Optics Letters

Optimal estimator for tomographic fluorescence lifetime multiplexing

STEVEN S. HOU,[1,2] BRIAN J. BACSKAI,[2] AND ANAND T. N. KUMAR[1,*]

[1]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts 02129, USA
[2]Alzheimer's Disease Research Unit, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts 02129, USA
*Corresponding author: ankumar@nmr.mgh.harvard.edu

Received 4 January 2016; revised 16 February 2016; accepted 16 February 2016; posted 18 February 2016 (Doc. ID 256639); published 17 March 2016

We use the model resolution matrix to analytically derive an optimal Bayesian estimator for multiparameter inverse problems that simultaneously minimizes inter-parameter cross talk and the total reconstruction error. Application of this estimator to time-domain diffuse fluorescence imaging shows that the optimal estimator for lifetime multiplexing is identical to a previously developed asymptotic time-domain (ATD) approach, except for the inclusion of a diagonal regularization term containing decay amplitude uncertainties. We show that, while the optimal estimator and ATD provide zero cross talk, the optimal estimator provides lower reconstruction error, while ATD results in superior relative quantitation. The framework presented here is generally applicable to other multiplexing problems where the simultaneous and accurate relative quantitation of multiple parameters is of interest.

OCIS codes: (170.3010) Image reconstruction techniques; (170.3650) Lifetime-based sensing; (170.3880) Medical and biological imaging; (170.6920) Time-resolved imaging; (170.6960) Tomography.

http://dx.doi.org/10.1364/OL.41.001352

Optical techniques offer the unique ability for tracking multiple disease markers or pathways simultaneously *in vivo* (also termed as "multiplexing") through a variety of intrinsic and extrinsic contrast mechanisms. These mechanisms include spectral contrast between oxy- and deoxyhemoglobin to quantify tissue oxygenation [1], multispectral or fluorescence lifetime (FL) labeling of intrinsic or extrinsic fluorophores [2–5], and intrinsic tissue absorption and scattering contrast [6]. The ability to capitalize on these contrast mechanisms strongly relies on inversion algorithms that can efficiently localize and quantify multiple optical parameters from tomography measurements. An important quantity that affects the quality of reconstructions in multiplexing is inter-parameter cross talk [7], which can be defined as the influence of one parameter in the spatial location of another parameter. Cross talk can be a significant problem for multiplexing applications and can lead to errors in localization and quantification of the reconstructions. For instance, in tomographic fluorescence lifetime multiplexing (TFLM), cross talk between distinct lifetimes results in poor quantitation and spatial localization of fluorophores located within a few mm, especially when using the early time points of time-domain (TD) fluorescence data [7,8] or with frequency domain inversion [9].

Inversion methods in optical imaging have so far primarily focused on minimizing reconstruction error but have not systematically addressed cross-talk minimization. In this Letter, we present a novel methodology to address cross talk in ill-posed, multiparameter inverse problems with special attention to FL multiplexing using TD measurements. The methodology employs a Bayesian formulation with a zero cross-talk constraint in addition to minimizing the mean square error (MSE) cost function. Cross talk is quantitatively represented by the off-diagonal blocks of the model resolution matrix (the columns of which are the imaging point-spread functions [PSFs]). While the traditional approach effectively minimizes for MSE, our approach focuses on achieving zero cross talk at the expense of a higher MSE. Application of the optimal estimator for the specific case of TFLM leads to rigorous statistical generalization of a previously derived asymptotic time domain (ATD) approach for tomographic FL multiplexing. We present simulations to compare the cross talk and error performance of the optimal estimator with both the ATD approach and standard reconstruction algorithms for TD imaging and show that the optimal approach and ATD provide better localization and relative quantitation compared with the standard MSE-based approach.

We address the problem in the context of FL multiplexing, although the results are applicable to any linear multiparameter inverse problem. Consider a turbid medium of volume Ω with N fluorophores of distinct lifetimes τ_n and yield distributions $\eta_n(r)$, $r \in \Omega$. Discretizing the medium into V voxels, the TD fluorescence signal for L time points and M pairs of sources and detectors located at points r_s and r_d on the boundary can be represented by the matrix equation:

$$y = W\eta, \tag{1}$$

where y is a $(\mathrm{ML} \times 1)$ measurement vector, $W = [W_1, \ldots, W_N]$ is a $(\mathrm{ML} \times NV)$ TD weight matrix, and $\eta = [\eta_1, \ldots, \eta_N]^T$ is a $(NV \times 1)$ vector of unknown fluorescence yields of all lifetimes. As previously discussed [8], $W_n = G^x(r_s, r, t) * \exp(-t/\tau_n) * G^m(r, r_d, t)$, where G^x and G^m are the excitation and emission Green's functions for light transport in the turbid medium. The central quantities of interest are the fluorescence yield distributions, $\eta_n(r)$, of each fluorophore with the assumption that the discrete lifetimes τ_n are either known or are independently retrieved through global analysis methods [3]. A straightforward solution to the above linear problem, which we refer to as the direct TD (DTD) approach, uses Tikhonov inversion of Eq. (1), resulting in a reconstructed yield distribution:

$$\hat{\eta}_{\mathrm{DTD}} = \widehat{W}_{\mathrm{DTD}} y, \tag{2}$$

where the inverse operator $\widehat{W}_{\mathrm{DTD}}$ is given by the standard Tikhonov expression with regularization parameter λ:

$$\widehat{W}_{\mathrm{DTD}} = W^T (W W^T + \lambda I)^{-1}. \tag{3}$$

The DTD inversion can be interpreted statistically by using the well-known connection between Tikhonov regularization and Bayesian inversion [10]. Given common assumptions that the measurement noise, n, and unknown yield, η, can be modeled as white Gaussian random vectors, η_{DTD} is equivalent to the minimum mean square error (MMSE) solution if the regularization parameter is chosen as $\lambda = (\sigma_n/\sigma_\eta)^2$, where σ_n and σ_η are the variances of the measurement noise and yield, respectively. Although providing an MMSE solution, DTD does not explicitly restrict the cross talk between the η_n's and can lead to severe cross talk between yield distributions of distinct lifetimes. The cross talk translates into poor spatial localization and quantitation [7] (see Fig. 2). The DTD approach is, thus, not the method of choice for multiplexing applications where the relative amounts of multiple overlapping parameters (which, in the present case, are the yield distributions) are of interest.

To systematically incorporate inter-parameter cross talk into the inverse problem, we consider the model resolution matrix, which takes the following form for a general linear imaging system with forward operator W and inverse operator (estimator) $\widehat{W}$:

$$R = \widehat{W} W. \tag{4}$$

For single parameter problems, R can be interpreted in the standard way [10], i.e., the c'th column of R represents the point-spread function for the c'th voxel, and the j'th row represents the contribution at the j'th voxel due to all the other voxels in the medium. To see the usefulness of the resolution matrix for the multiparameter case, we consider the case of two lifetimes, τ_1 and τ_2, with corresponding yield distributions, η_1 and η_2, although the results can be readily generalized to any number of parameters. Both W and $\widehat{W}$ can be split into two separate submatrices for each FL:

$$W = \begin{bmatrix} W_1 & W_2 \end{bmatrix}, \qquad \widehat{W} = \begin{bmatrix} \widehat{W}_1 \\ \widehat{W}_2 \end{bmatrix}, \tag{5}$$

where W_1 and W_2 are each of dimension $\mathrm{ML} \times V$. The resolution matrix R is then a block matrix with four quadrants:

$$R = \begin{bmatrix} \widehat{W}_1 W_1 & \widehat{W}_1 W_2 \\ \widehat{W}_2 W_1 & \widehat{W}_2 W_2 \end{bmatrix} \equiv \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}. \tag{6}$$

The diagonal blocks, R_{11} and R_{22}, can be interpreted similar to the single fluorophore-resolution matrix, as described above. The off-diagonal blocks have special significance for multiplexing. The columns of R_{12} can be interpreted as the cross talk into the η_1 distribution from a point inclusion with FL τ_2 (and vice versa for R_{21}). Because R_{12} and R_{21} provide a complete and quantitative measure of the intuitive notion of cross talk, we directly incorporate these terms into the optimization problem.

We next employ a novel Bayesian inversion algorithm to derive a general linear estimator, termed the cross-talk constrained TD (CCTD) estimator, which provides optimal cross-talk performance and minimal reconstruction error. This is achieved by finding an MMSE solution with an imposed zero-cross-talk constraint on R_{12} and R_{21}. Let the first- and second-order moments for noise, n and η be given by $E[n] = 0$, $E[\eta] = 0$, $\mathrm{cov}(n) = C_n$, and $\mathrm{cov}(\eta) = C_\eta$. The optimization problem takes the form:

$$\widehat{W}_{\mathrm{CCTD}} = \arg \min_{\widehat{W}} E[\|\eta - \hat{\eta}\|^2], \tag{7}$$

with the constraints:

$$R_{12} = 0 \quad \text{and} \quad R_{21} = 0. \tag{8}$$

Proceeding to solve Eqs. (7) and (8), we recognize that the optimization problem can be categorized as a quadratic programming problem with linear equality constraints. The detailed steps in the derivation will be presented elsewhere, but, briefly, we eliminate the constraints using a null space method [11] and solve the resulting unconstrained optimization problem using standard methods. The final result for the CCTD estimator is found to be ($j = 1, 2$):

$$\widehat{W}_j = C_{\eta_j} W_j^T N_j [N_j^T (W_j C_{\eta_j} W_j^T + C_n) N_j]^{-1} N_j^T, \tag{9}$$

where $N_j = \mathrm{null}(W_j^T)$ is the matrix whose columns span the null space of W_j^T, and C_{η_j} is the covariance matrix for the unknowns η_j. It is clear from Eq. (9) that a necessary condition for the nontrivial solution for the optimal estimator is that $N_j \neq 0$ for all j, i.e., the matrices W_j^T must have a nonzero null space. We label this as the "nullity condition for multiplexing" (NCM). Because no assumptions about the nature of W_1 and W_2 have been made thus far, the solutions in Eq. (9) and the NCM are applicable for general multiparameter inverse problems and not just for FL multiplexing.

We next apply the general estimator in Eq. (9) to the TD fluorescence forward problem [Eq. (1)]. It is first clear that, for overdetermined problems ($\mathrm{ML} > V$), the NCM is readily satisfied, because, in this case, $\mathrm{rank}(W_j) = V = \mathrm{rank}(W_j^T)$ so that by the rank-nullity theorem, $\mathrm{nullity}(W_j^T) = \mathrm{ML} - \mathrm{rank}(W_j^T) = \mathrm{ML} - V > 0$. For the underdetermined case ($\mathrm{ML} < V$), which is more common in tomography applications, the NCM is not generally satisfied when W_j is full rank because $\mathrm{rank}(W_j) = \mathrm{ML} = \mathrm{rank}(W_j^T)$, so that $\mathrm{nullity}(W_j^T) = \mathrm{ML} - \mathrm{rank}(W_j^T) = 0$. However, when $\tau_n > \tau_D$, where τ_D is the time scale for light diffusion in the medium [8], we can show that NCM is satisfied in the asymptotic region of the TD signal, defined for $t \gg \tau_D$. In this region, the TD weight matrix factorizes into a product of purely spatial and temporal terms [7,8]:

$$W = A\overline{W}, \tag{10}$$

where $A = [\exp(-t/\tau_1) \otimes I, \exp(-t/\tau_2) \otimes I]$ is a $(\mathrm{ML} \times 2M)$, overdetermined, basis matrix containing Kronecker products of exponential decay terms and the $(M \times M)$ identity matrix, I, and $\overline{W} = \mathrm{diag}(\overline{W_1}, \overline{W_2})$ is a $(2M \times 2V)$ block diagonal matrix containing CW weight matrices for each FL evaluated at a reduced medium absorption [8]. The factorization in Eq. (10) also holds for the individual FL matrices, so that $W_j = A_j \overline{W}_j$. Given this factorization, we have $\mathrm{nullity}(W_j^T) = \mathrm{nullity}(\overline{W}_j^T A_j^T)$. Because A_j is overdetermined, A_j^T is underdetermined with $\mathrm{nullity}(A_j^T) = \mathrm{ML} - \mathrm{rank}(A_j^T) = \mathrm{ML} - M > 0$, provided $L > 1$, i.e., more than one time point is included in the measurement. The nonzero nullity of A_j^T ensures that the product $W^T = \overline{W}_j^T A_j^T$ also has a nonzero null space (i.e., nontrivial solutions z exist that satisfy $\overline{W}_j^T A_j^T z = 0$), thus satisfying the NCM. In other words, the factorization in the asymptotic region in Eq. (10) splits the underdetermined matrix W_j into a product of the overdetermined A_j and underdetermined $\overline{W}_j$, thereby ensuring the nonzero nullity of the transpose, W_j^T. Below (Fig. 1), we will numerically illustrate the transition of the TD weight matrix from zero nullity (i.e., NCM not satisfied) at early time points to a nonzero value (NCM satisfied) as the asymptotic region is approached.

We can now obtain the CCTD estimator for the TD problem by exploiting the factorization in the asymptotic region. Substituting Eq. (10) into Eq. (9) and assuming model and data covariance matrices, $C_\eta = \sigma_\eta^2 I$ and $C_n = \sigma_n^2 I$, the CCTD estimator for TFLM takes the form:

$$\widehat{W}_{\mathrm{CCTD}} = \overline{W}^T \left(\overline{W}\,\overline{W}^T + \lambda \frac{\mathrm{diag}(\mathbf{C}_a)}{C_n} \right)^{\dagger} A^{\dagger}, \qquad (11)$$

where $\mathbf{C}_a = (A^T A/\sigma_n^2)^{-1}$, $A^\dagger$ is the Moore–Penrose pseudo-inverse [12] of the well-conditioned matrix A, $\mathrm{diag}(X)$ sets all off-diagonal blocks of a matrix X to zero and $\lambda = (\sigma_n/\sigma_\eta)^2$. Equation (11) is the central result of this Letter and offers a compact expression for a novel estimator for tomographic FL multiplexing that achieves zero cross talk between multiple lifetimes while also minimizing MSE. The matrix $\mathbf{C}_a$ is immediately recognized from linear regression theory [12] as the covariance matrix for the decay amplitudes in a linear multiexponential analysis with basis functions A. The diagonal terms of $\mathbf{C}_a$ are the uncertainties of each amplitude, while the off-diagonal terms correspond to the covariances between the amplitudes for distinct lifetimes. Equation (11), therefore, has the remarkable interpretation that the optimal estimator achieves zero cross talk by simply setting the off-diagonal elements of the decay amplitude covariance, $\mathbf{C}_a$, to zero. To further appreciate the significance of this result, we write the DTD inverse operator [Eq. (3)] in the asymptotic region. Using Eq. (10), Eq. (3) takes the alternate form:

$$\widehat{W}_{\mathrm{DTD}} = \overline{W}^T \left(\overline{W}\,\overline{W}^T + \lambda \frac{\mathbf{C}_a}{C_n} \right)^{-1} A^{\dagger}. \qquad (12)$$

Direct comparison of Eqs. (11) and (12) shows that the only difference between the optimal estimator and the DTD approach is that the optimal estimator sets the off-diagonal blocks of the matrix of the $\mathbf{C}_a$ to zero, while the DTD retains the full covariance of the decay amplitudes, thereby resulting in higher cross talk between multiple lifetimes.

It is also interesting to compare the CCTD estimator with a previously derived asymptotic TD (ATD) estimator for FL multiplexing [7,8]. The ATD approach capitalizes on the factorization in Eq. (10) to perform the tomographic recovery of the yield distribution in two steps. First, the forward problem in Eq. (1) takes the form $y = A\overline{W}\eta$ in the asymptotic region, using Eq. (10). We can then write:

$$A^\dagger y (= a) = \overline{W}\eta. \qquad (13)$$

In the above equation, $a = A^\dagger y$ simply represents the least-squares solution for recovering the decay amplitudes (a) from the TD data y using a multiexponential analysis [12]. In the second step, the standard Tikhonov regularization is applied to recover the yield distribution η from the vector of decay amplitudes a. The two steps of the ATD approach can be represented as a single estimator [7], $\widehat{W}_{\mathrm{ATD}}$:

$$\widehat{W}_{\mathrm{ATD}} = \overline{W}^T (\overline{W}\,\overline{W}^T + \lambda I)^{-1} A^\dagger. \qquad (14)$$

Thus, we see that the ATD approach is identical to the CCTD estimator to within a quantitative correction due to the amplitude uncertainties (diagonal elements on $\mathbf{C}_n$). The incorporation of the amplitude uncertainties in the regularization reduces the MSE of the CCTD approach relative to the ATD.

Finally, we compare the resolution matrices for DTD, ATD, and CCTD methods, given by $R_{\mathrm{method}} = \widehat{W}_{\mathrm{method}} W$. Using Eqs. (11), (12), and (14) in Eq. (6), we have

$$R_{\mathrm{DTD}} = \overline{W}^T (\overline{W}\,\overline{W}^T + \lambda \mathbf{C}_a / C_n)^{-1} \overline{W}, \qquad (15)$$

$$R_{\mathrm{CCTD}} = \overline{W}^T (\overline{W}\,\overline{W}^T + \lambda \mathrm{diag}(\mathbf{C}_a) / C_n)^{-1} \overline{W}, \qquad (16)$$

$$R_{\mathrm{ATD}} = \overline{W}^T (\overline{W}\,\overline{W}^T + \lambda I)^{-1} \overline{W}. \qquad (17)$$

Here, we have used the identity $A^\dagger A = I$, given A is full column rank. The key difference between the above three estimators resides in the fact that the resolution matrices for the CCTD and ATD estimators are fully block-diagonal (recall that $\overline{W}$ is block diagonal), therefore providing zero cross-talk for an arbitrary distribution of fluorophores. However, the DTD estimator has off-diagonal terms, resulting in significant cross talk and poor localization [7]. We note that, although ATD has previously been shown to reduce cross talk compared with DTD [7,8], Eq. (17) is the first rigorous demonstration that ATD provides zero cross talk, irrespective of the fluorophore distributions or measurement conditions.

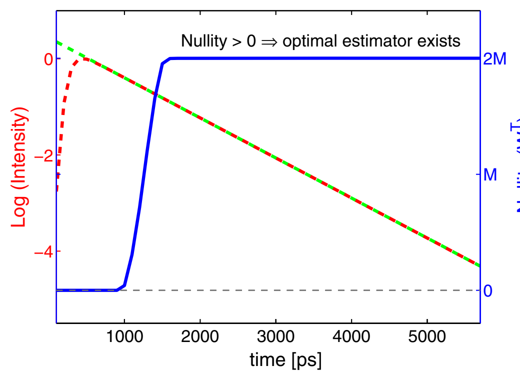

Fig. 1. Simulation to show that the asymptotic region of TD data satisfies the positive nullity condition [$\mathrm{nullity}(W_1^T) > 0$ for each FL component τ_j] for the existence of an optimal estimator in Eq. (11). $\mathrm{Nullity}(W_1^T)$ is plotted (blue, solid line, right axis) along with the diffuse fluorescence signal (red, dashed line) and a 1.2 ns pure exponential decay (green, dashed line). The nullity is zero for the early TD data but sharply rises to $2M$ (M = number of source-detector pairs) at the onset of the asymptotic region.

We next illustrate the theoretical results of the Letter using numerical simulations. First, we show in Fig. 1 the transition of the TD weight matrix from zero to positive nullity toward the asymptotic region where it satisfies the NCM condition. We consider a diffusive slab (2 cm × 2 cm × 2 cm, bulk absorption of $\mu_a = 0.6\ \text{cm}^{-1}$, reduced scattering of $\mu_s' = 10\ \text{cm}^{-1}$) with a single fluorophore inclusion (FL $\tau = 1.2$ ns) at the center, and with 42 sources and 42 detectors ($M = 1764$) arranged in a transillumination geometry and with three adjacent time points ($L = 3$) per measurement set. The time-dependent nullity of the single FL TD weight matrix, W_1^T, plotted in Fig. 1, is zero for early time points but rapidly transitions to $2M$ as the asymptotic region is approached. Thus, late time points in the TD fluorescence problem satisfy the NCM, thereby allowing optimal estimators with zero cross talk, while early time points do not allow for zero cross-talk estimators.

We next compare the performance of DTD, ATD, and CCTD methods for the diffuse medium used in Fig. 1, with voxel size of 1 mm^3 and with $\mu_a = 0.1\ \text{cm}^{-1}$ and $\mu_s' = 10\ \text{cm}^{-1}$. Two 1 mm^3 fluorescent inclusions of equal yield but distinct lifetimes of $\tau_1 = 0.8$ ns and $\tau_2 = 1.2$ ns were separated by 2 mm [Figs. 2(a)–2(f)] or 4 mm [Figs. 2(g)–2(l)] along the X axis, at depth $Z = 1$ cm. The forward TD data was generated using the Monte Carlo photon transport model [13] for 12 times points in the asymptotic TD region, with 100 ps separation, and with 3% Gaussian noise [equivalent to 12.73 and 12.81 dB for Figs. 2(a)–2(c) and 2(g)–2(i)]. The regularization parameter λ for each method was chosen to produce the lowest reconstruction error ($E = \|\eta_{\text{true}} - \eta_{\text{recon}}\|^2$). Figure 2 shows the X–Z slices of the DTD, ATD, and CCTD reconstructed yields for both lifetimes and the corresponding line plots of the yield along the true location of the inclusions. Figure 2 shows that, while both the ATD and CCTD accurately localize the inclusions for both the 2 and 4 mm cases, the DTD reconstructions show significant cross talk, which leads to poor localization, with the 2 mm separated inclusions nearly indistinguishable. While DTD provides the least reconstruction error (e.g., for the 4 mm case, $E = 1.9983$ compared with 1.9989 for ATD and 1.9987 for CCTD), it resulted in an average cross talk of 55.5% [calculated as $(R_{21}(j_1, j_1)/R_{11}(j_1, j_1) + R_{12}(j_2, j_2)/R_{22}(j_2, j_2))/2$, where j_1 and j_2 are the linear indices corresponding to the locations of the two inclusions], while ATD and CCTD provide zero cross talk. The high cross talk of DTD results in incorrect relative quantitation, leading to a nearly 2∶1 ratio of the yields for both cases compared with the true ratio of 1∶1. Both ATD and CCTD provided better relative quantitation than DTD. The CCTD estimator provides lower reconstruction error, at the expense of poorer relative quantitation, thus resulting in a slightly higher reconstructed yield for the 1.2 ns FL [(green line, Figs. 2(f) and 2(l)]. This can be attributed to the distinct regularization level for each FL component in the diagonal covariance matrix diag($\mathbf{C}_a$) in Eq. (11).

We have presented a novel framework for tomographic multiplexing problems, based on model resolution matrix constraints, to derive an optimal estimator that provides zero inter-parameter cross talk with guaranteed minimal error. We also showed that, for tomographic TD fluorescence data, the optimal estimator exists only in the asymptotic region, where it offers a rigorous statistically generalized version of a previously developed ATD approach. The resolution matrix constraint-based framework presented here can be readily applied to reduce inter-parameter cross talk in other multiplexing applications such as multispectral imaging and absorption-scattering contrast in diffuse optical imaging.

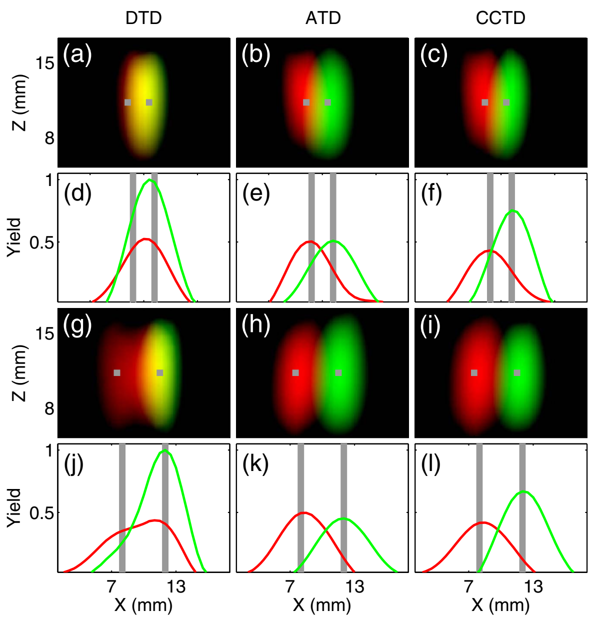

Fig. 2. Tomographic reconstructions to compare the DTD, ATD, and CCTD estimators for FL multiplexing of 0.8 and 1.2 ns fluorophore inclusions (filled gray) at the center in a 2 cm thick turbid medium. The inclusions are separated by 2 mm in (a)–(f) and 4 mm in (g)–(l). The X–Z slices of reconstructed yields of 0.8 and 1.2 ns are shown for DTD [(a), (g)], ATD [(b), (h)] and CCTD [(c), (i)] as the red (0.8 ns) and green (1.2 ns) components of a single RGB image. Yellow indicates cross talk. (d)–(f) and (j)–(l) show the corresponding line plots of the yield with the gray bars indicating the true locations.

Funding. National Institute of Biomedical Imaging and Bioengineering (NIBIB) (R01 EB000768, R01 EB015325).

REFERENCES

1. D. Boas, D. H. Brooks, E. L. Miller, C. DiMarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang, IEEE Signal Process. Mag. **18**(7), 57 (2001).
2. R. M. Levenson, D. T. Lynch, H. Kobayashi, J. M. Backer, and M. V. Backer, ILAR J. **49**, 78 (2008).
3. S. B. Raymond, D. A. Boas, B. J. Bacskai, and A. T. N. Kumar, J. Biomed. Opt. **15**, 046011 (2010).
4. V. Venugopal, J. Chen, M. Barroso, and X. Intes, Biomed. Opt. Express **3**, 3161 (2012).
5. W. L. Rice, D. M. Shcherbakova, V. V. Verkhusha, and A. T. Kumar, Cancer Res. **75**, 1236 (2015).
6. S. R. Arridge and J. C. Schotland, Inverse Prob. **25**, 123010 (2009).
7. S. S. Hou, W. L. Rice, B. J. Bacskai, and A. T. Kumar, Opt. Lett. **39**, 1165 (2014).
8. A. T. Kumar, S. B. Raymond, G. Boverman, D. A. Boas, and B. J. Bacskai, Opt. Express **14**, 12255 (2006).
9. A. T. Kumar, S. B. Raymond, B. J. Bacskai, and D. A. Boas, Opt. Lett. **33**, 470 (2008).
10. M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging* (CRC Press, 1998).
11. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University, 2004).
12. W. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University, 1992).
13. D. Boas, J. Culver, J. Stott, and A. Dunn, Opt. Express **10**, 159 (2002).