Universal Strongly Secure Network Coding with Dependent and Non-Uniform Messages

Ryutaroh Matsumoto, Member, IEEE, and Masahito Hayashi, Member, IEEE

Abstract—We construct a secure network coding that is strongly secure in the sense of Harada and Yamamoto [19] and universal secure in the sense of Silva and Kschischang [28], [29], while allowing arbitrary small but nonzero mutual information to the eavesdropper. Our secure network coding allows statistically dependent and non-uniform multiple secret messages, while all previous constructions of weakly or strongly secure network coding assumed independent and uniform messages, which is difficult to be ensured in practice.

Index Terms—information theoretic security, network coding, secure multiplex coding, strongly secure network coding

I. INTRODUCTION

Network coding [11] attracts much attention recently because it can offer improvements in several metrics, such as throughput and energy consumption, see [16], [17]. On the other hand, the information theoretic security [4], [25] also attracts much attention because it offers security that does not depend on a conjectured difficulty of some computational problem.

A juncture of the network coding and the information theoretic security is the secure network coding [7], [10], which prevents an eavesdropper, called Eve, from knowing the message from the legitimate sender, called Alice, to the multiple legitimate receivers by eavesdropping intermediate links up to a specified number in a network. It can be seen [13], [14] as a network coding counterpart of the traditional wiretap channel coding problem considered by Wyner [30] and subsequently others [25]. In both secure network coding and coding for wiretap channels, the secrecy is realized by including random bits into the transmitted signal by Alice so that the secret message becomes ambiguous to Eve. The inclusion of random bits, of course, decreases the information rate. In order to get rid of the decrease in the information rate, Yamamoto et al. [22] proposed the secure multiplex coding for wiretap channels, in which there is no loss of information, in which case the mutual information is zero for each i. Independently and simultaneously, Bhattacharjee and Narayanan [3] proposed the weakly secure network coding based on the same idea as [22], whose goal is also to get rid of the loss of information rate in the secure network coding. Their method [3] ensures that the mutual information between $S_i$ and Eve’s information is zero for each i. Recall that Eve’s knowledge on secret information $S_i$ is usually measured by the mutual information in the information theoretic security [4], [25]. As drawbacks, the construction depends on the network topology and coding at intermediate nodes, and the computational complexity of code construction is large.

Harada and Yamamoto [19] defined a stronger security requirement on the weakly secure network coding, which will be reviewed later, and called it as the strongly secure network coding. Then they showed its construction procedure. As [3], the construction depends on the network topology and coding at intermediate nodes, and the computational complexity of code construction is large.

In order to remove these drawbacks, Silva and Kschischang [28] proposed the universal weakly secure network coding, in which they showed an efficient code construction that can support up to two $F_q$-symbols in each $S_i$ and is independent of the network topology and coding at intermediate nodes, where $F_q$ denotes the finite field with q elements throughout this paper. The independence of coding at the source node from network topology and coding at intermediate nodes is termed universal by Silva and Kschischang in [28], [29]. They [28] also showed the existence of universal weakly secure network coding with more than two $F_q$-symbols in $S_i$, but have not shown an explicit construction.

Cai [5] removed most of drawbacks mentioned earlier. Cai proved that random linear network coding [21] gives the strongly secure network coding in the sense of [19] with arbitrarily high probability with sufficiently large finite fields. However, he did not provide evaluation of the required field size, and it seems huge. Moreover, for some applications (e.g. [8], [31]) we want to choose coding at intermediate nodes in non-random fashion.

There exists a common difficulty in all the previous constructions reviewed above. In practice, we are not sure if the multiple messages are uniform and statistically independent. However, all the previous studies [4, 25] assumed the uniformity and the independence, and without both of them their security proofs do not seem to hold. It is important to provide a security

proof for weakly and strongly secure network coding without uniformity or independence assumption. On the other hand, non-uniformity of secret messages has been considered in the ordinary secure network coding \cite{9, 32} (see also the survey \cite{6}). In \cite{6, 9, 32}, the randomness to hide a secret message was assumed to be statistically independent of the secret message, while our present study allows them to be statistically dependent.

We shall propose a construction of secure network coding. Our construction is strongly secure in the sense of \cite{19} and universal secure in the sense of \cite{28, 29}. Moreover, we do not make the uniformity nor the independence assumption on multiple secret messages. The optimality of our construction is verified under the uniformity and independence assumption at the end of Remark 7.

However, we relax an aspect of the security requirements traditionally used in the secure network coding. In previous proposals of secure network coding \cite{3, 7, 19, 28, 29}, it is required that the mutual information to the eavesdropper is exactly zero. We relax this requirement by regarding sufficiently small mutual information to be acceptable. This relaxation is similar to requiring the bit error rate to be sufficiently small instead of strictly zero. Also observe that our relaxed criterion is much stronger than one commonly used in the information theoretic security \cite{25}. Our construction can realize arbitrary small mutual information if coding over sufficiently many symbols in single packet is allowed.

This paper is organized as follows: Section II reviews related results used in this paper. Section III introduces the strengthened version of the privacy amplification theorem and the proposed scheme for secure network coding. Section IV concludes the paper.

Part of this paper was reported as earlier proceedings papers \cite{26, 27}. We substantially rewrote our security proof in \cite{27} so that we can analyze the security with dependent and non-uniform multiple secret messages, which was not done in \cite{27}. We borrowed ideas from \cite{26} so that we can prove Lemma 5.

II. Preliminary

A. Model of network and network coding

As in \cite{3, 7, 10, 19, 28, 29} we consider the single source multicast, and assume the linear network coding \cite{23, 24}. The source node is assumed to have at least \( n \) outgoing links. For \( i = 1, \ldots, n \), the source node prepares a packet \( P_i \) consisting of \( m \) symbols in \( F_q \), and transmits an \( F_q \)-linear combination of \( P_1, \ldots, P_n \) to each outgoing link, as explained in \cite{15} Section 2.1. At an intermediate node, only packets generated at the same time by the source node are linearly combined, as explained in \cite{15} Section 2.5. The linear combination coefficients at each node are fixed so that all the legitimate receivers can decode \( n \) packets \( P_1, \ldots, P_n \) from the source node.

If the random linear network coding \cite{21} is employed, we have to also include so-called encoding vectors in each packet \( P_i \). \cite{15} Section 2.2. We ignore those encoding vectors because they do not carry secret information.

Hereafter, we shall only consider the eavesdropper Eve and forget about the multiple legitimate receivers. The \( n \) packets \( P_1, \ldots, P_n \) carry in total \( mn \) symbols in \( F_q \). We shall propose a method encoding secret information into \( mn \) symbols by the source node. The \( mn \) symbols obtained by the proposed method are distributed to packets \( P_1, \ldots, P_n \).

Eve can eavesdrop \( \mu \) links. We assume \( \mu \leq n \) throughout this paper. The total number of eavesdropped symbols is therefore \( m\mu \). The set of \( \mu \) eavesdropped links is assumed to be fixed during packets \( P_1, \ldots, P_n \) are traveling on the network, as assumed in \cite{28, 29}. The situation considered here also includes the conventional store-and-forward network as a special case.

B. Security definitions

\textbf{Definition 1 (Strongly secure network coding):} \cite{19} Let \( m = 1 \), and \( S_1, \ldots, S_n \in F_q^n \) be messages. Let \( I \subset \{1, \ldots, n\} \), and \( Z \) be Eve’s observation by eavesdropping \( \mu \) links. A network coding is said to be \((\eta, \mu)\)-strongly secure if
\[
I(S_I; Z) = 0
\]
for all \( I \subset \{1, \ldots, n\} \) with \( |I| \leq \eta \), where \( S_I = [S_i : i \in I] \) and \( I(S_I; Z) \) denotes their mutual information as defined in \cite{12}.

Harada and Yamamoto \cite{19} showed a procedure to construct \((n-\mu, \mu)\)-strongly secure network coding under the uniformity and independence assumption on the messages \( S_1, \ldots, S_n \). Note that in the definition of \((\eta, \mu)\)-strong security above, we have an extra parameter \( \eta \) in addition to \( \mu \). The parameter \( \eta \) was implicit in \cite{19}.

We want to consider the universal security studied in \cite{28, 29}, and also want to use multiple symbols in a single packet \( P_i \), that is, \( m > 1 \). So we introduce our version of universal strong security, by following the approach initiated by Silva and Kschischang \cite{28, 29}.

\textbf{Definition 2:} Assume that we are given a linear network coding for single source multicast. Assume also that linear coding at intermediate nodes and the set of \( \mu \) eavesdropped links are fixed when packets \( P_1, \ldots, P_n \) travel from the source node to all the legitimate receivers. Suppose that we have \( T+1 \) messages \( S_1, \ldots, S_{T+1} \) and \( S_i \in F_q^{e_i} \). \( S_{T+1} \) denotes randomness not intended as a message. We assume \( \sum_{k=1}^{T+1} e_k = mn \). Let \( I \subset \{1, \ldots, T\} \), and \( Z \) be Eve’s observation. A linear transformation of \( S_1, \ldots, S_{T+1} \) at the source node is said to be a universal \((\eta, \mu)\)-strongly secure network coding if for all linear coding at intermediate nodes and all sets of \( \mu \) eavesdropped links we have
\[
I(S_I; Z) = 0, \quad (1)
\]
for all \( I \subset \{1, \ldots, n\} \) with \( \sum_{i \in I} e_i < m\eta \), where \( S_I = [S_i : i \in I] \).

C. Two-universal hash functions

We shall use a family of two-universal hash functions \cite{11} for the privacy amplification theorem introduced later.
Definition 3: Let \( \mathcal{F} \) be a set of functions from a finite set \( S_1 \) to another finite set \( S_2 \), and \( F \) a random variable on \( \mathcal{F} \). If for any \( x_1 \neq x_2 \in S_1 \) we have
\[
\Pr[F(x_1) = F(x_2)] \leq \frac{1}{|S_2|},
\]
then \( \mathcal{F} \) with the probability distribution of \( F \) is said to be a family of two-universal hash functions.

III. Construction of universal strongly secure network coding

A. Strengthened privacy amplification theorem

In order to evaluate the mutual information to Eve when the sum rate of multiple secret information is large, we need to strengthen the privacy amplification theorem originally appeared in [2], [20] as follows. The following theorem is a slightly enhanced version of [27, Theorem 2].

Theorem 4: Let \( X \) and \( Z \) be discrete random variables on finite sets \( X \) and \( Z \), respectively, and \( \mathcal{F} \) a family of functions from \( X \) to \( Z \). Let \( F \) be a random variable on \( \mathcal{F} \). Assume that \( X \) and \( F \) are conditionally independent given \( Z \), and that for any fixed realization \( z \) of \( Z \), the probability distribution of \( F \) given \( z \) satisfies the condition for a family of two-universal hash functions. Then we have
\[
\mathbb{E}_f \exp(\rho I(F(X); Z) = f)) \leq 1 + |S|^q \mathbb{E}_f [P_{XZ}(X|Z)]
\]
for all \( 0 \leq \rho \leq 1 \). We use the natural logarithm for all the logarithms in this paper, which include ones implicitly appearing in entropy and mutual information. Otherwise we have to adjust the above inequality.

Proof: Proof is given in Appendix B.

B. Description of the proposed scheme and analysis of its mutual information

In this section, we shall provide a universal \((n - \mu - \delta_\rho, \mu)\)-strongly secure network coding in a slightly modified sense of Definition 2, where \( \delta_\rho \) is a parameter measuring conditional non-uniformity to be defined in Eq. (10). We assume that we have \( T \) secret messages, which can be dependent or non-uniform, and that the \( i \)-th secret message is given as a random variable \( S_i \) whose realization is a row vector in \( \mathbf{F}_q^k \). The sizes \( k_i \) are determined later. We shall also use a supplementary random message \( S_{T+1} \) taking values in \( \mathbf{F}_q^{k_{T+1}} \), when the randomness in the encoder is insufficient to make \( S_i \) secret from Eve. By \( S \) we denote the entire collection \( (S_1, \ldots, S_{T+1}) \) of messages. We assume \( mn = k_1 + \cdots + k_{T+1} \).

Let \( \mathcal{L} \) be the set of all bijective \( \mathbf{F}_q \)-linear maps from \( \prod_{i=1}^{T+1} \mathbf{F}_q^k \) to itself, and \( L \) the uniform random variable on \( \mathcal{L} \) statistically independent of \( S = (S_1, \ldots, S_{T+1}) \), and arbitrary fix nonempty \( I \subseteq \{1, \ldots, T\} \). The source node store \( LS^i \) into packets \( P_1, \ldots, P_n \) defined in Section II-A and send them via its \( n \) outgoing links, where \( t \) denotes the transpose of a vector. Our construction just attaches a bijective linear function to an existing network coding. This coding scheme is illustrated in Fig. 1.

The legitimate sender and all the legitimate receivers agree on the choice of \( L \). The eavesdropper Eve may also know their choice of \( L \). Choice of \( L \) is part of protocol specification, the chosen \( L \) is repeatedly used, and agreement on its choice among legitimate senders and receivers is not counted as consumption of the network bandwidth. A legitimate receiver can recover \( S_1, \ldots, S_T, S_{T+1} \) by multiplying \( L^{-1} \) to his/her received information. By the assumption on Eve, her information can be expressed as \( BLS^i \) by using a \( mn \times mn \) matrix \( B \) over \( \mathbf{F}_q \), as in [28], [29].

For the nonempty \( I \subseteq \{1, \ldots, T\} \), denote the collection of random variables \( \{S_i : i \in I\} \) by \( S_I \), denote \( \{S_i : i \in \{1, \ldots, T+1\} \setminus I\} \) by \( S_{\bar{I}} \), and let \( k_I = \sum_{i \in I} k_i \).

For a fixed realization \( \ell \) of \( L \), the information gained by Eve is measured by the mutual information \( I(S_I ; BLS^i | L = \ell) \) as a common practice in the information theoretic security [4], [25]. Since its average \( \mathbb{E}_L I(S_I ; BLS^i | L = \ell) \) is the conditional mutual information \( I(S_I ; BLS^i | L) \), we will upper bound \( I(S_I ; BLS^i | L) \). After upper bounding the average \( I(S_I ; BLS^i | L) \) in Eq. (5), we can ensure that for most choices of \( \ell \) and all possible \( B \), \( I(S_I ; BLS^i | L = \ell) \) is small, as done in Eq. (9).

In order to use Theorem 4 we introduce a lemma.

Lemma 5: For fixed \( B \), the family of mapping \( S \mapsto BLS^i \) is a family of two-universal hash functions to the \( \text{rank}(B) \)-dimensional \( \mathbf{F}_q \)-linear space.

Proof: See Appendix B.

We can upper bound \( I(S_I ; BLS^i | L) \) as follows, by applying Theorem 4 with \( X = S, Z = S_I \), and \( F(X) = BLS^i \).

\[
\begin{align*}
\mathbb{E}_f \exp(\rho I(S_I ; BLS^i | L = \ell)) & \leq 1 + q^{\text{rank}(B)} \mathbb{E}[E[S_{S_I}^F(S_I)^\ell]] \\
& = 1 + q^{\text{rank}(B)} \mathbb{E}[E[S_{S_I}^F(S_I)^\ell]] \\
& \leq 1 + q^{\text{rank}(B)} \mathbb{E}[E[S_{S_I}^F(S_I)^\ell]].
\end{align*}
\]

From Eq. (4) we have
\[
\rho I(S_I ; BLS^i | L).
\]
there exists an $\mu \in \mathbb{R}$ to
\[ \ell \leq \frac{\ln \text{Pr}(S_{1} \mid \text{BLS}^{\ell})(S_{1} \cap S_{T}^{\mu})}{\rho} \]
(5)

Fix a real number $C_{1} > 1$. Equation (5) and the Markov inequality yield that
\[ \text{Pr}(L_{r,1} \leq 1/C_{1}) \]
for any single nonempty $I \subseteq \{1, \ldots, T\}$, where $L_{r,1} := \{\ell \mid I(S_{r} \mid \text{BLS}^{\ell})(L = \ell) \geq C_{1}/m \text{I}(I(S_{r} \mid \text{BLS}^{\ell})(L = \ell))\}$. Thus,
\[ \text{Pr}(\cup_{I \neq T} L_{r,1} < (2^{T} - 1)/C_{1}). \]
This means that a realization $\ell$ of $L$ satisfies
\[ I(S_{r} \mid \text{BLS}^{\ell})(L = \ell) \leq C_{1}/m \text{I}(I(S_{r} \mid \text{BLS}^{\ell})(L = \ell)) \]
\[ \leq C_{1}q^{\mu}(P_{S_{r} \mid S_{T}^{\mu}}(S_{1} \cap S_{T}^{\mu}))/\rho \]
(6)
for all the $(2^{T} - 1)$ nonempty subsets $I$ of $\{1, \ldots, T\}$ with probability at least $(1 - (2^{T} - 1)/C_{1})$. Defining another subset $L_{r,2} := \{\ell \mid \text{Pr}(I(S_{r} \mid \text{BLS}^{\ell})(L = \ell) > C_{1}/m \text{I}(I(S_{r} \mid \text{BLS}^{\ell})(L = \ell)))\}$, by Eq. (4) and the Markov inequality we obtain
\[ \text{Pr}(\cup_{I \neq T} (L_{r,1} \cup L_{r,2}) < (2^{T} - 1)/C_{1}). \]
Therefore, a realization $\ell$ of $L$ satisfies both Eq. (6) and
\[ \exp(\rho(I(S_{r} \mid \text{BLS}^{\ell})(L = \ell))) \leq C_{1}(1 + q^{\mu}P_{S_{r} \mid S_{T}^{\mu}}(S_{1} \cap S_{T}^{\mu}))/\rho \]
(7)
with probability at least $1 - 2 \times (2^{T} - 1)/C_{1}$.

Equation (7) implies
\[ I(S_{r} \mid \text{BLS}^{\ell})(L = \ell) \]
\[ = \frac{\ln m}{m} \ln \exp(I(S_{r} \mid \text{BLS}^{\ell})(L = \ell)) \]
\[ \leq \frac{\ln C_{1}}{m\rho} + \frac{1 + \ln q + \ln \text{E}P_{S_{r} \mid S_{T}^{\mu}}(S_{1} \cap S_{T}^{\mu})}{m\rho} \]
(8)
for $\mu \ln q + \frac{\ln \text{E}P_{S_{r} \mid S_{T}^{\mu}}(S_{1} \cap S_{T}^{\mu})}{m\rho} \geq 0$, where in Eq. (8) we used
\[ \ln(1 + \exp(x)) \leq 1 + x \text{ for } x \geq 0. \]

Up to now we considered fixed $B$. We need to ensure that the mutual information is small for any $B$ in order to show the universal security in [28], [29]. Let $x_{1}, \ldots, x_{1}$ be the $j$-th symbol in the $i$-th packet $P_{i}$ defined in Section II-A. Then there exists an $\mu \times n$ matrix $B_{\mu \times n}$ such that what is observed by Eve at the $j$-th symbols in her eavesdropped $\mu$ packets is expressed as $B_{\mu \times n}(x_{1}, \ldots, x_{m}) \cap S_{1} \cap S_{T}^{\mu}$ for $j = 1, \ldots, n$. The matrix $B$ is completely determined by $B_{\mu \times n}$. Let $A$ be any nonsingular $\mu \times \mu$ matrix over $\mathbb{F}_{q}$, and the $\mu \times \mu$ matrix $B'$ correspond to $AB_{\mu \times n}$. Then
\[ I(S_{r} \mid \text{BLS}^{\ell})(L = \ell) = I(S_{r} \mid \text{BLS}^{\ell})(L = \ell). \]

The number of nonsingular $\mu \times \mu$ matrices over $\mathbb{F}_{q}$ is
\[ \prod_{i=0}^{\mu-1}(q^{2} - q^{i}) > (q^{\mu} - (q - 1))^{\mu} = q^{2\mu}(q - 1)\mu. \]

Therefore, the number of matrices $B$’s that give different mutual information $I(S_{r} \mid \text{BLS}^{\ell})(L = \ell)$ is upper bounded by
\[ \frac{q^{\mu}}{q^{2\mu}(q - 1)\mu} \approx q^{\mu(1 + 1 - \mu)(q - 1)\mu}. \]

Therefore, the probability of $L$ satisfying Eqs. (6) and (7) simultaneously for all possible $B$ is at least
\[ 1 - 2 \times (2^{T} - 1)q^{\mu(1 + 1 - \mu)(q - 1)\mu}/C_{1}. \]
(9)

Recall that chosen $L$ is part of protocol specification and repeatedly used. Because Eqs. (6), (7) and (8) are independent of realization of the random variable $S$ representing secret information, Eqs. (6) and (8) are satisfied in every repeated use of $L$ with probability at least Eq. (9).

We need to clarify under what condition Eq. (6) converges to zero as $m \to \infty$. To do so, we shall introduce a version of conditional Rényi entropy introduced in [20]. There seems to be no standard definition for the conditional Rényi entropy, for example, definitions in [2] and [18] disagree and our definition in [20] is different from [2], [18]. For discrete random variables $X, Y$, define conditional Rényi entropy of order 1 $+ \rho$ as
\[ H_{1+\rho}(X|Y) = -\frac{\ln \text{E}P_{X|Y}(X|Y)}{\rho} \]
For $\rho = 0$, we define $H_{1}(X|Y) = \lim_{\rho \to 0} H_{1+\rho}(X|Y)$. By using l’Hôpital’s rule we see that $H_{1}(X|Y)$ is equal to the conditional Shannon entropy. Observe also that $H_{1+\rho}(X|Y) = \ln |X|$ if $X$ is conditionally uniform given $Y$, where $X$ denotes the alphabet of $X$.

In order to clarify under what condition Eq. (6) converges to zero, we need to assume some knowledge on $P_{S_{r} \mid \mu \times n}(S_{1} \cap S_{T}^{\mu})$. We consider the situation with which each message $S_{r}$ originates from a different organization and it is compressed before network coded. Under such situation, we assume that $S_{1} \cap S_{T}^{\mu}$ is nearly conditionally uniform given $S_{r}$. We assume that we have a nonnegative constant $\delta_{p}$ such that
\[ (n - k_{T}/m) \ln q - \frac{H_{1+\rho}(S_{1} \cap S_{T}^{\mu})}{m} \leq \delta_{p} \ln q \]
(10)
for some $0 < \rho \leq 1$, for all $T$, and for sufficiently large $m$. Observe that if all messages $S_{r}$’s are uniform and independent then $\delta_{p} = 0$. The parameter $\delta_{p}$ captures the deviation from the uniform and independent situation in terms of conditional Rényi entropy per the number of symbols in single packet. By taking the natural logarithm of Eq. (9), we see
\[ \ln \text{RHS of Eq. (9)} = \frac{C_{1}}{\rho} + \frac{m \rho(1 + \ln \text{E}P_{S_{r} \mid S_{T}^{\mu}}(S_{1} \cap S_{T}^{\mu}))}{m} \ln q \]
(9)
\[ = \frac{C_{1}}{\rho} + \frac{m \rho - \frac{H_{1+\rho}(S_{1} \cap S_{T}^{\mu})}{m} \ln q}{m} \]
(11)
When
\[ \mu < (n - k_{T}/m) - \delta_{p} \text{ i.e. } k_{T}/m < n - \mu - \delta_{p}, \]
(12)
(*) in Eq. (11) becomes negative as $m \to \infty$. Under such condition Eq. (11) converges to $-\infty$ as $m \to \infty$, which means
that the upper bound Eq. (6) can be made arbitrary small by letting
$m$ be large.

We will analyze how much information Eve can gain when
Eq. (12) does not hold. In such case we use the other upper
bound Eq. (8). We can rewrite Eq. (8) as
\[
\frac{\text{RHS of Eq. (8)}}{m} = 1 + \ln C_1 + \frac{\mu \ln q - \frac{H_{1\epsilon}(S_{1}\|S_{T})}{m}}{m} \leq \frac{1 + \ln C_1}{m} + \left(\mu - \frac{1}{m} \right) \ln q. \tag{10}
\]
We see that we can make the upper bound Eq. (8) on the
equivocation rate \(\frac{\text{RHS of Eq. (8)}}{m}\) arbitrary close to
\[\left(\mu + \delta_\rho - \left(\frac{n - k_1}{m}\right)\ln q\right)\tag{13}\]
by letting \(m\) be large.

By the above construction and evaluation of mutual
information, we provide a universal \((n - \mu - \delta_\rho, \mu)\)-strongly
secure network coding in the sense of Definition 2 with the
existence of \(C_1\) and the RHS of Eq. (6) is < \(\frac{1}{m}\) for all \(\alpha\) < 
\(\frac{1}{m}\). Through this, we ensure the
arbitrary small one (see also Remark 7).

Remark 6: The meaning of \(C_1\) is as follows: At Eqs. (4) and
(5), there might not exist a realization \(\ell\) of \(L\) that satisfies Eqs. (4) and (5) for all subsets \(I\) of \(\{1, \ldots, T\}\) simultaneously.

By sacrificing the tightness of the upper bounds, we ensure the
existence of \(\ell\) satisfying Eqs. (6) and (7).

Remark 7: Under the assumption that all messages \(S_1, \ldots, S_{T+1}\)
are independent, the mutual information can be made exactly zero for every eavesdropping matrix \(B\). The
reason is as follows: For fixed \(B\) and \(L = \ell\), we have
\[I(S_{1}\|BLS_{1}|L = \ell) = H(S_{1}|L = \ell) - H(S_{1}\|BLS_{1}, L = \ell).\tag{14}\]

The first term \(H(S_{1}|L = \ell)\) is an integer multiple of \(\ln q\) since
\(S_{1}\) is assumed to have the uniform distribution. Let \(\alpha_{r}\) be the projection from \(\prod_{i=1}^{\ell} F_{q}\) to \(\prod_{i \in I} F_{q}\) for \(\emptyset \neq I \subseteq \{1, \ldots, T\}\). For fixed \(B\) and \(\ell\), and a given realization \(z\) of \(BLS_{1}\), the set of solutions \(s\) such that \(z = B\ell s\) is written as \(\ker(B)\)\+ some vector \(v\). This means that the set of possible candidates of \(S_{1}\) given realization \(z\) of \(BLS_{1}\) is written as \(\alpha_{r}(\ker(B)) + \alpha_{r}(v)\), and \(S_{1}\) given realization \(z\) is uniformly distributed on \(\alpha_{r}(\ker(B)) + \alpha_{r}(v)\). Since the cardinality of \(\alpha_{r}(\ker(B)) + \alpha_{r}(v)\) is independent of \(S_{1}\) for fixed \(B\) and \(\ell\), the second term \(H(S_{1}\|BLS_{1}, L = \ell)\) is also an integer multiple of \(\ln q\). Therefore, if Eq. (6) holds for every \(B\) as verified in Eq. (9)
and the RHS of Eq. (9) is < \(\ln q\), then the LHS of Eq. (9) must be zero. Observe that under this assumption our construction is a universal \((n - \mu, \mu)\)-strongly secure network coding in the
exact sense of Definition 2. The parameter \((n - \mu, \mu)\) is optimal according to [6].

C. Numerical example of explicit computation of required
block size \(m\)

In this section we give a numerical example of computing
required block length \(m\) in order to ensure the mutual
information is below some value. In order to do so, we need an estimate of \(E[P_{S_{2}\|S_{1}, L = \ell}(S_{1}|S_{T})]\). We assume to have \(\delta_{0.5} = 5\) in Eq. (10) at \(\rho = 0.5\).

Let \(q = 256, n = 100, \mu = 30, T = 5, k_1 = 20m\) for all \(i\). We do not have \(S_{T+1}\). We want to ensure that we choose \(\ell\) with
probability at least \(1 - 10^{-12}\) such that \(I(S_{1}\|BLS_{1}|L = \ell) < \frac{1}{m}\) for all \(i = 1, \ldots, 5\). By Eq. (9) we choose \(C_1\) as
\[2 \times q^{\mu(n - 1 - \mu)}(q - 1)^{-\mu}(2^T - 1)/C_1 \leq 10^{-12}\]
\[C_1 = 2 \times 256^{\mu(n - 1 - \mu)}255^{-\mu}(2^T - 1)10^{12}\]
By using the \(\delta_\rho\), we can upper bound the RHS of Eq. (9) as follows:
\[C_1 q^{\mu}E[P_{S_{2}\|S_{1}, L = \ell}(S_{1}|S_{T})]/\rho = C_1 \exp_\rho(\mu + \frac{H_{1\epsilon}(S_{1}\|S_{T})}{m\ln q}) \leq C_1 \exp_\rho(\mu(n - k_1/m + \delta_\rho))/\rho \tag{15}\]
In order to keep the above upper bound to be below \(10^{-6}\) we have to choose
\[C_1 \exp_\rho(\mu(n - k_1/m + \delta_\rho))/\rho < 10^{-6}\]
\[m > -\log_{256}(10^6 \times 2 \times 256^{171}10^{30}255^{20}25^{12}/0.5)\leq 94.8 \times 9400/9400 \times 65110^{12}/0.5\]
\[m > \frac{0.5(30 - 100 + 20 + 5)}{9400} \geq 94\]
This means that we choose the matrix \(L\) at least as large as
9400×9400 over \(F_{256}\), which seems implementable. Recall that we assumed \(n = 100\) outgoing (logical) links from the source
node and that each outgoing link carries \(m = 94\) symbols in
single coding block in this example. With this example, we can also see that the exponent \(n\mu\) in Eq. (9) has large influence in
the required size of \(m\).

Remark 8: A vector in \(F_{q}^{mn}\) can be identified with an
element in \(F_{q^{*}}^{\mu}\), and multiplication by a nonzero element in
\(F_{q^{*}}^{\mu}\) is a \(F_{q}\)-linear mapping and can be identified with an
element in \(L\). Let \(L_{k}\) be a commutative subgroup of \(L\)
whose elements can be identified with nonzero elements in
\(F_{q}\). By looking at the proof of Lemma 5 in Appendix B
we can see that \(L_{k}\) can be used in place of \(L\) in our
construction. Necessary storage space to record choice of an
element in \(L_{k}\) is that of \(m q\) symbols and is smaller than
that of \(L\). Matrix multiplication by an element in \(L_{k}\) is at
least as fast as that in \(L\).

IV. Conclusion

In the secure network coding, there was loss of information
rate due to inclusion of random bits at the source node. Weakly
and strongly secure network coding remove that loss of information rate by using multiple messages to be kept secret from an eavesdropper, which require huge computational complexity in code construction or huge finite field size. In addition to this, the previous studies assumed
uniform and independent multiple messages, which seems too
strong assumption in practice. In this paper, we have shown
that random linear transform of multiple messages at the
source node realizes the strongly secure network coding with
arbitrary high probability with sufficiently large block length.
We did not assume uniformity nor independence in multiple
messages. Our numerical example in Section III-C showed that “sufficiently large block length” can be small.

**APPENDIX A**

**Proof of Theorem 4**

In order to show Theorem 4, we introduce the following lemma.

**Lemma 9:** Under the same assumption as Theorem 4, we have

\[ E_f \exp(-\rho H(F(X); Z|F = f)) \leq |S|^{-\rho} + E[P_{X|Z}(X|Z')^\rho] \]  

for 0 \leq \rho \leq 1.

**Proof of Theorem 4**

\[ E_f \exp(\rho H(F(X); Z|F = f)) = E_f \exp(\rho H(F(X)|F = f) - \rho H(F(X); Z|F = f)) \leq E_f |S|^{-\rho} \exp(-\rho H(F(X); Z|F = f)) \leq E_f |S|^{-\rho} + E[P_{X|Z}(X|Z')^\rho] \]  

(by Eq. (15))

\[ = 1 + |S|^{-\rho} E[P_{X|Z}(X|Z')^\rho]. \]

**Proof of Lemma 9**

Fix \( z \in \mathbb{Z} \). With a fixed realization \( z \) of \( Z \), by the assumption in Theorem 4, two random variables \( F \) and \( X \) are statistically independent. The concavity of \( x^\rho \) for \( 0 \leq \rho \leq 1 \) implies

\[ E_f \sum_{s \in S} P_{f|X}(s|z)^{1+\rho} = \sum_{x \in X} P_{X|Z}(x|z) E_f \left( \sum_{x' \in f^{-1}(f(x))} P_{X|Z}(x'|z) \right)^\rho \leq \sum_{x \in X} P_{X|Z}(x|z) \left( E_f \left( \sum_{x' \in f^{-1}(f(x))} P_{X|Z}(x'|z) \right) \right)^\rho. \]  

(17)

Since \( f \) is chosen from a family of two-universal hash functions defined in Definition 3, we have

\[ (**) \leq P_{X|Z}(x|z) + \sum_{x' \in f^{-1}(f(x))} P_{X|Z}(x'|z) \leq P_{X|Z}(x|z) + |S|^{-1}. \]

Since any two positive numbers \( x \) and \( y \) satisfy \((x+y)^\rho \leq x^\rho + y^\rho \) for \( 0 \leq \rho \leq 1 \), we have

\[ (P_{X|Z}(x|z) + |S|^{-1})^\rho \leq P_{X|Z}(x|z)^\rho + |S|^{-\rho}. \]  

(18)

By Eqs. (17) and (18) we can see

\[ E_f \sum_{s \in S} P_{f|X}(s|z)^{1+\rho} \leq \sum_{s \in S} P_{X|Z}(s|z)^{1+\rho} + |S|^{-\rho}. \]

Taking the average over \( Z \) of the both sides of the last equation, we have

\[ E_f E_{XZ} P_{f|XZ}(f(X)|Z)^\rho \leq E_{XZ} P_{X|Z}(X|Z)^\rho + |S|^{-\rho}. \]  

(19)

Define \( g(\rho) = E_{XZ} P_{f|XZ}(f(X)|Z)^\rho \ln P_{f|XZ}(f(X)|Z) \) as a function of \( \rho \) with fixed \( f \) and \( P_{XZ} \), and \( h(\rho) = \ln g(\rho) \). We have

\[ g'(\rho) = E_{XZ} P_{f|XZ}(f(X)|Z)^\rho \ln P_{f|XZ}(f(X)|Z), \]

\[ g''(\rho) = E_{XZ} P_{f|XZ}(f(X)|Z)^\rho (\ln P_{f|XZ}(f(X)|Z))^2, \]

\[ h'(\rho) = g'(\rho)/g(\rho), \]

\[ h''(\rho) = \frac{g''(\rho)g(\rho) - [g'(\rho)]^2}{g(\rho)^2}. \]

Define \( (X', Z') \) to be the random variables that have the same joint distribution as \( (X, Z) \) and statistically independent of \( X \) and \( Z \). To examine the sign of \( h''(\rho) \) we compute

\[ g''(\rho)g(\rho) - [g'(\rho)]^2 \]

\[ = E_{XX'ZZ'} P_{f|XZ}(f(X), Z) P_{f|XZ}(f(X'), Z') \left[ \ln P_{f|XZ}(f(X)|Z) - \ln P_{f|XZ}(f(X)|Z) \ln P_{f|XZ}(f(X')|Z') \right] \]

\[ - \frac{1}{2} E_{XX'ZZ'} P_{f|XZ}(f(X), Z) P_{f|XZ}(f(X'), Z') \left[ \ln P_{f|XZ}(f(X)|Z)^2 + \ln P_{f|XZ}(f(X')|Z')^2 - 2 \ln P_{f|XZ}(f(X)|Z) \ln P_{f|XZ}(f(X')|Z') \right] \]

\[ \geq 0. \]

This means that \( h''(\rho) \geq 0 \) and \( h(\rho) \) is convex. We can see

\[ E_{XZ} P_{f|XZ}(f(X)|Z)^\rho = \exp(h(\rho)) \geq \exp(\mathcal{h}(0) + \rho h'(0)) \]

\[ = \exp(-\rho H(f(X)|Z)). \]  

(20)

By Eqs. (19) and (20) we see that Eq. (16) holds.

**APPENDIX B**

**Proof of Lemma 5**

We shall prove Lemma 5 in this Appendix. Let \( \mathcal{L} \) be a subgroup of the group of all bijective linear maps on \( F_q^{mn} \). For \( \vec{x} \in F_q^{mn} \), the orbit \( O(\vec{x}) \) of \( \vec{x} \) under the action of \( \mathcal{L} \) is defined by

\[ O(\vec{x}) = \{ L\vec{x} \mid L \in \mathcal{L} \}. \]

**Lemma 10:** Let \( \vec{x}, \vec{y} \) be two different vectors belonging to \( O(\vec{z}) \). We have

\[ ||L \in \mathcal{L} \mid L\vec{x} = \vec{x}|| = ||L \in \mathcal{L} \mid L\vec{z} = \vec{y}||. \]

**Proof:** Let \( K \in \mathcal{L} \) such that \( K\vec{x} = \vec{y} \). We have

\[ ||L \in \mathcal{L} \mid L\vec{z} = \vec{x}|| = ||L \in \mathcal{L} \mid K\vec{z} = \vec{x}|| = ||L \in \mathcal{L} \mid L\vec{z} = \vec{x}|| = ||L \in \mathcal{L} \mid L\vec{z} = \vec{x}||. \]

**Lemma 11:** Let \( B \) be an \( m \times m \) matrix, \( \ker(B) = \{ \vec{x} \in F_q^{mn} \mid B\vec{x} = \vec{0} \} \), and \( \text{im}(B) = \{ B\vec{x} \mid \vec{x} \in F_q^{mn} \} \). The family of functions \( \{ BL \mid L \in \mathcal{L} \} \) with uniformly distributed \( L \) is a family of two-universal hash functions from \( F_q^{mn} \) to \( \text{im}(B) \) if and only if

\[ \frac{|O(\vec{v}) \cap \ker(B)|}{|O(\vec{v})|} \leq \frac{1}{|\text{im}(B)|} \]

for all \( \vec{v} \in F_q^{mn} \setminus \{ \vec{0} \} \).
Proof: With the uniform distribution on \( L \), LHS of Eq. (2) is equal to
\[
\frac{|\{L \in L \mid BL(x_1 - x_2) = \bar{0}\}|}{|L|} = \frac{|\{L \in L \mid L(x_1 - x_2) \in O(x_1 - x_2) \cap \ker(B)\}|}{|L|} = \frac{|O(x_1 - x_2) \cap \ker(B)|}{|O(x_1 - x_2)|} \quad \text{(by Lemma 10)}.
\]
Renaming \( \bar{x}_1 - \bar{x}_2 \) to \( \bar{v} \) proves the lemma. ■

**Proposition 12:** If \( L \) is the set of all bijective linear maps on \( F_q^{mn} \), then \( \{BL \mid L \in L\} \) uniformly distributed \( L \) is a family of two-universal hash functions from \( F_q^{mn} \) to \( \text{im}(B) \).

**Proof:** For a nonzero \( \bar{v} \in F_q^{mn} \), we have \( O(\bar{v}) = F_q^{mn} \setminus \{\bar{0}\} \), which implies
\[
O(\bar{v}) = F_q^{mn} - 1, \quad O(\bar{v}) \cap \ker(B) = \frac{F_q^{mn}}{\text{im}(B)} - 1.
\]
By Lemma 11 we can see that the proposition is true. ■

**Proof of Lemma 5:** Lemma 5 is equivalent to Proposition 12. ■

**Acknowledgment**

The authors thank anonymous reviewers of NetCod 2011 for carefully reading the initial manuscript and pointing out its shortcomings. The first author would like to thank Prof. H. Yamamoto to teach him the secure multiplex coding, Dr. S. Watanabe to point out the relation between the proposed scheme and [19], J. Kurihara to point out the relation between the proposed scheme and [28], Dr. J. Muramatsu and Prof. T. Ogawa for the helpful discussion on the universal coding. A part of this research was done during the first author’s stay at the Institute of Network Coding, the Chinese University of Hong Kong, and Department of Mathematical Sciences, Aalborg University. He greatly appreciates the hospitality by Prof. R. Yeung and Prof. O. Geil.

**References**


