

# A haemodynamic response function model in spatio-temporal diffuse optical tomography

Yiheng Zhang<sup>1</sup>, Dana H Brooks<sup>2</sup> and David A Boas<sup>3</sup>

<sup>1</sup> Department of Radiology, University of Michigan Hospital, 1500 E Medical Center Drive, CGC B2109, Ann Arbor, MI 48109, USA

<sup>2</sup> Electrical and Computer Engineering Department, 442 DANA, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, USA

<sup>3</sup> Athinoula A Martinos Center, Massachusetts General Hospital, Harvard Medical School, 149 13th Street, Room 2301, Charlestown, MA 02129, USA

E-mail: [yihzhang@med.umich.edu](mailto:yihzhang@med.umich.edu)

Received 6 May 2005, in final form 5 July 2005

Published 21 September 2005

Online at [stacks.iop.org/PMB/50/4625](http://stacks.iop.org/PMB/50/4625)

## Abstract

Diffuse optical tomography (DOT) is a new and effective technique for functional brain imaging. It can detect local changes in both oxygenated and deoxygenated haemoglobin concentrations in tissue based on differential absorption at multiple wavelengths. Traditional methods in spatio-temporal analysis of haemoglobin concentrations in diffuse optical tomography first reconstruct the spatial distribution at different time instants independently, then look at the temporal dynamics on each pixel, without incorporating any temporal information as a prior in the image reconstruction. In this work, we present a temporal haemodynamic response function model described by a basis function expansion, in a joint spatio-temporal DOT reconstruction of haemoglobin concentration changes during simulated brain activation. In this joint framework, we simultaneously employ spatial regularization, spectral information and temporal assumptions. We also present an efficient algorithm for solving the associated large-scale systems. The expected improvements in spatial resolution and contrast-to-noise ratio are illustrated with simulations of human brain activation.

## 1. Introduction

Diffuse optical tomography (DOT) is a relatively new functional imaging technique that can provide three-dimensional mapping of the blood volume and oxygenation in biological tissues to depths of up to several centimetres, based on the interaction of near-infrared light and human tissue. This is of significant benefit for the non-invasive monitoring of cerebral haemodynamics and oxygenation (Villringer *et al* 1993, Hoshi and Tamura 1993, Villringer and Chance 1997).

Compared with functional magnetic resonance imaging (fMRI), DOT has several advantages, including higher temporal resolution and lower cost, and also disadvantages, in particular lower spatial resolution.

In DOT, one needs to solve a tomographic inverse problem. This inverse problem is generally ill-posed because of the highly scattering nature of tissue in the near-infrared wavelength range and the limited number of measurements, which result in serious amplification of the data noise in the solution. In addition, although the haemodynamic (and therefore the optical) properties of interest may change in time, the model used in tomographic reconstructions is usually spatial, not temporal. Traditional methods used to analyse this spatio-temporal process follow separate spatial and temporal processing (Barbour *et al* 1999, 2001). This strategy usually includes three steps: first the spatial distribution of the optical properties, such as absorption and/or scattering coefficients, is reconstructed at each time instant independently based on measurement data and the system model; then the results are transformed to spatial maps of haemoglobin concentration changes using spectroscopic analysis (Delpy *et al* 1988); finally temporal analysis is applied to the time courses at each pixel or on some region of interest. To address the ill-posed inverse problem in the spatial reconstruction, regularization methods have been widely used, e.g. Arridge *et al* (1993), O'Leary (1996), Pogue *et al* (1999) and Gaudette *et al* (2000). These methods tend to employ spatial prior knowledge or assumptions to restrict the solution space. Although some work has suggested combining the first two steps together, resulting in a direct reconstruction of haemoglobin concentration changes from optical data (Hillman 2002, Li *et al* 2004, 2005), spatial and temporal processing have still been done separately.

In contrast, some work has been reported recently employing temporal information for joint spatio-temporal reconstruction in DOT. In Kolehmainen *et al* (2003), temporal evolution of the optical properties was modelled by a state-space model, and the whole spatio-temporal process was estimated using a fixed-interval Kalman smoother algorithm. In Prince *et al* (2003), temporal changes of the optical properties were first modelled with the sum of several sinusoidal functions with known frequencies, chosen to correspond to different physiological factors, then the linear coefficients of these sinusoids became the new unknowns which evolved in time according to a state-space model. A comparison work among several joint spacetime regularization methods for discrete, linear inverse problems can be found in Zhang *et al* (2005).

In fMRI data analysis, haemodynamic response function (HRF) models have been used for some time as a convenient and flexible means to directly model the temporal dynamics of the haemoglobin concentration (Friston *et al* 1994, 1998, Rajapakse *et al* 1998, Gossel *et al* 2001, Marrelec *et al* 2002). These models are based on the assumption that the underlying 'brain activation' system is linear and time invariant, and the HRF is the transfer function (in the frequency domain) or the impulse response function (in the time domain) of the system. The haemodynamic changes are modelled as the convolution of the HRF and the stimulus function, which is generally a boxcar function and corresponds to the experimental stimulus paradigm. Pixelwise estimation of the HRF parameters leads to either a linear or a nonlinear optimization problem, depending on the parametrized model used for the haemodynamic response function.

The assumption of stationarity and linearity of the underlying physiological process in HRF is a good approximation to the real system as long as the inter-stimulus interval does not decrease beyond a few seconds (Dale and Buckner 1997, Buckner 1998). In addition, this assumption has been justified by the observation of additivity in the fMRI signal (Boynton *et al* 1996), and it is believed that the assumption holds in a wide range of experimental conditions

(Glover 1999). The HRF is thought to vary across different subjects, different areas of the brain and different stimulations (Aguirre *et al* 1998).

The HRF has been estimated in fMRI data analysis using several methods. Among them, different sets of basis functions have been used to span the HRF subspace, resulting in a least-squares solution for the linear coefficients (Friston *et al* 1998); parametric methods have assumed a family of functions for the shape of the HRF and estimated the parameters (examples including Poisson functions (Friston *et al* 1994), Gaussian functions (Rajapakse *et al* 1998), truncated Gaussian functions (Gossl *et al* 2001) and joined half-cosine functions (Woolrich *et al* 2004)); and a Bayesian framework was generated to directly estimate the HRF or its parameters (Marrelec *et al* 2002, 2004).

Among these methods, the generalized linear model (GLM) is perhaps the most popular. In this work, we explore the use of the GLM framework with basis function expansions in a spatio-temporal DOT reconstruction. To use such a model in DOT, we face several obstacles, due to differences between fMRI data analysis and DOT reconstruction. First, in optical tomography, we need to solve an ill-posed inverse problem and the forward model is a spatial, but not a temporal, mapping. In contrast, in fMRI, the pixelwise time series of the blood-oxygenation-level-dependent (BOLD) signal can be directly used to model haemodynamic changes. Second, in optical techniques, we need to translate temporal changes of optical properties into changes of haemoglobin concentrations by using chromophore spectral information at multiple wavelengths, while in fMRI this is not necessary. Third, fMRI selectively detects only deoxy-haemoglobin concentration variations, while optical methods are sensitive to both oxy-haemoglobin (HbO) and deoxy-haemoglobin (HbR), as they are the two dominant absorbers in the near-infrared wavelength range. Thus, for DOT separate HRFs for HbO and HbR are required.

Based on these considerations, in this work we derive an ordinary least-squares solution, using a linearized forward model, in which we simultaneously employ spatial regularization, spectral information and temporal basis assumptions for reconstruction. The resulting framework leads to a large system of equations, even if the spatial dimension of the problem is small or moderately sized. We have developed an efficient algorithm that can reduce the computational complexity back to that of the spatial dimension only by iteratively solving the system while exploiting its embedded blockwise Kronecker product structure.

The paper is arranged as follows. In section 2, we will describe the use of an HRF model in fMRI time-series analysis, and derive the framework exploiting HRF for DOT reconstruction. Also our efficient computational algorithm will be presented. In section 3, we illustrate the improvement in simulations of human brain activation. A discussion section and a conclusion section follow.

## 2. Methods

In this section, we will first describe the use of an HRF model in fMRI data analysis, then derive how to use such a model in DOT spatio-temporal reconstructions of haemoglobin concentration changes.

### 2.1. HRF model in fMRI data analysis

In fMRI data analysis, we can model the BOLD time-series signal  $x_i(t)$  at the  $i$ th pixel as the convolution of the unknown haemodynamic response function  $h_i(t)$  with the stimulus function  $s(t)$  as

$$x_i(t) = (h_i * s)(t) + u_i(t) \quad (1)$$

where  $u_i(t)$  is the model noise. If we sample the stimulus function and the HRF at time points  $q = 1, 2, \dots, Q$ , and the BOLD signal at time points  $n = 1, 2, \dots, N$ , the resulting discrete convolution model is

$$x_{n,i} = \sum_{q=n}^Q s_{Q-q+n} h_{q,i} + u_{n,i} \quad (2)$$

for each pixel  $i$  and time index  $n$ . This can be rewritten in a matrix–vector form as

$$\mathbf{x}_i = \mathbf{S}\mathbf{h}_i + \mathbf{u}_i \quad (3)$$

where  $\mathbf{x}_i$  is the column vector of the BOLD signal at the  $i$ th pixel,  $\mathbf{S} \in R^{N \times Q}$  is the stimulus matrix, whose rows are stimulus time series with shifted lags, and  $\mathbf{h}_i, \mathbf{u}_i$  are the HRF time series and the measurement noise vector, respectively.

Assuming we use  $L$  temporal basis functions,  $\{b_1(t), \dots, b_L(t)\}$  to span the HRF space and supposing the same set of basis functions is used for all pixels, we can write

$$\mathbf{h}_i = [\mathbf{b}_1 \cdots \mathbf{b}_L] \cdot \begin{bmatrix} c^{1,i} \\ \vdots \\ c^{L,i} \end{bmatrix} = \mathbf{B}\mathbf{c}_i, \quad (4)$$

where the matrix  $\mathbf{B}$  contains the basis function time series in its columns and  $\mathbf{c}_i$  is the linear coefficient vector for the  $i$ th pixel. Substituting equation (4) into equation (3) and setting  $\mathbf{F} = \mathbf{S}\mathbf{B}$ ,  $\mathbf{F} \in R^{N \times L}$ , we have

$$\mathbf{x}_i = \mathbf{F}\mathbf{c}_i + \mathbf{u}_i. \quad (5)$$

This is the basic GLM model in fMRI time-series analysis, first developed by Friston *et al* (1994). The matrix  $\mathbf{F}$  is known as the *design matrix* and represents the convolution of the stimulus function with the HRF temporal basis functions. For the time series for each pixel, the parameter vector  $\mathbf{c}_i$  can be obtained from an ordinary linear least-squares solution. Note that in the original work, an unknown constant variable was added as a parameter to take the constant baseline level into account, and thus the design matrix was augmented by a constant column vector whose entries were all 1's. In this work, we will not consider the constant level.

As examples of the selection of the basis functions for the HRF model (or avoiding the detour of the convolution and directly selecting the columns of the design matrix  $\mathbf{F}$ ), discrete cosine transform basis functions were used in Friston *et al* (2000), and mixtures of gamma functions with their derivatives were used in Friston *et al* (1998).

## 2.2. HRF model in DOT reconstruction

As noted in section 2, the use of an HRF model in DOT reconstruction differs from that in fMRI data analysis in three major aspects:

- (i) In fMRI, the data analysis is based on a series of 2D brain pseudo-anatomical images, where the brightness values of the pixels represent the BOLD signals. Thus, the time series at each imaging pixel can be used to directly model the haemodynamic changes. In contrast, in DOT we measure the transmitted light intensity at detectors on the boundary and need to solve an inverse problem using a forward propagation model. This forward model maps and mixes the optical properties of the brain tissue to the light measurements, and is a mapping in space, not in time.

- (ii) In general, we are more interested in functional parameters such as changes of haemoglobin concentrations than optical parameters, and thus the reconstructed spatial distributions of the optical properties at multiple wavelengths need to be translated to those of concentration changes of HbO and HbR, given extinction coefficients or spectral information on the biological tissues (Boas *et al* 2002).
- (iii) The BOLD signal in fMRI is a function of only one metabolic quantity: concentration changes of deoxy-haemoglobin  $\Delta[\text{HbR}]$ , due to the fact that deoxy-haemoglobin is paramagnetic while oxy-haemoglobin is diamagnetic. So one needs only the HRF of deoxy-haemoglobin. But in DOT, concentration changes of both oxy- and deoxy-haemoglobin will affect the absorption of the tissue, and thus the detected light intensity. The temporal dynamics of HbO and HbR are known to be different (Villringer and Chance 1997, Franceschini *et al* 2000), so one needs to employ two different HRFs, one for HbO and one for HbR, respectively.

To include these changes, we must first model the relationship between the light intensity measurement and the linear coefficients of HRF basis functions of both HbO and HbR, then solve the inverse problem. Assume we have two sets of temporal basis functions, where one has  $L_1$  basis functions for the HRF of HbO,  $\{b_1(t), \dots, b_{L_1}(t)\}$ , the other has  $L_2$  basis functions for the HRF of HbR,  $\{d_1(t), \dots, d_{L_2}(t)\}$ :

$$\begin{aligned} \mathbf{hbo}_i &= [\mathbf{b}_1 \cdots \mathbf{b}_{L_1}] \cdot \begin{bmatrix} \gamma^{1,i} \\ \vdots \\ \gamma^{L_1,i} \end{bmatrix} = \mathbf{B} \cdot \boldsymbol{\gamma}_i \\ \mathbf{hbr}_i &= [\mathbf{d}_1 \cdots \mathbf{d}_{L_2}] \cdot \begin{bmatrix} \sigma^{1,i} \\ \vdots \\ \sigma^{L_2,i} \end{bmatrix} = \mathbf{D} \cdot \boldsymbol{\sigma}_i \end{aligned} \tag{6}$$

where  $\mathbf{hbo}_i$  ( $\mathbf{hbr}_i$ ) is the HbO (HbR) HRF time series of the  $i$ th pixel, matrix  $\mathbf{B} \in R^{Q \times L_1}$  ( $\mathbf{D} \in R^{Q \times L_2}$ ) contains the HbO (HbR) basis functions in its columns and  $\boldsymbol{\gamma}_i$  ( $\boldsymbol{\sigma}_i$ ) is the linear coefficient vector at  $i$ th pixel for HbO (HbR). Accordingly, the time series of oxy-haemoglobin concentration changes,  $\Delta[\text{HbO}]$ , and deoxy-haemoglobin concentration changes,  $\Delta[\text{HbR}]$ , at the  $i$ th pixel can be represented as

$$\begin{aligned} \Delta[\text{HbO}]_i &= \mathbf{F}_{\text{HbO}} \cdot \boldsymbol{\gamma}_i + \mathbf{u}_i \\ \Delta[\text{HbR}]_i &= \mathbf{F}_{\text{HbR}} \cdot \boldsymbol{\sigma}_i + \mathbf{m}_i, \end{aligned} \tag{7}$$

where  $\mathbf{F}_{\text{HbO}} = \mathbf{S}\mathbf{B} \in R^{N \times L_1}$  and  $\mathbf{F}_{\text{HbR}} = \mathbf{S}\mathbf{D} \in R^{N \times L_2}$  are the design matrices for HbO and HbR, respectively. That is, these design matrices are the convolution of the same impulse function matrix  $\mathbf{S}$  with each basis function matrix,  $\mathbf{B}$  or  $\mathbf{D}$ .

The next step is to translate concentration changes of HbO and HbR to changes in absorption coefficient  $\mu_a$ , denoted here by  $x$ . We assume in this work that only absorption coefficients are of interest in DOT reconstructions, and that the scattering coefficients are constant and known. At one pixel at one time instant, the changes of  $\mu_a$  at two different wavelengths  $\lambda_1, \lambda_2$  are related to  $\Delta[\text{HbO}]$  and  $\Delta[\text{HbR}]$  by

$$\begin{bmatrix} x^{\lambda_1} \\ x^{\lambda_2} \end{bmatrix} = \underbrace{\begin{bmatrix} \epsilon_{\text{HbO}}^{\lambda_1} & \epsilon_{\text{HbR}}^{\lambda_1} \\ \epsilon_{\text{HbO}}^{\lambda_2} & \epsilon_{\text{HbR}}^{\lambda_2} \end{bmatrix}}_{\mathbf{E}} \begin{bmatrix} \Delta[\text{HbO}] \\ \Delta[\text{HbR}] \end{bmatrix} \tag{8}$$

where  $\mathbf{E}$  is the extinction coefficient matrix. Substituting equation (7) into (8), and stacking all the time-series values for one pixel together, we have

$$\begin{bmatrix} \mathbf{X}_i^{\lambda_1} \\ \mathbf{X}_i^{\lambda_2} \end{bmatrix} = (\mathbf{E} \otimes \mathbf{I}_N) \begin{bmatrix} \Delta[\text{HbO}]_i \\ \Delta[\text{HbR}]_i \end{bmatrix} = (\mathbf{E} \otimes \mathbf{I}_N) \begin{bmatrix} \mathbf{F}_{\text{HbO}} \cdot \boldsymbol{\gamma}_i + \mathbf{u}_i \\ \mathbf{F}_{\text{HbR}} \cdot \boldsymbol{\sigma}_i + \mathbf{m}_i \end{bmatrix}, \quad (9)$$

where  $\otimes$  denotes the Kronecker product, defined so that the  $ij$ th block of  $[\mathbf{A} \otimes \mathbf{B}] = a_{ij} \mathbf{B}$  (Graham 1981); and  $\mathbf{I}_N$  denotes the  $N$ -dimensional identity matrix. Now we arrange all the time series and the parameters of all the pixels together to obtain

$$\begin{bmatrix} \mathbf{X}^{\lambda_1} \\ \mathbf{X}^{\lambda_2} \end{bmatrix} = (\mathbf{E} \otimes \mathbf{I}_N) \begin{bmatrix} \mathbf{F}_{\text{HbO}} \boldsymbol{\Gamma} + \mathbf{U} \\ \mathbf{F}_{\text{HbR}} \boldsymbol{\Sigma} + \mathbf{M} \end{bmatrix}. \quad (10)$$

The dimensionalities are:  $\mathbf{X}^{\lambda_i} \in R^{N \times P}$  is the time-by-space  $\mu_a$  distribution matrix at wavelength  $\lambda_i$ , where  $N$  and  $P$  are the number of time points and the number of imaging pixels, respectively; and  $\boldsymbol{\Gamma} \in R^{L_1 \times P}$  and  $\boldsymbol{\Sigma} \in R^{L_2 \times P}$  are the basis-by-space parameter matrices for HRFs of HbO and HbR, respectively.  $\mathbf{U}$ ,  $\mathbf{M}$  are the model noise matrices.

To relate the absorption coefficient distributions to the measurement data, we employ a linearized forward model in the reconstruction. Such a linear model can be obtained by linearization, e.g. using a Born or Rytov approximation (Kak and Slaney 1988), and discretization of the underlying differential equation model, based on the diffusion approximation for light propagation in a highly scattering medium such as human tissues (Ishimaru 1978).

Note that it will be convenient to apply the forward model if we change  $\mathbf{X}^{\lambda_i}$  from a form of time-by-space to a form of space-by-time. To achieve this, we write equation (11) separately at each wavelength as

$$\begin{aligned} \mathbf{X}^{\lambda_1} &= \epsilon_{\text{HbO}}^{\lambda_1} [\mathbf{F}_{\text{HbO}} \boldsymbol{\Gamma} + \mathbf{U}] + \epsilon_{\text{HbR}}^{\lambda_1} [\mathbf{F}_{\text{HbR}} \boldsymbol{\Sigma} + \mathbf{M}] \\ \mathbf{X}^{\lambda_2} &= \epsilon_{\text{HbO}}^{\lambda_2} [\mathbf{F}_{\text{HbO}} \boldsymbol{\Gamma} + \mathbf{U}] + \epsilon_{\text{HbR}}^{\lambda_2} [\mathbf{F}_{\text{HbR}} \boldsymbol{\Sigma} + \mathbf{M}] \end{aligned} \quad (11)$$

then transpose both  $\mathbf{X}^{\lambda_i}$ s and stack them together again as

$$\begin{bmatrix} \mathbf{X}^{\lambda_1 T} \\ \mathbf{X}^{\lambda_2 T} \end{bmatrix} = (\mathbf{E} \otimes \mathbf{I}_P) \begin{bmatrix} \boldsymbol{\Gamma}^T \mathbf{F}_{\text{HbO}}^T + \mathbf{U}^T \\ \boldsymbol{\Sigma}^T \mathbf{F}_{\text{HbR}}^T + \mathbf{M}^T \end{bmatrix}. \quad (12)$$

Let  $\mathbf{Y}^{\lambda_i} \in R^{M \times N}$  represents the space-by-time light intensity measurement matrix at wavelength  $\lambda_i$  where  $M$  is the dimension of the measurements and  $\mathbf{A}^{\lambda_i} \in R^{M \times P}$  the forward matrix at wavelength  $\lambda_i$ , we apply the forward models, at both wavelengths, as

$$\begin{aligned} \begin{bmatrix} \mathbf{Y}^{\lambda_1} \\ \mathbf{Y}^{\lambda_2} \end{bmatrix} &= \begin{bmatrix} \mathbf{A}^{\lambda_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{\lambda_2} \end{bmatrix} \begin{bmatrix} \mathbf{X}^{\lambda_1 T} \\ \mathbf{X}^{\lambda_2 T} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}^{\lambda_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{\lambda_2} \end{bmatrix} (\mathbf{E} \otimes \mathbf{I}_P) \begin{bmatrix} \boldsymbol{\Gamma}^T \mathbf{F}_{\text{HbO}}^T \\ \boldsymbol{\Sigma}^T \mathbf{F}_{\text{HbR}}^T \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{U}} \\ \tilde{\mathbf{M}} \end{bmatrix} \end{aligned} \quad (13)$$

where  $\tilde{\mathbf{U}}$ ,  $\tilde{\mathbf{M}}$  are measurement noise.

Now the new inverse problem is to estimate parameter matrices  $\boldsymbol{\Gamma}$  and  $\boldsymbol{\Sigma}$  given measurements  $\mathbf{Y}^{\lambda_i}$ , forward models  $\mathbf{A}^{\lambda_i}$ , extinction coefficients in  $\mathbf{E}$  and design matrices  $\mathbf{F}_{\text{HbO}}$ ,  $\mathbf{F}_{\text{HbR}}$ .

If we write the two block equations in equation (13) separately, in a form similar to equation (11), use the property of Kronecker products that  $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A}) \text{vec}(\mathbf{X})$

(Graham 1981) and then put the two separate equations back together, we can vectorize equation (13) as

$$\begin{bmatrix} \text{vec}(\mathbf{Y}^{\lambda_1}) \\ \text{vec}(\mathbf{Y}^{\lambda_2}) \end{bmatrix} = \begin{bmatrix} \epsilon_{\text{HbO}}^{\lambda_1} \mathbf{F}_{\text{HbO}} \otimes \mathbf{A}^{\lambda_1} & \epsilon_{\text{HbR}}^{\lambda_1} \mathbf{F}_{\text{HbR}} \otimes \mathbf{A}^{\lambda_1} \\ \epsilon_{\text{HbO}}^{\lambda_2} \mathbf{F}_{\text{HbO}} \otimes \mathbf{A}^{\lambda_2} & \epsilon_{\text{HbR}}^{\lambda_2} \mathbf{F}_{\text{HbR}} \otimes \mathbf{A}^{\lambda_2} \end{bmatrix} \begin{bmatrix} \text{vec}(\mathbf{\Gamma}^T) \\ \text{vec}(\mathbf{\Sigma}^T) \end{bmatrix} + \begin{bmatrix} \text{vec}(\tilde{\mathbf{U}}) \\ \text{vec}(\tilde{\mathbf{M}}) \end{bmatrix}. \quad (14)$$

We now have a joint spatio-temporal reconstruction framework in equation (14), where the new unknowns are the linear coefficients of temporal basis functions for both oxygenated and deoxygenated haemoglobin concentrations, at all imaging pixels, and the augmented new forward model involves the original spatial forward mapping, the spectrum information encoded in the extinction coefficients and the temporal assumptions presented by the design matrices. An ordinary least-squares solution can be obtained to estimate  $\text{vec}(\mathbf{\Gamma}^T)$  and  $\text{vec}(\mathbf{\Sigma}^T)$ .

### 2.3. Gradient-type Tikhonov regularization

Although we incorporate temporal basis functions as prior knowledge in regularization, the augmented forward model for the whole spatio-temporal reconstruction may still be seriously ill-conditioned. In this case, we use a Tikhonov regularization method to solve equation (14).

For the regularization matrix, a standard zero-order Tikhonov regularization using an identity matrix is not suitable, because it will penalize the sum of all the linear coefficients of the basis functions over all imaging pixels, and thus produce over-smoothed results. In the absence of prior assumptions about the behaviours of different basis functions, one should use regularization for each distinct basis, while imposing spatial smoothness on the coefficient distribution. In particular, we expect the neighbouring pixels to have similar behaviour for each basis function, and therefore introduce a gradient-type Tikhonov regularization, in which the regularization matrix has the form as  $\mathbf{I}_{L_1+L_2} \otimes \mathbf{G}$ , where  $\mathbf{G}$  is an approximate gradient operator which estimates the first variation of the spatial distribution of the linear coefficients along both  $X$  and  $Y$  directions. We recall that  $L_1$  and  $L_2$  are the number of temporal basis functions for the HRF of HbO and of HbR, respectively.

### 2.4. An efficient algorithm for this large-scale system

After inclusion of the gradient-type Tikhonov regularization into equation (14), we have the following normal equation to solve:

$$\begin{bmatrix} \text{vec}(\tilde{\mathbf{Y}}^{\lambda_1}) \\ \text{vec}(\tilde{\mathbf{Y}}^{\lambda_2}) \end{bmatrix} = \begin{bmatrix} \mathbf{F}_1 \otimes \mathbf{P}_1 + \mathbf{I}_{L_1} \otimes \lambda^2 \mathbf{G}^T \mathbf{G} & \mathbf{F}_2 \otimes \mathbf{P}_2 \\ \mathbf{F}_2^T \otimes \mathbf{P}_2 & \mathbf{F}_3 \otimes \mathbf{P}_3 + \mathbf{I}_{L_2} \otimes \lambda^2 \mathbf{G}^T \mathbf{G} \end{bmatrix} \begin{bmatrix} \text{vec}(\mathbf{\Gamma}^T) \\ \text{vec}(\mathbf{\Sigma}^T) \end{bmatrix} \quad (15)$$

where

$$\begin{aligned} \mathbf{F}_1 &= \mathbf{F}_{\text{HbO}}^T \mathbf{F}_{\text{HbO}} \\ \mathbf{F}_2 &= \mathbf{F}_{\text{HbO}}^T \mathbf{F}_{\text{HbR}} \\ \mathbf{F}_3 &= \mathbf{F}_{\text{HbR}}^T \mathbf{F}_{\text{HbR}} \\ \mathbf{P}_1 &= (\epsilon_{\text{HbO}}^{\lambda_1})^2 \mathbf{A}_{\lambda_1}^T \mathbf{A}_{\lambda_1} + (\epsilon_{\text{HbO}}^{\lambda_2})^2 \mathbf{A}_{\lambda_2}^T \mathbf{A}_{\lambda_2} \\ \mathbf{P}_2 &= \epsilon_{\text{HbO}}^{\lambda_1} \epsilon_{\text{HbR}}^{\lambda_1} \mathbf{A}_{\lambda_1}^T \mathbf{A}_{\lambda_1} + \epsilon_{\text{HbO}}^{\lambda_2} \epsilon_{\text{HbR}}^{\lambda_2} \mathbf{A}_{\lambda_2}^T \mathbf{A}_{\lambda_2} \\ \mathbf{P}_3 &= (\epsilon_{\text{HbR}}^{\lambda_1})^2 \mathbf{A}_{\lambda_1}^T \mathbf{A}_{\lambda_1} + (\epsilon_{\text{HbR}}^{\lambda_2})^2 \mathbf{A}_{\lambda_2}^T \mathbf{A}_{\lambda_2} \\ \tilde{\mathbf{Y}}^{\lambda_1} &= \epsilon_{\text{HbO}}^{\lambda_1} \mathbf{A}_{\lambda_1}^T \mathbf{Y}^{\lambda_1} \mathbf{F}_{\text{HbO}} + \epsilon_{\text{HbO}}^{\lambda_2} \mathbf{A}_{\lambda_2}^T \mathbf{Y}^{\lambda_2} \mathbf{F}_{\text{HbO}} \\ \tilde{\mathbf{Y}}^{\lambda_2} &= \epsilon_{\text{HbR}}^{\lambda_1} \mathbf{A}_{\lambda_1}^T \mathbf{Y}^{\lambda_1} \mathbf{F}_{\text{HbR}} + \epsilon_{\text{HbR}}^{\lambda_2} \mathbf{A}_{\lambda_2}^T \mathbf{Y}^{\lambda_2} \mathbf{F}_{\text{HbR}}. \end{aligned} \quad (16)$$

Equation (15) is very large scale, with dimension of (*total number of basis functions*)  $\times$  (*number of imaging pixels*). Directly solving the system will involve high computational complexity and a huge demand on memory. However, the blocks of the normal equation have a Kronecker product structure, which can be used to reduce the computation. We thus develop a blockwise iteration method which reduces solving the normal equation to the order of spatial dimension only, i.e. the number of imaging pixels.

Specifically, we alternatively solve for the coefficient matrices of HbO or HbR,  $\text{vec}(\Gamma^T)$ ,  $\text{vec}(\Sigma^T)$ : suppose after the  $(k-1)$ th step, we have  $\Sigma_{k-1}^T$ , then at the  $k$ th step we solve

$$\begin{aligned} (\mathbf{F}_1 \otimes \mathbf{P}_1 + \mathbf{I}_{L_1} \otimes \lambda^2 \mathbf{G}^T \mathbf{G}) \text{vec}(\Gamma_k^T) &= \text{vec}(\widetilde{\mathbf{Y}}^{\lambda_1} - \mathbf{P}_2 \Sigma_{k-1}^T \mathbf{F}_2^T) \\ (\mathbf{F}_3 \otimes \mathbf{P}_3 + \mathbf{I}_{L_2} \otimes \lambda^2 \mathbf{G}^T \mathbf{G}) \text{vec}(\Sigma_k^T) &= \text{vec}(\widetilde{\mathbf{Y}}^{\lambda_2} - \mathbf{P}_2 \Gamma_k^T \mathbf{F}_2). \end{aligned} \quad (17)$$

This is actually a  $2 \times 2$  block Gauss–Seidel iterative method. During each iteration, the two systems in equation (17) can be further decoupled into a sequence of independent, even smaller size systems. To see that, for the first step in equation (17), suppose we have the singular value decomposition (SVD) of the symmetric matrix  $\mathbf{F}_1$ , as  $\mathbf{F}_1 = \mathbf{V}_{\text{HbO}} \Sigma_{\text{HbO}} \mathbf{V}_{\text{HbO}}^T$ , we can make variable changes as

$$\begin{aligned} \mathbf{Z} &= \Gamma_k^T \mathbf{V}_{\text{HbO}} \\ \mathbf{W} &= (\widetilde{\mathbf{Y}}^{\lambda_1} - \mathbf{P}_2 \Sigma_{k-1}^T \mathbf{F}_2^T) \mathbf{V}_{\text{HbO}}, \end{aligned} \quad (18)$$

and again use properties of Kronecker products that  $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A}) \text{vec}(\mathbf{X})$  and  $\mathbf{AB} \otimes \mathbf{CD} = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})$ . The first equation in equation (17) then reduces to  $L_1$  independent systems which have the size of the number of imaging pixels:

$$([\Sigma_{\text{HbO}}]_{i,i} \mathbf{P}_1 + \lambda^2 \mathbf{G}^T \mathbf{G}) \mathbf{Z}_{:,i} = \mathbf{W}_{:,i} \quad (19)$$

where  $\mathbf{Z}_{:,i}$  and  $\mathbf{W}_{:,i}$  denote the  $i$ th column of matrices  $\mathbf{Z}$  and  $\mathbf{W}$ , respectively, and  $[\Sigma_{\text{HbO}}]_{i,i}$  denotes the  $i$ th diagonal component of diagonal matrix  $\Sigma_{\text{HbO}}$ ,  $i = 1, 2, \dots, L_1$ . After solving all  $L_1$  systems, the estimate of  $\Gamma_k^T$  can be obtained by

$$\Gamma_k^T = \mathbf{Z} \mathbf{V}_{\text{HbO}}^T. \quad (20)$$

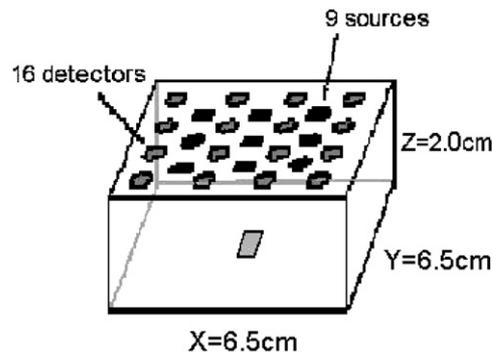
The transform  $\mathbf{V}_{\text{HbO}}$  is orthogonal and thus will keep optimality in the least-squares sense for both  $\mathbf{Z}$  and  $\Gamma_k^T$ . The second equation in equation (17) can be solved in a similar manner, with orthogonal transform  $\mathbf{V}_{\text{HbR}}$  obtained from SVD of  $\mathbf{F}_3$ :  $\mathbf{F}_3 = \mathbf{V}_{\text{HbR}} \Sigma_{\text{HbR}} \mathbf{V}_{\text{HbR}}^T$ .

In summary, the iteration method only needs to solve  $L_1 + L_2$  smaller systems which have the size of the number of imaging pixels at each iteration. In addition, all the matrices in equation (16), the SVDs of  $\mathbf{F}_1$  and  $\mathbf{F}_3$ , and all  $L_1 + L_2$  inverse matrices (in the case that the regularization parameter  $\lambda$  is fixed) can be calculated once at the beginning and used for all iterations, to save both in computation and memory. The augmented matrix in equation (15) has to have block diagonal dominance to guarantee convergence, and the convergence speed is a function of the degree to which the diagonal block matrices dominate. This dominance turned out to be the case for all our simulations.

### 3. Simulation results

To illustrate this framework of joint spatio-temporal DOT reconstructions with an HRF model, and explore possible improvement over the traditional reconstruction method, we present some simulations of human brain activation.

Simulations were performed in a semi-infinite geometry, with 9 sources and 16 detectors placed on the air–tissue plane to collect reflection measurements, as shown schematically in figure 1.

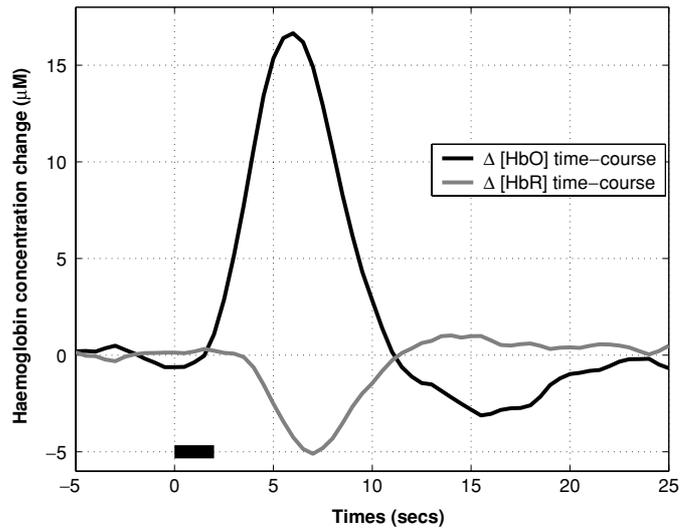


**Figure 1.** Geometric arrangement of sources and detectors: 9 sources and 16 detectors are put on the air–tissue boundary plane in a standard ‘9 within 16’ interlaced grid pattern. The minimum source–source and source–detector separations are 2.0 cm and 1.41 cm, respectively. Continuous-wave measurements are simulated at two wavelengths: 690 nm and 830 nm.

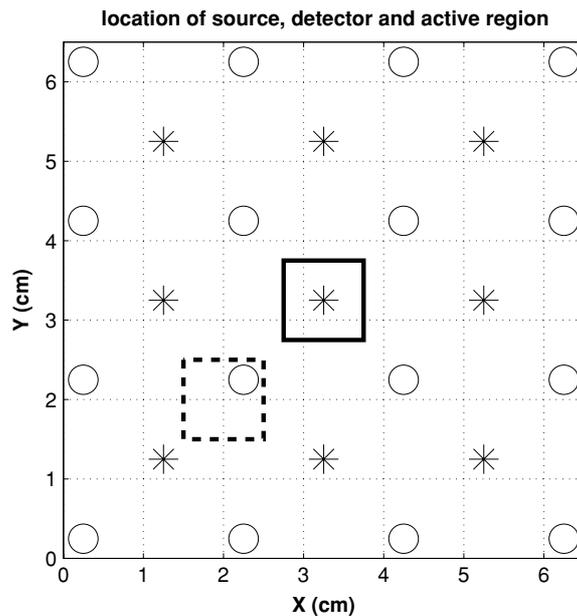
The minimum source–source and source–detector separations are 2.0 cm and 1.41 cm, respectively. In real applications, due to the limited dynamic range of the detectors, measurements on channels (source–detector pairs) with too large a separation are dominated by noise. So here we only use the measurement of the first (1.41 cm) and the second (3.16 cm) nearest neighbouring channels, altogether 84 channels. Continuous-wave (CW) measurements are simulated at two wavelengths: 690 nm and 830 nm. Only the two-dimensional  $X$ – $Y$  slice where an anomalous region is assumed to be centred was reconstructed. The thickness of this slice is 0.25 cm. The area imaged is  $6.5 \times 6.5 \text{ cm}^2$ , with the size of the imaging pixel  $0.25 \times 0.25 \text{ cm}^2$ , resulting in a grid of  $26 \times 26 = 676$  pixels. The forward matrix  $\mathbf{A}$  is obtained from a first-order Born approximation, and the same model is used to generate simulated measurements (i.e., scattering photon fluence caused by the anomalous region) and in reconstruction. White Gaussian noise at  $\text{SNR} = 25 \text{ dB}$  was added to the simulated measurements.

The optical properties of the homogeneous medium are  $\mu'_s = [10, 10] \text{ cm}^{-1}$  and  $\mu_a = [0.0732, 0.0760] \text{ cm}^{-1}$  for 690 and 830 nm, respectively. A cubic anomaly with length 1 cm was positioned in the plane 2 cm below the boundary plane. The shape of the anomaly is fixed, but its  $\Delta[\text{HbO}]$  and  $\Delta[\text{HbR}]$  vary with time, as shown in figure 2. These assumed true concentration changes were obtained by averaging adult brain activation evoked by a finger tapping task with 2 s duration (indicated by the solid bar in figure 2), estimated from multiple-subject measurements, as described in Boas *et al* (2003, 2004). Note we set the shape of the haemoglobin concentration changes as the same as those presented in (Boas *et al* 2003), while making the amplitude comparable to the haemoglobin responses in (Boas *et al* 2004). These haemoglobin concentration changes represent a typical response on the part of the motor cortex, showing the typical pattern with an event-related increase in  $\Delta[\text{HbO}]$  accompanied by a smaller decrease in  $\Delta[\text{HbR}]$ , delayed approximately 2 s from stimulus onset. Following stimulus onset, the  $\Delta[\text{HbO}]$  peaked at 6 s and the  $\Delta[\text{HbR}]$  peaked at 7 s. In addition, to test the effect of the anomaly location, we set two cases of the anomaly position, centred and off-centred, as shown in figure 3.

We designed three different sets of basis function. In the first case, we used two Gaussian curves with ‘correct’ mean and standard deviation as the basis for HbO: one for the large positive peak and the other for the later undershoot peak. Similar settings were used for HbR: one for the decreasing part and the other for the overshoot part. These basis functions are

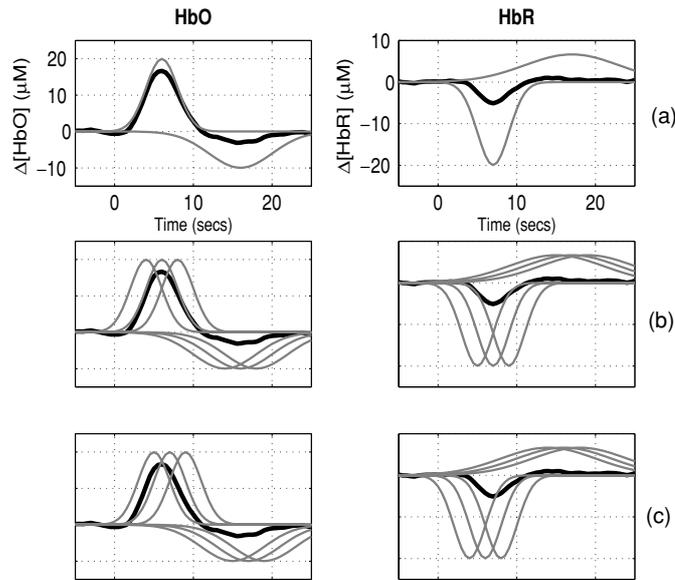


**Figure 2.** Preset concentration changes for both  $\Delta[\text{HbO}]$  and  $\Delta[\text{HbR}]$  on the active region. The assumed true concentration changes were obtained by averaging adult brain activation evoked by a finger tapping task with 2 s duration (indicated by the solid bar), which represent a typical response on the part of the motor cortex.



**Figure 3.** X–Y location of centred (solid square) and off-centred (dashed square) active regions, with source and detector positions shown by asterisks and open circles, respectively.

shown in the panels in row (a) in figure 4 as thin grey lines (thick dark lines in all panels indicate the true  $\Delta[\text{HbO}]$  and  $\Delta[\text{HbR}]$  time courses). These settings represent the ‘best’ case, in which we know the temporal pattern of haemoglobin concentration changes very well,

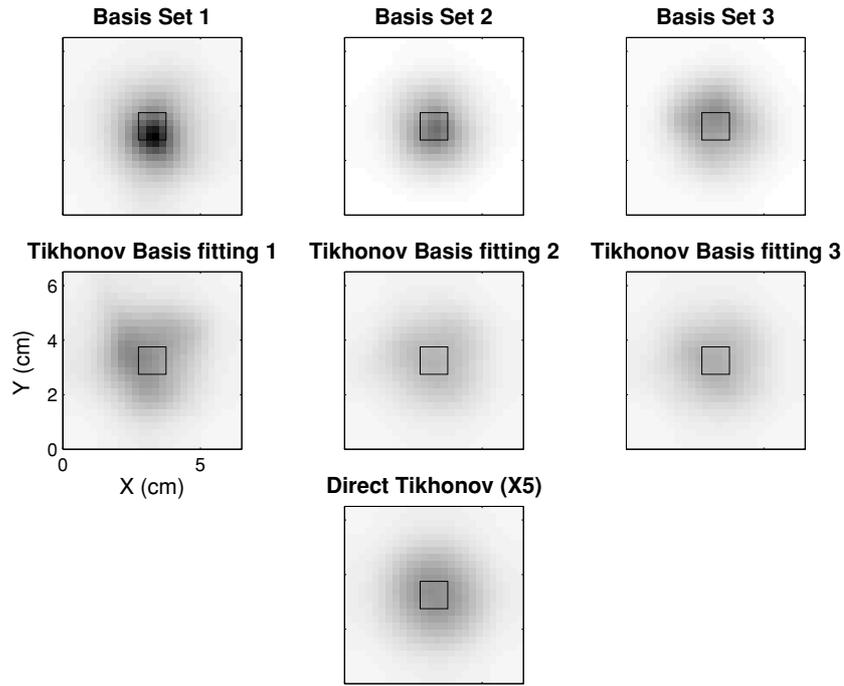


**Figure 4.** Different sets of basis functions (thin grey lines) with true concentration changes (thick dark lines). Basis functions for sets 1, 2 and 3 are shown in rows (a), (b) and (c), respectively. Panels on left column are for HbO and right for HbR response.

except for a scalar amplitude factor. In the second case, we used six Gaussian curves as basis functions for HbO, which included the basis in set 1 and four other signals by shifting  $\pm 2$  s, as shown in row (b) of figure 4. This case mimics the situation in which the peak-value time of the haemodynamic response will vary across subjects, but the shapes of the response are quite similar and known. We ‘guess’ the true peak-value time by using multiple similar basis with different means, and one of the several basis functions happens to be correct. In the third case, we varied the setting by shifting the basis functions in set 2 by 1 s, such that none of the basis functions has the correct peak-value time, as shown in row (c) of figure 4. The corresponding reconstructions are referred to in the sequence as *basis sets 1, 2 and 3*.

To compare with the proposed framework, we also used traditional Tikhonov regularization to solve the variations of the haemoglobin concentrations at each time instant independently, while employing the same gradient-type regularization matrix as a spatial constraint. Note that we use a direct reconstruction scheme which imposes prior spectral information into the image reconstruction rather than a two-step process of reconstructing absorption coefficient and then calculating haemodynamic concentration changes (Li *et al* 2004, 2005). In this way, the improvement of the proposed framework can be directly attributed to the temporal information. Afterward, the temporal dynamics of the haemoglobin concentrations are obtained by extracting time series from the results and are referred to as *Direct Tikhonov*.

To further illustrate the performance of the proposed joint spatio-temporal reconstruction framework, we fit the results of *Direct Tikhonov* by the three sets of temporal basis functions described above, with a least-squares criterion. This procedure used a posterior temporal smoothing based on the same level of prior temporal information, represented by the basis functions, in contrast with the joint spatio-temporal scheme. These results are referred to as *Tikhonov basis fitting 1, 2 and 3*.



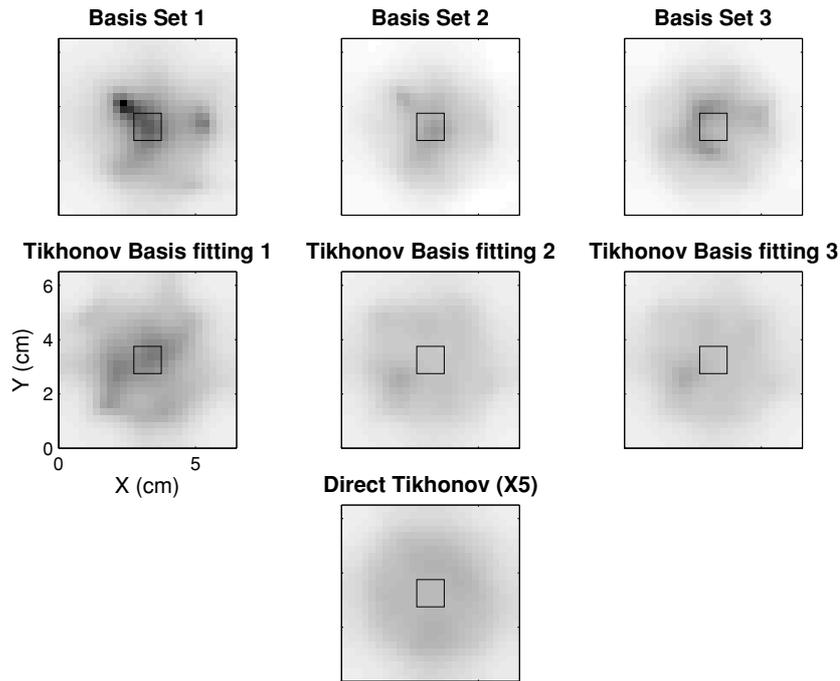
**Figure 5.** Spatial CNR maps of the centred active region for HbO, comparing different sets of basis functions, Direct Tikhonov regularization with and without posterior basis fitting. All panels have same colourmap, but CNR values of Direct Tikhonov regularization are amplified by a factor of 5. The true active region is indicated by the solid square.

To quantitatively compare the basis function model to the Direct Tikhonov method, with and without posterior basis function fitting, we calculate a spatial map of contrast-to-noise (CNR) ratio. For each imaging pixel, the CNR is given by the mean difference between the pre-stimulus and the post-stimulus responses, divided by the mean standard deviation across different noise realizations. Specifically, suppose at one pixel, we have reconstructed time courses of either  $\Delta[\text{HbO}]$  or  $\Delta[\text{HbR}]$ , for multiple noise realizations, and we denote the selected period of pre-stimulus and post-stimulus responses by  $\text{pre}_i^n$  and  $\text{post}_i^n$ , then the CNR is given by

$$\text{CNR} = \frac{\widetilde{\text{pre}} - \widetilde{\text{post}}}{\sqrt{[\text{mean}(\text{Std}_{\text{pre}})]^2 + [\text{mean}(\text{Std}_{\text{post}})]^2}}, \quad (21)$$

where  $\widetilde{\text{pre}}$  and  $\widetilde{\text{post}}$  are the average value over all time instants and all noise realizations, and  $\text{Std}_{\text{pre}}$  and  $\text{Std}_{\text{post}}$  are the standard deviation of the same intervals of the time courses over noise realizations. In this simulation, we generated ten noise realizations for each active region case (centred and off-centred). The duration of the pre-stimulus response was 5 s (−5–0 s) and the duration of the post-stimulus response was also 5 s (3–8 s for HbO and 4–9 s for HbR), where 0 s indicates the onset of the stimulation.

Figures 5 and 6 show examples of the CNR spatial map for the centred active region cases of HbO and HbR, respectively. The solid square indicates the location of the active region. In this case, the regularization parameter used in all three sets of basis functions was  $\lambda = 5 \times 10^{-7}$ , while  $\lambda = 1$  was used for Direct Tikhonov regularization. Since the regularization parameter could affect the resolution of imaging procedures differently, we altered the regularization



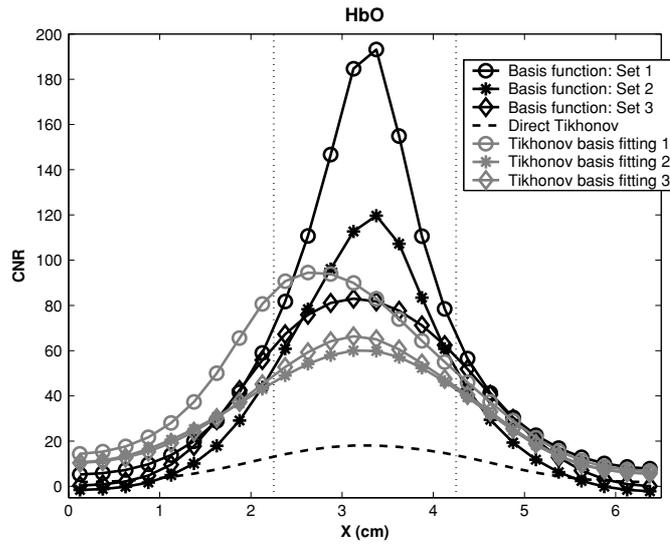
**Figure 6.** Comparison of the spatial CNR map of centred active region for HbR, using different sets of basis functions, Direct Tikhonov regularization with and without posterior basis fitting.

parameter for each to approximately equalize the full width at half-maximum (FWHM) of the reconstructed HbO or HbR object at peak-value time. For clear comparison of the results, we multiply the CNR values of the Direct Tikhonov method by a factor of 5 for both HbO and HbR results. For all simulation results, we have similar observations for both centred and off-centred active region cases, so we only show the results of centred one.

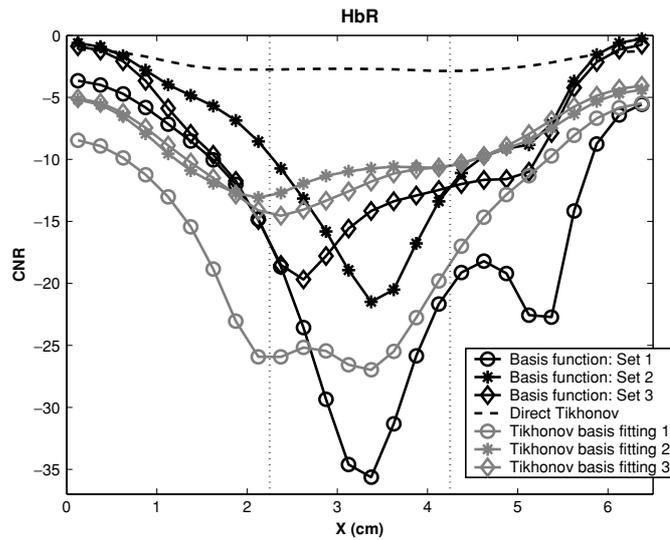
One can see that for both HbO and HbR, the basis function model gives a much better CNR map than the Direct Tikhonov method, which did not use any temporal information, in terms of magnitude and concentration. The posterior basis fitting on the Direct Tikhonov solution removed some noisy variation and thus resulted in significantly improved CNR maps, but not to the same degree as the basis function model, when the same level of temporal information was used.

Clearer observations can be obtained from figures 7 and 8 where we plot a one-dimensional CNR profile along the horizontal line  $Y = 3.5$  cm that passes through the active region centre. The thin dotted lines indicate the correct range of the active region. For figures 7 and 8, we do NOT multiply the results of the Direct Tikhonov method by a factor of 5, as was done in figures 5 and 6, to clearly show the magnitude gap between different methods. With either the basis function model or posterior basis fitting, basis set 1 gave the best results as expected because it used the most accurate temporal prior information.

To demonstrate the effect of the regularization parameter, we tested different parameter values for Direct Tikhonov regularization and the basis function model, and plotted curves of mean CNR value, FWHM and mean-square-error (MSE) as a function of  $\lambda$  in figure 9, for HbO results of centred active region. Mean CNR value was calculated over the true active region area in the spatial CNR maps; FWHM was estimated from the reconstructed object at



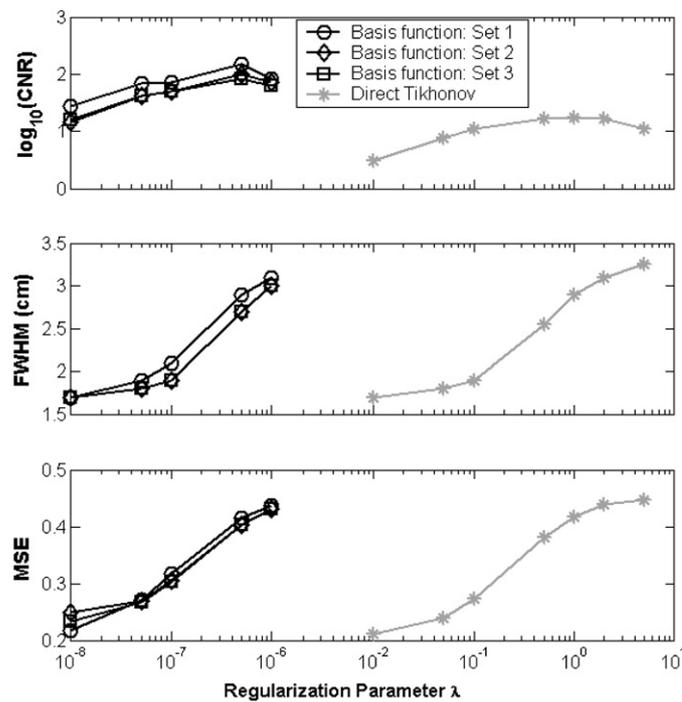
**Figure 7.** Comparison of CNR values on horizontal line for  $Y = 3.5$  cm, for HbO results. The true active region range is indicated by the thin dotted lines.



**Figure 8.** Comparison of CNR values on horizontal line for  $Y = 3.5$  cm, for HbR results. The true active region range is indicated by the thin dotted lines.

peak-value time; MSE was defined by the difference between the reconstructed time series at the pixel of true active region centre and the true temporal changes in figure 2. For FWHM and MSE values, the reconstructed object and temporal courses are average over results of 10 noise realizations.

From figure 9, one can see that the two different imaging methods, Direct Tikhonov regularization and the basis function model, used very different ranges of regularization parameters. For the mean CNR value curve, when the regularization level was low, CNR



**Figure 9.** Curves of mean CNR value over active region area, FWHM of reconstructed object and MSE of time courses at active region centre as a function of regularization parameters, for HbO of centred active region case.

values were small due to the large variation across noise realizations. As more regularization was used, solutions became more stable and resulted in an increasing CNR. After an optimal regularization level, the solutions became too smooth and CNR values decreased due to the reduced contrast between pre-stimulus and post-stimulus responses. This observation was true for both basis function models and the Direct Tikhonov method. When an appropriate  $\lambda$  was applied, basis function models can give much better mean CNR values than the Direct Tikhonov method.

In the FWHM curves, when regularization level was very low, the reconstructed spatial map was very noisy and resulted in a small FWHM, or a map containing multiple FWHM contours which were far from the true active region shape so that one cannot calculate the full width (this case corresponds to regularization parameter smaller than  $10^{-8}$ ). As more regularization was applied, the reconstructed objects became more smooth, resulting in a larger full width at half-maximum level. (In the extreme case of infinite regularization, the reconstructed image is totally homogeneous and the full width is also infinite.) This observation is true for both basis function models and the Direct Tikhonov method.

For the MSE curves of time courses at the active region centre, there is an optimal level for regularization associated with the minimum MSE. Away from this level, the MSE will increase because of either noisy (corresponding to under-regularization) or over-smoothed (corresponding to over-regularization) time courses. Here, the optimal regularization level for the MSE curve is at the left boundary, for both basis function models and the Direct Tikhonov method.

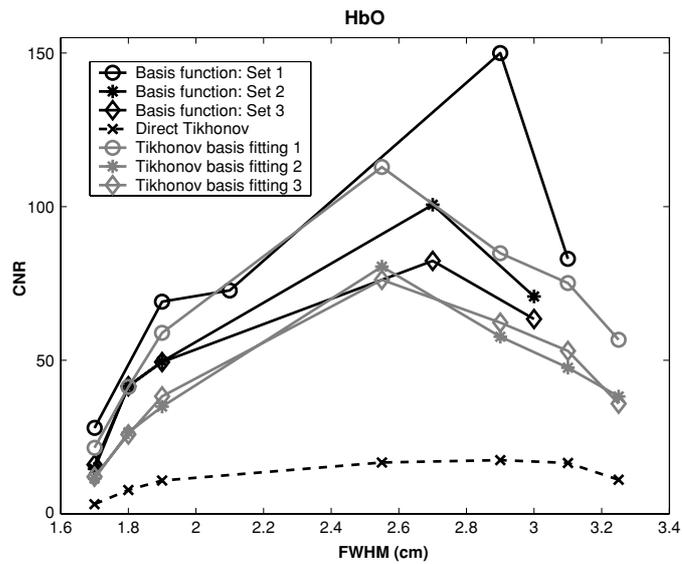


Figure 10. CNR versus FWHM for HbO reconstruction of centred active region.

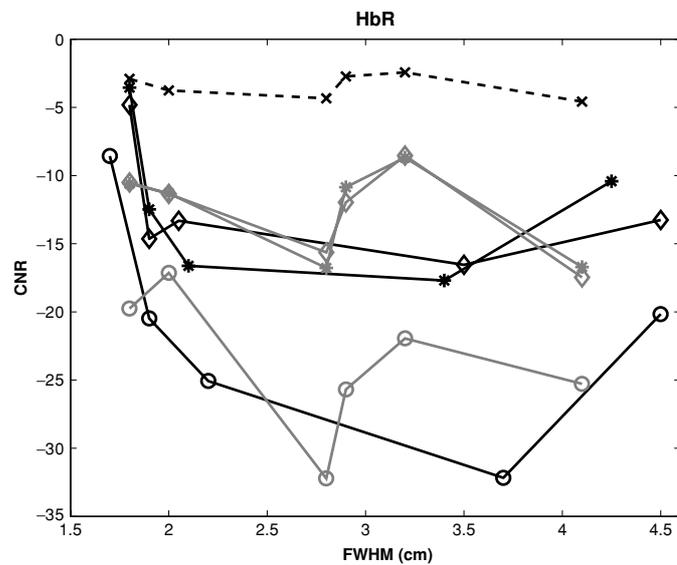


Figure 11. CNR versus FWHM for HbR reconstruction of centred active region.

To evaluate the spatial resolution of overall reconstruction results, we plotted FWHM against mean CNR value for HbO and HbR in figures 10 and 11, respectively. One can see that for fixed spatial resolution (FWHM), the basis function models gave a much larger CNR value for HbO, with relatively small improvement for HbR, than Direct Tikhonov results. The smaller improvement of CNR values for HbR is due to the weaker contrast in HbR response. And as expected, results of the basis model with set 1 were better than those of sets 2 and 3, while the latter two sets gave similar results. For the Tikhonov solution with temporal basis fitting, the contrasts at fixed spatial resolution were worse than the basis function model, when

the same set of basis function were used. Note that the actual FWHM for the true active region area is  $\sqrt{2} \approx 1.4$  cm, but all methods gave bad CNR values at this resolution, due to the ill-posedness of the inverse problem. Regularization based on gradient-type constraint was used to suppress the amplification of the measurement noise but it simultaneously added smoothness to the solutions, resulting in a larger FWHM.

#### 4. Discussion

From the simulation results, we see that when appropriate regularization has been applied on the linear coefficients space, the basis function model can give better solutions in terms of much higher CNR values and better spatial resolution, compared with the Direct Tikhonov method, with or without posterior basis fitting. We have similar observations for both centred and off-centred active region cases. Note that the reconstructed time courses of the basis function model are more stable (or smooth) than those of the Direct Tikhonov method without posterior basis fitting, because they are linear combinations of the basis functions and are smooth due to the smoothness of the basis. In contrast, results of the Direct Tikhonov method generally will have more variability in time due to the fact that there is no smoothness assumption or constraint applied on the time course. In the bottom panel of figure 9, we show the mean-square-error of reconstructed time courses at the active region centre, but there we compared the averaged time courses over ten noise realizations to the true haemoglobin concentration time series, so it is hard to judge the smoothness of individual time course.

The efficient algorithm we developed employs a very simple blockwise iteration method and takes advantage of Kronecker product structure in the blocks of the normal equation. We note that the Kronecker product and its properties have been used in optical reconstruction algorithms, see Ren and Dekany (2004) for an example. It significantly reduces the high computational complexity and huge memory burden of the proposed framework. Simulations with different noise realizations show stable, fast convergence results requiring less than 20 iterations (approximately 1 min on a 1.50 GHz Pentium computer).

Selecting the regularization parameters in both the Direct Tikhonov method and the proposed joint spatio-temporal framework is still an open problem. This regularization parameter controls the confidence level between data fidelity and constraint satisfaction, and thus will significantly affect solutions, as the simulation results show.

Basis functions in set 1 gave the best results among the three different choices of base, as expected, while sets 2 and 3 gave comparably good results, showing that the proposed framework worked well in these simulations. This indicates that in the realistic situation when the timing of the peak value of haemoglobin changes is unknown or varies across subjects, but the basic shape of the response is known, we can still get good estimates with basis function model by introducing several shifts of the basis functions.

HRF with a basis function expansion, on one hand, can incorporate prior knowledge of the response shape into the reconstruction; on the other hand, it may impose constraints or add bias to the solution, and it is hard to reconstruct solutions which happen to be far from the basis function space. One possible solution to address this problem is to use parametric models, solving a nonlinear optimization problem, an alternative also frequently used in fMRI data analysis. This model can add more flexibility in the reconstruction of more complex shape response (Izzetoglu *et al* 2005). Another solution is to design HRF model directly from the underlying physiological process, e.g. the Balloon model proposed by Buxton *et al* (1998) to describe the dynamics of the deoxy-haemoglobin, cerebral blood volume and cerebral blood flow.

Note that the basis functions in either set cannot approximate the true haemoglobin concentration changes exactly, and a small error will remain in the least-squares solution in fitting the true curves. This intrinsic error cannot be reduced by any regularization, but the basis functions used in all three sets in our simulation only result in a very small error, i.e., the true haemodynamic changes mostly fall into the basis function subspace.

For real experimental data in a functional brain imaging study, we need to solve several problems before using this framework. First, the measurements are sensitive to cerebral haemodynamic fluctuations of systemic origin. Such systemic fluctuations are associated, for instance, with arterial pulse, respiration and heart rate fluctuations (Franceschini *et al* 2000, Obrig *et al* 2000, Toronov *et al* 2000), which appear as interfering signals, obscuring or even overwhelming the activation-evoked response. We need to reduce this physiological interference before, or while, performing the reconstruction. Second, one needs more overlapping channels for the inverse tomographic problem. In the case that only a few channels are available, a simplified linear model, such as the modified Beer–Lambert law (MBLL), can be employed to estimate the haemodynamic response measured with individual channels and then interpolate them to produce spatial images (Franceschini *et al* 2000, Fantini *et al* 1999). Third, to deal with real experimental data, the accuracy of the head modelling is also crucial to reconstruction results. In this simulation, we simplify the scenario to a case that a homogeneous active region with time-varying haemoglobin concentration variation is embedded in a different homogeneous background medium. With this model, we can only obtain relative haemoglobin changes in settings where we use real measured data. To get absolute changes, more complicated and accurate modelling, or the inhomogeneity of the head such as the use of a layered model structures, should be included into the method (Strangman *et al* 2003, Boas and Dale 2005). Finally, in real experiments, accurate reconstructions need to model different basis functions, since in principle the HRF may vary across different subjects, different areas on the brain or different tasks. This in turn requires more prior information about the HRF basis.

In spite of these difficulties, we believe the proposed method is very promising based on our simulation results. The simulation scenario, although it makes some favourable assumptions and uses measurements free of physiological interferences, has taken some practical considerations into account, such as not knowing the true peak-value time and using shifted basis functions. Moreover, the true HRF used is actually calculated from measured data (Boas *et al* 2003, 2004), and this is not exactly matched by the Gaussian-bump model employed. The geometrical arrangement of the source–detector pairs is very similar to a real experimental setting and only the measurement of the first and the second nearest neighbouring channels is used in reconstruction, which is also close to a realistic case. In addition, the efficient algorithm successfully keeps the computational complexity to the order of the spatial dimension and is linear in the number of total basis functions. So in practice, one can try quite a large number of different bases while not worrying overly about heavy computation.

## 5. Conclusion

In this work, we present a joint spatio-temporal reconstruction scheme which incorporates spatial regularization, temporal information and spectral information together for diffuse optical tomography in functional brain imaging. The proposed method shows significant improvements in contrast-to-noise ratio and spatial resolution compared with traditional frame-by-frame Tikhonov regularization solutions in which no temporal information was used, based on simulation results. Moreover, when the same level of temporal information in

basis functions is available, it is shown that the proposed method is better than posterior basis fitting of the Tikhonov solution.

## Acknowledgments

The work of the first two authors was supported in part by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (award number EEC-9986821), the third author by National Institutes of Health grants R01-EB002482 and P41-RR14075. The authors thank Sol Diamond for helpful discussions.

## References

- Aguirre G K, Zarahn E and D'Esposito M 1998 The variability of human, BOLD hemodynamic responses *Neuroimage* **8** 360–9
- Arridge S R, Schweiger M, Hiraoka M and Delpy D T 1993 Performance of an iterative reconstruction algorithm for near infrared absorption imaging *Proc. SPIE* **1888** 360–71
- Barbour R L, Graber H L, Pei Y, Zhong S and Schmitz C H 2001 Optical tomographic imaging of dynamic features of dense-scattering media *J. Opt. Soc. Am. A* **18** 3018–36
- Barbour R L, Graber H L, Schmitz C H, Pei Y, Zhong S, Barbour S S, Blattman S and Panetta T 1999 Spatio-temporal imaging of vascular reactivity by optical tomography *Proc. Inter-Institute Workshop on In Vivo Optical Imaging at the NIH* pp 161–6
- Boas D A and Dale A M 2005 Simulation study of magnetic resonance imaging-guided cortically constrained diffuse optical tomography of human brain function *Appl. Opt.* **44** 1957–68
- Boas D A, Dale A M and Franceschini M A 2004 Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution, and accuracy *Neuroimage* **23** 275–88
- Boas D A, Franceschini M A, Dunn A K and Strangman G 2002 Noninvasive imaging of cerebral activation with diffuse optical tomography *In Vivo Optical Imaging of Brain Function* ed R D Frostig chapter 8 (Boca Raton, FL: CRC Press) pp 193–221
- Boas D A, Strangman G, Culver J P, Hoge R D, Jaszewski G, Poldrack R A, Rosen B R and Mandeville J B 2003 Can the cerebral metabolic rate of oxygen be estimated with near-infrared spectroscopy? *Phys. Med. Biol.* **48** 2405–18
- Boynton G, Engel S, Glover G and Heeger B 1996 Linear systems analysis of functional magnetic resonance imaging in human V1 *J. Neurosci.* **16** 4207–21
- Buckner R L 1998 Event-related fMRI and the hemodynamic response *Hum. Brain Mapp.* **6** 373–7
- Buxton R B, Wong E C and Frank L R 1998 Dynamics of blood flow and oxygenation changes during brain activation: the balloon model *Magn. Reson. Med.* **39** 855–64
- Dale A M and Buckner R L 1997 Selective averaging of rapidly presented individual trials using fMRI *Hum. Brain Mapp.* **5** 329–40
- Delpy D T, Cope M, Van der Zee P, Arridge S, Wray S and Wyatt J 1988 Estimation of optical pathlength through tissue from direct time of flight measurement *Phys. Med. Biol.* **33** 1433–42
- Fantini S, Franceschini M A, Gratton E, Hueber D, Rosenfeld W, Maulik D, Stubblefield P G and Stankovic M R 1999 Non-invasive optical mapping of the piglet brain in real time *Opt. Express* **4** 308–14
- Franceschini M A, Fantini S, Toronov V, Filiaci M E and Gratton E 2000 Cerebral hemodynamics measured by near-infrared spectroscopy at rest and during motor activation *OSA In Optical Imaging Workshop* ed A Gandjbakhche 73–80
- Franceschini M A, Toronov V, Filiaci M E, Gratton E and Fantini S 2000 On-line optical imaging of the human brain with 160 ms temporal resolution *Opt. Express* **6** 49–57
- Friston K J, Fletcher P, Josephs O, Holmes A, Rugg M D and Turner R 1998 Event-related fMRI: characterizing differential responses *Neuroimage* **7** 30–40
- Friston K J, Jezzard P and Turner R 1994 Analysis of functional MRI time series *Hum. Brain Mapp.* **1** 153–71
- Friston K J, Josephs O, Zarahn E, Holmes A, Rouquette S and Poline J 2000 To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis *Neuroimage* **12** 196–208
- Gaudette R J, Brooks D H, DiMarzio C A, Kilmer M E, Miller E L, Gaudette T and Boas D A 2000 A comparison study of linear reconstruction techniques for diffuse optical tomographic imaging of absorption coefficient *Phys. Med. Biol.* **45** 1051–70

- Glover G H 1999 Deconvolution of impulse response in event-related bold fMRI *Neuroimage* **9** 416–29
- Gossl C, Fahrmeir L and Auer D P 2001 Bayesian modeling of the hemodynamic response function in bold fMRI *Neuroimage* **14** 140–8
- Graham A 1981 *Kronecker Products and Matrix Calculus: With Applications* (New York: Halsted Press/Wiley)
- Hillman E M C 2002 Experimental and theoretical investigations of near infrared tomographic imaging methods and clinical applications *PhD Thesis* University College London, London
- Hoshi Y and Tamura M 1993 Detection of dynamic changes in cerebral oxygenation coupled to neuronal function during mental work in man *Neurosci. Lett.* **150** 5–8
- Ishimaru A 1978 *Wave Propagation and Scattering in Random Media* (San Diego: Academic)
- Izzetoglu M, Nioka S, Chance B and Onaral B 2005 Single trial hemodynamic response estimation in a block anagram solution study using fNIR spectroscopy *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)* pp 633–6
- Kak A C and Slaney M 1988 *Principles of Computerized Tomographic Imaging* (New York: IEEE)
- Kolehmainen V, Prince S, Arridge S R and Kaipio J P 2003 State-estimation approach to the nonstationary optical tomography problem *J. Opt. Soc. Am. A* **20** 876–89
- Li A, Boverman G, Zhang Y, Brooks D H, Miller E L, Kilmer M E, Zhang Q, Hillman E M C and Boas D A 2005 An optimal linear inverse solution given multiple priors in diffuse optical tomography *Appl. Opt.* **44** 1948–56
- Li A, Zhang Q, Culver J P, Miller E L and Boas D A 2004 Reconstructing chromosphere concentration images directly by continuous-wave diffuse optical tomography *Opt. Lett.* **29** 256–8
- Marrelec G, Benali H, Ciuciu P and Poline J-B 2002 Bayesian estimation of the hemodynamic response function in functional MRI *AIP Conf. Proc.* **617** 229–47
- Marrelec G, Ciuciu P, Plgrini-Issac M and Benali H 2004 Estimation of the hemodynamic response in event-related functional MRI: Bayesian networks as a framework for efficient Bayesian modeling and inference *IEEE Trans. Med. Imaging* **23** 959–67
- Obrig H, Neufang M, Wenzel R, Kohl M, Steinbrink J, Einhaupl K and Villringer A 2000 Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults *Neuroimage* **12** 623–39
- O’Leary M 1996 *Imaging with diffuse photon density waves PhD Thesis* University of Pennsylvania, Philadelphia, Pennsylvania
- Pogue B W, McBride T O, Prewitt J, Osterberg Ulf L and Paulsen K D 1999 Spatially variant regularization improves diffuse optical tomography *Appl. Opt.* **38** 2950–61
- Prince S, Kolehmainen V, Kaipio J P, Franceschini M A, Boas D A and Arridge S R 2003 Time-series estimation of biological factors in optical diffusion tomography *Phys. Med. Biol.* **48** 1491–504
- Rajapakse J C, Kruggel F, Maisog J M and Cramon D Yves von 1998 Modeling hemodynamic response for analysis of functional MRI time-series *Hum. Brain Mapp.* **6** 283–300
- Ren H and Dekany R 2004 Fast wave-front reconstruction by solving the Sylvester equation with the alternating direction implicit method *Opt. Express* **12** 3279–96
- Strangman G, Franceschini M A and Boas D A 2003 Factors affecting the accuracy of near-infrared spectroscopy concentration calculations for focal changes in oxygenation parameters *Neuroimage* **18** 865–79
- Toronov V, Franceschini M A, Filiaci M, Fantini S, Wolf M, Michalos A and Gratton E 2000 Near-infrared study of fluctuations in cerebral hemodynamics during rest and motor stimulation: temporal analysis and spatial mapping *Med. Phys.* **27** 801–15
- Villringer A and Chance B 1997 Non-invasive optical spectroscopy and imaging of human brain function *Trends Neurosci.* **20** 435–42
- Villringer A, Planck J, Hock C, Schleinkofer L and Dirnagl U 1993 Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults *Neurosci. Lett.* **154** 101–4
- Woolrich M W, Jenkinson M, Brady J M and Smith S M 2004 Fully Bayesian Spatio-temporal modeling of fMRI data *IEEE Trans. Med. Imaging* **23** 213–31
- Zhang Y, Ghodrati A and Brooks D H 2005 An analytical comparison of three spatio-temporal regularization methods for dynamic linear inverse problems in a common statistical framework *Inverse Problems* **21** 357–82