Network Equivalence in the Presence of an Eavesdropper

Theodoros K. Dikaliotis[1], Hongyi Yao[2] Tracey Ho[1], Michelle Effros[1], Joerg Kliewer[3]
[1,2]California Institute of Technology [3]New Mexico State University
[1]{tdikal, tho, effros}@caltech.edu [2]yaohongyi03@gmail.com [3]jkliewer@nmsu.edu

***Abstract*—We consider networks of noisy degraded wiretap channels in the presence of an eavesdropper. For the case where the eavesdropper can wiretap at most one channel at a time, we show that the secrecy capacity region, for a broad class of channels and any given network topology and communication demands, is equivalent to that of a corresponding network where each noisy wiretap channel is replaced by a noiseless wiretap channel. Thus in this case there is a separation between wiretap channel coding on each channel and secure network coding on the resulting noiseless network. We show with an example that such separation does not hold when the eavesdropper can access multiple channels at the same time, for which case we provide upper and lower bounding noiseless networks.**

I. Introduction

Information theoretically secure (secret) communication in the presence of an eavesdropper has been studied under various models. One body of literature studies the wiretap channel, introduced by Wyner [1], where the intended receiver and the eavesdropper observe outputs of a physical layer channel. Another body of literature investigates the secure capacity of networks of noise-free links. Under this model, introduced by Cai and Yeung in [2], an eavesdropper perfectly observes all information traversing a restricted but unknown subset of links. The first paper on the secure capacity of a network of noisy channels is [3], which finds upper and lower bounds on the unicast capacity of a network of independent broadcast erasure channels when the output observed by the eavesdropper equals that of the intended receiver on all wiretapped channels.

Our work considers the problem of secure communication over a network of independent wiretap channels which are physically degraded and "simultaneously maximizable" (see Definition 1 in Section II), and broadens consideration to general capacity regions specifying vectors of simultaneously achievable rates. We require asymptotically negligible decoding error probability and information leakage to the eavesdropper, as defined formally in Section II. In the case where the eavesdropper has access to only one link, the identity of which is unknown to the code designer, we show that the secrecy capacity region is identical to that of a corresponding noiseless network, for any network topology and connection types. Thus in this case capacity can be achieved by separate design of wiretap channel codes converting each channel to a pair of public and confidential noiseless links, and a secure network code on the resulting noiseless network. We show with an example that such separation does not hold when the eavesdropper can access multiple channels at the same time, for which case we provide upper and lower bounding noiseless networks. Our results bring together and generalize the wiretap channel and secure network coding literature, allowing application of existing results on secure network coding capacity to characterize or bound the secure capacity of networks of such wiretap channels. Our work builds on and generalizes the techniques developed by Koetter, Effros, and Medard in [4], [5], which show similar capacity bounds in the absence of secrecy constraints. We provide below outlines of all proofs, details of which are given in the full version of this paper [6].

II. Model and Preliminaries

Consider a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V} \times \mathbb{N}$ is a set of directed edges between pairs of nodes in the network. Edge (i, j, k) represents the k^{th} wiretap channel through which node i communicates to node j and through which an eavesdropper may or may not be listening. The total number of nodes in the network is m. The channel inputs and outputs for node i at time t are given by

$$X_t^{(i)} = \left(X_t^{(e)} : e \in \mathcal{E}_{\text{out}}(i)\right) \quad \text{and} \quad Y_t^{(i)} = \left(Y_t^{(e)} : \mathcal{E}_{\text{in}}(i)\right)$$

where $X_t^{(e)}$ and $Y_t^{(e)}$ denote the input to and the output from edge e respectively, and $\mathcal{X}^{(e)}$ and $\mathcal{Y}^{(e)}$ denote their alphabets, which may be discrete or continuous. We define

$$\mathcal{E}_{\text{in}}(i) = \{(u, v, w) \in \mathcal{E} : v = i\}$$
$$\mathcal{E}_{\text{out}}(i) = \{(u, v, w) \in \mathcal{E} : u = i\}$$

$$\mathcal{X}^{(i)} = \prod_{e \in \mathcal{E}_{\text{out}}(i)} \mathcal{X}^{(e)} \quad \text{and} \quad \mathcal{Y}^{(i)} = \prod_{e \in \mathcal{E}_{\text{in}}(i)} \mathcal{Y}^{(e)}.$$

Let $\mathcal{P}(\mathcal{E})$ denote the power set of the set of all edges. In a secure communication problem, an adversarial set $A \subseteq \mathcal{P}(\mathcal{E})$ is specified. Each set $E \in A$ describes a subset of channels over which an eavesdropper may be listening. The code is designed to be secure against eavesdropping on the set of channels E for every $E \in A$. When the eavesdropper listens to edge $e = (i, j, k)$, the eavesdropper receives, at each time t, a degraded version $Z_t^{(e)}$ of the channel output $Y_t^{(e)}$ observed by the intended recipient, which is the output node j of edge $e = (i, j, k)$. If the eavesdropper has eavesdropping set $E \in A$, then at time t it receives the set of random variables $\left(Z_t^{(e)} : e \in E\right)$, which we compactly write as $Z_t^{(E)}$.

The vector $\left(Z_1^{(E)},\ldots,Z_n^{(E)}\right)$ of observations from all edges $e\in E$ over time steps $t\in\{1,\ldots,n\}$ is denoted by $\left(Z^{(E)}\right)^n$. Similarly we define $\left(X^{(E)}\right)^n = \left(X_1^{(E)},\ldots,X_n^{(E)}\right)$ and $\left(Y^{(E)}\right)^n = \left(Y_1^{(E)},\ldots,Y_n^{(E)}\right)$ where $X_t^{(E)} = \left(X_t^{(e)} : e\in E\right)$ and $Y_t^{(E)} = \left(Y_t^{(e)} : e\in E\right)$.

For each $e\in\mathcal{E}$, channel e is a memoryless, time-invariant, physically degraded wiretap channel described by a conditional distribution

$$p(y^{(e)},z^{(e)}|x^{(e)}) = p(y^{(e)}|x^{(e)})\cdot p(z^{(e)}|y^{(e)}).$$

All wiretap channels are independent by assumption, giving

$$p\big(y^{(\mathcal{E})},z^{(\mathcal{E})}|x^{(\mathcal{E})}\big) = \prod_{e\in\mathcal{E}} p\big(y^{(e)},z^{(e)}|x^{(e)}\big)$$
$$= \prod_{e\in\mathcal{E}} p\big(y^{(e)}|x^{(e)}\big)p\big(z^{(e)}|y^{(e)}\big).$$

We further restrict our attention to channels that are "simultaneously maximizable," as defined below.

Definition 1: Wiretap channel e is called simultaneously maximizable if

$$\arg\left[\max_{p(x)} I(X^{(e)};Y^{(e)})\right] = \arg\left[\max_{p(x)} I(X^{(e)};Z^{(e)})\right]$$

and

$$\max_{p(x^{(e)})}\Big[I(X^{(e)};Y^{(e)}) - I(X^{(e)};Z^{(e)})\Big]$$
$$= \max_{p(x^{(e)})} I(X^{(e)};Y^{(e)}) - \max_{p(x^{(e)})} I(X^{(e)};Z^{(e)}).$$

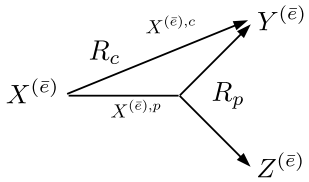
The about maximization is subject to any constraints on the channel input (*e.g.,* an input power constraint for a Gaussian channel) associated with the communication network of interest. Examples of simultaneously maximizable wiretap channels include weakly symmetric channels and Gaussian channels [7], [8]. Intuitively, restriction to simultaneously maximizable channels simplifies our analysis since the same input distribution maximizes both the total and secure capacity.

A code of blocklength n operates over n time steps to reliably communicate message

$$W^{(i\to\mathcal{B})}\in\mathcal{W}^{(i\to\mathcal{B})}\stackrel{\text{def}}{=}\{1,\ldots,2^{nR^{(i\to\mathcal{B})}}\}$$

from each source node $i\in\mathcal{V}$ to each nonempty set $\mathcal{B}\subseteq\mathcal{V}\backslash\{i\}$ of sink nodes in a manner that guarantees information theoretic security in the presence of any eavesdropper $E\in A$. This constitutes a unicast connection if $|\mathcal{B}|=1$ and a multicast connection if $|\mathcal{B}|>1$. Constant $R^{(i\to\mathcal{B})}$ is called the transmission rate from source i to sink set $\mathcal{B}$. The vector of all rates $R^{(i\to\mathcal{B})}$ is denoted by $R=\big(R^{(i\to\mathcal{B})} : i\in\mathcal{V},\mathcal{B}\in\mathcal{B}^{(i)}\big)$, where set $\mathcal{B}^{(i)}=\{\mathcal{B}:\mathcal{B}\subseteq\mathcal{V}\backslash\{i\},\mathcal{B}\neq\emptyset\}$ is the set of nonempty receiver sets to which node i may wish to transmit. Similarly, the vector of all messages is denoted by $W=\big(W^{(i\to\mathcal{B})} : i\in\mathcal{V},\mathcal{B}\in\mathcal{B}^{(i)}\big)$.

Each node $i\in\mathcal{V}$ also has access to a random variable $T^{(i)}\in\mathcal{T}^{(i)}\stackrel{\text{def}}{=}\{1,\ldots,2^{nC^{(i)}}\}$ for use in randomized coding



$X^{(\bar{e})} = (X^{(\bar{e}),c}, X^{(\bar{e}),p})$
$Y^{(\bar{e})} = (Y^{(\bar{e}),c}, Y^{(\bar{e}),p})$
$Y^{(\bar{e}),c} = X^{(\bar{e}),c}$
$Y^{(\bar{e}),p} = X^{(\bar{e}),p}$
$Z^{(\bar{e})} = Y^{(\bar{e}),p}$

Fig. 1. A noiseless degraded broadcast channel with confidential rate R_c and public rate R_p.

for secrecy, where

$$C^{(i)} = \sum_{e\in\mathcal{E}_{\text{out}}(i)} \max_{p(x^{(e)})} I\big(X^{(e)};Y^{(e)}\big) \tag{1}$$

is the sum of the outgoing channel capacities from node i. Each $T^{(i)}$ is uniformly distributed on its alphabet and independent of all messages and channel noise.

Definition 2: Let a network

$$\mathcal{N}\stackrel{\text{def}}{=}\Big(\prod_{e\in\mathcal{E}}\mathcal{X}^{(e)}, \prod_{e\in\mathcal{E}}\Big(p(y^{(e)}|x^{(e)})p\Big(z^{(e)}|y^{(e)}\Big)\Big),$$
$$\prod_{e\in\mathcal{E}}\Big(\mathcal{Y}^{(e)}\times\mathcal{Z}^{(e)}\Big)\Big)$$

be given corresponding to a graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$. A blocklength n solution $\mathcal{S}(\mathcal{N})$ is defined as a set of encoding functions

$$X_t^{(i)} : \Big(\mathcal{Y}^{(i)}\Big)^{t-1}\times\prod_{\mathcal{B}\in\mathcal{B}^{(i)}}\mathcal{W}^{(i\to\mathcal{B})}\times\mathcal{T}^{(i)}\longrightarrow\mathcal{X}^{(i)}$$

mapping $\Big(Y_1^{(i)},\ldots,Y_{t-1}^{(i)},\big(W^{(i\to\mathcal{B})}:\mathcal{B}\in\mathcal{B}^{(i)}\big),T^{(i)}\Big)$ to $X_t^{(i)}$ for each $i\in\mathcal{V}$ and $t\in\{1,\ldots,n\}$, and a set of decoding functions

$$\breve{W}^{(j\to\mathcal{K},i)} : \Big(\mathcal{Y}^{(i)}\Big)^{n}\times\prod_{\mathcal{B}\in\mathcal{B}^{(i)}}\mathcal{W}^{(i\to\mathcal{B})}\times\mathcal{T}^{(i)}\longrightarrow\mathcal{W}^{(j\to\mathcal{K})}$$

mapping $\Big(Y_1^{(i)},\ldots,Y_n^{(i)},\big(W^{(i\to\mathcal{B})}:\mathcal{B}\in\mathcal{B}^{(i)}\big),T^{(i)}\Big)$ to $\breve{W}^{(j\to\mathcal{K},i)}$ for each $j\in\mathcal{V}$, $\mathcal{K}\in\mathcal{B}^{(j)}$, and $i\in\mathcal{K}$. The solution $\mathcal{S}(\mathcal{N})$ is called a $(\lambda,\varepsilon,A,R)$–solution, denoted $(\lambda,\varepsilon,A,R)$–$\mathcal{S}(\mathcal{N})$, if $\Pr\Big(\breve{W}^{(j\to\mathcal{K},i)}\neq W^{(j\to\mathcal{K})}\Big)<\lambda$ for every $j\in\mathcal{V}$, $\mathcal{K}\in\mathcal{B}^{(j)}$ and $i\in\mathcal{K}$, and $I\left(\left(Z^{E}\right)^{n};W\right)<n\varepsilon$ for every $E\in A$.

Definition 3: The A–secure rate region $\mathcal{R}(\mathcal{N},A)\subseteq\mathbb{R}_+^{m(2^{m-1}-1)}$ of a network $\mathcal{N}$ is the closure of all rate vectors R such that for any $\lambda>0$ and $\varepsilon>0$, a solution $(\lambda,\varepsilon,A,R)$–$\mathcal{S}(\mathcal{N})$ exists.

Given a network $\mathcal{N}$ and a channel $\bar{e}\in\mathcal{E}$, the model $\mathcal{N}_{\bar{e}}(R_c,R_p)$ replaces $\bar{e}$ with noiseless bit pipes as defined below and illustrated in Figure 1.

Definition 4: Given a network

$$\mathcal{N} \stackrel{\text{def}}{=} (\prod_{e\in\mathcal{E}} \mathcal{X}^{(e)}, \prod_{e\in\mathcal{E}} \left(p\left(y^{(e)}|x^{(e)}\right) p\left(z^{(e)}|y^{(e)}\right)\right), \prod_{e\in\mathcal{E}} \left(\mathcal{Y}^{(e)} \times \mathcal{Z}^{(e)}\right))$$

and some $\bar{e} \in \mathcal{E}$, the model $\mathcal{N}_{\bar{e}}(R_c, R_p)$ replaces the degraded wiretap channel

$$\mathcal{C}_{\bar{e}} = \left(\mathcal{X}^{(\bar{e})}, p(y^{(\bar{e})}|x^{(\bar{e})})p(z^{(\bar{e})}|y^{(\bar{e})}), \mathcal{Y}^{(\bar{e})} \times \mathcal{Z}^{(\bar{e})}\right)$$

with the noiseless degraded wiretap channel

$$\mathcal{C}(R_c, R_p) = (\{0,1\}^{R_c+R_p}, \delta\big(y^{(\bar{e})} - (x^{(\bar{e}),c}, x^{(\bar{e}),p})\big)$$
$$\delta\big(z^{(\bar{e})} - y^{(\bar{e}),p}\big), \{0,1\}^{R_c+R_p} \times \{0,1\}^{R_p})$$

that delivers the rate-R_c confidential portion $x^{(\bar{e}),c}$ of channel input $x^{(\bar{e})} = (x^{(\bar{e}),c}, x^{(\bar{e}),p})$ to the intended receiver and the rate-R_p public portion $x^{(\bar{e}),p}$ of that input to both the intended receiver and eavesdropper. The resulting network is given by

$$\mathcal{N}_{\bar{e}}(R_c, R_p) \stackrel{\text{def}}{=} (\{0,1\}^{R_c+R_p} \times \prod_{e\in\mathcal{E}\setminus\{\bar{e}\}} \mathcal{X}^{(e)},$$
$$\delta(y^{(\bar{e})} - (x^{(\bar{e}),c}, x^{(\bar{e}),p}))\delta(z^{(\bar{e})} - y^{(\bar{e}),p})$$
$$\cdot \prod_{e\in\mathcal{E}\setminus\{\bar{e}\}} (p(y^{(e)}|x^{(e)}).p(z^{(e)}|y^{(e)})),$$
$$\{0,1\}^{R_c+R_p} \times \{0,1\}^{R_p} \times \prod_{e\in\mathcal{E}\setminus\{\bar{e}\}} (\mathcal{Y}^{(e)} \times \mathcal{Z}^{(e)})).$$

As in [4], [5], we allow non-integer values of R_c and R_p to denote noiseless bit pipes that require multiple channel uses to deliver some integer number of bits.

Many of the subsequent proofs use the notion of a "stacked network" introduced in [4], [5], extended here by adding an eavesdropper. Informally, the N-fold stacked network $\underline{\mathcal{N}}$ contains N copies of network $\mathcal{N}$. The N copies of each node $i \in \mathcal{V}$ use the outgoing messages and channel outputs from all N layers of the network to form the channel inputs in each layer of the stack. Likewise, each node uses the channel outputs and messages from all layers in the stack in building its message reconstructions. An eavesdropper $E \in A$ overhears all copies of channel e for each $e \in E$.

As defined formally below following [4], [5], a solution for N-fold stacked network $\underline{\mathcal{N}}$ must securely and reliably transmit, for each $i \in \mathcal{V}$ and $\mathcal{B} \in \mathcal{B}^{(i)}$, N independent messages $\underline{W}^{(i\to\mathcal{B})}(1), \ldots, \underline{W}^{(i\to\mathcal{B})}(N)$ from node i to all the receivers in set $\mathcal{B}$. We underline the variable names from $\mathcal{N}$ to denote variables for the stacked network $\underline{\mathcal{N}}$. Therefore $\underline{W}^{(i\to\mathcal{B})} \in \underline{\mathcal{W}}^{(i\to\mathcal{B})} \stackrel{\text{def}}{=} \left(\mathcal{W}^{(i\to\mathcal{B})}\right)^N$, $\underline{T}^{(i)} \in \underline{\mathcal{T}}^{(i)} \stackrel{\text{def}}{=} \left(\mathcal{T}^{(i)}\right)^N$, $\underline{X}_t^{(i)} \in \underline{\mathcal{X}}^{(i)} \stackrel{\text{def}}{=} \left(\mathcal{X}^{(i)}\right)^N$, $\underline{Y}_t^{(i)} \in \underline{\mathcal{Y}}^{(i)} \stackrel{\text{def}}{=} \left(\mathcal{Y}^{(i)}\right)^N$, and $\underline{Z}_t^{(e)} \in \underline{\mathcal{Z}}^{(e)} \stackrel{\text{def}}{=} \left(\mathcal{Z}^{(e)}\right)^N$ denote N-dimensional vectors of messages, channel inputs, channel outputs, and eavesdropper outputs, respectively, in network $\mathcal{N}$. The variables in the ℓ^{th} layer of the stack are denoted by an argument ℓ. Finally, we define the rate $R^{(i\to\mathcal{B})}$ for a stacked network to be $(\log_2 |\underline{\mathcal{W}}^{(i\to\mathcal{B})}|)/(nN)$ since any solution of blocklength n for N-fold stacked network $\underline{\mathcal{N}}$ can be operated as a rate-R solution of blocklength nN for network $\mathcal{N}$ under this definition [4, Theorem 1]. A similar argument, given in Theorem 1 below, justifies the security constraint imposed below. Definitions 5-7 are analogous to Definitions 4-6 in [4].

Definition 5: Let a network

$$\mathcal{N} \stackrel{\text{def}}{=} (\prod_{e\in\mathcal{E}} \mathcal{X}^{(e)}, \prod_{e\in\mathcal{E}} \left(p_e\left(y^{(e)}|x^{(e)}\right) p_e\left(z^{(e)}|y^{(e)}\right)\right), \prod_{e\in\mathcal{E}} \left(\mathcal{Y}^{(e)} \times \mathcal{Z}^{(e)}\right))$$

be given corresponding to a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, and let an eavesdropper set $A \subseteq P(\mathcal{E})$ be defined on network $\mathcal{N}$. Let $\underline{\mathcal{N}}$ be the N-fold stacked network for $\mathcal{N}$. A blocklength-n solution $\mathcal{S}(\underline{\mathcal{N}})$ to this network is defined as a set of encoding functions

$$\underline{X}_t^{(i)} : \left(\underline{\mathcal{Y}}^{(i)}\right)^{t-1} \times \prod_{\mathcal{B}\in\mathcal{B}^{(i)}} \underline{\mathcal{W}}^{(i\to\mathcal{B})} \times \underline{\mathcal{T}}^{(i)} \longrightarrow \underline{\mathcal{X}}^{(i)}$$

mapping $\left(\underline{Y}_1^{(i)}, \ldots, \underline{Y}_{t-1}^{(i)}, (\underline{W}^{(i\to\mathcal{B})} : \mathcal{B} \in \mathcal{B}^{(i)}), \underline{T}^{(i)}\right)$ to $\underline{X}_t^{(i)}$ for each $i \in \mathcal{V}$ and $t \in \{1, \ldots, n\}$, and decoding functions

$$\underline{\breve{W}}^{(j\to\mathcal{K},i)} : \left(\underline{\mathcal{Y}}^{(i)}\right)^n \times \prod_{\mathcal{B}\in\mathcal{B}^{(i)}} \underline{\mathcal{W}}^{(i\to\mathcal{B})} \times \underline{\mathcal{T}}^{(i)} \longrightarrow \underline{\mathcal{W}}^{(j\to\mathcal{K})}$$

mapping $\left(\underline{Y}_1^{(i)}, \ldots, \underline{Y}_n^{(i)}, (\underline{W}^{(i\to\mathcal{B})} : \mathcal{B} \in \mathcal{B}^{(i)}), \underline{T}^{(i)}\right)$ to $\underline{\breve{W}}^{(j\to\mathcal{K},i)}$ for each $j \in \mathcal{V}$, $\mathcal{K} \in \mathcal{B}^{(j)}$, and $i \in \mathcal{K}$. The solution $\mathcal{S}(\underline{\mathcal{N}})$ is called a $(\lambda, \varepsilon, A, R)$–solution for stacked network $\underline{\mathcal{N}}$, denoted $(\lambda, \varepsilon, A, R)$–$\mathcal{S}(\underline{\mathcal{N}})$, if $\left(\log_2 \left|\underline{W}^{(i\to\mathcal{B})}\right|\right) / (nN) = R^{(i\to\mathcal{B})}$, $I\left(\left(\underline{Z}^{(E)}\right)^n; \underline{W}\right) < nN\varepsilon$ for every $E \in A$, and $\Pr\left(\underline{\breve{W}}^{(j\to\mathcal{K},i)} \neq \underline{W}^{(j\to\mathcal{K})}\right) < \lambda$ for the specified encoding and decoding functions.

Definition 6: The A-secure rate region $\mathcal{R}(\underline{\mathcal{N}}, A) \subseteq \mathbb{R}_+^{m(2^{m-1}-1)}$ of stacked network $\underline{\mathcal{N}}$ is the closure of all rate vectors R such that for any $\lambda > 0$ and any $\varepsilon > 0$, a solution $(\lambda, \varepsilon, A, R)$–$\mathcal{S}(\underline{\mathcal{N}})$ exists for sufficiently large N.

Definition 7: Let a network

$$\mathcal{N} \stackrel{\text{def}}{=} (\prod_{e\in\mathcal{E}} \mathcal{X}^{(e)}, \prod_{e\in\mathcal{E}} \left(p_e\left(y^{(e)}|x^{(e)}\right) p_e\left(z^{(e)}|y^{(e)}\right)\right), \prod_{e\in\mathcal{E}} \mathcal{Y}^{(e)} \times \prod_{e\in\mathcal{E}} \mathcal{Z}^{(e)})$$

be given corresponding to a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Fix positive integers n and N as the blocklength and stack size, respectively. For each $i \in \mathcal{V}$ and $\mathcal{B} \in \mathcal{B}^{(i)}$, let $R^{(i\to\mathcal{B})}$ and $\tilde{R}^{(i\to\mathcal{B})}$ be constants with $\tilde{R}^{(i\to\mathcal{B})} \geqslant R^{(i\to\mathcal{B})}$. Define $W^{(i\to\mathcal{B})} = \{1, \ldots, 2^{nR^{(i\to\mathcal{B})}}\}$ and $\tilde{W}^{(i\to\mathcal{B})} = \{1, \ldots, 2^{n\tilde{R}^{(i\to\mathcal{B})}}\}$. Let $\underline{\mathcal{N}}$ be the N-fold stacked network for $\mathcal{N}$. A blocklength-n stacked

solution $\underline{\mathcal{S}}(\underline{\mathcal{N}})$ to this network is defined as a set of mappings

$$\underline{\tilde{W}}^{(i\to\mathcal{B})}:\underline{\mathcal{W}}^{(i\to\mathcal{B})}\to\underline{\tilde{\mathcal{W}}}^{(i\to\mathcal{B})}$$
$$X_t^{(i)}:\left(\mathcal{Y}^{(i)}\right)^{t-1}\times\prod_{\mathcal{B}\in\mathcal{B}^{(i)}}\tilde{\mathcal{W}}^{(i\to\mathcal{B})}\times\mathcal{T}^{(i)}\longrightarrow\mathcal{X}^{(i)}$$
$$\breve{\tilde{W}}^{(j\to\mathcal{K},i)}:\left(\mathcal{Y}^{(i)}\right)^{n}\times\prod_{\mathcal{B}\in\mathcal{B}^{(i)}}\tilde{\mathcal{W}}^{(i\to\mathcal{B})}\times\mathcal{T}^{(i)}\longrightarrow\tilde{\mathcal{W}}^{(j\to\mathcal{K})}$$
$$\underline{\breve{W}}^{(j\to\mathcal{K},i)}:\underline{\tilde{\mathcal{W}}}^{(j\to\mathcal{K})}\to\underline{\mathcal{W}}^{(j\to\mathcal{K})},$$

where the other channel encoder $\underline{\tilde{W}}^{(i\to\mathcal{B})}(\cdot)$ encodes message $\underline{W}^{(i\to\mathcal{B})}$ to $\underline{\tilde{W}}^{(i\to\mathcal{B})}\big(\underline{W}^{(i\to\mathcal{B})}\big)$, encoder $X_t^{(i)}(\cdot)$ independently encodes each dimension $\ell\in\{1,\dots,N\}$ of outgoing messages $\underline{\tilde{W}}^{(i\to\mathcal{B})}$, received channel outputs $\underline{Y}_1^{(i)},\dots,\underline{Y}_{t-1}^{(i)}$, and random keys $\underline{T}^{(i)}$ to channel input

$$X_t^{(i)}(\underline{Y}_1^{(i)}(\ell),\dots,\underline{Y}_{t-1}^{(i)}(\ell),\big(\underline{\tilde{W}}^{(i\to\mathcal{B})}(\ell):\mathcal{B}\in\mathcal{B}^{(i)}\big),\underline{T}^{(i)}(\ell)),$$

node decoder $\breve{\tilde{W}}^{(j\to\mathcal{K},i)}(\cdot)$ independently decodes each dimension of the reconstruction

$$\breve{\tilde{W}}^{(j\to\mathcal{K},i)}(\underline{Y}_1^{(i)}(\ell),\dots,\underline{Y}_n^{(i)}(\ell),\big(\underline{\tilde{W}}^{(i\to\mathcal{B})}(\ell):\mathcal{B}\in\mathcal{B}^{(i)}\big),\underline{T}^{(i)}(\ell))$$

of $\underline{\tilde{W}}^{(j\to\mathcal{K})}$ at node i, and channel decoder $\underline{\breve{W}}^{(j\to\mathcal{K},i)}(\cdot)$ reconstructs message vector $\underline{\breve{W}}^{(j\to\mathcal{K},i)}(\underline{\breve{\tilde{W}}}^{(j\to\mathcal{K},i)})$.

The following theorem extends [4, Theorem 2] from traditional to secure capacity.

Theorem 1: The rate regions $\mathcal{R}(\mathcal{N},A)$ and $\mathcal{R}(\underline{\mathcal{N}},A)$ are identical. Further, for any $R\in\text{int}\big(\mathcal{R}(\mathcal{N},A)\big)$, there exists a sequence of $(2^{-N\delta},\varepsilon,A,R)$–$\underline{\mathcal{S}}(\underline{\mathcal{N}})$ stacked solutions for the stacked network $\underline{\mathcal{N}}$ for some $\delta>0$.

Sketch of the proof: The argument to show $\mathcal{R}(\underline{\mathcal{N}},A)\subseteq\mathcal{R}(\mathcal{N},A)$ follows [4, Theorem 1]: given any $R\in\text{int}(\mathcal{R}(\underline{\mathcal{N}},A))$, a blocklength-$n$ $(\lambda,\varepsilon,A,R)-\mathcal{S}(\underline{\mathcal{N}})$ solution for network $\underline{\mathcal{N}}$ is unraveled across time to achieve a blocklength-nN solution for network $\mathcal{N}$. Since the given code satisfies the causality constraints and precisely implements the operations of $\mathcal{S}(\underline{\mathcal{N}})$, the solution $\mathcal{S}(\mathcal{N})$ achieves the same rate, error probability, and secrecy on $\mathcal{N}$ as the solution $\mathcal{S}(\underline{\mathcal{N}})$ achieves on $\underline{\mathcal{N}}$, which gives the forward result.

The converse follows [4, Theorem 2]. Again, fix $\varepsilon>0$, and for any $R\in\text{int}\big(\mathcal{R}(\mathcal{N},A)\big)$ choose $\tilde{R}\in\text{int}\big(\mathcal{R}(\mathcal{N},A)\big)$ with $\tilde{R}^{(i\to\mathcal{B})}>R^{(i\to\mathcal{B})}$ for all $(i,\mathcal{B})$ with $R^{(i\to\mathcal{B})}>0$. Define $\rho=\min_{i\in\mathcal{V}}\min_{\mathcal{B}\in\mathcal{B}^{(i)}}\big(\tilde{R}^{(i\to\mathcal{B})}-R^{(i\to\mathcal{B})}\big)$ and choose constant $\lambda>0$ satisfying

$$\max_{i\in\mathcal{V}}\max_{\mathcal{B}\in\mathcal{B}^{(i)}}\tilde{R}^{(i\to\mathcal{B})}\lambda+h(\lambda)<\rho.$$

This is possible by choosing λ small enough so that $\lambda<\rho/\big(3\max_{i\in\mathcal{V}}\max_{\mathcal{B}\in\mathcal{B}^{(i)}}\tilde{R}^{(i\to\mathcal{B})}\big)$ and $h(\lambda)<\rho/(3\rho)$. Since $\tilde{R}^{(i\to\mathcal{B})}>R^{(i\to\mathcal{B})}$, there exists a blocklength n such that a $(\lambda,\frac{\varepsilon}{3},A,\tilde{R})$–$\mathcal{S}(\mathcal{N})$ single-layer solution exists. A stacked solution is built using this same $(\lambda,\frac{\varepsilon}{3},A,R)$–$\mathcal{S}(\mathcal{N})$ single-layer solution in each layer and a randomly chosen channel code across the layers of the stack. □

III. Main Results

In Theorem 2, we show that for any network $\mathcal{N}$ of wiretap channels and any edge $\bar{e}\in\mathcal{E}$, replacing channel $\mathcal{C}_{\bar{e}}$ with a noiseless degraded wiretap channel of appropriate capacities R_c and R_p, as shown in Figure 1, yields a network $\mathcal{N}_{\bar{e}}(R_c,R_p)$ (Definition 4) whose secure capacity region contains the secure capacity region of $\mathcal{N}$. Theorem 2 extends [5, Theorem 5] from traditional to secure capacity.

Theorem 2: Consider a network $\mathcal{N}$ and an adversarial set $A\subseteq\mathcal{P}(\mathcal{E})$. $\mathcal{R}(\mathcal{N},A)\subseteq\mathcal{R}(\mathcal{N}_{\bar{e}}(R_c,R_p),A)$ for

$$R_c>\max_{p(x^{(\bar{e})})}I(X^{(\bar{e})};Y^{(\bar{e})})-\max_{p(x^{(\bar{e})})}I(X^{(\bar{e})};Z^{(\bar{e})})$$
$$R_p>\max_{p(x^{(\bar{e})})}I(X^{(\bar{e})};Z^{(\bar{e})}).$$

Sketch of the proof: By Theorem 1 it suffices to prove $\mathcal{R}(\underline{\mathcal{N}},A)\subseteq\mathcal{R}(\underline{\mathcal{N}}_{\bar{e}}(R_c,R_p),A)$. We employ a channel code across layers of the stack to emulate a secure code for network $\underline{\mathcal{N}}$ on network $\underline{\mathcal{N}}_{\bar{e}}(R_c,R_p)$. Typical inputs $\underline{X}_t$ to $\bar{e}$ are mapped to jointly typical outputs from a random codebook. It can be shown that the induced probability distribution p' is close to the probability distribution p of the original secure code for $\underline{\mathcal{N}}$, and that mutual information values under both probability distributions are similar. The bits transmitted over the noiseless channel correspond to the codeword index, and thus reveal a similar amount of information to the wiretapper as its observations of the original noisy channel. □

Theorem 3 shows cases where the upper bound shown in Theorem 2 is tight.

Theorem 3: Consider a network $\mathcal{N}$, an adversarial set $A\subseteq\mathcal{P}(\mathcal{E})$, and a single link $\bar{e}\in\mathcal{E}$. Let

$$R_c=\max_{p(x^{(\bar{e})})}I(X^{(\bar{e})};Y^{(\bar{e})})-\max_{p(x^{(\bar{e})})}I(X^{(\bar{e})};Z^{(\bar{e})})$$
$$R_p=\max_{p(x^{(\bar{e})})}I(X^{(\bar{e})};Z^{(\bar{e})}).$$

If $\bar{e}$ is invulnerable to wiretapping ($\bar{e}\notin E$ for all $E\in A$) or is not simultaneously wiretapped with other links ($\bar{e}\in E$ implies $|E|=1$), then $\mathcal{R}(\mathcal{N},A)=\mathcal{R}(\mathcal{N}_{\bar{e}}(R_c,R_p),A)$.

Sketch of the proof: We outline the proof for the case where $\bar{e}$ is wiretapped but not simultaneously with other links; the case where it is invulnerable to wiretapping is a simpler version.

We first show that $\mathcal{R}(\mathcal{N}_{\bar{e}}(R_c-\epsilon,R_p-\epsilon),A)\subseteq\mathcal{R}(\mathcal{N},A)$ for any $\epsilon>0$, by starting with a secure code of rate R for network $\mathcal{N}_{\bar{e}}(R_c,R_p)$ and constructing a corresponding secure code for network $\mathcal{N}$. Denote by C_t and P_t the transmissions across the confidential and public links, respectively, of edge $\bar{e}\in\mathcal{E}$ at time t. Let $C^n=(C_1,\dots,C_n)$, $P^n=(P_1,\dots,P_n)$ and denote by C_j^i and P_j^i for any $j<i$ the vectors $C_j^i=(C_j,C_{j+1},\dots,C_i)$ and $P_j^i=(P_j,P_{j+1},\dots,P_i)$. We define networks **I** and **II** shown in Figure 2 that are identical to networks $\mathcal{N}_{\bar{e}}(R_c,R_p)$ and $\mathcal{N}$ respectively with the addition of an auxiliary receiver that observes the wiretap output of $\bar{e}$, messages W and a noiseless side channel of capacity $C_{\bar{e}}$ (defined below) from a "super-source" that has access to (W,C^n,P^n). In network **I** (**II**) the auxiliary receiver is

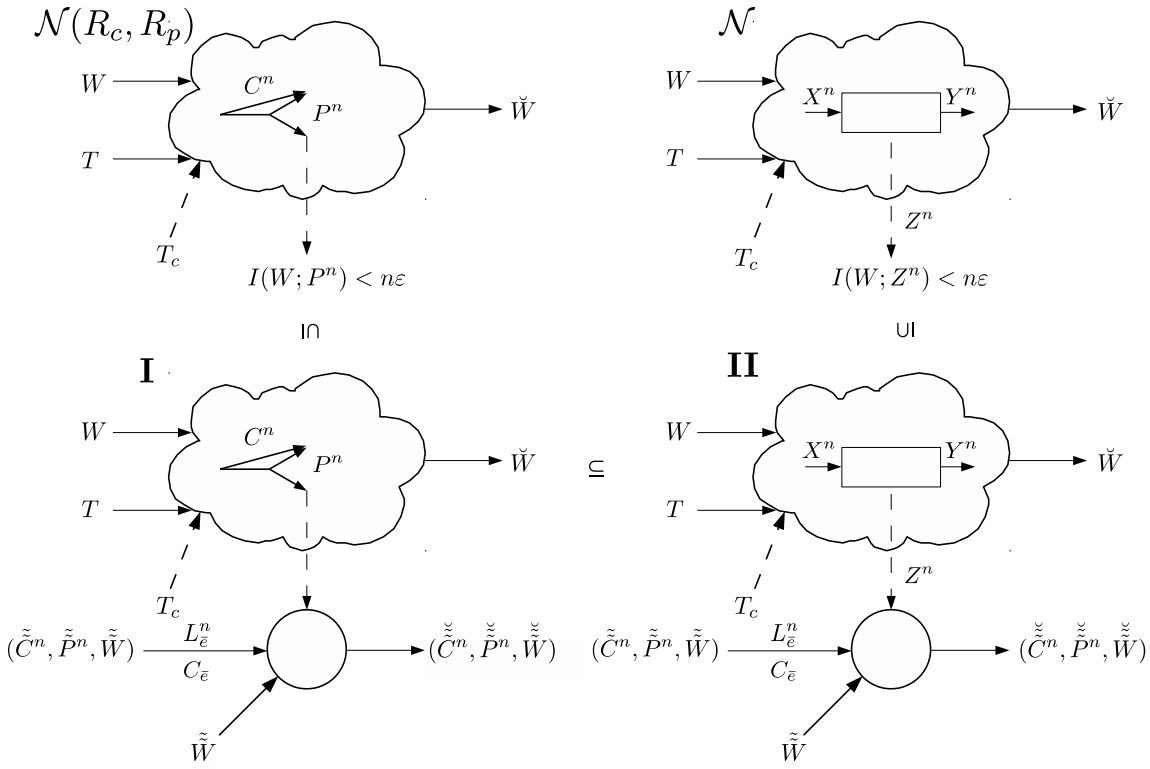


Fig. 2. Network $\mathcal{N}_{\bar{e}}(R_c, R_p)$ along with networks **I**, **II** and $\mathcal{N}$ that assist proving Theorem 3.

required to decode the confidential bits C^n.

We construct a code for a stacked version of network **I** with N_1 layers in which the auxiliary receiver is able to decode the confidential bits C^n. The constructed coded for the stacked version of network **I** can be seen as a code of blocklength $n_1 = nN_1$ for the non-stacked version of network **I**. To move the proof from network **I** to network **II** we use a stacked version of network **II** with N_2 layers. The code used at each layer of the stacked version of network **II** is the code of blocklength n_1 constructed above. We need to use a stacked version of network **II** to use a channel code at edge $\bar{e}$ of network **II** to emulate the noiseless edge $\bar{e}$ of network **I**.

In the following we show that the communication code of network **II** gives a secure code of network $\mathcal{N}$. These auxiliary receivers assist in the proof of the secrecy of the code for the eavesdropping set $\{e\} \in A$ in the following manner: capacity $C_{\bar{e}}$ is defined such that the sum of capacities of $(W, Z^n, L^n_{\bar{e}})$ (where $L^n_{\bar{e}}$ are the bits in the noiseless bit pipe of capacity$C_{\bar{e}}$) that are all the incoming links to the auxiliary receivers is almost equal to the entropy of (P^n, C^n, W) that correspond to the decoded message at the auxiliary receivers and therefore all links are filled up to capacity. Therefore there is no spare capacity at links Z^n to carry any information about message W and therefore the code is secure.

On the other hand, the upper bound result in Theorem 2 implies that $\mathcal{R}(\mathcal{N}, A) \subseteq \mathcal{R}(\mathcal{N}_{\bar{e}}(R_c + \epsilon, R_p + \epsilon), A)$ for any $\epsilon > 0$. We then prove a continuity result on the rate region $\mathcal{R}(\mathcal{N}_{\bar{e}}(R_c, R_p), A)$ with respect to (R_c, R_p) when $R_c > 0$ and $R_p > 0$. The lower bound result, the upper bound result, and the continuity result together prove Theorem 3. □

Example 1 demonstrates applications of Theorem 2 and 3 and shows that while Theorem 2 is tight in many cases, it is not always tight when the replaced link appears in one or more eavesdropping sets of size greater than 1.

Example 1: In the network of Figure 3(a), channels $e_1 = (1, 2, 1)$, $e_2 = (1, 4, 1)$, $e_3 = (1, 3, 1)$, $e_4 = (4, 2, 1)$, and $e_5 = (4, 3, 1)$ are independent degraded binary wiretap channels. Channels e_1 and e_3 have erasure probability 0 at each intended receiver and erasure probability $\frac{1}{2}$ at each wiretap output, as shown in Figure 3(e). Channels e_2, e_4, and e_5 have erasure probability $\frac{1}{2}$, with identical outputs for their intended and eavesdropped outputs, as shown in Figure 3(f). We consider a single multicast from source $\mathcal{S}$ at node 1 to terminals $\mathcal{T}_1$ and $\mathcal{T}_2$ at nodes 2 and 3. We therefore set $R^{(i\to\mathcal{B})} = 0$ for all $(i, \mathcal{B}) \neq (1, \{2, 3\})$ and then consider the point $R \in \mathcal{R}(\mathcal{N}, A)$ that maximizes $R^{(1\to\{2,3\})}$. The eavesdropper can listen in on either both e_1 and e_3 or just e_2, *i.e.*, $A = \big\{\{e_1, e_3\}, \{e_2\}\big\}$. The network $\breve{\mathcal{N}}$ shown in Figure 3(b) has secrecy capacity under adversarial set $A = \big\{\{e_1, e_3\}, \{e_2\}\big\}$ identical to that of the network in Figure 3(a) $\big(\mathcal{R}(\mathcal{N}, A) = \mathcal{R}(\breve{\mathcal{N}}, A)\big)$ and is obtained by three applications of Theorem 2. Here channel $\mathcal{C}_{e_4}$ and $\mathcal{C}_{e_5}$ have been replaced by channel $\mathcal{C}(\frac{1}{2}, 0)$ since channels e_4 and e_5 are invulnerable to eavesdropping ($e_4, e_5 \notin E$ for all $E \in A$). Likewise $\mathcal{C}_{e_2}$ has been replaced by $\mathcal{C}(0, \frac{1}{2})$ since e_2 cannot be simultaneously eavesdropped with any other channel ($e_2 \in E$ implies $|E| = 1$) and has 0 confidential

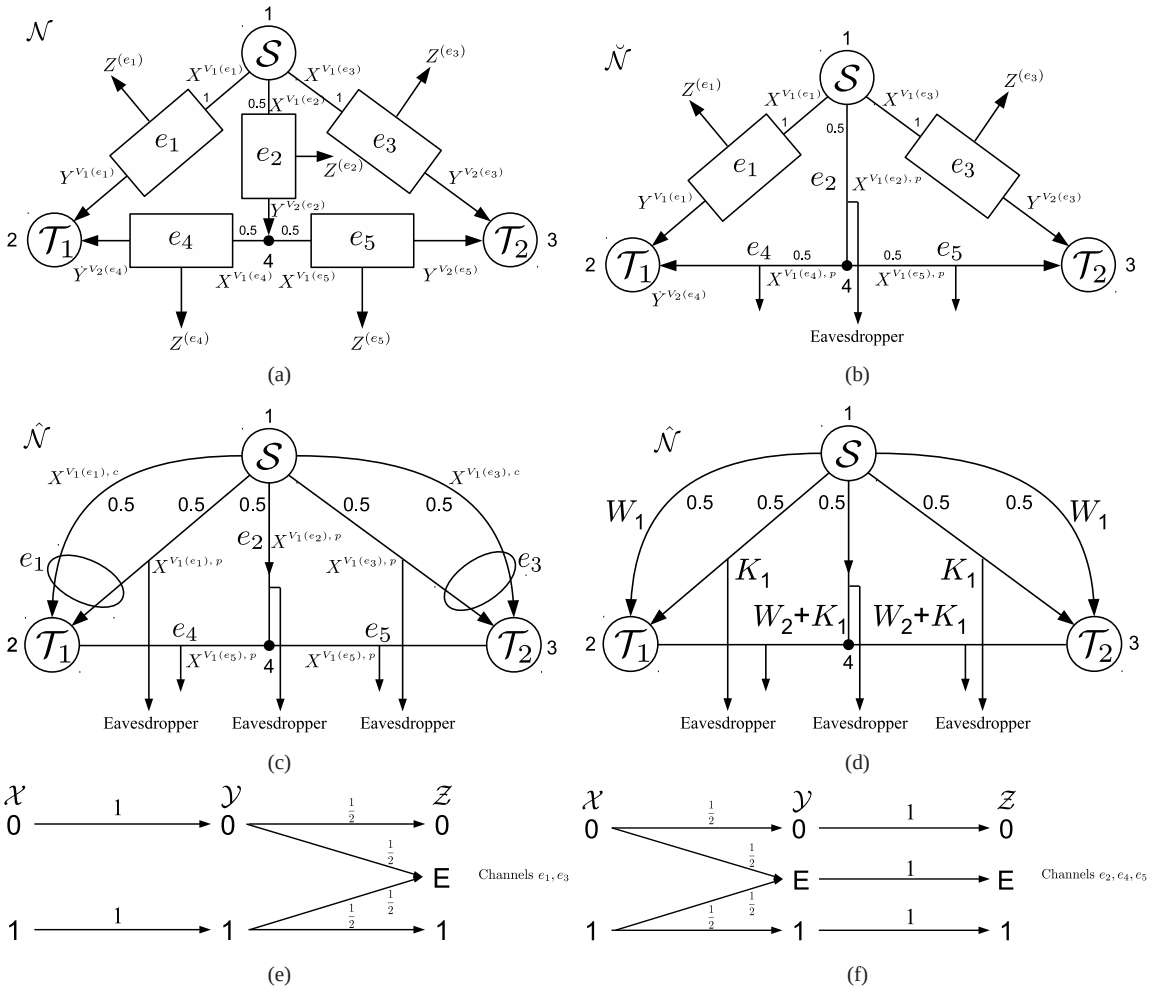


Fig. 3. (a) The network for Example 1 and (b) its equivalent model by replacing channels e_2, e_4, and e_5 by their equivalent noiseless links by Theorem 3 (rate-0 links are omitted from the model). (c) The noiseless model of (a) by applying Theorem 2 and (d) the secrecy capacity achieving code for the network in (c). (e), (f) The channel distributions for independent degraded wiretap channels e_1, e_3 and e_2, e_4, e_5 respectively.

bits. The noiseless network $\hat{\mathcal{N}}$ is an upper bounding model for the network in Figure 3(b) (and therefore also an upper bounding model for the network in Figure 3(a), giving $\mathcal{R}(\mathcal{N}, A) = \mathcal{R}(\breve{\mathcal{N}}, A) \subseteq \mathcal{R}(\hat{\mathcal{N}}, A)$), and is obtained by two applications of Theorem 2, replacing channels e_1 and e_3 by their upper bounding models.

A rate-1 blocklength-2 code for network $\hat{\mathcal{N}}$ is shown in Figure 3(d). The message $W^{(1\to\{2,3\})} \in \{0,1\}^2$ is broken into a pair of messages $W^{(1\to\{2,3\})} = \big(W_1, W_2\big) \in \{0,1\}^2$ with $H\big(W_1\big) = H\big(W_2\big) = 1$ and $H\big(W_1, W_2\big) = 2$. Random key $K_1 \in \{0,1\}$ is chosen uniformly at random and independently of $\big(W_1, W_2\big)$. The code is secure since $I\big(W_1, W_2; K_1\big) = 0$ and $I\big(W_1, W_2; W_2 + K_1\big) = 0$. In [6] we prove using information inequalities that the noisy network $\mathcal{N}$ of Figure 3(a) has multicast secrecy capacity at most 0.875.

To provide some intuition, notice that our capacity-achieving code for $\hat{\mathcal{N}}$ transmits the same key over a pair of noiseless links (e_1 and e_3 in $\hat{\mathcal{N}}$). Direct emulation of this solution in $\breve{\mathcal{N}}$ network in Figure 3(a) fails to maintain security. Specifically, if the same input is transmitted over channels e_1 and e_3 ($X_t^{(e_1)} = X_t^{(e_3)}$ for all $t \in \{1, \ldots, n\}$), then an eavesdropper accessing $E = \{e_1, e_3\}$ sees independent channel outputs $Z_t^{(e_1)}$ and $Z_t^{(e_3)}$ resulting from the same channel input $X_t^{(e_1)} = X_t^{(e_3)}$ at each time t. Since each transmitted bit is erased with probability $\frac{1}{2}$ and the erasure events are independent by assumption, an eavesdropper that wiretaps both e_1 and e_3 is expected to receive roughly 75% of the transmitted information bits. Consequently, a key of rate 0.5 is not enough to completely protect $W^{(1\to\{2,3\})}$ from the eavesdropper in this case. The problem here is that transmitting correlated information on multiple channels may be necessary to achieve the secure capacity in the noiseless case, but the same strategy may fail in the noisy case owing to independent realizations of probabilistic noise on different channels.

Theorems 4 and 5 provide two different lower bounds for the case of multiple wiretapped channels. These bounds correspond to achievable schemes that ensure all links to the eavesdropper are filled to capacity with independent randomness.

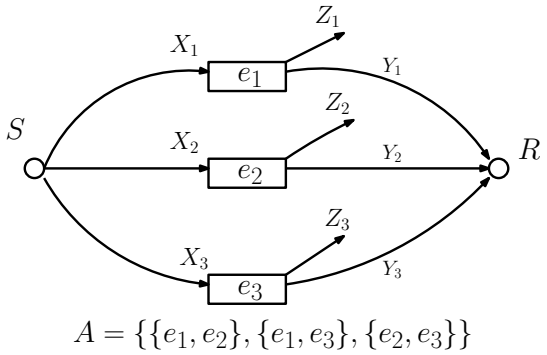


Fig. 4. An example wiretap network for which lower bound model-II is not tight but lower bound model-II is tight.

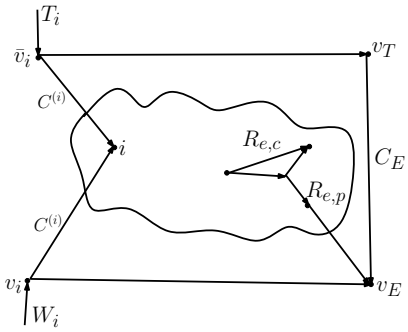


Fig. 5. The A-enhanced network $\mathcal{N}(A)$.

Lower bound model-I. The first lower bound results from removing the public portion of the upper bounding model. The lower bound is achievable since it is always possible to simply avoid the transmission of any rate on channel $\bar{e}$ that can be overheard by the eavesdropper.

Theorem 4: Consider a network $\mathcal{N}$, an adversarial set $A \subseteq \mathcal{P}(\mathcal{E})$, and a single link $\bar{e} \in \mathcal{E}$. $\mathcal{R}(\mathcal{N}_{\bar{e}}(R_c, 0), A) \subseteq \mathcal{R}(\mathcal{N}, A)$ for

$$R_c < \max_{p(x^{(\bar{e})})} I(X^{(\bar{e})}; Y^{(\bar{e})}) - \max_{p(x^{(\bar{e})})} I(X^{(\bar{e})}; Z^{(\bar{e})}).$$

Sketch of the proof: The proof of this theorem is similar to the proof of Theorem 3 except that in the noisy network we transmit independent random bits in place of public bits. □

The lower bound model-I of Theorem 4 is not tight in general. As a result, we do not use it to bound all channels but instead apply it to a selective sequence of channels from $\mathcal{E}$. Notice that the model $\mathcal{C}_{\bar{e}}(R_c, 0)$ for channel $\mathcal{C}_{\bar{e}}$ in Theorem 4 sets the public rate R_p to zero. This effectively removes $\bar{e}$ from all eavesdropping sets $E \in A$, giving a new adversarial set $A' = \big\{E \backslash \{\bar{e}\} : E \in A\big\}$. Repeated application of Theorem 4 on a carefully chosen sequence of channels enable us to reduce all eavesdropping sets to size at most one. Once this is accomplished, we can use the equivalence result of Theorem 3 to replace the remaining noisy channels.

To show that lower bound model-I is not tight, consider the network of Figure 4, where each i in $\{1, 2, 3\}$, $\max I(Y_i; X_i) = 2$ and $\max I(Z_i; X_i) = 1$. The adversary can eavesdrop any two of $\{e_1, e_2, e_3\}$. Since for each link in $\{e_1, e_2, e_3\}$ the confidential capacity is 1, and the public rate on two of the three links must be set to zero, the capacity of lower bound model-I is 3. In the following we introduce lower bound model-II, using which we get a tighter lower bound, 4, for this network.

Lower bound model-II. In this model we bound the secrecy capacity region of network $\mathcal{N}$ with adversarial set $A \subseteq \mathcal{P}(\mathcal{E})$ by deriving a relationship with the traditional capacity of a noiseless communication network called the A-enhanced network $\mathcal{N}(A)$ defined below and illustrated by Figure 5.

Definition 8: Consider network $\mathcal{N}$ on graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Define rate vector $R_{c,p} = \big((\check{R}_{e,c}, R_{e,p}) : e \in \mathcal{E}\big)$, and fix an adversarial set $A \subseteq \mathcal{P}(\mathcal{E})$. The A-enhanced network $\mathcal{N}(R_{c,p}, A)$ on graph $\check{\mathcal{G}} = (\check{\mathcal{V}}, \check{\mathcal{E}})$ is defined as follows:

1) $\check{\mathcal{V}} = \mathcal{V} \cup \big\{v_i : i \in \mathcal{V}\big\} \cup \big\{\bar{v}_i : i \in \mathcal{V}\big\} \cup \big\{v_E : E \in A\big\} \cup \{v_T\}$. For each $i \in \mathcal{V}$ we call v_i and $\bar{v}_i$ the i^{th} message node and random key node of network $\mathcal{N}(R_{c,p}, A)$. For each $E \in A$, node v_E is called an eavesdropper node. Node v_T is called the overall key node.
2) $\check{\mathcal{E}} = \big\{h_i : i \in \mathcal{V}\big\} \cup \big\{\bar{h}_i : i \in \mathcal{V}\big\} \cup \big\{\check{\mathcal{C}}_e : e \in \mathcal{E}\big\} \cup \big\{h_e : e \in \mathcal{E}\big\} \cup \big\{(v_T, v_E, 1) : E \in A\big\}$.

For each $i \in \mathcal{V}$, h_i is a noiseless hyperarc of capacity $C^{(i)}$ (or alternatively a set of bit pipes each of capacity $C^{(i)}$) from node v_i to all of the nodes in $\big\{i\big\} \cup \big\{v_E : E \in A\big\}$, and $\bar{h}_i$ is a noiseless hyperarc also of capacity $C^{(i)}$ (or alternatively a pair of bit pipes each of capacity $C^{(i)}$) from node $\bar{v}_i$ to both of the nodes in $\big\{i, v_T\big\}$, where $C^{(i)}$ is defined in (1) as the sum of the outgoing channel capacities from node i.

For each $e = (i, j, k) \in \mathcal{E}$, channel $\check{\mathcal{C}}_e$ in network is a bit pipe of capacity $R_{e,c}$ from node i to node j, and hyperarc h_e is a noiseless hyperarc of capacity $R_{e,p}$ from node i to all of the nodes in $\big\{j\big\} \cup \big\{v_E : E \in A, e \in E\big\}$. For every $E \in A$ channel $\mathcal{C}_{(v_T, v_E, 1)}$ is noiseless bit pipe of capacity

$$C_E = \sum_{i \in \mathcal{V}} C^{(i)} - \sum_{e \in E} R_{e,p}$$

from node v_T to node v_E.

The A-enhanced network is used for traditional (rather than secure) communication with a collection of reconstruction constraints that depend on both $\mathcal{N}$ and A.

Definition 9: Let $\mathcal{N}(R_{c,p}, A)$ be the A-enhanced network for network $\mathcal{N}$ and adversarial set $A \subseteq \mathcal{P}(\mathcal{E})$. A blocklength-$n$ solution $\mathcal{S}(\mathcal{N}(R_{c,p}, A))$ to network $\mathcal{N}(R_{c,p}, A)$ is defined as a set of encoding functions for each node v in $\check{\mathcal{V}}$

$$(X^{(v)})^n : (\mathcal{Y}^{(v)})_1^{n-1} \times (\mathcal{W}^{(v)})_1^{n-1} \times (\mathcal{T}^{(v)})_1^{n-1} \longrightarrow (\mathcal{X}^{(v)})^n$$

and decoding functions

$$(\hat{W^{(v)}})^n : (\mathcal{Y}^{(v)})_1^{n-1} \times (\mathcal{W}^{(v)})_1^{n-1} \times (\mathcal{T}^{(v)})_1^{n-1} \longrightarrow (\mathcal{W}^{(v)})$$
$$(\hat{T^{(v)}})^n : (\mathcal{Y}^{(v)})_1^{n-1} \times (\mathcal{W}^{(v)})_1^{n-1} \times (\mathcal{T}^{(v)})_1^{n-1} \longrightarrow (\mathcal{T}^{(v)}).$$

such that for each $i \in \mathcal{V}$ and $\mathcal{B} \in \mathcal{B}^{(i)}$, message $W^{(v_i \to \mathcal{B})}$ from node v_i is delivered to all of the nodes in $\mathcal{B} \in \mathcal{B}^{(i)}$, where $\mathcal{B}^{(i)}$ is the receivers set for node $i \in \mathcal{V}$ in network $\mathcal{N}$,

and random keys $T^{(i)} \in \mathcal{T}^{(i)} = \{1, \ldots, 2^{nC^{(i)}}\}$ are delivered from node $\bar{v}_i$ to nodes $\{v_E : E \in A\}$.

Definition 10: The rate region $\mathcal{R}(\mathcal{N}(R_{c,p}, A)) \subseteq \mathbb{R}_+^{m(2^{m-1}-1)}$ of the A-enhanced network $\mathcal{N}(R_{c,p}, A)$ of network $\mathcal{N}$ is the closure of all rate vectors R such that for any $\lambda > 0$, a solution (λ, R)–$\mathcal{S}(\mathcal{N}(R_{c,p}, A))$ exists.

Theorem 5: Consider network $\mathcal{N}$ on graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and an adversarial set $A \subseteq \mathcal{P}(\mathcal{E})$. Let $\mathcal{N}(R_{c,p}, A)$ be the A-enhanced network of network $\mathcal{N}$. If for every $e \in \mathcal{E}$

$$R_{e,p} < \max_{p(x)} I(X^{(e)}; Z^{(e)})$$
$$R_{c,p} < \max_{p(x)} I(X^{(e)}; Y^{(e)}) - \max_{p(x)} I(X^{(e)}; Z^{(e)}),$$

then $\mathcal{R}(\mathcal{N}(R_{c,p}, A)) \subseteq \mathcal{R}(\mathcal{N}, A)$.

Sketch of the proof: We start with a code for network $\mathcal{N}(R_{c,p}, A)$ and we will construct a secure code for network $\mathcal{N}$. We make use of an auxiliary network **I** which is the same as the A-enhanced network except that the noiseless bit pipes in $\{\check{\mathcal{C}}_e : e \in \mathcal{E}\} \cup \{h_e : e \in \mathcal{E}\}$ are changed back to the original noisy channels. We show that we can emulate the given code on network **I** such that the auxiliary receivers are still able to decode the required messages. Since the total capacity of all incoming links to the auxiliary receivers is almost equal to the entropy of $(P^n, C^n, W, (Z^{E\setminus\{\bar{e}\}})^n)$, there is no spare capacity at links $((Z^{E\setminus\{\bar{e}\}})^n, Z^n)$ to carry any information about message W and this corresponds to a secure code for network $\mathcal{N}$. □

Unlike the rest of the results, where changing a single wire-tap channel $\mathcal{C}_{\bar{e}}$ to its noiseless counterpart $\mathcal{C}_{\bar{e}}(R_c, R_p)$ results in an equivalent or bounding network, Theorem 5 requires all wiretap channels in the noisy network $\mathcal{N}$ to be changed to noiseless channels in order to obtain a lower bounding network. Intuitively, this is because our construction requires the eavesdropper $E \in A$ to decode all sources of randomness in the network, which is not possible generally for noisy networks where the entropy of the noise can be potentially infinite. If we wish to replace only some noisy channels by their noiseless counterparts then Theorem 4 should be used. When all channels are to be replaced Theorem 5 can be used, potentially leading to a tighter bound.

For example, we consider the network in Figure 4 where model-I gives a lower bound of 3. Here, we show that lower bound model-II gives a tighter lower bound, 4. The A-enhanced network is shown in Figure 6. For simplicity, we combine the three direct links (with capacity 1) from S to R into a single link with capacity 3. The following code achieves rate $(R_W, R_T) = (4, 6)$ in the A-enhanced network. Let $W = \{W_1, \ldots, W_4\}$ and $T = \{T_1, \ldots, T_6\}$. The outgoing link of S with capacity 3 directly delivers $\{W_1, W_2, W_3\}$ to R. Each of other outgoing links of S transmits a linearly independent combination of $\{W_4, T_5, T_6\}$. Node $\bar{V}_S$ transmits $\{T_1, T_2, T_3, T_4\}$ to each of $\{V_{\{1,2\}}, V_{\{1,3\}}, V_{\{2,3\}}\}$. Node $\{V_S\}$ transmits $\{W_1, \ldots, W_4\}$ to each of $\{V_{\{1,2\}}, V_{\{1,3\}}, V_{\{2,3\}}\}$. R can decode W_4 from the three linearly independent combinations of $\{W_4, T_5, T_6\}$. At

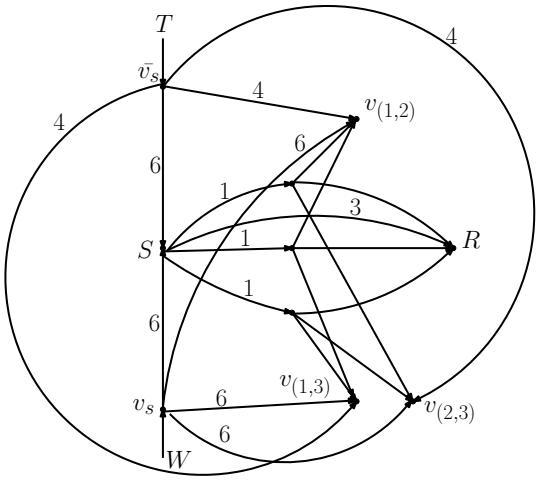


Fig. 6. The A-enhanced network for the network in Figure 4. The number on top of each link represents the link capacity.

$V_{\{1,2\}}$, messages $\{T_1, T_2, T_3, T_4\}$ and $\{W_1, W_2, W_3, W_4\}$ are directly received from $\bar{V}_S$ and V_S, respectively. By using W_4 and two linearly independent combinations of $\{W_4, T_5, T_6\}$, node $V_{\{1,2\}}$ can decode $\{T_5, T_6\}$. $V_{\{1,3\}}, V_{\{2,3\}}$ decode similarly.

ACKNOWLEDGMENTS

This work has been supported in part by NSF grants CNS 0905615, CCF 0830666, and CCF 1017632.

REFERENCES

[1] A. Wyner, "The wire-tap channel," *Bell Systems Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[2] N. Cai and R. W. Yeung, "Secure network coding," in *Proc. 2002 IEEE Int. Symp. Information Theory (ISIT 2002)*, Lausanne, Switzerland, Jun./Jul. 2002, p. 323.

[3] A. Mills, B. Smith, T. Clancy, E. Soljanin, and S. Vishwanath, "On secure communication over wireless erasure networks," in *Proc. of IEEE ISIT*, July 2008, pp. 161–165.

[4] R. Koetter, M. Effros, and M. Médard, "A Theory of Network Equivalence–Part I: Point-to-Point Channels," *Information Theory, IEEE Transactions on*, vol. 57, no. 2, pp. 972 –995, February 2011.

[5] ——, "A theory of network equivalence, Part II: Multiterminal Channels."

[6] T. Dikaliotis, H. Yao, T. Ho, M. Effros, and J. Kliewer, "Network equivalence in the presence of an eavesdropper," 2012. [Online]. Available: http://www.its.caltech.edu/$\sim$tho/eav-equi.pdf

[7] Leung-Yan-Cheong and M. S. Hellman, "The wire-tap channel," *Information Theory, IEEE Transactions on*, vol. 24, no. 4, pp. 451–456, Jul. 1978.

[8] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *Information Theory, IEEE Transactions on*, vol. 54, no. 4, pp. 2470–2492, June 2008.