# Secure Network Coding via Filtered Secret Sharing*


Jon Feldman† Dept. of IEOR

Tal Malkin‡ Dept. of CS

Rocco A. Servedio§ Dept. of CS

Cliff Stein¶ Dept. of IEOR

Columbia University, New York, NY

{jonfeld@ieor, tal@cs, rocco@cs, cliff@ieor}.columbia.edu



**Abstract**

We study the problem of using a multicast network code to transmit information securely in the presence of a "wire-tap" adversary who can eavesdrop on a bounded number of network edges. We establish a close connection between secure linear network coding and a new variant of the secret sharing problem, which we call *filtered secret sharing*. Using this connection, we establish new trade-offs between security, capacity, and bandwidth of secure linear network coding schemes. Our positive results show that by giving up a small amount of capacity, it is possible to dramatically reduce the bandwidth requirements of secure linear network coding. Our negative results show that within the framework we consider, unless capacity is relaxed, the bandwidth requirements can be prohibitively high.


## 1 Introduction

Networks that carry information are now ubiquitous, and so the problem of using them efficiently is critical. One of the most exciting new ideas of the last few years in the theoretical study of information networks is *network coding*. This line of research (e.g., [20, 14, 22, 23, 12], see also [21]), introduced by Ahlswede *et al.* [1], differs from traditional work on routing in networks in the following way. A packet sent through a network consists of routing information and data. Traditionally, routers manipulate the routing information, and just pass along the data. In network coding, we allow the routers to manipulate the data, i.e. we allow the network to do computation on the data. It has been shown [1, 23, 14] that by doing so, we can increase the effective capacity of the network. Network coding has been suggested as a practical tool for use in content distribution networks over the Internet [13, 24], as well as for wireless networks [7, 29].

In a traditional multicast situation with a single source and multiple destinations, the amount of information that can be transmitted from the source $s$ to a particular destination $t_i$ is equal to the minimum cut $\kappa_i$ between the source and destination. If we allow coding at the routers, we obtain the surprising result that we can *simultaneously* transmit $n = \min_i \kappa_i$ symbols of information to *every* destination [23]. Furthermore, given a network, we can construct such a network code in polynomial time [18]. In contrast, there are simple examples of networks in which this is not possible with traditional routing [15].

As any user of the Internet is painfully aware, it is imperative to consider security issues in any network scenario. To that end, several researchers have considered security issues in network coding. The problem of making a linear network code secure was first studied by Cai & Yeung [5], who considered a "wire-tap" adversary that can look at a bounded number of network edges. Jain [19] also considers this model, and gives more precise security conditions in certain cases. Ho *et al.* [16] consider the related problem of network coding in the presence of a *Byzantine attacker* who can modify data sent from a node in the network.


*A preliminary version of this work [11] appeared as an invited presentation at Allerton 2004.

†J. Feldman was supported by an NSF Postdoctoral Research Fellowship DMS-0303407.

‡T. Malkin was partially supported by NSF Early Career Development (CAREER) Grant CCF-0347839.

§R. Servedio was partially supported by NSF Early Career Development (CAREER) Grant CCF-0347282.

¶C. Stein was partially supported by NSF Grant DMI-9970063.

In this paper we study secure multicast network coding against a wire-tap adversary where perfect (information-theoretic) security is required. We abstract away the network topology and reduce the problem of information-theoretically secure linear network coding to a new variant of secret sharing, which we call *filtered secret sharing* and believe to be of independent interest. Informally, while in classical threshold secret sharing security is maintained against an adversary who receives at most $k$ of the $n$ shares, in filtered secret sharing the adversary receives at most $k$ among a set of $N \geq n$ *fixed linear combinations of all* $n$ *shares*. In other words, the shares of the secret are passed through some fixed $n$-by-$N$ linear filter, and then $k$ out of $N$ of these combinations are given to the adversary. This filtered secret sharing problem is investigated using techniques from secret sharing and from classical coding theory.

**1.1 Motivation For Our Work** Making a system secure always comes at a cost. For example, if one uses cryptography, one pays a cost in computation time. In network coding, the cost is that less information can be transmitted in each time step. More precisely we will study trade-offs between *security, bandwidth and capacity* in linear multicast network coding schemes. We will later define each of these terms more precisely, but give an informal definition here. *Security* is characterized by how many edges an adversary can observe without obtaining any information about the message in the network. Information is transmitted as elements of a finite field $\mathbb{F}_q$. The logarithm of the field size is the edge *bandwidth*, or how much information (in bits) needs to travel through an edge in one step. In many applications, an edge will have a physical upper limit on bandwidth; this will force us to make the bandwidth of our code small. For security, random symbols will be transmitted along with the information symbols; we measure the *capacity* of the network code as the number of information symbols transmitted in each step. The overall goal is to operate at a capacity close to the minimum cut value $n = \min_i \kappa_i$ and be secure against an adversary who can view many edges, under possibly limited edge bandwidth.

Cai & Yeung [5] considered one particular setting of security, bandwidth and capacity. Specifically, if $n$ is the minimum cut value in the underlying network of $N$ edges, and $k < n$ is the bound on the number of edges available to the adversary, they demonstrate the existence of a scheme with capacity $n - k$ as long as the edge bandwidth is greater than $\log \binom{N}{k}$. This result has two main drawbacks: (i) the construction of the scheme takes $\binom{N}{k}$ steps, and (ii) the bandwidth requirement is prohibitive for large $k$. Note that in the absence of security considerations the bandwidth requirement is at most the logarithm of the number of terminals in the network, and hence is at most $\log N$ [23, 18].

**1.2 Our Results** We exhibit new trade-offs between security, bandwidth and capacity of secure linear network coding schemes. We give positive results on achievable parameters that are more powerful than those previously known. We also give new negative results showing that filtered secret sharing is unsolvable in certain cases.

We first show that by giving up a little bit of capacity (namely, sending $n - \Theta(k)$ symbols instead of $n - k$), we can efficiently construct a scheme that is secure with high probability, where the required bandwidth is only $\Theta(\log N)$, independent of $k$. This bound is superior to the bound of $\log \binom{N}{k}$ in most cases, and allows a trade-off between capacity and field size. For very large $k = \Theta(N)$, our bandwidth requirement becomes $\Theta(1)$, independent of both $N$ and $k$.

Our negative result gives further support to our approach of giving up capacity in order to achieve security with a small bandwidth. We show that if one insists upon sending $n - k$ message symbols, then there are cases where the bandwidth must be almost as large as $\Theta(\sqrt{k} \log N)$. (We give more precise statements of both our positive and negative results later in the paper.)

**1.3 Techniques** As mentioned above, we reduce the secure network coding problem to a variant of secret sharing, which we call filtered secret sharing. We then show that filtered secret sharing is actually equivalent to a certain generalized (classical) code construction problem. More precisely, we study the problem of designing a code that has large distance from a given code. Within this framework, we derive positive results using methods similar to those used in a proof of the Gilbert-Varshamov bound (see [17]), and negative results using a bound [9] on the covering radius [8] that linear codes can achieve.

Our method for constructing a secure network code has a nice feature that makes it more useful when the network code is fixed (in hardware, say). If we are given a network and a network code, we can make this code secure *without changing the network code*, but only by applying a linear transformation to the input. Our ability to do this follows from a linear algebraic approach to network coding which actually abstracts away the network topology, along the lines of [20].

**1.4 Organization of the paper** Section 3 describes filtered secret sharing. Our main positive and negative results on filtered secret sharing are given in Section 3.3. Section 4 introduces the network coding model, exhibits a reduction from secure network coding to filtered secret sharing, and states our main positive and negative results for secure network coding (these results follow immediately from the results of Section 3.3 using the reduction).

In Sections 5 through 8 we prove the results of Section 3.3. Section 5 gives basic results on filtered secret sharing, and a characterization of linear solutions. In Section 6, we prove that filtered secret sharing is equivalent to a generalized (error-correcting) code construction problem. Finally, in Sections 7 and 8, we give our positive and negative results, respectively.

## 2 Preliminaries

**2.1 Notation** Throughout the paper all vectors $v$ are row vectors unless otherwise indicated, and we write $v^T$ to denote the corresponding column vector. If $v$ is an $n$-dimensional row vector and $w$ is an $m$-dimensional row vector we write $(v, w)$ to denote the $(n+m)$-dimensional row vector obtained by concatenating $v$ and $w$. We use $[n]$ to denote the set $\{1, \ldots, n\}$. Given $x \in \mathbb{F}_q^N$, the *ball of radius* $d$ around $x$ is the set of all vectors in $\mathbb{F}_q^N$ which differ from $x$ in at most $d$ coordinates. We write $\mathrm{Vol}_q(d, N)$ to denote the number of vectors in this ball.

**2.2 Information-theoretic security against a $k$-threshold adversary** We define a general "threshold" security condition that we use for both secret-sharing and secure network coding. Suppose we have some information source that produces arbitrary $x \in \mathbb{F}_q^t$. Let $f(x, r)$ be a function $f : \mathbb{F}_q^t \times \mathbb{F}_q^\ell \to \mathbb{F}_q^N$, where we think of the input $r \in \mathbb{F}_q^\ell$ as random. For $I \subseteq [N]$, we write $f_I(x, r) \in \mathbb{F}_q^{|I|}$ to denote the vector $f(x, r)$ restricted to the indices in $I$.

For $I \subseteq [N]$, we say that $f$ is *secure against* $I$ if for all $x, x' \in \mathbb{F}_q^t$, we have that the random variables $f_I(x, r)$ and $f_I(x', r)$ are identically distributed (here the randomness in each case is over the uniform choice of $r$ from $\mathbb{F}_q^\ell$). We say that $f$ is *secure against a $k$-threshold adversary* if $f$ is secure against $I$ for all $I \subseteq [N]$ with $|I| \leq k$. In other words, for any adversary who has access to at most $k$ indices of $f(x, r)$, the view of the adversary is independent of the information $x$.

## 3 Filtered Secret Sharing

**3.1 Secret Sharing** Informally, a secret sharing scheme allows a dealer to share a secret to $n$ parties, such that any set of at least $l$ parties can reconstruct the secret from their shares, but any adversary controlling at most $k$ parties cannot gain any information about the secret from their shares.[1] The vast majority of studied schemes are *linear*, namely the secret is viewed as an element in a finite field, and the shares are constructed by applying a linear transformation to the secret and some random field elements. The main measure of efficiency for secret sharing schemes is the total size of the shares, though computational efficiency of the

[1]While more general access structures are possible, we focus here on the most commonly studied threshold secret sharing case.

secret generation and reconstruction is also often required.

Secret sharing schemes were first introduced in [3, 25], who gave linear schemes with $l = k + 1$. Secret sharing schemes for any $k < l \leq n$ are introduced in [4], who call them *ramp schemes*, and consider longer secrets consisting of $l - k$ (rather than one) field elements.

Of particular interest to us is the case of $l = n$, namely all $n$ shares allow reconstruction of the secret, but any set of up to $k$ shares give no information about the secret. We focus on this case when introducing the generalized notion of filtered secret sharing below, since this is the case needed for our secure network coding application. Filtered secret sharing can be similarly defined for the more general case.

Since their introduction, secret sharing schemes were extensively studied, and found a variety of applications, e.g., for general secure computation [2, 6, 10]. We refer the reader to [26, 27, 28] for surveys and further references.

**3.2 Filtered Secret Sharing** We introduce the following generalization of the basic secret-sharing problem described above. Here the adversary gets her shares through some fixed linear "filter;" i.e., instead of having access to at most $k$ shares, she receives at most $k$ among a fixed set of $N > n$ *linear combinations* of all $n$ shares.

**Definition 1 Filtered Secret Sharing.** *The input consists of a prime power $q$, a number of shares $n$, a "filter" length $N \geq n$, an $n$-by-$N$ full-rank filter matrix $V$ over elements in $\mathbb{F}_q$ and a threshold $k < n$. The problem is to find a number $\ell$, an "information length" $t \leq n$, and a function $S : \mathbb{F}_q^t \times \mathbb{F}_q^\ell \to \mathbb{F}_q^n$ such that the function $S(x, r)V$ (which is $\mathbb{F}_q^t \times \mathbb{F}_q^\ell \to \mathbb{F}_q^N$) is information-theoretically secure against a $k$-threshold adversary. Moreover, for any given $y = S(x, r)$, the information $x$ must be uniquely and efficiently recoverable from $y$.*

Note that classical threshold secret sharing as described in Section 3.1 is the special case where the filter $V$ is the $n$-by-$n$ identity matrix.

In solutions to filtered secret sharing it is desirable for the information length $t$ to be large. It is straightforward to show that no solution $(t, \ell, S)$ exists to filtered secret sharing unless $t \leq n - k$ (see Appendix A). Cai and Yeung [5] implicitly show that a solution always exists with $t = n - k$, as long as $q > \binom{N}{k}$.

**3.3 Our Results on Filtered Secret Sharing** Our main results about filtered secret sharing are the following two theorems. (As will be clear later, we actually achieve more general results; here we highlight particular cases of interest.)

Our first theorem shows that by reducing the information length (compared to $t = n - k$), it is possible to solve the filtered secret sharing problem with a much smaller field size requirement:

**Theorem 1** *Let $(q, n, N, V, k)$ be an arbitrary instance of filtered secret sharing as in Definition 1. If $t \leq n - \sigma k$ with $\sigma > 1$ and $q = N^{\Omega(\frac{1}{\sigma - 1})}$, then there exists a poly(N)-time randomized algorithm that outputs (with high probability) a solution $(\ell, t, S)$ with $\ell = \sigma k$ and a linear function $S : \mathbb{F}_q^t \times \mathbb{F}_q^\ell \to \mathbb{F}_q^n$.*

A more precise characterization of our bound on $q$ can be found in the discussion of Section 7.1. We also note that for the case $k = \Theta(N)$, the bound on $q$ above can be replaced by $q = 2^{\Omega(\frac{1}{\sigma - 1})}$, a constant independent of $N$.

Our second theorem shows that if we insist on the information length $t$ being as large as $n - k$, then for a wide range of parameters filtered secret sharing is not possible with a linear function $S$ unless the field size $q$ is very large:

**Theorem 2** *For all sufficiently large $N$ and all constants $0 < \xi < 1$, there exist values $k, n$ with $k = \Theta(n^\xi)$, $k < n \leq N$, such that for any $q \leq N^{O(\sqrt{k/\log k})}$, there is an $n$-by-$N$ full rank filter matrix $V$ over $\mathbb{F}_q$ such that the filtered secret sharing instance $(q, n, N, V, k)$ has no solution $(\ell, t, S)$ with $t = n - k$ and $S$ linear.*

## 4 Network Coding

**4.1 The Network Coding Model** An instance of the *multicast network coding* problem consists of a directed acyclic graph $G = (V_G, E_G)$, a source node $s_G$, a set $T_G$ of sink nodes, a message length $n$, and a field $\mathbb{F}_q$. The edges of $G$ are used to transmit information through the graph; each edge carries one element of $\mathbb{F}_q$ per time step. (Timing issues within the network are not considered in this model; the information travels across the entire network in one "time step.") The *bandwidth* of an edge is $\log q$, i.e., the number of bits carried by the edge in each time step.

A solution to the multicast network coding problem is a scheme whereby an arbitrary message vector $\mathbf{m} \in \mathbb{F}_q^n$, which originates at the source node $s_G$, is communicated over this network so that each sink can recover the entire vector $\mathbf{m}$. More precisely, a solution consists of a collection of $|E_G|$ many functions $f_{(u,v)}$, one for each edge $(u, v)$ in $E_G$, with the following properties: (i) For each edge $(u, v)$ the symbol transmitted over $(u, v)$ is the value of $f_{(u,v)}$ applied to the symbols that are available at node $u$. If $u$ is the source, the entire message vector $\mathbf{m}$ is available; otherwise the symbols transmitted on edges $(w, u)$ into node $u$ are available. (ii) For each sink node $v$ in $T_G$ there must be some function $f_v$ which, if applied to the symbols received at node $v$, yields the original message $\mathbf{m}$.

In a *linear* network code, each of the functions described above is a linear function. Thus a linear solution to the network coding problem is given by a list of vectors $(\mathbf{v}[e])_{e \in E_G}$ describing which linear combination of the original $n$ messages is transmitted on each edge (the symbol $\mathbf{v}[e] \cdot \mathbf{m}$ is carried on edge $e$). Condition (i) above implies that for all edges $(u, v)$, where $u \neq s_G$, the symbol $\mathbf{v}[u, v] \cdot \mathbf{m}$ may be computed as a linear combination of the symbols $\mathbf{v}[w, u] \cdot \mathbf{m}$ carried on edges $(w, u)$ into node $u$.

It is now well-known that given an instance of the multicast linear network coding problem with $q \geq |T_G|$, a solution exists if and only if the minimum cut between the source and each sink is of size at least $n$ [1, 23]; moreover, efficient algorithms are known for constructing feasible solutions [18].

**4.2 Security Against a Wire-Tap Adversary.** A computationally unbounded "wire-tap" adversary against a network code has access to the symbols which are transmitted over some unknown set of at most $k$ edges of the network. Additionally, the adversary has full knowledge of the network code itself and of whatever protocol we use for security. We would like to transmit some information $x \in \mathbb{F}_q^t$, where $t \leq n$, over the network in a way that is information-theoretically secure against this adversary.

Our approach (introduced by Cai & Yeung [5]) is to use multicast linear network coding as described above, where we choose a random vector $r \in \mathbb{F}_q^\ell$ at the source node $s_G$, and let the message $\mathbf{m}$ be some function of $r$ and the information $x$. Then, a formal security requirement can be defined as follows. If we regard the length-$|E_G|$ sequence of symbols carried over all the edges in the network as being the output of a function $f(x, r) : \mathbb{F}_q^t \times \mathbb{F}_q^\ell \to \mathbb{F}_q^N$ of the information $x$ and the random $r$ chosen at the source, we would like this function to be secure against a $k$-threshold adversary.

Note that we have weakened our goal of sending $n$ symbols to sending $t \leq n$ symbols for the purposes of achieving security. The *capacity* of the secure network code is $t$, the dimension of the information vector $x$. Our main goal in this work is to balance the capacity $t$ against the adversarial threshold $k$, and the required edge bandwidth $\log q$.

**4.3 From Filtered Secret Sharing to Secure Network Coding** We can reduce the problem of constructing a secure linear network code with min-cut $n$, capacity $t$, field size $q$ and threshold $k$ to the filtered secret sharing problem, as follows.[2] Given an instance $(G = (V, E), s_G, T_G, n, q)$ of the multicast linear network coding problem, we first solve the original network coding problem (ignoring the security condition; i.e.,

[2]This reduction is implicit in the work of Cai & Yeung [5], although they do not explicitly give a secret sharing abstraction. They also suggest altering the network code itself by applying a linear transformation to the coding vectors, rather than the input as we suggest here. The two methods can be shown to be equivalent.

$k = 0$) using the known polynomial-time algorithm. Now let $N = |E|$, and let $V$ be an $n$-by-$N$ matrix where the columns of $V$ are the vectors $(\mathbf{v}^T[e])_{e \in E}$ from the solution to the network coding problem. Solve the filtered secret sharing problem on $V$ (using the parameters $n$, $q$, $k$) to obtain a function $S : \mathbb{F}_q^t \times \mathbb{F}_q^\ell \to \mathbb{F}_q^n$. Use this function to encode $(x, r)$ at the source, where $x \in \mathbb{F}_q^t$ is the information, and $r \in \mathbb{F}_q^\ell$ is chosen randomly. Clearly this will satisfy the $k$-threshold security condition on the network code, by the guaranteed properties of $S$. For feasibility, note that the vectors $v[e]$ entering each terminal span $\mathbb{F}_q^n$, by the feasibility of the original network code. Hence, each terminal can solve a system of equations to obtain $S(x, r)$, and use the recoverability condition on $S$ to obtain $x$.

We note that this reduction also has the advantage of being able to make an existing network code secure: if a network code has already been built (in hardware, say), then one can make it secure (at the expense of some capacity, of course) by sending $S(x, r)$ through the network instead of $x$.

**4.4 Our Main Secure Network Coding Results** Combining the reduction from Section 4.3 with the results of Section 3.3 gives our main results on secure network coding.

The following positive result, which follows from Theorem 1, shows that we can achieve very low bandwidth in secure network coding at the expense of a small loss of capacity:

**Theorem 3** *Let $G$ be a graph with $N$ edges, a source node $s_G$, and a set $T_G$ of sink nodes such that the minimum-cut between the source and any sink is at least $n$. Let $k$ be a threshold where $k < n$. Then for any $\sigma > 1$, as long as the desired bandwidth $\log q$ satisfies $q \geq \max\{N^{\Omega(\frac{1}{\sigma-1})}, |T_G|\}$, there exists a feasible linear network coding scheme with capacity $t = n - \sigma k$ and bandwidth $\log q$ which is secure against a $k$-threshold wire-tap adversary. (And there is a randomized $\mathrm{poly}(N)$-time algorithm that, given ($G$,$s_G$,$T_G$,$k$), outputs such a scheme with high probability.)*

This result enables us to trade off a capacity of $n - \sigma k$ against a field size requirement of $N^{\Omega(\frac{1}{\sigma-1})}$. (The requirement $q \geq |T_G|$, which also exists in Cai & Yeung [5], comes from the algorithm to construct the network code [18].) Note that $\sigma$ need not be close to 1; for example taking $\sigma = 3$ allows a field size as low as (roughly) $\sqrt{N}$ at a capacity of $n - 3k$.

For a negative result, if we consider Theorem 2, we see that there are instances of filtered secret sharing when $t = n - k$ and $q \leq N^{O(\sqrt{k/\log k})}$ that have no linear solutions. This implies that if we demand capacity $t = n - k$ with field size $q \leq N^{O(\sqrt{k/\log k})}$, then we cannot achieve security with this method. It is reasonable to conjecture in such a case that either there is no secure solution, or we need to take network topology into consideration to construct our secure code from scratch (rather than alter a given network code), perhaps using methods along the lines of Jain [19].

## 5 Necessary and Sufficient Conditions for Filtered Secret Sharing

The remainder of the paper is devoted to proving Theorems 1 and 2. In this section we derive necessary and sufficient independence conditions for linear functions $S(x, r)$ to be secure solutions to the filtered secret sharing problem. One direction (sufficiency) of the main theorem in this section (Theorem 6) is implicit in work of Cai & Yeung [5]. We offer a simpler presentation, a small generalization, and a proof that the condition given is also necessary. The proofs of the theorems in this section can be found in Appendix B.

We assume that we are solving an instance $(q, n, N, V, k)$ of filtered secret sharing where in the solution $(\ell, t, S)$, the function $S : \mathbb{F}_q^t \times \mathbb{F}_q^\ell \to \mathbb{F}_q^n$ must be linear. Since $S$ is linear we can write it as $S(x, r) = (x, r)T$ with $T$ a $(t+\ell)$-by-$n$ matrix, so we use the notation $(\ell, t, T)$ to refer to a linear filtered secret sharing solution.

**Lemma 4** *Without loss of generality, the rows of $T$ are linearly independent.*

This lemma also implies that wlog, we have $\ell \leq n - t$. In fact by the following we may assume $\ell = n - t$:

**Lemma 5** *If there is a solution $(\ell, t, T)$ with $\ell < n - t$, then there is a solution with $\ell = n - t$.*

We henceforth assume $T$ is an invertible $n$-by-$n$ matrix. Letting $M = T^{-1}$, the main theorem of this section is:

**Theorem 6** *The function $S(x, r) = (x, r)T = (x, r)M^{-1}$ gives a secure solution $(\ell, t = n - \ell, S)$ to the filtered secret sharing problem $(q, n, N, V, k)$ if and only if any set consisting of*
*(a) at most $k$ linearly independent columns of $V$, and*
*(b) any number of vectors from the first $n - \ell$ columns of $M$*
*is linearly independent.*

**5.1 The Existence of a Secure Matrix.** We say that a matrix $M$ which meets the conditions of Theorem 6 is *secure*. Implicit in the work of Cai & Yeung [5] is a proof that a secure matrix $M$ exists with $\ell = k$, as long as the field size $q$ satisfies $q > \binom{|E_G|}{k}$. In the application to network coding, having an alphabet of this size may well be a prohibitive bandwidth requirement for certain networks. Moreover, the algorithm they give in their proof for finding a secure matrix $M$ takes at least $\binom{|E_G|}{k}$ time steps.

## 6 Equivalence to a Coding Problem

In this section we show that finding a secure matrix $M$ meeting the independence conditions of Theorem 6 is equivalent to finding a linear error-correcting code with certain generalized distance properties. Roughly speaking, the code we are looking for must have all its codewords far away from any word in a linear subspace defined by the matrix $V$. In Sections 7 and 8 we will use this equivalence to establish upper and lower bounds on the field size required to find secure solutions to filtered secret sharing.

**6.1 Preliminaries.** In an instance of filtered secret sharing, the $n$-by-$N$ matrix $V$ is presumed to be full rank. Let $A$ be an $(N - n) \times N$ generator matrix for the null space of $V$. (Equivalently, $A$ is the parity-check matrix if $V$ is regarded as a generator for a code.)

We will henceforth use the notation $M$ to mean the first $n - \ell$ columns of the invertible square matrix "$M$" in Theorem 6. With this new notation, a matrix $M$ is secure[3] iff

$$Mx^T + Vw^T \neq 0 \text{ for all } x \in \mathbb{F}_q^{n-\ell}, w \in \mathbb{F}_q^N \text{ s.t. } x \neq 0, |w| \leq k. \tag{1}$$

We define a notion of "distance" between two matrices that is (roughly) the minimum distance between two vectors in the span of their rows. More precisely, for an $\alpha \times n$ matrix $P$ and a $\beta \times n$ matrix $Q$, we define

$$\delta(P, Q) \equiv \min_{x \in \mathbb{F}_q^\alpha, y \in \mathbb{F}_q^\beta, y \neq 0} \Delta(xP, yQ),$$

where $\Delta$ is the Hamming distance. Note the slight asymmetry in the treatment of $P$ and $Q$, namely that $x$ can be $0^\alpha$ but $y$ cannot be $0^\beta$; this makes the minimum distance of the code generated by $Q$ an upper bound on $\delta(P, Q)$.

**6.2 Main Theorem.** Now we present our main theorem relating the above notion of distance to the existence of a secure matrix $M$:

**Theorem 7** *There exists a secure $n \times (n - \ell)$ matrix $M$ if and only if there exists an $(t = n - \ell) \times N$ matrix $B$ with $\delta(A, B) > k$, where $A$ is the generator matrix for the null space of $V$.*

[3] Since the security of "$M$" (as in Theorem 6) depends only on its first $n - \ell$ columns, we may extend a matrix $M$ from (1) to be square and invertible using an arbitrary extension of $M$ to a basis, as suggested in [5].

**Proof:** Suppose there is some $n - \ell \times N$ matrix $B$ with $\delta(A, B) > k$. Let $M = VB^T$. Note that $B$ must have rank $n - \ell$, since otherwise it could not have $\delta(A, B) > 0$.

Because $\delta(A, B) > k$, and $A$ generates the null space of $V$, we have that $\Delta(y^T, B^T x^T) > k$ for all $x \in \mathbb{F}_q^{n-\ell}$, $x \neq 0$ and $y : Vy^T = 0$. Therefore, $V(B^T x^T + w^T) \neq 0$ for all $x \in \mathbb{F}_q^{n-\ell}$, $x \neq 0$ and $w \in \mathbb{F}_q^N$ where $|w| \leq k$. This implies $VB^T x^T + Vw^T = Mx^T + Vw^T \neq 0$ for all such $x, w$, which are exactly the security conditions in (1).

For the other direction, if we suppose there is a secure $M$, we construct $B$ as follows. For each column $M_i$ of $M$, let the $i$th column of $B^T$ be an arbitrary member of the coset $\{y \in \mathbb{F}_q^N : Vy = M_i\}$. Note that we again have $M = VB^T$.

Since $M$ is secure, we have $Mx + Vw \neq 0$ for all $x \neq 0$, $|w| \leq k$ (from (1)), and so $V(B^T x + w) \neq 0$ for all such $x, w$. Thus for all $y : Vy = 0$, and $x \neq 0$, we have $\Delta(B^T x, y) > k$. This implies $\delta(A, B) > k$. ■

**6.3 A Generalized Coding Problem.** Having proved Theorem 7, we now turn to the following problem:

> **Span Distance Problem:** Given an $\alpha$-by-$N$ matrix $A$ with rank $\alpha$, whose entries belong to $\mathbb{F}_q$, find a $\beta$-by-$N$ matrix $B$ over $\mathbb{F}_q$ such that $\delta(A, B) > k$.

We can regard this question as a generalization of the classical code construction problem: if $\{xB\}_{x \in \mathbb{F}_q^\beta}$ is regarded as a code, then every non-zero codeword must have good distance not only from the all-zeros codeword, but also from every other word generated by $A$.

In the following sections, we consider the span distance problem abstractly. When we apply this problem to Theorem 7, we have $\alpha = N - n$ and $\beta = n - \ell$. Setting $\sigma \geq 1$ such that $\ell = \sigma k$, we are now interested in the case of the span distance problem where $k = \frac{N - \alpha - \beta}{\sigma}$. In the application to network coding, the value $k(\sigma - 1)$ measures the amount of capacity we are willing to give up in order to reduce the field size necessary to achieve security.

## 7 A Positive Result: giving up information length to save on field size

The main theorem in this section gives a bound on the probability that a random code will solve the relevant case of the span distance problem:

**Theorem 8** *Let $A$ be an arbitrary $\alpha$-by-$N$ matrix with rank $\alpha$ over $\mathbb{F}_q$, and let $B$ be a random $\beta$-by-$N$ matrix over $\mathbb{F}_q$. Let $k, \sigma$ be such that $k = \frac{N-\alpha-\beta}{\sigma}$, where $\sigma \geq 1$. Then we have $\delta(A, B) > k$ with probability at least $1 - P_{BAD}$, where $P_{BAD} = q^{-\sigma k}\mathrm{Vol}_q(k, N)$.*

**Proof:** The argument follows along the same lines as the classical argument that random linear codes meet the Gilbert-Varshamov bound (see [17]). Let BAD be the set of words in $\mathbb{F}_q^N$ with distance at most $k$ from some linear combination of the rows of $A$. Using the bound $|\mathrm{BAD}| \leq q^\alpha \mathrm{Vol}_q(k, N)$, we have that for a particular $x_1 \in \mathbb{F}_q^\beta$, the probability (over choices of $B$) that $x_1 B \in \mathrm{BAD}$ is at most $\frac{q^\alpha \mathrm{Vol}_q(k,N)}{q^N}$. Applying a union bound over $x_1 \in \mathbb{F}_q^\beta$, we have the probability of some $x_1 B$ being in BAD is at most $P_{BAD} \leq q^{\alpha+\beta-N}\mathrm{Vol}_q(k, N) = q^{-\sigma k}\mathrm{Vol}_q(k, N)$. ■

**7.1 Applying Theorem 8.** Here we show that Theorem 8 allows us to use fields of quite modest size and still achieve a good probability bound in Theorem 8. It is easy to see that $\mathrm{Vol}_q(k, N) = \sum_{i=0}^{k}(q-1)^i\binom{N}{i}$. We consider two different ranges of values for $k$ (these are $k = o(N)$ and $k = \Theta(N)$) and use different upper bounds on $\mathrm{Vol}_q(k, N)$ in these two cases. We use the following facts in the bounds.

**Fact 9** *[17] For any $q \geq 2$, if $0 < k < (1 - 1/q)N$, then*

*(a) the largest term in the sum $\sum_{i=0}^{k}(q-1)^i\binom{N}{i}$ is the $i = k$ term,*

*(b) $\frac{\log \mathrm{Vol}_q(k,N)}{\log q} = (H_q(k/N) \pm o(1))N$, where $H_q(\delta) = \delta \log_q(q-1) - \delta \log_q \delta - (1-\delta)\log_q(1-\delta)$.*

We first consider the case $k = o(N)$. In this case, from Fact 9(a), it follows that $\mathrm{Vol}_q(k,N) \leq (k+1)q^k N^k$, and so we have that the probability $1 - P_{BAD}$ is positive as long as $q > (k+1)^{\frac{1}{(\sigma-1)k}} \cdot N^{\frac{1}{\sigma-1}}$. To obtain a high probability result such as $P_{BAD} < N^{-c}$ for some constant $c > 0$, it suffices to take $q > (k+1)^{\frac{1}{(\sigma-1)k}} \cdot N^{\frac{1+c/k}{\sigma-1}}$. This establishes Theorem 1.

Our lower bounds on $q$ are easily seen to be much less restrictive than the $q > \binom{N}{k}$ lower bound of Cai & Yeung, at the cost of only a small loss in capacity (number of information symbols we can transmit). As one example, if we take $\sigma = 2$ then we achieve capacity $n - 2k$ (as opposed to Cai & Yeung's $n - k$), but we require only that the field size $q$ be (roughly) at least $N^{1+c/k}$ which is close to $N$ for moderate $k$ and small constant $c$. (Of course, even smaller lower bounds on $q$ can be achieved by taking $\sigma > 2$.) Thus, if $k$ satisfies both $k = \omega(1)$ and $k = o(N)$, we lose only a $(1 - o(1))$ factor in capacity while obtaining a superpolynomial savings in field size.

We now consider the case $k = \Theta(N)$, and show that here we can achieve even more dramatic savings in field size. Taking $k = \delta N$ where $\delta = \Theta(1)$ is some constant and plugging Fact 9(b) into Theorem 8, we have that $P_{BAD} \leq q^{-\sigma\delta N} q^{N(H_q(\delta)+o(1))}$. It is easy to see that in general, $H_q(\delta) < \delta + \frac{1}{\log q}$, and thus (b) above implies that $P_{BAD} < q^{(-\delta(\sigma-1)/2+o(1))N}$ provided that $\delta(\sigma-1)/2 > 1/\log q$, i.e. $q > 2^{2/(\delta(\sigma-1))}$. Since $\delta$ is a fixed constant independent of $N$, the field size lower bound is $2^{\Omega(\frac{1}{\sigma-1})}$ which is independent of $N$.

## 8 A Negative Result: Maximum information length can require large field size

The main result in this section is the following theorem:

**Theorem 10** *Let $\alpha = N - \frac{\log N}{\log q} - \frac{\log \mathrm{Vol}_q(k,N)}{\log q} + 2\log N + \log q + \log\ln q$. If $\alpha, \beta$ satisfy*

$$k + \beta < N - \alpha = \frac{\log N}{\log q} + \frac{\log \mathrm{Vol}_q(k,N)}{\log q} - 2\log N - \log q - \log\ln q \tag{2}$$

*then there is an $\alpha \times N$ matrix $A$ over $\mathbb{F}_q$ such that there is no $\beta \times N$ matrix $B$ over $\mathbb{F}_q$ for which $\delta(A,B) > k$.*

In words, this theorem says that for certain values of $\alpha$ and $\beta$, if $q$ is too small then there exists an $\alpha \times N$ matrix $A$ over $\mathbb{F}_q$ for which the span distance problem cannot be solved if we take $k = N - \alpha - \beta$. This translates into the existence of instances of filtered secret sharing that are unsolvable with linear functions $S$. In the network coding application, taking $k = N - \alpha - \beta$ corresponds to taking $\sigma = 1$; this means that if we do not give up some capacity then there does not exist a secure matrix $M$ unless the field size $q$ is quite large.

**8.1 Establishing Theorem 10: using a code with good covering radius.** To establish Theorem 10, we need to find a full-rank $\alpha$-by-$N$ matrix $A$ which is such that for all full-rank $\beta$-by-$N$ matrices $B$, there is a point $x_1 B$, $x_1 \neq 0$ and a point $x_2 A$ where $\Delta(x_1 B, x_2 A) \leq k = N - \alpha - \beta$.

For the case $\beta = 1$, this is exactly a question of constructing a code $A$ with small *covering radius*. The covering radius [8] of a code is the minimum value $d$ such that the union of the spheres of radius $d$ around the points in the code cover the entire space $\mathbb{F}_q^n$. Suppose $A$ had covering radius of at most $N - \alpha - \beta$. Then, no matter what $B$ is ($B$ is a single vector, since $\beta = 1$), it has distance at most $N - \alpha - \beta$ to some point $x_2 A$. Now suppose $A$ has covering radius $d > N - \alpha - \beta$. Then there is some vector $B$ where $\Delta(B, x_2 A) > N - \alpha - \beta$ for all $x_2$. Moreover, any scalar multiple of $B$ will also have distance at least

$d$ from any $x_2 A$ (to see this, note that if $\Delta(aB, x_2 A) < d$, then $\Delta((1/a)aB, (1/a)(x_2 A)) < d$, and so $\Delta(B, ((1/a)x_2)A) < d$, a contradiction).

Thus for $\beta = 1$, a construction of $A$ with covering radius of at most $N - \alpha - \beta$ is necessary and sufficient for a negative result. Additionally, for $\beta > 1$, showing that there exists an $\alpha \times N$ matrix $A$ with covering radius at most $N - \alpha - \beta$ is sufficient for a negative result.

Cohen and Frankl [9] gave upper bounds on the covering radius of linear codes over $\mathbb{F}_q$. Their analysis can be used to obtain the following result[4], which then implies Theorem 10:

**Theorem 11** *For any value $1 \leq d \leq N$, there is a D-dimensional linear code over $\mathbb{F}_q$ with block length $N$ (i.e. a vector subspace of $\mathbb{F}_q^N$) which has covering radius at most d, where*

$$D \equiv N - \frac{\log N}{\log q} - \frac{\log \mathrm{Vol}_q(d,N)}{\log q} + 2\log N + \log q + \log \ln q.$$

**8.2 Applying Theorem 10.** We give one example here of how Theorem 10 can be applied, and this example establishes Theorem 2. Other interesting examples are possible, but we omit them for space reasons.

Let $\tau$ be any constant satisfying $0 < \tau < 1/2$. Let $c = N^\tau$, let $k = \gamma c^2 \log N$ (we will specify $\gamma$ shortly) and let $q = N^c$. By Fact 9(a), we can get a fairly good lower bound on $\mathrm{Vol}_q(k,N)$ just by considering the last term of the sum. We have $\mathrm{Vol}_q(k,N) \geq (q-1)^k \binom{N}{k} \geq (q/2)^k \left(\frac{N}{k}\right)^k$. We thus have that $\frac{\log \mathrm{Vol}_q(k,N)}{\log q} \geq k + \frac{-1 + k\log N - k\log k}{\log q}$. Plugging the above parameter settings for $k$ and $q$ into Equation (2) (but not substituting in yet for $c$), we have that Equation (2) is satisfied if

$$\begin{aligned}\gamma c^2 \log N + \beta \quad < \quad & \frac{1}{c} + \gamma c^2 \log N + \frac{-1 + \gamma c^2 \log^2 N - \gamma c^2 \log N \log(\gamma c^2 \log N)}{c \log N} \\ & -2\log N - c\log N - \log(c \ln N).\end{aligned}$$

This inequality is equivalent to

$$\beta < \frac{1}{c} - \frac{1}{c\log N} + \gamma c \log N - \gamma c \log(\gamma c^2 \log N) - (c+2)\log N - \log(c \ln N).$$

Now since $c = N^\tau$, it can be verified that taking $\gamma = \frac{2}{1-2\tau}$ (a fixed constant since $0 < \tau < 1/2$ is a fixed constant) makes the right-hand side of this last inequality at least $(c-3)\log N$ (for sufficiently large $N$), so $\beta$ can be any value smaller than this bound. This example shows that for a wide range of values of $k$ the lower bound on field size required for secure linear filtered secret sharing, if no capacity is given up, can be as large as $N^{\Omega(\sqrt{k/\log k})}$. It is interesting to contrast this lower bound with the upper bound of $\binom{N}{k}$ of Cai & Yeung.

## 9 Future Work

Several interesting directions for future research suggest themselves. Are there other application areas where the notion of filtered secret sharing will be useful? Are there other variants of secret sharing besides filtered secret sharing which are useful for secure network coding?

Within the arena of network coding, there are many promising avenues for future work. The results of Jain [19] are obtained by exploiting the topology of the network in question, while our techniques are independent of the topology. Can stronger results be obtained by combining the two approaches? Another natural question is whether network codes for information transmission problems other than multicast can be made secure using our techniques. Finally, it is also of interest to consider secure network coding in a framework where only statistical security or security against computationally bounded adversaries is required, as opposed to the information-theoretic security criterion of this paper.

[4]The expression for $D$ in Theorem 11 is slightly different from the result stated in [9]. Their analysis implicitly assumes that $q$ is independent of $N$, but this need not hold for us. We give a complete derivation of Theorem 11 in the full version of the paper.

## 10 Acknowledgment

We thank Amos Beimel for useful discussions.

## References

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

[2] M. Ben-Or, S. Goldwasser, and A. Wigderson. Completeness theorems for noncryptographic fault-tolerant distributed computations. In *Proc. of the 20th Annu. ACM Symp. on the Theory of Computing*, pages 1–10, 1988.

[3] G. R. Blakley. Safeguarding cryptographic keys. In R. E. Merwin, J. T. Zanca, and M. Smith, editors, *Proc. of the 1979 AFIPS National Computer Conference*, volume 48 of *AFIPS Conference proceedings*, pages 313–317. AFIPS Press, 1979.

[4] G. R. Blakley and C. Meadows. The security of ramp schemes. In G. R. Blakley and D. Chaum, editors, *Advances in Cryptology – CRYPTO '84*, volume 196 of *LNCS*, pages 242–268. Springer-Verlag, 1985.

[5] N. Cai and R. W. Yeung. Secure network coding. In *International Symposium on Information Theory (ISIT '02)*, June 2002.

[6] D. Chaum, C. Crépeau, and I. Damgård. Multiparty unconditionally secure protocols. In *Proc. of the 20th Annu. ACM Symp. on the Theory of Computing*, pages 11–19, 1988.

[7] P. A. Chou, Y. Wu, and K. Jain. Practical network coding. In *Proc. 41st Annual Allerton Conference on Communication, Control, and Computing*, October 2003.

[8] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein. *Covering Codes*. Elsevier, Amsterdam, 1997.

[9] G. D. Cohen and P. Frankl. Good coverings of hamming spaces with spheres. *Discrete Mathematics*, 56:125–131, 1985.

[10] R. Cramer, I. Damgård, and U. Maurer. General secure multi-party computation from any linear secret-sharing scheme. In B. Preneel, editor, *Advances in Cryptology – EUROCRYPT 2000*, volume 1807 of *LNCS*, pages 316–334. Springer, 2000.

[11] J. Feldman, T. Malkin, R. A. Servedio, and C. Stein. On the capacity of secure network coding. In *Proc. 42nd Annual Allerton Conference on Communication, Control, and Computing*, 2004.

[12] C. Fragouli and E. Soljanin. Information flow decomposition for network coding. Submitted to IEEE Transactions on Information Theory, June 2004.

[13] C. Gkantsidis and P. R. Rodriguez. Network coding for large scale content distribution. Technical Report MSR-TR-2004-80, Microsoft Research, 2004.

[14] T. Ho, D. Karger, M. Médard, and R. Koetter. Network coding from a network flow perspective. In *International Symposium on Information Theory (ISIT '03)*, June 2003.

[15] T. Ho, R. Koetter, M. Médard, D. Karger, and M. Effros. The benefits of coding over routing in a randomized setting. In *IEEE International Symposium on Information Theory (ISIT)*, June 2003.

[16] T. Ho, B. Leong, R. Koetter, M. Médard, M. Effros, and D. R. Karger. Byzantine modification detection in multicast networks using randomized network coding. In *IEEE International Symposium on Information Theory (ISIT 2004)*, June 2004.

[17] W. C. Huffman and V. Pless. *Fundamentals of Error Correcting Codes*. Cambridge University Press, 2003.

[18] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen. Polynomial time algorithms for multicast network code construction. Submitted to IEEE Transactions on Information Theory, July 2003.

[19] K. Jain. Security based on network topology against the wiretapping attack. *IEEE Wireless Communications*, pages 68–71, February 2004.

[20] R. Koetter and M. Médard. An algebraic approach to network coding. *Transactions on Networking*, October 2003.

[21] Ralf Koetter. Network coding home page, `http://tesla.csl.uiuc.edu/˜koetter/nwc`, 2004.

[22] A. R. Lehman and E. Lehman. Complexity classification of network information flow problemss. In *Symposium on Discrete Algorithms (SODA '04)*, January 2004.

[23] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear network coding. *IEEE Transactions on Information Theory*, 49:371–381, Februray 2003.

[24] D. S. Lun, M. Médard, and M. Effros. On coding for reliable communication over packet networks. In *Proceedings 42nd Annual Allerton Conference on Communication, Control, and Computing*, October 2004.

[25] Adi Shamir. How to share a secret. *Communications of the ACM*, 22:612–613, 1979.

[26] G. J. Simmons. An introduction to shared secret and/or shared control and their application. In G. J. Simmons, editor, *Contemporary Cryptology, The Science of Information Integrity*, pages 441–497. IEEE Press, 1992.

[27] D. R. Stinson. An explication of secret sharing schemes. *Designs, Codes and Cryptography*, 2:357–390, 1992.

[28] Douglas Stinson and Ruizhong Wei. Bibliography on secret sharing schemes, 1998, `http://www.cacr.math.uwaterloo.ca/˜dstinson/ssbib.html`.

[29] Y. Wu, P. A. Chou, and S.-Y. Kung. Information exchange in wireless networks with network coding and physical-layer broadcast. Technical Report MSR-TR-2004-78, Microsoft Research, 2004.

# Appendix

## A A lower bound on information length

**Theorem 12** *Any instance of filtered secret sharing requires $t \leq n - k$.*

**Proof:** (Sketch) By the fact that the filter matrix $V$ is full rank, there must be a full-rank $n$-by-$n$ submatrix $V'$ of $V$. Let $V''$ be a submatrix of $k$ columns from $V'$, and suppose the adversary has access to $a = S(x,r)V''$. The set of vectors $A = \{y \in \mathbb{F}_q^n : yV'' = a\}$ that could have been the secret vector $S(x,r)$ has cardinality exactly $q^{n-k}$, since $V''$ has full rank.

By recoverability, we have some function $I : \mathbb{F}_q^n \to \mathbb{F}_q^t$ that maps secret vectors $S(x,r) \in \mathbb{F}_q^n$ to information vectors $x \in \mathbb{F}_q^t$. For this function $S$ to be secure against the adversary just described, it must be the case that for all $x \in \mathbb{F}_q^t$, we have $I(y) = x$ for some $y \in A$. It follows that $q^t \leq q^{n-k}$, since $|A|$ = $q^{n-k}$. ■

## B Proofs for Section 5

**Lemma 4** *Without loss of generality, the rows of $T$ are linearly independent.*

**Proof:** If $T$ has a set of linearly dependent rows that includes a row among the first $t$, then $(x,r)M = (x',r')M$ for some $r, r', x, x'$ s.t. $x \neq x'$. This violates the unique recoverability condition on $S$. Otherwise, if $T$ has a set of linearly dependent rows all of which are among the last $\ell$, then removing one of those rows and setting $\ell' := \ell - 1$ does not change the distribution of $S(x,r)$ over random $r$ for any $x$. ■

**Lemma 5** *If there is a solution $(\ell, t, T)$ with $\ell < n - t$, then there is a solution with $\ell = n - t$.*

**Proof:** Take the solution $(\ell, t, xT)$ (wlog, $T$ has l.i. rows) and add $n - \ell - t$ l.i. rows to the end of $T$ to obtain an invertible square matrix $T'$. Setting $\ell' = n - t$, we have a new solution that still satisfies the security condition. For recoverability, note that since $T'$ is invertible, the full vector $(x,r)$ can be obtained from a given vector $(x,r)T'$. ■

**Theorem 6** *The function $S(x,r) = (x,r)T = (x,r)M^{-1}$ gives a secure solution $(\ell, t = n - \ell, S)$ to the filtered secret sharing problem $(q, n, N, V, k)$ if and only if any set consisting of*

*(a) at most $k$ linearly independent columns of $V$, and*

*(b) any number of vectors from the first $n - \ell$ columns of $M$*

*is linearly independent.*

**Proof:** We first show that the encoding is secure if the independence condition is met. Suppose that the adversary has access to $S_I(x,r)$, where $I \subseteq [N]$ is arbitrary and $k' = |I| \leq k$. Let $\overline{V}$ be a $n$-by-$k'$ submatrix of $V$ consisting of the columns indexed by $I$. We may assume that $\overline{V}$ has rank $k'$, since otherwise the adversary could drop an index and not lose any information. Note that what the adversary sees is the length-$k'$ vector $\mathbf{a} = S_I(x,r) = (x,r)M^{-1}\overline{V}$.

For a particular "guess" at the information $\hat{x}$, and the observed vector $\mathbf{a}$, let $R(\hat{x}, \mathbf{a})$ be the set of all possible random vectors $\hat{r}$ such that $\mathbf{a} = S_I(\hat{x}, \hat{r})$. It is easily shown that the function $S$ is secure if for all $x', x'' \in \mathbb{F}_q^{n-\ell}$, we have $|R(x', \mathbf{a})| = |R(x'', \mathbf{a})|$.

Given an information vector $\hat{x} \in \mathbb{F}_q^{t=n-\ell}$, the set $R(\hat{x}, \mathbf{a})$ has one member for every solution to the system of $n - \ell + k'$ equations in $n$ unknowns described by $\hat{\mathbf{y}}\overline{V}' = (\hat{x}, \mathbf{a})$, where $\hat{\mathbf{y}} \in \mathbb{F}_q^n$ is unknown, and $\overline{V}'$ is the $n$ by $(n - \ell + k')$ coefficient matrix $\overline{V}' = \left[\begin{array}{c|c} I_{n-\ell} & M^{-1}\overline{V} \\ 0 & \end{array}\right]$. Now suppose that $\overline{V}'$ has full rank; then, for all $(\mathbf{a}, \hat{x})$ pairs on the right hand side, the system of equations has exactly the same number of solutions. It follows that $R(x', \mathbf{a}) = R(x'', \mathbf{a})$ for all distinct information vectors $x', x''$, and thus the encoding is secure.

To prove that $\overline{V}'$ has full rank, we consider the matrix $M\overline{V}'$ (recall that $M$ has full rank by definition). This matrix $M\overline{V}'$ has its first $n-\ell$ columns matching the first $n-\ell$ columns of $M$, and the last $k$ columns are the matrix $MM^{-1}\overline{V} = \overline{V}$. Thus, if the conditions on $M$ in the theorem hold, the matrix $M\overline{V}'$ has full rank for all possible choices of $\overline{V}$, and thus the encoding is secure.

For the other direction, suppose the independence condition is not met. This means that there is some nontrivial linear combination of some l.i. set of at $k' \leq k$ columns of $V$ that equals some nontrivial linear combination of the first $t = n - \ell$ (l.i.) columns of $M$. If we define $\overline{V}$ as the submatrix consisting of those $k'$ columns, and $\overline{V}'$ in terms of $\overline{V}$ and $M$ as above, then this is equivalent to saying that $M\overline{V}'$ is not full rank, and thus $\overline{V}'$ is not full rank, since $M$ is full rank by assumption. We may conclude that $\overline{V}'(z_1, z_2)^T = 0$ for some $z_1 \in \mathbb{F}_q^{n-\ell}$, $z_2 \in \mathbb{F}_q^{k'}$. Also, we know that $z_1 \neq 0$ and $z_2 \neq 0$ by looking at the structure of $\overline{V}'$ (using the fact that $I_{n-\ell}$, $\overline{V}$ and $M$ are all full rank). So, we have $\begin{bmatrix} I_{n-\ell} \\ 0 \end{bmatrix} z_1^T + M^{-1}\overline{V} z_2^T = 0$.

Fix some information vector $x \in \mathbb{F}_q^{n-\ell}$ and random vector $r \in \mathbb{F}_q^{\ell}$. Let $\mathbf{a} = (x, r)M^{-1}\overline{V}$ be the vector of observed symbols on the edges $I$. Since $\mathbf{a}$ is the result of a possible choice of $r$, we have that $R(x, \mathbf{a}) > 0$. Note that $\mathbf{a} \cdot z_2 = (x, r)M^{-1}\overline{V} z_2^T = -(x, r) \begin{bmatrix} I_{n-\ell} \\ 0 \end{bmatrix} z_1^T = -x \cdot z_1$. In fact, the relation $\mathbf{a} \cdot z_2 = -x \cdot z_1$ holds for any pair of possible information vectors $x$ and observed vectors $\mathbf{a}$ which satisfy $R(x, \mathbf{a}) > 0$.

Let $x' = x + e_i$ where $i$ is an index such that $(z_1)_i \neq 0$. Thus we have that $x' \cdot z_1 \neq x \cdot z_1 = -\mathbf{a} \cdot z_2$. We conclude that $\mathbf{a}$ is not a possible observed vector for $x'$, and thus $R(x', \mathbf{a}) = 0$. We have demonstrated an $\mathbf{a}, x, x'$ where $R(x, \mathbf{a}) \neq R(x', a)$. This implies that the distributions given $x$ and $x'$ are different, and so the function $S$ is not secure. ■