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

Functional MRI deconvolution algorithms are gaining popularity to study the dynamicss of functional brain activity and connectivity at short timescales. This work sheds light on our understanding of two state-of-the-art approaches based on L1-norm regularized estimators: Paradigm Free Mapping (synthesis model) and Total Activation (analysis model). Through simulations with varying signal-to-noise ratios, and an experimental motor task dataset, we demonstrate that both formulations produce identical estimates of the innovation and activity-inducing signals underlying BOLD events when identical hemodynamic response and regularization parameters are used. These observations open up the possibility for future developments without questioning their core formulation and performance.

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.

Functional magnetic resonance imaging (fMRI) 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 blood oxygenation level-dependent (BOLD) signal, the appropriate approximation of the neuronal activity can be obtained by means of deconvolution with an assumed hemodynamic response [6]. Simply put, deconvolution methods are capable of blindly estimating neuronal activity with no prior information of the timing of the BOLD events.

which makes them especially attractive for exploring time-varying activity of resting-state fluctuations [7, 8, 9, 10, 11], naturalistic paradigms [1], or clinical conditions such as the study of interictal events in epilepsy.

Paradigm Free Mapping (PFM) [12] — which is available as *3dPFM* in AFNI — and Total Activation (TA) [13] 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) [14] and Akaike Information Criterion (AIC) [15], and b) a selection based on the estimated median absolute deviation (MAD) 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

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 [6]:

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

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

where y_t is the measured BOLD signal on a given voxel. This equation can be reformulated in matrix notation as $\mathbf{y} = \mathbf{Hs} + \epsilon$, where $\mathbf{y}, \mathbf{s} \in \mathbb{R}^N$, $\mathbf{H} \in \mathbb{R}^{NxN}$ is the HRF in Toeplitz matrix form, N is the number of frames of the fMRI acquisition, and ϵ represents additional white Guassian noise. The signal model in (1) can be extended to represent the neuronal signal $\mathbf{s}$ in terms of its innovation signal $\mathbf{u}$, i.e. its derivative, and can be described as $\mathbf{s} = \mathbf{Lu}$ where $\mathbf{L} \in \mathbb{R}^{N \times N}$ is an integration operator.

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 [6]. 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 ($\mathbf{Hs}$) and measured ($\mathbf{y}$) signals. Yet, the estimates of the neuronal activity $\mathbf{s}$ must be constrained with a regularization term to attenuate the collinearity and high variability of the design matrix $\mathbf{H}$.

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 regularized least-squares problem [12, 16, 17]:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{y} - \mathbf{Hs}\|_2^2 + \Omega(\mathbf{s}) \tag{2}$$

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 $\mathbf{s}$ must be a sparse vector. Thus, sparse estimates of $\mathbf{s}$ 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 $\mathbf{H}$. 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.

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{y} - \mathbf{Hs}\|_2^2 + \lambda\|\mathbf{s}\|_1 \tag{3}$$

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 $\mathbf{L}$ into the design matrix $\mathbf{H}$ allows the recovery of the innovation signal $\mathbf{u}$; i.e., the derivative of the activity-inducing signal $\mathbf{s}$. Therefore, the innovation signal can be estimated by solving the following optimization problem [17, 20]:

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} \frac{1}{2}\|\mathbf{y} - \mathbf{HLu}\|_2^2 + \lambda\|\mathbf{u}\|_1 \tag{4}$$

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 $\mathbf{s}$ [13, 21, 22]:

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

where x is the neuronal-related signal; i.e., the activity inducing signal $\mathbf{s}$ 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:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \mathcal{R}(\mathbf{x}) \tag{8}$$

where $\mathbf{y}$ 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.

This work evaluates the core of the two techniques, i.e. the regularized least-squares problem with temporal regularization, which corresponds to the generalized total-variation operator in Total Activation. Thus, we do not study the impact of spatial constraints, as we assume that spatial regularization terms should perform identically on both methods.

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) [14] and Akaike Information Criterion (AIC) [15], 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) activity-inducing signal with five neuronal events, convolved it with the SPMG1 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.

Furthermore, we compared the two techniques on a finger-tapping task where the ground-truth was unknown. 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 mm^3, TR=1.5 s, TEs=10.6/28.69/46.78/64.87/82.96 ms, flip angle=70^o, MB factor=4, 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

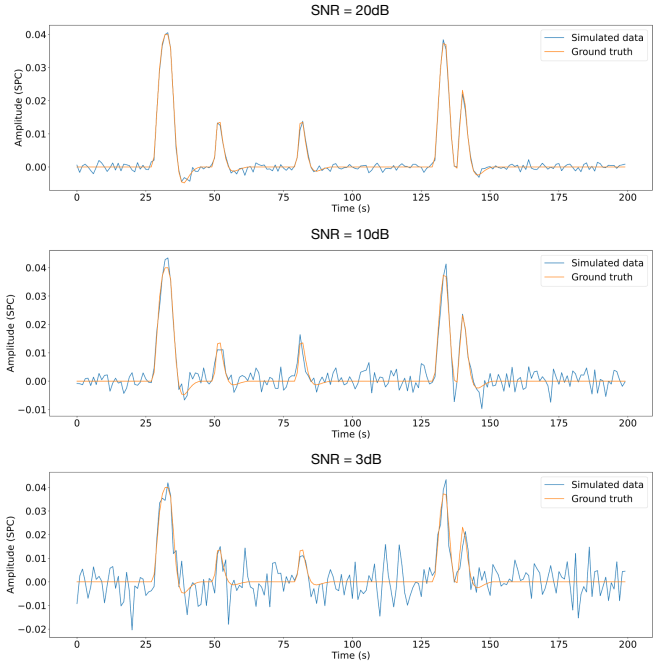

Figure 1: Simulated signal with different SNRs (20 dB, 10 dB and 3 dB).

the entire set of stimuli were presented. Data preprocessing consisted of optimally combining the echo time datasets, detrending of up to 5^{th}-order Legendre polynomials, spatial smoothing (3 mm FWHM) and normalization to signal percentage change. For this comparison, we selected a voxel that best represented the right-hand finger-tapping paradigm as shown in Figure 2.

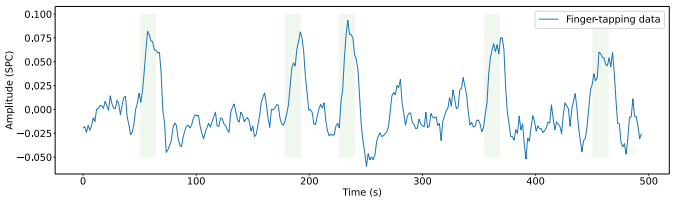

Figure 2: Most representative voxel of the finger-tapping task. Green blocks indicate the onsets and the duration of it.

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 3A shows the difference in the hemodynamic response function that PFM and TA use by default for TR = 0.1 s and TR = 1 s adjusted to a peak amplitude of one; i.e. the SPMG1 and the HRF resulting from the linear differential operator. The most observable difference between the two HRF is the time to peak: the HRF used by Total Activation does not begin at zero while the one used by PFM does.

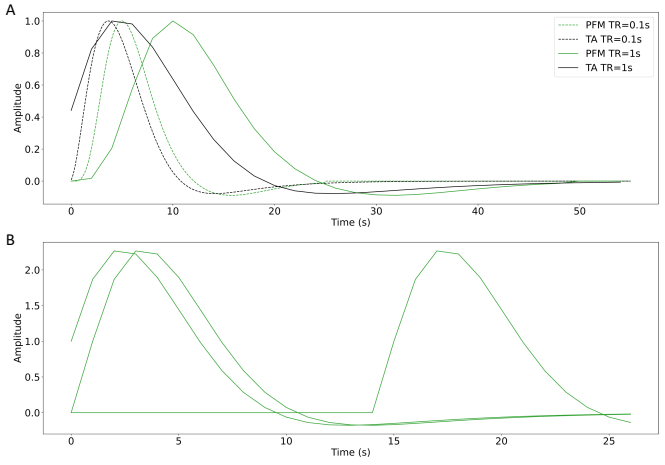

Figure 3: A) Canonical HRF models typically used by PFM (green) and TA (black) at TR = 0.1 s (dashed lines) and TR = 1 s (solid lines). Without loss of generality, the waveforms are scaled to unit amplitude for visualization. B) Representation of three shifted HRFs at TR=1 s (onsets=0, 1, and 15 s) that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

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 (Figure 3B).

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 converge to a previously estimated noise level of the data fit $\tilde{\sigma}$, where this pre-estimated noise is calculated from the median absolute deviation of fine-scale wavelet coefficients (Daubechies, order 3) [13]:

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

Thus, we calculated the regularization path with PFM (as described in 3.4) 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 4 depicts the estimated activity-inducing, innovation and activity-related signals when updating λ as in (10) in the three simulated SNR settings using the spike model (left) and the block model (right). Figure 4 (left) shows nearly identical results between PFM (left) and TA (right). 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 convergence of the residuals to have the same variances as the MAD estimate of the noise yields results that are identical in practice as shown in Figure 4 (right).

In addition, we performed the same comparison on experimental data as introduced in 3.1. Figure 5 (row 3) illustrates that the estimated activity-inducing, innovation and activity-related signals with PFM and TA are once again practically identical both for the spike model (left) and the block model (right).

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) [14] and the Akaike Information Criterion (AIC) [15]. Hence, we calculated the regularization path with PFM by means of the least angle regression (LARS) algorithm [23] and used the λ in the path to solve the deconvolution problem with Total Activation.

Figure 6 (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 6 (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 8). Thus, Figure 6 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 $\mathbf{s}$ 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 mainly undistinguishable results for both PFM and TA methods as shown in Figure 7 (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 7 right). Figure 8 (right) demonstrates that the 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.

Furthermore, we performed the same analysis on experimental data as shown in Figure 5 (rows 1-2, 4). Row 1 demonstrates that the PFM and TA regularization paths

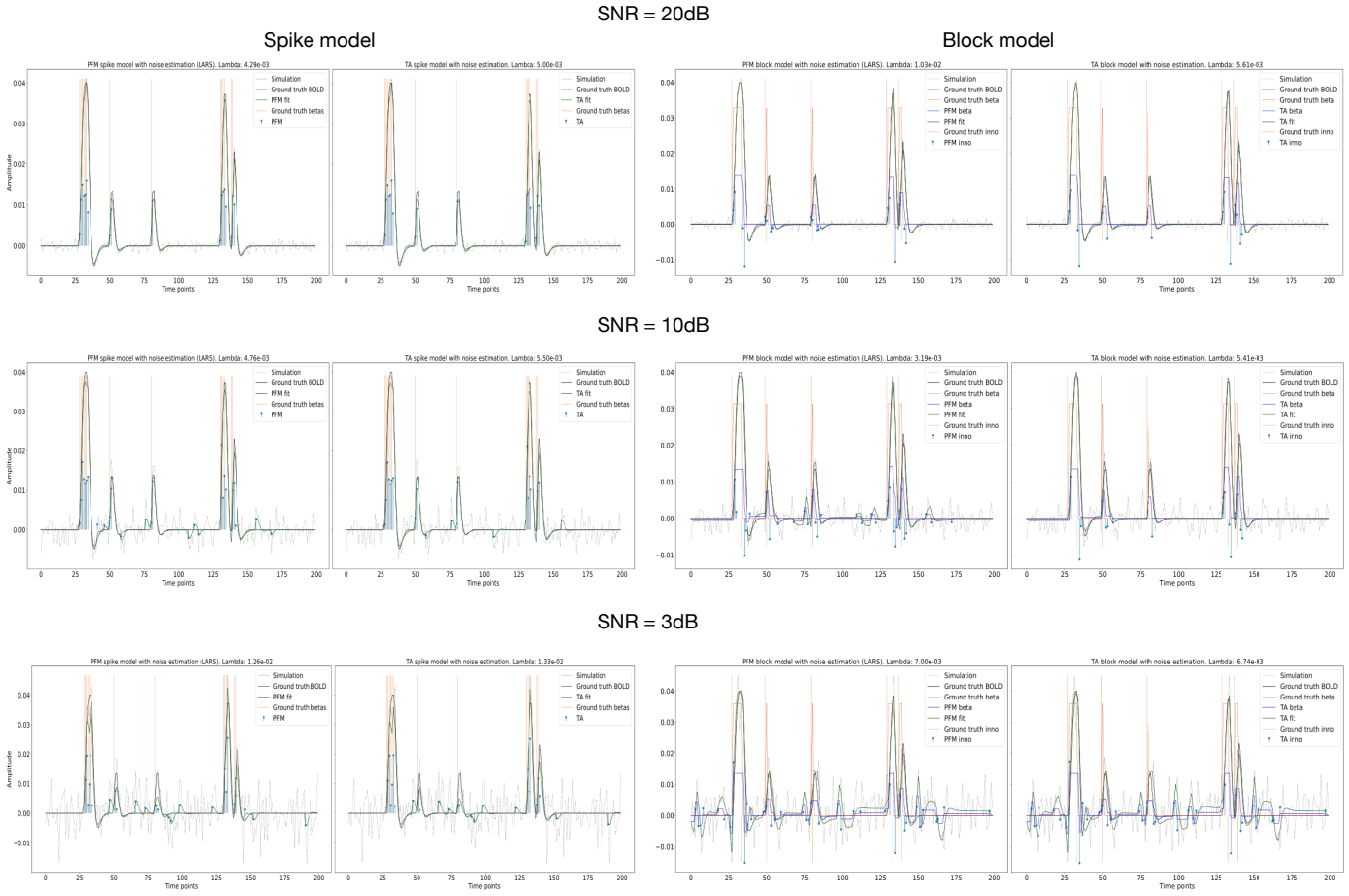

Figure 4: Estimated activity-inducing, innovation and activity-related (fit, $\mathbf{x}$) signals when λ is selected based on convergence of residuals to have same variance as MAD estimate of noise with spike model (left, with PFM on the left and TA on the right) and block model (right, with PFM on the left and TA on the right) on SNR = 20dB (top), SNR = 10dB (middle) and SNR = 3dB (bottom).

are identical when deconvolving experimental data, regardless of the deconvolution model (spike or block). Even though tiny differences can be seen between the two methods in the BIC and AIC selection of λ in row 4, row 2

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 [24], ISBI 2012 by César [25], and Farouj with a different formulation. Say conceptual differences among those.
  - Mention stability-selection [26]
  - Debiasing

Synthesis (PFM)	Analysis (TA)
+	+ Both the spike and block models solve the regularization problem with the same HRF.
	–

Table 1: Pros and cons of Paradigm Free Mapping and Total Activation.

  - Connected to debiasing other deconvolution algorithms that are based on a norm lower than 1.

This work demonstrates that Paradigm Free Mapping and Total Activation yield practically identical results when the same HRF model and regularization parameter are employed, demonstrating that synthesis and analysis models

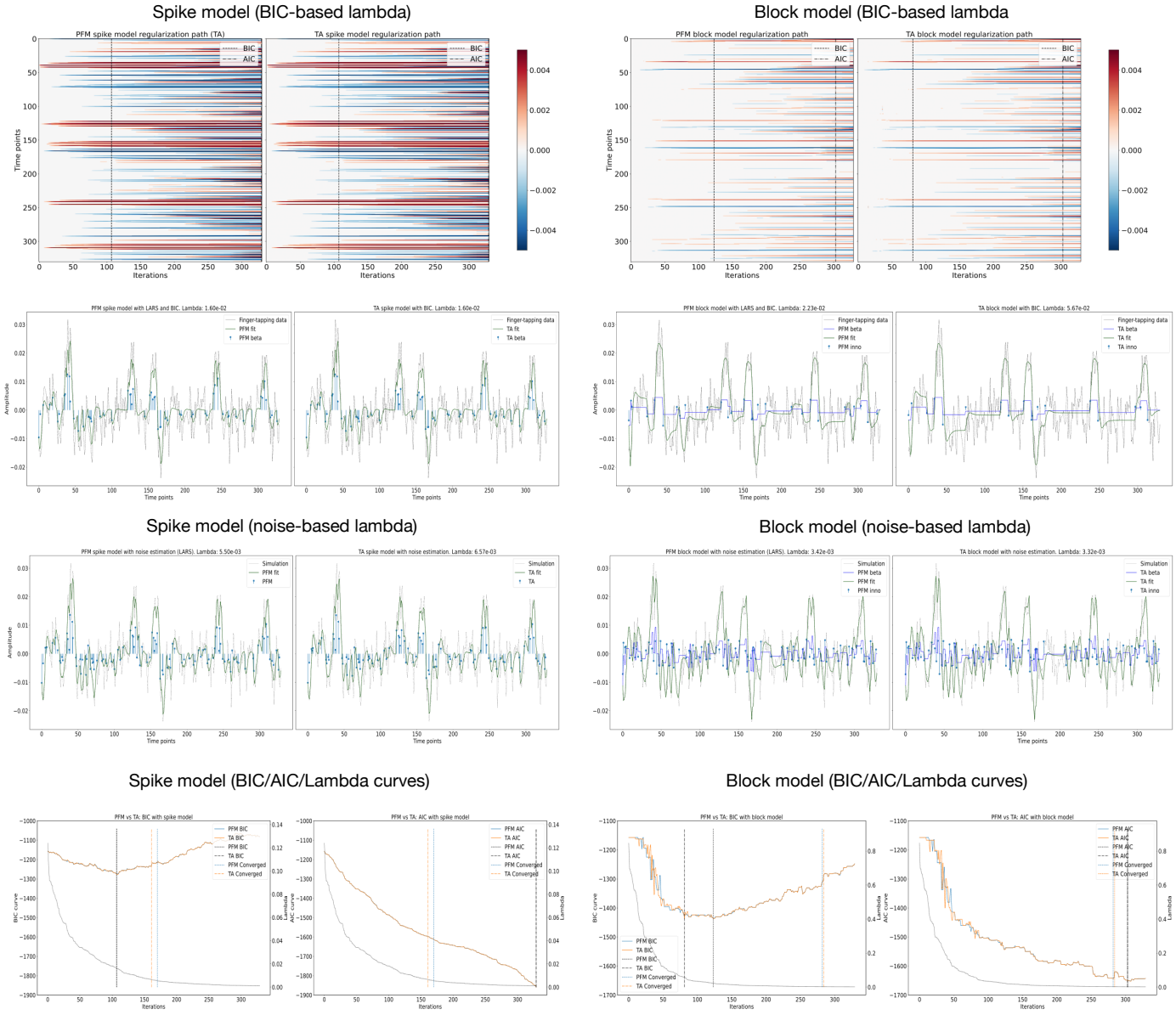

Figure 5: (Row 1) Regularization paths of the estimated activity-inducing signal (spike model — left) and innovation signal (block model — right); (Row 2) activity-inducing, innovation and activity-related (fit, $\mathbf{x}$) signals when λ is selected based on BIC, or (Row 3) based on convergence of residuals to have same variance as MAD estimate of noise; (Row 4) Corresponding curves of BIC and AIC, where the vertical lines indicate the three options to select λ (BIC, AIC and Converged/MAD).

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 can be assumed 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) [27] or the Generalized Forward-Backward Splitting [28] techniques.

Future work should focus on investigating the appropriate formulation depending on data acquisition (e.g. single-echo vs multi-echo), accounting for HRF variability, robust methods to select the regularization parameter, and other potential $\ell_{p,q}$-norm regularization terms (e.g. $p < 1$) or debiasing approaches.

5. Code availability

The code and materials used in this work can be found in the following GitHub repository: `https://github.com/`

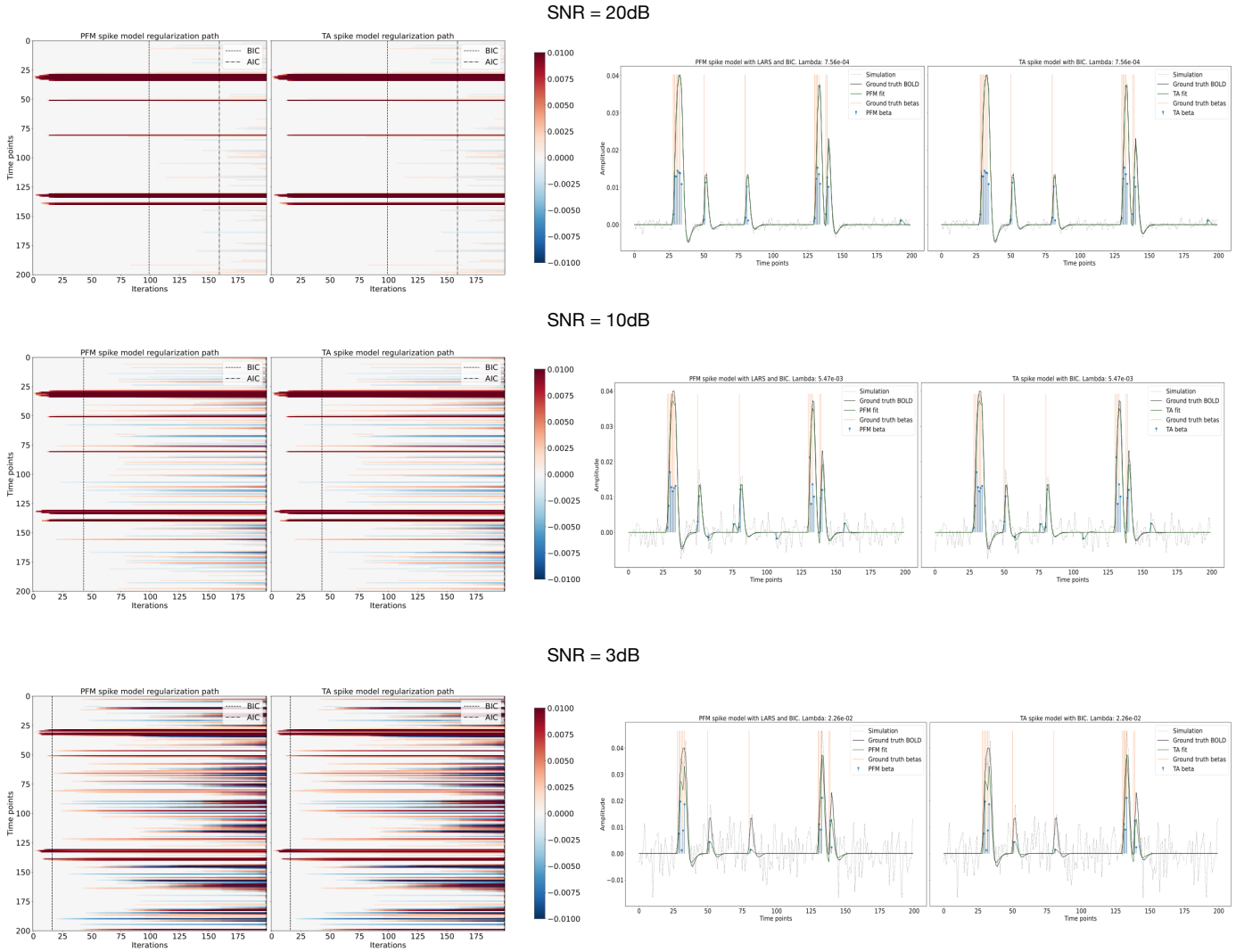

Figure 6: Spike model simulations. (Left) Heatmap of the regularization paths of the activity-inducing signal estimated with PFM and TA as a function of (increasing number of iterations in x-axis), whereas each row in the y-axis shows one time-point. Vertical lines denote iterations corresponding to the Akaike and Bayesian Information Criteria (AIC and BIC). (Right) Estimated activity-inducing (blue) and activity-related (green) signals when is set based on BIC. All estimates of are identical, regardless of SNR.

`eurunuela/pfm_vs_ta`. We encourage the reader to play with the parameters (e.g. SNR, varying HRF options and mismatch between algorithms, TR, number of events, onsets, and durations) in the provided Jupyter notebooks.

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] 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.

[7] 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.

[8] 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

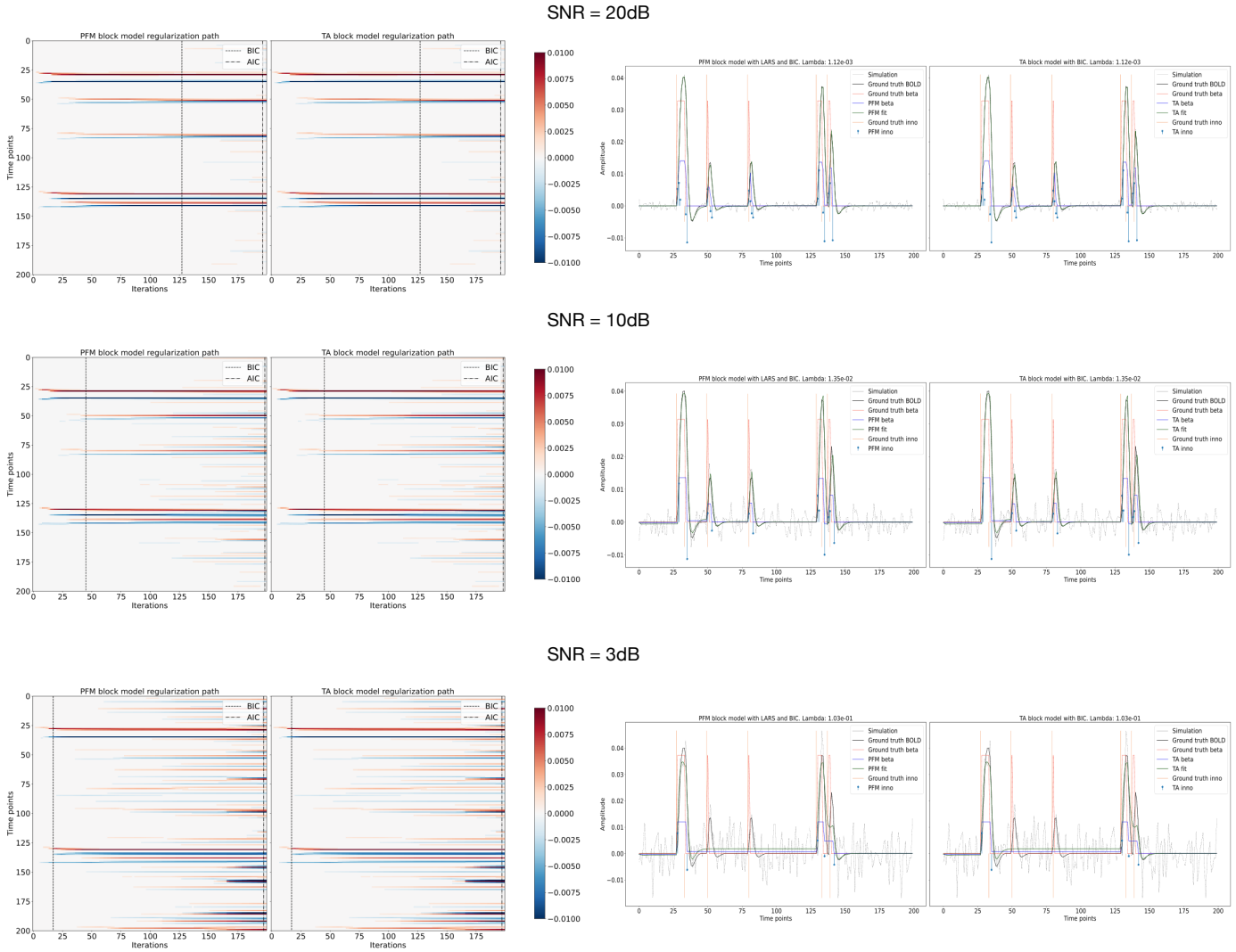

Figure 7: Block model simulations. (Left) Heatmap of the regularization paths of the innovation signal estimated with PFM and TA as a function of (increasing number of iterations in x-axis), whereas each row in the y-axis illustrates one time-point. Vertical lines denote iterations corresponding to the Akaike and Bayesian Information Criteria (AIC and BIC). (Right) Estimated innovation (blue) and activity-related (green) signals when is set based on BIC. All the estimates of are identical, regardless of SNR.

6 (2015) 1–10.

[9] 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.

[10] 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).

[11] 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.

[12] 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.

[13] 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.

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

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

[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,

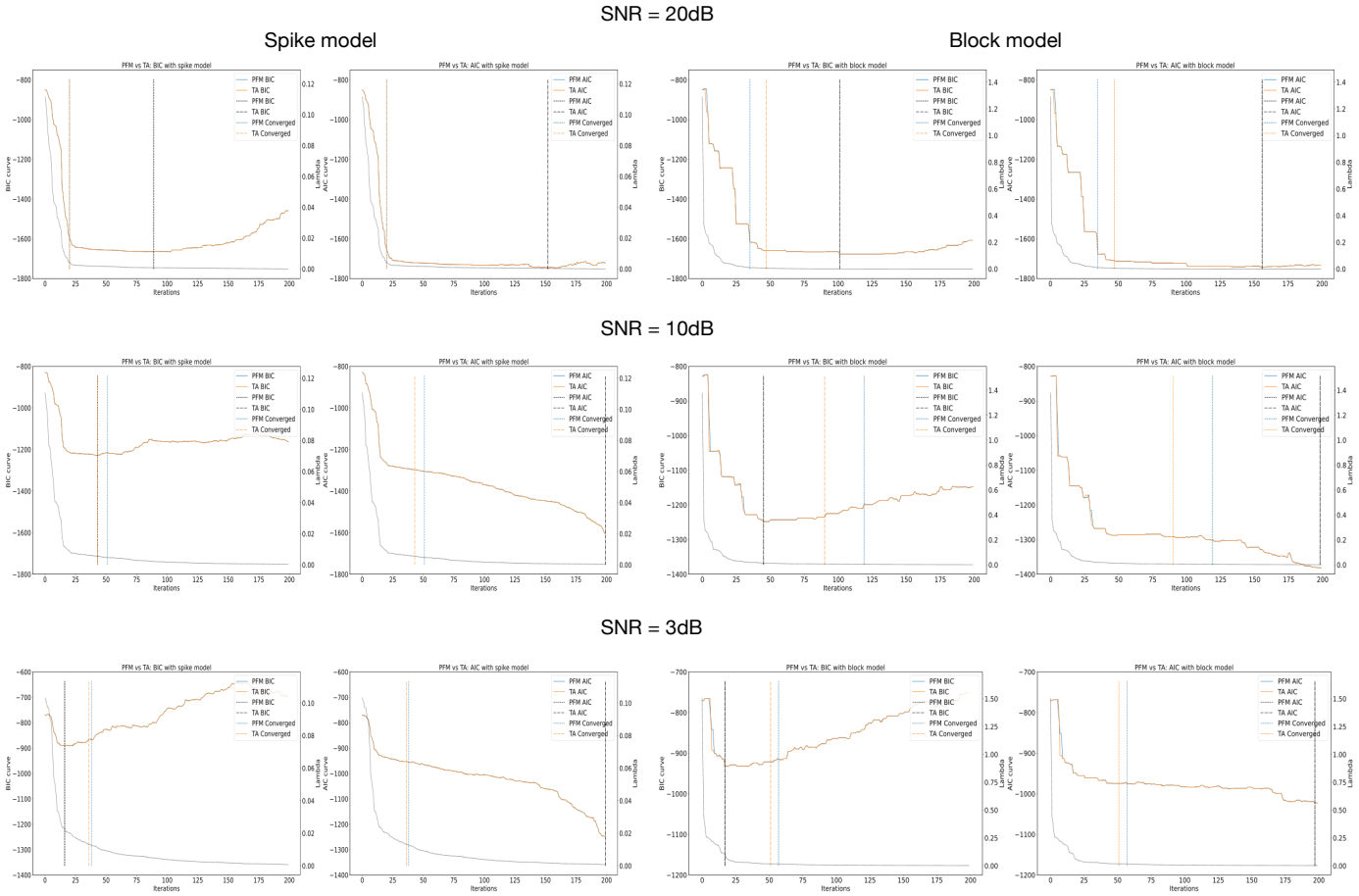

Figure 8: Lambdas and cost.

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] F. I. Karahanoglu, İ. Bayram, D. V. D. Ville, A Signal Processing Approach to Generalized 1-D Total Variation, IEEE Transactions on Signal Processing 59 (2011) 5265–5274.

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

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

[25] 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.

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

[27] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm, Society for Industrial and Applied Mathematics Journal on Imaging Sciences 2 (2009) 183–202.

[28] H. Raguet, J. Fadili, G. Peyré, A Generalized Forward-Backward Splitting, SIAM Journal on Imaging Sciences 6 (2013) 1199–1226.