Group Secret Key Generation over Broadcast Erasure Channels

Mahdi Jafari Siavoshani, Christina Fragouli, Suhas Diggavi, Uday Pulleti, Katerina Argyraki
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Switzerland

***Abstract*—We consider a group of m trusted nodes that aim to create a shared secret key $\mathcal{K}$ over a wireless channel in the presence an eavesdropper Eve. We assume an erasure 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 public channel which is observed by Eve. For this setup we characterize the secret key generation capacity and propose an achievability scheme that is computationally efficient and employs techniques from network coding. Surprisingly, whether we have $m = 2$ nodes, or an arbitrary number m of nodes, we can establish a shared secret key among them at the same rate, independently of m.**[1]

I. Introduction

We consider the problem of generating a secret key $\mathcal{K}$ among $m \geq 2$ honest nodes that communicate over a wireless channel in the presence of a passive eavesdropper, Eve. We restrict our attention to the case where communication occurs either through a broadcast erasure channel, where erasures are independent among all receivers of the broadcast transmissions (including Eve), or, through a no-cost public channel.

We characterize the *secret key generation* capacity and propose a computationally efficient achievability scheme that employs techniques from network coding. Surprisingly, we show that the rates at which we can generate a secret key among the m nodes, is not affected by the number of nodes m; that is, whether we try to establish a secret key between two nodes, or an arbitrary number, we can do this at the same rate. This result is reminiscent of the main theorem in network coding, where a source can multicast information to a set of receivers at the same rate, independently of the number of receivers [1].

Secret key generation over wireless channels is a problem that has attracted significant interest. In a seminal paper on "wiretap" channels, Wyner [2] 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 [3] 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. The problem of key agreement between a set of terminals with access to noisy broadcast channel and public discussion channel (visible to the eavesdropper) was studied in [6], where some achievable secrecy rates were established, assuming Eve does not have access to the noisy broadcast transmissions. This was generalized in [7], [8] by developing (non-computable) outer bounds for secrecy rates. To the best of our knowledge, ours is the first work to consider multi-terminal secret key agreement over erasure networks, when Eve also has access to the noisy broadcast transmissions. Moreover, unlike the information-theoretic works in [2], [3], [4], [6], [8] that assume infinite complexity operations, our scheme is computationally efficient.

The paper is organized as follows. In §II we introduce our notation and the problem formulation. In §III we describe our achievability scheme and analyze its performance. In §IV we state an upper bound for the secret key generation capacity and show that it matches the achievable rates.

II. Notation and Setup

A. Notation

We use uppercase letters (*e.g.*, X) to represent random variables (or more generally random objects). Every random variable in this paper takes values in a finite set. Given random variables $X_1, \ldots, X_m$, we write $X_{1:m}$ to denote $(X_1, \ldots, X_m)$. We also use $X^{1:t}$ or more simply X^t to denote $(X[1], \ldots, X[t])$. Bold capital letters (*e.g.*, $\boldsymbol{A}$) are reserved for deterministic matrices. Finally, we use $\mathsf{Uni}(\mathbb{F}_q^L)$ to denote the uniform distribution over vectors of length L that are defined over finite field $\mathbb{F}_q$.

B. Problem Statement

We consider a set of $m \geq 2$ honest nodes, $T_0, \ldots, T_{m-1}$ (T for "terminal") that aim to share a secret key $\mathcal{K}$ among themselves while keeping it concealed from a passive adversary, Eve. Eve does not perform any transmissions, but is trying to eavesdrop on (overhear) the communications between the honest nodes. For convenience, sometimes we will refer to node T_0 as "Alice."

We assume that Alice has access to an erasure broadcast channel such that the rest of the terminals (including Eve) receive independent erased version of what she broadcasts. The input and output symbols to the erasure channel are packets of length L of elements from a finite filed $\mathbb{F}_q$. When

[1]This work was funded in part by the Swiss National Science Foundation Award No PP00P2128639, and the ERC Starting Grant Project NOWIRE ERC-2009-StG-240317.

Alice transmits a packet, node T_i correctly overhears it with probability $1-\delta_i$, where δ_i is the "erasure probability" of the channel. Similarly, Eve correctly receives the packets with probability $1-\delta_E$. The erasure events happen independently over time and across different channels. For simplicity in this paper we will focus on the symmetric case where we have $\delta_i = \delta$.

We also assume that all of the honest terminals can discuss over a cost-free public channel where everybody (including Eve) can hear the discussion.

In the following we define a protocol that abstracts the interactive communication between terminals that aim to share a common secret key $\mathcal{K}$ (see also [3], [4], [6], [8]).

***Definition** 1:* The *secret key generating protocol* is defined as follows:

1) For $t=0$, all of the honest terminals generate independent random variables $W_0, \ldots, W_{m-1}$.
2) (i) For time $1 \le t \le n$, Alice transmits $X_0[t]$ over the broadcast channel where
$$X_0[t] = X_{0,t}\left(W_0, \mathcal{D}^{t-1}\right).$$
We will define the random variables $\mathcal{D}[t]$ in the following. Then the other terminals receive $X_1[t], \ldots, X_{m-1}[t]$ and Eve receives $X_E[t]$.
(ii) Following each of the broadcast transmissions, there is the possibility for the terminals to discuss over a cost-free public channel. This discussion continues for $r[t]$ rounds and is represented by the random variables $\mathcal{D}[t] = \left(\mathcal{D}_0[t], \ldots, \mathcal{D}_{r[t]}[t]\right)$, where
$$\mathcal{D}_i[t] = \mathcal{D}_{i,t}\left(W_j, X_j^t, \mathcal{D}^{t-1}, \mathcal{D}_{0:i-1}[t]\right)$$
is the public message revealed by the jth terminal with $j = i \mod m$ (in other words, the indexing of the discussion is done in a round robin order).
3) Finally, the ith terminal creates a key $\mathcal{K}_i$ where
$$\mathcal{K}_i = \mathcal{K}_i(W_i, X_i^n, \mathcal{D}^n).$$

***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
$$\mathbb{P}\left[\mathcal{K}_i \neq \mathcal{K}_j\right] < \epsilon, \quad \forall i,j : i \neq j, \tag{1}$$
$$I(\mathcal{K}_0; X_E^n, \mathcal{D}^n) < \epsilon, \tag{2}$$
and
$$\frac{1}{n} H(\mathcal{K}_0) > R_s - \epsilon. \tag{3}$$
The supremum of the achievable key rate as $n \to \infty$ and $\epsilon \to 0$ is called the *key generation capacity* C_s.

III. Lower Bound for the Key Generation Capacity

Here we describe and analyze our achievability scheme.

Private Phase:

1) Alice broadcasts n packets, $x_1, \ldots, x_n$, where $x_i \in \mathbb{F}_q^L$ and $x_i \sim \mathsf{Uni}(\mathbb{F}_q^L)$ (we will call them "x-packets"). Of these, n^* packets are received by at least one honest node. This set is denoted by N^* where $n^* = |N^*|$.

Public Discussion (Initial Phase):

1) Each honest node sends Alice publicly a feedback message specifying which x-packets it received. Let I_{T_i} denotes the set of packets' indices received by the ith terminal T_i.
2) Alice constructs $h = \delta_E \cdot n^*$ linear combinations of the x-packets, $y_1, \ldots, y_h$ (we will call them "y-packets"), as follows:
(i) She divides the set N^* of x-packets that were received by at least one honest node into non-overlapping subsets, such that each subset consists of all the packets that were commonly received by a different subset of honest nodes. To be more precise, let S be an arbitrary non-empty subset of $\{1, \ldots, m-1\}$ and let us define the set
$$N_{S,\overline{S}} \triangleq \{x_i | i \in I_{T_j} : \forall j \in S, \text{ and } i \notin I_{T_j} : \forall j \notin S\}.$$
Then we have
$$N^* = \bigcup_{\emptyset \neq S \subseteq \{1,\ldots,m-1\}} N_{S,\overline{S}}.$$
(ii) From each such subset of packets $N_{S,\overline{S}}$, she creates $\delta_E \cdot n_{S,\overline{S}}$ linear combinations using the construction provided in Lemma 2 (provided in the Appendix), where $n_{S,\overline{S}} \triangleq |N_{S,\overline{S}}|$.
Then she publicly reveals the coefficients she used to create all the y-packets.
3) Each node T_i reconstructs as many (say h_i) of the y-packets as it can (based on the x-packets it received in step #1). For h_i we can write
$$h_i = \sum_{\substack{\emptyset \neq S \subseteq \{1,\ldots,m-1\}: \\ i \in S}} \delta_E \cdot n_{S,\overline{S}}.$$

Public Discussion (Reconciliation Phase):

1) Alice creates $h - \min_i h_i$ linear combinations of the y-packets (we will call them "z-packets"), using the construction provided in Lemma 3 (provided in the Appendix). She publicly reveals both the contents and the coefficients of the z-packets, such that each node T_i receives at least $h - h_i$ of them.
2) Each node T_i combines the $h - h_i$ z-packets it received with the h_i y-packets it recreated in phase 1, and reconstructs all the y-packets.
3) Alice creates $l = \min_i h_i$ linear combinations of the y-packets, $k_1, \ldots, k_l$ (we will call them "k-packets"), using the construction provided in Lemma 4 (provided in the Appendix). She publicly reveals the coefficients she used to create all the k-packets.
4) Each node T_i reconstructs all the k-packets. The common secret key is the concatenation of all the k-packets, $\mathcal{K} = \langle k_1, \ldots, k_l \rangle$.

Now we may summarize the above achievability scheme as follows based on Definition 1. At $t = 0$ Alice generates the random variable $W_0 = X_0^n$ where $X_0[i] \sim \mathsf{Uni}(\mathbb{F}_q^L)$. We have also $W_{1:m-1} = \emptyset$. For each time t, $1 \le t < n$, she broadcasts $X_0[t]$ and there is no public discussions afterwards; namely we have $\mathcal{D}[t] = \emptyset$. After the nth transmission by Alice there is a public discussion in many rounds; simplified as follows. We have $\mathcal{D}[n] = (P_1, P_2, P_3, P_4)$ where P_1 denotes the set of indices I_{T_i} that have been sent back by the honest terminals, P_2 denotes the coefficients of the y-packets, P_3 denotes the z-packets and their coefficients, and finally P_4 represents the coefficients of k-packets.

***Theorem** 1:* The achievable secret key generation rate of the above scheme is

$$R_s = (1-\delta)\delta_E \left(L \log_2 q\right).$$

Proof: The way that the achievability scheme is proposed, constructively satisfies Condition (1).

To prove Condition (3) we proceed as follows. Let us define $\bar{l} \triangleq l/n$. Then we can write

$$\begin{aligned} H(\mathcal{K}) &= H(\mathcal{K}, \bar{l}) = H(\mathcal{K}|\bar{l}) + H(\bar{l}) \\ &\ge H(\mathcal{K}|\bar{l}) \\ &= H(\mathcal{K}|\alpha < \bar{l} < \beta) \cdot \mathbb{P}\left[\alpha < \bar{l} < \beta\right] \\ &\quad + H(\mathcal{K}|\bar{l} \ge \beta) \cdot \mathbb{P}\left[\bar{l} \ge \beta\right] + H(\mathcal{K}|\bar{l} \le \alpha) \cdot \mathbb{P}\left[\bar{l} \le \alpha\right] \\ &\ge H(\mathcal{K}|\alpha < \bar{l} < \beta) \cdot \mathbb{P}\left[\alpha < \bar{l} < \beta\right] \\ &\ge n\alpha \left(L \log_2 q\right) \left[1 - \mathbb{P}\left[\bar{l} \le \alpha\right] - \mathbb{P}\left[\bar{l} \ge \beta\right]\right], \end{aligned}$$

where $\alpha = \mu - \delta$ and $\beta = \mu + \delta$ for $0 < \delta \le \mu$. Then by applying the concentration result of Lemma 5 (see Appendix) we have

$$\mathbb{P}\left[\bar{l} \le \alpha\right] \le (m-1) \exp\left(-\frac{\delta^2}{2\mu} n\right) \triangleq a,$$

and

$$\mathbb{P}\left[\bar{l} \ge \beta\right] \le \exp\left(-\frac{(m-1)\delta^2}{3\mu} n\right) \triangleq b.$$

So we can write

$$\frac{1}{n} H(\mathcal{K}) > R_s - \epsilon,$$

where $R_s = \mu(L \log_2 q)$, $\mu = (1-\delta)\delta_E$, and

$$\epsilon = \mu \left(L \log_2 q\right)[a+b] + \delta \left(L \log_2 q\right)[1-a-b].$$

Then we have $\epsilon \to 0$ if $\delta \to 0$.

To prove Condition (2) we need to show that

$$I(K; X_E^n, P_1, P_2, P_3, P_4) < \epsilon.$$

By using a similar technique that we used above to bound $H(\mathcal{K})$, (using Lemma 2 and some concentration results for $n_{S,\overline{S}}$) we can show that

$$I(Y; X_E^n, P_1, P_2) < \epsilon. \tag{4}$$

Using Lemma 4, by construction we have also

$$I(K; P_3, P_4) = 0. \tag{5}$$

Now we know that the coefficients of the z-packets and k-packets form a basis (see Lemma 4) so the random variable Y and the random variable (K, P_3, P_4) are equivalent (having one we have the other). Then we can write (4) as follows

$$\begin{aligned} I(Y; X_E^n, P_1, P_2) &= I(K, P_3, P_4; X_E^n, P_1, P_2) \\ &= I(P_3, P_4; X_E^n, P_1, P_2) \\ &\quad + I(K; X_E^n, P_1, P_2 | P_3, P_4) < \epsilon, \end{aligned}$$

so

$$I(K; X_E^n, P_1, P_2 | P_3, P_4) < \epsilon. \tag{6}$$

Now we can expand

$$\begin{aligned} I(K; X_E^n, P_1, P_2, P_3, P_4) &= I(K; P_3, P_4) \\ &\quad + I(K; X_E^n, P_1, P_2 | P_3, P_4), \end{aligned}$$

where the first term is zero by (5) and second term is very small because of (6), so we are done. ■

IV. Upper Bound for the Key Generation Capacity

In this section we use some of the notation introduced in [5], [6]. In [5], Csiszar and Narayan consider the secrecy capacity among m terminals who have access to a multi-terminal correlated source where the ith terminal observes random variable X_i. The terminals also can discuss over a cost-free public channel. They assume that a subset $A \subseteq M$ of terminals want to share a common key which is possibly required to be concealed from a subset $D \subseteq A^c$ where $M = \{0, \cdots, m-1\}$ and also from an eavesdropper who listens to all the public discussions. For this setup the secrecy capacity is referred to as *private key* capacity $C_P^{\mathsf{sr}}(A, M|D)$ and for $D = \emptyset$ as *secret key* capacity $C_S^{\mathsf{sr}}(A, M)$.

They also consider another setup where there exists a broadcast channel from T_0 to the rest of terminals and the possibility of discussion over a public channel for all terminals [6]. As before, they assume that a subset $A \subseteq M$ of terminals want to share a common key which is possibly required to be concealed from a subset $D \subseteq A^c$ and from an eavesdropper who listens to all the public discussions (but she doesn't have any observation from the broadcast channel output). Let us denote the secrecy capacity in this case by $C_P^{\mathsf{ch}}(A, M|D)$ and for $D = \emptyset$ by $C_S^{\mathsf{ch}}(A, M)$. Then [6, Theorem 4.1] relates the source model and the channel model as follows.

***Theorem** 2 ([6, Theorem 4.1]):* The PK (private key) capacity $C_P^{\mathsf{ch}}(A, M|D)$ for $A \subseteq M$ with privacy from a set of terminals $D \subseteq A^c$ is equal to the maximum over of PK capacity of the corresponding emulated source model. Specifically

$$C_P^{\mathsf{ch}}(A, M|D) = \max_{P_{X_0}} C_P^{\mathsf{sr}}(A, M|D).$$

The same result holds for SK (secret key) capacity by setting $D = \emptyset$. In the source model the distribution of multi-terminal source is

$$P_{X_0, \ldots, X_{m-1}}(x_1, \ldots, x_m) = P_{X_0}(x_0) Q(x_1, \ldots, x_{m-1} | x_0),$$

where $Q(x_1,\ldots,x_{m-1}|x_0)$ is the transfer probability of the broadcast channel.

In all the previous models the eavesdropper does not have access to any side information from the multi-terminal source or from the output of the broadcast channel. However, what we are interested here is the case where she has also some side information shown by random variable X_E. In these cases, for the channel setup, the secrecy capacity is denoted by $C^{\mathsf{ch}}_{WP}(A,M|D)$ and $C^{\mathsf{ch}}_{WS}(A,M)$ in which "W" stands for wiretap. The capacity $C^{\mathsf{ch}}_{WS}(M,M)$ matches to the capacity defined in Definition 2. Finding $C^{\mathsf{ch}}_{WP}(A,M|D)$ for general broadcast channels is still an open problem but [6, Lemma 5.1] gives an upper bound for $C^{\mathsf{ch}}_{WP}(A,M|D)$ and $C^{\mathsf{ch}}_{WS}(A,M)$ as follows.

***Lemma** 1 ([6, Lemma 5.1]):* The PK (private key) capacity $C^{\mathsf{ch}}_{WP}(A,M|D)$, or the SK (secret key) capacity $C^{\mathsf{ch}}_{WS}(A,M)$ if $D=\emptyset$, for the model with *wiretap side information* (WSI) (which means that Eve observes a random variable X_E from the broadcast channel output) is bounded above by the PK capacity of the associated model without WSI. The associated model contains an extra terminal T_m which has access the the eavesdropper's side information and participates in secrecy generation. So the set A does not change but the set D becomes $D\cup\{m\}$.

***Corollary** 1:* Using Lemma 1 and Theorem 2 we may write

$$\begin{aligned} C^{\mathsf{ch}}_{WS}(M,M) &\le C^{\mathsf{ch}}_{P}\left(M,M\cup\{m\}|\{m\}\right) \\ &= \max_{P_{X_0}} C^{\mathsf{sr}}_{P}\left(M,M\cup\{m\}|\{m\}\right). \end{aligned}$$

In this work we are interested in an independent erasure broadcast channel where

$$Q(x_1,\ldots,x_{m-1},x_E|x_0)=q_E(x_E|x_0)\prod_{j=1}^{m-1}q_j(x_j|x_0),$$

and each $q_j(\cdot|\cdot)$ is an erasure channel with parameter δ and $q_E(\cdot|\cdot)$ is an erasure channel with parameter δ_E. As it is explained in [5, Example 7] the input and output random variables of this channel form a Markov chain on a tree (for definitions and notation refer to [5, Section V]). So as it explained in that example for this special case we have

$$C^{\mathsf{sr}}_{P}(A,M|D)=\min_{(i,j)\in E(G(A))} I(X_i;X_j|X_D),$$

where $G(A)$ denotes the smallest sub-tree of G whose vertex set contains A. For the specific problem in this paper $G(A)$ is a star rooted at Alice (terminal T_0).

Then, combining Corollary 1 and [5, Example 7] (as discussed above) we can conclude Theorem 3 as follows.

***Theorem** 3:* The key generation capacity defined in Definition 2 can be upper bounded as follows

$$\begin{aligned} C_s &\le \max_{p(x_0)}\min_{j\in M/\{0\}} I(X_0,X_j|X_E) \\ &\le (1-\delta)\delta_E\left(L\log_2 q\right). \end{aligned}$$

Proof: As mentioned before, the first upper bound mentioned in the theorem is obtained by applying Corollary 1 and [5, Example 7]. To obtain the second bound, we can write

$$\begin{aligned} I(X_0,X_j|X_E) &= H(X_0|X_E)-H(X_0|X_E,X_j) \\ &= \left[\delta_E-\delta_E\delta\right]H(X_0) \\ &\le (1-\delta)\delta_E\left(L\log_2 q\right), \end{aligned}$$

where X_j, X_E are random variables which are "erased" versions of X_0, with erasure probabilities δ and δ_E respectively. This concludes the theorem. ∎

Combining Theorem 1 and Theorem 3 we have our main result.

***Corollary** 2:* The key generation capacity among m terminals stated in §II-B is

$$C_s=(1-\delta)\delta_E\left(L\log_2 q\right),$$

where δ is the erasure probability from Alice to the rest of terminals and δ_E is the erasure probability from Alice to Eve.

REFERENCES

[1] R. Ahlswede, N. Cai, S-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, Jul. 2000.

[2] A. D. Wyner, "The wire-tap channel," *Bell System Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.

[3] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, pp. 733–742, 1993.

[4] 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.

[5] I. Csiszar and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, Dec. 2004.

[6] I. Csiszar and P. Narayan, "Secrecy capacities for multiterminal channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2437–2452, Jun. 2008.

[7] 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.

[8] A. A. Gohari and V. Anantharam, "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.

[9] M. Mitzenmacher and E. Upfal, "Probability and computing, randomized algorithm and probabilistic analysis," *Cambridge University Press*, 2006.

[10] F. J. Macwilliams and N. J. A. Sloane, "The theory of error correcting codes," *North-Holland*, 2006.

[11] C. Fragouli and E. Soljanin, "Network coding fundamentals, foundations and trends in networking," *Now Publishers*, Jun. 2007.

APPENDIX

***Lemma** 2:* Consider a set of n packets $x_1,\ldots,x_n$, $x_i\in\mathbb{F}_q^L$, where $x_i\sim\mathsf{Uni}(\mathbb{F}_q^L)$ and all packets x_i are independent from each other. Assume that Eve has overheard n_E of these packets. Call the packets Eve has $w_1,\ldots,w_{n_E}$. Then it is possible to create $n'=n-n_E$ linear combinations of the $x_1,\ldots,x_n$ packets over the finite field $\mathbb{F}_q$, say $y_1,\ldots,y_{n'}$, in polynomial time, so that these are secure from Eve, *i.e.*,

$$I(y_1,\ldots,y_{n'};w_1,\ldots,w_{n_E})=0.$$

The same result holds with high probability (of order $1-O(q^{-1})$) if the linear combinations are selected uniformly at random over $\mathbb{F}_q$.

Proof: Construct matrix X that has as rows the packets $x_1,\ldots,x_n$. Similarly, construct matrices Y and W that have as rows the packets $y_1,\ldots,y_{n'}$ and $w_1,\ldots,w_{n_E}$.

Note that because the packets $w_1, \ldots, w_{n_E}$ are by definition a subset of the packets $x_1, \ldots, x_n$, we can write $W = \boldsymbol{A}_E X$, with $\boldsymbol{A}_E \in \mathbb{F}_q^{n_E \times n}$ that has zeros and ones as elements. We will also construct the y packets as linear combinations of the x packets over a field $\mathbb{F}_q$. We will then have that $Y = \boldsymbol{A}X$, where $\boldsymbol{A} \in \mathbb{F}_q^{(n-n_E)\times n}$ is the matrix we are interested in designing. Thus we can write

$$\left[\begin{array}{c} Y \\ W \end{array} \right] = \left[\begin{array}{c} \boldsymbol{A} \\ \boldsymbol{A}_E \end{array} \right] X.$$

We now proceed by expanding $H(Y|W)$. We have

$$H(Y|W) = H(Y,W) - H(W) =$$
$$= \left[\mathrm{rk}\left(\boldsymbol{B}\right) - \mathrm{rk}(\boldsymbol{A}_E)\right] L \log_2 q = \left[\mathrm{rk}\left(\boldsymbol{B}\right) - n_E\right] L \log_2 q,$$

where $\boldsymbol{B} = \left[\begin{array}{c} \boldsymbol{A} \\ \boldsymbol{A}_E \end{array} \right]$ and L is the length of each packet x_i. Now the only way that we have $H(Y|W) = H(Y)$ is that $\boldsymbol{B}$ becomes a full rank matrix.

Using coding theory we will construct such a matrix $\boldsymbol{B}$, *without* knowing $\boldsymbol{A}_E$. All we know is that in each row of $\boldsymbol{A}_E$ there is only one 1 and the remaining elements are zero; so all of the vectors in the row span of $\boldsymbol{A}_E$ have Hamming weight less than or equal to n_E. Now, if we choose $\boldsymbol{A}$ to be a generator matrix of an maximum distance separable (MDS) linear code with parameters $[n, n-n_E, n_E+1]_q$ then each codeword has Hamming weight larger than or equal to n_E+1 [10]. So the row span of $\boldsymbol{A}$ and $\boldsymbol{A}_E$ are disjoint (except for the zero vector) and the matrix $\boldsymbol{B}$ becomes full-rank for all of matrices $\boldsymbol{A}_E$ that have the aforementioned structure. For example, we may select to use a generator matrix of a Reed-Solomon code [10], which is an MDS code, over a field of size $q = n+1$.

To prove the second assertion of the lemma, we note that creating vectors y_i uniformly at random is equivalent to selecting the elements of matrix $\boldsymbol{A}$ independently uniformly at random from the field $\mathbb{F}_q$. In this case we can write

$$\mathbb{P}\left[\boldsymbol{B} \text{ is full-rank}\right] = \frac{(q^n - q^{n_E}) \cdots (q^n - q^{n-1})}{q^{n(n-n_E)}}$$
$$= 1 - O(q^{-1}),$$

which goes to 1 as q increases. ∎

Lemma 3: Consider packets $y_1, \ldots, y_h$ and assume that each one of $m-1$ receivers has observed a different subset of these packets of size l. We can find $h-l$ linear combinations of the y packets, say $z_1, \ldots, z_{h-l}$ such that, each receiver can use its observations and the z packets to decode all the y packets.

Proof: This is a standard problem formulation in the network coding literature, and any of the standard polynomial-time approaches for network code design can be used [11]. ∎

Lemma 4: Consider a set of h packets $y_1, \ldots, y_h$ where $y_i \sim \mathsf{Uni}(\mathbb{F}_q^L)$ and assume that an eavesdropper Eve has overheard linear combinations of $h-l$ of these packets. Call the packets Eve has $z_1, \ldots, z_{h-l}$. Then it is possible to create l linear combinations of the $y_1, \ldots, y_h$ packets, say $k_1, \ldots, k_l$, in polynomial time, so that these are secure from Eve, *i.e.*,

$$I(k_1, \ldots, k_l; z_1, \ldots, z_{h-l}) = 0.$$

The same result holds with high probability (probability of order $1 - O(q^{-1})$) if the l packets k_i are created uniformly at random over $\mathbb{F}_q$.

Proof: Similarly to the proof of Lemma 2, let Y, Z and K be matrices that have as rows the packets $y_1, \ldots, y_h$, $z_1, \ldots, z_{h-l}$ and $k_1, \ldots, k_l$. We can then write

$$\left[\begin{array}{c} K \\ Z \end{array} \right] = \left[\begin{array}{c} \boldsymbol{A}_K \\ \boldsymbol{A}_Z \end{array} \right] Y,$$

where $\boldsymbol{A}_Z$ is a given known matrix, since we know the transmitted linear combinations, and we seek a matrix $\boldsymbol{A}_K$ such that, the matrix $\left[\begin{array}{c} \boldsymbol{A}_K \\ \boldsymbol{A}_Z \end{array} \right]$ is full rank. Equivalently, we seek vectors $k_1, \ldots, k_l$ that together with $z_1, \ldots, z_{h-l}$ form a basis; we can do this using any of standard methods, such as Gram-Schmidt orthogonalization. ∎

Lemma 5: The value of the parameter l in Theorem 1 converges exponentially fast in n to its expected value.

Proof: Let us consider the *random* variables h, h_i, and l defined in §III. For convenience, we will work with the normalized random variables $\overline{h} \triangleq h/n$, $\overline{h}_i \triangleq h_i/n$, and $\bar{l} \triangleq l/n$. Let us also define the random variable $\eta_j^{(i)}$ as follows

$$\eta_j^{(i)} = \begin{cases} 1 & \text{if the } j\text{th } x\text{-packet is received} \\ & \text{by } T_i \text{ but not by Eve,} \\ 0 & \text{otherwise.} \end{cases}$$

Then we can write $\overline{h}_i = \frac{1}{n}\sum_{j=1}^n \eta_j^{(i)}$ and we have $\mu = \mu_i \triangleq \mathbb{E}\left[\overline{h}_i\right] = (1-\delta)\delta_E$. As defined before, we have also $\bar{l} = \min_i \overline{h}_i$.

To bound $\bar{l}$, first observe that

$$\mathbb{E}\left[\bar{l}\right] = \mathbb{E}\left[\min_i \frac{1}{n}\sum_{j=1}^n \eta_j^{(i)}\right] = (1-\delta)\delta_E = \mu.$$

Then for $0 < \delta \leq \mu$ we can write

$$\mathbb{P}\left[\bar{l} \geq \mu + \delta\right] = \mathbb{P}\left[\overline{h}_i \geq \mu + \delta : \forall i\right]$$
$$= \mathbb{P}\left[\overline{h}_1 \geq \mu + \delta\right]^{(m-1)}$$
$$\leq \exp\left(-\frac{(m-1)\delta^2}{3\mu} n\right),$$

where in the last inequality we use Chernoff bound [9, Chapter 4]. On the other hand we can also write for $0 < \delta \leq \mu$

$$\mathbb{P}\left[\bar{l} \leq \mu - \delta\right] \leq (m-1)\mathbb{P}\left[\overline{h}_1 \leq \mu - \delta\right]$$
$$\leq (m-1)\exp\left(-\frac{\delta^2}{2\mu} n\right),$$

so we are done. ∎