Construction of Convolutional Codes over large fields

Paulo Almeida, Diego Napp, Raquel Pinto, Paolo Vettori & Rita Simoes

University of Aveiro, Portugal

BIRS

October 15, 2015

Overview

1. Fast review of Roxana's presentation: Basics
2. Performance in streaming
3. Issues in Constructions

Let $\mathbb{F}$ be a finite field and $\mathbb{F}[D]$ the polynomial ring. Let $G \in \mathbb{F}^{k\times n}$ be a matrix

$$\ldots u_2,\ u_1,\ u_0 \xrightarrow{G} \ldots v_2 = u_2 G,\ v_1 = u_1 G,\ v_0 = u_0 G$$

represented in a polynomial fashion

$$\cdots + u_2 D^2 + u_1 D + u_0 \xrightarrow{G} \cdots + \underbrace{u_2 G}_{v_2} D^2 + \underbrace{u_1 G}_{v_1} D + \underbrace{u_0 G}_{v_0}$$

why not substitute G by $G(D) = G_0 + G_1 D + \cdots + G_s D^s$?

$$\ldots u_2 D^2 + u_1 D + u_0 \xrightarrow{G(D)} \ldots \underbrace{(u_2 G_0 + u_1 G_1 + u_0 G_2)}_{v_2} D^2 + \underbrace{(u_1 G_0 + u_0 G_1)}_{v_1} D + \underbrace{u_0 G_0}_{v_0}$$

- Block codes: $\mathcal{C} = \{uG\} = \mathrm{Im}_{\mathbb{F}} G \sim \{u(D)G\} = \mathrm{Im}_{\mathbb{F}} G(D)$
- Convolutional codes: $\mathcal{C} = \{u(D)G(D)\} = \mathrm{Im}_{\mathbb{F}[D]} G(D)$

Definition

A convolutional code $\mathcal{C}$ is a $\mathbb{F}[D]$-submodule of $\mathbb{F}^n[D]$.

Since $\mathbb{F}[D]$ is a principal ideal domain, $\mathcal{C}$ is free and there exists a matrix $G(D)$ called encoder whose rows form a bases for $\mathcal{C}$. If $\mathcal{C}$ has rank k then we say the $\mathcal{C}$ has rate k/n.

$$\mathcal{C} \quad = \quad \mathrm{Im}_{\mathbb{F}[D]} G(D) = \left\{ u(D)G(D) : \ u(D) \in \mathbb{F}^k[D] \right\}$$

Definition

A convolutional code $\mathcal{C}$ is a $\mathbb{F}[D]$-submodule of $\mathbb{F}^n[D]$.

Since $\mathbb{F}[D]$ is a principal ideal domain, $\mathcal{C}$ is free and there exists a matrix $G(D)$ called encoder whose rows form a bases for $\mathcal{C}$. If $\mathcal{C}$ has rank k then we say the $\mathcal{C}$ has rate k/n.

$$\mathcal{C} \quad = \quad \mathrm{Im}_{\mathbb{F}[D]} G(D) = \left\{ u(D)G(D) : \ u(D) \in \mathbb{F}^k[D] \right\}$$

Remark: We could consider rational functions $\mathbb{F}(D)$ (or even better, Laurent series $\mathbb{F}((D))$) instead of polynomials $\mathbb{F}[D]$ and define $\mathcal{C}$ as a $\mathbb{F}(D)$-subspace of $\mathbb{F}^n(D)$.

Definition

A convolutional code $\mathcal{C}$ is a $\mathbb{F}[D]$-submodule of $\mathbb{F}^n[D]$.

Since $\mathbb{F}[D]$ is a principal ideal domain, $\mathcal{C}$ is free and there exists a matrix $G(D)$ called encoder whose rows form a bases for $\mathcal{C}$. If $\mathcal{C}$ has rank k then we say the $\mathcal{C}$ has rate k/n.

$$\mathcal{C} \quad = \quad \mathrm{Im}_{\mathbb{F}[D]} G(D) = \left\{ u(D)G(D) : \ u(D) \in \mathbb{F}^k[D] \right\}$$

Remark: We could consider rational functions $\mathbb{F}(D)$ (or even better, Laurent series $\mathbb{F}((D))$) instead of polynomials $\mathbb{F}[D]$ and define $\mathcal{C}$ as a $\mathbb{F}(D)$-subspace of $\mathbb{F}^n(D)$.

Multidimensional convolutional codes

One can very naturally extend the ideas used to convolutional codes to 2-dimensional convolutional codes if the data in distributed in a plane.

Definition

A 2D convolutional code $\mathcal{C}$ is a $\mathbb{F}[D_1, D_2]$-submodule of $\mathbb{F}^n[D_1, D_2]$.

Two encoders $G(D), G'(D)$ generate the same code if there exist a unimodular (i.e. its determinant is in $\mathbb{F}$) matrix $U(D)$ such that $G(D) = U(D)G'(D)$.

- We assume $G(D)$ is in *row reduced form* with row degrees $\{\nu_1, \ldots, \nu_k\}$
- The set $\{\nu_1, \ldots, \nu_k\}$, called Forney indices, is the same for all reduced encoders $G(D)$ of $\mathcal{C}$.
- The degree (the size of the memory) is defined as

$$\delta = \sum_{i=1}^{k} \nu_i$$

.

- The degree δ is also equal to the largest degree of the full size minors of $G(D)$.

Two encoders $G(D), G'(D)$ generate the same code if there exist a unimodular (i.e. its determinant is in $\mathbb{F}$) matrix $U(D)$ such that $G(D) = U(D)G'(D)$.

- We assume $G(D)$ is in *row reduced form* with row degrees $\{\nu_1, \ldots, \nu_k\}$
- The set $\{\nu_1, \ldots, \nu_k\}$, called Forney indices, is the same for all reduced encoders $G(D)$ of $\mathcal{C}$.
- The degree (the size of the memory) is defined as

$$\delta = \sum_{i=1}^{k} \nu_i$$

.

- The degree δ is also equal to the largest degree of the full size minors of $G(D)$.

Remark

A block code is a convolutional code without memory ($\delta = 0$).

Block codes vs Convolutional codes

- In block coding it is normally considered n and k large.
- Convolutional codes are tipically studied for n and k small and fixed ($n = 2$ and $k = 1$ is common) and for several values of δ.
- Roughly speaking: What matters in block codes is the block length and what matters for convolutional codes is the degree.

Block codes vs Convolutional codes

- In block coding it is normally considered n and k large.
- Convolutional codes are tipically studied for n and k small and fixed ($n = 2$ and $k = 1$ is common) and for several values of δ.
- Roughly speaking: What matters in block codes is the block length and what matters for convolutional codes is the degree.

Convolutional codes

- Decoding over the symmetric channel is difficult.
- The field is typically $\mathbb{F}_2$. The degree cannot be too large so that the Viterbi decoding algorithm is efficient.
- In [Tomas, Rosenthal, Smarandache 2012]:
  - Decoding over the erasure channel is *easy*.
  - Viterbi is not needed, just linear algebra.
- Codes with large field sizes $|\mathbb{F}|$ and degrees δ perform very well.

Parity-check matrices

Definition

$\mathcal{C}$ is called observable, basic or non-catastrophic if one and hence every generator matrix $G(D)$ of $\mathcal{C}$ is prime, i.e., admits a (polynomial) right inverse.

If $\mathcal{C}$ is observable then there exists a matrix $H(D) \in \mathbb{F}^{n-k\times n}$, called parity-check matrix such that

$$\begin{aligned}\mathcal{C} &= \mathrm{Im}_{\mathbb{F}[D]} G(D) = \left\{ u(D)G(D) : \ u(D) \in \mathbb{F}^k[D] \right\} \\ &= \ker_{\mathbb{F}[D]} H(D) = \left\{ v(D) \in \mathbb{F}^n[D] : \ H(D)v(D)^\top = 0 \right\}\end{aligned}$$

The Hamming weight of a polynomial vector

$$v(D) = \sum_{i \in \mathbb{N}} v_i D^i = v_0 + v_1 D + v_2 D^2 + \cdots + v_\nu D^\nu \in \mathbb{F}[D]^n,$$

defined as

$$\mathrm{wt}(v(D)) = \sum_{i=0}^{\nu} \mathrm{wt}(v_i).$$

Definition

The free distance of a convolutional code $\mathcal{C}$ is given by,

$$d_{\mathrm{free}}(\mathcal{C}) \quad = \quad \min \{\mathrm{wt}(v(D)) \quad | \quad v(D) \in \mathcal{C} \quad \text{and} \quad v(D) \neq 0\}$$

Theorem

Rosenthal and Smarandache (1999) showed that the free distance of convolutional code of rate k/n and degree δ must be upper bounded by

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

This bound was called the generalized Singleton bound since it generalizes in a natural way the Singleton bound for block codes (when $\delta = 0$). A code achieving (1) is called Maximum Distance Separable (MDS).

Theorem

Rosenthal and Smarandache (1999) showed that the free distance of convolutional code of rate k/n and degree δ must be upper bounded by

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

This bound was called the generalized Singleton bound since it generalizes in a natural way the Singleton bound for block codes (when $\delta = 0$). A code achieving (1) is called Maximum Distance Separable (MDS).

Theorem [Climent, Napp, Perea, Pinto (submitted)]

Let $\mathcal{C}$ be a 2D convolutional code of rate k/n and degree δ. Then

$$\text{dist}(\mathcal{C}) \leq \frac{(\lfloor \delta/k \rfloor + 1)(\lfloor \delta/k \rfloor + 2)}{2} n - k - \delta + k\lfloor \delta/k \rfloor + 1$$

This bound is called the 2D generalized Singleton bound and constructions achieving this bound were given (based on circulant Cauchy matrices).

Definition

Another important distance measure for a convolutional code is the jth column distance $d_j^c(\mathcal{C})$, (introduced by Costello), given by

$$d_j^c(\mathcal{C}) \;=\; \min\left\{\mathrm{wt}(v_{[0,j]}(D)) \;\mid\; v(D)\in\mathcal{C} \;\text{ and }\; v_0\neq 0\right\}$$

where $v_{[0,j]}(D) = v_0 + v_1 D + \ldots + v_j D^j$ represents the j-th truncation of the codeword $v(D)\in\mathcal{C}$. If $H(D) = H_0 + H_1 D + H_2 D^2 + \cdots + H_\nu D^\nu$ is a parity-check of $\mathcal{C}$.

$$d_j^c(\mathcal{C}) \;=\; \min\left\{\mathrm{wt}((v_0,\ldots,v_j)) \;\mid\; (v_0,\ldots,v_j)\ \textit{satisfies}\ (2),\ v_0\neq 0\right\}$$

$$\underbrace{\begin{pmatrix} H_0 & & & \\ H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \\ H_j & H_{j-1} & \cdots & H_0 \end{pmatrix}}_{(j+1)(n-k)\times(j+1)n} \begin{pmatrix} v_0 \\ \vdots \\ v_j \end{pmatrix} = 0 \qquad (2)$$

where $H_j = 0$, for $j > \nu$.

The column distances are invariants of the code, i.e., they do not depend on the choice of generator matrix and satisfy

$$d_0^c \leq d_1^c \leq \cdots \leq \lim_{j\to\infty} d_j^c(\mathcal{C}) = d_{\text{free}}(\mathcal{C}) \leq (n-k)\left(\left\lfloor \frac{\delta}{k} \right\rfloor + 1\right) + \delta + 1.$$

The j-th column distance is upper bounded as following

$$d_j^c(\mathcal{C}) \leq (n-k)(j+1) + 1, \tag{3}$$

The column distance can achieve bound (3) for $j \leq L$ where

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

and the earliest time instant that can achieve the Singleton bound is

$$M = \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lceil \frac{\delta}{n-k} \right\rceil.$$

Definition (Gluesing-Luerssen,Rosenthal,Smadandache (2006))

A convolutional code $\mathcal{C}$ of rate k/n and degree δ with every $d_j^c(\mathcal{C})$ maximal, for each $j \leq L$ is said to have a maximum distance profile (MDP), i.e., if

$$d_j^c = (n-k)(j+1)+1, \quad \text{for } j = 0, \ldots, L.$$

And it is called strongly MDS (sMDS) if it is MDS at time M.

Remark

MDS $\nRightarrow$ sMDP and sMDS $\nLeftarrow$ MDP

Remark

When $(n-k)|\delta$ (i.e. all Forney indices are equal), then

$$\text{MDS} \Leftrightarrow \text{sMDP}$$

Characterization of MDP

Let $\mathcal{C}$ be an (n, k, δ) convolutional code with parity-check matrix $H(D) = H_0 + H_1 D + \cdots + H_\nu D^\nu$. Then $\mathcal{C}$ is MDP if and only if every $(L+1)(n-k) \times (L+1)(n-k)$ full-size minors of the sliding parity-check matrix

$$H_j^c = \begin{pmatrix} H_0 & & & \\ H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \\ H_j & H_{j-1} & \cdots & H_0 \end{pmatrix} \in \mathbb{F}^{(j+1)(n-k)\times(j+1)n}, \tag{4}$$

with no zeros in the diagonal, is non-zero.

Remark

Note that the remaining full-size minors are trivially zero.

- Allen conjecture (1999) the existence of convolutional codes that are both sMDS and MDP when $k = 1$ and $n = 2$.
- Smarandache et. al (2001), provided the first concrete construction of MDS convolutional codes for all rates and degrees
- Rosenthal et. al (2005), provided a non-constructive proof (using algebraic geometry) of the existence of MDP convolutional codes.
- Gluessing-Luerssen et. al (2006), provided the first concrete construction of MDP convolutional codes for all parameters. And conjectured the existence of sMDS convolutional codes that are also MDP (proved in the case $(n-k)|\delta$).
- Hutchinson (2008) gave a non-constructive proof of the existence of conv. codes both MDP and sMDS for all rates and degrees.
- Napp and Smarandache (to appear) provided the first concrete construction of convolutional codes that are both sMDS and MDP for all rates and degrees.

Performance over the erasure channel

Theorem

Let $\mathcal{C}$ be an (n, k, δ) convolutional code and $d_{j_0}^c$ the $j = j_0$ -th column distance. If in any sliding window of length $(j_0 + 1)n$ at most $d_{j_0}^c - 1$ erasures occur then we can recover completely the transmitted sequence.

Performance over the erasure channel

Theorem

Let $\mathcal{C}$ be an (n, k, δ) convolutional code and $d_{j_0}^c$ the $j = j_0$ -th column distance. If in any sliding window of length $(j_0 + 1)n$ at most $d_{j_0}^c - 1$ erasures occur then we can recover completely the transmitted sequence.

Remark

The best scenario happens when the convolutional code is MDP. In this case full error correction *from left to right* is possible as soon as the fraction of erasures is not more than $\frac{n-k}{n}$ in any sliding window of length $(L+1)n$. Recovery rate $R_\omega = \frac{(j+1)(n-k)}{(j+1)n}$

Performance over the erasure channel

Theorem

Let $\mathcal{C}$ be an (n, k, δ) convolutional code and $d_{j_0}^c$ the $j = j_0$ -th column distance. If in any sliding window of length $(j_0 + 1)n$ at most $d_{j_0}^c - 1$ erasures occur then we can recover completely the transmitted sequence.

Remark

The best scenario happens when the convolutional code is MDP. In this case full error correction *from left to right* is possible as soon as the fraction of erasures is not more than $\frac{n-k}{n}$ in any sliding window of length $(L+1)n$. Recovery rate $R_\omega = \frac{(j+1)(n-k)}{(j+1)n}$

Corollary

If in any sliding window of length $(L+1)n$ at most $(n-k)(L+1)$ erasures occur then an MDP can completely recover the transmitted sequence.

Example

$$\cdots vv\Big|\overbrace{\star\star\cdots\star\star}^{60}\overbrace{vvv\cdots vv}^{80}\overbrace{\star\star\cdot\star\star}^{60}vv\Big|vv\cdots$$

A $[202, 101]$ MDS blok code can correct 101 erasures in a window of 202 symbols (recovering rate $\frac{101}{202}$): $\Rightarrow$ cannot correct this window.

A $(2, 1, 50)$ MDP convolutional code has also 50% error capability. $(L+1)n = 101 \times 2 = 202$. Take a window of 120 symbols, correct and continue until you correct the whole window.
We have flexibility in choosing the size and position of the sliding window.

Example

$$\cdots vv\,|\overbrace{\star\star\cdots\star\star}^{60}\overbrace{vvv\cdots vv}^{80}\overbrace{\star\star\cdot\star\star}^{60}vv\,|\,vv\cdots$$

A $[202, 101]$ MDS blok code can correct 101 erasures in a window of 202 symbols (recovering rate $\frac{101}{202}$): $\Rightarrow$ cannot correct this window.

A $(2, 1, 50)$ MDP convolutional code has also 50% error capability. $(L+1)n = 101 \times 2 = 202$. Take a window of 120 symbols, correct and continue until you correct the whole window.
We have flexibility in choosing the size and position of the sliding window.

Getting lost in the decoding

$$\cdots vv\,|\overbrace{\star\star\cdots\star\star}^{22}\overbrace{vv\star\star vv\star\star\cdots vv\star\star}^{180(90)}|\overbrace{vvv\cdots vv}^{202}\overbrace{\star\star\cdot\star\star}^{60}vv\,|\,vv\cdots$$

MDP cannot correct it from left to right...

Reverse: From right to left

Let $\mathcal{C}$ with parity-check $H(D) = H_0 + H_1 D + \cdots H_\nu D^\nu$. Then $\overline{H}(D) = H_\nu + H_{\nu-1} D + \cdots + H_0 D^\nu$ defines a (reverse) conv. code $\overline{\mathcal{C}}$ with the property that

$$v_0 + v_1 D + \cdots v_s D^s \in \mathcal{C}$$

if and only if

$$v_s + v_{s-1} D + \cdots v_0 D^s \in \overline{\mathcal{C}}$$

A MDP convolutional code $\mathcal{C}$ if called reverse-MDP if $\overline{\mathcal{C}}$ is also MDP.

Reverse: From right to left

Let $\mathcal{C}$ with parity-check $H(D) = H_0 + H_1 D + \cdots H_\nu D^\nu$. Then $\overline{H}(D) = H_\nu + H_{\nu-1} D + \cdots + H_0 D^\nu$ defines a (reverse) conv. code $\overline{\mathcal{C}}$ with the property that

$$v_0 + v_1 D + \cdots v_s D^s \in \mathcal{C}$$

if and only if

$$v_s + v_{s-1} D + \cdots v_0 D^s \in \overline{\mathcal{C}}$$

A MDP convolutional code $\mathcal{C}$ if called reverse-MDP if $\overline{\mathcal{C}}$ is also MDP.

Let $\mathcal{C}$ be an (n, k, δ) convolutional code with parity-check matrix $H(D) = H_0 + H_1 D + \cdots + H_\nu D^\nu$. Then $\mathcal{C}$ is reverse- MDP if and only if every $(L+1)(n-k) \times (L+1)(n-k)$ full-size minors of

$$H_j^c = \begin{pmatrix} H_0 & & & \\ H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \\ H_j & H_{j-1} & \cdots & H_0 \end{pmatrix} \text{ and } \overline{H}_j^c = \begin{pmatrix} H_\nu & & & \\ H_{\nu-1} & H_\nu & & \\ \vdots & \vdots & \ddots & \\ H_{\nu-j} & H_{\nu-j+1} & \cdots & H_\nu \end{pmatrix} \quad (5)$$

with no zeros in the diagonal, is non-zero.

Example

$$\cdots vv\,|\overbrace{\star\star\cdots\star\star}^{22}\overbrace{vv\star\star vv\star\star\cdots vv\star\star}^{180(90)}|\overbrace{vvv\cdots vv}^{202}\overbrace{\star\star\cdot\star\star}^{60}vv|vv\cdots$$

reverse-MDP can correct it from right to left.

Getting again lost in the decoding

Consider a $[75, 50]$ MDS blok code with recovering rate $\frac{25}{75}$.
And a $(3, 2, 16)$ reverse-MDP convolutional code.

$$\ldots\star\star|\overbrace{\star\star\cdots\star\star}^{14}\overbrace{vv\cdots vv}^{21}\overbrace{\star\star\cdots\star\star}^{12}\overbrace{vv\cdots vv}^{28}|\overbrace{vv\cdots vv}^{19}\overbrace{\star\star\cdots\star\star}^{13}\overbrace{vv\cdots vv}^{30}\overbrace{\star\star\cdots\star\star}^{13}|\star\star\cdots$$

Neither one can correct it. The reverse-MDP cannot find guard space of at least $n\nu = 48$ correct symbols...we need to compute a guard space.

Definition

Let $\mathcal{C}$ be an (n, k, δ) convolutional code with parity-check matrix $H(D) = H_0 + H_1 D + \cdots + H_\nu D^\nu$. Then $\mathcal{C}$ is complete-MDP if and only if every $(L+1)(n-k) \times (L+1)(n-k)$ full-size minors of

$$\begin{pmatrix} H_m & \cdots & H_0 & & & \\ & H_m & \cdots & H_0 & & \\ & & \ddots & \vdots & \ddots & \\ & & & H_m & \cdots & H_0 \end{pmatrix} \in \mathbb{F}^{(L+1)(n-k)\times(\nu+L+1)n} \tag{6}$$

with no zeros in the diagonal, is non-zero.

Theorem (Computing a new guard space)

Given a code sequence from some complete MDP convolutional code. If in a window of size $(\nu + L + 1)n$ there are not more than $(L+1)(n-k)$ erasures, and if they are distributed in such a way that between position 1 and sn and between positions $(\nu + L + 1)n$ and $(\nu + L + 1)n - s(n-k)$ for $s = 0, 1, \ldots, L+1$, there are not more than $s(n-k)$ erasures, then full correction of all symbols in this interval will be possible.

In the previous example: Consider a $(3, 2, 16)$ complete-MDP convolutional code, that has also 50% error capability.

Getting again lost in the decoding

$$\ldots\star\star\,|\overbrace{\star\star\cdots\star\star}^{14}\overbrace{vv\cdots vv}^{21}\overbrace{\star\star\cdots\star\star}^{12}\overbrace{vv\cdots vv}^{28}\,|\overbrace{vv\cdots vv}^{19}\overbrace{\star\star\cdots\star\star}^{13}\overbrace{vv\cdots vv}^{30}\overbrace{\star\star\cdots}^{1}$$

Consider

$$\overbrace{vv\cdots vv}^{21}\overbrace{\star\star\cdots\star\star}^{12}\overbrace{vv\cdots vv}^{28}\,|\overbrace{vv\cdots vv}^{19}\overbrace{\star\star\cdots\star\star}^{13}\overbrace{vv\cdots vv}^{30}$$

Then, $(\nu + L + 1)n = (24 + 1 + 16)3 = 123$ there are not more than $(L + 1)(n - k) = 25$ and the distribution of the erasures satisfies the condition of the theorem: We can compute a guard space.

Constructions

H. Gluesing-Luerssen, J. Rosenthal, and R. Smarandache (2006)

Let $\mathcal{C} = \{v(D) \in \mathbb{F}((D))^n \mid H(D)v(D) = 0\}$

$$H(D) = \sum_{i=0}^{m} H_i D^i = \sum_{i=0}^{m} \begin{bmatrix} A_i & B_i \end{bmatrix} D^i = \begin{bmatrix} A(D) & B(D) \end{bmatrix} \in \mathbb{F}[D]^{(n-k)\times n},$$

where $m = \lceil \frac{\delta}{n-k} \rceil$. Assume in addition that A_0 is invertible and let

$$A(D)^{-1}B(D) = \sum_{i=0}^{\infty} P_i D^i \in \mathbb{F}((D))^{(n-k)\times k}$$

be the Laurent expansion of $A(D)^{-1}B(D)$ over the field $\mathbb{F}((D))$. Define

$$\widehat{H}_j^c = \begin{bmatrix} I_{(j+1)(n-k)} & P_j^c \end{bmatrix} \text{ with } P_j^c = \begin{pmatrix} P_0 & & & \\ P_1 & P_0 & & \\ \vdots & \vdots & \ddots & \\ P_j & P_{j-1} & \cdots & P_0 \end{pmatrix}.$$

Theorem cont.

Then, the following conditions are equivalent, for all $j \in \{1, \ldots, L\}$:

1. $d_j^c = (n-k)(j+1)+1$;, i.e., $\mathcal{C}$ is MDP;
2. every nontrivial $(n-k)(j+1) \times (n-k)(j+1)$ full-size minor of H_j^c is nonzero.
3. P_j^c is lower triangular LT-superregular;

The construction of MDP convolutional codes boils down to the construction of LT-superregular matrices.

LT-Superregular matrices

Definition [Gluesing-Luerssen,Rosenthal,Smadandache (2006)]

A lower triangular matrix

$$B = \begin{pmatrix} a_0 & & & \\ a_1 & a_0 & & \\ \vdots & \vdots & \ddots & \\ a_j & a_{j-1} & \cdots & a_0 \end{pmatrix} \tag{7}$$

is *LT-superregular* if all of its minors, with no zeros in the diagonal, are nonsingular.

Remark

Note that due to such a lower triangular configuration the remaining minors are necessarily zero.

Example

$$\beta^3+\beta+1=0 \Rightarrow \begin{pmatrix} 1 & & & & \\ \beta & 1 & & & \\ \beta^3 & \beta & 1 & & \\ \beta & \beta^3 & \beta & 1 & \\ 1 & \beta & \beta^3 & \beta & 1 \end{pmatrix} \in \mathbb{F}_{2^3}^{5\times 5} \text{ is LT-superregular}$$

Example

$$\epsilon^5+\epsilon^2+1=0 \Rightarrow \begin{pmatrix} 1 & & & & & & \\ \epsilon & 1 & & & & & \\ \epsilon^6 & \epsilon & 1 & & & & \\ \epsilon^9 & \epsilon^6 & \epsilon & 1 & & & \\ \epsilon^6 & \epsilon^9 & \epsilon^6 & \epsilon & 1 & & \\ \epsilon & \epsilon^6 & \epsilon^9 & \epsilon^6 & \epsilon & 1 & \\ 1 & \epsilon & \epsilon^6 & \epsilon^9 & \epsilon^6 & \epsilon & 1 \end{pmatrix} \in \mathbb{F}_{2^5}^{7\times 7} \text{ is}$$

LT-superregular

Remarks

- Construction of classes of LT-superregular matrices is very difficult due to their triangular configuration.
- Only two classes exist:

1. Rosenthal et al. (2006) presented the first construction. For any n there exists a prime number p such that

$$\begin{pmatrix} \binom{n}{0} & & & \\ \binom{n-1}{1} & \binom{n}{0} & & \\ \vdots & \ddots & \ddots & \\ \binom{n-1}{n-1} & \cdots & \binom{n-1}{1} & \binom{n}{0} \end{pmatrix} \in \mathbb{F}_p^{n \times n}$$

is LT-superregular. Bad news: Requires a field with very large characteristic.

Remarks

2. Almeida, Napp and Pinto (2013) first construction over any characteristic: Let α be a primitive element of a finite field $\mathbb{F}$ of characteristic p. If $|\mathbb{F}| \geq p^{2^M}$ then the following matrix

$$\begin{bmatrix} \alpha^{2^0} & & & & \\ \alpha^{2^1} & \alpha^{2^0} & & & \\ \alpha^{2^2} & \alpha^{2^1} & \alpha^{2^0} & & \\ \vdots & & \ddots & \ddots & \\ \alpha^{2^{M-1}} & \cdots & & \cdots & \alpha^{2^0} \end{bmatrix}.$$

is LT-superregular. Bad news: $|\mathbb{F}|$ very large.

Construction of Reverve-MDP

In order to construct (n, k, δ) reverse-MDP (for $(n-k)|\delta$) we need to construct

$$\begin{pmatrix} a_0 & & & \\ a_1 & a_0 & & \\ \vdots & \vdots & \ddots & \\ a_j & a_{j-1} & \cdots & a_0 \end{pmatrix} \text{ and } \begin{pmatrix} a_j & & & \\ a_{j-1} & a_0 & & \\ \vdots & \vdots & \ddots & \\ a_0 & a_1 & \cdots & a_j \end{pmatrix}$$

both LT-superregular.

Reverse-MDP

1. We do not have a characterization in terms of LT-superregular matrices when $(n-k) \nmid \delta$.
2. Although there are some clever ideas and several particular examples, only the construction of Almeida, Napp and Pinto (2013) gives a general construction of reverse superregular.

Complete-MDP

1. The existence of complete-MDP has not been proven.
2. No direct relation with LT-superregular matrices...a new notion of superregularity is needed.
3. We conjecture that the construction of Almeida, Napp and Pinto (2013) will generate complete-MDP but to-date no general construction of complete-MDP has given.
4. More efficient decoding algorithms.

Things to do

1. Better constructions of superregular matrices.
2. More efficient decoding algorithms.
3. Adapt these constructions to the requirements of the applications, e.g., delay (Ahmed, Ashish and et.al.).