Open access • Journal Article • DOI:10.1109/TIFS.2011.2151188

Secret-Key Agreement With Channel State Information at the Transmitter

— Source link

Ashish Khisti, Suhas Diggavi, Gregory W. Wornell

Institutions: University of Toronto, University of California, Los Angeles, Massachusetts Institute of Technology

Published on: 01 Sep 2011 - IEEE Transactions on Information Forensics and Security (IEEE Press)

Topics: Channel state information, Channel capacity, Communication channel, Upper and lower bounds and Transmitter

Secret-Key Agreement With Channel State Information at the Transmitter

Ashish Khisti, *Member, IEEE*, Suhas N. Diggavi, *Member, IEEE*, and Gregory W. Wornell, *Fellow, IEEE*

***Abstract*—We study the capacity of secret-key agreement over a wiretap channel with state parameters. The transmitter, the legitimate receiver, and the eavesdropper are connected by a discrete memoryless wiretap channel with a memoryless state sequence. The transmitter and the legitimate receiver generate a secret-key that must be concealed from the eavesdropper. We assume that the state sequence is known noncausally to the transmitter and no public discussion channel is available. We derive lower and upper bounds on the secret-key capacity. The lower bound involves a source-channel codebook for constructing a common reconstruction sequence at the legitimate terminals and then mapping this sequence to a secret-key using a secret-key codebook. For the special case of Gaussian channels with additive interference (*secret-keys from dirty paper channel*) our bounds differ by 0.5 bit/symbol and coincide in the high signal-to-noise-ratio and high interference-to-noise-ratio regimes. In another special case—*symmetric channel state information (CSI)*—when the legitimate receiver is also revealed the state sequence, we establish optimality of our lower bound. In addition, only causal side information at the transmitter and the receiver suffices to attain the secret-key capacity in the case of symmetric CSI.**

***Index Terms*—Channel reciprocity, channels with state parameters, fading channels, information-theoretic security, secret-key generation, secret sharing, wireless channels, wire-tap channels.**

I. Introduction

SECRET-KEYS are a fundamental requirement for any application involving secure communication or computation. An information theoretic approach to secret-key generation between two or more terminals was pioneered in [3] and [4] and subsequently extended in [5]–[8]. There has also been a significant interest in developing practical approaches for generating shared secret-keys between two or more terminals based on such techniques; see, e.g., [9]–[15] and references therein.

We study the secret-key agreement capacity over a wiretap-channel with a state variable. Channels with state variables [16]–[18] are used in several important applications such as fading channels [19], broadcast channels [20], and digital watermarking systems [21]. In fading channels, the state variable captures the instantaneous fading coefficient of the channel. In broadcast channels, the state sequence models an interfering message to another receiver. In watermarking systems, the state sequence represents a host sequence on which the information message needs to be embedded. In fading channels, we assume that the state sequence is revealed to the terminals causally while in the other two applications, the entire state sequence is known to the transmitter in advance. Unless otherwise stated, we assume the noncausal model. As we discuss in the sequel, the seemingly more general case when each receiver also has a side information can be easily incorporated in this model. We also treat the case when the transmitter and receiver have symmetric and causal channel state information (CSI) and establish the secret-key capacity. This setup is motivated by the application to fading channels where there has been a significant interest already. In yet another application, lower bounds on secret-message transmission over channels with state parameters that exploit secret-key agreement as a building block have been recently proposed [22].

In the present paper we only focus on the case when there is no discussion channel available. We point the reader to [1], [2] for some results on the case when a public discussion channel is available. Notice that a different setup, the *wiretap channel with side information*, is studied in [23]–[25]. Our results indicate that the achievable rates are higher than those reported in [23]–[25], albeit for a different problem.

After the conference papers [1], [2] appeared, on which this paper is based, the authors also became aware about a recent work [26] which studies a similar secret-key agreement scheme for establishing a trade-off between secret-key and secret-message transmission for a problem involving correlated sources. To our knowledge, our results on the upper bound on the secret-key capacity, the asymptotic optimality of the lower bound for the Gaussian case, and the secret-key capacity for the case of symmetric CSI are not included in [26].

The rest of the paper is structured as follows. Section II introduces the statement of the main problem whereas Section III introduces the main results of the paper. Proofs for the case of noncausal CSI, Gaussian model, and the symmetric CSI appear in Sections IV–VI, respectively. A numerical example involving an on–off fading channel is provided in Section VII.

Manuscript received September 16, 2010; revised January 16, 2011; accepted January 19, 2011. Date of publication June 07, 2011; date of current version August 17, 2011. This work was presented in part at the European Wireless Conference 2010, Lucca, Italy, April 2010, and at the IEEE International Symposium on Information Theory (ISIT), Seoul Korea, June 2009. The work of A. Khisti was supported by a Natural Science and Engineering Research Council (NSERC) Discovery Grant. This work was also supported by the National Science Foundation (NSF) under Grant CCF-1017772. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Wade Trappe.

A. Khisti is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5A 2N4, Canada (e-mail: akhisti@comm.utoronto.ca).

S. N. Diggavi is with the University of California, Los Angles (UCLA), Los Angeles, CA 90095 USA (e-mail: suhas@ee.ucla.edu).

G. Wornell is with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: gww@mit.edu).

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

Digital Object Identifier 10.1109/TIFS.2011.2151188

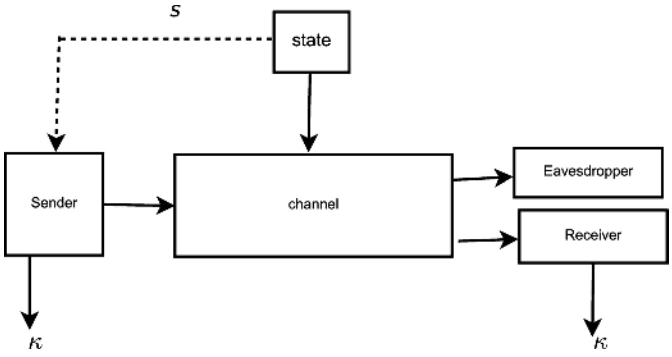


Fig. 1. Wiretap channel controlled by a state parameter. The channel transition probability $p_{y_r,y_e|x,s}$ is controlled by a state parameter s. The entire source sequence s^n is known to the sender but not to the receiver or the eavesdropper. The sender and receiver generate a secret-key κ at the end of the transmission.

II. Problem Statement

A. Channel Model

The channel model has three terminals—a sender, a receiver, and an eavesdropper. The sender communicates with the other two terminals over a discrete-memoryless channel controlled by a random state parameter. The transition probability of the channel is $p_{y_r,y_e|x,s}(\cdot)$, where x denotes the channel input symbol, whereas y_r and y_e denote the channel output symbols at the receiver and the eavesdropper, respectively. The symbol s denotes a state variable that controls the channel transition probability. This is illustrated in Fig. 1. We assume that it is independent and identically distributed (i.i.d.) with the distribution $p_s(\cdot)$. Further, the entire sequence s^n—the channel state information (CSI)—is known to the sender before the communication begins.

B. Secret-Key Capacity

A length n encoder is defined as follows. The sender samples a random variable m from some conditional distribution $p_{m|s^n}(\cdot|s^n)$. The random variable m is a source of external randomness. The encoding function produces a channel input sequence

$$x^n = f_n(m, s^n) \tag{1}$$

and transmits it over n uses of the channel. At time i the symbol x_i is transmitted and the legitimate receiver and the eavesdropper observe output symbols y_{ri} and y_{ei}, respectively, sampled from the conditional distribution $p_{y_r,y_e|x,s}(\cdot)$. The sender and receiver compute secret-keys

$$\kappa = g_n(m, s^n), \qquad l = h_n\left(y_r^n\right). \tag{2}$$

A rate R is achievable if there exists a sequence of encoding functions such that for some sequence ε_n that vanishes as $n \to \infty$, we have that $\Pr(\kappa \neq l) \leq \varepsilon_n$ and

$$\frac{1}{n}H(\kappa) \geq R - \varepsilon_n \tag{3}$$

and

$$\frac{1}{n}I\left(\kappa; y_e^n\right) \leq \varepsilon_n. \tag{4}$$

The largest achievable rate is the secret-key capacity.

C. Extended Model

In our proposed model, we are assuming the state variable is only known to the transmitter and not to the receiving terminals. A more general model involves a state variable that can be decomposed into $s = (s_t, s_r, s_e, s_0)$, where the sequence s_t^n is revealed noncausally to the sender, whereas s_r^n and s_e^n are revealed to the legitimate receiver and the eavesdropper, respectively, while s_0^n is not revealed to any of the terminals. It turns out that the model in Section II-A includes this extended model. The secret-key capacity for this new model is identical to the secret-key capacity of a particular model in Section II-A defined by: $\bar{y}_r = (y_r, s_r)$ and $\bar{y}_e = (y_e, s_e)$ and the channel transition probability

$$p(\bar{y}_r, \bar{y}_e|s_t, x) = \sum_{s_0} p(y_r, y_e|s_0, s_r, s_e, s_t, x)p(s_0, s_r, s_e|s_t). \tag{5}$$

The equivalence can be established by noting that the modified channel preserves the same knowledge of the side information sequences as the original problem, the rate and equivocation terms only depend on the joint distribution $p\left(\bar{y}_r^n, \bar{y}_e^n, x^n, s_t^n\right)$ and for any input distribution $p\left(x^n|s_t^n\right)$, the extended channel satisfies

$$p\left(\bar{y}_r^n, \bar{y}_e^n|x^n, s_t^n\right) = \prod_{i=1}^{n} p(\bar{y}_{ri}, \bar{y}_{ei}|x_i, s_{ti}) \tag{6}$$

where each term on the right-hand side of (6) obeys (5).

We omit a detailed proof in interest of space and point to the reader to [27, pp. 17–25] and [28, Ch. 7, pp. 7–54] for an analogous observation.

III. Main Results

We summarize the main results of this paper in this section.

A. Capacity Bounds

We first provide an achievable rate (lower bound) on the secret-key capacity.

Theorem 1: A lower bound on the secret-key capacity, defined in Section II-B, is

$$R^{-} = \max_{p_{u|s}, p_{x|s,u}} I(u; y_r) - I(u; y_e) \tag{7}$$

where the maximization is over all auxiliary random variables u that satisfy the Markov condition $u \to (x, s) \to (y_r, y_e)$ and furthermore satisfy the constraint that

$$I(u; y_r) - I(u; s) \geq 0. \tag{8}$$

The intuition behind the coding scheme is as follows. Upon observing s^n, the sender communicates the best possible reproduction u^n of the state sequence to the receiver. The receiver produces u^n with high probability provided that (8) is satisfied.

The set of all codewords u^n is binned into 2^{nR^-} bins and the bin-index is declared to be the secret-key.

We next provide an upper bound to the secret-key capacity that is amenable to numerical evaluation.

Theorem 2: The secret-key capacity is upper bounded by $C \leq R^+$, where

$$R^+ = \min_{p_{y_r,y_e|x,s} \in \mathcal{P}} \max_{p_{x|s}} I(x,s;y_r|y_e) \tag{9}$$

where $\mathcal{P}$ denotes all the joint distributions $p^\star_{y_r,y_e|x,s}$ that have the same marginal distribution as the original channel.

The intuition behind the upper bound is as follows. We create a degraded channel by revealing the output of the eavesdropper to the legitimate receiver. We further assume a channel with two inputs (x^n, s^n), i.e., the state sequence s^n is not arbitrary, but rather a part of the input codeword with distribution p_s. The secrecy capacity of the resulting wiretap channel is then given by $I(x,s;y_r|y_e)$.

In related literature, we note that our coding scheme is similar to [29] and [30] that studies the problem of communicating a state sequence, under common knowledge, and under a distortion constraint. The slight difference here is that we use the common reconstruction sequence in this problem only as an intermediate step to generate a common secret-key and do not impose any distortion constraint.

It is also interesting to compare our lower bound in Theorem 1 with the secret-message transmission problem [23]–[25] over wiretap channels with state parameters. While the secret-key can be an arbitrary function of the state sequence (known only to the transmitter), the secret-message must be independent of the state sequence, thus imposing a stricter constraint. A lower bound on the secret message transmission capacity is (cf. [23]–[25])

$$R = \max_{p_{u|s}, p_{x|s,u}} I(u;y_r) - \max\left(I(u;s), I(u;y_e)\right). \tag{10}$$

B. Secret-Keys From Dirty Paper Coding

We study the Gaussian case under an average power constraint. The channel to the legitimate receiver and the eavesdropper is expressed as

$$\begin{aligned} y_r &= x + s + z_r \\ y_e &= x + s + z_e \end{aligned} \tag{11}$$

where $z_r \sim \mathcal{N}(0,1)$ and $z_e \sim \mathcal{N}(0, 1+\Delta)$ denote the additive white Gaussian noise and are assumed to be sampled independently. The state parameter $s \sim \mathcal{N}(0,Q)$ is also sampled i.i.d. at each time instance and is independent of both z_r and z_e. Furthermore, the channel input satisfies an average power constraint $E[x^2] \leq P$. We assume that the CSI s^n is revealed noncausally to the sender but not to any other terminals.

In this model, P also denotes the signal-to-noise ratio (SNR), Q denotes the interference-to-noise ratio (INR), whereas Δ denotes the degradation level of the eavesdropper. While we do not elaborate, possible applications of this model include secret-key generation in multimedia communication systems and in broadcast channels, where connections to the dirty paper coding channel already exist.

Propositions 1 and 2, stated next, provide the lower and upper bounds on the secret-key capacity. For the lower bound, we limit our analysis to the case when $P \geq 1$.[1]

Proposition 1: Assuming that $P \geq 1$, a lower bound on the secret-key agreement capacity is

$$R^- = \frac{1}{2}\log\left(1 + \frac{\Delta(P+Q+2\rho\sqrt{PQ})}{P+Q+1+\Delta+2\rho\sqrt{PQ}}\right) \tag{12}$$

where $|\rho| < 1$ and

$$P(1-\rho^2) = 1 - \frac{1}{P+Q+1}. \tag{13}$$

Proposition 2: An upper bound on the secret-key capacity is given by

$$R^+ = \frac{1}{2}\log\left(1 + \frac{\Delta(P+Q+2\sqrt{PQ})}{P+Q+1+\Delta+2\sqrt{PQ}}\right). \tag{14}$$

It can be readily verified that the upper and lower bounds are close in several interesting regimes. In Fig. 2, we numerically plot these bounds as a function of SNR and the degradation level at the eavesdropper. Proposition 3, whose proof is omitted due to space constraints, states that the bounds are close in several regimes.

Proposition 3: Our bounds on the secret-key capacity in Propositions 1 and 2 satisfy the following:

$$R_+ - R_- \leq \frac{1}{2} \text{ bits/symbol} \tag{15}$$

$$\lim_{P\to\infty} R_+ - R_- = 0 \tag{16}$$

$$\lim_{Q\to\infty} R_+ - R_- = 0. \tag{17}$$

C. Symmetric CSI

In the special case where the state sequence s is also revealed to the legitimate receiver, we have a complete characterization of the secret-key capacity, as stated below.

Theorem 3: The secret-key capacity for the channel model in Section II-A when the state sequence s^n is also revealed to the decoder is given by

$$C_{\text{sym}} = \max_{p_{u|s(\cdot)} p_{x|u,s(\cdot)}} I(u;y_r|s) - I(u;y_e|s) + H(s|y_e) \tag{18}$$

where the maximization is over all auxiliary random variables u that obey the Markov chain $u \to (x,s) \to (y_r, y_e)$. Additionally, it suffices to limit the cardinality of the auxiliary variable to $|\mathcal{S}|(1+|\mathcal{X}|)$.

The achievability in (18) follows from (7) by augmenting $\bar{y}_r = (y_r, s)$. Observe that the condition in (8) is redundant, as $I(u;y_r,s) - I(u;s) \geq 0$ holds for any input distribution $p_{u,x|s}(\cdot,\cdot)$. The expression in (7) simplifies as follows:

$$\begin{aligned} R^- &= \max_{p_{u|s}, p_{x|u,s}} I(u;y_r,s) - I(u;y_e) \\ &= \max_{p_u, p_{x|s,u}} I(u;y_r|s) - I(u;y_e|s) + I(s;u|y_e) \quad (19) \\ &= \max_{p_u, p_{x|s,u}} I(u;y_r|s) - I(u;y_e|s) + H(s|y_e) \quad (20) \end{aligned}$$

[1] The choice that $P \geq 1$ simply guarantees that (13) has a solution in ρ. The lower bound is valid for any P and Q for which (13) has a solution in ρ.

where the last relation follows by noting that if u is an optimal choice in (19) then by selecting $u^\star = (u, s)$ will leave the difference in the two mutual information terms unchanged but increase the second term $H(s|y_e)$ as specified in (20). In this case, the expression in (20) is identical to (18). The converse follows by an application of Csiszar's Lemma [28, Ch. 2] and is provided in Section VI-B.

It is also possible to provide another coding scheme for Theorem 3 that only requires a causal knowledge of s^n at the encoder. The scheme is based on the following interpretation of (18). The term $I(u;y_r|s) - I(u;y_e|s)$ is the rate of a multiplexed wiretap codebook constructed assuming that all the three terminals have knowledge of s^n. The second term $H(s|y_e)$ is the rate of the additional secret-key that can be produced by exploiting the fact that s^n is only known to the sender and the legitimate terminal. This scheme only requires that the CSI sequence be revealed *causally* to the terminals.

Observe that the capacity expression (18) involves a tension between two competing effects. To maximize the contribution of the rate obtained from the multiplexed wiretap codebook, it is desirable to select u to be correlated with s in a certain positive manner. However, doing so will leak more information about s to the wiretapper. To maximize $H(s|y_e)$, we need an input that masks the state sequence from the eavesdropper. The optimal distribution is required to strike a balance between the two terms. We illustrate this trade-off via an example in Section VII. Finally, it can be easily verified that the expression (18) simplifies in the following special case.

Corollary 1: Suppose that for each $s \in \mathcal{S}$ the channel $p_{y_r,y_e|s=s,x}(y_r, y_e|s, x)$ is such that the eavesdropper's channel is less noisy compared to the legitimate receiver's channel. Then the secret-key capacity with s^n revealed to both the legitimate terminals is

$$C = \max_{p_{x|s}} H(s|y_e). \tag{21}$$

Intuitively, when the wiretap channel cannot contribute to the secrecy, (21) states that transmitter should select an input that masks the state from the output as much as possible. We omit a formal proof of Corollary 1 due to space constraints.

IV. Noncausal CSI

In this section we provide proofs of Theorems 1 and 2.

Proof of Theorem 1

The coding theorem involves constructing a common sequence u^n at the legitimate terminals and using it to generate a secret-key.

1) Codebook Generation: Assume that the input distribution is such that $I(u;y_r) > I(u;s)$ as required in Theorem 1. Let ε_n be a sequence of nonnegative numbers that goes to zero such that $2\varepsilon_n < I(u;y_r) - I(u;s)$.

- Generate a total of $T = 2^{n(I(u;y_r)-2\varepsilon_n)}$ sequences. Each sequence is sampled i.i.d. from a distribution $p_u(\cdot)$. Label them $u_1^n, \ldots, u_T^n$.
- Select a rate $R = I(u;y_r) - I(u;y_e) - \varepsilon_n$ and randomly partition the set of sequences selected in the previous step into 2^{nR} bins. There will be $2^{n(I(u;y_e)-\varepsilon_n)}$ sequences in each bin. This is illustrated in Fig. 3.

2) Encoding:

- Given a state sequence s^n, the encoder selects a sequence u^n randomly from the list of all possible sequences that are jointly typical with s^n. Let the index of this sequence be L.
- At time $i = 1, 2, \ldots, n$, the encoder transmits symbol x_i generated by sampling the distribution $p_{x|u,s}(\cdot|u_i, s_i)$.

3) Secret-Key Generation:

- The decoder upon observing y_r^n finds a sequence u^n jointly typical with y_r^n.
- Both encoder and the decoder declare the bin-index of u^n to be the secret-key.

4) Error Probability Analysis: An error occurs only if one of the following events occur:

$$\mathcal{E}_1 = \{(u^n(l), s^n) \notin \mathcal{T}_\varepsilon^n(u,s) \text{ for all } 1 \le l \le T\} \tag{22}$$

$$\mathcal{E}_2 = \{(u^n(L), y_r^n) \notin \mathcal{T}_\varepsilon^n(u,y_r)\} \tag{23}$$

$$\mathcal{E}_3 = \{(u^n(l), y_r^n) \in \mathcal{T}_\varepsilon^n(u,y_r) \text{ for some } l \ne L\}. \tag{24}$$

Since the number of sequences $T > 2^{nI(u;s)}$, it follows from the Covering Lemma [28, Ch. 3] that $\Pr(\mathcal{E}_1) \to 0$ as $n \to \infty$. Furthermore, let $\mathcal{E}_1^c = \{(u^n, s^n, x^n) \in \mathcal{T}_{\varepsilon'}^n(u,s,x)\}$ and $\Pr(\mathcal{E}_1^c) \to 1$ as $n \to \infty$ for any $\varepsilon' > \varepsilon$. Since $p(y_r^n|u^n(L), x^n, s^n) = \prod_{i=1}^n p(y_{ri}|u_i, x_i, s_i)$ it follows from the conditional typicality Lemma [28, Ch. 2] that $\Pr(\mathcal{E}_2 \cap \mathcal{E}_1^c) \to 0$ as $n \to \infty$. Finally, since every $u^n(l)$ is generated i.i.d. $p_u(u_i)$ and is independent of y_r^n for $l \ne L$, it follows from the Packing Lemma [28, Ch. 3] that $\Pr(\mathcal{E}_3) \to 0$ if $T < 2^{nI(u;y_r)}$.

5) Secrecy Analysis: We need to show that for the proposed encoder and decoder, the equivocation at the eavesdropper satisfies

$$\frac{1}{n}H(\kappa|y_e^n) = I(u;y_r) - I(u;y_e) + o_n(1) \tag{25}$$

where $o_n(1)$ is a term that goes to zero as $n \to \infty$.

Note that while the key κ in general can be a function of (s^n, m) as indicated in (1), in our coding scheme the secret-key is a deterministic function of u^n and hence we have

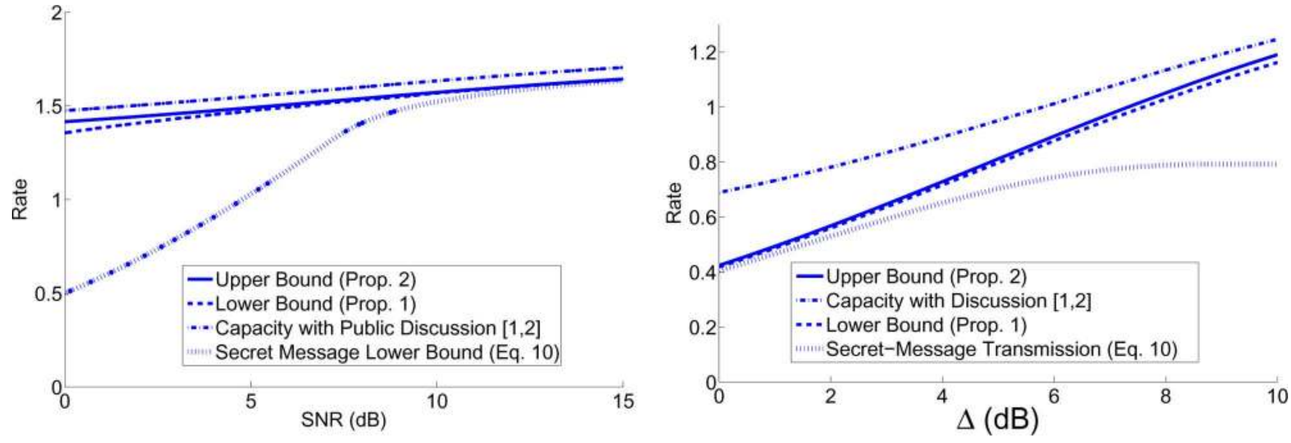
$$\begin{aligned}\frac{1}{n}H(\kappa|y_e^n) &= \frac{1}{n}H(\kappa, u^n|y_e^n) - \frac{1}{n}H(u^n|y_e^n, \kappa)\\ &= \frac{1}{n}H(u^n|y_e^n) - \frac{1}{n}H(u^n|y_e^n, \kappa)\\ &= \frac{1}{n}H(u^n|y_e^n) - \varepsilon_n\end{aligned}$$

where the last step follows from the fact that there are $T_0 = 2^{n(I(u;y_e)-\varepsilon_n)}$ sequences in each bin. Again applying the packing lemma we can show that with high probability the eavesdropper uniquely finds the codeword $u^n(L)$ jointly typical with y_e^n in this set and hence Fano's Inequality implies that

$$\frac{1}{n}H(u^n|y_e^n, \kappa) \le \varepsilon_n.$$

It remains to show that

$$\frac{1}{n}H(u^n|y_e^n) \ge I(u;y_r) - I(u;y_e) - o_n(1).$$



(a)

(b)

Fig. 2. Bounds on the capacity of the "secret-keys from dirty paper" channel. In the left figure, we plot the bounds on capacity as a function of SNR (decibels) where the INR $Q = 10$ dB and the degradation level at the eavesdropper, $\Delta = 10$ dB. The uppermost curve is the capacity with public-discussion [1] whereas the successively lower curves denote the upper and lower bounds on the capacity as stated in Proposition 2 and Proposition 1. The lowermost curve is the secret message transmission lower bound (10) evaluated for a jointly Gaussian input distribution. In the right figure, we vary the degradation level at the eavesdropper Δ (in decibels) and compute the secret-key rates for $P = 3$ (dB) and $Q = 2$ (dB). The uppermost curve is the secret-key capacity with public discussion [1], the successively lower curves are the upper and the lower bounds, whereas the lowermost curve is the secret-message transmission rate evaluated for Gaussian inputs.

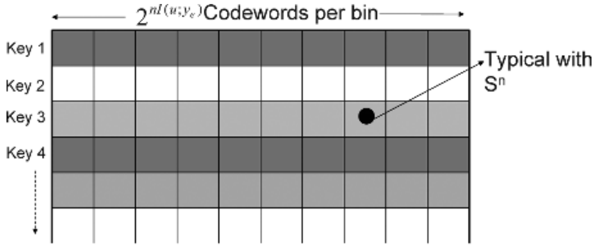


Fig. 3. Codebook for the secret-key agreement problem. A total of $\approx 2^{nI(u;y_r)}$ codewords are generated i.i.d. $p_u(\cdot)$ and partitions into 2^{nR} bins so that there are $2^{nI(u;y_e)}$ sequences in each bin. Given s^n, a jointly typical sequence u^n is selected and its bin index constitutes the secret-key.

Using the chain rule of the joint entropy, we have

$$\frac{1}{n}H(u^n|y_e^n) = \frac{1}{n}H(u^n) + \frac{1}{n}H(y_e^n|u^n) - \frac{1}{n}H(y_e^n) \quad (26)$$

$$= \frac{1}{n}H(u^n) + \frac{1}{n}H(y_e^n|u^n, s^n) - \frac{1}{n}H(y_e^n) + \frac{1}{n}I(s^n; y_e^n|u^n). \quad (27)$$

We now appropriately bound each term in (27). First note that since the sequence u^n is uniformly distributed among the set of all possible codeword sequences, it follows that

$$\frac{1}{n}H(u^n) = \frac{1}{n}\log_2|\mathcal{C}| = I(u; y_r) - 2\varepsilon_n. \quad (28)$$

Next, as verified below, the channel to the eavesdropper $(u^n, s^n) \rightarrow y_e^n$, is memoryless

$$p_{y_e^n|u^n,s^n}(y_e^n|u^n,s^n)$$
$$= \sum_{x^n\in\mathcal{X}^n} p_{y_e^n|u^n,s^n,x^n}(y_e^n|u^n,s^n,x^n)p_{x^n|u^n,s^n}(x^n|u^n,s^n)$$
$$= \sum_{x^n\in\mathcal{X}^n} \prod_{i=1}^{n} p_{y_e|u,s,x}(y_{e,i}|u_i,s_i,x_i)p_{x|u,s}(x_i|u_i,s_i)$$
$$= \prod_{i=1}^{n} \sum_{x_i\in\mathcal{X}} p_{y_e|u,s,x}(y_{e,i}|u_i,s_i,x_i)p_{x|u,s}(x_i|u_i,s_i)$$
$$= \prod_{i=1}^{n} p_{y_e|u,s}(y_{e,i}|u_i,s_i).$$

The second step above follows from the fact that the channel is memoryless and the symbol x_i at time i is generated as a function of (u_i, s_i). Hence we have that

$$\frac{1}{n}H(y_e^n|s^n,u^n) = \frac{1}{n}\sum_{i=1}^{n} H(y_{e,i}|s_i,u_i). \quad (29)$$

Furthermore, note that

$$\frac{1}{n}H(y_e^n) \leq \frac{1}{n}\sum_{i=1}^{n} H(y_{ei}). \quad (30)$$

Finally, in order to lower bound the last term in (27), we let J be a random variable which equals 1 if (s^n, u^n) are jointly typical. Note that $\Pr(J = 1) = 1 - o_n(1)$.

$$\frac{1}{n}I(s^n; y_e^n|u^n) = \frac{1}{n}H(s^n|u^n) - \frac{1}{n}H(s^n|u^n, y_e^n)$$
$$\geq \frac{1}{n}H(s^n|u^n, J = 1)\Pr(J = 1) - \frac{1}{n}H(s^n|u^n, y_e^n)$$
$$= \frac{1}{n}H(s^n|u^n, J = 1) - \frac{1}{n}H(s^n|u^n, y_e^n) - o_n(1)$$
$$\geq H(s|u) - \frac{1}{n}H(s^n|u^n, y_e^n) - o_n(1) \quad (31)$$
$$\geq H(s|u) - \frac{1}{n}\sum_{i=1}^{n} H(s_i|u_i, y_{e,i}) - o_n(1) \quad (32)$$

where (31) follows from the fact that s^n is an i.i.d. sequence and hence conditioned on the fact that (s^n, u^n) is a pair of typical sequence there are $2^{nH(s|u)-no_n(1)}$ possible sequences s^n.

Substituting (28), (29), (30), and (32) in the lower bound (27) and using the fact that as $n \to \infty$, the summation converges to the mean values

$$\begin{aligned}\frac{1}{n}H(\kappa|y_e^n) &= I(u; y_r) + H(y_e|u, s) - H(y_e)\\ &\quad + H(s|u) - H(s|u, y_e) - o_n(1)\\ &= I(u; y_r) - I(y_e; s|u) - I(y_e; u)\\ &\quad + I(y_e; s|u) - o_n(1)\\ &= I(u; y_r) - I(y_e; u) - o_n(1)\end{aligned}$$

as required.

Proof of Theorem 2

A sequence of length-n code satisfies

$$\frac{1}{n}H(\kappa|y_r^n) \le \varepsilon_n \tag{33}$$

$$\frac{1}{n}H(\kappa|y_e^n) \ge \frac{1}{n}H(\kappa) - \varepsilon_n \tag{34}$$

where (33) follows from the Fano's inequality since the receiver is able to recover the secret-key κ given y_r^n and (34) is a consequence of the secrecy constraint. Furthermore, note that $\kappa \to (x^n, s^n) \to (y_r^n, y_e^n)$ holds as the encoder generates the secret-key κ. Thus we can bound the rate $R = (1/n)H(\kappa)$ as below

$$\begin{aligned}nR &\le I(\kappa; y_r^n|y_e^n) + 2n\varepsilon_n\\ &\le I(\kappa, s^n, x^n; y_r^n|y_e^n) + 2n\varepsilon_n\\ &\le I(s^n, x^n; y_r^n|y_e^n) + H(\kappa|s^n, x^n) + 2n\varepsilon_n\\ &= I(s^n, x^n; y_r^n|y_e^n) + 3n\varepsilon_n && (35)\\ &\le \sum_{i=1}^{n} I(s_i, x_i; y_{r,i}|y_{e,i}) + 3n\varepsilon_n && (36)\\ &\le nI(x, s; y_r|y_e) + 3n\varepsilon_n && (37)\end{aligned}$$

where (35) follows from the Fano Inequality because κ can be obtained from (x^n, s^n), (36) from the fact that the channel is memoryless and the last step follows from the concavity $I(x, s; y_r|y_e)$ in the input distribution $p_{x,s}$ (see, e.g., [31]).

Finally, since the secret-key capacity only depends on the marginal distribution of the channel and not on the joint distribution, we can minimize over all joint distributions with fixed marginal distributions.

V. Gaussian Case

We develop the lower and upper bounds on secret-key agreement capacity for the Gaussian channel model.

Proof of Proposition 1

Recall that $s \sim \mathcal{N}(0, Q)$. Choose $x \sim \mathcal{N}(0, P)$ to be a Gaussian random variable and let $E[xs] = \rho\sqrt{PQ}$. The lower bound follows by selecting $u = x + \alpha s$:

$$\begin{aligned}R &= I(u; y_r) - I(u; y_e)\\ &= h(u|y_e) - h(u|y_r).\end{aligned}$$

Further evaluating each of the terms above with $u = x + \alpha s$, note that

$$\begin{aligned}h(u|y_e) &= h(x + \alpha s|x + s + z_e)\\ &= \frac{1}{2}\log 2\pi e\Bigg(P + \alpha^2 Q + 2\alpha\rho\sqrt{PQ}\\ &\qquad - \frac{(P + \alpha Q + (1+\alpha)\rho\sqrt{PQ})^2}{P + Q + 1 + \Delta + 2\rho\sqrt{PQ}}\Bigg)\end{aligned}$$

and

$$\begin{aligned}h(u|y_r) &= h(x + \alpha s|x + s + z_r)\\ &= \frac{1}{2}\log 2\pi e\Bigg(P + \alpha^2 Q + 2\alpha\rho\sqrt{PQ}\\ &\qquad - \frac{(P + \alpha Q + \rho(1+\alpha)\sqrt{PQ})^2}{P + Q + 1 + 2\sqrt{PQ}}\Bigg).\end{aligned}$$

This yields that

$$\begin{aligned}R = \frac{1}{2}\log&\left(1 + \frac{\Delta}{1 + \frac{PQ(\alpha-1)^2(1-\rho^2)}{P+\alpha^2 Q+2\rho\alpha\sqrt{PQ}}}\right)\\ &+\frac{1}{2}\log\left(\frac{P + Q + 1 + 2\rho\sqrt{PQ}}{P + Q + 1 + \Delta + 2\rho\sqrt{PQ}}\right).\end{aligned} \tag{38}$$

Note that the first term in the expression above is maximized when $\alpha = 1$. In this case we have that

$$\begin{aligned}R &= \frac{1}{2}\log\left(\frac{(1+\Delta)(P + Q + 1 + 2\rho\sqrt{PQ})}{P + Q + 1 + \Delta + 2\rho\sqrt{PQ}}\right) && (39)\\ &= \frac{1}{2}\log\left(1 + \frac{\Delta(P + Q + 2\rho\sqrt{PQ})}{P + Q + 1 + \Delta + 2\rho\sqrt{PQ}}\right) && (40)\end{aligned}$$

as required.

To complete the proof, we show that the choice $\alpha = 1$ is indeed feasible when $P \ge 1$ and (P, ρ) satisfy (13).

In particular, the constraint (8) requires that

$$\begin{aligned}&h(u|s) \ge h(u|y_r)\\ &\Rightarrow h(x|s) \ge h(x + s|x + s + z_r)\\ &\Rightarrow \frac{1}{2}\log P(1-\rho^2) \ge \frac{1}{2}\log\left(\frac{P + Q + 2\rho\sqrt{PQ})}{P + Q + 1 + 2\rho\sqrt{PQ}}\right).\end{aligned}$$

Rearranging,

$$P(1-\rho^2) \ge 1 - \frac{1}{P + Q + 1 + 2\rho\sqrt{PQ}} \ge 1 - \frac{1}{P + Q + 1} \tag{41}$$

as required.

It is worth comparing the choice of the auxiliary variable $u = x + s$ in the present problem with the choice of optimal u in the dirty paper coding problem [32]. While the input x is independent of s in [32], as illustrated in Fig. 4, the optimal x

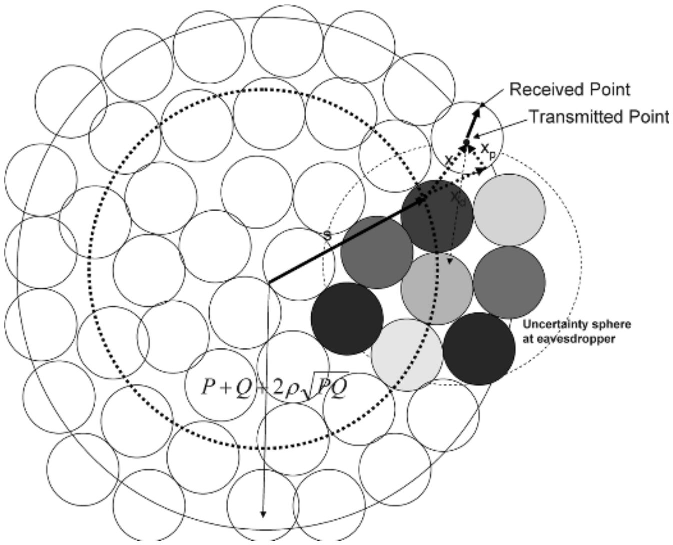


Fig. 4. Secret-key agreement codebook for the dirty paper channel. The transmit sequence x^n is selected so that $u^n = x^n + s^n$ is a sequence in the codebook $\mathcal{C}$. The smaller spheres above denote the noise uncertainty at the legitimate receiver. Their centers are the codewords in $\mathcal{C}$. Our binning of smaller spheres guarantees that the noise uncertainty sphere of the eavesdropper has all possible messages, resulting in (asymptotically) perfect equivocation.

in the secret-key problem has a component along s. This is because scaling the interference sequence increases the secret-key rate. Second, recall that in [32] we find the auxiliary codeword u^n that is closest to αs^n, where $\alpha = P/(P+N)$. In contrast this MMSE scaling is not performed in the secret-key problem.

Proof of Proposition 2

We evaluate the upper bound in Theorem 2 for the choice $z_e = z_r + z_\delta$, where $z_\delta \sim \mathcal{N}(0, \Delta)$ is independent of z_r

$$\begin{aligned} I(s, x; y_r|y_e) &= h(y_r|y_e) - h(y_r|y_e, x, s) \\ &= h(y_r|y_e) - h(z_r|z_e) \\ &\leq \frac{1}{2}\log\Bigg(P + Q + 1 + 2\sqrt{PQ} \\ &\qquad - \frac{(P+Q+1+2\sqrt{PQ})^2}{P+Q+1+\Delta+2\sqrt{PQ}}\Bigg) - \\ &\quad - \frac{1}{2}\log\left(1 - \frac{1}{1+\Delta}\right) \end{aligned}$$

where we have used the fact that the conditional entropy $h(y_r|y_e)$ is maximized by a Gaussian distribution [33]. The above expression gives (14).

VI. Symmetric CSI

We establish the secret-key capacity for the case of symmetric CSI, i.e., when s^n is revealed to both the transmitter and the legitimate receiver.

A. Achievability for Theorem 3

As explained in Section III-C, the achievability result follows directly from Theorem 1 by replacing y_r with $\bar{y_r} = (y_r, s)$ in the lower bound expression. We nevertheless provide an alternate scheme that only requires the knowledge of causal CSI at the transmitter. The idea is to use a different wiretap codebook for each realization of the state variable. In particular, suppose that $\mathcal{S} = \{s_1, \ldots, s_M\}$ denote the set of available states. Since the encoder and the decoder are both aware of the state realization s_i and can use this common knowledge to select the appropriate codebook for transmission. These codebooks are constructed for an eavesdropper who is also aware of the state sequence realization. Suppose that we fix the distribution $p_{u,x|s=s_i}(\cdot)$ in (18). Let

$$R_i = I(u; y_r|s = s_i) - I(u; y_e|s = s_i) \qquad (42)$$

and $p_i = \Pr(s = s_i)$. For each $i = 1, 2\ldots, M$, a wiretap codebook of length np_i and rate R_i is constructed and used to transmit a message κ_i. Another independent key κ_s of rate $R_s = H(s|y_e)$ is then generated by exploiting the fact that s^n is not known to the eavesdropper.

1) Codebook Construction:

- For each $i = 1, \ldots, M$ generate a codebook $\mathcal{C}_i$ of rate $R_i - 2\varepsilon_n$ and length $n_i = n(p_i - \varepsilon_n)$ by sampling the codewords i.i.d. from the distribution $p_{u|s}(\cdot|s_i)$.
- Construct a codebook $\mathcal{C}_s$, where the set of all typical sequences s^n of size $2^{n(H(s)-2\varepsilon_n)}$ is partitioned into $2^{n(R_s-\varepsilon_n)}$ bins each containing $2^{n(I(s;y_e)-\varepsilon_n)}$ sequences.

2) Encoding:

- For each $i = 1, \ldots, M$ the transmitter selects a random message κ_i and a random codeword sequence $t_i^{n_i}$ in the corresponding bin of $\mathcal{C}_i$.
- Upon observing $s(j) = s_i$ at time $t = j$, it selects the next available symbol of $t_i^{n_i}$ and samples the channel input symbol from the distribution $p_{x|s,u}$.
- At the end of the transmission, it looks for the bin index of s^n in $\mathcal{C}_s$ and declares this to be κ_s.
- The overall secret-key is $(\kappa_1, \ldots, \kappa_M, \kappa_s)$.

3) Decoding:

- The decoder divides y_r^n into subsequences $(y_1^{n_1}, \ldots, y_M^{n_M})$, where the subsequences $y_i^{n_i}$ is obtained by collecting the symbols of y_r^n when $s = s_i$.
- For $i = 1, \ldots, M$, it searches for a codeword $t_i^{n_i}$ in $\mathcal{C}_i$ that is jointly typical with $y_i^{n_i}$. If no such codeword or multiple codewords is found an error is declared. Otherwise the bin index of $t_i^{n_i}$ is set to equal $\hat{\kappa}_i$.

Through standard arguments it can be shown that the error probability in decoding at the legitimate receiver vanishes as $n \to \infty$ provided we select the rates according to (42). We omit the details due to space constraints.

4) Secrecy Analysis: First, consider splitting $y_e^n = (y_{e1}^{n_1}, \ldots, y_{eM}^{n_M})$, where the subsequence $y_{ej}^{n_j}$ is obtained by grouping the symbols of y_e^n when $s = s_j$. From the construction of the wiretap codebook $\mathcal{C}_j$ it follows that

$$\frac{1}{n}H\left(\kappa_j|y_{ej}^{n_j}\right) \geq \frac{1}{n}H(\kappa_j) - \varepsilon_n, \qquad j = 1, \ldots, M. \qquad (43)$$

Next since the messages are selected independently and the encoding functions are also independent, it follows that

$$\begin{aligned} &\frac{1}{n}H\left(\kappa_j|\kappa_1, \ldots, \kappa_{j-1}, \kappa_{j+1}, \ldots, \kappa_M, y_e^n, s^n\right) \\ &\qquad = \frac{1}{n}H(\kappa_j|y_{ej}^{n_j}) \geq \frac{1}{n}H(\kappa_j) - \varepsilon_n. \qquad (44) \end{aligned}$$

Thus by the chain rule we have that

$$\frac{1}{n}H\left(\kappa_1,\ldots,\kappa_M|y_e^n,s^n\right)\geq R_0-\varepsilon_n \tag{45}$$

where $R_0 = H(\kappa_1,\ldots,\kappa_M) = I(u;y_r|s) - I(u;y_e|s)$. The secrecy analysis can be completed by combining (45) and (52) which is established in Lemma 1 at the end of this section.

$$\begin{aligned}\frac{1}{n}H(\kappa_1^M,\kappa_s|y_e^n) &= \frac{1}{n}H(\kappa_1^M|\kappa_s,y_e^n)+\frac{1}{n}H(\kappa_s|y_e^n) && (46)\\ &\geq \frac{1}{n}H(\kappa_1^M|s^n,y_e^n)+\frac{1}{n}H(\kappa_s|y_e^n) && (47)\\ &\geq I(u;y_r|s)-I(u;y_e|s)\\ &\quad+\frac{1}{n}H(\kappa_s|y_e^n)-o_n(1) && (48)\\ &\geq I(u;y_r|s)-I(u;y_e|s)+\frac{1}{n}H(s^n|y_e^n)\\ &\quad-\frac{1}{n}H(s^n|y_e^n,\kappa_s)-o_n(1) && (49)\\ &\geq I(u;y_r|s)-I(u;y_e|s)+H(s|y_e)\\ &\quad-\frac{1}{n}H(s^n|y_e^n,\kappa_s)-o_n(1) && (50)\\ &= I(u;y_r|s)-I(u;y_e|s)\\ &\quad+H(s|y_e)-o_n(1) && (51)\end{aligned}$$

where (47) and (49) follow from the fact that κ_s is a deterministic function of s^n while (48) follows by substituting (45) and (50) follows by substituting (52) while (51) follows from the fact that $1/nH(s^n|y_e^n,\kappa_s)\to 0$ as $n\to\infty$, since from the construction of $\mathcal{C}_s$ there are at most $2^{n(I(s;y_e)-\varepsilon_n)}$ sequences associated with any given bin. Hence the decoder can decode s^n with high probability and hence Fano's inequality applies.

It only remains to establish the following.

Lemma 1: For any input distribution $p_{u,x|s}$ such that $I(u;y_r|s) > I(u;y_e|s)$, we have that

$$\frac{1}{n}H(s^n|y_e^n)\geq\frac{1}{n}H(s|y_e)-o_n(1). \tag{52}$$

Proof: First observe that we can write

$$\begin{aligned}\frac{1}{n}H(s^n|y_e^n) &= \frac{1}{n}H(y_e^n|s^n)+\frac{1}{n}H(s^n)-\frac{1}{n}H(y_e^n) && (53)\\ &= \frac{1}{n}H(y_e^n|s^n,u^n)+\frac{1}{n}I(u^n;y_e^n|s^n)\\ &\quad+\frac{1}{n}H(s^n)-\frac{1}{n}H(y_e^n). && (54)\end{aligned}$$

We now observe the following. Since the channel from $(u^n,s^n)\to y_e^n$ is memoryless

$$\frac{1}{n}H(y_e^n|s^n,u^n)=\frac{1}{n}\sum_{i=1}^{n}H(y_{ei}|s_i,u_i)\to H(y_e|s,u) \tag{55}$$

as $n\to\infty$. Next note that by construction

$$\frac{1}{n}H(u^n|s^n)=I(u;y_r|s)-2\varepsilon_n \tag{56}$$

and since $I(u;y_r|s) > I(u;y_e|s)$, it follows through standard calculations that[2]

$$\frac{1}{n}H(u^n|s^n,y_e^n)\leq I(u;y_r|s)-I(u;y_e|s)-o_n(1). \tag{57}$$

Combining the above two inequalities

$$\frac{1}{n}I(u^n;y_e^n|s^n)\geq I(u;y_e|s)-o_n(1). \tag{58}$$

Since the sequence s^n is sample i.i.d., we have

$$\frac{1}{n}H(s^n)=H(s) \tag{59}$$

and finally from the chain rule

$$\frac{1}{n}H(y_e^n)\leq\frac{1}{n}H(y_{ei})\to H(y_e) \tag{60}$$

as $n\to\infty$. Substituting (55), (58), (59), and (60) into (54) completes the claim. ■

B. Converse

For any sequence of codes indexed by the codeword length n, we show that the secret-key rate is upper bounded by the capacity expression (18) plus a term that vanishes to zero as the block length goes to zero. By applying the Fano inequality on the secret-key rate, we have that for some sequence ε_n that approaches zero as n goes to infinity that

$$nR\leq I(\kappa;l)+n\varepsilon_n\leq I(\kappa;s^n,y_r^n)+n\varepsilon_n \tag{61}$$

where the last step follows from the data processing inequality since $l=h_n(s^n,y_r^n)$. Furthermore, from the secrecy condition $I(\kappa;y_e^n)\leq n\varepsilon_n$ and hence

$$\begin{aligned}nR &\leq I(\kappa;s^n,y_r^n)-I(\kappa;y_e^n)+2n\varepsilon_n && (62)\\ &\leq \sum_{i=1}^{n}I(\kappa;y_{ri},s_i|y_e^{i-1}y_{r,i+1}^n,s_{i+1}^n)\\ &\quad-I(\kappa;y_{e,i}|y_e^{i-1}y_{r,i+1}^n,s_{i+1}^n) && (63)\end{aligned}$$

where the second step follows from the Csiszar sum-identity [28, Ch. 2] applied to difference of mutual informations. The derivation is analogous to [35] and is omitted. If we let $v_i = (y_e^{i-1}y_{r,i+1}^n,s_{i+1}^n)$ and $u_i=(\kappa,v_i)$ note that $v_i\to u_i\to(x_i,s_i)\to(y_{r,i},y_{e,i})$ holds. Maximizing over each term in the summation, we obtain that

$$\begin{aligned}R &\leq \max_{p_{u,v,x}}I(u;y_r,s|v)-I(u;y_e|v)+2\varepsilon_n && (64)\\ &= \max_{p_{u,x}}I(u;y_r,s)-I(u;y_e)+2\varepsilon_n && (65)\end{aligned}$$

where the second step follows from the fact that the maximizing over v is redundant since (64) involves a convex combination of $I(u;y_r,s|v=v_i)-I(u;y_e|v=v_i)$ and hence we can replace

[2] Intuitively for any typical s^n, the total number of sequences u^n is $2^{nI(u;y_r|s)}$. The probability that a sequence u^n is jointly typical with y_e^n is $2^{-nI(u;y_e|s)}$. A precise argument involves bounding the expected size of the list and invoking a concentration result. See cf. [34, Lemma 1] for an analogous calculation.

with the term that results in the largest value. We recover (18) from (65) by using an approach similar to (20).

VII. Symmetric CSI: Numerical Example

In this section, we provide numerical computations of the achievable secret-key rate for an on–off channel

$$\begin{aligned} y_r &= s_r x + z_r \\ y_e &= s_e x + z_e \end{aligned} \tag{66}$$

where both s_r, $s_e \in \{0,1\}$, the random variables are mutually independent and equiprobable. Furthermore, we assume that s_r is revealed to the legitimate terminals, whereas the eavesdropper is revealed $\tilde{y}_e = (\mathsf{s}_e, y_e)$. The noise random variables are mutually independent and Gaussian with zero mean and unit variance. The power constraint $E[x^2] \leq P$ also holds.

We evaluate the rate expression for Gaussian inputs, i.e., $x \sim \mathcal{N}(0, P_0)$ when $\mathsf{s}_r = 0$ and $u = x \sim \mathcal{N}(0, P_1)$ when $\mathsf{s}_r = 1$. Further to satisfy the average power constraint we have that $P_0 + P_1 \leq 2P$. From Theorem 3, the following rate is achievable:

$$\begin{aligned} R &= I(u; y_r|\mathsf{s}_r) - I(u; \tilde{y}_e|\mathsf{s}_r) + H(\mathsf{s}_r|\tilde{y}_e) && (67) \\ &= I(u; y_r|\mathsf{s}_r) - I(u; y_e, \mathsf{s}_e|\mathsf{s}_r) + H(\mathsf{s}_r|\mathsf{s}_e, y_e) && (68) \\ &= \frac{1}{8}\log(1+P_1) \\ &\quad + \frac{1}{2}E_{y_e}[H(p(y_e), 1-p(y_e))] + \frac{1}{2} && (69) \end{aligned}$$

where

$$p(y_e) = \frac{\mathcal{N}_{y_e}(0, P_0+1)}{\mathcal{N}_{y_e}(0, P_0+1) + \mathcal{N}_{y_e}(0, P_1+1)} \tag{70}$$

indicates the posterior distribution $\Pr(\mathsf{s}_r = 0|y_e)$.

In Fig. 5, we numerically evaluate this rate for SNR $= 17$ dB. For comparison, we also plot the corresponding rate with public discussion [2]

$$R_{\text{disc}} = \frac{1}{8}\log(1+2P_1) + \frac{1}{2}E_{y_e}[H(p(y_e), 1-p(y_e))] + \frac{1}{2}. \tag{71}$$

In Fig. 5, the solid curves show the secret-key rate with and without public discussion, while the dashed curve is the entropy $H(\mathsf{s}_r|\mathsf{s}_e = 1, y_e)$ and the dotted curve denotes contribution of the wiretap code. Note that in general there is a trade-off between these two terms. To maximize the conditional entropy, we set $P_0 = P_1 = P/2$, while to maximize the wiretap codebook rate we need to set $P_0 = 0$ and $P_1 = P$. The resulting secret-key rate is maximized by selecting a power allocation that balances these two terms.

VIII. Conclusions

We investigated the secret-key agreement capacity over a wiretap channel controlled by a state parameter. Lower and upper bounds on the capacity are established when the state sequence is known noncausally to the encoder. The lower bound is obtained by creating a common reconstruction sequence at the legitimate terminals and binning the set of reconstruction sequences to generate a secret-key. Our bounds coincide in several special cases establishing the capacity results. While the present paper only treats the case without public discussion, we refer to the reader to [1] and [2] for some results on public discussion.

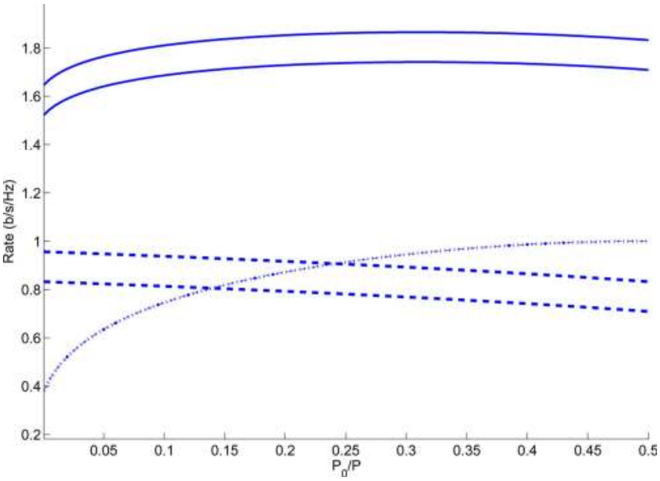


Fig. 5. Achievable secret-key rate as a fraction of power allocated to the state $\mathsf{s}_r = 0$ and SNR $= 17$ dB. The solid curves denote the secret-key rate, the dashed curve denotes the rate of the secret-message, while the dotted curve denotes the conditional entropy term $H(\mathsf{s}_r|\mathsf{s}_e = 1, y_e = y_e)$ in (69). The upper solid and dashed curves denote the case of public discussion while the other solid and dashed curves denote the case of no public discussion.

References

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

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

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

[4] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography—Part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 7, pp. 1121–1132, Jul. 1993.

[5] I. Csiszar and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004.

[6] I. Csiszar and P. Narayan, "Secrecy generation for multiple input multiple output channel models," in *Proc. Int. Symp. Inform. Theory*, 2009, pp. 2447–2451.

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

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

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

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

[11] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, "Robust key generation from signal envelopes in wireless networks," in *Proc. 14th ACM Conf. Computer and Communications Security*, 2007, pp. 401–410.

[12] 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. Antennas Propag.*, vol. 53, no. 11, pp. 3776–3784, Nov. 2005.

[13] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, “Fingerprints in the ether: Using the physical layer for wireless authentication,” in *Proc. IEEE Int. Conf. Comm.*, 2007, pp. 4646–4651.

[14] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. McLaughlin, “Wireless information-theoretic security,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.

[15] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, “Radiotelepathy: Extracting a secret key from an unauthenticated wireless channel,” in *Proc. 14th ACM Int. Conf. Mobile Computing and Networking*, 2008, pp. 128–139.

[16] C. E. Shannon, “Channels with side information at the transmitter,” *IBM J. Res. Devel.*, vol. 2, pp. 289–293, Oct. 1958.

[17] S. I. Gel’fand and M. S. Pinsker, “Coding for channels with random parameters,” *Problems of Control and Information Theory*, vol. 9, pp. 19–31, 1980.

[18] J. Wolfowitz*, Coding Theorems of Information Theory*. New York: Springer Verlag, 1978.

[19] G. Caire and S. Shamai, “On the capacity of some channels with channel state information,” *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.

[20] G. Caire and S. Shamai, “On achievable rates in a multi-antenna Gaussian broadcast channel,” in *Proc. Int. Symp. Inform. Theory*, Washington, DC, Jun. 2001, pp. 147–147.

[21] B. Chen and G. W. Wornell, “Quantization index modulation: A class of provably good methods for digital watermarking and information embedding,” *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[22] Y. Chia and A. E. Gamal, “Wiretap channel with causal state information,” in *Proc. Int. Symp. Inf. Theory*, Austin, TX, Jun. 2010.

[23] C. Mitrapant, H. Vinck, and Y. Luo, “An achievable region for the Gaussian wiretap channel with side information,” *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2181–2190, May 2006.

[24] Y. Chen and H. Vinck, “Wiretap channel with side information,” in *Proc. Int. Symp. Inf. Theory*, Seattle, WA, Jun. 2006.

[25] W. Liu and B. Chen, “Wiretap channel with two-sided state information,” in *Proc. 41st Asilomar Conf. Signals, Systems and Comp.*, Nov. 2007.

[26] V. Prabhakaran, K. Eswaran, and K. Ramchandran, “Secrecy via sources and channels,” *IEEE Trans. Inf. Theory*, Nov. 2009, submitted for publication.

[27] G. Keshet, Y. Steinberg, and N. Merhav, “Channel coding in the presence of side information,” *Found. Trends Commun. Inf. Theory*, vol. 4, Jun. 2007.

[28] A. E. Gamal and Y. H. Kim, Lecture Notes on Network Information Theory 2010 [Online]. Available: http://arxiv.org/abs/1001.3404

[29] Y. Steinberg, “Simultaneous transmission of data and state with common knowledge,” in *Proc. Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 935–939.

[30] Y. Steinberg, “Coding and common reconstruction,” *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4995–5010, Nov. 2009.

[31] A. Khisti, A. Tchamkerten, and G. W. Wornell, “Secure broadcasting over fading channels,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, Jun. 2008.

[32] M. H. Costa, “Writing on dirty paper,” *IEEE Trans. Inf. Theory*, vol. 29, no. 5, pp. 439–441, May 1983.

[33] A. Khisti and G. W. Wornell, “Secure transmission with multiple antennas I: The MISOME wiretap channel,” *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2009.

[34] Y. Chia and A. E. Gamal, “3-receiver broadcast channels with common and confidential messages,” *IEEE Trans. Inf. Theory*, 2009, submitted for publication.

[35] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” *IEEE Trans. Inf. Theory*, vol. IT-24, no. 10, pp. 339–348, Oct. 1978.

Ashish Khisti (S’01–M’08) received the B.A.Sc. degree in engineering sciences from the University of Toronto, Toronto, ON, Canada, and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, MA.

He is an assistant professor in the Electrical and Computer Engineering (ECE) Department, University of Toronto. His research interests span the areas of information theory, wireless physical layer security, and multimedia communication systems. At the University of Toronto, he heads the Signals, Multimedia and Security Laboratory.

For his graduate studies, Dr. Khisti was a recipient of the NSERC postgraduate fellowship, HP/MIT alliance fellowship, Harold H. Hazen Teaching award, and the Morris Joseph Levin Masterworks award.

Suhas N. Diggavi (S’93–M’99) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1998.

After completing the Ph.D. degree, he was a Principal Member Technical Staff in the Information Sciences Center, AT&T Shannon Laboratories, Florham Park, NJ. After that, he was on the faculty at the School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, where he directed the Laboratory for Information and Communication Systems (LICOS). He is currently a Professor in the Department of Electrical Engineering, University of California, Los Angeles. His research interests include wireless communications networks, information theory, network data compression, and network algorithms. He has eight issued patents.

Dr. Diggavi is a recipient of the 2006 IEEE Donald Fink prize paper award, 2005 IEEE Vehicular Technology Conference Best Paper Award, and the Okawa Foundation Research Award. He is currently an editor for ACM/IEEE TRANSACTIONS ON NETWORKING and the IEEE TRANSACTIONS ON INFORMATION THEORY.

Gregory W. Wornell (S’83–M’88–SM’01–F’04) received the B.A.Sc. degree (with honors) from the University of British Columbia, Canada, and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, all in electrical engineering and computer science, in 1985, 1987 and 1991, respectively.

Since 1991, he has been on the faculty at MIT, where he is Professor of Electrical Engineering and Computer Science. At MIT he leads the Signals, Information, and Algorithms Laboratory within the Research Laboratory of Electronics, and codirects the MIT Center for Wireless Networking. He is also chair of Graduate Area I (Systems, Communication, Control, and Signal Processing) within the Electrical Engineering and Computer Science Department’s doctoral program, and a member of the MIT Computational and Systems Biology Initiative. He has held visiting appointments at the Department of Electrical Engineering and Computer Science at the University of California, Berkeley, in 1999–2000, at Hewlett-Packard Laboratories, Palo Alto, CA, in 1999, and at AT&T Bell Laboratories, Murray Hill, NJ, in 1992–1993. His research interests and publications span the areas of signal processing, digital communication, and information theory, and include algorithms and architectures for wireless and sensor networks, broadband systems, and multimedia environments.

Prof. Wornell has been involved in the Signal Processing and Information Theory societies of the IEEE in a variety of capacities, and maintains a number of close industrial relationships and activities. He has won a number of awards for both his research and teaching.