Secure Network Coding in the Setting in Which a Non-Source Node May Generate Random Keys

Debaditya Chaudhuri
University at Buffalo
Email: debadity@buffalo.edu

Michael Langberg
University at Buffalo
Email: mikel@buffalo.edu

Michelle Effros
California Institute of Technology
Email: effros@caltech.edu

***Abstract*—It is common in the study of secure multicast network coding in the presence of an eavesdropper that has access to z network links, to assume that the source node is the only node that generates random keys. In this setting, the secure multicast rate is well understood. Computing the secure multicast rate, or even the secure unicast rate, in the more general setting in which all network nodes may generate (independent) random keys is known to be as difficult as computing the (non-secure) capacity of multiple-unicast network coding instances — a well known open problem. This work treats an intermediate model of secure unicast in which only one node can generate random keys, however that node need not be the source node. The secure communication rate for this setting is characterized again with an eavesdropper that has access to z network links.**

I. Introduction

In this work, we study secure network communication over a directed acyclic network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ having a single source node S, a single terminal node T, and a single node K, which is capable of generating random "keys" independent of the messages generated by S. We employ a notion of secure "wiretap" communication networks introduced by Cai and Yeung in [1] and studied further in, for example [2]–[6]. Under this notion of security, given a communication scheme over $\mathcal{G}$, we consider an edge $e \in \mathcal{E}$ of the network to be secure in the presence of a wiretap adversary if and only if $I(M; X_e) = 0$, where M denotes the source message and X_e denotes the information communicated on edge e.[1] To be secure in the presence of an adversary that wiretaps any size-z subset $\mathcal{W} = \{e_1, \cdots, e_z\} \subset \mathcal{E}$ of edges, we require that $I(M; X_{\mathcal{W}}) = 0$, where $X_{\mathcal{W}} = (X_{e_1}, \cdots, X_{e_z})$.

Given integers R and z, we define a secure network code over the network $\mathcal{G}$ to be (R, z)-feasible if it allows information to be communicated from the source S to the terminal T at rate R and, in addition, it secures the network against a wiretap adversary that eavesdrops on up to z edges of the network. Our work entails determining, for each z, the closure of the set of rates that are (R, z)-feasible, thereby deriving the capacity-security region.

When $K = S$, the capacity-security region for secure multicast network codes is well understood [1], [2] with several follow up works [3]–[6] that address various methods to alter any given non-secure linear network code into a new code that is secure. In contrast, determining the capacity-security region for secure network codes over a single-source single-terminal network, where every node can generate random keys, is as hard as the problem of characterizing the (non-secure) capacity region of the k-unicast problem as shown by [7]. Results of a similar nature are also presented in [8]. The k-unicast problem is a well known open problem in the study of network codes [8]–[12].

In this work, we seek to make progress in the apparently difficult generalization from the scenario where only the source can generate random keys to the scenario where all nodes can generate keys by studying the case where only a single node can generate keys but allowing that single node to be arbitrary. Our central result is a characterization of the capacity-security region in the unicast (single-source single-terminal) setting when only a single network node $K \neq S \in \mathcal{V}$ can generate random keys.

The remainder of the paper is organized as follows. In Section II, we present our model and preliminary notation. Our main result, the capacity-security characterization of the networks at hand, appears in Section III. The characterization is combinatorial in nature and involves different cut-set bounds between the source node, the key generating node, and the terminal node. Achievability is proven in Section IV via a reduction from secure communication over $\mathcal{G}$ to (non-secure) multi-source multi-cast network coding over a modified network $\mathcal{G}^*$ as shown in Figure 1b. The converse proof, which is based on cutset bounds, appears in Section V. An additional converse proof, in the more general context of cyclic networks, is presented in Appendix A. The proofs of some one of our lemmas and claims are presented in Appendix B and Appendix C, respectively.

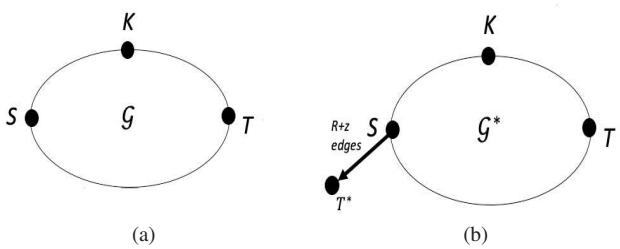
II. Network Model

Our system model consists of the following components:

(a) A finite directed acyclic graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$. We assume that each edge $e \in \mathcal{E}$ noiselessly transmits one unit of information (i.e., one field element in a given field $\mathbb{F}_q$) per unit time. We use multiple edges to model an edge with the ability to communicate more than one information symbol per unit time.

(b) A source node S, which generates a source message vector of length R, $M = \begin{bmatrix} M_1 & M_2 & \cdots & M_R \end{bmatrix}^T$, with $M_1, M_2, \cdots, M_R$ independently and uniformly distributed over the field $\mathbb{F}_q$ of size q.

[1] Detailed definitions of all concepts discussed here and below appear in Section II.



(a) (b)

Fig. 1: (a) Network model $\mathcal{G}$, and (b) the modified network $\mathcal{G}^*$ obtained from $\mathcal{G}$ by adding T^* and setting the demands at T and T^* to (M, N).

(c) A terminal node $T \in \mathcal{V}$, which is required to decode all the messages generated by the source S with zero error.
(d) A node $K \in \mathcal{V}$, which generates a random "key" vector, $N = \left[N_1, \cdots, N_{|N|}\right]^T$ with $N_1, \cdots, N_{|N|}$ independently and uniformly distributed over the field $\mathbb{F}_q$ with N independent of M.
(e) An eavesdropper that can access any subset $\mathcal{W} \subset \mathcal{E}$ of edges for which $|\mathcal{W}| \leq z$.

In the following subsections, we introduce our definition of a network code and discuss the notions of topological order and cut sets.

A. Network Code

We define a scalar linear network code $\mathcal{N}$ for the network $\mathcal{G}$ to be an assignment of a linear encoding function f_e to each edge $e \in \mathcal{E}$ and a linear decoding function g_T to terminal T. For $e \in \mathcal{E}$, we denote the edge message on e by X_e, and for any set $\mathcal{A} \subseteq \mathcal{E}$, we define $X_{\mathcal{A}} = \{X_e : e \in \mathcal{A}\}$. If $e \in \mathcal{E}$ and $e = (u, v)$ then the edge message X_e is a linear combination of all the messages carried by the edges in $\mathrm{In}(u) = \{(w, u) : (w, u) \in \mathcal{E}\}$, the incoming edges of u. The edge message at e is obtained using local encoding at u. We define X_e using the local encoding function $\bar{f}_e$ on $e = (u, v)$ as

$$X_e = \bar{f}_e(X_{\mathrm{In}(u)}) = \sum_{e' \in \mathrm{In}(u)} \bar{c}_{e',e} X_{e'}. \tag{1}$$

Here, X_e denotes the message on edge e, for each edge $e' \in \mathrm{In}(u)$, $X_{e'}$ denotes the messages on edges e' and $\bar{c}_{e',e}$ is the coefficient acting on each message $X_{e'}$. If edge e is an outgoing edge of S (or K), then X_e is a function of the source messages (or keys) as well. Given, such a network code, an adversary that wiretaps any size-z subset of edges $\mathcal{W} \subset \mathcal{E}$ would obtain the information $X_{\mathcal{W}}$ on the wiretapped edges. A network code is said to be (R, z)-feasible if

$$g_T(X_{\mathrm{In}(T)}) = M \tag{2}$$
$$\mathrm{I}(M; X_{\mathcal{W}}) = 0, \tag{3}$$

where T is the terminal node and M is the R-dimensional message vector generated by the source S.

B. Topological Order

To achieve secure communication over the network $\mathcal{G}$, the source S must "mix" the message symbols in M with the (received) random key symbols in N. This mixture of messages and keys is communicated to the terminal T, which must decode correctly to reconstruct message M. Let $\mathcal{V} = \{v_0, ..., v_{n-1}\}$. Since $\mathcal{G}$ is directed and acylic, we assume, without loss of generality, that the nodes $v_i \in \mathcal{V}$ are indexed according to their topological order in $\mathcal{G}$. This implies that the node v_i receives its incoming information only from nodes $v_0, \cdots, v_{i-1}$. We also assume that the index of K in this topological order is less than that of S which in turn is less than that of the terminal T. More specifically, we assume $K = v_0$, $S = v_m$, and $T = v_{n-1}$ for $v_0, v_m, v_{n-1} \in \mathcal{V}$ and $0 < m < n-1$. There is no loss of generality in these assumptions as otherwise, either transmissions on outgoing edges of S cannot be secure or the communication rate R between S and T is zero. This implies that nodes $\{v_0, \ldots, v_{m-1}\}$ only transmit, on their outgoing edges, functions of the information generated by K while nodes $\{v_m, \ldots, v_{n-1}\}$ may potentially transmit functions of the information generated at both S and K.

C. The Cut Sets

For any pair of nodes $u, v \in \mathcal{V}$, a cut is a set of edges in $\mathcal{E}$ which, when removed, disconnects all paths from u to v. The cut with the minimum capacity that separates u and v is denoted as $\mathrm{mincut}_{\mathcal{G}}(u, v)$. Since each edge in $\mathcal{E}$ is assumed to be of unit capacity, $|\mathrm{mincut}_{\mathcal{G}}(u, v)|$ represents the total capacity of all the edges in $\mathrm{mincut}_{\mathcal{G}}(u, v)$. The cuts as defined above may also separate sets of nodes in the network $\mathcal{G}$. For a subset of nodes $\mathcal{A}$, the set $\mathrm{mincut}_{\mathcal{G}}(\mathcal{A}, v)$ is the minimum capacity cut that separates the set of nodes in $\mathcal{A} \subset \mathcal{V}$ from the node $v \in \mathcal{V}$. For the network $\mathcal{G}$, we use the following notation

$$C_{K-S} = |\mathrm{mincut}_{\mathcal{G}}(K, S)|$$
$$C_{K-T} = |\mathrm{mincut}_{\mathcal{G}}(K, T)|$$
$$C_{S-T} = |\mathrm{mincut}_{\mathcal{G}}(S, T)|$$
$$C_{KS-T} = |\mathrm{mincut}_{\mathcal{G}}(\{K, S\}, T)|$$

III. Results

In this work we prove the following theorem.

Theorem 1. *Given the directed acyclic network $\mathcal{G}$ and integers R and z such that $R > 0$, there exists an (R, z)-feasible network code $\mathcal{N}$ over $\mathcal{G}$ if and only if,*

$$z \leq \min(C_{K-S}, C_{K-T}) \tag{4}$$
$$R \leq C_{S-T} \tag{5}$$
$$R + z \leq C_{KS-T} \tag{6}$$

The proof of Theorem 1 is divided into two parts, the achievability proof, shown in Section IV, and the converse proof shown in Section V.

IV. Proof of Theorem 1: Achievability

Proof. For the network $\mathcal{G} = (\mathcal{E}, \mathcal{V})$ with source node S and key generating node K holding R message symbols M and z key symbols N respectively, we set the values of integers R and z such that they satisfy the bounds (4), (5), and (6). We implement a random linear network code $\mathcal{N}$ over $\mathcal{G}$ and over a sufficiently large field $\mathbb{F}_q$ such that, for any edge $e = (u, v) \in \mathcal{E}$, the local encoding coefficients $\{\bar{c}_{e',e}\}_{e' \in \mathrm{In}(u)}$ associated with edge e, as described in (1), are i.i.d. and uniform over $\mathbb{F}_q$.

The network code $\mathcal{N}$ is said to be decodable at rate R over network $\mathcal{G}$, if it satisfies the condition of (2). We consider the following lemma which we prove in Section VI-A.

Lemma 1. *Given integers R, z that satisfy (4)-(6) of Theorem 1, the random linear network coding scheme $\mathcal{N}$ is decodable at rate R with probability at least $1 - \dfrac{2(|\mathcal{E}| + R + z)^2}{q}$.*

We now consider a wiretapping adversary that can eavesdrop on any subset of edges $\mathcal{W} \subset \mathcal{E}$ such that $|\mathcal{W}| = z$. We denote the information gleaned by the adversary as $X_{\mathcal{W}}$ which may be expressed as

$$X_{\mathcal{W}} = \begin{bmatrix} \mathbf{A}_{\mathcal{W}} & \mathbf{B}_{\mathcal{W}} \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix} \tag{7}$$

Here, $\mathbf{A}_{\mathcal{W}}$ and $\mathbf{B}_{\mathcal{W}}$ are $z \times R$ and $z \times z$ matrices whose rows are global encoding vectors associated with each edge in $\mathcal{W}$, acting on M and N, respectively. We consider the network coded information to be secure if and only if (3) holds for any $\mathcal{W} \subset \mathcal{E}$ of size z, i.e. the adversary gains no information about the source message symbols M even after wiretapping a z-sized subset of edges in the network. In [3], Cai and Yeung show that a linear network coding scheme is secure if and only if the following condition holds.

$$\mathrm{rk}(\begin{bmatrix} \mathbf{A}_{\mathcal{W}} & \mathbf{B}_{\mathcal{W}} \end{bmatrix}) = \mathrm{rk}(\mathbf{B}_{\mathcal{W}}) \tag{8}$$

Here, $\mathrm{rk}(.)$ denotes the *rank* of a matrix.

The following lemma is proven in Section VI-B by analyzing the matrices $\mathbf{A}_{\mathcal{W}}$ and $\mathbf{B}_{\mathcal{W}}$.

Lemma 2. *Given integers R, z that satisfy (4)-(6) of Theorem 1, the random linear network coding scheme $\mathcal{N}$ over $\mathcal{G}$ is z-secure with probability at least $1 - \dfrac{\binom{|\mathcal{E}|}{z} 2z}{q}$ for all wiretap sets $\mathcal{W} \subset \mathcal{E}$ of size z.*

A network code is said to be (R, z)-feasible if it is both R-feasible and z-secure. It now follows that, given integers R and z that satisfy (4), (5), and (6), the suggested network code is (R, z)-feasible with probability at least

$$\Big(1 - \frac{2(|\mathcal{E}| + R + z)^2 + \binom{|\mathcal{E}|}{z} 2z}{q}\Big),$$

which, for sufficiently large q, implies our achievability with high probability. □

V. Proof of Theorem 1: Converse

Proof. We prove the converse for any (not necessarily linear) (R, z)-feasible network code $\mathcal{N}$ over the network $\mathcal{G}$. We start with an (R, z)-feasible coding scheme and show that R and z satisfy the bounds of (4), (5) and (6). Here, we give a partial proof in which we only address bound (4). Proofs of a similar nature apply to the other bounds as well. Details of the converse proof, in the more general context of cyclic networks, appear in Appendix A.

We denote by $\mathbb{C}_{K-S}$ the minimum cut separating K and S, and by C_{K-S} the total capacity of the edges in $\mathbb{C}_{K-S}$. The random variable X_{K-S}, over the support set $\mathcal{X}_{K-S}$, represents the information on all edges of $\mathbb{C}_{K-S}$. We denote by $\mathcal{W}$ any subset of z edges in $\mathcal{E}$ that is wiretapped by an eavesdropping adversary. Then $X_{\mathcal{W}}$ denotes the encoded information on all the edges in $\mathcal{W}$. We denote the set of edges that are incoming to S as $\mathrm{In}(S)$, and the encoded information on all of the edges in $\mathrm{In}(S)$ as $X_{\mathrm{In}(s)}$ with support set $\mathcal{X}_{\mathrm{In}(S)}$. Similarly, for $\mathrm{Out}(S)$.

For the bound $z \leq \min(C_{K-S}, C_{K-T})$ we consider two cases. First, assume by contradiction that $z > C_{K-S}$. Specifically set $z = C_{K-S} + 1$. This implies that the eavesdropping adversary may choose to wiretap all the edges in $\mathbb{C}_{K-S}$ and an edge $e \in \mathrm{Out}(S)$ to obtain the wiretap set $\mathcal{W} = \mathbb{C}_{K-S} \cup \{e\}$ of size z. Then the wiretapped information is $X_{\mathcal{W}} = (X_{K-S}, X_e)$, where X_e is the information on the chosen edge e. Note that $X_e = \bar{f}_e(X_S)$, where, $X_S := (M, X_{\mathrm{In}(S)})$ is the information present at the source S.

For z-security, we require that the mutual information $\mathrm{I}(M; X_{\mathcal{W}}) = 0$. Therefore,

$$\mathrm{I}(M; X_{\mathcal{W}}) = \mathrm{I}(M; X_{K-S}) + \mathrm{I}(M; X_e | X_{K-S}) = 0,$$

implying that, $\mathrm{I}(M; X_{K-S}) = 0$ and $\mathrm{I}(M; X_e | X_{K-S}) = 0$. Thus, we conclude that $\mathrm{H}(X_e | X_{K-S}) = \mathrm{H}(X_e | X_{K-S}, M)$.

Suppose that cut $\mathbb{C}_{K-S}$ partitions $\mathcal{G}$ into disjoint subnetworks $\mathcal{A}$ and $\bar{\mathcal{A}}$, where $\mathcal{A}$ includes the key generating node K. Note that any information communicated through edges in $\bar{\mathcal{A}}$ must be a function of X_{K-S}. In addition, $\mathrm{In}(S) \subset \mathbb{C}_{K-S} \cup \mathcal{E}_{\bar{\mathcal{A}}}$, implying that all information reaching S is a function of X_{K-S}. We conclude, for any edge $e \in \mathrm{Out}(S)$, that

$$X_e = h_e(M, X_{K-S}), \tag{9}$$

where, h_e is some deterministic function. Equation (9) implies that $\mathrm{H}(X_e | X_{K-S}, M) = 0$ which in turn implies $\mathrm{H}(X_e | X_{K-S}) = 0$. This means that to be z-secure the information X_{K-S} must completely determine X_e for all $e \in \mathrm{Out}(S)$. Therefore, the information $X_{\mathrm{Out}(S)} := \{X_e\}_{e \in \mathrm{Out}(S)}$ is also a deterministic function of X_{K-S}. As $\mathrm{I}(M; X_{K-S}) = 0$ shows that X_{K-S} is independent of M, it follows that $X_{\mathrm{Out}(S)}$ is also independent of M and thus $\mathrm{I}(M; X_{\mathrm{Out}(S)}) = 0$. This, in turn, implies that the rate realizable by the network code $\mathcal{N}$ is $R = 0$ which is a contradiction.

A similar proof holds for $z \leq C_{K-T}$, in which we study the set $\mathcal{W} = \mathbb{C}_{K-T} \cup \{e\}$ for any edge $e \in \mathrm{In}(T)$. □

VI. Proof of Lemmas

A. Proof of Lemma 1

We begin by considering the modified network $\mathcal{G}^* = (\mathcal{V}^*, \mathcal{E}^*)$, obtained from $\mathcal{G}$ as shown in Figure 1b. Specifically, $\mathcal{G}^*$ is obtained from $\mathcal{G}$ by adding a new node T^* and $R+z$ parallel edges from S to T^*. As in $\mathcal{G}$, the network $\mathcal{G}^*$ has nodes S and K holding R symbols of M and z symbols of N, respectively. Here, the outgoing edges of S include those in the original network $\mathcal{G}$, denoted as $\mathrm{Out}(S)$, and the additional $R+z$ edges. Both terminals T and T^* want to decode all R symbols of M and z symbols of N. A network code, over $\mathcal{G}^*$, that satisfies the demands of terminals T and T^* is a multi-source multicast network code which is $\mathbf{R}$-feasible, where $\mathbf{R} = (R, z)$.

We use a random linear multi-source multicast network code $\mathcal{N}^*$ over network $\mathcal{G}^*$ and the finite field $\mathbb{F}_q$. In what follows, we set some notation.

1. Let $O_K \triangleq |\mathrm{Out}(K)|$, $I_S \triangleq |\mathrm{In}(S)|$ and $O_S \triangleq |\mathrm{Out}(S)|$.
2. The node K transmits z linear combinations of N through $\mathrm{Out}(K)$. We express the information on these edges as $X_{\mathrm{Out}(K)} = \mathbf{B}_K N$. Here, the rows of $\mathbf{B}_K$, which is an $O_K \times z$ matrix, are the local encoding vectors associated with each edge in $\mathrm{Out}(K)$. The entries of $\mathbf{B}_K$ are i.i.d. and uniform over the field $\mathbb{F}_q$.
3. The message source S receives I_S linear combinations of N through the edges in $\mathrm{In}(S)$. We express the information on these edges as $X_{\mathrm{In}(S)} = \mathbf{V}_{In(S)}\mathbf{B}_K N$. $\mathbf{V}_{\mathrm{In}(S)}$ is an $I_S \times O_K$ matrix, and the rows of $\mathbf{V}_{In(S)}\mathbf{B}_K$ are the global encoding vectors, associated with each edge in $\mathrm{In}(S)$, acting on N.
4. S "mixes" the received I_S symbols of $X_{\mathrm{In}(S)}$ with the R symbols of M and transmits the resulting combinations through $\mathrm{Out}(S)$ and to T^*. We express the information on $\mathrm{Out}(S)$ as

$$\begin{aligned} X_{\mathrm{Out}(S)} &= \begin{bmatrix} \mathbf{A_S} & \mathbf{B}_S \end{bmatrix} \begin{bmatrix} M \\ \mathbf{V}_{In(S)}\mathbf{B}_K N \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A_S} & \mathbf{B}_S\mathbf{V}_{In(S)}\mathbf{B}_K \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix}. \end{aligned}$$

   Here, the rows of the matrix $\begin{bmatrix} \mathbf{A_S} & \mathbf{B}_S \end{bmatrix}$ are the local encoding vectors associated with the edges in $\mathrm{Out}(S)$. $\mathbf{A}_S$ and $\mathbf{B}_S$ are $O_S \times R$ and $O_S \times I_S$ matrices respectively. The entries of $\mathbf{A}_S$ and $\mathbf{B}_S$ are i.i.d. and uniform over $\mathbb{F}_q$.

We now consider the following claims. Claim 2 is proven in Appendix C-A.

Claim 1. *The multi-source multicast random linear network code $\mathcal{N}^*$, as described above, is $\mathbf{R}$-feasible over the network $\mathcal{G}^*$ with probability at least $1 - \dfrac{2(|\mathcal{E}|+R+z)^2}{q}$.*

Proof of Claim 1. Given integers R and z, we start by observing the min-cut capacities in $\mathcal{G}^*$ between the subsets of the node set $\{S, K\}$ and each terminal T and T^* as follows.

$$|\mathrm{mincut}_{\mathcal{G}^*}(K, T)| = C_{K-T} \geq z \tag{10}$$
$$|\mathrm{mincut}_{\mathcal{G}^*}(K, T^*)| = \min(R+z, C_{K-S}) \geq z \tag{11}$$
$$|\mathrm{mincut}_{\mathcal{G}^*}(S, T^*)| = R+z \geq R \tag{12}$$
$$|\mathrm{mincut}_{\mathcal{G}^*}(S, T)| = C_{S-T} \geq R \tag{13}$$
$$|\mathrm{mincut}_{\mathcal{G}^*}(\{K, S\}, T^*)| = R+z \tag{14}$$
$$|\mathrm{mincut}_{\mathcal{G}^*}(\{K, S\}, T)| = C_{KS-T} \geq R+z \tag{15}$$

From (10)-(15), we see that for all source-terminal pairs in $\mathcal{G}^*$, the corresponding Min-Cut Max-Flow bounds are satisfied.

Let L be the total number of encoding coefficients employed over all the edges in $\mathcal{E}^*$. We can bound L by $\sum_{e\in\mathcal{E}^*} |\mathcal{E}^*| \leq |\mathcal{E}^*|^2 = (|\mathcal{E}|+R+z)^2$. Using Theorem 8 of [13] and Theorem 5.4 of [14] (derived from [15]), we have that the network code $\mathcal{N}^*$ is $\mathbf{R}$-feasible over the network $\mathcal{G}^*$ with probability at least

$$\left(1 - \frac{2}{q}\right)^L > 1 - \frac{2L}{q} > 1 - \frac{2(|\mathcal{E}|+R+z)^2}{q}$$

This proves the claim. □

Claim 2. *The $\mathbf{R}$-feasible network code $\mathcal{N}^*$ over $\mathcal{G}^*$, when restricted to $\mathcal{G}$, implies that $\mathcal{N}$ is R-decodable over $\mathcal{G}$.*

From Claim 1 and Claim 2, we have that the network code $\mathcal{N}$ is R-decodable over $\mathcal{G}$ with probability at least

$$1 - \frac{2(|\mathcal{E}|+R+z)^2}{q}$$

This proves the lemma. ■

B. Proof of Lemma 2

We use the notation introduced in the proof of Lemma 1. For any edge $e \in \mathcal{E}$, we express the information on e as,

$$X_e = u_e \begin{bmatrix} X_{\mathrm{Out}(K)} \\ X_{\mathrm{Out}(S)} \end{bmatrix} = u_e \begin{bmatrix} \mathbf{0} & \mathbf{B}_K \\ \mathbf{A}_S & \mathbf{B}_S\mathbf{V}_{\mathrm{In}(S)}\mathbf{B}_K \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix} \tag{16}$$

Here, u_e is an edge-e encoding vector of dimension $O_K + O_S$, acting on $X_{\mathrm{Out}(K)}$ and $X_{\mathrm{Out}(S)}$. We partition $u_e = \begin{bmatrix} u_K & u_S \end{bmatrix}$ such that the O_K-dimensional vector u_K acts on the information from $\mathrm{Out}(K)$ and the O_S-dimensional vector u_S acts on the information from $\mathrm{Out}(S)$. Thus, we rewrite (16) as follows.

$$X_e = \begin{bmatrix} u_K & u_S \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{B}_K \\ \mathbf{A}_S & \mathbf{B}_S\mathbf{V}_{\mathrm{In}(S)}\mathbf{B}_K \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix} \tag{17}$$

We now consider an adversary that wiretaps any subset $\mathcal{W} \subset \mathcal{E}$ of edges such that $|\mathcal{W}| = z$. Then, using (17), we obtain the information observed by the adversary as follows.

$$X_{\mathcal{W}} = \begin{bmatrix} \mathbf{U}_K & \mathbf{U}_S \end{bmatrix} \begin{bmatrix} X_{\mathrm{Out}(K)} \\ X_{\mathrm{Out}(S)} \end{bmatrix} \tag{18}$$

Here, $\begin{bmatrix} \mathbf{U}_K & \mathbf{U}_S \end{bmatrix}$ is a $z \times (O_K + O_S)$ matrix where $\mathbf{U}_K$ is a $z \times O_K$ matrix and $\mathbf{U}_S$ is a $z \times O_S$ matrix. We assume that $\begin{bmatrix} \mathbf{U}_K & \mathbf{U}_S \end{bmatrix}$ has full row-rank of z, as otherwise, the adversary

could simply drop an edge in $\mathcal{W}$ and not lose any information. Using (17), we rewrite (18) as follows.

$$X_{\mathcal{W}} = \begin{bmatrix}\mathbf{U}_S\mathbf{A}_S & \mathbf{U}_K\mathbf{B}_K + \mathbf{U}_S\mathbf{B}_S\mathbf{V}_{\mathrm{In}(S)}\mathbf{B}_K\end{bmatrix}\begin{bmatrix}M\\N\end{bmatrix} \quad (19)$$

From (7) and (19), we have that

$$\mathbf{A}_{\mathcal{W}} = \mathbf{U}_S\mathbf{A}_S \quad \text{and} \quad \mathbf{B}_{\mathcal{W}} = \left[\mathbf{U}_K + \mathbf{U}_S\mathbf{B}_S\mathbf{V}_{\mathrm{In}(S)}\right]\mathbf{B}_K$$

Let,

$$\mathbf{\Phi} \triangleq \left[\mathbf{U}_K + \mathbf{U}_S\mathbf{B}_S\mathbf{V}_{\mathrm{In}(S)}\right]. \quad (20)$$

From our decodability proof, we know that $\mathrm{rk}(\mathbf{V}_{\mathrm{In}(S)}) = z$, as otherwise, T^* could not have decoded the keys N. For the security condition of (8) to hold, we show that $\mathrm{rk}(\mathbf{B}_{\mathcal{W}}) = \mathrm{rk}(\mathbf{\Phi}\mathbf{B}_K) = z$. Therefore, we compute the following.

$$\Pr_{\mathbf{B}_K,\mathbf{B}_S}\{\mathrm{rk}(\mathbf{B}_{\mathcal{W}}) = z\} = \Pr_{\mathbf{B}_S}\{\mathrm{rk}(\mathbf{\Phi}) = z\}\Pr_{\mathbf{B}_K}\{\mathrm{rk}(\mathbf{\Phi}\mathbf{B}_K) = z|\mathrm{rk}(\mathbf{\Phi}) = z\}. \quad (21)$$

We now consider the following claims proven in Appendix C-B and Appendix C-C, respectively.

Claim 3. $\Pr_{\mathbf{B}_S}\{\mathrm{rk}(\mathbf{\Phi}) = z\} > 1 - \frac{z}{q}$

Claim 4. *Given an $n \times m$ matrix $\mathbf{A}$ and an $m \times n$ matrix $\mathbf{B}$ such that $\mathrm{rk}(\mathbf{A}) = n$ and the entries of $\mathbf{B}$ are i.i.d. and uniform over the field $\mathbb{F}_q$, then $\mathrm{rk}(\mathbf{AB}) = n$ with probability at least $1 - \frac{n}{q}$, over $\mathbf{B}$.*

Let us consider the following event.

- $\mathbb{E}_{\mathcal{W}}$: The condition of (8) holds for a given wiretap set $\mathcal{W}$ of size z.

Using Claim 3 and Claim 4 we conclude from (21) that

$$\Pr_{\mathbf{B}_K,\mathbf{B}_S}\{\mathbb{E}_{\mathcal{W}}\} > \left(1 - \frac{z}{q}\right)^2 > 1 - \frac{2z}{q} \quad (22)$$

Denoting the complementary event of $\mathbb{E}_{\mathcal{W}}$ by $\bar{\mathbb{E}}_{\mathcal{W}}$ and using the union bound over event $\bar{\mathbb{E}}_{\mathcal{W}}$ for any $\mathcal{W} \subset \mathcal{E}$ of size z, we have the following.

$$\Pr\{\bigcup_{\mathcal{W}\subset\mathcal{E}}\bar{\mathbb{E}}_{\mathcal{W}}\} \leq \sum_{\mathcal{W}\subset\mathcal{E}}\frac{2z}{q} = \frac{\binom{|\mathcal{E}|}{z}2z}{q}.$$

Namely, the probability over the i.i.d. entries of $\mathbf{B}_S$ and $\mathbf{B}_K$, of the network code being secure against an adversary with a wiretap set $\mathcal{W}$ of size z is at least $1 - \frac{\binom{|\mathcal{E}|}{z}2z}{q}$. This proves the lemma. ■

VII. Conclusion

In this paper, we characterize the capacity-security region for single unicast network codes over a directed acyclic network in which only one node, which is not necessarily the source node, can generate random keys. We present a random linear achievability proof and a matching coverse proof. Our converse can be extended to cyclic networks as well. (Details appear in Appendix A.) Our work establishes an intermediate step between the well understood problem of characterizing the capacity-security region in which only the source node generates random keys and the problem of characterizing the capacity-security region when every node can generate random keys.

Several problems are left open. An extension of our result to the context of multicast network coding is within reach and the subject of future research. It would also be interesting to extend our achievability to single unicast network coding over networks with cycles. Additional possible extensions include the study of single unicast networks in which more than one node can independently generate random keys.

Acknowledgements

Work supported in part by NSF grants CCF-1526771 and CCF-1817241.

References

[1] N. Cai and R. W. Yeung, "Secure network coding," *IEEE International Symposium on Information Theory*, p. 323, 2002.

[2] J. Feldman, T. Malkin, C. Stein, and R. Servedio, "On the capacity of secure network coding," *42nd Annual Allerton Conference on Communication, Control, and Computing*, pp. 63–68, 2004.

[3] N. Cai and R. W. Yeung, "A security condition for multi-source linear network coding," *IEEE International Symposium on Information Theory*, pp. 561–565, 2007.

[4] ——, "On the optimality of a construction of secure network codes," *IEEE International Symposium on Information Theory*, pp. 166–170, 2008.

[5] S. El Rouayheb, E. Soljanin, and A. Sprintson, "Secure network coding for wiretap networks of type II," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1361–1371, 2012.

[6] D. Silva and F. R. Kschischang, "Universal secure network coding via rank-metric codes," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 1124–1135, 2011.

[7] W. Huang, T. Ho, M. Langberg, and J. Kliewer, "Single-unicast secure network coding and network error correction are as hard as multiple-unicast network coding," *IEEE Transactions on Information Theory*, vol. 64, no. 6, pp. 4496–4512, 2018.

[8] T. H. Chan and A. Grant, "Network coding capacity regions via entropy functions," *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5347–5374, 2014.

[9] T. Cui, T. Ho, and J. Kliewer, "On secure network coding with nonuniform or restricted wiretap sets," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 166–176, 2013.

[10] M. Langberg and M. Médard, "On the multiple unicast network coding, conjecture," *47th Annual Allerton Conference on Communication, Control, and Computing*, pp. 222–227, 2009.

[11] S. Jalali and T. Ho, "On capacity region of wiretap networks," *arXiv preprint arXiv:1212.3859*, 2012.

[12] T. Chan and A. Grant, "Capacity bounds for secure network coding," *Australian Communications Theory Workshop*, pp. 95–100, 2008.

[13] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking (TON)*, vol. 11, no. 5, pp. 782–795, 2003.

[14] C. Fragouli and E. Soljanin, *Network Coding Fundamentals*. NOW publishers, 2007.

[15] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.

Appendix A
Proof of Theorem 1: Converse (For cyclic networks)

Proof. We prove the converse for the more general setting of directed networks $\mathcal{G}_{\mathrm{cyc}} = (\mathcal{V}_{\mathrm{cyc}}, \mathcal{E}_{\mathrm{cyc}})$ that may contain cycles.

As before, $\mathcal{G}_{\text{cyc}}$ has the message generating source node S, the random key generating node K and the terminal T, as shown in Figure 2. We show that for such a network and for any network coding scheme, the bounds given in (4), (5) and (6) are upper bounds for the capacity-security region.

As we address networks with cycles, we consider the notion of *time* in our definition of a network code. Namely, we consider an n-time step system. In such a system, we assume that communication starts at time step $i = 1$. The source S holds a message M uniformly distributed in $[q^{nR}]$ and node K holds random keys N uniformly distributed in $[q^{R_K}]$, where R_K is not restricted in any way. For any edge $e \in \mathcal{E}$ such that $e = (u, v)$, where $u, v \in \mathcal{V}_{\text{cyc}}$, we define the information on e at the i-th time step, for all $i \in [n] = \{1, \cdots, n-1\}$, as

$$X_e^{(i)} = \bar{f}_e^{(i)}(\{X_{e'}^{(j)}\}_{e' \in \text{In}(u), j \in [i-1]}) \tag{23}$$

Here, $\bar{f}_e^{(i)}$ is the time-variant local encoding function at edge e at the i-th time step, $\text{In}(u)$ denotes the set of incoming edges in $\mathcal{E}$ to node u and $[i-1] = \{1, \cdots, i-1\}$. In our work, we consider $X_e^{(i)}$ to be a random variable with the support set $\mathcal{X}_e$, for all $i \in [n]$. For a given cut $\mathbb{C}$, we denote by $X_{\mathbb{C}}^{(i)}$ the composite of the variables corresponding to edges $e \in \mathbb{C}$ at time step i, i.e. $X_{\mathbb{C}}^{(i)} = (\{X_e^{(i)}\}_{e \in \mathbb{C}})$. The support set of $X_{\mathbb{C}}^{(i)}$ for all $i \in [n]$ is denoted by $\mathcal{X}_{\mathbb{C}}$. We use the notation $X_e^{[n]}$ to denote the information on edge e for the n-time step system, i.e. $X_e^{[n]} := \{X_e^{(j)}\}_{j \in [n]}$.

For the network model $\mathcal{G}_{\text{cyc}}$, the following definitions are useful for the discussions that follow:

- For any cut $\mathbb{C}_{u-v}$, separating any two nodes $u, v \in \mathcal{V}$, we define two sub-networks $\mathcal{A} = (\mathcal{V}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}})$ and $\bar{\mathcal{A}} = (\mathcal{V}_{\bar{\mathcal{A}}}, \mathcal{E}_{\bar{\mathcal{A}}})$, with $u \in \mathcal{V}_{\mathcal{A}}$ and $v \in \mathcal{V}_{\bar{\mathcal{A}}}$, as shown in Figure 2. Here, $\mathcal{V} = \mathcal{V}_{\mathcal{A}} \cup \mathcal{V}_{\bar{\mathcal{A}}}$ and $\mathcal{E} = \mathcal{E}_{\mathcal{A}} \cup \mathcal{E}_{\bar{\mathcal{A}}} \cup \mathbb{C}_{u-v}$.
- We denote by $\mathbb{C}_{K-S}$, the minimum cut separating K and S and by C_{K-S}, the total capacity of the edges in $\mathbb{C}_{K-S}$. The random variable $X_{K-S}^{(i)}$, over the support set $\mathcal{X}_{K-S}$, represents the information, at the i-th time step on all the edges of $\mathbb{C}_{K-S}$. $X_{K-S}^{[n]} = \{X_{K-S}^{[j]}\}_{j \in [n]}$ represents the information on all the edges of $\mathbb{C}_{K-S}$ for the n-time step system. We use similar notations for the time variant random variables which represent the information on the edges in $\mathbb{C}_{K-T}$, $\mathbb{C}_{S-T}$, and $\mathbb{C}_{KS-T}$.
- We denote by $\mathcal{W}$ any subset of z edges in $\mathcal{E}$ that is wiretapped by an eavesdropping adversary. $X_{\mathcal{W}}^{(i)}$ denotes the encoded information on all the edges in $\mathcal{W}$ at the i-th time step. We assume that the wiretap set $\mathcal{W}$ is time invariant, i.e. it does not change with the time step i. Thus, the information obtained by the adversary for the n-time step system is $X_{\mathcal{W}}^{[n]}$.
- We denote by $\text{In}(S)$, the set of edges that are incoming to S. We denote the encoded information at the i-th time step on all the edges in $\text{In}(S)$ as $X_{\text{In}(s)}^{(i)}$ with support set $\mathcal{X}_{\text{In}(S)}$. Similarly, for $\text{Out}(S)$.
- For any sub-graph $\mathcal{A} = (\mathcal{V}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}}) \subset \mathcal{G}_{\text{cyc}}$, let

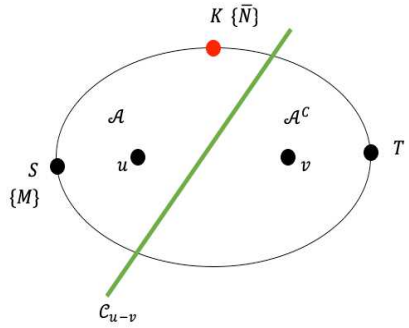


Fig. 2: The partitioning of a network $\mathcal{G}$ due to the cut set $\mathbb{C}_{u-v}$.

$$\mathcal{S}_{\mathcal{A}} := \{S, K\} \cap \mathcal{V}_{\mathcal{A}}. \tag{24}$$

Given the definitions above, we start with an (R, z)-feasible coding scheme and show that R and z satisfy the bounds of (4), (5) and (6). Here, a scheme is (R, z)-feasible with blocklength n if T can decode $M \in [q^{nR}]$ and any wiretapped subset of edge $\mathcal{W}$ hold no information on M. We shall now consider each of the bounds separately in the following subsections.

A. *Bound on* z: $z \leq \min(C_{K-S}, C_{K-T})$

1) $\mathbf{z \leq C_{K-S}}$*:* Suppose, by contradiction, that $z > C_{K-S}$. Particularly, assume $z = C_{K-S} + 1$ This implies that the eavesdropping adversary may choose to wiretap all the edges in $\mathbb{C}_{K-S}$ and an edge $e \in \text{Out}(S)$ to obtain the wiretap set $\mathcal{W} = \mathbb{C}_{K-S} \cup \{e\}$ of size z. Then the wiretapped information is $X_{\mathcal{W}}^{[n]} = (X_{K-S}^{[n]}, X_e^{[n]})$, where $X_e^{[n]}$ is the information on the chosen edge e for the n-time step system. From (23), we see that

$$\begin{aligned} X_e^{(i)} &= \bar{f}_e^{(i)}(\{X_S^{(j)}\}_{j \in [i-1]}) \\ &= \bar{f}_e^{(i)}(\{X_S^{[i-1]}\}). \end{aligned} \tag{25}$$

Where, $X_S^{[i-1]} := (M, X_{\text{In}(S)}^{[i-1]})$ is the information present at the source S for all time steps up to the $(i-1)$-th time-step. For z-security, we require that the mutual information $\text{I}(M; X_{\mathcal{W}}^{(i)}) = 0$. Therefore,

$$\text{I}(M; X_{\mathcal{W}}^{[n]}) = \text{I}(M; X_{K-S}^{[n]}) + \text{I}(M; X_e^{[n]} | X_{K-S}^{[n]}) = 0. \tag{26}$$

Implying that,

$$\text{I}(M; X_{K-S}^{[n]}) = 0, \tag{27}$$

$$\text{I}(M; X_e^{[n]} | X_{K-S}^{[n]}) = 0. \tag{28}$$

From (28), we obtain the following.

$$\text{H}(X_e^{[n]} | X_{K-S}^{[n]}) = \text{H}(X_e^{[n]} | X_{K-S}^{[n]}, M). \tag{29}$$

Suppose the cut $\mathbb{C}_{K-S}$ partitions $\mathcal{G}_{\text{cyc}}$ into disjoint sub-networks $\mathcal{A}$ and $\bar{\mathcal{A}}$. Then, as per the definition in (24), $\mathcal{S}_{\bar{\mathcal{A}}} =$

$\{S\}$. We denote by $X_{\bar{\mathcal{A}}}$, the source message M and/or key N held by the nodes in $\bar{\mathcal{A}}$. We see that $\mathrm{In}(S) \subset \mathbb{C}_{K-S} \cup \mathcal{E}_{\bar{\mathcal{A}}}$, which implies that any edge $e' \in \mathrm{In}(S)$ either belongs to the set $\mathrm{In}(S) \cap \mathbb{C}_{K-S}$ or the set $\mathrm{In}(S) \cap \mathcal{E}_{\bar{\mathcal{A}}}$.

For any edge $e' \in \mathrm{In}(S) \cap \mathbb{C}_{K-S}$, we observe that

$$X_{e'}^{[n]} = h_{e'}(X_{K-S}^{[n]}), \tag{30}$$

For $e' \in \mathrm{In}(S) \cap \mathcal{E}_{\bar{\mathcal{A}}}$, we consider the following lemma which we prove in Appendix B.

Lemma 3. *For any cut $\mathbb{C}$ that partitions graph $\mathcal{G}_{\mathrm{cyc}}$ into disjoint sub-networks $\mathcal{A}$ and $\bar{\mathcal{A}}$, there exists, for any edge $e \in \mathcal{E}_{\bar{\mathcal{A}}}$ and any time step $i \in [n]$, a deterministic mapping $g_e^{(i)}$ such that $g_e^{(i)}(\{X_{\mathbb{C}}^{(j)}\}_{j\in[i-1]}, X_{\bar{\mathcal{A}}}) = X_e^{(i)}$.*

Therefore, using Lemma 3 for edge $e' \in \mathrm{In}(S) \cap \mathcal{E}_{\bar{\mathcal{A}}}$:

$$\begin{aligned} X_{e'}^{[j]} &= g_{e'}^{(j)}(M, X_{K-S}^{[j-1]}) \\ &= \bar{h}_{e'}(M, X_{K-S}^{[j-1]}), \end{aligned} \tag{31}$$

where, $\bar{h}_{e'}$ is a deterministic function.

Then, using (31) we obtain the information on $\mathrm{In}(S)$ as follows.

$$\begin{aligned} X_{\mathrm{In}(S)}^{[j-1]} &= \{X_{e'}^{[j-1]}\}_{e'\in\mathrm{In}(S)\cap\mathbb{C}_{K-S}} \cup \{X_{e'}^{[j-1]}\}_{e'\in\mathrm{In}(S)\cap\mathcal{E}_{\bar{\mathcal{A}}}} \\ &= \{h_{e'}(X_{K-S}^{[j-1]})\}_{e'\in\mathrm{In}(S)\cap\mathbb{C}_{K-S}} \\ &\quad \cup \{\bar{h}_{e'}(M, X_{K-S}^{[j-2]})\}_{e'\in\mathrm{In}(S)\cap\mathcal{E}_{\bar{\mathcal{A}}}} \\ &= \bar{h}_{\mathrm{In}(S)}(M, X_{K-S}^{[j-1]}). \end{aligned} \tag{32}$$

Thus, for the chosen edge $e \in \mathrm{Out}(S)$, using (32) and (25) we have

$$\begin{aligned} X_e^{[n]} &= \{X_e^{(j)}\}_{j\in[n]} \\ &= \{\bar{f}_e^{(i)}(M, X_{\mathrm{In}(S)}^{[j-1]})\}_{j\in[n]} \\ &= \{\bar{f}_e^{(i)}(M, \bar{h}_{\mathrm{In}(S)}(M, X_{K-S}^{[j-1]}))\}_{j\in[n]} \\ &= f(M, X_{K-S}^{[n-1]}). \end{aligned} \tag{33}$$

Thus, (33) shows that $X_e^{[n]}$ is a deterministic function of M and $X_{K-S}^{[n-1]}$. As $\mathrm{H}(X_e^{[n]}|X_{K-S}^{[n]}, M) \leq \mathrm{H}(X_e^{[n]}|X_{K-S}^{[n-1]}, M)$, this implies that

$$\mathrm{H}(X_e^{[n]}|X_{K-S}^{[n]}, M) = 0. \tag{34}$$

Therefore by (29) and (34), we have,

$$\mathrm{H}(X_e^{[n]}|X_{K-S}^{[n]}) = 0. \tag{35}$$

Thus, to be z-secure, (35) shows that for all $e \in \mathrm{Out}(S)$, the random variable $X_e^{[n]}$ must be completely determined by $X_{K-S}^{[n]}$. Therefore, the information $X_{\mathrm{Out}(S)}^{[n]} := \{X_e^{[n]}\}_{e\in\mathrm{Out}(S)}$ is also a deterministic function of $X_{K-S}^{[n]}$. As (27) shows that $X_{K-S}^{[n]}$ is independent of the message symbols M, it follows that $X_{\mathrm{Out}(S)}^{[n]}$ is also independent of M and thus

$$\mathrm{I}(M; X_{\mathrm{Out}(S)}^{[n]}) = 0. \tag{36}$$

Equation (36), in turn implies that the rate realizable by the network code $\mathcal{N}$ is $R = 0$ which is a contradiction.

2) $\mathbf{z \leq C_{K-T}}$: Suppose, by contradiction, that $z > C_{K-T}$, specifically assuming that $z = C_{K-T}+1$. This implies that the eavesdropping adversary may choose to wiretap all the edges in $\mathbb{C}_{K-T}$ and any edge $e \in \mathrm{In}(T)$ to obtain the wiretapped set $\mathcal{W} = \mathbb{C}_{K-T} \cup \{e\}$ of size z. Then the wiretapped information is $X_{\mathcal{W}}^{[n]} = (X_{K-T}^{[n]}, X_e^{[n]})$, where $X_e^{[n]}$ is the information on the chosen edge e. For z-security, we require that the mutual information $\mathrm{I}(M; X_{\mathcal{W}}^{[n]}) = 0$. Therefore,

$$\mathrm{I}(M; X_{\mathcal{W}}^{[n]}) = \mathrm{I}(M; X_{K-T}^{[n]}) + \mathrm{I}(M; X_e^{[n]}|X_{K-T}^{[n]}) = 0. \tag{37}$$

Further implying that,

$$\mathrm{I}(M; X_{K-T}^{[n]}) = 0, \tag{38}$$

$$\mathrm{I}(M; X_e^{[n]}|X_{K-T}^{[n]}) = 0. \tag{39}$$

From (39),

$$\mathrm{H}(X_e^{[n]}|X_{K-T}^{[n]}) = \mathrm{H}(X_e^{[n]}|X_{K-T}^{[n]}, M). \tag{40}$$

We now consider the cut $\mathbb{C}_{K-T}$ and the corresponding partitions $\mathcal{A}$ and $\bar{\mathcal{A}}$. Note that corresponding to the cut $\mathbb{C}_{K-T}$, the set of information and key generating source nodes in $\mathcal{G}_{\mathrm{cyc}}$ which are also present in $\bar{\mathcal{A}}$ is $\mathcal{S}_{\bar{\mathcal{A}}}$ where $\mathcal{S}_{\bar{\mathcal{A}}} \subseteq \{S, K\}$.

Note that $\mathrm{In}(T) \subset \mathbb{C}_{K-T} \cup \mathcal{E}_{\bar{\mathcal{A}}}$. Due to the cut $\mathbb{C}_{K-T}$, it follows that either $S \in \mathcal{V}_{\bar{\mathcal{A}}}$ or $S \in \mathcal{V}_{\mathcal{A}}$. For any edge $e' \in \mathrm{In}(T) \cap \mathbb{C}_{K-T}$, we have the following.

$$X_{e'}^{[n]} = h_{e'}(X_{K-T}^{[n-1]}, M), \tag{41}$$

where, $h_{e'}$ is a deterministic function.

For any edge $e' \in \mathrm{In}(T) \cup \mathcal{E}_{\bar{\mathcal{A}}}$, by Lemma 3, we have the following.

$$\begin{aligned} X_{e'}^{[n]} &= \{g_{e'}^{(j)}(X_{K-T}^{[j-1]}, M)\}_{j\in[n]} \\ &= h_{e'}(X_{K-T}^{[n-1]}, M). \end{aligned} \tag{42}$$

Equation (42) shows that for any edge $e \in \mathrm{In}(T)$, the random variable $X_e^{[n]}$ is completely determined by $X_{K-T}^{[n-1]}$ and M. As $\mathrm{H}(X_e^{[n]}|X_{K-T}^{[n]}, M) \leq \mathrm{H}(X_e^{[n]}|X_{K-T}^{[n-1]}, M)$, we have.

$$\mathrm{H}(X_e^{[n]}|X_{K-T}^{[n]}, M) = 0, \tag{43}$$

which implies by (40) that $\mathrm{H}(X_e^{[n]}|X_{K-T}^{[n]}) = 0$. This holds for all $e \in \mathrm{In}(T)$ and thus $\mathrm{H}(X_{\mathrm{In}(T)}|X_{K-T}^{[n]}) = 0$. As (38) shows that $X_{K-T}^{[n]}$ is independent of M, therefore we conclude that,

$$\mathrm{I}(M; X_{\mathrm{In}(T)}^{[n]}) = 0. \tag{44}$$

This in turn implies that the rate realizable by the network code $\mathcal{N}$ is $R = 0$ which is a contradiction. Thus, for $R > 0$, an (R, z)-feasible network code exists only if $z \leq \min(C_{K-S}, C_{K-T})$, i.e., bound (4) holds.

B. Upper Bound of R

The bound (5) is a direct consequence of Theorem 2.1 of [14] and therefore the proof is not included here.

C. Upper Bound of $R+z$: $R+z \leq C_{KS-T}$

To show that an (R,z)-feasible network code exists only if bound (6) holds, we start by considering the following cases:

- **Case 1:** $z \geq C_{KS-T}$
- **Case 2:** $z < C_{KS-T}$

For **Case 1**, we see that the eavesdropping adversary has the option of wiretapping all the edges in $\mathbb{C}_{KS-T}$. Therefore, we set $\mathbb{C}_{KS-T} \subseteq \mathcal{W}$ thereby forcing $\mathrm{I}(M; X^{[n]}_{KS-T}) = 0$. This, implies that $X^{[n]}_{KS-T}$ is independent of the message symbols M. We also observe that $\mathrm{In}(T) \subset \mathbb{C}_{KS-T} \cup \mathcal{E}_{\bar{\mathcal{A}}}$. From our previous discussions, we note that the random variable $X^{[n]}_{\mathrm{In}(T)}$ is a deterministic function of $X^{[n]}_{KS-T}$ and therefore is also independent M. Thus, the terminal T receives no information regarding the message symbols M and therefore the rate realizable by the network code in this case is $R=0$ which is a contradiction.

For **Case 2**, let $\mathcal{W} \subset \mathbb{C}_{KS-T}$. Then, $\mathbb{C}_{KS-T} = \mathcal{W} \cup \mathcal{W}^C$ where, $\mathcal{W}^C = \mathbb{C}_{KS-T} \setminus \mathcal{W}$. We denote the information on the edges of the set $\mathcal{W}^C$ as $X^{[n]}_{\mathcal{W}^C}$ and thus we have that $X^{[n]}_{KS-T} = (X^{[n]}_{\mathcal{W}}, X^{[n]}_{\mathcal{W}^C})$ where $\mathrm{H}(X^{[n]}_{\mathcal{W}^C}) \leq n(C_{KS-T}-z)$. Here, our measure $\mathrm{H}(.)$ of entropy equals 1 for a uniform random variable in $\mathbb{F}_q$. Thus, we have the following.

$$
\begin{aligned}
nR &= \mathrm{I}(M; X^{[n]}_{KS-T}) && (45)\\
&= \mathrm{I}(M; X^{[n]}_{\mathcal{W}}, X^{[n]}_{\mathcal{W}^C}) \\
&= \mathrm{I}(M; X^{[n]}_{\mathcal{W}}) + \mathrm{I}(M; X^{[n]}_{\mathcal{W}^C} | X^{[n]}_{\mathcal{W}}) \\
&= \mathrm{I}(M; X^{[n]}_{\mathcal{W}^C} | X^{[n]}_{\mathcal{W}}) && (46)\\
&= \mathrm{H}(X^{[n]}_{\mathcal{W}^C} | X^{[n]}_{\mathcal{W}}) - \mathrm{H}(X^{[n]}_{\mathcal{W}^C} | X^{[n]}_{\mathcal{W}}, M) \\
&\leq \mathrm{H}(X^{[n]}_{\mathcal{W}^C} | X^{[n]}_{\mathcal{W}}) \\
&\leq \mathrm{H}(X^{[n]}_{\mathcal{W}^C}) \\
&\leq n(C_{KS-T}-z). && (47)
\end{aligned}
$$

Here, (45) is due to our assumption of correctly decoding M and the min-cut max-flow theorem as the cut $\mathbb{C}_{KS-T}$ is an $(S-T)$-cut. (46) is due to the security condition. Thus, one may realize an (R,z)-feasible network code over the network $\mathcal{G}_{\mathrm{cyc}}$ only if the bound (6) holds for integers $R>0$ and z.

Combining our analysis for bounds (4), (5) and (6) proves the theorem.

□

Appendix B
Proof of Lemma 3

We prove this lemma using an induction hypothesis on the time step parameter i. At time $i=0$, we assume that the network edges do not carry any information. Thus, at time $i=1$, the information on all network edges in $\bar{\mathcal{A}}$ are solely a function of the random variable $X_{\bar{\mathcal{A}}}$.

We assume by induction that the hypothesis holds for $1 < i \leq I-1$, i.e. for $i = I-1$, we have the following.

$$X^{(I-1)}_e = g^{(I-1)}_e(\{X^{(j)}_{\mathbb{C}}\}_{j\in[I-2]}, X_{\bar{\mathcal{A}}}). \tag{48}$$

We now consider time step $i=I$. For an edge $e=(u,v)$, $X^{(N)}_e$ is a function of the incoming edges to u and $X_{\bar{\mathcal{A}}}$ (the latter only if $u \in \mathcal{S}_{\bar{\mathcal{A}}}$). Namely,

$$X^{(I)}_e = f^{(I)}_e(\{X^{(I-1)}_{e'}\}_{e'\in\mathrm{In}(u)}, X_{\bar{\mathcal{A}}}).$$

For $u \in \mathcal{V}_{\bar{\mathcal{A}}}$, the incoming edges of u are either included in the cut $\mathbb{C}$ or are in $\mathcal{E}_{\bar{\mathcal{A}}}$. Thus,

$$X^{(I)}_e = f^{(I)}_e(X^{[I-1]}_{\mathbb{C}}, \{X^{(I-1)}_{e'}\}_{e'\in\mathrm{In}(u)}, X_{\bar{\mathcal{A}}}).$$

By induction, as $X^{(I-1)}_{e'}$ is a function of $X^{[I-2]}_{\mathbb{C}}$ and $X_{\bar{\mathcal{A}}}$ for $e' \in \mathcal{E}_{\bar{\mathcal{A}}}$. We conclude that there exists a function $g^{(I)}_e$ such that,

$$X^{(I)}_e = g^{(I)}_e(X^{[I-1]}_{\mathbb{C}}, X_{\bar{\mathcal{A}}}).$$

This proves the lemma. ■

Appendix C
Proof of Claims

A. Proof of Claim 2

To prove that the network code $\mathcal{N}$ is R-decodable over network $\mathcal{G}$, given that $\mathcal{N}^*$ is $\mathbf{R}$-feasible over $\mathcal{G}^*$, we consider the following steps.

1. We disconnect terminal T^* from S by removing the edges connecting S to T^*.
2. We keep the random assignment of the local coding coefficients for each edge in $\mathcal{E}$ unchanged.
3. As in $\mathcal{N}^*$, the source S does not decode the keys N but "mixes" the incoming combinations of the keys in N with the message M that it holds, and transmits the resulting combinations through $\mathrm{Out}(S)$.
4. The information that terminal T wants to decode also remains unchanged.

By initiating the steps above, we obtain the network code $\mathcal{N}$ from $\mathcal{N}^*$. It also follows that the network code $\mathcal{N}$ allows T to decode all R symbols of M as the $\mathbf{R}$-feasible $\mathcal{N}^*$ allows T to decode all R symbols of M and z symbols of N, thereby satisfying condition (2). This proves the claim. □

B. Proof of Claim 3

Since $\mathrm{rk}(\begin{bmatrix}\mathbf{U}_K & \mathbf{U}_S\end{bmatrix}) = z$, we assume, without loss of generality, that the first σ_K columns of $\mathbf{U}_K$ and σ_S columns of $\mathbf{U}_S$ are jointly linearly independent with $\sigma_K + \sigma_S = z$. Then, we have that $\mathbf{U}_S = \begin{bmatrix}\bar{\mathbf{U}}_S & \bar{\mathbf{U}}_S\mathbf{\Gamma}\end{bmatrix}$, where $\bar{\mathbf{U}}_S$ is $z \times \sigma_S$ matrix of full column-rank σ_S, and $\mathbf{\Gamma}$ is a $\sigma_S \times (O_S - \sigma_S)$ matrix. Let $\bar{\mathbf{U}}_K$ be the sub-matrix of $\mathbf{U}_K$ containing the σ_K linearly independent columns of $\mathbf{U}_K$, and $\mathbf{\Delta}$ be a $\sigma_K \times O_K$ matrix such that $\mathbf{U}_K = \bar{\mathbf{U}}_K\mathbf{\Delta}$. Then, as per our assumption, the first σ_K columns $\mathbf{\Delta}$ form a $\sigma_K \times \sigma_K$ identity matrix. Therefore, we have that $\mathrm{rk}(\mathbf{\Delta}) = \sigma_K$

We now consider the matrix $\mathbf{V}_{\mathrm{In}(S)}$. Since, $\mathrm{rk}(\mathbf{V}_{\mathrm{In}(S)}) = z$, we may express $\mathbf{V}_{\mathrm{In}(S)}$ as follows.

$$\mathbf{V}_{\mathrm{In}(S)} = \begin{bmatrix} \mathbf{V} & \mathbf{V}_1 \\ \mathbf{V}_2 & \mathbf{V}_3 \end{bmatrix} \tag{49}$$

Here, $\mathbf{V}$ is a $z \times z$ invertible sub-matrix of $\mathbf{V}_{\mathrm{In}(S)}$. The columns of the $z \times (O_K - z)$ sub-matrix $\mathbf{V}_1$ and the rows of the $(I_S - z) \times z$ sub-matrix $\mathbf{V}_2$ are spanned by the columns and rows of $\mathbf{V}$ respectively, while the rows of the sub-matrix $\mathbf{V}_3$ are spanned by the rows of $\mathbf{V}_1$. Thus we may rewrite (49) as follows

$$\mathbf{V}_{\mathrm{In}(S)} = \begin{bmatrix} \mathbf{V} & \mathbf{V}_1 \\ \mathbf{AV} & \mathbf{V}_3 \end{bmatrix} \tag{50}$$

Here, $\mathbf{A}$ is an $(I_S - z) \times z$ matrix. We now partition the matrix $\mathbf{B}_S = \begin{bmatrix} \mathbf{B}_{S,1} & \mathbf{B}_{S,2} \end{bmatrix}$ such that $\mathbf{B}_{S,1}$ and $\mathbf{B}_{S,2}$ are $O_S \times z$ and $O_S \times (I_S - z)$ matrices respectively. Let $\mathbf{B} \triangleq \mathbf{B}_S \mathbf{V}_{\mathrm{In}(S)}$, then by using (50) we obtain the following.

$$\begin{aligned} \mathbf{B} &= \begin{bmatrix} \mathbf{B}_{S,1}\mathbf{V} + \mathbf{B}_{S,2}\mathbf{AV} & \mathbf{B}_{S,1}\mathbf{V}_1 + \mathbf{B}_{S,2}\mathbf{V}_3 \end{bmatrix} \\ &\triangleq \begin{bmatrix} \bar{\mathbf{B}} & \tilde{\mathbf{B}} \end{bmatrix} \end{aligned} \tag{51}$$

Here, the matrices $\bar{\mathbf{B}}$ and $\tilde{\mathbf{B}}$ are $O_S \times z$ and $O_S \times (I_S - z)$, respectively. Furthermore, we partition $\bar{\mathbf{B}} = \begin{bmatrix} \bar{\mathbf{B}}^1 \\ \bar{\mathbf{B}}^2 \end{bmatrix}$, where $\bar{\mathbf{B}}^1$ and $\bar{\mathbf{B}}^2$ are $\sigma_S \times z$ and $(O_S - \sigma_S) \times z$ sub-matrices respectively. Likewise, we partition $\tilde{\mathbf{B}} = \begin{bmatrix} \tilde{\mathbf{B}}^1 \\ \tilde{\mathbf{B}}^2 \end{bmatrix}$. Then, using (50) and (51), we may rewrite (20) as follows.

$$\mathbf{\Phi} = \begin{bmatrix} \bar{\mathbf{U}}_K & \bar{\mathbf{U}}_S \end{bmatrix} \begin{bmatrix} \mathbf{\Delta} \\ \\ \bar{\mathbf{B}}^1 + \mathbf{\Gamma}\bar{\mathbf{B}}^2 & \tilde{\mathbf{B}}^1 + \mathbf{\Gamma}\tilde{\mathbf{B}}^2 \end{bmatrix} \tag{52}$$

We now consider partition $\mathbf{\Delta} = \begin{bmatrix} \mathbf{\Delta}_z & \hat{\mathbf{\Delta}} \end{bmatrix}$, where $\mathbf{\Delta}_z$ consists of the first z columns of $\mathbf{\Delta}$. Then, we have that $\mathbf{\Delta}_z = \begin{bmatrix} \mathbf{I}_{\sigma_K} & \bar{\mathbf{\Delta}}_z \end{bmatrix}$ where $\mathbf{I}_K$ is the $\sigma_K \times \sigma_K$ identity matrix and $\bar{\mathbf{\Delta}}_z$ is a $\sigma_K \times (z - \sigma_K)$ matrix whose columns are spanned by $\mathbf{I}_{\sigma_K}$. Thus, we see that the σ_K rows of $\mathbf{\Delta}_z$ are linearly independent. We may rewrite (52) as follows.

$$\begin{aligned} \mathbf{\Phi} &= \begin{bmatrix} \bar{\mathbf{U}}_K & \bar{\mathbf{U}}_S \end{bmatrix} \begin{bmatrix} \mathbf{\Delta}_z & \hat{\mathbf{\Delta}} \\ \\ \bar{\mathbf{B}}^1 + \mathbf{\Gamma}\bar{\mathbf{B}}^2 & \tilde{\mathbf{B}}^1 + \mathbf{\Gamma}\tilde{\mathbf{B}}^2 \end{bmatrix} \\ &\triangleq \begin{bmatrix} \bar{\mathbf{U}}_K & \bar{\mathbf{U}}_S \end{bmatrix} \mathbf{Q} \end{aligned} \tag{53}$$

Since $\begin{bmatrix} \bar{\mathbf{U}}_K & \bar{\mathbf{U}}_S \end{bmatrix}$ is invertible, $\mathrm{rk}(\mathbf{Q}) = z$ implies $\mathrm{rk}(\mathbf{\Phi}) = z$. To prove that $\mathrm{rk}(\mathbf{Q}) = z$, we show that the σ_S rows of $\bar{\mathbf{B}}^1 + \mathbf{\Gamma}\bar{\mathbf{B}}^2$, having dimension z, are linearly independent and not spanned by the σ_K rows of $\mathbf{\Delta}_z$.

Given that the entries of the matrix $\mathbf{B}_S$ are i.i.d and uniform in $\mathbb{F}_q$, for any $\psi \in \mathbb{F}_q^z$ and for $i \in [O_S]$, we compute the probability $\Pr_{\mathbf{B}_S}\{(\bar{b})^i = \psi\}$, where $(\bar{b})^i$ denotes the i-th row of $\bar{\mathbf{B}}$ of dimension z. From (51), we see that

$$(\bar{b}^i) = (b_{S,1})^i \mathbf{V} + (\mathbf{B}_{S,2}\mathbf{A})^i \mathbf{V}. \tag{54}$$

Here, $(b_{S,1})^i$ and $(\mathbf{B}_{S,2}\mathbf{A})^i$ denotes the i-th row of $\mathbf{B}_{S,1}$ and $\mathbf{B}_{S,2}\mathbf{A}$ respectively. As $\mathbf{V}$ is invertible, we have the following.

$$\begin{aligned} \Pr_{\mathbf{B}_S}\{(\bar{b})^i = \psi\} &= \Pr_{(b_{S,1})^i, \mathbf{B}_{S,2}}\{(b_{S,1})^i + (\mathbf{B}_{S,2}\mathbf{A})^i = \psi\mathbf{V}^{-1}\} \\ &= \sum_{\bar{\psi}} \Pr_{\mathbf{B}_{S,2}}\{(\mathbf{B}_{S,2}\mathbf{A})^i = \bar{\psi}\}. \\ &\qquad \Pr_{(b_{S,1})^i}\{(b_{S,1})^i = (\psi\mathbf{V}^{-1} - \bar{\psi})\} \\ &= \frac{1}{q^z} \sum_{\bar{\psi}} \Pr_{\mathbf{B}_{S,2}}\{(\mathbf{B}_{S,2}\mathbf{A})^i = \bar{\psi}\} \qquad (55) \\ &= \frac{1}{q^z} \qquad (56) \end{aligned}$$

Here, (55) is due to that fact that the entries of $\mathbf{B}_{S,1}$, which is a sub-matrix of $\mathbf{B}_S$, are i.i.d. and uniform in $\mathbb{F}_q$. From (56), we see that the rows of $\bar{\mathbf{B}}$ are uniform in $\mathbb{F}_q^z$. For a fixed matrix $\mathbf{B}_{S,2}$ in (54) and due to the invertibility of matrix $\mathbf{V}$, there exists a 1-1 map between $(b_{S,1})^i \in \mathbb{F}_q^z$ and $(\bar{b}^i) \in \mathbb{F}_q^z$, for all $i \in [O_S]$. Now, as the vectors $(b_{S,1})^i$ are chosen independently for each $i \in [O_S]$, it follows that the corresponding vectors $(\bar{b}^i)$ must also be independent for all $i \in [O_S]$. This implies that the rows of $\bar{\mathbf{B}}^1$, which is a sub-matrix of $\bar{\mathbf{B}}$ containing its first σ_S rows, are also i.i.d and uniform in $\mathbb{F}_q^z$.

Let $\mathbf{C} \triangleq \bar{\mathbf{B}}^1 + \mathbf{\Gamma}\bar{\mathbf{B}}^2$ and c^i denote the i-th row of $\mathbf{C}$ for $i \in [\sigma_S]$. For any $\rho \in \mathbb{F}_q^z$, we compute the probability $\Pr_{\bar{\mathbf{B}}}\{c^i = \rho\}$. Denoting the i-th rows of $\bar{\mathbf{B}}^1$ and $\mathbf{\Gamma}\bar{\mathbf{B}}^2$ as $(\bar{b}^1)^i$ and $(\mathbf{\Gamma}\bar{\mathbf{B}}^2)^i$, respectively, we have that,

$$c^i = (\bar{b}^1)^i + (\mathbf{\Gamma}\bar{\mathbf{B}}^2)^i.$$

Note that the rows of $\bar{\mathbf{B}}^1$ form the first σ_S rows of $\bar{\mathbf{B}}$ and therefore are i.i.d. and uniform in $\mathbb{F}_q^z$. Thus, by applying the same argument as in (56), we obtain $\Pr_{\bar{\mathbf{B}}}\{c^i = \rho\} = \frac{1}{q^z}$. The vectors $\{c^i\}_{i \in [O_S]}$ are mutually independent due to the fact that the vectors $\{(\bar{b}^1)^i\}_{i \in [O_S]}$ are mutually independent. Thus, as $\sigma_K + \sigma_S = z$, we have the following.

$$\begin{aligned} \Pr_{\mathbf{B}_S}\{\mathrm{rk}(\Phi) = z\} &= \frac{\prod_{l=0}^{\sigma_S - 1} q^z - q^{\sigma_K + l}}{q^{z\sigma_S}} \\ &= \prod_{l=0}^{\sigma_S - 1} \left(1 - \frac{1}{q^{z - \sigma_k - l}}\right) \\ &> \left(1 - \frac{1}{q}\right)^{\sigma_S} \\ &> 1 - \frac{\sigma_S}{q} \\ &> 1 - \frac{z}{q} \end{aligned} \tag{57}$$

This proves our claim.

□

C. Proof of Claim 4

Let $\mathbf{AB} = \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_n \end{bmatrix}$, where $\lambda_j \in \mathbb{F}_q^n$ for $j \in [n]$, $\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix}$, where $a_i \in \mathbb{F}_q^n$ for $i \in [m]$ and

$\mathbf{B} = \{b_{i,j}\}_{i\in[m],j\in[n]}$, where $b_{i,j}$'s are i.i.d. and uniform over $\mathbb{F}_q$.

For any $j \in [n]$, we have

$$\lambda_j = \sum_{i\in[m]} a_i b_{i,j} \tag{58}$$

For any $\omega \in \mathbb{F}_q^n$, we first compute $\Pr_{b_j}\{\lambda_j = \omega\}$, where b_j is the m-dimensional j-th column of $\mathbf{B}$. Since $\mathrm{rk}(\mathbf{A}) = n$, we assume, without loss of generality, that the last n columns of $\mathbf{A}$ are linearly independent. We also partition b_j such that $b_j = \begin{bmatrix}\bar{b}_j \\ \tilde{b}_j\end{bmatrix}$ where $\tilde{b}_j$ consists of the last n entries of b_j. Then, we may rewrite (58) as

$$\lambda_j = \sum_{i\in[m-n]} a_i b_{i,j} + \sum_{i=m-n+1}^{m} a_i b_{i,j} \tag{59}$$

Then,

$$\begin{aligned}\Pr_{b_j}\{\lambda_j = \omega\} &= \Pr_{b_j}\{\sum_{i\in[m]} a_i b_{i,j} = \omega\} \\ &= \sum_{\hat{\omega}} \Pr_{\bar{b}_j}\{\sum_{i\in[m-n]} a_i b_{i,j} = \hat{\omega}\}. \\ &\qquad \Pr_{\tilde{b}_j}\{\sum_{i=m-n+1}^{m} a_i b_{i,j} = \omega - \hat{\omega}\}\end{aligned} \tag{60}$$

Since the n-dimensional columns $\{a_i\}_{i=m-n+1}^{m}$ are linearly independent, the n-system of equations $\sum_{i=m-n+1}^{m} a_i b_{i,j} = \omega - \hat{\omega}$ must have a unique solution for each $\tilde{b}_j \in \mathbb{F}_q^n$, and as the entries of $\tilde{b}_j$ are i.i.d. uniform in $\mathbb{F}_q$, we have that $\Pr_{\tilde{b}_j}\{\sum_{i=m-n+1}^{m} a_i b_{i,j} = \omega - \hat{\omega}\} = 1/q^n$. Thus, we may rewrite (60) as follows.

$$\begin{aligned}\Pr_{b_j}\{\lambda_j = \omega\} &= \frac{1}{q^n}\sum_{\hat{\omega}} \Pr_{\bar{b}_j}\{\sum_{i\in[m-n]} a_i b_{i,j} = \hat{\omega}\} \\ &= \frac{1}{q^n}\end{aligned} \tag{61}$$

Equation (61) implies that the columns $\{\lambda_j\}_{j\in[n]}$ are uniform in $\mathbb{F}_q^n$. The columns λ_j are also mutually independent due to the fact that the columns b_j are independent for all $j \in [n]$. Thus, we have

$$\begin{aligned}\Pr_{\mathbf{B}}\{\mathrm{rk}(\mathbf{AB}) = n\} &= \frac{\prod_{l=0}^{n-1}\left(q^n - q^l\right)}{q^{n^2}} \\ &= \prod_{l=0}^{n-1}\left(1 - \frac{1}{q^{n-l}}\right) \\ &> \left(1 - \frac{1}{q}\right)^n \\ &> 1 - \frac{n}{q}\end{aligned} \tag{62}$$

This proves the claim.

□