



Simultaneous retrieval of optical and geometrical parameters of multilayered turbid media via state-estimation algorithms

HÉCTOR GARCÍA, GUIDO BAEZ,* AND JUAN POMARICO

Centro de Investigaciones en Física e Ingeniería del Centro de la Provincia de Buenos Aires -CIFICEN- (UNCPBA - CONICET - CICPBA)

*gbaez@exa.unicen.edu.ar

Abstract: In the present paper we propose an implementation of the Kalman filter algorithm, which allows simultaneous recovery of the absorption coefficient, the reduced scattering coefficient and the thicknesses of multi-layered turbid media, with the deepest layer taken as semi-infinite. The approach is validated by both Monte Carlo simulations and experiments, showing good results in structures made up of four layers. As it is a Bayesian algorithm, prior knowledge can be included to improve the accuracy of the retrieved unknowns. One of the most promising applications of this approach is the capability of real-time monitoring of living organs by near infrared spectroscopy. In particular, determination of blood perfusion in the adult head is one of the desired goals, allowing continuous control of stroke patients. This demands accurate measurement of the optical properties, especially absorption, of the head layers, from scalp to the cortex.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

OCIS codes: (290.7050) Turbid media; (290.1990) Diffusion; (170.3660) Light propagation in tissue; (170.2655) Functional monitoring and imaging.

References and links

1. M. Patterson, B. Chance, and B. Wilson, "Time resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties," *App. Opt.* **28**, 2331–2336 (1989).
2. M. Hiraoka, M. Firbank, M. Essenpreis, M. Cope, S. Arridge, P. van der Zee, and D. Delpy, "A Monte Carlo investigation of optical pathlength in inhomogeneous tissue and its application to near-infrared spectroscopy," *Phys. Med. Biol.* **38**, 1859–1876 (1993).
3. D. Contini, F. Martelli, and G. Zaccanti, "Photon migration through a turbid slab described by a model based on diffusion approximation. I. Theory," *App. Opt.* **36**, 4587–4599 (1997).
4. B. Montcel, R. Chabrier, and P. Poulet, "Detection of cortical activation with time-resolved diffuse optical methods," *App. Opt.* **44**, 1942–1947 (2005).
5. B. Tromberg, A. Cerussi, N. Shah, M. Compton, A. Durkin, D. Hsiang, J. Butler, and R. Mehta, "Diffuse optics in breast cancer: detecting tumors in pre-menopausal women and monitoring neoadjuvant chemotherapy," *Br. Can Res.* **7**, 279–285 (2005).
6. B. Wassermann, A. Kummrow, K. Moesta, D. Grosenick, J. Mucke, H. Wabnitz, M. Möller, R. Macdonald, P. Schlag, and H. Rinneberg, "In-vivo tissue optical properties derived by linear perturbation theory for edge-corrected time-domain mammograms," *Opt. Exp.* **13**, 8571–8583 (2005).
7. D. Comelli, A. Bassi, A. Pifferi, P. Taroni, A. Torricelli, R. Cubeddu, F. Martelli, and G. Zaccanti, "In vivo time-resolved reflectance spectroscopy of the human forehead," *App. Opt.* **46**, 1717–1725 (2007).
8. C. Sato, T. Yamaguchi, M. Seida, Y. Ota, I. Yu, Y. Iguchi, M. Nemoto, and Y. Hoshi, "Intraoperative monitoring of depth-dependent hemoglobin concentration changes during carotid endarterectomy by time-resolved spectroscopy," *App. Opt.* **46**, 2785–2792 (2007).
9. E. Hillman, "Optical brain imaging in vivo: techniques and applications from animal to man," *J. Biomed. Opt.* **12**, 051402 (2007).
10. P. Jones, H. Shin, D. Boas, B. Hyman, M. Moskowitz, C. Ayata, and A. Dunn, "Simultaneous multispectral reflectance imaging and laser speckle flowmetry of cerebral blood flow and oxygen metabolism in focal cerebral ischemia," *J. Biomed. Opt.* **13**, 044007 (2008).
11. T. W. L. Scheeren, P. H. Schober, and L. A. Schwarte, "Monitoring tissue oxygenation by near infrared spectroscopy (NIRS): background and current applications," *J. Clin. Monit. Comput.* **26**, 279–287 (2012).
12. A. Ishimaru, "Diffusion of light in turbid material," *Appl. Opt.* **28**, 2210–2215 (1989).
13. S. R. Arridge, "Optical tomography in medical imaging," *Inverse Probl.* **15**, R41 (1999).

14. A. Jelzow, "In vivo quantification of absorption changes in the human brain by time-domain diffuse near-infrared spectroscopy," Ph.D. thesis, Technischen Universität Berlin (2013).
15. A. Liemert and A. Kienle, "Light diffusion in N-layered turbid media: frequency and time domains," *J. Biomed. Opt.* **15**, 025002 (2010).
16. F. Martelli, A. Sassaroli, S. del Bianco, and G. Zaccanti, "Solution of the time-dependent diffusion equation for a three-layer medium: application to study photon migration through a simplified human head model," *Phys. Med. Biol.* **52**, 2827–2843 (2007).
17. H. García, D. Iriarte, J. Pomarico, D. Grosenick, and R. Macdonald, "Retrieval of the optical properties of a semiinfinite compartment in a layered scattering medium by single-distance, time-resolved diffuse reflectance measurements," *J. Quant. Spec. Rad. Trans.* **189**, 66–74 (2017).
18. R. A. Willoughby, "Solutions of ill-posed problems (a. n. tikhonov and v. y. arsenin)," *SIAM Rev.* **21**, 266–267 (1979).
19. C. Sato, M. Shimada, Y. Tanikawa, and Y. Hoshi, "Estimating the absorption coefficient of the bottom layer in four-layered turbid mediums based on the time-domain depth sensitivity of near-infrared light reflectance," *J. Biomed. Opt.* **18**, 097005 (2013).
20. Y.-K. Liao and S.-H. Tseng, "Reliable recovery of the optical properties of multi-layer turbid media by iteratively using a layered diffusion model at multiple source-detector separations," *Biomed. Opt. Express* **5**, 975–989 (2014).
21. F. Martelli, S. D. Bianco, L. Spinelli, S. Cavalieri, P. D. Ninni, T. Binzoni, A. Jelzow, R. Macdonald, and H. Wabnitz, "Optimal estimation reconstruction of the optical properties of a two-layered tissue phantom from time-resolved single-distance measurements," *J. Biomed. Opt.* **20**, 115001 (2015).
22. R. Re, D. Contini, L. Zucchelli, A. Torricelli, and L. Spinelli, "Effect of a thin superficial layer on the estimate of hemodynamic changes in a two-layer medium by time domain NIRS," *Biomed. Opt. Express* **7**, 264–278 (2016).
23. D. Milej, A. Abdalmalak, P. McLachlan, M. Diop, A. Liebert, and K. S. Lawrence, "Subtraction-based approach for enhancing the depth sensitivity of time-resolved NIRS," *Biomed. Opt. Express* **7**, 4514–4526 (2016).
24. D. Simon, *Optimal State Estimation: Kalman, H infinity, and Nonlinear Approaches* (John Wiley and Sons, 2006).
25. J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, vol. 160 (Springer Science and Business Media, 2006).
26. G. R. Baez, J. A. Pomarico, and G. E. Elicabe, "An improved extended Kalman filter for diffuse optical tomography," *Biomed. Phys. Eng. Express* **3**, 015013 (2017).
27. D. Poularikas, *The Transforms and Applications Handbook: Second edition* (CRC Press, Boca Raton, 2000).
28. D. Contini, F. Martelli, and G. Zaccanti, "Photon migration through a turbid slab described by a model based on diffusion approximation. I. Theory," *Appl. Opt.* **36**, 4587–4599 (1997).
29. A. Liemert and A. Kienle, "Light diffusion in a turbid cylinder. ii. layered case," *Opt. Express* **18**, 9266–9279 (2010).
30. R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Fluids Eng.* **82**, 35–45 (1960).
31. S. G. Diamond, T. J. Huppert, V. Kolehmainen, M. A. Franceschini, J. P. Kaipio, S. R. Arridge, and D. A. Boas, "Physiological system identification with the kalman filter in diffuse optical tomography," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005*, (Springer, 2005), pp. 649–656.
32. M. Schweiger, S. R. Arridge, and I. Nissila, "Gauss-newton method for image reconstruction in diffuse optical tomography," *Phys. Medicine Biol.* **50**, 2365 (2005).
33. H. Dehghani and D. T. Delpy, "Near-infrared spectroscopy of the adult head: effect of scattering and absorbing obstructions in the cerebrospinal fluid layer on light distribution in the tissue," *Appl. Opt.* **39**, 4721–4729 (2000).
34. T. Vo-Dinh, ed., *Biomedical Photonics Handbook* (CRC Press, 2003).
35. A. Custo, W. M. Wells, A. H. Barnett, E. M. C. Hillman, and D. A. Boas, "Effective scattering coefficient of the cerebral spinal fluid in adult head models for diffuse optical imaging," *Appl. Opt.* **45**, 4747–4755 (2006).
36. G. E. Strangman, Z. Li, and Q. Zhang, "Depth sensitivity and source-detector separations for near infrared spectroscopy based on the Colin27 brain template," *PLOS ONE* **8**, 1–13 (2013).
37. B. Hallacoglu, A. Sassaroli, and S. Fantini, "Optical characterization of two-layered turbid media for non-invasive, absolute oximetry in cerebral and extracerebral tissue," *PLOS ONE* **8**, 1–15 (2013).
38. *NVIDIA CUDA C programming guide, version 3.0* (2010).
39. E. Alerstan, T. Svensson, and S. Anderson-Engels, "Parallel computing with graphics processing units for high speed Monte Carlo simulation of photon migration," *J. Biomed. Opt.* **13**, 060504 (2008).
40. R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini, "A solid tissue phantom for photon migration studies," *Phys. Med. Biol.* **42**, 1971–1979 (1997).
41. J. M. Bernardo and A. F. M. Smith, *Bayesian Theory* (John Wiley & Sons, 1994).
42. V. Rohatgi, *Statistical Inference*, Dover Books on Mathematics (Dover Publications, 2003).
43. D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari, "Tikhonov regularization and the L-curve for large discrete ill-posed problems," *J. Comput. Appl. Math.* **123**, 423 – 446 (2000).

1. Introduction

During the past twenty years, Near Infrared Spectroscopy (NIRS) has become an increasingly interesting tool for Biomedical applications, due to the ability of light in this spectral region

(between 600 nm and 900 nm), to penetrate several centimeters into biological tissue before being absorbed, in contrast to light of other wavelengths, outside this window. [1–4].

Using this property, NIRS has found many applications in the field of Biomedical Optics, becoming a valuable complementary technique to other well established ones, such as MRI, X-Ray Tomography and Ultrasound. Its most important applications range from tumor localization and characterization in mammography [5, 6] to real time monitoring of vital organs [7–10], with the advantage of providing a relative inexpensive and portable equipment for functional analysis. For example, for the case of the brain, this can be achieved by means of the detection and monitoring of blood perfusion in the cortex [11].

Generally speaking, obtaining useful information from experimental data requires modelling of the light propagation in the tissue of interest, which corresponds to a multiple scattering problem [12] usually modeled by the Diffusion Approximation (DA), a simplified option to the more complete theoretical approach given by the Radiative Transfer Equation (RTE) [13]. In particular, it is possible to design models which take into account light interaction with tissue made out of several layers, each of these presenting their own absorption and scattering properties. These parameters, namely μ_a for the absorption coefficient and μ'_s for the reduced scattering coefficient [13], are needed in order to mimic photon propagation in the human head, which is build up from different media such as scalp, skull, cerebrospinal fluid, white matter and gray matter. However, this kind of modelling is not always an easy task. The first and simplest models of the human head considered this region of the body as an optically homogeneous medium [14]. This is, of course, a rather unreal approximation. Since then, several multilayered models have been introduced which take into account the presence of layers of different optical properties and thicknesses [2, 15, 16]. In a recent work [17] some of the authors presented a model for light propagation in layered, turbid cylinders in which the last layer is semi-infinite. This geometry is more similar to the biological system we wish to simulate, and having an analytical model for this geometry helps reducing time calculations when compared to other models.

But, even having an appropriate model, if the purpose is to simultaneously recover several parameters from a given set of boundary measurements, we face the so-called inverse problem. Because of the ill-posedness of this kind of problems and due to the unavoidable presence of noise, traditional parameter estimation techniques can not be used and some regularization is needed. In deterministic inverse problems, such as Tikhonov regularization [18], this problem can be tackled. Nevertheless, there may be situations when it is possible to use prior information, and Bayesian techniques provide a framework where this information can be naturally incorporated.

Several authors [19–23] have considered different modelling approaches to recover simultaneous information of multi-layered media using, generally, a Levenberg-Marquardt approach [19, 20, 22], and a Weighted Maximum Likelihood Estimation [21] and a Moment-based reconstruction [23]. However, these approaches correspond to deterministic-like ones, which are limited in the information retrieved, i.e., the provided result are, usually, point estimates, while in the Bayesian case the resulting solutions are given as probability distributions.

The main goal of this work is thus to propose and validate an algorithm based on the Extended Kalman Filter (EKF) [24, 25], which is a Bayesian method for Observation-Evolution problems that has been adapted for use in Optical Tomography as a parameter estimation tool [26]. It can incorporate information about the measurement noise, as well as prior knowledge about the variables of interest. As will be shown, it is capable of simultaneously retrieving the optical (absorption and reduced scattering), as well as geometrical (thickness), parameters in layered structures. A multivariate distribution of all the parameters of interest is obtained, which allows to perform different statistical analysis, such as marginalization, from which it is possible to obtain information regarding one or more parameters. Also, interval estimation is possible, allowing the acquisition of information about the credibility of a specific parameter. It is also shown that information about correlation between layers and parameters can be obtained in this way.

The proposal is tested with both, Monte Carlo (MC) simulations and multi-distance diffuse reflectance experiments on phantoms consisting of four layers (three slices of different thickness and a semi-infinite one). Simultaneous retrieval of all optical parameters, namely $[\mu_{a,j}, \mu'_{s,j}]$, $j = 1, 2, 3, 4$, as well as the thicknesses l_j , for all layers but the last semi-infinite one, is achieved. It is also shown how the inclusion of prior knowledge improves the results.

This publication is structured as follows: Section 2 introduces the theoretical model used to describe light diffusion in layered, turbid media, as well as the EKF formalism. In Section 3 we describe the most important features of the MC simulations and the experimental considerations for the construction of the layered phantoms; the setup used for the validation experiments is also shown. Section 4 presents the results of applying our algorithm to both, simulations and experiments. In the last Section we summarize the main conclusions, together with some ideas for future work. For the seek of clarity in the main text we have included the detailed derivation of the probability distributions in an Appendix.

2. Theory

2.1. Light diffusion in a semi-infinite turbid cylinder

We consider an N -layered turbid cylinder of radius R , where layer $j = 1, \dots, N$ has absorption coefficient $\mu_{a,j}$, reduced scattering coefficient $\mu'_{s,j}$, refractive index n_j and thickness l_j , except for the last layer, which is considered to be of infinite thickness ($l_N = \infty$). The whole construction contains thus a given number of layers of different thicknesses on top of a semi-infinite one, as sketched in Fig. 1. The origin of the coordinates system is located at the center of the top surface. A beam-like source impinges on the cylinder at the point $(\rho, \phi, z) = (0, 0, 0)$; according to the diffusion approximation, together with the extrapolated boundary condition (EBC) [1], it can be considered as an isotropic point source at a depth $z_0 = 1/\mu'_{s,1}$, i.e., the actual point source position is $\mathbf{r}_0 = (0, 0, z_0)$. The diffusion equation (DE) for the fluence $\Phi_j(\rho, z, \omega)$ within the j^{th} layer, in frequency domain and in cylindrical coordinates is given by:

$$\left[\left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{\partial^2}{\partial z^2} \right) - \left(\frac{\mu_{a,1}}{D_1} + i \frac{\omega}{D_1 c_1} \right) \right] \Phi_1(\rho, z, \omega) = -\frac{\delta(z - z_0)}{D_1 \rho}, \quad 0 \leq z \leq L_1 \quad (1)$$

$$\left[\left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{\partial^2}{\partial z^2} \right) - \left(\frac{\mu_{a,j}}{D_j} + i \frac{\omega}{D_j c_j} \right) \right] \Phi_j(\rho, z, \omega) = 0, \quad L_{j-1} \leq z \leq L_j \quad (2)$$

where $j = 2, \dots, N-1$, $D_j = 1/[3(\mu'_{s,j} + \mu_{a,j})]$, c_j are the diffusion coefficient and the speed of light in the j^{th} layer, respectively, ω is the modulation frequency of the light source and $L_j = \sum_{i=1}^j l_i$ where $L_1 > z_0$. We have omitted the angular dependence in Eqs. (1) and (2) because of the cylindrical geometry. Applying the EBC together with the zero-order Hankel transform [27] and calling $\hat{\Phi}_i$ to Φ_i after the Hankel transform, it is possible to convert the previous system of partial differential equations into a system of ordinary differential equations of the form:

$$\left(\frac{d^2}{dz^2} - \alpha_1^2 \right) \hat{\Phi}_1 = -\frac{1}{D_1} \delta(z - z_0), \quad 0 \leq z \leq L_1 \quad (3)$$

$$\left(\frac{d^2}{dz^2} - \alpha_j^2 \right) \hat{\Phi}_j = 0, \quad L_{j-1} \leq z \leq L_j \quad (4)$$

being $\alpha_j^2 = \frac{\mu_{a,j}}{D_j} + s_n^2 + i \frac{\omega}{D_j c_j}$, while s_n satisfies the relation $J_0(R_{EB} s_n) = 0$ for the Bessel function of the first kind and order zero and for the extrapolated radius $R_{EB} = 2AD_1$ where A is a coefficient that takes into account air-media refractive index mismatch [28].

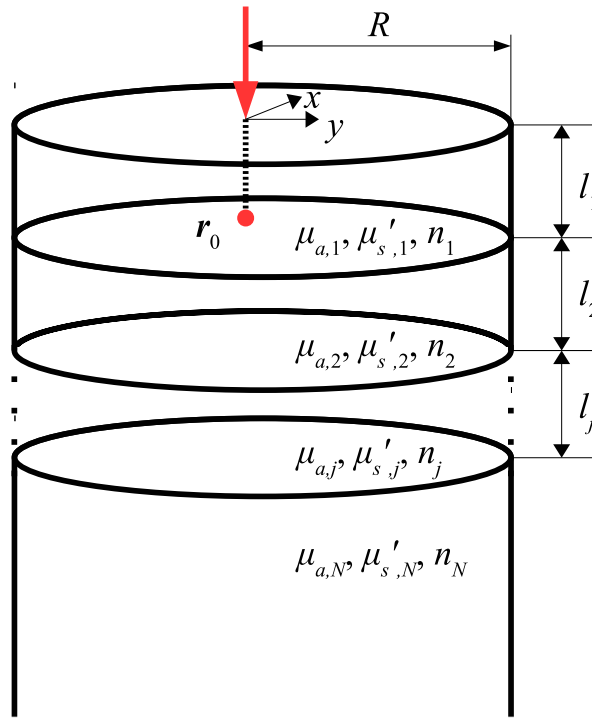


Fig. 1. Scheme for the modelling of light diffusion in a multi-layered cylinder with the last layer of infinite thickness.

The problem reduces now to solving Eqns. (3) and (4), together with the following boundary conditions:

$$\hat{\Phi}_1(s_n, z = -z_{b,1}, \omega) = 0 \tag{5}$$

$$n_{j+1}^2 \hat{\Phi}_j(s_n, z = L_j, \omega) = n_j^2 \hat{\Phi}_{j+1}(s_n, z = L_j, \omega) \tag{6}$$

$$D_j^2 \frac{d\hat{\Phi}_j(s_n, z, \omega)}{dz} \Big|_{z=L_j} = D_{j+1}^2 \frac{d\hat{\Phi}_{j+1}(s_n, z, \omega)}{dz} \Big|_{z=L_j} \tag{7}$$

$$\hat{\Phi}_N(s_n, z \rightarrow \infty, \omega) = 0. \tag{8}$$

Since we are interested in the diffuse reflectance on the top surface, the corresponding Green's

function for that specific layer is:

$$\begin{aligned}
 G_1(s_n, z, \omega) = & \frac{e^{-\alpha(z-z_0)} - e^{-\alpha(z+z_0+z_{b,1})}}{2D_1\alpha_1} + \\
 & + \frac{\sinh[\alpha_1(z_0+z_{b,1})] \sinh[\alpha_1(z+z_{b,1})]}{D_1 e^{\alpha_1(l_1+z_{b,1})}} \times \\
 & \times (D_1\alpha_1 n_1^2 \beta_3 - D_2\alpha_2 n_2^2 \gamma_3) \times \\
 & \times (D_1\alpha_1 n_1^2 \beta_3 \cosh[\alpha_1(l_1+z_{b,1})] + \\
 & D_2\alpha_2 n_2^2 \gamma_3 \sinh[\alpha_1(l_1+z_{b,1})])^{-1},
 \end{aligned} \tag{9}$$

where β_3 and γ_3 can be calculated from the quantities β_N and γ_N :

$$\begin{aligned}
 \beta_N = & D_{N-1}\alpha_{N-1}n_{N-1}^2 \cosh[\alpha_1(\alpha_{N-1}l_{N-1})] + \\
 & + D_N\alpha_N n_N^2 \sinh[\alpha_1(\alpha_{N-1}l_{N-1})]
 \end{aligned} \tag{10}$$

$$\begin{aligned}
 \gamma_N = & D_{N-1}\alpha_{N-1}n_{N-1}^2 \sinh[\alpha_1(\alpha_{N-1}l_{N-1})] + \\
 & + D_N\alpha_N n_N^2 \cosh[\alpha_1(\alpha_{N-1}l_{N-1})]
 \end{aligned} \tag{11}$$

and the recurrence relations:

$$\begin{aligned}
 \beta_k = & D_{k-1}\alpha_{k-1}n_{k-1}^2 \cosh[\alpha_1(\alpha_{k-1}l_{k-1})] \beta_k + \\
 & + D_k\alpha_k n_k^2 \sinh[\alpha_1(\alpha_{k-1}l_{k-1})] \gamma_k
 \end{aligned} \tag{12}$$

$$\begin{aligned}
 \gamma_k = & D_{k-1}\alpha_{k-1}n_{k-1}^2 \sinh[\alpha_1(\alpha_{k-1}l_{k-1})] \beta_k + \\
 & + D_k\alpha_k n_k^2 \cosh[\alpha_1(\alpha_{k-1}l_{k-1})] \gamma_k
 \end{aligned} \tag{13}$$

with $k = 3, 4, \dots, N - 1$. Please note that this minimum value of k is required by the recurrence expressions. For the particular case of a single layer (semi-infinite homogeneous medium), the second term in Eq. (9) vanishes; for the case of two layers, γ_3 and β_3 must be set to 1 [29].

Finally, the reflectance $R(\rho, \omega)$ at the surface $z = 0$ and at a distance ρ from the light source can be obtained as follows [17]:

$$R(\rho, \omega) = \frac{1}{4A\pi^2 R_{EB}^2} \sum_{n=1}^{\infty} G_1(s_n, z=0, \omega) \frac{J_0(s_n \rho)}{J_1^2(s_n R_{EB})}, \tag{14}$$

where A is a factor that takes into account the index mismatch between the cylinder and the surrounding medium [3].

2.2. Parameter estimation procedures

In this subsection, we present the methodologies used to perform the estimation of optical properties.

2.2.1. Extended Kalman filter

The EKF [24] is a technique for state-estimation in observation-evolution problems. It is a non-linear extension to the linear Kalman Filter [30], which has been used, for example, in

estimation of physiological parameters [31]. In [26], the EKF was employed to perform optical tomography on phantoms with good results. Considering a state-evolution model written in the form:

$$x_{t+1} = F(x_t, u_t) + \eta_t, \quad \eta_t \sim N(0, Q_t) \quad (15)$$

$$y_{t+1} = G(x_t, u_t) + \nu_t, \quad \nu_t \sim N(0, R_t) \quad (16)$$

where F is the state-transition operator, G is the observation operator, η_t and ν_t are noise terms following a Normal distribution $N(0, X)$ with mean 0 and covariance matrix X and u_t is a control variable. The EKF uses a linearization of the non-linear operators F and G , and it is suitable for problems with weak non-linearities. The method consist of two steps: the prediction step, where the unknown x_{t+1} evolves according to the transition operator F , and the updating step, where the update is performed according to the predicted state and corrected by the observation operator G . The entire process contemplates the presence of additive Gaussian noise, in the evolution as well as in the measurement. However, it is also possible to consider non-additive noise [24]. The expressions for the predict and update equations for step $t \geq 1$ are given by [25]:

Predict step:

$$x_{t+1|t} = F(x_{t|t}) + Q_t$$

$$\Gamma_{t+1|t} = F'_t \Gamma_{t|t} (F'_t)^T$$

Update step:

$$K_t = \Gamma_{t+1|t} H'_t \left(H'_t \Gamma_{t+1|t} (H'_t)^T + R_t \right)^{-1}$$

$$x_{t+1|t+1} = x_{t+1|t} + K_t [y_t - H(x_{t+1|t})]$$

$$\Gamma_{k+1|k+1} = (I - K_t H'_t) \Gamma_{t+1|t}$$

where $F'_t = \frac{\partial F}{\partial x}(x_{t|t})$, $H'_t = \frac{\partial G}{\partial x}(x_{t+1|t})$ and Γ is the evolution covariance matrix, which evolves by accumulating information of prior steps given that the evolution relates the transition and observation operators (specifically, linearizations of them). The subscript $*|t$ is reserved for variables in the prediction step, while the subscript $*|t + 1$ is used for variables in the update step.

2.2.2. Parameter scaling

In order to improve the parameter estimation procedure, several transformations are employed to reduce the condition number of the matrices R_t and Γ_t . These transformations were suggested in [32]. First, we must consider that the inputs of our model are $\mu_{a,j}$, D_j —although the parameter of interest is μ'_s , the EKF will be performed on the parameter D and then we will revert the transformations to obtain information on μ'_s — and L_i for $j = 1, \dots, N$ and $i = 1, \dots, N - 1$ where N is, as before, the number of layers. We perform the following transformations:

$$x_{a,j} = \log \left(\frac{\mu_{a,j}}{\mu_{a,j,0}} \right) \quad (17)$$

$$x_{D,j} = \log \left(\frac{D_j}{D_{j,0}} \right) \quad (18)$$

$$x_{L,i} = \log \left(\frac{l_i}{l_{i,0}} \right) \quad (19)$$

where $x_{\eta,j,0}$ ($\eta = a, D, l$) corresponds to the initial point of the corresponding variable in the j^{th} layer. For the data transformations, as we are using a frequency-domain model, we use the conventional data divided into logarithm of the amplitude and phase for each source-detector

separation. The model can be seen as an operator in the form:

$$\hat{H}(x, x_0) = H(x_0 \exp(x))$$

The matrix Γ defined in the previous section can be adjusted to incorporate prior knowledge about the parameters of interest. However, the parameters of the EKF are scaled and the prior knowledge must be scaled as well, in order to be used in our framework. It can be seen [25] that this methodology is equivalent to finding a Gaussian approximation of the non-linear inverse problem, i.e. at step k , the resulting distribution is normal with mean $x_{k+1|k+1}$ and covariance $\Gamma_{k+1|k+1}$.

2.2.3. Distribution characterization

After updating the extended Kalman Filter, we obtain the posterior distribution:

$$p(x|y) \propto \exp\left(-\frac{1}{2}\left(\|H(x, x_0) - y\|_{R_{t+1}}^2\right) - \frac{1}{2}\|x - x_{t+1|t+1}\|_{\Gamma_{t+1|t+1}}^2\right), \quad (20)$$

where $\|x\|_A = x^T A^{-1} x$.

If we wish to study the parameters marginally, i.e. each parameter separately or by pairs, we have to consider that the Kalman Filter output is a joint distribution. Nevertheless, it is possible to obtain the marginal distribution [25], which is also Gaussian, where the probability density function (pdf) can be written as:

$$\pi(x_i) \propto \exp\left(-\frac{1}{2} \frac{(x_i - \hat{x}_i)^2}{\Gamma_{t+1|t+1}(i, i)}\right) \quad (21)$$

for the univariate marginal distribution (in the appendix the multivariate distribution is shown) where $\Gamma_{t+1|t+1}(i, i)$ is the i -th element of the diagonal of the covariance matrix $\Gamma_{t+1|t+1}$ obtained in the Kalman Filter. To retrieve the marginal distribution for the original parameters, we perform the inverse transformation by steps and see how the distribution changes. First, we see that as x (the scaled parameter) follows a normal distribution, $x \sim N(x_{k+1}, \Gamma_{k+1})$, then $\exp(x)$ follows a log-normal distribution with the same parameters, $\exp(x) \sim LN(x_{k+1}, \Gamma_{k+1})$. Finally, it is straightforward to check that if $\exp(x) \sim LN(x_{k+1}, \Gamma_{k+1})$ then the posterior density is:

$$\pi(\mu_a | y_N) = \frac{1}{\mu_{a,0}} \pi_{\exp(x_a)} \left(\frac{\mu_a}{\mu_{a,0}} \right) \quad (22)$$

$$\pi(l | y_N) = \frac{1}{l_0} \pi_{\exp(x_l)} \left(\frac{l}{l_0} \right) \quad (23)$$

$$\pi(\mu'_s | y_N) = \frac{\mu_{a,0} + \mu'_{s,0}}{(\mu_a + \mu'_s)^2} \pi_{\exp(x_s)} \left(\frac{\mu_{a,0} + \mu'_{s,0}}{\mu_a + \mu'_s} \right) \quad (24)$$

where $\pi_{\exp(x)}$ is the density of the log-normal distribution of each parameter. Using these functions, and the information of the log-normal distribution, it is possible to calculate point estimators, which the most popular is the mode, also called Maximum a Posteriori estimator (MAP). However, for μ'_s , it is not possible to get an analytical expression for the mode and, for this reason, numerical methods should be used instead. The MAP for the parameters are:

$$\hat{\mu}_a = \mu_{a,0} \exp(x_a - \sigma_a^2) \quad (25)$$

$$\hat{l} = l_0 \exp(x_l - \sigma_l^2) \quad (26)$$

$$\hat{\mu}'_s = \operatorname{argmax}_{\mu'_s > 0} \pi(\mu'_s) \quad (27)$$

In the Bayesian paradigm, the solution is given by the resulting distributions (22), (23) and (24), in this case, the ones obtained via the Kalman Filter. Interval estimators can also be described using the distributions obtained before. If we use a false positive probability of α percent and denoting by P the cumulative distribution function, then:

$$1 - \alpha = P(a < x - \mu < b) \quad (28)$$

$$= P(\exp(a + \mu) < \exp(x) < \exp(b + \mu)) \quad (29)$$

$$= P(x_0 \exp(a + \mu) < x_0 \exp(x) < x_0 < \exp(b + \mu)). \quad (30)$$

Therefore, the interval estimates for μ_a and l can be described by using the interval estimates of the Gaussian distribution, scaled accordingly. For μ'_s , further transformations are needed and the resulting distribution is:

$$1 - \alpha = P\left(\frac{\exp(-b - \mu)}{\mu_{a,0} + \mu'_{s,0}} - \mu_a < \mu'_s < \frac{\exp(-a - \mu)}{\mu_{a,0} + \mu'_{s,0}} - \mu_a\right) \quad (31)$$

In this way, it is possible to analyze the state of information regarding the parameters and how the information we may have is updated in the light of data. It is also possible to determine the initial distributions using the calculations shown in the appendix. There, we use the standard deviations of the scaled parameters to approximate our knowledge regarding the unknowns.

It is also possible to obtain *a posteriori* information about the correlation of different properties by studying the *a posteriori* bivariate marginal distribution among different sets of optical parameters. The bivariate distribution for two parameters is:

$$\pi(x_1, x_2 | y_k) = g_1(x_1)g_2(x_2)f_X(h_1(x_1), h_2(x_2) | x_1, x_2, D_{1,2}) \quad (32)$$

where the functions $g_i, h_i, i = 1, 2$ depends on the chosen parameter and f is the multivariate log-normal distribution with mean and covariance obtained in the Kalman Filter (see Appendix). For each parameter, the functions are

$$g_i(\mu_{a,j}) = \frac{1}{\mu_{a,j,0}}, \quad h_i(\mu_{a,j}) = \frac{\mu_{a,j}}{\mu_{a,j,0}} \quad (33)$$

$$g_i(l_j) = \frac{1}{l_{j,0}}, \quad h_i(l_j) = \frac{l_j}{l_{j,0}} \quad (34)$$

$$g_i(\mu'_{s,j}) = \frac{1}{3\kappa_0(\mu_{a,j} + \mu'_{s,j})^2}, \quad h_i(\mu'_{s,j}) = \frac{1}{3\kappa_0(\mu_{a,j} + \mu'_{s,j})} \quad (35)$$

where $\kappa_0 = \frac{1}{3(\mu_{a,0} + \mu'_{s,0})}$. For example, if we consider estimation in $\mu_{a,1}$ and $\mu_{a,2}$, the resulting bivariate distribution is

$$\pi(\mu_{a,1}, \mu_{a,2} | y_k) = \frac{1}{\mu_{a,1,0}\mu_{a,2,0}} f_X\left(\frac{\mu_{a,1}}{\mu_{a,1,0}}, \frac{\mu_{a,2}}{\mu_{a,2,0}} | x_1, x_2, D_{1,2}\right) \quad (36)$$

3. Materials and methods

3.1. Monte Carlo simulations

In order to validate our algorithm, Monte Carlo (MC) simulations were carried out on a laterally infinite medium—the radius of the corresponding analytical solution used was of $R = 200\text{mm}$ —for eight different source-detector distances (ranging from $\rho = 10$ mm to 45 mm in steps $\Delta\rho$ of 5 mm), over four-layered media. The geometrical aspects are shown in Fig. 3 and the details of their thicknesses and optical properties can be found in Table 1. These parameters were taken

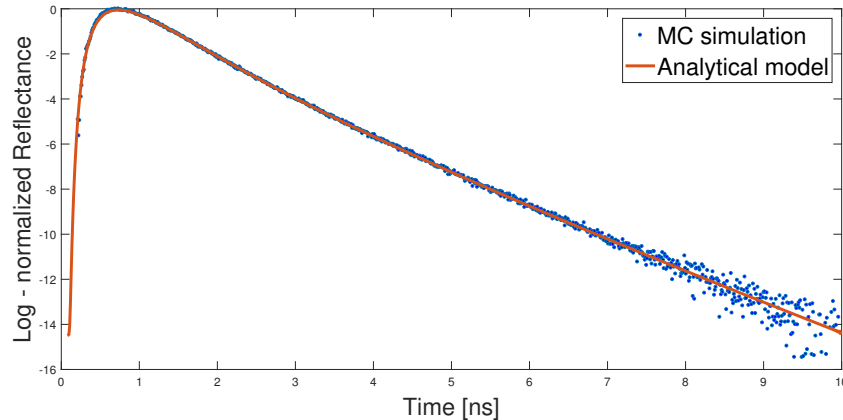


Fig. 2. Normalized Log-reflectance of a Monte-Carlo simulation vs. the corresponding analytical model. The optical and geometrical properties correspond to those presented in Table 1 and the source-detector distance is $\rho = 30\text{mm}$.

from several different sources [20, 22, 33–37]. There is little agreement in the literature about the proper choice of parameters that best suits the behavior of light propagation in biological systems like the human head. Whether our choice may not be appropriate, it is still useful as a proof of concept to validate our methodology.

Table 1. Optical properties and thicknesses of the four layers construction for the MC simulations.

Layer #	$\mu_a[\text{mm}^{-1}]$	$\mu'_s[\text{mm}^{-1}]$	$l[\text{mm}]$
1	0.009	0.86	8
2	0.003	1	3
3	0.0075	2	4
4	0.005	1.1	∞

In each simulation, 2×10^9 photons were launched at the top surface, and the corresponding distributions of times of flight (DTOFs) were acquired at several distances from the source. To accelerate the calculations, we used CUDA [38] technology with a NVIDIA GTX 480 graphics card, each simulation taking between three and four hours, depending on the geometrical and optical parameters. The used code is a customized version of that presented in [39]. An example of MC DTOF is shown in Fig. 2, together with a theoretical DTOF calculated from Eq. (14) (after applying the proper Fourier antitransform), for $\rho = 30\text{mm}$ and optical and geometrical parameters corresponding to those of Table 1. Please note that these DTOFs were plotted in order to show that both agree very well over a large range of times. For both, theory and MC it is possible to represent many orders of magnitude for the Diffuse Reflectance.

3.2. Phantom experiments

Time-resolved reflectance experiments were performed on a four-layered phantom, as shown in Fig. 3. Note that in order to make the measurements compatible with our model, a Fourier transform on the data was performed. A pulsed diode laser Becker & Hickl BHL700 ($\lambda = 785\text{nm}$, $P = 1\text{mW}$, $\text{FWHM} = 270\text{ps}$, 50MHz repetition rate, acquisition time = 30s) impinged onto the upper surface of the phantom, and the diffusely reflected radiation was collected by a

fiber connected to a photomultiplier (Hamamatsu PCM-100-20), which sent the signal to a Time Correlated Single Photon Counting (Becker&Hickl SPC-130) card. In this way, corresponding DTOFs were acquired at several separations between source and detector ranging from $\rho = 35$ mm to $\rho = 70$ mm, in steps of $\Delta\rho = 5$ mm. For these measurements, we tried to avoid changing the collection parameters and/or configuration between them; for example, collection times adequate for $\rho = 70$ mm, will produce signal levels which are too high for source-detector distances shorter than 35 mm. In Fig. 4 we show an example of a DTOF acquired at a interoptode distance of $\rho = 70$ mm using the Tikhonov deconvolution procedure described in the appendix. Please note that, although the shown DTOF has a very low signal-to-noise ratio, the deconvolution procedure is able to retrieve useful information from it.

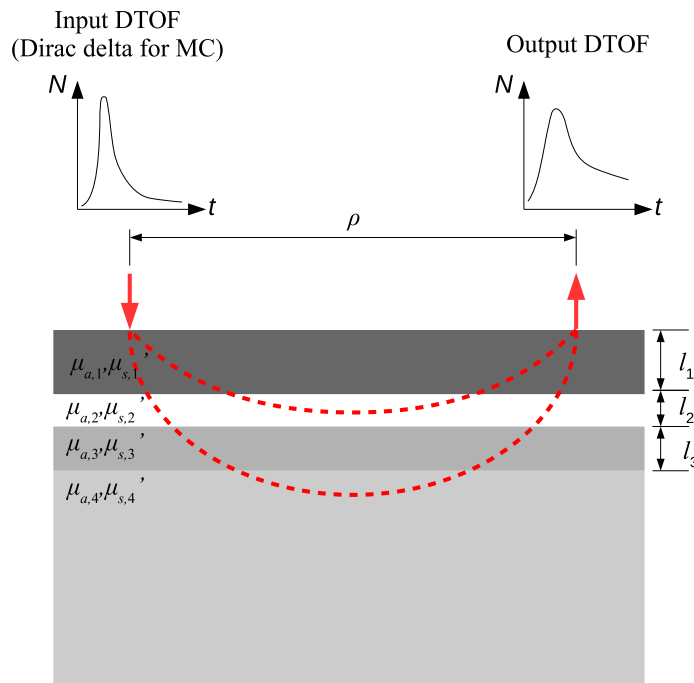


Fig. 3. Scheme showing the geometry used for DTOF acquisition in both, MC simulations and phantom experiments. For the case of MC simulations, the input pulse is taken as a Dirac delta.

The four-layered phantom was prepared with a mixture of water (carrier agent), milk (scattering agent) and ink (absorbing agent). In this case the optical and geometrical parameters were prepared using the same references cited in the previous Subsection. Solidification was achieved by adding agarose to the liquid solution [40], heating up to 95°C and stirring while cooling.

The first layer was prepared with a specific proportion of these components and poured over the bottom of a cuvette while still liquid. Once it was solid, we prepared the solution for the second layer, with different concentrations of the mentioned components, and poured over the first, solid, layered. The same process was repeated for layers three and four.

In order to retrieve the reference optical parameters of the four layers, we measured the corresponding DTOFs on homogeneous phantoms prepared from the same mixture used to produce each layer. These results, as well as the thicknesses, are shown in Table 2. Actual thicknesses were measured directly from the phantom by cutting it along the Z-axis after acquisition of the DTOF. The effect of the IRF is being considered, as for the passage from

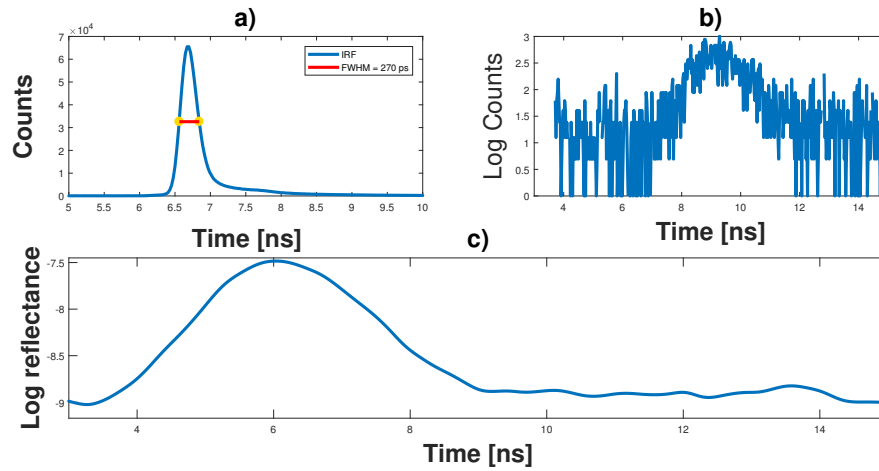


Fig. 4. Information retrieved by the deconvolution procedure. a) IRF with its FWHM of 270 ps. b) DTOF obtained from the four layers medium, at $\rho = 70$ mm. c) Deconvolution solution using Eq. (42) with the information defined in a) and b).

the Time-Domain measurement to the Frequency-Domain measurement, it is necessary to deconvolve the signal in order to apply the Fourier Transform, this was approached by a Tikhonov regularization, the regularization parameter takes into account the noise level of the measurement, here we used the L-Curve method but, the Morozov's principle [25] also works well (see Appendix).

Table 2. Optical properties and thicknesses of the four layers construction for the phantom experiments.

Layer #	$\mu_a [mm^{-1}]$	$\mu'_s [mm^{-1}]$	$l [mm]$
1	0.0124	0.51	5
2	0.0023	0.67	4
3	0.0077	1.29	7
4	0.0060	0.82	50

4. Results

4.1. Monte Carlo simulations results

For the Monte-Carlo simulations, several reconstruction schemes were considered, considering a semi-infinite medium with four layers of optical properties and layer thicknesses given in Table 1 and different source-detector separations. The input of the Kalman Filter is an initial distribution, which is completely described by its mean and covariance. This covariance matrix can be used to describe the *a priori* knowledge about the variables of interest. We assume that there is no correlation between different optical properties and different layers. This is represented by a diagonal matrix, with the entries of the diagonal representing the *a priori* variance of each parameter.

This information corresponds to the scaled parameters, defined in Eqns. (17), (18) and (19), which we assumed are normal. However, it is possible to translate the initial distribution into the scaled parameter using Eqns (29) and (31), where the mean value is the initial value in the unscaled space of parameters. Then, in the normalized space the mean is zero, according to Eqns

(22), (23) and (24) with the initial values $\mu_{a,0}$, $\mu'_{s,0}$, L_0 . Besides, the initial $1 - \alpha$ credibility interval [41] is $a = -Z_{1-\alpha/2}$ and $b = Z_{1-\alpha/2}$ where $Z_\beta = F^{-1}(\beta)$ and F denotes the normal cumulative distribution with zero mean and a given covariance. The chosen initial distributions can be seen in Fig. 5 which represents our initial knowledge. It is important to observe that we can use this information for a wide range of situations, if we let the covariance tend to zero for a given parameter, the distribution tends to an hyperplane with a fixed value. However, this is not recommended because the Kalman gain matrix could become numerically singular if not treated properly, leading to undesired results. This can be avoided in two ways: first, one can use the matrix Q_t as a regularizer; second, the parameter can be fixed in the model to reduce the dimensionality and disregarding the known parameter. By letting the variance of a given parameter tend to infinity, the Gaussian model resembles more to a uniform distribution. The Kalman Filter updates the covariance matrix using the information provided by the user, the data and the model. The scaling is used in order to improve the condition number of the Jacobian matrix used in the Extended Kalman Filter, but also of the covariance matrix.

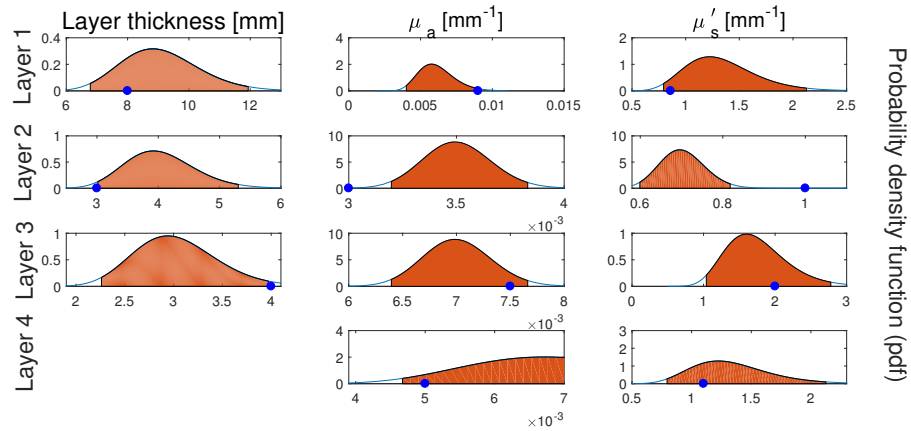


Fig. 5. Initial distributions for each parameter and each layer. The shaded area represents the 99% credibility area. The blue dots represents the true value of the parameter.

Since we are working in a parameter estimation problem (instead of a time-series one), the EKF will stop after reaching a predefined maximum number of iterations or after a given stopping criterion of the form:

$$\epsilon_{k+1} = \frac{\|x_{k+1} - x_k\|_\infty}{\|x_k\|_\infty} < \text{tol}, \quad \text{or} \quad \epsilon_{k+1} \geq \epsilon_k, \quad (37)$$

where “tol” is the desired tolerance, set by the user. After about 600 iterations, the algorithm converges to a solution, as can be seen in Fig. 6, where the MAP of every iteration is shown in real (unscaled) parameters. Using Eqns (25), (26) and (27) with the corresponding 0.99 credibility interval limits, the Kalman Filter evolves marginally. In Fig. 7 we show the distribution of the last iteration with the 0.99 credibility intervals where the reference values are highlighted to show that they are contained in every *a posteriori* interval (we call it *a posteriori* because our state of knowledge was improved by the presence of data).

In Fig. 8, we can see how our knowledge evolved after applying the filtering algorithms. In Fig. 9 the relevance on the amount of information is shown. By using more source-detector separations, our state of knowledge gets improved and this can be seen as a reduction of the width of the intervals. This may be used to determine how many source-detector separations should be used, given a certain level of credibility in the recovery.

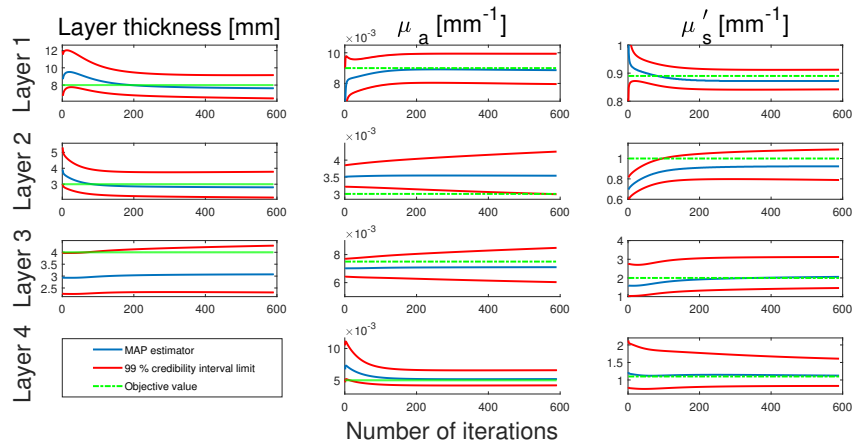


Fig. 6. Evolution of the distributions with respect to the iteration number for each parameter and each layer. The green line corresponds to the reference value.

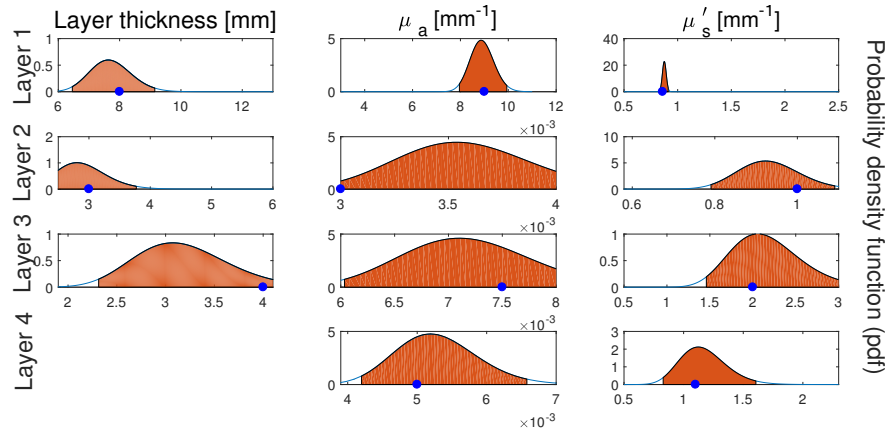


Fig. 7. Final distributions for each parameter and each layer. The shaded area represents the 99% credibility area. The x axis was preserved from Fig. 5, in this way, the effect of the algorithm can be better visualized.

In Fig. 13 joint distributions for different combinations of parameters are shown. This can be extended to the three dimensional case or even more dimensions (see Appendix).

4.2. Phantom experiments results

The EKF was also used for retrieving the optical properties and thicknesses of the four-layered phantom introduced in Section 3.2. Figure 10 shows the initial distributions of l , μ_a and μ'_s for each layer used to recover every parameter, while Fig. 11 shows the corresponding final distributions. In each case the colored regions represent the 99% credibility area, while the blue dots represent the reference values. Figure 12 merges both plots to show the evolution of the distributions for each parameter. It can be seen that these values always lay inside the credibility interval of the final distributions. This figure only shows the results obtained using all of the eight interoptode distances. Should we use more distances, the distributions would tend to sharpen, as it happens in the case of MC simulations.

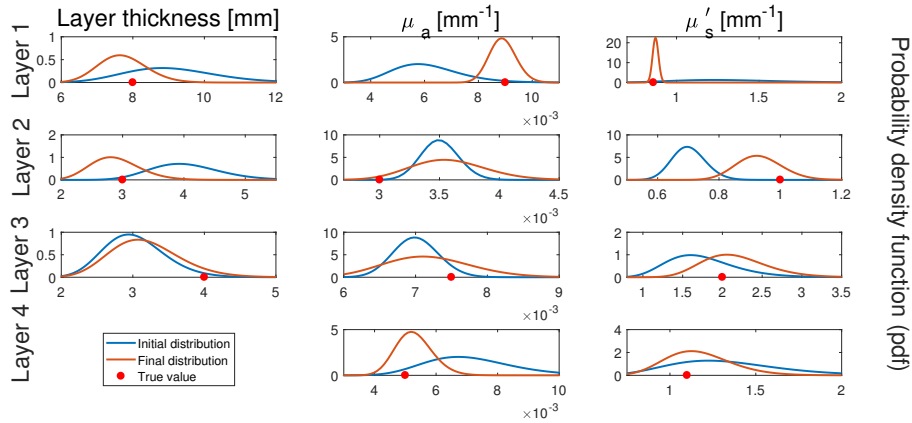


Fig. 8. Comparison of the initial and final distributions for each parameter and each layer. The final distribution contains the reference value in its 99% credibility interval.

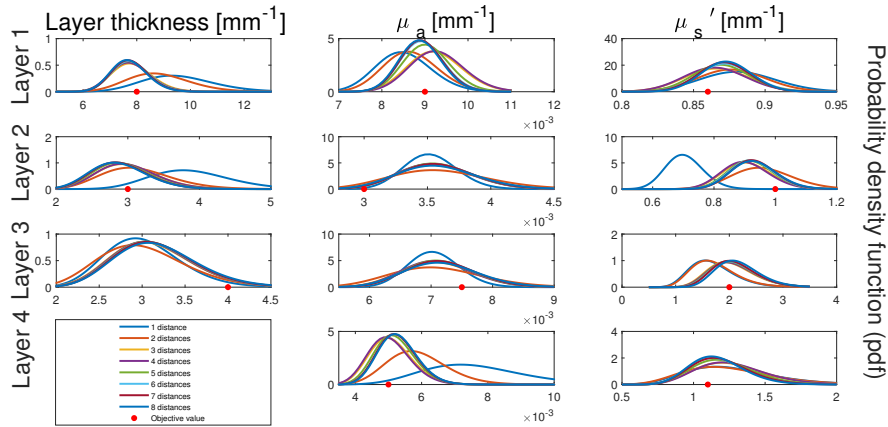


Fig. 9. Distributions obtained using an increasing number of source-detector separations for each parameter and each layer. The red dot corresponds to the reference value.

Finally, the EKF algorithm permits to study how correlated are the retrieved parameters. Figure 13 shows marginalized bivariate distributions for l_3 vs. l_1 , $\mu_{a,3}$ vs. $\mu_{a,1}$, $\mu'_{s,3}$ vs. $\mu'_{s,1}$ and $\mu'_{s,2}$ vs. $\mu_{a,1}$, which were chosen as examples. These distributions provide information about both, the degree of correlation and dispersion of the involved parameters. For example, Fig. 13(a) shows that l_1 and l_3 present low correlation and Fig. 13(b) indicates that $\mu_{a,1}$ and $\mu_{a,3}$ are correlated in such a way that if $\mu_{a,1}$ increases, $\mu_{a,3}$ tends to decrease and Fig. 13(c) shows that when $\mu_{s,1}$ increases, so does $\mu_{s,1}$ with higher correlation.

As shown in the Appendix, it is possible to obtain distributions of higher dimensions, depending on the needs of the user. Moreover, an analytical expression for the joint distribution for any pair of parameters was derived. This allows to study the degree of correlation between parameters of different layers, such as, for example, $\mu_{a,1}$ and $\mu'_{s,3}$, as shown in Fig. 13(d). Finally, in the case where this methodology should be used in another scheme, such as tomographic reconstruction — which could need an initial guess provided by these type of techniques —, it is important to analyze the importance of the punctual estimator. Although the MAP is usually used, some other

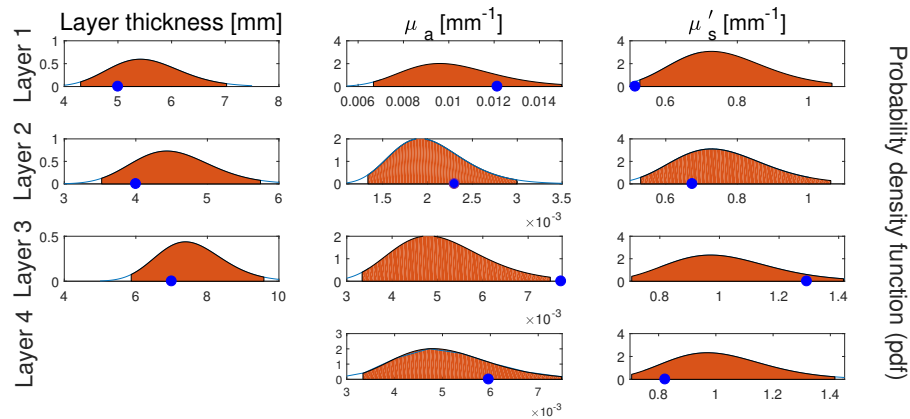


Fig. 10. Initial distributions of the optical properties and thicknesses of the four-layered phantom. The blue dots represent the reference values listed in Table 2.

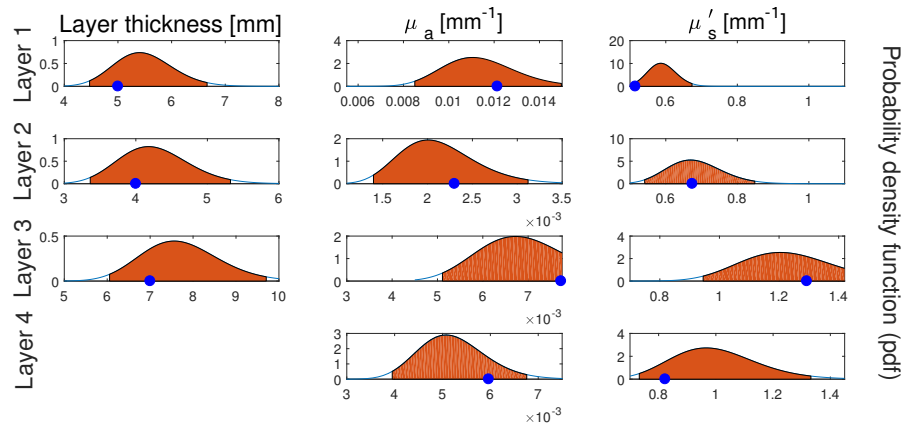


Fig. 11. Final distributions of the retrieved optical properties and thicknesses of the four-layered phantom using eight source-detector separations. The blue dots represent the reference values listed in Table 2. The x axis is preserved from Fig. 10, in order to better visualize the effect of the algorithm.

estimation may be taken in consideration, such as the median or the mean.

5. Discussion

Determination of optical properties in multi-layered media is challenging because of the intrinsic ill-posedness of the operators modelling light propagation through turbid media translating in multiple outputs that are almost equal for very different combination of optical and geometrical parameters. In this study, we proposed a method that allows to provide prior information and is also able to update the information retrieved (for example, if the statistics of the measurement equipment is better characterized). We take advantage of that information to obtain factibility regions for each parameter, as well as different combinations of them, coded in the full covariance matrix. The presented inverse method is flexible in the sense that, should we obtain another model (for example, a Monte Carlo simulator or a numerical scheme, such as the Finite Element Method),

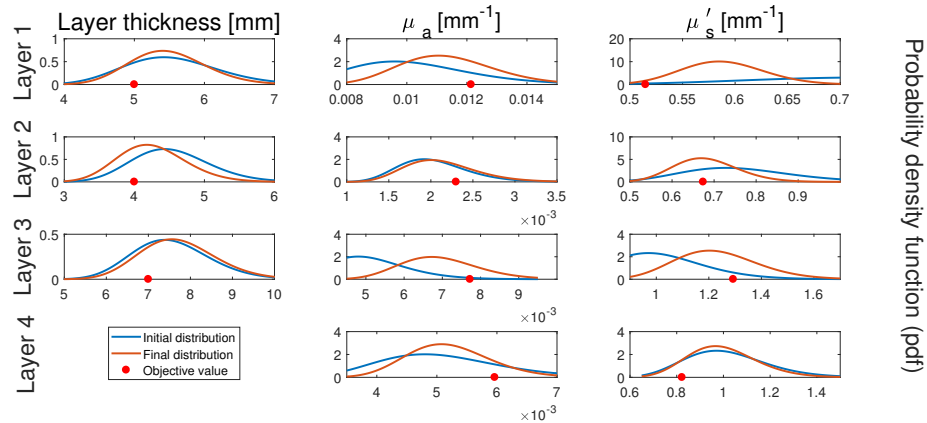


Fig. 12. Initial and final distributions of the optical properties and thicknesses of the four-layered phantom using eight source-detector separations. The red dots represent the reference values listed in Table 2.

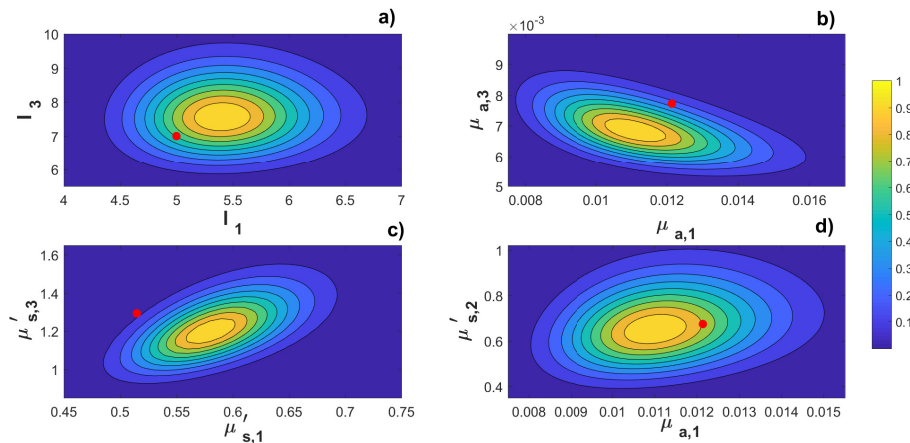


Fig. 13. Bivariate distribution contours for, from left to right and up to down: a) l_1 and l_3 , b) $\mu_{a,1}$ and $\mu_{a,3}$, c) $\mu'_{s,1}$ and $\mu'_{s,3}$, and d) $\mu_{a,1}$ and $\mu'_{s,3}$.

its incorporation to the scheme is straightforward, as an observation operator. This also applies for situations where we know the evolution of a related parameter, as it may be, for example, the case of fluorescence, where the evolution can be seen in terms of the pharmacokinetics of the fluorescent molecule, such as Indocyanine green.

Our method was tested on Monte Carlo simulations in order to avoid the so-called inverse crime, in which the same model is used for both, data generating and reconstruction, and also on phantoms. The success of the methodology can suggest studies to be performed in ex-vivo or in-vivo situations.

The solutions of our method are probability distributions, this means that it is not necessary to perform multiple reconstructions in order to obtain statistics for the parameters, they are already calculated in a single recovery. However, we must remark that this probability distributions are conditional in the data collected as well as the prior information stated by the user, meaning that different prior information will probably lead to a different posterior distribution.

Although we collect data in the Time-Domain, we take it to the Frequency-Domain because, as we have stated, the EKF allows the use of different observation operators. However, there are some numerical issues that must be taken care of: if we used the entire DTOF, which contains all the information regarding photon times of flight, the jacobian matrix of the corresponding operator would be of size $O(N^2)$, where N is the number of time channels, increasing memory requirements, as we would also need to calculate this matrix where, for the central finite-difference scheme, would be of complexity of $O(2N^2)$. For analytical problems, this may not seem expensive but, if we were to use a numerical solver or a Monte Carlo simulation, calculation times become critical, as well as memory requirements. The Time-Domain measurement can also be transformed to Frequency-Domain using multiple modulation frequencies, in this sense we can choose the size of our measurement and it may be possible to reduce the amount of source-detector separations.

Any observation model can be used, such as a time-gated or even CW measurement models, but its implementation must be appropriately studied. In the case of CW measurements, given the high ill-posedness of the CW operator, the prior knowledge (especially with the reduced scattering coefficients) must be quite certain (if not knowing the exact values of the "known" variables) in order to guarantee a convergence, and this may not even be the case. For Time-Domain observation models, we must consider the computational burden as stated before, especially for the numerical models. In that case, a Frequency-Domain numerical implementation (in the sense of, for example, Finite Element Methods) is cheaper than its Time-Domain counterpart—an elliptic PDE against a parabolic PDE—. As a proof of concept, we decided to implement the Frequency-Domain approach and leave the implementation of Time-Domain models for future work. Another important aspect in the use of the Frequency-Domain instead of the Time-Domain regime, is the elimination of the "time zero" (t_0) parameter. In TD, it must be estimated in order to perform the search of the optical properties (simultaneously, leading to non-uniqueness of the resulting operator); in FD it is a scale factor which can be eliminated through an appropriate calibration.

Regarding the credibility intervals, the results show (for example, Fig. 7) that, for the μ_a parameter on the first layer we have a rough uncertainty. As our update of information lies in the observation operator, the prior information and the measurement, and, the prior information is being corrected by the data, the main provider of updating is the measurement and the model. The results might suggest that we can not improve further the information of the first layer, considering that our minimum source-detector distance is of 35 mm. A deep knowledge of the model is necessary in order to obtain better results but also to interpret them; most importantly, the method stabilizes and does not diverge, as it gives all the information it can. This can be compared to the Monte-Carlo simulation where we had lower source-detector distances and the information of the first layer is more certain.

Regarding the limitations of the method, as a Bayesian inverse problem that relies on prior information, the quality of this information, as well as the data collected, influence on the posterior distribution retrieved. This can be translated in that if we do choose, as a starting distribution, one that is way too far from the objective distribution, there is no guarantee of the convergence of it to the true distribution. This may be due to the existence of multiple local modes of the resulting distribution which can appear because of the model (which is the other source of information). It also needs proper tuning of the *a priori* distributions as well as the measurement covariance, which information may, or may not, be available. It can be adapted to allow for broader areas by extending the methodology to a non-parametric approach such as Particle Filter or Gaussian Sums [24], which, also, does not depend on the assumption of Gaussian distributions in the noise as well as the parameter uncertainty.

One of the conclusions that can be achieved while comparing the results from the Monte-Carlo simulations and the experimental results is that the incorporation of information, by itself, may not be enough to achieve better credibility intervals, not only by amount of information (more

source-detector distances) but also the kind of retrievable information provided by that new information. It can be seen, for example, that when smaller distances are incorporated in the MC simulation (see Fig. 9), we were able to obtain much more and precise results than the counterpart in the experimental setup, where we do not have that kind of short distance measurement. In conclusion, it's not only the amount, but the quality of the information (to incorporate information that complements our knowledge, but not that much to become redundant with what we already have). It must be studied, according to the experimental setup, the number of unknowns, and the model to decide an appropriate number of source-detector distances, as well as their location, used in order to optimize the posterior distribution in the sense that the marginalized variance is minimized for each unknown.

6. Conclusions

In this paper, we have developed an inverse problem methodology, based on the extended Kalman filter, for the retrieval of the optical and geometrical parameters in multi-layered turbid media. This approach allowed us to incorporate measurement noise information as well as prior knowledge about the parameters of interest. As it is a Bayesian method, the obtained solutions are distributions of the parameters of interest. Their analytical expressions allow for fast determination of moments and estimators, providing more information than deterministic methods, which usually provide only punctual estimators. Using this distributions, it is possible to show how the state of information improves when there is more information available, such as more source-detector distances. This methodology also allows to perform correlation analysis of different pairs of parameters; this is because the solutions are joint distributions of all the parameters, which can be marginalized to analyze the parameters of interest. We have validated the proposed approach in four-layered media with a semi-infinite bottom layer by both, Monte Carlo simulations and phantom experiments, showing that it is possible to simultaneously retrieve all the parameters (i.e., four absorption coefficients, four reduced scattering coefficients and three thicknesses) within the credibility interval. These results suggest that our method could be successfully applied to biological systems such as the human head, consisting of several layers, in order to study blood activity in the cortical region of the brain.

It is important to remark that, given an adequate theoretical model, there are no limitations concerning the number of layers or the number of parameters to recover. However, an appropriate analysis of the number of source-detector distances must be performed to achieve better results, as could be seen in this work. Given the measured data, the resulting distributions of the parameters can be seen as a first order approximation of the actual ones. In this work, the information provided to the method was codified via the model (to incorporate interoptode distances), or through the description of the initial distributions represented by the distribution parameters. Nevertheless, if we had precise information of the layer thickness (for example, tomographic reconstructions of the head of a patient), it could be excluded from the inverse model. This implies a reduction of the dimensionality of the problem, which improves the resulting distributions and reduces calculation times as well. In a future work we will adapt the extended Kalman Filter to a Gaussian sum in order to improve the speed and convergence of the algorithm.

Appendix

Derivation of the distributions

In the univariate case, as the transformations for μ_a and L are analogous, we will derive it directly:

Let $X \sim N(\mu, \sigma)$ then $\exp(X) \sim LN(\mu, \sigma)$ we wish to determine the distribution of αX , where $\alpha > 0$. In this case, we can write this as $Y = r(X)$ where $r(X) = \alpha X$, by using the method of

transformations for monotone functions [42] the probability distribution is

$$f_Y(y) = f_X(r^{-1}(y)) \left| \frac{d(r^{-1})}{dy} \right| \quad (38)$$

using the distribution of the variable X . Then the distribution is

$$f_Y(y) = \frac{1}{\alpha} f_X \left(\frac{y}{\alpha} \middle| \hat{x}, \sigma \right) \quad (39)$$

which leaves the argument of the log-normal distribution unitless. For μ'_s the function is $Y = g(X)$ with $g(X) = \frac{1}{3\kappa_0 X} - \mu_a$ which is also monotone. Using the identity (38), the resulting distribution is

$$f_Y(\mu'_s) = \frac{1}{3\kappa_0(\mu_a + \mu'_s)^2} f_X \left(\frac{1}{3\kappa_0(\mu_a + \mu'_s)} \middle| \hat{x}, \sigma \right) \quad (40)$$

For the multidimensional case, the distribution is obtained in a similar fashion by stating

$$f_Y(y_1, \dots, y_n) = f_X(g_1^{-1}(y_1), \dots, g_n^{-1}(y_n) | x_{t+1}, \Gamma_{t+1|t+1}) | \det J(y) \quad (41)$$

where $J(y) = \frac{\partial g^{-1}}{\partial y}$ is the Jacobian of the inverse function. In our case, X corresponds to the exponential of the output of the Kalman Filter.

Deconvolution procedure

The problem of deconvolution is ill-posed [25], but this can be tackled with the Tikhonov regularization whose expression for our problem is:

$$f_\lambda = \mathcal{F}^{-1} \left(\frac{\mathcal{F}[g]\mathcal{F}[y]}{|\mathcal{F}[g]|^2 + \lambda|\mathcal{F}[l]|^2} \right) \quad (42)$$

where the regularization parameter λ was found using the L-curve criterion [43].

About the initial distributions

The initial distribution is a Normal distribution with zero mean and covariance $diag(\sigma_a, \sigma_D, \sigma_l)$. The mean is zero because we are working in the scaled parameter space and, as we start from the initial guess we get $x_0 = \log(p_0/p_0)$ where $p_0 = [\mu_{a,0}, D_0, l_0]'$ the logarithm is taken component-wise. For the initial covariance, we assumed that the scaled parameters follow a Normal distribution which is independent of the other parameters. In this sense, we know that

$$P(-3\sigma_a < x_a < 3\sigma_a) = 0.95 \quad (43)$$

where σ_a is the standard deviation to translate this into our original space we perform the calculations made in Eqs. (28) - (30), leading to

$$P(\mu_{a,0} \exp(-3\sigma_a) < \mu_a < \mu_{a,0} \exp(3\sigma_a)) = 0.95 \quad (44)$$

Because we performed monotonic transformations only. Then, we can state our prior information by finding σ_a such that $\mu_{a,0} \exp(-3\sigma_a) < \mu_a < \mu_{a,0} \exp(3\sigma_a)$ better represents our knowledge. Note that our information is stated only as a one single parameter (σ_a) problem and, as our distribution is non-symmetrical, we might not be able to incorporate the prior information exactly, but only approximately. The prior information for the other parameters can be described analogously.

Funding

Authors would like to thank financial support from FCCIC 2016 Program, CICPBA, Buenos Aires, Argentina.

Disclosure

The authors declare that there are no conflicts of interest related to this article.