This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

Multi-party secret key agreement over state-dependent wireless broadcast channels

Mahdi Jafari Siavoshani, *Member, IEEE,* Shaunak Mishra, *Member, IEEE,* Christina Fragouli, *Fellow, IEEE,* Suhas N. Diggavi, *Fellow, IEEE*

***Abstract*—We consider a group of m trusted and authenticated nodes that aim to create a shared secret key K over a wireless channel in the presence of an eavesdropper Eve. We assume that there exists a state dependent wireless broadcast channel from one of the honest nodes to the rest of them including Eve. All of the trusted nodes can also discuss over a cost-free, noiseless and unlimited rate public channel which is also overheard by Eve.**

For this setup, we develop an information-theoretically secure secret key agreement protocol. We show the optimality of this protocol for "linear deterministic" wireless broadcast channels. This model generalizes the packet erasure model studied in literature for wireless broadcast channels. Here, the main idea is to convert a deterministic channel to multiple independent erasure channels by using superposition coding.

For "state-dependent Gaussian" wireless broadcast channels, by using insights from the deterministic problem, we propose an achievability scheme based on a multi-layer wiretap code. By using the wiretap code, we can mimic the phenomenon of converting the wireless channel to multiple independent erasure channels. Then, finding the best achievable secret key generation rate leads to solving a non-convex power allocation problem over these channels (layers). We show that using a dynamic programming algorithm, one can obtain the best power allocation for this problem. Moreover, we prove the optimality of the proposed achievability scheme for the regime of high-SNR and large-dynamic range over the channel states in the (generalized) degrees of freedom sense.

Keywords

Secret key sharing, multi-terminal secrecy, information theoretical secrecy, wireless channel, public discussion.

I. Introduction

We consider the problem of generating a secret key K among $m \geq 2$ honest (trusted and authenticated) nodes that communicate over a wireless channel in the presence of a passive eavesdropper Eve (for example consider a scenario where all people in a conference room aim to generate a common secret key in the presence of one or multiple adversaries behind the doors). We restrict our attention to the case where communication occurs either through a broadcast channel, where the received symbols are independent among all receivers of the broadcast transmissions including Eve (given that the transmitted symbols is known), or, through a no-cost noiseless public channel.

Here, extending our earlier partial results appeared in [1], we focus on the group secret key agreement over a *state-dependent Gaussian broadcast channel*. This model can be motivated by fading wireless channels, where the channel states vary over time; i.e., the variation of SNR[1] level is modeled by the state of the channel. The use of state-dependent channels for secrecy has been of interest recently (see for example [2], [3], [4], [5] and references therein). To gain insight into our problem, we first investigate a *deterministic approximation* of the wireless channel as introduced in [6].

For the deterministic broadcast channel we will show that using a superposition based secrecy scheme [7], we can develop a group key agreement protocol that can be shown to be information-theoretically optimal. This can be done by converting the deterministic channel to multiple independent erasure channels. In particular, we show that we can get the same key agreement rate for the entire group as we would get for a single pair of nodes. Therefore this result demonstrates that in the presence of an unlimited public channel, we get secret key-agreement rates for linear deterministic channels, that is invariant to network size. Similar to the case of erasure broadcast channel [8], a key idea to get this is a connection to network coding (NC), which allows efficient (in the block length) reconciliation of the group secret (also, refer to [9, Appendix A] for a review of our previous results on the group secret key agreement over erasure broadcast channels).

We use the deterministic achievability scheme to get some insight about the Gaussian wireless broadcast channel with state. To this end, we use a multi-layer (nested message set, degraded channel) wiretap code based on the broadcast approach of [7], [10] to develop a key-agreement protocol for the noisy broadcast problem. This enables a scheme that converts the wireless channel with state to behave similar to the deterministic case. In particular, by using this technique we obtain a number of independent erasure channels. As a result, we show that the achievable secret key generation rate is given by a non-convex optimization problem that determines the power allocation over different layers of the wiretap code (e.g., different erasure channels).

Although the power allocation optimization problem is *non-convex*, by investigating and exploiting its special structure, we provide a dynamic programming based algorithm that finds the optimal solution to this optimization problem. The final

M. Jafari Siavoshani is with the Department of Computer Engineering, Sharif University of Technology, Tehran, Iran (e-mail: mjafari@sharif.edu).

S. Mishra, S. N. Diggavi, and C. Fragouli are with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA (e-mails: shaunakmishra@ucla.edu, suhas@ee.ucla.edu, christina.fragouli@ucla.edu).

This work was funded in part by NSF grant 1321120 and 1514531.

[1]Signal to noise ratio.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

solution is hard to be written in a closed form expression for the general case. However, the output of our algorithm should not be considered as a numerical approximation but an exact solution. The devised algorithm enables us to evaluate the performance of the proposed group secret key-agreement protocol for various situations.

Finally, we derive an upper bound on the secrecy rate and compare it with the achievable rate by the proposed scheme. Furthermore, we show that although the proposed achievability scheme is not optimal, it can be proved that for the high-SNR regime when there is a large-dynamic range between the channel states, this scheme is optimal in the (generalized) degrees of freedom sense.

A. Related Work

Secret key generation over wireless channels is a problem that has attracted significant interest. In a seminal paper on "wiretap" channels, Wyner [11] pioneered the notion that one can establish information-theoretic secrecy between Alice and Bob by utilizing the noisy broadcast nature of wireless transmissions. However, his scheme works only if we have perfect knowledge of Eve's channel and moreover, only if Eve has a worse channel than Bob. In a subsequent seminal work, Maurer [12] showed the value of feedback from Bob to Alice, even if Eve hears all the feedback transmissions (i.e., the feedback channel is public). He showed that even if the channel from Alice to Eve is better than that to Bob, feedback allows Alice and Bob to create a key which is information-theoretically secure from Eve (also see [13]). The problem of key agreement between a set of terminals having access to a noisy broadcast channel and a public discussion channel (visible to the eavesdropper) was studied in [14], where the secret key generation capacity is completely characterized, assuming Eve does not have access to the noisy broadcast transmissions. The case when the eavesdropper also had access to the broadcast channel was the main focus of recent work in [15], [16] which developed upper and lower bounds for secrecy rates. If the trusted nodes have access to a multi-terminal channel instead of a broadcast channel, [17] and [18], independently, derived upper and lower bounds for secret key generation capacity under the assumption that Eve has only access to the public channel.

The best achievable secrecy rate by our scheme for the Gaussian state-dependent channel is given by a non-convex optimization problem (see (20)) which can be reformulated as a *generalize linear fractional program* [19]. In [20], the *weighted throughput maximization* problem have been studied which involves a similar optimization problem to (20) and the authors employs numerical techniques introduced in [19] to solve this problem. In our case, however, the convergence time of such numerical method is not practical and we have to develop an approach in Section VII to solve optimization problem (20) analytically.

To the best of our knowledge, ours is the first work to consider multi-terminal secret key agreement over erasure networks and wireless broadcast channels with state, when Eve also has access to the noisy broadcast transmissions. Moreover, unlike the information-theoretic works (e.g., [11], [12], [13], [14], [16]) that assume infinite complexity operations, our schemes for the deterministic broadcast channels (that includes the erasure channel case [8]) as well as for the Gaussian broadcast channels are computationally efficient. It is worth mentioning that following a conference version of this work on the packet erasure channel [8], there has been some attempts to bring those ideas into practical scenarios, e.g., [21], [22], [23], [24].

The rest of the paper is organized as follows. In Section II, we introduce our notation and the problem formulation. Section III summarizes the main results of the paper. Our general upper bound on the secret key generation capacity for an independent broadcast channel is presented in Section IV. Each of the "deterministic," and "state-dependent Gaussian" models will be discussed in Section V and Section VI, respectively. The solution of the non-convex optimization problem is derived in Section VII. Finally, open questions and future directions will be discussed in Section VIII.

II. Notation and Problem Statement

For convenience, during the paper, we use $[i:j]$ to denote the set of integers $\{i, i+1, \ldots, j\}$. Given random variables $X_1, \ldots, X_m$, we write $X_{1:m}$ to denote $(X_1, \ldots, X_m)$. We use also X^t to denote $(X[1], \ldots, X[t])$ where t is the discreet time index. All the logarithms are in base two unless otherwise stated. We write $f(x) \dot{=} g(x)$ to denote that $\log f(x) = \log g(x) + o(\log x)$. The notation "$\dot{\leq}$" and "$\dot{\geq}$" are defined similarly.

A. Problem Statement

We consider a set of $m \geq 2$ honest nodes $\{0, \ldots, m-1\}$ that aim to share a secret key K among themselves while keeping it concealed from a passive adversary Eve, denoted by "E". Eve does not perform any transmissions, but is trying to eavesdrop on (overhear) the communications between the honest nodes[2].

We assume that Alice (terminal 0) has access to a broadcast channel such that the rest of the terminals (including Eve) receive *independent* noisy version of what she broadcasts (see Figure 1a), where the input and output symbols of the channel are from some arbitrary sets. We also assume that all of the honest terminals can discuss over a cost-free noiseless public channel where everybody (including Eve) can hear the discussion (see Figure 1b).

The protocol stated in Definition 1 introduces the most general form of an interactive communication between terminals aiming to share a common secret key K (e.g., see also [12], [14], [16], [13]).

Definition 1 (Secret key generating protocol)**.**
1) For $t = 0$, all of the honest terminals generate independent random variables $Q_0, \ldots, Q_{m-1}$.
2) (i) For time $1 \leq t \leq n$, Alice transmits $X_0[t]$ over

[2]For convenience, sometimes we will refer to legitimate terminals $0, 1, 2, \ldots$, as "Alice," "Bob," "Calvin," and so on. So for example, we use X_0, X_1, X_2, etc. interchangeably with $X_{\mathsf{A}}, X_{\mathsf{B}}, X_{\mathsf{C}}$, etc.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

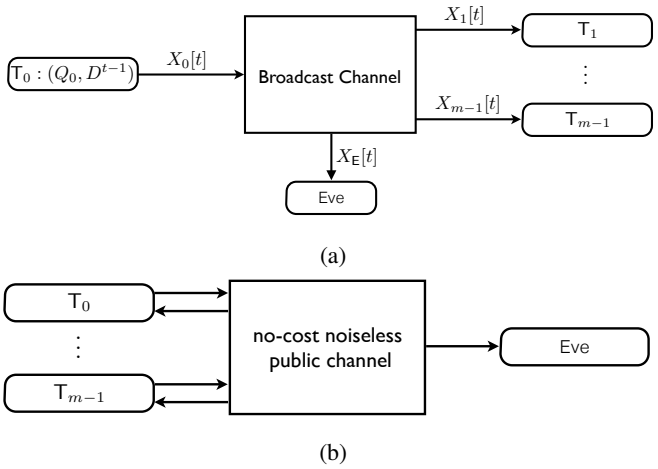

Fig. 1: (a) The broadcast channel that Alice (terminal T_0) has access to. (b) The cost-free noiseless public channel that all of the trusted node can discuss over. Eve overhears all of the public discussions completely.

the broadcast channel. Then the other terminals receive $X_1[t], \ldots, X_{m-1}[t]$, and Eve receives $X_{\mathsf{E}}[t]$.
(ii) Following each of the broadcast transmissions, there is the possibility for the legitimate terminals to discuss over a cost-free noiseless public channel. The discussion continues in a round robin order for an arbitrary number of rounds. The whole public discussion at time t is denoted by $D[t]$. Notice that what Alice broadcasts at time t depends on Q_0 and D^{t-1}.
3) Finally, the ith terminal creates a key K_i where $K_i = K_i(Q_i, X_i^n, D^n)$.

Definition 2. *A number R_s is called an achievable key generation rate if for every $\epsilon > 0$ and sufficiently large n there exists a key generating protocol as defined in Definition 1 such that we have*

$$\mathbb{P}\left[K_i \neq K_j\right] < \epsilon, \quad \forall i, j \in [0 : m-1], \quad i \neq j, \tag{1}$$

$$I(K_0; X_{\mathsf{E}}^n, D^n) < \epsilon, \tag{2}$$

$$\frac{1}{n} H(K_0) > R_s - \epsilon. \tag{3}$$

The supremum of the achievable key rate as $n \to \infty$ and $\epsilon \to 0$ is called the secret key generation (SKG) capacity *C_s.*

Remark 1. *It is worth mentioning that in contrast to [14] where the secret key should be generated by only a subset of terminals $[0 : m-1]$ and the rest of nodes can act as helpers, in the above model all the terminals need to have the shared secret key at the end of the protocol. Moreover, in [14], some of the helpers though participating in the key generation protocol, are wire-tapped by the eavesdropper and the secret key should be kept secret from them as well. Also, they do not consider eavesdroppers who do not take participate in the protocol. However, it should be emphasized that our proposed scheme can be easily modified so that only a subset of nodes share a secret and the rest act as helpers.*

B. State-Dependent Gaussian Broadcast Channels

Here, we introduce the state-dependent additive white Gaussian broadcast channel model which is the main focus of this paper. In this model, we assume that for each receiver the channel state remains unchanged during a block of symbols of length L and changes independently from one block to another block. We also assume L is large enough so that enables us to apply information theoretical arguments within each block. The transmitted vector sent by Alice is denoted by $X_{\mathsf{A}} \in \mathbb{R}^L$. The received vector at each receiver (including Eve) depend on its channel state at a particular time instant. We define a random variable $S_i[t] \in [0 : s]$ corresponding to the channel state for the ith terminal at time t and similarly define the random variable $S_{\mathsf{E}}[t] \in [0 : s]$ for Eve. For the channel state of a receiver $r \in \{1, \ldots, m-1, \mathsf{E}\}$ we assume that[3] $\mathbb{P}\left[S_r[t] = k\right] = \delta_k, \forall k \in [0 : s]$, where $\sum_{k=0}^{s} \delta_k = 1$. The received vector at the receiver r is modelled by a state-dependent white Gaussian channel as follows

$$\hat{X}_r[t] = \sqrt{h_{S_r[t]}} X_{\mathsf{A}}[t] + Z_r[t], \quad \forall r \in \{1, \ldots, m-1, \mathsf{E}\}, \tag{4}$$

where $\hat{X}_r[t] \in \mathbb{R}^L$ and $Z_r[t] \in \mathbb{R}^L$. For the additive noise of each receiver we assume $Z_r[t] \sim N(0, \boldsymbol{I}_L)$ and the noise vectors are also independent over time. The channel gains $\sqrt{h_i}$ are some real constants such that $h_0 < \cdots < h_s$. Additionally, the channel input is subject to an average power constraint $P_{\max}$, i.e., $\frac{1}{L}\mathbb{E}\left[\|X_{\mathsf{A}}\|^2\right] \leq P_{\max}$.

Moreover, we assume that the CSI[4] is completely known by each receiver. So we define a composite received vector for each receiver r as $X_r[t] = (\hat{X}_r[t], S_r[t])$.

C. Deterministic Broadcast Channel

Now, following the idea proposed in [6], we introduce the deterministic approximation model for our Gaussian channel. We assume that the transmitted vector (packet) sent by Alice is denoted by $X_{\mathsf{A}} \in \mathbb{F}_q^L$ where $\mathbb{F}_q$ is a finite field of size q. Then, the received vector at the receiver r is modeled by a state-dependent deterministic broadcast channel as follows

$$\hat{X}_r[t] = \boldsymbol{F}_{S_r[t]} X_{\mathsf{A}}[t], \qquad \forall r \in \{1, \ldots, m-1, \mathsf{E}\}, \tag{5}$$

where $\boldsymbol{F}_i \in \mathbb{F}_q^{L \times L}$ for $i \in [0 : s]$. As defined in Section II-B, $S_r[t]$ denotes the channel state at a receiving node r at time t with a distribution $\mathbb{P}\left[S_r[t] = k\right] = \delta_k, \forall k \in [0 : s]$, where $\sum_{k=0}^{s} \delta_k = 1$. Moreover, similar to the Gaussian model, we define a composite received vector for the receiver r as $X_r[t] = (\hat{X}_r[t], S_r[t])$.

In order to capture and model the different SNR level for the Gaussian channel, we use the shift matrix model developed in [6]. To this end, we consider matrices $\boldsymbol{F}_i$ such that they satisfy the following nested structure

$$\vec{0} = \ker \boldsymbol{F}_s \subset \ker \boldsymbol{F}_{s-1} \subset \cdots \subset \ker \boldsymbol{F}_0 = \mathbb{F}_q^L, \tag{6}$$

$$\operatorname{rank}(\boldsymbol{F}_i - \boldsymbol{F}_{i-1}) = \operatorname{rank}(\boldsymbol{F}_i) - \operatorname{rank}(\boldsymbol{F}_{i-1}). \tag{7}$$

[3]For simplicity of demonstration and without loss of generality, here we only consider a symmetric problem where the probability distribution over the states are the same for all of the receivers (including Eve). Moreover, we focus on a finite number of states. Both of these restrictions can be relaxed.

[4]Channel state information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

For convenience we assume that $\boldsymbol{F}_s = \boldsymbol{I}_L$ where $\boldsymbol{I}_L$ is the identity matrix of size L. The two extreme states "0" and "s" correspond to complete erasure and complete reception of the transmitted vector (packet) X_{A}. The deterministic model is indeed an extension to the packet erasure broadcast channel, studied in [8], [24], which has only two channel states, i.e., $s = 1$, (see also [9, Appendix A]).

D. Discussion on the Cryptographic Analysis

We do not make any assumption on the computational ability of the adversary, i.e., this leads to the unconditional secrecy in the cryptography terminology. Additionally, it is assumed that the channel state is random and not known by any party including the adversary a-priori. The adversary can hear the Alice's broadcast through a fading channel, and also the acknowledgments through a noise-free public discussion channel, i.e., we consider a passive eavesdropper that does not tamper with the communications of legitimate nodes. Furthermore, we assume all the legitimate nodes are authenticated so only these nodes can participate in the public discussion. Our *security model* for the adversary is defined in Definition 2, Eq. (2), i.e., $I(K_0; X_{\mathsf{E}}^n, D^n) < \epsilon$. This definition provides an unconditional secrecy guarantee which is more powerful than the computational secrecy guarantee. The *communication model* is also discussed in Section II.

We have demonstrated the value of some of these ideas in test-bed implementations in [22] and [24].

III. Main Results

The main results of this paper are summarized in the following. For the secret key generation scenario among m terminals that have access to a "deterministic broadcast channel," we completely characterize the key generation capacity. This result can be considered as the generalization of the result of [8], [24] for "packet erasure broadcast channels" (see Theorem 1). For a "state-dependent Gaussian broadcast channel," we provide upper and lower bounds for the key generation capacity and show that these bounds will match in the high-dynamic range, high-SNR regime. Furthermore, the achievable secrecy rate by our proposed scheme for the Gaussian model is described by a non-convex power optimization problem. Although this problem is non-convex, by exploiting its special structure, we find the optimal power allocation that leads to the best secrecy rate achievable by the proposed scheme.

Theorem 1. *The SKG capacity among m terminals that have access to a state-dependent deterministic broadcast channel, defined in Section II-C, is given by*

$$C_s^{\mathsf{det}} = \sum_{i=1}^{s} \left[\operatorname{rank} \boldsymbol{F}_i - \operatorname{rank} \boldsymbol{F}_{i-1}\right] \theta_i (1-\theta_i) \log q,$$

where $\theta_i \triangleq \sum_{j=0}^{i-1} \delta_j$.

Theorem 1 is proved in Section V (see Lemma 2 and Lemma 3). Notice that the result of [8] is a special case of Theorem 1 when $s = 1$.

Theorem 2. *The SKG capacity among m terminals that have access to a state-dependent Gaussian broadcast channel, as defined in Section II-B, is upper bounded by*

$$C_s^{\mathsf{gaus}} \le \frac{1}{2} L \sum_{i=0}^{s} \sum_{j=0}^{s} \delta_i \delta_j \log\left(1 + \frac{h_i P_{\max}}{1 + h_j P_{\max}}\right).$$

Moreover, the secrecy capacity can be lower bounded by the solution of the following (non-convex) optimization problem

$$C_s^{\mathsf{gaus}} \ge \begin{cases} \max & \sum_{i=1}^{s} \Delta_i L R_i \\ \text{subject to} & \sum_{i=1}^{s} P_i = P_{\max} \\ & P_i \ge 0, \quad \forall i \in [1:s], \end{cases}$$

where $\Delta_i \triangleq (1-\theta_i)\theta_i$. Also $\forall i \in [1:s]$ we have

$$R_i \triangleq \frac{1}{2}\left[\log\left(1 + \frac{h_i P_i}{1 + h_i I_i}\right) - \log\left(1 + \frac{h_{i-1} P_i}{1 + h_{i-1} I_i}\right)\right],$$

where $I_i \triangleq \sum_{j=i+1}^{s} P_j$. Additionally, for the high-dynamic range case where $h_i \gg h_{i-1}, \forall i \in [1:s]$, and when we are in high SNR regime, we can write

$$C_s^{\mathsf{gaus}} \doteq \frac{1}{2} L \sum_{i=1}^{s} \Delta_i \log \frac{h_i}{h_{i-1}},$$

where "$\doteq$" defined in Section II, is used to denote for the exponential equality with respect to some scaling parameter Q. Here, as $Q \to \infty$, we asymptotically approach to the high-dynamic, high-SNR regime (for more details refer to Section VII-A).

Theorem 2 is proved in Section VI and Section VII-A. In particular see Lemma 4, Lemma 5, and Lemma 7.

It is worth mentioning that the power optimization problem stated in Theorem 2 is a non-convex problem. Although the closed-form solution of the this problem is not easy to derive explicitly, but by using dynamic programming it can be easily found numerically. In Section VII, based on the structure of this optimization problem and by exploiting special properties of its KKT necessary conditions for the optimality, we propose a dynamic programming algorithm that finds the optimal power allocation (see Algorithms 1 and 2). More specifically, we have the following theorem.

Theorem 3. *Algorithms 1 and 2 find the optimal solution of the optimization problem stated in Theorem 2.*

The whole Section VII is devoted to proving Theorem 3.

IV. Upper Bound for the Key Generation Capacity of Independent Broadcast Channels

The secret key generation capacity among multiple terminals (without eavesdropper having access to the broadcast channel) is completely characterized in [14]. By using this result, it is possible to state an upper bound for the secrecy capacity of the key generation problem among multiple terminals where the eavesdropper has also access to the broadcast channel. This can be done by adding a dummy terminal to the first problem and giving all the eavesdropper's information to this dummy node and let it to participate in the key generation

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

protocol (we refer the interested readers to [14, Section V] for more details). By doing so, the secret key generation rate does not decrease. Hence by combining [14, Theorem 4.1] and [14, Lemma 5.1], the following result can be stated.

Lemma 1. *The secret key generation capacity among m terminals as defined in Definition 2, is upper bounded as follows*

$$C_s \le \max_{P_{X_0}} \min_{\lambda\in\Lambda([0:m-1])} \Bigg[H(X_{[0:m-1]}|X_{\mathsf{E}}) - \sum_{B\subsetneq[0:m-1]} \lambda_B H(X_B|X_{B^c}, X_{\mathsf{E}}) \Bigg],$$

where $\Lambda([0:m-1])$ is the set of all collections $\lambda = \{\lambda_B : B \subsetneq [0:m-1], B\neq\varnothing\}$ of weights $0 \le \lambda_B \le 1$, satisfying

$$\sum_{B\subsetneq[0:m-1],\ i\in B} \lambda_B = 1, \qquad \forall i \in [0:m-1]. \tag{8}$$

Note that in the above expression for the upper bound, it is possible to change the order of maximization and minimization, see [14, Theorem 4.1].

Remark 2. *The upper bound stated in Lemma 1 is not the best known upper bound for the secret key sharing capacity of the multi-terminal secret key sharing problem (see also [15], [16] for alternative improved bounds). However in this work, we use Lemma 1 to derive an upper bound for our problem. This bound is good enough that in addition to the proposed achievability scheme, completely characterize the secret key sharing capacity for the "state-dependent deterministic channels" scenario.*

Now, back to our problem where the channel from Alice to the other terminals are assumed to be independent, we can further simplify the upper bound given in Lemma 1, as stated in Corollary 1.

Corollary 1 ([8],[9])**.** *If the channels from Alice to the other terminals are independent, then the upper bound stated in Lemma 1 for the SKG capacity is simplified to*

$$C_s \le \max_{P_{X_0}} \min_{j\in[1:m-1]} I(X_0; X_j|X_{\mathsf{E}}) \tag{9}$$

$$\le \min_{j\in[1:m-1]} \max_{P_{X_0}} I(X_0; X_j|X_{\mathsf{E}}). \tag{10}$$

Remark 3. *Using [12, Theorem 7] or [13, Theorem 2], we observe that the bound given in* (10) *is indeed tight for the two terminals problem where we have the Markov chains $X_{\mathsf{B}} \leftrightarrow X_{\mathsf{A}} \leftrightarrow X_{\mathsf{E}}$, i.e., when the channels are independent or $X_{\mathsf{A}} \leftrightarrow X_{\mathsf{B}} \leftrightarrow X_{\mathsf{E}}$, i.e., when the channels are degraded. In Section V, we will further show that the above upper bound is also tight for the stated-dependent deterministic broadcast channels.*

V. Group Secret Key Agreement over Deterministic Broadcast Channels

In this section, we prove Theorem 1 that characterizes the secret key generation capacity for a deterministic broadcast channel defined in Section II-C. The proof of this theorem, as an underlying machinery, uses the achievability technique for the packet erasure broadcast channel that is appeared in [8], [24] (for more details, also see [9, Appendix A]).

A. Upper Bound for the Key Generation Capacity

Using Corollary 1, the SKG capacity C_s^{det} for the independent broadcast channel can be upper bounded by (10). Then we can state the following result, Lemma 2.

Lemma 2. *The SKG capacity of the deterministic broadcast channel, introduced in Section II-C, is upper bounded by $C_s^{\mathsf{det}} \le \sum_{i=1}^{s} [\operatorname{rank} \boldsymbol{F}_i - \operatorname{rank} \boldsymbol{F}_{i-1}]\, \theta_i(1-\theta_i)\log q$, where $\theta_i = \sum_{j=0}^{i-1}\delta_j$.*

Proof. From (10) and because of the symmetry of the problem, we have $C_s^{\mathsf{det}} \le \max_{P_{X_{\mathsf{A}}}} I(X_{\mathsf{A}}; X_{\mathsf{B}}|X_{\mathsf{E}})$ where we use "A" and "B" to denote for terminal 0 and terminal 1. Then, we can write

$$\begin{aligned} H(X_{\mathsf{A}}|X_{\mathsf{E}}) &= \sum_{i=0}^{s-1} \delta_i \Big[H(X_{\mathsf{A}}|\hat{X}_{\mathsf{E}}, S_{\mathsf{E}}=i) \Big] \\ &= \sum_{i=0}^{s-1} \delta_i \Big[H(X_{\mathsf{A}}, \hat{X}_{\mathsf{E}}|S_{\mathsf{E}}=i) - H(\hat{X}_{\mathsf{E}}|S_{\mathsf{E}}=i) \Big] \\ &= \sum_{i=0}^{s-1} \delta_i \left[H(X_{\mathsf{A}}) - H(\boldsymbol{F}_i X_{\mathsf{A}}) \right], \end{aligned}$$

and similarly

$$H(X_{\mathsf{A}}|X_{\mathsf{B}}, X_{\mathsf{E}}) = \sum_{i=0}^{s-1} \kappa_i \left[H(X_{\mathsf{A}}) - H(\boldsymbol{F}_i X_{\mathsf{A}}) \right],$$

where $\kappa_i \triangleq 2\delta_i(\delta_0 + \cdots + \delta_{i-1})\mathbb{1}_{\{i>0\}} + \delta_i^2$. Thus, we have $I(X_{\mathsf{A}}; X_{\mathsf{B}}|X_{\mathsf{E}}) = \sum_{i=0}^{s-1} \rho_i \left[H(X_{\mathsf{A}}) - H(\boldsymbol{F}_i X_{\mathsf{A}}) \right]$ where $\rho_i \triangleq \delta_i - \kappa_i$. Now, by observing that $H(\boldsymbol{F}_i X_{\mathsf{A}}) = H(\boldsymbol{F}_i X_{\mathsf{A}}, \boldsymbol{F}_{i-1} X_{\mathsf{A}})$ and applying the chain rule recursively, we get $H(\boldsymbol{F}_i X_{\mathsf{A}}) = \sum_{j=1}^{i} H(\boldsymbol{F}_j X_{\mathsf{A}}|\boldsymbol{F}_{j-1} X_{\mathsf{A}})$. So $I(X_{\mathsf{A}}; X_{\mathsf{B}}|X_{\mathsf{E}})$ can be expanded as follows

$$\begin{aligned} I(X_{\mathsf{A}}; X_{\mathsf{B}}|X_{\mathsf{E}}) &= \sum_{i=0}^{s-1} \rho_i [H(X_{\mathsf{A}}) - H(\boldsymbol{F}_i X_{\mathsf{A}})] \\ &= \sum_{j=1}^{s} H(\boldsymbol{F}_j X_{\mathsf{A}}|\boldsymbol{F}_{j-1} X_{\mathsf{A}}) \sum_{i=0}^{j-1} \rho_i. \end{aligned}$$

Hence we can upper bound C_s^{det} as follows

$$\begin{aligned} C_s^{\mathsf{det}} &\le \max_{P_{X_{\mathsf{A}}}} \sum_{j=1}^{s} H(\boldsymbol{F}_j X_{\mathsf{A}}|\boldsymbol{F}_{j-1} X_{\mathsf{A}}) \sum_{i=0}^{j-1} \rho_i \\ &= \max_{P_{X_{\mathsf{A}}}} \sum_{j=1}^{s} H\left([\boldsymbol{F}_j - \boldsymbol{F}_{j-1}] X_{\mathsf{A}}|\boldsymbol{F}_{j-1} X_{\mathsf{A}} \right) \sum_{i=0}^{j-1} \rho_i \\ &\overset{(a)}{\le} \max_{P_{X_{\mathsf{A}}}} \sum_{j=1}^{s} H\left([\boldsymbol{F}_j - \boldsymbol{F}_{j-1}] X_{\mathsf{A}} \right) \sum_{i=0}^{j-1} \rho_i \\ &\overset{(b)}{=} \sum_{j=1}^{s} \operatorname{rank}\left(\boldsymbol{F}_j - \boldsymbol{F}_{j-1} \right) \left(\sum_{i=0}^{j-1} \rho_i \right) \log q \end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

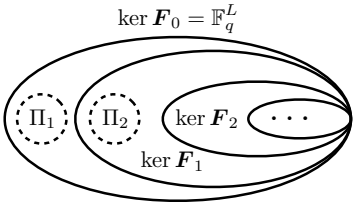

Fig. 2: Demonstration of Proposition 1. Here, we have $\vec{0} = \ker \boldsymbol{F}_s \subset \ker \boldsymbol{F}_{s-1} \subset \cdots \subset \ker \boldsymbol{F}_0 = \mathbb{F}_q^L$ and $\Pi_1, \ldots, \Pi_s$ satisfy (12).

$$\overset{\text{(c)}}{=} \sum_{j=1}^{s} [\operatorname{rank} \boldsymbol{F}_j - \operatorname{rank} \boldsymbol{F}_{j-1}] \left(\sum_{i=0}^{j-1} \rho_i \right) \log q, \quad (11)$$

where (a) is true because conditioning reduces the entropy, (b) is true because uniform distribution on X_{A} achieves the maximum values for all the entropies in the summation, and finally (c) is true because of the assumption we have made in (7). Also, note that $\sum_{i=0}^{j-1} \rho_i = \theta_j(1-\theta_j) \geq 0$, where $\theta_j \triangleq \sum_{i=0}^{j-1} \delta_i$. This completes the proof. □

B. Lower Bound for the Key Generation Capacity

In this section, we will present a scheme that achieves the same secret key generation rate as we derived in the upper bound stated in Lemma 2. But before that, let us state the following proposition.

Proposition 1. *Suppose s subspaces $\ker \boldsymbol{F}_i \subseteq \mathbb{F}_q^L$ satisfy the nested condition* (6)*, i.e., $\vec{0} = \ker \boldsymbol{F}_s \subset \ker \boldsymbol{F}_{s-1} \subset \cdots \subset \ker \boldsymbol{F}_0 = \mathbb{F}_q^L$. Then it is possible to find subspaces $\Pi_1, \ldots, \Pi_s$, such that $\cap_{i \in \mathcal{V}} \Pi_i = \vec{0}$ for all $\mathcal{V} \subseteq [1:s]$ where $|\mathcal{V}| \geq 2$ and they also satisfy*

$$\begin{aligned} \Pi_1 \oplus \ker \boldsymbol{F}_1 &= \mathbb{F}_q^L, \\ \Pi_2 \oplus \Pi_1 \oplus \ker \boldsymbol{F}_2 &= \mathbb{F}_q^L, \\ &\vdots \\ \Pi_s \oplus \cdots \oplus \Pi_1 \oplus \ker \boldsymbol{F}_s &= \mathbb{F}_q^L \end{aligned} \quad (12)$$

where "$\oplus$" is the direct sum of two disjoint subspaces. Moreover for $i \in [1:s]$ we have $\dim \Pi_i = \operatorname{rank} \boldsymbol{F}_i - \operatorname{rank} \boldsymbol{F}_{i-1}$. For more clarification, Figure 2 demonstrates the proposition.

In our proposed achievability scheme, Alice uses superposition coding where she creates a vector

$$X_{\mathsf{A}}[t] = X_{\mathsf{A},1}[t] + \cdots + X_{\mathsf{A},s}[t], \quad (13)$$

such that $X_{\mathsf{A},i}[t] \in \Pi_i$. Because of (12), $\{\Pi_1, \ldots, \Pi_s\}$ is a set of disjoint sub-spaces that span the whole space $\mathbb{F}_q^L$. So every vector $X_{\mathsf{A}}[t] \in \mathbb{F}_q^L$ can be uniquely decomposed as (13). Now each $X_{\mathsf{A},i}[t] \in \Pi_i$ can be considered as a vector that is transmitted by Alice and will be received independently by each trusted terminal or Eve with erasure probability $\theta_i = \sum_{j=0}^{i-1} \delta_i$. Note that the vector $X_{\mathsf{A},i}[t]$ is correctly received by the rth receiver only if $S_r \geq i$.

So we may view the broadcast channel from Alice to the rest of terminals as s independent *packet erasure channels*; where Π_i is the set of messages transmitted over the ith channel (layer) and the erasure probability of the ith channel is θ_i.

Then we can proceed as follows. On the kth layer, we run independently the scheme propose in [8], [24] for the secret key sharing problem over an erasure broadcast channel (see also [9, Appendix A] for more details). Then, we can state the following result.

Lemma 3. *The achievable SKG rate of the above scheme for each layer k is given by $R_k^{\mathsf{det}} = (1-\theta_k)\theta_k \dim(\Pi_k) \log q$. So for the total achievable secrecy rate we have*

$$\begin{aligned} R_s^{\mathsf{det}} &= \sum_{i=1}^{s} \theta_i(1-\theta_i) \dim(\Pi_i) \log q \\ &= \sum_{i=1}^{s} [\operatorname{rank} \boldsymbol{F}_i - \operatorname{rank} \boldsymbol{F}_{i-1}] \, \theta_i(1-\theta_i) \log q. \end{aligned}$$

Observe that this matches the upper bound stated in Lemma 2, and therefore yields a characterization of the group key-agreement rate for deterministic channels, i.e., this completes the proof of Theorem 1.

Remark 4. *This result can be easily extended to the asymmetric case where the channels to the legitimate users are not statistically identical (but still independent). Moreover, notice that the key-generation rate is the same for any $m \geq 2$. This is in fact similar to the erasure channel case [8] (see also [9, Appendix A]), where the critical difference between $m = 2$ and $m > 2$ is that the key-reconciliation necessitated the use of ideas from* Network Coding.

VI. Group Secret Key Agreement over State-dependent Gaussian Broadcast Channels

In this section, by using the results derived in the previous sections, we will study the secret key generation capacity among multiple terminals having access to a state-dependent Gaussian broadcast channel. We will derive upper and lower bounds for the secret key generation capacity. Although, the proposed bounds are not matched in general, we will show that they will match in the high-dynamic range, high-SNR regime in a degree of freedom sense.

A. Upper Bound for the Key Generation Capacity

In order to upper bound the secrecy capacity for the Gaussian broadcast channel, we cannot apply the result of Corollary 1 directly because this result has been derived under the assumption that the transmitted and received symbols are discreet. However, the work in [18] has extended the results of [14] for continuous channels. So by using [18, Theorem 3.2], we can write an upper bound for the secrecy capacity similar to Lemma 1 with the addition of a power constraint over the transmitted symbols. Then we can state the following result, as stated in Lemma 4.

Lemma 4. *The key generation capacity of the Gaussian broadcast channel given in* (4) *using public discussions is upper bounded as follows*

$$C_s^{\mathsf{gaus}} \leq \frac{1}{2} L \sum_{i=0}^{s} \sum_{j=0}^{s} \delta_i \delta_j \log \left(1 + \frac{h_i P_{\max}}{1 + h_j P_{\max}} \right). \quad (14)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

Proof. Using [18, Theorem 3.2] and by proceeding similar steps to the proof of Corollary 1 (see [8] and [9]), we can write

$$C_s^{\mathsf{gaus}} \leq \min_{j\in[1:m-1]} \sup_{\substack{P_{X_0}:\\ \mathbb{E}\left[\|X_0\|^2\right]\leq LP_{\max}}} I(X_0;X_j|X_{\mathsf{E}}) \overset{\text{(a)}}{=} \sup_{\substack{P_{X_{\mathsf{A}}}:\\ \mathbb{E}\left[\|X_{\mathsf{A}}\|^2\right]\leq LP_{\max}}} I(X_{\mathsf{A}};X_{\mathsf{B}}|X_{\mathsf{E}}),$$

where (a) is true because of the symmetry. Hence, there exists an input distribution $P_{X_{\mathsf{A}}}$ such that $\mathbb{E}\left[\|X_{\mathsf{A}}\|^2\right] \leq LP_{\max}$ where the secrecy capacity is upper bounded as follows

$$\begin{aligned}
C_s^{\mathsf{gaus}} &\leq I(X_{\mathsf{A}};X_{\mathsf{B}}|X_{\mathsf{E}})\\
&= I(X_{\mathsf{A}};\hat{X}_{\mathsf{B}},S_{\mathsf{B}}|\hat{X}_{\mathsf{E}},S_{\mathsf{E}})\\
&= h(\hat{X}_{\mathsf{B}},S_{\mathsf{B}}|\hat{X}_{\mathsf{E}},S_{\mathsf{E}}) - h(\hat{X}_{\mathsf{B}},S_{\mathsf{B}}|\hat{X}_{\mathsf{E}},S_{\mathsf{E}},X_{\mathsf{A}})\\
&\overset{\text{(a)}}{=} h(\hat{X}_{\mathsf{B}},S_{\mathsf{B}}|\hat{X}_{\mathsf{E}},S_{\mathsf{E}}) - h(\hat{X}_{\mathsf{B}},S_{\mathsf{B}}|X_{\mathsf{A}})\\
&= h(\hat{X}_{\mathsf{B}},S_{\mathsf{B}}|\hat{X}_{\mathsf{E}},S_{\mathsf{E}}) - h(S_{\mathsf{B}}|X_{\mathsf{A}}) - h(\hat{X}_{\mathsf{B}}|S_{\mathsf{B}},X_{\mathsf{A}})\\
&\overset{\text{(b)}}{=} h(\hat{X}_{\mathsf{B}},S_{\mathsf{B}}|\hat{X}_{\mathsf{E}},S_{\mathsf{E}}) - h(S_{\mathsf{B}}) - h(Z_{\mathsf{B}})\\
&= h(\hat{X}_{\mathsf{B}},\hat{X}_{\mathsf{E}}|S_{\mathsf{E}},S_{\mathsf{B}}) + h(S_{\mathsf{E}},S_{\mathsf{B}})\\
&\quad - h(\hat{X}_{\mathsf{E}},S_{\mathsf{E}}) - h(S_{\mathsf{B}}) - h(Z_{\mathsf{B}})\\
&= h(\hat{X}_{\mathsf{B}},\hat{X}_{\mathsf{E}}|S_{\mathsf{E}},S_{\mathsf{B}}) - h(\hat{X}_{\mathsf{E}}|S_{\mathsf{E}}) - h(Z_{\mathsf{B}}).
\end{aligned}$$

where (a) is true since we have the Markov chain $X_{\mathsf{B}} \leftrightarrow X_{\mathsf{A}} \leftrightarrow X_{\mathsf{E}}$ and (b) follows from the fact that the state variables are independent of X_{A} and given X_{A} and S_{B} the only uncertainty left in $\hat{X}_{\mathsf{B}}$ is that of noise Z_{B}. Now the above relation can be more simplified as follows

$$\begin{aligned}
C_s^{\mathsf{gaus}} &\leq \sum_{i=0}^{s}\sum_{j=0}^{s}\delta_i\delta_j h(\hat{X}_{\mathsf{B}},\hat{X}_{\mathsf{E}}|S_{\mathsf{E}}=j,S_{\mathsf{B}}=i)\\
&\quad - \sum_{k=0}^{s}\delta_k h(\hat{X}_{\mathsf{E}}|S_{\mathsf{E}}=k) - h(Z_{\mathsf{B}})\\
&= \sum_{i=0}^{s}\sum_{j=0}^{s}\delta_i\delta_j h(\sqrt{h_i}X_{\mathsf{A}}+Z_{\mathsf{B}},\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}})\\
&\quad - \sum_{k=0}^{s}\delta_k h(\sqrt{h_k}X_{\mathsf{A}}+Z_{\mathsf{E}}) - h(Z_{\mathsf{B}})\\
&= \sum_{i=0}^{s}\sum_{j=0}^{s}\delta_i\delta_j h(\sqrt{h_i}X_{\mathsf{A}}+Z_{\mathsf{B}}|\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}}) - h(Z_{\mathsf{B}})\\
&\overset{\text{(a)}}{\leq} \sum_{i=0}^{s}\sum_{j=0}^{s}\frac{\delta_i\delta_j}{2}\log\Big[(2\pi e)^L\times\\
&\det\Big(\operatorname{cov}(\sqrt{h_i}X_{\mathsf{A}}+Z_{\mathsf{B}}|\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}})\Big)\Big] - h(Z_{\mathsf{B}}), \quad (15)
\end{aligned}$$

where (a) follows from the fact that for a fixed variance, Gaussian distribution maximizes the entropy.

The inequality (a) in (15) is achieved when $(\sqrt{h_i}X_{\mathsf{A}} + Z_{\mathsf{B}}|\sqrt{h_j}X_{\mathsf{A}} + Z_{\mathsf{E}})$ has a Gaussian distribution. A sufficient condition for this to be satisfied is when X_{A}, Z_{B}, and Z_{E} are Gaussian and independent, namely, $X_{\mathsf{A}} \sim N(\vec{0}, P_{\max}\boldsymbol{I}_L)$, $Z_{\mathsf{B}} \sim N(\vec{0}, \boldsymbol{I}_L)$, and $Z_{\mathsf{E}} \sim N(\vec{0}, \boldsymbol{I}_L)$. This observation makes the calculation of

$$\frac{1}{2}\log\Big[(2\pi e)^L\det\Big(\operatorname{cov}(\sqrt{h_i}X_{\mathsf{A}}+Z_{\mathsf{B}}|\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}})\Big)\Big]$$

much easier as it is equivalent to the evaluation of $h(\sqrt{h_i}X_{\mathsf{A}}+Z_{\mathsf{B}},\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}}) - h(\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}})$ when X_{A}, Z_{B}, and Z_{E} are Gaussian and independent as shown below,

$$\begin{aligned}
&\frac{1}{2}\log\Big[(2\pi e)^L\det\Big(\operatorname{cov}(\sqrt{h_i}X_{\mathsf{A}}+Z_{\mathsf{B}}|\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}})\Big)\Big] =\\
&= h\big(\sqrt{h_i}X_{\mathsf{A}}+Z_{\mathsf{B}},\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}}\big) - h\big(\sqrt{h_j}X_{\mathsf{A}}+Z_{\mathsf{E}}\big)\\
&= \sum_{k=1}^{L} h\big(\sqrt{h_i}X_{\mathsf{A},k}+Z_{\mathsf{B},k},\sqrt{h_j}X_{\mathsf{A},k}+Z_{\mathsf{E},k}\big)\\
&\quad - h\big(\sqrt{h_j}X_{\mathsf{A},k}+Z_{\mathsf{E},k}\big)\\
&= \frac{L}{2}\Big[\log\big((2\pi e)^2(1+h_iP_{\max}+h_jP_{\max})\big)\\
&\quad - \log\big(2\pi e(1+h_jP_{\max})\big)\Big],
\end{aligned}$$

where $\mathbb{E}\left[X_{\mathsf{A},k}^2\right] = P_{\max}$ and $\mathbb{E}\left[Z_{\mathsf{B},k}^2\right] = \mathbb{E}\left[Z_{\mathsf{E},k}^2\right] = 1$ for all $k \in [1:L]$.

Hence, the upper bound on the secrecy capacity reads as follows

$$\begin{aligned}
C_s^{\mathsf{gaus}} &\leq \sum_{i=0}^{s}\sum_{j=0}^{s}\frac{\delta_i\delta_j L}{2}\log\left[(2\pi e)^2(1+h_iP_{\max}+h_jP_{\max})\right]\\
&\quad - \sum_{i=0}^{s}\sum_{j=0}^{s}\frac{\delta_i\delta_j L}{2}\log\left[2\pi e(1+h_jP_{\max})\right] - \frac{L}{2}\log(2\pi e)\\
&= \frac{1}{2}L\sum_{i=0}^{s}\sum_{j=0}^{s}\delta_i\delta_j\log\left(1+\frac{h_iP_{\max}}{1+h_jP_{\max}}\right),
\end{aligned}$$

where we are done. □

B. Lower Bound for the Key Generation Capacity

Before stating our achievability scheme, let us first define a "nested message set, degraded channel" wiretap scenario.

Definition 3. *Assume a wiretap channel scenario where there is a transmitter called Alice who broadcasts X_{A} and there are $s+1$ receivers Y_i where the ith receiver receives Y_i according to the broadcast channel $(\mathcal{X}_{\mathsf{A}}, p(y_0,\ldots,y_s|x), \mathcal{Y}_0\times\cdots\times\mathcal{Y}_s))$ such that*

$$p(y_0,\ldots,y_s|x_{\mathsf{A}}) = p(y_s|x_{\mathsf{A}})\cdot p(y_{s-1}|y_s)\cdots p(y_0|y_1).$$

Suppose that Alice has s messages $W_1,\ldots,W_s$ where $W_i \in \{1,\ldots,2^{LR_i}\}$ and $W_i \sim \mathsf{Uni}([1:2^{LR_i}])$. The goal is that she wants to broadcast these messages such that $\forall i$:
(i) *each message W_i should be decodable by the receivers $Y_i,\ldots,Y_s$ with a negligible error probability, and*
(ii) *all the receivers $Y_0,\ldots,Y_{i-1}$ should be ignorant about the message W_i, namely for the leakage rate we should have*

$$R_{\text{leak},i}^{(L)} \triangleq \frac{1}{L}I(W_{i+1},\ldots,W_s;Y_i^{1:L}) \leq \epsilon_L, \forall i\in[0:s]. \quad (16)$$

Now suppose that a multi-receiver wiretap scenario as defined in Definition 3 consists of $s+1$ independent Gaussian channels where the rth channel is defined as follows

$$Y_r[t] = \sqrt{h_r}X[t] + Z_r[t], \qquad \forall r\in[0:s], \quad (17)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

where $Z_r[t] \sim N(0,1)$ and h_r are some fixed constant representing the channel gains such that $h_0 < \cdots < h_s$. We also assume that the channel input is subject to an average power constraint $P_{\max}$, i.e., $\frac{1}{L}\sum_{t=1}^{L}\mathbb{E}\left[X^2[t]\right] \le P_{\max}$. Then we can state the following result.

Lemma 5. *Using a properly designed layered wiretap code similar to [7], [10], we can achieve the following set of rates for the "nested message set, degraded Gaussian wiretap channel,"*

$$R_i = \frac{1}{2}\left[\log\left(1+\frac{h_i P_i}{1+h_i I_i}\right) - \log\left(1+\frac{h_{i-1}P_i}{1+h_{i-1}I_i}\right)\right], \tag{18}$$

$\forall i \in [1:s]$, *where* $I_i \triangleq \sum_{j=i+1}^{s} P_j$.

Proof. In the following, we will describe a code construction that achieves the rates stated in Lemma 5. Because it is very similar to [7] and also due to space limit, we only present a sketch of the proof for the theorem.

Assume that the code has s layers that correspond to each channel where they are indexed from 1 up to s (the channel Y_0 should decode nothing). To each layer a power constraint P_i is assigned such that $\sum_{i=1}^{s} P_i \le P_{\max}$. The transmitter uses superposition coding to encode each message W_i that corresponds to layer i; namely, it broadcasts

$$X[t] = \sum_{i=1}^{s} X_i[t],$$

over the channel described by (17). Then by receiving Y_r, the rth receiver uses successive decoding, that starts from the layer 1 to decode X_1 assuming the rest of the layers as noise and subtracting X_1 from the received vector after decoding. Then it continues this process to decode the rest of layers.

More precisely we construct s codebooks $\hat{\mathcal{C}}_i(2^{L\hat{R}_i}, L)$ each contains $2^{L\hat{R}_i}$ codewords X_i^L of length L by choosing in total $L2^{L\hat{R}_i}$ symbols independently from the Gaussian distribution $N(0,P_i)$ where

$$\hat{R}_i = \frac{1}{2}\log\left(1+\frac{h_i P_i}{1+h_i I_i}\right),$$

and $I_i = \sum_{j=i+1}^{s} P_j$. Each codebook $\hat{\mathcal{C}}_i$, $0 < i \le s$, is divided into 2^{LR_i} bins where

$$R_i = \frac{1}{2}\left[\log\left(1+\frac{h_i P_i}{1+h_i I_i}\right) - \log\left(1+\frac{h_{i-1}P_i}{1+h_{i-1}I_i}\right)\right].$$

At every layer i, each message is mapped into one bin, and one codeword in the bin is randomly chosen. So, layer i can transmits 2^{LR_i} messages. Following a similar argument as stated in [7], [10], it can be shown that the above codebook satisfies the requirement of Definition 3. □

Remark 5. *Note that all of the above discussions are also valid for complex channels. The only difference is that there will be no* $\frac{1}{2}$ *coefficient before rates given by* (18) *and other expressions should be updated accordingly.*

Now, as described in the proof of Lemma 5, by using a properly designed layered coding for the nested message set, degraded channel wiretap scenario, we can convert the Gaussian channel given in (4) to a set of s independent erasure channels where the erasure of the messages for each channel (layer) depends on the receiver channel state. In fact using the layered coding scheme for the wiretap channel, we mimic the orthogonality behaviour that we have for the deterministic channel as described by (6) and (12).

To be more specific, we assume that Alice broadcasts an L-length vector

$$X_{\mathsf{A}}[t] = \sum_{i=1}^{s} X_{\mathsf{A},i}[t],$$

where she maps W_i (the messages corresponding to the ith layer) to $X_{\mathsf{A},i}[t]$ according to the codebook described in the proof of Lemma 5. From the proof we know that the receiver r which observes the channel state $S_r = i$ can decode messages up to layer i and is ignorant about messages of layers above i. So, equivalently, we can say that the message W_i experiences erasure probability $\theta_i = \sum_{j=0}^{i-1}\delta_j$, when it passes through the channel (4).

Now for each layer i, we run the interactive secret key sharing scheme introduced in [8], [24] (also, see [9, Appendix A]) where Alice broadcasts an n-length sequence of random messages, i.e., W_i^n. Then, by discussing over the public channel, the trusted terminals reconcile their secret messages to build a common key. The key generation rate for each layer is $\Delta_i L R_i$, so for a fixed power allocation we achieve the following secrecy rate

$$R_s^{\mathsf{gaus}} \le \sum_{i=1}^{s} \Delta_i L R_i,$$

where R_i is defined in (18) and $\Delta_i = (1-\theta_i)\theta_i$.

The maximum secrecy rate is obtained by optimizing the above rate over the power allocations $\{P_i\}_{i=1}^{s}$. Thus we can write

$$R_s^{\mathsf{gaus}} = \begin{cases} \max & \sum_{i=1}^{s}\Delta_i L R_i \\ \text{subject to} & \sum_{i=1}^{s} P_i \le P_{\max} \\ & P_i \ge 0, \quad \forall i \in [1:s]. \end{cases} \tag{19}$$

Because R_1 is an increasing function of P_1 when other P_i are kept fixed and R_i does not depend on P_1 for $i > 1$ we can write the power constant inequality as an equality. We also apply a change of variables to Problem 19 from $\{P_i\}$ to $\{I_k\}$. So we can rewrite (19) in the canonical form (see [25]) as follows

$$R_s^{\mathsf{gaus}} = \begin{cases} \min & -\sum_{i=1}^{s}\Delta_i L R_i \\ \text{subject to} & -[I_{k-1} - I_k] \le 0, \ \forall k \in [1:s], \end{cases} \tag{20}$$

where for convenience we define $I_0 \triangleq P_{\max}$, $I_s \triangleq 0$, and we have also

$$R_i = \frac{1}{2}\log\left(\frac{1+h_i I_{i-1}}{1+h_i I_i}\cdot\frac{1+h_{i-1}I_i}{1+h_{i-1}I_{i-1}}\right).$$

In Section VII, we will focus on solving the optimization problem (20).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

C. Discussion about the Complexity of the Proposed Scheme

In this section, we briefly discuss about the complexity of the proposed secret key sharing algorithm. In more detail, the algorithm consists of two main parts. First, as discussed in Section VI-B, by applying a multi-layer wiretap code [7], [10], a Gaussian broadcast channel is converted to a number of different message-level erasure broadcast channels. Then in the second part, for each layer, assuming a message-level broadcast erasure channel, legitimate terminals create a shared secret key among themselves where this scheme presented in [8], [9], [22], [24]. Hence, we can break up the complexity analysis of the proposed scheme into two parts.

For the complexity analysis of the second part, referring to [8], [9], [22], [24], we can easily observe that each of the legitimate nodes needs to perform $O(n^2L)$ operations where n is the number of packets and L is the packet length. Additionally, the total communication complexity, i.e., number of transmitted bits, of this algorithm is $O(n(n+L))$ bits. For the first part of the algorithm one needs to consider a practical implementation of wiretap codes. For example we can use the result of [26] where constructs such a wiretap code. The encoding and decoding complexity of the proposed wiretap code is linear in the packet length (i.e., L in our setup) since it is an LDPC code. So, this part of the scheme can at most add $O(nL)$ operations to the complexity of each node. Hence, to summarize, the total computation complexity of all nodes remain the same as each of them needs $O(n^2L)$ operations. Moreover, the total communication complexity of the algorithm $O(n(n+L))$ bits.

VII. Solving the Non-convex Power Allocation Problem

Here, we present how the optimization problem (20) can be solved. Our final result is not in closed form but instead we propose a recursive algorithm (i.e., a dynamic program) that finds all the possible solutions of KKT[5] conditions (which provide necessary conditions for an optimal solution to the optimization problem (20)) and find an optimum solution by searching among them. By using the proposed algorithm, we reduce the search space of the optimization problem (20) from a multi-dimensional continuous space to a finite elements set; i.e., the set of solutions to the KKT conditions. In this sense, the final result is exact (the proposed algorithm is not a numerical approximation), but it is hard to describe the solution in a single closed form equation for all possible parameters involved in the secrecy problem (e.g., channel gains, probability distribution over states, etc.). However, note that for each set of given problem parameters, it is possible to state the final solution in terms of these given parameters (but here we only focus on deriving the final "value" of the solution, not its "expression"). To find the optimal solutions of the above-mentioned optimization problem, we proceed as follows.

Because the constraints of optimization problem (20) are affine, we can use the KKT conditions to derive a set of necessary conditions for the optimum power allocation (e.g., see [27, Chapter 5]). By defining the Lagrangian $\mathfrak{L}$ as

$$\mathfrak{L}(P_1,\ldots,P_s,\lambda_1,\ldots,\lambda_s) = -\sum_{i=1}^{s}\Delta_i LR_i + \sum_{i=1}^{s}\lambda_i[I_i - I_{i-1}],$$

and applying the KKT theorem (e.g., see [25, Chapter 5]), we write a set of necessary conditions for the optimal solution of (20) as follows

$$\begin{cases} \frac{\partial \mathfrak{L}}{\partial I_k} = 0, & \forall k \in [1:s-1], \\ \lambda_k[I_k - I_{k-1}] = 0, & \forall k \in [1:s], \\ I_k \le I_{k-1}, & \forall k \in [1:s], \\ \lambda_k \ge 0, & \forall k \in [1:s]. \end{cases} \tag{21}$$

By taking the derivative of $\mathfrak{L}$ with respect to I_k, $\forall k \in [1:s-1]$, and doing some algebra we get

$$0 = \frac{\partial \mathfrak{L}}{\partial I_k} = \frac{(h_{k+1}\beta_k - h_{k-1}\alpha_k)I_k - (\alpha_k - \beta_k)}{(\ln 2)(1+h_{k-1}I_k)(1+h_kI_k)(1+h_{k+1}I_k)} + (\lambda_k - \lambda_{k+1}) = F_k^{(1)}(I_k) + (\lambda_k - \lambda_{k+1}), \tag{22}$$

where $\alpha_k \triangleq (h_{k+1}-h_k)\Delta_{k+1}$, $\beta_k \triangleq (h_k - h_{k-1})\Delta_k$ and $F_k^{(1)}(I_k)$ is defined accordingly. Notice that because I_k's are positive variables the denominator of $F_k^{(1)}(I_k)$ is strictly positive. For the ease of reference, some of the important variables of our problem are gathered in Table I.

Variable	Definition
θ_k $(\forall k \in [1:s])$	The effective erasure probability that the message W_k (message of kth layer) experience which is $\sum_{i=0}^{k-1}\delta_i$.
Δ_k $(\forall k \in [1:s])$	A dummy variable which is $\theta_k(1-\theta_k)$.
I_k $(\forall k \in [0:s])$	The contribution of the interference of all layers above k to the decoding of the kth layer which is $\sum_{i=k+1}^{s}P_i$. For convenience we define $I_0 = P_{\max}$ and $I_s = 0$.
h_k $(\forall k \in [0:s])$	The square of channel gains. Remember that $h_0 < \cdots < h_s$.
λ_k $(\forall k \in [1:s])$	The Lagrangian multipliers of optimization problem (20).
α_k $(\forall k \in [1:s-1])$	$= (h_{k+1}-h_k)\Delta_{k+1}$.
β_k $(\forall k \in [1:s-1])$	$= (h_k - h_{k-1})\Delta_k$.
$\{I_k^*\}_{k=0}^s$	With an abuse of notation denotes any solution to the set of KKT conditions stated in (21).
$\{I_k^{**}\}_{k=0}^s$	The optimum power allocation of the optimization problem (20) that also satisfies (21).
$r_k^{(1)}$ $(\forall k \in [0:s])$	The root of the numerator of $F^{(1)}(x)$ defined in (23). Note that we define $r_0^{(1)} \triangleq P_{\max}$ and $r_s^{(1)} \triangleq 0$.
$r_{k,1}^{(2)}, r_{k,2}^{(2)}$	The real roots (if exist) of the numerator of $F^{(2)}(x)$ defined in (24).

TABLE I: Explanation for some of the important variables.

The main idea of our proof is to propose a recursive algorithm that first finds all the solutions of the KKT equations (21), (with an abuse of notation) each is denoted by $\{I_k^*\}_{k=0}^s$. Then among these solutions finds the one that maximizes the secrecy rate given by (20), which is denoted by $\{I_k^{**}\}_{k=0}^s$.

To this end, in every iteration, the proposed algorithm picks some k (for that I_k^* is not determined yet) and then determine the sign of $F_k^{(1)}(x)$ (or as will be discussed later,

[5]Karush–Kuhn–Tucker conditions (e.g., see [25]).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

for some cases determine the sign of $F_k^{(2)}(x)$ which will be defined in (24)). Then using (22) (or in some cases using (24)) in addition to the complementary slackness condition, it determines whether we have to examine the following three cases: (i) $I_k^* = I_{k-1}^*$, (ii) $I_k^* = I_{k+1}^*$, or (iii) the value of I_k^* is determined in this iteration. Hence, at the end of each iteration the size of the optimization problem is reduced by one (either I_k^* is determined in this iteration or is equal to the I_{k-1}^* or I_{k+1}^*). Ignoring the details for a moment, we can repeat the above procedure until all values of I_k^* are determined. These sets of I_k^*'s are the solutions to the KKT conditions (21).

Considering more details, we can proceed as follows. First, let $r_k^{(1)}$ to be the root of the numerator of $F_k^{(1)}(x)$ (assuming $h_{k+1}\beta_k - h_{k-1}\alpha_k \neq 0$), namely,

$$r_k^{(1)} \triangleq \frac{\alpha_k - \beta_k}{h_{k+1}\beta_k - h_{k-1}\alpha_k}, \qquad \forall k \in [1:s-1], \tag{23}$$

and by convention set $r_0^{(1)} \triangleq P_{\max}$ and $r_s^{(1)} \triangleq 0$. Then we can observe the following different situations, as stated in Case 1.

Case 1 (Linear Case[6])**.** *Based on different values of problem parameters, for each $k \in [1:s-1]$, we have various cases for a solution I_k^* that satisfies* (21) *as follows:*
(1) If $h_{k+1}\beta_k - h_{k-1}\alpha_k = 0$ (which means that the numerator of $F_k^{(1)}(x)$ is a constant and $r_k^{(1)}$ is not defined), then because of the ordering over channel gains we should have $\alpha_k > \beta_k$. Now, because of (22) *and since λ_k's are non-negative, we should have $\lambda_k^* > 0$ which by using the complementary slackness condition leads to $I_k^* = I_{k-1}^*$. This is equivalent to $P_k^* = 0$.*
(2) If $\alpha_k < \beta_k$ then we get $h_{k+1}\beta_k > h_{k-1}\alpha_k$, so we have $r_k^{(1)} < 0$ and $F_k^{(1)}(x) > 0$ for $x \geq 0$. Because of (22)*, we conclude that $\lambda_{k+1}^* > 0$ which by using the complementary slackness condition results in $I_k^* = I_{k+1}^*$, i.e., $P_{k+1}^* = 0$.*
(3) If $h_{k+1}\beta_k < h_{k-1}\alpha_k$ we can conclude that $\alpha_k > \beta_k$ so we have $r_k^{(1)} < 0$ and $F_k^{(1)}(x) < 0$ for $x \geq 0$. Because of (22)*, we conclude that $\lambda_k^* > 0$ which by using the complementary slackness condition results in $I_k^* = I_{k-1}^*$, i.e., $P_k^* = 0$.*
(4) If $\alpha_k > \beta_k$ and $h_{k+1}\beta_k > h_{k-1}\alpha_k$ then we have $r_k^{(1)} > 0$. Moreover, we have $F_k^{(1)}(x) > 0$ for $x > r_k^{(1)}$ and $F_k^{(1)}(x) < 0$ for $x < r_k^{(1)}$. Now, there exists the following different cases:

(a) *If $P_{\max} < r_k^{(1)}$ then we have $F_k^{(1)}(x) < 0$ for $x \leq P_{\max}$. From* (22)*, we conclude that $\lambda_k^* > 0$ which leads to $I_k^* = I_{k-1}^*$, i.e., $P_k = 0$.*
(b) *If $0 < r_k^{(1)} \leq P_{\max}$ then we have $F_k^{(1)}(x) < 0$ for $x < r_k^{(1)}$ and $F_k^{(1)}(x) > 0$ for $x > r_k^{(1)}$. Now I_k^* can be equal to $r_k^{(1)}$ without any further requirement. However, if we have $r_k^{(1)} < I_k^*$ then we should have $I_k^* = I_{k+1}^*$. Similarly if we have $I_k^* < r_k^{(1)}$ then we have to have $I_k^* = I_{k-1}^*$.*

The above different cases are derived under the assumption that we do not have any extra information about a solution $\{I_k^*\}$. In particular, prior to solving the KKT conditions, we do not know whether we have $I_k^* = I_{k\pm l}^*$ (for some valid k and l) or not. Note that due to the ordering on the optimal solution I_k^*'s imposed by (20), if we have $I_k^* = I_{k+l}^*$ then we should also have $I_k^* = I_{k+1}^* = \cdots = I_{k+l}^*$.

Now suppose that, by some mean (e.g., from the previous iterations of our proposed algorithm for finding the solutions of KKT equations (21)), we know that $I_k^* = I_{k+l}^*$. This knowledge enables us to reduce the size of the optimization problem (20). Notice that after having this information, the derivative of the Lagrangian $\mathfrak{L}$ with respect to I_k is not given by (22) anymore. More precisely, let us assume that $I_k^* = I_{k+l}^*$. Then, taking the derivative of $\mathfrak{L}$ with respect to I_k and by doing some algebra, we can write

$$\begin{aligned} 0 = \frac{\partial \mathfrak{L}}{\partial I_k} = \\ = \frac{\Delta_k(1+h_{k+l}I_k)(1+h_{k+l+1}I_k)(h_k - h_{k-1})}{(\ln 2)(1+h_{k-1}I_k)(1+h_kI_k)(1+h_{k+l}I_k)(1+h_{k+l+1}I_k)} \\ - \frac{\Delta_{k+l+1}(1+h_{k-1}I_k)(1+h_kI_k)(h_{k+l+1} - h_{k+l})}{(\ln 2)(1+h_{k-1}I_k)(1+h_kI_k)(1+h_{k+l}I_k)(1+h_{k+l+1}I_k)} \\ + (\lambda_k - \lambda_{k+l+1}) = F_k^{(2)}(I_k) + (\lambda_k - \lambda_{k+l+1}). \end{aligned} \tag{24}$$

Notice that for all values of $l \in [1:s-k]$, the numerator of $F_k^{(2)}(I_k)$ is a quadratic function in I_k and the denominator is strictly positive for $I_k \geq 0$ because h_k's are positive real quantities.

Similar to the Case 1, here in this case, we can also find the real roots $r_{k,1}^{(2)}$ and $r_{k,2}^{(2)}$ of the numerator of $F_k^{(2)}(x)$ and find the sign of the function $F_k^{(2)}(x)$ for different values of $x \in [0, P_{\max}]$ based on the place of these roots. Hence, based on the real roots of the numerator of $F_k^{(2)}(x)$, we can write a set of different conditions similar to Case 1, as stated in Case 2.

Case 2 (Quadratic Case[7])**.** *Here, we do not write all the different possibilities for* (24) *because the idea is very similar to Case 1. Instead, we explain the main part of the procedure in the following. In general, based on the position of the roots of numerator of $F_k^{(2)}(x)$, we can potentially have up to five different cases. To clarify the method, for example, consider the situation where the numerator of $F_k^{(2)}(x)$ has two distinct real roots $r_{k,1}^{(2)} < r_{k,2}^{(2)}$ where $r_{k,1}^{(2)}, r_{k,2}^{(2)} \in [0, P_{\max}]$. Then we have to consider the following five different cases for the solution I_k^*: (1) $I_k^* \in \left[0, r_{k,1}^{(2)}\right)$, (2) $I_k^* = r_{k,1}^{(2)}$, (3) $I_k^* \in \left(r_{k,1}^{(2)}, r_{k,2}^{(2)}\right)$, (4) $I_k^* = r_{k,2}^{(2)}$, and (5) $I_k^* \in \left(r_{k,2}^{(2)}, P_{\max}\right]$. In the items (1), (3), and (5) one can find the sign of $F_k^{(2)}(x)$ in the corresponding interval and based on that determine whether we should have $I_k^* = I_{k-1}^*$ or $I_k^* = I_{k+l+1}^*$.*

Remark 6. *It is worth to emphasize that the structure of the optimization problem* (20) *is such that the denominator of $F_k^{(1)}(x)$ and $F_k^{(2)}(x)$ are always strictly positive for $x \geq 0$. Moreover, the numerator of $F_k^{(1)}(x)$ is always at most a linear function of x and that of $F_k^{(2)}(x)$ is always at most a quadratic function in x. This fact significantly simplifies*

[6]Here by the linear case, we mean that the numerator of $F_k^{(1)}(I_k)$ in (22) is a linear function of I_k.

[7]Here by the quadratic case, we mean that the numerator of $F_k^{(2)}(I_k)$ in (24) is a quadratic function of I_k.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

finding the solutions of KKT equations (21) *and hence solving the optimization problem* (20).

The above discussion for different possible cases based on the numerator roots of $F_k^{(1)}$ and $F_k^{(2)}$ can be applied to any specific instance of the optimization problem (20). As briefly explained before, the main idea is to apply a recursive algorithm that finds the set of solutions of (21) iteratively by determining variables I_k's one by one. This procedure can also be considered as reducing the size of the optimization problem (20) (i.e., number of states) by one in every iteration.

To better explain our proposed method, we present the pseudo code of our algorithm in Algorithm 1 and Algorithm 2. For more clarification, in addition to the comments inside the pseudo code, the important variables of the pseudo code are explained separately in Table II.

Putting it together, we can describe our algorithm as follows (see also Algorithm 1 and Algorithm 2). First, Algorithm 1 initializes a data structure $\mathrm{d}(i)$ for each $i \in [0:s]$ which contains the required information about I_i^* (see Lines 1 to 8 of Algorithm 1). Then it calls Algorithm 2 that is a recursive function.

Starting from the original KKT conditions, at every iteration, Algorithm 2 picks a number k (such that I_k is not determined yet) and apply Case 1 or Case 2 (depending if it is linear or quadratic case) to that particular I_k (see Lines 3, 6 and 16 of Algorithm 2). By doing so, the size of the original KKT conditions (i.e., number of undetermined variables I_k) is reduced by one and we may have up to five (in fact up to three if Case 1 holds and up to five if Case 2 holds) new set of KKT conditions to be solved.

Now, we can repeat the above process on each of these new set of conditions and go forward iteratively (see Lines 13 and 23 of Algorithm 2). This procedure is like discovering a tree starting from some point as root (the root is determined by the first $k \in [1:s-1]$ picked up by the algorithm). Note that many of these new set of conditions do not lead to valid solutions that satisfy the original KKT conditions (21). This will be determined later as the algorithm proceeds by observing some contradictions on the intervals of I_k's. This process continues until we obtain problems of zero size (that have all variable I_k's determined and satisfy the original KKT conditions (21); Line 29 of Algorithm 2) or at some point in the middle of the algorithm the determined I_k's up to that point violate the KKT conditions (so this particular branch will be discarded; Lines 12 and 22 of Algorithm 2).

The above-mentioned algorithm enables us to find all of the solutions to KKT conditions (21). Then because the KKT equations provide a necessary condition on the optimal solution, it is sufficient to check among all of the solutions of KKT equations to find the optimal power allocation for the optimization problem (20) (Line 10 of Algorithm 2). Consequently, the search space of the original optimization problem is reduced from the continuous space $\mathbb{R}^{s-1}$ to a set of size at most 5^{s-1} elements. Note that this is the worst case analysis and in practice the size of the set can be much smaller than this number[8].

Variable	Description
$\mathrm{d}(i)$	An array that contains the available information about the solution I^* of (21) at every step of the algorithm. At the beginning, we have $i \in [0, s]$. However, the size of the problem becomes smaller in every iteration.
$\mathrm{d}(i).\mathrm{l}$ and $\mathrm{d}(i).\mathrm{u}$	Lower and Upper bounds on the indices of states such that we have $I^*_{\mathrm{d}(i).\mathrm{l}} = \cdots = I^*_{\mathrm{d}(i).\mathrm{u}}$.
$\mathrm{d}(i).\min$ and $\mathrm{d}(i).\max$	Determine the interval that I_k^*'s belongs to, i.e., $I_k^* \in \big[\mathrm{d}(i).\min, \mathrm{d}(i).\max\big]$ where $k \in \big[\mathrm{d}(i).\mathrm{l} : \mathrm{d}(i).\mathrm{u}\big]$.
$\mathrm{d}(i).\mathrm{determined}$	Let $k = \mathrm{d}(i).\mathrm{l}$. If the value of I_k^* is completely determined, we have $\mathrm{d}(i).\mathrm{determined}$ = "true" otherwise it is equal to "false".
SolSet	A set that at the end of the algorithm contains all of the solutions (data structures d) that satisfy the KKT conditions (21).

TABLE II: Description of the important variables in Algorithms 1 and 2.

Algorithm 1 Finding all of the solutions that satisfy KKT conditions (21).

Require: s, $\{h_i\}_{i=0}^s$, $\{\Delta_k\}_{k=1}^s$, $P_{\max}$
1: **for all** $i \in [0:s]$ **do** ▷ Initialization
2: $\mathrm{d}(i).\mathrm{l} = \mathrm{d}(i).\mathrm{u} = i$ ▷ In general we may have $I^*_{\mathrm{d}(i).\mathrm{l}} = \cdots = I^*_{\mathrm{d}(i).\mathrm{u}}$
3: $\mathrm{d}(i).\min = 0$ ▷ Initializing the lower bound on I_i^*
4: $\mathrm{d}(i).\max = P_{\max}$ ▷ Initializing the upper bound on I_i^*
5: $\mathrm{d}(i).\mathrm{determined} = \mathrm{false}$ ▷ At the beginning, the value of I_i^* is not determined
6: **end for**
7: $\mathrm{d}(0).\min = P_{\max}$; $\mathrm{d}(0).\mathrm{determined} = \mathrm{true}$ ▷ Also part of the initialization
8: $\mathrm{d}(s).\max = 0$; $\mathrm{d}(s).\mathrm{determined} = \mathrm{true}$ ▷ Also part of the initialization
9: $\mathrm{SolSet} = \textsc{Recursion}(\mathrm{d}, \{h_i\}_{i=0}^s, \{\Delta_k\}_{k=1}^s)$
10: For all $I \in \mathrm{SolSet}$ find the one which maximizes the achievable rate R; call it I^{**}
11: **return** I^{**}
12: **end**

In the following, in Lemma 6 and in Section VII-A, we present two special cases for the optimization problem (20) that is insightful.

Lemma 6. *Consider the set of* $\{r_k^{(1)}\}_{k=0}^s$ *as defined in* (23). *If we have* $0 = r_s^{(1)} < r_{s-1}^{(1)} < \cdots < r_1^{(1)} < r_0^{(1)} = P_{\max}$, *then the KKT conditions given by* (21) *have a unique solution. Moreover, the optimal power allocation is determined by* $I_k^{**} = r_k^{(1)}$ *for all* $k \in [1:s-1]$.

Proof. For the proof of this lemma refer to [9, Appendix B]. □

[8]In the examples discussed in the following, the number of solutions is something like 2 or 3.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

A. High-dynamic range, high-SNR regime

In this section, we show that our proposed achievability scheme, stated in Section VI-B, is optimal for the "high-dynamic range, high-SNR regime" in a degrees of freedom sense. We first give a formal definition of degrees of freedom in our setup as follows. The degrees of freedom for secret key sharing over a state dependent Gaussian broadcast channel is defined as

$$\mathrm{DoF}_s = \lim_{Q\to\infty} \frac{C_s^{\mathsf{gaus}}}{\frac{1}{2}\log Q}$$

where $h_i = Q^{\gamma_i}$ for $i \in [0:s]$, $\gamma_i > 0$ and $\gamma_i > \gamma_{i-1}$.

Clearly, as $Q \to \infty$, $h_i \gg h_{i-1}$ (high-dynamic range) and $h_i \gg 1$ (high-SNR). The following theorem completely characterizes DoF_s, and hence proves the optimality of our proposed achievability scheme in the high-dynamic range and high-SNR regime.

Lemma 7. *The degrees of freedom (DoF_s) for secret key sharing over a state dependent Gaussian broadcast channel is given by:* $\mathrm{DoF}_s = L\sum_{i=1}^{s}(\gamma_i - \gamma_{i-1})\,\Delta_i$.

Proof. We prove the theorem in two steps. First, we show a lower bound on DoF_s using the proposed achievability scheme in Section VI-B, and then we show a matching upper bound on DoF_s using the upper bound on the secret key generation capacity as stated in Lemma 4.

Lower bound on DoF_s*:* If $h_i \gg h_{i-1}$ for all i, we have $r_i^{(1)} \doteq Q^{-\gamma_i}$. Then the ordering condition stated in Lemma 6 is satisfied and as a result we have $I_i^{**} = r_i^{(1)}$. Using this observation, we can derive a lower bound on DoF_s as shown below,

$$\begin{aligned}
\mathrm{DoF}_s &= \lim_{Q\to\infty} \frac{C_s^{\mathsf{gaus}}}{\frac{1}{2}\log Q} \\
&\overset{(a)}{\geq} \lim_{Q\to\infty} \frac{L\sum_{i=1}^{s}\Delta_i\left(\frac{1}{2}\log\left(\frac{1+h_i r_{i-1}^{(1)}}{1+h_i r_i^{(1)}}\cdot\frac{1+h_{i-1}r_i^{(1)}}{1+h_{i-1}r_{i-1}^{(1)}}\right)\right)}{\frac{1}{2}\log Q} \\
&= L\sum_{i=1}^{s}\Delta_i\,(\gamma_i-\gamma_{i-1}) \qquad (25)
\end{aligned}$$

where (a) follows from $h_i \gg h_{i-1}$ and Lemma 6.

Upper bound on DoF_s*:* An upper bound on DoF_s can be derived as shown below,

$$\begin{aligned}
\mathrm{DoF}_s &= \lim_{Q\to\infty} \frac{C_s^{\mathsf{gaus}}}{\frac{1}{2}\log Q} \\
&\overset{(a)}{\leq} L\sum_{i>j}\delta_i\delta_j\,(\gamma_i-\gamma_j) \\
&= L\sum_{i=1}^{s}\sum_{j=0}^{i-1}\sum_{k=j+1}^{i}(\gamma_k-\gamma_{k-1})\,\delta_i\delta_j \\
&\overset{(b)}{=} L\sum_{k=1}^{s}\sum_{i=k}^{s}\sum_{j=0}^{k-1}(\gamma_k-\gamma_{k-1})\delta_i\delta_j \\
&= L\sum_{k=1}^{s}\Delta_k(\gamma_{k-1}-\gamma_k)
\end{aligned}$$

where (a) follows from Lemma 4 and (b) follows by exchanging the order of the summations. The above upper bound on DoF_s matches the lower bound in (25) and this completes the proof of the theorem. □

Algorithm 2 The recursive core part of the algorithm that is called by Algorithm 1.

function RECURSION(d,$\{h_i\}_{i=0}^{s}$, $\{\Delta_k\}_{k=1}^{s}$)
  SolutionSet $= \varnothing$
  Find an index j such that $\mathrm{d}(j).\mathrm{determined} = \mathrm{false}$
  if there is such j **then**
    $k = \mathrm{d}(j).\mathrm{l}$
    if $\mathrm{d}(j).\mathrm{l} = \mathrm{d}(j).\mathrm{u}$ **then** ▷ Linear case ($I_k^* \neq I_{k+1}^*$)
      Find the root $r^{(1)}$ of the numerator of $F_k^{(1)}(x)$ defined in (23)
      Based on the value of $r^{(1)}$, break the interval $[\mathrm{d}(j).\min, \mathrm{d}(j).\max]$ if necessary
      for all possible subinterval of the interval $[\mathrm{d}(j).\min, \mathrm{d}(j).\max]$ **do**
        Find the sign of $F_k^{(1)}(x)$ in this subinterval
        According to the sign of $F_k^{(1)}(x)$ (and based on Case 1), update $\mathrm{d}(j).\mathrm{l}$, $\mathrm{d}(j).\mathrm{u}$, $\mathrm{d}(j).\min$, $\mathrm{d}(j).\max$, and $\mathrm{d}(j).\mathrm{determined}$, but to a new data structure d'
        if d' is consistent up to this point **then** ▷ if $\forall i$ we have $\mathrm{d}'(i).\min \leq \mathrm{d}'(i).\max$
          SolutionSet $\leftarrow$ SolutionSet $\cup$ RECURSION(d', $\{h_i\}_{i=0}^{s}$, $\{\Delta_k\}_{k=1}^{s}$)
        end if
      end for
    else ▷ Quadratic case
      Find the roots $r_1^{(2)}$ and $r_2^{(2)}$ of the numerator of $F_k^{(2)}(x)$
      Based on the values of $r_1^{(2)}$ and $r_2^{(2)}$, break the interval $[\mathrm{d}(j).\min, \mathrm{d}(j).\max]$ if necessary
      for all possible subinterval of the interval $[\mathrm{d}(j).\min, \mathrm{d}(j).\max]$ **do**
        Find the sign of $F_k^{(2)}(x)$ in this subinterval
        According to the sign of $F_k^{(2)}(x)$, update $\mathrm{d}(j).\mathrm{l}$, $\mathrm{d}(j).\mathrm{u}$, $\mathrm{d}(j).\min$, $\mathrm{d}(j).\max$, and $\mathrm{d}(j).\mathrm{determined}$, but to a new data structure d'
        if d' is consistent up to this point **then** ▷ if $\forall i$ we have $\mathrm{d}'(i).\min \leq \mathrm{d}'(i).\max$
          SolutionSet $\leftarrow$ SolutionSet $\cup$ RECURSION(d', $\{h_i\}_{i=0}^{s}$, $\{\Delta_k\}_{k=1}^{s}$)
        end if
      end for
    end if
  else ▷ If all $\mathrm{d}(j)$'s are determined
    if the found solution is consistent **then** ▷ if $\forall i$ we have $\mathrm{d}(i).\min \leq \mathrm{d}(i+1).\min$
      SolutionSet $= \{\mathrm{d}\}$
    end if
  end if

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

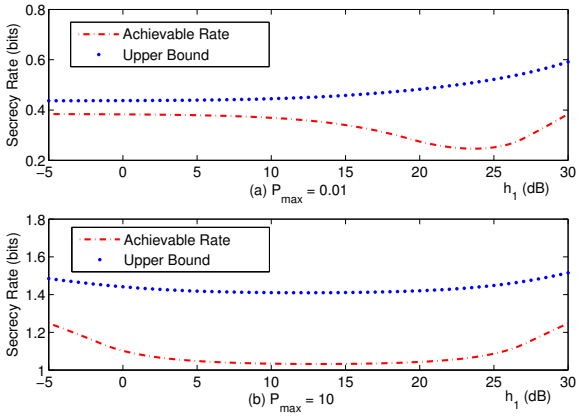

Fig. 3: The achievable rate and the upper bound as a function of h_1 with $P_{\max}$: (a) $P_{\max} = 0.01$, (b) $P_{\max} = 10$ (see Example 1).

Algorithm 3 Algorithm 2, continue

32: **return** SolutionSet
33: **end function**

B. Numerical Evaluations

In this section, we numerically evaluate the performance of the secret key sharing scheme proposed in Section VII for a few examples and compare it with the upper bound stated in Lemma 4.

Example 1. *Consider a setup with* 3 *states* ($s = 2$) *where* $h_0 = -5\text{dB}$, $-5\text{dB} < h_1 < 30\text{dB}$ *and* $h_2 = 30\text{dB}$. *The probability distribution across the states is assumed to be uniform. Figure 3 shows the achievable rate and the upper bound as a function of* h_1 *with the following choices of* $P_{\max}$*: (a)* $P_{\max} = 0.01$ *and (b)* $P_{\max} = 10$. *Clearly, there is a gap between the upper bound and the achievable rate. As it is mentioned before, the proposed scheme is not optimal in an absolute sense, but only in a degrees of freedom sense as proved in Section VII-A.*

Example 2. *Consider a setup with* 4 *states where* $h_0 = -5\text{dB}$, $h_3 = 30\text{dB}$, $h_1 = \min[g_1, g_2]$ dB *and* $h_2 = \max[g_1, g_2]$ dB *where* $-5\text{dB} < g_1, g_2 < 30\text{dB}$. *The probability distribution across the states is assumed to be uniform. Figure 4 shows the achievable rate and the upper bound as a function of* g_1 *and* g_2 *with* $P_{\max} = 10$. *Similar to Example 1, this illustrates the absolute gap between the upper bound and the achievable rate.*

Example 3. *Consider a setup with* 36 *(equiprobable) states, uniformly spaced in the range* -5dB *to* 30dB *(i.e.,* $h_i = (-5 + i)\text{dB}$, $i \in [0 : s]$*). Figure 5 shows the fraction of* $P_{\max}$ *allocated to each state by the proposed scheme for* $P_{\max} \in \{0.1, 1, 10, 100\}$.

The above examples illustrate different aspects of the proposed scheme: the gap with respect to the upper bound and the distribution of power across the states. In general, these aspects depend on the setup parameters. For a given setup, the numerical implementation of our proposed scheme can be used for efficient evaluations, even with a large number of states (e.g., 36 states in Example 3).

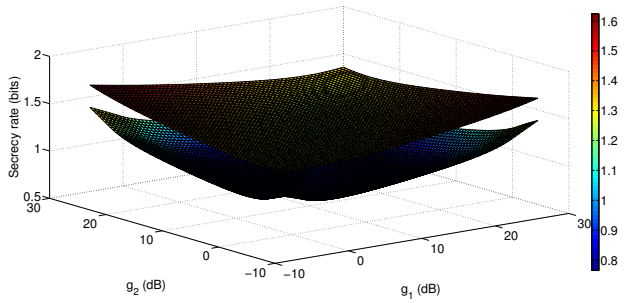

Fig. 4: The achievable rate (lower surface) and the upper bound (upper surface) as a function of g_1 and g_2 with $P_{\max} = 10$ in a setup with 4 equiprobable states (see Example 2).

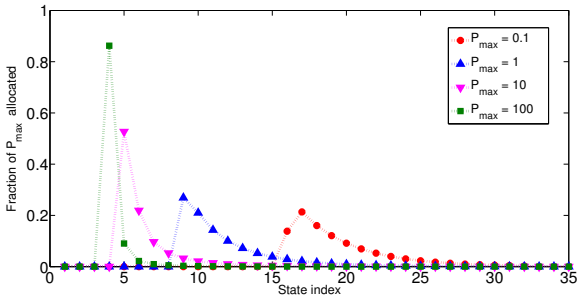

Fig. 5: Fraction of $P_{\max}$ allocated to each layer by the proposed scheme as explained in Example 3.

VIII. Discussion, Open Questions and Future Directions

Here, in this section we bring forward discussion about multi-party secret key sharing problem, open questions and possible future directions.

First, the SKG capacity problem among multiple terminals over a state-dependent Gaussian channel in the presence of a passive eavesdropper is still unsolved. But, the optimality of the proposed multi-party secret key sharing scheme has been shown for deterministic channels (that includes the erasure channels as a special case). By having intuition from this result, the achievability scheme for the Gaussian state-dependent channel is based on the message level erasure, simulated by using the wiretap code. However, in our outer bound on the SKG capacity, we do not have such an assumption and this can be a reason that explains the gap between our achievability scheme and outer bound.

Similar ideas used in this work for the secret key sharing problem over erasure channels can also be applied for the secret communication over these channels, e.g., see [28]. However, in our work, we go beyond and used these ideas to propose a coding scheme for multi-terminal secret key sharing over the Gaussian state-dependent broadcast channel (in the presence of public discussion). On the other hand, this is still open whether the same connection can be obtained between

secret communication over erasure and state-dependent Gaussian channels or not.

In our achievability scheme, we use public channel to send feedback from all the receivers to Alice. However, it is worth mentioning that although the public channel is available and without cost, we use it to communicate only the channel state which is a limited feedback; but not to transmit all the output feedback. Hence, it is possible to adopt our protocol to use ACK/NAK (e.g., similar to [24]) instead of public channel. However, the resulting protocol maybe not optimal even for the deterministic channels.

We would like to emphasize that this thread of work is not pure theoretical and there have been some attempts to implement these ideas (e.g., see [21], [22], [23], [24]). As an example, [24] reports to create shared secret key in a test-bed containing 5 nodes at rate 10 kbit/sec, with their secrecy being independent of the adversary's computational capabilities.

Finally, in this work we do not claim that our proposed scheme is a complete replacement of existing crypto-systems that rely on the adversarys computational limitations. However, if it is used in collaboration with such systems it can add an extra layer of security to the system in the physical layer.

REFERENCES

[1] M. Siavoshani, S. Mishra, S. Diggavi, and C. Fragouli, "Group secret key agreement over state-dependent wireless broadcast channels," in *IEEE International Symposium on Information Theory Proceedings*, Jul. 2011, pp. 1960–1964.

[2] A. Khisti, S. Diggavi, and G. W. Wornell, "Secret-Key Agreement With Channel State Information at the Transmitter," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 672–681, Sep. 2011.

[3] Y. K. Chia and A. E. Gamal, "Wiretap Channel With Causal State Information," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2838–2849, May 2012.

[4] P. Xu, Z. Ding, X. Dai, and K. K. Leung, "A General Framework of Wiretap Channel With Helping Interference and State Information," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 2, pp. 182–195, Feb. 2014.

[5] A. Sonee and G. A. Hodtani, "On the Secrecy Rate Region of Multiple-Access Wiretap Channel With Noncausal Side Information," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 6, pp. 1151–1166, Jun. 2015.

[6] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless Network Information Flow: A Deterministic Approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.

[7] Y. Liang, L. Lai, H. Poor, and S. Shamai, "The broadcast approach over fading Gaussian wiretap channels," in *IEEE Information Theory Workshop, 2009. ITW 2009*, Oct. 2009, pp. 1–5.

[8] M. Siavoshani, C. Fragouli, S. Diggavi, U. Pulleti, and K. Argyraki, "Group secret key generation over broadcast erasure channels," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2010, pp. 719–723.

[9] M. Jafari Siavoshani, S. Mishra, S. Diggavi, and C. Fragouli, "Group secret key agreement over state-dependent wireless broadcast channels," *CoRR*, vol. arXiv:1604.02380 [cs.CR], 2016.

[10] Y. Liang, L. Lai, H. Poor, and S. Shamai, "A Broadcast Approach for Fading Wiretap Channels," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 842–858, Feb. 2014.

[11] A. Wyner, "The wire-tap channel," *Bell System Technical Journal, The*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[12] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[13] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography - part I: Secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[14] I. Csiszar and P. Narayan, "Secrecy Capacities for Multiterminal Channel Models," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2437–2452, Jun. 2008.

[15] A. A. Gohari and V. Anantharam, "Information-Theoretic Key Agreement of Multiple Terminals – Part I," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3973–3996, Aug. 2010.

[16] ——, "Information-Theoretic Key Agreement of Multiple Terminals – Part II: Channel Model," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3997–4010, Aug. 2010.

[17] I. Csiszar and P. Narayan, "Secrecy Generation for Multiaccess Channel Models," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 17–31, Jan. 2013.

[18] C. Chan and L. Zheng, "Multiterminal Secret Key Agreement," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3379–3412, Jun. 2014.

[19] N. T. H. Phuong and H. Tuy, "A unified monotonic approach to generalized linear fractional programming," *Journal of Global Optimization*, vol. 26, no. 3, pp. 229–259, 2003.

[20] L. P. Qian, Y. J. Zhang, and J. Huang, "Mapel: Achieving global optimality for a non-convex wireless power control problem," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1553–1563, Mar 2009.

[21] I. Safaka, M. J. Siavoshani, U. Pulleti, E. Atsan, C. Fragouli, K. Argyraki, and S. Diggavi, "Exchanging Secrets without Using Cryptography," *arXiv:1105.4991 [cs, math]*, May 2011.

[22] I. Safaka, C. Fragouli, K. Argyraki, and S. Diggavi, "Exchanging pairwise secrets efficiently," in *2013 Proceedings IEEE INFOCOM*, Apr. 2013, pp. 2265–2273.

[23] E. Atsan, I. Safaka, L. Keller, and C. Fragouli, "Low cost security for sensor networks," in *2013 International Symposium on Network Coding (NetCod)*, Jun. 2013, pp. 1–6.

[24] K. Argyraki, S. Diggavi, M. Duarte, C. Fragouli, M. Gatzianas, and P. Kostopoulos, "Creating Secrets out of Erasures," in *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*. New York, NY, USA: ACM, 2013, pp. 429–440.

[25] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004.

[26] D. Klinc, J. Ha, S. W. McLaughlin, J. Barros, and B. J. Kwak, "LDPC Codes for the Gaussian Wiretap Channel," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 532–540, Sep. 2011.

[27] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Athena Scientific, 2003.

[28] L. Czap, V. M. Prabhakaran, C. Fragouli, and S. N. Diggavi, "Secret communication over broadcast erasure channels with state-feedbac," *IEEE Transactions on Information Theory*, vol. 61, pp. 4788–4808, Sep. 2015.

Mahdi Jafari Siavoshani is an Assistant Professor in the Department of Computer Engineering, Sharif University of Technology (SUT), Tehran, Iran, from 2013. He received his B.S. degrees in Communication Systems as well as Physics both from SUT in 2005. He was awarded an Excellency scholarship from Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, where he received the M.Sc. degree in 2007 and the Ph.D. degree in 2012, both in Computer, Communication, and Information Sciences. After completing his Ph.D., he joined the Institute of Network Coding (INC) at the Chinese University of Hong Kong (CUHK) as a postdoctoral fellow from 2012 to 2013. His research interests include finding fundamental limits of Communication Systems and Networks. Moreover, he is interested in the connection between Communication and Computer Sciences. In particular, he is interested in theory and algorithms for network information flow, network coding, network algorithms, information and coding theory, wireless communication networks.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2612649, IEEE Transactions on Information Forensics and Security

Shaunak Mishra received the B. Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India in 2010, and the M.S. degree in electrical engineering from the University of California, Los Angeles in 2011. He is currently a Ph.D. student in the department of Electrical Engineering at University of California, Los Angeles. His research interests include statistics and information theory with applications in security and machine learning. He is a recipient of the Henry Samueli fellowship at UCLA (2010).

Christina Fragouli is a Professor at UCLA in the Electrical Engineering Department. She received the B.S. degree in Electrical Engineering from the National Technical University of Athens, Athens, Greece, in 1996, and the M.Sc. and Ph.D. degrees in Electrical Engineering from the University of California, Los Angeles, in 1998 and 2000, respectively. She has worked at the Information Sciences Center, AT&T Labs, Florham Park New Jersey, and the National University of Athens. She also visited Bell Laboratories, Murray Hill, NJ, and DIMACS, Rutgers University. Between 2006-2007, 2007-2012 and 2012-2015 she was an FNS Assistant Professor, an Assistant Professor and an Associate Professor, respectively, in the School of Computer and Communication Sciences, EPFL, Switzerland.

She is an IEEE fellow. She served as an Associate Editor for IEEE Communications Letters, for Elsevier Computer Communication, for IEEE Transactions on Communications, for IEEE Transactions on Information Theory, and for IEEE Transactions on Mobile Communications. She received the Fulbright Fellowship for her graduate studies, the Outstanding Ph.D. Student Award 2000-2001, UCLA, Electrical Engineering Department, the Zonta award 2008 in Switzerland, the Starting Investigator ERC award in 2009, the Mobihoc 2013 best paper award, and the MASCOTS 2011 best paper award. Her research interests are in network coding, wireless communications, algorithms for networking and network security.

Suhas N. Diggavi received the B. Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, India, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1998. After completing his Ph.D., he was a Principal Member Technical Staff in the Information Sciences Center, AT&T Shannon Laboratories, Florham Park, NJ. After that he was on the faculty of the School of Computer and Communication Sciences, EPFL, where he directed the Laboratory for Information and Communication Systems (LICOS). He is currently a Professor, in the Department of Electrical Engineering, at the University of California, Los Angeles, where he directs the Information Theory and Systems laboratory. His research interests include wireless network information theory, wireless networking systems, network data compression and network algorithms; more information can be found at http://licos.ee.ucla.edu. He has received several recognitions for his research including the 2013 IEEE Information Theory Society & Communications Society Joint Paper Award, the 2013 ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc) best paper award, the 2006 IEEE Donald Fink prize paper award. He is currently a Distinguished Lecturer and also serves on board of governors for the IEEE Information theory society. He is a Fellow of the IEEE. He has been an associate editor for IEEE Transactions on Information Theory, ACM/IEEE Transactions on Networking, IEEE Communication Letters, a guest editor for IEEE Selected Topics in Signal Processing and in the program committees of several IEEE conferences. He has also helped organize IEEE conferences including serving as the Technical Program Co-Chair for 2012 IEEE Information Theory Workshop (ITW) and the Technical Program Co-Chair for the 2015 IEEE International Symposium on Information Theory (ISIT). He has 8 issued patents.