# Interactive Secret Key Generation over Reciprocal Fading Channels

Ashish Khisti
Dept. of Electrical and Computer Engineering
University of Toronto
Toronto, ON, Canada
Email: akhisti@comm.utoronto.ca

***Abstract*—We study a two-terminal secret-key generation problem over a two-way, approximately reciprocal, block-fading channel. The channel gains between the legitimate terminals are not revealed to any terminal, whereas the channel gains of the eavesdropper are revealed perfectly to the eavesdropper. We study a separation based scheme that involves a training phase followed by a communication phase within each block. The training phase generates correlated estimates of the channel state sequence between the two terminals, whereas the communication phase generates correlated source sequences between the two terminals. A portion of the secret-key is generated from the correlated channel state sequences by creating omniscience between the legitimate terminals and the remainder of the secret-key is generated from the correlated source sequences. An upper bound on the secret-key capacity is also established by reducing the setup to a coherent channel by providing genie-aided side information to the receivers. We observe that the diffrence between the upper and lower bound decreases as $\frac{1}{T}$ in the high signal-to-noise-ratio (SNR) regime. Numerical results indicate that the proposed scheme achieves significant gains over the commonly used training-only schemes even for moderate SNR.**

## I. Introduction

Channel reciprocity is an inherent feature in time-division-duplex systems. In recent years there has been a growing interest in using reciprocity for secret-key generation. Indeed a number of experimental as well as analytical studies [1]–[6], demonstrate its viability in various wireless systems. However the associated information-theoretic limits still remain to be established.

While the information theoretic framework for secure communication was established by Shannon [7], the problem of secret-key generation from common randomness between two legitimate terminals has been studied relatively recently [8], [9]. The common randomness can arise due to observations of correlated source signals or due to communication over a noisy channel. In the *source model*, the legitimate terminals, say $A$ and $B$, and an eavesdropping terminal, say $E$, observe $N$ i.i.d. copies of random variables $x_A$, $x_B$ and $x_E$ respectively, which are sampled from a joint distribution $p(x_A, x_B, x_E)$. In addition, a public channel of unlimited capacity is also available. Terminals $A$ and $B$ are required to generate a shared secret-key at the end of the communication which must be concealed from $E$. In the *channel model*, terminal $A$ communicates with terminal $B$ (and is eavesdropped by terminal $E$) over a (one-way) discrete memoryless broadcast channel. Note that this model can be reduced to the source model by transmitting i.i.d. source symbols over the channel. This method is referred to as *source-emulation* and is known to be optimal in some special cases. See also e.g., [10]–[14] and references therein for further extensions on secret-key generation.

There has been a growing interest in applying the secret-key generation framework in wireless fading channels. See e.g., [15]–[25]. However only reference [18] studies a setup directly relevant to the channel reciprocity systems. The authors study a two-way, non-coherent, block fading channel with perfect reciprocity in uplink and downlink. Under the restricted class of training-only scheme, the authors establish the maximum secret-key rate. Furthermore they also study a two-phase scheme where part of the coherence block is used for training and the remainder is used for secure-message transmission using a wiretap code. However the message transmission should be done without any rate or power allocation, which significantly restricts the achievable rate.

In the present paper we study a setup similar to [18] but assume only *imperfect* reciprocity, i.e., we assume that the channel gains in uplink and downlink are correlated, but the correlation coefficient is not unity. In general achieving per-

This work was supported by a Natural Science Engineering Research Council (NSERC) Discovery Grant.

fect channel reciprocity in baseband is challenging because different terminals use different I/Q mixers, amplifiers and path lengths in the RF chains. While closed-loop calibration can be performed (see e.g. [26]) such methods can become challenging if the calibration needs to be done in the open air. Hence we believe that our assumption of imperfect reciprocity may perhaps be more realistic in practice. Nevertheless we note that while our lower bound can also be applied to the case of perfect reciprocity, and improves upon [18], our proposed upper bound becomes degenerate in that case.

We propose a separation based transmission scheme,f where part of the coherence block is used for training and the remaining part is used for source-emulation. Unlike message transmission, the use of source emulation yields positive rates even when power adaptation across the channel gains is not performed. This observation allows us to improve the results in [18]. The upper bound is based on a genie aided scheme where we reveal the respective channel gains to the legitimate terminals and then obtain an upper bound in the resulting coherent model. We further show that the gap between our upper and lower bounds decays as $\frac{1}{T}$ in the high signal-to-noise-ratio regime, thus establishing the near optimality of the proposed two-phase scheme in this operating regime. To our knowledge this provides the first asymptotic capacity result on secret-key generation in a two-way non-coherent channel model.

## II. System Model and Main Results

We consider a setup with two legitimate terminals $A$ and $B$ and one eavesdropper $E$. The terminals $A$ and $B$ communicate over a two-way non-coherent wireless channel:

$$y_A(t) = h_{BA}(t)x_B(t) + n_A(t) \quad (1)$$

$$y_B(t) = h_{AB}(t)x_A(t) + n_B(t) \quad (2)$$

$$z_{AE}(t) = g_{AE}(t)x_A(t) + n_{AE}(t) \quad (3)$$

$$z_{BE}(t) = g_{BE}(t)x_B(t) + n_{BE}(t) \quad (4)$$

where $t \in \{1, \ldots, N\}$ denotes the time index, $y_A(t)$ and $y_B(t)$ denote the output symbols at terminals $A$ and $B$ and $\{z_{AE}(t), z_{BE}(t)\}$ denote the output symbols at terminal $E$. The input symbols generated by terminals $A$ and $B$ at time $t$ are denoted by $x_A(t)$ and $x_B(t)$ respectively and are required to satisfy the average power constraints:

$$\frac{1}{N}\sum_{t=1}^{N} E[|x_A(t)|^2] \le P, \quad \frac{1}{N}\sum_{t=1}^{N} E[|x_B(t)|^2] \le P$$

For convenience we assume that the channel gains $(g_{AE}, g_{BE})$ are sampled independently of $(h_{AB}, h_{BA})$. In the proposed block-fading model, the channel gains $(h_{AB}, h_{BA}, g_{AE}, g_{BE})$ are drawn from a distribution $p(h_{AB}, h_{BA})p(g_{AE}, g_{BE})$ once every $T$ consecutive symbols and stay constant over this period i.e., the channel gains remain constant in the interval $t \in [iT + 1, (i + 1)T]$ for $i = 0, 1, \ldots, \frac{N}{T} - 1$. We assume that the channel gains $h_{AB}(t)$ and $h_{BA}(t)$ are not revealed to any of the terminals apriori, whereas the channel gains $(g_{AE}(t), g_{BE}(t))$ are revealed to the eavesdropper terminal.

*Remark 1:* While we assume that the channel gains of the eavesdropper are independent, our upper and lower bounds naturally extend when the channel gains are drawn from a joint distribution $p(h_{AB}, h_{BA}, g_{AE}, g_{BE})$. Note that we assume a full duplex channel model in (1), (2) as both terminals $A$ and $B$ transmit and receive at the same time. The "half-duplex" case where only one of the terminals can transmit at a give time is also of interest, but will not be treated in this paper. Finally note that the eavesdropper observes the transmission from the two terminals $A$ and $B$ over non-interfering links (c.f. (3), (4)). An eavesdropper who observes a superposition of the two signals can only be weaker than the proposed eavesdropper.

We will denote the sequence of length $N$ observed by terminal $A$ by $y_A^N$ and use a similar notation for other input and output channel symbols. The sequence $\bar{h}_{AB}^K$ is used to denote the $K = \frac{N}{T}$ unique coefficients on the channel between $A$ to $B$ and a similar notation is used for the other channel gains. However we will also use the notation $h_{AB}(i)$ to denote the channel gain at time $i$, which maps to the $\lfloor \frac{i}{T} \rfloor + 1$ index in the sequence $\bar{h}_{AB}^K$. We will use the notation $\mathbf{z}(t) = \{z_{AE}(t), z_{BE}(t)\}$ to denote both the output symbols at terminal $E$ and $\mathbf{g}(i) \triangleq (g_{AE}(i), g_{BE}(i))$.

*Definition 1 (Secret-Key Capacity):* A feasible secret-key generation protocol is defined as follows. Terminals $A$ and $B$ sample independent random variables $m_A$ and $m_B$ from a product distribution $p(m_A)p(m_B)$. At time $t$, terminals $A$ and $B$ generate the symbols $x_A(t) = f_A(m_A, y_A^{t-1})$ and $x_B(t) = f_B(m_B, y_B^{t-1})$. At the end of $N$ channel uses, the terminals $A$ and $B$ generate secret keys $k_A$ and $k_B$ respectively using the functions $k_A = \mathcal{K}_A(y_A^N, m_A)$ and $k_B = \mathcal{K}_B(y_B^N, m_B)$. We require that $\Pr(k_A \neq k_B) \le \varepsilon_N$ and furthermore $\frac{1}{N} I(k_A; \mathbf{z}^N, \mathbf{g}^K) \le \varepsilon_N$ for some sequence $\varepsilon_N$ that goes to zero as $N \to \infty$. The largest achievable rate $R = \frac{1}{N} H(k_A)$ is denoted as the *secret-key capacity*.

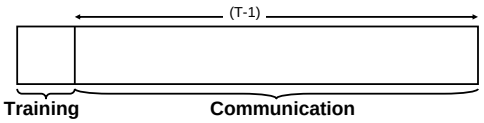


Fig. 1. Separation-based approach for secret-key generation. In each coherence block of length $T$ symbols, the first symbol is dedicated to training and a power $P_1$ is used. The remaining $T-1$ symbols are used for source emulation. Each symbol is sampled i.i.d. from $\mathcal{CN}(0, P_2)$ in this phase.

## III. Main Results

We establish upper and lower bounds on the secret-key capacity for the two-way non-coherent channel model in section II.

*Theorem 1:* An upper-bound on the secret-key capacity of the two-way non-coherent public fading channel is given by $C \leq R^+$ where

$$\begin{aligned} R^+ = &\frac{1}{T} I(h_{AB}; h_{BA}) + \\ &\max_{\{P(h_{AB})\} \in \mathcal{P}_{AB}} I(x_A; y_B | z_{AE}, g_{AE}, h_{AB}) \\ + &\max_{\{P(h_{BA})\} \in \mathcal{P}_{BA}} I(x_B; y_A | z_{BE}, g_{BE}, h_{BA}), \end{aligned} \quad (5)$$

where the above expression is evaluated for

$$x_A \sim \mathcal{CN}(0, P(h_{AB})),\ x_B \sim \mathcal{CN}(0, P(h_{BA})),$$

and the set $\mathcal{P}_{AB}$ is the set of all non-negative power allocation functions $P_{h_{AB}}$ that satisfy the average power constraint $E[P(h_{AB})] \leq P$ and the set $\mathcal{P}_{BA}$ is defined similarly.

*Remark 2:* The upper bound expression (5) in Theorem 1 has a natural interpretation. The term $\frac{1}{T} I(h_{AB}; h_{BA})$ is the contribution arising from reciprocal channel gains. The conditional mutual information terms are the contributions arising from the fading channel model in forward and reverse links respectively, when the fading gains $(h_{AB}, h_{BA})$ are revealed to the eavesdropper. The upper bound is obtained by taking the sum of these three terms.

Our proposed coding scheme involves a separation based approach. As shown in Fig. 1, the first symbol in each coherence block is reserved for sending a training symbol and a total power of $P_1$ is used in this phase. This allows the terminals $B$ and $A$ to generate linear minimum mean squared error estimates of the channel gains $h_{AB}$ and $h_{BA}$ respectively i.e.,

$$\hat{h}_{AB} = \alpha h_{AB} + n_{AB} \quad (6)$$

$$\hat{h}_{BA} = \alpha h_{BA} + n_{BA}, \quad (7)$$

where $\alpha = \frac{P_1}{P_1+1}$ and $n_{AB}$ and $n_{BA}$ both have the distribution $\mathcal{CN}(0, 1-\alpha)$ and are independent of the channel gains $h_{AB}$ and $h_{BA}$ respectively.

For the reminder of coherence block, both the terminals transmit i.i.d. symbols from $\mathcal{CN}(0, P_2)$ to generate correlated source sequences. At the end of this communication secret-keys are generated from the estimated channel sequences and the correlated source sequences.

We first provide a lower bound by assuming that the public-messages for secret-key generation are transmitted over an external public discussion channel. The resulting rate-expression is simpler and has a form similar to (5).

*Theorem 2:* An achievable secret-key rate when an additional public discussion channel of arbitrarily high rate is available communication is given by

$$\begin{aligned} R^-_{PD} = &\frac{1}{T} I(\hat{h}_{AB}; \hat{h}_{BA}) + \\ &\frac{T-1}{T} I(x_A; y_B | \hat{h}_{AB}, \hat{h}_{BA}, z_{AE}, g_{AE}) + \\ &\frac{T-1}{T} I(x_B; y_A | \hat{h}_{AB}, \hat{h}_{BA}, z_{BE}, g_{BE}), \end{aligned} \quad (8)$$

where we evaluate the expression for $x_A \sim \mathcal{CN}(0, P_2)$ and $x_B \sim \mathcal{CN}(0, P_2)$ and $\hat{h}_{AB}$ and $\hat{h}_{BA}$ are specified in (6) and (7) respectively and the variables $P_1$ and $P_2$ are non-negative and satisfy the relation

$$P_1 + (T-1)P_2 \leq TP. \quad (9)$$

*Remark 3:* The expression in (8) has a form that is very similar to the upper bound expression in (5). It differs in the following respects: (i) the channel gains $h_{AB}$ and $h_{BA}$ are replaced by their estimates $\hat{h}_{AB}$ and $\hat{h}_{BA}$ respectively; (ii) the conditional mutual information terms are additionally scaled by a factor of $\left(1 - \frac{1}{T}\right)$; (iii) power allocation accross the channel gains is not performed in the lower bound expression.

An explicit evaluation of the lower bound when that the channel gains $h_{AB}$ and $h_{BA}$ are jointly Gaussian, zero mean random variables with a cross-correlation of $\rho$ is as follows.

*Proposition 1:* An achievable secret-key rate when terminals $A$ and $B$ have access to a public discussion channel is given by:

$$R^-_{PD} = \frac{1}{T} R^-_{P,T} + \frac{T-1}{T} \left( R^-_{P,AB} + R^-_{P,BA} \right) \quad (10)$$

where the expressions for $R^-_{P,T}$, $R^-_{P,AB}$ and $R^-_{P,BA}$ are given by:

$$R^-_{P,T} = -\log\left(1 - \alpha^2 \rho^2\right) \quad (11)$$

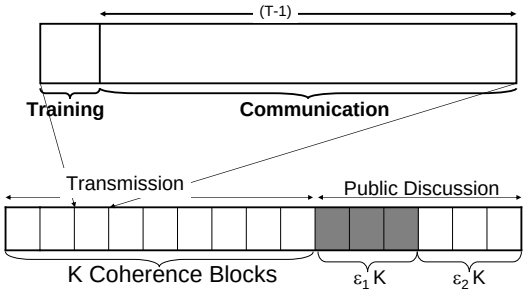


Fig. 2. Extension of the proposed coding scheme in absence of public discussion. A total of $K$ coherence blocks are used for training and source emulation. Thereafter a total of $\varepsilon_1 K$ coherence blocks are used for transmission of the public message associated with the channel sequences and another $\varepsilon_2 K$ coherence blocks are used for the transmission of the public message associated with the source sequences.

$$R^-_{P,AB} = E\left[\log\left(1+\frac{P_2|h_{AB}|^2}{1+P_2|g_{AE}|^2}\right)\right] - \log\left(1+\frac{P_2}{1+P_1}\right) \quad (12)$$

$$R^-_{P,BA} = E\left[\log\left(1+\frac{P_2|h_{BA}|^2}{1+P_2|g_{BE}|^2}\right)\right] - \log\left(1+\frac{P_2}{1+P_1}\right). \quad (13)$$

□

We next consider the case when the public discussion channel is not present. In this scenario, we need to use the communication channel for sending the public messages during the key-generation step. As shown in Fig. 2, we accomplish this by using additional coherence blocks for sending the public messages of the channel and source sequences. We also need to quantize the source sequences to satisfy the rate constraint imposed by the channel.

*Theorem 3:* An achievable secret-key rate in the absence of public discussion for the two-way reciprocal fading channel is:

$$R^- = \frac{1}{1+\varepsilon_1+\varepsilon_2}\left(\frac{1}{T}R_{\rm I} + \frac{T-1}{T}R_{\rm II}\right). \quad (14)$$

We provide the expressions for $R_{\rm I}$ and $R_{\rm II}$ below, both of which depend on $\varepsilon_1$ and $\varepsilon_2$.

$$R_{\rm I} = I(u_A;\hat{h}_{BA}) + I(u_{BA};\hat{h}_{AB}) - I(u_{AB};u_{BA}) \quad (15)$$

where the random variables $u_{AB}$ and $u_{BA}$ satisfy the Markov chain

$$u_{AB} \to \hat{h}_{AB} \to \hat{h}_{BA} \to u_{BA}, \quad (16)$$

and the rate constraints

$$I(u_{AB};\hat{h}_{AB}|\hat{h}_{BA}) \le (T-1)\varepsilon_1 R_c(P), \quad (17)$$

$$I(u_{BA};\hat{h}_{BA}|\hat{h}_{AB}) \le (T-1)\varepsilon_1 R_c(P), \quad (18)$$

where $R_c(P)$ is an achievable rate for the non-coherent block fading channel [27].

The rate expression for $R_{\rm II}$ is expressed as $R_{\rm II} = R^-_{\rm AB} + R^-_{\rm BA}$, where

$$R^-_{\rm AB} = I(v_B;x_A|u) - I(v_B;z_{AE},g_{AE}|u), \quad (19)$$

$$R^-_{\rm BA} = I(v_A;x_B|u) - I(v_A;z_{BE},g_{BE}|u) \quad (20)$$

where the random variable $u \triangleq (u_{AB}, u_{BA})$. The random variables $v_A$ and $v_B$ in (20) and (19) satisfy the following Markov Conditions

$$v_A \to y_A \to (u, x_B),\ v_B \to y_B \to (u, x_A). \quad (21)$$

as well as the rate constraints:

$$I(v_A;y_A|u,x_B) \le \varepsilon_2 R_c(P), \quad (22)$$

$$I(v_B;y_B|u,x_A) \le \varepsilon_2 R_c(P). \quad (23)$$

□

We further evaluate the rate expression in Theorem 3 for a Gaussian test channel:

$$u_{AB} = \hat{h}_{AB} + q_{AB}, \quad u_{BA} = \hat{h}_{BA} + q_{BA} \quad (24)$$

where $q_{AB}, q_{BA} \sim \mathcal{CN}(0, Q_1)$ are Gaussian random variables independent of all other variables. Similarly we let

$$v_A = y_A + w_A, \quad v_B = y_B + w_B \quad (25)$$

where $w_A, w_B \sim \mathcal{CN}(0, Q_2)$ are Gaussian random variables independent of all other variables.

*Proposition 2:* An achievable secret-key rate in the absence of public discussion for the two-way reciprocal fading channel is given by:

$$R^- = \frac{1}{1+\varepsilon_1+\varepsilon_2}\left\{\frac{1}{T}R_{\rm I}(\varepsilon_1,P_1) + \frac{T-1}{T}R_{\rm II}(\varepsilon_2,P_2)\right\} \quad (26)$$

where $P_1$ and $P_2$ are non-negative and satisfy (9) and $\varepsilon_1$ and $\varepsilon_2$ are non-negative constants that will be specified in the sequel. The rate expressions $R_{\rm I}$ and $R_{\rm II}$ are as follows.

$$R_{\rm I}(\varepsilon_1,P_1) = \left\{-2\log\left(1-\frac{\alpha^2\rho^2}{1+\frac{Q_1}{\alpha}}\right)+\log\left(1-\frac{\alpha^2\rho^2}{\left(1+\frac{Q_1}{\alpha}\right)^2}\right)\right\} \quad (27)$$

where $\alpha = \frac{P_1}{1+P_1}$ and $Q_1$ satisfies

$$\log\left(1+\frac{\alpha(1-\alpha^2\rho^2)}{Q_1}\right) \le \varepsilon_1 T R_c(P). \quad (28)$$

The expression for $R_{\rm II}$ satisfies

$$R_{\rm II}(\varepsilon_2,P_2) = \left\{R^-_{AB}(\varepsilon_2,P_2) + R^-_{BA}(\varepsilon_2,P_2)\right\}, \quad (29)$$

$$R_{AB}^{-}(\varepsilon_2,P_2)=E\left[\log\left(1+\frac{P_2|h_{AB}|^2}{(1+Q_2)(1+P_2|g_{AE}|^2)}\right)\right] - \log\left(\frac{\sigma_{AB}^2P_2}{1+Q_2}+1\right), \quad (30)$$

$$R_{BA}^{-}(\varepsilon_2,P_2)=E\left[\log\left(1+\frac{P_2|h_{BA}|^2}{(1+Q_2)(1+P_2|g_{BE}|^2)}\right)\right] - \log\left(\frac{\sigma_{BA}^2P_2}{1+Q_2}+1\right), \quad (31)$$

where

$$\sigma_{AB}^2=\sigma_{BA}^2=1-\frac{\alpha^2}{Q_1+\alpha} \quad (32)$$

and $Q_2$ satisfies

$$\log\left(1+\frac{\sigma_{AB}^2P_2+1}{Q_2}\right)\le\varepsilon_2R_c(P). \quad (33)$$

□

*Remark 4:* The rate achieved in Prop 2 reduces to the rate achieved using public discussion in Prop. 1 when we take $Q_1=Q_2=0$. In particular when $Q_1=0$ note that the expression for $R_{\rm I}$ in (27) immediately reduces to (10). Furthermore (32) reduces to

$$\sigma_{AB}^2=\sigma_{BA}^2=\frac{1}{1+P_1}. \quad (34)$$

Substituting $Q_2=0$, and (34) in (30) and (31) we obtain (12) and (13) respectively. Thus the rate expressions in Prop. 2 are consistent with the expressions in Prop. 1 when quantization noise is introduced in (24) and (25).

We also observe that the upper and lower bounds are close in the high signal-to-noise-ratio (SNR) regime.

*Corollary 1:* In the high SNR regime the upper and lower bounds satisfy the following relation:

$$\lim_{P\to\infty}R^{+}(P)-R_{\rm PD}^{-}(P)=\frac{1}{T}\gamma \quad (35)$$

$$\lim_{P\to\infty}R^{+}(P)-R^{-}(P)=\frac{1}{T}\gamma \quad (36)$$

where $R^+$, $R_{\rm PD}^-$ and $R^-$ are given by (5), (10) and (26) respectively and

$$\gamma\triangleq E\left[\log\left(1+\frac{|h_{AB}|^2}{|g_{AE}|^2}\right)\right]+E\left[\log\left(1+\frac{|h_{BA}|^2}{|g_{BE}|^2}\right)\right] \quad (37)$$

is a constant that only depends on the distributions $p(h_{AB})p(g_{AE})$ and $p(h_{BA})p(g_{BE})$.

We omit the proof of Corollary 1 due to space constraints.

*Numerical Comparison:*

Fig. 3 shows the bounds on secret-key capacity as a function of SNR when the coherence period $T=10$, while Fig. 4 shows the bounds as a function of the coherence period $T$ when SNR = 35 dB. In Fig. 3 we fix the correlation coefficient in uplink and downlink gains to $\rho=0.95$ while in Fig. 4 it is fixed to $\rho=0.99$. In both cases, we observe that the gap between the upper and lower bound (with public discussion channel) is small. Intuitively the range of SNR considered is sufficiently large so that the penalty due to channel estimation vanishes. However the penalty arising from the overhead associated with the transmission of public messages is not negligible in this range and hence there is a significant gap between the upper bound and the lower bound in Theorem 3. The lowermost (horizontal) curve is an upper bound on the rate of a training based scheme $R_{\rm training}=-\frac{1}{T}\log(1-\rho^2)$. It is clear that training-only schemes are in general inefficient even when the overhead information for transmitting public messages is accounted for. In particular, we see that while the rate of the training based scheme decays to zero as the coherence period goes to infinity, whereas the rate of the remaining schemes does not exhibit this phenomenon.

## IV. Proof of Theorem 1

In this section we provide a proof of Theorem 1. From the Fano's inequality and the secrecy constraint we have that

$$NR\le I(k_A;k_B)-I(k_A;\mathbf{z}^N,\bar{\mathbf{g}}^K)+2n\varepsilon_n \quad (38)$$

$$\le I(k_A;k_B|\mathbf{z}^N,\bar{\mathbf{g}}^K)+2n\varepsilon_n \quad (39)$$

where $\varepsilon_n\to0$ as $n\to\infty$. In the following steps the term $\varepsilon_n$ will be suppressed. Since $k_A=f_A(m_A,y_A^N)$ and $k_B=f_B(m_B,y_B^N)$, using the data-processing inequality and the chain rule of mutual information, we have:

$$\begin{aligned}
&NR\le I(m_A,\bar{h}_{BA}^K,y_A^N;m_B,\bar{h}_{AB}^K,y_B^N|\mathbf{z}^N,\bar{\mathbf{g}}^K)\\
&=I(m_A,\bar{h}_{BA}^K,y_A^N;m_B,\bar{h}_{AB}^K,y_B^{N-1}|\mathbf{z}^N,\bar{\mathbf{g}}^K)+\\
&I(m_A,\bar{h}_{BA}^K,y_A^N;y_B(N)|\mathbf{z}^N,\bar{\mathbf{g}}^K,m_B,\bar{h}_{AB}^K,y_B^{N-1})\\
&=I(m_A,\bar{h}_{BA}^K,y_A^{N-1};m_B,\bar{h}_{AB}^K,y_B^{N-1}|\mathbf{z}^N,\bar{\mathbf{g}}^K)+\\
&I(y_A(N);m_B,\bar{h}_{AB}^K,y_B^{N-1}|\mathbf{z}^N,\bar{\mathbf{g}}^K,m_A,\bar{h}_{BA}^K,y_A^{N-1})+\\
&I(m_A,\bar{h}_{BA}^K,y_A^{N-1};y_B(N)|\mathbf{z}^N,\bar{\mathbf{g}}^K,m_B,\bar{h}_{AB}^K,y_B^{N-1})+\\
&I(y_A(N);y_B(N)|\mathbf{z}^N,\bar{\mathbf{g}}^K,m_B,\bar{h}_{AB}^K,y_B^{N-1},m_A,\bar{h}_{BA}^K,y_A^{N-1}).
\end{aligned} \quad (40)$$

We bound each of the four terms in (40).

We can show that:

$$I(m_A,\bar{h}_{BA}^K,y_A^{N-1};m_B,\bar{h}_{AB}^K,y_B^{N-1}|\mathbf{z}^N,\bar{\mathbf{g}}^K)$$

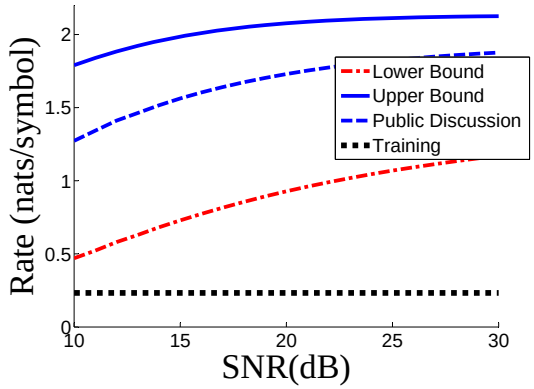


Fig. 3. Bounds on the SK capacity as a function of SNR for a coherence period of $T = 10$.

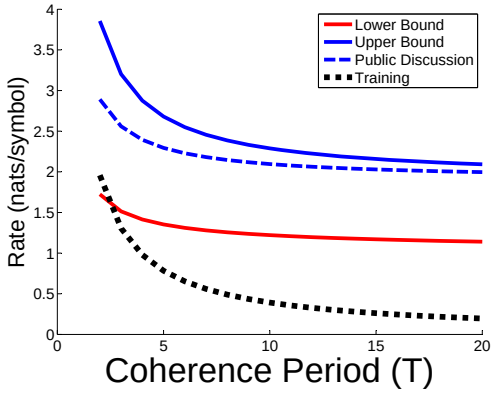


Fig. 4. Bounds on the SK capacity as a function of coherence period for SNR = 30dB.

$$\leq I(m_A, \bar{h}_{BA}^K, y_A^{N-1}; m_B, \bar{h}_{AB}^K, y_B^{N-1} | \mathbf{z}^{N-1}, \bar{\mathbf{g}}^K) \tag{41}$$

by following an argument as in [8]. We omit the details due to space constraint. Since $(y_A(N), z_{BE}(N))$ are independent of all other random variables given $(x_B(N), h_{BA}(N), g_{BE}(N))$ we can show that

$$I(y_A(N); m_B, \bar{h}_{AB}^K, y_B^{N-1} | \mathbf{z}^N, \bar{\mathbf{g}}^K, m_A, \bar{h}_{BA}^K, y_A^{N-1}) \leq I(y_A(N); x_B(N) | z_{BE}(N), g_{BE}(N), h_{BA}(N)). \tag{42}$$

and likewise

$$I(m_A, \bar{h}_{BA}^K, y_A^{N-1}; y_B(N) | \mathbf{z}^N, \bar{\mathbf{g}}^K, m_B, \bar{h}_{AB}^K, y_B^{N-1}) \leq I(x_A(N); y_B(N) | z_{AE}(N), h_{AB}(N), g_{AE}(N)). \tag{43}$$

Finally using the fact that $x_A(N)$ and $x_B(N)$ are functions of $(m_A, y_{BA}^{N-1})$ and $(m_B, y_{AB}^{N-1})$ respectively we can show that the last term in (40) equals zero. Thus substituting (41), (42) and (43) into (40) we have that

$$\begin{aligned} NR \leq\ & I(m_A, \bar{h}_{BA}^K, y_A^{N-1}; m_B, \bar{h}_{AB}^K, y_B^{N-1} | \mathbf{z}^{N-1}, \bar{\mathbf{g}}^K) \\ & + I(y_A(N); x_B(N) | z_{BE}(N), g_{BE}(N), h_{BA}(N)) \\ & + I(x_A(N); y_B(N) | z_{AE}(N), h_{AB}(N), g_{AE}(N)) \end{aligned} \tag{44}$$

Recursively applying the same steps we have that

$$\begin{aligned} NR \leq\ & I(m_A, \bar{h}_{BA}^K; \bar{h}_{AB}^K, m_B | \bar{\mathbf{g}}^K) + \\ & \sum_{i=1}^N I(x_B(i); y_A(i) | z_{BE}(i), g_{BE}(i), h_{BA}(i)) + \\ & \sum_{i=1}^N I(x_A(i); y_B(i) | z_{AE}(i), g_{AE}(i), h_{AB}(i)) \end{aligned} \tag{45}$$

$$= KI(h_{BA}; h_{AB}) + \sum_{i=1}^N I(x_B(i); y_A(i) | z_{BE}(i), g_{BE}(i), h_{BA}(i)) + \sum_{i=1}^N I(x_A(i); y_B(i) | z_{AE}(i), g_{AE}(i), h_{AB}(i)) \tag{46}$$

where in (46) we use the fact that $(m_A, m_B)$ are mutually independent and independent of the channel gains and furthermore $(\bar{h}_{AB}^K, \bar{h}_{BA}^K)$ are independent of $\bar{\mathbf{g}}^K$. Finally to establish the upper bound (5) suppose that we assign a power $P_i(h_{AB})$ when the channel gain at time $i$ equals $h_{AB}$. Using the fact that a Gaussian input distribution maximizes the conditional mutual information terms in (46) we have:

$$I(x_A(i); y_B(i) | z_{AE}(i), g_{AE}(i), h_{AB}(i)) \leq E\left[\log\left(1 + \frac{P_i(h_{AB})|h_{AB}|^2}{1 + P_i(h_{AB})|g_{AE}|^2}\right)\right]. \tag{47}$$

Thus we have that

$$\sum_{i=1}^N I(x_A(i); y_B(i) | z_{AE}(i), g_{AE}(i), h_{AB}(i)) \leq \sum_{i=1}^N E\left[\log\left(1 + \frac{P_i(h_{AB})|h_{AB}|^2}{1 + P_i(h_{AB})|g_{AE}|^2}\right)\right] \tag{48}$$

$$\leq E\left[\sum_{i=1}^N \log\left(1 + \frac{P_i(h_{AB})|h_{AB}|^2}{1 + P_i(h_{AB})|g_{AE}|^2}\right)\right] \tag{49}$$

$$\leq NE\left[\log\left(1 + \frac{\frac{1}{N}\sum_{i=1}^N P_i(h_{AB})|h_{AB}|^2}{1 + \frac{1}{N}\sum_{i=1}^N P_i(h_{AB})|g_{AE}|^2}\right)\right] \tag{50}$$

$$= NE\left[\log\left(1 + \frac{P(h_{AB})|h_{AB}|^2}{1 + P(h_{AB})|g_{AE}|^2}\right)\right], \tag{51}$$

$$= NI(x_A; y_B | z_{AE}, g_{AE}, h_{AB}) \tag{52}$$

where $P(h_{AB}) \triangleq \frac{1}{N}\sum_{i=1}^N P_i(h_{AB})$ is the average

power allocated when the fading state equals $h_{AB}$ and (50) uses the fact that the function $f(x) = \log\left(1+\frac{ax}{1+bx}\right)$ is a concave function in $x$ and hence Jensen's inequality applies. In a similar fashion we can show that

$$\sum_{i=1}^{N} I(x_B(i); y_A(i)|z_{BE}(i), g_{BE}(i), h_{BA}(i))$$
$$\leq NE\left[\log\left(1+\frac{P(h_{BA})|h_{BA}|^2}{1+P(h_{BA})|g_{BE}|^2}\right)\right]. \quad (53)$$
$$= NI(x_B; y_A|z_{BE}, g_{BE}, h_{BA}) \quad (54)$$

The upper bound (5) follows by substituting in (52) and (54) into (46). This completes the proof of Theorem 1.

## V. Proof of Theorem 2

We first sketch the proof of Theorem 2 where we assume the availability of a public discussion channel. As stated before the coding scheme is a separation based scheme. It divides each coherence block into two parts as follows: (a) **Training**: In each coherence block, the first symbol is used to transmit a pilot symbol:

$$x_A(iT+1) = x_B(iT+1) = \sqrt{P_1}, \quad i = 0, \ldots, \frac{N}{T} - 1. \quad (55)$$

The legitimate receivers $B$ and $A$ use the corresponding output symbols $y_B(iT+1)$ and $y_A(iT+1)$ for estimating the underlying channel gains $\hat{h}_{AB}$ and $\hat{h}_{BA}$ respectively in (6) and (7) respectively.

(b) **Source Emulation**: The remainder of the coherence block is used to transmit i.i.d. Gaussian symbols i.e.,

$$x_A(t) \sim \mathcal{CN}(0, P_2), x_B(t) \sim \mathcal{CN}(0, P_2), \quad \forall t \in [1, N], t \neq iT+1. \quad (56)$$

For our analysis we use the notation $\bar{\mathbf{x}}_A(t)$ to denote the (vector) sequence of $T-1$ transmitted Gaussian symbols in coherence block $t$. and let $\bar{\mathbf{y}}_B(t)$ denote the corresponding output symbol vector in block $t$. We execute this transmission over $K$ coherence blocks. At the end of this we observe that terminal $A$ has an estimate $\hat{h}_{BA}^K$ of the channel sequence, transmitted a source sequence $\bar{\mathbf{x}}_A^K$ on the forward channel and observed $\bar{\mathbf{y}}_A^K$ over the reverse channel. Likewise terminal $B$ has estimated $\hat{h}_{AB}^K$, observed $\bar{\mathbf{y}}_A^K$ and transmitted $\bar{\mathbf{x}}_B^K$. The eavesdropper has been revealed $(g_A^K, g_B^K)$, and observes $(\bar{\mathbf{z}}_A^K, \bar{\mathbf{z}}_B^K)$ over its two channels in (3), (4).

The secret-key generation proceeds in the following phases. In the first phase both the terminals transmit the bin index of their respective channel sequences over the public discussion channel, in order to generate a common knowledge of $(\hat{h}_{AB}^K, \hat{h}_{BA}^K)$. A secret-key of rate

$$R_T^- = \frac{1}{T} I(\hat{h}_{AB}; \hat{h}_{BA}) \quad (57)$$

is generated in this phase. In the second phase terminal $A$ sends the bin index of the sequence $\bar{\mathbf{y}}_A^K$ over the public discussion channel so that terminal $B$ can reconstruct this sequence using $(\bar{\mathbf{x}}_B^K, \hat{h}_{AB}^K, \hat{h}_{BA}^K)$ as side information. A secret-key of rate

$$R_{BA}^- = \frac{T-1}{T}\Big\{ I(y_A; x_B, \hat{h}_{AB}, \hat{h}_{BA}) - I(y_A; z_B, g_B, \hat{h}_{AB}, \hat{h}_{BA})\Big\} \quad (58)$$
$$= \frac{T-1}{T} I(y_A; x_B|\hat{h}_{AB}, \hat{h}_{BA}, z_B, g_B) \quad (59)$$

is achieved. In a similar fashion by binning the sequence $\bar{\mathbf{y}}_B^K$ we can achieve the rate of

$$R_{AB}^- = \frac{T-1}{T} I(y_B; x_A|\hat{h}_{AB}, \hat{h}_{BA}, z_A, g_A). \quad (60)$$

The overall-key rate that is achieved is given by the sum of (57), (59) and (60).

The proof of the secrecy analysis is omitted due to space constraints. The proof of Prop. 1 follows by evaluating or bounding the associated mutual information expression for Gaussian input distributions. We skip the derivations due to space constraints.

## VI. Proof of Theorem 3

In absence of the discussion channel, we use the separation based scheme involving training and communication phases as in (55) and (56) respectively over $K$ coherence blocks. At the end of this we have the same set of source and channel sequences as in the previous section. Thereafter we use $\varepsilon_1 \cdot K$ coherence blocks for generating the secret-key from the channel sequences and another $\varepsilon_2 \cdot K$ coherence blocks for generating the secret-key from the source sequences. Due to this overhead, the total rate achieved is scaled by a factor of $\frac{1}{1+\varepsilon_1+\varepsilon_2}$ as in (26).

In the secret-key generation phase we suitably quantize each of the channel-state and source sequences to satisfy the rate constraint for public messages. Terminal $A$ quantizes the sequence $\hat{h}_{BA}^K$ to a codeword $u_{BA}^K$ using a Gaussian codebook associated with the test channel for $u_{BA}$ in (24). A Wyner-Ziv codebook is applied such that terminal $B$ can recover the codeword sequence given the side information $\hat{h}_{AB}^K$ and the bin-index. The rate of

this codebook must satisfy (18) and the associated secret-key rate is given by

$$R^{-}_{\mathrm{T,A}} = \frac{1}{T} I(u_{BA}; \hat{h}_{AB}). \tag{61}$$

Likewise, terminal $B$ quantizes the sequence $\hat{h}^K_{AB}$ to a codeword $u^K_{AB}$ in a Gaussian codebook associated with the test channel for $u_{AB}$ in (24) and a Wyner-Ziv codebook of rate (17) is applied. The resulting secret-key rate is given by

$$R^{-}_{\mathrm{T,B}} = \frac{1}{T}\Big\{ I(u_{AB}; \hat{h}_{BA}) - I(u_{AB}; u_{BA}) \Big\}. \tag{62}$$

The rate-expression for $R_{\mathrm{I}}$ in (15) follows by combining the contributions from (19) and (20). To generate secret-key from the source sequences, terminal $A$ maps the sequence $\bar{\mathbf{y}}^K_A$ to a codeword $\mathbf{v}^K_A$ using the test channel for $v_A$ in (25). A Wyner-Ziv codebook is applied to generate a bin index of $\mathbf{v}^K_A$, such that terminal $B$ can use it and the side information sequence $(\bar{\mathbf{x}}^K_B, u^K)$ to decode $\mathbf{v}^K_A$. The rate of the Wyner-Ziv codebook must satisfy the rate constraint (22) and the associated secret-key rate is given by (19). Likewise the Wyner-Ziv codebook on the reverse channel must satisfy the rate constraint (23) and the associated rate of the secret-key is given by (20)

We omit the secrecy analysis and a more detailed description of the proposed scheme due to space constraints.

The proof of Prop. 2 follows by explicitly evaluating or bounding the mutual information terms in Theorem 3 and will not be included due to space constraints.

## REFERENCES


[1] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radiotelepathy: extracting a secret key from an unauthenticated wireless channel," in *14th ACM international conference on Mobile computing and networking*, 2008, pp. 128—139.

[2] R. Wilson, D. Tse, and R. A. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *IEEE Trans. Information Forensics and Security*, vol. 3, no. 2, pp. 364–375, 2007.

[3] A. A. Hassan, W. E. Stark, J. E. Hershey, and S. Chennakeshu, "Cryptographic key agreement for mobile radio," *Digital Signal Processing*, vol. 6, no. 2, pp. 207–212, 1996.

[4] C. Ye, A. Reznik, and Y. Shah, "Extracting secrey from jointly Gaussian random variables," in *Proc. Int. Symp. Inform. Theory*, Seattle, WA, June 2006.

[5] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, "Robust key generation from signal envelopes in wireless networks," in *14th ACM conference on Computer and communications security*, 2007, pp. 401–410.

[6] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka, "Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels," *IEEE Transactions on Antennas and Propagation,*, vol. 53, no. 11, pp. 3776—3784, 2005.

[7] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[8] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory*, vol. 39, pp. 733–742, Mar. 1993.

[9] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography – Part I: Secret sharing," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1121–1132, July 1993.

[10] I. Csiszar and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, 2004.

[11] A. A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals - part I: Source model," *IEEE Trans. Inform. Theory*, vol. 56, no. 8, pp. 3973–3996, 2010.

[12] ——, "Information-theoretic key agreement of multiple terminals - part II: Channel model," *IEEE Trans. Inform. Theory*, vol. 56, no. 8, pp. 3997–4010, 2010.

[13] V. Prabhakaran, K. Eswaran, and K. Ramchandran, "Secrecy via sources and channels," *IEEE Trans. Inform. Theory*, submitted, Nov 2009. [Online]. Available: http://www.ifp.illinois.edu/~vinodmp/publications/Secrecy09.pdf

[14] A. Khisti, S. N. Diggavi, and G. W. Wornell, "Secret-key agreement using asymmetry in channel state information," in *Proc. Int. Symp. Inform. Theory*, Seoul, Korea, June 2009.

[15] ——, "Secret-key agreement with channel state information at the transmitter," *IEEE Trans. Information Forensics and Security*, vol. 6, no. 3, 2011.

[16] A. Khisti and S. N. Diggavi, "A remark on secret-key generation over correlated fading channels," in *Globecom, Workshop on Physical Layer Security*, Houston, TX, 2011.

[17] A. Khisti, "Secret-key agreement over wiretap channels with transmitter side information," in *European Wireless*, Lucca, Italy, Apr. 2010.

[18] L. Lai, Y. Liang, and H. V. Poor, "A unified framework for key agreement over wireless fading channels," *IEEE Trans. Information Forensics and Security*, vol. 7, no. 2, pp. 480–490, Apr. 2012.

[19] L. Lai, Y. Liang, and W. Du, "Cooperative key generation in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1578–1588, 2012.

[20] A. Agrawal, Z. Rezki, A. Khisti, and M. S. Alouini, "Noncoherent secret-key agreement with public discussion," *IEEE Trans. Information Forensics and Security*, vol. 6, no. 3, pp. 565–574, 2011.

[21] Y. Abdallah, M. A. Latif, M. Youssef, A. Sultan, and H. E. Gamal, "Keys through arq: Theory and practice," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3-1, pp. 737–751, 2011.

[22] O. O. Koyluoglu and H. E. Gamal, "Polar coding for secure transmission and key agreement," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1472–1483, 2012.

[23] M. A. Latif, A. Sultan, and H. E. Gamal, "Arq-based secret key sharing," in *ICC*, 2009, pp. 1–6.

[24] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "A broadcast approach to secret key generation over slow fading channels," *CoRR*, vol. abs/1103.3113, 2011.

[25] T. F. Wong and J. M. Shea, "Secret sharing over fast-fading mimo wiretap channels," *Special Issue on Wireless Physical Layer Security, EURASIP J. Wireless Commun. Netw.*, 2009.

[26] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems," in *Future Netw. Mobile Summit*, 2010.

[27] B. Hassibi and B. M. Hochwald, "How much training is needed in a multiple-antenna wireless link?" *IEEE Trans. Inform. Theory*, Apr. 2011.