A Useful Analogy Between Wiretap and Gelfand-Pinsker Channels

Ziv Goldfeld
MIT
zivg@mit.edu

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. The analogy is exploited to characterize the secrecy-capacity regions of the semi-deterministic WTBC (an open problem until this work) and a class of physically degraded WTBC. The derivations are based on known solutions for the corresponding GPBCs.**

I. Introduction

Two fundamental, yet seemingly unrelated, information-theoretic models are 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 1975 [1] and initiated the study of physical layer security. Csiszár and Körner characterized the WTC's secrecy-capacity 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]; we henceforth referred to it as the GP channel (GPC). The capacity of a GPC $q_{Y|X,Z}$ with state distribution q_Z is:

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

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 upper bounds on the GPC capacity to upper 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. The secrecy-capacity region of a certain class of cooperative physically-degraded (PD) WTBCs is also derived.

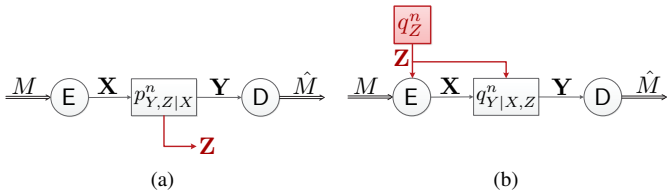

Fig. 1: (a) The WTC with transition probability $p_{Y,Z|X}$; (b) The GPC with state distribution q_Z and channel transition probability $q_{Y|X,Z}$.

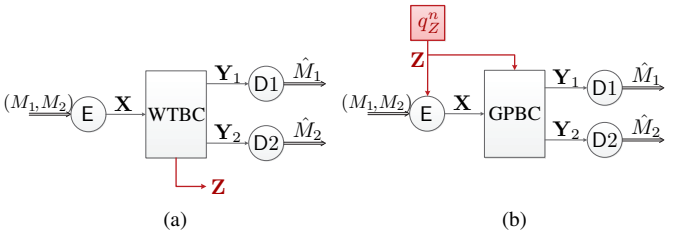

Fig. 2: (a) The WTBC with transition probability $p_{Y_1,Y_2,Z|X}$; (b) An analogous GPBC is obtained by replacing the eavesdropper's observation with a state sequence $\mathbf{Z} \sim q_Z^n$, non-causally revealing $\mathbf{Z}$ to the encoder and setting the channel law to $p_{Y_1,Y_2|X,Z}$.

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 by:

1) Converting the eavesdropper's observation Z^n to an independently and identically distributed (i.i.d.) state sequence with some appropriate distribution;
2) Non-causally revealing the state sequence to the encoder;
3) Setting $p_{Y_1,Y_2|X,Z}$ (the conditional marginal of the WTBC's transition probability, with Z in the role of the state) as the GPBC's transition kernel.

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

The GPBC has been widely studied in the literature and the capacity region is known for various cases, such as PD-GPBC without and with cooperative receivers [3], [4], SD-GPBCs [5], and more. WTBC also received considerable attention in

the literature [6]–[8]; however, solutions are known only for some special cases. In particular, the secrecy-capacity of SD-WTBC is known only under the assumption that the stochastic receiver is less noisy than the eavesdropper [8]. The coding scheme therein does not rely on this less-noisy property; the converse proofs, however, do. Since no corresponding assumption was imposed on the SD-GPBC from [5], our analogy-based proof characterizes the SD-WTBC's secrecy-capacity region without assuming the aforementioned ordering between the sub-channels. As a natural extension to the analogy for the base case (WTCs versus GPCs), the obtained secrecy-capacity region is described by the same rate bounds as these in the capacity characterization of the SD-GPBC from [9].

An important ingredient in proving the analogy is to adopt the definition of WTC achievability from, e.g., [7], [10], [11], that merges the reliability and security requirements into a single demand on the joint distribution induced by the 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 (TV) from a target measure under which:

1) The message M and its estimate $\hat{M}$ almost surely coincide (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 $P^{(c_n)}_{M,\hat{M},\mathbf{Z}} \approx P^{(c_n)}_{M} \mathbb{1}_{\{\hat{M}=M\}} q_Z^n$, for large block lengths. We highlight that this security requirement is twofold. First, it dictates that the message should be asymptotically independent of the eavesdropper's observed signal - a secrecy requirement. Second, the marginal distribution of the eavesdropper's signal should be asymptotically indistinguishable from a product distribution q_Z^n, a feature known as 'stealth'. The latter property plays an important role in establishing the proposed analogy, as it allow to relate the eavesdropper's signal to the i.i.d. state sequence in the GPC.

With this notion of achievability, we then 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 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 [12], 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 [12] are equivalent.

II. Wiretap Broadcast Channels

We use notations from [13, Section II]. Throughout, the sets $\mathcal{X}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{Z}$ are assumed to be finite. The $\left(\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Z}, p_{Y_1,Y_2,Z|X}\right)$ DM-WTBC is shown in Fig. 2(a). The message pair (M_1, M_2) is uniformly distributed over $\left[1\colon 2^{nR_1}\right] \times \left[1\colon 2^{nR_2}\right]$, and a WTBC code is defined as follows.

Definition 1 (WTBC Code) *An (n, R_1, R_2) WTBC code c_n with a product message set $\mathcal{M}_1^{(n)} \times \mathcal{M}_2^{(n)}$, where $\mathcal{M}_j^{(n)} \triangleq [1 : 2^{nR_j}]$, for $j = 1, 2$, is a triple of maps $\left(f_n, \phi_1^{(n)}, \phi_2^{(n)}\right)$ 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 decoder at Receiver j, for $j = 1, 2$.*

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

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

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 this unification by adopting the achievability definition from [7], [10], [11]. This definition merges the reliability and security requirements into a single demand on $P^{(c_n)}$, phrased in terms of TV.

Definition 2 (WTBC Achievability) *A pair $(R_1, R_2) \in \mathbb{R}_+^2$ is called achievable if there exists a $\gamma > 0$, a 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 all sufficiently large n*

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

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

Remark 1 (Equivalence to Standard Definitions) *The above definition of achievability is equivalent to the more standard definition from [12]. Therein, achievability was defined in terms of a vanishing average error probability and the* effective secrecy *metric that requires*

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

is made arbitrarily small. See [13, Section III-B] for details.

Remark 2 (Target i.i.d. Distribution) *The structure of the 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)-(5) *cannot always be a priori determined 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 direct*

proof.[1] *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 first describe the analogy principles for the base case of the classic wiretap and GP channels. Extensions to multiuser (namely, broadcasting) scenarios are given afterwards.

A. The Base Case - A Unified Perspective

Consider the classic WTC and GPC. 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) wiretap code $c_n=(f_n,\phi_n)$ 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 $g_n:\mathcal{M}_n\times\mathcal{Z}^n\to\mathcal{P}(\mathcal{X}^n)$ is a stochastic encoder and $\psi_n:\mathcal{Y}^n\to\mathcal{M}_n$ is the decoder, is

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

With respect to Definition 2, if R is an achievable rate for the WTC, then there exist a $q_Z\in\mathcal{P}(\mathcal{Z})$ and a sequence of (n,R) wiretap codes $\{c_n\}_{n\in\mathbb{N}}$, with

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

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

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

See [13, Section IV-A-1] for details.

As seen from (8)-(9), 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 ($\left\{\tilde{P}^{(c_n)}\right\}_{n\in\mathbb{N}}$ and $\left\{\tilde{Q}^{(b_n)}\right\}_{n\in\mathbb{N}}$ 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.

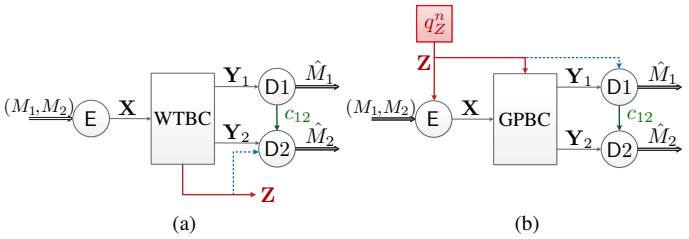

Fig. 3: (a) A PD-WTBC with cooperative receivers, where Receiver 1 has access to the eavesdropper's observation $\mathbf{Z}$; (b) The analogous PD-GPBC with cooperative receivers, where Receiver 1 has access to the state sequence $\mathbf{Z}$.

B. Analogy Between Multiuser Setups

Consider a WTBC $\left(\mathcal{X},\mathcal{Y}_1,\mathcal{Y}_2,\mathcal{Z},p_{Y_1,Y_2,Z|X}\right)$. An analogous GPBC is constructed in three steps (shown in Fig. 2):

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

The produced $\left(\mathcal{Z},\mathcal{X},\mathcal{Y}_1,\mathcal{Y}_2,q_Z,p_{Y_1,Y_2|X,Z}\right)$ GPBC inherits the properties the WTBC possesses (e.g., 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 output signals of the GPBC. A GPBC receiver that observes the state sequence translates to a WTBC receiver that observes the eavesdropper's channel output. The analogy also accounts for WTBC settings with cooperative components. Namely, if the WTBC's receivers are connected by, e.g., a bit-pipe, then the same applies for the receivers of the analogous GPBC. Fig. 3(a) shows a PD-WTBC with cooperative receivers and where Receiver 1 also observes the eavesdropper's output; the analogous PD-GPBC with an informed Receiver 1 and cooperative receivers is shown in Fig. 3(b).

As for the base case, the admissible 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.

Capacity results for GPBCs are available for numerous cases [3]–[5]. The analogy enables leveraging these results to study corresponding WTBCs. This is done by relating the performance of two analogous models as follows.

Proposition 1 (Good Wiretap Codes and Good GP Codes) *Consider a $\left(\mathcal{X},\mathcal{Y}_1,\mathcal{Y}_2,\mathcal{Z},p_{Y_1,Y_2,Z|X}\right)$ 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) WTBC codes $\{c_n\}_{n\in\mathbb{N}}$, where $c_n=\left(f_n,\phi^{(n)}_1,\phi^{(n)}_2\right)$, for each $n\in\mathbb{N}$.*

[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 $p_L<p_B$, 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 $h_b(p_E)-h_b(p_L)$ being $\left(\mathsf{Ber}\left(\frac{1}{2}\right)\right)^n$.

[2] We adhere to the standard definitions for GPCs, see, e.g., [14, Setion 7.6].

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

For each $n \in \mathbb{N}$, set $g_n \triangleq P^{(c_n)}_{\mathbf{X}|\mathbf{Z},M_{[1:2]}}$ and $\psi^{(n)}_j \triangleq \phi^{(n)}_j$, 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) $b_n \triangleq \left(g_n, \psi^{(n)}_1, \psi^{(n)}_2\right)$ is an (n, R_1, R_2)-code for the $\left(\mathcal{Z}, \mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, q_Z, p_{Y_1,Y_2|X,Z}\right)$ 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 $\left|\left|P^{(c_n)} - Q^{(b_n)}\right|\right|_{\mathsf{TV}} \leq e^{-n\gamma}$, for any n large enough.*

3) The codes $\{b_n\}_{n\in\mathbb{N}}$ attain $\mathsf{P_e}(b_n) \xrightarrow[n\to\infty]{} 0$; consequently, (R_1, R_2) is achievable 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 $\mathcal{M}_{12} \triangleq \mathcal{M}^{(n)}_1 \times \mathcal{M}^{(n)}_2$. The first claim is straightforward: for each $n \in \mathbb{N}$, $P^{(c_n)}_{\mathbf{X}|\mathbf{Z},M_{12}}$ and $\psi^{(n)}_j$, for $j = 1, 2$, are valid (stochastic) encoder and decoders for the GPBC. For (2), fix $n \in \mathbb{N}$, and first observe

$$P^{(c_n)}_{M_{12},\mathbf{X},\mathbf{Y}_{[1:2]},\mathbf{Z},\hat{M}_{12}} = P^{(c_n)}_{M_{12},\mathbf{Z}} \cdot Q^{(b_n)}_{\mathbf{X},\mathbf{Y}_{[1:2]},\hat{M}_{12}|M_{12},\mathbf{Z}}, \quad (10)$$

which follows from the structure of $P^{(c_n)}$ and $Q^{(b_n)}$ (see (3) and (7)) and because $b_n = \left(g_n, \psi^{(n)}_1, \psi^{(n)}_2\right)$. Recalling that $Q^{(b_n)}_{\mathbf{Z},M_{12}} = q^n_Z \cdot p^{(U)}_{\mathcal{M}_{12}}$, we have

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

Claim (3) follows by upper bounding $\mathsf{P_e}(b_n)$ as

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

where (a) is because $||p-q||_{\mathsf{TV}} = \sum_{x:\ p(x)>q(x)}\left[p(x)-q(x)\right]$ and since $m_{12} \neq \hat{m}_{12}$ if and only if $Q^{(c_n)}(m_{12},\hat{m}_{12}) \geq p^{(U)}_{\mathcal{M}_{12}}(m_{12})\mathbb{1}_{\left\{\hat{m}_{12}=m_{12}\right\}}$; (b) is the triangle inequality; (c) uses Property (3-a) from [13, Lemma 1]. Finally, the RHS above vanishes to 0 as $n \to \infty$ by (11) and our hypothesis. ∎

IV. Secrecy-Capacity Results

A. Semi-Deterministic WTBCs

We characterize the secrecy-capacity region of the SD-WTBC $p_{Y_1,Y_2,Z|X} = \mathbb{1}_{\{Y_1=y_1(X)\}}p_{Y_2,Z|X}$, where $y_1 : \mathcal{X} \times \mathcal{Z} \to \mathcal{Y}_1$ and $p_{Y_2,Z|X} : \mathcal{X} \to \mathcal{P}(\mathcal{Y}_2 \times \mathcal{Z})$.

Theorem 1 (SD-WTBC Secrecy-Capacity) *The secrecy-capacity region of the SD-WTBC is the union of rate pairs $(R_1, R_2) \in \mathbb{R}^2_+$ satisfying:*

$$R_1 \leq H(Y_1|Z), \quad (12a)$$
$$R_2 \leq I(U;Y_2) - I(U;Z), \quad (12b)$$
$$R_1 + R_2 \leq H(Y_1|Z) + I(U;Y_2) - I(U;Y_1,Z) \quad (12c)$$

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 U to take values in a set $\mathcal{U}$, with $|\mathcal{U}| \leq |\mathcal{X}| + 1$.

The direct proof of Theorem 1 uses the inner bound from [7, Theorem 3], where the performance criterion [7] corresponds to the definition of achievability used herein (Definition 2). Setting $Q = U_0 = 0$, $U_1 = Y_1$ and $U_2 = U$ reduces the rate bounds from [7, Theorem 3] to those from (12). The analogy-based converse proof is given in Section V.

Remark 3 *Until now, the secrecy-capacity region was known only under the assumption that the stochastic legitimate channel $p_{Y_2|X}$ is less noisy than the eavesdropper's channel $p_{Z|X}$ [8, Theorem 5]. Our analogy-based converse proof obviates this assumption. It is important to note that the analogy does rely on the achievability notion from Definition 2, and in particular on the asymptotic i.i.d. requirement (stealth) from $\mathbf{Z}$. With this definition of achievability one could furnish a converse proof for the SD-WTBC that directly arrives at the expressions from* (12) *without using the analogy. However, doing so will be nothing but reproducing the analogy-based proof without explicitly terming it that way. We are unaware of an alternative proof method for converse part of Theorem 1.*

B. Physically-Degraded WTBCs

To stress the versatility of the analogy framework, we also present the secrecy-capacity region of the PD-WTBC with cooperative receivers from Fig. 3(a). Formally, we consider a PD-WTBC $p_{Y_1,Z|X}p_{Y_2|Y_1}$, where $Y_2 \multimap Y_1 \multimap (X, Z)$ forms a Markov chain, with an informed receiver, i.e., when Receiver 1 observes the pair $(\mathbf{Y}_1, \mathbf{Z})$, and where a unidirectional noiseless link of capacity $c_{12} < \infty$ extends from (the informed) Receiver 1 to (the uninformed) Receiver 2. Codes and achievability for this channel are defined in accordance with Definitions 1 and 2.

Theorem 2 (PD-WTBC Secrecy-Capacity) *The secrecy-capacity region of the PD-WTBC with cooperative receivers and an informed Receiver 1 is the union of rate pairs $(R_1, R_2) \in \mathbb{R}^2_+$ satisfying:*

$$R_1 \leq I_p(X;Y_1|U,Z) \quad (13a)$$
$$R_2 \leq I_p(U;Y_2) - I_p(U;Z) + c_{12} \quad (13b)$$
$$R_1 + R_2 \leq I_p(X;Y_1|Z) \quad (13c)$$

where the union is over all $p_{U,X} \in \mathcal{P}(\mathcal{U}\times\mathcal{X})$, each inducing a joint distribution $p \triangleq p_{U,X}p_{Y_1,Z|X}p_{Y_2|Y_1}$. Furthermore, one may restrict U to take values in a set $\mathcal{U}$, with $|\mathcal{U}| \leq |\mathcal{X}| + 1$.

Due to space limitation we omit the proof of Theorem 2 and refer the reader to [13, Section VI-B and Remark 7].

V. Analogy-Based Converse Proof of Theorem 1

Let $(R_1, R_2) \in \mathbb{R}_+^2$ be achievable rate for the SD-WTBC and $\{c_n\}_{n\in\mathbb{N}}$ be the corresponding sequence of (n, R_1, R_2) WTBC 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 Items (2) and (3) from Proposition 1 holds. Furthermore, since the WTBC is SD with $p_{Y_1,Y_2,Z|X} = \mathbb{1}_{\{Y_1=y_1(X)\}} p_{Y_2,Z|X}$, the obtained GPBC is also SD. Namely, the GPBC's transition probability factors 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 [5].

The converse proof of [5, 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

$$R_1 \le \frac{1}{n}\sum_{i=1}^{n} H_Q(Y_{1,i}|Z_i) + \epsilon_n \tag{14a}$$

$$R_2 \le \frac{1}{n}\sum_{i=1}^{n} \Big[I_Q\big(M_2, Y_2^{i-1}, Z_{i+1}^n; Y_{2,i}\big) - I_Q\big(M_2, Y_2^{i-1}, Z_{i+1}^n; Z_i\big)\Big] + \epsilon_n \tag{14b}$$

$$R_1 + R_2 \le \frac{1}{n}\sum_{i=1}^{n} \Big[I_Q\big(M_2, Y_2^{i-1}, Z_{i+1}^n, Y_{1,i+1}^n; Y_{2,i}\big) + H_Q(Y_{1,i}|Z_i) - I_Q\big(M_2, Y_2^{i-1}, Z_{i+1}^n, Y_{1,i+1}^n; Z_i, Y_{2,i}\big)\Big] + \epsilon_n, \tag{14c}$$

where the subscript Q indicates that the underlying distribution is $Q^{(b_n)}$ and $\epsilon_n \triangleq \frac{2}{n} + \mathsf{P_e}(b_n)\sum_{j=1,2} R_j$.

Since the TV between two distribution upper bounds the TV between any of their marginals [13, Property (3-a), Lemma 1]),

$$\Big|\Big| P^{(c_n)}_{M_2,Y^i,Z_i^n} - Q^{(b_n)}_{M_2,Y^i,Z_i^n} \Big|\Big|_{\mathsf{TV}} \le e^{-n\gamma} \tag{15}$$

for large enough n, uniformly in $i \in [1:n]$. Now, over finite probability spaces, an exponentially decaying TV dominates the difference between two corresponding mutual information terms ([13, Lemma 3]). Combining this observation with (15), we may replace the information measures from (14), which are taken with respect to $Q^{(b_n)}$, with the same terms, but with an underlying distribution $P^{(c_n)}$ plus a vanishing term. Namely, there exists a $\delta > 0$, such that for n large enough

$$R_1 \le \frac{1}{n}\sum_{i=1}^{n} H_P(Y_{1,i}|Z_i) + \epsilon_n + e^{-n\delta} \tag{16a}$$

$$R_2 \le \frac{1}{n}\sum_{i=1}^{n} \Big[I_P(V_i; Y_{2,i}) - I_P(V_i; Z_i)\Big] + \epsilon_n + 2e^{-n\delta} \tag{16b}$$

$$R_1 + R_2 \le \frac{1}{n}\sum_{i=1}^{n} \Big[H_P(Y_{1,i}|Z_i) + I_P(V_i, T_i; Y_{2,i}) - I_P(V_i, T_i; Y_{1,i}, Z_i)\Big] + \epsilon_n + 3e^{-n\delta} \tag{16c}$$

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 (16), we see that the WTBC's secrecy-capacity region is contained in the convex closure of the union of rate pairs (R_1, R_2) satisfying:

$$R_1 \le H_p(Y_1|Z) \tag{17a}$$

$$R_2 \le I_p(V; Y_2) - I_p(V; Z) \tag{17b}$$

$$R_1 + R_2 \le H_p(Y_1|Z) + I_p(V, T; Y_2) - I_p(V, T; Y_1, Z) \tag{17c}$$

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):

$$I_p(V; Y_2) - I_p(V; Z) \le I_p(U; Y_2) - I_p(U; Z) \tag{18a}$$

$$H_p(Y_1|Z) + I_p(V, T; Y_2) - I_p(V, T; Y_1, Z) \le H_p(Y_1|Z) + I_p(U; Y_2) - I_p(U; Y_1, Z). \tag{18b}$$

This is established by closely following the arguments from the end of the converse proof of the analogous SD-GPBC [5, Section III]. Namely, setting $U = V$ if p is such that $I_p(T; Y_2|V) - I_p(T; Y_1, Z|V) \le 0$, and $U = (V, T)$ if $I_p(T; Y_2|V) - I_p(T; Z|V) \ge 0$ suffices. Noting that every distribution p must satisfy at least one of these information inequalities concludes the proof.

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] 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.

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

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

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

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

[12] 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.

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

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