The Role of Feedback in Two-Way Secure Communications

Xiang He, *Member, IEEE*, and Aylin Yener, *Member, IEEE*

***Abstract*—Most practical communication links are bidirectional. In these models, since the source node also receives signals, its encoder has the option of computing its output based on the signals it received in the past. On the other hand, from a practical point of view, it would also be desirable to identify the cases where such an encoder design may not improve communication rates. This question is particularly interesting for the case where the transmitted messages and the feedback signals are subject to eavesdropping. In this paper, we investigate the question of how much impact the feedback has on the secrecy capacity by studying two fundamental models. First, we consider the Gaussian two-way wiretap channel and derive an outer bound for its secrecy capacity region. We show that the secrecy rate loss can be unbounded when feedback signals are not utilized except for a special case we identify, and thus conclude that utilizing feedback can be highly beneficial in general. Second, we consider a half-duplex Gaussian two-way relay channel where the relay node is also an eavesdropper, and find that the impact of feedback is less pronounced compared to the previous scenario. Specifically, the loss in secrecy rate, when ignoring the feedback, is quantified to be less than 0.5 bit per channel use when the relay power goes to infinity. This achievable rate region is obtained with simple time sharing along with cooperative jamming, which, with its simplicity and near optimum performance, is a viable alternative to an encoder that utilizes feedback signals.**

***Index Terms*—Cooperative jamming, feedback, information theoretic security, two-way relay channel, two-way wiretap channel.**

I. Introduction

MOST communication links are bidirectional, where the backward channel can carry information and/or provide some form of feedback. For example, in ARQ schemes, the backward channel provides the acknowledgment of receipt of the packets. In peer-to-peer networks, information is communicated in both directions. The impact of the existence of bidirectionality on the channel capacity has been considered extensively up to date. Shannon proposed the two-way channel model in [1] where communication took place in both directions, and derived the inner bound and the outer bound on its capacity region. These bounds were shown to match for the full-duplex Gaussian two-way channel in [2]. An interesting implication of this result is that the signals received in the past, i.e., the feedback signals, are not needed for encoding to achieve the capacity region for this model. Though this feature is desirable in practice for simpler encoder design, it is also known that this approach is suboptimal in general, for example, as was proved in [3] for a two-way channel where the two nodes share a common output from the channel.

In secure communication, the question of whether feedback signals should be used for encoding has been studied in several special scenarios. Shannon showed that a completely secure backward channel can be used to send a one-time pad to increase the secrecy capacity of the forward channel [4]. In [5], it was proved that such a strategy, where the source node decodes the key from the destination, is optimal for a degraded wiretap channel with a secure rate limited noiseless feedback link. Another achievable scheme, which does not require decoding of the feedback, was first proposed in [6] in the setting of secret key generation and later in [7]. The scheme proves even if the forward channel and backward channel each has zero secrecy capacity, a positive secrecy rate can still be achieved when these two channels are used together. This is done by combining multiple channel uses and designing codes for the resulting equivalent broadcast channel in which the eavesdropper is eventually put at a disadvantage because of its lack of side information. Reference [8] combines this scheme with the key strategy in [4] and shows that a higher secrecy rate is achievable for the model in [7].

In [5], [7], and [8], the destination has the freedom to design the feedback signals. References [8] and [9] considered the scenario where the destination was restricted to sending its observation of the channel output, and could not manipulate the feedback signal to its advantage. It was shown that feedback helped to achieve a higher secrecy rate even in this case.

One feature that is common to the coding schemes in [5], [7], and [8] is that the eavesdropper always receives two separate sets of received signals: one from the forward channel and a second set of signals from the backward channel if it is not secure. While this is more inline with the conventional information theoretic models with feedback [10, Sec. 7.12], [11], letting the eavesdropper receiving the signals of the forward and the backward channel separately might inadvertently give the eavesdropper an advantage, as compared to superimposing them

Manuscript received November 23, 2009; revised May 31, 2012; accepted April 16, 2013. Date of publication September 16, 2013; date of current version November 19, 2013. This work was supported in part by the National Science Foundation under Grants CCF-051483, CNS-0716325, and CCF-0964362, and in part by the DARPA ITMANET Program under Grant W911NF-07-1-0028. This paper was presented in part at the 42nd Asilomar Conference on Signals, System and Computers, Pacific Grove, CA, USA, October 2008, and in part at the 2010 IEEE International Conference on Communications.

X. He was with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802 USA. He is now with Microsoft, Redmond, WA 98052 USA (e-mail: xianghe@microsoft.com).

A. Yener is with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802 USA (e-mail: yener@ee.psu.edu; yener@engr.psu.edu).

Communicated by M. Motani, Associate Editor for Communication Networks.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2013.2281711

together. Specifically, when the eavesdropper receives the sum of the outputs from the forward and the backward channels, introducing artificial noise into the backward channel at the time when the forward channel is in use can interfere with the eavesdropper's observation of the forward channel and hence reduce its recognizance of the message being transmitted on it. This so-called *cooperative jamming* scheme has been shown to improve secrecy rates in a Gaussian two-way channel with an external eavesdropper, i.e., the Gaussian two-way wiretap channel [12]. Yet in [12], the source node does not take advantage of the signals it received from the backward channel when encoding its transmission signals. The question remains, therefore, in such a cooperative jamming scheme, whether the achievable rates can be improved by utilizing these signals.

In this paper, we consider the wireless communication scenario where the eavesdropper observes the sum of the outputs of the forward and the backward channel, and hence the legitimate nodes in the network can potentially utilize feedback *and* cooperative jamming to protect the confidential message. We focus on two models where both techniques are potentially useful: 1) a class of Gaussian full-duplex two-way wiretap channels, and 2) a Gaussian half-duplex two-way relay channel with an untrusted relay. These two models represent two distinct communication scenarios allowing us to provide a comprehensive analysis on the merits of these techniques. They are also intimately related: we shall see that the results we obtain for the former are instrumental for the latter.

For the first model, we derive a computable outer bound to its secrecy capacity region. We then compare it to the achievable rates when the feedback is ignored at both nodes. Interestingly, when the ratio of the power constraint of the two legitimate nodes is fixed and the channel is fully connected with independent link noise, the gap between the achieved secrecy rate and the outer bound is bounded by a constant, which only depends on the channel gains.

On the other hand, when the ratio of the power constraints is not fixed, we show that ignoring feedback signals leads to unbounded loss in the secrecy rate when the power increases. The loss is measured as the gap between the achievable rate when the feedback is used and the upper bound when the feedback is not used, hence is not caused by the potential suboptimality of the achievable scheme. This result shows that utilizing the feedback for encoding at the legitimate nodes is highly beneficial for this model in general.

In the second model, we consider the case where the eavesdropper is part of the network rather than being external to it. In this model, two nodes wish to exchange information via a relay node from whom the information needs to be kept secret. Here the relay node is "honest but curious" [13], in that it will faithfully carry out the designated relaying scheme, but is not trusted to decode the message it is relaying. This setting was first considered in [14] for the three node relay channel and later thoroughly studied in [15] and [16]. Later, in [17], we considered a restricted version of the model in this paper, by studying the case when the feedback signals were not used at the source or the destination for encoding purposes. In this paper, we identify one case where doing so will not incur much loss in secrecy rate. In establishing this result, we utilize the outer bound found for the first model. Our analysis proves that if the power of the relay goes to ∞, then the loss in the secrecy rates caused by ignoring the feedback is bounded by 0.5 bit per channel use. Interestingly, a simple TDMA scheme with cooperative jamming yields the achievable rate.

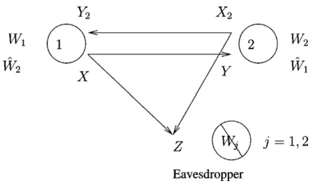

Fig. 1. Two-way wiretap channel.

The channel models in this paper are closely related to the channel-type model in the secret key generation literature; see [6], [18]–[21] for example. The major difference from these works is that our model accepts two inputs, one from the source, the other from the destination. The eavesdropper observes a noisy superposition of these two inputs. This is more complicated than the channel-type model where the noisy part of the channel is a wiretap channel which only accepts one input from the source node, and any input from the destination can only be transmitted over a noiseless public discussion link which is orthogonal to the wiretap channel. Recently, [22] has considered a channel-type secret key generation model where the channel component in the model accepts inputs from multiple nodes. Yet, these nodes only receive from the noiseless public discussion link [22, Sec. II], which is a fundamentally different model from those considered in this paper.

The remainder of the paper is organized as follows. In Section II, we describe the two models considered in this work. Section III focuses on the Gaussian two-way wiretap channel. Section IV focuses on the two-way relay channel with an untrusted relay. Section V concludes the work.

Throughout the paper, the notation $C(x)$ is defined as $C(x) = \frac{1}{2}\log_2(1 + x)$. Also x_i denotes the ith component of vector x, while x^i denotes $\{x_1, \ldots x_i\}$. $\mathcal{N}(0, \sigma^2)$ denotes a Gaussian distribution with zero mean and variance σ^2.

II. Channel Models

In this section, we describe the two channel models considered in this paper. Both models involve information exchange between two nodes: Node 1 and Node 2. Node 1 wants to send a message W_1 to Node 2. Node 2 wants to send a message W_2 to Node 1. Both messages must be kept secret from the eavesdropper. The encoding functions used at the two nodes are allowed to be stochastic. Without loss of generality, we use M_j to model the local randomness in the encoding function used by Node j, $j = 1, 2$.

A. Two-Way Wiretap Channel

The first model we consider in this paper is a two-way wiretap channel model. The channel model is shown in Fig. 1. The channel description is given by[1]

$$\Pr(Y, Y_2, Z|X, X_2) = \Pr(Z|X, X_2)\Pr(Y|X, X_2, Z)\Pr(Y_2|X_2, X, Z). \quad (1)$$

[1] Equation (1) includes the case where Y is correlated with Z when conditioned on X, X_2. This could be useful, for example, when the channel is fading, its states are correlated and the state values are not known to the transmitters or receivers.

From (1), we observe

$$Y_2 - \{X_2, X, Z\} - Y \tag{2}$$

is a Markov chain.

At each channel use, Node 1 and Node 2 transmit simultaneously. At the ith channel use, the encoding function of Node 1 is defined as

$$X_i = f_i({Y_2}^{i-1}, W_1, M_1). \tag{3}$$

The encoding function of Node 2 is defined as

$$X_{2,i} = g_i(Y^{i-1}, W_2, M_2). \tag{4}$$

Note that with the introduction of M_j, $j = 1, 2$, we can define f_i, g_i as deterministic encoders. Also note that another way to define f_i is $X_i = f_i(X^{i-1}, {Y_2}^{i-1}, M_1)$. It is easy to see that this definition is equivalent to the definition given in (3).

Let n be the total number of channel uses. Node 2 must decode W_1 reliably from ${X_2}^n$, Y^n, M_2, W_2. Node 1 must decode W_2 reliably from ${Y_2}^n$, X^n, M_1, W_1. Let the decoding results be $\hat{W}_1$ and $\hat{W}_2$, respectively. Then, we require

$$\lim_{n\to\infty} \Pr(W_j \neq \hat{W}_j) = 0, \quad j = 1, 2. \tag{5}$$

From Fano's inequality [10], we have

$$H(W_1|{X_2}^n, Y^n, M_2, W_2) < n\varepsilon_1 \tag{6}$$

$$H(W_2|{Y_2}^n, X^n, M_1, W_1) < n\varepsilon_2 \tag{7}$$

where $\varepsilon_j > 0$ and $\lim_{n\to\infty} \varepsilon_j = 0, j = 1, 2$.

In addition, both messages must be kept secret from the eavesdropper. Hence, we need

$$I(W_1, W_2; Z^n) < n\varepsilon_3 \tag{8}$$

where $\varepsilon_3 > 0$ and $\lim_{n\to\infty} \varepsilon_3 = 0$.

Define R_j, $j = 1, 2$ as

$$R_j = \lim_{n\to\infty} \frac{1}{n} H(W_j), \quad j = 1, 2. \tag{9}$$

The secrecy rate region is defined as all rate pairs $\{R_1, R_2\}$ for which (5) and (8) holds.

The Gaussian case of the two-way wiretap channel model was first proposed in [12] and [23] and is shown in Fig. 2. Formally, the channel is described as

$$Y_2 = X_2 + N_3 + \sqrt{\alpha}X \tag{10}$$

$$Y = X + N_1 + \sqrt{\beta}X_2 \tag{11}$$

$$Z = \sqrt{h_1}X + \sqrt{h_2}X_2 + N_2 \tag{12}$$

where $\sqrt{\alpha}$, $\sqrt{\beta}$, $\sqrt{h_1}$, $\sqrt{h_2}$ are the channel gains. N_i, $i = 1, 2, 3$ are Gaussian random variables with zero mean and unit variance, representing the channel noise. We assume that given N_2, N_1 is independent from N_3:

$$p(N_1, N_2, N_3) = p(N_2)p(N_1|N_2)p(N_3|N_2). \tag{13}$$

We use ρ to denote the correlation between N_1 and N_2. η denotes the correlation between N_2 and N_3. Obviously, $-1 \leq \rho \leq 1$, and $-1 \leq \eta \leq 1$.

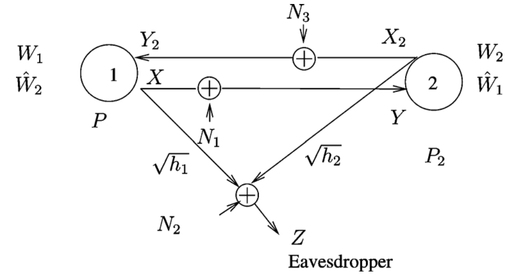

Fig. 2. Gaussian two-way wiretap channel.

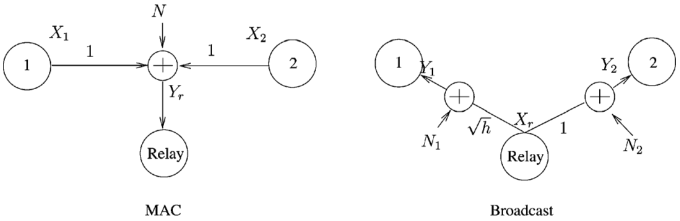

Fig. 3. Gaussian two-way half-duplex relay channel with an untrusted relay.

From (1) and (13), we readily see this channel belongs to the class of channels described by (1) and shown in Fig. 1.

Observe that the terms $\sqrt{\alpha}X$ and $\sqrt{\beta}X_2$ are not shown in Fig. 2. This is because each node knows its own transmitted signal and $\sqrt{\alpha}$, $\sqrt{\beta}$, $\sqrt{h_1}$, $\sqrt{h_2}$, and can always subtract the interference caused by its own transmitted signals. Hence, we can remove $\sqrt{\alpha}X$ and $\sqrt{\beta}X_2$ from (10) and (11). The channel is thus equivalent to

$$Y_2 = X_2 + N_3 \tag{14}$$

$$Y = X + N_1 \tag{15}$$

$$Z = \sqrt{h_1}X + \sqrt{h_2}X_2 + N_2. \tag{16}$$

In the sequel, we shall focus on this equivalent model instead.

Let the power constraint of Node 1 be P, and of Node 2 be P_2, i.e.,

$$\frac{1}{n}\sum_{k=1}^{n} E\left[X_k^2\right] \leq P \tag{17}$$

$$\frac{1}{n}\sum_{k=1}^{n} E\left[{X_{2,k}}^2\right] \leq P_2. \tag{18}$$

Remark 1: When Y_2 is a constant, or, the feedback is ignored by Node 1, the model reduces to the relay channel with a confidential message to the relay, which was considered in [16], [24], and [25]. □

B. Two-Way Relay Channel With an Untrusted Relay

The second model we consider in this paper is the Gaussian two-way relay channel with an untrusted relay node. The channel model is shown in Fig. 3. At any time slot, the channel either behaves as a multiple-access channel (MAC), shown on the left, or as a broadcast channel (BC), shown on the right. After normalizing the channel gains, the MAC can be expressed as

$$Y_r = X_1 + X_2 + N. \tag{19}$$

The BC can be expressed as

$$Y_1 = \sqrt{h}X_r + N_1 \tag{20}$$

$$Y_2 = X_r + N_2 \tag{21}$$

where $\sqrt{h}$ is the channel gain, $h \neq 0$. N, N_1, N_2 are independent zero mean Gaussian random variables with unit variance.

We assume Node 1 and Node 2 transmit simultaneously during the MAC mode. $X_{j,i}$, $j = 1, 2$ denote the signals transmitted by Node j during the ith channel use such that the channel is in MAC mode, $i \geq 1$. We use ϕ_i to denote the number of channel uses that the channel was in the broadcast mode before this channel use. The notation X_j^i denotes the set of signals: $\{X_{j,k}, k = 1 \dots i\}$.

Similarly, $X_{r,i}$ denotes the signal transmitted by the relay node during the ith channel use that the channel is in broadcast mode, $i \geq 1$. We use ψ_i to denote the number of channel uses that the channel was in the MAC mode before this channel use.

$Y_{1,i}$, $Y_{2,i}$, $Y_{r,i}$ are received signals defined in the same fashion.

The channel switches between the MAC mode and the broadcast mode according to a globally known schedule. We assume the schedule is independent from the local randomness at each node, the messages and the channel noise. The first mode is assumed to be the MAC mode. The case where the first mode is a broadcast mode can be viewed as a special case of invoking the MAC mode first by transmitting nothing during the first MAC mode. The rate loss caused by the wasted channel use is negligible as the number of channel uses goes to ∞.

Suppose the MAC mode is activated for n channel uses. The broadcast mode is activated for m channel uses. Hence, the communication spans over $n + m$ channel uses. It should be noted that, in general, neither the n channel uses of the MAC mode, nor the m channel uses of the broadcast mode have to be consecutive. We assume the schedule is stable, in the sense that the following limit exists:

$$\alpha = \lim_{n+m\to\infty} \frac{n}{m+n}. \tag{22}$$

For a given α, we use $\{T(\alpha)\}$ to denote a sequence of schedules with increasing total number of channel uses $n + m$ such that (22) holds, and α is the limit of the time sharing factor of the MAC mode in the schedule $T(\alpha)$ as $n + m \to \infty$.

The average power constraints for Node 1, Node 2, and the relay can be expressed as

$$\frac{1}{m+n}\sum_{k=1}^{n} E\left[X_{i,k}^2\right] \leq \bar{P}_i, \quad i = 1, 2, \tag{23}$$

$$\frac{1}{m+n}\sum_{k=1}^{m} E\left[X_{r,k}^2\right] \leq \bar{P}_r. \tag{24}$$

For the purpose of completeness, we also introduce the notation $P_i, i = 1, 2$, to denote the average power of Node i during the MAC mode. Since these two nodes are only transmitting during the MAC mode, P_i and $\bar{P}_i$ are related as

$$P_i = \bar{P}_i/\alpha, \quad i = 1, 2. \tag{25}$$

Similarly, we use P_r to denote the average power of the relay node during the broadcast mode. Since the relay node only transmits during the broadcast mode, P_r is related to $\bar{P}_r$ as follows:

$$P_r = \bar{P}_r/(1-\alpha). \tag{26}$$

For the ith channel use in which the channel operates in the MAC mode, the encoding functions at Node 1, $f_{1,i}$, is defined as

$$X_{1,i} = f_{1,i}(Y_1^{\phi_i}, W_1, M_1). \tag{27}$$

Similarly, the encoding functions at Node 2, $f_{2,i}$, is defined as

$$X_{2,i} = f_{2,i}(Y_2^{\phi_i}, W_2, M_2). \tag{28}$$

Note that $f_{1,i}, f_{2,i}$ are deterministic functions, and we use M_r to model the local randomness at the relay. For the ith channel use in which the channel operates in broadcast mode, the encoding function of the relay node g_i is defined as:

$$X_{r,i} = g_i(Y_r^{\psi_i}, M_r) \tag{29}$$

where g_i is a deterministic function.

The eavesdropper knows Y_r^n, X_r^m, M_r. Therefore, the secrecy constraint is expressed as

$$\lim_{m+n\to\infty} \frac{1}{m+n} H(W_1, W_2|Y_r^n, X_r^m, M_r) = \lim_{m+n\to\infty} \frac{1}{m+n} H(W_1, W_2). \tag{30}$$

Since $W - \{X_r^m, Y_r^n\} - M_r$ is a Markov chain, we have

$$\lim_{m+n\to\infty} \frac{1}{m+n} H(W_1, W_2|Y_r^n, X_r^m, M_r) = \lim_{m+n\to\infty} \frac{1}{m+n} H(W_1, W_2|Y_r^n, X_r^m). \tag{31}$$

Therefore, the secrecy constraint can be expressed as

$$\lim_{m+n\to\infty} \frac{1}{m+n} H(W_1, W_2|Y_r^n, X_r^m) = \lim_{m+n\to\infty} \frac{1}{m+n} H(W_1, W_2). \tag{32}$$

Let $\hat{W}_j$, $j = 1, 2$ be the decoding result computed by the intended receiver of $W_j, j = 1, 2$. Then, the reliable communication requirement is expressed as

$$\lim_{m+n\to\infty} \Pr(\hat{W}_j \neq W_j) = 0, j = 1, 2. \tag{33}$$

Define R_1, R_2 as

$$R_j = \lim_{m+n\to\infty} \frac{1}{n+m} H(W_j), j = 1, 2. \tag{34}$$

The secrecy capacity region is defined as the union of all rate pairs (R_1, R_2) such that there is an α, a sequence of schedules $\{T(\alpha)\}$ and a choice of encoding function for which (32) and (33) are satisfied.

Remark 2: In general,

$$\lim_{n+m\to\infty} \frac{1}{n+m} H(W|X_r^m, Y_r^n) \neq \lim_{n+m\to\infty} \frac{1}{n+m} H(W|Y_r^n). \tag{35}$$

This can be proved by a counterexample: consider the communication protocol:

1) First the relay node randomly generates and broadcasts a key via X_r to Node 1 and Node 2 using a channel code.
2) Node 1 uses the key as a one-time pad [4] to encrypt its confidential message W and sends it to the relay using a channel code. The other nodes remain silent.
3) The relay decodes the codeword sent by Node 1 and encodes and forwards it to the destination.
4) The destination recovers the codeword sent by Node 1 by decoding the signals from the relay. It then decrypts it with the key it received in step 1 and recovers W.

Since the one-time pad is a perfectly secure cipher [4], for this communication protocol, we have

$$H(W) = H(W|Y_r^n). \tag{36}$$

However, since the key is determined by X_r^m, given the key, W is uniquely determined by Y_r^n. Therefore, we have

$$H(W|X_r^m, Y_r^n) = 0 \neq H(W|Y_r^n). \tag{37}$$

□

III. Feedback in the Two-Way Wiretap Channel

A. Improvement on the Known Achievable Secrecy Rate: A Motivating Example

For the two-way wiretap channel, reference [12] derived an achievable rate using Gaussian codebooks. However, in this scheme, the signal Y_2 received by Node 1 is not used to compute the signal X transmitted by Node 1. Likewise, the signal Y received by Node 2 is not used to compute the signal X_2 transmitted by Node 2. We next show that this scheme can be improved upon with respect to the achievable secrecy rate. To show this, it is sufficient to show that a larger R_1 is achievable for Node 1 for a set of channel gains. In the following, we provide such an example.

We assume $\rho = 0$, $\eta = 0$, which means N_1, N_2, N_3 are all independent, which was the setting considered by [12]. The largest rate for Node 1 achievable with the scheme of [12] is given by

$$R_1 = [C(P) - C\left(\frac{h_1 P}{h_2 P_2 + 1}\right)]^+ \tag{38}$$

which is achieved by letting Node 2 transmit an i.i.d. Gaussian sequence with variance P_2. When $\frac{h_1}{h_2 P_2 + 1} \geq 1$, we observe from (38) that the secrecy rate is always 0. Below, we choose $P = 3$, $P_2 = 1$, $\sqrt{h_1} = \sqrt{2}$, $\sqrt{h_2} = 1$ such that this condition is fulfilled and prove a positive secrecy rate is achievable with our scheme.

The coding scheme we use is similar to that of [6]. It is composed of one channel use described in Fig. 4, followed by one channel use described in Fig. 5. In an odd step, Node 1 sends a signal denoted by J_1 and Node 2 sends a signal denoted by X_2. After this step, Node 1 adds its received signal $X_2 + N_3$ to a new signal X_1 and transmits it in the following even step. At the same time, Node 2 sends a signal denoted by J_2. We use the notation N_i to denote the channel noise in the odd step and N_i' to denote the channel noise in the even step.

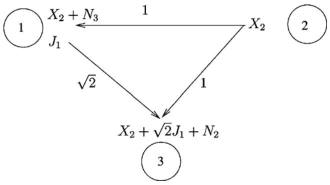

Fig. 4. Odd step.

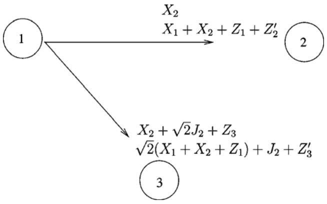
1
X_2
$X_1 + X_2 + Z_1 + Z_2'$
2
$X_2 + \sqrt{2}J_2 + Z_3$
$\sqrt{2}(X_1 + X_2 + Z_1) + J_2 + Z_3'$
3

Fig. 5. Even step.

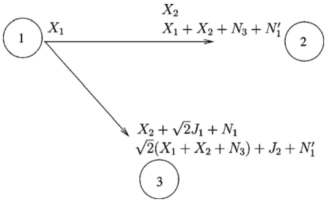
1
X_1
X_2
$X_1 + X_2 + N_3 + N_1'$
2
$X_2 + \sqrt{2}J_1 + N_1$
$\sqrt{2}(X_1 + X_2 + N_3) + J_2 + N_1'$
3

Fig. 6. Equivalent channel.

Combining these two steps, we obtain an equivalent memoryless channel shown in Fig. 6. The achievable secrecy rate for this channel is given by [26]

$$[I(X_1; Y) - I(X_1; Y_{e,1}, Y_{e,2})]^+ \tag{39}$$

where

$$Y = X_1 + N_3 + N_1' \tag{40}$$
$$Y_{e,1} = X_2 + \sqrt{2}J_1 + N_2 \tag{41}$$
$$Y_{e,2} = \sqrt{2}(X_1 + X_2 + N_3) + J_2 + N_2'. \tag{42}$$

We then choose X_1, X_2, J_1, J_2 as zero mean independent Gaussian random variables with unit variance. From Figs. 4 and 5, this choice satisfies the average power constraints. Evaluating (39) for this distribution, we obtain

$$C\left(\frac{1}{2}\right) - C\left(\frac{a^2}{2a^2 + 2 - \frac{a^2}{a^2+2}}\right) > 0 \tag{43}$$

where $a = \sqrt{2}$.

Since the original channel takes twice as many channel uses to implement this scheme, the actual secrecy rate is half the value indicated by (43). Nevertheless, the achievable secrecy rate is positive.

This means that the scheme that utilizes feedback signals leads to higher achievable secrecy rate for this channel compared to the scheme of [12].

B. Utilizing Feedback is Beneficial: A Definitive Answer

Although we have shown that using feedback can improve the secrecy rate, it remains unclear whether this can only be done by letting Node 1 use the signal Y to compute X, or whether there are smarter schemes that can outperform those with feedback. In this section, we settle this question by showing that the secrecy rate that is achieved by utilizing feedback can exceed an upper bound on the secrecy rate when feedback is ignored.

We begin with the achievable rates. Let us use $[x]^+$ to denote $\max\{x, 0\}$. Then, we have the following theorem.

Theorem 1: Define R_1^* as

$$R_1^* = \max_{0\le\alpha\le1} \alpha \left[\begin{array}{l} C(P) - \\ \left[\begin{array}{l} C\left(\frac{h_1 P}{h_2 P_2+1}\right) - \\ \frac{1-\alpha}{\alpha}\left[C(P_2) - C\left(\frac{h_2 P_2}{h_1 P+1}\right)\right]^+ \end{array}\right]^+ \end{array}\right]^+ \tag{44}$$

and R_2^* as

$$R_2^* = \max_{0\le\alpha\le1} \alpha \left[\begin{array}{l} C(P_2) - \\ \left[\begin{array}{l} C\left(\frac{h_2 P_2}{h_1 P+1}\right) - \\ \frac{1-\alpha}{\alpha}\left[C(P) - C\left(\frac{h_1 P}{h_2 P_2+1}\right)\right]^+ \end{array}\right]^+ \end{array}\right]^+ . \tag{45}$$

Define the region $\mathbf{R}$ as the convex hull of the following three rate pairs of (R_1, R_2):

$$(0, 0), \quad (R_1^*, 0), \quad (0, R_2^*). \tag{46}$$

The rate region $\mathbf{R}$ is achievable.

Proof: The proof is given in Appendix C. ■

Remark 3: The achievable scheme is composed of two phases. During phase 1, with a time sharing factor of $1 - \alpha$, Node 2 sends a key to Node 1. During phase two, Node 1 utilizes this key to encrypt its message and transmits the result to Node 2. Hence, $\alpha = 1$ corresponds to the case when both nodes ignore their received signals when computing their transmitting signals. □

Fig. 7 plots the achievable secrecy rate R_1 with $\alpha = 0.5$. Also plotted in the figure is an upper bound on R_1 found without utilizing Y_2, i.e., without feedback in the model. This upper bound, given later by Theorem 5, is derived from a more general result in Theorem 3 and hence is relegated to Section III-C. Fig. 7 shows that the achieved secrecy rate with feedback can exceed the upper bound without feedback by an amount that increases with power, a fact we shall state formally in Theorem 8. This means that it is impossible to achieve the same secrecy rate by designing the encoder at Node 2 if Node 1 ignores the feedback signal Y_2.

Having made our point, we now provide the derivation of an outer bound on the secrecy capacity region of Gaussian two-way wiretap channel, for which the upper bound in Fig. 7 is a special case.

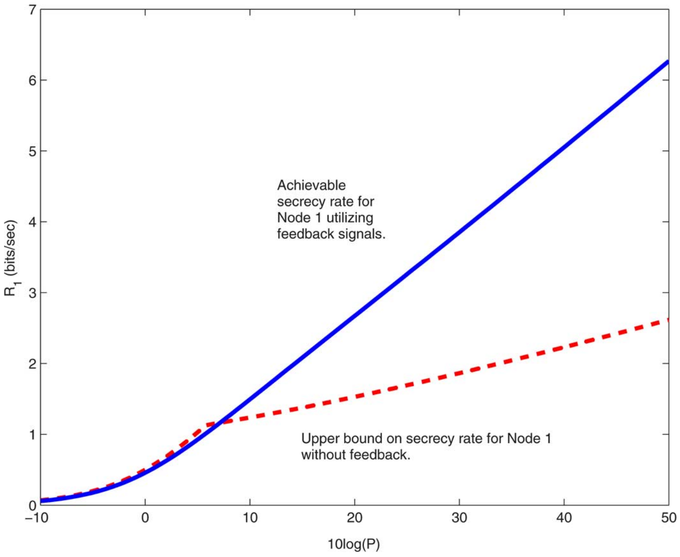

Fig. 7. Comparison of the sum secrecy rate when $P_2 = P^{1/4}$, $h_1 = h_2 = 1$, $\rho = \eta = 0$, $\alpha = 0.5$. Achievable secrecy rate at Node 1 is computed according to (44). Upper bound on the secrecy rate at Node 1 when Y_2 is ignored is computed with Theorem 5 in Section III-C.

C. Outer Bound

We begin by deriving an upper bound on R_1.

Theorem 2: For the channel model in Fig. 1, R_1 is upper bounded by

$$\max_{\Pr(X,X_2)} \min\{I(X;Y), I(X;Y|Z,X_2) + I(X_2;Y_2,Z|X)\}. \tag{47}$$

Proof: See Appendix A. ■

Remark 4: Ignoring Y_2 at Node 1 is equivalent to viewing Y_2 as a constant. From (47), R_1, in this case, is upper bounded by

$$\max_{\Pr(X,X_2)} \min\{I(X;Y), I(X;Y|Z,X_2) + I(X_2;Z|X)\} \tag{48}$$

which is the upper bound proved in [25]. □

Theorem 3: The secrecy capacity region of the channel model in Fig. 1 is bounded by

$$\cup_{\Pr(X,X_2)} \{(R_1, R_2) : (50)\ (51)\ (52) \text{ holds}\} \tag{49}$$

$$0 \le R_1 \le I(X;Y) \tag{50}$$

$$0 \le R_2 \le I(X_2;Y_2) \tag{51}$$

$$R_1 + R_2 \le \min\left\{\begin{array}{l} I(X;Y|Z,X_2) + I(X_2;Z,Y_2|X), \\ I(X_2;Y_2|Z,X) + I(X;Z,Y|X_2) \end{array}\right\}. \tag{52}$$

Proof: The proof is provided in Appendix B. ■

For a deterministic binary wire-tap channel, Theorem 3 leads to the equivocation capacity region, as shown by the following theorem.

Theorem 4: When X, X_2 are binary and $Y = X \oplus X_2, Y_2 = X_2 \oplus X, Z = X \oplus X_2$, the secrecy capacity region is given by

$$R_j \ge 0, j = 1, 2 \tag{53}$$

$$R_1 + R_2 \le 1. \tag{54}$$

Proof: The achievability follows from [23, Th. 2]. The converse follows from Theorem 3. The sum rate bound specializes as follows:

$$\begin{aligned} & I\left(X;Y|Z,X_2\right)+I\left(X_2;Y_2,Z|X\right) && (55)\\ &= I\left(X;X|X\oplus X_2,X_2\right)+I\left(X_2;X_2,X\oplus X_2|X\right) && (56)\\ &= I\left(X;X|X,X_2\right)+I\left(X_2;X_2,X|X\right) && (57)\\ &= I\left(X_2;X_2,X|X\right) && (58)\\ &\le H\left(X_2\right)\le 1. && (59)\end{aligned}$$

■

We next consider the Gaussian channel.

Theorem 5: When Y_2 is a constant, i.e., Y_2 is ignored by Node 1, the secrecy rate R_1 is upper bounded by

$$\inf_{\sigma^2\ge 0} C\left(\frac{P\left(1+\sigma^2-\sqrt{h_1}\rho\right)^2}{\left(1+\sigma^2-\rho^2\right)\left(h_1P+1+\sigma^2\right)}\right)+C\left(\frac{h_2P_2}{1+\sigma^2}\right). \quad (60)$$

Proof: Define N_4 as a Gaussian random variable such that $N_4\sim\mathcal{N}(0,\sigma^2)$ and is independent from N_i, $i=1,2,3$. Recall that Z is the signal received by the eavesdropper. We next consider a channel where the eavesdropper receives $Z+N_4$. Since $Z+N_4$ is a degraded version of Z, we can find an upper bound of the original channel by deriving an upper bound for this new channel. This upper bound is found by applying the bound (48).

We next prove that all terms in the upper bound (48) is maximized when X, X_2 are independent and each has a Gaussian distribution with zero mean and maximum possible variance: $I(X;Y)$ is obviously maximized by this distribution. For the other two terms, we have

$$\begin{aligned} & I\left(X;Y|X_2,Z+N_4\right) && (61)\\ &= I\left(X;X+N_1|X_2,\sqrt{h_1}X+\sqrt{h_2}X_2+N_2+N_4\right) && (62)\\ &= h\left(X+N_1|X_2,\sqrt{h_1}X+\sqrt{h_2}X_2+N_2+N_4\right)\\ &\quad - h\left(N_1|N_2+N_4\right) && (63)\\ &\le h\left(X+N_1|\sqrt{h_1}X+N_2+N_4\right)-h\left(N_1|N_2+N_4\right) && (64)\end{aligned}$$

and

$$\begin{aligned} & I\left(X_2;Z+N_4|X\right) && (65)\\ &= I\left(X_2;\sqrt{h_2}X_2+N_2+N_4|X\right) && (66)\\ &= h\left(\sqrt{h_2}X_2+N_2+N_4|X\right)-h\left(N_2+N_4\right) && (67)\\ &\le h\left(\sqrt{h_2}X_2+N_2+N_4\right)-h\left(N_2+N_4\right). && (68)\end{aligned}$$

Equations (64) and (68) show that the second term in (48) is maximized when X and X_2 are independent. Moreover, (64) is known to be maximized when X has a Gaussian distribution with the maximum possible variance; see [27]. Equation (68) is also maximized when X_2 has a Gaussian distribution with the maximum possible variance. Hence, we have shown the optimal input distribution for X, X_2 is an independent Gaussian distribution. For this distribution, it can be verified the second term in (48) becomes (60).

Hence, we have proved the theorem. ■

Remark 5: When $\sigma^2\to\infty$, (60) converges to $C(P)$, which corresponds to the first term in (48). Thus, (60) is written as one term instead of the two terms as in (48). □

Remark 6: We introduce N_4 to further tighten the bound. For example, consider the case where $\rho=\eta=0$. In this case, the upper bound can be expressed as

$$\min_{0\le\alpha\le 1} C\left(\frac{P}{\alpha h_1P+1}\right)+C\left(\alpha h_2P_2\right) \quad (69)$$

where $\alpha=1/(1+\sigma^2)$. Consider choosing the remaining parameters as $h_1=1$, $h_2=10$, $P=100$, $P_2=5$. It can be verified that the minimum is attained around $\alpha=0.09$, and not at $\sigma^2=0$. Hence, the bound presented here is tighter than the bound in [25]. □

Next, we present the following theorem.

Theorem 6: The secrecy capacity region of the Gaussian two-way wiretap channel is outer bounded by

$$\begin{aligned} & 0\le R_1\le C(P),0\le R_2\le C(P_2) && (70)\\ & R_1+R_2\le && (71)\end{aligned}$$

$$\min\left\{\begin{array}{l}\inf_{\sigma^2\ge 0}\left(\begin{array}{l}C\left(\frac{P\left(1+\sigma^2-\sqrt{h_1}\rho\right)^2}{(1+\sigma^2-\rho^2)(h_1P+1+\sigma^2)}\right)\\ +C\left(\frac{P_2\left(h_2+1+\sigma^2-2\sqrt{h_2}\eta\right)}{1+\sigma^2-\eta^2}\right)\end{array}\right),\\ \inf_{\sigma^2\ge 0}\left(\begin{array}{l}C\left(\frac{P_2\left(1+\sigma^2-\sqrt{h_2}\eta\right)^2}{(1+\sigma^2-\eta^2)(h_2P_2+1+\sigma^2)}\right)\\ +C\left(\frac{P\left(h_1+1+\sigma^2-2\sqrt{h_1}\rho\right)}{1+\sigma^2-\rho^2}\right)\end{array}\right)\end{array}\right\}. \quad (72)$$

Proof: Again we consider a channel where the eavesdropper receives $Z+N_4$ and derive an outer bound for this new channel. N_4 is as defined in the proof of Theorem 5.

To prove the theorem, we first show $I(X;Y)$, $I(X;Y|Z,X_2)$, $I(X_2;Y_2,Z|X)$, $I(X_2;Y_2)$, $I(X_2;Y_2|Z,X_2)$ and $I(X;Z,Y|X_2)$ are maximized simultaneously when X and X_2 are independent, $X\sim\mathcal{N}(0,P)$, and $X_2\sim\mathcal{N}(0,P_2)$.

Due to the symmetry of the channel model, we only need to show $I(X;Y)$, $I(X;Y|Z,X_2)$, and $I(X_2;Y_2,Z|X)$ are maximized by this distribution.

The case of $I(X;Y|Z,X_2)$ was shown in the proof of Theorem 5.

For $I(X_2;Y_2,Z|X)$, we have

$$\begin{aligned} & I\left(X_2;Y_2,Z|X\right) && (73)\\ &= I\left(X_2;X_2+N_3,\sqrt{h_2}X_2+N_2+N_4|X\right) && (74)\\ &= h\left(\sqrt{h_2}X_2+N_2+N_4,X_2+N_3|X\right)-h\left(N_2+N_4,N_3\right) && (75)\\ &\le h\left(\sqrt{h_2}X_2+N_2+N_4,X_2+N_3\right)-h\left(N_2+N_4,N_3\right). && (76)\end{aligned}$$

Hence, $I(X_2;Y_2,Z|X)$ is maximized when X and X_2 are independent, $X\sim\mathcal{N}(0,P)$ and $X_2\sim\mathcal{N}(0,P_2)$. The theorem then is a consequence of Theorem 3 when evaluated at this input distribution. ■

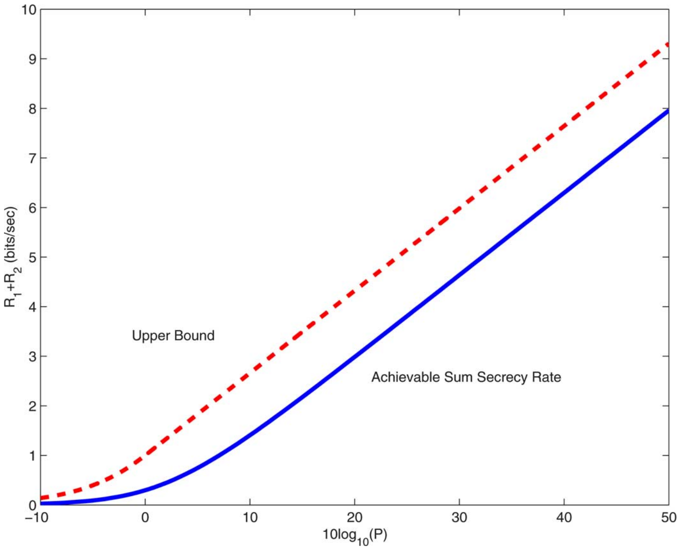

Fig. 8. Comparison of the sum secrecy rate when $P = P_2, h_1 = h_2 = 1, \rho = \eta = 0$. The sum secrecy rate computed with (218), when Y_2 is ignored at Node 1. The upper bound on the sum rate is computed with (72).

Remark 7: The introduction of N_4 is again useful in tightening the bound. For example, consider the case where $\rho = \eta = 0$, $h_1 = 1$, $h_2 = 10$, $P = 100$, $P_2 = 5$.

In this case the upper bound on R_1, which is $C(P)$, is about 3.3291. The first term inside the minimum in (72), which is also an upper bound on R_1 takes the form

$$\min_{0\leq\alpha\leq1} C\left(\frac{P}{\alpha h_1 P + 1}\right) + C\left(P_2\left(\alpha h_2 + 1\right)\right) \tag{77}$$

where $\alpha = 1/(1+\sigma^2)$. It can be verified that the minimum is smaller than 3.24 and is attained around $\alpha = 0.32$. Hence, the upper bound on R_1 is dominated by the first term inside the minimum in (72) and is not attained at $\sigma^2 = 0$. □

D. Comparing the Achievable Rates and the Outer Bound

1) $\rho = \eta = 0$: First let us consider the case with independent link noise, i.e., the model considered in [12].

Theorem 7: When $\rho = \eta = 0$, $P_2 = kP$, k is a positive constant, and $h_j \neq 0$, $j = 1, 2$, the loss in secrecy rates when received signals are not used to compute transmitting signals at Node j, $j = 1, 2$ is bounded by a constant, which is only a function of h_1 and h_2.

Proof: The proof is given in Appendix D. ■

Fig. 8 illustrates the case described in Theorem 7. In this figure, we fix $P = P_2$, increase P, and compare the achievable sum secrecy rate and its upper bound. It is observed that the gap between these two does not increase with P.

Next, the following theorem states the fact illustrated in Fig. 7.

Theorem 8: Even in the case where cooperative jamming is possible ($h_j \neq 0, j = 1, 2$), when P is not linearly increasing with P_2, ignoring Y_2 at Node 1 can lead to *unbounded* loss in the secrecy rate.

Proof: The proof is given in Appendix E. ■

2) $h_1 \leq 1$, $\rho = \sqrt{h_1}$ and Y_2 is a Constant: We next consider a special case of the model that attracted some interest in the past, see for example, [25], [28]. In this model, Z is a degraded version of Y given X_2, and Y_2 is ignored by Node 1.

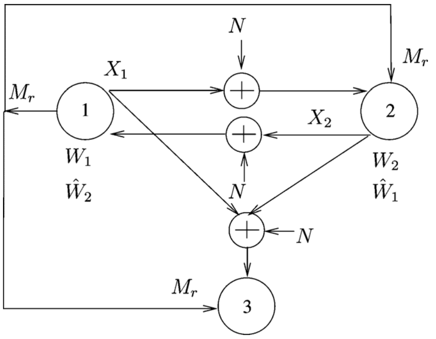

Fig. 9. Two-way wiretap channel with additional public noiseless forward link.

In this case, N_2 can be written as $\sqrt{h_1}N_1 + N_2'$, where N_2' is independent from N_1, N_3, and $N_2' \sim \mathcal{N}(0, 1 - h_1)$. Then, the signals received by the eavesdropper Z can be expressed as

$$Z = \sqrt{h_1}X + \sqrt{h_2}X_2 + \sqrt{h_1}N_1 + N_2' \tag{78}$$

$$= \sqrt{h_1}\left(X + N_1\right) + N_2' + \sqrt{h_2}X_2. \tag{79}$$

From this, we observe that, given X_2, Z is a degraded version of $Y = X + N_1$.

Corollary 1: When $h_1 \leq 1$, $\rho = \sqrt{h_1}$, and $h_2 \neq 0$, Y_2 is a constant, then the achievable rate of R_1 using cooperative jamming is at most 0.5 bit per channel use from the secrecy capacity.

Remark 8: Corollary 1 was first proposed in [28] and later appeared in [25]. Here we describe the approach of [25]. The approach of [28] is different and uses results on the wiretap channel with noisy feedback, which can be found in [29].

From Theorem 1 and Remark 3, the achievable rate for R_1 in this case is obtained by letting $\alpha = 1$ and evaluating R_1^*. In this case,

$$R_1 = C(P) - C(\frac{h_1 P}{h_2 P_2 + 1}). \tag{80}$$

The upper bound proposed in [25] on R_1 is

$$\min\{C(P), C(P) - C(h_1 P) + C(h_2 P_2)\}. \tag{81}$$

Here we observe (81) can be obtained from (60) when evaluated with $\sigma^2 \to \infty$ and $\sigma^2 = 0$. Reference [25] proves Corollary 1 by comparing (80) and (81). It can be readily verified that the gap between (80) and (81) is less than 0.5 bit per channel use.

IV. Feedback in Half-Duplex Two-Way Relay Channel With an Untrusted Relay

In this section, we derive an outer bound for the secrecy capacity region of the two-way relay channel with an untrusted relay in Section II-B (see Fig. 3). To find the outer bound, we first consider the channel in Fig. 9.

We assume X_1 and X_2 have the same power constraint as the X_1, X_2 in Fig. 3. M_r is now accessible to Node 1 and delivered to the other nodes via a public noiseless link. The remaining part of the channel is activated when the original two-way relay channel is in the MAC mode, and is inactive when the original two-way relay channel model is in the broadcast mode. Doing so ensures the overall number of channel uses to be the same between these two models.

Recall that $M_j, j = 1, 2$ still models the local randomness at Node j, $j = 1, 2$. The encoding function of Node 1 at the ith channel use when the channel is active can be defined as

$$X_{1,i} = \tilde{f}_{1,i}(Y_1^{i-1}, W_1, M_1, M_r). \tag{82}$$

Similarly, the encoding function of Node 2 at the ith channel use when the channel is active can be defined as

$$X_{2,i} = \tilde{f}_{2,i}(Y_2^{i-1}, W_2, M_2, M_r). \tag{83}$$

With these preparations, we present the following theorem.

Theorem 9: The secrecy rate region of the channel in Fig. 9 includes the secrecy capacity region of the two-way relay channel in Fig. 3.

Proof: Consider the model in Fig. 3. Suppose during a MAC mode, a genie reveals X_1+X_2+N to Node 1 and Node 2. We also add a public noiseless link that takes inputs from Node 1 and provides outputs to Node 2 and the relay. We make M_r accessible to Node 1 and use the public noiseless link to deliver M_r to Node 2 and the relay. This side information does not increase the knowledge of the relay and hence will not decrease the secrecy capacity region of the channel.

During a broadcast mode, a genie reveals the link noise level N_2 to Node 2. Similarly, the link noise N_1 is revealed to Node 1. This side information will not decrease the secrecy capacity region of the channel either.

With the side information provided to the nodes, the links from the relay to Nodes 1 and 2 can be removed. This is because

1) Nodes 1 and 2 have the signal received by the relay $X_1 + X_2 + N$.
2) Node 1 sends M_r via the public noiseless forward link. With M_r available at Node 2, it can compute the signal transmitted by the relay node. Due to the same reason, Node 1 knows the signal transmitted by the relay node as well.
3) With noise N_2 available at Node 2, Node 2 can compute the signal it received from the relay. For similar reasons, Node 1 can compute the signal it received from the relay as well.

Since N_1, N_2, N are independent, N_1 and N_2 can be incorporated as the local randomness at Nodes 1 and 2, respectively.

After removing the links from the relay to Nodes 1 and 2, the channel indeed becomes that which is described by Fig. 9, where Node 3 corresponds to the relay node whose output broadcast link to Nodes 1 and 2 is removed. Since, every step we took during this transformation could only expand the secrecy capacity region, we have proved the theorem. ■

To derive an outer bound for the secrecy capacity of the channel in Fig. 9, we first consider the case when the channel is active regardless of whether the two-way relay channel is in MAC mode or broadcast mode. We recognize that in this case, the channel becomes a special case of the two-way wiretap channel defined in Section II. Utilizing this connection leads to the following corollary.

Corollary 2: The secrecy capacity region of the channel in Fig. 9 is outer bounded by

$$R_1 + R_2 \leq \min\left\{C\left(\bar{P}_1\right), C\left(\bar{P}_2\right)\right\} \tag{84}$$

$$R_1 \geq 0, R_2 \geq 0 \tag{85}$$

where $\bar{P}_i$ is the average power constraint of Node i.

Proof: The channel in Fig. 9 is a special case of the channel defined in (1), where

$$Y, Y_2, Z, X, X_2 \tag{86}$$

in (1) correspond to

$$\{X_1 + N, M_r\}, X_2 + N, \{X_1 + X_2 + N, M_r\}, \{X_1, M_r\}, X_2 \tag{87}$$

in Fig. 9, respectively, and $\Pr(Y, Y_2, Z|X, X_2)$ becomes $\Pr(N)$.

Therefore, the corollary follows as a direct consequence of Theorem 6 with $\eta = \rho = 1, h_1 = h_2 = 1, \sigma^2 = 0$. ■

Remark 9: It is interesting to note that despite the fact that the two models considered in this paper are distinct, the outer bound we obtained for the two-way wiretap channel is useful and in fact necessary to obtain the outer bound for the two-way untrusted relay channel. □

Note that to apply Corollary 2 to the half-duplex two-way relay channel, we need to take into account the channel uses when the channel in Fig. 9 is inactive during the channel uses when the original two-way relay channel is in the broadcast mode. Hence, the outer bound in Corollary 2 becomes the following region $\mathbf{A}$:

$$R_1 + R_2 \leq \alpha \min\left\{C\left(\bar{P}_1/\alpha\right), C\left(\bar{P}_2/\alpha\right)\right\} \tag{88}$$

$$R_1 \geq 0, R_2 \geq 0 \tag{89}$$

which reflects the number of channel uses during which some nodes are inactive.

Define region $\mathbf{B}$ as

$$0 \leq R_1 \leq (1-\alpha)C(\bar{P}_r/(1-\alpha)) \tag{90}$$

$$0 \leq R_2 \leq (1-\alpha)C(h\bar{P}_r/(1-\alpha)). \tag{91}$$

Then, we have the following theorem.

Theorem 10: An outer bound for the secrecy capacity of two-way relay channel is given by

$$\cup_{0\leq\alpha\leq 1}\{\mathbf{A} \cap \mathbf{B}\}. \tag{92}$$

Proof: Region $\mathbf{A}$ follows by applying Corollary 2 and taking into account the fact that the channel is inactive when the original two-way relay channel is in broadcast mode as described above. Region $\mathbf{B}$ follows from removing the secrecy constraint and applying the cut-set bound [10, Th. 15.10.1], by considering the cut where the set $\mathcal{T}$ includes the relay node and Node 2, which leads to (91). Equation (90) is derived similarly due to the symmetry of the channel model. ■

Remark 10: When $\bar{P}_r \to \infty$, and $h \neq 0$, then the region is maximized when $\alpha \to 1$. The outer bound thus becomes

$$R_1 + R_2 \leq \min\left\{C\left(\bar{P}_1\right), C\left(\bar{P}_2\right)\right\} \tag{93}$$

$$R_1 \geq 0, R_2 \geq 0. \tag{94}$$

□

A. Comparison With Achievable Rates

In this section, we compare the outer bound with the achievable secrecy rate region. We begin by restating an achievable rate for R_1 from [17]. The achievable strategy involves Node 2 to act as a cooperative jammer whenever Node 1 transmits the message to be kept secret from the relay. The relay uses compress-and-forward. Both Nodes 1 and 2 ignore their received signals when computing the transmitted signals.

Theorem 11 [17, Th. 1]: The following secrecy rate of R_1 is achievable for the model in Fig. 3:

$$0 \le R_1 \le \max_{0 \le P_1' \le \bar{P}_1/\alpha, 0<\alpha<1} \alpha \left[C\left(\frac{P_1'}{(1+\sigma_c^2)}\right) - C\left(\frac{P_1'}{(1+P_2)}\right)\right]^+ \tag{95}$$

where σ_c^2 is the variance of the Gaussian quantization noise determined by

$$\alpha C\left(\frac{P_1'+1}{\sigma_c^2}\right) = (1-\alpha) C\left(P_r\right). \tag{96}$$

P_2 was defined in (25), P_r was defined in (26).

The rate region (R_1, R_2) then follows from time sharing.

Remark 11: For any fixed α such that $0 < \alpha < 1$, if the power of the relay $\bar{P}_r \to \infty$, then $\sigma_c^2 \to 0$, the achievable rate converges to

$$\alpha(C(P_1) - C(\frac{P_1}{1+P_2})). \tag{97}$$

Equation (97) is a monotonically increasing function of α. Hence, as long as $\alpha < 1$, we can always increase α and increase the achievable secrecy rate. Therefore, when $\bar{P}_r \to \infty$, the optimal time sharing factor $\alpha \to 1$. The achievable rate then converges to

$$C(\bar{P}_1) - C(\frac{\bar{P}_1}{1+\bar{P}_2}). \tag{98}$$

The secrecy rate region is obtained with time sharing and it converges to

$$R_1 + R_2 \le C(\bar{P}_1) - C(\frac{\bar{P}_1}{1+\bar{P}_2}) \tag{99}$$

$$R_1 \ge 0, R_2 \ge 0. \tag{100}$$

Hence, the achieved sum secrecy rate goes to ∞ when $\bar{P}_r \to \infty$ and $\bar{P}_1 = \bar{P}_2 \to \infty$. □

Utilizing this result, we have the following corollary.

Corollary 3: When $\bar{P}_r \to \infty$, the gap between the outer bound and the achievable rate is bounded by 0.5 bit per channel use.

To prove this corollary, we use the following simple fact. Define the following functions:

$$f(x,y) = \frac{1}{2}\log_2\left(\frac{(1+x)(1+y)}{1+x+y}\right) \tag{101}$$

$$g(x,y) = \min\{C(x), C(y)\}. \tag{102}$$

Let $h(x,y) = g(x,y) - f(x,y)$. Then, $0 \le h(x,y) \le 0.5$.

Corollary 3 can then be proved by letting $x = \bar{P}_1, y = \bar{P}_2$. The upper bound on the sum rate and the achievable sum secrecy rate then become $g(x,y)$ and $f(x,y)$ when $\bar{P}_r \to \infty$, which means the gap between the upper bound and lower bound of the sum secrecy rate is less than 0.5 bit per channel use. Since the achievable region and the outer bound are only different on the bounds for the sum rate, this proves the gap between the inner bound and outer bound of the secrecy capacity region is also less than 0.5 bit per channel use when $\bar{P}_r \to \infty$. Hence, we have proved Corollary 3.

V. Conclusion

In this paper, we have investigated the merit of using the signals received by the source node, i.e., the feedback, for encoder design on achieving a larger secrecy rate region. In order to answer this question, we have studied two models: the Gaussian two-way wiretap channel, and the Gaussian half-duplex two-way relay channel with an untrusted relay. For each model, we have derived a computable outer bound for the secrecy capacity region. For the first model, by measuring the gap between the outer bound and the achievable rate region, we have found that the loss in secrecy rate due to ignoring the feedback signals can be unbounded. Hence, the use of feedback can be highly beneficial in this model. For the second model, we have found that the feedback can be safely ignored if the power of the relay is abundant. In particular, the gap between the achievable rate region and the outer bound is bounded by 0.5 bit per channel use when the power of the relay goes to ∞. It is worth mentioning that the achievable rate region in this case is attained via a time sharing cooperative jamming scheme, which, with its simplicity and near optimum performance, is a viable alternative to an encoding scheme that utilizes feedback signals.

Appendix A
Proof of Theorem 2

Let $\varepsilon = \varepsilon_1 + \varepsilon_3$, where ε_1 was defined in (6), and ε_3 was defined in (8). To simplify the notation, we use M_2' to denote $\{M_2, W_2\}$. Then, we have

$$H\left(W_1\right) - n\varepsilon \tag{103}$$
$$\le H\left(W_1|Z^n\right) - H\left(W_1|Z^n, {X_2}^n, Y^n, M_2'\right) \tag{104}$$
$$= I\left(W_1; M_2', {X_2}^n, Y^n|Z^n\right) \tag{105}$$
$$= I\left(W_1; {X_2}^n|Z^n, Y^n, M_2'\right) + I\left(W_1; M_2', Y^n|Z^n\right) \tag{106}$$
$$= I\left(W_1; M_2', Y^n|Z^n\right) \tag{107}$$
$$\le I\left(W_1, M_1, {Y_2}^n; M_2', Y^n|Z^n\right) \tag{108}$$
$$= I\left(W_1, M_1, {Y_2}^n; M_2', Y^n, Z^n\right) - I\left(W_1, M_1, {Y_2}^n; Z^n\right) \tag{109}$$

where (104) follows from (6) and (8). Note that since, in this proof, we are only bounding the rate of W_1, we omit W_2 from the condition term of (6). Equation (107) follows from the fact that ${X_2}^n$ is a deterministic function of Y^{n-1} and M_2', as shown in (4).

Then, we rewrite the first term in (109) as

$$I\left(W_1, M_1, {Y_2}^n; M_2', Y^n, Z^n\right) \tag{110}$$
$$= I\left(W_1, M_1, {Y_2}^n; Y_n|Z_n, M_2', Y^{n-1}, Z^{n-1}\right) + I\left(W_1, M_1, {Y_2}^n; Y^{n-1}, Z^n, M_2'\right). \tag{111}$$

For the first term in (111), we have

$$\begin{align}
&I\left(W_1, M_1, Y_2{}^n; Y_n|Z_n, M_2', Y^{n-1}, Z^{n-1}\right) \tag{112}\\
&= I\left(W_1, M_1, Y_2{}^n; Y_n|X_{2,n}, Z_n, M_2', Y^{n-1}, Z^{n-1}\right) \tag{113}\\
&\leq h\left(Y_n|Z_n, X_{2,n}\right) - \nonumber\\
&\quad h\left(Y_n|X_{2,n}, M_2', Y^{n-1}, Z^n, W_1, M_1, Y_2{}^n\right) \tag{114}\\
&= h\left(Y_n|Z_n, X_{2,n}\right) - \nonumber\\
&\quad h\left(Y_n|X_{2,n}, X_n, M_2', Y^{n-1}, Z^n, W_1, M_1, Y_2{}^n\right) \tag{115}\\
&= h\left(Y_n|Z_n, X_{2,n}\right) - h\left(Y_n|X_{2,n}, X_n, Z_n\right) \tag{116}\\
&= I\left(X_n; Y_n|Z_n, X_{2,n}\right). \tag{117}
\end{align}$$

In (113), we use the fact that $X_{2,n}$ is a deterministic function of $\{M_2', Y^{n-1}\}$, as shown by (4). In (115), we use the fact that X_n is a deterministic function of $\{W_1, M_1, Y_2{}^{n-1}\}$, as shown by (3). In (116), we use the fact that

$$Y_n - \{X_{2,n}, X_n, Z_n\} - \{M_2', Y^{n-1}, Z^{n-1}, W_1, M_1, Y_2{}^n\} \tag{118}$$

is a Markov chain, due to (1), the channel being memoryless and the fact that encoding functions are causal. In particular, (1) allows us to remove $Y_{2,n}$ from the condition term. Applying this result, we find that (109) is upper bounded by

$$I\left(X_n; Y_n|Z_n, X_{2,n}\right) + I\left(W_1, M_1, Y_2{}^n; Y^{n-1}, M_2'|Z^n\right). \tag{119}$$

The second term in (119) can be rewritten as

$$\begin{align}
&I\left(W_1, M_1, Y_2{}^n; Y^{n-1}, M_2'|Z^n\right) \tag{120}\\
&= I\left(W_1, M_1, Y_2{}^{n-1}; Y^{n-1}, M_2'|Z^n\right) + \nonumber\\
&\quad I\left(Y_{2,n}; Y^{n-1}, M_2'|W_1, M_1, Y_2{}^{n-1}, Z^n\right). \tag{121}
\end{align}$$

The second term in (121) can be upper bounded as

$$\begin{align}
&I\left(Y_{2,n}; Y^{n-1}, M_2'|W_1, M_1, Y_2{}^{n-1}, Z^n\right) \tag{122}\\
&= I\left(Y_{2,n}; Y^{n-1}, M_2'|X_n, W_1, M_1, Y_2{}^{n-1}, Z^n\right) \tag{123}\\
&= h\left(Y_{2,n}|X_n, W_1, M_1, Y_2{}^{n-1}, Z^n\right) - \nonumber\\
&\quad h\left(Y_{2,n}|X_n, W_1, M_1, Y_2{}^{n-1}, Z^n, Y^{n-1}, M_2'\right) \tag{124}\\
&\leq h\left(Y_{2,n}|X_n, Z_n\right) - \nonumber\\
&\quad h\left(Y_{2,n}|X_n, W_1, M_1, Y_2{}^{n-1}, Z^n, Y^{n-1}, M_2'\right) \tag{125}\\
&= h\left(Y_{2,n}|X_n, Z_n\right) - \nonumber\\
&\quad h\left(Y_{2,n}|X_{2,n}, X_n, Z_n, W_1, M_1, Y_2{}^{n-1}, Z^{n-1}, Y^{n-1}, M_2'\right) \tag{126}\\
&= h\left(Y_{2,n}|X_n, Z_n\right) - h\left(Y_{2,n}|X_{2,n}, X_n, Z_n\right) \tag{127}\\
&= I(X_{2,n}; Y_{2,n}|X_n, Z_n). \tag{128}
\end{align}$$

In (123), we use the fact that X_n is a deterministic function of $\{W_1, M_1, Y_2{}^{n-1}\}$, as shown by (3). In (126), we use the fact that $X_{2,n}$ is a deterministic function of M_2', Y^{n-1}, as shown by (4). In (127), we use the fact that

$$Y_{2,n} - \{X_{2,n}, X_n, Z_n\} - \{W_1, M_1, Y_2{}^{n-1}, Z^{n-1}, Y^{n-1}, M_2'\} \tag{129}$$

is a Markov chain. This is because the encoding functions are causal and the channel is memoryless.

Applying this result, we find that (119) is now upper bounded by

$$\begin{align}
&I\left(X_n; Y_n|Z_n, X_{2,n}\right) + I(X_{2,n}; Y_{2,n}|X_n, Z_n) \nonumber\\
&\quad + I\left(W_1, M_1, Y_2{}^{n-1}; Y^{n-1}, M_2'|Z^n\right). \tag{130}
\end{align}$$

The last term in (130) can be rewritten as

$$\begin{align}
&I\left(W_1, M_1, Y_2{}^{n-1}; Y^{n-1}, M_2'|Z^{n-1}\right) + \nonumber\\
&I\left(W_1, M_1, Y_2{}^{n-1}; Z_n|Y^{n-1}, M_2', Z^{n-1}\right) \nonumber\\
&\quad - I\left(W_1, M_1, Y_2{}^{n-1}; Z_n|Z^{n-1}\right). \tag{131}
\end{align}$$

The second term and the last term in (131) can be upper bounded together

$$\begin{align}
&I\left(W_1, M_1, Y_2{}^{n-1}; Z_n|M_2', Y^{n-1}, Z^{n-1}\right) \nonumber\\
&\quad - I\left(W_1, M_1, Y_2{}^{n-1}; Z_n|Z^{n-1}\right) \tag{132}\\
&= -I\left(Z_n; M_2', Y^{n-1}|Z^{n-1}\right) \nonumber\\
&\quad - h\left(Z_n|W_1, M_1, Y_2{}^{n-1}, M_2', Y^{n-1}, Z^{n-1}\right) \nonumber\\
&\quad + h\left(Z_n|Z^{n-1}, W_1, M_1, Y_2{}^{n-1}\right) \tag{133}\\
&\leq -h\left(Z_n|W_1, M_1, Y_2{}^{n-1}, M_2', Y^{n-1}, Z^{n-1}\right) + \nonumber\\
&\quad h\left(Z_n|Z^{n-1}, W_1, M_1, Y_2{}^{n-1}\right) \tag{134}\\
&= -h\left(Z_n|X_n, X_{2,n}, W_1, M_1, Y_2{}^{n-1}, M_2', Y^{n-1}, Z^{n-1}\right) \nonumber\\
&\quad + h\left(Z_n|X_n, Z^{n-1}, W_1, M_1, Y_2{}^{n-1}\right) \tag{135}\\
&\leq -h\left(Z_n|X_n, X_{2,n}, W_1, M_1, Y_2{}^{n-1}, M_2', Y^{n-1}, Z^{n-1}\right) \nonumber\\
&\quad + h\left(Z_n|X_n\right) \tag{136}\\
&= -h\left(Z_n|X_n, X_{2,n}\right) + h\left(Z_n|X_n\right) \tag{137}\\
&= I(X_{2,n}; Z_n|X_n). \tag{138}
\end{align}$$

In (135), we use the fact that X_n is a deterministic function of $\{W_1, M_1, Y_2{}^{n-1}\}$, and $X_{2,n}$ is a deterministic function of $\{M_2', Y^{n-1}\}$. In (137), we use the fact that

$$Z_n - \{X_n, X_{2,n}\} - \{W_1, M_1, Y_2{}^{n-1}, M_2', Y^{n-1}, Z^{n-1}\} \tag{139}$$

is a Markov chain. This is due to the fact that the channel is memoryless and the encoding functions (3) and (4) are causal.

Applying this result to (131), we find that (130) is now upper bounded by

$$\begin{align}
&I\left(X_n; Y_n|X_{2,n}, Z_n\right) + I(X_{2,n}; Y_{2,n}, Z_n|X_n) \nonumber\\
&\quad + I\left(W_1, M_1, Y_2{}^{n-1}; Y^{n-1}, M_2'|Z^{n-1}\right). \tag{140}
\end{align}$$

Hence, we have shown that

$$\begin{align}
H(W_1) - n\varepsilon &\leq I\left(W_1, M_1, Y_2{}^n; Y^n, M_2'|Z^n\right) \nonumber\\
&\leq I\left(X_n; Y_n|X_{2,n}, Z_n\right) + I\left(X_{2,n}; Y_{2,n}, Z_n|X_n\right) \nonumber\\
&\quad + I\left(W_1, M_1, Y_2{}^{n-1}; Y^{n-1}, M_2'|Z^{n-1}\right). \tag{141}
\end{align}$$

Applying this result repeatedly for $n-1, n-2, \ldots, 1$, we have

$$\begin{align}
&\frac{1}{n}H(W_1) - \varepsilon \tag{142}\\
&\leq \frac{1}{n}\sum_{i=1}^{n}(I(X_i; Y_i|X_{2,i}, Z_i) + I(X_{2,i}; Y_{2,i}, Z_i|X_i)). \tag{143}
\end{align}$$

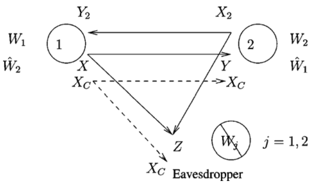

Fig. 10. Two-way wiretap channel with a public noiseless forward link.

Define Q as a random variable that is uniformly distributed over $\{1, 2, \ldots, n\}$. Define $X = X_Q$, $Y = Y_Q$, $Z = Z_Q$, $X_2 = X_{2,Q}$, $Y_2 = Y_{2,Q}$. Then, the right-hand side of (143) equals

$$I(X; Y|Z, X_2, Q) + I(X_2; Y_2, Z|X, Q) \tag{144}$$

$$\leq I(X; Y|Z, X_2) + I(X_2; Y_2, Z|X) \tag{145}$$

where we use the fact that $Y - \{Z, X_2, X\} - Q$ is a Markov chain and $\{Y_2, Z\} - \{X, X_2\} - Q$ is a Markov chain. Applying this result in (143) and letting $n \to \infty$, we obtained the upper bound in the theorem.

Appendix B
Proof of Theorem 3

Equation (50) follows from removing the eavesdropper and applying the bounds of two-way channel from [1]. Equation (51) can be derived similarly thanks to the symmetry of the channel model.

We next derive (52). We focus on the first term inside the minimum in (52). The second term can be derived similarly thanks to the symmetry of the channel model.

First we add a public noiseless broadcast channel to the channel in Fig. 1. The new channel model is shown in Fig. 10. The broadcast channel takes the input from Node 1. Its outputs are received by Node 2 and the eavesdropper. Since the channel is noiseless, the outputs equal the input, and is denoted by X_C. X_C is continuous. The introduction of the public noiseless broadcast channel certainly does not decrease the secrecy capacity region. Hence, to upper bound the secrecy capacity region of the original channel, we can consider this new model instead. We next apply Theorem 2 to this channel, which says R_1 is bounded by

$$I(X, X_C; Y, X_C|Z, X_C, X_2) + I(X_2; Y_2, Z, X_C|X, X_C). \tag{146}$$

The first term in (146) is upper bounded by

$$\begin{aligned}
&I(X, X_C; Y, X_C|Z, X_C, X_2) && (147)\\
&= I(X; Y|Z, X_C, X_2) && (148)\\
&= h(Y|Z, X_C, X_2) - h(Y|Z, X, X_C, X_2) && (149)\\
&\leq h(Y|Z, X_2) - h(Y|Z, X, X_C, X_2) && (150)\\
&= h(Y|Z, X_2) - h(Y|Z, X, X_2) && (151)\\
&= I(X; Y|Z, X_2). && (152)
\end{aligned}$$

In (151), we use the fact that $Y - \{Z, X, X_2\} - X_C$ is a Markov chain.

The second term is (146) is upper bounded by

$$\begin{aligned}
&I(X_2; Y_2, Z, X_C|X, X_C) && (153)\\
&= I(X_2; Y_2, Z|X, X_C) && (154)\\
&\leq h(Y_2, Z|X) - h(Y_2, Z|X, X_2, X_C) && (155)\\
&= h(Y_2, Z|X) - h(Y_2, Z|X, X_2) && (156)\\
&= I(X_2; Y_2, Z|X). && (157)
\end{aligned}$$

In (156), we use the fact that $\{Y_2, Z\} - \{X, X_2\} - X_C$ is a Markov chain.

Hence, (146) is upper bounded by

$$I(X; Y|Z, X_2) + I(X_2; Y_2, Z|X). \tag{158}$$

This means introducing a public noiseless forward channel brings no change in the expression of the upper bound of R_1.

We next prove (158) is also an upper bound on $R_1 + R_2$. This is done by showing if $R_1 = r_1$, $R_2 = r_2$ is achievable, then $R_1 = r_1 + r_2$ is also achievable.

Construct a message set $\{W_a\}$ which has the same cardinality of the message set $\{W_2\}$. Let part of the secret message be transmitted via W_a. The remaining part of the secret message is transmitted via W_1. The role of W_2 is to serve as a secret key. Let W_2 be taken from the set $\{W_2\}$ according to a uniform distribution. W_2 is independent from W_a and W_1.

Let $\oplus$ be the modulus addition defined over $\{1, \ldots |W_2|\}$, where $|W_2|$ is the cardinality of the set $\{W_2\}$. Recall that $\hat{W}_2$ denotes the result obtained by Node 1 when it tries to decode W_2. We let Node 1 transmit $\hat{W}_2 \oplus W_a$ over the public channel. Since the public channel is noiseless with continuous input, it can transmit $\hat{W}_2 \oplus W_a$ with a single channel use. Because Node 2 knows W_2, it can recover W_a from $\hat{W}_2 \oplus W_a$ when $W_2 = \hat{W}_2$.

The signal available to the eavesdropper now becomes the output of the wiretap channel Z^n, and the output of the public link, which is $W_a \oplus \hat{W}_2$. Conditioned on these signals, the equivocation of W_1, W_a can be computed as

$$\begin{aligned}
&H\left(W_1, W_a|Z^n, W_a \oplus \hat{W}_2\right) && (159)\\
&\geq H\left(W_1, W_a|Z^n, W_a \oplus \hat{W}_2, W_a \oplus W_2\right) && (160)\\
&= H\left(W_1, W_a, W_a \oplus \hat{W}_2|Z^n, W_a \oplus W_2\right)\\
&\quad - H\left(W_a \oplus \hat{W}_2|Z^n, W_a \oplus W_2\right) && (161)\\
&= H(W_1, W_a|Z^n, W_a \oplus W_2)\\
&\quad + H\left(W_a \oplus \hat{W}_2|W_1, W_a, Z^n, W_a \oplus W_2\right)\\
&\quad - H\left(W_a \oplus \hat{W}_2|Z^n, W_a \oplus W_2\right) && (162)\\
&\geq H(W_1, W_a|Z^n, W_a \oplus W_2)\\
&\quad - H\left(W_a \oplus \hat{W}_2|Z^n, W_a \oplus W_2\right) && (163)\\
&\geq H(W_1, W_a|Z^n, W_a \oplus W_2)\\
&\quad - H\left(W_a \oplus \hat{W}_2|W_a, Z^n, W_a \oplus W_2\right) && (164)\\
&= H(W_1, W_a|Z^n, W_a \oplus W_2) - H\left(\hat{W}_2|W_a, Z^n, W_2\right) && (165)\\
&\geq H(W_1, W_a|Z^n, W_a \oplus W_2) - H\left(\hat{W}_2|W_2\right) && (166)\\
&\geq H(W_1, W_a|Z^n, W_a \oplus W_2) - n\varepsilon. && (167)
\end{aligned}$$

In (167), we use the fact that W_2 can be reliably decoded by Node 1. Hence, (167) follows from Fano's inequality.

The first term in (167) can be bounded as follows:

$$
\begin{aligned}
&H\left(W_1, W_a | Z^n, W_a \oplus W_2\right) && (168)\\
&= H\left(W_a | Z^n, W_a \oplus W_2\right) + H\left(W_1 | Z^n, W_a, W_a \oplus W_2\right) && (169)\\
&= H\left(W_a | Z^n, W_a \oplus W_2\right) + H\left(W_1 | Z^n, W_a, W_2\right) && (170)\\
&= H\left(W_a | W_a \oplus W_2\right) + H\left(W_1 | Z^n, W_a, W_2\right) && (171)\\
&= H\left(W_a | W_a \oplus W_2\right) + H\left(W_1 | Z^n, W_2\right) && (172)\\
&= H\left(W_a\right) + H\left(W_1 | Z^n, W_2\right) && (173)\\
&\geq H\left(W_a\right) + H\left(W_1\right) - n\varepsilon && (174)\\
&\geq H\left(W_1, W_a\right) - n\varepsilon. && (175)
\end{aligned}
$$

Equation (171) is due to the fact that Z^n is independent from W_a, W_2, which leads to

$$
\begin{aligned}
&I\left(W_a; Z^n | W_a \oplus W_2\right) \leq I\left(W_a, W_a \oplus W_2; Z^n\right) && (176)\\
&= I\left(W_a, W_2; Z^n\right) = 0. && (177)
\end{aligned}
$$

Equation (172) follows from the fact that W_a is independent from Z^n, W_1, W_2. Equation (174) follows from the fact that collective secrecy implies one message is secure even if the other message is revealed to the eavesdropper [12].

The argument above shows the rate of W_1, W_a is the secrecy rate R_1. Since W_a is chosen from the message set $\{W_a\}$ according to a uniform distribution, we have $R_1 = r_1 + r_2$.

Therefore, $R_1 + R_2$ is upper bounded by (158).

Hence, we have proved the theorem.

Appendix C
Proof of Theorem 1

We prove $R_1 = R_1^*, R_2 = 0$ is achievable. The achievability of $R_1 = 0$, $R_2 = R_2^*$ can be proved similarly due to the symmetry of the channel model.

The communication is divided into two phases:

1) The first phase lasts n channel uses. During it, Node 2 sends a key K to Node 1. At the same time, Node 1 performs cooperative jamming by transmitting an i.i.d. Gaussian noise sequence with power P.
2) The second phase lasts $\bar{n}$ channel uses, during which Node 1 encrypts the confidential message W with K, and sends the result back to Node 2. At the same time, Node 2 performs cooperative jamming by transmitting an i.i.d. Gaussian noise sequence with power P_2.

Let $\alpha = n/(n+\bar{n})$ be the time sharing factor of the first phase. $0 \leq \alpha \leq 1$ and α is a constant.

The following notation is used in the remainder of the proof: $\bar{x}$ denotes any signal x which is related to the second phase. Otherwise, the signal is related to the first phase. With this notation, the signals received by the eavesdropper during the two phases are given by

$$
\begin{aligned}
Z^n &= \sqrt{h_1} X^n + \sqrt{h_2} {X_2}^n + N_2^n && (178)\\
\bar{Z}^{\bar{n}} &= \sqrt{h_1} \bar{X}^{\bar{n}} + \sqrt{h_1} {\bar{X}_2}^{\bar{n}} + \bar{N}_2^{\bar{n}}. && (179)
\end{aligned}
$$

The codebooks used by Nodes 1 and 2 are denoted by $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively, and are generated in the following way: $\mathcal{C}_2$ is composed of i.i.d. sequences sampled from the Gaussian distribution $\mathcal{N}(0, P_2)$. The codebook is then randomly binned into several bins. The size of the codebook depends on the number of bins needed to represent the key K and the size of the bin necessary to confuse the eavesdropper. Specifically, the size of the bin is chosen to be

$$2^{\lfloor n\left(C\left(\frac{h_2 P_2}{h_1 P + 1}\right) - \epsilon\right)\rfloor} \qquad (180)$$

where $\lfloor x \rfloor$ denotes the largest integer smaller or equal to x, $\epsilon > 0$ and $\lim_{n\to\infty} \epsilon = 0$.

Let R_K be the rate of the secret key. Then, there are 2^{nR_K} bins. R_K is given by

$$
\begin{aligned}
0 < R_K &= \frac{1}{n} H\left(K | \mathcal{C}_1, \mathcal{C}_2\right) && (181)\\
&< \min\left\{\left[C\left(P_2\right) - C\left(\frac{h_2 P_2}{h_1 P + 1}\right)\right]^+, C\left(\frac{h_1 P}{h_2 P_2 + 1}\right)\right\}. && (182)
\end{aligned}
$$

Observe that the key rate is chosen to be smaller than $\left[C\left(P_2\right) - C\left(\frac{h_2 P_2}{h_1 P + 1}\right)\right]^+$ to keep the key K secret from the eavesdropper. As will be shown later, the key is used to compensate the rate loss of the forward channel needed to confuse to eavesdropper. Hence, the rate of the key is chosen not to exceed this rate loss, which leads to the term $C\left(\frac{h_1 P}{h_2 P_2 + 1}\right)$ in (182).

$\mathcal{C}_1$ is composed of 2^{nR_K} codebooks. Each codebook is composed of i.i.d. sequences sampled from the Gaussian distribution $\mathcal{N}(0, P)$, and is composed of $2^{\bar{n}C(P)}$ i.i.d. Gaussian sequences. The sequences of each codebook are randomly binned into several bins. The size of each bin is chosen to be

$$2^{\lfloor\left(\bar{n}C\left(\frac{h_1 P}{h_2 P_2 + 1}\right) - nR_K - \bar{n}\epsilon_1\right)\rfloor} \qquad (183)$$

where $\epsilon_1 > 0$ and $\lim_{n\to\infty} \epsilon_1 = 0$.

During the first phase, Node 2 generates a secret key K according to a uniform distribution over $\{1, \ldots, 2^{nR_K}\}$ and selects the bin from $\mathcal{C}_2$ according to K. Then, it chooses a codeword from this bin according to a uniform distribution and transmits it to Node 1.

Since Node 1 is transmitting an i.i.d. Gaussian noise sequence during the first phase, the channel model in this phase is equivalent to the Gaussian wiretap channel [30], which uses the same codebook and encoding scheme as we do here. Reference [30] proves that, by doing so, K is kept secret from the eavesdropper and can be reliably decoded by Node 1. That is,

$$
\begin{aligned}
&\frac{1}{n} I\left(K; Z^n | \mathcal{C}_1, \mathcal{C}_2\right) \leq \varepsilon && (184)\\
&\lim_{n\to\infty} E[\Pr(\hat{K} \neq K | \mathcal{C}_1, \mathcal{C}_2)] = 0 && (185)
\end{aligned}
$$

where $\varepsilon \geq 0$, $\lim\limits_{n\to\infty} \varepsilon = 0$.

Let $\hat{K}$ be the estimate of K Node 1 decodes from its received signal. Node 1 computes its transmitted signals as follows: it first chooses the codebook according to the key $\hat{K}$ it decoded from the first phase. Then, it chooses the bin from the codebook according to the secret message W. Finally, it chooses the transmitted codeword from this bin according to a uniform distribution.

If $\hat{K} = K$, then Node 2 knows the subcodebook used by Node 1. The subcodebook is composed of i.i.d. Gaussian sequences and its rate is within the AWGN channel capacity between Nodes 1 and 2. This observation, along with (185), leads to the following fact:

$$\lim_{\bar{n}\to\infty} E[\Pr(\hat{W} \neq W|\mathcal{C}_1,\mathcal{C}_2)] = 0. \tag{186}$$

We next bound the equivocation

$$H\left(W|Z^n,\bar{Z}^{\bar{n}},\mathcal{C}_1,\mathcal{C}_2\right). \tag{187}$$

It is understood that $\mathcal{C}_1$, $\mathcal{C}_2$ is always on the condition term. Hence, we omit it in the sequel to simplify the notation and reinstate it only when necessary.

The equivocation rate is then bounded as follows:

$$\begin{aligned}
&H\left(W|Z^n,\bar{Z}^{\bar{n}}\right)\\
&=H\left(\bar{X}^{\bar{n}},W|Z^n,\bar{Z}^{\bar{n}}\right)-H\left(\bar{X}^{\bar{n}}|W,Z^n,\bar{Z}^{\bar{n}}\right) && (188)\\
&\geq H\left(\bar{X}^{\bar{n}},W|Z^n,\bar{Z}^{\bar{n}}\right)-\bar{n}\varepsilon && (189)\\
&=H\left(W|Z^n,\bar{Z}^{\bar{n}},\bar{X}^{\bar{n}}\right)+H\left(\bar{X}^{\bar{n}}|Z^n,\bar{Z}^{\bar{n}}\right)-\bar{n}\varepsilon && (190)\\
&=H\left(\bar{X}^{\bar{n}}|Z^n,\bar{Z}^{\bar{n}}\right)-\bar{n}\varepsilon && (191)\\
&=H\left(\bar{X}^{\bar{n}}|Z^n,\bar{Z}^{\bar{n}}\right)-H\left(\bar{X}^{\bar{n}}\right)+H\left(\bar{X}^{\bar{n}}\right)-\bar{n}\varepsilon && (192)\\
&=H\left(\bar{X}^{\bar{n}}\right)-I\left(\bar{X}^{\bar{n}};Z^n,\bar{Z}^{\bar{n}}\right)-\bar{n}\varepsilon && (193)\\
&=H\left(\bar{X}^{\bar{n}}\right)-I\left(\bar{X}^{\bar{n}};\bar{Z}^{\bar{n}}\right)-I\left(\bar{X}^{\bar{n}};Z^n|\bar{Z}^{\bar{n}}\right)-\bar{n}\varepsilon. && (194)
\end{aligned}$$

Here, (189) follows from the fact that given W, the number of possible $\bar{X}^{\bar{n}}$ equals the cardinality of the bin that corresponds to W from all the 2^{nR_K} codebooks, which is $2^{\bar{n}\left(C\left(\frac{h_1P}{h_2P_2+1}\right)-\epsilon_1\right)}$. Note that these candidates of $\bar{X}^{\bar{n}}$ form a Gaussian codebook by itself with a rate of $C\left(\frac{h_1P}{h_2P_2+1}\right)-\epsilon_1$. Since Node 2 is transmitting i.i.d. Gaussian noise, the channel between Node 1 and the eavesdropper is an AWGN channel whose capacity is $C\left(\frac{h_1P}{h_2P_2+1}\right)$. Therefore, given W, the eavesdropper can determine $\bar{X}^{\bar{n}}$ from $\bar{Z}^{\bar{n}}$ using joint typical decoding. Equation (189) then follows by applying Fano's inequality.

Equation (191) follows since W is a deterministic function of $\bar{X}^{\bar{n}}$.

The third term in (194) can then be bounded as follows:

$$\begin{aligned}
&I\left(\bar{X}^{\bar{n}};Z^n|\bar{Z}^{\bar{n}}\right) && (195)\\
&=h\left(Z^n|\bar{Z}^{\bar{n}}\right)-h\left(Z^n|\bar{Z}^{\bar{n}},\bar{X}^{\bar{n}}\right) && (196)\\
&=h\left(Z^n|\bar{Z}^{\bar{n}}\right)-h\left(Z^n|\bar{X}_f^{\bar{n}}+\bar{N}_2^{\bar{n}},\bar{X}^{\bar{n}}\right) && (197)\\
&=h\left(Z^n|\bar{Z}^{\bar{n}}\right)-h\left(Z^n|\bar{X}^{\bar{n}}\right) && (198)\\
&\leq h\left(Z^n|\bar{Z}^{\bar{n}}\right)-h\left(Z^n|\bar{X}^{\bar{n}},K\right) && (199)\\
&=h\left(Z^n|\bar{Z}^{\bar{n}}\right)-h\left(Z^n|K\right) && (200)\\
&=h\left(Z^n|\bar{Z}^{\bar{n}}\right)-h\left(Z^n\right)-h\left(Z^n|K\right)+h\left(Z^n\right) && (201)\\
&=I\left(Z^n;K\right)-I\left(Z^n;\bar{Z}^{\bar{n}}\right) && (202)\\
&\leq I\left(Z^n;K\right)\leq n\varepsilon_2. && (203)
\end{aligned}$$

Equation (198) is because $\bar{X}_2^{\bar{n}}+\bar{N}_2^{\bar{n}}$ is a sequence of i.i.d. Gaussian noise, which is independent from Z^n and $\bar{X}^{\bar{n}}$.

Equation (200) follows from the fact that $Z^n - K - \bar{X}^{\bar{n}}$ is a Markov chain. Equation (203) follows from (184).

Substituting (203) into (194), we have

$$\begin{aligned}
&H\left(W|Z^n,\bar{Z}^{\bar{n}}\right) && (204)\\
&\geq H\left(\bar{X}^{\bar{n}}\right)-I\left(\bar{X}^{\bar{n}};\bar{Z}^{\bar{n}}\right)-\left(\bar{n}\varepsilon+n\varepsilon_2\right). && (205)
\end{aligned}$$

The second term in (205) can be bounded as follows. For this purpose, we reinstate the $\mathcal{C}_1,\mathcal{C}_2$ on the condition term

$$\begin{aligned}
&I\left(\bar{X}^{\bar{n}};\bar{Z}^{\bar{n}}|\mathcal{C}_1,\mathcal{C}_2\right) && (206)\\
&\leq h\left(\bar{Z}^{\bar{n}}\right)-h\left(\bar{Z}^{\bar{n}}|\bar{X}^{\bar{n}},\mathcal{C}_1,\mathcal{C}_2\right) && (207)\\
&=h\left(\bar{Z}^{\bar{n}}\right)-h\left(\bar{Z}^{\bar{n}}|\bar{X}^{\bar{n}},\mathcal{C}_1,\mathcal{C}_2\right) && (208)\\
&=h\left(\bar{Z}^{\bar{n}}\right)-h\left(\bar{Z}^{\bar{n}}|\bar{X}^{\bar{n}}\right) && (209)\\
&=I\left(\bar{X}^{\bar{n}};\bar{Z}^{\bar{n}}\right) && (210)\\
&=\bar{n}I\left(\bar{X};\bar{Z}\right). && (211)
\end{aligned}$$

Equation (209) follows from the fact that given $\bar{X}^{\bar{n}}$, $\bar{Z}^{\bar{n}}$ only depends on the jamming signal and channel noise. Therefore, we can drop codebooks $\mathcal{C}_1,\mathcal{C}_2$ from the conditioning term. Equation (211) follows from the fact that Node 2 transmits i.i.d. Gaussian noise during the second phase, and the codebook used by Node 1 is composed of i.i.d. Gaussian sequences.

Since

$$I\left(\bar{X};\bar{Z}\right)=C\left(\frac{h_1P}{h_2P_2+1}\right) \tag{212}$$

$$H\left(\bar{X}^{\bar{n}}|\mathcal{C}_1,\mathcal{C}_2\right)=nR_K+\bar{n}C\left(P\right), \tag{213}$$

we have

$$\begin{aligned}
&H\left(W|Z^n,\bar{Z}^{\bar{n}},\mathcal{C}_1,\mathcal{C}_2\right) && (214)\\
&=\left(nR_K+\bar{n}C\left(P\right)\right)-\bar{n}C\left(\frac{h_1P}{h_2P_2+1}\right)-\left(\bar{n}\varepsilon+n\varepsilon_2\right) && (215)\\
&\geq H(W|\mathcal{C}_1,\mathcal{C}_2)-\left(\bar{n}(\varepsilon+\epsilon_1)+n\varepsilon_2\right). && (216)
\end{aligned}$$

Therefore, $0 \leq I(W;Z^n,\bar{Z}^{\bar{n}}|\mathcal{C}_1,\mathcal{C}_2) < (\bar{n}(\varepsilon+\epsilon_1)+n\varepsilon_2)$. This, along with (186), gives us

$$\begin{aligned}
&\lim_{n,\bar{n}\to\infty}\frac{1}{n+\bar{n}}I(W;Z^n,\bar{Z}^{\bar{n}}|\mathcal{C}_1,\mathcal{C}_2)+E[\Pr(\hat{W}\neq W)|\mathcal{C}_1,\mathcal{C}_2]\\
&=0. && (217)
\end{aligned}$$

From the linearity of expectation and nonnegativity of mutual information and probability, we see that there must exists codebooks $\mathcal{C}_1=\mathcal{C}_1^*$, $\mathcal{C}_2=\mathcal{C}_2^*$ such that both terms on the left-hand side of (217) go to 0 as $n,\bar{n}\to\infty$. This observation, along with that fact that $n+\bar{n}$ channel uses are involved, proves that the secrecy rate pair $(R_1^*,0)$ is achievable.

Hence, we have proved the theorem.

APPENDIX D
PROOF OF THEOREM 7

Since received signals are not used to compute transmitting signals at Node j, $j=1,2$, we let $\alpha=1$ in Theorem 1. In this case, when $P=kP_2$, R_j^* becomes

$$R_1^*=C\left(P\right)-C\left(\frac{h_1}{h_2k+1/P}\right) \tag{218}$$

$$R_2^*=C\left(kP\right)-C\left(\frac{h_2k}{h_1+1/P}\right). \tag{219}$$

The sum rate bound given by Theorem 6 is upper bounded by

$$\min\left\{\begin{array}{l}C\left(\frac{P}{h_1P+1}\right)+C\left(\left(h_2+1\right)kP\right),\\ C\left(\frac{kP}{h_2kP+1}\right)+C\left(\left(h_1+1\right)P\right)\end{array}\right\}. \tag{220}$$

To prove Theorem 7, it is sufficient to show both R_1^* and R_2^* are within constant gaps of (220), as we show below.

$$C\left(\frac{P}{h_1P+1}\right)+C\left(\left(h_2+1\right)kP\right)-R_1^* \tag{221}$$

$$=C\left(\frac{P}{h_1P+1}\right)+C\left(\left(h_2+1\right)kP\right)-C\left(P\right)$$
$$+C\left(\frac{h_1}{h_2k+1/P}\right) \tag{222}$$

$$\leq C\left(\frac{P}{h_1P+1}\right)+C\left(\left(h_2+1\right)kP\right)-C\left(P\right)+C\left(\frac{h_1}{h_2k}\right) \tag{223}$$

$$\leq C\left(\frac{1}{h_1}\right)+C\left(\left(h_2+1\right)kP\right)-C\left(P\right)+C\left(\frac{h_1}{h_2k}\right) \tag{224}$$

$$=C\left(\frac{1}{h_1}\right)+\frac{1}{2}\log_2\left(\frac{1+\left(h_2+1\right)kP}{1+P}\right)+C\left(\frac{h_1}{h_2k}\right) \tag{225}$$

$$\leq C\left(\frac{1}{h_1}\right)+\frac{1}{2}\log_2\left(\frac{1+\max\left\{1,\left(h_2+1\right)k\right\}P}{1+P}\right)$$
$$+C\left(\frac{h_1}{h_2k}\right) \tag{226}$$

$$\leq C\left(\frac{1}{h_1}\right)+\frac{1}{2}\log_2\left(\max\left\{1,\left(h_2+1\right)k\right\}\right)+C\left(\frac{h_1}{h_2k}\right). \tag{227}$$

For R_2^*, we have

$$C\left(\frac{P}{h_1P+1}\right)+C\left(\left(h_2+1\right)kP\right)-R_2^* \tag{228}$$

$$=C\left(\frac{P}{h_1P+1}\right)+C\left(\left(h_2+1\right)kP\right)-C\left(kP\right)$$
$$+C\left(\frac{h_2k}{h_1+1/P}\right) \tag{229}$$

$$\leq C\left(\frac{P}{h_1P+1}\right)+C\left(\left(h_2+1\right)kP\right)-C\left(kP\right)+C\left(\frac{h_2k}{h_1}\right) \tag{230}$$

$$\leq C\left(\frac{1}{h_1}\right)+C\left(\left(h_2+1\right)kP\right)-C\left(kP\right)+C\left(\frac{h_2k}{h_1}\right) \tag{231}$$

$$=C\left(\frac{1}{h_1}\right)+\frac{1}{2}\log_2\left(\frac{1+\left(h_2+1\right)kP}{1+kP}\right)+C\left(\frac{h_2k}{h_1}\right) \tag{232}$$

$$\leq C\left(\frac{1}{h_1}\right)+\frac{1}{2}\log_2\left(h_2+1\right)+C\left(\frac{h_2k}{h_1}\right). \tag{233}$$

Hence, we have proved the theorem.

Appendix E
Proof of Theorem 8

To prove this theorem, we only need show that it is possible to achieve a secrecy rate for Node 1 that exceeds the upper bound given by Theorem 5. Consider the case when $h_1 = h_2 = 1$. Then, by evaluating (60) at $\sigma^2 = 0$ and $\sigma^2 \to \infty$ with $\rho = \eta = 0$, we find the secrecy rate R_1 is bounded by

$$\min\{C(P), C(P_2)+0.5\} \tag{234}$$

when Y_f is ignored by Node 1. Choose P_2 and P such that

$$C(P_2)+0.5 < 0.4C(P). \tag{235}$$

For this power configuration, from (234), we observe that R_1 is upper bounded by $0.4C(P)$.

Let the α in Theorem 1 be 0.5. R_1^* then becomes

$$0.5C\left(P\right)-0.5\left[C\left(\frac{P}{P_2+1}\right)-C\left(P_2\right)+C\left(\frac{P_2}{P+1}\right)\right]^+. \tag{236}$$

A sufficient condition for $R_1^* = 0.5C(P)$ is that

$$C\left(\frac{P}{P_2+1}\right)+C\left(\frac{P_2}{P+1}\right)>C\left(P_2\right). \tag{237}$$

It can be verified that this condition is equivalent to

$$\frac{\left(\frac{P}{P_2+1}+1\right)^2}{P+1}>1. \tag{238}$$

A sufficient condition for it to hold is

$$\frac{\left(\frac{P}{P_2+1}+1\right)^2}{\left(\sqrt{P}+1\right)^2}>1 \tag{239}$$

which means

$$\sqrt{P}>P_2+1. \tag{240}$$

Choose $P_2 = P^{1/4}$. For sufficiently large P, both (235) and (240) can be fulfilled. In this case, the achievable rate is $0.5C(P)$, which is greater than the upper bound $0.4C(P)$. The difference is $0.1C(P)$, which is not a bounded function of P. Hence, we have proved the theorem.

References

[1] C. E. Shannon, “Two-way communication channels,” in *Proc. 4th Berkeley Symp. Math. Stat. Probab.*, 1961, vol. 1, pp. 351–384.

[2] T. Han, “A general coding scheme for the two-way channel,” *IEEE Trans. Inf. Theory*, vol. 30, no. 1, pp. 35–44, Jan. 1984.

[3] G. Dueck, “The capacity region of the two-way channel can exceed the inner bound,” *Inf. Control.*, vol. 40, no. 3, pp. 258–266, 1979.

[4] C. E. Shannon, “Communication theory of secrecy systems,” *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Sep. 1949.

[5] E. Ardetsanizadeh, M. Franceschetti, T. Javidi, and Y. H. Kim, “Wiretap channel with secure rate-limited feedback,” *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5353–5361, Dec. 2009.

[6] U. M. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[7] G. T. Amariucai and S. Wei, “Strictly positive secrecy rates of binary wiretapper channels using feedback schemes,” in *Proc. 42nd Annu. Conf. Inf. Sci. Syst.*, Mar. 2008, pp. 624–629.

[8] D. Gunduz, D. R. Brown, III, and H. V. Poor, “Secure communication with feedback,” in *Proc. IEEE Int. Symp. Inf. Theory Appl.*, Dec. 2008.

[9] R. Ahlswede and N. Cai, “Transmission, Identification and Common Randomness Capacities for Wire-Tape Channels With Secure Feedback From the Decoder,” in *General Theory of Information Transfer and Combinatorics*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer, 2006, vol. 4123, pp. 258–275.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley-Interscience, 2006.

[11] T. M. Cover and C. Leung, “An achievable rate region for the multiple-access channel with feedback,” *IEEE Trans. Inf. Theory*, vol. IT-27, no. 3, pp. 292–298, May 1981.

[12] E. Tekin and A. Yener, “The general Gaussian multiple access and two-way wire-tap channels: Achievable rates and cooperative jamming,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[13] L. Lima, M. Medard, and J. Barros, “Random linear network coding: A free cipher?,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 546–550.

[14] Y. Oohama, “Relay channels with confidential messages,” Mar. 2007 [Online]. Available: http://arxiv.org/abs/cs/0611125

[15] X. He and A. Yener, "Cooperation with an untrusted relay: A secrecy perspective," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3801–3827, Aug. 2010.

[16] E. Ekrem and S. Ulukus, "Secrecy in cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 137–155, Jan. 2011.

[17] X. He and A. Yener, "Two-hop secure communication using an untrusted relay," *Eurasip J. Wireless Commun. Network., Spec. Issue Wireless Phys. Layer Security*, no. 305146, pp. 13–13, Mar. 2009.

[18] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography—Part I. Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[19] A. Khisti, S. Diggavi, and G. Wornell, "Secret-key generation using correlated sources and channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 652–670, Feb. 2012.

[20] V. M. Prabhakaran, K. Eswaran, and K. Ramchandran, "Secrecy via sources and channels—A secret key—Secret message rate tradeoff region," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 1010–1014.

[21] A. A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals—Part II: Channel model," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3997–4010, Aug. 2010.

[22] I. Csiszár and P. Narayan, "Secrecy generation for multiple input multiple output channel models," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2009, pp. 2447–2451.

[23] E. Tekin and A. Yener, "Achievable rates for two-way wire-tap channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 941–945.

[24] M. Bloch and A. Thangraj, "Confidential messages to a cooperative relay," in *Proc. IEEE Inf. Theory Workshop*, May 2008, pp. 154–158.

[25] M. Bloch, "Channel scrambling for secrecy," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2009, pp. 2452–2456.

[26] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[27] A. Khisti and G. Wornell, "Secure transmission with multiple antennas-I: The MISOME Wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[28] X. He and A. Yener, "On the role of feedback in two way secure communication," in *Proc. 42nd Annu. Asilomar Conf. Signals, Systems, Computing*, Oct. 2008, pp. 1093–1097.

[29] X. He, "Cooperation and information theoretic security in wireless networks," Ph.D. dissertation, The Pennsylvania State University, University Park, PA, USA, 2010.

[30] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, Jul. 1978.

Xiang He (S'08–M'10) received B.S. and M.S. degrees in Electrical Engineering from Shanghai Jiao Tong University, Shanghai, China in 2003 and 2006 respectively. His master's study was about high speed FPGA implementation of channel encoder, decoder and MIMO detectors. He received his Ph.D. degree in 2010 from the Department of Electrical Engineering at the Pennsylvania State University and joined Microsoft in that year. In 2010, he received Melvin P. Bloom Memorial Outstanding Doctoral Research Award from the Department of Electrical Engineering at the Pennsylvania State University and the best paper award from the Communication Theory Symposium in IEEE International Conference on Communications (ICC). In 2011, he was named as one of the exemplary reviewers by IEEE Communication Letters. His research interests include information theoretic secrecy, coding theory, queuing theory, optimization techniques, distributed detection and estimation.

Aylin Yener (S'91–M'00) received the B.Sc. degree in electrical and electronics engineering, and the B.Sc. degree in physics, from Boğaziçi University, Istanbul, Turkey; and the M.S. and Ph.D. degrees in electrical and computer engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, New Brunswick, NJ. Commencing fall 2000, for three semesters, she was a P. C. Rossin Assistant Professor at the Electrical Engineering and Computer Science Department, Lehigh University, PA. In 2002, she joined the faculty of The Pennsylvania State University, University Park, PA, where she was an Assistant Professor, then Associate Professor, and is currently Professor of Electrical Engineering since 2010. During the academic year 2008–2009, she was a Visiting Associate Professor with the Department of Electrical Engineering, Stanford University, CA. Her research interests are in information theory, communication theory and network science, with recent emphasis on green communications and information security. She received the NSF CAREER award in 2003.

Dr. Yener previously served as a technical program chair or co-chair for various conferences for the IEEE Communications Society, as an associate editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, as an associate editor and an editorial advisory board member for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She served as the student committee chair for the IEEE Information Theory Society 2007–2011, and was the co-founder of the Annual School of Information Theory in North America co-organizing the school in 2008, 2009 and 2010. Dr. Yener currently serves on the board of governors of the IEEE Information Theory Society as its treasurer.