Efficient whole brain estimation of the haemodynamic response function for TV-regularized semi-blind deconvolution of neural activity in fMRI

Hamza Cherkaoui[a,b]

[a] Parietal - NeuroSpin - CEA, [b] Neuroimagerie pharmacologique - BioMaps - CEA

Paris-Sud University - Doctoral school n°575 - EOBE

Supervisors:

Philippe Ciuciu[1], Claire Leroy[2] and Thomas Moreau[3]

[1] NeuroSpin CEA,
[2] Service Hospitalier Frédéric Joliot CEA,
[3] Centre Saclay Inria

Jury:

Dr. Julien Mairal
Pr. Dimitri Van de Ville
Pr. Charles Soussen
Dr. Myriam Edjlali-Goujon
Dr. Florence Forbes

Outline

The need to better understand drugs' action

- The buprenorphine is an opioid – class of analgesic molecule.
- Such medication induces a patient-dependent analgesic effect

Synchropioid project: randomized, double-blinded, study design

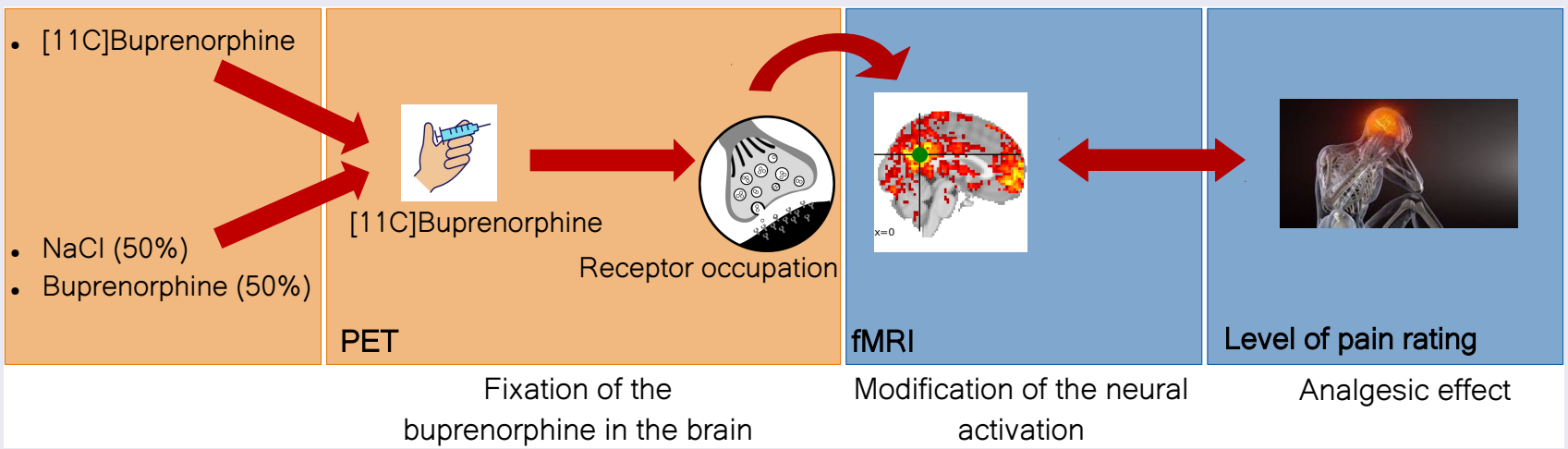


Figure: Illustration of the different steps studied in the synchropioid protocol.

fMRI signal description

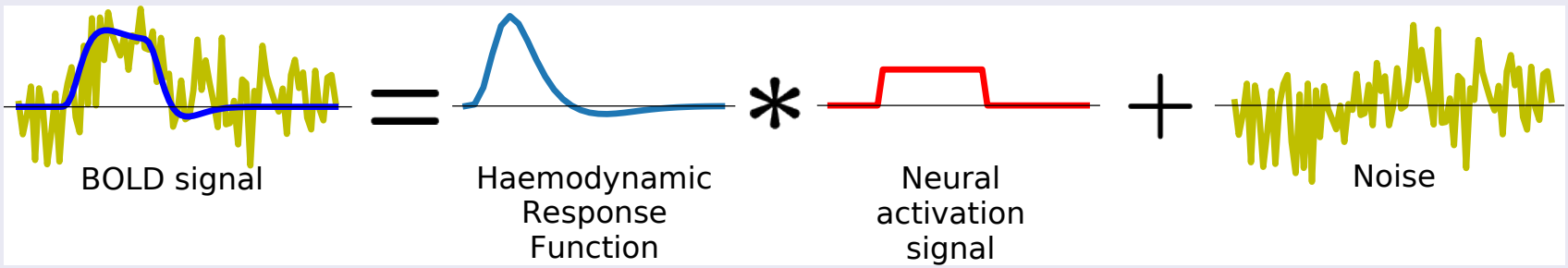


Figure: Illustration of BOLD signal modeling (Ogawa et al., 1998).

The interest of estimating the HRF

- The HRF models a complex cascade of events produced notably by the glial cells and the vascular system.
- Its shape (the haemodynamic delay) will be affected by drugs.
- Its estimation for the whole brain will characterize the effect of drugs in the brain.

fMRI signal description

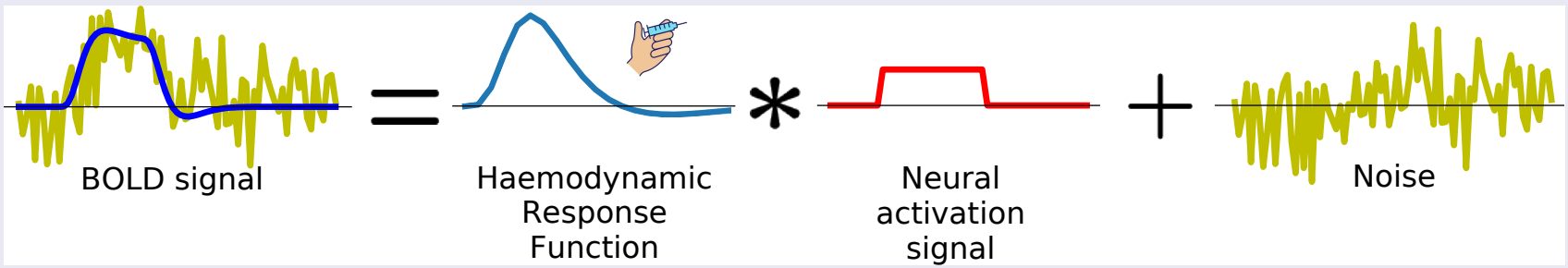


Figure: Illustration of BOLD signal modeling (Ogawa et al., 1998).

The interest of estimating the HRF

- The HRF models a complex cascade of events produced notably by the glial cells and the vascular system.
- Its shape (the haemodynamic delay) will be affected by drugs.
- Its estimation for the whole brain will characterize the effect of drugs in the brain.

BOLD signal modeling

Main features

- Disentangle the neurovascular coupling from neural activity
- Paradigm-free approach
- Model the whole brain (multivariate model)

BOLD signal modeling

Main features

- Disentangle the neurovascular coupling from neural activity
- Paradigm-free approach
- Model the whole brain (multivariate model)

Illustration of the BOLD signal decomposition

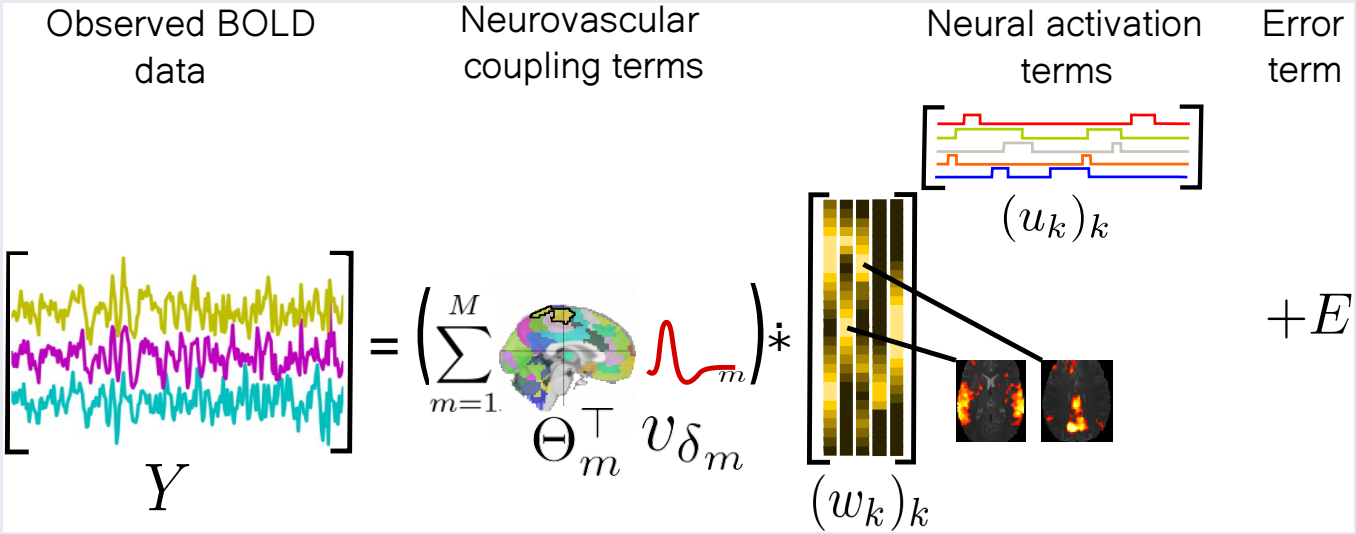


Figure: Illustration of the low-rank multivariate BOLD signal model (Cherkaoui et al., under review, NeuroImage, 2021).

HRF modeling

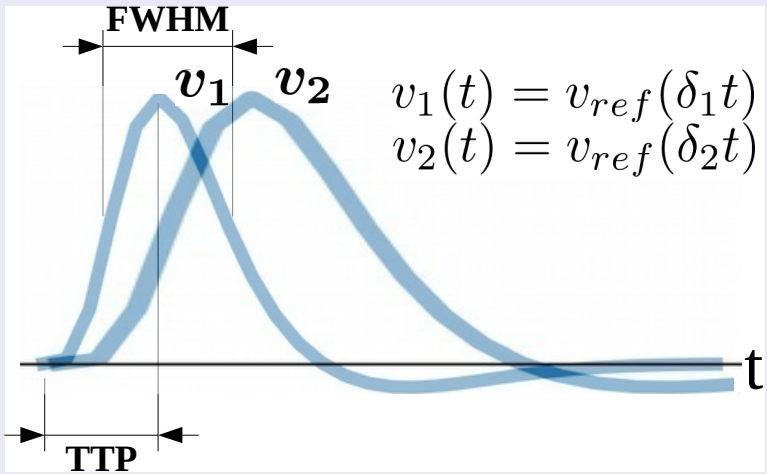


Figure: Illustration the HRF modeling.

- We dilate the time axis of a reference HRF (Friston et al., 1998)
- The Time-To-Peak (TTP) and the Full Width at Half Max (FWHM) evolve jointly

Formalization of our model

$$\boldsymbol{Y} = \left(\sum_{m=1}^{M} \boldsymbol{\Theta}_m^\top \boldsymbol{v}_{\delta_m} \right) \dot{*} \left(\sum_{k=1}^{K} \boldsymbol{w_k}^\top \boldsymbol{u_k} \right) + \boldsymbol{E} \tag{1}$$

Parameter estimation

$$\underset{(\boldsymbol{W},\boldsymbol{U},\boldsymbol{\delta})}{\arg\min}\ \frac{1}{2}\left\|\boldsymbol{Y}-\left(\sum_{m=1}^{M}\boldsymbol{\Theta}_m^{\top}\boldsymbol{v}_{\delta_m}\right)\ \dot{*}\ \left(\sum_{k=1}^{K}\boldsymbol{w}_k^{\top}\boldsymbol{u}_k\right)\right\|_F^2+\lambda\sum_{k=1}^{K}\|\nabla\boldsymbol{u}_k\|_1 \tag{2}$$

subject to $\forall k, \|\boldsymbol{w_k}\|_1 = 10, \quad \forall j, w_{kj} \geq 0, \quad \forall m, \delta_m \in [0.5, 2]$

with $\lambda = \lambda_f \lambda_{\max}$ such that $\lambda_f \in [0, 1]$

Alternated minimization strategy:

- Estimation of $\boldsymbol{u_k}$: TV regularization – Proximal Gradient Descent (PGD)
- Estimation of $\boldsymbol{w_k}$: projected gradient descent
- Estimation of $\boldsymbol{\delta_m}$: projected gradient descent

Single subject illustration (task-fMRI)

Preprocessing and acquisition parameters:

- 1 motor task-fMRI 3 min 34 s acquisition drawn at random from the Human Connectome Project (HCP) dataset (Van Essen et al., 2013).
- temporal resolution: $TR = 0.753$ s.
- classical preprocessing done with `fmriprep`.

Decomposition parameters:

- $K = 30$ (number of spatio-temporal components)
- $M = 96$ (number of regions based on the 'Harvard-Oxford' parcellation, Desikan et al., 2006)
- $\lambda_f = 0.8$ (TV regularization parameter for the temporal components)

Single subject illustration (task-fMRI)

Experimental condition: *left-hand* motor action

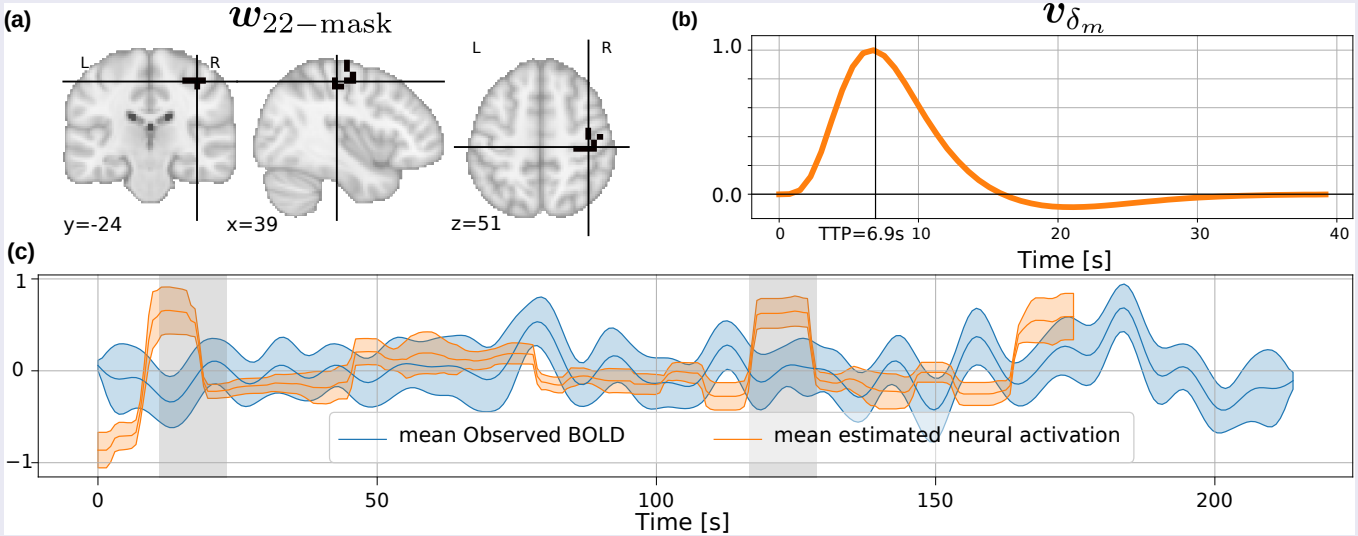


Figure: (a) *right motor cortex* spatial map – (b) Estimated HRF (for those voxels) – (c) Estimated neural signal (for the selected voxels)

Remarks

- The multivariate modeling allows to recover:
  - coherent neural activation signals (w.r.t. the experimental paradigm).
  - well known functional networks.

Single subject illustration (rs-fMRI)

Preprocessing and acquisition parameters:

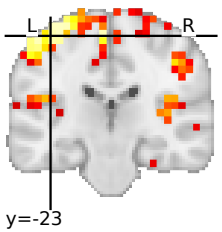
- 1 rs-fMRI 6min acquisition drawn at random from the UK Bio Bank (UKBB) dataset (Sudlow et al., 2015).
- temporal resolution: $TR = 0.735$ s.
- classical preprocessing done with `fmriprep`.

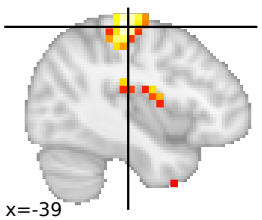
Decomposition parameters:

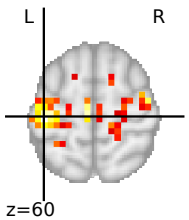
- $K = 20$
- $M = 96$ ('Harvard-Oxford' parcellation)
- $\lambda_f = 0.8$

Neural activation estimation:

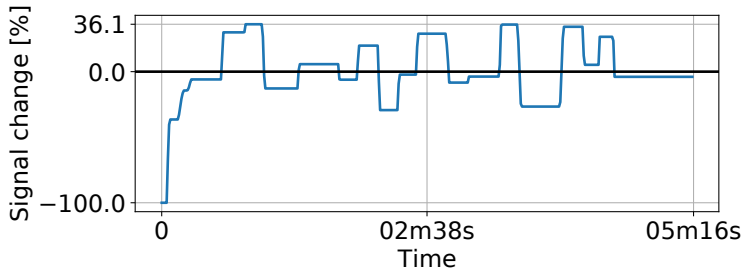
- *Left* motor cortex:







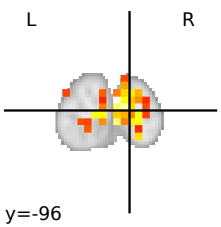
$\boldsymbol{w}_{14}$

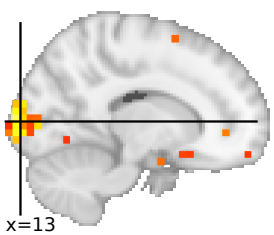


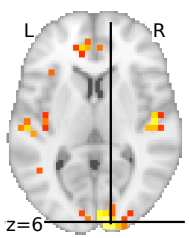
$\boldsymbol{u}_{14}$

Figure: Component n°14 ($\boldsymbol{w}_{14}, \boldsymbol{u}_{14}$) - *left* motor cortex

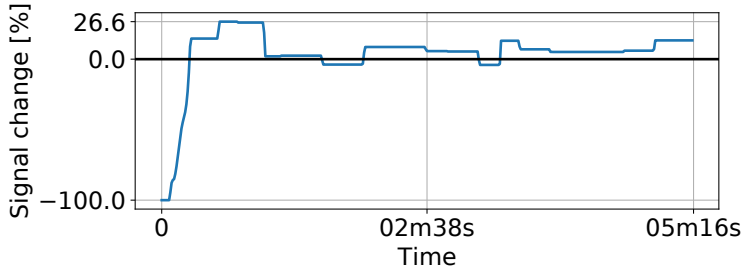
- Visual cortex:







$\boldsymbol{w}_{17}$



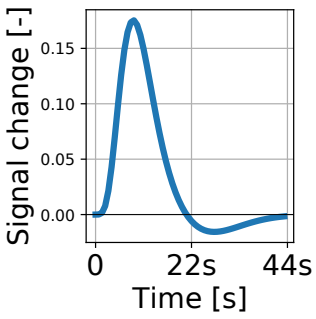
$\boldsymbol{u}_{17}$

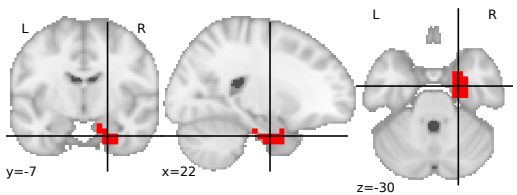
Figure: Component n°17 ($\boldsymbol{w}_{17}, \boldsymbol{u}_{17}$) - Visual cortex

Single subject illustration (rs-fMRI)

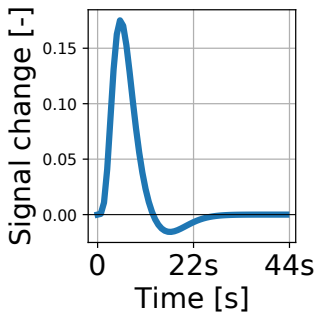
Neurovascular coupling estimation:

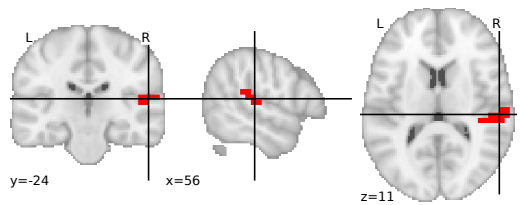
- HRF example ($\boldsymbol{v}_\delta$, $\boldsymbol{\Theta}$):





$\boldsymbol{v}_{\delta_7}$ $\boldsymbol{\Theta}_7$





$\boldsymbol{v}_{\delta_{13}}$ $\boldsymbol{\Theta}_{13}$

Figure: HRF examples

- Neurovascular map:

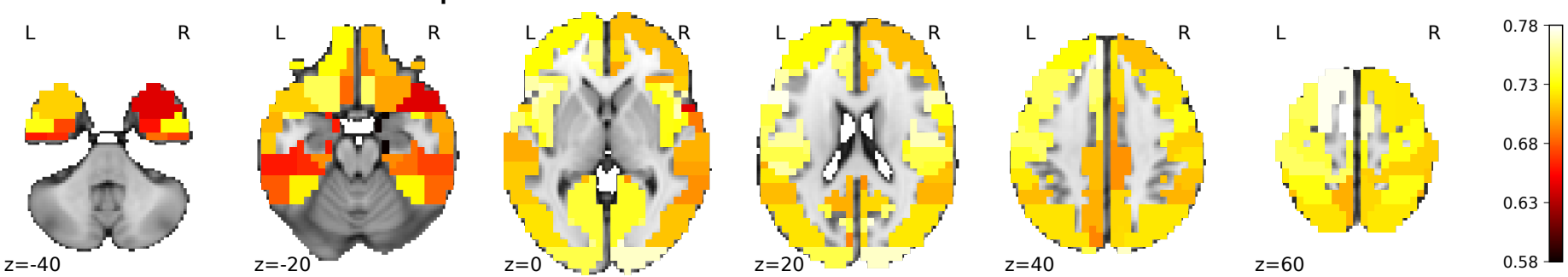


Figure: Spatial distribution of the HRF parameter $\boldsymbol{\delta}$

fMRI temporal resolution limitation of our method

Temporal resolution limitation:

- Time-To-Peak map:

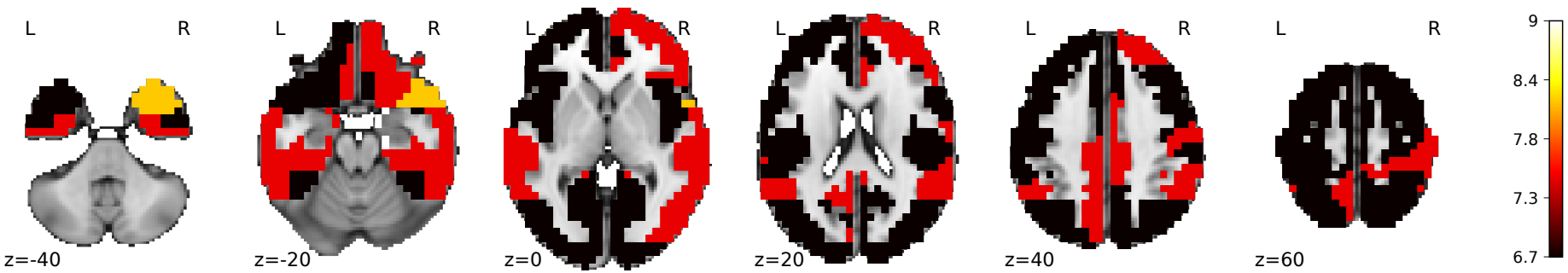


Figure: Spatial distribution of the TTPs.

- Time-To-Peak characteristic: standard deviation = 0.43 s, max-min = 2.24 s.
- The model needs a temporal resolution about one second: Repetition Time (TR) $\leq$ 1.0 s.

Single subject illustration (rs-fMRI)

Summary of the single subject decomposition:

Neural activity estimation

- The spatial maps feature symmetric sparse maps with compact activated regions.
- We recover **well known functional networks**: motor network, visual network, auditory network, Default Mode Network (DMN), Control Executive Network (CEN), etc.

Neurovascular coupling estimation

- We recover a **smooth** spatial neurovascular coupling map.
- The visual cortex features fast haemodynamic delays (Taylor et al., 2018).

Experimental validation on the UKBB dataset

Goal:

- Characterization of patients with an history of stroke.

Preprocessing and acquisition parameters:

- 48 rs-fMRI 6min acquisitions drawn at random from the UKBB dataset separated in two groups:
  - subjects who suffered from a stroke in the past
  - healthy subjects.

Decomposition parameters:

- $K = 20$
- $M = 96$ ('Harvard-Oxford' parcellation)
- $\lambda_f \in \{0.001, 0.22, 0.45, 0.67, 0.9\}$

Quantification of the asymmetry:

$$\mathrm{IHD}(\boldsymbol{\delta}^s_{\mathrm{R}}, \boldsymbol{\delta}^s_{\mathrm{L}}) = \frac{\|\boldsymbol{\delta}^s_{\mathrm{L}} - \boldsymbol{\delta}^s_{\mathrm{R}}\|_2}{\|\boldsymbol{\delta}^s_{\mathrm{L+R}}\|_2}, \quad \forall s = 1, \ldots, 24.$$

Experimental validation on the UKBB dataset

Distribution of the neurovascular asymmetry:

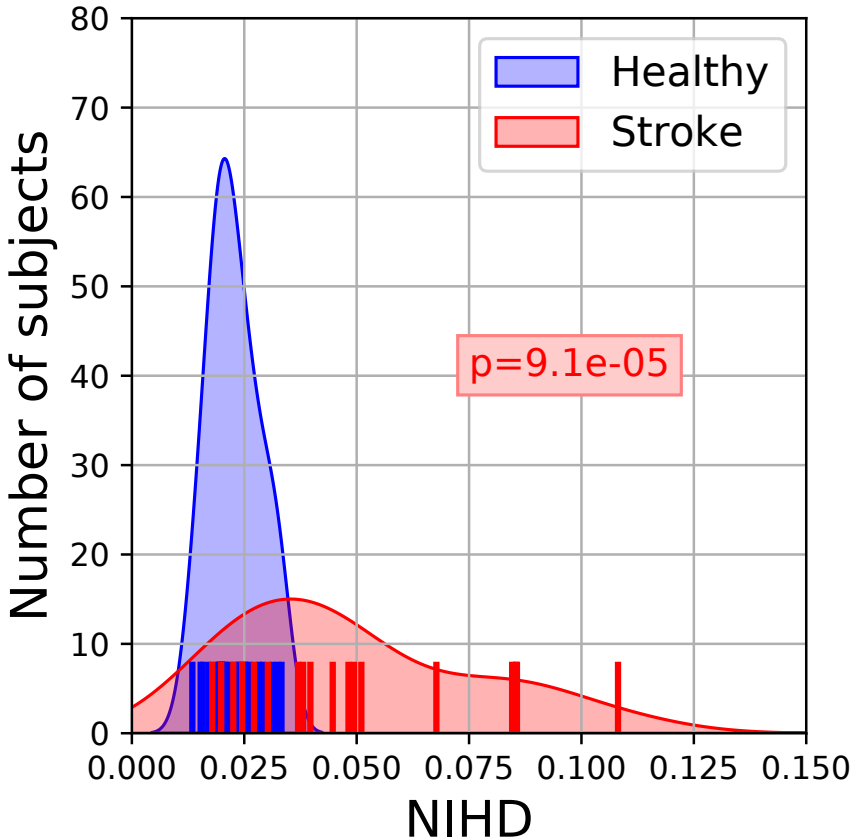


Figure: IHD distribution for each group.

Conclusions

- The spatial distribution of the HRF parameters $((\boldsymbol{\delta_m})_m)$ is more asymmetric for subjects who suffered from a stroke in the past.
- We observe an inter-subject variability within this group.

Experimental validation on the UKBB dataset

Goal:

- Classify each subject to its corresponding age group using his haemodynamic estimates.

Preprocessing and acquisition parameters:

- 486 rs-fMRI 6min acquisitions drawn at random from the UKBB dataset separated in two age-groups.

Decomposition parameters:

- $K = 20$
- $M = 96$ ('Harvard-Oxford' parcellation)
- $\lambda_f \in \{0.001, 0.22, 0.45, 0.67, 0.9\}$

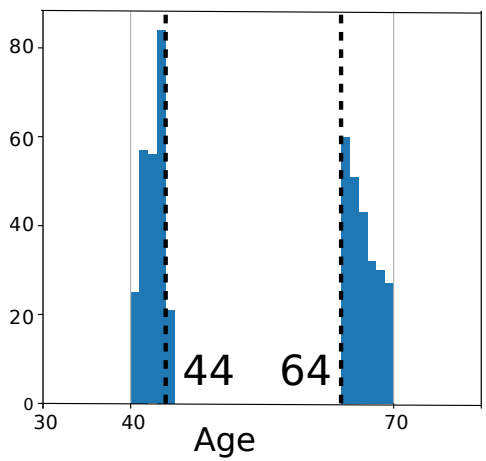


Figure: Age histogram

Experimental validation on the UKBB dataset

Classification score:

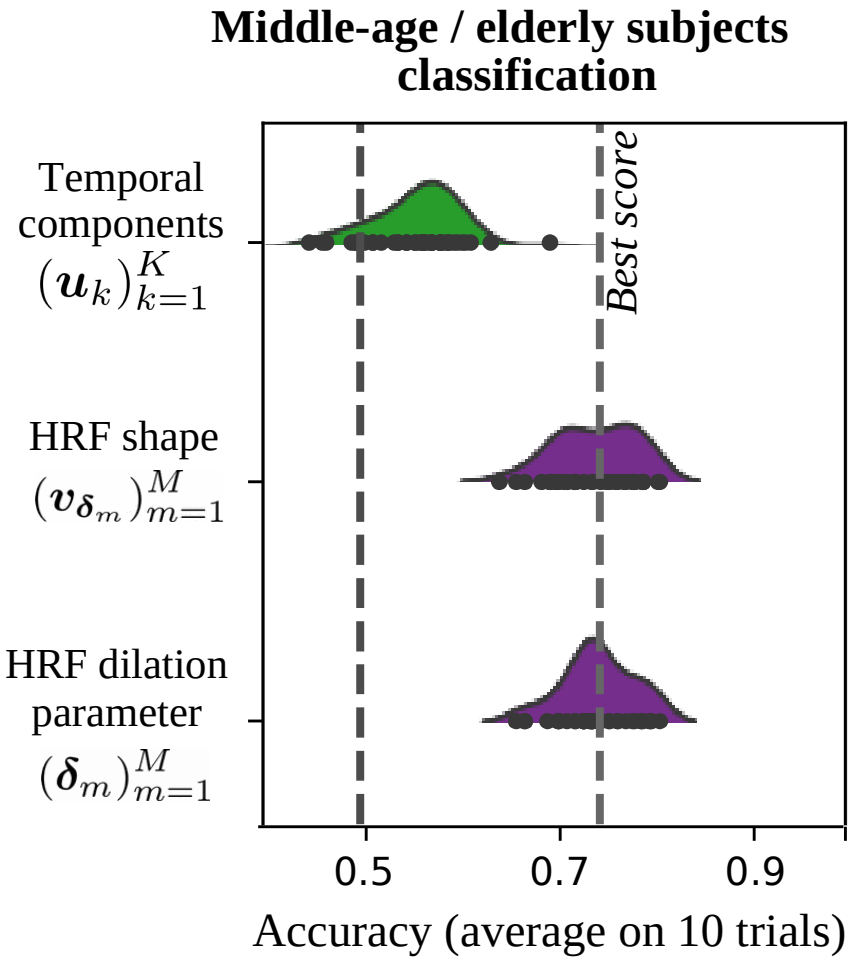


Figure: Classification score (best score: 0.74).

Conclusion

- The HRF parameters $((\delta_m)_m)$ and the HRF shape $((v_{\delta_m})_m)$ predict the age group.
- Our model captures the degradation of the neurovascular coupling induced by aging (West et al., 2020).

Experimental validation on the UKBB dataset

Running times

- In average one decomposition corresponds to: 500 time-frames and 8500 voxels (after a spatial sub-sampling to increase the signal-to-noise ratio).
- For a single subject decomposition time: 30 s on 1 CPU.
- For the UKBB dataset decomposition time for the age experiment: approximately 12 hours on 40 CPUs.

Remark

I made an effort to provide fast algorithms and an efficient Python implementation.

`https://github.com/hcherkaoui/hemolearn`

Minimization of

$$\min_{\boldsymbol{u}\in\mathbb{R}^T} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{v} * \boldsymbol{u}\|_2^2 + \lambda\|\boldsymbol{u}\|_{\mathrm{TV}}. \tag{3}$$

Related usage: Estimation of the neural activation signal $\boldsymbol{u}$ from the BOLD signal $\boldsymbol{y}$ with a fixed HRF $\boldsymbol{v}$.

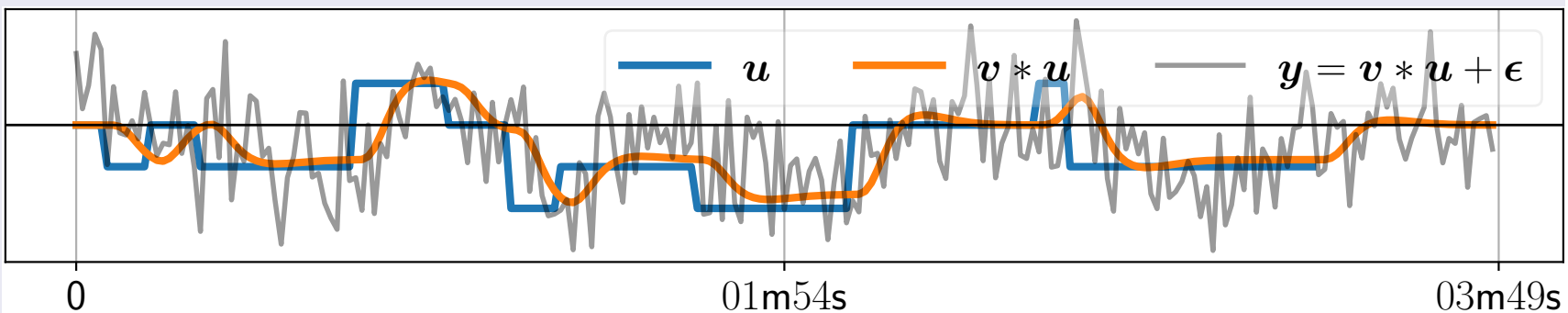


Figure: 1D fMRI signal deconvolution performed using TV regularization.

Generalization of the optimization problem:

Original problem

$$\min_{\boldsymbol{u}\in\mathbb{R}^T} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{v} * \boldsymbol{u}\|_2^2 + \lambda\|\boldsymbol{u}\|_{\mathrm{TV}}. \tag{4}$$

We replace $\boldsymbol{v} * \boldsymbol{u}$ by $\boldsymbol{A}\boldsymbol{u}$ to consider a more general case.

Minimization of

$$\min_{\boldsymbol{u}\in\mathbb{R}^T} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{u}\|_2^2 + \lambda\|\boldsymbol{u}\|_{\mathrm{TV}}. \tag{5}$$

Problem formulation

Equivalent re-formulation of the problem in 1D

Analysis formulation	**Synthesis** formulation
$\min_{\boldsymbol{u}\in\mathbb{R}^T} \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{u}\|_2^2 + \lambda \underbrace{\|\boldsymbol{D}\boldsymbol{u}\|_1}_{\|\boldsymbol{u}\|_{\mathrm{TV}}}$	$\min_{\boldsymbol{z}\in\mathbb{R}^T} \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{L}\boldsymbol{z}\|_2^2 + \lambda\|\boldsymbol{z}\|_1$

with $\boldsymbol{u} = \boldsymbol{L}\boldsymbol{z}$ and $\boldsymbol{L} = \boldsymbol{D}^{-1}$

Which formulation to choose:

We demonstrated, in Cherkaoui et al., NeurIPS, 2020, that the convergence rate of Analysis is much faster than the Synthesis.

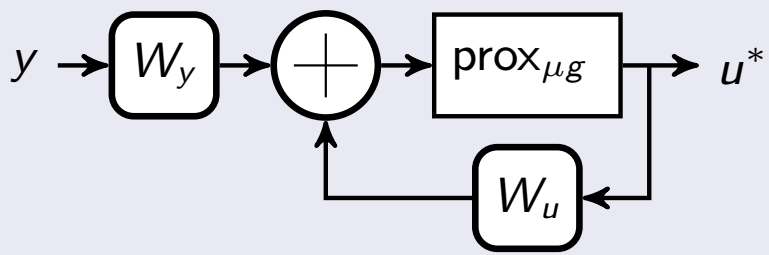
Classical minimization of the problem

Proximal Gradient Descent (PGD):

$$\boldsymbol{u}^{t+1} = \text{prox}_{\mu\|\cdot\|_{TV}}(\boldsymbol{u}^t - \mu \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{u}^t - \boldsymbol{y})). \tag{6}$$

Equivalent to a Recurrent Neural Network:

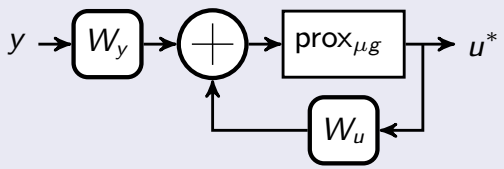
$$\boldsymbol{u}^{t+1} = \text{prox}_{\mu\|\cdot\|_{TV}}(\boldsymbol{W}_u\boldsymbol{u}^t + \boldsymbol{W}_y\boldsymbol{y}). \tag{7}$$



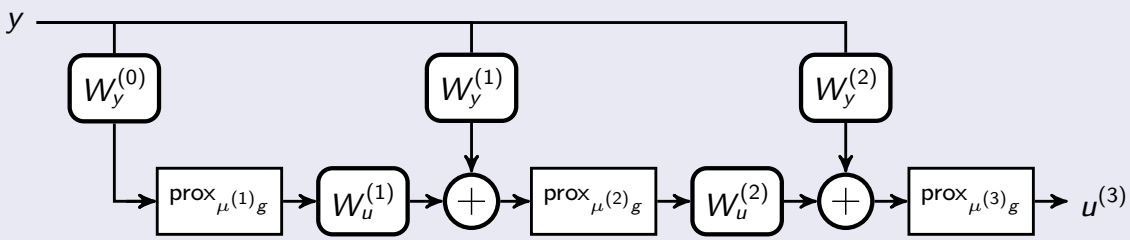
(a) **Proximal Gradient Descent (PGD)** - Recurrent Neural Network

Unrolling the analysis formulation

How to unroll proximal gradient descent?



(a) **PGD** - Recurrent Neural Network



(b) **LPGD** - Unfolded network for Learned PGD with $T = 3$

(Gregor and Le Cun, 2010)

How to compute the $\text{prox}_{\mu\|\cdot\|_{TV}}$ for each layer in a differential way?

Calculation of the proximal operator

Two approaches for differential computations of prox TV:

$$\text{prox}_{\mu\|\cdot\|_{TV}}(x) = \underset{u\in\mathbb{R}^T}{\arg\min}\ \frac{1}{2}\|x-u\|_2^2 + \mu\|u\|_{TV}. \tag{8}$$

Approximate the operator

- Use the equivalent synthesis formulation of Eq. (8) and a LISTA network (Gregor and Le Cun, 2010) to approximate the operator.
- Use the back-propagation to compute the gradient of the $\text{prox}_{\mu\|\cdot\|_{TV}}(\cdot)$ approximation.

Compute numerically the operator

- Solve the proximal operator numerically (Condat, 2013).
- Use the formula provided by Cherkaoui et al., NeurIPS, 2020 to compute the gradient of $\text{prox}_{\mu\|\cdot\|_{TV}}(\cdot)$.

Experimental validation on real fMRI data

Preprocessing, acquisition and decomposition parameters:

- 1 rs-fMRI acquisition drawn at random from the UKBB dataset.
- We retain only 8000 cropped time-series of 3 minute 3 seconds.
- We fix the HRF $\boldsymbol{v}$ and estimate the neural activity signal $\boldsymbol{u}$ for each voxel.

Performance comparison

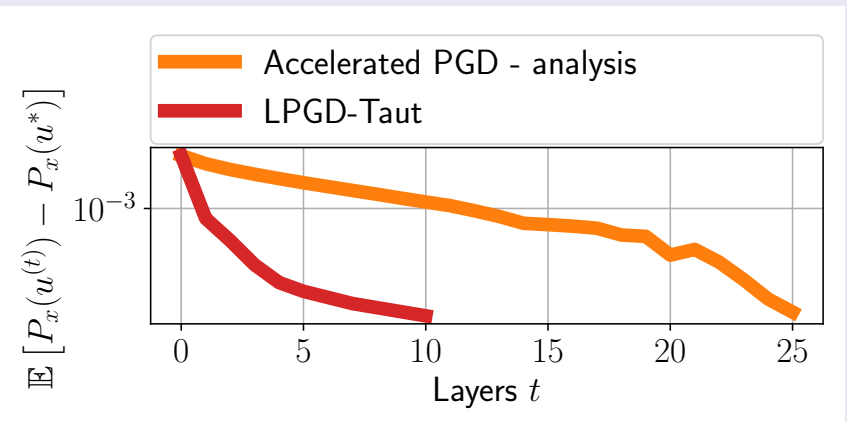


Figure: **Performance comparison:** Our analytic prox-TV derivative method outperforms the PGD in the analysis formulation.

The Synchropioid cohort

- Opioid-naive healthy volunteers: 30 subjects under placebo; 30 subjects under analgesic dose of buprenorphine.
- 90 min of [^{11}C]-buprenorphine PET imaging acquisition: localize the distribution of buprenorphine in the brain (partial agonist of μ-opioid receptors).
- 2 rs-fMRI sessions of 14 min: characterize the effect on the neurovascular coupling simultaneously to PET imaging.

The Synchropioid protocol

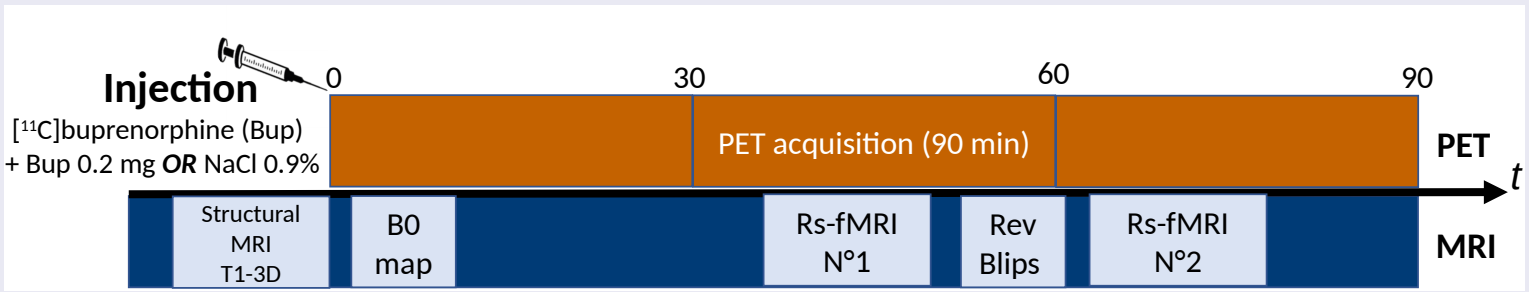


Figure: Illustration of the imaging protocol of the Synchropioid project.

Preprocessing and acquisition parameters:

- 2 rs-fMRI 14 min acquisition for each (2 volunteers on placebo condition and 2 volunteers on analgesic dose of buprenorphine).
- temporal resolution: $TR = 0.8$ s.
- classical preprocessing done with `fmriprep`.

Decomposition parameters:

- $K = 20$
- $M = 122$ ('BASC' parcellation, Bellec et al., 2013)
- $\lambda_f = 0.1$

Neurovascular coupling estimation

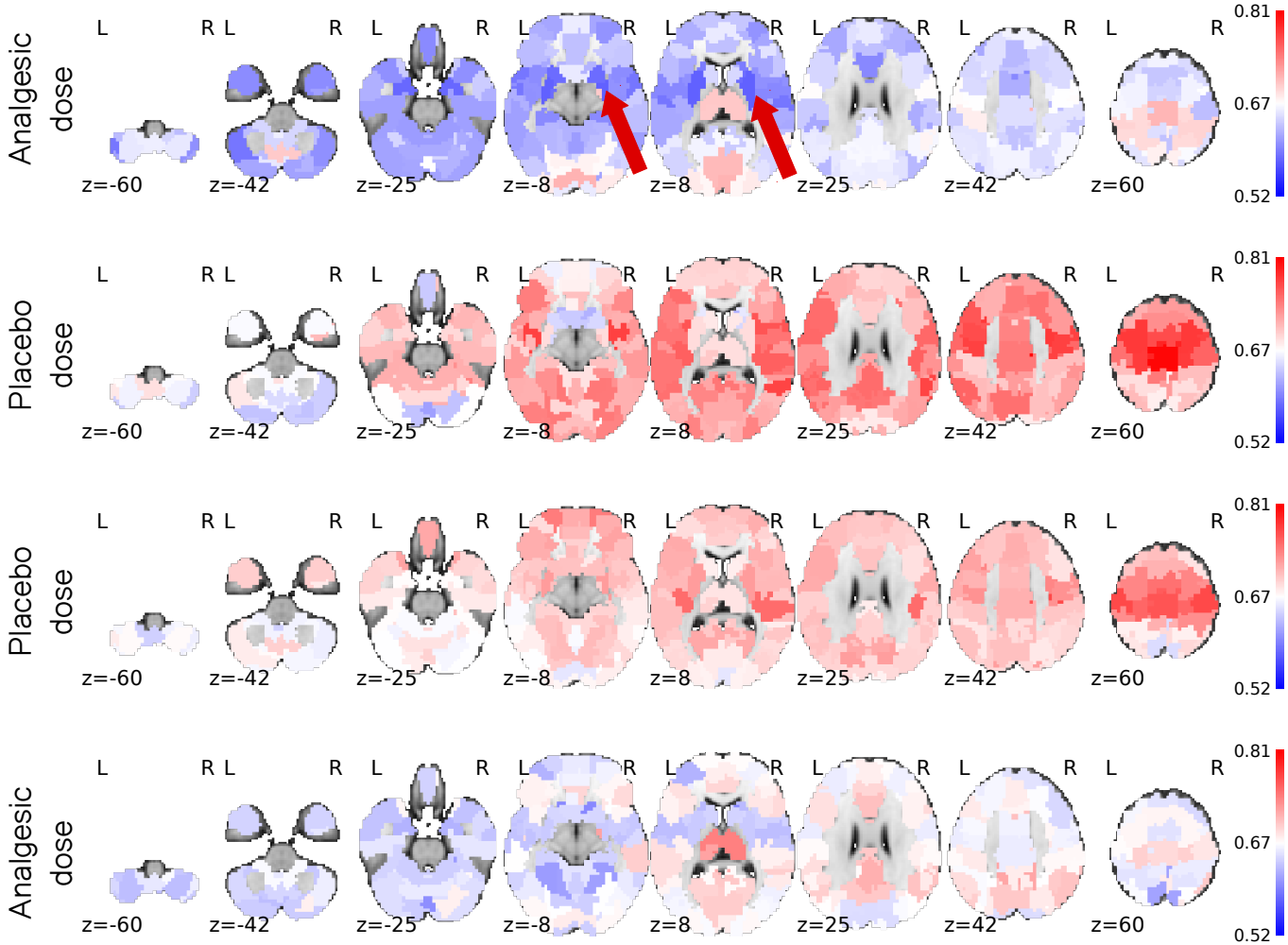


Figure: Maps of HRF dilation parameters $\boldsymbol{\delta}$ for rs-fMRI N°1 (Adriaens et al., 2014).

Buprenorphine spatial distribution

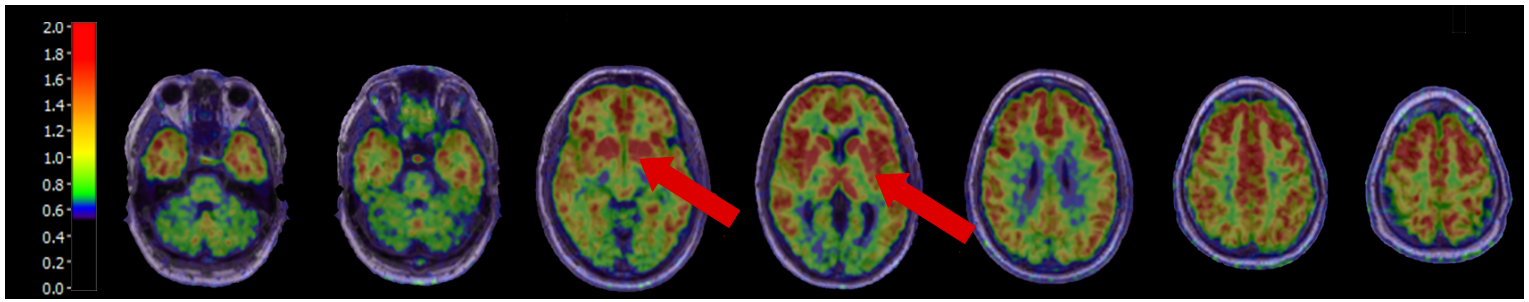


Figure: Spatial distribution of [^{11}C]buprenorphine uptake (SUVr maps).

Conclusion

- The spatial distribution of [^{11}C]buprenorphine is concordant with the known distribution of μ-opioid receptors (highest concentration in the putamen and the insula), Zubieta et al., 2000 and Greenwald et al., 2003.
- The highest [^{11}C]buprenorphine uptakes are concordant with the slowest haemodynamic delays (Cherkaoui et al., submitted to OHBM, 2021).

Variability over time across subjects:

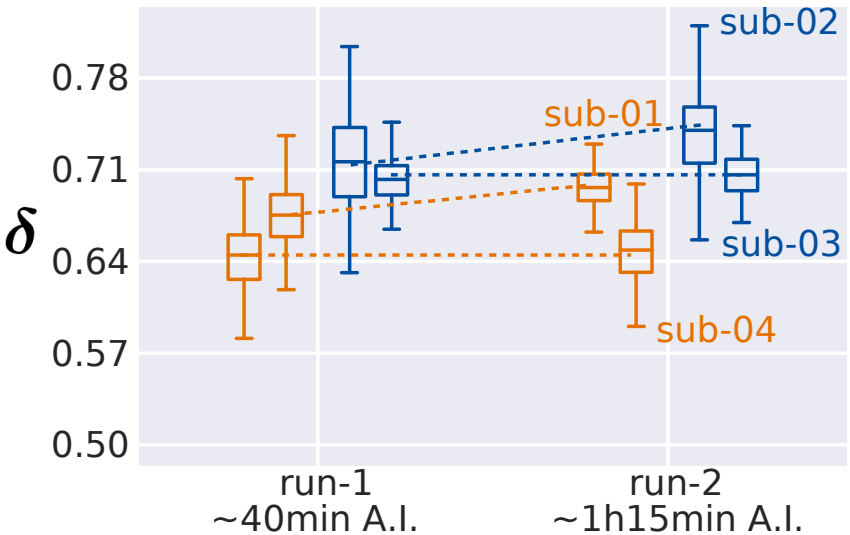


Figure: Evolution of the haemodynamic responses δ in each participant.

Conclusion

- We capture the effect of the buprenorphine on the neurovascular coupling.
- We observe a significant variability across time and subjects.
- This inter-subject variability needs to be better investigated in regards to the buprenorphine spatial fixation and the analgesic effect for each patient.

Conclusion

Take home message:

The joint estimation of the Haemodynamic Response Function and the Neural activation signal from fMRI is possible in both rs-fMRI and task fMRI as long as data is collected with a short TR.

Main contributions:

- A multivariate semi-blind deconvolution approach.
- Experimental validations on large cohort.
- Validation on a pharmacological context.
- New approach to minimize TV regularized problems.

Perspectives

Future developments:

- Estimation of shared spatial maps across subjects.
- Investigate other possible regularization and constraints.

Future clinical investigations:

- Expand the analysis of the Synchropioid cohort.
- Apply `HemoLearn` to the EpiTEP project (Dr Bouilleret).

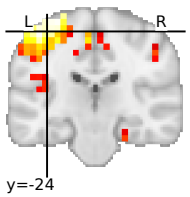
Thank you for listening !

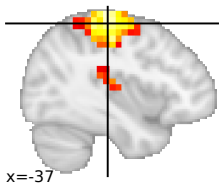
First author publications

- **Cherkaoui, H.** and Moreau, T. and Ciuciu, P. and Fernandez, B. and Bottlaender, M. and Tournier, N. and Leroy, C., *Characterization of the haemodynamic response function after a buprenorphine challenge study in Human healthy volunteer*, submitted to **OHBM**, 2021.
- **Cherkaoui, H.** and Moreau, T. and Halimi, A. and Leroy, C. and Ciuciu, P., *Multivariate semi-blind deconvolution of fMRI time series*, under review, **NeuroImage**, 2021.
- **Cherkaoui, H.** and Sulam, J. and Moreau, T., *Learning to solve TV regularised problems with unrolled algorithms*, **NeurIPS**, 2020.
- **Cherkaoui, H.** and Moreau, T. and Halimi, A. and Ciuciu, P., *Sparsity-based blind deconvolution of neural activation signal in fMRI*, **ICASSP**, 2019.
- **Cherkaoui, H.** and Moreau, T. and Halimi, A. and Ciuciu, P., *fMRI BOLD signal decomposition using a multivariate low-rank model*, **EUSIPCO**, 2019.
- **Cherkaoui, H.** and Gueddari, L. and Lazarus, C. and Grigis, A. and Poupon, F. and Vignaud, A. and Farrens, S. and Starck, J.-L.. and Ciuciu, P. *Analysis vs synthesis-based regularization for combined compressed sensing and parallel MRI reconstruction at 7 tesla*, **EUSIPCO**, 2018.

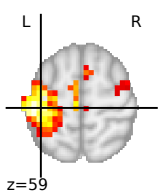
Comparison to ICA

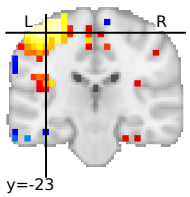
- *Left* motor cortex:

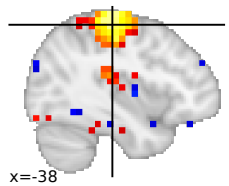


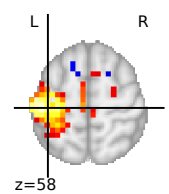


$\boldsymbol{w}_{13}$





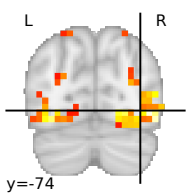


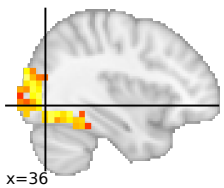


$\boldsymbol{w}_{03}$ ICA

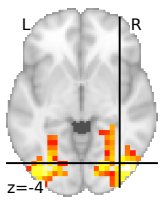
Figure: *left* motor cortex

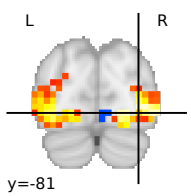
- Visual cortex:

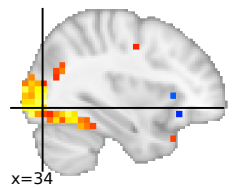


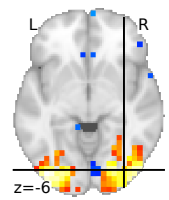


$\boldsymbol{w}_{16}$









$\boldsymbol{w}_{18}$ ICA

Figure: Visual cortex

Comparison to Wu et al., 2013

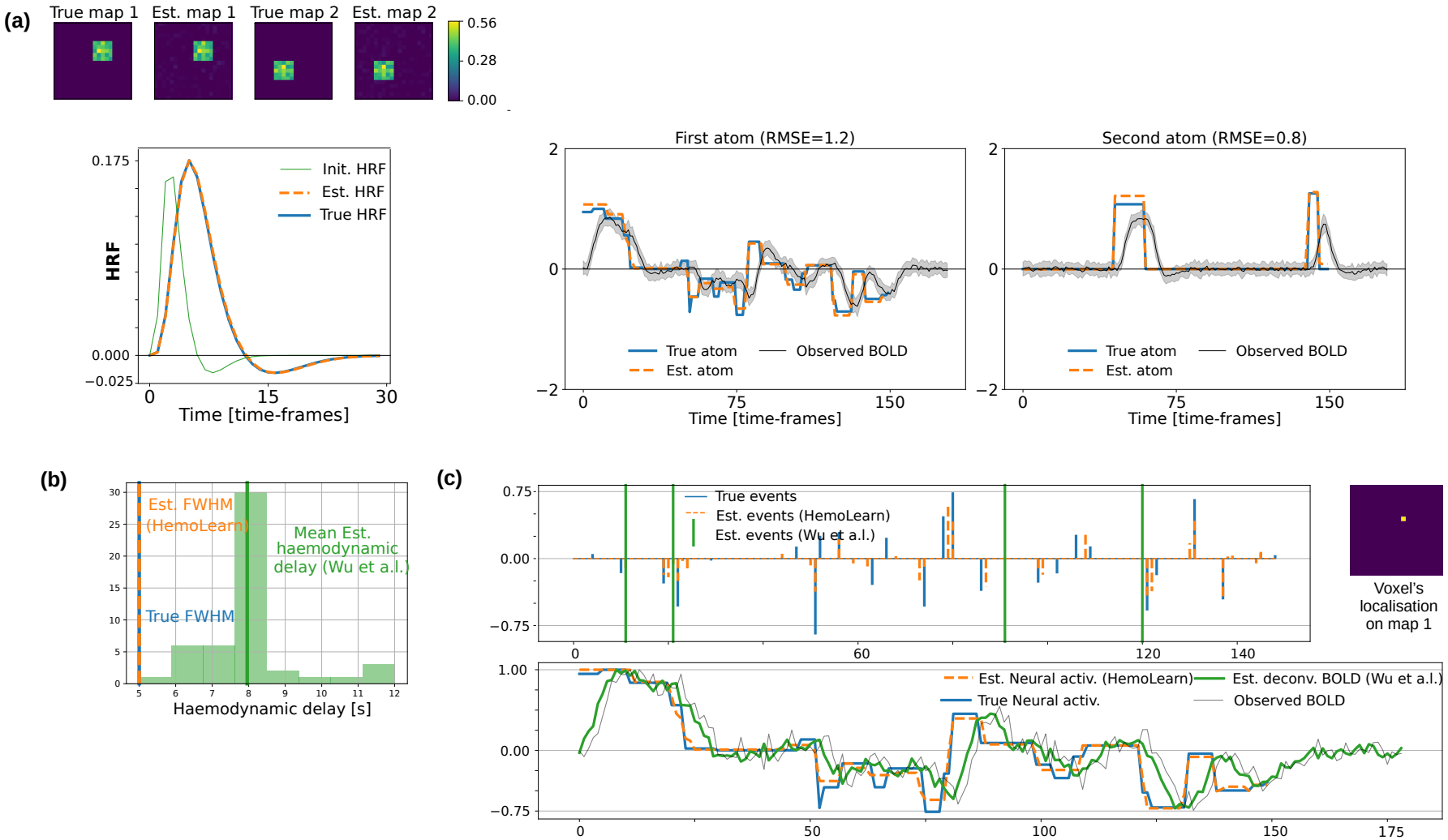


Figure: Comparison to Wu et al., 2013 on a synthetic case:
(a) Components estimation – (b) Haemodynamic delays estimation comparison – (c) Voxelwise semi-blind deconvolution example.

Convergence rate comparison

Analysis formulation convergence rate

$$P(u^{(t)}) - P(u^*) \leq \frac{\rho}{2t}\|u^{(0)} - u^*\|_2^2, \tag{9}$$

Synthesis formulation convergence rate

$$P(u^{(t)}) - P(u^*) \leq \frac{2\widetilde{\rho}}{t}\|u^{(0)} - u^*\|_2^2, \tag{10}$$

Theorem (Lower bound for the ratio $\frac{\|AL\|_2^2}{\|A\|_2^2}$ expectation)

Let A be a random matrix in $\mathbb{R}^{m\times k}$ with i.i.d normal entries. The expectation of $\|AL\|_2^2/\|A\|_2^2$ is asymptotically lower bounded when k tends to ∞ by

$$\mathbb{E}\left[\frac{\|AL\|_2^2}{\|A\|_2^2}\right] \geq \frac{2k+1}{4\pi^2} + o(1)$$

Convergence rate comparison

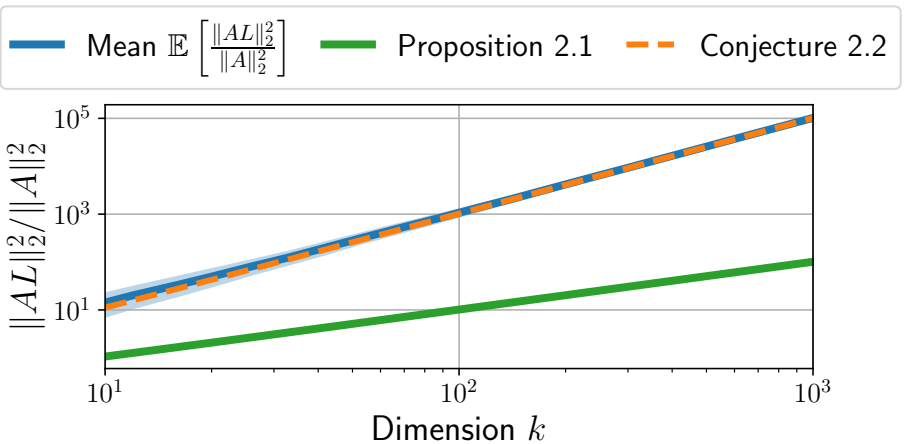


Figure: Evolution of $\mathbb{E}\left[\frac{\|AL\|_2^2}{\|A\|_2^2}\right]$ *w.r.t* the dimension k for random matrices A with *i.i.d* normal entries. In light blue is the confidence interval [0.1, 0.9] computed with the quantiles.

So, we can expect that $\widetilde{\rho}/\rho$ scales as $\Theta(k^2)$.
Which leads to $\frac{\widetilde{\rho}}{2} \gg \rho$ in large enough dimension.

The analysis formulation should be much more efficient in terms of iterations than the synthesis formulation.

Derivative of prox-TV

Theorem (Jacobian of prox-TV)

Let $x \in \mathbb{R}^k$ and $u = prox_{\mu\|\cdot\|_{TV}}(x)$, and denote by $\mathcal{S}$ the support of $z = \widetilde{D}u$. Then, the Jacobian J_x and J_μ of the prox-TV relative to x and μ can be computed as

$$J_x(x, \mu) = L_{:,\mathcal{S}}(L_{:,\mathcal{S}}^\top L_{:,\mathcal{S}})^{-1} L_{:,\mathcal{S}}^\top$$

and

$$J_\mu(x, \mu) = -L_{:,\mathcal{S}}(L_{:,\mathcal{S}}^\top L_{:,\mathcal{S}})^{-1} \operatorname{sign}(Du)_\mathcal{S}$$

Process summary

- Forward pass: use the Taut-string algorithm ($\Theta(k)$ complexity in most cases).
- Back-propagation pass: use the automatic-differentiation along with the analytic formulas of J_x and J_μ.