Signal source separation and localization in the analysis of dynamic near–infrared optical tomographic time series

Harry L. Graber[*a], Yaling Pei[b], Randall L. Barbour[a,b],
David K. Johnston[c], Ying Zheng[c], John E. Mayhew[c]
[a]Department of Pathology, SUNY Downstate Medical Center,
450 Clarkson Avenue, Brooklyn, NY, USA 11203
[b]NIRx Medical Technologies LLC, 15 Cherry Lane, Glen Head, NY, USA 11545
[c]Department of Psychology, University of Sheffield, Western Bank, Sheffield, UK S10 2TP

ABSTRACT

The emerging sub-field of *dynamic* medical optical tomography shows great potential for conferring significantly enhanced early diagnosis and treatment monitoring capabilities upon researchers and clinicians. In previous reports we have showed that adoption of elementary time-series analysis techniques can bring about large improvements in localization and contrast in optical tomographic images. Here we build upon the earlier work, and show that well-known techniques for extraction and localization of signals embedded in a noisy background, and for deconvolution of signal mixtures, also can be successfully applied to the problem of interpreting dynamic optical tomography data sets. A general linear model computation is used for the signal extraction/localization problem, while the deconvolution problem is addressed by means of a blind source separation technique extensively reported.

Keywords: Medical and biological imaging, pattern recognition and feature extraction, tomographic image processing, image analysis

1. INTRODUCTION

Practical realization of diffuse optical tomography (DOT) using visible or near-infrared tomographic (NIR) light – currently being developed as a diagnostic modality for noninvasive mapping of hemoglobin concentration and oxygen saturation in tissue – requires solution of several problems that do not arise for superficially similar modalities such as x-ray CT imaging. Chief among these and the first to be recognized [1,2] is that the net signal recorded in any source-detector channel is a function of some weighted average of optical coefficients within a large volume of tissue that lacks well-defined boundaries. This difficulty was surmounted by means of the image reconstruction techniques (many varieties derived from a few basic strategies) that have been developed and implemented during the past ~15 years [3-9]. More recently it was recognized that the presence of prominent $\lesssim 1$ Hz modulations in data from NIR spectroscopic (NIRS) tissue function measurements [10] implies that the optical coefficients of tissue fluctuate appreciably on short (seconds to minutes) time scales. In that case it is important that the full set of data used to reconstruct each image be taken in a time period sufficiently short that the properties of the target tissue may be regarded as static [11]. Performance of rapid measurements affords one the opportunity to acquire data for a large series of images in a reasonably short time period. The resulting images then can be post-processed with the same sorts of time-series analysis techniques as are used in the interpretation of NIRS [12] or topographic image series [13]; the ability of such post-processing to yield temporal feature maps that have contrast and/or spatial resolution superior to those of the individual images in the series has been demonstrated [11,14-16]. More important, with regard to clinical applications of DOT, is that said temporal features may relate more directly to the health and functional state of the tissue than the optical coefficients do [11,14].

Among the clinically valuable abilities gained through performance of a rapid time-series measurement is that of separately specifying the contribution of each component of the vasculature (arteries, microvessels, veins) to the overall temporal fluctuation seen in a given image pixel or source-detector channel. That is, the use of temporal information

[*] harry.graber@downstate.edu; phone (718) 270-1286; fax (718) 270-3313; downstate.edu

allows one to distinguish among structures whose physical size and proximity may be smaller than the spatial resolution of a DOT image. This ability stems from the fact that in the absence of exogenous lumiphores and contrast agents, hemoglobin, which ordinarily is present only within the vascular compartment of tissue, is the principal DOT contrast agent. As it is well established that blood vessel diameters and/or flow rates vary with characteristic periods [17], it can be expected that the hemoglobin concentration and oxygen saturation also will exhibit fluctuations at these frequencies. Finally, the blood vessel types are associated with fluctuations in distinct frequency bands that have little or no overlap [18].

There is, however, an additional "dimension" for which the problem of resolution must be addressed if optical tomography is to achieve the degree of clinical utility that is held out for it. Namely, any spatiotemporal feature or pattern that is revealed by time-series analysis, using either linear [11,14,19] or nonlinear [14,20] approaches, is likely to be a mixture of multiple underlying physiological rhythms that are not directly observable [21]. The problem of separating, or "unmixing," detected or reconstructed patterns of fluctuation into their underlying physiological components has previously been the subject of significant research efforts in a number of biological and medical fields [22-26] (and in many other areas of science as well). These efforts have resulted in the development of a large corpus of useful signal separation methods. To our knowledge, however, until now there has been no demonstration of their applicability in DOT. The current report supplies clear evidence that the strategies initially developed in other areas can be successfully adopted to analysis of DOT data, when the tomographic measurement is performed in a dynamic (i.e., time-series) mode; that their use can improve the contrast and spatial resolution of DOT data, while also in some cases dramatically reducing the overall computational effort that is required; and that it is reasonable to expect that their application can assist in the interpretation of DOT data, improving its clinical usefulness.

For testing and demonstration purposes, the studies described below were conducted on simulations of DOT measurements. By this means, the experimenters had complete control over the number, size and locations of inclusions (which represent blood-vessel cross sections or small pathologies), the form and amplitude of dynamic behavior assigned to the optical coefficients of each one, and the data noise-to-signal ratio. Two analytic strategies, the mechanics of which are subsequently described, have been successfully applied to the detector and image data. The first of these, an adaptation of an algorithm first described by Molgedey and Schuster (MSA) [27], is an example of a "data driven" analytic strategy. It seeks to identify the underlying patterns of behavior of which the detected or reconstructed time series are mixtures, in the absence of *a priori* knowledge of them. As such, it is a type of what have come to be called "blind source separation" (BSS) algorithms. (It is perhaps unfortunate that the term BSS has become established in the literature, because no such approach is ever absolutely "blind," or knowledge-free [28].) The second, complementary, technique employed is an elementary form of General Linear Model (GLM) computation, which is used in the present study to identify those locations in the medium where the various modeled dynamic functions are found, under the assumption that *a priori* knowledge of these is available. That is, GLM is an example of a "model based" analytic approach. (It should be noted that the data-driven/model-based distinction is not absolute. Intermediate forms of both MSA and GLM, in which "weak models" are used to guide the solution to which the methods converge, have been described by some of the authors of the current report [29].)

2. METHODS

2.1 Structure and Properties of Model Medium

The target medium explored, shown in Figure 1, is a geometrically simple 2–D structure consisting of an 8–cm circle with eight embedded 0.6-cm diameter inclusions. The absorption (μ_a) and reduced scattering (μ_s') coefficients of the light gray background in the Figure were static, with numerical values of $\mu_a = 0.06$ cm^{-1} and $\mu_s' = 10$ cm^{-1}. Four time-varying functions — $\mu_a(t)_Q$, $\mu_a(t)_{C1}$, $\mu_a(t)_{C2}$, and $\mu_a(t)_S$ — were assigned to the absorption coefficient of pairs of inclusions. That is, four different non-periodic fluctuations were present in the medium simultaneously, with each present in two of the eight inclusions.

The particular functions chosen were easily generated representations of the types of dynamics that are known to occur in tissue vascular structures – quasiperiodic ($Q(t)$), chaotic ($C1(t)$, $C2(t)$), and stochastic ($S(t)$) fluctuations [30]. $Q(t)$ was generated by adding two sinusoidal functions with incommensurate frequencies. $C1(t)$ and $C2(t)$ were generated by assigning randomly generated initial values to the Hénon equation, $x_n = 1.0 - 1.4x^2_{n-1} + 0.3x_{n-2}$. While the same equation was used for both cases, by choosing different initial values the two chaotic time series produced are uncorrelated. $S(t)$ was generated by drawing independent samples from a random variable uniformly distributed between -1 and +1. A total of 1000 values were computed for each function. The residual correlations between the

finite-length model function pairs are: $\rho(Q,C1) = -0.030$, $\rho(Q,C2) = 0.043$, $\rho(Q,S) = 0.016$, $\rho(C1,C2) = 0.101$, $\rho(C1,S) = -0.012$, $\rho(C2,S) = -0.063$. The range of each time-varying μ_a function amounted to a 20% fluctuation about the mean value of 0.06 cm^{-1}. Thus the absorptions assigned to the inclusions were computed from the formulas

$$\mu_a(t)_Q = 0.06\left[1 + 0.2Q(t)\right], \qquad \mu_a(t)_{C1} = 0.06\left[1 + 0.2C1(t)\right],$$
$$\mu_a(t)_{C2} = 0.06\left[1 + 0.2C2(t)\right], \qquad \mu_a(t)_S = 0.06\left[1 + 0.2S(t)\right], \tag{1}$$

where $Q(t)$ is a quasiperiodic time series, $C1(t)$ and $C2(t)$ are chaotic time series and $S(t)$ is a stochastic time series, and the range of each of these four functions includes all values from -1 to +1 (see Ref. [31] for function plots, and Figures 6 and 8 in Results, below).

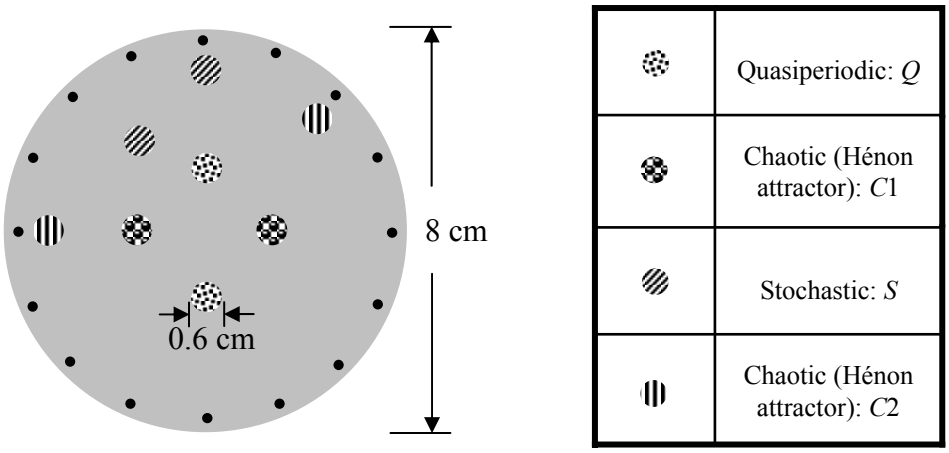


Fig. 1. Indicated dynamics were imposed on the inclusions' μ_a, which ranged from 0.048 cm^{-1} to 0.072 cm^{-1} over time. The remainder of the target had a constant μ_a of 0.06 cm^{-1}, and the entire target had constant μ_s = 10 cm^{-1}. Black dots denote source/detector locations.

A significant feature of the functions $C1(t)$ and $C2(t)$ is that while their asymptotic (i.e., in the limit as their durations increase without bound) cross-correlation function is identically zero, their statistical distributions are exactly the same. The latter fact is shown graphically in Figure 2, where eleven-bin histograms of the 1000-time-point model

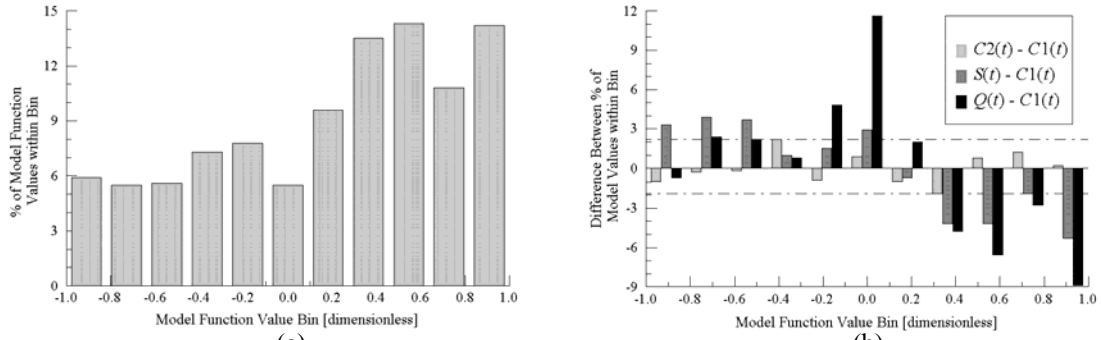


Fig. 2. (a) Value-frequency histogram of the temporal model function $C1(t)$; (b) differences between histograms of $C1(t)$ and of $C2(t)$ (light gray bars), $S(t)$ (dark gray), and $Q(t)$ (black). Histograms were generated by dividing the full (–1 to +1) range of possible values into eleven equal-sized bins and expressing the number of computed function values within each bin as a percentage of the total. In (b), all 11 light gray bars lie within the band defined by the two horizontal dot-dash lines, while 7 of 11 dark gray and 7 of 11 black bars fall outside it.

function $C1(t)$ and of the frequency differences $C2(t) - C1(t)$, $S(t) - C1(t)$, and $Q(t) - C1(t)$ are plotted. Taking the $C1(t)$ and $C2(t)$ frequencies as the "expected" and "observed" scores, respectively, results in a value of 1.713 for the χ^2

statistic for the comparison between $C1(t)$ and $C2(t)$. (Reversing the roles of $C1(t)$ and $C2(t)$ gives a slightly *lower* χ^2 value.) On the other hand, taking the $S(t)$ and $Q(t)$ frequencies as the "observed" scores results in $\chi^2 = 20.5$ and $\chi^2 = 48.5$, respectively. That is, the probability that $C2(t)$ has the same distribution as $C1(t)$ is >0.99, while the corresponding probabilities for $S(t)$ and $Q(t)$ [i.e., that they were drawn from the same distribution as was $C1(t)$] are ~0.025 and <0.001, respectively. The significance of the statistical close similarity between $C1(t)$ and $C2(t)$ is that this represents the very case in which it is most difficult for many of the BSS algorithms, such as independent component analysis (ICA) [28], that are commonly employed in attempts to separate distinct dynamic behaviors. The nearly complete success, as shown subsequently, of the PCA-MSA approach at distinguishing them is a testament to that method's power.

2.2 Solution of Forward Problem

Tomographic data for the simulated tissue models were acquired by using the finite element method to solve the diffusion equation with Dirichlet boundary conditions for a DC source [31-33]. For a spatial domain Λ with boundary $\partial\Lambda$, this is represented by the expression

$$\nabla\cdot\left[D(\mathbf{r})\nabla\phi(\mathbf{r})\right] - \mu_a(\mathbf{r})\phi(\mathbf{r}) = -\delta(\mathbf{r}-\mathbf{r}_s), \qquad \mathbf{r}\in\Lambda, \tag{2}$$

where $\phi(\mathbf{r})$ is the photon intensity at position $\mathbf{r}$, $\mathbf{r}_s$ is the position of a DC point source, and $D(\mathbf{r})$ and $\mu_a(\mathbf{r})$ are the position-dependent diffusion and absorption coefficients, respectively. Here the definition used for the diffusion coefficient was $D(\mathbf{r}) = 1/\{3[\mu_a(\mathbf{r}) + \mu_s'(\mathbf{r})]\}$, where $\mu_s'(\mathbf{r})$ is the position-dependent reduced scattering coefficient. Forward-problem solutions were computed for each of sixteen sources positioned about the target at 22.5° intervals (see Fig. 1), with each source located at a depth of 2 mm in from the extended boundary (*i.e.*, within the strip lying between the physical and extended boundaries) [31-33]. Intensity values at the same locations were used as detector readings. Sets of detector readings were computed for each value of the time-varying optical coefficients in the target's inclusions. These calculations were carried out on a fine, 1,488-element, 789-node, finite element mesh [34].

Imaging operators were computed, in the manner described in Ref. [35], for each of the 256 source/detector channels. A single set of imaging operators, computed for an 8-cm-diameter homogeneous medium whose properties are equal to the temporal mean values of the target, was used for all inverse problem computations. The same algorithm as described in the preceding paragraph was used for the operator computations, but here the medium was discretized into a 950-element, 403-node mesh — the same one used for computing solutions to the inverse problem (see Subsection 2.4) — that is considerably coarser than that used for the detector readings computations.

2.3 Detector Noise Model

The robustness of our signal-separation and -localization algorithms to additive detector noise was studied by substituting $\hat{\phi}(s,d;t) = \phi(s,d;t) + n(s,d;t)$, where at each time point and for each source-detector channel $n(s,d;t)$ is independently sampled from a Gaussian distribution with mean zero and a channel-dependent standard deviation, for $\phi(s,d;t)$ prior to all subsequent image reconstruction, PCA-MSA, or GLM computations. The manner in which the standard deviation of the noise varies with the central angle between source and detector is shown in Figure 3.

The premise that underlay this measurement noise model was that the noise-to signal ratio is primarily a function of source-detector distance (i.e., angle). Accordingly, the noise-to-signal ratio, as a function of the central angle θ, was obtained from the formula $N/S = K_0 + (K_{180} - K_0)(\theta/180^\circ)^\beta$, where K_0 is the noise-to-signal ratio at $\theta = 0^\circ$ (source and detector co-located; backscattering), K_{180} is the noise-to-signal ratio at $\theta = 180^\circ$ (source and detector on opposite sides of the medium; transmission), and β is the exponent that defines the functional form. For the results presented below, the values of K_0 and β were always 0.01 and 4, respectively, while K_{180} was varied over a range from 0.03 to 0.5. (In Fig. 3, the boxed numbers associated with the curves are 100 times K_{180}.)

2.4 Solution of Inverse Problem

The reconstruction algorithm that was used to generate the results presented below seeks to solve a modified perturbation equation whose form is

$$\mathbf{W}_r \cdot \delta\mathbf{x} = \delta\mathbf{I}_r, \tag{3}$$

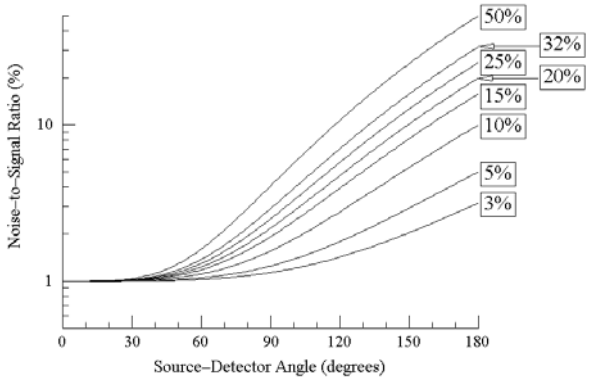


Fig. 3. Variation in noise-to-signal (*N*/*S*) ratio with central angle between source and detector locations, for cases in which (independent, Gaussian) noise was added to all detector readings prior to subsequent analyses. Modeled *N/S* ratio increases with increasing angle (distance), in agreement with usual experimental or clinical experience.

where $\delta\mathbf{x}$ is the vector of differences between the optical properties (*e.g.*, absorption and scattering (diffusion) coefficients) of a target (measured) and a "background" medium, $\mathbf{W}_r$ is the weight matrix describing the influence that each voxel or element has on the surface detectors for the selected reference medium, and $\delta\mathbf{I}_r$ represents a normalized difference between detector readings obtained from the target in two distinct states. The difference between Eq. (3) and a standard linear perturbation equation lies in the structure of the right–hand side. Here we used the previously described Normalized Difference Method (NDM) [32], in which the right–hand side of Eq. (3) is defined by

$$\left(\delta\mathbf{I}_r\right)_i = \frac{\left(\mathbf{I}-\mathbf{I}_0\right)_i}{\left(\mathbf{I}_0\right)_i}\left(\mathbf{I}_r\right)_i . \tag{4}$$

In Eq. (4), $\mathbf{I}_r$ is the computed detector readings corresponding to a selected reference medium, and $\mathbf{I}$ and $\mathbf{I}_0$ represent, in the examples considered here, the intensity at a specific time point and the time-averaged mean, respectively.

A truncated singular value decomposition (TSVD) algorithm [33] was used to compute numerical solutions to Eq. (3). The reconstructions were carried out on a finite-element mesh containing 950 elements and 403 nodes (for a total of 806 unknowns); the same mesh was used at all time points. The truncation parameter (i.e., number of singular vectors retained) was at most 110, while values of 100, 90 and 80 also were used. The principal reason for adjusting the truncation parameter was to determine whether going to lower values could have a beneficial effect on image quality when the detector noise-to-signal ratio is high.

Computations were performed on a Dell Inspiron 7500 laptop computer (Red Hat Linux v. 7.2 O/S). A time period of less than one minute was required for computation of the weight matrix and the subsequent computation of its singular values and singular vectors; from the total elapsed time between successive image file writes, it can be deduced that after the preceding one-time-only steps were completed, no more than 100 milliseconds was required for each image reconstruction.

2.5 General Linear Model

It is frequently the case that the time courses of parameter fluctuations are known (in the present case because a defined medium is the test object; in experimental practice from simultaneous vital-sign measurements, or from simultaneous optical tomographic measurements performed at a different anatomical site) or suspected (e.g., the goal of the analysis is to determine the extent to which detected time series track the designed time courses of imposed provocations). These ancillary functions are referred to as *models*, and the degree to which each is present in the time series at each image pixel can be quantified. The mathematical framework within which said quantification is carried out is the general linear model (GLM) [36]. Let K be the number of model functions considered, L be the number of recovered absorption coefficient time series (i.e., number of detector channels or of image pixels), and N the number of time points. Then according to the GLM, the model functions are related to the recovered time series via $\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where the $N\times L$ matrix $\boldsymbol{\mu}$ contains the L recovered N-point time series of μ_a values, $\mathbf{A}$ is the $N\times K$ matrix of model functions, $\boldsymbol{\beta}$ is a $K\times L$

matrix of fitting parameters, and $\boldsymbol{\varepsilon}$ is a $N\times L$ matrix of residual errors. Solution of the model involves finding that $\boldsymbol{\beta}$ which minimizes $\boldsymbol{\varepsilon}$.

Any degree and type of interaction among models, and any variety of uni- or multivariate analysis of variance or covariance, can be incorporated into the GLM framework [36,37]. The specific GLM code used to generate the results presented here, however, performed only a straightforward multiple linear regression computation, finding the numerical values of the coefficients affording the best fit of the spatiotemporally varying reconstructed $\mu_a(x,y,t)$ to the previously described model functions: $\mu_a(x,y,t) = \beta_0(x,y) + \beta_1(x,y)\mu_a(t)_Q + \beta_2(x,y)\mu_a(t)_{C1} + \beta_3(x,y)\mu_a(t)_{C2} + \beta_4(x,y)\mu_a(t)_S$.

At each position and for each coefficient, the *t*-statistic appropriate for testing the hypothesis that the observed coefficient value could be obtained purely by chance is estimated. The formula employed for computation of the *t*-statistic estimate, for the k^{th} model function in the j^{th} image pixel or detector channel is [38]:

$$\mathbf{c}_k^{\mathrm{T}}\boldsymbol{\beta}_j \sqrt{\frac{N-K}{\boldsymbol{\varepsilon}_j^{\mathrm{T}}\boldsymbol{\varepsilon}_j \mathbf{c}_k^{\mathrm{T}}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{c}_k}} \approx t_{N-K}, \tag{5}$$

where $N - K$ is the number of degrees of freedom, $\boldsymbol{\beta}_j$ is the j^{th} column of $\boldsymbol{\beta}$, $\boldsymbol{\varepsilon}_j$ is the j^{th} column of $\boldsymbol{\varepsilon}$, and $\mathbf{c}_k$ is the k^{th} column of $\mathbf{C}$, a K-row matrix of contrasts [36]; for all computations reported on here, $\mathbf{C}$ was simply the $K\times K$ identity matrix, which is the default value (if $\mathbf{c}_k$ has, for example, two elements equal to ±1, then GLM will test the significance of the sum of or difference between two model functions for $\boldsymbol{\mu}_j$). The corresponding *p*-value, or probability that a value of *t* at least as large as that obtained from Eq. (5) could occur purely by chance, is subsequently computed via the formula [39]:

$$p\left(t_{N-K}\right) = I_{\frac{N-K}{N-K+t^2}}\left(\tfrac{N-K}{2}, \tfrac{1}{2}\right), \tag{6}$$

where $I_x(a,b)$ is the incomplete beta function [39], the definition of which is $I_x(a,b) \equiv \int_0^x t^{a-1}\left(1-t\right)^{b-1}dt \Big/ \int_0^1 t^{a-1}\left(1-t\right)^{b-1}dt$. The in-house software package that performs the GLM computations also can generate spatial maps of *t*-values and/or the corresponding *p*-values.

2.6 Principal Component Analysis – Molgedey-Schuster Algorithm

As elsewhere described [19], the effect of PCA is to express the function $\mu_a(x,y,t)$ as a sum of products of purely spatial and purely temporal basis functions — i.e., $\mu_a(x,y,t) = \sum_m \alpha_m S_m(x,y) T_m(t)$, where the number of terms in the summation is at most equal to the smaller of N and L — that are mutually orthogonal (i.e., $S_m \cdot S_n = 0$, $T_m \cdot T_n = 0$, when $m \neq n$) and that account for the largest possible percentage of the spatiotemporal variance in $\mu_a(x,y,t)$, for any decomposition of the indicated type and irrespective of the point at which the summation is truncated [36]. In practice, we perform PCA by means of singular value decomposition on the $N\times L$ matrix $\boldsymbol{\mu}$: $\boldsymbol{\mu} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}$ [40]. As $N > L$ in our dynamic DOT detector and image sets, the dimensions of the factors on the right-hand side of the preceding equation are $N\times L$ for $\mathbf{U}$ and $L\times L$ for both $\boldsymbol{\Sigma}$ and $\mathbf{V}$. The main-diagonal elements of $\boldsymbol{\Sigma}$ are the non-zero singular values σ_l of $\boldsymbol{\mu}$, and are conventionally arranged so that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_L$. The columns of $\mathbf{U}$ and $\mathbf{V}$ are the normalized eigenvectors of $\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}$ and $\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\mu}$, respectively, with all eigenvectors corresponding to $\sigma_l = 0$ excluded. In practice, eigenvectors for which $0 < \sigma_l << \sigma_1$ typically also are excluded. Then the dimensions of $\mathbf{U}$, $\boldsymbol{\Sigma}$ and $\mathbf{V}$ become $N\times L'$, $L'\times L'$ and $L\times L'$, respectively, where L' is at most equal to L and frequently $L' << L$. If each column of $\boldsymbol{\mu}$ is mean-subtracted prior to the SVD computation, then it follows that the SVD result is equivalent to the previously reported procedure [41] of finding the eigenvectors of $\boldsymbol{\mu}$'s covariance matrix and then projecting the original set of time series onto each eigenvector in turn.

An important caveat is that there is no guarantee that a one-to-one correspondence exists between a set of principal components and the set of physiological processes that gave rise to the observed spatiotemporal variability of $\mu_a(x,y,t)$. If, for example, two or more sources of variability have nearly the same variance, PCA cannot separate them [41], and the mathematical solution produced by the SVD computation could be any linear combination of the underlying functions. In many biological contexts, however, there are substantive grounds for expecting that the physiological processes that ultimately generate the observed spatiotemporal variability are not only orthogonal, but are statistically independent [42]. Phenomenologically, independence means that knowledge of one process, no matter how perfect, gives one no information whatever about the others [28]. Mathematically, the cross-correlation of two independent functions (e.g., sinusoids of different frequencies) is equal to zero for all values of the time-lag parameter τ, while

orthogonal functions (e.g., sin$t$ and cos$t$) do not necessarily have zero cross-correlation for any time lag other than $\tau = 0$ [28]. Thus the two-part premise from which we proceed is: 1) $\mathbf{U} = \mathbf{PM}$, where $\mathbf{U}$ ($N \times L'$) is the matrix of normalized time series that results from applying PCA to $\boldsymbol{\mu}$, $\mathbf{P}$ ($N \times L'$) is the normalized underlying physiological temporal functions of interest, and $\mathbf{M}$ ($L' \times L'$) is the *mixing matrix*, for which the element in row i, column j is the contribution of the i^{th} function of interest to the j^{th} principal component (PC); 2) the expectation value for the cross-correlation between any two columns of $\mathbf{P}$ is zero for all values of τ. From the latter it follows that $\mathbf{P}^{\mathrm{T}}\mathbf{P} = \mathbf{I}_{(L' \times L')}$, which further implies that $\mathbf{M}$ is an orthogonal matrix.

Because $\mathbf{M}$ is unknown in practice, it is not possible to "unmix" the PCs by computing $\mathbf{P} = \mathbf{UM}^{\mathrm{T}}$ (the inverse of an orthogonal matrix is equal to its transpose [40]). Instead, we seek an orthogonal matrix $\mathbf{W}$ such that the columns of $\mathbf{P}' = \mathbf{UW}$ are uncorrelated at L' distinct non-zero values of τ, and take $\mathbf{P}'$ as our working estimate of $\mathbf{P}$. This procedure can be called extended temporal decorrelation (ETD) (by analogy with the extended spatial decorrelation technique described by Stetter et al. [41], who operated upon the spatial parts of the PCs of their data). Several strategies for selecting appropriate time lags to use in an ETD computation have been proposed [43]; in generating the results presented in the current report, we have followed Zheng et al. [29] and used the earliest time at which each PC has a negative local minimum in its autocorrelation function.

The $\tau = 0$ cross-correlations between all pairs of columns of $\mathbf{P}'$ are found by computing the matrix product $\mathbf{P}'^{\mathrm{T}}\mathbf{P}'$ ($L' \times L'$), the ij^{th} element of which is the correlation between the i^{th} and j^{th} time series. The cross-correlations for a time lag of n time steps are the elements of $\mathbf{P}'^{\mathrm{T}}\mathbf{S}_n\mathbf{P}' = \mathbf{W}^{\mathrm{T}}(\mathbf{U}^{\mathrm{T}}\mathbf{S}_n\mathbf{U})\mathbf{W}$, where $\mathbf{S}_n$ ($N \times N$) is the n-step *shift matrix*. The definition for $\mathbf{S}_n$ is $(\mathbf{S}_n)_{ij} = \delta_i^{j+n}$, $1 \le i,j \le N$, where δ_i^{j+n} is the Kronecker delta function (Ref. [40], §5.10); it is a simple matter to show that $\mathbf{S}_n = (\mathbf{S}_1)^n$, and that $\mathbf{S}_{+n}$ and $\mathbf{S}_{-n}$ are a pseudoinverse pair [40]. The original formulation of the ETD concept presented by Molgedey and Schuster [27] made use of the symmetric shift matrix $\mathbf{S}_n^{\,*} = \mathbf{S}_{+n} + \mathbf{S}_{-n}$, in which case $\mathbf{W}^{\mathrm{T}}(\mathbf{U}^{\mathrm{T}}\mathbf{S}_n^{\,*}\mathbf{U})\mathbf{W}$ is a diagonal matrix when n is one of the time-lag values for which the columns of $\mathbf{P}'$ are uncorrelated. Then $\mathbf{W}$ can be found by solving the eigenvalue problem $(\mathbf{U}^{\mathrm{T}}\mathbf{S}_n^{\,*}\mathbf{U})\mathbf{W} = \mathbf{WD}$; the preceding formula generalizes to the case of decorrelations at multiple simultaneous time lags, as it is evident that $\mathbf{W}^{\mathrm{T}}\left[\mathbf{U}^{\mathrm{T}}\left(\sum_i \mathbf{S}_{n_i}^{*}\right)\mathbf{U}\right]\mathbf{W} = \sum_i \mathbf{W}^{\mathrm{T}}\left(\mathbf{U}^{\mathrm{T}}\mathbf{S}_{n_i}^{*}\mathbf{U}\right)\mathbf{W} = \sum_i \mathbf{D}_i \equiv \mathbf{D}$. However, empirical studies performed at Sheffield have led to the conclusion that for the analysis of experimental data, which invariably contain noise, superior results are obtained when unsymmetric shift matrices are used. In that case, $\mathbf{W}^{\mathrm{T}}\left[\mathbf{U}^{\mathrm{T}}\left(\sum_i \mathbf{S}_{n_i}^{*}\right)\mathbf{U}\right]\mathbf{W}$ is not a diagonal matrix, and a more computation-intensive iterative cost-function minimization procedure is required for finding $\mathbf{W}$. In generating the results presented below, we have used a constant power constraint [29], i.e., the solver attempted to find that $\mathbf{W}$ which minimized the cost function

$$C(\mathbf{W}) = \sum_{i=1}^{L'}\sum_{j>i}^{L'}\sum_{n=0}^{L'} \rho\left(p_i(t), p_j(t+\tau_n)\right)^2, \tag{7}$$

where $\rho(a,b)^2$ is the square of the correlation between a and b, and τ_n is the n^{th} time-lag value, with $\tau_0 = 0$. It is noteworthy that the minimization algorithm does not *force* orthogonality upon $\mathbf{W}$, but that the $\mathbf{W}$ produced typically *is* nearly orthogonal nevertheless.

It also should be noted that using an iterative procedure to solve for $\mathbf{W}$ allows us to incorporate *a priori* knowledge of the form of one or more time series in $\mathbf{P}'$ into the computation, in a straightforward manner. As discussed in Subsection 2.5 above, time-series optical tomographic studies frequently involve collection of such ancillary information. Then a penalty term equal to one minus the square of the correlation between a model function and the first column of $\mathbf{P}'$ can be added to the cost function in Eq. (7), and $\mathbf{W}$s corresponding to $\mathbf{P}'$s whose first columns resemble the model function would be selected for. (A different way of augmenting the cost function with *a priori* information is described in Ref. [29].) However, no use was made of this capability in generating any of the results shown in the current report.

3. RESULTS

3.1 PCA-MSA Applied to Noise-free Data

PCA was carried out on the 1000-point time series of absorption-coefficient images, $\mu_a(x,y,t)$, reconstructed from the noise-free sets of detector readings corresponding to the successive states of the target medium's dynamic inclusions. As seen in Figure 4, the first four singular values are at least two orders of magnitude larger than all successive ones (Fig 4(a)), which means that essentially 100% of all spatiotemporal variability in the $\mu_a(x,y,t)$ time series resides in just

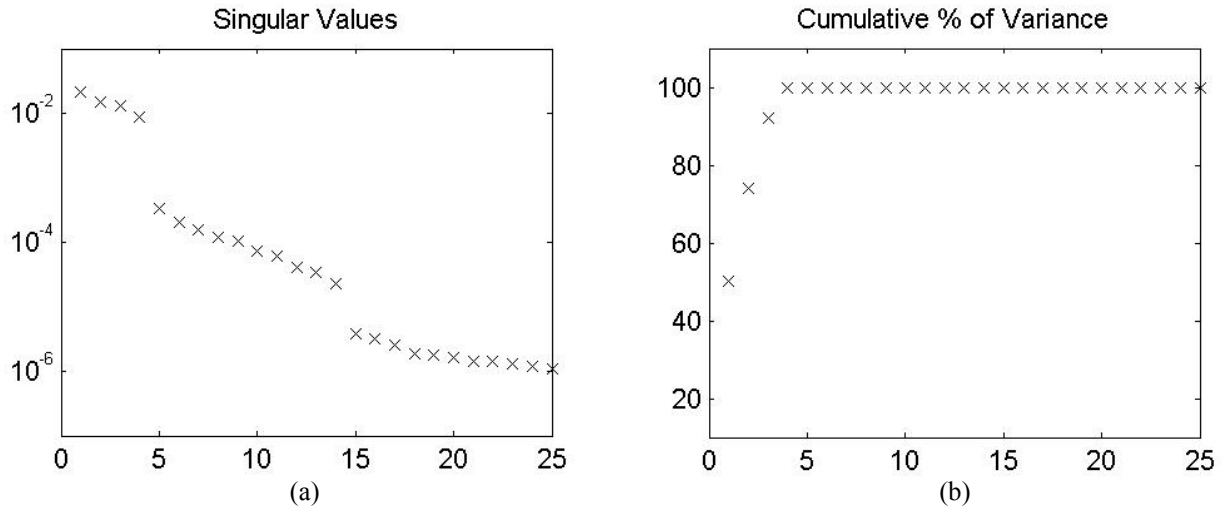
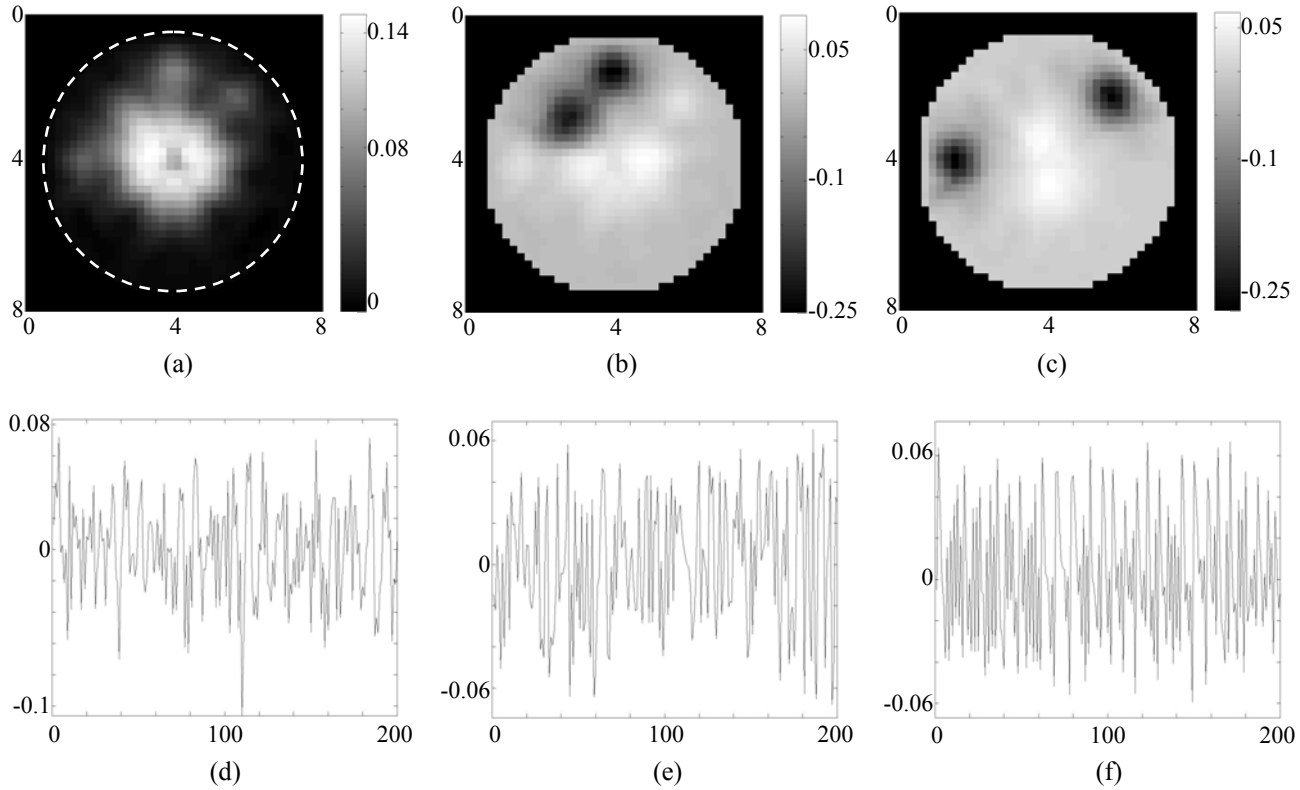


Fig. 4. (a) Singular value spectrum obtained by carrying out PCA on the image time series $\mu_a(x,y,t)$ reconstructed from noise-free computed detector data; the σ_l are arranged in order of decreasing magnitude. (b) Cumulative distribution of the singular values; quantity plotted on the ordinate axis is $100\sigma_l \Big/ \sum_{l'=1}^{L} \sigma_{l'}$.

the first four PCs (Fig. 4(b)). That information, however, consists of a mixture of the four model functions in each PC, as is readily apparent in the data shown in Figure 5. Comparison of the spatial parts of the PCs (Fig. 5(a)-(c), (g)) to Fig. 1 shows that several pairs of inclusions are present in each PC. The same fact is demonstrated quantitatively by computing the coefficients of determination (i.e., 100 times the squared correlation coefficients) between the four temporal model functions and the temporal parts of the PCs (Fig. 5(e)-(f), (h)), which are given in Table 1. The entries in the Table are the percentage of variability in each PC that is linearly attributable to each model function. Inspection



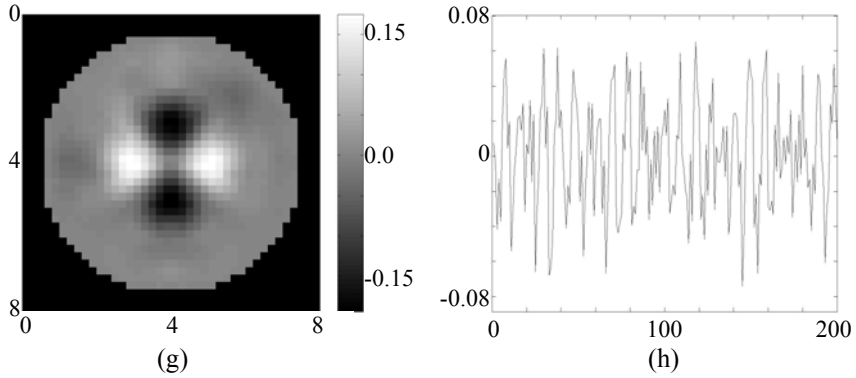


Fig. 5. Spatial [(a)-(c), (g)] and temporal [(d)-(f), (h)] factors of the first four PCs (i.e., the components corresponding to the four largest singular values) of the image time series $\mu_a(x,y,t)$ reconstructed from noise-free computed detector data. For enhanced readability, only the first 200 of the 1000 time points in the temporal factors are shown.

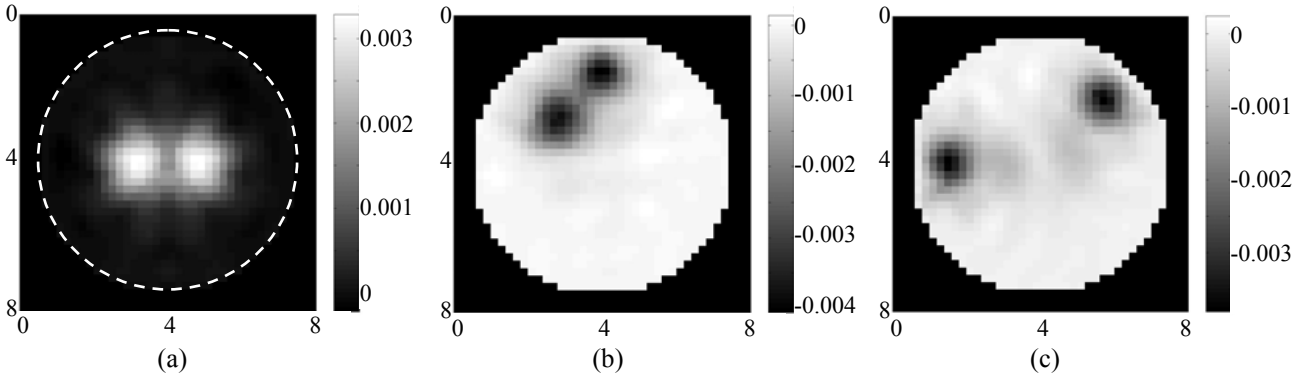
Table 1. Percentage of Variability Attributable to the Temporal Model Functions, in Temporal Factors of PCs Computed from Noise-Free Reconstructed Images

	PC1 (Fig. 5(d))	PC2 (Fig. 5(e))	PC3 (Fig. 5(f))	PC4 (Fig. 5(h))
$\mu_a(t)_{C1}$	47.2	10.8	0.534	41.4
$\mu_a(t)_S$	12.9	86.3	0.149	0.592
$\mu_a(t)_{C2}$	6.53	3.86	88.4	1.20
$\mu_a(t)_Q$	34.1	1.84	5.33	58.7

of the PC time series reveals suggestions of the model functions $\mu_a(t)_S$, $\mu_a(t)_{C2}$, and $\mu_a(t)_Q$ in Fig. 5(e), 5(f), and 5(h), respectively (see Figs. 6 and 7), but it is evident that PCA does not isolate them.

Subsequently, ETD was performed by applying the MSA to the first four PCs of the reconstructed $\mu_a(x,y,t)$ (i.e., $L' = 4$). The resulting unmixed patterns of spatiotemporal dynamics are shown in Figure 6. Comparison of the MS components' spatial portions (Fig. 6(a)-(c), (g)) to Fig. 1 shows that each pair of dynamic inclusions is accurately located and sized, and is almost completely isolated from the others. The temporal parts (Fig. 6(d)-(f), (h)) of the MS components also are seen to reproduce the four isolated modeled dynamic functions (see Ref. [31] for time-domain plots of the latter) almost perfectly. The preceding assertion is corroborated by the coefficients of determination between the MS time series and the model functions, which are given in Table 2.

The preceding result, while nearly ideal, has two limitations. The first is that the PCA-MSA operations were not performed until after the complete $\mu_a(x,y,t)$ time series was reconstructed. The singular value spectrum (Fig. 4) suggests, however, that it might be possible to obtain a result comparable to that in Fig. 6 with the reconstruction of only four images, thereby achieving a 99.6% data reduction. (This lowering of computational burden for the inverse problem can be important. For example, it could allow us to use reconstruction algorithms based on conjugate gradient descent (CGD) instead of SVD-based one used here, without having to give up the high data-processing speed associated with SVD [34]. While the CGD strategy is more computation-intensive than SVD, it can incorporate range constraints and



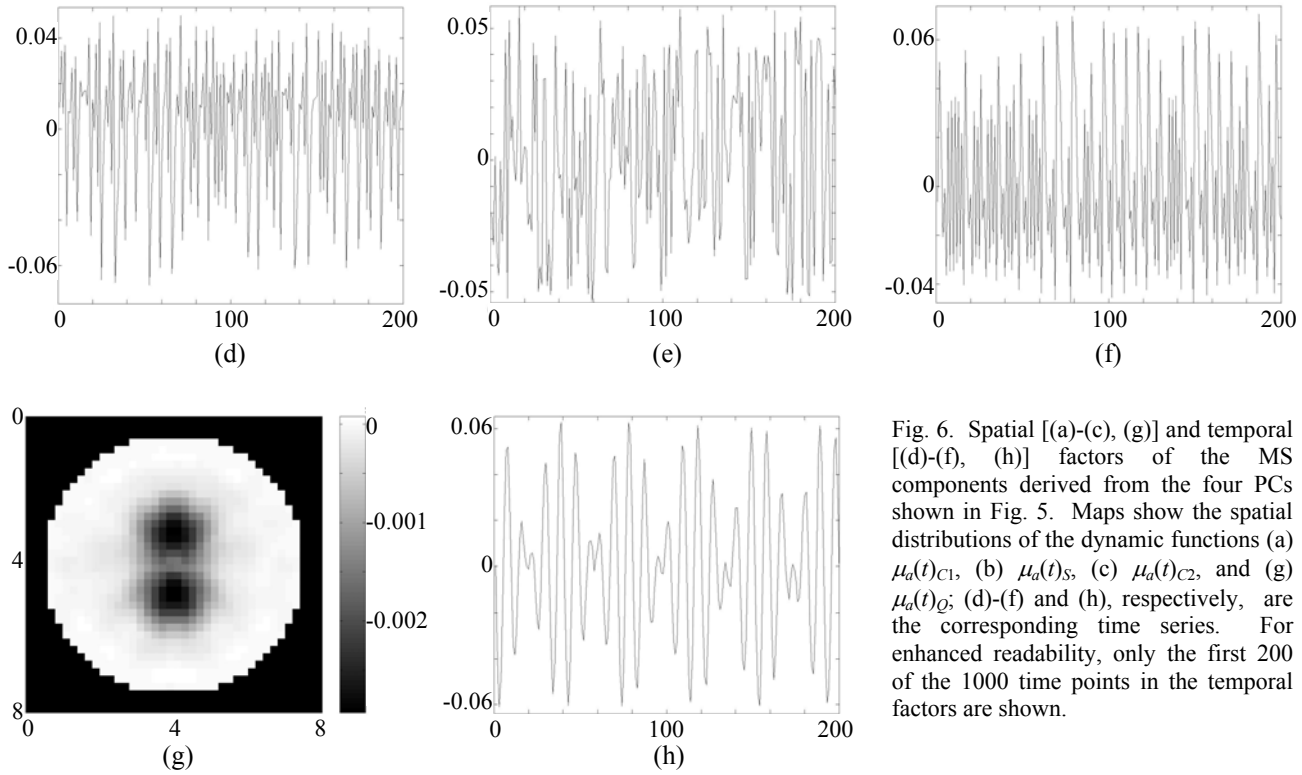


Fig. 6. Spatial [(a)-(c), (g)] and temporal [(d)-(f), (h)] factors of the MS components derived from the four PCs shown in Fig. 5. Maps show the spatial distributions of the dynamic functions (a) $\mu_a(t)_{C1}$, (b) $\mu_a(t)_S$, (c) $\mu_a(t)_{C2}$, and (g) $\mu_a(t)_Q$; (d)-(f) and (h), respectively, are the corresponding time series. For enhanced readability, only the first 200 of the 1000 time points in the temporal factors are shown.

Table 2. Percentage of Variability Attributable to the Temporal Model Functions, in Temporal Factors of MS Components Computed from Noise-Free Reconstructed Images

	MSC1 (Fig. 6(d))	MSC2 (Fig. 6(e))	MSC3 (Fig. 6(f))	MSC4 (Fig. 6(h))
$\mu_a(t)_{C1}$	97.9	6.0×10^{-3}	2.0	3.2×10^{-2}
$\mu_a(t)_S$	5.4×10^{-4}	99.8	0.14	2.5×10^{-2}
$\mu_a(t)_{C2}$	0.14	5.9×10^{-2}	99.7	2.4×10^{-2}
$\mu_a(t)_Q$	2.3×10^{-5}	9.5×10^{-2}	3.3×10^{-2}	99.8

other types of *a priori* knowledge more easily [33].) Second and more fundamental, the addition of even a small amount of noise to the detector data produces the situation in which the noise variance is comparable to the model-function variance in many regions of the image, resulting in a singular value spectrum in which noise may account for a larger percentage of the spatiotemporal variability than one or more models. In such a case, in which the spectrum also loses the classical "staircase" appearance seen in Fig. 4(a), it can be difficult to impossible to select a subset of PCs from which any of the model functions can be isolated.

Both above-mentioned limitations were addressed, and resolved with different degrees of success, by applying the PCA-MSA operations directly to the detector readings and subsequently reconstructing images from only the unmixed PC time series. Note that the PCA and image reconstruction operations are *not* commutative, so that the order of operation is a nontrivial consideration. Therefore it is necessary to determine whether temporal and spatial properties of the target medium can be recovered, when the data-reduction and BSS operations are performed first, with at least the same degree of accuracy as is seen in Fig. 6 above. Accordingly, PCA was carried out on the 1000-point time series of noise-free sets of detector readings corresponding to the successive states of the target medium's dynamic inclusions. The resulting spectrum of singular values is, as seen in Figure 7, qualitatively similar to that obtained from the time series of reconstructed images (Fig. 4). Because it is clear from Fig. 7 that the first four PCs account for essentially all of the spatiotemporal variability in the data set, only they were selected for further processing via the MSA. The "spatial" (row=source position index, column=detector position index) and temporal parts of the resulting unmixed com-

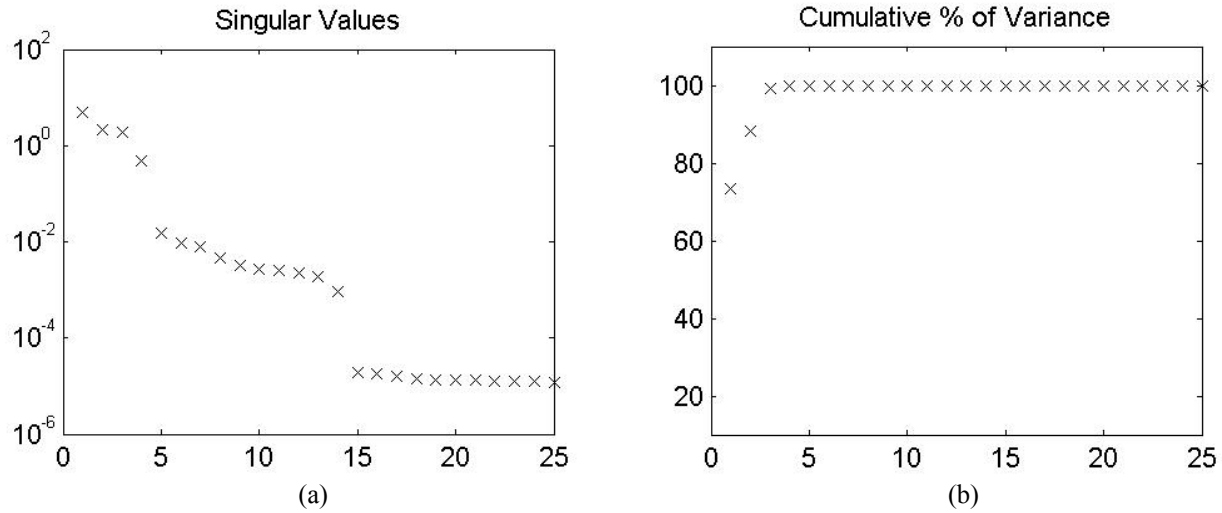
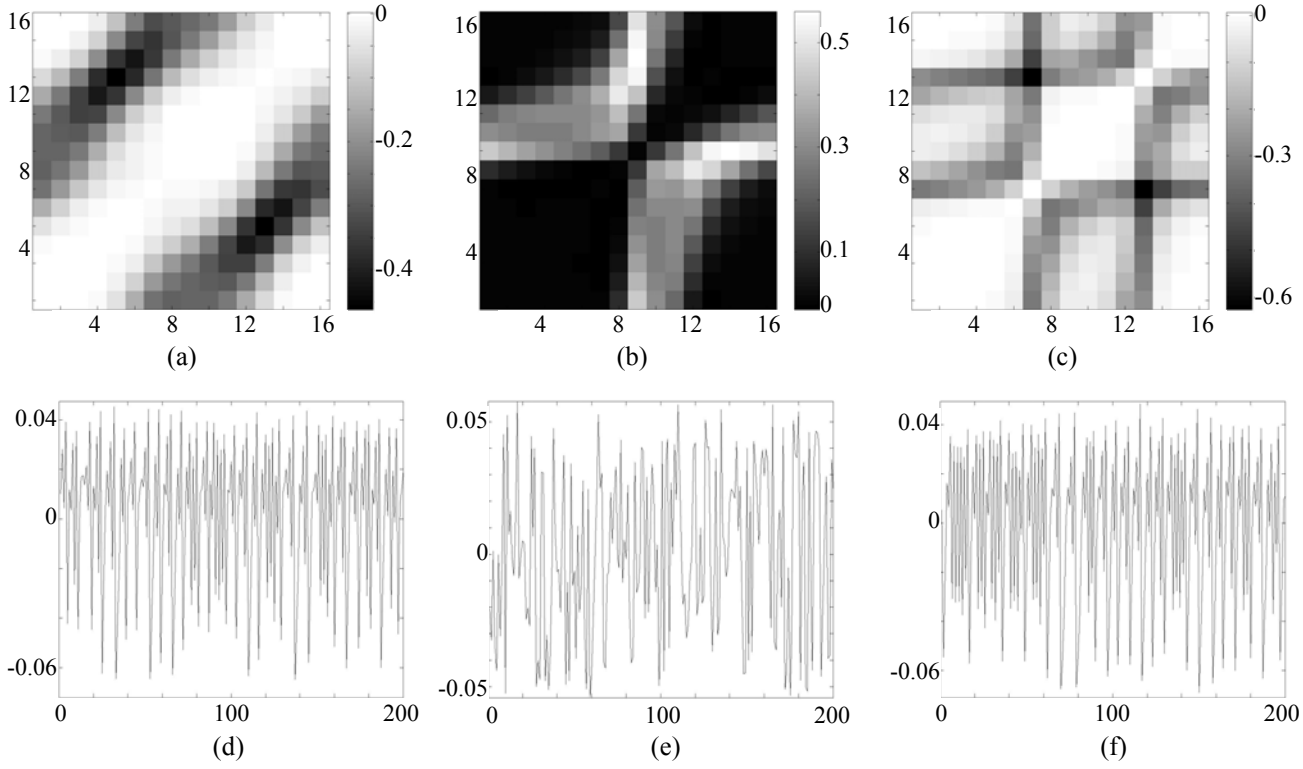


Fig. 7. (a) Singular value spectrum obtained by carrying out PCA on the time series of noise-free computed detector data; the σ_l are arranged in order of decreasing magnitude. (b) Cumulative distribution of the singular values.

ponents, shown in Figure 8, demonstrate that each of the four model functions is present in the reduced data set, and that each (Fig. 8(d)-(f), (h), and Table 3) has been completely isolated. In the "spatial" maps (Fig. 8(a)-(c), (g)), each subset of pixels running parallel to the main diagonal corresponds to channels with a fixed angle between source and detector, while each row corresponds to a fixed source location and each column to a fixed detector location. Then careful inspection of these maps also gives insight into the geometrical structure of target medium: the patterns seen in Fig. 8(a) and 8(g) suggest the presence of a dynamic structure lying near the center of the medium, while those in 8(b) and 8(c) each suggest the presence of localized structures lying near the surface, in the vicinity of one or more particular



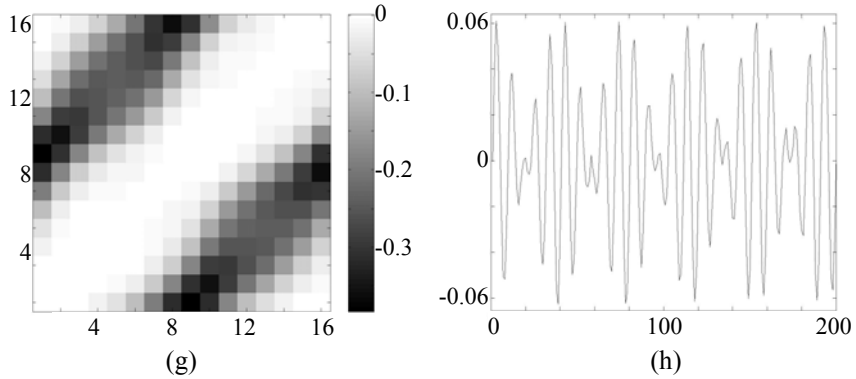


Fig. 8. Spatial [(a)-(c), (g)] and temporal [(d)-(f), (h)] factors of the MS components derived from the first four PCs of the noise-free detector data time series. Maps show the spatial distributions of the dynamic functions (a) $\mu_a(t)_{C1}$, (b) $\mu_a(t)_S$, (c) $\mu_a(t)_{C2}$, and (g) $\mu_a(t)_Q$; (d)-(f) and (h), respectively, are the corresponding time series.

source/detector positions. Further, the pattern in 8(c) clearly suggests that the corresponding dynamics are present in two widely separated near-surface locations. (Another point worth noting is that there is an irreducible two-fold degeneracy in PCA-MSA computation results such as those in Fig. 8, because the algebraic signs of both the spatial and temporal factors can be changed without affecting their product. Consequently, there is no real significance to the observation that the maps in Fig. 8(a), (c) and (g) have one background color and that in Fig. 8(b) has another.)

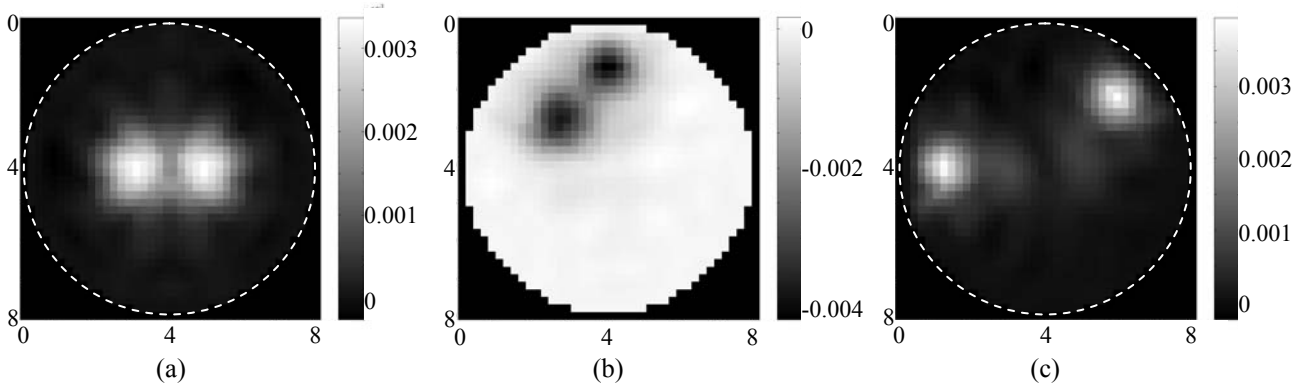
Table 3. Percentage of Variability Attributable to the Temporal Model Functions, in Temporal Factors of MS Components Computed from Noise-Free Detector Data

	MSC1 (Fig. 8(d))	MSC2 (Fig. 8(e))	MSC3 (Fig. 8(f))	MSC4 (Fig. 8(h))
$\mu_a(t)_{C1}$	94.5	1.3×10^{-3}	5.5	2.4×10^{-2}
$\mu_a(t)_S$	5.7×10^{-4}	99.9	0.12	4.4×10^{-3}
$\mu_a(t)_{C2}$	1.8	8.1×10^{-2}	98.1	3.7×10^{-2}
$\mu_a(t)_Q$	1.7×10^{-5}	4.2×10^{-2}	5.9×10^{-2}	99.8

Images were reconstructed using the spatial parts of the four MS components shown in Fig. 8 as the input data. The images obtained are shown in Figure 9. Comparison of these maps to Fig. (1) shows that each pair of dynamic inclusions is accurately located and sized, and is almost completely isolated from the others, while comparison to Fig. 6(a)-(c) and (g) shows that the order of application of the PCA-MSA and image reconstruction operations does not noticeably affect the quality of the final result.

3.2 PCA-MSA Applied to Noisy Data

The addition of a small amount of noise (see curve labeled "3%" in Fig. 3) to the detector data changes the singular value spectrum, qualitatively and quantitatively, to that shown in Figure 10. The principal differences between this result and that in Fig. 7 are that the former has only three PCs containing spatiotemporal information that is not



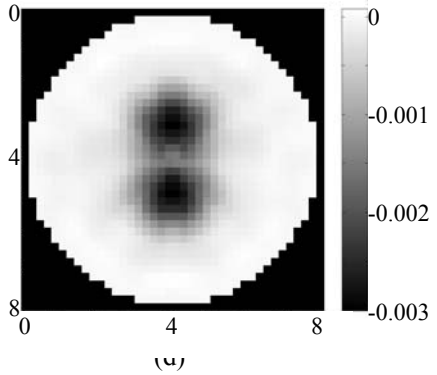


Fig. 9. Images reconstructed from the "spatial" part of the four MS components shown in Fig. 8. Maps show the spatial distributions of the dynamic functions (a) $\mu_a(t)_{C1}$, (b) $\mu_a(t)_S$, (c) $\mu_a(t)_{C2}$, and (d) $\mu_a(t)_Q$.

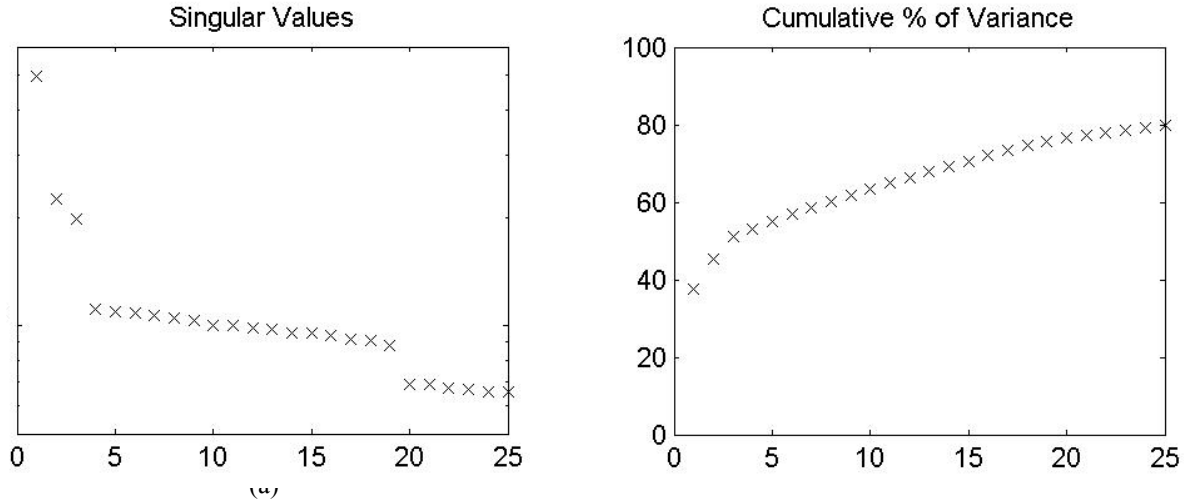
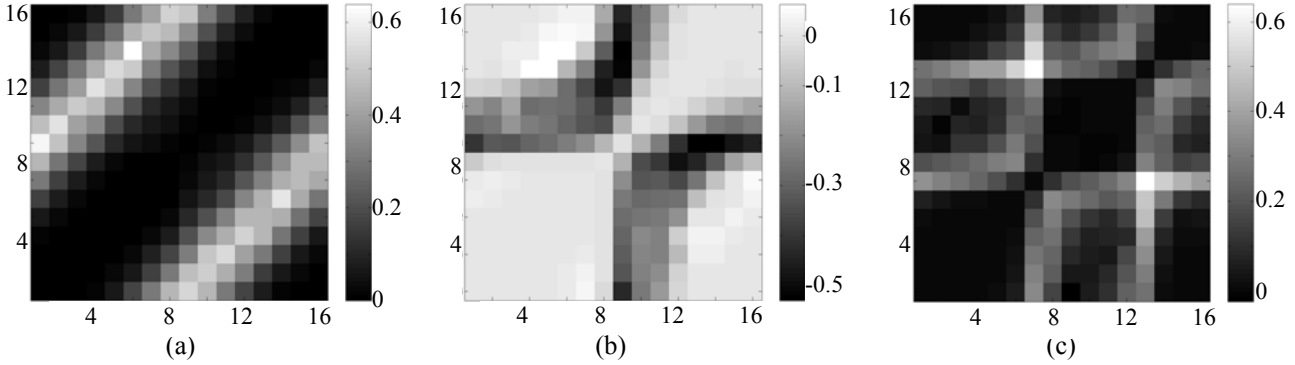


Fig. 10. (a) Singular value spectrum obtained by carrying out PCA on the time series of computed detector data contaminated with a low level of additive Gaussian noise; the σ_l are arranged in order of decreasing magnitude. (b) Cumulative distribution of the singular values.

primarily noise, and that, because of the added noise, these three components account for little more than half of all the variability in the data set. Applying the MSA to the first three PCs of the noisy detector data produces the result shown in Figure 11. The "spatial" maps (Fig. 11(a)-(c)) indicate that the first effect that noise has on the ability to separate the model functions is that the two associated with the deep inclusions become indistinguishable. This is not surprising, because the temporal variations caused by the deep inclusions is distributed over all source/detector channels and so are more thoroughly disrupted by the additive noise than are patterns of variation that are seen by only a few channels, and because the amplitude of the detector variations that they create is small in all channels. The correlations between the



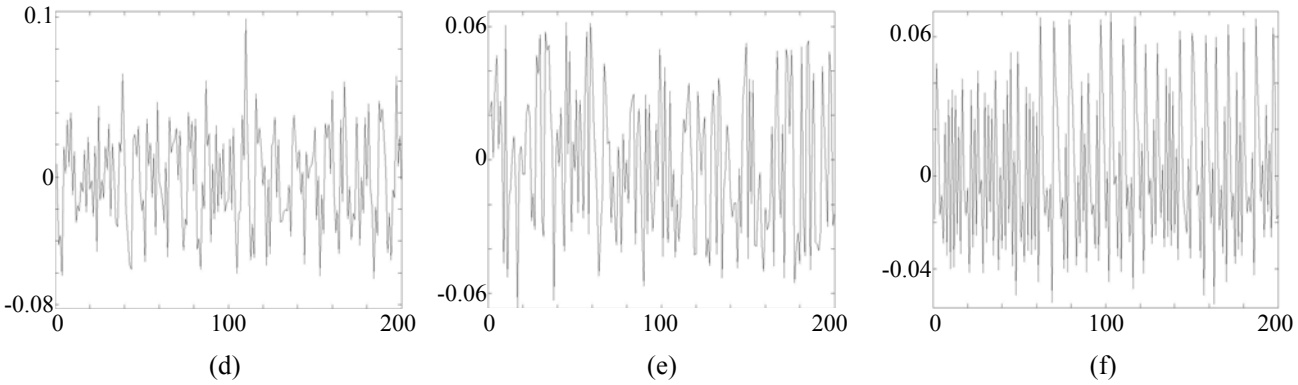


Fig. 11. Spatial [(a)-(c)] and temporal [(d)-(f)] factors of the MS components derived from the first three PCs of the noise-corrupted detector data time series. Maps show the spatial distributions of the dynamic functions (a) $\mu_a(t)_{C1}+\mu_a(t)_Q$, (b) $\mu_a(t)_S$, and (c) $\mu_a(t)_{C2}$; (d)-(f) are the corresponding time series.

The correlations between the four temporal model functions and the temporal parts of the MS components in Fig. 11(d)-(f) are given in Table 4. These confirm that two of the model functions are effectively isolated, while the remaining two are combined in a single component.

Table 4. Percentage of Variability Attributable to the Temporal Model Functions, in Temporal Factors of MS Components Computed from 3% Noise-Added Detector Data

	MSC1 (Fig. 11(d))	MSC2 (Fig. 11(e))	MSC3 (Fig. 11(f))
$\mu_a(t)_{C1}$	50.4	1.11	1.07
$\mu_a(t)_S$	0.960	95.9	5.99×10^{-4}
$\mu_a(t)_{C2}$	0.228	1.98×10^{-2}	94.7
$\mu_a(t)_Q$	39.9	3.04×10^{-2}	0.127

The images reconstructed from the data in Fig. 11(a)-(c) are shown in Figure 12, and again show that the addition of a small amount of noise to the detector data limits our ability to effect a separation of the model functions. Other results not shown have led to the determination that when the degree of added noise is 10% or higher (see Fig. 3), then no accurate spatial or temporal information about the target medium can be extracted from the detector data when the maimally "blind" BSS method described here is used. It can be expected that real experimental data will exhibit less sensitivity to noise than he model used here. Because of the large degree of spectral overlap between the added noise and $\mu_a(t)_{C1}$, $\mu_a(t)_{C2}$, and $\mu_a(t)_S$, and because of the small number of samples per cycle (<10) in $\mu_a(t)_Q$, the noise sensitivity

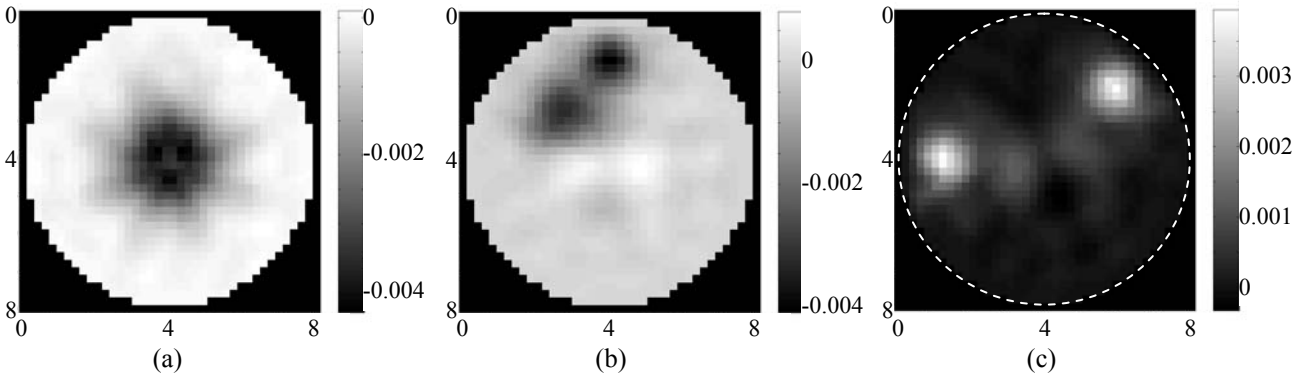


Fig. 12. Images reconstructed from the "spatial" part of the three MS components shown in Fig. 11. Maps show the spatial distributions of the dynamic functions (a) $\mu_a(t)_{C1}+\mu_a(t)_Q$, (b) $\mu_a(t)_S$, and (c) $\mu_a(t)_{C2}$.

of these model functions is predictably high. That sensitivity, however, makes them particularly useful test functions for algorithms that purport to make signal separation more robust.

3.3 GLM; Impact of Detector Noise

Two basic strategies for decreasing the sensitivity to noise currently are available to us, both based upon incorporation of *a priori* knowledge regarding temporal patterns of fluctuations in the medium's properties into the analysis. (As discussed in Subsection 2.4 above, it can be expected that information of this type will typically be available in a dynamic optical tomographic measurement.) One is to modify the MSA cost function so that components whose temporal parts resemble the model functions are selected for. The other, which was used for the current report, is to reconstruct a complete time series of images, then use the GLM for image post-processing. While the latter option is more computation-intensive, it has the advantage of producing useful statistical parameter maps that assist in evaluating the reliability of the images of temporal models' spatial distributions. An example of this is given in Figure 13, in which the result of applying GLM to the noise-free reconstructed $\mu_a(x,y,t)$ is shown. Here the best-fit parameters

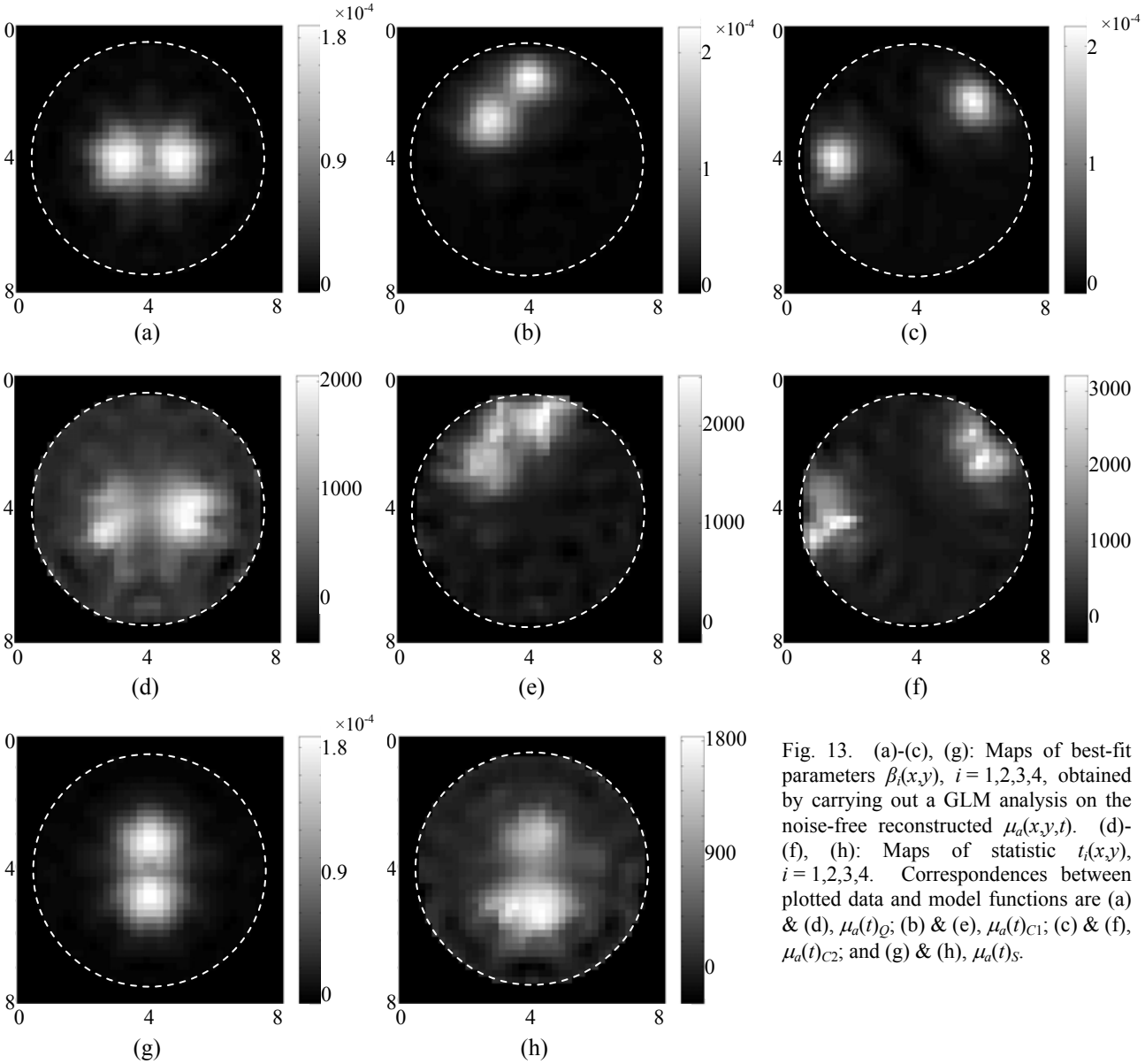
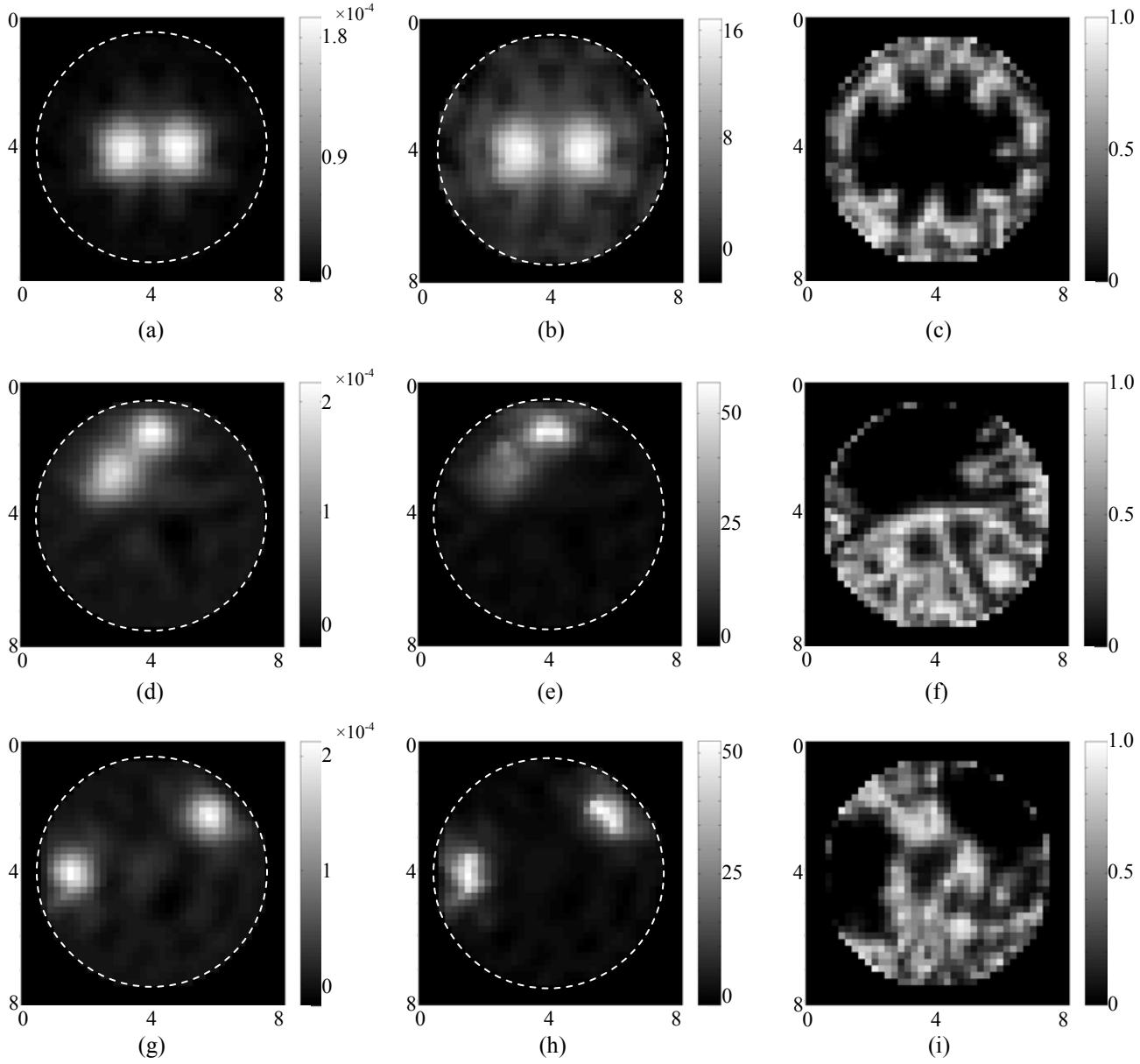


Fig. 13. (a)-(c), (g): Maps of best-fit parameters $\beta_i(x,y)$, $i = 1,2,3,4$, obtained by carrying out a GLM analysis on the noise-free reconstructed $\mu_a(x,y,t)$. (d)-(f), (h): Maps of statistic $t_i(x,y)$, $i = 1,2,3,4$. Correspondences between plotted data and model functions are (a) & (d), $\mu_a(t)_Q$; (b) & (e), $\mu_a(t)_{C1}$; (c) & (f), $\mu_a(t)_{C2}$; and (g) & (h), $\mu_a(t)_S$.

$\beta_i(x,y)$, $i = 1,2,3,4$, are plotted in Fig. 13(a), (b), (c) and (g). Corresponding maps of the statistical parameter $t_i(x,y)$, $i = 1,2,3,4$, computed from Eq. (5), are plotted in Fig. 13(d), (e), (f) and (h). The latter are spatial maps of the *t*-statistic appropriate for testing the hypothesis that values of $\beta_i(x,y)$ at least as large as those seen in the former set of images could occur purely by chance. While the information in Figs. 6 and 8 is of the same quality as that seen in the $\beta_i(x,y)$ maps of Fig. 13, the PCA-MSA computation does not produce estimates of $t_i(x,y)$. The greater robustness of the GLM approach to noise is demonstrated by the data plotted in Figure 14, for which the GLM analysis was carried out on the $\mu_a(x,y,t)$ time series reconstructed from detector data corrupted with 3% (see Figure 3) additive white noise; this is the same level of noise that produced the results shown in Figs. 10-12. GLM, it is seen, retains the ability to correctly locate and size the inclusions associated with each of the four model functions. The third column in Fig. 14 contains the same information as the second, but in a different form; these are maps of $p_i(x,y)$, $i = 1,2,3,4$, the position-dependent probability that values of β_i at least as large as those seen in the first row could occur purely by chance. Comparison of Fig. 14 to Fig. 13 shows further that the effect of a low level of detector noise on the best-fit parameter maps is hardly noticeable. Noise has a much larger effect on the *t*-statistic maps, but the values obtained still correspond to essentially zero probability of chance occurrence of β_is at least as large as those in Fig. 14, within all eight dynamic inclusions.



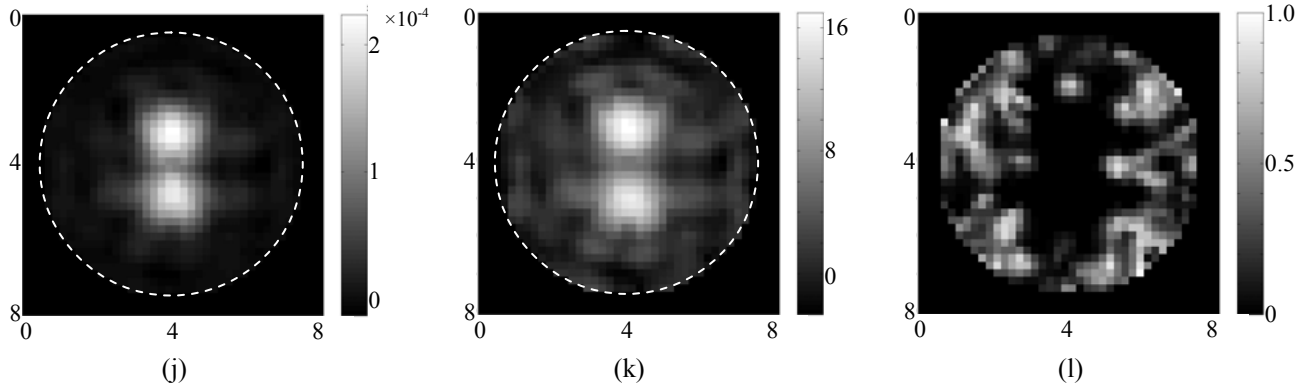
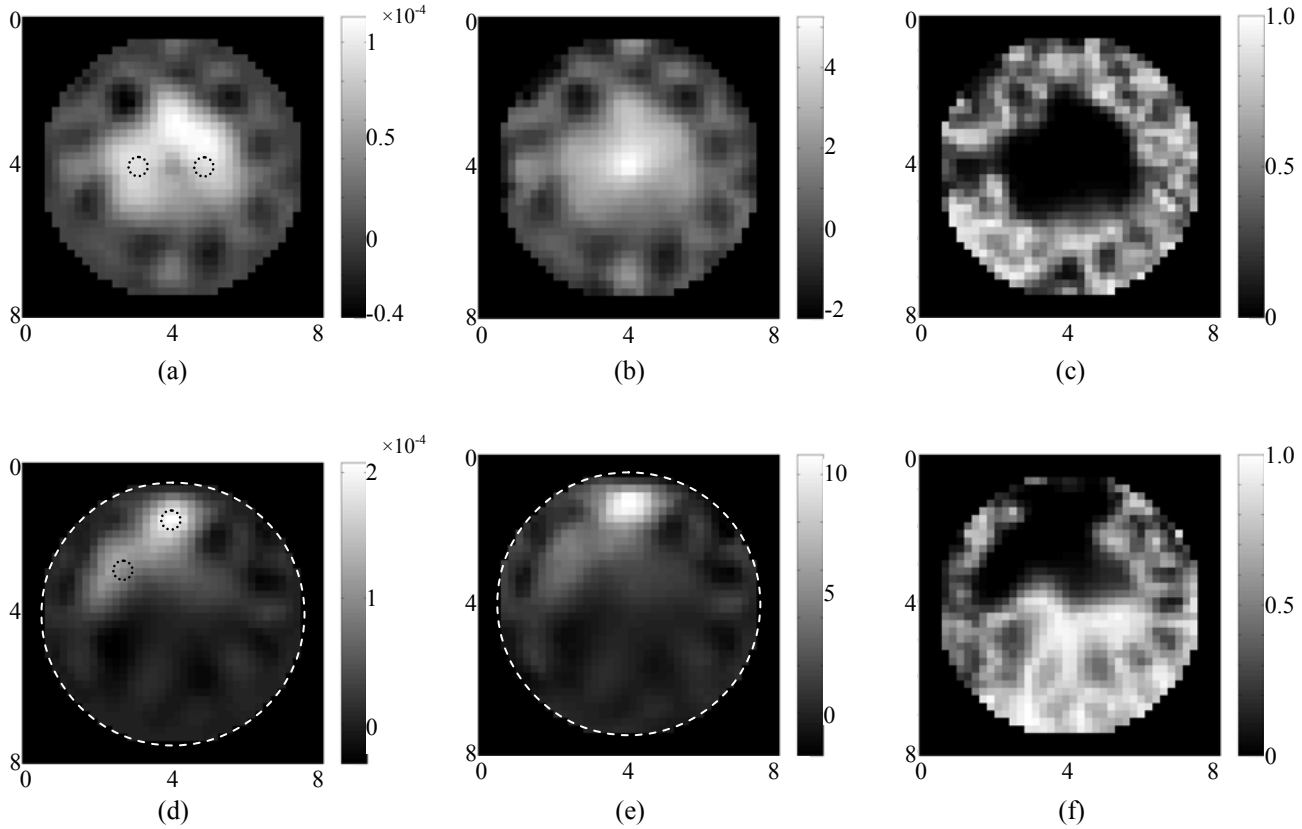


Fig. 14. First column [(a), (d), (g), (j)] – maps of best-fit parameters $\beta_i(x,y)$, $i = 1,2,3,4$, obtained by carrying out a GLM analysis on the 3%-noise-added reconstructed $\mu_a(x,y,t)$. Second column [(b), (e), (h), (k)] – maps of statistic $t_i(x,y)$, $i = 1,2,3,4$. Third column [(c), (f), (i), (l)] – maps of statistic $p_i(x,y)$, $i = 1,2,3,4$. Reading from top to bottom, in any column, plotted data correspond to the model functions $\mu_a(t)_Q$, $\mu_a(t)_{C1}$, $\mu_a(t)_{C2}$, and $\mu_a(t)_S$, respectively.

By progressively increasing the noise-to-signal ratio of the detector data, we have demonstrated that the GLM analysis we have used is highly robust. This can be seen by inspection of the results plotted in Figure 15, for which the $\mu_a(x,y,t)$ time series was reconstructed from detector data corrupted with 50% (see Figure 3) additive white noise. It is seen that even in this extreme case, while significant distortions occur in the spatial distribution of each model function, the presence and general location of each one is correctly identified, with high statistical significance.



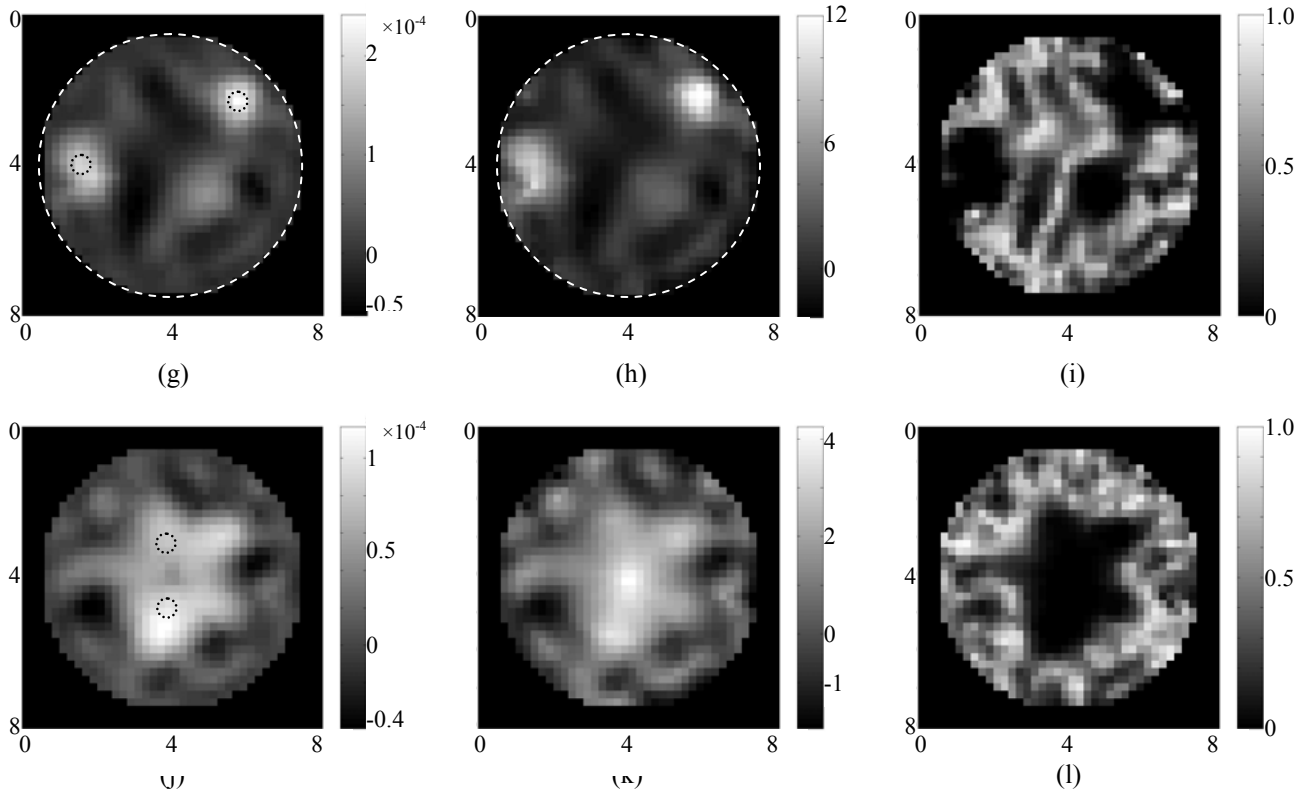


Fig. 15. First column [(a), (d), (g), (j)] – maps of best-fit parameters $\beta_i(x,y)$, i = 1,2,3,4, obtained by carrying out a GLM analysis on the 50%-noise-added reconstructed $\mu_a(x,y,t)$. Second column [(b), (e), (h), (k)] – maps of statistic $t_i(x,y)$, i = 1,2,3,4. Third column [(c), (f), (i), (l)] – maps of statistic $p_i(x,y)$, i = 1,2,3,4. Reading from left to right, in either row, plotted data correspond to the model functions $\mu_a(t)_Q$, $\mu_a(t)_{C1}$, $\mu_a(t)_{C2}$, and $\mu_a(t)_S$, respectively. Small dotted circles in first column maps indicate actual locations and sizes of the inclusions.

4. DISCUSSION

Several authors of the current report have, in recent years, emphasized the use of dynamic optical tomography for diagnostic medical imaging [11,14-16,19,20,31-33]. This interest is based on: 1) knowledge of a number of disease conditions in which structural derangements of microvessels occur, but these are clinically "silent" until irreparable tissue destruction has already taken place; 2) the logical near-necessity that in the development of almost any disease process there is an incipient phase in which the behavior (i.e., dynamics) of cells or cell assemblies becomes abnormal prior to the development of any macroscopic anatomical abnormality; 3) empirical observations of large, characteristic, and repeatable changes in the measured optical tomographic signal occurring in response to mild physiological provocation. It is important to recognize that the indetectibility issue that arises in connection with the first two of the preceding points is a problem of technology, not of physics or biology. All of the tools presently in clinical use for tissue study, whether invasive or not, are either intrinsically insensitive to the relevant functional or microanatomical phenomena, or do not possess spatial and/or temporal resolution high enough to enable their detection. The third enumerated point has led us to hypothesize that dynamic optical tomography may be the technology best suited to redressing these difficulties.

Interpretation of the detector and image data produced by dynamic optical tomography studies will, however, require the development of an analytic capability for partitioning the observed dynamics between vascular and nonvascular effects, and for assigning the former to influences of the various humoral, autonomic, and local metabolic mechanisms that contribute to temporal fluctuations in blood volume and blood oxygenation. Likewise, a capability for partitioning the observed dynamics into factors coordinated with and independent of imposed provocations, and for accurately

quantifying the former even when they account for only a small fraction of the total variability, is needed. Remarkably similar problems are encountered in other areas of research on noninvasive physiological monitoring, such as interpretation of electroencephalographic (EEG) and functional magnetic resonance imaging (fMRI) data [22,24]. It is perhaps ironic that in these latter fields the vascular contributions to the total signal variability usually is part of the noise to be filtered out, leaving behind only the small effects of local neuronal activity [29,41]. Even so, the results presented above have demonstrated that the same mathematical tools frequently used for signal separation and identification in EEG and fMRI studies are potentially useful in analyzing the vascular responses measured via dynamic optical tomography as well.

The examples, presented above, of results obtained when the Molgedey-Schuster algorithm was used to perform blind source separation show that the MSA can separate mixtures of independent temporal functions, provided that the signal-to-noise ratio of the data is sufficiently high. This ability extends to the case of distinguishing two behaviors with identical histograms, which alternative BSS strategies such as ICA cannot accomplish. While the threshold level for acceptable noise reported here might seem low (depending on the application, it is not, in fact, unrealistic), we reiterate that the specific functions that were used as temporal models in this study confer a predictably high degree of noise sensitivity on the computations. It can be expected that the MSA will prove more robust to detector noise when the goal is to effect BSS of physiological rhythms that one would expect to find in real data.

A more important point with regard to the impact of noise, however, is that it can be effectively countered through the incorporation into the MSA cost function of *a priori* information, also called *weak models*, derived from ancillary measurements or from the design time-course of experimental provocations, as explained in Subsection 2.6 above. This we infer from the experience that some of the current report's authors have amassed in using the algorithm to analyze topographic optical image data [29]. Significantly, these successes, and others obtained without use of weak models [41], have occurred notwithstanding the fact that the statistical independence of the dynamic behaviors whose mixtures we observe likely will not be absolute. This further suggests that ETD-based algorithms not sensitive to small violations of their underlying theoretical assumptions. Thus, while we certainly are mindful of the possibility that use of these techniques can produce time series just as thoroughly mixed as the original data if the unobservable behaviors do not possess the independence property, we are not overly concerned by it.

The GLM computations reported on here essentially are a way of attempting to answer the questions: 1) "Where, and with what amplitude, are known temporal patterns present in the medium under study?" and 2) "How much confidence can we have in our answer to the first question?" (The PCA-MSA method, by contrast, essentially is an attempt to answer the complementary question "*What are* the spatiotemporal patterns of behavior present in the medium?") The results presented here indicate that these two questions can be usefully answered even when the fraction of overall variability in the image time series that is associated with optical parameter fluctuations in the medium is comparable to the fraction that results from detector noise. It can reasonably be projected that adoption of this technique as a standard component of data analysis for dynamic optical tomographic data will allow practitioners to achieve spatial resolution at least as good as that obtained with modalities such as functional MRI, while retaining the much higher temporal resolution that has previously been reported [14-16,19,20].

5. ACKNOWLEDGEMENT

This research was supported in part by the US Army Medical Command (Contract No. DAMD17-03-0017, the New York State Department of Health, and by the National Institutes of Health under Grants R21-HL67387, R21-DK63692, R41-CA96102, and R43-CA91725.

6. REFERENCES

1. B. Chance, ed., *Photon Migration in Tissue* (Plenum, New York, 1988).
2. R. F. Bonner, R. Nossal, S. Havlin, and G. H. Weiss, "Model for photon migration in turbid biological media," J. Opt. Soc. Am. A **4**, 423-432 (1987).
3. R. L. Barbour, H. L. Graber, R. Aronson, J. Lubowsky, "Model for 3-D optical imaging of tissue," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, 1395-1399 (1990).
4. S. R. Arridge, P. van der Zee, M. Cope, and D. T. Delpy, "reconstruction methods for infra-red absorption imaging," in *Time-Resolved Spectroscopy and Imaging of Tissues*, B. Chance, ed., Proc. SPIE **1431**, 204-215 (1991).

5. K. D. Paulsen and H. Jiang, "Spatially–varying optical property reconstruction using a finite element diffusion equation approximation," Med. Phys. **22**, 691-701 (1995).
6. B. W. Pogue and M. S. Patterson, "Forward and inverse calculations for near-infrared imaging using a multigrid finite difference method," in *Advances in Optical Imaging and Photon Migration*, R. R. Alfano, ed., Vol. 21 of OSA Proceedings Series (Optical Society of America, Washington, D.C., 1994), pp. 176-180.
7. S. B. Colak, D. G. Papaioannou, G. W. 't Hooft, M. B. van der Mark, H. Schomberg, J. C. J. Paasschens, J. B. M. Melissen, and N. A. A. J. van Asten, "Tomographic image reconstruction from optical projections in light–diffusing media," Appl. Opt. **36**, 180-213 (1997).
8. M. V. Klibanov, T. R. Lucas, and R. M. Frank, "A fast and accurate imaging algorithm in optical/diffuion tomography," Inv. Prob. **13**: 1341-1361 (1997).
9. A. D. Klose and A. H. Hielscher, "Optical tomography sing the time-independent equation of radiative transfer. Part II: Inverse model," J. Quant. Spectrosc. Radiat. Transfer **72**, 715-732 (2002).
10. C. E. Elwell, R. Springett, E. Hillman, and D. T. Delpy, "Oscillations in cerebral haemodynamics," in *Oxygen Transport to Tissue XXI (Advances in Experimental Medicine and Biology, Vol. 471)*, A. Eke and D. T. Delpy, eds. (Kluwer Academic/Plenum, New York, 1999), 57-65.
11. R. L. Barbour, H. L. Graber, Y. Pei, S. Zhong, C. H. Schmitz, "Optical tomographic imaging of dynamic features of dense-scattering media," J. Opt. Soc. Am. A **18**, 3018-3036 (2001).
12. A. Zourabian, A. Siegel, B. Chance, N. Ramanujan, M. Rode, and D. A. Boas, "Trans-abdominal monitoring of fetal arterial blood oxygenation using pulse oximetry," J. Biomed. Opt. **5**, 391-405 (2000).
13. J. E. W. Mayhew, S. Askew, Y. Zheng, J. Porrill, G. W. M. Westby, P. Redgrave, D. M. Rector, and R. M. Harper, "Cerebral vasomotion: a 0.1-Hz oscillation in reflected light imaging of neural activity," Neuroimage **4**, 183-193 (1996).
14. H. L. Graber, C. H. Schmitz, Y. Pei, S. Zhong, S.-L. S. Barbour, S. Blattman, T. Panetta, and R. L. Barbour, "Spatio-temporal imaging of vascular reactivity," in *Physiology and Function from Multidimensional Imaging*, A. V. Clough and C.-T. Chen, eds., Proc. SPIE **3978**, 364-376 (2000).
15. H. L. Graber, S. Zhong, Y. Pei, I. Arif, J. Hira, and R. L. Barbour, "Dynamic imaging of muscle activity by optical tomography," in *Biomedical Topical Meetings*, OSA Technical Digest (Optical Society of America, Washington DC, 2000), pp. 407-408.
16. S. Blattman, H. L. Graber, S. Zhong, Y. Pei, J. Hira, I. Arif, and R. L. Barbour, "Imaging of differential reactivity of the vascular tree in the human forearm by optical tomography," in *Biomedical Topical Meetings*, OSA Technical Digest (Optical Society of America, Washington DC, 2000), pp. 458-460.
17. M Intaglietta, "Arteriolar vasomotion; implications for tissue ischemia" Blood Vessels **28**, 1-7 (1991).
18. M. Pagani, F. Lombardi, S. Guzzetti, O. Rimoldi, R. Furlan, P. Pizzinelli, G. Sandrone, G. Malfatto, S. Dell'Orto, E. Piccaluga, M. Turiel, G. Baselli, S. Cerutti, and A. Malliani, "Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog," Circ. Res. **59**, 178-193 (1986).
19. G. S. Landis, T. F. Panetta, S. B. Blattman, H. L. Graber, Y. Pei, C. H. Schmitz, and R. L. Barbour, "Clinical applications of dynamic optical tomography in vascular disease," in *Optical Tomography and Spectroscopy of Tissue IV*, B. Chance, R. R. Alfano, B. J. Tromberg, M. Tamura, and E. M. Sevick-Muraca, eds., Proc. SPIE **4250**, 130-141 (2001).
20. R. L. Barbour, H. L. Graber, Y. Pei, and C. H. Schmitz, "Imaging of vascular chaos," in *Optical Tomography and Spectroscopy of Tissue IV*, B. Chance, R. R. Alfano, B. J. Tromberg, M. Tamura, and E. M. Sevick-Muraca, eds., Proc. SPIE **4250**, 577-590 (2001).
21. G. Baselli, S. Cerutti, S. Civardi, A. Malliani, and M. Pagani, "Cardiovascular variability signals: towards the identification of a closed-loop model of the neural control mechanisms," IEEE Trans. Biomed. Eng. **35**, 1033-1045 (1988).
22. K. J. Worsley, J.-B. Poline, K. J. Friston, and A. C. Evans, "Characterizing the response of PET and fMRI data using multivariate linear models (MLM)," Neuroimage **6**, 305-319 (1998).
23. R. Vetter, J.-M. Vesin, P. Celka, and U. Scherrer, "Observer of the human cardiac sympathetic nerve activity using noncausal blind source separation," IEEE Trans. Biomed. Eng. **46**, 322-330 (1999).
24. T.-W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources," Neural Comput. **11**, 409-433 (1999).
25. L. De Lathuawer, B. De Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by blind source subspace separation," IEEE Trans. Biomed. Eng. **47**, 567-572 (2000).
26. I. Schießl, M. Stetter, J. E. W. Mayhew, N. McLoughlin, J. S. Lund, and K. Obermayer, "Blind source separation from optical imaging recordings with extended spatial decorrelation," IEEE Trans. Biomed. Eng. **47**, 573-577 (2000).
27. L. Molgedey and H. Schuster, "Separation of independent signals using time-delayed correlations," Phys. Rev. Lett. **72**, 3634-3637 (1994).
28. J. V. Stone, "Independent component analysis: an introduction," Trends Cogn. Sci. **6**, 59-64 (2002).
29. Y. Zheng, D. Johnston, J. Berwick, and J. Mayhew, "Signal source separation in the analysis of neural activity in brain," Neuroimage **13**, 447-458 (2001).
30. T. M. Griffith, "Chaos and fractals in vascular biology," Vasc. Med. Rev. **5**, 161-182 (1994).

31. H. L. Graber, Y. Pei, and R. L. Barbour, “Imaging of spatiotemporal coincident states by DC optical tomography,” IEEE Trans. Med. Imag. **21**, 852-866 (2002).
32. Y. Pei, H. L. Graber, and R. L. Barbour, “Influence of systematic errors in reference states on image quality and on stability of derived information for DC optical imaging,” Appl. Opt. **40**, 5755-5769 (2001).
33. Y. Pei, H. L. Graber, and R. L. Barbour, “Normalized constraint algorithm for minimizing inter-parameter crosstalk in DC optical tomography,” Opt. Express **9**, 97-109 (2001), http://www.opticsexpress.org/abstract.cfm?URI=OPEX-9-2-97.
34. Y. Pei, H. L. Graber, and R. L. Barbour, “A system for imaging functional features of the vascular bed: (II) A method for real-time imaging,” in preparation.
35. R.L. Barbour, S.S. Barbour, P.C. Koo, H.L. Graber, R. Aronson, J. Chang, "MRI-guided optical tomography: Prospects and computation for a new imaging method" IEEE Comput. Sci. Eng. **2**(4), 63-77 (1995).
36. R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 3rd Ed. (Prentice-Hall, 1992).
37. “General linear models (GLM),” *Electronic Statistics Textbook* (StatSoft Inc., Tulsa OK, 2002), http://www.statsoftinc.com/textbook/stathome.html.
38. A. Holmes, “The general linear model and statistical parametric mapping,” http://www.fil.ion.ucl.ac.uk/spm/course/notes01/glm/img22.htm.
39. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN*, 2nd Ed. (Cambridge, 1992), §6.4.
40. D. Zwillinger, ed., *CRC Standard Mathematical Tables and Formulae*, 31st Ed. (Chapman & Hall/CRC, Boca Raton, FL, 2002), §2.6.
41. M. Stetter, I. Schießl, T. Otto, F. Sengpiel, M. Hübener, T. Bonhoeffer, and K. Obermayer, “Principal component analysis and blind separation of sources for optical imaging of intrinsic signals,” Neuroimage **11**, 482-490 (2000).
42. G. D. Brown, S. Yamada, and T. J. Sejnowski, “Independent component analysis at the neural cocktail party,” Trends Neurosci. **24**, 54-63 (2001).
43. M. Stetter, *Exploration of Cortical Function* (Kluwer Academic, Boston MA, 2002).