On Cooperative Wireless Network Secrecy

EPFL Technical Report LICOS-REPORT-2008-009

Etienne Perron, Suhas Diggavi, Emre Telatar
EPFL, Lausanne, Switzerland
Email: {etienne.perron,suhas.diggavi,emre.telatar}@epfl.ch

Abstract—**Given that wireless communication occurs in a shared and inherently broadcast medium, the transmissions are vulnerable to undesired eavesdropping. This occurs even when a point-to-point communication is sought, and hence a fundamental question is whether we can utilize the wireless channel properties to establish secrecy. In this paper we consider secret communication between two special nodes ("source" and "destination") in a wireless network with authenticated relays: the message communicated to the destination is to be kept information-theoretically (unconditionally) secret from any eavesdropper within a class. Since the transmissions are broadcast and interfere with each other, complex signal interactions occur. We develop cooperative schemes which utilize these interactions in wireless communication over networks with arbitrary topology, and give provable unconditional secrecy guarantees.**

I. Introduction

A fundamental aspect of wireless communication is its broadcast nature, *i.e.,* transmission from a node can be overheard (albeit through different channels) at several locations. This property makes wireless communication inherently vulnerable to eavesdropping by an adversary. As the use of wireless networks grows, this would be an important concern to be addressed. This issue has been identified by recent discoveries that the wireless 802.11 networks are vulnerable to eavesdropping [4]. Therefore, a fundamental question is how to ensure secrecy in wireless communication.

To ask this question we first need to address the notion of secrecy that we seek. In 1948, Shannon introduced the notion of information theoretic secrecy [15], in which he studied whether communication from a source to a destination can be kept secret from an eavesdropper, who has *complete* access to the transmission, without any assumptions on computational capabilities. As one would expect, such a question resulted in the pessimistic answer that unless the source and destination had somehow a large amount of shared common randomness (key) kept secret from the eavesdropper, the task was impossible. In fact, the shared common randomness needed was essentially of the same rate as the source message itself, making it completely impractical. This observation led to the *computational* approach pioneered by Diffie and Hellman [8], where instead of having such a long shared secret key, a shorter key is exploded into a larger one. The goal of the design is to guarantee that there is no efficient algorithm for the eavesdropper to discover the information transmitted. For example, the security of the well-known RSA public-key cryptosystem [14] is based on the difficulty of factoring large integers, and other cryptographic protocols are based on the difficulty of computing discrete logarithms over groups. Both these protocols are based on the (as of yet) unproven computational intractability hypothesis for these algorithmic problems.

Clearly the secrecy of any system can be enhanced if one could have even a small amount of shared key between the source–destination pair, which can be kept unconditionally secret from an eavesdropper[1]. In particular one of our motivations for the formulation in this paper is that we can potentially generate such secret key using physical wireless channel properties. Such a functionality can be incrementally deployed in networks by passing this secret key to the higher layers in the network protocol stack where it could then be used to enhance secrecy. In this respect, the two approaches to secrecy can be complementary, with the information-theoretic secrecy used to provide further secrecy opportunities. Of course, in order to make this more practical, we need to utilize some distinguishing property between the destination and the eavesdropper in order to provide secrecy. In this paper, we study this problem for a wireless network, where a source node is transmitting to a destination node, with the help of (authenticated) relay nodes, when an unknown (passive) eavesdropper is also present in the network. We call this setup *cooperative secrecy*, since the (authenticated) nodes in the network cooperate to provide secrecy against potential eavesdroppers.

In wireless communication, even though the signal from the source is broadcast, it is received at the destination and the potential eavesdropper through *different* (fading) channels. It is this distinction that is exploited in information-theoretic secrecy for broadcast channels in the seminal work of wire-tap channels [20], [6]. However, not much is known about the cooperative secrecy setup, where there are relay nodes facilitating secure communication between a source and a destination[2]. To the best of our knowledge, our work here is the first to examine this problem for an *arbitrary* wireless network and provide provable secrecy guarantees. The main difficulties in dealing with arbitrary relay networks are (i) the

[1] Another motivation to study information-theoretic (or unconditional) secrecy might be that theoretically, quantum computers could make difficult algorithmic problems tractable [16].

[2] Notable exceptions are some recent studies of techniques when there is a *single* relay node present, as an extension of the classical relay channel to the secrecy problem [11], [10].

broadcast nature of wireless communications, (ii) the fact that signals from simultaneously transmitting nodes interfere with one another at other nodes. These give rise to complex signal interactions making the understanding of wireless networks difficult. Though there has been some recent understanding in terms of scaling laws for asymptotically large wireless networks [9], [21], [12], there has not been a complete understanding of communication for the non-asymptotic regime. Our work develops on the recent understanding of wireless network information flow in [2], [3], which (approximately) characterized the unicast (or multicast) capacity without secrecy constraints. Our formulation asks a natural question of additionally keeping such a cooperative communication secure against a class of potential eavesdroppers.

The main result of our paper is an achievable trade-off between the reliable transmission rate from the source to the legitimate destination and the amount of information leaked to an eavesdropper over an *arbitrary* wireless relay network. We assume that the exact location of the eavesdropper is unknown, and therefore there is a *class* of potential eavesdroppers. Roughly speaking, the trade-off is related to the information-theoretic min-cuts[3] between the source-destination and source-eavesdropper pairs. In particular, one extreme point of the trade-off is that we can ensure perfect secrecy (zero rate of information leakage to the eavesdropper) for an information rate of (approximately) the "difference" between these two min-cut values. The other extreme point is to transmit information to the legitimate destination at its min-cut while leaking an information rate related to the "difference" between these two min-cut values to the eavesdropper. The precise statements and implications of our results are given in Section IV. We also propose a noise insertion strategy by the authenticated nodes in the network, which help the legitimate receiver for secure communication. We illustrate many of the ideas in this paper through a *deterministic* model for wireless networks proposed in [1], [2]. The main results are proved for arbitrary wireless networks for both deterministic as well as noisy signal interaction models. These models are described more precisely in Section II.

The paper is organized as follows. We give the precise problem statement and motivate the wireless network models studied in this paper in Section II. We give a series of examples in Section III to illustrate some of the ideas in this paper. We state the main results in Section IV and discuss some of their implications. We sketch the overall proof in Section V, while giving the proof details in the appendix. Apart from this, the appendix also contains an extension of Theorem 1 and a corresponding numerical computation. We conclude in Section VI with a discussion about possible extensions and open questions raised by this work.

II. Problem Statement

We consider transmission over a wireless relay network $\mathcal{G} = (\mathcal{V}, \mathcal{L})$, where $\mathcal{V}$ is the set of vertices representing the communication nodes in the relay network and $\mathcal{L}$ is the set of annotated channels between the nodes, which describe the signal interactions. Note that these channels are not point-to-point links, rather, they model how the transmitted signals are superimposed and received at the receiving nodes (*i.e.,* there is broadcast and interference). We consider a special node $S \in \mathcal{V}$ as the source of the message which wants to securely communicate to another special node $D \in \mathcal{V}$ (the destination) with the help of a set of (authenticated) relay nodes $\mathcal{A} \subset \mathcal{V}$ in the network. The secrecy is with respect to a set of possible (passive) eavesdropper nodes $\mathcal{E} \subset \mathcal{V}$ where $\mathcal{E}$ is disjoint from $\mathcal{A} \cup \{S, D\}$. We want to keep all or part of the message secret if any one of the possible eavesdropper nodes $E \in \mathcal{E}$ listens to the wireless transmissions in the relay network[4].

Wireless interaction model: In this well-accepted model [18], transmitted signals get attenuated by (complex) gains to which independent (Gaussian) receiver noise is added. More formally, the received signal y_j at node $j \in \mathcal{V}$ at time t is given by,

$$y_j[t] = \sum_{i \in \mathcal{N}_j} h_{ij} x_i[t] + z_j[t], \tag{1}$$

where h_{ij} is the complex channel gain between node i and j which is the annotation of the channels in $\mathcal{L}$, x_i is the signal transmitted by node i, and $\mathcal{N}_j$ are the set of nodes that have non-zero channel gains to j. We assume that the average transmit power constraints for all nodes is 1 and the additive receiver Gaussian noise is of unit variance. We use the terminology *Gaussian wireless network* when the signal interaction model is governed by (1).

Deterministic interaction model: In [1], a simpler deterministic model which captures the essence of wireless interaction was developed. The advantage of this model is its simplicity, which gives insight to strategies for the noisy wireless network model in (1). We will utilize this model in Section III to develop intuition for the wireless network secrecy problem, by giving illustrative examples for this model. Our main results are developed for both this deterministic model as well as the model in (1). The deterministic model of [1] simplifies the wireless interaction model in (1) by eliminating the noise and discretizing the channel gains through a binary expansion of q bits. Therefore, the received signal $\mathbf{y}_j^{(d)}$ which is a binary vector of size q is modeled as

$$\mathbf{y}_j^{(d)}[t] = \sum_{i \in \mathcal{N}_j} \mathbf{G}_{ij} \mathbf{x}_i^{(d)}[t], \tag{2}$$

where $\mathbf{G}_{ij}$ is a $q \times q$ binary matrix representing the (discretized) channel transformation between nodes i and j and $\mathbf{x}_i^{(d)}$ is the (discretized) transmitted signal. All operations in (2) are done over the binary field. We use the terminology *deterministic wireless network* when the signal interaction model is governed by (2). An illustration of this deterministic

[3] The information-theoretic min-cut is related to the cooperative information transfer rate [17], [3].

[4] Our results apply also when *all* eavesdroppers in $\mathcal{E}$ are present, but cannot collude since they are not allowed to transmit anything. In this paper we restrict ourselves to the passive eavesdropper model, motivated by possible secure authentication protocols [19] which could potentially discover an active eavesdropper.

model is given in Figure 1 for the broadcast and multiple access networks. Figure 1(a) shows a deterministic model of the broadcast channel, where the channel from the transmitter to Receiver 1 is stronger than that to Receiver 2. This is represented by the deterministic model developed in [1] with 5 most significant bits (MSB) of the transmitted signal captured by Rx 1 and only 2 MSB of the transmitted signal captured by Rx 2. The deterministic model of the multiple access channel shown in Figure 1(b) adds one more ingredient, which is how the bits from two transmitting nodes interact at a receiver. In Figure 1(b) the channel from Tx 1 to Rx is stronger than that of Tx 2. Therefore, the interaction is between the 2 MSBs of Tx 2 with the lower significant bits of Tx 1, and the interaction is modeled with an addition over the binary field (*i.e.,* `xor`). This interaction captures the dynamic range of the signal interactions. It was shown in [1], that this model *approximately*[5] captures the wireless interaction model of (1) for the broadcast and multiple access channels. For general networks the deterministic model yields insights which, when translated to the noisy wireless network, lead one to develop cooperative strategies for the model in (1), which are (provably) approximately[6] optimal [3].

Given the relay network with the above signal interaction models given in (1) and (2), we want to ensure that we can communicate reliably *and* secretly between S and D. The notion of reliability is the standard information-theoretic notion that the destination can decode the source message of rate R with arbitrarily small probability of error. More formally:

Definition 1: A (T, ϵ)-code is given by a (possibly) probabilistic source encoding function f_S mapping W to $\mathbf{X}_S$, a set of deterministic relay encoding functions f_i, mapping $\mathbf{Y}_i$ to $\mathbf{X}_i$ at each relay $i \in \mathcal{A}$, and a deterministic decoding function f_D, mapping $\mathbf{Y}_D$ into an estimate $\hat{W}$ of the message. Here, $\mathbf{X}_i = (x_i[1], \ldots, x_i[T])$ and $\mathbf{Y}_i$ are blocks of T symbols (each symbol being a real or complex number or a binary vector), and W is the information message, which is uniformly distributed in the set $\{1, \ldots, 2^{TR}\}$. The quantity R is the rate of the code. The probability of error of a (T, ϵ)-code is required to be bounded by ϵ: $\mathbf{P}(W \neq \hat{W}) < \epsilon$.

The notion of information-theoretic secrecy is defined through the *equivocation* rate R_e, which is the residual uncertainty about the message when the observation of the strongest eavesdropper is given. More formally, [20], [6]:

Definition 2: Given a (T, ϵ)-code, the equivocation rate is

$$\frac{1}{T} \min_{E \in \mathcal{E}} H(W|\mathbf{Y}_E), \tag{3}$$

where W is the uniformly distributed source message, $\mathbf{Y}_E$ is the sequence of observations at eavesdropper E and $H(\cdot|\cdot)$ denotes the (conditional) entropy [17].

In this paper, we also refer to the equivocation rate as simply the "equivocation".

Definition 3: A rate-equivocation pair (R, R_e) is called "achievable" if for any $\epsilon > 0$, there exists a blocklength T and a (T, ϵ)-code of rate R and equivocation R_e.

By "perfect secrecy", we mean a situation where a rate-equivocation pair (R, R_e) such that $R = R_e$ is achievable. If we are only interested in such cases, we can use the following quantity:

Definition 4: The perfect secrecy rate R_s of a network is defined as

$$R_s \triangleq \max\{R : (R, R) \text{ is achievable}\}.$$

III. Examples

Through the examples in this section we seek to illustrate the following ideas: (i) The operation of the relays to set up cooperation; (ii) A source coding scheme to ensure secrecy; (iii) Handling of a class of potential eavesdroppers; (iv) A noise insertion/jamming scheme by relays to help secrecy. All four examples are based on the deterministic wireless network model developed in [1]. Though the examples illustrate these ideas, proving the performance of a scheme for general networks in IV needs methods that are technically more difficult. These are sketched in Section V and in the appendix .

A. Deterministic examples

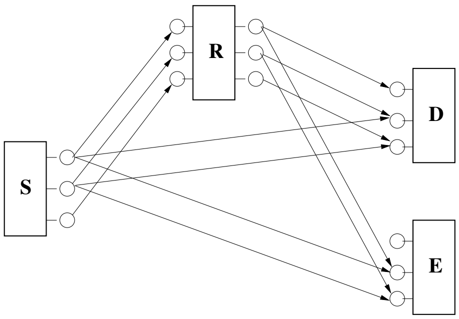


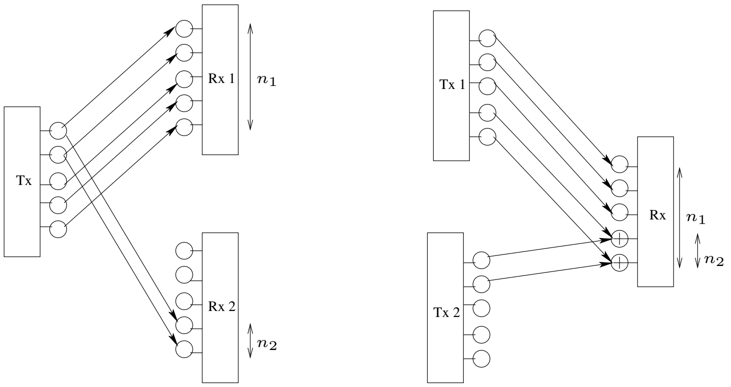
Fig. 2. Example 1: The relay network with a single helper node providing secrecy against a single eavesdropper.

Example 1: Consider the deterministic wireless network in Figure 2 with a source S, a relay R and destination D which wants secrecy from an eavesdropper E[7]. It is clear that the maximum communication rate (with no concern for secrecy) between the source and the destination is 3 bits. As the channel from the source to the eavesdropper E is as strong as that to the destination, we cannot ensure any secrecy in the absence of a relay [6]. Therefore, the cooperation of the relay is crucial to ensure secrecy. Suppose the source transmits bits $(W_1[t], W_2[t], W_3[t])$ at time t. Let the relay transmit bits $(W_2[t-1], W_1[t-1], W_3[t-1])$ via its outputs,

[5] The approximation is in the sense that the capacity region of the deterministic model is within 1 bit of the capacity region of the Gaussian counterparts.

[6] It has been shown for single unicast there is an *approximate* max-flow, min-cut result where the difference is within a constant number of bits, which depends on the topology of the network, but not the values of the channel gains [3].

[7] This is a replacement for a similar example that existed in a previous version of this technical report, and that did not clearly illustrate the use of a relay, since it implicitly assumed that there was no delay in the operation of the relay R.



(a) Broadcast channel (b) Multiple access channel

Fig. 1. The linear deterministic model for a Gaussian broadcast channel (BC) is shown in (a) and for a Gaussian multiple access channel (MAC) is shown in (b).

where the symbols are delayed because the relay needs to first hear the incoming signal before transmitting. The mixing at the destination (and at the eavesdropper) is therefore of bits transmitted at different times by the source. This is the reason for the time index notion in this example. Assume that the transmission takes place over $T+1$ timeslots, but S only transmits during the first T timeslots and therefore sends $3T$ bits over $T+1$ timeslots, resulting in a rate of $\frac{3T}{T+1} \to 3$ bits per timeslot. R transmits information during timeslots 2 through $T+1$. At the end of timeslot t, we have the received signals

$$y_D[t] = \begin{bmatrix} W_2[t-1] \\ W_1[t-1] \oplus W_1[t] \\ W_3[t-1] \oplus W_2[t] \end{bmatrix}$$

and

$$y_E[t] = \begin{bmatrix} 0 \\ W_2[t-1] \oplus W_1[t] \\ W_1[t-1] \oplus W_2[t] \end{bmatrix}.$$

At time 1, R does not transmit anything yet, and the destination receives $(0, W_1[1], W_2[1])$. At time 2, D receives $(W_2[1], W_1[1] \oplus W_1[2], W_3[1] \oplus W_2[2])$. At time 3, D receives $(W_2[2], W_1[2] \oplus W_1[3], W_3[2] \oplus W_2[3])$, and it has now enough linearly independent equations to decode $W_1[1]$, $W_1[2]$, $W_1[3]$, $W_2[1]$, $W_2[2]$ and $W_3[1]$. In general, at timeslot $t \in \{3, \ldots, T\}$, D can decode $(W_1[1], \ldots, W_1[t])$, $(W_2[1], \ldots, W_2[t-1])$ and $(W_3[1], \ldots, W_3[t-2])$. At timeslot $T+1$, the source S remains silent. Hence, D receives $(W_2[T], W_1[T], W_3[T])$. It can use $W_2[T]$ to decode $W_3[T-1]$ which was still unknown, yielding full knowledge of all transmitted bits. E receives $(0, W_1[1], W_2[t])$ during timeslot 1, and $(0, W_2[1] \oplus W_1[2], W_1[1] \oplus W_2[2])$ during timeslot 2, which allows E to decode $W_1[1], W_1[2], W_2[1], W_2[2]$ at the end of timeslot 2. This argument also holds for all timeslots $t = 3, \ldots, T$, and thus, E can decode $(W_1[1], \ldots, W_1[T])$ and $(W_2[t], \ldots, W_2[T])$. However, it is impossible for E to decode any of the $W_3[t]$, $t = 1, \ldots, T$, because it never receives any information related to these bits. Thus, the scheme described here achieves $(R, R_e) = (3, 1)$, and hence ensures 1 bit of secrecy. The remaining examples of this section use layered networks, where the notion of time can be ignored.

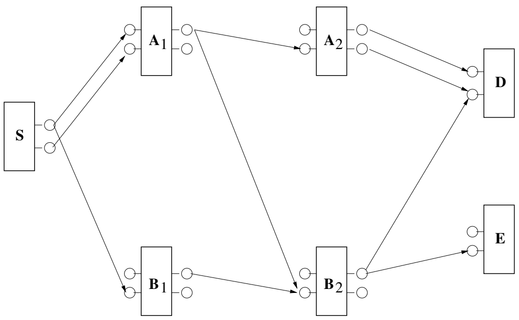


Fig. 3. Example 2: Two-layer network with single eavesdropper, with a map-forward relaying strategy.

Example 2: We now study a larger deterministic wireless network with $\mathcal{A} = \{A_1, A_2, B_1, B_2\}$. The network is shown in Figure 3. The information-theoretic cutset from S to D is 2 bits. This example illustrates that the relays do not have to decode the source message to ensure secure and reliable communication. The source S sends (W_1, W_2). Each relay operates by taking a linear combination of its received signal and transmitting it. One possible operation is when A_1 sends $(0 \cdot W_1) \oplus (1 \cdot W_2) = W_2$ and B_1 forwards its received signal W_1. Therefore A_2 receives (W_2) and B_2 receives the linear combination $(W_1 \oplus W_2)$. Both A_2 and B_2 simply forward what they receive via all of their outputs. As a result, D receives $(W_2, W_2 \oplus W_1 \oplus W_2)$, and it can solve this equation system to obtain both information bits. E receives only $(W_1 \oplus W_2)$, which reveals one bit of information. To see this, we can prepare the source bits (W_1, W_2) such that $W_1 = U_1, W_2 = U_1 \oplus U_2$ and therefore since E receives $W_1 \oplus W_2 = U_2$, it has no knowledge of U_1 which is secret. This example therefore also illustrates a source-encoding scheme to ensure secrecy[8]. ■

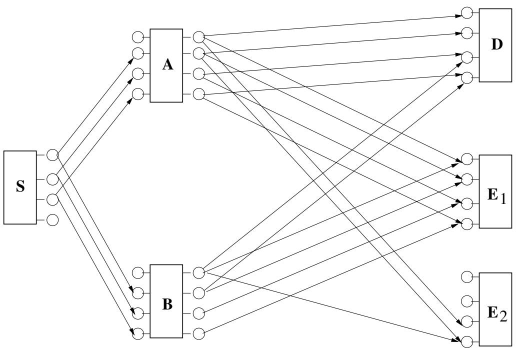


Fig. 4. Example 3, showing the conflict created by multiple potential eavesdroppers.

Example 3: In Figure 4, we illustrate the conflicting needs for secrecy against a class of eavesdroppers, by having $\mathcal{E} = \{E_1, E_2\}$. The information-theoretic cutset bound for the pair (S, D) is 3 bits. If there was only E_1, we could have the following scheme: S sends (W_1, W_2, W_3), A and B both send $(W_1, W_2, W_3, 0)$. In this case, D receives $(W_1, W_2, W_3 \oplus W_1, W_2)$, while E_1 receives $(0, 0, 0, 0)$. Hence, we could have a perfect secrecy situation, achieving $(R, R_e) = (3, 3)$. If there was only E_2, we could have the scheme: S sends (W_1, W_2, W_3), A sends $(0, W_1, W_2, W_3)$, while B sends $(W_1, 0, 0, 0)$. In this case, D receives $(0, W_1, W_2 \oplus W_1, W_3)$, while E_2 receives $(0, 0, 0, W_1 \oplus W_1) = (0, 0, 0, 0)$, again leading to perfect secrecy. Now, when $\mathcal{E} = \{E_1, E_2\}$, *i.e.,* either eavesdropper (or both eavesdroppers) could be present, we need to ensure secrecy against the set. A possible strategy is: S sends (W_1, W_2, W_3), A sends $(0, W_1, W_2, W_3)$, and B sends $(0, W_1, W_2, W_3)$ or (W_1, W_1, W_2, W_3). In this case, D receives $(0, W_1, W_2, W_3 \oplus W_1)$ or $(0, W_1, W_2 \oplus W_1, W_3 \oplus W_1)$, E_1 receives $(0, 0, 0, 0)$ or $(W_1, 0, 0, 0)$, and E_2 receives $(0, 0, 0, W_1)$ or $(0, 0, 0, 0)$. In either case, we only achieve $(R, R_e) = (3, 2)$ bits. Some additional thought also reveals that it is not possible to make R_e larger than 2 when E_1 and E_2 can both be present. We thus see the tension created by the uncertainty about the eavesdropper. ■

Example 4: For the network given in Figure 5, we introduce a new relaying strategy, where some of the authenticated relay nodes actively help secrecy by inserting random bits into the communication[9]. In a sense these relays are "jamming" the eavesdropper and helping secure communication. The cutset bound between (S, D) is 2 bits. Let the source send (W_1, W_2) at its outgoing nodes. One strategy is that A sends $(0, W_1, 0)$, and B sends (W_2). In this case, D receives $(0, W_1, W_2)$, while E receives $(0, 0, W_2)$, yielding $(R, R_e) = (2, 1)$. Any strategy where the relay does not insert randomness cannot ensure more than 1 bit of secrecy. Now we allow A to generate a random bit b before every transmission. Then, it transmits $(b, W_1, 0)$. The transmitter at B remains unchanged. Now, D receives (b, W_1, W_2). The eavesdropper E receives $(0, 0, W_2 \oplus b)$, and therefore we obtain perfect secrecy of 2 bits. This example illustrates that active noise insertion by the relays can enhance secrecy. ■

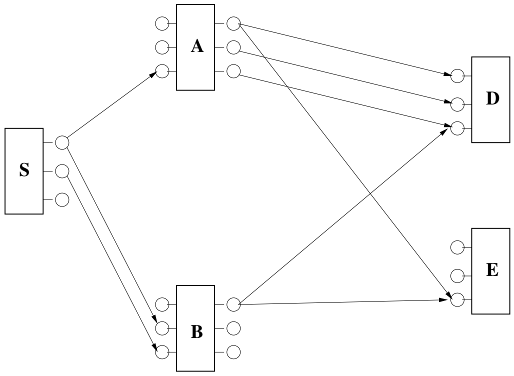


Fig. 5. Example 4: Example showing how noise insertion at a relay can improve secrecy.

B. A Gaussian example

In this section, we give an example of a scheme that achieves perfect secrecy in a Gaussian wireless network. Consider the diamond network shown in Figure 6, whose channels are defined as follows.

Definition 5: The *Gaussian diamond network* is a network with two relays A and B and one eavesdropper E, whose channels are given by the equations:

$$y_A[t] = h_{SA} x_S[t] + z_A[t] \tag{4}$$

$$y_B[t] = h_{SB} x_S[t] + z_B[t] \tag{5}$$

$$y_D[t] = h_{AD} x_A[t] + h_{BD} x_B[t] + z_D[t] \tag{6}$$

$$y_E[t] = h_{AE} x_A[t] + h_{BE} x_B[t] + z_E[t]. \tag{7}$$

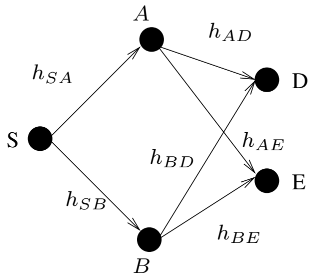


Fig. 6. Gaussian diamond wireless network.

For simplicity assume that all channel gains are real, and that $|h_{SA}| > |h_{SB}|$. Equations (4) and (5) together describe

[8] Those who are familiar with secure network coding [5] might notice some connections with this example. However, in network coding, the communication is over a graph and hence the medium does not impose broadcast and interference, allowing for a larger range of strategies. In a wireless channel the broadcast and multiple access properties are imposed by nature, making the problem more difficult.

[9] A related strategy was developed in [10] for the Gaussian (single) relay channel where the relay forwarded Gaussian noise along with decoded information.

a stochastically degraded Gaussian broadcast channel, while equations (6) and (7) each describe a Gaussian multiple-access channel.

For an arbitrary $\theta \in [0,1]$, define the following two functions:

$$R^{\text{BC}}_{m_B}(\theta) \triangleq \frac{1}{2}\log\left(1+\frac{\theta h^2_{SB}}{1+(1-\theta)h^2_{SB}}\right) \tag{8}$$

$$R^{\text{BC}}_{m_A}(\theta) \triangleq \frac{1}{2}\log\left(1+(1-\theta)h^2_{SA}\right). \tag{9}$$

It is well known [17] that for any $\theta \in [0,1]$, one can reliably transmit a message m_B of rate $R^{\text{BC}}_{m_B}$ to B and simultaneously transmit (m_A, m_B) to A, where m_A is of rate $R^{\text{BC}}_{m_A}$. We also define the following region of rates:

Definition 6: For $k \in \{D, E\}$, we define $\mathcal{R}^{\text{MAC}}_k$ as the region of all pairs (R_A, R_B) satisfying

$$\begin{aligned} R_A &< \frac{1}{2}\log(1+h^2_{Ak}) \\ R_B &< \frac{1}{2}\log(1+h^2_{Bk}) \\ R_A + R_B &< \frac{1}{2}\log(1+h^2_{Ak}+h^2_{Bk}). \end{aligned}$$

Note that $\mathcal{R}^{\text{MAC}}_k$ is the achievable rate-region of the multiple-access channel from A and B to k, for $k \in \{D, E\}$.

Now, we describe an achievable perfect secrecy rate for the Gaussian diamond network.

Theorem 1: In the Gaussian diamond network,

$$R_s \geq \max\left[R^{\text{BC}}_{m_A}(\theta) + R^{\text{BC}}_{m_B}(\theta) - \frac{1}{2}\log(1+h^2_{AE}+h^2_{BE})\right], \tag{10}$$

where the maximization is over all $\theta \in [0,1]$ for which

$$(R^{\text{BC}}_{m_A}(\theta), R^{\text{BC}}_{m_B}(\theta)) \in \mathcal{R}^{\text{MAC}}_D$$

and for which the set

$$\begin{gathered}\left\{(R_{j_A}, R_{j_B}) \in \mathcal{R}^{\text{MAC}}_E : R_{j_A} + R_{j_B} = \frac{1}{2}\log(1+h^2_{AE}+h^2_{BE})\right\} \\ \cap \left([0, R^{\text{BC}}_{m_A}(\theta)] \times [0, R^{\text{BC}}_{m_B}(\theta)]\right)\end{gathered}$$

is non-empty.

The proof of Theorem 1 is given in the appendix. A similar theorem for the case when the relay nodes are allowed to insert noise is given in Appendix VII-F. ■

First note that the Gaussian diamond network of Figure 6 is very closely related to the deterministic wireless network shown in Figure 5. In fact, the noise insertion strategy in Example 4 can also be incorporated into the Gaussian case to enhance the secrecy. Moreover, all the deterministic examples given have Gaussian counterparts, and the strategies suggested by the deterministic examples can also be implemented in the Gaussian wireless model of (1). This is the connection in the insight that is used in the results presented in Section IV. However, for general networks, we need a strategy that works for *any* network, making the use of more sophisticated probabilistic methods outlined in Section V necessary.

IV. Main Results

First, we introduce the notion of an information-theoretic min-cut in a network.

Definition 7: We denote by Λ_i the set of all cutsets in the network that separates the source S from node i:

$$\Lambda_i = \{\Omega \subset \mathcal{V} : S \in \Omega, i \in \Omega^c\}.$$

For a subset $\Omega \subset \mathcal{V}$, we denote by X_Ω the tuple of all X_i for $i \in \Omega$.

Definition 8: The quantity $\min_{\Omega\in\Lambda_i} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c})$ is called the min-cut between the pair (S, i), where $I(\cdot;\cdot|\cdot)$ is the (conditional) mutual information.

Now, we state our two main results.

Definition 9: Let a distribution $p(\{x_i\}_{i\in\mathcal{V}})$ over all transmit alphabets be given. We define $\mathcal{R}(p)$ to be the set of all pairs (R, R_e) satisfying

$$\begin{aligned} R &< \min_{\Omega\in\Lambda_D} H(Y_{\Omega^c}|X_{\Omega^c}), \\ R_e &< \min\big\{R, \min_{\Omega\in\Lambda_D} H(Y_{\Omega^c}|X_{\Omega^c}) \\ &\qquad - \max_{E\in\mathcal{E}} \min_{\Omega\in\Lambda_E} H(Y_{\Omega^c}|X_{\Omega^c})\big\}. \end{aligned}$$

Theorem 2: In an arbitrary network with deterministic signal interaction as given in (2), all rate-equivocation pairs in the region

$$\mathcal{R} = \cup_{\prod_{i\in\mathcal{V}} p(x_i)} \mathcal{R}(p)$$

are achievable, where the union is taken over all possible product-distributions on the transmit alphabets.

Theorem 2 guarantees the existence of codes for the region of rate-equivocation pairs $\mathcal{R}$. However, this is not a complete characterization in that, we do not have a matching converse stating that no strategy can achieve pairs (R, R_e) that lie outside this region. Also note that the result in Theorem 2 can be generalized to include arbitrary deterministic interaction functions, rather than the linear model given in (2), by using techniques similar to those in [2].

Corollary 1: In a network with deterministic signal interaction as given in (2), the perfect secrecy rate R_s is lower bounded as follows:

$$R_s \geq \max_{\prod_{i\in\mathcal{V}} p(x_i)} \big[\min_{\Omega\in\Lambda_D} H(Y_{\Omega^c}|X_{\Omega^c}) - \max_{E\in\mathcal{E}} \min_{\Omega\in\Lambda_E} H(Y_{\Omega^c}|X_{\Omega^c})\big].$$

Proof: From Theorem 2, whenever $(R, R_e) \in \mathcal{R}$, so is (R_e, R_e), and perfect secrecy at rate R_e is possible. Corollary 1 then directly follows from Definition 4. ■

Definition 10: Let a distribution $p(\{x_i\}_{i\in\mathcal{V}})$ over all transmit alphabets be given. We define $\tilde{\mathcal{R}}(p)$ to be the set of all pairs (R, R_e) satisfying

$$\begin{aligned} R_S &< \min_{\Omega\in\Lambda_D} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) - \gamma, \\ R_e &< \min\big\{R_S - \gamma, \min_{\Omega\in\Lambda_D} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) \\ &\qquad - \max_{E\in\mathcal{E}} \min_{\Omega\in\Lambda_E} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) - \gamma\big\} \end{aligned}$$

where γ is a constant which depends on the topology of the network but *not* on the channel gains in $\mathcal{L}$ or the signal-to-noise ratio (SNR) of operation.

Theorem 3: In a Gaussian wireless network, all rate-equivocation pairs in the region

$$\tilde{\mathcal{R}} = \cup_{\prod_{i\in\mathcal{V}} p(x_i)} \tilde{\mathcal{R}}(p)$$

are achievable, where the union is taken over all possible product-distributions on the transmit alphabets.

Remark 1: The rate-equivocation pairs developed in the examples 1–3 of Section III-A, can indeed be found by a direct application of Theorem 2. Therefore, we can use this result to find the performance for an *arbitrary* network with deterministic signal interactions.

Remark 2: Note that since the gap γ is a constant and the rates given in Definition 10 can grow with the SNR, this gap can be made small with respect to the rates of operation. Therefore, even though there is a gap of γ bits in the bound on the equivocation in Theorem 3, it can be small with respect to the total rate, *i.e.,* we might leak only a small number of bits for Gaussian wireless networks. However, we also believe that this gap is a purely technical issue in the proof, and not a fundamental problem arising in Gaussian wireless networks. Indeed, the diamond example given in Section III-B shows that such a gap may not exist and we can obtain perfect secrecy.

V. Proof Outlines

In this section, we outline the proofs of Theorems 2 and 3, *i.e.*, we prove the achievability of any $(R, R_e) \in \mathcal{R}(p)$ for the deterministic wireless network and the achievability of any (R, R_e) in $\tilde{\mathcal{R}}(p)$ for the Gaussian wireless network for any fixed product distribution p. The proof is sketched for the case when $|\mathcal{E}| = 2$, *i.e.*, there are only two eavesdroppers, denoted by $\mathcal{E} = \{E_1, E_2\}$. The generalization to an arbitrary number of eavesdroppers is straightforward.

In both theorems, we claim the existence of a code that operates at rate R and achieves equivocation R_e. It was shown in [2] that if there is no secrecy requirement, then a coding strategy for the deterministic wireless network is the following: The mapping f_S is a mapping from $\{1, \ldots, 2^{TR}\}$ to a set of 2^{TR} sequences, where each sequence is X_S-typical as defined in [7]. Similarly, the mapping f_i at relay $i \in \mathcal{A}$ is a mapping from the set of all possible receive sequences to a subset of the X_i-typical transmit sequences. For Gaussian wireless networks, there are infinitely many possible receive sequences $\mathbf{y}_i$, and thus, relay i first uses a quantizer to quantize $\mathbf{Y}_i$ to a representation sequence $\hat{\mathbf{Y}}_i$. Then, the mapping is applied to $\hat{\mathbf{Y}}_i$ instead of $\mathbf{Y}_i$. The destination D declares $\hat{w}$ if it is the unique member of $\{1, \ldots, 2^{TR}\}$ jointly typical with receive sequence $\mathbf{Y}_D$. In [2], it is shown for deterministic wireless networks that if we select the mappings f_S and f_i, $i \in \mathcal{A}$ at random, and then use the randomly selected code, then the expected error probability is small as long as $R < \min_{\Omega\in\Lambda_D} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c})$ and the blocklength T is large enough. The expectation is taken over all random codes. Hence, it follows that there exists at least one code which has the desired error probability.

Note that although the existence of a good code is proved via a random coding argument, the selected code itself is deterministic. Now, we introduce the secrecy requirement. To guarantee secrecy, the source-encoder at S needs to be probabilistic. At the beginning of a block, when observing W, S will randomly generate two independent junk-messages J_1 and J_2, which are uniformly distributed in $\{1, \ldots, 2^{TB_1}\}$ and $\{1, \ldots, 2^{TB_2}\}$, respectively. In the following, Lemma 1 states that if a code of this type satisfies certain conditions given in Definition 11, then the region of rate-equivocation pairs (19) and (20) is guaranteed to be achievable. The proof uses the fact that for such a code, each eavesdropper is forced to decode part of the junk (J_1, J_2) before being able to decode W. Then, by making the rate of (J_1, J_2) large enough, we can make it impossible for the eavesdroppers to decode W at all. Finally, Lemmas 4 and 5 state that such secrecy-codes actually exist for the networks we consider. Their proofs are via a random code construction similar to the one in [2] and [3], respectively.

Definition 11: A network is called *securable* with information rate R_D, junk rates R_{E_1}, R_{E_2} and gap Δ if for any $\epsilon_0 > 0$ and for any (R, B_1, B_2) such that

$$R + B_1 + B_2 < R_D, \tag{11}$$

$$B_1 + B_2 < R_{E_1} \tag{12}$$

and

$$B_2 < R_{E_2}, \tag{13}$$

there exists a code that encodes a message triple $(W, J_1, J_2) \in \{1, \ldots, 2^{TR}\} \times \{1, \ldots, 2^{TB_1}\} \times \{1, \ldots, 2^{TB_2}\}$ into blocks of length T, and which is such that

$$I(\mathbf{X}_S; \mathbf{Y}_{E_1}) \leq T(R_{E_1} + \Delta), \tag{14}$$

$$I(\mathbf{X}_S; \mathbf{Y}_{E_2}) \leq T(R_{E_2} + \Delta), \tag{15}$$

and

$$\mathbf{P}(\text{decoding } (W, J_1, J_2) \text{ wrongly at } D) \leq \epsilon_0, \tag{16}$$

$$\frac{1}{2^{TR}} \sum_{w=1}^{2^{TR}} \mathbf{P}(A_w^{(1)}) \leq \epsilon_0 \tag{17}$$

and

$$\frac{1}{2^{T(R+B_1)}} \sum_{w=1}^{2^{TR}} \sum_{j_1=1}^{2^{TB_1}} \mathbf{P}(A_{w,j_1}^{(2)}) \leq \epsilon_0. \tag{18}$$

Here, $\mathbf{X}_S$, $\mathbf{Y}_{E_1}$ and $\mathbf{Y}_{E_2}$ are the transmitted and received sequences for this code when the messages W, J_1 and J_2 are chosen uniformly and independently at random. Further, $A_w^{(1)}$ is the event that E_1 makes an error when decoding (J_1, J_2) given that $W = w$ and assuming that W is available at E_1. Similarly, $A_{w,j_1}^{(2)}$ is the event that E_2 makes an error when decoding J_2 given that $(W, J_1) = (w, j_1)$ and assuming that W and J_1 are available at E_2.

Definition 11 is used in the following lemma, which gives an achievable rate-equivocation region for any securable network.

Lemma 1: If a network is securable with information rate R_D, junk rates R_{E_1}, R_{E_2} and gap Δ, then all rate-equivocation pairs (R, R_e) that satisfy

$$0 \leq R < R_D, \tag{19}$$

$$0 \leq R_e < \min\left\{R, R_D - \max_{E\in\mathcal{E}} R_E - \Delta\right\} \tag{20}$$

are achievable.

Proof: Assume that a network is securable with information rate R_D and junk rates R_{E_1} and R_{E_2}. Without loss of generality, assume that $R_{E_1} \geq R_{E_2}$. Choose B_1 and B_2 as

$$B_1 + B_2 = R_{E_1} - \epsilon_1 \tag{21}$$

and

$$B_2 = R_{E_2} - \epsilon_1, . \tag{22}$$

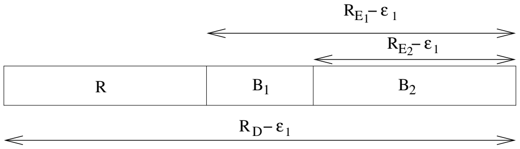


Fig. 7. Illustration of the rate allocation in (21), (22) and (23).

Before each transmission block, the source S generates junk messages (J_1, J_2) with rates B_1 and B_2. Assume that $R = R_D - R_{E_1} - \epsilon_1$. In fact, if R is smaller than this quantity, we artificially increase it by adding additional randomness. If R is larger than this quantity, then we communicate part of it through the junk messages. It follows that

$$R + B_1 + B_2 = R_D - \epsilon_1. \tag{23}$$

Then, since the network is assumed to satisfy Definition 11, we are assured the existence of a code with properties (14) through (18). Let S use this code.

Then, (16) tells us that reliable decoding of (W, J_1, J_2) is possible at the destination D. It remains to show that this code also achieves an equivocation rate R_e arbitrarily close to the upper bound in (20). Using Lemma 2 given at the end of this proof, the remaining uncertainty at eavesdropper E_1 and E_2 can be bounded as

$$\frac{1}{T}H(W|\mathbf{Y}_{E_1}) \geq R + B_1 + B_2 - \frac{1}{T}H(\mathbf{X}_S|W, \mathbf{Y}_{E_1}) - \frac{1}{T}I(\mathbf{X}_S; \mathbf{Y}_{E_1}) \tag{24}$$

and

$$\frac{1}{T}H(W|\mathbf{Y}_{E_2}) \geq R + B_1 + B_2 - \frac{1}{T}H(\mathbf{X}_S|W, J_1, \mathbf{Y}_{E_2}) - \frac{1}{T}I(\mathbf{X}_S; \mathbf{Y}_{E_2}) - \frac{1}{T}H(J_1), \tag{25}$$

respectively. In the application of Lemma 2, we set (a, b) to be $(W, 0)$ and (W, J_1) for E_1 and E_2, respectively.

Note that together with a variation of Fano's inequality, stated in Lemma 3, (17), (18) imply that

$$\frac{1}{T}H(\mathbf{X}_S|W, \mathbf{Y}_{E_1}) \leq \epsilon_2 \tag{26}$$

and

$$\frac{1}{T}H(\mathbf{X}_S|W, J_1, \mathbf{Y}_{E_2}) \leq \epsilon_3. \tag{27}$$

From (23), (24), (14) and (26), we get the bound

$$\frac{1}{T}H(W|\mathbf{Y}_{E_1}) \geq R_D - R_{E_1} - \Delta - \epsilon_1 - \epsilon_2.$$

For E_2, we combine (23), (25), (15) and (27) with the identity $H(J_1) = B_1$, and obtain

$$\frac{1}{T}H(W|\mathbf{Y}_{E_2}) \geq R_D - R_{E_2} - \Delta - B_1 - \epsilon_1 - \epsilon_3. \tag{28}$$

From (21) and (22) we get that $R_{E_2} + B_1 = R_{E_1}$, hence

$$\frac{1}{T}H(W|\mathbf{Y}_{E_2}) \geq R_D - R_{E_1} - \Delta - \epsilon_1 - \epsilon_3.$$

It follows that by choosing ϵ_k, $k = 1, 2, 3$ sufficiently small, the uncertainties at each of the eavesdroppers can be made arbitrarily close to the right hand side of (20). ■

The following two lemmas are used in the proof of Lemma 1.

Lemma 2: Consider a set of random variables denoted by (a, b), $\mathbf{X}_S$ and $\mathbf{Y}_E$. Then,

$$H(a|\mathbf{Y}_E) \geq H(\mathbf{X}_S) - I(\mathbf{X}_S; \mathbf{Y}_E) - H(\mathbf{X}_S|a, b, \mathbf{Y}_E) - H(b).$$

Lemma 3: Consider a deterministic code for a network that transmits a pair of messages (W, J), that are uniformly distributed in $\{1, \ldots, 2^{TR}\} \times \{1, \ldots, 2^{TB}\}$. Let T be the blocksize of the code, and assume that

$$\frac{1}{2^{TR}} \sum_{w=1}^{2^{TR}} \mathbf{P}(A_w) \leq \epsilon_0, \tag{29}$$

where A_w is the event that a certain node E makes an error when decoding J given that $W = w$ and assuming that W is available at E. Then, it follows that

$$H(\mathbf{X}_S|W, \mathbf{Y}_E) \leq 1 + TB\epsilon_0.. \tag{30}$$

The main results of this paper are consequences of the following two lemmas:

Lemma 4: Given a deterministic wireless network and a product distribution p on its transmit alphabets, the network is securable with gap $\Delta = 0$ and the following information and junk rates:

$$R_D = \min_{\Omega\in\Lambda_D} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c})$$

$$R_{E_1} = \min_{\Omega\in\Lambda_{E_1}} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c})$$

$$R_{E_2} = \min_{\Omega\in\Lambda_{E_2}} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}).$$

■

Lemma 5: Given a Gaussian wireless network and a product distribution p on its transmit alphabets, the network is securable with gap $\Delta = \gamma$ and the following information and junk rates:

$$\begin{aligned} R_D &= \min_{\Omega \in \Lambda_D} I(X_\Omega; Y_{\Omega^c} | X_{\Omega^c}) - \gamma \\ R_{E_1} &= \min_{\Omega \in \Lambda_{E_1}} I(X_\Omega; Y_{\Omega^c} | X_{\Omega^c}) - \gamma \\ R_{E_2} &= \min_{\Omega \in \Lambda_{E_2}} I(X_\Omega; Y_{\Omega^c} | X_{\Omega^c}) - \gamma, \end{aligned}$$

where γ is as in Definition 10.

■

The proof of Theorem 2 follows from Lemma 1 and Lemma 4 by noting that for deterministic wireless networks, $I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) = H(Y_{\Omega^c}|X_{\Omega^c})$. The proof of Theorem 3 follows from Lemma 1 and Lemma 5.

VI. Discussion

In this paper we have formulated a communication scenario with secrecy requirement for wireless networks with cooperating relays which help to enhance the secrecy against an eavesdropper. The eavesdropper's wireless channel is not exactly known but only known to belong to a class. The way we generate secrecy uses the characteristic properties of the wireless channel: the eavesdroppers channel is statistically distinct from that of the destination and our schemes exploit this difference to set up cooperation mechanisms to provide secrecy. It is possible to interpret equivocation as secret key generation, and therefore we can use the techniques outlined in this paper to generate a unconditionally secure key. The translation of message secrecy to key-generation is as follows. Through proper coding, a part of the message of size equal to the equivocation can be kept secret from any eavesdropper, and thus play the role of a secret key that is communicated to the destination. It should be noted the rate-equivocation results presented in this paper are only achievability results, and not a complete characterization of the rate-equivocation region. To obtain such a characterization we need a matching converse stating that no scheme can do better, and this is part of further research. Even though through examples in Section III, we discussed insertion of noise by the relays, the results presented in Section IV are for deterministically operating relays. The study of noise-inserting relays for arbitrary networks is also a subject for further work.

References

[1] A. Avestimehr, S. Diggavi, and D. Tse, *A deterministic approach to wireless relay networks*, in Proceedings of Allerton Conference on Communication, Control, and Computing, Illinois, USA, Sept. 2007. See: http://licos.epfl.ch/index.php?p=research_projWNC.

[2] ———, *Wireless network information flow*, in Proceedings of Allerton Conference on Communication, Control, and Computing, Illinois, USA, Sept. 2007. See: http://licos.epfl.ch/index.php?p=research_projWNC.

[3] ———, *Approximate capacity of gaussian relay networks*, in Proc. of the IEEE Int. Symposium on Inform. Theory, Toronto, Canada, July 2008. See: http://licos.epfl.ch/index.php?p=research_projWNC.

[4] L. Buttyan and J.-P. Hubaux, *Security and Cooperation in Wireless Networks*, Cambridge University Press, Cambridge, 2008.

[5] N. Cai and R. Yeung, *Secure network coding*, in Proc. of the IEEE Int. Symposium on Inform. Theory, Lausanne, Switzerland, June 2002.

[6] I. Csiszár and J. Korner, *Broadcast channels with confidential messages*, IEEE Trans. Inform. Theory, 24 (1978).

[7] ———, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

[8] W. Diffie and M. E. Hellman, *New directions in cryptography*, IEEE Transactions on Information Theory, 22 (1976), pp. 644–654.

[9] P. Gupta and P. Kumar, *The capacity of wireless networks*, IEEE Trans. Inform. Theory, 46 (2000), pp. 388–404.

[10] L. Lai and H. E. Gamal, *Cooperative secrecy: The relay-eavesdropper channel*, in Proc. of the IEEE Int. Symposium on Inform. Theory, Nice, France, June 2007.

[11] Y. Oohama, *Relay channels with confidential messages*, IEEE Trans. Inform. Theory, (2006). submitted.

[12] A. Ozgur, O. Leveque, and D. Tse, *Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks*, IEEE Transactions on Information Theory, 53 (2007), pp. 3549–3572.

[13] E. Perron, S. Diggavi, and E. Telatar, *On cooperative wireless network secrecy*, Tech. Rep. LICOS-REPORT-2008-009, École Polytechnique Fédérale de Lausanne, Switzerland, Aug. 2008. http://infoscience.epfl.ch/record/126166.

[14] R. Rivest, A. Shamir, and L. Adleman, *A method for obtaining digital signatures and public-key cryptosystems*, Communications of the ACM, 21 (1978), pp. 120–126.

[15] C. E. Shannon, *Communication theory of secrecy systems*, Bell System Tech. J., 28 (1949), pp. 656–715.

[16] P. Shor, *Algorithms for quantum computation: Discrete logarithms and factoring*, in Proc. 35nd Annual Symposium on Foundations of Computer Science, 1994, pp. 124–134.

[17] Th.M.Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[18] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, May 2005.

[19] M. Wegman and J. Carter, *New hash functions and their use in authentication and set equality*, Journal of Computer and System Sciences, 22 (1981), pp. 265–279.

[20] A. D. Wyner, *The wire-tap channel*, Bell System Tech. J., 54 (1975), pp. 1355–1387.

[21] L.-L. Xie and P. Kumar, *A network information theory for wireless communication: Scaling laws and optimal operation*, IEEE Trans. Inform. Theory, 50 (2004), pp. 748–767.

VII. Appendix

A. Proof of Lemma 2

The proof follows through the following chain of inequalities:

$$\begin{aligned} H(a|\mathbf{Y}_E) &\overset{(a)}{\geq} H(a|\mathbf{Y}_E, b) \\ &= H(\mathbf{X}_S) + H(a, \mathbf{Y}_E, b|\mathbf{X}_S) \\ &\quad - H(\mathbf{X}_S|a, \mathbf{Y}_E, b) - H(\mathbf{Y}_E, b) \end{aligned}$$

$$\begin{aligned} &\overset{(b)}{\geq} H(\mathbf{X}_S) + H(\mathbf{Y}_E|\mathbf{X}_S) \\ &\quad - H(\mathbf{X}_S|a, \mathbf{Y}_E, b) - H(\mathbf{Y}_E, b) \\ &\overset{(a)}{\geq} H(\mathbf{X}_S) + H(\mathbf{Y}_E|\mathbf{X}_S) \\ &\quad - H(\mathbf{X}_S|a, \mathbf{Y}_E, b) - H(\mathbf{Y}_E) - H(b) \\ &= H(\mathbf{X}_S) - I(\mathbf{X}_S; \mathbf{Y}_E) \\ &\quad - H(\mathbf{X}_S|a, \mathbf{Y}_E, b) - H(b), \end{aligned}$$

where the items marked with (a) are true because conditioning decreases the entroyp, and (b) follows by dropping the variables a and b from the second term.

B. *Proof of Lemma 3*

Let the assumptions of Lemma 3 be true. To show (30), we start with the following inequality:

$$\begin{aligned} H(\mathbf{X}_S|W,\mathbf{Y}_E) &\le H(\mathbf{X}_S,J|W,\mathbf{Y}_E) \\ &= H(J|W,\mathbf{Y}_E) + \underbrace{H(\mathbf{X}_S|J,W,\mathbf{Y}_E)}_{=0} \\ &= H(J|W,\mathbf{Y}_E), \end{aligned}$$

where the indicated term is zero because the transmit sequence $\mathbf{X}_S$ is a function of (J,W). We proceed to upper bound $H(J|W,\mathbf{Y}_E)$. First, define the binary random variable

$$\xi \triangleq \mathbf{1}_{\{A_W\}},$$

where A_w is as defined in Lemma 3. In words, ξ indicates whether J was wrongly decoded at E, given that W is available at E. We expand the following joint entropy in two different ways:

$$\begin{aligned} H(\xi,J|W,\mathbf{Y}_E) &= H(J|W,\mathbf{Y}_E) + H(\xi|J,W,\mathbf{Y}_E) \\ &= H(\xi|W,\mathbf{Y}_E) + H(J|W,\xi,\mathbf{Y}_E). \end{aligned} \tag{31}$$

Knowing $\mathbf{Y}_E$ and W, we know the decision $\hat{J} = f_D(\mathbf{Y}_E, W)$ made by E, and hence

$$H(\xi|J,W,\mathbf{Y}_E) = H(\xi|J,\hat{J},W,\mathbf{Y}_E) = 0.$$

Also

$$\begin{aligned} H(\xi|W,\mathbf{Y}_E) &\le H(\xi) \\ &= h(\mathbf{P}(\xi=1)) \\ &\le 1. \end{aligned}$$

Using this, (31) implies

$$\begin{aligned} H(J|W,\mathbf{Y}_E) &\le 1 + H(J|\xi,\mathbf{Y}_E,W) \\ &= 1 + \sum_{w=1}^{2^{TR}} \mathbf{P}(W=w) H(J|\xi,\mathbf{Y}_E,W=w) \\ &= 1 + \sum_{w=1}^{2^{TR}} \frac{1}{2^{TR}} H(J|\xi,\mathbf{Y}_E,W=w) \\ &= 1 + \sum_{w=1}^{2^{TR}} \frac{1}{2^{TR}} \big[\mathbf{P}(\xi=0|W=w) \\ &\quad \cdot \underbrace{H(J|\xi=0,\mathbf{Y}_E,W=w)}_{=0} \\ &\quad + \mathbf{P}(\xi=1|W=w) \\ &\quad \cdot \underbrace{H(J|\xi=1,\mathbf{Y}_E,W=w)}_{\le \log(2^{TB}-1) \le TB}\big] \\ &\le 1 + TB\frac{1}{2^{TR}} \sum_{w=1}^{2^{TR}} \mathbf{P}(\xi=1|W=w) \\ &= 1 + TB\frac{1}{2^{TR}} \sum_{w=1}^{2^{TR}} \mathbf{P}(A_w) \\ &\le 1 + TB\epsilon_0, \end{aligned}$$

where the last inequality is true because of (29).

C. *Proof of Lemma 4*

Let a network, a transmit product distribution p, and rates R, B_1 and B_2 be given, with

$$R + B_1 + B_2 < \min_{\Omega \in \Lambda_D} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}), \tag{32}$$

$$B_1 + B_2 < \min_{\Omega \in \Lambda_{E_1}} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) \tag{33}$$

and

$$B_2 < \min_{\Omega \in \Lambda_{E_2}} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}). \tag{34}$$

We show that there exists a code that satisfies (14) through (18) in Definition 11.

Let the block-length be T. We generate a random codebook in the following way. For the source node S, each (W, J_1, J_2) is mapped to a sequence $\mathbf{x}_S$, drawn uniformly at random from the set of all X_S-typical sequences (with respect to the distribution p). For each relay node $i \in \mathcal{A}$, each possible input sequence $\mathbf{y}_i$ is mapped to a sequence $\mathbf{x}_i$, drawn uniformly at random from the set of X_i-typical sequences. Once this random code generation process is finished, the mappings are deterministic and fixed for all time. From (32) and the proof of Theorem 2.1 in [2], it is clear that for T large enough,

$$\mathrm{E}\left[\mathbf{P}(\text{decoding } (W,J_1,J_2) \text{ wrongly at } D)\right] \le \frac{\epsilon_0}{3}, \tag{35}$$

where the expectation is taken over all (randomly generated) codes. On the other hand, for a fixed $W = w$, the number of codewords used at S is $2^{T(B_1+B_2)}$, and there is (statistically) no difference between this collection of codewords and a randomly generated codebook of size $2^{T(B_1+B_2)}$. Hence, if both S and E_1 know W, then the error probability analysis from [2] can be applied again for every fixed $W = w$, and we find that as long as (33) holds, we have $\mathrm{E}\left[\mathbf{P}(A_w^{(1)})\right] \le \frac{\epsilon_0}{3}$ when T is large enough. Since this is true for any $w \in \{1, \ldots, 2^{TR}\}$, we obtain

$$\mathrm{E}\left[\frac{1}{2^{TR}} \sum_{w=1}^{2^{TR}} \mathbf{P}(A_w^{(1)})\right] = \frac{1}{2^{TR}} \sum_{w=1}^{2^{TR}} \mathrm{E}\left[\mathbf{P}(A_w^{(1)})\right] \le \frac{\epsilon_0}{3}. \tag{36}$$

We apply the same argument once more for E_2 and for each fixed $(W, J_1) = (w, j_1)$, to obtain

$$\mathrm{E}\left[\frac{1}{2^{T(R+B_1)}} \sum_{w=1}^{2^{TR}} \sum_{j_1=1}^{2^{TB_1}} \mathbf{P}(A_{w,j_1}^{(2)})\right] \le \frac{\epsilon_0}{3}. \tag{37}$$

Note that if networks are non-layered as defined in [2], than some care must be taken when using the proof of Theorem 2.1 in [2]. In a non-layered network, some relay nodes may, during a certain block of time-slots, receive a signal which depends on sequences $\mathbf{x}_S^{(k_1)}$ and $\mathbf{x}_S^{(k_2)}$ that were transmitted during different past blocks k_1 and k_2. To resolve this, the scheme needs to be operated over a sequence of transmission blocks, and the decoding is done over the sequence of blocks rather than over a single block. Details of this proof technique for non-layered networks can be found in [2].

Summing (35), (36) and (37), we see that

$$\mathrm{E}\Bigg[\mathbf{P}(\text{decoding } (W,J_1,J_2) \text{ wrongly at } D)+ \frac{1}{2^{TR}}\sum_{w=1}^{2^{TR}}\mathbf{P}(A_w^{(1)})+ \frac{1}{2^{T(R+B_1)}}\sum_{w=1}^{2^{TR}}\sum_{j_1=1}^{2^{TB_1}}\mathbf{P}(A_{w,j_1}^{(2)})\Bigg] \le \epsilon_0,$$

where the expectation is over all codes. From this, we conclude that for at least one of the codes, (16), (17) and (18) are true at the same time.

It remains to show that the code satisfies (14) and (15). Let $E \in \{E_1, E_2\}$, and fix any cut $\Omega \in \Lambda_E$.

Assume that the network is layered. Note that whenever we write $I(\mathbf{X}_S;\mathbf{Y}_E)$, what we mean is $I(\mathbf{X}_S(M_1);\mathbf{Y}_E(M_1))$, where $M_1 = (W_1, J_{1,1}, J_{2,1})$ is the message transmitted during block 1 (information and junk messages combined), which is uniformly distributed in $\{1,\ldots,2^{T(R+B_1+B_2)}\}$. So far, we have not mentioned that if E cannot directly receive the signal transmitted by S, then $\mathbf{X}_S(M_1)$ and $\mathbf{Y}_E(M_1)$ are blocks that occur at different instances in time. If, for instance, there is one layer of relays between S and E, then a relay node A would receive $\mathbf{Y}_A(M_1)$ during the transmission of $\mathbf{X}_S(M_1)$ by S, but it would transmit $\mathbf{X}_A(M_1)$ during next block of length T, during the reception of $\mathbf{Y}_A(M_2)$. E would therefore start receiving the block $\mathbf{Y}_E(M_1)$ after S has finished transmitting $\mathbf{X}_S(M_1)$. Let K be a large positive integer. We assume that transmission takes place over many time-blocks, but we focus on a window of K time-blocks, each of length T. We define a new set of vectors $(\underline{\mathbf{X}}_{\mathcal{V}}, \underline{\mathbf{Y}}_{\mathcal{V}})$ of length $(K+L)T$, where L is the number of layers of relay-nodes that connect S to E in the network. Recall that $\mathcal{V}$ is the set of all nodes in the network, and $\mathcal{A}$ is the set of relay nodes. Let $\mathcal{A}_l \subseteq \mathcal{A}$ be the set of relay nodes that lie in layer l. For $l = 0,\ldots,L$, define

$$\underline{\mathbf{X}}_{\mathcal{A}_l}(\underline{M}) \triangleq \begin{bmatrix} \mathbf{X}_{\mathcal{A}_l}(M_{1-l}) \\ \vdots \\ \mathbf{X}_{\mathcal{A}_l}(M_0) \\ \mathbf{X}_{\mathcal{A}_l}(M_1) \\ \vdots \\ \mathbf{X}_{\mathcal{A}_l}(M_K) \\ \mathbf{X}_{\mathcal{A}_l}(M_{K+1}) \\ \vdots \\ \mathbf{X}_{\mathcal{A}_l}(M_{K+L-l}) \end{bmatrix}$$

and

$$\underline{\mathbf{Y}}_{\mathcal{A}_l}(\underline{M}) \triangleq \begin{bmatrix} \mathbf{Y}_{\mathcal{A}_{l+1}}(M_{1-l}) \\ \vdots \\ \mathbf{Y}_{\mathcal{A}_{l+1}}(M_0) \\ \mathbf{Y}_{\mathcal{A}_{l+1}}(M_1) \\ \vdots \\ \mathbf{Y}_{\mathcal{A}_{l+1}}(M_K) \\ \mathbf{Y}_{\mathcal{A}_{l+1}}(M_{K+1}) \\ \vdots \\ \mathbf{Y}_{\mathcal{A}_{l+1}}(M_{K+L-l}) \end{bmatrix},$$

where we set $\mathcal{A}_0 \triangleq \{S\}$ and $\mathcal{A}_{L+1} \triangleq \{E\}$, and $\underline{M} = \{M_{1-L},\ldots,M_{K+2L}\}$ is the set of all messages involved during $K+3L$ blocks. Assume that transmission of the first message M_1 starts at $t = 1$. We then have that the t-th component of each super-block $\underline{\mathbf{X}}_i$ or $\underline{\mathbf{Y}}_i$, $i \in \mathcal{V}$ is actually received or transmitted at time t. We have the following chain of inequalities, which closely follows the proof of Theorem 14.10.1 in [17]:

$$\begin{aligned}
I(\underline{\mathbf{X}}_S;\underline{\mathbf{Y}}_E) &\overset{(a)}{\le} I(\underline{\mathbf{X}}_\Omega;\underline{\mathbf{Y}}_{\Omega^c}) \\
&= \sum_{t=1}^{(K+L)T} I(\underline{\mathbf{X}}_\Omega;\underline{Y}_{\Omega^c}[t]|\underline{Y}_{\Omega^c}[1],\ldots,\underline{Y}_{\Omega^c}[t-1]) \\
&= \sum_{t=1}^{(K+L)T} \Big[H(\underline{Y}_{\Omega^c}[t]|\underline{Y}_{\Omega^c}[1],\ldots,\underline{Y}_{\Omega^c}[t-1]) \\
&\quad - H(\underline{Y}_{\Omega^c}[t]|\underline{Y}_{\Omega^c}[1],\ldots,\underline{Y}_{\Omega^c}[t-1],\underline{\mathbf{X}}_\Omega)\Big] \\
&\overset{(b)}{\le} \sum_{t=1}^{(K+L)T} \Big[H(\underline{Y}_{\Omega^c}[t]|\underline{Y}_{\Omega^c}[1],\ldots,\underline{Y}_{\Omega^c}[t-1],\underline{X}_{\Omega^c}[t]) \\
&\quad - H(\underline{Y}_{\Omega^c}[t]|\underline{Y}_{\Omega^c}[1],\ldots,\underline{Y}_{\Omega^c}[t-1],\underline{\mathbf{X}}_\Omega,\underline{X}_{\Omega^c}[t])\Big] \\
&\overset{(c)}{\le} \sum_{t=1}^{(K+L)T} \Big[H(\underline{Y}_{\Omega^c}[t]|\underline{X}_{\Omega^c}[t]) \\
&\quad - H(\underline{Y}_{\Omega^c}[t]|\underline{X}_\Omega[t],\underline{X}_{\Omega^c}[t])\Big] \\
&= \sum_{t=1}^{(K+L)T} I(\underline{X}_\Omega[t];\underline{Y}_{\Omega^c}[t]|\underline{X}_{\Omega^c}[t]) \\
&\overset{(d)}{\to} (K+L)T\; I(X_\Omega;Y_{\Omega^c}|X_{\Omega^c}),
\end{aligned}$$

as T and K grow large. The steps are justified below:

- (a) The mutual information is made larger by adding more super-blocks on each side.
- (b) $\underline{X}_{\Omega^c}[t]$ is what is transmitted by all nodes in Ω^c at time t and thus is a function of everything that was received by all nodes in Ω^c up to time $t-1$. Hence, we have equality in the first term. The second term is dicreased by further conditioning on $\underline{X}_{\Omega^c}[t]$.
- (c) The first term is increased by dropping a part of the conditioning. For the second term, we use the fact that $\underline{Y}_{\Omega^c}[t]$ depends only on the current transmit values $\underline{X}_\Omega[t], \underline{X}_{\Omega^c}[t]$.

(d) Assume that $t \in \{(k-1)T+1, \ldots, kT\}$ for some $k \in \{1, \ldots, K+L\}$, *i.e.*, assume that t lies in the k-th block of the super-block. Consider the first layer: $(\underline{X}_S[t], \underline{Y}_{\mathcal{A}_1}[t])$ is a fixed component of $(\mathbf{X}_S(M_k), \mathbf{Y}_{\mathcal{A}_1}(M_k))$, which is a set of jointly typical sequences with respect to $p_{X_S} p_{Y_{\mathcal{A}_1}|X_S}$, picked uniformly at random from a codebook of size $2^{T(R+B_1+B_2)}$. Hence, as T grows large, the joint distribution of $(\underline{X}_S[t], \underline{Y}_{\mathcal{A}_1}[t])$ converges to $p_{X_S} p_{Y_{\mathcal{A}_1}|X_S}$. We advance one layer in the network: The tuple $(\underline{X}_{\mathcal{A}_1}[t], \underline{Y}_{\mathcal{A}_2}[t])$ is a fixed component of the tuple of sequences $(\mathbf{X}_{\mathcal{A}_1}(M_{k-1}), \mathbf{Y}_{\mathcal{A}_2}(M_{k-1}))$, which is a member of a collection of jointly typical sequences with respect to $p_{X_{\mathcal{A}_1}} p_{Y_{\mathcal{A}_2}|X_{\mathcal{A}_1}}$. In addition, the tuple of sequences is independent of the tuple in the first layer, because the message M_{k-1} is independent of M_k. Nevertheless, it follows that the distribution of $(\underline{X}_{\mathcal{A}_1}[t], \underline{Y}_{\mathcal{A}_2}[t])$ converges to $p_{X_{\mathcal{A}_1}} p_{Y_{\mathcal{A}_2}|X_{\mathcal{A}_1}}$. We do so for each layer, and we find that the distribution of $(\underline{X}_{\mathcal{V}}[t], \underline{Y}_{\mathcal{V}}[t])$ converges to

$$\prod_{l=1}^{L+1} p_{X_{\mathcal{A}_{l-1}}} p_{Y_{\mathcal{A}_l}|X_{\mathcal{A}_{l-1}}}. \tag{38}$$

The product comes from the fact that all involved messages M_k through M_{k+L} are independent. But because $p_{\{X_i\}_{i \in \mathcal{V}}}$ was chosen to be a product distribution, (38) is nothing but $p_{X_{\mathcal{V}}} p_{Y_{\mathcal{V}}|X_{\mathcal{V}}}$. Then, because the mutual information is a continuous function of the underlying distribution, it holds that $I(\underline{X}_\Omega[t]; \underline{Y}_{\Omega^c}[t] | \underline{X}_{\Omega^c}[t]) \to I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c})$ for all t.

Finally, we can write

$$\begin{aligned} I(\underline{\mathbf{X}}_S; \underline{\mathbf{Y}}_E) &\overset{(a)}{\geq} \\ &I(\mathbf{X}_S(M_1), \ldots, \mathbf{X}_S(M_K); \mathbf{Y}_E(M_1), \ldots, \mathbf{Y}_E(M_K)) \\ &\overset{(b)}{=} \sum_{k=1}^{K} I(\mathbf{X}_S(M_k); \mathbf{Y}_E(M_k)) \\ &= K I(\mathbf{X}_S(M_1); \mathbf{Y}_E(M_1)), \end{aligned} \tag{39}$$

where we obtain (a) by dropping all terms in $\underline{\mathbf{X}}_S$ and $\underline{\mathbf{Y}}_E$ that depend on messages different from $M_1, \ldots, M_K$, and (b) is true by the independence of the M_k. Combining (39) with the chain of inequalities, we get

$$\begin{aligned} I(\mathbf{X}_S(M_1); \mathbf{Y}_E(M_1)) &\leq \frac{K+L}{K} T I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) \\ &\to T I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) \end{aligned} \tag{40}$$

as K grows large.

Since (40) is true for any $\Omega \in \Lambda_E$, we have that

$$I(\mathbf{X}_S(M_1); \mathbf{Y}_E(M_1)) \leq T \min_{\Omega \in \Lambda_E} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}),$$

which proves that the code satisfies (14) and (15), with $\Delta = 0$.

For non-layered networks, a similar argument can be made, by considering communication over many super-blocks.

D. *Proof of Lemma 5*

The proof for the Gaussian case is very similar to the proof of Lemma 4, except that this time, we use results in [3]. For a given network $\mathcal{G}$ and a given product distribution p, we have that when we use a randomly generated codebook for (W, J_1, J_2) at the source node S and vector quantizers followed by randomly generated mappings at the relay nodes, then the three error probabilities of interest, averaged over all codes, are all small as long as

$$R + B_1 + B_2 < \min_{\Omega \in \Lambda_D} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) - \gamma, \tag{41}$$

$$B_1 + B_2 < \min_{\Omega \in \Lambda_{E_1}} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) - \gamma \tag{42}$$

and

$$B_2 < \min_{\Omega \in \Lambda_{E_2}} I(X_\Omega; Y_{\Omega^c}|X_{\Omega^c}) - \gamma. \tag{43}$$

We can conclude that a code that satisfies (16), (17) and (18) exists, with the information- and junk rates given by the right hand sides of (41), (42) and (43), respectively.

We use exactly the same upper bounding technique as in Appendix VII-C, and we obtain that the gap between the rates and the corresponding bounds is $\Delta = \gamma$.

E. *Proof of Theorem 1*

The claim of the theorem is that communication at a rate arbitrarily close to the one given in the optimization (10) is possible with perfect secrecy. It suffices to show the existence of a code that achieves this:

Let θ be such that

$$(R^{\text{BC}}_{m_A}(\theta), R^{\text{BC}}_{m_B}(\theta)) \in \mathcal{R}^{\text{MAC}}_D \tag{44}$$

and such that

$$\mathcal{B}_E \cap \left([0, R^{\text{BC}}_{m_A}(\theta)] \times [0, R^{\text{BC}}_{m_B}(\theta)]\right) \tag{45}$$

is non-empty, where

$$\begin{aligned} \mathcal{B}_E = \big\{ &(R_{j_A}, R_{j_B}) \in \mathcal{R}^{\text{MAC}}_E : \\ &R_{j_A} + R_{j_B} = \frac{1}{2}\log(1 + h^2_{AE} + h^2_{BE}) \big\}. \end{aligned} \tag{46}$$

Then, define $R_D = R^{\text{BC}}_{m_A}(\theta) + R^{\text{BC}}_{m_B}(\theta)$ and $R_E = \frac{1}{2}\log(1 + h^2_{AE} + h^2_{BE})$, and let $R = R_D - R_E - \epsilon_1$. Let W be the information message, uniformly distributed in $\{1, \ldots, 2^{TR}\}$. We split W into two parts, W_A and W_B, of rate R_{w_A} and R_{w_B}, respectively, such that W_k is uniformly distributed in $\{1, \ldots, 2^{TR_{w_k}}\}$, for $k \in \{A, B\}$. Before each transmission block, the source S generates two junk messages J_A and J_B, uniformly distributed in $\{1, \ldots, 2^{TR_{j_A}}\}$ and $\{1, \ldots, 2^{TR_{j_B}}\}$, respectively. Define $M_A = (W_A, J_A)$ and $M_B = (W_B, J_B)$. The junk rates R_{j_A} and R_{j_B} are picked arbitrarily such that

$$(R_{j_A}, R_{j_B}) \in \mathcal{B}_E \cap \left([0, R^{\text{BC}}_{m_A}(\theta)] \times [0, R^{\text{BC}}_{m_B}(\theta)]\right).$$

The information rates are chosen as

$$R_{w_k} = R^{\text{BC}}_{m_k}(\theta) - R_{j_k} - \frac{\epsilon_1}{2},$$

for $k \in \{A, B\}$. Note that this choice ensures

$$R_{m_k} = R_{w_k} + R_{j_k} < R^{\text{BC}}_{m_k}(\theta), \tag{47}$$

and hence, by the Gaussian broadcast channel achievability result [17], there exists a broadcast code from S to A and B such that A can decode (M_A, M_B) and B can decode M_B, with arbitrarily small error probability.

The relays operate as follows: A discards message M_B and encodes M_A only, while B encodes M_B. The encoding function used at k is a randomly, uniformly generated mapping from $\{1, \ldots, 2^{TR_{m_k}}\}$ to the set of X_k-typical sequences of length T. This code is generated once at random, and is fixed and deterministic thereafter. It can be shown, similarly to the proof of Lemma 4, that the two properties

$$(R_{m_A}, R_{m_B}) \in \mathcal{R}^{\text{MAC}}_D$$

and

$$(R_{j_A}, R_{j_B}) \in \mathcal{R}^{\text{MAC}}_E$$

imply that

$$\text{E}\Bigg[\mathbf{P}(\text{wrong decoding of } (M_A, M_B) \text{ at } D) + \frac{1}{2^{T(R_{w_A}+R_{w_B})}} \sum_{w_A=1}^{2^{TR_{w_A}}} \sum_{w_B=1}^{2^{TR_{w_B}}} \mathbf{P}(\tilde{A}_{w_A,w_B})\Bigg] \le \epsilon_0,$$

where $\tilde{A}_{w_A,w_B}$ is the event that E wrongly decodes (J_A, J_B) given that $(W_A, W_B) = (w_A, w_B)$ is available at E. From this, using an adaption of Fano's inequality similar to Lemma 3, it follows that there exists a code for which the error probability at D can be made small, and for which

$$H(J_A, J_B | W_A, W_B, \mathbf{Y}_E) \le \epsilon_2.$$

Now, we are ready to bound the equivocation at E. From Lemma 2, we have

$$\begin{aligned} H(W|\mathbf{Y}_E) &= H(W_A, W_B|\mathbf{Y}_E) \\ &\ge H(\mathbf{X}_A, \mathbf{X}_B) - I(\mathbf{X}_A, \mathbf{X}_B; \mathbf{Y}_E) \\ &\quad - H(\mathbf{X}_S | W_A, W_B, \mathbf{Y}_E) \\ &= H(M_A, M_B) - I(\mathbf{X}_A, \mathbf{X}_B; \mathbf{Y}_E) \\ &\quad - H(\mathbf{X}_S | W_A, W_B, \mathbf{Y}_E) \\ &= T\big(R^{\text{BC}}_{m_A}(\theta) + R^{\text{BC}}_{m_B}(\theta)\big) \\ &\quad - I(\mathbf{X}_A, \mathbf{X}_B; \mathbf{Y}_E) - T\epsilon_1 - T\epsilon_2 \\ &\ge T\big(R^{\text{BC}}_{m_A}(\theta) + R^{\text{BC}}_{m_B}(\theta)\big) \\ &\quad - T\frac{1}{2}\log(1 + h^2_{AE} + h^2_{BE}) - T\epsilon_1 - T\epsilon_2, \end{aligned}$$

Where we upper bounded $I(\mathbf{X}_A, \mathbf{X}_B; \mathbf{Y}_E)$ by the sum-rate capacity of the multiple access channel from (A, B) to E. Therefore, by choosing ϵ_1 and ϵ_2 small enough, the equivocation $\frac{1}{T}H(W|\mathbf{Y}_E)$ can be made arbitrarily close to $R_D - R_E$, hence we have perfect secrecy.

F. The Gaussian Diamond Network with Noise Insertion

In this section, we guarantee a certain secrecy rate in the Gaussian diamond network (Definition 5) when the relay nodes A and B are allowed to insert noise. In particular, both relay nodes are permitted to select a junk message of arbitrary rate uniformly at random before each transmission of a block and to use this junk message during the encoding. The following theorem states an achievable secrecy rate under this additional assumption.

Theorem 4: In the Gaussian diamond network with noise insertion,

$$R_s \ge \max\left[R_{m_A} + R_{m_B} - \frac{1}{2}\log(1 + h^2_{AE} + h^2_{BE})\right], \tag{48}$$

where the indicated maximization is over all $(\theta, R_{m_A}, R_{m_B})$ for which

$$(R_{m_A}, R_{m_B}) \in \mathcal{R}^{\text{MAC}}_D$$

and $\exists (R_{j_A}, R_{j_B}) \in \mathcal{B}_E$ such that

$$(R_{m_A}, R_{m_B}) \in [R_{j_A}, R_{j_A} + R^{\text{BC}}_{m_A}(\theta)] \times [R_{j_B}, R_{j_B} + R^{\text{BC}}_{m_B}(\theta)],$$

where $\mathcal{B}_E$ is defined in (46).

The proof of Theorem 4 is very similar to that of Theorem 1. The perfect secrecy rate is achievable by the same code construction, with the difference that the junk messages J_A and J_B are generated at A and B, respectively. Thanks to this, the broadcast channel between S and (A, B) can be fully used to send the information messages W_A and W_B. The secrecy rate is then limited by the capacity of this broadcast channel, as well as by the amount of secrecy the two multiple access channels (from (A, B) to D and to E) can provide. Note that this result resembles the "noise forwarding" strategy for the relay network in [10]. However, our scheme for the Gaussian diamond network can be extended to a number of different networks, since we are in general interested in results that guarantee secrecy for any network.

G. A Numerical Example for the Gaussian Diamond Network

In this section, we provide a numerical example for the achievable perfect secrecy rates in Theorems 1 and 4.

Let $h_{SA} = h_{AD} = 2$ be the stronger channel gains. All other channels gains are $h_{SB} = h_{BD} = h_{AE} = h_{BE} = 1$. The multiple access channel (MAC) capacity regions are then given by

$$\begin{aligned} \mathcal{R}^{\text{MAC}}_D = \{(R_A, R_B) &: \\ R_A &< \frac{1}{2}\log 5 \simeq 1.16 \text{ bits}, \\ R_B &< \frac{1}{2}\log 2 = 0.5 \text{ bits}, \\ R_A + R_B &< \frac{1}{2}\log 6 \simeq 1.29 \text{ bits}\} \end{aligned}$$

and

$$\begin{aligned}\mathcal{R}_E^{\text{MAC}} &= \{(R_A, R_B) : \\ R_A &< \frac{1}{2}\log 2 = 0.5 \text{ bits}, \\ R_B &< \frac{1}{2}\log 2 = 0.5 \text{ bits}, \\ R_A + R_B &< \frac{1}{2}\log 3 \simeq 0.79 \text{ bits}\}\end{aligned}$$

These two capacity regions are shown in Figure 8.

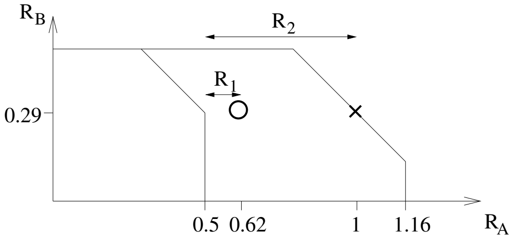


Fig. 8. The MAC capacity regions $\mathcal{R}_E^{\text{MAC}}$ (small pentagon) and $\mathcal{R}_D^{\text{MAC}}$ (large pentagon). The overall rate pair (R_{m_A}, R_{m_B}) is marked by a cross and a circle for the strategies with and without noise insertion, respectively. The achievable secrecy rate without noise insertion (R_1) and with noise insertion (R_2) are indicated. The transmit power as well as the noise variance are 1, and the channel gains are as given in the beginning of this section.

1) Case without Noise Insertion: If the relay nodes do not insert noise, we compute the perfect secrecy rate given by Theorem 1, *i.e.*, we maximize

$$\begin{aligned}&R_{m_A}^{\text{BC}}(\theta) + R_{m_B}^{\text{BC}}(\theta) = \\ &\quad \frac{1}{2}\log\big(1 + 4(1-\theta)\big) + \frac{1}{2}\log\Big(1 + \frac{\theta}{1 + (1-\theta)}\Big)\end{aligned}$$

under the conditions that $(R_{m_A}^{\text{BC}}(\theta), R_{m_B}^{\text{BC}}(\theta)) \in \mathcal{R}_D^{\text{MAC}}$ and that $\mathcal{B}_E \cap \big([0, R_{m_A}^{\text{BC}}(\theta)] \times [0, R_{m_B}^{\text{BC}}(\theta)]\big)$ should be non-empty. We find that the optimal value of θ is $\theta^{\text{opt}} \simeq 0.66$, and

$$\begin{aligned}R_{m_A} &= R_{m_A}^{\text{BC}}(\theta^{\text{opt}}) \simeq 0.62 \text{ bits}, \\ R_{m_B} &= R_{m_B}^{\text{BC}}(\theta^{\text{opt}}) \simeq 0.29 \text{ bits}.\end{aligned}$$

This point is marked with a circle in Figure 8. The achievable perfect secrecy rate is

$$\begin{aligned}R_1 &= R_{m_A} + R_{m_B} - \frac{1}{2}\log(1 + h_{AE}^2 + h_{BE}^2) \\ &\simeq 0.62 + 0.29 - 0.79 = 0.12 \text{ bits}. \qquad (49)\end{aligned}$$

2) Case with Noise Insertion: For the case when noise insertion at the relay nodes is permitted, we compute the perfect secrecy rate given by Theorem 4. It is easy to see that by choosing for instance $R_{j_A} = 0.5$ and $R_{j_B} = 0.29$, we can achieve an overall rate of

$$\begin{aligned}R_{m_B} &\simeq 0.29 \text{ bits} \\ R_{m_A} &= \frac{1}{2}\log(1 + h_{AD}^2 + h_{BD}^2) - R_{m_B} \\ &\simeq 1.29 - 0.29 = 1 \text{ bit}.\end{aligned}$$

This point is marked with a cross in Figure 8. In this case, the achievable perfect secrecy rate is

$$\begin{aligned}R_2 &= R_{m_A} + R_{m_B} - \frac{1}{2}\log(1 + h_{AE}^2 + h_{BE}^2) \\ &\simeq 1 + 0.29 - 0.79 = 0.5 \text{ bits}.\end{aligned}$$

This secrecy rate is considerably larger than (49). The two achievable secrecy rates R_1 and R_2 are also shown in Figure 8. Note that since all the rates are measured in bits, all logarithms are base 2.