Learning-based models for gravitational wave analysis

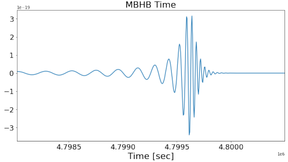
Elie Leroy
Supervision J. Bobin and H. Moutarde

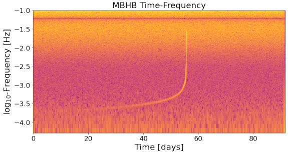
CEA Saclay - IRFU

12 October 2022

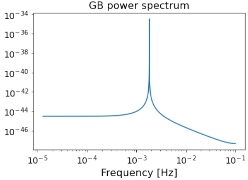
Expected sources & signals in the LISA data

Massive Black Hole Binaries (MBHB)
- Loud chirps
- 1 every 3 days





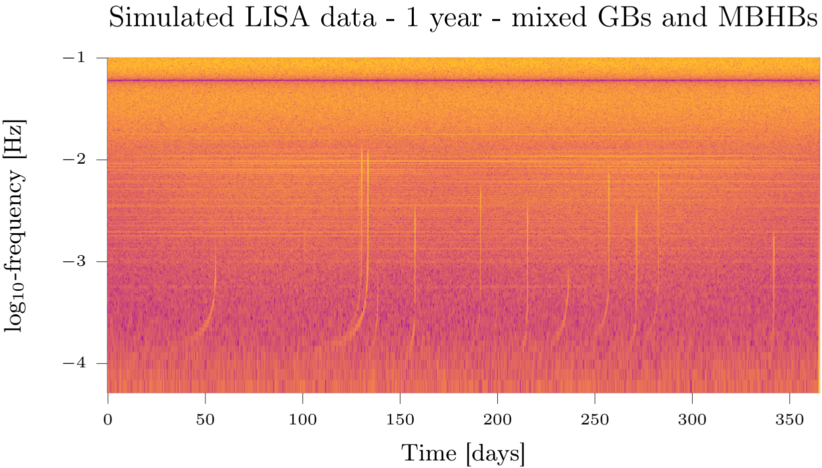
Galactic Binaries (GB)
- Stationary, quasi-monochromatic
- Tens of thousands detectable by LISA



Other sources
- EMRIs
- Confusion background
- Stochastic background

Noises & artifacts
- Instrumental
- Glitches
- Gaps

Lisa Data Challenge - LDC2a



Signal unmixing : getting back to separate sources

Unmixing problem : exploiting an adapted representation

State of the art

- Parametric methods : MCMC
  - \+ Physical relevance, parameter space exploration, uncertainty quantification
  - – Slow, require efficient signal generative model, sensitive to initialization
- Template matching
  - \+ Fast, good looking extracted signal
  - – Need for big template basis, bias
- Non-parametric methods : wavelet transform, PCA
  - \+ Fast, don't rely on generative model
  - – Linear models w.r.t. input signal

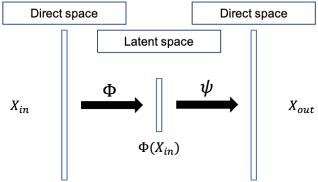
Our approach : Learn low-dimension non-linear representation

Low dimension representations

Work well for :

- high dimension signal described by few parameters
- tackling multiple problems (e.g. detection, extraction, ...)
- Galactic Binaries signal analysis [1]

AutoEncoders

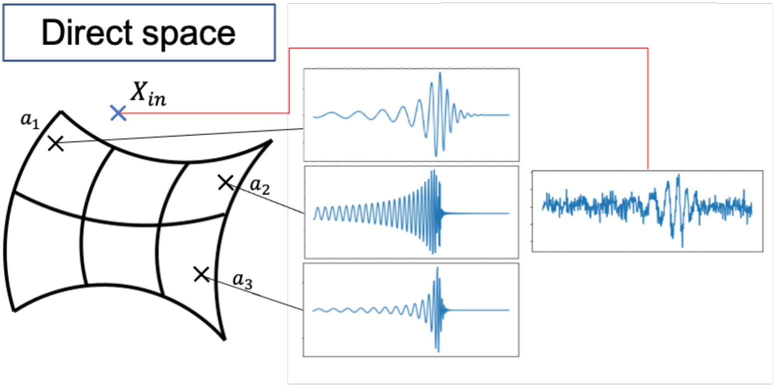


Unsupervised learning

$$\underset{\Phi,\Psi}{\text{minimize}} \left(||X_{in}^{\mathcal{T}} - X_{out}^{\mathcal{T}}||_2^2 \right)$$

1. Blelly, A., Moutarde, H., & Bobin, J. (2020). Sparsity-based recovery of Galactic-binary gravitational waves.

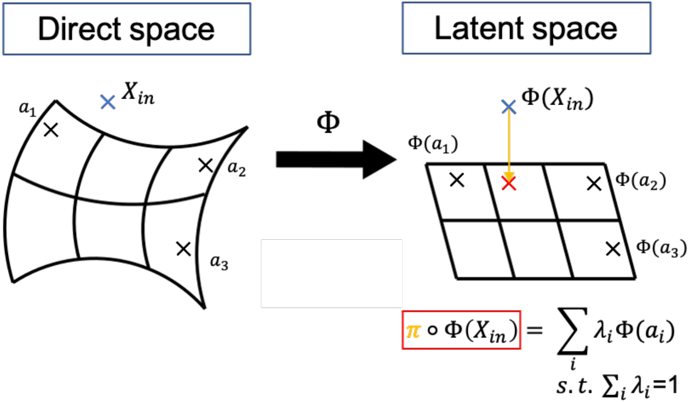
Interpolatory AutoEncoder : Direct space & manifold



Direct space : $\mathbb{R}^N$

Anchor points : $(a_i)_{1 \leq i \leq m}$ with $m \ll N$

Interpolatory AutoEncoder : Latent space & interpolation



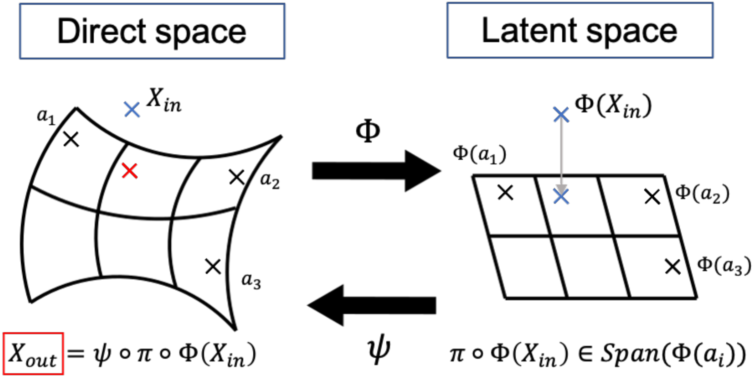
Fast interpolation

$\text{Argmin}_{(\lambda_i)} ||\Phi(X_{in}) - \sum_i \lambda_i \Phi(a_i)||_2^2$

Barycentric span projection

$\text{Argmin}_{(\lambda_i)} ||X_{in} - \Psi(\sum_i \lambda_i \Phi(a_i))||_2^2$

Interpolatory AutoEncoder [2] : Learning & output



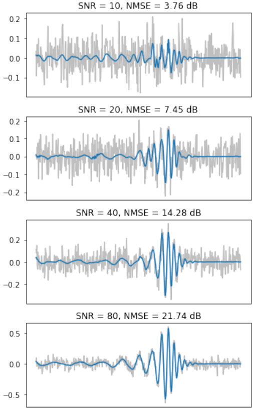
Unsupervised learning

$$\text{minimize}_{\Phi,\Psi} \left(||X_{in}^{\mathcal{T}} - X_{out}^{\mathcal{T}}||_2^2 + \mu||\Phi(X_{in}^{\mathcal{T}}) - \pi_{FI}(\Phi(X_{in}^{\mathcal{T}}))||_2^2 \right)$$

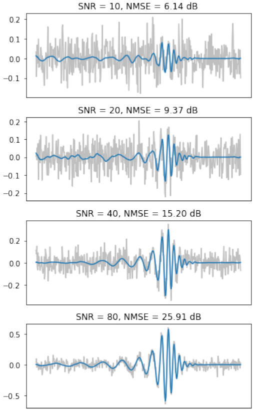
2. Bobin, J., Gertosio, R., Bobin, C., & Thiam, C. (2021). Non-linear interpolation learning for example-based inverse problem regularization. github.com/jbobin/IAE

Results : Signal extraction

Fast Interpolation

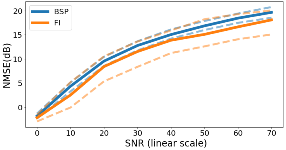


Barycentric Span Projection



Reconstruction quality criterion

$$-20\log_{10}\left(\frac{||X_{out}-X^*||_2}{||X^*||_2}\right)$$

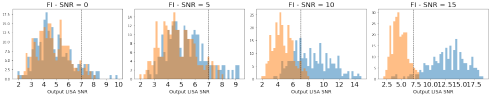


Results : Signal detection

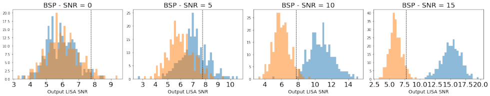
Hypothesis testing

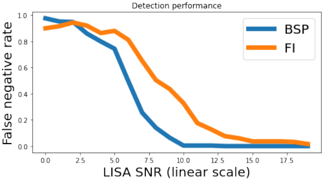
- Generate MBHB+noise and noise-only signals
- Attempt to extract MBHB and compute a metric on X_{out}
- Thresholding based on fixed acceptable false positive rate

Fast Interpolation



Barycentric Span Projection

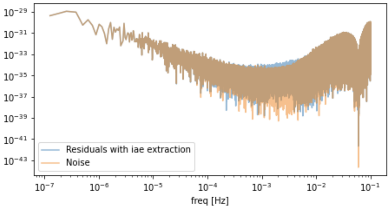




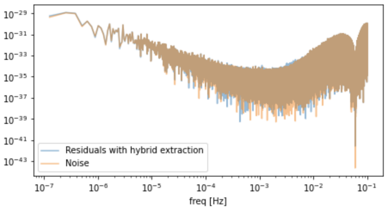
Work in progress : adaptive STFT

- Remark : loud MBHB can leak into residual during inspiral
- Not possible to extend IAE to arbitrary lengths
- Developed a hybrid method combining
  - IAE to capture coalescence
  - An adaptive Time-Frequency decomposition to adapt window size to instantaneous frequency and its derivative $\dot{f}$

Residuals IAE alone



Residuals hybrid method



Take home message and perspectives

Signal analysis is about finding adapted representations to make its features stand out from its environment.

MBHB analysis :

- Tested a model of convolutive Interpolatory AutoEncoder
- Currently working on benchmarks to compare state of the art methods with hybrid method
- Collaboration with L2IT to investigate MBHB parameter estimation

Thank you !