

Hemodynamic Deconvolution Demystified: Sparsity-Driven Regularization at Work

Eneko Uruñuela^{a,b,*}, Thomas A.W. Bolton^{c,d}, Younes Farouj^d, Dimitri Van De Ville^{d,e}, César Caballero-Gaudes^a

^a*Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastián, Spain.*

^b*University of the Basque Country (EHU/UPV), Donostia-San Sebastián, Spain.*

^c*Gamma Knife Center, Department of Clinical Neuroscience, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland*

^d*Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.*

^e*Faculty of Medicine, University of Geneva, Geneva, Switzerland*

Abstract

Deconvolution of the hemodynamic response is an important step to access short timescales of brain activity recorded by functional magnetic resonance imaging (fMRI). Albeit conventional deconvolution algorithms have been around for a long time (e.g., Wiener deconvolution), recent state-of-the-art methods based on sparsity-pursuing regularization are attracting increasing interest to investigate brain dynamics. This technical note revisits the main concepts underlying two main methods, Paradigm Free Mapping and Total Activation, in the most accessible way. Despite their apparent differences, these methods are theoretically equivalent as they represent the synthesis and analysis sides of the same problem. We demonstrate this equivalence in practice with their best-available implementations using both simulations, with different signal-to-noise ratios, and experimental data of motor task and resting-state fMRI. We evaluate the parameter settings that lead to equivalent results, and benchmark the computational speed of both algorithms. This note is useful for practitioners interested in having a better understanding of state-of-the-art hemodynamic deconvolution, and who want to make use of the most efficient implementation.

Keywords: fMRI deconvolution, paradigm free mapping, total activation

1. Introduction

Functional magnetic resonance imaging (fMRI) data analysis is often directed to identify and disentangle the neural processes that occur in different brain regions during task or at rest. As the blood oxygenation level-dependent (BOLD) signal of fMRI is only a proxy for neuronal activity mediated through neurovascular coupling, an intermediate step that estimates the activity-inducing signal, at the timescale of fMRI, from the BOLD timeseries can be useful. Conventional analysis of task fMRI data relies on the general linear models (GLM) to establish statistical parametric maps of brain activity by regression of the empirical timecourses against hypothetical ones built from

*Corresponding author

Email address: e.urunuela@bcbl.eu (Eneko Uruñuela)

the knowledge of the experimental paradigm. However, timing information of the paradigm can be
10 unknown, inaccurate, or insufficient in some scenarios such as naturalistic stimuli.

Deconvolution and methods alike are aiming to estimate neuronal activity by undoing the effect
15 of the hemodynamic response function (HRF). Given the inherently ill-posed nature of hemodynamic
deconvolution, due to the strong temporal low-pass characteristics of the HRF, the key is
to introduce additional constraints that are typically expressed as regularizers. For instance, the
so-called Wiener deconvolution is expressing a “minimal energy” constraint on the deconvolved
20 signal, and has been used in the framework of psychophysiological interactions analysis to compute
the interaction between a seed’s activity timecourse and an experimental modulation (Glover
1999; Gitelman et al. 2003). More recently, the interest in deconvolution has increased to explore
time-varying activity in resting-state fMRI data. In that case, the aim is to gain better insights
25 into the signals that drive functional connectivity and its temporal dynamics, as well as the organizational
principles of brain function. In particular, the spatio-temporal structure of functional
components that dynamically construct resting-state networks (Petridou et al. 2013; Karahanoglu
and Van De Ville 2015, 2017; Kinany et al. 2020; Gonzalez-Castillo et al. 2019; Allan et al. 2015).
30 Deconvolution techniques can also prove to be helpful in clinical conditions to characterize functional
alterations of patients with a progressive stage of multiple sclerosis at rest (Bommarito et al.
2020), to find functional signatures of prodromal psychotic symptoms and anxiety at rest on patients
suffering from schizophrenia (Zöller et al. 2019), to detect the foci of interictal events in epilepsy patients without an EEG recording (Lopes et al. 2012), or to study functional dissociations
35 observed during non-rapid eye movement sleep that are associated with reduced consolidation of information and impaired consciousness (Tarun et al. 2020).

A series of recent studies have attempted to understand neural processes by studying the interactions
35 between BOLD responses without estimating the underlying neuronal activity. For instance, co-activation patterns have been used to replicate seed correlation-based resting-state functional
networks with a small portion of the data (Liu and Duyn 2013; Liu et al. 2013, 2018; Majeed et al.
2009, 2011; Cifre et al. 2020a,b; Zhang et al. 2020). Likewise, the dynamics of functional connectivity
have recently been investigated with the use of co-fluctuations and edge-centric techniques
40 on tasks (Faskowitz et al. 2021), resting-state (Esfahlani et al. 2020) and naturalistic paradigms
(Faskowitz et al. 2020; Betzel et al. 2020). Methods based on the multiplication of temporal derivatives
have also been presented for the estimation of dynamic functional connectivity on task fMRI
data (Shine et al. 2015, 2016).

This note revisits synthesis- and analysis-based deconvolution methods for fMRI data and comprises four sections. In the first, we present the theory behind two state-of-the-art deconvolution approaches based on estimators that promote sparsity: Paradigm Free Mapping (PFM) (Cabanturo Gaudes et al. 2013) — available as *3dPFM* and *3dMEPFM* in AFNI — and Total Activation (TA) (Karahanoglu et al. 2013) — available as part of the *iCAPs toolbox*. We then assess their performance controlling for a fair comparison on simulated and experimental data. Finally, we discuss the benefits and shortcomings of each technique and conclude with our vision on potential extensions and developments.

2. Theory

I think we need a short paragraph here introducing: 1) our notations (vectors, matrices, discrete, continuous); 2) basic math definitions, in particular for norms.

Add more detailed refs.

revise next part—too self-centered.
I think we should just cite some general dynFC review papers first, and then go into more detail about the proposed deconvolution methods papers by us.

Conventional general linear model (GLM) analysis puts forward a number of regressors incorporating knowledge about the paradigm or behavior. For instance, the timing of epochs for a certain condition can be modeled as an indicator function $p(t)$, convolved with the HRF $h(t)$, and sampled at TR resolution (Friston et al. 1994, 1998; Boynton et al. 1996; Cohen 1997):

$$p(t) \rightarrow p * h(t) \rightarrow x[k] = p * h(k \cdot \text{TR}).$$

Rather some
Friston pa-
per(s) here

The vector $\mathbf{x} = [x[k]]_{k=1,\dots,N}$ then constitutes the hypothetical response, and several of them can be stacked as the columns of the design matrix $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_L]$, leading to the celebrated GLM:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad (1)$$

where the empirical timecourse \mathbf{y} is explained by a linear combination of the regressors in \mathbf{X} weighted by the parameter weights in $\boldsymbol{\beta}$ and corrupted by additive noise \mathbf{e} . Under independent and identically distributed Gaussian assumptions of the latter, the maximum likelihood estimate of the parameter weights reverts to the ordinary least-squares estimator; i.e., minimizing the residual sum of squares between the fitted model and measurements. The number of regressors L is typically much less than the number of measurements N , and thus the regression problem is over-determined and does no require additional constraints or assumptions.

In the deconvolution approach, no prior knowledge is taken into account, and the purpose is to estimate the deconvolved activity-inducing signal \mathbf{s} from the measurements \mathbf{y} , which can be formulated as the signal model

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{e}, \quad (2)$$

where \mathbf{H} is an $N \times N$ Toeplitz matrix that represents the discrete convolution with the HRF, and \mathbf{s} a length- N vector with the unknown activity-inducing signal. Despite the apparent similarity with the GLM equation, there are two important differences. First, the multiplication with the design matrix of the GLM is an expansion as a weighted linear combination of its columns, while the multiplication with the HRF matrix represents a convolution with its shifted rows. Second, determining \mathbf{s} is an ill-posed problem given the nature of the HRF; e.g., as can be seen intuitively, the rows of \mathbf{H} are highly correlated due to large overlap between shifted HRFs (see Figure 3B), thus introducing ambiguity and instability in the estimates of \mathbf{s} . Therefore, additional assumptions under the form of regularization are needed.

2.1. Paradigm-free mapping

The first approach for deconvolution is based on the forward model formulation of Eq. (2); i.e., the to-be-recovered activity-inducing signal $\hat{\mathbf{s}}$ is convolved by the HRF and compared against the measured timecourse in the least-squares sense:

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2 + \Omega(\mathbf{s}). \quad (3)$$

The first term quantifies data fitness, which can be justified as the log-likelihood term derived from Gaussian noise assumptions, while the second term $\Omega(\mathbf{s})$ brings in regularization and be interpreted as a prior on the activity-inducing signal. In classical Wiener deconvolution, the (squared) ℓ_2 -norm of \mathbf{s} is imposed (i.e., $\Omega(\mathbf{s}) = \lambda \|\mathbf{s}\|_2^2$), which introduces a trade-off, controlled by the regularization parameter λ , between data fit and “energy” of the solution. In paradigm-free mapping (PFM), the formulation of Eq. (3) was considered equivalently as fitting the measurements using the atoms

of a dictionary (columns of \mathbf{H}) with corresponding weights (entries of \mathbf{s}) (Gaudes et al. 2011; Caballero Gaudes et al. 2013; Uruñuela et al. 2020). In addition, sparsity-pursuing regularization was introduced on \mathbf{s} , which in a strict way reverts to choosing $\Omega(\mathbf{s}) = \lambda\|\mathbf{s}\|_0$ and solving the optimization problem (Bruckstein et al. 2009). However, due to the convolution model defined in (3), finding the optimal solution to the problem demands an exhaustive search across all possible combinations of the columns of \mathbf{H} . Hence, a pragmatic solution is to solve the convex-relaxed optimization problem for the l_1 -norm, commonly known as the LASSO (Tibshirani 1996):

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2 + \lambda\|\mathbf{s}\|_1, \quad (4)$$

which provides fast convergence to a global solution. From the neuronal perspective, imposing sparsity on the activity-inducing signal expresses that activity, at the fMRI timescale of seconds,
70 is supposed to be short and typically can be explained by a small number of large entries. Further on, we refer to this assumption as the spike model.

2.2. Total activation

Alternatively, deconvolution can be formulated as a denoising problem where the signal to be recovered is directly fitting the measurements and at the same time satisfying some suitable regularization, which leads to

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \Omega(\mathbf{x}). \quad (5)$$

Well-known regularized regression techniques such as ridge regression (i.e., $\Omega(\mathbf{x}) = \lambda\|\mathbf{x}\|_2^2$) and elastic net (i.e., $\Omega(\mathbf{x}) = \lambda_1\|\mathbf{x}\|_2^2 + \lambda_2\|\mathbf{x}\|_1$) fall under this category [REF]. One other powerful regularizer is total variation (TV), which is the l_1 -norm of the derivative, $\Omega(\mathbf{x}) = \lambda\|\mathbf{Dx}\|_1$, and favors recovery of piecewise-constant signals [REF]. The approach of generalized TV introduces an additional differential operator \mathbf{D}_H in the regularizer that can be tailored as the inverse operator of a linear system (Karahanoglu et al. 2011), that is, $\Omega(\mathbf{x}) = \lambda\|\mathbf{DD}_H\mathbf{x}\|_1$. In the context of hemodynamic deconvolution, total activation is proposed for which the discrete operator \mathbf{D}_H is derived from the inverse of the continuous-domain linearized balloon-windkessel model. Exchanging the poles and zeros of the latter's linear-system characterization leads to a differential operator of the form

$$D_H = \prod_{i=1}^{M_1} (D - \alpha_i I) \left(\prod_{j=1}^{M_2} (D - \gamma_j I) \right)^{-1}, \quad (6)$$

where D is the derivative operator, α_i the zeros, and γ_j the poles. The interested reader is referred to (Karahanoglu et al. 2013) for a detailed description.

Therefore, the solution of the total-activation problem

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\|\mathbf{DD}_H\mathbf{x}\|_1 \quad (7)$$

75 will render the activity-related signal \mathbf{x} for which the activity-inducing signal $\mathbf{s} = \mathbf{D}_H\mathbf{x}$ and so-called innovation signal $\mathbf{u} = \mathbf{Ds}$ will also be available, as they are required for the regularization. We refer to the assumption for the activity-inducing signal as the block model.

Even without would be the case, right?

2.3. Unifying both perspectives

I would go for a schematic block-flowchart that goes between activity-related/inducing/innovation signals with between each step the two operators; i.e., \mathbf{H} and $\mathbf{D}_\mathbf{H}$ for the first, \mathbf{L} and \mathbf{D} for the second.

PFM and TA are based on the synthesis- and analysis-based formulation of the deconvolution problem, respectively. In the first case, the recovered deconvolved signal is synthesized to be matched to the measurements, while in the second case, the recovered signal is directly matched to the measurements but needs to satisfy its analysis in terms of deconvolution. This also corresponds to using the forward or backward model of the hemodynamic system, respectively. Both approaches are equivalent¹ [REF Elad]. First, TA can be made equivalent to PFM by removing the derivative operator of TA's regularizer. It can then be readily verified that replacing in that case $\mathbf{x} = \mathbf{H}\mathbf{s}$ leads to identical equations and thus both assume a spike model. Second, PFM can also be made equivalent to TA, by considering the modified forward model

$$\mathbf{y} = \mathbf{HLu} + \mathbf{e},$$

Would be nice to check a bit closer, this is what I concluded.

where the activity-inducing signal \mathbf{s} is rewritten in terms of the innovation signal \mathbf{u} ; i.e., the derivative $\mathbf{Ds} = \mathbf{u}$ of \mathbf{s} (Cherkaoui et al. 2019; Uruñuela et al. 2020). This way, PFM will solve for the innovation signal \mathbf{u} :

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \frac{1}{2} \|\mathbf{y} - \mathbf{HLu}\|_2^2 + \lambda \|\mathbf{u}\|_1, \quad (8)$$

and becomes equivalent to TA by replacing $\mathbf{u} = \mathbf{DD}_\mathbf{H}\mathbf{x}$, and thus adopting the block model.

2.4. Algorithms

I think we need a subsection here that gives a brief overview of the different computational approaches to solve PFM and TA. In particular, I think that for the first Figure mentioned above, we could have a second panel (b) in which we put a flowchart of LARS and FGP to illustrate the main idea. – maybe the parameter selection can be actually merged into this subsection as it will depend on the algorithm.

Despite their algebraic equivalence, both techniques are solved with their own type of techniques. XXX Here, we evaluate PFM and TA solved using XX and XX, respectively. We do not introduce additional spatial regularization.

2.5. Parameter selection

The correct selection of the regularization parameter λ is a critical decision for the accurate performance of deconvolution methods. Even though many techniques have been proposed in the literature, the optimal strategy that selects λ is yet to be found. Algorithms like least angle regression (LARS) (Efron et al. 2004) provide all the possible solutions to the optimization problem and their corresponding value of λ , i.e., the regularization path, but don't provide the optimal solution. Therefore, strategies that exploit the regularization path can provide a selection of λ that is close to the optimal. For instance, in PFM, the optimal result is given by the Bayesian

¹Without dwelling into technicalities, this equivalence is correct up to the constant, which is in the null space of the derivative operator.

Information Criterion (BIC) (Schwarz 1978), i.e., the regularization path solution with the minimum BIC is considered optimal. Another approach could be to update the regularization parameter λ on every iteration n like TA does, so that the residuals converge to a previously estimated noise level of the data fit $\tilde{\sigma}$. The pre-estimated noise level is calculated from the median absolute deviation of fine-scale wavelet coefficients (Daubechies, order 3) (Karahanoglu et al. 2013):

$$\lambda^{n+1} = \frac{N\tilde{\sigma}}{\frac{1}{2}\|\mathbf{y} - \mathbf{x}^n\|_F^2}\lambda^n, \quad (9)$$

where x^n is the n^{th} iteration estimate, ...

Here, we compare the performance of the two deconvolution algorithms with both selection criteria and in terms of the estimation of the activity-inducing signal \mathbf{s} using the *spike model* in (4) and the innovation signal \mathbf{u} using the *block model* in (8).

3. Methods

3.1. Simulated data

In order to compare the two methods while controlling for their correct performance, we simulated a 400 seconds (TR = 2 s) activity-inducing signal with five neuronal events, convolved it with the canonical HRF, and we added noise of different sources (physiological, thermal, and motion-related) with different signal-to-noise ratios (SNR = [20 dB, 10 dB, 3 dB]) that represent low, medium and high levels of noise as shown in Figure 1. Noise was created following the procedure in (Caballero Gaudes et al. 2013) as the sum of uncorrelated Gaussian noise and sinusoidal signals to simulate a realistic noise model with thermal noise, cardiac and respiratory physiological fluctuations. We generated the sinusoidal term as

$$\sum_{i=1}^2 \frac{1}{2^{i-1}} (\sin(2\pi f_{r,i}t + \phi_{r,i}) + \sin(2\pi f_{c,i}t + \phi_{c,i})), \quad (10)$$

with up to second-order harmonics per cardiac ($f_{c,i}$) and respiratory ($f_{r,i}$) component that were randomly generated following normal distributions with variance 0.04 and mean if_r and if_c , for $i = [1, 2]$. We set the fundamental frequencies to $f_r = 0.3$ Hz for the respiratory component (Birn et al. 2006) and $f_c = 1.1$ Hz for the cardiac component (Shmueli et al. 2007). The phases of each harmonic ϕ were randomly selected from a uniform distribution between 0 and 2π radians. In order to simulate physiological noise that is proportional to the change in BOLD signal, a variable ratio between the physiological (σ_P) and the thermal (σ_0) noise was modeled as $\sigma_P/\sigma_0 = a(tSNR)^b + c$, where $a = 5.01 \times 10^{-6}$, $b = 2.81$, and $c = 0.397$. The physiological-thermal noise model was extracted following the experimental measures of the physiological-to-thermal noise ratio at 7T in Table 3 in (Triantafyllou et al. 2005). The code used to simulate the data can be found in the GitHub repository shared in section 6.

3.2. Experimental data

Motor task dataset: One healthy subject was scanned in a 3T MR scanner (Siemens) as part of a larger experiment under a Basque Center on Cognition, Brain and Language Review Board-approved protocol. T2*-weighted multi-echo fMRI data was acquired with a multiband (MB) multi-echo gradient echo-planar imaging sequence (340 scans, 52 slices, Partial-Fourier = 6/8, voxel

size = 2.4x2.4x3 mm³, TR = 1.5 s, TEs = 10.6/28.69/46.78/64.87/82.96 ms, multiband factor = 4, flip angle = 70°, GRAPPA = 2).

During the fMRI acquisition, subjects performed a motor task consisting of five different movements (left-hand finger tapping, right-hand finger tapping, moving the left toes, moving the right toes and moving the tongue). These conditions were randomly intermixed every 16 seconds, and were only repeated once the entire set of stimuli were presented. Data preprocessing consisted of optimally combining the echo time datasets, detrending of up to 5th-order Legendre polynomials, spatial smoothing (3 mm FWHM) and normalization to signal percentage change. Figure 2 shows the time-series of a representative voxel in the motor cortex, where the colored bands illustrate the onset and duration of the right-hand finger-tapping condition of the paradigm.

Resting-state datasets: One healthy subject was scanned in a 3T MR scanner (Siemens) as part of a larger experiment under a Basque Center on Cognition, Brain and Language Review Board-approved protocol. Two runs of T2*-weighted fMRI data were acquired during resting-state, each with 10 min duration, with 1) a standard gradient-echo echo-planar imaging sequence (monoband) (TR = 2000 ms, TE = 29 ms, flip-angle = 78°, matrix size = 64x64, voxel size = 3x3x3 mm³, 33 axial slices with interleaved acquisition, slice gap = 0.6 mm) and 2) a simultaneous multislice gradient-echo echo-planar imaging sequence (multiband factor = 3) developed by the Center of Magnetic Resonance Research (University of Minnesota, USA; TR = 800 ms, TE = 29 ms, flip-angle = 60°, matrix size = 64×64, voxel size = 3x3x3 mm³, 42 axial slices with interleaved acquisition, no slice gap). Single-band reference images were also collected in both resting-state acquisitions for head motion realignment. Field maps were also obtained to correct for field distortions.

During both acquisitions, participants were instructed to keep their eyes open, fixating a white cross that they saw through a mirror located on the head coil, and not think about anything specific. The data was pre-processed using AFNI (Cox 1996). First, volumes corresponding to the initial 10 seconds were removed to allow for a steady-state magnetization. Then, the voxel time-series were despiked to reduce large-amplitude deviations and slice-time corrected. Inhomogeneities caused by magnetic susceptibility were corrected with FUGUE (FSL) using the field map images (Jenkinson et al. 2012). Next, functional images were realigned to a base volume (monoband: volume with the lowest head motion; multiband: single-band reference image). Finally, a simultaneous nuisance regression step was performed comprising up to 6th-order Legendre polynomials, low-pass filtering with a cutoff frequency of 0.25 Hz (only on multiband data to match the frequency content of the monoband), 6 realignment parameters plus temporal derivatives, 5 principal components of white matter (WM), 5 principal components of lateral ventricle voxels (anatomical CompCor) and 5 principal components of the brain's edge voxels. WM, CSF and brain's edge-voxel masks were obtained from FreeSurfer tissue and brain segmentations. In addition, scans with potential artifacts were identified and censored as the euclidean norm of the temporal derivative of the realignment parameters (ENORM) was larger than 0.4, and the proportion of voxels adjusted in the despiking step exceeded 10%.

3.3. Selection of the hemodynamic response function

With the aim of making a fair comparison between both methods, we first compared their HRFs. Figure 3A shows the difference in the HRF that PFM and TA use by default for TR = 0.1 s and TR = 1 s adjusted to peak amplitude of one; i.e., the canonical HRF and the HRF resulting from the linear differential operator. The most observable difference between the two HRFs is the time to peak: the HRF used by Total Activation does not begin at zero while the one used by PFM does.

Fig. seems corrupted?

I don't understand the x-axis in the figure, is then index, not time?

But that's not related to TTP

While PFM allows for the use of any HRF—the columns of the design matrix \mathbf{H} are composed by shifted versions of the HRF—the linear differential operator in TA is tailored for a fixed HRF.
 Hence, for practical reasons, we reproduced the HRF in the Total Activation filter and incorporated it into the PFM formulation (Figure 3B).

4. Results

4.1. Selection of the regularization parameter by solving the regularization path

Paradigm Free Mapping bases its selection of the regularization parameter on the BIC. Hence, we calculated the regularization path with PFM by means of the least angle regression (LARS) algorithm (Efron et al. 2004) and used the λ s from the regularization path to solve the deconvolution problem with Total Activation.

Figure 4A (left) shows the regularization paths of PFM and TA side by side for the three SNR conditions for the spike model; i.e., the inverse problem described in (4). Each iteration of LARS reduces the value of λ ; i.e., reduces the sparsity promoted by the l_1 -norm, and reveals new non-zero coefficients as shown in the x-axis of the heatmaps. Vertical black lines depict the selection of the regularization parameter based on BIC, and thus, the colored coefficients indicated by the vertical lines depict the estimated activity-inducing signal $s(t)$. Figure 4A (right) illustrates the resulting estimation of the activity-inducing and neuronal-related signals when basing the selection of λ on BIC for the three simulated SNR conditions. Given that the regularization paths of both techniques are identical, the BIC-based selection of the regularization parameter and the results of deconvolving with said λ are identical too (see Figure ??). Thus, Figure 4A demonstrates that, regardless of the simulated SNR condition, both deconvolution algorithms produce identical regularization paths when the same HRF and regularization parameters are applied, and hence, identical estimates of the activity-inducing signal s and neuronal-related signal x .

The regularization path to estimate innovation signals yields mainly undistinguishable results for both PFM and TA methods as shown in Figure 4A (left). Again, the BIC-based selection of λ is identical for both PFM and TA, and the estimation of the innovation signal u shows no distinguishable differences between the algorithms (see Figure 4A right). Therefore, both Paradigm Free Mapping and Total Activation yield nearly identical regularization paths and estimates of the innovation signal regardless of the simulated SNR condition when applying the same HRF and regularization parameters with the block model.

4.2. Selection of the regularization parameter based on the estimation of the noise

We calculated the regularization path with PFM (as described in 4.1) and selected the λ corresponding to the residuals that were closest to the estimated noise level of the data. We applied Total Activation with temporal regularization in its original form. Figure 4B depicts the estimated activity-inducing, innovation, and activity-related signals when updating λ following (9) in the three simulated SNR settings using the spike model (left) and the block model (right). Figure 4B (left) shows nearly identical results between PFM (left) and TA (right) with the use of the spike model. The minimal differences are the result of slight dissimilarities in the convergence of the residuals to the estimated noise level of the data. Likewise, the use of the block model with a selection of λ based on the MAD estimate of the noise yields results that are identical in practice as shown in Figure 4B (right).

It's maybe not so simple: what I take from Elad's paper is that any convolution operator \mathbf{H} (for synthesis) can be turned in an equivalent analysis operator.

To be explained above in Algorithms.

195 *4.3. Performance on experimental data*

Additionally, in order to describe the extent of the discrepancies between the techniques, we calculated the residual sum of squares (RSS); i.e., the sum of squares of the differences between the estimated activity-inducing or innovation signals of PFM and TA as:

$$RSS = \frac{\sum_k (\hat{s}_{\text{PFM}}[k] - \hat{s}_{\text{TA}}[k])^2}{N}, \quad (11)$$

where N is the number of frames of the acquisition.

Figure 5 depicts the RSS of the spike (Figure 5A) and block (Figure 5B) models for the three experimental datasets introduced: i.e., motor, monoband and multiband. It is clear that RSS values are lower than those of the activity-inducing and innovation signals, suggesting that the differences between PFM and TA are negligible. Furthermore, the largest differences can be seen in regions with high vasculature and are probably a result of differences in the amplitude of the estimated activity-inducing and innovation signals.

What about computation time? Do something similar as Hamza in his PhD?

What about Figs. 6 and 7? Why is PFM and TA not equivalent there?

Expand more on interpretation of this Figure.

205 **5. Discussion**

Should be expanded to provide a clear message.

This work demonstrates that PFM and TA yield practically identical results when the same HRF model and equivalent regularization parameter are employed, demonstrating that synthesis and analysis models are equivalent for temporal fMRI deconvolution. Thus, previously observed differences in performance must be due to differences in usage options. With the equivalence in the temporal deconvolution demonstrated, it is reasonable to assume that additional regularization terms in the spatial or temporal domains would not modify this equivalence when convex operators are employed; e.g., when the regularization problem can be solved by means of the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) (Beck and Teboulle 2009) or the Generalized Forward-Backward Splitting (Raguet et al. 2013) techniques. Our findings are in line with the equivalence of analysis and synthesis methods in under-determined cases ($N \leq V$) as demonstrated in (Elad et al. 2007).

Taking into account the advantages and disadvantages of the presented techniques shown in Table 1, future work will improve and extend deconvolution methods for fMRI. For instance, the appropriate formulation depending on data acquisition (i.e., single-echo vs multi-echo) could be studied and compared with existing methods (Caballero-Gaudes et al. 2019), or formulations that account for HRF variability could be investigated too (Badillo et al. 2013; Gaudes et al. 2012; Farouj et al. 2019). Furthermore, robust methods to select the regularization parameter (Uruñuela et al. 2020; Meinshausen and Bühlmann 2010) and other potential $\ell_{p,q}$ -norm regularization terms (e.g., $p < 1$) or debiasing approaches could be explored.

Again, move up

Ola, what is V ? :-)
Move up in Section on analysis-synthesis equivalence

Folks, I think we need to polish the message for the reader further. If the methods are equivalent, either the reader has learned something about different types of optimization (but this is not yet well described in the paper), or there is a clear practical take-home message.

Paradigm Free Mapping	Total Activation
+ Its formulation can be extended straightforwardly for deconvolution of multiple signals with a common neuronal-related signal, e.g., for multi-echo formulations Caballero-Gaudes et al. 2019.	+ Both the spike and block models solve the regularization problem with the same HRF.
+ The model can implement any HRF shape very easily since it only requires the coefficients at the required temporal resolution.	-
-	-

Table 1: Advantages (+) and disadvantages (-) of Paradigm Free Mapping and Total Activation with respect to each other.

6. Code and data availability

The code and materials used in this work can be found in the following GitHub repository:
230 https://github.com/eurunuela/pfm_vs_ta. We encourage the reader to explore the parameters
(e.g. SNR, varying HRF options and mismatch between algorithms, TR, number of events, onsets,
and durations) in the provided Jupyter notebooks. Likewise, the data used to produce the figures
can be found in <https://osf.io/f3ryg/>.

7. Acknowledgements

This research was funded by the European Union's Horizon 2020 research and innovation
235 program (agreement No. 713673 of the Marie Skłodowska-Curie grant), La Caixa Foundation
(ID 100010434, fellowship code LCF/BQ/IN17/11620063), the Spanish Ministry of Economy and
Competitiveness (RYC-2017-21845), the Basque Government (BERC 2018-2021, PIBA_2019_104,
PRE_2019_1_005 4), and the Spanish Ministry of Science, Innovation and Universities (PID2019-
105520GB-100).

240 References

- Allan, T.W., Francis, S.T., Caballero-Gaudes, C., Morris, P.G., Liddle, E.B., Liddle, P.F., Brookes, M.J., Gowland, P.A., 2015. Functional connectivity in mri is driven by spontaneous bold events. *PloS one* 10, e0124577.
- Badillo, S., Vincent, T., Ciuci, P., 2013. Group-level impacts of within-and between-subject hemodynamic variability in fMRI. *Neuroimage* 82, 433–448.
- Beck, A., Teboulle, M., 2009. A fast iterative shrinkage-thresholding algorithm. *Society for Industrial and Applied Mathematics Journal on Imaging Sciences* 2, 183–202. doi:10.1137/080716542.
- Betzel, R.F., Byrge, L., Esfahlani, F.Z., Kennedy, D.P., 2020. Temporal fluctuations in the brain's modular architecture during movie-watching. *NeuroImage* 213, 116687. doi:10.1016/j.neuroimage.2020.116687.
- 250 Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *Neuroimage* 31, 1536–1548.
- Bommarito, G., Tarun, A., Farouj, Y., Preti, M.G., Petracca, M., Droby, A., El Mendili, M.M., Inglese, M., Van De Ville, D., 2020. Functional network dynamics in progressive multiple sclerosis. *medRxiv* .
- Boynton, G.M., Engel, S.A., Glover, G.H., Heeger, D.J., 1996. Linear systems analysis of functional magnetic resonance imaging in human v1. *Journal of Neuroscience* 16, 4207–4221.
- Bruckstein, A.M., Donoho, D.L., Elad, M., 2009. From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images. *SIAM Review* 51, 34–81. doi:10.1137/060657704.
- Caballero-Gaudes, C., Moia, S., Panwar, P., Bandettini, P.A., Gonzalez-Castillo, J., 2019. A deconvolution algorithm for multi-echo functional mri: Multi-echo sparse paradigm free mapping. *NeuroImage* 202, 116081.

- 265 Caballero Gaudes, C., Petridou, N., Francis, S.T., Dryden, I.L., Gowland, P.A., 2013. Paradigm free
mapping with sparse regression automatically detects single-trial functional magnetic resonance
imaging blood oxygenation level dependent responses. *Human Brain Mapping* doi:10.1002/hbm.
21452.
- 270 Cherkaoui, H., Moreau, T., Halimi, A., Ciuciu, P., 2019. Sparsity-based Blind Deconvolution of
Neural Activation Signal in fMRI, in: ICASSP 2019 - 2019 IEEE International Conference on
Acoustics, Speech and Signal Processing (ICASSP), pp. 1323–1327. doi:10.1109/ICASSP.2019.
8683358.
- 275 Cifre, I., Flores, M.T.M., Ochab, J.K., Chialvo, D.R., 2020a. Revisiting non-linear functional brain
co-activations: directed, dynamic and delayed. *arXiv preprint arXiv:2007.15728*.
- Cifre, I., Zarepour, M., Horovitz, S., Cannas, S., Chialvo, D., 2020b. Further results on why a
280 point process is effective for estimating correlation between brain regions. *Papers in Physics* 12,
120003–120003.
- Cohen, M.S., 1997. Parametric analysis of fMRI data using linear systems methods. *Neuroimage*
6, 93–103.
- 285 Cox, R., 1996. AFNI: Software for analysis and visualization of functional magnetic resonance
neuroimages. *Computers and Biomedical Research* 29, 162–173.
- Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least Angle Regression. *The Annals of
Statistics* 32, 407–499. doi:10.1214/009053604000000067.
- Elad, M., Milanfar, P., Rubinstein, R., 2007. Analysis versus synthesis in signal priors. *Inverse
problems* 23, 947.
- 290 Esfahlani, F.Z., Jo, Y., Faskowitz, J., Byrge, L., Kennedy, D.P., Sporns, O., Betzel, R.F., 2020.
High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proceedings of
the National Academy of Sciences* 117, 28393–28401.
- Farouj, Y., Karahanoğlu, F.I., Van De Ville, D., 2019. Bold signal deconvolution under uncer-
tain haemodynamics: A semi-blind approach, in: 2019 IEEE 16th International Symposium on
295 Biomedical Imaging (ISBI 2019), IEEE. pp. 1792–1796.
- Faskowitz, J., Esfahlani, F.Z., Jo, Y., Sporns, O., Betzel, R.F., 2020. Edge-centric functional
network representations of human cerebral cortex reveal overlapping system-level architecture.
Technical Report. Nature Publishing Group.
- 295 Faskowitz, J., Tanner, J.C., Mišić, B., Betzel, R.F., 2021. An Edge-Centric Model for Harmonizing
Multi-Relational Network Datasets. Preprint. *Neuroscience*. doi:10.1101/2021.01.07.425450.
- Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M., Turner, R., 1998. Event-related
fMRI: characterizing differential responses. *Neuroimage* 7, 30–40.
- 300 Friston, K.J., Jezzard, P., Turner, R., 1994. Analysis of functional MRI time-series. *Human brain
mapping* 1, 153–171.

- ³⁰⁰ Gaudes, C.C., Karahanoğlu, F.I., Lazeyras, F., Van De Ville, D., 2012. Structured sparse deconvolution for paradigm free mapping of functional MRI data, in: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 322–325.
- Gaudes, C.C., Petridou, N., Dryden, I.L., Bai, L., Francis, S.T., Gowland, P.A., 2011. Detection and characterization of single-trial fMRI bold responses: Paradigm free mapping. Human Brain Mapping doi:10.1002/hbm.21116.
- ³⁰⁵ Gitelman, D.R., Penny, W.D., Ashburner, J., Friston, K.J., 2003. Modeling regional and psychophysiological interactions in fMRI: The importance of hemodynamic deconvolution. NeuroImage 19, 200–207. doi:10.1016/S1053-8119(03)00058-2.
- Glover, G.H., 1999. Deconvolution of impulse response in event-related bold fMRI. Neuroimage 9, 416–429.
- ³¹⁰ Gonzalez-Castillo, J., Caballero-Gaudes, C., Topolski, N., Handwerker, D.A., Pereira, F., Bandettini, P.A., 2019. Imaging the spontaneous flow of thought: Distinct periods of cognition contribute to dynamic functional connectivity during rest. NeuroImage 202, 116129. doi:10.1016/j.neuroimage.2019.116129.
- ³¹⁵ Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. Fsl. Neuroimage 62, 782–790.
- Karahanoğlu, F.I., Bayram, İ., Ville, D.V.D., 2011. A Signal Processing Approach to Generalized 1-D Total Variation. IEEE Transactions on Signal Processing 59, 5265–5274. doi:10.1109/TSP.2011.2164399.
- ³²⁰ Karahanoğlu, F.I., Caballero-Gaudes, C., Lazeyras, F., Van De Ville, D., 2013. Total activation: fMRI deconvolution through spatio-temporal regularization. NeuroImage doi:10.1016/j.neuroimage.2013.01.067.
- ³²⁵ Karahanoğlu, F.I., Van De Ville, D., 2015. Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. Nature Communications 6, 7751. doi:10.1038/ncomms8751.
- Karahanoğlu, F.I., Van De Ville, D., 2017. Dynamics of large-scale fMRI networks: Deconstruct brain activity to build better models of brain function. Current Opinion in Biomedical Engineering 3, 28–36. doi:10.1016/j.cobme.2017.09.008.
- ³³⁰ Kinany, N., Pirondini, E., Micera, S., Van De Ville, D., 2020. Dynamic Functional Connectivity of Resting-State Spinal Cord fMRI Reveals Fine-Grained Intrinsic Architecture. Neuron 108, 424–435.e4. doi:10.1016/j.neuron.2020.07.024.
- Liu, X., Chang, C., Duyn, J.H., 2013. Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns. Frontiers in systems neuroscience 7, 101.
- ³³⁵ Liu, X., Duyn, J.H., 2013. Time-varying functional network information extracted from brief instances of spontaneous brain activity. Proceedings of the National Academy of Sciences 110, 4392–4397.

- Liu, X., Zhang, N., Chang, C., Duyn, J.H., 2018. Co-activation patterns in resting-state fMRI signals. *Neuroimage* 180, 485–494.
- Lopes, R., Lina, J.M., Fahoum, F., Gotman, J., 2012. Detection of epileptic activity in fMRI without recording the eeg. *Neuroimage* 60, 1867–1879.
- Majeed, W., Magnuson, M., Hasenkamp, W., Schwarb, H., Schumacher, E.H., Barsalou, L., Keilholz, S.D., 2011. Spatiotemporal dynamics of low frequency bold fluctuations in rats and humans. *Neuroimage* 54, 1140–1150.
- Majeed, W., Magnuson, M., Keilholz, S.D., 2009. Spatiotemporal dynamics of low frequency fluctuations in bold fMRI of the rat. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 30, 384–393.
- Meinshausen, N., Bühlmann, P., 2010. Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72, 417–473.
- Petridou, N., Gaudes, C.C., Dryden, I.L., Francis, S.T., Gowland, P.A., 2013. Periods of rest in fMRI contain individual spontaneous events which are related to slowly fluctuating spontaneous activity. *Human Brain Mapping* 34, 1319–1329. doi:10.1002/hbm.21513.
- Raguet, H., Fadili, J., Peyré, G., 2013. A Generalized Forward-Backward Splitting. *SIAM Journal on Imaging Sciences* 6, 1199–1226. doi:10.1137/120872802.
- Schwarz, G., 1978. Estimating the Dimension of a Model. *Annals of Statistics* 6, 461–464. doi:10.1214/aos/1176344136.
- Shine, J.M., Bissett, P.G., Bell, P.T., Koyejo, O., Balsters, J.H., Gorgolewski, K.J., Moodie, C.A., Poldrack, R.A., 2016. The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron* 92, 544–554.
- Shine, J.M., Koyejo, O., Bell, P.T., Gorgolewski, K.J., Gilat, M., Poldrack, R.A., 2015. Estimation of dynamic functional connectivity using multiplication of temporal derivatives. *NeuroImage* 122, 399–407.
- Shmueli, K., van Gelderen, P., de Zwart, J.A., Horovitz, S.G., Fukunaga, M., Jansma, J.M., Duyn, J.H., 2007. Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI bold signal. *Neuroimage* 38, 306–320.
- Tarun, A., Wainstein-Andriano, D., Sterpenich, V., Bayer, L., Perogamvros, L., Solms, M., Axmacher, N., Schwartz, S., Van De Ville, D., 2020. Nrem sleep stages specifically alter dynamical integration of large-scale brain networks. *Iscience* 24, 101923.
- Tibshirani, R., 1996. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 58, 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x.
- Triantafyllou, C., Hoge, R.D., Krueger, G., Wiggins, C.J., Potthast, A., Wiggins, G.C., Wald, L.L., 2005. Comparison of physiological noise at 1.5 t, 3 t and 7 t and optimization of fMRI acquisition parameters. *Neuroimage* 26, 243–250.

- Uruñuela, E., Jones, S., Crawford, A., Shin, W., Oh, S., Lowe, M., Caballero-Gaudes, C., 2020.
375 Stability-Based Sparse Paradigm Free Mapping Algorithm for Deconvolution of Functional MRI
Data. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine
and Biology Society, EMBS 2020-July, 1092–1095. doi:10.1109/EMBC44109.2020.9176137.
- Zhang, X., Pan, W.J., Keilholz, S.D., 2020. The relationship between bold and neural activity
arises from temporally sparse events. *Neuroimage* 207, 116390.
- 380 Zöller, D., Sandini, C., Karahanoğlu, F.I., Padula, M.C., Schaer, M., Eliez, S., Van De Ville, D.,
2019. Large-scale brain network dynamics provide a measure of psychosis and anxiety in 22q11.2
deletion syndrome. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 4, 881–892.

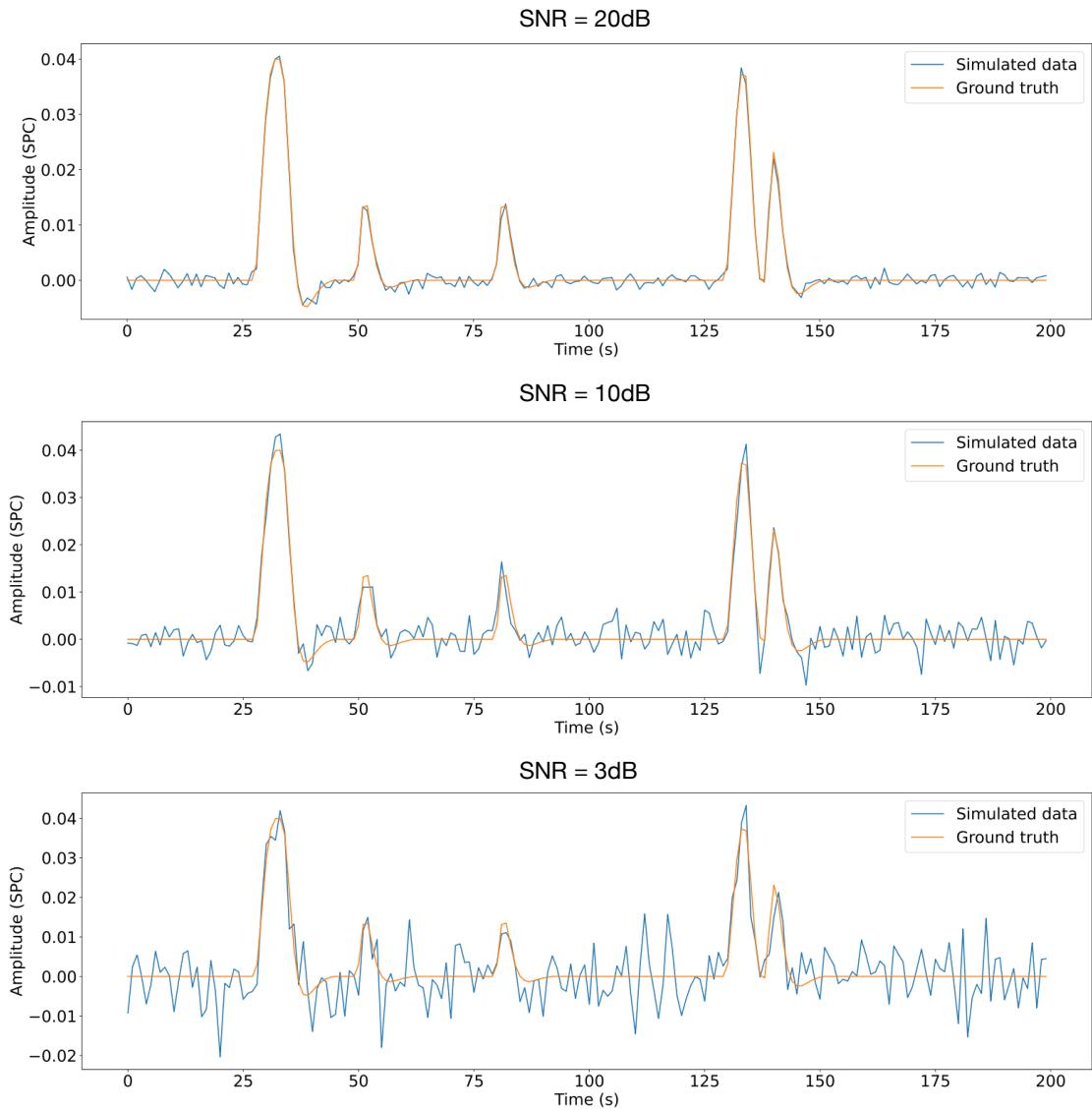


Figure 1: Simulated signal with different SNRs (20 dB, 10 dB and 3 dB)

Why not combine this in a single plot with different lines/colors for the SNR levels?

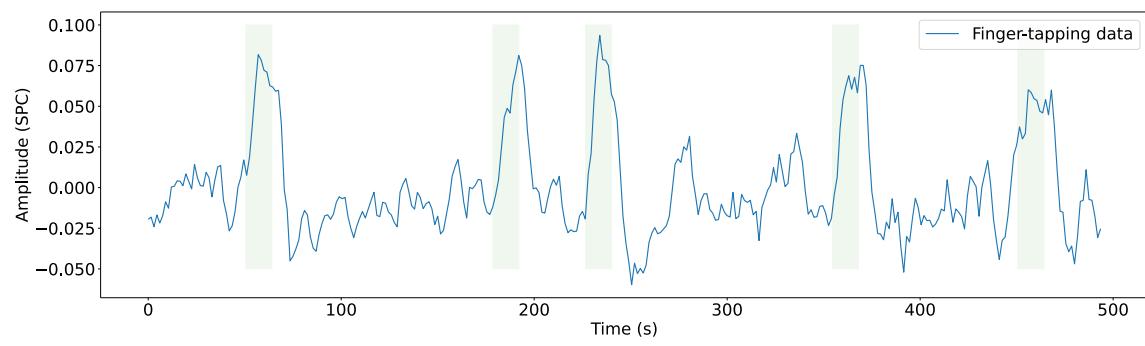


Figure 2: Most representative voxel of the finger-tapping task. Green blocks indicate the onsets and the duration of it.

This image does not look complete here?

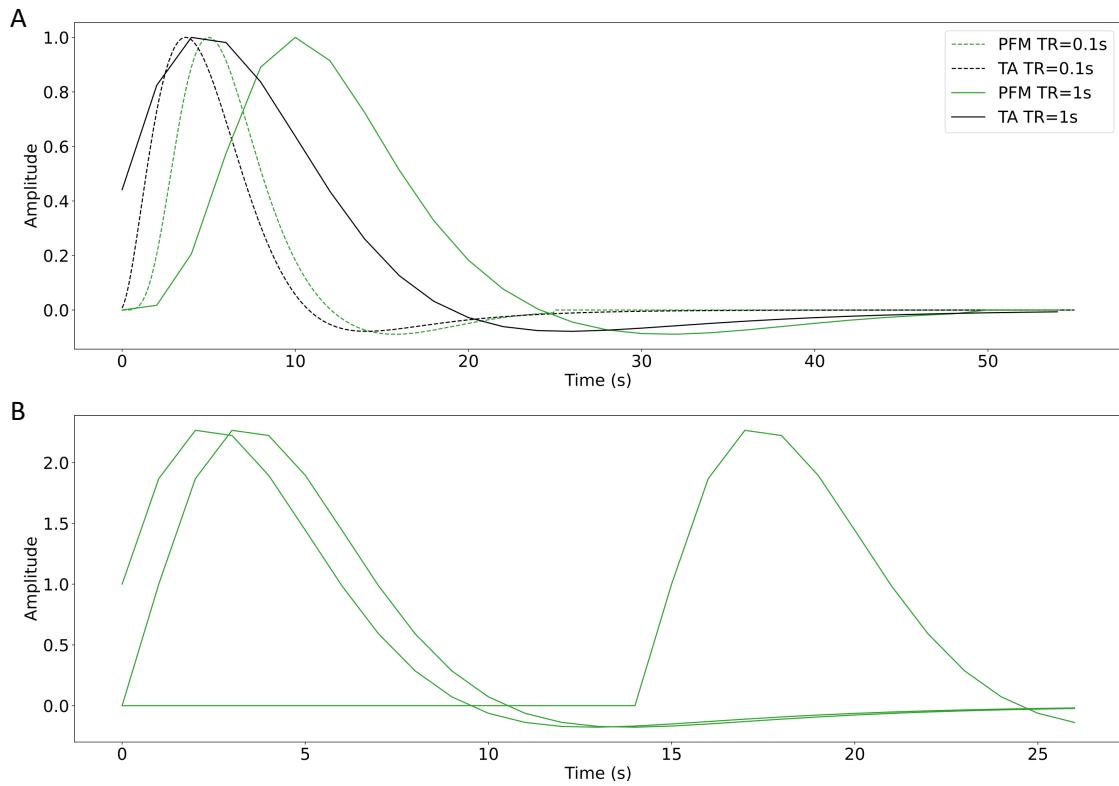


Figure 3: A) Canonical HRF models typically used by PFM (green) and TA (black) at TR = 0.1 s (dashed lines) and TR = 1 s (solid lines). Without loss of generality, the waveforms are scaled to unit amplitude for visualization. B) Representation of three shifted HRFs at TR=1 s (onsets=0, 1, and 15 s) that build the design matrix for PFM when the HRF model has been matched to that in TA.

If I understand well, you plot them as function of index? But I would map them on top of one another as function of time...

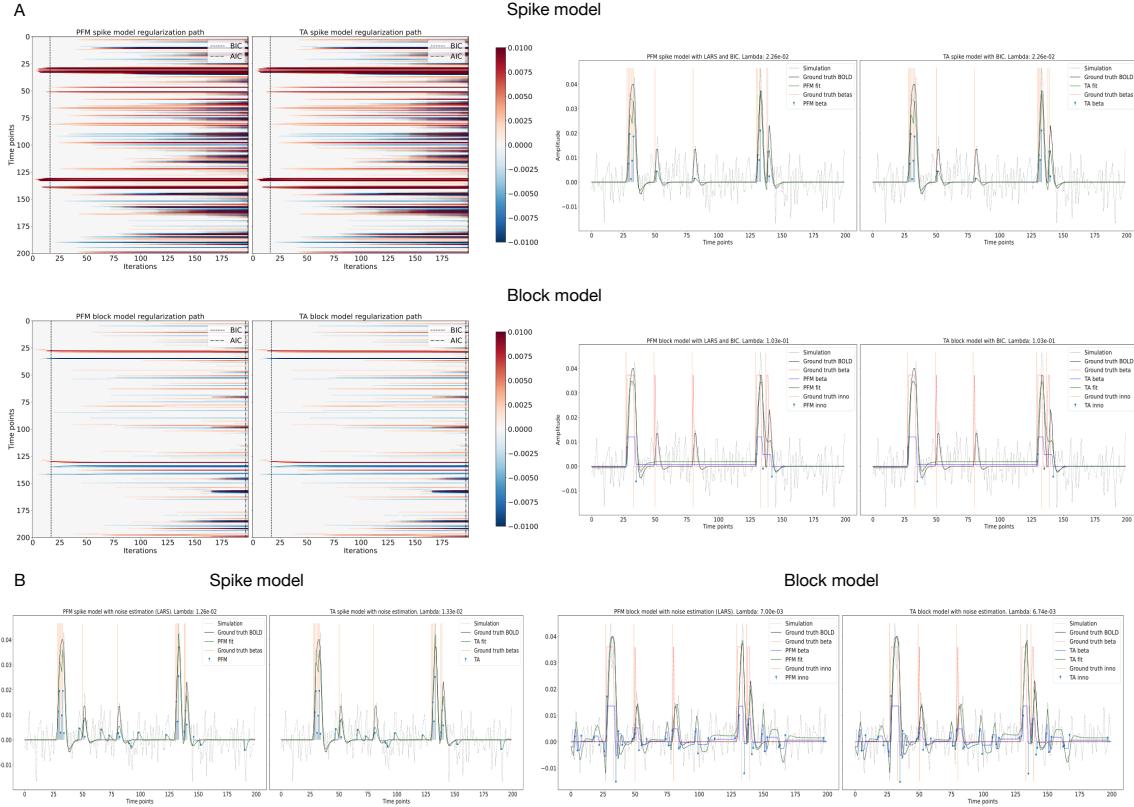


Figure 4: A) (Left) Heatmap of the regularization paths of the activity-inducing (top) and innovation (bottom) signals estimated with PFM and TA as a function of λ for the simulated data with SNR = 3 dB (x-axis: increasing number of iterations or λ as given by LARS; y-axis: time; color: amplitude). Vertical lines denote iterations corresponding to the Akaike and Bayesian Information Criteria (AIC and BIC) optima. (Right) Estimated activity-inducing (blue) and activity-related (green) signals λ is selected based on BIC. B) Estimated activity-inducing, innovation and activity-related (fit, \mathbf{x}) signals when λ is selected based on the MAD method with the spike model (left, with PFM on the left and TA on the right) and block model (right, with PFM on the left and TA on the right) for the simulated data with SNR = 3 dB.

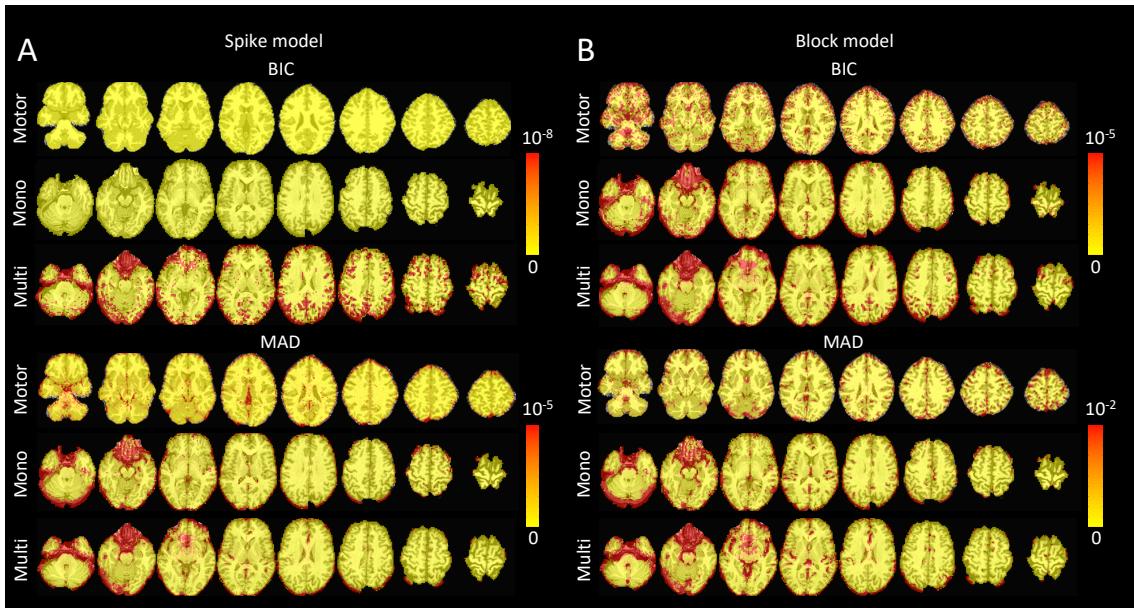


Figure 5: Sum of squares of the differences of the activity-inducing signals estimated with Paradigm Free Mapping and Total activation for the different selections of the regularization parameter: BIC (top), and MAD (bottom). The sum of square difference maps are shown for the three experimental datasets introduced in Section 3: the motor task (Motor), the monoband resting-state (Mono), and the multiband resting-state (Multi) datasets. A) Sum of squares of the differences when using the spike model. B) Sum of squares of the differences when using the block model.

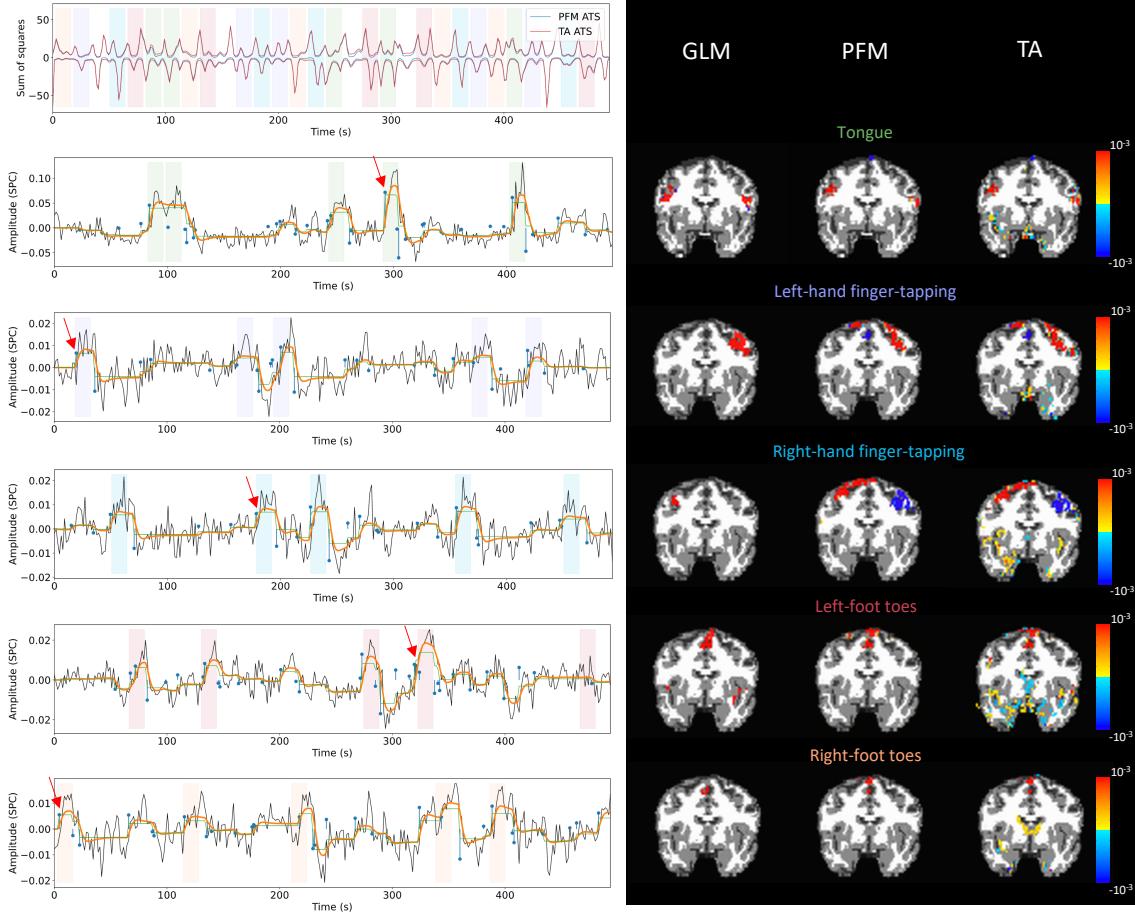


Figure 6: Row 1: Activation time-series of the innovation signals estimated by PFM (in blue) or TA (in red) calculated as the sum of squares at every time-frame. Positive-valued and negative-valued contributions were separated into two distinct timecourses. Color-bands indicate the onset and duration of each condition in the task (green: tongue, purple: left-hand finger-tapping, blue: right-hand finger-tapping, red: left-foot toes, orange: right-foot toes). Rows 2-6: time-series of a representative voxel for each task with the PFM-estimated innovation (blue), PFM-estimated activity-inducing (green), and activity-related (i.e., fitted, orange) signals, with their corresponding GLM, PFM, and TA maps on the right. The maps shown on the right are sampled at the time-point labeled with the red arrows and display the innovation signals at that moment across the whole brain.

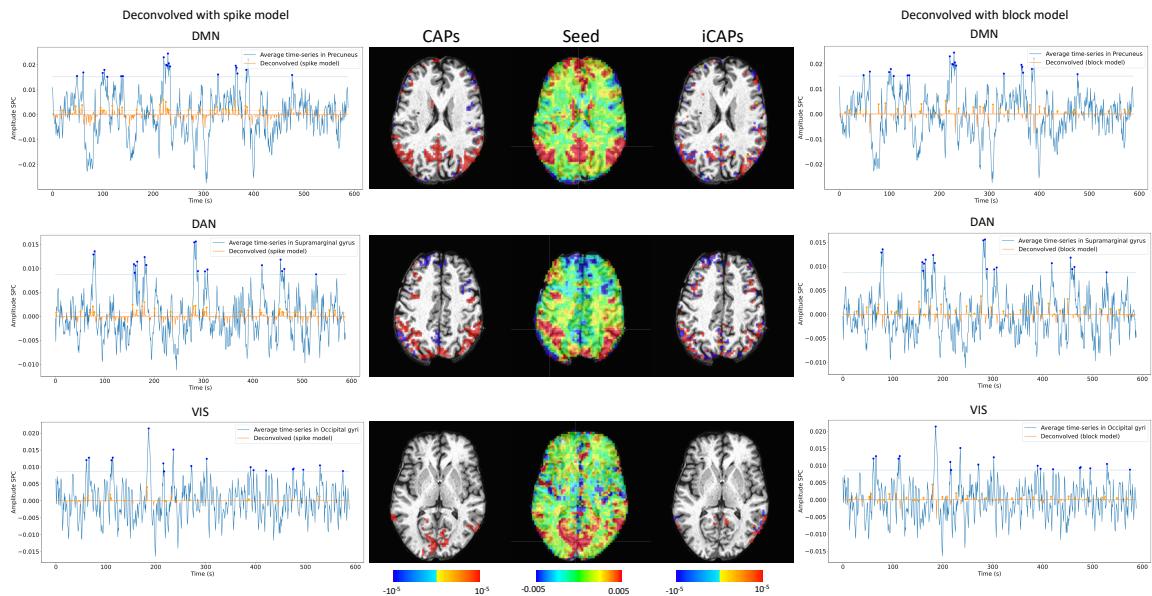


Figure 7: CAPs (left) and iCAPs (right) obtained with the PFM-estimated activity-inducing and innovation signals, respectively. Time-points selected with a 95th percentile threshold are shown over the average time-series (blue) in the seed region (white-cross) and the deconvolved signal (orange). CAPs and iCAPs maps of the seed and deconvolved signals obtained by averaging the selected time-points are illustrated in the center.