Wiretap and Gelfand-Pinsker Channels Analogy and its Applications

Conference Paper

Author(s):
Goldfeld, Ziv; Permuter, Haim H.

Publication date:
2018-02

Permanent link:
https://doi.org/10.3929/ethz-b-000245061

Wiretap and Gelfand-Pinsker Channels Analogy and its Applications

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

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

***Abstract*—A framework of analogy between wiretap channels (WTCs) and state-dependent point-to-point channels with non-causal encoder channel state information (referred to as Gelfand-Pinker channels (GPCs)) is proposed. A good (reliable and secure) sequence of wiretap codes is shown to induce a good (reliable) sequence of codes for a corresponding GPC. Consequently, the framework enables exploiting existing results for GPCs to produce converse proofs for their wiretap analogs. The fundamental limits of communication of two analogous wiretap and GP models are characterized by the same rate bounds; the optimization domains may differ. The analogy readily extends to multiuser broadcasting scenarios, encompassing broadcast channels (BCs) with deterministic components, degradation ordering between users, and BCs with cooperative receivers. Given a wiretap BC (WTBC) with two receivers and one eavesdropper, an analogous Gelfand-Pinsker BC (GPBC) is constructed by converting the eavesdropper's observation sequence to a state sequence with an appropriate product distribution, and non-causally revealing the states to the encoder. The transition matrix of the (state-dependent) GPBC is the appropriate conditional marginal of the WTBC's transition law, with the eavesdropper's output playing the role of the channel state. The analogy is exploited to characterize the secrecy-capacity regions of the SD-WTBC, which was an open problem until this work, based on the corresponding solution of the SD-GPBC.**

I. Introduction

Two fundamental, but seemingly unrelated, information-theoretic models are that of the wiretap channel (WTC) and the state-dependent point-to-point channel with non-causal encoder channel state information (CSI). The discrete and memoryless (DM) WTC (Fig. 1(a)) was introduced by Wyner in his celebrated 1975 paper [1] that initiated the study of physical layer security. Csiszár and Körner characterized the secrecy-capacity of the WTC as

$$C_{\mathsf{WT}}(p_{Y,Z|X}) = \max_{p_{U,X}} \Big[I(U;Y) - I(U;Z) \Big], \qquad (1)$$

where $p_{Y,Z|X}$ is the WTC's transition matrix and the underlying distribution is $p_{U,X}p_{Y,Z|X}$. The state-dependent channel with non-causal encoder CSI is due to Gelfand and Pinsker (GP) [2], and is henceforth referred to as the GP channel (GPC). A single-letter capacity formula for any GPC $q_{Y|X,Z}$ with state distribution q_Z was derived in [2]:

$$C_{\mathsf{GP}}(q_Z, q_{Y|X,Z}) = \max_{q_{U,X|Z}} \Big[I(U;Y) - I(U;Z) \Big], \qquad (2)$$

where the joint distribution is $q_Z q_{U,X|Z} q_{Y|X,Z}$. An interesting question is whether the resemblance of (1) and (2) is coincidental or is there an inherent relation between these problems.

This paper shows that an inherent relation is indeed the case, by proposing a rigorous framework that links the WTC and the GPC, establishing these two problems as analogous to one another. Specifically, we prove that any good (reliable and secure) sequence of codes for the WTC induces a good (reliable) sequence of codes of the same rate for a corresponding GPC. This observation enables exploiting known outer bounds on the GPC capacity to outer bound the secrecy-capacity of an analogous WTC. While the solutions to the base cases from Fig. 1 have been known for decades, many multiuser extensions of these models remain open problems. Through the analogy we derive a converse proof for the semi-deterministic (SD) wiretap broadcast channel (WTBC), an open problem until this work, thus characterizing its secrecy-capacity region.

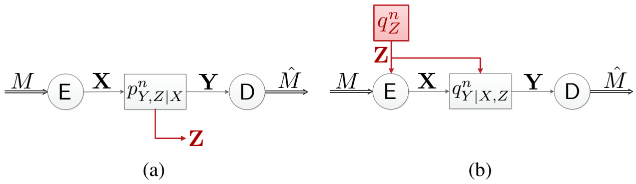


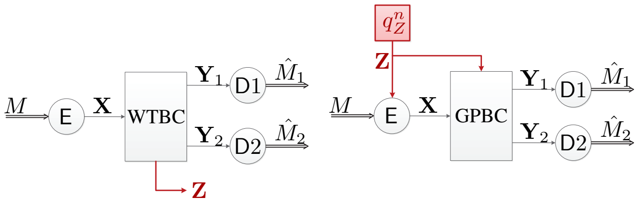
Fig. 1: (a) The WTC with transition probability $p_{Y,Z|X}$, where X is the channel input and Y and Z are the channel outputs observed by the legitimate receiver and the eavesdropper, respectively; (b) The GPC with state distribution $Z \sim q_Z$, and channel transition probability $q_{Y|X,Z}$, where X is the input and Y is the output.

To this end we extend the wiretap-GP analogy to multiuser broadcasting scenarios. Given a WTBC $p_{Y_1,Y_2,Z|X}$ (Fig. 2(a)), with two legitimate receivers observing Y_1 and Y_2 and one eavesdropper that intercepts Z, an analogous GP broadcast channel (GPBC), shown in Fig. 2(b), is constructed as follows:

1) Converting the eavesdropper's observation sequence Z^n to an independently and identically distributed (i.i.d.) state sequence with some appropriate distribution;
2) Revealing the state sequence in a non-causal manner to the encoder;
3) Setting the state-dependent BC $p_{Y_1,Y_2|X,Z}$ (the conditional marginal of the WTBC's transition probability) with Z in the role of the state.

The aforementioned relation between good sequences of codes for analogous WTBCs and GPBCs remains valid, which allows capitalizing on known GPBC capacity results to derive converse proofs for their analogous WTBC.

The GPBC has been widely studied in the literature and the capacity region is known for various cases [3]–[5]. Of particular interest is the capacity derivation of the SD-GPBC from [4]. WTBC also received considerable attention in the literature [6]–[8]; however, solutions are known only for some



(a) (b)

Fig. 2: (a) The WTBC with transition probability $p_{Y_1,Y_2,Z|X}$, where X is the channel input and Y_1, Y_2 and Z are the channel outputs observed by the legitimate receivers and the eavesdropper, respectively; (b) An analogous GPBC is obtained from the WTBC by replacing the eavesdropper's observation with a state random variable $Z \sim q_Z$, revealing $\mathbf{Z}$ in a non-causal manner to the encoder and setting the state-dependent BC $p_{Y_1,Y_2|X,Z}$ as the conditional marginal distribution of the WTBC's transition probability $p_{Y_1,Y_2,Z|X}$.

special cases. To the best of our knowledge, the widest framework of DM WTBCs for which tight secrecy-capacity results are available is due to [8], where, in particular, the region for the SD-WTBC was derived under a further assumption that the eavesdropper is less noisy than the stochastic receiver. The coding scheme therein remains feasible without this less-noisy property; the converse proofs, however, relies on it. Since no corresponding assumption was imposed while deriving the SD-GPBC result from [4], our analogy-based proof method characterizes the SD-WTBC secrecy-capacity regions without assuming this ordering between the sub-channels. As a natural extension to the analogy for the base case (WTCs versus GPCs), the obtained secrecy-capacity regions are described by the same rate bounds as their GPBC counterparts.

An important ingredient in proving the analogy is to adopt the definition of WTC achievability from, e.g., [7], [9], [10], that merges the reliability and security requirements into a single demand on the joint distribution induced by a wiretap code. Specifically, we require that a good sequence of wiretap codes induces a sequence of joint distributions (on the message, its estimate and the eavesdropper's observation) that is asymptotically indistinguishable in total variation from a target measure under which:

1) The message M and its estimate $\hat{M}$ are almost surely equal (a reliability requirement);
2) The eavesdropper's observation is independent of the message and is distributed according to some product measure, say q_Z (a security requirement).

Denoting by $P^{(c_n)}_{M,\hat{M},\mathbf{Z}}$ the joint distribution of M, $\hat{M}$ and $\mathbf{Z}$ induced by a wiretap code c_n, the above requirements mean that for large block lengths $P^{(c_n)}_{M,\hat{M},\mathbf{Z}} \approx P^{(c_n)}_{M}\mathbb{1}_{\{\hat{M}=M\}}q_Z^n$, where the approximation is in total variation.

With that notion of achievability, we then use distribution approximation arguments to show that such a sequence of wiretap codes induces a sequence of reliable codes for the analogous GPC. The GP encoder and decoder(s) are distilled from the joint distribution induced by the wiretap code by appropriately inverting it. Under this inversion, the asymptotic i.i.d. distribution of the eavesdropper's observation $\mathbf{Z}$ becomes the state distribution in the corresponding GPC. The asymptotic independence of $\mathbf{Z}$ and the message(s) in the WTC's target distribution corresponds to the independence of the message(s) and the state in a GP coding scenario. The performance metric described above strongly related to the more standard notion of achievability used in [11], where performance of a wiretap code was measured via the error probability and the effective secrecy metric. We show that under mild conditions (namely, a super-linear decay of the involved quantities), our definition of achievability and the one from [11] are equivalent.

II. Preliminary Definitions

We set up the problem of a WTBC, which is used in the next section for developing the analogy paradigm. The notations we use are from [12, Section II]. Let $\mathcal{X}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{Z}$ be finite sets (all alphabets throughout this work are assumed to be finite) and let $p_{Y_1,Y_2,Z|X} : \mathcal{X} \to \mathcal{P}(\mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Z})$ be a transition probability distribution from $\mathcal{X}$ to $\mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Z}$. The $\left(\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Z}, p_{Y_1,Y_2,Z|X}\right)$ DM-WTBC is illustrated in Fig. 2(a). The sender chooses a pair of messages (m_1, m_2) uniformly at random from product set $\left[1 : 2^{nR_1}\right] \times \left[1 : 2^{nR_2}\right]$ and maps it onto a sequence $\mathbf{x} \in \mathcal{X}^n$ (the mapping may be random). The sequence $\mathbf{x}$ is transmitted over the DM-WTBC with transition probability $p_{Y_1,Y_2,Z|X}$. The output sequences $\mathbf{y}_1 \in \mathcal{Y}_1^n$, $\mathbf{y}_2 \in \mathcal{Y}_2^n$ and $\mathbf{z} \in \mathcal{Z}^n$ are observed by Receiver 1, Receiver 2 and the eavesdropper, respectively. Based on $\mathbf{y}_j$, $j = 1, 2$, Receiver j produces an estimate $\hat{m}_j$ of m_j. The eavesdropper tries to glean whatever it can about the transmitted messages (m_1, m_2) from $\mathbf{z}$.

Definition 1 (WTBC Code) *An (n, R_1, R_2)-code c_n for the WTBC with a product message set $\mathcal{M}_1^{(n)} \times \mathcal{M}_2^{(n)}$, where for $j = 1, 2$ we set $\mathcal{M}_1^{(n)} \triangleq [1 : 2^{nR_j}]$, is a triple of functions $\left(f_n, \phi_1^{(n)}, \phi_2^{(n)}\right)$ such that $f_n : \mathcal{M}_1^{(n)} \times \mathcal{M}_2^{(n)} \to \mathcal{P}(\mathcal{X}^n)$ is a stochastic encoder, and $\phi_j^{(n)} : \mathcal{Y}_j^n \to \mathcal{M}_j^{(n)}$ is the decoding function for Receiver j, for $j = 1, 2$.*

For any (n, R_1, R_2)-code $c_n = \left(f_n, \phi_1^{(n)}, \phi_2^{(n)}\right)$, the induced joint distribution is:

$$P^{(c_n)}(m_{[1:2]}, \mathbf{x}, \mathbf{y}_{[1:2]}, \mathbf{z}, \hat{m}_{[1:2]}) = \frac{1}{\left|\mathcal{M}_1^{(n)}\right|\left|\mathcal{M}_2^{(n)}\right|} f_n(\mathbf{x}|m_{[1:2]}) \times p^n_{Y_1,Y_2,Z|X}(\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}|\mathbf{x}) \mathbb{1}_{\bigcap\limits_{j=1,2}\left\{\hat{m}_j = \phi_j^{(n)}(\mathbf{y}_j)\right\}}, \quad (3)$$

where $m_{[1:2]} \triangleq (m_1, m_2)$ and similarly for $\mathbf{y}_{[1:2]}$ and $\hat{m}_{[1:2]}$.

Our analogy relies on developing a unified perspective on two different problems. We arrive at the desired unification by defining achievability in a manner that is slightly different from typical definitions. Adopting the definition of achievability from [7], [9], [10], we merge the reliability and security requirements into a single requirement on the induced distribution from (3) phrased in terms of total variation.

Definition 2 (WTBC Achievability) *A pair of non-negative real numbers $(R_1, R_2) \in \mathbb{R}^2_+$ is called achievable if there exists a $\gamma > 0$, a probability distribution $q_Z \in \mathcal{P}(\mathcal{Z})$ and* *a sequence of (n, R_1, R_2)-codes $\{c_n\}_{n\in\mathbb{N}}$ such that for any sufficiently large n*

$$\left|\left|P^{(c_n)}_{M_{[1:2]},\hat{M}_{[1:2]},Z^n} - p^{(U)}_{\mathcal{M}_1^{(n)}\times\mathcal{M}_2^{(n)}}\mathbb{1}_{\{\hat{M}_{[1:2]}=M_{[1:2]}\}}q_Z^n\right|\right|_{\mathsf{TV}} \leq e^{-n\gamma}, \tag{4}$$

where $p^{(U)}_{\mathcal{A}}$ is the uniform distribution over a finite set $\mathcal{A}$.

Remark 1 (Rate of Convergence) *The exponential rate of convergence in* (4) *is not necessary. Any super-linear convergence rate is sufficient for the purposes of this work.*

Remark 2 (Equivalence to Standard Definitions) *The achievability definition in this work is equivalent to the more standard notion of achievability used in [11]. Therein, achievability was defined in terms of a vanishing average error probability and the* effective secrecy *metric that requires*

$$\begin{aligned}&\mathsf{D}\left(P^{(c_n)}_{M_1,M_2,Z^n}\middle|\middle|p^{(U)}_{\mathcal{M}_1^{(n)}\times\mathcal{M}_2^{(n)}}q_Z^n\right)\\ &\quad= \underbrace{I_{P^{(c_n)}}(M_1,M_2;Z^n)}_{\textit{Strong secrecy measure}} + \underbrace{\mathsf{D}\left(P^{(c_n)}_{Z^n}\middle|\middle|q_Z^n\right)}_{\textit{Stealth measure}} \end{aligned}\tag{5}$$

is made arbitrarily small. See [12, Section III-B] for details.

Remark 3 (Target i.i.d. Distribution) *The exact identity of target i.i.d. distribution q_Z^n that approximates the $P^{(c_n)}_{Z^n|M_{[1:2]},\hat{M}_{[1:2]}}$ in* (4) *and* (5) *cannot always be a priori determined solely based on the WTBC's transition kernel $p_{Y_1,Y_2,Z|X}$. The structure of q_Z depends on the sequence of codes $\{c_n\}_{n\in\mathbb{N}}$, and, typically, it can be understood from the proof of achievability.*[1] *Accordingly, the definition of achievability (Definition 2) does not shoot for a specific q_Z; rather, it just requires the existence of any q_Z satisfying* (4).

As usual, the *secrecy-capacity region* $\mathcal{C}_{\mathsf{WT}}(p_{Y_1,Y_2,Z|X})$ is the convex closure of the set of achievable rate pairs.

III. Wiretap and Gelfand-Pinsker Analogy

We describe the analogy principle for the base case of the classic wiretap and GP channels. As a first simple example, the analogy is used to derive a converse proof for the WTC's secrecy-capacity theorem. Then, we outline extensions of this idea to multiuser (namely, broadcasting) scenarios. These extension are subsequently used to prove the main secrecy-capacity results of this work that are stated in Section IV.

A. The Base Case - A Unified Perspective

For simplicity of presentation consider the classic wiretap and GPCs. These problems are related through the fact that their target joint distributions share the same structure. To see this, consider the $p_{Y,Z|X}$ WTC, for which achievability is defined similarly to Definition 2, and the point-to-point GPC with state distribution q_Z and channel transition probability $q_{Y|X,Z}$.[2] The joint distribution induced by an (n, R)-code $c_n = (f_n, \phi_n)$ for the wiretap channel is (see (3))

$$\tilde{P}^{(c_n)}(m,\mathbf{x},\mathbf{y},\mathbf{z},\hat{m}) = \frac{1}{|\mathcal{M}_n|}f_n(\mathbf{x}|m)p^n_{Y,Z|X}(\mathbf{y},\mathbf{z}|\mathbf{x})\mathbb{1}_{\left\{\hat{m}=\phi_n(\mathbf{y})\right\}} \tag{6}$$

while the induced distribution for the GPC with respect to an (n, R)-code $b_n = (g_n, \psi_n)$, where $g_n : \mathcal{M}_n \times \mathcal{Z} \to \mathcal{P}(\mathcal{X})$ is a stochastic encoder and $\phi_n : \mathcal{Y}^n \to \mathcal{M}_n$ is the decoder, is

$$\tilde{Q}^{(b_n)}(\mathbf{z}, m, \mathbf{x}, \mathbf{y}, \hat{m}) = q_Z^n(\mathbf{z})\frac{1}{|\mathcal{M}_n|}g_n(\mathbf{x}|\mathbf{z}, m)q^n_{Y|X,Z}(\mathbf{y}|\mathbf{x}, \mathbf{z}) \times \mathbb{1}_{\left\{\hat{m}=\psi_n(\mathbf{y})\right\}}. \tag{7}$$

With respect to Definition 2, a non-negative real number R is achievable for the WTC if there exist a distribution $q_Z \in \mathcal{P}(\mathcal{Z})$ and a sequence of (n, R)-codes $\{c_n\}_{n\in\mathbb{N}}$, such that

$$\left|\left|\tilde{P}^{(c_n)}_{M,\hat{M},Z^n} - p^{(U)}_{\mathcal{M}_n}\mathbb{1}_{\{\hat{M}=M\}}q_Z^n\right|\right|_{\mathsf{TV}} \xrightarrow[n\to\infty]{} 0. \tag{8}$$

For the GPC, it can be shown that under mild conditions,[3] a vanishing error probability is equivalent to

$$\left|\left|\tilde{Q}^{(c_n)}_{M,\hat{M},Z^n} - p^{(U)}_{\mathcal{M}_n}\mathbb{1}_{\{\hat{M}=M\}}q_Z^n\right|\right|_{\mathsf{TV}} \xrightarrow[n\to\infty]{} 0. \tag{9}$$

For details, see [12, Section IV-A-1].

Having (8) and (9), it is evident that while each problem has its own induced joint distribution, their target measures share the same structure. In both problems, a "good" sequence of codes induces a sequence of distributions ($\left\{\tilde{P}^{(c_n)}\right\}_{n\in\mathbb{N}}$ or $\left\{\tilde{Q}^{(b_n)}\right\}_{n\in\mathbb{N}}$ for the WTC or the GPC, respectively) that approximates a target distribution where: (i) $M = \hat{M}$ almost surely; (ii) $\mathbf{Z}$ is independent of M. The first item is a consequence of the reliability requirement in both problems. For the second item, note that, while the independence of $\mathbf{Z}$ and M is the security requirement in the WTC scenario, it is actually part of the problem definition for the GPC. The above described correspondence between the WTC and the GPC stands at the heart of the analogy between them.

B. Analogy Between Multiuser Setups

As a natural extension to the ideas from Section III-A, we now describe the analogy between WTBCs and GPBCs. Consider a WTBC $\left(\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Z}, p_{Y_1,Y_2,Z|X}\right)$ as defined in Section II. An analogous GPBC is constructed in three steps (see Fig. 2):

1) Replace the eavesdropper of the WTBC with a state sequence $\mathbf{Z} \sim q_Z^n$, where q_Z^n is the target product measure from the definition of WTBC achievability (see Definition 2);
2) Non-causally reveal $\mathbf{Z}$ to the encoder;
3) Set the GPBC's transition probability as the conditional marginal distribution $p_{Y_1,Y_2|X,Z}$.

The produced analogous $\left(\mathcal{Z}, \mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, q_Z, p_{Y_1,Y_2|X,Z}\right)$ GPBC inherits the properties the WTBC possesses (e.g.,

[1] For instance, for the degraded binary symmetric WTBC with crossover probabilities p_L and p_E for the legitimate and eavesdropper channels, respectively, where $p_L < p_B$, one may verify that q_Z may be chosen as a product $\mathsf{Ber}\left(\frac{1}{2}\right)$ measure. This is a consequence of the optimal input distribution that attains that secrecy-capacity $h(p_E) - h(p_L)$ being $\left(\mathsf{Ber}\left(\frac{1}{2}\right)\right)^n$.

[2] We adhere to the standard definitions for GPCs, see, e.g., [13, Setion 7.6].

[3] namely, a super-linear decay of the error probability

deterministic components, order of degradeness, etc). For example, if the WTBC is SD $p_{Y_1,Y_2,Z|X} = \mathbb{1}_{\{Y_1=y_1(X)\}} p_{Y_2,Z|X}$, then so is the GPBC since $p_{Y_1,Y_2|X,Z} = \mathbb{1}_{\{Y_1=y_1(X)\}} p_{Y_2|X,Z}$. If one of the observed signals of the legitimate receivers is a degraded version of the other, then the same ordering applies for the signal intercepted by the receivers of the GPBC. The analogy also accounts for WTBC settings with cooperative components. Namely, if the receivers of the WTBC are connected by, e.g., a finite-capacity bit-pipe, then the same applies for the receivers of the analogous GPBC.

As for the base case, the capacity regions of two analogous wiretap and GP BCs are described by rate bounds of the same structure. The underlying distribution and the part thereof over which we take the union is, however, different. This relation between the regions is emphasized in Section IV.

Since GPBCs have been extensively treated in the literature and capacity results are available for numerous cases [3]–[5], the analogy allows leveraging these results to study corresponding WTBCs. This is done by relating the performance of two analogous models as follows. Due to lack of space, the proof of the proposition is omitted; the reader is referred to [12] for details.

Proposition 1 (Good Wiretap Codes and Good GP Codes) *Consider a* $\big(\mathcal{X},\mathcal{Y}_1,\mathcal{Y}_2,\mathcal{Z},p_{Y_1,Y_2,Z|X}\big)$ *WTBC. Let* $(R_1,R_2)\in\mathbb{R}_+^2$ *be an achievable rate pair for the WTBC, with a corresponding sequence of* (n,R_1,R_2)*-codes* $\{c_n\}_{n\in\mathbb{N}}$*, where* $c_n = \Big(f_n,\phi_1^{(n)},\phi_2^{(n)}\Big)$*, for each* $n\in\mathbb{N}$*. For every* $n\in\mathbb{N}$*, define* $g_n \triangleq P^{(c_n)}_{\mathbf{X}|\mathbf{Z},M_{[1:2]}}$ *and* $\psi_j^{(n)} \triangleq \phi_j^{(n)}$*, for* $j=1,2$*, where* $P^{(c_n)}_{\mathbf{X}|\mathbf{Z},M_{[1:2]}}$ *is the conditional marginal distribution of* $\mathbf{X}$ *given* $(\mathbf{Z},M_1,M_2)$ *with respect to* $P^{(c_n)}$ *from* (3) *induced by the* n*-th wiretap code* c_n*. Then:*

1) $b_n \triangleq \Big(g_n,\psi_1^{(n)},\psi_2^{(n)}\Big)$ *is an* (n,R_1,R_2)*-code for the* $\big(\mathcal{Z},\mathcal{X},\mathcal{Y}_1,\mathcal{Y}_2,q_Z,p_{Y_1,Y_2|X,Z}\big)$ *GPBC.*

2) The distribution $Q^{(b_n)}_{\mathbf{Z},M_{[1:2]},\mathbf{X},\mathbf{Y}_{[1:2]},\hat{M}_{[1:2]}}$ *induced by* b_n *(analogous to* $\tilde{Q}^{(b_n)}$ *from* (7) *with* $M_{[1:2]}$*,* $\mathbf{Y}_{[1:2]}$ *and* $\hat{M}_{[1:2]}$ *in the roles of* M*,* $\mathbf{Y}$ *and* $\hat{M}$ *therein, respectively) satisfies* $\big|\big|P^{(c_n)} - Q^{(b_n)}\big|\big|_{\mathsf{TV}} \le e^{-n\gamma}$*, for any* n *large enough.*

3) The sequence of codes $\{b_n\}_{n\in\mathbb{N}}$ *attains* $\mathsf{P_e}(b_n) \xrightarrow[n\to\infty]{} 0$*, and consequently,* (R_1,R_2) *is an achievable rate pair for the aforementioned GPBC.*

Proof: For simplicity of notation, throughout the proof we denote $M_{12}\triangleq M_{[1:2]}$*,* $m_{12}\triangleq m_{[1:2]}$*,* $\hat{M}_{12}\triangleq \hat{M}_{[1:2]}$*,* $\hat{m}_{12}\triangleq\hat{m}_{[1:2]}$ *and* $\mathcal{M}_{12}\triangleq\mathcal{M}_1^{(n)}\times\mathcal{M}_2^{(n)}$*. The first claim is straightforward as for each* $n\in\mathbb{N}$*,* $P^{(c_n)}_{\mathbf{X}|\mathbf{Z},M_{12}}$ *and* $\psi_j^{(n)}$*, for* $j=1,2$*, are valid (stochastic) encoder and decoders for the GPBC. For (2), fix* $n\in\mathbb{N}$*, and first observe*

$$\begin{aligned} &P^{(c_n)}_{M_{12},\mathbf{X},\mathbf{Y}_{[1:2]},\mathbf{Z},\hat{M}_{12}} \\ &\quad\stackrel{(a)}{=} P^{(c_n)}_{M_{12},\mathbf{Z}}\cdot g_n\cdot p^n_{Y_1,Y_2|X,Z}\cdot \mathbb{1}_{\bigcap_{j=1,2}\left\{\hat{M}_j=\psi_j^{(n)}(\mathbf{Y}_j)\right\}} \\ &\quad\stackrel{(b)}{=} P^{(c_n)}_{M_{12},\mathbf{Z}}\cdot Q^{(b_n)}_{\mathbf{X},\mathbf{Y}_{[1:2]},\hat{M}_{12}|M_{12},\mathbf{Z}} \end{aligned} \tag{10}$$

where (a) follows by the factorization of $P^{(c_n)}$ *from* (3)*, while (b) is because* $b_n=\Big(g_n,\psi_1^{(n)},\psi_2^{(n)}\Big)$ *and due to the structure of* $Q^{(b_n)}$*. Recalling that* $Q^{(c_n)}_{\mathbf{Z},M_{12}} = q_Z^n\cdot p^{(U)}_{\mathcal{M}_{12}}$*, we have*

$$\big|\big|P^{(c_n)} - Q^{(b_n)}\big|\big|_{\mathsf{TV}} = \Big|\Big|P^{(c_n)}_{M_{12},\mathbf{Z}} - p^{(U)}_{\mathcal{M}_{12}}\cdot q_Z^n\Big|\Big|_{\mathsf{TV}} \xrightarrow[n\to\infty]{} 0. \tag{11}$$

Claim (3) follows because $\mathsf{P_e}(b_n)$ *is upper bounded as*

$$\begin{aligned} \mathsf{P_e}(b_n) &= \sum_{\substack{m_{12},\hat{m}_{12}:\\ m_{12}\neq\hat{m}_{12}}}\Big[Q^{(c_n)}(m_{12},\hat{m}_{12}) - p^{(U)}_{\mathcal{M}_{12}}(m_{12})\mathbb{1}_{\{\hat{m}_{12}=m_{12}\}}\Big] \\ &\stackrel{(a)}{=} \Big|\Big|Q^{(c_n)}_{M_{12},\hat{M}_{12}} - p^{(U)}_{\mathcal{M}_1^{(n)}\times\mathcal{M}_2^{(n)}}\mathbb{1}_{\left\{\hat{M}_{12}=M_{12}\right\}}\Big|\Big|_{\mathsf{TV}} \\ &\stackrel{(b)}{\le} \Big|\Big|Q^{(b_n)}_{M_{12},\hat{M}_{12}} - P^{(c_n)}_{M_{12},\hat{M}_{12}}\Big|\Big|_{\mathsf{TV}} \\ &\qquad\qquad + \Big|\Big|P^{(c_n)}_{M_{12},\hat{M}_{12}} - p^{(U)}_{\mathcal{M}_{12}}\mathbb{1}_{\{\hat{M}_{12}=M_{12}\}}\Big|\Big|_{\mathsf{TV}} \\ &\stackrel{(c)}{\le} \Big|\Big|Q^{(b_n)} - P^{(c_n)}\Big|\Big|_{\mathsf{TV}} + \Big|\Big|P^{(c_n)}_{M_{12},\hat{M}_{12},\mathbf{Z}} - p^{(U)}_{\mathcal{M}_{12}}\mathbb{1}_{\{\hat{M}_{12}=M_{12}\}}q_Z^n\Big|\Big|_{\mathsf{TV}} \end{aligned}$$

where (a) is because $||p-q||_{\mathsf{TV}} = \sum_{x:\ p(x)>q(x)}\big[p(x)-q(x)\big]$ *and since* $m_{12}\neq\hat{m}_{12}$ *if and only if* $Q^{(c_n)}(m_{12},\hat{m}_{12}) \ge p^{(U)}_{\mathcal{M}_{12}}(m_{12})\mathbb{1}_{\left\{\hat{m}_{12}=m_{12}\right\}}$*; (b) is the triangle inequality; (c) uses Property (3-a) from [12, Lemma 1]. Finally, the RHS above vanishes to 0 as* $n\to\infty$ *by* (11) *and our hypothesis.* ■

IV. The Secrecy-Capacity Region of the SD-WTBC

We give a single-letter characterization of the secrecy-capacity region of the SD-WTBC. A WTBC is SD if $p_{Y_1,Y_2,Z|X} = \mathbb{1}_{\{Y_1=y_1(X)\}} p_{Y_2,Z|X}$, where $y_1:\mathcal{X}\times\mathcal{Z}\to\mathcal{Y}_1$ and $p_{Y_2,Z|X}:\mathcal{X}\to\mathcal{P}(\mathcal{Y}_2\times\mathcal{Z})$. Until now, the secrecy-capacity region of this setup was known only under the assumption that the stochastic channel is less noisy than the channel to the eavesdropper [8, Theorem 5]. Our analogy-based converse proof makes this assumption unnecessary.

Theorem 1 (Secrecy-Capacity) *The secrecy-capacity region of the* $\big(\mathcal{X},\mathcal{Y}_1,\mathcal{Y}_2,\mathcal{Z},\mathbb{1}_{\{Y_1=y_1(X)\}}p_{Y_2,Z|X}\big)$ *SD-WTBC is given by the union of rate pairs* $(R_1,R_2)\in\mathbb{R}_+^2$ *satisfying:*

$$R_1 \le H(Y_1|Z), \tag{12a}$$

$$R_2 \le I(U;Y_2) - I(U;Z), \tag{12b}$$

$$R_1+R_2 \le H(Y_1|Z) + I(U;Y_2) - I(U;Y_1,Z) \tag{12c}$$

where the union is over all $p_{U,X}\in\mathcal{P}(\mathcal{U}\times\mathcal{X})$*, each inducing a joint distribution* $p_{U,X}\mathbb{1}_{\{Y_1=y_1(X)\}}p_{Y_2,Z|X}$*. Furthermore, one may restrict the auxiliary random variable* U *to take values in a set* $\mathcal{U}$ *whose cardinality is bounded by* $|\mathcal{U}|\le|\mathcal{X}|+1$*.*

The direct part of Theorem 1 relies on a specialization of the inner bound on the secrecy-capacity region of the WTBC derived in [7, Theorem 3]. As the performance criterion in that work corresponds to the definition of achievability used herein (Definition 2), the result from [7] applies for our setup. Setting $Q=U_0=0$, $U_1=Y_1$ and recasting U_2 as U reduces the rate bounds from [7, Theorem 3] to those from (12). Since $Y_1 = y_1(X)$, this choice of the auxiliaries $\big(Q,U_{[0:2]}\big)$ is feasible. The analogy-based converse proof is given next.

Converse Proof: Let $(R_1,R_2)\in\mathbb{R}_+^2$ be an achievable rate pair for the SD-WTBC and $\{c_n\}_{n\in\mathbb{N}}$ be the corresponding sequence of (n,R_1,R_2)-codes satisfying (4) for some $\gamma>0$ and $q_Z\in\mathcal{P}(\mathcal{Z})$, and any n large enough. By Proposition 1, $\{c_n\}_{n\in\mathbb{N}}$ gives rise to a sequence of (n,R_1,R_2)-codes $\{b_n\}_{n\in\mathbb{N}}$ for the $\big(\mathcal{Z},\mathcal{X},\mathcal{Y}_1,\mathcal{Y}_2,q_Z,p_{Y_1,Y_2|X,Z}\big)$ GPBC, each inducing a joint distribution $Q^{(b_n)}$, such that:

1) $\big|\big|P^{(c_n)}-Q^{(b_n)}\big|\big|_{\mathsf{TV}}\leq e^{-n\gamma}$, for any large enough n, where $P^{(c_n)}$ is the distribution from (3) induced by c_n.
2) $\mathsf{P}_{\mathsf{e}}(b_n)\xrightarrow[n\to\infty]{}0$.

Furthermore, note that since the WTBC is SD, i.e., its transition probability factors as $p_{Y_1,Y_2,Z|X}=\mathbb{1}_{\{Y_1=y_1(X)\}}p_{Y_2,Z|X}$, the obtained GPBC is also SD. Namely, the GPBC's transition probability decomposes as $p_{Y_1,Y_2|X,Z}=\mathbb{1}_{\{Y_1=y_1(X)\}}p_{Y_2|X,Z}$, which falls under the framework of [4, Theorem 1].

The converse proof of [4, Theorem 1] for the SD-GPBC shows that if $\{b_n\}_{n\in\mathbb{N}}$ is a sequence of (n,R_1,R_2)-codes with a vanishing error probability, then

$$R_1\leq\frac{1}{n}\sum_{i=1}^{n}H_Q(Y_{1,i}|Z_i)+\epsilon_n \tag{13a}$$

$$R_2\leq\frac{1}{n}\sum_{i=1}^{n}\Big[I_Q\big(M_2,Y_2^{i-1},Z_{i+1}^n;Y_{2,i}\big)-I_Q\big(M_2,Y_2^{i-1},Z_{i+1}^n;Z_i\big)\Big]+\epsilon_n \tag{13b}$$

$$R_1+R_2\leq\frac{1}{n}\sum_{i=1}^{n}\Big[I_Q\big(M_2,Y_2^{i-1},Z_{i+1}^n,Y_{1,i+1}^n;Y_{2,i}\big)+H_Q(Y_{1,i}|Z_i)-I_Q\big(M_2,Y_2^{i-1},Z_{i+1}^n,Y_{1,i+1}^n;Z_i,Y_{2,i}\big)\Big]+\epsilon_n, \tag{13c}$$

where the subscript Q indicates that the underlying distribution is $Q^{(b_n)}$ and $\epsilon_n\triangleq\frac{2}{n}+\mathsf{P}_{\mathsf{e}}(b_n)\sum_{j=1,2}R_j$. Since the total variation of two distribution upper bounds the total variation between their marginals [12, Property (3-a), Lemma 1]),

$$\Big|\Big|P^{(c_n)}_{M_2,Y^i,Z_i^n}-Q^{(b_n)}_{M_2,Y^i,Z_i^n}\Big|\Big|_{\mathsf{TV}}\leq e^{-n\gamma} \tag{14}$$

for large n, uniformly in $i\in[1:n]$. Recall that over finite probability spaces an exponentially decaying total variation dominates the difference between two corresponding mutual information terms (see [12, Lemma 3]). Combining this observation with (14), we may replace the information measures from the RHS of (13) that are taken with respect to $Q^{(b_n)}$ with the same terms, but with an underlying distribution $P^{(c_n)}$ (which we denote by a subscript P) plus a vanishing term. Namely, there exists a $\delta>0$, such that for n large enough

$$R_1\leq\frac{1}{n}\sum_{i=1}^{n}H_P(Y_{1,i}|Z_i)+\epsilon_n+e^{-n\delta} \tag{15a}$$

$$R_2\leq\frac{1}{n}\sum_{i=1}^{n}\Big[I_P(V_i;Y_{2,i})-I_P(V_i;Z_i)\Big]+\epsilon_n+2e^{-n\delta} \tag{15b}$$

$$R_1+R_2\leq\frac{1}{n}\sum_{i=1}^{n}\Big[H_P(Y_{1,i}|Z_i)+I_P(V_i,T_i;Y_{2,i})-I_P(V_i,T_i;Y_{1,i},Z_i)\Big]+\epsilon_n+3e^{-n\delta} \tag{15c}$$

where, for every $i\in[1:n]$, we have defined $V_i\triangleq\big(M_2,Y_2^{i-1},Z_{i+1}^n\big)_P$ and $T_i\triangleq\big(Y_{1,i+1}^n\big)_P$, with the subscript P indicating that the underlying distribution is $P^{(c_n)}$.

Letting n tend to infinity in (15), we see that any achievable rate pair (R_1,R_2) must be contained in the convex closure of the union of rate pairs satisfying:

$$R_1\leq H_p(Y_1|Z) \tag{16a}$$

$$R_2\leq I_p(V;Y_2)-I_p(V;Z) \tag{16b}$$

$$R_1+R_2\leq H_p(Y_1|Z)+I_p(V,T;Y_2)-I_p(V,T;Y_1,Z) \tag{16c}$$

where the union is over all $p_{V,T,X}\in\mathcal{P}(\mathcal{V}\times\mathcal{T}\times\mathcal{X})$, each inducing a joint distribution $p\triangleq p_{V,T,X}p_{Y_1,Y_2,Z|X}$, i.e., $(Y_1,Y_2,Z)\multimap X\multimap(V,T)$ forms a Markov chain. This Markov relation follows because $(Y_{1,i},Y_{2,i},Z_i)\multimap X_i\multimap\big(M_2,Y_{1,i+1}^n,Y_2^{i-1},Z_{i+1}^n\big)$, for all $i\in[1:n]$, under $P^{(c_n)}$.

To conclude the proof it remains to show that there exists an auxiliary random variable U, such that for any (V,T):

$$I_p(V;Y_2)-I_p(V;Z)\leq I_p(U;Y_2)-I_p(U;Z) \tag{17a}$$

$$H_p(Y_1|Z)+I_p(V,T;Y_2)-I_p(V,T;Y_1,Z)\leq H_p(Y_1|Z)+I_p(U;Y_2)-I_p(U;Y_1,Z). \tag{17b}$$

This is established by closely following the arguments from the end of the converse proof of the analogous SD-GPBC [4, Section III], as outlined next. Setting $U=V$ if p is such that $I_p(T;Y_2|V)-I_p(T;Y_1,Z|V)\leq 0$, and $U=(V,T)$ if $I_p(T;Y_2|V)-I_p(T;Z|V)\geq 0$ suffices. Finally, noting that every distribution p must satisfy at least one of these information inequalities concludes the converse.

References

[1] A. D. Wyner. The wire-tap channel. *Bell Sys. Techn.*, 54(8):1355–1387, Oct. 1975.

[2] 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.

[3] Y. Steinberg. Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information. *IEEE Trans. Inf. Theory*, 51(8):2867–2877, Aug. 2005.

[4] A. Lapidoth and L. Wang. The state-dependent semideterministic broadcast channel. *IEEE Trans. Inf. Theory*, 59(4):2242–2251, 2013.

[5] L. Dikstein, H. H. Permuter, and Y. Steinberg. On state-dependent broadcast channels with cooperation. *IEEE Trans. Inf. Theory*, 62(5):2308–2323, May 2016.

[6] E. Ekrem and S. Ulukus. Multi-receiver wiretap channel with public and confidential messages. *IEEE Trans. Inf. Theory*, 59(4):2165–2177, Apr. 2013.

[7] M. H. Yassaee, M. R. Aref, and A. Gohari. Achievability proof via output statistics of random binning. *IEEE Trans. Inf. Theory*, 60(11):6760–6786, Nov. 2014.

[8] M. Benammar and P. Piantanida. Secrecy capacity region of some classes of wiretap broadcast channels. *IEEE Trans Inf. Theory*, 61(10):5564–5582, Oct. 2015.

[9] H. Tyagi and S. Watanabe. Converses for secret key agreement and secure computing. *IEEE Trans. Inf. Theory*, 61(9):4809–4827, 2015.

[10] M. H. Yassaee. One-shot achievability via fidelity. In *Proc. Int. Symp. Inf. Theory (ISIT-2015)*, pages 301–305, Hong Kong, China, Jun. 2015.

[11] J. Hou and G. Kramer. Effective secrecy: Reliability, confusion and stelth. In *IEEE Int. Symp. Inf. Theory (ISIT-2014)*, Honolulu, HI, USA, Jun.-Jul. 2014.

[12] Z. Goldfeld, , and H. H. Permuter. Wiretap and gelfand-pinsker channels analogy and its applications. *Submitted for publication to IEEE Trans. Inf. Theory*, 2017.

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