{B}_n}\mathsf{F}\left(P_{I|\mathsf{B}_n}, p_{\mathcal{I}_n}^{(U)}\right) \nearrow 1 \qquad (59)
$$

as $n \to \infty$. The relation (10) now establishes the result of Lemma 3.

B. Proof of Theorem 1

Fix $n \in \mathbb{N}$, $(L_1, L_2) \in \mathbb{R}_+^2$, $\epsilon, \delta > 0$, a PMF $Q_{U_0,U_1,U_2,X} \in \mathcal{P}(\mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{X})$ and denote $Q_{U_0,U_1,U_2,X,Y_1,Y_2} \triangleq Q_{U_0,U_1,U_2,X}W_{Y_1,Y_2|X}$. In the following we omit the blocklength n from our notations of the sets of indices, e.g., we write $\mathcal{M}_0$ instead of $\mathcal{M}_0^{(n)}$. Furthermore, we assume that quantities of the form 2^{nR}, where $n \in \mathbb{N}$ and $R \in \mathbb{R}_+$, are integers.[5]

1) Message Splitting: Split each message $m_j \in \mathcal{M}_j$, $j = 1, 2$, into a pair of messages denoted by (m_{j0}, m_{jj}). The triple $m_p \triangleq (m_0, m_{10}, m_{20})$ is referred to as a *public message* while m_{jj}, $j = 1, 2$, serves as *private message* j. The rates associated with m_{j0} and m_{jj}, $j = 1, 2$, are denoted by R_{j0} and R_{jj}, while the corresponding alphabets are $\mathcal{M}_{j0}$ and $\mathcal{M}_{jj}$, respectively. The partial rates R_{j0} and R_{jj}, $j = 1, 2$, satisfy

$$
\begin{aligned}
R_j &= R_{j0} + R_{jj} && \text{(60a)}\\
0 &\leq R_{j0} \leq R_j && \text{(60b)}\\
R_{j0} &\leq L_j. && \text{(60c)}
\end{aligned}
$$

Let M_{j0} and M_{jj} be independent random variables uniformly distributed over $\mathcal{M}_{j0}$ and $\mathcal{M}_{jj}$, respectively. We use the notations $M_p \triangleq (M_0, M_{10}, M_{20})$, $\mathcal{M}_p \triangleq \mathcal{M}_0 \times \mathcal{M}_{10} \times \mathcal{M}_{20}$ and $R_p \triangleq R_0 + R_{10} + R_{20}$. Note that M_p is uniformly distributed over $\mathcal{M}_p$ and that $|\mathcal{M}_p| = 2^{nR_p}$. Moreover, let (W_1, W_2) be a pair of independent random variables, where W_j, $j = 1, 2$, is uniformly distributed over $\mathcal{W}_j = \left[1 : 2^{n\tilde{R}_j}\right]$ and independent of (M_0, M_1, M_2) (which implies their independence of (M_p, M_{11}, M_{22}) as well).

[5]Otherwise simple modifications of some of subsequent expressions using floor and ceiling operations are required.

$$\mu(\mathcal{C}_n) = \prod_{m_p \in \mathcal{M}_p} Q^n_{U_0}\big(\mathbf{u}_0(m_p)\big) \prod_{j=1,2} \prod_{\substack{\left(m_p^{(j)}, m_{jj}, w_j, i_j\right) \\ \in \mathcal{M}_p \times \mathcal{M}_{jj} \times \mathcal{W}_j \times \mathcal{I}_j}} Q^n_{U_j|U_0}\Big(\mathbf{u}_j\big(m_p^{(j)}, m_{jj}, w_j, i_j\big)\Big|\mathbf{u}_0\big(m_p^{(j)}\big)\Big) \tag{61}$$

$$\begin{aligned} &P^{(\mathcal{C}_n)}\Big(m_p, m_{11}, m_{22}, w_1, w_2, i_1, i_2, \mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2, \big(\hat{m}_0, \hat{m}_1\big), \big(\hat{m}_0, \hat{m}_2\big)\Big) \\ &\quad = 2^{-n(R_p+R_{11}+R_{11}+\tilde{R}_1+\tilde{R}_2)} P^{(\mathcal{C}_n)}_{\mathsf{LE}}(i_1, i_2|m_p, m_{11}, m_{22}, w_1, w_2) \mathbb{1}_{\left\{\mathbf{u}_0=\mathbf{u}_0(m_p)\right\} \cap \bigcap_{j=1,2} \left\{\mathbf{u}_j=\mathbf{u}_j(m_p, m_{jj}, w_j, i_j)\right\}} \\ &\qquad \times Q^n_{X|U_0,U_1,U_2}(\mathbf{x}|\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2) W^n_{Y_1,Y_2|X}(\mathbf{y}_1, \mathbf{y}_2|\mathbf{x}) \mathbb{1}_{\bigcap_{j=1,2} \left\{\left(\hat{m}_0, \hat{m}_j\right)=\phi_j^{(\mathcal{C}_n)}(\mathbf{y}_j)\right\}} \end{aligned} \tag{65}$$

$$\begin{aligned} &P\Big(\mathcal{C}_n, m_p, m_{11}, m_{22}, w_1, w_2, i_1, i_2, \mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2, \big(\hat{m}_0, \hat{m}_1\big), \big(\hat{m}_0, \hat{m}_2\big)\Big) \\ &\quad = \mu(\mathcal{C}_n) P^{(\mathcal{C}_n)}\Big(m_p, m_{11}, m_{22}, w_1, w_2, s_1, s_2, \mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2, \big(\hat{m}_0, \hat{m}_1\big), \big(\hat{m}_0, \hat{m}_2\big)\Big) \end{aligned} \tag{66}$$

2) Codebook $\mathcal{C}_n$: Let $\mathsf{C}_0^{(n)} \triangleq \left\{\mathbf{U}_0(m_p)\right\}_{m_p \in \mathcal{M}_p}$ be a random public message codebook that comprises 2^{nR_p} i.i.d. random vectors $\mathbf{U}_0(m_p)$, each distributed according to $Q^n_{U_0}$. A realization of $\mathsf{C}_0^{(n)}$ is denoted by $\mathcal{C}_0^{(n)} \triangleq \left\{\mathbf{u}_0(m_p)\right\}_{m_p \in \mathcal{M}_p}$.

Fix a public message codebook $\mathcal{C}_0^{(n)}$. For every $m_p \in \mathcal{M}_p$ and $j = 1, 2$, let $\mathsf{C}_j^{(n)}(m_p) \triangleq \{\mathbf{U}_j(m_p, m_{jj}, w_j, i_j)\}_{(m_{jj}, w_j, i_j) \in \mathcal{M}_{jj} \times \mathcal{W}_j \times \mathcal{I}_j}$, where $(m_{jj}, w_j, i_j) \in \mathcal{M}_{jj} \times \mathcal{W}_j \times \mathcal{I}_j$ and $\mathcal{I}_j \triangleq \left[1 : 2^{nR'_j}\right]$, be a random codebook of private messages j, consisting of conditionally independent random vectors each distributed according to $Q^n_{U_j|U_0=\mathbf{u}_0(m_p)}$. A realization of $\mathsf{C}_j^{(n)}(m_p)$ is denoted by $\mathcal{C}_j^{(n)}(m_p) \triangleq \left\{\mathbf{u}_j(m_p, m_{jj}, w_j, i_j)\right\}_{(m_{jj}, w_j, i_j) \in \mathcal{M}_{jj} \times \mathcal{W}_j \times \mathcal{I}_j}$.

We denote $\mathsf{C}_j^{(n)} \triangleq \left\{\mathsf{C}_j^{(n)}(m_p)\right\}_{m_p \in \mathcal{M}_p}$, and its realization by $\mathcal{C}_j^{(n)}$. A random codebook is denoted by $\mathsf{C}_n = \left\{\mathsf{C}_0^{(n)}, \mathsf{C}_1^{(n)}, \mathsf{C}_2^{(n)}\right\}$, while $\mathcal{C}_n = \left\{\mathcal{C}_0^{(n)}, \mathcal{C}_1^{(n)}, \mathcal{C}_2^{(n)}\right\}$ denotes a fixed codebook (a possible outcome of C_n). Denoting the set of all possible realizations of C_n by $\mathfrak{C}_n$, the above codebook construction induces a PMF $\mu \in \mathcal{P}(\mathfrak{C}_n)$ over the codebook ensemble. For every $\mathcal{C}_n \in \mathfrak{C}_n$, we have (61) at the top of this page.

For a fixed codebook $\mathcal{C}_n \in \mathfrak{C}_n$ we now describe its associated encoding function $f^{(\mathcal{C}_n)}$ and decoding functions $\phi_j^{(\mathcal{C}_n)}$, for $j = 1, 2$.

3) Encoder $f^{(\mathcal{C}_n)}$: Fix a codebook $\mathcal{C}_n \in \mathfrak{C}_n$. To transmit the message pair (m_0, m_1, m_2) the encoder transforms it into the triple $\big(m_p, m_{11}, m_{22}\big)$, and draws W_j uniformly from $\mathcal{W}_j$, $j = 1, 2$; denote the realization of W_j by $w_j \in \mathcal{W}_j$. Given $(m_p, m_{11}, m_{22}, w_1, w_2)$, a pair of indices $(i_1, i_2) \in \mathcal{I}_1 \times \mathcal{I}_2$ is randomly selected by the likelihood encoder according to

$$P^{(\mathcal{C}_n)}_{\mathsf{LE}}(i_1, i_2|m_p, m_{11}, m_{22}, w_1, w_2) \triangleq \frac{2^{i_{Q^n}\left(\mathbf{u}_1(m_p, m_{11}, w_1, i_1); \mathbf{u}_2(m_p, m_{22}, w_2, i_2)\middle|\mathbf{u}_0(m_p)\right)}}{\sum\limits_{\substack{(i'_1, i'_2) \\ \in \mathcal{I}_1 \times \mathcal{I}_2}} 2^{i_{Q^n}\left(\mathbf{u}_1(m_p, m_{11}, w_1, i'_1); \mathbf{u}_2(m_p, m_{22}, w_2, i'_2)\middle|\mathbf{u}_0(m_p)\right)}} \tag{62}$$

where i_{Q^n} stands for the information density with respect to the conditional product distribution $Q^n_{U_1,U_2|U_0}$ (and its marginals). The structure of $P^{(\mathcal{C}_n)}_{\mathsf{LE}}$ adheres to the setup of Lemma 3 from Section III and, in particular, to the stochastic choice of indices therein as described in (14). Replacing the commonly used joint typicality encoder with $P^{(\mathcal{C}_n)}_{\mathsf{LE}}$, we are able to establish several important properties of the chosen codewords and their induced distribution.

Let (i_1, i_2) be the selected pair of indices. The channel input sequence is randomly generated according to the conditional product distribution

$$Q^n_{X|U_0=\mathbf{u}_0(m_p), U_1=\mathbf{u}_1(m_p, m_{11}, w_1, i_1), U_2=\mathbf{u}_2(m_p, m_{22}, w_2, i_2)}.$$

4) Decoder $\phi_j^{(\mathcal{C}_n)}$: Decoder $j = 1, 2$ operates in two stages. First, it searches for a unique $\hat{m}_p \in \mathcal{M}_p$ such that

$$\Big(\mathbf{u}_0(\hat{m}_p), \mathbf{y}_j\Big) \in \mathcal{T}^n_\delta(Q_{U_0,Y_j}). \tag{63}$$

If no such unique index is found, set $\phi_j^{(\mathcal{C}_n)} = (1, 1)$. Otherwise, having $\hat{m}_p \in \mathcal{M}_p$, Decoder $j = 1, 2$ proceeds by looking for a unique pair $(\hat{m}_{jj}, \hat{w}_j) \in \mathcal{M}_{jj} \times \mathcal{W}_j$ for which there exists an index $\hat{i}_j \in \mathcal{I}_j$ such that

$$\Big(\mathbf{u}_0(\hat{m}_p), \mathbf{u}_j(\hat{m}_p, \hat{m}_{jj}, \hat{w}_j, \hat{i}_j), \mathbf{y}_j\Big) \in \mathcal{T}^n_\delta(Q_{U_0,U_j,Y_j}). \tag{64}$$

Recall that each $m_p \in \mathcal{M}_p$ specifies a triple $(m_0, m_{10}, m_{20}) \in \mathcal{M}_0 \times \mathcal{M}_{10} \times \mathcal{M}_{20}$. If the second stage is also executed successfully, then the decoder has a triple $(\hat{m}_p, \hat{m}_{jj}, \hat{w}_j) \in \mathcal{M}_p \times \mathcal{M}_{jj} \times \mathcal{W}_j$ with $\hat{m}_p$ and $(\hat{m}_{jj}, \hat{w}_j)$ being the unique indices satisfying (63) and (64), respectively. In this case we set $\phi_j^{(\mathcal{C}_n)}(\mathbf{y}_j) = \big(\hat{m}_0, \hat{m}_j\big)$, where $\hat{m}_j$ is assembled from $(\hat{m}_{j0}, \hat{m}_{jj})$; otherwise, set $\phi_j^{(\mathcal{C}_n)} = (1, 1)$.

5) Induced Code and Joint Distribution: The triple $\left(f^{(\mathcal{C}_n)}, \phi_1^{(\mathcal{C}_n)}, \phi_2^{(\mathcal{C}_n)}\right)$ defined with respect to the codebook $\mathcal{C}_n \in \mathfrak{C}_n$ constitutes an (n, R_0, R_1, R_2) code c_n for the BC with privacy leakage constraints. Thus, for every codebook $\mathcal{C}_n \in \mathfrak{C}_n$, the induced joint distribution is given in (65) at the top of this page, where the random variables $\mathbf{U}_0$, $\mathbf{U}_1$ and $\mathbf{U}_2$ are the chosen codewords at the conclusion of the encoding process (from which the input $\mathbf{X}$ to the BC is generated).

Taking the random codebook generation into account, we also have (66) at the top of the previous page, where $\mu \in \mathcal{P}(\mathfrak{C}_n)$ is described in (61). The PMF P induces a probability measure $\mathbb{P} \triangleq \mathbb{P}_P$, with respect to which the subsequent analysis is performed. Specifically, all the multi-letter information measures in the sequel are taken with respect to P from (66), while single-letter information terms are calculated with respect to $Q_{U_0,U_1,U_2,X,Y_1,Y_2}$.

6) Average Error Probability Analysis: The output sequences of $P_{\mathsf{LE}}^{(\mathcal{C}_n)}$ from (62) are jointly typical with high probability as long as the sum of the rates of the product bin is greater than the mutual information between the coding random variables [32, Th. 3]. The rest of the error probability analysis goes through via classic joint typicality arguments. The details of the analysis are relegated to Appendix B, where it is shown that

$$\mathbb{E}P_{\mathsf{e}}(\mathsf{C}_n) \leq \eta(n, \delta, \delta') \tag{67}$$

where $\delta' \in (0, \delta)$ and $\lim_{n\to\infty} \eta(n, \delta, \delta') = 0$ for all $0 < \delta' < \delta$, if

$$R_1' + R_2' > I(U_1; U_2|U_0) \tag{68a}$$
$$R_0 + R_{10} + R_{20} < I(U_0; Y_1) - \tau_\delta \tag{68b}$$
$$R_0 + R_{10} + R_{20} < I(U_0; Y_2) - \tau_\delta \tag{68c}$$
$$R_{11} + \tilde{R}_1 + R_1' < I(U_1; Y_1|U_0) - \tau_\delta \tag{68d}$$
$$R_{22} + \tilde{R}_2 + R_2' < I(U_2; Y_2|U_0) - \tau_\delta \tag{68e}$$

with $\tau_\delta \to 0$ as $\delta \to 0$. Furthermore, setting $\eta_n \triangleq \eta(n, \delta_n, \delta_n')$ where $\{\delta_n\}_{n\in\mathbb{N}}$ and $\{\delta_n'\}_{n\in\mathbb{N}}$ are sequences that converge sufficiently slowly to zero as n grows, we have $\lim_{n\to\infty} \eta_n = 0$. To clarify, the δ' that appears in (67) and in upper bounds below is a consequence of the Conditional Typicality Lemma [20, Sec. 2.5]. This lemma considers conditioning on sequences that are jointly letter-typical with respect to a slightly smaller gap than the original δ. This smaller gap is δ'.

7) Properties for Leakage Analysis: In contrast to previous works, we do not analyse the expected leakages of the random code. Instead, we establish certain properties that the random code possesses and then extract a specific sequence of codes that satisfies these properties as well as reliability. It is then shown that the extracted sequence of codes admits the leakage constraints.

By symmetry, we consider only the properties required for the analysis of the rate-leakage from M_1 to the 2nd receiver. The corresponding derivations for M_2 follows similar lines and the resulting rate constraints match up to changing some indices.

We first need a decodability property. Specifically, Decoder 2 should be able to decode (W_1, I_1) with a low error probability based on $(M_p, M_{11}, M_{22}, W_2, I_2, \mathbf{Y}_2)$. We consider a decoding rule based on a joint typicality test: Decoder 2 searches for a unique pair $(\check{w}_1, \check{i}_1) \in \mathcal{W}_1 \times \mathcal{I}_1$ such that

$$\Big(\mathbf{u}_0(m_p), \mathbf{u}_1(m_p, m_{11}, \check{w}_1, \check{i}_1), \mathbf{u}_2(m_p, m_{22}, w_2, i_2), \mathbf{y}_2\Big) \in \mathcal{T}_\delta^n(Q_{U_0,U_1,U_2,Y_2}). \tag{69}$$

For a fixed codebook $\mathcal{C}_n \in \mathfrak{C}_n$ (which specifies a code c_n), let $P_1^{(\mathsf{Leak})}(\mathcal{C}_n)$ denote the probability that Decoder 2 fails in this decoding process. As explained in Appendix B, we have

$$\mathbb{E}P_1^{(\mathsf{Leak})}(\mathsf{C}_n) \leq \kappa(n, \delta, \delta') \tag{70}$$

where $\delta' \in (0, \delta)$ and $\lim_{n\to\infty} \kappa(n, \delta, \delta') = 0$ for all $0 < \delta' < \delta$, if

$$\tilde{R}_1 < I(U_1; Y_2|U_0, U_2) - \xi_\delta \tag{71a}$$
$$\tilde{R}_1 + R_1' < I(U_1; U_2, Y_2|U_0) - \xi_\delta \tag{71b}$$

with $\xi_\delta \to 0$ as $\delta \to 0$. Again, by allowing δ and δ' from (70) to converge to zero sufficiently slow with n, $\kappa(n, \delta, \delta')$ may be replaced by a κ_n with $\lim_{n\to\infty} \kappa_n = 0$.

We are now ready to state Lemmas 4-6. Proofs are given in Appendices C-E.

Lemma 4: If (71) is valid with $\xi_\delta \to 0$ as $\delta \to 0$, then there exists $\zeta_1(n, \delta, \delta')$, where $\delta' \in (0, \delta)$, such that

$$H(W_1, I_1|M_p, M_{11}, M_{22}, W_2, I_2, \mathbf{Y}_2, \mathsf{C}_n) \leq n\zeta_1(n, \delta, \delta') \tag{72}$$

and $\lim_{n\to\infty} \zeta_1(n, \delta, \delta') = 0$ for all $0 < \delta' < \delta$. Furthermore, setting $\zeta_{1,n} \triangleq \zeta_1(n, \delta_n, \delta_n')$ where $\{\delta_n\}_{n\in\mathbb{N}}$ and $\{\delta_n'\}_{n\in\mathbb{N}}$ are sequences that decay sufficiently slow to zero as n grows, we have $\lim_{n\to\infty} \zeta_{1,n} = 0$.

Lemma 5: There exist $\zeta_2(n, \delta, \delta')$ that satisfies the same properties as $\zeta_1(n, \delta, \delta')$ from Lemma 4, such that

$$I(\mathbf{U}_1; \mathbf{Y}_2|\mathbf{U}_0, \mathbf{U}_2, \mathsf{C}_n) \leq nI(U_1; Y_2|U_0, U_2) + n\zeta_2(n, \delta, \delta'). \tag{73}$$

Lemma 6: There exists $\zeta_3(n, \delta, \delta')$ that satisfies the same properties as $\zeta_1(n, \delta, \delta')$ from Lemma 4, such that

$$I(\mathbf{U}_1; M_{22}, W_2, I_2|M_p, \mathsf{C}_n) \leq nI(U_1; U_2|U_0) + n\zeta_3(n, \delta, \delta'). \tag{74}$$

The Uniform Approximation Lemma from Section III further implies that if

$$R_2' + \min\left\{R_1', R_{22} + \tilde{R}_2\right\} > I(U_1; U_2|U_0) + \delta \tag{75}$$

then there exist $\zeta_{4,n}$ with $\lim_{n\to\infty} \zeta_{4,n}$ such that

$$\mathbb{E}_{\mathsf{C}_n}\left|\left|P_{M_p,M_{11},W_1,I_1|\mathsf{C}_n} - p^{(U)}_{\mathcal{M}_p\times\mathcal{M}_{11}\times\mathcal{W}_1\times\mathcal{I}_1}\right|\right|_{\mathsf{TV}} \leq \zeta_{4,n} \tag{76}$$

where $p^{(U)}_{\mathcal{M}_p\times\mathcal{M}_{11}\times\mathcal{W}_1\times\mathcal{I}_1}$ is the uniform distribution on $\mathcal{M}_p \times \mathcal{M}_{11} \times \mathcal{W}_1 \times \mathcal{I}_1$. To see this, observe that by symmetry we have

$$\begin{aligned}
&\mathbb{E}_{\mathsf{C}_n}\left|\left|P_{M_p,M_{11},W_1,I_1|\mathsf{C}_n} - p^{(U)}_{\mathcal{M}_p\times\mathcal{M}_{11}\times\mathcal{W}_1\times\mathcal{I}_1}\right|\right|_{\mathsf{TV}} \\
&\quad= \sum_{\substack{(m_p,m_{11},w_1)\\ \in\mathcal{M}_p\times\mathcal{M}_{11}\times\mathcal{W}_1}} \frac{1}{|\mathcal{M}_p||\mathcal{M}_{11}||\mathcal{W}_1|} \\
&\qquad\times\mathbb{E}_{\mathsf{C}_n}\left|\left|P_{I_1|M_p=m_p,M_{11}=m_{11},W_1=w_1,\mathsf{C}_n} - p^{(U)}_{\mathcal{I}_1}\right|\right|_{\mathsf{TV}} \\
&\quad= \mathbb{E}_{\mathsf{C}_n}\left|\left|P_{I_1|M_p=1,M_{11}=1,W_1=1,\mathsf{C}_n} - p^{(U)}_{\mathcal{I}_1}\right|\right|_{\mathsf{TV}}.
\end{aligned} \tag{77}$$

Note that $(M_p, M_{11}, W_1) = (1, 1, 1)$ fixes a single u_1-bin (comprising $2^{nR_1'}$ codewords), while the pair (M_{22}, W_2) (of total rate $R_{22}+\tilde{R}_2$) uniformly chooses a u_2-bin (comprising $2^{nR_2'}$ codewords). Lemma 3 now gives the desired relation because bins are generated conditionally independent given

$\mathbf{U}_0$ and the chosen codeword pair is drawn according to $P_{\mathsf{LE}}^{(\mathcal{C}_n)}$ from (62) which adheres to the structure of (14).

We now invoke the Selection Lemma [45, Lemma 5] to extract a specific sequence of codes that satisfies several desired properties. We restate this lemma next.

Lemma 7 (Selection Lemma): Let $\{A_n\}_{n\in\mathbb{N}}$ be a sequence of random variables, where A_n takes values in $\mathcal{A}_n$. Let $\left\{f_n^{(1)}, f_n^{(2)}, \ldots, f_n^{(\mathrm{J})}\right\}_{n\in\mathbb{N}}$ be a collection of $\mathrm{J} < \infty$ sequences of bounded functions $f_n^{(i)} : \mathcal{A}_n \to \mathbb{R}_+$, $j \in [1 : \mathrm{J}]$. If

$$\mathbb{E} f_n^{(j)}(A_n) \xrightarrow[n\to\infty]{} 0, \quad \forall j \in [1 : \mathrm{J}] \tag{78a}$$

then there exists a sequence $\{a_n\}_{n\in\mathbb{N}}$, where $a_n \in \mathcal{A}_n$ for every $n \in \mathbb{N}$, such that

$$f_n^{(j)}(a_n) \xrightarrow[n\to\infty]{} 0, \quad \forall j \in [1 : \mathrm{J}]. \tag{78b}$$

Consider the sequence of random codes $\{\mathsf{C}_n\}_{n\in\mathbb{N}}$, the functions[6]

$$f_n^{(1)}(\mathcal{C}_n) \triangleq P_{\mathsf{e}}(\mathcal{C}_n) \tag{79a}$$

$$f_n^{(2)}(\mathcal{C}_n) \triangleq H(W_1, I_1|M_p, M_{11}, M_{22}, W_2, I_2, \mathbf{Y}_2, \mathsf{C}_n = \mathcal{C}_n) \tag{79b}$$

$$f_n^{(3)}(\mathcal{C}_n) \triangleq \frac{1}{n} I(\mathbf{U}_1; \mathbf{Y}_2|\mathbf{U}_0, \mathbf{U}_2, \mathsf{C}_n = \mathcal{C}_n) - I(U_1; Y_2|U_0, U_2) \tag{79c}$$

$$f_n^{(4)}(\mathcal{C}_n) \triangleq \frac{1}{n} I(\mathbf{U}_1; M_{22}, W_2, I_2|M_p, \mathsf{C}_n = \mathcal{C}_n) - I(U_1; U_2|U_0) \tag{79d}$$

$$f_n^{(5)}(\mathcal{C}_n) \triangleq \left|\left| P_{M_p, M_{11}, W_1, I_1|\mathsf{C}_n=\mathcal{C}_n} - p^{(U)}_{\mathcal{M}_p\times\mathcal{M}_{11}\times\mathcal{W}_1\times\mathcal{I}_1} \right|\right|_{\mathsf{TV}} \tag{79e}$$

as well as the functions $f_n^{(6)}$, $f_n^{(7)}$ and $f_n^{(8)}$ that correspond to $f_n^{(2)}$, $f_n^{(3)}$ and $f_n^{(4)}$, respectively, with respect to the analysis for M_2. We also impose constraints on the rates that arise from repeating the above steps for M_2. Namely, we set

$$\tilde{R}_2 < I(U_2; Y_1|U_0, U_1) - \xi(\delta) \tag{80a}$$

$$\tilde{R}_2 + R_2' < I(U_2; U_1, Y_1|U_0) - \xi(\delta) \tag{80b}$$

and

$$R_1' + \min\left\{R_2', R_{11} + \tilde{R}_1\right\} > I(U_1; U_2|U_0) + \delta \tag{80c}$$

in accordance with (71) and (75), respectively. This implies that results analog to those of Lemmas 4-6 and (76) hold for M_2.

Replacing δ and δ' in the definitions of $\eta_j(n, \delta, \delta')$, for $j = 1, 2, 3$, with $\{\delta_n\}_{n\in\mathbb{N}}$ and $\{\delta_n'\}_{n\in\mathbb{N}}$ that decay to zero sufficiently slow, we have

$$\mathbb{E}_{\mathsf{C}_n} f_n^{(j)}(\mathsf{C}_n) \xrightarrow[n\to\infty]{} 0, \quad j \in [1 : 7]. \tag{81}$$

Lemma 7 now implies the existence of a sequence of codebooks $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ (each inducing an (n, R_1, R_1, R_2) code c_n) and another sequence of numbers $\{\eta_n\}_{n\in\mathbb{N}}$ with $\lim_{n\to\infty}\eta_n = 0$, such that for $j \in [1 : 7]$ we have

$$f_n^{(j)}(\mathcal{C}_n) \le \eta_n, \quad \forall n \in \mathbb{N}. \tag{82}$$

[6]We slightly abuse notation in the definition of $f_n^{(1)}$ because P_{e} is a function of the code c_n rather than the codebook $\mathcal{C}_n$. However, since $\mathcal{C}_n$ uniquely defines c_n we prefer this presentation for the sake of simplicity.

8) Leakage Analysis of M_1 Under $\mathcal{C}_n$: All subsequent information measures are calculated with respect to $P^{(\mathcal{C}_n)}$ from (65). We emphasize this by using $H_{\mathcal{C}_n}$ and $I_{\mathcal{C}_n}$ as the notation of such entropy or mutual information terms, respectively.

First, because $f_n^{(5)}(\mathcal{C}_n) \xrightarrow[n\to\infty]{} 0$ and by the continuity of entropy, there exists a sequence $\{\theta_n\}_{n\in\mathbb{N}}$ with $\lim_{n\to\infty}\theta_n = 0$, such that

$$\left| H_{\mathcal{C}_n}(M_{11}, W_1, I_1|M_p) - \log\left(|\mathcal{M}_{11}||\mathcal{W}_1||\mathcal{I}_1|\right)\right| \le \theta_n \tag{83}$$

for every $n \in \mathbb{N}$. Next, since

$$\ell_1(c_n) = \frac{1}{n} I_{\mathcal{C}_n}(M_1; \mathbf{Y}_2) = R_1 - \frac{1}{n} H_{\mathcal{C}_n}(M_1|\mathbf{Y}_2) \tag{84}$$

we can upper bound the leakage of M_1 to the second receiver by lower bounding the conditional entropy term from the RHS of (84). We have

$$\begin{aligned}
&H_{\mathcal{C}_n}(M_1|\mathbf{Y}_2)\\
&\overset{(a)}{\ge} H_{\mathcal{C}_n}(M_{11}|M_p, M_{22}, W_2, I_2, \mathbf{Y}_2)\\
&= H_{\mathcal{C}_n}(M_{11}, \mathbf{Y}_2|M_p, M_{22}, W_2, I_2)\\
&\quad - H_{\mathcal{C}_n}(\mathbf{Y}_2|M_p, M_{22}, W_2, I_2)\\
&\overset{(b)}{\ge} H_{\mathcal{C}_n}(M_{11}, \mathbf{Y}_2|M_p, M_{22}, W_2, I_2) - H_{\mathcal{C}_n}(\mathbf{Y}_2|\mathbf{U}_0, \mathbf{U}_2)\\
&= H_{\mathcal{C}_n}(M_{11}, W_1, I_1, \mathbf{Y}_2|M_p, M_{22}, W_2, I_2)\\
&\quad - H_{\mathcal{C}_n}(W_1, I_1|M_p, M_{11}, M_{22}, W_2, I_2, \mathbf{Y}_2) - H(\mathbf{Y}_2|\mathbf{U}_0, \mathbf{U}_2)\\
&= H_{\mathcal{C}_n}(M_{11}, W_1, I_1|M_p, M_{22}, W_2, I_2)\\
&\quad + H_{\mathcal{C}_n}(\mathbf{Y}_2|M_p, M_{11}, W_1, I_1, M_{22}, W_2, I_2)\\
&\quad - H_{\mathcal{C}_n}(W_1, I_1|M_p, M_{11}, M_{22}, W_2, I_2, \mathbf{Y}_2)\\
&\quad - H_{\mathcal{C}_n}(\mathbf{Y}_2|\mathbf{U}_0, \mathbf{U}_2)\\
&\overset{(c)}{=} H_{\mathcal{C}_n}(M_{11}, W_1, I_1|M_p, M_{22}, W_2, I_2)\\
&\quad - H_{\mathcal{C}_n}(W_1, I_1|M_p, M_{11}, M_{22}, W_2, I_2, \mathbf{Y}_2)\\
&\quad - I_{\mathcal{C}_n}(\mathbf{U}_1; \mathbf{Y}_2|\mathbf{U}_0, \mathbf{U}_2)\\
&= H_{\mathcal{C}_n}(M_{11}, W_1, I_1|M_p) - I_{\mathcal{C}_n}(\mathbf{U}_1; M_{22}, W_2, I_2|M_p)\\
&\quad - H_{\mathcal{C}_n}(W_1, I_1|M_p, M_{11}, M_{22}, W_2, I_2, \mathbf{Y}_2)\\
&\quad - I_{\mathcal{C}_n}(\mathbf{U}_1; \mathbf{Y}_2|\mathbf{U}_0, \mathbf{U}_2)
\end{aligned} \tag{85}$$

where:

(a) is because conditioning cannot increase entropy and since M_1 corresponds to the pair (M_{10}, M_{11}) while $M_p = (M_0, M_{10}, M_{20})$;

(b) follows because $\mathbf{U}_0$ and $\mathbf{U}_2$ are specified by (M_p, M_{22}, W_2, I_2) and since conditioning cannot increase entropy;

(c) uses the deterministic relations stated in (b) along with $\mathbf{U}_1$ being determined by (M_p, M_{11}, W_1, I_1) and the Markov relation $\mathbf{Y}_2 - (\mathbf{U}_0, \mathbf{U}_1, \mathbf{U}_2) - (M_p, M_{11}, W_1, I_1, M_{22}, W_2, I_2)$.

Inserting (82) (for $j \in [2 : 4]$), (83) and $R_{11} = R_1 - R_{10}$ into (85) further gives

$$\begin{aligned}
&H_{\mathcal{C}_n}(M_1|\mathbf{Y}_2)\\
&\ge n\left(R_1 - R_{10} + \tilde{R}_1 + R_1' - I(U_1; U_2, Y_2|U_0) - 3\eta_n - \theta_n\right)\\
&\overset{(a)}{\ge} nR_1 - n\left(L_1 + 3\eta_n + \theta_n\right)
\end{aligned}$$

where (a) follows by taking

$$\tilde{R}_1 + R_1' - R_{10} > I(U_1; U_2, Y_2|U_0) - L_1 \tag{86a}$$
$$R_1' + L_1 - R_{10} > I(U_1; U_2|U_0). \tag{86b}$$

The bound in (86b) ensures the feasibility of an $\tilde{R}_1 > 0$ that satisfies (71a) and (86a) simultaneously. The corresponding rate bounds for the analysis of $\ell_2(c_n)$ are

$$\tilde{R}_2 + R_2' - R_{20} > I(U_2; U_1, Y_1|U_0) - L_2 \tag{87a}$$
$$R_2' + L_2 - R_{20} > I(U_1; U_2|U_0). \tag{87b}$$

Recalling that η_n and θ_n can be made arbitrarily small with n, there exists $n_0(\epsilon) \in \mathbb{N}$, such that for all $n > n_0(\epsilon)$

$$P_{\mathsf{e}}(c_n) \leq \epsilon \tag{88a}$$
$$\ell_1(c_n) \leq L_1 + \epsilon \tag{88b}$$
$$\ell_2(c_n) \leq L_2 + \epsilon \tag{88c}$$

as required.

Our last step is to apply Fourier-Motzkin Elimination (FME) on (68), (75), (80) and (86)-(87), while using (60) and the non-negativity of the involved terms, to eliminate R_{j0}, R_j' and $\tilde{R}_j$, for $j = 1, 2$. Since all the above linear inequalities have constant coefficients, the FME can be performed by a computer program, e.g., by the FME-IT software [46]. This shows the sufficiency of (23).

C. Proof of Corollary 2

Fix $(L_1, L_2) \in \mathbb{R}_+^2$ and $Q_{U_0,U_1,U_2,X} \in \mathcal{P}(\mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{X})$. The rate bounds describing $\tilde{\mathcal{R}}_{\mathsf{I}}(L_1, L_2, Q_{U_0,U_1,U_2,X})$ are:

$$R_1 \leq I(U_1; Y_1|U_0) - I(U_1; U_2, Y_2|U_0) + L_1 \tag{89a}$$
$$R_1 \leq I(U_1; Y_1|U_0) + \min\Big\{I(U_0; Y_1), I(U_0; Y_2)\Big\} \tag{89b}$$
$$R_2 \leq I(U_2; Y_2|U_0) - I(U_2; U_1, Y_1|U_0) + L_2 \tag{89c}$$
$$R_2 \leq I(U_2; Y_2|U_0) + \min\Big\{I(U_0; Y_1), I(U_0; Y_2)\Big\} \tag{89d}$$
$$\begin{aligned} R_1 + R_2 \leq\ & I(U_1; Y_1|U_0) + I(U_2; Y_2|U_0) - I(U_1; U_2|U_0) \\ & + \min\Big\{I(U_0; Y_1), I(U_0; Y_2)\Big\}. \end{aligned} \tag{89e}$$

To prove the first claim, assume that $L_1 \geq L_1^\star(Q_{U_0,U_1,U_2,X})$. Consequently, the term inside the positive part function from the RHS of (89a) is non-negative as it satisfies

$$\begin{aligned} & I(U_1; Y_1|U_0) - I(U_1; U_2, Y_2|U_0) + L_1 \\ & \quad \geq I(U_1; Y_1|U_0) + \min\Big\{I(U_0; Y_1), I(U_0; Y_2)\Big\} \end{aligned} \tag{90}$$

which makes (89a) inactive due to (89b), and therefore, $\tilde{\mathcal{R}}_{\mathsf{O}}(L_1, L_2, Q_{U_0,U_1,U_2,X}) = \tilde{\mathcal{R}}_{\mathsf{O}}(\infty, L_2, Q_{U_0,U_1,U_2,X})$.

An analogous argument with respect to L_2 proves the second claim (essentially by showing that if $L_2 \geq L_2^\star$ then (89c) is inactive due (89d)). The third claim follows by combining both preceding arguments.

D. Proof of Theorem 3

We show that given an (L_1, L_2)-achievable rate triple (R_0, R_1, R_2), there is a PMF $Q_{W,U,V,X} \in \mathcal{P}(\mathcal{W} \times \mathcal{U} \times \mathcal{V} \times \mathcal{X})$, such that (27) holds when the information measures are calculated with respect to $Q_{W,U,V,X} W_{Y_1,Y_2|X}$. Due to the symmetric structure of the rate bounds defining $\mathcal{R}_{\mathsf{O}}(L_1, L_2)$, we present only the derivation of (27a)-(27d) and (27h). The other inequalities from (27) are established by similar arguments.

Since (R_0, R_1, R_2) is (L_1, L_2)-achievable, for every $\epsilon > 0$ there is a sufficiently large $n \in \mathbb{N}$ and an (n, R_0, R_1, R_2) code c_n for which (22) holds. We note that all subsequent entropy and mutual information terms are calculated with respect to the PMF from (19) that is specified by c_n.

Fix $\epsilon > 0$ and find the corresponding blocklength $n \in \mathbb{N}$. By Fano's inequality we have

$$H(M_0, M_j|Y_j^n) \leq 1 + n\epsilon R_j \triangleq n\delta_{n,\epsilon}^{(j)}, \quad j = 1, 2. \tag{91}$$

Define $\delta_{n,\epsilon} = \max\left\{\delta_{n,\epsilon}^{(1)}, \delta_{n,\epsilon}^{(2)}\right\}$. Next, by (22b), we write

$$\begin{aligned} n(L_1 + \epsilon) &\geq I(M_1; Y_2^n) \\ &= I(M_1; M_0, M_2, Y_2^n) - I(M_1; M_0, M_2|Y_2^n) \\ &\overset{(a)}{\geq} I(M_1; Y_2^n|M_0, M_2) - H(M_0, M_2|Y_2^n) \\ &\overset{(b)}{\geq} I(M_1; Y_2^n|M_0, M_2) - n\delta_{n,\epsilon} \end{aligned} \tag{92}$$

where (a) uses the independence of M_1 and (M_0, M_2) and the non-negativity of entropy, while (b) is by (91). (92) implies

$$I(M_1; Y_2^n|M_0, M_2) \leq nL_1 + n(\epsilon + \delta_{n,\epsilon}). \tag{93}$$

Similarly, we have

$$I(M_1; Y_2^n|M_0) \leq nL_1 + n(\epsilon + \delta_{n,\epsilon}). \tag{94}$$

The common message rate R_0 satisfies

$$\begin{aligned} nR_0 &= H(M_0) \\ &\overset{(a)}{\leq} I(M_0; Y_1^n) + n\delta_{n,\epsilon} \\ &= \sum_{i=1}^n I(M_0; Y_{1,i}|Y_1^{i-1}) + n\delta_{n,\epsilon} \\ &\leq \sum_{i=1}^n I(M_0, Y_1^{i-1}; Y_{1,i}) + n\delta_{n,\epsilon} \qquad (95a) \\ &\overset{(b)}{\leq} \sum_{i=1}^n I(W_i; Y_{1,i}) + n\delta_{n,\epsilon} \qquad (95b) \end{aligned}$$

where (a) uses (91) and (b) defines $W_i \triangleq (M_0, Y_1^{i-1}, Y_{2,i+1}^n)$. By reversing the roles of Y_1^n and Y_2^n and repeating similar steps, we also have

$$\begin{aligned} nR_0 &\leq \sum_{i=1}^n I(M_0, Y_{2,i+1}^n; Y_{2,i}) + n\delta_{n,\epsilon} \qquad (96a) \\ &\leq \sum_{i=1}^n I(W_i; Y_{2,i}) + n\delta_{n,\epsilon}. \qquad (96b) \end{aligned}$$

For R_1, it follows that

$$\begin{aligned} nR_1 &= H(M_1|M_0, M_2) \\ &\overset{(a)}{\leq} I(M_1; Y_1^n|M_0, M_2) - I(M_1; Y_2^n|M_0, M_2) \\ &\quad + nL_1 + n\xi_{n,\epsilon} \\ &\overset{(b)}{=} \sum_{i=1}^{n} \Big[I(M_1; Y_1^i, Y_{2,i+1}^n|M_0, M_2) \\ &\quad - I(M_1; Y_1^{i-1}, Y_{2,i}^n|M_0, M_2) \Big] + nL_1 + n\xi_{n,\epsilon} \\ &= \sum_{i=1}^{n} \Big[I(M_1; Y_{1,i}|M_2, W_i) \\ &\quad - I(M_1; Y_{2,i}|M_2, W_i) \Big] + nL_1 + n\xi_{n,\epsilon} \\ &\overset{(c)}{=} \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}|W_i, V_i) - I(U_i; Y_{2,i}|W_i, V_i) \Big] \\ &\quad + nL_1 + n\xi_{n,\epsilon} \end{aligned} \tag{97}$$

where (a) uses (91) and (92) and $\xi_{n,\epsilon} = 2\delta_{n,\epsilon} + \epsilon$, (b) follows from a telescoping identity [47, eqs. (9) and (11)], and (c) uses $U_i \triangleq (M_1, W_i)$ and $V_i \triangleq (M_2, W_i)$.

R_1 is also upper bounded as

$$\begin{aligned} nR_1 &= H(M_1|M_0) \\ &\overset{(a)}{\leq} I(M_1; Y_1^n|M_0) - I(M_1; Y_2^n|M_0) + nL_1 + n\xi_{n,\epsilon} \\ &\overset{(b)}{=} \sum_{i=1}^{n} \Big[I(M_1; Y_1^i, Y_{2,i+1}^n|M_0) - I(M_1; Y_1^{i-1}, Y_{2,i}^n|M_0) \Big] \\ &\quad + nL_1 + n\xi_{n,\epsilon} \\ &\overset{(c)}{=} \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}|W_i) - I(U_i; Y_{2,i}|W_i) \Big] + nL_1 + n\xi_{n,\epsilon} \end{aligned} \tag{98}$$

where (a) is by (91) and (94), (b) uses a telescoping identity, while (c) follows by the definition of (W_i, U_i).

For the sum $R_0 + R_1$, we have

$$\begin{aligned} n(R_0 + R_1) &= H(M_0, M_1) \\ &\overset{(a)}{\leq} I(M_0, M_1; Y_1^n) + n\delta_{n,\epsilon} \\ &\overset{(b)}{\leq} \sum_{i=1}^{n} I(W_i, U_i; Y_{1,i}) + n\delta_{n,\epsilon} \end{aligned} \tag{99}$$

where (a) follows from (91) and (b) follows by the definition of (W_i, U_i). Moreover, consider

$$\begin{aligned} n(R_0 + R_1) &= H(M_1|M_0) + H(M_0) \\ &\overset{(a)}{\leq} I(M_1; Y_1^n|M_0) + I(M_0; Y_2^n) + n\delta_{n,\epsilon} \\ &\leq \sum_{i=1}^{n} \Big[I(M_1, Y_{2,i+1}^n; Y_{1,i}|M_0, Y_1^{i-1}) \\ &\quad + I(M_0; Y_{2,i}|Y_{2,i+1}^n) \Big] + n\delta_{n,\epsilon} \\ &= \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}|W_i) + I(Y_{2,i+1}^n; Y_{1,i}|M_0, Y_1^{i-1}) \\ &\quad + I(M_0; Y_{2,i}|Y_{2,i+1}^n) \Big] + n\delta_{n,\epsilon} \\ &\overset{(b)}{=} \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}|W_i) + I(Y_1^{i-1}; Y_{2,i}|M_0, Y_{2,i+1}^n) \\ &\quad + I(M_0; Y_{2,i}|Y_{2,i+1}^n) \Big] + n\delta_{n,\epsilon} \\ &\overset{(c)}{\leq} \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}|W_i) + I(W_i; Y_{2,i}) \Big] + n\delta_{n,\epsilon} \end{aligned} \tag{100}$$

where (a) is by (91), (b) is Csiszár's sum identity, while (c) uses the definition of (W_i, U_i).

To bound the sum $R_0 + R_1 + R_2$, we start by writing

$$\begin{aligned} &H(M_1|M_0, M_2) \\ &\overset{(a)}{\leq} I(M_1; Y_1^n|M_0, M_2) + n\delta_{n,\epsilon} \\ &= \sum_{i=1}^{n} I(M_1; Y_{1,i}|M_0, M_2, Y_1^{i-1}) + n\delta_{n,\epsilon} \\ &\leq \sum_{i=1}^{n} I(M_1, Y_{2,i+1}^n; Y_{1,i}|M_0, M_2, Y_1^{i-1}) + n\delta_{n,\epsilon} \\ &\overset{(b)}{=} \sum_{i=1}^{n} \Big[I(U_i; Y_{1,i}|W_i, V_i) \\ &\quad + I(Y_{2,i+1}^n; Y_{1,i}|M_0, M_2, Y_1^{i-1}) \Big] + n\delta_{n,\epsilon} \end{aligned} \tag{103}$$

where (a) uses (91) and (b) follows by the definition of (W_i, U_i, V_i). Moreover, we have

$$\begin{aligned} &H(M_2|M_0) \\ &\overset{(a)}{\leq} I(M_2; Y_2^n|M_0) + n\delta_{n,\epsilon} \\ &\overset{(b)}{=} \sum_{i=1}^{n} \Big[I(M_2; Y_{2,i}^n|M_0, Y_1^{i-1}) \\ &\quad - I(M_2; Y_{2,i+1}^n|M_0, Y_1^i) \Big] + n\delta_{n,\epsilon} \\ &\overset{(c)}{=} \sum_{i=1}^{n} \Big[I(M_2; Y_{2,i+1}^n|M_0, Y_1^{i-1}) + I(V_i; Y_{2,i}|W_i) \\ &\quad - I(M_2; Y_{1,i}, Y_{2,i+1}^n|M_0, Y_1^{i-1}) \\ &\quad + I(M_2; Y_{1,i}|M_0, Y_1^{i-1}) \Big] + n\delta_{n,\epsilon} \\ &\overset{(d)}{=} \sum_{i=1}^{n} \Big[I(V_i; Y_{2,i}|W_i) - I(V_i; Y_{1,i}|W_i) \\ &\quad + I(M_2; Y_{1,i}|M_0, Y_1^{i-1}) \Big] + n\delta_{n,\epsilon} \end{aligned} \tag{104}$$

where:

(a) follows from (91);
(b) is a telescoping identity;
(c) is by the mutual information chain rule and the definition of (V_i, U_i);

(d) uses the mutual information chain rule again. Combining (103) and (104) yields

$$\begin{aligned} n(R_1+R_2) &\le \sum_{i=1}^{n}\Big[I(U_i;Y_{1,i}|W_i,V_i)+I(V_i;Y_{2,i}|W_i) \\ &\quad -I(V_i;Y_{1,i}|W_i)+I(M_2,Y_{2,i+1}^n;Y_{1,i}|M_0,Y_1^{i-1})\Big] \\ &\quad +2n\delta_{n,\epsilon} \\ &= \sum_{i=1}^{n}\Big[I(U_i;Y_{1,i}|W_i,V_i)+I(V_i;Y_{2,i}|W_i) \\ &\quad +I(Y_{2,i+1}^n;Y_{1,i}|M_0,Y_1^{i-1})\Big]+2n\delta_{n,\epsilon} \end{aligned} \tag{105a}$$

Applying Csiszár's sum identity on the last term in (105a) gives

$$\begin{aligned} n(R_1+R_2) = \sum_{i=1}^{n}\Big[&I(U_i;Y_{1,i}|W_i,V_i)+I(V_i;Y_{2,i}|W_i) \\ &+I(Y_1^{i-1};Y_{2,i}|M_0,Y_{2,i+1}^n)\Big]+2n\delta_{n,\epsilon}. \end{aligned} \tag{105b}$$

Combining (95a) with (105a) and (96a) with (105b) yields

$$\begin{aligned} n(R_0+R_1+R_2) \le \sum_{i=1}^{n}\Big[&I(U_i;Y_{1,i}|W_i,V_i) \\ &+I(V_i;Y_{2,i}|W_i)+I(W_i;Y_{1,i})\Big]+3n\delta_{n,\epsilon} \end{aligned} \tag{106}$$

and

$$\begin{aligned} n(R_0+R_1+R_2) \le \sum_{i=1}^{n}\Big[&I(U_i;Y_{1,i}|W_i,V_i) \\ &+I(V_i;Y_{2,i}|W_i)+I(W_i;Y_{2,i})\Big]+3n\delta_{n,\epsilon}, \end{aligned} \tag{107}$$

respectively.

By repeating similar steps, we obtain bounds related to the remaining rate bounds in (27) as given in (108a)-(108f) at the bottom of this page.

The bounds are rewritten by introducing a time-sharing random variable Q that is uniformly distributed over the set $[1:n]$ and is independent of all the other random variables whose distribution is described in (19). For instance, the bound (97) is rewritten as

$$\begin{aligned} R_1 &\le \frac{1}{n}\sum_{q=1}^{n}\Big[I(U_q;Y_{1,q}|W_q,V_q)-I(U_q;Y_{2,q}|W_q,V_q)\Big] \\ &\quad +L_1+\xi_{n,\epsilon} \\ &= \sum_{i=q}^{n}\mathbb{P}(Q=q)\Big[I(U_Q;Y_{1,Q}|W_Q,V_Q,Q=q) \\ &\quad -I(U_Q;Y_{2,Q}|W_Q,V_Q,Q=q)\Big]+L_1+\xi_{n,\epsilon} \\ &\le I(U_Q;Y_{1,Q}|W_Q,V_Q,Q)-I(U_Q;Y_{2,Q}|W_Q,V_Q,Q) \\ &\quad +L_1+n\xi_{n,\epsilon} \end{aligned} \tag{109}$$

Denote $Y_1 \triangleq Y_{1,Q}$, $Y_2 \triangleq Y_{2,Q}$, $W \triangleq (W_Q,Q)$, $U \triangleq (U_Q,Q)$ and $V \triangleq (V_Q,Q)$. We thus have the bounds from (27) with the added terms $\delta_{n,\epsilon}$ and $\xi_{n,\epsilon}$, which can be made arbitrarily small by increasing the blocklength n while decreasing ϵ.

To complete the converse proof, note that since the channel is memoryless and without feedback, and because $U_q = (M_1,W_q)$ and $V_q=(M_2,W_q)$, the chain

$$(Y_{1,q},Y_{2,q})-X_q-(U_q,V_q)-W_q \tag{110}$$

is Markov, for every $q\in[1:n]$. This implies that $(Y_1,Y_2)-X-(U,V)-W$ forms a Markov chain, which establishes Theorem 3.

E. Proof of Theorem 4

The direct part of Theorem 4 follows by setting $U_0=W$, $U_1=Y_1$ and $U_2=V$ into $\mathcal{C}_{\mathsf{SD}}(L_1,L_2)$, which establishes its inclusion in $\mathcal{R}_{\mathsf{I}}(L_1,L_2)$.

$$nR_2 \le \sum_{i=1}^{n}\Big[I(V_i;Y_{2,i}|W_i,U_i)-I(V_i;Y_{1,i}|W_i,U_i)\Big]+nL_2+n\xi_{n,\epsilon} \tag{108a}$$

$$nR_2 \le \sum_{i=1}^{n}\Big[I(V_i;Y_{2,i}|W_i)-I(V_i;Y_{1,i}|W_i)\Big]+nL_2+n\xi_{n,\epsilon} \tag{108b}$$

$$n(R_0+R_2) \le \sum_{i=1}^{n}I(W_i,V_i;Y_{2,i})+n\delta_{n,\epsilon} \tag{108c}$$

$$n(R_0+R_2) \le \sum_{i=1}^{n}\Big[I(V_i;Y_{2,i}|W_i)+I(W_i;Y_{1,i})\Big]+n\delta_{n,\epsilon} \tag{108d}$$

$$n(R_0+R_1+R_2) \le \sum_{i=1}^{n}\Big[I(U_i;Y_{1,i}|W_i)+I(V_i;Y_{2,i}|W_i,U_i)+I(W_i;Y_{1,i})\Big]+3n\delta_{n,\epsilon} \tag{108e}$$

$$n(R_0+R_1+R_2) \le \sum_{i=1}^{n}\Big[I(U_i;Y_{1,i}|W_i)+I(V_i;Y_{2,i}|W_i,U_i)+I(W_i;Y_{2,i})\Big]+3n\delta_{n,\epsilon} \tag{108f}$$

For the converse we prove the reverse inclusion, i.e., $\mathcal{R}_{\mathsf{O}}(L_1, L_2) \subseteq \mathcal{C}_{\mathsf{SD}}(L_1, L_2)$. First we remove the restriction from $\mathcal{R}_{\mathsf{O}}(L_1, L_2)$ that $X - (U, V) - W$ forms a Markov chain; this can only enlarge the region. Fix a PMF $Q_{W,U,V,X} \in \mathcal{P}(\mathcal{W} \times \mathcal{U} \times \mathcal{V} \times \mathcal{X})$, which induces a joint distribution $Q_{W,U,V,X}\mathbb{1}_{\{Y_1 = y_1(X)\}}W_{Y_2|X}$. Each of the bounds defining $\mathcal{R}_{\mathsf{O}}(L_1, L_2)$ and $\mathcal{C}_{\mathsf{SD}}(L_1, L_2)$ are taken with respect to $Q_{W,U,V,X}\mathbb{1}_{\{Y_1 = y_1(X)\}}W_{Y_2|X}$.

We start by noting that (29a) and (27a) are the same. Next, the RHS of (27b) is upper bounded by the RHS of (29b) since

$$\begin{aligned} R_1 &\le I(U; Y_1|W, V) - I(U; Y_2|W, V) + L_1 \\ &= 7H(Y_1|W, V) - H(Y_1|W, V, U) - I(U; Y_2|W, V) + L_1 \\ &\overset{(a)}{\le} H(Y_1|W, V) - I(Y_1; Y_2|W, V, U) - I(U; Y_2|W, V) + L_1 \\ &= H(Y_1|W, V) - I(U, Y_1; Y_2|W, V) + L_1 \\ &\overset{(b)}{\le} H(Y_1|W, V, Y_2) + L_1 \end{aligned} \tag{111}$$

where (a) is by the non-negativity of entropy and (b) is because conditioning cannot increase entropy.

For (27d) we clearly have

$$\begin{aligned} R_0 + R_1 &\le I(U; Y_1|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \\ &\le H(Y_1|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \end{aligned} \tag{112}$$

which coincides with (29c). Furthermore, inequalities (29d) and (29e) are the same as (27f) and (27g), respectively. For the sum of rates, the RHS of (29f) upper bounds that of (27h) because

$$I(U; Y_1|W, V) \le H(Y_1|W, V). \tag{113}$$

Removing the other bounds from (27) can only enlarge $\mathcal{R}_{\mathsf{O}}(L_1, L_2)$, which shows its inclusion in $\mathcal{C}_{\mathsf{SD}}(L_1, L_2)$. This characterizes $\mathcal{C}_{\mathsf{SD}}(L_1, L_2)$ as the (L_1, L_2)-leakage-capacity region of the SD-BC.

VIII. Summary and Concluding Remarks

We considered the BC with privacy leakage constraints. Under this model, all four scenarios concerning secrecy (i.e., when both, either or neither of the private messages are secret) are special cases by appropriate choices for the leakage thresholds. Inner and outer bounds on the leakage-capacity region were derived and shown to be tight for SD and PD BCs, as well as for BCs with a degraded message set. The coding strategy that achieved the inner bound is based on a Marton-like codebook construction with a common message supplemented by an extra layer of binning. Splitting each private message into a public and a private part, a public message that comprises the public parts and the common message was constructed. To correlate the codewords for the private parts, we used the likelihood encoder. Its simple structure enabled a rigorous analysis of performance for the proposed scheme. Theorem 1 fixes a weakness of previous work by letting the eavesdropper know the codebook. The main tool needed was the likelihood encoder (Lemma 3).

Our results include various past works as special cases. Large leakage thresholds reduce our inner and outer bounds to Marton's inner bound with a common message [21] and the UVW-outer bound [22], respectively. The leakage-capacity region of the SD-BC without a common message recovers the capacity regions where both [5], either [6], [27], or neither [21] private message is secret. The result for the BC with a degraded message set and a privacy leakage constraint captures the capacity regions for the BC with confidential messages [3] and the BC with a degraded message set (without secrecy) [28]. Furthermore, we derived conditions on the allowed leakage values that differentiates whether a further increase of each leakage threshold induces a larger inner bound or not. The conditions effectively let one (numerically) calculate privacy leakage threshold values above which the inner bound saturates. This idea was visualized by means of a BW-BC example that showed the transition of the leakage-capacity region from secrecy-capacity regions for different scenarios to the capacity region without secrecy.

Appendix A
Proof of Corollary 8

The region $\mathcal{C}_{\mathsf{D}}(L_1, L_2)$ is obtained from $\mathcal{C}^0_{\mathsf{SD}}(L_1, L_2)$ by setting $W = 0$ and $V = Y_2$, which implies that $\mathcal{C}_{\mathsf{D}}(L_1, L_2) \subseteq \mathcal{C}^0_{\mathsf{SD}}(L_1, L_2)$. For the converse, the RHS of (32a) is upper bounded by

$$R_1 \le H(Y_1|W, V, Y_2) + L_1 \le H(Y_1|Y_2) + L_1. \tag{114}$$

For (32c), we have

$$\begin{aligned} &I(V; Y_2|W) - I(V; Y_1|W) + L_2 \\ &\quad\le I(V; Y_1, Y_2|W) - I(V; Y_1|W) + L_2 \\ &\quad= I(V; Y_2|W, Y_1) + L_2 \\ &\quad\le H(Y_2|Y_1) + L_2. \end{aligned} \tag{115}$$

The RHSs of (32b) and (32d) are clearly upper bounded as

$$H(Y_j|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \le H(Y_j), \quad j = 1, 2. \tag{116}$$

Finally, (42c) is implied by (32e) since

$$\begin{aligned} &R_1 + R_2 \\ &\le H(Y_1|W, V) + I(V; Y_2|W) + \min\Big\{I(W; Y_1), I(W; Y_2)\Big\} \\ &\le H(Y_1|W, V) + I(W, V; Y_2) \\ &\le H(Y_1, Y_2|W, V) + I(W, V; Y_1, Y_2) \\ &= H(Y_1, Y_2). \end{aligned} \tag{117}$$

Appendix B
Error Probability Analysis for the Proof of Theorem 1

By the symmetry of the codebook construction with respect to $(M_p, M_{11}, W_1, M_{22}, W_2)$ and due to their uniformity, we may assume that $(M_p, M_{11}, W_1, M_{22}, W_2) = (1, 1, 1, 1, 1)$.

1) Encoding Errors: Fix any $\delta' \in (0, \delta)$. An encoding error event is described as given in (118) at the top of the next page.

2) Decoding Errors: To account for decoding errors, define (119) from the top of the next page, where $j = 1, 2$.

For any event $\mathcal{A}$ from the σ-algebra over which $\mathbb{P}$ is defined, denote $\mathbb{P}_1 = \mathbb{P}\big(\mathcal{A}\big|M_p = 1, M_{11} = 1, W_1 = 1, M_{22} = 1, W_2 = 1\big)$. By the union bound, the expected error probability

$$\mathcal{E} = \left\{\left(\mathbf{U}_0(1), \mathbf{U}_1(1,1,1,I_1), \mathbf{U}_2(1,1,1,I_2)\right) \notin \mathcal{T}_{\delta'}^n(Q_{U_0,U_1,U_2})\right\} \tag{118}$$

$$\mathcal{D}_0 = \left\{\left(\mathbf{U}_0(1), \mathbf{U}_1(1,1,1,I_1), \mathbf{U}_2(1,1,1,I_2), \mathbf{Y}_1, \mathbf{Y}_2\right) \in \mathcal{T}_{\delta}^n(Q_{U_0,U_1,U_2,Y_1,Y_2})\right\} \tag{119a}$$

$$\mathcal{D}_0^{(j)}(m_p) = \left\{\left(\mathbf{U}_0(m_p), \mathbf{Y}_j\right) \in \mathcal{T}_{\delta}^n(Q_{U_0,Y_j})\right\} \tag{119b}$$

$$\mathcal{D}_1^{(j)}(m_{jj}, w_j, i_j) = \left\{\left(\mathbf{U}_0(1), \mathbf{U}_j(1, m_{jj}, w_j, i_j), \mathbf{Y}_j\right) \in \mathcal{T}_{\delta}^n(Q_{U_0,U_j,Y_j})\right\} \tag{119c}$$

$$\begin{aligned}
\mathbb{E}P_{\mathsf{e}}(\mathsf{C}_n) &\le \mathbb{P}_1\left(\mathcal{E} \cup \mathcal{D}_0^c \cup \bigcup_{j=1,2} \mathcal{D}_0^{(j)}(1)^c \cup \left\{\bigcup_{\tilde{m}_p \neq 1} \mathcal{D}_0^{(j)}(\tilde{m}_p)\right\} \cup \mathcal{D}_1^{(j)}(1,1,I_j)^c \cup \left\{\bigcup_{(\tilde{m}_{jj},\tilde{w}_j)\neq(1,1)} \mathcal{D}_0^{(j)}(\tilde{m}_{jj}, \tilde{w}_j, I_j)\right\}\right) \\
&\le \underbrace{\mathbb{P}_1\left(\mathcal{E}\right)}_{P_0^{[1]}} + \underbrace{\mathbb{P}_1\left(\mathcal{D}_0^c \cap \mathcal{E}^c\right)}_{P_0^{[2]}} + \sum_{j=1,2} \Bigg[\underbrace{\mathbb{P}_1\left(\mathcal{D}_0^{(j)}(1)^c \cap \mathcal{D}_0\right)}_{P_j^{[0]}} + \underbrace{\mathbb{P}_1\left(\mathcal{D}_1^{(j)}(1,1,I_j)^c \cap \mathcal{D}_0\right)}_{P_j^{[1]}} + \underbrace{\mathbb{P}_1\left(\bigcup_{\tilde{m}_p \neq 1} \mathcal{D}_0^{(j)}(\tilde{m}_p)\right)}_{P_J^{[2]}} \\
&\quad + \underbrace{\mathbb{P}_1\left(\bigcup_{\substack{(\tilde{m}_{jj},\tilde{w}_j)\neq(1,1),\\ \tilde{i}_j \in \mathcal{I}_j}} \mathcal{D}_1^{(j)}(\tilde{m}_{jj}, \tilde{w}_j, \tilde{i}_j)\right)}_{P_j^{[3]}} \Bigg]
\end{aligned} \tag{120}$$

is bounded as in (120) given at the top of this page.[7] Note that with respect to the notation in (120) $P_0^{[1]}$ is the probability of an encoding error, while $P_j^{[k]}$, for $k \in [0:3]$, are the decoding errors of Decoder j. We proceed with the following steps:

1) By [32, Th. 3], we have $P_0^{[1]} \to 0$ as $n \to \infty$ if

$$R_1' + R_2' > I(U_1; U_2|U_0). \tag{121}$$

2) The Conditional Typicality Lemma [20, Section 2.5] implies that $P_0^{[2]} \to 0$ as n grows. More precisely, there exists a function $\beta(n, \delta, \delta')$ with $\lim_{n\to\infty} \beta(n, \delta, \delta') = 0$ for any $0 < \delta' < \delta$, such that $P_0^{[2]} \le \beta(n, \delta, \delta')$. Furthermore, replacing δ and δ' with properly chosen decaying sequences $\{\delta_n\}_{n\in\mathbb{N}}$ and $\{\delta_n'\}_{n\in\mathbb{N}}$, respectively, and setting $\beta_n \triangleq \beta(n, \delta_n, \delta_n')$, we have $\lim_{n\to\infty} \beta_n = 0$.
3) The definitions in (VII-E) clearly give $P_j^{[0]} = P_j^{[1]} = 0$, for $j = 1, 2$ and every $n \in \mathbb{N}$. This is since $\mathcal{D}_0^{(j)}(1)^c \cap \mathcal{D}_0 = \mathcal{D}_1^{(j)}(1,1,I_j)^c \cap \mathcal{D}_0 = \emptyset$, for $j = 1, 2$.
4) For $P_j^{[2]}$, $j = 1, 2$, we have

$$\begin{aligned}
P_j^{[2]} &\overset{(a)}{\le} \sum_{\tilde{m}_p \neq 1} 2^{-n\left(I(U_0;Y_j) - \tau_j^{[2]}(\delta)\right)} \\
&\le 2^{nR_p} 2^{-n\left(I(U_0;Y_j) - \tau_j^{[2]}(\delta)\right)} \\
&= 2^{n\left(R_p - I(U_0;Y_j) + \tau_j^{[2]}(\delta)\right)}
\end{aligned} \tag{122}$$

where (a) follows since $\mathbf{U}_0(\tilde{m}_p)$ is independent of $\mathbf{Y}_j$, for any $\tilde{m}_p \neq 1$. Thus, for $P_j^{[2]}$ to vanish as $n \to \infty$, we take:

$$R_p < I(U_0; Y_j) - \tau_j^{[2]}(\delta), \quad j = 1, 2 \tag{123}$$

where $\tau_j^{[2]}(\delta) \to 0$ as $\delta \to 0$.

5) For $P_j^{[3]}$, $j = 1, 2$, we have

$$\begin{aligned}
P_j^{[3]} &\overset{(a)}{\le} \sum_{\substack{(\tilde{m}_{jj},\tilde{w}_j)\neq (1,1),\\ \tilde{i}_j \in \mathcal{I}_j}} 2^{-n\left(I(U_j;Y_j|U_0) - \tau_j^{[3]}(\delta)\right)} \\
&\le 2^{n(R_{jj} + R_j' + \tilde{R}_j)} 2^{-n\left(I(U_j;Y_j|U_0) - \tau_j^{[3]}(\delta)\right)} \\
&= 2^{n\left(R_{jj} + R_j' + \tilde{R}_j - I(U_j;Y_j|U_0) + \tau_j^{[3]}(\delta)\right)}
\end{aligned} \tag{124}$$

where (a) follows since $\mathbf{U}_j(1, \tilde{m}_{jj}, \tilde{w}_j, \tilde{i}_j)$ is independent of $\mathbf{Y}_j$, for any $(\tilde{m}_{jj}, \tilde{w}_j) \neq (1, 1)$ and $\tilde{i}_j \in \mathcal{I}_j$, while both of them are drawn conditioned on $\mathbf{U}_0(1)$. We have $P_j^{[3]} \to 0$ as $n \to \infty$ if

$$R_{jj} + R_j' + \tilde{R}_j < I(U_j; Y_j|U_0) - \tau_j^{[3]}(\delta), \quad j = 1, 2 \tag{125}$$

where, as before, $\tau_j^{[3]}(\delta) \to 0$ as $\delta \to 0$.

Summarizing the above results, while substituting $R_p = R_0 + R_{10} + R_{20}$ and setting

$$\tau_\delta \triangleq \max\left\{\tau_j^{[k]}(\delta)\right\}_{\substack{j=1,2,\\ k=2,3}} \tag{126}$$

we find that

$$\mathbb{E}P_{\mathsf{e}}(\mathsf{C}_n) \le \eta(n, \delta, \delta'), \quad \forall n \in \mathbb{N} \tag{127}$$

[7] As in Section VII-B, we abuse notation in writing $\mathbb{E}P_{\mathsf{e}}(\mathsf{C}_n)$ because P_e is actually a function of the code c_n rather than the codebook $\mathcal{C}_n$. We favor this notation for its simplicity and remind the reader that $\mathcal{C}_n$ uniquely defines c_n.

where $\lim_{n\to\infty}\eta(n,\delta,\delta')=0$ for all $0<\delta'<\delta$, if the conditions in (68) are met. As mentioned before, if we replace δ and δ' with properly chosen sequences $\{\delta_n\}_{n\in\mathbb{N}}$ and $\{\delta'_n\}_{n\in\mathbb{N}}$, respectively, that decay sufficiently slowly to zero and set $\eta_n \triangleq \eta(n,\delta_n,\delta'_n)$, we have $\lim_{n\to\infty}\eta_n=0$.

A. Leakage Associated Errors

This subsection shows how (71) ensures $\mathbb{E}\lambda^{(1)}_{m_{11}}(\mathsf{C}_n)\to 0$ as $n\to\infty$, for any $m_{11}\in\mathcal{M}_{11}$. As before, by the symmetry of the underlying random code with respect to the messages, we have

$$\mathbb{E}\lambda^{(1)}_{m_{11}}(\mathsf{C}_n)=\mathbb{E}\lambda^{(1)}_{1}(\mathsf{C}_n),\quad \forall m_{11}\in\mathcal{M}_{11} \tag{128}$$

and we may further assume that $(M_p,W_1,M_{22},W_2)=(1,1,1,1)$. By arguments similar to those presented in the encoding and decoding error probability analysis, one can verify that (71) implies the existence of a function $\kappa(n,\delta)$ with $\lim_{n\to\infty}\kappa(n,\delta)=0$ for any $\delta>0$, such that $\mathbb{E}\lambda^{(1)}_{1}(\mathsf{C}_n)\le\kappa(n,\delta)$. Furthermore, replacing δ with a sequence $\{\delta_n\}_{n\in\mathbb{N}}$ that decays sufficiently slow to zero as n grows and setting $\kappa_n\triangleq\kappa(n,\delta_n)$, we have $\kappa_n\to 0$ as $n\to\infty$.

This essentially follows by the law of large numbers and the Conditional Typicality Lemma that ensure the joint typicality of the transmitted sequences and the outputs. If $\tilde{w}_1$ is incorrect but I_1 is the true index chosen by the likelihood encoder, $\mathbf{U}_1(1,1,\tilde{w}_1,I_1)$ is conditionally independent $\mathbf{Y}_2$ given $(\mathbf{U}_0(1),\mathbf{U}_2(1,1,1,I_2))$. The correlation between $\mathbf{U}_0(1)$, $\mathbf{U}_1(1,1,\tilde{w}_1,I_1)$ and $\mathbf{U}_2(1,1,1,I_2)$ is a consequence of the likelihood encoder's operation. Since the search space in this case is of size $2^{n\tilde{R}_1}$, taking

$$\tilde{R}_1<I(U_1;Y_2|U_0,U_2)-\xi(\delta) \tag{129a}$$

where $\xi(\delta)\to 0$ as $\delta\to 0$, results in a vanishing probability of the event that this u_1-sequence satisfies the typicality test from (69).

Furthermore, if $\tilde{w}_1$ and $\tilde{i}_1$ are both incorrect, we have that $\mathbf{U}_1(1,1,\tilde{w}_1,\tilde{i}_1)$ is conditionally independent of $\big(\mathbf{U}_2(1,1,1,I_2),\mathbf{Y}_2\big)$ given $\mathbf{U}_0(1)$. The search space is now of size $2^{n(\tilde{R}_1+R'_1)}$, and therefore, taking

$$\tilde{R}_1+R'_1<I(U_1;U_2,Y_2|U_0)-\xi(\delta) \tag{129b}$$

implies a vanishing probability of this second error event.

Finally, note that the error event where $W_1=1$ is correct but $\tilde{i}_1$ is wrong has arbitrarily small probability if $R'_1<I(U_1;U_2,Y_2|U_0)-\xi(\delta)$ (the structure of the mutual information term is the same as in (129b) because an incorrect $\tilde{i}_1$ produces the same statistical relations as an incorrect pair $(\tilde{w}_1,\tilde{i}_1)$). Evidently, the latter constraint is redundant due to (129b).

Appendix C
Proof of Lemma 4

Recall that $P_1^{(\mathsf{Leak})}(\mathcal{C}_n)$ denotes the error probability in decoding (W_1,I_1) from $(M_p,M_{11},M_{22},W_2,I_2,\mathbf{Y}_2)$ by means of the typicality test from (69) with respect to the fixed code $\mathcal{C}_n\in\mathfrak{C}_n$. The analysis in Appendix B shows that as long as (71) holds, we have

$$\mathbb{E}P_1^{(\mathsf{Leak})}(\mathsf{C}_n)\le\kappa(n,\delta,\delta') \tag{130}$$

where $\lim_{n\to\infty}\kappa(n,\delta,\delta')=0$ for all $0<\delta'<\delta$. As a consequence, we have

$$\begin{aligned}
&H(W_1,I_1|M_p,M_{11},M_{22},W_2,I_2,\mathbf{Y}_2,\mathsf{C}_n)\\
&\quad\le H(W_1,I_1|M_p,M_{11},M_{22},W_2,I_2,\mathbf{Y}_2)\\
&\quad\overset{(a)}{\le} 1+n\cdot\kappa(n,\delta,\delta')n(\tilde{R}_1+R'_1)
\end{aligned} \tag{131}$$

where (a) is because conditioning cannot increase entropy, while (b) uses Fano's inequality and (130). Setting $\zeta_1(n,\delta,\delta')\triangleq\frac{1}{n}+\kappa(n,\delta,\delta')(\tilde{R}_1+R'_1)$ completes the proof.

Appendix D
Proof of Lemma 5

Define the indicator function $E=\mathbb{1}_{\mathcal{A}}$, where

$$\mathcal{A}=\Big\{(\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2,\mathbf{Y}_2)\notin\mathcal{T}^n_\delta(Q_{U_0,U_1,U_2,Y_2})\Big\} \tag{132}$$

and note that $\mathbb{P}\big(E=1\big)\le\mathbb{P}_1\big(\mathcal{E}\big)+\mathbb{P}_1\big(\mathcal{D}_0^c\cap\mathcal{E}^c\big)$, where $\mathcal{E}$ and $\mathcal{D}_0$ are defined in (118) and (119a), respectively, from Appendix B. The analysis in Appendix B shows the existence of a function $\tilde{\beta}(n,\delta,\delta')$, such that

$$\mathbb{P}\big(E=1\big)=\mathbb{P}_1\big(\mathcal{E}\big)+\mathbb{P}_1\big(\mathcal{D}_0^c\cap\mathcal{E}^c\big)\le\tilde{\beta}(n,\delta,\delta') \tag{133}$$

where $0<\delta'<\delta$ and $\lim_{n\to\infty}\tilde{\beta}(n,\delta,\delta')=0$ for all such values of δ and δ'. Furthermore, $\lim_{n\to\infty}\tilde{\beta}(n,\delta_n,\delta'_n)=0$ for sequences $\{\delta_n\}_{n\in\mathbb{N}}$ and $\{\delta'_n\}_{n\in\mathbb{N}}$ that decay sufficiently slow to zero with n.

We now expand the mutual information term from the LHS of (73) as follows

$$\begin{aligned}
&I(\mathbf{U}_1;\mathbf{Y}_2|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n)\\
&\quad\le I(\mathbf{U}_1,E;\mathbf{Y}_2|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n)\\
&\quad= I(E;\mathbf{U}_2|\mathbf{U}_0,,\mathbf{U}_2,\mathsf{C}_n)+I(\mathbf{U}_1;\mathbf{Y}_2|\mathbf{U}_0,\mathbf{U}_2,E,\mathsf{C}_n)\\
&\quad\overset{(a)}{\le} 1+\sum_{j=0}^{1}\mathbb{P}\big(E=j\big)I(\mathbf{U}_1;\mathbf{Y}_2|\mathbf{U}_0,\mathbf{U}_2,E=j,\mathsf{C}_n).
\end{aligned} \tag{134}$$

where (a) is because E is binary and the entropy function is non-negative. Note that

$$\begin{aligned}
&\mathbb{P}\big(E=1\big)I(\mathbf{U}_1;\mathbf{Y}_2|\mathbf{U}_0,\mathbf{U}_2,E=1,\mathsf{C}_n)\\
&\quad\le\mathbb{P}\big(E=1\big)H(\mathbf{Y}_2|E=1,\mathsf{C}_n)\\
&\quad\le\tilde{\beta}(n,\delta,\delta')\cdot n\log|\mathcal{Y}_2|
\end{aligned} \tag{135}$$

where (a) uses (133).

For the mutual information term conditioned on $E=0$, we first have

$$\begin{aligned}
&H(\mathbf{Y}_2|\mathbf{U}_0,\mathbf{U}_2,E=0,\mathsf{C}_n)\\
&\le H(\mathbf{Y}_2|\mathbf{U}_0,\mathbf{U}_2,E=0)\\
&=\sum_{\substack{(\mathbf{u}_0,\mathbf{u}_2)\\ \in\mathcal{T}^n_\delta(Q_{U_0,U_2})}} Q_{\mathbf{U}_0,\mathbf{U}_2|E}(\mathbf{u}_0,\mathbf{u}_2|0)H(\mathbf{Y}_2|\mathbf{U}_0=\mathbf{u}_0,\mathbf{U}_2=\mathbf{u}_2,E=0)\\
&\le nH(Y_2|U_0,U_2)
\end{aligned} \tag{136}$$

where the last inequality is because for every $(\mathbf{u}_0,\mathbf{u}_2)\in\mathcal{T}^n_\delta(Q_{U_0,U_2})$ the support of the conditional PMF $P_{\mathbf{Y}_2|\mathbf{U}_0=\mathbf{u}_0,\mathbf{U}_2=\mathbf{u}_2,E=0}$ is upper bounded by the size of the conditional typical set $\mathcal{T}^n_\delta(Q_{U_0,U_2,Y_2}|\mathbf{u}_0,\mathbf{u}_2)$, which is

$$
\begin{aligned}
&H(\mathbf{Y}_1|\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2,E=0,\mathsf{C}_n)\\
&\overset{(a)}{=} H(\mathbf{Y}_1|\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2,E=0)\\
&= \sum_{\substack{(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2)\\ \in\mathcal{T}_\delta^n(Q_{U_0,U_1,U_2})}} P_{\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2|E}(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2|0)H(\mathbf{Y}_2|\mathbf{U}_0=\mathbf{u}_0,\mathbf{U}_1=\mathbf{u}_1,\mathbf{U}_2=\mathbf{u}_2,E=0)\\
&= \sum_{\substack{(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2)\\ \in\mathcal{T}_\delta^n(Q_{U_0,U_1,U_2})}} P_{\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2|E}(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2|0)\sum_{i=1}^n H(Y_{2,i}|\mathbf{U}_0=\mathbf{u}_0,\mathbf{U}_1=\mathbf{u}_1,\mathbf{U}_2=\mathbf{u}_2,Y_2^{i-1},E=0)\\
&\overset{(b)}{=} \sum_{\substack{(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2)\\ \in\mathcal{T}_\delta^n(Q_{U_0,U_1,U_2})}} P_{\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2|E}(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2|0)\sum_{i=1}^n H(Y_{2,i}|U_{0,i}=u_{0,i},U_{1,i}=u_{1,i},U_{2,i}=u_{2,i})\\
&= \sum_{\substack{(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2)\\ \in\mathcal{T}_\delta^n(Q_{U_0,U_1,U_2})}} P_{\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2|E}(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2|0)\sum_{\substack{(u_0,u_1,u_2)\\ \in\mathcal{U}_0\times\mathcal{U}_1\times\mathcal{U}_2}} \nu_{\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2}(u_0,u_1,u_2)H(Y_2|U_0=u_0,U_1=u_1,U_2=u_2)\\
&\overset{(c)}{\geq} n\cdot \sum_{\substack{(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2)\\ \in\mathcal{T}_\delta^n(Q_{U_0,U_1,U_2})}} P_{\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2|E}(\mathbf{u}_0,\mathbf{u}_1,\mathbf{u}_2|0)(1-\delta)H(Y_2|U_0,U_1,U_2)\\
&= n(1-\delta)H(Y_2|U_0,U_1,U_2)
\end{aligned}
\tag{137}
$$

$$
\tilde{\mathcal{D}} \triangleq \left\{\exists\ (\tilde{m}_p,\tilde{m}_{22},\tilde{w}_2,\tilde{i}_2)\neq\ (M_p,M_{22},W_2,I_2),\quad \mathbf{U}_0(\tilde{m}_p)=\mathbf{U}_0\ \text{ and }\ \mathbf{U}_2(\tilde{m}_p,\tilde{m}_{22},\tilde{w}_2,\tilde{i}_2)=\mathbf{U}_2\right\} \tag{143}
$$

upper bounded by $2^{nH(Y_2|U_0,U_2)(1+\delta)}$. This step also relies on the entropy being maximized by the uniform distribution and the logarithm being a monotonically increasing function.

For the other (subtracted) entropy term, we have (137) given at the top of this page, where (a) is because $\mathbf{Y}_2 - (\mathbf{U}_0,\mathbf{U}_1,\mathbf{U}_2) - \mathsf{C}_n$ forms a Markov chain, (b) follows since given $(U_{0,i},U_{1,i},U_{2,i})$, $Y_{2,i}$ is independent of all other random variables, while (c) is by the definition of letter-typical sequences from (3).

Inserting (135), (136) and (137) into (134) gives

$$
I(\mathbf{U}_1;\mathbf{Y}_2|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n) \leq nI(U_1;Y_2|U_0,U_2)+n\zeta_2(n,\delta,\delta') \tag{138}
$$

where

$$
\zeta_2(n,\delta,\delta') = \frac{1}{n}+\delta H(Y_2|U_0,U_1,U_2)+\tilde{\beta}(n,\delta,\delta') \tag{139}
$$

as needed.

Appendix E
Proof of Lemma 6

Rewriting the mutual information term of interest as a difference of entropies, we have

$$
\begin{aligned}
&I(\mathbf{U}_1;M_{22},W_2,I_2|M_p,\mathsf{C}_n)\\
&\quad = H(\mathbf{U}_1|M_p,\mathsf{C}_n)-H(\mathbf{U}_1|M_p,M_{22},W_2,I_2,\mathsf{C}_n).
\end{aligned}
\tag{140}
$$

Since $\mathbf{U}_0$ is defined by (M_p,C_n), we clearly have,

$$
H(\mathbf{U}_1|M_p,\mathsf{C}_n) \leq H(\mathbf{U}_1|\mathbf{U}_0,\mathsf{C}_n). \tag{141}
$$

Next, since $(M_p,M_{22},W_2,I_2,\mathsf{C}_n)$ determines both $\mathbf{U}_0$ and $\mathbf{U}_2$, we write the subtracted entropy term as

$$
\begin{aligned}
&H(\mathbf{U}_1|M_p,M_{22},W_2,I_2,\mathsf{C}_n)\\
&= H(\mathbf{U}_1|\mathbf{U}_0,\mathbf{U}_2,M_p,M_{22},W_2,I_2,\mathsf{C}_n)\\
&= H(\mathbf{U}_1|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n)-I(\mathbf{U}_1;M_p,M_{22},W_2,I_2|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n)\\
&\geq H(\mathbf{U}_1|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n)-H(M_p,M_{22},W_2,I_2|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n).
\end{aligned}
\tag{142}
$$

We now upper bound $H(M_p,M_{22},W_2,I_2|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n)$ by a vanishing term times the blocklength n. Let $F=\mathbb{1}_{\tilde{\mathcal{D}}}$ be the indicator function of the event $\tilde{\mathcal{D}}$ defined in (143) at the top of this page. Standard error probability analysis of random codes shows that

$$
\mathbb{P}\big(F=1\big)=\mathbb{P}\big(\tilde{\mathcal{D}}\big) \leq 2^{n\left(R_p+R_{22}+\tilde{R}_2+R_2'-H(U_0,U_2)+\tilde{\alpha}(\delta)\right)} \tag{144}
$$

where $\tilde{\alpha}(\delta)\to 0$ as $\delta\to 0$. Consequently, taking

$$
R_p+R_{22}+\tilde{R}_2+R_2' < H(U_0,U_2)-\tilde{\alpha}(\delta) \tag{145}
$$

results in $\mathbb{P}\big(F=1\big)\leq\tilde{\kappa}(n,\delta)$ with $\lim_{n\to\infty}\tilde{\kappa}(n,\delta)=0$ for every $\delta>0$. Next, note that (145) holds on account of (68c) and (68e) (adding (68c) and (68e) results in a tighter bound on the same rates) and consider the following:

$$
\begin{aligned}
&H(M_p,M_{22},W_2,I_2|\mathbf{U}_0,\mathbf{U}_2,\mathsf{C}_n)\\
&\quad\leq H(M_p,M_{22},W_2,I_2|\mathbf{U}_0,\mathbf{U}_2)\\
&\quad\overset{(a)}{\leq} 1+H(M_p,M_{22},W_2,I_2|\mathbf{U}_0,\mathbf{U}_2,F)\\
&\quad= 1+\mathbb{P}\big(F=0\big)H(M_p,M_{22},W_2,I_2|\mathbf{U}_0,\mathbf{U}_2,F=0)\\
&\qquad+\mathbb{P}\big(F=1\big)H(M_p,M_{22},W_2,I_2|\mathbf{U}_0,\mathbf{U}_2,F=1)
\end{aligned}
$$

$$\begin{aligned}
&\overset{(b)}{\leq} 1 + H(M_p, M_{22}, W_2, I_2|\mathbf{U}_0, \mathbf{U}_2, F=0) \\
&\quad + \mathbb{P}\big(F=1\big)\cdot n(R_p + R_{22} + \tilde{R}_2 + R_2') \\
&\overset{(c)}{\leq} 2\Big[1 + n\cdot\tilde{\kappa}(n,\delta)(R_p + R_{22} + \tilde{R}_2 + R_2')\Big] \\
&= n\tilde{\zeta}_3^{(1)}(n,\delta) \qquad (146)
\end{aligned}$$

where

$$\tilde{\zeta}_3^{(1)}(n,\delta) \triangleq \frac{1}{n} + \tilde{\kappa}(n,\delta)(R_p + R_{22} + \tilde{R}_2 + R_2'). \qquad (147)$$

In the above derivation (a) follows because the uniform distribution maximizes entropy and since F is binary, (b) upper bounds the first entropy term by the logarithm of the support size, while (c) uses Fano's inequality.

Inserting (141), (142) and (146) into (140) gives

$$I(\mathbf{U}_1; M_{22}, W_2, I_2|M_p, \mathsf{C}_n) \leq I(\mathbf{U}_1; \mathbf{U}_2|\mathbf{U}_0, \mathsf{C}_n) + n\tilde{\zeta}_3^{(1)}(n,\delta). \qquad (148)$$

To complete the proof, it suffices to show that there exists a function $\tilde{\zeta}_3^{(2)}(n,\delta,\delta')$ that satisfies the same properties as $\zeta_3(n,\delta,\delta')$ from the statement of Lemma 6 for which

$$I(\mathbf{U}_1; \mathbf{U}_2|\mathbf{U}_0, \mathsf{C}_n) \leq nI(U_1; U_2|U_0) + n\tilde{\zeta}_3^{(2)}(n,\delta,\delta'). \qquad (149)$$

This can be established by arguments similar to those presented in the proof of Lemma 5 and we therefore omit the details. Combining (148) with (149) and setting $\zeta_3(n,\delta,\delta') \triangleq \tilde{\zeta}_3^{(1)}(n,\delta) + \tilde{\zeta}_3^{(2)}(n,\delta,\delta')$ completes the proof.

ACKNOWLEDGEMENTS

The authors would like to thank the Associate Editor and the anonymous reviewers for helping to improve the presentation of this paper. We also kindly thank Ido B. Gattegno for his work on the FME-IT software [46] that assisted us with technical details of proofs.

REFERENCES

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.
[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.
[4] 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.
[5] Y. Zhao, P. Xu, Y. Zhao, W. Wei, and Y. Tang, "Secret communications over semi-deterministic broadcast channels," in *Proc. 4th Int. Conf. Commun. Netw. China (CHINACOM)*, Xi'an, China, Aug. 2009, pp. 1–4.
[6] W. Kang and N. Liu, "The secrecy capacity of the semi-deterministic broadcast channel," in *Proc. Int. Symp. Inf. Theory*, Seoul, South Korea, Jun./Jul. 2009, pp. 2767–2771.
[7] Z. Goldfeld, G. Kramer, H. H. Permuter, and P. Cuff, "Strong secrecy for cooperative broadcast channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 469–495, Jan. 2017.
[8] E. Ekrem and S. Ulukus, "Secrecy in cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 137–155, Jan. 2011.
[9] R. Liu and H. V. Poor, "Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1235–1249, Mar. 2009.
[10] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.
[11] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4215–4227, Sep. 2010.
[12] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.
[13] E. Ekrem and S. Ulukus, "The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2083–2114, Apr. 2011.
[14] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.
[15] E. Ekrem and S. Ulukus, "Secrecy capacity of a class of broadcast channels with an eavesdropper," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, pp. 1–29, Mar. 2009.
[16] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "Secrecy capacity region of Gaussian broadcast channel," in *Proc. 43rd Annu. Conf. Inf. Sci. Syst. (CISS)*, Baltimore, MD, USA, Mar. 2009, pp. 152–157.
[17] M. Benammar and P. Piantanida, "Secrecy capacity region of some classes of wiretap broadcast channels," *IEEE Trans Inf. Theory*, vol. 61, no. 10, pp. 5564–5582, Oct. 2015.
[18] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
[19] A. El Gamal and E. van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 120–122, Jan. 1981.
[20] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[21] S. I. Gel'fand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Inf. Transmiss.*, vol. 16, no. 1, pp. 17–25, Jan./Mar. 1980.
[22] C. Nair, "A note on outer bounds for broadcast channel," presented at the Int. Zurich Seminar, Jan. 2011. [Online]. Available: http://arXiv.org/abs/1101.0640.
[23] Y. Liang, G. Kramer, and S. Shamai (Shitz), "Capacity outer bounds for broadcast channels," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Porto, Portugal, May 2008, pp. 2–4.
[24] Y. Liang, "Multiuser communications with relaying and user cooperation," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Illinois Urbana–Champaign, Champaign, IL, USA, 2005.
[25] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 350–355, Jan. 2007.
[26] C. Nair and V. W. Zizhou, "On the inner and outer bounds for 2-receiver discrete memoryless broadcast channels," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, USA, Jan./Feb. 2008, pp. 226–229.
[27] Z. Goldfeld, G. Kramer, and H. H. Permuter, "Cooperative broadcast channels with a secret message," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Hong Kong, Jun. 2015, pp. 1342–1346.
[28] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 60–64, Jan. 1977.
[29] E. van der Meulen, "Random coding theorems for the general discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 21, no. 2, pp. 180–190, Mar. 1975.
[30] S. I. Gel'fand, "Capacity of one broadcast channel," *Probl. Peredachi Inf.*, vol. 13, no. 3, pp. 106–108, Jul./Sep. 1977.
[31] J. L. Massey, "Applied digital information theory," ETH Zurich, Zürich, Switzerland, Tech. Rep., pp. 1980–1998.
[32] M. H. Yassaee, "One-shot achievability via fidelity," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Hong Kong, Jun. 2015, pp. 301–305.
[33] 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.
[34] E. C. Song, P. Cuff, and H. V. Poor, "The likelihood encoder for lossy compression," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1836–1849, Apr. 2016.
[35] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Wiretap channels with random states non-causally available at the encoder," *IEEE Trans. Inf. Theory*, to be published. [Online]. Available: https://arxiv.org/pdf/1608.06057
[36] M. H. Yassaee, M. R. Aref, and A. Gohari, "A technique for deriving one-shot achievability results in network information theory," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 1287–1291. [Online]. Available: http://arxiv.org/abs/1303.0696.

[37] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.
[38] Y. Liang, G. Kramer, and H. V. Poor, "On the equivalence of two achievable regions for the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 95–100, Jan. 2011.
[39] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
[40] E. Ekrem and S. Ulukus, "Capacity region of Gaussian MIMO broadcast channels with common and confidential messages," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5669–5680, Sep. 2012.
[41] Y. Geng and C. Nair, "The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2087–2104, Apr. 2014.
[42] Z. Goldfeld and H. H. Permuter, "MIMO Gaussian broadcast channels with common, private and confidential messages," *IEEE Trans. Inf. Theory*, to be published.
[43] A. Gohari, C. Nair, and V. Anantharam, "Improved cardinality bounds on the auxiliary random variables in Marton's inner bound," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 1272–1276.
[44] J. Hou and G. Kramer, "Effective secrecy: Reliability, confusion and stealth," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 601–605.
[45] 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.
[46] I. B. Gattegno, Z. Goldfeld, and H. H. Permuter, "Fourier–Motzkin elimination software for information theoretic inequalities," *IEEE Inf. Theory Soc. Newslett.*, vol. 65, no. 3, pp. 25–28, Sep. 2015. [Online]. Available: http://www.ee.bgu.ac.il/~fmeit/
[47] G. Kramer, "Teaching IT: An identity for the Gelfand–Pinsker converse," *IEEE Inf. Theory Soc. Newslett.*, vol. 61, no. 4, pp. 4–6, Dec. 2011.

Ziv Goldfeld (S'13) received his B.Sc. (summa cum laude) and M.Sc. (summa cum laude) degrees in Electrical and Computer Engineering from the Ben-Gurion University, Israel, in 2012 and 2014, respectively. He is currently a student in the direct Ph.D. program for honor students in Electrical and Computer Engineering at that same institution.

Between 2003 and 2006, he served in the intelligence corps of the Israeli Defense Forces.

Ziv is a recipient of several awards, among them are the Dean's List Award, the Basor Fellowship, the Lev-Zion fellowship, IEEEI-2014 best student paper award, a Minerva Short-Term Research Grant (MRG), and a Feder Family Award in the national student contest for outstanding research work in the field of communications technology.

Gerhard Kramer (S'91–M'94–SM'08–F'10) received the Dr. sc. techn. (Doktor der technischen Wissenschaften) degree from the Swiss Federal Institute of Technology (ETH), Zurich, in 1998.

From 1998 to 2000, he was with Endora Tech AG, Basel, Switzerland, as a Communications Engineering Consultant. From 2000 to 2008, he was with Bell Labs, Alcatel-Lucent, Murray Hill, NJ, as a Member of Technical Staff. He joined the University of Southern California (USC), Los Angeles, in 2009. Since 2010, he has been a Professor and Head of the Institute for Communications Engineering at the Technical University of Munich (TUM), Munich, Germany.

Dr. Kramer served as the 2013 President of the IEEE Information Theory Society. He has won several awards for his work and teaching, including an Alexander von Humboldt Professorship in 2010 and a Lecturer Award from the Student Association of the TUM Electrical and Computer Engineering Department in 2015. He has been a member of the Bavarian Academy of Sciences and Humanities since 2015.

Haim H. Permuter (M'08–SM'13) received his B.Sc. (summa cum laude) and M.Sc. (summa cum laude) degrees in Electrical and Computer Engineering from the Ben-Gurion University, Israel, in 1997 and 2003, respectively, and the Ph.D. degree in Electrical Engineering from Stanford University, California in 2008.

Between 1997 and 2004, he was an officer at a research and development unit of the Israeli Defense Forces. Since 2009 he is with the department of Electrical and Computer Engineering at Ben-Gurion University where he is currently an associate professor.

Prof. Permuter is a recipient of several awards, among them the Fullbright Fellowship, the Stanford Graduate Fellowship (SGF), Allon Fellowship, and and the U.S.-Israel Binational Science Foundation Bergmann Memorial Award. Haim is currently serving on the editorial board of the IEEE TRANSACTIONS ON INFORMATION THEORY.