A Remark on Secret-Key Generation over Correlated Fading Channels

Ashish Khisti
ECE Dept.
Univ. of Toronto, CANADA
akhisti@comm.utoronto.ca

Suhas N. Diggavi
Electrical Engineering Dept.
UCLA, USA
suhas@ee.ucla.edu

***Abstract*— We study secret-key agreement with public discussion over a flat-fading wiretap channel model. The fading gains are correlated across the receivers and sampled independently at each time. Perfect receiver channel state information (CSI) is assumed, whereas a noisy CSI of the main channel is also available to the transmitter. We propose lower and upper bounds on the capacity. Our lower bound is achieved by a coding scheme that involves a separate binning of the receiver CSI sequence and its channel output sequence. In general it improves upon the joint-binning schemes considered in earlier works. Our upper and lower bounds coincide, establishing the capacity, when either the transmitter has no CSI or when the channel gains of the legitimate receiver and the eavesdropper are statistically independent.**

I. Introduction

In recent years there has been a significant interest in developing secret-key agreement protocols over fading channels, see e.g., [1]–[7] and the references therein. In time-division duplex (TDD) wireless systems, a natural reciprocity between uplink and downlink exists, which is clearly a valuable resource for generating a shared secret key. In frequency division duplex (FDD) systems, such a reciprocity does not exist, but *public interaction* between the remote terminals can still be used to generate a shared secret-key that remains concealed from an eavesdropper.

While a significant body of literature exists for practical protocol designs for secret-key generation, surprisingly little attention has been devoted towards understanding information theoretic limits. The pioneering works in [8], [9] introduce a channel-wiretapper model (CW) where the sender and receiver communicate over a wiretap channel. A public discussion channel (of unlimited capacity) is also available for communication. A characterization of the secret-key capacity of the CW model remains open. However it has been solved for the practically important case of *independent noise* channels. When the output symbols at the receiver (y_r) and eavesdroppers (y_e) are conditionally independent given the input symbol x, i.e., $y_\mathrm{r} \leftrightarrow x \leftrightarrow y_\mathrm{e}$ holds then $C = \max_{p(x)} I(x; y_\mathrm{r}|y_\mathrm{e})$. Building upon these results, reference [10] establishes the secret-key capacity for a class of fading channels. The fading coefficients are sampled i.i.d. both in time and across the receivers and the channel gains are revealed the respective receivers. Since the channel gains can be viewed as additional outputs at the receiver [11], the model essentially reduces to a continuous valued and cost constrained extension of the CW model [8], [9]. The secret-key capacity is characterized in an analogous manner and Gaussian inputs are shown to be optimal. Reference [12] studies a non-coherent i.i.d. Rayleigh fading CW model and establishes that (i) the capacity achieving distribution is discrete and (ii) the secret-key capacity remains bounded in the signal-to-noise ratio (SNR) regardless of the number of antennas at each terminal.

While the capacity results in [10], [12] provide useful fundamental limits, they crucially depend on the fading channel gains of the receiver and the eavesdropper being independent. When this condition does not hold, the proposed coding schemes may not be optimal. In realistic scattering environments, correlation between the channel gains could be observed, see e.g., [13], [14]. As such the correlation depends on a number of factors such as the altitude of the base-station, the number of scatteres and the position of the receivers. Secondly the results assume that no channel state information (CSI) is available at the transmitter and are applicable only to FDD systems. In TDD systems, the transmitter may have access to a noisy version of of the legitimate receiver's channel state information (CSI), which again is not considered in earlier works.

In this paper, we first study correlated fading channels with receiver only CSI and establish the secret-key capacity using a two stage scheme where the receiver channel gains are first revealed to the transmitter over the discussion channel. The secret-key generation codebook is then used conditioned on this knowledge at all the terminals. We observe that the capacity achieving technique in [10], [12] that involves joint binning of the receiver output and channel gains is sub-optimal. We then extend these results to the case when the transmitter also has access to a noisy CSI of the legitimate receiver. We propose a natural extension of our two-step coding scheme, an upper bound on the secret-key capacity, as well as the capacity when the channel gains of the receiver and eavesdropper are independent.

II. Channel Model

The channel model is an i.i.d. fading channel model described by

$$\begin{aligned} y_\mathrm{r}(t) &= h_\mathrm{r}(t)x(t) + z_\mathrm{r}(t) \\ y_\mathrm{e}(t) &= h_\mathrm{e}(t)x(t) + z_\mathrm{e}(t) \end{aligned}, \quad t = 1, 2, \ldots, n \tag{1}$$

where the noise random variables $z_\mathrm{r}(t)$ and $z_\mathrm{e}(t)$ are mutually independent and sampled from $\mathcal{CN}(0,1)$ independently for each t. The fading gains $(h_\mathrm{r}(t), h_\mathrm{e}(t))$ are sampled from a joint distribution $p_{h_\mathrm{r},h_\mathrm{e}}(h_r, h_e)$, independently for each t. The input symbols are complex-valued and satisfy an average power constraint $\frac{1}{n}\sum_{t=1}^{n} E[|x(t)|^2] \leq P$. The realizations of $h_\mathrm{r}(t)$ and $h_\mathrm{e}(t)$ are revealed to the legitimate receiver and the eavesdropper. For our numerical results we consider the case of Gaussian fading where h_r and h_e are each zero mean, unit variance, jointly Gaussian random variables with a correlation coefficient of ρ_e.

In addition, we assume that the transmitter is revealed an i.i.d. sequence $h_\mathrm{t}(t)$, which is a noisy version of $h_\mathrm{r}(t)$. The transmitter state h_t satisfies the Markov chain $h_\mathrm{t} \to (x, h_\mathrm{r}) \to (y_\mathrm{r}, y_\mathrm{e}, h_\mathrm{e})$ indicating that the channel outputs at the receiver and eavesdropper are independent of h_t given (x, h_r). In the case of Gaussian channels we let $h_\mathrm{t}(t) = \rho_t h_\mathrm{r}(t) + w(t)$ where $w(t)$ is zero mean Gaussian random variable with variance $1 - \rho_t^2$ and independent of everything else.

III. Main Results

Our main results are as follows.

Theorem 1: For the case of receiver-only CSI i.e., when $h_\mathrm{t} = 0$ the secret-key capacity is given by

$$C = E_{h_\mathrm{r},h_\mathrm{e}}\left[\log\left(1 + \frac{P|h_\mathrm{r}|^2}{1 + P|h_\mathrm{e}|^2}\right)\right] \tag{2}$$

The capacity achieving scheme involves a two-step process. First the receiver reveals h_r^n to all the terminals using the public discussion channel. Thereafter a conditional secret-key generation codebook is used to achieve a rate of

$$C = \max_{p(x)} \{I(x; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{r}; y_\mathrm{e}, h_\mathrm{e}|h_\mathrm{r})\} \tag{3}$$

$$= \max_{p(x)} I(x; y_\mathrm{r}|y_\mathrm{e}, h_\mathrm{r}, h_\mathrm{e}) \tag{4}$$

It is interesting to compare the proposed coding scheme with the joint-binning scheme [10].

Proposition 1: An achievable rate using the joint-binning scheme in [10] for the Gaussian fading channel with receiver-only CSI and a Gaussian input distribution $x \sim \mathcal{CN}(0, P)$ is:

$$R_\mathrm{JB} = E_{h_\mathrm{r},h_\mathrm{e}}\left[\log\left(1 + \frac{P|h_\mathrm{r}|^2}{1 + P|h_\mathrm{e}|^2}\right)\right] + \log(1 - \rho_e^2) \tag{5}$$

if $R_\mathrm{JB} \geq 0$. The rate is zero otherwise.

We note that the loss in (5) with respect to the capacity expression (2) is the $\log(1 - \rho_e^2)$ term. This can be interpreted as the penalty arising from the eavesdropper CSI being correlated with the receiver CSI. The joint-binning scheme loses secret-key bits as the eavesdropper can learn more information about y_r^n, which is jointly binned with h_r^n. In contrast our proposed scheme only bins y_r^n and reveals h_r^n and thus avoids this leakage.

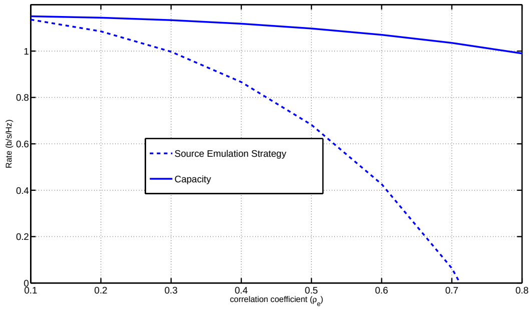

Fig. 1. A comparison of capacity achieving Scheme (2) and joint-binning scheme (5) for SNR = 10 dB and $\rho_t = 0$.

Fig. 1 provides a numerical comparison between the capacity and the joint-binning scheme as a function of the correlation parameter ρ_e. We assume SNR = 10 dB and $\rho_t = 0$. We see that even a small amount of correlation can result in a significant penalty in the joint-binning scheme.

When the transmitter has access to a side-information sequence h_t we have the following results.

Proposition 2: An achievable secret-key rate for the fading-wiretap channel with transmitter CSI is

$$\begin{aligned} R^- = \max_{p_{x|h_\mathrm{t}}} \Big\{ & I(x, h_\mathrm{t}; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r}) \\ & + \max\{I(h_\mathrm{r}; h_\mathrm{t}) - I(h_\mathrm{r}; y_\mathrm{e}, h_\mathrm{e}), 0\}\Big\}. \end{aligned} \tag{6}$$

The secret-key rate is achieved by a natural extension of the capacity achieving scheme with receiver-only CSI. In the first step the legitimate receiver bins the sequence h_r^n so that a secret key of rate $R_1 = \max\{I(h_\mathrm{r}; h_\mathrm{t}) - I(h_\mathrm{r}; y_\mathrm{e}, h_\mathrm{e}), 0\}$ can be achieved. In the second phase the sequence y_r^n is binned using a conditional secret-key generation codebook so that a rate of $R_2 = I(x, h_\mathrm{t}; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r})$ can be achieved. The total secret-key rate is $R^- = R_1 + R_2$.

In contrast, the joint-binning scheme yields a lower rate as stated below:

Proposition 3: An achievable secret-key rate for the fading wiretap channel with transmitter CSI using the joint-binning scheme is:

$$\begin{aligned} R_\mathrm{JB}^- = \max_{p_{x|h_\mathrm{t}}} \Big\{ & I(x, h_\mathrm{t}; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r}) \\ & + I(h_\mathrm{r}; h_\mathrm{t}) - I(h_\mathrm{r}; y_\mathrm{e}, h_\mathrm{e})\Big\} \end{aligned} \tag{7}$$

The joint-binning scheme provides an advantage to the eavesdropper if its state sequence h_e^n is strongly correlated with h_r^n. This is manifested in the fact that the second term $I(h_\mathrm{r}; h_\mathrm{t}) - I(h_\mathrm{r}; y_\mathrm{e}, h_\mathrm{e})$ in (7) becomes negative. The strategy in Prop. 2 discussed earlier alleviates this problem by separately binning h_r^n and y_r^n.

We next state an upper bound on the secret-key rate.

Proposition 4: An upper bound on the secret-key rate for the fading wiretap channel with public discussion is

$$R^+ = \max_{p_{x|h_\mathrm{t}}} I(x, h_\mathrm{t}; y_\mathrm{r}, h_\mathrm{r} | y_\mathrm{e}, h_\mathrm{e}) \tag{8}$$

The above upper bound equals the secret-key capacity if the channel also satisfies $(y_\mathrm{r}, h_\mathrm{r}) \to (x, h_\mathrm{t}) \to (y_\mathrm{e}, h_\mathrm{e})$.

We remark that when $(y_\mathrm{r}, h_\mathrm{r}) \to (x, h_\mathrm{t}) \to (y_\mathrm{e}, h_\mathrm{e})$ holds, the capacity can be achieved using only a joint-binning scheme, even though it is in general sub-optimal (c.f. Prop. 2).

Theorem 2: When the channel gains h_r and h_e are independent, the secret-key capacity for the Gaussian fading wiretap channel is lower and upper bounded by $C^- \leq C \leq C^+$ where

$$C^- = \max_{P(h_\mathrm{t})} E_{h_\mathrm{r}, h_\mathrm{t}, h_\mathrm{e}} \left[\log \left(1 + \frac{P(h_\mathrm{t})|h_\mathrm{r}|^2}{1 + P(h_\mathrm{t})|h_\mathrm{e}|^2} \right) \right] + \log \frac{1}{1 - \rho_t^2} \tag{9}$$

where the maximum is over all power allocation policies $P(h_\mathrm{t})$ that satisfy $E[P(h_\mathrm{t})] \leq P$ and where[1]

$$C^+ = E_{h_\mathrm{r}, h_\mathrm{e}} \left[\log \left(1 + \frac{|h_\mathrm{r}^\dagger h_\mathrm{e}|^2}{|h_\mathrm{e}|^4} \right) \right] + \log \frac{1}{1 - \rho_t^2}. \tag{10}$$

IV. Proof of Main Results

We provide a proof of the main results in this section. In the analysis of our coding schemes we assume that the fading gains h_r and h_e are discrete valued and belong to a set $\{h_1, h_2, \ldots, h_D\}$. We let $p_j = \Pr(h_\mathrm{r} = h_j)$. The result can be extended to continuous valued channel gains using quantization arguments. We omit the details in this paper, but refer the reader to [10], [15] for a similar analysis.

A. Proof of Theorem 1

We first establish that the rate expression in (3) is achievable. In our proposed coding scheme, the sender samples an i.i.d. sequence x^n from the distribution $p_x(\cdot)$ and sends it over n channel uses. The receiver observes $(y_\mathrm{r}^n, h_\mathrm{r}^n)$ whereas the eavesdropper observes $(y_\mathrm{e}^n, h_\mathrm{e}^n)$. At the end of the source transmission the receiver transmits h_r^n over the public discussion channel. At this point all the terminals have access to h_r^n. The sender partitions the sequence x^n into subsequences $(x_1^{n_1}, \ldots, x_D^{n_D})$ where $x_j^{n_j}$ denotes the subsequence of x^n corresponding to the indices where $h_{r,j} = h_j$. Likewise the receiver partitions y_r^n into subsequences $(y_{r1}^{n_1}, \ldots, y_{rD}^{n_D})$. The receiver applies an independent secret-key generation codebook [8], [9] on each of the subsequences $y_{r,j}^{n_j}$ of rate:

$$R_j = I(x; y_\mathrm{r} | h_\mathrm{r} = h_j) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r} | h_\mathrm{r} = h_j) \tag{11}$$

and generates a key k_j. The overall key $k = (k_1, \ldots, k_D)$ has a rate $R = \sum_{j=1}^{D} p_j R_j$ which equals the expression in (3).

To establish the rate in (4) we observe that because the noise variables z_r and z_e are independent we have that $y_\mathrm{r} \leftrightarrow (x, h_\mathrm{r}) \leftrightarrow (y_\mathrm{e}, h_\mathrm{e})$ and hence

$$\begin{aligned} C &= I(x; y_\mathrm{r} | h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r} | h_\mathrm{r}) && (12) \\ &= I(y_\mathrm{e}, h_\mathrm{e}, x; y_\mathrm{r} | h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r} | h_\mathrm{r}) && (13) \\ &= I(x; y_\mathrm{r} | h_\mathrm{r}, y_\mathrm{e}, h_\mathrm{e}). && (14) \end{aligned}$$

Furthermore it follows from [9, Theorem 2] that an upper bound on the secret-key capacity with outputs $(y_\mathrm{r}, h_\mathrm{r})$ and $(y_\mathrm{e}, h_\mathrm{e})$ at the legitimate terminals and eavesdropper respectively is

$$C^+ = \max_{p(x)} I(x; y_\mathrm{r}, h_\mathrm{r} | y_\mathrm{e}, h_\mathrm{e}) = \max_{p(x)} I(x; y_\mathrm{r} | y_\mathrm{e}, h_\mathrm{e}, h_\mathrm{r}), \tag{15}$$

where we use the fact that h_r is independent of (y_e, x) given h_e in the second step. This upper bound coincides with (4).

To establish Theorem 1 it only remains to show that the expression in (4) is maximized by a Gaussian input i.e., $x \sim \mathcal{CN}(0, P)$. Let $p_x(\cdot)$ be any distribution with $E[x^2] = P_1 \leq P$. For each fixed (h_r, h_e), the estimation error of y_r given y_e is

$$\begin{aligned} \sigma^2_{y_\mathrm{r}|y_\mathrm{e}} &= 1 + |h_r|^2 P_1 - \frac{P_1^2 |h_r|^2 |h_e|^2}{1 + P_1 |h_e|^2} && (16) \\ &= 1 + \frac{P_1 |h_r|^2}{1 + P_1 |h_e|^2}. && (17) \end{aligned}$$

Thus we have

$$\begin{aligned} h(y_\mathrm{r} | h_\mathrm{r}, y_\mathrm{e}, h_\mathrm{e}) &= E_{h_\mathrm{r}, h_\mathrm{e}} [h(y_\mathrm{r} | y_\mathrm{e}, h_\mathrm{r} = h_r, h_\mathrm{e} = h_e)] && (18) \\ &\leq E_{h_\mathrm{r}, h_\mathrm{e}} \left[\log 2\pi e \left(1 + \frac{P_1 |h_\mathrm{r}|^2}{1 + P_1 |h_\mathrm{e}|^2} \right) \right] && (19) \\ &\leq E_{h_\mathrm{r}, h_\mathrm{e}} \left[\log 2\pi e \left(1 + \frac{P |h_\mathrm{r}|^2}{1 + P |h_\mathrm{e}|^2} \right) \right] && (20) \end{aligned}$$

where (19) follows from (17) and the fact that a Gaussian input distribution maximizes the differential entropy among all distributions with a fixed variance and the last step follows from the fact that the objective function is increasing in P_1 and so we maximize it by

[1] We use $h_\mathrm{r}^\dagger$ to denote the conjugate of h_r.

setting $P_1 = P$. Thus we have

$$I(x; y_\mathrm{r}|h_\mathrm{r}, y_\mathrm{e}, h_\mathrm{e}) = h(y_\mathrm{r}|h_\mathrm{e}, h_\mathrm{r}, y_\mathrm{e}) - h(y_\mathrm{r}|y_\mathrm{e}, h_\mathrm{r}, h_\mathrm{e}, x) \tag{21}$$

$$= h(y_\mathrm{r}|h_\mathrm{e}, h_\mathrm{r}, y_\mathrm{e}) - h(z_\mathrm{r}) \tag{22}$$

$$= h(y_\mathrm{r}|h_\mathrm{e}, h_\mathrm{r}, y_\mathrm{e}) - \log 2\pi e \tag{23}$$

$$\le E_{h_\mathrm{r}, h_\mathrm{e}}\left[\log\left(1 + \frac{P|h_\mathrm{r}|^2}{1 + P|h_\mathrm{e}|^2}\right)\right] \tag{24}$$

where the last step follows from (20). Since equality holds by selecting a Gaussian input distribution, this complete the proof of Theorem 1.

B. Proof of Prop. 1

The joint-binning scheme proposed in [8], [9] involves joint binning of $(y_\mathrm{r}^n, h_\mathrm{r}^n)$ such that the transmitter can reproduce these sequences with high probability given x^n. The rate that can be achieved is,

$$R_\mathrm{JB} = I(x; y_\mathrm{r}, h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}, h_\mathrm{r}) \tag{25}$$

$$= I(x; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r}) + I(x; h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; h_\mathrm{r}) \tag{26}$$

$$= I(x; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r}) - I(h_\mathrm{e}; h_\mathrm{r}) \tag{27}$$

where the last relation follows from the fact that (x, h_r) are independent and $y_\mathrm{e} \to h_\mathrm{e} \to h_\mathrm{r}$ holds. Evaluating (27) with $x \sim \mathcal{CN}(0, P)$ we have that

$$I(x; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r})$$

$$= h(y_\mathrm{r}|h_\mathrm{r}, y_\mathrm{e}, h_\mathrm{e}) - h(y_\mathrm{r}|h_\mathrm{r}, x) \tag{28}$$

$$= E_{h_\mathrm{r}, h_\mathrm{e}}\left[\log\left(1 + \frac{P|h_\mathrm{r}|^2}{1 + P|h_\mathrm{e}|^2}\right)\right] \tag{29}$$

and using the jointly Gaussian fading model we have

$$I(h_\mathrm{r}; h_\mathrm{e}) = -\log(1 - \rho_e^2). \tag{30}$$

This establishes (5).

C. Proof of Prof. 2

The coding scheme is an extension of the scheme in the proof of Theorem 1. In particular we propose a layered coding scheme as follows:

- The sender samples x_i from the distribution $p_{x|h_\mathrm{t}}(x_i|h_{ti})$ for $i = 1, 2, \ldots, n$ and transmits it at time $t = i$. The receiver and eavesdropper are revealed $(y_{r,i}, h_{r,i})$ and $(y_{e,i}, h_{e,i})$ respectively.
- Upon receiving $(h_\mathrm{r}^n, y_\mathrm{r}^n)$, the receiver applies a Slepian-Wolf code [8], [9] of rate $R_{s,0} = H(h_\mathrm{r}|h_\mathrm{t})$ to h_r^n and transmits the corresponding bin index over the public discussion channel. By virtue of the Slepian-Wolf coding theorem the transmitter is able to recover sequence h_r^n with high probability upon observing h_t^n and the bin index.
- The sender and receiver apply a secret-key agreement codebook [8], [9] to h_r^n to generate a secret key of rate
$$R_0 = \max\left(0, I(h_\mathrm{r}; h_\mathrm{t}) - I(h_\mathrm{r}; h_\mathrm{e}, y_\mathrm{e})\right) \tag{31}$$
If the expression in (31) is zero, no secret-key is produced in this step.
- With the common knowledge of h_r^n between the transmitter and receiver, the sequences x^n, h_t^n and y_r^n are partitioned into D sub-sequences. The sender partitions the sequences (x^n, h_t^n) into D subsequences $\{(x_1^{n_1}, h_{t1}^{n_1}), \ldots, (x_j^{n_j}, h_{tj}^{n_j}), \ldots, (x_{tD}^{n_D}, h_{tD}^{n_D})\}$ where $(x_j^{n_j}, h_{tj}^{n_j})$ corresponds to those indices $i \in [1, n]$ where $h_{r,i} = h_j$. Likewise the receiver partitions y_r^n into $(y_{r,1}^{n_1}, \ldots, y_{r,D}^{n_D})$.
- A separate secret-key generation codebook in [8] is then applied to each the D subsequences and a key k_j of rate
$$R_j = I(x, h_\mathrm{t}; y_\mathrm{r}|h_\mathrm{r} = j) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r} = j) \tag{32}$$
is produced.
- The overall secret-key is obtained by concatenating each of the $D + 1$ keys in the above steps. The secret-key has a rate
$$R = R_0 + \sum_{j=1}^{D} \Pr(h_\mathrm{r} = j) R_j, \tag{33}$$
which reduces to (6).

D. Proof of Prop. 3

A straightforward extension of the joint-binning scheme gives

$$R_\mathrm{JB} = I(x, h_\mathrm{t}; y_\mathrm{r}, h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}, h_\mathrm{r}) \tag{34}$$

$$= I(x, h_\mathrm{t}; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r}) + I(x, h_\mathrm{t}; h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; h_\mathrm{r}) \tag{35}$$

$$= I(x, h_\mathrm{t}; y_\mathrm{r}|h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; y_\mathrm{r}|h_\mathrm{r}) + I(h_\mathrm{t}; h_\mathrm{r}) - I(y_\mathrm{e}, h_\mathrm{e}; h_\mathrm{r}) \tag{36}$$

where the last step follows from the Markov condition $x \leftrightarrow h_\mathrm{t} \leftrightarrow h_\mathrm{r}$.

E. Proof of Prop. 4

In [16, Theorem 4], it is shown that an upper bound on secret-key agreement capacity for the wiretap channel $p_{y_\mathrm{r}, y_\mathrm{e}|h_\mathrm{t}, x}(\cdot)$ with non-causal transmitter CSI h_t^n is given by

$$C \le \max_{p_{x|h_\mathrm{t}}} I(x, h_t; y_\mathrm{r}|y_\mathrm{e}) \tag{37}$$

Following the discussion in [11], [17], the channel with two-sided CSI is equivalent to a channel with transmitter only CSI but with outputs $(h_\mathrm{r}, y_\mathrm{r})$ and $(h_\mathrm{e}, y_\mathrm{e})$ at the legitimate receiver and the eavesdropper respectively. Hence the above upper bound can also be applied to the case of two-sided CSI:

$$C \le \max_{p_{x|h_\mathrm{t}}} I(x, h_t; y_\mathrm{r}, h_\mathrm{r}|y_\mathrm{e}, h_\mathrm{e}) \tag{38}$$

thus establishing (8).

When the Markov condition $(h_{\rm r}, y_{\rm r}) \leftrightarrow (x, h_{\rm t}) \leftrightarrow (y_{\rm e}, h_{\rm e})$ is satisfied, we have

$$
\begin{aligned}
& I(x, h_t; y_{\rm r}, h_{\rm r}|y_{\rm e}, h_{\rm e}) \\
& = I(y_{\rm e}, h_{\rm e}, x, h_t; y_{\rm r}, h_{\rm r}) - I(y_{\rm e}, h_{\rm e}; y_{\rm r}, h_{\rm r}) \qquad (39) \\
& = I(x, h_t; y_{\rm r}, h_{\rm r}) - I(y_{\rm e}, h_{\rm e}; y_{\rm r}, h_{\rm r}) \qquad (40)
\end{aligned}
$$

which equals (34). Thus the capacity can be achieved by a joint-binning scheme in this special case.

F. Proof of Theorem 2

When the channel gains $h_{\rm r}$ and $h_{\rm e}$ are independent, the Markov condition $(h_{\rm r}, y_{\rm r}) \leftrightarrow (x, h_{\rm t}) \leftrightarrow (y_{\rm e}, h_{\rm e})$ is satisfied. Thus we have that

$$C = \max_{p(x|h_{\rm t})} I(x, h_{\rm t}; y_{\rm r}, h_{\rm r}|y_{\rm e}, h_{\rm e}) \tag{41}$$

To establish the lower bound (9) we select $x \sim \mathcal{CN}(0, P(h_{\rm t}))$ and evaluate (7).

$$
\begin{aligned}
& I(x, h_{\rm t}; y_{\rm r}|h_{\rm r}) - I(y_{\rm r}; y_{\rm e}, h_{\rm e}|h_{\rm r}) \\
& = h(y_{\rm r}|h_{\rm r}, y_{\rm e}, h_{\rm e}) - h(y_{\rm r}|h_{\rm r}, x, h_{\rm t})
\end{aligned}
$$

Since $(y_{\rm r}, y_{\rm e})$ are jointly Gaussian random variables it follows that

$$h(y_{\rm r}|h_{\rm r}, y_{\rm e}, h_{\rm e}) \tag{42}$$

$$= E\left[\log 2\pi e\left(1 + |h_{\rm r}|^2 P(h_{\rm t}) - \frac{P^2(h_{\rm t})|h_{\rm r}|^2|h_{\rm e}|^2}{1 + P(h_{\rm t})|h_{\rm e}^2|}\right)\right] \tag{43}$$

$$= E\left[\log 2\pi e\left(1 + \frac{P(h_{\rm t})|h_{\rm r}|^2}{1 + P(h_{\rm t})|h_{\rm e}^2|}\right)\right] \tag{44}$$

Thus using $h(z_{\rm r}) = \log 2\pi e$, we have,

$$I(x, h_{\rm t}; y_{\rm r}|h_{\rm r}) - I(y_{\rm r}; y_{\rm e}, h_{\rm e}|h_{\rm r}) \tag{45}$$

$$= E\left[\log\left(1 + \frac{P(h_{\rm t})|h_{\rm r}|^2}{1 + P(h_{\rm t})|h_{\rm e}^2|}\right)\right] \tag{46}$$

Furthermore since $h_{\rm r}$ and $h_{\rm e}$ are independent and $h_{\rm r} \sim \mathcal{CN}(0, 1)$ and $h_{\rm t}$ are jointly Gaussian with a correlation coefficient of ρ_t

$$I(h_{\rm r}; h_{\rm t}) - I(h_{\rm e}; h_{\rm r}) = h(h_{\rm r}) - h(h_{\rm r}|h_{\rm t}) = -\log(1 - \rho_t^2) \tag{47}$$

Substituting (46) and (47) we obtain the desired lower bounds in (9).

To establish the upper bound note that

$$
\begin{aligned}
R_+ &= I(h_{\rm t}, x; y_{\rm r}, h_{\rm r}|y_{\rm e}, h_{\rm e}) \\
&= I(h_{\rm t}, x; y_{\rm r}|y_{\rm e}, h_{\rm e}, h_{\rm r}) + I(h_{\rm t}, x; h_{\rm r}|y_{\rm e}, h_{\rm e}) \qquad (48)
\end{aligned}
$$

The seco nd term in (48) can be upper bounded by observing that when $h_{\rm r}$ and $h_{\rm e}$ are independent, we have that $(x, y_{\rm e}, h_{\rm e}) \to h_{\rm t} \to h_{\rm r}$ and hence

$$I(h_{\rm t}, x; h_{\rm r}|y_{\rm e}, h_{\rm e}) \le I(h_{\rm t}; h_{\rm r}) = -\log(1 - \rho_t^2). \tag{49}$$

The first term in (48) can be upper bounded as follows

$$
\begin{aligned}
& I(h_{\rm t}, x; y_{\rm r}|y_{\rm e}, h_{\rm e}, h_{\rm r}) = h(y_{\rm r}|y_{\rm e}, h_{\rm e}, h_{\rm r}) - h(z_{\rm r}) \qquad (50) \\
& \le h\left(y_{\rm r} - \frac{h_{\rm r} h_{\rm e}^\dagger}{|h_{\rm e}|^2} y_{\rm e} \,\middle|\, y_{\rm e}, h_{\rm e}, h_{\rm r}\right) - h(z_{\rm r}) \qquad (51) \\
& \le E\left[\log\left(1 + P\frac{|h_{\rm r}^\dagger h_{\rm e}|^2}{|h_{\rm e}|^4}\right)\right] \qquad (52)
\end{aligned}
$$

Note that the upper bound expression in (10) follows by substituting (49) and (52) into (48).

References

[1] A. A. Hassan, W. E. Stark, J. E. Hershey, and S. Chennakeshu, “Cryptographic key agreement for mobile radio,” *Elsevier Digital Signal Processing Magazine*, vol. 6, pp. 207–212, 1996.

[2] C. Ye, A. Reznik, and Y. Shah, “Extracting secrecy from jointly gaussian random variables,” in *Proc. Int. Symp. Inform. Theory*, Jul. 2006.

[3] D. T. R. Wilson and R. A. Scholtz, “Channel identification: secret sharing using reciprocity in uwb channels,” *IEEE Trans. Information Forensics and Security*, vol. 2, pp. 364–375, 2007.

[4] A. Sayeed and A. Perrig, “Secure wireless communications: secret keys through multipath,” in *Proc. Int. Conf. Acoust. Speech, Signal Processing*, Apr. 2008.

[5] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, “Radiotelepathy: extracting a secret key from an unauthenticated wireless channel,” in *ACM SigMobile Intl Conf. Mobile Computing and Networking (Mobicom)*, 2008.

[6] S. Jana, S. P. Nandha, M. Clark, S. Kasera, N. Patwari, and S. Krishnamurty, “On the effectiveness of secret key extraction using wireless signal strength in real environments,” in *ACM SigMobile Intl Conf. Mobile Computing and Networking (Mobicom)*, 2009.

[7] 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 Trans. Information Forensics and Security*, vol. 53, pp. 3776–3784, Nov. 2005.

[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, Jul. 1993.

[10] T. Wong, M. Bloch, and J. Shea, “Secret sharing over fast-fading mimo wireless channels,” *EURASIP Journal on Wireless Communications and Networking, Special Issue on Wireless Physical Layer Security*, Sep. 2009.

[11] G. Caire and S. Shamai, “On the capacity of some channels with channel state information,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 2007–2019, 1999.

[12] A. Agrawal, Z. Rezki, A. Khisti, and M. S. Alouini, “Non-coherent secret-key agreement with public discussion,” *To Appear, IEEE. Trans. on Inf. Forensics and Security, Special Issue on Physical Layer Security*, Sep. 2011.

[13] W. C. Y. Lee, “Effects on correlation between two mobile radio base-station antennas,” *IEEE. Trans. on Comm.*, vol. 21, pp. 1214–1224, Nov. 1973.

[14] S. B. Rhee and G. I. Zysman, “Results of suburban base-station spatial diversity measurements on the uhf bands,” *IEEE Trans. Commun.*, vol. 22, pp. 1630–1634, Oct. 1974.

[15] A. Khisti, A. Tchamkerten, and G. W. Wornell, “Secure Broadcasting,” *IEEE Trans. Inform. Theory, Special Issue on Information Theoretic Security*, 2008.

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

[17] A. Khisti, S. N. Diggavi, and G. W. Wornell, “Secret-key generation using correlated sources and channels,” *IEEE. Trans. on Inf. Forensics and Security, Special Issue on Physical Layer Security*, Sep. 2011.