

Publication IV

P. Hiltunen, S. Särkkä, I. Nissilä, A. Lajunen, and J. Lampinen. State space regularization in the nonstationary inverse problem for diffuse optical tomography. *Inverse problems*, 27(2), 025009 (15 pages), 2011.

© 2011 IOP Publishing Ltd.
Reprinted with permission.

State space regularization in the nonstationary inverse problem for diffuse optical tomography

P Hiltunen¹, S Särkkä¹, I Nissilä^{1,2}, A Lajunen¹ and J Lampinen¹

¹ Department of Biomedical Engineering and Computational Science, Aalto University School of Science and Technology, PO Box 12200, FI-00076 AALTO, Finland

² Biomag Laboratory, HUSLAB, Helsinki University Central Hospital, PO Box 340, FI-00029 HUS, Helsinki, Finland

E-mail: petri.hiltunen@tkk.fi, simo.sarkka@tkk.fi, ilkka.nissila@tkk.fi, atte.lajunen@tkk.fi and jouko.lampinen@tkk.fi


Received 9 July 2010, in final form 2 December 2010

Published 21 January 2011

Online at stacks.iop.org/IP/27/025009

Abstract

In this paper, we present a regularization method in the nonstationary inverse problem for diffuse optical tomography (DOT). The regularization is based on a choosing time evolution process such that in a stationary state it has a covariance function which corresponds to a process with similar smoothness properties as the first-order smoothness Tikhonov regularization. The proposed method is computationally more lightweight than the method where the regularization is augmented as a measurement. The method was tested in the case of the inverse problem of DOT. A solid phantom with optical properties similar to tissue was made, incorporating two moving parts that simulate two different physiological processes: a localized change in absorption and a surrounding rotating two-part shell which simulates slow oscillations in the tissue background physiology. A sequence of measurements of the phantom was made and the reconstruction of the image sequence was computed using this method. It allows the recovery of the full time series of images from relatively slow measurements with one source active at a time. In practice, this allows instruments with a larger dynamic range to be applied to the imaging of functional phenomena using DOT.

 Online supplementary data available from stacks.iop.org/IP/27/025009/mmedia

(Some figures in this article are in colour only in the electronic version)

1. Introduction

1.1. Diffuse optical tomography

Diffuse optical tomography (DOT) is a functional imaging method which is used to study the physiology of living tissue non-invasively. Typical applications include the imaging of hemoglobin concentration and oxygen saturation in the breast, muscle and brain tissue during interventions, stress or sensory stimulation of the brain. The time scale of hemodynamic events in these experiments is from hundreds of milliseconds to minutes.

To obtain good sensitivity to the deepest parts of the tissue being imaged, it is essential to include fully transmissive measurements where the source and detector optodes are positioned at opposite sides of the tissue. However, if high resolution of the superficial parts of the tissue is desired, a high density of optodes over the region of interest is needed, with relatively short source-to-detector separations. Combining both types of measurements is technically challenging as a large dynamic range is needed. In practical implementations of DOT systems, there is a trade-off between imaging time, dynamic range and the number of source optodes. To minimize cross-talk, most instruments use the time-multiplexed approach, in which each source is activated sequentially. In order to take advantage of the superior 3D imaging capability of the relatively slow instruments that have the highest sensitivity at low light levels and the largest dynamic range for practical applications that involve the imaging of relatively fast functional phenomena, a method of reconstructing the time series of 3D images so that the reconstruction is updated based on the data corresponding to a single active source is needed. In the state space method, the image corresponding to a time step is reconstructed based on the modelled previous state of the tissue and the data corresponding to the current time step, which includes data for each detector but only one source. This allows the reconstructed time series to have a higher temporal resolution and greater accuracy of interpretation of the time-multiplexed data than if the full data corresponding to all the source optodes were required to reconstruct a single time step. If only the data corresponding to one source is used to reconstruct images for each time step, it will lead to images with reduced contrast and incorrect reconstruction of the position of the perturbation, as shown by Kolehmainen *et al* (2003).

Furthermore, it is possible to include auxiliary data (such as global physiological signals) in the state space model, as suggested in Diamond *et al* (2006).

1.2. State space method

The state space method consists of two models, an evolution model and a measurement model. The evolution model describes the temporal dynamics of the process under study. In DOT we are mostly interested in the time evolution of the optical properties x_k , such as the absorption coefficient μ_a . We can assume that it is a stochastic Markov process $x_k = Ax_{k-1} + q_{k-1}$, $q_k \sim N(0, Q_k)$, where q_k is a state noise and k is a state index. The transition matrix A describes how optical properties evolve during time instances t_{k-1} and t_k . The model could include a physiological model of the process x_k and auxiliary inputs, such as blood pressure and heart rate, as in Diamond *et al* (2006), or it could be a random walk process, i.e. $A = I$, as in Kolehmainen *et al* (2003) and Prince *et al* (2003).

The measurement model gives a physical observation of the state x_k . At every time instance t_k a measurement y_k of the process is made. In DOT the forward model is a nonlinear function. In this section we present the state space method using a linear model for simplicity. In this case the observation model is a stochastic equation $y_k = F_k x_k + r_k$, $r_k \sim N(0, R_k)$, where F_k is an observation matrix and r_k is an observation noise.

This *evolution-observation model* has a well-known solution, the Kalman filter (Grewal and Andrews 2001), which consists of a prediction step

$$\begin{aligned} m_k^- &= Am_{k-1} \\ P_k^- &= AP_{k-1}A^T + Q_{k-1}, \end{aligned} \quad (1)$$

and an estimation step

$$\begin{aligned} S_k &= F_k P_k F_k^T + R_k \\ K_k &= P_k^- F_k^T S_k^{-1} \\ m_k &= m_k^- + K_k (y_k - F_k m_k^-) \\ P_k &= P_k^- - K_k S_k K_k^T, \end{aligned} \quad (2)$$

where m_k and P_k are the estimate of the mean and the covariance of the process x_k , respectively. We can use a more general nonlinear version of the evolution model and the measurement model. Then the solution can be sought using nonlinear methods, such as the extended Kalman filter (Grewal and Andrews 2001).

Regularization methods are required, because the inverse problem in DOT is ill-posed. Kaipio and Somersalo (1999) have presented an augmented measurement model to include a spatial regularization into the state space model. Kolehmainen *et al* (2003) and Prince *et al* (2003) applied this method to DOT. The augmented measurement model $\tilde{y}_k = \tilde{F}_k x_k + \tilde{r}_k$,

$$\begin{bmatrix} y_k \\ \sqrt{\alpha} L \bar{x} \end{bmatrix} = \begin{bmatrix} F_k \\ \sqrt{\alpha} L \end{bmatrix} x_k + \begin{bmatrix} r_k \\ \hat{r}_k \end{bmatrix}, \quad \hat{r}_k \sim N(0, \hat{R}_k), \quad (3)$$

contains a pseudo measurement $\sqrt{\alpha} L \bar{x}$ of the regularization term $\alpha L x_k$, where α is a regularization parameter, L is a regularization operator and \bar{x} is an *a priori* estimate of x_k . This formulation has a well-known connection to the Tikhonov regularized solution (Kaipio and Somersalo 1999).

The augmented measurement method has a drawback in 3D space. The discretized solution vector m_k tends to have a large number of elements. The discretization used in this paper has 15 711 degrees of freedom. Therefore, solution of the linear system in the estimation step is computationally impractical (the dimension of the matrix S_k is $15\,743 \times 15\,743$, sum of the physical measurement vector y_k dimension (32) and the degrees of freedom (15 711)). We could overcome this issue by decreasing the degrees of freedom, e.g. using wavelet domain reconstruction (Rantala *et al* 2006), or using approximative Kalman filters, such as the variational Kalman filter (Auvinen *et al* 2009b) or the BFGS-Kalman filter (Auvinen *et al* 2009a). Instead of that, we present a method where the smoothness assumptions are included in the evolution process. This method does not use the augmented measurement model; therefore, the matrix S_k is of size 32×32 (the size of the measurement vector). As shown by Kaipio and Somersalo (2004), pp 134–5, the augmented model combined with the random walk time evolution can be equivalently represented as a modified time evolution model. However, our method is based on the known spatially smooth time evolution model rather than the model derived from the augmented model.

In section 2.1 we go through a statistical reconstruction method and compare it to the general Tikhonov reconstruction method when the object is static. After that, in section 2.2.1 we present a state space method which has similar smoothness properties to the statistical reconstruction method. The forward method used in DOT is described briefly in section 2.3. Results from simulated data and real phantom measurement are shown in section 3.

2. Material and methods

2.1. Stationary reconstruction

We assume that the measured data $y \in \mathbb{R}^m$ are corrupted with an additive noise

$$y = f(x) + v, \quad v \sim N(0, V), \quad (4)$$

where the covariance matrix $V \in \mathbb{R}^{m \times m}$ and m is a number of the measurements. The measurement model $f(x)$ in DOT is described in section 2.3. The unknown optical properties $x \in \mathbb{R}^n$ are a discrete vector, such that $x_i = x(\mathbf{r}_i)$, $i \in \{1, \dots, n\}$, where \mathbf{r}_i are discretization grid points. We assume that the optical properties are realizations from the normal distribution

$$x \sim N(m_s, C_s), \quad (5)$$

where the mean and the covariance are $m_{s,i} = m_s(\mathbf{r}_i)$ and

$$C_{s,ij} = C_s(\mathbf{r}_i, \mathbf{r}_j). \quad (6)$$

We selected the covariance function from the Matérn family, e.g. Rasmussen and Williams (2005),

$$C_s(\mathbf{r}, \mathbf{r}') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} \|\mathbf{r} - \mathbf{r}'\|}{\ell} \right)^\nu K_\nu \left(\frac{\sqrt{2\nu} \|\mathbf{r} - \mathbf{r}'\|}{\ell} \right), \quad (7)$$

where Γ is the gamma function and K_ν is the modified Bessel function of the second kind. Parameters σ , ν and ℓ are positive constants. For $\nu > k$ the sample paths $x(\mathbf{r})$ are k times differentiable, if the process is Gaussian (Gneiting and Guttorp 2010). See more analysis of the Matérn family covariance function in appendix A.

The stationary inverse problem can be written in the Bayesian framework, where the solution of the inverse problem is summarized by the posterior density

$$p(x|y) \propto p(y|x)p(x), \quad (8)$$

where the likelihood function $p(y|x) = N(y - f(x), V)$ and the prior density $p(x) = N(m_s, C_s)$. A *maximum a posteriori* (MAP) estimate is equivalent to the Tikhonov regularized solution of the inverse problem

$$x_{\text{MAP}} = \arg \min_{x \in \mathbb{R}^n} \|L_y(y - f(x))\|^2 + \alpha \|L_x(x - m_s)\|^2, \quad (9)$$

where $V^{-1} = L_y^T L_y$ and $C_s^{-1} = \alpha L_x^T L_x$.

In the classical inverse problem theory, smooth solutions are achieved using the smoothness regularizations, where the regularization matrix L_x is a difference operator, such as the discretized Laplace operator $L_x u \approx \Delta u$. It corresponds to an improper prior density $\propto \exp(-0.5\alpha \|L_x(x - m_s)\|^2)$, where L_x is a sparse matrix. This has been a more computationally attractive method, because C_s is a dense matrix.

2.2. State space approach

In this section we build a state space model, a stationary stochastic process, which has similar spatial smoothness properties as the stationary reconstruction has. The presented process has the property that it will decay back to the average value if it has been disturbed and any measurements have not been made.

2.2.1. Mean reverting process. In the following sections we study a finite-dimensional continuous-time space stochastic process $x_i(t) = x(\mathbf{r}_i, t)$, where $\mathbf{r}_i, i \in \{1, \dots, n\}$, are grid points. We could do the same analysis on a spatio-temporal random field $x(\mathbf{r}, t)$, but for simplicity we restrain ourselves in the finite-dimensional model. See the complete analysis of the discretization of the continuous infinite-dimensional stochastic process in Pikkarainen (2006).

We assume that the time evolution process is a mean reverting process, Ornstein–Uhlenbeck process (Uhlenbeck and Ornstein 1930). This process can be written as a stochastic differential equation (SDE). Some properties of the more general SDE are shown in appendix B. If the measurements are made at $t_k, k \in \mathbb{Z}^+$, then the evolution-observation model is

$$\frac{dx(t)}{dt} = \lambda(\mu - x(t)) + w(t), \quad (10)$$

$$y_k = f_k(x(t_k)) + v_k \quad v_k \sim N(0, V_k), \quad (11)$$

where λ is a positive scalar constant and $\mu \in \mathbb{R}^n$ is a constant vector. The state noise $w(t) \in \mathbb{R}^n$ is a *time-white noise* process. The state noise process has the following properties:

$$E(w(t)) = 0 \quad (12)$$

$$E(w(t)w(t')^T) = \delta(t - t')C_w \quad (13)$$

$$E(w(t)x(t')^T) = 0, \quad (14)$$

where δ is Dirac's delta function and $C_w \in \mathbb{R}^{n \times n}$ is a spatial covariance matrix, to be defined later.

Next we show how the spatial covariance matrix is selected. The process (10) is easier to analyse if we do the change of the variable $z(t) = x(t) - \mu$, which gives

$$\frac{dz(t)}{dt} = -\lambda z(t) + w(t). \quad (15)$$

Due to subtraction of the constant the covariance is same and the mean is shifted by the constant. In appendix C we have presented the solution of the mean and the covariance of the process $z(t)$. The covariance matrix is

$$\text{cov}(x(t), x(t')) = e^{-\lambda(t+t')} \left(\text{cov}(x(0), x(0)) + \frac{1}{2\lambda} C_w (e^{2\lambda \min(t,t')} - 1) \right), \quad (16)$$

where $\text{cov}(x(0), x(0))$ is an initial marginal spatial covariance matrix. The process will converge to the stationary process

$$E(x(t)) \longrightarrow \mu, \quad t \rightarrow \infty \quad (17)$$

$$\text{cov}(x(t), x(t')) \longrightarrow \frac{1}{2\lambda} C_w e^{-\lambda|t-t'|}, \quad t, t' \rightarrow \infty. \quad (18)$$

If we select $\text{cov}(x(0), x(0)) = C_w/(2\lambda)$ and $x(0) = \mu$, then the process is stationary $t > 0$. We require that the process will have similar spatial smoothness properties as the stationary reconstruction in section 2.1. Therefore we select

$$C_w = 2\lambda C_s, \quad (19)$$

where C_s is the covariance matrix in equation (6), and the marginal covariance matrix will be

$$\text{cov}(x(t), x(t)) = C_s \quad (20)$$

at every time step t . Even if the process starts from other state it will converge to the stationary state, when $t \rightarrow \infty$.

2.2.2. *Time discretization.* The continuous state equation (10) can be discretized using equation (B.5) from appendix B. The process $x(t)$ at t_{k+1} is

$$x(t_{k+1}) = e^{-\lambda(t_{k+1}-t_k)}x(t_k) + \int_{t_k}^{t_{k+1}} e^{-\lambda(t_{k+1}-\tau)}\lambda\mu d\tau + \int_{t_k}^{t_{k+1}} e^{-\lambda(t_{k+1}-\tau)}w(\tau) d\tau \quad (21)$$

if the state $x(t_k)$ is known. Measurements are made at time interval $[t_k, t_{k+1}]$ during which we assume that the process is constant. The evolution-observation model can be written as

$$x_k = A_{k-1}x_{k-1} + b_{k-1} + q_{k-1} \quad q_{k-1} \sim N(0, Q_{k-1}) \quad (22)$$

$$y_k = f_k(x_k) + v_k \quad v_k \sim N(0, V_k), \quad (23)$$

where

$$A_k = e^{-\lambda(t_{k+1}-t_k)} \quad (24)$$

$$b_k = \int_{t_k}^{t_{k+1}} e^{-\lambda(t_{k+1}-\tau)}\lambda\mu d\tau \quad (25)$$

$$q_k = \int_{t_k}^{t_{k+1}} e^{-\lambda(t_{k+1}-\tau)}w(\tau) d\tau. \quad (26)$$

The covariance matrix $Q_k = E(q_k q_k^T)$ is

$$\begin{aligned} Q_k &= \int_{t_k}^{t_{k+1}} e^{-\lambda(t_{k+1}-\tau)}C_w e^{-\lambda(t_{k+1}-\tau)} d\tau \\ &= \frac{1}{2\lambda}(C_w - e^{-2\lambda\Delta t_{k+1}}C_w), \end{aligned} \quad (27)$$

where $\Delta t_{k+1} = t_{k+1} - t_k$ (cf equation (B.7) in appendix B). A solution of the discrete state space model can be solved using the extended Kalman filter (Grewal and Andrews 2001) or after the linearization of the measurement model using the Kalman filter described in section 1.2.

2.3. Forward model in diffuse optical tomography

The photon density $u(\mathbf{r})$ in the scattering dominant medium can be approximated by the diffusive approximation (DA) of the radiative transfer equation. The DA is a partial differential equation

$$-\nabla \cdot \kappa(\mathbf{r})\nabla u(\mathbf{r}) + \left(\mu_a(\mathbf{r}) + \frac{i\omega}{c}\right)u(\mathbf{r}) = 0, \quad \mathbf{r} \in \Omega, \quad (28)$$

where ω is the modulation frequency, c is the speed of the light and the absorption coefficient $\mu_a(\mathbf{r})$ and the diffusion coefficient $\kappa(\mathbf{r})$ are the optical properties. The boundary condition is

$$u(\mathbf{r}) + 2\xi\kappa(\mathbf{r})\frac{\partial u(\mathbf{r})}{\partial \mathbf{n}} = J^-(\mathbf{r}), \quad \mathbf{r} \in \partial\Omega, \quad (29)$$

where \mathbf{n} is the outward normal vector and $J^-(\mathbf{r})$ is the incoming photon current. The coefficient ξ is due to the refractive index mismatch at the boundary $\partial\Omega$.

Measurable quantity, the outgoing photon current, at the boundary, also called the exitance, is

$$J^+(\mathbf{r}) = -\kappa(\mathbf{r})\frac{\partial u(\mathbf{r})}{\partial \mathbf{n}}, \quad \mathbf{r} \in \partial\Omega. \quad (30)$$

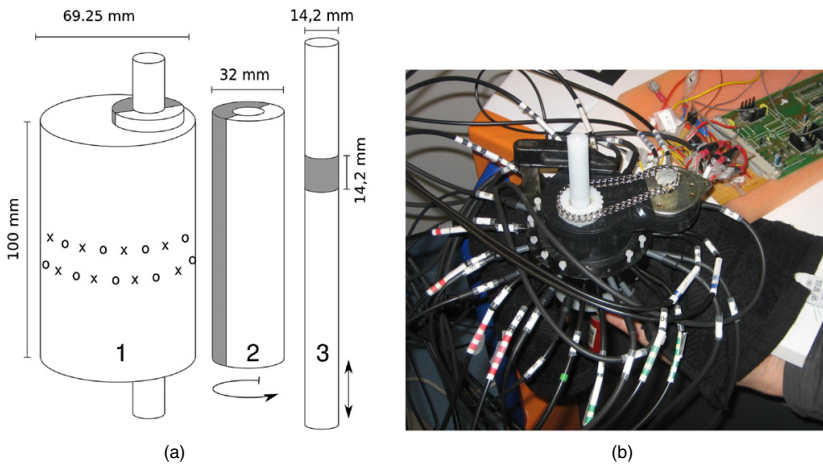


Figure 1. (a) The phantom consists of a cylindrical body and two movable parts. The body (part 1) has homogeneous optical properties. The sources (x) and detectors (o) are attached on the surface of the body. Half of the shell (part 2) has a higher absorption coefficient than the body. A small section of the rod (part 3) has a higher absorption coefficient. Otherwise, movable parts have same optical properties as the body. (b) The phantom measurement in action.

The complex-valued exitance is presented using the logarithm of the moduli and the argument, $y_{jq} = [\ln |J_q^+(\mathbf{r}_j)| \quad \arg J_q^+(\mathbf{r}_j)]$, $j \in \{1, \dots, n_m\}$, when the source q is active at \mathbf{r}_q , $q \in \{1, \dots, n_s\}$.

We assume that the diffusion coefficient is constant and we are only interested in changes in the absorption coefficient. The forward model is $y_t = f_t(x_t)$, where $x_{t,i} = \mu_a(\mathbf{r}_i)$ and the source q is active at t . The model can be linearized at every time step $\delta y_t = J_t(x_t)\delta x_t$, where $J_t(x_t)$ is the Jacobian matrix of f_t at x_t . For computational reasons we linearized the model only once at the background optical properties x_{bg} and used the model $\delta y_t = J_t(x_{bg})\delta x_t$.

The forward model is implemented in the TOAST package and we used its Matlab interface (Schweiger and Arridge 2008). The package uses the finite element method to solve equation (28) and the adjoint method to calculate the Jacobian matrix (Arridge 1999).

2.4. Phantom

The phantom consists of a cylindrical main body (part 1) and two movable parts (a rod and a shell; parts 2 and 3) that go through the phantom. The main body itself is homogeneous but the movable parts have optically differing perturbations. The shell is divided longitudinally into two differing regions and the rod has a small perturbation in the middle. By moving these parts, different physiological phenomena can be simulated. The phantom was made from XOR crystal polyester resin (CREARTEC trend-design-gmbh), which is optically non-scattering and clear. Its scattering and absorption properties were modified with titanium dioxide powder and Pro Jet 900 NP ink (Avecia Biologics Limited), which were mixed in the resin with the help of a little ethanol. To achieve tissue-like optical properties, the concentration of titanium dioxide is 191.5 mg dl^{-1} and the concentration of the ink is 0.14 mg dl^{-1} . The dimensions of the phantom can be seen in figure 1.

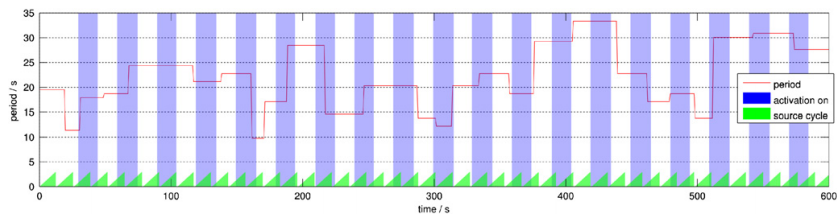


Figure 2. Time course of the translating and the rotating part of the phantom and sources. The red line presents the period of the rotating part. Jumps in the line show the time points where the period changes. Blue bars show when the activation is on, i.e. the perturbation in the translating part is inside the phantom. Green triangles present the cycle of the sources.

The rod was used to simulate hemodynamic responses by moving it up and down inside the phantom. The shell was rotated continuously by a stepper motor and it simulated systemic slow oscillations. The stepper motor was connected to the cylinder with two cogwheels and a chain, and was controlled with a computer. The systemic slow oscillations in human adults range from low frequency (~ 0.1 Hz) to very low frequency (~ 0.02 Hz) (Obrig *et al* 2000). The cycle times were determined according to systemic slow oscillations and they ranged from 9.80 to 34.03 s with the average of 20.49 s. The rod was moved by hand with 15 s intervals so that the perturbation was either completely outside the phantom or in the middle of it. See figure 2.

2.5. Measurement setup

The measurements were performed using the frequency-domain (FD) optical tomography system developed at Aalto University (Nissilä *et al* 2002, 2005). In FD, the optical power of the light source is modulated with a radiofrequency signal (100 MHz) and the phase shift and amplitude of the modulated component of the detected light is measured. This corresponds approximately to the measurement of intensity and mean time using time-domain systems.

Optical fibres were positioned in two rings of 16 positions each. Detector 16 was not used due to a technical problem. The position of the source and detector fibres is shown in figure 1. In the first step, the phantom was measured without movement in the rod or the shell. Calibration according to the three-step procedure described in Nissilä *et al* (2005) and Tarvainen *et al* (2005) was applied to the data and the initial (homogeneous) guess for the optical properties of the phantom and the amplitude and phase coupling coefficient was determined by manually adjusting the global parameters until an adequate match between the simulated and measured slopes of $\arg J^+$ and $\ln |J^+|$ as a function of distance was achieved. The background optical properties were $\mu_{a,bg} = 0.016 \text{ mm}^{-1}$ and $\kappa_{bg} = 0.49 \text{ mm}$.

In the next step, the phantom was measured with the rotating shell controlled by computer and the moving rod (the activation) was manually moved in and out of the phantom with 15-s intervals, so that the inter-stimulus-onset interval was approximately 30 s. The measurement time was 600 ms per source position. Measurement without rotation of the shell was also performed.

The forward problem was solved on an unstructured mesh with 48 127 nodes and 33 813 tetrahedral elements. The optical parameters were discretized on a structured grid with 15 711 nodes.

3. Results

First we compare the presented regularization model and the state space method where the regularization is implemented using the augmented measurement model in the case of simulated data. Then we compare two different reconstruction methods for the measured data: (i) the stationary reconstruction and (ii) the state space reconstruction. In the stationary reconstruction we ignore the fact that the object is under motion. The data are averaged before we solve equation (9). In the state space reconstruction whole time series is solved.

3.1. Comparison of state space models in 2D

First we compare the presented model and the state model where the regularization is implemented using the augmented measurement model in a 2D simulated case. The object was a circle of diameter 50 mm. Optical properties of the background were $\mu_a = 0.025 \text{ mm}^{-1}$ and $\kappa = 0.1646 \text{ mm}$. Around the boundary of the object, 32 sources and 32 detectors were placed evenly. Sources were active once in random order, i.e. data consist of 64 measurements at 32 time instances ($\Delta t = 1$ arbitrary units). The object contains a perturbation, diameter 14 mm, located at 12.5 mm from the centre of the object. Optical properties of the perturbation were $\mu_a = 0.05 \text{ mm}^{-1}$ and $\kappa = 0.1646 \text{ mm}$. The perturbation translates 2.8° around the centre of the object at every time step. Total translation is 90° .

The data were simulated in an unstructured mesh which contains 13 051 nodes and 25 752 triangular elements. 1% multiplicative noise was added to the measurement. The state space models were calculated in a structured grid which contains 786 nodes. The forward model was linearized in an unstructured grid which contains 3511 nodes and 6840 triangular elements.

First the presented state space model was run. Parameters of the steady state covariance $C_s(\mathbf{r}, \mathbf{r}')$ were $\sigma^2 = 0.01$, $\nu = 5/2$ and $\ell = 10$. This choice should prefer twice differentiable solutions like the first-order Tikhonov regularization. We tested several values of the parameter controlling the mean reverting process and selected $\lambda = 0.5$ and $\mu = 0$. Then we run the state space model with augmented measurement model, such that $Q = 10I$, $R = 0.01I$, $\alpha = 10^4$, and the regularization matrix L was the discretized Laplace operator. In both cases, Kalman filter was utilized with the linearized measurement model.

The results are presented in figure 3 and a movie (movie 1) is available in the supplementary material at stacks.iop.org/IP/27/025009/mmedia. As expected, both methods gave very similar results. The root mean square error was ranged from 5.2×10^{-3} to 5.9×10^{-3} in the presented state space model and from 5.1×10^{-3} to 5.8×10^{-3} with the augmented measurement model. In both cases, the reconstructions followed the true target without significant lag, but the contrast was low.

3.2. Stationary reconstruction

First we tested the stationary reconstruction. We calculated the Tikhonov regularized solution using equation (9), where L_x was the discretized Laplace operator. The regularization parameter was 0.05. Linear drift was first removed from the data. Then the difference data between two states, averaged data when the perturbation was inside and averaged data when the perturbation was outside were calculated. This calculation includes the assumption that the object does not change during the measurement. This is not actually true since the rotating part and the translating part are in different phase when a source is next time active. This adds the noise to the measurement. Figure 4 presents the reconstructed absorption coefficient

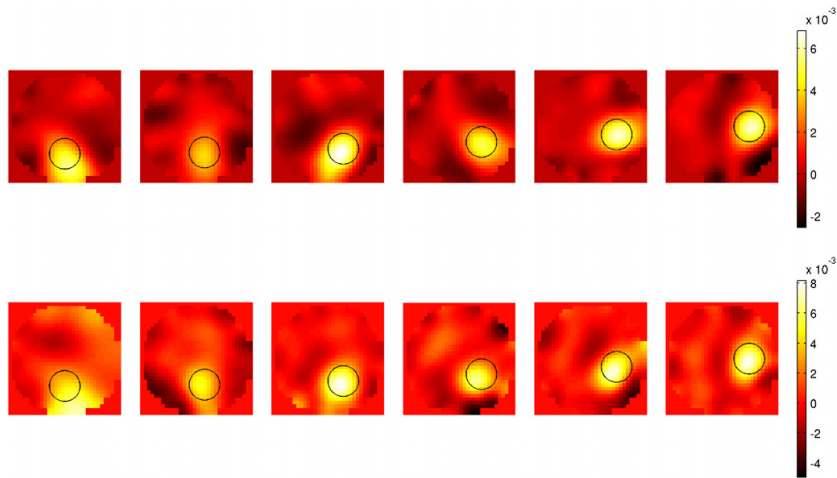


Figure 3. Comparison of the presented regularization method (top row) and the state space method where regularization is implemented using the augmented measurement model (bottom row) in the 2D case with the simulated data. Each row presents reconstructions at time instants $t = \{1, 7, 13, 20, 26, 32\}$ (arbitrary units). The black circle presents true location of the perturbation ($\delta\mu_a = 0.025 \text{ mm}^{-1}$).

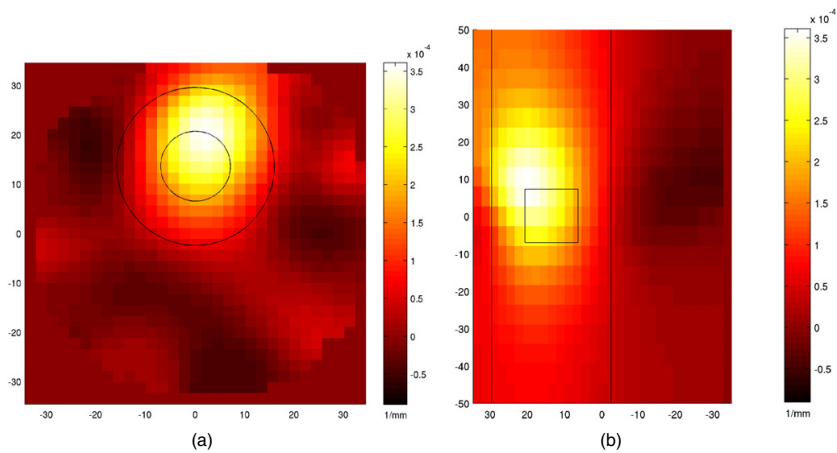


Figure 4. The reconstruction of the absorption coefficient solved using the first-order Tikhonov regularization from the averaged difference data. Images present horizontal (a) and vertical (b) planes through location where the absorption has maximum value. Approximative locations of the rotating part and the translating part of the phantom are drawn by black lines.

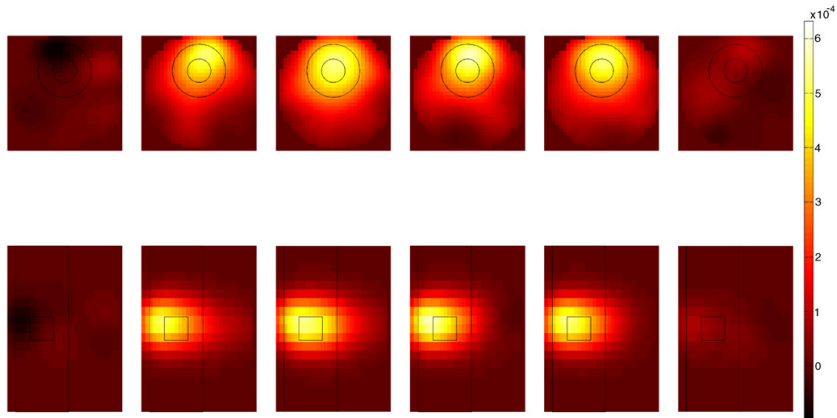


Figure 5. Columns show the state space reconstruction of the absorption coefficient at time instants $t = \{-1.6, 2.4, 6.5, 10.6, 14.6, 18.7\}$ s after the beginning of the activation. The upper row is the horizontal plane through location where the absorption has maximum value at time instant 6.5 s. The lower row is the vertical plane respectively. The movie (movie 2) is available in the supplementary material at stacks.iop.org/IP/27/025009/mmedia. Approximative locations of the rotating part and the translating part of the phantom are drawn by black lines.

from vertical and horizontal planes which go through the location where the absorption gets its maximum value. The results do not show any significant artefacts but the maximum of the absorption coefficient is higher than expected.

3.3. State space reconstruction

Next we tested the dynamic reconstruction method developed in this paper. The parameters of the steady state covariance $C_s(\mathbf{r}, \mathbf{r}')$ were $\sigma^2 = 0.05$, $\nu = 5/2$, and $\ell = 5$. This choice should prefer twice differentiable solutions like the first-order Tikhonov regularization. We tested several values of the parameter controlling the mean reverting process and selected $\lambda = 0.61$. If a small value is used, the successive states have a large correlation and the rise time of the activation is too long.

The data used were difference data. First the linear drift was removed. Then the average, when the perturbation was outside the main body of the phantom, was subtracted. We are interested in changes in the optical properties as a function of time. Therefore, we set $\mu = 0$ and the initial state was $x(0) = 0$.

First the Kalman-filtered solution was calculated from the linearized model. Then the time-locked average of the time series was calculated relative to the beginning of the activation. Figure 5 shows the time series of the averaged over intervals between insertions of the perturbation in the translating part. The movie (movie 2) is available in the supplementary material at stacks.iop.org/IP/27/025009/mmedia. Figure 6(a) presents activation at a voxel where the absorption coefficient had the highest value at $t = 6.5$ s. For comparison the corresponding time course for the measurement where the shell was not rotated is given in figure 6(b).

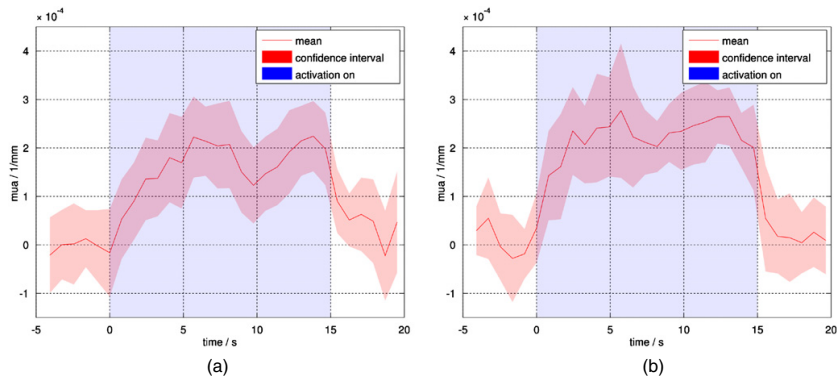


Figure 6. (a) The red line shows the average of the absorption coefficient at a voxel where it got its highest value at time instant 6.5 s after the beginning of the activation. The shaded red area shows the $\pm 2\sigma$ confidence interval. The shaded blue area shows time when the perturbation in the translating part is inside of the main body. (b) Is the same but the rotating part of the phantom was turned off.

4. Discussion

In this paper we presented a regularization method for the nonstationary inverse problem. Our evolution model was the Ornstein–Uhlenbeck process with the spatially smooth noise process. One could use another spatially smooth process, such as the random walk with appropriately chosen covariance matrix. A difference between the Ornstein–Uhlenbeck process and the random walk is that the first one has a stationary state and the latter has not.

The problem was studied in 2D DOT by Kolehmainen *et al* (2003). They used the augmented measurement model, which was introduced in Kaipio and Somersalo (1999). Our method, which is based on a spatially smooth time evolution model, performs equally well as the augmented measurement model in the 2D case, with similar smoothness properties and regularization effect. On the other hand, the suggested Matérn family of the covariance functions gives us much more liberty to control the smoothness and the spatial and the temporal correlation length of the process.

The presented method outperforms the augmented measurement model in cases where the spatial discretization is large, such as in 3D problems. We studied the case of the reconstruction of the time series of 3D images from data measured from a dynamic optical tomography phantom. The calculation of the Kalman gain matrix in the augmented measurement model method took about 9 min, leading to over 90 h of computation for the complete time series using the Kalman filter (not performed). One step of the prediction and the estimation in the Kalman filter using the presented method took only about 3 s, and the total reconstruction time was about 45 min, including the linearization of the forward model and the calculation of the covariance matrix. We estimated that the extended Kalman filter, which solves the forward problem at every step, would take about 24 h. Therefore we used the linearized model instead.

The stationary reconstruction gave reasonable results even though the rotating part of the phantom interferes with the measurement. The state space method gave more information than the stationary reconstruction, allowing the reconstruction of the raw and averaged time series of the localized absorption change in the phantom. Using the presented method, it is

easier to apply slower, but highly sensitive, instruments to the study of dynamic processes in tissue. Activating the rotating shell which simulates physiological confounds in tissue reduced the reconstructed contrast and led to a more noisy time series. To minimize this loss in image quality, auxiliary signals which are coupled to the background physiology should be included in the time evolution model. This is a subject of further study.

Acknowledgments

The research was supported by the Emil Aaltonen foundation and the Academy of Finland (project 120946).

Appendix A. Matérn family of the covariance functions

We use the Matérn class covariance function 7 as it is defined by Rasmussen and Williams (2005). The parameter ν controls the smoothness properties of the process drawn from the Gaussian process with Matérn covariance. If $\nu > k$, the process will be k times mean square differentiable (Gneiting and Guttorp 2010). The covariance can be expressed as a product of a polynomial of order n and an exponential function

$$C_{\nu=\frac{3}{2}}(\mathbf{r}, \mathbf{r}') = \left(1 + \frac{\sqrt{3}r}{\ell}\right) \exp\left(-\frac{\sqrt{3}r}{\ell}\right) \quad (\text{A.1})$$

$$C_{\nu=\frac{5}{2}}(\mathbf{r}, \mathbf{r}') = \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right) \quad (\text{A.2})$$

...

where $r = \|\mathbf{r} - \mathbf{r}'\|$, if $\nu = n + 1/2$. See the complete function in Rasmussen and Williams (2005). The covariance function has two special cases: the exponential covariance function

$$C_{\nu=\frac{1}{2}}(\mathbf{r}, \mathbf{r}') = \sigma^2 \exp\left(-\frac{r}{\ell}\right), \quad (\text{A.3})$$

when $\nu = 1/2$, and the squared exponential covariance function

$$C_{\nu \rightarrow \infty}(\mathbf{r}, \mathbf{r}') = \sigma^2 \exp\left(-\frac{r^2}{2\ell^2}\right), \quad (\text{A.4})$$

when $\nu \rightarrow \infty$ (Rasmussen and Williams 2005). Therefore the parameter ℓ can be defined as the length scale of the process. The parameter σ^2 is simply the variance.

Appendix B. Stochastic differential equations

In this paper we study SDEs of the form

$$\frac{dx(t)}{dt} = Ax(t) + b + w(t), \quad (\text{B.1})$$

where $A \in \mathbb{R}^{n \times n}$ is a constant matrix. The state noise process $w(t)$ has the following properties:

$$E(w(t)) = 0 \quad (\text{B.2})$$

$$E(w(t)w(t')^T) = C_w \delta(t - t') \quad (\text{B.3})$$

$$E(w(t)x(t')^T) = 0, \quad (\text{B.4})$$

where $C_w(t, t') \in \mathbb{R}^{n \times n}$ is a noise covariance matrix.

The solution of the SDE is similar to the solution of the deterministic differential equation (Bar-Shalom *et al* 2001)

$$x(t) = e^{(t-t_0)A} x(t_0) + \int_{t_0}^t e^{(t-\tau)A} (b + w(\tau)) d\tau. \quad (\text{B.5})$$

The expectation value $m(t) = E(x(t)) = e^{(t-t_0)A} E(x(t_0)) + \int_{t_0}^t e^{(t-\tau)A} b d\tau$ can be computed using property (B.2). Similarly one can solve the covariance matrix

$$C(t, t') = E((x(t) - m(t))(x(t') - m(t'))^T) \quad (\text{B.6})$$

$$= e^{(t-t_0)A} C(t_0, t_0) e^{(t'-t_0)A^T} + \int_{t_0}^{\min(t, t')} e^{(t-\tau)A} C_w e^{(t'-\tau)A^T} d\tau, \quad (\text{B.7})$$

where $C(t_0, t_0)$ is the initial marginal covariance matrix (Bar-Shalom *et al* 2001).

Appendix C. Mean reverting process

The mean reverting process (10) can be transformed into

$$\frac{dx(t)}{dt} = -\lambda x(t) + w(t) \quad (\text{C.1})$$

by using the change of the variable. The state noise process $w(t)$ has same properties as in section 2.2.1 and λ is a positive constant. Next we study the process multiplied by $e^{\lambda t}$

$$\frac{d(x(t) e^{\lambda t})}{dt} = e^{\lambda t} \left(\frac{dx(t)}{dt} + \lambda x(t) \right) = w(t) e^{\lambda t}. \quad (\text{C.2})$$

The solution of the above process is (see appendix B)

$$x(t) e^{\lambda t} = x(0) + \int_0^t w(\tau) e^{\lambda \tau} d\tau. \quad (\text{C.3})$$

The expectation value of the process (C.2) is

$$E(x(t) e^{\lambda t}) = E(x(0)) + \int_0^t E(w(\tau)) e^{\lambda \tau} d\tau \quad (\text{C.4})$$

$$= E(x(0)), \quad (\text{C.5})$$

from which we can solve the expectation value of the mean reverting process

$$E(x(t)) = e^{-\lambda t} E(x(0)). \quad (\text{C.6})$$

The covariance function of the process (C.2) is

$$E((x(t) e^{\lambda t} - E(x(t) e^{\lambda t}))(x(t') e^{\lambda t'} - E(x(t') e^{\lambda t'}))^T) \quad (\text{C.7})$$

$$= E((x(0) - E(x(0)))(x(0) - E(x(0)))^T) + \int_0^t \int_0^{t'} e^{\lambda \tau} e^{\lambda \nu} E(w(\tau) w(\nu)^T) d\nu d\tau \quad (\text{C.8})$$

$$= \text{cov}(x(0), x(0)) + \int_0^{\min(t, t')} e^{2\lambda \tau} C_w d\tau \quad (\text{C.9})$$

$$= \text{cov}(x(0), x(0)) + \frac{1}{2\lambda} C_w (e^{2\lambda \min(t, t')} - 1) \quad (\text{C.10})$$

from which we can solve the covariance matrix of the mean reverting process

$$\text{cov}(x(t), x(t')) = e^{-\lambda(t+t')} \left(\text{cov}(x(0), x(0)) + \frac{1}{2\lambda} C_w (e^{2\lambda \min(t,t')} - 1) \right). \quad (\text{C.11})$$

If $\text{cov}(x(0), x(0)) = C_w/(2\lambda)$, then the covariance function is

$$\text{cov}(x(t), x(t')) = \frac{1}{2\lambda} C_w e^{-\lambda|t-t'|}, \quad (\text{C.12})$$

which is the stationary covariance matrix of the process (C.2).

References

- Arridge S R 1999 Optical tomography in medical imaging *Inverse Problems* **15** R41–93
- Auvinen H, Bardsley J M, Haario H and Kauranne T 2009a Large-scale Kalman filtering using the limited memory BFGS method *Electron. Trans. Numer. Anal.* **35** 217–33
- Auvinen H, Bardsley J M, Haario H and Kauranne T 2009b The variational Kalman filter and an efficient implementation using limited memory BFGS *Int. J. Numer. Methods Fluids* **64** 314–35
- Bar-Shalom Y, Li X, Li X and Kirubarajan T 2001 *Estimation with Applications to Tracking and Navigation* (New York: Wiley Interscience)
- Diamond S G, Huppert T J, Kolehmainen V, Franceschini M A, Kaipio J P, Arridge S R and Boas D A 2006 Dynamic physiological modeling for functional diffuse optical tomography *Neuroimage* **30** 88–101
- Gneiting T and Guttorp P 2010 Continuous parameter stochastic process theory *Handbook of Spatial Statistics* ed A E Gelfand, P J Diggle, M Fuentes and P Guttorp (Boca Raton: Chapman and Hall/CRC) chapter 2
- Grewal M and Andrews A 2001 *Kalman Filtering: Theory and Practice Using MATLAB* (New York: Wiley)
- Kaipio J and Somersalo E 1999 Nonstationary inverse problems and state estimation *J. Inverse Ill-Posed Problems* **7** 273–82
- Kaipio J and Somersalo E 2004 *Statistical and Computational Inverse Problems* (Berlin: Springer)
- 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
- Nissilä I, Kotilahti K, Fallström K and Katila T 2002 Instrumentation for the accurate measurement of phase and amplitude in optical tomography *Rev. Sci. Instrum.* **73** 3306
- Nissilä I, Noponen T, Kotilahti K, Katila T, Lipiäinen L, Tarvainen T, Schweiger M and Arridge S 2005 Instrumentation and calibration methods for the multichannel measurement of phase and amplitude in optical tomography *Rev. Sci. Instrum.* **76** 044302
- Obrig H, Neufang M, Wenzel R, Kohl M, Steinbrink J, Einhäupl K and Villringer A 2000 Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults *Neuroimage* **12** 623–39
- Pikkarainen H K 2006 State estimation approach to nonstationary inverse problems: discretization error and filtering problem *Inverse Problems* **22** 365
- Prince S, Kolehmainen V, Kaipio J P, Franceschini M A, Boas D and Arridge S R 2003 Time-series estimation of biological factors in optical diffusion tomography *Phys. Med. Biol.* **48** 1491–504
- Rantala M, Vänskä S, Järvenpää S, Kalke M, Lassas M, Moberg J and Siltanen S 2006 Wavelet-based reconstruction for limited-angle X-ray tomography *IEEE Trans. Med. Imaging* **25** 210–7
- Rasmussen C E and Williams C K I 2005 *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)* (Cambridge, MA: MIT Press)
- Schweiger M and Arridge S R 2008 Toast reconstruction package <http://web4.cs.ucl.ac.uk/research/vis/toast/>
- Tarvainen T, Kolehmainen V, Vauhkonen M, Vanne A, Gibson A P, Schweiger M, Arridge S R and Kaipio J P 2005 Computational calibration method for optical tomography *Appl. Opt.* **44** 1879–88
- Uhlenbeck G E and Ornstein L S 1930 On the theory of the Brownian motion *Phys. Rev.* **36** 823–41