

Error Correction Capability of Random Network Error Correction Codes

Huseyin Balli, Xijin Yan and Zhen Zhang
 Dept. of Electrical Engineering - Systems
 University of Southern California
 Los Angeles, CA, U.S.A.
 {balli, xyan, zzhang}@usc.edu

Abstract—In this paper, we study the error correction capability of random linear network error correction codes [7]. We derive bounds on the probability mass function of the minimum distance of a random network error correction code and the field size required for the existence of a network error correction code with a given degradation, which is the difference between the highest possible minimum distance in the Singleton bound and the minimum distance of the code. The main tool that we use to study these problems is an improved bound on the failure probability of random linear network codes that at one or more sinks, the source messages are not decodable. This problem was originally studied in [6].

I. BASIC DEFINITIONS AND STATEMENTS OF MAIN RESULTS

Consider a communication network modeled by a finite acyclic directed graph $\{V, E\}$ where V is the set of nodes and E is the set of communication channels of the network, respectively. Each directed edge $e = (i, j) \in E$ represents a channel leading from the node i to the node j . We call node i the tail of e and node j the head of e . Furthermore, the channel e is called an outgoing channel of node i and an incoming channel of node j . For a node i , define $Out(i) = \{e \in E : e \text{ is an outgoing channel of } i\}$ and $In(i) = \{e \in E : e \text{ is an incoming channel of } i\}$. In this paper, we allow multiple channels between two nodes and assume that all channels have unit capacity. Let $\{s\}$ and T be two disjoint subsets of V . Node s is called the source node, and the elements of T are called sink nodes. Other nodes in $J = V - \{s\} - T$ are called internal nodes. Let \mathcal{F} be a finite field. We define source messages to be ω random variables $\underline{X} = (X_i : i = 1, \dots, \omega)$ where $X_i \in \mathcal{F}, \forall i$. They are transmitted to the source node s through ω imaginary channels in the set $In(s)$ and are required to be decoded at all sink nodes. At each node $i \in V - T$, there is a local kernel matrix $K_i = (k_{de} : d \in In(i), e \in Out(i))$ where $k_{de} \in \mathcal{F}$. K_i defines the local coding operations at the node i . Denote the message transmitted over the i -th imaginary channel d_i by $U_{d_i} = X_i$, then the message transmitted over a channel e , denoted by U_e , is calculated by the following formula inductively

$$U_e = \sum_{d \in In(i)} k_{de} U_d.$$

We define the global kernel f_e for a channel e as a \mathcal{F} -valued ω -dimensional vector such that $U_e = \underline{X} f_e$. Thus it can be

calculated by

$$f_e = \sum_{d \in In(i)} k_{de} f_d.$$

Under this formulation, the message received at each sink $t \in T$ includes a vector $A_t = (U_d : d \in In(t))$ and a matrix $F_t = (f_d : d \in In(t))$ for which $\underline{X} F_t = A_t$. This equation is called the decoding equation at sink t when channels are error-free. In this error-free case, the source messages are decodable at t if and only if the rank of F_t is ω .

Since this work is a continuation of [7], we reproduce some basic concepts for network error correction codes from [7]. Suppose there is an error in channel e , then the output of the channel can be modified such that $\tilde{U}_e = U_e + Z_e$, where the error $Z_e \in \mathcal{F}$ is treated as a message received through channel e . Thus, for each channel e , we introduce an imaginary channel e' connected to the tail of e to provide error messages. A linear network code for the original network can be amended to a code with these added imaginary channels by letting $k_{e'e} = 1$ and $k_{e'd} = 0$ for all other channels. The global kernels $\tilde{f}_e : e \in E$ for this network are of dimension $\omega + |E|$ and are called the extended global kernel for the original network. Let $\underline{Z} = \{Z_e : e \in E\}$ be the error message vector. An error pattern is a set of channels in the original network in which channel errors occur. For an error pattern ρ , we have $Z_e = 0$ for $e \notin \rho$. The message transmitted on each channel e can be represented by $\tilde{U}_e = (\underline{X}, \underline{Z}) \tilde{f}_e$. The matrix $\tilde{F}_t = (\tilde{f}_e : e \in In(t))$ is called the decoding matrix at sink t . Let the row vectors of \tilde{F}_t be $row_t(d) : d \in In(s) \cup E$ of dimension $|In(t)|$. Let $\langle L \rangle$ stand for the subspace spanned by the vectors in a collection of vectors L . We define two linear spaces

$$\Delta(t, \rho) = \langle \{row_t(e) : e \in \rho\} \rangle,$$

$$\Psi(t) = \langle \{row_t(e) : e \in In(s)\} \rangle.$$

We call $\Delta(t, \rho)$ the error space of error pattern ρ and $\Phi(t)$ the message space.

Definition 1: We say that an error pattern ρ_1 is dominated by another error pattern ρ_2 with respect to a sink t , denoted by $\rho_1 \prec_t \rho_2$, if $\Delta(t, \rho_1) \subseteq \Delta(t, \rho_2)$ for any linear code.

We use $|\rho|$ to denote the number of erroneous channels in an error pattern ρ .

Definition 2: The rank of an error pattern with respect to a sink t is defined as

$$\text{rank}_t(\rho) = \min\{|\tilde{\rho}| : \rho \prec_t \tilde{\rho}\}.$$

Definition 3: A code is called regular if $\dim(\Phi(t)) = \omega$. For regular codes, the minimum distance at sink node t is defined by

$$d_{min}^t = \min\{\text{rank}_t(\rho) : \Phi(t) \cap \Delta(t, \rho) \neq \{\phi\}\}.$$

In [7], the error correction capability of a network error correction code for several kinds of errors is characterized in terms of the minimum distance of the code. This concept plays exactly the same role as it plays in classical coding theory.

A random network code is a collection of random local kernel values k_{de} for pairs d, e for which $\text{tail}(e) = \text{head}(d)$. They are independently, uniformly distributed random variables taking values in the base field \mathcal{F} . All other parameters of a random network code are functions of the random local kernel values. This includes the minimum distance of the code at each sink node $t \in T$. Let this random variable be denoted by D_{min}^t . In this paper, we are interested in the probability mass function of D_{min}^t . Clearly, D_{min}^t takes values in $\{0, \dots, \delta_t + 1\}$ by the Singleton bound. We have the following result.

Theorem 1: For any single source multicast over an acyclic network $G = \{V, E\}$, let the minimum cut capacity for the sink node $t \in T$ be C_t , the source information rate be ω symbols per unit time, and the redundancy of the code be $\delta_t = C_t - \omega$. For a given $d \geq 0$, called the degradation of a code, the random network code defined above satisfies:

$$\Pr(D_{min}^t < \delta_t + 1 - d) \leq \frac{\binom{|E|}{\delta_t + 1 - d} \binom{d + |J| + 1}{|J|}}{(|\mathcal{F}| - 1)^{d+1}}. \quad (1)$$

Let E_t be the event such that $D_{min}^t < \delta_t + 1 - d$ at sink $t \in T$. Then the inequality (1) implies:

$$\Pr(\cup_{t \in T} E_t) \leq \frac{\sum_{t \in T} \binom{|E|}{\delta_t + 1 - d} \binom{d + |J| + 1}{|J|}}{(|\mathcal{F}| - 1)^{d+1}},$$

which in turn implies the following corollary.

Corollary 1: If the field size satisfies

$$|\mathcal{F}| > 1 + \left(\sum_{t \in T} \binom{|E|}{\delta_t + 1 - d} \binom{d + |J| + 1}{|J|} \right)^{\frac{1}{d+1}}, \quad (2)$$

then there exists a network code having degradations at most d at all sinks $t \in T$.

In [7], it is proved that if $|\mathcal{F}| > \sum_{t \in T} \binom{|E|}{\delta_t}$, then there exists an MDS code. The above corollary says, the field size required for a code with some degradation is much smaller than that required for the existence of a MDS code.

To prove these results, we need an improved upper bound for the failure probability at a sink node $t \in T$ of a random linear network code. We define the failure probability P_e^t as the probability that the source messages are not decodable at sink t which is equivalently the probability that the rank of the matrix F_t is lower than ω . We use P_e to denote the probability that there exists at least one sink node in T at which decoding

fails. This problem was first considered in [6] where it was proved that

- 1) when there is no redundancy, the failure probability is upper bounded by

$$1 - \left(1 - \frac{|T|}{|\mathcal{F}|}\right)^{|E|}, \quad (3)$$

- 2) when the redundancy is r , the failure probability is upper bounded by

$$1 - \sum_{x=0}^r \binom{r + \omega}{x} \left(1 - \frac{1}{|\mathcal{F}|}\right)^{L(r + \omega - x)} \left(1 - \left(1 - \frac{1}{|\mathcal{F}|}\right)^L\right)^x, \quad (4)$$

where L is the longest length of the source to sink paths.

The following theorems are improvements over their results. Theorem 2: When $\min\{C_t : t \in T\} = \omega$, the failure probabilities for a random network code satisfy

- 1) for each $t \in T$,

$$P_e^t \leq 1 - \left(1 - \frac{1}{|\mathcal{F}| - 1}\right)^{|J|+1}, \quad (5)$$

- 2) for all $t \in T$,

$$P_e \leq 1 - \left(1 - \frac{|T|}{|\mathcal{F}| - 1}\right)^{|J|+1}. \quad (6)$$

Theorem 2 is a special case of the following Theorem 3. We omit its proof for brevity.

Theorem 3: Let the redundancy for a sink node t be $\delta_t = C_t - \omega$. The failure probability P_e^t is upper bounded by

$$P_e^t \leq \frac{\binom{\delta_t + |J| + 1}{|J|}}{(|\mathcal{F}| - 1)^{\delta_t + 1}}. \quad (7)$$

Theorem 3 immediately implies the following corollary.

Corollary 2: The probability that the messages are decodable at all sinks is lower bounded by

$$1 - \sum_{t \in T} \frac{\binom{\delta_t + |J| + 1}{|J|}}{(|\mathcal{F}| - 1)^{\delta_t + 1}}. \quad (8)$$

In a separate work, we have shown that the ratio of the bound in [6] over our bound, when applied to a block code, grows exponentially with block length n . Therefore, our result can be viewed as an improvement of theirs. Furthermore, when our result is applied to network error correction codes, it leads to stronger results.

II. PROOFS OF RESULTS

Let $s = v_0 < v_1 < \dots < v_m$, where $m \leq |J|$, be an upstream to downstream order of nodes of the network. For a fixed sink $t \in T$, there exist C_t channel disjoint paths leading from s to t . Let these paths be denoted by $\mathcal{P}_i : 1 \leq i \leq C_t$, where

$$\mathcal{P}_i = \{e_{i,k} : k = 1, \dots, m_i\},$$

$e_{i,k}$ is the head of $e_{i,k-1}$, s is the tail of $e_{i,1}$, and t is the head of e_{i,m_i} . Let the set of channels in all these paths be denoted by $E_p \subset E$. Define cut $CUT_0 = \{e_{i,1} : 1 \leq i \leq C_t\}$ which consists of the first channels of all C_t paths. Once CUT_k is

defined, CUT_{k+1} is formed from CUT_k by replacing channels in $In(v_{k+1}) \cap CUT_k$ by channels in $Out(v_{k+1}) \cap E_p$. By induction, all cuts CUT_k for $k = 0, \dots, m$ can be defined. We notice that CUT_m is a subset of $In(t)$. We say, a failure occurs in CUT_k if the rank of the matrix $F^{(k)} = (f_e : e \in CUT_k)$ is lower than ω . Let the probability that a failure occurs in CUT_k be denoted by $p_e^{(k)}$. The decoding failure probability at sink t denoted by P_e^t is at most $p_e^{(m)}$. For CUT_k , we partition the cut into two parts:

- 1) the inside part: $CUT_k^i = CUT_k \cap In(v_{k+1})$,
- 2) the outside part: $CUT_k^o = CUT_k - In(v_{k+1})$.

Let $CUT_k^* = CUT_{k+1} \cap Out(v_{k+1})$. We have $|CUT_k^*| = |CUT_k^i|$. In fact, CUT_{k+1} is obtained from CUT_k by replacing CUT_k^i by CUT_k^* .

Lemma 1: The failure probability at cut CUT_k is upper bounded by

$$p_e^{(k)} \leq \frac{\binom{\delta_t + k + 1}{k}}{(|\mathcal{F}| - 1)^{\delta_t + 1}}. \quad (9)$$

Apparently, Theorem 3 is a special case of this lemma. Let

$$r_k = |CUT_k^o| - Rank((f_e : e \in CUT_k^o)),$$

and Γ_k be the event that no failure occurs at CUT_k . Under the condition of Γ_k , we have $r_k \leq \delta_t$. Obviously Γ_{k+1}^c is the event that failure occurs at CUT_{k+1} . We have the next lemma.

Lemma 2:

$$Pr(\Gamma_{k+1}^c | \Gamma_k, r_k = l) \leq \frac{1}{(|\mathcal{F}| - 1)^{\delta_t + 1 - l}}. \quad (10)$$

Proof of Lemma 2. If $Rank((f_e : e \in CUT_k^o)) = \omega$, then $Pr(\Gamma_{k+1}^c | \Gamma_k, r_k = l) = 0$. The bound is valid. Therefore, we consider only the case that $Rank((f_e : e \in CUT_k^o)) < \omega$. Let O be the linear space spanned by $\{f_e : e \in CUT_k^o\}$ and I be the linear space spanned by $\{f_e : e \in In(v_{k+1})\}$. Under the condition Γ_k , $(f_e : e \in CUT_k)$ has rank ω . Therefore, $\{f_e : e \in CUT_k\}$ spans the whole ω dimensional space. This implies that the dimension of I is at least $\omega - Rank((f_e : e \in CUT_k^o))$. The vectors in $\{f_e : e \in CUT_k^o\}$ are independently identically distributed in the linear space I with uniform distribution by the assumption that the local kernels k_{de} are independently and identically distributed in \mathcal{F} with uniform distribution. Thus we have

$$\begin{aligned} h &= dim(I) - dim(I \cap O) \\ &= \omega - Rank((f_e : e \in CUT_k^o)) \\ &> 0, \end{aligned}$$

where $dim(\cdot)$ stands for the dimension of a linear space. The number of channels in CUT_k^* is

$$g = C_t - |CUT_k^o| = \delta_t + h - l.$$

Let $CUT_k^* = \{e_1, \dots, e_g\}$. Let O_i be the linear space spanned by vectors $\{f_{e_j} : 1 \leq j \leq i\}$ and global kernel vectors for channels in CUT_k^o . We consider the sequence

$$Z_i = dim(O_i) - dim(O_{i-1})$$

where $O_0 = O$. Z_i takes values either 0 or 1. The event Γ_{k+1}^c corresponds to the set of sequences $Z = (Z_1, \dots, Z_g)$ of weights at most $h - 1$. Therefore,

$$Pr(\Gamma_{k+1}^c | \Gamma_k, r_k = l) = \sum_{Z \in \{0,1\}^g : wt(Z) \leq h-1} Pr(Z),$$

where $wt(\cdot)$ is the weight of a binary sequence. We have

$$Pr(Z) = \prod_{i=1}^g Pr(Z_i | Z_1, \dots, Z_{i-1}),$$

where

$$Pr(Z_i = 0 | Z_1, \dots, Z_{i-1}) = \frac{1}{|\mathcal{F}|^{h-wt(Z_1, \dots, Z_{i-1})}}, \quad (11)$$

and

$$\begin{aligned} &Pr(Z_i = 1 | Z_1, \dots, Z_{i-1}) \\ &= 1 - \frac{1}{|\mathcal{F}|^{h-wt(Z_1, \dots, Z_{i-1})}} \quad (\leq 1) \\ &= Pr(Z_i = 0 | Z_1, \dots, Z_{i-1}) + Pr(Z_i = 1 | Z_1, \dots, Z_{i-1}). \end{aligned} \quad (12)$$

The equation (11) is proved as follows. The random variable Z_i takes value 0 if and only if f_{e_i} is in O_{i-1} , the linear space spanned by the vectors in

$$\{f_{e_j} : 1 \leq j \leq i-1\} \cup \{f_e : e \in CUT_k^o\}.$$

Then we have

$$dim(I) - dim(I \cap O_{i-1}) = h - wt(Z_1, \dots, Z_{i-1}).$$

Since f_{e_i} is uniformly distributed in I , it falls in $I \cap O_{i-1}$ with probability

$$\frac{|I \cap O_{i-1}|}{|I|} = \frac{1}{|\mathcal{F}|^{h-wt(Z_1, \dots, Z_{i-1})}}.$$

We use a different method to characterize a binary sequence Z . We consider the location of the i -th 0 in the sequence. This can be characterized by the number of 1's before the i -th 0. Let this be $t_Z(i)$. Obviously

$$t_Z = (t_Z(i) : i = 1, \dots, g - wt(Z))$$

is a non-decreasing sequence with maximum entry value $t_Z(i) \leq h - 1$ and length $g - h + 1$. By using this sequence, we proceed as follows:

$$\begin{aligned} &Pr(\Gamma_{k+1}^c | \Gamma_k, r_k = l) \\ &= \sum_{Z : wt(Z) \leq h-1} Pr(Z) \\ &\stackrel{(*)}{\leq} \sum_{t_Z \in \mathcal{T}_{g, h-1}} \prod_{i=1}^{g-h+1} \frac{1}{|\mathcal{F}|^{h-t_Z(i)}} \end{aligned}$$

In step (*), $\mathcal{T}_{g, h-1}$ consists of all t_Z sequences corresponding to Z -sequences of length g and weight $h - 1$. We use bound (11) for the first $g - h + 1$ 0's in the Z sequence and the upper bound 1 in (12) for all other cases which include the case that the position is after the $(g - h + 1)$'s 0 no matter what is the value of the Z sequence on the position. Replace

$\mathcal{T}_{g,h-1}$ by a bigger set $\mathcal{T}_{g,h-1}^*$ which consists of all sequences t of length $g-h+1$ with maximum entry value $h-1$ without the non-decreasing monotony requirement. That is, $t \in \mathcal{T}_{g,h-1}^*$ satisfies two conditions:

- 1) the length of t is $g-h+1$, i.e. $t = (t_1, \dots, t_{g-h+1})$;
- 2) for each $i : 1 \leq i \leq g-h+1$, $t_i \in \{0, \dots, h-1\}$.

We obtain an upper bound as below:

$$\begin{aligned} Pr(\Gamma_{k+1}^c | \Gamma_k, r_k = l) &\leq \sum_{t \in \mathcal{T}_{g,h-1}^*} \prod_{i=1}^{g-h+1} \frac{1}{|\mathcal{F}|^{h-t(i)}} \\ &= \left(\sum_{t=0}^{h-1} \frac{1}{|\mathcal{F}|^{h-t}} \right)^{g-h+1} \\ &\leq \frac{1}{(|\mathcal{F}| - 1)^{\delta_{t-l+1}}}. \end{aligned} \quad (13)$$

The lemma is proved. \square

Proof of Lemma 1. We prove this lemma by induction on k . For $k = 0$, Lemma 2 gives the desired result

$$p_e^{(0)} \leq \frac{1}{(|\mathcal{F}| - 1)^{\delta_{t+1}}}.$$

Assuming that the result of the lemma is proved for $k = 0, \dots, k'$ for all acyclic networks, we now prove it for $k = k' + 1$. We have

$$\begin{aligned} p_e^{(k'+1)} &\leq Pr(\Gamma_{k'}^c) + Pr(\Gamma_{k'}) Pr(\Gamma_{k'+1}^c | \Gamma_{k'}) \\ &= p_e^{(k')} + Pr(\Gamma_{k'}) \sum_{l=0}^{\delta_t} Pr(r_{k'} = l | \Gamma_{k'}) \\ &\quad \cdot Pr(\Gamma_{k'+1}^c | \Gamma_{k'}, r_{k'} = l). \end{aligned} \quad (14)$$

If $Rank(f_e : e \in COU_{k'}^o) = \omega$ then $Pr(\Gamma_{k'+1}^c | \Gamma_{k'}, r_{k'} = l) = 0$ for all l . The bound in Lemma 1 holds. In case that $Rank(f_e : e \in COU_{k'}^o) < \omega$, it is obvious that $r_{k'}$ takes values from 0 to δ_t . We prove the following observation.

Observation *The induction hypothesis, which says that for $k \leq k'$, the upper bound in Lemma 1 is valid for any acyclic network, implies*

$$Pr(r_{k'} = l | \Gamma_{k'}) \leq \frac{\binom{l+k-1}{k-1}}{(|\mathcal{F}| - 1)^l}.$$

This is proved as follows. Consider a subset N of $CUT_{k'}$ which includes all channels in $CUT_{k'}^o$ and $\omega - Rank(f_e : e \in CUT_{k'}^o) - 1$ channels from $CUT_{k'}^i$. Remove all paths that do not intersect with the set N . Thus we obtain a new network with minimum cut capacity $|N| = \omega + l - 1$. We use a subscript *new* to distinguish the quantities for this network and the same quantities for the original network. We can see that for this network,

$$Pr(r_{k'} = l | \Gamma_{k'}) \leq p_{e,new}^{(k')} \leq \frac{\binom{l+k}{k}}{(|\mathcal{F}| - 1)^l}.$$

Using this observations, we proceed as follows:

$$\begin{aligned} p_e^{(k'+1)} &\leq p_e^{(k')} + Pr(\Gamma_{k'}) \sum_{l=0}^{\delta_t} Pr(r_{k'} = l | \Gamma_{k'}) \\ &\quad \cdot Pr(\Gamma_{k'+1}^c | \Gamma_{k'}, r_{k'} = l) \\ &\stackrel{(*)}{\leq} \frac{\binom{\delta_t+k'+1}{k'}}{(|\mathcal{F}| - 1)^{\delta_{t+1}}} \\ &\quad + \sum_{l=0}^{\delta_t} \frac{\binom{l+k'}{k'}}{(|\mathcal{F}| - 1)^l} \frac{1}{(|\mathcal{F}| - 1)^{\delta_{t-l+1}}} \\ &= \frac{\sum_{l=0}^{\delta_t+1} \binom{l+k'}{k'}}{(|\mathcal{F}| - 1)^{\delta_{t+1}}} \\ &\stackrel{(**)}{=} \frac{\binom{\delta_t+k'+2}{k'+1}}{(|\mathcal{F}| - 1)^{\delta_{t+1}}}, \end{aligned} \quad (15)$$

which gives the desired result. In step (*), we used

- the induction hypothesis that the result of the lemma is valid for $k = k'$, i.e.,

$$p_e^{(k')} \leq \frac{\binom{\delta_t+k'+1}{k'}}{(|\mathcal{F}| - 1)^{\delta_{t+1}}};$$

- the Lemma 2, which gives

$$Pr(\Gamma_{k'+1}^c | \Gamma_{k'}, r_{k'} = l) \leq \frac{1}{(|\mathcal{F}| - 1)^{\delta_{t-l+1}}};$$

- and the second observation proved above which implies

$$Pr(r_{k'} = l | \Gamma_{k'}) \leq \frac{\binom{l+k'}{k'}}{(|\mathcal{F}| - 1)^l}.$$

In step (**), we used the formula $\sum_{i=0}^k \binom{n+i}{n} = \binom{k+n+1}{n+1}$. Thus the lemma is proved. \square

As a special case of the lemma, Theorem 3 is also proved.

Proof of Theorem 1. Note from [7], we have:

Lemma 3: If for any error pattern ρ satisfying $rank_t(\rho) = \delta$, the linear space $\{\text{row}_t(e) : e \in In(s) \cup \rho\}$ has dimension $\omega + \delta$, then $d_{min}^t \geq \delta + 1$.

In this lemma, the statement that the linear space $\{\text{row}_t(e) : e \in In(s) \cup \rho\}$ has dimension $\omega + \delta$, is equivalent to any of the following statements:

- The matrix $(\tilde{f}_e^\rho : e \in In(t))$ has rank $\omega + rank_t(\rho)$, where \tilde{f}_e^ρ is a $\omega + |\rho|$ dimensional column vector obtained from \tilde{f}_e by removing all entries $\tilde{f}_e(d)$ for $d \notin In(s) \cup \rho$. This is called the global kernel vector for channel e restricted to error pattern ρ .
- There exists $\omega + rank_t(\rho)$ linearly independent global kernel vectors restricted to ρ among channels in $In(t)$.
- If $rank_t(\rho) = |\rho| = \delta$, and the error pattern is known at the sink t , then both the source messages and the error messages from the channels in ρ can be decoded at t .

An erasure error for sink t is an error with the error pattern known by the decoder at sink t . Lemma 3 implies that if a code has erasure correction capability δ at t , then the minimum distance of the code at t is at least $\delta + 1$.

For each error pattern ρ , there exists an error pattern ρ' such that 1) $\rho \prec_t \rho'$, 2) $\text{rank}_t(\rho) = \text{rank}_t(\rho') = |\rho'|$. If the source messages and the error messages from the imaginary channels for channels in ρ can be decoded at t for any error pattern $\text{rank}_t(\rho) \leq |\rho| \leq \delta$ (when the error pattern is known at the decoder), then $d_{\min}^t \geq \delta + 1$. This becomes almost the same problem as in the case without errors.

Suppose that the bound derived for failure probability without channel errors can be applied to this case, then we can derive a bound for the probability in the theorem. We proceed as follows: First, the redundancy now is $d = \delta_t - \text{rank}_t(\rho)$ instead of δ_t since we add $|\rho| = \text{rank}_t(\rho)$ error messages. For fixed $d \geq 0$, consider all error patterns of $\text{rank}_t(\rho) = |\rho| = \delta_t - d$. For each such error pattern, the probability that $\langle \{\text{row}_t(e) : e \in \text{In}(s) \cup \rho\} \rangle$ has dimension lower than $\omega + \delta$ is upper bounded by $\frac{\binom{d+|J|+1}{|J|}}{(|\mathcal{F}|-1)^{d+1}}$. Then the probability that there exists a $t \in T$ and an error pattern ρ satisfying $\text{rank}_t(\rho) = |\rho| = \delta_t - d$ for which $\langle \{\text{row}_t(e) : e \in \text{In}(s) \cup \rho\} \rangle$ has dimension lower than $\omega + \delta_t - d$ is at most

$$\sum_{t \in T} \binom{|E|}{\delta_t - d} \frac{\binom{d+|J|+1}{|J|}}{(|\mathcal{F}|-1)^{d+1}}.$$

This implies that

$$\Pr(D_{\min}^t < \delta_t + 1 - d) \leq \frac{\binom{|E|}{\delta_t + 1 - d} \binom{d+|J|+1}{|J|}}{(|\mathcal{F}|-1)^{d+1}}.$$

Thus the theorem is proved.

We now prove that Theorem 3 can be applied in this case. In [7], the following lemma is proved.

Lemma 4: For any error pattern ρ , there exist at least C_t channel disjoint paths from either s or $\{\text{tail}(e) : e \in \rho\}$ to t having the properties that 1) there are exactly $\text{rank}_t(\rho)$ paths from $\{\text{tail}(e) : e \in \rho\}$ to t and 2) each of these $\text{rank}_t(\rho)$ paths from $\{\text{tail}(e) : e \in \rho\}$ to t starts with an erroneous channel in ρ .

If we consider only error pattern satisfying $\text{rank}_t(\rho) = |\rho|$ as we discussed above. The set of first channels of the paths that start with channels in ρ is exactly ρ . Define an imaginary node i_e for each $e \in \rho$. Use two channels $e_1 = (i, i_e)$ and $e_2 = (i_e, j)$ to replace channel e and define a channel $e' = (s, i_e)$. Channels $e' : e \in \rho$ provide error messages. Then this is a single source multicast problem of transmitting $\omega + |\rho|$ message symbols from s to t . For any code with local kernels k_{de} , amend the code by letting $k_{de_1} = k_{de}$, $k_{e_1 e_2} = 1$ and $k_{e' e_2} = 1$. Then the extended global kernel vector f_e^ρ will be exactly the global kernel vector for this new network. To apply Theorem 3 to this case, we need to consider:

- The encoding at i_e is no longer random but deterministic;
- The channels $e' : e \in \rho$ transmit error messages to node $i_e : e \in \rho$. The coding for these channels are also deterministic.
- In Theorem 3, the number of internal nodes $|J|$ occurs in the bound. If Theorem 3 can be applied, whether this number remains the same.

In Theorem 3, we assume that at all nodes the coding is random. In this case, coding for some nodes is deterministic.

We take the cuts for paths in Lemma 4 with the erroneous channels in ρ replaced by channels $e' = (s, i_e)$. The first cut CUT_0 includes all channels (s, i_e) whose global kernel vectors are the projection vectors of channels in ρ . Therefore, as long as there exist ω linearly independent global kernel vectors for other channels in the CUT_0 , there exist $\omega + |\rho|$ linearly independent global kernel vectors for channels in CUT_0 . This is because the values of global kernel vectors $f_e^\rho : e \in CUT_0 - \{(s, i_e) : e \in \rho\}$ at position $d \in \rho$ are all zero. Then at s , the problem is the same as in Theorem 3. This takes care of our second concern.

At node $v_{k+1} = i_e$, given Γ_k , from CUT_k to CUT_{k+1} , the only thing we do is to replace $e' = (s, i_e)$ by $e_2 = (i_e, j)$ (suppose that $e = (i, j)$). Since the global kernel vector for e' is the only vector among all global kernel vectors for channels in CUT_k which has a non-zero entry at channel e and $k_{e'e_2} = 1$, $f_{e_2}^\rho(e) \neq 0$. From the discussion above, this is the only global kernel vector with a non-zero entry at e among all such vectors for channels in CUT_{k+1} . It is apparent that Γ_k implies Γ_{k+1} if $v_{k+1} = i_e$. This implies that Theorem 3 can be applied to this case and nodes $i_e : e \in \rho$ should not be counted when we consider the number of nodes where we do coding. The theorem is proved. \square

III. CONCLUDING REMARKS

In this paper, we derived upper bounds on the failure probability of random linear network codes which can be viewed as improvements of similar results in [6]. The bound implies that for random linear block codes, the failure probability decays to 0 exponentially with increasing block length for a fixed redundancy rate and a fixed field. This bound is applied to network error correction codes where we derived a lower bound on the probability that a random linear network error correction code has degradation d . This bound implies an upper bound on the field size $|\mathcal{F}|$ required for the existence of a linear network error correction code with a given degradation. Many problems remain open in this research direction such as the design algorithm for network error correction codes with a given degradation when field size is not sufficient for the existence of MDS codes.

REFERENCES

- [1] R. W. Yeung and N. Cai, "Network error correction, part i: Basic concepts and upper bounds," *submitted to IEEE Trans. Inform. Theory*, Jan. 2006.
- [2] N. Cai and R. W. Yeung, "Network error correction, part ii: Lower bounds," *submitted to IEEE Trans. Inform. Theory*, Jan. 2006.
- [3] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.
- [4] R. W. Yeung, N. Cai, S.-Y. R. Li, and Z. Zhang, "Theory of Network Coding," *Foundations and Trends in Communications and Information Theory*, vol. 2, nos. 4 and 5, pp. 241–381, 2005.
- [5] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.
- [6] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and N. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.
- [7] Z. Zhang, "Network error correction coding in packetized networks," *to appear in IEEE Trans. Inform. Theory*, Mar. 2006.