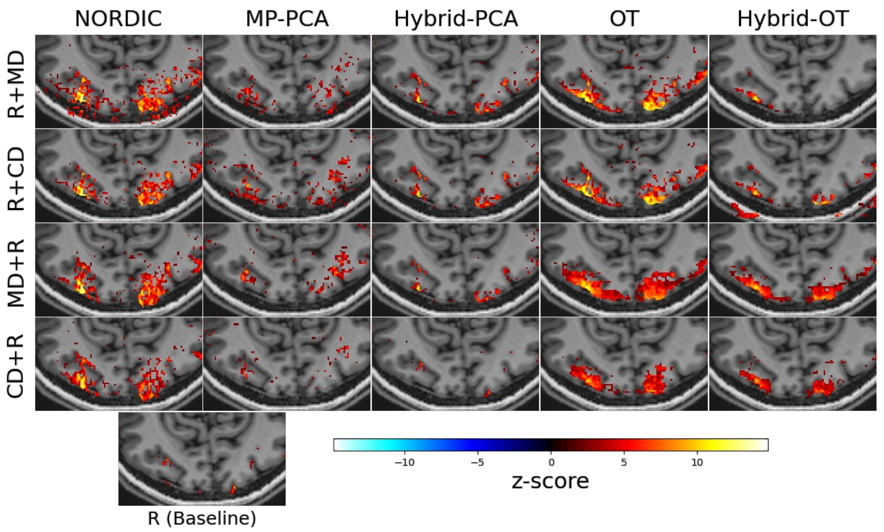
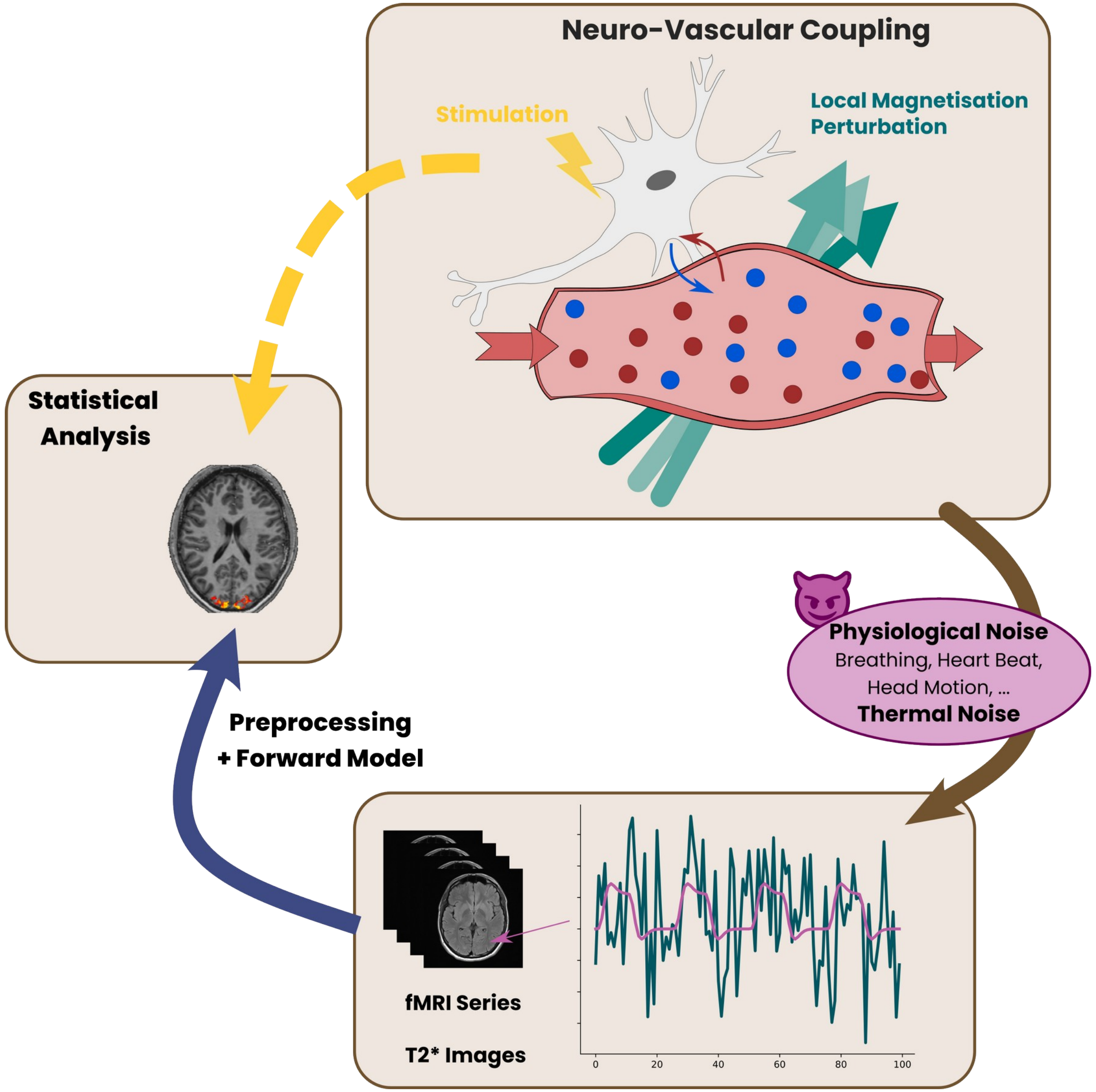
Denoising of fMRI Volumes using Local Low Rank Methods

Pierre-Antoine Comby
Supervisors: Philippe Ciuciu & Alexandre Vignaud

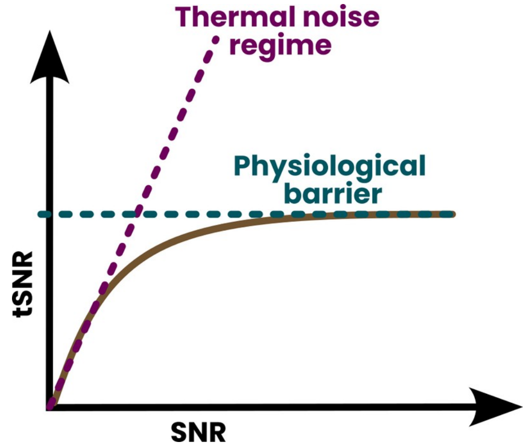


fMRI : Physics & Challenges



- Best Image Quality (SNR) does not imply best statistical analysis

$$\mathrm{SNR} \propto \frac{B_0^{1.94}(\Delta x)^3}{g\sqrt{R}}$$



- Two Main Source of Noise
  - Thermal noise (Complex Gaussian)
  - Physiological Noise (Breathing, heart beat, motion ...)
- Physiological Noise can be dealt with using GLM regressors and external acquisition

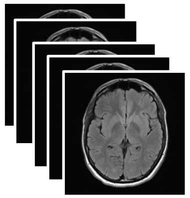
At high resolution (= low Spatial SNR), Noise in fMRI data is mostly thermal

From global to local low rank

- Let the sequence of fMRI Images:

$$\mathbf{Y} = \mathbf{X} + \mathbf{N} \in \mathbb{C}^{N_x N_y N_z \times N_t} \quad N_{ij} \sim \mathcal{N}(0, \boldsymbol{\Sigma}^2)$$

- Denoising using a low-rank *a priori*
  - fMRI Image share a lot of common information

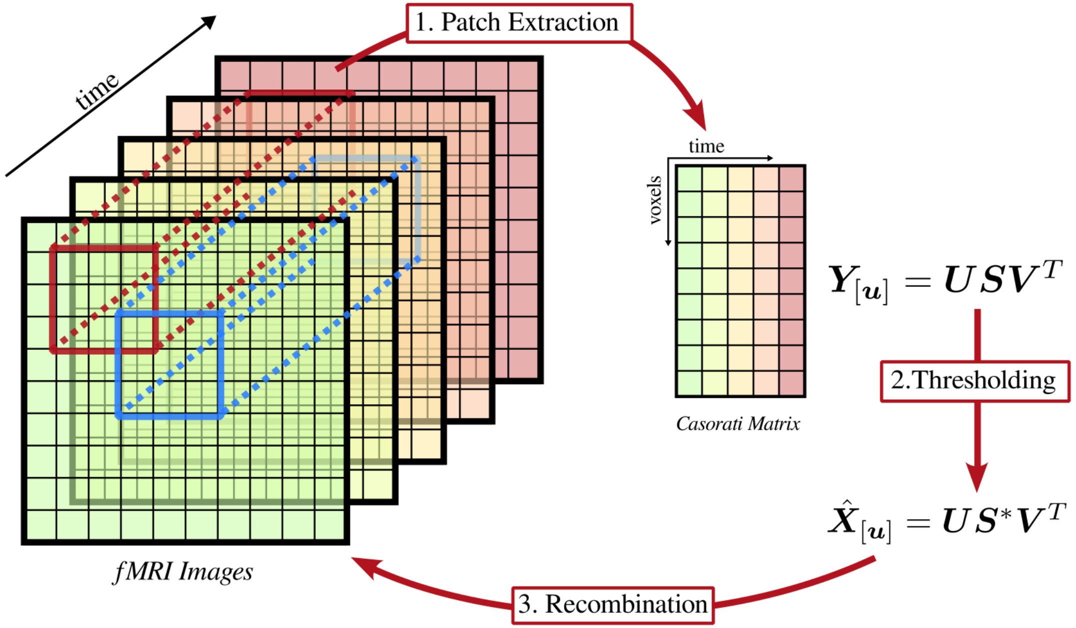


$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|_*$$

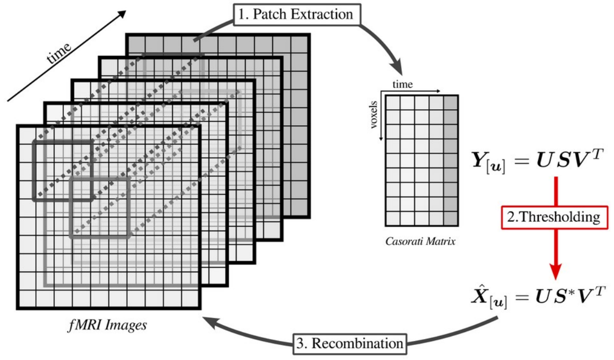
- $N_x N_y N_z \sim 10^5 \; N_t$: **DoF limited for the rank constraint**
- Noise Level is **spatially heterogeneous**

→ Use a *Local* formalism

$$\mathbf{Y}_{[\mathbf{u}]} = \mathbf{X}_{[\mathbf{u}]} + \mathcal{N}(0, \sigma_{[\mathbf{u}]})$$



Singular Value Thresholding



- Solution of

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|_*$$

- Given by the **singular value hard/soft thresholding**

$$\mathbf{Y} = \mathbf{USV}^T = \sum_{i=1}^{N} s_i \mathbf{u}_i \mathbf{v}_i^T \longrightarrow \hat{\mathbf{X}} = \sum_{i=1}^{N} \eta(s_i)\mathbf{u}_i \mathbf{v}_i^T$$

- Main Challenges
  - Choice of the threshold (noise level dependent)
    - Estimation of Noise Variance
  - Efficient Computation (lots of patches)
    - Complexity: *O(N_patches x SVD)*
    - *Solution: Use a Mask to reduced the number of relevant patches*

→

Highlight two methods

- NORDIC
- Optimal Threshold

NORDIC : an empirical approach

g-factor: Estimated noise variance

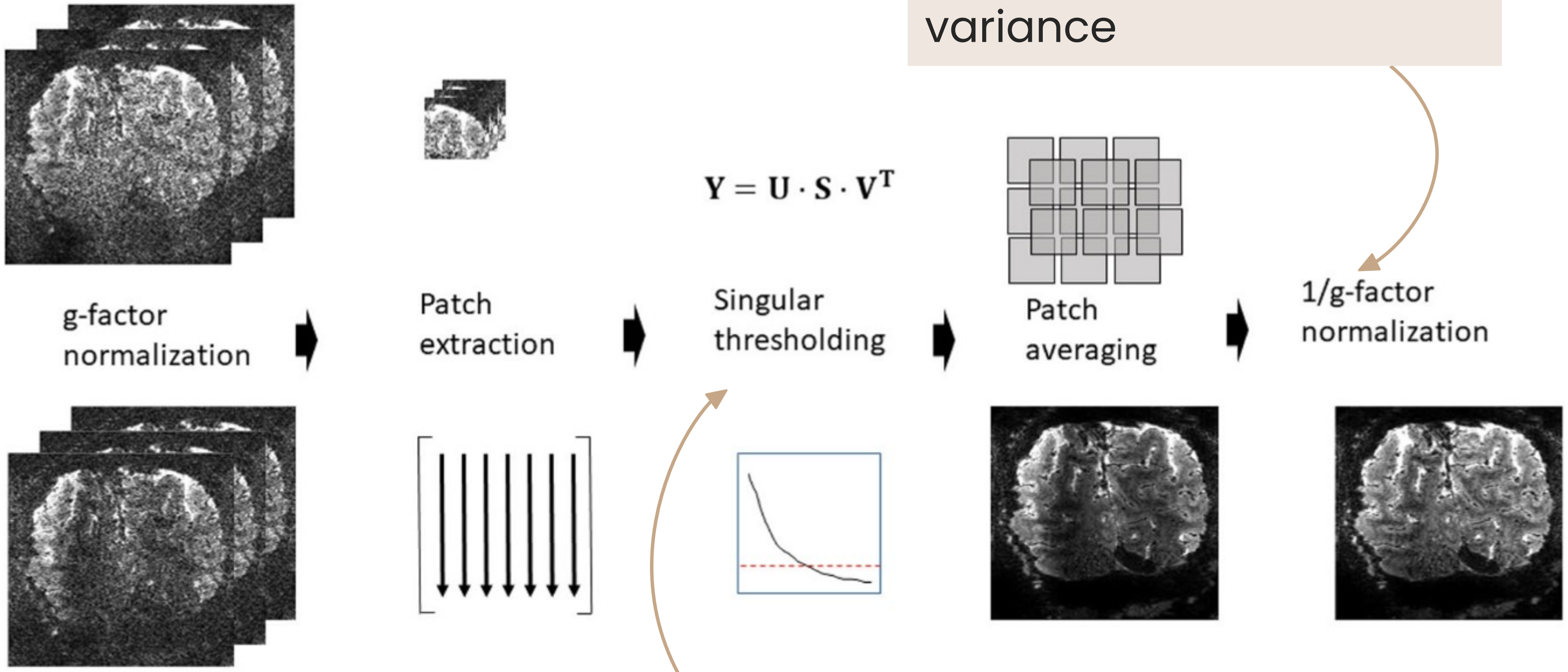
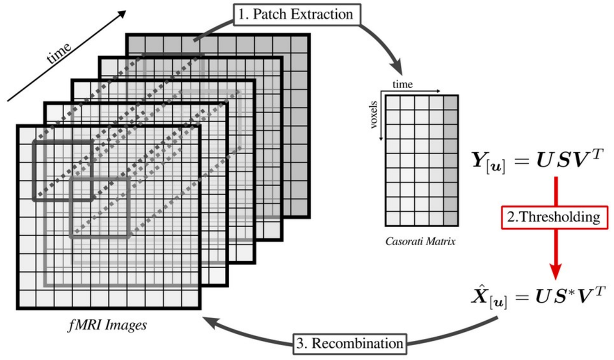


Fig. 7 of [Viozili et al. 2021]
NORDIC Workflow
(Annotations added).

Threshold: largest singular value of a *simulated* noise-only patch



Optimal Threshold : A Mathematical approach

Mathematical Formulation

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}/\sqrt{N}$$

Data
$\mathbf{Y} \in \mathbb{C}^{M\times N}$
$\beta = M/N$

Low Rank Matrix *Unknown*

iid Gaussian Matrix

Find the best low rank matrix:

$$\arg\min_{\hat{\mathbf{X}}} \|\mathbf{X} - \hat{\mathbf{X}}\|_2^2$$

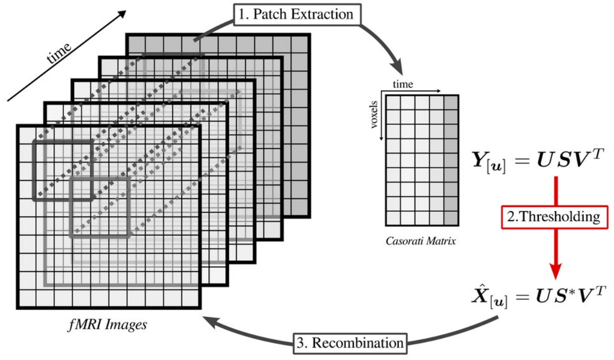
Classical Solution:

Truncated SVD

$$\mathbf{Y} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \sum_{i=1}^{N} s_i \mathbf{u}_i {\mathbf{v}_i}^T$$

Solution:
LR approx.

$$\hat{\mathbf{X}} = \sum_{i=1}^{N} \eta(s_i) \mathbf{u}_i {\mathbf{v}_i}^T$$



Thresholding Functions

Hard $\eta(s) = s1_{s>\theta}$

Soft $\eta(s) = \max(0, s - \theta)$

Optimal θ
[Gavish et al. 2014]

Optimal
[Gavish et al. 2017]

$$\eta^*(s) = \begin{cases} \frac{1}{s}\sqrt{(s^2 - \beta - 1)^2 - 4\beta} & s \geq 1 + \sqrt{\beta} \\ 0 & s \leq 1 + \sqrt{\beta} \end{cases}$$

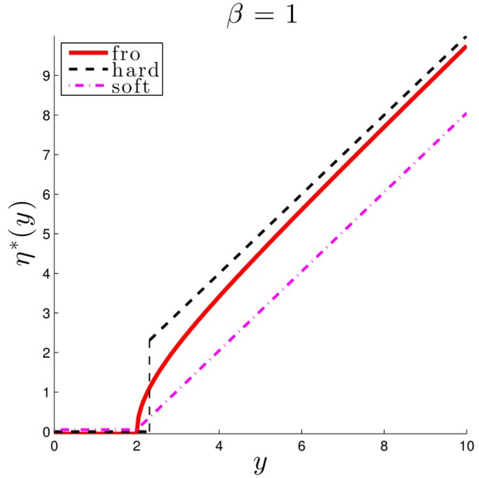
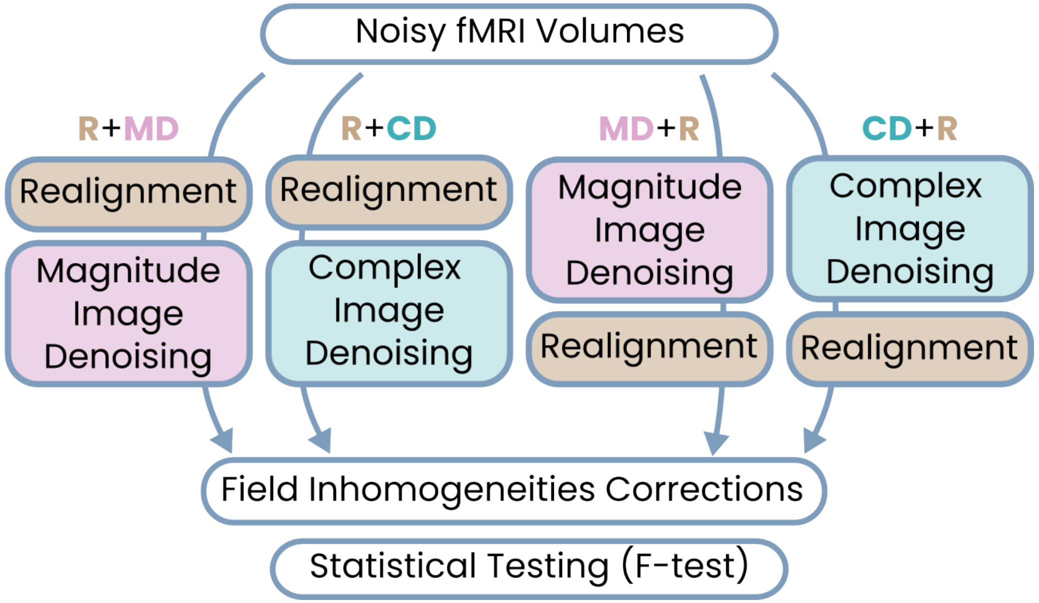


Fig 2 of [Gavish and Donoho, 2014]
Hard, Soft and Optimal Thresholding function

Materials and Methods

- Retinotopy Task
  - 1mmiso TR=2.4s
  - 3D-EPI Sequence
  - ***6 Volunteers***
  - *External noise map available*
- F-test on global effect

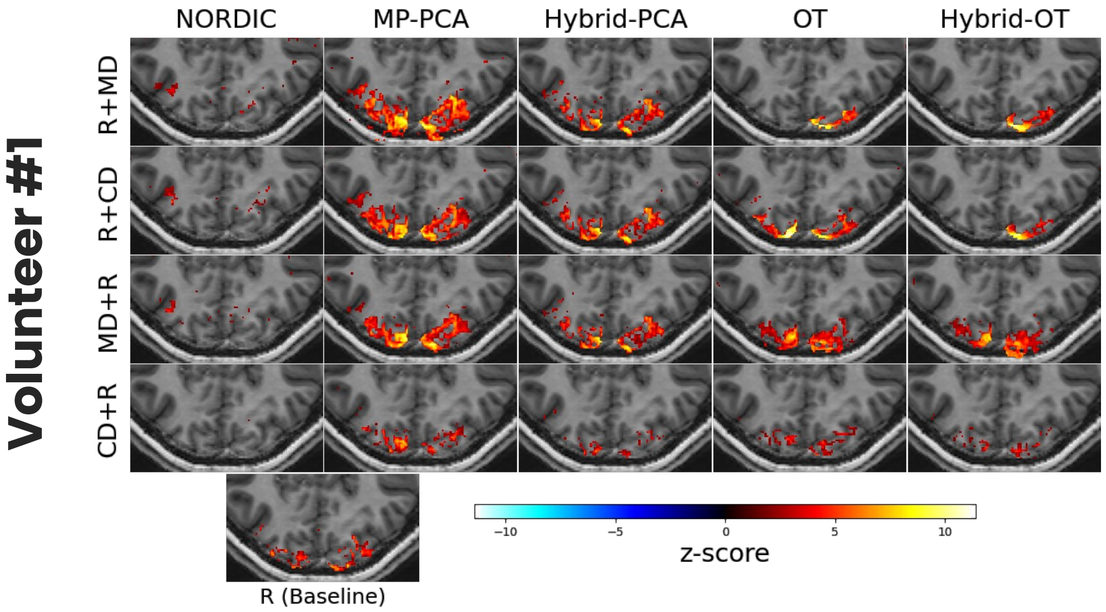
4 Processing Scenarios

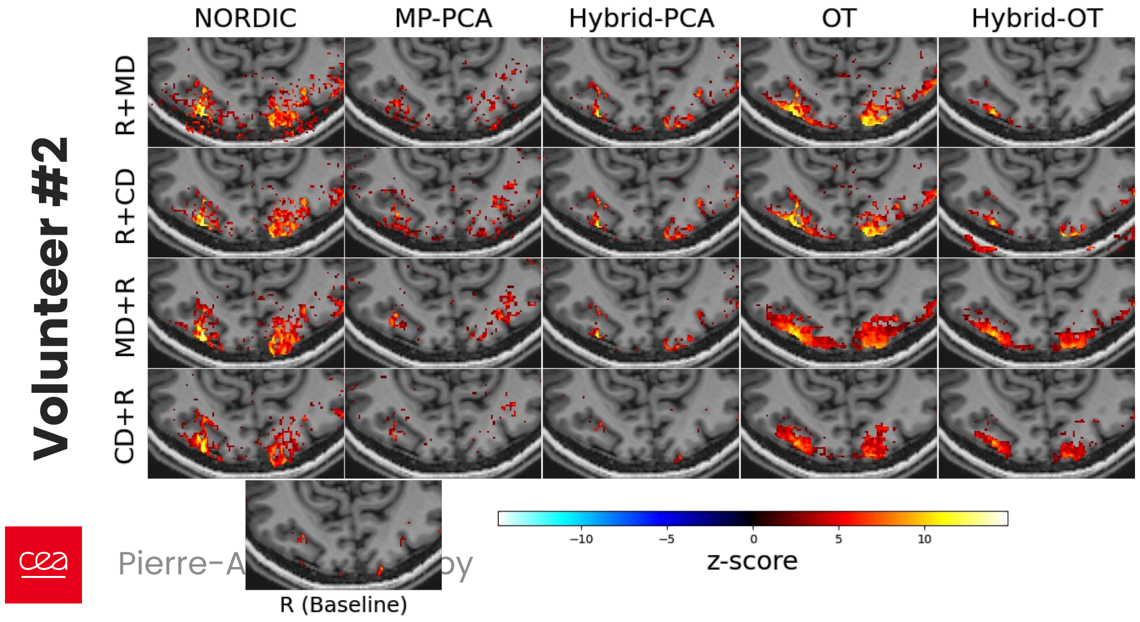


5 Methods Benchmark

Name	Threshold Function	Reference
NORDIC	$\max(0, \lambda - \theta_{global})$	*[Vizioli et al. 2021]*
MP-PCA	$\max(0, \lambda - \theta_{[u]})$	*[Veraart et al. 2016]*
HYBRID-PCA	$\max(0, \lambda - \theta_{[u]})$	*[Henriques et al. 2022]*
OPTIMAL THRESH	$\frac{N_t\sigma^2}{\lambda}\sqrt{\left(\frac{\lambda^2}{N_t\sigma^2} - \beta - 1\right)^2 - 4\beta}\,\mathbf{1}_{\frac{\lambda}{\sqrt{N_t}\sigma}\geq 1+\sqrt{\beta}}$	*[Gavish et al. 2017]*
HYBRID-OT		$\beta = \frac{N_v}{N_t}$

Results





Average gain of activation detection (N=6)

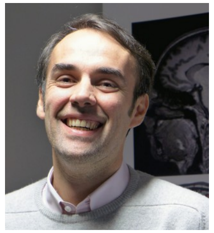
Denoiser	NORDIC	MP-PCA	Hybrid-PCA	OT	Hybrid-OT
R+MD	×3.52 ×3.64	×6.02 ×4.92	×**6.09** ×**4.93**	×3.73 ×4.33	×0.91 ×1.00
R+CD	×0.57 ×0.55	×2.98 ×2.70	×3.27 ×2.91	×**8.03** ×**5.04**	×5.29 ×4.01
MD+R	×3.36 ×3.48	×6.32 ×4.77	×**7.55** ×**6.10**	×3.10 ×3.19	×1.22 ×4.90
CD+R	×2.57 ×2.59	×5.53 ×4.39	×4.97 ×4.27	×**7.91** ×**6.26**	×5.45 ×5.00

- Inter volunteer variability
  - Motion, response to stimuli, …
- R+CD with OT provides best results
- Gains even for magnitude-only image

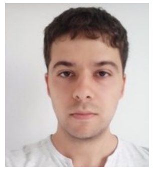
Conclusion & Future work

- High Resolution fMRI benefits from de-noising: **8x increase in voxel activations**
- Complex-valued data is not mandatory to get improved results.
- Local low-rank methods are now easily accessible for fMRI and other contrasts
  - Early results for CEST Imaging at Neurospin
- Future Work
  - Add other methods to the package (NLM, SURE thresholding)
  - Faster Computation with GPU processing
  - Application to other Paradigms and Constrast
    - First results in CEST imaging
    - Resting State data









Thank you!

Any Questions?

ISBI Paper

https://hal.science/hal-03895194

Code

https://github.com/paquiteau/patch-denoising

```
$ pip install patch-denoise
$ patch-denoise input.nii denoised.nii
```

Pierre-Antoine Comby
pierre-antoine.comby@cea.fr
https://github.com/paquiteau

Alexandre Vignaud
alexandre.vignaud@cea.fr

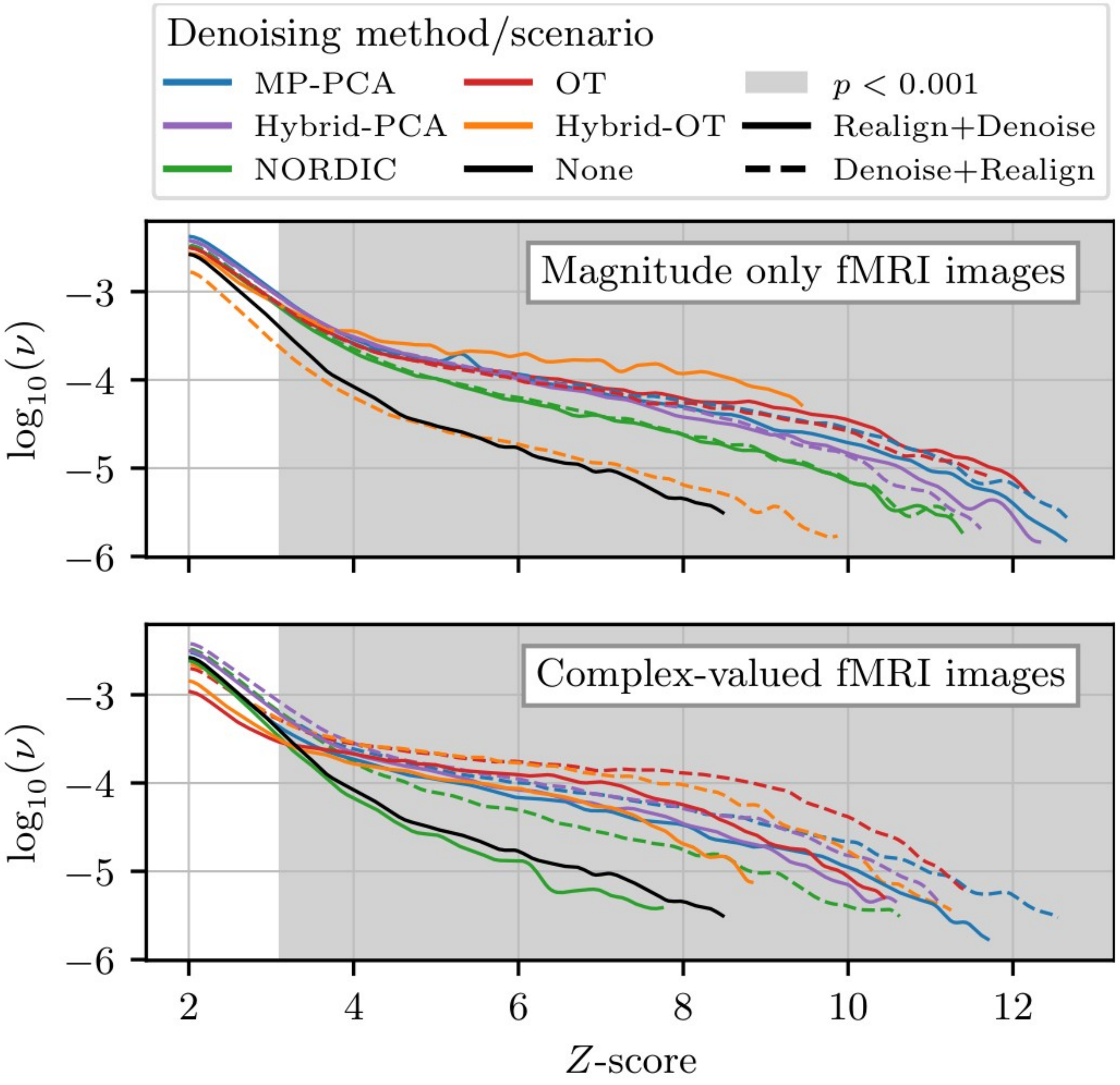
Philippe Ciuciu
philippe.ciuciu@cea.fr
https://philippeciuciu.fr

A Boost on Z-Score

- Compare to baseline (no de-noising), all denoising methods shift the z-score distribution $\nu(Z)$

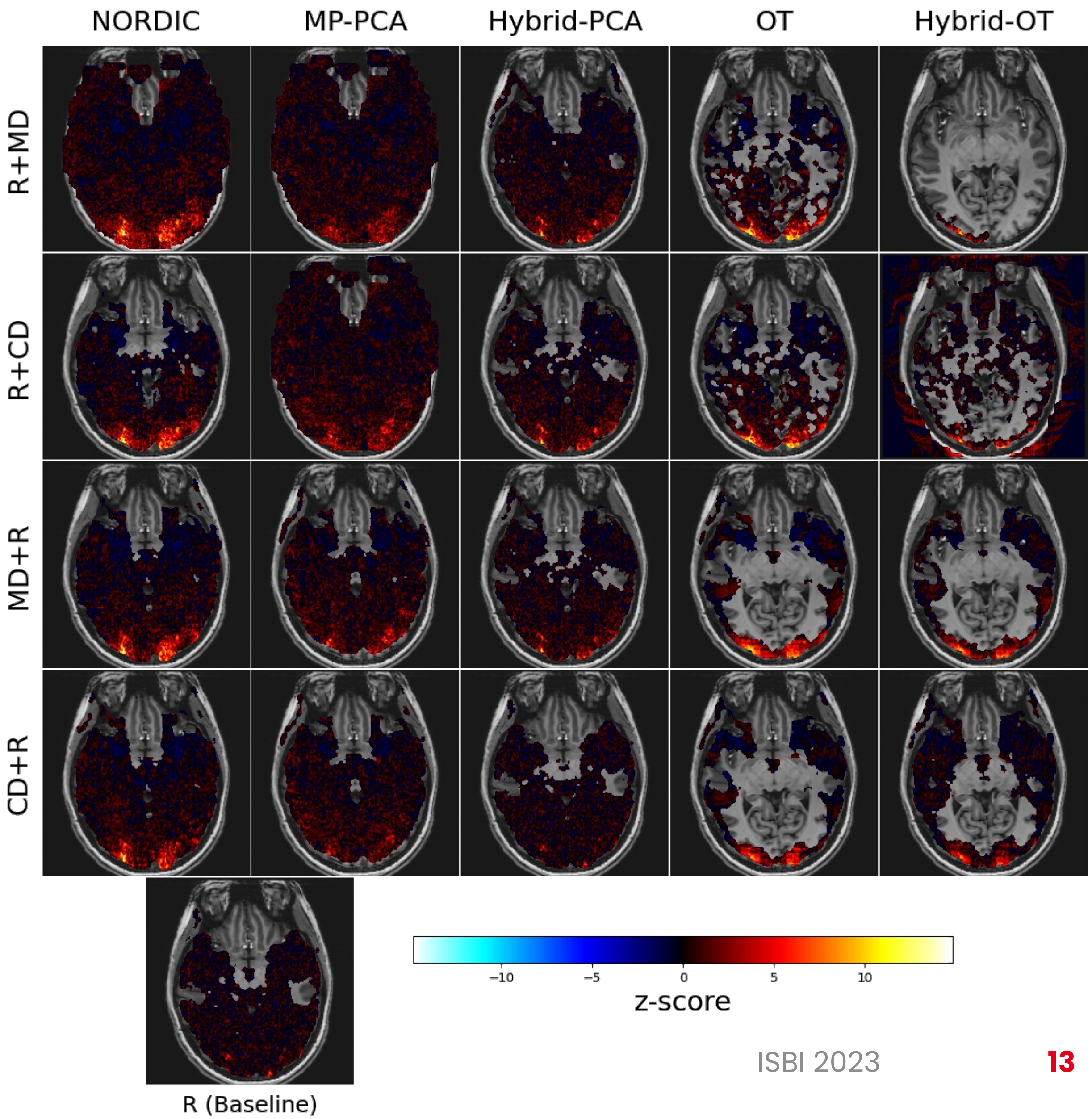
- Complex Image de-noising perform best when denoising is done after motion correction

  - Motion Correction is a trade-off between reinterpolation of the data (changing noise statistics) and lower rank signals



Z-Score Maps

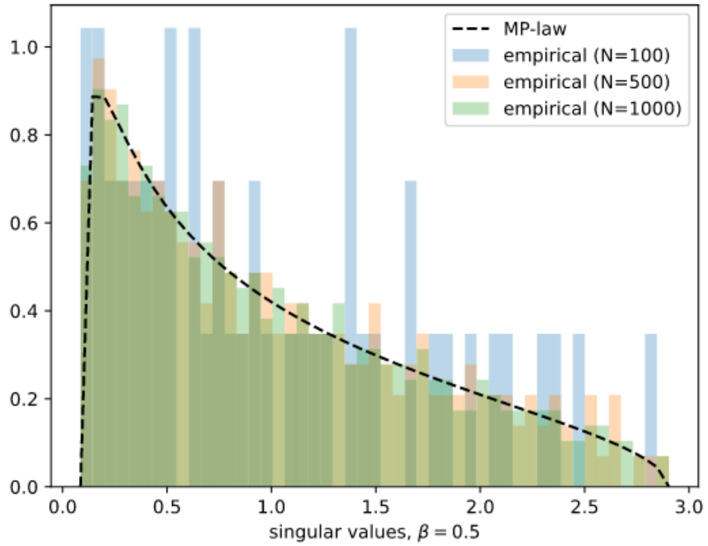
- Remove variability in White Matter
  - Best method for this is OT
- Preserve Contrast in Gray Matter
- Overall boost of z-scores



Marchenko–Pastur Law

- Hypothesis

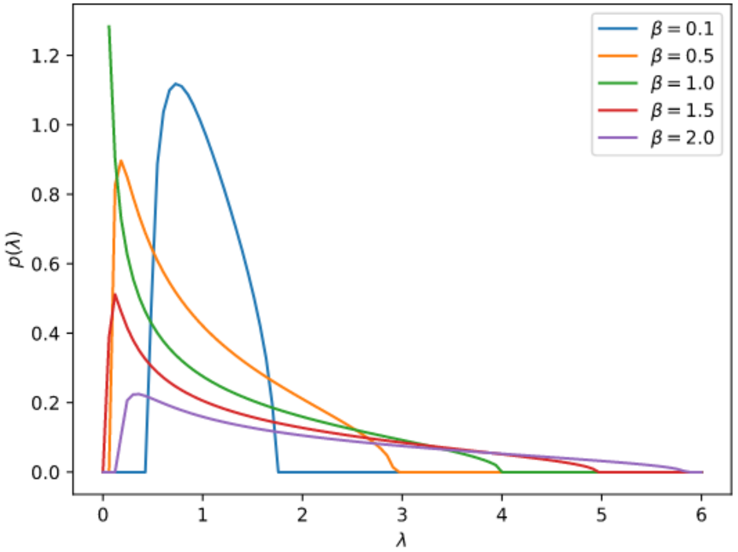
$$\mathbf{X}_\mathrm{N} \in \mathbb{C}^{\beta \mathrm{N} \times \mathrm{N}} \hookrightarrow \mathcal{N}(0, \sigma^2)$$



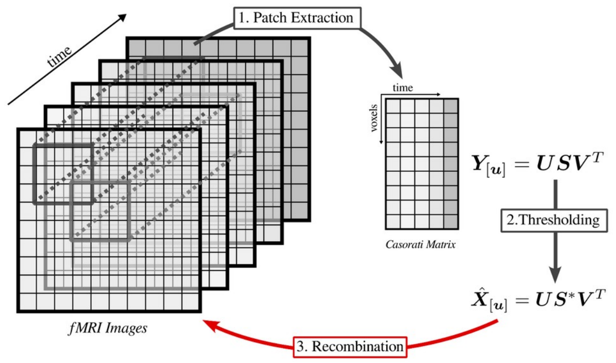
- Results $\mathrm{N} \to \infty$
  - The eigenvalues of $\mathbf{Y}_\mathrm{N} = \frac{1}{\mathrm{N}}\mathbf{X}\mathbf{X}^\top$ follows the distribution:

$$\mathrm{p}(\lambda|\sigma, \beta) = \frac{\sqrt{(\beta_+ - \lambda)(\lambda - \beta_-)}}{2\pi\beta\lambda\sigma^2}\mathbf{1}_{\lambda\in[\beta_-,\beta_+]}$$

  - Support $\beta_\pm = \sigma^2(1 \pm \sqrt{\beta})^2$
  - Expectation $\mathbb{E}[\lambda] = \sigma^2$



Recombination Strategies



1) Average of Patches

2) Weighted Average of patches *[Majon et al. 2013]*

$$\hat{\mathbf{X}}_i = \frac{\sum_{j=1}^{P} w_j \mathbf{X}_{[\mathbf{u}_j]}}{\sum_{j=1}^{P} w_j} \qquad w_j = \frac{1}{1 + \|S^*_{[\mathbf{u}_j]}\|_0}$$

- Lower rank patches are promoted

3) Select Center value of patches

- Requires a maximum overlap