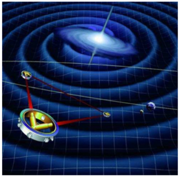
DE LA RECHERCHE À L'INDUSTRIE

CEA - Saclay

www.cea.fr

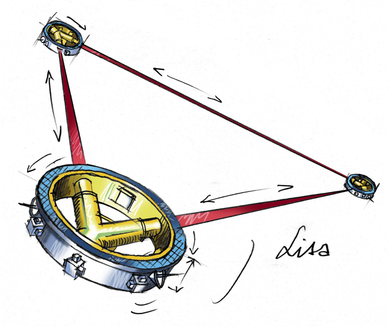
Learning-based representation of gravitational wave signals for LISA data analysis



Cinquième Assemblée Générale du GdR Ondes Gravitationnelles | Aurore BLELLY

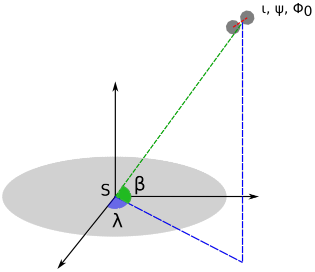
Oct. 12, 2021

- 3 satellites 2.5 millions km one from another
- Follows the movement of the Earth on its orbit
- Probes a **frequency range that is still unexplored** $\Rightarrow$ high potential for **new discoveries**
- Launching: 2034



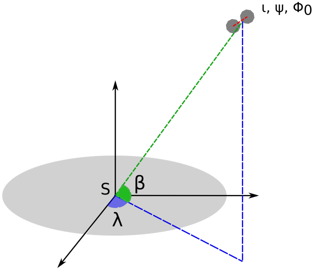
[1]

[1] https://www.esa.int/Science_Exploration/Space_Science/LISA



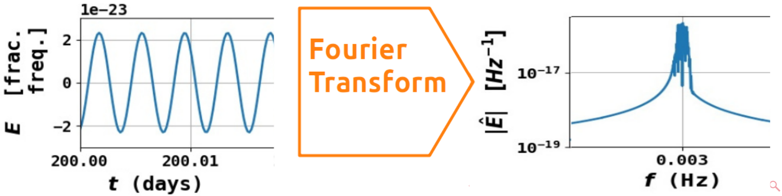
$+\mathcal{A}, f_0, \dot{f}$

- 20.000 expected observable systems, **8** parameters/system
- Standard approach: parameter estimation (**MCMC** / Bayesian approach)
- **High computing cost** both for waveform production and parameter estimation.



$+\mathcal{A}, f_0, \dot{f}$

- 20.000 expected observable systems, **8** parameters/system
- Standard approach: parameter estimation (**MCMC** / Bayesian approach)
- **High computing cost** both for waveform production and parameter estimation.

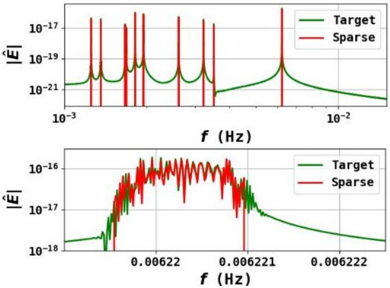


Objectives

- 2 main objectives:
  - Detection
  - Signal estimation/extraction
- Develop a method that can be adapted for **any type of sources**
  - GBs, MBHMs,MBHBs, glitches
- Fast **generative** parameter-free model
  - good fast representation of sought signals

Sparse representation in Fourier Basis [2]

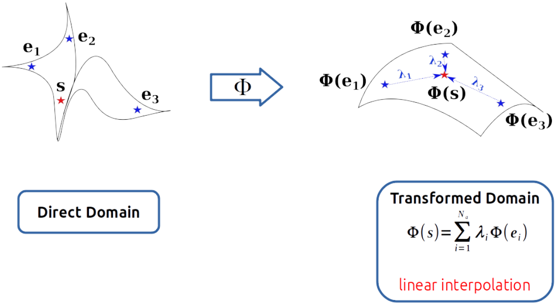
Previous work



Signal estimation using a sparsity prior

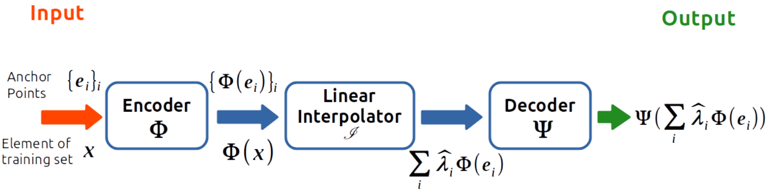
- Took part in LDC1-3: all sources detected
- Gaps: sources deterioration
- Recovered signal only partially represents the sought source, no possibility to separate sources.

[2] A. Blelly, J. Bobin, H. Moutarde, *Sparsity Based Recovery of Galactic Binaries Gravitational Waves*



- Find a space where the signal can be represented **simply**.
- Build a representation upon the availability of samples: "Anchor Points"
- $\sim$ Non-linear dictionary learning

[3] J.Bobin, R.Carloni Gertosio, C.Bobin, C.Thiam, *Non-linear interpolation learning for example-based inverse problem regularization*

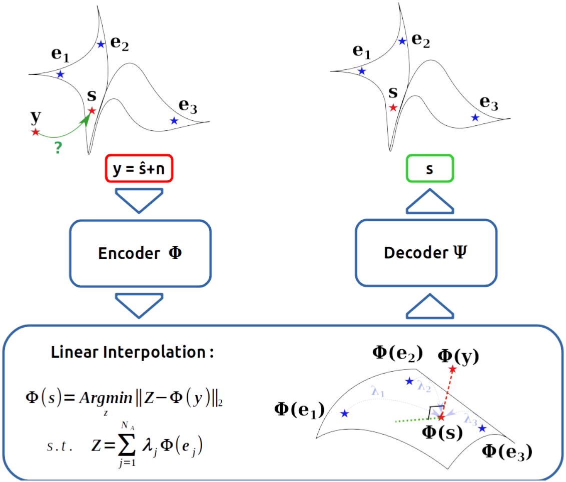


- **Linear interpolator**:
  $\{\widehat{\lambda}_i\} = \mathrm{Argmin}_{\{\lambda_i\}_i} \|\Phi(x) - \sum_i \lambda_i \Phi(e_i)\|_2^2$
- If all goes well: $\exists\{\lambda_i\}_i$ s.t. $x \simeq \Psi\left(\sum_i \widehat{\lambda}_i \Phi(e_i)\right)$
- Encoder / Decoder learned as:

$$\Phi, \Psi = \underset{\Phi,\Psi}{\mathrm{Argmin}} \sum_{x \in \mathcal{T}} \left\| x - \Psi\left(\sum_i \widehat{\lambda}_i \Phi(e_i)\right)\right\|^2$$

Interpolatory Auto Encoder (IAE)

Manifold projection for inverse problem



Introduction
LISA mission
Galactic Binaries
IAE
Preliminary Results
Conclusion

- 2 dense hidden layers both for Φ and Ψ
  - 1 layer $= 128$ neurons
  - input size $= 128$ frequency bins (waveforms in Fourier domain over 1 year observation, cut around main frequency)

[4] A. Blelly, J. Bobin, H. Moutarde, *Sparse data inpainting for the recovery of Galactic-binary gravitational wave signals from gapped data*

- 2 dense hidden layers both for Φ and Ψ
  - 1 layer $= 128$ neurons
  - input size $= 128$ frequency bins (waveforms in Fourier domain over 1 year observation, cut around main frequency)
- Training set:
  - 5000 waveforms (fixed frequency for now)
  - can be reduced down to 1000 samples with no major quality degradation

[4] A. Blelly, J. Bobin, H. Moutarde, *Sparse data inpainting for the recovery of Galactic-binary gravitational wave signals from gapped data*

Interpolatory Auto Encoder (IAE)

Implementation details

- 2 dense hidden layers both for Φ and Ψ
  - 1 layer $= 128$ neurons
  - input size $= 128$ frequency bins (waveforms in Fourier domain over 1 year observation, cut around main frequency)
- Training set:
  - 5000 waveforms (fixed frequency for now)
  - can be reduced down to 1000 samples with no major quality degradation
- Anchor points: between 7 and 50 (ideally: manifold dimension $+ 1$)

[4] A. Blelly, J. Bobin, H. Moutarde, *Sparse data inpainting for the recovery of Galactic-binary gravitational wave signals from gapped data*

- 2 dense hidden layers both for Φ and Ψ
  - 1 layer $= 128$ neurons
  - input size $= 128$ frequency bins (waveforms in Fourier domain over 1 year observation, cut around main frequency)
- Training set:
  - 5000 waveforms (fixed frequency for now)
  - can be reduced down to 1000 samples with no major quality degradation
- Anchor points: between 7 and 50 (ideally: manifold dimension + 1)
- **Fast training**: $\sim$ 20 min on a **laptop**

[4] A. Blelly, J. Bobin, H. Moutarde, *Sparse data inpainting for the recovery of Galactic-binary gravitational wave signals from gapped data*

- 2 dense hidden layers both for Φ and Ψ
  - 1 layer $=128$ neurons
  - input size $=128$ frequency bins (waveforms in Fourier domain over 1 year observation, cut around main frequency)
- Training set:
  - 5000 waveforms (fixed frequency for now)
  - can be reduced down to 1000 samples with no major quality degradation
- Anchor points: between 7 and 50 (ideally: manifold dimension $+ 1$)
- **Fast training**: $\sim$ 20 min on a **laptop**
- **"Plug and Play"** method: can directly be used in algos already developed. (ex: inpainting algorithm [4])

[4] A. Blelly, J. Bobin, H. Moutarde, *Sparse data inpainting for the recovery of Galactic-binary gravitational wave signals from gapped data*

Our quality estimator: Normalized Mean Square Error (NMSE)

$$NMSE = -10\log_{10}\left[\frac{\left\|y^{true} - y^{est}\right\|^2}{\left\|y^{true}\right\|^2}\right]$$

- High ($>> 1$) when y^{est} is close to y^{true}
- Low / negative when estimation is bad.

Our quality estimator: Normalized Mean Square Error (NMSE)

$$NMSE = -10\log_{10}\left[\frac{\left\|y^{true} - y^{est}\right\|^2}{\left\|y^{true}\right\|^2}\right]$$

- High ($>> 1$) when y^{est} is close to y^{true}
- Low / negative when estimation is bad.

We will look at its **distribution** on a given test set (1000 waveforms with different parameters).

Our quality estimator: Normalized Mean Square Error (NMSE)

$$NMSE = -10\log_{10}\left[\frac{\left\|y^{true} - y^{est}\right\|^2}{\left\|y^{true}\right\|^2}\right]$$

- High ($>> 1$) when y^{est} is close to y^{true}
- Low / negative when estimation is bad.

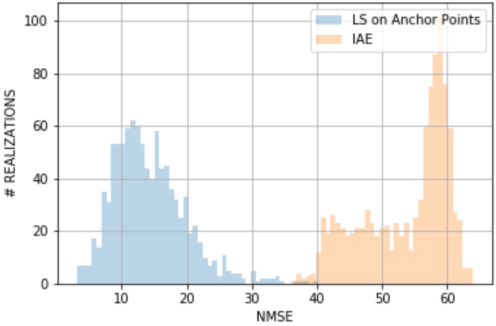
We will look at its **distribution** on a given test set (1000 waveforms with different parameters).

We will compare the quality of signal:

- estimated by IAE
- directly estimated on the span of anchor points in direct domain.

A first trial with 20 APs

Distribution of quality of restored signal on test set

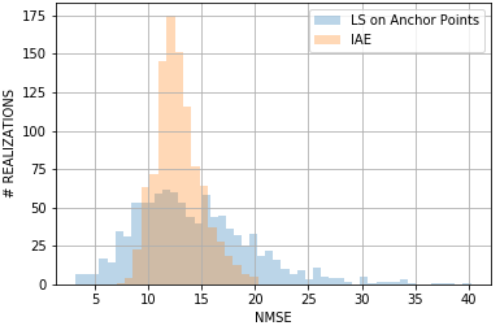


Reconstruction of **non-noisy** signals

- LS: Least Squares (simple projection on APs)
- **NMSE**: quality of signal estimate

A first trial with 20 APs

Distribution of quality of restored signal on test set



Reconstruction of **noisy** signals

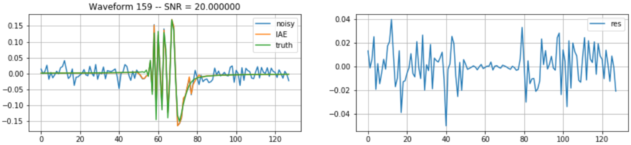
- LS: Least Squares (simple projection on APs)
- **NMSE**: quality of signal estimate

$\Longrightarrow$ Noise robustness issue?

Over-fitting noisy signal

A criterion to set the number of anchor points

Noisy signals:



Proposed solutions:

- Decrease the number of anchor points:

$$dim(\text{Manifold}) = \#\text{Parameters}$$

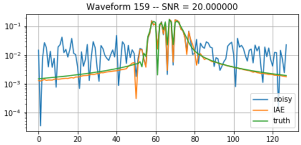
$$\#\text{APs} = dim(\text{Manifold}) + 1$$

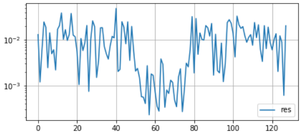
- Change model / interpolation function

Over-fitting noisy signal

A criterion to set the number of anchor points

Noisy signals:





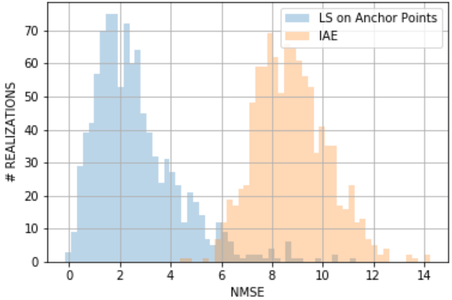
Proposed solutions:

- Decrease the number of anchor points:

$$dim(\text{Manifold}) = \#\text{Parameters}$$

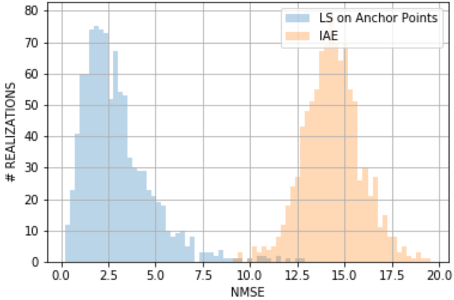
$$\#\text{APs} = dim(\text{Manifold}) + 1$$

- Change model / interpolation function



Quality of reconstructed signal for $SNR = 5$

- For $\#AP = 8$
- Stronger constraint on sought signal

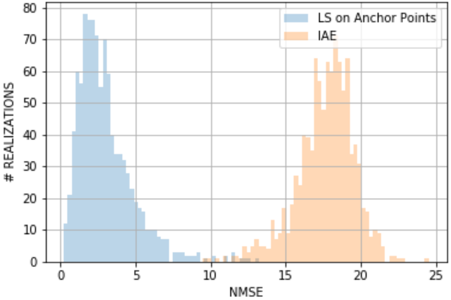


Quality of reconstructed signal for $SNR = 20$

- For $\#AP = 8$
- Stronger constraint on sought signal

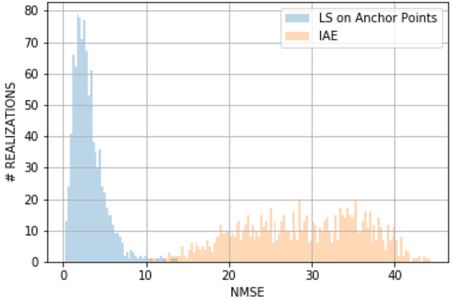
Drastically reducing the number of anchor points

Toward a "best" adapted representation



Quality of reconstructed signal for $SNR = 50$

- For $\#AP = 8$
- Stronger constraint on sought signal

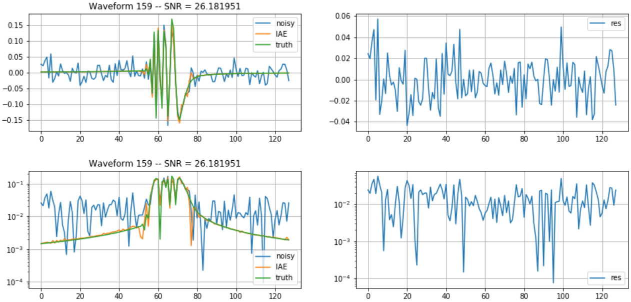


Quality of reconstructed signal for $SNR = \infty$ (= no noise)

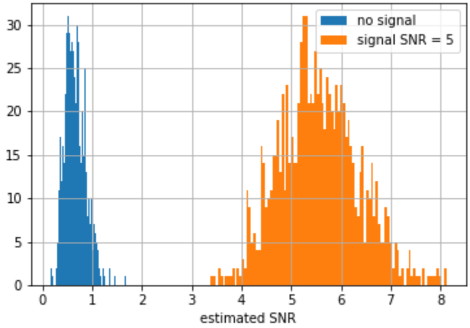
- For $\#AP = 8$
- Stronger constraint on sought signal

Recovered waveforms for 7 APs

Reduced noise over-fitting



- Greatly reduced noise over-fitting !
- Really good recovery of waveform tails, even below noise level



Estimated signal SNR for noise-only inputs and for $SNR = 5$ inputs.

- **P-value test:** compare estimated SNR with distribution obtained for noise-only inputs.
- seems to work even for weak amplitude inputs signals!

Prospects and future works

- **Work in Progress:**
  - Fast generative model to extract waveforms from measurements
- **Further prospects:**
  - Current results shown at **fixed frequency**. Preliminary tests on a whole frequency range are encouraging.
  - Link representation between different data channels to improve robustness against noise
  - Trying to link the learned space to parameter space in order to estimate parameters.

Commissariat à l'énergie atomique et aux énergies alternatives
Centre de Saclay | 91191 Gif-sur-Yvette Cedex
| F. +33(0)1 69 08 75 84

DRF
Irfu
DPhN

Etablissement public à caractère industriel et commercial | R.C.S. Paris B 775 685 019

Anchor Points choice

- Choose the best anchor points using $\sim$ **reverse basis-pursuit** algorithm. We start from a wide set of anchor points and we eliminate the ones carrying the **least information**.

- Choose the best anchor points using ∼ **reverse basis-pursuit** algorithm. We start from a wide set of anchor points and we eliminate the ones carrying the **least information**.
- Lots of APs = lots of information redundancy.

- Choose the best anchor points using $\sim$ **reverse basis-pursuit** algorithm. We start from a wide set of anchor points and we eliminate the ones carrying the **least information**.
- Lots of APs = lots of information redundancy.
- We regularize by adding a sparsity constraint to the interpolator:

$$\{\widehat{\lambda}_i\} = \underset{\{\lambda_i\}_i}{\mathrm{Argmin}} \left\| \Phi(x) - \sum_i \lambda_i \Phi(e_i) \right\|_2^2 + \sum_i \underbrace{\gamma_i}_{\text{learned threshold}} |\lambda_i|$$

Anchor Points choice

- Choose the best anchor points using $\sim$ **reverse basis-pursuit** algorithm. We start from a wide set of anchor points and we eliminate the ones carrying the **least information**.
- Lots of APs = lots of information redundancy.
- We regularize by adding a sparsity constraint to the interpolator:

$$\{\widehat{\lambda}_i\} = \underset{\{\lambda_i\}_i}{\mathrm{Argmin}} \left\| \Phi(x) - \sum_i \lambda_i \Phi(e_i) \right\|_2^2 + \sum_i \underbrace{\gamma_i}_{\text{learned threshold}} |\lambda_i|$$

- **REMARKABLE OUTCOME:** the most used anchor points are always the same!

Anchor Points choice

- Choose the best anchor points using $\sim$ **reverse basis-pursuit** algorithm. We start from a wide set of anchor points and we eliminate the ones carrying the **least information**.
- Lots of APs = lots of information redundancy.
- We regularize by adding a sparsity constraint to the interpolator:

$$\{\widehat{\lambda}_i\} = \underset{\{\lambda_i\}_i}{\mathrm{Argmin}} \left\| \Phi(x) - \sum_i \lambda_i \Phi(e_i) \right\|_2^2 + \sum_i \underbrace{\gamma_i}_{\text{learned threshold}} |\lambda_i|$$

- **REMARKABLE OUTCOME:** the most used anchor points are always the same!
- Compromise between over-fitting and under-fitting the data

- **SNR**: Signal-to-Noise Ratio

$$SNR(X) = \mathbf{E}\left[\frac{|X|^2}{\sigma^2}\right]$$

The greater it is, the more the signal is above noise level / detectable

- **P-Value:**
  - Test H_0: *there is no signal* against H_1: *there is at least one signal*
  - Fixing a threshold for decision on **estimated SNR**