

Statistical molecular target detection framework for multispectral optoacoustic tomography

Stratis Tzoumas, Andrii Kravtsiv, Yuan Gao, Andreas Buehler, Vasilis Ntziachristos

Abstract— Statistical sub-pixel detection via the adaptive matched filter (AMF) has been shown to improve the molecular imaging sensitivity and specificity of optoacoustic (photoacoustic) imaging. Applied to multispectral optoacoustic tomography (MSOT), AMF assumes that the spatially-varying tissue spectra follow a multivariate Gaussian distribution, that the spectrum of the target molecule is precisely known and that the molecular target lies in “low probability” within the data. However, when these assumptions are violated, AMF may result in considerable performance degradation. The objective of this work is to develop a robust statistical detection framework that is appropriately suited to the characteristics of MSOT molecular imaging. Using experimental imaging data, we perform a statistical characterization of MSOT tissue images and conclude to a detector that is based on the t-distribution. More importantly, we introduce a method for estimating the covariance matrix of the background-tissue statistical distribution, which enables robust detection performance independently of the molecular target size or intensity. The performance of the statistical detection framework is assessed through simulations and experimental *in vivo* measurements and compared to previously used methods.

Index Terms— multispectral optoacoustic tomography, photoacoustic tomography, molecular imaging, statistical sub-pixel detection, spectral unmixing, covariance contamination.

Copyright (c) 2010 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

The work of S. Tzoumas was supported by the DFG GRK 1371 grant. V. Ntziachristos acknowledges support from an ERC Advanced Investigator Award 233161 and the European Union project FAMOS (FP7 ICT, Contract 317744).

*S. Tzoumas is with the Institute for Biological and Medical Imaging (IBMI), and Chair for Biological Imaging, Technische Universität München and Helmholtz Zentrum München, 85764 Neuherberg, Germany (e-mail: strtzoumas@gmail.com).

A. Kravtsiv is with the Institute for Biological and Medical Imaging (IBMI), Helmholtz Zentrum München, 85764 Neuherberg, Germany (e-mail: kravaan@gmail.com).

Y. Gao is with the Karlsruhe School of Optics and Photonics, Karlsruhe Institute of Technology, Germany (email: gyolive@gmail.com)

A. Buehler is with the Institute for Biological and Medical Imaging (IBMI), Helmholtz Zentrum München, 85764 Neuherberg, Germany (e-mail: andreas.buehler@helmholtz-muenchen.de).

*V. Ntziachristos is with the Institute for Biological and Medical Imaging (IBMI), and Chair for Biological Imaging, Technische Universität München and Helmholtz Zentrum München, 85764 Neuherberg, Germany (e-mail: v.ntziachristos@tum.de).

I. INTRODUCTION

MULTISPECTRAL Optoacoustic Tomography (MSOT) offers high-resolution detection of optically absorbing reporter molecules at depths beyond the ones reached by optical microscopy and it is now increasingly considered for visualizing near-infrared fluorescent proteins, fluorescent dyes or absorbing nanoparticles, *in vivo* in small animal or human studies [1-3]. MSOT molecular imaging relies on the identification of the spectral signature of a reporter molecule from the absorbing tissue background and it defines a multispectral target detection problem [4].

Multispectral target detection is essentially a binary classification problem where each pixel has to be identified as target or background. Since the spatial extent of the target is typically only a small fraction of the whole image, binary classification algorithms that are based on the minimization of the misclassification rate are not well suited for this problem [5, 6]. Practical multispectral target detection algorithms are commonly based on the Neyman-Pearson criterion (maximize the probability of detection while retaining the probability of false alarm under a constant threshold [7]) and are derived using a generalized likelihood ratio test (GLRT) [5].

Recently, it was shown that the MSOT molecular imaging sensitivity and specificity can be considerably enhanced through the utilization of GLRT statistical detection methods, like the adaptive matched filter (AMF) [8], for extracting weak spectral contributions stemming from molecular agents of biological significance [4], [9]. AMF allows for statistically modeling and suppressing the MSOT tissue background, i.e. it overcomes the challenge of modelling the spectral variability of tissue using exact spectra. Moreover, by offering a constant false alarm rate (CFAR), AMF enables molecular imaging of high specificity, potentially eliminating the need for *ex vivo* validation.

AMF would be optimal for MSOT applications if the spatially-varying background tissue spectra followed a multivariate Gaussian distribution with known parameters and the spectrum of the target molecule within tissue would be known precisely. However, a number of challenges, typically presented in MSOT molecular imaging applications, may compromise the detection performance of AMF. *First*, the MSOT spectral background may not follow a multivariate Gaussian distribution. *Second*, deep seated molecular targets typically exhibit different spectrum than the one obtained from libraries or individual measurements (*spectral mismatch*) due to the spectral coloring introduced to optical absorbers in deep tissue [4]. *Third*, when the molecular agent is present in high

amount within tissue, the presence of the signal of interest (SOI) within the data compromises the accurate computation of the background statistics, i.e. the mean and the covariance matrix, which are typically computed from all available data with maximum likelihood. This effect, commonly referred to as *covariance contamination* in literature, may cause considerable performance degradation [10-12]. The two latter effects (i.e. the *spectral mismatch* and the *covariance contamination*) are closely linked since the *covariance contamination* does not affect significantly the performance of adaptive processing if the target signal is precisely known [11, 13].

In this work we introduce a novel statistical detection framework for resolving molecular agents within tissue with multispectral optoacoustic tomography. Using experimental imaging data, we statistically characterize MSOT tissue images and conclude to a GLRT detector that is based on the t-distribution. More importantly, using an archive of agent-free MSOT tissue images we introduce a novel approach for estimating the covariance matrix of the background-tissue statistical distribution. The covariance matrix estimated through this approach is uncontaminated, thus offering robust detection performance independently of the molecular target size or intensity. We validate the performance of the proposed method statistically using simulations of artificial target implantation on experimental MSOT data. We further demonstrate the detection performance in purely experimental MSOT data.

The rest of this manuscript is organized as follows: In Sec. II, we offer background on the characteristics of MSOT molecular imaging and introduce statistical sub-pixel detection in the context of MSOT. In Sec III, we describe the collection of the imaging dataset and the target implantation framework that is used for statistically evaluating the detection performance under different conditions. In Sec. IV we introduce the proposed statistical detection framework. In Sec. V we quantitatively evaluate the performance of the proposed statistical detection framework and compare it with currently utilized methods. Concluding remarks are offered in Sec. VI.

II. BACKGROUND

A. MSOT Imaging of Molecular Targets

The wavelength (λ) dependent and position (\mathbf{r}) dependent multispectral optoacoustic images $P(\mathbf{r}, \lambda)$ can be associated to the concentrations of the absorbing molecules through the following relation:

$$P(\mathbf{r}, \lambda) = C(\mathbf{r})\Phi(\mathbf{r}, \lambda) \sum_i c_i(\mathbf{r})\varepsilon_i(\lambda) + n(\mathbf{r}, \lambda) \quad (1)$$

In Eq. (1), $\Phi(\mathbf{r}, \lambda)$ is the space and wavelength dependent light fluence, $\varepsilon_i(\lambda)$ are the wavelength dependent molar absorption coefficients of the optical absorbers (i.e. the absorption spectra) and $c_i(\mathbf{r})$ the unknown associated concentrations at a position \mathbf{r} . $C(\mathbf{r})$ is a scaling factor associated with system effects (i.e. the system's spatial sensitivity field) and the Grüneisen parameter. Finally, $n(\mathbf{r}, \lambda)$ represents the cumulative effect of various types of superimposed noise, such as electronic noise and image reconstruction artifacts.

Quantitative MSOT imaging aims at estimating the concentrations of all absorbers within tissue. For achieving this goal, the unknown optical fluence $\Phi(\mathbf{r}, \lambda)$ must be estimated and accounted for. Using a light propagation model, the light fluence can be related to tissue absorption [$\mu_a(\mathbf{r}, \lambda) = \sum_i c_i(\mathbf{r})\varepsilon_i(\lambda)$] and scattering coefficients [$\mu_s'(\mathbf{r}, \lambda)$], i.e. $\Phi(\mathbf{r}, \lambda) = \Phi(\mathbf{r}, \lambda; c_i(\mathbf{r}), \mu_s'(\mathbf{r}, \lambda))$. Using this model, non-linear inversion schemes have been proposed for quantitatively recovering $c_i(\mathbf{r})$ and $\mu_s'(\mathbf{r}, \lambda)$ [14]. Such non-linear inversion schemes have been investigated in simulations and phantoms [15]. Nevertheless, their *in vivo* application is still limited, possibly due to the increased complexity of *in vivo* tissues. Therefore, the MSOT quantification problem is considered an unmet challenge [16].

In vivo molecular imaging often seeks to accurately detect the position and area occupied by a target molecule within tissue. In this case, the spectral analysis problem formulates as a detection problem. By reformulating (1) to separate the contribution of the tissue background and the target molecule we obtain:

$$\begin{aligned} P(\mathbf{r}, \lambda) &= B(\mathbf{r}, \lambda) + c_t(\mathbf{r})T(\mathbf{r}, \lambda), \\ B(\mathbf{r}, \lambda) &= C(\mathbf{r})\Phi(\mathbf{r}, \lambda) \sum_i c_b^i(\mathbf{r})\varepsilon_b^i(\lambda) + n(\mathbf{r}, \lambda), \\ T(\mathbf{r}, \lambda) &= C(\mathbf{r})\Phi(\mathbf{r}, \lambda)\varepsilon_t(\lambda). \end{aligned} \quad (2)$$

In Eq. (2) $c_b^i(\mathbf{r})$ and $\varepsilon_b^i(\lambda)$ are the concentrations and absorption spectra, respectively, of tissue-intrinsic molecules, while $c_t(\mathbf{r})$ and $\varepsilon_t(\lambda)$ are the concentration and absorption spectrum of the extrinsic molecular target. The goal of molecular target detection is to decide between two competing hypotheses for each pixel in the image, namely $c_t(\mathbf{r})=0$ (non-existing target), or $c_t(\mathbf{r})>0$ (existing target), for extracting the position and area occupied by the molecular agent within tissue. For achieving this, models that capture the background and the target spectral variability ($B(\mathbf{r}, \lambda)$ and $T(\mathbf{r}, \lambda)$) are required. Molecular target detection implies that the amount of the target to be detected is substantially lower than the contribution of background molecules. Therefore, throughout this work it is assumed that the unknown light fluence field is not affected by the molecular target but only by the background tissue molecules (Born-type approximation): $\Phi(\mathbf{r}, \lambda) = \Phi(\mathbf{r}, \lambda; c_b^i(\mathbf{r}), \mu_s'(\lambda))$.

$B(\mathbf{r}, \lambda)$ and $T(\mathbf{r}, \lambda)$ are difficult to be modeled due to the effects of the light fluence $\Phi(\mathbf{r}, \lambda)$, which is hereby assumed an unknown function of space and wavelength, and the noise term $n(\mathbf{r}, \lambda)$. Early spectral optoacoustic methods assumed a constant optical fluence and solved a system of linear equations for spectrally unmixing all photo-absorbing molecules [17]. Typically, in this linear unmixing approach, the spectra of oxygenated (Hb-O₂) and deoxygenated (Hb) hemoglobin are used for modeling the background tissue absorption $B(\mathbf{r}, \lambda)$.

B. Statistical sub-pixel detection of molecular targets

More recently statistical sub-pixel detection methods were shown to outperform previous approaches in molecular imaging cases where the target is sparsely present within the data [4], [9]. Statistical sub-pixel detection methods model the

background spectral variability $B(\mathbf{r}, \lambda)$ using a multivariate statistical distribution, while the detector is typically derived through a generalized likelihood ratio test (GLRT) [18]. A well-known algorithm in this category is the AMF [8]:

$$D_{AMF}(\mathbf{x}_i) = \frac{1}{N\mathbf{s}^T\mathbf{G}^{-1}\mathbf{s}} (\mathbf{s}^T\mathbf{G}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}))^2 \quad (3)$$

In Eq. (3), \mathbf{x}_i is an L -dimensional vector corresponding to the MSOT intensity of the pixel under test (L being the number of wavelengths), \mathbf{s} is an L -dimensional vector representing the absorption spectrum of the target molecule (i.e. $\mathcal{E}_i(\lambda)$). $\boldsymbol{\mu}$ and \mathbf{G} are the mean and the covariance matrix ($L \times L$) of the multivariate Gaussian distribution that models the spectral background. Under the assumption that the molecular target lies in “low probability” within the data, the parameters $\boldsymbol{\mu}$ and \mathbf{G} of the background statistical distribution can be computed from the all available pixels with maximum likelihood, i.e. $\boldsymbol{\mu} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$ and $\mathbf{G} = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T$, where N is the number of pixels in the MSOT image.

The performance of statistical detection algorithms depends on the accurate computation of its parameters, i.e. the mean $\boldsymbol{\mu}$ and the covariance matrix \mathbf{G} [19]. If the true covariance matrix \mathbf{G} was known, an inaccurate estimate of the mean value only affects the scaling of the detection image. Conversely, inaccurate computation of the covariance matrix can have a major adverse impact on the target detectability [10, 12, 13]. Inaccurate computation of \mathbf{G} may result from the prominent presence of the target spectrum within the pixels that are used for its computation (*covariance contamination*). A method previously considered for mitigating this effect is *diagonal loading* [20], i.e. the superposition of a diagonal matrix with constant entries to the sample covariance matrix multiplied by a scalar b called *loading factor*, i.e. $\mathbf{G}_{dl} = \mathbf{G} + b\mathbf{I}$. If b is appropriately selected, diagonal loading has been shown to mitigate the effects of covariance contamination in hyperspectral imaging [12]. However, the optimal value of b , depends on the level of contamination, which is generally unknown and may vary substantially depending on the molecular imaging application at hand. For this reason, diagonal loading does not present a robust solution for the MSOT molecular target detection problem.

III. IMAGING DATA AND TARGET IMPLANTATION

A. Synthetic MSOT data

The quantitative evaluation of detection algorithms in the context of MSOT molecular imaging is performed using artificial target implantation on experimental MSOT data, a concept that was previously described in [4].

1) Experimental MSOT dataset of tissue background.

Experimental MSOT images of *in vivo* background tissue were obtained using an In Vision 256-TF MSOT system (iThera Medical GmbH, Munich Germany). A previous version of this system is analytically described elsewhere [21]. The MSOT scanner acquires 2D images at a transverse slice in the focal plane of the ultrasound array. By translating the animal in the imaging chamber, multiple 2D images can be

acquired, covering the whole body. Imaging was performed at 21 different excitation wavelengths from 700 nm to 900 nm with a step size of 10 nm. Image reconstruction was performed using a model-based inversion algorithm [22] with Tikhonov regularization. After image reconstruction the tissue area was manually segmented in each image.

The imaging dataset includes three nude mice, imaged *in vivo* at 41-45 different 2D slices covering the entire area of the brain and the area of the liver, kidneys and the lower abdomen area. Each mouse was imaged at two different physiological conditions (breathing 100% O₂ and 20% O₂, i.e. medical air). Two different oxygenation conditions were employed to consider changes in the spectra of hemoglobin in tissue that affect the background spectral variability. The imaging dataset is composed of 292 multispectral images in total.

2) Target implantation 1: small targets

Target implantation is achieved through the artificial superposition of targets with the spectra of Indocyanine Green (ICG) or IntegriSense750 (PerkinElmer Inc. Massachusetts, U.S.) at randomly selected positions upon the MSOT background-tissue image. The introduced disk-shaped targets (radius 4 pixels; 37 pixels in total) are implanted at 20 simulated peak absorption coefficients decreasing exponentially from 3 to 0.02 cm⁻¹. The intensity of the implanted target at each pixel is defined by the simulated absorption coefficient and the optical fluence at this pixel [$\Phi^{sim}(\mathbf{r}, \lambda)$] which is simulated using a 2D finite element solution of the diffusion approximation and uniform tissue optical properties stemming from literature [4]. The simulated optical fluence is wavelength dependent introducing changes in the spectral signature of the implanted molecular targets, which is computed per pixel as in $s^{imp}(\mathbf{r}, \lambda) = s^{orig}(\mathbf{r}, \lambda)\Phi^{sim}(\mathbf{r}, \lambda)$. We assumed a Born-type approximation where we neglect the effect of the molecular target itself on the simulation of $\Phi^{sim}(\mathbf{r}, \lambda)$. A more analytical description of the target implantation on MSOT images is provided in Ref. [4] and Supp. Fig. 1. We note that, due to a number of simplifying assumptions the simulated absorption coefficient of the implanted targets may not be accurate in absolute value. However, this does not impair the quantitative evaluation since it is only the relative detection performance that is of interest.

A target is considered detectable if there exists a detection threshold T_s that allows for more than 70% of true positives (the true positive set is defined as $TP = \{i \in D_t : D(\mathbf{x}_i) > T_s\}$) and less than 0.045% false positives (the false positive set is defined as $FP = \{i \in D_b : D(\mathbf{x}_i) > T_s\}$):

$$\exists T_s : |TP| > 0.7|D_t| \wedge |FP| < 0.00045|D_b|, \quad (4)$$

where i is the pixel index, D_b and D_t are the sets of pixels corresponding to the background and target, respectively, and $|\cdot|$ denotes the cardinality of the set. $D(\mathbf{x}_i)$ is the detection value attributed to the pixel i . We note that the percentage of 0.045% for false positives was selected because the background area is much larger than the target area. The value $0.00045|D_b|$ is equivalent to $0.5|D_t|$.

In the case of small implanted targets, the detection performance is assessed in terms of minimum detectable agent

amount (MDAA; until cm^{-1}) that is required for successful detection, according to inequality (4). Lower MDAA values indicate higher molecular imaging sensitivity.

2) Target implantation 2: large targets

For quantifying the effects of covariance contamination due to the presence of the molecular target, the target implantation framework was adapted for further considering targets of varying sizes. A large rectangular target occupying a total area of 4096 pixels is initially implanted in tissue background at a randomly selected position. This initial target is iteratively divided into 2, 4, 8 and 16 equally sized rectangular regions corresponding to 2048, 512 and 256 pixels, respectively. In each division stage the target is sequentially implanted at all positions covering the whole initial area of 4096 pixels. In each case the target absorber is implanted at 7 exponentially decreasing simulated absorption coefficient from 3 to 0.03 cm^{-1} . The spectrum of the target at each pixel is computed as in $s^{\text{imp}}(\mathbf{r}, \lambda) = s^{\text{orig}}(\mathbf{r}, \lambda) \Phi^{\text{sim}}(\mathbf{r}, \lambda)$. $\Phi^{\text{sim}}(\mathbf{r}, \lambda)$ is computed as described in (III-A-2) considering only background tissue absorbers.

For large targets the visibility metric defined in (4) does not offer appropriate intuition on the detection result. Instead, the percentage of the detected pixels, i.e. the percentage of pixels within the target area that are attributed detection values larger than a threshold T_b , is computed. This threshold is defined as the average of the 500 larger false-positive detection values (i.e. $T_b = (1/500) \sum_{i=1}^{500} D_b^s(i)$, where D_b^s is a vector containing the detection values attributed to the background sorted in descend order). The average of 500 pixels was selected for avoiding the dominant influence of small outliers in the quantitative analysis. While the region of interest (ROI) corresponding to the target changes according to the target size and position, the same ROI was always used for the computation of false positives. The same false positive ROI was used in order to facilitate a straightforward comparison where any performance differences between large and smaller targets are directly attributed to the effect of covariance contamination (see Suppl. Fig. 2). In each target division stage the total percentage of detected pixels within the whole initial area of 4096 pixels is computed.

B. Experimental imaging dataset

For confirming observations stemming from synthetic data, a purely experimental MSOT dataset of mice containing fluorescent molecular targets was compiled:

1) *In vivo localized insertion dataset*: A capillary tube (diam. 0.86 mm) containing a fluorescent dye was rectally inserted into an anesthetized CD1 mouse and the animal was imaged in the lower abdominal area. The capillary tube was iteratively filled with the fluorochrome Alexa Fluor 750 (AF750) at an optical density ranging from 6.6 to approx. 0.2 OD for achieving different target intensities. The same experiment was performed using the fluorochrome Alexa Fluor 790 (AF790). The dataset is described in Ref. [9]. Imaging was performed at 21 wavelengths from 700 to 900 nm with a step size of 10 nm. The fluorescent target appeared at an imaging depth of ~ 1 cm in the transverse MSOT image.

2) *Brain tumor dataset*: A nude mouse bearing a brain tumor expressing a near-infrared fluorescent protein (U87-iRFP) was imaged at 12 different wavelengths (690, 695, 700, 705, 710,

715, 720, 730, 750, 770, 800, and 830 nm) in the area of the brain. Fluorescence cryoslicing imaging was employed for validating the location of iRFP *post-mortem*. The study is described in Ref. [3].

3) *Bain injection dataset*: Different amounts of fluorescently labelled macrophages were injected in the left and the right brain hemisphere of a euthanized nude mouse. Cells were labeled with the near-infrared fluorescent cyanine dye 1,1'-Diiodo-3,3,3',3'-Tetramethylindotricarbocyanine Iodide (DiI). The exact position of the implanted cells was identified *post-mortem* through cryoslicing fluorescent imaging. Imaging was performed at 21 wavelengths from 700 to 900 nm with a step size of 10 nm. The dataset is analytically described in Ref. [23].

4) *Systemic injection of ICG*: 40 nmoles of Indocyanine Green (ICG) diluted in 200 μL saline were injected in the tail-vein of a nude mouse and in the following the animal was imaged in the area of the liver. ICG has been previously documented to accumulate in the liver [24]. Imaging was performed at 21 wavelengths from 700 to 900 nm with a step size of 10 nm.

5) *Systemic injection of AF750*. 1.2 nmoles of Alexa Fluor 750 (AF750) diluted in 200 μL saline were injected in the tail-vein of a nude mouse and the animal was imaged in the area of the bladder, where AF750 accumulates [24]. Imaging was performed at 21 wavelengths from 700 to 900 nm.

All animal procedures were approved by the District Government of Upper Bavaria.

IV. ROBUST STATISTICAL DETECTION FRAMEWORK

In this section we describe the statistical detection framework developed. In (IV-A), we consider the issue of MSOT background-tissue statistical modeling and in (IV-B) we introduce a robust covariance estimation approach for mitigating the effects of covariance contamination. The two concepts are combined in (IV-C) for deriving the proposed statistical detection framework.

A. Statistical characterization of MSOT data and EC-GLRT detection

Identifying a statistical distribution that accurately describes the background spectral variability $B(\mathbf{r}, \lambda)$ is an important factor in the design of a target detection algorithm [19]. We hereby compare the multivariate Gaussian assumption, followed by AMF, to the multivariate t-distribution which offers better ability to model multidimensional data with longer tails [25]. Both the Gaussian and the t-distribution belong to the family of the elliptically contoured distributions that were introduced by Manolakis et. al. in hyperspectral imaging for the purpose of synthetic data generation [25]. For statistically characterizing the multidimensional MSOT data we use univariate statistics of their Mahalanobis distance distribution as proposed in [25, 26].

The Mahalanobis distance of L -dimensional data that follow the multivariate Gaussian distribution (m_G see Appendix A) is distributed as a χ^2 distribution with L degrees of freedom. By contrast, the Mahalanobis distance of data following the

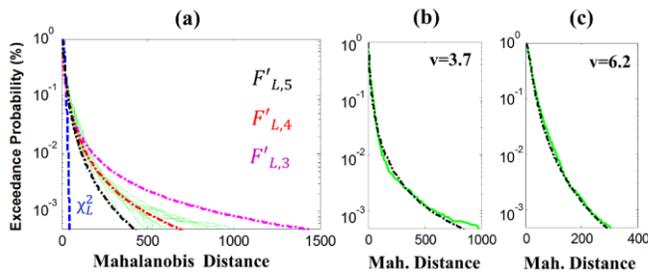


Figure 1. Statistical characterization of MSOT data through the exceedance probability of their Mahalanobis distance distribution. (a) Probability of exceedance of the Mahalanobis distance of 10 different MSOT images (green solid lines). The theoretical exceedance probability of the χ^2 distribution (blue dashed line) and the F distribution with different degrees of freedom (dash-dotted lines) are also presented for comparison. (b, c) Comparison of the exceedance probability of the Mahalanobis distance of two different MSOT images (b, c solid green lines) with the $F_{L,\nu}$ distribution (dash-dotted lines), where ν is selected by minimizing the exceedance metric.

multivariate t distribution with ν degrees of freedom (m_t , see Appendix B) is distributed as an F distribution with parameters L and ν . By studying the Mahalanobis distance distribution of experimental MSOT data, an appropriate multivariate statistical distribution can be potentially identified for modeling the MSOT tissue background.

The probability of exceedance, which corresponds to the probability that the data will exceed a certain value (Appendix C), has been proposed for statistically characterizing the Mahalanobis distance of multidimensional data [26]. The exceedance metric (Eq. (5)), compares the distribution of the Mahalanobis distance \mathbf{m} (Appendix B) to a known probability density function (PDF) f , allowing for the identification of appropriate distributions as well as unknown parameters of such distribution. The exceedance metric is defined through the inverse exceedance probability (Appendix C) as in:

$$M_{exc}(\mathbf{m}, f) = \sum_{i=1}^K \left| E_m^{-1}(P_i) - E_f^{-1}(P_i) \right|, \quad (5)$$

In Eq. (5), P_i are K logarithmically spaced values of the exceedance probability spanning the range from 1 to 10^{-4} , and E^{-1} is the inverse exceedance probability (Appendix C). E_m is the exceedance probability of the Mahalanobis distance distribution of the data, computed numerically, and E_f is the theoretical exceedance probability of the PDF f , which is computed from the analytical expression of f .

Fig. 1(a) presents the probability of exceedance of the Mahalanobis distance for 10 different MSOT background images corresponding to different anatomical areas of a mouse ranging from the brain to the abdominal region (green solid lines). The theoretical exceedance probabilities of the associated χ^2 distribution (blue dashed line) and the F distribution for different values of the parameter ν (dash-dotted lines) are also presented for comparison. From the exceedance probability plot it is obvious that the Mahalanobis distance of MSOT data does not follow a χ^2 distribution, while the F distribution provides a far better fit. This indicates that the t-distribution may be more appropriate for modeling

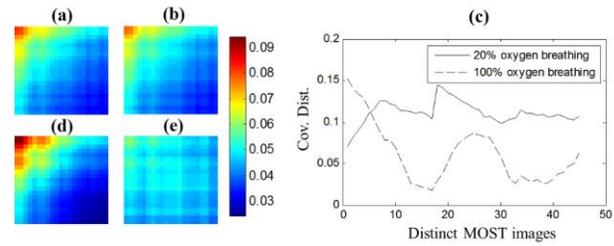


Figure 2. Comparison of MSOT background-tissue covariance matrices. (a, b) Covariance matrices stemming from two different mice (\mathbf{G}_{GL}^1 and \mathbf{G}_{GL}^2 , respectively) using data from the whole body and two different breathing conditions. (c) Graph presenting the distance between \mathbf{G}_{GL}^1 and the covariance matrices \mathbf{G}_s^k corresponding to all anatomical slices of mouse 3. Cov. Dist. = $\left\| \frac{\mathbf{G}_{GL}^1}{\|\mathbf{G}_{GL}^1\|_F} - \frac{\mathbf{G}_s^k}{\|\mathbf{G}_s^k\|_F} \right\|_F$. (d, e) Covariance matrices produced from single anatomical slices of mouse 3 corresponding to the median (d) and maximum (e) distance as compared to \mathbf{G}_{GL}^1 .

the MSOT tissue background. Fig. 1(b), (c) present the distribution of the Mahalanobis distance for two different MSOT images (solid green lines). The dash-dotted black lines correspond to the exceedance probability of the $F_{L,\nu}$, where the parameter ν is computed as the value that minimizes the exceedance metric of (5). We observed that different MSOT images may correspond to different degrees of freedom ν . Fig. 1(b), (c) further suggest that the F distribution with optimally selected parameter ν appears capable of modeling well both the main body as well as the tails of the Mahalanobis distance distribution.

Theiler et. al. have derived a generalized likelihood ratio test (GLRT) detector for multivariate data following the t-distribution with ν degrees of freedom ($\nu > 2$), termed EC-GLRT detector [27]:

$$D_{EC-GLRT}(\mathbf{x}_i) = \frac{(\nu-1)}{(\nu-2) + (\mathbf{x}_i - \boldsymbol{\mu})^T \mathbf{G}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})} \frac{(\mathbf{s}^T \mathbf{G}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}))^2}{\mathbf{s}^T \mathbf{G}^{-1} \mathbf{s}} \quad (6)$$

In Eq. (6), \mathbf{G} is the covariance matrix which can be calculated with maximum likelihood and ν are the degrees of freedom of the t-distribution. For $\nu = \infty$ the EC-GLRT detector becomes equivalent to AMF. EC-GLRT requires an estimate of the parameter ν , which is hereby estimated for each dataset under test through the minimization of the exceedance metric of (5):

$$\hat{\nu} = \underset{\nu}{\text{Arg min}} M_{exc}(\mathbf{m}, F'_{L,\nu}), \quad (7)$$

B. Robust covariance matrix estimation through quasi-local covariance shrinkage.

A major challenge in achieving simultaneously sensitive and robust detection performance is the estimation of a covariance matrix that is always uncontaminated from the SOI and also close to the true covariance matrix. To achieve this we introduce a covariance estimation scheme where the sample covariance matrix is appropriately merged with an uncontaminated, global covariance matrix derived from a an archive of agent-free MSOT images of tissue background. Our covariance estimation approach follows three distinct steps. In a first step, an uncontaminated covariance matrix is estimated

by combining the sample covariance matrix of the MSOT image under test and a global covariance matrix that is computed using uncontaminated training data. In a second step, the level of contamination of the image under test is estimated by comparing its sample covariance matrix with a dictionary of uncontaminated ones. Finally, in a third step, the uncontaminated covariance matrix and the sample covariance matrix are appropriately merged based on the estimated level of covariance contamination. Step one is described in (IV-B-1) and steps two and three are described in (IV-B-2).

1) *Estimation of uncontaminated covariance matrix using training data*

In MSOT imaging an un-contaminated global covariance matrix can be computed using training data stemming from animals where no molecular agents have been introduced. Fig. 2(a), (b) present two global covariance matrices (\mathbf{G}_{GL}^1 and \mathbf{G}_{GL}^2 , respectively) computed from two different mouse datasets (mouse 1 and 2, respectively). Each global covariance matrix was computed as the sample covariance matrix over all pixels of MSOT images spanning the whole body of each mouse and from two different *in vivo* breathing conditions in each case (see Sec. III-A-1). We observe close consistency in the structure of the two covariance matrices. A third mouse dataset was used for investigating differences in the structure of the covariance matrix per anatomical slice and physiological condition. Fig. 2(c) presents the Frobenius norm distance between the normalized global covariance matrix $\mathbf{G}_{GL}^1/\|\mathbf{G}_{GL}^1\|$ and the normalized sample covariance matrices $\mathbf{G}_s^k/\|\mathbf{G}_s^k\|$ corresponding to different anatomical slices and physiological conditions of mouse 3. Normalization is performed to consider changes in the structure rather than the intensity of the covariance matrix. The covariance matrices of mouse 3 corresponding to the median and the maximum norm distance are presented in Fig. 2(d), (e), respectively. Despite the dramatic differences in anatomy and tissue physiology we observe that the covariance matrix of background-tissue retains a rather consistent structure, indicating the possibility of using training data for covariance estimation.

While \mathbf{G}_{GL} is uncontaminated, its deviation from the sample covariance matrix of the MSOT image under test may lead to ill background modeling. An alternative quasi-local covariance matrix can be computed through the combination of the global and the sample covariance matrix of the MSOT image under test through the following formula:

$$\mathbf{G}_{QL} = \mathbf{U}_{GL} (\text{diag}(\mathbf{U}_{GL}^T \mathbf{G}_s \mathbf{U}_{GL})) \mathbf{U}_{GL}^T, \quad (8)$$

In (8) \mathbf{U}_{GL} is the matrix containing the eigenvectors of \mathbf{G}_{GL} and \mathbf{G}_s is the sample covariance matrix of the multispectral image under test. *diag* indicates the operator that retains the diagonal elements of the matrix while setting the rest to zero.

\mathbf{G}_{QL} is uncontaminated by the SOI, as its eigenvectors are not influenced by the target spectrum [13], while it further preserves characteristics of the sample covariance of the data under test. We note that similar covariance estimators have been previously proposed in literature for cases of ill covariance computation due to limited samples [28].

2) *Covariance shrinkage*

\mathbf{G}_{QL} may still deviate from \mathbf{G}_s , possibly leading to a reduced detection performance when the data are uncontaminated. Ideally the estimated covariance matrix would vary between \mathbf{G}_s and \mathbf{G}_{QL} depending on the level of SOI contamination. To achieve this we employ the covariance shrinkage scheme:

$$\mathbf{G}_{est} = (1-a)\mathbf{G}_s + a\mathbf{G}_{QL}, \quad (9)$$

where the shrinkage parameter a can vary between 0 (optimal in the case of uncontaminated data) and 1 (optimal in the case of highly contaminated data) adapting to the level of covariance contamination.

The level of contamination could be theoretically estimated through the distance between the sample covariance matrix and the true background covariance matrix, $\|\mathbf{G}_s - \mathbf{G}_{true}\|_F$. Assuming that the unknown, true covariance matrix lies within a dictionary \mathbf{D}_G of uncontaminated covariance matrices stemming from training MSOT images, the following metric can offer an insight on the level of contamination of the sample covariance matrix:

$$m_c = \underset{\mathbf{G}_{true} \in \mathbf{D}_G}{\text{Min}} \left\| \frac{\mathbf{G}_s}{\|\mathbf{G}_s\|_F} - \frac{\mathbf{G}_{true}}{\|\mathbf{G}_{true}\|_F} \right\|_F \quad (10)$$

By means of cross-validation in simulated data, we empirically concluded to the following rule for estimating the shrinkage level using the contamination metric m_c :

$$a = \min\{1, \kappa \sqrt{m_c}\} \quad (11)$$

The empirical rule of Eq. (11) and the selection of the parameter $\kappa=4$ is explained in (V-B-2). The value of parameter κ was retained constant in all evaluations.

C. *Robust statistical detection framework*

The hereby proposed statistical detection framework combines the EC-GLRT detector of IV-A with the covariance matrix estimation scheme of IV-B.

In a first step the sample covariance matrix \mathbf{G}_s of the MSOT image under test is estimated using all available pixels and the level of contamination is assessed through metric m_c of Eq. (10). The covariance matrix dictionary \mathbf{D}_G is computed from the archive of agent-free MSOT images corresponding to the experimental data of Sec. III-A-1. We note that in the cross-validation presented in Sec. V, all covariance matrices corresponding to the mouse under test were excluded from the dictionary \mathbf{D}_G , which typically includes ~200 different covariance matrices. Using the metric m_c , the loading level a is computed according to Eq. (11). The covariance matrix \mathbf{G}_{est} is estimated through Eq. (9) using the \mathbf{G}_{QL} covariance matrix of Eq. (8). The \mathbf{G}_{GL} covariance matrix is estimated from all available agent-free MSOT data of Sec. III-A. Similarly to the case of \mathbf{D}_G , in the cross-validation presented in Sec. V, all MSOT data corresponding to the mouse under test were excluded from the computation of \mathbf{G}_{GL} .

In a last step, the final detection result is computed. In this step, the MSOT image is initially analysed using the AMF

formula of Eq. (3) using \mathbf{G}_{est} and the detection result $D_{\text{AMF}}(\mathbf{x}_i)$ is compared to a predefined threshold T_{AMF} that is associated with a small false alarm probability (T_{AMF} was selected such that $|\{i \in A : D(\mathbf{x}_i) > T_{\text{AMF}}\}| = 0.00002|A|$, where A is the complete set of pixels of the training dataset). If $D_{\text{AMF}}(\mathbf{x}_i) < T_{\text{AMF}} \forall \mathbf{x}_i$, the AMF has not detected a strong target in the image. In this case, the parameter ν is computed according to Eq. (7) and in the following $D_{\text{EC-GLRT}}(\mathbf{x}_i)$ is computed according to Eq. (6) using \mathbf{G}_{est} and returned as the output of the detection scheme. If $D_{\text{AMF}}(\mathbf{x}_i) > T_{\text{AMF}}$, the AMF has recognized a strong target. In this case, $D_{\text{AMF}}(\mathbf{x}_i)$ is returned as the output of the detection scheme. This latter test is performed for facilitating a better target visualization, as, in the case of strong targets EC-GLRT does not offer optimal visualization due to its intensity normalization property. Algorithm 1 summarizes the hereby proposed detection scheme.

Algorithm 1: Robust statistical detection framework

1. Compute the sample mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{G}_s of the MSOT image \mathbf{x} under test:

$$\boldsymbol{\mu} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i, \mathbf{G}_s = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T$$

2. Compute \mathbf{G}_{GL} as the sample covariance matrix over all training data. Compute \mathbf{G}_{QL} from \mathbf{G}_{GL} and \mathbf{G}_s according to (8). Compute the covariance matrix dictionary \mathbf{D}_G from each available training MSOT image independently.
3. Estimate the covariance matrix as in:

$$\mathbf{G}_{\text{est}} = (1-\alpha)\mathbf{G}_s + \alpha\mathbf{G}_{\text{QL}}$$

where α is computed from (11) with respect to the metric m_c of (10).

4. If $D_{\text{AMF}}(\mathbf{x}_i) < T_{\text{AMF}} \forall i$, $\hat{\nu} = \text{Arg min}_{\nu} M_{\text{exc}}(\mathbf{m}, F'_{L,\nu})$,
else $\hat{\nu} = \infty$.

5. Return: $D(\mathbf{x}_i) = \frac{(v-1)}{(v-2) + (\mathbf{x}_i - \boldsymbol{\mu})^T \mathbf{G}_{\text{est}}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})} \frac{(\mathbf{s}^T \mathbf{G}_{\text{est}}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}))^2}{\mathbf{s}^T \mathbf{G}_{\text{est}}^{-1} \mathbf{s}}$

V. RESULTS

A. Performance of EC-GLRT

Figs. 3(a), (b) present a quantitative performance comparison, in terms of MDAA, of AMF vs. EC-GLRT as evaluated using the target implantation framework of Sec. III-A-2. Two different spectral targets are considered, namely ICG (Fig. 3(a)) and IntegriSense750 (Fig. 3(b)). The statistical evaluation was performed using target implantation on all background MSOT images of the imaging dataset (Sec. III-A-1), and 6 different target positions per image. The six grey error-bars correspond to statistics stemming from the three individual mice imaged at two different physiological conditions each. The black error-bars correspond to statistics stemming from all data. The bar height corresponds to the mean MDAA and the error-bar to the standard deviation. According to the quantitative evaluation, EC-GLRT allows for detecting the implanted targets at 20%-30% lower simulated absorbance, therefore offering a statistical performance enhancement that is consistent for all mice and for both implanted targets. Examples of three representative detection cases where the detection result is overlaid onto the

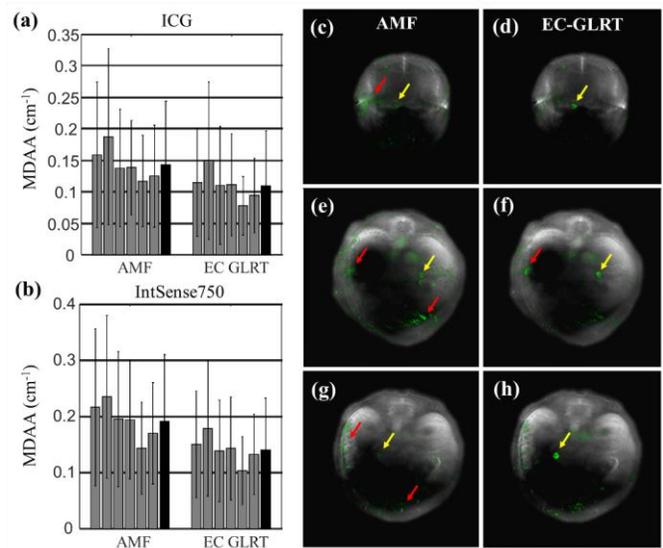


Figure 3. (a, b) Quantitative comparison of AMF and EC-GLRT in terms of MDAA (see Sec. III-A-2) for the cases of ICG (a) and Integrisense750 implanted targets (b). The bar height corresponds to the mean MDAA and the errorbar to the standard deviation over all images employed for evaluation. The six grey errorbars correspond to three different mice imaged at two different physiological conditions each (Sec. III-A-1). The black errorbars correspond to statistics stemming from all data. EC-GLRT with appropriate selection of parameter ν offered consistently a statistically enhanced performance. (c-h) Examples of detection result using AMF (c, e, g) and EC-GLRT (d, f, h) for the same target intensity. The detection result is overlaid onto the anatomical optoacoustic image with green pseudocolor. The yellow arrows point the target positions and the red arrows point the false positives.

anatomical image with green pseudo-color are presented in Fig. 3(c)-(h). The first column (Fig. 3(c), (e), (g)) corresponds to AMF, while the second column (Fig. 3(d), (f), (h)) corresponds to EC-GLRT. For the same target intensity, EC-GLRT resolved the targets (yellow arrows) with less false positives (red arrows) as compared to AMF.

Fig. 4 presents examples of AMF vs EC-GLRT comparison in the case of purely experimental data. Fig. 4(a) corresponds to a case where fluorescently labelled cells have been introduced in the left and right brain hemispheres of a mouse *post mortem* (Sec. III-B-3). Fig. 4(b) corresponds to a mouse brain containing a tumor that is expressing a near infrared fluorescent protein (Sec. III-B-2). Fig. 4(c) corresponds to the dataset described in Sec. III-B-1, where a capillary tube, containing AF750 at an optical density of 0.53 OD has been introduced within tissue. In each case yellow arrows indicate the position of the fluorescent target. In all three cases the detection performance of AMF (first column) is compromised by false positive detection artifacts. Conversely, EC-GLRT (second column) achieves detection of the molecular target with considerably reduced false positives. In each case the detection result is overlaid onto the anatomical image with green pseudocolor (left) and is also presented independently (right) for facilitating a straightforward comparison.

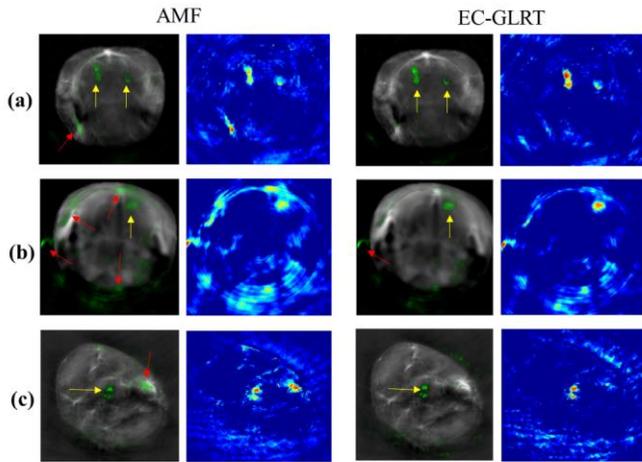


Figure 4 Comparison of AMF and EC-GLRT in purely experimental data. (a) Mouse brain with two lesions of DiR labelled cells. (b) Mouse brain with a tumor expressing iRFP fluorescent protein. (c) Mouse abdomen with a fluorescent target corresponding to a capillary tube containing AF750. In each case, the detection result is presented standalone on the right side and overlaid onto the anatomical image with green pseudocolor on the left side. Yellow arrows indicate the position of the targets, while red arrows indicate false positives. The standalone detection images corresponding to AMF are presented in square root for facilitating a better visual comparison.

B. Effects of covariance contamination and quasi-local covariance shrinkage

1) Effects of covariance contamination

Fig. 5 presents the effects of covariance contamination on the detection performance of AMF. A statistical evaluation of the detection performance of AMF, in terms of percentage of detected pixels, was performed through the implantation of targets of varying sizes, as described in Sec. III-A-3. Statistics on the percentage of detected pixels (y axis) for different target sizes (colors) and different target intensities (x axis) are presented in Fig. 5(a). Small targets of 256 pixels (green errorbars) appear fully detectable above a certain simulated agent absorbance. Conversely, in the case of larger targets of 2048 pixels the detection performance is compromised when these targets appear in high intensities, as an effect of covariance contamination. The effects of covariance contamination are more evident in the case of the largest implanted targets (4064 pixels), where the targets are no longer fully detectable at any simulated target intensity. Fig. 5(b)-(d) present the detection result of AMF in the case of a large implanted target for three different simulated intensities. For comparison purposes, Fig. 5(e)-(g) present the ideal detection result of AMF where the uncontaminated background covariance matrix is used instead. Evidently, due to the covariance contamination caused by the large target size, parts of the target are not detectable by AMF at any simulated target intensity. The percentage of detected pixels according to the quantitative evaluation is presented in the lower left part of the individual images.

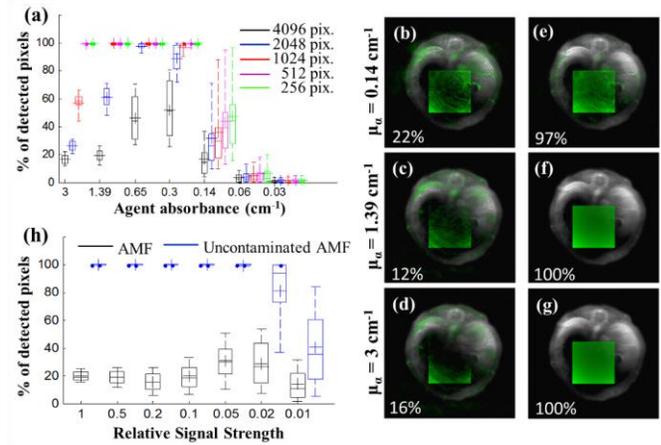


Figure 5. Effects of covariance contamination on the detection performance of AMF. (a) Statistical evaluation of AMF detection performance in terms of the % of detected pixels (y-axis) for different target sizes (colors) and different target intensities (x-axis), using the target implantation of Sec. III-A-3. Statistics are derived from target implantation on 6 different MSOT images and 6 different target positions per MSOT image. (b-d) Examples of the AMF detection performance in the case of a large target implanted at three different intensities. (e-g) Detection result of uncontaminated AMF for the same targets. (h) Statistical evaluation of the detection performance of standard (black) and uncontaminated AMF (blue) in the case of large targets (4096 pixels) as a function of relative signal strength. Statistics are derived from target implantation on all MSOT images of the experimental dataset (Sec. III-A-1).

A more straightforward approach for assessing the performance degradation due to covariance contamination is by investigating the detection performance as a function of the relative signal strength (RSS), i.e. $\|T(\mathbf{r}, \lambda)\|_2 / \|B(\mathbf{r}, \lambda)\|_2$, where $T(\mathbf{r}, \lambda)$ represents the implanted target and $B(\mathbf{r}, \lambda)$ represents the background. Fig. 5(h) presents the detection performance of AMF in the case of large targets (4064 pixels) as a function of RSS (black errorbars). The performance of uncontaminated AMF (blue errorbars) is also presented for comparison. Statistics correspond to targets implanted on all MSOT images of the experimental dataset (Sec. III-A-1).

2. Performance of quasi-local covariance shrinkage

Fig. 6 presents the performance of the proposed quasi-local covariance shrinkage scheme (QL shr. AMF; see Supp. Table I, Appendix). Fig. 6(a) presents the values of $\sqrt{m_c}$ (Eq. (10)) obtained in the case of large ICG implanted targets of 4096 pixels as a function of the RSS. The three different color-bars correspond to wavelength sampling of 21, 10 and 7 equally spaced wavelengths spanning the range of 700-900 nm. The metric m_c appears capable of identifying high levels of covariance contamination (RSS \geq 20%), but there is considerable overlap in its values in low contamination levels (RSS $<$ 20%). The empirical rule of Eq. (11) and the value of parameter κ were selected in such a way so that $\alpha=1$ when the value of m_c indicates a substantial level of contamination (RSS \geq 20%). Due to the overlap of m_c values in the case of low or no contamination, α is typically larger than zero even in uncontaminated cases. The values obtained by m_c under

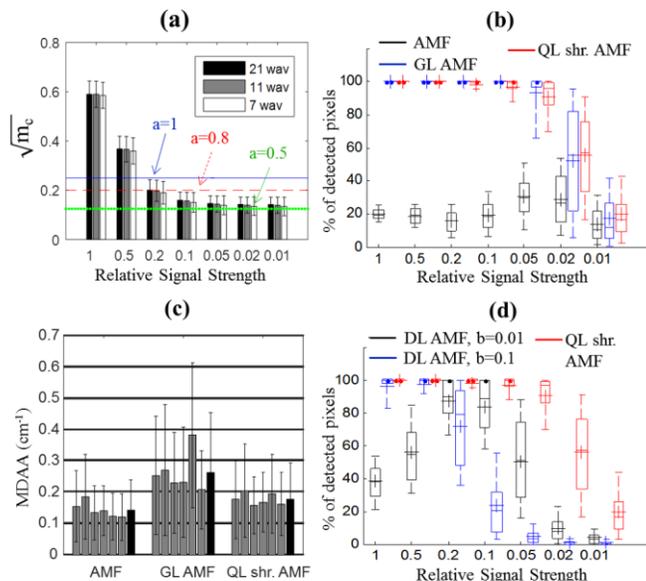


Figure 6. Evaluation of the QL shrinkage approach. (a) Values obtained by $\sqrt{m_c}$ (Eq. (10)) under different levels of RSS (x-axis) and under different wavelength sampling approaches (color-bars). The values obtained by α [Eq. (11)] for three different $\sqrt{m_c}$ values are presented with horizontal lines. (b) Performance comparison of AMF, GL AMF, and QL shr. AMF in terms of % of detected pixels in the case of large targets (4098 pixels). Statistics correspond to ICG targets implanted on all background MSOT images. (c) Performance comparison of AMF, GL AMF, and QL shr. AMF in terms of MDAA for the case of small ICG implanted targets. The six grey errorbars correspond to three different mice imaged at two different physiological conditions each. The black errorbars correspond to statistics stemming from all data. (d) Comparison of the detection performance of QL shrinkage (red) and diagonal loading using two different constant loading levels, i.e. $b=0.1$ (blue) and $b=0.01$ (black). Statistics in (a)-(d) correspond to target implantation on all experimental data of Sec. III-A-1.

different levels of contamination are relatively unaffected by different wavelength sampling strategies.

Fig. 6(b) presents a quantitative performance comparison of GL AMF (Supp. Table I, Appendix) and QL shr. AMF in the case of large ICG implanted targets of 4096 pixels (high contamination), in terms of % of detected pixels (y-axis) and as a function of RSS (x-axis). The performance degradation of AMF, caused by covariance contamination (black error-bars in Fig. 6(b)), is mitigated by the use of an uncontaminated covariance matrix. In the case of large implanted targets, both GL AMF (blue) and QL shr. AMF (red) offer a comparable performance, successfully mitigating the effects of covariance contamination.

Fig. 6(c) presents a quantitative performance comparison in the case of small ICG targets (uncontaminated case), in terms of MDAA (Sec. III-A-2). In the case of small targets GL AMF offers worse detection performance as compared to AMF, increasing the MDAA from 1.5 to 3 times. This performance degradation is expected since in the uncontaminated case, the sample covariance matrix offers better modeling abilities than the global one. QL shr. AMF offers an improved performance

with respect to GL AMF, which is only 1.1-1.6 times worse than the one of AMF.

Finally Fig. 6(d) presents a performance comparison of the proposed quasi-local covariance shrinkage approach (QL shr. AMF) and the previously considered diagonal loading scheme (DL AMF; Supp. Table I, Appendix) [12]. A high level of diagonal loading ($b=0.1$) offers good performance in the case of high target intensities but a considerably reduced performance in the case of low target intensities (blue errorbars). Conversely a lower level of diagonal loading ($b=0.01$) offers reduced detection performance in the case of high target intensities (black errorbars). Evidently, QL shr. AMF (red errorbars) offers a substantially enhanced performance as compared to DL AMF.

C. Evaluation of statistical detection framework over existing approaches

Fig. 7 compares the performance of the proposed robust statistical detection framework (RSDF) to the one of AMF and to least-squares (LS) spectral fitting [17] which have been previously used in MSOT molecular imaging. In the case of LS fitting the spectra of oxygenated and deoxygenated hemoglobin were used for modeling the tissue background. Fig. 7(a) presents a quantitative comparison in the case of small targets, using the target implantation framework of Sec. III-A-2. According to the simulations, LS fitting offers, on average, 5 times reduced detection performance as compared to AMF and RSDF. AMF and RSDF offer comparable performance. Statistics correspond to all experimental data of Sec. III-A-1. Fig. 7(b) presents the detection results of LS fitting, AMF and RSDF for a representative simulation where the small ICG target has been implanted onto the experimental MSOT image at two different simulated intensities. The target position is indicated by the yellow arrows while red arrows indicate false positives.

Fig. 7(c) presents a comparison of LS fitting, AMF and RSDF in the case of large targets (4096 pixels), in terms of % of detected pixels. LS fitting offers good detection performance in the case of high target intensities ($\mu_a > 0.65 \text{ cm}^{-1}$) but a considerably reduced performance in the case of low target intensities. AMF offers reduced detection ability in all simulated target intensities due to the effects of covariance contamination. Conversely, RSDF offers a considerably enhanced performance as compared to AMF and LS fitting. Fig. 7(d) presents the detection results of LS fitting, AMF and RSDF for a representative simulation where the ICG target has been implanted at two different intensities.

Fig. 8 presents detection examples stemming from purely experimental data. The first column corresponds to LS fitting, the second to AMF and the third to RSDF. In the latter case, the values of α and ν are further presented alongside the images. Fig. 8(a) corresponds to brain images with DiR labelled macrophages (Sec. III-B-3). The position of the cells (yellow arrows) is accurately detected by AMF and RSDF but not by LS fitting which yields substantially stronger false positive detection artifacts. Fig. 8(b-c) corresponds to the dataset described in Sec. III-B-1 where a capillary tube

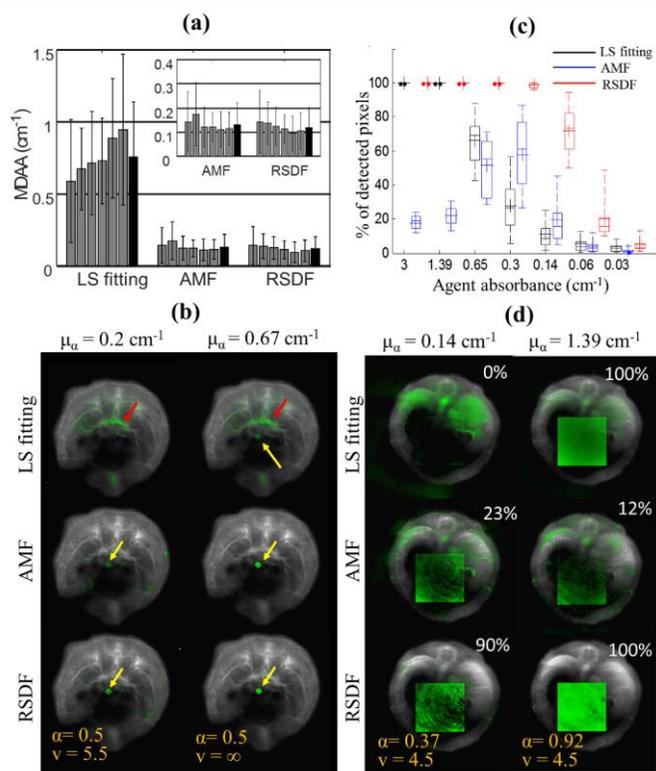


Figure 7. Quantitative comparison of LS fitting, AMF and RSDF in simulated data. (a) Performance comparison in terms of MDAA for the case of small ICG targets (Sec. III-A-2). The six grey errorbars correspond to three different mice imaged at two different physiological conditions each (Sec. III-A-1). The black errorbars correspond to statistics stemming from all data. (b) Detection examples of LS fitting, AMF and RSDF in two cases where a small ICG target has been implanted in a low (left) and a high (right) intensity. Yellow arrows indicate the position of the target, while red arrows indicate false positives. The parameters α and ν of RSDF are further presented in each case. (c) Performance comparison in terms of % of detected pixels for the case of large (4096 pixel) targets (Sec. III-A-3). Statistics are derived from target implantation on 6 different MSOT images and 6 different target positions per image. (d) Detection examples of LS fitting, AMF and RSDF in two cases where a large ICG target has been implanted at a low (left) and a high (right) intensity. The percentage of detected pixels is presented in each case in the upper right part of the image.

containing AF750 (AF790, respectively) at an optical density of approx. 0.25 OD has been introduced in deep tissue. The location of the fluorochrome (yellow arrows) is detectable only by RSDF (c), or by both AMF and RSDF (b), but not by LS fitting which yields false positives at higher intensities. Fig. 8(d) presents an MSOT image of the same experiment, but in this case the capillary tube has been filled with a high concentration of AF790, corresponding to 6.6 OD. In this case LS fitting and RSDF accurately detect the position of the target, while the performance of AMF is considerably compromised due to the effect of covariance contamination. Fig. 8(e) corresponds to the dataset described in Sec. III-B-5, where the accumulation of AF750 in the bladder is monitored. Again, the detection performance of AMF is compromised due to the large target size, while LS fitting and RSDF accurately detect the fluorochrome bio-distribution in the bladder (dashed line). Finally Fig. 8(f) corresponds to the dataset describe in

Sec. III-B-4 where MSOT monitors the accumulation of ICG in the liver. Both RSDF and LS fitting resolve an extensive distribution of ICG in the area of the liver, whereas AMF fails due to the effects of covariance contamination. The detection results of Fig. 8 demonstrate that RSDF combines high molecular imaging sensitivity (Fig. 8(a)-(c)) with a robust performance that is not affected by the target size or intensity (Fig. 8(d)-(f)).

VI. DISCUSSION AND CONCLUSIONS

Spectral analysis methods play a fundamental role in the ability to extract valuable molecular information from multispectral optoacoustic images. The ultimate goal of spectral unmixing in MSOT imaging is the quantification of the concentrations of all absorbing molecules within tissue. Despite substantial prior work, the quantification of the absorbers' concentrations in experimental *in vivo* images is considered still an open challenge [16].

This work considers the problem of multispectral detection of extrinsic molecular agents in MSOT imaging. As opposed to spectral unmixing, the spectral detection problem seeks to identify the position and area occupied by molecular targets within the tissue [29] and it has a direct application on MSOT molecular imaging. While the multispectral detection problem can be formulated as a binary classification problem, it has some characteristic properties that largely affect the development of appropriate algorithms. *First*, while sufficient data are typically available for accurately training the background class, the sparsity of the targets typically hinders the successful training of the target class. *Second*, the minimization of the misclassification rate is not a good performance metric when the goal is to detect targets that are considerably smaller than the background [5, 6]. For these reasons, practical multispectral detectors typically follow a GLRT approach which seeks to maximize the probability of detection while retaining the probability of false alarm under a predefined threshold (Neyman Pearson criterion) [7]. Moreover, most practical GLRT detectors do not train a target class but assume a linear model of a known target spectrum superimposed on a background that is modeled statistically [5].

GLRT statistical sub-pixel detection has been shown to offer an enhancement in the molecular imaging sensitivity of MSOT [4]. However, common statistical sub-pixel detection algorithms are not optimally suited to the particular characteristics of MSOT molecular imaging. In this work we introduced a robust statistical detection scheme designed according to the characteristics of MSOT molecular imaging. We studied both the aspects of background statistical modeling and covariance matrix estimation that are key parameters for the design of statistical sub-pixel detectors [19]. This work did not study the effects of spectral mismatch in the detection performance, which will be assessed in future work.

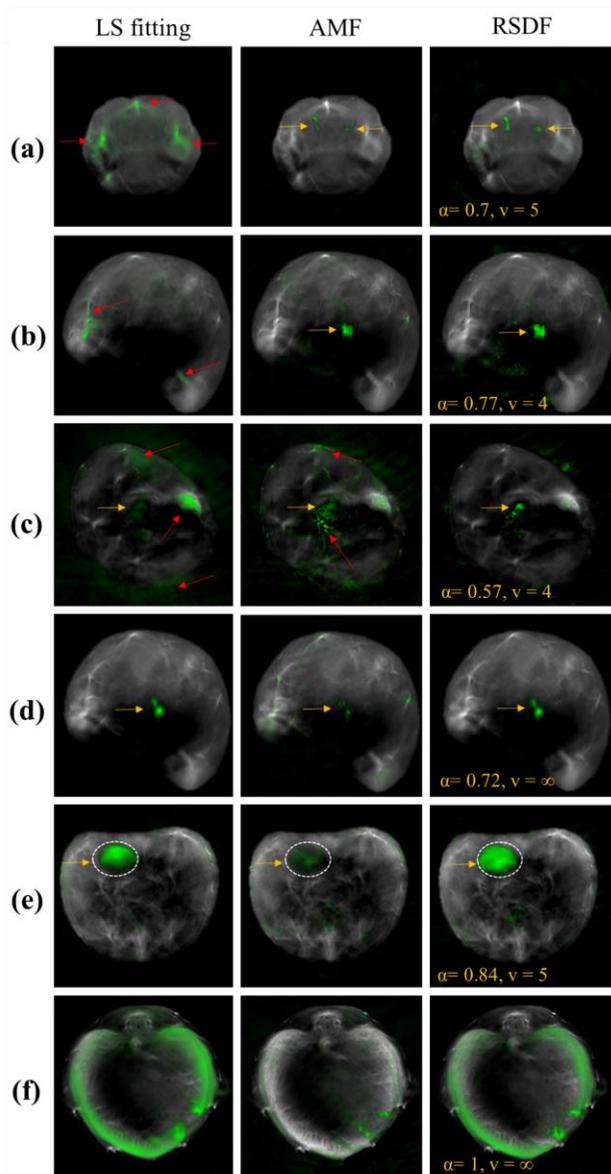


Figure 8. Comparison of the detection result of LS fitting (left column), AMF (middle column) and RSDF (right column) in the case of purely experimental data. In the latter case the parameters α and v are presented within the image. (a) DiR labelled macrophages introduced in the left and right brain hemisphere (arrows). (b-d) MSOT images with an insertion of AF750 at an optical density of approx. 0.25OD (b), AF790 at a similar optical density (c) and AF790 at 6.6 OD (d) introduced within tissue (arrows), as described in Sec. III-B-1. (e) MSOT images of the abdominal area after systemic injection of 1.2 nmoles of AF750. AF750 accumulates in the area of the bladder (dashed line). (f) MSOT images of the liver after systemic injection 40 nmoles of ICG. In all cases the detection result in overlaid onto the anatomical image with green pseudocolor.

Through a statistical characterization of experimental background-tissue MSOT data we found that the t-distribution appears more accurate in modeling both the main body and the tails of the MSOT data. This observation led to the utilization of a data-adaptive EC-GLRT detector, which in turn offered an enhancement in the detection performance observed both in simulated and experimental data (Fig. 3, 4).

The most significant pitfall of statistical sub-pixel detection in the context of MSOT is the effect of covariance contamination in the cases of extensive presence of the molecular target. Using target implantation simulations of varying target sizes we quantitatively assessed this effect and found that the performance of AMF degrades substantially in the case of large molecular targets (Fig. 5). This effect was also demonstrated in experimental MSOT data (Fig. 8(e), (f)).

For mitigating the effect of covariance contamination we introduced a covariance estimation approach that exploits an archive of training background-tissue MSOT images. Through the comparison of covariance matrices stemming from different anatomical regions and physiological conditions we observed a consistency in the structure, which indicated the possibility of using training data for covariance estimation. However, the simplistic exchange of the sample covariance matrix with a global covariance matrix computed from training data, caused a substantial negative effect on the detection performance in the case of low contamination (Fig. 6), reducing the detection performance up to 3 times. For overcoming this effect we designed a covariance estimation approach, where a quasi-local uncontaminated covariance matrix is appropriately merged with the sample covariance matrix based on the estimated level of the covariance contamination. The proposed covariance estimation scheme offers robust performance that is not affected by covariance contamination, while also retaining good performance in the uncontaminated case. The proposed covariance estimation scheme considerably outperformed diagonal loading, a method previously proposed for mitigating the effects of covariance contamination in hyperspectral remote sensing [12].

The performance of the proposed robust statistical detection framework was evaluated statistically using extensive target implantation simulations and considering targets of different sizes and intensities. The method was found to considerably outperform adaptive matched filter in the case of large targets and the LS fitting approach in the case of weakly absorbing targets. Importantly, RSDF offers an automatic and universal performance independently of the target size or intensity. A number of purely experimental imaging studies further confirmed the direct applicability of the method in experimental molecular imaging data.

ACKNOWLEDGMENT

The authors would like to thank Uwe Klemm and Sarah Glasl for assisting in animal handling and image acquisition.

REFERENCES

- [1] A. Taruttis and V. Ntziachristos, "Advances in real-time multispectral optoacoustic imaging and its applications," *Nat. Photonics*, vol. 9, no. 4, pp. 219-227, 2015.
- [2] V. Ntziachristos and D. Razansky, "Molecular imaging by means of multispectral optoacoustic tomography (MSOT)," *Chem. Rev.*, vol. 110, no. 5, pp. 2783-2794, 2010.
- [3] N. C. Deliolanis, A. Ale, S. Morscher, N. C. Burton, K. Schaefer, K. Radrich, *et al.*, "Deep-Tissue Reporter-Gene Imaging with Fluorescence and Optoacoustic Tomography: A Performance Overview," *Mol. Imaging Biol.*, vol. 16, no. 5, pp. 652-660, 2014.
- [4] S. Tzoumas, N. Deliolanis, S. Morscher, and V. Ntziachristos, "Unmixing Molecular Agents From Absorbing Tissue in Multispectral Optoacoustic Tomography," *IEEE Trans. Med. Imag.*, vol. 33, no. 1, pp. 48-60, 2014.

- [5] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29-43, 2002.
- [6] N. M. Nasrabadi, "Hyperspectral target detection: An overview of current and future challenges," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 34-44, 2014.
- [7] S. M. Kay, *Fundamentals of statistical signal processing, Vol. II: Detection Theory*. Upper Saddle River, NJ : Prentice Hall, 1998.
- [8] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208-216, 1992.
- [9] S. Tzoumas, A. Nunes, N. C. Deliolanis, and V. Ntziachristos, "Effects of multispectral excitation on the sensitivity of molecular optoacoustic imaging," *J. Biophotonics*, vol. 8, no. 8, pp. 629-637, 2014.
- [10] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397-401, 1988.
- [11] J. Theiler and B. R. Foy, "Effect of signal contamination in matched-filter detection of the signal on a cluttered background," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 98-102, 2006.
- [12] S. Niu, S. E. Golowich, V. K. Ingle, and D. G. Manolakis, "Implications and mitigation of model mismatch and covariance contamination for hyperspectral chemical agent detection," *Opt. Eng.*, vol. 52, no. 2, pp. 026202-026202, 2013.
- [13] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing*. Norwood: Artech House, 2005.
- [14] B. Cox, S. Arridge, and P. Beard, "Estimating chromophore distributions from multiwavelength photoacoustic images," *JOSA A*, vol. 26, no. 2, pp. 443-455, 2009.
- [15] J. Laufer, B. Cox, E. Zhang, and P. Beard, "Quantitative determination of chromophore concentrations from 2D photoacoustic images using a nonlinear model-based inversion scheme," *Appl. Opt.*, vol. 49, no. 8, pp. 1219-1233, 2010.
- [16] B. Cox, J. G. Laufer, S. R. Arridge, and P. C. Beard, "Quantitative spectroscopic photoacoustic imaging: a review," *J. Biomed. Opt.*, vol. 17, no. 6, pp. 0612021-0612022, 2012.
- [17] M. Li, J. Oh, X. Xie, G. Ku, W. Wang, C. Li, *et al.*, "Simultaneous molecular and hypoxia imaging of brain tumors in vivo using spectroscopic photoacoustic tomography," *Proc. IEEE*, vol. 96, no. 3, pp. 481-489, 2008.
- [18] E. J. Kelly, "An Adaptive Detection Algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 2, pp. 115-127, 1986.
- [19] S. Matteoli, M. Diani, and J. Theiler, "An Overview of Background Modeling for Detection of Targets and Anomalies in Hyperspectral Remotely Sensed Imagery," *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, vol. 7, no. 6, pp. 2317-2336, 2014.
- [20] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1365-1376, 1987.
- [21] A. Buehler, E. Herzog, D. Razansky, and V. Ntziachristos, "Video rate optoacoustic tomography of mouse kidney perfusion," *Opt. Lett.*, vol. 35, no. 14, pp. 2475-2477, 2010.
- [22] A. Rosenthal, D. Razansky, and V. Ntziachristos, "Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1275-85, Jun 2010.
- [23] S. Tzoumas, A. Zaremba, U. Klemm, A. Nunes, K. Schaefer, and V. Ntziachristos, "Immune cell imaging using multi-spectral optoacoustic tomography," *Opt Lett*, vol. 39, no. 12, pp. 3523-3526, 2014.
- [24] A. Taruttis, S. Morscher, N. C. Burton, D. Razansky, and V. Ntziachristos, "Fast multispectral optoacoustic tomography (MSOT) for dynamic imaging of pharmacokinetics and biodistribution in multiple organs," *PLoS One*, vol. 7, no. 1, p. e30491, 2012.
- [25] D. G. Manolakis, D. Marden, J. P. Kerekes, and G. A. Shaw, "Statistics of hyperspectral imaging data," in *Aerospace/Defense Sensing, Simulation, and Controls*, 2001, pp. 308-316.
- [26] D. B. Marden and D. G. Manolakis, "Using elliptically contoured distributions to model hyperspectral imaging data and generate statistically similar synthetic data," in *Defense and Security*, 2004, pp. 558-572.
- [27] J. Theiler and B. R. Foy, "EC-GLRT: Detecting weak plumes in non-Gaussian hyperspectral clutter using an elliptically-contoured generalized likelihood ratio test," in *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, 2008, pp. 1-221-I-224.
- [28] N. Gorelik, D. Blumberg, S. R. Rotman, and D. Borghys, "Nonsingular approximations for a singular covariance matrix," in *SPIE Defense, Security, and Sensing*, 2012, pp. 839021-839021-8.
- [29] P. Mohajerani, S. Tzoumas, A. Rosenthal, and V. Ntziachristos, "Optical and Optoacoustic Model-Based Tomography: Theory and current challenges for deep tissue imaging of optical contrast," *IEEE Signal Process. Mag.*, vol. 32, no. 1, pp. 88-100, 2015.