A time-domain wavelet-based approach for fluorescence diffuse optical tomography

Nicolas Ducros[a)]
Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci, 20133 Milano, Italy; CREATIS, CNRS 5220, 69621 Villeurbanne, France; INSERM U630, 69621 Villeurbanne, France; Université de Lyon I, 69621 Villeurbanne, France; INSA Lyon, 69621 Villeurbanne, France; and CEA-LETI MINATEC, 38054 Grenoble, France

Anabela Da Silva
Institut Fresnel, CNRS, 13397 Marseille, France; École Centrale de Marseille, 13397 Marseille, France; and Université de Marseille, 13397 Marseille, France

Jean-Marc Dinten
CEA-LETI MINATEC, 38054 Grenoble, France

Chandra Sekhar Seelamantula
Department of Electrical Engineering, Indian Institute of Science, Bangalore, India

Michael Unser
Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Françoise Peyrin
CREATIS, CNRS 5220, 69621 Villeurbanne Cedex, France; INSERM U630, 69621 Villeurbanne Cedex, France; Université de Lyon I, 69621 Villeurbanne Cedex, France; and INSA Lyon, 69621 Villeurbanne Cedex, France

(Received 1 February 2010; revised 28 April 2010; accepted for publication 29 April 2010; published 24 May 2010)

Purpose: In the context of fluorescence diffuse optical tomography, determining the optimal way to exploit the time-resolved information has been receiving much attention and different *features* of the time-resolved signals have been introduced. In this article, the authors revisit and generalize the notion of feature, considering the projection of the measurements onto some basis functions. This leads the authors to propose a novel approach based on the wavelet transform of the measurements.
Methods: A comparative study between the reconstructions obtained from the proposed wavelet-based approach and the reconstructions obtained from the reference temporal moments is provided. An inhomogeneous cubic medium is considered. Reconstructions are performed from synthetic measurements assuming Poisson noise statistics. In order to provide fairly comparable reconstructions, the reconstruction scheme is associated with a particular procedure for selecting the regularization parameter.
Results: In the noise-free case, the reconstruction quality is shown to be mainly driven by the number of selected features. In the presence of noise, however, the reconstruction quality depends on the type of the features. In this case, the wavelet approach is shown to outperform the moment approach. While the optimal time-resolved reconstruction quality, which is obtained considering the whole set of time samples, is recovered using only eight wavelet functions, it cannot be attained using moments. It is finally observed that the time-resolved information is of limited utility, in terms of reconstruction, when the maximum number of detected photons is lower than 10^5.
Conclusions: The wavelet approach allows for better exploiting the time-resolved information, especially when the number of detected photons is low. However, when the number of detected photons decreases below a certain threshold, the time-resolved information itself is shown to be of limited utility. © *2010 American Association of Physicists in Medicine*.
[DOI: 10.1118/1.3431571]

Key words: fluorescence diffuse optical tomography, time-resolved imaging, data features, wavelets, multiresolution

I. INTRODUCTION

Traditional approaches in fluorescence diffuse optical tomography (FDOT) can be broadly classified into three groups: (i) continuous wave (CW), (ii) frequency domain (FD), and (iii) time-resolved (TR) techniques. These three approaches differ in the modes of excitation and detection. Specifically, CW FDOT is based on the measurement of the attenuation of a steady-state excitation light.[1–5] Frequency-domain FDOT is based on the measurement of the phase and demodulation of an amplitude-modulated excitation.[6,7] TR FDOT is based on the temporal measurement of the distortion of an excitation

light pulse.[8–14] The CW and FD approaches are inexpensive and easy to develop. TR FDOT, on the other hand, is more costly and requires careful handling. It, however, allows one to separate absorption and diffusion,[15] to derive fluorescence lifetime,[16] and to better reconstruct deeply embedded markers.[8] FD FDOT, unlike CW FDOT, also offers this advantage. However, in practice, the FD measurements can only be performed for few modulation frequencies. Thus, the information content of the TR measurements, which contains all frequencies, is richer than the FD measurement one.

An important aspect is to know how to select useful information contained in the TR signals. This issue is more commonly known as the problem of selecting *data features* or *data types*. Initial developments focused on early-arriving photons of temporal signals.[9] This approach, however, suffers from low signal-to-noise ratio (SNR) and provides poor depth resolution. A more recent trend has been to exploit the global features of time-resolved signals.[10–13,17] Specifically, the Laplace transform of the TR signals[12] and above all the temporal moments of the fluorescence signals have been intensively studied and employed.[10,11,13,18,19] The moment approach is of particular interest since it allows for a physical interpretation of the features, in terms of the number of photons and their time of flight, while being easy to calculate.[20] The moment approach is the most studied one and is therefore often recognized as the gold standard. As an alternative, local features have also been proposed, with particular attention on photon counts within time windows[21,22] as well as on the photon peak value and time.[14]

While various types of data features have been introduced and are used in practical applications, few studies have compared their performances. The first contribution in this direction is by Grosenick *et al.*,[21] who compared reconstructions from photon counts within different time windows in the context of TR-DOT. Comparisons of reconstructions obtained from moments of different orders have been provided by Gao *et al.* in the context of TR DOT (Ref. 23) and by Lam *et al.*[11] in the context of TR FDOT. Riley *et al.*[14] showed that local features such as the photon peak value and time can be superior to moments in terms of noise robustness. Recently, we established that the domain of interest of the moments approach in FDOT is mainly determined by the SNR of the measurements.[19] When photonic noise is considered, the SNR of a measurement depends on the number of detected photons; the more photons detected, the better the SNR. When enough photons can be detected, the moments-based approach outperforms the classical CW approach. Below a certain threshold, however, the moments-based approach is severely limited by the low SNRs of the measurements and it is then comparable to the CW approach in terms of reconstruction quality. Since it can be experimentally difficult to collect many photons, especially when thick or highly absorbing media are considered, features allowing for exploiting the TR information at lower SNRs are desirable.

We introduce in this paper a new kind of localized analysis based on the wavelet decomposition of the TR FDOT signal. The success of wavelet algorithms for a large number of applications in biomedical imaging mainly lies in the multiresolution capabilities and the ease with which regularization criteria can be incorporated.[24] Specific applications to the field of DOT/FDOT are that of Zhu *et al.*, who demonstrated a reduction in the computational costs for the same reconstruction quality,[25,26] and Kanmani *et al.*[27] and Frassati *et al.*,[28] who employed wavelets in the forward problem of DOT and FDOT, respectively.

In this paper, we present a novel wavelet-based technique for the exploitation of the TR information; i.e., primarily as a tool to aid in the choice of the data features. We evaluate, in terms of reconstructions quality, the benefit of the proposed wavelet-based approach and provide a comparative study that shows its advantage over the temporal moments approach.

This paper is organized as follows. In Sec. II, we briefly review the different models involved in the theory of FDOT. In Sec. III, we introduce a general formalism that describes the notion of a data feature/type. The wavelet decomposition, which fits naturally in this formalism, is then introduced as well as the procedure adopted to solve the new inverse problem. Section IV contains the presentation of our numerical simulations. This presentation comprises the description of the chosen phantom, the setting of the number of detected photons, as well as the implementation of the forward model, wavelet transform, and reconstruction procedure. In Sec. V, we present the results of the numerical simulations for a specific reconstruction configuration. We also define quantitative criteria to facilitate the comparison of the reconstruction. Finally, we evaluate the influence of the number of detected photons.

I.A. Notations

We consider real-valued continuously defined functions $f(t)$, $t \in \mathbb{R}$. The time-domain convolution of the two functions f and g is written as $f*g$ and is defined as $(f*g)(t) = \int_{-\infty}^{+\infty} f(\tau) g(t-\tau) \mathrm{d}\tau$. The Fourier transform of $f(t)$ is denoted by $\hat{f}(\omega)$ and is defined as $\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t) \exp(-\mathrm{j}\omega t) \mathrm{d}t$, j being the square root of -1. The moment of order k of the function f is given by $\int f(t) t^k \mathrm{d}t$.

Matrices are denoted in upper-case bold letters, while vectors are marked with lower-case bold letters. Vectors are seen as column matrices: $\mathbf{x} = (x_1, \ldots, x_n)^T$. The notation $\|\cdot\|$ denotes the ℓ_2 norm defined by $\|\mathbf{x}\| = (\mathbf{x}^T \mathbf{x})^{1/2}$. The weighted ℓ_2 norm induced by the positive-definite matrix $\mathbf{Q}$ is denoted as $\|\mathbf{x}\|_{\mathbf{Q}} = (\mathbf{x}^T \mathbf{Q} \mathbf{x})^{1/2}$.

II. FDOT THEORY

We consider an absorbing and diffusing medium Ω that embeds fluorescence markers. Let $\partial\Omega$ be the boundary of Ω. FDOT consists of exciting the medium with a laser beam at some positions $\mathbf{s} \in \partial\Omega$ and exploiting the fluorescence light measured at some positions $\mathbf{d} \in \partial\Omega$ to infer the internal structure of the medium. As shown in Fig. 1, the underlying principle of FDOT can be explained by a three-step process.

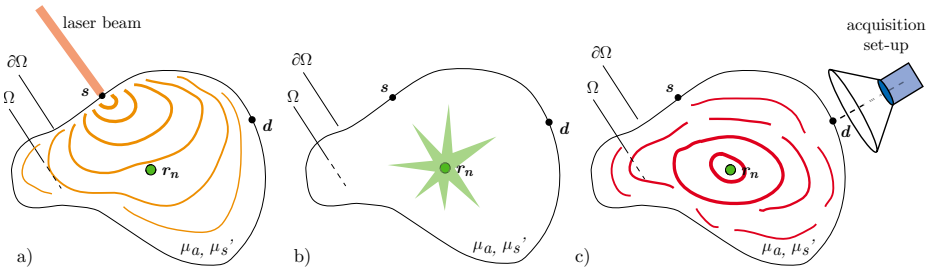

FIG. 1. Illustration of the mechanism of FDOT. (a) Excitation: The light emitted by the source at position **s** propagates through the medium. (b) Fluorescence: The fluorescent marker absorbs a fraction of the excitation light and then emits light at a higher wavelength. (c) Emission: The light emitted by the fluorescent marker at position $\mathbf{r_n}$ propagates through the medium; the fluorescent marker thus acts as a secondary source of light.

It involves light propagation at excitation wavelength λ_x, fluorescence of the markers, and propagation at fluorescence wavelength λ_f.

II.A. Underlying physical models

II.A.1. Light propagation

Light propagation within biological tissues is strongly dominated by absorption and scattering of light. Among the different models of light propagation developed for the so-called *turbid* media, the deterministic models are obtained by solving partial differential equations (see Secs. II.D–II.F of Ref. 29 for an overview of the different models). Although general models such as the telegraph equation[30] or the radiative transfer equation[31] can be used, the light propagation in FDOT is classically assumed to follow the diffusion equation. Within this framework, the photon density ϕ (photons s^{-1} cm^{-2}, refer to Sec. II.C of Ref. 29 for the definition) at position $\mathbf{r} \in \Omega$ and time t satisfies the following derivative equation:[29]

$$\left[-\nabla\left(\frac{1}{3\mu_s'(\mathbf{r})}\nabla\right)+\frac{1}{\nu}\frac{\partial}{\partial t}+\mu_a(\mathbf{r})\right]\phi_x(\mathbf{r},t)=S(\mathbf{r},t), \tag{1}$$

where μ_a (cm^{-1}) is the absorption coefficient, μ_s' (cm^{-1}) is the reduced scattering coefficient, ν (cm s^{-1}) is the speed of light within the medium, and S (photons s^{-1} cm^{-3}) is the source term. Typically, the source model consists of placing a virtual source at position $\mathbf{s}' \in \Omega$, typically chosen at $1/\mu_s'$ beneath the physical source position $\mathbf{s} \in \partial\Omega$.[32] Thus, we have

$$S(\mathbf{r},t)=N_0\delta(\mathbf{r}-\mathbf{s}')\delta(t), \tag{2}$$

where N_0 is the number of injected photons.

To complete the description of light propagation, note that boundary conditions must be incorporated into Eq. (1) (refer to Refs. 29 and 33 for details on the treatment of boundaries). As recommended in Ref. 34, the photon density is considered as the measurable quantity.

II.A.2. Fluorescence

Let us consider a fluorescent marker concentration $c(\mathbf{r})$ (μM) within an infinitesimal volume $d^3\mathbf{r}$ centered at position $\mathbf{r}$. The fluorescence marker is characterized by its fluorescence lifetime τ (ns), quantum yield η (dimensionless), and cross section γ (cm^2). The fluorescence pulse response f in both space and time (cm^{-1} s^{-1}) is given by

$$f(\mathbf{r},t)=e(t)c(\mathbf{r})\delta(\mathbf{r})\mathrm{d}^3\mathbf{r}, \tag{3}$$

where

$$e(t)=\eta\gamma/\tau\exp(-t/\tau). \tag{4}$$

More complicated multiexponential decays could also be incorporated but this level of sophistication is generally not necessary in FDOT.

II.B. Forward model

In what follows, the spatial dependence of a function is systematically indicated as a subscript. With this convention, the light measured at detection point **d** resulting from excitation at source point **s** is denoted by $u_{\mathbf{s},\mathbf{d}}(t)$. The medium Ω is discretized into N voxels of volume v_n, centered at positions $\mathbf{r_n}$, $n \in \{1,\ldots,N\}$. Considering a constant marker concentration within voxels, we have

$$u_{\mathbf{s},\mathbf{d}}(t)=\sum_{n=1}^{N} c_n v_n N_x[g_{\mathbf{s},\mathbf{r_n}} * e * g_{\mathbf{r_n},\mathbf{d}}](t), \tag{5}$$

where c_n is the local marker concentration at position $\mathbf{r_n}$, and $g_{\mathbf{s},\mathbf{r_n}}$ and $g_{\mathbf{r_n},\mathbf{d}}$ are the Green's function for Eq. (1) at position $\mathbf{r_n}$, considering the Dirac source functions located in **s** and **d**, respectively.

Considering a distribution of S source point $\{\mathbf{s}\}$ and D detector point $\{\mathbf{d}\}$, a set of $S\times D$ measurement pair $\{\mathbf{s}_j,\mathbf{d}_j\}$, $j \in \{1,2,3,\ldots,S\times D\}$ is formed. Considering any measurement on the model of Eq. (5), we obtain the following matrix system:

$$\mathbf{u}(t)=\mathbf{W}(t)\mathbf{c}. \tag{6}$$

The vector $\mathbf{u}(t)=[u_1(t),u_2(t),\ldots,u_{SD}(t)]$ is the measurement vector. The vector $\mathbf{c}=[c_1,c_2,\ldots,c_N]$ is the concentration vector. The matrix $\mathbf{W}(t)$ of size $SD\times N$, which maps the concentration vector onto the measurement vector, is referred to as the weight matrix. Its (j,n)th entry is given by $w_{j,n}(t)=v_n N_x[g_{\mathbf{s}_j,\mathbf{r}_n} * e * g_{\mathbf{r}_n,\mathbf{d}_j}](t)$. Note that the Green's functions for any other light propagation equation could have been used in Eqs. (5) and (6).

II.C. Noise model

Due to the quantum nature of light, light measurements are intrinsically corrupted by Poisson noise. Consequently, the variance of a set of light measurements equals the mean value of this set. This was observed by Selb *et al.*[35] in experimental situation. However, on the basis of CW FDOT measurements, Hyde *et al.*[36] observed the noise variance to be rather *proportional* to the mean value of the measurement. Riley *et al.*[14] found the *standard deviations* of TR FDOT measurements to be proportional their mean values. The explanation for the last two observations is unclear but has been attributed to the presence of some nonphotonic sources of noise.

In this study, the number of detected photons is hypothesized to be large enough to neglect any nonphotonic sources of noise. Subsequently, the classical Poisson noise model can be used. Moreover, since the number of detected photons is assumed to be large, the Poisson distribution is approximated by a Gaussian distribution whose variance equals the mean. Hence, the noisy version of the measurement $u_{\mathbf{s},\mathbf{d}}(t)$, denoted by $\tilde{u}$, is given by

$$\tilde{u}_{\mathbf{s},\mathbf{d}}(t) = u_{\mathbf{s},\mathbf{d}}(t) + \mathcal{N}\{\mu = 0, \sigma^2 = u_{\mathbf{s},\mathbf{d}}(t)\}, \tag{7}$$

where $\mathcal{N}(\mu,\sigma^2)$ denotes a Gaussian distribution of mean μ and variance σ^2.

III. BASIS FUNCTIONS

In this section, we refer to *data types*,[17] which are also referred to as *measurement types*[37] or *data features*[12] in literature. All these terms are used to describe the transformations applied to the TR measurements in order to reduce the redundancy, the computational cost of the inverse problem, and possibly the computational cost of the forward model. As Schweiger and Arridge[17] already suggested in their pioneering work, a good data type should possess the following properties:

- *Efficiency*: Maximize the information captured from the TR measurements in order to yield high-accuracy reconstructions, both in terms of sensitivity and resolution.
- *Sparseness*: Provide a parsimonious representation.
- *Robustness*: The quality of reconstruction must be robust and any performance degradation with respect to noise must be at an acceptable level.

Here, the concept of data types is formalized by considering the projections of the TR measurements on some TR functions, referred to as *basis functions* (BFs). Let $u(t)$ denote a TR measurement and $p(t)$ a BF. The data feature of $u(t)$ associated with $p(t)$ is defined by the dot product: $\langle u,p\rangle = \int u(t)p(t)\mathrm{d}t$. In the following, we make the distinction between data *features*, which are projection coefficients, and data *types*, which describe a type of BFs.

Note that within this BF-based framework, the classical kth order temporal moments of the measurements are easily obtained, considering a nonorthonormal basis of monomial, i.e., choosing $p(t)=t^k$, as illustrated in Fig. 2.

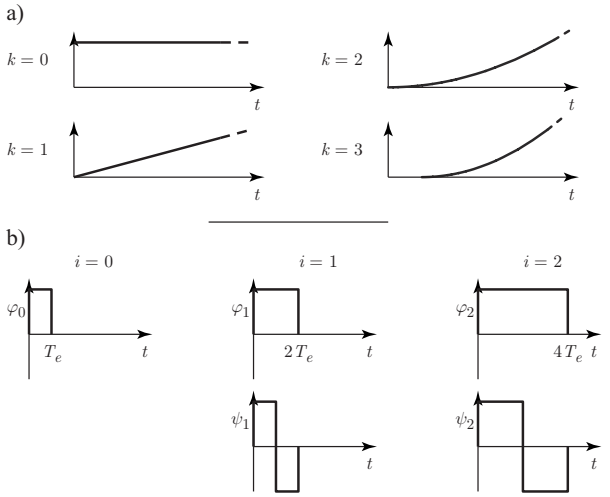

FIG. 2. Basis functions. Any data type can be conceived as the projection of the measurements on such functions. (a) BFs corresponding to the moments and (b) BFs corresponding to the Haar wavelet decomposition. Moments are nonlocalized data features, whereas the wavelet features are localized one.

III.A. Temporal wavelet transform

In this paper, we introduce a new data type, choosing the BFs as wavelet bases. The motivation for this choice is to benefit from the well-known data compression and noise reduction properties of the wavelet bases. The proposed approach results in a temporal multiresolution analysis (MRA), wherein a given function is projected onto translated and dilated versions of the so-called scaling function $\varphi(t)$ (which yields smooth approximations) and wavelet function $\psi(t)$ (which gives rise to the detail coefficients).[38]

A large number of wavelet bases are available. In this study, the Haar wavelets are considered. On top of its implementation simplicity, this choice offers a natural link between the new Haar features and the classical features based on the number of photons detected within time windows. In the Haar case, the scaling function φ is a B-spline of degree zero. The scaling function φ_i at scale i is given by

$$\varphi_i(t) = h(t) - h(t - 2^i T_e), \tag{8}$$

where $h(t)$ is the Heaviside unit-step function and T_e is the time step. Thus, the analysis is performed at a coarser scale as i increases. The finest scale is obtained for $i=0$. The wavelet ψ_i at scale i is given by

$$\psi_{i+1}(t) = \varphi_i(t) - \varphi_i(t - 2^i T_e). \tag{9}$$

The approximation coefficients at scale i are given by $a_i[k] = \langle u(t), \varphi_i(t-k2^iT_e)\rangle$ and the corresponding detail coefficients are given by $d_i[k]=\langle u(t), \psi_i(t-k2^iT_e)\rangle$.

III.B. Inverse problem within the BF framework

We consider a set of P BFs $\{p_1(t),\ldots,p_P(t)\}$. Projecting both sides of the time-resolved forward model $\mathbf{u}=\mathbf{W}\mathbf{c}$ [see Eq. (6)] on this set of BFs leads to the following *time-featured* forward model:

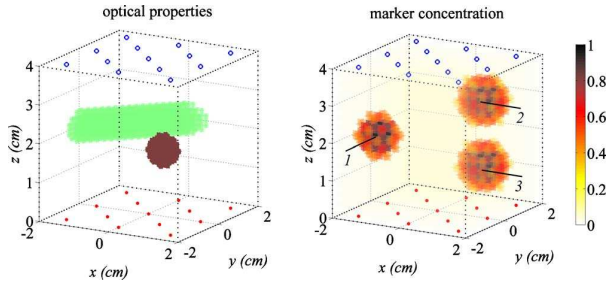

FIG. 3. Synthesized phantom. The red dots represent the source points, the blue rings the detector points, and the thin black dots the mesh nodes (for clarity, only those at the border are represented). On the left are represented the optical properties of the medium. The background values are $\mu_a = 0.2\ \text{cm}^{-1}$, $\mu_s' = 10\ \text{cm}^{-1}$, $\tau = 0.3$ ns, and $n = 1.4$. The absorbing homogeneity ($\delta\mu_a = 0.02\ \text{cm}^{-1}$) is depicted in green, and the diffusing inhomogeneity ($\delta\mu_s' = 1\ \text{cm}^{-1}$) in brown. The marker concentration distribution is shown on the right side.

$$\mathbf{y} = \mathbf{A}\mathbf{c}. \tag{10}$$

In Eq. (10), $\mathbf{y}$ is the *time-featured* measurement vector, of size $SDP \times 1$, such that $\mathbf{y} = [\langle \tilde{u}_{1,1}, p_1 \rangle, \ldots, \langle \tilde{u}_{1,1}, p_P \rangle, \ldots, \langle \tilde{u}_{I,J}, p_P \rangle]^T$. The matrix $\mathbf{A}$ is the *featured* weight matrix of size $IJP \times N$. Its nth column is given by $\mathbf{a_n} = [\langle w_{(1,1),n}, p_1 \rangle, \ldots, \langle w_{(1,1),n}, p_P \rangle, \ldots, \langle w_{(I,J),n}, p_P \rangle]^T$.

The reconstruction procedure used in this study to invert the featured forward model was introduced in a previous work.[19] It is chosen here for its ability to provide fairly comparable reconstructions. This reconstruction procedure belongs to the Tikhonov regularization framework.[39] The regularized solution $\mathbf{c}_\alpha^*$ is thus chosen as

$$\mathbf{c}_\alpha^* = \arg\min_{\mathbf{c}} \|\mathbf{y} - \mathbf{A}\mathbf{c}\|_{\mathbf{Q}}^2 + \alpha \|\mathbf{c}\|^2, \tag{11}$$

where $\mathbf{Q}$ is the matrix that allows for down-weighting some measurements and α is the regularization parameter. The choice of α is crucial since the solution $\mathbf{c}_\alpha^*$ dramatically depends on its value. It is even more crucial if, as we intend to do, one is interested in getting fairly comparable solutions from the inversion of different forward models—i.e., obtained from the projection of Eq. (6) on different sets of BFs. Here, as justified in Ref. 19, the choice of α is based on the variability parameter $\mathcal{V}$ of the reconstruction, which is defined as

$$\mathcal{V}(\alpha) = \frac{\|\sigma[\mathbf{c}_\alpha^*]\|}{\|\mathrm{E}[\mathbf{c}_\alpha^*]\|}, \tag{12}$$

where $\mathrm{E}[\mathbf{c}_\alpha^*]$ and $\sigma[\mathbf{c}_\alpha^*]$ are two vectors corresponding to the mean and standard deviation of the components of $\mathbf{c}_\alpha^*$. The regularization parameter α is chosen such that the variability parameter $\mathcal{V}$ of the solution satisfies a prescribed value.

IV. NUMERICAL SIMULATIONS

IV.A. Description of the phantom

We consider the synthetic phantom depicted in Fig. 3. This synthetic phantom is a cube of $4 \times 4 \times 4\ \text{cm}^3$, whose optical properties are inhomogeneous. The background absorption coefficient μ_a is set to $0.2\ \text{cm}^{-1}$, the background reduced scattering coefficient μ_s' to $10\ \text{cm}^{-1}$, and the refractive index n to 1.4. The choice of these values is based on experimental measurements of breast-tissue properties.[40] We consider two inhomogeneities depicted in Fig. 3. The first inhomogeneity is a cylinder with a diameter of 1 cm, whose absorption coefficient is 10% higher than the background. The axis of this cylinder lies in the plane $z = 2$ cm and intersects the borders of the cube at positions $[-2,0,2]$ and $[0,2,2]$ cm. The second inhomogeneity is a sphere with a diameter of 1 cm, whose reduced scattering coefficient is 10% higher than the background. The center of the sphere is located at position $[1,-1,2]$ cm. The described optical properties are assumed to be the same at both excitation and fluorescence wavelengths.

The absorbing and diffusing medium is excited by $S = 13$ point sources and probed by $D = 13$ detectors. The sources and detectors are uniformly arranged onto two grids of side 3 cm placed 5 cm apart. The medium lies between the detection and the excitation planes, in the so-called transmission geometry.

We consider a unique fluorescent marker type with a lifetime τ of 0.3 ns. The fluorescent markers concentrate preferentially around the three positions $\mathbf{r}_1 = [-1.125, -1.125, 2.125]$, $\mathbf{r}_2 = [0.875, 0.875, 2.875]$, and $\mathbf{r}_3 = [0.875, 0.875, 1.125]$, which are referred to as fluorescence inclusions 1, 2, and 3, respectively (see Fig. 3 for illustration). The marker concentration is set to 1 μM at the center of the inclusions. Away from the center, the concentration decreases. A Gaussian distribution, whose standard deviation σ can be adjusted to tune the spatial extent of the inclusion, is used to account for the decrease in the concentration. Here, inclusion 1 is chosen to be narrower than inclusions 2 and 3. Specifically, the standard deviations of inclusions 1, 2, and 3 are $\sigma_1 = 0.15$ cm, $\sigma_2 = 0.2$ cm, and $\sigma_3 = 0.2$ cm, respectively. To simulate the autofluorescence of the tissue—natural fluorescence in the absence of markers—as well as the unperfect uptake of the marker, a fluorescent background, i.e., a homogeneous distribution of markers within the medium, is also considered. The concentration of the background is set to 0.01 μM.

We have chosen this particular configuration of inclusions in order to assess two of the weaknesses of FDOT: First, the poor ability to detect an inclusion (here, inclusion 1) located far away from the sources and the detectors, [41] and second, the limited resolution between two inclusions (here, inclusions 2 and 3) aligned along the z axis.[23] It is of interest to analyze how the TR approaches cope with these situations.

IV.B. Number of detected photons

When Poisson noise is considered, the SNR of a measurement increases with respect to the value measured. Indeed, if few photons are detected, the SNR of the measurement is low, whereas if many photons are detected, the SNR of the measurement is high. In the following, the SNR is tuned by rescaling the TR measurements before applying the Poisson noise. Specifically, we impose the value of the *maximum*

number of photons detected by *one* of the measurement pairs. We denote by $C_{\max}$ this number of photons. Increasing the prescribed value for $C_{\max}$ is experimentally equivalent to increasing the integration time until $C_{\max}$ photons are collected by one of the detectors. In the present study, $C_{\max}$ lies in the range $[10^5, 10^9]$, which is consistent with the assumption made in Sec. II C concerning the number of detected photons.

IV.C. Computation of the TR weight matrix

The time-varying entries $w_{j,n}(t)=v_n N_x[g_{\mathbf{s}_j,\mathbf{r}_n} * d * g_{\mathbf{r}_n,\mathbf{d}_j}](t)$ of the weight matrix are computed as follows.

Step 1: Calculation of the product of the Fourier transform of the Green's functions: $\hat{g}_{\mathbf{s}_j,\mathbf{r}}(\omega)\hat{g}_{\mathbf{r},\mathbf{d}_j}(\omega)$.

The Green's functions $\hat{g}_{\mathbf{s}_j,\mathbf{r}}(\omega)$ and $\hat{g}_{\mathbf{r},\mathbf{d}_j}(\omega)$ are computed by solving the Fourier transform of Eq. (1) by means of finite element method. Specifically, we used the function toastSysmat.m (MATLAB package TOAST,[42] see Ref. 43 for details) to calculate the propagation operator

$$\mathcal{P}(\mathbf{r}) = -\frac{1}{3}\nabla\left[\frac{1}{\mu_s'(\mathbf{r})}\nabla\right] + \frac{j\omega}{\nu} + \mu_a(\mathbf{r})$$

and then inverse the systems $\mathcal{P}(\mathbf{r})g_{\mathbf{r},\mathbf{d}_j} = \delta(\mathbf{r}-\mathbf{d}_j)$ and $\mathcal{P}(\mathbf{r})g_{\mathbf{r},\mathbf{s}_j} = \delta(\mathbf{r}-\mathbf{s}_j)$. This procedure is repeated for $N_\omega = 128$ frequencies chosen so as to get a time step T_e of 125 ps over an observation range of 16 ns. For this calculation, we used a regular grid of 35,937 nodes with a voxel size of $0.125 \times 0.125 \times 0.125$ cm^3. For a given frequency, the computation time of the propagation operator is less than 1 s, the Green's functions for any source and detector positions are then obtained in 45 s, and their multiplications finally require about 1 s (Dell precision workstation running LINUX 2.6.18 with 2.83 GHz Intel Xeon processor and 4 Gbyte RAM).

Step 2: The previous results are multiplied by the Fourier transform of the decaying exponential,

$$\hat{e}(\omega) = \frac{\eta\gamma}{1 + j\omega\tau}. \tag{13}$$

Step 3: The $\hat{w}_{j,n}$'s are inverse-Fourier transformed.

At this point, the weight matrix is of size $169 \times 35\,937 \times 128$. The discretization of the TR weight matrix along its space and time dimensions has been chosen so as to limit numerical errors. However, such precision is not necessary for the purpose of forward modeling and reconstruction. As a result, the two following operations were carried out. (1) The TR weight matrix is subsampled to a coarser grid along its space dimension. The coarser grid, which is fine enough with regard to the resolution of FDOT, consists of 4096 voxels with size of $0.25 \times 0.25 \times 0.25$ cm^3. (2) The weight matrix is restricted to the $N_e = 32$ time samples covering the range [0.5;4.5] ns. The value of the lower bound is chosen large enough to ensure that the light propagation model does not break down. The upper bound is chosen so as to reject time samples of low value. As a result, the final TR weight matrix is of size $169 \times 4096 \times 32$.

IV.D. Computation of the wavelet transforms

The wavelet transform of the measurements along the time axis is performed by means of the filter-bank implementation described by Mallat.[44] This choice allows for significantly alleviating the computational cost of the procedure. The filter-bank implementation of the wavelet transform relies on the iterative discrete filtering and downsampling of the input signal. Two filters are used: The first one is a low-pass filter whose outputs are the approximation coefficients; the second one is a high-pass filter whose outputs are the detail coefficients.

When Haar wavelets are considered, the wavelet coefficient at scale i reduces to sums and differences of the approximation coefficients at scale $i-1$. Specifically, the approximation and detail Haar wavelet coefficients at scale i are given by

$$a_i[q] = \tfrac{1}{2}(a_{i-1}[2q-1] + a_{i-1}[2q]), \quad q \in \{1, \ldots, 2^{-i}N_e\} \tag{14a}$$

and

$$d_i[q] = \tfrac{1}{2}(a_{i-1}[2q-1] - a_{i-1}[2q]), \quad q \in \{1, \ldots, 2^{-i}N_e\}. \tag{14b}$$

To initialize the algorithm, the approximation coefficients at scale 0 are chosen to equal the time samples. Thus, $a_0[q] = u(qT_e)$. The wavelet decomposition of the TR weight matrix up to scale 5 is performed in 2 s.

IV.E. Computation of the solution of the inverse problem

The reconstructed concentration $\mathbf{c}_\alpha^*$ is calculated for 50 noise realizations of the measurements $\mathbf{u}(t)$. So as to downweigh measurements, the resulting features with low SNR, $\mathbf{Q}$ is chosen as a diagonal matrix whose jth component is the inverse of the variance of the jth component of $\mathbf{y}$. This is equivalent to choosing $\mathbf{Q}$ as the inverse of the covariance matrix of the featured measurements $\mathbf{y}$, keeping only the diagonal entries. The retained α is chosen so as to respect a stability of 10% ($\mathcal{V}=0.1$), which is observed to be an acceptable level. The regularization parameter α is calculated iteratively. The variability $\mathcal{V}$ is first evaluated for $\alpha=10$ and $\alpha=1000$. Then, α is obtained by the bisection method since the variability monotonically decreases with respect to the regularization parameter. Typically, this iterative process converges to the prescribed value of $\mathcal{V}$, with an accuracy of 0.01%, in less than 20 iterations.

The reconstruction time mainly depends on the size of the featured weight matrix, i.e., on the number of the chosen BFs. The reconstruction time is 1.6 s for 1 BF, 2.7 s for 2 BFs, 6.3 s for 4 BFs, 25 s for 8 BFs, 201 s for 16 BFs, and 1675 s for 32 BFs.

V. RESULTS

We present now fluorescence marker concentrations reconstructed from the following data features:

- Moments up to the order $k \in \{0,1,3,7\}$. Higher-order moments are sensitive to noise and of limited use in reconstruction.[18] Hence, reconstructions obtained from moments up to order 7 can be considered as the best reconstructions achievable. In the following, we use $m_{0\to k}$ to denote the set of moments from order 0 to order k.
- Haar MRA from scale $i=0$ (the finest) to scale $i=5$ (the coarsest). Precisely, the Haar MRA at scale i consists of the set of Haar approximation coefficients at scale i, i.e., the set $h_i=\{a_i[q]\}, q \in \{1,\ldots,2^{-i}N_e\}$.

The three following remarks are important for interpreting the results. (1) The Haar MRA at scale $i=5$ is equivalent to the moment at order $n=0$. Thus, $m_0=h_5$. (2) When the TR signals are restricted to their Haar MRA at scale $i=5$ (or, equivalently, to their 0th order moments), the TR FDOT reduces to the CW FDOT. (3) The Haar MRA at scale $i=0$ is equivalent to directly considering the TR signal (all the samples are regarded). Thus, reconstructions based on Haar MRA at scale $i=0$ are considered as the best reconstructions achievable with TR measurements.

V.A. Performance metrics

In order to quantitatively characterize the quality of a reconstruction $\mathbf{c}^*$, we use the following performance metrics:

- *Reconstruction error* ϵ_r: This is a global measure that depends on the distance between the phantom concentration $\mathbf{c_p}$ and the reconstructed concentration $\mathbf{c}^*$. The normalized definition of this metric is

$$\epsilon_r = \frac{\|\mathbf{c}^* - \mathbf{c_p}\|_2^2}{\|\mathbf{c_p}\|_2^2}. \tag{15}$$

The closer to zero this metric is, the better is the quality of reconstruction.

- *Reconstruction energy* E_r: This global measure accounts for the ability of preserving the energy of the concentration during the reconstruction process. Indeed, because of the ill-posedness of the inverse problem, the energy of the reconstructed concentration can be severely reduced so as to preserve its stability. The reconstruction energy is defined by

$$E_r = \frac{\|\mathbf{c}^*\|_2^2}{\|\mathbf{c_p}\|_2^2}. \tag{16}$$

The closer to one this metric is, the better is the reconstruction.

We also introduce local criteria, based on the reconstruction accuracy of the three inclusions. Specifically, the reconstructed inclusion concentrations c_1^*, c_2^*, and c_3^* are calculated by searching the positions of the local maximum concentrations $\mathbf{r_1^*}$, $\mathbf{r_2^*}$, and $\mathbf{r_3^*}$ in the vicinity of the three ground-truth positions $\mathbf{r_1}$, $\mathbf{r_2}$, and $\mathbf{r_3}$, respectively. The three reconstructed inclusion concentrations are then defined by $c_i^*=c(\mathbf{r_i^*})$, $i \in \{1,2,3\}$. After that, the two following criteria can be evaluated.

- *z-axis sensitivity* z_S: This criterion assesses the ability of reconstructing an inclusion located far away from sources and detectors. This criterion is important since inclusions are hardly detectable in such zones. The z-axis sensitivity z_S is defined by

$$z_S = \frac{c_1^*}{c_p(\mathbf{r_1^*})}. \tag{17}$$

The closer to one the z-axis sensitivity is, the better the reconstruction is.

- *z-axis contrast* z_C: This is a measure of the ability to resolve two adjacent inclusions along the z axis. It is defined as

$$z_C = \frac{c_{\text{peak}} - c_{\text{valley}}}{(c_{\text{peak}}^2 + c_{\text{valley}}^2)^{1/2}}, \tag{18}$$

where c_{peak} is the peak concentration chosen as the mean of the concentration of inclusions 2 and 3, thus $c_{\text{peak}}=[c_2^* + c_3^*]/2$; c_{valley} is the valley concentration chosen as the minimum concentration among the voxels joining the inclusion positions $\mathbf{r_2}$ and $\mathbf{r_3}$. The closer to one the z-axis contrast is, the better is the reconstruction.

V.B. Noise-free case

We first consider noise-free scenario and present the 3D reconstructed concentrations obtained from the different data features. The reconstructions obtained from moments and Haar MRA are depicted in Fig. 4. To facilitate the comparison of reconstructions obtained from the same number of data features, the latter is placed along the same column. The four performance metrics, evaluated for every data features, are shown in Fig. 5(a). In this representation, the results are grouped depending on the number of BFs used. For the Haar MRA, the scale $i=0$ involves 32 BFs; the scale $i=1$, 16; the scale $i=2$, 8; the scale $i=3$, 4; the scale $i=4$, 2; and the scale $i=5$ only 1.

V.C. Noisy case

We next present the reconstruction results obtained from the different data features in the presence of noise. The maximum number of detected photons is first set to $C_{\max}=10^7$. A quantitative evaluation, based on the performance metrics defined in Sec. V A, is available in Fig. 5. The reconstruction profiles along the z axis, when the moments are considered and when the Haar MRA is considered, are plotted in Figs. 6(a) and 6(b), respectively.

The influence of noise on the reconstructed concentration is investigated next. To this end, we consider a maximum number of detected photons $C_{\max}$ varying from 10^5 to 10^{10}. The results are plotted in Fig. 7. The analysis focuses on the z-axis sensitivity and contrast since these two local metrics are of particular interest in practical applications. Moreover, for clarity, the investigated data features are restricted to the moments up to orders from 0 to 3 and to the Haar MRA at scales for which the number of BFs is the same that for the moments, i.e., scales $i=3,4,5$.

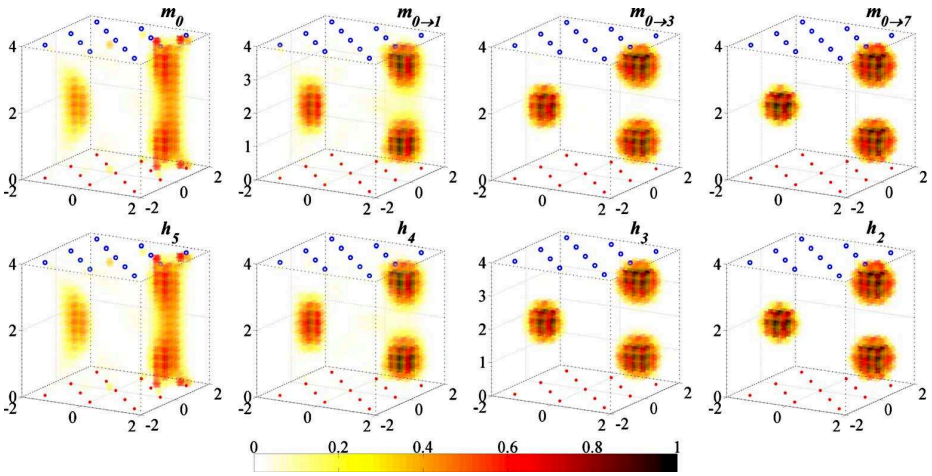

FIG. 4. Three-dimensional representation of the reconstructed fluorescence marker concentrations. No noise was considered. On the upper row moments up to orders 0, 1, 3, and 7 considered. On the bottom row Haar MRA at scales i from 1 to 5 are considered. In a given column, reconstructions are obtained from the same number of BFs (from left to right: 1, 2, 4, and 8).

VI. DISCUSSION

The discussion of the results presented in Sec. V is divided into two parts. First, we compare and analyze the reconstructions obtained from noise-free measurements. Then, we focus on the reconstructions in the presence of noise.

VI.A. Noise-free case

In the noise-free case, the number of BFs used for the reconstruction has a significant impact on the quality of reconstruction. As can be observed on the 3D fluorescence marker reconstructions of Fig. 4, the reconstruction quality is enhanced for increasing numbers of BFs. For instance, inclusion 1 gets better resolved when more moments are considered. The separation between inclusions 2 and 3 also gradually improves. The same trend can be observed for Haar MRA at finer scales. However, for a given number of BFs, the reconstructions obtained from different types of BFs are hardly distinguishable.

These visual observations are confirmed by the performance metrics shown in Fig. 5(a). The global performance metrics as well as the local ones indicate a significant improvement of the reconstruction quality for increasing numbers of BFs. All the previously mentioned points suggest that, in the absence of noise, the TR information is able to significantly improve the reconstruction quality, regardless of the type of BFs chosen to analyze the signals.

VI.B. Noisy case

Comparing the performance metrics of the noisy reconstructions to those of the noise-free reconstructions [compare Fig. 5(a) with Fig. 5(b)], it can be readily seen that the presence of noise on the measurements significantly degrades the quality of reconstruction. This is not surprising since the inverse problem is ill-posed and thus the reconstruction is sensitive to noise. As a result, the conclusions derived from the noise-free case must be carefully rechecked in the presence of noise.

First, it can be seen that the reconstructions obtained from many BFs are severely penalized by the presence of noise on the measurements, which limits the benefit of using the TR information. A more accurate analysis, in terms of the influence of the data type, can be carried out from performance metrics given by Fig. 5(b) together with the reconstruction profiles in Fig. 6. Let us compare how the quality of reconstructions is improved by incrementing the number of BFs. The color code in Figs. 6(a) and 6(b) is the same when the same number of BFs is considered, facilitating the comparison. Let us start with the reconstruction profiles obtained from m_0 and h_5 that required the projection onto only one BF and let us increase the number of BFs to 2. Comparing $m_{0\to1}$ to m_0 and h_4 to h_5, an improvement of the reconstruction quality is observed, notably for the reconstruction contrast z_C. A further improvement is observed, increasing the number of BFs to 4. Interestingly, the benefit is higher with the Haar MRA approach (in Fig. 5 compare ϵ_r and z_C for h_3, h_4, $m_{0\to3}$, and $m_{0\to1}$). While the quality of the Haar MRA-based reconstruction is still improved using eight BFs, the moments approach provides little improvement (compare, for instance, $m_{0\to7}$ to $m_{0\to3}$ on the reconstruction profiles of Fig. 6). The reconstruction quality is not further improved considering more BFs (Haar MRA with $i<2$).

We observe that the optimal TR reconstruction quality, which is obtained considering the Haar MRA at scale $i=0$, is attained using a Haar MRA at scales $i\leq2$. Therefore, it can be concluded that the whole TR information is compressed, in the reconstruction sense, using a Haar MRA at scale $i=2$. However, the optimal TR reconstruction quality cannot be reached using the moments approach.

Now, let us inspect how the total number of detected photons $C_{\max}$—equivalently, the level of noise—affects the pre-

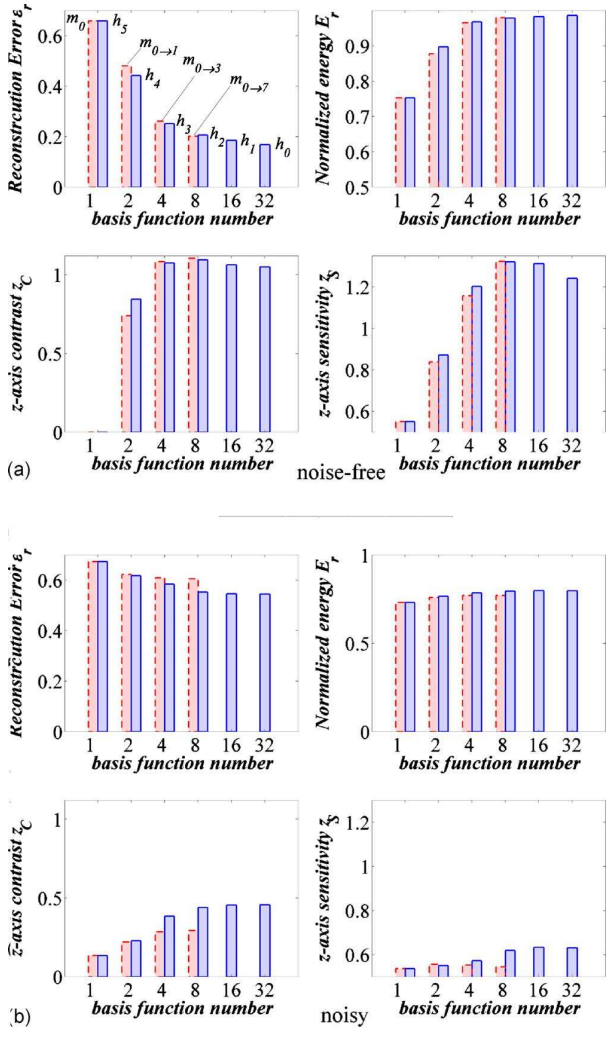

FIG. 5. Performance metrics for both moment-based and Haar-MRA-based reconstructions. The red dash-line bars indicate the metrics of the moments; the blue solid-line bars the metrics of the Haar MRA. C_{max} is set to 10^7 photons. The performance metrics are plotted with respect to the number of BFs used for the reconstruction. The use of 1 BF corresponds to either m_0 or h_5, 2 BFs to either $m_{0\to1}$ or h_4, 4 BFs to either $m_{0\to3}$ or h_3, 8 BFs to either $m_{0\to7}$ or h_2, 16 BFs to h_1, and 32 BFs to h_0.

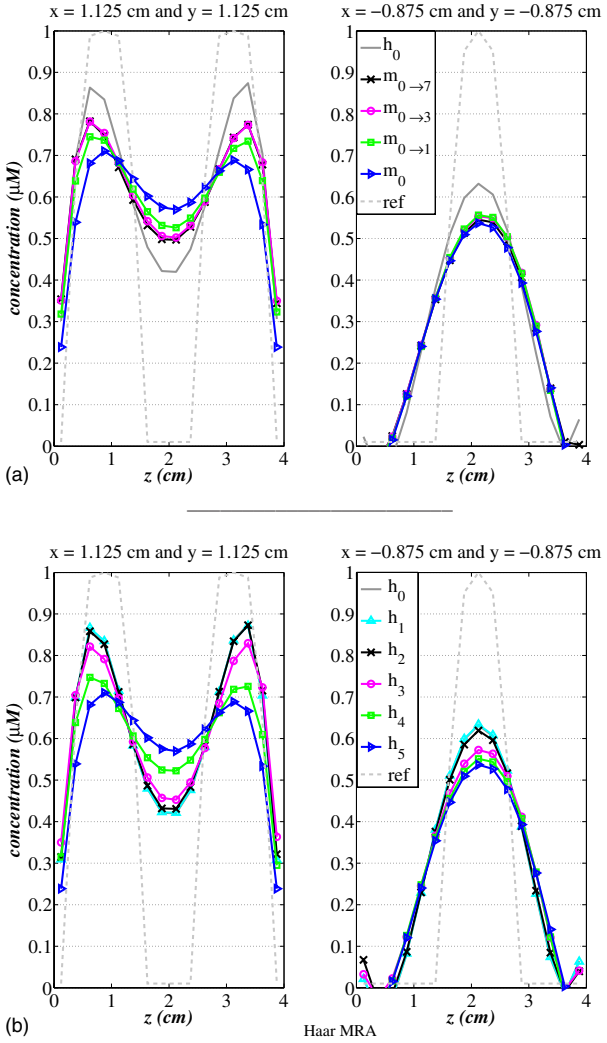

FIG. 6. Fluorescence marker concentration profiles reconstructed from (a) moments up to orders 0–3. (b) Haar MRA at scales 0–5. In both cases, the maximum count number C_{max} is set to 10^7 photons.

vious conclusions. In Fig. 7, we plot the z-axis contrast z_C and sensitivity z_S with respect to C_{max} ranging from 10^5 to 10^{10} photons. Both performance metrics exhibit almost the same pattern.

First, we observe that the Haar MRA approach outperforms the moments approach, regardless of the number of BFs chosen, on the whole range of C_{max} even though the benefit is more limited in terms of sensitivity for $C_{max} < 10^7$ photons. Second, the Haar MRA at scale $i=2$ is shown to provide reconstructions very close to those obtained from the full TR signals within the whole range of C_{max}, which allows for substantially alleviating the reconstruction cost. Third, it can be seen that the TR information is increasingly more beneficial than the CW information for increasing C_{max}. The larger C_{max} is—and thus the better the signal-to-noise ratio of the measurements—the more beneficial the TR information is. However, the TR information is found to be of limited interest in terms of contrast for $C_{max} < 10^5$ photons and in terms of sensitivity for $C_{max} < 10^6$. It should be noted that the provided domains of interest *a priori* depend on different factors such as the optical properties of the medium, the geometry of the medium, or the acquisition configuration.

VII. CONCLUSION

In this paper, we have addressed the problem of choosing the best data features in time-resolved fluorescence diffuse optical tomography. The data feature problem has been reinterpreted from the point of view of the projection of measurements onto some basis functions and a wavelet approach has been proposed. A time-resolved forward model and its projection onto wavelet basis functions have been imple-

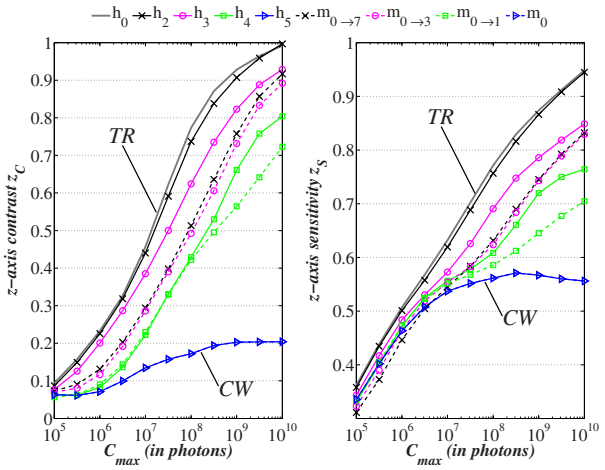

FIG. 7. Evaluation of the z-axis contrast and sensitivity for maximum numbers of detected photons ranging from 10^5 to 10^{10}. The solid lines indicate the use of Haar MRA, while the dashed lines indicate the use of moments. The gray solid line without marker is the reference reconstruction obtained from all the TR samples. The number of BFs employed can be 1 (blue line marked with △), 2 (green line marked with □), 4 (magenta line marked with ○), or 8 (black line marked with ×).

mented. For reconstruction purpose, multiresolution approximations of the time-resolved signals have been considered at different scales. The reference temporal moments of the signal, up to order three, have also been considered. Then, the reconstruction of the marker concentration from the different data features has been performed with a methodology that ensures the comparability of the reconstructions.

For the particular configurations investigated in this paper, our conclusions are the following:

- In the noise-free case, the reconstruction quality is improved with respect to the number of basis functions, regardless of the type of basis functions.
- In the noisy case, the number of basis functions is still an important aspect but the type of basis function becomes important.
- In terms of reconstruction, the wavelet approach outperforms the moment approach for maximum number of detected photons ranging from 10^5 to 10^{10}. In particular, the wavelet approach allows for extending the domain of interest of the TR approach down to 10^5 photons in terms of reconstruction contrast and down to 10^6 photons in terms of reconstruction sensitivity.
- The TR information is compressed with limited reconstruction degradation using only few basis functions. The TR information cannot be compressed, in terms of reconstruction quality, using only temporal moments.
- Whatever the chosen basis functions, the TR approach provides a reconstruction quality equivalent to the CW one if the maximum number of detected photons remains below 10^5.

With these conclusions, we provide theoretical domains of interest for the use of the wavelet transforms of TR measurements. We also derive the domain of interest of the TR modality with respect to the CW modality. Since the domains of interest are expressed in terms of the number of detected photons, they can be of interest from the experimental point of view. In an experimental context, indeed, no benefit of the wavelet or TR approaches can be expected if the number of detected photons remains below the provided thresholds. It should be noted that the experimental domains of interest are most likely to be shifted to larger values of detected photons due to the unavoidable presence of model mismatch and/or extra sources of noise.

Finally, it is important to note that the domains of interest derived here are not universal. Indeed, they have been obtained for a particular acquisition geometry (the so-called transmission geometry) and for a particular reconstruction approach (the Tikhonov framework). Evaluating the domain of interest of the wavelet approach, which comprises the full TR approach, for other acquisition geometries (such as the reflexion geometry or the transmission geometry rotating the medium) and within other reconstruction frameworks (notably the sparse reconstruction one) could be some relevant directions for further studies.

ACKNOWLEDGMENTS

The authors would like to acknowledge the fruitful discussions with Lionel Hervé. This work was supported in part by the Région Rhône-Alpes in the context of project I3M "Multiscale medical imaging and modeling: From the small animal to the human being," of the cluster ISLE.

[a]Electronic mail: nicolas.ducros@polimi.it

[1]V. Ntziachristos and R. Weissleder, "Charge-coupled-device based scanner for tomography of fluorescent near-infrared probes in turbid media," Med. Phys. **29**(5), 803–809 (2002).
[2]R. Roy, A. Godavarty, and E. M. Sevick-Muraca, "Fluorescence-enhanced optical tomography using referenced measurements of heterogeneous media," IEEE Trans. Med. Imaging **22**(7), 824–836 (2003).
[3]E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos, "A submillimeter resolution fluorescence molecular imaging system for small animal imaging," Med. Phys. **30**(5), 901–911 (2003).
[4]A. Corlu, R. Choe, T. Durduran, M. A. Rosen, M. Schweiger, S. R. Arridge, M. D. Schnall, and A. G. Yodh, "Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans," Opt. Express **15**(11), 6696–6716 (2007).
[5]A. Koenig, L. Hervé, V. Josserand, M. Berger, J. Boutet, A. da Silva, J.-M. Dinten, P. Peltie, J.-L. Coll, and P. Rizo, "In vivo mice lung tumor follow-up with fluorescence diffuse optical tomography," J. Biomed. Opt. **13**(1), 011008 (2008).
[6]D. Y. Paithankar, A. U. Chen, B. W. Pogue, M. S. Patterson, and E. M. Sevick-Muraca, "Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media," Appl. Opt. **36**(10), 2260–2272 (1997).
[7]A. B. Milstein, J. J. Stott, S. Oh, D. A. Boas, R. P. Millane, C. A. Bouman, and K. J. Webb, "Fluorescence optical diffusion tomography using multiple-frequency data," J. Opt. Soc. Am. A Opt. Image Sci. Vis **21**(6), 1035–1049 (2004).
[8]J. Selb, A. Dale, and D. Boas, "Linear 3d reconstruction of time-domain diffuse optical imaging differential data: Improved depth localization and lateral resolution," Opt. Express **15**(25), 16400–16412 (2007).
[9]J. Wu, L. Perelman, R. R. Dasari, and M. S. Feld, "Fluorescence tomographic imaging in turbid media using early-arriving photons and laplace transforms," Proc. Natl. Acad. Sci. U.S.A. **94**(16), 8783–8788 (1997).
[10]A. Liebert, H. Wabnitz, J. Steinbrink, H. Obrig, M. Moller, R. Macdonald, A. Villringer, and H. Rinneberg, "Time-resolved multidistance near-infrared spectroscopy of the adult head: Intracerebral and extracerebral absorption changes from moments of distribution of times of flight of

photons," Appl. Opt. **43**(15), 3037–3047 (2004).
[11]S. Lam, F. Lesage, and X. Intes, "Time domain fluorescent diffuse optical tomography: Analytical expressions," Opt. Express **13**(7), 2263–2275 (2005).
[12]F. Gao, H. J. Zhao, Y. Tanikawa, and Y. Yamada, "A linear, featured-data scheme for image reconstruction in time-domain fluorescence molecular tomography," Opt. Express **14**(16), 7109–7124 (2006).
[13]A. Laidevant, A. da Silva, M. Berger, J. Boutet, J.-M. Dinten, and A. C. Boccara, "Analytical method for localizing a fluorescent inclusion in a turbid medium," Appl. Opt. **46**(11), 2131–2137 (2007).
[14]J. Riley, M. Hassan, V. Chernomordik, and A. Gandjbakhche, "Choice of data types in time resolved fluorescence enhanced diffuse optical tomography," Med. Phys. **34**, 4890–4900 (2007).
[15]S. R. Arridge and W. R. B. Lionheart, "Nonuniqueness in diffusion-based optical tomography," Opt. Lett. **23**(11), 882–884 (1998).
[16]A. T. N. Kumar, S. B. Raymond, A. K. Dunn, B. J. Bacskai, and D. A. Boas, "A time domain fluorescence tomography system for small animal imaging," IEEE Trans. Med. Imaging **27**(8), 1152–1163 (2008).
[17]M. Schweiger and S. R. Arridge, "Application of temporal filters to time resolved data in optical tomography," Phys. Med. Biol. **44**(7), 1699–1717 (1999).
[18]N. Ducros, L. Herve, A. da Silva, J.-M. Dinten, and F. Peyrin, "A comprehensive study of the use of temporal moments in time-resolved diffuse optical tomography: Part I. Theoretical material," Phys. Med. Biol. **54**(23), 7089–7105 (2009).
[19]N. Ducros, A. da Silva, L. Herve, J.-M. Dinten, and F. Peyrin, "A comprehensive study of the use of temporal moments in time-resolved diffuse optical tomography: Part II. Three-dimensional reconstructions," Phys. Med. Biol. **54**(23), 7107–7119 (2009).
[20]S. R. Arridge and M. Schweiger, "Direct calculation of the moments of the distribution of photon time of flight in tissue with a finite-element method," Appl. Opt. **34**(15), 2683–2687 (1995).
[21]D. Grosenick, H. Wabnitz, H. H. Rinneberg, K. T. Moesta, and P. M. Schlag, "Development of a time-domain optical mammograph and first in vivo applications," Appl. Opt. **38**(13), 2927–2943 (1999).
[22]D. Grosenick, K. T. Moesta, M. Moller, J. Mucke, H. Wabnitz, B. Gebauer, C. Stroszczynski, B. Wassermann, P. M. Schlag, and H. Rinneberg, "Time-domain scanning optical mammography: I. Recording and assessment of mammograms of 154 patients," Phys. Med. Biol. **50**(11), 2429–2449 (2005).
[23]F. Gao, Y. Tanikawa, H. Zhao, and Y. Yamada, "Semi-three-dimensional algorithm for time-resolved diffuse optical tomography by use of the generalized pulse spectrum technique," Appl. Opt. **41**(34), 7346–7358 (2002).
[24]M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," Proc. IEEE **84**(4), 626–638 (1996).
[25]W. Zhu, Y. Wang, Y. Deng, Y. Yao, and R. L. Barbour, "A wavelet-based multiresolution regularized least squares reconstruction approach for optical tomography," IEEE Trans. Med. Imaging **16**(2), 210–217 (1997).
[26]W. Zhu, Y. Wang, and J. Zhang, "Total least-squares reconstruction with wavelets for optical tomography," J. Opt. Soc. Am. A Opt. Image Sci. Vis **15**(10), 2639–2650 (1998).
[27]B. Kanmani, P. Bansal, and R. M. Vasu, "Computationally efficient optical tomographic reconstructions through waveletizing the normalized quadratic perturbation equation," Saratov Fall Meeting 2005: Optical Technologies in Biophysics and Medicine VII, 2006 (unpublished).
[28]A. Frassati, A. da Silva, J. Dinten, and D. Georges, "Model reduction by multiresolution method applied to fluorescence diffuse optical tomography," Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007 (EMBS 2007), 2007, pp. 767–770 (unpublished).
[29]A. Ishimaru, "Theory and application of wave propagation and scattering in random media," Proc. IEEE **65**(7), 1030–1061 (1977).
[30]V. Y. Soloviev and L. V. Krasnosselskaia, "Dynamically adaptive mesh refinement technique for image reconstruction in optical tomography," Appl. Opt. **45**(12), 2828–2837 (2006).
[31]A. D. Klose, U. Netz, J. Beuthan, and A. H. Hielscher, "Optical tomography using the time-independent equation of radiative transfer—Part 1: Forward model," J. Quant. Spectrosc. Radiat. Transf. **72**(5), 691–713 (2002).
[32]M. S. Patterson, B. Chance, and B. C. Wilson, "Time resolved reflectance and transmittance for the noninvasive measurement of tissue optical-properties," Appl. Opt. **28**(12), 2331–2336 (1989).
[33]R. C. Haskell, L. O. Svaasand, T.-T. Tsay, T.-C. Feng, M. S. McAdams, and B. J. Tromberg, "Boundary conditions for the diffusion equation in radiative transfer," J. Opt. Soc. Am. A Opt. Image Sci. Vis **11**(10), 2727–2741 (1994).
[34]N. Ducros, A. da Silva, J.-M. Dinten, and F. Peyrin, "Approximations of the measurable quantity in diffuse optical problems: Theoretical analysis of model deviations," J. Opt. Soc. Am. A Opt. Image Sci. Vis **25**(5), 1174–1180 (2008).
[35]J. Selb, D. K. Joseph, and D. A. Boas, "Time-gated optical system for depth-resolved functional brain imaging," J. Biomed. Opt. **11**(4), 044008 (2006).
[36]D. Hyde, E. Miller, D. Brooks, and V. Ntziachristos, "A statistical approach to inverting the born ratio," IEEE Trans. Med. Imaging **26**(7), 893–905 (2007).
[37]S. R. Arridge, "Optical tomography in medical imaging," Inverse Probl. **15**(2), R41–R93 (1999).
[38]M. Unser and T. Blu, "Wavelet theory demystified," IEEE Trans. Signal Process. **51**(2), 470–483 (2003).
[39]A. N. Tikhonov and V. A. Arsenin, *Solution of Ill-Posed Problems* (Winston-Wiley, New York, 1977).
[40]J. Mobley and T. Vo-Dinh, in *Biomedical Photonics Handbook*, edited by T. Vo-Dinh (CRC, Boca Raton, FL, 2003).
[41]D. S. Kepshire, S. C. Davis, H. Dehghani, K. D. Paulsen, and B. W. Pogue, "Subsurface diffuse optical tomography can localize absorber and fluorescent objects but recovered image sensitivity is nonlinear with depth," Appl. Opt. **46**(10), 1669–1678 (2007).
[42]Available online at http://web4.cs.ucl.ac.uk/research/vis/toast/.
[43]M. Schweiger and S. R. Arridge, "The finite-element method for the propagation of light in scattering media: Frequency domain case," Med. Phys. **24**(6), 895–902 (1997).
[44]S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell. **11**(7), 674–693 (1989).