



ELSEVIER

Contents lists available at ScienceDirect

Medical Image Analysis

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

A 3-D spatio-temporal deconvolution approach for MR perfusion in the brain



Carole Frindel*, Marc C. Robini, David Rousseau

Université de Lyon, CREATIS; CNRS UMR5220; Inserm U1044; INSA-Lyon; Université Lyon 1, France
 INSA Lyon, 7 Avenue Jean Capelle, 69621 Villeurbanne, France

ARTICLE INFO

Article history:

Received 8 February 2013

Received in revised form 12 September 2013

Accepted 7 October 2013

Available online 16 October 2013

Keywords:

Acute stroke

Perfusion weighted MRI

Deconvolution

Spatio-temporal model

Tissue outcome prediction

ABSTRACT

We propose an original spatio-temporal deconvolution approach for perfusion-weighted MRI applied to cerebral ischemia. The regularization of the underlying inverse problem is achieved with spatio-temporal priors and the resulting optimization problem is solved by half-quadratic minimization. Our approach offers strong convergence guarantees, including when the spatial priors are non-convex. Moreover, experiments on synthetic data and on real data collected from subjects with ischemic stroke show significant performance improvements over the standard approaches—namely, temporal deconvolution based on either truncated singular-value decomposition or l_2 -regularization—in terms of various performance measures.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

With the growing development of multicomponent-imaging for biomedical practice, image data-structures usually combine spatial components with components of a different nature (e.g., spectral Toronov et al., 2007, temporal Jähne, 1993, tensorial LeBihan et al., 2001, or colorimetric Brekke et al., 2006). Simple processing techniques consider the spatial components independently, but more accurate results can be achieved by joint approaches which consider the spatial components simultaneously along with their dependencies. Such joint approaches have been proposed in functional MRI for image restoration (Descombes et al., 1998), in contrast-enhanced MRI for spatio-temporal reconstruction (Schmid et al., 2006), in diffusion-tensor imaging for tensor-field denoising (Heim et al., 2007), and in imaging spectroscopy for classification and segmentation (Plaza et al., 2009).

In this paper, we propose a joint approach for the temporal deconvolution problem that arises in cerebral perfusion-weighted imaging (PWI) of acute ischemic stroke—in the western world, stroke represents a real public health issue for it is the second leading cause of death (Donnan et al., 2008). The amount of tissue damage caused by acute ischemic stroke is a continuum between irreversibly damaged tissues (infarct core) and benign oligemia.

* Corresponding author at: Université de Lyon, France. Tel.: +33 4 72 43 75 48; fax: +33 4 72 43 85 26.

E-mail address: carole.frindel@creatis.insa-lyon.fr (C. Frindel).

Within this continuum, the tissues in the penumbra (magnetically silent, yet possibly viable tissues) present a risk of infarction but may recover proper irrigation using specific therapeutics. The clinical challenge lies in a two-class classification problem: identifying viable tissue regions from infarct core to aid clinical decision-making and to improve long-term patient outcome. PWI provides critical real-time information about ongoing tissue injury by tracking the first pass of an injected contrast-agent (e.g., gadolinium) using T2*-weighted MRI. The data consists of the temporal signals of contrast-agent concentration in the voxels of a volume of interest. After deconvolution, these signals are post-processed to obtain maps of perfusion parameters which are used for interpretation; typical parameters are the cerebral blood volume (CBV), the cerebral blood flow (CBF), the mean transit time (MTT) and the time-to-peak of the residue function (T^{\max}) (Østergaard et al., 2009; Ritzenthaler et al., 2011). Several processing techniques have been proposed to generate perfusion-parameter maps. They are based on the deconvolution of the concentration signals by the local arterial input function (AIF) to account for the dispersion of the injected bolus and for individual differences. In this context, the most popular deconvolution method is based on truncated singular value decomposition (TSVD) (Østergaard et al., 1996b,a). However, state-of-the-art TSVD techniques for PWI do not exploit the spatio-temporal nature of the data (Østergaard et al., 1996b; Østergaard et al., 1996a; Andersen et al., 2002; Wu et al., 2003; Calamante et al., 2003; Mouridsen et al., 2006); they treat each concentration signal independently from the others, thus

neglecting the spatial correlation due to the non-random organization of microvascular circulation in the brain. Efforts have been made recently to fill this gap: spatially-correlated image models have been used in perfusion CT quantification of rectal tumors and in cardiovascular perfusion MRI (He et al., 2010; Schmid, 2011) (these methods are reviewed in Section 3). Yet these approaches do not have strong convergence guarantees—namely, convergence of the iterates to a critical point of the cost function for the deterministic approach in He et al. (2010), and equilibrium of the Markov chain for the stochastic approach in Schmid (2011)—which is important for clinical applications of perfusion quantification.

Our contributions are the following: (i) we propose a theoretically grounded spatio-temporal model for the PWI deconvolution problem, (ii) we provide a globally convergent algorithm to solve the associated optimization problem, and (iii) we show that our approach outperforms the standard (temporal-only) deconvolution methods using both synthetic and real data. Our experiments demonstrate that the proposed method gives more accurate estimates of the residue functions than do the reference TSVD approach (Østergaard et al., 1996b,a) and temporal-only regularization—this is all the more true when the signal-to-noise ratio (SNR) is low, as is the case in necrotic and partially damaged tissues. In the synthetic data case, the validation is performed in terms of the peak SNRs of the computed residue functions and of the estimated CBF, MTT and T^{\max} maps. In the real data case, we perform a receiver operating characteristic (ROC) analysis using final infarct regions delineated from follow-up imaging one month after the stroke.

The paper is organized as follows. Section 2 outlines the medical application field, Section 3 reviews spatio-temporal image models in the context of PWI, and Section 4 describes the theoretical background for perfusion parameter estimation. Our spatio-temporal model is presented in Section 5, where the PWI deconvolution problem is formulated in terms of the minimization of a specific cost function, and Section 6 is devoted to the construction of the optimization algorithm along with the study of its convergence properties. Section 8 describes the experimental setup and the criteria considered for performance evaluation. Our experiments on synthetic and real data are presented in Section 9, followed by concluding remarks.

2. Medical context

Blood flow is critical to the functioning of any organ for it provides oxygen and essential nutrients. In the case of blood flow disruption in the brain, cerebral autoregulation seeks to maintain the circulation by altering either the blood flow or the blood volume, or both. There are some well-defined thresholds on the CBF in normal, partially damaged and necrotic tissues. For instance, in the gray matter, the CBF ranges between 50 and 60 mL/100 g/min in normal tissues (Gonzalez, 2006) and falls below 20 mL/100 g/min in damaged regions. Severe reduction of the blood flow in a vessel, or ischemia, is often caused by thrombosis or arterial embolism; it leads to an interruption of the oxygen supply of the tissue perfused by the clotted vessel and subsequently to a stroke. Consequently, the primary goal of acute stroke treatment is to restore normal blood flow as soon as possible. This is performed by the use of thrombolytic drugs which must be administered as soon as possible to ensure the best chance of restoring the function of the damaged tissues. Since the use of such drugs is contraindicated in hemorrhagic stroke, early and accurate diagnosis is extremely important.

PWI is used in combination with other imaging techniques for guiding therapeutic decisions. It allows to determine which

areas of the brain are irreversibly injured from a stroke and which areas remains at risk. By way of illustration, Fig. 1 shows parameter maps extracted from a PWI exam of the brain of an 80-year-old male stroke patient (a CT angiography scan indicated an occlusion of the sylvian segment of the middle cerebral artery on the right side). The T^{\max} map reveals a damaged region that corresponds to the most severely affected tissues, which is also clearly visible in the MTT map. The interpretation of these parametric maps guides the recanalization of the occluded vessel.

PWI can be performed with or without contrast agent (Luypaert et al., 2001). PWI with contrast agent often uses gadolinium-based tracers and is generally preferred over PWI without contrast agent for its lower sensitivity to motion and its higher SNR. Both spin echo (SE) and gradient echo (GRE) sequences have been applied successfully to PWI, but GRE is used more frequently because it yields higher SNR (Heiland et al., 1998; Wintermark et al., 2005). Perfusion signals vary over a 5–10 s period once the contrast-agent is released into the brain tissues (which occurs 7–10 s after intravenous injection). Therefore, fast imaging is required for accurate perfusion measurements with sufficient time resolution. The PWI processing pipeline is the following: first, the imaging sequence captures a temporal series of perfusion volumes which constitutes a 4-D spatio-temporal data-structure; second, the data structure is used to generate functional maps of the in vivo hemodynamics of the brain (Rosen et al., 1991), some examples of which are displayed in Fig. 1; and third, the functional maps are used to measure and to characterize the physiological parameters of interest.

3. Related work

Spatio-temporal image models have been introduced recently in the context of perfusion parameter estimation (He et al., 2010; Schmid, 2011). These methods are briefly described below with an emphasis on the differences with our approach.

In He et al. (2010), He et al. model the 4-D spatio-temporal data-structure as a piecewise-smooth function without distinction between spatial and temporal components. They use two regularization terms: the first one penalizes the gradient within homogeneous 4-D regions and the second one penalizes the edge-field representing the boundaries between these regions. Their approach is inspired from that of Mumford and Shah (1989), but is eventually quite different from it and has no convergence guarantee. By contrast, our method—described in the next three sections—has the following advantages: (i) it distinguishes the temporal dimension from the spatial dimensions, which allows to adjust the strength of the regularization in the spatial domain independently from that in the temporal domain; and (ii) it provides strong convergence guarantees.

In Schmid (2011), Schmid models the spatial dependencies between contrast-agent concentration signals via a Bayesian model, and he proposes to sample the resulting posterior distribution using a Markov chain Monte Carlo (MCMC) method. This stochastic approach is intrinsically different from ours, which is deterministic, and is computationally much more expensive. To fix ideas, the computation time reported in Schmid (2011) is 3–5 min for processing a 108×256 slice on a standard quad-core PC without parallelization, whereas the Matlab implementation of our algorithm takes about one minute to process an entire $128 \times 128 \times 20$ volume on a single core of a machine with an Intel Core 8 CPU (2.20 GHz) and 8 GB of RAM. Besides, there is no practical bound on the mixing time of the Markov chain generated by the MCMC algorithm proposed by Schmid, which makes

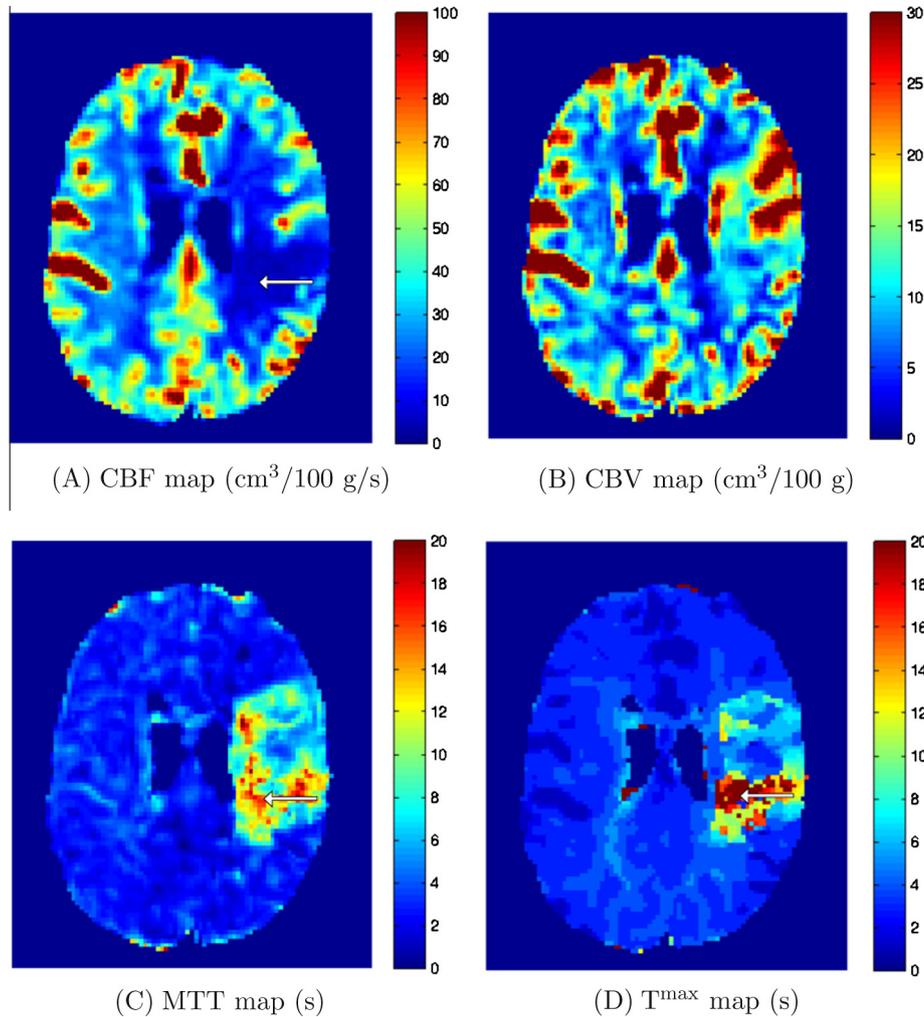


Fig. 1. Examples of perfusion parameter maps. The ischemic stroke lesion is pointed out by arrows.

the computation time very difficult to control. Our approach does not have this drawback: the computation time is efficiently controlled by adjusting the upper bound on the difference between subsequent iterates which is used for stopping the algorithm.

4. Data formation model

4.1. Noise-free model

We use the standard perfusion model proposed in Østergaard et al. (1996b). More precisely, the temporal signal C_v of the average contrast-agent concentration in a volume element v is obtained by convolving the AIF C_a with the residue function R_v representing the fraction of contrast agent still present in the capillaries in v :

$$C_v(t) = \alpha \rho_v \text{CBF}_v \int_0^t C_a(\tau) R_v(t - \tau) d\tau, \quad (1)$$

where ρ_v is the brain tissue density in v , CBF_v the regional blood flow in v , and the constant α depends on the hematocrit levels in the arteries and in the capillaries. Strictly speaking, α is a function of v , but this dependency is usually ignored because it cannot be determined in practice (Schreiber et al., 1998).

Eq. (1) is central for determining perfusion parameters in cerebral tissues—its variables are summarized in Table 1. The objective

Table 1

Variables of the temporal contrast-agent concentration model given in (1).

Variable	Known	Unit	Denomination
$C_v(t)$	Yes	mM	Average contrast-agent concentration
ρ_v	Yes	g/cm ³	Mean density
CBF_v	No	cm ³ /100 g/s	Blood flow
$C_a(t)$	Yes	mM	Arterial input function
$R_v(t)$	No	No unit	Residue function

is to estimate the signal $\text{CBF}_v \cdot R_v$ by deconvolving its observation C_v for each voxel v in a volume of interest. Given such an estimate, CBF_v is obtained by setting $t = 0$ (as $R_v(0) = 1$), and we can compute the MTT using the central volume theorem (Meier and Zierler, 1954):

$$\text{MTT}_v = \frac{\text{CBV}_v}{\text{CBF}_v} \quad \text{with} \quad \text{CBV}_v = \frac{\int_0^\infty C_v(t) dt}{\int_0^\infty C_a(t) dt}. \quad (2)$$

These parameters are supplemented by T_v^{max} —the time needed for the arterial blood to arrive in v —which is obtained by tracking the maximum of the estimate of $\text{CBF}_v \cdot R_v$. We refer to Fieselmann et al. (2011) for more details on the meaning and the computation of perfusion parameters.

4.2. Noise statistics

In practice, the observation of the contrast-agent concentration signal C_v is obtained from the MR signal S_v in voxel v by assuming a linear relationship with the change in the transverse relaxation rate (Østergaard et al., 1996a):

$$D_v(t) = -\frac{\kappa}{TE} \ln \left(\frac{S_v(t)}{S_v(0)} \right), \quad (3)$$

where D_v denotes the observation of C_v , κ is a constant, and TE is the echo time. Fig. 2 displays typical observations of contrast-agent concentration signals along with gamma-fits.

According to Gudbjartsson and Patz (1995), $S_v(t)$ is the realization of a random variable

$$\mathcal{X}_v(t) \sim \text{Rice}(A_v(t), \sigma), \quad (4)$$

where $A_v(t)$ is the intensity in the absence of noise and σ is the standard deviation of the Gaussian noise in the real and imaginary parts of the complex MR signal (we assume that σ does not vary with time). It can be observed that $\mathcal{X} \sim \text{Rice}(v, \sigma)$ is approximately Gaussian with mean v and standard deviation σ if $v \geq 3\sigma$, and that $\ln \mathcal{X}$ is approximately Gaussian with mean $\ln(v)$ and standard deviation σ/v if $v \geq 10\sigma$. Therefore, if $A_v(t) \geq 10\sigma$ for all $t \geq 0$, we can use the model

$$D_v(t) = -\frac{\kappa}{TE} \left(\ln \left(\frac{A_v(t)}{A_v(0)} \right) + \zeta_v(t) + \zeta_v(0) \right), \quad (5)$$

where $\zeta_v(t)$ is the realization of a random variable $\mathcal{Y}_v(t) \sim \mathcal{N}(0, \sigma^2/A_v^2(t))$, or equivalently,

$$D_v(t) = C_v(t) + \eta_v(t), \quad (6)$$

where $\eta_v(t)$ is the realization of a random variable

$$\mathcal{Z}_v(t) \sim \mathcal{N}(0, \zeta_v^2(t)) \quad \text{with} \quad \zeta_v^2(t) \propto \sigma^2 \left(\frac{1}{A_v^2(t)} + \frac{1}{A_v^2(0)} \right). \quad (7)$$

In the case of our application, $A_v(t) \geq 10\sigma$ for all the voxels in the background and independently of the sample time, which justifies the approximation of the noise by Gaussian random variables. The validity of this approximation is also supported by (i) the graphical comparison of the histogram of the noise in healthy tissues with a Gaussian distribution, as depicted in Fig. 3, and by

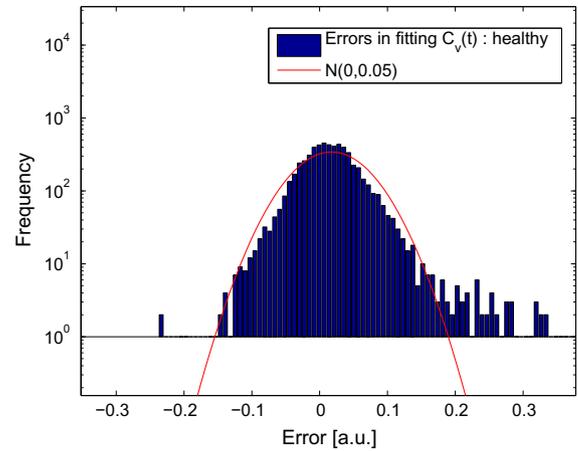


Fig. 3. Histogram of the residuals (i.e., fitting errors) associated with the gamma-fit of the “healthy” contrast-agent concentration signal depicted in blue in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

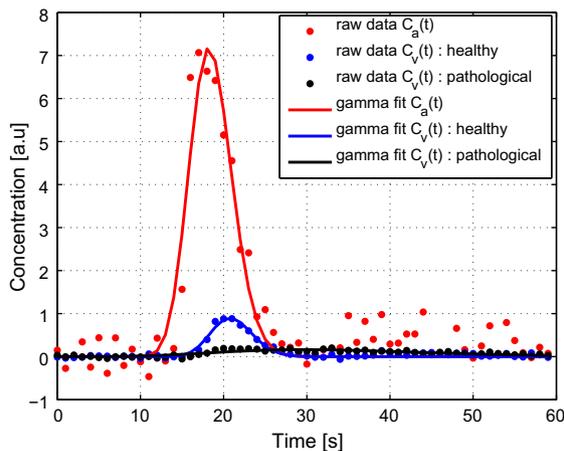
(ii) Lilliefors tests: for a significance level of 0.05, the p -value is about 0.4 in healthy tissues and about 0.5 in damaged tissues.

5. Construction of the cost function

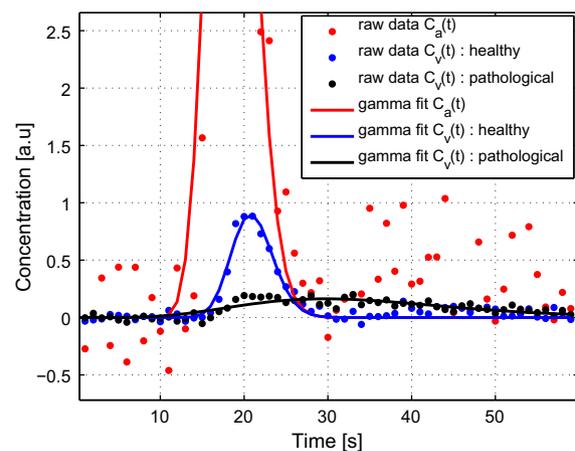
This section is devoted to the construction of the cost function whose global minima define the set of solutions to our spatio-temporal reconstruction approach to PWI; so we naturally begin with the discretization of the data formation model (Section 5.1). The proposed cost function is the sum of a data-fidelity term, a temporal-regularization term and a spatial-regularization term which are successively described in Sections 5.2, 5.3 and 5.4 along with compact notations that will be used in the sequel. The global expression of the cost function is given in Section 5.5, where we also establish useful properties for studying the convergence of the proposed optimization algorithm.

5.1. Discretization

The spatial domain is partitioned into a set of voxels $\{v_k; k \in \llbracket 1, K \rrbracket\}$ ($\llbracket a, b \rrbracket$ is a shorthand notation for $\{n \in \mathbb{Z} | a \leq n \leq b\}$), and the contrast-agent concentration signals C_{v_1}, \dots, C_{v_K} associated



(A) Signal examples



(B) Magnification of (A)

Fig. 2. Typical AIF and contrast-agent concentration signals in healthy and damaged tissues (in arbitrary units). The red curve represents the AIF (that is, the contrast-agent concentration in the arterial input), the blue curve corresponds to normal brain perfusion, and the black curve illustrates decreased and delayed perfusion in damaged tissues. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with these voxels (see (1)) are sampled with a temporal sampling period Δ_t . The temporal samples are indexed by n , and we let N denote the number of samples of each signal C_{v_k} . For simplicity, we set $C_k := C_{v_k}$, $R_k := R_{v_k}$ and $F_k := \text{CBF}_{v_k}$ for all k . Then, letting $t_n := n\Delta_t$ for all n , the discrete signal $(C_k(t_n))_{n \in \llbracket 1, N \rrbracket}$ associated to a voxel v_k is defined by

$$C_k(t_n) := C_{v_k}(n\Delta_t) = \alpha \rho_k F_k \int_0^{n\Delta_t} C_a(\tau) R_k(n\Delta_t - \tau) d\tau. \quad (8)$$

We consider the discrete data formation model defined by the trapezoidal approximation to this integral: since $C_a(0) = 0$ and $R_k(0) = 1$, we have

$$\begin{cases} C_k(t_1) \approx \frac{1}{2} \alpha \rho_k F_k \Delta_t C_a(t_1), \\ \forall n \in \llbracket 2, N \rrbracket, C_k(t_n) \approx \alpha \rho_k F_k \Delta_t \left(\frac{1}{2} C_a(t_n) + \sum_{i \in \llbracket 1, n-1 \rrbracket} C_a(t_i) R_k(t_n - t_i) \right), \end{cases} \quad (9)$$

or, equivalently, in matrix notation,

$$\mathbf{c}_k \approx \alpha \rho_k F_k \Delta_t \mathbf{C}_a \mathbf{r}_k \quad (10)$$

with $\mathbf{c}_k = [C_k(t_1), \dots, C_k(t_N)]^T$, $\mathbf{r}_k = [1, R_k(t_1), \dots, R_k(t_{N-1})]^T$ and

$$\mathbf{C}_a = \begin{pmatrix} \frac{1}{2} C_a(t_1) & 0 & 0 & \cdots & 0 \\ \frac{1}{2} C_a(t_2) & C_a(t_1) & 0 & \cdots & 0 \\ \frac{1}{2} C_a(t_3) & C_a(t_2) & C_a(t_1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} C_a(t_N) & C_a(t_{N-1}) & C_a(t_{N-2}) & \cdots & C_a(t_1) \end{pmatrix}. \quad (11)$$

5.2. Data fidelity

As emphasized in Section 4.2, we assume that the data are corrupted by additive Gaussian noise. Therefore, for each voxel $k \in \llbracket 1, K \rrbracket$, we have the discrete temporal signal

$$\mathbf{d}_k = \mathbf{c}_k + \boldsymbol{\eta}_k, \quad (12)$$

where \mathbf{c}_k is defined in (10) and the components $\eta_{k,1}, \dots, \eta_{k,N}$ of $\boldsymbol{\eta}_k$ are the realizations of identically distributed zero-mean Gaussian random variables. Since our goal is to estimate the signals $F_k \mathbf{r}_k =: \mathbf{f}_k$, a natural choice for the data-fidelity term is given by

$$\Phi(\mathbf{f}_1, \dots, \mathbf{f}_K) = \sum_{k \in \llbracket 1, K \rrbracket} \|\alpha \rho_k \Delta_t \mathbf{C}_a \mathbf{f}_k - \mathbf{d}_k\|_2^2, \quad (13)$$

where $\|\cdot\|_2$ denotes the ℓ_2 -norm. Assuming equality in (10), any minimizer of Φ is a maximum-likelihood estimate of $(F_1 \mathbf{r}_1, \dots, F_K \mathbf{r}_K)$, but such solutions are unacceptable because the convolution matrix \mathbf{C}_a is ill-conditioned. This is a consequence of the particular structure of \mathbf{C}_a (triangular and almost Toeplitz) together with the fact that the first few coefficients of the signal $C_a(t_1), \dots, C_a(t_N)$ are close to zero and that this signal is increasing on $\llbracket 1, n_0 \rrbracket$ with $C_a(t_{n_0}) = \sup_{n \in \llbracket 1, N \rrbracket} C_a(t_n) \gg C_a(t_1)$.

For simplicity, we will use the following compact notation in the remainder of the paper:

$$\Phi(\mathbf{f}) = \|\mathbf{H}\mathbf{f} - \mathbf{d}\|_2^2, \quad (14)$$

where \mathbf{f} is the vertical concatenation of the \mathbf{f}_k 's, \mathbf{d} is the vertical concatenation of the \mathbf{d}_k 's, and the ‘‘data-fidelity matrix’’ \mathbf{H} is the block-diagonal concatenation of the matrices $\alpha \rho_k \Delta_t \mathbf{C}_a$; that is,

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_K \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_K \end{pmatrix} \quad \text{and} \quad \mathbf{H} = \alpha \Delta_t \text{diag}(\rho_1 \mathbf{C}_a, \dots, \rho_K \mathbf{C}_a) \quad (15)$$

(\mathbf{f} and \mathbf{d} are of size $KN \times 1$ and \mathbf{H} is of size $KN \times KN$).

5.3. Temporal regularization

From a physiological standpoint, the true (continuous) residue functions are smooth, which suggests to use ℓ_2 -regularization along the temporal dimension. Therefore, we define the temporal-regularization term as the sum of the squared ℓ_2 -norms of the temporal-derivative signals:

$$\Psi_t(\mathbf{f}) = \sum_{k \in \llbracket 1, K \rrbracket} \sum_{n \in \llbracket 2, N \rrbracket} \left(\frac{f_{k,n} - f_{k,n-1}}{\Delta_t} \right)^2, \quad (16)$$

where $f_{k,n}$ denotes the n th sample (or component) of \mathbf{f}_k . Besides the fact that temporal ℓ_2 -regularization is fully justified, it will be made clear in Section 5.5 that it has two important advantages: first, it ensures the existence of a solution to our reconstruction problem, and second, it guarantees that this problem is well-posed if the spatial-regularization term is convex.

The compact expression of Ψ_t is

$$\Psi_t(\mathbf{f}) = \|\mathbf{T}\mathbf{f}\|_2^2, \quad (17)$$

where the ‘‘temporal-differences matrix’’ \mathbf{T} is defined by

$$\mathbf{T} = \text{diag}(\underbrace{\mathbf{D}_t, \dots, \mathbf{D}_t}_{K \text{ times}}) \quad \text{with} \quad \mathbf{D}_t = \frac{1}{\Delta_t} \begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & 0 \\ & & \ddots & \ddots & \\ 0 & & & -1 & 1 \end{pmatrix} \quad (18)$$

(\mathbf{D}_t is of size $(N-1) \times N$ and thus \mathbf{T} is of size $K(N-1) \times KN$).

5.4. Spatial regularization

The construction of the spatial-regularization term is based on the following observations: on the one hand, the time-concentration signals at some neighboring voxels v_k and v_l are similar if both v_k and v_l belong to a healthy tissue region, and on the other hand, these signals can be very different if either v_k or v_l or both are in a damaged tissue region. In other words, we have to deal with two opposed situations: (i) when $F_k \mathbf{r}_k \approx F_l \mathbf{r}_l$, as observed in healthy tissues, and (ii) when $F_k \mathbf{r}_k$ is significantly different from $F_l \mathbf{r}_l$ for most n , which occurs in damaged tissues or at the interfaces between healthy and damaged tissues. This suggests to use discontinuity-preserving regularization (rather than ℓ_2 -regularization) in the spatial domain. Another argument in favor of this choice is that piecewise-smoothness in the spatial domain is consistent with clinical practice; indeed, brain tissues are usually classified by thresholding perfusion parameters, which produces homogeneous spatially-connected regions reflecting the cerebral hemodynamics.

Let \mathfrak{N} be a neighborhood system on the set of voxel indices—that is, $\mathfrak{N} = \{\mathfrak{N}(k); k \in \llbracket 1, K \rrbracket\}$ is a collection of subsets of $\llbracket 1, K \rrbracket$ such that, for all $(k, l) \in \llbracket 1, K \rrbracket^2$, $k \notin \mathfrak{N}(k)$ and $l \in \mathfrak{N}(k) \iff k \in \mathfrak{N}(l)$ —and let $\mathfrak{N} = \{(k, l) \in \llbracket 1, K \rrbracket^2 \mid l \in \mathfrak{N}(k)\}$ be the set of neighboring voxel pairs. The proposed spatial-regularization term is of the form

$$\Psi_s(\mathbf{f}) = \sum_{(k,l) \in \mathfrak{N}} \sum_{n \in \llbracket 1, N \rrbracket} \psi \left(\frac{f_{k,n} - f_{l,n}}{\|\gamma_k - \gamma_l\|_2} \right), \quad (19)$$

where γ_k denotes the center of voxel k and $\psi: \mathbb{R} \rightarrow \mathbb{R}$ is even and increasing in \mathbb{R}_+ . In the case of isotropic spatial sampling, a natural way to define the neighborhood system \mathfrak{N} is to let $\mathfrak{N}(k)$ be the set of indices of the voxels whose center is in a closed ball centered at γ_k . In the case of anisotropic spatial sampling, this generalizes to

$$\mathfrak{N}(k) = \left\{ l \in \llbracket 1, K \rrbracket \mid 0 < \|\text{diag}(A_x^{-1}, A_y^{-1}, A_z^{-1})(\gamma_k - \gamma_l)\| \leq r \right\}, \quad (20)$$

where A_x , A_y and A_z denote the resolutions along the spatial dimensions, and where the norm $\|\cdot\|$ and the radius $r > 0$ are

independent of k . In our experiments, we will use the 26-nearest neighborhood system, which is defined by $\|\cdot\| = \|\cdot\|_\infty$ (the maximum norm on \mathbb{R}^3) and $r = \sqrt{3}$.

Borrowing from the terminology of Markov random field modeling, we call ψ a *potential function* (PF). From the optimization point of view, a very convenient choice is $\psi(u) = u^2$, which corresponds to ℓ_2 -regularization. However, the quadratic PF produces unsatisfactory results because it unduly penalizes large spatial gradients and thus precludes the formation of discontinuities at the interfaces between healthy and damaged regions. Many forms of PF have been proposed in the literature to overcome this limitation, with a dichotomy between convex and non-convex functions. Convex PFs preserve discontinuities in the sense that they reduce (but not eliminate) smoothing in their vicinity (Charbonnier et al., 1997); they are usually nearly affine beyond a neighborhood of the origin. A well-known example is

$$\psi_{1,\delta}(u) = \sqrt{u^2 + \delta^2} - \delta, \quad (21)$$

where $\delta > 0$ (this function is often used for obtaining a differentiable approximation to total-variation regularization (Rodríguez and Wohlberg, 2009), in which case δ is small compared to the width of the range of the original image). Non-convex PFs detect discontinuities in the sense that they yield solutions whose gradient magnitudes are either below a threshold $\theta_0 > 0$ or above another threshold $\theta_1 > \theta_0$, but not in between (Nikolova, 2005). However, the price to pay is an increase in optimization difficulty compared to the convex case. Two typical examples of non-convex PFs are

$$\psi_{2,\delta}(u) = \ln(1 + (u/\delta)^2) \quad \text{and} \quad \psi_{3,\delta}(u) = \frac{u^2}{\delta^2 + u^2}, \quad (22)$$

which were originally proposed in Hebert and Leahy (1989) and Geman et al. (1985), respectively. The main difference between these two functions lie in their behavior at infinity: we have $\lim_{u \rightarrow +\infty} \psi_{2,\delta}(u) = +\infty$, whereas $\psi_{3,\delta}$ is bounded. Consequently, $\psi_{2,\delta}$ should be preferred over $\psi_{3,\delta}$ in terms of optimization difficulty, but $\psi_{3,\delta}$ is expected to produce sharper discontinuities than $\psi_{2,\delta}$ does.

The compact expression of Ψ_s is less obvious than that of Ψ_t . For any $(k, l) \in \mathfrak{Z}$, we let $\mathbf{D}_{(k,l)}$ be the $N \times KN$ finite-difference matrix defined by

$$\mathbf{D}_{(k,l)} \mathbf{f} = \frac{1}{\|\gamma_k - \gamma_l\|_2} \begin{pmatrix} f_{k,1} - f_{l,1} \\ \vdots \\ f_{k,N} - f_{l,N} \end{pmatrix}, \quad (23)$$

and we define the “spatial-differences matrix” \mathbf{S} to be the vertical concatenation of the $\mathbf{D}_{(k,l)}$ ’s:

$$\mathbf{S} = \begin{pmatrix} \mathbf{D}_{\pi(1)} \\ \vdots \\ \mathbf{D}_{\pi(J)} \end{pmatrix}, \quad (24)$$

where $J = \frac{1}{2} \sum_{k \in [1,K]} |\mathfrak{R}(k)|$ is the cardinality of \mathfrak{Z} and π is any bijection from $[1, J]$ to \mathfrak{Z} . Then,

$$\Psi_s(\mathbf{f}) = \sum_{i \in [1, JN]} \psi(\mathbf{S}_i \mathbf{f}), \quad (25)$$

where \mathbf{S}_i denotes the i th row of \mathbf{S} .

5.5. Global cost function

Using the compact notations, the global cost function $\Omega: \mathbb{R}^{KN} \rightarrow \mathbb{R}$ is given by

$$\begin{aligned} \Omega(\mathbf{f}) &= \Phi(\mathbf{f}) + \lambda_t \Psi_t(\mathbf{f}) + \lambda_s \Psi_s(\mathbf{f}) \\ &= \|\mathbf{H}\mathbf{f} - \mathbf{d}\|_2^2 + \lambda_t \|\mathbf{T}\mathbf{f}\|_2^2 + \lambda_s \sum_{i \in [1, JN]} \psi(\mathbf{S}_i \mathbf{f}), \end{aligned} \quad (26)$$

where the so-called *smoothing parameters* $\lambda_t > 0$ and $\lambda_s \geq 0$ adjust the regularization strength in the temporal dimension and in the spatial domain, respectively. As mentioned in Section 5.3, temporal ℓ_2 -regularization imparts important properties to Ω . These properties are stated in Propositions 1–3 below, whose proofs are given in Appendix A.

Proposition 1. Ω is coercive, that is, $\lim_{\|\mathbf{f}\| \rightarrow \infty} \Omega(\mathbf{f}) = +\infty$.

Proposition 2. If ψ is convex, then Ω is strictly convex.

Proposition 3. Assume that ψ is non-convex and twice differentiable. Let

$$\Gamma(\psi) = -\inf_{u \in \mathbb{R}} \psi''(u) \quad (27)$$

be the maximal concavity of ψ , and let $\tilde{\mathbf{H}}$ be the vertical concatenation of \mathbf{H} and $\sqrt{\lambda_t} \mathbf{T}$. Then, Ω is strictly convex if

$$\lambda_s < \frac{2}{\Gamma(\psi)} \left(\frac{\mu_{\inf}(\tilde{\mathbf{H}})}{\mu_{\sup}(\mathbf{S})} \right)^2 =: \lambda_s^*, \quad (28)$$

where $\mu_{\inf}(\tilde{\mathbf{H}})$ and $\mu_{\sup}(\mathbf{S})$ denote the smallest and largest singular values of $\tilde{\mathbf{H}}$ and \mathbf{S} , respectively.

In the remainder of this paper, we assume that ψ is differentiable. Consequently, Ω is differentiable, and thus the set of solutions to the minimization of Ω belongs to the set of stationary points of Ω , that is,

$$\Omega(\mathbf{f}^*) = \inf_{\mathbf{f} \in \mathbb{R}^{KN}} \Omega(\mathbf{f}) \quad \Rightarrow \quad \nabla \Omega(\mathbf{f}^*) = 0. \quad (29)$$

Proposition 1 ensures that Ω has a global minimum, which means that our optimization problem has a solution. If ψ is convex, it follows from Proposition 2 that Ω has a unique stationary point and that this point is the global minimum \mathbf{f}^* of Ω , that is, $\nabla \Omega(\mathbf{f}) = 0 \Leftrightarrow \mathbf{f} = \mathbf{f}^*$. Therefore, we propose to search for a minimum of Ω using a fixed-point iteration scheme to solve the equation $\nabla \Omega(\mathbf{f}) = 0$. We describe our deterministic relaxation algorithm in the next section, where we also study its convergence properties in both the convex and non-convex cases. But before that, let us briefly discuss the implications of Proposition 3. Simply speaking, this proposition states that for any non-convex, twice-differentiable PF ψ and any $\lambda_t > 0$, there exists $\lambda_s^* > 0$ such that Ω is strictly convex for all $\lambda_s \in [0, \lambda_s^*]$. Hence, the ratio λ_s/λ_s^* can be interpreted as a measure of the difficulty in minimizing Ω : the larger λ_s/λ_s^* , the higher the difficulty, and conversely, the closer λ_s/λ_s^* is to zero, the easier the optimization problem. This leads to the following relationships between the optimization difficulty and the components of Ω (the convergence analysis for the convex-case in Section 6.2.1 leads to similar conclusions in terms of convergence speed). First, and obviously, the difficulty increases with the spatial-regularization strength. Second, the greater the maximum concavity of ψ , the greater the difficulty. For instance, for the non-convex PFs defined in (22), we have $\Gamma(\psi_{2,\delta}) = 1/(4\delta^2)$ and $\Gamma(\psi_{3,\delta}) = 1/(2\delta^2)$, which suggests that $\psi_{2,\delta}$ should be preferred over $\psi_{3,\delta}$ and that the difficulty decreases with increasing δ (this latter point is not surprising, since δ adjusts the size of the symmetric intervals where $\psi_{2,\delta}$ and $\psi_{3,\delta}$ behave like quadratic functions). Third, the difficulty increases with $\mu_{\inf}(\tilde{\mathbf{H}})$, or, equivalently, the better the conditioning of \mathbf{C}_a and the larger λ_t , the easier the minimization of Ω .

6. Deterministic relaxation

6.1. Construction of the algorithm

We assume that the PF ψ satisfies the following conditions:

- C1.** ψ is strictly increasing on \mathbb{R}_+ , even, and differentiable.
C2. $\psi''(0^+)$ exists and is positive, and $\lim_{u \rightarrow +\infty} \psi'(u)/u = 0$.

Conditions of type C1 are standard in regularized reconstruction and conditions of type C2 guarantee edge-preservation properties (Charbonnier et al., 1997; Li, 1998); they are all satisfied by the PFs $\psi_{1,\delta}$, $\psi_{2,\delta}$ and $\psi_{3,\delta}$ defined in (21) and (22). We let $\psi^\dagger: \mathbb{R} \rightarrow (0, +\infty)$ be the so-called *interaction function* defined by

$$\psi^\dagger(u) = \begin{cases} \psi'(u)/u & \text{if } u \neq 0, \\ \psi''(0^+) & \text{if } u = 0. \end{cases} \quad (30)$$

It is easy to check that

$$\nabla \Omega(\mathbf{f}) = 0 \iff \underbrace{(2\mathbf{H}^T \mathbf{H} + 2\lambda_t \mathbf{T}^T \mathbf{T} + \lambda_s \mathbf{S}^T \mathbf{E}(\mathbf{f}) \mathbf{S})}_{=: \mathbf{M}(\mathbf{f})} \mathbf{f} = 2\mathbf{H}^T \mathbf{d}, \quad (31)$$

$$\text{where } \mathbf{E}(\mathbf{f}) = \text{diag}(\psi^\dagger(\mathbf{S}_1 \mathbf{f}), \dots, \psi^\dagger(\mathbf{S}_N \mathbf{f})). \quad (32)$$

By (50) (see Appendix A), and because $\psi^\dagger(u) > 0$ for all u , the symmetric matrix $\mathbf{M}(\mathbf{f})$ is positive definite and hence invertible. Therefore, $\nabla \Omega(\mathbf{f}) = 0 \iff \Xi(\mathbf{f}) = \mathbf{f}$ with $\Xi: \mathbb{R}^{KN} \rightarrow \mathbb{R}^{KN}$ defined by $\Xi(\mathbf{f}) = 2(\mathbf{M}(\mathbf{f}))^{-1} \mathbf{H}^T \mathbf{d}$. This suggests the following iterative relaxation algorithm:

$$\mathbb{A}_\Omega \equiv \begin{cases} \mathbf{f}^{(0)} \in \mathbb{R}^{KN}, \\ \forall p \in \mathbb{N}, \quad \mathbf{f}^{(p+1)} = 2(\mathbf{M}(\mathbf{f}^{(p)}))^{-1} \mathbf{H}^T \mathbf{d}. \end{cases} \quad (33)$$

This fixed-point iteration scheme can be equivalently written as

$$\mathbf{f}^{(p+1)} = \arg \inf_{\mathbf{f} \in \mathbb{R}^{KN}} \Omega_0(\mathbf{f}, \mathbf{e}(\mathbf{f}^{(p)})), \quad (34)$$

where the functions $\mathbf{e}: \mathbb{R}^{KN} \rightarrow (0, +\infty)^{JN}$ and $\Omega_0: \mathbb{R}^{KN} \times (0, +\infty)^{JN} \rightarrow \mathbb{R}$ are defined by

$$\Omega_0(\mathbf{f}, \mathbf{e}) = 2\|\mathbf{H}\mathbf{f} - \mathbf{d}\|_2^2 + 2\lambda_t \|\mathbf{T}\mathbf{f}\|_2^2 + \lambda_s \sum_{i \in [1, JN]} \varepsilon_i (\mathbf{S}_i \mathbf{f})^2 \quad (35)$$

$$\text{and } \mathbf{e}(\mathbf{f}) = [\psi^\dagger(\mathbf{S}_1 \mathbf{f}), \dots, \psi^\dagger(\mathbf{S}_N \mathbf{f})]^T. \quad (36)$$

For any $\varepsilon \in (0, +\infty)^{JN}$, the map $\Omega_0(\cdot, \varepsilon): \mathbb{R}^{KN} \rightarrow \mathbb{R}$ is a positive definite quadratic function, and thus \mathbb{A}_Ω belongs to the well-known class of half-quadratic regularization algorithms (Charbonnier et al., 1997; Nikolova and Ng, 2005; Allain et al., 2006; Robini et al., 2013). (Even more specifically, \mathbb{A}_Ω belongs to the subclass of half-quadratic algorithms of the multiplicative form originally introduced in Geman and Reynolds (1992).)

Each iteration of \mathbb{A}_Ω solves a $KN \times KN$ linear system with symmetric positive matrix of the form of $\mathbf{M}(\mathbf{f})$ defined in (31). These matrices have at most $KN(N + n - 1)$ non-zero coefficients, where $n := \sup_k |\mathfrak{N}(k)|$ is the maximum size of the neighborhoods used in the spatial-regularization term. Because of this very sparse structure (the sparsity ratio is of the order of $1/K$), and because N is not too large in clinical applications (the contrast-agent concentration signals are measured over a period of about one minute with a sampling period close to one second), the update equation of \mathbb{A}_Ω can be solved exactly and efficiently using Cholesky decomposition. Furthermore, stability is ensured by the fact that the matrix $\mathbf{T}^T \mathbf{T}$ is weakly diagonally dominant.

6.2. Convergence properties

Let $\tilde{\mathbf{H}}$ be as in Proposition 3 and let $\tilde{\mathbf{d}}$ be the vertical concatenation of \mathbf{d} with the all-zero column vector of length $K(N-1)$, that is,

$$\tilde{\mathbf{H}} = \begin{pmatrix} \mathbf{H} \\ \sqrt{\lambda_t} \mathbf{T} \end{pmatrix} \quad \text{and} \quad \tilde{\mathbf{d}} = \begin{pmatrix} \mathbf{d} \\ \mathbf{0} \end{pmatrix}. \quad (37)$$

We can write Ω in the standard form

$$\Omega(\mathbf{f}) = \|\tilde{\mathbf{H}}\mathbf{f} - \tilde{\mathbf{d}}\|_2^2 + \lambda_s \sum_i \psi(\mathbf{S}_i \mathbf{f}), \quad (38)$$

and it follows that the convergence results derived in Charbonnier et al. (1997), Nikolova and Ng (2005), Allain et al. (2006), and Robini et al. (2013) apply to \mathbb{A}_Ω . The two following subsections examine the cases when ψ is convex and non-convex.

6.2.1. Convex case

Convergence results for the convex case are given in Charbonnier et al. (1997), Nikolova and Ng (2005), and Allain et al. (2006) under smoothness assumptions on ψ (in addition to C1 and C2) and if one of the following conditions hold: (i) $\tilde{\mathbf{H}}$ has full rank Charbonnier et al. (1997); (ii) $\tilde{\mathbf{H}}^T \tilde{\mathbf{H}}$ is invertible Nikolova and Ng (2005); and (iii) $\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} + \lambda_s \mathbf{S}^T \mathbf{S}$ is invertible Allain et al. (2006). In fact, since $\tilde{\mathbf{H}}^T \tilde{\mathbf{H}}$ is positive definite by the proof of Proposition 1 (see Appendix A), all these conditions are satisfied. The global convergence result with the weakest assumptions on ψ is given by Propositions 8 and 9 in Allain et al. (2006), from which we readily deduce the following theorem.

Theorem 1. Assume that conditions C1 and C2 are satisfied and that

- C3.** ψ is convex and C^1 ;
C4. ψ^\dagger is decreasing on \mathbb{R}_+ .

Then, any sequence $(\mathbf{f}^{(p)})_p$ generated by \mathbb{A}_Ω converges to the global minimum of Ω .

If, in addition to C1–C4, ψ is C^2 and ψ is C^3 in a neighborhood of zero, then Corollary 2.5 in Nikolova and Ng (2005) gives an upper bound on the root-convergence factor

$$R(\mathbb{A}_\Omega) = \sup \left\{ \limsup_{p \rightarrow \infty} \|\mathbf{f}^{(p)} - \mathbf{f}^*\|^{1/p} : (\mathbf{f}^{(p)})_p \text{ generated by } \mathbb{A}_\Omega \right\}, \quad (39)$$

where \mathbf{f}^* denotes the global minimum of Ω . If $\psi = \psi_{1,\delta}$, this upper bound is given by

$$R(\mathbb{A}_\Omega) \leq \left(\frac{\lambda_s \mu_{\sup}^2(\mathbf{S})}{2\delta \mu_{\inf}^2(\tilde{\mathbf{H}}) + \lambda_s \mu_{\sup}^2(\mathbf{S})} \right) \left(\frac{\kappa^2}{\delta^2 + \kappa^2} \right), \quad (40)$$

where $\kappa = \sup_{i \in [1, JN]} \|\mathbf{S}_i \mathbf{f}^*\|$. The consequences of this inequality agree with the implications of Proposition 3 for the non-convex case: (i) the convergence speed decreases with increasing spatial-regularization strength, (ii) the convergence is faster when the scale parameter δ is large, and (iii) the convergence speed increases when the condition number of \mathbf{C}_a decreases and when the temporal-regularization strength increases.

6.2.2. Non-convex case

The convergence results for the non-convex case are summarized in Theorem 2 below, whose proof can be found in (Robini et al., 2013). We denote by \mathfrak{S} the set of stationary points of Ω , that is,

$$\mathfrak{S} = \{\mathbf{f} \in \mathbb{R}^{KN} \mid \nabla \Omega(\mathbf{f}) = \mathbf{0}\}. \quad (41)$$

We call $\mathbf{g} \in \mathfrak{S}$ *isolated* if

$$\exists a > 0, \quad \forall \mathbf{f} \in \mathfrak{S} \setminus \{\mathbf{g}\}, \quad \|\mathbf{f} - \mathbf{g}\| \geq a, \quad (42)$$

and we say that \mathfrak{S} is *discrete* if all its points are isolated.

Theorem 2. Assume that conditions C1 and C2 are satisfied and that

- C3'.** ψ is non-convex and twice differentiable;
C4'. $\psi'''(0) = 0$ and ψ is four times differentiable at zero;
C5'. ψ^\dagger is strictly decreasing on \mathbb{R}_+ .

Let $(\mathbf{f}^{(p)})_p$ be any sequence generated by \mathbb{A}_Ω .

- (I) If Ω is strictly convex, then $(\mathbf{f}^{(p)})_p$ converges to the global minimum of Ω .
- (II) If \mathfrak{E} is discrete, then there exists $\mathbf{g} \in \mathfrak{E}$ such that $\lim_{p \rightarrow \infty} \mathbf{f}^{(p)} = \mathbf{g}$.
- (III) Let \mathbf{g} be an isolated stationary point of Ω . If \mathbf{g} is a local minimum of Ω , then there exists an open neighborhood U of \mathbf{g} such that $(\mathbf{f}^{(p)})_p$ converges to \mathbf{g} if $\mathbf{f}^{(0)} \in U$.
- (IV) There exists $\mathbf{g} \in \mathfrak{E}$ such that $\lim_{p \rightarrow \infty} \Omega(\mathbf{f}^{(p)}) = \Omega(\mathbf{g})$, and for any $p \in \mathbb{N}$, $\Omega(\mathbf{f}^{(p+1)}) \leq \Omega(\mathbf{f}^{(p)})$ with equality if and only if $\mathbf{f}^{(p)} \in \mathfrak{E}$.
- (V) $\lim_{p \rightarrow \infty} \inf_{\mathbf{g} \in \mathfrak{E}} \|\mathbf{f}^{(p)} - \mathbf{g}\| = 0$.

(Remark. It is further assumed in Robini et al. (2013) that $\lim_{u \rightarrow +\infty} \psi(u) = +\infty$, so that Ω is coercive, but this condition is not needed here because coercivity is ensured by temporal ℓ_2 -regularization, as shown in the proof of Proposition 1 (see Appendix A).)

Assume that C1, C2 and C3'–C5' hold, which is the case for $\psi_{2,\delta}$ and $\psi_{3,\delta}$ in (22). Then it follows from (i) and Proposition 3 that (28) is a sufficient condition for global convergence. If Ω is not strictly convex, we have from (ii) that the algorithm converges to a stationary point if \mathfrak{E} is discrete, but it can be objected that $\lim_{p \rightarrow \infty} \Omega(\mathbf{f}^{(p)})$ may be a maximum or a saddle point. However, this situation is very unlikely because any isolated stationary point that is a minimum is an attractor (by (iii)), whereas maxima and saddle points are unstable. Indeed,

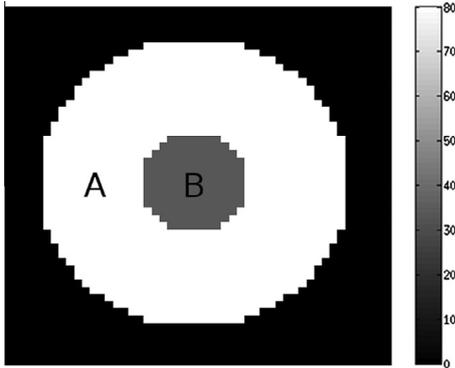


Fig. 4. Simulated CBF map ($\text{cm}^3/100 \text{ g/s}$). Regions A and B respectively correspond to healthy and damaged tissues.

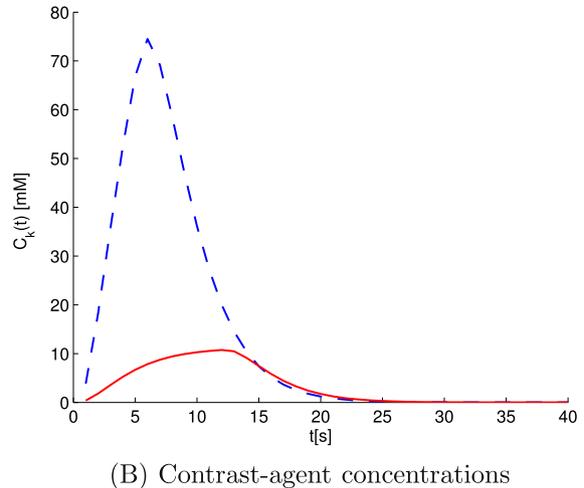
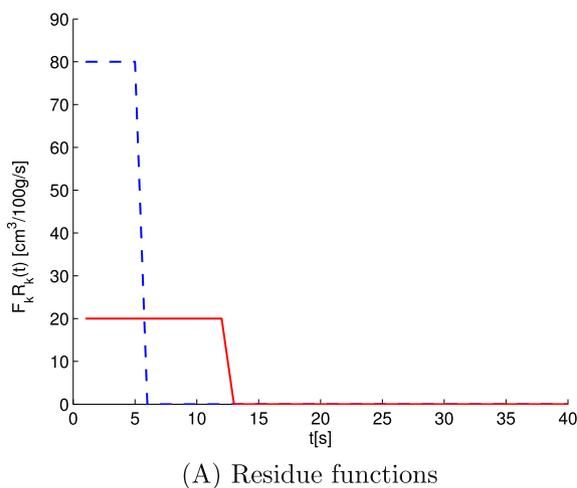


Fig. 5. Simulated residue functions and contrast-agent concentrations in the healthy (dashed blue line) and damaged (red plain line) regions.

from (iv), any small perturbation away from an isolated stationary point that is not a minimum—including the round-off errors in floating-point arithmetic—will eventually move the iterates away from this point. In this light, we can agree that the algorithm behaves well if $(\mathbf{f}^{(p)})_p$ gets arbitrarily close to \mathfrak{E} , which is guaranteed by (v). This includes the worst-case scenario when \mathfrak{E} is not discrete and $(\mathbf{f}^{(p)})_p$ does not converge, but this specific situation can be interpreted as a failure in the design of the spatial-regularization term, as it can be shown then that \mathfrak{E} contains a non-empty continuum.

7. Experimental setup

This section describes the synthetic and real data considered in our experiments together with the measures used to compare the performances of the competing algorithms.

7.1. Synthetic data

We generated a 50×50 slice of perfusion signals with two homogeneous regions corresponding to healthy and damaged brain tissues, as illustrated in Figs. 4 and 5 (the spatial resolution is 1.875 mm in both directions and the temporal resolution is 1 s). The contrast-agent concentration signals C_k are simulated using (1), and the AIF C_a is modeled by the gamma-variate function used in Mouridsen et al. (2006) and He et al. (2010), that is,

$$C_a(t) = \begin{cases} 0 & \text{if } t < 0, \\ at^b \exp(-t/c) & \text{if } t \geq 0, \end{cases} \quad (43)$$

where $a = 1$, $b = 3$, and $c = 1.5 \text{ s}$ (this setting gives an AIF similar to that obtained with a standard injection scheme). The residue functions R_k are modeled by boxcar functions, as in Østergaard et al. (1996b):

$$R_k(t) = \begin{cases} 1 & \text{if } t \leq \text{MTT}_k, \\ 0 & \text{if } t > \text{MTT}_k, \end{cases} \quad (44)$$

where MTT_k can take two different values depending on whether the voxel v_k is in the healthy region A or in the damaged region B. The CBV is set to 4 mL/100 g, and CBF_k is set to 80 mL/100 g/min if $v_k \in A$ and 20 mL/100 g/min if $v_k \in B$. Therefore, according to (2), $\text{MTT}_k = 3 \text{ s}$ if $v_k \in A$ and $\text{MTT}_k = 12 \text{ s}$ if $v_k \in B$.

Finally, the synthetic data is obtained by adding Gaussian noise to the simulated average contrast-agent concentration signals. The noise level is measured by the SNR in the data defined by

$$\text{SNR}_{\text{dB}} = 20 \log_{10} \left(\frac{C_{\max}}{\sigma} \right), \quad (45)$$

where $C_{\max} = \sup\{C_k(t_n); k \in \llbracket 1, K \rrbracket, n \in \llbracket 1, N \rrbracket\}$ is the maximum contrast-agent concentration and σ is the noise standard-deviation. In our experiments on synthetic data, the SNR is 22.6 dB unless stated otherwise (this is the average noise level observed in our real data).

7.2. Clinically acquired human MRI data

Our real data consist of perfusion-weighted MR volumes obtained from patients with acute ischemic stroke (due to middle cerebral artery occlusion) within 3 h of symptom onset. They were acquired on a 1.5 T clinical whole-body scanner (Siemens Avanto) using a gradient echo sequence (TR = 1540 ms, TE = 30 ms). For each patient, the data set is a $128 \times 128 \times 20$ voxel volume representing a $240 \times 240 \times 214 \text{ mm}^3$ region (the 128×128 slices are 5 mm thick and consecutive slices are separated by 6 mm gaps), each voxel being associated to a discrete contrast-agent concentration signal of length $N = 60$ with sampling period $\Delta_t = 1.2$ s. We also have the final infarct region delineated by a physician from one month follow-up T2 FLAIR imaging (T2 FLAIR images are obtained via an inversion-recovery pulse sequence that removes the signal from fluids and hence renders damaged tissues hyperintense). This final infarct region will play the role of perfect classification for performing a ROC analysis, which is justified by the study reported in [Rivers et al. \(2006\)](#).

7.3. Performance evaluation

We compare the performances of three algorithms: TSVD deconvolution, deconvolution using temporal regularization, and reconstruction with spatio-temporal regularization. For fair comparison, the hyper-parameters of these algorithms (namely, the SVD threshold and the temporal and spatial smoothing parameters λ_t and λ_s) are selected to obtain the best solutions in terms of peak SNR (PSNR) in the case of the synthetic data and in terms of ROC analysis in the case of the real data. The performances are evaluated at each stage of the processing pipeline using (i) the PSNR of the computed solutions; (ii) the PSNRs of the CBF, MTT and T^{\max} maps estimated from the computed solutions; and (iii) ROC parameters measuring closeness to perfect classification.

7.3.1. PSNR measurements

In the case of the synthetic data, the true CBF values and the true residue functions are known, and it is therefore possible to measure directly the quality of the computed solution $\hat{\mathbf{f}} = (\hat{\mathbf{f}}_1, \dots, \hat{\mathbf{f}}_K)$. We consider two measures. The first one is the PSNR of $\hat{\mathbf{f}}$ in a given region of interest (i.e., healthy tissues, damaged tissues, or healthy + damaged tissues), which is defined by

$$\text{PSNR}_{\text{dB}} = 10 \log_{10} \left(\frac{|\mathcal{M}| N f_{\max}^2}{\sum_{k \in \mathcal{M}} \|\hat{\mathbf{f}}_k - F_k \mathbf{r}_k\|_2^2} \right), \quad (46)$$

where $\mathcal{M} \subset \llbracket 1, K \rrbracket$ is the set of indices of the voxels in the region of interest, F_k and \mathbf{r}_k are the true CBF value and residue function at voxel k , and $f_{\max} = \sup\{F_k \mathbf{r}_k(t_n); k \in \mathcal{M}, n \in \llbracket 1, N \rrbracket\}$. The second measure is the PSNR of the estimate $\hat{\mathbf{P}}$ (computed from $\hat{\mathbf{f}}$) of a parametric map $\mathbf{P} = \{P_{v_k}; k \in \mathcal{M}\}$:

$$\text{PSNR}_{\text{dB}} = 10 \log_{10} \left(\frac{|\mathcal{M}| (\sup_{k \in \mathcal{M}} P_{v_k})^2}{\sum_{k \in \mathcal{M}} (\hat{P}_{v_k} - P_{v_k})^2} \right), \quad (47)$$

where $\mathbf{P} = \text{CBF, MTT or } T^{\max}$.

7.3.2. ROC analysis

In the case of the real data, we perform a ROC analysis to compare the predictive power of the T^{\max} maps estimated from the solutions computed by TSVD and spatio-temporal reconstruction (ROC analysis has previously been used for acute stroke MRI study in [Christensen et al. \(2009\)](#)). The performance is measured from the ROC curve, which plots the true positive rate (TPR) versus the false positive rate (FPR) associated with a moving threshold on the T^{\max} values. More precisely, let \mathcal{I} be the set of voxels in the final infarct region obtained from follow-up T2 FLAIR imaging, let \mathcal{V} be the set of voxels in the whole tissue region, and let $\hat{\mathcal{I}}(\tau)$ be the set of voxels whose estimated T^{\max} value is greater than τ . The ROC curve is the parametric curve $\tau \in [0, +\infty) \mapsto (\text{FPR}(\tau), \text{TPR}(\tau)) \in [0, 1]^2$, where

$$\text{FPR}(\tau) = \frac{|\hat{\mathcal{I}}(\tau) \cap (\mathcal{V} \setminus \mathcal{I})|}{|\mathcal{V} \setminus \mathcal{I}|} \quad \text{and} \quad \text{TPR}(\tau) = \frac{|\hat{\mathcal{I}}(\tau) \cap \mathcal{I}|}{|\mathcal{I}|}. \quad (48)$$

A perfect classifier would yield $\text{FPR} = 0$ and $\text{TPR} = 1$, whereas random guessing gives $\text{FPR} = \text{TPR}$. Our performance measures are the area under the ROC curve (AUC) and the minimum distance to perfect classification

$$\text{MDPC} = \inf_{\tau \geq 0} d^*(\tau), \quad d^*(\tau) = \|(\text{FPR}(\tau), \text{TPR}(\tau)) - (0, 1)\|_2. \quad (49)$$

The closer the AUC is to one and/or the closer MDPC is to zero, the higher the predictive power.

8. Results

8.1. Synthetic data

In this section, we compare the performances of the competing algorithms via PSNR measurements reflecting the quality of the estimated residue functions and of the associated parametric maps. We show (both quantitatively and qualitatively) that temporal regularization outperforms TSVD and that spatio-temporal regularization outperforms temporal regularization, and we compare convex and non-convex spatial regularization. The computation times given in the sequel correspond to Matlab implementations run on a single core of an Intel core i7 2.2 GHz laptop with 8 GB RAM.

8.1.1. Comparison in terms of residue functions

For accurate evaluation of the performance of the proposed spatio-temporal reconstruction approach, we first need to find appropriate values for the spatial and temporal smoothing parameters λ_s and λ_t . To do so, we set these parameters to the values that give the best PSNR in terms of residue functions. From now on, we consider the convex PF ψ_1 and the non-convex PF ψ_2 defined in (21) and (22).

[Fig. 6](#) displays the PSNR as a function of (λ_s, λ_t) , and [Fig. 7](#) shows the PSNR as a function of λ_s for a few fixed values of λ_t . Predictably enough, too large values of λ_s give over-smooth solutions, as is the case for too large values of λ_t when λ_s is fixed. Compared to temporal regularization (i.e., $\lambda_s = 0$), spatio-temporal regularization yields up to a 9 dB improvement in PSNR and produces better solutions for any λ_s in an interval of the form $(0, \mathcal{A})$, where \mathcal{A} increases as λ_t decreases.

[Table 2](#) gives the PSNR values associated with the estimates computed by the competing algorithms. Our spatio-temporal approach produces solutions in excellent agreement with the true residues in both healthy and damaged tissue regions, and it significantly outperforms TSVD deconvolution which has difficulty in recovering low-flow residue signals. In all cases, the PSNR in dam-

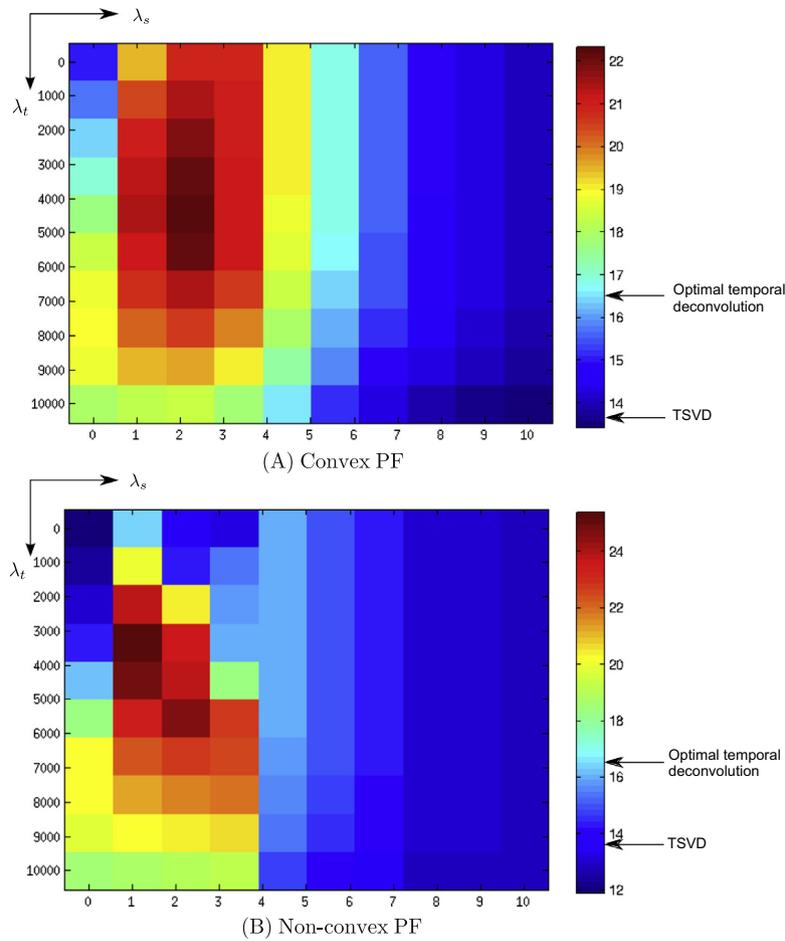


Fig. 6. PSNR of the residues as a function of the smoothing parameters λ_s and λ_t (the SNR in the data is 22.6 dB). The PSNR values obtained by TSVD and by optimal temporal deconvolution are indicated on the color bars. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

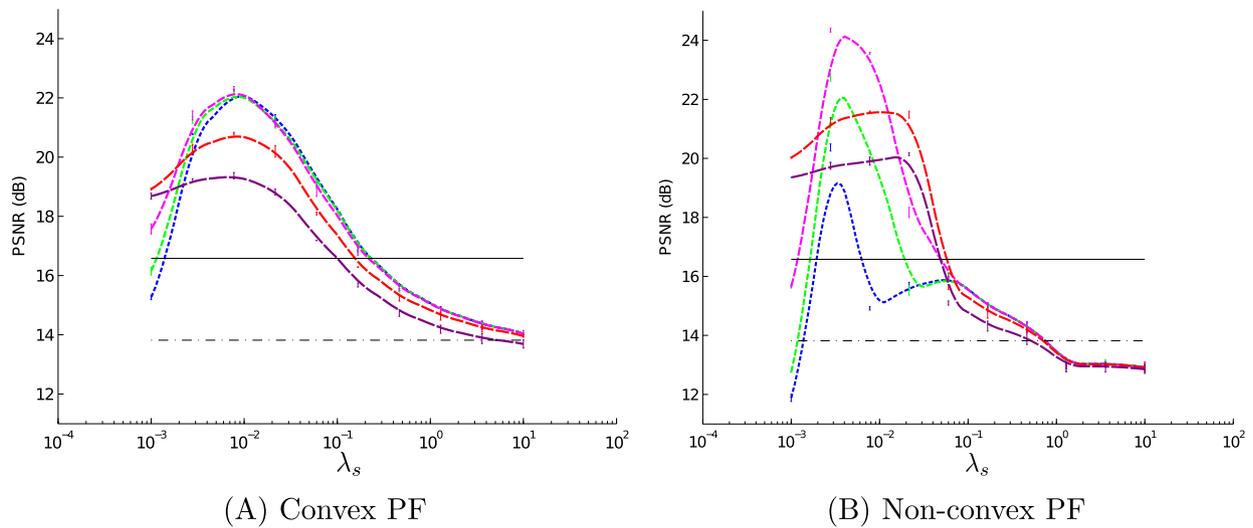


Fig. 7. PSNR vs. λ_s curves for different values of λ_t (the SNR in the data is 22.6 dB): $\lambda_t = 1$ (blue), $\lambda_t = 5$ (green), $\lambda_t = 50$ (magenta), $\lambda_t = 1000$ (red), and $\lambda_t = 5000$ (purple). The dotted and solid lines correspond to TSVD and to optimal temporal deconvolution, respectively.

aged tissues is lower than that in healthy tissues; the reason is that the blood flow is lower in damaged than in healthy tissues, which translates to lower SNR in damaged regions. Nevertheless, the PSNR obtained by spatio-temporal regularization in the damaged

region is about 9 dB higher than that obtained by TSVD in the healthy region.

Fig. 8 displays the optimal PSNR values as a function of the SNR in the data. Spatio-temporal regularization significantly outper-

Table 2
PSNR (dB) of the residue functions (the SNR in the data is 22.6 dB).

Method	Healthy tissue	Damaged tissue	All tissue
TSVD	14.41	10.82	13.82
Temporal deconvolution	18.40	16.72	16.58
Spatio-temporal reconstruction with ψ_1	24.02	21.52	22.51
Spatio-temporal reconstruction with ψ_2	27.15	23.37	25.13

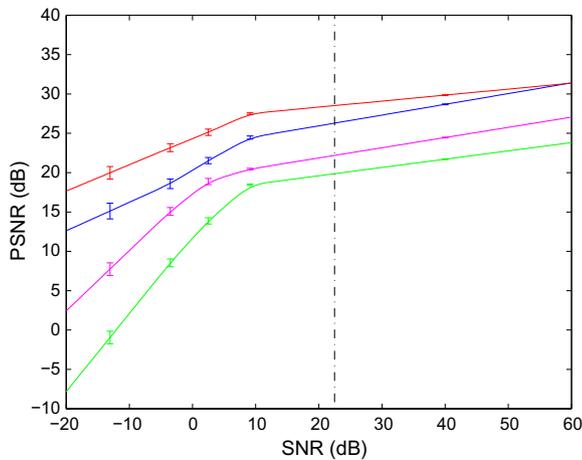


Fig. 8. Optimal PSNR of the residues as a function of the SNR in the data: TSVD (green), temporal deconvolution (magenta), spatio-temporal reconstruction with ψ_1 (blue), spatio-temporal reconstruction with ψ_2 (red). The dotted line indicates the SNR in our clinical data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

forms TSVD over the whole range of SNR values, and, as one would expect, the PSNR increases with the SNR in all cases. Compared to convex reconstruction, non-convex reconstruction is more interesting when the SNR is below 40 dB, but there is a price for this: the computation time ranges between 5 s (SNR = 60 dB) and 25 s (SNR = -20 dB) when using the convex PF φ_1 , whereas it ranges

between 35 s and 135 s when using the non-convex PF φ_2 . Therefore, non-convex reconstruction should be preferred in low SNR situations which occur when increasing temporal and/or spatial resolutions. For the noise level observed in our clinical data (SNR \approx 22.60 dB), the convex and non-convex reconstruction approaches perform similarly, and thus we will stick to the convex case for our experiments on real data in Section 8.2.

We complete our quantitative performance analysis by a qualitative inspection of the solutions produced by TSVD deconvolution and by our spatio-temporal reconstruction algorithm. Figs. 9 and 10 show the estimated flow residues at time $t = 5$ s and plots of their middle rows. The solutions obtained by spatio-temporal reconstruction (either convex or non-convex) are of much higher quality than those obtained by TSVD. We also observe that non-convex reconstruction outperforms convex reconstruction in terms of edge preservation: compared to the convex case, non-convex reconstruction produces sharper edges and does not fail to recover edges when the SNR is low.

Finally, Fig. 11 displays typical examples of flow residue signals in the healthy and damaged regions estimated from the data with 22.60 dB SNR. In both tissue classes, the signal computed by TSVD shows large oscillations which translate to large negative values without physical meaning. This unwanted effect is much less prominent in the case of temporal deconvolution and further attenuated when using spatio-temporal reconstruction (the example shown corresponds to the convex case).

8.1.2. Comparison in terms of parametric maps

Fig. 12 shows the CBF, MTT and T^{\max} maps computed from the perfusion volume estimates produced by the competing algorithms from the data with 22.60 dB SNR. In the case of TSVD deconvolution, all three maps are very noisy in both the healthy and damaged tissue regions, and it is almost impossible to distinguish these two tissue classes in terms of CBF or MTT. By contrast, the CBF and T^{\max} maps computed from the temporal deconvolution estimate are less noisy than those obtained via TSVD, and temporal deconvolution allows us to discriminate damaged from healthy tissues in terms of CBF. As the maps associated with spatio-temporal reconstruction show, the quality of the parametric estimates is significantly improved by exploiting the coherence between the flow

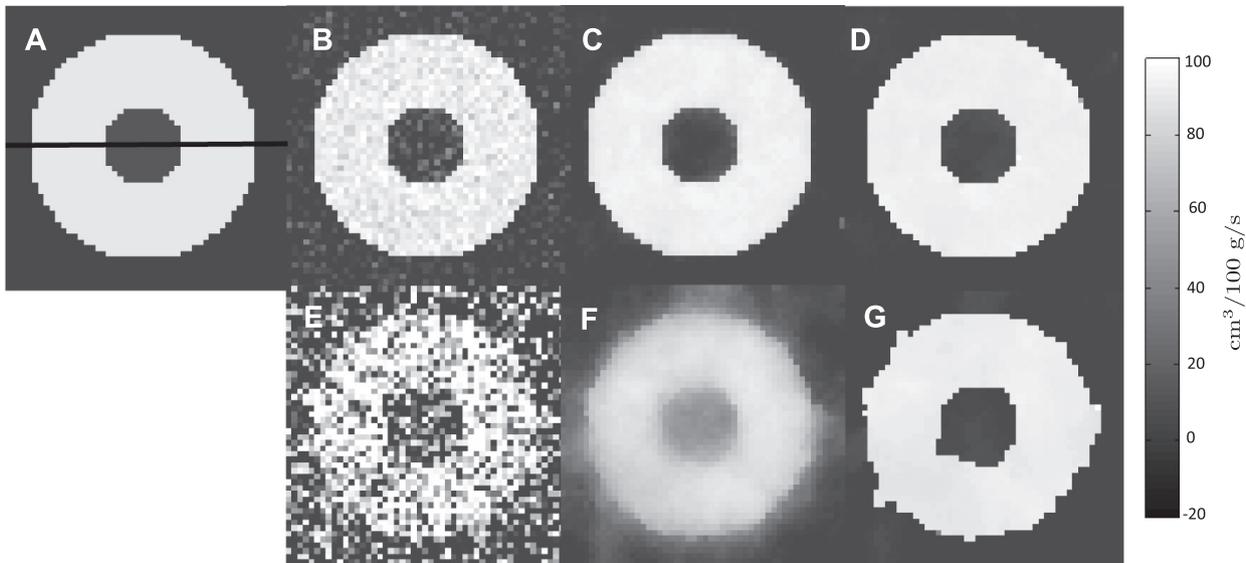


Fig. 9. Estimation of the CBF distribution displayed in (A) ($t = 5$,s) for two different values of the SNR in the data: (B–D) 22.6 dB; (E–G) 2.5 dB. Figures (B) and (E) show the distributions obtained by TSVD; (C) and (F) were obtained by convex spatio-temporal reconstruction; and (D) and (G) were obtained by non-convex spatio-temporal reconstruction. Note that negative values are reconstruction artifacts without physical meaning—they are set to zero in practice.

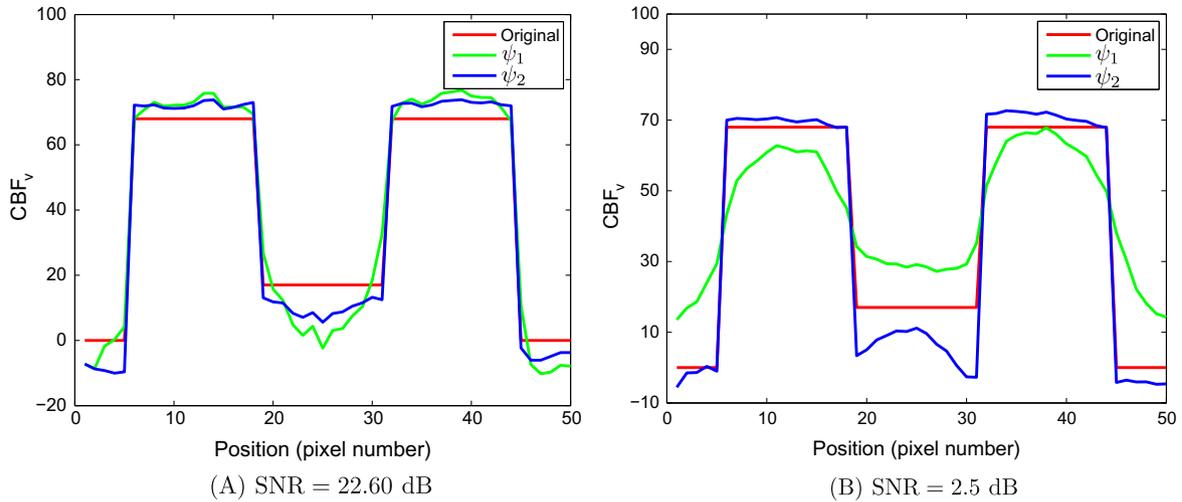


Fig. 10. Intensity profiles (middle rows) of the CBF distributions shown Fig. 9(A, C, D, F, and G): true CBF distribution (red); convex spatio-temporal reconstruction (green); and non-convex spatio-temporal reconstruction (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

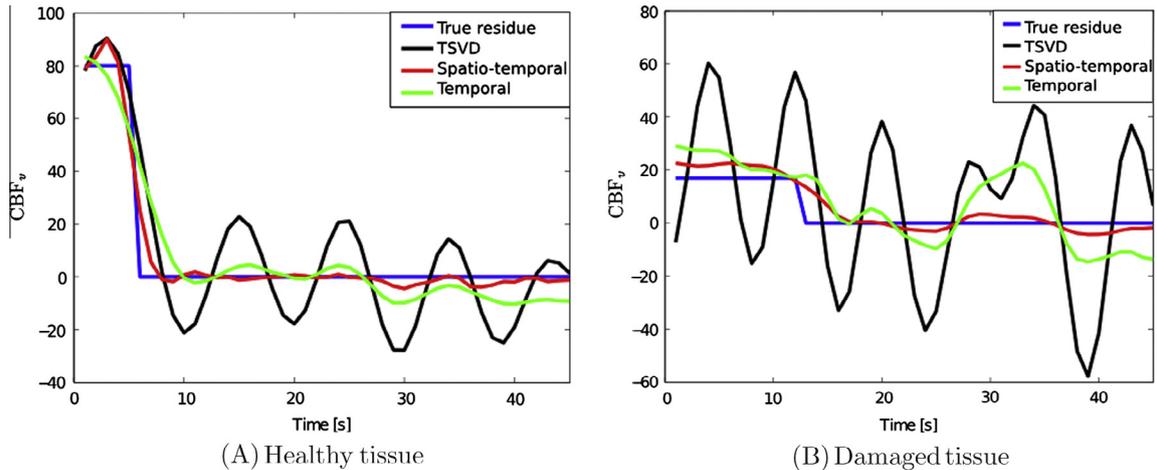


Fig. 11. Example residue functions obtained by the competing algorithms.

residue signals of neighboring voxels: (i) the healthy and damaged regions are smooth in all three cases; (ii) healthy and damaged tissues are separated by a sharp discontinuity in the CBF and T^{\max} maps; and (iii) damaged tissues can be distinguished from the healthy ones in the MTT map even though its contrast is low. These qualitative observations are supported by the PSNR measurements given in Table 3: for the three considered parameters, spatio-temporal reconstruction outperforms both TSVD and temporal deconvolution.

8.2. Real data

In the real data case, the true residues are unknown, and thus we cannot rely on PSNR measurements to tune the hyper-parameters of the algorithms. We circumvent this difficulty by selecting the values that best predict the final infarct in terms of ROC analysis (see Section 7.3.2). Besides, the AIF must be estimated: we do so by averaging the six sharpest contrast-agent signals in the anterior cerebral artery, where by sharp we mean small T^{\max} , narrow peak width, and large peak amplitude.

Fig. 13 shows the estimated flow residues at time $t = 9.6$ s for different data sets obtained using the protocol outlined in Sec-

tion 8.2. (For each data set, the processing of the entire $128 \times 128 \times 20$ volume of contrast-agent concentration signals takes 1–2 min using TSVD and 3–4 min using spatio-temporal reconstruction.) We observe that spatio-temporal reconstruction yields significant improvements over TSVD: not only the healthy tissue regions are smoother, but also the ischemic injury regions are better delineated.

In common practice, the final infarct region is defined as the set of voxels whose T^{\max} value is above a given threshold (Nagakane et al., 2012). Instead of setting this threshold empirically, we explore a continuous spectrum of threshold values using ROC curves and we choose the value τ_{roc} that gives the closest point to perfect classification, that is, $\tau_{roc} = \arg \inf_{\tau \geq 0} d^*(\tau)$ with d^* defined in (49). This allows us to measure the quality of a solution by comparing the associated estimated infarct region $\hat{\mathcal{I}}(\tau_{roc})$ to the infarct region \mathcal{I} delineated manually from follow-up T2 FLAIR imaging. The ROC curves and the classification results obtained for the different data sets are displayed in Fig. 14. We observe that the ROC curves associated with spatio-temporal reconstruction are almost always above those obtained via TSVD, which shows that spatio-temporal reconstruction yields better predictive power than TSVD deconvolution. The superiority of spatio-temporal reconstruction over

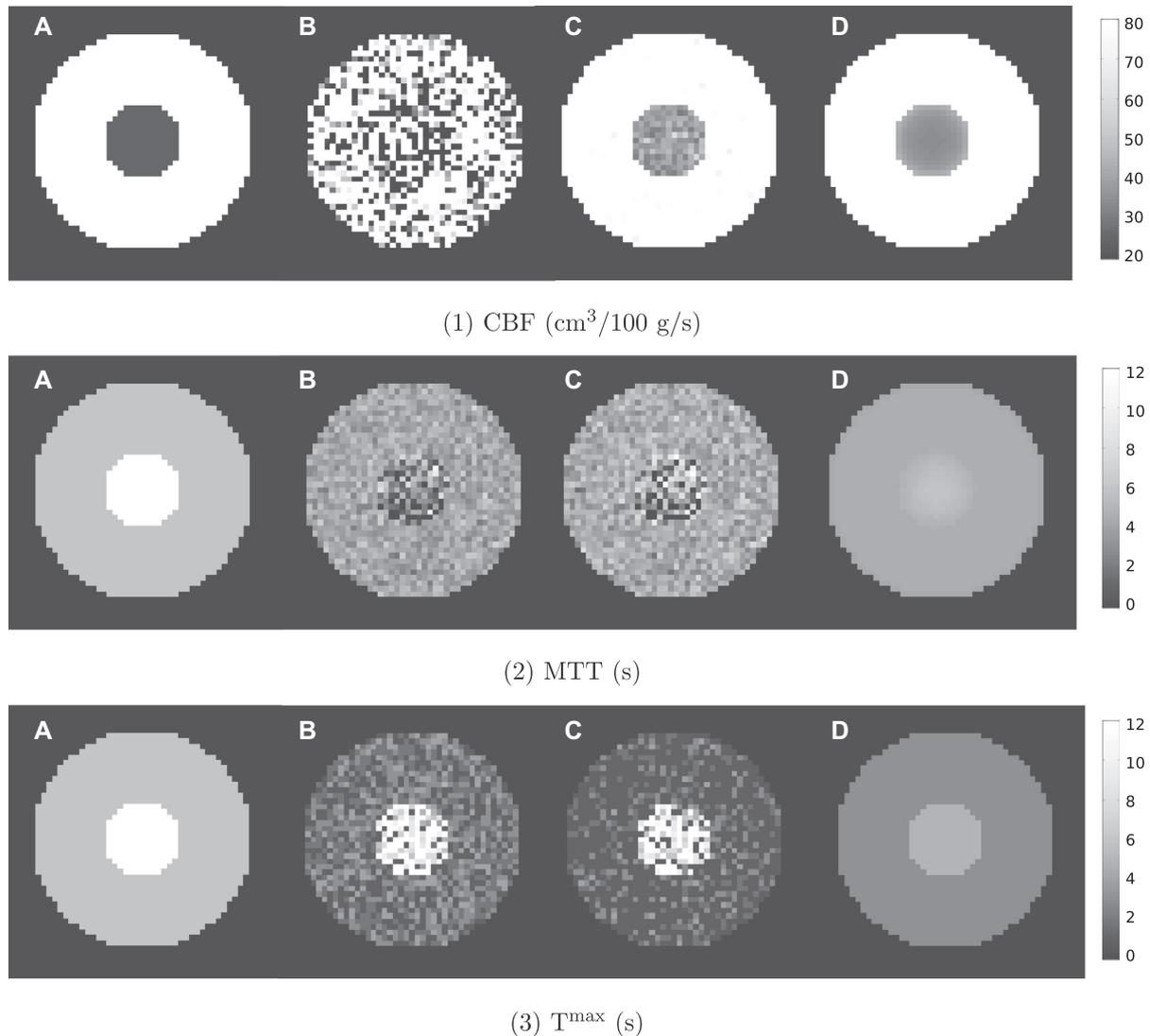


Fig. 12. Parametric maps (CBF, MTT and T^{\max}) computed from the estimates obtained by (B) TSVD, (C) temporal deconvolution, and (D) spatio-temporal reconstruction. The leftmost images (A) are the true parametric maps.

Table 3
PSNR (dB) of parametric maps.

Method	Healthy tissues	Damaged tissues	All tissues
<i>CBF</i>			
TSVD	10.96	1.21	7.99
Temporal deconvolution	26.35	28.40	26.55
Spatio-temporal reconstruction	28.53	29.78	28.66
<i>MTT</i>			
TSVD	36.99	18.58	27.46
Temporal deconvolution	37.52	18.99	27.82
Spatio-temporal reconstruction	45.71	21.66	30.87
<i>T^{\max}</i>			
TSVD	26.27	15.71	22.83
Temporal deconvolution	24.89	14.50	21.55
Spatio-temporal reconstruction	28.57	21.21	26.76

TSVD is also clearly seen from the classifications shown in Fig. 14 and from the measurements given in Table 4: in all cases, spatio-temporal reconstruction outperforms TSVD deconvolution in terms of false positives, false negatives, AUC, and MDPC.

9. Conclusion

We proposed a globally convergent spatio-temporal reconstruction approach to improve the quality of parametric maps in PWI. Our algorithm comes with strong convergence guarantees—convergence to the global minimum in the convex-case and convergence to a local minimum in the non-convex case—and experiments on both synthetic and real data show that it significantly outperforms TSVD deconvolution in terms of PSNR, ROC analysis and visual inspection. Hence our approaches definitively improves the localization of ischemic tissues. We focused on hemodynamic parameters that are routinely tested for binary classification (salvageable/necrotic tissue) in clinical studies. In the case where clinical practice would require three or more classes to improve diagnosis, we could perform hierarchical ROC analysis (Ishwaran and Gatsonis, 2000) to study multi-level classifications obtained from hemodynamic parameter maps. Then, an interesting perspective would be to assess the performance of our reconstruction algorithm in terms of fine classifications involving arterial, venous, and other tissue classes (as in Chou et al. (2007), Martel et al. (2001), and Wu and Liu (2007)).

We did not address the issue of selecting the hyper-parameters λ_t and λ_s which adjust the strengths of temporal and spatial regu-

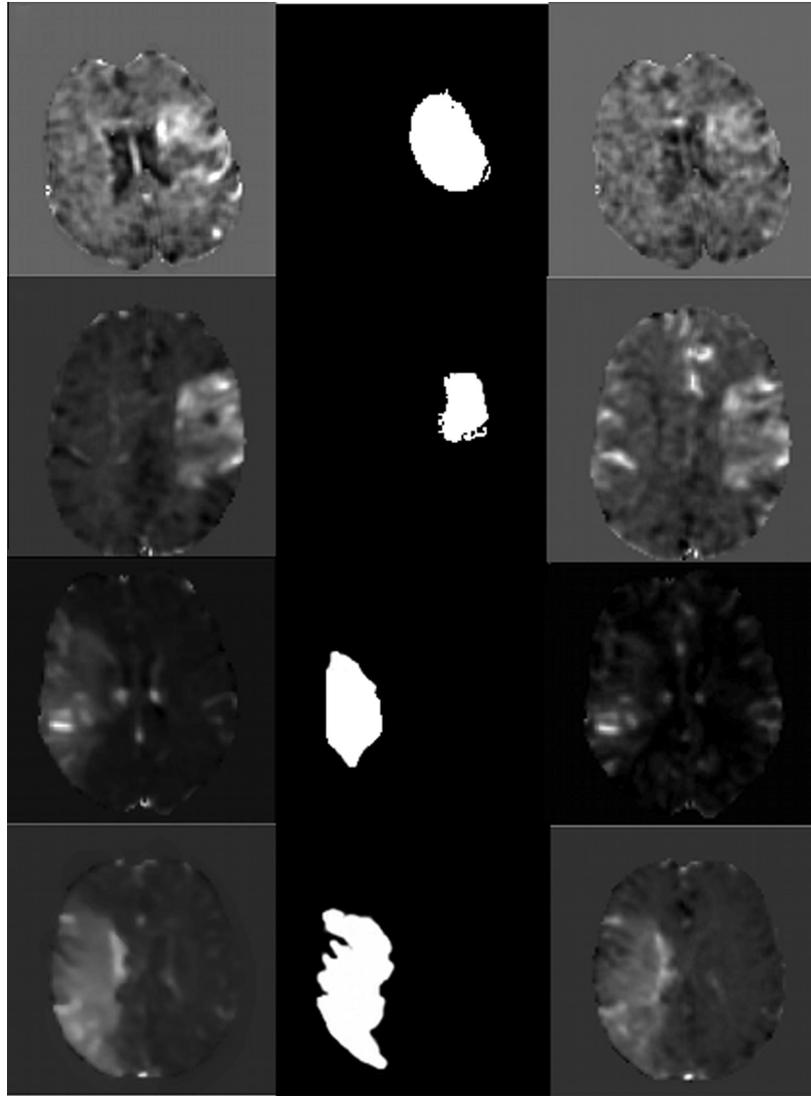


Fig. 13. Deconvolution of PWI data from ischemic subjects (each row corresponds to a different subject): spatio-temporal reconstruction (left column); FLAIR mask (middle column); and TSVD deconvolution (right column).

larization. Good starting points are the L-hypersurface method (Belge et al., 2002), the Monte Carlo SURE approach (Ramani et al., 2008), the use of the no-reference measure proposed in Zhu and Milanfar (2010), and the zero-crossing method developed in Ito et al. (2011). Another interesting possibility is to choose the values of λ_t and λ_s via ROC analysis. This requires a small database containing both PWI data acquired shortly after the stroke and final infarcts delineated by physicians from follow-up imaging. The resulting estimates of the optimal values of λ_t and λ_s are then tied to the MR scanner and the acquisition sequence used to construct the database, as a fixed acquisition protocol guarantees that the optimal values of λ_t and λ_s vary only slightly from one patient to the other.

Acknowledgements

The authors would like to thank Fabien Chauveau from the CRNL (UMR CNRS 5292, INSERM U1028) and Marlène Wiart, Tae-Hee Cho, Yves Berthezène and Norbert Nighoghossian from the CREATIS (UMR CNRS 5220, INSERM U1044) for providing the real data and for their encouragement in this work. This work was per-

formed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon.

Appendix A

A.1. Proof of Proposition 1

Let $\mathbf{f} \in \mathbb{R}^{KN} \setminus \{\mathbf{0}\}$ such that $\mathbf{T}\mathbf{f} = \mathbf{0}$. Then, there exists $k \in \llbracket 1, K \rrbracket$ such that $f_{k,1} = \dots = f_{k,N} \neq 0$. Consequently, since \mathbf{C}_a is non-zero and non-negative, we have $\mathbf{C}_a \mathbf{f}_k \neq \mathbf{0}$, and thus $\mathbf{H}\mathbf{f} \neq \mathbf{0}$. This shows that

$$\ker(\mathbf{H}) \cap \ker(\mathbf{T}) = \{\mathbf{0}\}. \quad (50)$$

Since $\lambda_t > 0$, it follows that the quadratic form $Q : \mathbf{f} \mapsto \|\mathbf{H}\mathbf{f}\|_2^2 + \lambda_t \|\mathbf{T}\mathbf{f}\|_2^2$ is positive definite, which is equivalent to say that there exists $c > 0$ such that $Q(\mathbf{f}) \geq c \|\mathbf{f}\|_2^2$ for all \mathbf{f} . Therefore, the quadratic function $\Phi + \lambda_t \Psi_t : \mathbf{f} \mapsto Q(\mathbf{f}) - 2\mathbf{d}^T \mathbf{H}\mathbf{f} + \|\mathbf{d}\|_2^2$ is coercive. Then, since ψ is increasing on \mathbb{R}_+ and even, Ψ_s is lower bounded, and thus Ω is coercive.

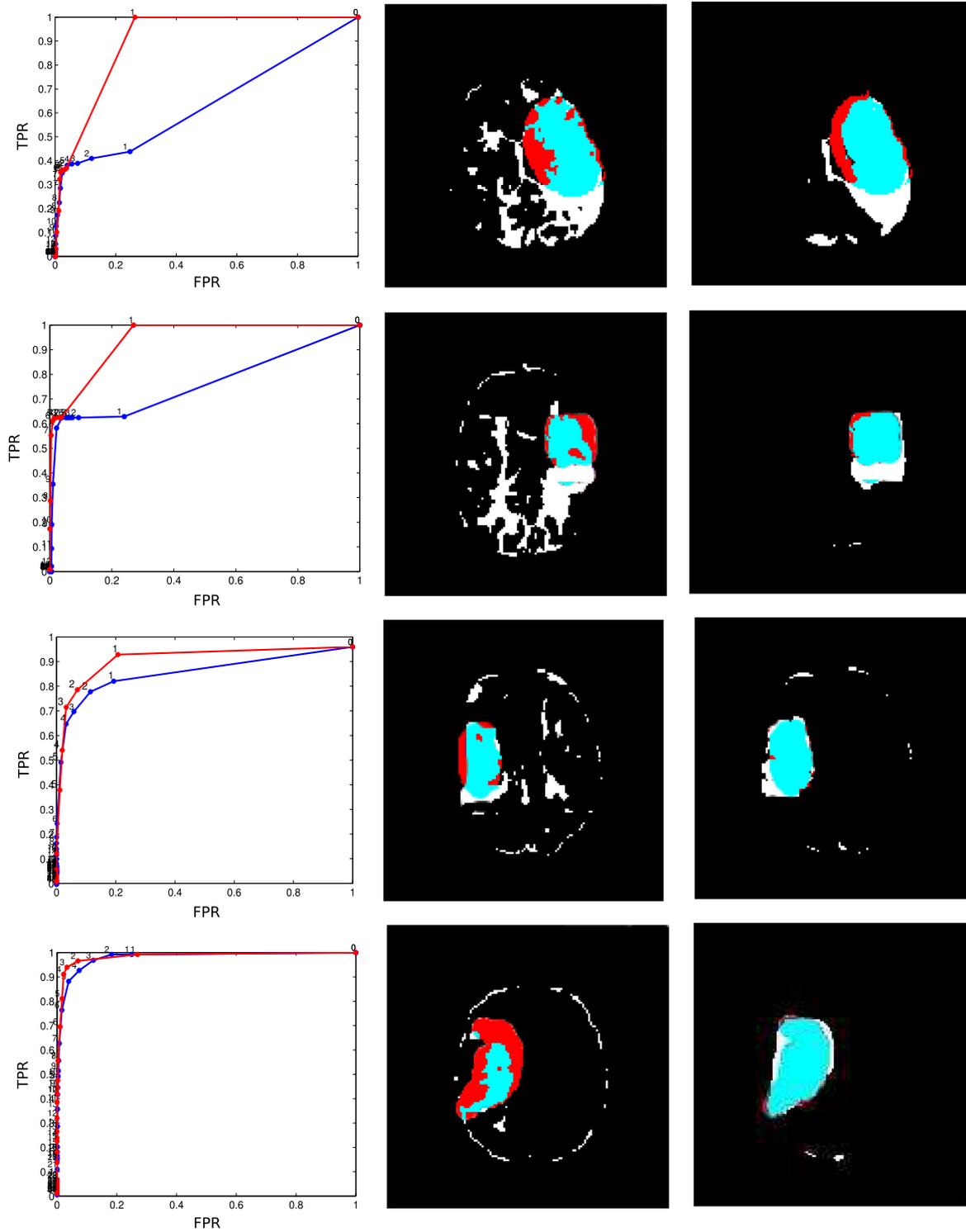


Fig. 14. Tissue outcome prediction (each row corresponds to a different subject). Left column : ROC curves obtained by varying the threshold on the T_{max} map computed from the TSVD estimates (blue dashed line) and from the spatio-temporal reconstruction estimates (red plain line). Middle column: false positives (white), true positives (light blue), false negatives (dark red) and true negatives (black) associated with the TSVD estimates. Right column: same as middle column but for spatio-temporal estimates.

Table 4
Area under curve (AUC) and minimum distance to perfect classification (MDPC) for the subjects considered in Fig. 13.

Patient	Spatio-temporal		TSVD	
	AUC	MDPC	AUC	MDPC
#1	0.90	0.26	0.64	0.61
#2	0.94	0.27	0.76	0.38
#3	0.94	0.19	0.89	0.22
#4	0.98	0.06	0.91	0.08

A.2. Proof of Proposition 2

Assume that ψ is convex. Then the maps $f \mapsto \psi(\mathbf{S}_t f)$ are convex (because the composition of a convex function with a linear map is convex), and thus so is Ψ_s as a sum of convex functions. In addition, according to the proof of Proposition 1, the quadratic function $\Phi + \lambda_t \Psi_t$ is positive definite and hence strictly convex. Therefore, Ω is strictly convex as the sum of a strictly convex function and of a convex function.

A.3. Proof of Proposition 3

The Hessian matrix of Ω at \mathbf{f} is

$$\nabla^2 \Omega(\mathbf{f}) = 2\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} + \lambda_s \mathbf{S}^T \text{diag}(\psi''(\mathbf{S}\mathbf{f}), \dots, \psi''(\mathbf{S}_j \mathbf{f})) \mathbf{S}.$$

For any $\mathbf{g} \in \mathbb{R}^{KN}$, we have

$$\begin{aligned} \mathbf{g}^T \nabla^2 \Omega(\mathbf{f}) \mathbf{g} &= 2\mathbf{g}^T \tilde{\mathbf{H}}^T \tilde{\mathbf{H}} \mathbf{g} + \lambda_s \sum_{i \in [1, JN]} \psi''(\mathbf{S}_i \mathbf{f}) (\mathbf{S}_i \mathbf{g})^2 \\ &\geq \mathbf{g}^T \underbrace{(2\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} - \lambda_s \Gamma(\psi) \mathbf{S}^T \mathbf{S})}_{=: \mathbf{L}} \mathbf{g}. \end{aligned} \quad (51)$$

Given a real symmetric matrix \mathbf{A} , we denote its smallest and largest eigenvalues by $v_{\inf}(\mathbf{A})$ and $v_{\sup}(\mathbf{A})$, respectively. Using the Weyl's inequality for eigenvalues, we have

$$\begin{aligned} v_{\inf}(\mathbf{L}) &\geq v_{\inf}(2\tilde{\mathbf{H}}^T \tilde{\mathbf{H}}) + v_{\inf}(-\lambda_s \Gamma(\psi) \mathbf{S}^T \mathbf{S}) \\ &= 2v_{\inf}(\tilde{\mathbf{H}}^T \tilde{\mathbf{H}}) - \lambda_s \Gamma(\psi) v_{\sup}(\mathbf{S}^T \mathbf{S}) \\ &= 2\mu_{\inf}^2(\tilde{\mathbf{H}}) - \lambda_s \Gamma(\psi) \mu_{\sup}^2(\mathbf{S}). \end{aligned}$$

Now assume that (28) holds. Then, the above inequality gives that $v_{\inf}(\mathbf{L}) > 0$, which means that \mathbf{L} is positive definite. It follows from (51) that $\nabla^2 \Omega(\mathbf{f})$ is positive definite for all $\mathbf{f} \in \mathbb{R}^{KN}$, and thus Ω is strictly convex.

References

- Allain, M., Idier, J., Goussard, Y., 2006. On global and local convergence of half-quadratic algorithms. *IEEE Trans. Image Process.* 15 (5), 1130–1142.
- Andersen, I.K., Have, A.S., Rasmussen, C.E., Hansson, L., Marstrand, J.R., Larsson, H.B., Hansen, L.K., 2002. Perfusion quantification using gaussian process deconvolution. *Magn. Reson. Med.* 48 (2), 351–361.
- Belge, M., Kilmer, M., Miller, E., 2002. Efficient determination of multiple regularization parameters in a generalized L-curve framework. *Inverse Problems* 18 (4), 1161–1183.
- Brekke, C., Lundervold, A., Enger, P.O., Brekken, C., Stalset, E., Pedersen, T.B., Haraldseth, O., Krueger, P.G., Bjerkvig, R., Chekenya, M., 2006. NG2 expression regulates vascular morphology and function in human brain tumours. *Neuroimage* 29 (3), 965–976.
- Calamante, F., Gadian, D.G., Connelly, A., 2003. Quantification of bolus-tracking MRI: improved characterization of the tissue residue function using Tikhonov regularization. *Magn. Reson. Med.* 50 (6), 1237–1247.
- Charbonnier, P., Blanc-Féraud, L., Aubert, G., Barlaud, M., 1997. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Process.* 6 (2), 298–311.
- Chou, Y.-C., Teng, M.M.H., Guo, W.-Y., Hsieh, J.-C., Wu, Y.-T., 2007. Classification of hemodynamics from dynamic-susceptibility-contrast magnetic resonance (DSC-MR) brain images using noiseless independent factor analysis. *Med. Image Anal.* 11 (3), 242–253.
- Christensen, S., Mouridsen, K., Wu, O., Hjort, N., Karstoft, H., Thomalla, G., Rother, J., Fiehler, J., Kucinski, T., Østergaard, L., 2009. Comparison of 10 perfusion MRI parameters in 97 sub-6-hour stroke patients using voxel-based receiver operating characteristics analysis. *Stroke* 40 (6), 2055–2061.
- Descombes, X., Kruggel, F., von Cramon, D.Y., 1998. Spatio-temporal fMRI analysis using Markov random fields. *IEEE Trans. Med. Imag.* 17 (6), 1028–1039.
- Donnan, G.A., Fisher, M., Macleod, M., Davis, S.M., 2008. *Stroke*. *Lancet* 371 (5), 1612–1623.
- Fieselmann, A., Kowarschik, M., Ganguly, A., Hornegger, J., Fahrig, R., 2011. Deconvolution-based CT and MR brain perfusion measurement: theoretical model revisited and practical implementation details. *Int. J. Biomed. Imaging* 2011 (14), 1–20.
- Geman, D., Reynolds, G., 1992. Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Machine Intell.* 14 (3), 367–383.
- Geman, S., McClure, D., 1985. Bayesian image analysis: an application to single photon emission tomography. In: *Proceedings of the Statistical Computation Section: Annual Meeting of the American Statistical Association*, August. Las Vegas, Nevada, pp. 12–18.
- Gonzalez, R.G., 2006. *Acute Ischemic Stroke: Imaging And Intervention*. Springer, <http://books.google.fr/books?id=Qbac1ljoDSoC>.
- Gudbjartsson, H., Patz, S., 1995. The Rician distribution of noisy MRI data. *Magn. Reson. Med.* 34 (6), 910–914.
- Hebert, T., Leahy, R., 1989. A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Trans. Med. Imag.* 8 (2), 194–202.
- He, L., Orten, B., Do, S., Karl, W.C., Kambadakone, A., Sahani, D., Pien, H., 2010. A spatio-temporal deconvolution method to improve perfusion CT quantification. *IEEE Trans. Med. Imag.* 29 (5), 1182–1191.
- Heiland, S., Kreibich, W., Reith, W., Benner, T., Döfler, A., Forsting, M., Sartor, K., 1998. Comparison of echo-planar sequences for perfusion-weighted MRI: which is best? *Neuroradiology* 40 (4), 216–221.
- Heim, S., Fahrmeir, L., Eilers, P.H.C., Marx, B.D., 2007. 3D space-varying coefficient models with application to diffusion tensor imaging. *Comput. Stat. Data Anal.* 51 (12), 6212–6228.
- Ishwaran, H., Gatsonis, C.A.W., 2000. A general class of hierarchical ordinal regression models with applications to correlated ROC analysis. *Stroke* 28, 731–750.
- Ito, K., Jin, B., Takeuchi, T., 2011. A regularization parameter for nonsmooth Tikhonov regularization. *SIAM J. Sci. Comput.* 33 (3), 1415–1438.
- Jähne, B., 1993. *Spatio-Temporal Image Processing: Theory and Scientific Applications*. Lecture Notes in Computer Science, vol. 751. Springer, <http://books.google.fr/books?id=gO6V5gh4IXsC>.
- LeBihan, D., Mangin, J.F., Poupon, C., Clark, C.A., Pappata, S., Molko, N., Chabriat, H., 2001. Diffusion tensor imaging: concepts and applications. *J. Magn. Reson. Imaging* 13 (4), 534–546.
- Li, S., 1998. Close-form solution and parameter selection for convex minimization-based edge-preserving smoothing. *IEEE Trans. Pattern Anal. Machine Intell.* 20 (9), 916–932.
- Luybaert, R., Boujraf, S., Sourbron, S., Osteaux, M., 2001. Diffusion and perfusion MRI: basic physics. *Eur. J. Radiol.* 38 (1), 19–27.
- Martel, A.L., Moody, A.R., Alder, S.J., Delay, G.S., Morgan, P.S., 2001. Extracting parametric images from dynamic contrast-enhanced MRI studies of the brain using factor analysis. *Med. Image Anal.* 5 (1), 29–39.
- Meier, P., Zierler, K., 1954. On the theory of the indicator-dilution method for measurement of blood flow and volume. *J. Appl. Phys.* 6, 731–744.
- Mouridsen, K., Friston, K., Hjort, N., Gyldensted, L., Østergaard, L., Kiebel, S., 2006. Bayesian estimation of cerebral perfusion using a physiological model of microvasculature. *NeuroImage*.
- Mumford, D., Shah, J., 1989. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.* 42 (5), 577–685.
- Nagakane, Y., Christensen, S., Ogata, T., Churilov, L., Ma, H., Parsons, M.W., Desmond, P.M., Levi, C.R., Butcher, K.S., Davis, S.M., Donnan, G.A., 2012. Moving beyond a single perfusion threshold to define penumbra: a novel probabilistic mismatch definition. *Stroke* 43, 1548–1555.
- Nikolova, M., 2005. Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares. *Multiscale Model. Simul.* 4 (3), 960–991.
- Nikolova, M., Ng, M., 2005. Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM J. Sci. Comput.* 27 (3), 937–966.
- Østergaard, L., Sorensen, A., Kwong, K., Weisskoff, R., Gyldensted, C., Rosen, B., 1996a. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part II: experimental comparison and preliminary results. *Magn. Reson. Med.* 36 (5), 726–736.
- Østergaard, L., Weisskoff, R., Chesler, D., Gyldensted, C., Rosen, B., 1996b. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: mathematical approach and statistical analysis. *Magn. Reson. Med.* 36 (5), 715–725.
- Østergaard, L., Jonsdottir, K.Y., Mouridsen, K., 2009. Predicting tissue outcome in stroke: new approaches. *Curr. Opin. Neurol.* 22 (1), 54–59.
- Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J.C., Trianni, G., 2009. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* 113 (Supplement 1), S110–S122.
- Ramani, S., Blu, T., Unser, M., 2008. Monte-Carlo SURE: a black-box optimization of regularization parameters for general denoising algorithms. *IEEE Trans. Image Process.* 17 (9), 1540–1554.
- Ritzenthaler, T., Cho, T.H., Wiart, M., Berthiller, J., Østergaard, L., Hermier, M., Derex, L., Berthezène, Y., Nighoghossian, N., 2011. Assessment of baseline hemodynamic parameters within infarct progression areas in acute stroke patients using perfusion-weighted MRI. *Neuroradiology* 53 (8), 571–576.
- Rivers, K., Wardlaw, J., Armitage, P., Bastin, M., Carpenter, T., Cvorov, V., P.J., H., M.S., D., 2006. Do acute diffusion- and perfusion-weighted MRI lesions identify final infarct volume in ischemic stroke? *Stroke* 37, 98–104.
- Robini, M.C., Zhu, Y., Luo, J., 2013. Edge-preserving reconstruction with contour-line smoothing and non-quadratic data-fidelity. *Inv. Probl. Imaging* 7 (4), 2013.
- Rodríguez, P., Wohlberg, B., 2009. Efficient minimization method for a generalized total variation functional. *IEEE Trans. Image Process.* 18 (2), 322–332.
- Rosen, B.R., Belliveau, J.W., Aronen, H.J., Kennedy, D., Buchbinder, B.R., Fischman, A., Gruber, M., Glas, J., Weisskoff, R.M., Cohen, M.S., 1991. Susceptibility contrast imaging of cerebral blood volume: human experience. *Magn. Reson. Med.* 22 (2), 293–299.
- Schmid, V.J., 2011. Voxel based adaptive spatio-temporal modelling of perfusion cardiovascular MRI. *IEEE Trans. Med. Imag.* 30 (7), 1305–1313.
- Schmid, V.J., Whitcher, B., Padhani, A.R., Taylor, N.J., Yang, G.-Z., 2006. Bayesian methods for pharmacokinetic models in dynamic contrast-enhanced magnetic resonance imaging. *IEEE Trans. Med. Imag.* 25 (12), 1627–1636.
- Schreiber, W.G., Guckel, F., Stritzke, P., Schmiedek, P., Schwartz, A., Brix, G., 1998. Cerebral blood flow and cerebrovascular reserve capacity: estimation by dynamic magnetic resonance imaging. *J. Cereb. Blood Flow Metab.* 18, 1143–1156.
- Toronov, V.Y., Zhang, X., Webb, A.G., 2007. A spatial and temporal comparison of hemodynamic signals measured using optical and functional Magnetic

- Resonance Imaging during activation in the human primary visual cortex. *Neuroimage* 34 (3), 1136–1148.
- Wintermark, M., Sesay, M., Barbier, E., Borbély, K., Dillon, W.P., Eastwood, J.D., Glenn, T.C., Grandin, C.B., Pedraza, S., Soustiel, J.F., Nariái, T., Zaharchuk, G., Caillé, J.M., Dousset, V., Yonas, H., 2005. Comparative overview of brain perfusion imaging techniques. *J. Neuroradiol.* 32 (5), 294–314.
- Wu, X.Y., Liu, G.R., 2007. Application of independent component analysis to dynamic contrast-enhanced imaging for assessment of cerebral blood perfusion. *Med. Image Anal.* 11 (3), 254–265.
- Wu, O., Østergaard, L., Weisskoff, R.M., Benner, T., Rosen, B.R., Sorensen, A.G., 2003. Tracer arrival timing-insensitive technique for estimating flow in mr perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix. *Magn. Reson. Med.* 50 (1), 164–174.
- Zhu, X., Milanfar, P., 2010. Automatic parameter selection for denoising algorithms using a no-reference measure of image content. *IEEE Trans. Image Process.* 19 (12), 3116–3132.