A Broadcast Approach for Fading Wiretap Channels

Yingbin Liang, *Member, IEEE*, Lifeng Lai, *Member, IEEE*, H. Vincent Poor, *Fellow, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

***Abstract*— A (layered) broadcast approach is studied for the fading wiretap channel without the channel state information (CSI) at the transmitter. Two broadcast schemes, based on superposition coding and embedded coding, respectively, are developed to encode information into a number of layers and use stochastic encoding to keep the corresponding information secret from an eavesdropper. The layers that can be successfully and securely transmitted are determined by the channel states to the legitimate receiver and the eavesdropper. The advantage of these broadcast approaches is that the transmitter does not need to know the CSI to the legitimate receiver and the eavesdropper, but the scheme still adapts to the channel states of the legitimate receiver and the eavesdropper. Three scenarios of block fading wiretap channels with stringent delay constraints are studied, in which either the legitimate receiver's channel, the eavesdropper's channel, or both channels are fading. For each scenario, the secrecy rate that can be achieved via the broadcast approach developed in this paper is derived, and the optimal power allocation over the layers (or the conditions on the optimal power allocation) is also characterized. A notion of probabilistic secrecy, which characterizes the probability that a certain secrecy rate of decoded messages is achieved during one block, is also introduced and studied for scenarios when the eavesdropper's channel is fading. Numerical examples are provided to demonstrate the impact of the CSI at the transmitter and the channel fluctuations of the eavesdropper on the average secrecy rate. These examples also demonstrate the advantage of the proposed broadcast approach over the compound channel approach.**

***Index Terms*— Channel state information, fading channel, layered broadcast approach, secrecy rate, wiretap channel.**

Manuscript received October 27, 2011; revised December 29, 2012; accepted August 1, 2013. Date of publication December 18, 2013; date of current version January 15, 2014. Y. Liang was supported in part by the National Science Foundation CAREER Award under Grant CCF-10-26565 and in part by the National Science Foundation under Grants CCF-10-26566 and CNS-11-16932. L. Lai was supported in part by the National Science Foundation CAREER Award under Grant CCF-13-18980 and in part by the National Science Foundation under Grant CNS-13-21223. H. V. Poor was supported by the National Science Foundation under Grant CCF-10-16671. S. Shamai (Shitz) was supported in part by the Israel Science Foundation (ISF) and in part by the European Commission in the framework of the Network of Excellence in Wireless COMmunications NEWCOM#. This paper was presented in part at the 2009 IEEE Information Theory Workshop and the 2012 IEEE International Symposium on Information Theory.

Y. Liang is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: yliang06@syr.edu).

L. Lai is with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA 01609 USA (e-mail: llai@wpi.edu).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

Communicated by R. Yates, Associate Editor for Communications Networks.

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/TIT.2013.2293756

I. Introduction

As a complement to cryptographic techniques, physical layer security exploits physical channel randomness for secure transmissions without the use of secret keys. Such an approach was first introduced by Wyner in [1] via the wiretap channel model, and was further extended to a more general broadcast scenario by Csiszár and Körner in [2]. More recently, there has been a surge in interest in applying this approach to wireless networks (see the recent monographs [3] and [4] for overviews of recent work). As physical layer security exploits physical channel statistics to achieve secure communication, successful implementation of this approach depends crucially on the transmitter's knowledge about the channel state information (CSI). Previous studies have been mostly focused on scenarios in which the CSI is available to the transmitter although there are some exceptions, e.g., [5]–[8] and the references mentioned below. However, in wireless networks, such CSI may not be available to the transmitter possibly due to limited feedback resources. (The receivers, however, may be able to estimate the channel states.) More specifically for security concerns, eavesdroppers do not generally have incentive to feed their channel states back to transmitters. A reasonable approach to model channel uncertainty is the compound wiretap channel, e.g., [9]–[13], and arbitrary varying channel [14], [15], which guarantee secure message transmission under any channel state, in particular under the worst channel state. However, in order to guarantee the performance for the worst case which may occur only rarely, the channel resources are not used in an efficient manner if the actual channel state is better than the worst case. The focus of this paper is on the design of schemes that achieve as high a secrecy rate as the legitimate receiver's channel supports, and as the eavesdropper's channel permits. Since the channel state is unknown to the transmitter, the problem we address here is to design communication schemes that do not exploit channel state realizations but still adapt to the actual channel state that occurs in order to achieve as good a secrecy performance as possible.

Towards this end, a novel (layered) broadcast approach is especially appealing. This approach has been introduced for wireless systems without secrecy constraints in [16] based on superposition coding first introduced in [17] for broadcast channels. In this strategy, the transmitter splits the entire message into a number of components with each component being transmitted via one layer of input. These layers of inputs are then combined into one channel input using superposition encoding. The receiver decodes the layers one after another via successive interference cancelation. The realization of the channel state of the receiver determines up to which layer

the receiver can decode. More layers of messages can be decoded if the receiver's channel state is better. Hence, with a fixed coding scheme that does not require the transmitter to know the receiver's channel state, such an approach still allows the receiver to obtain as many layers of messages as its instantaneous channel state supports. We also note that the notion of the broadcast approach addressed in [16] has been conceptually extended and streamlined by introducing variable-to-fixed channel coding in [18].

In this paper, we generalize the broadcast approach in [16] to the fading wiretap channel, in which both the legitimate receiver's and eavesdropper's channels are time-varying block fading channels, i.e., the channel states are constant over one block and change ergodically across blocks. In particular, the CSI, i.e., the instantaneous channel realizations, are not known at the transmitter, and are known only at the corresponding receivers. A delay constraint is assumed so that messages must be transmitted within one block, i.e., coding across blocks as in [5] is not allowed. Our goal is to design transmission schemes such that the legitimate receiver decodes more information as its channel gets better, and out of information decoded at the legitimate receiver, more information is kept secure from the eavesdropper, as the eavesdropper's channel gets worse. We wish to characterize the rate of information that is decodable at the legitimate receiver and is secure from the eavesdropper. In particular, the performance measure of interest here is the delay-limited secrecy rate averaged over a long time range. This is different from the outage performance studied in [19]–[21], which focused on the delay-limited rate only over a short time range (say one coherence block).

We first develop two types of broadcast approaches respectively for two simpler fading channel scenarios in which only one of the channels is fading. These two approaches are then combined to study the general scenario in which both channels are fading. In the first scenario, only the legitimate receiver's channel is fading and the eavesdropper's channel is constant. For this scenario, the entire message is split into a number of layers with each layer employing stochastic encoding [1], [2] (also see [3, Section 2.3]) to achieve secrecy. These layers are then combined using superposition coding. Depending on its channel state, the legitimate receiver can decode messages up to a certain layer. Since the eavesdropper's channel is constant, all layers of messages are guaranteed to be kept secure from the eavesdropper via the stochastic encoding. Based on this approach, we derive the average secrecy rate over a large number of blocks for a given power allocation across layers of messages. We then employ the Euler equation derived in the calculus of variations to characterize the optimal power allocation to achieve an optimal average secrecy rate.

In the second scenario, only the eavesdropper's channel is fading and the legitimate receiver's channel is constant. In contrast to the first scenario, in which layers of messages are encoded into codewords in different subcodes, here all layers of messages are encoded into one codeword in an embedded fashion as in [22]. Each layer of messages corresponds to one index that identifies the codeword. In particular, lower layers of messages serve as randomization for protecting higher layers of messages from the eavesdropper. Depending on the eavesdropper's channel state, all messages down to certain layers are kept secure from the eavesdropper. We then derive the average secrecy rate over a large number of blocks. We further show that the secrecy rate achieved by this broadcast approach is the best secrecy rate that the instantaneous channel allows although the transmitter does not know the eavesdropper's CSI. The only sacrifice due to the lack of the CSI at the transmitter is that some lower layer messages may not be kept secure from the eavesdropper. This is in contrast to the first type broadcast approach developed for the case when the legitimate receiver has a fading channel, for which all messages transmitted over the channel are guaranteed to be kept secure from the eavesdropper, but the secrecy rate achieved may not be optimal.

For the third scenario, in which both channels to the legitimate receiver and the eavesdropper undergo fading, we combine the two types of broadcast approaches developed above. In particular, the entire message is split into layers identified by two-dimensional index pairs (say along horizonal and vertical index directions). For a given state of the legitimate receiver (i.e., a fixed horizonal index), all layers of messages are encoded via the vertical indices into one codeword in an embedded fashion via the broadcast approach developed for the second scenario, and codewords with different horizonal indices are then encoded together via the broadcast approach developed for the first scenario. Depending on its channel state, the legitimate receiver can decode messages up to a certain layer indexed by a horizonal index. Also depending on the eavesdropper's channel state, messages down to a certain layer indexed by a vertical index can be kept secure from the eavesdropper. Based on this scheme, we derive the average secrecy rate over a large number of blocks for a given power allocation across the layers of messages. We also employ the Euler equation to characterize necessary conditions for an optimal power allocation to achieve the optimal average secrecy rate. We further illustrate the structure of the optimal power allocation via a numerical example.

We note that from the three scenarios mentioned above, it is clear that the broadcast approach does not guarantee that all transmitted messages are kept secure from the eavesdropper for all eavesdropper states for the scenarios when the eavesdropper experiences a fading channel. The eavesdropper's actual channel state realization determines which layers of messages are secure, and the probability that such a state occurs determines the probability of achieving the corresponding secrecy rate. We hence introduce and study a notion of *probabilistic secrecy*, which characterizes the probability that certain layers of decoded messages are kept secure, i.e., the probability that the corresponding secrecy rate is achievable. Furthermore, the probabilistic secrecy also suggests that our broadcast approach protects different layers of messages unequally with higher layers of messages being more likely to be secure. Hence, for scenarios in which multiple messages with heterogeneous security demands need to be simultaneously transmitted over the channel, the messages with higher levels of security demands should be encoded into layers with larger indices so that these messages are less likely to be learned by the eavesdropper. We also note that

the probabilistic secrecy is different from the deterministic secrecy required for the classical wiretap channel [1], the fading wiretap channel (see, e.g., [5], [23] and [24]), and the compound wiretap channel (see, e.g., [9]–[13]), in which all messages decoded by the legitimate receiver are guaranteed to be secure (with probability one).

We then provide numerical examples to demonstrate the impact of the CSI at the transmitter and the channel fluctuations of the eavesdropper on the average secrecy rate. These numerical results suggest that the legitimate receiver's CSI affects the secrecy rates much more than the eavesdropper's CSI. Without the legitimate receiver's CSI, the transmitter has to spread its power to accommodate a number of possible state realizations, and such power spreading reduces the secrecy rate. However, the eavesdropper's CSI affects mostly the legitimate receiver's knowledge about which layers of messages are secure, but does not affect much the amount of information that is kept secure from the eavesdropper. Another important factor that affects the secrecy rate are the channel fluctuations (i.e., fading) of the eavesdropper, which create opportunities for achieving better secrecy rates.

We finally note that this study is different from the recent study in [25]. This study applies the conceptual idea of the original broadcast approach in [16] of transmitting layers of messages, but the actual coding scheme is different from that in [16] by incorporating stochastic coding either for each layer of messages or in an embedded fashion to guarantee secrecy for messages. Hence, secrecy is achieved solely via the broadcast approach, and no further feedback from the legitimate receiver is allowed to assist secrecy achievement. However, the study in [25] uses the original coding scheme in [16] for signal transmission, which does not guarantee secrecy, and secrecy of messages is instead achieved by allowing feedback from the legitimate user.

The organization of the paper is as follows. In Section II, we introduce our system model. In Sections III, IV, and V, we study the three scenarios in more detail. In Section VI, we provide numerical examples. Finally, in Section VII, we conclude the paper with some comments on future directions. We note that although the first two scenarios are special cases of the third scenario, they are presented separately for developing two types of broadcast approaches that are useful for the general scenario. Including these two scenarios also helps to understand better the combined approach for the third scenario.

II. System Model

In this paper, we study the fading wiretap channel (see Fig. 1), in which a transmitter sends a message to a legitimate receiver and wishes to keep this message secret from an eavesdropper. Both the legitimate receiver's and the eavesdropper's channels are corrupted not only by additive complex Gaussian noise, but also by multiplicative fading gain coefficients. The channel input-output relationship for one channel use is given by

$$Y = HX + W \quad \text{and} \quad Z = GX + V \tag{1}$$

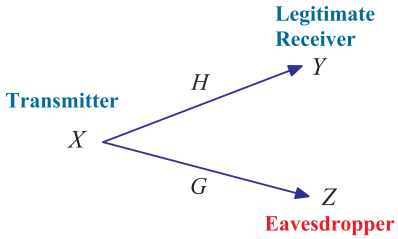

Fig. 1. An illustration of the fading wiretap channel.

where X is the input from the transmitter, Y and Z are outputs at the legitimate receiver and the eavesdropper respectively, H and G are fading channel gain coefficients, and the noise variables W and V are independent proper complex Gaussian random variables with zero means and unit variances. The noise variables are independent and identically distributed (i.i.d.) over channel uses. The fading gain H and G are assumed to experience block fading, i.e, they are constant within a block and change ergodically across blocks. The block length are assumed to be sufficiently long such that one codeword can be successfully transmitted if properly constructed.[1] The channel input is subject to an average power constraint P over each block, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[|X_i|^2\right] \le P \tag{2}$$

where i denotes the symbol time (i.e., channel use) index, and where n is the blocklength.

It is assumed that the transmitter does not know the instantaneous channel state information, and each receiver knows its own channel state. Each message is required to be transmitted within one block, i.e., the message is transmitted under a delay constraint. The legitimate receiver is required to decode the transmitted message with a small probability of error at the end of each block, and the message needs to be kept as secure as possible from the eavesdropper. The measure of security is based on the equivocation rate given by

$$\frac{1}{n}H(W|Z^n) \tag{3}$$

where Z^n denotes the received outputs at the eavesdropper over one block, and hence depends on the channel state realization of the eavesdropper during this block. The message W is kept secure from the eavesdropper during one block if there exists a positive ϵ_n that approaches zero as n goes to infinity such that

$$\frac{1}{n}H(W|Z^n) \ge \frac{1}{n}H(W) - \epsilon_n.$$

In this paper, it is not required that all messages transmitted over the channel be perfectly secure. However, our performance measure is the secrecy rate, which is the rate of the

[1] We note that here we implicitly assume that the block length is the same for both the legitimate receiver's channel and the eavesdropper's channel. Such an assumption is not strictly needed. In fact, as long as the common divisor of the block lengths of the legitimate receiver and the eavesdropper is sufficiently large to accommodate one codeword, all the results developed in the paper hold. In this case, each codeword length is the same as the common divisor, and each actual block may contain multiple codewords.

messages that are kept secure from the eavesdropper. If all messages transmitted over one block are viewed as a single message, then our performance measure can be interpreted as the level of secrecy achieved for this message, i.e., the equivocation of the message. Furthermore, we are interested in characterizing the secrecy rate under the delay constraint, but averaged over a large number of blocks. Namely, if we let $R(H, G)$ to denote the achievable secrecy rate corresponding to the channel state pair (H, G), then the average secrecy rate is given by $E[R(H, G)]$.

We also introduce the notion of probabilistic secrecy, which characterizes the probability that a certain secrecy rate of decoded messages can be achieved during a block, i.e., decoded messages at a certain rate can be kept secure from the eavesdropper. Such a probabilistic manner arises because the eavesdropper's channel is random and unknown to the transmitter, and hence encoding at the transmitter may not guarantee all messages decoded by the legitimate receiver to be secure from the eavesdropper at any eavesdropper's state. The state of the eavesdropper determines which messages are kept secure, and the probability that such a state occurs determines the probability of achieving the corresponding secrecy rate.

III. Fading Channel to Legitimate Receiver

In this section, we study the case in which only the legitimate receiver experiences a block fading channel, i.e., H is a constant over one block and changes independently to another realization from one block to another. The channel to the eavesdropper is assumed to be a constant, i.e., G is fixed and is hence known to every node. The transmitter does not know the instantaneous channel state to the legitimate user, but the legitimate receiver is assumed to know the channel state. In the sequel, we first develop a layered broadcast approach for the case with a discrete fading state and then generalize the approach to the case with a continuous fading state.

A. Discrete Legitimate Channel States

We first consider the case in which the legitimate receiver has a finite number of channel states, i.e., H may take L values, say $H_1, \ldots, H_L$ with $|H_1| \leq |H_2| \leq \cdots \leq |H_L|$. For this channel, we propose a (layered) broadcast approach, which generalizes the approach introduced in [16] for the broadcast channel without secrecy constraints. More specifically, the entire message is split into L parts, i.e., L layers of messages denoted by W_l for $l = 1, \ldots, L$.

Definition 1: A secrecy rate tuple $(R_1, \ldots, R_L)$ is *achievable* if there exists a coding scheme that encodes the messages $W_1, \ldots, W_L$ at the rate tuple $(R_1, \ldots, R_L)$ such that for $l = 1, \ldots, L$, the legitimate receiver decodes W_l with a small probability of error if its channel realization is H_l, and all messages $W_1, \ldots, W_L$ are kept secure from the eavesdropper.

We note that the above definition does not involve the probabilities assigned to the legitimate receiver's channel states, and involves only specification of the messages that should be decoded and kept secure at each state. In this sense, the model here can be viewed as a system with multiple legitimate receivers with each satisfying a decoding-secrecy requirement specified in the above definition.

The following theorem characterizes secrecy rate tuples that can be achieved by a broadcast approach.

Theorem 1: For the fading wiretap channel with the legitimate receiver having one of the L fading states $H_1, \ldots, H_L$, where $|G| < |H_1| \leq |H_2| \leq \cdots \leq |H_L|$, and with the eavesdropper having a fixed channel state G, the following secrecy rate tuples $(R_1, \ldots, R_L)$ are achievable:

$$R_l = \log\left(1 + \frac{|H_l|^2 P_l}{1 + |H_l|^2 \sum_{k=l+1}^{L} P_k}\right) - \log\left(1 + \frac{|G|^2 P_l}{1 + |G|^2 \sum_{k=l+1}^{L} P_k}\right), \quad l = 1, \ldots, L \quad (4)$$

where P_l denotes the transmission power assigned for transmitting W_l and satisfies the power constraint $\sum_{l=1}^{L} P_l \leq P$.

Remark 1: For the case when the legitimate receiver also has a fixed fading state (i.e., the channel now is the Gaussian wiretap channel), the total secrecy rate of all messages following from Theorem 1 is optimal. Hence, the broadcast approach that we develop (see the proof of Theorem 1) is optimal for the Gaussian wiretap channel.

We note that in this degraded setting, since messages decoded by a receiver with a worse channel state should also be decoded by the receiver with a better channel state, the legitimate receiver at the channel state H_l can decode $W_1, \ldots, W_l$ if $(R_1, \ldots, R_L)$ is achievable. Hence, the total rate of the messages that the legitimate receiver at the state H_l can decode is given by

$$\begin{aligned}\sum_{j=1}^{l} R_j &= \sum_{j=1}^{l}\left[\log\left(1 + \frac{|H_j|^2 P_j}{1 + |H_j|^2 \sum_{k=j+1}^{L} P_k}\right) - \log\left(1 + \frac{|G|^2 P_j}{1 + |G|^2 \sum_{k=j+1}^{L} P_k}\right)\right] \\ &= \left[\sum_{j=1}^{l} \log\left(1 + \frac{|H_j|^2 P_j}{1 + |H_j|^2 \sum_{k=j+1}^{L} P_k}\right)\right] - \log\left(1 + \frac{|G|^2 \sum_{k=1}^{l} P_k}{1 + |G|^2 \sum_{k=l+1}^{L} P_k}\right). \quad (5)\end{aligned}$$

We also note that the second term in (4) seems to suggest that the eavesdropper may also decode the current layer by removing interference caused by the layers that it has decoded. However, this interpretation is misleading. We will show below that the eavesdropper does not obtain any information about the messages $W_1, \ldots, W_L$, i.e., perfect secrecy is achieved for all layers of messages.

We next provide the proof of the above theorem, which describes the layered broadcast approach in more detail.

Proof of Theorem 1: We consider a codebook that contains L subcodebooks corresponding to L layers (see Fig. 2). For each layer l, the subcodebook $\mathcal{C}_l$ contains $2^{n\tilde{R}_l}$ codewords $x_l^n(w_l)$ indexed by $w_l = 1, \ldots, 2^{n\tilde{R}_l}$, where

$$\tilde{R}_l = \log\left(1 + \frac{|H_l|^2 P_l}{1 + |H_l|^2 \sum_{k=l+1}^{L} P_k}\right), \quad (6)$$

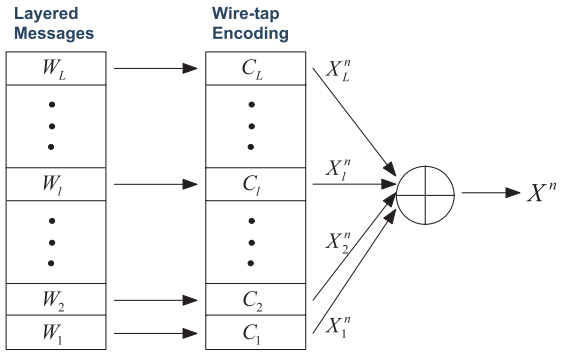

Fig. 2. A codebook for the broadcast approach.

$\frac{1}{n}\sum_{i=1}^{n} x_{li}^2(w_l) \le P_l$ for $w_l = 1, \ldots, 2^{n\tilde{R}_l}$, and $\sum_{l=1}^{L} P_l \le P$. These codewords are divided into 2^{nR_l} bins, where

$$R_l = \log\left(1 + \frac{|H_l|^2 P_l}{1 + |H_l|^2 \sum_{k=l+1}^{L} P_k}\right) - \log\left(1 + \frac{|G|^2 P_l}{1 + |G|^2 \sum_{k=l+1}^{L} P_k}\right). \quad (7)$$

The encoding scheme is described as follows. In order to transmit a message tuple $(w_1, \ldots, w_L)$, for each l, the message w_l is mapped into the w_lth bin in the subcodebook $\mathcal{C}_l$, and one codeword x_l^n in the bin is randomly chosen with a uniform distribution over the entire bin. The final input transmitted over the channel is given by

$$x^n = \sum_{l=1}^{L} x_l^n.$$

Following steps similar to those in [3, Section 2.3], it can be shown that there exists a codebook as described above such that if this codebook and the encoding scheme as described above are applied to the Gaussian wiretap channel with the channels to the legitimate receiver and the eavesdropper respectively being at the state H_l and G, the legitimate receiver can successfully decode $W_1, \ldots, W_l$ with a small probability of error. Furthermore, the eavesdropper can successfully decode X_l^n with a small probability of error if it knows $W_1, \ldots, W_l$ for all $l = 1, \ldots, L$. Based on this property, it can be shown that

$$\frac{1}{n} H(W_1, \ldots, W_L | Z^n) \ge \frac{1}{n} H(W_1, \ldots, W_L) - \epsilon_n \quad (8)$$

where ϵ_n goes to zero as n approaches infinity. The above equation implies that perfect secrecy is achieved asymptotically as n approaches infinity. The details can be specialized from the proof for Theorem 4 in Section V-A.

B. Continuous Legitimate Channel State

In this subsection, we generalize our result for the discrete fading channel to the continuous fading channel. We still assume that only the legitimate receiver's channel is block fading and the eavesdropper's channel is fixed. Hence, the legitimate receiver's channel gain H can take continuous values. For each channel state $H = h$, we let $s = |h|^2$, and use s as an index for the layer of the message that is intended for the legitimate receiver at the state h to decode. For each layer s, we assume that the transmitter allocates power $\rho(s)ds$. We use $\Sigma(s)$ to denote the total power allocated to the layers corresponding to better channel states, i.e., the states $\hat{s}$ such that $\hat{s} > s$. Hence,

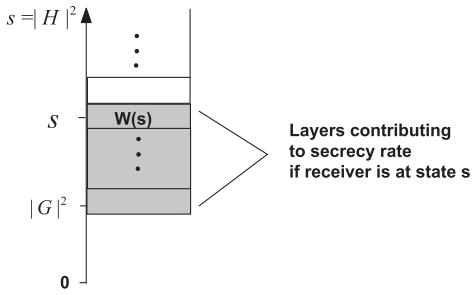

Fig. 3. An illustration of layers of messages.

$$\Sigma(s) = \int_s^{\infty} \rho(x)dx, \quad (9)$$

and

$$\rho(s) = -\Sigma'(s). \quad (10)$$

We note that in this paper we consider only the case in which $\Sigma(s)$ is differentiable.

The following result on the average secrecy rate follows directly by applying Theorem 1.

Corollary 1: For the fading wiretap channel with the legitimate receiver having a block fading channel with continuous states and the eavesdropper having a fixed channel state G, the average secrecy rate under the delay constraint achieved via a broadcast approach is given by

$$R = \log e \max_{\Sigma(x)} \int_{|G|^2}^{\infty} (1 - F(x)) \left[\frac{-x\Sigma'(x)}{1 + x\Sigma(x)} + \frac{|G|^2 \Sigma'(x)}{1 + |G|^2 \Sigma(x)}\right] dx \quad (11)$$

where $f(\cdot)$ is the probability density function of the fading state s, and $F(\cdot)$ is the cumulative distribution function of s.

Proof of Corollary 1: Following from (4), we obtain the following secrecy rate corresponding to layer $s = |h|^2$. It is clear that if $s \le |G|^2$, then $dR = 0$. If $s > |G|^2$, then the secrecy rate is given by

$$dR = \log\left(1 + \frac{s\rho(s)ds}{1 + s\Sigma(s)}\right) - \log\left(1 + \frac{|G|^2 \rho(s)ds}{1 + |G|^2 \Sigma(s)}\right) \approx \log e \left[\frac{s\rho(s)ds}{1 + s\Sigma(s)} - \frac{|G|^2 \rho(s)ds}{1 + |G|^2 \Sigma(s)}\right] \quad (12)$$

where the second approximate equation follows because ds approaches zero. This approximation can be made rigorous because for any given $\delta > 0$, there must exist an interval $[-a, a]$ centered around $x = 0$ such that

$$x \le \ln(1 + x) \le (1 + \delta)x, \quad \forall x \in [-a, a].$$

It can be seen that if the legitimate receiver's channel is at state s, then it can decode messages corresponding to all layers x if $x \le s$ (see Fig. 3). Hence, the total secrecy rate

achievable if the legitimate receiver's channel is at state s is given by

$$R(s) = \log e \int_{|G|^2}^{s} \frac{x\rho(x)dx}{1+x\Sigma(x)} - \frac{|G|^2\rho(x)dx}{1+|G|^2\Sigma(x)}. \quad (13)$$

Averaging the above rate over all fading state realizations of the legitimate receiver's channel, we obtain

$$\begin{aligned} R &= \int_{|G|^2}^{\infty} f(s)R(s)ds \\ &= \log e \int_{|G|^2}^{\infty} (1-F(x)) \left[\frac{x\rho(x)}{1+x\Sigma(x)} - \frac{|G|^2\rho(x)}{1+|G|^2\Sigma(x)} \right] dx \end{aligned}$$

where $f(\cdot)$ is the probability density function of the fading state s, and $F(\cdot)$ is the cumulative distribution function of s. The above average rate can be further improved by optimizing over power allocations $\rho(\cdot)$, or equivalently $\Sigma(\cdot)$. We can also use (10) to replace $\rho(x)$ in the final equation for the average rate, which completes the proof. ∎

To obtain the optimal average rate R given in (11) and the corresponding optimal power allocation function $\Sigma(\cdot)$, we study the following optimization problem. In particular, we focus on continuous power allocation functions, i.e., $\Sigma(\cdot)$ is a continuous function defined over $[0, \infty)$.

$$\begin{aligned} &\max_{\Sigma(x)} \int_{|G|^2}^{\infty} S(x, \Sigma(x), \Sigma'(x))dx \\ &\text{subject to } \ 0 \le \Sigma(x) \le P, \ \Sigma'(x) \le 0, \quad \text{for } x \ge 0 \end{aligned} \quad (14)$$

where

$$\begin{aligned} &S(x, \Sigma(x), \Sigma'(x)) \\ &\quad = (1-F(x)) \left[\frac{-x\Sigma'(x)}{1+x\Sigma(x)} + \frac{|G|^2\Sigma'(x)}{1+|G|^2\Sigma(x)} \right]. \end{aligned} \quad (15)$$

The following theorem characterizes the structure of the optimal power allocation function.

Theorem 2: Let

$$\eta(x) = \frac{1-F(x)-(x-|G|^2)f(x)}{xf(x)(x-|G|^2)-(1-F(x))|G|^2}. \quad (16)$$

An optimal solution to (14), if one exists, has the following structure. There exist $0 \le x_1 < y_1 < x_2 < y_2 < \cdots < x_m < y_m$, such that $\eta(x)$ is strictly decreasing over $[x_i, y_i]$ for $i = 1, \ldots, n$, $\eta(x_1) = P$, $\eta(y_n) = \eta(x_0) = 0$, $\eta(y_i) = \eta(x_{i+1})$ for $i = 1, \ldots, m-1$, and

$$\Sigma^*(x) = \begin{cases} P & 0 \le x \le x_1; \\ \eta(x) & x_i \le x \le y_i, \\ & \text{for } \ i = 1, \ldots, m; \\ \eta(y_i) = \eta(x_{i+1}), & y_i < x < x_{i+1}, \\ & \text{for } \ i = 1, \ldots, m-1; \\ 0 & y_m \le x. \end{cases} \quad (17)$$

Remark 2: The functions $\Sigma(x)$ that satisfy the conditions given in Theorem 2 may not be unique.

Remark 3: In Theorem 2, y_m may be infinity.

Proof of Theorem 2: Based on the property that $0 \le \Sigma(x) \le P$, and the fact that $\Sigma(x)$ is continuous and nonincreasing, it is clear that any optimal $\Sigma^*(x)$ if one exists must have the following form:

$$\Sigma^*(x) = \begin{cases} P & 0 \le x \le x_1; \\ \text{a strictly decreasing function} & x_i \le x \le y_i, \\ & \text{for } \ i = 1, \ldots, m; \\ \text{a constant,} & y_i < x < x_{i+1}, \\ & \text{for } \ i = 1, \ldots, m-1; \\ 0 & y_m \le x. \end{cases} \quad (18)$$

where $0 \le x_1 < y_1 < x_2 < y_2 < \cdots < x_m < y_m$.

The optimization problem (14) is a problem of the constrained calculus of variation. We thus apply the technique in [26] to provide a necessary condition that $\Sigma^*(x)$ satisfies. Over the intervals $(x_1, y_1]$, $[x_i, y_i]$ for $i = 2, \ldots, m-1$, and $[x_m, y_m)$, since $\Sigma^*(x)$ is strictly decreasing, it does not satisfy the inequality constraints in (14) with equality, i.e., it is not on the boundary of the constraint set. Due to the complementary slackness conditions [26], the following Euler equation must be satisfied:

$$S_\Sigma - \frac{d}{dx} S_{\Sigma'} = 0, \quad (19)$$

where

$$\begin{aligned} S_\Sigma &= \frac{\partial S(x, \Sigma(x), \Sigma'(x))}{\partial \Sigma}, \\ S_{\Sigma'} &= \frac{\partial S(x, \Sigma(x), \Sigma'(x))}{\partial \Sigma'}. \end{aligned}$$

For the function $S(x, \Sigma(x), \Sigma'(x))$ given in (15), we obtain

$$\begin{aligned} S_\Sigma &= (1-F(x)) \left[\frac{x^2\Sigma'(x)}{(1+x\Sigma(x))^2} - \frac{|G|^4\Sigma'(x)}{(1+|G|^2\Sigma(x))^2} \right] \\ S_{\Sigma'} &= (1-F(x)) \left[\frac{-x}{1+x\Sigma(x)} + \frac{|G|^2}{1+|G|^2\Sigma(x)} \right] \\ \frac{d}{dx} S_{\Sigma'} &= \frac{xf(x)}{1+x\Sigma(x)} - \frac{f(x)|G|^2}{1+|G|^2\Sigma(x)} \\ &\quad + (1-F(x)) \left[\frac{x^2\Sigma'(x)-1}{(1+x\Sigma(x))^2} - \frac{|G|^4\Sigma'(x)}{(1+|G|^2\Sigma(x))^2} \right]. \end{aligned}$$

We substitute the above equations into the Euler equation and obtain

$$\Sigma^*(x) = \eta(x) = \frac{1-F(x)-(x-|G|^2)f(x)}{xf(x)(x-|G|^2)-(1-F(x))|G|^2}, \quad (20)$$

over the intervals $(x_1, y_1]$, $[x_i, y_i]$ for $i = 2, \ldots, m-1$, and $[x_m, y_m)$. This also implies that $\eta(x)$ must be strictly decreasing over these intervals. Due to the continuity of $\Sigma^*(x)$, the values of $\Sigma^*(x)$ over (y_i, x_{i+1}) are given by $\eta(y_i) = \eta(x_{i+1})$ for $i = 1, \ldots, m-1$, which also implies that $\eta(x)$ must satisfy $\eta(y_i) = \eta(x_{i+1})$ for $i = 1, \ldots, m-1$. Also due to the continuity of $\Sigma^*(x)$, $\eta(x_1) = P$, and $\eta(y_m) = 0$.

We note that in order to obtain $\Sigma^*(x)$ and the intervals $[x_i, y_i]$ for $i = 1, \ldots, m$, one can start from setting $\Sigma(0) = P$, and gradually increase x to identify the intervals during which the right-hand-side (RHS) of equation (20) strictly decreases until the RHS of (20) drops to zero. These intervals are then $[x_i, y_i]$ for $i = 1, \ldots, m$. The function $\Sigma^*(x)$ is clearly obtained during this process.

Example 1: In this example, we consider the case when the channel to the legitimate receiver experiences Rayleigh fading. Hence, $s = |H|^2$ is exponentially distributed, and

$$f(x) = \frac{1}{\sigma_1} e^{-\frac{x}{\sigma_1}} \quad \text{and} \quad F(x) = 1 - e^{-\frac{x}{\sigma_1}}, \quad x \geq 0 \tag{21}$$

where σ_1 is the parameter of the exponential distribution.

Substituting (21) into (16), we obtain

$$\eta(x) = \frac{\sigma_1 - x + |G|^2}{x(x - |G|^2) - \sigma_1 |G|^2}. \tag{22}$$

By solving $\eta(x_1) = P$ and $\eta(y_1) = 0$, we obtain

$$\begin{aligned} &y_1 = \sigma_1 + |G|^2, \quad \text{and} \\ &x_1 = \tfrac{(P|G|^2-1)+\sqrt{(P|G|^2-1)^2+4P(P\sigma_1|G|^2+|G|^2+\sigma_1)}}{2P}. \end{aligned} \tag{23}$$

It is easy to check that $|G|^2 < x_1 < y_1$. We also note that $\eta(x)$ is strictly decreasing over the range $[x_1, y_1]$, because the numerator of $\eta(x)$ is decreasing, and the denominator of $\eta(x)$ is increasing over the interval $[x_1, y_1]$. Since x_1 and y_1 are both unique solutions to $\eta(x_1) = P$ and $\eta(y_1) = 0$, respectively, and $\eta(x)$ is strictly decreasing over $[x_1, y_1]$, the optimal $\Sigma^*(x)$ is thus given by

$$\Sigma^*(x) = \begin{cases} P & 0 \leq x \leq x_1; \\ \eta(x) & x_1 \leq x \leq y_1; \\ 0 & y_1 \leq x. \end{cases} \tag{24}$$

Since the above $\Sigma^*(x)$ is the unique function that satisfies the conditions given in Theorem 2, it is the only possible optimal solution for the power allocation function.

By taking the derivative of $\Sigma^*(x)$, we obtain

$$\begin{aligned} \rho^*(x) &= -\Sigma^{*\prime}(x) \\ &= \frac{-x^2 + 2\sigma_1 x - 2\sigma_1|G|^2 + 2|G|^2 x - |G|^4}{(x(x - |G|^2) - \sigma_1|G|^2)^2}. \end{aligned} \tag{25}$$

By substituting $\Sigma^*(x)$ and $\Sigma^{*\prime}(x)$ to (11), we can obtain the optimal average secrecy rate via a broadcast approach for the Rayleigh fading channel. Numerical results are provided in Section VI.

IV. Fading Channel to Eavesdropper

In this section, we study the case in which only the eavesdropper experiences a block fading channel, i.e., G is a constant over each block, and changes independently from one block to another. The legitimate receiver's channel gain H is assumed to be a constant, and is thus known to all nodes. As for the case in which only the legitimate receiver's channel is fading, it is assumed that the transmitter does not know the instantaneous channel state to the eavesdropper, but the eavesdropper knows its own channel state. In the rest of this section, we first study the case with a discrete fading state, and then generalize our result to the case with a continuous fading state.

A. Discrete Eavesdropping Channel States

We first consider the case in which the eavesdropper has a finite number of channel states, i.e., G may take L values, say $G_1, \ldots, G_L$ with $|G_1|^2 < |G_2|^2 < \cdots < |G_L|^2 < |H|^2$. We note that if the eavesdropper's best state is better than the legitimate receiver's state $|H|^2$, then no message can be made secure for the eavesdropper's states that are better than $|H|^2$. However, our following results are still applicable by considering only the eavesdropper's states that are worse than $|H|^2$.

For this case, we develop a second type of broadcast approach that is different from the one developed in Section III. To proceed, we start by splitting the entire message into L layers of messages $W_1, W_2, \ldots, W_L$.

Definition 2: A secrecy rate tuple $(R_1, \ldots, R_L)$ is achievable if there exists a coding scheme that encodes $W_1, \ldots, W_L$ at the rate tuple $(R_1, \ldots, R_L)$ such that the legitimate receiver can decode all messages with a small probability of error, and message W_l is kept secure from the eavesdropper if the eavesdropper's channel state is G_l for $l = 1, \ldots, L$.

The following theorem characterizes achievable secrecy rate tuples via a broadcast approach.

Theorem 3: Consider the fading wiretap channel with the legitimate receiver having a fixed channel state H and the eavesdropper possibly having one of L fading states $G_1, \ldots, G_L$ with $|G_1|^2 < |G_2|^2 < \cdots < |G_L|^2 < |H|^2$. The following secrecy rate tuples $(R_1, \ldots, R_L)$ are achievable:

$$R_l = \log\left(1 + |G_{l+1}|^2 P\right) - \log\left(1 + |G_l|^2 P\right), \quad \textit{for} \quad l = 1, \ldots, L-1; \tag{26}$$

$$R_L = \log\left(1 + |H|^2 P\right) - \log\left(1 + |G_L|^2 P\right). \tag{27}$$

Since the messages that are secure from the eavesdropper with the state G_j are also secure from the eavesdropper with the state G_l if $|G_j| > |G_l|$, all $W_l, \ldots, W_L$ are secure from the eavesdropper at the state G_l if $(R_1, \ldots, R_L)$ is achievable. Hence, the total rate of the messages that are secure from the eavesdropper at the channel state G_l is given by

$$R_l + R_{l+1} + \cdots + R_L = \log\left(1 + |H|^2 P\right) - \log\left(1 + |G_l|^2 P\right). \tag{28}$$

We note that the secrecy rate in (28) is equal to the secrecy capacity of the channel with the state pair (H, G_l). Hence, the second type broadcast approach (described in the proof for Theorem 3) achieves the best secrecy rate that the instantaneous channel allows although the transmitter does not know the eavesdropper's CSI. The only sacrifice due to the lack of the CSI at the transmitter is that some lower layer messages may not be kept secure from the eavesdropper. This is in contrast to the first type of broadcast approach developed for the case when the legitimate receiver has a fading channel, for which all messages transmitted over the channel are guaranteed to be kept secure from the eavesdropper, but the secrecy rate achieved may not be optimal.

We note that although the legitimate receiver does not know the eavesdropper's channel state, the broadcast approach still prevents the eavesdropper from knowing certain layers of

information with these layers determined by the eavesdropper's channel state. However, without knowing the eavesdropper's channel state, the legitimate receiver knows only the probability that certain layers of messages are kept secure, which is referred to as probabilistic secrecy and is studied in the following subsection.

We next provide the details of the proof for Theorem 3, in which the second type broadcast approach is developed in detail.

Proof of Theorem 3: In contrast to the broadcast approach developed for proving Theorem 1 that employs a subcodebook for each layer of messages, the broadcast approach here generalizes the embedding code structure proposed in [22] that uses only one codebook. Each codeword is indexed by a random index and all layers of messages. Depending on the channel state of the eavesdropper, up to certain layers of messages jointly with the random index serve as randomness to protect the remaining higher-layer messages. In this way, these higher-layer messages can be viewed as a vector bin number, and the lower-layer messages and the random index can be viewed as the index (vector) of the codeword within each bin. In particular, the entire code can be viewed in an embedded fashion in that each layer of messages serves as a bin number with the corresponding bins being embedded into larger bins indexed by messages one layer higher. We describe this codebook in more detail as follows.

We construct a codebook that contains $2^{n\log(1+|H|^2P)}$ codewords x^n, which are indexed by $(q, w_1, \ldots, w_{L-1}, w_L)$ with

$$\begin{aligned} q &= 1, 2, \ldots, 2^{n\log(1+|G_1|^2P)}, \\ w_1 &= 1, 2, \ldots, 2^{n\left[\log(1+|G_2|^2P)-\log(1+|G_1|^2P)\right]}, \\ w_2 &= 1, 2, \ldots, 2^{n\left[\log(1+|G_3|^2P)-\log(1+|G_2|^2P)\right]}, \\ &\vdots \\ w_{L-1} &= 1, 2, \ldots, 2^{n\left[\log(1+|G_L|^2P)-\log(1+|G_{L-1}|^2P)\right]}, \\ w_L &= 1, 2, \ldots, 2^{n\left[\log(1+|H|^2P)-\log(1+|G_L|^2P)\right]}. \end{aligned} \tag{29}$$

Using this codebook, to transmit a message tuple $(w_1, w_2, \ldots, w_L)$, the encoder randomly selects an index q with the uniform distribution and transmits $x^n(q, w_1, w_2, \ldots, w_L)$. To connect this approach to the wiretap binning scheme, here, for an eavesdropper's channel state G_l, the codewords in the codebook can be viewed as being assigned to the bins indexed by $(w_l, \ldots, w_L)$.

Using the codebook structure specified in (29) and following steps similar to those in [3, Section 2.3], it can be shown that there exists a codebook with the above structure such that if this codebook and the above encoding scheme are applied, then the legitimate receiver can decode X^n, and hence $W_1, \ldots, W_L$, with a small probability of error. Furthermore, for $l = 1, \ldots, L$, if the eavesdropper's channel state is G_l, then the eavesdropper can decode the channel input X^n with a small probability of error if it knows $W_l, \ldots, W_L$. Based on this property, it can be shown that

$$\frac{1}{n}H(W_l, \ldots, W_L|Z_l^n) \geq \frac{1}{n}H(W_l, \ldots, W_L) - \delta_n \tag{30}$$

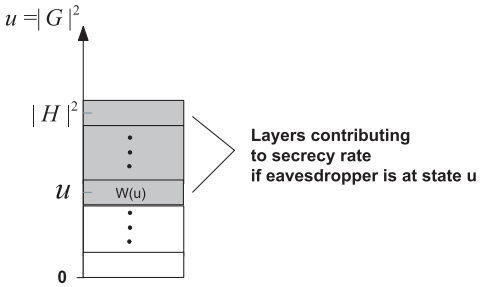

Fig. 4. An illustration of the layers of messages that are secure from the eavesdropper.

where δ_n goes to zero as n approaches infinity. The above equation implies that perfect secrecy is achieved asymptotically as n approaches infinity. The details can be specialized from the proof for Theorem 4 in Section V-A. ■

We note that the broadcast approach developed above is different from the original broadcast approach [16] and the one developed in Section III in that the power is not spread over layers of messages because one codebook that contains information about all layers of messages is employed. Furthermore, our scheme generalizes the embedding scheme in [22] (that treats the scenario with two eavesdropper's channel states) to the broadcast approach with multiple-layer embedding to accommodate multiple eavesdropper's channel states. This scheme is further extended for the case with an infinite number of layers in the following subsection. More importantly, our scheme with multiple-layer embedding does not result in reduction in the secrecy rate due to the single codebook design and no power spreading over layers.

B. Continuous Eavesdropping Channel State

We now generalize the result in the preceding subsection to the case in which the eavesdropper has a continuous channel state, i.e., the channel gain G takes continuous values. In this case, the message should be encoded correspondingly to a continuum of layers. For each state $G = g$, we let $u = |g|^2$, and use u as an index for the layer of the message that needs to be kept secure from the eavesdropper in the state g. The following result follows directly from Theorem 3.

Corollary 2: For the fading wiretap channel with the legitimate receiver having a fixed channel state H and the eavesdropper having a block fading channel, the average secrecy rate under the delay constraint achieved via the broadcast approach based on embedded coding is given by

$$R = Q(|H|^2)\log\left(1+|H|^2P\right) - \int_0^{|H|^2} q(u)\log(1+uP)du \tag{31}$$

where $q(\cdot)$ and $Q(\cdot)$ are the probability density function and cumulative distribution function of $|G|^2$, respectively.

We note that the above rate R can be easily computed numerically.

Proof of Corollary 2: Following from (28), the total secrecy rate when the eavesdropper's channel state is $u = |G|^2$ is given as follows (see Fig. 4):

$$R(u) = \begin{cases} \log\left(1+|H|^2P\right) - \log(1+uP), & \text{if } u < |H|^2 \\ 0, & \text{otherwise.} \end{cases} \tag{32}$$

Averaging the above rate over all eavesdropper's channel state realizations, we obtain

$$R = \int_0^{|H|^2} q(u)R(u)du$$
$$= \int_0^{|H|^2} q(u)\left[\log\left(1+|H|^2P\right) - \log(1+uP)\right]du \quad (33)$$
$$= Q(|H|^2)\log\left(1+|H|^2P\right) - \int_0^{|H|^2} q(u)\log(1+uP)du, \quad (34)$$

which concludes the proof. ∎

Based on the above proof, we now characterize the probabilistic secrecy for this scenario, i.e., the probability that a given secrecy rate R is achievable, denoted by $Pr(R)$. It is clear from (32) that if R is greater than the maximum rate $\log\left(1+|H|^2P\right)$ decodable at the legitimate receiver, $Pr(R)=0$. Otherwise, in (32), we set $R(u_R)=R$ to obtain

$$u_R = \frac{2^{\log(1+|H|^2P)-R}-1}{P}$$

which is the best eavesdropper's state such that messages with the rate R are still secure. Since these messages are also secure for any eavesdropper's state $u \le u_R$, $Pr(R)$ should be equal to $Pr\{u \le u_R\}$, which is $Q(u_R)$, i.e., the cumulative probability distribution of u evaluated at u_R. In summary, $Pr(R)$ is given by

$$Pr(R) = \begin{cases} Q(u_R) & \text{for} \quad R \le \log\left(1+|H|^2P\right); \\ 0 & \text{otherwise.} \end{cases}$$

V. Fading Channels to Both Legitimate Receiver and Eavesdropper

In this section, we study the general case, in which both the legitimate receiver and the eavesdropper experience block fading channels, i.e., H and G are constant over each block, and change independently to other realizations from one block to another. It is assumed that the transmitter knows neither the instantaneous channel state to the legitimate receiver nor the channel state to the eavesdropper, but the legitimate receiver and the eavesdropper know their corresponding channel states. As in the previous sections, we start with the case when the channel gains have finite numbers of states. We then study the case with continuous channel states.

A. Discrete Legitimate and Eavesdropping Channel States

We first consider the case in which both the legitimate receiver and the eavesdropper have finite numbers of channel states, i.e., H and G take one of $H_1,\ldots,H_L$ values and one of $G_1,\ldots,G_K$ values, respectively, where $|H_1|<\cdots<|H_L|$ and $|G_1|<\cdots<|G_K|$. For each $1\le l\le L$, we use K_l to denote the largest index of the state level of G that is below H_l, i.e., $K_l=\max_{|G_k|\le|H_l|} k$. We develop a broadcast approach that combines the two broadcast approaches developed in Sections III and IV. We first split the entire message into a number of components $W_{l[1,K_l]}$ for $1\le l\le L$, where $W_{l[1,K_l]}$ denotes $W_{l1},\ldots,W_{lK_l}$.

Definition 3: A secrecy rate tuple $\{R_{l[1,K_l]}\}_{l=1,\ldots,L}$ is achievable if there exists a coding scheme that encodes the messages $W_{l[1,K_l]}$ at the rates $R_{l[1,K_l]}$ for $1\le l\le L$ such that if the legitimate receiver's channel is at H_l and the eavesdropper's channel is at G_k for $1\le l\le L$ and $1\le k\le K_l$, then the legitimate receiver decodes the message W_{lk} with a vanishing probability of error and the eavesdropper is kept ignorant of the message W_{lk}.

We note that in the above definition, we consider only rate components with state indices (l,k) such that $1\le l\le L$ and $1\le k\le K_l$. This implies that for each fixed k, we consider only those l such that $|G_k|\le|H_l|$, and equivalently for each given l, we consider only those $1\le k\le K_l$. Consequently, Definition 3 specializes to Definitions 1 and 2 for the corresponding scenarios, respectively. For example, for scenario 1, in which the eavesdropper is at a fixed state k, we consider those states l of the legitimate receiver such that $|G_k|\le|H_l|$. Then, the requirement in Definition 3 becomes that the legitimate receiver decodes the messages W_{lk} if its state is at H_l, and the eavesdropper is kept ignorant of the messages W_{lk} for all l under consideration for a fixed k (i.e., those l such that $|G_k|\le|H_l|$). This precisely reduces to Definition 1. The same is true for scenario 2.

The following theorem characterizes achievable secrecy rate tuples via a broadcast approach.

Theorem 4: For the fading wiretap channel with the legitimate receiver having one of L fading states $H_1,\ldots,H_L$ with $|H_1|<\cdots<|H_L|$ and the eavesdropper having one of K fading states $G_1,\ldots,G_K$ with $|G_1|<\cdots<|G_K|$, the following secrecy rate tuples $(R_{1,[1,K_1]},\ldots,R_{L[1,K_L]})$ are achievable:

$$R_{lk} = \begin{cases} \log\left(1+\frac{|G_{k+1}|^2P_l}{1+|G_{k+1}|^2\sum_{j=l+1}^{L}P_j}\right) \\ \quad -\log\left(1+\frac{|G_k|^2P_l}{1+|G_k|^2\sum_{j=l+1}^{L}P_j}\right), \\ \qquad \textit{for} \quad 1\le l\le L, \quad 1\le k\le K_l-1 \\ \log\left(1+\frac{|H_l|^2P_l}{1+|H_l|^2\sum_{j=l+1}^{L}P_j}\right) \\ \quad -\log\left(1+\frac{|G_{K_l}|^2P_l}{1+|G_{K_l}|^2\sum_{j=l+1}^{L}P_j}\right), \\ \qquad \textit{for} \quad 1\le l\le L, \quad k=K_l \end{cases} \quad (35)$$

where P_l denotes the transmission power assigned to state l and satisfies the power constraint $\sum_{l=1}^{L}P_l\le P$.

We note that since the messages that are decodable by the legitimate receiver at any state H_j can also be decoded by the legitimate receiver at the state H_l if $|H_j|<|H_l|$, the legitimate receiver at the state H_l can decode all messages $W_{1[1,K_1]},\ldots,W_{l,[1,K_l]}$ for $l=1,\ldots,L$. And since the messages that are secure from the eavesdropper with any state G_j are also secure from the eavesdropper with the state G_k if $|G_j|>|G_k|$, all $W_{1[k,K_1]},\ldots,W_{L[k,K_L]}$ are secure from the eavesdropper at the state G_k.

We also note that similarly to the case in which only the channel to the eavesdropper is fading, employment of the broadcast approach does not require that the legitimate receiver know the channel state to the eavesdropper. However, without knowing the eavesdropper's channel state, the legitimate

receiver understands only the probability that certain layers of messages are kept secure, which is studied in the following subsection as probabilistic secrecy.

Proof of Theorem 4: The basic idea is to combine the two types of broadcast approaches developed in Sections III and IV. The details are as follows.

We consider a codebook that contains L subcodebooks corresponding to L layers of the legitimate receiver's channel. For each layer l, the subcodebook $\mathcal{C}_l$ contains $2^{n\log\left(1+|H_l|^2P_l/(1+|H_l|^2\sum_{j=l+1}^{L}P_j)\right)}$ codewords x_l^n indexed by $(q_l, w_{l1}, w_{l2}, \ldots, w_{lK_l})$, where

$$\begin{aligned}
q_l &= 1, 2, \ldots, 2^{n\log\left(1+\frac{|G_1|^2P_l}{1+|G_1|^2\sum_{j=l+1}^{L}P_j}\right)},\\
w_{l1} &= 1, 2, \ldots, 2^{n\left[\log\left(1+\frac{|G_2|^2P_l}{1+|G_2|^2\sum_{j=l+1}^{L}P_j}\right)-\log\left(1+\frac{|G_1|^2P_l}{1+|G_1|^2\sum_{j=l+1}^{L}P_j}\right)\right]},\\
w_{l2} &= 1, 2, \ldots, 2^{n\left[\log\left(1+\frac{|G_3|^2P_l}{1+|G_3|^2\sum_{j=l+1}^{L}P_j}\right)-\log\left(1+\frac{|G_2|^2P_l}{1+|G_2|^2\sum_{j=l+1}^{L}P_j}\right)\right]},\\
&\vdots\\
w_{l(K_l-1)} &= 1, 2, \ldots, 2^{n\left[\log\left(1+\frac{|G_{K_l}|^2P_l}{1+|G_{K_l}|^2\sum_{j=l+1}^{L}P_j}\right)-\log\left(1+\frac{|G_{K_l-1}|^2P_l}{1+|G_{K_l-1}|^2\sum_{j=l+1}^{L}P_j}\right)\right]},\\
w_{lK_l} &= 1, 2, \ldots, 2^{n\left[\log\left(1+\frac{|H_l|^2P_l}{1+|H_l|^2\sum_{j=l+1}^{L}P_j}\right)-\log\left(1+\frac{|G_{K_l}|^2P_l}{1+|G_{K_l}|^2\sum_{j=l+1}^{L}P_j}\right)\right]}. \qquad (36)
\end{aligned}$$

The encoding scheme is described as follows. To transmit a set of messages $w_{1[1,K_1]}, \ldots, w_{L,[L,K_L]}$, for each $l = 1, \ldots, L$, the transmitter randomly and uniformly selects q_l, and q_l together with $w_{l[1,k_l]}$ determines a codeword $x_l^n(q_l, w_{l1}, \ldots, w_{lK_l})$. The input transmitted over the channel is then given by

$$x^n = \sum_{l=1}^{L} x_l^n(q_l, w_{l1}, \ldots, w_{lK_l}).$$

Following steps similar to those in [3, Section 2.3], it can be shown that there exists a codebook as described above such that if the legitimate receiver has the channel state H_l, then it can decode $X_1^n, \ldots, X_l^n$, and hence the messages $W_{1[1,K_1]}, \ldots, W_{l[1,K_l]}$, with a small probability of error, and if the eavesdropper's channel is at G_k, then the eavesdropper can successfully decode X_l^n with a small probability of error if it knows $W_{l[k,K_l]}$ and $X_1^n, \ldots, X_{l-1}^n$, for $l = 1, \ldots, L$. More formally, this property implies that there exists a positive δ_n which approaches zero as n goes to infinity such that for $k = 1, \ldots, K$,

$$\begin{aligned}
&H(X_1^n|Z_k^n, W_{1[k,K_1]}) \le n\delta_n\\
&H(X_2^n|Z_k^n, W_{2[k,K_2]}, X_1^n) \le n\delta_n\\
&\qquad\vdots\\
&H(X_L^n|Z_k^n, W_{L[k,K_L]}, X_1^n, \ldots, X_{L-1}^n) \le n\delta_n \qquad (37)
\end{aligned}$$

where Z_k^n denotes the channel output received by the eavesdropper if its channel state is G_k.

From the codebook construction, it is clear that if the legitimate receiver has a channel realization H_l, it can decode $X_1^n, \ldots, X_l^n$, and hence the messages $W_{1[1,K_1]}, \ldots, W_{l[1,K_l]}$. It is then sufficient to show that if the eavesdropper is in the state G_k, the messages $W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}$ are kept secure from the eavesdropper. Towards this end, we compute the following equivocation rate:

$$\begin{aligned}
&H(W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}|Z_k^n)\\
&= H(W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}, Z_k^n) - H(Z_k^n)\\
&= H(W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}, Z_k^n, X_1^n, \ldots, X_L^n)\\
&\quad -H(X_1^n, \ldots, X_L^n|W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}, Z_k^n) - H(Z_k^n)\\
&= H(W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}, X_1^n, \ldots, X_L^n)\\
&\quad +H(Z_k^n|W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}, X_1^n, \ldots, X_L^n)\\
&\quad -H(X_1^n, \ldots, X_L^n|W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}, Z_k^n) - H(Z_k^n)\\
&\ge H(X_1^n, \ldots, X_L^n) + H(Z_k^n|X_1^n, \ldots, X_L^n)\\
&\quad -H(X_1^n, \ldots, X_L^n|W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}, Z_k^n) - H(Z_k^n), \qquad (38)
\end{aligned}$$

where in the last step, the first term follows from the chain rule and nonnegativity of the entropy, and the second term follows due to the Markov chain relationship $(W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}) \to (X_1^n, \ldots, X_L^n) \to Z_k^n$.

Following from the codebook construction and the encoding scheme, it is clear that $X_1^n, \ldots, X_L^n$ are independently and uniformly distributed over their corresponding subcodebooks. Hence, we obtain

$$H(X_1^n, \ldots, X_L^n) = n\sum_{l=1}^{L}\log\left(1+\frac{|H_l|^2P_l}{1+|H_l|^2\sum_{j=l+1}^{L}P_j}\right). \qquad (39)$$

Using (37), we obtain

$$H(X_1^n, \ldots, X_L^n|W_{1[k,K_1]}, \ldots, W_{L[k,K_L]}, Z_k^n) < n\epsilon_n \qquad (40)$$

where ϵ_n approaches zero if n goes to infinity.

We also compute

$$\begin{aligned}
&\frac{1}{n}\left[H(Z_k^n|X_1^n, \ldots, X_L^n) - H(Z_k^n)\right]\\
&\qquad \ge -\frac{1}{n}\sum_{i=1}^{n}\log\left(|G_k|^2\sum_{l=1}^{L}\mathrm{E}[|X_{li}|^2]+1\right) \qquad (41)\\
&\qquad \ge -\log\left(\frac{|G_k|^2}{n}\sum_{i=1}^{n}\sum_{l=1}^{L}\mathrm{E}[|X_{li}|^2]+1\right) \qquad (42)\\
&\qquad \ge -\log\left(1+|G_k|^2\sum_{l=1}^{L}P_l\right). \qquad (43)
\end{aligned}$$

Hence,

$$\begin{aligned}\frac{1}{n}H(W_{1[k,K_1]},\ldots,W_{L[k,K_L]}|Z_k^n) &\geq \sum_{l=1}^{L}\log\left(1+\frac{|H_l|^2P_l}{1+|H_l|^2\sum_{j=l+1}^{L}P_j}\right)\\ &\quad-\log\left(1+|G_k|^2\sum_{j=1}^{L}P_j\right)-\epsilon\\ &=\sum_{l=1}^{L}\left[\log\left(1+\frac{|H_l|^2P_l}{1+|H_l|^2\sum_{j=l+1}^{L}P_j}\right)\right.\\ &\quad\left.-\log\left(1+\frac{|G_k|^2P_l}{1+|G_k|^2\sum_{j=l+1}^{L}P_j}\right)\right]-\epsilon, \quad (44)\end{aligned}$$

where the last step applies

$$\begin{aligned}&\log\left(1+|G_k|^2\sum_{j=1}^{L}P_j\right)\\ &\qquad=\sum_{l=1}^{L}\log\left(1+\frac{|G_k|^2P_l}{1+|G_k|^2\sum_{j=l+1}^{L}P_j}\right). \quad (45)\end{aligned}$$

Comparing equation (44) with the rates of the messages given in (36), we conclude that perfect secrecy is achieved asymptotically as n approaches infinity. ■

B. Continuous Channel States

We now generalize our result in the preceding subsection to the case in which the channel states take continuous values. For each channel state pair $(H,G)=(h,g)$, we let $(s,u)=(|h|^2,|g|^2)$, and use (s,u) to index layers of messages. For each layer s, we assume that the transmitter allocates power $\rho(s)ds$, and we use $\Sigma(s)$ to denote the total power allocated to the layers with better channel states, i.e., the states $\hat{s}$ such that $\hat{s}>s$. Hence,

$$\Sigma(s)=\int_s^{\infty}\rho(x)dx \quad (46)$$

and

$$\rho(s)=-\Sigma'(s). \quad (47)$$

Following from Theorem 4, we obtain the following result on the average secrecy rate.

Corollary 3: For the fading wiretap channel with both the legitimate receiver and the eavesdropper having block fading channels with continuous channel states, the average secrecy rate under the delay constraint achieved via the broadcast approach described above is given by

$$R=\log e\max_{\Sigma(x)}\int_0^{\infty}dx(1-F(x))\Sigma'(x)\left[\frac{-xQ(x)}{1+x\Sigma(x)}+\int_0^{x}du\frac{uq(u)}{1+u\Sigma(x)}\right] \quad (48)$$

where $F(\cdot)$ and $Q(\cdot)$ are cumulative distribution functions for s and u, respectively.

Proof of Corollary 3: Consider the case when the legitimate receiver and the eavesdropper have the channel states $(s,u)=(|h|^2,|g|^2)$. Following from (35), if $s>u$, then the rate of the messages that can be decoded by the legitimate receiver at the state s while being kept secure from the eavesdropper at the state u is given by

$$\begin{aligned}dR&=\log\left(1+\frac{s\rho(s)ds}{1+s\Sigma(s)}\right)-\log\left(1+\frac{u\rho(s)ds}{1+u\Sigma(s)}\right)\\ &\approx\log e\left[\frac{s\rho(s)ds}{1+s\Sigma(s)}-\frac{u\rho(s)ds}{1+u\Sigma(s)}\right] \quad (49)\end{aligned}$$

where the second equation follows because ds approaches zero. If $s\leq u$, then $dR=0$. Since all messages corresponding to the legitimate receiver's state x such that $x<s$ can be decoded by the legitimate receiver at state s, the total rate of the messages that can be decoded by the legitimate receiver at the state s and also be kept secure from the eavesdropper at the state u is given by

$$R(s,u)=\log e\int_u^{s}\left[\frac{x\rho(x)}{1+x\Sigma(x)}-\frac{u\rho(x)}{1+u\Sigma(x)}\right]dx \quad (50)$$

if $s>u$, and $R(s,u)=0$ if $s\leq u$. An illustration of the layers of messages that contribute to the secrecy rate $R(s,u)$ is depicted in Fig. 5.

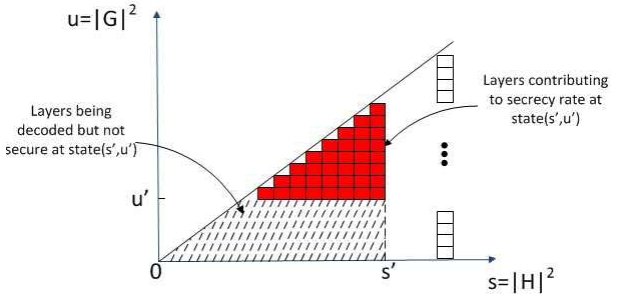

Fig. 5. An illustration of the layers of messages that are decodable at the legitimate receiver and secure from the eavesdropper.

Averaging the above rate over all fading state realizations of the legitimate receiver's channel and the eavesdropper's channel, we obtain

$$\begin{aligned}R&=\int_0^{\infty}ds\int_0^{s}duf(s)q(u)R(s,u)\\ &=\log e\int_0^{\infty}du\int_u^{\infty}dsf(s)q(u)\\ &\quad\times\int_u^{s}dx\left[\frac{x\rho(x)}{1+x\Sigma(x)}-\frac{u\rho(x)}{1+u\Sigma(x)}\right] \quad (51)\\ &=\log e\int_0^{\infty}duq(u)\\ &\quad\times\int_u^{\infty}dx\rho(x)\left[\frac{x}{1+x\Sigma(x)}-\frac{u}{1+u\Sigma(x)}\right]\int_x^{\infty}dsf(s)\\ &=\log e\int_0^{\infty}duq(u)\\ &\quad\times\int_u^{\infty}dx(1-F(x))\rho(x)\left[\frac{x}{1+x\Sigma(x)}-\frac{u}{1+u\Sigma(x)}\right]\\ &=\log e\int_0^{\infty}dx(1-F(x))\rho(x)\\ &\quad\times\left[\frac{x}{1+x\Sigma(x)}\int_0^{x}duq(u)-\int_0^{x}du\frac{uq(u)}{1+u\Sigma(x)}\right]\\ &=\log e\int_0^{\infty}dx(1-F(x))\rho(x)\\ &\quad\times\left[\frac{xQ(x)}{1+x\Sigma(x)}-\int_0^{x}du\frac{uq(u)}{1+u\Sigma(x)}\right]. \quad (52)\end{aligned}$$

The average rate R given above can be further improved by optimizing over all possible power allocation functions $\rho(\cdot)$, or equivalently, over all possible cumulative power allocation functions $\Sigma(\cdot)$. We can also use (47) to replace $\rho(x)$ with $-\Sigma'(x)$, which concludes the proof. ■

As in Section IV, we can also characterize the probability that a given secrecy rate R is achievable, denoted by $Pr(R)$. By setting $u = 0$ in (50), we obtain the following total rate of the messages that the legitimate receiver at the state s can decode:

$$R(s) = \log e \int_0^s \frac{x\rho(x)}{1 + x\Sigma(x)} dx.$$

We set $R(s_T) = R$, and can numerically obtain s_T, which represents the lowest state of the legitimate receiver that can decode the messages at the rate R. If $s < s_T$, the probability of achieving the secrecy rate R when the legitimate receiver's state is in s is zero, i.e., $Pr(R|s) = 0$. Otherwise, for any state $s \geq s_T$, we characterize the probability that the given secrecy rate R is achievable. Towards this end, we set $R(s, u_R) = R$ in (50), and then fix s and solve the equation to obtain $u_R(s)$, which is a function of s. Such $u_R(s)$ exists because $R(s, u)$ in (50) is monotonic as a function of u, and can be found numerically. It is clear that $u_R(s)$ is the best eavesdropper's state such that messages with the rate R are secure. Since these messages are also secure in any eavesdropper's state $\hat{u} \leq u_R(s)$, $Pr(R|s) = Q(u_R(s))$. Thus, the total probability $Pr(R)$, which is the probability that the messages with the given rate R are secure from the eavesdropper, can be obtained by averaging $P(R|s)$ over all states $s \geq s_T$, and is given by

$$Pr(R) = \int_{s_T}^{\infty} f(s) Q(u_R(s)) ds.$$

From the legitimate receiver's point of view, since it knows its own channel state, the conditional probability $P(R|s) = Q(u_R(s))$ characterizes the probability to achieve a certain secrecy rate R at the current block with the state $s = |H|^2$.

In order to obtain the optimal average secrecy rate R given in (48), we need to solve the following optimization problem:

$$\max_{\Sigma(x)} \int_0^{\infty} S(x, \Sigma(x), \Sigma'(x)) dx$$
$$\text{subject to } 0 \leq \Sigma(x) \leq P, \quad \Sigma'(x) \leq 0, \quad \text{for } x \geq 0; \quad (53)$$

where

$$S(x, \Sigma(x), \Sigma'(x)) = (1 - F(x)) Q(x) \frac{-x\Sigma'(x)}{1 + x\Sigma(x)} + (1 - F(x))\Sigma'(x) \int_0^x \frac{uq(u)}{1 + u\Sigma(x)} du. \quad (54)$$

Theorem 5: An optimal solution to (53), if one exists, has the following structure. There exist $0 \leq x_1 < y_1 < x_2 < y_2 < \cdots < x_m < y_m$, and a function $\eta(x)$, such that $\eta(x)$ satisfies

$$\frac{(1 - F(x)) Q(x)}{(1 + x\eta(x))^2} = \frac{x f(x) Q(x)}{1 + x\eta(x)} - f(x) \int_0^x \frac{uq(u)}{1 + u\eta(x)} du \quad (55)$$

and is strictly decreasing over $[x_i, y_i]$ for $i = 1, \ldots, m$, $\eta(x_1) = P$, $\eta(y_m) = 0$, $\eta(y_i) = \eta(x_{i+1})$ for $i = 1, \ldots, m-1$, and an optimal $\Sigma^*(x)$ is given by

$$\Sigma^*(x) = \begin{cases} P & 0 \leq x \leq x_1; \\ \eta(x) & x_i \leq x \leq y_i, \\ & \text{for } i = 1, \ldots, m; \\ \eta(y_i) = \eta(x_{i+1}), & y_i < x < x_{i+1}, \\ & \text{for } i = 1, \ldots, m-1; \\ 0 & y_m \leq x. \end{cases} \quad (56)$$

Proof: The argument is similar to that for proving Theorem 2. Hence, we here provide only details for obtaining the Euler condition (55). Due to the complementary slackness conditions, over the intervals $(x_1, y_1]$, $[x_i, y_i]$ for $i = 2, \ldots, m-1$, and $[x_m, y_m)$, since $\Sigma^*(x)$ does not satisfy the inequality constraints with equality, i.e., it is not on the boundary of the constraint set, then the following Euler equation must be satisfied:

$$S_{\Sigma} - \frac{d}{dx} S_{\Sigma'} = 0. \quad (57)$$

For the function $S(x, \Sigma(x), \Sigma'(x))$ given in (54), we obtain

$$S_{\Sigma} = (1 - F(x)) Q(x) \frac{x^2 \Sigma'(x)}{(1 + x\Sigma(x))^2} + (1 - F(x))\Sigma'(x) \int_0^x \frac{-u^2 q(u)}{(1 + u\Sigma(x))^2} du, \quad (58)$$

$$S_{\Sigma'} = (1 - F(x)) Q(x) \frac{-x}{1 + x\Sigma(x)} + (1 - F(x)) \int_0^x \frac{uq(u)}{1 + u\Sigma(x)} du, \quad (59)$$

$$\frac{d}{dx} S_{\Sigma'} = [-f(x)Q(x) + (1 - F(x))q(x)] \frac{-x}{1 + x\Sigma(x)} + (1 - F(x)) Q(x) \frac{-1 + x^2\Sigma'(x)}{(1 + x\Sigma(x))^2} - f(x) \int_0^x \frac{uq(u)}{1 + u\Sigma(x)} du + (1 - F(x)) \frac{xq(x)}{1 + x\Sigma(x)} - (1 - F(x)) \int_0^x \frac{u^2 q(u)\Sigma'(x)}{(1 + u\Sigma(x))^2}. \quad (60)$$

We substitute the above equations into the Euler equation and obtain the condition given in (55). ■

Example 2: Consider the case when the channels to the legitimate receiver and the eavesdropper experience independent Rayleigh fading, i.e., s and u are exponentially distributed as characterized by

$$f(x) = \frac{1}{\sigma_1} e^{-\frac{x}{\sigma_1}} \quad \text{and} \quad F(x) = 1 - e^{-\frac{x}{\sigma_1}}, \quad x \geq 0, \quad (61)$$

$$q(x) = \frac{1}{\sigma_2} e^{-\frac{x}{\sigma_2}} \quad \text{and} \quad Q(x) = 1 - e^{-\frac{x}{\sigma_2}}, \quad x \geq 0. \quad (62)$$

where σ_1 and σ_2 are parameters for the exponential distributions of s and u, respectively.

The Euler condition (55) now becomes

$$\frac{1 - e^{-\frac{x}{\sigma_2}}}{(1 + x\Sigma(x))^2} - \frac{x(1 - e^{-\frac{x}{\sigma_2}})}{\sigma_1(1 + x\Sigma(x))} + \frac{1}{\sigma_1\sigma_2} \int_0^x \frac{ue^{-\frac{u}{\sigma_2}}}{1 + u\Sigma(x)} du = 0. \quad (63)$$

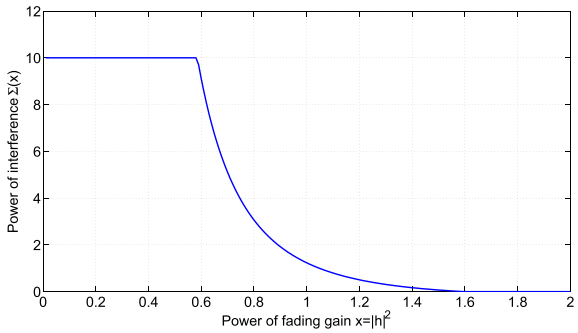

Fig. 6. An optimal function $\Sigma(x)$ for the Rayleigh fading channel with $P = 10dB$ and $\sigma_1 = \sigma_2 = 1$.

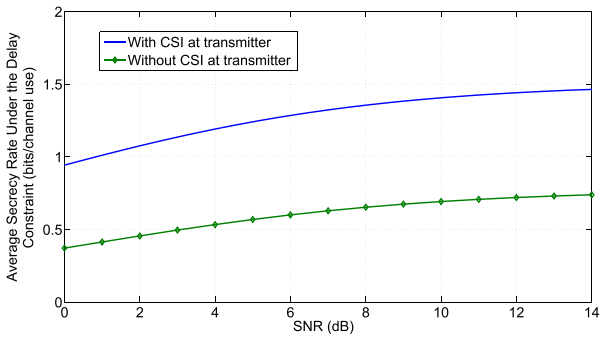

Fig. 7. Comparison of rates for scenario 1: only the channel to the legitimate receiver is fading.

Consider the case with $\sigma_1 = \sigma_2 = 1$. Following from the above condition, if $\Sigma(y_1) = 0$, then y_1 satisfies

$$2 - 2e^{-y_1} - y_1 = 0$$

whose root can be computed numerically and is equal to

$$y_1 = 1.5936.$$

Using the condition (63), it is easy to find a $\Sigma^*(x)$ function that satisfies the necessary condition given in Theorem 5. We plot the function $\Sigma^*(x)$ in Fig. 6 for the case with the power $P = 10dB$ and $\sigma_1 = \sigma_2 = 1$. We note that this function $\Sigma^*(x)$ is strictly decreasing over the interval $[x_1, y_1]$, which suggests that the optimal solution is unique if it exists.

This example also demonstrates the impact of probabilistic secrecy, under which we achieve a positive secrecy rate under delay constraints for certain channel state realizations as demonstrated in Section VI. However, under a deterministic secrecy constraint that requires all transmitted messages be secure from the eavesdropper, only zero secrecy rate can be achieved for any block. Even over a large number of blocks, the secrecy rate is zero under a deterministic secrecy constraint if the legitimate receiver and the eavesdropper have the same channel statistics whereas the secrecy rate is positive under the probabilistic secrecy for the same scenario as for the above example.

VI. Numerical Results

In this section, we provide numerical examples to demonstrate the impact of the CSI at the transmitter on the average secrecy rate. We also compare the average secrecy rates for the three scenarios studied in the paper.

We first consider scenario 1 as studied in Example 1, in which only the legitimate receiver's channel is fading with the Rayleigh distribution and the eavesdropper's channel is constant. The distribution of $s = |H|^2$ is exponential with the parameter $\sigma_1 = 2$, i.e., $p(s) = \frac{1}{\sigma_1} e^{-s/\sigma_1}$. The eavesdropper's channel state is at $|G|^2 = 0.5$. In Fig. 7, we plot the average secrecy rates achieved via the broadcast approach and compare them with the rates achievable when the legitimate receiver's CSI is known at the transmitter and the eavesdropper. With the legitimate receiver's CSI at the transmitter, the average secrecy rate (which is also the capacity) can be obtained by averaging the secrecy rate for each channel state over the state distribution and optimizing over all possible power allocation over the channel states as given below

$$\bar{R} = \max_{P(s):E_s[P(s)]\leq P} \int_{|G|^2}^{\infty} \Big[\log\left(1+sP(s)\right) - \log\left(1+|G|^2P(s)\right)\Big]\rho(s)ds \quad (64)$$

where the optimizing power allocation can be obtained by using the Lagrangian multiplier method as in [21] and [23].

It is clear from Fig. 7 that the knowledge of the legitimate receiver's CSI provides a great advantage in achieving better secrecy rates. Due to the lack of the CSI, the transmitter's power is spread over many layers of messages in order to accommodate possibly occurring channel states. However, when the CSI is available, the transmitter spends all its power for the particular state realization at each coherence block. In this way, the CSI helps to use the transmitter's power more efficiently. We also note that if one adopts the compound channel approach [12] that requires secrecy no matter which legitimate receiver's state occurs, then the secrecy rate for this example is zero. Hence, the broadcast approach offers an attractive alternative to the compound channel approach [12] although the transmitter does not have the CSI. We also note that for this scenario, if there is no delay constraint, even if the transmitter does not know the CSI, it can exploit the statistics of the legitimate receiver's channel to achieve a better secrecy rate. Here, the channel statistics help to avoid power spreading whereas the broadcast approach inherently degrades the rates due to power spreading over layers.

We then study scenario 2, in which only the eavesdropper's channel is fading with the Rayleigh distribution and the legitimate receiver's channel is constant. The distribution of $u = |G|^2$ is exponential with the parameter $\sigma_2 = 0.5$, i.e., $p(u) = \frac{1}{\sigma_2} e^{-u/\sigma_2}$. The legitimate receiver's channel state is at $|H|^2 = 2$. In Fig. 8, we plot the average secrecy rates achieved via the broadcast approach and compare them with the rates achievable when the eavesdropper's CSI is known at the transmitter. With the eavedropper's CSI at the transmitter, the average secrecy rate (which is also the capacity) under the delay constraint is given by

$$\bar{R} = \max_{P(u):E_u[P(u)]\leq P} \int_{0}^{|H|^2} \Big[\log\left(1+|H|^2P(u)\right) - \log\left(1+uP(u)\right)\Big]q(u)du \quad (65)$$

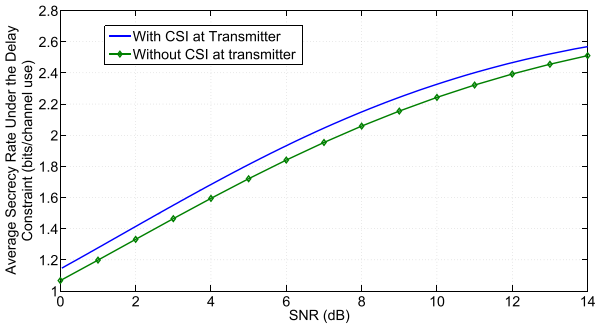

Fig. 8. Comparison of rates for scenario 2: only the channel to the eavesdropper is fading.

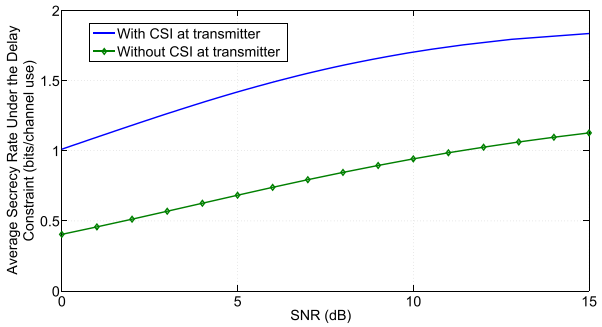

Fig. 9. Comparison of rates for scenario 3: the channels to both the legitimate receiver and the eavesdropper are fading.

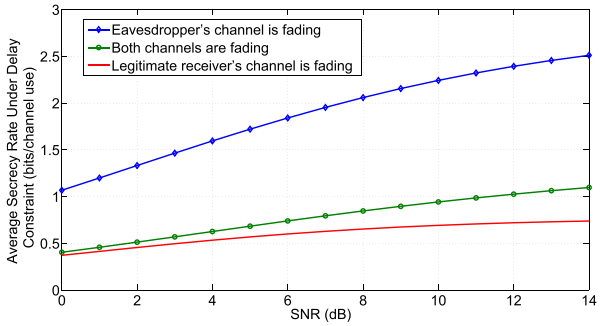

Fig. 10. Comparison of rates for the three scenarios.

where the optimizing power allocation can be obtained by using the Lagrangian multiplier method as in [21] and [23].

It is clear from Fig. 8 that the rates corresponding to the two cases are very close, suggesting that the knowledge of the eavesdropper's CSI does not provide much advantage to achieve better secrecy rates. This is not surprising, as we have seen in Section IV that the broadcast approach already achieves the maximum possible secrecy rate for each block. The small gap between the two rates is because with the CSI, the transmitter can adapt its power allocation over the channel states to achieve a better rate. Another role that the CSI plays is that with the CSI the transmitter guarantees secrecy for all transmitted messages, whereas without the CSI the transmitter does not guarantee secrecy for all transmitted messages, and the legitimate receiver knows only the probability that a certain secrecy rate is achievable without the eavesdropper's CSI. Hence, without the CSI, a certain amount of power is wasted over some layers of messages that happens not to be secure for certain realizations of the channel state; whereas with the CSI, this amount of power can be used to convey information securely. We note that this does not mean that the gap between the two cases in Fig. 8 bounds the information leakage, because the same amount of power may generate different values of rates if it is used differently.

We further note that the secrecy rate that can be achieved using the compound channel approach is zero for this example due to the assumption that all transmitted messages must be secure no matter which eavesdropper's state occurs. Therefore the broadcast approach adopted here again offers an attractive alternative to the compound channel approach. However, unlike the first scenario and the compound channel approach, the broadcast approach developed here does not guarantee the secrecy of the entire message and achieves only probabilistic secrecy.

We now consider scenario 3 as studied in Example 2, in which both the channel to the legitimate receiver and the channel to the eavesdropper are fading. The distributions of $s = |H|^2$ and $u = |G|^2$ are independent and are both exponential with the parameters $\sigma_1 = 2$ and $\sigma_2 = 0.5$, i.e., $p(s) = \frac{1}{\sigma_1} e^{-s/\sigma_1}$ and $p(u) = \frac{1}{\sigma_2} e^{-u/\sigma_2}$, respectively. In Fig. 8, we plot the average secrecy rates achieved via the broadcast approach and compare them with the rates achievable when both channels' CSI is known at the transmitter and the eavesdropper. With the CSI at the transmitter, the average secrecy rate (which is also the capacity) under the delay constraint is given by

$$\bar{R} = \max_{P(u,s):E_{s,u}[P(s,u)]\le P} \int_0^\infty ds \int_0^s du \rho(s) q(u) \times \left[\log\left(1 + sP(s,u)\right) - \log\left(1 + uP(s,u)\right)\right] \quad (66)$$

where the optimizing power allocation can be obtained by using the Lagrangian multiplier method as in [21] and [23]. From our understanding of scenarios 1 and 2, the gap between the rates corresponding to the two cases is mainly due to the lack of the legitimate receiver's CSI which results in the transmitter's power being spread over states. Similar to scenario 2, the secrecy rate that can be achieved using the compound channel approach is zero for this example. Therefore the broadcast approach adopted here again improves the secrecy rate although the entire message may not be fully kept secure. We also note that Figs. 7–9, study only the impact of the CSI at the transmitter on the average secrecy rate. If we include the information revealing rate (i.e., the leakage rate) to the eavesdropper, such impact may be different under this metric. Hence, it would be of interest to study the information revealing rate, which would provide a complementary view to the achievable secrecy rate studied here.

We further compare the average secrecy rates for the three scenarios in Fig. 10, none of which have the CSI at the transmitter. It is clear from the figure that scenario 2 has the best rate, and scenario 3 has a better rate than scenario 1. It is easy to understand that scenario 3 has worse rates than scenario 2 because the transmitter's power is spread over

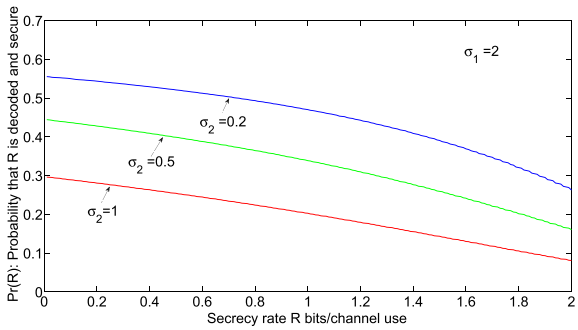

Fig. 11. Comparison of the probabilities of a rate R is decoded and secure for scenario 3.

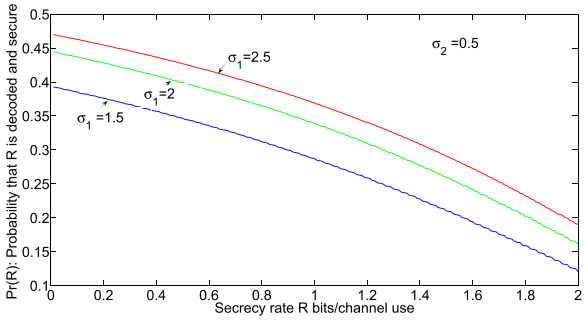

Fig. 12. Comparison of the probabilities of a rate R is decoded and secure for scenario 3.

the states due to no knowledge of the legitimate receiver's CSI. However, it may seem counter-intuitive that scenario 3 has better rates than scenario 1. This is due to the fact that when the eavesdropper's channel is fading, there is a good chance that its state is below the channel average, and such channel fluctuation facilitates achievement of a better secrecy rate and overcomes the effect of no eavesdropper's CSI at the transmitter. Therefore, the two major factors that affect the secrecy rate are the knowledge of the legitimate receiver's CSI and the channel fluctuation of the eavesdropper. The knowledge of the eavesdropper's CSI only weakly affects the secrecy rate.

We now study the performance of the probabilistic secrecy for scenario 3. Figs. 11 and 12 plot and compare the probabilities of achieving a certain target secrecy rate over a large number of blocks for a number of values of σ_1 and σ_2. It can be seen from Fig. 11 that with σ_1 fixed, the probability of secrecy becomes larger as σ_2 decreases. This is reasonable because a smaller σ_2 implies less chance that the eavesdropper's channel can be in a good state, which results in a higher probability of achieving a certain secrecy rate. It can also be seen from Fig. 12 that with σ_2 fixed, the probability of secrecy becomes larger as σ_1 increases. This is also reasonable because a larger σ_1 implies a better legitimate receiver's channel, and hence results in higher probability of achieving a target secrecy rate.

VII. Discussion and Conclusion

In this paper, we have studied a (layered) broadcast approach for fading wiretap channels. We have developed two broadcast approaches for the cases in which either the legitimate receiver's or the eavesdropper's channel is fading, respectively, and have combined these two approaches for the general case in which both nodes' channels are fading. For each case, we have obtained the average secrecy rate achieved under the delay constraint by using the broadcast approach and have derived the optimal power allocation across layers. We have also introduced a notion of probabilistic secrecy, and characterized the probability that a given secrecy rate is achievable for the valid scenarios when the eavesdropper's channel is fading. Moreover, we have provided numerical examples to demonstrate how the CSI at the transmitter and the channel fluctuations of the eavesdropper affect the average secrecy rate.

Here, we provide some further notes on the notion of probabilistic secrecy. The design goal of this paper is to maximize the secrecy rate, i.e., to transmit as many bits as possible in secret. In order to achieve this goal, it is unavoidable to leak information, because there is no CSI at the transmitter, and the information designed to be secure when the eavesdropper's state is weak can be leaked if the eavesdropper's state turns out to be strong. In this sense, the notion of probabilistic secrecy can be defined in a more general way to include both the average secrecy rate and the average leakage rate as performance measures. In this case, the average secrecy versus leakage rates dictate a rate region, which captures the tradeoff between the two rates. Characterization of such a rate region depends on the actual power allocations across layers if the layered approach is applied, i.e., each point in the region is achieved by a specific power allocation scheme. Optimizing various combinations of the secrecy and leakage rates achieves boundary points of the region. From this more general viewpoint, the problem we study is a special case, optimizing only the secrecy rate without the leakage rate constraint.

We note that if the focus is on the leakage rate, power allocation schemes can be significantly different from what we study in this paper. We use a simple example to illustrate this as follows. Suppose the legitimate receiver has two states H_1 and H_2, and the eavesdropper has two states G_1 and G_2. We further assume that $|H_2| > |G_2| > |H_1| > |G_1|$. For such a system, if we want to minimize the leakage rate, we should assign power only to one layer corresponding to the state pair (H_2, G_2) so that no information can be leaked (and hence the leakage rate is zero), although now the receiver decodes only at the state H_2. However, if we aim at maximizing the secrecy rate and do not consider the leakage rate, we should assign power to three layers respectively corresponding to the state pairs (H_2, G_2), (H_2, G_1) and (H_1, G_1). Clearly for this case, there is inevitable information leakage associated with layers (H_2, G_1) and (H_1, G_1) if the state turns out to be (H_2, G_2). This example suggests that the leakage rate can vary significantly depending on how we design the power allocation scheme.

It is clear that the more general notion of probabilistic secrecy facilitates the evaluation of the secrecy rate and the leakage rate as functions of power allocations across layers (i.e., across the fading parameters at the legitimate user and the eavesdropper assuming both are unavailable at the transmitter). What to optimize depends on the application, and we have focused on the secrecy rate in this paper. Our study can

be viewed as a generalization of the broadcast approach in [16] and the variable-to-fixed channel rate coding in [18], where the later becomes a special case when the eavesdropper enjoys no received signal (i.e., zero fading realization). Other choices of the performance measure are also possible, which result in different optimization problems (which are not always guaranteed to have analytical solutions) and consequently different transmission schemes as demonstrated above.

Several directions are interesting to explore in the future. The development of upper bounds on the average achievable secrecy rate remains open. Although the secrecy capacity for the case in which the CSI is known at the transmitter can serve as an upper bound, such a bound in general is not tight. It is thus interesting to provide tighter upper bounds in order to better understand how well the broadcast approach performs. Moreover, since under the system model studied here, not all information is secure from the eavesdropper, it is interesting to study how much information is revealed to the eavesdropper jointly with the average secrecy rate. The tradeoff between the rates would provide a more general characterization of the performance. For the case with a delay constraint, it is of interest to explore the broadcast approach jointly with a key-based technique recently proposed in [27]. It is also of interest to study the broadcast approach for the case with a relaxed delay constraint, in which coding over a few blocks is allowed. Some ideas in [28] may be further explored for the case with a secrecy constraint. It is also of great importance to evaluate the penalty incurred by delay constraints, in particular, a stringent one-block constraint, by comparing the secrecy rate under a delay constraint and the ergodic secrecy rate. As a final comment, it is of interest to explore whether progressive encoding of layering as appeared in, e.g., [29] and [30], could be useful for the fading wiretap channel. The idea is that layers would be progressively (but not independently) encoded, which can be useful for studying the channel with secrecy constraints, and may be closely related to the embedded coding proposed in [22].

ACKNOWLEDGMENT

The authors would like to thank Shaofeng Zou of Syracuse University for providing the numerical results in Figs. 11 and 12, and would also like to thank Dr. Tie Liu of Texas A&M University for his helpful discussions about the broadcast approach based on embedded coding.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell. Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[3] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, Now Publisher, Hanover, MA, USA, 2008.

[4] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[5] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[6] Z. Li, R. D. Yates, and W. Trappe, "Secret communication with a fading eavesdropper channel," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Jun. 2007, pp. 1296–1300.

[7] A. Khisti, A. Tchamkerten, and G. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inf. Theory, Special Issue Inf. Theoretic Security*, vol. 54, no. 6, pp. 2453–2469, Jun. 2008.

[8] X. He and A. Yener, "The role of channel states in secret key generation," in *Proc. IEEE 21st Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2010, pp. 2681–2686.

[9] E. Ekrem and S. Ulukus, "Degraded compound multi-receiver wiretap channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5681–5698, Sep. 2012.

[10] A. Khisti, "On the MISO compound wiretap channel," in *Proc. Inform. Theory Applic. Workshop (ITA)*, Jan. 2010, pp. 1–7.

[11] T. Liu, V. Prabhakaran, and S. Vishwanath, "The secrecy capacity of a class of parallel Gaussian compound wiretap channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Jul. 2008, pp. 116–120.

[12] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wire-tap channels," *EURASIP J. Wireless Commun. Netw.*, vol. 3, pp. 1–8, Jan. 2009.

[13] M. Kobayashi, Y. Liang, S. Shamai (Shitz), and M. Debbah, "On the compound MIMO broadcast channels with confidential messages," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Jun./Jul. 2009, pp. 1283–1287.

[14] X. He and A. Yener, "Providing secrecy when the eavesdropper channel is arbitrarily varying: A case for multiple antennas," in *Proc. 48th Annu. Allerton Conf. Commun., Control Comput.*, Sep./Oct. 2010, pp. 1228–1235.

[15] X. He and A. Yener, "MIMO wiretap channels with arbitrarily varying eavesdropper channel states," *submitted to IEEE Trans. Inf. Theory*, Jul. 2010.

[16] S. Shamai (Shitz) and A. Steiner, "A broadcast approach for a single-user slowly fading MIMO channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2617–2635, Oct. 2003.

[17] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[18] S. Verdú and S. Shamai (Shitz), "Variable-rate channel capacity," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2651–2667, Jun. 2010.

[19] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Sep. 2005, pp. 2152–2155.

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

[21] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory, Special Issue Inf. Theoretic Security*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[22] H. D. Ly, T. Liu, and Y. Blankenship, "Security embedding codes," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 1, pp. 148–159, Feb. 2012.

[23] Y. Liang and H. V. Poor, "Secure communication over fading channels," in *Proc. 44th Annu. Allerton Conf. Commun., Control Comput.*, Sep. 2006, pp. 1–10.

[24] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," in *Proc. Annu. Allerton Conf. Commun., Control Comput.*, Sep. 2006, pp. 10–15.

[25] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "Secret-key sharing based on layered broadcast coding over fading channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Jun./Jul. 2009, pp. 2762–2766.

[26] J. Gregory and K. Pericak-Spector, "New methods of solving general constrained calculus of variations problems involving PDEs," *Utilitas Math.*, vol. 58, pp. 215–224, Nov. 2000.

[27] K. Khalil, M. Youssef, O. O. Koyluoglu, and H. El Gamal, "On the delay limited secrecy capacity of fading channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Jun./Jul. 2009, pp. 2617–2621.

[28] P. A. Whiting and E. M. Yeh, "Broadcasting over uncertain channels with decoding delay constraints," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 904–921, Mar. 2006.

[29] D. Gunduz and O. Simeone, "On the capacity region of a multiple access channel with common messages," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Jun. 2010, pp. 470–474.

[30] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz), "Multi-layer transmission and hybrid relaying for relay channels with multiple out-of-band relays," *Trans. Emerging Telecommun. Technol.*, Aug. 2013.

Yingbin Liang (S'01–M'05) received the Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 2005. In 2005-2007, she was working as a postdoctoral research associate at Princeton University. In 2008-2009, she was an assistant professor at the Department of Electrical Engineering at the University of Hawaii. Since December 2009, she has been a faculty member at the Department of Electrical Engineering and Computer Science at the Syracuse University. Dr. Liang's research interests include communications, wireless networks, information theory, and machine learning.

Dr. Liang was a Vodafone Fellow at the University of Illinois at Urbana-Champaign during 2003-2005, and received the Vodafone-U.S. Foundation Fellows Initiative Research Merit Award in 2005. She also received the M. E. Van Valkenburg Graduate Research Award from the ECE department, University of Illinois at Urbana-Champaign, in 2005. In 2009, she received the National Science Foundation CAREER Award, and the State of Hawaii Governor Innovation Award. She is currently serving as an Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY.

Lifeng Lai (M'07) received the B.E. and M.E. degrees from Zhejiang University, Hangzhou, China in 2001 and 2004 respectively, and the PhD degree from The Ohio State University at Columbus, OH, in 2007. He was a postdoctoral research associate at Princeton University from 2007 to 2009, and was an assistant professor at University of Arkansas, Little Rock from 2009 to 2012. Since August 2012, he has been an assistant professor at the Worcester Polytechnic Institute. Dr. Lai's research interests include information theory, stochastic signal processing and their applications in wireless communications, security and other related areas.

Dr. Lai was a Distinguished University Fellow of the Ohio State University from 2004 to 2007. He is a co-recipient of the Best Paper Award from IEEE Global Communications Conference (Globecom) in 2008, the Best Paper Award from IEEE International Conference on Communications (ICC) in 2011 and the Best Paper Award from IEEE Smart Grid Communications (SmartGridComm) in 2012. He received the National Science Foundation CAREER Award in 2011, and Northrop Young Researcher Award in 2012. He served as a Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on Signal Processing Techniques for Wireless Physical Layer Security. He is currently serving as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. He has also held visiting appointments at several other institutions, most recently at Imperial College and Stanford. Dr. Poor's research interests are in the areas of information theory, stochastic analysis, and statistical signal processing, and their applications in wireless networks and related fields including social networks and smart grid. Among his publications in these areas are *Smart Grid Communications and Networking* (Cambridge University Press, 2012) and *Principles of Cognitive Radio* (Cambridge University Press, 2013).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, a Fellow of the American Academy of Arts and Sciences, an International Fellow of the Royal Academy of Engineering (U. K.), and a Corresponding Fellow of the Royal Society of Edinburgh. He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, in 2004-07 as the Editor-in-Chief of these TRANSACTIONS, and in 2009 as General Co-chair of the IEEE International Symposium on Information Theory, held in Seoul, South Korea. He received a Guggenheim Fellowship in 2002 and the IEEE Education Medal in 2005. Recent recognition of his work includes the 2010 IET Ambrose Fleming Medal for Achievement in Communications, the 2011 IEEE Eric E. Sumner Award, the 2011 IEEE Information Theory Paper Award, and honorary doctorates from Aalborg University, the Hong Kong University of Science and Technology, and the University of Edinburgh.

Shlomo Shamai (Shitz) (S'80–M'82–SM'88–F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion-Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975-1985 he was with the Communications Research Labs, in the capacity of a Senior Research Engineer. Since 1986, he has been with the Department of Electrical Engineering, Technion-Israel Institute of Technology, where he is now a Technion Distinguished Professor, and holds the William Fondiller Chair of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shlomo Shamai is an IEEE Fellow, a member of the Israeli Academy of Sciences and Humanities and a Foreign Associate of the U.S. National Academy of Engineering. He is the recipient of the 2011 Claude E. Shannon Award and the 2014 Rothschild Prize in Mathematics/Computer Sciences and Engineering. He has been awarded the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and is a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++) Best Paper Award, and the 2010 Thomson Reuters Award for International Excellence in Scientific Research. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served twice on the Board of Governors of the Information Theory Society. He is a member of the Executive Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY.