COMPRESSED SENSING:

FROM ALGORITHMS TO CIRCUITS

Gianluca Setti

gianluca.setti@unife.it

SSIGPRO – Statistical Signal Processing Group

ENDIF-University of Ferrara, ARCES –University of Bologna

Politecnico di Torino

July 9th, 2014

University of Bologna

University of Ferrara

ACKNOWLEDGEMENTS : OUR GROUP ON COMPRESSED SENSING

Full-time Faculty Members:

Fabio Pareschi, UNIFE, Assistant Professor (EMI Reduction, RNG, Analog Electronics, *Compressed Sensing*)

Riccardo Rovatti, UNIBO, Professor (Digital Electronics and Signal Processing)

Gianluca Setti, UNIFE/UNIBO, Professor (Analog Electronics, Circuit Theory and Signal Processing)

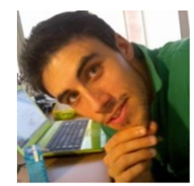
Post-Docs, present and former PhD Students:

Mauro Mangia, UNIBO, Post-Doc (*Compressed Sensing*, DS-CDMA communication)

Valerio Cambareri, UNIBO, PhD Student (*Compressed Sensing*, Hyperspectral Imaging)

Nicola Bertoni, UNIBO, PhD Student (*Compressed Sensing*, High frequency DC/DC converter design)

Javier Haboba, UNIBO, former PhD Student (*Compressed Sensing*, thermal-aware DC/DC converter design) – now Researcher @YPF (Argentina)



- Theory/Introduction
  1. E. Candès, J. Romberg, T Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information", *IEEE Trans. on Information Theory*, 52(2) pp. 489 - 509, February 2006
  2. E. Candès, J. Romberg, "Quantitative robust uncertainty principles and optimally sparse decompositions", *Foundations of Comput. Math.*, 6(2), pp. 227 - 254, April 2006
  3. D. Donoho, "Compressed sensing", *IEEE Trans. on Information Theory*, 52(4), pp. 1289 - 1306, April 2006
  4. E. J. Candes ,M. B. Wakin, “An Introduction to Compressive Sampling,” *IEEE Signal Processing Magazine*, 25 (2) pp. 21–30, March 2008
  5. E. Candes, “The restricted isometry property and its implications for compressed sensing,” *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.
  6. D. Donoho and J. Tanner, “Precise undersampling theorems,” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 913 –924, June 2010.
- Decoding/Reconstruction
  1. E. J. Candes, T. Tao, “Decoding by linear programming,” *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005
  2. E. J. Candes, J. K. Romberg, K. Justin, T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
  3. D. Needell, J. Tropp, “Cosamp: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
  4. A. Maleki and D. Donoho, “Optimally tuned iterative reconstruction algorithms for compressed sensing,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 330–341, 2010.
- Circuits and Systems Implementations
  1. Jason Laska, Sami Kirolos, Marco Duarte, Tamer Ragheb, Richard Baraniuk, and Yehia Massoud, Theory and implementation of an analog-to-information converter using random demodulation. (IEEE Int. Symp. on Circuits and Systems (ISCAS), New Orleans, Louisiana, 2007)

...AND SOME BIBLIOGRAPHY - II

1. Robucci, R.; Gray, J.D.; Leung Kin Chiu; Romberg, J.; Hasler, P.; , "Compressive Sensing on a CMOS Separable-Transform Image Sensor," *Proceedings of the IEEE* , vol.98, no.6, pp.1089-1101, June 2010
2. Xi Chen; Zhuizhuan Yu; Hoyos, S.; Sadler, B.M.; Silva-Martinez, J.; , "A Sub-Nyquist Rate Sampling Receiver Exploiting Compressive Sensing," *Circuits and Systems I: Regular Papers, IEEE Transactions on* , vol.58, no.3, pp.507-520, March 2011
3. Chen, F.; Chandrakasan, A.P.; Stojanovic, V.M.; , "Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," *Solid-State Circuits, IEEE Journal of* , vol.47, no.3, pp.744-756, March 2012
4. Yenduri, P. K.; Rocca, A. Z.; Rao, A. S.; Naraghi, S.; Flynn, M. P.; Gilbert, A. C.; , "A Low-Power Compressive Sampling Time-Based Analog-to-Digital Converter," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on,* September 2012
5. Mamaghanian, H.; Khaled, N.; Atienza, D.; Vandergheynst, P.; , "Design and Exploration of Low-Power Analog to Information Conversion Based on Compressed Sensing," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on,* September 2012

- Applications

1. Suarez, M.; Brea, V. M.; Fernandez-Berni, J.; Carmona-Galan, R.; Linan, G.; Cabello, D.; Rodriguez-Vazquez, A.; , "CMOS-3D Smart Imager Architectures for Feature Detection," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on,* September 2012
2. Hyder, Md. D.; Mahata, K.; , "Maximum a Posteriori Based Approach for Target Detection in MTI Radar," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*
3. Patel, V.M.; Easley, G.R.; Healy, D.M.; Chellappa, R.; , "Compressed Synthetic Aperture Radar," *Selected Topics in Signal Processing, IEEE Journal of* , vol.4, no.2, pp.244-254, April 2010
4. Potter, L.C.; Ertin, E.; Parker, J.T.; Cetin, M.; , "Sparsity and Compressed Sensing in Radar Imaging," *Proceedings of the IEEE* , vol.98, no.6, pp.1006-1020, June 2010
5. Starck, J.-L.; Bobin, J.; , "Astronomical Data Analysis and Sparsity: From Wavelets to Compressed Sensing," *Proceedings of the IEEE* , vol.98, no.6, pp.1021-1030, June 2010

- Journals

1. M. Mangia, R. Rovatti, G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 5, pp. 1001 –1014, May 2012, (**2013 IEEE CAS Society Guillemin-Cauer Best Paper Award**)
2. J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A pragmatic look at some compressive sensing architectures with saturation and quantization," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, Sep. 2012.
3. A. Caprara, F. Furini, A. Lodi, M. Mangia, Rovatti, G. Setti, "Generation of Antipodal Random Vectors With Prescribed Non-Stationary 2-nd Order Statistics," *IEEE Transactions on Signal Processing*, vol.62, no.6, pp.1603,1612, March, 2014
4. V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, "Low-Complexity Multiclass Encryption by Compressed Sensing Part I: Definition and Main Properties," *IEEE Transactions on Signal Processing*, under revision
5. V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, "Low-Complexity Multiclass Encryption by Compressed Sensing Part I: Known-Plaintext Attacks," *IEEE Transactions on Signal Processing*, under revision
6. P. Albertini, G. Frattini, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, "An Analog-to-Information converter for Biomedical Signals based on Compressive Sensing," *IEEE Transactions on Biomedical Circuits and Systems,* submitted

- Conferences

1. M. Mangia, R. Rovatti, and G. Setti, "Analog-to-information conversion of sparse and non-white signals: Statistical design of sensing waveforms," in Proceedings of *2011 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2011, pp. 2129–2132 (**best student paper award**)
2. M. Mangia, J. Haboba, R. Rovatti, G. Setti, "Rakeness-Based Approach to Compressed Sensing of ECGs", *IEEE Biomedical Circuits and Systems Conference (BioCAS2011)*, San Diego, November 2011.

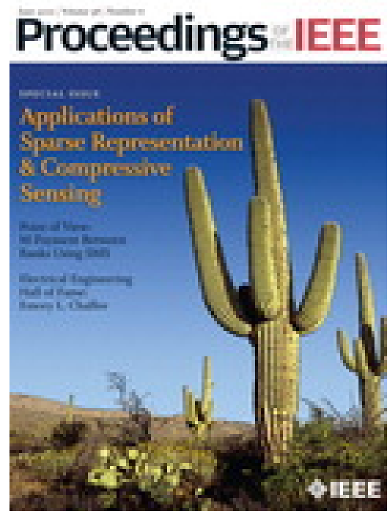
- Conferences (cont)

3. J. Haboba, M. Mangia, R. Rovatti, G. Setti “An Architecture for 1-Bit Localized Compressive Sensing with Applications to EEG”, *IEEE Biomedical Circuits and Systems Conference (BioCAS2011)*, San Diego, November 2011.
4. M. Mangia, F. Pareschi, R. Rovatti, G. Setti, G. Frattini, “Coping with Saturating Projection Stages in RMPI-Based Compressive Sensing,” *2012 IEEE International Symposium on Circuits and Systems (ISCAS2012*), Seoul, May 2012
5. V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, "A Rakeness-based Design Flow for Analog-to-Information Conversion by Compressive Sensing", *2013 IEEE International Symposium on Circuits and Systems (ISCAS)*, Beijing, 2013
6. V. Cambareri, J. Haboba, F. Pareschi, R. Rovatti, G. Setti, K. Wong, "A Two-Class Information Concealing System based on Compressed Sensing", *2013 IEEE International Symposium on Circuits and Systems (ISCAS)*, Beijing, 2013
7. N. Bertoni, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, “Correlation Tuning in Compressive Sensing Based on Rakeness: a Case Study,” *2013 IEEE International Conference on Circuits and Systems (ICECS2013)*, Abu Dhabi, December 2013
8. V. Cambareri, M. Mangia, R. Rovatti, G. Setti, “Joint Analog-to-Information Conversion of Heterogeneous Biosignals,” *IEEE Biomedical Circuits and Systems Conference (BioCAS2013)*, Rotterdam, October 2013
9. D. Bellasi, R. Rovatti, L. Benini, G. Setti, “An Architecture for Low-Power Compressed Sensing and Estimation in Wireless Sensor Nodes,” *IEEE International Symposium on Circuits and Systems (ISCAS2014)*, Melbourne, June 2014
10. M. Mangia, R. Rovatti, G. Setti, P. Vandergheynst, “Combining Spread Spectrum Compressive sensing with Rakeness for low frequency modulation in RMPI architecture,” *IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP),* Florence, May 2014
11. V. Cambareri, R. Rovatti, G. Setti, “Maximum Entropy Hadamard Sensing of Sparse and Localized Signals,” *IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP)*, Florence, May 2014

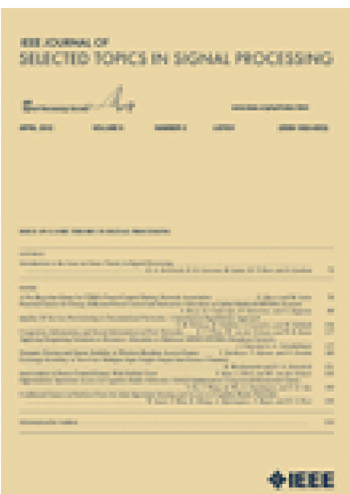
...AND SOME BIBLIOGRAPHY - I

- A few important special issues

Special Issue on Compressive Sampling
Gust Editors: R. Baraniuk; E. Candes; R. Nowak; M. Vetterli
IEEE Signal Processing Magazine, **March 2008**



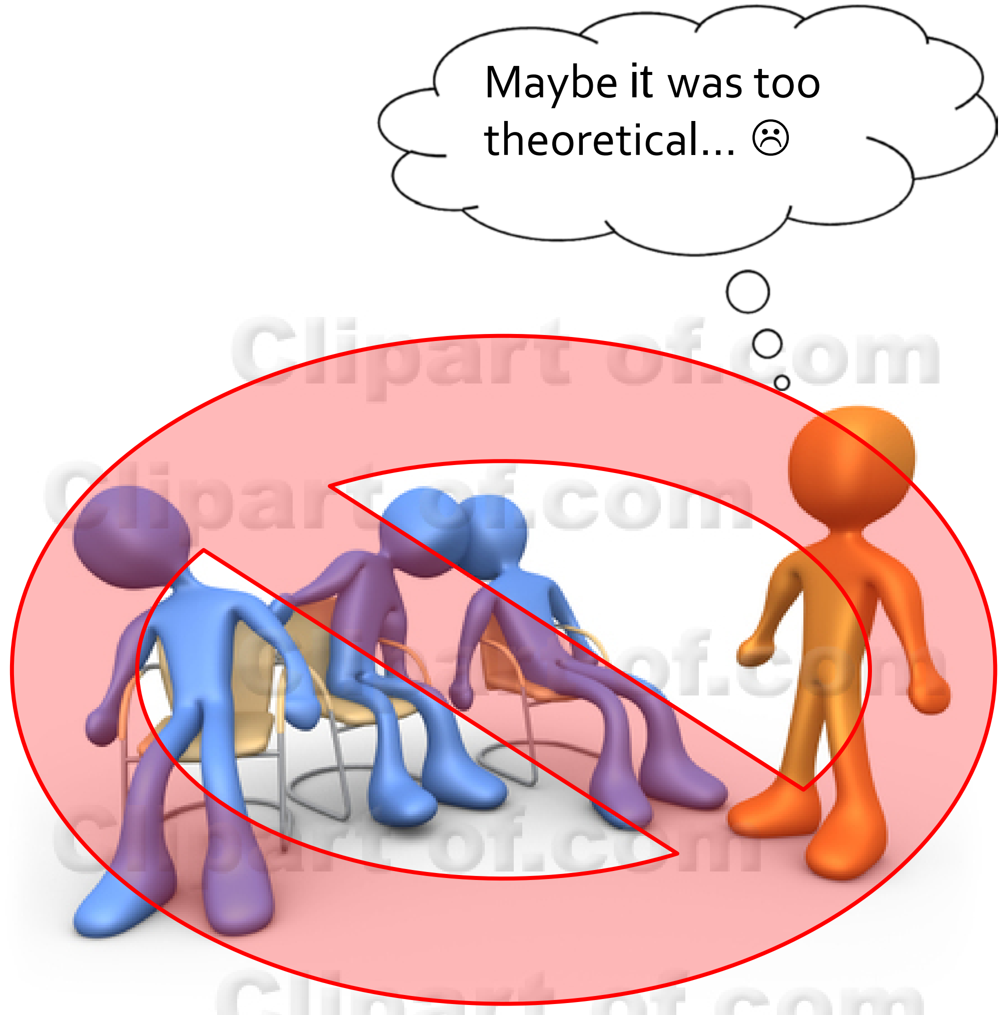
Special Issue on Applications of Sparse Representation & Compressive Sensing
Gust Editors: R. Baraniuk; E. Candes; M. Elad; Yi Ma
Proceedings of the IEEE, **June 2010**



Special Issue on Compressive Sensing
Gust Editors: R. Chartrand, R. Baraniuk, Y. Eldar, M. Figueiredo, J. Tanner
IEEE J. on Selected Topics in Signal Processing, **April 2010**

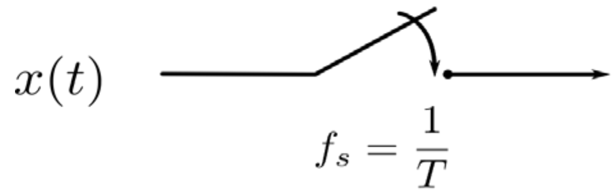
Special Issue on Circuits, Systems and Algorithms for Compressive Sensing
Gust Editors: D. Allstot, R.Rovatti, G. Setti
IEEE J. on Emerging and Selected Topics in Circuits and Systems, **September 2012**

TIME FOR A DECISION….



- Foreword / On Sampling
- An Introduction to Compressed Sensing
  - Sparse Signals and Basis Expansions
  - Compressed Sensing (and its intuitive view in the "land of vectors")
  - Incoherent acquisition and Sparse Signal Recovery by Linear Programming
  - Improving Compressed Sensing with *Rakeness*
- Applications
  - Surface ElectroMyoGraphy signal acquisition (with IIT@PoliTo)
  - Embedded Security in CS
- Analog-to-Information Conversion
  - Random Modulation Pre-Integration (RMPI) implementation in 180nm TI CMOS technology (hardware-algorithm co-design)
  - A smart way to deal with saturation
  - Measured Performances
- Digital CS Front End and Signal Feature Extraction in CS domain
- Conclusion

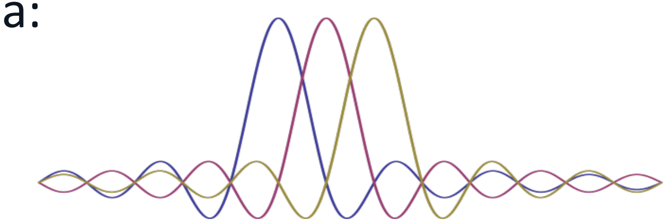
- *Nyquist – Shannon Sampling Theorem:* to perfectly reconstruct a finite – bandwidth signal $x(t)$ from its uniform samples $x(j\,T)$, acquire them at sampling rate $f_s > 2W$ (Nyquist rate)



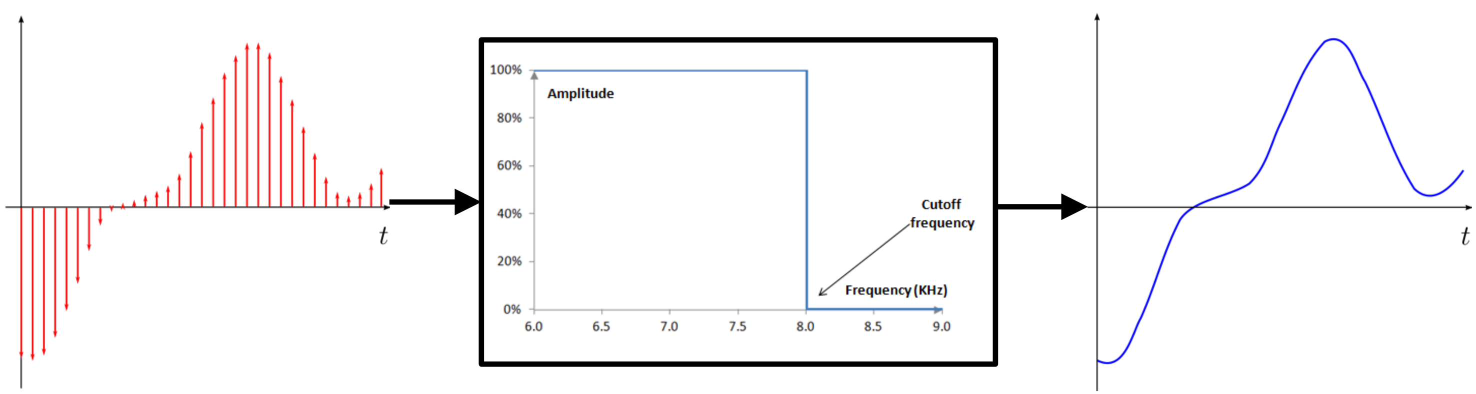
$$x_j = x(jT)$$

- The theorem guarantees (**<u>sufficient condition!!</u>**) that $x(t)$ can be reconstructed by the Shannon – Whittaker interpolation formula:

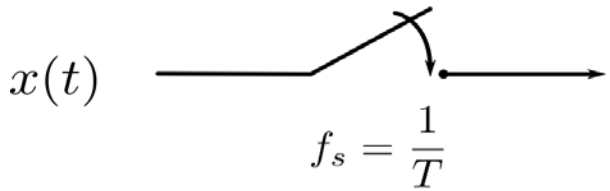
$$x(t) = \sum_{j=-\infty}^{+\infty} x(jT)\,\mathrm{sinc}\left(\frac{t - jT}{T}\right)$$



- Constructive and general approach to signal reconstruction (**no a-priori information** on the signal structure!!):



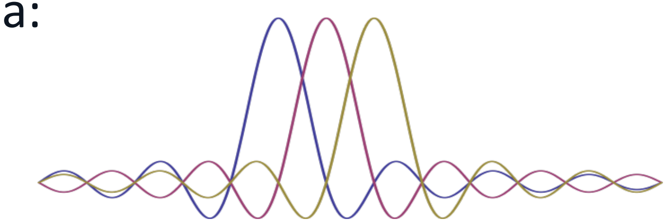
- *Nyquist – Shannon Sampling Theorem:* to perfectly reconstruct a finite – bandwidth signal $x(t)$ from its uniform samples $x(j\,T)$, acquire them at sampling rate $f_s > 2W$ (Nyquist rate)



$$x_j = x(jT)$$

- The theorem guarantees (**<u>sufficient condition!!</u>**) that $x(t)$ can be reconstructed by the Shannon – Whittaker interpolation formula:

$$x(t) = \sum_{j=-\infty}^{+\infty} x(jT)\,\mathrm{sinc}\left(\frac{t-jT}{T}\right)$$

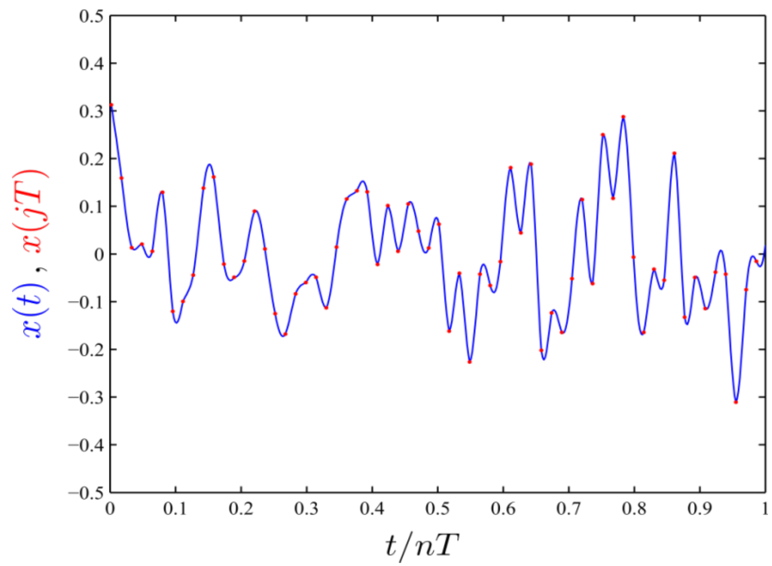


- Constructive and general approach to signal reconstruction (**no a-priori information** on the signal structure!!):

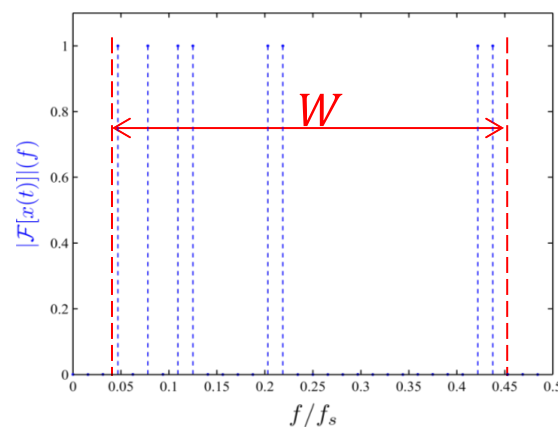
Most intuitive and used approach to sampling.

<u>Most simple and general decoding/reconstructing solution</u>

- **A problem:** High sampling rates ⇒ *high dynamic power consumption* – a motivating hardware issue to try and *lower the sample rate as much as possible!*
- … but Nyquist-Shannon theorem gives only **a sufficient condition** for general signal reconstruction….. So we trade *generality* for *lower sampling rates* when acquiring signals whose *structural properties* allow sampling below the Nyquist rate.
- Consider this toy example: $x(t) = \sum_{k=1}^{8} \sin(2\pi f_k)$

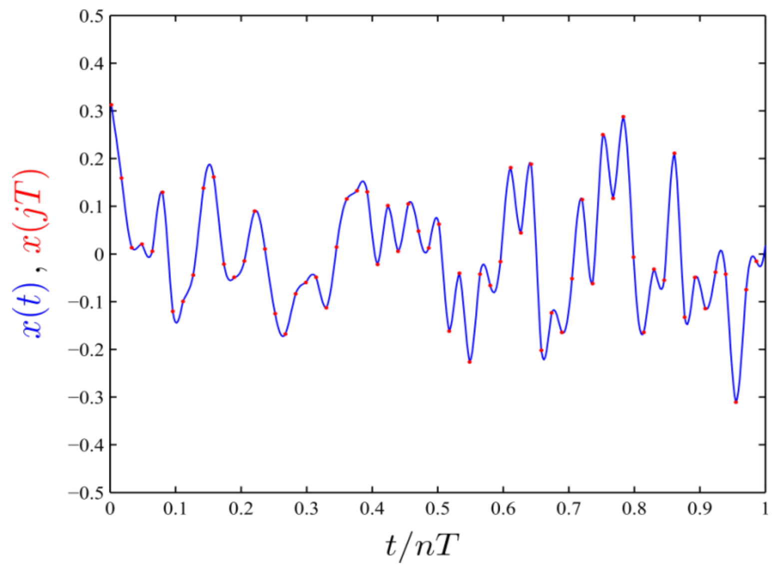


$$\{f_k/f_s\} = \{0.05, 0.075, 0.11, 0.12, 0.2, 0.22, 0.43, 0.44\}$$

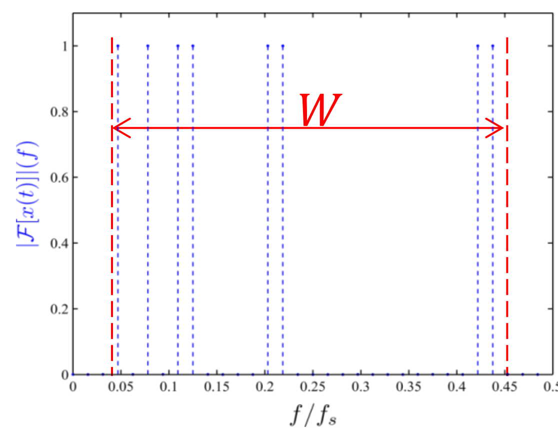


- Although $x(t)$ is a sum of eight tones, its bandwidth is large / it would require a sampling rate much higher than its actual information content.
- $x(t)$ has a very simple representation (very few terms) w.r.t. the Fourier basis in $[0, nT]$.

- **A problem:** High sampling rates ⇒ *high dynamic power consumption* – a motivating hardware issue to try and *lower the sample rate as much as possible!*
- … but Nyquist-Shannon theorem gives only **a sufficient condition** for general signal reconstruction….. So we trade *generality* for *lower sampling rates* when acquiring signals whose *structural properties* allow sampling below the Nyquist rate.
- Consider this toy example: $x(t) = \sum_{k=1}^{8} \sin(2\pi f_k)$



$$\{f_k/f_s\} = \{0.05, 0.075, 0.11, 0.12, 0.2, 0.22, 0.43, 0.44\}$$



- Although $x(t)$ is a sum of eight tones, its bandwidth is large / it would require a sampling rate much higher than its actual information content.
- $x(t)$ has a v basis in $[0, nT]$.

The information content of this signal is much less than what W indicates

- *How many **<u>significant terms</u>** are there in the basis representation of $x(t)$?*
- Let a generic basis expansion:

$$x(t) = \sum_{j=0}^{n-1} \alpha_j \psi_j(t) \qquad \alpha_j = \langle x(t), \psi_j(t) \rangle_{[0,nT]} = \int_0^{nT} x(t)\psi_j(t)dt$$

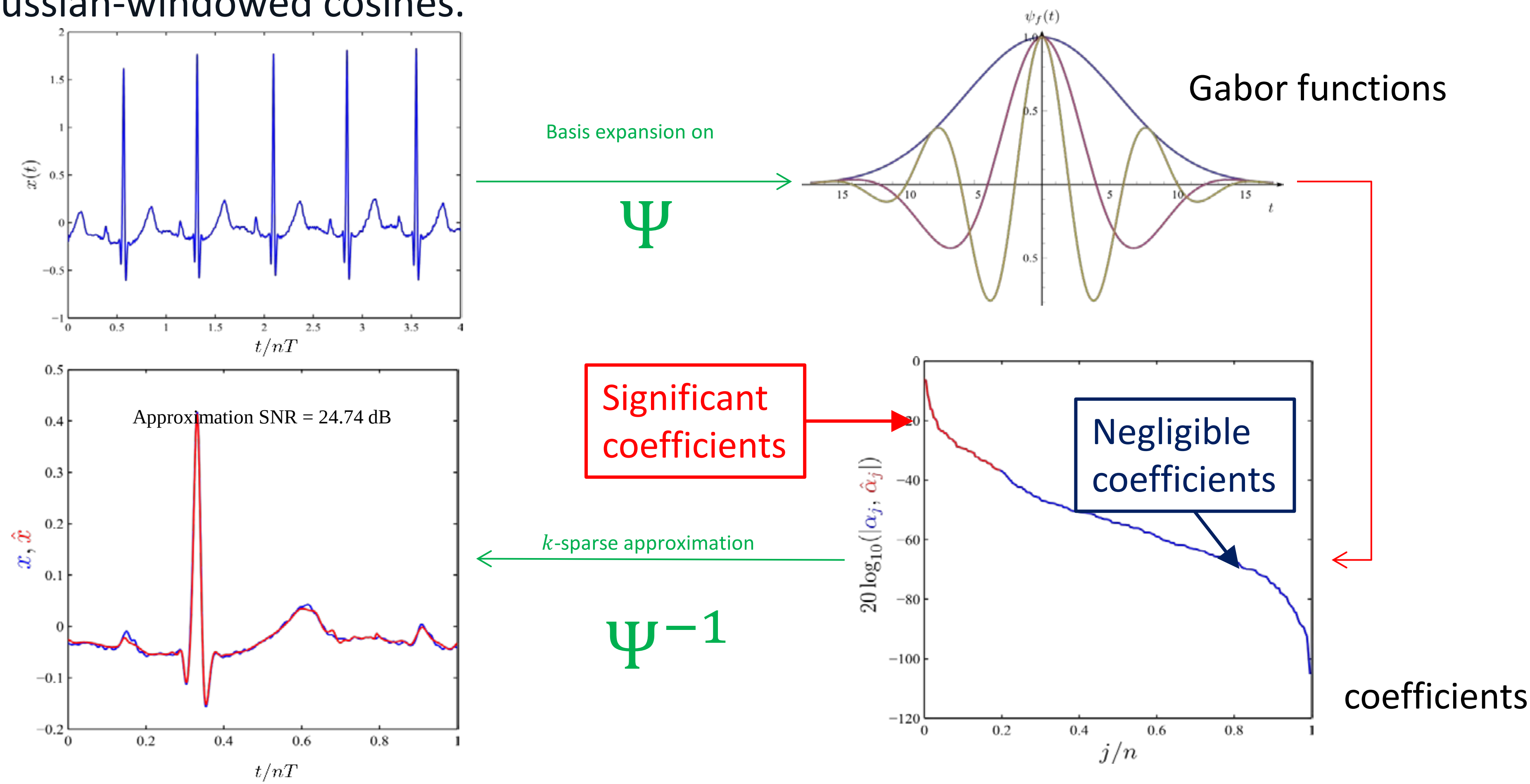
- We define **sparse** any signal **with very few nonzero coefficients** in its basis expansion w.r.t. the dimensionality of the sparsity basis $\Psi_{n \times n}$.
- Our example $x(t)$ is 8-sparse. In general, a k-sparse signal is such that

$$x = \Psi\alpha\,,\ \alpha \in \mathbb{R}^n\,, k = \#\{j \in [0, n-1]\ :\ \alpha_j \neq 0\}$$

<u>Note:</u>

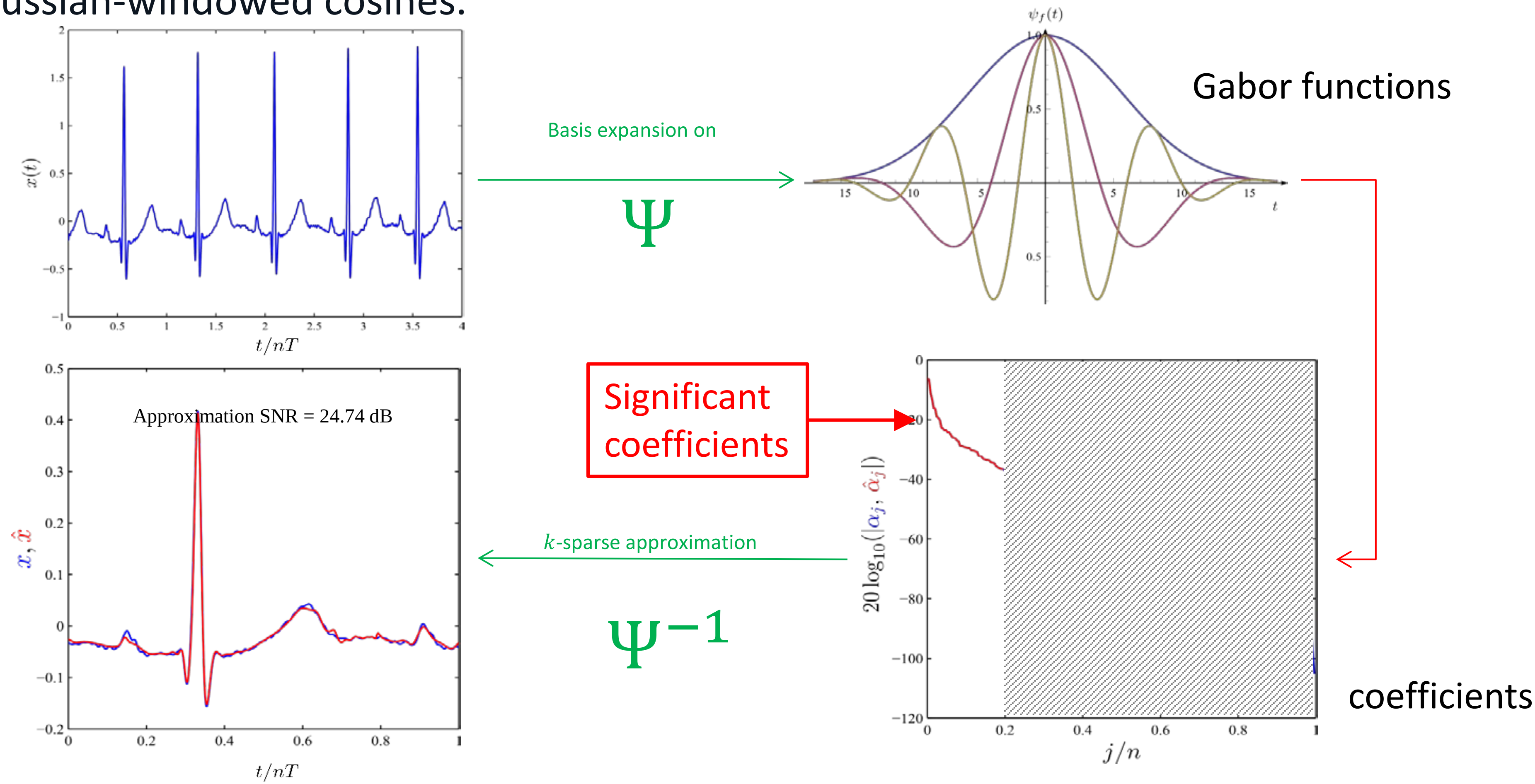
- A signal may be *sparse* in some bases and *dense* in others. Finding a basis or dictionary such that $k < n$ coefficients represent ***most* of the information/energy** in a signal also means you have found **a way to *compress* it after acquisition** (*trivial compression: discard $n - k$ negligible coefficients*).

- Examples: wavelet domain compression of images, biosignals; audio coding.
- **ECG signals** are sparse w.r.t. a *(real)* *Gabor dictionary* – a set of scaled and shifted, Gaussian-windowed cosines.



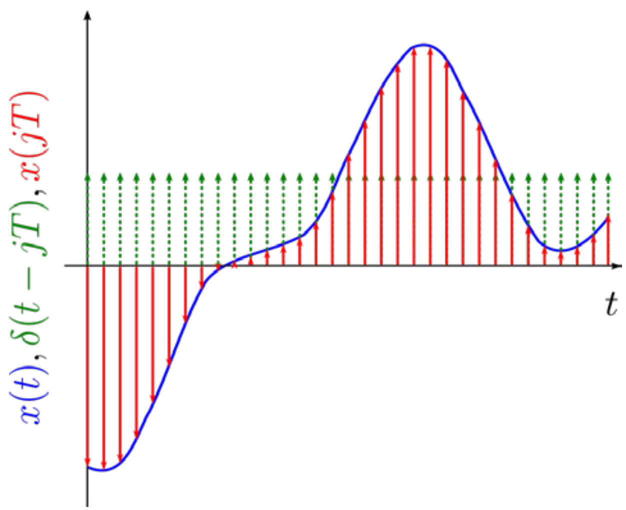
- If one could "sample" only the most significant (red) coefficients, the sampling operation would be much more efficient….

- Examples: wavelet domain compression of images, biosignals; audio coding.
- **ECG signals** are sparse w.r.t. a *(real)* *Gabor dictionary* – a set of scaled and shifted, Gaussian-windowed cosines.



- If one could "sample" only the most significant (red) coefficients, the sampling operation would be much more efficient….

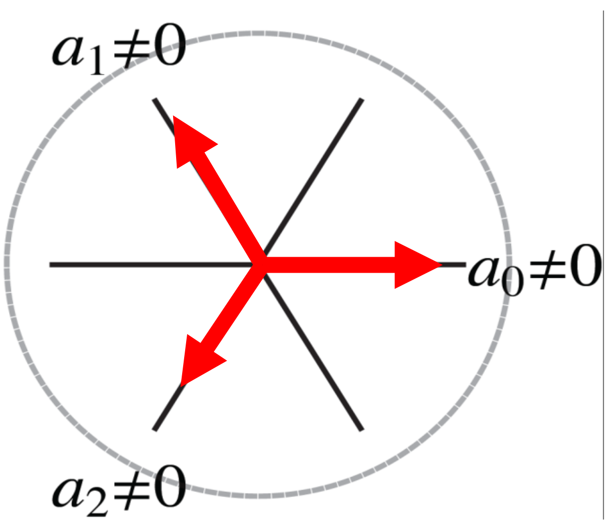
- **Good News**: most signals are sparse when expressed in a proper basis (all biosignals, even images, radar, radiotelescope…)… **Bad News**: but how can we acquire them?
- **Remark**: the sampling operation is a linear projection



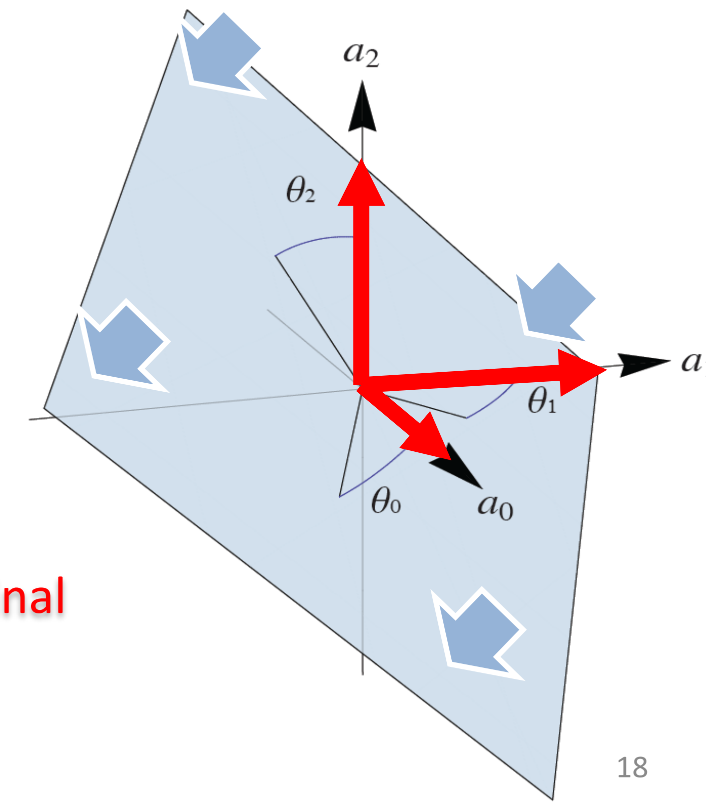
$$x(jT) = \int_0^{nT} x(t)\delta(t-jT)dt$$
$$= \langle x(t), \delta(t-jT)\rangle_{[0,nT]}$$
$$= \langle x(t), \phi_j(t)\rangle_{[0,nT]}$$

Projection onto $\phi_j(t)$ to which one needs to **add (projection independent) noise**

- How should we choose $\phi_j(t)$ to get the information?

- **Good News**: most signals are sparse when expressed in a proper basis (all biosignals, even images, radar, radiotelescope…)… **Bad News**: but how can we acquire them?
- **Remark**: the sampling operation is a linear projection
- Assume that the signal to acquire is represented by a vector $\boldsymbol{x}$ which is sparse since it has only one component wrt the coordinate system (a_0, a_1, a_2)
- If I know that only one component is non-null, I do not acquire all 3 components, but only the projection (i.e. 2 components only) on a plane which must be aligned to (a_0, a_1, a_2) to have non-zero projections ($\theta_j \neq \pi/2$)
- 

…and reconstruct the orginal signal from projections.



Compressive Sensing (E. Candès, D. Donoho 2006)

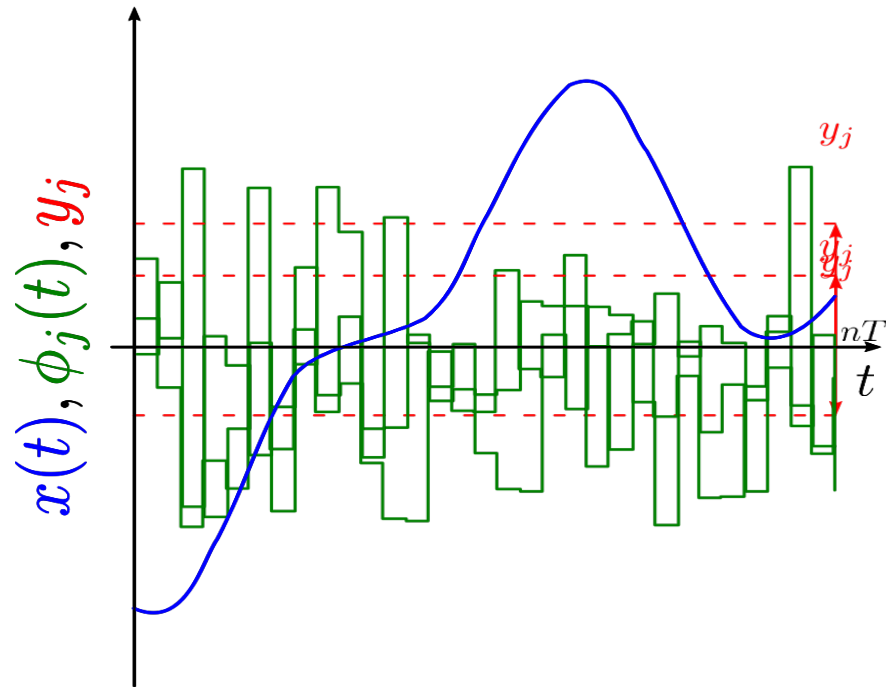
- Signal $x(t)$defined for $0 \leq t \leq nT$
- Capturing signal information content by projections on a proper set of different waveforms

$$y_j = \int_0^{nT} x(t)\phi_j(t)dt = \langle x(t), \phi_j(t) \rangle_{[0,nT]}$$

- ***Compressive Sensing***: by using a proper set of sampling function $\phi_j(t)$, we can get signal information content by using only ***m*** waveforms (with $\underline{m < n}$)

CS

$$y_j = \int_0^{nT} x(t)\phi_j(t)dt$$
$$= \langle x(t), \phi_j(t) \rangle_{[0,nT]}$$



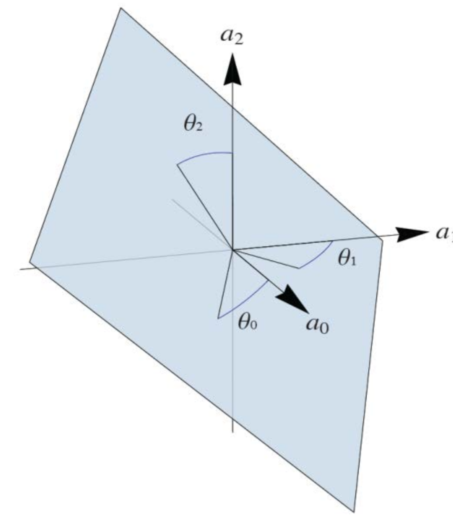
- *Compressed Sensing* (E. Candès, D. Donoho 2006) proposes a radically new approach to sampling:

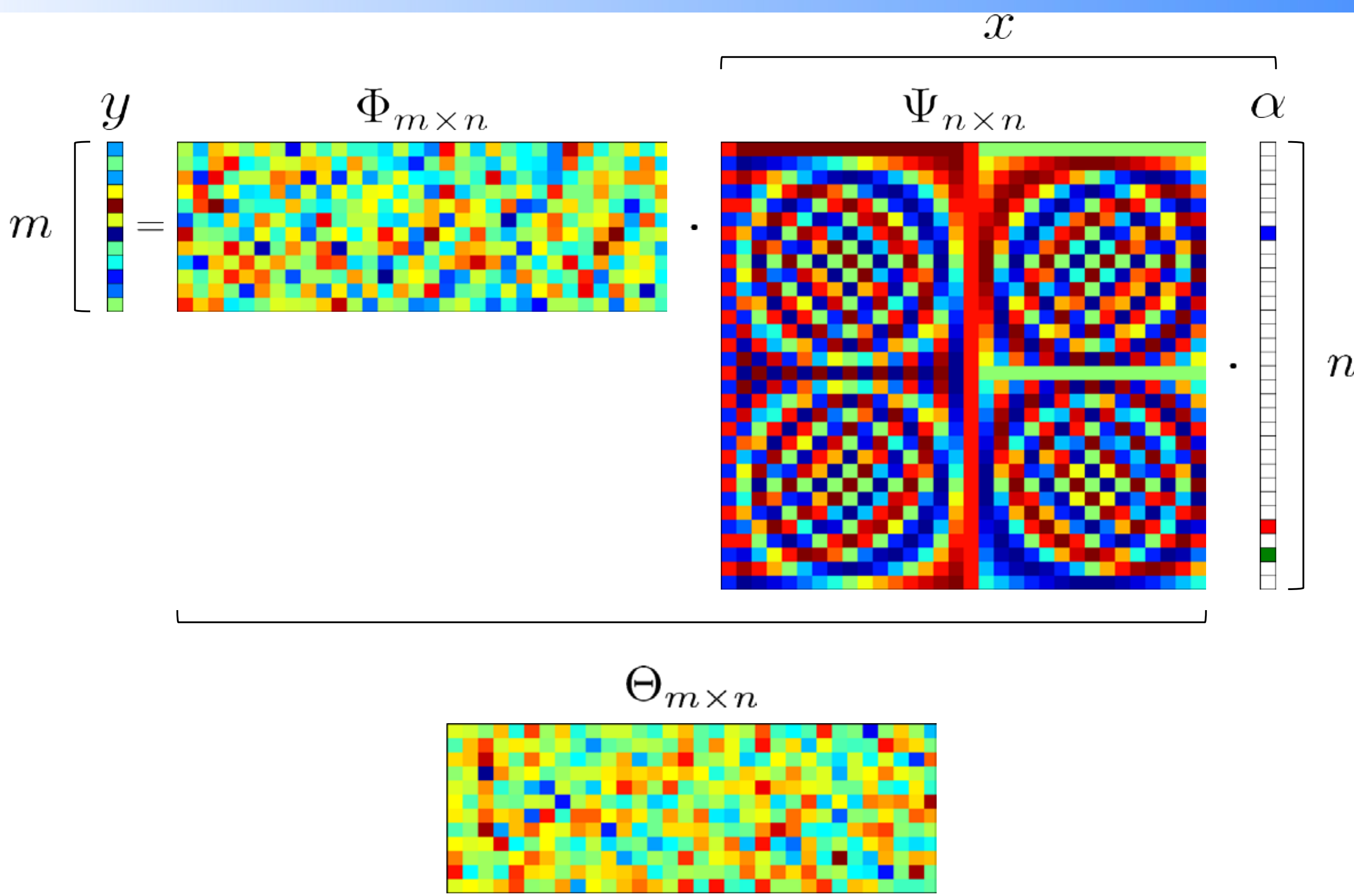
> a n-dimensional, k-sparse signal may be *acquired in a compressed form directly at the A/D interface*, thus allowing a sample rate reduction.

- **Input Vector**: $x \in \mathbb{R}^n$ (e.g. Nyquist rate samples), $x = \Psi_{n \times n}\alpha$, where $\Psi_{n \times n}$ is the *sparsity basis, α* is k-sparse.
- **Sampling Operator**: Let $\Phi_{m \times n}$ be a linear operator with $m < n$ that is *incoherent* to Ψ, i.e., the *coherence*:

$$\mu(\Phi, \Psi) = \sqrt{n} \max_{j,l} |\langle \phi_{j,\cdot} \psi_{\cdot,l} \rangle|$$

is as small as possible, i.e. *the rows of Φ are properly aligned to the sparsity basis vectors to always collect information at every sample*



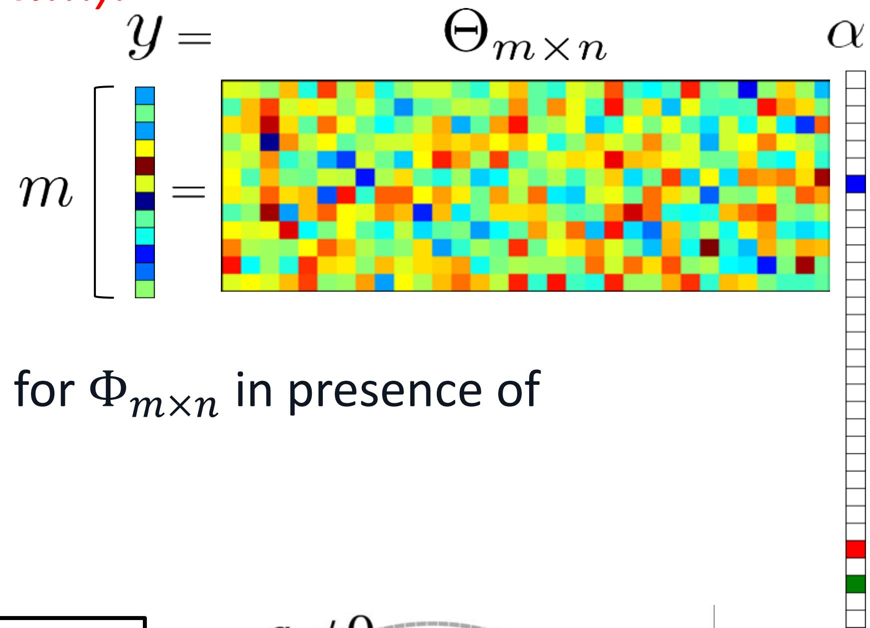


- **Sampling=Projection:** We will measure $y = \Phi_{m\times n}\, x = \Phi_{m\times n}\, \Psi_{n\times n}\alpha = \Theta_{m\times n}\alpha$
  - The linear operator $\Theta_{m\times n}$ clearly performs a dimensionality reduction since it shrinks $\mathbb{R}^n \rightarrow \mathbb{R}^m$.

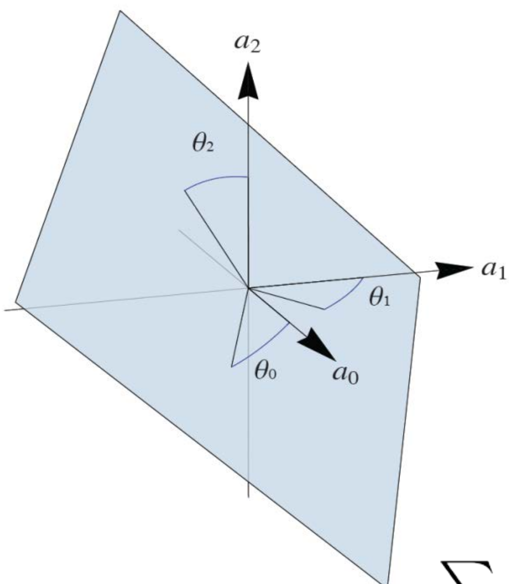
Compressed Sensing: a few issues

1. Solving $y = \Theta\alpha$ in x has infinitely many solutions in general because it is an ***underdetermined*** *system of equations ($m < n$!!!).*

How can one solve this????

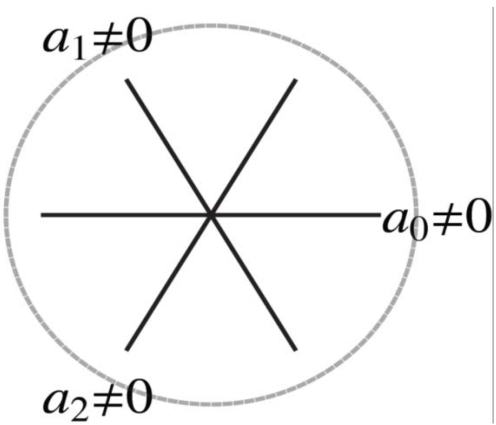


2. Is there a way for ensuring the "best choice" for $\Phi_{m\times n}$ in presence of a projection independent noise?



One needs to choose the projection plane to maximize the projection of each vector (i.e. "conserve" its length and energy)!!

choose $\theta_0, \theta_1, \theta_2$ such that
$$\sum_j \sin^2\theta_j = 1 \text{ and } \cos\theta_j = \max\ \forall j$$

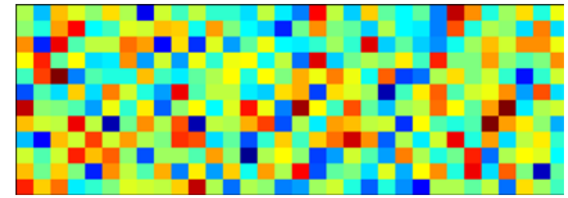


- **<u>Second issue:</u>** (*Restricted Isometry Property - RIP*) one requires that matrix $\Phi_{m\times n}$ is "almost an isometry", i.e. formally that, it $\exists\ \delta_k \in [0,1]$ such that

$$(1-\delta_k)\|\alpha\|_2^2 \le \|\Phi\alpha\|_2^2 \le (1+\delta_k)\|\alpha\|_2^2$$

i.e. the "energy of α" is almost conserved by the projection

$\Phi_{m\times n}$



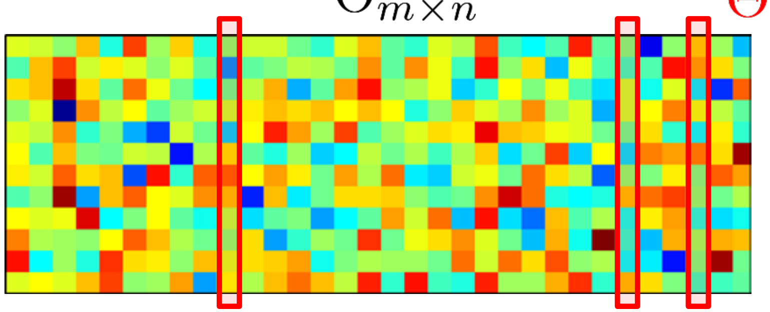
- **<u>First issue:</u>** (*Sparse Signal Recovery Problem*) one can obtain the solution by solving the following *optimization problem*

$$\min_{\alpha\in\mathbb{R}^n} \|\alpha\|_0 = \min_{\alpha\in\mathbb{R}^n} \sum_{l=0}^{n-1} \#\{\alpha_l | \alpha_l \neq 0\}\ s.t.\ y = \Theta\alpha$$

Is this solvable?

<u>Remark</u>: When x is k-sparse in Ψ, only k of n columns of Θ are considered: let S be the support of α, $y = \Theta\alpha = \Theta_S\,\alpha_S$.

$\Theta_{m\times n}$ $\Theta_{\cdot,S}$



If an oracle **told us the support S** and $k < m < n$, we would have to solve an ***overdetermined*** *system of equations* by least squares (easy and polynomial-time at worst)

- **Unfortunately S is unknown in general**, and guessing it would take $\binom{n}{k}$ attempts – given that k is known.
  - This is a *combinatorial* problem, **it is NP-hard in general**.
  - Take $n = 1024, k = 8$.
    - Worst-case sparse support search (at $10^{-2}\ s/\text{instance}$): $2.917 \times 10^{17}\ s$.
    - Age of the universe: $4.339 \times 10^{17}\ s$ ... it is "a bit" unpractical!!

- **"Engineering approach"**:

1. Which problem can we solve?
2. Are there cases in which **the problem we can solve has** the same solution of **the problem we must solve**?

- Consider the *convex optimization* problem:

$$\min_{\alpha \in \mathbb{R}^n} \|\alpha\|_1 = \min_{\alpha \in \mathbb{R}^n} \sum_{l=0}^{n-1} |\alpha_l| \ s.t.\ y = \Theta\alpha \quad \text{(basis pursuit problem)}$$

- *Linear programming* problem (easy and polynomial-time).
- Is there *a set of cases* for which this solves the sparse signal recovery problem?

- **Good News**: *The minimum* 1*-norm solution is also the sparsest* (minimum 0-norm) *solution*, i.e.,

$$\hat{\alpha}_{(\ell_1)} = \hat{\alpha}_{(\ell_0)} = \alpha$$

provided that:

$$m \geq C(\Theta) k \log\left(\frac{n}{k}\right) \qquad \text{(in practice } m \approx 4k\text{)}$$

where $C(\Theta)$ depends on the *structure* of $\Theta = \Phi\Psi$.

- There is a flourishing literature on algorithms that find solutions to basis pursuit/sparse signal recovery problems.

Basis Pursuit (BP): Basis Pursuit with Denoising (BPDN):

$$\min_{\alpha \in \mathbb{R}^n} \sum_{l=0}^{n-1} |\alpha_l| \quad \text{subject to} \quad y = \Theta\alpha$$

$$\min_{\alpha \in \mathbb{R}^n} \sum_{l=0}^{n-1} |\alpha_l| \quad \text{subject to} \quad ||y - \Theta\alpha||_2^2 \leq \epsilon^2$$

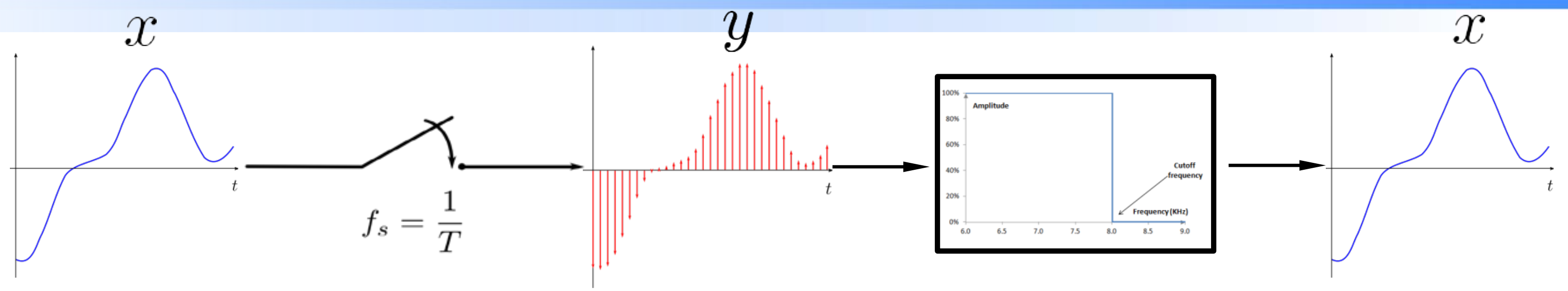
1. In addition to Min-L_1, we have *CoSaMP, Orthogonal Matching Pursuit, Iterative Hard/Soft Thresholding, Smoothed* L_0, … They converge, but may be slow…
2. Designing **real-time algorithms** which rapidly converge to the sparsest solution at very large problem sizes (large n) is **an *open challenge*.**

Decoding strategies divided in two family

- **minimization based procedures**
  among all the possible counterimages of the vector $y = \Theta\alpha$ we look for the "most sparse", i.e., the one with is minimum $\|\alpha\|_{0/1}$.
  reference algorithm: Min-L_1
- **iterative support guessing procedures**
  Alternate a rough, non-necessarily sparse, solution of $y = \Theta\alpha$ (pseudoinverse) with an estimate of support of α (for example by thresholding on the magnitudes of the components of the temporary solution) that is then exploited in a pseudoinverse-based step refining the value.
  reference algorithm: CoSaMP


Is anybody
awake?

Shannon



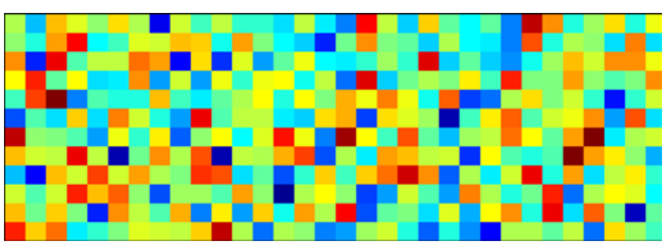
CS

$x = \Psi\alpha \rightarrow \boxed{\Phi} \rightarrow y \rightarrow \boxed{\min\ \ell_1} \rightarrow \hat{\alpha}\ s.t.\ \hat{x} = \Psi\hat{\alpha} \simeq x$

- must be sparse
- must preserve energy (RIP)
- Ψ and Φ must be incoherent
- $m \ll n$, i.e. less resources for acquisition
- must be "solvable" (i.e. Min-L_1); m must be large enough wrt k

- **Note**: The most universal choice is to choose Φ at random, since it guarantees that Φ is incoherent for every Ψ. "Most convenient": are Gaussian RV

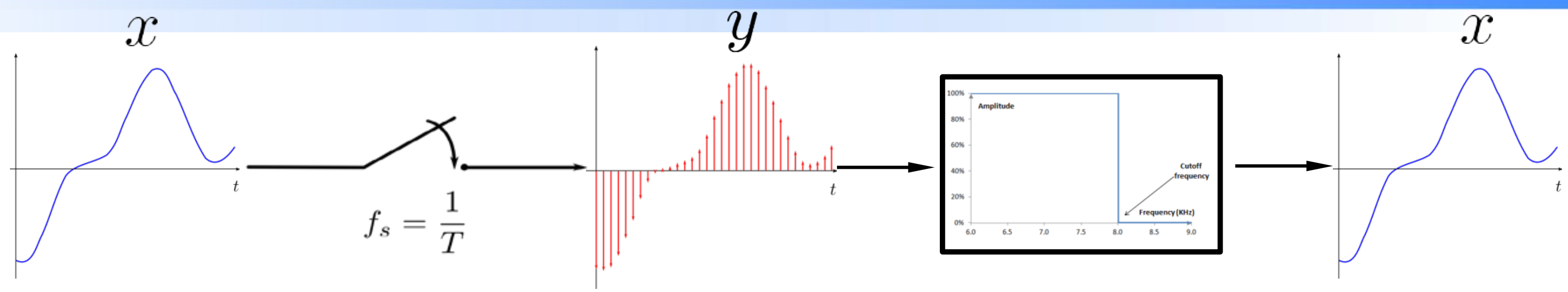
Φ



$$\phi_{j,l} \sim \mathcal{N}(0, \Sigma_\Phi)\ (i.i.d.)$$
$$\Sigma_\Phi = \mathbf{E}\left(\Phi^\dagger \Phi\right) = I_n/n$$

RIP can be proven

Shannon



CS

$x = \Psi\alpha \rightarrow \boxed{\Phi} \rightarrow y \rightarrow \boxed{\min\ \ell_1} \rightarrow \hat{\alpha}\ s.t.\hat{x} = \Psi\hat{\alpha} \simeq x$

must be sparse

must preserve energy (RIP)

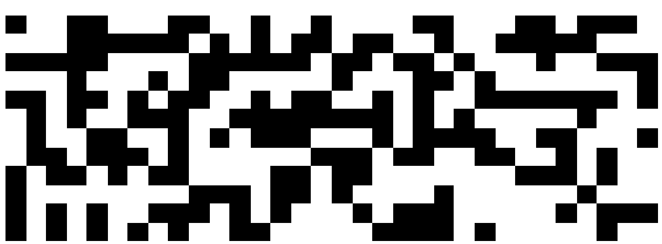
Ψ and Φ must be incoherent

$m \ll n$, i.e. less resources for acquisition

must be "solvable" (i.e. Min-L_1); m must be large enough wrt k

- *Binary antipodal* ($\{+1, -1\}$) are also used as hardware-friendly solutions (more later)

Φ

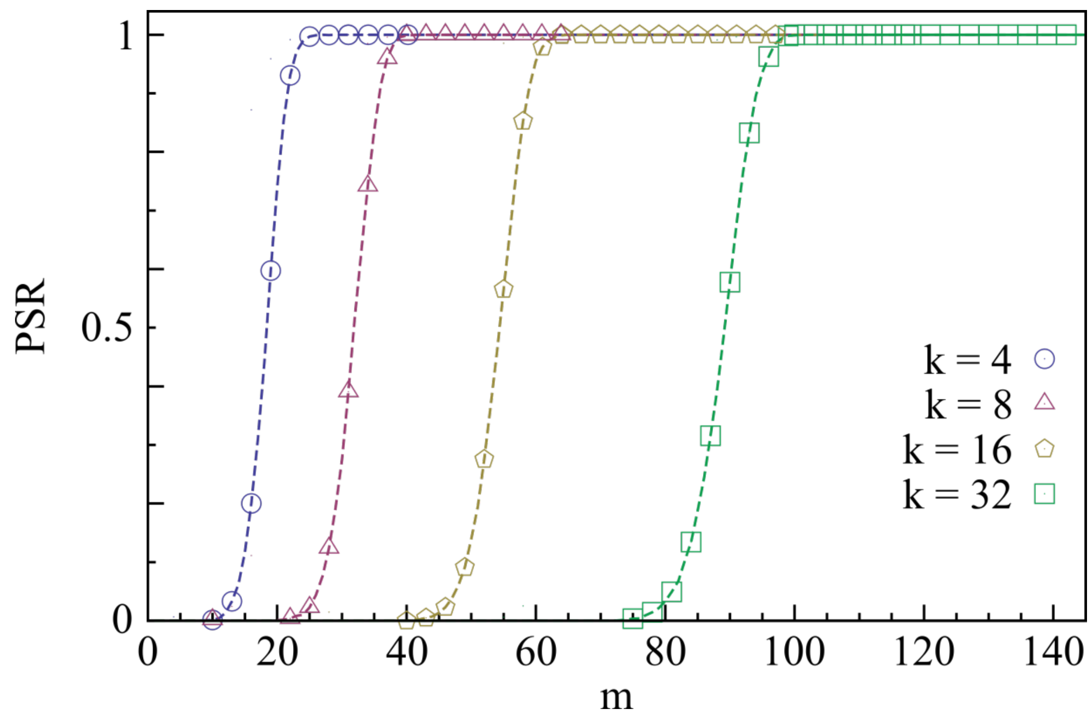


$$\phi_{j,l} = \begin{cases} 1/\sqrt{n}, & p = 1/2 \\ -1/\sqrt{n}, & 1 - p = 1/2 \end{cases}$$

- **Input signal**: unit-energy, k-sparse vectors $x = \Psi\alpha$ (Non-zero entries in α are i.i.d. Gaussian RVs with zero mean and variance $1/k$)
- $\Psi_{n\times n}$ is the **discrete Fourier basis** ($n = 256$)
- **Measurement matrix**: $\Phi \in \mathbb{R}^{m\times n}$ ~ i.i.d Gaussian RVs with zero mean and variance $1/n$.
- **Reconstruction Algorithm:**

$$\hat{\alpha} = \arg\min\left|\left|\alpha\right|\right|_1 \; s.t.\, y = \Phi\Psi\alpha$$

i.e. Min-L_1



- **Figure of Merit**:

$$\mathrm{PSR} = \mathrm{P}\left[\frac{\|\alpha\|_2^2}{\|\hat{\alpha}-\alpha\|_2^2} > 10^6\right]$$

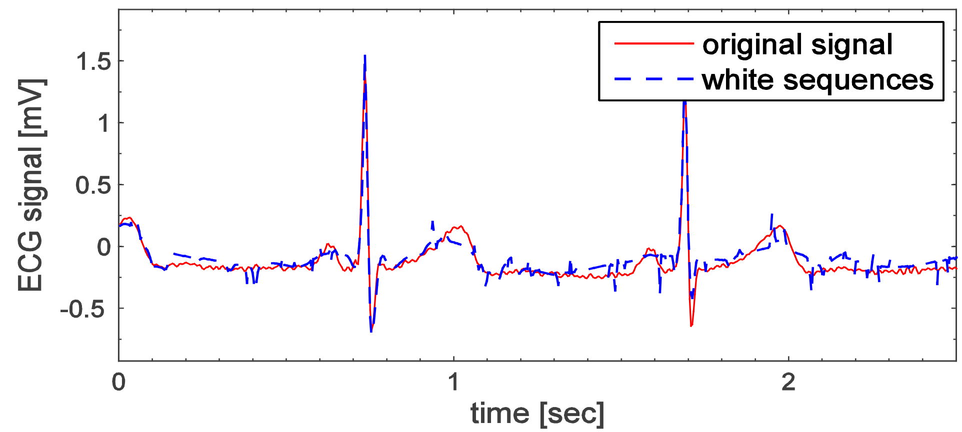
Probability of Successful Reconstruction

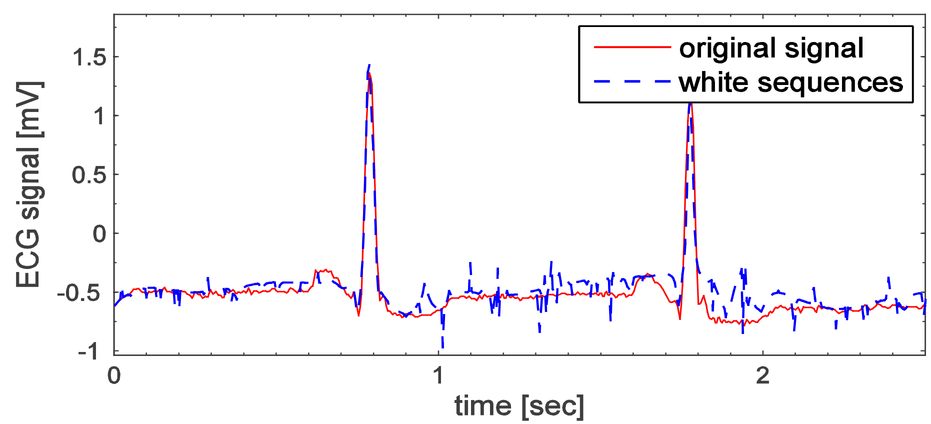
- **CS of ECG signals**:
  $m = 55$ measurements
  Nyquist-rate samples:
  $$n = 256\ (f_n = 256\ \mathrm{Hz})$$
- **Compression rate** (n/m): 4.7.
  - Ψ is a dictionary of Gabor time-frequency atoms
  - $\Phi_{55\times256}$ is an antipodal (pseudo)random measurement matrix
  - Intrinsic **SNR=17 dB**
- **Figure of Merit**:

$$\mathrm{ARSNR_{dB}} = \mathbf{E}\left[20\log_{10}\left(\frac{||\alpha||_2}{||\alpha-\hat{\alpha}||_2}\right)\right]$$

(Average Reconstruction SNR)

This case: **ARSNR ≈ 15 dB**.





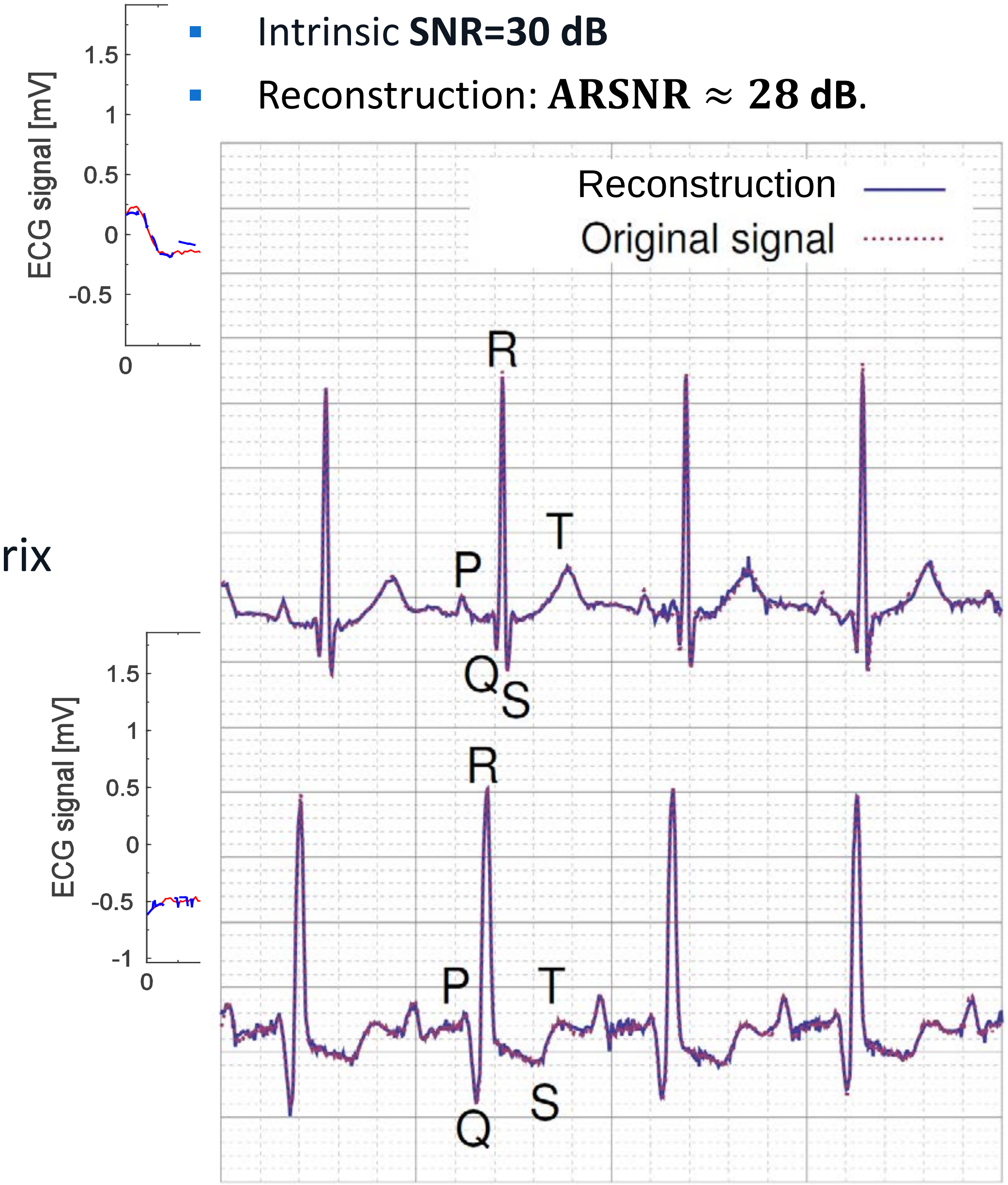
- **CS of ECG signals**:
  $m = 55$ measurements
  Nyquist-rate samples:
  $$n = 256\ (f_n = 256\ \text{Hz})$$
- **Compression rate** (n/m): 4.7.
  - Ψ is a dictionary of Gabor time-frequency atoms
  - $\Phi_{55\times 256}$ is an antipodal (pseudo)random measurement matrix
  - Intrinsic **SNR=17 dB**
- **Figure of Merit**:

$$\mathrm{ARSNR_{dB}} = \mathbf{E}\left[20\log_{10}\left(\frac{||\alpha||_2}{||\alpha-\hat{\alpha}||_2}\right)\right]$$

(Average Reconstruction SNR)

This case: $\mathbf{ARSNR} \approx \mathbf{15}$ **dB**.

- $m = 80$; **Compression rate** (n/m): 3.5.
- Intrinsic **SNR=30 dB**
- Reconstruction: $\mathbf{ARSNR} \approx \mathbf{28}$ **dB**.

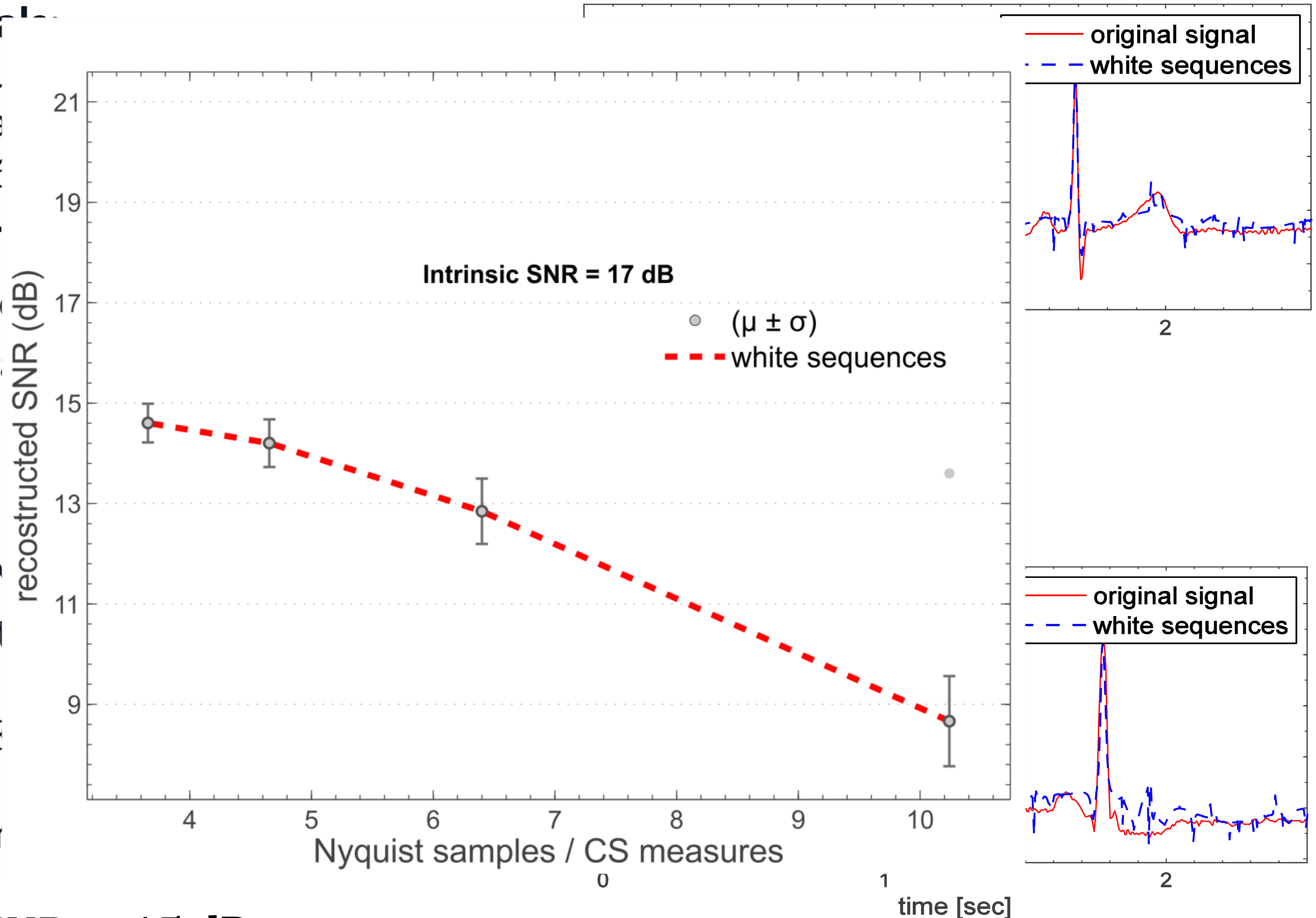


A Real World Example

- **CS of ECG signa**
  $m = 55$ measu
  Nyquist-rate sa
  $n = 25($
- **Compression r**
  - Ψ is a dicti
    frequency
  - $\Phi_{55\times256}$ is
    (pseudo)ra
  - Intrinsic **SN**
- **Figure of Meri**

$$\mathrm{ARSNR}_{\mathrm{dB}} = \mathbf{E}$$

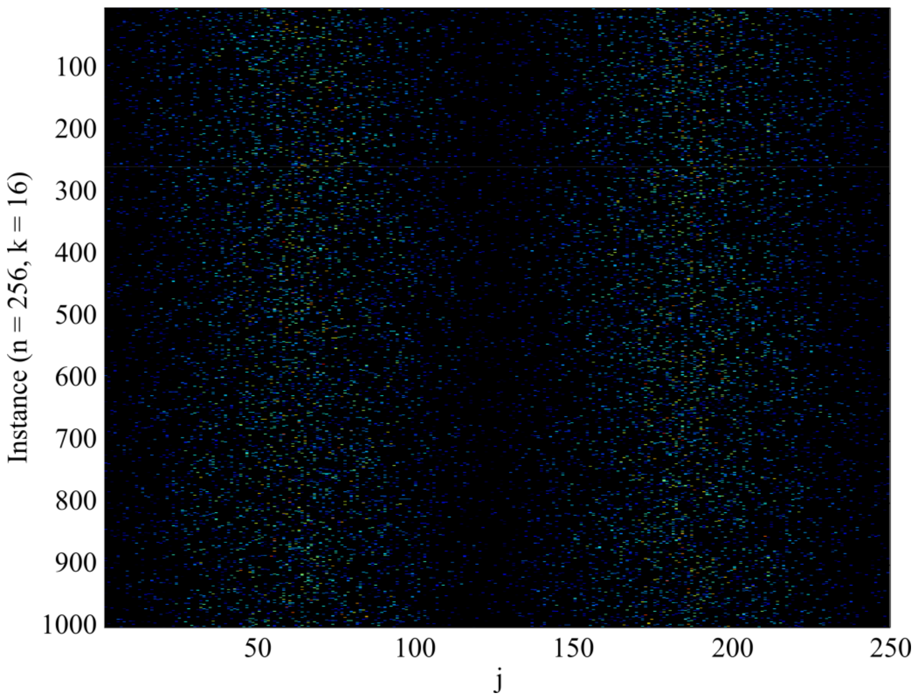
(Average Recor

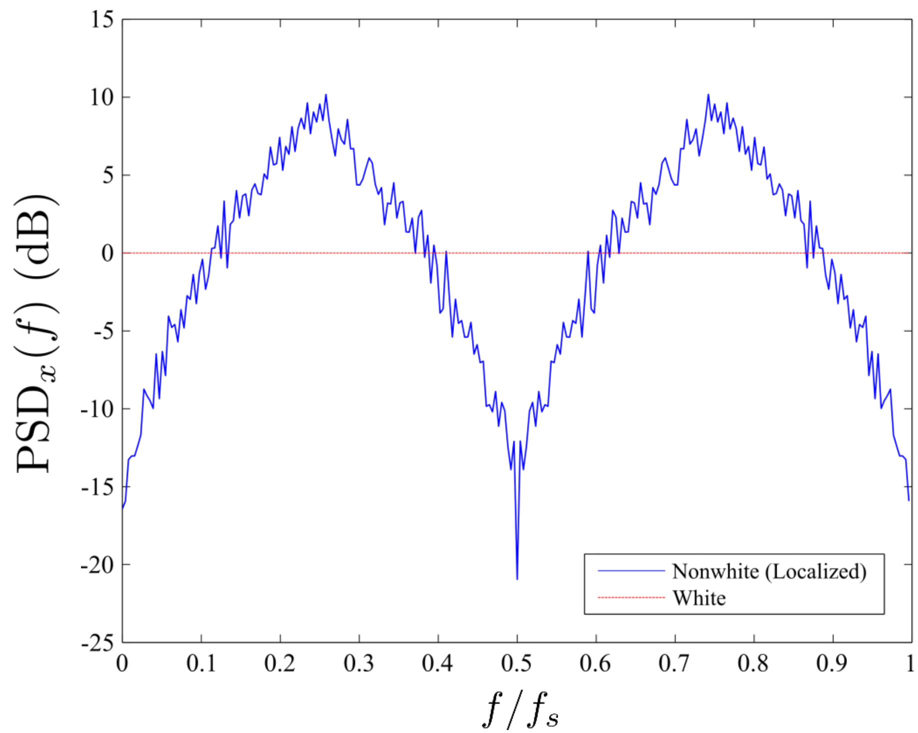


This case: **ARSNR ≈ 15 dB.**

- Standard CS is *universal* for all k-sparse signals in Ψ *regardless of their support*: The implicit assumption is that the signal is white – this is almost always false!

- Consider a wideband natural signal generated from a *nonwhite* process, whose instances are k-sparse (w.r.t. Fourier):

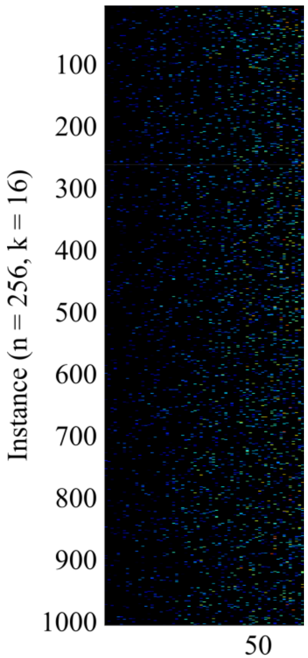


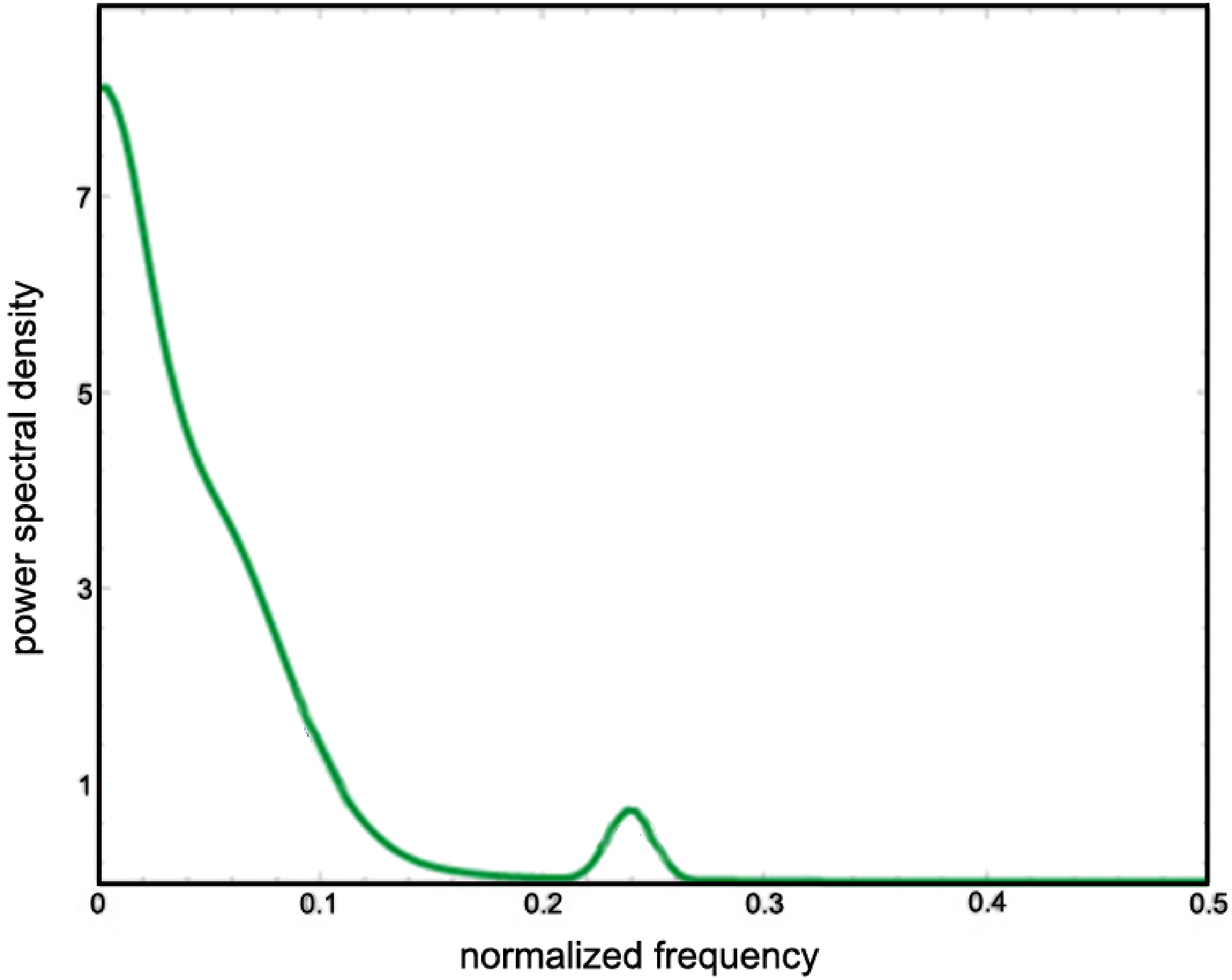


- Some indices are evidently more frequent in the k-sparse signal representation!

CAN WE DO BETTER? LOCALIZED SIGNALS I

- Standard CS is *universal* for all k-sparse signals in Ψ *regardless of their support*: The implicit ass … se!
- Consider a … , whose instances a …





- Some indices are evidently more frequent in the k-sparse signal representation!

- Intuition:
  - *Best* random projection vectors for *white* processes (worst case, universal) → *white* (i.i.d.) Gaussian or antipodal symbols.
  - *Best* random projection vectors for **localized** and stationary processes → *nonwhite* Gaussian or antipodal symbols **chosen to maximize the raked energy**.

- We define **rakeness** as:

$$\rho(\phi, x) = \mathrm{E}_{\phi_j, x}\left[\left|\langle \phi_j, x\rangle\right|^2\right]$$

i.e., the **expectation** of the affinity of random projection vectors ϕ_j **to collecting the energy** in x.

- And introduce the *maximum rakeness* optimization:

$$\max_{\phi_{j,\cdot}} \rho(\phi_{j,\cdot}, x)$$

$$\text{s.t.} \begin{cases} \langle \phi_{j,\cdot}, \phi_{j,\cdot} \rangle = e \\ \rho(\phi_{j,\cdot}, \phi_{j,\cdot}) \leq r^2 e \end{cases}$$

- **Remark:** The first constraint ensures that the vectors ϕ_j have constant energy, while the second ensures that they **are still random enough** (to guarantee incoherence)

Why this is needed?

Rakeness: Intuition in the "Land of Vectors"

- We define **rakeness** as:

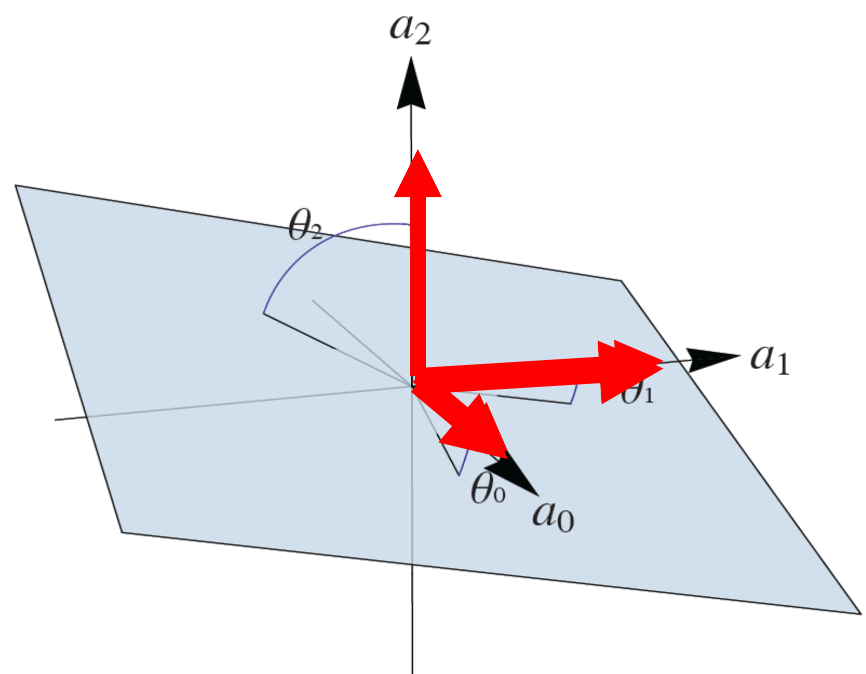
$$\rho(\phi, x) = \mathrm{E}_{\phi_j, x}\left[\left|\langle \phi_j, x \rangle\right|^2\right]$$

i.e., the **expectation** of the affinity of random projection vectors ϕ_j **to collecting the energy** in x.

- And introduce the *maximum rakeness* optimization:

$$\max_{\phi_{j,\cdot}} \rho(\phi_{j,\cdot}, x) \quad \text{s.t.} \begin{cases} \langle \phi_{j,\cdot}, \phi_{j,\cdot} \rangle = e \\ \rho(\phi_{j,\cdot}, \phi_{j,\cdot}) \leq r^2 e \end{cases}$$

- The signal is localized since the probability that the non-zero component is aligned to a_0 or a_1 is larger than it is aligned to a_2 (the first 2 happen more frequently)
- To maximize rakeness one should project on a plane aligned with a_0 and a_1, but this would spoil the capability to collect information along a_2



Remark: Incoherence between input signal and sampling vectors is the equivalent of guaranteeing to collect information in every direction (generalization of duality time-frequency for sampling with δ in the time domain)

Sensing matrix generator used on CS for signal localized in Fourier Domain

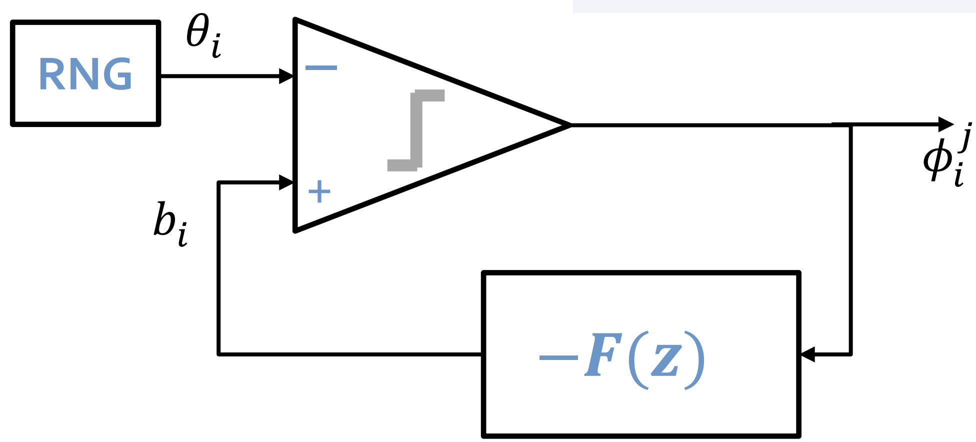
- The power spectrum of the generated sequence is fixed by the coefficient of the feedback liter.
- Slices of length T of this process can be used as projection waveforms in an RMPI architecture.

features:

- time-invariant linear filter
- finite number of coefficient m in the filter:

$$\theta_i \to U(-1,1)$$
$$b_i \in [-1,1]$$
$$F(z) = \sum_{j=0}^{B-1} f_j z^{-j}$$
$$F(z) = \sum_{j=0}^{B-1} |f_j| \leq 1$$

$$\hat{\phi}(f) = \frac{\left|1 + F(e^{2\pi i f})\right|^{-2}}{\int_{-1/2}^{1/2} \left|1 + F(e^{2\pi i f})\right|^{-2} \mathrm{d}f}$$

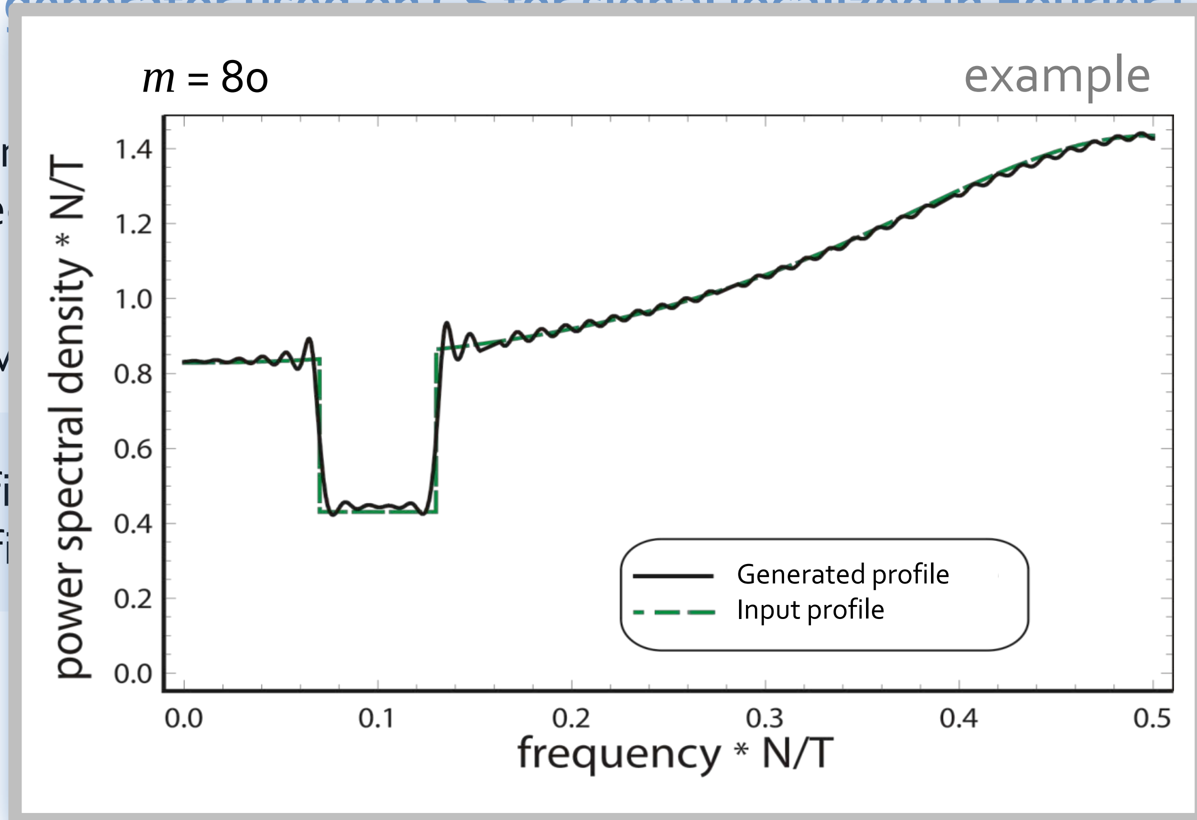


Sensing matrix generator used on CS for signal localized in Fourier Domain

- The power spectrum ... coefficient of the fe...
- Slices of length T ... waveforms in an RM...

features:

- time-invariant linear fi...
- finite number of coeffi...



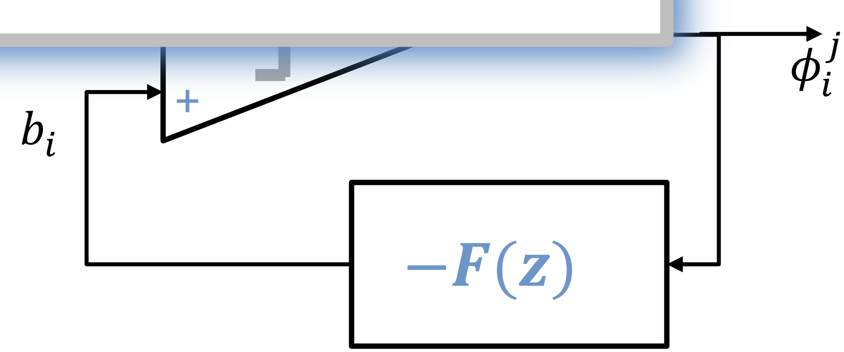
$U(-1,1)$

$-1,1]$

$\sum_{j=0}^{B-1} f_j z^{-j}$

$\sum_{j=0}^{B-1} |f_j| \leq 1$

$$\hat{\phi}(f) = \frac{\left|1 + F(e^{2\pi i f})\right|^{-2}}{\int_{-1/2}^{1/2} \left|1 + F(e^{2\pi i f})\right|^{-2} \mathrm{d}f}$$



- **Input signal**: unit-energy, k-sparse vectors $x = \Psi\alpha$ (Non-zero entries in α are i.i.d. Gaussian RVs with zero mean and variance $1/k$)
- $\Psi_{n\times n}$ is the **discrete Fourier basis** ($n = 256$)
- **Measurement matrix**: $\Phi \in \mathbb{R}^{m\times n}$ ~ i.i.d Gaussian RVs with zero mean and variance $1/n$.
- **Reconstruction Algorithm:**

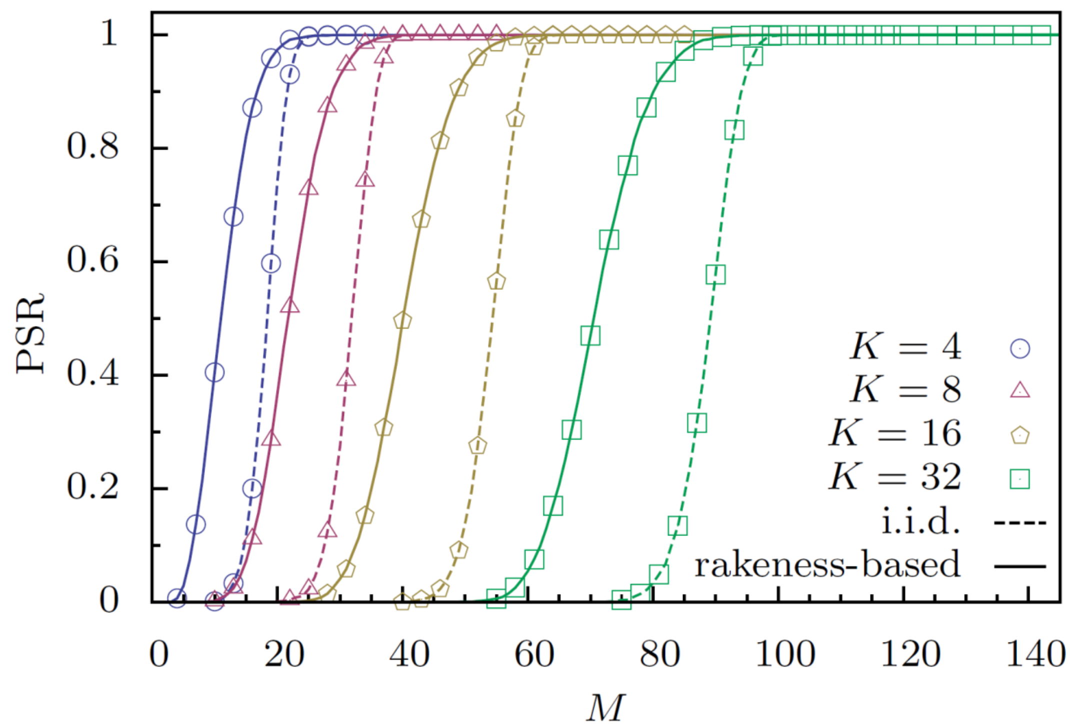
  $$\hat{\alpha} = \arg\min ||\alpha||_1 \; s.t.\, y = \Phi\Psi\alpha$$

  i.e. Min-L_1

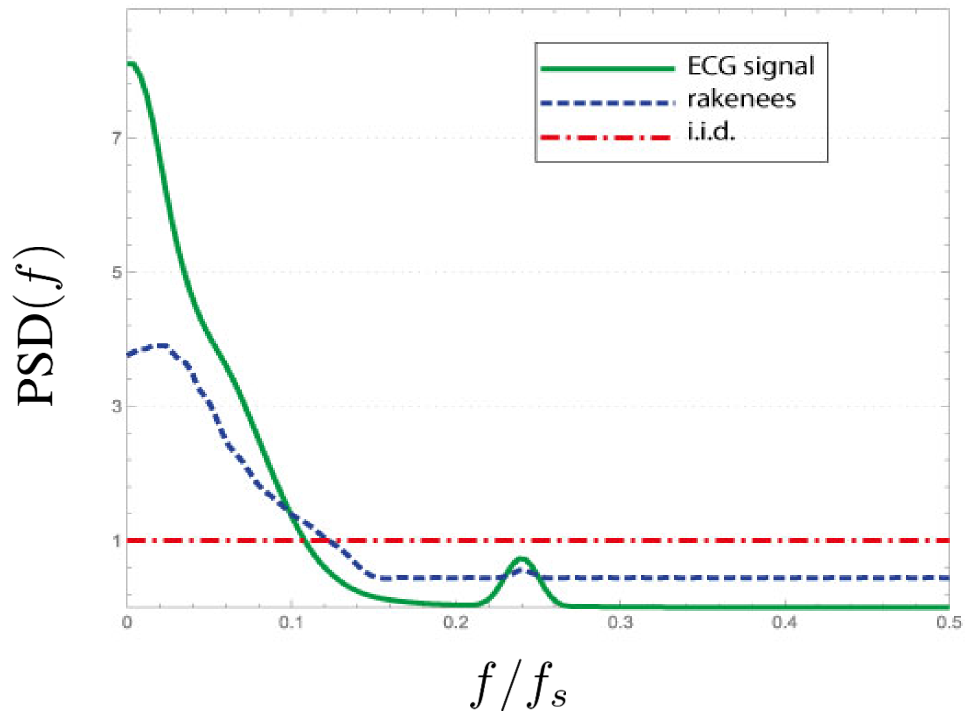
- **Figure of Merit**:

  $$\mathrm{PSR} = \mathrm{P}\left[\frac{\|\alpha\|_2^2}{\|\hat{\alpha}-\alpha\|_2^2} > 10^6\right]$$

  Probability of Successful Reconstruction

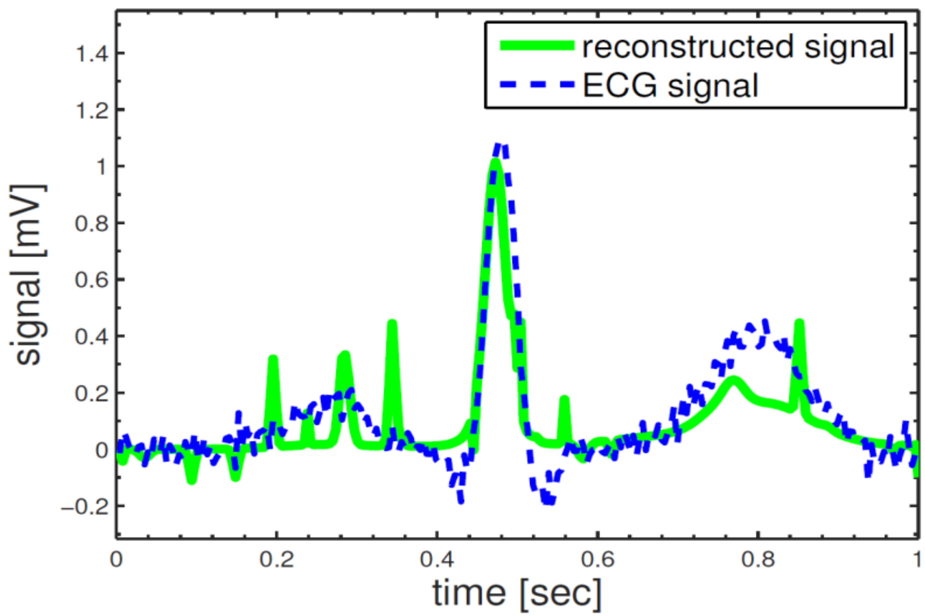


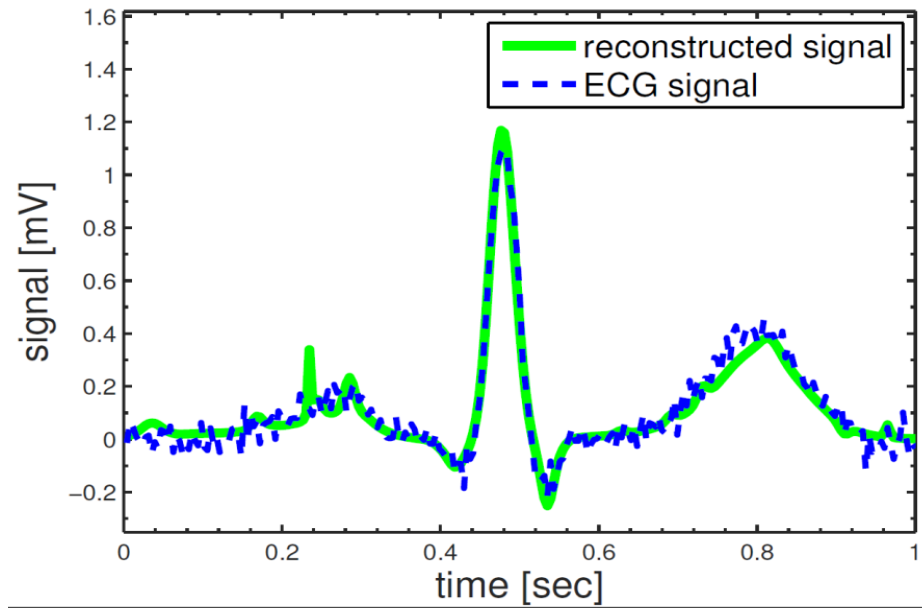
- We apply the rakeness-based design flow to the CS of ECG signals with antipodal random sequences:



n/m

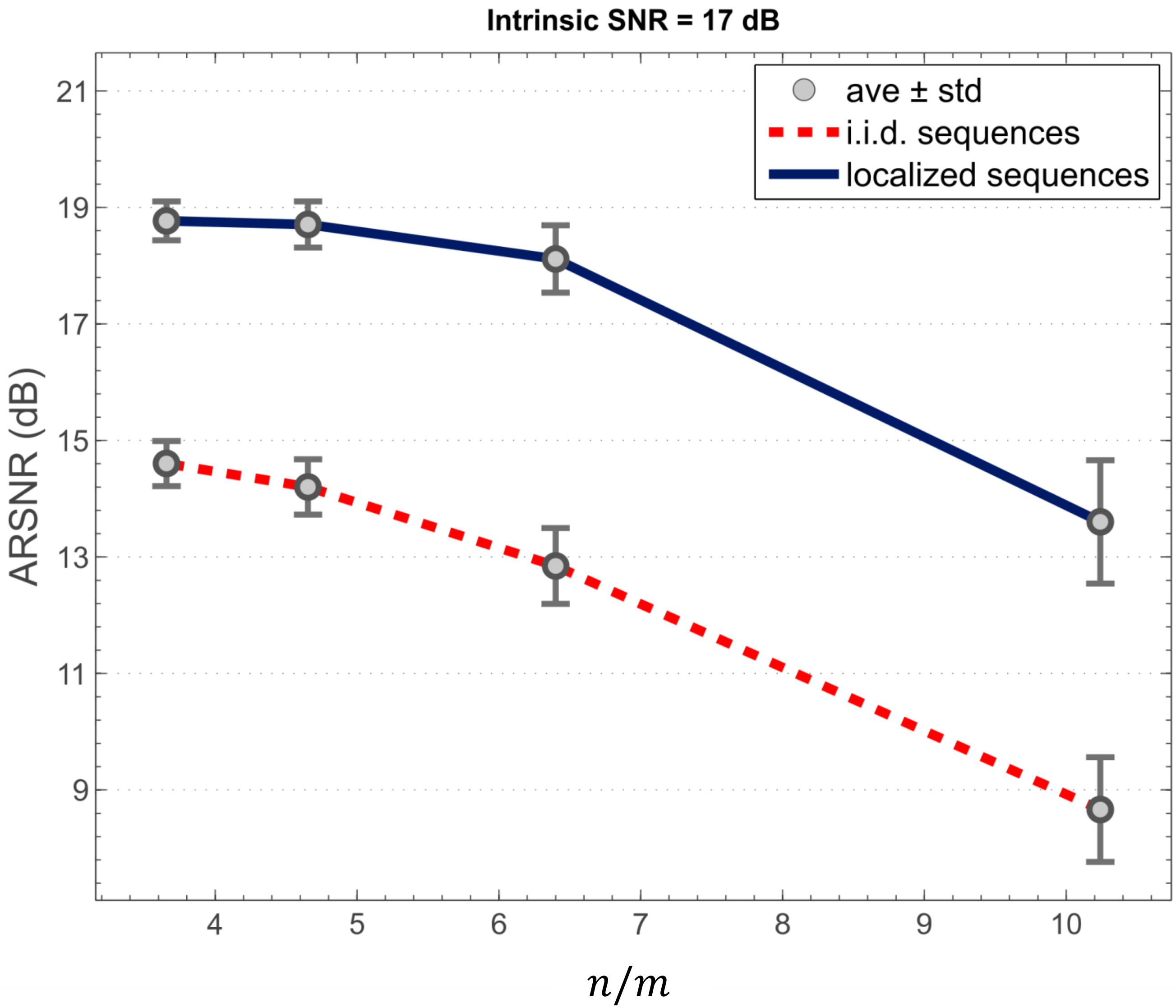
- We apply the rakeness-based design flow to the CS of ECG signals with antipodal random sequences:



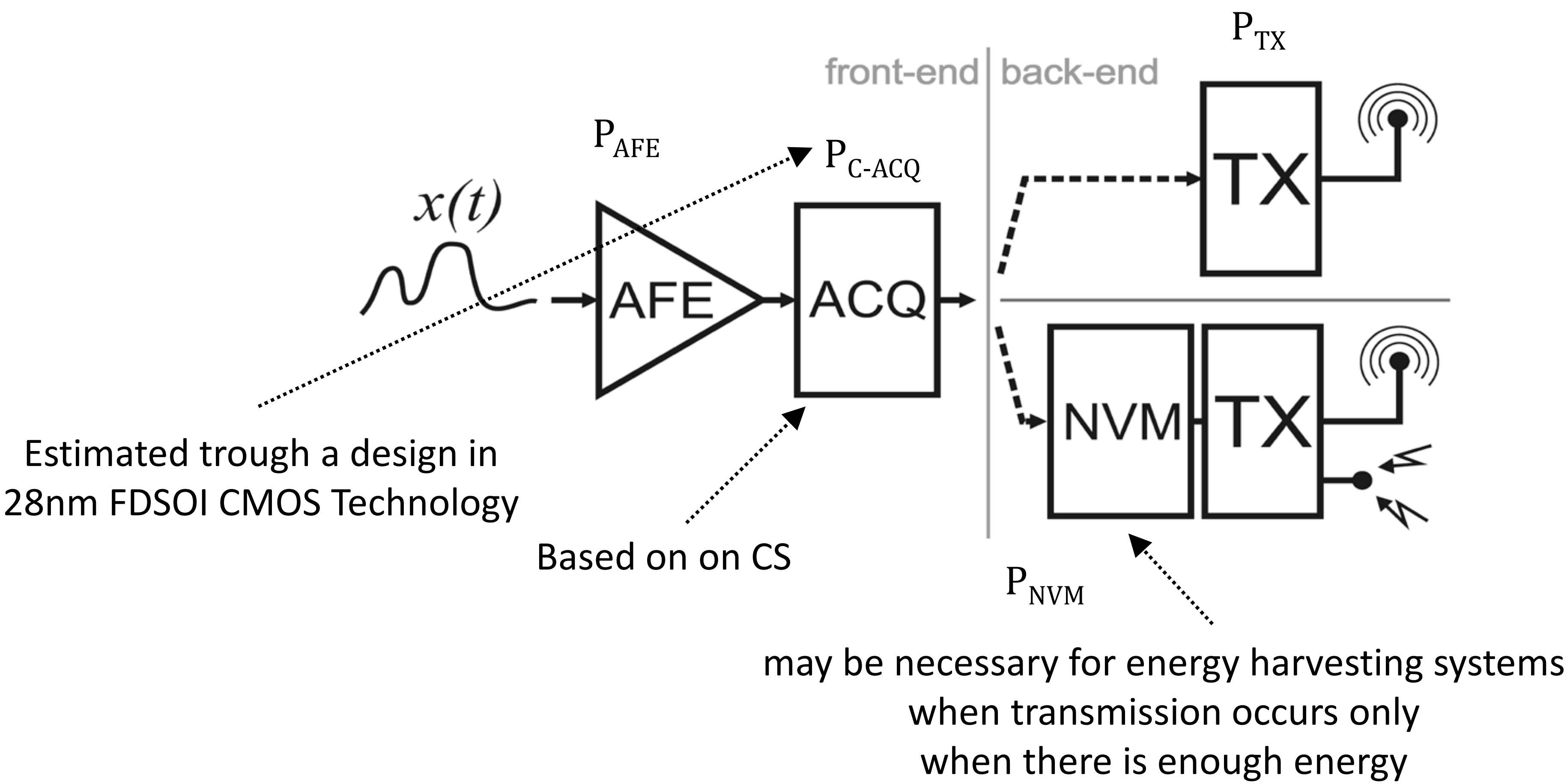


n/m

- We appl… ECG signals v…



Most of real (Bio)signal are sparse and compressed sensing is an excellent way to reduce the energy necessary for signal acquisition → natural application in body area network nodes!

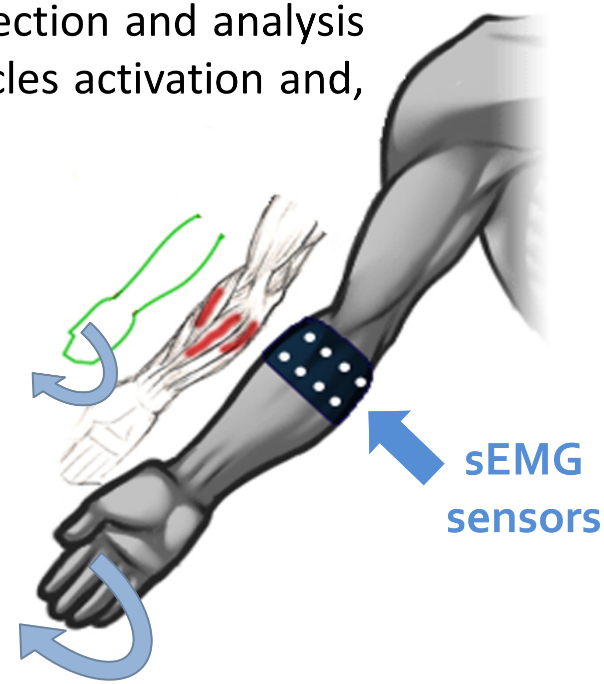


Hand Movement Recognition (collab with IIT@PoliTo)

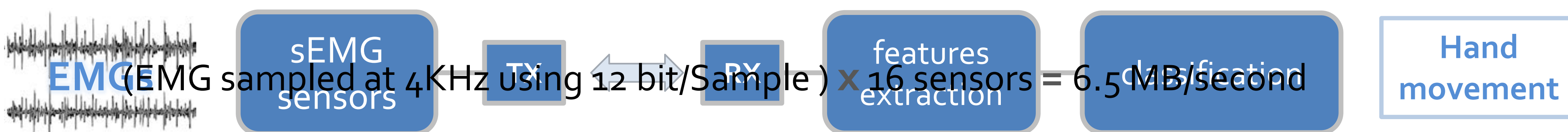
Surface ElectroMyoGraphy (sEMG) is the gold standard for the detection and analysis of muscle activation. EMGs gives access to information about muscles activation and, therefore, to movements which are being performed.

sEMG movement recognition is performed following 3 major phases:

- signal acquisition and digitalization
- features extraction
- classification



Adopting a Wireless BAN is particularly interesting to bring sEMG as an application in the real world. Which implicitly means that a huge amount of data must be transmitted every time.



(EMG sampled at 4KHz using 12 bit/Sample) x 16 sensors = 6.5MB/second

A technique to reduce the size of the transmitted data must be consider!!

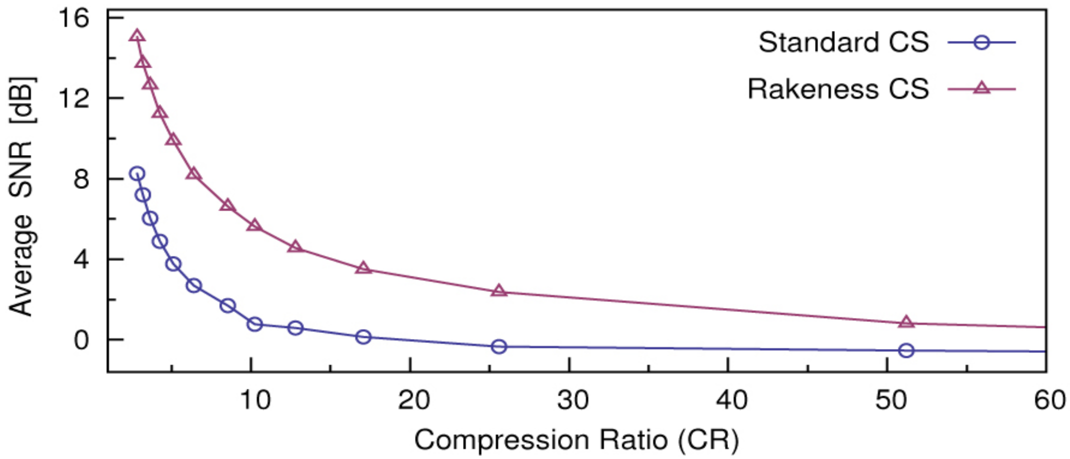
Compressed Sensing can reach this goal!!

Test On EMG data from Physionet

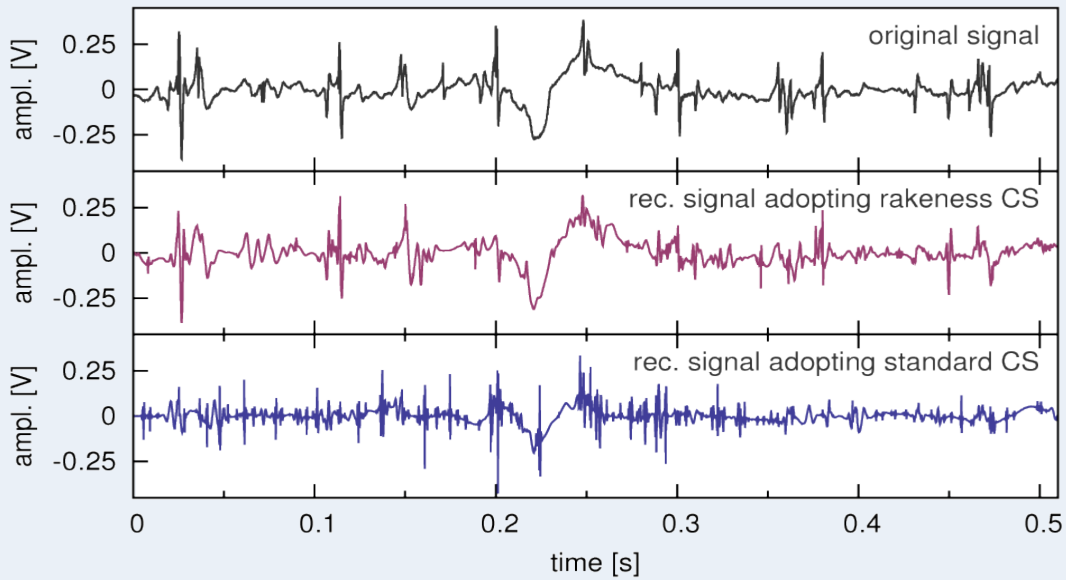
$$\text{Compression Ratio} = \frac{n}{m}$$

$$\text{Average SNR} = \mathrm{E}\left[\left(\frac{\|x\|_{l2}}{\|x-\tilde{x}\|_{l2}}\right)_{\mathrm{dB}}\right]$$

The rakeness-based approach presents a Average SNR increase up to 7 dB.



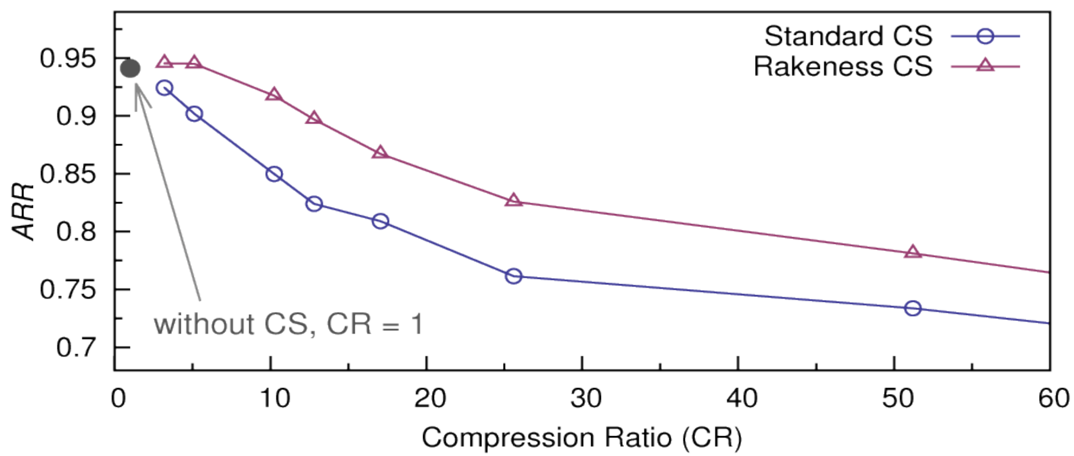
first 2048 samples of the original signal (top), the associated reconstructed signals adopting both rakeness-based CS (middle) and standard CS (bottom) with $m = 60$ which correspond to a CR= 8.5



sEMG Hand Movement Recognition Task

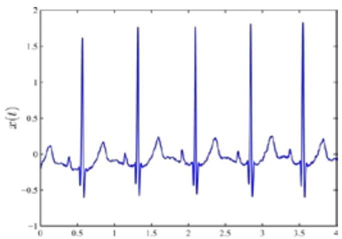
This task consisted in recognizing three well separated movements: 1. hand in neutral and relaxed position, 2. finger flexion (hand close as in a power grip), and 3. finger extension (hand opening).
To test the quality of the recognition the Average Recognition Rate was computed as:

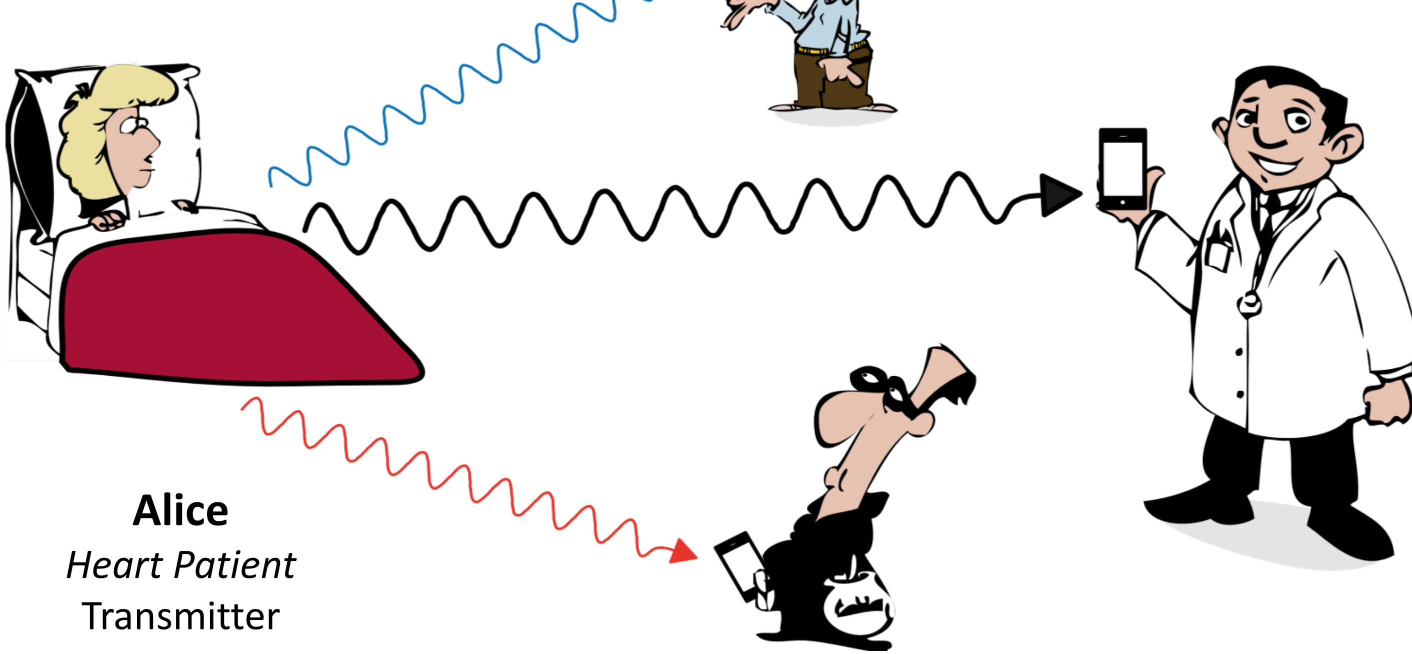
$$ARR = \frac{\text{number}_{correct\_recognition}}{\text{number}_{total\_samples}}$$



Results confirm the quality of the CS approach for the task of hand motion recognition. Indeed, depending on the specific application under study, very high levels of CR could be accepted. CS is particularly interesting when coupled with rakeness as **it improves the performance of the hand motion recognition system of about 7% with all CR from 5 to 60**. Furthermore, the **rakeness-based CS seems to maintain the performance at their top level with a CR up to 5**.

Remote ECG Monitoring



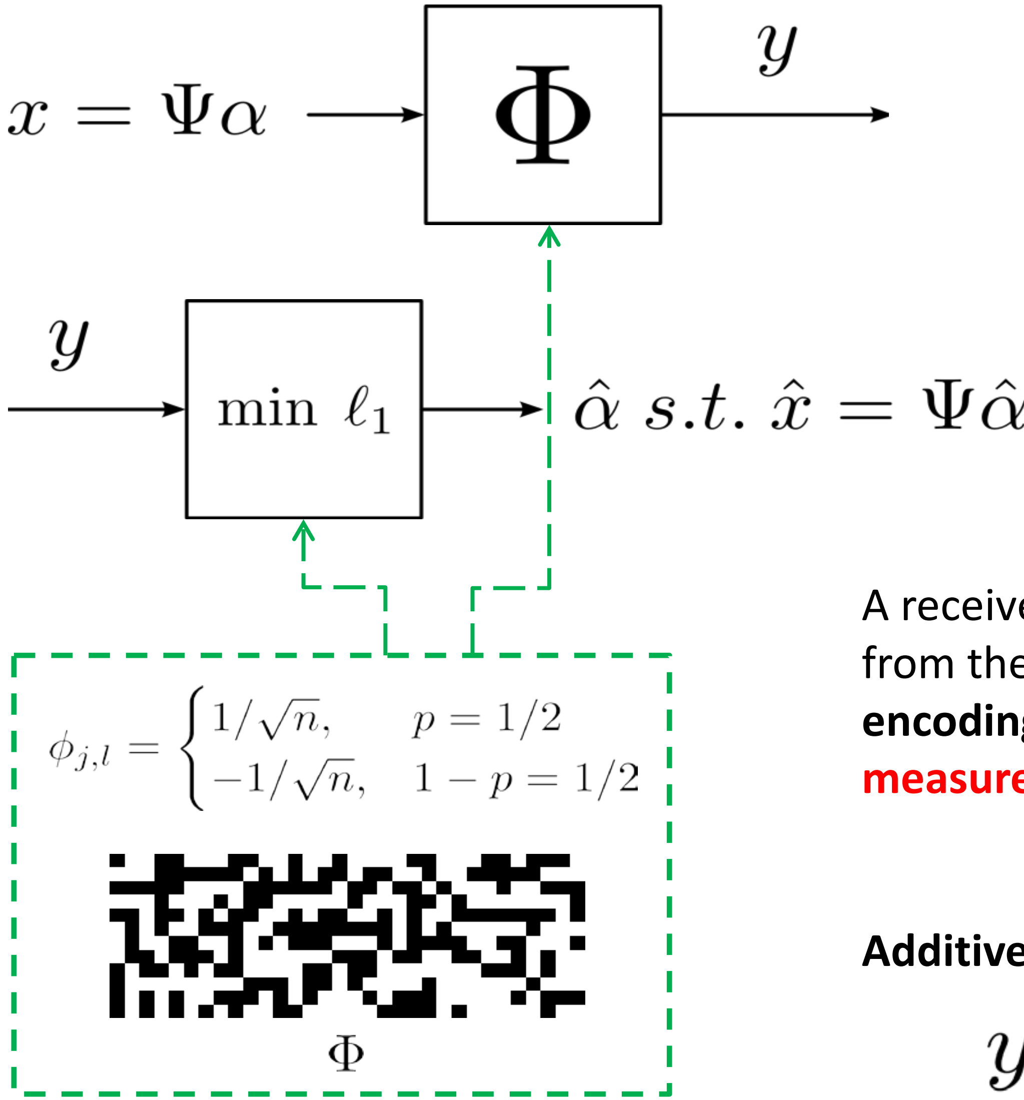


Steve
A Relative
Low-quality,
non-diagnostic ECG

Bob
Medical Doctor
High-quality ECG

Alice
Heart Patient
Transmitter

Marvin
Eavesdropper
No recoverable information/noise



Encoding Matrix

Acquisition = Encoding Stage = Transmitter

Min-L1 = Decoding Stage = Receiver

A receiver who wants **exact signal recovery** from the measurements y must know the **exact encoding matrix** Φ, or is otherwise subject to **measurement noise** as:

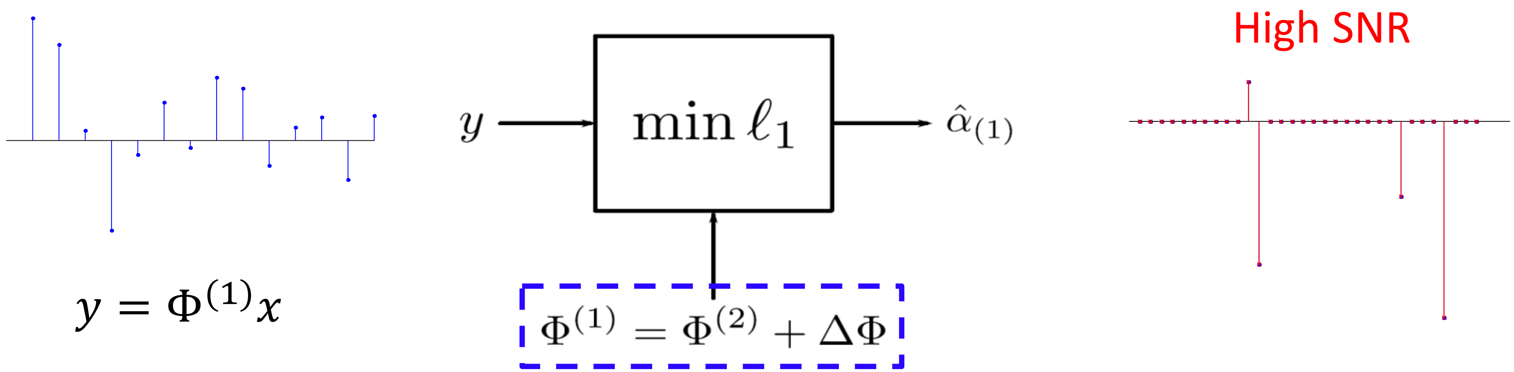
$$y = \Phi x + \nu$$

Additive perturbation model:

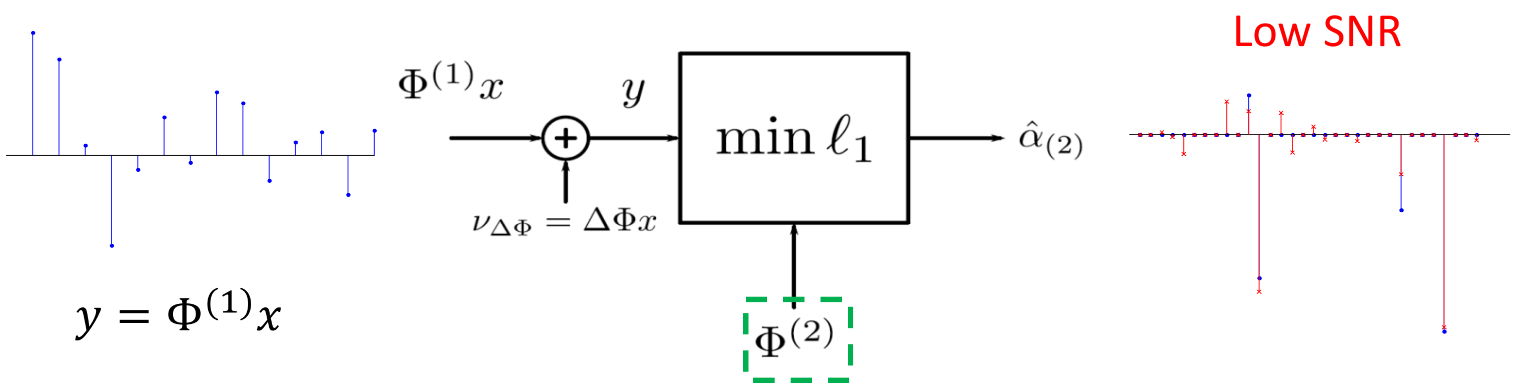
$$y = (\Phi x + \Delta\Phi)x$$

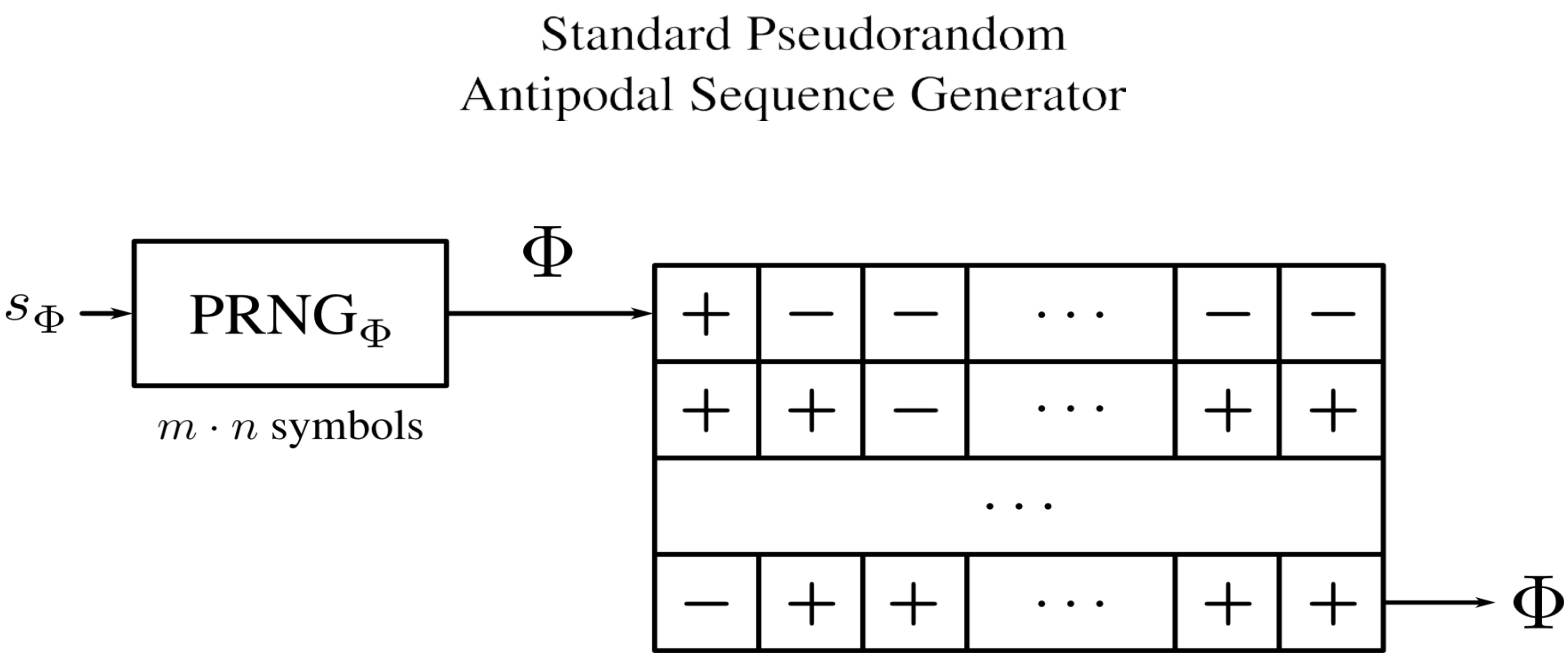
Two-class Compressed Sensing

- A **controlled** (pseudorandom) **perturbation $\Delta\Phi$** can be **exploited** to devise two decoder classes:
  - **First-class**: completely knows the true encoding matrix $\Phi^{(1)} = \Phi^{(2)} + \Delta\Phi$.



  - **Second-class**: knows $\Phi^{(2)}$ **but doesn't know $\Delta\Phi$!**



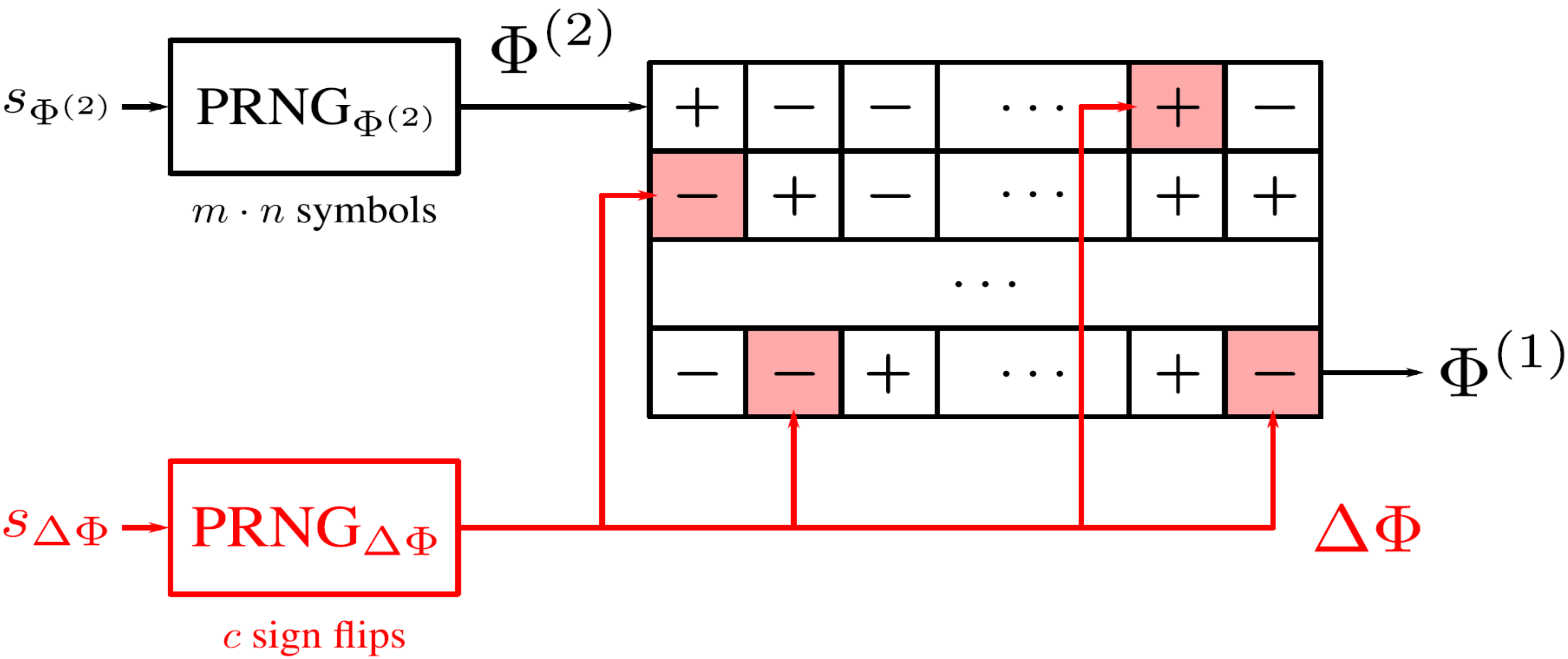


- A simple and effective choice to differentiate decoder classes is to **pseudorandomly invert the sign of c symbols in an $m \cdot n$ i.i.d. sequence**.
- **First-class decoders** are *unaffected*; **second class decoders** can only decode **noisy approximations** of the input. **The noise power is controlled by $\gamma = \frac{c}{m \cdot n}$.**

Pseudorandom Antipodal Sequence Generator

+

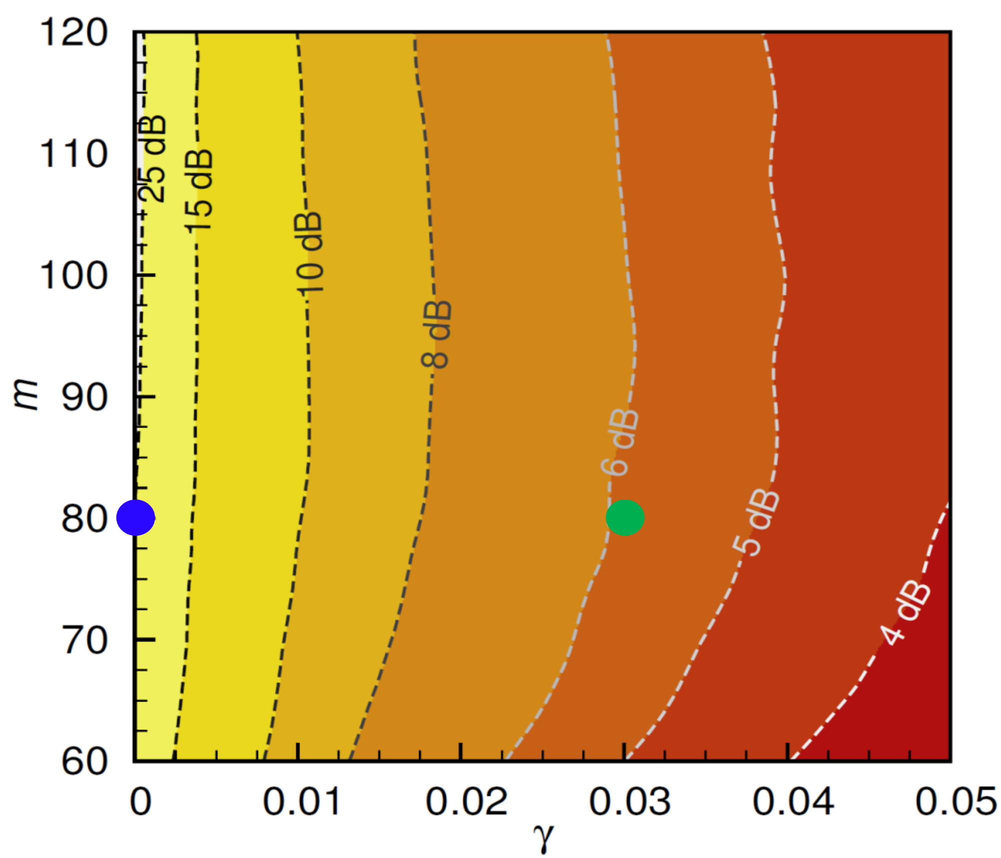
Pseudorandom Pattern of Sign Flips

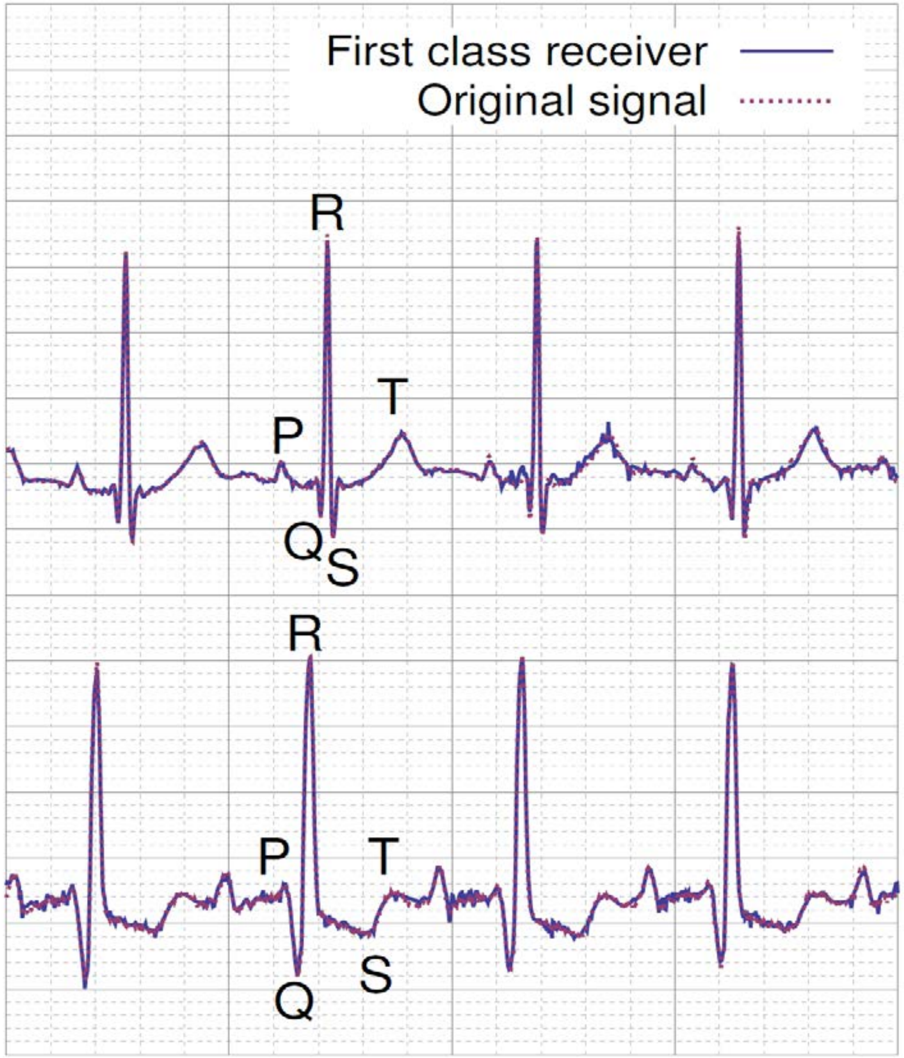


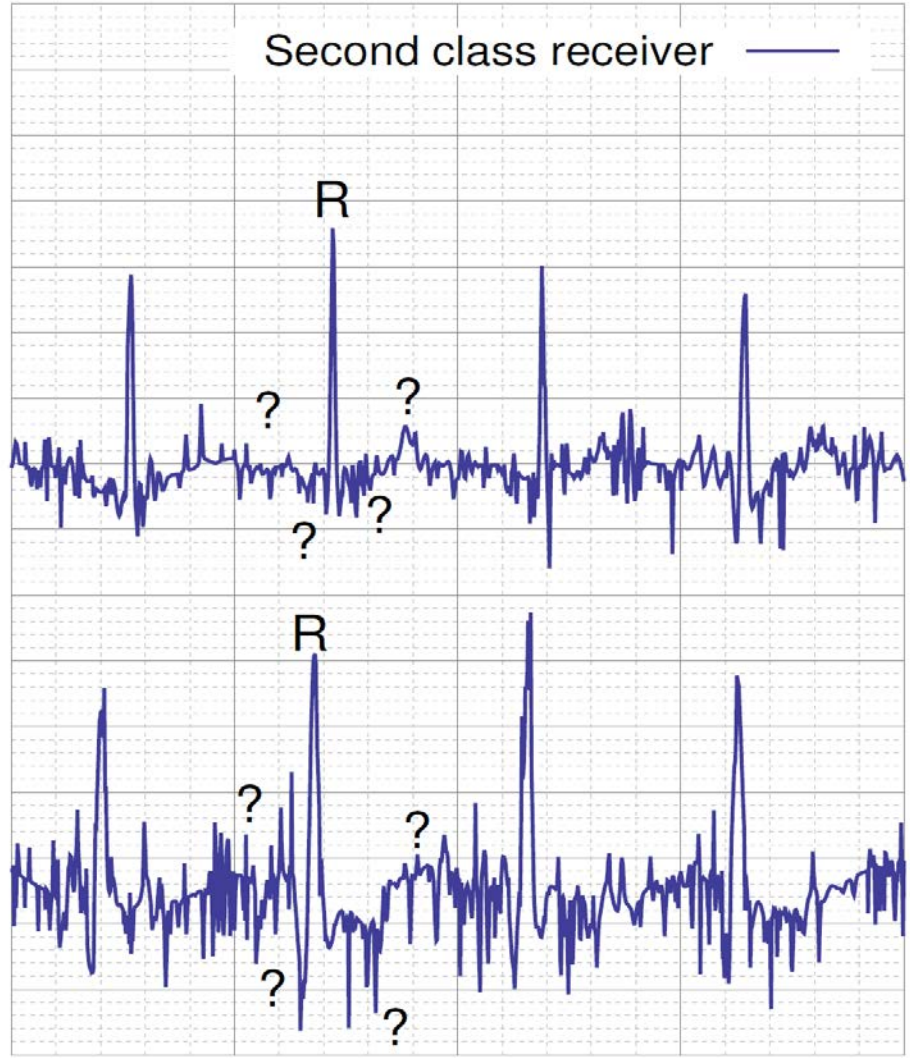
A Case Study: Secure CS of ECGs I

- Example: CS of ECG signals.
  - First-class user: medical equipe
  - Second-class user: relatives
- Let $n = 256$ Nyquist rate samples. For varying m and γ we measure the average reconstruction SNR.
  - $\Psi_{256\times1440}$ is a dictionary of Gabor time-frequency atoms.
  - $\Phi_{m\times256}$ is the devised measurement matrix generator.
  - Intrinsic SNR=30 dB
- We choose $m = 80, \gamma = 0.03$ such that $\frac{m}{n} \approx 3$ and $\mathrm{ARSNR}^{(1)} \approx 25$ dB, $\mathrm{ARSNR}^{(2)} \approx 6$ dB.

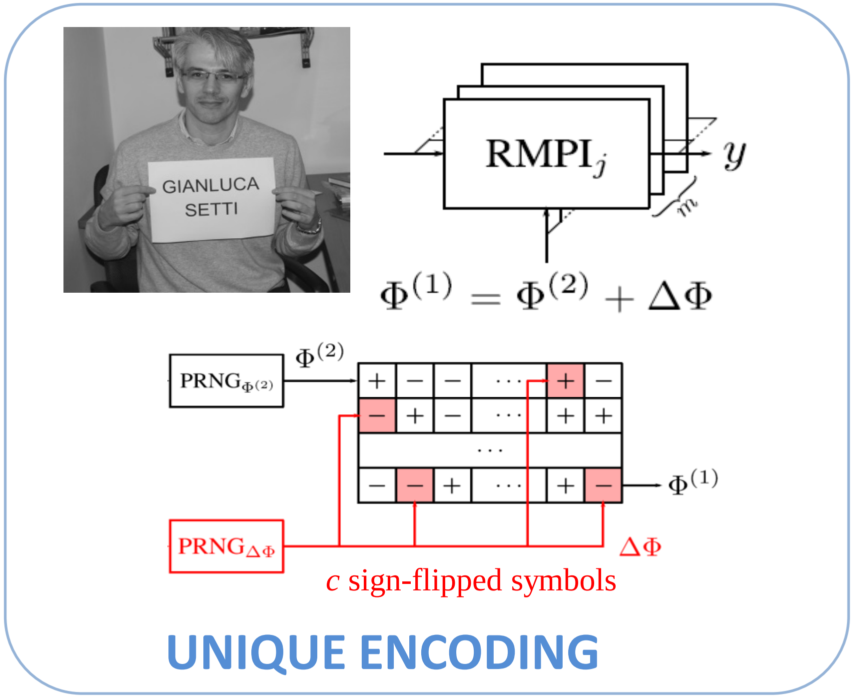
$$\mathrm{ARSNR_{dB}} = \mathbf{E}\left[20\log_{10}\left(\frac{\|\alpha\|_2}{\|\alpha-\hat{\alpha}\|_2}\right)\right]$$

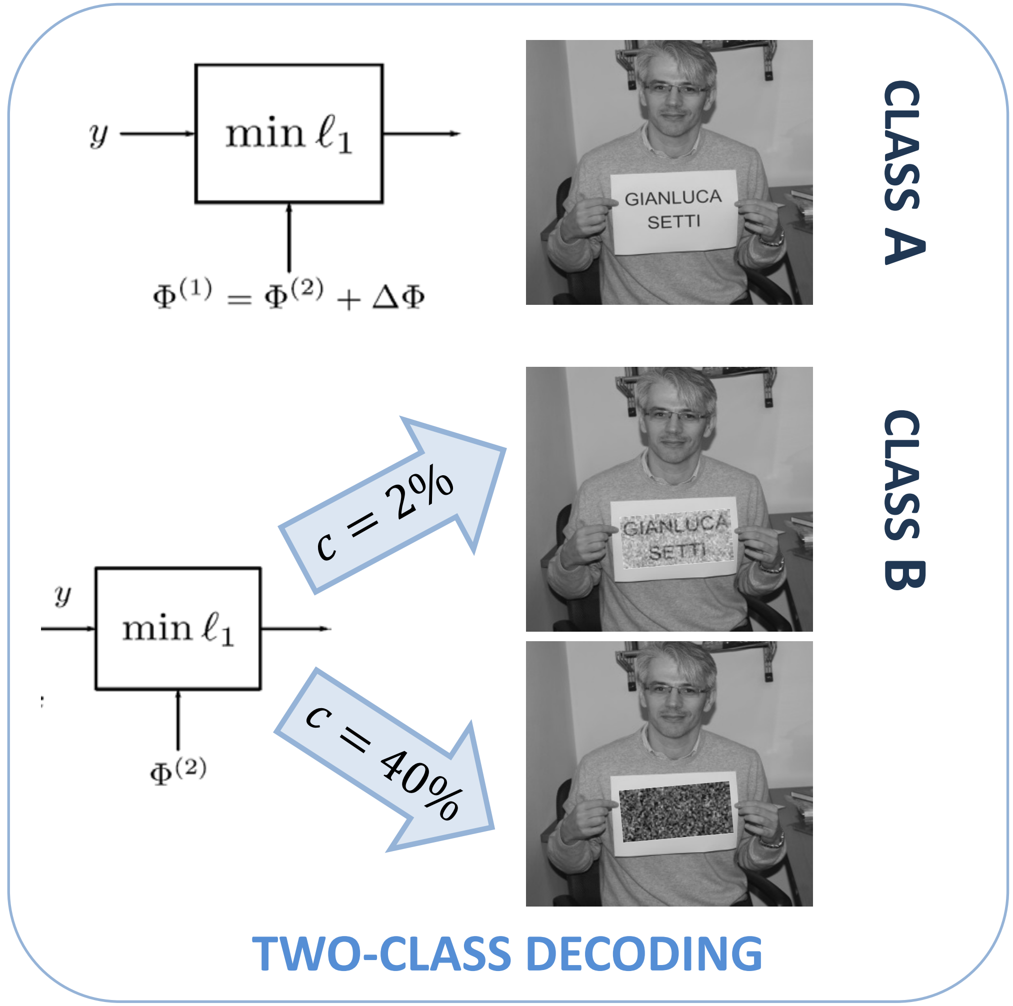



First class receiver
Original signal
R
P
T
Q
S
R
P
T
Q
S


Second class receiver
R
?
?
?
?
R
?
?
?
?

- Two **user classes** with different information access levels





- The full image frame is block-processed; each block is encoded and transmitted.
- When a block contains high-level information, a class B user is only able to reconstruct the transmitted block with low SNR.
- $c \geq 40\%$ ensures **total corruption** of the information contained in the image.

- It can be shown that the CS measurements are such that

$$y_j \xrightarrow[n\to\infty]{d} \mathcal{N}(0, W)$$

where W is the input signal power.

- All W-power signals have *the same* measurement PDF $f_y(y) \sim N(0,1)$ so for large n the system is asymptotically secret in the Shannon sense (except for $W = 0$).

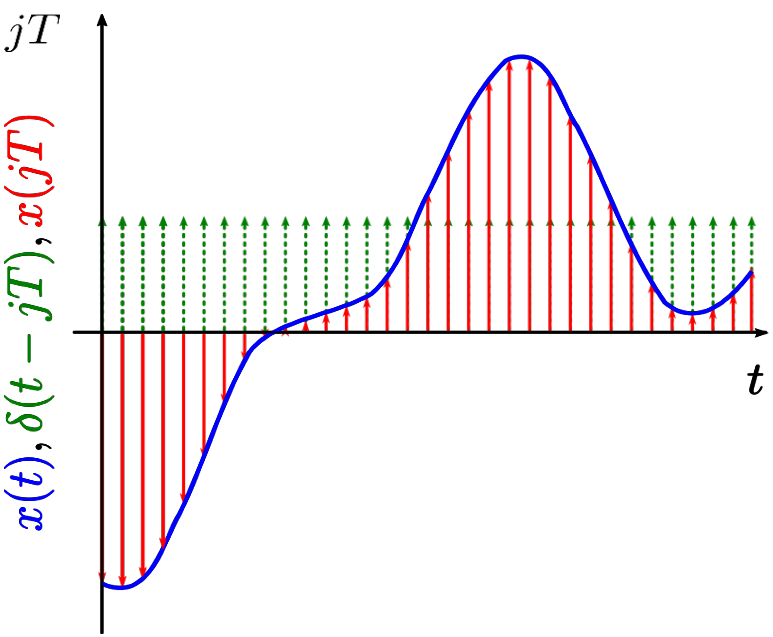
$$f_y(y) \equiv f_{y|x}(y|x)$$

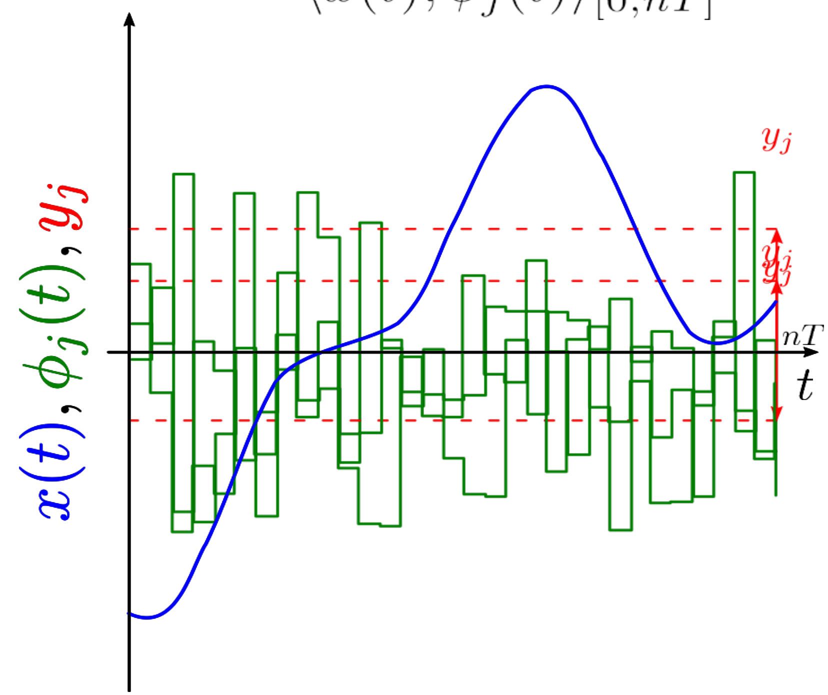
- A malicious eavesdropper cannot infer any information from the statistical analysis of y, which is *practically identical to Gaussian noise for* $n \to \infty$.

- How do we translate these algorithms into analog circuitry for signal acquisition?
- In general, sampling *is* the projection of $x(t)$ over a set of waveforms:

$$x(jT) = \int_0^{nT} x(t)\delta(t - jT)dt = \langle x(t), \delta(t - jT)\rangle_{[0,nT]}$$

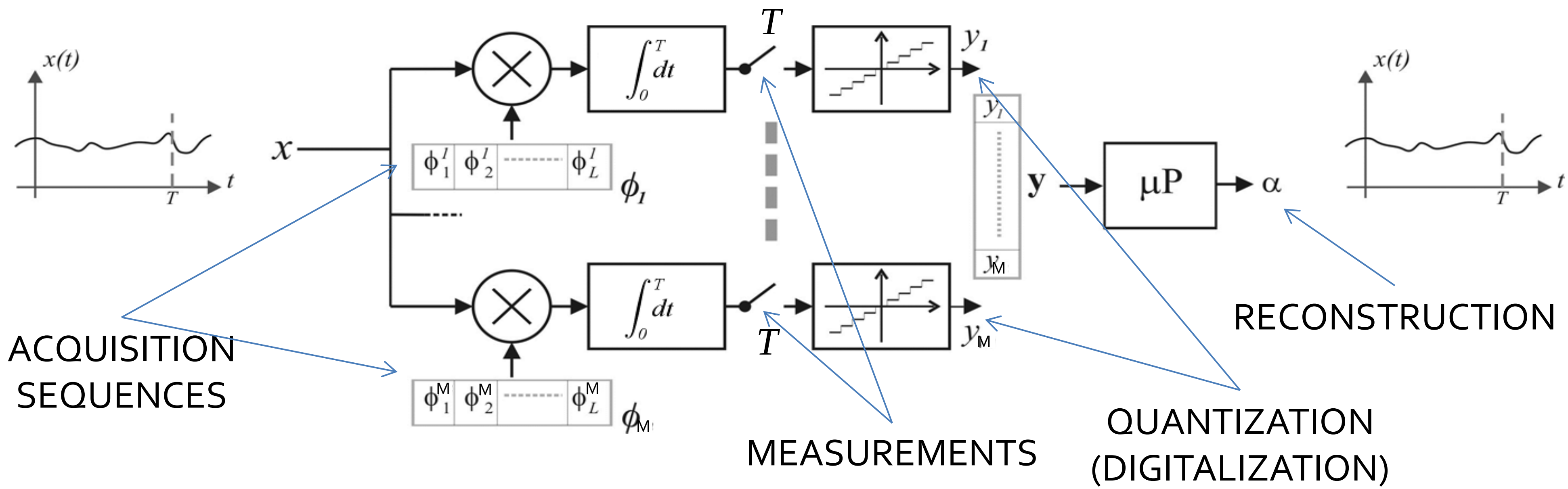
$$y_j = \int_0^{nT} x(t)\phi_j(t)dt = \langle x(t), \phi_j(t)\rangle_{[0,nT]}$$





Main architecture: Random Modulation Pre Integration

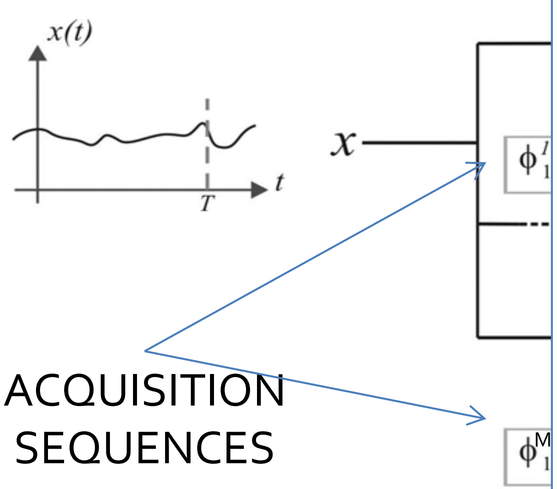
- RMPI



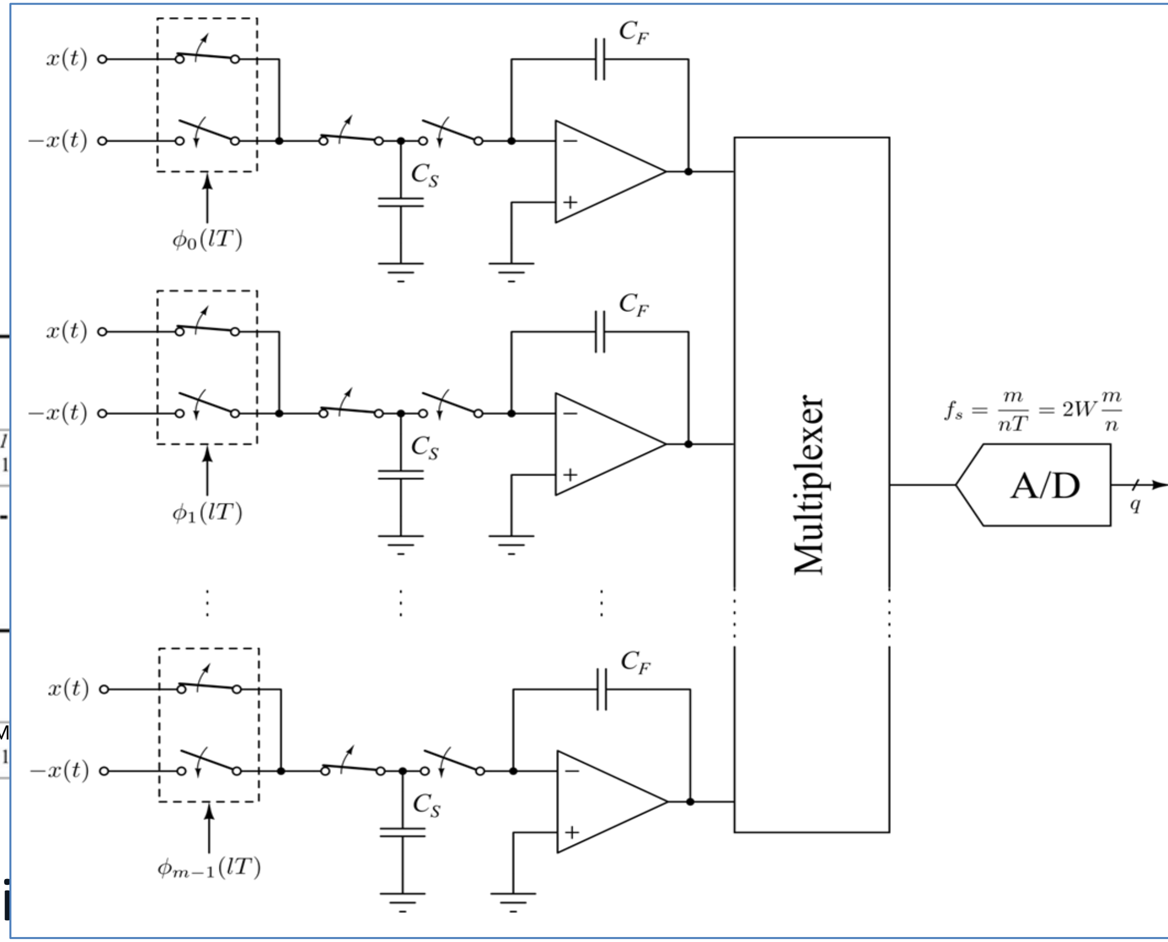
Acquisition sequences: either Gaussian random sequences or binary antipodal random sequences

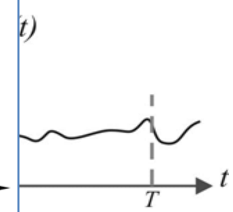
Main architecture: Random Modulation Pre Integration

- RMPI



ACQUISITION SEQUENCES





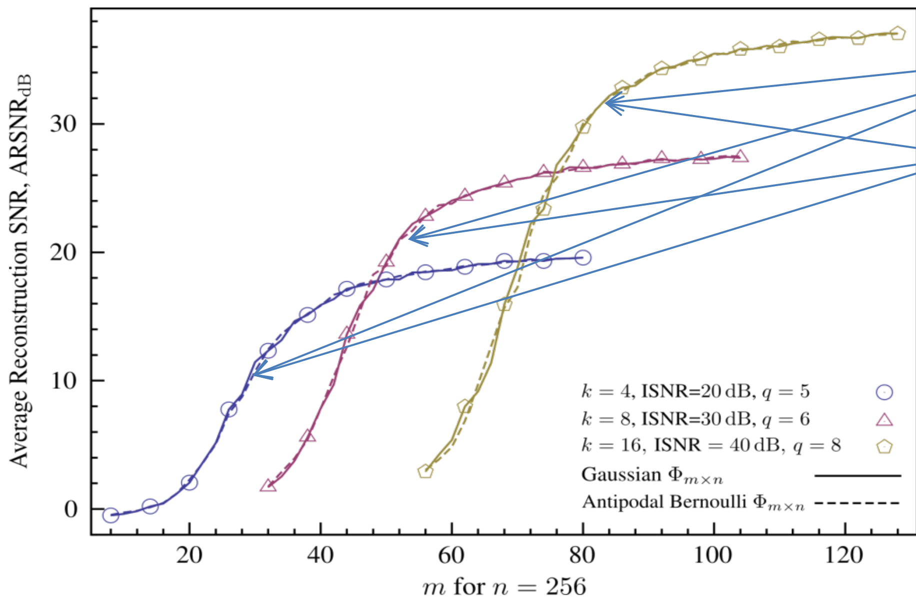
RUCTION

Acquisiti

sequences or binary antipodal random sequences

RMPI: Gaussian vs Binary sequences

RMPI performance comparison



Gaussian acquisition sequences *(solid lines)*

Binary antipodal acquisition sequences *(dashed lines)*

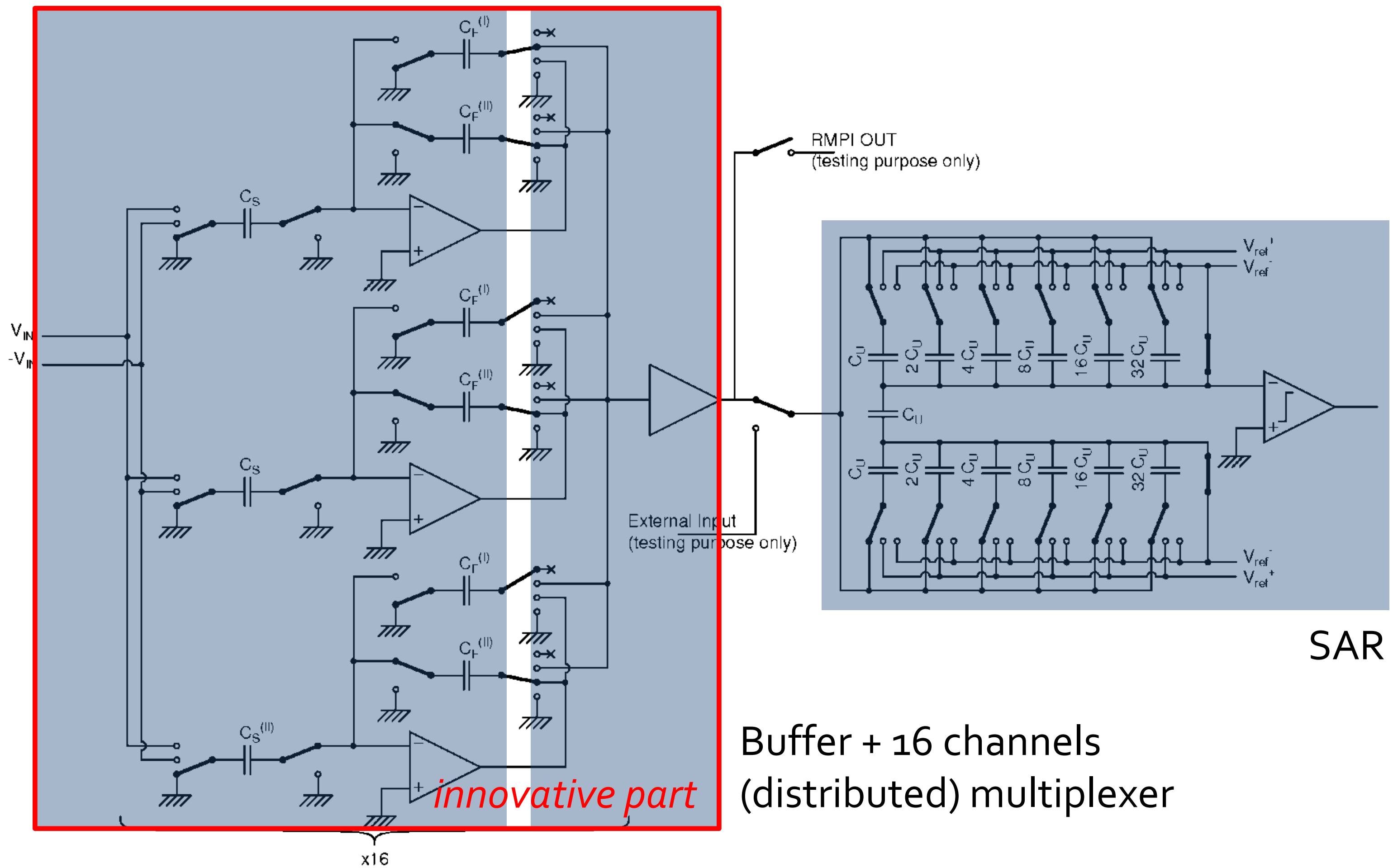
From J. Haboba et al, *"A pragmatic look at some compressive sensing architectures with saturation and quantization," IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 2, no. 3, Sep. 2012.*

SSIGPRO

Implementation overview

- Switched capacitor based architecture – *low voltage, fully differential*
- Additional I/O pins for testing purpose
- Detailed schematic (single-ended version):

Binary antipodal RMPI, 16 channels
(external ϕ for maximum flexibility)

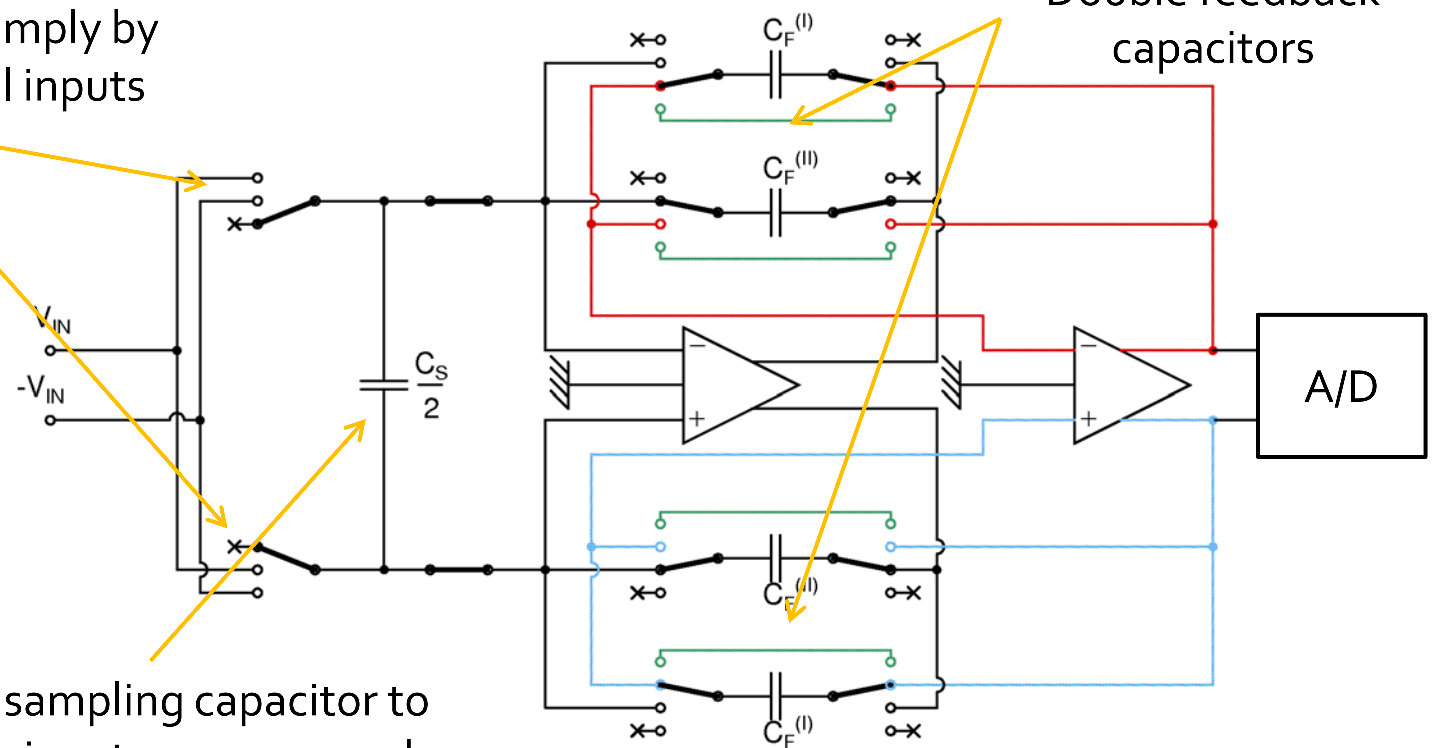


Switched capacitor fully differential integrator with embedded +1/-1 multiplication and buffer/multiplexer: main challenges in the design

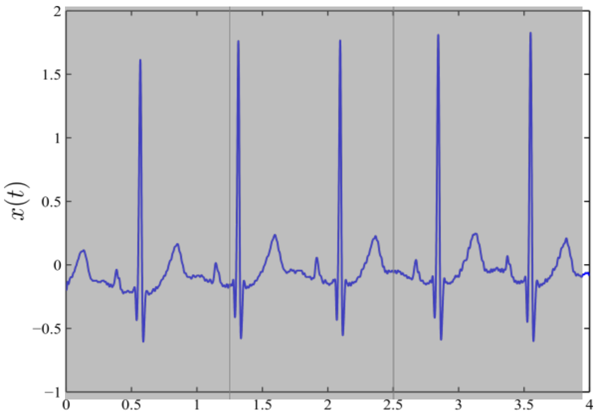
- Minimization of **leakage currents** (long integration time)
- Ensuring that **no pieces of signal are lost** between two time frames (two feedback capacitors)
- Coping with **saturation of OP-AMPS**

+1/-1 multiplication simply by swapping differential inputs

Double feedback capacitors



Single sampling capacitor to remove input common mode



Frame #1 Frame #2 Frame #3

time

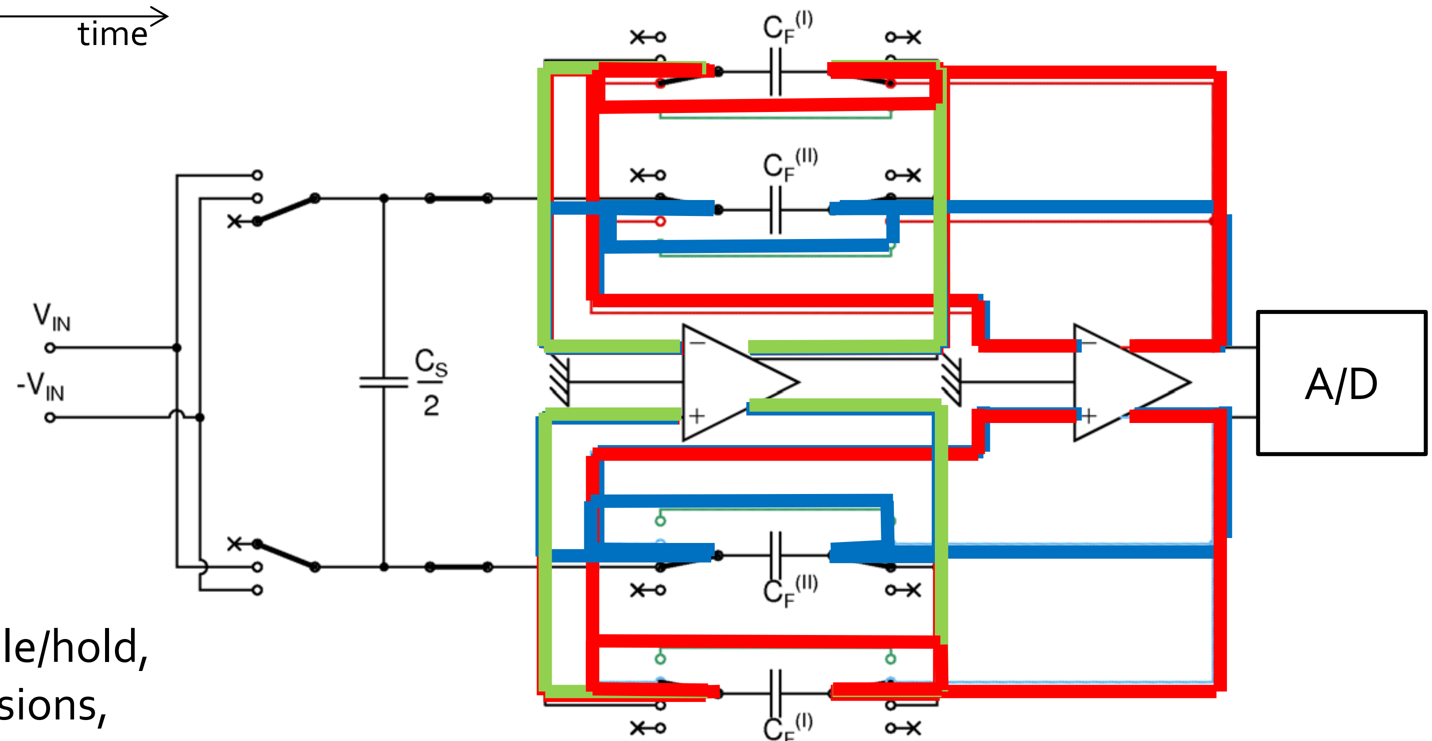
$$y_j^{(\#1)} = \int_0^{nT} \phi_j(t)x(t)dt$$

$$y_j^{(\#2)} = \int_{nT}^{(n+1)T} \phi_j(t)x(t)dt$$

$$y_j^{(\#3)} = \int_{(n+1)T}^{(n+2)T} \phi_j(t)x(t)dt$$

Event sequence:

- Time frame #1: $C_F{}^{(I)}$ is used to accumulate $y^{(\#1)}$
- Time frame #2: $C_F{}^{(I)}$ is used by the buffer, then resetted;
  $C_F{}^{(II)}$ is used to accumulate $y^{(\#2)}$
- Time frame #3: $C_F{}^{(II)}$ is used by the buffer, then resetted;
  $C_F{}^{(I)}$ is used again to accumulate $y^{(\#3)}$



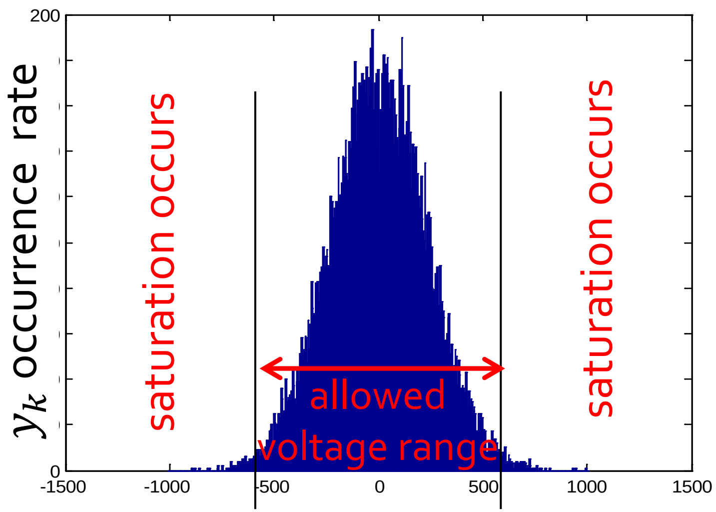
Results:
No need for external sample/hold,
No lag between CS conversions,
Time for A/D conversion and reset capacitor charge

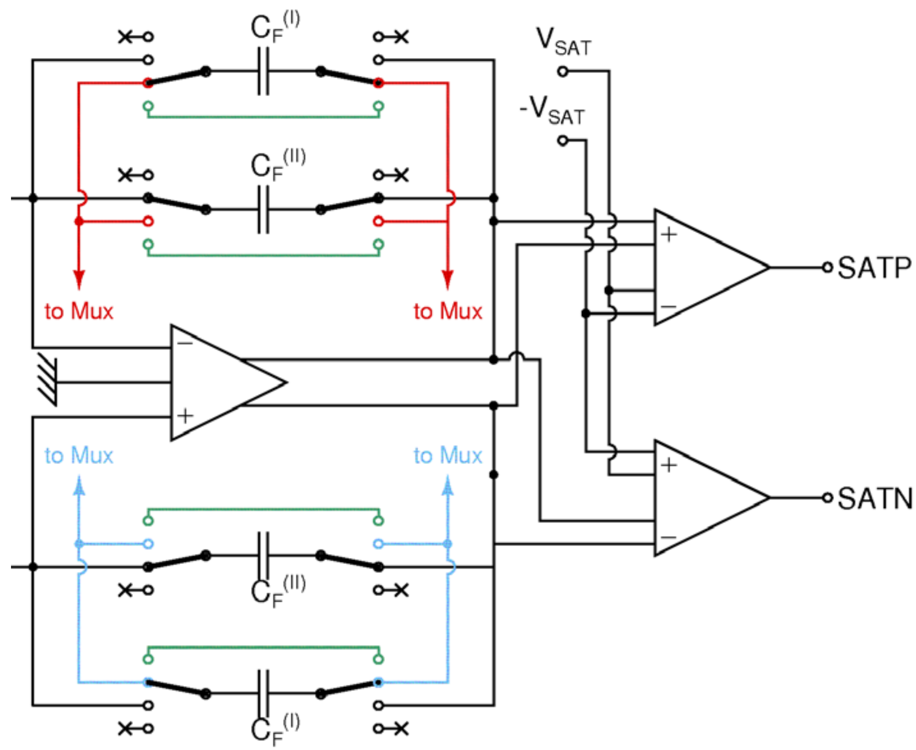
Additional feature: saturation checking

- each integrator has two comparators for continuous checking for both positive and negative saturation

$$y_k \cong \frac{C_S}{C_F} \sum_{j=0}^{n-1} \phi_j^k \, x(jT)$$

Central Limit Theorem

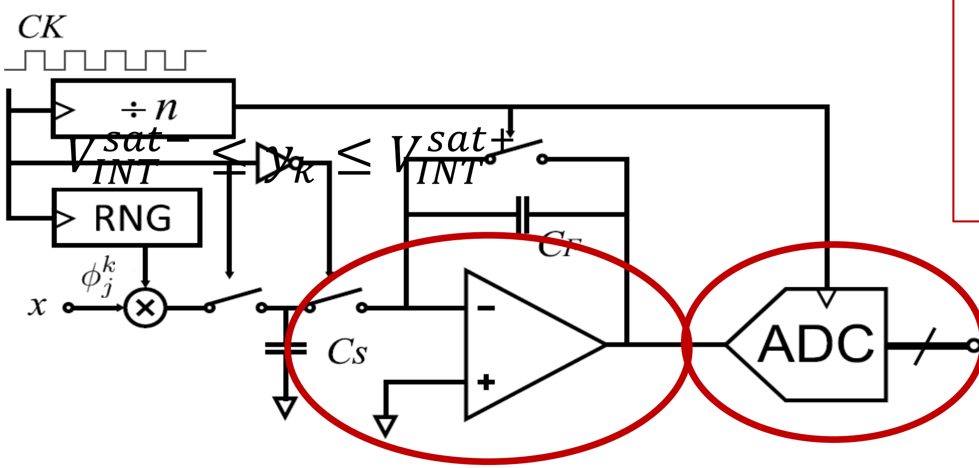




theoretically: measurements are Gaussian and not bounded;

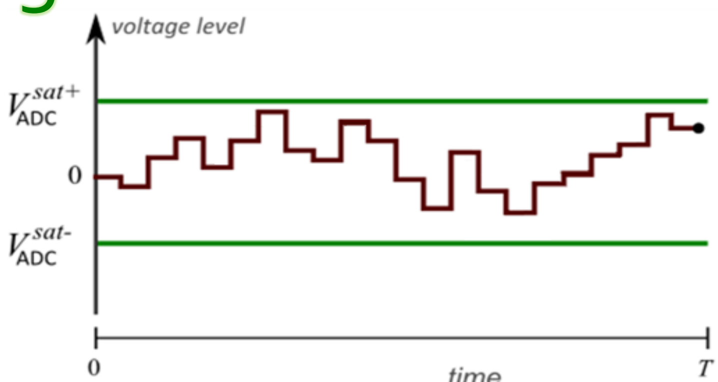
practically the peak value is <u>very large</u> compared with common values

Basic integrator schematic:



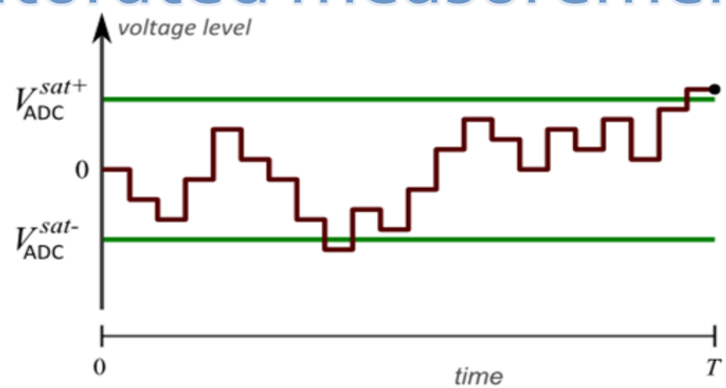
The ADC suffers from
SATURATION!
$V_{ADC}^{sat-} \leq y_k \leq V_{ADC}^{sat+}$

good measurement



$V_{ADC}^{sat-} \leq y_k \leq V_{ADC}^{sat+}$

saturated measurement



$y_k > V_{ADC}^{sat+}$

*Saturated measurements can be **detected** and **removed** by the reconstruction algorithm…*

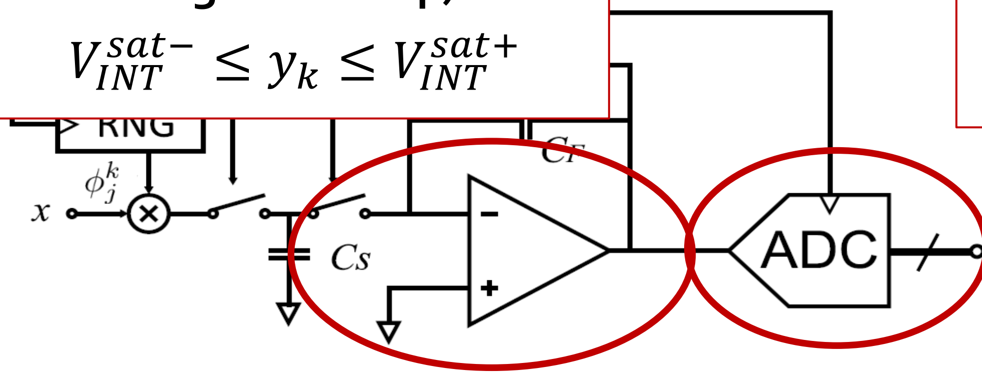
Basi

Also the ***integrator*** **(OP-AMP)** may suffer from saturation! **(saturation may occur at ANY integration step)**

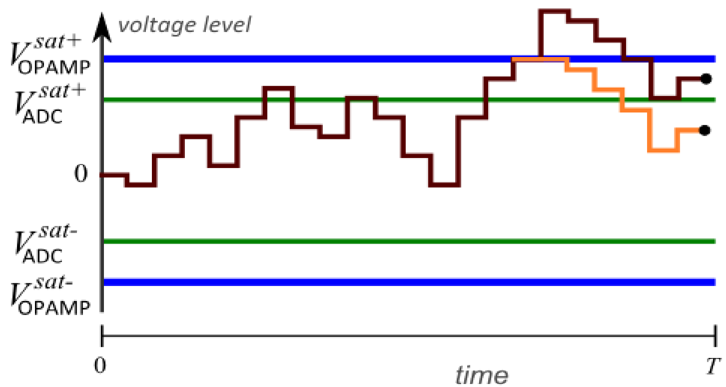
$$V_{INT}^{sat-} \leq y_k \leq V_{INT}^{sat+}$$

The ADC suffers from **SATURATION!**

$$V_{ADC}^{sat-} \leq y_k \leq V_{ADC}^{sat+}$$



corrupted measurement



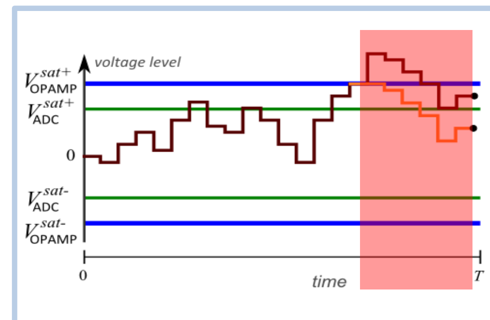
*The **OPAMP reaches saturation.** The measurement is invalid, and the final value **<u>could be in the ADC range</u>**!*

It is IMPOSSIBLE to detect that a measurement is corrupted only by observing the converted value

our proposal: RMPI Smart Saturation Checking, RMPI-SSC

If an OP-AMP saturation is detected, integration is stopped, and the time in which saturation has occurred is used into the reconstruction algorithm

+1	-1	+1	+1	-1	-1	-1	+1	-1	-1	-1	+1	-1	+1	+1	-1	
-1	-1	+1	-1	+1	+1	+1	-1	+1	-1	+1	-1	+1	0	0	0	***OPAMP SAT!!***
+1	-1	-1	+1	-1	-1	+1	+1	-1	+1	-1	+1	+1	-1	-1	+1	
+1	-1	+1	+1	-1	+1	+1	+1	-1	0	0	0	0	0	0	0	***OPAMP SAT!!***
-1	+1	-1	-1	+1	+1	-1	+1	0	0	0	0	0	0	0	0	***OPAMP SAT!!***
-1	-1	+1	+1	-1	-1	+1	+1	+1	-1	-1	+1	+1	-1	+1	+1	
+1	+1	-1	+1	+1	-1	+1	+1	+1	-1	+1	-1	-1	-1	+1	+1	
+1	-1	+1	+1	-1	-1	-1	+1	-1	-1	-1	+1	-1	+1	+1	-1	

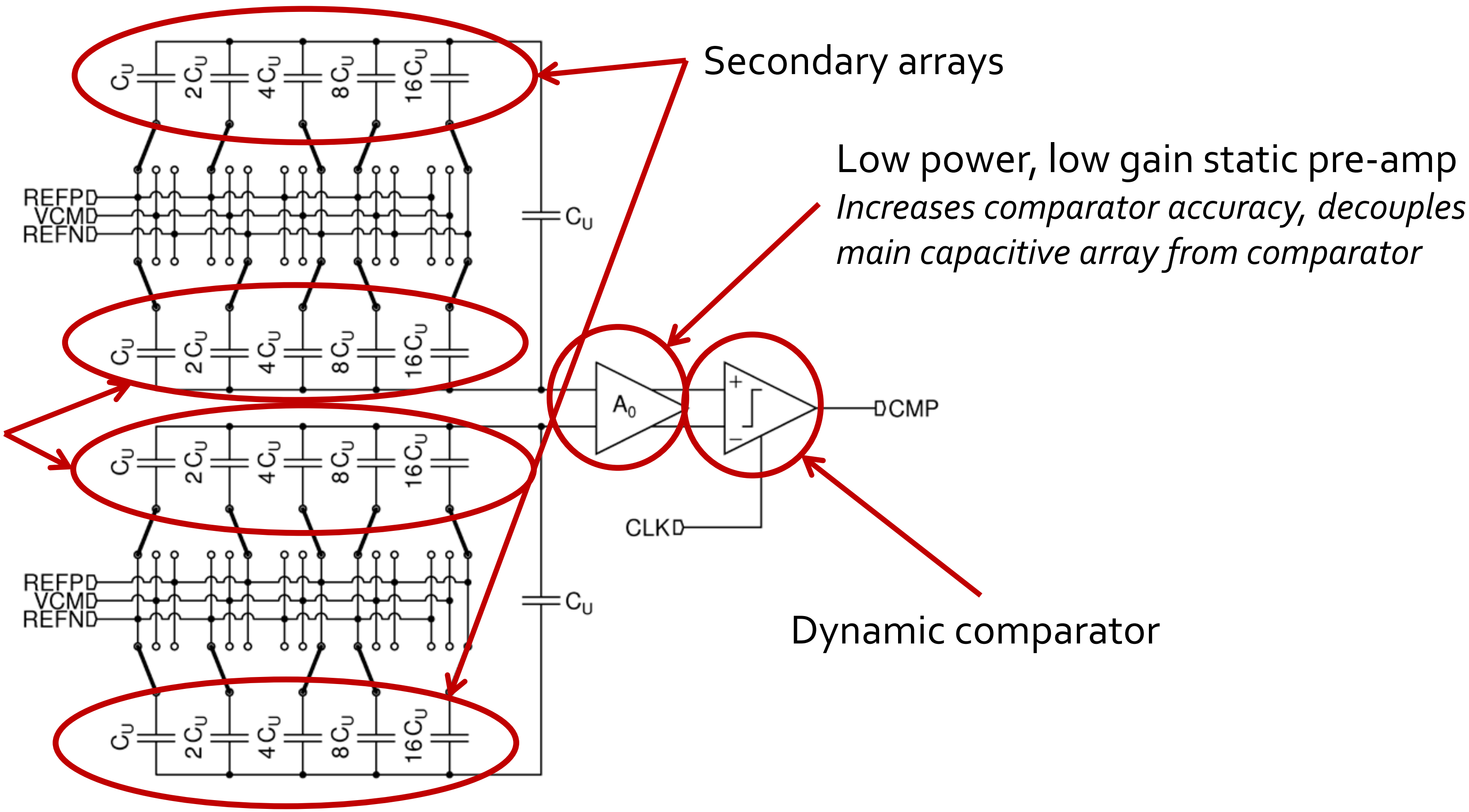


Ternary (0,+1,-1) Random Modulation pre-Integration

This approach is perfectly compatible

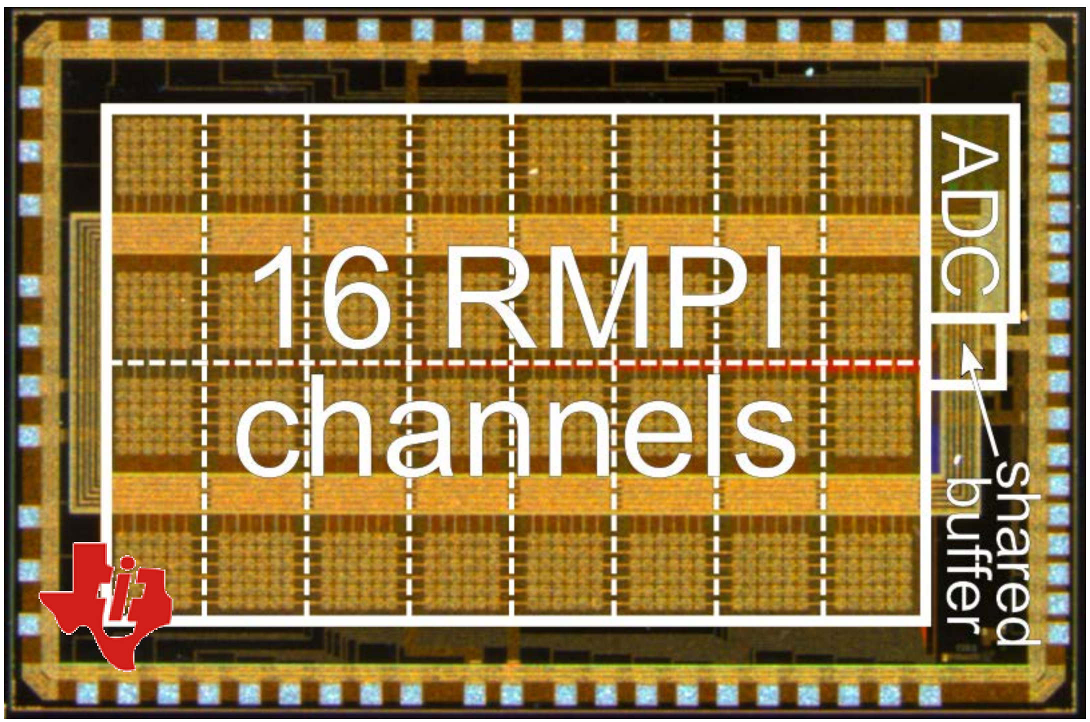
with the CS guidelines on Φ (Incoherence and RI)

A/D: standard split capacitor SAR (main caps array + secondary caps array), 11 bit



The BioCOMP prototype embeds all *analog circuits* for **Compressed Sensing** based acquisition of Biomedical Signals:

- 16 RMPI channels, switched-capacitors binary antipodal modulation, saturation checking capabilities
- Single, shared buffer
- Successive Approximation Register (SAR) based ADC



Technology:

TI CMOS 180 nm 9T5V 1.8V

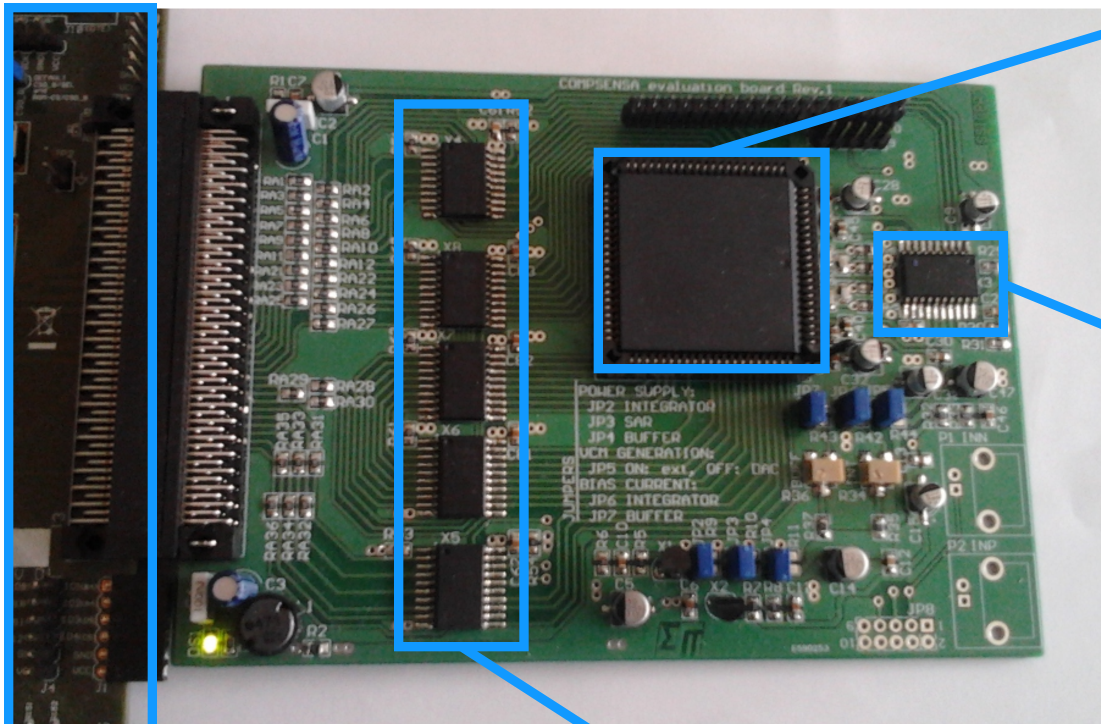
Total size (incl. pads)

2.6 x 4.0 mm^2

Circuit size

2.3 x 3.7 mm^2

The testing board:



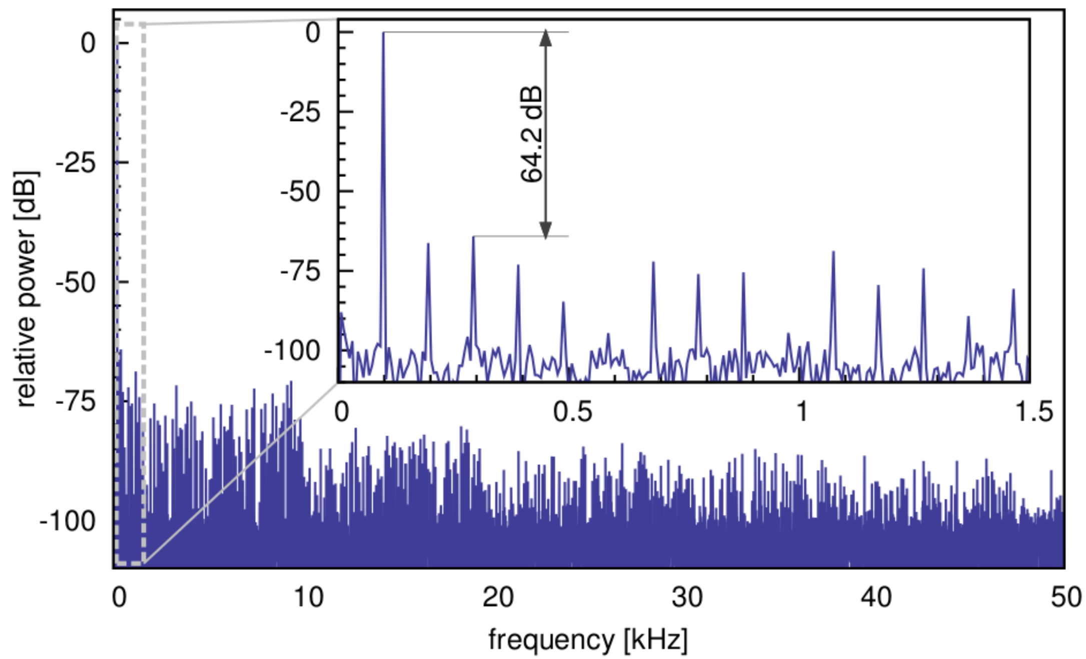
BioCOMP chip
(PLCC84)

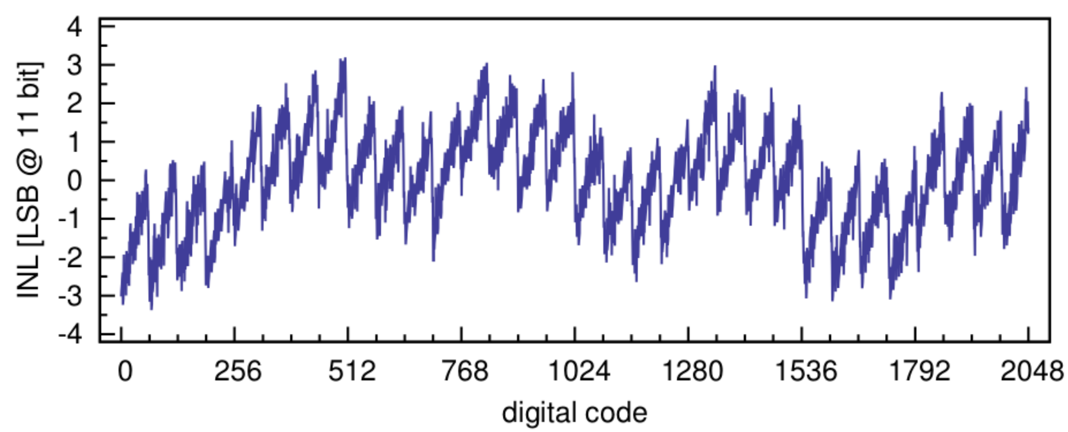
External DAC
(TLV5630)
Test signal and references voltage generation

Spartan 3E FPGA,
embedding necessary digital logic

Some digital control logic…

Parameter	Value
Area	$320 \times 770\ \mu m^2$
Power consumption $(100kS/s)$	$10.08\ \mu W$
SFDR $(100kS/s,\ f_{in} = 100\ Hz)$	$64.2\ dB$
Effective number of bits	$8.99\ bit$
INL *(best fit)*	$< 3.4\ LSB$
Max. conversion speed	$200\ kS/s$
Figure of merit $(100kS/s)$	$198\ fJ/c.-l.$

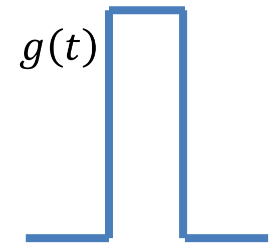




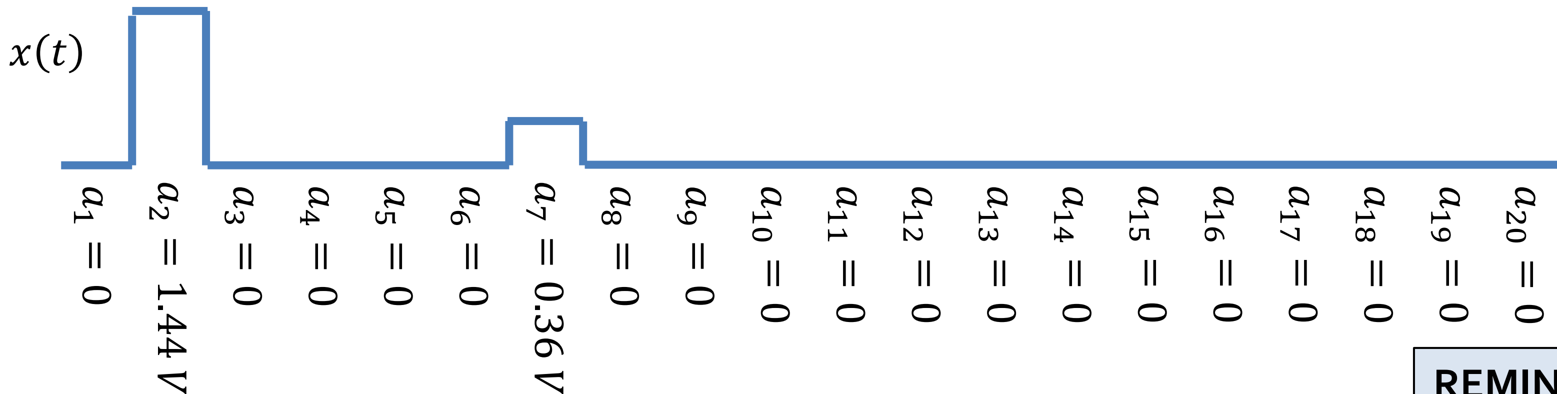
Artificial sparse signal #1: signal sparse on the canonical basis (*unit pulse step)*

- 2-sparse signal (k=2) (sum of 2 pulses at random position)
- Time-window length: 20 steps (n=20)
- Limited number of used channels (m=8)

$$x(t) = \sum_{i=1}^{n} a_i \, g\left(\frac{t-i}{T}\right)$$



Example:



Measurements:

$$y_j = \sum_{i=1}^{n} a_i \phi_{j,i} = \begin{cases} 1.8\,V & \phi_{j,2} = 1, \phi_{j,7} = 1 \\ 1.08\,V & \phi_{j,2} = 1, \phi_{j,7} = -1 \\ -1.08\,V & \phi_{j,2} = -1, \phi_{j,7} = 1 \\ -1.8\,V & \phi_{j,2} = -1, \phi_{j,7} = -1 \end{cases}$$

Only 4 possible values
since K=2

REMINDER:
the RMPI computes

$$y_j = \int_0^{nT} x(t)\phi_j(t)dt$$

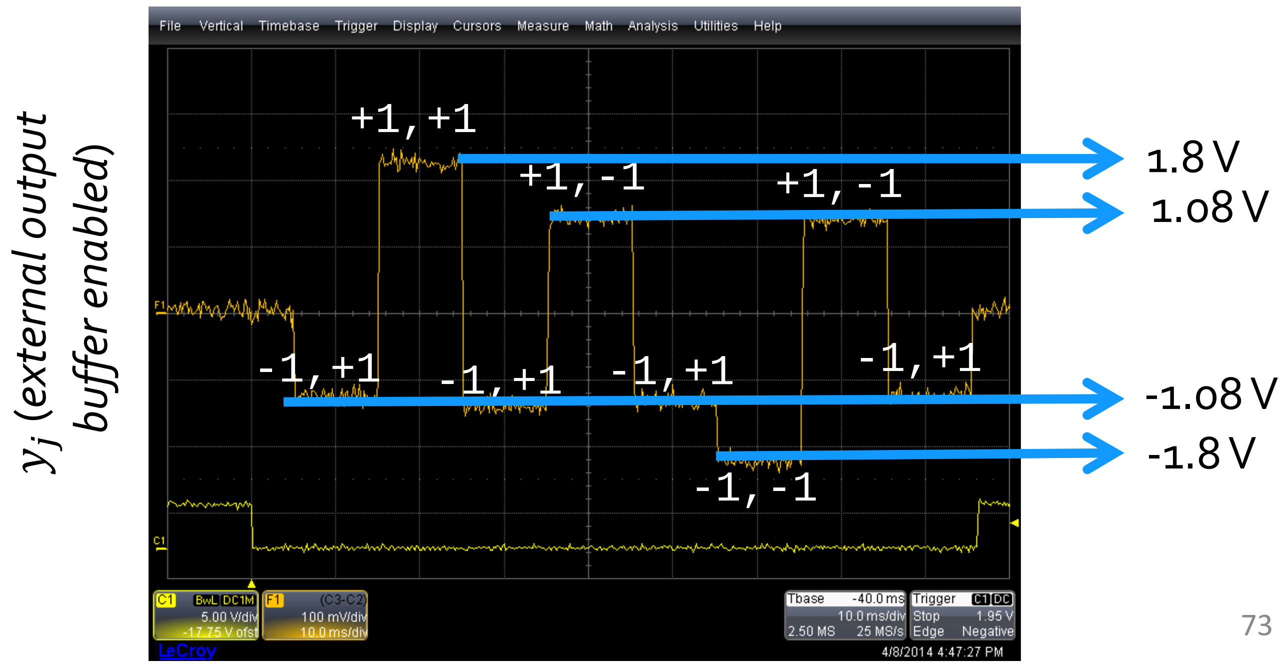
$$\approx \sum_{i=1}^{n} x(iT)\,\phi_{j,i}$$

SSIGPRO

… when using the sensing vectors

ϕ_1: +1, -1, -1, -1, +1, -1, +1, -1, -1, -1, -1, +1, +1, -1, -1, -1, +1, -1, -1, +1
ϕ_2: -1, +1, +1, +1, +1, -1, +1, +1, -1, -1, -1, -1, +1, +1, -1, -1, -1, +1, +1, -1
ϕ_3: +1, -1, +1, -1, +1, +1, +1, -1, -1, +1, -1, +1, -1, -1, +1, -1, -1, -1, +1, +1
ϕ_4: +1, +1, -1, +1, +1, +1, -1, -1, -1, -1, +1, -1, +1, +1, +1, +1, -1, +1, +1, +1
ϕ_5: -1, -1, -1, -1, -1, +1, +1, +1, +1, +1, -1, -1, +1, -1, -1, -1, +1, +1, -1, -1
ϕ_6: -1, -1, +1, -1, +1, -1, -1, +1, +1, +1, -1, -1, +1, +1, -1, +1, +1, -1, -1, -1
ϕ_7: +1, +1, -1, -1, +1, +1, -1, +1, +1, -1, +1, +1, +1, +1, +1, +1, -1, -1, -1, +1
ϕ_8: +1, -1, -1, -1, +1, -1, +1, +1, +1, -1, -1, -1, +1, -1, -1, -1, +1, -1, -1, -1

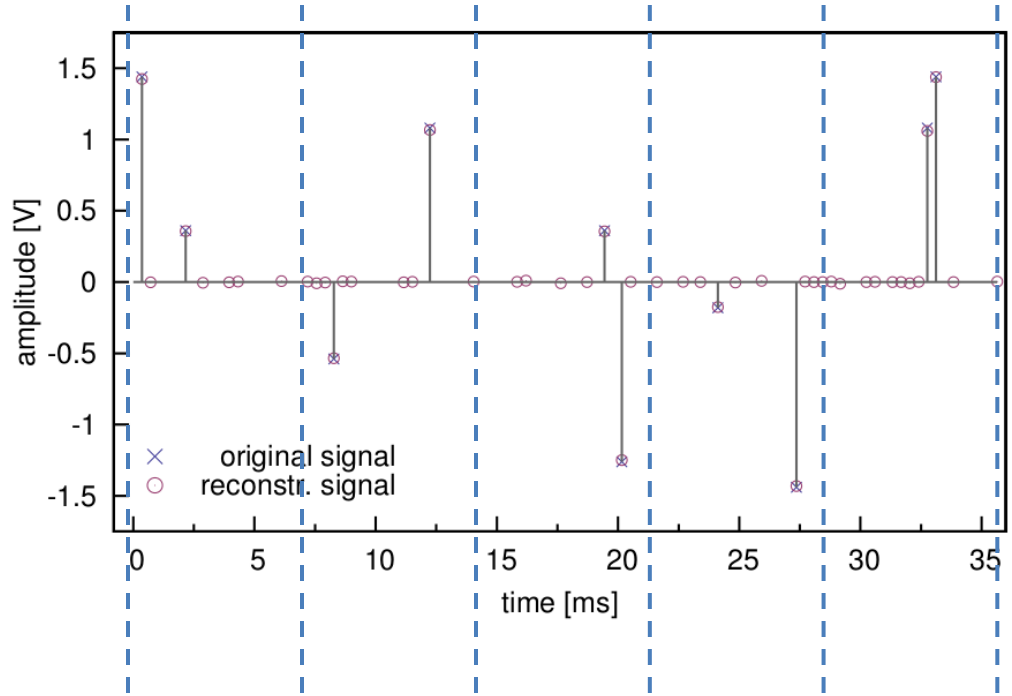
$\phi_{j,2}, \phi_{j,7}$:
-1, +1
+1, +1
-1, +1
+1, -1
-1, +1
-1, -1
+1, -1
-1, +1



Performance summary

(5 different time windows averaging)

Parameter	Value
Sparsity level	$n = 20, k = 2$
Sampling frequency	$f_s = 2.78\ kHz$
Time window length	$T_W = 7.2\ ms$
Power consumption	$251\ \mu W$
Measurement SNR	$39.6\ dB$
Reconstruction SNR	$37.7\ dB$



Each time windows has only
K=2 pulses at different position

<u>Average reconstruction SNR: 37.7 dB</u>

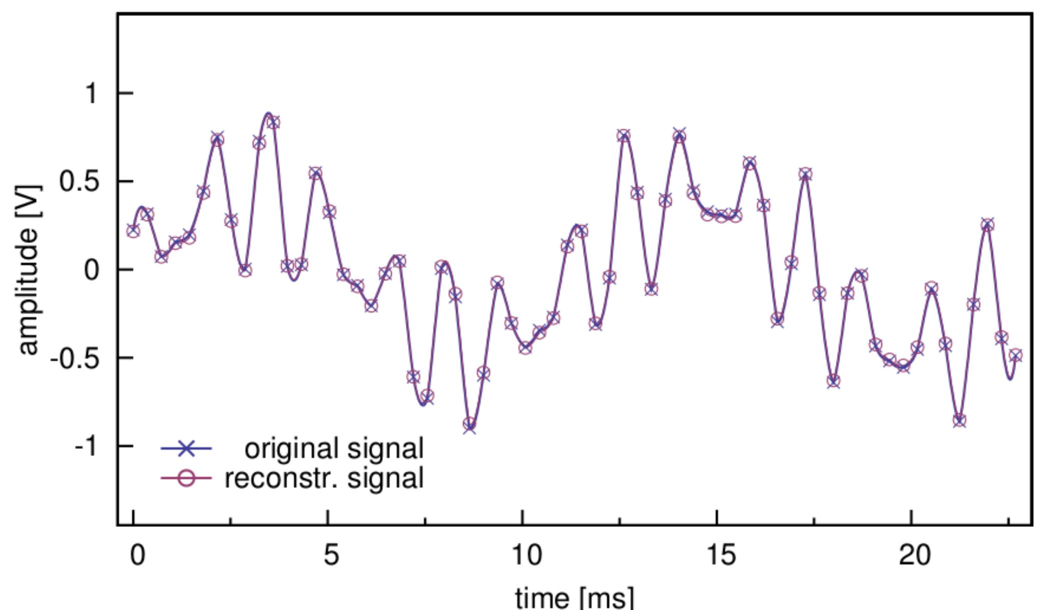
Basic system testing – IV

Artificial sparse signal #2: signal sparse on the canonical basis (*unit pulse step)*

- 3-sparse signal (k=3) (sum of 3 sinusoidal tones)
- Time-window length: 64 steps (n=64)
- All channels used (m=16)

$$x(t) = \sum_{i=1}^{n/2-1} a_i \sin(2\pi i t) + \sum_{i=n/2n} a_i \cos(2\pi(n-i)t)$$

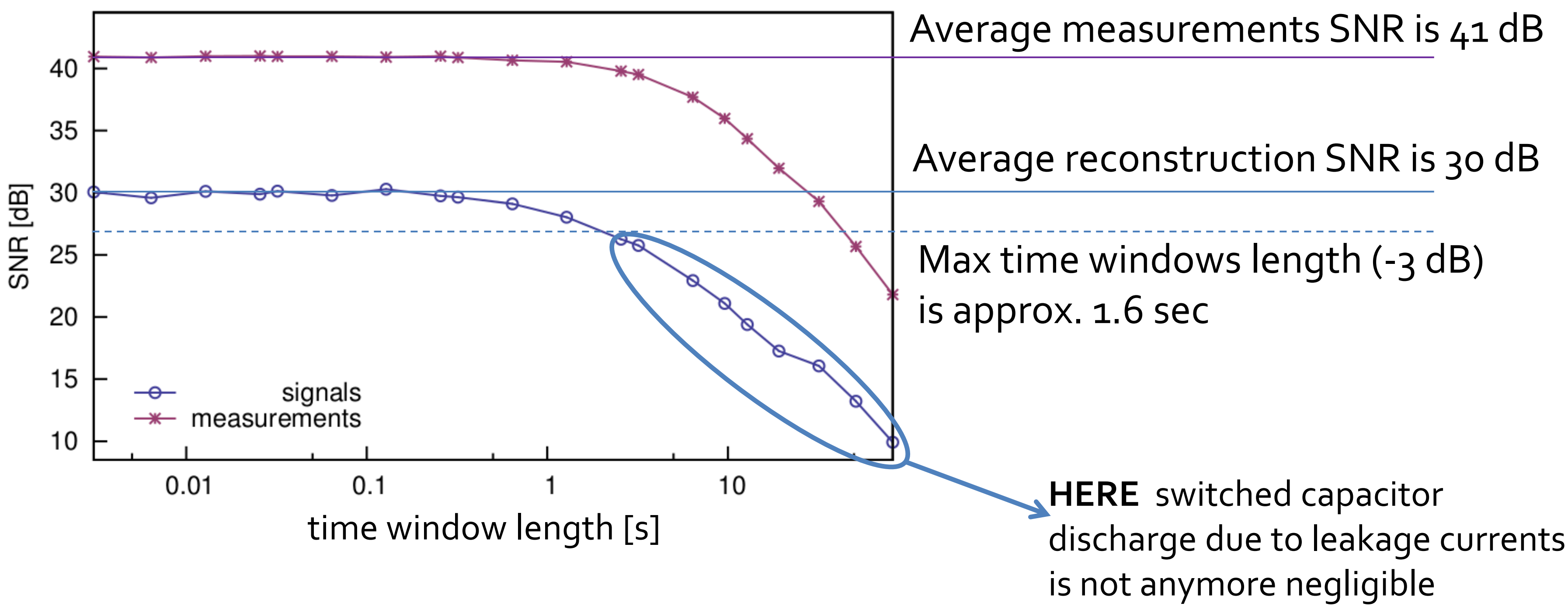
Parameter	Value
Sparsity level	$n = 64, k = 3$
Sampling frequency	$f_s = 2.78\ kHz$
Time window length	$T_W = 23\ ms$
Power consumption	$495\ \mu W$
Measurement SNR	$40.9\ dB$
Reconstruction SNR	$30.0\ dB$



Average reconstruction SNR: 30.0 dB

By using **artificial sparse signal #2** working speed range of the circuit has been investigated

- **Upper bound**: maximum switched capacitor frequency: **120 KHz (8.3usec)**
- **Lower bound**: minimum switched capacitor frequency: **0.625Hz (1.6 sec)**



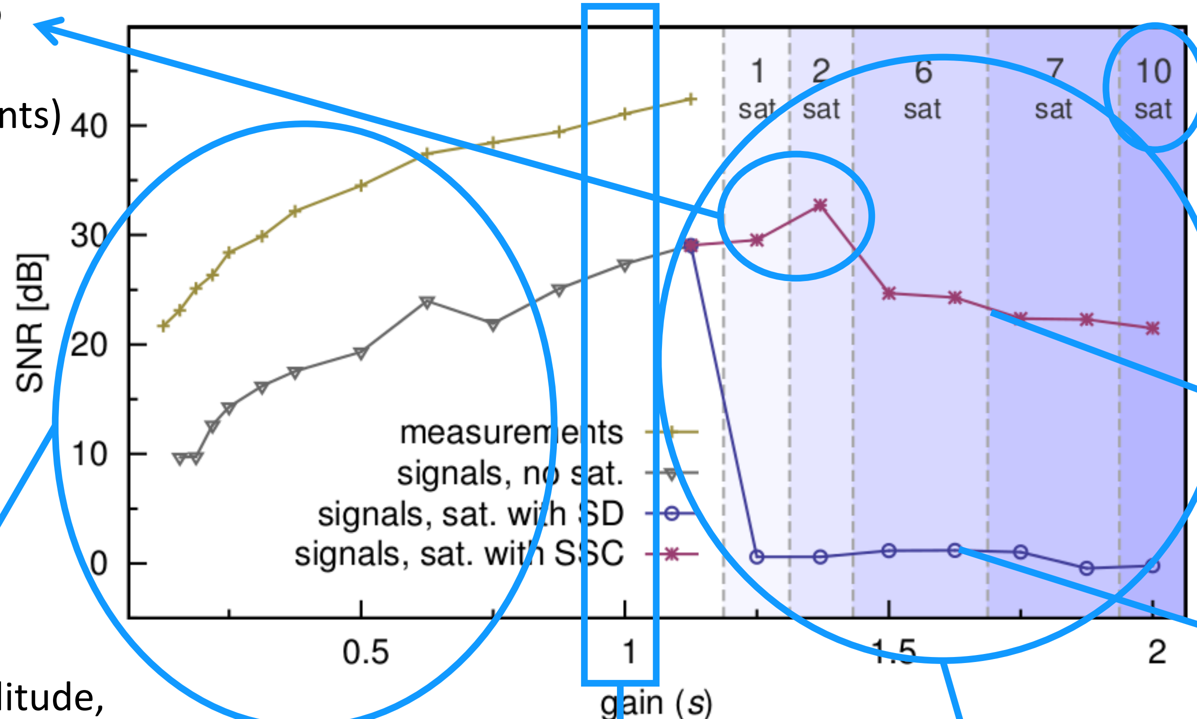
Maximum time window length: 1.6 sec

(enough for all Biomedical signals acquisition)

The sinusoidal test signal (**artificial sparse signal #2**) has been scaled by a factor *s* to test saturation checking capabilities

$$x(t) = s\left(\sum_{i=1}^{n/2-1} a_i \sin(2\pi i t) + \sum_{i=n/2}^{n} a_i \cos(2\pi(N-i)t)\right)$$

Better performance with respect to *standard* case (no saturation events)



We can reconstruct the input signal with **more than 60%** saturated measurements

By **using saturation time information**, we correctly reconstruct the signal

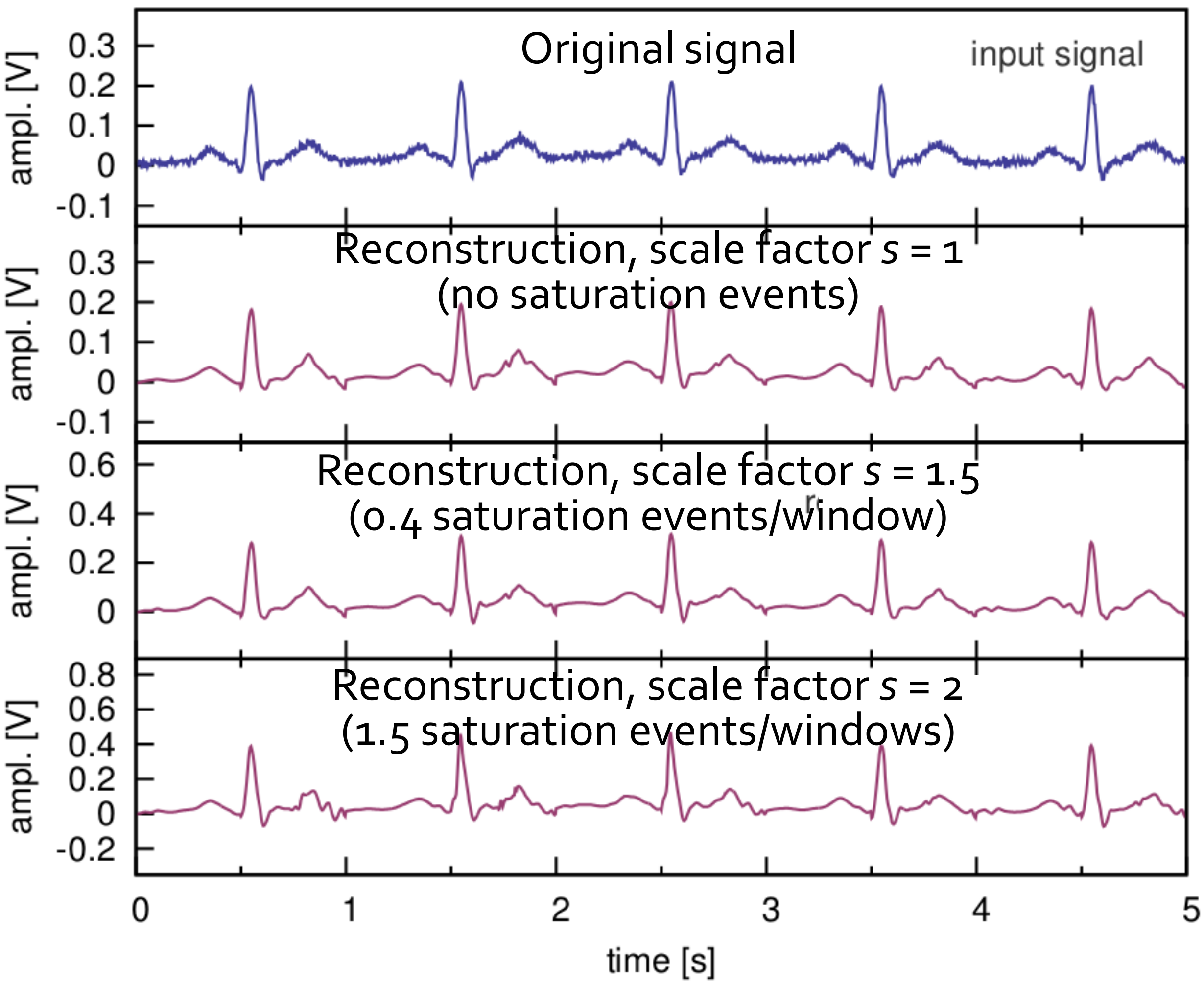
Not using saturation time information, reconstruction is not possible

low *s*: low amplitude, quantization noise is high. Low measurement SNR, low reconstruction SNR

Standard case: s=1

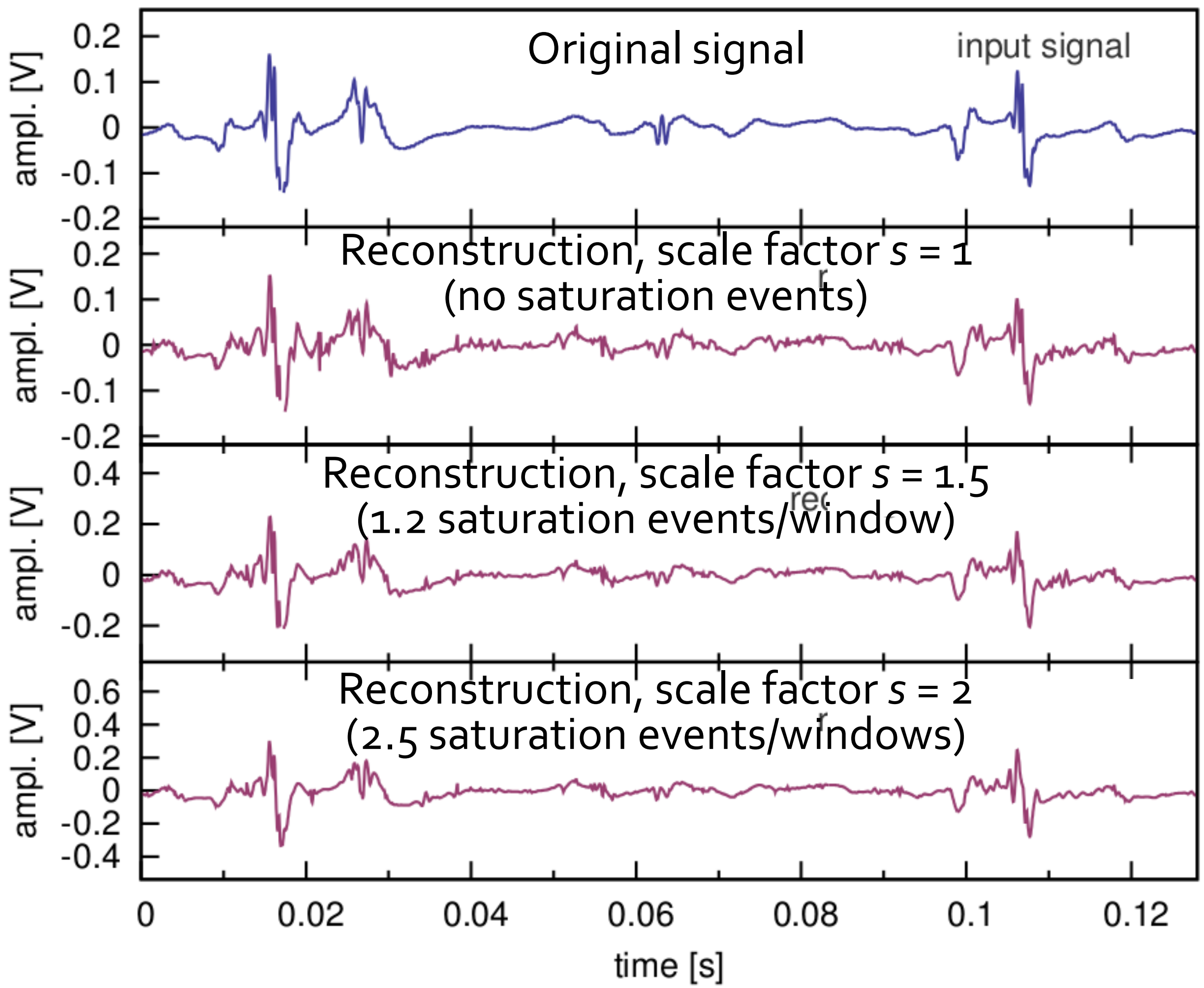
High *s*: some saturation events are detected

Test on a **real ECG signal** (available on PhysioNet database)

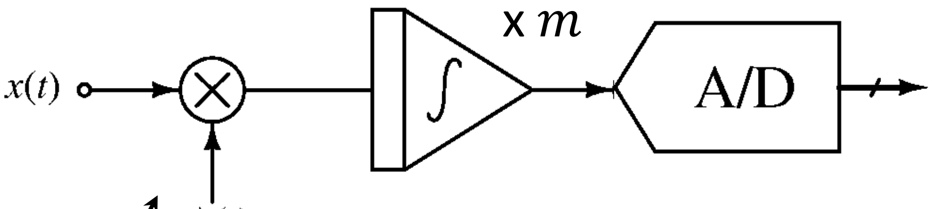


- *Signal length*: 5 s
- *Approx. heart beat*: 1 beat/s
- *Sampling frequency*: 256 Hz
- *Time windows length*: 0.5 s (n = 128)
- *Channel used*: m=16, compression ratio is 8
- ***Rakeness***
  (more than 32 measurements required for standard CS approach)
- *Sparisty basis:* Wavelet
- *Three different amplitudes*
  (for testing saturation behavior)

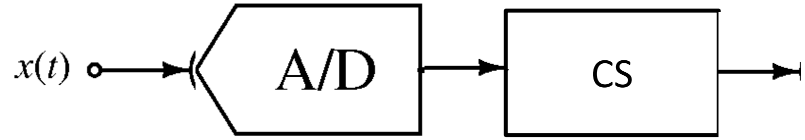
Test on a **real EMG signal** (available on PhysioNet database)



- *Signal length*: 128 ms
- *Sampling frequency*: 20 kHz
- *Time windows length*: 12.8 ms (n= 256)
- *Channel used*: m=24, compression ratio is ≈ 10 (two circuits are used)
- ***Rakeness***
  (more than 64 measurements required for standard CS approach)
- *Three different amplitudes for testing saturation behavior*



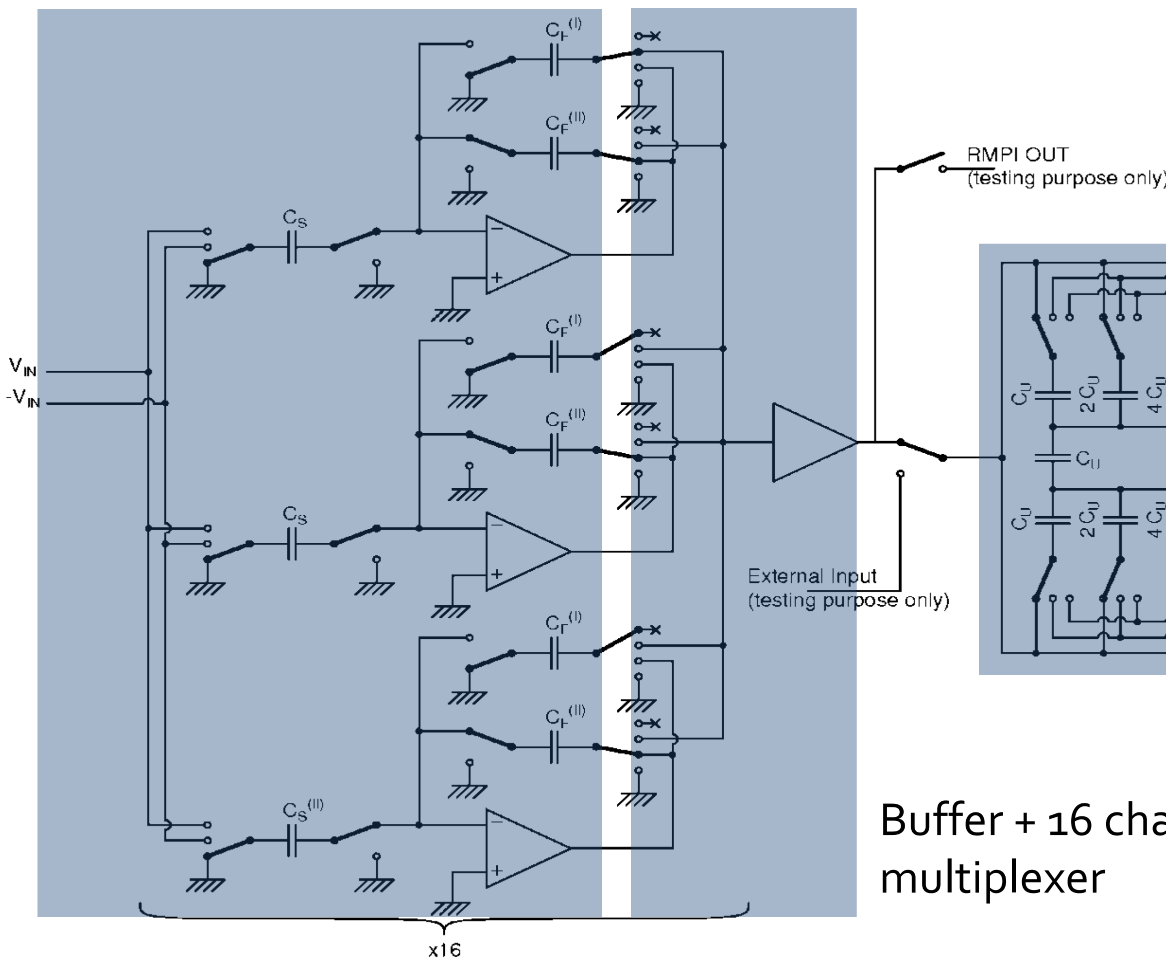
$j = 0,1, \ldots m-1$

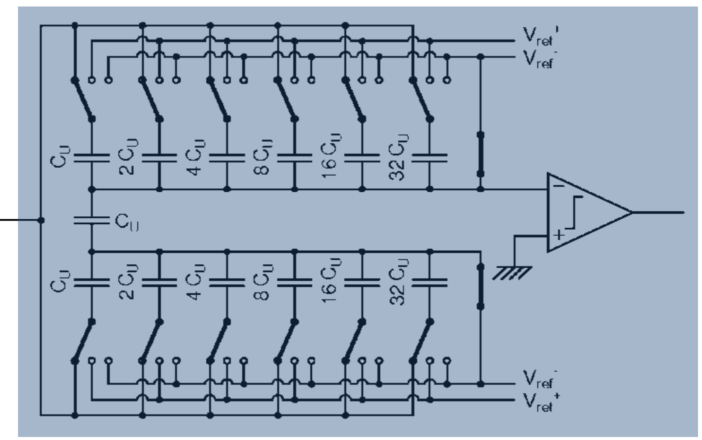


Analo

e:
on

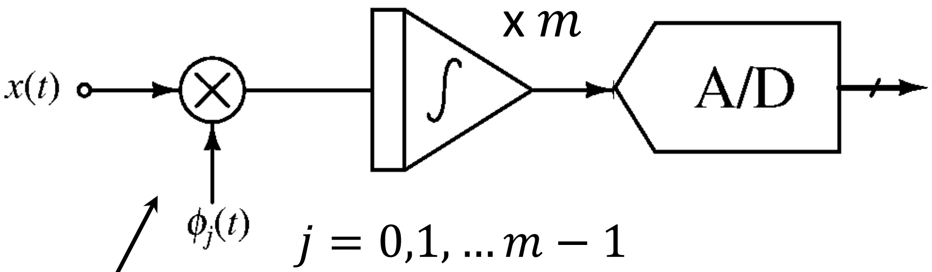
Binary antipodal RMPI, 16 channels
(external ϕ for maximum flexibility)

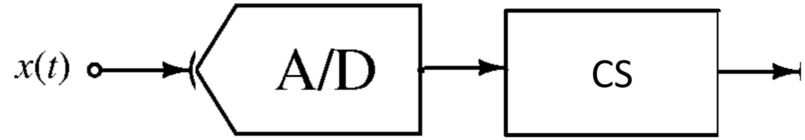




SAR

Buffer + 16 channel (distributed) multiplexer





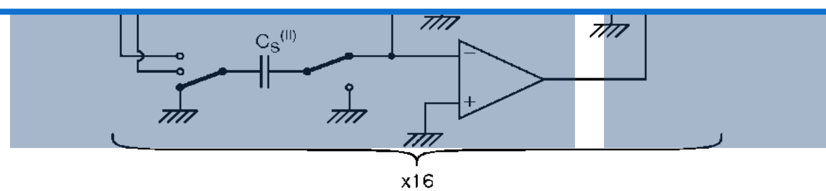
As reported in [1] a Digital Implementation Scheme of CS is more efficient in terms of energy consumption

$\Longrightarrow$

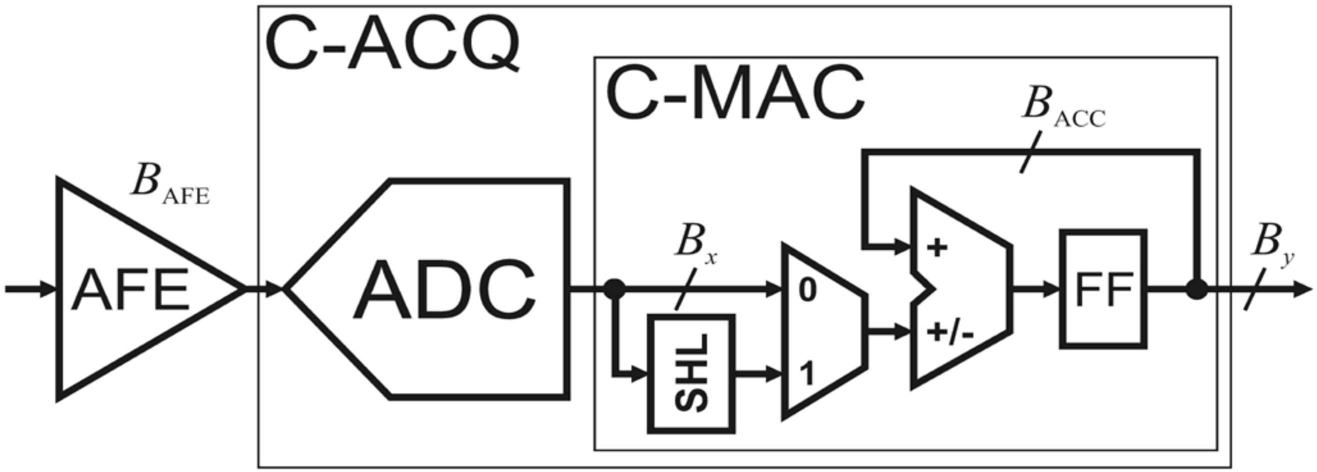
Main aim of this contribution is 1) to present an enhanced C-ACQ block based on a "puncturing" algorithm that yields the optimal sampling positions for a fixed number of sample and 2) and use it to estimate a single waveform parameter in the CS-domain

[1] F. Chen, A. P. Chandrakasan, V. M. Stojanovic, "Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," *IEEE Journal of Solid-State Circuits*, vol.47, no.3, pp.744,756, March 2012

Binar
(exte



Buffer + 16 channel (distributed) multiplexer



- The innovative part is in C-MAC (enhanced compressed multiply and accumulate) which computes the quantity

$$y_j = \sum_{k=0}^{n-1} \phi_{jk} x_k \qquad \phi_{jk} = \{-2, -1, 1, 2\} \cup \{0\}$$

- Multiplication by 2 correspond to SHL and by 0 correspond to avoid sampling
- Let B_{AFE}, B_x, B_{ACC} and B_y the (resolution) bits of the analog front-end the ADC, the accumulator and the C-MAC output
- In general m parallel stages are necessary (i.e. $j = 1,2, \dots, m$), but for a single parameter estimation we may assume $j = 1$

- Assume that the signal to acquire can be written as

$$x(t) = \alpha z(t) + \eta(t)$$

with $z(t)$ is known and represented by the samples $z = (z_0, \ldots, z_{n-1})^\top(t)$ is a realization of a zero mean disturbance processes producing the samples $\eta = (\eta_0, \ldots, \eta_{n-1})^\top$

- Starting from the variance of x, **we want to compute an estimate $\hat{\alpha}$ of α**

- Set $\hat{\alpha} = \frac{y}{\phi^\top z}$ and since $\mathbf{E}[\hat{\alpha}] = \frac{\alpha\phi^\top z}{\phi^\top z} + \frac{\phi^\top \mathbf{E}[\eta]}{\phi^\top z} = \alpha$ **we must minimize the estimator variance**

$$\min_{\phi} \sigma_{\hat{\alpha}}^2 = \min_{\phi} \mathbf{E}[(\hat{\alpha} - \alpha)^2] = \min_{\phi} \sigma_\eta^2 \frac{\phi^\top \phi}{(\phi^\top z)^2},$$

- Clearly ϕ cannot be chosen randomly since it may let to $\phi^\top z \mapsto 0$ and also $\phi = z$ (which would minimize $\hat{\alpha}$) is not possible, since $\phi \in \{-2, -1, 0, 1, 2\}^n$ and only s elements must be non zero to have s samples

PUNCTURING ALGORITHM

- A possible solution of the discrete optimization problem is given by

$J \leftarrow$ permute $\{0, \ldots, n-1\}$ s.t. $|z_{J(j)}| \geq |z_{J(j+1)}|$
for $j = 0, \ldots, n-2$

Initialize $a \leftarrow (0, \ldots, 0)^\top$ and $\bar{\omega} \leftarrow \infty$

for $j = 0, \ldots, s-1$ **do**

    $\bar{A} \leftarrow 0$

    for $A = 2, \ldots, 1$ **do**

$$\omega \leftarrow \frac{\phi^\top \phi + A^2}{(\phi^\top z + A|z_{J(j)}|)^2}$$

        if $\omega < \bar{\omega}$ **then**

            $\bar{\omega} \leftarrow \omega$ and $\bar{A} \leftarrow \mathrm{sign}(z_{J(j)})A$

        end if

    end for

    $\phi_{J(j)} \leftarrow \bar{A}$

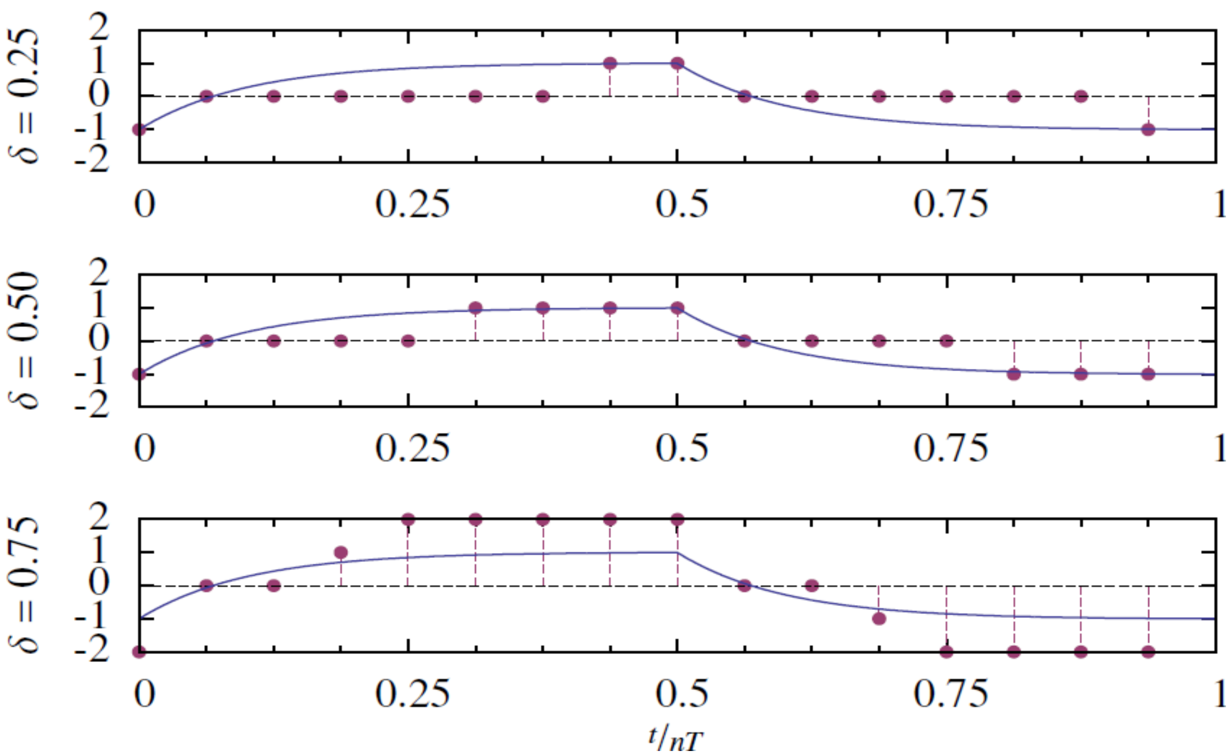
end for

$$\phi \leftarrow \frac{\phi}{\min\{|\phi_j| \text{ s.t. } \phi_j \neq 0, \;\; j = 0, \ldots, n-1\}}$$

which simply consider that when a large-modulus component of z is present the denominator tend to be larger than the numerator $\Rightarrow$ one orders z in non-decreasing order and associate it with the coefficient of ϕ minimizing ω (i.e. $\sigma^2_{\hat{\alpha}}$). Note that n and s are assumed known.

- Assume $z(t)$ and RC-type charge-discharge waveform (in the normalized $t \in [0,1]$ interval) and assume that $n = 16$ samples of which only $s = 4,8,12$ are used (i.e. the sample density $\delta = 0.25,0.5,0.75$)

$$z(t) = -1 + 2\begin{cases} \dfrac{1-e^{-\frac{t}{\tau}}}{1-e^{-\frac{1}{2\tau}}} & \text{if } t \leq \frac{1}{2} \\ 1 - \dfrac{1-e^{-\frac{(t-1/2)}{\tau}}}{1-e^{-\frac{1}{2\tau}}} & \text{if } t > \frac{1}{2} \end{cases}$$



Optimal sampling (in terms of noise rejection)

1. Concentrate when z is largest
2. The use of 4 instead of 2 values for ϕ becomes useful when δ increases and small values of z are sampled

- **Goal: estimate α with a resolution of B_α minimizing the energy necessary for the operation**. Of course $B_{\mathrm{AFE}} < B_\alpha$ and $B_\alpha \leq B_x$ (assume $B_\alpha = B_x$ at the maximum acquisition frequency f_s)
- The estimator will reduce the noise on α with respect to η, i.e.

$$\frac{\sigma_{\hat{\alpha}}^2}{\sigma_\eta^2} \leq 2^{2(B_{\mathrm{AFE}} - B_\alpha)}$$

so that $B_\alpha - B_{\mathrm{AFE}}$ is the **processing gain in the estimation**

- Six different sensor nodes as combination of:

1. Full sample (n samples) or sub-samples ($s < n$ samples) (**_FS** or **_SS**)
2. On-board or off-board estimation (**ON_** or **OFF_**)
3. Immediate transmission or storage in memory (**_TX** or **_NV**)

- **OFF_** means immediate transmission of all samples, i.e. $m = s$ and $B_y = B_\alpha$

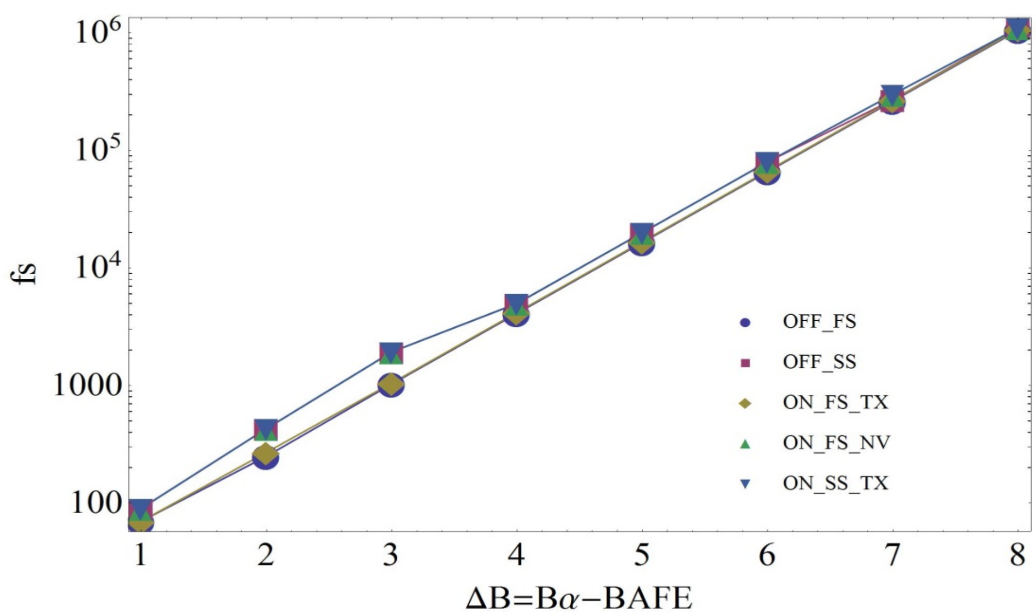
$$\text{Total Power} = \mathrm{P}_{\mathrm{AFE}} + \mathrm{P}_{\mathrm{ADC}} + \mathrm{P}_{\mathrm{TX}}$$

- **ON_** rely on local estimation, i.e. $m = 1$ and $B_y = B_\alpha$

$$\text{Total Power} = \mathrm{P_{AFE}} + \mathrm{P_{ADC}} + \mathrm{P_{C\text{-}MAC}} + \mathrm{P_{TX}}$$ **_TX**

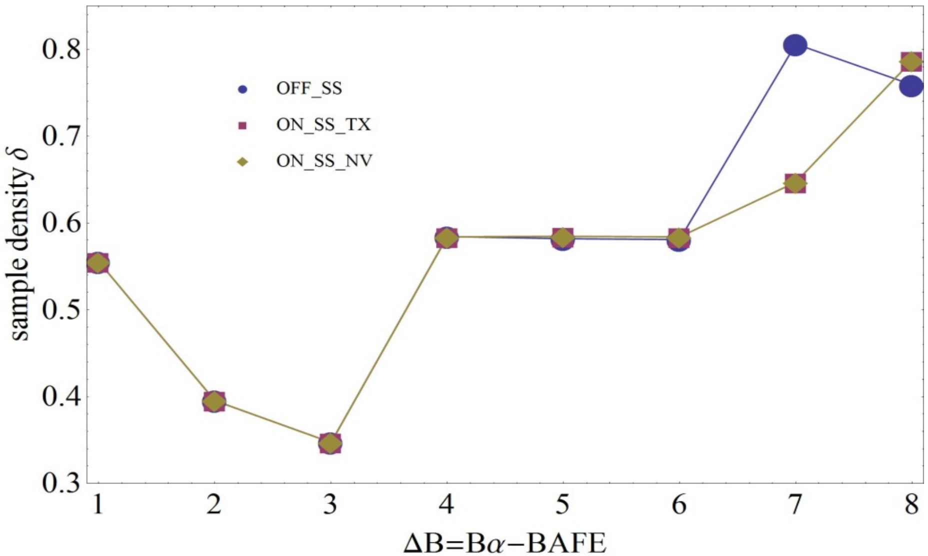
$$\text{Total Power} = \mathrm{P_{AFE}} + \mathrm{P_{ADC}} + \mathrm{P_{C\text{-}MAC}} + \mathrm{P_{NV}}$$ **_NV**

(no $\mathrm{P_{TX}}$ is present since the data is stored in memory till enough energy is harvested)

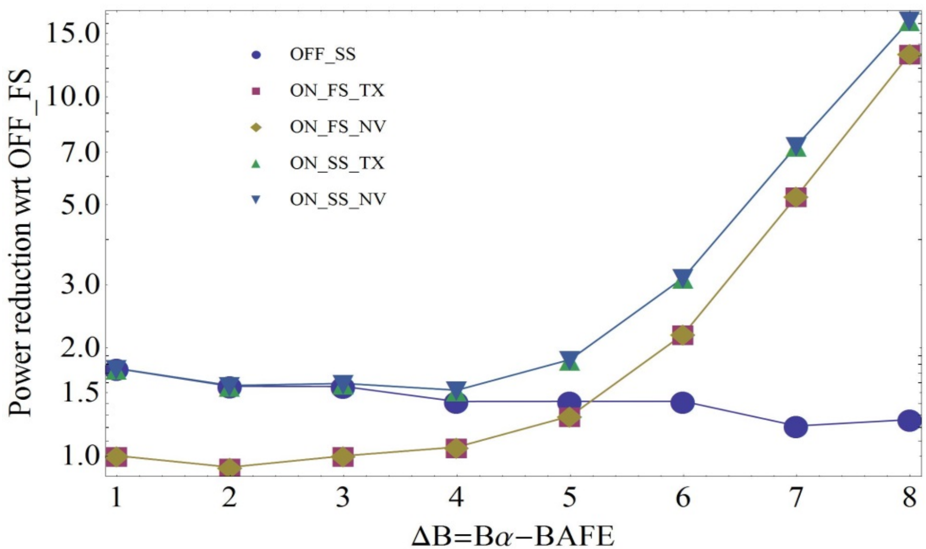


Determined using the Algorithm to find the couple n, s (i.e. f_s and δ) which gives a gain ΔB with minimum power consumption

$f_s\,(n)$ has practically the same behavior in all cases



Larger δ must be used to obtain larger ΔB i.e. (as seen before) use all the coefficient values that the architecture allow (i.e. 5 possible values for ϕ are better than 3)



Maximum reduction is obtained with ON_SS. The impact of the difference between TX and NV is not so important

- Compressed Sensing (CS) is a set of theoretical tools and algorithms which has found many applications when *sparsity* can be exploited, i.e., when we want to acquire ***signals with a structure***.

- This technique has been successfully applied to
  - Single-pixel Camera (CS Imager)
  - Distributed/wireless sensor networks
  - Magnetic Resonance Imaging (MRI, fMRI)
  - CS of ECG/EEG signals
  - Analog-to-Information Conversion (**RMPI**, Random Sampling,….)

- It is a field in which CASS community can strongly contribute since it requires ***algorithms/circuits co-design***, something on which our community is particularly strong

- **Rakeness-based Projection Waveforms**
  - I.i.d. random matrices Φ may not be the best choice in general.
  - When signals being sampled are *localized* in energy (e.g. their power spectrum is wideband but non-white, as it is almost always the case), one can use sequence maximizing rakeness to optimize performances

  M. Mangia, R. Rovatti, G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," IEEE Transactions on Circuits and Systems I, May 2012 (**2013 IEEE CAS Society Guillemin-Cauer Best Paper Award**)

  A. Caprara, F. Furini, A. Lodi, M. Mangia, Rovatti, G. Setti, "Generation of Antipodal Random Vectors With Prescribed Non-Stationary 2-nd Order Statistics," IEEE Transactions on Signal Processing, March, 2014

- **Information Concealing and Compressed Sensing**
  - Since Φ is sensitive to perturbation noise, we might change some coefficients (by sign flipping) and embed a mild form of cryptography directly in the A2I conversion.

  V. Cambareri, J.Haboba, F. Pareschi, R. Rovatti, G. Setti, K.W. Wong, «A Two-Class Information Concealing System based on Compressed Sensing", ISCAS 2013

  V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, "Low-Complexity Multiclass Encryption by Compressed Sensing Part I and Part II: Definition and Main Properties," IEEE Transactions on Signal Processing, under revision

SSIGPRO

- **Analog A2I Implementation**
  - First implementation of an Analog A2I including saturation checking and exploiting rakenss to achieve a truly hardware-software co-design

    P. Albertini, G. Frattini, M. Mangia, F. Pareschi, R. Rovatti, G. Setti, "An Analog-to-Information converter for Biomedical Signals based on Compressive Sensing," IEEE Transactions on Biomedical Circuits and Systems, submitted

    M. Mangia, F. Pareschi, R. Rovatti, G. Setti, G. Frattini, "Coping with Saturating Projection Stages in RMPI-Based Compressive Sensing," ISCAS2012, May 2012

- **Digital CS Front-end for BAN and WSN**
  - Design of a simple Digital CS-front end exploiting on-board feature extraction in BAN and WSN

    D. Bellasi, R. Rovatti, L. Benini, G. Setti, "An Architecture for Low-Power Compressed Sensing and Estimation in Wireless Sensor Nodes," ISCAS2014, June 2014

Thank you for your attention.

Questions?