Decoding of MDP Convolutional Codes over the Erasure Channel

Virtudes Tomás*
Department of Computational Science and Artificial Intelligence
University of Alicante
Alicante, Spain
Email: vtomas@dccia.ua.es

Joachim Rosenthal°
Mathematics Institute
University of Zürich
Winterthurerstr 190
CH-8057 Zürich, Switzerland
www.math.uzh.ch/aa

Roxana Smarandache‡
Department of Mathematics and Statistics
San Diego State University
San Diego, CA 92182-7720, USA
Email: rsmarand@sciences.sdsu.edu

***Abstract*—This paper studies the decoding capabilities of maximum distance profile (MDP) convolutional codes over the erasure channel and compares them with the decoding capabilities of MDS block codes over the same channel. The erasure channel involving large alphabets is an important practical channel model when studying packet transmissions over a network, e.g, the Internet.**

Keywords — Convolutional codes, maximum distance separable codes, parity check matrix, decoding, erasure channel, Reed-Solomon codes.

I. Introduction

When transmitting over an erasure channel like the Internet, one of the problems encountered is the delay experienced on the received information which is due to the possible retransmission of lost packets. One way to eliminate these delays is by using forward error correction.

Until now mainly block codes have been used for such a task, see e.g. [2] and the references therein. The use of convolutional codes over the erasure channel has been studied much less and we are aware of the work of Epstein [1] and the more recent work by Arai et al. [3]. In this paper we demonstrate how maximum distance profile (MDP) convolutional codes provide an attractive alternative to block codes.

Convolutional codes have a certain flexibility given by the "sliding window" characteristic. This means that the received information can be grouped in blocks or windows in many ways, depending on the erasure bursts, and then be decoded by decoding the "easy" blocks first. This flexibility in grouping information brings certain freedom in the handling of sequences; we can split the blocks in smaller windows, we can overlap windows, etc., we can proceed to decode in a less strict order. The blocks are not fixed as in the block code case, i.e., they do not have a fixed grouping of a fixed length. We can slide along the transmitted sequence and decide the place where we want to start our decoding depending on the erasure occurrence. This property allows us to correct in a given block more erasures than a block code of that same length could do.

An $[N, K]$ block code used for transmission over an erasure channel can correct up to $N-K$ erasures in a given block. The optimal error capability of $N - K$ is achieved by an $[N, K]$ maximum distance separable (MDS) code.

As an alternative consider now a class of (n, k, δ) convolutional codes, i.e., a class of rate k/n convolutional codes having degree δ. We will demonstrate that for this class, the maximum number of errors which can be corrected in some sliding window of appropriate size is achieved by the subclass of MDP convolutional codes. In this paper, we will study the maximum number of erasures that such a class of codes can decode and the conditions under which this happens. Moreover we will show that over the erasure channel this class of codes can decode efficiently. To be more specific we will show that a code over a very large alphabet (e.g q in the range of 2^{1000}) and the number of states in the order of q^{50} can still be decoded in practical terms on a personal computer. This is of course something which cannot be achieved by trellis decoding.

The paper is organized as follows. Section II provides the necessary background for the development of the paper. Thus, subsection II-A explains the assumptions on the channel model; subsection II-B provides all the necessary concepts about MDP convolutional codes and their characterizations. Section III is the main part of the paper. It contains our main result and describes in detail the decoding procedure. It also provides examples and special concerns to be noticed when comparing with MDS block codes, and in particular with Reed-Solomon codes. Section IV shows a decoding method in which the transmitted information is recovered directly.

II. Preliminaries

A. Erasure channel

An erasure channel is a communication channel where the symbols sent either arrive correctly or the receiver knows that a symbol has not been received or was received incorrectly. An important example of an erasure channel is the Internet, where

* Partially supported by Spanish grant MTM2008-06674-C02-01 and a grant of the Vicerectorat d'Investigació, Desenvolupament i Innovació of the Universitat d'Alacant for PhD students during a stay at Zürich Universität on charge to the same program.

° Supported by the Swiss National Science Foundation under Project no. 113251.

‡ Supported by NSF Grants DMS-0708033 and TF-0830608.

packet sizes are upper bounded by 12,000 bits - the maximum that the Ethernet protocol allows (that everyone uses at the user end). In many cases, this maximum is actually used. Due to the nature of the TCP part of the TCP/IP protocol stack, most sources need an acknowledgment confirming that the packet has arrived at the destination; these packets are only 320 bits long. So if everyone were to use TCP/IP, the packet size distribution would be as follows: 35% –320 bits, 35% – 12,000 bits and 30% – in between the two, uniform. Real-time traffic used, e.g., in video calling does not need an acknowledgment since that would take too much time; overall, the following is a good assumption of the packet size distribution: 30% – 320 bits, 50% – 12,000 bits, 20% –in between, uniform.

We can model each packet as an element or sequence of elements from a large alphabet. Since packets over the Internet are usually protected by a cyclic redundancy check (CRC) code the receiver knows when a packet is in error or has not arrived. For the purpose of illustration we could employ as alphabet the finite field $\mathbb{F} := \mathbb{F}_{2^{1,000}}$. If a packet has less than 1,000 bits then one uses simply the corresponding element of $\mathbb{F}$. If the packet is larger one uses several alphabet symbols to describe the packet. Even if one uses some interleaving, such an encoding scheme results in the property that errors tend to occur in bursts and this is a phenomena observed about many channels modeled via the erasure channel. This point is important to keep in mind when designing codes which are capable of correcting many errors over the erasure channel.

B. MDP convolutional codes

Let $\mathbb{F}$ be a finite field. We view a convolutional code $\mathcal{C}$ with rate k/n as a submodule of $\mathbb{F}^n[z]$ (see [5], [12], [13]) that can be described as

$$\mathcal{C} = \left\{ \boldsymbol{v}(z) \in \mathbb{F}^n[z] | \boldsymbol{v}(z) = G(z)\boldsymbol{u}(z) \quad \text{with} \quad \boldsymbol{u}(z) \in \mathbb{F}^k[z] \right\}$$

where $G(z)$ is a $n \times k$ polynomial matrix called a **generator matrix** for $\mathcal{C}$, $\boldsymbol{u}(z)$ is the **information vector** and $\boldsymbol{v}(z)$ is the **code vector** or **codeword**.

We define the **degree** of a convolutional code $\mathcal{C}$, and we denote it by δ, as the maximum of the degrees of the determinants of the $k \times k$ sub-matrices of any generator matrix of $\mathcal{C}$. Then we say that $\mathcal{C}$ is an (n, k, δ) convolutional code [11].

In case the convolutional code $\mathcal{C}$ is also observable (see, e.g., [12], [15]) then $\mathcal{C}$ can be equivalently described through a parity check matrix. In other words, there exists in this case an $(n-k) \times n$ full rank polynomial matrix $H(z)$ such that

$$\mathcal{C} = \left\{ \boldsymbol{v}(z) \in \mathbb{F}^n[z] \;\; | \;\; H(z)\boldsymbol{v}(z) = \mathbf{0} \in \mathbb{F}^{n-k}[z] \right\}.$$

If we write $\boldsymbol{v}(z) = \mathbf{v}_0 + \mathbf{v}_1 z + \ldots + \mathbf{v}_l z^l$ (with $l \geq 0$) and we represent $H(z)$ as a matrix polynomial

$$H(z) = H_0 + H_1 z + \ldots + H_\nu z^\nu.$$

we can expand the kernel representation in the following way

$$\begin{bmatrix} H_0 & & & & & \\ \vdots & \ddots & & & & \\ H_\nu & \ldots & H_0 & & & \\ & \ddots & & \ddots & & \\ & & H_\nu & \ldots & H_0 & \\ & & & \ddots & \vdots & \\ & & & & H_\nu & \end{bmatrix} \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_l \end{bmatrix} = \mathbf{0}. \tag{1}$$

An important distance measure for convolutional codes is the **free distance**:

$$d_{\text{free}}(\mathcal{C}) := \min \left\{ \text{wt}(\boldsymbol{v}(z)) \;\; | \;\; \boldsymbol{v}(z) \in \mathcal{C} \quad \text{and} \quad \boldsymbol{v}(z) \neq 0 \right\}.$$

The following lemma shows the importance of the free distance as a performance measure of a code used over the erasure channel.

Lemma 2.1: If $\mathcal{C}$ is a convolutional code with free distance $d := d_{\text{free}}$ and if during transmission at most $d-1$ erasures occur then these erasures can be uniquely decoded. Moreover, there exist patterns of d erasures which cannot be uniquely decoded.

Proof: Let $\boldsymbol{v}(z) = \mathbf{v}_0 + \mathbf{v}_1 z + \ldots + \mathbf{v}_l z^l$ be a received vector with $d-1$ symbols erased. Let the erasures be in positions $i_1, \ldots, i_{d-1}$. The homogeneous system (1) of $(\nu+l+1)(n-k)$ equations with $(l+1)n$ unknowns can be changed into an equivalent nonhomogeneous system

$$\hat{H} \begin{bmatrix} v_{i_1} \\ v_{i_2} \\ \vdots \\ v_{i_{d-1}} \end{bmatrix} = \mathbf{b}$$

of $(\nu + l + 1)(n - k)$ equations with $d - 1$ unknowns $v_{i_1}, \ldots, v_{i_{d-1}}$.

This nonhomogeneous system has a solution, because of the assumption that the channel allows only erasures. In addition the columns of the system matrix are linearly independent, because $d = d_{\text{free}}(\mathcal{C})$, so the matrix $\hat{H}$ is full column rank. It follows from these two facts that the solution must be unique. ∎

Rosenthal and Smarandache [14] showed that an (n, k, δ) convolutional code has a free distance upper bounded by

$$d_{\text{free}}(\mathcal{C}) \leq (n-k) \left(\left\lfloor \frac{\delta}{k} \right\rfloor + 1 \right) + \delta + 1. \tag{2}$$

This bound is known as the **generalized Singleton bound** [14] since it generalizes in a natural way the Singleton bound for block codes. Analogously, we say that an (n, k, δ) code is a **maximum distance separable** convolutional code (MDS) [14] if its free distance achieves the generalized Singleton bound.

Another local distance measure, important as well for decoding and related with the previous one, is the **column distance** [8], $d_j^c(\mathcal{C})$, given by the expression

$$d_j^c(\mathcal{C}) = \min \left\{ \text{wt}(\mathbf{v}_{[0,j]}(z)) \;\; | \;\; \boldsymbol{v}(z) \in \mathcal{C} \;\; \text{and} \;\; \mathbf{v}_0 \neq 0 \right\}$$

where $\mathbf{v}_{[0,j]}(z) = \mathbf{v}_0 + \mathbf{v}_1 z + \ldots + \mathbf{v}_j z^j$ represents the jth truncation of the codeword $\boldsymbol{v}(z) \in \mathcal{C}$. It is related with the $d_{\text{free}}(\mathcal{C})$ in the following way

$$d_{\text{free}}(\mathcal{C}) = \lim_{j\to\infty} d_j^c(\mathcal{C}). \tag{3}$$

The j-th column distance is then upper bounded by

$$d_j^c(\mathcal{C}) \leq (n-k)(j+1)+1 \tag{4}$$

and the maximality of any of the column distances implies the maximality of all the previous ones, that is, if $d_j^c(\mathcal{C}) = (n-k)(j+1)+1$ for some j, then $d_i^c(\mathcal{C}) = (n-k)(i+1)+1$ for $i \leq j$, see [4], [6]. The $(m+1)$-tuple $(d_0^c(\mathcal{C}), d_1^c(\mathcal{C}), \ldots, d_m^c(\mathcal{C}))$ is called the **column distance profile** of the code [8].

Since no column distance can achieve a value greater than the generalized Singleton bound, the largest integer for which that bound can be attained is

$$L = \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lfloor \frac{\delta}{n-k} \right\rfloor. \tag{5}$$

An (n,k,δ) convolutional code $\mathcal{C}$ is **maximum distance profile** (MDP) [4], [6], if $d_L^c(\mathcal{C}) = (n-k)(L+1)+1$. In this case, every $d_j^c(\mathcal{C})$ for $j \leq L$ is maximal, so we can say that the column distances of MDP codes increase as rapidly as possible for as long as possible.

In order to characterize the column distances as well as MDP codes algebraically assume the parity check matrix is given as $H(z) = \sum_{i=0}^{\nu} H_i z^i$. For each $j > \nu$ define $H_j = 0$ and define:

$$\mathcal{H}_j = \begin{bmatrix} H_0 & & & \\ H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \\ H_j & H_{j-1} & \cdots & H_0 \end{bmatrix} \in \mathbb{F}^{(j+1)(n-k)\times(j+1)n}. \tag{6}$$

Then we have:

Theorem 2.2: ([4, Proposition 2.1]) Let $d \in \mathbb{N}$. Then the following properties are equivalent.

(a) $d_j^c = d$;

(b) none of the first n columns of $\mathcal{H}_j$ is contained in the span of any other $d-2$ columns and one of the first n columns of $\mathcal{H}_j$ is in the span of some other $d-1$ columns of that matrix.

As a consequence we have the algebraic characterization of MDP convolutional codes:

Theorem 2.3: ([6, Theorem 3.1]) The j-th column distance attains the maximum value

$$d_j^c = (n-k)(j+1)+1, \tag{7}$$

if and only if, every $(j+1)(n-k)\times(j+1)(n-k)$ full-size minor of $\mathcal{H}_j$ formed from the columns with indices $1 \leq i_1 < \cdots < i_{(j+1)(n-k)}$, where $i_{s(n-k)} \leq sn$ for $s = 1, \ldots, j$, is nonzero.

In particular when $j = L$, then $H(z)$ represents an MDP code, if and only if, every $(L+1)(n-k)\times(L+1)(n-k)$ full-size minor of $\mathcal{H}_L$ formed from the columns with indices $1 \leq i_1 < \cdots < i_{(L+1)(n-k)}$, where $i_{s(n-k)} \leq sn$ for $s = 1, \ldots, L$, is nonzero.

MDP convolutional codes can be thought to be like an MDS block code within windows of size $(L+1)n$. The nonsingular full-size minors property given in the previous theorem ensures that if we truncate a codeword at iterations up to L it will have weight higher or equal than the bound (7).

III. Decoding over an erasure channel

Let us suppose that we use an MDP convolutional code $\mathcal{C}$ to transmit over an erasure channel. Then we can state the following result.

Theorem 3.1: Let $\mathcal{C}$ be an (n,k,δ) MDP convolutional code. If in any sliding window of length $(L+1)n$ at most $(L+1)(n-k)$ erasures occur then we can recover the whole sequence.

Proof: Assume that we have been able to correctly decode up to an instant $t-1$. Then we have the following homogeneous system :

$$\begin{bmatrix} H_\nu & H_{\nu-1} & \ldots & H_0 & & & \\ & H_\nu & \ldots & H_1 & \ddots & & \\ & & \ddots & & & H_0 & \\ & & & H_L & \ldots & H_1 & H_0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_{t-\nu} \\ \vdots \\ \mathbf{v}_{t-1} \\ \star \\ \star \\ \vdots \\ \star \end{bmatrix} = 0 \tag{8}$$

where $\star$ takes the place of a vector that had some of the components erased. Let the positions of the erased field elements be $i_1, \ldots, i_e$, $e \leq (n-k)(L+1)$, where $i_1, \ldots, i_s$, $s \leq n$, are the erasures occurring in the first erased n-vector. We can compute the syndrome and get a nonhomogeneous system with $(L+1)(n-k)$ equations and e, at most $(L+1)(n-k)$, variables.

We claim that there is an extension $\{\tilde{\mathbf{v}}_t, \ldots, \tilde{\mathbf{v}}_{t+L}\}$ such that the vector $(\mathbf{v}_{t-\nu} \ \ldots \ \mathbf{v}_{t-1} \ \tilde{\mathbf{v}}_t, \ldots, \tilde{\mathbf{v}}_{t+L})$ is a codeword and such that $\tilde{\mathbf{v}}_t$ is unique.

Indeed, we know that a solution of the system exists since we assumed only erasures occur. To prove the uniqueness of $\tilde{\mathbf{v}}_t$, or equivalently, of the erased elements $\tilde{v}_{i_1}, \ldots, \tilde{v}_{i_s}$, let us suppose there exist two such good extensions $\{\tilde{\mathbf{v}}_t, \ldots, \tilde{\mathbf{v}}_{t+L}\}$ and $\{\tilde{\tilde{\mathbf{v}}}_t, \ldots, \tilde{\tilde{\mathbf{v}}}_{t+L}\}$. Let $\mathbf{h}_{i_1}, \ldots, \mathbf{h}_{i_e}$, be the column vectors of the sliding parity-check matrix in (8) which correspond to the erasure elements. We have:

$$\tilde{v}_{i_1}\mathbf{h}_{i_1} + \ldots \tilde{v}_{i_s}\mathbf{h}_{i_s} + \ldots + \tilde{v}_{i_e}\mathbf{h}_{i_e} = \tilde{\mathbf{b}}$$

and

$$\tilde{\tilde{v}}_{i_1}\mathbf{h}_{i_1} + \ldots + \tilde{\tilde{v}}_{i_s}\mathbf{h}_{i_s} + \ldots + \tilde{\tilde{v}}_{i_e}\mathbf{h}_{i_e} = \tilde{\tilde{\mathbf{b}}},$$

where the vectors $\tilde{\mathbf{b}}$ and $\tilde{\tilde{\mathbf{b}}}$ correspond to the known part of the system. Subtracting these equations and observing that $\tilde{\mathbf{b}} = \tilde{\tilde{\mathbf{b}}}$, we obtain:

$$(\tilde{v}_{i_1} - \tilde{\tilde{v}}_{i_1})\mathbf{h}_{i_1} + \ldots + (\tilde{v}_{i_s} - \tilde{\tilde{v}}_{i_s})\mathbf{h}_{i_s} + \ldots + (\tilde{v}_{i_e} - \tilde{\tilde{v}}_{i_e})\mathbf{h}_{i_e} = 0.$$

Using Theorem 2.2 for $j = L$, and using that the code is MDP, so $d_L^c = (L+1)(n-k)+1$, we obtain by part (b) that the system is full rank and necessarily,

$$\tilde{v}_{i_1} - \tilde{\tilde{v}}_{i_1} = 0, \ \ldots, \ \tilde{v}_{i_s} - \tilde{\tilde{v}}_{i_s} = 0,$$

This concludes the proof of our claim.

In order to find the value of this unique vector, we solve the full column rank system, find a solution and retain the part which is unique. Then we slide n bits to the next $n(L+1)$ window and proceed as above.

■

A. Examples and Remarks

Remark 3.2: The decoding algorithm requires only simple linear algebra. For every $(n-k)$ erasures a matrix of size at most $(L+1)(n-k)$ has to be inverted over the base field $\mathbb{F}$. This is easily achieved even over fairly large fields.

In addition one should notice that for a rate $\frac{k}{n}$ MDP convolutional code, $100 \cdot \frac{n-k}{n}$ percent of the erasures can be corrected.

Remark 3.3: Theorem 3.1 is optimal in a certain sense: One can show that for any (n, k, δ) code there exist patterns of $(L+2)(n-k)$ erasures in a sliding window of length $(L+2)n$ which cannot be uniquely decoded.

The following illustrative example compares the size of a particular MDP convolutional code with an MDS block code which would perform similarly.

Example 3.4: Let us take a $(2, 1, 50)$ MDP convolutional code to decode over an erasure channel. In this case the decoding can be completed if in any sliding window of length 202 there are not more than 101 erasures; 50% of the erasures can be recovered.

The MDS block code which achieves a comparable performance is a $[200, 100]$ MDS block code. In a block of 200 symbols we can recover 100 erasures, that is again 50%.

Remark 3.5: It has been noticed that the parameter L gives us an upper bound on the length of the window we can take to correct, but it should be noticed as well that the property of Theorem 2.3 holds for every $j < L$. This means that we can take smaller windows to set our systems (the size will be conveniently decided by the distribution of the erasures in the sequence). Then in a window of size $(j+1)n$ symbols we can recover at most $(j+1)(n-k)$ erasures.

This property allows us to recover the erasures in situations where the MDS block codes cannot do it. For example, assume that we have been able to correctly decode up to an instant t and then it comes a block of 200 symbols where 2 bursts of 60 erasures occur separated by a block of 80 clean symbols, and after it, clean symbols again.

$$\overbrace{\star\star\ldots\star\star}^{60} v_{61} v_{62} \ldots v_{140} \overbrace{\star\star\ldots\star\star}^{60} v_{201} v_{202} \ldots$$

In this situation 120 erasures happen in a block of 200 symbols and the MDS block code is not able to recover them. In the block code situation one has to skip the whole block losing that information, and go on with the decoding.

However, the MDP convolutional code can deal with this situation. Let us set a 120 symbols length window; in these windows we can correct up to 60 erasures. We can take 100 previous decoded symbols, then set a window with the first 60 erasures and 60 more clean symbols. In this way we can recover the first block of erasures. Then we can slide through the received sequence with this 120 symbols window until we set the rest of the erasures in the same way.

$$v_{40} \ldots v_{140} \overbrace{\star\star\ldots\star\star}^{60} v_{201} v_{202} \ldots v_{260}$$

After this we have correctly decoded the sequence.

Remark 3.6: Another advantage to remark is related to the storage and to the field size required to construct the codes. In the example, we propose we have a $[200, 100]$ MDS block code. If we take, for example, a Reed-Solomon code (one of the most widely used MDS block codes) then we need to store the 200 roots of a 200 degree polynomial to set the code. That is, we need at least 200 field elements.

However, to set the $(2, 1, 50)$ MDP convolutional code we need to store the coefficients of 2 polynomials of degree 50, that is at most 100 different elements.

Nevertheless there are some disadvantages. On the one hand, the storage and the field size are smaller, but on the other hand, there are not direct constructions for the case of MDP convolutional codes. This is still an open problem.

The construction of MDP convolutional codes has been developed somewhat [7], however there exists still no efficient algorithm to construct this class of codes. In relation to this problem, special type of matrices called *superregular matrices* proved to be relevant during this study and this topic has become of main importance when trying to construct MDP convolutional codes [6], [9].

If we denote by $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_r}$ the $r \times r$ submatrix obtained from a matrix $T \in \mathbb{F}^{n\times n}$ by taking the rows with indices $i_1, \ldots, i_r$ and the columns with indices $j_1, \ldots, j_r$, then we can define a superregular matrix as follows.

Definition 3.7: [6] A lower triangular Toepliz matrix T

$$T = \begin{bmatrix} t_1 & 0 & \ldots & 0 \\ t_2 & t_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ t_n & \ldots & t_2 & t_1 \end{bmatrix} \in \mathbb{F}^{n\times n} \tag{9}$$

is said to be superregular if $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_r}$ is nonsingular for all $1 \le r \le n$ and all indices $1 \le i_1 < \ldots < i_r \le n$, $1 \le j_1 < \ldots < j_r \le n$ which satisfy $j_s \le i_s$ for $s = 1, \ldots, r$. The submatrices obtained by picking such indices are called the *proper submatrices* and their determinants the *proper minors* of T.

Unfortunately, the characterization or construction of these matrices is a hard problem and more research is needed in this direction in order to come up with a construction for MDP

convolutional codes. It was however shown in [6] that over an infinite field MDP convolutional codes form a generic set. As a result we can expect that over a very large field (e.g. $\mathbb{F}_q = \mathbb{F}_{2^{1000}}$) a randomly constructed convolutional code is MDP.

IV. Decoding with the help of the generator matrix

In this section we explain how the use of the generator matrix of the MDP code can make our decoding process more efficient and faster.

We know that the encoding process is represented by $G(z)\boldsymbol{u}(z) = \boldsymbol{v}(z)$, so the idea is to use this relation to recover directly the original message $\boldsymbol{u}(z)$ instead of computing first the code sequence and then decode it into the original sequence $\boldsymbol{u}(z)$, as we did before when working with the parity check matrix.

In an analogous way to the parity check matrix we can expand the generator matrix into

$$G(z) = G_0 + G_1 z + \ldots + G_m z^m$$

and define $\mathcal{G}_j$ as

$$\mathcal{G}_j = \begin{bmatrix} G_0 & & & \\ G_1 & G_0 & & \\ \vdots & & \ddots & \\ G_j & G_{j-1} & \ldots & G_0 \end{bmatrix}. \tag{10}$$

Then the equivalences in the following theorem give us the properties to improve the decoding algorithm.

Theorem 4.1: ([4, Theorem 2.4]) Let $\mathcal{H}_j$ and $\mathcal{G}_j$ be as in (6) and (10). Then the following are equivalent:

1) $d_j^c = (n-k)(j+1)+1$
2) every $(j+1)k \times (j+1)k$ full-size minor of $\mathcal{G}_j^T$ formed from the columns with indices $1 \leq t_1 < \ldots < t_{(j+1)k}$, where $t_{sk+1} > sn$ for $s = 1, \ldots, j$, is nonzero.

One notices that for an MDP convolutional code the maximum size of the matrix $\mathcal{G}_j$ we can construct is again given by the parameter L. This tells us that the maximum number of original symbols we are able to recover in one time is $(L+1)k$. Since in any sliding window of length $(L+1)n$ not more than $(L+1)(n-k)$ erasures occur we can set a full rank system with at least $(L+1)k$ equations to recover the $(L+1)k$ symbols of the original sequence. We will leave the details to the reader.

V. Conclusion

In this paper, we propose MDP convolutional codes as an alternative to block codes when decoding over an erasure channel. We have seen that the step-by-step-MDS property of the MDP codes lets us recover the maximum number of erasures at every step. Even over large field sizes the complexity of decoding is polynomial for a fixed window size since the decoding algorithm requires the solving of some linear system only. Moreover, the sliding window property allows us to adapt the decoding process to the distribution of the erasures in the sequence. We have shown how the possibility of taking smaller windows lets us recover erasures that the block codes cannot recover.

Acknowledgments

We would like to thank Martin Haenggi who explained us the distribution of packet sizes when transmitting files over the Internet.

References

[1] M. A. Epstein. Algebraic decoding for a binary erasure channel. Technical Report 340, Massachusetts Institute of Technology, March 1958. Reprinted from the 1958 IRE National Convention Record, Part 4.

[2] S. Fashandi, S. O. Gharan, and A. K. Khandani. Coding over an erasure channel with a large alphabet size, 2008. arXiv:0805.4440.

[3] M. Arai, A. Yamamoto, A. Yamaguchi, S. Fukumoto and K. Iwasaki. Analysis of Using Convolutional Codes to Recover Packet Losses over Burst Erasure Channel. *Proc. of the 2001 Pacific Rim International Symposium on Dependable Computing*, 258 - 265, 2001.

[4] H. Gluesing-Luerssen, J. Rosenthal, and R. Smarandache. Strongly MDS convolutional codes. *IEEE Trans. Inform. Theory*, 52(2):584–598, 2006.

[5] H. Gluesing-Luerssen and F.-L. Tsang. A matrix ring description for cyclic convolutional codes. *Adv. Math. Commun.*, 2(1):55–81, 2008.

[6] R. Hutchinson, J. Rosenthal, and R. Smarandache. Convolutional codes with maximum distance profile. *Systems & Control Letters*, 54(1):53–63, 2005.

[7] R. Hutchinson, R. Smarandache, and J. Trumpf. On superregular matrices and MDP convolutional codes. *Linear Algebra Appl.*, 428(11-12):2585–2596, 2008.

[8] R. Johannesson and K. Sh. Zigangirov. *Fundamentals of Convolutional Coding*. IEEE Press, New York, 1999.

[9] G. Kéri. Types of superregular matrices and the number of n-arcs and complete n-arcs in $\mathrm{PG}(r, q)$. *J. Combin. Des.*, 14(5):363–390, 2006.

[10] S. Lin and D. J. Costello Jr. *Error Control Coding: Fundamentals and Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1983.

[11] R. J. McEliece. The algebraic theory of convolutional codes. In V. Pless and W.C. Huffman, editors, *Handbook of Coding Theory*, volume 1, pages 1065–1138. Elsevier Science Publishers, Amsterdam, The Netherlands, 1998.

[12] J. Rosenthal. Connections between linear systems and convolutional codes. In B. Marcus and J. Rosenthal, editors, *Codes, Systems and Graphical Models*, IMA Vol. 123, pages 39–66. Springer-Verlag, 2001.

[13] J. Rosenthal, J. M. Schumacher, and E. V. York. On behaviors and convolutional codes. *IEEE Trans. Inform. Theory*, 42(6, part 1):1881–1891, 1996.

[14] J. Rosenthal and R. Smarandache. Maximum distance separable convolutional codes. *Appl. Algebra Engrg. Comm. Comput.*, 10(1):15–32, 1999.

[15] J. Rosenthal and E. V. York. BCH convolutional codes. *IEEE Trans. Inform. Theory*, 45(6):1833–1844, 1999.