Secret Key Agreement Using Asymmetry in Channel State Knowledge

Ashish Khisti
Deutsche Telekom Inc. R&D Lab USA
Los Altos, CA, 94040
Email: ashish.khisti@telekom.com

Suhas Diggavi
LICOS, EPFL
Lausanne, Switzerland
Email: suhas.diggavi@epfl.ch

Gregory Wornell
EECS Dept., MIT
Cambridge, MA, 02139
Email: gww@mit.edu

***Abstract*—We study secret-key agreement protocols over a wiretap channel controlled by a state parameter. The secret-key capacity is established when the wiretap channel is discrete and memoryless, the sender and receiver are both revealed the underlying state parameter, and no public discussion is allowed. An optimal coding scheme involves a two step approach — (i) design a wiretap codebook assuming that the state parameter is also known to the eavesdropper (ii) generate an additional secret key by exploiting the uncertainty of the state parameter at the eavesdropper. When unlimited public discussion is allowed between the legitimate terminals, we provide an upper bound on the secret-key capacity and establish its tightness when the channel outputs of the legitimate receiver and eavesdropper satisfy a *conditional independence* property. Numerical results for an on-off fading model suggest that the proposed coding schemes significantly outperform naive schemes that either disregard the contribution of the common state sequence or the contribution of the underlying channel.**

I. Introduction

Generating a shared secret-key between two terminals by exploiting the reciprocity in the physical wireless channel has received a lot of recent attention. See e.g, [1] and the references therein. The sender and receiver exchange pilot signals to learn the channel gains in uplink and downlink respectively. When the channels are reciprocal, the uplink and downlink gains are close to one another and this correlation is exploited to generate shared secret keys.

Motivated by these works we study the information theoretic problem of secret-key agreement over a channel controlled by one state parameter. This state parameter is revealed to both the sender and the receiver and not to the eavesdropper. In the fading model discussed above, this state parameter models the fading gain between the sender and the receiver. The sender and receiver can learn this value over reciprocal wireless channels by exchanging pilot signals, whereas the eavesdropper cannot directly learn this value. A good coding scheme for this problem exploits two sources of uncertainty at the eavesdropper — one due to the lack of knowledge of state parameter at the eavesdropper, and the other due to the equivocation introduced by the channel. As our capacity expression illustrates, there is in fact a balance between the gains from the two uncertainties.

In other related works, the case when an *independent message* needs to be transmitted over the wiretap channel with state parameters (and no public discussion) has been studied in [2], [3], [4]. Achievable rate-equivocation regions are provided for the case of either transmitter-side information or two-sided state information, but the complete characterization of this region remains open. In contrast to sending independent messages, the formulation studied in the present paper, allows the secret key to arbitrarily depend on the state sequence.

This work was supported in part by NSF under Grant No. CCF-0515109

II. Problem Statement

The problem setup is described in Fig. 1. The sender and receiver communicate over a discrete-memoryless-wiretap-channel with input symbol x, the output at the legitimate receiver y_{r} and the output at the eavesdropper y_{e}. The channel transition probability is conditioned on state parameter s_r is specified by

$$\Pr(y_{\mathrm{r}}^n = y_r^n, y_{\mathrm{e}}^n = y_e^n | x^n = x^n, \mathsf{s}_r^n = s_r^n) = \prod_{i=1}^{n} p_{y_r y_{\mathrm{e}}|x \mathsf{s}_r}(y_{ri}, y_{ei}|x_i, s_{ri}) \tag{1}$$

where the state parameter sequence s_r^n is sampled i.i.d. from the distribution $p_{\mathsf{s}_r}(\cdot)$.

In defining a length-n encoder and decoder, we will assume that the state sequence s_r^n is known non-causally to the sender *and* the receiver. However our coding theorems only require a causal knowledge of the state sequence. We first separately consider the case when no public discussion is allowed between the encoder and the decoder and when unlimited discussion is allowed.

A. No Public Discussion

A length n encoder is defined as follows. The sender samples a random variables u from the conditional distribution $p_{u|\mathsf{s}_r^n}(\cdot|s_r^n)$. The encoding function produces a channel input sequence $x^n = f_n(u, \mathsf{s}_r^n)$ 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 $p_{y_r, y_{\mathrm{e}}|x,\mathsf{s}_r}(\cdot)$. The sender and receiver compute secret keys $\kappa = g_n(u, \mathsf{s}_r^n)$ and $l = h_n(\mathsf{s}_r^n, y_{\mathrm{r}}^n)$. 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 $\frac{1}{n} H(\kappa) \geq R - \varepsilon_n$ and $\frac{1}{n} I(\kappa; y_{\mathrm{e}}^n) \leq \varepsilon_n$. The largest achievable rate is the secret-key capacity.

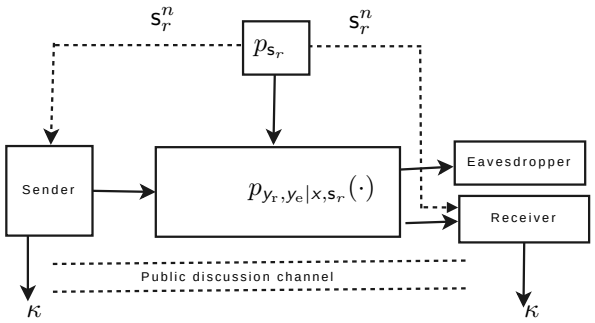

Fig. 1. Secret-key agreement over a wiretap channel controlled with a state parameter. The channel is a discrete-memoryless-broadcast channel. The state parameter s_r is sampled i.i.d. and revealed to the sender and receiver. We separately consider two cases (a) unlimited interactive public discussion is allowed between the sender and the receiver and (b) no such discussion is allowed.

Remark 1: Note that the formulation can be easily extended to include the case when the eavesdropper is also revealed the state parameter s_r. This can be done by re-defining the output symbol $\tilde{y_{\mathrm{e}}} = (y_{\mathrm{e}}, \mathsf{s}_r)$. More generally, our formulation also extends to the case when the eavesdropper observes a state parameter s_e correlated with s_r and the channel transition probability is $p_{y_{\mathrm{r}},y_{\mathrm{e}}|x,\mathsf{s}_r,\mathsf{s}_e}(\cdot)$. In this case it sufficies to consider the channel where the eavesdropper observes $\tilde{y_{\mathrm{e}}} = (y_{\mathrm{e}}, \mathsf{s}_e)$ with transition probability

$$p_{y_{\mathrm{r}},\tilde{y}_{\mathrm{e}}|x,\mathsf{s}_r}(y_r, \tilde{y}_e|x, s_r) = p_{y_{\mathrm{r}},y_{\mathrm{e}}|x,\mathsf{s}_r,\mathsf{s}_e}(y_r, y_e|x, s_r, s_e) p_{\mathsf{s}_e|\mathsf{s}_r}(s_e|s_r), \tag{2}$$

which reduces to the present formulation.

B. Presence of Public Discussion

When a public discussion channel is present, the described protocol follows closely the interactive communication protocol in [5]. The sender transmits symbols $x_1, \ldots, x_n$ at times $0 < i_1 < i_2 < \ldots < i_n$ over the wiretap channel. At these times the receiver and the eavesdropper observe symbols $y_{r1}, \ldots, y_{rn}$ and $y_{e1}, \ldots, y_{en}$ respectively. In the remaining times the sender and receiver exchange messages ψ_t and ϕ_t where $1 \leq t \leq k$. For convenience we let $i_{n+1} = k + 1$. The eavesdropper observes both ψ_t and ϕ_t.

More specifically the sender and receiver sample random variables u and v from conditional distributions $p_{u|\mathsf{s}_r^n}(\cdot|s_r^n)$ and $p_{v|\mathsf{s}_r^n}(\cdot|s_r^n)$ and observe that $u \to \mathsf{s}_r^n \to v$.

- At times $0 < t < i_1$, the sender generates $\phi_t = \Phi_t(u, \mathsf{s}_r^n, \psi^{t-1})$ and the receiver generates $\psi_t = \Psi_t(v, \mathsf{s}_r^n, \phi^{t-1})$. These messages are exchanged over the public channel.
- At times i_j , $1 \leq j \leq n$, the sender generates $x_j = X_j(u, \mathsf{s}_r^n, \psi^{i_j-1})$ and sends it over the channel. The receiver and eavesdropper observe $y_{r,j}$ ad $y_{e,j}$ respectively. For these times we set $\psi_{i_j} = \phi_{i_j} = 0$.
- For times $i_j < t < i_{j+1}$, where $1 \leq j \leq n$, the sender and receiver compute $\phi_t = \Phi_t(u, \mathsf{s}_r^n, \psi^{t-1})$ and $\psi_t = \Psi_t(v, \mathsf{s}_r^n, y_{\mathrm{r}}^j, \phi^{t-1})$ respectively and exchange them over the public channel.
- At time $k + 1$, the sender and receiver compute $\kappa = g_n(u, \mathsf{s}_r^n, \psi^k)$ and the receiver computes $l = h_n(v, \mathsf{s}_r^n, y_{\mathrm{r}}^n, \phi^k)$.

We require that for some sequence ε_n that vanishes as $n \to \infty$, $Pr(\kappa \neq l) \leq \varepsilon_n$ and $\frac{1}{n} I(\kappa; y_{\mathrm{e}}^n, \psi^k, \phi^k) \leq \varepsilon_n$.

III. Summary of Results

A. No Public Discussion

The following theorem characterizes the secret-key capacity when no public-discussion channel is available between the legitimate terminals.

Theorem 1: The secret-key capacity for the channel model in section II-A is

$$C = \max_{\mathcal{P}} \left\{ I(t; y_{\mathrm{r}}|\mathsf{s}_r) - I(t; y_{\mathrm{e}}|\mathsf{s}_r) + H(\mathsf{s}_r|y_{\mathrm{e}}) \right\}, \tag{3}$$

where $\mathcal{P}$ is the set of all joint distributions $p_{t,x,\mathsf{s}_r,y_{\mathrm{r}},y_{\mathrm{e}}}(\cdot)$ that satisfy the Markov chain $t \to (x, \mathsf{s}_r) \to (y_{\mathrm{r}}, y_{\mathrm{e}})$. Furthermore in (3) it suffices to maximize over the auxiliary random variables t whose cardinality is bounded by $|\mathcal{S}_r|(1 + |\mathcal{X}|)$.□

Remark 2: The expression in (3) can be interpreted as generating two independent keys. The first key at rate $R_{\mathrm{ch}} = I(t; y_{\mathrm{r}}|\mathsf{s}_r) - I(t; y_{\mathrm{e}}|\mathsf{s}_r)$ is achieved by transmitting an independent message with perfect secrecy using a wiretap codebook for a modified channel where s_r^n is revealed to the eavesdropper (in addition to legitimate terminals). The second key, which is independent of the first key and has a rate of $R_{\mathrm{src}} = H(\mathsf{s}_r|y_{\mathrm{e}})$ is produced by exploiting the common knowledge of s_r^n at the legitimate terminals. This intuition is formalized in the achievability scheme in section V-A.

Next, consider the case when for each $\mathsf{s}_r \in \mathcal{S}_{\mathrm{r}}$, the channel of the eavesdropper is less noisy than the channel of the legitimate receiver i.e., $\max_{\mathcal{P}} I(t; y_{\mathrm{r}}|\mathsf{s}_r) - I(t; y_{\mathrm{e}}|\mathsf{s}_r) = 0$. In this case, the secret-key capacity reduces to

$$C = \max_{p_{x|\mathsf{s}_r}} H(\mathsf{s}_r|y_{\mathrm{e}}).$$

It is achieved by generating the secret-key based on the common knowledge of s_r^n between the legitimate terminals and choosing an input distribution that leaks minimal information about s_r^n to the eavesdropper. More generally, there is a balance between the amount of information leaked to the eavesdropper and the ability to transmit information over the wiretap channel in the capacity achieving scheme. This balance is reflected in the maximization in (3).

B. Unlimited Public Discussion

When unlimited public discussion is allowed between the sender and the receiver, as described in section II-B, we have the following result on the secret-key capacity.

Theorem 2: The secret-key capacity in the presence of unlimited public discussion between the sender and the receiver is

$$C = \max_{p_{x|\mathsf{s}_r}} I(x; y_{\mathrm{r}}|y_{\mathrm{e}}, \mathsf{s}_r) + H(\mathsf{s}_r|y_{\mathrm{e}}). \tag{4}$$

when the channel satisfies the Markov condition $y_{\mathrm{r}} \rightarrow (x, \mathsf{s}_r) \rightarrow y_{\mathrm{e}}$. For any discrete memoryless channel (4) provides an upper bound to the secret-key capacity.

Remark 3: The Markov condition in Theorem 2 can be interpreted as requiring that the noise on the legitimate and the eavesdroppers channel be mutually independent. Furthermore, analogous to the capacity expression in Theorem 1 the expression in (4) also involves a sum of two terms and accordingly the lower bound is constructed by generating two separate keys.

IV. Numerical Example

We consider the the following on-off channel for the receivers:

$$\begin{aligned} y_{\mathrm{r}} &= \mathsf{s}_r x + z_{\mathrm{r}} \\ y_{\mathrm{e}} &= \mathsf{s}_e x + z_{\mathrm{e}}, \end{aligned} \tag{5}$$

where both $\mathsf{s}_r, \mathsf{s}_e \in \{0, 1\}$, the random variables are mutually independent and $\Pr(\mathsf{s}_r = 0) = \Pr(\mathsf{s}_e = 0) = 0.5$. Furthermore we assume that s_r is revealed to the legitimate terminals, whereas the eavesdropper is revealed $\tilde{y_{\mathrm{e}}} = (\mathsf{s}_e, y_{\mathrm{e}})$. The noise random variables are mutually independent, zero mean and unit variance Gaussian random variables and the power constraint is that $E[x^2] \leq P$.

We evaluate the secret-key rate expression for Gaussian inputs i.e., $t = x \sim \mathcal{N}(0, P_0)$ when $\mathsf{s}_{\mathrm{r}} = 0$ and $t = x \sim \mathcal{N}(0, P_1)$ when $\mathsf{s}_{\mathrm{r}} = 1$. Further to satisfy the average power constraint we have that $P_0 + P_1 \leq 2P$. An achievable rate from Theorem 1 and Remark 1 is,

$$\begin{aligned} R &= I(x; y_{\mathrm{r}}|\mathsf{s}_r) - I(x; \tilde{y_{\mathrm{e}}}|\mathsf{s}_r) + H(\mathsf{s}_r|\tilde{y_{\mathrm{e}}}) && (6) \\ &= I(x; y_{\mathrm{r}}|\mathsf{s}_r) - I(x; y_{\mathrm{e}}, \mathsf{s}_e|\mathsf{s}_r) + H(\mathsf{s}_r|\mathsf{s}_e, y_{\mathrm{e}}) && (7) \end{aligned}$$

Substituting (5) above and simplifying, we have that

$$R = \frac{1}{8}\log(1 + P_1) + \frac{1}{2}E_{y_{\mathrm{e}}}[H(p(y_e), 1 - p(y_e))] + \frac{1}{2}, \tag{8}$$

where

$$p(y_e) = \frac{\mathcal{N}_{y_e}(0, P_0 + 1)}{\mathcal{N}_{y_e}(0, P_0 + 1) + \mathcal{N}_{y_e}(0, P_1 + 1)} \tag{9}$$

is the aposterior distribution $\Pr(\mathsf{s}_e = 0|y_e)$ which is used to numerically evaluate the second term in (8). Similarly by choosing Gaussian inputs in Theorem 2, the secret-key rate reduces to

$$R = \frac{1}{8}\log(1 + 2P_1) + \frac{1}{2}E_{y_{\mathrm{e}}}[H(p(y_e), 1 - p(y_e))] + \frac{1}{2}. \tag{10}$$

Fig. 2 illustrates the secret-key rate in (8),(10) as a function of the power allocation when SNR = 17 dB. There are three curves — the solid curve is the resulting secret key rate in (8), while the dashed curve is the entropy $H(\mathsf{s}_r|\mathsf{s}_e = 1, y_{\mathrm{e}})$ and the dotted curve denotes the secret-message rate. The upper solid and dashed curves denote the case of public discussion whereas the lower curves denote the case of no public discussion. Note that in general there is a tradeoff between these two terms. To maximize the conditional entropy we set $P_0 = P_1 = P/2$, while to maximize the secret-message rate we 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. The optimum fraction of power transmitted in the state $\mathsf{s}_r = 0$ as a function of the signal to noise ratio is shown in Fig. 3. Note that no power is transmitted when the signal-to-noise ratio is below $\approx -2.5dB$. In this regime the channels are sufficiently noisy so that $H(\mathsf{s}_r|y_{\mathrm{e}}, \mathsf{s}_e = 1) \approx 1$ even with $P_0 = 0$ and hence all the available power is used for transmitting the secret-message. As the signal-to-noise ratio increases more information regarding s_r gets leaked to the eavesdropper and to compensate for this effect, a non-zero fraction of power is transmitted when $\mathsf{s}_r = 0$.

V. Proofs

A. Coding Theorem for Theorem 1

The lower bound involves separately constructing two independent keys κ_{ch} and κ_{src} at rates $R_{\mathrm{ch}} = I(t; y_{\mathrm{r}}|\mathsf{s}_r) - I(t; y_{\mathrm{e}}|\mathsf{s}_r)$ and $R_{\mathrm{src}} = H(\mathsf{s}_r|y_{\mathrm{e}})$ respectively.

The key κ_{ch} is constructed by using a multiplexed coding scheme as follows. Let $\mathcal{S}_r = \{s_1, \ldots, s_m\}$ and let $p_i = \Pr(\mathsf{s}_i = s_i)$. For each i construct a wiretap codebook [6] of rate $R_i = I(t; y_{\mathrm{r}}|\mathsf{s}_r = s_i) - I(t; y_{\mathrm{e}}|\mathsf{s}_r = s_i) - 2\varepsilon_n$ consisting of $2^{np_i I(t; y_{\mathrm{r}}|\mathsf{s}_r = s_i) - \varepsilon_n}$ codewords each of length np_i and sampled i.i.d. from the distribution $p_{t|\mathsf{s}_r = s_i}(\cdot)$. For each codebook an independent message w_i uniformly distributed over the set $\{1, 2, \ldots, 2^{np_i R_i}\}$ is selected and corresponding codeword symbols are transmitted whenever $\mathsf{s} = s_i$. The resulting key is the collection of these messages i.e., $\kappa_{\mathrm{ch}} = (w_1, \ldots, w_m)$. Clearly the multiplexed codebook has a rate of

$$\begin{aligned} R_{\mathrm{ch}} &= \sum_{i=1}^{m} p_i R_i && (11) \\ &= I(t; y_{\mathrm{r}}|\mathsf{s}_r) - I(t; y_{\mathrm{e}}|\mathsf{s}_r) - 2\varepsilon_n && (12) \end{aligned}$$

as required. We further show below that

$$\frac{1}{n}H(\kappa_{\mathrm{ch}}|y_{\mathrm{e}}^n, \mathsf{s}_r^n) = R_{\mathrm{ch}} - o_n(1), \tag{13}$$

where $o_n(1)$ decays to zero as $n \rightarrow \infty$. From the analysis of each wiretap codebook $\mathcal{C}_i$, we have that

$$\frac{1}{n}H(w_i|y_{\mathrm{e|i}}^n) = \frac{1}{n}H(w_i) - \varepsilon_n$$

where $y_{\mathrm{e|i}}^n$ is the projection of y_{e}^n onto those time indices where the state parameter takes the value $\mathsf{s}_r = s_i$. Furthermore, since the messages and codewords are selected independently in

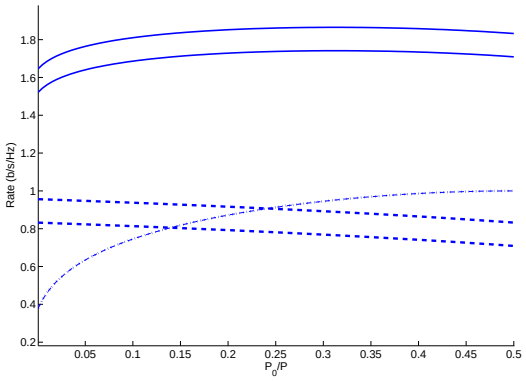

Fig. 2. The achievable secret-key rate as a fraction of power allocated to the state $\mathsf{s}_r = 0$ and SNR = 17 dB. The solid curve denotes the secret-key rate, the dashed curve denotes the rate of the secret-message, while the dotted curve denotes the conditional entropy term $H(\mathsf{s}_r|\mathsf{s}_e = 1, y_e = y_e)$ in (8). 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.

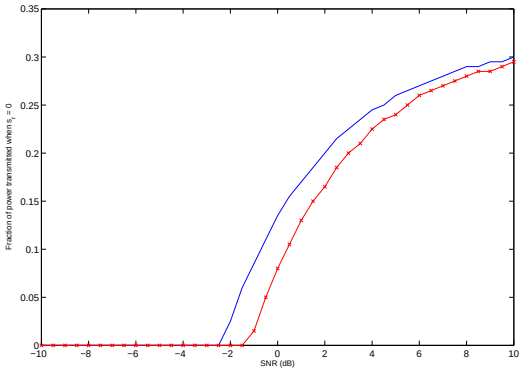

Fig. 3. Optimal fraction of power that must be allocated to the state $\mathsf{s}_r = 0$ to maximize the secret-key rate with Gaussian inputs. The curve marked with a ($\times$) denotes the case of public discussion while the other curve denotes the case of no public discussion.

each codebook

$$\frac{1}{n}H(w_i|y_{\rm e}^n, \mathsf{s}_r^n, w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_m) = \frac{1}{n}H(w_i|y_{{\rm e}|{\rm i}}^n) = \frac{1}{n}H(w_i) - \varepsilon_n. \quad (14)$$

Finally we have that

$$\frac{1}{n}H(\kappa_{\rm ch}|y_{\rm e}^n, \mathsf{s}_r^n) = \frac{1}{n}H(w_1, \ldots, w_m|y_{\rm e}^n, \mathsf{s}_r^n)$$
$$\geq \frac{1}{n}\sum_{i=1}^{m} H(w_i|w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_m, y_{\rm e}^n, \mathsf{s}_r^n)$$
$$\geq \frac{1}{n}\sum_{i=1}^{m} H(w_i) - m\varepsilon_n = R_{\rm ch} - m\varepsilon_n$$

as required.

The remaining key $\kappa_{\rm src}$ is obtained from the sequence s_r^n via random binning. No additional communication between the sender and receiver is required in this step. To generate this key, the set of all typical sequences s_r^n is partitioned into $2^{nH(\mathsf{s}_r|y_{\rm e})}$ bins, each consisting $2^{n(I(\mathsf{s}_r;y_{\rm e})-\varepsilon_n)}$ sequences. The secret-key $k_{\rm src}$ is the bin index of the sequence s_r^n. Using standard arguments it can be shown that

$$\frac{1}{n}H(\kappa_{\rm src}|y_{\rm e}^n) = H(\mathsf{s}_r|y_{\rm e}) - o_n(1). \quad (15)$$

To complete the secrecy analysis of our codebook, note that

$$\frac{1}{n}H(\kappa_{\rm src}, \kappa_{\rm ch}|y_{\rm e}^n)$$
$$= \frac{1}{n}H(\kappa_{\rm src}|y_{\rm e}^n) + \frac{1}{n}H(\kappa_{\rm ch}|y_{\rm e}^n, \kappa_{\rm src})$$
$$\geq \frac{1}{n}H(\kappa_{\rm src}|y_{\rm e}^n) + \frac{1}{n}H(\kappa_{\rm ch}|y_{\rm e}^n, \kappa_{\rm src}, \mathsf{s}_r^n)$$
$$= \frac{1}{n}H(\kappa_{\rm src}|y_{\rm e}^n) + \frac{1}{n}H(\kappa_{\rm ch}|y_{\rm e}^n, \mathsf{s}_r^n) \quad (16)$$
$$= H(\mathsf{s}_r|y_{\rm e}) + I(t; y_{\rm r}|\mathsf{s}_r) - I(t; y_{\rm e}|\mathsf{s}_r) - o_n(1) \quad (17)$$

where (16) follows from the fact that $\kappa_{\rm src}$ is a deterministic function of s_r^n and the last step is obtained by substituting (13) and (15) for the two equivocation terms.

B. Converse for Theorem 1

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 (3) plus a term that vanishes to zero as the block length goes to zero. Apply Fano's inequality on the secret-key rate, we have that for some sequence ε_n that approaches zero as n goes to infinity

$$nR \leq I(\kappa; l) + n\varepsilon_n \leq I(\kappa; \mathsf{s}_r^n, y_{\rm r}^n) + n\varepsilon_n \quad (18)$$

where the last step follows from the data processing inequality since $l = h_n(\mathsf{s}_r^n, y_{\rm r}^n)$. Furthermore from the secrecy condition we have that

$$nR \leq I(\kappa; \mathsf{s}_r^n, y_{\rm r}^n) - I(\kappa; y_{\rm e}^n) + 2n\varepsilon_n \quad (19)$$
$$= \sum_{i=1}^{n} I(\kappa, y_{\rm e}^{i-1}\mathsf{s}_{r,i+1}^n y_{r,i+1}^n; \mathsf{s}_{r,i}, y_{r,i}) - \sum_{i=1}^{n} I(\kappa, y_{\rm e}^{i-1}\mathsf{s}_{r,i+1}^n y_{r,i+1}^n; y_{e,i}) + 2n\varepsilon_n \quad (20)$$

where the second step follows from the well known chain rule applied to difference of mutual informations (see e.g., [6]). For each $i = 1, 2, \ldots, n$, define the random variable $t_i = (\kappa, y_{\rm e}^{i-1}\mathsf{s}_{r,i+1}^n y_{r,i+1}^n)$. By noting that the encoding functions are defined by $\kappa = g_n(u, \mathsf{s}_r^n)$ and $x^n = f_n(u, \mathsf{s}_r^n)$, it can be verified that the joint distribution, $p_{t_i, x_i, \mathsf{s}_{r,i}, y_{r,i}, y_{e,i}} \in \mathcal{P}$. Thus we have from (20) that

$$R - 2\varepsilon_n \leq \frac{1}{n}\sum_{i=1}^{n} I(t_i; \mathsf{s}_{r,i}, y_{r,i}) - I(t_i; y_{e,i})$$
$$\leq \max_{\mathcal{P}} \{I(t; \mathsf{s}_r, y_r) - I(t; y_e)\}$$

$$= \max_{\mathcal{P}} \{I(t; y_r, s_r) - I(t; y_e, \mathsf{s}_r) + I(t; \mathsf{s}_r|y_\mathrm{e})\} \quad (21)$$

$$= \max_{\mathcal{P}} \{I(t; y_r|s_r) - I(t; y_e|\mathsf{s}_r) + I(t; \mathsf{s}_r|y_\mathrm{e})\} \quad (22)$$

$$= \max_{\mathcal{P}} \{I(t; y_r|s_r) - I(t; y_e|\mathsf{s}_r) + H(\mathsf{s}_r|y_\mathrm{e})\}, \quad (23)$$

where (21) and (22) both follow from the chain rule of mutual information and the last expression follows by observing that if $t^\star$ is any random variable that maximizes (22), then selecting $t = (t^\star, \mathsf{s}_r)$, we have

$$I(t^\star; y_r|\mathsf{s}_r) - I(t^\star; y_e|\mathsf{s}_r) = I(t; y_r|s_r) - I(t; y_e|\mathsf{s}_r)$$

while the term $I(t^\star; \mathsf{s}_r|y_\mathrm{e})$ increases to $H(\mathsf{s}_r|y_\mathrm{e})$. Hence for the maximizing distribution, we can replace $I(t; \mathsf{s}_r|y_\mathrm{e})$ by the entropy term as in (23).

C. Coding Theorem for Theorem 2

The proposed coding scheme is a direct extension of the coding schemes in [7], [5] that consider the channel with no state parameters. The sender chooses a distribution $p_{x|\mathsf{s}_r}$ and given s_r, samples the channel input symbol from this distribution. Upon observing $(y_\mathrm{r}^n, \mathsf{s}_r^n)$, the receiver transmits the bin-index from a Slepian-Wolf code over a public channel. The resulting secret-key rate is

$$\begin{aligned} R &= \max_{p_{x|\mathsf{s}_r}} I(x, \mathsf{s}_r; y_\mathrm{r}, \mathsf{s}_r) - I(y_\mathrm{e}; y_\mathrm{r}, \mathsf{s}_r) \\ &= \max_{p_{x|\mathsf{s}_r}} I(x; y_\mathrm{r}|\mathsf{s}_r) - I(y_\mathrm{e}; y_\mathrm{r}|\mathsf{s}_r) + H(\mathsf{s}_r|y_\mathrm{e}). \end{aligned}$$

which matches the upper bound (4) when the Markov condition $y_\mathrm{r} \to (x, \mathsf{s}_r) \to y_\mathrm{e}$ is satisfied.

D. Upper Bound for Theorem 2

For any sequence of encoding and decoding functions, we have from Fano's inequalty that

$$\begin{aligned} nR &\le I(\kappa; l) + n\varepsilon_n \\ &\le I(\kappa; v, \mathsf{s}_r^n, y_\mathrm{r}^n, \phi^k) + n\varepsilon_n \quad (24) \\ &\le I(\kappa; v, \mathsf{s}_r^n, y_\mathrm{r}^n, \phi^k) - I(\kappa; y_\mathrm{e}^n, \phi^k, \psi^k) + 2n\varepsilon_n \quad (25) \\ &\le I(\kappa; v, \mathsf{s}_r^n, y_\mathrm{r}^n|\psi^k, \phi^k, y_\mathrm{e}^n) + 2n\varepsilon_n \\ &\le I(\kappa; v, y_\mathrm{r}^n|\psi^k, \phi^k, y_\mathrm{e}^n, \mathsf{s}_r^n) + I(\kappa; \mathsf{s}_r^n|\psi^k, \phi^k, y_\mathrm{e}^n) + 2n\varepsilon_n \\ &\le I(\kappa; v, y_\mathrm{r}^n|\psi^k, \phi^k, y_\mathrm{e}^n, \mathsf{s}_r^n) + H(\mathsf{s}_r^n|y_\mathrm{e}^n) + 2n\varepsilon_n \\ &\le I(\kappa; v, y_\mathrm{r}^n|\psi^k, \phi^k, y_\mathrm{e}^n, \mathsf{s}_r^n) + \sum_{i=1}^n H(\mathsf{s}_{r,i}|y_{e,i}) + 2n\varepsilon_n \end{aligned}$$

where (24) follows from the fact that $l = h_n(v, \mathsf{s}_r^n, y_\mathrm{r}^n, \phi^k)$, and (25) follows from secrecy constraint. It sufficies to show that

$$I(\kappa; v, y_\mathrm{r}^n|\psi^k, \phi^k, y_\mathrm{e}^n, \mathsf{s}_r^n) \le \sum_{i=1}^n I(x_i; y_{r,i}|\mathsf{s}_{r,i}, y_{e,i}) \quad (26)$$

since,

$$\begin{aligned} nR &\le \sum_{i=1}^n I(x_i; y_{r,i}|\mathsf{s}_{r,i}, y_{e,i}) + H(\mathsf{s}_{r,i}|y_{e,i}) + 2n\varepsilon_n \\ &\le n\left(\max_{p_{x|\mathsf{s}_r}} I(x; y_\mathrm{r}|\mathsf{s}_r, y_\mathrm{e}) + H(\mathsf{s}_r|y_\mathrm{e}) + 2\varepsilon_n\right) \end{aligned}$$

as required. Thus it only remains to establish (26). The proof follows closely the upper bound derived in [5] with some modifications to take into account the state parameters. Since $\kappa = g_n(u, \psi^k, \mathsf{s}_r^n)$ we have that

$$\begin{aligned} &I(\kappa; v, y_\mathrm{r}^n|\psi^k, \phi^k, y_\mathrm{e}^n, \mathsf{s}_r^n) \le I(u; v, y_\mathrm{r}^n|\psi^k, \phi^k, \mathsf{s}_r^n, y_\mathrm{e}^n) \\ &= I(u; v, y_\mathrm{r}^n, \psi^k, \phi^k, \mathsf{s}_r^n, y_\mathrm{e}^n) - I(u; \mathsf{s}_r^n, \psi^k, \phi^k, y_\mathrm{e}^n) \\ &= I(u; v, \psi^{i_1-1}, \phi^{i_1-1}, \mathsf{s}_r^n) - I(u; \mathsf{s}_r^n, \psi^{i_1-1}, \phi^{i_1-1}) \\ &\quad + I(u; y_\mathrm{r}^n, y_\mathrm{e}^n, \psi_{i_1+1}^k, \phi_{i_1+1}^k|v, \psi^{i_1-1}, \phi^{i_1-1}, \mathsf{s}_r^n) \\ &\quad - I(u; y_\mathrm{e}^n, \psi_{i_1+1}^k, \phi_{i_1+1}^k|\mathsf{s}_r^n, \psi^{i_1-1}, \phi^{i_1-1}) \\ &= I(u; v|\mathsf{s}_r^n, \phi^{i_1-1}, \psi^{i_1-1}) + \sum_{j=1}^n F_{r,j} - F_{e,j} + \sum_{j=1}^n G_{r,j} - G_{e,j} \end{aligned} \quad (27)$$

where we have introduced

$$\begin{aligned} F_{r,j} &= I(u; \mathsf{y}_{rj}, \mathsf{y}_{ej}|\mathsf{s}_r^n, v, \phi^{i_j-1}, \psi^{i_j-1}, y_\mathrm{e}^{j-1}, y_\mathrm{r}^{j-1}) \\ F_{e,j} &= I(u; \mathsf{y}_{ej}|\phi^{i_j-1}, \psi^{i_j-1}, y_\mathrm{e}^{j-1}, \mathsf{s}_r^n), \end{aligned} \quad (28)$$

$G_{r,j} = I(u; \phi_{i_j+1}, \ldots, \phi_{i_{j+1}-1}, \psi_{i_j+1} \ldots, \psi_{i_{j+1}-1}|\phi^{i_j-1}, \psi^{i_j-1}, y_\mathrm{r}^j, y_\mathrm{e}^j, v, \mathsf{s}_r^n)$ and $G_{\mathrm{e,j}} = I(u; \phi_{i_j+1}, \ldots, \phi_{i_{j+1}-1}, \psi_{i_j+1}, \ldots, \psi_{i_{j+1}-1}|\phi^{i_j-1}, \psi^{i_j-1}, y_\mathrm{e}^j, \mathsf{s}_r^n)$.

To complete the proof it suffices to show that

$$I(u; v|\mathsf{s}_r^n, \phi^{i_1-1}, \psi^{i_1-1}) \le I(u; v|\mathsf{s}_r^n) = 0 \quad (29)$$

$$F_{r,j} - F_{e,j} \le I(x_j; \mathsf{y}_{rj}|\mathsf{y}_{ej}, s_{rj}) \quad (30)$$

$$G_{r,j} - G_{e,j} \le 0. \quad (31)$$

Due to space limitations, we only establish the first relation. The remaining two relations can be established in an analogous manner. Since $\phi_{i_1-1} = \Phi_{i_1-1}(u, \mathsf{s}_r^n, \psi^{i_1-2})$ and $\psi_{i_1-1} = \Psi_{i_1-1}(u, \mathsf{s}_r^n, \phi^{i_1-2})$ we have that

$$\begin{aligned} &I(u; v|\mathsf{s}_r^n, \phi^{i_1-1}, \psi^{i_1-1}) \\ &\le I(u, \phi_{i_1-1}; v, \psi_{i_1-1}|\mathsf{s}_r^n, \phi^{i_1-2}, \psi^{i_1-2}) \\ &= I(u; v|\mathsf{s}_r^n, \phi^{i_1-2}, \psi^{i_1-2}). \end{aligned}$$

Continuing this process, $I(u; v|\mathsf{s}_r^n, \phi^{i_1-1}, \psi^{i_1-1}) \le I(u; v|\mathsf{s}_r^n) = 0$.

REFERENCES

[1] R. Wilson, D. Tse, and R. Scholtz, "Channel Identification: Secret Sharing using Reciprocity in UWB Channels," *submitted to IEEE Transactions on Information Forensics and Security*, Mar. 2006.

[2] Y. Chen and H. Vinck, "Wiretap channel with side information," in *Proc. Int. Symp. Inform. Theory*, Jun. 2006.

[3] C. Mitrapant, H. Vinck, and Y. Luo, "An achievable region for the gaussian wiretap channel with side information," *IEEE Trans. Inform. Theory*, vol. 52, pp. 2181–2190, May 2006.

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

[5] 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.

[6] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, pp. 339–348, 1978.

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