Signal Processing 91 (2011) 2810–2821

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Activelets: Wavelets for sparse representation of hemodynamic responses

Ildar Khalidov [a], Jalal Fadili [b], François Lazeyras [c], Dimitri Van De Ville [c,d,*], Michael Unser [a]

[a] Biomedical Imaging Group, Ecole Polytechnique Fédérale de Lausanne, Switzerland
[b] GREYC, ENSICAEN, Caen, France
[c] Department of Radiology and Medical Informatics, University of Geneva, Switzerland
[d] Medical Image Processing Lab, Ecole Polytechnique Fédérale de Lausanne, IBI-STI/EPFL, Station 17, CH-1015 Lausanne (VD), Switzerland

ARTICLE INFO

Article history:
Received 15 January 2010
Received in revised form
3 March 2011
Accepted 11 March 2011
Available online 21 March 2011

Keywords:
BOLD fMRI
Hemodynamic response
Wavelet design
Sparsity
ℓ_1 minimization

ABSTRACT

We propose a new framework to extract the activity-related component in the BOLD functional magnetic resonance imaging (fMRI) signal. As opposed to traditional fMRI signal analysis techniques, we do not impose any prior knowledge of the event timing. Instead, our basic assumption is that the activation pattern is a sequence of short and sparsely distributed stimuli, as is the case in slow event-related fMRI.

We introduce new wavelet bases, termed "activelets", which sparsify the activity-related BOLD signal. These wavelets mimic the behavior of the differential operator underlying the hemodynamic system. To recover the sparse representation, we deploy a sparse-solution search algorithm.

The feasibility of the method is evaluated using both synthetic and experimental fMRI data. The importance of the activelet basis and the non-linear sparse recovery algorithm is demonstrated by comparison against classical B-spline wavelets and linear regularization, respectively.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

1.1. fMRI and time-course analysis

Functional magnetic resonance imaging (fMRI) is being used increasingly in modern neuroscience. It allows non-invasive measurements of the evoked neural activity through neurovascular coupling and the blood-oxygenation-level-dependent (BOLD) effect, first observed by Ogawa et al. [1]. Few seconds after the stimulation, a decrease in deoxyhemoglobin (dHb) level in the implicated brain regions occurs, making the so-called local T2* MR signal stronger. The typical spatio-temporal resolution of fast fMRI echo-planar imaging (EPI) techniques varies between 1–50 mm^3/voxel and 0.5–5 s/volume.

In traditional fMRI, the subject's brain is scanned while performing a task or being exposed to stimuli that are relevant to some cognitive function [2]. There are two general types of experimental design. Within the block-based paradigm, each stimulus lasts for a non-negligible period of time (several scans) and is modeled as a boxcar function. In the event-based framework, a stimulus is short in time and can be mathematically represented with a Dirac delta-function. Event-related fMRI activations are weaker and more variable than block-based responses, which makes them harder to detect.

Modeling the hemodynamic system has been (and still is) a subject of active research [3]. One popular framework to link neural activation and BOLD response makes use of the so-called balloon model by Buxton and Frank [4].

* Corresponding author at: Medical Image Processing Lab, Ecole Polytechnique Fédérale de Lausanne, IBI-STI/EPFL, Station 17, CH-1015 Lausanne (VD), Switzerland. Tel.: +41 21 6939669; fax: +41 21 6936990.
E-mail address: dimitri.vandeville@epfl.ch (D. Van De Ville).

0165-1684/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2011.03.008

Essential components of their system include regional blood flow, vessel volume, dHb extraction fraction and the perfusion-inducing signal by the neuronal response to stimulus. Although the underlying differential equations are non-linear, it is common to assume that the hemodynamic system is linear and time-invariant. In this setting, a model for the task-related BOLD response can be obtained as a convolution between the stimulus pattern, describing the task, and the *hemodynamic response function* (HRF). This is the type of model that is used in all standard fMRI analysis packages.

The measured functional MR signal is degraded due to various sources: physiological factors such as respiratory and (aliased) cardiac components, subject's movements, scanning artifacts due to field inhomogeneity, image reconstruction and post-processing. Given the noisy measurements, traditional fMRI data analysis tries to find evidence for the presence of a hypothetical task-related BOLD response [2]. If such evidence is found (on statistical grounds), voxels are declared as "active". The most popular framework proposes a general linear model (GLM) that contains regressors of interest (e.g., task responses for various conditions) and other variates (baseline, low-frequency drifts, and so on). Given the noise statistics, the parameters (weights of each regressor) are then fitted to the data. The parameters' strength (or their statistical significance) is then evaluated by taking into account the residual error.

In addition to confirmatory analyses, researchers have also proposed data-driven exploratory methods that do not (or only partially) rely on the prior knowledge. The most popular ones are subspace methods, such as principal components analysis (PCA) [5] and independent component analysis (ICA) [6,7]. These methods have the capability to reveal unmodeled trends in the data. However, manual intervention is often needed to distinguish noise-related from neurophysiological-relevant components. Semi-blind approaches use the knowledge of the stimulus timings. Glover et al. include a calibration trial from which they estimate the HRF [8]. This provides them with a Wiener filter that they apply to subsequent measurements to estimate the neural activity pattern through a deconvolution process. Makni et al. proposed a Bayesian framework for a joint HRF estimation-detection task [9]. Other studies have also investigated in detail the trial-by-trial variability of event-related BOLD responses [10–12] or have used sparsity to constrain the fitting of a GLM with many regressors [13,14].

In slow event-related fMRI experiments, the activity-inducing events are well separated in time from each other. This means that the corresponding BOLD responses do not overlap and that the deconvolved stimulus pattern is sparse in time. In this paper, we assume that the activations are distributed sparsely and develop a new wavelet-based framework that is able to find the activity-related signal component in an fMRI time-course, without prior knowledge on the positions of the activity-inducing time onsets. While conventional fMRI experiments heavily rely on the knowledge of the timing of the events (typically, a stimulus or a recorded feedback), it should be noted that in some cases this information is imprecise or even unavailable. Possible reasons for this could be related to the type of subject (e.g., feedback in a decision-making task with small children) or because the "task" is implicit (e.g., epileptic patients with sparsely distributed interictal epileptic discharges [15]). Our aim is to extract the activity-related signal component from time-courses, which can then be further analyzed depending on the neurological question at hand.

1.2. Wavelets and fMRI

Over the last decade, wavelets have become an essential tool in mathematics, engineering, and physics [16]. The wavelet function, when applied to the data, behaves as a multiscale derivative operator. The order of the derivative operator is directly linked to the wavelet's number of vanishing moments. Therefore, singularities, such as edges, only have a local influence in the decomposition and are well approximated by a few (large) coefficients. This property has been the driving force behind many applications, such as coding [17] and denoising [18]. More recently, this feature is exploited in more general applications (e.g., reconstruction and deconvolution) by optimizing a sparsity-promoting penalty on the wavelet coefficients, typically via ℓ_1 regularization [19,20]. This principle has also been extended to a wider range of (non-orthogonal) dictionaries such as (redundant) wavelet frames or combined transformations, see [21,22] for an overview.

The wavelet transform has also been applied to fMRI time-series processing. In the spatial domain, the activity maps can be compactly represented by the wavelet decomposition; for an overview, see [23,24]. In the temporal domain, the transform's decorrelating property can be used advantageously to fit the GLM's parameters in the presence of colored noise, which is the case in fMRI with fast repetition timing [25]. Also, the Hurst exponent—the characterizing parameter of a self-similar process—can be efficiently estimated in the wavelet domain and used to distinguish between healthy subjects and patients suffering from Alzheimer disease [26]. Wavelet-based characterization of multi-fractal behavior for fMRI has been proposed [27]. Finally, estimation with the use of penalized partial linear models and classical wavelets has been investigated in [28].

1.3. Contributions of this paper

While traditional wavelets, with their derivative-like behavior, offer good energy compaction, we believe that they are not well suited for the activation-related signal in fMRI. In this paper, we design a new type of wavelets, named "activelets", driven by a characterization of the hemodynamic system. Basically, we (still) consider the hemodynamic system as linear and stationary and derive the differential operator L that links the hemodynamic response with the stimulus. To that end, we make use of a linearized set of differential equations that model the hemodynamic system [29]. Embarking from there, we design the exponential-spline wavelets that essentially invert the system's response. Therefore, these wavelets,

coined activelets, will form a dictionary in which the hemodynamic response is sparsely represented. The essential feature that characterizes activelets from traditional wavelets is that they have a number of (well-chosen) exponential vanishing moments. Short activations then induce a cone of influence in the activelet decomposition of the BOLD signal, the same way as singularities do in the classical wavelet representation. Consequently, the use of a pursuit algorithm in the activelet dictionary will help in finding the fMRI "activations", without prior knowledge of onset timings nor activity strengths. To get the sparsest representation, we formulate the problem as an ℓ_1-regularized optimization problem and we deploy iterative-thresholding algorithms to solve it. To satisfy the sparsity hypothesis, we only focus on activity that is sparsely distributed in time; e.g., such as in slow event-related designs.

1.4. Organization of the paper

After a brief preliminary on our fMRI time-course model in Section 2, we fully detail the activelets method in Section 3. Then, in Section 4, we derive an alternative linear approach—the minimum mean-squared error (MMSE) estimator that uses the same operator L as a whitening operator—that will then be used as reference method in the experimental section. In Section 5.1, we compare the techniques on the simulated fMRI data. In Section 5.2, we show how the method can be applied to real data. Finally, in Section 6, we discuss the main aspects of our method, as well as its potential for neuroscience.

2. fMRI time-course modeling

In this paper, we consider a Dirac spike train $s(t)$ as the source of activity in the brain. It should be noted that each impulse represents an event (spike) at the fMRI timescale (i.e., resolution of seconds) and is not be confused with the underlying neuronal spiking at the millisecond resolution. Mathematically, we can write the activity-inducing signal

$$s(t)=\sum_l c_l\delta(t-t_l), \tag{1}$$

where c_l and t_l denote the activity strengths and activity onsets, respectively; $\delta(t)$ denotes the Dirac impulse. Due to neurovascular coupling and a subtle interplay of blood flow, blood volume, and overcompensation of oxygenated hemoglobin, the BOLD response can be measured in fMRI time-series. We represent the hemodynamic system that links activity to BOLD signal as

$$x(t)=\mathcal{H}\{s\}(t), \tag{2}$$

where operator $\mathcal{H}$ is usually modeled as a system of differential [30,29].

We assume the hemodynamic system to be linear and shift-invariant—an approximation that is common in literature and very reasonable when the events are sufficiently spaced in time. In that case, the activity-related BOLD signal can be modeled as a weighted sum of shifted hemodynamic response functions

$$x(t)=\sum_l c_l h(t-t_l). \tag{3}$$

The problem that we face is to estimate the presence of activity-related signal $x(t)$ in the noisy fMRI measurements:

$$y[k]=x(k)+n[k], \tag{4}$$

where we assume a uniform sampling step without loss of generality, and $n[k]$ is the disturbance term that includes noise, baseline, drifts, aliased cardiac and respiratory contributions (cf. Fig. 1).

Our problem setting is different from a classical "linear model" analysis, which looks for the presence of a *fixed* regressor; i.e., a linear combination of a few known elementary signals. In our case, we just assume that the response will be of the form of (3), without fixing the weights or onset times.

3. The activelets framework

3.1. From hemodynamic system to operator

As we assume that the hemodynamic response function $h(t)$ originates from an underlying linear shift-invariant system, we can write the differential equation

$$\mathrm{L}\{h\}(t)=\delta(t), \tag{5}$$

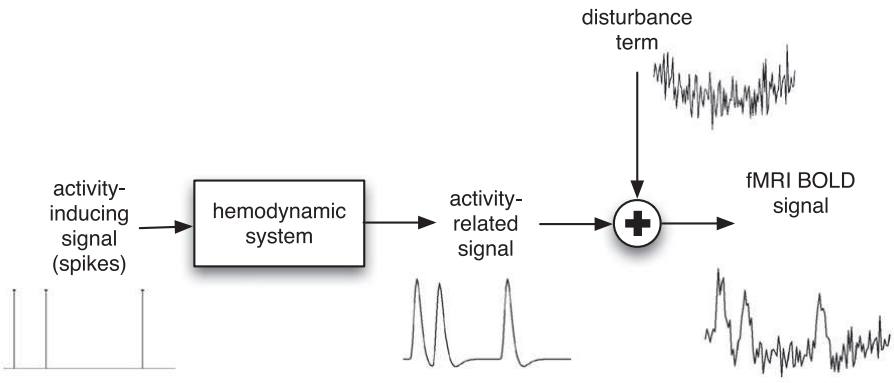

Fig. 1. The hemodynamic system links neuronal activity to fMRI BOLD response, which is due to neurovascular coupling and a complex interplay between various physiological parameters.

where the differential operator L is also linear, time-invariant. Mathematically, $h(t)$ is a Green function of L and the impulse response of the causal inverse operator L^{-1}.

In general, the operator L can be characterized by its Fourier transform

$$\hat{\mathrm{L}}(\omega)=\frac{\prod_{n=1}^{N}(j\omega-\alpha_n)}{\prod_{m=1}^{M}(j\omega-\gamma_m)}. \tag{6}$$

For notational simplicity, we group the parameters in $\vec{\alpha}=(\alpha_1,\ldots,\alpha_N)$ and $\vec{\gamma}=(\gamma_1,\ldots,\gamma_M)$.

Most authors assume a linear, shift-invariant model of the hemodynamic system, which can be justified even when the underlying set of coupled partial differential equations (PDEs) are non-linear. In the latter case, we identify the operator L that is linked to the hemodynamic system by linearizing the system of non-linear differential equations. This procedure is equivalent to considering the first-order Volterra series approximation [29]. From the derivation that can be found in the Appendix, we obtained the following parameters:

$$\vec{\alpha}=\left(-\frac{1}{\tau_0},-\frac{1}{\alpha\tau_0},-\frac{1}{2\tau_s}\left(1\pm j\sqrt{\frac{4\tau_s^2}{\tau_f}-1}\right)\right),$$

$$\vec{\gamma}=\left(-\frac{(k_1+k_2)\left(\frac{1-\alpha}{\alpha\tau_0}-\frac{c}{\alpha}\right)-(k_3-k_2)\frac{1}{\tau_0}}{-(k_1+k_2)c\tau_0-k_3+k_2}\right). \tag{7}$$

There is an additional scaling factor $(V_0\varepsilon/\tau_0)(-(k_1+k_2)c\tau_0-k_3+k_2)$. An overview of the numerous constants with their meanings and typical values can be found in Table 2.

3.2. Wavelet representations

In our model, the activity-inducing signal is a Dirac impulse train that passes through a system characterized by the operator L^{-1}. If we applied L to the ideal continuous BOLD signal, we would recover a sparse representation. In practice, given the noisy samples, we would like to have a transform that compacts the activation energy on a few large coefficients. We would then estimate activations by identifying the most significant coefficients with the help of sparsity-pursuit algorithms.

The classical wavelet transform is known to provide a sparse representation of piecewise-smooth functions with a finite number of singularities. Traditional wavelets act as N th-order derivatives. Whenever the number N of vanishing moments is greater than the smoothness of the function, such a wavelet would, at each point, annihilate the first N terms of the Taylor series expansion, essentially leaving significant coefficients only close to the discontinuity points.

In our case, from (3) and (5), the activity-related signal $x(t)$ satisfies $\mathrm{L}\{x\}(t)=0$ on the intervals $t_l<t<t_{l+1}$; i.e., the operator L annihilates $x(t)$ at all points except the activation points t_l. For a sparse representation, we would like the wavelet to "behave like L", thus leaving non-zero coefficients only around t_l. We now provide an explicit construction of such wavelets.

3.3. Operator-like wavelet basis

As we saw in Section 3.1, the operator L in our problem can be defined by a linear time-invariant differential equation. For this class of operators, the corresponding family of wavelets $\Psi=\{\psi_{i,k}\}_{i,k\in\mathbb{Z}}$ has been defined in [31], where ψ_i is the scale-dependent mother wavelet and $\psi_{i,k}(t)=\psi_i(t-2^ik)$. The key to the construction of these wavelets is the operator's Green function $\rho_{\mathrm{L}}(t)$ for which $\mathrm{L}\{\rho_{\mathrm{L}}\}(t)=\delta(t)$ holds. In particular, the multiresolution analysis $\{V_i\}_{i=-\infty}^{\infty}$ of $L_2(\mathbb{R})$ can be obtained by (conceptually) deploying these functions at the locations $t=2^ik$, $k\in\mathbb{Z}$. In general, the Green function is not a Riesz basis generator and therefore we need to introduce B-spline-like scaling functions $\varphi_i(t)$, which are shortest localized versions of $\rho_{\mathrm{L}}(t)$, to form a Riesz basis of $L_2(\mathbb{R})$. For example, in case of a regular derivative $\mathrm{L}=\mathrm{D}$, ρ_{D} would be the Heaviside function and φ_i the zeroth-degree B-spline dilated with a factor 2^i; for a general differential operator L, however, the localization will be scale-dependent and so will be φ_i. Once the multiresolution analysis has been defined, the wavelets can be derived as the generators of the orthogonal complement between two subsequent spaces. These wavelets will be of the form $\psi_i(t)=\mathrm{L}^*\{\phi_i\}$, where ϕ_i is a low-pass smoothing function at scale i. Consequently, the operator-like wavelets annihilate the null space components of L:

Property 1 (*Vanishing exponential moments*). *For each scale* $i\in\mathbb{Z}$, *shift* $t'\in\mathbb{R}$, $\alpha\in\vec{\alpha}$ *of multiplicity* $m\geq k$:

$$\int_{-\infty}^{\infty}t^k e^{\alpha t}\psi_i^*(t-t')\,\mathrm{d}t=0.$$

(*For proof, see* [31].)

For our purpose, this multiscale operator-like behavior of the wavelets is most important; i.e., for an input function $f(t)$, the dyadic wavelet transform results into wavelet coefficients

$$d_i[k]=\langle f,\psi_{i,k}\rangle=(\mathrm{L}\{f\}*\phi_i^*)(2^ik).$$

Given the linearized model for the BOLD signal, $x(t)$ in (3) satisfies $\mathrm{L}\{x\}(t)=\sum_k c_k\delta(t-t_k)$. The properties of the corresponding activelet basis Ψ guarantee that the wavelet coefficients for $x(t)$ decay rapidly, eventually turning to 0 as the support of the wavelet does not cover t_k. By setting the operator L parameters to 0, we get back the traditional B-spline wavelets and pure (or classical) derivatives. In Fig. 2, we show the evolution of the wavelet at scale 0 as the parameters of L evolve linearly from 0 to the values in the balloon model.

Given the embedding of the approximation spaces, $V_{i+1}\in V_i$, we can derive the scaling relation

$$\varphi_{i+1}(t)=\sum_k h_i[k]\varphi_i(t-2^ik),$$

where h_i is the scaling filter. Similarly, due to $V_i\oplus_\perp W_i=V_{i-1}$, the wavelet filter can be derived:

$$\psi_{i+1}(t)=\sum_k g_i[k]\varphi_i(t-2^ik).$$

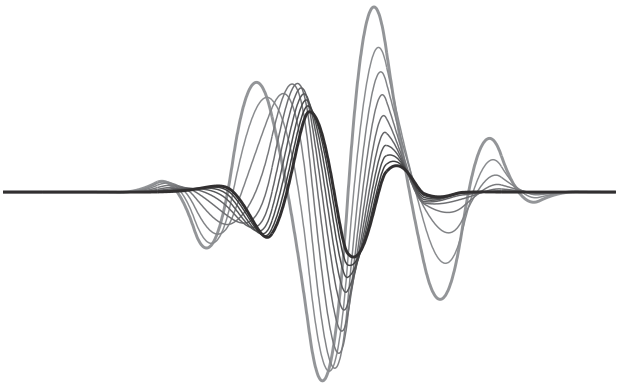

Fig. 2. From classical B-spline wavelets (thick gray) to activelets (thick black).

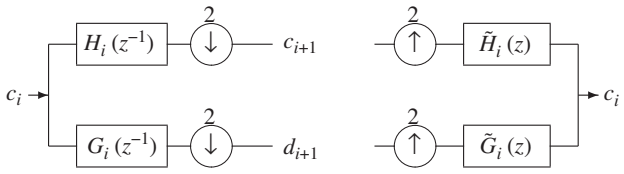

Fig. 3. Filterbank implementation of the activelet basis decomposition for a single decomposition level. Notice that the scaling and wavelet filters are scale-dependent. The reconstruction is performed using the dual filters.

Explicit forms of the scaling and wavelet filters can be found in [31], including the dual filters for the synthesis side. Given these filters, we can adapt Mallat's fast decomposition-reconstruction filterbank algorithm with scale-dependent filters as shown for one decomposition level in Fig. 3. Our implementation performs filtering in the FFT domain since closed-form expressions of the filters are known in the Fourier domain, similar to [32]. We can also obtain the undecimated activelet transform by using the "à trous" algorithm; i.e., the filters G_i and H_i are upsampled with a factor 2^i instead of downsampling the signal.

3.4. Algorithms for the sparse recovery problem

The activelet transform guarantees a sparse representation for the activity-related signal. The activelet basis can easily be extended to an activelet frame by performing the undecimated activelet transform (UDAT). The frame expansion being overcomplete, it brings shift-invariance and can lead to even sparser representations. Given the noisy data $y[k]$, $k=1,\ldots,T$ in (4), we would like to identify (to a good approximation because of noise) the signal x that has a representation as sparse as possible (ideally the sparsest one) in the undecimated activelet dictionary.

Let Φ be the $T\times K$ overcomplete dictionary matrix whose columns include the UDAT basis functions, normalized to a unit ℓ_2 norm. $x=\Phi w_0$ is the ideal noiseless fMRI signal in (4) assumed to have a sparse representation vector w_0 in Φ (synthesis-type prior), but we observe a noisy version of it y. We seek to identify the components of w_0 by solving the convex ℓ_1 optimization problem

$$(P_{1,\varepsilon}): \min_w \|w\|_1 \quad \text{subject to } \|W(y-\Phi w)\|_2 \le \varepsilon, \tag{8}$$

where W is a linear weighting operator, and $\varepsilon>0$. The weighting operator is problem-dependent and may account for our prior statistical knowledge on the noise. Problem $(P_{1,\varepsilon})$ is equivalent[1] to the Lagrangian form

$$(Q_{\lambda,\ell_1}): \min_w \|y-\Phi w\|_2^2+\lambda\|w\|_1. \tag{9}$$

(Q_{λ,ℓ_1}) is the well-known BPDN, or the popular Lasso in the statistical literature.

Problems $(P_{1,\varepsilon})$ and (Q_{λ,ℓ_1}) have been extensively studied in the recent years both in terms of their theoretical guarantees and convergent algorithms to solve them. We will not delve into these details here and the interested reader may refer to e.g., [33,22].

Among the algorithms to solve (Q_{λ,ℓ_1}), in the statistical literature, the Lasso, LARS and homotopy methods were proposed to track its regularization path by solving it for *all* $\lambda\in[0,\|\Phi^{\mathrm{T}}y\|_\infty]$ [34–36] (in fact in the overdetermined case). These methods associate to each problem (Q_{λ,ℓ_1}), $\lambda\in[0,\|\Phi^{\mathrm{T}}y\|_\infty]$, a solution $w_\lambda^\star$, and follow the entire solution path $\{w_\lambda^\star\}$ starting at $w_\lambda^\star=0$ when $\lambda=\|\Phi^{\mathrm{T}}y\|_\infty$ [36]. The key observation is that the solution subset is piecewise-constant as a function of λ, changing only at critical values of λ; i.e., the solution path is polygonal.

The LARS method computes the solution by considering one coordinate at a time as a candidate to enter the active set. Inspired by the notion of path-following, an accelerated algorithm (IT-LARS) was proposed [37]. The IT-LARS follows the solution path approximately by successively selecting *groups* of atoms at each iteration using a stagewise iterative-thresholding (IT) variant of LARS. Now, the sequence $\{\lambda_\ell\}_{\ell\ge 0}$ is not data-adapted as in LARS, but allowed to be strictly decreasing. Let A^+ be the Moore–Penrose pseudo-inverse of a matrix A. For a subset $I\subset\{1,\ldots,T\}$, $\bar{I}$ is its complement. A_I is the restriction of A to the columns indexed by I. For a vector d, $d[I]$ is the subvector indexed by I. With these notations, the main steps of IT-LARS are summarized in Algorithm 1.

Algorithm 1. IT-LARS algorithm.

Initialization: Iteration counter $\ell=0$, $w^{(0)}=0$, $r^{(0)}=y$, $\lambda_\ell=\|\Phi^{\mathrm{T}}y\|_\infty$.
Choose τ, e.g., $\tau=\sqrt{T}\sqrt{1+2\sqrt{2/T}}$ (see the text).
for $\|r^{(\ell)}\|_2>\tau\sigma_n$ **do**
- Residual : $r^{(\ell)}=y-\Phi w^{(\ell)}$,
- Correlation : $c^{(\ell)}=\Phi^{\mathrm{T}}r^{(\ell)}$,
- Hard Thresholding : $I^{(\ell)}=\{i:|c^{(\ell)}[i]|>\lambda_\ell\}$,
- Update direction : $d^{(\ell)}[I^{(\ell)}]=\Phi_{I^{(\ell)}}^{+}r^{(\ell)}$,
and $d^{(\ell)}[?I^{(\ell)}]=0$,
- Update the solution : $w^{(\ell+1)}=w^{(\ell)}+\gamma d^{(\ell)}, 0<\gamma\le 1$,
- $\lambda_{\ell+1}=g(\lambda_\ell)$,
- Next iteration : $\ell\leftarrow\ell+1$.

Output : Reconstruct x from $w^{(\ell)}$.

[1] There is a bijection between ε and λ such that (P_{ε,ℓ_1}) and (Q_{λ,ℓ_1}) share the same solution set.

In case of noisy data, the LARS iterations are applied until the residual $r^{(\ell)}$ at iteration ℓ satisfies $\|r^{\ell}\|_2 \le \tau\sigma_n$, where $\tau > 1$. The noise standard deviation σ_n is either known in advance or can be estimated from the data, e.g., from the resting periods in the fMRI setting. For additive white Gaussian noise, by standard concentration inequalities on χ^2_T variables, the just-mentioned inequality on the residual norm holds with probability higher than $1-e^{-v\sqrt{T/8}}$ if $\tau = \sqrt{T}\sqrt{1+v\sqrt{2/T}}$.

The strictly decreasing function $g(\cdot)$ reflects the update schedule of λ_k, typically exponential or linear. The computation bottleneck of IT-LARS lies in solving the least-squares projection step to get d^k which can be done efficiently using a conjugate gradient (CG) solver involving multiplications by Φ or Φ^{T}. It can be shown that the computational complexity of IT-LARS is $O(S(l+2)V)$, where l is the number of CG iterations (10 were sufficient in our experiments), S is the number of IT-LARS iterations and V is the computational complexity of an application of the linear operator Φ^{T} or its adjoint.

3.5. Implementation

The analysis and synthesis operators Φ^{T} and Φ are never constructed explicitly. Rather, their application to a vector is implemented with fast implicit algorithms using the activelet filters at either the analysis or synthesis side of the efficient filterbank implementation as explained in Section 3.3, see also Fig. 3. Moreover, the scale-dependent filters of the activelet transform can be precomputed and stored. During this preliminary step, it is also easy to normalize the rows of the decomposition (analysis) matrix to have their ℓ_2-norm equal to 1, as required by the IT-LARS.

4. An alternative (linear) method

To establish a point of comparison, we propose in this section the best linear estimator that optimizes the minimum mean-squared error (MMSE). To that aim, we start by presenting an appropriate stochastic formulation for our problem. The optimal (linear) solution can then be found as the L*L smoothing spline.

4.1. Stochastic formulation

Suppose that the inter-event timings $t_l - t_{l-1}$ are independently and identically distributed (i.i.d.) random variables that follow an exponential distribution with parameter $\sigma_0 > 0$. Let the stimulus amplitudes c_l be i.i.d. random variables with known first- and second-order moments $\mathbf{E}\{c_l\} = \mu_1$ and $\mathbf{E}\{c_l^2\} = \mu_2$. We assume that c_l are independent of t_l. The *random Dirac spike train* process

$$s(t) = \sum_l c_l \delta(t - t_l) \tag{10}$$

is wide-sense stationary with mean $\mu_1\sigma_0$ and autocorrelation function $C_{ss}(\tau) = \mathbf{E}\{s(t)s(t+\tau)\} = \mu_1^2\sigma_0^2 + \mu_2\sigma_0 \cdot \delta(\tau)$. The power spectrum is then given by $P_{ss}(\omega) = \mathcal{F}\{C_{ss}\}(\omega) = \mu_1^2\sigma_0^2\delta(\omega) + \mu_2\sigma_0$. By subtracting the mean from (10), we get a zero-mean, white *innovation signal* $i(t) = s(t) - \mu_1\sigma_0$, with $C_{ii}(\tau) = \mu_2\sigma_0 \cdot \delta(\tau)$.

If the random Dirac impulse train process drives the (linear, shift-invariant) hemodynamic system, we can rewrite the activity-related BOLD response as

$$x(t) = (s * h)(t) \tag{11}$$

$$x(t) = \sum_l c_l h(t - t_l) \tag{12}$$

$$x(t) = (i * h)(t) + \underbrace{\mu_1\sigma_0 \int_{-\infty}^{\infty} h(t)}_{h_0}. \tag{13}$$

The key observation to make is that the re-centered process $x(t) - h_0$ is regular; i.e., it admits a whitening operator that turns it into an uncorrelated, zero-mean process. Indeed, using the operator L associated with the hemodynamic system, we have $\mathrm{L}\{i*h\}(t) = (i*\mathrm{L}\{h\})(t) = (i*\delta)(t) = i(t)$, which is white and zero-mean. Equivalently, we can also say $\mathrm{L}^*\mathrm{L}\{C_{(x-h_0)(x-h_0)}(\tau)\} = \mu_2\sigma_0 \cdot \delta(\tau)$, where L* is the adjoint of the operator L.

4.2. L*L smoothing spline solution

We measure the noisy samples $y[k] = x(k) + n[k]$, where the fMRI noise $n[k]$ follows $AR(\rho)$—a first-order autoregressive model with autoregressive parameter ρ and associated power spectrum $P_{nn}(\omega) = \sigma^2/|1-\rho e^{-j\omega}|^2$, $\rho = 0.2$. We can define the re-centered measurements as $y_0[k] = y[k] - h_0$. The best linear estimator $\tilde{x}(t') = \sum_k b_{t_0}[k]y_0[k]$ of $x(t')$, given the samples $\{y_0[k]\}_k$, will then be the one that minimizes the mean-squared error

$$\min_{\tilde{x}(t')} E[|x(t') - \tilde{x}(t')|^2]. \tag{14}$$

We can make use of the generalized smoothing spline theory [38] to solve (14). All operators of the form (6) with $N > M$ are spline-admissible [39], hence the operator L is associated to the hemodynamic system. Consequently, according to [38], the best linear MMSE estimator $\tilde{x}(t)$ of (14) given the noisy and re-centered measurements $y_0[k]$ of the underlying continuous-time stationary process $x(t) - h_0$ with whitening operator L is obtained by the L*L smoothing spline, corresponding to [38, Theorem 5]

$$\tilde{x}(t) = \sum_{k\in\mathbb{Z}} (b_{\gamma,\sigma} * y_0)[k]\varphi(t-k). \tag{15}$$

Here, $\varphi(t)$ is the L*L exponential B-spline with parameters $\{\vec{\alpha}, -\vec{\alpha}^*\}$, $\{\vec{\gamma}, -\vec{\gamma}^*\}$, and the digital filter $b_{\gamma,\sigma}$ is given by its z-transform

$$B_{\gamma,\sigma}(z) = \frac{1}{\sum_{k\in\mathbb{Z}}\varphi(k)z^{-k} + \dfrac{\sigma^2}{\mu_2\sigma_0|1-\rho z^{-1}|^2}\Delta_L(z)\Delta_L(z^{-1})}, \tag{16}$$

where we used the exponential B-spline localization filter $\Delta_{\mathrm{L}}(z) = \prod_{k=1}^{N}(1 - e^{\alpha_k}z^{-1})$. In other words, the smoothing spline solution that solves (14) reverts to linear filtering of the re-centered measurements. We implemented the digital filters in (15) using the causal–anticausal decomposition, as described in [38, Appendix II].

The smoothing spline solution is the best linear MMSE estimator, and it would even be the best possible estimator if the underlying process were Gaussian. As this is not the case, we can expect non-linear estimator to do better. Nevertheless, the linear solution can be considered as a gold standard for signal processing problems, and thus it is a worthful point of comparison. Notice that the linear solution can be applied both to traditional B-splines (only zeros at the origin) or the exponential splines associated to the activelet operator.

5. Results

5.1. Synthetic data

We first evaluated the performance of our method on synthetic fMRI data set, consisting of 100 sequences generated according to the model (3). Five short events of length 0.8 s were modeled with inter-event delays $t_{l+1}-t_l$ following an exponential distribution with mean interstimulus delay of 40 s. Each event was convolved with a hemodynamic response (HRF), for which the parameters of the underlying balloon model were randomly generated following Gaussian distributions; the mean and standard deviation were taken from the experimentally measured histograms in [29], and are listed in Table 2 in the Appendix. Correlated noise following the model $AR(0.2)$, with $\sigma_n = 0.3$, leading to an average SNR of -7 dB, and a baseline consisting of a constant and a slowly varying sinusoid of random amplitude (frequency around 0.01 Hz) were added to the data.

We compared the activelet-domain non-linear sparse recovery algorithm with two linear techniques: one corresponds to the linear MMSE signal estimation as described in Section 4, another is the "wavelet Wiener" solution which is the unique minimizer of

$$\tilde{x} = \underset{x}{\operatorname{argmin}} \|y-x\|_2^2 + \sum_i \gamma_i \|T_i x\|_2^2,$$

where T_i computes the i-th scale of the wavelet transform coefficients and the scale-dependent regularization coefficients γ_i are chosen by an Oracle:

$$\{\gamma_1, \ldots, \gamma_J\} = \underset{\{\gamma_1,\ldots,\gamma_J\}}{\operatorname{argmin}} \|x-\tilde{x}\|_2.$$

MMSE and wavelet Wiener techniques also use an Oracle to remove the baseline. Additionally we compared each method with its traditional B-spline counterpart.

In the activelet non-linear sparse recovery method, we propose to use prior information on the second-order statistics of the signal and the noise to select an appropriate weighting operator W in (8). As a first approach, W may be taken as the inverse of the square-root of the noise covariance matrix resulting in a Mahalanobis-like data fidelity term in (8). As the noise here is wide-sense stationary, this is equivalent to taking $\mathcal{F}(W)(\omega) = P_{nn}^{-1/2}(\omega)$. Another idea is to minimize the residual in the Wiener domain, where the SNR of the measurements is maximized. Specifically, we have

$$\mathcal{F}(W)(\omega) = \frac{P_{hh}(\omega)}{P_{hh}(\omega)+P_{nn}(\omega)},$$

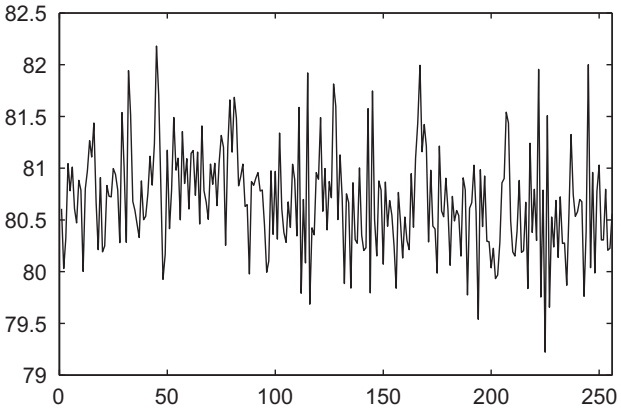

Fig. 4. A sample synthetic time-course.

where $P_{hh}(\omega)$ and $P_{nn}(\omega)$ correspond to the power spectra of the HRF and the AR noise model, respectively. We use the activelets dictionary with three decomposition levels. Additionally, we add low-scale B-splines to the dictionary to capture the baseline.

The computations were done on a 2 GHz Intel Core Duo MacBook Pro computer using Matlab 7 of MathWorks, WaveLab 802 [40] and SparseLab 100 [41]. In Fig. 4, we show a sample synthetic time-course. The results of the applied methods are shown in Fig. 5. The corresponding SNR levels are given in Table 1.

Interestingly, the sparse recovery with the non-linear method performs the best when the basis functions are well-tuned to the system response, as is the case with the activelets dictionary. However, it strongly degrades otherwise. The single-trial variability that was introduced in the synthetic data is dealt with in a satisfactory way by our approach. Apparently, the sparsifying properties of the operator are not too much affected by these deviations. Among the linear methods, the MMSE estimator performs the best when the model is well chosen. The performance of the Oracle-driven wavelet Wiener is the worst even when using the activelets, and gets worse with standard wavelets.

5.2. fMRI experimental data

We have applied our method to the dataset from a traditional fMRI experiment. The subject was scanned in a Siemens Magnetom 3T Scanner. The visual stimulation consisted of 10 flashing checkerboard excitations (duration = 500 ms) with varying interstimulus timings, followed by a resting period, during which the subject closed the eyes. Two hundred and fifty six scans were performed with TE = 30 ms, TR = 1 s and voxel size $2.6 \times 1.8 \times 5$ mm. We used the Statistical Parameter Mapping (SPM, http://www.fil.ion.ucl.ac.uk/spm/) Matlab package to do standard pre-processing of the dataset. This operation included realignment, co-registration and spatial Gaussian smoothing (FWHM = 8 mm) of the data.

We computed the SPM parameter map for the F-test of the effects-of-interest (see Fig. 6, right column) and picked three time-courses that corresponded to a strongly

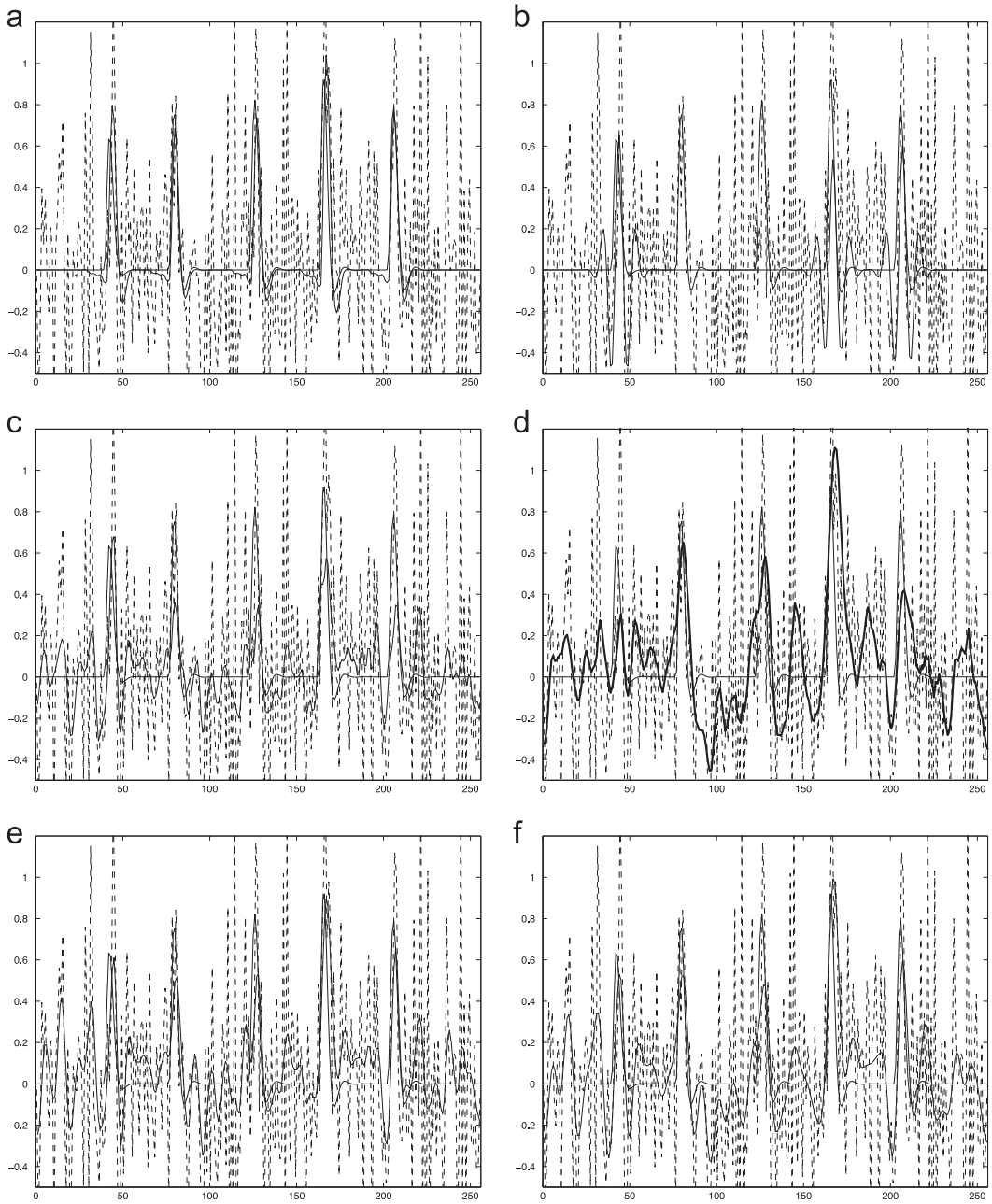

Fig. 5. Example of synthetic data: activelets versus B-spline wavelets; sparse recovery with wavelets, linear wavelet and linear MMSE solutions. Thin line: original signal, dotted line: noisy measurements, bold line: estimation. (a) Activelets with non-linear sparse recovery. (b) B-spline wavelets with non-linear sparse recovery. (c) Activelet Wiener solution. (d) B-spline wavelet Wiener solution. (e) Activelet-spline MMSE solution. (f) B-spline MMSE solution.

Table 1
SNR values (mean $\pm$ standard deviation over several realizations of noise, measured in dB) for different estimation methods.

	MMSE	Wavelet Wiener	Sparse recovery wavelet
Activelets	406 ± 0.78	3.76 ± 0.89	6.62 ± 1.66
B-spline wavelets	3.56 ± 0.77	1.08 ± 0.81	2.27 ± 1.24

active voxel, a weakly active voxel (around 5% familywise error threshold) and a non-active voxel.

Finally, we have computed the energy of the detected activity signal for the three most active slices. To this end, the activelet algorithm was optimized for real-data processing—the estimated baseline was updated at each iteration by projecting the residue on a set of low-scale B-splines. The signal from the resting period was used to estimate the noise level. The results are shown in Fig. 6. Note that the activelets are able to identify the spatial location of the active voxels from the data itself, while SPM uses the knowledge of the event timing.

6. Discussion

The problem of "blind" activity detection in fMRI is difficult due to the presence of strong disturbance

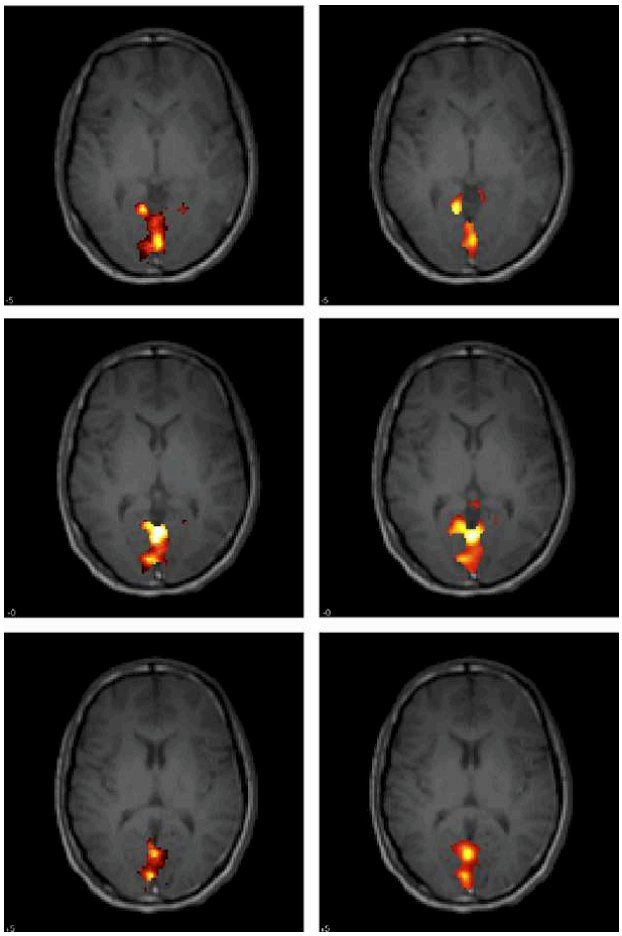

Fig. 6. Activation maps for the three most active slices superposed on the corresponding anatomical T_1 images. Left column: energy of the activity signal as detected by activelets. Right column: corresponding contrast obtained by standard regression analysis in SPM2 with known onset times, threshold = 20%.

components. Even in traditional fMRI data analysis, where the onset times are known, an important compromise on the flexibility of the model has to be made. As an extreme option, one could test the measurement for the presence of activity-related signal of the form (3) with fixed weights and onsets. Obviously, this approach is robust against false positive (FP) detections, but it loses all temporal resolution. At the same time, the variability of the HRF (which is known not only to vary over space and time for the same subject, but also between subjects) might lead to failures in detecting the activity that is actually present but different from the model; this situation is known as a false negative (FN) response. State-of-art fMRI analysis software (e.g., SPM toolbox for Matlab) employ additional regressors, such as derivatives of the model with respect to its parameters, to account for the BOLD variability and shift invariance [42]. More regressors in the GLM decreases the degrees of freedom and thus boils down to trading some of the FP rate for the improved FN rate. Friman et al. use the "optimal" FN/FP trade-off as a criterion to choose appropriate regressors [43].

To validate the proposed paradigm, we did compare the signal recovery capability of the activelets, the wavelet Wiener and MMSE algorithms. The latter was applied with a wide range of regularization parameter γ, which can be linked to the specificity level of GLM methods. Large values of γ force the solution to zero, resulting in no sensitivity but high specificity. As γ decreases, the algorithm eventually fits the measurements. The MMSE approach for the activelets or B-spline wavelet bases shows low specificity. Nevertheless, the activelets basis maintains some advantage compared to traditional wavelets (cf. Table 1, MMSE results), while, as we see in Fig. 5, there is no visual improvement in the recovered signal.

One of the key results of our paper is that we improve the FN/FP trade-off by employing the non-linear sparse recovery technique. Comparing Fig. 5(a) to Fig. 5(b), we note that, in this case, the use of the proper (activelets) basis becomes crucial; the B-spline wavelets cannot fit the signal well, and consequently show a very low SNR. Importantly, the non-linear activelet-based sparse recovery reveals no activity-related signal in the areas that contain pure noise. This is not the case for the linear methods, meaning that they suffer from a high FP rate.

The improvement in the FN/FP trade-off is best seen on the ROC curve that is plotted in Fig. 8. Given the ground truth, we divide the signal into active and non-active intervals. The performance of the result is then measured by the mean-squared error (ε_k) for each interval k with respect to the ground truth. After normalization with the true mean signal μ during activation, we obtain the values $1-\varepsilon_k/\mu$, which can be related to the sensitivity (for active intervals) and the specificity (for non-active intervals) within the framework of a statistical decision taken for each interval. The two markers show the Oracle-driven wavelet Wiener results. All MMSE estimators use an Oracle to remove the baseline. Despite this non-negligible advantage, the results from the proposed activelets method with non-linear optimization-based sparse recovery are superior. Moreover, the parameters of the HRFs varied randomly while the activelets operator L remained fixed. The performance achieved by the sparse recovery algorithm with the activelet dictionary clearly demonstrates the versatility of our approach with respect to the shape of the HRF.

The results for fMRI experiments in Fig. 7 further demonstrate the suitability of our approach for the analysis of real-world data. For the most active voxel in Fig. 7(a) and (b), nine activations were detected out of 10; the undetected activation coincides with a sudden drop in the baseline BOLD signal. The additional detection (last peak of the thin line) happened right after the subject had closed the eyes. The activity-related signal in the weakly active voxel in Fig. 7(c) and (d) is almost entirely masked by noise; however, the algorithm still detects six activations. Importantly, the detection in the non-active voxel is zero. By comparing the spatial distribution of detected activity to the SPM contrast in Fig. 6, we see that our method is able to correctly identify the most active spots. In the areas where the SPM correlation is weaker, the maps are not necessarily identical: activelets "react" to individual activations separately, while regression analysis looks for voxels that respond at all onset times.

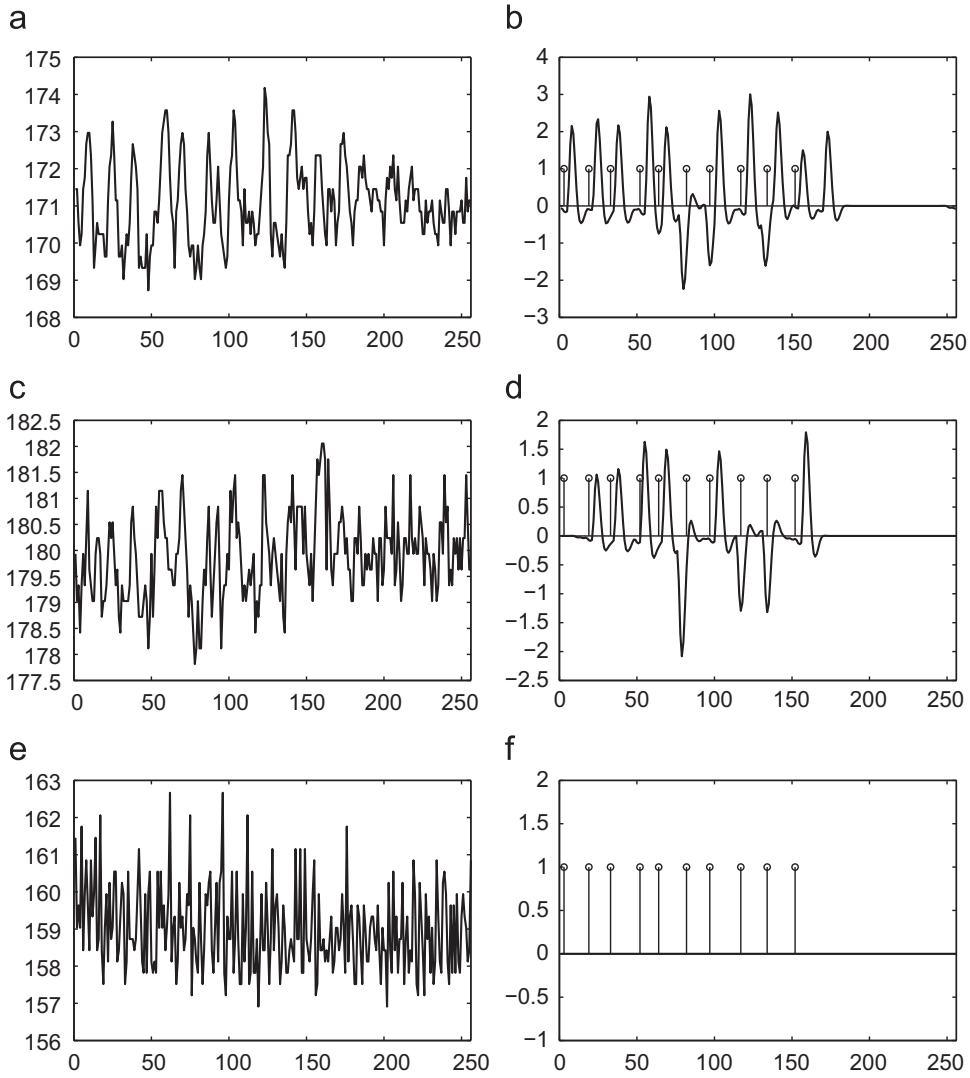

Fig. 7. FMRI experimental data and the activity-related signal extracted by the activelet method. The vertical lines correspond to onsets times. (a) Most active voxel: time-course. (b) Most active voxel: estimated activity-related signal. (c) Weakly active voxel: time-course. (d) Weakly active voxel: estimated activity-related signal. (e) Non-active voxel: time-course. (f) Non-active voxel: estimated activity-related signal.

The ℓ_1 minimization pursuit used here is an exploratory estimation approach. We deployed it to demonstrate the appropriateness of the activelets dictionary for hemodynamic signals. One future research direction is to develop a statistical detection step that would produce activation maps based on the activity-related signal estimated by the activelet framework. This would entail to elaborate a hypothesis testing step based on some score derived from the (asymptotic) statistical properties of the estimates provided by the non-linear method. In this respect, the work of Knight and Fu [44] gives some interesting indications. Another approach may be to combine estimation and inference (by traditional stochastic samples) in a Bayesian framework that explicitly uses the prior distribution of the discrete time input signal [45–47].

7. Conclusions

We proposed a framework for the recovery of activity-related signals in the fMRI measurements without knowledge of the actual onset times by exploiting their sparsity in time. The two key components of our method are: (1) specially tailored wavelet basis that allows sparse approximation of the activity-related signal and (2) a signal recovery procedure that is based on the optimizes the ℓ_1 of the expansion coefficients. We have demonstrated that both components are beneficial, and they should be combined for best results. The underlying assumption for the framework to be valid is that the events are sparsely distributed over time which is appropriate for slow event-related design. We have found the method to be robust to modeling errors; i.e., deviations from the idealized HRF response did not strongly affect the sparsifying properties of the differential operator L.

The next practical step is to apply the method to the elucidation of specific neurological questions. There is also room for improvement for making the algorithm faster so that it can be applied to large-scale data on a routine basis. It should be noted that the interest from the neuroimaging for paradigm-free mapping is increasing; e.g., in recent work [48] a dictionary of hemodynamic

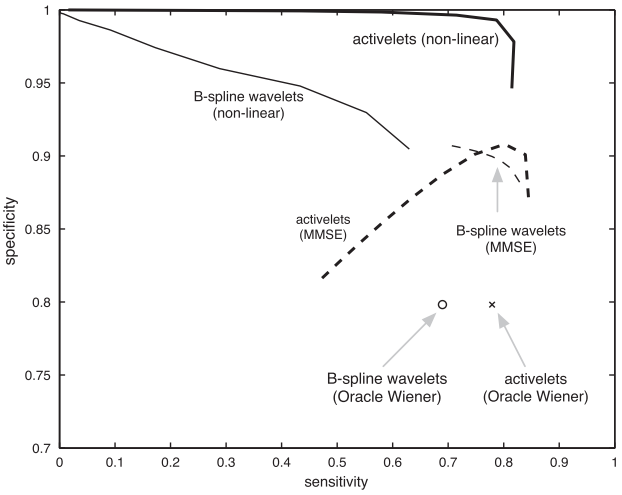

Fig. 8. ROC curves for the different methods. The importance of both the activelets dictionary and the non-linear sparse recovery are demonstrated; i.e., their combination leads to the best results that outperform the others.

response functions was proposed in combination with ridge regression.

Acknowledgments

The authors would like to thank Cédric Vonesch for the interesting discussions. This work was supported by the Center for Biomedical Imaging (CIBM) of the Geneva—Lausanne Universities and the EPFL, the foundations Leenaards and Louis-Jeantet, as well as by the Swiss National Science Foundation under Grants 200020-109415 and PP00P2-123438.

Appendix A. BOLD response and hemodynamic system

The hemodynamic model that we use has been described by Friston et al. [29]. It is a combination of the balloon/windkessel model [30,49] with a model that links synaptic activity and changes in regional blood flow. Mathematically, this model corresponds to the non-linear state-space definition of a system with four state variables $\{s,f_{\text{in}},v,q\}$

$$\begin{cases} \dot{s} = \varepsilon u - \dfrac{s}{\tau_s} - \dfrac{f_{\text{in}}-1}{\tau_f}, \\ \dot{f}_{\text{in}} = s, \\ \dot{v} = \dfrac{1}{\tau_0}(f_{\text{in}} - v^{1/\alpha}), \\ \dot{q} = \dfrac{1}{\tau_0}\left(f_{\text{in}}\dfrac{1-(1-E_0)^{1/f_{\text{in}}}}{E_0} - v^{1/(\alpha-1)}q\right), \end{cases} \tag{17}$$

and one observed quantity (the BOLD signal)

$$\text{BOLD}_{\text{non-linear}} = V_0(k_1(1-q) + k_2\left(1-\frac{q}{v}\right) + k_3(1-v)). \tag{18}$$

The meaning of the various variables and parameters are summarized in Table 2. Note that all variables are expressed in normalized form; i.e., relative to resting values. Full details can be found in [29].

We define the variables $\{x_1,x_2,x_3,x_4\} = \{s, 1-f_{\text{in}}, 1-v, 1-q\}$. Linearization of (17) around the resting point $\{x_1, x_2, x_3, x_4\} = (0,0,0,0)$ gives

$$\begin{cases} \dot{x}_1 = \varepsilon u - \dfrac{x_1}{\tau_s} + \dfrac{x_2}{\tau_f}, \\ \dot{x}_2 = -x_1, \\ \dot{x}_3 = \dfrac{1}{\tau_0}\left(x_2 - \dfrac{x_3}{\alpha}\right), \\ \dot{x}_4 = cx_2 - \dfrac{1-\alpha}{\alpha\tau_0}x_3 - \dfrac{1}{\tau_0}x_4, \end{cases} \tag{19}$$

with $c = (1+(1-E_0)\ln(1-E_0)/E_0)/\tau_0$. We diagonalize the system by making use of the Gauss method:

$$\begin{cases} \left(\text{D}^2 + \dfrac{1}{\tau_s}\text{D} + \dfrac{1}{\tau_f}\text{I}\right)\{x_2\} = -\varepsilon u, \\ \left(\text{D}^2 + \dfrac{1}{\tau_s}\text{D} + \dfrac{1}{\tau_f}\text{I}\right)\{x_1\} = \varepsilon \text{D}u, \\ \left(\text{D} + \dfrac{1}{\alpha\tau_0}\text{I}\right)\left(\text{D}^2 + \dfrac{1}{\tau_s}\text{D} + \dfrac{1}{\tau_f}\text{I}\right)\{x_3\} = -\dfrac{\varepsilon}{\tau_0}u, \\ \left(\text{D} + \dfrac{1}{\tau_0}\text{I}\right)\left(\text{D} + \dfrac{1}{\alpha\tau_0}\text{I}\right)\left(\text{D}^2 + \dfrac{1}{\tau_s}\text{D} + \dfrac{1}{\tau_f}\text{I}\right)\{x_4\} = -c\varepsilon\text{D}u + \left(\dfrac{1-\alpha}{\alpha\tau_0^2} - \dfrac{c}{\alpha\tau_0}\right)\varepsilon u. \end{cases} \tag{20}$$

Additionally, the linearized equation for the BOLD signal is

$$\text{BOLD}_{\text{linear}}(t) = V_0((k_1+k_2)x_4(t) + (k_3-k_2)x_3(t)).$$

Finally, the HRF $h(t)$ is obtained by setting $u(t) = \delta(t)$. It satisfies the differential equation

$$\begin{aligned} &\left(\text{D} + \frac{1}{\tau_0}\text{I}\right)\left(\text{D} + \frac{1}{\alpha\tau_0}\text{I}\right)\left(\text{D}^2 + \frac{1}{\tau_s}\text{D} + \frac{1}{\tau_f}\text{I}\right)\{h\} \\ &\quad = V_0(k_1+k_2)\left(-c\varepsilon\text{D}u + \left(\frac{1-\alpha}{\alpha\tau_0^2} - \frac{c}{\alpha\tau_0}\right)\varepsilon u\right) \\ &\quad -V_0(k_3-k_2)\left(\frac{\varepsilon}{\tau_0}\text{D}u + \frac{\varepsilon}{\tau_0^2}u\right) \end{aligned}$$

Table 2
List of the variables and constants involved in the hemodynamic model with their typical values.

Symbol	Meaning	Typical value	Standard deviation
u	Stimulus	–	–
s	Flow inducing signal	–	–
f_{in}	Blood flow	–	–
v	Normalized venous volume	–	–
q	Normalized deoxyhemoglobine content	–	–
ε	Neuronal efficacy	0.54	0.085
τ_s	Signal decay	1.54	0.169
τ_f	Autoregulation	2.46	0.212
τ_0	Transit time	098	0.169
α	Balloon stiffness	0.33	0034
E_0	Oxygen extraction fraction	0.34	0.043
V_0	Resting blood volume fraction	1	–
k_1	BOLD constant 1	$7E_0$	–
k_2	BOLD constant 2	2	–
k_3	BOLD constant 3	$2E_0-0.2$	–

The right-hand side can be further developed as

$$\frac{V_0\varepsilon}{\tau_0}\left((-(k_1+k_2)c\tau_0-k_3+k_2)\mathrm{D}u+\left((k_1+k_2)\left(\frac{1-\alpha}{\alpha\tau_0}-\frac{c}{\alpha}\right)-(k_3-k_2)\frac{1}{\tau_0}\right)u\right),$$

yielding the linear differential operator L of the form (6) with parameters (7).

References

[1] S. Ogawa, T.M. Lee, Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation, Magnetic Resonance in Medicine 16 (1990) 9–18.

[2] R. Frackowiak, K. Friston, C. Frith, R. Dolan, J. Mazziotta, Human Brain Function, Academic Press, 1997.

[3] M.E. Raichle, Behind the scenes of functional brain imaging: a historical and physiological perspective, Proceedings of the National Academy of Sciences 95 (3) (1998) 765–772.

[4] R.B. Buxton, L.R. Frank, A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation, Journal of Cerebral Blood Flow and Metabolism 17 (January) (1997) 64–72.

[5] K.J. Friston, C.D. Frith, P.F. Liddle, R.S.J. Frackowiak, Functional connectivity: the principal component analysis of large (pet) datasets, Journal of Cerebral Blood Flow and Metabolism 13 (1993) 5–14.

[6] M. McKeown, S. Makeig, G. Brown, T. Jung, S. Kindermann, A. Bell, T. Sejnowski, Analysis of fMRI data by blind separation into independent spatial components, Human Brain Mapping 6 (1998) 160–188.

[7] T. Adali, V.D. Calhoun, Complex ICA of brain imaging data, IEEE Signal Processing Magazine 24 (5) (2007) 136–139.

[8] G.H. Glover, Deconvolution of impulse response in event-related BOLD fMRI, NeuroImage 9 (4) (1999) 416–429.

[9] S. Makni, P. Ciuciu, J. Idier, J.-B. Poline, Joint detection-estimation of brain activity in functional MRI: a multichannel deconvolution solution, IEEE Transactions on Signal Processing 53 (9) (2005) 3488–3502.

[10] J.R. Duann, T.P. Jung, W.J. Kuo, T.C. Yeh, S. Makeig, J.C. Hsieh, T.J. Sejnowski, Single-trial variability in event-related BOLD signals, NeuroImage 15 (4) (2002) 823–835.

[11] S. Donnet, M. Lavielle, J.-B. Poline, Are fMRI event-related response constant in time? A model selection answer, NeuroImage (2006) 1169–1176.

[12] P. Baraldi, A. Manginelli, M. Maieron, D. Liberati, A. Porro, An ARX model-based approach to trial by trial identification of fMRI-BOLD responses, NeuroImage 37 (2007) 189–201.

[13] H. Luo, S. Puthusserypady, A sparse Bayesian method for determination of flexible design matrix for fMRI data analysis, IEEE Transactions on Circuits and Systems I 52 (12) (2005) 2699–2706.

[14] X. Li, D. Coyle, L. Maguire, T.M. McGinnity, D.R. Watson, H. Benali, A least angle regression method for fMRI activation detection in phase-encoded experimental designs, NeuroImage 52 (4) (2010) 1390–1400.

[15] A. Salek-Haddadi, B. Diehl, K. Hamandi, M. Merschhemke, A. Liston, K. Friston, J.S. Duncan, D.R. Fish, L. Lemieux, Hemodynamic correlates of epileptiform discharges: an EEG-fMRI study of 63 patients with focal epilepsy, Brain Research 1088 (1) (2006) 148–166.

[16] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, San Diego, CA, 1998.

[17] M. Vetterli, J. Kovacevic, Wavelets and Subband Coding, second ed. ⟨http://www.waveletsandsubbandcoding.org⟩: Open-Access Edition, 2007.

[18] D.L. Donoho, I.M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika 81 (3) (1994) 425–455[Online]. Available: citeseer.ist.psu.edu/donoho93ideal.html⟩.

[19] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM Journal on Scientific Computing 20 (1) (1999) 33–61.

[20] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Communications on Pure and Applied Mathematics 57 (2004) 1413–1457.

[21] 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 (1) (2009) 34–81.

[22] J.-L. Starck, F. Murtagh, M.J. Fadili, Sparse Signal and Image Processing: Wavelets, Curvelets and Morphological Diversity, Cambridge University Press, Cambridge, UK, 2010.

[23] D. Van De Ville, T. Blu, M. Unser, Surfing the brain: an overview of wavelet-based techniques for fMRI data analysis, IEEE Engineering in Medicine and Biology Magazine 25 (2) (2006) 65–78.

[24] M.J. Fadili, E.T. Bullmore, A comparative evaluation of wavelet-based methods for multiple hypothesis testing of brain activation maps, NeuroImage 23 (3) (2004) 1112–1128.

[25] M.J. Fadili, E. Bullmore, Wavelet-generalised least squares: a new BLU estimator of linear regression models with 1/f errors, NeuroImage 15 (2002) 217–232.

[26] V. Maxim, L. Sendur, M.J. Fadili, J. Suckling, R. Gould, R. Howard, E. Bullmore, Fractional Gaussian noise, functional MRI and Alzheimer's disease, NeuroImage 25 (1) (2005) 141–158.

[27] P. Ciuciu, P. Abry, C. Rabrait, H. Wendt, Log wavelet leaders cumulant based multifractal analysis of EVI fMRI time series: evidence of scaling in ongoing and evoked brain activity, IEEE Journal of Selected Topics in Signal Processing 2 (6) (2008) 929–943.

[28] M.J. Fadili, E. Bullmore, Penalized partially linear models using sparse representations with an application to fMRI time series, IEEE Transactions on Signal Processing 53 (9) (2005) 9.

[29] K.J. Friston, A. Mechelli, R. Turner, C.J. Price, Nonlinear responses in fMRI: the balloon model, Volterra kernels and other hemodynamics, NeuroImage 12 (2000) 466–477.

[30] R.B. Buxton, E. Wong, L.R. Frank, Dynamics of blood flow and oxygenation changes during brain activation: the balloon model, Magnetic Resonance in Medicine 39 (June) (1998) 855–864.

[31] I. Khalidov, M. Unser, From differential equations to the construction of new wavelet-like bases, IEEE Transactions on Signal Processing 54 (4) (2006) 1256–1267.

[32] T. Blu, M. Unser, A complete family of scaling functions: the (α,τ)-fractional splines, in: Proceedings of the 28th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03), IEEE, Hong Kong, April 2003.

[33] S. Mallat, A Wavelet Tour of Signal Processing, third ed., Elsevier, 2009.

[34] R. Tibshirani, Regression shrinkage and selection via the Lasso, Journal of the Royal Statistical Society 58 (1) (1996) 267–288.

[35] B. Efron, T. Hastie, I.M. Johnstone, R. Tibshirani, Least angle regression, Annals of Statistics 32 (2) (2004) 407–499.

[36] M.R. Osborne, B. Presnell, B.A. Turlach, A new approach to variable selection in least squares problems, IMA Journal of Numerical Analysis 20 (2000) 389–403.

[37] M.J. Fadili, J.-L. Starck, Sparse representation-based image deconvolution by iterative thresholding, in: F. Murtagh, J.-L. Starck (Eds.), Astronomical Data Analysis IV, Marseille, France, 2006.

[38] M. Unser, T. Blu, Generalized smoothing splines and the optimal discretization of the Wiener filter, IEEE Transactions on Signal Processing 53 (6) (2005) 2146–2159.

[39] M. Unser, Cardinal exponential splines: Part II—think analog, act digital, IEEE Transactions on Signal Processing 53 (4) (2005) 1439–1449.

[40] J. Buckheit, D. Donoho, Wavelab and reproducible research, in: Wavelets and Statistics, Springer-Verlag, Berlin, New York, 1995[Online]. Available: ⟨citeseer.ist.psu.edu/article/buckheit95 wavelab.html⟩.

[41] D. Donoho, I. Drori, V. Stodden, Y. Tsaig, Sparselab, ⟨http://sparselab.stanford.edu/⟩, December 2005. [Online]. Available: ⟨http://sparselab.stanford.edu/⟩.

[42] K. Friston, P. Fletcher, O. Josephs, A. Holmes, M. Rugg, R. Turner, Event-related fMRI: characterizing differential responses, NeuroImage 7 (1998) 30–40.

[43] O. Friman, M. Borga, P. Lundberg, H. Knutsson, Adaptive analysis of fMRI data, NeuroImage 19 (3) (2003) 837–845.

[44] K. Knight, W. Fu, Asymptotics for Lasso-type estimators, Annals of Statistics 28 (5) (2000) 1356–1378.

[45] J.J. Kormylo, J.M. Mendel, Maximum-likelihood detection and estimation of Bernoulli–Gaussian processes, IEEE Transactions on Information Theory 28 (1982) 482–488.

[46] Q. Cheng, R. Chen, T.-H. Li, Simultaneous wavelet estimation and deconvolution of reflection seismic signals, IEEE Transactions on Geosciences and Remote Sensing 34 (2) (1996) 377–384.

[47] F. Champagnat, Y. Goussard, J. Idier, Unsupervised deconvolution of sparse spike trains using stochastic approximation, IEEE Transactions on Signal Processing 44 (12) (1996) 2988–2998.

[48] 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, in press, doi:10.1002/hbm.21116.

[49] J.B. Mandeville, J.J. Marota, C. Ayata, G. Zaharchuk, M.A. Moskowitz, B.R. Rosen, R.M. Weisskoff, Evidence of a cerebrovascular post-arteriole windkessel with delayed compliance, Journal of Cerebral Blood Flow and Metabolism 19 (6) (1999) 679–689.