Secret Key Generation over Noisy Channels with Correlated Sources

Germán Bassi, *Member, IEEE*, Pablo Piantanida, *Senior Member, IEEE*, and
Shlomo Shamai (Shitz), *Fellow Member, IEEE*

Abstract—This paper investigates the problem of secret key generation over a wiretap channel when the terminals observe correlated sources. These sources are independent of the main channel and the users overhear them before the transmission takes place. A novel achievable scheme is proposed, and its optimality is shown under certain less-noisy conditions. This result improves upon the existing literature where the more stringent condition of degradedness is required. Furthermore, numerical evaluation of the proposed scheme and previously reported results for a binary model are presented; a comparison of the numerical bounds provides insights on the benefit of the novel scheme.

I. INTRODUCTION

THE WIRETAP channel, introduced by Wyner [2], is the basic model for analyzing secrecy in wireless communications. In this model, the transmitter, named Alice, wants to communicate reliably with Bob while keeping the transmitted message –or part of it– secret from an eavesdropper, named Eve, overhearing the communication through another channel. Secrecy is characterized by the amount of information that is not *leaked*, which can be measured by the equivocation rate –the remaining uncertainty about the message at the eavesdropper. The secrecy capacity of the wiretap channel is thus defined as the maximum transmission rate that can be attained with zero leakage. In their influential paper [3], Csiszár and Körner determine the rate-equivocation region of a general broadcast channel with any arbitrary level of security, which also establishes the secrecy capacity of the wiretap channel. These schemes guarantee secrecy by exploiting an artificial random noise that saturates the eavesdropper's decoding capabilities.

On the other hand, Shannon [4] shows that it is also possible to achieve a positive secrecy rate by means of a *secret key*.

G. Bassi is with the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm 100 44, Sweden (e-mail: germanb@kth.se).

P. Piantanida is with CentraleSupélec–French National Center for Scientific Research (CNRS)–Université Paris-Sud, 3 Rue Joliot-Curie, F-91192 Gif-sur-Yvette, France, and with Montreal Institute for Learning Algorithms (MILA) at Université de Montréal, 2920 Chemin de la Tour, Montréal, QC H3T 1N8, Canada (e-mail: pablo.piantanida@centralesupelec.fr).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa, 32000, Israel (e-mail: sshlomo @ee.technion.ac.il).
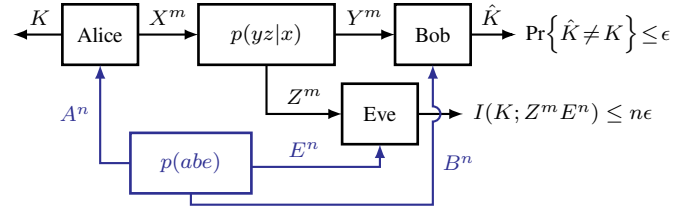
Fig. 1. System model for the problem of secret key generation.

Alice and Bob can safely communicate over a noiseless public broadcast channel as long as they share a secret key. The rate of this key, however, must be at least as large as the rate of the message to attain zero leakage. The main question that arises in this scenario is therefore: how do the legitimate users safely share the secret key? The answer is that the users should not communicate the key itself, which would then be compromised. Instead, they should only convey enough information to allow themselves to *agree* upon a key without disclosing, at the same time, any relevant information about it to the eavesdropper (for further discussion the reader is referred to [5], [6]).

In this work, we study the problem of secret key generation over a wiretap channel with correlated sources at each terminal. These sources are assumed to be independent of the main channel and there is no additional public broadcast channel of finite or infinite rate, as seen in Fig. 1. It is assumed that each node acquires the n-sequence observation of its corresponding source before the communication begins.

A. Related Work

Maurer [7] and Ahlswede and Csiszár [8] are among the first to study the use of correlated observations available at the legitimate users as a means to agree upon a key. In addition to the correlated observations, the terminals may communicate over a public broadcast channel of infinite capacity to which the eavesdropper has also access. Two models are proposed in [8]: the "source model", where the users observe correlated random sources controlled by nature, and the "channel model", where the users observe inputs and outputs of a noisy channel controlled by one of the users. In [9], Csiszár and Narayan study the first model but assume that the public broadcast channel has finite capacity and there is a third "helper" node who is not interested in recovering the key but rather helping Alice and Bob. The same authors also analyze the channel model with only one [10] or with multiple channel inputs [11].

Capacity results are presented in [9]–[11] assuming that there is only one round of communication over the public channel. General inner and outer bounds for both source and channel models with interaction over the public channel are introduced by Gohari and Anantharam in [12], [13].

More recently, Khisti *et al.* [14] investigate the situation where there is no helper node, the users communicate over a wiretap channel, and a separate public discussion channel may or may not be available. The simultaneous transmission of a secret message along with a key generation scheme using correlated sources is analyzed by Prabhakaran *et al.* [15]. The authors obtain a simple expression that reveals the trade-off between the achievable secrecy rate and the achievable rate of the secret key. The corresponding Gaussian channel with correlated Gaussian sources but independent of the channel components is recently studied in [16]. Closed form expressions for both secret key generation and secret message transmission are derived. On the other hand, Salimi *et al.* [17] consider simultaneous key generation of two independent users over a multiple access channel with feedback, where each user eavesdrops the other. In addition, the receiver can actively send feedback, through a private noiseless (or noisy) link, to increase the size of the shared keys.

The authors of [14]–[16] do not assume interactive communication, i.e., there is only one round of communication. Salimi *et al.* [17], however, allow the end user to respond once through the feedback link. Other authors have analyzed key generation schemes that rely on several rounds of transmissions. Tyagi [18] characterizes the minimum communication rate required to generate a maximum-rate secret key with r rounds of interactive communication. He shows that this rate is equal to the *interactive common information* (a quantity he introduces) minus the secret key capacity. In his model, two users observe i.i.d. correlated sources and communicate over an error-free channel. Hayashi *et al.* [19] study a similar problem but consider general (not necessarily i.i.d.) source sequences of finite length. Their proposed protocol attains the secret key capacity for general observations as well as the second-order asymptotic term of the maximum feasible secret key length for i.i.d. observations. They also prove that the standard one-way communication protocol fails to attain the aforementioned asymptotic result. Courtade and Halford [20] analyze the related problem of how many rounds of public transmissions are required to generate a specific number of secret keys. Their model assumes that there are n terminals connected through an error-free public channel, where each terminal is provided with a number of messages before transmission that it uses to generate the keys.

As previously mentioned, the focus of the present work is on sources that are independent of the main channel; nonetheless, some works have addressed the general situation of correlated sources and channels. Prior work on secrecy for channels with state include Chen and Vinck's [21] and Liu and Chen's [22] analysis of the wiretap channel with state. These works employ Gelfand and Pinsker's scheme [23] to correlate the transmitted codeword with the channel state at the same time that it saturates the eavesdropper's decoding capabilities. A single-letter expression of the secrecy capacity for this model is still unknown, although a multi-letter bound is provided by Muramatsu [24] and a novel lower bound is recently reported in [25]. As a matter of fact, the complexity of this problem also lies in the derivation of an outer bound that can handle simultaneously secrecy and channels with state.

To the best of our knowledge, only a handful of works have studied the problem of key generation for channels with state. The previously mentioned result of Prabhakaran *et al.* [15] is one of these examples. Zibaeenejad [26] analyzes a similar scenario where there is also a public channel of finite capacity between the users and he provides an inner and an outer bound for this model. Although the inner bound is developed for a channel with state, it is possible to apply it to the model used in the present work, i.e., sources independent of the main channel. However, some steps of the proof reported in [26] appear to be obscure and a constraint seems to be missing in the final expression; the resulting achievable rate was recently shown in [27] to be in certain cases unachievable. As a consequence, we have decided not to compare our inner bound to this previously reported scheme.

The works found in the literature that are closely related to the problem dealt here [14]–[17], [26] derive the equivocation of their schemes using a *weak secrecy* condition. In line with these works, we use the same measure in the analysis of our proposed scheme; however, it can be shown that our strategy also fulfills the *strong secrecy* criterion (see Remark 7) which has become more frequent nowadays. Recent works on the wiretap channel employ this approach, e.g., [28], [29], where in particular [29] does not assume that the messages have a uniform distribution.

Finally, it is worth noting that the problem of secure source transmission with side information [30]–[32] is closely related to the present work, since the reconstructed source may serve as a key as long as it has been reliably and securely transmitted. It is not surprising that some of the techniques developed in those works may be found here as well.

B. Contributions and Organization of the Paper

In this work, we introduce a novel coding scheme (Theorem 2) for the problem of secret key generation over a wiretap channel with correlated sources at each terminal. The correlated sources are assumed to be independent of the main channel and, thanks to a previously reported outer bound [33], this scheme is shown to be optimal (Propositions 1, 2, and 3) whenever the channel and/or source components satisfy the specific *less-noisy* conditions described in Table I. In contrast, the proposed schemes in [14]–[17] were optimal only when the stronger *degradedness* condition holds true for the channel and source components.

The main improvement of our scheme with respect to the literature is to introduce a two-layer codebook for describing the source. Although a two-layer scheme is not a new technique for the "source model" (cf. [8, Thm. 1]), it introduces considerable difficulty and has not been investigated in the framework of the combined model of Fig. 1. Difficulty arises in the derivation of Eve's equivocation, as shown by Lemma 2 in Section V-F. However, a scheme that is developed with

two description layers can achieve higher secret key rates than those of a single-layer scheme.

This paper is organized as follows. Section II provides some definitions and our previously reported outer bound. In Section III, we first present the inner bound for the problem of secret key agreement and then we enumerate the cases where said achievable scheme is optimal. Section IV illustrates with a binary example the improvement of the present work over a previously reported scheme. In Section V, we give the detailed proof of the inner bound. Finally, Section VI summarizes and concludes the work.

Notation and Conventions

Throughout this work, we use the standard notation of [34]. Specifically, given two integers i and j, the expression $[i : j]$ denotes the set $\{i, i+1, \ldots, j\}$, whereas for real values a and b, $[a, b]$ denotes the closed interval between a and b. We use the notation $x_i^j = (x_i, x_{i+1}, \ldots, x_j)$ to denote the sequence of length $j - i + 1$ for $1 \leq i \leq j$. If $i = 1$, we drop the subscript for succinctness, i.e., $x^j = (x_1, x_2, \ldots, x_j)$. Lowercase letters such as x and y are mainly used to represent constants or realizations of random variables, capital letters such as X and Y stand for the random variables in itself, and calligraphic letters such as $\mathcal{X}$ and $\mathcal{Y}$ are reserved for sets, codebooks, or special functions.

The set of nonnegative real numbers is denoted by $\mathbb{R}_+$. The probability distribution (PD) of the random vector X^n, $p_{X^n}(x^n)$, is succinctly written as $p(x^n)$ without subscript when it can be understood from the argument x^n. Given three random variables X, Y, and Z, if its joint PD can be decomposed as $p(xyz) = p(x)p(y|x)p(z|y)$, then they form a Markov chain, denoted by $X \multimap Y \multimap Z$. The random variable Y is said to be *less noisy* than Z w.r.t. X if $I(U; Y) \geq I(U; Z)$ for each random variable U such that $U \multimap X \multimap (Y, Z)$; this relation is denoted by $Y \succeq_x Z$. Entropy is denoted by $H(\cdot)$ and mutual information, $I(\cdot; \cdot)$. The expression $[x]^+$ denotes $\max\{x, 0\}$. Given $u, v \in [0, 1]$, the function $h_2(u) \triangleq -u \log_2 u - (1 - u) \log_2(1 - u)$ is the binary entropy function and $u * v \triangleq u(1 - v) + v(1 - u)$. We denote typical and conditional typical sets by $\mathcal{T}_\delta^n(X)$ and $\mathcal{T}_\delta^n(Y|x^n)$, respectively.

II. PRELIMINARIES

A. Problem Definition

Consider the wiretap channel with correlated sources at every node (A, B, E), as shown in Fig. 1. The legitimate users (Alice and Bob) want to agree upon a secret key $K \in \mathcal{K}$ while an eavesdropper (Eve) is overhearing the communication. Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{E}$, $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be six finite sets. Alice, Bob, and Eve observe the random sequences (sources) A^n, B^n, and E^n, respectively, drawn i.i.d. according to the joint distribution $p(abe)$ on $\mathcal{A} \times \mathcal{B} \times \mathcal{E}$. Alice communicates with Bob through m instances of a discrete memoryless channel with input $X \in \mathcal{X}$ and output $Y \in \mathcal{Y}$. Eve is listening the communication through another channel with input $X \in \mathcal{X}$ and output $Z \in \mathcal{Z}$. This channel is defined by its transition probability $p(yz|x)$ and it is independent of the sources' distribution.

Definition 1 (Code): A $(2^{nR_k}, n, m)$ secret key code c_n for this model consists of:
- a key set $\mathcal{K}_n \triangleq [1 : 2^{nR_k}]$,
- a source of local randomness $R_r \in \mathcal{R}_r$ at Alice,
- an encoding function $\varphi \colon \mathcal{A}^n \times \mathcal{R}_r \to \mathcal{X}^m$,
- a key generation function $\psi_a \colon \mathcal{A}^n \times \mathcal{R}_r \to \mathcal{K}_n$, and
- a key generation function $\psi_b \colon \mathcal{B}^n \times \mathcal{Y}^m \to \mathcal{K}_n$.

The rate of such a code is defined as the number of channel uses per source symbol $\frac{m}{n}$.

Given a code, let $K = \psi_a(A^n, R_r)$ and $X^m = \varphi(A^n, R_r)$; then, the performance of the $(2^{nR_k}, n, m)$ secret key code c_n is measured in terms of its average probability of error

$$\mathsf{P}_e(\mathsf{c}_n) \triangleq \Pr\{\psi_b(B^n, Y^m) \neq K | \mathsf{c}_n\}, \qquad (1)$$

in terms of the information leakage

$$\mathsf{L}_k(\mathsf{c}_n) \triangleq I(K; E^n Z^m | \mathsf{c}_n), \qquad (2)$$

and in terms of the uniformity of the keys

$$\mathsf{U}_k(\mathsf{c}_n) \triangleq nR_k - H(K | \mathsf{c}_n). \qquad (3)$$

Definition 2 (Achievability): A tuple $(\eta, R_k) \in \mathbb{R}_+^2$ is said to be *achievable* for this model if, for every $\epsilon > 0$ and sufficiently large n, there exists a $(2^{nR_k}, n, m)$ secret key code c_n such that

$$\frac{m}{n} \leq \eta + \epsilon, \quad \mathsf{P}_e(\mathsf{c}_n) \leq \epsilon, \quad \frac{1}{n}\mathsf{L}(\mathsf{c}_n) \leq \epsilon, \quad \frac{1}{n}\mathsf{U}(\mathsf{c}_n) \leq \epsilon. \quad (4)$$

The set of all achievable tuples is denoted by $\mathcal{R}^\star$ and is referred to as the *secret key region*.

B. Outer Bound

The following theorem gives an outer bound on $\mathcal{R}^\star$, i.e., it defines the region $\mathcal{R}_{\mathrm{out}} \supseteq \mathcal{R}^\star$.

Theorem 1: An *outer bound* on the secret key region for this channel model is given by

$$R_k \leq \max_{p \in \mathcal{P}} \big\{ \eta \big[I(T; Y) - I(T; Z) \big]$$
$$+ I(V; B|U) - I(V; E|U) \big\} \quad (5)$$

subject to

$$I(V; A|B) \leq \eta \, I(X; Y), \qquad (6)$$

where $\mathcal{P}$ is the set of input probability distributions given by

$$\mathcal{P} = \big\{ p(txyzuvabe) =$$
$$p(tx)p(yz|x)p(abe)p(v|a)p(u|v) \big\} \quad (7)$$

with $|\mathcal{T}| \leq |\mathcal{X}|$, $|\mathcal{U}| \leq |\mathcal{A}| + 1$, and $|\mathcal{V}| \leq (|\mathcal{A}| + 1)^2$.

Proof: Refer to Appendix A for details. ∎

Theorem 1 shows that the secret key generated between Alice and Bob has two components. The first two terms on the r.h.s. of (5) represent the part of the key that is securely transmitted through the noisy channel (given by the random variable T) as in the wiretap channel. On the other hand, the last two terms on the r.h.s. of (5) characterize the part of the key that is securely extracted from the correlated sources (given by the random variables U and V). Since the source and

channel variables are independent in the model, it should not be surprising that the variable T is independent of (U, V). However, given that the users need to agree on common extracted bits from the source, the noisy channel imposes the restriction (6) on the amount of information exchanged during that agreement.

Remark 1: The calculation of the bounds (5) and (6) is done using the probability distribution (7). However, we note that (7) is an uncommon single-letter expression of the source and channel variables since the sequences have different lengths. This remark is also applicable to all the regions presented in the sequel.

III. MAIN RESULTS

In this section, we first introduce a key generation scheme for the aforementioned model that leads to a novel inner bound for the secret key region (Theorem 2). We then study some special cases where this scheme turns out to achieve the (optimal) secret key region (Propositions 1, 2, and 3).

A. Inner Bound

The following theorem gives an inner bound on $\mathcal{R}^\star$, i.e., it defines the region $\mathcal{R}_{\text{in}} \subseteq \mathcal{R}^\star$.

Theorem 2: A tuple $(\eta, R_k) \in \mathbb{R}_+^2$ is achievable if there exist random variables U, V, Q, T, X on finite sets $\mathcal{U}$, $\mathcal{V}$, $\mathcal{Q}$, $\mathcal{T}$, $\mathcal{X}$, respectively, with joint distribution $p(uvqtxyzabe) = p(q|t)p(tx)p(yz|x)p(abe)p(v|a)p(u|v)$, which verify

$$R_k \leq \eta \big[I(T; Y|Q) - I(T; Z|Q) \big] \\ + I(V; B|U) - I(V; E|U) \quad (8)$$

subject to

$$I(U; A|B) \leq \eta \, I(Q; Y), \quad (9a)$$
$$I(V; A|B) \leq \eta \, I(T; Y). \quad (9b)$$

Moreover, it suffices to consider sets $\mathcal{U}$, $\mathcal{V}$, $\mathcal{Q}$, and $\mathcal{T}$ such that $|\mathcal{U}| \leq |\mathcal{A}| + 2$, $|\mathcal{V}| \leq (|\mathcal{A}| + 1)(|\mathcal{A}| + 2)$, $|\mathcal{Q}| \leq |\mathcal{X}| + 2$, and $|\mathcal{T}| \leq (|\mathcal{X}| + 1)(|\mathcal{X}| + 2)$.

Proof: Alice employs the two-layer description (U, V) to compress the source A and it transmits it through the two-layer channel codeword (Q, T). Each layer of the description must fit in the corresponding layer of the channel codeword according to (9). The achievable secret key rate (8) is a combination of the secret bits transmitted through the noisy channel in the manner of the wiretap channel and the secret bits obtained by the reconstruction of the source at Bob. The full proof is deferred to Section V. ∎

Remark 2: The regions $\mathcal{R}_{\text{out}}$ and $\mathcal{R}_{\text{in}}$ do not coincide in general. This is due to the presence of the condition (9a) in the inner bound, and the looser condition (6) in the outer bound with respect to (9b). We present in Section III-B a few special cases where these differences disappear and both regions coincide.

Remark 3: By setting $U = \emptyset$, the region in Theorem 2 recovers the results in [14, Thm. 1 and 4], when the eavesdropper has access to a correlated source, and [15, Thm. 2], when

there is no secret message to be transmitted. In these works, there was only one layer to encode the source A^n while our coding scheme allows for two layers, introducing considerable difficulty in the derivation of Eve's equivocation (see e.g. the multiple binning stages in the proof and (40)). The advantage of having two layers of description is that Theorem 2 can potentially achieve higher secret key rates (see Section IV) and it recovers the result of Csiszár and Narayan [9] (see Remark 8).

Remark 4: The region in Theorem 2 also recovers the result in [35, Thm. 1], which was published after the original submission of this manuscript. In that work, Alice and Bob communicate over a public noiseless channel of rate R_1 and a secure noiseless channel of rate R_2. The proposed achievable scheme in [35] sends the codeword Q through the public channel, i.e., $I(Q; Y) = R_1$, and the codeword T through the secure channel, i.e., $I(T; Y|Q) = R_2$ and $I(T; Z|Q) = 0$. The reader may verify that, by using the aforementioned quantities and $\eta = 1$, both regions are equal.

Remark 5: Theorem 2 improves upon our previous work in [33, Sec. IV-A] since (9) replaces the more stringent condition: $I(V; A|B) \leq \eta \, I(Q; Y)$.

Remark 6: The problem of key generation dealt with in the present work is intimately connected to the problem of secure source transmission with side information, at both receiver and eavesdropper [31], [32], since the part of the source that can be reliably and securely transmitted serves as key which is a function of the source. It is thus not surprising that Theorem 2 bears a resemblance to our previous result in [32, Thm. 2].

Remark 7: Theorem 2 is obtained using the *weak* secrecy and uniformity conditions in (4). However, employing the method introduced in [36], we can show that the *strong* secrecy and uniformity conditions, i.e., $\mathsf{L}(\mathsf{c}_n) \leq \epsilon$ and $\mathsf{U}(\mathsf{c}_n) \leq \epsilon$, also hold true. The proof relies on using l times a secret key code c_n to generate l independent keys. We then interpret these l keys as l realizations of a random source in the "source model", which allows us to distill strong secret keys by means of a one-way direct reconciliation protocol and privacy amplification with extractors. These two steps involve the transmission of additional information through the channel; nonetheless, the cost of these additional channel uses is negligible compared to the total transmission time for large l, m, and n. We omit the details of the proof here due to the similarities with [37, Sec. 4.5] and [33, App. B-C].

B. Optimal Characterization of the Secret Key Rate

The inner bound $\mathcal{R}_{\text{in}}$ is optimal under certain less-noisy conditions in channel and/or source components. These special cases are summarized in Table I.

1) Eve Has a Less Noisy Channel: If Eve has a less noisy channel than Bob, i.e., $Z \succeq_x Y$, the information transmitted over the channel is compromised. Therefore, the amount of secret key that can be generated only depends on the statistical differences between sources.

Proposition 1: If $Z \succeq_x Y$, a tuple $(\eta, R_k) \in \mathbb{R}_+^2$ is achievable if and only if there exist random variables U, V,

	$E \succeq_A B$	$B \succeq_A E$
$Z \succeq_X Y$	$R_k = 0$	Proposition 1
$Y \succeq_X Z$	Proposition 2	Proposition 3

TABLE I
REGIMES WHERE THEOREM 2 IS OPTIMAL. NO SECRET KEY IS
ACHIEVABLE IF $Z \succeq_X Y$ AND $E \succeq_A B$.

X on finite sets $\mathcal{U}$, $\mathcal{V}$, $\mathcal{X}$, respectively, with joint distribution $p(uvabexyz) = p(u|v)p(v|a)p(abe)p(x)p(yz|x)$, which verify

$$R_k \leq I(V;B|U) - I(V;E|U) \tag{10a}$$

$$\text{subject to} \quad I(V;A|B) \leq \eta\, I(X;Y). \tag{10b}$$

Proof: Given the less-noisy condition on Eve's channel, i.e., $I(T;Y) \leq I(T;Z)$ for any RV T such that $T \multimap X \multimap (YZ)$, the bound (5) is maximized with $T = \emptyset$. On the other hand, the region (10) is achievable by setting auxiliary RVs $Q = T = X$ in $\mathcal{R}_{\text{in}}$. ∎

Remark 8: The secret key capacity of the wiretap channel with a public noiseless channel of rate R [9, Thm. 2.6] turns out to be a special case of Proposition 1, where $X = Y = Z$ and defining $\eta\, H(X) = \eta \log|\mathcal{X}| \equiv R$.

2) Eve Has a Less Noisy Source: If Eve has a less noisy source than Bob, i.e., $E \succeq_A B$, the amount of secret key that can be generated depends on the amount of secure information transmitted through the wiretap channel.

Proposition 2: If $E \succeq_A B$, a tuple $(\eta, R_k) \in \mathbb{R}_+^2$ is achievable if and only if there exist random variables T, X on finite sets $\mathcal{T}$, $\mathcal{X}$, respectively, with joint distribution $p(txyz) = p(tx)p(yz|x)$, which verify

$$R_k \leq \eta\big[I(T;Y) - I(T;Z)\big]. \tag{11}$$

Proof: Given the less-noisy condition on Eve's source, i.e., $I(V;B) \leq I(V;E)$ for any RV V such that $V \multimap A \multimap (BE)$, the bound (5) is maximized with $U = V$ and independent of the sources. The region (11) is achievable by using the same auxiliary RVs in the inner bound as in the outer bound. ∎

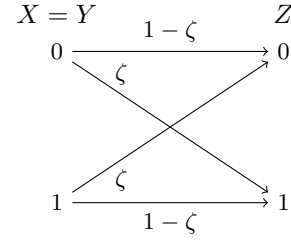*Remark 9:* The bound (11) is equal to the secrecy capacity of the wiretap channel.

Remark 10: Even though the bound (11) becomes independent of the sources sequences (A^n, B^n, E^n), we assume $n \neq 0$, and thus the rate η is finite.

3) Bob Has a Less Noisy Channel and Source: If Bob has a less noisy channel and source than Eve, i.e., $Y \succeq_X Z$ and $B \succeq_A E$, the lower layers of the channel and source codewords are not needed any more.
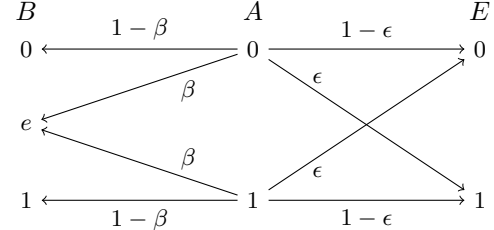
Proposition 3: If $Y \succeq_X Z$ and $B \succeq_A E$, a tuple $(\eta, R_k) \in \mathbb{R}_+^2$ is achievable if and only if there exist random variables V, X on finite sets $\mathcal{V}$, $\mathcal{X}$, respectively, with joint distribution $p(vabexyz) = p(v|a)p(abe)p(x)p(yz|x)$, which verify

$$R_k \leq \eta\big[I(X;Y) - I(X;Z)\big] + I(V;B) - I(V;E) \tag{12a}$$

$$\text{subject to} \quad I(V;A|B) \leq \eta\, I(X;Y). \tag{12b}$$

(a) Main channel.

(b) BEC/BSC sources.

Fig. 2. System model for the wiretap channel with BEC/BSC sources.

Proof: Given the less-noisy conditions on Bob's channel and source, the bound (5) is maximized with $U = \emptyset$ and $T = X$. The region (12) is achievable by also setting auxiliary RVs $U = Q = \emptyset$ and $T = X$ in the inner bound. ∎

Remark 11: Proposition 3 extends the results from [14, Thm. 4] and [15, Thm. 3] which assumed the more stringent conditions of degradedness: $A \multimap B \multimap E$ and $X \multimap Y \multimap Z$.

IV. SECRET KEY AGREEMENT OVER A WIRETAP CHANNEL WITH BEC/BSC SOURCES

As mentioned in Remark 3, the inner bound introduced in Section III-A employs two layers of description, and thus it is an improvement over previously reported results. In this section, we compare the performance of our achievable scheme with that of [14] for a specific binary source and channel model.

A. System Model

Consider the communication system depicted in Fig. 2. The main channel consists of a noiseless link from Alice to Bob and a binary symmetric channel (BSC) with crossover probability $\zeta \in \big[0, \frac{1}{2}\big]$ from Alice to Eve (see Fig. 2a). Additionally, the three nodes have access to correlated sources; in particular, Alice observes a binary uniformly distributed source, i.e., $A \sim \mathcal{B}\big(\frac{1}{2}\big)$, which is the input of two parallel channels as shown in Fig. 2b. Bob observes the output of a binary erasure channel (BEC) with erasure probability $\beta \in [0, 1]$, and Eve, a BSC with crossover probability $\epsilon \in \big[0, \frac{1}{2}\big]$. For simplicity, we assume $\eta = 1$ in the sequel.

Remark 12: The sources (A, B, E) satisfy different properties according to the values of the parameters (β, ϵ) [38], specifically:

- if $0 \leq \beta < 2\epsilon$, E is a *degraded* version of B, i.e., $A \multimap B \multimap E$;

- if $2\epsilon \leq \beta < 4\epsilon(1 - \epsilon)$, B is *less noisy* than E, i.e., $B \succeq_A E$; and,
- if $4\epsilon(1 - \epsilon) \leq \beta < h_2(\epsilon)$, B is *more capable* than E.

B. Performance of the Coding Scheme

The following proposition provides a simple expression of the inner bound from Theorem 2. The expression is obtained by simplifying the maximization process of the input distribution, and thus it might not be optimal. However, this suffices to show the higher rates achieved by this scheme as we see later.

Proposition 4: The tuple $(\eta = 1, R_k) \in \mathcal{R}_{\text{in}}$ if there exist $u, v, q \in \left[0, \frac{1}{2}\right]$ such that:

$$R_k \leq (1 - \beta)\big[h_2(v * u) - h_2(v)\big] - h_2(v * u * \epsilon) \tag{13a}$$
$$+ h_2(v * \epsilon) + h_2(\zeta) + h_2(q) - h_2(\zeta * q),$$
$$\text{subject to} \quad \beta\big[1 - h_2(v * u)\big] \leq 1 - h_2(q). \tag{13b}$$

Proof: The bound (13) is directly calculated from (8) and (9a) with the following choice of input random variables: $T = X$, $Q = X \oplus Q'$, $V = A \oplus V'$, and $U = V \oplus U'$. Here, $X \sim \mathcal{B}(\frac{1}{2})$, $Q' \sim \mathcal{B}(q)$, $V' \sim \mathcal{B}(v)$, and $U' \sim \mathcal{B}(u)$, and each random variable is independent of each other and (A, B, E). The condition (9b) in the inner bound becomes redundant with the mentioned choice of input distribution. ∎

As previously mentioned, we provide next the inner bound presented in [14, Thm. 4][1] as a means of comparison. This inner bound is similar to Theorem 2 but with only one layer of description for the source A; thus, its achievable region is denoted $\mathcal{R}_{\text{in-1L}}$.

Proposition 5 ([14, Thm. 4]): The tuple $(\eta = 1, R_k) \in \mathcal{R}_{\text{in-1L}}$ if and only if

$$R_k \leq \big[h_2(\epsilon) - \beta\big]^+ + h_2(\zeta). \tag{14}$$

Proof: See Appendix B. ∎

Remark 13: Proposition 5 is a special case of Proposition 4 with $u = q = \frac{1}{2}$, and $v = 0$ or $v = \frac{1}{2}$. As mentioned in Remark 3, the inner bound [14, Thm. 4] is a special case of our Theorem 2 with $U = \emptyset$ (thus $u = \frac{1}{2}$). Moreover, given that in this model the Markov chain $X \multimap Y \multimap Z$ holds, the channel codebook of Proposition 5 has only one layer (thus $q = \frac{1}{2}$). On the other hand, there are two layers of description in Proposition 4, and whenever $U \neq \emptyset$ (i.e., $u < \frac{1}{2}$), we have that $Q \neq \emptyset$ (i.e., $q < \frac{1}{2}$). This relationship is determined by (13b).

We perform numerical optimization of the bound (13) for different values of β while fixing $\zeta = 0.01$ and $\epsilon = 0.05$; the results are shown in Fig. 3 along with the bound (14). We see in the figure the advantage of having two layers of description for the source A. Our proposed scheme, Proposition 4, attains higher secret key rates than the scheme with only one layer of description (Proposition 5) for intermediate values of β. It is in this regime, when the source B is no longer *less noisy* than E, that two layers of description are needed.

[1]Theorem 4 from [14] is actually a capacity result assuming that $A \multimap B \multimap E$ and $X \multimap Y \multimap Z$. In our present example, only the second Markov chain holds independently of the value of the parameters β and ϵ, but this does not invalidate the use of the inner bound.

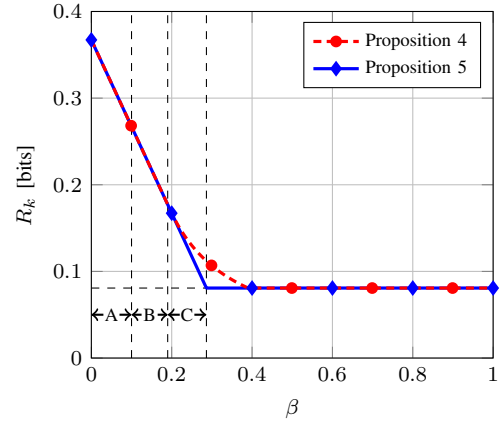Fig. 3. Achievable secret key rates for the wiretap channel with BEC/BSC sources, with $\zeta = 0.01$ and $\epsilon = 0.05$. In region A, $A \multimap B \multimap E$, while in region B, $B \succeq_A E$, and finally in region C, B is more capable than E. The horizontal dotted line corresponds to the secrecy capacity of the main channel, i.e., $h_2(\zeta)$.

V. Proof of Theorem 2

We begin by presenting a high-level description of the coding strategy before properly developing the proof. In this scheme, the secret key is learned by extracting common bits from the correlated sources and from exchanging other bits through the noisy channel. In particular, Alice compresses the source observation A^n using a two-layer source codebook (determined by U^n and V^n). Alice then transmits the corresponding bin indices r_1 and (r_2, r_p) to Bob with the aid of a code for the wiretap channel. Using his side-information B^n, Bob recovers the codewords U^n and V^n and he further obtains the bin indices (r_2, k_1), where k_1 is independent of r_p provided that the conditions of Lemma 2 (Section V-F) are met. The key is finally generated using bits from the indices k_1 and k_2, where the latter was sent over the noisy channel along with r_1, r_2, and r_p. We provide a detailed proof in the following.

A. Codebook Generation

Let us define the quantity

$$R_f < (\eta + \epsilon)I(T; Z|Q) - \epsilon_f, \tag{15}$$

and fix the following joint probability distribution:

$$p(qtxyzuvabe) =$$
$$p(q|t)p(tx)p(yz|x)p(u|v)p(v|a)p(abe). \tag{16}$$

Then, proceed as follows:
1) Randomly pick 2^{nS_1} sequences $u^n(s_1)$ from $\mathcal{T}_\delta^n(U)$ and divide them into 2^{nR_1} equal-size bins $\mathcal{B}_1(r_1)$, $r_1 \in [1 : 2^{nR_1}]$.
2) For each codeword $u^n(s_1)$, randomly pick 2^{nS_2} sequences $v^n(s_1, s_2)$ from $\mathcal{T}_\delta^n(V|u^n(s_1))$[2] and divide

[2]As a matter of fact, the sequences $v^n(s_1, s_2)$ should be chosen from $\mathcal{T}_{\delta'}^n(V|u^n(s_1))$, $\delta < \delta'$, in order to assure that $(u^n(s_1), v^n(s_1, s_2)) \in \mathcal{T}_{\delta'}^n(UV)$ (see e.g. Conditional Typicality Lemma [34]). This remark also applies to the generation of the channel codewords $q^m(\cdot)$ and $t^m(\cdot)$ in this part of the proof. However, we omit this detail throughout the proof to simplify the notation and ease the reading.
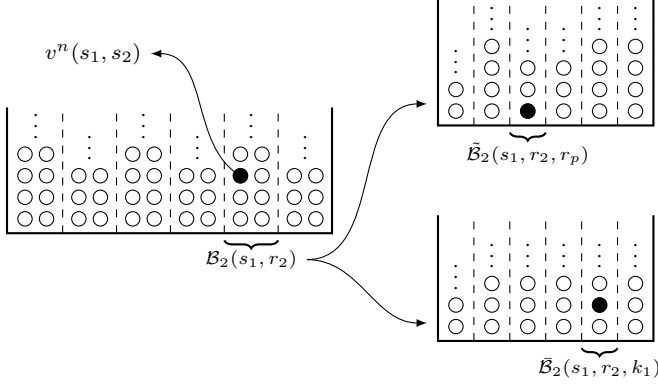
Fig. 4. Multiple binning stages of the codewords $v^n(s_1, s_2)$, where each circle represents a codeword. The fact that a codeword belongs to different sub-bins $\tilde{\mathcal{B}}_2$ and $\bar{\mathcal{B}}_2$ is shown through a black circle, which depicts the same codeword.

them into 2^{nR_2} equal-size bins $\mathcal{B}_2(s_1, r_2)$, $r_2 \in [1 : 2^{nR_2}]$. Furthermore, distribute the sequences inside each bin in two different types of sub-bin:

- 2^{nR_p} equal-size sub-bins $\tilde{\mathcal{B}}_2(s_1, r_2, r_p)$, $r_p \in [1 : 2^{nR_p}]$; and,
- $2^{nR_{k_1}}$ equal-size sub-bins $\bar{\mathcal{B}}_2(s_1, r_2, k_1)$, $k_1 \in [1 : 2^{nR_{k_1}}]$.

Note that a sequence $v^n(s_1, s_2)$ belongs to sub-bins $\tilde{\mathcal{B}}_2(s_1, r_2, r_p)$ and $\bar{\mathcal{B}}_2(s_1, r_2, k_1)$ where r_p and k_1 are independent. See Fig. 4 for a schematic representation.
3) Randomly pick $2^{n(R_1+R_2)}$ sequences $q^m(r_1, r_2)$ from $\mathcal{T}_\delta^m(Q)$.
4) For each $q^m(r_1, r_2)$, randomly pick $2^{n(R_p+R_{k_2}+R_f)}$ sequences $t^m(r_1, r_2, r_p, k_2, r_f)$ from $\mathcal{T}_\delta^m(T|q^m(r_1, r_2))$.
5) Randomly distribute the set of $2^{n(R_{k_1}+R_{k_2})}$ indices (k_1, k_2) into 2^{nR_k} equal-size bins $\mathcal{B}_k(k)$, $k \in [1 : 2^{nR_k}]$.

B. Encoding

Given a sequence a^n, and the indices k_2 and r_f chosen uniformly at random in $[1 : 2^{nR_{k_2}}]$ and $[1 : 2^{nR_f}]$, the encoder proceeds as follows:

1) It looks for an index $s_1 \equiv \hat{s}_1$ such that $(u^n(\hat{s}_1), a^n) \in \mathcal{T}_{\delta'}^n(UA)$. If more than one index is found, choose one uniformly at random among them, whereas if there is no such index, choose one uniformly at random in $[1 : 2^{nS_1}]$. The probability of not finding such an index is arbitrarily small as $n \to \infty$ if $\delta' < \epsilon_1$ and

$$S_1 > I(U; A) + \epsilon_1. \qquad (17)$$

2) Then, it looks for an index $s_2 \equiv \hat{s}_2$ such that $(v^n(s_1, \hat{s}_2), a^n) \in \mathcal{T}_{\delta'}^n(VA|u^n(s_1))$. If more than one index is found, choose one uniformly at random among them, whereas if there is no such index, choose one uniformly at random in $[1 : 2^{nS_2}]$. The probability of not finding such an index is arbitrarily small as $n \to \infty$ if $\delta' < \epsilon_2$ and

$$S_2 > I(V; A|U) + \epsilon_2. \qquad (18)$$

3) Let $\mathcal{B}_1(r_1)$ and $\tilde{\mathcal{B}}_2(s_1, r_2, r_p)$ be the bins of $u^n(s_1)$ and $v^n(s_1, s_2)$, respectively.
4) The encoder selects the codeword $t^m(r_1, r_2, r_p, k_2, r_f)$. It then transmits the associated jointly typical sequence $x^m \sim \prod_{i=1}^m p(x_i|t_i(r_1, r_2, r_p, k_2, r_f))$, generated on the fly.

C. Decoding

Given a sequence b^n and the channel output y^m, the decoder proceeds as follows:

1) It starts by looking for the unique set of indices $(r_1, r_2, r_p, k_2, r_f) \equiv (\hat{r}_1, \hat{r}_2, \hat{r}_p, \hat{k}_2, \hat{r}_f)$ such that

$$\left(q^m(\hat{r}_1, \hat{r}_2), t^m(\hat{r}_1, \hat{r}_2, \hat{r}_p, \hat{k}_2, \hat{r}_f), y^m\right) \in \mathcal{T}_\delta^m(QTY).$$

The probability of error in decoding can be made arbitrarily small as $(n, m) \to \infty$ provided that

$$R_1 + R_2 + R_p + R_{k_2} + R_f < (\eta + \epsilon)I(T; Y) - \delta,$$
$$R_p + R_{k_2} + R_f < (\eta + \epsilon)I(T; Y|Q) - \delta.$$

2) The decoder looks for the unique index $s_1 \equiv \hat{s}_1$ such that $u^n(\hat{s}_1) \in \mathcal{B}_1(r_1)$ and $(u^n(\hat{s}_1), b^n) \in \mathcal{T}_\delta^n(UB)$. The probability of error in decoding can be made arbitrarily small as $n \to \infty$ provided that

$$S_1 - R_1 < I(U; B) - \delta. \qquad (19)$$

3) Then, it looks for the unique index $s_2 \equiv \hat{s}_2$ such that $v^n(s_1, \hat{s}_2) \in \tilde{\mathcal{B}}_2(s_1, r_2, r_p)$ and $(v^n(s_1, \hat{s}_2), b^n) \in \mathcal{T}_\delta^n(VB|u^n(s_1))$. The probability of error in decoding can be made arbitrarily small as $n \to \infty$ provided that

$$S_2 - R_2 - R_p < I(V; B|U) - \delta. \qquad (20)$$

D. Key Generation

According to the preceding steps and with increasing high probability as $(n, m) \to \infty$, Bob correctly decodes the index k_2 and both Alice and Bob possess the same sequence $v^n(s_1, s_2) \in \bar{\mathcal{B}}_2(s_1, r_2, k_1)$. Therefore, they both agree on the same secret key k, which is the bin where the pair (k_1, k_2) belongs, i.e., $(k_1, k_2) \in \mathcal{B}_k(k)$.

E. Key Uniformity

Consider the following chain of inequalities:

$$H(K|\mathcal{C}) = H(K_1K_2|\mathcal{C}) - H(K_1K_2|\mathcal{C}K) \qquad (21a)$$
$$\geq H(K_1|\mathcal{C}) + nR_{k_2} - n(R_{k_1} + R_{k_2} - R_k) \qquad (21b)$$
$$\geq H(K_1|\mathcal{C}U^n) - n(R_{k_1} - R_k) \qquad (21c)$$
$$= H(V^n|\mathcal{C}U^n) - H(V^n|\mathcal{C}U^nK_1) - n(R_{k_1} - R_k) \qquad (21d)$$
$$\geq H(V^n|\mathcal{C}U^n) - n(S_2 - R_{k_1}) - n(R_{k_1} - R_k), \qquad (21e)$$

where

- (21b) follows from K_2 being chosen uniformly in $[1 : 2^{nR_{k_2}}]$ and independently of K_1, and that there are $2^{n(R_{k_1}+R_{k_2}-R_k)}$ pairs (K_1, K_2) in each bin K;
- (21d) is due to K_1 being a function of $(V^n, \mathcal{C})$; and,

- (21e) is due to the number of sequences V^n associated with sub-bin index K_1 being $2^{n(S_2-R_{k_1})}$, i.e., $\log \sum_{r_2} |\bar{\mathcal{B}}_2(s_1, r_2, k_1)| = n(S_2 - R_{k_1})$.

Before continuing the analysis, we introduce the random variable Υ, such that

$$\Upsilon \triangleq \mathbb{1}\{(U^n, A^n) \in \mathcal{T}_\delta^n(UA)\}. \tag{22}$$

Moreover, in order to improve readability, we drop the index from the codeword U^n, and thus the codebook $\mathcal{C}$ is composed of: $U^n \in \mathcal{T}_\delta^n(U)$ and $V^n(s) \in \mathcal{T}_\delta^n(V|U^n)$, where $s \in \mathcal{S} \triangleq [1 : 2^{nS_2}]$. Finally, we note that, conditioned on the codebook $\mathcal{C}$, the entropy of V^n is given by the entropy of its index S. Therefore,

$$\begin{align}
H(V^n|\mathcal{C}U^n) &= H(S|\mathcal{C}U^n) \tag{23a} \\
&\geq H(S|\mathcal{C}U^n\Upsilon) \tag{23b} \\
&\geq H(S|\mathcal{C}U^n, \Upsilon = 1)(1 - \epsilon), \tag{23c}
\end{align}$$

where the last step is due to $\Pr\{\Upsilon = 1\} \geq 1 - \epsilon$.

Now, for a specific codebook $\mathcal{C} = \mathsf{c}_n$ (which determines the codeword $U^n = u^n$), let us define the random variable S_c with distribution

$$p_{S_c} \triangleq p_{S|\mathcal{C}=\mathsf{c}_n, U^n=u^n, \Upsilon=1}. \tag{24}$$

Therefore,

$$H(S_c) = H(S|\mathcal{C} = \mathsf{c}_n, U^n = u^n, \Upsilon = 1), \tag{25}$$

and

$$\begin{align}
H(S|\mathcal{C}U^n, \Upsilon = 1) &= \mathbb{E}_\mathcal{C}\big[H(S_c)\big] \tag{26a} \\
&= \sum_{s \in \mathcal{S}} \mathbb{E}_\mathcal{C}\big[-p_{S_c}(s) \log p_{S_c}(s)\big] \tag{26b} \\
&= |\mathcal{S}|\, \mathbb{E}_\mathcal{C}\big[-p_{S_c}(1) \log p_{S_c}(1)\big], \tag{26c}
\end{align}$$

where the last step is due to the symmetry of the random codebook generation and encoding procedure, i.e., the probability p_{S_c} is independent of the specific value of the index. This is addressed in the following lemma.

Lemma 1: Let $\varepsilon_1, \varepsilon_2, \xi > 0$ and let χ be a function of the codebook c_n defined as

$$\chi(\mathsf{c}_n) \triangleq \mathbb{1}\big\{\big|p_{S_c}(1) - |\mathcal{S}|^{-1}\big| \geq \varepsilon_1 |\mathcal{S}|^{-1}\big\}. \tag{27}$$

Then, $\Pr\{\chi(\mathcal{C}) = 1\} \leq \varepsilon_2$ for large n if $S_2 < H(A) - \xi$.

Proof: This lemma is similar to the one introduced in [33, Lemma 5] and its proof is reproduced in Appendix C for completeness. ∎

Using the previous lemma we may continue (26),

$$\begin{align}
H(S|&\mathcal{C}U^n, \Upsilon = 1) \\
&\geq |\mathcal{S}|\, \mathbb{E}_\mathcal{C}\big[-p_{S_c}(1) \log p_{S_c}(1) \mid \chi(\mathcal{C}) = 0\big](1 - \varepsilon_2) \tag{28a} \\
&\geq (1 - \varepsilon_1)\big[\log |\mathcal{S}| - \log(1 + \varepsilon_1)\big](1 - \varepsilon_2) \tag{28b} \\
&\geq n(S_2 - \varepsilon'), \tag{28c}
\end{align}$$

for some $\varepsilon' > 0$. Putting together (21), (23), and (28), we obtain

$$H(K|\mathcal{C}) \geq n(S_2 - \varepsilon')(1 - \epsilon) - n(S_2 - R_k) \geq n(R_k - \epsilon'), \tag{29}$$

for some $\epsilon' > 0$. Finally, the uniformity of the keys, as defined in (3), averaged over all codebooks is

$$\mathbb{E}[\mathsf{U}_k(\mathcal{C})] = nR_k - H(K|\mathcal{C}) \leq n\epsilon', \tag{30}$$

and thus the key is asymptotically uniform.

Remark 14: It is worth noting that the preceding steps show that the probability of V^n is almost uniformly distributed on the codebook,

$$H(V^n|\mathcal{C}U^n) \geq n(S_2 - \varepsilon')(1 - \epsilon), \tag{31}$$

which follows from (23) and (28). A lower bound on $H(U^n|\mathcal{C})$ may be obtained using a similar analysis. Given that the sequences U^n and V^n are divided randomly and independently on equal-size bins and sub-bins, the bin and sub-bin indices (e.g. r_p) are also distributed almost uniformly on their respective sets.

F. Key Leakage

We may first relate the entropy of K to that of (K_1, K_2) as in (21),

$$\begin{align}
H(K&|\mathcal{C}E^nZ^m) \\
&= H(K_1K_2|\mathcal{C}E^nZ^m) - H(K_1K_2|\mathcal{C}E^nZ^mK) \tag{32a} \\
&\geq H(K_1K_2|\mathcal{C}E^nZ^m) - n(R_{k_1} + R_{k_2} - R_k). \tag{32b}
\end{align}$$

Then, consider the following chain of inequalities:

$$\begin{align}
H(K_1&K_2|\mathcal{C}E^nZ^m) \\
&\geq H(K_1K_2|\mathcal{C}E^nZ^mr_1r_2) \\
&= H(K_2U^nV^n|\mathcal{C}E^nZ^mr_1r_2) \\
&\quad - H(U^nV^n|\mathcal{C}E^nZ^mr_1r_2K_1K_2) \tag{33a} \\
&\geq H(K_2U^nV^n|\mathcal{C}E^nZ^mr_1r_2) - H(U^n|\mathcal{C}E^nr_1) \\
&\quad - H(V^n|\mathcal{C}E^nZ^mU^nr_2K_1K_2) \\
&\geq H(K_2U^nV^n|\mathcal{C}E^nZ^mr_1r_2) - 2n\epsilon_n \tag{33b} \\
&= H(K_2U^nV^nA^n|\mathcal{C}E^nZ^mr_1r_2) \\
&\quad - H(A^n|\mathcal{C}E^nZ^mU^nV^nK_2) - 2n\epsilon_n \tag{33c} \\
&\geq H(K_2A^n|\mathcal{C}E^nZ^mr_1r_2) - H(A^n|U^nV^nE^n) - 2n\epsilon_n \\
&\geq H(A^n|\mathcal{C}E^nZ^mr_1r_2K_2) + H(K_2|\mathcal{C}E^nZ^mr_1r_2) \\
&\quad - n[H(A|UVE) + 2\epsilon_n] \\
&= \underbrace{H(A^n|\mathcal{C}E^nZ^mr_1r_2r_pK_2)}_{\triangleq E_s} \\
&\quad + \underbrace{I(A^n; r_p|\mathcal{C}E^nZ^mr_1r_2K_2) + H(K_2|\mathcal{C}E^nZ^mr_1r_2)}_{\triangleq E_c} \\
&\quad - n[H(A|UVE) + 2\epsilon_n], \tag{33d}
\end{align}$$

where

- (33a) is due to K_1 being a function of $(V^n, \mathcal{C})$;
- (33b) follows from Lemma 2 below; and,
- (33c) is due to (r_1, r_2) being functions of $(U^n, V^n, \mathcal{C})$.

Lemma 2: Let $\epsilon_n, \delta, \delta', \varepsilon_1 > 0$, then, given the codebook generation and encoding procedure of the scheme,

$$H(U^n|\mathcal{C}E^nr_1) \leq n\epsilon_n \tag{34a}$$

if $S_1 - R_1 < I(U;E) - \delta$, and

$$H(V^n|\mathcal{C}E^n Z^m U^n r_2 K_1 K_2) \leq n\epsilon_n \tag{34b}$$

if $S_2 - R_2 - R_{k_1} + R_f < I(V;E|U) + (\eta + \epsilon)I(T;Z|Q) - \delta'$ and $R_p + R_f > (\eta + \epsilon)I(T;Z|Q) + \varepsilon_1$.

Proof: See Appendix D. ∎

In the last step of (33), we split up the equivocation into two parts as in [32]. The "source" term E_s writes:

$$E_s = H(A^n|\mathcal{C}E^n r_1 r_2 r_p) \tag{35a}$$
$$= H(A^n r_2 r_p|\mathcal{C}E^n r_1) - H(r_2 r_p|\mathcal{C}E^n r_1) \tag{35b}$$
$$= H(A^n|\mathcal{C}E^n r_1) + H(r_2 r_p|\mathcal{C}A^n E^n r_1) - H(r_2 r_p|\mathcal{C}) $$
$$\quad + I(r_2 r_p; E^n r_1|\mathcal{C}) \tag{35c}$$
$$\geq H(A^n|U^n E^n) + H(r_p|\mathcal{C}A^n E^n r_1 r_2) - n(R_2 + R_p) $$
$$\quad + I(r_2 r_p; E^n|\mathcal{C}r_1) \tag{35d}$$
$$\geq n[H(A|UE) - \varepsilon] - n(R_2 + R_p) + H(r_p|\mathcal{C}A^n E^n r_1 r_2) $$
$$\quad + I(r_2 r_p; E^n|\mathcal{C}r_1), \tag{35e}$$

where

- (35a) follows from the Markov chain $(A^n E^n) \,\multimapdotbothvert\, (\mathcal{C}r_1 r_2 r_p) \,\multimapdotbothvert\, (K_2 Z^m)$;
- (35d) is due to the Markov chain $(A^n E^n) \,\multimapdotbothvert\, U^n \,\multimapdotbothvert\, (r_1 \mathcal{C})$, the fact that the indices r_2 and r_p belong to sets of cardinality 2^{nR_2} and 2^{nR_p}, and the non-negativity of mutual information; and,
- (35e) stems from the lower bound found on Lemma 3 below.

Lemma 3: Given the codebook generation and encoding procedure of the scheme,

$$H(A^n|U^n E^n) \geq n[H(A|UE) - \varepsilon]. \tag{36}$$

Proof: Using well-known properties of typical sets, we have

$$H(A^n|U^n E^n) = - \sum_{\forall (u^n a^n e^n)} p(u^n a^n e^n) \log p(a^n|u^n e^n)$$
$$\geq - \sum_{(u^n a^n e^n) \in \mathcal{T}_\delta^n(UAE)} p(u^n a^n e^n) \log p(a^n|u^n e^n)$$
$$\geq \sum_{(u^n a^n e^n) \in \mathcal{T}_\delta^n(UAE)} p(u^n a^n e^n) n[H(A|UE) - \varepsilon^{(1)}]$$
$$\geq (1 - \varepsilon^{(2)}) n[H(A|UE) - \varepsilon^{(1)}]$$
$$\geq n[H(A|UE) - \varepsilon^{(3)}],$$

where in the last step we choose $\varepsilon^{(3)}$ large enough to have a lower bound. ∎

On the other hand, the "channel" term E_c writes:

$$E_c = H(r_p K_2|\mathcal{C}E^n Z^m r_1 r_2) - H(r_p|\mathcal{C}A^n E^n Z^m r_1 r_2 K_2)$$
$$= H(r_p K_2|\mathcal{C}Z^m r_1 r_2) - I(r_p K_2; E^n|\mathcal{C}Z^m r_1 r_2)$$
$$\quad - H(r_p|\mathcal{C}A^n E^n Z^m r_1 r_2 K_2). \tag{37}$$

The first term on the r.h.s. of (37) corresponds to the equivocation (of the *private* message, given the *common* message and the output of the channel) in the wiretap channel setting. Following the arguments of [3, Sec. IV] and [39, Sec. 2.3],

together with constraint (15) and Remark 14, we can easily prove the following lower bound[3]:

$$H(r_p K_2|\mathcal{C}Z^m r_1 r_2)$$
$$\geq n[R_p + R_{k_2} + R_f - (\eta + \epsilon)I(T;Z|Q) - \varepsilon'], \tag{38}$$

for sufficiently large n.

Gathering (32), (33), (35), (37), and (38), we have that

$$H(K|\mathcal{C}E^n Z^m)$$
$$\geq n[I(V;A|UE) - R_{k_1} + R_k - R_2 + R_f$$
$$\quad - (\eta + \epsilon)I(T;Z|Q) - \varepsilon''] + I(r_p; Z^m K_2|\mathcal{C}A^n E^n r_1 r_2)$$
$$\quad + I(r_2 r_p; E^n|\mathcal{C}r_1) - I(r_p K_2; E^n|\mathcal{C}Z^m r_1 r_2), \tag{39}$$

for some $\varepsilon'' > 0$. We now study the last two multi-letter terms on the r.h.s. of (39):

$$I(r_2 r_p; E^n|\mathcal{C}r_1) - I(r_p K_2; E^n|\mathcal{C}Z^m r_1 r_2)$$
$$= I(r_2 r_p; E^n|\mathcal{C}r_1) - I(r_2 r_p Z^m K_2; E^n|\mathcal{C}r_1)$$
$$\quad + I(r_2 Z^m; E^n|\mathcal{C}r_1) \tag{40a}$$
$$= -I(Z^m K_2; E^n|\mathcal{C}r_1 r_2 r_p) + I(r_2 Z^m; E^n|\mathcal{C}r_1) \tag{40b}$$
$$= I(r_2 Z^m; E^n|\mathcal{C}r_1) \tag{40c}$$
$$\geq 0, \tag{40d}$$

where

- (40c) stems from the Markov chain $E^n \,\multimapdotbothvert\, (\mathcal{C}r_1 r_2 r_p) \,\multimapdotbothvert\, (Z^m K_2)$; and,
- (40d) is due to the non-negativity of mutual information.

Inequality (39) may then be lower bounded as

$$H(K|\mathcal{C}E^n Z^m)$$
$$\geq n[I(V;A|UE) - R_{k_1} + R_k - R_2 + R_f$$
$$\quad - (\eta + \epsilon)I(T;Z|Q) - \varepsilon''] \tag{41a}$$
$$\geq n(R_k - \varepsilon''), \tag{41b}$$

where the last inequality holds if

$$R_{k_1} + R_2 - R_f \leq I(V;A|UE) - (\eta + \epsilon)I(T;Z|Q). \tag{42}$$

Finally,

$$\mathbb{E}[\mathsf{L}_k(\mathcal{C})] = I(K;E^n Z^m|\mathcal{C})$$
$$= H(K|\mathcal{C}) - H(K|\mathcal{C}E^n Z^m)$$
$$\leq n\varepsilon'',$$

and the key is asymptotically secure.

G. Sufficient Conditions

Putting all pieces together, we have proved that the proposed scheme allows the legitimate users to agree upon a key of rate R_k, while keeping it secret from the eavesdropper if

$$R_1 \leq S_1,$$
$$R_2 \leq S_2,$$
$$R_p \leq S_2 - R_2,$$
$$R_{k_1} \leq S_2 - R_2,$$
$$R_k \leq R_{k_1} + R_{k_2},$$

---

[3]Remark 14 assures that the indices r_1, r_2, and r_p are distributed almost uniformly, a condition that is necessary to invoke the result from the wiretap channel setting.

$$R_f < (\eta + \epsilon)I(T;Z|Q) - \epsilon_f \,,$$
$$S_1 > I(U;A) + \epsilon_1 \,,$$
$$S_2 > I(V;A|U) + \epsilon_2 \,,$$
$$R_p + R_{k_2} + R_f < (\eta + \epsilon)I(T;Y) - \delta - R_1 - R_2 \,,$$
$$R_p + R_{k_2} + R_f < (\eta + \epsilon)I(T;Y|Q) - \delta \,,$$
$$S_1 - R_1 < I(U;B) - \delta \,,$$
$$S_2 - R_2 - R_p < I(V;B|U) - \delta \,,$$
$$S_2 < H(A) - \xi \,,$$
$$S_1 - R_1 < I(U;E) - \delta \,,$$
$$S_2 - R_2 - R_{k_1} + R_f < I(V;E|U) + (\eta + \epsilon)I(T;Z|Q) - \delta' \,,$$
$$R_p + R_f > (\eta + \epsilon)I(T;Z|Q) + \varepsilon_1 \,,$$
$$R_{k_1} + R_2 - R_f \le I(V;A|UE) - (\eta + \epsilon)I(T;Z|Q) \,.$$

After applying Fourier-Motzkin elimination to this set of inequalities and taking $(n,m) \to \infty$, we obtain (8) subject to the conditions (9) and

$$I(T;Z|Q) \le I(T;Y|Q) \,, \tag{43a}$$
$$I(U;A|E) \le \eta I(Q;Y) \,, \tag{43b}$$
$$I(U;A|E) + I(V;A|UB) \le \eta I(T;Y) \,. \tag{43c}$$

The achievable region $\mathcal{R}_{\text{in}}$ is the convex hull of the union of this region over all joint probability distributions $p \in \mathcal{P}$, where the elements of $\mathcal{P}$ are defined in (16). We show next that the same achievable region is obtained by the convex hull of the union of the region defined by (8) and (9) over all $p \in \mathcal{P}$; therefore we prefer this more compact version.

The conditions (43b) and (43c) are redundant whenever $I(U;B) \le I(U;E)$, whereas if $(U,V) \sim p(u,v)$ are such that $I(U;B) > I(U;E)$ while satisfying (9) and (43), we see that

$$I(V;B) - I(V;E)$$
$$= I(V;B|U) - I(V;E|U) + I(U;B) - I(U;E)$$
$$> I(V;B|U) - I(V;E|U) \,.$$

This implies that a larger achievable secret key rate is obtained with $U = \emptyset$ and $V \sim p(v) = \sum_u p(u,v)$, which still satisfies (9) and (43). Similarly, we see that if (Q,T) are such that $I(T;Z|Q) > I(T;Y|Q)$ while satisfying (9) and (43),

$$\eta\big[I(T;Y|Q) - I(T;Z|Q)\big] + I(V;B|U) - I(V;E|U)$$
$$< I(V;B|U) - I(V;E|U) \,.$$

This implies that the achievable secret key rate is increased by choosing $Q = T$, which still satisfies (9) and (43). Therefore, the conditions (43) are redundant after the maximization and may be discarded.

We have shown thus far that, *averaged* over all possible codebooks, the probability of error (1), the key leakage (2), and the uniformity of the keys (3) become negligible as $(n,m) \to \infty$ if the conditions (8) and (9) hold true. Nonetheless, by applying the selection lemma [37, Lemma 2.2], we may conclude that there exists a *specific* sequence of codebooks such that the probability of error, the key leakage, and the uniformity of the keys tend to zero as $(n,m) \to \infty$.

The bounds on the cardinality of the alphabets $\mathcal{U}$, $\mathcal{V}$, $\mathcal{Q}$, and $\mathcal{T}$ follow from Fenchel–Eggleston–Carathéodory's theorem and the standard cardinality bounding technique [34, Appendix C]; therefore their proof is omitted. This concludes the proof of Theorem 2. ∎

VI. SUMMARY AND CONCLUDING REMARKS

In this work, we investigated the problem of secret key generation over a noisy channel in presence of correlated sources (independent of the main channel) at all terminals. We introduced a novel coding scheme using separate source and channel components –which shares common roots with our previous works [32], [33]. With the use of two description layers on the source observed at the encoder, this scheme improves upon the existing works in the literature which only rely on one layer of description.

The corresponding achievable secret key rate was shown to be optimal for all classes of less-noisy sources and channels (Propositions 1, 2, and 3). In Section IV, we compared the performance of the proposed scheme with a previously reported result for a simple binary model. Numerical computation of the corresponding bounds provided interesting insights on the regimes where the novel scheme outperforms the previous one.

This work, however, does not address the scenario where the sources and the noisy channel are correlated. The extension of the above mentioned result of Prabhakaran *et al.* [15] by using two description layers is a natural consequence. Indeed, this extension –posterior to the short version of the present work in [1]– has been recently addressed in [27]. Using two description layers as introduced here, the proposed achievable scheme recovers the present inner bound for $\eta = 1$ provided that the sources are independent of the channel.

APPENDIX A
PROOF OF THEOREM 1 (OUTER BOUND)

The outer bound is derived by following similar steps to those in [33, Thm. 4], which assumed $\eta = 1$. It is reproduced here for completeness.

Let (η, R_k) be an achievable tuple according to Definition 2, and $\epsilon > 0$. Then, there exists a $(2^{nR_k}, n, m)$ secret key code c_n with functions $\varphi(\cdot)$, $\psi_a(\cdot)$, and $\psi_b(\cdot)$ such that

$$X^m = \varphi(A^n, R_r) \,, \tag{44a}$$
$$K = \psi_a(A^n, R_r) \,, \tag{44b}$$
$$\hat{K} = \psi_b(B^n, Y^m) \,, \tag{44c}$$

that verify

$$\frac{m}{n} \le \eta + \epsilon \,, \tag{45a}$$
$$\Pr\big\{K \ne \hat{K}\big\} \le \epsilon \,, \tag{45b}$$
$$I(K; E^n Z^m) \le n\epsilon \,, \tag{45c}$$
$$nR_k - H(K) \le n\epsilon \,, \tag{45d}$$

where we have dropped the conditioning on the codebook c_n from (45b)–(45d) and all subsequent calculations for clarity. Before continuing, we present the following remark that is

useful to establish Markov chains between the random variables.

Remark 15: From the fact that random variables A_i, B_i, E_i are independent across time and the channel $X \mapsto (Y, Z)$ is memoryless, the joint distribution of $(K, A^n, B^n, E^n, X^m, Y^m, Z^m)$ can be written as follows. For each $i \in [1 : n]$ and each $j \in [1 : m]$, we have

$$
\begin{aligned}
&p(k, a^n, b^n, e^n, x^m, y^m, z^m) \\
&= p(a^{i-1}, b^{i-1}, e^{i-1}) \, p(a_i, b_i, e_i) \, p(a_{i+1}^n, b_{i+1}^n, e_{i+1}^n) \\
&\quad p(k, x^m | a^n) \, p(y^{j-1}, z^{j-1} | x^{j-1}) \, p(y_j, z_j | x_j) \\
&\quad p(y_{j+1}^m, z_{j+1}^m | x_{j+1}^m),
\end{aligned}
\tag{46}
$$

where $P_\varphi(x^m | a^n) = \sum_{\forall k} p(k, x^m | a^n)$ and $P_{\psi_a}(k | a^n) = \sum_{\forall x^m} p(k, x^m | a^n)$.

We may now carry on with the derivation of the outer bound. First consider,

$$
\begin{aligned}
nR_k &\leq H(K) + n\epsilon \tag{47a} \\
&= H(K | E^n Y^m) + I(K; E^n Y^m) + n\epsilon \tag{47b} \\
&\leq H(K | E^n Y^m) + I(K; E^n Y^m) \\
&\quad - I(K; E^n Z^m) + 2n\epsilon \tag{47c} \\
&= H(K | E^n Y^m) + I(K; Y^m | E^n) \\
&\quad - I(K; Z^m | E^n) + 2n\epsilon \tag{47d} \\
&\leq H(K | E^n Y^m) - H(K | B^n Y^m) \\
&\quad + I(K; Y^m | E^n) - I(K; Z^m | E^n) + 3n\epsilon \tag{47e} \\
&= \underbrace{I(K; B^n | Y^m) - I(K; E^n | Y^m)}_{R_s} \\
&\quad + \underbrace{I(K; Y^m | E^n) - I(K; Z^m | E^n)}_{R_c} + 3n\epsilon, \tag{47f}
\end{aligned}
$$

where

- (47a) stems from the uniformity of the keys (45d);
- (47c) is due to the security condition (45c); and,
- (47e) follows from (44), (45b), and Fano's inequality, $H(K | B^n Y^m) \leq n\epsilon$.

We now study separately the "source" term R_s and the "channel" term R_c. Hence,

$$
\begin{aligned}
R_s &= \sum_{i=1}^n I(K; B_i | Y^m B^{i-1}) - I(K; E_i | Y^m E_{i+1}^n) \\
&= \sum_{i=1}^n I(K; B_i | Y^m B^{i-1} E_{i+1}^n) \\
&\quad - I(K; E_i | Y^m B^{i-1} E_{i+1}^n) \tag{48a} \\
&= \sum_{i=1}^n I(V_i; B_i | U_i) - I(V_i; E_i | U_i) \tag{48b} \\
&= n[I(V_J; B_J | U_J J) - I(V_J; E_J | U_J J)] \tag{48c} \\
&= n[I(V; B | U) - I(V; E | U)], \tag{48d}
\end{aligned}
$$

where

- (48a) is due to Csiszár sum identity;
- (48b) follows from the definition of the auxiliary RVs $U_i = (Y^m B^{i-1} E_{i+1}^n)$ and $V_i = (KU_i)$;
- (48c) introduces the auxiliary RV J uniformly distributed over $[1 : n]$ and independent of all the other variables; and,

- (48d) stems from the definition of random variables $U = (U_J J)$, $V = (V_J J)$, $B = B_J$, and $E = E_J$.

This establishes the "source" term in (47f) with auxiliary RVs (U, V) that satisfy the following Markov chain

$$
U_i \; \circ\!\!-\!\!\circ \; V_i \; \circ\!\!-\!\!\circ \; A_i \; \circ\!\!-\!\!\circ \; (B_i E_i). \tag{49}
$$

The first part of (49) is trivial given the definition $V_i = (KU_i)$, whereas the second part follows from the i.i.d. nature of the sources and that they are correlated to the main channel only through the encoder's input (44a), see (46),

$$
(KY^m B^{i-1} E_{i+1}^n) \; \circ\!\!-\!\!\circ \; A_i \; \circ\!\!-\!\!\circ \; (B_i E_i). \tag{50}
$$

The "channel" term R_c can be single-letterized similarly,

$$
R_c = m[I(T; Y | Q) - I(T; Z | Q)], \tag{51}
$$

where we first define the auxiliary RVs $Q_i = (E^n Y^{i-1} Z_{i+1}^m)$ and $T_i = (KQ_i)$, we then introduce the auxiliary RV L uniformly distributed over $[1 : m]$, and we finally define $Q = (Q_L L)$, $T = (T_L L)$, $Y = Y_L$, and $Z = Z_L$. The auxiliary RVs in this term, i.e., (Q, T), satisfy the following Markov chain

$$
Q_i \; \circ\!\!-\!\!\circ \; T_i \; \circ\!\!-\!\!\circ \; X_i \; \circ\!\!-\!\!\circ \; (Y_i Z_i), \tag{52}
$$

where the nontrivial part is due to the memoryless property of the channel and (44b), provided the joint probability distribution satisfies (46). Since neither Q nor T appear on other parts of the outer bound, we may expand R_c as

$$
\begin{aligned}
R_c &= m \sum_{q \in \mathcal{Q}} p_Q(q) \left[I(T; Y | Q = q) - I(T; Z | Q = q) \right] \tag{53a} \\
&\leq m \max_{q \in \mathcal{Q}} \left[I(T; Y | Q = q) - I(T; Z | Q = q) \right] \tag{53b} \\
&= m[I(T^\star; Y) - I(T^\star; Z)], \tag{53c}
\end{aligned}
$$

where in the last step we choose auxiliary RV $T^\star \sim p_{T|Q}(\cdot | q)$.

Gathering (47), (48), (51), and (53), the rate of the secret key writes

$$
R_k \leq I(V; B | U) - I(V; E | U) + \frac{m}{n} \left[I(T; Y) - I(T; Z) \right] + 3\epsilon. \tag{54}
$$

If we let $(n, m) \to \infty$ and take arbitrarily small ϵ, we obtain the bound (5).

In order to obtain (6), we use the following Markov chain that is a consequence of (44a), provided the joint probability satisfies (46):

$$
(B^n E^n) \; \circ\!\!-\!\!\circ \; A^n \; \circ\!\!-\!\!\circ \; X^m \; \circ\!\!-\!\!\circ \; (Y^m Z^m). \tag{55}
$$

Due to the data processing inequality, we have

$$
I(A^n; Y^m) \leq I(X^m; Y^m) \leq m \, I(X; Y), \tag{56}
$$

where in the last inequality we use the memoryless property of the channel. Next, consider

$$
\begin{aligned}
I(A^n; Y^m) &= I(A^n B^n; Y^m) \tag{57a} \\
&\geq I(A^n; Y^m | B^n) \tag{57b} \\
&= I(A^n; KY^m | B^n) - I(A^n; K | B^n Y^m) \tag{57c} \\
&\geq I(A^n; KY^m | B^n) - n\epsilon \tag{57d} \\
&\geq n[I(A; V | B) - \epsilon], \tag{57e}
\end{aligned}
$$

where
- (57a) follows from the Markov chain (55); and,
- (57d) stems from $H(K|B^n Y^m) \leq n\epsilon$ due to (44) and (45b), and $H(K|A^n B^n Y^m) \geq 0$.

For the last step, i.e., (57e), consider

$$
\begin{aligned}
&I(KY^m; A^n|B^n) \\
&= I(KY^m; A^n E^n|B^n) && \text{(58a)} \\
&= \sum_{i=1}^{n} I(KY^m; A_i E_i|B^n A_{i+1}^n E_{i+1}^n) && \text{(58b)} \\
&\geq \sum_{i=1}^{n} I(KY^m B^{i-1} E_{i+1}^n; A_i E_i|B_i) && \text{(58c)} \\
&= \sum_{i=1}^{n} I(V_i; A_i E_i|B_i) && \text{(58d)} \\
&\geq \sum_{i=1}^{n} I(V_i; A_i|B_i) && \text{(58e)} \\
&= n\, I(V_J; A_J|B_J J) && \text{(58f)} \\
&= n\, I(V_J J; A_J|B_J) && \text{(58g)} \\
&= n\, I(V; A|B), && \text{(58h)}
\end{aligned}
$$

where
- (58a) stems from the Markov chain $(B^n E^n) \,\text{–}\!\circ\!\text{–}\, A^n \,\text{–}\!\circ\!\text{–}\, (KY^m)$;
- (58c) follows from the sources being i.i.d., i.e., $(A_i E_i) \,\text{–}\!\circ\!\text{–}\, B_i \,\text{–}\!\circ\!\text{–}\, (B^{i-1} B_{i+1}^n A_{i+1}^n E_{i+1}^n)$;
- (58d) is due to the auxiliary RV $V_i = (KY^m B^{i-1} E_{i+1}^n)$;
- (58f) introduces the auxiliary RV J uniformly distributed over $[1:n]$ and independent of all the other variables;
- (58g) follows from the independence of J and $(A_J B_J)$; and,
- (58h) stems from the definition of random variables $V = (V_J J)$, $B = B_J$, and $A = A_J$.

Putting (56) and (57) together, we obtain:

$$
I(V; A|B) \leq \frac{m}{n} I(X; Y) + \epsilon, \qquad (59)
$$

which gives the condition (6) as we let $(n, m) \to \infty$ and take an arbitrarily small ϵ.

Although the definition of the auxiliary RVs (TUV) used in the proof makes them arbitrarily correlated, the bounds (5) and (6) only depend on the *marginal* PDs $p(tx)$ and $p(uv|a)$. Consequently, we can restrict the set of possible joint PDs to (7), i.e., independent source and channel variables, and still achieve the maximum.

The bound on the cardinality of the alphabets $\mathcal{T}$, $\mathcal{U}$, and $\mathcal{V}$ follow from Fenchel–Eggleston–Carathéodory's theorem and the standard cardinality bounding technique [34, Appendix C]; therefore their proof is omitted. This concludes the proof of Theorem 1. ∎

APPENDIX B
PROOF OF PROPOSITION 5

For completeness, we first present the inner bound from [14, Thm. 4] but rewritten using the notation of the present work:

$$
R_k \leq \max_{p(x)p(v|a)} \left\{ I(V; B) - I(V; E) + \eta\, I(X; Y|Z) \right\} \quad \text{(60a)}
$$
$$
\text{subject to} \quad I(V; A|B) \leq \eta\, I(X; Y). \qquad \text{(60b)}
$$

In the sequel, we assume $\eta = 1$.

The main channel in the system model depicted in Fig. 2a is not only degraded but also Y equals X; thus, the last term on the r.h.s. of (60a) may be expanded as follows

$$
I(X; Y|Z) = H(X|Z) = H(X) + H(Z|X) - H(Z). \quad (61)
$$

Since X is the input of a BSC of parameter ζ and output Z, it is clear that

$$
I(X; Y|Z) \leq H(Z|X) = h_2(\zeta), \qquad (62)
$$

with equality if and only if $X \sim \mathcal{B}\left(\frac{1}{2}\right)$. Moreover, this choice of X maximizes the r.h.s. of (60b) and makes the condition redundant:

$$
I(V; A|B) \leq H(A|B) = \beta H(A) = \beta \leq 1 = H(X), \quad (63)
$$

given that $A \sim \mathcal{B}\left(\frac{1}{2}\right)$ and $0 \leq \beta \leq 1$.

It remains to be determined what the maximizing value of the first two terms on the r.h.s. of (60a) is. Let us first assume that B is *more capable* than E, i.e., $0 \leq \beta < h_2(\epsilon)$ according to Remark 12. Then, we may write

$$
\begin{aligned}
&I(V; B) - I(V; E) \\
&= I(A; B) - I(A; E) - \big[I(A; B|V) - I(A; E|V)\big] \\
&\leq I(A; B) - I(A; E) && \text{(64a)} \\
&= H(A|E) - H(A|B) && \text{(64b)} \\
&= h_2(\epsilon) - \beta, && \text{(64c)}
\end{aligned}
$$

where the inequality is due to $I(A; B|V) \geq I(A; E|V)$ for all $p(v, a)$ given the more capable assumption. The bound (64) holds with equality if and only if $V = A$. We also note that (64) is a monotonically decreasing function of β and it is zero when $\beta = h_2(\epsilon)$. For $\beta > h_2(\epsilon)$, the bound (64) is no longer valid; however, we can rightfully argue that as Bob's source degrades while Eve's remains the same, it is not possible to obtain more secret bits from the sources than for $\beta = h_2(\epsilon)$. Therefore, for $\beta > h_2(\epsilon)$,

$$
I(V; B) - I(V; E) \leq 0, \qquad (65)
$$

which holds with equality if and only if $V = \emptyset$.

Combining (60), (62), (64), and (65), we obtain the bound in (14). This concludes the proof of Proposition 5. ∎

APPENDIX C
PROOF OF LEMMA 1

According to the encoding procedure detailed in Section V-B, the index S is chosen uniformly among all the jointly typical codewords or, if there is no jointly typical codeword, uniformly on the whole codebook. We may thus characterize $p_{S_c}(1)$ as

$$
p_{S_c}(1) = \sum_{a^n \in \mathcal{T}_\delta^n(A)} \frac{p(a^n)}{\Pr\{\mathcal{T}_\delta^n(A)\}} \, \Upsilon_{a^n}, \qquad (66)
$$

where

$$
\Upsilon_{a^n} = \frac{\nu_1}{1 + \sum_{i=2}^{|\mathcal{S}|} \nu_i} + |\mathcal{S}|^{-1} \prod_{i=1}^{|\mathcal{S}|} (1 - \nu_i), \qquad (67)
$$

and ν_i is the event that the codeword $v^n(i)$ is jointly typical with the source sequence a^n, i.e.,

$$\nu_i \triangleq \mathbb{1}\big\{v^n(i) \in \mathcal{T}_\delta^n(V|u^n, a^n) \\ | \, v^n(i) \in \mathcal{T}_\delta^n(V|u^n), u^n \in \mathcal{T}_\delta^n(U|a^n)\big\} . \quad (68)$$

The first term in (67) distributes the probability of each sequence $a^n \in \mathcal{T}_\delta^n(A)$ uniformly among all the jointly typical codewords, while the second term in (67) distributes this probability uniformly among all codewords in $\mathcal{S}$, given that no one was jointly typical with a^n. It is not hard to see that the expected value of ν_i is

$$\mathbb{E}_\mathcal{C}[\nu_i] = \frac{|\mathcal{T}_\delta^n(V|u^n, a^n)|}{|\mathcal{T}_\delta^n(V|u^n)|} \triangleq \gamma , \quad (69)$$

for some $(u^n, a^n) \in \mathcal{T}_\delta^n(UA)$.

The expected value of (66) depends on the behavior of Υ_{a^n}. Each ν_i is a Bernoulli RV with $\mathbb{E}_\mathcal{C}[\nu_i] = \gamma$ and it is independent of the other ν_i's. Let us define

$$\nu = \sum_{i=2}^{|\mathcal{S}|} \nu_i , \quad (70)$$

then ν is a Binomial RV, and thus, for $j \in [0 : |\mathcal{S}| - 1]$,

$$p_\nu(j) = \binom{|\mathcal{S}| - 1}{j} \gamma^j (1 - \gamma)^{|\mathcal{S}| - 1 - j} . \quad (71)$$

After some manipulations, it is possible to show that

$$\mathbb{E}_\mathcal{C}\left[\frac{1}{1 + \nu}\right] = \frac{1 - (1 - \gamma)^{|\mathcal{S}|}}{\gamma |\mathcal{S}|} . \quad (72)$$

Hence,

$$\mathbb{E}_\mathcal{C}[\Upsilon_{a^n}] = \mathbb{E}_\mathcal{C}\left[\frac{\nu_1}{1 + \nu} + \frac{1}{|\mathcal{S}|} \prod_{i=1}^{|\mathcal{S}|} (1 - \nu_i)\right] = \frac{1}{|\mathcal{S}|} , \quad (73)$$

and consequently, the expected value of (66) is

$$\mathbb{E}_\mathcal{C}[p_{S_c}(1)] = \mathbb{E}_\mathcal{C}[\Upsilon_{a^n}] = |\mathcal{S}|^{-1} . \quad (74)$$

Noting that Υ_{a^n} and $\Upsilon_{a^{n\prime}}$ are independent variables given different sequences a^n and $a^{n\prime}$, and that $(\Upsilon_{a^n})^2 \leq \Upsilon_{a^n}$, we obtain

$$\mathbb{E}_\mathcal{C}[(p_{S_c}(1))^2] \leq 2^{-n[H(A) - \xi]} |\mathcal{S}|^{-1} + |\mathcal{S}|^{-2} , \quad (75)$$

for some $\xi > 0$. Therefore,

$$\mathrm{Var}[p_{S_c}(1)] \leq 2^{-n[H(A) - \xi]} |\mathcal{S}|^{-1} , \quad (76)$$

and in view of Chebyshev's inequality,

$$\Pr\big\{\big|p_{S_c}(1) - |\mathcal{S}|^{-1}\big| \geq \varepsilon_1 |\mathcal{S}|^{-1}\big\} \leq \varepsilon_1^{-2} 2^{-n[H(A) - \xi]} |\mathcal{S}| \\ = \varepsilon_1^{-2} 2^{-n[H(A) - S_2 - \xi]} .$$

This probability converges exponentially fast towards zero if $S_2 < H(A) - \xi$. This concludes the proof of Lemma 1. ∎

APPENDIX D
PROOF OF LEMMA 2

Let us modify the problem definition and then extend the scheme of Theorem 2 by introducing two virtual users who observe the source sequence E^n. The first user has access to the index r_1 as side information and we require that it decodes the codeword U^n. On the other hand, the second user has access to a different side information (which contains U^n) and we require that it decodes the codeword V^n. The keen reader can immediately see that we may bound the entropies in the statement of the lemma using Fano's inequality if the probability of error at the virtual users tend to zero.

Before proceeding, we note that the entropy in (34b) has (Z^m, K_2) in the conditioning. These variables are related to the *channel* and they affect the entropy of the *source*-related variable V^n through the knowledge they provide about the index r_p. In the sequel, we first characterize the decrease on the entropy of r_p and we then proceed to analyze the probability of error of the virtual users.

Let us introduce the random variable Υ, such that

$$\Upsilon \triangleq \mathbb{1}\{(Q^m, Z^m) \in \mathcal{T}_\delta^m(QZ)\} . \quad (77)$$

Then, using the binary variable Υ, it follows that

$$H(V^n|\mathcal{C}E^n Z^m U^n r_2 K_1 K_2) \\ \leq 1 + H(V^n|\mathcal{C}E^n Z^m U^n r_2 K_1 K_2 \Upsilon) \\ \leq 1 + H(V^n|\mathcal{C}E^n Z^m U^n r_2 K_1 K_2, \Upsilon = 1) + n S_2 \delta , \quad (78)$$

where the last inequality is due to $H(V^n|\mathcal{C}U^n) \leq n S_2$ and $\Pr\{\Upsilon = 0\} \leq \delta$.

In order to bound (78), we observe that, although $r_p \in [1 : 2^{n R_p}]$, the index has only a non-zero probability in a smaller subset of indices given the conditioning on Z^m, r_1 (known through U^n and $\mathcal{C}$), r_2, K_2, and $\Upsilon = 1$. For a specific codebook c_n (which defines the codewords $q^m(\cdot)$ and $t^m(\cdot)$), a channel output z^m, and the indices r_1, r_2, and k_2, let us define the set of possible indices r_p as

$$\mathcal{S}_R \triangleq \{r_p : t^m(r_1, r_2, r_p, k_2, r_f) \in \mathcal{T}_\delta^m(T|q^m(r_1, r_2), z^m) \\ \text{for some } r_f\} . \quad (79)$$

In principle, the size of this set depends on the particular codebook, channel output, and indices chosen. However, for sufficiently large n, the following lemma shows that the cardinality of $\mathcal{S}_R$ is close to its mean value for almost all codebooks.

Lemma 4: Let $\varepsilon_1, \varepsilon_5, \varepsilon_6 > 0$, and let χ be a function of the codebook c_n, the sequence z^m, and the indices r_1, r_2, and k_2 (not shown explicitly) defined as

$$\chi(\mathsf{c}_n, z^m) \triangleq \\ \mathbb{1}\big\{\big|S_R - [R_p + R_f - (\eta + \epsilon)I(T; Z|Q)]\big| \geq \varepsilon_5\big\} , \quad (80)$$

where $S_R \triangleq \frac{1}{n} \log |\mathcal{S}_R|$ and the set $\mathcal{S}_R$ is defined in (79). Then, $\Pr\{\chi(\mathcal{C}, Z^m) = 1\} \leq \varepsilon_6$ for sufficiently large n if $R_p + R_f > (\eta + \epsilon)I(T; Z|Q) + \varepsilon_1$.

Proof: See Appendix D-A. ∎

We may thus write,

$$H(V^n|\mathcal{C}E^nZ^mU^nr_2K_1K_2, \Upsilon = 1)$$
$$\le H(V^n|\mathcal{C}E^nZ^mU^nr_2K_1K_2, \Upsilon = 1, \chi = 0)$$
$$+ nS_2\varepsilon_6, \quad (81a)$$
$$= H(V^n|\mathcal{C}E^nZ^mU^nr_2K_1K_2, \Upsilon = 1, r_p \in \mathcal{S}_R, \chi = 0)$$
$$+ nS_2\varepsilon_6, \quad (81b)$$
$$\le H(V^n|\mathcal{C}E^nU^nr_2K_1, r_p \in \mathcal{S}_R, \chi = 0) + nS_2\varepsilon_6, \quad (81c)$$

where

- (81a) follows from $H(V^n|\mathcal{C}U^n) \le nS_2$ and Lemma 4, where χ denotes $\chi(\mathcal{C}, Z^m)$; and,
- (81b) is due to $r_p \in \mathcal{S}_R$ being a function of $(\mathcal{C}, Z^m, r_1, r_2, K_2, \Upsilon = 1)$.

In light of (81c), we define the side information of the second virtual user as $(u^n(s_1), r_2, k_1, r_p \in \mathcal{S}_R)$. According to the random codebook generation procedure, the number of codewords $V^n(\cdot)$ in a particular sub-bin $\tilde{\mathcal{B}}_2$ is $|\tilde{\mathcal{B}}_2(s_1, r_2, r_p)| = 2^{n(S_2-R_2-R_p)}$; thus, conditioned on $(u^n(s_1), r_2, r_p \in \mathcal{S}_R, \chi = 0)$, there are at most

$$\sum_{r_p \in \mathcal{S}_R} |\tilde{\mathcal{B}}_2(s_1, r_2, r_p)| = 2^{\log|\mathcal{S}_R|}2^{n(S_2-R_2-R_p)}$$
$$\le 2^{n[S_2-R_2+R_f-(\eta+\epsilon)I(T;Z|Q)+\varepsilon_5]}$$

distinct codewords $V^n(\cdot)$. These codewords will be evenly distributed in the sub-bins $\tilde{\mathcal{B}}_2$, given the symmetry of the random codebook generation and the independence in the creation of the sub-bins $\tilde{\mathcal{B}}_2$ and $\bar{\mathcal{B}}_2$, if

$$R_{k_1} < S_2 - R_2 + R_f - (\eta + \epsilon)I(T;Z|Q) + \varepsilon_5. \quad (82)$$

The reader may verify that this is true due to (18) and (42). Therefore, using the side information $(u^n(s_1), r_2, k_1, r_p \in \mathcal{S}_R)$, the second virtual user can construct a set of possible codewords $V^n(\cdot)$ defined as

$$\mathcal{S}_V \triangleq \bigcup_{r_p \in \mathcal{S}_R} \tilde{\mathcal{B}}_2(s_1, r_2, r_p) \cap \bar{\mathcal{B}}_2(s_1, r_2, k_1), \quad (83)$$

where the number of codewords is at most

$$|\mathcal{S}_V| \le 2^{n[S_2-R_2-R_{k_1}+R_f-(\eta+\epsilon)I(T;Z|Q)+\varepsilon_5]}. \quad (84)$$

We are finally ready to state the modified problem definition. Let virtual user 1 decode the codeword $u^n(s_1) \in \mathcal{B}_1(r_1)$ using the source sequence e^n, i.e., it looks for the unique index $s_1 \equiv \hat{s}_1$ such that $u^n(\hat{s}_1) \in \mathcal{B}_1(r_1)$ and

$$(u^n(\hat{s}_1), e^n) \in \mathcal{T}^n_\delta(UE). \quad (85)$$

The probability of error in decoding is arbitrarily small as $n \to \infty$ if

$$S_1 - R_1 < I(U;E) - \delta'. \quad (86)$$

On the other hand, let virtual user 2 decode the codeword $v^n(s_1, s_2) \in \mathcal{S}_V$ using the source sequence e^n, i.e., it looks for the unique index $s_2 \equiv \hat{s}_2$ such that $v^n(s_1, \hat{s}_2) \in \mathcal{S}_V$ and

$$(v^n(s_1, \hat{s}_2), e^n) \in \mathcal{T}^n_\delta(VE|u^n(s_1)). \quad (87)$$

The probability of error in decoding is arbitrarily small as $n \to \infty$ if

$$S_2 - R_2 - R_{k_1} + R_f - (\eta + \epsilon)I(T;Z|Q) + \varepsilon_5$$
$$< I(V;E|U) - \delta'. \quad (88)$$

To sum up, if (86) and (88) hold true, the probability of error in decoding at the virtual users is arbitrarily small as $n \to \infty$. Therefore, using Fano's inequality, we have

$$H(U^n|\mathcal{C}E^nr_1) \le n\epsilon_n, \quad (89a)$$
$$H(V^n|\mathcal{C}E^nU^nr_2K_1, r_p \in \mathcal{S}_R, \chi = 0) \le n\epsilon_n, \quad (89b)$$

where ϵ_n denotes a sequence such that $\epsilon_n \to 0$ as $n \to \infty$. Joining (78), (81), and (89), we recover the statement of the lemma. This concludes the proof of Lemma 2. ∎

A. Proof of Lemma 4

Before analyzing the set of possible indices r_p, let us first concentrate on characterizing the set of possible codewords $t^m(\cdot)$. Since the indices r_1, r_2, and k_2 are fixed, there are only $2^{n(R_p+R_f)}$ codewords to choose from. Moreover, given a specific codebook c_n generated according to the procedure from Section V-A, the indices r_1 and r_2 fix the codeword $q^m(r_1, r_2)$; thus, we simplify the notation and the codebook is composed of $q^m \in \mathcal{T}^m_\delta(Q)$ and $t^m(r) \in \mathcal{T}^m_\delta(T|q^m)$, where $r \in [1 : 2^{n(R_p+R_f)}]$. The set $\mathcal{S}_T$ of possible codewords $t^m(r)$ is then defined as

$$\mathcal{S}_T \triangleq \{t^m(r) : t^m(r) \in \mathcal{T}^m_\delta(T|q^m, z^m)\}. \quad (90)$$

Then, according to the random codebook generation,

$$\mathbb{E}_{\mathcal{C}Z^m}[|\mathcal{S}_T|] = \sum_{r=1}^{2^{n(R_p+R_f)}} \mathbb{E}_{\mathcal{C}Z^m}[\mathbb{1}\{T^m(r) \in \mathcal{T}^m_\delta(T|q^m, z^m)\}]$$
$$= 2^{n(R_p+R_f-\alpha)}, \quad (91)$$

where,

$$2^{-n\alpha} \triangleq \Pr\{T^m(1) \in \mathcal{T}^m_\delta(T|q^m, z^m)\} = \frac{|\mathcal{T}^m_\delta(T|q^m, z^m)|}{|\mathcal{T}^m_\delta(T|q^m)|},$$

for some $(q^m, z^m) \in \mathcal{T}^m_\delta(QZ)$. The value of α may be bounded using standard bounds for the cardinality of typical sets, yielding

$$(\eta + \epsilon)I(T;Z|Q) - \varepsilon_1 \le \alpha \le (\eta + \epsilon)I(T;Z|Q) + \varepsilon_1, \quad (92)$$

for some $\varepsilon_1 > 0$.

Similarly, we may calculate

$$\mathbb{E}_{\mathcal{C}Z^m}[|\mathcal{S}_T|^2] = 2^{2n(R_p+R_f-\alpha)} + 2^{n(R_p+R_f-\alpha)}(1 - 2^{-n\alpha}),$$

and finally,

$$\text{Var}[|\mathcal{S}_T|] \le 2^{n(R_p+R_f-\alpha)}. \quad (93)$$

We may now use Chebyshev's inequality to bound the value of $|\mathcal{S}_T|$,

$$\Pr\{||\mathcal{S}_T| - \mathbb{E}_{\mathcal{C}Z^m}[|\mathcal{S}_T|]| \ge \varepsilon_2 \mathbb{E}_{\mathcal{C}Z^m}[|\mathcal{S}_T|]\}$$
$$\le \varepsilon_2^{-2}2^{-n(R_p+R_f-\alpha)}, \quad (94)$$

for some $\varepsilon_2 > 0$. This probability tends to zero exponentially fast with n if $R_p + R_f > (\eta + \epsilon)I(T; Z|Q) + \varepsilon_1$. Taking the logarithm in the argument of the probability of (94) we obtain

$$\Pr\left\{\left|\frac{1}{n}\log|\mathcal{S}_T| - \beta\right| \geq \varepsilon_3\right\} \leq \varepsilon_4, \quad (95)$$

for some $\varepsilon_3 \geq \varepsilon_1 + \frac{1}{n}\log(1 + \varepsilon_2)$ and $\varepsilon_4 \geq \varepsilon_2^{-2}2^{-n(\beta-\varepsilon_1)}$, where

$$\beta \triangleq R_p + R_f - (\eta + \epsilon)I(T; Z|Q). \quad (96)$$

We note that (95) provides an estimate on the second virtual receiver's uncertainty on the actual transmitted codeword $T^m(\cdot)$, i.e., the set $\mathcal{S}_T$, rather than the index r_p, i.e., the set $\mathcal{S}_R$. In order to bound the latter, consider the following

$$\mathbb{E}_{\mathcal{C}Z^m}[\log|\mathcal{S}_R|]$$
$$\leq \mathbb{E}_{\mathcal{C}Z^m}\left[\log|\mathcal{S}_R|\,\big|\,|\mathcal{S}_R| \leq 2^{n(\beta+\varepsilon_3)}\right] + nR_p\varepsilon_4 \quad (97a)$$
$$\leq n(\beta - \varepsilon_5)p_s + n(\beta + \varepsilon_3)(1 - p_s) + nR_p\varepsilon_4 \quad (97b)$$
$$= n[\beta + \varepsilon_3 + \varepsilon_4 R_p - p_s(\varepsilon_3 + \varepsilon_5)], \quad (97c)$$

where

- (97a) follows from having at most 2^{nR_p} indices r_p, the fact that $|\mathcal{S}_R| \leq |\mathcal{S}_T|$ (e.g. some indices might be repeated), and (95); and,
- (97b) is due to the definition $p_s \triangleq \Pr\{|\mathcal{S}_R| \leq 2^{n(\beta-\varepsilon_5)}\}$, for some $\varepsilon_5 > 0$.

On the other hand, consider the following lower bound derived from (38):

$$\mathbb{E}_{\mathcal{C}Z^m}[\log|\mathcal{S}_R|] \geq H(r_p|\mathcal{C}Z^m r_1 r_2 K_2) \geq n(\beta - \varepsilon'), \quad (98)$$

where the first inequality is due to the definition of the set $\mathcal{S}_R$ in (79) and the fact that the uniform distribution maximizes the entropy. Joining (97) and (98) we obtain,

$$p_s \leq \frac{\varepsilon' + \varepsilon_3 + \varepsilon_4 R_p}{\varepsilon_3 + \varepsilon_5} < \varepsilon_6, \quad (99)$$

where the last inequality holds if $\varepsilon_5 \gg \max\{\varepsilon', \varepsilon_3, \varepsilon_4 R_p\}$. For a sufficiently large n, it is always possible to find such a ε_5. Finally, the lemma's statement is recovered using (95) (jointly with the fact that $|\mathcal{S}_R| \leq |\mathcal{S}_T|$) and (99). This concludes the proof of Lemma 4. ∎

REFERENCES

[1] G. Bassi, P. Piantanida, and S. Shamai, "Secret Key Generation over Noisy Channels with Common Randomness," in *2016 IEEE International Symposium on Information Theory (ISIT)*, Jul. 2016, pp. 510–514.

[2] A. D. Wyner, "The Wire-Tap Channel," *Bell System Technical Journal*, 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] C. E. Shannon, "Communication Theory of Secrecy Systems," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[5] A. Chorti, C. Hollanti, J.-C. Belfiore, and H. V. Poor, "Physical Layer Security: A Paradigm Shift in Data Confidentiality," in *Physical and Data-Link Security Techniques for Future Communication Systems*, ser. Lecture Notes in Electrical Engineering, M. Baldi and S. Tomasin, Eds. Springer International Publishing, 2016, vol. 358, pp. 1–15.

[6] P. Narayan and H. Tyagi, "Multiterminal Secrecy by Public Discussion," in *Foundations and Trends® in Communications and Information Theory*. Hanover, MA, USA: Now Publishers Inc., 2016, vol. 13, no. 2–3, pp. 129–275.

[7] U. M. Maurer, "Secret Key Agreement by Public Discussion from Common Information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[8] R. Ahlswede and I. Csiszár, "Common Randomness in Information Theory and Cryptography—Part I: Secret Sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[9] I. Csiszár and P. Narayan, "Common Randomness and Secret Key Generation with a Helper," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 344–366, Mar. 2000.

[10] ——, "Secrecy Capacities for Multiterminal Channel Models," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2437–2452, Jun. 2008.

[11] ——, "Secrecy Generation for Multiaccess Channel Models," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 17–31, Jan. 2013.

[12] A. A. Gohari and V. Anantharam, "Information-Theoretic Key Agreement of Multiple Terminals—Part I," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3973–3996, Aug. 2010.

[13] ——, "Information-Theoretic Key Agreement of Multiple Terminals—Part II: Channel Model," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3997–4010, Aug. 2010.

[14] A. Khisti, S. N. Diggavi, and G. W. Wornell, "Secret-Key Generation Using Correlated Sources and Channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 652–670, Feb. 2012.

[15] V. M. Prabhakaran, K. Eswaran, and K. Ramchandran, "Secrecy via Sources and Channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6747–6765, Nov. 2012.

[16] A. Bunin, P. Piantanida, and S. Shamai, "The Gaussian Wiretap Channel with Correlated Sources at the Terminals: Secret Communication and Key Generation," in *2016 ICSEE International Conference on the Science of Electrical Engineering*, Nov. 2016, pp. 1–5.

[17] S. Salimi, M. Skoglund, J. D. Golic, M. Salmasizadeh, and M. R. Aref, "Key Agreement over a Generalized Multiple Access Channel Using Noiseless and Noisy Feedback," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1765–1778, Sep. 2013.

[18] H. Tyagi, "Common Information and Secret Key Capacity," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5627–5640, Sep. 2013.

[19] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret Key Agreement: General Capacity and Second-Order Asymptotics," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3796–3810, Jul. 2016.

[20] T. A. Courtade and T. R. Halford, "Coded Cooperative Data Exchange for a Secret Key," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3785–3795, Jul. 2016.

[21] Y. Chen and A. J. H. Vinck, "Wiretap Channel with Side Information," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 395–402, Jan. 2008.

[22] W. Liu and B. Chen, "Wiretap Channel with Two-Sided Channel State Information," in *Conference Record of the 41st Asilomar Conference on Signals, Systems and Computers, 2007. ACSSC 2007*, Nov. 2007, pp. 893–897.

[23] S. I. Gelfand and M. S. Pinsker, "Coding for Channel with Random Parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[24] J. Muramatsu, "General Formula for Secrecy Capacity of Wiretap Channel with Noncausal State," in *2014 IEEE International Symposium on Information Theory*, Jun. 2014, pp. 21–25.

[25] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Wiretap Channels with Random States Non-Causally Available at the Encoder," *arXiv:1608.00743 [cs, math]*, Aug. 2016. [Online]. Available: http://arxiv.org/abs/1608.00743

[26] A. Zibaeenejad, "Key Generation over Wiretap Models with Non-Causal Side Information," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 7, pp. 1456–1471, Jul. 2015.

[27] A. Bunin, Z. Goldfeld, H. H. Permuter, S. Shamai, P. Cuff, and P. Piantanida, "Key and Message Semantic-Security over State-Dependent Channels," *IEEE Trans. Inf. Forensics Security*, pp. 1–1, 2018.

[28] M. Wiese, J. Nötzel, and H. Boche, "A Channel Under Simultaneous Jamming and Eavesdropping Attack—Correlated Random Coding Capacities Under Strong Secrecy Criteria," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3844–3862, Jul. 2016.

[29] 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.

[30] E. Ekrem and S. Ulukus, "Secure Lossy Source Coding with Side Information," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, Sep. 2011, pp. 1098–1105.

[31] J. Villard and P. Piantanida, "Secure Multiterminal Source Coding With Side Information at the Eavesdropper," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3668–3692, Jun. 2013.

[32] J. Villard, P. Piantanida, and S. Shamai, "Secure Transmission of Sources Over Noisy Channels With Side Information at the Receivers," *IEEE Trans. Inf. Theory*, vol. 60, pp. 713–739, Jan. 2014.

[33] G. Bassi, P. Piantanida, and S. Shamai, "The Wiretap Channel with Generalized Feedback: Secure Communication and Key Generation," *arXiv:1507.07091 [cs, math]*, Jul. 2015. [Online]. Available: http://arxiv.org/abs/1507.07091

[34] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[35] D. Cao and W. Kang, "Secret key generation from correlated sources and secure link," in *2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)*, Oct. 2017, pp. 1–5.

[36] U. Maurer and S. Wolf, "Information-Theoretic Key Agreement: From Weak to Strong Secrecy for Free," in *Advances in Cryptology — EUROCRYPT 2000*, B. Preneel, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 351–368.

[37] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[38] C. Nair, "Capacity Regions of Two New Classes of 2-Receiver Broadcast Channels," in *2009 IEEE International Symposium on Information Theory*, Jun. 2009, pp. 1839–1843.

[39] Y. Liang, H. V. Poor, and S. Shamai, "Information Theoretic Security," in *Foundations and Trends® in Communications and Information Theory*. Hanover, MA, USA: Now Publishers Inc., 2008, vol. 5, no. 4–5, pp. 355–580.