An information theory for erasure channels

Kavé Salamatian
Université Pierre et Marie Curie
LIP6
Email: salamat@rp.lip6.fr

Ramin Khalili
Université Pierre et Marie Curie
LIP6
Email: khalili@rp.lip6.fr

I. Introduction

In this paper we consider a special class of communication channels called wireless erasure channels. In these channels, symbols sent over the channel are received errorless or erased and replace by the symbols e. These channels are very relevant for modelling communication networks from the viewpoint of higher layer where applications stand. In these networks the channel appear as erasure channels where a packet arrives at destination without errors or are erased by link layer error-detection mechanism because of transmission errors or collisions. Moreover, erasures might occur because of buffer overflows caused by network congestion.

Classically most of multi-user information theory researches have dealt with error channels and the erasure channel has been left out. However this class of channel is very important on different perspectives. Computer network uses erasure channels for communicating; meaning that information theory over erasure channels is relevant to these communication systems. Moreover the relative analytic simplicity of erasure channels make them attractive as an initial playground where essential characteristics of a communication problem could be extracted and the insight gained be used to tackle with the more complex error channels.

In this paper we present a rapid review of known results in the context of different scenario of the erasure channel. We will also give a converse capacity bound for relay channel showing that the cut-set bound is not attainable in general by any fixed coding scheme for the single sender-single relay case. This bound shows the difference between the degraded situation where the cut-set bound is attainable and some specific case of the non-degraded situation where the cut-set bound is not attainable. The obtained reverse bound could be used for deriving tighter cut-set type bound for general multi-terminal erasure channels. However, these tighter bound will not be presented here because of lack of space.

II. Point to point erasure channels

Erasure can be characterised by a conditional distribution function $\mathscr{E}^n(\mathrm{e}|x_n)$, giving the probability that an erasure is observed at the output of the channel given that a symbol x_n was sent over the channel at time t. If the probability of erasure is not a function of the sent symbol x_n, we might characterize an erasure channel by just an erasure probability p_n and we might drop the subscript if the channel is memoryless.

The erasure process $\{Z\}$ is a helpful notation that is defines as $Z_n = 1$, if an erasure was observed at the output of the channel at time n; 0 otherwise. The loss process completely defines the erasure channel. Let $\mathbf{X}^n(z^n)$ denote the subsequence of random variables X^n, such that $z^n(i) = 0$.

The following theorem is very helpful in the context of erasure channels.

Theorem 1 (Shearer Theorem [1]). *Let X^n be a collection of n random variables and Z^n be a collection of n boolean random variable, such that for each i, $1 \leq i \leq n$, $\mathbb{E}\{Z_i\} = p$. If $X^n(Z^n)$ is a sub-collection containing the i^{th} random variable X_i if $Z_i = 1$. Then $\mathbb{E}\{H(X^n(Z^n))\} = (1-p)H(X^n)$* □

The theorem can be extended to conditional entropy as well. This results in the fact that the mutual information over a stationary and ergodic point to point erasure channel with an erasure process Z have a very simple form given by

$$I(X^n;Y^n) = n(1-\mathbb{E}\{p\})H(X)$$

This directly led to the following theorem :

Theorem 2. *The capacity of a stationary and ergodic point-to- point erasure channel is given by :*

$$C = (1-\mathbb{E}\{Z\})\ symbols/trans.$$

Where $\{Z\}$ is the erasure process of the channel. □

This capacity result might be even extended to channels where the loss proportion accepts a law of large number. It can be proven than an erasure channel have a strong converse if $\frac{1}{n}\sum_{i\leq n} Z^n(i)$ converges in probability to a fixed value [2].

Proposition 1 ([2])**.** *The capacity of the erasure channel defined by the loss process $\{Z\}$ is :*

$$C = 1 - \limsup \frac{1}{n}\sum_{i\leq n} Z^n(i)$$

It is noteworthy that even if the capacity of the erasure channel only depends on the stationary loss probability but the memory structure may dramatically change the reliability function of the channel and the empirical behaviour of coding scheme for finite block length. In particular error exponent of erasure correcting codes can be obtained through the analysis of large deviation of the empirical proportion of losses [3].

One important characteristic of erasure channel is the availability of practical optimal codes. Let's suppose an erasure coding scheme (ϕ,ψ), where the coding function ϕ maps the message set $\mathscr{M} = \{1,\ldots,2^{nR}\}$ to a codeword set $\mathscr{C} \subseteq \mathscr{X}^n$. The decoding mapping is defined as $\psi : \mathscr{Y}^n \to \mathscr{P}(\mathscr{C})$ where $\mathscr{Y} = \mathscr{X} \cup \{\mathrm{e}\}$ and $\mathscr{P}(\mathscr{C})$ is the set of subsets of $\mathscr{C}$. A decoder for an erasure channel ψ implements a list decoding that will put in the coset $\psi(\mathbf{y}^n)$, all sequences in $\mathscr{C}$ that are in agreement with the symbols received in $\mathbf{y}^n$. If the coset contains a single codeword, the sent message is decoded correctly. If the coset contains more than a single codeword, ambiguity will subsists about the sent message. A decoding based on a random choice among the coset will results in a decoding error probability $(1-\frac{1}{|\psi(\mathbf{y}^n)|})$, *i.e* any ambiguity leads to a non-vanishing error probability. The straightforward relationship between the probability of erroneous decoding and the probability of ambiguity is particular to erasure channels and cannot be extended to error channels. Nevertheless, the given description is not particular to a specific coding scheme and different coding schemes differs on the codeword set $\mathscr{S}$ used.

Performance of a coding scheme depends on Hamming distance properties of the codeword set $\mathscr{C}$; if an erasure code have a minimal Hamming distance $d_{\min}$, it would able to decode up to $d_{\min}-1$ erasures. By Singleton bound, it is known that the largest possible minimal Hamming distance in a set containing K n-dimensional codewords is $n - \log(K) + 1$. A code attaining the Singleton Bound is called a Maximal Distance Separable code (MDS). If an MDS codeword suffers $e(\mathbf{y}^n)$ erasure and $e(\mathbf{y}^n) < d_{\min}$, one can still decode the sent codeword.

Such codes attain the capacity of the erasure channel. This can be proven by observing that typical erasure pattern for erasure channels that accept a law of large numbers for erasure proportion , have a proportion of erasure that is within a distance $\delta(\epsilon)$ of the erasure probability, *i.e.* $\forall \mathbf{y}^n \in \mathscr{A}(\epsilon)$, $|\frac{e(\mathbf{y}_R^n)}{n} - p| < \delta(\epsilon)$ with $\lim_{\epsilon\to 0}\delta(\epsilon) = 0$. In other terms the proportion of erasure converges asymptocally to a value p, the typical set is a subset of the sets of erasure patterns that have a proportion of erasure close to p and the probability than the number of erasures in a received sequence $\mathbf{y}^n$ exceeds np goes to zero with larger block size n. Therefore for MDS codes with rate $R<(1-p)$ the probability that the number of erasures exceed the minimal distance goes to 0 with larger block size. Clearly because of the maximal distance property of the MDS code, if these codes exist they are optimal. However, it is well know that Reed Solomon codes are MDS and for every finite block length n it is possible to choose an alphabet size q and to build a Reed Solomon code (n,k,q) with rate $R=\frac{k}{n}$.

III. BROADCAST ERASURE CHANNELS

In this section we consider a broadcast erasure channel with *degraded message set*. This channel have a single source that want to transmit information to a set of k receivers. We will present the analysis for the transmission of "*Degraded Message Set*" (DMS) [4] as it gives the largest general capacity region known up to now. For such a message set the capacity region is the set of rate tuples $(R_1,\cdots,R_k)$ such that the source can reliably transmit with rate R_i to the receiver i the i^{th} level of the DMS. A k-ary broadcast channel $\mathscr{B}$ consists in a sequence of joint probability transitions $\mathbb{P}\mathrm{rob}\{Y_0^n,\cdots,Y_{k-1}^n|X^n\}$ from $\mathscr{X}^n$ toward $\mathscr{Y}^{n\times k}$. The marginal probability transmission $\mathbb{P}\mathrm{rob}\{Y_i^n|X^n\}$ are called the component channels. An erasure broadcast channels could be characterized by the erasure probability only.

The capacity region for the erasure broadcast has been derived in [5] where it was shown that the set of ϵ-achievable rates $(R_0,\cdots,R_{k-1})$ over a broadcast channel $\mathscr{B}$ with the degraded message set is the closure of :

$$\bigcup_{u_0,\cdots,u_{k-1}\in\mathscr{U}} R_B(\epsilon,U_{i<k})$$

where $R_B(\epsilon,U_{i<k})$ is the set of tuple rates $(R_0,\cdots,R_{k-1})$ satisfying :

$$\begin{aligned}
&0\le R_i \le I(U_i;X(Z_i)|U_{i-1},\cdots,0)\\
&0\le R_{k-1} \le H(X(Z_{k-1})|U_{k-2},\cdots,0)\\
&0\le \sum_{j\le i} R_j \le I(U_i;X(Z_i))\;\; i<k-1\\
&0\le \sum_{i<k} R_i \le H(X(Z_{k-1}))
\end{aligned}$$

where Z is the loss process defined as $Z^n(i)=1$ if $Y^n(i)$ is a loss, 0 otherwise. $X^n(z^n)$ denotes the subsequence of random variable $X^n(i)$, such that $Z^n(i)=0$. Using the Shearer theorem we have therefore the remarkably simple expression form for the capacity region :

Theorem 3 ([5]). *A tuple of rates $(R_0,\cdots,R_{k-1})$ is achievable over a memoryless broadcast erasure channel with degraded message set if and only if :*

$$\sum_{i=1}^{k-1}\frac{R_i}{1-p_i}<1$$

where p_i is the packet loss rate between sender and receiver i. □

The capacity region of the channel can be attained by a very simple time-sharing method to design broadcast codes. We partition the information flow in different priority levels. Each priority level is encoded using an MDS code with a rate equal to $(1-p_i)$. The different MDS encoders use the same alphabet and block size, but they have different encoding rates. The encoded data are interleaved and encapsulated in symbols where each resulted symbol contains k different levels of

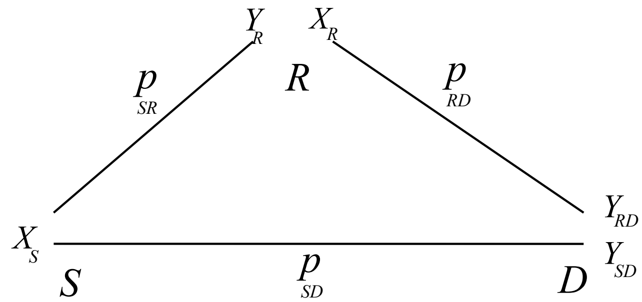


Fig. 1. Single relay erasure channel

protection. These encapsulated symbols are sent over the channel and if a receiver received $(1-p_i)$ percent of the packets sent over the channel it is able to decode i^{th} level of priority.

It is moreover proven in [5] that this encoding scheme is optimal, *i.e.* no other coding scheme can have better error exponents that this scheme over erasure broadcast channels.

However the broadcast in it simplest version do not make use of any node collaboration. In next section we will begin to analyse situation where node collaboration enter in the game.

IV. Single relay erasure channel

The simplest scenario of node collaboration is the single relay channel. In this setting we have three nodes: one node (S) acting as a sender , one destination node (D) that receives symbols from sender and a relay node (R) and decode them to figure out the message send the sender, and one relay node receiving information from sender and collaborating with it to transmit information to receiver through forwarding encoded symbols to the receiver. The relay channel is characterized by three erasures probabilities : p_{SR}, p_{SD}, P_{RD}.

A. Erasure coding for the relay channel

The operation of any relay coding scheme can be described by the list decoding approach described for point to point channels. Let's assume that the source in a relay channel is using an encoding function $\phi_S : \{1, \dots, 2^{nR}\} \to \mathscr{C}_S$ and the relay in this channel is using an encoding function $\phi_R : \mathscr{Y}^n \to \mathscr{C}_R$, where $|\mathscr{C}_R| = 2^{nR'}$. The mapped sequence $\phi_R(\mathbf{y}_R^n)$ will be sent to the destination by the relay.

Without loss of generality, the encoding function at relay could be decomposed into two functions $\phi_R = \phi_{RD} \circ \psi_R$, where $\psi_R : \mathscr{Y}^n \to \mathscr{P}(\mathscr{C}_S)$ is a list decoding mapping characterizing the ambiguity about the message sent by the source. The coset $\psi_R(\mathbf{y}_R^n)$ will contain all sequences in $\mathscr{C}_S$ that are in agreement with $\mathbf{y}_R^n$. If the message sent by source is decoded at relay, the coset $\psi_R(\mathbf{y}_R^n)$ will contain only one codeword. $\phi_{RD} : \mathscr{P}(\mathscr{C}_S) \to \mathscr{C}_R$ is an encoding function mapping a subset of the codeword space to a single codeword in $\mathscr{C}_R$. To ensure that ϕ_R is a deterministic function mapping each received symbol sequence $\mathbf{y}_R^n$ to a single value, $\phi_{RD}^{-1}(x^n)$ for any $x \in \mathscr{C}_R$ consists of a union of cosets $\psi_R(\mathbf{z}^n) \subseteq \mathscr{C}_S$ for different received sequences $\mathbf{z}^n$.

At the destination node two sequences of symbols are received: $\mathbf{y}_{SD}^n$ from source, $\mathbf{y}_{RD}^n$ from relay. The destination implements a joint decoding that will be described thereafter. By using only the

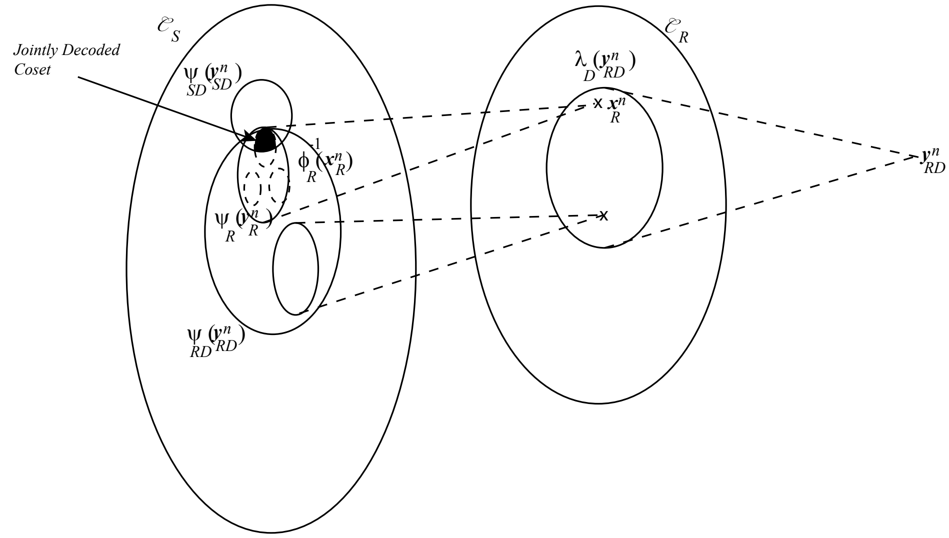


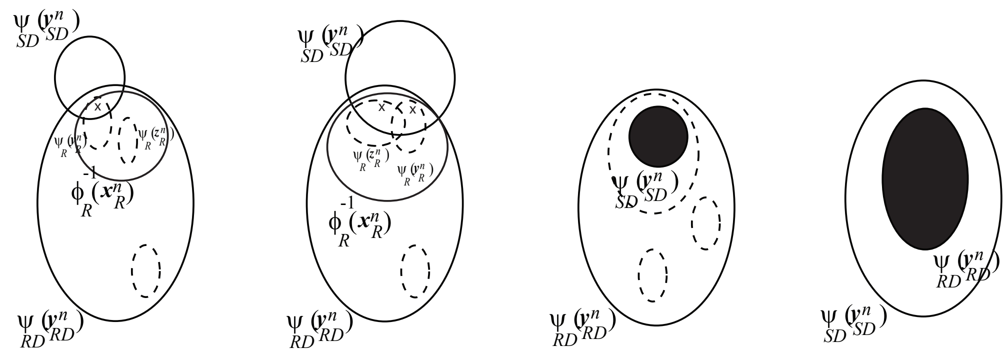
Fig. 2. Structure of joint erasure decoding of $\mathbf{y}^n_{RD}$ and $\mathbf{y}^n_{SD}$ code

sequence coming directly from the source $\mathbf{y}^n_{SD}$ and using a list decoding, the coset $\psi_{SD}(\mathbf{y}^n_{SD})$ will contain all sequences in $\mathscr{C}_S$ that are in agreement with the symbols received in $\mathbf{y}^n_{SD}$.

The sequence coming from the relay is more complex to exploit. Here also a list decoding can be applied. The list decoding $\psi_{RD} : \mathscr{Y}^n \to \mathscr{P}(\mathscr{C}_S)$ maps a sequence $\mathbf{y}^n_{RD}$ received from relay at destination to a coset $\psi_{RD}(\mathbf{y}^n_{RD})$ of all codewords in $\mathscr{C}_S$ compatible with a received sequence $\mathbf{y}^n_{RD}$. Without loss of generality, the list decoding mapping ψ_{RD} can be decomposed into two main function : $\psi_{RD} = \phi^{-1}_{RD} \circ \lambda_D$, where $\lambda_D : \mathscr{Y}^n \to \mathscr{P}(\mathscr{C}_R)$ is a list decoding mapping characterizing the ambiguity about the codeword sent by relay at destination. The coset $\lambda_D(\mathbf{y}^n)$ will contain all sequences in $\mathscr{C}_R$ that are in agreement with the symbols received in $\mathbf{y}^n_{RD}$. This coset $\lambda_D(\mathbf{y}^n)$ is thereafter mapped by ϕ^{-1}_{RD} to the subset of $\mathscr{C}_S$ that might have led to the transmission of any elements of $\lambda_D(\mathbf{y}^n_{RD})$ by the relay, *i.e* $\psi_{RD}(\mathbf{y}^n_{RD}) = \bigcup_{\mathbf{x}^n \in \lambda_D(\mathbf{y}^n_{RD})} \phi^{-1}_{RD}(\mathbf{x}^n)$.

The joint decoding of $\mathbf{y}^n_{RD}$ and $\mathbf{y}^n_{SD}$ is also a list decoding function $\psi_D : \mathscr{Y}^n \times \mathscr{Y}^n \to \mathscr{P}(\mathscr{C}_S)$. The decoded message list $\psi_D(\mathbf{y}^n_{RD}, \mathbf{y}^n_{SD})$ contains codewords that are jointly compatible with the two received symbols sequences $\mathbf{y}^n_{RD}$ and $\mathbf{y}^n_{SD}$. In other terms $\psi_D(\mathbf{y}^n_{RD}, \mathbf{y}^n_{SD}) = \psi_{SD}(\mathbf{y}^n_{SD}) \bigcap \psi_{RD}(\mathbf{y}^n_{RD})$. We show in Fig. 2 an illustration explaining the decoding in simple relay channel. It is noteworthy that because of the particular nature of erasure channels, the previous description of erasure decoding is generic and is not specific to a particular coding scheme. Any coding scheme for erasure might be reformulated in the previous framework. In a deterministic coding scheme, for each possible combination of received sequences $\mathbf{y}^n_{RD}$ and $\mathbf{y}^n_{SD}$, the sets $\psi_{SD}(\mathbf{y}^n_{SD})$ and $\psi_{RD}(\mathbf{y}^n_{RD})$ are fixed and chosen before transmission. The difference between different coding schemes is relative to the difference in the content of $\mathscr{C}_S$, $\mathscr{C}_R$ and $\phi^{-1}_{RD}(\mathbf{x}^n)$ for all $\mathbf{x}^n \in \mathscr{C}_R$ and their performance are function of minimal distance properties these sets .

Because the erasure nature of the channel and the coset structure, the intersection of the two decoding sets is never empty $\psi_{SD}(\mathbf{y}^n_{SD}) \bigcap \psi_{RD}(\mathbf{y}^n_{RD}) \neq \emptyset$. The joint decoding could therefore results in one of the four different situations depicted in Fig. 3. The situation depicted in (c) and (d) are easy to explain. These are situation where all the symbols received at relay (resp. destination) have also been received at destination (resp. relay). Under this situation decoding of the message should be done using a single symbol sequence and joint decoding is not possible. Nevertheless without spatial correlation of the loss process this situation is non-typical and very unlikely to happen.



(a) Situation where the joint decoding works

(b) Problematic situation

(c) All symbol received at destination are also received at relay

(d) All symbols received at relay are also received at destination

Fig. 3. Different possible situation of the joint decoding at destination

As explained, the set $\psi_{RD}(\mathbf{y}_{RD}^n)$ consists of an union of cosets $\psi_R(\mathbf{z}^n)$ for different received sequences $\mathbf{z}^n$. We have therefore the two possible situation depicted in Fig. 3(a) and (b). The only situation that might results in a correct decoding without ambiguity of the sent message is the situation in Fig. 3(a) (we have also to ensure that only one codeword of $\mathscr{C}_S$ is in $\psi_{SD}(\mathbf{y}_{SD}^n)\bigcap\psi_{RD}(\mathbf{y}_{RD}^n)$). The situation in Fig. 3(b) will also results in ambiguity at decoding. We emphasize one more time that the described situation is not specific to any coding scheme but is rather generic for every coding scheme used in the erasure relay channel.

B. Capacity bounds

By extension we call "*virtually degraded*" the situation where the relay can decode the message

C. Performance analysis of erasure codes over relay channels

Performance of a specific coding scheme will depend of the properties of codewords chosen in each one of the cosets. As a codeword set with a minimal distance $d_{\min}$ can correct up to $d_{\min}-1$ erasures, it is desirable to increase as possible the minimal distance in the codeword sets. However, by Singleton bound the minimum distance of a codeword set containing 2^{nR} symbols is bounded by $d_{\min} < n(1-R)$. MDS codes attains this bound and are therefore very suitable for erasure channels as there are sphere packing codes for the erasure channels, *i.e.* if MDS code exists no coding scheme can have better performance than these codes. As we wish here to provide reverse bound we will assume in the forthcoming that all codeword set ($\mathscr{C}_S$ and $\mathscr{C}_R$) are MDS.

The coset $\psi_{SD}(\mathbf{y}_{SD}^n)$ have a simple structure. It contains all codeword that are compatible with a received sequence $\mathbf{y}_{SD}^n$. However under typical erasure pattern with high probability ($> 1-\epsilon$) $n(1-p_{SD})+\delta(n)$[1] symbols will be received in $\mathbf{y}_{SD}^n$ at destination. The coset $\psi_{SD}(\mathbf{y}_{SD}^n)$ contains all codewords that have agrement with the $n(1-p_{SD})+\delta(n)$ symbols received in $\mathbf{y}_{SD}^n$ in common. The maximal miminum distance in $\psi_{SD}(\mathbf{y}_{SD}^n)$ can be easily derived. All codewords in $\psi_{SD}(\mathbf{y}_{SD}^n)$ have $n(1-p_{SD})+\delta(n)$ symbols in common. Moreover, if the initial codewords set $\mathscr{C}_S$ was chosen to be an MDS set, there is $2^{nR-n(1-p_{SD})+\delta(n)}$ codewords in $\psi_{SD}(\mathbf{y}_{SD}^n)$. Therefore, by Singleton bound the largest minimal distance in $\psi_{SD}(\mathbf{y}_{SD}^n)$ would be $d_{\min}(\psi_{SD}(\mathbf{y}_{SD}^n)) \leq n(1-R)+1$. This property can be generalized for any coset $\psi_R(\mathbf{z}^n)$, $d_{\min}(\psi_R(\mathbf{z}^n)) < n(1-R)+1$ (because of the constraint that all codeword compatible with a received sequence $\mathbf{z}^n$ have to be in the coset).

[1]In the forthcoming we will use $\delta(n)$ as the generic term with the property that $\lim_{n\to\infty}\delta(n)=0$

Analysis of the coset $\psi_{RD}(\mathbf{y}^n_{RD})$ is more challenging. As explained, the set $\psi_{RD}(\mathbf{y}^n_{RD})$ consists of an union of cosets $\psi_R(\mathbf{z}^n)$ for different received sequences $\mathbf{z}^n$. However, for transmission rate $R < (1-p_{SR})$, the coset $\psi_R(\mathbf{y}^n_R)$ consists of a single point (as the message sent by source can be decoded by relay). If the transmission rate $R > (1-p_{SR})$, the coset $\psi_R(\mathbf{y}^n_R)$ will contains $2^{nR-n(1-p_{SR}+\delta(n))}$ codewords that are all compatible with $\mathbf{y}^n_R$. Moreover if $(1-p_{SR}) < (1-p_{RD})$, each codeword in $\mathscr{C}_R$ can be related to only one coset $\psi_R(\mathbf{y}^n_R)$; if $(1-p_{SR}) > (1-p_{RD})$, each codeword in $\mathscr{C}_R$ should be related to several cosets $\psi_R(\mathbf{y}^n_R)$.

The following theorem gives a reverse coding bound for the relay channel.

Theorem 4 (Converse theorem for relay erasure channel). *No erasure coding scheme can exceed the following bound over an erasure relay channel :*

$$\begin{cases} R & < 1-p_{SD}p_{SR}, & if\ (1-p_{SR}) \leq (1-p_{RD}) \\ R & < \max\{T,(1-p_{SD})\}, & if\ (1-p_{SR}) > (1-p_{RD}) \end{cases}$$

with $T = \min\{(1-p_{SR}),(1-p_{SD})+(1-p_{RD})\}$ □

Proof. *The proof proceeds by using the list decoding interpretation for erasure coding schemes and using properties of the decoding set $\psi_{RD}(\mathbf{y}^n_{RD})$ under different situations. The keystone argument is based on the universality of the list decoding based joint decoding and the fact that that no code can have a better performance in term of probability of decoding error than an MDS code*

- $(1-p_{SR}) < (1-p_{RD})$:
  Under this hypothesis $\psi_{RD}(\mathbf{y}^n_{RD})$ contains a single coset but with $\min\{2^{nR-n(1-p_{SR}+\delta(n))},1\}$ codewords inside (for $R < (1-p_{SR})$ the coset contains only one point). We will be in the situation depicted in Fig. 3(a). All points in the intersection $\psi_{SD}(\mathbf{y}^n_{SD})\bigcap\psi_{RD}(\mathbf{y}^n_{RD})$ should be compatible with $\mathbf{y}^n_{SD}$ and $\mathbf{y}^n_R$. All codewords in $\psi_{SD}(\mathbf{y}^n_{SD})\bigcap\psi_{RD}(\mathbf{y}^n_{RD})$ will therefore have (under typical erasure pattern and without any spatial correlation in erasures) $n(1-p_{SD}p_{SR})+\delta(n)$ symbols in common. To ensure that there is a single codeword in the intersection we should have $d_{\min}(\psi_{SD}(\mathbf{y}^n_{SD})) > n-(n(1-p_{SD}p_{SR})+\delta(n))$, i.e. $R < (1-p_{SD}p_{SR})-\delta(n)$.
- $(1-p_{SR}) > (1-p_{RD})$ *and* $R < (1-p_{SR})$*:*
  Under this hypothesis $\psi_{RD}(\mathbf{y}^n_{RD})$ contains $2^{nR-n(1-p_{RD})}$ cosets with a single codeword inside each one. The codewords are not constrained to be compatible with each other. If we assume $R' = (1-p_{RD})$, the maximal minimum distance in the set $\psi_{RD}(\mathbf{y}^n_{RD})$ is therefore $d_{\min}(\psi_{RD}(\mathbf{y}^n_{RD})) =$ $n-(nR-n(1-p_{RD}))+1$. *To ensure that under such extreme condition there is a single codeword in the intersection we should have $d_{\min}(\psi_{RD}(\mathbf{y}^n_{RD}) > n-(n(1-p_{SD})+\delta(n))$,* i.e. $R < (1-p_{SD})+(1-p_{RD})-\delta(n)$.
- $(1-p_{SR}) > (1-p_{RD})$ *and* $R > (1-p_{SR})$*:*
  Under this hypothesis $\psi_{RD}(\mathbf{y}^n_{RD})$ contains several cosets with $2^{nR-n(1-p_{SR})}$ codewords in each. These cosets might have intersection or being separated. For this case we should avoid being in the situation depicted in Fig. 3 where ambiguity will remains between the correct codeword in $\psi_R(\mathbf{y}^n_R)$ and another codeword not compatible with $\mathbf{y}^n_R$ but still in $\psi_{SD}(\mathbf{y}^n_{SD})$, i.e. *the codeword is in agreement with the $n(1-p_{SD})+\delta(n)$ symbols received in $\mathbf{y}^n_{SD}$ under typical erasure pattern, and not be in agreement with at least one symbols out of the $np_{SR}(1-p_{SR})+\delta(n)$ symbols received at relay but not at destination (under typical erasure patterns).*
  If the source S is sending at a rate $R > (1-p_{SD})$, the maximal minimum distance of $\mathscr{C}_S$ will be smaller than $np_{SD}+1$. the set $\phi^{-1}_{RD}(\lambda^{-1}(\mathbf{y}^n_{RD}))-\psi_R(\mathbf{y}^n_D)$ have also a minimal distance smaller than $np_{SD}+\delta(n)$ (as it contains at least another coset $\psi_R(\mathbf{z}^n)$ and it was shown that the minimal distance in any coset is smaller than $n(1-R)$). This means that it is impossible to design a set $\phi^{-1}_{RD}(\mathbf{x}^n)-\psi_R(\mathbf{y}^n_D)$ where one can guarantee that for all possible reception patterns containing $n(1-p_{SD})+\delta(n)$ received symbols there is no codeword in $\psi_{SD}(\mathbf{y}^n_R)$ (as this will means that

the minimal distance in $\phi_{RD}^{-1}(\mathbf{x}^n)$ is larger than np_{SD}). There exists therefore a typical erasure pattern containing $n(1-p_{SD})+\delta(n)$ received symbols such that there is at least one codewords in $\phi_{RD}^{-1}(\lambda^{-1}(\mathbf{y}_{RD}^n))-\psi_R(\mathbf{y}_D^n)$ that is compatible with $\mathbf{y}_{SD}^n$. This means that whenever transmission rate goes higher than $(1-p_{SD})$ and the set $\phi_{RD}^{-1}(\lambda^{-1}(\mathbf{y}_{RD}^n))$ contains more than a single coset with more than one codeword inside, we will be in situation depicted in Fig. 3(b) with a probability larger than 0. In summary, it is impossible to transfer reliably over an erasure relay channel using a fixed coding scheme with a rate higher than $(1-p_{SD})$ if $\phi_{RD}^{-1}(\mathbf{x}^n)$ contain more than a single coset.

The last item show also that the transmission rate over the channel from relay to destination (R') should never exceed $(1-p_{RD})$ as it will lead directly to the set $\psi_{RD}(\mathbf{y}_{RD}^n)$ containing more than one cosets. ∎

Remarks

- Linear coding schemes for the erasure relay channel that can be applied in practice have been presented in [6], [7]. These codes can solve the joint decoding problem described previously by solving a linear system of equation with nR variables with a complexity $\mathcal{O}(n\log(n))$. Almost-MDS [8] codes with linear decoding complexity are also applicable in this context.
- The physically degraded channel situation is when all symbols received at destination are also received at relay. In this situation $\psi_{SD}(\mathbf{y}_{SD}^n)\subseteq\psi_R(\mathbf{y}_R^n)$ and the only situation joint decoding succeed is when $\psi_R(\mathbf{y}_R^n)$ contains a single codeword, *i.e.* when $R<(1-p_{SR})$. The converse bound for physically degraded channel conditions is therefore $R<\min\{(1-p_{SR}),(1-p_{SD})+(1-p_{RD})\}$. In [7] a milder situation that supersed the physical degraded condition and is enough to make the cut-set bound tight defined as "*virtually degraded*" condition. Under this condition the relay should be able to decode the message sent by source.
- The proof of the converse shows three types of collaboration for the relay node. The first type of collaboration we will call "*active collaboration*" is possible if we are in virtually degraded condition,*i.e.* when the relay can decode the message sent by source. Active collaboration of relay consists of arranging points in $\phi_{RD}^{-1}(\mathbf{x}^n)$ such that it makes a MDS set. This type of collaboration is possible when $R<(1-p_{SR})$. The second type of collaboration we call "*passive collaboration*" is when the relay cannot decode the message sent by sender, but nevertheless can forward all its received symbols to destination without rearranging the points in $\phi_{RD}^{-1}(\mathbf{x}^n)$. This type of collaboration occurs when $(1-p_{SR})<(1-p_{RD})$ and $R>(1-p_{SR})$. The last type of collaboration is the "*no collaboration*" state that occurs when $(1-p_{SR})>(1-p_{RD})$ and $R>(1-p_{SR})$. Under this setting the sender cannot ensure any fruitful collaboration for sending information to final destination and it is useless to forward any information.
- The theorem shows that the well-known and classical cut-set bound is not achievable when $(1-p_{SR})>(1-p_{RD})$. A precise examination of the proof show that the bottleneck is when the decoding set $\phi_{RD}^{-1}(\mathbf{x}^n)$ contains more than one coset. In fact the theorem says that when $(1-p_{SR})>(1-p_{RD})$, the information coming from the relay are not useful for decoding purposes. Two solutions have been proposed two overcome this shortage. One solution proposed in [9] uses a side information in the form of the erasure pattern over the source to relay channel. This side information could be used when the decoding set $\psi_{RD}(\mathbf{y}_{RD}^n)$ is designed to contain different cosets relative to different erasure patterns over the source to relay channel, *i.e.* for every erasure pattern there is only one coset in $\psi_{RD}(\mathbf{y}_{RD}^n)$ compatible with this pattern. We therefore fall back to the situation depicted in Fig. 3(a) and the cut-set bound can be one more time attained. Nevertheless, the amount of extra information needed to transfer the side information should be assessed. A simple evaluation show that one needs $n(1-h(p))$ bits of extra information to transfer as side information the erasure pattern over the sender to relay

channel. Accounting the information rate needed for side information, the scheme proposed in [9] attains a proportion $\frac{R\log(|\mathscr{X}|)}{R\log(|\mathscr{X}|)+(1-h(p))}$ of the cut-set bound.

- Another solution to expand the converse bound is proposed in [10]. Classically in information theory, one choose a random code at the beginning of the communication and inform the receiver about the used coding scheme. The used coding scheme might be seen as a side information that is given at the beginning of communication. However, in [10], the use of a randomly changing coding scheme is proposed, *i.e.* the relay node choose randomly at each transmission how to mix the received symbols and the decoding set $\phi_{RD}^{-1}(\mathbf{x}^n)$ is not fixed (as in classical settings) but will changes randomly during the transmission. By averaging over erasure channel as well as random code statistics, the probability that $\psi_{RD}(\mathbf{y}_{RD}^n)$ becomes a MDS set goes to 1. A naïve evaluation of the amount of extra information needed to transfer the side information leads to $\mathscr{O}(n)$ bits per transferred symbols that results to an asymptotic information rate (accounting the amount of information needed for side information) of transfer from source to destination equal to 0.

REFERENCES

[1] F. R. K. Chung, R. L. Graham, P. Frankl, and J. B. Shearer, "Some intersection theorem for ordered set and graphs," *Journ. of Comb. Theory*, vol. 43, no. Series A, no. Series A, pp. 23–37, 1986.

[2] K. Salamatian, *Transmission Multimédia Fiable Sur Internet*. PhD thesis, Université Paris SUD, 1999.

[3] S. Boucheron and K. Salamatian, "On the error exponent of some markovian channels," *Actes du 31ème GRETSI*, September 1999.

[4] J. Korner and K. Marton, "General broadcast channels with degraded messages sets," *IEEE. Trans. Inform. Theory*, vol. IT-23, pp. 60–64, 1977.

[5] S.Boucheron and K.Salamatian, "About priority encoding transmission," *IEEE. Trans. Inform. Theory*, vol. on processing.

[6] R. Khalili and K. Salamatian, "On the achievability of cut-set bound for a class of relay channels," *IWWAN*, 2004.

[7] R. Khalili and K. Salamatian, "On the achievability of cut-set bound for a class of erasure relay channels: The non-degraded case," *ISITA conference*, 2004.

[8] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, D. A. Spielman, and V. Stemann, "Practical loss-resilient codes," pp. 150–159, 1997.

[9] R. Gowaikar, A. F. Dana, R. Palanki, B. Hassibi, and M. Effros, "On the capacity of wireless erasure networks," *ISIT*, 2004.

[10] D. S. Lun, M. Médard, and M. Effros, "On coding for reliable communication over packet networks," *Allerton*, 2004.