1

# Second order adjoint sensitivities for fluorescence optical tomography based on the $SP_N$ approximation

NISHIGANDHA PATIL<sup>1,3</sup> AND NAREN NAIK<sup>1,2,4</sup>

<sup>1</sup>Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur-208016, INDIA <sup>2</sup>Center for Lasers and Photonics, Indian Institute of Technology Kanpur, Kanpur-208016, INDIA <sup>3</sup>nipat@iitk.ac.in <sup>4</sup>nnaik@iitk.ac.in

Compiled May 9, 2019

Use of second order sensitivity information has been shown in the literature to yield faster convergence, better noise tolerance and localisation besides enhanced post-reconstruction analysis capabilities. In this paper we derive adjoint based second-order derivatives for  $SP_N$  approximation modeled fluorescence optical tomography. We modify the regularizing Levenberg-Marquardt method to use the second-order sensitivity information through a predictor-corrector framework. Reconstruction studies presented for the fluorophore absorption coefficient in low as well as high scattering tissue-mimicking phantoms in both, ideal and differential fluorophore-uptake settings show consistently superior noise tolerance and contrast recovery with the second order scheme as compared to its first-order counterpart. © 2019 Optical Society of America

http://dx.doi.org/10.1364/ao.XX.XXXXXX

#### 1. INTRODUCTION

The inverse problem in optical tomography, to recover an optical parameter 'p', is often cast as a residual minimisation problem, or a model-fitting problem involving a non-linear system of equations. The central idea in derivative based approaches to solve this problem is to model the residual (or the cost function) in terms of its first order (FO) or second order (SO) Taylor series approximation about the current iterate  $p^{(k)}$ . Several works in the literature make use of the FO derivative [1–3] or at most approximations to the SO derivative [4]. Two reasons are often cited in support of the general reluctance in using SO derivatives. The first is that close to the solution, the small residual approximation [5] or the assumption of mild non-linearity [3, 6] can be used which renders the contribution of the higher order terms negligible. The second is that the evaluation of the SO derivatives is computationally expensive.

However, studies in the literature [7–9] suggest that there may be more to gain by dropping the mild non-linearity and small residual assumptions. Convergence issues related to the use of FO derivative based schemes in such scenarios are well established in the literature [3, 7, 10]. When the initial guess is not too close to the actual solution, the contributions from the second order terms are significant, especially in the initial iterates when the residual is large [8]. Previous works have shown that the use of the SO derivative or approximations to it, offers several advantages in the solution of the inverse problem, namely, faster convergence [4, 9, 11], better contrast recovery [8], and in some cases better reconstructions [4, 9, 12]. In [4], Klose and Hielscher demonstrate that the Quasi Newton (QN) schemes exhibit early convergence as compared to the Conjugate Gradient (CG) method using FO derivatives. For the fluorescence diffuse optical tomography problem, Roy and Muraca [7] demonstrate the use of a truncated Newton scheme. They directly compute the required Hessian-vector products using a finite difference scheme over the gradient. A second order scheme for solving the inverse problem using a predictor-corrector approach, developed by Hettlich and Rundell [9] (referred to as the second degree method by the authors, and also by us in the rest of the text), has been used by Kanmani and Vasu [8] in a non-linear CG framework, which afforded better contrast recovery of the absorption coefficient in the diffusion optical tomography problem. In [11], Kress and Lee observe that this second degree method allows flexibility in choice of initial guess as well as the regularisation parameter in an inverse obstacle scattering problem.

The second degree method of Hettlich and Rundell in essence attempts to improve on the FO predicted update by making use of the SO information in the corrector. A similar idea has been proposed by Transtrum and Setna in [5] to improve the performance of the Levenberg-Marquardt scheme in certain situations. They prescribe a SO correction for the FO update which significantly improves the success rate and fit quality in the broad class of optimisation problems studied by them. In a more recent work, Gould et al [13] describe a higher order method which makes use of SO derivatives of the residual for solving nonlinear least squares problems. They observe that the higher order scheme is more robust in terms of the choice of initial guess and requires fewer function evaluations on average in the test cases considered by them. While these studies suggest that considerable gains may be afforded by use of SO schemes in the reconstruction problem, the chief impediment to their use is the computational effort required to evaluate the SO derivatives. The development of a computationally efficient scheme for evaluation of the exact SO derivatives, thus, has potential to improve the quality of reconstructions in optical tomography through facilitating the use of full Newton type schemes. In addition, the availability of the SO sensitivities, makes it feasible to perform numerical post-reconstruction analysis to check the iterates for convergence [14].

The main contributions of the present paper are (i) the explicit evaluation of the SO adjoint sensitivities of exiting partial current for the  $SP_N$  approximation modeling fluorescent radiation transfer, and (ii) demonstration of the efficacy of using a regularizing Levenberg-Marquardt method modified to use SO sensitivity information through a predictor-corrector framework, to solve the fully non-linear fluorescence optical tomography (FOT) inverse problem.

For a tomographic problem with  $N_d$  detectors,  $N_s$  sources and  $N_p$  parameters, the evaluation of the FO and SO sensitivity requires  $O(N_pN_s)$  and  $O((N_pN_s)^2/2)$  large scale computations (forward solves) respectively using a finite difference scheme. The FO sensitivity can be efficiently computed by using adjoints with only  $O(N_d + N_s)$  large scale computations (forward + adjoint solves) [15, 16]. This is useful when  $N_p$  is much larger than the total number of measurements  $N_dN_s$ . One approach to evaluate the SO sensitivity is to compute it as the first difference of the FO adjoint sensitivity at the additional cost of  $O(N_p(N_d + N_s))$ computations. This scheme is particularly amenable to evaluation of Hessian-vector products [7]. An alternate approach amenable to parallelization is to evaluate the SO sensitivity directly using an adjoint based scheme; this is the focus of the present work.

Adjoint based schemes have been popularly used in the evaluation of the FO sensitivities ([1, 15, 16] and references therein). In the field of reactor physics a framework for the SO adjoint perturbation theory has been developed by Greenspan [17] for time-independent and by Gilli et al [18] for time-dependent problems. Using a scheme similar to [18], SO adjoint sensitivities have also been presented by Kanmani and Vasu [8] for the elastic scattering problem under the diffusion approximation (DA). A general theory for development of SO adjoint sensitivity for large-scale linear systems with mixed boundary conditions is presented by Cacuci in [19]. In this scheme, the SO sensitivity is defined as the first variation of the FO sensitivity. Two additional adjoints are then defined that allow the efficient computation of the SO sensitivity at the cost of  $O(N_p(N_d + N_s))$  large scale computations.

We use the scheme developed by Cacuci [19] to derive an explicit expression for the SO adjoint sensitivities. While SO adjoint sensitivities have been derived in earlier works relevant to the elastic scattering problem under the DA [8, 20], to the best of our knowledge, no such works exist for the FOT problem under the  $SP_N$  approximation. Besides the forward model employed, our scheme differs from that of Kanmani and Vasu [8] in the nature of the adjoints employed, since the parameter variations on the boundary are accounted for in our work, while they are neglected in [8]. The  $SP_3$  approximation has wider applicability

[16] than the more popular DA and requires minimal increase in computation over it [21]. It is also computationally less demanding than solving the full radiative transfer equation and its higher order approximations while being more accurate than the DA [21]. In [22], Lu et al have demonstrated that the  $SP_N$  approximation is better suited than the DA for the linearised problem in FOT. Considerable reduction in the computational time has been demonstrated with a vectorized implementation of the adjoint scheme for evaluating the FO sensitivities in [15, 16]. Vectorization in MATLAB refers to the process of modifying loops in the code to use matrix and vector operations efficiently. Since MATLAB is optimised for using matrices, an appropriate vectorization strategy can lead to considerable speedup in implementation time. In our work, vectorized implementations for global FEM matrix assembly as well as FO adjoint sensitivity computations as proposed in [15], are used to speed up the SO adjoint sensitivity computations wherever possible; albeit at the cost of memory.

To exhibit the feasibility of using the evaluated sensitivities, we consider the FOT inverse problem to reconstruct the fluorophore absorption coefficient in the medium. We use a regularizing Levenberg-Marquardt (LM) method [23], [24], [16] modified to make use of the second derivative information through the predictor-corrector approach of Hettlich and Rundell [9]. We contrast the reconstructions so obtained with a scheme using a conventional first-derivative based regularizing LM method. Numerical results are presented for phantoms in low scattering and high scattering settings for noiseless and noisy data sets with both ideal (no background fluorescence) and differential (with background fluorescence) uptake of fluorophore. Validation of the adjoint-based SO sensitivities with respect to finite differences for the elastic scattering problem in optical tomography has been reported by us earlier in [25].

This manuscript is structured as follows, in section 2 we describe the forward model used. A detailed derivation of the SO adjoint sensitivity for the FOT Problem is presented in section 3. The finite element method calculations and vectorized implementation are given in section 4. We briefly describe the inverse problem and discuss the numerical studies in section 5 and section 6 has the conclusions. The appendix contains the details of the FEM calculations for the additional adjoint sources used in our formulation.

#### 2. THE SP3 APPROXIMATION

The  $SP_3$  approximation to the coupled radiative transfer equation modeling the generation and propagation of fluorescent radiation is given in matrix form as [16]:

$$-\nabla \cdot C^{\nabla} \nabla \phi + C \phi = 0$$
 (1)

$$C^{\nabla b}(n \cdot \nabla \varphi) + C^b \phi = C^S$$
<sup>(2)</sup>

where

$$C^{\nabla} = \begin{pmatrix} C^{\nabla x} & 0 \\ 0 & C^{\nabla m} \end{pmatrix} \quad C = \begin{pmatrix} C^{x} & 0 \\ -C^{\beta} & C^{m} \end{pmatrix} \quad C^{b} = \begin{pmatrix} C^{bx} & 0 \\ 0 & C^{bm} \end{pmatrix}$$
$$C^{\nabla b} = \begin{pmatrix} C^{\nabla bx} & 0 \\ 0 & C^{\nabla bm} \end{pmatrix} \quad n \cdot \nabla \phi = \begin{pmatrix} \underline{n} \cdot \underline{\nabla} \phi^{x} \\ \underline{n} \cdot \underline{\nabla} \phi^{m} \end{pmatrix} \quad \nabla = \begin{pmatrix} \underline{\nabla} & 0 \\ 0 & \underline{\nabla} \end{pmatrix}$$
$$C^{S} = \begin{bmatrix} C^{Sx}, 0 \end{bmatrix}^{T} \quad \phi = \begin{bmatrix} \phi^{x}, \phi^{m} \end{bmatrix}^{T} \quad n = \text{diag}\left([\underline{n}, \underline{n}]\right)$$

Here '*x*, *m*' denote quantities at excitation and emission wavelength respectively.  $\phi^{x/m} = [\phi_1^{x/m}, \phi_2^{x/m}]^T$  denotes the composite moments of radiance [21]. The 2 × 2 coefficient matrices depend on the modulation frequency of the source  $\omega$  (Hz) as well as the optical properties of the medium and fluorophore namely the intrinsic absorption coefficient  $\mu_{ai}^{x/m}$ (cm<sup>-1</sup>), the scattering coefficient  $\mu_s^{x/m}$ (cm<sup>-1</sup>), the fluorophore absorption coefficient  $\mu_{af}^{x/m}$ (cm<sup>-1</sup>), the fluorescent yield  $\eta$ , the fluorescence lifetime  $\tau$ (ns), the anisotropy factor *g* and the refractive index of the medium  $n_m$  and are defined as,

$$C^{\nabla x/m} = \begin{pmatrix} \frac{1}{3\mu_{a1}^{x/m}} I & 0\\ 0 & \frac{1}{7\mu_{a3}^{x/m}} I \end{pmatrix} C^{\nabla bx/m} = \begin{pmatrix} \frac{1+B_1}{3\mu_{a1}^{x/m}} & -\frac{D_1}{\mu_{a3}^{x/m}}\\ -\frac{D_2}{\mu_{a1}^{x/m}} & \frac{1+B_2}{7\mu_{a3}^{x/m}} \end{pmatrix}$$
$$C^{x/m} = \begin{pmatrix} \mu_{a0}^{x/m} & -\frac{2}{3}\mu_{a0}^{x/m}\\ -\frac{2}{3}\mu_{a0}^{x/m} & (\frac{4}{9}\mu_{a0}^{x/m} + \frac{5}{9}\mu_{a2}^{x/m}) \end{pmatrix} I = \text{diag}([1,1])$$
$$C^{bx/m} = \begin{pmatrix} (\frac{1}{2}+A_1) & -(\frac{1}{8}+C_1)\\ -(\frac{1}{8}+C_2) & (\frac{7}{24}+A_2) \end{pmatrix} C^{\beta} = \begin{pmatrix} \beta & -\frac{2}{3}\beta\\ -\frac{2}{3}\beta & \frac{4}{9}\beta \end{pmatrix}$$
with  $\mu_{a1}^{x/m} = \mu_{a1}^{x/m} + \mu_{af}^{x/m} + (1-g^n)\mu_s^{x/m} + \frac{i\omega}{c}$ 
$$\beta = \frac{\eta\mu_{af}^x}{1-i\omega\tau}, \ \underline{n} \cdot \underline{\nabla}\varphi^{(x/m)} = \begin{pmatrix} \underline{\underline{n}} \cdot \underline{\nabla}\varphi_1^{x/m}\\ \underline{\underline{n}} \cdot \underline{\nabla}\varphi_2^{x/m} \end{pmatrix}$$
$$\underline{\underline{n}} = [n_x, n_y]^T, \ \underline{\nabla} = \text{diag}\left([\underline{\underline{\nabla}}, \underline{\underline{\nabla}}]\right), \ \underline{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)^T$$

For a source distribution  $S(\mathbf{r}, \Omega)$ , with  $\Omega$  denoting the direction vector, r the spatial coordinate and  $\underline{n}$  the unit normal, the source vector at excitation is evaluated as

$$C^{Sx} = \begin{pmatrix} \int_{\Omega \cdot n < 0} S(\mathbf{r}, \Omega) 2 |\Omega \cdot \underline{n}| d\Omega \\ \int_{\Omega \cdot n < 0} S(\mathbf{r}, \Omega) (5 |\Omega \cdot \underline{n}|^3 - 3 |\Omega \cdot \underline{n}|) d\Omega \end{pmatrix}$$

The measurement at the  $j^{th}$  detector is defined as the exiting partial current on the boundary, and is expressed in terms of the composite moments as

$$j^{+}(r_{j}) = C^{J}\phi - C^{\nabla J}(n \cdot \nabla \phi) = \tilde{C}^{J}(\mathbf{r}_{j})\phi|_{\mathbf{r}_{j}}$$
(3)

with  $\tilde{C}^{J} = C^{J} + C^{\nabla J} (C^{\nabla b})^{-1} C^{b}$ . Here  $j^{+} = (j^{+x}, j^{+m})^{T}$ ,  $C^{J} =$ diag ([ $C^{Jx}$ ,  $C^{Jm}$ ]) and  $C^{\nabla J} =$  diag ([ $C^{\nabla Jx}, C^{\nabla Jm}$ ]) with  $C^{Jx/m} = \left( (\frac{1}{4} + J_{0}) \quad [(-\frac{2}{3})(\frac{1}{4} + J_{0}) + \frac{1}{3}(\frac{5}{16} + J_{2})] \right) \delta(\mathbf{r} - \mathbf{r}_{\mathbf{j}}),$  $C^{\nabla Jx/m} = \left( \frac{(0.5+J_{1})}{3\mu_{a1}^{x/m}} \quad \frac{J_{3}}{7\mu_{a3}^{x/m}} \right) \delta(\mathbf{r} - \mathbf{r}_{\mathbf{j}})$  and the terms { $J_{i}$ } are

evaluated as in [21]. In our notation the explicit dependence of the various quantities involved on the spatial coordinate 'r' is suppressed for brevity.

# 3. EVALUATION OF THE SECOND ORDER ADJOINT SENSITIVITY

In the present work we use the method described by Cacuci [19] to derive the SO adjoint sensitivities. We begin by briefly summarising the derivation of the analytical FO sensitivity from our earlier work [16]. The SO sensitivity is defined as a perturbation

of the FO adjoint sensitivity. Adjoint variables  $\Theta$  and  $\Lambda$  are then appropriately defined to obtain an analytical expression for the SO adjoint sensitivity.

#### A. The first order adjoint sensitivity

Let  $p_i$  denote an optical property such as  $(\mu_{ai}^x, \mu_{af}^x, \mu_s^x, g, ...)$ . From 3, we see that the exiting partial current  $j^+$  is a function of  $(p_i, \phi)$ . Therefore, the first variation of the measurement when the optical parameter  $p_i$  is perturbed is given by

$$\delta j^{+} = \frac{\partial \left( \tilde{C}^{J} \phi \right)}{\partial p_{i}} \delta p_{i} + \frac{\partial \left( \tilde{C}^{J} \phi \right)}{\partial \phi} \delta \phi = \tilde{C}^{J} \delta \phi$$
(4)

where  $\delta \phi \equiv \delta \phi [p_i, \delta p_i]$ . Here, we neglect the variation of the parameter value at the detector location.

The relationship between the variation in parameter value  $\delta p_i$  and the variation in fluence  $\delta \phi$  is obtained from 1 and 2 as the first order perturbation equation [16],

$$-\nabla \cdot C^{\nabla} \nabla \delta \phi + C \delta \phi = \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_i} \delta p_i \phi - \frac{\partial C}{\partial p_i} \delta p_i \phi$$
(5)

and the perturbed boundary condition,

$$C^{\nabla b}\left(n\cdot\nabla\delta\phi\right) + C^{b}\delta\phi = -\frac{\partial C^{\nabla b}}{\partial p_{i}}\delta p_{i}(n\cdot\nabla\phi)$$
(6)

We now define an adjoint  $\Psi \equiv \begin{pmatrix} \Psi^{xx} & \Psi^{xm} \\ \Psi^{mx} & \Psi^{mm} \end{pmatrix}$  with  $\Psi^{\alpha\gamma} =$ 

 $\begin{pmatrix} \Psi_1^{\alpha\gamma}, \\ \Psi_2^{\alpha\gamma} \end{pmatrix}$ ,  $(\alpha, \gamma = x/m)$ , through the adjoint equation,

$$-\nabla \cdot (C^{\nabla})^T \nabla \Psi + C^T \Psi = C_{int}^R$$
<sup>(7)</sup>

with adjoint boundary conditions

$$(C^{\nabla})^T (n \cdot \nabla \Psi) + (C^{ba})^T \Psi = C^R_{bnd}$$
(8)

where  $C^{ba} = \text{diag}\left(\left[C^{bax}, C^{bam}\right]\right)$  and  $C^{bax/m} = C^{\nabla x/m}(C^{\nabla bx/m})^{-1}C^{bx/m}$  and  $\Psi_{1,2}^{\alpha\gamma}$  are the composite moments of the adjoint field analogous to  $\phi_{1,2}^{x,m}$ . As pointed by Fedele et al [15], since the radiance at emission does not affect that at excitation, the component  $\Psi^{mx} = 0$ .  $C_{int}^R$  and  $C_{bnd}^R$  are as yet undefined adjoint sources in the interior and on the boundary respectively.

We now operate on the perturbation equation 5 with  $\int_{V} \Psi^{T}(\cdot) dV$ ,

$$\int_{V} \Psi^{T} \left( -\nabla \cdot C^{\nabla} \nabla \delta \phi + C \delta \phi \right) dV$$
$$= \int_{V} \Psi^{T} \left( \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_{i}} \nabla \phi - \frac{\partial C}{\partial p_{i}} \phi \right) \delta p_{i} dV$$
(9)

Using integration by parts twice, we get

$$\int_{V} \left( -\nabla \cdot \left( C^{\nabla} \right)^{T} \nabla \Psi + C^{T} \Psi \right)^{T} \delta \phi dV$$
$$- \int_{\partial V} \left( \Psi^{T} C^{\nabla} (n \cdot \nabla \delta \phi) + \left( (C^{\nabla})^{T} (n \cdot \nabla \Psi) \right)^{T} \delta \phi \right) d\sigma$$
$$= \int_{V} \Psi^{T} \left( \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_{i}} \nabla \phi - \frac{\partial C}{\partial p_{i}} \phi \right) \delta p_{i} dV$$
(10)

We use the forward (eq. 2) and perturbed boundary conditions (eq. 6) to simplify the above expression to,

$$\int_{V} \left( -\nabla \cdot \left( C^{\nabla} \right)^{T} \nabla \Psi + C^{T} \Psi \right)^{T} \delta \phi dV$$
  
$$- \int_{\partial V} \left( (C^{\nabla})^{T} (n \cdot \nabla \Psi) + (C^{ba})^{T} \Psi \right)^{T} \delta \phi d\sigma$$
  
$$= \int_{V} \Psi^{T} \left( \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_{i}} \nabla \phi - \frac{\partial C}{\partial p_{i}} \phi \right) \delta p_{i} dV$$
  
$$+ \int_{\partial V} \Psi^{T} C_{i}^{\nabla bp} \phi \delta p_{i} d\sigma$$
(11)

with  $C_i^{\nabla bp} = \operatorname{diag}(\left[C_i^{\nabla bpx}, C_i^{\nabla bpm}\right])$  and  $C_i^{\nabla bpx/m} = C^{\nabla x/m} (C^{\nabla bx/m})^{-1} \frac{\partial C^{\nabla bx/m}}{\partial p_i} (C^{\nabla bx/m})^{-1} C^{bx/m}.$ 

In order to obtain the adjoint sensitivity w.r.t the exiting partial current, we set in 7 and 8

$$C_{int}^{R} = 0$$
 and  $C_{bnd}^{R} = \left(\tilde{C}^{J}(\mathbf{r}_{j})\right)^{T} \delta(\mathbf{r} - \mathbf{r}_{j})$  (12)

Using 7 and 8, with the adjoint sources as defined above in 12, we can now write an expression for FO partial sensitivity as,

$$\delta j_i^+ = \int_V \Psi^T \left( \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_i} \nabla \phi - \frac{\partial C}{\partial p_i} \phi \right) \delta p_i dV + \int_{\partial V} \Psi^T C_i^{\nabla b p} \phi \delta p_i d\sigma$$
(13)

$$\equiv J_i \delta p_i = \begin{pmatrix} J_i^x \\ J_i^m \end{pmatrix} \delta p_i$$
 (14)

where

$$J_{i}^{x/m} = \int_{V} (\Psi^{x/m})^{T} \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_{i}} \nabla \phi dV - \int_{V} (\Psi^{x/m})^{T} \frac{\partial C}{\partial p_{i}} \phi dV + \int_{\partial V} (\Psi^{x/m})^{T} C_{i}^{\nabla bp} \phi d\sigma$$
(15)

with  $\Psi^x = \begin{pmatrix} \Psi^{xx} \\ 0 \end{pmatrix}$  and  $\Psi^m = \begin{pmatrix} \Psi^{xm} \\ \Psi^{mm} \end{pmatrix}$ .

Since the adjoint  $\Psi$  depends on the nominal value  $p_i$  but not on the perturbation  $\delta p_i$ , the set of equations 7 and 8 need to be solved only once to evaluate all the partial sensitivities,  $J_i$ .

#### B. Second order adjoint sensitivity

We will now proceed to apply the method of adjoints to derive the  $(i, j)^{th}$  component,  $\frac{\partial j^+}{\partial p_i \partial p_j}$ , of the SO sensitivity. The analytical FO sensitivity  $J_i^{x/m}$  can be viewed as a functional of  $(\phi, p_i, \Psi^{x/m})$ . We can write the first variation of  $J_i^{x/m}$  w.r.t perturbation of the optical parameter  $p_j$  as

$$\delta J_i^{x/m} = \frac{\partial J_i^{x/m}}{\partial p_j} \delta p_j + \frac{\partial J_i^{x/m}}{\partial \phi^x} \delta \phi^x + \frac{\partial J_i^{x/m}}{\partial \phi^m} \delta \phi^m + \left(\delta \Psi^{x/m}\right)^T \frac{\partial J_i^{x/m}}{\partial (\Psi^{x/m})^T}$$
(16)

$$\equiv H_{ij}^{x/m} \delta p_j \tag{17}$$

where  $\delta \phi \equiv \delta \phi \left[ p_j, \delta p_j \right]$ ,  $\delta \Psi \equiv \delta \Psi \left[ p_j, \delta p_j \right]$ . The first term in the summation can be directly obtained by taking the derivatives of the relevant coefficient matrices in eq. 15, as

$$\frac{\partial J_i^{x/m}}{\partial p_j} = \int_V \left( \Psi^{x/m} \right)^T \nabla \cdot \frac{\partial^2 C^{\nabla}}{\partial p_j \partial p_i} \nabla \phi dV + \int_{\partial V} \left( \Psi^{x/m} \right)^T \frac{\partial C_i^{\nabla bp}}{\partial p_j} \phi d\sigma$$
(18)

We discuss the evaluation of the partial variations w.r.t  $\phi^{x/m}$  and  $\Psi^{x/m}$  below.

#### B.1. Variation with respect to $\phi$

We define adjoint  $\Theta = [\Theta^x, \Theta^m]$ , with  $\Theta^x \equiv \begin{pmatrix} \Theta^{xx} \\ 0 \end{pmatrix}$ ,  $\Theta^m \equiv$ 

$$\begin{pmatrix} \Theta^{xm} \\ \Theta^{mm} \end{pmatrix}$$
, and  $\Theta^{\alpha\gamma} = \begin{pmatrix} \Theta_1^{\alpha\gamma} \\ \Theta_2^{\alpha\gamma} \end{pmatrix}$ ,  $(\alpha, \gamma = x/m)$ , through the adjoint equations

$$-\nabla \cdot (C^{\nabla})^T \nabla \Theta_i + C^T \Theta_i = S_{\theta}$$
(19)

with adjoint boundary conditions

$$(C^{\nabla})^T (n \cdot \nabla \Theta_i) + (C^{ba})^T \Theta_i = 0$$
(20)

We set the source  $S_{\theta} = diag\left(\left[\frac{\partial (J_i^x)^T}{\partial \phi}, \frac{\partial (J_i^m)^T}{\partial \phi}\right]\right) \delta(\mathbf{r} - \mathbf{r_u}).$ This enables us to pick out the variation of  $J_i^{x/m}$  w.r.t  $\phi$  at a point  $\mathbf{r}_u$  in the domain.

To obtain the source strengths  $\frac{\partial J_i^{x/m}}{\partial \phi}$ , we use integration by parts on the first term in 15, and rearrange the terms to obtain  $J_i^{x/m}$  in an alternate form as,

$$J_{i}^{x/m} = \int_{V} \left( \nabla \cdot \left( \frac{\partial C^{\nabla}}{\partial p_{i}} \right)^{T} \nabla \Psi^{x/m} - \frac{\partial C^{T}}{\partial p_{i}} \Psi^{x/m} \right)^{T} \phi dV$$
$$- \int_{\partial V} \left( \left( \frac{\partial C^{\nabla}}{\partial p_{i}} \right)^{T} (n \cdot \nabla \Psi^{x/m}) \right)^{T} \phi d\sigma$$
$$+ \int_{\partial V} \left( \left( C_{i}^{\nabla bpx} - \frac{\partial C^{\nabla}}{\partial p_{i}} (C^{\nabla b})^{-1} C^{b} \right)^{T} \Psi^{x/m} \right)^{T} \phi d\sigma$$
(21)

We can now write,

$$\frac{\partial \left(J_{i}^{x/m}\right)^{T}}{\partial \phi} = \int_{V} \nabla \cdot \left(\frac{\partial C^{\nabla}}{\partial p_{i}}\right)^{T} \nabla \Psi^{x/m} dV$$
$$- \int_{V} \frac{\partial C^{T}}{\partial p_{i}} \Psi^{x/m} dV - \int_{\partial V} \left(\frac{\partial C^{\nabla}}{\partial p_{i}}\right)^{T} (n \cdot \nabla \Psi^{x/m}) d\sigma$$
$$+ \int_{\partial V} \left(C_{i}^{\nabla bpx} - \frac{\partial C^{\nabla}}{\partial p_{i}} (C^{\nabla b})^{-1} C^{b}\right)^{T} \Psi^{x/m} d\sigma$$
(22)

We now operate on the perturbation eq. 5 corresponding to parameter  $p_i$ , with  $\int_V (\Theta_i^x)^T (\cdot) dV$ ,

$$\int_{V} \left( \Theta_{i}^{x} \right)^{T} \left( -\nabla \cdot C^{\nabla} \nabla \delta \phi + C \delta \phi \right) dV$$
$$= \int_{V} \left( \Theta_{i}^{x} \right)^{T} \left( \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_{j}} \phi - \frac{\partial C}{\partial p_{j}} \phi \right) \delta p_{j} dV$$
(23)

As with the evaluation of  $\delta j_i^+$ , we work through this equation using integration by parts and use the definition of the adjoint

5

(eq. 19 and eq. 20) along with the perturbed boundary condition (eq. 6), to obtain

$$\left(\frac{\partial J_i^x}{\partial \phi^x} \delta \phi^x\right)_j = \int_V (\Theta_i^x)^T \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_j} \delta p_j \nabla \phi dV - \int_V (\Theta_i^x)^T \frac{\partial C}{\partial p_j} \delta p_j \phi dV + \int_{\partial V} (\Theta_i^x)^T C_j^{\nabla b p} \phi d\sigma$$
(24)

Note:  $\frac{\partial J_i^x}{\partial \phi^m} = 0$ 

Similarly by operating on 5 with  $\int_{V} (\Theta_{i}^{m})^{T} (\cdot) dV$  and simplifying we can write

$$\begin{pmatrix} \frac{\partial J_i^m}{\partial \phi^x} \delta \phi^x + \frac{\partial J_i^m}{\partial \phi^m} \delta \phi^m \end{pmatrix}_j = \int_{\partial V} (\Theta_i^m)^T C_j^{\nabla bp} \phi \delta p_j d\sigma$$
$$+ \int_V (\Theta_i^m)^T \left( \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_j} \nabla \phi - \frac{\partial C}{\partial p_j} \phi \right) \delta p_j dV$$
(25)

#### B.2. Variation with respect to $\Psi$

As with the forward problem, we can map the change in parameter value,  $\delta p_j$ , to the change in the adjoint field,  $\delta \Psi^{x/m}$ , through the perturbation of the adjoint system of equations 7, 8

$$-\nabla (C^{\nabla})^{T} \cdot \nabla \delta \Psi^{x/m} + C^{T} \delta \Psi^{x/m}$$
$$= \nabla \cdot \left(\frac{\partial C^{\nabla}}{\partial p_{j}}\right)^{T} \delta p_{j} \nabla \Psi^{x/m} - \frac{\partial C^{T}}{\partial p_{j}} \delta p_{j} \Psi^{x/m}$$
(26)

and the corresponding perturbed boundary conditions

$$(C^{\nabla})^{T}n \cdot \nabla \delta \Psi^{x/m} + (C^{ba})^{T} \delta \Psi^{x/m}$$
  
=  $-\left(\frac{\partial C^{\nabla}}{\partial p_{j}}\right)^{T} \delta p_{j}(n \cdot \nabla \Psi^{x/m}) - \left(\frac{\partial C^{ba}}{\partial p_{j}}\right)^{T} \delta p_{j} \Psi^{x/m}$  (27)

We now define another set of adjoints  $\Lambda = \begin{pmatrix} \Lambda^x \\ \Lambda^m \end{pmatrix}$ , satisfying

$$-\nabla \cdot C^{\nabla} \nabla \Lambda_i + C \Lambda_i = S_{\Lambda}$$
<sup>(28)</sup>

with boundary conditions

$$C^{\nabla b} \left( n \cdot \nabla \Lambda_i \right) + C^b \Lambda_i = 0 \tag{29}$$

Analagous to the adjoint source  $S_{\theta}$ , we define the adjoint source  $S_{\Lambda} = diag\left(\left[\frac{\partial J_i^x}{\partial \Psi^x}, \frac{\partial J_i^m}{\partial \Psi^m}\right]\right) \delta(\mathbf{r} - \mathbf{r_u})$ , with the source strength

$$\frac{\partial J_i^{x/m}}{\partial (\Psi^{x/m})^T} = \int_V \left( \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_j} \nabla \phi - \frac{\partial C}{\partial p_j} \phi \right) dV + \int_{\partial V} C_j^{\nabla bp} \phi d\sigma$$
(30)

We note that, operating on the adjoint perturbation equation 26 with  $\int_V (\cdot)^T (\Lambda_i) dV$  instead of  $\int_V (\Lambda_i (\cdot)^T) dV$ , as in the previous cases, allows us to directly obtain the partial sensitivity in the form of  $\int_V \Psi^T (\cdots) \Lambda_i dV$  as opposed to  $\int_V \Lambda_i (\cdots) \Psi^T dV$ . We thus obtain,

$$\int_{V} \left( -\nabla \cdot \left( C^{\nabla} \right)^{T} \nabla \delta \Psi^{x/m} + C^{T} \delta \Psi^{x/m} \right)^{T} \Lambda_{i} dV$$
$$= \int_{V} \left( \nabla \cdot \left( \frac{\partial C^{\nabla}}{\partial p_{j}} \right)^{T} \nabla \Psi^{x/m} - \frac{\partial C^{T}}{\partial p_{j}} \Psi^{x/m} \right)^{T} \delta p_{j} \Lambda_{i} dV \qquad (31)$$

As in the previous cases, this can be simplified by using integration by parts and the perturbed boundary condition (eq. 27) along with the definition of the adjoints in eq. 28 and eq. 29 to obtain

$$\left(\delta \Psi^{x/m}\right)^T \frac{\partial J_i^x}{\partial (\Psi^{x/m})^T} = \int_V \left(\Psi^{x/m}\right)^T \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_j} \delta p_j \Lambda_i dV - \int_V \left(\Psi^{x/m}\right)^T \frac{\partial C}{\partial p_j} \delta p_j \Lambda_i dV - \int_{\partial V} \left(\Psi^{x/m}\right)^T \left(\frac{\partial C^{\nabla}}{\partial p_j} (C^{\nabla})^{-1} C^{ba} + \frac{\partial C^{ba}}{\partial p_j}\right) \delta p_j \Lambda_i d\sigma$$
(32)

#### B.3. An expression for the second order adjoint sensitivity

Using the partial variations derived in B.1 and B.2 along with the direct term evaluated in eq. 18, we can now write the (i, j)<sup>th</sup> component of the SO sensitivity as

$$\begin{split} H_{ij}^{x/m} &= \int_{V} \left( \Psi^{x/m} \right)^{T} \left( \nabla \cdot \frac{\partial^{2} C^{\nabla}}{\partial p_{i} \partial p_{j}} \nabla \phi \right) dV \\ &+ \int_{\partial V} \left( \Psi^{x/m} \right)^{T} \frac{\partial C_{i}^{\nabla bp}}{\partial p_{j}} \phi d\sigma \\ &+ \int_{V} \left( \Theta_{i}^{x/m} \right)^{T} \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_{j}} \nabla \phi dV \\ &- \int_{V} \left( \Theta_{i}^{x/m} \right)^{T} \frac{\partial C}{\partial p_{j}} \phi dV + \int_{\partial V} \left( \Theta_{i}^{x/m} \right)^{T} C_{j}^{\nabla bp} \phi d\sigma \\ &+ \int_{V} \left( \Psi^{x/m} \right)^{T} \left( \nabla \cdot \frac{\partial C^{\nabla}}{\partial p_{j}} \nabla \Lambda_{i} - \frac{\partial C}{\partial p_{j}} \Lambda_{i} \right) dV \\ &- \int_{\partial V} \left( \Psi^{x/m} \right)^{T} \left( \frac{\partial C^{\nabla x}}{\partial p_{j}} (C^{\nabla})^{-1} C^{ba} + \frac{\partial C^{ba}}{\partial p_{j}} \right) \Lambda_{i} d\sigma \end{split}$$
(33)

#### 4. FINITE ELEMENT FORMULATION

The domain of interest V, with boundary  $\partial V$ , is discretised into a structured grid with  $N_n$  nodes and  $N_e$  triangular elements with the local basis function  $N^T \equiv [N_1, N_2, N_3]$  where  $N_i$  denote the standard 2D linear shape functions. The Galerkin finite element method (GFEM), which employs the trial function  $N^T$ , is then used to solve the forward and adjoint problems. The forward and adjoint composite quantities are approximated as

$$u \approx \tilde{u} = \underbrace{\begin{pmatrix} N^T & 0\\ 0 & N^T \end{pmatrix}}_{\mathbf{N}^T} \begin{pmatrix} [u_1]\\ [u_2] \end{pmatrix}$$
(34)

where  $u = \phi^{x/m}$ ,  $\Psi^{x/m}$ ,  $\Theta^{x/m}$ ,  $\Lambda^{x/m}$  and  $[\cdot]$  denote nodal quantities. The optical parameters, and hence the coefficient matrices, are considered piece-wise constant over the elements.

The weighted residual form of the system of equations 1 and 2 can be written in block format as [16],

$$\underbrace{\begin{pmatrix} [\mathcal{A}_{x}] & 0\\ -[\mathcal{A}_{\beta}] & [\mathcal{A}_{m}] \end{pmatrix}}_{[\mathcal{A}]} \underbrace{\begin{pmatrix} [\phi^{x}]\\ [\phi^{m}] \end{pmatrix}}_{[\phi]} = \underbrace{\begin{pmatrix} [S_{F}^{x}]\\ 0 \end{pmatrix}}_{S}$$
(35)

where

$$\mathcal{A}_{x/m} = \int_{V} \left( \left[ \nabla \mathbf{N}^{T} \right] C^{\nabla x/m} \left[ \nabla \mathbf{N} \right] + \left[ \mathbf{N}^{T} \right] C^{x/m} \left[ \mathbf{N} \right] \right) dV$$
$$- \int_{\partial V} \left[ \mathbf{N} \right]^{T} C^{\nabla x/m} \left( C^{\nabla bx/m} \right)^{-1} C^{bx/m} \left[ \mathbf{N} \right]^{T} d\sigma$$
(36)

$$\mathcal{A}_{\beta} = \int_{V} \left[ \mathbf{N}^{T} \right] C^{\beta} \left[ \mathbf{N} \right] dV$$
(37)

$$S_F^x = Q \left[ \mathbf{N}(\mathbf{r}_s) \right]^T C^{\nabla x} \left( C^{\nabla bx} \right)^{-1} C^{Sx}$$
(38)

The excitation source is modeled as a point source at  $\mathbf{r}_s$  with strength *Q*.

As with the forward problem, we can write the block matrix form for the weighted residual of the adjoint system of equations 7 and 8 as [16],

$$[\mathcal{A}]^T[\Psi] = [S_{\Psi}] \tag{39}$$

with  $[S_{\Psi}] = diag\left(\left[[S_{\Psi}^{x}], [S_{\Psi}^{m}]\right]\right)$  and

$$[S_{\Psi}^{x/m}] = \left[\mathbf{N}(\mathbf{r}_j)\right]^T \left(\tilde{C}^{Jx/m}\right)^T$$
(40)

The finite element form for the sensitivity  $J_i^{x/m}$  can then be written as [16]

$$\begin{pmatrix} \begin{bmatrix} J_i^x \\ \end{bmatrix} \\ \begin{bmatrix} J_i^m \end{bmatrix} \end{pmatrix} = \underbrace{\begin{pmatrix} \begin{bmatrix} \Psi^{xx} \\ 0 \end{bmatrix} \begin{bmatrix} \Psi^{xm} \end{bmatrix}}_{[\Psi]}^T \underbrace{\begin{pmatrix} \begin{bmatrix} \delta \mathcal{A}_x \end{bmatrix} & 0 \\ -\begin{bmatrix} \delta \mathcal{A}_\beta \end{bmatrix} \begin{bmatrix} \delta \mathcal{A}_m \end{bmatrix}}_{[\delta \mathcal{A}]} \underbrace{\begin{pmatrix} \begin{bmatrix} \phi^x \end{bmatrix}}_{[\phi^m]} \end{pmatrix}_{[\phi]}$$
(41)

where

$$\delta \mathcal{A}_{x/m} = -\int_{V} \left( \nabla [\mathbf{N}]^{T} \frac{\partial C^{\nabla x/m}}{\partial p_{i}} \nabla [\mathbf{N}] dV + [\mathbf{N}]^{T} \frac{\partial C^{x/m}}{\partial p_{i}} [\mathbf{N}] \right) dV + \int_{\partial V} [\mathbf{N}]^{T} \underbrace{\left( C_{i}^{\nabla bpx/m} - \frac{\partial C^{\nabla x/m}}{\partial p_{i}} \left( C^{\nabla bx/m} \right)^{-1} C^{bx/m} \right)}_{C^{bx/m}} [\mathbf{N}] d\sigma \quad (42)$$

$$\delta \mathcal{A}_{\beta} = \int_{V} [\mathbf{N}]^{T} \frac{\partial C^{\beta}}{\partial p_{i}} [\mathbf{N}] dV$$
(43)

We note that the adjoint operator defining  $\Theta$  (resp.  $\Lambda$ ) is the same as that defining  $\Psi$  (resp.  $\phi$ ). Thus, we can write

$$[\mathcal{A}]^T[\Theta_i] = [S_{\Theta}] \tag{44}$$

(45)

and

$$[\mathcal{A}][\Lambda_i] = [S_\Lambda]$$

where  $[\Theta] = [[\Theta^x], [\Theta^m]], [\Lambda] = [[\Lambda^x]^T, [\Lambda^m]^T]$  and  $[S_{\Theta}] = diag([[S_{\Theta}^x], [S_{\Theta}^m]])$  with

$$[S_{\Theta}^{x/m}] = [\delta \mathcal{A}]^T [\Psi^{x/m}]$$
(46)

and

$$[S_{\Lambda}] = [\delta \mathcal{A}] [\phi] \tag{47}$$

A detailed derivation of these adjoint sources is presented in the appendix.

For the  $SP_3$  approximation the  $u^{th}$  column of the sources  $[S_{\Theta}^{x/m}]$  and  $[S_{\Lambda}^{x/m}]$  is of length  $2N_n$ , with non-zero entries only at rows corresponding to the  $u^{th}$  element. Since  $[S_{\Theta}^{x/m}]$ ,  $[S_{\Lambda}^{x/m}]$  depend on  $\Psi$  and  $\phi$  respectively in addition to the perturbed quantities, to obtain the elemental SO sensitivity they are defined as 3D matrices of size  $2N_n \times N_e \times N_d$  and  $2N_n \times N_e \times N_s$  respectively. The adjoints  $[\Theta_i]$  and  $[\Lambda_i]$  need to be solved only once to compute SO sensitivity w.r.t  $p_i$ . The computational solution

of the system of equations 44 and 45 can be easily parallelised over the adjoint source vectors by simultaneously computing  $[\Theta_i]$  (resp.  $[\Lambda_i]$ ) either for all the detectors (resp. sources) or elements depending on the computational resources available. This allows for a faster implementation of the SO sensitivities as compared to that using a finite difference scheme over the FO adjoint Jacobians, since the total overhead associated with distributing the task over multiple processors is lower with the SO adjoint scheme.

Applying the GFEM to eq. 33, we can write,

$$[H_{ij}] = [\Psi]^T [\delta^2 \mathcal{A}] [\phi] + [\Theta]^T [\delta \mathcal{A}] [\phi] + [\Psi]^T [\delta \mathcal{A}^T] [\Lambda]$$
(48)

where

$$\delta^{2} \mathcal{A} = -\int_{V} \nabla [\mathbf{N}]^{T} \frac{\partial^{2} C^{\nabla}}{\partial p_{j} \partial p_{i}} \nabla [\mathbf{N}] dV$$
$$-\int_{\partial V} [\mathbf{N}]^{T} \left( \frac{\partial^{2} C^{\nabla}}{\partial p_{j} \partial p_{i}} \left( C^{\nabla b} \right)^{-1} C^{b} - \frac{\partial C_{i}^{\nabla bp}}{\partial p_{j}} \right) [\mathbf{N}] d\sigma \qquad (49)$$

and

$$\delta \mathcal{A}^{T} = -\int_{V} (\nabla[\mathbf{N}])^{T} \frac{\partial C^{\nabla}}{\partial p_{j}} \nabla[\mathbf{N}] dV - \int_{V} [\mathbf{N}]^{T} \frac{\partial C}{\partial p_{j}} [\mathbf{N}] dV - \int_{\partial V} [\mathbf{N}]^{T} \frac{\partial C^{ba}}{\partial p_{j}} [\mathbf{N}] d\sigma$$
(50)

#### Vectorized global FEM matrix assembly

In the *SP*<sub>3</sub> approximation, a generic global FEM matrix *G* (contributing to formation of  $A_{x/\beta/m}$ ) of size  $2N_n \times 2N_n$  can be decomposed into four global FEM sub-matrices  $G_{ij}$  of size  $N_n \times N_n$  (corresponding to the  $(i, j)^{th}$  component-block of the C-type coefficient matrices). Each of these sub-matrices can be assembled using the following pseudo-code:

$$L = \operatorname{repmat}(C^*(i, j, :), 3, 3) \cdot K;$$
(51)

$$G_{ij} = \text{sparse}(Iv, Jv, L(:), N_n, N_n);$$
(52)

Note: We use MATLAB's notation for matrix repetition ('repmat'), creation of sparse matrix ('sparse'), elementwise product ('.\*') and matrix slicing (':').

'*C*<sup>\*</sup>' denotes any of the  $2 \times 2 \times N_e$  C-type coefficient matrices and '*K*' denotes the  $3 \times 3 \times N_e$  2-D counterpart of the K-type kernel matrices defined in [15]. The index vectors  $I_v$ ,  $J_v$  are 2-D analogues of those described in [15]. For example, each global FEM sub-matrix corresponding to  $A_x$  can be generated using the pseudo-code in Code-box 1. The index vectors  $I_s$ ,  $J_s$  are 2-D analogues of those described in [15].

#### Vectorization of global FEM matrix-field products

In the course of evaluating the SO adjoint sensitivity, we come across products of analytical derivatives with field vectors  $'u' (= \phi, \Psi, \Lambda_i, \Theta_i)$  at several instances of the type

$$T = \left[\delta G\right]\left[u\right] \tag{53}$$

The matrix  $\delta G$  is built from terms of the type  $\frac{\partial C^*}{\partial p}$ , and evaluating the product '*T*' in a conventional loop-based implementation would require looping over all the elements (or all the nodes for a nodal basis implementation). This computation is sped up by a vectorized implementation. For example, we can compute the matrix-field product  $[\delta A_x][u]$  using the pseudo-code in Code-box 2. Here D = 3(resp.4) for 2-dimensional (resp. 3-dimensional) and  $u^e = [u_1^e, u_2^e]^T$  with  $u_i^e$ , a  $D \times N_e$  matrix with

#### Research Article

7



**Code-box 2.** Pseudo-code for vectorization of global FEM matrixfield products

for 
$$i = 1 : D$$
  
for  $j = 1 : D$   
 $dG_v(2i - 1 : 2i, 2j - 1 : 2j, :) = \operatorname{repmat}(K_{stiffness}(i, j, :), 2, 2) . * \frac{\partial C^{\nabla x}}{\partial p}$   
 $+ \operatorname{repmat}(K_{mass}(i, j, :), 2, 2) . * \frac{\partial C^x}{\partial p};$   
end  
 $T_v(2i - 1 : 2i, :) = \sum_{j=1}^{D} dG_v(2i - 1 : 2i, 2j - 1 : 2j, :) . * u^e(2j - 1 : 2j, :);$   
end  
for  $i = 1 : D - 1$   
for  $j = 1 : D - 1$   
 $dG_b(2i - 1 : 2i, 2j - 1 : 2j, :) = \operatorname{repmat}(K_{boundary}(i, j, :), 2, 2) . * \tilde{C}^{bx};$   
end

$$T_b(2i-1:2i,:) = \sum_{j=1}^{(D-1)} (dG_b(2i-1:2i,2j-1:2j,:).*u^e(2j-1:2j,:));$$
  
end  
$$T = T_v + T_b;$$

the  $j^{th}$  row containing nodal values of the  $u_i$  for the  $j^{th}$  element. This code is then looped over the number of sources / detectors depending on the field vector u.

#### Vectorization of sensitivity computations

From 41 and 48, we see that the expression for adjoint sensitivity consists of terms of the type  $d\mathcal{G} = [v]^T [\delta G][u] \equiv [v]^T [T]$ . The products of any field quantity v with the matrix '*T*' can be easily computed as

$$d\mathcal{G} = \sum_{i=1}^{\tilde{D}} v^e (2i-1:2i,:) \cdot * T(2i-1:2i,:)$$
(54)

where  $v^e$  is defined similar to  $u^e$ ,  $\tilde{D} = D, D - 1$ . While it is feasible to use this approach to evaluate the terms of the type  $[\Psi]^T[\delta G][\phi]$ , it cannot be used for terms involving  $[\Theta]$  and  $[\Lambda]$ since their large size makes prohibitive memory demands for such an implementation. Instead, vectorization of FEM matrix-field products as described earlier is used to compute the products of type  $T \equiv [\delta G][u]$ ,  $u = \phi$ ,  $\Psi$  and a conventional loop based implementation is used to compute the product  $[v]^T[T]$ with  $[v] = [\Theta]$ ,  $[\Lambda]$  as described in the pseudo-code in Code-box 3: where  $N_{e_i}$  is the vector of nodes corresponding to element  $e_i$ .

**Code-box 3.** Pseudo-code for vectorization of sensitivity computations

for  $e_1 = 1 : N_e$ for  $e_2 = 1 : N_e$   $d\mathcal{G}(e_1, e_2, :) = v_1(N_{e_2}, e_1, :) . * T(1 : 2 : 2D, e_1, :)$   $+ v_2(N_{e_2}, e_1, :) . * T(2 : 2 : 2D, e_1, :);$ end end

This code is then looped over the sources (resp. detectors) for  $v = \Lambda(\text{resp. }\Theta)$ .

We summarize the numerical evaluation of the SO sensitivities (48) as:

Evaluation of SO adjoint sensitivity

- Step 1: Assemble A<sub>x/m/β</sub> as in pseudo-code in Codebox 1.
- Step 2: Solve for  $\phi$ ,  $\Psi$  using eqs. 35 and 39 resp.
- Step 3: Evaluate  $S_{\Theta}^{x/m}$ ,  $S_{\Lambda}^{x}$  as in eqs. 46 and 47 resp. using pseudo-code in Code-box 2.
- Step 4: For each detector (resp. source) evaluate  $\Theta_i$  (resp.  $\Lambda_i$ ) using eq 44 (resp. 45)
- Step 5: For each source-detector pair,
  - Step 5a: Use vectorization of sensitivity computations as in 54 to evaluate  $d\mathcal{G}_1 = [\Psi]^T [\delta^2 \mathcal{A}][\phi]$
  - Step 5b: Compute  $d\mathcal{G}_2 = [\Theta]^T [\delta \mathcal{A}][\phi]$  and  $d\mathcal{G}_3 = [\Psi]^T [\delta \mathcal{A}^T][\Lambda]$  using pseudo-code in Code-box 3.
  - Step 5c: Put together  $[H_{ij}] = \text{diag}[d\mathcal{G}_1] + d\mathcal{G}_2 + d\mathcal{G}_3$

# 5. NUMERICAL STUDIES

#### A. Phantom description

We present numerical investigations on a square domain of size  $2 \times 2$  cm with 10 detectors located along each edge at a spacing of 0.2 cm. Measurements are taken on all sides for each source of strength 1mW modulated at 100MHz located at the centre of each edge. Each dataset consists of 160 complex measurements. We scale the data using a logarithmic transformation, which separates the amplitude and the phase component to give 320 real measurements. For the reconstruction studies, a structured mesh with spacing h = 0.05 cm (3200 elements) is used for the reconstruction, while the simulated data is generated on a finer mesh with spacing h = 0.025 cm (12800 elements). The simulations are performed on MATLAB with a system powered by a 6-core Intel Xeon(R) processor with 64 GB RAM. The optical properties of the medium and the fluorophore for the high scattering ( $\mu_a \ll \mu'_s$ ) and low scattering ( $\mu_a \approx \mu'_s$ ) phantoms

considered are listed in Table 1. In each case, the phantom consists of two circular inhomogeneities embedded in the medium, of radius 0.2cm each, centered at (-0.5, 0) cm and (0.5, 0) cm with different values of  $\mu_{af}^x$ . We present numerical results for ideal ( i.e. no background fluorescence) and differential uptake of fluorophore wherein the background  $\mu_{af}^x$  is taken to be (1/50) times the value of the inhomogeneity with higher  $\mu_{af}^x$ .

## B. Validation of the second order sensitvity

To demonstrate the accuracy and efficacy of the current scheme, we present comparisons for the SO sensitivity,  $\frac{\partial^2 j^+}{\partial (\mu_{a_f}^x)^2}$  evaluated using the adjoint scheme with that evaluated using the finite difference scheme (with perturbation  $\delta \mu_{a_f}^x$ ) over the adjoint Jacobians,

$$H_{FD} = \frac{\left(\frac{\partial j^+}{\partial \mu_{af}^x}(\mu_{af}^x + \delta \mu_{af}^x) - \frac{\partial j^+}{\partial \mu_{af}^x}(\mu_{af}^x)\right)}{\delta \mu_{af}^x}$$
(55)

For the validation study, the optical parameters are taken as that of Phantom 1 described in Table 1. Two circular inhomogeneities of radius 0.2 cm are embedded in the medium at (0.5,0) and (-0.5,0) with  $\mu_{af}^x = 0.006 \text{ cm}^{-1}$ . We use a single excitation source located at (-1,0) modulated at 100 MHz and consider 10 detectors on the same edge as the source. In figure 1, we plot the real (left) and imaginary (right) components of the SO sensitivity using both the schemes for variation in parameter values at the locations p1 = (0,0), p2 = (-0.5,0.1), p3 = (0.5,0), p4 = (1,0.6). The points p2, p3 lie within the circular inclusions. The plot pipj corresponds to  $\frac{\partial^2 j^+}{\partial \mu_{af}^*(pi) \partial \mu_{af}^*(pi)}$ . The evaluated SO sensitivities



**Fig. 1.** The real(left) and imaginary (right) component of the SO sensitivity for adjoint based ('A') and finite difference 'F' schemes.

with both the schemes are in close agreement with a mean difference of the  $\mathcal{O}(10^{-15})$  between both evaluations. For a single source-detector pair, the adjoint scheme takes 160 seconds to compute the  $N_e \times N_e$  SO sensitivity matrix while the finite difference scheme takes 2000 seconds. This substantial speedup in the computation time makes it feasible to use the SO sensitivities in the solution of tomographic inverse problems. A similar comparison has been presented for the elastic scattering problem for the evaluation of the SO sensitivities  $\frac{\partial^2 j^+}{\partial (\mu_a)^2}$  in our earlier work [25].

#### C. Reconstruction studies

We consider the non-linear inverse problem, to reconstruct the fluorophore absoprtion coefficient  $\mu_{af}^x$  in the medium. All other optical properties as well as the background fluorescence are assumed known. Given measurements  $j_{meas}^{+x/m}$ , we define the cost function as

$$\zeta(p) = \frac{1}{2} ||\mathbf{R}(p)||^2 = \frac{1}{2} ||[R_1, R_2, \cdots R_{N_s}]^T||^2$$
(56)

where the residual vector for the *i*<sup>th</sup> source is  $R_i = [r_{i,1}, r_{i,2}, \cdots r_{i,N_d}]^T$ , with

$$r_{i,j} \equiv r_{i,j}(p, \mathbf{r}_j) = j_{meas}^{+x/m}(\mathbf{r}_j) - j^{+x/m}(p, \mathbf{r}_j)$$
(57)

We solve this problem using the second degree scheme of Hettlich and Rundell [9], implemented within the framework of the regularising Levenberg-Marquardt method [16, 23, 24].

The update vector  $s^k$ , at the  $k^{th}$  iterate, is evaluated using a predictor-corrector approach wherein, first a predictor  $\tilde{s}^k$  is generated by solving the set of equations

$$\left( (J^k)^T J^k + \lambda^k L^T L \right) \tilde{s}^k = - (J^k)^T R^k$$
(58)

where  $J^k$ ,  $R^k$  denote the  $N_e \times M$  Jacobian matrix and  $M \times 1$  residual vector,  $N_e$  being the total number of elements and M the total number of measurements.  $\lambda^k$  is the Levenberg-Marquardt parameter. The matrix 'L' is the graph Laplacian of the spatial discretisation scheme as in [2].

We then define a matrix  $T^k \equiv \left[T_1^k, T_2^k...T_M^k\right]^T$ , where the  $T_i^k$ s are evaluated as

$$T_i^k = J^{k,i} + (\tilde{s}^k)^T H^{k,i}$$
(59)

 $J^{k,i}$ ,  $H^{k,i}$  denote the respective components of the Jacobian and the Hessian corresponding to the  $i^{th}$  measurement. The corrected update vector  $s^k$  is then computed by solving

$$\left( (T^k)^T T^k + \lambda^k L^T L \right) s^k = -(T^k)^T R^k$$
(60)

Though it is possible to use different values of  $\lambda^k$  in the predictor (eq. 58) and corrector (eq. 60), we choose to use the same  $\lambda^k$  in our implementation. For the test-cases considered in our work, we froze the SO sensitivity, *H*, at the first iterate as also done in earlier works [8, 9].

The implementation of the FO regularising Levenberg-Marquardt scheme involves only the predictor step given by eq. 58 i.e. the update vector  $s_{FO}^k = \tilde{s_k}$ .

For the phantoms considered, reasonable reconstructions were obtained till SNR's of about 25dB, and we present results for noiseless data, and noisy data with SNR 25dB. We threshold the reconstructed values at  $0.2 \max \mu_{af}^{x,rec}$ .

To qualify the reconstructions, we use two metrics (i) the correlation coefficient  $\rho_c$  and (ii) the deviation factor defined as [4]

$$\rho_{c} = \frac{\sum_{i=1}^{N_{e}} (\mu_{af,i}^{x,rec} - \bar{\mu}_{af}^{x,rec})(\mu_{af,i}^{x,act} - \bar{\mu}_{af}^{x,act})}{(N_{e} - 1)\Delta\mu_{af}^{x,rec}\Delta\mu_{af}^{x,act}}$$
(61)

$$\rho_d = \frac{\sqrt{(1/N_e) \sum_{i=1}^{N_e} (\mu_{af,i}^{x,rec} - \mu_{af,i}^{x,act})^2}}{\Delta \mu_{af}^{x,act}}$$
(62)

Here  $\bar{\mu}_{af}^{x,rec}$ ,  $\bar{\mu}_{af}^{x,act}$  are the mean values and  $\Delta \mu_{af}^{x,rec}$ ,  $\Delta \mu_{af}^{x,act}$  are the standard deviations of the reconstructed and original spatial

**Table 1.** Description of the optical properties of the medium and fluorophore for both phantoms. The subscripts i/f denote quantities related to the background/fluorophore and the superscripts x/m indicate quantities at excitation/emission wavelength respectively.

|           | $\mu_{ai}^{x}$ | $\mu_{af}^{x}$ | $\mu^m_{ai}$        | $\mu^m_{af}$         | $\mu_s^{x}$ | $\mu_s^m$       | $\eta_f$ | $\tau_f(ns)$ |
|-----------|----------------|----------------|---------------------|----------------------|-------------|-----------------|----------|--------------|
|           |                |                |                     |                      |             |                 |          |              |
| Phantom 1 | 0.031          | 0.006          | $0.7987 \mu_{ai}^x$ | $0.0846\mu_{af}^{x}$ | 54.75       | $0.732 \mu_s^x$ | 0.016    | 0.56         |
| Phantom 2 | 0.45           | 0.087          | 0.45                | $0.043\mu_{af}^{x}$  | 10.0        | 10.0            | 0.019    | 0.4          |

**Table 2.** Error measures for the First order (FO) and Second order (SO) reconstruction schemes. For respective sets *A*, *S*, the letters *I*, *D* distinguish ideal and differential uptake cases with noiseless (N0) and noisy data with SNR 25dB(N1).

|         | ρ <sub>c</sub>  |      | $\rho_d$ |      | $\mu_{af}^{x,rec}$        |               |  |  |  |
|---------|---|------|----------|------|---------------------------|---------------|--|--|--|
|         | FO  | SO   | FO       | SO   | FO                        | SO            |  |  |  |
|         |   |      |          |      |                           |               |  |  |  |
| Dataset | High scattering setting, $\mu_{af}^{x,act} = (0.004, 0.006)cm^{-1}$ |      |          |      |                           |               |  |  |  |
|         |   |      |          |      |                           |               |  |  |  |
| SI-N0   | 0.81  | 0.85 | 0.59     | 0.51 | (.0045,.0056)             | (.0056,.0065) |  |  |  |
| SI-N1   | 0.81  | 0.84 | 0.59     | 0.54 | (.0042,.0054) (.0048,.006 |               |  |  |  |
|         | $\mu_{af}^{x,act} = (0.18, 0.3)cm^{-1}$                             |      |          |      |                           |               |  |  |  |
|         |   |      |          |      |                           |               |  |  |  |
| SD-N0   | 0.78  | 0.90 | 0.67     | 0.49 | (.0899,.1524)             | (.1364,.2259) |  |  |  |
| SD-N1   | 0.67  | 0.86 | 0.79     | 0.55 | (.0638,.1132)             | (.1260,.2266) |  |  |  |
|         |   |      |          |      |                           |               |  |  |  |
| Dataset | Low scattering setting $\mu_{af}^{x,act} = (0.058, 0.087)cm^{-1}$   |      |          |      |                           |               |  |  |  |
|         |   |      |          |      |                           |               |  |  |  |
| AI-N0   | 0.81  | 0.90 | 0.62     | 0.49 | (.0333,.0507)             | (.0426,.0624) |  |  |  |
| AI-N1   | 0.76  | 0.83 | 0.67     | 0.60 | (.0306,.0489)             | (.0353,.0531) |  |  |  |
|         | $\mu_{af}^{x,act} = (0.0522, 0.087) cm^{-1}$                        |      |          |      |                           |               |  |  |  |
|         |   |      | 3        |      |                           |               |  |  |  |
| AD-N0   | 0.64  | 0.79 | 0.83     | 0.66 | (.0158,.0270)             | (.0271,.0449) |  |  |  |
| AD-N1   | 0.62  | 0.77 | 0.82     | 0.66 | (.0142,.0312) (.0226,.051 |               |  |  |  |

parameter distributions in the region of interest, which we define as a rectangle of  $2 \times 1$ cm centered at the origin. A higher correlation coefficient and a lower deviation factor indicate a good match between the actual and reconstructed parameter values.

#### High scattering phantoms

In these phantoms, we use a homogenous initialisation of  $0.001 \text{ cm}^{-1}$  for the ideal uptake case, and of  $0.006 \text{ cm}^{-1}$  (which is the value of  $\mu_{af}^x$  of the background) for the differential uptake case for both the schemes. The optical parameter maps after thresholding are plotted for the FO and SO scheme in the left and right columns of Figure 2 respectively. The cross sections along y = 0 are plotted in Figures 4 (a,b). As is evident from the error metrics and the figures 2 (a-d), both the schemes per-

form at par for the ideal uptake case. In the differential uptake setting, the SO scheme performs distinctly better than the FO scheme in terms of both, the reconstructed parameter value and localisation (refer Table 2 and figure 2 (e-h)). The difference is particularly stark with the noisy data set SD-N1.

#### Low scattering phantoms

For this case we use a homogeneous initialisation of  $\mu_{af}^{x,init} = 0.01 \text{ cm}^{-1}$  for the ideal uptake case, and  $\mu_{af}^{x,init} = 0.0017 \text{ cm}^{-1}$  (which is the  $\mu_{af}^{x}$  value of the background) for the differential uptake case. In Figure 3 the reconstructed parameter maps are plotted for the FO scheme (left column) and the SO scheme (right column). In the low scattering setting, the SO scheme works better than the FO scheme for both ideal as well as differential uptake cases. For the differential uptake case the SO scheme clearly distinguishes better between the two objects. The cross sections along y = 0 for this set are plotted in Figure 4 (c,d).

Our numerical studies suggest that the SO scheme has a wider applicability than the FO scheme. It performs better than the FO scheme for the more realistic setting of differential uptake of fluorophore as well as for low-scattering test cases in addition to demonstrating better noise tolerance. A closer examination of the residuals,  $R^k$ , also revealed that the drop in the first few iterates is substantially larger with the SO than the FO scheme. The evaluation of the SO sensitivity was frozen at the initial guess in our work; however, using the updated SO sensitivities beyond the first iterate did not substantially improve the reconstructions for this problem.

# 6. CONCLUSIONS

We have presented an adjoint based scheme for the evaluation of SO sensitivities with respect to exiting partial current measurements at excitation and emission wavelength under the  $SP_3$ approximation for the coupled equations modeling fluorescence radiation transfer. We also describe a vectorized implementation for the SO adjoint sensitivity computations. By appropriately redefining the coefficient matrices, this can be extended to arbitrary order N of the  $SP_N$  approximation. Validation of the derived SO sensitivities with respect to a finite difference scheme demonstrates the accuracy of the scheme as well as highlights the speed-up achieved using the same. This massive speed-up makes it feasible to explore exact SO derivative based reconstruction and analysis schemes for the inverse problem in fluorescence optical tomography.

We have demonstrated the use of the SO sensitivity, through a SO regularising Levenberg Marquardt scheme to reconstruct the fluorophore absorption coefficient in various types of phantoms and contrasted it with a FO scheme. The evaluation of the SO sensitivities is frozen at the initial guess enabling precomputation of the SO sensitivities to allow faster reconstructions. It is observed that the SO scheme has wider applicabil-



**Fig. 2.** Reconstructions in high scattering setting for 1. Ideal uptake (a,b) data-set SI-N0, (c,d) data-set SI-N1, 2. Differential uptake (e,f) data set SD-N0, (g,h) data-set SD-N1. Reconstructions obtained using the first order (FO) scheme, and second order (SO) scheme are placed in the left and right columns respectively. The dashed red circles indicate the actual inhomogeneities.



**Fig. 3.** Reconstructions in low scattering setting for 1. Ideal uptake (a,b) data-set AI-N0, (c,d) data-set AI-N1, 2. Differential uptake (e,f) data set AD-N0, (g,h) data-set AD-N1. The reconstructions on the left are obtained using the first order (FO) scheme, while those on the right are obtained using the second order (SO) scheme. The dashed red circles indicate the actual inhomogeneity.



**Fig. 4.** Cross sectional plots along y = 0 for 1. High scattering setting (a,b) and 2. Low scattering setting (c,d). The left column plots the cross-sections for the ideal uptake cases (\*I-N0/N1), while the right column plots the cross-sections for the differential uptake cases (\*D-N0/N1)

ity and consistently provides better estimates of the parameter value. It distinctly outperforms the FO scheme for noisy data sets. These observations are consistent with other works in the literature and underscore the importance of making use of the SO sensitivities.

The objective of this study has been primarily to motivate an exploration of the efficacy and usability of SO schemes in the inverse problem of FOT, which is promoted by the speedup in evaluation of SO sensitivity obtained through the adjoint based scheme. While this scheme involves nearly as many number of computations, ( $O(N_p(N_d + N_s)))$ ), as the finite difference scheme over the adjoint FO sensitvities, as pointed in the text, it is more amenable to parallelisation owing to the structure of the sub-problems involved. The evaluated SO sensitivities also find potential use in numerical schemes for post-reconstruction analysis of the iterates.

The substantial speedup demonstrated however comes at the cost of memory, and the present implementation is memory intensive. The storage of the SO sensitivities also makes considerable demands on the memory. In the present work, we have not investigated optimising the performance of the scheme to make it memory efficient and work in this direction is ongoing.

Acknowledgement: The authors thank and acknowledge support of the Board of Research in Nuclear Sciences (BRNS), India in a part of this research.

# 7. APPENDIX

Applying the GFEM to 19, we can evaluate the components of the adjoint source  $S_{\Theta}^{x/m} = \frac{\partial (f_i^{x/m})^T}{\partial \phi} \delta(\mathbf{r} - \mathbf{r}_u)$  as

$$[S_{\Theta}^{x/m}]_{u} \equiv \int_{V} [\mathbf{N}]^{T} S_{\Theta}^{x/m} dV = [\mathbf{N}(\mathbf{r}_{u})]^{T} \frac{\partial \left(J_{i}^{x/m}\right)^{T}}{\partial \phi}$$
(63)

Using the definition of  $\frac{\partial (I_i^{x/m})^T}{\partial \phi}$  from eq. 21, we can write

$$\begin{split} [S_{\Theta}^{x/m}]_{u} &= \int_{V_{u}} [\mathbf{N}]^{T} \nabla \cdot \left(\frac{\partial C^{\nabla}}{\partial p_{i}}\right)^{T} \nabla [\mathbf{N}] [\Psi^{x/m}] dV \\ &- \int_{V_{u}} [\mathbf{N}]^{T} \frac{\partial C^{T}}{\partial p_{i}} [\mathbf{N}] [\Psi^{x/m}] dV \\ &- \int_{\partial V_{u} \cap \partial V} [\mathbf{N}]^{T} \left(\frac{\partial C^{\nabla}}{\partial p_{i}}\right)^{T} [n \cdot \nabla \Psi^{x/m}] d\sigma \\ &+ \int_{\partial V_{u} \cap \partial V} [\mathbf{N}]^{T} \left(C_{i}^{\nabla bp} - \frac{\partial C^{\nabla}}{\partial p_{i}} (C^{\nabla b})^{-1} C^{b}\right)^{T} [\mathbf{N}] [\Psi^{x/m}] d\sigma \end{split}$$
(64)

Using integration by parts, this simplifies to,

Similarly, applying GFEM to 28, we can evaluate  $S_{\Lambda}^{x/m}\delta(\mathbf{r} - \mathbf{r}_u)$  as

$$[S_{\Lambda}^{x/m}]_{u} \equiv \int_{V} [\mathbf{N}]^{T} S_{\Lambda}^{x/m} dV = [\mathbf{N}(\mathbf{r}_{u})]^{T} \frac{\partial J_{i}^{x/m}}{\partial (\Psi^{x/m})^{T}}$$
(67)

Using definition of  $\frac{\partial l_i^{x/m}}{\partial (\Psi^{x/m})^T}$  from 30, as with the expression above, we can derive,

$$\begin{split} [S_{\Lambda}^{x}]_{u} &= -\int_{V_{u}} [\nabla \mathbf{N}]^{T} \frac{\partial C^{\nabla x}}{\partial p_{i}} [\nabla \mathbf{N}] [\phi^{x}] dV \\ &- \int_{V_{u}} [\mathbf{N}]^{T} \frac{\partial C^{x}}{\partial p_{i}} [\mathbf{N}] [\phi^{x}] dV \\ &+ \int_{\partial V_{u} \cap \partial V} [\mathbf{N}]^{T} \left( C_{i}^{\nabla bpx} - \frac{\partial C^{\nabla x}}{\partial p_{i}} (C^{\nabla bx})^{-1} C^{bx} \right) [\mathbf{N}] [\phi^{x}]_{u} d\sigma \qquad (68) \\ &= [\delta \mathcal{A}_{x}]_{u} [\phi^{x}] \qquad (69) \end{split}$$

$$\begin{split} [S_{\Lambda}^{m}]_{u} &= -\int_{V_{u}} [\nabla \mathbf{N}]^{T} \frac{\partial C^{\nabla m}}{\partial p_{i}} [\nabla \mathbf{N}] [\phi^{m}] dV \\ &- \int_{V_{u}} [\mathbf{N}]^{T} \frac{\partial C^{m}}{\partial p_{i}} [\mathbf{N}] [\phi^{m}] dV + \int_{V_{u}} [\mathbf{N}]^{T} \frac{\partial C^{\beta}}{\partial p_{i}} [\mathbf{N}] [\phi^{x}] dV \\ &+ \int_{\partial V_{u} \cap \partial V} [\mathbf{N}]^{T} C_{i}^{\nabla bpm} [\mathbf{N}] [\phi^{m}] d\sigma \\ &- \int_{\partial V_{u} \cap \partial V} [\mathbf{N}]^{T} \frac{\partial C^{\nabla m}}{\partial p_{i}} (C^{\nabla bm})^{-1} C^{bm} [\mathbf{N}] [\phi^{m}] d\sigma \end{split}$$
(70)

$$= \left[\delta \mathcal{A}_{m}\right]_{u} \left[\phi^{m}\right]_{u} - \left[\delta \mathcal{A}_{\beta}\right]_{u} \left[\phi^{x}\right]_{u}$$
(71)

where  $[\cdot]_u$  indicates the corresponding quantity relevant to the  $u^{th}$  element. We can thus write the sources  $[S_{\Theta}]$ ,  $[S_{\Lambda}]$  in block form as in eq.46 and eq. 47.

## REFERENCES

1. S. R. Arridge, "Optical tomography in medical imaging," Inverse problems **15**, R41 (1999).

- M. Schweiger, S. R. Arridge, and I. Nissilä, "Gauss-newton method for image reconstruction in diffuse optical tomography," Phys. medicine biology 50, 2365 (2005).
- 3. J. E. Dennis Jr and R. B. Schnabel, <u>Numerical methods for</u> <u>unconstrained optimization and nonlinear equations</u>, vol. 16 (Siam, 1996).
- A. D. Klose and A. H. Hielscher, "Quasi-newton methods in optical tomographic image reconstruction," Inverse problems 19, 387 (2003).
- M. K. Transtrum and J. P. Sethna, "Improvements to the levenbergmarquardt algorithm for nonlinear least-squares minimization," arXiv preprint arXiv:1201.5885 (2012).
- S. Wright and J. Nocedal, "Numerical optimization," Springer Sci. 35, 7 (1999).
- R. Roy and E. Sevick-Muraca, "Truncated newton's optimization scheme for absorption and fluorescence optical tomography: Part i theory and formulation," Opt. express 4, 353–371 (1999).
- B. Kanmani and R. Vasu, "Diffuse optical tomography through solving a system of quadratic equations: theory and simulations," Phys. Medicine & Biol. 51, 981 (2006).
- F. Hettlich and W. Rundell, "A second degree method for nonlinear inverse problems," SIAM J. on Numer. Analysis 37, 587–620 (1999).
- R. Meyer, "Theoretical and computational aspects of nonlinear regression," in Nonlinear programming, (Elsevier, 1970), pp. 465–486.
- R. Kress and K.-M. Lee, "A second degree newton method for an inverse obstacle scattering problem," J. Comput. Phys. 230, 7661– 7669 (2011).
- R. Pierri, G. Leone, and R. Persico, "Second-order iterative approach to inverse scattering: numerical results," JOSA A 17, 874–880 (2000).
- N. Gould, T. Rees, and J. Scott, "A higher order method for solving nonlinear least-squares problems," Tech. rep., STFC Rutherford Appleton Laboratory (2017).
- N. Naik, J. Eriksson, P. de Groen, and H. Sahli, "A nonlinear iterative reconstruction and analysis approach to shape-based approximate electromagnetic tomography," IEEE Transactions on Geosci. Remote. Sens. 46, 1558–1574 (2008).
- F. Fedele, J. P. Laible, and M. J. Eppstein, "Coupled complex adjoint sensitivities for frequency-domain fluorescence tomography: Theory and vectorized implementation," J. Comput. Phys. **187**, 597–619 (2003).
- N. Naik, N. Patil, Y. Yadav, J. Eriksson, and A. Pradhan, "Fully nonlinear sp<sub>3</sub> approximation based fluorescence optical tomography," IEEE transactions on medical imaging **36**, 2308–2318 (2017).
- E. Greenspan, D. Gilai, and E. Oblow, "Second-order generalized perturbation theory for source-driven systems," Nucl. Sci. Eng. 68, 1–9 (1978).
- L. Gilli, D. Lathouwers, J. Kloosterman, and T. van der Hagen, "Applying second-order adjoint perturbation theory to time-dependent problems," Annals Nucl. Energy 53, 9–18 (2013).
- D. G. Cacuci, "Second-order adjoint sensitivity analysis methodology (2nd-asam) for computing exactly and efficiently first- and second-order sensitivities in large-scale linear systems: 1. computational methodology," J. Comput. Phys. 284, 687 – 699 (2015).
- D. G. Cacuci, "Second-order adjoint sensitivity analysis methodology (2nd-asam) for computing exactly and efficiently first-and second-order sensitivities in large-scale linear systems: 2. illustrative application to a paradigm particle diffusion problem," J. Comput. Phys. 284, 700–717 (2015).
- A. D. Klose and E. W. Larsen, "Light transport in biological tissue based on the simplified spherical harmonics equations," J. Comput. Phys. 220, 441–470 (2006).
- Y. Lu, B. Zhu, C. Darne, I.-C. Tan, J. C. Rasmussen, and E. M. Sevick-Muraca, "Improvement of fluorescence-enhanced optical tomography with improved optical filtering and accurate model-based reconstruction algorithms," J. biomedical optics 16, 126002–1260024 (2011).
- A. Doicu, F. Schreier, and M. Hess, "Iterative regularization methods for atmospheric remote sensing," J. Quant. Spectrosc. Radiat. Transf. 83, 47–61 (2004).
- 24. M. Hanke, "A regularizing levenberg-marquardt scheme, with applica-

tions to inverse groundwater filtration problems," Inverse problems **13**, 79 (1997).

 N. Patil and N. Naik, "Adjoint based hessian evaluation for sp n modeled optical tomography," in <u>Biomedical Imaging and Sensing Conference</u>, , vol. 10711 (International Society for Optics and Photonics, 2018), p. 107111M.