Explicit Wiretap Channel Codes via Source Coding, Universal Hashing, and Distribution Approximation, When the Channels' Statistics are Uncertain

Rémi A. Chou

***Abstract*— We consider wiretap channels with uncertainty on the eavesdropper channel under (i) noisy blockwise type II, (ii) compound, or (iii) arbitrarily varying models. We present explicit wiretap codes that can handle these models in a uni-fied manner and only rely on three primitives, namely source coding with side information, universal hashing, and distrib-ution approximation. Our explicit wiretap codes achieve the best known single-letter achievable rates, previously obtained non-constructively, for the models considered. Our results are obtained for strong secrecy, do not require a pre-shared secret between the legitimate users, and do not require any symmetry properties on the channel. An extension of our results to compound main channels is also derived via new capacity-achieving polar coding schemes for compound settings.**

***Index Terms*— Compound wiretap channel, arbitarily varying wiretap channel, polar codes, source coding, universal hashing.**

I. Introduction

THE wiretap channel [2] is a fundamental primitive to model eavesdropping at the physical layer [3], [4]. Beyond theoretical results that characterize the secrecy capacity for this model, significant progress has been made in the development of explicit wiretap codes for Wyner's wiretap channel [2]. Specifically, coding schemes based on low-density parity-check codes, e.g., [5], [6], and [7], polar codes, e.g. [8], [9], [10], [11], [12], [13], and [14], and invertible extractors, e.g., [15], [16], and [17], have been successfully developed for Wyner's model [2] or some of its special cases.

An assumption made by all the above references is that the eavesdropper channel statistics are perfectly known by the legitimate users. To model uncertainty, several models have been introduced: Type II models [18], [19], [20], where the eavesdropper can learn an arbitrary and unknown part of the legitimate sender codeword, and models where the eavesdropper channel statistics are not perfectly known but only known to belong to a given set. These latter models are useful when the physical location of the eavesdropper is uncertain from the point of view of the legitimate users, and include compound models [21], [22], where the channel statistics are known to be fixed for all channel uses, and arbitrarily varying models [23], [24], where the channel statistics change at each channel use.

Our contributions are summarized as follows. (i) We construct explicit wiretap codes that achieve the best known single-letter achievable rates, previously obtained non-constructively, when uncertainty holds on the eavesdropper channel under a noisy blockwise type II, compound, or arbitrarily varying model. (ii) We prove the sufficiency of three primitives to construct such wiretap codes: source coding with side information, universal hashing, and distribution approximation. (iii) We extend our results to the case where uncertainty holds on the main channel according to a compound model. (iv) We demonstrate that all the models considered in this paper can be handled in a unified manner by the same encoding and decoding schemes, up to an appropriate choice of parameters. We stress that our results are obtained for strong secrecy, do not require a pre-shared secret between the legitimate users, and do not require any symmetry properties on the channel.

Our approach consists in separately handling the reliability constraint and the security constraints. The reliability constraint is handled via a combination of source coding with side information and distribution approximation implemented with polar codes. The security constraints are handled with a combination of universal hashing and distribution approximation implemented via two-universal hash functions and polar codes, respectively. The main difficulty in our approach is to combine universal hashing and source coding with side information such that (i) non-symmetric and non-degraded channels can be handled, and (ii) the analysis of the security of the overall coding scheme is possible. (i) is performed via the idea of block-Markov coding as introduced in [25] and [26] with the following two important modifications to enable (ii): (1) Each encoding block of the block-Markov construction is constructed from L sub-blocks in which all the involved random variables have the same joint distribution across all sub-blocks. (2) The construction of each encoding block is such that the encoder output distribution approaches a fixed target distribution. In particular, these two points are key to analyzing the security of universal hashing via the leftover hash lemma [27], whose application in our analysis raises several additional challenges. First, while the leftover hash lemma proves a security guarantee on the output of a hash function, in our coding scheme, we need to prove a security guarantee on a message M that is not obtained as the output

Manuscript received 8 November 2021; revised 28 March 2022 and 28 July 2022; accepted 15 October 2022. Date of publication 31 October 2022; date of current version 7 December 2022. This work was supported by NSF under Grant CCF-1850227 and Grant CCF-2047913. An earlier version of this paper was presented in part at the 2018 IEEE International Symposium on Information Theory (ISIT) [DOI: 10.1109/ISIT.2018.8437777]. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Rafael Felix Schaefer.

The author is with the Department of Electrical Engineering and Computer Science, Wichita State University, Wichita, KS 67260 USA (e-mail: remi.chou@wichita.edu).

Digital Object Identifier 10.1109/TIFS.2022.3218414

of a hash function. To circumvent this difficulty, we prove the statistical equivalence between our coding scheme and another coding scheme where the message M is obtained as the output of a hash function. Second, because of the block-Markov construction, a precise study of the inter-dependencies between the encoding blocks is needed to evaluate the overall leakage when considering all the blocks jointly.

In Section III, we formally describe the model considered in this paper. In Section IV, we state our main results. In Section V, we describe our proposed coding scheme. The analysis of our coding scheme is presented in Sections VI, VII, VIII. In Section IX, we present an extension of our results to the case where uncertainty holds on the legitimate user channel under a compound model [3], [22]. Finally, in Section X, we provide concluding remarks.

II. Notation

For $a, b \in \mathbb{R}_+$, define $[a, b] \triangleq [\lfloor a \rfloor, \lceil b \rceil] \cap \mathbb{N}$. The components of a vector $X^{1:N}$ of size N are denoted with superscripts, i.e., $X^{1:N} \triangleq (X^i)_{i \in [1,N]}$. For any set $\mathcal{A} \subseteq [1, N]$, let $X^{1:N}[\mathcal{A}]$ be the components of $X^{1:N}$ whose indices are in $\mathcal{A}$. For two distributions p_{XY} and q_{XY} defined over $\mathcal{X} \times \mathcal{Y}$, define the variational distance between p_X and q_X as $\mathbb{V}(p_X, q_X) \triangleq \sum_{x \in \mathcal{X}} |p_X(x) - q_X(x)|$, the Kullback-Leibler divergence between p_X and q_X as $\mathbb{D}(p_X \| q_X)$, and the conditional Kullback-Leibler divergence between $p_{Y|X}$ and $q_{Y|X}$ as $\mathbb{E}_{p_X}[\mathbb{D}(p_{Y|X} \| q_{Y|X})] \triangleq \sum_{x \in \mathcal{X}} p_X(x) \mathbb{D}(p_{Y|X=x} \| q_{Y|X=x})$. Unless otherwise specified, capital letters denote random variables, whereas lowercase letters designate realizations of associated random variables, e.g., x is a realization of the random variable X. Let $\mathbb{1}\{\omega\}$ be the indicator function, which is equal to 1 if the predicate ω is true and 0 otherwise. For any $x \in \mathbb{R}$, define $[x]^+ \triangleq \max(0, x)$. Finally, $\mathrm{GF}(2^N)$ denotes a finite field of order 2^N.

III. Model and Known Results

Consider the finite alphabets $\mathcal{X} \triangleq \{0, 1\}$, $\mathcal{Y}$, and $(\mathcal{Z}_s)_{s \in \mathcal{S}}$, where $\mathcal{S}$ is a finite set. Consider also the conditional probabilities $(p_{YZ(s)|X})_{s \in \mathcal{S}}$. A wiretap channel is defined as a discrete memoryless channel with transition probability for one channel use $p_{YZ(s)|X}(y, z(s)|x)$ where $x \in \mathcal{X}$ is the channel input from the transmitter, $y \in \mathcal{Y}$ is the channel output observed by the legitimate receiver, $z(s) \in \mathcal{Z}_s$ is the channel output observed by the eavesdropper, $s \in \mathcal{S}$ is arbitrary, unknown to the legitimate users, and can potentially change for each channel use. In the following, we omit the index $s \in \mathcal{S}$ whenever $|\mathcal{S}| = 1$. Moreover, when the codeword $X^{1:N}$ is sent over the channel, in addition to the channel output $Z^{1:N}(\mathbf{s})$, $\mathbf{s} \in \mathcal{S}^N$, the eavesdropper has access to $X^{1:N}[\mathcal{S}] \triangleq (X^i)_{i \in \mathcal{S}}$, where $\mathcal{S} \subseteq [1, N]$ is chosen by the eavesdropper and such that $|\mathcal{S}| \leq \alpha N$ for some $\alpha \in [0, 1]$.

Definition 1: For $B \in \mathbb{N}$, define $\mathcal{B} \triangleq [1, B]$. For $b \in \mathcal{B}$ and $R_b \geq 0$, define $R \triangleq \sum_{b \in \mathcal{B}} R_b / B$. A $(2^{NR}, N, B)$ code has a rate R, operates over B encoding blocks, and consists for each encoding Block $b \in \mathcal{B}$ of

- *A message set $\mathcal{M}_b \triangleq [1, 2^{NR_b}]$.*
- *A stochastic encoding function $f_b : \mathcal{M}_b \to \mathcal{X}^N$, used by the transmitter to encode a message M_b, uniformly distributed over $\mathcal{M}_b$, into $X_b^{1:N} \triangleq f_b(M_b)$. The messages $M_{1:B} \triangleq (M_b)_{b \in \mathcal{B}}$ are assumed mutually independent.*
- *A deterministic decoding function used by the legitimate receiver $g_b : \mathcal{Y}^N \to \mathcal{M}_b$, to form M_b, an estimate of M_b, given the channel outputs $Y_b^{1:N}$. We write $M_{1:B} \triangleq (M_b)_{b \in \mathcal{B}}$.*

Definition 2: A rate R is achievable if there exists a sequence of $(2^{NR}, N, B)$ codes such that

$$\mathbb{P}[M_{1:B} = M_{1:B}] \xrightarrow{N \to \infty} 0,$$
$$\max_{\mathbf{s} \in \mathcal{S}^{NB}, \mathcal{A} \in \mathcal{A}} I\left(M_{1:B}; Z_{1:B}^{1:N}(\mathbf{s}), X_{1:B}^{1:N}[\mathcal{A}]\right) \xrightarrow{N \to \infty} 0,$$

where $\mathcal{A} \triangleq \{(\mathcal{A}_b)_{b \in \mathcal{B}} : \mathcal{A}_b \subseteq [1, N] \text{ and } |\mathcal{A}_b| = \alpha N, \forall b \in \mathcal{B}\}$, $(Z_b^{1:N}(\mathbf{s}_b), X_b^{1:N}[\mathcal{A}_b])$ corresponds to the random variables in Block $b \in \mathcal{B}$ for $\mathcal{A} = (\mathcal{A}_b)_{b \in \mathcal{B}} \in \mathcal{A}$ and $\mathbf{s}_b \in \mathcal{S}^N$, $X_{1:B}^{1:N}[\mathcal{A}] \triangleq (X_b^{1:N}[\mathcal{A}_b])_{b \in \mathcal{B}}$, and $Z_{1:B}^{1:N}(\mathbf{s}) \triangleq (Z_b^{1:N}(\mathbf{s}_b))_{b \in \mathcal{B}}$ for $\mathbf{s} = (\mathbf{s}_b)_{b \in \mathcal{B}} \in \mathcal{S}^{NB}$.

The supremum of such achievable rates is called secrecy capacity and denoted by C_s.

When $\alpha = 0$ and $|\mathcal{S}| = 1$, our model recovers Wyner's wiretap channel [2]. When $\alpha = 0$ and $\mathbf{s} = (\mathbf{s}_b)_{b \in \mathcal{B}} \in \mathcal{S}^{NB}$ is unknown to the legitimate users but all the components of $\mathbf{s}_b$, $b \in \mathcal{B}$, are identical, our model recovers a wiretap channel with a compound model for the eavesdropper's channel [21], [22]; the general model, as introduced in [21], with compound models for both the eavesdropper's channel and the main channel is treated in Section IX. When $\alpha = 0$ and $\mathbf{s} = (\mathbf{s}_b)_{b \in \mathcal{B}} \in \mathcal{S}^{NB}$ is unknown to the legitimate users, our model recovers a wiretap channel with an arbitrarily varying eavesdropper's channel [23]. When $\alpha > 0$ and $|\mathcal{S}| = 1$, our model recovers a special case of the wiretap channel of type II [18] when $p_{Z|X} = p_Z$ and $p_{Y|X}(y|x) = \mathbb{1}\{y = x\}$, $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$, a special case of the wiretap channel of type II with noisy main channel [19] when $p_{Z|X} = p_Z$, and a special case of the hybrid Wyner's/type II wiretap channel [20]. Specifically, the difference between our model and the models in [18], [20], and [19] is that, in our model, the eavesdropper observes a fraction α of each codeword $X_b^{1:N}$, $b \in \mathcal{B}$, whereas in [18], [20], and [19], the eavesdropper would be able to observe a fraction α of all the codewords considered jointly, i.e., $(X^{1:N})_{b \in \mathcal{B}}$. While the original type II constraint of [18] is stronger than a blockwise type II constraint, the latter constraint is relevant to model side-channel attacks where the eavesdropper is able to learn a bounded fraction of each codeword sent over the channel.

We now review the best known achievable rates for special cases of our model.

Theorem 1 [2], [28]: Suppose that $|\mathcal{S}| = 1$, and $\alpha = 0$. Then, the secrecy capacity is

$$C_s = \max_{U - X - (Y, Z) \atop |\mathcal{U}| \leq |\mathcal{X}|} [I(U; Y) - I(U; Z)]^+.$$

Theorem 2 [18]: Suppose that $|\mathcal{S}| = 1$, $p_{Z|X} = p_Z$, and for any $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $p_{Y|X}(y|x) = \mathbb{1}\{y = x\}$. Then,

$$C_s = 1 - \alpha.$$

Theorem 3 [19]: Suppose that $|\mathcal{S}| = 1$, *and* $p_{Z|X} = p_Z$. *Then,*

$$C_s = \max_{\substack{U-X-Y \\ |\mathcal{U}| \le |\mathcal{X}|}} [I(U;Y) - \alpha I(U;X)]^+.$$

Theorem 4 [20]: Suppose that $|\mathcal{S}| = 1$. *Then,*

$$C_s = \max_{\substack{U-X-(Y,Z) \\ |\mathcal{U}| \le |\mathcal{X}|}} [I(U;Y) - \alpha I(U;X) - (1-\alpha) I(U;Z)]^+.$$

Theorem 5 [21], [22]: Consider the wiretap channel with compound eavesdropper channel statistics, i.e., assume that $\mathbf{s} = (\mathbf{s}_b)_{b \in \mathcal{B}} \in \mathcal{S}^{N\mathcal{B}}$ *is unknown to the legitimate users but all the components of* $\mathbf{s}_b$, $b \in \mathcal{B}$, *are identical. Assume also that* $\alpha = 0$. *Then,*

$$C_s \ge \max_{\substack{p_{U,X}: \\ U-X-(Y,Z(s)) \ \forall s \in \mathcal{S} \\ |\mathcal{U}| \le |\mathcal{X}|}} \min_{s \in \mathcal{S}} [I(U;Y) - I(U;Z(s))]^+.$$

Moreover, for a degraded channel, i.e., when for all $s \in \mathcal{S}$, $X - Y - Z(s)$, *we have*

$$C_s = \max_{p_X} \min_{s \in \mathcal{S}} I(X;Y|Z(s)).$$

Theorem 6 [23], [24]: Consider the wiretap channel with arbitrarily varying eavesdropper channel, i.e., assume that $\mathbf{s} \in \mathcal{S}^{N\mathcal{B}}$ *is unknown to the legitimate users. Assume also that* $\alpha = 0$. *Define* $\overline{\mathcal{S}}$ *as the set of all the convex combinations of elements of* $\mathcal{S}$. *If there exists a best channel for the eavesdropper, i.e.,* $\exists s^* \in \overline{\mathcal{S}}$, $\forall s \in \mathcal{S}$, $X - Z(s^*) - Z(s)$, *then*

$$C_s \ge \max_{\substack{p_{U,X}: \\ U-X-(Y,Z(\bar{s})) \ \forall \bar{s} \in \overline{\mathcal{S}} \\ |\mathcal{U}| \le |\mathcal{X}|}} \min_{\bar{s} \in \overline{\mathcal{S}}} [I(U;Y) - I(U;Z(\bar{s}))]^+.$$

Moreover, if there exists a best channel for the eavesdropper and for all $\bar{s} \in \overline{\mathcal{S}}$, $X - Y - Z(\bar{s})$, *then*

$$C_s = \max_{p_X} \min_{\bar{s} \in \overline{\mathcal{S}}} I(X;Y|Z(\bar{s})).$$

Note that [20], [21], [22], [23], and [24] prove the existence of coding schemes that achieve the rates in Theorems 3-6 but do not provide explicit coding schemes. To the best of our knowledge, no explicit coding schemes that achieve the secrecy rates in Theorems 3-6 have been previously proposed.

More specifically, [12], [13], and [29] provided polar coding schemes that achieve the strong secrecy capacity for Wyner's wiretap channel, i.e., Theorem 1, with the following caveats: a pre-shared secret with negligible rate is required in [13] and [29], no efficient method is known to construct the codebooks in [12], and the existence of certain deterministic maps is needed in [13], similar to [30, Theorem 3]. Note that a main tool in [13] and [29] is block-Markov coding to support non-degraded and non-symmetric channels. Using techniques similar to [12], [13], and [29], including block-Markov coding, and ideas for compound channels without security constraints in [25] and [31], it is unclear to us how to extend existing polar coding schemes to the wiretap channel models of Theorems 2-5 and Theorems 7, 8, 11 because of the uncertainty on the eavesdropper's observations.

A different approach than polar coding to obtain wiretap codes for Wyner's wiretap channel is provided in [15], [16], and [32]. Specifically, these works construct wiretap codes using (i) capacity-achieving channel codes (without security constraint), and (ii) universal hashing [33], and have been the first works to provide efficient codes that asymptotically achieve optimal secrecy rates and strong secrecy for additive or symmetric and degraded wiretap channels. Reference [34] subsequently extended these constructions to any wiretap channels as in Theorem 1.

It is also worth noting that [35] proposed wiretap channel coding for Wyner's model using source coding with side information and universal hashing. It is, however, unclear to us how to directly translate the scheme of [35] to an efficient code construction without employing block-Markov coding for the part of the coding scheme that involves source coding with side information.

Our approach in this paper departs from the works in [15], [16], [32], and [34] because, instead of relying on channel codes, we rely on source codes to handle the reliability constraint, which allows us to use a block-Markov coding approach to handle non-symmetric and non-degraded channels. Our approach also departs from existing polar coding schemes, as our construction solely relies on polar coding results for source coding with side information, does not require the existence of certain maps, and does not require a pre-shared key to ensure strong secrecy. In addition to proposing the first explicit coding schemes that achieve the secrecy rates in Theorems 3-6 and Theorems 7, 8, 11, our coding approach also proves that all the models considered in this paper can be treated under a unified framework that only requires three primitives: (i) source coding with side information, (ii) universal hashing, and (iii) distribution approximation.

IV. Statement of Main Results

Our main results are the following theorems.

Theorem 7: If all the components of $\mathbf{s}_b$, $b \in \mathcal{B}$, *are identical, then the coding scheme of Section V achieves the secrecy rate*

$$\max_U \ I(U;Y) - \alpha I(U;X) - (1-\alpha) \max_{s \in \mathcal{S}} I(U;Z(s))^+,$$

where the maximum is taken over U *such that* $\forall s \in \mathcal{S}$, $U - X - (Y, Z(s))$ *and* $|\mathcal{U}| \le |\mathcal{X}|$.

Theorem 8: Assume that the components of $\mathbf{s}_b$, $b \in \mathcal{B}$, *are arbitrary. If there exists a best channel for the eavesdropper, then the coding scheme of Section V achieves the secrecy rate*

$$\max_U \ I(U;Y) - \alpha I(U;X) - (1-\alpha) \max_{\bar{s} \in \overline{\mathcal{S}}} I(U;Z(\bar{s}))^+,$$

where the maximum is taken over U *such that* $\forall \bar{s} \in \overline{\mathcal{S}}$, $U - X - (Y, Z(\bar{s}))$ *and* $|\mathcal{U}| \le |\mathcal{X}|$.

The proof of Theorem 7 is presented in two parts. First, in Section VI, the initialization phase, i.e., Algorithms 1, 2, is ignored and Theorem 7 is proved under the assumption that the legitimate users have a pre-shared key whose rate is negligible. Next, in Section VII, Theorem 7 is proved without this assumption by considering the initialization phase combined with Algorithms 3, 4. The proof of Theorem 8 is similar to the one of Theorem 7 and is discussed in Section VIII.

Finally, from Theorems 7 and 8, we conclude that the secrecy rates of Theorems 1-6 are achieved.

Note that we will also extend Theorem 7 to the case of a compound main channel in Theorem 11.

V. Coding Scheme

Our coding scheme consists of two phases: An initialization phase presented in Section V-B, and the actual secure communication phase presented in Section V-C. The initialization phase allows the legitimate users to share a secret key which is used in the second phase of the coding scheme. Both phases rely on three primitives presented in Section V-A.

In this section, for $s \in \mathcal{S}$, we consider an arbitrary joint distribution $q_{UXYZ(s)}$, $q_{UX}p_{YZ(s)|X}$ with $|\mathcal{U}| = |\mathcal{X}| = 2$ and such that $U - X - (Y, Z(s))$. Let K be a power of two, let $(U^{1:K}, X^{1:K})$ be distributed according to $q_{U^{1:K}X^{1:K}}$, $\prod_{i=1}^{K} q_{UX}$, and define $A^{1:K} := U^{1:K}G_K$, $V^{1:K} := X^{1:K}G_K$, where $G_K := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{\otimes \log K}$ is the matrix defined in [36]. Define also for $\delta_K := 2^{-K^\beta}$, $\beta \in]0, 1/2[$, the sets

$$\begin{aligned}
\mathcal{V}_U &:= \{i \in [1, K] : H(A^i | A^{1:i-1}) > 1 - \delta_K\},\\
\mathcal{H}_U &:= \{i \in [1, K] : H(A^i | A^{1:i-1}) > \delta_K\},\\
\mathcal{V}_{U|Y} &:= \{i \in [1, K] : H(A^i | A^{1:i-1}Y^{1:K}) > 1 - \delta_K\},\\
\mathcal{H}_{U|Y} &:= \{i \in [1, K] : H(A^i | A^{1:i-1}Y^{1:K}) > \delta_K\},\\
\mathcal{V}_X &:= \{i \in [1, K] : H(V^i | V^{1:i-1}) > 1 - \delta_K\},\\
\mathcal{V}_{X|U} &:= \{i \in [1, K] : H(V^i | V^{1:i-1}U^{1:K}) > 1 - \delta_K\}.
\end{aligned}$$

A. Primitives Used in the Coding Scheme

1) Primitive 1: Source coding (**SC**) with side information for the source $(\mathcal{U} \times \mathcal{Y}, q_{UY})$ [36]. Define the encoder $f^{\mathbf{SC}} := (f_1^{\mathbf{SC}}, f_2^{\mathbf{SC}})$ with

$$\begin{aligned}
f_1^{\mathbf{SC}}(A^{1:K}) &:= A^{1:K}[\mathcal{V}_{U|Y}],\\
f_2^{\mathbf{SC}}(A^{1:K}) &:= A^{1:K}[\mathcal{H}_{U|Y} \setminus \mathcal{V}_{U|Y}].
\end{aligned}$$

Then, define $g^{\mathbf{SC}}$ as the successive cancellation decoder of [36] such that if $A^{1:K} := g^{\mathbf{SC}}(f_1^{\mathbf{SC}}(A^{1:K}), f_2^{\mathbf{SC}}(A^{1:K}), Y^{1:K})$, then

$$\mathbb{P}[A^{1:K} = A^{1:K}] \leq K\delta_K. \tag{1}$$

Remark 1: We decompose $f^{\mathbf{SC}}$ in two parts $f_1^{\mathbf{SC}}$ and $f_2^{\mathbf{SC}}$ because $f_1^{\mathbf{SC}}(A^{1:K})$ can be shown to be almost uniform in divergence, e.g., [37, Lemma 8], which will be a useful property in our coding scheme analysis. Note, however, that the distribution of $(f_1^{\mathbf{SC}}(A^{1:K}) \| f_2^{\mathbf{SC}}(A^{1:K}))$ is not necessarily close to a uniform distribution.

2) Primitive 2: Universal hashing (**UH**) [38]. Let $c, d \in \mathbb{N}$ such that $d \leq c$, and define $\mathcal{S} := \{0, 1\}^c \setminus \{\mathbf{0}\}$. Then, define for $S \in \mathcal{S}$, $T \in \{0, 1\}^c$, $R \in \{0, 1\}^d$, $R^0 \in \{0, 1\}^{c-d}$

$$\begin{aligned}
f_S^{\mathbf{UH}}(R, R^0) &:= S^{-1} \cdot (R \| R^0),\\
g_S^{\mathbf{UH}}(T, d) &:= (S \cdot T)_d,
\end{aligned}$$

where $\cdot$ is the multiplication in $\mathrm{GF}(2^c)$ and $(\cdot)_d$ selects the d most significant bits, such that

$$g_S^{\mathbf{UH}}(f_S^{\mathbf{UH}}(R, R^0), d) = R.$$

By [16], $\mathcal{F} := \{g_S^{\mathbf{UH}}\}_{S \in \mathcal{S}}$ is a family of two-universal hash functions.

3) Primitive 3: Distribution approximation (**DA**) for $q_{A^{1:K}}$, the distribution of $A^{1:K} := U^{1:K}G_K$, where $U^{1:K}$ follows $q_{U^{1:K}} := \prod_{i=1}^{K} q_U$. Let $T^{1:|\mathcal{V}_U|}$ be a sequence of uniformly distributed bits over $\{0, 1\}^{|\mathcal{V}_U|}$. Then, define $A^{1:K}$ according to the distribution $p_{A^{1:K}} := \prod_{j=1}^{K} p_{A^j|A^{1:j-1}}$ with

$$p_{A^j|A^{1:j-1}}(a^j|a^{1:j-1}) := \begin{cases} \mathbb{1}\{a^j = T^j\} & \text{if } j \in \mathcal{V}_U \\ q_{A^j|A^{1:j-1}}(a^j|a^{1:j-1}) & \text{if } j \in \mathcal{V}_U^c \end{cases} \tag{2}$$

We write $A^{1:K} = f^{\mathbf{DA}}(T^{1:|\mathcal{V}_U|})$. Moreover, we have

$$\begin{aligned}
\mathbb{D}(q_{A^{1:K}} \| p_{A^{1:K}}) &\overset{(a)}{=} \mathbb{E}_{q_{A^{1:j-1}}}^{K} \mathbb{D}(q_{A^j|A^{1:j-1}} \| p_{A^j|A^{1:j-1}})_{j=1}\\
&\overset{(b)}{=} \sum_{j \in \mathcal{V}_U} (1 - H(A^j|A^{1:j-1})) \overset{(c)}{\leq} K\delta_K,
\end{aligned} \tag{3}$$

where *(a)* holds by the chain rule, *(b)* holds by (2), *(c)* holds by the definition of $\mathcal{V}_U$.

4) Variant of Primitive 3: Channel prefixing (**CP**) for the distribution $q_{X^{1:K}U^{1:K}} := \prod_{i=1}^{K} q_{XU}$. Given $U^{1:K}$ distributed according to $q_{U^{1:K}}$, define $V^{1:K}$ according to the distribution $p_{U^{1:K}V^{1:K}} := q_{U^{1:K}} \prod_{j=1}^{K} p_{V^j|V^{1:j-1}U^{1:K}}$ with

$$p_{V^j|V^{1:j-1}U^{1:K}}(v^j|v^{1:j-1}u^{1:K}) := \begin{cases} 1/2 & \text{if } j \in \mathcal{V}_{X|U} \\ q_{V^j|V^{1:j-1}U^{1:K}}(v^j|v^{1:j-1}u^{1:K}) & \text{if } j \in \mathcal{V}_{X|U}^c \end{cases} \tag{4}$$

We write $V^{1:K} = f^{\mathbf{CP}}(U^{1:K})$. Moreover, we have

$$\begin{aligned}
\mathbb{D}(q_{U^{1:K}V^{1:K}} \| p_{U^{1:K}V^{1:K}}) &\overset{(a)}{=} \mathbb{E}_{q_{U^{1:K}V^{1:j-1}}}^{K} \mathbb{D}(q_{V^j|V^{1:j-1}U^{1:K}} \| p_{V^j|V^{1:j-1}U^{1:K}})_{j=1}\\
&\overset{(b)}{=} \sum_{j \in \mathcal{V}_{X|U}} (1 - H(V^j|V^{1:j-1}U^{1:K})) \overset{(c)}{\leq} K\delta_K,
\end{aligned} \tag{5}$$

where *(a)* holds by the chain rule, *(b)* holds by (4), *(c)* holds by the definition of $\mathcal{V}_{X|U}$.

B. Coding Scheme: Phase I - Initialization

The legitimate users create a secret key with length l_{key}, which will be specified later in Section VII-B, with Algorithms 1 and 2, which operate over B_0 blocks of length $N := KL$, where $L, K \in \mathbb{N}$, and K is a power of two. We define $\mathcal{B}_0 := [1, B_0]$ and $\mathcal{L} := [1, L]$. In each Block $b \in \mathcal{B}_0$, the encoder forms the key Key_b with length $l := l_{\text{key}}/B_0$, as described in Algorithm 1. The encoder uses the following randomization sequences: $R_b^{\text{init}^0} := (R_{b,l}^{\text{init}^0})_{l \in \mathcal{L}}$, where $R_{b,l}^{\text{init}^0}$, $l \in \mathcal{L}$, is a sequence of uniformly distributed bits over $\{0, 1\}^{|\mathcal{H}_{U|Y}| - |\mathcal{V}_{U|Y}|}$, R_b^{init} is a sequence of uniformly distributed

Algorithm 1 Initialization at the Transmitter

Require: Randomization sequences $(R_b^{\text{init}})_{b\in[B_0]}$ and $(R_b^{\text{init}'})_{b\in[B_0]}$

1: **for** Block $b \in [B_0]$ **do**
2: **for** Sub-block $l \in [L]$ **do**
3: Define $A_{b,l}^{1:K}$, $f^{\mathbf{DA}}(R_{b,l}^{\text{loc}})$
4: Define $U_{b,l}^{1:K}$, $A_{b,l}^{1:K} G_K$
5: Define $V_{b,l}^{1:K}$, $f^{\mathbf{CP}}(U_{b,l}^{1:K})$
6: Define $X_{b,l}^{1:K}$, $V_{b,l}^{1:K} G_K$
7: **end for**
8: Transmit $X_b^{1:N}$, $\mathsf{k}_{l\in[L]} X_{b,l}^{1:K}$ over the channel
9: Let $Y_b^{1:N}$, $\mathsf{k}_{l\in[L]} Y_{b,l}^{1:K}$, $Z_b^{1:N}(\mathbf{s}_b)$, $\mathsf{k}_{l\in[L]} Z_{b,l}^{1:K}(\mathbf{s}_{b,l})$ denote the channel outputs
10: Transmit with a channel code [39] D_b, $\mathsf{k}_{l\in[L]} f_2^{\mathbf{SC}}(A_{b,l}^{1:K}) \oplus R_{b,l}^{\text{init}'} \; \mathsf{k}\, f_1^{\mathbf{SC}}(A_{b,l}^{1:K})$, where $\oplus$ denotes modulo 2 addition
11: Define $U_b^{1:N}$, $\mathsf{k}_{l\in[L]} U_{b,l}^{1:K}$
12: Define Key_b, $g^{\mathbf{UH}}_{R_b^{\text{init}}}(U_b^{1:N}, l^c_{\text{key}})$
13: **end for**

Algorithm 2 Initialization Phase at the Receiver

Require: $(R_b^{\text{init}})_{b\in[B_0]}$ and $(R_b^{\text{init}'})_{b\in[B_0]}$

1: **for** Block $b \in [B_0]$ **do**
2: Form an estimate $\widehat{D}_b$ of D_b
3: **for** Sub-block $l \in [L]$ **do**
4: Given $(\widehat{D}_b, R_b^{\text{init}'})$ and Line 10 of Algorithm 1, form an estimate of $(f_1^{\mathbf{SC}}(A_{b,l}^{1:K}), f_2^{\mathbf{SC}}(A_{b,l}^{1:K}))$ and denote this estimate by $(A_{b,l}^{1:K}[\mathcal{V}_{U|Y}], A_{b,l}^{1:K}[\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|Y}])$
5: Form an estimate of $A_{b,l}^{1:K}$ as
$$A_{b,l}^{1:K}, \; g^{\mathbf{SC}}(A_{b,l}^{1:K}[\mathcal{V}_{U|Y}], A_{b,l}^{1:K}[\mathcal{H}_{U|Y}\backslash\mathcal{V}_{U|Y}], Y_{b,l}^{1:K})$$
6: Form $U_{b,l}^{1:K}$, $A_{b,l}^{1:K} G_K$ an estimate of $U_{b,l}^{1:K}$
7: **end for**
8: Form $U_b^{1:N}$, $\mathsf{k}_{l\in[L]} U_{b,l}^{1:K}$ an estimate of $U_b^{1:N}$
9: Form $\widehat{\text{Key}}_b = g^{\mathbf{UH}}_{R_b^{\text{init}}}(U_b^{1:N}, l^c_{\text{key}})$ an estimate of Key_b
10: **end for**

bits over $\mathcal{R}^{\text{init}}$, $\{0,1\}^N \backslash \{\mathbf{0}\}$. The encoder also uses the local randomness $(R_{b,l}^{\text{loc}})_{l\in[L]}$, where $R_{b,l}^{\text{loc}}$, $l \in [L]$, is a sequence of uniformly distributed bits over $\{0,1\}^{|\mathcal{V}_U|}$.

Remark 2: In Line 10 of Algorithm 1, note that the channel code [39] requires a uniformly distributed message. While $\mathsf{k}_{l\in[L]} A_{b,l}^{1:K}[\mathcal{H}_{U|Y}]$ *is not a sequence of uniformly distributed bits,* D_b *is a sequence of uniformly distributed bits over* $\{1, 2^{L|\mathcal{H}_{U|Y}|}\}$.

High-level description of the initialization phase: The initialization phase is depicted in Figure 1 and consists in B_0 communication blocks. All the communication blocks are independent, and each Block $b \in [B_0]$ will lead to the exchange of a key Key_b between the legitimate users, which will be shown to be secret from the eavesdropper. Additionally,

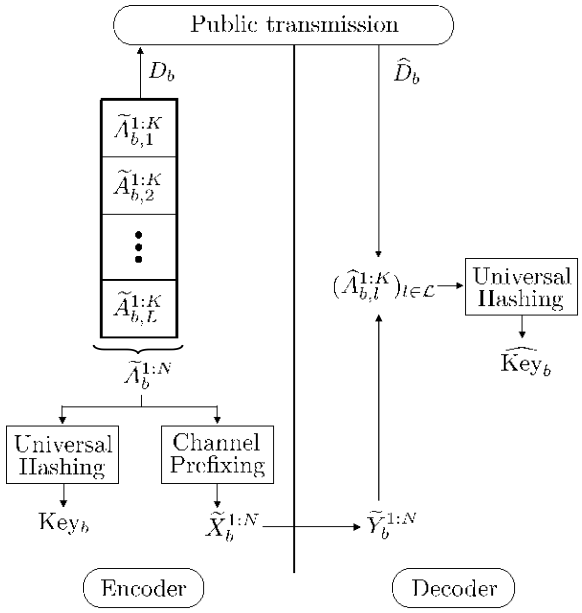

Fig. 1. Initialization phase for Block $b \in [B_0]$. The encoder creates $A_b^{1:N}$, which is made of L sub-blocks $(A_{b,l}^{1:K})_{l\in[L]}$. Then, from $A_b^{1:N}$, the encoder creates Key_b (by universal hashing), and the codeword $X_b^{1:N}$ (via channel prefixing), which is sent over the channel and whose noisy observation by the legitimate receiver is $Y_b^{1:N}$. The decoder creates an estimate of $A_b^{1:N}$ from $Y_b^{1:N}$ and an estimate of D_b, which is sent to him via a channel code, as described in Line 10 of Algorithm 1. Finally, the decoder creates $\widehat{\text{Key}}_b$, an estimate of Key_b, from his estimate of $A_b^{1:N}$.

B_0 is chosen such that the length of the keys $(\text{Key}_b)_{b\in[B_0]}$ is sufficiently large to be used in the main coding scheme, which is described in the next section and allows the exchange of a secret message between the legitimate users. It will also be shown that the initialization phase considered jointly with the main coding scheme has a negligible effect on the overall communication rate and the overall information leakage to the eavesdropper.

Consider Block $b \in [B_0]$ in Algorithm 1. As described in Lines 3-4, the encoder creates $(U_{b,l}^{1:K})_{l\in[L]}$ such that the distribution of $(U_{b,l}^{1:K})_{l\in[L]}$ is close to the product distribution $q_{U^{1:N}}$. Then, as described in Lines 5-6, channel prefixing is performed to create from $(U_{b,l}^{1:K})_{l\in[L]}$ the codewords $(X_{b,l}^{1:K})_{l\in[L]}$ that are sent over the channel, and whose noisy observations at the legitimate receiver are $(Y_{b,l}^{1:K})_{l\in[L]}$. Additionally, the key Key_b is formed from $(U_{b,l}^{1:K})_{l\in[L]}$ through universal hashing, as described in Line 12. As shown later, secrecy of the key is ensured via an appropriate choice of the hash function output length. As described in Line 10, the encoder sends D_b to the legitimate receiver using a regular channel code (without security guarantees) - see also Remark 2.

Finally, as described in Lines 2-7 of Algorithm 2, upon estimating D_b, the legitimate receiver forms an estimate of $(U_{b,l}^{1:K})_{l\in[L]}$ from $(Y_{b,l}^{1:K})_{l\in[L]}$. Then, as described in Line 9 of Algorithm 2, from the estimate of $(U_{b,l}^{1:K})_{l\in[L]}$, the legitimate receiver creates an estimate of Key_b.

C. Coding Scheme: Phase II - Secure Communication

The encoding scheme operates over B blocks of length $N = KL$, where $L, K \in \mathbb{N}$ and K is a power of two.

We define $\mathcal{B} \triangleq [1, \lfloor B \rfloor]$ and $\mathcal{L} \triangleq [1, L]$. Encoding at the transmitter and decoding at the receiver are described in Algorithms 3 and 4, respectively. In each block $b \in \mathcal{B}$, the transmitter encodes, as described in Algorithm 3, a message M_b uniformly distributed over $[1, 2^{|M_b|}]$ and represented by a binary sequence with length

$$|M_b| \triangleq \begin{cases} |M_1| \text{ if } b = 1 \\ |M_1| - L|V_{U|Y}| \text{ otherwise} \end{cases}.$$

Algorithms 3 and 4 depend on the parameter

$$r \triangleq |M_1|, \tag{6}$$

which will be specified later.

In each block $b \in \mathcal{B}$, as described in Algorithm 3, the encoder uses the local randomness R_b^0, a binary randomization sequence uniformly distributed over $[1, 2^{|R_b^0|}]$. The sequences $R_{1:B}^0 \triangleq (R_b^0)_{b\in\mathcal{B}}$ are mutually independent. The length of the sequences R_b^0 is defined for $b \in \mathcal{B}$ as $|R_b^0| \triangleq L|V_U| - r$. In each block $b \in \mathcal{B}$, the encoder also uses, as described in Algorithm 3, R_b, a binary randomization sequence with length $L|V_U|$, uniformly distributed over $\mathcal{R} \triangleq \{0,1\}^{L|V_U|} \setminus \{\mathbf{0}\}$. The sequences $R_{1:B} \triangleq (R_b)_{b\in\mathcal{B}}$ are mutually independent. Moreover, it is assumed that $M_{1:B}$, $R_{1:B}$, and $R_{1:B}^c$ are mutually independent.

Remark 3: In Algorithm 3, observe that $T_b^{1:|V_U|L}$, $b \in \mathcal{B}$, is uniformly distributed over $\{0,1\}^{|V_U|L}$ because $(M_b \| M_b^0 \| R_b^0)$ is uniformly distributed over $\{0,1\}^{|V_U|L}$ and independent of R_b. Hence, the L random variables $(T_{b,l}^{1:|V_U|})_{l\in\mathcal{L}}$ are uniformly distributed over $\{0,1\}^{|V_U|}$ and independent. When the elements of $\mathbf{s}_b$ are all equal to s, then, by construction, the conditional probability $p_{T_{b,l}^{1:|V_U|}}(\cdot|s)$ is the same for all $l \in \mathcal{L}$, and the L pairs $((T_{b,l}^{1:|V_U|}, Z_{b,l}^{1:K}(s)))_{l\in\mathcal{L}}$ are independently and identically distributed according to the joint distribution $p_{T_{b,1}^{1:|V_U|} Z_{b,1}^{1:K}(s)}$.

Remark 4: In Algorithm 3, consider $X_{b,l}^{1:K}[\mathcal{A}_{b,l}]$, $b \in \mathcal{B}$, $l \in \mathcal{L}$, where for all $l \in \mathcal{L}$, $\mathcal{A}_{b,l} \in [1, K]$ and $|\mathcal{A}_{b,l}| = \alpha N$ such that $X^{1:N}[\mathcal{A}_b] \triangleq (X_{b,l}^{1:K}[\mathcal{A}_{b,l}])_{l\in\mathcal{L}}$ corresponds to the αN symbols of the codewords emitted at the transmitter that the eavesdropper has chosen to have access to. Similar to Remark 3, the L triplets $((T_{b,l}^{1:|V_U|}, X_{b,l}^{1:K}[\mathcal{A}_{b,l}], Z_{b,l}^{1:K}(\mathbf{s}_{b,l})))_{l\in\mathcal{L}}$ are independent, however, they are not necessarily identically distributed because the components of $\mathbf{s}_{b,l}$ are arbitrary, and because the sets $(\mathcal{A}_{b,l})_{l\in\mathcal{L}}$ are arbitrarily chosen by the eavesdropper.

High-level description of the coding scheme: We depict in Figure 2 how codewords are created at the transmitter. Note that there exists an interdependence between two consecutive encoding blocks since M_b^0, $b \in [2, B]$, used in Block b, is obtained from Block $b-1$, as described in Line 3 of Algorithm 3.

Consider Block $b \in \mathcal{B}$ of Algorithm 3. The encoder starts by creating $T_b^{1:|V_U|L}$ via universal hashing applied on the sequence created by M_b^0, the secret message M_b, and the local randomness R_b, as described in Line 4. Next, $T_b^{1:|V_U|L}$ is broken down into L pieces with same length in Line 6, from which the encoder creates L sub-blocks $(A_{b,l}^{1:K})_{l\in\mathcal{L}}$, as described in Line 7. Then, from $(A_{b,l}^{1:K})_{l\in\mathcal{L}}$, the codewords $(X_{b,l}^{1:K})_{l\in\mathcal{L}}$, are obtained via channel prefixing, as described in Lines 8-10. The codewords $(X_{b,l}^{1:K})_{l\in\mathcal{L}}$ are sent over the channel and their noisy observations at the legitimate receiver are denoted by $(Y_{b,l}^{1:K})_{l\in\mathcal{L}}$. Note that the L sub-blocks $(A_{b,l}^{1:K})_{l\in\mathcal{L}}$ are created such that their distri-bution is close to the product distribution $q_{A^{1:N}}$. A crucial

Algorithm 3 Encoding

Require: Randomization sequences $(R_b)_{b\in\mathcal{B}}$, $(R_b^0)_{b\in\mathcal{B}}$, and messages $(M_b)_{b\in\mathcal{B}}$

1: Define $M_1^c \triangleq \emptyset$
2: **for** Block $b \in \mathcal{B}$ **do**
3: Define $M_b^0 \triangleq \| f_1^{\mathbf{SC}}(A_{b-1,l}^{1:K})_{l\in\mathcal{L}}$ if $b = 1$
4: Define $T_b^{1:|V_U|L} \triangleq f_{R_b}^{\mathbf{UH}}(M_b, M_b^0 \| R_b^0)$
5: **for** Sub-block $l \in \mathcal{L}$ **do**
6: Consider the notation $T_{b,l}^{1:|V_U|} \triangleq T_b^{(l-1)|V_U|+1:l|V_U|}$
7: Define $A_{b,l}^{1:K} \triangleq f^{\mathbf{DA}}(T_{b,l}^{1:|V_U|})$
8: Define $U_{b,l}^{1:K} \triangleq A_{b,l}^{1:K} G_K$
9: Define $V_{b,l}^{1:K} \triangleq f^{\mathbf{CP}}(U_{b,l}^{1:K})$
10: Define $X_{b,l}^{1:K} \triangleq V_{b,l}^{1:K} G_K$
11: **end for**
12: Transmit $X_b^{1:N} \triangleq (X_{b,l}^{1:K})_{l\in\mathcal{L}}$ over the channel
13: Let $Y_b^{1:N} \triangleq (Y_{b,l}^{1:K})_{l\in\mathcal{L}}$, $Z_b^{1:N}(\mathbf{s}_b) \triangleq (Z_{b,l}^{1:K}(\mathbf{s}_{b,l}))_{l\in\mathcal{L}}$ denote the channel outputs
14: **end for**
15: Using a pre-shared secret, apply a one-time pad to $(f_2^{\mathbf{SC}}(A_{b,l}^{1:K}))_{l\in\mathcal{L}, b\in\mathcal{B}}$, and $(f_1^{\mathbf{SC}}(A_{B,l}^{1:K}))_{l\in\mathcal{L}}$, then transmit the result with a channel code [39].

Algorithm 4 Decoding

Require: $(R_b)_{b\in\mathcal{B}}$, $(f_2^{\mathbf{SC}}(A_{b,l}^{1:K}))_{l\in\mathcal{L}, b\in\mathcal{B}}$, $(f_1^{\mathbf{SC}}(A_{B,l}^{1:K}))_{l\in\mathcal{L}}$

1: Define $A_{B,l}^{1:K}[\mathcal{V}_{U|Y}] \triangleq f_1^{\mathbf{SC}}(A_{B,l}^{1:K})$ for any $l \in \mathcal{L}$
2: **for** Block $b \in \mathcal{B}$ from $b = B$ to $b = 1$ **do**
3: **for** $l \in \mathcal{L}$ **do**
4: Form an estimate of $A_{b,l}^{1:K}$ as

$$A_{b,l}^{1:K} \triangleq g^{\mathrm{SC}}(A_{b,l}^{1:K}[\mathcal{V}_{U|Y}], f_2^{\mathbf{SC}}(A_{b,l}^{1:K}), Y_{b,l}^{1:K})$$

5: **end for**
6: From Line 7 in Algorithm 3, determine an estimate of $T_b^{1:|V_U|L}$ as

$$T_b^{1:|V_U|L} \triangleq \| A_{b,l}^{1:K}[\mathcal{V}_U]$$

7: From Line 4 in Algorithm 3, form an estimate of $(M_b \| M_b^0 \| R_b^0)$ as

$$(M_b \| M_b^0 \| R_b^c) \triangleq R_b \, T_b^{1:|V_U|L}$$

8: From Line 3 in Algorithm 3 and M_b^c, form $A_{b-1,l}^{1:K}[\mathcal{V}_{U|Y}]_{l\in\mathcal{L}}$ an estimate of $f_1^{\mathbf{SC}}(A_{b-1,l}^{1:K})_{l\in\mathcal{L}}$
9: **end for**

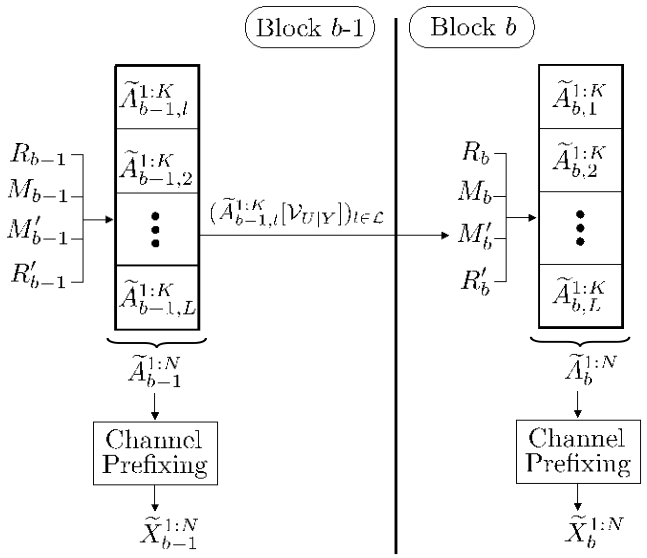

Fig. 2. In Block $b \in \mathcal{B}$, $A_b^{1:N}$ is made of L sub-blocks $(A_{b,l}^{1:K})_{l\in\mathcal{L}}$, which are constructed from R_b (randomness for universal hashing), M_b (secret message), M_b^0 (a part of $A_{b-1}^{1:N}$ from Block $b-1$), and R_b^0 (local randomness). The construction of M_b^0 in Line 3 of Algorithm 3 creates a dependency between Block $b \in \{2, \ldots, B\}$ and Block $b-1$. In Block $b \in \mathcal{B}$, the codeword $X_b^{1:N}$, to be sent over the channel, is then obtain via channel prefixing from $A_b^{1:N}$.

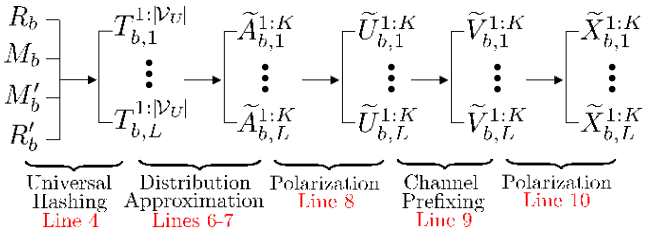

Fig. 3. Summary of the steps in Algorithm 3 to obtain the codeword $X_b^{1:N}$, $\mathsf{k}_{l\in\mathcal{L}}(X_{b,l}^{1:K})$ from R_b (randomness for universal hashing), M_b (secret message), M_b^0 (a part of Block $b-1$), and R_b^0 (local randomness). Line 7 (distribution approximation) describes the creation of the L sub-blocks $(A_{b,l}^{1:K})_{l\in\mathcal{L}}$ from $T_b^{1:|\mathcal{V}_U|L}$, and ensures that their distribution is close to the product distribution $q_{A^{1:N}}$, which will be a crucial fact to analyze the information leakage of the coding scheme.

point to ensure this property comes from the uniformity of M^0, i.e., the uniformity of $A^{1:K}_{b,l}[\mathcal{V}_{U|Y}]$, $l \in \mathcal{L}$, which follows from Line 7 and the property $\mathcal{V}_{U|Y} \supseteq \mathcal{V}_U$. Finally, as described in Line 15, using a pre-shared secret (obtained from the initialization phase in Section V-B), the encoder applies a one-time pad to $(A_{b,l}^{1:K}[H_{U|Y}\backslash\mathcal{V}_{U|Y}])_{l\in\mathcal{L},b\in\mathcal{B}}$, and $(A_{B,l}^{1:K}[\mathcal{V}_{U|Y}])_{l\in\mathcal{L}}$, and sends the result to the legitimate receiver with a channel code [39]. This step is done for technical reasons: $(A_{b,l}^{1:K}[H_{U|Y}\backslash\mathcal{V}_{U|Y}])_{l\in\mathcal{L},b\in\mathcal{B}}$ are not uniformly distributed and could not be included in the definition of M_b, $b \in \mathcal{B}$, as our analysis relies on the uniformity of M^0, $b \in \mathcal{B}$. However, as shown later, the length of $(A_{b,l}^{1:K}[H_{U|Y}\backslash\mathcal{V}_{U|Y}])_{l\in\mathcal{L},b\in\mathcal{B}}$ is negligible compared to NB such that the overall communication rate is not affected. It will also be shown that this has a negligible effect on the overall information leakage to the eavesdropper.

In a given block $b \in \mathcal{B}$, we depict in Figure 3 a summary of the different phases in Algorithm 3 through which the encoder output is obtained from the local randomness R_b^0, the secret message M_b, the randomness R_b used for universal hashing, and M_b^0. Note that R_b needs to be shared between the legitimate users but does not need to be secret from the eavesdropper, and can be recycled over several blocks so that the exchange of necessary randomness for universal hashing between the legitimate users does not affect the overall communication rate.

At the decoder, the legitimate receiver first estimates $(A_{B,l}^{1:K})_{l\in\mathcal{L}}$ from $(A_{B,l}^{1:K}[H_{U|Y}])_{l\in\mathcal{L}}$ and $(Y_{B,l}^{1:K})_{l\in\mathcal{L}}$, as described in Lines 2-4 of Algorithm 4 for Block B. Then, from this esti-mate of $(A_{B,l}^{1:K})_{l\in\mathcal{L}}$, the legitimate receiver forms an estimate of M_B and M_B^0, as described in Lines 6-7 of Algorithm 4 for Block B. Next, to estimate the message M_{B-1} and M_{B-1}^0, the legitimate receiver uses the estimate of M_B^0 along with $(A_{B-1,l}^{1:K}[H_{U|Y}\backslash\mathcal{V}_{U|Y}])_{l\in\mathcal{L}}$, and $(Y_{B-1,l}^{1:K})_{l\in\mathcal{L}}$, as described in Lines 2-7 of Algorithm 4 for Block $B-1$. Hence, the legitimate receiver can estimate all the messages $(M_b)_{b\in\mathcal{B}}$ starting from the last block and iterating through the previous blocks via the loop in Line 2 of Algorithm 4.

Note that in the analysis of the coding scheme secrecy rate, one needs to account for *(i)* the one-time pad in Line 15 of Algorithm 3, *(ii)* the transmission of the randomness $(R_b)_{1:B}$ that is used in Algorithms 3 and 4, and *(iii)* the initialization phase (Algorithms 1 and 2). We will show that *(i)*, *(ii)*, and *(iii)* are done with a negligible impact on the secrecy rate in Sections VI-C, VI-G, and VII-C, respectively.

VI. Proof of Theorem 7 With a Pre-Shared Key

In this section, we prove Theorem 7 when the legitimate users have access to a pre-shared secret key whose rate is negligible. Hence, we ignore in this section the initialization phase, i.e., Algorithms 1, 2. We also assume in this section that all the components of $\mathbf{s}_b$, $b \in \mathcal{B}$, are identical and equal to s. To simplify notation, we write s instead of $\mathbf{s}_b$, $b \in \mathcal{B}$.

A. Characterization of the Distribution Induced by the Encoder

Let $p_{U_b^{1:N}X_b^{1:N}Y_b^{1:N}Z_b^{1:N}(s)}$ denote the distribution induced by the encoding scheme described in Algorithm 3. Lemma 1 gives an approximation of $p_{U_b^{1:N}X_b^{1:N}Y_b^{1:N}Z_b^{1:N}(s)}$ in terms of the distribution $q_{UXYZ(s)}$ defined in Section V-A. This result will be useful in our subsequent analysis.

Lemma 1: For $b \in \mathcal{B}$, we have

$$\mathsf{D}(q_{U^{1:N}X^{1:N}Y^{1:N}Z^{1:N}(s)} \| p_{U_b^{1:N}X_b^{1:N}Y_b^{1:N}Z_b^{1:N}(s)}) \le 2LK\delta_K,$$

where $q_{U^{1:N}X^{1:N}Y^{1:N}Z^{1:N}(s)} \coloneqq \prod_{i=1}^{N} q_{UXYZ(s)}$.

Proof: See Appendix A. ∎

B. Reliability

We now show that the receiver is able to recover the original message with a vanishing error probability. Define $M_{1:B} \coloneqq (M_b)_{b\in\mathcal{B}}$. Define for $b \in \mathcal{B}$, $A_B^{1:N} \coloneqq \mathsf{k}_{l\in\mathcal{L}}(A_{B,l}^{1:K}, A_B^{1:N})$, $A_b^{1:N} \coloneqq \mathsf{k}_{l\in\mathcal{L}}(A_{b,l}^{1:K}, A_b^{1:N})$, $A_b^{1:LK} \coloneqq (A_{b,l}^{1:K})_{l\in\mathcal{L}}$, $E_{b-1} \coloneqq \{A_{b-1}^{1:N} = A_{b-1}^{1:N}\}$, and $E_{A_b} \coloneqq \{(Y_b^{1:N}, A_b^{1:N}) = (Y_b^{1:N}, A_b^{1:N})\}$. For $b \in \mathcal{B}$, consider a coupling [40, Lemma 3.6] between $p_{Y_b^{1:N}A_b^{1:N}}$ and

$q_{Y_b^{1:N}A_b^{1:N}}$ such that $\mathsf{P}[E_{A_b}] = \mathsf{V}(p_{Y_b^{1:N}A_b^{1:N}}, q_{Y_b^{1:N}A_b^{1:N}})$. For $b \in B$, consider $(A_b^{1:N}, Y_b^{1:N}, A_b^{1:N}, Y_b^{1:N})$ distributed according to this coupling, then

$$\begin{aligned}
&\mathsf{P}\left[M_{1:B} = M_{1:B}\right] \\
&\leq \mathsf{P}\left[M_b = M_b\ b \in B\right] \\
&\overset{(a)}{\leq} \mathsf{P}\left[T_b^{1:|V_U|L} = T_b^{1:|V_U|L}\ b \in B\right] \\
&\overset{(b)}{\leq} \mathsf{P}\left[A_b^{1:N} = A_b^{1:N}\ b \in B\right] \\
&\leq \sum_{b \in B} \mathsf{P}\left[A_b^{1:N} = A_b^{1:N} | E_{A_b}^c \cap E_b^c\right] + \mathsf{P}[E_{A_b} \cup E_b] \\
&\overset{(c)}{\leq} \mathsf{P}\left[A_{b,l}^{1:K} = A_{b,l}^{1:K} | E_b^c \cap E_A^c\right] \\
&\overset{(d)}{\leq} \sum_{b \in B} \sum_{l=1}^{L} \mathsf{P}[E_{A_b}] + \mathsf{P}[E_b] \leq \sum_{b=1}^{B} KL\delta_K + \sqrt{2\ln 2}\sqrt{2LK\delta_K} + \mathsf{P}\left[A_b^{1:N} = A_b^{1:N}\right] \\
&\overset{(e)}{\leq} \sum_{b \in B} (KL\delta_K + \sqrt{2\ln 2}\sqrt{2LK\delta_K})(B - b + 1) \\
&= (KL\delta_K + \sqrt{2\ln 2}\sqrt{2LK\delta_K})B(B+1)/2, \qquad (7)
\end{aligned}$$

where (a) holds by Line 7 in Algorithm 4, (b) holds by Line 6 in Algorithm 4, (c) holds by the union bound, (d) holds because $\mathsf{P}\left[A_{b,l}^{1:K} = A_{b,l}^{1:K} | E_{b,l}^c \cap E_A^c\right] \leq K\delta_K$ by (1) and because $\mathsf{P}[E_{A_b}] = \mathsf{V}(p_{Y_b^{1:N}A_b^{1:N}}, q_{Y_b^{1:N}A_b^{1:N}}) \leq \sqrt{2\ln 2}\sqrt{2LK\delta_K}$ by Lemma 1 and Pinsker's inequality, (e) holds by induction.

C. Pre-Shared Key Rate

The coding scheme described in Algorithms 3 and 4 involves a one-time pad to securely transmit $(f_2^{\mathbf{SC}}(A_{b,l}^{1:K}))_{l \in L, b \in B}$, and $(f_1^{\mathbf{SC}}(A_{B,l}^{1:K}))_{l \in L}$, which requires a pre-shared key with length $l_{\mathrm{OTP}} = LB|H_{U|Y}\backslash V_{U|Y}| + L|V_{U|Y}|$ and rate

$$\begin{aligned}
\frac{l_{\mathrm{OTP}}}{NB} &= \frac{|H_{U|Y}| - |V_{U|Y}|}{K} + \frac{|V_{U|Y}|}{KB} \\
&\leq \frac{|H_{U|Y}| - |V_{U|Y}|}{K} + \frac{1}{B} \\
&= \delta(K) + 1/B,
\end{aligned}$$

where $\delta(K)$ is such that $\lim_{K\to\infty}\delta(K) = 0$ since $\lim_{K\to\infty}|H_{U|Y}|/K = H(U|Y)$ [36], and $\lim_{K\to\infty}|V_{U|Y}|/K = H(U|Y)$ [30], [41].

D. Blockwise Security Analysis

We prove in this section that security holds in each block $b \in B$ individually. We use a series of lemmas to obtain this result and determine acceptable values for the parameter r defined in (6). For (X, Z) distributed according to p_{XZ}, defined over the finite alphabet $\mathcal{X} \times \mathcal{Z}$, recall that the ε-smooth min-entropy of X given Z is defined as [27]

$$H_\infty(p_{XZ}|p_Z) \coloneqq \max_{r_{XZ} \in \mathcal{B}(p_{XZ})} \min_{z \in \mathrm{Supp}(p_Z)} \min_{x \in \mathcal{X}} \log \frac{p_Z(z)}{r_{XZ}(x,z)},$$

where $\mathrm{Supp}(p_Z) \coloneqq \{z \in \mathcal{Z} : p_Z(z) > 0\}$ and $\mathcal{B}(p_{XZ}) \coloneqq \{(r_{XZ} : \mathcal{X} \times \mathcal{Z} \to [0,1]) : \mathsf{V}(p_{XZ}, r_{XZ}) \leq \varepsilon\}$. We will also need the following version of the leftover hash lemma.

Lemma 2 [27]: Let T and Z be distributed according to p_{TZ} over $\mathcal{T} \times \mathcal{Z}$. Consider $F : \mathcal{R} \times \{0,1\}^k \to \{0,1\}^r$, where the first input, denoted by R, is uniformly distributed over $\mathcal{R}$ to indicate that F is chosen uniformly at random in a family of two-universal hash functions. Then, for any $\varepsilon \in [0,1[$,

$$\mathsf{V}(p_{F(R,T),R,Z}, p_{U_K}p_{U_R}p_Z) \leq \varepsilon + \sqrt{2^{r - H_\infty(p_{TZ}|p_Z)}}, \qquad (8)$$

where p_{U_K} and p_{U_R} are the uniform distribution over $\{0,1\}^r$ and $\mathcal{R}$, respectively.

We now would like to use Lemma 2 to make $(M_b \| M_b^0)$ almost independent from the eavesdropper channel observations. However, in the encoding scheme described in Algorithm 3, $(M_b \| M_b^0)$ is not defined as the output of a two-universal hash function as required in Lemma 2. To overcome this challenge, we show in the following lemma that the distribution p induced by the encoder in Algorithm 3 also describes a process for which $(M_b \| M_b^0)$ is defined as $(M_b \| M_b^0) \coloneqq g_{R_b}^{\mathbf{UH}}(T_b^{1:|V_U|L}, r)$ where r is defined in (6). For convenience, we write in the following $F(R_b, T_b^{1:|V_U|L}) \coloneqq g_{R_b}^{\mathbf{UH}}(T_b^{1:|V_U|L}, r)$.

Lemma 3: Fix $b \in B$. To simplify notation, we write T_b instead of $T_b^{1:|V_U|L}$, $Z_b(s)$ instead of $Z_b^{1:N}(s)$, X_b instead of $X_b^{1:N}$, and $Z_b(s)$ instead of $Z_b^{1:N}(s)$. We also define $\tilde{M}_b \coloneqq (M_b \| M_b^0)$ such that $T_b \coloneqq R_b^{-1}(\tilde{M}_b \| R_b)$. Next, define

$$q_{\tilde{M}_b T_b X_b Z_b(s) R_b} \coloneqq p_{X_b Z_b(s)|T_b} q_{T_b} q_{R_b} q_{\tilde{M}_b|T_b R_b}, \qquad (9)$$

with q_{T_b} the uniform distribution over $\{0,1\}^{|V_U|L}$, q_{R_b} the uniform distribution over $\mathcal{R}$, and $\forall \tilde{m}_b, \forall t_b, \forall r_b$, $q_{\tilde{M}_b|T_b R_b}(\tilde{m}_b|t_b, r_b) \coloneqq \mathbb{1}\{\tilde{m}_b = F(r_b, t_b)\}$. Then, we have

$$p_{\tilde{M}_b T_b X_b Z_b(s) R_b} = q_{\tilde{M}_b T_b X_b Z_b(s) R_b}.$$

Proof: See Appendix B. ∎

Let $A_b \in \llbracket 1, N \rrbracket$ such that $|A_b| \geq \alpha N$ and consider $X_b^{1:N}[A_b]$, the αN symbols that the eavesdropper has chosen to have access to in Block $b \in B$. We study, by combining Lemmas 2, 3, the independence between $(R_b, Z_b^{1:N}(s), X_b^{1:N}[A_b])$, i.e., all the knowledge at the eavesdropper in Block $b \in B$, and $(M_b \| M_b^0)$ as follows.

Lemma 4: Fix $b \in B$. We adopt the same notation as in Lemma 3 and also write $X_b[A_b]$ instead of $X_b^{1:N}[A_b]$ for convenience. We have for any $\gamma \in]0,1[$

$$\begin{aligned}
&\mathsf{V}(p_{\tilde{M}_b R_b Z_b(s) X_b[A_b]}, p_{\tilde{M}_b} p_{R_b Z_b(s) X_b[A_b]}) \\
&\leq 2^{1-L^\gamma} + \sqrt{2^{r - H(T_b|Z_b(s)X_b[A_b]) + N\delta^{(1)}(K,L)}}, \qquad (10)
\end{aligned}$$

where $\delta^{(1)}(K,L) \coloneqq (K^{-1}+1)\sqrt{2L^{\gamma-1}}$.

Proof: See Appendix C. ∎

Next, using Lemma 1, we lower bound the conditional entropy in (10) in the following lemma.

Lemma 5: Fix $b \in B$. We adopt the same notation as in Lemmas 3, 4. We have

$$\begin{aligned}
&H(T_b|Z_b(s)X_b[A_b]) \\
&\geq N[(1-\alpha)H(U|Z(s)) + \alpha H(U|X) - \delta^{(2)}(K,L)],
\end{aligned}$$

with $\delta^{(2)}(K,L) \triangleq 2^{\sqrt{2\ln 2}\sqrt{2N\delta_K}(\log(|\mathcal{X}|^2 \max_{s\in\mathcal{S}}|\mathcal{Z}_s|) - N^{-1}\log(\sqrt{2\ln 2}\sqrt{2N\delta_K})) + N^{-1}H_b(N\delta_K) + N\delta_K + o(1)}$, *and* $H_b(\cdot)$ *the binary entropy.*

Proof: See Appendix D. ∎

By combing Lemma 4 and Lemma 5 we obtain the following result.

Lemma 6: Fix $b \in \mathcal{B}$. *We adopt the same notation as in Lemma 5. We have for any* $\gamma \in]0,1[$

$$\mathbb{V}(p_{\tilde{M}_b R_b Z_b(s) X_b[\mathcal{A}_b]}, p_{\tilde{M}_b} p_{R_b Z_b(s) X_b[\mathcal{A}_b]}) \le 2^{1-L^\gamma} + 2^{-N[(1-\alpha)H(U|Z(s))+\alpha H(U|X)-\delta^{(3)}(K,L)]},$$

where $\delta^{(3)}(K,L) \triangleq \delta^{(1)}(K,L) + \delta^{(2)}(K,L)$, *with* $\delta^{(1)}(K,L)$ *defined in Lemma 4 and* $\delta^{(2)}(L,K)$ *defined in Lemma 5.*

Finally, we obtain security in a given block as follows.

Lemma 7: Fix $b \in \mathcal{B}$ *and* $\xi > 0$. *We choose*

$$r \triangleq (1-\alpha)\min_{s\in\mathcal{S}} H(U|Z(s)) + \alpha H(U|X) - \delta^{(3)}(K,L) - \xi$$

with $\delta^{(3)}(K,L)$ *defined in Lemma 6. Then, for* L *large enough*

$$I(M_b M_b^0; Z_b(s) X_b[\mathcal{A}_b] R_b) \le \delta^{(4)}(K,L,\xi),$$

where $\delta^{(4)}(K,L,\xi) \triangleq (2^{1-L^\gamma} + \sqrt{2^{-N\xi}})\log\frac{2^N}{2^{1-L^\gamma}+\sqrt{2^{-N\xi}}}$.

Proof: We adopt the same notation as in the previous lemmas. By definition of r and by Lemma 6, we have

$$\mathbb{V}(p_{M_b R_b Z_b(s) X_b[\mathcal{A}_b]}, p_{M_b} p_{R_b Z_b(s) X_b[\mathcal{A}_b]}) \le 2^{1-L^\gamma} + \sqrt{2^{-N\xi}}. \tag{11}$$

We thus have

$$\begin{aligned} & I(M_b M_b^0; Z_b(s) X_b[\mathcal{A}_b] R_b) \\ &= I(\tilde{M}_b; Z_b(s) X_b[\mathcal{A}_b] R_b) \\ &\overset{(a)}{\le} f(\mathbb{V}(p_{\tilde{M}_b R_b Z_b(s) X_b[\mathcal{A}_b]}, p_{\tilde{M}_b} p_{R_b Z_b(s) X_b[\mathcal{A}_b]})) \\ &\overset{(b)}{\le} f(2^{1-L^\gamma} + \sqrt{2^{-N\xi}}), \end{aligned} \tag{12}$$

where *(a)* holds by [42, Lemma 2.7] with $f : x \to x\log(2^N/x)$, *(b)* holds for L large enough since f is increasing for small enough values. ∎

E. Analysis of Security Over All Blocks Jointly

We obtain security over all blocks jointly from Lemma 7 as follows.

Lemma 8: For convenience, we define for $i, j \in \mathcal{B}$, $Z_{1:i}(s) \triangleq (Z_b^{1:N}(s))_{b\in[\![1,i[\![}$, $X_{1:i}[\mathcal{A}] \triangleq (X_b^{1:N}[\mathcal{A}_b])_{b\in[\![1,i[\![}$, $R_{i:j} \triangleq (R_b)_{b\in[\![i,j[\![}$, *and* $M_{i:j} \triangleq (M_b)_{b\in[\![i,j[\![}$. *We have*

$$\max_{s\in\mathcal{S}}\max_{\mathcal{A}\in\mathbb{A}} I(M_{1:B}; Z_{1:B}(s) X_{1:B}[\mathcal{A}] R_{1:B}) \le 2B\delta^{(4)}(L,K,\xi),$$

where $\delta^{(4)}(L,K,\xi)$ *is defined in Lemma 7.*

Proof: For convenience, define for $i \in \mathcal{B}$, $L_i \triangleq (Z_i(s), X_i[\mathcal{A}_i], R_i)$ and $L_{1:i} \triangleq (Z_{1:i}(s), X_{1:i}[\mathcal{A}], R_{1:i})$. Then,

$$\begin{aligned} I(M_{1:B}; L_{1:B}) &= \sum_{i=0}^{B-1} I(M_{1:B}; L_{i+1}|L_{1:i}) \\ &\overset{(a)}{=} \sum_{i=0}^{B-1} I(M_{1:i+1}; L_{i+1}|L_{1:i}) \\ &\le \sum_{i=0}^{B-1} I(M_{1:i+1} L_{1:i}; L_{i+1}) \\ &= \sum_{i=0}^{B-1} I(M_{i+1}; L_{i+1}) \\ &\quad + I(M_{1:i} L_{1:i}; L_{i+1}|M_{i+1}) \\ &\overset{(b)}{\le} B\delta^{(4)}(K,L,\xi) \\ &\quad + \sum_{i=0}^{B-1} I(M_{1:i} M_{i+1} L_{1:i}^{\mathsf{c}}; L_{i+1} M_{i+1}) \\ &\overset{(c)}{=} B\delta^{(4)}(K,L,\xi) + \sum_{i=0}^{B-1} I(M_{i+1}^0; L_{i+1} M_{i+1}) \\ &\overset{(d)}{=} B\delta^{(4)}(K,L,\xi) + \sum_{i=0}^{B-1} I(M_{i+1}^0; L_{i+1}|M_{i+1}) \\ &\le B\delta^{(4)}(K,L,\xi) + \sum_{i=0}^{B-1} I(M_{i+1} M_{i+1}^0; L_{i+1}) \\ &\overset{(e)}{\le} 2B\delta^{(4)}(K,L,\xi), \end{aligned} \tag{13}$$

where *(a)* holds by the chain rule and since we have $I(M_{i+2:B}; L_{i+1}|L_{1:i} M_{1:i+1}) \le I(M_{i+2:B}; L_{1:i+1} M_{1:i+1}) = 0$, *(b)* holds by Lemma 7, *(c)* holds by the chain rule and because $(M_{1:i}, L_{1:i}) - M_{i+1}^0 - (L_{i+1}, M_{i+1})$ forms a Markov chain, *(d)* holds by independence between M_{i+1}^0 and M_{i+1}, *(e)* holds by Lemma 7. The lemma holds since (13) holds for any $s \in \mathcal{S}$ and any $\mathcal{A} \in \mathbb{A}$. ∎

F. Secrecy Rate

The rate of the transmitted messages is

$$\begin{aligned} \frac{\sum_{b\in\mathcal{B}}|M_b|}{BN} &\overset{(a)}{=} \frac{r + (B-1)(r - L + |\mathcal{V}_{U|Y}|)}{BN} \\ &\ge \frac{r}{N} - \frac{|\mathcal{V}_{U|Y}|}{K} \\ &\overset{(b)}{=} I(U;Y) - \alpha I(U;X) - (1-\alpha)\max_{s\in\mathcal{S}} I(U;Z(s)) \\ &\quad - \delta^{(3)}(K,L) - \xi + o(1), \end{aligned}$$

where *(a)* holds by (6), *(b)* holds by the choice of r in Lemma 7 and because $\lim_{K\to\infty} |\mathcal{V}_{U|Y}|/K = H(U|Y)$ by [41].

G. Randomness Amortization

The randomness $(R_b)_{1:B}$ in the coding scheme of Section V-C needs to be shared between the legitimate users. This can be done with negligible impact on the overall communication rate similar to [16] using an hybrid argument by repeating the coding scheme of Section V-C with the same randomness $(R_b)_{1:B}$.

VII. Proof of Theorem 7 Without Pre-Shared Key

The coding scheme of Section V-C requires a pre-shared secret key between the legitimate users. We now consider the initialization phase, described in Algorithms 1, 2, to generate such a key with negligible impact on the overall communication rate. We study the reliability and the secrecy of the

generated key in Sections VII-A and VII-B, respectively, the impact of the initialization phase on the overall communication rate in Section VII-C, and the joint secrecy of the initialization phase and the coding scheme of Section V-C in Section VII-D. We adopt the same notation as in Section VI.

A. Key Reliability

Similar to Lemma 1, we have the following result.

Lemma 9: For $b \in \mathcal{B}_0$, the distribution p induced by the encoder of Algorithm 1 is approximated as follows.

$$\mathsf{D}(q_{U^{1:N}X^{1:N}Y^{1:N}Z^{1:N}(s)} \| p_{U_b^{1:N}X_b^{1:N}Y_b^{1:N}Z_b^{1:N}(s)}) \leq 2LK\delta_K.$$

Then, we have

$$\begin{aligned}\mathsf{P}\left[(\mathrm{Key}_b)_{b\in\mathcal{B}_0} = (\mathrm{Key}_b)_{b\in\mathcal{B}_0}\right] &\leq \mathsf{P}[(U_b^{1:N})_{b\in\mathcal{B}_0} = (U_b^{1:N})_{b\in\mathcal{B}_0}] \\ &\leq B_0 L\left(\sqrt{2\ln 2}\sqrt{2K\delta_K} + 2K\delta_K\right),\end{aligned}$$

where the last inequality holds similar to (7).

B. Key Secrecy

We first show secrecy in a given Block $b \in \mathcal{B}_0$. Let $\mathcal{A}_b \subseteq [1, N]$ such that $|\mathcal{A}_b| = \alpha N$ and consider $X_b^{1:N}[\mathcal{A}_b]$, the αN symbols that the eavesdropper has chosen to have access to in Block $b \in \mathcal{B}_0$. Define $p_{U_{\mathrm{Key}}}$ the uniform distribution over $\{0,1\}^{l_{\mathrm{key}}^{\mathsf{c}}}$. We have

$$\begin{aligned}&\mathbb{V}\left(p_{\mathrm{Key}\,b R_b^{\mathrm{init}} Z_b(s) X_b[\mathcal{A}_b] D_b R_b^{\mathrm{init}0}},\ p_{U_{\mathrm{Key}}} p_{R_b^{\mathrm{init}} Z_b(s) X_b[\mathcal{A}_b] D_b R_b^{\mathrm{init}0}}\right) \\ &\overset{(a)}{\leq} 2 + 2^{l_{\mathrm{key}}^{\mathsf{c}} - H_\infty\left(p_{U_b Z_b(s) X_b[\mathcal{A}_b] D_b R_b^{\mathrm{init}0}} \,|\, p_{Z_b(s) X_b[\mathcal{A}_b] D_b R_b^{\mathrm{init}0}}\right)} \\ &\overset{(b)}{\leq} 2\cdot 2^{-L\gamma} + 2^{l_{\mathrm{key}}^{0} - H(U_b|Z_b(s) X_b[\mathcal{A}_b] D_b R_b^{\mathrm{init}0}) + N\delta^{(1)}(K,L)},\end{aligned} \tag{14}$$

where *(a)* holds by Lemma 2, *(b)* holds by Lemma 14 with $\gamma \in]0,1[$ as in the proof of Lemma 4 with $\delta^{(1)}(K,L)$ defined in Lemma 4.

Lemma 10: For $b \in \mathcal{B}_0$, we have

$$\begin{aligned}&H(U_b|Z_b(s) X_b[\mathcal{A}_b] D_b R_b^{\mathrm{init}\mathsf{c}}) \\ &\geq N[I(U;Y) - \alpha I(U;X) - (1-\alpha) I(U;Z(s)) - \delta^{(5)}(K,L)],\end{aligned}$$

where $\delta^{(5)}(K,L) \triangleq 2^{\sqrt{2\ln 2}\sqrt{2N\delta_K}}(\log(|\mathcal{X}|^2 \max_{s\in\mathcal{S}} |\mathcal{Z}_s|) - N^{-1}\log(\sqrt{2\ln 2}\sqrt{2N\delta_K})) + o(1)$.

Proof: We have

$$\begin{aligned}&H(U_b|Z_b(s) X_b[\mathcal{A}_b] D_b R_b^{\mathrm{init}\mathsf{c}}) \\ &= H(U_b|Z_b(s) X_b[\mathcal{A}_b]) - I(D_b R_b^{\mathrm{init}\mathsf{c}}; U_b|Z_b(s) X_b[\mathcal{A}_b]) \\ &\geq H(U_b|Z_b(s) X_b[\mathcal{A}_b]) - L(|\mathcal{H}_{U|Y}| + |\mathcal{H}_{U|Y} \setminus \mathcal{V}_{U|Y}|) \\ &\overset{(a)}{\geq} H(U_b|Z_b(s) X_b[\mathcal{A}_b]) - NH(U|Y) - o(KL) \\ &\overset{(b)}{\geq} N(1-\alpha)H(U|Z(s)) + N\alpha H(U|X) - NH(U|Y) \\ &\quad - N\delta^{(5)}(K,L),\end{aligned}$$

where *(a)* holds because $\lim_{K\to\infty} |\mathcal{H}_{U|Y}|/K = H(U|Y)$ [36], and $\lim_{K\to\infty} |\mathcal{V}_{U|Y}|/K = H(U|Y)$ [30], [41], *(b)* holds similar to the proof of Lemma 5. ∎

Next, we choose

$$l_{\mathrm{key}}^{\mathsf{c}} \triangleq N[I(U;Y) - \alpha I(U;X) - (1-\alpha)\max_{s\in\mathcal{S}} I(U;Z(s)) - \delta^{(1)}(K,L) - \delta^{(5)}(K,L) - \xi],$$

with $\xi > 0$. By (14) and Lemma 10, we obtain for $b \in \mathcal{B}_0$,

$$\begin{aligned}&\mathbb{V}\left(p_{\mathrm{Key}_b R_b^{\mathrm{init}} Z_b(s) X_b[\mathcal{A}_b] D_b R^{\mathrm{init}0}},\ p_{U_{\mathrm{Key}}} p_{R_b^{\mathrm{init}} Z_b(s) X_b[\mathcal{A}_b]}^{\otimes D} p_{R_b^{\mathrm{init}0}}\right) \\ &\leq 2\cdot 2^{-L\gamma} + \sqrt{2^{-N\xi}}.\end{aligned} \tag{15}$$

Lemma 11: We have for L large enough

$$\begin{aligned}&I(\mathrm{Key}_b^b; Z_b(s) X_b[\mathcal{A}_b] D_b R_b^{\mathrm{init}} R_b^{\mathrm{init}0}) \leq \delta^{(4)}(K,L,\xi), \\ &\log|\mathcal{K}_b| - H(\mathrm{Key}_b) \leq \delta^{(4)}(K,L,\xi),\end{aligned}$$

with $\mathcal{K}_b \triangleq \{0,1\}^{l_{\mathrm{key}}^{\mathsf{c}}}$ and $\delta^{(4)}(K,L,\xi)$ defined in Lemma 7.

Proof: The first inequality holds as the proof of Lemma 7 by using (15) in place of (11). The second inequality holds by [42, Lemma 2.7] and (15). ∎

By mutual independence of all the B_0 blocks of the initialization phase, we obtain from Lemma 11 the following result.

Lemma 12: Define $\mathrm{Key} \triangleq (\mathrm{Key}_b)_{b\in\mathcal{B}_0}$ and $\mathcal{K} \triangleq \mathcal{K}_b^{B_0}$. Let $Z^{\mathrm{init}}(s)$ denote all the knowledge of the eavesdropper related to the initialization phase, i.e., $Z^{\mathrm{init}}(s) \triangleq (Z_b(s), X_b[\mathcal{A}_b], D_b, R_b^{\mathrm{init}}, R_b^{\mathrm{init}})_{b\in\mathcal{B}_0}$. Then, for K large enough

$$\begin{aligned}&\max_{s\in\mathcal{S}} \max_{\mathcal{A}\in\mathcal{A}} I(\mathrm{Key}; Z^{\mathrm{init}}(s)) \leq B_0\delta^{(4)}(K,L,\xi), \\ &\log|\mathcal{K}| - H(\mathrm{Key}) \leq B_0\delta^{(4)}(K,L,\xi).\end{aligned}$$

C. Impact of the Initialization Phase on the Overall Communication Rate

The initialization phase requires $\rho N B_0$ channel uses, for some fixed $\rho \in \mathbb{N}$, to generate the secret key and transmit $(D_b, R_b^{\mathrm{init}}, R_b^{\mathrm{init}\mathsf{c}})_{b\in\mathcal{B}_0}$. We choose B_0 such that

$$B_0 = \frac{l_{\mathrm{OTP}}}{l_{key}^0},$$

where $l_{\mathrm{OTP}} = o(NB)$ represents the key length necessary to perform the one-time pad that appears in Algorithms 1, 2. Hence, the impact of the initialization phase on the overall communication is

$$\rho N B_0 < \rho N\left(1 + \frac{l_{\mathrm{OTP}}}{l_{key}^0}\right) = \rho\frac{o(NB)}{l_{key}^{\mathsf{c}}/N} = o(NB). \tag{16}$$

We deduce from (16) that the communication rate of the coding scheme of Section V-C and the initialization phase (considered jointly) is the same as the communication rate of the coding scheme of Section V-C alone.

D. Security of Algorithms 3, 4 and the Initialization Phase When Considered Jointly

Let M_{OTP} be the sequence that needs to be secretly transmitted with a one-time pad in Algorithm 3. Let $C \triangleq M_{\mathrm{OTP}} \oplus \mathrm{Key}$ be the encrypted version of M_{OTP} using Key, obtained in the initialization phase. Let $Z_B(s) \triangleq (Z_{1:B}(s), X_{1:B}[\mathcal{A}], R_{1:B})$ denote all the observations of the eavesdropper related to the coding scheme of Section V-C, excluding C. Let $Z^{\mathrm{init}}(s)$,

defined as in Lemma 12, denote all the observations of the eavesdropper related to the initialization phase. The following lemma shows that strong secrecy holds for the coding scheme of Section V-C and the initialization phase considered jointly.

Lemma 13: We have

$$\max_{s \in \mathcal{S}, A \in \mathcal{A}} I(M_{1:B}; C Z_B(s) Z^{\text{init}}(s)) \leq 2(B+B_0)\delta^{(4)}(K, L, \xi),$$

where $\delta^{(4)}(K, L, \xi)$ is defined in Lemma 7.

Proof: We have

$$\begin{aligned} & I(M_{1:B}; C Z_B(s) Z^{\text{init}}(s)) \\ & \overset{(a)}{=} I(M_{1:B}; Z_B(s)) + I(M_{1:B}; C \mid Z_B(s) Z^{\text{init}}(s)) \\ & \leq I(M_{1:B}; Z_B(s)) + I(M_{1:B} Z_B(s) Z^{\text{init}}(s); C) \\ & = I(M_{1:B}; Z_B(s)) + I(C; M_{1:B} Z_B(s)) \\ & \quad + I(C; Z^{\text{init}}(s) \mid M_{1:B} Z_B(s)), \end{aligned} \tag{17}$$

where *(a)* holds by the chain rule and because $I(M_{1:B}; Z^{\text{init}}(s) \mid Z_B(s)) \leq I(M_{1:B} Z_B(s); Z^{\text{init}}(s)) = 0$. Next, we have

$$\begin{aligned} & I(C; M_{1:B} Z_B(s)) \\ & \leq \log |K| - H(C \mid M_{1:B} Z_B(s)) \\ & \leq \log |K| - H(\text{Key} \oplus M_{\text{OTP}} \mid M_{\text{OTP}} M_{1:B} Z_B(s)) \\ & = \log |K| - H(\text{Key} \mid M_{\text{OTP}} M_{1:B} Z_B(s)) \\ & = \log |K| - H(\text{Key}). \end{aligned} \tag{18}$$

We also have

$$\begin{aligned} & I(C; Z^{\text{init}}(s) \mid M_{1:B} Z_B(s)) \\ & \leq I(C M_{\text{OTP}}; Z^{\text{init}}(s) \mid M_{1:B} Z_B(s)) = \\ & I(\text{Key} M_{\text{OTP}}; Z^{\text{init}}(s) \mid M_{1:B} Z_B(s)) \\ & \overset{(b)}{=} I(\text{Key}; Z^{\text{init}}(s) \mid M_{\text{OTP}} M_{1:B} Z_B(s)) \\ & \leq I(\text{Key} M_{\text{OTP}} M_{1:B} Z_B(s); Z^{\text{init}}(s)) \\ & \overset{(c)}{=} I(\text{Key}; Z^{\text{init}}(s)), \end{aligned} \tag{19}$$

where *(b)* holds by the chain rule and because $I(M_{\text{OTP}}; Z^{\text{init}}(s) \mid M_{1:B} Z_B(s)) \leq I(M_{\text{OTP}} M_{1:B} Z_B(s); Z^{\text{init}}(s)) = 0$, *(c)* holds by the chain rule and because $I(M_{\text{OTP}} M_{1:B} Z_B(s); Z^{\text{init}}(s) \mid \text{Key}) \leq I(M_{\text{OTP}} M_{1:B} Z_B(s); Z^{\text{init}}(s) \text{Key}) = 0$. By combining (17), (18), and (19), we obtain $I(M_{1:B}; C Z_B(s) Z^{\text{init}}(s)) \leq I(M_{1:B}; Z_B(s)) + I(\text{Key}; Z^{\text{init}}(s)) + \log |K| - H(\text{Key})$. Finally, we obtain the lemma with Lemmas 8 and 12. ∎

VIII. Proof of Theorem 8

We assume in the following that there exists a best channel for the eavesdropper [23], i.e., $\exists s^* \in \overline{\mathcal{S}}$, $\forall s \in \mathcal{S}$, $X - Z(s^*) - Z(s)$. Similar to the proof of Theorem 7, we proceed in two steps. We first ignore the initialization phase and assume that the legitimate users have access to a secret key to perform the one-time pad in Algorithms 3, 4. We only show blockwise security as the remainder of the proof is similar to the proof in Section VI. We also omit the second step that consists in analyzing the initialization phase jointly with Algorithms 3, 4, as it is similar to the analysis in Section VII.

A. Blockwise Security Analysis

We adopt the same notation as in Section VI. We have the following inequality, whose proof is identical to the proof of Lemma 1. For $b \in \mathcal{B}$, we have

$$\mathsf{D}(q_{U^{1:N} X^{1:N} Y^{1:N} Z^{1:N}(\mathbf{s}_b)} \| p_{U_b^{1:N} X_b^{1:N} Y_b^{1:N} Z_b^{1:N}(\mathbf{s}_b)}) \leq 2N\delta_K, \tag{20}$$

where we have defined $q_{U^{1:N} X^{1:N} Y^{1:N} Z^{1:N}(\mathbf{s}_b)} \triangleq \prod_{i=1}^{N} q_{UXYZ(s_{b,i})}$. Next, similar to Lemma 4 using (20) in place of Lemma 1, we have for any $\gamma \in]0, 1[$

$$\begin{aligned} & \mathsf{V}(p_{\tilde{M}_b R_b Z_b(\mathbf{s}_b) X_b[A_b]}, p_{\tilde{M}_b} p_{R_b Z_b(\mathbf{s}_b) X_b[A_b]}) \\ & \leq 2^{1-L^\gamma} + 2\sqrt{2^{r - H_T^b(Z_b(\mathbf{s}_b) X_b[A_b]) + N\delta^{(1)}(K,L)}}, \end{aligned} \tag{21}$$

where $\delta^{(1)}(K, L)$ is defined in Lemma 4. We then have

$$\begin{aligned} & H(T_b \mid Z_b(\mathbf{s}_b) X_b[A_b]) \\ & \overset{(a)}{\geq} H(U_b \mid Z_b(\mathbf{s}_b) X_b[A_b]) - N\delta^{(2)}(K, L) \\ & \geq H(U_b \mid Z_b(s^*) Z_b(\mathbf{s}_b) X_b[A_b]) - N\delta^{(2)}(K, L) \\ & \overset{(b)}{=} H(U_b \mid Z_b(s^*) X_b[A_b]) - N\delta^{(2)}(K, L) \\ & \overset{(c)}{=} N(1-\alpha) H(U \mid Z(s^*)) + N\alpha H(U \mid X) - N\delta^{(2)}(K, L), \end{aligned} \tag{22}$$

where *(a)* holds as in the proof of Lemma 5 with $\delta^{(2)}(K, L)$ defined in Lemma 5, *(b)* holds because $(U_b, X_b) - Z_b(s^*) - Z_b(\mathbf{s}_b)$ forms a Markov chain, *(c)* holds as in the proof of Lemma 5. Finally, from (21) and (22), we can conclude as in Section VI-D that blockwise security holds.

IX. Extension to Uncertainty on the Main Channel

Assume now that uncertainty on the main channel also holds according to a compound model, i.e., the channel of Section III is now defined by the conditional probabilities $(p_{Y(t)Z(s)|X})_{s \in \mathcal{S}, t \in \mathsf{T}}$, where T is a finite set. Assume also that for all channel uses $s \in \mathcal{S}$ and $t \in \mathsf{T}$ are fixed. We extend Theorem 7 to this setting in Section IX-C using new polar coding schemes for source coding with compound side information and for compound channel coding described in Sections IX-A and IX-B, respectively.

A. Source Coding With Compound Side Information

[43] provides a polar coding scheme with optimal rate for lossless source coding with compound side information. However, for our purposes, we modify the coding scheme in [43] to ensure near uniformity of the encoder output.

Consider a compound source $(\mathcal{U} \times \mathcal{Y}_j)_{j \in \mathcal{J}}, (p_{UY_j})_{j \in \mathcal{J}}$, where $\mathcal{U} \subset \{0, 1\}$ and $\mathcal{J} \subset [\![1, J]\!]$. Let $(t_j)_{j \in \mathcal{J}} \in \mathbb{N}^J$ with $t_1 \geq 1$ and define for $j \in \mathcal{J}$, $T_j \triangleq \sum_{i=1}^{j} t_i$ and $N_j \triangleq NT_j$, where N is a power of two. Consider for $j \in \mathcal{J}$, $(U^{1:N_j}, Y_j^{1:N_j}) = (U_t^{1:N}, (Y_j)_t^{1:N})_{t \in [\![1, T_J]\!]}$ distributed according to the product distribution $p_{U^{1:N_J} Y_j^{1:N_J}}$. For $j_0 \in \mathcal{J}$, we also use the notation $Y_{j_0}^{1:N_j} = (Y_{j_0, t}^{1:N_{j-1}})_{t \in [\![1, t_j]\!]}$, $j \in [\![2, J]\!]$, to indicate that $Y_{j_0}^{1:N_j}$ is made of t_j blocks of length N_{j-1}.

Define for $t \in [1, T_J]$, $A_t^{1:N}$, $U_t^{1:N} G_N$ and for δ_N, 2^{-N^β}, $\beta \in]0, 1/2[$, and $j \in [J]$ define the sets

$$\mathcal{V}_U, \quad i \in [1, N] : H(A^i | A^{1:i-1}) > 1 - \delta_N,$$
$$\mathcal{H}_{U|Y_j}, \quad i \in [1, N] : H(A^i | A^{1:i-1}(Y_j)^{1:N}) > \delta_N, \mathcal{V}_{U|Y_j}$$
$$, \quad i \in [1, N] : H(A^i | A^{1:i-1}(Y_j)^{1:N}) > 1 - \delta_N.$$

We also use the notation $U_t^{1:N_j} = (U_t^{1:N_{j-1}})_{t \in [1, t_j]}$, $j \in [2, J]$, to indicate that $U^{1:N_j}$ is made of t_j blocks of length N_{j-1}. The encoding is described in Algorithm 5. By the successive cancellation decoder for polar source coding with side information [36], Decoder 1 with $[e^{(1)}(U^{1:N_1}), E^0] = A_1^{1:N}[\mathcal{H}_{U|Y_1}]$ and $Y_1^{1:N}$ can compute a good estimate $\hat{U}_1^{1:N_1}$ of $U^{1:N_1}$. Now, assume that when $L \in [1, J-1]$, for any Decoder $l \in [1, L]$, there is a function $g_l^{(L)}$ such that $\hat{U}_l^{1:N_L} = g_l^{(L)}(e^{(L)}(U^{1:N_L}), E_l^0, Y_l^{1:N_L})$ is a good estimate of $U^{1:N_L}$. Then, Algorithms 6 and 7 show that any decoder $l \in [1, L+1]$ can form a good estimate $\hat{U}_l^{1:N_{L+1}}$ of $U^{1:N_{L+1}}$ from $[e^{(L+1)}(U^{1:N_{L+1}}), E_l^0, Y_l^{1:N_{L+1}}]$.

The encoding and decoding algorithms for source cod-ing with compound side information are described in Algorithms 5, 6, 7, and yield the following result.

Theorem 9: The algorithms 5, 6, 7 perform source cod-ing with compound side information on sequences with length $T_J N$ with optimal rate $\max_{j \in [J]} H(U|Y_j)$ and encod-ing/decoding complexity $T_J N O(\log N)$.

Note that the encoding is different than in [43] as the encoder output is split into E and E^0, however, the decoder is equivalent to the one in [43]. Consequently, the probability of error in the reconstruction of the source asymptotically vanishes by [43]. Additionally, remark that the rate of E^0 is negligible compared to N_J because for any $j \in [J]$, $|\mathcal{H}_{U|Y_j} \backslash \mathcal{V}_{U|Y_j}| = |\mathcal{H}_{U|Y_j}| - |\mathcal{V}_{U|Y_j}| = o(N)$ by [36] and [29, Lemma 7]. Hence, the coding scheme rate is the same as in [43] but now can also be used to ensure a near uniform encoder output by one-time padding E^0 with a sequence of $|E^0|$ uniformly distributed bits shared by the encoder and decoder. Note that it generalizes the polar coding schemes for source coding with nearly uniform output [44] in [37] and [45].

B. Compound Channel Coding From Source Coding

We now propose a capacity-achieving compound channel coding scheme from source coding with compound side information via a technique similar to the one in [26] used to obtain channel coding from source coding with side information.

Consider a compound channel $\mathcal{X}, (p_{Y_j|X})_{j \in [J]}, (\mathcal{Y}_j)_{j \in [J]}$, where $\mathcal{X}$, $\{0, 1\}$ and J, $[1, J]$. Consider an arbitrary distribution p_X on $\mathcal{X}$ and define for $j \in [J]$, p_{XY_j}, $p_X p_{Y_j|X}$. Consider for $j \in [J]$, $(X^{1:N}, Y_j^{1:N})$ distributed according to the product distribution $p_{X^{1:N} Y_j^{1:N}}$. Define $V^{1:N}$, $X^{1:N} G_N$ and for δ_N, 2^{-N^β}, $\beta \in]0, 1/2[$, and $j \in [J]$, define the sets

$$\mathcal{V}_X, \quad i \in [1, N] : H(V^i | V^{1:i-1}) > 1 - \delta_N,$$
$$\mathcal{H}_{X|Y_j}, \quad i \in [1, N] : H(V^i | V^{1:i-1} Y_j^{1:N}) > \delta_N,$$

Algorithm 5 Encoding

Require: Assume that the sequence to compress is $U^{1:N_J}$

1: Define the function $e^{(1)} : U^{1:N_1} \to A_1^{1:N}[\mathcal{V}_{U|Y_1}]$
2: **for** $j = 1$ to $J - 1$ **do**
3: Define $f^{(j)} : U^{1:N_j} \to (A_t^{1:N}[\mathcal{V}_{U|Y_{j+1}}])_{t \in [1, T_j]}$
4: Define the function $e^{(j+1)}$ which maps $U^{1:N_{j+1}}$ to

$$[e^{(j)}(U_1^{1:N_j}), (e^{(j)}(U_{t+1}^{1:N_j}) \| f^{(j)}(U_t^{1:N_j}))_{t \in [1, t_{j+1}-1]}, f^{(j)}(U_{t_{j+1}}^{1:N_j})],$$

(if the two sequences have different lengths, then the shorter sequence is padded with zeros)
5: **end for**
6: Define E, $e^{(J)}(U^{1:N_J})$
7: For $j \in [J]$, define E_j^0, $(A_t^{1:N}[\mathcal{H}_{U|Y_j} \backslash \mathcal{V}_{U|Y_j}])_{t \in [1, T_j]}$, and E^0, $(E_j^0)_{j \in [J]}$.
8: **return** (E, E^0)

Algorithm 6 Decoder $j_0 \in [1, L]$

Require: (E, E^0) and $Y_{j_0}^{1:N_{L+1}}$

1: Form $U_{j_0,1}^{1:N_L}$, $g_{j_0}^{(L)}(e^{(L)}(U_1^{1:N_L}), E_{j_0}^0, Y_{j_0,1}^{1:N_L})$, where $e^{(L)}(U_1^{1:N_L})$ is obtained from $e^{(L+1)}(U^{1:N_{L+1}})$
2: **for** Block $t = 2$ to Block $t = t_{L+1}$ **do**
3: Form $U_{j_0,t}^{1:N_L}$, $g_{j_0}^{(L)}(e^{(L)}(U_t^{1:N_L}) \| f^{(L)}(U_{t-1}^{1:N_L}) \| f^{(L)}(U_{j_0,t-1}^{1:N}), E_{j_0}^0, Y_{j_0,t}^{1:N})$
4: **end for**
5: **return** $U_{j_0}^{L+1}$, $(U_{j_0,t}^L)_{t \in [1, t_{L+1}]}$, an estimate of $U^{1:N_{L+1}}$

Algorithm 7 Decoder $L + 1$

Require: (E, E^0) and $Y_{L+1}^{1:N_{L+1}}$

1: With the successive cancellation decoder for source coding with side information [36], form $U_{L+1,t_{L+1}}^{1:N_L}$ from

$$f^{(L)}(U_{t_{L+1}}^{1:N_L}), E_{L+1}^0, Y_{L+1,t_{L+1}}^{1:N_L}$$

2: **for** Block $t = t_{L+1} - 1$ to Block $t = 1$ **do**
3: Form an estimate $f^{(L)}(U_t^{1:N_L})$ of $f^{(L)}(U_t^{1:N_L})$ with

$$f^{(L)}(U_t^{1:N_L}), \quad f^{(L)}(U_t^{1:N_L}) \| e^{(L)}(U_{t+1}^{1:N_L}) \| e^{(L)}(U_{L+1,t+1}^{1:N})$$

4: With the successive cancellation decoder for source coding with side information [36], form $U_{L+1,t}^{1:N_L}$ from $f^{(L)}(U_t^{1:N_L}), E_{L+1}^0, Y_{L+1,t}^{1:N_L}$
5: **end for**
6: **return** $U_{L+1}^{1:N_{L+1}}$, $(U_{L+1,t}^{1:N_L})_{t \in [1, t_{L+1}]}$

$$\mathcal{V}_{X|Y_j}, \quad i \in [1, N] : H(V^i | V^{1:i-1} Y_j^{1:N}) > 1 - \delta_N.$$

Let $(t_j)_{j \in [J]} \in \mathbb{N}^J$ with $t_1 = 1$ and define for $j \in [J]$, $T_j = \sum_{i=1}^{j} t_i$ and $N_j = NT_j$. We use the same notation as in Section IX-A. Let $|E|$ be the length of the output E in

the encoder of source coding with compound side information described in Algorithm 5. By Euclidean division, there exist $q \in \mathbb{N}$ and $r \in [1, T_J]$ such that $|E| = T_J q + r$. For $t \in [1, r]$, consider an arbitrary set $A_t \subseteq V_X$ such that $|A_t| = q + 1$, and, for $t \in [r+1, T_J]$, consider an arbitrary set $A_t \subseteq V_X$ such that $|A_t| = q$. Hence, $\sum_{t=1}^{T_J} |A_t| = |E|$.

The encoding and decoding algorithms for compound chan-nel coding are described in Algorithms 8 and 9, and yield the following result, whose proof is similar to [46]. Note that other capacity-achieving polar coding schemes had also been proposed for compound symmetric channels in [25] and [31].

Theorem 10: Algorithms 8 and 9 perform compound chan-nel coding over B blocks of length $T_J N$ with optimal rate $\max_{p_X} \min_{j \in J} I(X; Y_j)$ and encoding/decoding complexity $O(B T_J N \log N)$.

Algorithm 8 Encoder

Require: E_0, $(E_{0,t})_{t \in [1, T_J]}$, where $E_{0,t}$, $t \in [1, T_J]$, is a sequence of $|A_t|$ uniformly distributed bits (local randomness). Messages $(M_{b,t})_{b \in [1,B], t \in [1,T_J]}$, where $M_{b,t}$, $b \in [1, B]$, $t \in [1, T_J]$, is a sequence of $|V_X \backslash A_t|$ uniformly distributed bits

1: **for** Block $b = 1$ to Block $b = B$ **do**
2: **for** Sub-block $t = 1$ to Sub-block $t = T_J$ **do**
3: Define $V_{b,t}^{1:N}$ according to $\prod_{j=1}^{N} p_{V_{b,t}^j | V_{b,t}^{1:j-1}}$ with

$$p_{V_{b,t}^j | V_{b,t}^{1:j-1}}(v_{b,t}^j | v_{b,t}^{1:j-1}) \in \begin{cases} \{1\{v_{b,t}^j = M_{b,t}^j\}\} & \text{if } j \in V_X \backslash A_t \\ \{1\{v_{b,t}^j = E_{b-1,t}^j\}\} & \text{if } j \in A_t \\ \{p_{V^j | V^{1:j-1}}(v_{b,t}^j | v_{b,t}^{1:j-1})\} & \text{if } j \in V_X^c \end{cases},$$

4: Send $X_{b,t}^{1:N}$, $V_{b,t}^{1:N} G_N$ over the channel.
5: **end for**
6: Define (E_b, E_b^0) as the output of the encoder described in Algorithm 5 (for the compound source $(p_{XY_j})_{j \in J}$) applied to $X_b^{1:N_J}$, $(X_{b,t}^{1:N})_{t \in [1, T_J]}$
7: Break down E_b into T_J sequences $(E_{b,t})_{t \in [1, T_J]}$, such that $|E_{b,t}| = |A_t|$, $t \in [1, T_J]$.
8: **end for**
9: Do a one-time pad with $(E_b^0)_{b \in [1, B]}$ and E_B to ensure uniformity (similar to Algorithm 3) and send it to the receiver via channel codes [39] for each $p_{Y_j|X}$, $j \in J$

Remark 5: We do not write the dependence of the estimates with respect to $j \in J$ in Algorithm 9 to simplify notation.

C. Extension to Compound Uncertainty on the Main Channel

Using the preliminary results of Section IX-A and IX-B, an immediate extension of Theorem 7 is as follows.

Theorem 11: Assume that in the coding scheme of Section V the primitive source coding with side information is replaced by source coding with compound side information from Section IX-A. Assume also that instead of channel coding in Lines 10 and 15 of Algorithm 1 and 3, respectively, we use compound channel coding from Section IX-B. Then,

Algorithm 9 Decoder $j \in J$

Require: Channel output $Y_j^{1:BN_J}$, estimate E_B of E_B, and estimate $(E_b^0)_{b \in [1, B]}$ of $(E_b^0)_{b \in [1, B]}$

1: **for** Block $b = B$ to Block $b = 1$ **do**
2: Use (E_b, E_b^0) with Decoder j in Algorithms 6, 7 to create an estimate $X_b^{1:N_J}$, $(X_{b,t}^{1:N})_{t \in [1, T_J]}$ of $X_b^{1:N_J}$, $(X_{b,t}^{1:N})_{t \in [1, T_J]}$.
3: **for** Sub-block $t = 1$ to Sub-block $t = T_J$ **do**
4: Form an estimate $V_{b,t}^{1:N}$, $X_{b,t}^{1:N} G_N$ of $V_{b,t}^{1:N}$
5: Form an estimate $M_{b,t}$, $V_{b,t}^{1:N}[V_X \backslash A_t]$ of $M_{b,t}$
6: Form an estimate $E_{b-1,t}$, $V_{b,t}^{1:N}[A_t]$ of $E_{b-1,t}$
7: **end for**
8: Form E_{b-1}, $(E_{b-1,t})_{t \in [1, T_J]}$ an estimate of E_{b-1}
9: **end for**
10: **return** $(M_{b,t})_{b \in [B], t \in [1, T_J]}$

the following secrecy rate is achieved

$$\max \min_{t \in T} I(U; Y(t)) - \alpha I(U; X) - (1 - \alpha) \max_{s \in S} I(U; Z(s))^+$$

where the maximum is over random variables U such that $\forall t \in T, \forall s \in S$, $U - X - (Y(t), Z(s))$, and $|U| \leq |X|$.

X. Concluding Remarks

We constructed explicit wiretap codes that achieve the best known single-letter achievable rates, previously obtained non-constructively, when uncertainty holds on the eavesdropper channel under a (i) noisy blockwise type II, (ii) compound, or (iii) arbitrarily varying model. Our construction solely relies on three primitives: source coding with side information, universal hashing, and distribution approximation. We also extended our result to the case where uncertainty holds on the legitimate user channel under a compound model. This extension can thus be applied to the problem of secret sharing from correlated randomness. Specifically, it can directly be applied to the case of a discrete channel model as in [47, Section II], and adapted to the case of a discrete source model with a single dealer, as in [48] and [49], for arbitrary access structures. The case of Gaussian channels or sources, e.g., [47] and [50], is, however, more challenging as quantization may be needed. The case of rate-limited communication for source models is also more challenging as vector quantization is needed and requires other proof techniques [51].

We anticipate that our code construction can be generalized to the broadcast channel with confidential messages and the multiple access wiretap channel when uncertainty holds on the eavesdropper's channel according to a compound model, using a distributed version of the leftover hash lemma akin to [52]. Such results would generalize known constructions based on polar codes, e.g., [11], [29], and [53], that require a seed for strong secrecy and assume perfect knowledge of the eavesdropper's channel statistics. An open problem is to provide explicit coding schemes to handle an arbitrarily varying main channel as, for instance, in the models in [23], [24], [54], and [55].

APPENDIX A PROOF OF LEMMA 1

Let $b \in \mathcal{B}$ and $l \in \mathcal{L}$. By (3), we have

$$\mathsf{D}(q_{A^{1:K}} \| p_{A^{1:K}_{b,l}}) \leq K\delta_K, \tag{23}$$

we can indeed apply (3) because the bits $A^{1:K}_{b,l}[\mathcal{V}_U]$ are uniformly distributed, which is a consequence of the definition of $A^{1:K}_{b,l}[\mathcal{V}_U]$ in Line 7 of Algorithm 3 using the fact that the bits $T_b^{1:|\mathcal{V}_U|L} = R_b^{-1}(M_b \| M_b^0 \| R_b^0)$ are uniformly distributed since the bits $(M_b \| M_b^0 \| R_b^0)$ are uniformly distributed. Next, we have

$$\begin{aligned}
&\mathsf{D}(q_{U^{1:K}V^{1:K}} \| p_{U^{1:K}_{b,l}V^{1:K}_{b,l}}) \\
&\overset{(a)}{=} \mathsf{E}_{q_{U^{1:K}}} \mathsf{D}(q_{V^{1:K}|U^{1:K}} \| p_{V^{1:K}_{b,l}|U^{1:K}_{b,l}}) + \mathsf{D}(q_{U^{1:K}} \| p_{U^{1:K}_{b,l}}) \\
&\overset{(b)}{\leq} \mathsf{E}_{q_{U^{1:K}}} \mathsf{D}(q_{V^{1:K}|U^{1:K}} \| p_{V^{1:K}_{b,l}|U^{1:K}_{b,l}}) + K\delta_K \\
&\overset{(c)}{\leq} 2K\delta_K,
\end{aligned} \tag{24}$$

where *(a)* holds by the chain rule for relative entropy [56], *(b)* holds by (23) because $\mathsf{D}(q_{U^{1:K}} \| p_{U^{1:K}_{b,l}}) = \mathsf{D}(q_{A^{1:K}} \| p_{A^{1:K}_{b,l}})$ by invertibility of G_K, *(c)* holds by (5). Then,

$$\begin{aligned}
&\mathsf{D}(q_{U^{1:N}X^{1:N}Y^{1:N}Z^{1:N}(s)} \| p_{U_b^{1:N}X_b^{1:N}Y_b^{1:N}Z_b^{1:N}(s)}) \\
&\overset{(a)}{=} \sum_{l\in\mathcal{L}} \mathsf{D}(q_{U^{1:K}X^{1:K}Y^{1:K}Z^{1:K}(s)} \| p_{U^{1:K}_{b,l}X^{1:K}_{b,l}Y^{1:K}_{b,l}Z^{1:K}_{b,l}(s)}) \\
&\overset{(b)}{=} \sum_{l\in\mathcal{L}} [\mathsf{D}(q_{U^{1:K}X^{1:K}} \| p_{U^{1:K}_{b,l}X^{1:K}_{b,l}}) \\
&\quad + \mathsf{E}[\mathsf{D}(q_{Y^{1:K}Z^{1:K}(s)|U^{1:K}X^{1:K}} \| p_{Y^{1:K}_{b,l}Z^{1:K}_{b,l}(s)|U^{1:K}_{b,l}X^{1:K}_{b,l}})]] \\
&\overset{(c)}{=} \sum_{l\in\mathcal{L}} \mathsf{D}(q_{U^{1:K}X^{1:K}} \| p_{U^{1:K}_{b,l}X^{1:K}_{b,l}}) \\
&\overset{(d)}{\leq} \sum_{l\in\mathcal{L}} 2K\delta_K = 2LK\delta_K,
\end{aligned}$$

where *(a)* holds because the random variables $(U^{1:K}_{b,l}, X^{1:K}_{b,l}, Y^{1:K}_{b,l}, Z^{1:K}_{b,l}(s))$ across the different sub-blocks $l \in \mathcal{L}$ are independent by construction (see Algorithm 3 and Remark 3), *(b)* holds by the chain rule for relative entropy [56] and the expectation is over $q_{U^{1:K}X^{1:K}}$, *(c)* holds because $p_{Y^{1:K}_{b,l}Z^{1:K}_{b,l}(s)|U^{1:K}_{b,l}X^{1:K}_{b,l}} = p_{Y^{1:K}_{b,l}Z^{1:K}_{b,l}(s)|X^{1:K}_{b,l}} = q_{Y^{1:K}Z^{1:K}(s)|X^{1:K}} = q_{Y^{1:K}Z^{1:K}(s)|U^{1:K}X^{1:K}}$, *(d)* holds by (24) because $\mathsf{D}(q_{U^{1:K}X^{1:K}} \| p_{U^{1:K}_{b,l}X^{1:K}_{b,l}}) = \mathsf{D}(q_{U^{1:K}V^{1:K}} \| p_{U^{1:K}_{b,l}V^{1:K}_{b,l}})$ by invertibility of G_K.

APPENDIX B PROOF OF LEMMA 3

For any $(\bar{m}_b, t_b, x_b, z_b(s), r_b)$, we have

$$\begin{aligned}
&p_{\bar{M}_b T_b X_b Z_b(s) R_b}(\bar{m}_b, t_b, x_b, z_b(s), r_b) \\
&\overset{(a)}{=} p_{X_b Z_b(s)|T_b}(x_b, z_b(s)|t_b) p_{\bar{M}_b}(\bar{m}_b) p_{R_B}(r_b) \\
&\quad \times \sum_{r_b^0} p^{R_b^0}(r_b^0) p_{T_b|R_b^0 \bar{M}_b R_b}(t_b|r_b^0, \bar{m}_b, r_b) \\
&\overset{(b)}{=} p_{X_b Z_b(s)|T_b}(x_b, z_b(s)|t_b) 2^{-r} |R|^{-1} \\
&\quad \times \sum_{r_b^0} \frac{\mathbb{1}\{t_b = r_b^{-1}(\bar{m}_b \| r_b^0)\}}{2^{-r+|\mathcal{V}_U|L}} \\
&= p_{X_b Z_b(s)|T_b}(x_b, z_b(s)|t_b) 2^{-|\mathcal{V}_U|L} |R|^{-1} \\
&\quad \times \sum_{r_b^0} \mathbb{1}\{r_b = t_b = (\bar{m}_b \| r_b)\} \\
&\overset{(c)}{=} p_{X_b Z_b(s)|T_b}(x_b, z_b(s)|t_b) 2^{-|\mathcal{V}_U|L} |R|^{-1} \mathbb{1}\{F(r_b, t_b) = \bar{m}_b\} \\
&\overset{(d)}{=} q_{\bar{M}_b T_b X_b Z_b(s) R_b}(\bar{m}_b, t_b, x_b, z_b(s), r_b),
\end{aligned}$$

where *(a)* holds because $p_{\bar{M}_b T_b X_b Z_b R_b} = p_{X_b Z_b|T^b} p_{\bar{M}_b} p_{R_b} p_{|M_b R^- b b}$ and R_b^c is independent of (M_b, R_b), *(b)* holds by uniformity of $\bar{M}_b$, R_b, R_b^c, and by definition of T_b, *(c)* holds because $(F(r_b, t_b) = \bar{m}_b) \Longleftrightarrow (\sum_{r_b^0} \mathbb{1}\{r_b\ t_b = (\bar{m}_b, r_b)\} = 1)$ (because $\exists! r^0 \in \{0,1\}^{|\mathcal{V}_U|L-r}$ such that $r_b\ t_b = (\bar{m}_b \| r^0)$) and $(F(r_b, t_b) = \bar{m}_b) \Longleftrightarrow (\sum_{r_b^0} \mathbb{1}\{r_b\ t_b = (\bar{m}_b, r_b^c)\} = 0)$, *(d)* holds by definition of q.

APPENDIX C PROOF OF LEMMA 4

We have

$$\begin{aligned}
&\mathsf{V}(p_{M_b R_b Z_b(s) X_b[A_b]}, p_{M_b} p_{R_b Z_b(s) X_b[A_b]}) \\
&\overset{(a)}{=} \mathsf{V}(q_{F(R_b,T_b) R_b Z_b(s) X_b[A_b]}, q_{\bar{M}_b} q_{R_b Z_b(s) X_b[A_b]}) \\
&\overset{(b)}{\leq} 2 + 2^{r - H_\infty(p_{T_b Z_b(s) X_b[A_b]} | p_{Z_b(s) X_b[A_b]})} \\
&\overset{(c)}{\leq} 2 \cdot 2^{-L\nu} + 2^{r - H(T_b|Z_b(s) X_b[A_b]) + L\delta^{(0)}(K,L)} \\
&\overset{(d)}{\leq} 2^{-L\nu} + 2^{r - H(T^b|Z_s^b X_b[A_b]) + N\delta^{(1)}(K,L)},
\end{aligned}$$

where *(a)* holds by Lemma 3 and the definition of q, *(b)* holds by Lemmas 2 and 3, *(c)* holds by Lemma 14 below, which can indeed be applied by Remark 4, with $\varepsilon = 2^{-L\nu}$, $\delta^{(0)}(K,L) = \sqrt{2L\nu^{-1}} \log(2^{|\mathcal{V}_U|} + 3)$, *(d)* holds by choosing $\delta^{(1)}(K,L) = (K^{-1} + 1)\sqrt{2L\nu^{-1}} \geq \delta^{(0)}(K,L)/K$.

Lemma 14 [57] : Let $p_{X^L Z^L} = \prod_{i=1}^L p_{X_i Z_i}$ be a probability distribution over $\mathcal{X}^L \times \mathcal{Z}^L$. For any $\delta > 0$, $H_\infty(p_{X^L Z^L} | p_{Z^L}) \geq H(X^L|Z^L) - L\delta$, where $\varepsilon = 2^{-\frac{L\delta^2}{2\log^2(|\mathcal{X}|+3)}}$.

Remark 6: An argument similar to [58, Lemma 10] to lower bound the min-entropy would require adding an extra round of reconciliation to the coding scheme as in [59]. Lemma 14 appears to be a simpler alternative here.

APPENDIX D PROOF OF LEMMA 5

We first introduce some notation for convenience. Define for any $I \subseteq \{1, K\}$, $A_b[I]$, $(A^{1:K}_{b,l}[I])_{l\in\mathcal{L}}$ and A_b, $(A^{1:K}_{b,l})_{l\in\mathcal{L}}$. For $b \in \mathcal{B}$, consider $(U^{1:K}_{b,l}, X^{1:K}_{b,l}, Z^{1:K}_{b,l}(s))_{l\in\mathcal{L}}$ distributed according to $q_{U^{1:N}X^{1:N}Z^{1:N}(s)}$, $\prod_{i=1}^N q_{UXZ(s)}$ and define for $l \in \mathcal{L}$, $A^{1:K}_{b,l}$, $U^{1:K}_{b,l} G_K$. Next, define for any $I \subseteq \{1, K\}$, $A_b[I]$, $(A^{1:K}_{b,l}[I])_{l\in\mathcal{L}}$ and A_b, $(A^{1:K}_{b,l})_{l\in\mathcal{L}}$. Define $U_b[A_b]$, $(U^{1:K}_{b,l}[A_{b,l}])_{l\in\mathcal{L}}$, $U_b[A_b^c]$, $(U^{1:K}_{b,l}[A^c_{b,l}])_{l\in\mathcal{L}}$, U_b, $(U^{1:K}_{b,l})_{l\in\mathcal{L}}$, $X_b[A_b]$, $(X^{1:K}_{b,l}[A_{b,l}])_{l\in\mathcal{L}}$, $X_b[A_b^c]$, $(X^{1:K}_{b,l}[A^c_{b,l}])_{l\in\mathcal{L}}$, X_b, $(X^{1:K}_{b,l})_{l\in\mathcal{L}}$, $Z_b(s)[A_b]$, $(Z^{1:K}_{b,l}(s)[A_{b,l}])_{l\in\mathcal{L}}$, $Z_b(s)[A_b^c]$, $(Z^{1:K}_{b,l}(s)[A^c_{b,l}])_{l\in\mathcal{L}}$, $Z_b(s)$, $(Z^{1:K}_{b,l}(s))_{l\in\mathcal{L}}$. Then, we have

$$\begin{aligned}
&H(A_b[\mathcal{V}_U]|Z_b(s)X_b[A_b]) - H(A_b[\mathcal{V}_U]|Z_b(s)X_b[A_b]) \\
&= H(A_b[\mathcal{V}_U]Z_b(s)X_b[A_b]) - H(A_b[\mathcal{V}_U]Z_b(s)X_b[A_b]) \\
&\quad + H(Z_b(s)X_b[A_b]) - H(Z_b(s)X_b[A_b]) \\
&\geq -2\sqrt{2\ln 2 \, LK\delta_K} \log \sqrt{\frac{(|\mathcal{X}|^2 |\mathcal{Z}_s|)^{LK}}{2\ln 2 \cdot 2LK\delta_K}}
\end{aligned}$$

$$\geq -2^{\sqrt{2\ln 2}\sqrt{LK\delta_K}\,(\frac{1}{K}K\log(|\mathcal{X}|^2\max_{s\in\mathcal{S}}|\mathcal{Z}|_s} - \log(\sqrt{2\ln 2}\ \sqrt{2LK\delta_K})), \quad -\delta', \tag{25}$$

where the first inequality holds by [42, Lemma 2.7] applied twice because for N large enough, $\sqrt{(q_{A_b[V_U]Z_b(s)X_b[A_b]}, p_{A_b[V_U]Z_b(s)X_b[A_b]})} \leq \sqrt{2\ln 2}\sqrt{D(q_{A_b[V_U]Z_b(s)X_b[A_b]}\|p_{A_b[V_U]Z_b(s)X_b[A_b]})} \leq \sqrt{2\ln 2}\sqrt{D(q_{U_b^{1:N}X_b^{1:N}Y_b^{1:N}Z_b^{1:N}(s)}\|p_{U_b^{1:N}X_b^{1:N}Y_b^{1:N}Z_b^{1:N}(s)})} \leq \sqrt{2\ln 2}\sqrt{2LK\delta_K}$ where we have used Pinsker's inequality, the chain rule for divergence, positivity of the divergence, and Lemma 1. Then, we have

$$\begin{aligned}
&H(T_b|Z_b(s)X_b[A_b]) \\
&\overset{(a)}{=} H(A_b[V_U]|Z_b(s)X_b[A_b]) \\
&\overset{(b)}{\geq} H(A_b[V_U]|Z_b(s)X_b[A_b]) - \delta' \\
&= H(A_b[H_U]|Z_b(s)X_b[A_b]) \\
&\quad - H(A_b[H_U\backslash V_U]|A_b[V_U]Z_b(s)X_b[A_b]) - \delta' \\
&\geq H(A_b[H_U]|Z_b(s)X_b[A_b]) - L|H_U\backslash V_U| - \delta' \\
&\overset{(c)}{=} H(A_b[H_U]|Z_b(s)X_b[A_b]) - o(LK) - \delta' \\
&= H(A_b[H_U]U_b|Z_b(s)X_b[A_b]) \\
&\quad - H(U_b|A_b[H_U]Z_b(s)X_b[A_b]) - o(LK) - \delta' \\
&\overset{(d)}{\geq} H(A_b[H_U]U_b|Z_b(s)X_b[A_b]) \\
&\quad - H_b(LK\delta_K) - (LK)^2\delta_K - o(LK) - \delta' \\
&\geq H(U_b|Z_b(s)X_b[A_b]) \\
&\quad - H_b(LK\delta_K) - (LK)^2\delta_K - o(LK) - \delta' \\
&\overset{(e)}{=} H(U_b|Z_b(s)[A_b^c]X_b[A_b]) \\
&\quad - H_b(LK\delta_K) - (LK)^2\delta_K - o(LK) - \delta',
\end{aligned} \tag{26}$$

where (a) holds by definition of $A_b[V_U]$, (b) holds by (25), (c) holds because $\lim_{K\to\infty}|H_U|/K = H(U)$ by [36], and $\lim_{K\to\infty}|V_U|/K = H(U)$ by [30], [41], (d) holds by Fano's inequality since the error probability in the reconstruction of U_b from $A_b[H_U]$ is upper-bounded by $LK\delta_K$ by the result for source coding with side information from [36], reviewed in (1), and the union bound, (e) holds because $U_b - (Z_b(s)[A_b^c], X_b[A_b]) - Z_b(s)[A_b]$ forms a Markov chain. Next, we have

$$\begin{aligned}
&H(U_b|Z_b(s)[A_b^c]X_b[A_b]) \\
&= H(U_b[A_b^c]|Z_b(s)[A_b^c]X_b[A_b]) \\
&\quad + H(U_b[A_b]|U_b[A_b^c]Z_b(s)[A_b^c]X_b[A_b]) \\
&\overset{(a)}{=} H(U_b[A_b^c]|Z_b(s)[A_b^c]) \\
&\quad + H(U_b[A_b]X_b[A_b]|U_b[A_b^c]Z_b(s)[A_b^c]) \\
&\quad - H(X_b[A_b]|U_b[A_b^c]Z_b(s)[A_b^c]) \\
&\overset{(b)}{=} H(U_b[A_b^c]|Z_b(s)[A_b^c]) + H(U_b[A_b]X_b[A_b]) \\
&\quad - H(X_b[A_b]) \\
&= H(U_b[A_b^c]|Z_b(s)[A_b^c]) + H(U_b[A_b]|X_b[A_b]) \\
&\overset{(c)}{=} N(1-\alpha)H(U|Z(s)) + N\alpha H(U|X),
\end{aligned} \tag{27}$$

where (a) holds because $X_b[A_b]$ is independent of $(U_b[A_b^c], Z_b(s)[A_b^c])$, (b) holds because $(U_b[A_b], X_b[A_b])$ is independent of $(U_b[A_b^c], Z_b(s)[A_b^c])$, (c) holds because $q_{U^{1:N}X^{1:N}Z^{1:N}(s)} = \prod_{i=1}^{N} q_{UXZ(s)}$. We obtain the lemma from (26) and (27).

REFERENCES

[1] R. A. Chou, "Explicit codes for the wiretap channel with uncertainty on the Eavesdropper's channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 476–480.

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

[3] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, 2009.

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

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

[6] A. Subramanian, A. Thangaraj, M. Bloch, and S. W. McLaughlin, "Strong secrecy on the binary erasure wiretap channel using large-girth LDPC codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 585–594, Sep. 2011.

[7] V. Rathi, R. Urbanke, M. Andersson, and M. Skoglund, "Rate-equivocation optimal spatially coupled LDPC codes for the BEC wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2011, pp. 2393–2397.

[8] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011.

[9] E. Sasoglu and A. Vardy, "A new polar coding scheme for strong security on wiretap channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 1117–1121.

[10] M. Andersson, R. F. Schaefer, T. J. Oechtering, and M. Skoglund, "Polar coding for bidirectional broadcast channels with common and confidential messages," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1901–1908, Sep. 2013.

[11] Y.-P. Wei and S. Ulukus, "Polar coding for the general wiretap channel with extensions to multiuser scenarios," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, pp. 278–291, Feb. 2016.

[12] J. M. Renes, R. Renner, and D. Sutter, "Efficient one-way secret-key agreement and private channel coding via polarization," in *Advances in Cryptology–(ASIACRYPT)*. Berlin, Germany: Springer, 2013, pp. 194–213.

[13] T. C. Gulcu and A. Barg, "Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1311–1324, Feb. 2017.

[14] R. A. Chou and M. R. Bloch, "Polar coding for the broadcast channel with confidential messages," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Apr. 2015, pp. 1–5.

[15] M. Hayashi and R. Matsumoto, "Construction of wiretap codes from ordinary channel codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2538–2542.

[16] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," in *Advances in Cryptology*. Berlin, Germany: Springer, 2012, pp. 294–311.

[17] H. Tyagi and A. Vardy, "Universal hashing for information-theoretic security," *Proc. IEEE*, vol. 103, no. 10, pp. 1781–1795, Oct. 2015.

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

[19] 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.

[20] M. Nafea and A. Yener, "A new wiretap channel model and its strong secrecy capacity," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 2804–2808.

[21] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wiretap channels," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, pp. 1–12, Oct. 2009.

[22] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy results for compound wiretap channels," *Problems Inf. Transmiss.*, vol. 49, no. 1, pp. 73–98, Mar. 2013.

[23] E. MolavianJazi, M. Bloch, and J. N. Laneman, "Arbitrary jamming can preclude secure communication," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2009, pp. 1069–1075.

[24] I. Bjelaković, H. Boche, and J. Sommerfeld, "Capacity results for arbitrarily varying wiretap channels," in *Proc. Inf. Theory, Combinatorics, Search Theory*. Berlin, Germany: Springer, 2013, pp. 123–144.
[25] S. H. Hassani and R. Urbanke, "Universal polar codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 1451–1455.
[26] M. Mondelli, R. Urbanke, and S. H. Hassani, "How to achieve the capacity of asymmetric channels," in *Proc. 52nd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2014, pp. 789–796.
[27] R. Renner, "Security of quantum key distribution," *Int. J. Quantum Inf.* vol. 6, no. 1, pp. 1–127, Feb. 2008.
[28] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.
[29] R. A. Chou and M. R. Bloch, "Polar coding for the broadcast channel with confidential messages: A random binning analogy," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2410–2429, May 2016.
[30] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric models," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, Dec. 2013.
[31] E. Şaşoğlu and L. Wang, "Universal polarization," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 2937–2946, Jun. 2016.
[32] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3989–4001, Jun. 2011.
[33] C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, Nov. 1995.
[34] M. Hayashi and R. Matsumoto, "Secure multiplex coding with dependent and non-uniform multiple messages," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2355–2409, May 2016.
[35] J. M. Renes and R. Renner, "Noisy channel coding via privacy amplification and information reconciliation," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7377–7385, Nov. 2011.
[36] E. Arikan, "Source polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 899–903.
[37] R. A. Chou, B. N. Vellambi, M. R. Bloch, and J. Kliewer, "Coding schemes for achieving strong secrecy at negligible cost," *IEEE Trans. Inf. Theory*, vol. 63, no. 3, pp. 1858–1873, Mar. 2017.
[38] J. L. Carter and M. N. Wegman, "Universal classes of hash functions," *J. Comput. Syst. Sci.*, vol. 18, no. 2, pp. 143–154, Apr. 1979.
[39] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jan. 2009.
[40] D. Aldous, "Random walks on finite groups and rapidly mixing Markov chains," in *Séminaire de Probabilités*. Berlin, Germany: Springer, 1983, pp. 243–297.
[41] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6213–6237, Nov. 2015.
[42] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge, U.K.: Cambridge Univ. Press, 1981.
[43] M. Ye and A. Barg, "Universal source polarization and an application to a multi-user problem," in *Proc. 52nd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2014, pp. 805–812.
[44] R. A. Chou and M. R. Bloch, "Data compression with nearly uniform output," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 1979–1983.
[45] R. A. Chou, M. R. Bloch, and A. Yener, "Universal covertness for discrete memoryless sources," *IEEE Trans. Inf. Theory*, vol. 67, no. 8, pp. 5432–5442, Aug. 2021.
[46] R. A. Chou and M. R. Bloch, "Using deterministic decisions for low-entropy bits in the encoding and decoding of polar codes," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2015, pp. 1380–1385.
[47] S. Zou, Y. Liang, L. Lai, and S. Shamai (Shitz) (Shitz), "An information theoretic approach to secret sharing," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3121–3136, Apr. 2015.
[48] R. A. Chou, "Secret sharing over a public channel from correlated random variables," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 991–995.
[49] R. A. Chou, "Distributed secret sharing over a public channel from correlated random variables," 2021, *arXiv:2110.10307*.
[50] V. Rana, R. A. Chou, and H. M. Kwon, "Information-theoretic secret sharing from correlated Gaussian random variables and public communication," *IEEE Trans. Inf. Theory*, vol. 68, no. 1, pp. 549–559, Jan. 2022.
[51] R. Sultana and R. A. Chou, "Low-complexity secret sharing schemes using correlated random variables and rate-limited public communication," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2021, pp. 970–975.
[52] R. A. Chou, "Private classical communication over quantum multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 68, no. 3, pp. 1782–1794, Mar. 2022.
[53] R. A. Chou and A. Yener, "Polar coding for the multiple access wiretap channel via rate-splitting and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 64, no. 12, pp. 7903–7921, Dec. 2018.
[54] J. Nötzel, M. Wiese, and H. Boche, "The arbitrarily varying wire-tap channel—Secret randomness, stability, and super-activation," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3504–3531, Jun. 2016.
[55] Y. Chen, D. He, C. Ying, and Y. Luo, "Strong secrecy of arbitrarily varying wiretap channel with constraints," *IEEE Trans. Inf. Theory*, vol. 68, no. 7, pp. 4700–4722, Jul. 2022.
[56] T. Cover and J. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1991.
[57] T. Holenstein and R. Renner, "On the randomness of independent experiments," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1865–1871, Apr. 2011.
[58] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Proc. Int. Conf. Theory Appl. Cryptograph. Techn.* (Lecture Notes in Computer Science). Berlin, Germany: Springer-Verlag, 2000, pp. 351–368.
[59] R. A. Chou and M. R. Bloch, "Separation of reliability and secrecy in rate-limited secret-key generation," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4941–4957, Aug. 2014.

Rémi A. Chou received the Engineering degree from Supélec, Gif-sur-Yvette, France, in 2011, and the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2015. From 2015 to 2017, he was a Post-Doctoral Scholar at The Pennsylvania State University, University Park, PA, USA. He is currently an Assistant Professor with the Department of Electrical Engineering and Computer Science, Wichita State University, Wichita, KS, USA.