ITA2017 - page 1

Monday, March 06, 2017 10:37 AM

Analog Codes & Good Frames

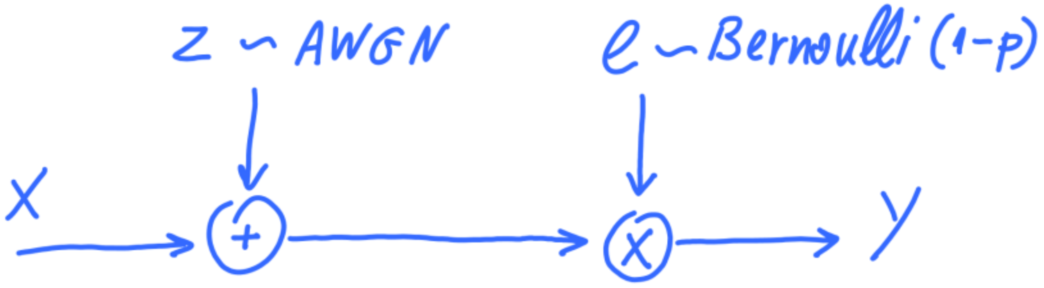
Rami Zamir, Tel Aviv University

Joint work with Marina Haikin, TAU
Matan Gavish, HUJI

Page 2

Monday, March 06, 2017 10:37 AM

Channel with Noise & Erasures

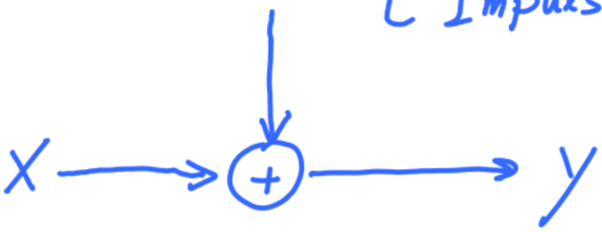


$$Y = \begin{cases} X + Z \,, & \text{w.p. } 1-p \\ 0 \;(\text{"erasure"}) \,, & \text{w.p. } p \end{cases}$$

[Tulino - Verdu - Caire - Shamai 2007]

* similar to an impulsive channel :

$$\tilde{Z} \sim \begin{cases} \text{AWGN} \,, & \text{w.p. } 1-p \\ \text{Impulse} \,, & \text{w.p. } p \end{cases}$$



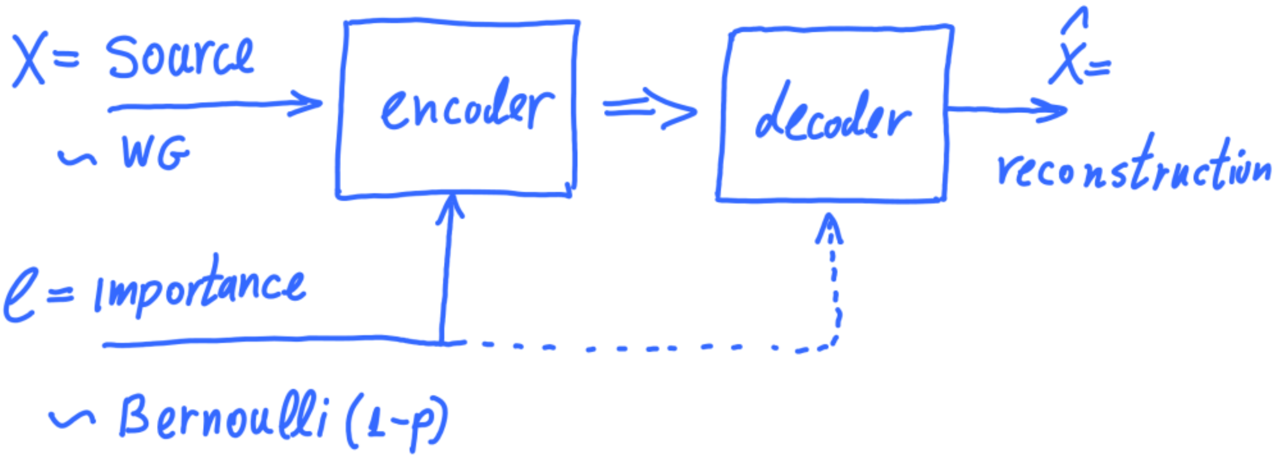
[Wolf 1983]

$$\Rightarrow \; C = (1-p) \cdot \frac{1}{2} \log(1 + \text{SNR}) \left[\frac{\text{bit}}{\text{channel use}}\right]$$

Page 3

Monday, March 06, 2017 10:37 AM

Lossy Source Coding with Erasures



$$d(x,\hat{x},e) = \begin{cases} (\hat{x}-x)^2, & \text{if } e=1 \\ 0, & \text{if } e=0 \ (\text{"erasure"}) \end{cases}$$

$$\Rightarrow \quad R(D) = (1-p)\cdot \frac{1}{2}\log\left(SDR\right) \quad \left[\frac{\text{bit}}{\text{source sample}}\right]$$

[Martinian - Wornell - Zamir 2006]

Page 4

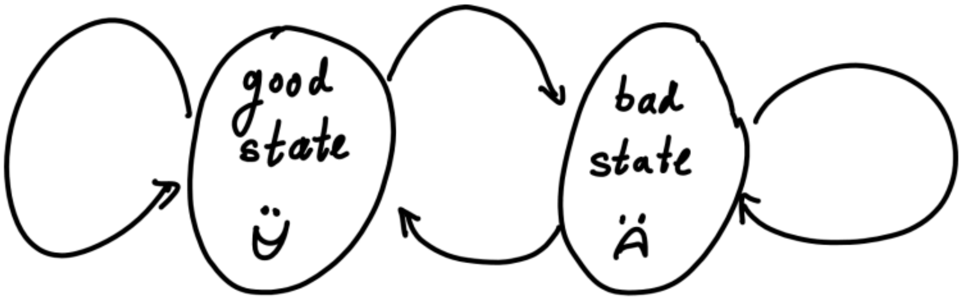
Monday, March 06, 2017 10:37 AM

Sounds simple, but is it easy to achieve ?...

1. practical (sub-optimal) decoders do not function well over state-varying channels

2. Effective coding dimension (against noise) = $(1-p) \cdot n$ is smaller than total block length = n

3. State with memory (runs of erasures)
   ⇒ n must be larger (to guarantee ergodicity)
   ⇒ adds complexity

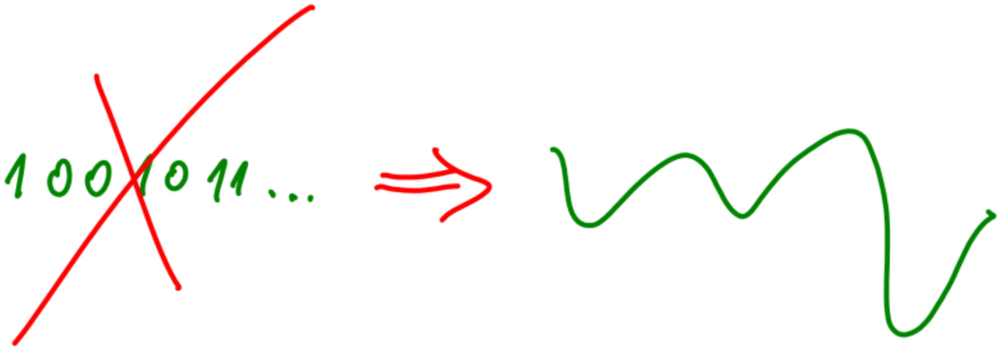


⇒ Goal : decouple noise & erasure protection !

Page 5

Monday, March 06, 2017 10:37 AM

Analog ("DFT") Code for Erasures



use redandancy in spectrum !

[Wolf 1983]

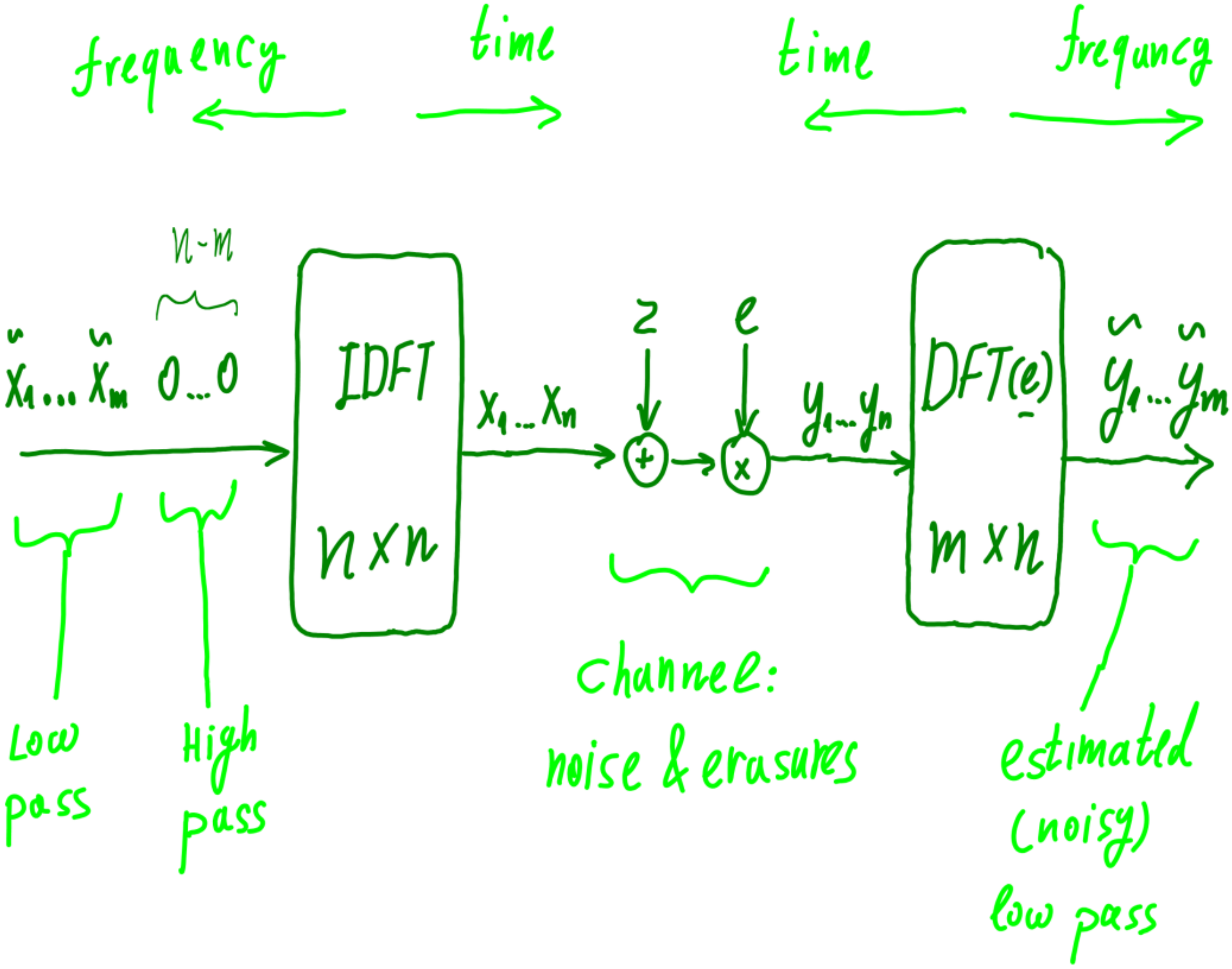
Page 6

Monday, March 06, 2017 10:37 AM

Analog ("DFT") Code for Erasures

Let $K \triangleq (1-p)\cdot n$ ($\approx$ # un-erased samples)

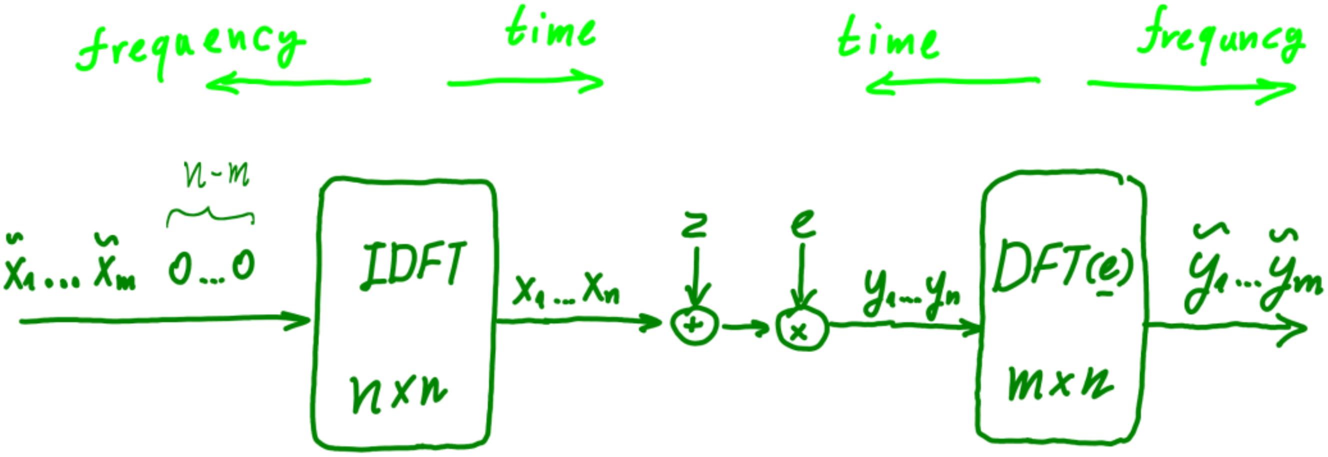
pick $1 \le m \le K$ (analog signal bandwidth)



Page 7

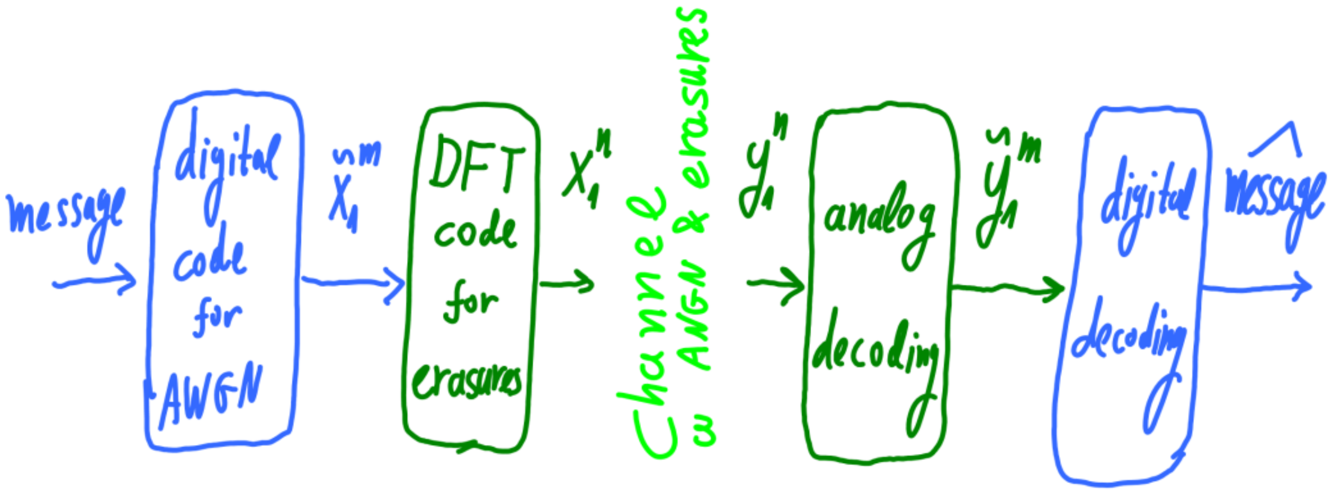
Monday, March 06, 2017 10:37 AM

Analog ("DFT") Code for Erasures



* Combined system:

digital code (for AWGN) analog code (for erasures):



Page 8

Monday, March 06, 2017 10:37 AM

Analog decoding:

Time-Frequency Inversion with Erasures

k un-erased outputs (in time)

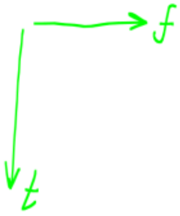
m inputs in frequency

$$\begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} = \overset{k\times n}{T} \cdot \overset{n\times n}{IDFT} \cdot \overset{n\times m}{F} \cdot \begin{pmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_m \end{pmatrix} + \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix}$$

k un-erased times

$$\begin{bmatrix} - 1 - & - & - \\ - & 1 & - \\ \vdots & & \ddots \\ - & - & 1 - \end{bmatrix}$$

$$\frac{1}{\sqrt{m}} \cdot \left\{ e^{\sqrt{-1}\,2\pi f \cdot t} \right\}_{f,t=1}^{n}$$



AWGN (σ^2)

m active frequencies

$$\begin{bmatrix} 1 & & & \\ & 1 & 0 & \\ & & \ddots & \\ & & & 1 \\ & 0 & & \end{bmatrix}$$

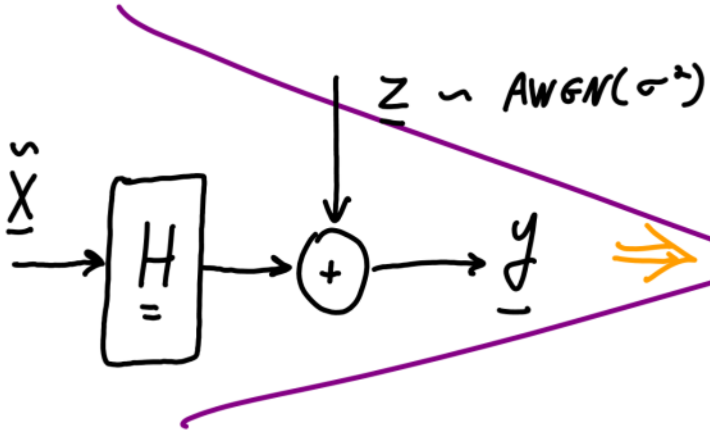
$\sqrt{m}$ factor ⇒
unit row norm of IDFT·F ⇒
frequency → time power preservation

Page 9

Monday, March 06, 2017 10:37 AM

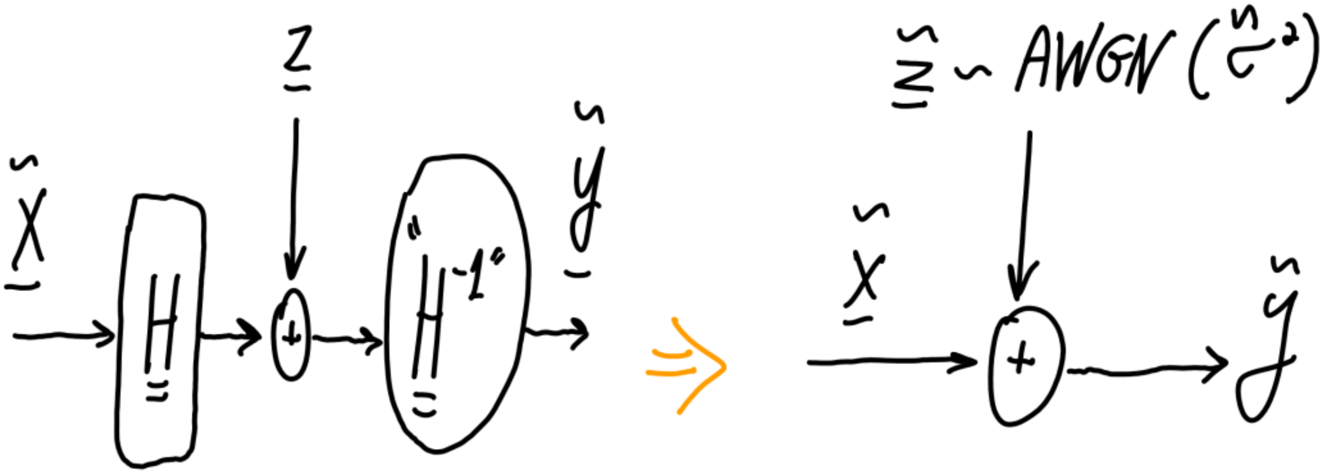
Time-Frequency Inversion with Erasures

* Linear channel:



$$\cancel{C = \max I(\underline{\tilde{x}}; \underline{y})}$$

* Equivalent AWGN Channel:



Page 10

Monday, March 06, 2017 10:37 AM

Time-Frequency Inversion with Erasures

un-erased outputs (in time)

inputs in frequency

$$\begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} = \overset{k\times n}{T} \cdot \overset{n\times n}{IDFT} \cdot \overset{n\times m}{F} \cdot \begin{pmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_m \end{pmatrix} + \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix}$$

k un-erased times

$$\begin{bmatrix} -1- & & & \\ & -1- & \ddots & \\ \vdots & & & \\ & & & 1- \end{bmatrix}$$

$$\frac{1}{\sqrt{m}} \cdot \left\{ e^{\sqrt{-1}\,2\pi f \cdot t} \right\}^{n}_{f,t=1}$$

m active frequencies

$$\begin{bmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots \\ & & & 1 \\ & 0 & \end{bmatrix}$$

⇒ Least-Squares estimate of $\underline{\hat{x}}$ (assume $k \geq m$):

$$\underline{\hat{y}} = \text{pseudo inverse}\Big(\underbrace{T \cdot IDFT \cdot F}_{\triangleq\, H \quad k\times m}\Big) \cdot \underline{y}$$

m dim

$$= \underset{m\times m}{\left(H^t H\right)^{-1}} \cdot \underset{m\times k}{H^t} \cdot \underline{y}$$

k dim

* if $m = k \leq n$ ⇒ perfect inverse ⇒ $\underline{\hat{y}} = H^{-1} \cdot \underline{y}$

* if $m \leq k = n$ ⇒ $T = I$ ⇒ Low Pass Filter ⇒ $\underline{\hat{y}} = F^t \cdot DFT \cdot \underline{y}$

Page 11

Monday, March 06, 2017 10:37 AM

Equivalent noise

Least-Squares estimate of $\underline{\tilde{x}}$ (assume $k \geq m$):

$$\underbrace{\underline{\tilde{y}}}_{m\text{-dim}} = \underbrace{(H^t \cdot H)^{-1}}_{m \times m} \cdot \underbrace{H^t}_{m \times k} \cdot \underbrace{\underline{y}}_{k\text{-dim}}, \qquad \text{where} \quad \underbrace{H}_{k \times m} \triangleq \underbrace{T}_{k \times n} \cdot \underbrace{IDFT}_{n \times n} \cdot \underbrace{F}_{n \times m}$$

$\Rightarrow$ equivalent noise: $\underline{\tilde{Z}} = (H^t H)^{-1} H^t \cdot \underline{Z}$

$\Rightarrow$ equivalent noise power: $\tilde{\sigma}^2 = \sigma^2 \cdot \frac{1}{m} \cdot \mathrm{trace}\,(H^t \cdot H)^{-1}$

Recall diagonalization of a symmetric matrix

$$H^t \cdot H = \mathcal{U}^t \cdot D \cdot \mathcal{U}, \qquad D = \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_m^2 \end{pmatrix}$$

unitary $\nearrow$

where $\lambda_1, \ldots, \lambda_m$ = singular values of H^t (same as non-zero singular values of H).

$\Rightarrow$ $\mathrm{trace}\,(H^t H)^{-1} = \mathrm{trace}\,(\mathcal{U}^t D^{-1} \mathcal{U}) = \sum_{i=1}^{m} \frac{1}{\lambda_i^2}$

$\Rightarrow$ equivalent noise power: $\tilde{\sigma}^2 = \sigma^2 \cdot \frac{1}{m} \cdot \sum_{i=1}^{m} \frac{1}{\lambda_i^2}$

Page 12

Monday, March 06, 2017 10:37 AM

Inversion - Amplification Lemma

If H is an $K \times m$ matrix, $m \leq K$, then

$$\frac{1}{m} \cdot \text{trace}\left\{ (H^t H)^{-1} \right\} \geq \frac{1}{\frac{1}{m} \cdot \text{trace}\left\{ H^t \cdot H \right\}}$$

with equality iff $H^t \cdot H$ is scaled identity

$\Longleftrightarrow$ rows of H^t are orthogonal and equi-norm.

Proof: Harmonic - arithmetic means inequality for singular values $\frac{1}{m} \cdot \sum_i \lambda_i^2 \geq 1 / \frac{1}{m} \sum_i 1/\lambda_i^2$.

Page 13

Monday, March 06, 2017 10:37 AM

Inversion - Amplification Lemma

If H is an $K \times m$ matrix, $m \leq K$, then

$$\frac{1}{m}\cdot \text{trace}\{(H^t H)^{-1}\} \geq \frac{1}{\frac{1}{m}\cdot \text{trace}\{H^t \cdot H\}}$$

with equality iff $H^t \cdot H$ is scaled identity

$\Longleftrightarrow$ rows of H^t are orthogonal and equi-norm.

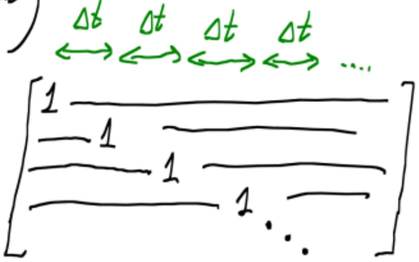
Remarks:

1. $\left|e^{j 2\pi \cdot f \cdot t}\right| = 1 \quad \forall f,t \Rightarrow \|\underbrace{\text{row}_f(F^t \cdot DFT \cdot T^t)}_{\substack{K \text{ elements} \\ 1/m \text{ normalization}}}\|^2 \underset{\forall f}{=} K/m$

   $\Rightarrow \frac{1}{m}\cdot \sum_{i=1}^{m} \lambda_i^2 = K/m$

2. $\text{row}_f(F^t \cdot DFT \cdot T^t) \perp \text{row}_{f'}(F^t \cdot DFT \cdot T^t)$

   iff T = uniform sampling =



$\therefore \quad \tilde{\sigma}^2 \geq \frac{K}{m}\cdot \sigma^2,$

equality iff the un-erased samples are uniform!

Page 14

Monday, March 06, 2017 10:37 AM

Random erasure pattern:
Typical / average noise amplification

Fix $p = K/n$ and $\beta \triangleq k/m$ $(\beta \geq 1)$

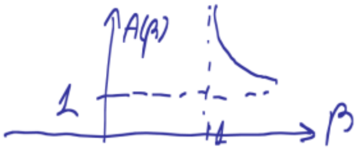
Observation: If $e_1, \dots, e_n \sim$ Bernoulli$(1-p)$ or uniform over all $\binom{n}{k}$ patterns, then the arithmetic-harmonic mean ratio of $H = T \cdot IDFT \cdot F$ has a typical asymptotic behavior:

$$\frac{1/m \sum_{i=1}^{m} 1/\lambda_i^2}{\left(1/m \sum_{i=1}^{m} \lambda_i^2\right)^{-1}} \xrightarrow[n \to \infty]{} \text{Constant} \triangleq A(\beta, p) \quad \text{a.s.}$$

∴ Capacity of analog code:

$$\hat{C} = \frac{m}{n} \cdot \frac{1}{2} \log(1 + \widetilde{SNR}) = \frac{1}{\beta} \cdot C\left(\frac{\beta \cdot SNR}{A(\beta)}\right)$$

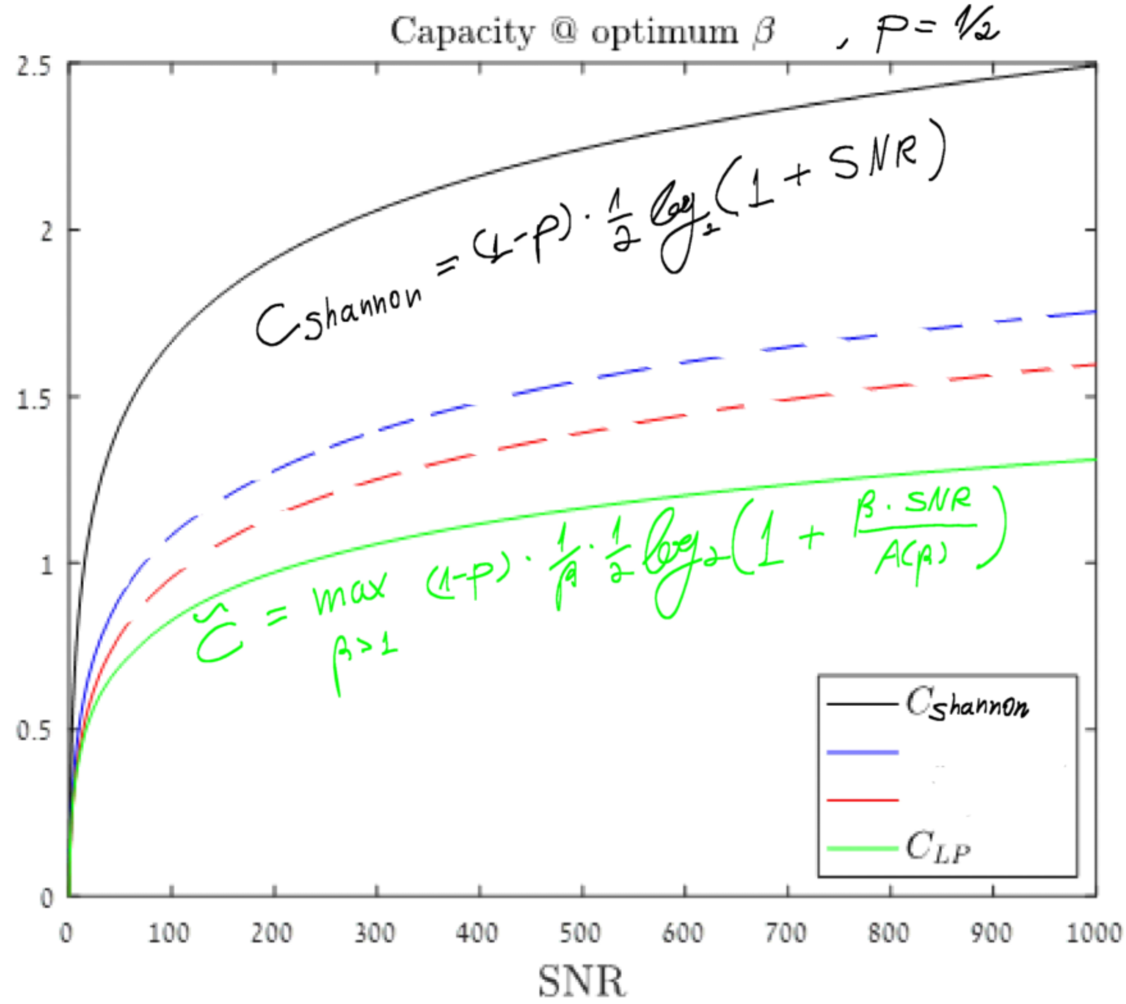
where optimum β depends on SNR



Page 15

Monday, March 06, 2017 10:37 AM

Performance with Band-Limited interpolation
(F = Low-pass)



$$C_{shannon} = (1-p)\cdot\frac{1}{2}\log_2(1+SNR)$$

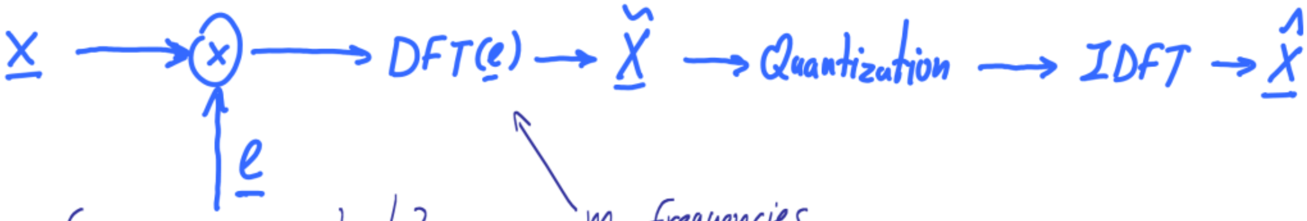
$$\hat{C} = \max_{\beta>1} (1-p)\cdot\frac{1}{\beta}\cdot\frac{1}{2}\log_2\left(1+\frac{\beta\cdot SNR}{A(\beta)}\right)$$

Page 16

Monday, March 06, 2017 10:37 AM

Analog coding for a Source with Erasures

[Haikin - Zamir ISIT 2016]



$T = \begin{cases} n-k \text{ non-important} \\ k \text{ important} \end{cases}$

m frequencies

$m \geq k$

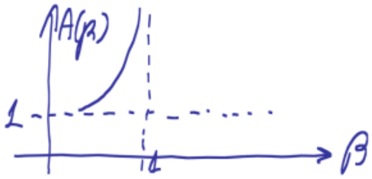
$\beta \triangleq k/m \leq 1$

Signal amplification ≥ 1, with equality iff T = uniform sampling

∴ Rate-distortion of analog coding:

$$\check{R} = \frac{m}{n} \cdot \frac{1}{2} \log\left(\widetilde{SDR}\right) = \frac{1}{\beta} \cdot R\left(1 + \beta \cdot [A(\beta) \cdot SDR - 1]\right)$$

optimum β is a function of SDR



Page 17

Monday, March 06, 2017 10:37 AM

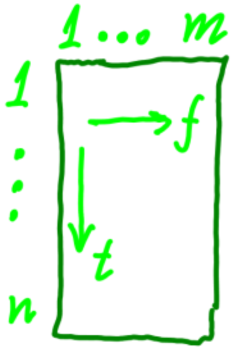
Can we do better ?

Yes, change F !

Page 18

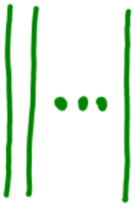
Monday, March 06, 2017 10:37 AM

From DFT to Frames and back...



- F = Band-limited DFT

- Can be viewed as



= m vectors in $\mathbb{C}^n$ = subspace of $\mathbb{C}^n$

- Or as

= n vectors in $\mathbb{C}^m$ = over-complete basis of $\mathbb{C}^m$
= frame

- Since $n > m$, frame vectors are not orthogonal

Page 19

Monday, March 06, 2017 10:37 AM

Robust Frames

Let frame $\mathcal{F} = \begin{array}{cc} & 1 \ \cdots \ m \\ 1 & \underline{\qquad\qquad} \\ 2 & \underline{\qquad\qquad} \\ \vdots & \vdots \\ n & \underline{\qquad\qquad} \end{array} = \{ \underline{x}_1, \dots, \underline{x}_n \}$

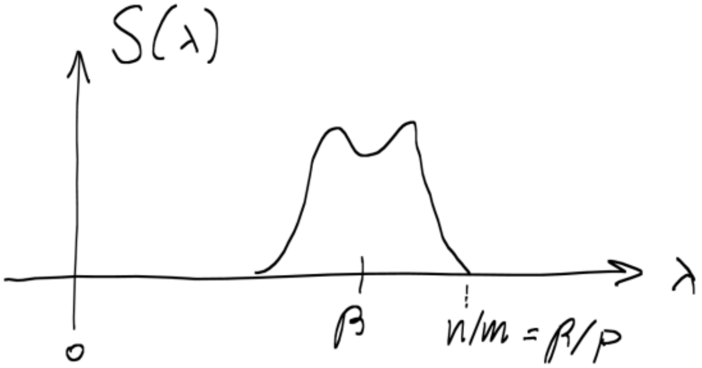
Fix $\|\underline{x}_i\| = 1$, $i = 1, \dots, n$.

* small noise amplification $A(\mathcal{F}) \simeq 1$

$\Leftrightarrow$ the $k \times m$ sub-matrix $\mathcal{F}_{\underline{e}} \triangleq T \cdot \mathcal{F}$

$\simeq$ orthonormal for "most" erasure patterns $e_1, \dots, e_n$

$\Leftrightarrow$ singular-value spectrum $\lambda_1, \dots, \lambda_m$ of $\mathcal{F}_{\underline{e}}$ concentrates around β for random erasures as $n \to \infty$

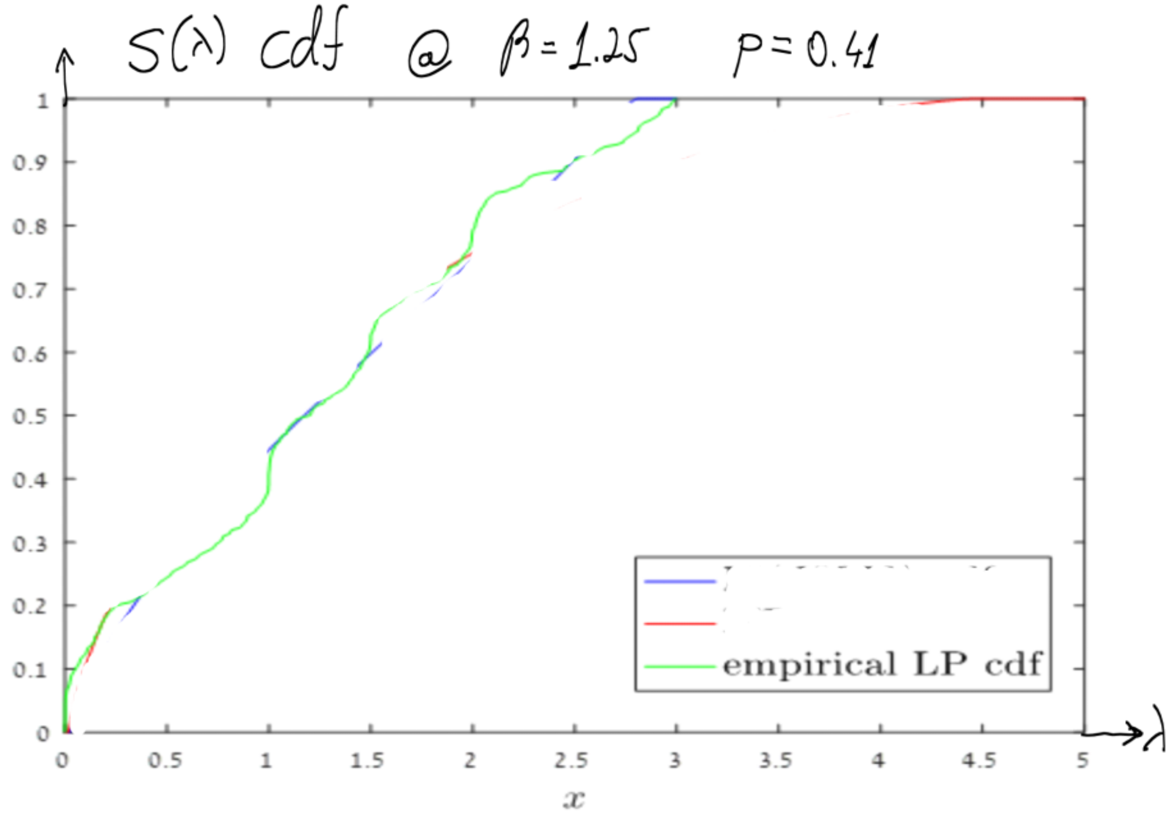


Page 20

Monday, March 06, 2017 10:37 AM

Example 1: low-pass frame

$$\underset{n\times m}{F} = \underset{n\times n}{IDFT} \cdot \underset{n\times m}{\begin{pmatrix} 1 & & 0 \\ & \ddots & 1 \\ 0 & & \end{pmatrix}}$$



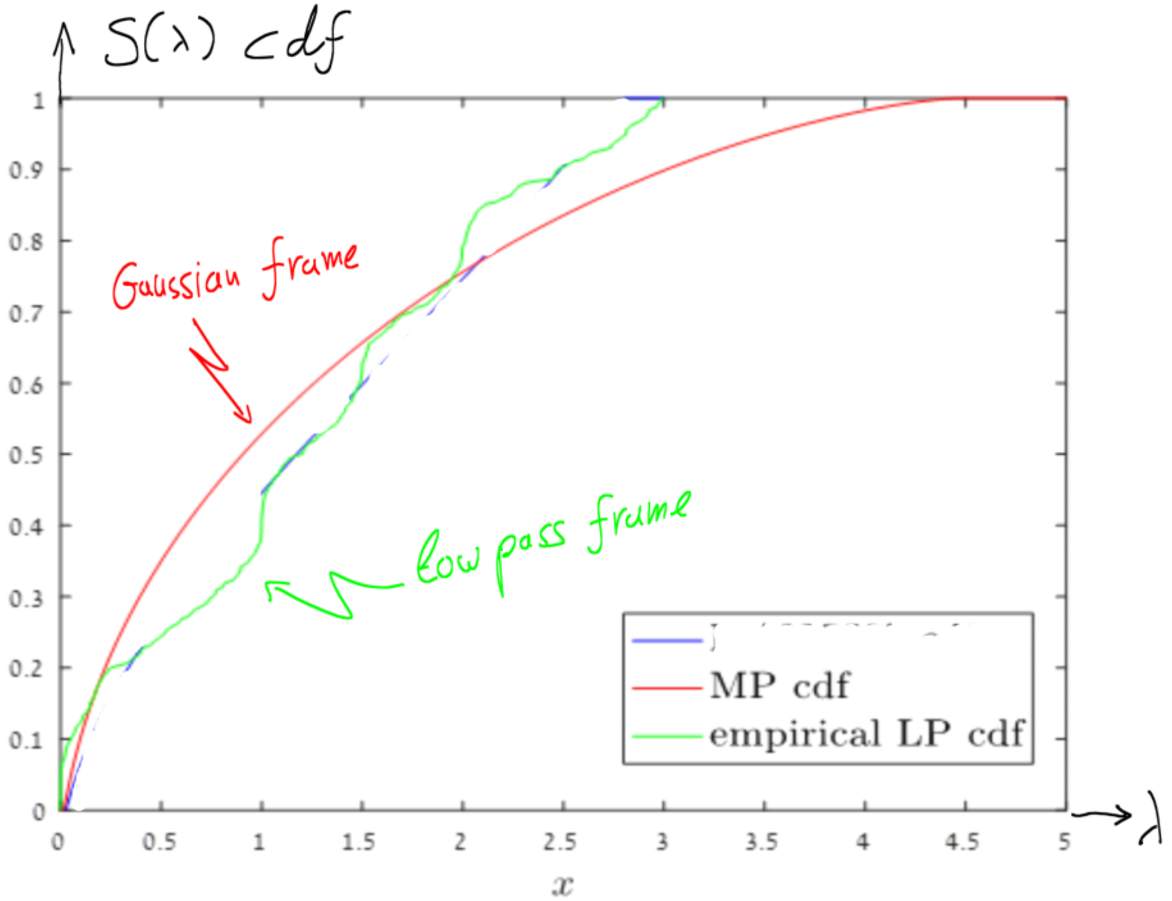
$\Rightarrow$ "truncated Hadamard" spectrum

Page 21

Monday, March 06, 2017 10:37 AM

Example 2: iid. Frame

$$\underset{n\times m}{F} \sim \text{iid } N(0, 1/m)$$



$\Longrightarrow$ Marcenko-Pastur distribution

Page 22

Monday, March 06, 2017 10:37 AM

Example 3: Difference-Set Spectrum Frame

$$\underset{m\times n}{F} = IDFT \cdot \begin{array}{c} \begin{array}{ccccc} f_1 & f_2 & f_3 & \cdots & f_m \end{array} \\ \left(\begin{array}{cccccc} 1 & — & — & — & — & — \\ — & 1 & — & — & — & — \\ — & — & 1 & — & — & — \\ & & & & \vdots & \\ — & — & — & — & 1 & — \\ & & 0 & & & \end{array}\right) \end{array}$$

$$= \left\{ e^{j2\pi \cdot f_i \cdot t} \right\}_{t=1,\dots n} \quad , \quad i=1,\dots,m$$

$f_i \in$ difference set of size m in the group Z_n *

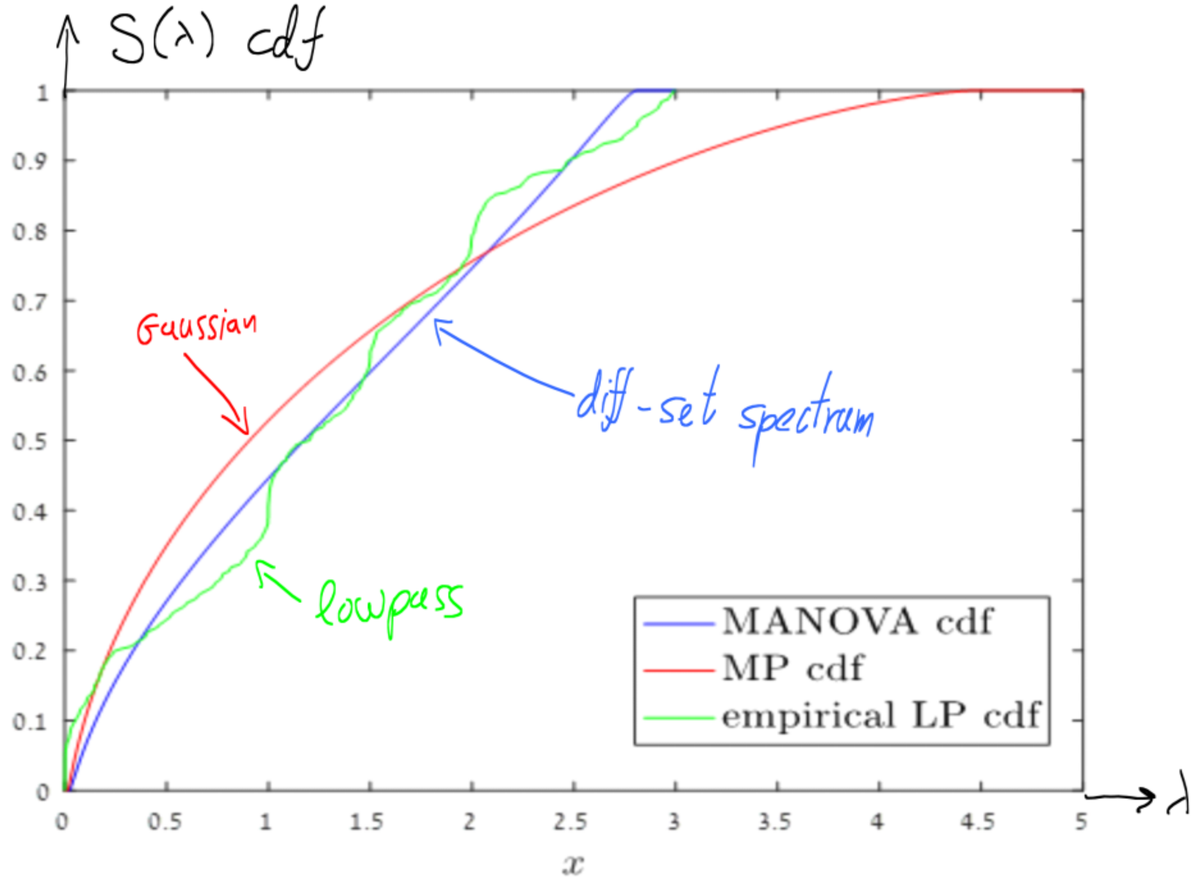
* every difference $f_i - f_j \bmod n$ appears the same number of times

Page 23

Monday, March 06, 2017 10:37 AM

Example 3: Difference-Set Spectrum Frame

$$F_{m \times n} = DSS$$



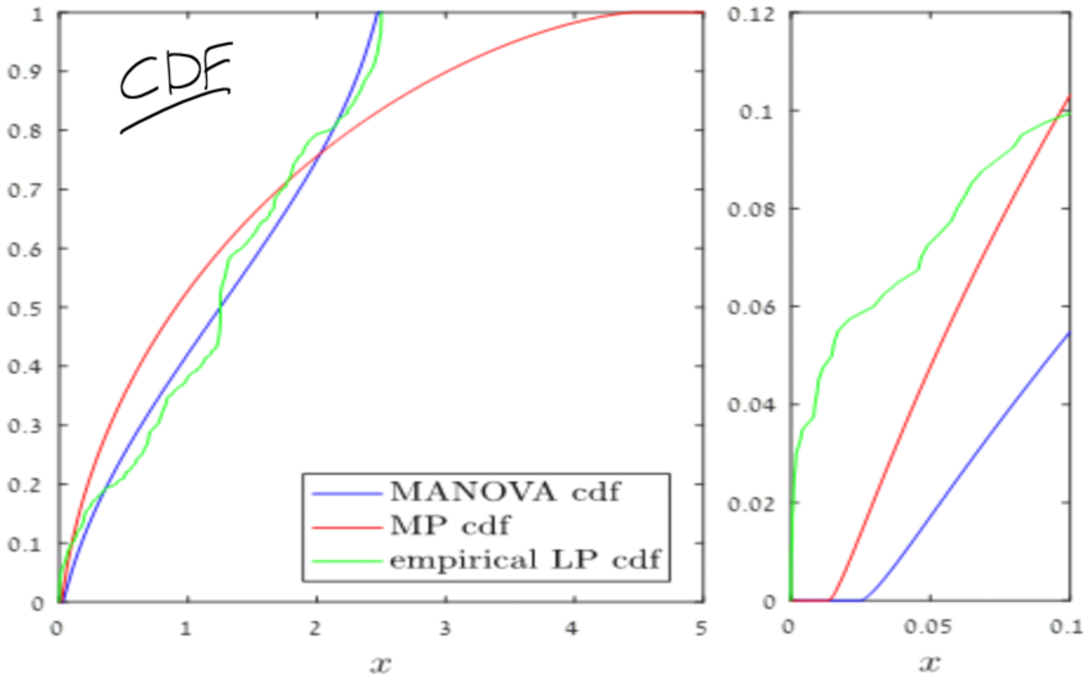
⇒ Manova distribution (experimental)

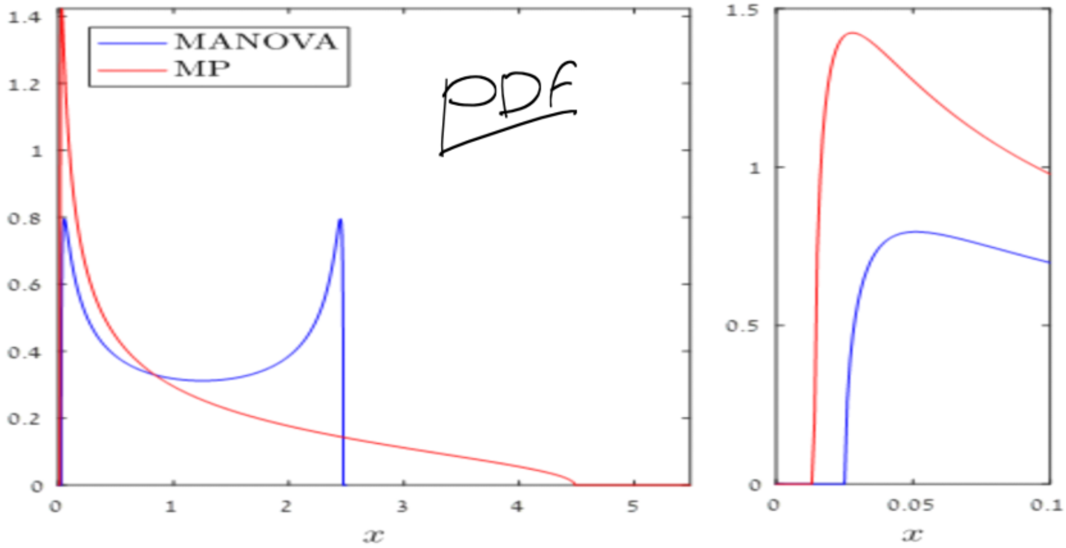
Page 24

Monday, March 06, 2017 10:37 AM

Zoom @ small singular values

$\beta = 1.25$ $p = 1/2$





Page 25

Monday, March 06, 2017 10:37 AM

Old & new results in Frame Theory

* Gaussian (iid) = Marcenko - Pastur [1967]

$H_{k \times m} \sim$ iid $N(0, 1/m)$, $\beta = k/m$ is fixed

$$\text{eigen values } \{H^t H\} \xrightarrow[m,k \to \infty]{} f_{MP}(x)$$

* Random Fourier = Manova / Jacobi [Farrell 2013]

$$H_{k \times m} = \begin{pmatrix} \text{Bernoulli}(p) & 0 \\ 0 & \ddots \end{pmatrix} \cdot IDFT \cdot \begin{pmatrix} \text{Bernoulli}(p/\beta) & 0 \\ 0 & \ddots \end{pmatrix} , \quad \begin{matrix} \beta = \frac{k}{m} \\ p = \frac{m}{n} \end{matrix} \text{ fixed}$$

$$\text{eigen values } \{H^t H\} \xrightarrow[m,k \to \infty]{} f_{Manova}(x)$$

* $\underline{\underline{E}}$qui-angular $\underline{\underline{T}}$ight $\underline{\underline{F}}$rames = Manova

[Haikin - Zamir - Gavish , ArXiv Jan 2017]

Note that DSS frame = ETF

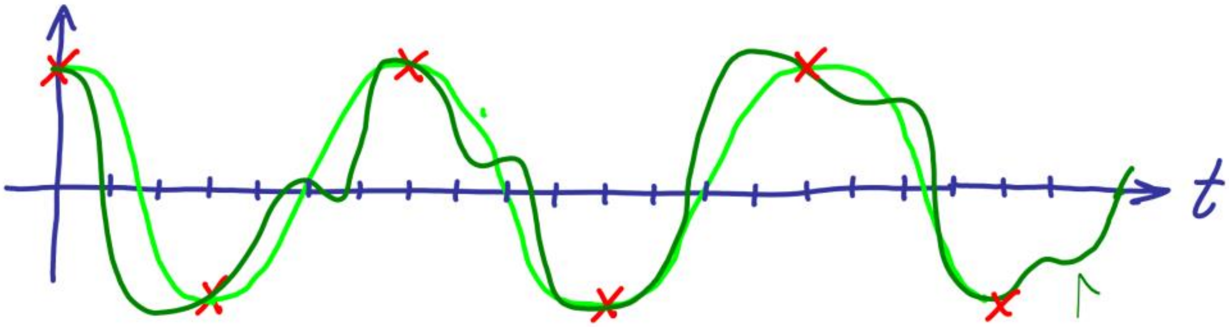
Page 26

Monday, March 06, 2017 10:37 AM

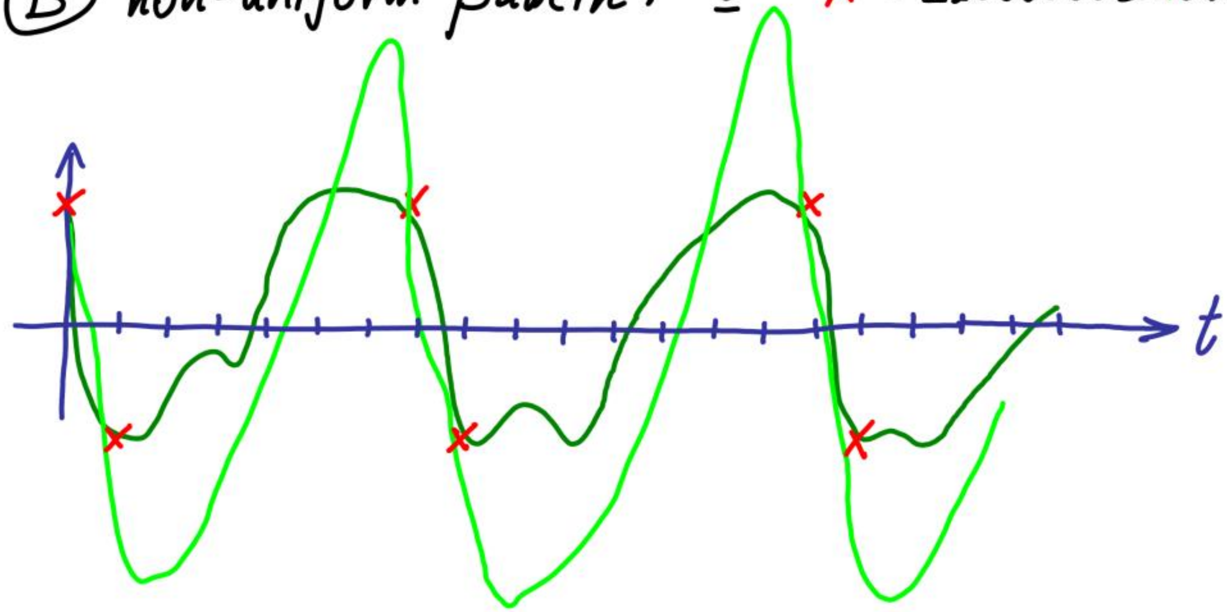
Why DSS is robust to erasure pattern ?

∼ = DSS   ∼ = LP

Ⓐ uniform pattern : $\underline{e}$ = X = 100010001000...



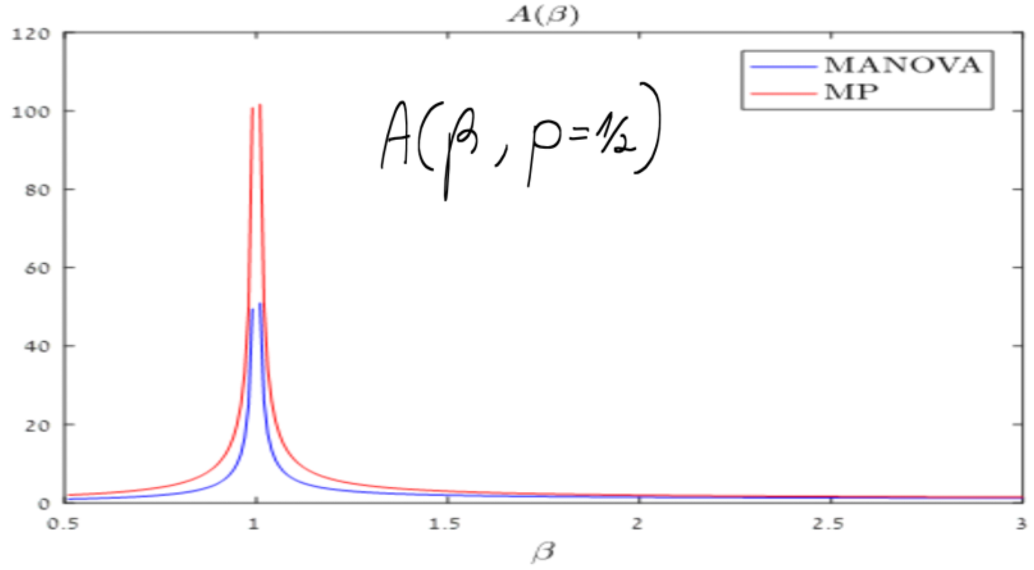
Ⓑ non-uniform pattern : $\underline{e}$ = X = 1100000011000000...

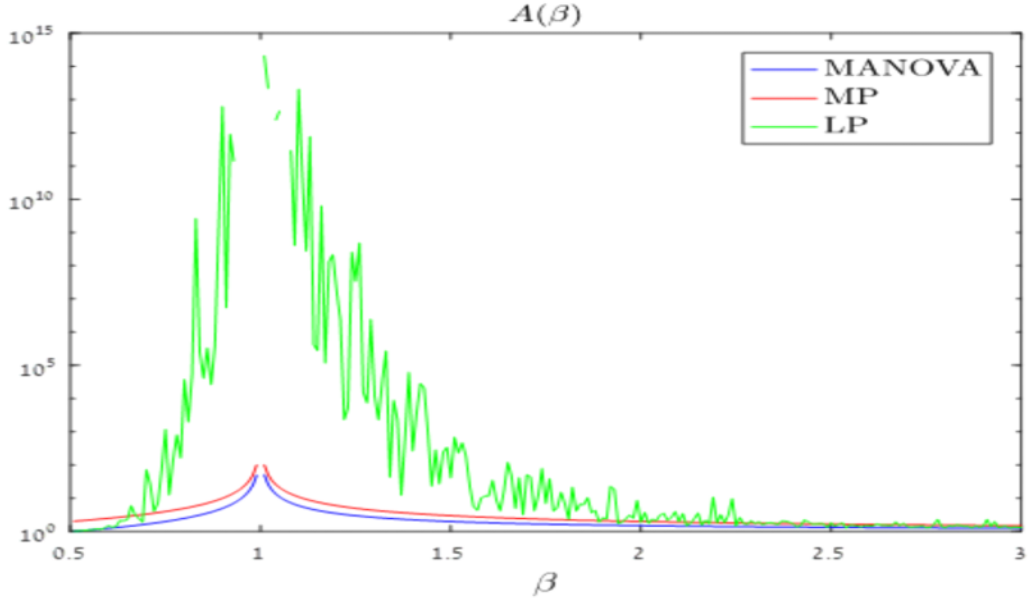


Page 27

Monday, March 06, 2017 10:37 AM

Amplification, Capacity & RDF for Analog Coding using (good) frames

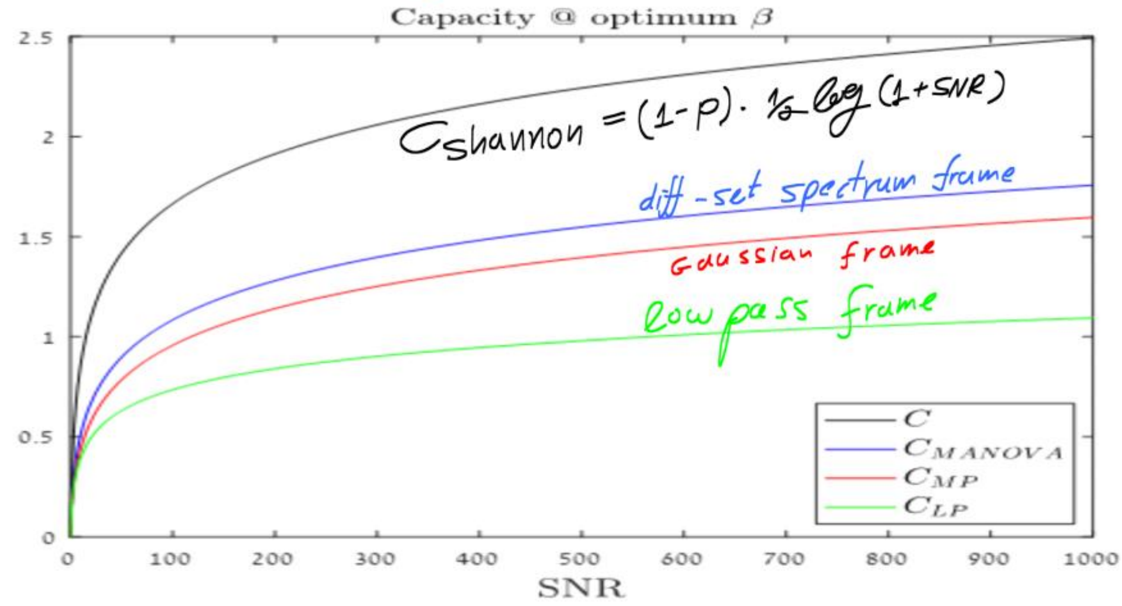


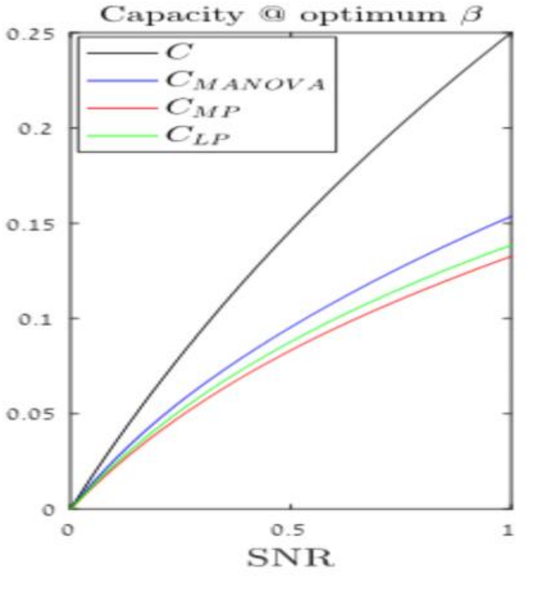


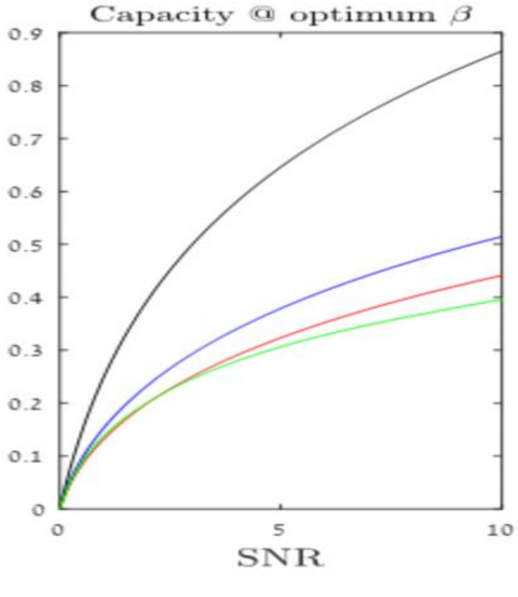
Page 28

Monday, March 06, 2017 10:37 AM

Amplification, Capacity & RDF for Analog Coding using (good) frames







Limiting capacity loss in analog coding

Let $A(F, \beta, \rho)$ be the limiting $(n \to \infty)$ amplification of a frame family F, with capacity

$$\check{C} \triangleq \sup_{\beta \geq 1} \left\{ \frac{1-\rho}{2} \cdot \frac{1}{\beta} \log\left(1 + \frac{\beta \cdot SNR}{A(\beta)}\right) \right\}.$$

HSNR) If $A(F, \beta, \rho) < \infty \quad \forall \beta > 1$

then $\check{C}/C \xrightarrow[SNR \to \infty]{} 1$

LSNR) If $\inf_{\beta \geq 1} A(F, \beta, \rho) = 1$,

then $\check{C}/C \xrightarrow[SNR \to 0]{} 1$.

Page 29

Monday, March 06, 2017 10:37 AM

Summary

- Digital-analog Coding for noise & erasures

- Goodness measure = noise amplification

- Manova Spectrum superiority

- Difference-Set Spectrum : a deterministic construction for good frames

Page 30

Monday, March 06, 2017 10:37 AM

Thank You

Page 31

Monday, March 06, 2017 10:37 AM