One-Shot Perfect Secret Key Agreement for Finite Linear Sources

Chung Chan, Navin Kashyap, Praneeth Kumar Vippathalla and Qiaoqiao Zhou

***Abstract*—We consider a non-asymptotic (one-shot) version of the multiterminal secret key agreement problem on a finite linear source model. In this model, the observation of each terminal is a linear function of an underlying random vector composed of finitely many i.i.d. uniform random variables. By restricting the public discussion to be a linear function of the terminals' observations, we obtain a characterization of the communication complexity (minimum number of symbols of public discussion) of generating a secret key of maximum length. More precisely, we show that the minimum discussion can be achieved by a non-interactive protocol in which each terminal first does a linear processing of its own private observations, following which the terminals all execute a discussion-optimal communication-for-omniscience protocol. The secret key can be chosen to be a linear function of the vector of all observations.**

I. Introduction

The problem of secret key agreement via public discussion was first formulated for two terminals by Maurer [1] and Ahlswede and Csiszár [2], and subsequently extended to multiple terminals by Csiszár and Narayan [3]. In the set-up of this problem, the terminals involved must agree upon a secret key based on correlated observations from a source, using interactive public discussion. The key must be kept information-theoretically secure from an eavesdropper having access to the public discussion. The conventional setting allows unlimited public discussion, and the aim is to agree upon a secret key of largest possible length. The problem formulation is in fact asymptotic in nature: the terminals observe an infinite sequence of i.i.d. realizations of the correlated source random variables, and the asymptotic secret key rate (number of symbols of secret key generated per source realization) must be as large as possible. The largest possible asymptotic key rate, termed the secrecy capacity, is by now quite well understood [3, 4].

A more difficult problem is to determine the secrecy capacity under a constraint on the amount or rate of public discussion allowed. Specifically, when the (asymptotic) rate of public discussion is bounded above by R, the problem is to determine the maximum achievable secret key rate $C_{\mathrm{S}}(R)$, which we term the rate-constrained secrecy capacity. This problem was considered in the case of two terminals by Tyagi [5] and Liu et al. [6]. The primary focus of Tyagi [5] was on the related problem of characterizing what we will call the communication complexity R_{S}, which is the least discussion rate needed to achieve the (unconstrained) secrecy capacity; he left open the rate-constrained secrecy capacity problem. Liu et al. [6] gave a characterization of the achievable region of key and discussion rate pairs using a notion of XY-concave envelopes that they develop. They used their methods to give a precise description of the ratio $\frac{C_{\mathrm{S}}(R)}{R}$ in the regime of $R \to 0$.

The multiterminal $C_{\mathrm{S}}(R)$ and R_{S} problems were considered in our prior works [7–10]. Among our contributions there were some general outer bounds on the achievable rate region, and upper and lower bounds on R_{S}; in the special case of the hypergraphical source model, we derived tighter upper bounds on R_{S} and the ratio $\frac{C_{\mathrm{S}}(R)}{R}$ valid for all $R > 0$. In the important special case of the pairwise independent network (PIN) model (see e.g. [11]), our bounds were good enough to precisely characterize R_{S} and $C_{\mathrm{S}}(R)$.

In this paper, we make further progress on these problems by focusing on the (multiterminal) finite linear source model [12], which generalizes the hypergraphical source and PIN models. In the finite linear model, the observation of each terminal is a linear function of an underlying random vector composed of finitely many i.i.d. uniform random variables. Furthermore, we consider a non-asymptotic, *single-shot* version of the secret key agreement problem as opposed to the asymptotic version in [12]. In this version, the terminals observe only one realization of the source, and after some public discussion, must agree (with probability 1) upon a secret key that is statistically independent of the public communication. Single-shot analogues of the R_{S} and $C_{\mathrm{S}}(R)$ problems can be formulated in this setting — see Section II. We study these problems with a view towards extending the results obtained for the single-shot setting to the asymptotic model.

Courtade and Halford [13] formulated and analyzed the single-shot secret key generation problem for hypergraphical sources. They made a key assumption to facilitate their analysis, namely, that the communication from each terminal is a linear function of its observations. Under this restriction, they effectively resolved the single-shot R_{S} and $C_{\mathrm{S}}(R)$ problems for hypergraphical sources. Note that linear discussion was also considered in [12, 14] for finite linear sources, but the objective there was to achieve the unconstrained secrecy capacity of the asymptotic model perfectly at a finite blocklength, so

C. Chan (corresponding author, email: chung.chan@cityu.edu.hk) is with the Department of Computer Science, City University of Hong Kong. His work is supported by a grant from the University Grants Committee of the Hong Kong Special Administrative Region, China (Project No. 21203318).

Q. Zhou (email: zq115@ie.cuhk.edu.hk) is with the Institute of Network Coding and the Department of Information Engineering, The Chinese University of Hong Kong.

N. Kashyap (nkashyap@iisc.ac.in) and Praneeth Kumar V. (praneethv@iisc.ac.in) are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012. Their work is supported in part by a Swarnajayanti Fellowship awarded to N. Kashyap by the Department of Science & Technology (DST), Government of India.

as to avoid excessive delay in generating the secret key.

Taking inspiration from [13], we too restrict the public discussion to be a linear function of the terminals' observations. Under the linear discussion model, for finite linear sources, we obtain a characterization (Corollary 2 in Section IV) of the communication complexity of generating a secret key of maximum length. The minimum discussion is achieved by a non-interactive protocol in which each terminal first does a linear processing of its own private observations, following which the terminals all execute a (single-shot) discussion-optimal communication-for-omniscience protocol on their linearly processed observations. At the end of this, each terminal is able to recover the observations of all the other terminals (omniscience), and it then applies a linear function to the entire vector of observations to obtain a maximum-length secret key.

The rest of the paper is organized as follows. Section II contains the formal problem formulation, Section III presents an illustrative example, and Section IV contains statements of the main results, complete proofs of which can be found in the full version of this paper [15]. The paper ends in Section V with a discussion of the possible ways in which the results could be extended to settings beyond that of our problem formulation.

II. Problem formulation

We use the sans serif font K to represent a random variable with distribution P_{K} and taking values from a set K. We use the boldface uppercase $\boldsymbol{M}$ for matrices and boldface lowercase san serif font $\mathbf{x} := \begin{bmatrix}\mathsf{x}_1 & \dots & \mathsf{x}_{\ell(\mathbf{x})}\end{bmatrix}$ for random row vectors, where $\ell(\mathbf{x})$ denotes the length of the vector. We assume all the entries take values from the same finite field $\mathbb{F}_q$ of order q. We take logarithm $\log$ to base q and so all the information quantities are in the units of $\log q$ bits. For a finite set B, we use $\mathbf{y}_B := \begin{bmatrix}\mathbf{y}_{i_1} & \cdots & \mathbf{y}_{i_{|B|}}\end{bmatrix}$ to denote a row vector obtained by concatenating the row vectors $\mathbf{y}_i$'s for some enumeration $i_1, \dots, i_{|B|}$ of the set B. We use the notation

$$\mathbf{x} \in \langle\!\langle \mathbf{y}_B \rangle\!\rangle \text{ or } \langle\!\langle \mathbf{y}_{i_1}, \dots, \mathbf{y}_{i_{|B|}} \rangle\!\rangle$$

to mean that there exists a deterministic matrix $\boldsymbol{M}$ such that $\mathbf{x} = \mathbf{y}_B \boldsymbol{M}$.

As in [3], the multiterminal secret key agreement problem consists of a finite set $V = \{1, 2, \dots, m\}$ of $m \geq 2$ users who want to share a secret key after some public discussion that can be eavesdropped by a wiretapper. The one-shot perfect linear secret key agreement (SKA) scheme consists of the following phases.

One-shot private observation: Each user $i \in V$ observes the component $\mathbf{z}_i$ of a given finite linear source $\mathbf{z}_V$ defined in [4] with the requirement that

$$\mathbf{z}_i \in \langle\!\langle \mathbf{x} \rangle\!\rangle \quad \forall i \in V \tag{1}$$

for some uniformly random vector $\mathbf{x}$ over $\mathbb{F}_q$. $\mathbf{x}$ is referred to as the base of $\mathbf{z}_V$. In the special case when $\mathbf{z}_i$ is a subvector of $\mathbf{x}$, $\mathbf{z}_V$ is called the hypergraphical source [4], which is the source model considered in [13]. Unlike the model in [3] and [12] where each user observes n i.i.d. samples of the source, we consider the one-shot model as in [13, 16] where each user only observes one sample.

Private randomization: Each user $i \in V$ privately generates a random vector $\mathbf{u}_i$ over $\mathbb{F}_q$ independent of the source $\mathbf{z}_V$, i.e.,

$$P_{\mathbf{u}_V|\mathbf{z}_V} = \prod_{i \in V} P_{\mathbf{u}_i}. \tag{2}$$

Note that there is no restriction on the length nor the distribution of $\mathbf{u}_i$, and so the requirement that it must be a vector over $\mathbb{F}_q$ does not lose generality. Note also that such a randomization was not explicitly considered in the formulations of [3, 12, 13].

Linear public discussion: Each user $i \in V$ publicly reveals the message

$$\mathbf{f}_i \in \langle\!\langle \mathbf{u}_i, \mathbf{z}_i \rangle\!\rangle. \tag{3}$$

Hence, everyone including the wiretapper observes $\mathbf{f}_V$. Unlike [3], the discussion above is non-interactive as interaction is unnecessary for linear discussion as explained in [17].[1]

Secret key agreement: After the public discussion, each user $i \in V$ attempts to agree on a secret key K satisfying

$$H(\mathsf{K}|\mathbf{u}_i, \mathbf{z}_i, \mathbf{f}_V) = 0 \quad \forall i \in V \tag{4}$$

$$\log|K| - H(\mathsf{K}|\mathbf{f}_V) = 0 \tag{5}$$

where (4) is the recoverability constraint that requires the secret key to be perfectly recoverable by every user and (5) is the secrecy constraint that requires the key to be uniformly random and perfectly independent of the entire public discussion. Note that we do not assume *a priori* that K is a linear function of the private source, and so the key length $\log|K|$ is not required to be an integer.[2]

The objective is to characterize the set of achievable key lengths and discussion lengths. In particular, a quantity of interest is the *constrained secrecy capacity* defined as

$$c_{\mathrm{S}}(r) := c_{\mathrm{S}}(\mathbf{z}_V, r) := \max\{\log|K| \mid \ell(\mathbf{f}_V) \leq r\}, \tag{6}$$

where the maximization is over all possible secret key agreement schemes subject to a constraint on the total public discussion length, r. (The dependency on $\mathbf{z}_V$ is implicit if there is no ambiguity.) Characterizing the entire curve of $c_{\mathrm{S}}(r)$ is difficult even in the case of linear discussion, but some points on the curve can be characterized, such as $c_{\mathrm{S}}(0)$ considered

[1] Suppose the discussion is interactive, i.e., a message, say f, revealed in public by some user $i \in V$ is a linear function $\psi(\mathbf{u}_i, \mathbf{z}_i, \tilde{\mathbf{f}})$ of the private observations of user i as well as all the previously discussed messages denoted by $\tilde{\mathbf{f}}$. By linearity, we can rewrite f as $\psi(\mathbf{u}_i, \mathbf{z}_i, \mathbf{0}) + \psi(\mathbf{0}, \mathbf{0}, \tilde{\mathbf{f}})$ where $\mathbf{0}$ denotes an all-zero vector of an appropriate length. Note that, given $\tilde{\mathbf{f}}$, there is a bijection between f and $\mathsf{f}' := \psi(\mathbf{u}_i, \mathbf{z}_i, \mathbf{0})$, and so user i can reveal f' instead of f in public without loss of generality, since f can be recovered from f' and other discussion messages $\tilde{\mathbf{f}}$. As f' does not depend on $\tilde{\mathbf{f}}$, we can convert any interactive discussion to a non-interactive discussion by replacing every discussion message f by the corresponding f'.

[2] Nevertheless, it will follow from Theorem 1 that K can be chosen to be a linear function of the private source without loss of optimality, and so the maximum key length is indeed an integer.

in [18]. As in [3, 12, 13], we also consider the *unconstrained secrecy capacity* defined as

$$c_{\mathrm{S}} := c_{\mathrm{S}}(\mathbf{z}_V) := \max\{c_{\mathrm{S}}(r) \mid r \geq 0\}, \tag{7}$$

which is the secrecy capacity without the constraint on the discussion length. The smallest discussion length required to achieve c_{S} is denoted by

$$r_{\mathrm{S}} := r_{\mathrm{S}}(\mathbf{z}_V) := \inf\{r \geq 0 \mid c_{\mathrm{S}}(\mathbf{z}_V, r) = c_{\mathrm{S}}(\mathbf{z}_V)\} \tag{8}$$

and referred to as the *communication complexity*. As in [3, 13], we will characterize c_{S} and r_{S} using the closely related problem of communication for omniscience defined as follows.

Omniscience: We say that the public discussion achieves omniscience of $\mathbf{z}_V$ if

$$H(\mathbf{z}_V|\mathbf{u}_i, \mathbf{z}_i, \mathbf{f}_V) = 0 \quad \forall i \in V. \tag{9}$$

The smallest length of communication for omniscience is defined as

$$r_{\mathrm{CO}} := r_{\mathrm{CO}}(\mathbf{z}) := \min \ell(\mathbf{f}_V), \tag{10}$$

where the minimization is over all public discussion schemes subject to (9) in place of (5) and (4). The problem under the one-shot model for hypergraphical and finite linear sources is proposed in [16, 19] and referred to as the *cooperative data exchange*. In [3], the secret key agreement scheme that achieves the capacity is by first achieving omniscience of $\mathbf{z}_V$ and then extracting the secret key as a function of $\mathbf{z}_V$, implying that the rate of communication for omniscience is no smaller than the communication complexity. We say that c_{S} can be achieved via omniscience of $\mathbf{z}_V$.

III. Motivating Example

We will use the following example to illustrate the problem formulation and motivate our main results. Consider $V = \{1, 2, 3, 4\}$ and a finite linear source $\mathbf{z}_V$ (see (1)) over the binary field $\mathbb{F}_2$ with a base $\mathbf{x}$ of length $\ell(\mathbf{x}) = 4$ as follows:

$$\begin{aligned} \mathbf{z}_1 &:= \begin{bmatrix} \mathsf{x}_1 & \mathsf{x}_2 \oplus \mathsf{x}_3 \end{bmatrix} \\ \mathbf{z}_2 &:= \begin{bmatrix} \mathsf{x}_1 & \mathsf{x}_2 \oplus \mathsf{x}_4 \end{bmatrix} \\ \mathbf{z}_3 &:= \begin{bmatrix} \mathsf{x}_1 \oplus \mathsf{x}_2 & \mathsf{x}_3 \end{bmatrix} \\ \mathbf{z}_4 &:= \begin{bmatrix} \mathsf{x}_1 \oplus \mathsf{x}_3 \oplus \mathsf{x}_4 \end{bmatrix}. \end{aligned} \tag{11}$$

A feasible secret key agreement scheme is to choose

$$\mathsf{K} = \mathsf{x}_1, \mathbf{f}_1 = \begin{bmatrix} \mathsf{x}_2 \oplus \mathsf{x}_3 \end{bmatrix}, \text{ and } \mathbf{f}_2 = \begin{bmatrix} \mathsf{x}_2 \oplus \mathsf{x}_4 \end{bmatrix}, \tag{12}$$

but without any private randomizations $\mathbf{u}_V$ and discussions $\mathbf{f}_3$ and $\mathbf{f}_4$ by users 3 and 4. The secret key K is perfectly recoverable by every user, i.e., satisfying (4), since users 1 and 2 directly observes the key bit x_1, which can also be computed by users 3 and 4 using their private sources and public discussion as follows

$$\begin{aligned} \mathsf{x}_1 &= \underbrace{\begin{bmatrix} \mathsf{x}_1 \oplus \mathsf{x}_2 & \mathsf{x}_3 \end{bmatrix}}_{\mathbf{z}_3} \underbrace{\begin{bmatrix} \mathsf{x}_2 \oplus \mathsf{x}_3 \end{bmatrix}}_{\mathbf{f}_1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} \mathsf{x}_1 \oplus \mathsf{x}_3 \oplus \mathsf{x}_4 \end{bmatrix}}_{\mathbf{z}_4} \underbrace{\begin{bmatrix} \mathsf{x}_2 \oplus \mathsf{x}_3 \end{bmatrix}}_{\mathbf{f}_1} \underbrace{\begin{bmatrix} \mathsf{x}_2 \oplus \mathsf{x}_4 \end{bmatrix}}_{\mathbf{f}_2} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \end{aligned}$$

The secrecy constraint (5) also holds because $\log|K| = 1 = H(\mathsf{K}|\mathbf{f}_V)$, which follows from the definition of the base $\mathbf{x}$ that x_1 is uniformly random and independent of x_2, x_3, and x_4.

Note that the above scheme does not achieve the omniscience condition in (9) because users 1, 2 and 4 cannot recover x_3 after the discussion. However, it is easy to show that omniscience can be achieved if we further set $\mathbf{f}_3 = \begin{bmatrix} \mathsf{x}_3 \end{bmatrix}$, i.e., with an additional bit of discussion by user 3. Since 1 bit of secret key can be achieved with 2 bits of public discussion and omniscience can be further achieved with an additional bit of discussion, we have

$$c_{\mathrm{S}}(r_{\mathrm{CO}}) \geq \begin{cases} 1 & r \geq 2 \\ 0 & r < 2 \end{cases}, \text{ and } \begin{cases} c_{\mathrm{S}} \geq 1 \\ r_{\mathrm{S}} \leq 2 \\ r_{\mathrm{CO}} \leq 3 \end{cases} \tag{13}$$

by the definitions (6), (7), (8) and (10). The challenge is to show the reverse inequalities and therefore establish the optimality of the achieving schemes.

IV. Main Results

We start with some general admissible conditions that simplify the secret key agreement scheme without loss of optimality.

Theorem 1 *$c_{\mathrm{S}}(r)$ remains unchanged by imposing the additional constraints that*

$$\ell(\mathbf{u}_i) = 0 \quad \forall i \in V, \quad \textit{and} \tag{14a}$$

$$\mathsf{K} = \mathbf{k} \in \langle\!\langle \mathbf{z}_V \rangle\!\rangle, \tag{14b}$$

which mean respectively that private randomization is not needed and that the secret key can be chosen to be a linear function of the private source. □

Corollary 1 *$c_{\mathrm{S}}(r)$ must be integer, non-decreasing and right continuous in r.* □

Proof The claim in the corollary that $c_{\mathrm{S}}(r)$ must be an integer follows from (14b) that the key can be linear and therefore a uniformly random vector by the secrecy constraint (5). Monontonicity and continuity follows directly from the definition (6). The proof of the theorem is more involved and is given in [15, Appendix A]. ■

The example in Section III considers such a secret key agreement scheme without private randomization. The secret key x_1 is also linear in the private source trivially because it is observed by users 1 and 2 directly. Note that our formulation allows the private randomization to have arbitrary length and

distribution, and the key to be arbitrary random variables that need not be linear in the private source. The above admissible constraints (14) makes the problem tractable as it significantly reduces the space of secret key agreement schemes we need to consider to characterize $c_{\mathrm{S}}(r)$. Indeed, since there is only a finite number of linear functions of $\mathbf{z}_V$, there is only a finite number of admissible secret key agreement scheme. It is worth noting that the constraints (14) were assumed in the formulation of [13] for the hypergraphical source model, and our result implies that such constraints are admissible since hypergraphical sources are special case of the finite linear sources.

For the general source model in [3], the admissible constraint (14a) that private randomization does not help improve $c_{\mathrm{S}}(r)$ remains a plausible conjecture. However, it is clear that the constraint (14b) is not admissible for some sources that are not finite linear. Nevertheless, this constraint is essential in bringing the existing characterizations of the capacity from the general source model to the one-shot finite linear source model as follows.

Theorem 2 *$c_{\mathrm{S}}(r)$ in the extreme cases with 0 and respectively unbounded discussion lengths are*

$$c_{\mathrm{S}}(0) = \max\{H(\mathbf{g}) \mid \mathbf{g} \in \langle\!\langle \mathbf{z}_i \rangle\!\rangle, \forall i \in V\} \tag{15}$$

$$c_{\mathrm{S}} = \left\lfloor \min_{\mathcal{P} \in \Pi'(V)} \frac{\sum_{C \in \mathcal{P}} H(\mathbf{z}_C) - H(\mathbf{z}_V)}{|\mathcal{P}| - 1} \right\rfloor, \tag{16}$$

where the maximization is over the choices of random vector $\mathbf{g}$, and the minimization is over the collection $\Pi'(V)$ of partitions of V into at least two non-empty disjoint sets. c_{S} can be achieved via communication for omniscience of $\mathbf{z}_V$ at the smallest length

$$r_{\mathrm{CO}} = H(\mathbf{z}_V) - c_{\mathrm{S}}, \tag{17}$$

which implies the upper bound $r_{\mathrm{S}} \le r_{\mathrm{CO}}$ on r_{S}. □

PROOF See [15, Appendix B]. ■

For the running example given in Section III, it is straightforward to evaluate the above expressions (15), (16) and (17) to yield $c_{\mathrm{S}}(0) = 0$, $c_{\mathrm{S}} = 1$ and $r_{\mathrm{CO}} = 3$. In particular, an optimal solution to (16) can be shown to be $\mathcal{P} = \{\{1,2,3\},\{4\}\}$. This implies the optimality of the omniscience scheme in Section III in achieving both c_{S} and r_{CO}.

The above result follows quite directly from existing works for the asymptotic model. For instance, the r.h.s. of (15) is the multivariate Gács–Körner common information evaluated for the finite linear source model. $\mathbf{g}$ is called the maximal common function of $\mathbf{z}_i$ for $i \in V$. $c_{\mathrm{S}}(0) = J_{\mathrm{GK}}(\mathbf{z}_V)$ was shown in [18] but for the asymptotic model instead, and the result has recently been extended to general sources under a very general setting in [20]. It is straightforward to extend this result to the current one-shot model. Indeed, the capacity result can be shown for general sources.

The duality (17) between secret key agreement and communication for omniscience also follows directly from the asymptotic model in [3], which is specialized to the asymptotic finite linear source model in [12]. The characterization (16) of c_{S} is the same as that of the asymptotic model [3, 4] except for the floor operation, since the minimization in (16) may not be integer but c_{S} must be integer by Corollary 1. The characterization of r_{CO} for the one-shot finite linear source model is given in [16, 21], which focus primarily on the omniscience problem instead of the secret key agreement problem.

Note that one can summarize the theorem by saying that $c_{\mathrm{S}}(0)$, c_{S} and r_{CO} for the one-shot model is the same as those of the asymptotic model for finite linear source but with an additional integer constraint: $C_{\mathrm{S}}(0)$ is already an integer for the asymptotic model while we can take the floor and the ceiling respectively for C_{S} and R_{CO} to turn them into integer achievable lengths. It therefore appears reasonable to conjecture that $c_{\mathrm{S}}(r)$ for the one-shot model is the same as the $C_{\mathrm{S}}(R)$ for the asymptotic model for finite linear source but with an additional floor operation as in (16) to satisfy the integer constraint in Corollary 1. The following result resolves this partially at the communication complexity r_{S}.

Theorem 3 *If $r_{\mathrm{S}} < r_{\mathrm{CO}}$, then there exists $\mathbf{z}'_V$ with*

$$\mathbf{z}'_i \in \langle\!\langle \mathbf{z}_i \rangle\!\rangle \quad \forall i \in V \tag{18}$$

such that

$$r_{\mathrm{S}}(\mathbf{z}'_V) = r_{\mathrm{S}}(\mathbf{z}_V) \tag{19}$$

$$r_{\mathrm{CO}}(\mathbf{z}'_V) < r_{\mathrm{CO}}(\mathbf{z}_V) \tag{20}$$

$$c_{\mathrm{S}}(\mathbf{z}'_V) = c_{\mathrm{S}}(\mathbf{z}_V). \tag{21}$$

$\mathbf{z}'_V$ is said to be reduced source of $\mathbf{z}_V$ (by linear processing), since the above implies $H(\mathbf{z}'_V) < H(\mathbf{z}_V)$. □

Corollary 2 *The communication complexity is*

$$r_{\mathrm{S}} = \min\{r_{\mathrm{CO}}(\mathbf{z}'_V) \mid \mathbf{z}'_i \in \langle\!\langle \mathbf{z}_i \rangle\!\rangle, c_{\mathrm{S}}(\mathbf{z}'_V) = c_{\mathrm{S}}(\mathbf{z}_V)\}, \tag{22}$$

which can be achieved via omniscience of the linearly reduced source $\mathbf{z}'_V$. □

PROOF The corollary follows immediately from theorem by repeatedly linearly reducing the source until $r_{\mathrm{S}} = r_{\mathrm{CO}}$. This is possible since the theorem guarantees a linear processing of the source exists that can reduce r_{CO} without changing $(c_{\mathrm{S}}, r_{\mathrm{S}})$ whenever $r_{\mathrm{S}} < r_{\mathrm{CO}}$. For the proof of the theorem, see [15, Appendix C]. ■

For the running example in Section III, the omniscience scheme does not achieve r_{S}, i.e., $r_{\mathrm{S}} < r_{\mathrm{CO}}$, and so the theorem above guarantees a linear processing of the source that reduces r_{CO} without changing $(c_{\mathrm{S}}, r_{\mathrm{S}})$. Such a linearly reduced source can be obtained with

$$\mathbf{z}'_3 = \mathbf{z}_3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathsf{x}_1 \oplus \mathsf{x}_2 \oplus \mathsf{x}_3 \in \langle\!\langle \mathbf{z}_3 \rangle\!\rangle$$

and $\mathbf{z}'_i = \mathbf{z}_i$ for $i \in 1, 2, 4$. It is straightforward to show that $c_{\mathrm{S}}(\mathbf{z}'_V) = 1$ by (16) and $r_{\mathrm{CO}}(\mathbf{z}'_V) = 2$ by (17), and the source is reduced in the sense that $H(\mathbf{z}'_V) = 3 < H(\mathbf{z}_V)$. By going through all possible independent linear processings

of $\mathbf{z}_i$'s, which is possible as there is only a finite number of possibilities, one can show that the above defined $\mathbf{z}'_V$ is optimal to (22), achieving the minimum $r_{\rm CO}$. Hence, $r_{\rm S} = 2$ as desired by the above corollary.

V. Extensions

In this work, we considered the one-shot secret key agreement problem under a finite linear source model with linear public discussion, perfect secrecy and recoverability. Indeed, due to the linearity of the source mode, it is plausible that the public discussion is linear without loss of optimality, i.e., non-linear discussion cannot improve the secrecy capacity. It is straightforward to show Theorem 2 without requiring linear discussion, as the converse parts follow from those of the asymptotic model without requiring the discussion to be linear. However, extending Theorem 1 and Theorem 3 appears challenging and we have to resort to techniques that rely on the linearity of the discussion to prove the results.

Another challenge is to extend Theorem 1 and Theorem 3 to the asymptotic model where users observe $n \geq 1$ i.i.d. samples $\mathbf{z}_V^n := \begin{bmatrix}\mathbf{z}_{V1} & \dots & \mathbf{z}_{Vn}\end{bmatrix}$ of the private source, and the constrained secrecy capacity and discussion rate is per sample of the observation, i.e.,

$$C_{\rm S}(R) := \max\left\{ \frac{\log|K|}{n} \,\middle|\, \frac{\ell(\mathbf{f}_V)}{n} \leq R \right\}.$$

The recoverability (4) and secrecy (5) constraints can also be relaxed to the asymptotic versions in [3], i.e., for some positive $\delta_n, \epsilon_n \to 0$ as $n \to \infty$,

$$\begin{aligned} &\frac{1}{n}\log|K| - H(\mathsf{K}|\mathbf{f}_V) \leq \delta_n \\ &\Pr\left(\exists i \in V, \mathsf{K} \neq \phi_i(\mathbf{u}, \mathbf{z}_i^n, \mathbf{f}_V)\right) \leq \epsilon_n \end{aligned}$$

for a sequence in n of secret key agreement schemes and some functions ϕ_i for $i \in V$ that user i uses to recover the secret key. As mentioned below Theorem 2, the characterizations of $C_{\rm S}(0)$, $C_{\rm S}$ and $R_{\rm CO}$ are known for the asymptotic model and they are indeed used to derive the corresponding characterizations for the one-shot model. We believe that the other results in Theorem 1 and Theorem 3 can also be extended. The current proofs can be directly extended if we impose perfect recoverability instead, i.e., with $\epsilon_n = 0$ for sufficiently large n. However, the proofs without assuming perfect recoverability remain elusive. What we desire is a proof that perfect recoverability is admissible and can therefore be assumed without loss of optimality. In a similar vein, we also desire a proof that $R_{\rm S}$ can be achieved exactly, i.e., for sufficiently large n, there exists a secret key agreement scheme with $\frac{1}{n}\log|K| = C_{\rm S}$ and $\frac{\ell(\mathbf{f}_V)}{n} = R_{\rm S}$, and that linear public discussion is admissible.

References

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

[2] R. Ahlswede and I. Csiszár, "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.

[3] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004.

[4] C. Chan and L. Zheng, "Mutual dependence for secret key agreement," in *Proceedings of 44th Annual Conference on Information Sciences and Systems*, 2010.

[5] H. Tyagi, "Common information and secret key capacity," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5627–5640, Sep. 2013.

[6] J. Liu, P. Cuff, and S. Verdú, "Secret key generation with limited interaction," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7358–7381, Nov. 2017.

[7] M. Mukherjee, N. Kashyap, and Y. Sankarasubramaniam, "On the public communication needed to achieve sk capacity in the multiterminal source model," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3811–3830, July 2016.

[8] C. Chan, M. Mukherjee, N. Kashyap, and Q. Zhou, "Secret key agreement under discussion rate constraints," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, June 2017, pp. 1519–1523.

[9] ——, "On the optimality of secret key agreement via omniscience," *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 3811–3830, April 2018.

[10] ——, "Upper bounds via lamination on the constrained secrecy capacity of hypergraphical sources," *IEEE Transactions on Information Theory*, pp. 1–1, 2019.

[11] S. Nitinawarat and P. Narayan, "Perfect omniscience, perfect secrecy, and Steiner tree packing," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6490–6500, Dec. 2010.

[12] C. Chan, "Linear perfect secret key agreement," in *Information Theory Workshop (ITW), 2011 IEEE*. IEEE, 2011, pp. 723–726.

[13] T. A. Courtade and T. R. Halford, "Coded cooperative data exchange for a secret key," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3785–3795, July 2016.

[14] C. Chan, "Delay of linear perfect secret key agreement," in *Forty-Ninth Annual Allerton Conference on Communication, Control, and Computing*, Sep. 2011.

[15] C. Chan, N. Kashyap, P. K. Vippathalla, and Q. Zhou, "One-shot perfect secret key agreement for finite linear sources," *CoRR*, vol. abs/1901.05817, 2019. [Online]. Available: http://arxiv.org/abs/1901.05817

[16] N. Milosavljevic, S. Pawar, S. E. Rouayheb, M. Gastpar, and K. Ramchandran, "Efficient algorithms for the data exchange problem," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1878 – 1896, Apr. 2016.

[17] C. Chan, "Generating secret in a network," Ph.D. dissertation, Massachusetts Institute of Technology, 2010.

[18] C. Chan, M. Mukherjee, N. Kashyap, and Q. Zhou, "Multiterminal secret key agreement at asymptotically zero discussion rate," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 2654–2658.

[19] N. Milosavljevic, S. Pawar, S. El Rouayheb, M. Gastpar, and K. Ramchandran, "Deterministic algorithm for the cooperative data exchange problem," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, Jul. 2011.

[20] C. Chan, M. Mukherjee, P. K. Vippathalla, and Q. Zhou, "Multiterminal Secret Key Agreement with Nearly No Discussion," *arXiv e-prints*, p. arXiv:1904.11383, Apr 2019.

[21] N. Ding, C. Chan, Q. Zhou, R. A. Kennedy, and P. Sadeghi, "Determining optimal rates for communication for omniscience," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1919–1944, Mar. 2018.