Wiretap Channels with Random States Non-Causally Available at the Encoder

Ziv Goldfeld
Ben Gurion University
gziv@post.bgu.ac.il

Paul Cuff
Princeton University
cuff@princeton.edu

Haim H. Permuter
Ben Gurion University
haimp@bgu.ac.il

***Abstract*—We study the state-dependent (SD) wiretap channel (WTC) with non-causal channel state information (CSI) at the encoder. This model subsumes all other instances of CSI availability as special cases, and calls for an efficient utilization of the state sequence both for reliability and security purposes. A lower bound on the secrecy-capacity, that improves upon the previously best known result by Chen and Han Vinck, is derived based on a novel superposition coding scheme. The improvement over the Chen and Han Vinck result is strict for some SD-WTCs. Specializing the lower bound to the case where CSI is also available to the decoder reveals that it is at least as good as the achievable formula by Chia and El-Gamal, which is already known to outperform the adaptation of the Chen and Han Vinck code to the encoder and decoder CSI scenario. The results are derived under the strict semantic-security metric that requires negligible information leakage for all message distributions.**

I. Introduction

Reliably transmitting a message over a noisy state-dependent (SD) channel with non-causal encoder channel state information (CSI) is one of the most fundamental settings in information theory. The formulation of the problem and the derivation of its capacity dates back to Gelfand and Pinsker's (GP's) celebrated paper [1]. While the original motivation for the problem, as presented in [1], stems from the memory with stuck-at faults example [2], the implications of the result were much broader. One such prominent implication is that viewing the state sequence (known to the encoder) as a codeword of some other message naturally relates the GP scenario to the problem of broadcasting. It is therefore of no surprise that GP coding achieves the corner points of the best known inner bound on the capacity region of the broadcast channel [3]. Another virtue of the GP model is its generality. Namely, it is the most general instance of a SD point-to-point channel in which any or all of the terminals have non-causal access to sequence of states. Motivated by the above as well as the indisputable importance of security in modern communication systems, we study the SD wiretap channel (WTC) with non-causal encoder CSI, which incorporates the notion of security in the presence of a wiretapper into the GP channel coding problem.

First to consider a discrete and memoryless (DM) WTC with random states were Chen and Han Vinck [4], that studied encoder CSI scenario. They established a lower bound on the secrecy-capacity based on a combination of wiretap coding with GP coding. This work was later generalized in [5] to a WTC that is driven by a pair of states, one available to the encoder and the other one to the decoder. However, as previously mentioned, since CSI at the encoder is the most general setup, the result of [5] is a special of [4]. A more sophisticated coding scheme was constructed by Chia and El-Gamal for the SD-WTC with causal encoder CSI and full decoder CSI [6]. Their idea was to explicitly extract a cryptographic key from the random state, and protect a part of the confidential message via a one-time-pad with that key. The remaining portion of the confidential message is protected using a wiretap code (whenever wiretap coding is possible). Although their code is restricted to utilize the state in a causal manner, the authors of [6] proved that it can strictly outperform the adaptations of the non-causal schemes from [4], [5] to the encoder and decoder CSI setup.

In this paper we study the SD-WTC with non-causal encoder CSI, for which we propose a novel superposition-based coding scheme. The scheme results in a new lower bound on the secrecy-capacity, which recovers the previously best known achievability formulas from [4] and [5] as special cases. The relation to the previous schemes can be strict, i.e., there are examples where our scheme achieves strictly higher secrecy rates than [4], [5].

When specializing to the case where the decoder also knows the state sequence, our achievability is shown to be at least as good as the scheme from [6]. In fact, [6] provided two separate coding schemes and stated their achievability result as the maximum between the two. Recovering [6] from our lower bound results in a compact and simplified (yet equivalent) characterization of their achievable formula. Thus, our superposition-based coding scheme encompasses a unification of the two schemes from [6]. Interestingly, while both schemes from [6] rely on generating the aforementioned cryptographic key, our code construction does not involve any explicit key generation/agreement phase. Instead, we use an over-populated superposition codebook and encode the entire confidential message at the outer layer. The transmission is correlated with the state sequence by means of the likelihood encoder [7], while security is ensured by making the eavesdropper decode the inner layer codeword that contains no confidential information. Having done so, the eavesdropper is lacking the resources to extract any information about the secret message.

Our results are derived under the strict metric of semantic-security (SS). The SS criterion is a cryptographic benchmark that was adapted to the information-theoretic framework (of computationally unbounded adversaries) in [8]. In that work,

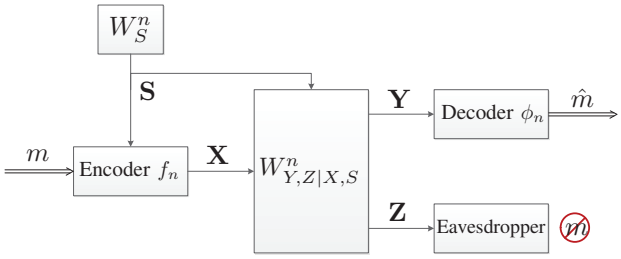

Fig. 1: The state-dependent wiretap channel with non-casual encoder channel state information.

SS was shown to be equivalent to a negligible mutual information between the message and the eavesdropper's observations for all message distributions. In contrast to this stringent security requirement, all the aforementioned secrecy results were derived under the weak-secrecy metric, i.e., a vanishing *normalized* mutual information with respect to a *uniformly distributed* message. Nowadays, however, weak-secrecy is widely regarded as being too loose, giving rise to the recent effort of upgrading information-theoretic secrecy results to strong-secrecy (namely, by removing the normalization factor but keeping the uniformity assumption on the message). SS is clearly a further strengthening of them both. Consequently, our achievability result outperforms the schemes from [4], [5] for the SD-WTC with non-causal encoder CSI not only in terms of the achievable secrecy rate, but also in the upgraded sense of security is provides. When CSI is also available at the decoder, our result implies that an upgrade to SS is possible, without inflicting any loss of rate compared to [6].

II. Wiretap Channels with Random States Non-Causally Available at the Encoder

We use notation from [9, Section II]. A less common notation this paper uses is of $\mathcal{P}(\mathcal{X})$, which denotes the set of all probability mass functions (PMFs) over a finite set $\mathcal{X}$.

We study the SD-WTC with non-causal encoder CSI, for which we establish a new and improved achievability formula that (in some cases) strictly outperforms the previously best known coding schemes for this scenario. The secrecy-capacity of a WTC with random states observed non-causally by some or all of the terminals is a highly challenging open problem in information-theoretic security that have received noticeable attention throughout the years (see, e.g., [4]–[6]. This interest in such secure communication scenarios stems from trying to understand how to optimally correlated the transmission with the state observation while exploiting the additional randomness offered by the knowledge of the state sequence to further enhance the secrecy rate. The optimal integration of these two ingredient is yet to be fully understood.

A. Problem Setup

Let $\mathcal{S}$, $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ be finite sets. The $\big(\mathcal{S},\mathcal{X},\mathcal{Y},\mathcal{Z},W_S,W_{Y,Z|X,S}\big)$ DMSD-WTC with non-causal encoder CSI is illustrated in Fig. 1. A state sequence $\mathbf{s}\in\mathcal{S}^n$ is generated in and i.i.d. manner according to W_S and is revealed in a non-causal fashion to the sender, who chooses a message m from the set $\big[1:2^{nR}\big]$. The sender then maps the observed state sequence $\mathbf{s}$ and the chosen message m into a sequence $\mathbf{x}\in\mathcal{X}^n$ (the mapping may be random). The sequence $\mathbf{x}$ is transmitted over the DMSD-WTC with transition probability $W_{Y,Z|X,W}$. The output sequences $\mathbf{y}\in\mathcal{Y}^n$ and $\mathbf{z}\in\mathcal{Z}^n$ are observed by the receiver and the eavesdropper, respectively. Based on $\mathbf{y}$, the receiver produces an estimate $\hat{m}$ of m. The eavesdropper tries to glean whatever it can about the message from $\mathbf{z}$.

Remark 1 (Most General Model) *Before rigorously defining the setup and stating the result, we note that the considered model is the most general instance of a SD-WTC with non-causal CSI known at some or all of the terminals. The broadest model one may consider is when the SD-WTC $W_{\tilde{Y},\tilde{Z}|X,S_1,S_2,S_3}$ is driven by a triple of correlated state random variables $(S_1,S_2,S_3)\sim W_{S_1,S_2,S_3}$, where S_1 is known to the transmitter, S_2 is known to the receiver and S_3 is available at the eavesdropper's site. However, setting $S=S_1$, $Y=(\tilde{Y},S_2)$, $Z=(\tilde{Z},S_3)$ in SD-WTC with non-causal encoder CSI and defining the channel's transition probability as*

$$W_{Y,Z|X,S}=W_{(\tilde{Y},S_2),(\tilde{Z},S_3)|X,S_1}=W_{S_2,S_3|S_1}W_{\tilde{Y},\tilde{Z}|X,S_1,S_2,S_3},$$

one clearly recovers this (prima facie) general SD-WTC from the model with non-causal encoder CSI only.

Definition 1 (Code) *An (n,R)-code c_n for the SD-WTC with non-causal encoder CSI has a message set $\mathcal{M}_n\triangleq\big[1:2^{nR}\big]$, a stochastic encoder $f_n:\mathcal{M}_n\times\mathcal{S}^n\to\mathcal{P}(\mathcal{X}^n)$ and a decoder $\phi_n:\mathcal{Y}^n\to\hat{\mathcal{M}}_n$, where $\hat{\mathcal{M}}_n=\mathcal{M}_n\cup\{e\}$ and $e\notin\mathcal{M}_n$.*

For any message distribution $P_M\in\mathcal{P}(\mathcal{M}_n)$ and any (n,R)-code c_n, the induced joint PMF on $\mathcal{S}^n\times\mathcal{M}_n\times\mathcal{X}^n\times\mathcal{Y}^n\times\mathcal{Z}^n\times\hat{\mathcal{M}}_n$ is:

$$P^{(c_n)}(\mathbf{s},m,\mathbf{x},\mathbf{y},\mathbf{z},\hat{m})=W_S^n(\mathbf{s})P_M(m)f_n(\mathbf{x}|m,\mathbf{s})\times W^n_{Y,Z|X,S}(\mathbf{y},\mathbf{z}|\mathbf{x},\mathbf{s})\mathbb{1}_{\{\hat{m}=\phi_n(\mathbf{y})\}}. \quad (1)$$

The performance of c_n is evaluated in terms of its rate R, the maximal decoding error probability and the SS-metric.

Definition 2 (Maximal Error Probability) *The maximal error probability of an (n,R)-code c_n is*

$$e(c_n)=\max_{m\in\mathcal{M}_n}e_m(c_n), \quad (2)$$

where $e_m(c_n)=\sum_{(\mathbf{s},\mathbf{x})\in\mathcal{S}^n\times\mathcal{X}^n}W_S^n(\mathbf{s})f_n(\mathbf{x}|m,\mathbf{s})\sum_{\substack{\mathbf{y}\in\mathcal{Y}^n:\\ \phi_n(\mathbf{y})\neq m}}W^n_{Y|X,S}(\mathbf{y}|\mathbf{x},\mathbf{s}).$

Definition 3 (Information Leakage and SS Metric) *Let c_n be an (n,R)-code for the SD-WTC with non-causal encoder CSI and $P_M\in\mathcal{P}(\mathcal{M}_n)$. The information leakage to the eavesdropper under the message PMF $P_M\in\mathcal{P}(\mathcal{M}_n)$ is*

$$\ell(P_M,c_n)=I_{c_n}(M;\mathbf{Z}), \quad (3)$$

where the subscript c_n denotes that the mutual information term is calculated with respect to the marginal $P^{(c_n)}_{M,\mathbf{Z}}$ of (1).

The SS metric with respect to c_n is

$$\ell_{\mathrm{Sem}}(c_n) = \max_{P_M \in \mathcal{P}(\mathcal{M}_n)} \ell(P_M, c_n). \tag{4}$$

Remark 2 *SS requires that the code c_n works well for all message PMFs. This means that the mutual information term in* (4) *is maximized over P_M when the code c_n is known. In other words, although not stated explicitly, the maximizing P_M may depend on c_n.*

Definition 4 (Achievability) *A number $R \in \mathbb{R}_+$ is called an achievable SS-rate for the SD-WTC with non-causal encoder CSI, if for every $\epsilon > 0$ and sufficiently large n, there exists a CR (n, R)-code c_n with $e(c_n) \leq \epsilon$ and $\ell_{\mathrm{Sem}}(c_n) \leq \epsilon$.*

Definition 5 (SS-Capacity) *The SS-capacity C_{Sem} of the SD-WTC with non-causal encoder CSI is the supremum of the set of achievable SS-rates.*

B. Main Results

The main result of this work is a novel lower bound on the SS-capacity of the SD-WTC with non-causal encoder CSI. Our achievability formula strictly outperforms the best previously known coding scheme for the considered scenario. To state our main result, let $\mathcal{U}$ and $\mathcal{V}$ be finite alphabets and for any $Q_{U,V,X|S} : \mathcal{S} \to \mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{X})$ define

$$R_{\mathrm{A}}\left(Q_{U,V,X|S}\right) \triangleq \min \left\{ \begin{array}{c} I(V;Y|U) - I(V;Z|U), \\ I(U,V;Y) - I(U,V;S), \\ I(U,V;Y) - I(U;S) - I(V;Z|U) \end{array} \right\}, \tag{5}$$

where the mutual information terms are calculated with respect to the joint distribution $W_S Q_{U,V,X|S} W_{Y,Z|X,S}$.

Theorem 1 (SD-WTC SS-Capacity Lower Bound) *The SS-capacity of the SD-WTC with non-causal encoder CSI is lower bounded as*

$$C_{\mathrm{Sem}} \geq R_{\mathrm{A}} \triangleq \max_{Q_{U,V,X|S}} R_{\mathrm{A}}\left(Q_{U,V,X|S}\right), \tag{6}$$

and one may restrict the cardinalities of U and V to $|\mathcal{U}| \leq |\mathcal{S}||\mathcal{X}| + 5$ and $|\mathcal{V}| \leq |\mathcal{S}|^2|\mathcal{X}|^2 + 5|\mathcal{S}||\mathcal{X}| + 3$.

An extended outline of the proof of Theorem 1 is given in Section IV (see [10, Section VI-B] for the full details), and is based on a novel superposition coding scheme for secrecy. The superposition codebook encodes the entire secret message in its *outer layer*, meaning that no information is carried by the inner layer of the code. Nonetheless, the rate of the lower layer codebook is chosen to allow the eavesdropper to decode it. This results in the eavesdropper 'wasting' his channel resources on decoding the lower layer, leaving it with insufficient resources to unveil the secret message. The legitimate decoder, on the other hand, decodes both layers of the codebook. The transmission is correlated with the observed state sequence by means of the likelihood encoder [7] and SS is established using the stronger SCL (both the superposition version from [10, Lemma 1] and the heterogeneous SCL from [11, Lemma 1]).

Remark 3 (Cardinality Bounds) *The cardinality bounds on the auxiliary random variables U and V in Theorem 1 are established by standard application of the Eggleston-Fenchel-Carathéodory theorem twice. The details are omitted.*

III. Special Cases and Examples

A. Comparison to the Encoder and Decoder CSI Case

Consider the case when the state sequence $\mathbf{S}$ is also available to the legitimate receiver, i.e., when Y is replaced with (Y, S). The scenario when the encoder CSI is causal was studies by Chia and El-Gamal in [6], where a lower bound on the weak-secrecy capacity $C_{\mathrm{Weak}}^{\mathrm{Enc-Dec-CSI}}$ was established. To restate their result, let $\mathcal{T}$ be a finite set and for any $P_T \in \mathcal{P}(\mathcal{T})$ and $P_{X|T,S} : \mathcal{S} \times \mathcal{T} \to \mathcal{P}(\mathcal{X})$ define

$$R_{\mathrm{CE}}\left(P_T P_{X|T,S}\right) \triangleq \min \left\{ \begin{array}{c} I(T;Y|S)\ ,\ H(S|T,Z) \\ \quad + \left[I(T;Y,S) - I(T;Z)\right]^+ \end{array} \right\}, \tag{7a}$$

where $[x]^+ = \max(0, x)$ and the mutual information terms are calculated with respect to $W_S P_T P_{X|T,S} W_{Y,Z|X,S}$. Theorem 1 in [6] states that

$$C_{\mathrm{Weak}}^{\mathrm{Enc-Dec-CSI}} \geq R_{\mathrm{CE}}^{\mathrm{Enc-Dec-CSI}} \triangleq \max_{P_T P_{X|T,S}} R_{\mathrm{CE}}\left(P_T P_{X|T,S}\right). \tag{7b}$$

The independence between T and S is an outcome of the causality restriction on encoder CSI.

In effect, the result of [6, Theorem 1] was not expressed as in (7). Rather, the authors derived two separate lower bounds on $C_{\mathrm{Weak}}^{\mathrm{Enc-Dec-CSI}}$ and stated their achievability result as the maximum between the two. Be it as it may, it is readily verified that (7) is an equivalent representation of [6, Theorem 1]. Furthermore, [6, Remark 3.1] effectively asserts that whenever $I(T;Y,S) \geq I(T;Z)$, allowing correlation between T and S does not outcome in higher secrecy-rates. However, no such claim was established when the inequality is reversed.

Although studying the causal model, the authors of [6] showed that their result is at least as good as the best previously known scheme for the non-causal encoder CSI scenario. The latter scheme is obtained from [4, Theorem 2] - an achievable weak-secrecy rate for the SC-WTC with non-causal CSI at the encoder only - by replacing Y with (Y, S) (see Remark 1). All the more so, an example was provided in [6] where it is shown that in some cases $R_{\mathrm{CE}}^{\mathrm{Enc-Dec-CSI}}$ achieves strictly higher rates than [4, Theorem 2] (see also [5]). As stated in the following proposition, our achievable formula R_{A} is at least as good as R_{CE}, when the legitimate receiver also has access to $\mathbf{S}$.

To formulate the relation between the result of Theorem 1 and [6, Theorem 1], note that when the legitimate receiver also observes the state sequence, R_{A} becomes

$$R_{\mathrm{A}}^{\mathrm{Enc-Dec-CSI}} = \max_{Q_{U,V,X|S}} R_{\mathrm{A}}^{\mathrm{Enc-Dec-CSI}}\left(Q_{U,V,X|S}\right), \tag{8a}$$

where for any $Q_{U,V,X|S} : \mathcal{S} \to \mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{X})$, $R_{\mathrm{A}}^{\mathrm{Enc-Dec-CSI}}\left(Q_{U,V,X|S}\right)$ is obtained from R_{A} in (5) by replacing Y therein with (Y, S).

Proposition 2 *The following relation holds:*

$$\begin{aligned} R_{\mathrm{CE}}^{\mathrm{Enc-Dec-CSI}} &\leq \max_{P_{T,X|S}} R_{\mathrm{CE}}^{\mathrm{Enc-Dec-CSI}}\left(P_{T,X|S}\right) \\ &\leq R_{\mathrm{A}}^{\mathrm{Enc-Dec-CSI}}. \end{aligned} \tag{9}$$

The proof of Proposition 2 shows that $R_{\mathrm{A}}^{\mathrm{Enc-Dec-CSI}}$ recovers $R_{\mathrm{CE}}^{\mathrm{Enc-Dec-CSI}}$ by either setting $U = T$ and $V = S$ or setting $U = 0$ and $V = (T, S)$ (the choice of the auxiliaries varies depending on whether $I(T; Y, S) \geq I(T; Z)$ or not). The full details are omitted (see [10, Appendix E]).

A few remarks are at hand regarding the result of Proposition 2:

1) As seen in (9), our formula reduces to a maximization of $R_{\mathrm{CE}}^{\mathrm{Enc-Dec-CSI}}\left(P_{T,X|S}\right)$ over a domain of distribution that allow correlation between T and S. This is since our coding scheme was tailored for the non-causal CSI scenario, in contrast to the causal construction from [6] that results in restricting T and S to be independent. Although, this correlation is unnecessary when $I(T; Y, S) \geq I(T; Z)$, it may be the case that a correlated T and S are better when $I(T; Y, S) < I(T; Z)$.

2) The coding scheme from [6] uses the state sequence to explicitly generate a key (of the largest rate possible while still keeping the eavesdropper ignorant of it), which is used to one-time-pad a part of the confidential message; the other part of the message is protected via a wiretap code (whenever wiretap coding is possible). Our coding scheme, however, does not involve any explicit key generation (nor key agreement) phase. Instead, our code is based on a superposition codebook that fully encodes the confidential message in its outer layer, and SS is ensured by making the eavesdropper 'waste' channel resources on decoding the inner layer codeword that carries no confidential information whatsoever. Nonetheless, the relation between our scheme (when adjusted to the encoder-decoder CSI scenario) and the one-time-pad-based scheme from [6] is observed as follows. Note that to recover $R_{\mathrm{CE}}^{\mathrm{Enc-Dec-CSI}}$ from $R_{\mathrm{A}}^{\mathrm{Enc-Dec-CSI}}$ we introduce the state random variable S as part as the auxiliary random variable V. Doing so, essentially, uses the state sequence to randomize the choice of the transmitted codeword for a prescribed confidential message m. Since $\mathbf{S}$ is also known to the decoder, it can reverse this randomized choice and backtrack to the transmitted message. The eavesdropper, being ignorant of the state sequence, cannot do the same. This is an alternative perspective of the one-time-pad operation: randomly choosing a codeword from a cluster of codewords associated with each confidential message. Making these clusters large enough (so that they overlap), allows only a party that has access to the key used for the randomized choice to isolate the original message.

3) Our coding scheme outcomes in SS and a vanishing maximal error probability, while achieving possibly higher rates than [6], where only weak-secrecy and a vanishing average error probability were guaranteed. Thus, an upgrade of both performance metrics from [6] is possible, without inflicting any loss of rate. Furthermore, our scheme is based on a *single* transmission block, while [6, Theorem 1] relies on transmitting many such blocks. The purpose of a multiple-block transmission is to generate the key at each block from the state sequence of the previous block, thus simplifying the security analysis as far as the independence of the generated key and the eavesdropper's channel observation.

B. Comparison to Previous Schemes for the SD-WTC with Non-Causal Encoder CSI

The result of Theorem 1 recovers the previously best known achievable formula for the SD-WTC with non-causal encoder CSI from [4, Theorem 2]. Moreover, our achievability is strictly better than [4, Theorem 2] for some SD-WTCs. In [4, Theorem 2] it is stated that the weak-secrecy capacity of the considered SD-WTC is lower bounded by

$$R_{\mathrm{CHV}}^{\mathrm{Enc-CSI}} \triangleq \max_{P_{U,X|S}} R_{\mathrm{CHV}}^{\mathrm{Enc-CSI}}\left(P_{U,X|S}\right), \tag{10a}$$

where for any $P_{U,X|S} : \mathcal{S} \to \mathcal{P}(\mathcal{U} \times \mathcal{X})$,

$$R_{\mathrm{CHV}}^{\mathrm{Enc-CSI}}\left(P_{U,X|S}\right) \triangleq \min \left\{ \begin{array}{l} I(U;Y) - I(U;Z), \\ I(U;Y) - I(U;S) \end{array} \right\}, \tag{10b}$$

and the mutual information terms are taken with respect to $W_S P_{U,X|S} W_{Y,Z|X,S}$, i.e., $U - (X, S) - (Y, Z)$ forms a Markov Chain.

First note that Theorem 1 recovers $R_{\mathrm{CHV}}^{\mathrm{Enc-CSI}}$ by setting $U = 0$ in R_{A} (while relabeling V as U). Consequently,

$$R_{\mathrm{A}} \geq R_{\mathrm{CHV}}^{\mathrm{Enc-CSI}}. \tag{11}$$

On top of that, in [10, Section V-C] we provide an example that shows that there exist SD-WTCs for which the inequality in (11) is strict. Due to lack of space, the example and the derivation of the strict relation is not included in this document. The derivation relies on the main idea behind the example from [6] used for showing the superiority of their schemes over previously known achievable result for the SD-WTC with non-causal CSI at both the encoder and the decoder. Our example falls outside of the framework considered in [6].

IV. Outline of Proof of Theorem 1

We give a detailed description of the codebook construction and of the encoding and decoding processes. Due to space limitation, the analysis of reliability and SS is omitted and only the resulting rate bounds accompanied by broad explenations are stated. The reader is referred to [10, Section VI-B] for the full details. Fix $\epsilon > 0$ and a conditional PMF $Q_{U,V,X|S} : \mathcal{S} \to \mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{X})$.

Codebook $\mathcal{B}_n$: We use a superposition codebook where the outer layer also encodes the confidential message. The codebook is constructed independently of the state sequence

$\mathbf{S}$, but with sufficient redundancy to correlate the transmission with $\mathbf{S}$.

Let I and J be two independent random variables uniformly distributed over $\mathcal{I}_n \triangleq \left[1 : 2^{nR_1}\right]$ and $\mathcal{J}_n \triangleq \left[1 : 2^{nR_2}\right]$, respectively. Let $\mathcal{B}_U^{(n)} \triangleq \left\{\mathbf{u}(i, \mathcal{B}_U)\right\}_{i\in\mathcal{I}_n}$ be an inner layer codebook generated as i.i.d. samples of Q_U^n. For every $i \in \mathcal{I}_n$, let $\mathcal{B}_V^{(n)}(i) \triangleq \left\{\mathbf{v}(i, j, m, \mathcal{B}_V)\right\}_{(j,m)\in\mathcal{J}_n\times\mathcal{M}_n}$ be a collection of $2^{n(R_2+R)}$ vectors of length n drawn according to the distribution $Q^n_{V|U=\mathbf{u}(i,\mathcal{B}_U)}$. We use $\mathcal{B}_n$ to denote our superposition codebook, i.e., the collection of the inner and all the outer layer codebooks. The encoder and decoder are described next for a fixed superposition codebook $\mathcal{B}_n$.

Encoder $f_n^{(\mathcal{B}_n)}$: The encoding phase is based on the likelihood-encoder [7], which, in turn, allows us to approximate the (rather cumbersome) induced joint distribution by a much simpler distribution which we use for the analysis.

To send $m \in \mathcal{M}_n$ upon observing the state sequence $\mathbf{s} \in \mathcal{S}^n$, the encoder randomly chooses $(i, j) \in \mathcal{I}_n \times \mathcal{J}_n$ according to

$$P^{(\mathcal{B}_n)}(i, j|m, \mathbf{s}) = \frac{Q^n_{S|U,V}\big(\mathbf{s}\big|\mathbf{u}(i, \mathcal{B}_U), \mathbf{v}(i, j, m, \mathcal{B}_V)\big)}{\sum\limits_{(i',j')} Q^n_{S|U,V}\big(\mathbf{s}\big|\mathbf{u}(i', \mathcal{B}_U), \mathbf{v}(i', j', m, \mathcal{B}_V)\big)}, \tag{12}$$

where $Q_{S|U,V}$ is the conditional marginal of $Q_{S,U,V}$ defined by $Q_{S,U,V}(s, u, v) = \sum_{x\in\mathcal{X}} W_S(s)Q_{U,V,X|S}(u, v, x|s)$, for every $(s, u, v) \in \mathcal{S} \times \mathcal{U} \times \mathcal{V}$. The channel input sequence is then generated by feeding the chosen u- and v-codewords along with the state sequence into a DMC $Q_{X|U,V,S}$.

Decoder $\phi_n^{(\mathcal{B}_n)}$: Upon observing $\mathbf{y} \in \mathcal{Y}^n$, the decoder searches for a unique triple $(\hat{i}, \hat{j}, \hat{m}) \in \mathcal{I}_n \times \mathcal{J}_n \times \mathcal{M}_n$ such that $\Big(\mathbf{u}(\hat{i}, \mathcal{B}_U), \mathbf{v}(\hat{i}, \hat{j}, \hat{m}, \mathcal{B}_V), \mathbf{y}\Big) \in \mathcal{T}^n_\epsilon(Q_{U,V,Y})$. If such a unique triple is found, then set $\phi_n^{(\mathcal{B}_n)}(\mathbf{y}) = \hat{m}$; otherwise, $\phi_n^{(\mathcal{B}_n)}(\mathbf{y}) = e$.

The triple $(\mathcal{M}_n, f_n^{(\mathcal{B}_n)}, \phi_n^{(\mathcal{B}_n)})$ defined with respect to the codebook $\mathcal{B}_n$ constitutes an (n, R)-code c_n.

Main Idea Behind Analysis: The key step is to show that the joint PMF induced by the above encoding and decoding scheme, say $P^{(\mathcal{B}_n)}$, is close in total variation to a new (and simpler) distribution $\Gamma^{(\mathcal{B}_n)}$, which is used for the reliability and security analyses. For any $P_M \in \mathcal{P}(\mathcal{M}_n)$, $\Gamma^{(\mathcal{B}_n)}$ is defined by

$$\begin{aligned}\Gamma^{(\mathcal{B}_n)}&(m, i, j, \mathbf{u}, \mathbf{v}, \mathbf{s}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \hat{m}) = P_M(m)2^{-n(R_1+R_2)}\\ &\times \mathbb{1}_{\left\{\mathbf{u}=\mathbf{u}(i,\mathcal{B}_U)\right\}\cap\left\{\mathbf{v}=\mathbf{v}(i,j,m,\mathcal{B}_V)\right\}} Q^n_{S|U,V}(\mathbf{s}|\mathbf{u}, \mathbf{v})\\ &\times Q^n_{X|U,V,S}(\mathbf{x}|\mathbf{u}, \mathbf{v}, \mathbf{s})W^n_{Y,Z|X,S}(\mathbf{y}, \mathbf{z}|\mathbf{x}, \mathbf{s})\mathbb{1}_{\left\{\phi_n^{(\mathcal{B}_n)}(\mathbf{y})=\hat{m}\right\}},\end{aligned} \tag{13}$$

Namely, with respect to $\Gamma^{(\mathcal{B}_n)}$, the indices $(i, j) \in \mathcal{I}_n \times \mathcal{J}_n$ are drawn independently and uniformly each over its alphabet. The, the sequence $\mathbf{s}$ is generated by feeding the corresponding u- and v-codewords into the DMC $Q^n_{S|U,V}$. It can be shown the with respect to a random superposition codebook $\mathbb{B}_n$, $P^{(\mathbb{B}_n)}$ and $\Gamma^{(\mathbb{B}_n)}$ are close in total variation in several senses (both in expectation and with high probability), if

$$R_1 > I(U; S) \tag{14a}$$

$$R_1 + R_2 > I(U, V; S), \tag{14b}$$

The proof of the approximation is based on the stronger SCL for superposition codes from [10, Lemma 1].

Having that, both the reliability and the security analysis are preformed with respect to $\Gamma^{(\mathcal{B}_n)}$ instead of $P^{(\mathcal{B}_n)}$. Standard joint-typicality decoding arguments fir superposition codes show that reliability follows provided that

$$R + R_2 < I(V; Y|U), \tag{15a}$$

$$R + R_1 + R_2 < I(U, V; Y). \tag{15b}$$

With the help of the heterogeneous stronger SCL from [11, Lemma 1], SS is ensured if

$$R_2 > I(V; W|U). \tag{16}$$

This rate bound essentially means that the rates of the codebooks are chosen so that the eavesdropper can decode the inner layer codeword. This makes him waste channel resources on decoding a codeword that carries no confidential information. The remaining resources are insufficient for extracting any information on the outer layer codeword, which, in turn, results in our code being semantically-secure. Finally, applying the Fourier-Motzkin Elimination on (14), (15) and (16) shows that $R_{\mathrm{A}}\left(Q_{U,V,X|S}\right)$ is achievable.

REFERENCES

[1] S. I. Gelfand and M. S. Pinsker. Coding for channel with random parameters. *Problemy Pered. Inform. (Problems of Inf. Trans.)*, 9(1):19–31, 1980.

[2] A. V. Kuznetsov and B. S. Tsybakov. Coding in a memory with defective cells. *Problemy Pered. Inform. (Problems of Inf. Trans.)*, 10(2):52–60, 1974.

[3] K. Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inf. Theory*, 25(3):306–311, May 1979.

[4] Y. Chen and A. J. Han Vinck. Wiretap channel with side information. *IEEE Trans. Inf. Theory*, 54(1):395–402, Jan. 2008.

[5] W. Liu and B. Chen. Wiretap channel with two-sided state information. In *Proc. 41st Asilomar Conf. Signals, Syst. Comp*, page 893897, Pacific Grove, CA, US, Nov. 2007.

[6] Y.-K. Chia and A. El Gamal. Wiretap channel with causal state information. *IEEE Trans. Inf. Theory*, 58(5):2838–2849, May 2012.

[7] E. Song, P. Cuff, and V. Poor. The likelihood encoder for lossy compression. *IEEE Trans. Inf. Theory*, 62(4):1836–1849, Apr. 2016.

[8] M. Bellare, S. Tessaro, and A. Vardy. A cryptographic treatment of the wiretap channe. In *Proc. Adv. Crypto. (CRYPTO 2012)*, Santa Barbara, CA, USA, Aug. 2012.

[9] Z. Goldfeld, P. Cuff, and H. H. Permuter. Semantic-security capacity for wiretap channels of type II. *IEEE Trans. Inf. Theory*, 62(7):1–17, Jul. 2016.

[10] Z. Goldfeld, P. Cuff, and H. H. Permuter. Wiretap channel with random states non-causally available at the encoder. *To be submitted to IEEE Trans. Inf. Theory*, 2016.

[11] Z. Goldfeld, P. Cuff, and H. H. Permuter. Arbitrarily varying wiretap channels with type constrained states. *Submitted to IEEE Trans. Inf. Theory*, 2016. Available on ArXiv at http://arxiv.org/abs/1601.03660.