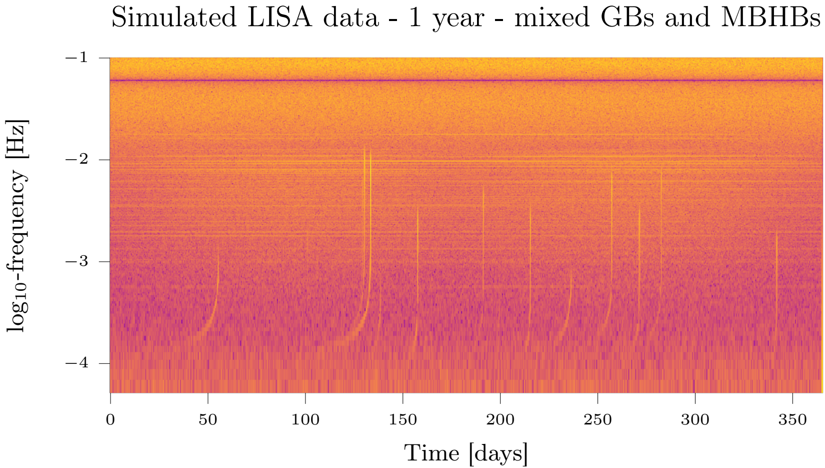
Learning-based models for gravitational wave analysis

Elie Leroy
with J. Bobin and H. Moutarde

CEA Saclay - IRFU

24 November 2022

Lisa Data Challenge - LDC2a



Unmixing problem : 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
  - − Costly, require efficient signal generative model, sensitive to initialization
- Match filtering
  - \+ Efficient, smooth extracted signal
  - − Need for big template basis, bias
- Dimension reduction : wavelet transform, PCA
  - \+ Fast, don't rely on generative model
  - − Linear models w.r.t. input signal

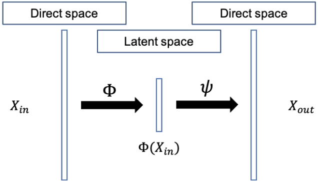
Our approach : Learn low-dimension non-linear representation

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

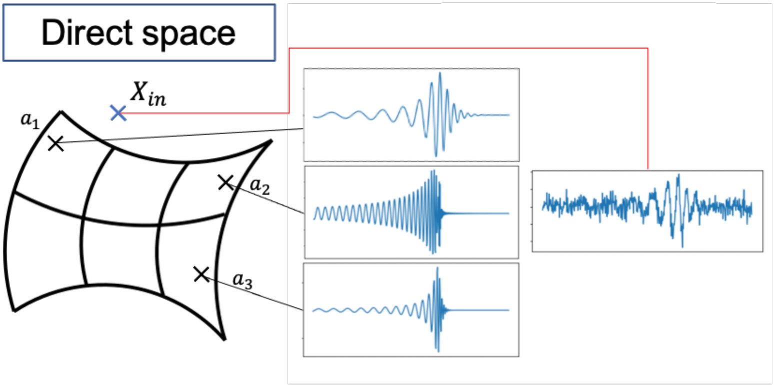


Unsupervised learning

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

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

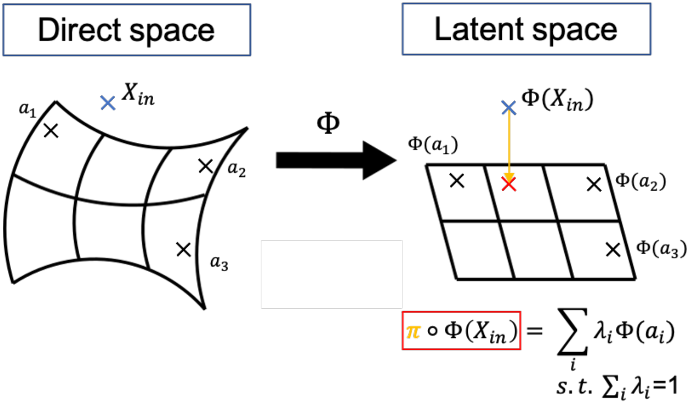
Interpolatory AutoEncoder : Direct space & manifold



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

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

Interpolatory AutoEncoder : Latent space & interpolation



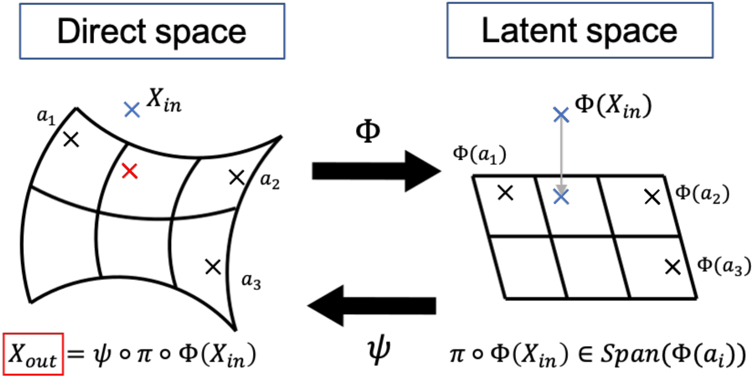
Fast interpolation

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

Barycentric span projection

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

Interpolatory AutoEncoder [2] : Learning & output



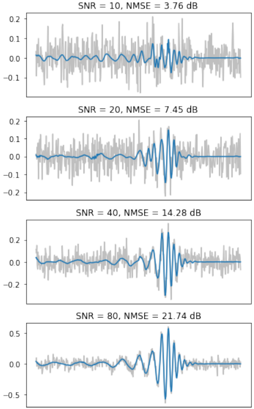
Unsupervised learning

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

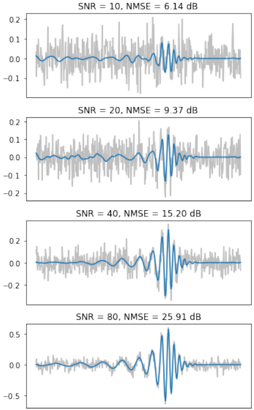
2. Bobin, J., Gertosio, R., Bobin, C., & Thiam, C. (2021). Non-linear interpolation learning for example-based inverse problem regularization. github.com/jbobin/IAE

Results : Signal extraction

Fast Interpolation

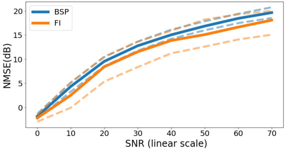


Barycentric Span Projection



Reconstruction quality criterion

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

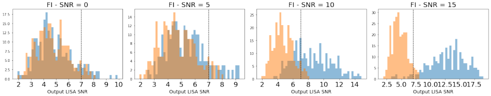


Results : Signal detection

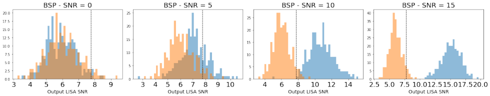
Hypothesis testing

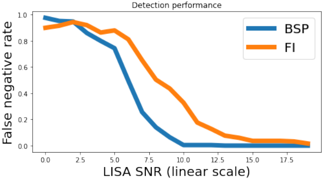
- Generate MBHB+noise and noise-only signals
- Run both through IAE and compute a metric on X_{out}
- For a fixed false positive rate measure false negative rate

Fast Interpolation



Barycentric Span Projection

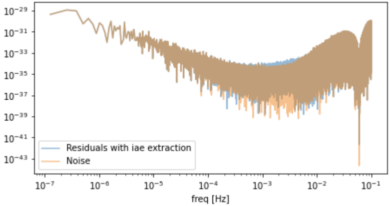




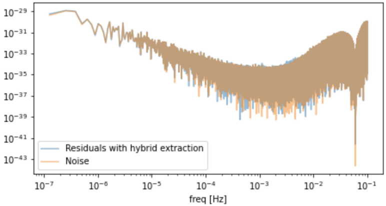
Adaptative STFT to capture inspiral

- 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 outlooks

Non parametric methods are efficient tools for problems ahead of parametric sampling such as detection and signal extraction

MBHB analysis :

- Tested a model of convolutive Interpolatory AutoEncoder
- Now finalizing benchmarking and comparisons
- Article in writing on this representation method for MBHB

Thank you !