

# Secure Network Coding

Ning Cai\*and Raymond W. Yeung<sup>†</sup>

February 29, 2008

In the paradigm of network coding, the nodes in a network are allowed to encode the information received from the input links. With network coding, the full capacity of the network can be utilized. In this paper, we propose a model which incorporates network coding and information security. In this model, a collection of subsets of the channels in the network is given, and a wiretapper is allowed to access any one (but not more than one) of these subsets without being able to obtain any information about the message transmitted. Our model includes secret sharing in classical cryptography as a special case. We present a construction of secure linear network codes provided a certain graph-theoretic condition is satisfied. We also prove the necessity of this condition for the special case that the wiretapper may choose to access any subset of channels of a fixed size. Finally, we extend our results to the scenario when the wiretapper is allowed to obtain a controlled amount of information about the message.

---

\*N. Cai is with The State Key Lab. of ISN, Xidian University, Xi'an, Shaanxi, 710071, China. He was with Department of Information Engineering, The Chinese University of Hong Kong, N.T., Hong Kong when this work was done. Email: caining@mail.xidian.edu.cn

<sup>†</sup>R. W. Yeung is with Department of Information Engineering, The Chinese University of Hong Kong, N.T., Hong Kong. Email: whyeung@ie.cuhk.edu.hk

# 1 Introduction

We start our discussion with the following classical problem in cryptography. Suppose a sender wants to send the output of a random message  $M$  with alphabet  $\mathcal{M} = \{0, 1, \dots, p-1\}$  to a receiver. The sender can send information via a “public” channel, whose output can be accessed by the receiver as well as a wiretapper who tries to obtain some information about  $M$ , or the sender can send information via a “secure” channel, whose output can be accessed only by the receiver. The usual way to protect  $M$  from the wiretapper is that the sender generates a “secret key”  $K$  independent of the source  $M$  according to the uniform distribution over  $\mathcal{M}$ . Let  $m$  be the outcome of  $M$ , and let  $k$  be the outcome of  $K$ . Then the sender sends the key  $k$  to the receiver via the secure channel, and sends  $m + k \pmod{p}$  via the public channel. Upon receiving both  $k$  and  $m + k$ , the receiver as the legal user can recover  $m$  because  $m = (m + k) - k$ . On the other hand, the wiretapper cannot obtain any information about  $m$  by knowing  $m + k$  alone because what he/she knows is a total randomization of the message  $m$ .

The main idea in the above scheme is that the sender has to randomize the message in order to protect it from the wiretapper, where in this case the alphabets of the random key and of the information source have the same size (the two alphabets are the same). Shannon showed in [12] that this protocol is optimal in the sense of minimizing the size of the random key. This result, known as the *perfect secrecy theorem*, has been generalized to the *imperfect secrecy theorem* by Yeung [13] (p. 116).

In the above scheme, if another wiretapper observes  $k$  but cannot observe  $m + k$ , he/she again cannot obtain any information about  $M$ . Thus the only thing we have to do for security is to make sure that an illegal user cannot obtain the outputs of both the public and the secure channels. This observation tells us that there is actually no logical difference between the public channel and the secure channel. This is the simplest example of a communication system over a wiretap network (CSWN) which will be studied in this paper.

In the next section, we present our model of a CSWN and define a secure network code, which in our terminology is called an admissible code. In Section 3, we construct a class

of linear codes based on the work of Li *et al.* [9] on linear network coding. In Section 4, we present a sufficient condition for the construction in Section 3 to be admissible. The proof of the sufficiency of this condition is deferred to Section 7. In Section 5, we prove the optimality of our construction in Section 3 for the special case that the wiretapper may choose to access any subset of channels of a fixed size. Two simple examples are given in Section 6 to illustrate the results. In Section 8, we extend our results to the scenario when the wiretapper is allowed to obtain a controlled amount of information about the message. The paper is concluded in Section 9.

## 2 Communication System on a Wiretap Network

In this section, we first present our model of a communication system on a wiretap network (CSWN), which subsumes the secret sharing model proposed independently by Blakley [2] and Shamir [11] (see also Ozarow and Wyner's wire-tap channel II [10], a special case of secret sharing). Then we will define a code for a CSWN.

A CSWN consists of the following components:

1) *Directed multigraph  $\mathcal{G}$* : The pair  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  is called a directed multigraph<sup>1</sup>, where  $\mathcal{V}$  and  $\mathcal{E}$  are the node set and the edge set of  $\mathcal{G}$ , respectively. In our model, we assume that  $\mathcal{G}$  is acyclic, i.e., it does not contain a directed cycle.

2) *Source node  $s$* : The node set  $\mathcal{V}$  contains a node  $s$ , called the source node, where a random message  $M$  taking values in an alphabet  $\mathcal{M}$  is generated.

3) *Set of user nodes  $\mathcal{U}$* : A user node is a node in  $\mathcal{V}$  which is fully accessed by a legal user who is required to receive the random message  $M$  with zero error. There is generally more than one user node in a network. The set of user nodes is denoted by  $\mathcal{U}$ .

4) *Collection of sets of wiretap edges  $\mathcal{A}$* :  $\mathcal{A}$  is a collection of subsets of the edge set  $\mathcal{E}$ . Each member of  $\mathcal{A}$  may be fully accessed by a wiretapper, but no wiretapper may access more than one member of  $\mathcal{A}$ .

---

<sup>1</sup>In a multigraph, there can be more than one edge from one node to another node.

We refer to the quadruple  $(\mathcal{G}, s, \mathcal{U}, \mathcal{A})$  as a CSWN. We also refer to the multigraph  $\mathcal{G}$  as a network and the edges in  $\mathcal{E}$  as channels. The random message  $M$  is generated at the source node  $s$  according to an arbitrary distribution on an alphabet  $\mathcal{M}$ , called the *message set*. On each channel in  $\mathcal{E}$ , an index taken from an alphabet  $F$ , called the *transmission alphabet*, can be transmitted. We are interested in the maximum value of  $|\mathcal{M}|$  for which the message  $M$  can be multicast from the source node  $s$  to the set of user nodes  $\mathcal{U}$  while being protected from wiretappers who can access any set of channels in  $\mathcal{A}$ .

The current work is a generalization of the work by Ahlswede *et al.* [1] and Li *et al.* [9] on network coding. In the paradigm of network coding, the nodes in a communication network are allowed to encode the information received from the input links before it is transmitted on the output links. The advantage of network coding is that it can utilize the full capacity of a network for multicasting information.

In the model we study in the current paper, in the absence of a wiretapper, i.e.,  $\mathcal{A} = \emptyset$ , a CSWN is reduced to the model studied in [1] and [9]. It was proved in [1] that information can be multicast from the source node  $s$  to all the user nodes in  $\mathcal{U}$  at rate  $\tau$  if and only if the value of a maximum flow from  $s$  to each user node is at least  $\tau$  in the graph  $\mathcal{G}$ . In general, information can be multicast from the source node to the user nodes at a higher rate with network coding than without network coding when there are at least two user nodes (see the example in [1]). Subsequently, it was proved in [9] by an explicit construction that this can be achieved by linear network codes. For a comprehensive treatment of network coding, we refer the reader to [14].

As we have discussed earlier, it is necessary to randomize the message in order to protect it from the wiretappers. This can be explained as follows. If there is no randomness in the network, the index transmitted on any channel is a function of the message  $M$  and hence is not independent of  $M$  unless the index takes a constant value. If this is the case, the channel becomes degenerate as it cannot transmit any useful information through.

Let  $K$  be an independent random variable, called the key, that takes values in an alphabet  $\mathcal{K}$  according to the uniform distribution. To facilitate our discussion, we denote the

sets of input and output channels of a given node  $a \in \mathcal{V}$  by  $\text{In}(a)$  and  $\text{Out}(a)$ , respectively. A code for a CSWN consists of a set of local encoding mappings  $\{\phi_e : e \in \mathcal{E}\}$  such that for all  $e$ ,  $\phi_e$  is a function from  $\mathcal{M} \times \mathcal{K}$  to  $F$  if  $e \in \text{Out}(s)$ , and is a function from  $F^{|\text{In}(t)|}$  to  $F$  if  $e \in \text{Out}(t)$  for  $t \neq s$ . For  $e \in \mathcal{E}$ , let  $Y_e$  be the random symbol in  $F$  transmitted on channel  $e$ , i.e., the value of  $\phi_e$ . For a subset  $B$  of  $\mathcal{E}$ , denote  $(Y_e : e \in B)$  by  $Y_B$ .

To complete the description of a code, we have to specify the order in which the channels send the indices, called the *encoding order*. Since the graph  $\mathcal{G}$  is acyclic, it defines a partial order on the node set  $\mathcal{V}$ . Then the nodes in  $\mathcal{V}$  can be indexed in a way such that for two nodes  $t$  and  $t'$ , if there is a channel from node  $t$  to node  $t'$ , then  $t < t'$ . According to this indexing, node  $t$  sends indices in its output channels before node  $t'$  if and only if  $t < t'$ . The order in which the channels within the set of output channels of a node send the indices is immaterial. The important point here is that whenever a channel sends an index, all the indices necessary for encoding have already been received. A code defined as such induces a function  $\Phi_u$  from  $\mathcal{M} \times \mathcal{K}$  to  $F^{|\text{In}(u)|}$  for all user nodes  $u \in \mathcal{U}$ , where the value of  $\Phi_u$  denotes the indices received by the user node  $u$  in its input channels.

A code  $\{\phi_e : e \in \mathcal{E}\}$  is admissible for a CSWN  $(\mathcal{G}, s, \mathcal{U}, \mathcal{A})$  if the following conditions are satisfied:

- 1) For all user nodes  $u \in \mathcal{U}$  and all  $\mathbf{m}, \mathbf{m}' \in \mathcal{M}$  with  $\mathbf{m} \neq \mathbf{m}'$ ,

$$\Phi_u(\mathbf{m}, \mathbf{k}) \neq \Phi_u(\mathbf{m}', \mathbf{k}')$$

for all  $\mathbf{k}, \mathbf{k}' \in \mathcal{K}$ . This guarantees that any two messages are distinguishable at every user node, and we refer to this as the *decodable condition*.

- 2) For all  $A \in \mathcal{A}$

$$H(M|Y_A) = H(M).$$

Here  $H(\cdot|\cdot)$  and  $H(\cdot)$  denote conditional entropy and entropy, respectively. In other words,  $M$  and  $Y_A$  are independent. This is referred to as the *secure condition*.

The model of a CSWN and the results in Sections 3 and 4 here have been presented as a conference paper in [3].

### 3 A Class of Linear Codes for Communication Systems on a Wiretap Network

In this section, we propose a class of linear codes for a CSWN. In defining a linear network code, we let the transmission alphabet  $F$  be a finite field  $GF(q)$ , where  $q$  is a sufficiently large power of a prime. In other words, a symbol in  $GF(q)$  can be transmitted on each channel in the network.

In the rest of the paper, we adopt the terminologies for linear network codes in [15]. In defining an  $n$ -dimensional linear network code on  $\mathcal{G}$ , we let  $\text{In}(s)$  consist of  $n$  imaginary channels terminating at the source node  $s$ .

**Definition 1 (Global description of a linear Network code)** *An  $n$ -dimensional linear network code on  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  consists of a column  $n$ -vector  $\mathbf{f}_e$  for every channel  $e \in \mathcal{E} \cup \text{In}(s)$  such that:*

1. for  $e \in \text{Out}(t)$ ,  $\mathbf{f}_e$  is a linear combination of  $\mathbf{f}_d$ ,  $d \in \text{In}(t)$ .
2.  $\mathbf{f}_e$ ,  $e \in \text{In}(s)$ , form the standard basis of the vector space  $F^n$ .

The vector  $\mathbf{f}_e$  is called the global encoding kernel for channel  $e$ .

We use  $\langle \cdot \rangle$  to denote the linear span of a set of vectors. For  $t \in \mathcal{V}$ , let

$$V_t = \langle \{\mathbf{f}_e : e \in \text{In}(t)\} \rangle.$$

For  $T \subset \mathcal{V}$ , let

$$V_T = \langle \{\cup_{t \in T} V_t\} \rangle,$$

and for  $B \in \mathcal{E}$ , let

$$V_B = \langle \{\mathbf{f}_e : e \in B\} \rangle.$$

For a node  $t \in \mathcal{V}$  where  $t \neq s$ , let  $\text{maxflow}(t)$  denote the value of a maximum flow from the source node  $s$  to node  $t$ .

The following existence theorem of a linear network code with the prescribed property is due to Jaggi *et al.* [8], where they proposed a polynomial-time algorithm, known as the Jaggi-Sanders algorithm, for constructing such a code.

**Theorem 1** *If  $\max\text{flow}(u) \geq n$  for all  $u \in \mathcal{U}$ , then there exists an  $n$ -dimensional linear network code on  $\mathcal{G}$  over  $GF(q)$  for  $q > |\mathcal{U}|$  such that  $\dim(V_u) = n$  for all  $u \in \mathcal{U}$ .*

We now define a class of linear codes for a CSWN by the following construction.

**Construction 1**

1) Choose suitable positive integers  $n$  and  $r$ , where  $r < n$ . The message  $M$  is randomly chosen from  $GF^{(n-r)}(q)$  (not necessarily uniformly distributed), while the independent random key  $K$  is distributed uniformly on  $GF^r(q)$ . Let the outcome  $\mathbf{m}$  of  $M$  be a row vector in  $GF^{(n-r)}(q)$  and the outcome  $\mathbf{k}$  of  $K$  be a row vector in  $GF^r(q)$ . Let  $X = (M, K)$ .

2) Choose a suitable  $n$ -dimensional linear network code on  $\mathcal{G}$ .

3) Encode the vector  $X$  by transmitting in each channel  $e$  the value  $X \mathbf{f}_e$ .

We will show later how  $n$ ,  $r$  and the linear network code can be chosen to make the code admissible, i.e., decodable and secure.

## 4 Construction of Admissible Codes

In this section, we first present Theorem 2 which states that if a certain condition is satisfied, then it is possible to obtain an admissible code by Construction 1. The proof of this theorem is deferred to Section 7. The sufficient condition in Theorem 2 depends on a linear network code satisfying certain properties whose existence is hard to verify. Nevertheless, a more explicit sufficient condition will be obtained.

**Theorem 2** *There exists an admissible code on  $\mathcal{G}$  over  $GF(q)$  for  $q > |\mathcal{A}|$  by Construction 1 if there exists an  $n$ -dimensional linear network code over  $GF(q)$  such that for all user nodes  $u \in \mathcal{U}$ ,*

$$\dim(V_u) = n, \tag{1}$$

and for all wiretap sets of channels  $A \in \mathcal{A}$ ,

$$\dim(V_A) \leq r. \quad (2)$$

In the directed graph  $\mathcal{G}$ , a path is a sequence of channels  $e_1, e_2, \dots, e_l$  such that for  $1 \leq i \leq l-1$ , there exists  $t_i \in \mathcal{V}$  such that  $e_i \in \text{In}(t_i)$  and  $e_{i+1} \in \text{Out}(t_i)$ . Two paths are disjoint if they do not share a common channel (but they may share a common node). The following theorem is similar to Theorem 2, condition therein depends only on the graph  $\mathcal{G}$  and the collection of wiretap channels  $\mathcal{A}$ .

**Theorem 3** *Let  $\mathcal{G}^* = (\mathcal{V}, \mathcal{E}^*)$ , where  $\mathcal{E}^* \subset \mathcal{E}$ , be a subgraph of  $\mathcal{G}$  satisfying the following:*

*i) For any  $u \in \mathcal{U}$ , there are  $n$  disjoint paths in  $\mathcal{G}^*$  from the source node  $s$  to the user node  $u$ .*

*ii) For any  $A \in \mathcal{A}$ , there are at most  $r$  disjoint paths in  $\mathcal{E}^*$  from the source node  $s$  to the channels in  $A \in \mathcal{E}^*$ .*

*If such a subgraph  $\mathcal{G}^*$  exists, then there exists an admissible code on  $\mathcal{G}$  over  $GF(q)$  by Construction 1 for  $q > \max\{|\mathcal{U}|, |\mathcal{A}|\}$ .*

This theorem is a simple consequence of Theorems 1 and 2 and the following lemma.

**Lemma 1** *For any  $A \in \mathcal{A}$ , let  $\text{maxflow}(A)$  denote the maximum number of disjoint paths from the source node  $s$  to the channels in  $A$ . For any linear network code defined on  $\mathcal{G}$ ,  $\dim(V_A) \leq \text{maxflow}(A)$ .*

**Proof** We label the channels in  $\mathcal{E}$  according to the encoding order, with the smallest label being 1. By induction on the largest label  $l$  among the channels in  $A$ , we will show that there exist  $\dim(V_A)$  disjoint paths from the source node  $s$  to the channels in  $A$ . The claim is trivially true for  $l \leq \text{Out}(s)$ . Assume the claim is true for all  $l \leq L-1$  for some  $L > \text{Out}(s)$ , and we will prove that it is true for  $l = L$ . Consider any  $A \in \mathcal{A}$  such that  $l = L$ , and let  $e_1, e_2, \dots, e_d$  be channels in  $A$  where  $d = \dim(V_A)$ , such that the labels of  $e_1, e_2, \dots, e_d$  are in ascending order and  $\mathbf{f}_{e_1}, \mathbf{f}_{e_2}, \dots, \mathbf{f}_{e_d}$  form a maximal set of linearly independent vectors in  $V_A$ .

We can assume without loss of generality that the label of  $e_d$  is equal to  $L$ , because otherwise the claim follows immediately by applying the induction hypothesis to  $A' = \{e_1, e_2, \dots, e_d\}$ . Let  $e_d \in \text{Out}(t)$ . Since  $\mathbf{f}_{e_d}$  is linearly independent of  $\mathbf{f}_{e_1}, \mathbf{f}_{e_2}, \dots, \mathbf{f}_{e_{d-1}}$ ,  $V_t \not\subseteq \langle \mathbf{f}_{e_1}, \mathbf{f}_{e_2}, \dots, \mathbf{f}_{e_{d-1}} \rangle$ . Then there exists a channel  $e'_d \in V_t$  such that  $\mathbf{f}_{e_1}, \mathbf{f}_{e_2}, \dots, \mathbf{f}_{e_{d-1}}, \mathbf{f}_{e'_d}$  are linearly independent and all their labels are smaller than  $L$ . Let  $A'' = \{e_1, e_2, \dots, e_{d-1}, e'_d\}$ . By the induction hypothesis, there exist  $d = \dim(V_A)$  disjoint paths from the source node  $s$  to the channels in  $A''$ . Then by appending  $e_d$  to the path from  $s$  to  $e'_d$ , we obtain a set of  $\dim(V_A)$  disjoint paths from the source node  $s$  to the channels in  $A$ . This proves the claim, which implies that  $\text{maxflow}(A) \geq \dim(V_A)$ . The lemma is proved.  $\square$

**Proof of Theorem 3** Assume the existence of the subgraph  $\mathcal{G}^*$  as prescribed and let  $q > \max\{|\mathcal{U}|, |\mathcal{A}|\}$ . We will confine our discussion to  $\mathcal{G}^*$ . The condition i) in the theorem implies that  $\text{maxflow}(u) \geq n$  for all  $u \in \mathcal{U}$ , and the condition ii) in the theorem implies that  $\text{maxflow}(A) \leq r$  for all  $A \in \mathcal{A}$ . Since  $q > |\mathcal{U}|$ , by Theorem 1, there exists an  $n$ -dimensional linear network code on  $\mathcal{G}^*$  such that  $\dim(V_u) = n$  for all  $u \in \mathcal{U}$ . Now for this network code, for any  $A \in \mathcal{A}$ , by Lemma 1,  $\dim(V_A) \leq \text{maxflow}(A) \leq r$ . Since  $q > |\mathcal{A}|$ , by invoking Theorem 2, we see the existence of an admissible code on  $\mathcal{G}^*$  by Construction 1. The theorem is proved.  $\square$

## 5 Optimality of Construction 1

Consider the case that the collection  $\mathcal{A}$  of wiretap sets consists of all the  $r$ -subsets of  $\mathcal{E}$  ( $A$  is an  $r$ -subset of  $\mathcal{E}$  means that  $A \subset \mathcal{E}$  and  $|A| = r$ ). In Theorem 3, let  $n = \min_{u \in \mathcal{U}} \text{maxflow}(u)$ , and let  $\mathcal{E}^* = \mathcal{E}$ , i.e.,  $\mathcal{G}^* = \mathcal{G}$ . Then the conditions i) and ii) are satisfied, and we obtain the following corollary of the theorem.

**Corollary 1** *Let  $\mathcal{A}$  consist of all the  $r$ -subsets of  $\mathcal{E}$ , and let  $n = \min_{u \in \mathcal{U}} \text{maxflow}(u)$ . Then there exists an admissible code on  $\mathcal{G}$  over  $GF(q)$  by Construction 1 for  $q > \max\{|\mathcal{U}|, \binom{|\mathcal{E}|}{r}\}$ .*

Now the code obtained by Construction 1 can transmit a message  $M$  consisting of  $n - r$  symbols in  $GF(q)$  to all the user nodes  $u \in \mathcal{U}$  securely. To achieve this, a key consisting of

$r$  symbols in  $GF(q)$  is used. In this section, we will establish the optimality of the code so constructed by proving two fundamental performance bounds and then showing that the tightness of these bounds are achievable.

Consider any admissible code on a CSWN. Let  $u \in \mathcal{U}$  be such that  $\text{maxflow}(u) = n$  and  $(W, W^c)$  be a minimum cut between the source node  $s$  and node  $u$ . Denote the set of channels on  $(W, W^c)$  by  $E_W$ . Then  $|E_W| = n$ . Since the message  $M$  can be decoded at node  $u$  and the symbols received at node  $u$  are functions of  $Y_{E_W}$ , we have

$$H(M|Y_{E_W}) = 0. \quad (3)$$

On the other hand, for any subset  $\mathcal{I}$  of  $E_W$  with cardinality  $r$ , since the code is secure, we have

$$H(M|Y_{\mathcal{I}}) = H(M). \quad (4)$$

It follows that

$$\begin{aligned} H(M) &= H(M|Y_{\mathcal{I}}) - H(M|Y_{E_W}) \\ &= I(M; Y_{E_W \setminus \mathcal{I}}|Y_{\mathcal{I}}) \\ &\leq H(Y_{E_W \setminus \mathcal{I}}|Y_{\mathcal{I}}) \\ &\leq H(Y_{E_W \setminus \mathcal{I}}) \\ &\leq (n - r) \log q. \end{aligned}$$

The tightness of this upper bound on  $H(M)$  is achievable by the code obtained by Construction 1 when  $M$  is distributed uniformly on  $GF^{(n-r)}(q)$ . In other words, the code multicasts the maximum possible amount of information to the user nodes securely.

In the rest of the section, we will prove that the code uses the minimum amount of randomness to achieve the required security when the message  $M$  is distributed uniformly. In establishing this result, we need a set of inequalities stated in the next lemma due to Han [7] (see also [4], Theorem 17.6.3).

**Lemma 2** *For a subset  $\alpha$  of  $\mathcal{N} = \{1, 2, \dots, n\}$ , let  $\bar{\alpha} = \mathcal{N} \setminus \alpha$  and denote  $(X_i, i \in \alpha)$  by*

$X_\alpha$ . For  $1 \leq r \leq n$ , let

$$h_r = \frac{1}{\binom{n-1}{r-1}} \sum_{\alpha:|\alpha|=r} H(X_\alpha|X_{\bar{\alpha}}). \quad (5)$$

Then

$$h_1 \leq h_2 \leq \dots \leq h_n.$$

Let  $u \in \mathcal{U}$  be any user node and consider any cut  $(W, W^c)$  between the source node  $s$  and node  $u$ . Let  $|E_W| = n' \geq n$ . For any  $\mathcal{I} \subset E_W$  such that  $|\mathcal{I}| = r$ , consider

$$\begin{aligned} H(M) &= H(M|Y_{E_W}) + I(Y_{E_W}; M) \\ &= I(Y_{\mathcal{I}}; M) + I(Y_{E_W \setminus \mathcal{I}}; M|Y_{\mathcal{I}}) \\ &= I(Y_{E_W \setminus \mathcal{I}}; M|Y_{\mathcal{I}}), \end{aligned} \quad (6)$$

where the second and the third equalities follow from (3) and (4), respectively. Summing over all  $\mathcal{I}$ , we have

$$\begin{aligned} \binom{n'}{r} H(M) &= \sum_{\mathcal{I}} I(Y_{E_W \setminus \mathcal{I}}; M|Y_{\mathcal{I}}) \\ &\leq \binom{n'-1}{n'-r-1} \left[ \frac{1}{\binom{n'-1}{n'-r-1}} \sum_{\mathcal{I}} H(Y_{E_W \setminus \mathcal{I}}|Y_{\mathcal{I}}) \right] \\ &\leq \binom{n'-1}{n'-r-1} H(Y_{E_W}), \end{aligned}$$

where the last inequality follows from Lemma 2. Hence,

$$H(Y_{E_W}) \geq \frac{n'}{n'-r} H(M). \quad (7)$$

Finally,

$$H(M) + H(K) \geq H(M, K) \quad (8)$$

$$= H(M, K, Y_{E_W}) \quad (9)$$

$$\geq H(Y_{E_W})$$

$$\geq \frac{n'}{n'-r} H(M),$$

where (9) follows from

$$H(Y_{E_W}|M, K) = 0, \quad (10)$$

or  $Y_{E_W}$  is a function of  $M$  and  $K$ . This implies

$$H(K) \geq \frac{r}{n' - r} H(M). \quad (11)$$

This lower bound on  $H(K)$  applies to every cut between the source node  $s$  and any user node  $u$ , in particular to a cut with size equal to  $n$ . Therefore, we conclude that

$$H(K) \geq \frac{r}{n - r} H(M). \quad (12)$$

The tightness of this lower bound on  $H(K)$  is achieved by the code obtained by Construction 1 when both  $M$  and  $K$  are uniformly distributed, i.e.,  $H(M) = (n - r) \log q$  and  $H(K) = r \log q$ . Under this condition, the code uses the minimum amount of randomness to achieve the required security. In Appendix A, we prove that

$$H(K) \geq H(Y_{\mathcal{I}}).$$

This lower bound on  $H(K)$  gives further insight into the problem.

We note that the inequality in (8) holds regardless of whether the message  $M$  and the key  $K$  are independent. In fact, toward establishing (12), this assumption in Construction 1 has not been invoked. Hence, (12) is valid even when  $M$  and  $K$  are not independent.

In obtaining (9) in the above, we have used the fact  $Y_{E_W}$  is a function of  $M$  and  $K$ . Close examination of the steps in our proof reveals that  $K$  can be more generally interpreted as the randomness introduced into the network at the “upstream” of the set of channels  $E_W$ . When  $n' > n$ , (11) is a looser lower bound on  $H(K)$  than (12). This means that it is not necessary for all the randomness  $K$  to be generated at the source node  $s$  as in Construction 1. As long as all the randomness  $K$  is generated at the “upstream” of any cut of size  $n$  between the source node  $s$  and any user node  $u$ , it is already good enough. This observation would be useful if the source node  $s$  does not have enough resource to generate all the required randomness.

Hence, we have proved that when the message is uniformly distributed, the code obtained by Construction 1 is optimal in terms of both the amount information that can be multicast in the network securely and the amount of randomness used for achieving the required security.

## 6 Two Examples

In this section, we give two examples to illustrate our results.

**Example 1 (Secret sharing)** Consider the CSWN shown in Figure 1 with

$$\mathcal{U} = \{u_1, u_2, u_3\}$$

and

$$\mathcal{A} = \{\{(s, a_1)\}, \{(s, a_2)\}, \{(s, a_3)\}\}.$$

This CSWN represents the (1,2)-threshold secret sharing scheme.

**Example 2** Consider the CSWN shown in the Figure 2 with

$$\mathcal{U} = \{u_1, u_2\}$$

and

$$\mathcal{A} = \{\{(a_1, u_1)\}, \{(a_2, u_2)\}, \{(a_0, b)\}\}.$$

Then  $\{(s, a_1), (a_1, u_1)\}$  and  $\{(s, a_2), (a_2, a_0), (a_0, b), (b, u_1)\}$  are two disjoint paths from the source node  $s$  to user node  $u_1$ , and  $\{(s, a_2), (a_2, u_2)\}$  and  $\{(s, a_1), (a_1, a_0), (a_0, b), (b, u_2)\}$  are two disjoint paths from the source node  $s$  to user node  $u_2$ . Here  $\mathcal{E}^* = \mathcal{E}$ , and each  $A \in \mathcal{A}$  is a subset of  $\mathcal{E}^*$ . Therefore,  $A \subset \mathcal{E}^*$  for all  $A \in \mathcal{A}$ .

It is clear that there is only one path from the source node  $s$  to any set of wiretap channels  $A$  (in fact in this case all the sets of wiretap channels contain only one channel). Thus by Theorem 3, for a sufficiently large power  $q$  of a prime, an admissible code exists for  $n = 2$ ,  $k = 1$  and  $m = 1$ .

For this example,  $q = 3$  is sufficiently large. In the following, we will work in the finite field  $GF(3)$ , and all operations are assumed to be in  $GF(3)$ . Let  $M$  be a ternary source taking values in  $GF(3)$ . At the source node  $s$ , an independent random key  $K$  is also generated according to the uniform distribution on  $GF(3)$ . Denote the values taken by  $M$  and  $K$  by  $m_1$  and  $k_1$ , respectively. The source node  $s$  sends  $m_1 - k_1$  and  $m_1 + k_1$  to nodes  $a_1$  and  $a_2$ , respectively. Then node  $a_1$  sends  $m_1 - k_1$  to nodes  $a_0$  and  $u_1$ , and node  $a_2$  sends  $m_1 + k_1$  to nodes  $a_0$  and  $u_2$ . Node  $a_0$ , upon receiving  $m_1 - k_1$  and  $m_1 + k_1$ , solves for  $k_1$  and sends it to node  $b$ . Finally, node  $b$  sends  $k_1$  to the two user nodes  $u_1$  and  $u_2$ . It is easy to check that such a code satisfies the decodable condition and the secure condition, and is therefore admissible.

## 7 Proof of Theorem 2

Assume the existence of the  $n$ -dimensional linear network code as prescribed in the theorem. Denote the code by  $\mathcal{C}$  and let  $\mathbf{f}_e, e \in \mathcal{E}$  be the global encoding kernels. For all  $A \in \mathcal{A}$ , let  $\dim(V_A) = r_A$ , and let  $\{\mathbf{a}_1(A), \mathbf{a}_2(A), \dots, \mathbf{a}_{r_A}(A)\}$  be a maximally independent set of vectors in  $\{\mathbf{f}_e, e \in A\}$ . Note that  $r_A \leq r$  by (2).

**Lemma 3** *If  $q > |\mathcal{A}|$ , there exist column  $n$ -vectors  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-r}$  such that for all  $A \in \mathcal{A}$ ,*

$$\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-r}, \mathbf{a}_1(A), \mathbf{a}_2(A), \dots, \mathbf{a}_{r_A}(A)$$

*are linearly independent.*

**Proof** It suffices to show that for  $1 \leq i \leq n - r$ , if  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{i-1}$  have been chosen such that for all  $A \in \mathcal{A}$ ,

$$\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{i-1}, \mathbf{a}_1(A), \mathbf{a}_2(A), \dots, \mathbf{a}_{r_A}(A) \tag{13}$$

are linearly independent, then it is possible to choose  $\mathbf{b}_i$  such that for all  $A \in \mathcal{A}$ ,

$$\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{a}_1(A), \mathbf{a}_2(A), \dots, \mathbf{a}_{r_A}(A) \tag{14}$$

are linearly independent. Specifically,  $\mathbf{b}_i$  is chosen such that it is linearly independent of the set of vectors in (13) for all  $A \in \mathcal{A}$ , i.e., we require that

$$\mathbf{b}_i \in GF^n(q) \setminus \bigcup_{A \in \mathcal{A}} \langle \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{i-1}, \mathbf{a}_1(A), \mathbf{a}_2(A), \dots, \mathbf{a}_{r_A}(A) \rangle.$$

Thus we need to show that the set above is nonempty. Since the vectors in (13) are linearly independent,

$$\begin{aligned} & \left| \bigcup_{A \in \mathcal{A}} \langle \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{i-1}, \mathbf{a}_1(A), \mathbf{a}_2(A), \dots, \mathbf{a}_{r_A}(A) \rangle \right| \\ & \leq \sum_{A \in \mathcal{A}} |\langle \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{i-1}, \mathbf{a}_1(A), \mathbf{a}_2(A), \dots, \mathbf{a}_{r_A}(A) \rangle| \\ & = \sum_{A \in \mathcal{A}} q^{r_A+i-1} \\ & \leq \sum_{A \in \mathcal{A}} q^{r+i-1} \\ & = |\mathcal{A}|q^{r+i-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} & \left| GF^n(q) \setminus \bigcup_{A \in \mathcal{A}} \langle \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{i-1}, \mathbf{a}_1(A), \mathbf{a}_2(A), \dots, \mathbf{a}_{r_A}(A) \rangle \right| \\ & \geq q^n - |\mathcal{A}|q^{r+i-1} \\ & = q^{r-i+1}(q^{n-r-i+1} - |\mathcal{A}|) \\ & \geq q^{r-i+1}(q - |\mathcal{A}|) \\ & > 0 \end{aligned}$$

since  $i \leq n - r$  and  $q > |\mathcal{A}|$ . Hence,  $\mathbf{b}_i$  can be chosen for all  $1 \leq i \leq n - r$ .  $\square$

Subsequent to [3], Feldman *et al.* [5] pointed out that the condition for  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-r}$  in Lemma 3 is equivalent to a Hamming distance property of a certain type of codes. They also derived a tradeoff between the size of the message set  $\mathcal{M}$  and the size of the transmission alphabet  $F$ . Specifically, they showed that it is sufficient to take  $q > |\mathcal{A}|^{\frac{1}{r\epsilon+1}}$  if we want to send a message consisting of  $\lfloor n - r(1 + \epsilon) \rfloor$  instead of  $(n - r)$   $q$ -ary symbols through the

network. We note that this tradeoff can readily be obtained by replacing  $i \leq n - r$  and  $q > |\mathcal{A}|$  by  $i \leq n - r(1 + \epsilon)$  and  $q^{r\epsilon+1} > |\mathcal{A}|$ , respectively in the above proof.

Let  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-r}$  be chosen according to the above lemma. Extend  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-r}$  to a basis  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-r}, \mathbf{b}_{n-r+1}, \dots, \mathbf{b}_n$  for  $GF^n(q)$ , and define the  $n \times n$  matrix

$$Q = [ \mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n ]. \quad (15)$$

Evidently,  $Q$  is nonsingular.

Let  $\mathcal{T} : GF^n(q) \rightarrow GF^n(q)$  be the linear transformation represented by the matrix  $Q^{-1}$ . Now transform the linear network code  $\mathcal{C}$  into another linear network code  $\mathcal{C}'$  with global encoding kernels  $\mathbf{f}'_e, e \in \mathcal{E}$  by  $\mathcal{T}$ , i.e.,  $\mathbf{f}'_e = Q^{-1}\mathbf{f}_e$  for all  $e \in \mathcal{E}$ . For  $u \in \mathcal{U}$  and  $A \in \mathcal{A}$ , the vector spaces  $V_u$  and  $V_A$  for the linear network code  $\mathcal{C}$  become  $V'_u = \mathcal{T}V_u$  and  $V'_A = \mathcal{T}V_A$  for the linear network code  $\mathcal{C}'$ , respectively. Since  $Q$  is invertible, by (1) and (2),

$$\dim(V'_u) = \dim(V_u) = n \quad (16)$$

and

$$\dim(V'_A) = \dim(V_A) \leq r. \quad (17)$$

With the global encoding kernels  $\mathbf{f}'_e, e \in \mathcal{E}$ , we obtain a linear network code  $\mathcal{C}_s$  for a CSWN by Construction 1. It follows from (16) by a straightforward argument that  $\mathcal{C}_s$  is decodable because at each user node  $u \in \mathcal{U}$ , both  $\mathbf{m}$  and  $\mathbf{k}$  can be decoded with zero error.

To complete the proof, we only have to check that  $\mathcal{C}_s$  is secure. For  $1 \leq j \leq r_A$ , let  $\mathbf{a}'_j(A) = Q^{-1}\mathbf{a}_j(A)$ . Let  $F_A$  and  $F'_A$  be  $n \times r_A$  matrices whose  $j$ th columns are  $\mathbf{a}_j(A)$  and  $\mathbf{a}'_j(A)$ , respectively. Then

$$F'_A = Q^{-1}F_A. \quad (18)$$

Let  $Y_A$  be the vector of symbols transmitted on the channels in the wiretap set  $A$ . Let  $\mathbf{y}_A$  be the value of  $Y_A$  when  $X = \mathbf{x}$ , i.e.,

$$\mathbf{y}_A = \mathbf{x}F'_A. \quad (19)$$

In other words, upon observing  $\mathbf{y}_A$ , the knowledge of the wiretapper is that  $\mathbf{x}$  is a solution of the above equation. For a row  $r_A$ -vector  $\mathbf{y} \in GF^{r_A}(q)$ , let

$$C(\mathbf{y}) = \{\mathbf{x} : \mathbf{x} \in GF^n(q), \mathbf{y} = \mathbf{x}F'_A\}. \quad (20)$$

Then the solution set of (19) is given by  $C(\mathbf{y}_A)$ , which is seen to be a coset of the null space  $C(\mathbf{0})$  under the linear transformation represented by  $F'_A$ . Therefore,  $GF^n(q)$  is partitioned into  $\{C(\mathbf{y}) : \mathbf{y} \in GF^{r_A}(q)\}$ .

For  $\mathbf{m} \in GF^{(n-r)}(q)$ , let

$$D(\mathbf{m}) = \{(\mathbf{m}, \mathbf{k}) : \mathbf{k} \in GF^r(q)\}.$$

We now show that for all  $\mathbf{m} \in GF^{(n-r)}(q)$  and  $\mathbf{y} \in GF^{r_A}(q)$ ,

$$|D(\mathbf{m}) \cap C(\mathbf{y})| = q^{r-r_A}, \quad (21)$$

which does not depend on  $\mathbf{m}$ . Let

$$G_A = [ \mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_{n-r} \ F'_A ],$$

where  $\mathbf{e}_j$  is the column  $n$ -vector whose  $j$ th component is 1 and all other components are 0. It follows from the definition of  $D(\mathbf{m})$  and  $C(\mathbf{y})$  that if a vector  $\mathbf{x}$  is in their intersection, then

$$\mathbf{x}G_A = [ \mathbf{m} \ \mathbf{y} ]. \quad (22)$$

Therefore, by (15) and (18), we have

$$QG_A = [ \mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_{n-r} \ \mathbf{a}_1(A) \ \mathbf{a}_2(A) \ \cdots \ \mathbf{a}_{r_A}(A) ].$$

By construction, the columns of  $QG_A$  are linearly independent, so that  $\text{rank}(QG_A) = n - r + r_A$ . Since  $Q$  is nonsingular,  $\text{rank}(G_A) = \text{rank}(QG_A) = n - r + r_A$ . It follows that for any  $(\mathbf{m}, \mathbf{y}) \in GF^{(n-r+r_A)}(q)$ , the solution set of (22) is nonempty and is an affine subspace with cardinality

$$\frac{q^n}{q^{n-r+r_A}} = q^{r-r_A}.$$

This proves (21). In other words, for each  $\mathbf{y} \in GF^{rA}(q)$  observed by the wiretapper, every message  $\mathbf{m} \in GF^{(n-r)}(q)$  is possible, and for each  $\mathbf{m}$ , the total number of keys  $\mathbf{k} \in GF^r(q)$  that can produce  $\mathbf{y}$  is equal to  $q^{r-rA}$ .

For all  $A \in \mathcal{A}$ ,  $\mathbf{y} \in GF^{rA}(q)$ , and  $\mathbf{m} \in GF^{(n-r)}(q)$  with  $\Pr\{M = \mathbf{m}\} > 0$ ,

$$\begin{aligned} \Pr\{Y_A = \mathbf{y} | M = \mathbf{m}\} &= \Pr\{(\mathbf{m}, K) \in D(m) \cap C(y) | M = \mathbf{m}\} \\ &= \Pr\{(\mathbf{m}, K) \in D(m) \cap C(y)\}, \end{aligned}$$

because for any fixed  $\mathbf{m}$ ,  $Y_A = \mathbf{y}$  if and only if  $(\mathbf{m}, K) \in D(m) \cap C(y)$ , and  $K$  is independent of  $M$ . Since  $K$  is uniformly distributed, by (21), we obtain

$$\Pr\{Y_A = \mathbf{y} | M = \mathbf{m}\} = q^{-r} q^{r-rA} = q^{-rA},$$

which does not depend on  $\mathbf{m}$ . Hence,  $Y_A$  is independent of  $M$ , and so the linear network code  $\mathcal{C}_s$  we have constructed by Construction 1 is secure.

We end this section with a remark. Since  $G_A$  has full rank, each column of  $F'_A$  cannot be a linear combination of  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{n-r}$ , or equivalently, the lower  $r$  components cannot be all zero. In other words, the key  $K$  is involved in each of the symbols transmitted on the channels of a wiretap set. In fact, if  $F'_A$  contains a column such that the lower  $r$  components are all zero, then the wiretapper receives a symbol which is a known linear combination of the symbols in the message  $M$ , making the code not secure.

## 8 Imperfect Secrecy

In this section, we extend our results in the previous sections for the special case that the collection  $\mathcal{A}$  of wiretap sets consists of all the  $r$ -subsets of  $\mathcal{E}$  by allowing the wiretapper to obtain a controlled amount of information about the message. Specifically, the secure condition is replaced by the condition that for all  $A \in \mathcal{A}$ ,

$$I(M; Y_A) \leq i \log q, \tag{23}$$

where  $i$  is a fixed integer satisfying  $0 \leq i \leq r$ . We will refer to this as the *imperfectly secure condition*. The integer  $i$  specifies how much information can be leaked to the wiretapper. When  $i = 0$ , the imperfectly secure condition reduces to the secure condition.

We first show that under the imperfectly secure condition, the message  $M$  can consist of at most  $(n - r + i)$   $q$ -ary symbols. Consider any code on a CSWN satisfying both the decodable condition and the imperfectly secure condition. Let  $u \in \mathcal{U}$  be such that  $\text{maxflow}(u) = n$  and  $(W, W^c)$  be a minimum cut between the source node  $s$  and node  $u$ . Consider

$$H(M) = I(Y_{\mathcal{I}}; M) + I(Y_{E_W \setminus \mathcal{I}}; M | Y_{\mathcal{I}}) \quad (24)$$

$$\leq I(Y_{\mathcal{I}}; M) + H(Y_{E_W \setminus \mathcal{I}}) \quad (25)$$

$$\leq (n - r + i) \log q, \quad (26)$$

where (24) follows from (6), and (26) follows from (23) and  $|E_W \setminus \mathcal{I}| = n - r$ , proving the claim. Note that (25) is equivalent to the imperfect secrecy theorem in [13] (p. 116).

Next, we show that tightness in (26) can be achieved by an “imperfectly secure” code obtained via Construction 1. In Construction 1, let  $K = (M^*, K')$ , where  $M^*$  is chosen randomly from  $GF^i(q)$ , and  $K'$  is independent of  $(M, M^*)$  and distributed uniformly on  $GF^{r-i}(q)$ . The input pair  $(M, K)$  in Construction 1 now becomes the triple  $(M, M^*, K')$ , where  $M' = (M, M^*)$  is regarded as the message of our imperfectly secure code which consists of  $(n - r + i)$   $q$ -ary symbols, and  $K'$  is regarded as the key of the code which consists of  $(r - i)$   $q$ -ary symbols. Since the code obtained by Construction 1 is secure when the input pair is  $(M, K)$ , we have

$$I(M; Y_A) = 0$$

for all  $A \in \mathcal{A}$ . Thus

$$I(M'; Y_A) = I(M; Y_A) + I(M^*; Y_A | M) \leq H(M^*) = i \log q,$$

i.e., the imperfectly secure condition is satisfied by the code we have constructed. Evidently, node  $u$  can decode the message  $M' = (M, M^*)$  because in the code obtained by Construction 1, the pair  $(M, K)$  can be recovered. Hence, we have obtained a code that multicasts

the maximum possible amount of information while satisfying the imperfect secure condition with the prescribed  $i$ .

Finally, we prove that the imperfectly secure code we have constructed above uses the minimum amount of randomness to achieve the required level of security when the message  $M'$  is uniformly distributed. The proof is a generalization of the corresponding proof in Section 4, so we only present the sketch here. Let  $u \in \mathcal{U}$  be such that  $\text{maxflow}(u) = n$  and  $(W, W^c)$  be a minimum cut between the source node  $s$  and node  $u$ . For any  $\mathcal{I} \subset E_W$  such that  $|\mathcal{I}| = r$ , by (6) and (23) with  $M$  replaced by  $M'$ , we have

$$H(M') \leq i \log q + I(Y_{E_W \setminus \mathcal{I}}; M' | Y_{\mathcal{I}}).$$

Summing over all  $\mathcal{I}$ , we have

$$\binom{n}{r} H(M') \leq \binom{n}{r} i \log q + \binom{n-1}{n-r-1} H(Y_{E_W}),$$

which implies

$$H(Y_{E_W}) \geq \frac{n}{n-r} (H(M') - i \log q).$$

It follows that

$$\begin{aligned} H(M') + H(K') &\geq H(Y_{E_W}) \\ &\geq \frac{n}{n-r} (H(M') - i \log q), \end{aligned}$$

or

$$H(K') \geq \frac{r}{n-r} H(M') - \frac{n}{n-r} (i \log q).$$

When the message  $M'$  is uniformly distributed,  $H(M') = (n-r+i) \log q$  and  $H(K') = (r-i) \log q$ , and it can readily be checked that the inequality above is tight. This completes the proof.

## 9 Conclusion

In this paper, we have introduced the communication system on a wiretap network (CSWN) as a model for multicasting on a network with information-theoretic security. Our model

subsumes secret sharing in classical cryptography. We have proposed a construction of a secure linear network code for a CSWN. The optimality of our construction is proved for the special case that the wiretapper may choose to access any subset of channels of a fixed size. Moreover, we have extended this construction to the scenario when the wiretapper is allowed to obtain a controlled amount of information about the message. This extended construction is also shown to be optimal.

## A A Lower Bound on $H(K)$

Assume that the collection  $\mathcal{A}$  of wiretap sets consists of all the  $r$ -subsets of  $\mathcal{E}$ . Let  $u \in \mathcal{U}$  be any user node and consider any cut  $(W, W^c)$  between the source node  $s$  and node  $u$ . Let  $|E_W| = n' \geq n$ . We will prove that

$$H(K) \geq H(Y_{\mathcal{I}}). \tag{27}$$

for any  $\mathcal{I} \subset E_W$  such that  $|\mathcal{I}| = r$ . Consider

$$H(Y_{\mathcal{I}}|M, K) \leq H(Y_{E_W}|M, K) = 0,$$

which implies

$$H(Y_{\mathcal{I}}|M, K) = 0.$$

Together with

$$H(Y_{\mathcal{I}}) = H(Y_{\mathcal{I}}|M)$$

from the security condition, we have

$$\begin{aligned} H(Y_{\mathcal{I}}) &= H(Y_{\mathcal{I}}|M) - H(Y_{\mathcal{I}}|M, K) \\ &= I(Y_{\mathcal{I}}; K|M) \\ &\leq H(K|M) \\ &= H(K), \end{aligned}$$

proving (27).

## Acknowledgment

The work of Raymond W. Yeung was partially supported by a grant from the Research Grant Committee of the Hong Kong Special Administrative Region, China (RGC Ref. No. CUHK 2/06C). The work of Ning Cai was partially supported by a grant from the National Natural Science Foundation of China (Ref. No. 60672119). Raymond Yeung would like to thank Prof. Ueli Maurer for suggesting the problem. Both authors would like to thank Prof. Te Sun Han for pointing out the references for his inequalities and for his comments on the draft.

## References

- [1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Info. Theory*, IT-46: 1204-1216, 2000.
- [2] G. R. Blakley, "Safeguarding cryptographic keys," in *Proceedings of the National Computer Conference*, 48: 313-317, 1979.
- [3] N. Cai and R. W. Yeung, "Secure network coding," IEEE International Symposium on Information Theory, Lausanne, Switzerland, Jun 30-Jul 5, 2002.
- [4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., Wiley, 2006.
- [5] J. Feldman, T. Malkin, C. Stein, and R. A. Servedio, "On the capacity of secure network coding," 42nd Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, Sept 29-Oct 1, 2004.
- [6] L. K. Ford, Jr. and D. K. Fulkerson, *Flows in Networks*, Princeton University Press, Princeton, New Jersey, 1962.
- [7] T. S. Han, "Nonnegative entropy measures of multivariate symmetric correlations," *Info. Contr.*, 36: 133-156, 1978.

- [8] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen, “Polynomial time algorithms for multicast network code construction,” *IEEE Trans. Info. Theory*, IT-51: 1973-1982, 2005.
- [9] S.-Y. R. Li, R. W. Yeung and N. Cai, “Linear network coding,” *IEEE Trans. Info. Theory*, IT-49: 371-381, 2003.
- [10] L. H. Ozarow and A. D. Wyner, “Wire-tap Channel II,” *AT&T Bell Labs. Tech. J.*, 63: 2135-2157, 1984.
- [11] A. Shamir, “How to share a secret,” *Comm. ACM*, 22: 612-613, 1979.
- [12] C. E. Shannon, “Communication theory of secrecy systems”, *Bell Sys. Tech. Journal* 28, pp. 656-715, 1949,
- [13] R. W. Yeung, *A First Course in Information Theory*, Kluwer Academic/Plenum Publishers, 2002.
- [14] R. W. Yeung, *Information Theory and Network Coding*, Springer 2008.
- [15] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, “Network coding Theory,” *Foundations and Trends in Comm. and Info. Theory*, vol. 2, nos. 4 and 5, 241-381, 2005.

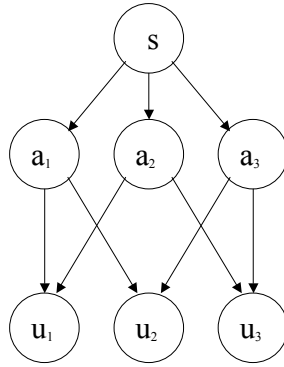


Figure 1: A CSWN representing the (1,2)-threshold secret sharing scheme.

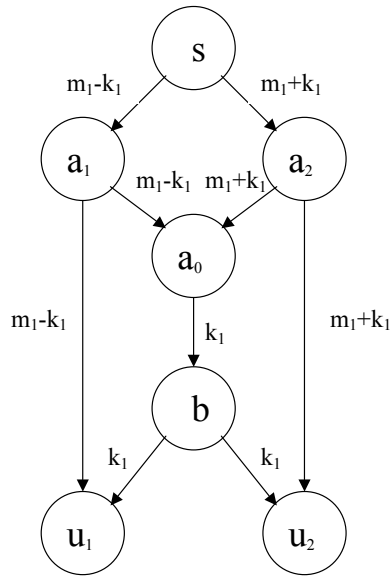


Figure 2: An example of an admissible code for a CSWN.