CoSeRa 2013
17-19 September 2013
Bonn, Germany

Oceanographic Data Transmission: a Compressed Sensing-based Approach

Stefano Fortunati[#*], Raffaele Grasso[*], Fulvio Gini[#], Maria S. Greco[#],

[#]Dipartimento di Ingegneria dell'Informazione, University of Pisa,
via G. Caruso 16, 56122, Italy.

[*]CMRE, Viale San Bartolomeo 400, 19126 La Spezia.

Outline

- **Main concepts of the Compressed Sensing theory:**
  1. Sparsity and incoherence
  2. Signal recovery
  3. RIP property
- **Compressed data transmission:**
  1. Motivations for a CS-based transmission scheme
  2. Temperature data
  3. Simulation results and performance description
  4. Ongoing works

The Compressed Sensing theory (1/3)

Main statement: Certain signals or images can be recovered from far fewer samples or measurements than traditional methods use.

To make it possible, CS relies on two principles:

- ***Sparsity***, which pertains to the signal of interest,
- ***Incoherence***, which pertains to the sensing modality.

Sparsity

Definition: A signal $\mathbf{f}$ is said to be *s-sparse* if it has at most *s* nonzero entries.

Typically, one must deal with signals that are not themselves sparse, but which admit a sparse representation in some orthonormal basis $\mathbf{\Psi}$. In other words, the coefficient vector $\mathbf{x}$ of $\mathbf{f}$ through the orthonormal basis $\mathbf{\Psi}$ is sparse:

$$\mathbf{f} = \mathbf{\Psi}\mathbf{x} = \sum_{i=1}^{n} x_i \boldsymbol{\psi}_i, \quad \text{where} \quad x_i = \langle \mathbf{f}, \boldsymbol{\psi}_i \rangle.$$

The Compressed Sensing theory (2/3)

The sensing problem

Theoretical results show that if the signal admits a sparse representation in a given representation basis (such as the Fourier or the wavelet bases), then it is possible to reconstruct exactly the original signal from a very low number of *linear* and *non-adaptive* measurements.

Measurement model

- $\mathbf{f} \in \mathbb{R}^n$: original finite-dimensional signal,
- $\mathbf{\Psi} \in \mathcal{M}_{\mathbb{R}}^{n\times n}$: representation (orthogonal) matrix,
- $\mathbf{\Phi} \in \mathcal{M}_{\mathbb{R}}^{m\times n}$: sampling or measurement matrix with $m \ll n$

$$\boxed{\mathbf{y} = \mathbf{\Phi f} = \mathbf{\Phi \Psi x}, \quad \mathbf{y} \in \mathbb{R}^m}$$

Incoherence

Definition: The coherence between the sensing basis $\mathbf{\Phi}$ and the representation basis $\mathbf{\Psi}$ is defined as:

$$\mu(\mathbf{\Phi},\mathbf{\Psi}) = \sqrt{n} \max_{1\le k,j\le n} \left|\left\langle \boldsymbol{\varphi}_k, \boldsymbol{\psi}_j \right\rangle\right|, \quad \mu(\mathbf{\Phi},\mathbf{\Psi}) \in \left[1, \sqrt{n}\right].$$

The Compressed Sensing theory (3/3)

Sparse signal recovery

Given the measurement vector $\mathbf{y}$ and the knowledge that the signal $\mathbf{f}$ is sparse at least in some representation basis, it is natural to attempt to recover $\mathbf{f}$ by solving the following optimization problem:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_0, \quad \text{s.t. } \mathbf{\Phi\Psi x} = \mathbf{y}.$$

However, the l_0-norm is a discrete and non convex function and hence is potentially very difficult to solve this optimization problem (strictly speaking, the l_0-norm is not a norm, not being positive homogeneous).

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{R}^n} F(\mathbf{x}) \quad \text{s. t. } \mathbf{\Phi\Psi x} = \mathbf{y}$$

Basis Pursuit (BP)

Substitute the l_0-norm with its continuous and convex approximation: the l_1-norm.

$$F(\mathbf{x}) = \|\mathbf{x}\|_1 \triangleq \sum_i |x_i|$$

Smoothed l_0 algorithm*

Substitute the l_0-norm with some continuous but non-convex approximation.

$$F(\mathbf{x}) = n - \sum_{i=1}^{n} f_\sigma(x_i) \text{ where } f_\sigma(x) = e^{-\frac{x^2}{2\sigma^2}}.$$

***From:**

Mohimani, H.; Babaie-Zadeh, M.; Jutten, C.; "A Fast Approach for Overcomplete Sparse Decomposition Based on Smoothed Norm," IEEE Transactions on Signal Processing, vol. 57, no. 1, pp. 289-301, Jan. 2009.

The Restricted isometry property

Definition: A matrix **A** satisfies the restricted isometry property (RIP) of order *c*, if there exist a constant $\delta_c \in (0,1)$ such that the following inequality chain holds true:

$$\boxed{(1-\delta_c)\|\mathbf{x}\|_2^2 \le \|\mathbf{A}\mathbf{x}\|_2^2 \le (1+\delta_c)\|\mathbf{x}\|_2^2.}$$

It can be shown that that sensing matrices that satisfy the RIP can be sempled from two families of random matrices:

1. Form $\mathbf{\Phi}$ by sampling i.i.d. entries from the normal distribution with zero mean and variance $1/m$:

$$[\mathbf{\Phi}]_{ij} \sim \mathcal{N}(0,1/m)$$

2. Form $\mathbf{\Phi}$ by sampling i.i.d. entries from symmetric Bernoulli distributions, e.g.:

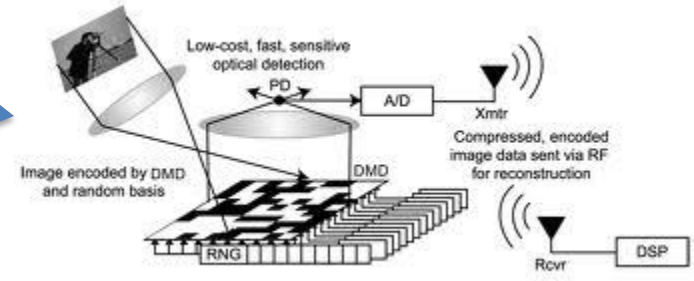
$$[\mathbf{\Phi}]_{ij} = \begin{cases} +1/\sqrt{m} & \text{with probability } 1/2, \\ -1/\sqrt{m} & \text{with probability } 1/2, \end{cases}$$

It must be noted that the identity matrix **I** does not satisfy the RIP.

Some application of CS for data acquisition (1/2)

Remote sensing: Improvement of the imaging devices that use CCD (charge coupled device) array technology. A CS camera that collects incoherent measurements using a digital micromirror array requires just one photosensitive element instead of millions.

One-pixel camera



• **Compressed sensing of hyperspectral data**: The application of CS to hyperspectral imaging has the potential for significantly reducing the sampling rate and hence the cost of the analog-to-digital sensors. Moreover, the dimension of the HS data cube could be significantly reduced.

• **SAR imaging**: Basically, a SAR measurement can be viewed as a sample of the spatial Fourier transform of the scattering field of interest. The CS theory could be very useful to reconstruct the entire scene with a very low number of SAR measurements.

Some application of CS for data acquisition (2/2)

- **Sonar and radar signal processing application:** CS-based space-time adaptive processing (STAP) algorithms, radar imaging, waveform selection and so on.
- **Mission design for mobile robot using CS**: Mobile robots are increasingly being used to survey and map spatial phenomena for large-scale environmental monitoring applications. In a lot of cases the domain to be sampled is very large, and the CS theory could provide a valuable tool to reduce the number of robots needed for the mission and to optimize the data gathering.

Application of our interest: Sampling and transmission of oceanographic data from autonomous underwater gliders to the control centre.

Compressed data transmission (1/2)

In all the data acquisition missions, in addition to the sampling problem, there is another fundamental practical issue: the transmission of the collected data.

Motivation: Since the data transmission process is an expansive procedure in terms of both battery consumption and monetary cost, an in order to minimize the time on surface, it could be useful to find an algorithm that is able to reconstruct the ocean field of interest using a small part of the acquired data.

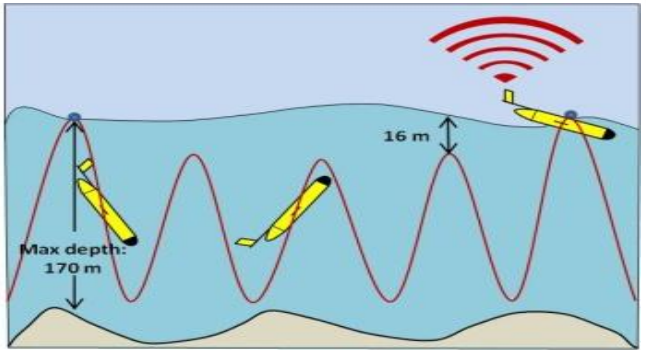
A possible approach: Use the CS framework to select a certain number of linear CS measurements of the acquired data to be transmitted to the receiver for the field reconstruction. The receiver could then reconstruct the original data by performing a sparse minimization algorithm (i.e. the Basis Pursuit).

Important observation: In order to exploit CS techiques at the maximum extent, the signal of interest should be sampled by using CS hardware devices (e.g. the *random demodulator*) operating directly at a sub-Nyquist rate. This part is left to future works.

Compressed data transmission (2/2)

Data acquisition

The CS framework is used to select a certain number of linear CS measurements $\mathbf{y}$ of the data $\mathbf{f}$, acquired by the glider, to be transmitted to the receiver for the field reconstruction.



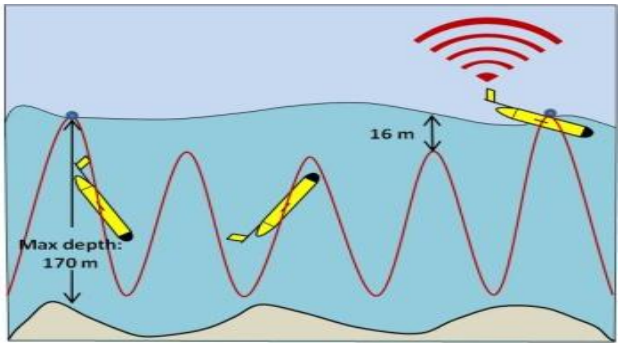
$\mathbf{y}=\Phi\mathbf{f}$

Data reconstruction

In the processing center of the receiver, where processing power and battery consumption is not a problem, the original field data $\mathbf{f}$ are reconstructed using a sparse minimization algorithm.

Two possible approaches to arrange the compressed measurement vector

First approach: Continuously sample the water, collect all the acquired data in a vector $\mathbf{f}$ and finally arrange the measurement vector $\mathbf{y}$ to be transmitted as:



$$\mathbf{y} = \mathbf{\Phi f}.$$

mx1 vector of the measurements to be transmitted

measurement matrix

nx1 vector of all the collected data

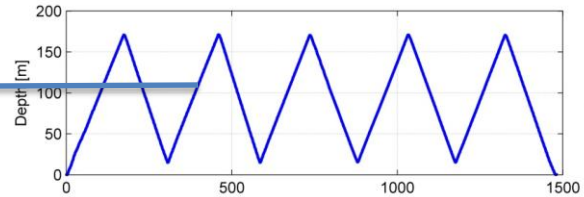
In this approach, it can be used all the possible sensing matrices (identity, Bernoulli and Gaussian matrices) since we have at our disposal all the data continuously acquired by the glider.

Second approach: Randomly sample the water in order to directly construct the measurement vector $\mathbf{y}$ to be transmitted in the data acquisition phase. This approach is equal to the first one in which the identity matrix is used:

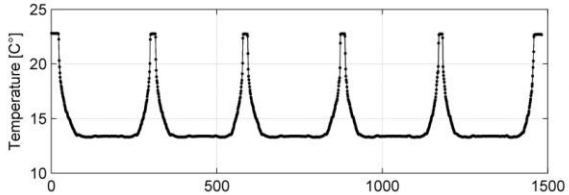
- Generate at random m rows of an nxn identity matrix,
- Collect in a vector $\mathbf{y}$ the m measurements relative to the non zero entries of the m rows,
- Transmit the vector $\mathbf{y}$ for the CS field reconstruction.

An example of collected data: the temperature

i^{th} diving-climbing path of the glider

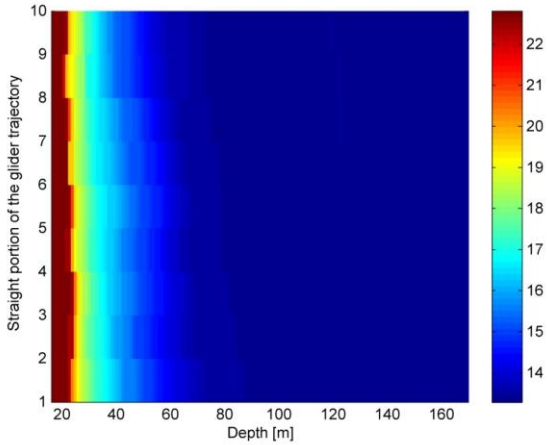


Saw tooth trajectory of the glider



1-dimensional signal of the temperature collected during a dive cycle of the glider

Because of the strong correlation between the data collected during two adjacent diving-climbing paths, it would be useful to rearrange the data as a matrix obtained interpolating the measured field values during the ith path in a regular depth grid of p points. More precisely, the ith row of the data matrix will contain the interpolated field value collected by the glider during the ith path.



Performance index

The performance analysis is performed in terms of mean value and standard deviation of the RMSE with respect to the random measurement matrix of the RMSE. More precisely, let $N_{\mathbf{f}}$ the dimension of the signal $\mathbf{f}$ and let $\hat{\mathbf{f}}$ be the estimated signal, we can define the root mean square of the reconstruction error, $\mathrm{RMSE}(\mathbf{\Phi})$, as:

$$\mathrm{RMSE}(\mathbf{\Phi}) = \sqrt{\frac{1}{N_{\mathbf{f}}}\sum_{i=1}^{N_{\mathbf{f}}}\left(\mathbf{f}_i - \hat{\mathbf{f}}_i(\mathbf{\Phi})\right)^2}$$

The $\mathrm{RMSE}(\mathbf{\Phi})$ is a function of the particular realization of the measurement matrix $\mathbf{\Phi}$. More precisely, $\mathrm{RMSE}(\mathbf{\Phi})$ is itself a random variable whose mean value μ is defined as:

$$\mu = E_{\mathbf{\Phi}}\left\{\mathrm{RMSE}(\mathbf{\Phi})\right\}$$

The mean value μ represents the root mean square reconstruction error averaged over all the possible realization of the measurement matrix $\mathbf{\Phi}$. The average RMSE will be evaluated for different values of the compression ratio (*CR*) simply defined as:

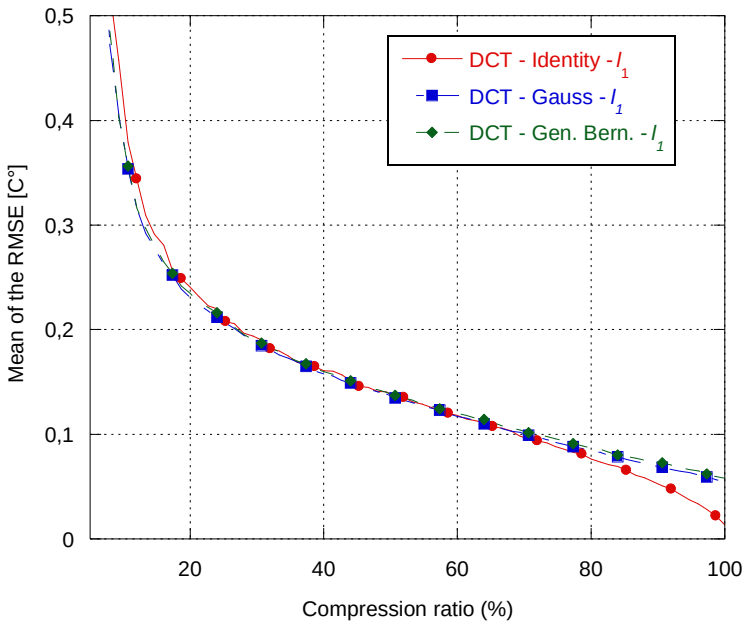
$$CR = \frac{m}{n}$$

Number of transmitted CS measurements

Total amount of samples in the original data

RMSE for three different types of sensing matrices

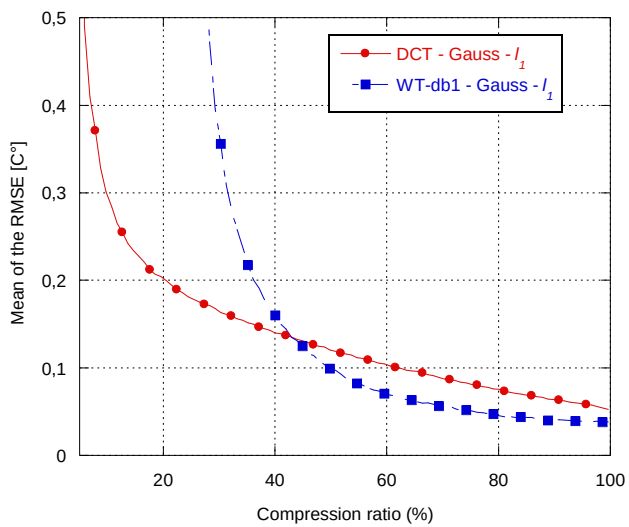
- Representation matrix $\mathbf{\Psi}$: 2D – Discrete Cosine Transform (DCT).
- Sensing matrices $\mathbf{\Phi}$:
  - Identity matrix (that do not satisfy the RIP),
  - Gaussian and Bernoulli matrices (that satisfy the RIP).
- Minimization: Basis Pursuit problem (constrained ℓ_1-norm minimization).

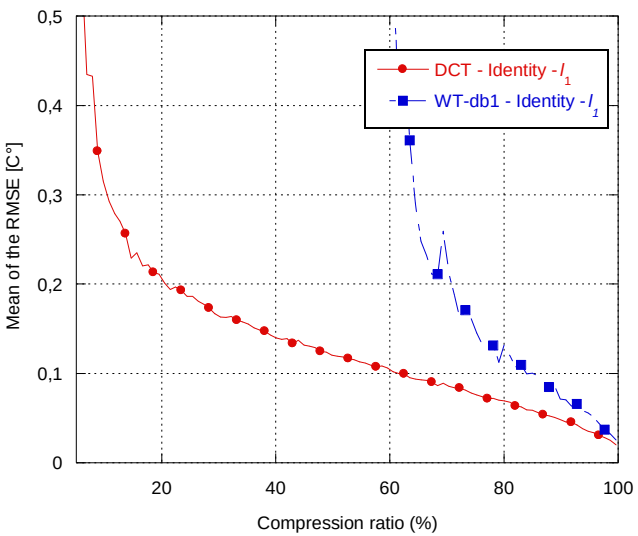


- With the 2D-DCT representation matrix, the identity matrix guarantees the same reconstruction performance as the Gaussian and Bernoulli matrices.
- The random sampling scheme can be implemented without performance loss providing that the 2D-DCT represenation basis is used.
- Using a compression ratio of 60%, the temperature field can be reconstructed with a RMSE of about 0.1°C.

RMSE for two representation matrices

- Representation matrices $\mathbf{\Psi}$: 2D – DCT and 2D – WT with a db1 wavelet.
- Sensing matrices $\mathbf{\Phi}$: Gaussian and Identity matrices.

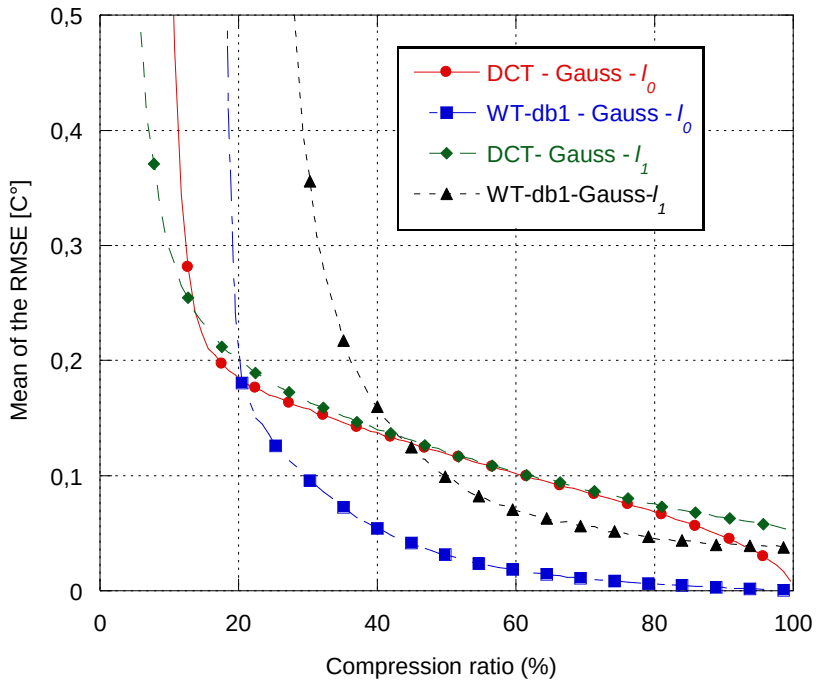




- With a Gaussian sensing matrix, the DCT works better than the WT for low CR values. As the CR increases, the WT outperforms the DCT basis.
- The WT basis is not suitable for the Identity sensing matrix.
- Then, if a random sampling strategy is implemented, the WT basis should not be used as representation basis.

Basis Pursuit vs smoothed ℓ_0-minimization

- Representation matrix $\mathbf{\Psi}$:
  - 2D – Discrete Cosine Transform (DCT),
  - 2D – Wavelet Transform (WT) with a db1 wavelet.
- Sensing matrices $\mathbf{\Phi}$: Gaussian matrix.



- When the DCT is used, the performance of the two minimization algorithms are similar,
- Using the WT, the smoothed ℓ_0-minimization has much better performance than the Basis Pursuit.
- The smoothed ℓ_0-minimization has a low computatinal cost that the Basis Pursuit.
- The better behaviour of the smoothed ℓ_0-minimization has been verified for all the available data, but not general statement can be drown from these results.

Concluding remarks and ongoing work

Goals of the proposed CS application

1. Pre-processing prior to transmission of the measured data using a limited bandwidth satellite link and saving the battery energy,
2. The possibility to transmit only a certain number of CS measurements (instead of all the collected data), and then to implement a reconstruction algorithm at the receiver, was investigated.
3. The performance of the proposed CS algorithm has been assessed in terms of RMSE vs compression ratio for different sensing matrices (Identity, Gaussian and Bernoulli matrices), for two representation bases (DCT and WT) and for two minimization algorithms (BP and smoothed ℓ_0-minimization).

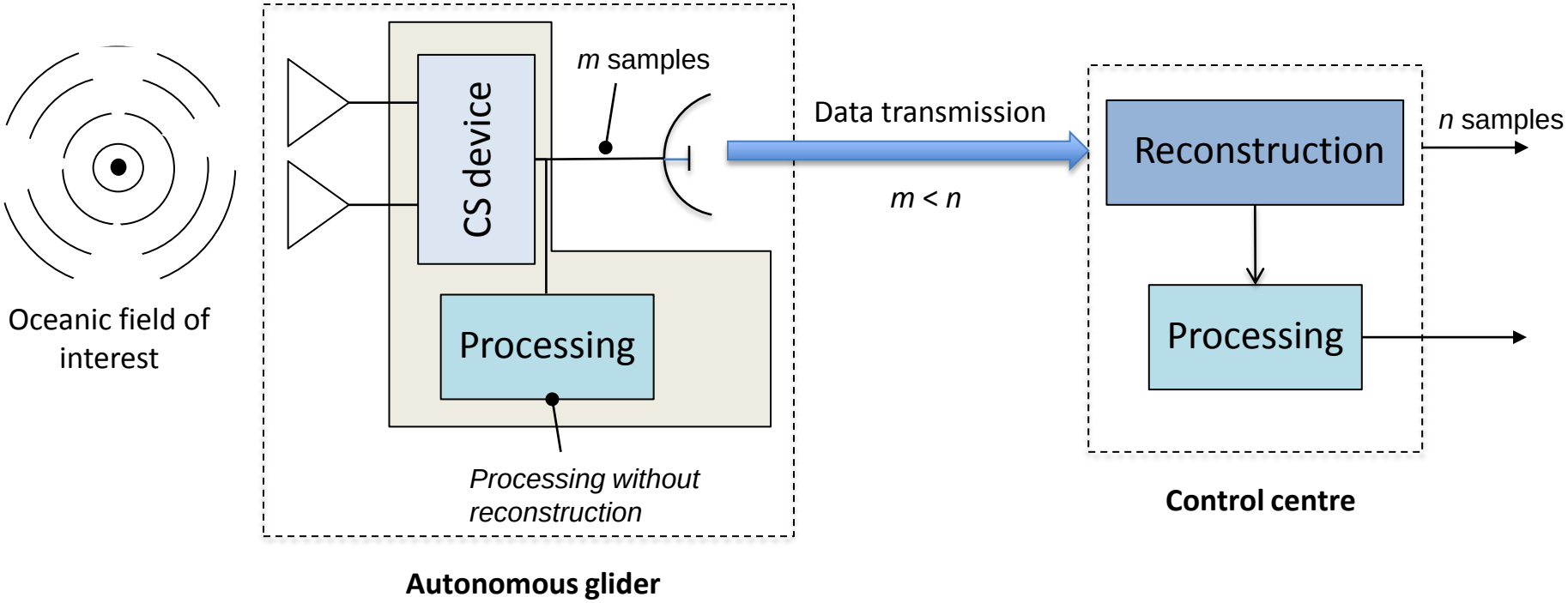
Main results

1. The random sampling (i.e. sensing with the Identity matrix) of the sea column is an effective scheme when the DCT representation matrix is used,
2. The smoothed ℓ_0-minimization provides equal or even better performance (in particular with the WT) than the Basis Pursuit.

Ongoing works (1/5)

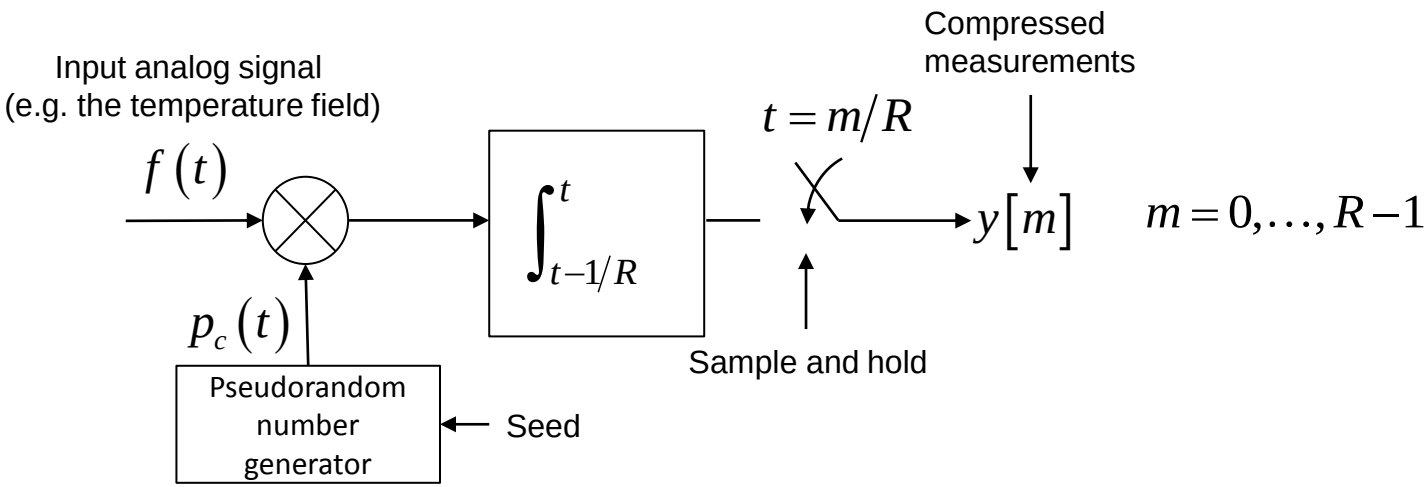
CS hardware devices for sub-Nyquist sampling

Development of sampling devices that are able to directly gather data of the field of interest without the need of a post-processing with a random measurement matrix.



Ongoing works (2/5)

An example of a possible CS device: the random demodulator



Matrix interpretation of the random demodulator

Compressed measurements $\mathbf{y} = \mathbf{\Phi f}$

$$\mathbf{H} = \begin{pmatrix} 111 & & \\ & 111 & \\ & & 111 \end{pmatrix}, R \times N \qquad \mathbf{D} = \mathrm{diag}(1,1,-1,1,\cdots,-1),\ N \times N$$

Random demodulator measurement matrix $\mathbf{\Phi} = \mathbf{HD\Psi}$

Representation matrix

Chip sequence

From:

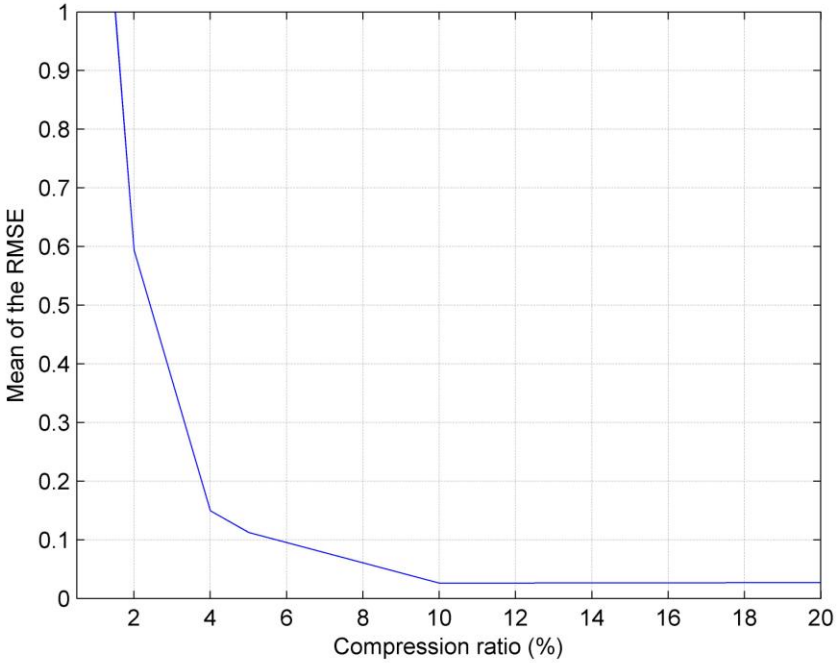
Tropp, J.A.; Laska, J.N.; Duarte, M.F.; Romberg, J.K.; Baraniuk, R.G., "Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals," *Information Theory, IEEE Transactions on* , vol.56, no.1, pp.520,544, Jan. 2010.

Ongoing works (3/5)

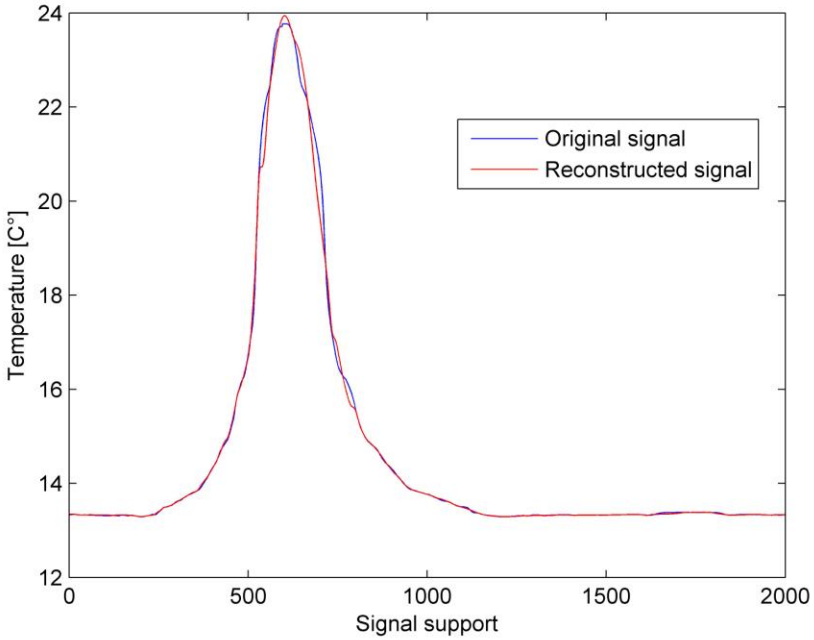
Original signal: One temperature profile (2000 sample),

Sampling device: random demodulator.

RMSE of the reconstructed signal for different values of the compression ratio



Example of the reconstructed profile using 80 compressed measurements (CR=4%)



Ongoing works (4/5)

Oceanic field reconstruction using a fleet of autonomous gliders

All the gliders implement a random demodulator. Then, for each sampling instant, the following **measurement equation** holds:

$$\mathbf{y}_k = \mathbf{\Phi}_k \mathbf{\Psi}_k \mathbf{x}_k + \mathbf{e}_k$$

Random demodulator measurement matrix

Coefficient vector (sparse vector)

Measurement noise

$$\mathbf{y}_k = \left(y_{1,k}, y_{2,k}, \cdots, y_{N_g,k} \right)^T$$

$$N_g = \text{num of gliders}$$

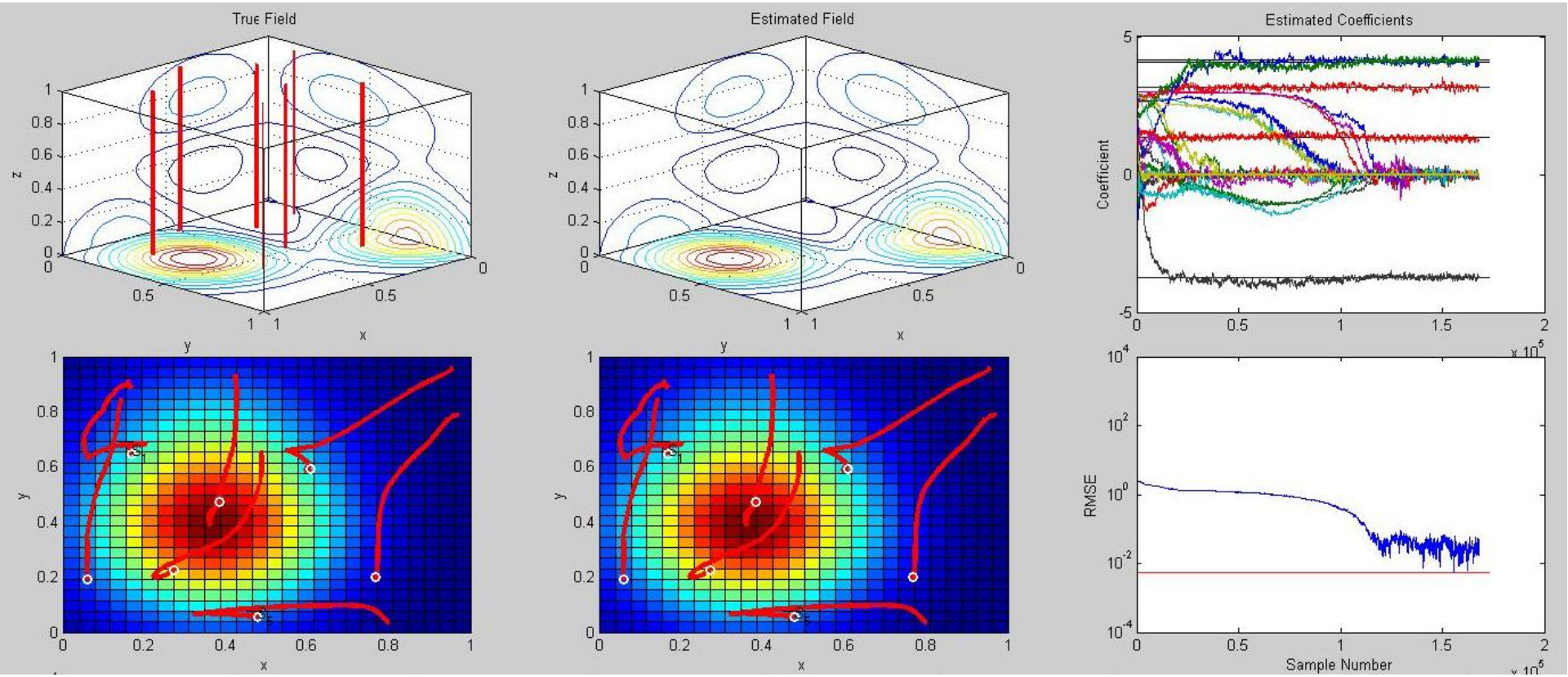
Each component of the measurement vector represents the measure of the *ith* glider at the *kth* sampling instant.

State space modeling (assumed constant or slow varying)

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{n}_k \qquad \mathbf{x}_k = \left(x_{1,k}, x_{2,k}, \cdots, x_{N_g,k} \right)^T$$

- The coefficient vector of the field of interest can be estimated with a *sparity-aware* (*centralized or decentralized*) *Kalman filter*.
- The trajectory of each glider of the fleet can be recursively set up in order to minimize the reconstruction error.

Ongoing works (5/5)



Thanks for your attention!