

Time resolved fluorescence tomography of turbid media based on lifetime contrast

Anand T. N. Kumar

*Department of Neurology and Department of Radiology, Massachusetts General Hospital,
Harvard Medical School, Charlestown MA 02129*

ankumar@nmr.mgh.harvard.edu

Scott B. Raymond

Harvard-MIT Division of Health Science and Technology, Boston MA, 02115

Gregory Boverman

Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy NY, 12180

David A. Boas

*Athinoula A Martinos center for Biomedical Imaging, Massachusetts General Hospital,
Harvard Medical School, Charlestown MA, 02129*

Brian J. Bacskai

*Alzheimer's Disease Research Unit, Department of Neurology, Massachusetts General
Hospital, Charlestown MA, 02129*

Abstract: A general linear model for time domain (TD) fluorescence tomography is presented that allows a lifetime-based analysis of the entire temporal fluorescence response from a turbid medium. Simulations are used to show that TD fluorescence tomography is optimally performed using two complementary approaches: A direct TD analysis of a few time points near the peak of the temporal response, which provides superior resolution; and an asymptotic multi-exponential analysis based tomography of the decay portion of the temporal response, which provides accurate localization of yield distributions for various lifetime components present in the imaging medium. These results indicate the potential of TD technology for biomedical imaging with lifetime sensitive targeted probes, and provide useful guidelines for an optimal approach to fluorescence tomography with TD data.

© 2006 Optical Society of America

OCIS codes: (170.3880) Medical and biological imaging, (170.3650) Lifetime-based sensing, (170.6920) Time-resolved imaging, (170.7050) Turbid media

References and links

1. Hintersteiner *et al.*, "In-vivo detection of amyloid-beta deposits by near-infrared imaging using an oxazine-derivative probe," *Nat. Biotechnol.* **23**, 577 (2005).
2. V. Ntziachristos, J. Ripoll, L. V. Wang, R. Weissleder, "Looking and listening to light," *Nat. Biotechnol.* **23**, 314 (2005).
3. E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos, "A submillimeter resolution for small animal imaging," *Med. Phys.* **30**, 901 (2003).
4. M. A. O'Leary, D. A. Boas, X.D. Li, B. Chance, and A. G. Yodh, "Fluorescence lifetime imaging in turbid media," *Opt. Lett.* **21**, 158 (1996).

5. A. Godavarty, E. M. Sevick-Muraca and M. J. Eppstein, "Three-dimensional fluorescence lifetime tomography," *Med. Phys.* **48**, 1701-1720 (2003).
6. B. B. Das, F. Liu and R. R. Alfano, "Time-resolved fluorescence and photon migration studies in biomedical and model random media," *Rep. Prog. Phys.* **60**, 227 (1997).
7. K. Chen, L. T. Perelman, Q. G. Zhang, R. R. Dasari, and M. S. Feld, "Optical computed tomography in a turbid medium using early arriving photons," *J. Biomed. Opt.* **5** 144 (2000).
8. 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. USA* **94**, 8783 (1997).
9. D. Hall, G. Ma, F. Lesage, Y. Wang, "Simple time-domain optical method for estimating the depth and concentration of a fluorescent inclusion in a turbid medium," *Opt. Lett.* **29**, 2258 (2004).
10. G. M. Turner, G Zacharakis, A. Sourbet, J. Ripoll, V. Ntziachristos, "Complete-angle projection diffuse optical tomography by use of early photons," *Opt. Lett.* **30**, 409 (2005).
11. G. Ma, N. Mincu, F. Lesage, P. Gallant, and L. McIntosh, "System irf impact on fluorescence lifetime fitting in turbid medium," *Proc. SPIE* **5699**, 263 (2005).
12. S. Bloch, F. Lesage, L. Mackintosh, A. Gandjbakche, K. Liang, S Achilefu, "Whole-body fluorescence lifetime imaging of a tumor-targeted near-infrared molecular probe in mice," *J. Biomed. Opt.* **10**, 054003-1 (2005).
13. A. T. N. Kumar, J. Skoch, B. J. Bacsikai, D. A. Boas and A. K. Dunn, "Fluorescence lifetime-based tomography for turbid media," *Opt. Lett.* **30**, 3347-3349 (2005).
14. D. Hattery, V. Chernomordik, M. Loew, I. Gannot, A. Gandjbakhche, *J. Opt. Soc. Am. A*, "Analytical solutions for time-resolved fluorescence lifetime imaging in a turbid medium such as tissue," *J. Opt. Soc. Am.* **18**, 1523-1530 (2001).
15. X. Lam, F. Lesage, X. Intes, "Time domain fluorescent diffuse optical tomography: analytical expressions," *Opt. Express* **13**, 2263-2275 (2005).
16. S. V. Apreleva, D. F. Wilson, and S. A. Vinogradov, "Feasibility of diffuse optical imaging with long-lived luminescent probes," *Opt. Lett.* **31**, 1082-1084 (2006)
17. P. R. Sevin, "The renaissance of fluorescence resonance energy transfer," *Nat. Struct. Biol.* **7**, 730-734, (2000).
18. P. I. H. Bastiaens, and A. Squire, "Fluorescence lifetime imaging microscopy: spatial resolution biochemical processes in the cell," *Trends Cell. Biol.* **9**, 48-52 (1999).
19. B. J. Bacsikai, J. Skoch, G. A. Hickey, R. Allen and B. T. Hyman, "Fluorescence resonance energy transfer determinations using multiphoton fluorescence lifetime imaging microscopy to characterize amyloid-beta plaques," *J. Biomed. Opt.* **8**, 368-375 (2003).
20. Equation (4) is a generalized form of a semi-analytical expression for an infinite medium given in Reference 9, as can be checked by setting $t' \rightarrow t - t'$ and using the analytical expression for the Greens functions with $\mu_{a,s}^x = \mu_{a,s}^m$.
21. S.R. Arridge, "Optical tomography in medical imaging," *Inverse Problems* **15**, R41 (1999).
22. S. Chandrasekhar, *Radiative Transfer* (Dover, New York, 1960).
23. F. Gao, H. Zhao, Y. Tanikawa, and Y. Yamada, "A linear, featured-data scheme for image reconstruction in time-domain fluorescence molecular tomography," *Opt. Express* **14**, 7109-7124 (2006).
24. M. S. Patterson, B. Chance, and B. C. Wilson, "Time-resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties," *Appl. Opt.* **28**, 2331-2336 (1989).
25. J. C. Haselgrove, J. C. Schotland, and J. S. Leigh, "Long-time behavior of photon diffusion in an absorbing medium: application to time-resolved spectroscopy," *Appl. Opt.* **31**, 2678-2683 (1992).
26. A. T. N. Kumar, L. Zhu, J. F. Christian, A. A. Demidov, and P. M. Champion, "On the Rate Distribution Analysis of Kinetic Data Using the Maximum Entropy Method: Applications to Myoglobin Relaxation on the Nanosecond and Femtosecond Timescales," *J. Phys. Chem. B.* **105**, 7847-7856 (2001).
27. A. 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* **21**, 1035-1049 (2004).
28. A. T. N. Kumar, J. Skoch, F. L. Hammond, A. K. Dunn, D. A. Boas and B. J. Bacsikai, "Time resolved fluorescence imaging in diffuse media," *Proc. of SPIE* **6009** 60090Y, (2005).
29. E. E. Graves, J. P. Culver, J. Ripoll, R. Wessleder, V. Ntziachristos, "Singular-value analysis and optimization of experimental parameters in fluorescence molecular tomography," *J. Opt. Soc. Am. A* **21**, 231-241 (2004).
30. J. P. Culver, V. Ntziachristos, M. J. Holboke and A. G. Yodh, "Optimization of optode arrangements for diffuse optical tomography: A singular-value analysis," *Opt. Lett.* **26**, 701-703 (2001).
31. K. Viswanath and M. Mycek, "Do fluorescence decays remitted from tissues accurately reflect intrinsic fluorophore lifetimes?," *Opt. Lett.* **29**, 1512-1514 (2004).

1. Introduction

The development of disease-specific fluorescent markers and genomic reporters has prompted concurrent advances in optical tomography techniques for the non-invasive diagnosis of disease in a living animal or human subject [1, 2]. The most common optical tomographic techniques for fluorescence are based on continuous wave (CW) excitation [3] and frequency modulated

(FD) excitation [4, 5]. Another measurement mode is in the time domain (TD) [7–15], where the excitation is performed using short laser pulses in conjunction with time resolved detection. Fluorescence lifetime reconstructions with turbid media have been discussed in previous works in conjunction with FD [4, 5] and TD [6, 11–14] measurements (CW measurements are incapable of distinguishing fluorescence lifetime from yield). A single TD measurement with a short laser pulse implicitly contains all modulation frequencies, and hence provides the most complete optical information. Moreover, the surface decay data from TD measurements can directly reveal the intrinsic fluorophore lifetime, without the need for reconstructions [11–13]. This feature could be of tremendous importance given the potential development of lifetime sensitive probes for in-vivo applications and the sensitivity of fluorescence lifetime to factors affecting local tissue environment such as pH, viscosity, oxygen concentration and tissue autofluorescence, in addition to molecular interactions such as Forster resonance energy transfer (FRET) [17]. Fluorescence lifetime imaging (FLIM) is already a well established microscopy technique that is used to probe lifetime contrast in thin tissue sections [18, 19].

Image reconstruction algorithms for TD fluorescence measurements have so far been primarily based on derived, or transformed, data types, such as the Laplace transform [8, 23], Mellin transform or moments [15] and the Fourier transform [27, 28]. One advantage of working in a transformed space of the TD data lies in the simplicity of the relationship between the fluorescence lifetime and the measured phase, as compared to the non-linear dependence on fluorophore lifetime (through the exponential decay factor) in the direct TD case. For instance, the phase measured in FD experiments is linearly related to the lifetime distribution [4]. Nevertheless, the lifetime is still in the form of a distribution, which can only be recovered using tomographic reconstructions to remove the contribution to the phase from diffusive propagation. Also, the measured phase at a certain modulation frequency (or imaginary frequency in the Laplace case) is an admixture of contributions due to all the lifetimes present in the medium. On the other hand, Laplace transforms have been applied to selectively reconstruct the early rise portion of the temporal response curve, since early arriving photons undergo minimal scattering, and are largely unaffected by the long lived fluorophore lifetimes [8].

We recently demonstrated experimentally [13] that analyzing the asymptotic decay portion of the diffuse fluorescence temporal response (DFTR) can by itself have distinct advantages: The yield distribution(s) for multiple lifetime component(s) present within the medium can be localized separately using the surface decay amplitudes extracted from multi-exponential fits. In what follows we will simply label this the “asymptotic” approach. The asymptotic approach reduces a cumbersome analysis of a large temporal data set in the decay portion of the DFTR into a multi-exponential curve fitting followed by simple CW reconstructions. This approach can be viewed as an application of the *inverse* Laplace transform (which is equivalent to multi-exponential fitting for a few discrete lifetimes) to reconstruct the decay portion of the DFTR. However, the restriction of this method to the decay portion excludes the information from the earlier portion of the DFTR, which is characterized by a better signal-to-noise (SNR) ratio, and may also contain useful spatial information. It is thus imperative to seek an approach that incorporates the rising and peak portions of the DFTR data into the analysis. In this work, we develop a theoretical formalism that allows a lifetime-based separation of fluorescence yield distributions using the entire TD data. In order to evaluate the optimal choice of temporal measurements for tomography using the direct TD approach, we use a singular value decomposition analysis [29, 30] of the TD weight matrix. To our knowledge, this optimization has not been carried out previously for TD fluorescence measurements, although optimization of multi-frequency data in FD has been studied previously [27, 28]. The relative merits of the optimized direct TD approach and the asymptotic approach are then compared using simulations and the advantages of TD data over CW, in the presence of lifetime contrast, are demonstrated.

This paper consists of three central parts. In Section 2, key integral expressions are presented that enable lifetime-specific tomographic reconstructions of the entire DFTR, along with a discussion of the conditions when the results are valid. In Section 3, the optimization of temporal measurements for fluorescence tomography is addressed numerically, using a SVD analysis. In Section 4, the theoretical formalism developed and the results of the SVD analysis are applied to simulated noisy data to more specifically determine the imaging performance of the rise and decay portions of the DFTR.

2. Theoretical development

In this section, we revisit the basic expressions involved in TD fluorescence tomography, and present a simplified expression under the specific condition of long lifetimes. Consider a turbid imaging medium embedded with fluorophores, described by yield and lifetime distributions (at one wavelength) $\eta(\mathbf{r})$ and $\tau(\mathbf{r}) = 1/\Gamma(\mathbf{r})$, respectively. For tomography, optical sources and detectors are arranged on the surface of the imaging medium. The fluorescence intensity measured at a detector point \mathbf{r}_d at time t for an impulsive excitation at source position \mathbf{r}_s and $t = 0$ can be written in the standard way as a double convolution over time, of the source and emission Green's functions (omitting experimental scaling factors for simplicity):

$$U_F(\mathbf{r}_s, \mathbf{r}_d, t) = \int_{\Omega} d^3r W(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t) \eta(\mathbf{r}), \quad (1)$$

where the weight function, or sensitivity function is given by

$$W(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t) = \int_0^t dt' \int_0^{t'} dt'' G^m(\mathbf{r}_d - \mathbf{r}, t - t') e^{-\Gamma(\mathbf{r})(t' - t'')} G^x(\mathbf{r} - \mathbf{r}_s, t''), \quad (2)$$

with G^x and G^m denoting the source and emission Green's functions, which depend on the net absorption and scattering coefficients (background + fluorophore) at the excitation and emission wavelengths, $\mu_a^{(x,m)}(\mathbf{r})$ and $\mu_s^{(x,m)}(\mathbf{r})$. The above expression assumes a single absorption and re-emission event due to the fluorophore. But this does not prevent the inclusion of multiple absorption of the excitation and emission light by fluorophores in the background medium, which can be incorporated by obtaining the $G^{(x,m)}$ as solutions to the diffusion or transport equations at the excitation and emission wavelengths with the net absorption including the fluorophore absorption at these wavelengths.

As it stands, the TD fluorescence weight function in Eq. (2) is a double convolution in time and is computationally intensive, especially for a tomographic measurement setup with a large number of sources and detectors. But a closer inspection reveals that Eq. (2) can be rewritten in a more manageable form. First, we define a background weight function as:

$$W^B(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t') = \int_0^{t'} dt'' G^x(\mathbf{r}_s, \mathbf{r}, t' - t'') G^m(\mathbf{r}, \mathbf{r}_d, t''), \quad (3)$$

which depends only on the medium optical properties and reduces to the weight function for an absorption perturbation when the excitation and emission wavelengths coincide. Using the commutativity of the convolution operation, we can now re-write Eq. (2) as,

$$W(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t) = \int_0^t dt' W^B(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t') e^{-\Gamma(\mathbf{r})(t - t')}. \quad (4)$$

Since W^B can be pre-calculated with a knowledge of background optical properties, the advantage of Eq. (4) over Eq. (2) is that only a single time integral is left for the tomographic recovery of the yield and lifetime distributions [20]. A more useful form of this expression is

realized if the fluorophores within the medium are described as multiple distributions, $\eta_n(\mathbf{r})$, corresponding to discrete lifetime components, $\tau_n = 1/\Gamma_n$. We then get, for the weight function of each lifetime component,

$$W_n(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t) = \int_0^t dt' W_n^B(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t') e^{-\Gamma_n(t-t')}, \quad (5)$$

so that the total fluorescence signal is expressed as

$$U_F(\mathbf{r}_s, \mathbf{r}_d, t) = \sum_n \int_{\Omega} d^3r W_n(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t) \eta_n(\mathbf{r}). \quad (6)$$

If it is further assumed that the lifetimes are longer than the absorption timescale, i.e., $\tau_n > \tau_a (= 1/\nu\mu_a(\mathbf{r}))$, (see Section 2.1) Eq. 5 can be expressed in a more elegant way that also reveals the connection with previously developed asymptotic lifetime-based-tomography [13]. To derive this most generally, consider the source-free radiative transport equation (RTE) for the Greens functions $G^{(x,m)}$, which is a rigorous description of light transport in a turbid medium [21, 22]:

$$\begin{aligned} \left(\mathbf{s} \cdot \nabla + \frac{1}{v} \frac{\partial}{\partial t} + \mu_a^{(x,m)}(\mathbf{r}) + \mu_s^{(x,m)}(\mathbf{r}) \right) G^{(x,m)}(\mathbf{r}, \mathbf{s}, t) \\ = \mu_s^{(x,m)}(\mathbf{r}) \int_{\Omega} \Theta(\mathbf{s}, \mathbf{s}') G^{(x,m)}(\mathbf{r}, \mathbf{s}', t) ds', \end{aligned} \quad (7)$$

where $\Theta(\mathbf{s}, \mathbf{s}')$ is the scattering phase function. The source terms are dropped from the excitation RTE on the basis that the fluence is calculated away from the source location, and similarly from the emission RTE, given that only a single fluorophore emission event is considered in accordance with the Born approximation initially made in Eq. 2. (Multiple absorption of the excitation and emission light by the fluorophore is still incorporated in the total absorption at these wavelengths viz., μ_a^s and μ_a^m .) Suppose now that we write (dropping the angular dependence for simplicity):

$$G^{(x,m)}(\mathbf{r}, t) = G_0^{(x,m)}(\mathbf{r}, t) e^{-\nu\mu_a^{(x,m)}(\mathbf{r})t}, \quad (8)$$

it can be verified by substituting the above solution into Eq. (7) that the functions $G_0^{(x,m)}$ are dependent only on the *gradient* of the absorption coefficient, $\nabla\mu_a^{(x,m)}(\mathbf{r})$, and independent of $\mu_a^{(x,m)}(\mathbf{r})$ itself. Thus, $G_0^{(x,m)}$ are invariant to constant shifts in the absorption. If we define the Green's functions $G_n^{(x,m)}$ evaluated using a reduced background medium absorption, $\mu_a^{(x,m)}(\mathbf{r}) - \Gamma_n/\nu$, which is positive under the long lifetime condition viz., $\Gamma_n < \nu\mu_a^{(x,m)}(\mathbf{r})$, it is then easily verified using Eq. (8) that

$$G_n^{(x,m)}(\mathbf{r}, t) = G^{(x,m)}(\mathbf{r}, t)|_{\mu_a - \Gamma_n/\nu} = G^{(x,m)}(\mathbf{r}, t)|_{\mu_a} e^{\Gamma_n t}. \quad (9)$$

Now, writing $e^{\Gamma_n t'} = e^{\Gamma_n(t' - t'')} e^{\Gamma_n t''}$, we can use Eq. (9) in Eq. (5) to show that:

$$W_n(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t) = e^{-\Gamma_n t} \int_0^t dt' W_n^B(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t'), \quad (10)$$

where W_n^B is given by Eq. (3) but with the reduced absorption Green's functions $G_n^{x,m}$.

The form of the weight function in Eq. (10) allows the fluorescence signal to be expressed as a multi-exponential sum, analogous to fluorescence lifetime imaging [18] (FLIM):

$$U_F(\mathbf{r}_s, \mathbf{r}_d, t) = \sum_n A_n(\mathbf{r}_s, \mathbf{r}_d, t) e^{-\Gamma_n t}, \quad (11)$$

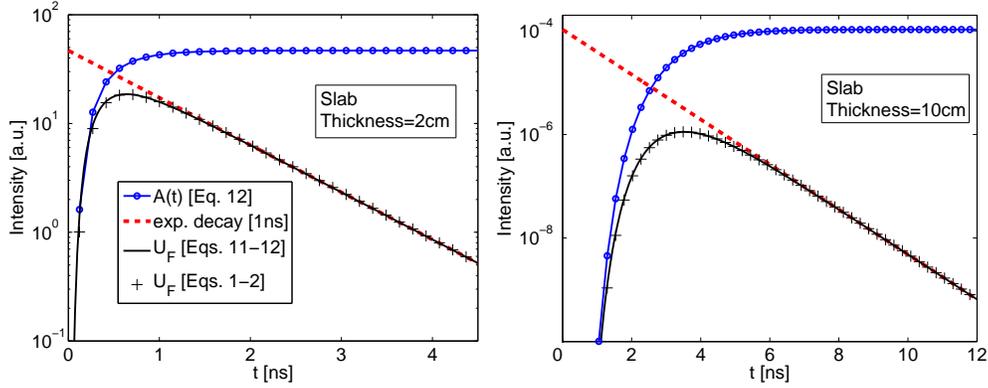


Fig. 1. Simulations to elucidate the diffuse and pure fluorescent decay components in the diffuse fluorescence temporal response, and to demonstrate the accuracy the time domain fluorescence model presented in Eq. (11) in the text. The medium was an infinite slab of thickness 2cm (left panel) and 10cm (right panel), with optical properties $\mu_s^x = \mu_s^m = 10/cm, \mu_a^x = \mu_a^m = 0.1/cm$. The fluorescence signal was calculated for a single source detector pair, with a small fluorescent inclusion at the center. The signal calculated using the conventional approach in Eqs.(1-2), (+ symbol) is compared with that calculated using an effective-absorption based model, viz., Eqs. (11-12) (solid black line). The convolved medium diffusion, $A(t)$ (dotted blue line) and asymptotic fluorescence decay (dashed red line) are also delineated for both cases.

where the decay amplitudes A_n depend on time, in addition to the source and detector locations. The amplitudes define a linear inversion problem for the yield distributions of each lifetime component:

$$A_n(\mathbf{r}_s, \mathbf{r}_d, t) = \int d^3r \left[\int_0^t dt' W_n^B(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t') \right] \eta_n(\mathbf{r}). \quad (12)$$

For fluorophores with lifetimes comparable to optical diffusion time scales in biological media (\approx nanoseconds), $A_n(t)$ has a non-trivial time evolution that is determined by the size and optical properties of the imaging medium. In Figure 1, the temporal evolution of $A(t)$ is shown for infinite slabs of thicknesses 2cm and 10cm, with a single 2mm^3 fluorophore inclusion of 1ns lifetime embedded at the center of the slab. Furthermore, the net fluorescence signal calculated using Eqs. (3) and (11-12) is compared with the fluorescence signal computed directly using Eqs. (1-2), to confirm the accuracy of the effective absorption model in Eq. (10). The above equations are also applicable to phosphorescence signals from diffuse media [16], where the lifetimes (\approx microseconds) are very large compared to the diffusion time scales. In this scenario, $A(t)$ can be approximated as a step function in time.

Asymptotic limit: From Eq. (10), it is clear that the weight function for each lifetime component is an average over a time t of the background sensitivity function W_n^B . Let τ_D denote the timescale for the evolution of W_n^B , which will depend on absorption, scattering and medium boundaries (see section 2.1 below). For $t > \tau_D$, the average over W_n^B will then become time independent and reduce to a CW sensitivity function, which we denote by \overline{W}_n . We are thus lead to asymptotic lifetime-based tomography, which was derived previously using complex integration (see Eqs. (3) and (4) in Reference 13):

$$\lim_{t > \tau_D} W_n(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t) \rightarrow e^{-\Gamma_n t} \overline{W}_n(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}). \quad (13)$$

Therefore, Eqs. (11-12) along with Eq. (3) constitute a TD generalization of asymptotic

lifetime-based tomography, that includes the early arriving photons in addition to the asymptotic decay portion corresponding to the late arriving fluorescent photons. Note from Fig. 1 that the amplitude of the asymptotic fluorescence decay (shown as dashed red lines) equals the long time value of $A(t)$, which is related to the CW sensitivity function \bar{W} .

The results presented in this section can be summarized as follows. With the lifetimes calculated from the asymptotic decay of the TD signal, Eq. (11) can be used to reconstruct the yield distribution for each lifetime using TD data. Since the amplitude for each lifetime is in general time dependent and cannot be separated, the reconstruction is performed directly on the measured data:

$$\begin{pmatrix} \dot{U}_F^j(t_k) \\ \dot{U}_F^j(t_{k+1}) \\ \vdots \end{pmatrix} = \begin{pmatrix} \dot{\cdot} & \dot{W}_n^j(t_k) & \dot{W}_{n+1}^j(t_k) & \dot{\cdot} \\ \dot{\cdot} & \dot{W}_n^j(t_{k+1}) & \dot{W}_{n+1}^j(t_{k+1}) & \dot{\cdot} \\ \dot{\cdot} & \dot{\cdot} & \dot{\cdot} & \dot{\cdot} \end{pmatrix} \begin{pmatrix} \dot{\eta}_n(\mathbf{r}) \\ \dot{\eta}_{n+1}(\mathbf{r}) \\ \dot{\cdot} \end{pmatrix}, \quad (14)$$

where for simplicity of notation we have dropped the source-detector co-ordinates and have instead used a single superscript j to denote measurement index that labels each source-detector (S-D) pair. For times longer than τ_D , the decay amplitude becomes time-independent, so that the amplitude for each lifetime component can be recovered asymptotically using multi-exponential fits. These amplitudes constitute a derived data set (inverse Laplace transform) for the inversion of the yield distributions:

$$\begin{pmatrix} \dot{A}_n^j \\ \dot{A}_{n+1}^j \\ \dot{\cdot} \end{pmatrix} = \begin{pmatrix} \dot{\cdot} & \dot{\bar{W}}_n^j & \dot{0} & \dot{\cdot} \\ \dot{\cdot} & \dot{0} & \dot{\bar{W}}_{n+1}^j & \dot{\cdot} \\ \dot{\cdot} & \dot{\cdot} & \dot{\cdot} & \dot{\cdot} \end{pmatrix} \begin{pmatrix} \dot{\eta}_n(\mathbf{r}) \\ \dot{\eta}_{n+1}(\mathbf{r}) \\ \dot{\cdot} \end{pmatrix}. \quad (15)$$

The key difference between Eqs. (14) and (15) is that the asymptotic weight function in Eq. (15) is block diagonal, whereas the TD weight function in Eq. (14) has off-diagonal terms. Thus, the direct TD reconstruction will be characterized by significantly more cross-talk between the lifetime distributions than the asymptotic reconstruction. At least two questions immediately arise, related to the practical application of the above results. Firstly, what is the optimal choice of time points for the TD reconstruction using Eq. (14)? Secondly, what are the relative merits of the direct TD and asymptotic approaches? We will address these questions in Sections 3 and 4.

2.1. Conditions for asymptotic recovery of intrinsic fluorophore lifetimes

A basis for the theoretical work presented in this paper is the direct estimation of the intrinsic fluorophore lifetimes from the decay of TD fluorescence signals. There are two different time scales involved in determining whether fluorophore lifetimes are revealed in the measured decay on the surface of the turbid medium. First is the intrinsic absorption time scale $\tau_a = (\nu\mu_a)^{-1}$, which is the asymptotic decay time of the diffuse temporal response (DTR) at the excitation wavelength, in the limit of homogenous semi-infinite media [24]. Second is the asymptotic decay time, τ_D , of the DTR from a finite sized imaging medium, which includes the effect of boundaries. It is known that the presence of boundaries reduces the decay time [24,25], so that $\tau_D < \tau_a$. Since the DFTR is a convolution of the fluorescence decay with the DTR, two scenarios emerge for a lifetime based analysis of TD fluorescence data from diffuse media:

- **Strong condition:** $\tau_n > \tau_a$. Since $\tau_a > \tau_D$, this guarantees that lifetimes can be measured asymptotically, irrespective of tissue optical properties and medium boundaries. Furthermore, the multi-exponential model presented in Eqs. (11-12) is valid.

- **Weak condition:** $\tau_a > \tau_n > \tau_D$. Lifetimes can still be measured asymptotically, but the reduced absorption model in Eq. (10) is no longer valid. The more general expression for the weight function, viz., Eq. (5), should instead be used for both the direct TD and asymptotic reconstructions.

The strong condition is easily satisfied for nanosecond lifetime fluorophores in biomedical applications ($\mu_a > 0.1\text{cm}^{-1}$ corresponds to $\tau_a < 0.5\text{ns}$). In applications with small volumes as in small animal imaging [12], with thicknesses of a few cm, the weak condition is almost always satisfied. [A numerical evaluation of τ_D for a range of tissue optical properties can be found in Reference 13.] Note that for heterogenous media, it is known that the decay time is relatively constant on the measurement surface [25], so that we can use the average, or bulk absorption in the medium to define τ_a , in evaluating the above conditions. The above two simple rules dictate the condition for measuring intrinsic lifetimes from surface fluorescence decays for arbitrary diffuse imaging media. Note that the *average* decay time on the surface might itself change due to factors that affect the amplitude of individual lifetime components (e.g., thickness of autofluorescence layers [31]), but the point is that individual lifetimes can still be recovered through multi-exponential fits, under the above conditions.

3. Singular value analysis of the time domain weight function

In this section, we will present an optimization study of the number and location of time points for a direct TD reconstruction. The general optimization of source-detector (S-D) pairs and time points is a complex multi-dimensional problem since each S-D pair could ideally be associated with a different time gate. It is, however, reasonable to view the temporal points and S-D arrangements as independent dimensions in the optimization, since for biomedical imaging applications, the length scales involved are not too large and the correspondingly small variations in the temporal response along the measurement surface can be assumed not to affect the results in a significant way. Therefore, in this paper, we consider a fixed S-D geometry and focus on optimizing the temporal measurements for fluorescent tomography. The optimization of S-D configuration has been discussed in previous works for CW fluorescence [29] and non-fluorescent [30] tomography, using a singular value decomposition (SVD) analysis. Here, we will apply SVD to the TD weight matrix W_n defined in Eq. (10) for optimization of the time points within the DFTR. SVD of a matrix W_n yields the three orthogonal matrices, U , S and V , defined as: $W_n = USV^T$. The columns of U and V represent the measurement space and image space modes, and the diagonal matrix S of singular values determines the extent to which these modes are coupled [29, 30]. The number of singular values above a pre-determined noise threshold is directly related to image resolution [30].

The weight matrix as defined in Eq. (10) was simulated for a diffusive slab medium of thickness 2cm in the transmission geometry, with a S-D arrangement as shown in Fig. 2, with 21 sources and 29 detectors arranged in a honeycomb pattern (yielding 609 S-D measurement pairs). The medium consisted of 3564 voxels of size 2mm^3 ($1\text{mm} \times 1\text{mm} \times 2\text{mm}$). The temporal points were chosen to be 200ps apart, corresponding to the typical minimum gate width in time-gated detection techniques [10, 13], and spread across a time range of 6ns . To begin with, consider performing tomography using Eq. 14 with all S-D pairs and a single time point. What is the location for this time point for an optimal reconstruction? To answer this question, the singular value spectra for W_n evaluated at various time points were calculated. The five spectra with the highest values are plotted in Fig 3(a), and the number of singular values, N_σ , above a chosen noise threshold of 10^{-14} is plotted in the inset of Fig. 3(b). It is seen that the spectrum for the time gate near the peak has the highest number of singular values above the noise threshold. It is noteworthy that the slope of the singular value spectrum is lowest for the earliest time, and increases for later times. This could be attributed to the narrower spatial sensitivity

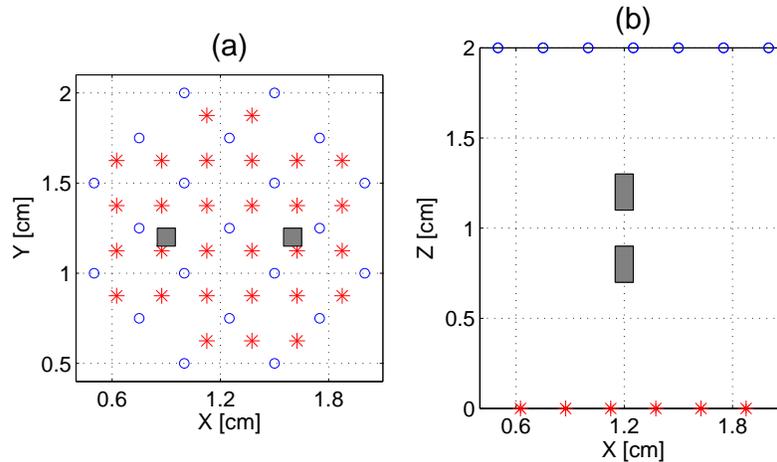


Fig. 2. Measurement geometry and arrangement of sources (*) and detectors (o) used for the simulations. The medium was assumed to be a diffusive slab of thickness 2cm. The targets used in the tomographic reconstructions are shown as gray shaded areas. (a) shows the top-down view and laterally separated (perpendicular to the source-detector axis) targets and (b) shows the side view, and targets located axially, i.e., along the source-detector axis.

profile sampled by the early arriving photons. However, the higher signal level (and the best SNR, in the presence of shot noise) near the peak of the DFTR overcomes the faster decay of the spectrum, resulting in a larger N_σ near the peak. We thus conclude that tomography with a single time gate is optimally performed with a time point near the peak of the DFTR. Note the linearity of N_σ in the exponential decay region (red curve in the inset of Fig 3(b)). This could be attributed to the fact that the SVD spectra are also approximately exponential, as evident from the log plot in Fig. 3(a), so that the intercept of $\text{diag}(S)$ at fixed noise threshold depends linearly on time.

Next, SVD was performed on the weight matrix calculated for all possible pairs of time points, and the pair of time points with the highest number of singular values, N_σ , was determined. It turns out that one of the time points was again near the peak of the DFTR and the other was located near the rise portion of the DFTR. In the same way, the location and N_σ for multiple combinations of time points were determined. In Figure 3(b), N_σ is plotted as a function of the number of time points used. It seen that the proportional increase in N_σ diminishes rapidly after the first 3 or 4 temporal measurements. Also shown in the inset in Fig 3(b) are the first five optimal time points on a representative DFTR on the surface, which are located near the initial portion of the DFTR before the beginning of the fluorescence decay. It was determined that additional time points were located near the decay portion of the DFTR and added little to N_σ .

While the exact location of the optimal time points might vary slightly depending on the specific medium geometry, S-D arrangement and the location of the heterogeneity, it is generally clear from the results in Fig. 3 that the most useful time points of the DFTR for a direct TD reconstruction are located near the rise and peak portions. This result is consistent with Eq. 13, which shows that the weight function is asymptotically factorized into a spatial and temporal component so that multiple time points in the decay region are redundant for tomography. In other words, a brute force direct TD approach is *not* ideal for tomography with the long time decay data. (When the lifetimes are widely separated, the shorter lived components may be

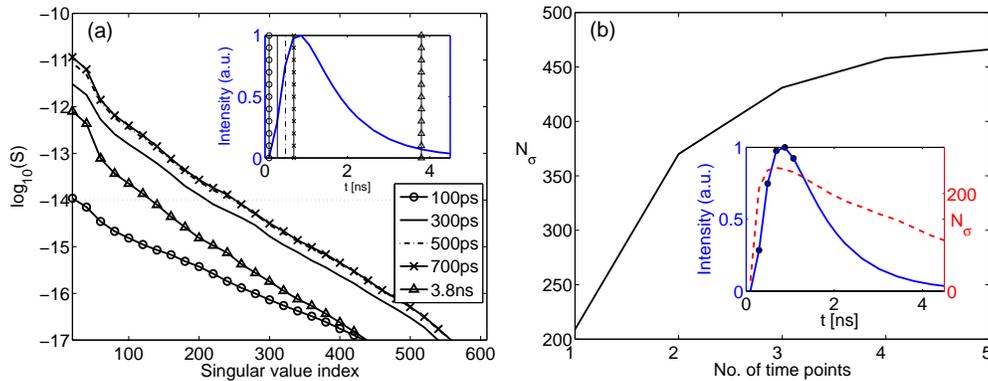


Fig. 3. (a) Singular value spectra of the TD weight matrix W_n [Eq. (10)] evaluated for a single temporal measurement. The time point for which the individual spectra are plotted are indicated as vertical lines in the inset, along with a representative DFTR. (b) Number of useful singular values, N_σ , as a function of the number of time points used in the weight matrix W_n . The weight matrix was optimized separately for each combination of time points. The noise threshold for evaluating N_σ was 10^{-14} (horizontal dotted line in panel (a)). The inset shows the optimal location of the five most significant time points on the DFTR (filled circles) along with N_σ for a single-time weight matrix as a function of the chosen time point along the DFTR (dashed line, right Y-axis).

suppressed by reconstructing later delays [16].) Instead, the asymptotic approach based on a derived data type, viz., the inverse laplace transform (i.e., multi-exponential fit) is more appropriate. In the next section, we will perform tomographic reconstructions with simulated data using realistic noise levels to more quantitatively study the imaging performance of the direct TD and asymptotic approaches.

4. Tomography using direct TD and asymptotic approaches

The results presented so far in the paper suggest that time domain fluorescence tomography can be comprehensively performed in three simple steps. (1) Obtain the *intrinsic* fluorescence lifetime(s) and the corresponding decay amplitude(s) from the asymptotic tail. (2) Reconstruct the individual yield distributions for each decay component using the decay amplitudes for all S-D pairs. (3) Reconstruct the yield distributions for each lifetime component using a few time points near the rise and the peak of the DFTR. These three steps reduce the computational complexity involved in a brute-force reconstruction of a large temporal data set, while retaining the most complete information possible from a TD experiment.

The question immediately arises as to the relative performance of the direct TD and asymptotic approaches. To address this, tomography was performed on simulated noisy data. The simulations employed the same slab geometry used in the SVD analysis above, with two fluorescent inclusions positioned symmetrically with respect to the medium geometry and S-D arrangement (Fig. 2) to remove any intrinsic bias due to the point spread function. The forward data was simulated with a shot noise model, which is characteristic of photon counting detection schemes, for three laterally placed inclusions (Fig. 2(a)) with center-to-center separations of 7mm, 5mm and 3mm. Also considered was the case where the inclusions had non-zero axial separation of 4mm, with zero lateral separation (Fig. 2(b)). The inclusions had distinct lifetimes of 1ns and 1.5ns. The regularization was carried out using a Moore-Penrose inversion algorithm, using the pre-calculated SVD matrices, U, S, V of the weight matrix W_n . Denoting

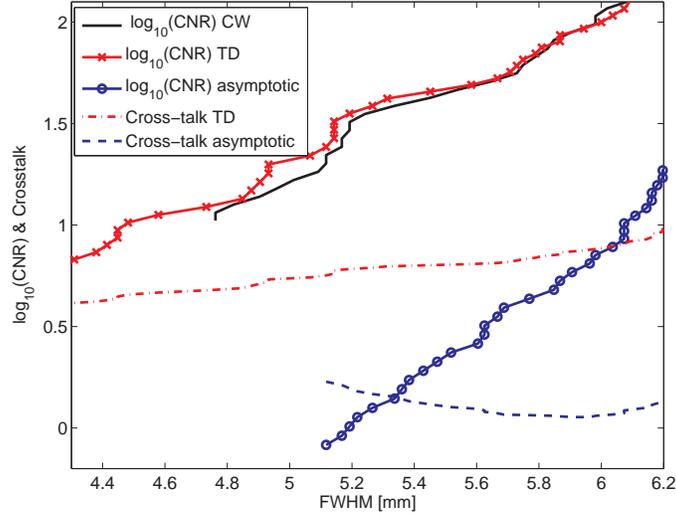


Fig. 4. Plot of the contrast-to-noise ratio (CNR) vs full-width-half-maximum (FWHM) for reconstructions using CW (solid black), optimal direct TD using 4 most significant time points near the rise and peak of the DFTR (red), and the asymptotic approach (blue circles). The cross talk for the 1ns inclusion, viz., the false yield amplitude at the location of the 1ns lifetime component due the 1.5ns component is also shown for the TD (red dash-dot line) and asymptotic (blue dashed line) cases. The simulations used the measurement setup shown in Fig 2, with two 2mm^3 fluorescent targets 7mm apart, having distinct lifetimes of 1ns and 1.5ns.

the measurement vector by Y and the image by X , the inversion takes the following typical form for under-determined problems $Y = W_n X$:

$$X = VS(S^2 + \alpha\lambda I)^{-1}U^T y \quad (16)$$

where $\alpha = \max\{\text{diag}(W_n^T W_n)\}$ and the regularization parameter λ is typically between 10^{-5} and 10^{-3} . Three different reconstructions were performed, namely, CW, direct TD, and asymptotic. The CW reconstructions were performed using the time integrated TD data. The direct TD reconstructions used Eq. (14) with a set of 4 time points on the rising edge of the DFTR, following the SVD analysis results of Fig. 3. The asymptotic TD reconstruction was performed using the amplitudes obtained from a multi-exponential analysis of the decay portion in Eq. 15.

The sensitivity of the reconstructed images to measurement noise was quantified by simulating 100 data sets with noise for each S-D pair and time gate. The contrast-to-noise ratio (CNR), and the full-width-half maximum (FWHM) were then calculated as a function of λ . The CNR is estimated as the ratio of the peak image intensity in a region of interest surrounding the known location of the inclusion, to the mean standard deviation of the voxels in that region. The FWHM was estimated as the cube-root of the total volume of the voxels within half the peak intensity. In addition, for the lifetime based TD and asymptotic reconstructions, which provide separate yield reconstructions η_1 and η_2 for the 1ns and 1.5ns lifetimes, the cross-talk X was estimated to quantify the separability of the two inclusions based on lifetimes. If Ω_1 denotes a chosen region-of-interest around the known location of the 1ns inclusion, then $X^{1ns} = \max[\eta_2(\Omega_1)]/\max[\eta_1(\Omega_1)]$. The yield cross talk for the 1.5ns component was similarly evaluated. The CNR vs FWHM plot is shown in Fig. 4 for the 1ns lifetime inclusion, for the case with 7mm lateral separation between the inclusions. It is clear that the TD reconstruction

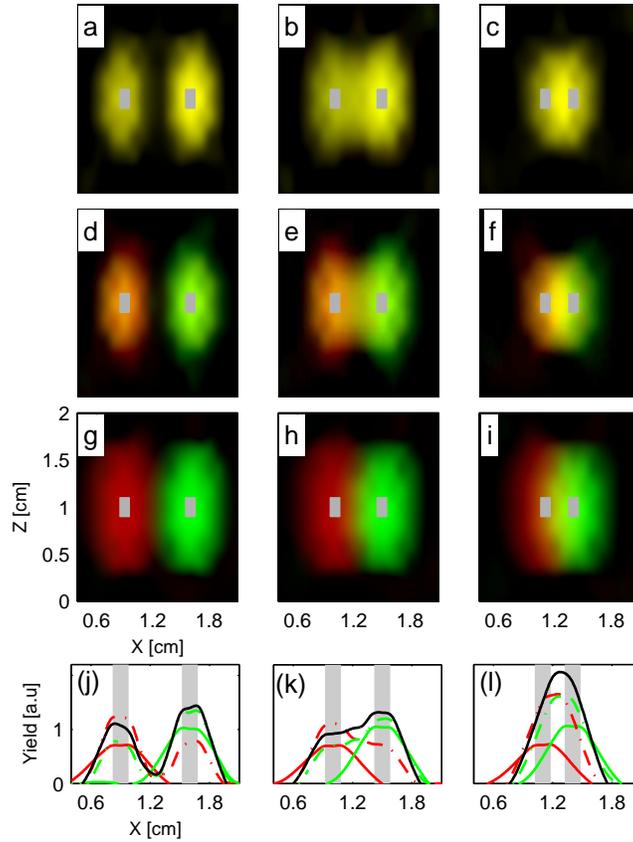


Fig. 5. X-Z slices of the 3-D Reconstructions of two targets with separations of 7mm, 5mm and 3mm, and with lifetimes of 1ns and 1.5ns, located transverse to the S-D axis, using CW (a-c), direct TD (d-f) and asymptotic (g-i) data sets. The measurement geometry used is shown in Fig. 2. The true location of the inclusions (2mm^3) in each case is indicated by the gray shaded area. The reconstructions were regularized such that $\log_{10}(\text{CNR})$ is near unity for all the cases. The images were generated by setting the red and green components of the RGB colormap to be the scaled yield reconstructions for the 1ns and 1.5ns lifetime components, respectively. This way, the cross-talk between the two components is easily visualized as mixture of the two colors (thus, yellow indicates 100% crosstalk). Quantitative plots of the yield along the X axis, at the fixed depth of the inclusions are shown for separations of 7mm (j), 5mm (k) and 3mm (l) (CW - black line; direct TD - 1ns, dashed-dot red and 1.5ns, dashed-dot green; asymptotic - 1ns, solid red and 1.5ns, solid green.)

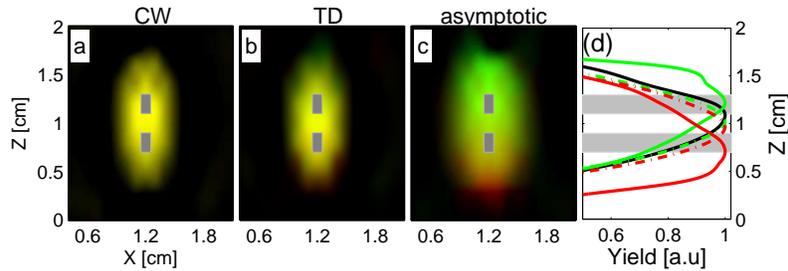


Fig. 6. Reconstructions for targets located axially, i.e., along the S-D axis, as shown in Fig. 2(b). The X-Z slice of the 3-D reconstructions are shown for (a) CW (b) lifetime-based direct TD and (c) lifetime-based asymptotic reconstructions. The colormap scheme used is the same as in Fig. 5. (d) shows quantitative plots of the yield along the depth Z, for the fixed X location of the inclusions. (CW - black line; direct TD - 1ns, dashed-dot red and 1.5ns, dashed-dot green; asymptotic - 1ns, solid red and 1.5ns, solid green.)

shows a dramatic improvement in the CNR and FWHM over the asymptotic reconstruction, and an improvement over the resolution of the CW case. The CNR improvement is evidently due to the better SNR of the peak portion of the DFTR compared to the asymptotic tail. The FWHM improvement of the TD over CW is due to the tomographic separation of the yield distributions for the lifetime components. Thus, for fixed CNR, the lifetime based TD reconstruction will have superior resolution compared to the asymptotic and CW reconstructions. However, the cross-talk, (which is the reconstructed amplitude of the 1.5ns inclusion at the location of the 1ns inclusion) is significantly higher for TD than the asymptotic case, and is attributable to the non-diagonal nature of the TD portion of the forward problem in Eq. (14). We note that the crosstalk for the asymptotic approach will depend on the separability of the lifetimes from the multi-exponential fits of raw experimental data, an aspect that will be explored in future work.

The effect of the crosstalk can be more clearly seen in the reconstructed tomographic images shown in Fig. 5, where the X-Z slices of the 3-D reconstructions for all three data sets and separations are displayed. Also shown are the plots of the yield at a fixed depth (Z) where the yield is maximum. It should be noted that the regularization λ was not identical for all the reconstructions but was rather determined by the condition that $\log_{10}(\text{CNR})$ was near 1. This is necessary to properly account for the difference in the noise characteristics of the different methods. (For example, CW has the best SNR, and should thus be the least regularized.) To visualize crosstalk easily, the yield images for the 1ns and 1.5ns components for the TD and asymptotic approaches were assigned red and green colors in an RGB colormap of a single image. The degree of crosstalk is thus revealed as a mixture of these two colors (e.g., yellow implies 100% crosstalk). Thus, CW reconstructions have no lifetime information so that they are shaded in yellow. It is clear from Fig. 5 that the TD reconstruction has superior resolution but significantly more cross-talk than the asymptotic reconstructions, as can also be seen in the intensity plots in the bottom panel. For small target separations, the cross talk of the TD method proves detrimental to its accuracy, whereas the asymptotic case recovers the localizations accurately even for 3mm separation. Thus it can be said that the direct TD approach using optimal time points provides more precise (better-resolved) reconstructions, and is useful when the targets are well separated, whereas the asymptotic reconstructions are more accurate but less precise (less-resolved). In Figure 6, the reconstructions are shown with the targets located axially, i.e., along the S-D axis. The advantage of the lifetime based asymptotic reconstruction is even more evident in this case, as the localizations of the two lifetimes are not reproduced either for the CW or the direct TD reconstructions.

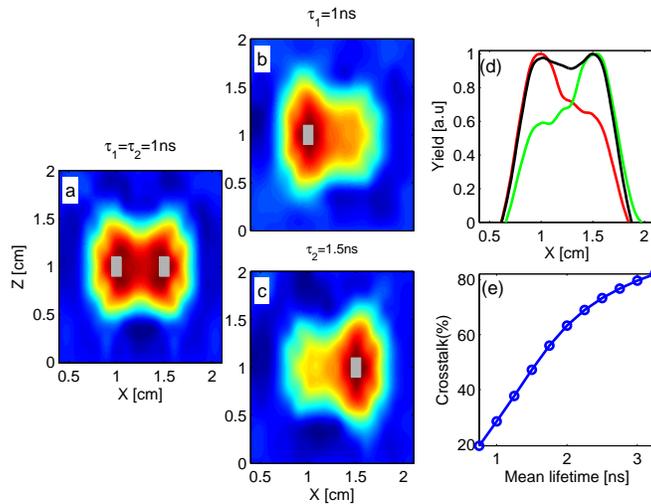


Fig. 7. Illustration of the enhancement of direct TD reconstructions in the presence of lifetime contrast. The colormap scheme used is direct and reflects the actual reconstructed yield, in contrast to that used in Figs (5) and (6). (a) Yield reconstructions with two inclusions separated by 7mm, with both having the same lifetime of 1ns and (b-c) yield reconstructions for inclusions with distinct lifetimes of 1ns and 1.5ns. (d) 1-D plot of the reconstructed yield along the X axis at the actual depth of the inclusions, for no lifetime contrast (black) and with the inclusions having 1ns (red) and 1.5ns (green) lifetimes. (e) Dependence of cross talk for the direct TD reconstructions on the mean lifetime. The two inclusions had a fixed lifetime separation of 0.5ns, while the mean lifetime was varied between 0.75ns and 3.25ns.

Although characterized by cross-talk, the presence of lifetime contrast enhances the images reconstructed using direct TD data. To delineate this effect clearly, Fig. 7 shows a comparison of the reconstructions of the two laterally placed targets using the direct TD approach, with and without lifetime contrast between the targets. The effect of lifetime contrast on the direct TD reconstruction is clearly seen as a significant reduction in the point-spread-function of the reconstructed yield distributions for the two lifetimes. Of course, this effect will depend on the difference between the fluorophore lifetimes and the diffuse propagation time τ_D , which is near 0.4ns for the present simulation. (corresponding to $\mu_a = 0.1$). As the mean lifetime becomes much larger than τ_D , the cross talk will also increase, diminishing the separability of the corresponding yield distributions. This is due to the fact that the elements of the first row of the TD weight matrix in Eq. (14) are almost identical for the early time points, when $\tau_n \gg \tau_D$. To study this quantitatively, the simulations in Fig. 7 (b) and (c) were repeated for a range of mean lifetimes, with fixed lifetime separation of 0.5ns and the crosstalk was estimated for each case. In Fig. 7(e), the cross talk of the direct TD approach is plotted as a function of the mean lifetime of the inclusions, indicating the large range of lifetimes for which direct TD reconstructions can benefit from lifetime contrast.

5. Conclusions

We have presented a theoretical formalism for TD fluorescence tomography with turbid media that allows a lifetime-based reconstruction of yield distributions using the entire TD data. Besides providing a comprehensive understanding of TD fluorescence signals from diffuse media, a key advance of this work from the previously presented asymptotic lifetime based tomog-

raphy is an algorithm for lifetime based tomographic separation using the peak and rise portions of the temporal diffuse fluorescence response. The formalism is generally valid provided the intrinsic fluorescence lifetimes are revealed in the long time decay, a condition well satisfied [11–13] given the typical optical response times of diffuse tissue and fluorophore lifetimes used in molecular imaging. This is important since the measured average lifetime on the surface can by itself provide useful diagnostic information, without the need for tomography. This shows the potential for lifetime sensing in diagnostic imaging and extends the application of this work to a wide range of biomedical imaging problems.

The results presented here can be viewed as an inevitable consequence of the long lifetime condition: The longer lived fluorescence decay effectively convolves over the intrinsic diffuse material response, resulting in a decay tail that is separated in space and time. This implies that for tomography with the long time data, a direct use of multiple time points in the decay portion is redundant. The optimal approach is to perform tomography using the amplitudes recovered from multi-exponential fits. This result was shown to be consistent with a SVD analysis of the time-dependent Born weight functions, which showed that the optimal time points to use in a direct TD reconstruction are located near the peak and rise portions of the DFTR.

Tomographic reconstructions with simulated noisy data also revealed the relative merits of optimized direct TD and the asymptotic approaches. It was found that the direct TD and asymptotic approaches yield complementary information: The asymptotic approach provides superior localization ability due to minimal cross-talk between the images for multiple lifetimes, while the optimal direct TD reconstructions yield better resolution due to superior SNR near the peak of the DFTR. Thus, when no lifetime contrast is present in the medium, the direct TD analysis should be the method of choice. For targets located along the S-D axis, it was shown that the asymptotic analysis is superior to both direct TD and CW in its ability to accurately localize the targets, provided they have different lifetimes. Axially located targets could occur for example in small animal brain imaging, where transillumination may be the preferred geometry when depth resolution along the various brain regions is desired.

The simulations presented here considered two distinct lifetime inclusions placed both lateral and axial to the measurement axis. Although simplistic, this example has highlighted key aspects of TD fluorescence tomography with lifetime contrast that can potentially be extended to more complicated spatial distributions of lifetimes. This work is also based on the assumption that lifetimes present in the medium are few and discrete, or at least can be described as sharp distributions centered around a mean lifetime. In the more general case when lifetimes are broadly distributed, a numerical inverse Laplace transform can be used to recover amplitude distributions [26]. The simulation analysis presented here was not meant to optimize for any particular TD detection technique (e.g., wide-field time gated, time correlated detection schemes), but was rather an attempt to explore the information content in a TD signal and to provide a recipe for an optimal approach for TD fluorescence tomography with turbid media. The purpose of the numerical simulations was also not to make a statement about the absolute resolution achievable by TD methods. This quantity can be optimized using better S-D arrangement and adjusting actual experimental conditions. Indeed, sub-mm resolution has recently been demonstrated using CW measurements [3]. Such optimization will enhance the resolution of all three approaches viz., CW, direct TD and asymptotic, so that the main results obtained here will not be affected.

In future work, we will attempt to extend the formalism developed here to a hybrid model that incorporates both the direct TD and the asymptotic approaches into a single inverse problem in a self-consistent fashion. This hybrid TD-asymptotic approach is expected to provide optimal localization and resolution. It should be reiterated that although a diffusive slab model was assumed for the simulations, the formalism developed here can readily incorporate Green's

functions calculated as solutions of either the diffusion or transport equations, as appropriate, for heterogenous media with complex boundaries. We are currently engaged in applying the formalism developed here to imaging complex shaped mouse phantoms and mouse models of disease.

Optical molecular imaging can immensely benefit from the use of biochemical reporter probes that are not merely disease specific, but also provide specific molecular signatures such as spectral and lifetime shifts that help isolate the disease from background tissue [2]. The unique advantages of time domain technology as explored in this work strongly motivates the development of fluorescent contrast agents that exhibit target specific lifetime shift upon binding. We also hope that this work will provide useful guidelines for biological imaging using time domain fluorescence tomography.

Acknowledgments

This work was supported by National Institutes of Health Grants EB000768, AG026240 and P41-RR14075.