Paradigm Free Mapping vs Total Activation

Eneko Uruñuela[a,b,*], Thomas A.W. Bolton[c], 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, Donostia-San Sebastián, Spain.*
[c] *Department of Decoded Neurofeedback, ATR Computational Neuroscience Laboratories, Kyoto, Japan*
[d] *Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland.*
[e] *Faculty of Medicine of the University of Geneva, Geneva, Switzerland*

Abstract

Here's where the fantastic abstract will go.

Keywords: fMRI deconvolution, paradigm free mapping, total activation

1. Introduction

- Talk about our motivation for this paper.
- We could mention iCAPs Neuron, and papers with applications like PFM, TA, clinical patient papers with iCAPs.
- Apart from [[Richard F. Betzel]]'s work [1, 2, 3], we could mention the connection with the [[Multiplication of Temporal Derivatives]] method [4, 5].
  - These are basically calculating the derivative, which is the same as applying a high-pass filter and calculating the correlation.

There is an increasing interest in methods that aim to recover the underlying neuronal activity from functional magnetic resonance imaging (fMRI) data with no prior information of the timing of the blood oxygenation level-dependent (BOLD) events. One of such techniques is deconvolution, which does not consider task-related stimulus functions or any other specific cause of the underlying neuronal activity. In other words, deconvolution methods are capable of blindly estimating the neuronal activity, which makes them especially attractive for exploring time-varying activity of resting-state fluctuations [6, 7, 8, 9, 10], naturalistic paradigms [1], or clinical conditions such as the study of interictal events in epilepsy.

Paradigm Free Mapping (PFM) [11] — which is available as *3dPFM* in AFNI — and Total Activation (TA) [12] are two of such deconvolution algorithms.

This note comprises three sections. In the first, we present the theory behind the Paradigm Free Mapping and Total Activation deconvolution algorithms. We then assess their performance using the same hemodynamic response function with different criteria for the selection of the regularization parameter: a) a selection based on the Bayesian (BIC) [13] and Akaike Information Criterion (AIC) [14], and b) a selection based on the estimated standard deviation of the noise in the data. We report that both methods produce identical results when estimating the underlying activity-inducing and innovation signals in different signal-to-noise ratio (SNR) settings. In the final section, we discuss the pros and cons of each of the described techniques and conclude with future steps.

2. Theory

- What is deconvolution and different formulations presented as a review.
- Analysis vs synthesis
  - TA paper but without the spatial regularization
  - PFM paper
  - In Gitelman it's an **H** multiplied by a Fourier term.
- Spikes and block models

The hemodynamic response to neuronal activity at time t can be modeled as the convolution with a finite impulse response function of the neuronal signal $s_{t-\tau}$ at time $t-\tau$ with the hemodynamic response function h_τ [15]:

$$y_t = \sum_t h_\tau s_{t-\tau}, \tag{1}$$

where y_t is the measured BOLD signal on a given voxel. This equation can be reformulated in matrix notation as $\mathbf{y} = \mathbf{Hs}$ where $\mathbf{H} \in \mathbb{R}^{NxN}$ is the HRF in Toeplitz matrix form, and N is the number of frames of the fMRI acquisition.

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

Functional MRI data analyses are often directed to disentangling and understanding the neural processes that occur among brain regions. However, interactions in the brain are expressed, not at the level of hemodynamic responses, but at the neural level. Thus, an intermediate step that estimates the underlying neuronal activity is necessary for such analyses. Given the nature of the fMRI BOLD signal, the appropriate approximation of the neuronal activity can be obtained by means of deconvolution with an assumed hemodynamic response [15]. Hence, the maximum likelihood estimate of the hemodynamic response to the underlying neural activity can be calculated using the ordinary least-squares estimator that minimizes the residual sum of squares between the modeled (**Hs**) and measured (**y**) signals. Yet, the estimates of the neuronal activity **s** must be constrained with a regularization term to attenuate the collinearity and high variability of the design matrix **H**.

2.1. Paradigm Free Mapping

Paradigm Free Mapping (PFM) builds upon the signal model introduced in (1); i.e., the BOLD signal is the result of convolving the underlying neural activity with the hemodynamic response, and proposes to estimate the activity-inducing signal by solving the following regularization problem [16, 11, 17]:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{s}\|_F^2 + \Omega(\mathbf{s}) \tag{2}$$

where $\Omega(\mathbf{s})$ is the regularization term.

Assuming that single-trial BOLD responses are the result of brief bursts of neuronal activation, the activity-inducing signal **s** must be a sparse vector. Thus, sparse estimates of **s** could be obtained by substituting $\Omega(\mathbf{s})$ in (3) with an l_0-norm and solving the optimization problem [18]. 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 the design matrix **H**. Hence, a pragmatic solution is to solve the optimization problem with the use of an l_1-norm, or LASSO [19], which is a convex function and therefore provides fast convergence to the optimal solution.

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{s}\|_F^2 + \lambda\|\mathbf{s}\|_1 \tag{3}$$

where λ regulates how sparse the optimal solution is.

Such formulation provides flexibility to expand the capabilities of PFM. For instance, incorporating the integration operator **L** into the design matrix **H** allows the recovery of the innovation signal **u**; i.e., the derivative of the activity-inducing signal **s**. Therefore, the innovation signal can be estimated by solving the following optimization problem [17, 20]:

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{L}\mathbf{u}\|_F^2 + \lambda\|\mathbf{u}\|_1 \tag{4}$$

2.2. Total Activation

Even though based on the same signal model as PFM, Total Activation (TA) proposes to use a linear differential operator L_h that inverts the hemodynamic system based on activelets to recover the activity-inducing signal **s** [21, 12]:

$$L_h\{x\}(t) = s(t) \tag{5}$$

where x is the neuronal-related signal; i.e., the activity inducing signal **s** convolved with the HRF, and L_h is defined as

$$L_h = \prod_{i=1}^{M_1}(D - \alpha_i I)(\prod_{j=1}^{M_2}(D - \gamma_j I))^{-1} \tag{6}$$

where D is the derivative operator, $\alpha_i (i = 1, \ldots, M_1)$ define the zeros of the filter, $\gamma_j (j = 1, \ldots, M_2)$ represent the poles, I is the identity matrix and $M_1 > M_2$. Given the relationship between the activity-inducing and the innovation signal, the latter can be recovered as:

$$L\{x\}(t) = D\{s\}(t) = u(t) \tag{7}$$

where $L = DL_h$ and D is the derivative.

Therefore, for a given voxel, the neuronal-related signal could be estimated by solving the following regularized least-squares problem:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_F^2 + \mathcal{R}(\mathbf{x}) \tag{8}$$

where **y** is the fMRI data and $\mathcal{R}(\mathbf{x})$ is the following l_1-norm regularization term:

$$\mathcal{R}(\mathbf{x}) = \lambda \sum_{t=1}^{N} \| \Delta_L \{\mathbf{x}\}\| \tag{9}$$

where λ is the regularization parameter.

3. Results

- Methods on how we're doing simulations and results (with simulations and experimental data)
  - Different SNRs and maybe even use CAPs
  - Selection of HRF explained if both use the same but it's different from what's used for simulating.
    - What happens? For example with gamma for simulating.
  - Selection of regularization parameter
    - Present with real data on a voxel

A critical decision with deconvolution methods is the selection of the regularization parameter λ, for which many techniques have been proposed in the literature; but an optimal is yet to be discovered. In fact, Paradigm Free Mapping and Total Activation base their selection of the regularization parameter on different criteria: the Bayesian

Information Criterion (BIC) [13] and Akaike Information Criterion (AIC) [14], and a selection based on the convergence of the residuals to a pre-estimated level of the noise respectively. Hence, we compare the performance of the two algorithms with both selection criteria. Furthermore, we explore the differences between the techniques in terms of the estimation of the activity-inducing signal $\mathbf{s}$ using the *spike model* in (3) and the innovation signal $\mathbf{u}$ using the *block model* in (4).

3.1. Simulated and experimental data

In order to compare the two methods while controlling for their correct performance, we simulated a 400 seconds (TR = 2 s) time series with five neuronal events 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.

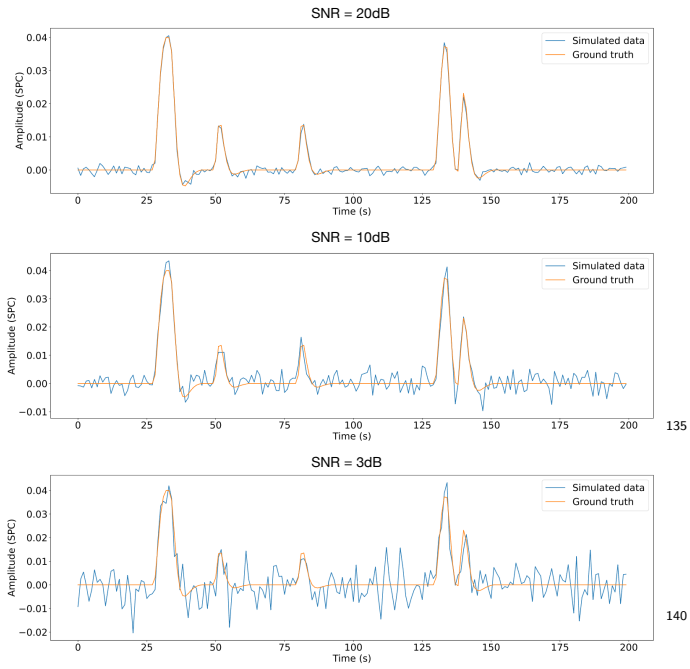

Figure 1: Simulated signal with different SNRs (20 dB, 10 dB and 3 dB).

Furthermore, we compared the two techniques on a finger-tapping task where the ground-truth is unknown. For this comparison, we selected a voxel that best represented the finger-tapping paradigm described in [17] as shown in Figure 2.

3.2. Selection of the hemodynamic response function

With the aim of making a fair comparison of the two methods, we first compared their hemodynamic response functions. Figure 3 shows the difference in the hemodynamic response function that PFM and TA use by default; the SPMG1 and the HRF resulting from the linear

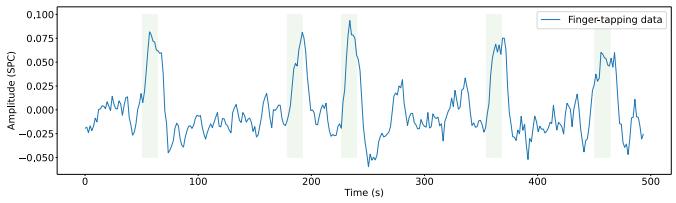

Figure 2: Most representative voxel of the finger-tapping task. Green blocks indicate the onsets and the duration of it.

differential operator respectively. A clear difference is observable in that the PFM hemodynamic response function begins at zero while the TA HRF starts at 1. Hence, the Total Activation HRF starts close to its peak, which is advanced around 2.5 frames with respect to PFM. Another difference worth mentioning is that PFM normalizes its HRF to a peak amplitude of 1, whereas the TA HRF is not normalized.

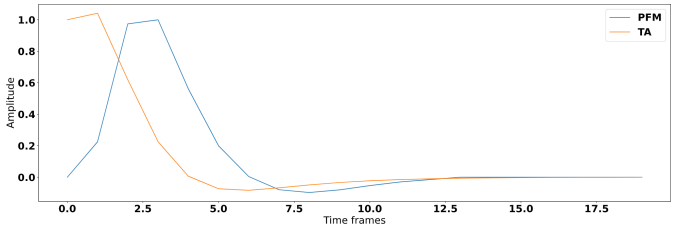

Figure 3: Diffence in the HRF of PFM (blue) and TA (orange).

While Paradigm Free Mapping allows for the use of any hemodynamic response function — 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.

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

Total Activation proposes to solve the inverse problem by updating the regularization parameter λ on every iteration n so that the residuals converged to a previously estimated noise level of the data fit $\tilde{\sigma}$, where the pre-estimated noise level is calculated from the median absolute deviation of fine-scale wavelet coefficients (Daubechies, order 3) [12]:

$$\lambda^{n+1} = \frac{N\tilde{\sigma}}{\frac{1}{2}\|\mathbf{y} - \mathbf{x}^n\|_F^2}\lambda^n \tag{10}$$

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

Paradigm Free Mapping bases its selection of the regularization parameter on the Bayesian Information Criterion (BIC) [13] and the Akaike Information Criterion (AIC) [14]. Hence, we calculated the entire regularization

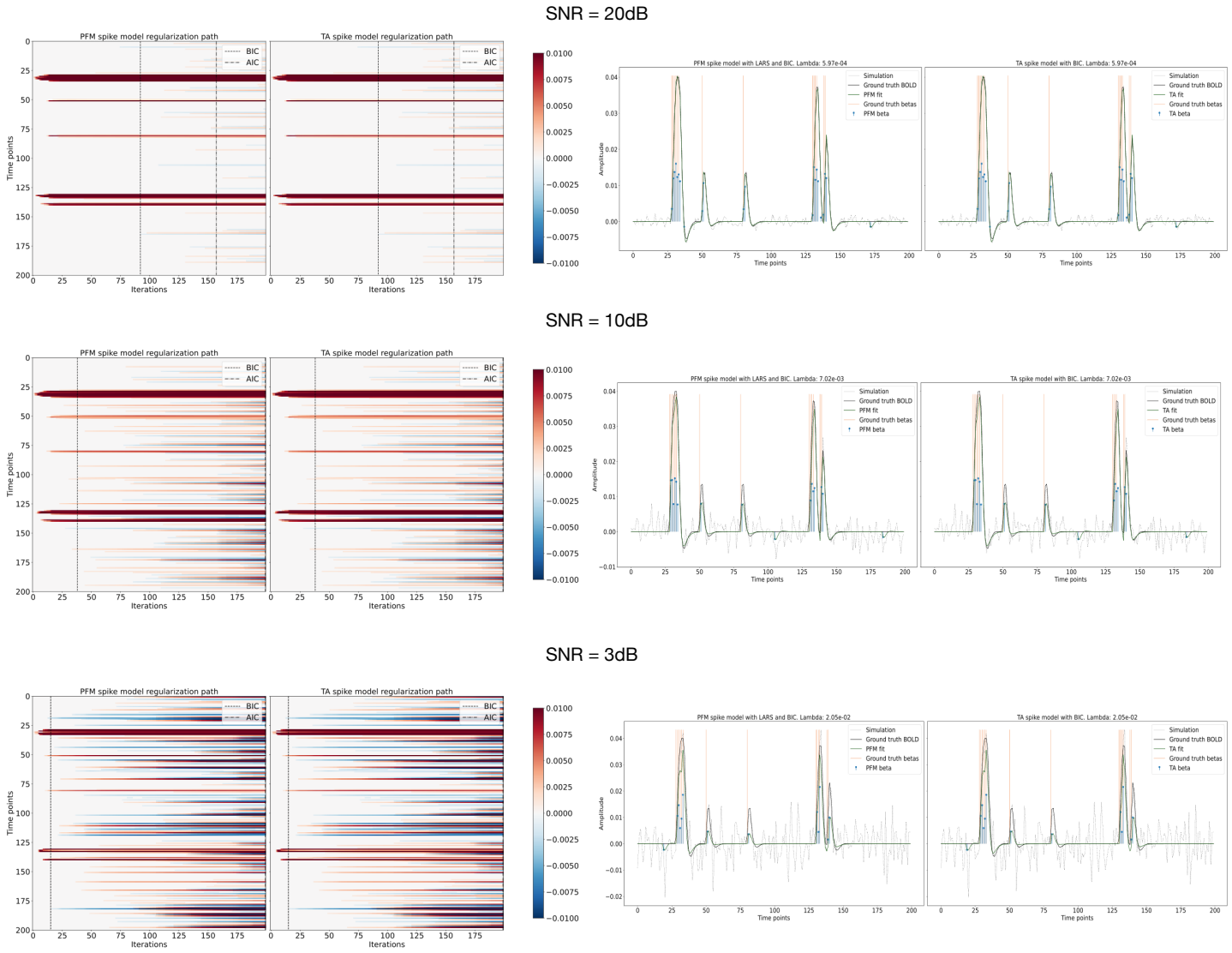

Figure 4: Regularization path and estimated activity-inducing and neuronal-related signals with the spike model and a selection of λ based on BIC. Each row of plots corresponds to a different SNR condition.

path with PFM by means of the least angle regression (LARS) algorithm [22] and used the λ in the path to solve the deconvolution problem with Total Activation.

Figure 4 (left) shows the regularization path of PFM and TA side by side for the three SNR conditions for the spike model; i.e., the inverse problem described in (3). Each iteration of LARS reduces the value of λ; i.e., reduces the sparsity promoted by the l_1-norm, and reveals a new non-zero coefficient as shown in the x axis of the heatmaps. Vertical black lines depict the selection of the regularization parameter based on BIC and AIC, and thus, the colored coefficients to the left of the vertical lines depict the estimated activity-inducing signal $s(t)$. Figure 4 (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. Thus, Figure 4 demonstrates that, regardless of the 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 $\mathbf{s}$ and neuronal-related signal $\mathbf{x}$.

The regularization path to estimate innovation signals; i.e., solving the optimization problem using the block model in (4), yields identical results for both PFM and TA methods as shown in Figure 5 (left). Again, the BIC-based selection of λ is identical for both PFM and TA and the estimation of the innovation signal $\mathbf{u}$ shows no distinguishable differences between the algorithms (see Figure 5 right).

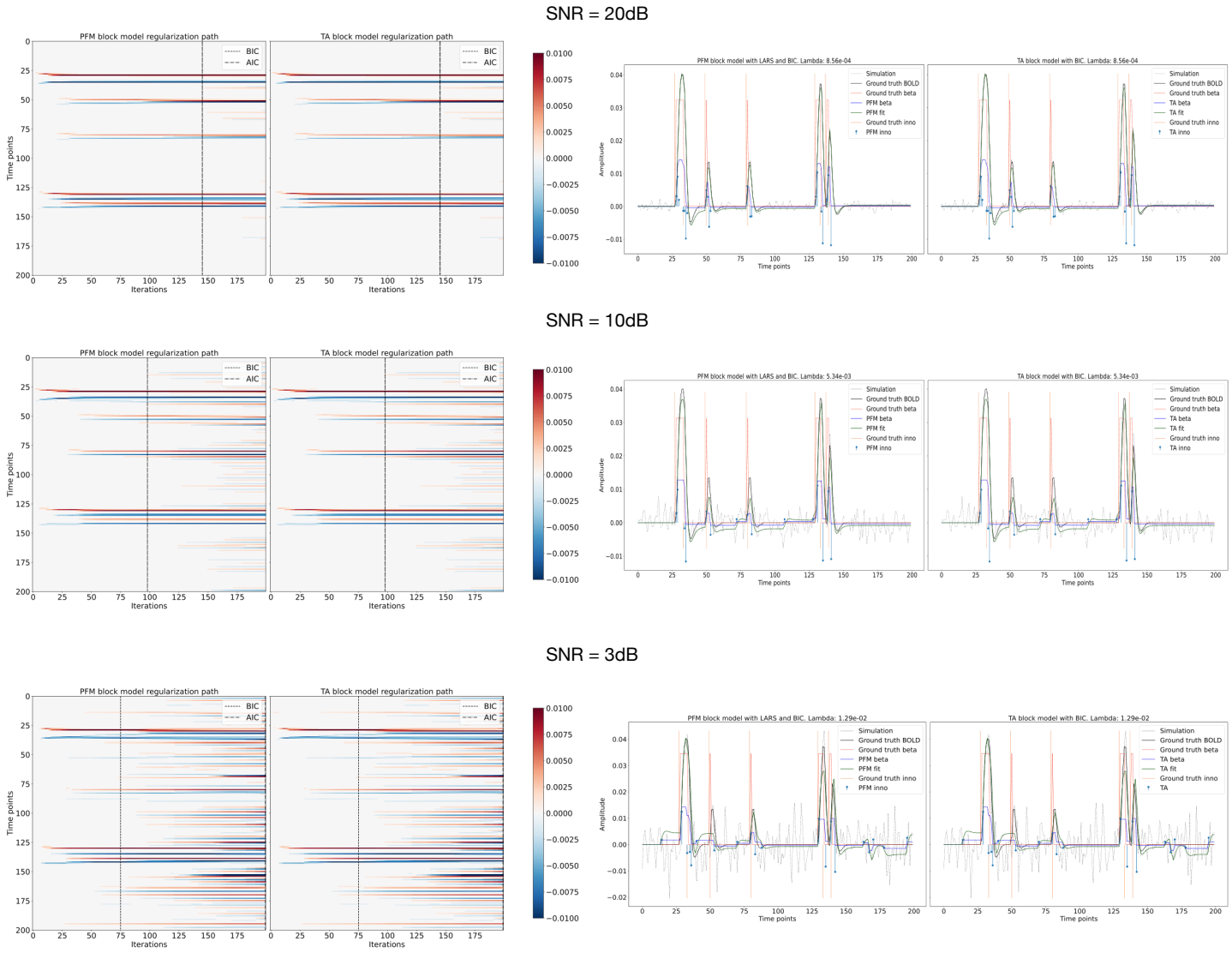

Figure 5: Regularization path and estimated activity-inducing and neuronal-related signals with the block model and a selection of λ based on BIC. Each row of plots corresponds to a different SNR condition.

Therefore, both Paradigm Free Mapping and Total Activation yield identical regularization paths and estimates of the innovation signal regardless of the SNR condition when applying the same HRF and regularization parameters with the block model.

4. Discussion

- Pros and cons of each formulation: analysis vs synthesis
- Link with other approaches
- Finish with conclusions and a moving forward
  - We have to refine the deconvolution
  - HRF variability there are three: conference proceeding by Philippe [23], ISBI 2012 by César [24], and Farouj with a different formulation. Say conceptual differences among those.
  - Mention stability-selection [25]
  - Debiasing
  - Connected to debiasing other deconvolution algorithms that are based on a norm lower than 1.

References

[1] R. F. Betzel, L. Byrge, F. Z. Esfahlani, D. P. Kennedy, Temporal fluctuations in the brain's modular architecture during movie-watching, NeuroImage (2020) 116687.

[2] F. Z. Esfahlani, Y. Jo, J. Faskowitz, L. Byrge, D. Kennedy, O. Sporns, R. Betzel, High-amplitude co-fluctuations in cortical activity drive functional connectivity, bioRxiv (2020) 800045.

[3] J. Faskowitz, F. Z. Esfahlani, Y. Jo, O. Sporns, R. F. Betzel, Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture, Technical Report, Nature Publishing Group, 2020.

[4] J. M. Shine, O. Koyejo, P. T. Bell, K. J. Gorgolewski, M. Gilat, R. A. Poldrack, Estimation of dynamic functional connectivity using multiplication of temporal derivatives, NeuroImage 122 (2015) 399–407.

[5] J. M. Shine, P. G. Bissett, P. T. Bell, O. Koyejo, J. H. Balsters, K. J. Gorgolewski, C. A. Moodie, R. A. Poldrack, The dynamics of functional brain networks: integrated network states during cognitive task performance, Neuron 92 (2016) 544–554.

[6] N. Petridou, C. C. Gaudes, I. L. Dryden, S. T. Francis, P. A. Gowland, Periods of rest in fmri contain individual spontaneous events which are related to slowly fluctuating spontaneous activity, Human brain mapping 34 (2013) 1319–1329.

[7] F. I. Karahanoğlu, D. Van De Ville, Transient brain activity disentangles fmri resting-state dynamics in terms of spatially and temporally overlapping networks, Nature communications 6 (2015) 1–10.

[8] F. I. Karahanoğlu, D. Van De Ville, Dynamics of large-scale fmri networks: Deconstruct brain activity to build better models of brain function, Current Opinion in Biomedical Engineering 3 (2017) 28–36.

[9] N. Kinany, E. Pirondini, S. Micera, D. Van De Ville, Dynamic functional connectivity of resting-state spinal cord fmri reveals fine-grained intrinsic architecture, Neuron (2020).

[10] J. Gonzalez-Castillo, C. Caballero-Gaudes, N. Topolski, D. A. Handwerker, F. Pereira, P. A. Bandettini, Imaging the spontaneous flow of thought: Distinct periods of cognition contribute to dynamic functional connectivity during rest, NeuroImage 202 (2019) 116129.

[11] C. C. Gaudes, N. Petridou, S. T. Francis, I. L. Dryden, P. A. Gowland, Paradigm free mapping with sparse regression automatically detects single-trial functional magnetic resonance imaging blood oxygenation level dependent responses, Human brain mapping 34 (2013) 501–518.

[12] F. I. Karahanoğlu, C. Caballero-Gaudes, F. Lazeyras, D. Van De Ville, Total activation: fmri deconvolution through spatio-temporal regularization, Neuroimage 73 (2013) 121–134.

[13] G. Schwarz, et al., Estimating the dimension of a model, The annals of statistics 6 (1978) 461–464.

[14] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: Selected papers of hirotugu akaike, Springer, 1998, pp. 199–213.

[15] D. R. Gitelman, W. D. Penny, J. Ashburner, K. J. Friston, Modeling regional and psychophysiologic interactions in fmri: the importance of hemodynamic deconvolution, Neuroimage 19 (2003) 200–207.

[16] C. C. Gaudes, N. Petridou, I. L. Dryden, L. Bai, S. T. Francis, P. A. Gowland, Detection and characterization of single-trial fmri bold responses: Paradigm free mapping, Human brain mapping 32 (2011) 1400–1418.

[17] E. Uruñuela, S. Jones, A. Crawford, W. Shin, S. Oh, M. Lowe, C. Caballero-Gaudes, Stability-based sparse paradigm free mapping algorithm for deconvolution of functional mri data, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 1092–1095.

[18] A. M. Bruckstein, D. L. Donoho, M. Elad, From sparse solutions of systems of equations to sparse modeling of signals and images, SIAM review 51 (2009) 34–81.

[19] R. Tibshirani, Regression shrinkage and selection via the lasso, Journal of the Royal Statistical Society: Series B (Methodological) 58 (1996) 267–288.

[20] H. Cherkaoui, T. Moreau, A. Halimi, P. Ciuciu, Sparsity-based blind deconvolution of neural activation signal in fmri, in: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2019, pp. 1323–1327.

[21] I. Khalidov, J. Fadili, F. Lazeyras, D. Van De Ville, M. Unser, Activelets: Wavelets for sparse representation of hemodynamic responses, Signal processing 91 (2011) 2810–2821.

[22] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, et al., Least angle regression, The Annals of statistics 32 (2004) 407–499.

[23] S. Badillo, T. Vincent, P. Ciuciu, Group-level impacts of within- and between-subject hemodynamic variability in fmri, Neuroimage 82 (2013) 433–448.

[24] C. C. Gaudes, F. I. Karahanoğlu, F. Lazeyras, D. Van De Ville, Structured sparse deconvolution for paradigm free mapping of functional mri data, in: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), IEEE, 2012, pp. 322–325.

[25] N. Meinshausen, P. Bühlmann, Stability selection, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72 (2010) 417–473.