UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

L1-norm Based Nonlinear Reconstruction Improves Quantitative Accuracy of Spectral Diffuse Optical Tomography

Lu, Wenqi; Lighter, Daniel; Styles, Iain B.

DOI:
10.1364/BOE.9.001423

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Lu, W, Lighter, D & Styles, IB 2018, 'L1-norm Based Nonlinear Reconstruction Improves Quantitative Accuracy of Spectral Diffuse Optical Tomography', *Biomedical Optics Express*, vol. 9, no. 4, pp. 1423-1444. https://doi.org/10.1364/BOE.9.001423

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 23. Apr. 2024

Research Article | Vol. 9, No. 4 | 1 Apr 2018 | BIOMEDICAL OPTICS EXPRESS 1423

Biomedical Optics EXPRESS

L_1-norm based nonlinear reconstruction improves quantitative accuracy of spectral diffuse optical tomography

Wenqi Lu,[1] Daniel Lighter,[2] and Iain B. Styles[1,*]

[1] *School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*
[2] *Physical Sciences for Health Centre for Doctoral Training, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*
[*] *I.B.Styles@bham.ac.uk*

Abstract: Spectrally constrained diffuse optical tomography (SCDOT) is known to improve reconstruction in diffuse optical imaging; constraining the reconstruction by coupling the optical properties across multiple wavelengths suppresses artefacts in the resulting reconstructed images. In other work, L_1-norm regularization has been shown to improve certain types of image reconstruction problems as its sparsity-promoting properties render it robust against noise and enable the preservation of edges in images, but because the L_1-norm is non-differentiable, it is not always simple to implement. In this work, we show how to incorporate L_1 regularization into SCDOT. Three popular algorithms for L_1 regularization are assessed for application in SCDOT: iteratively reweighted least square algorithm (IRLS), alternating directional method of multipliers (ADMM), and fast iterative shrinkage-thresholding algorithm (FISTA). We introduce an objective procedure for determining the regularization parameter in these algorithms and compare their performance in simulated experiments, and in real data acquired from a tissue phantom. Our results show that L_1 regularization consistently outperforms Tikhonov regularization in this application, particularly in the presence of noise.

Published by The Optical Society under the terms of the Creative Commons Attribution 4.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

OCIS codes: (000.1430) Biology and medicine; (000.4430) Numerical approximation and analysis; (100.3190) Inverse problems; (110.6955) Tomographic imaging.

References and links

1. A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," Physics in Medicine and Biology **50**(4), R1–R43 (2005).
2. S. Srinivasan, B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, J. J. Gibson, T. D. Tosteson, S. P. Poplack, and K. D. Paulsen, "Interpreting hemoglobin and water concentration, oxygen saturation, and scattering measured in vivo by near-infrared breast tomography," Proceedings of the National Academy of Sciences **100**(21), 12349–12354 (2003).
3. H. Dehghani, B. W. Pogue, S. P. Poplack, and K. D. Paulsen, "Multiwavelength three-dimensional near-infrared tomography of the breast: initial simulation, phantom, and clinical results," Applied Optics **42**(1), 135–145 (2003).
4. D. A. Boas, K. Chen, D. Grebert, and M. A. Franceschini, "Improving the diffuse optical imaging spatial resolution of the cerebral hemodynamic response to brain activation in humans," Optics Letters **29**(13), 1506–1508 (2004).
5. B. W. Zeff, B. R. White, H. Dehghani, B. L. Schlaggar, and J. P. Culver, "Retinotopic mapping of adult human visual cortex with high-density diffuse optical tomography," Proceedings of the National Academy of Sciences **104**(29), 12169–12174 (2007).
6. A. Custo, D. A. Boas, D. Tsuzuki, I. Dan, R. Mesquita, B. Fischl, W. E. L. Grimson, and W. Wells, "Anatomical atlas-guided diffuse optical tomography of brain activation," Neuroimage **49**(1), 561–567 (2010).
7. H. Niu, S. Khadka, F. Tian, Z. Lin, C. Lu, C. Zhu, and H. Liu, "Resting-state functional connectivity assessed with two diffuse optical tomographic systems," J. Biomed. Opt. **16**(4), 046006 (2011).
8. A. T. Eggebrecht, S. L. Ferradal, A. Robichaux-Viehoever, M. S. Hassanpour, H. Dehghani, A. Z. Snyder, T. Hershey, and J. P. Culver, "Mapping distributed brain function and networks with diffuse optical tomography," Nature Photonics **8**, 448–454 (2014).
9. H. Dehghani, B. W. Pogue, S. Jiang, B. Brooksby, and K. D. Paulsen, "Three-dimensional optical tomography: resolution in small-object imaging," Appl. Opt. **42**(16), 3117–3128 (2003).

#313263 https://doi.org/10.1364/BOE.9.001423
Journal © 2018 Received 27 Nov 2017; revised 22 Dec 2017; accepted 22 Dec 2017; published 2 Mar 2018

Biomedical Optics EXPRESS

10. X. Heng, R. Springett, H. Dehghani, B. W. Pogue, K. D. Paulsen, and J. F. Dunn, "Magnetic-resonance-imaging–coupled broadband near-infrared tomography system for small animal brain studies," Appl. Opt. **44**(11), 2177–2188 (2005).
11. S. R. Arridge and J. C. Schotland, "Optical tomography: Forward and inverse problems," Inverse Problems **25**(12), 123010 (2009).
12. H. Dehghani, M. E. Eames, P. K. Yalavarthy, S. C. Davis, S. Srinivasan, C. M. Carpenter, B. W. Pogue, and K. D. Paulsen, "Near infrared optical tomography using nirfast: Algorithm for numerical model and image reconstruction," Communications in Numerical Methods in Engineering **25**(6), 711–732 (2009).
13. S. Srinivasan, B. W. Pogue, S. Jiang, H. Dehghani, and K. D. Paulsen, "Spectrally constrained chromophore and scattering near-infrared tomography provides quantitative and robust reconstruction," Applied Optics **44**(10), 1858–1869 (2005).
14. S. Srinivasan, B. W. Pogue, B. Brooksby, S. Jiang, H. Dehghani, C. Kogel, W. A. Wells, S. P. Poplack, and K. D. Paulsen, "Near-infrared characterization of breast tumors in vivo using spectrally-constrained reconstruction," Technology in Cancer Research and Treatment **4**(5), 513–526, 2005.
15. B. Brooksby, S. Srinivasan, S. Jiang, H. Dehghani, B. W. Pogue, K. D. Paulsen, J. Weaver, C. Kogel, and S. P. Poplack, "Spectral priors improve near-infrared diffuse tomography more than spatial priors," Opt. Lett. **30**(15), 1968–1970 (2005).
16. D. A. Boas, A. M. Dale, and M. A. Franceschini, "Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution, and accuracy," Neuroimage **23**, S275–S288 (2004).
17. D. A. Boas, D. H. Brooks, E. L. Miller, C. A. DiMarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang, "Imaging the body with diffuse optical tomography," IEEE Signal Processing Magazine **18**(6), 57–75 (2001).
18. A. H. Hielscher, A. D. Klose, and K. M. Hanson, "Gradient-based iterative image reconstruction scheme for time-resolved optical tomography," IEEE Transactions on Medical Imaging **18**(3), 262–271 (1999).
19. P. K. Yalavarthy, B. W. Pogue, H. Dehghani, C. M. Carpenter, S. Jiang, and K. D. Paulsen, "Structural information within regularization matrices improves near infrared diffuse optical tomography," Opt. Express **15**(13), 8043–8058 (2007).
20. R. P. K. Jagannath and P. K. Yalavarthy, "Nonquadratic penalization improves near-infrared diffuse optical tomography," J. Opt. Soc. Am. A **30**(8), 1516–1523 (2013).
21. J. C. Ye, S. Y. Lee, and Y. Bresler, "Exact reconstruction formula for diffuse optical tomography using simultaneous sparse representation," in 5th International Symposium on Biomedical Imaging: From Nano to Macro (IEEE, 2008), pp. 1621–1624.
22. H. R. A. Basevi, K. M. Tichauer, F. Leblond, H. Dehghani, J. A. Guggenheim, R. W. Holt, and I. B. Styles, "Compressive sensing based reconstruction in bioluminescence tomography improves image resolution and robustness to noise," Biomedical Optics Express **3**(9), 2131–2141 (2012).
23. J. Prakash, C. B. Shaw, R. Manjappa, R. Kanhirodan, and P. K. Yalavarthy, "Sparse recovery methods hold promise for diffuse optical tomographic image reconstruction," IEEE J. Sel. Top. Quantum Electron. **20**(2), 74–82 (2014).
24. V. C. Kavuri, Z. Lin, F. Tian, and H. Liu, "Sparsity enhanced spatial resolution and depth localization in diffuse optical tomography," Biomed. Opt. Express **3**(5), 943–957 (2012).
25. C. B. Shaw and P. K. Yalavarthy, "Effective contrast recovery in rapid dynamic near-infrared diffuse optical tomography using L_1-norm-based linear image reconstruction method," J. Biomed. Opt. **17**(8), 0860091 (2012).
26. J-C. Baritaux, K. Hassler, M. Bucher, S. Sanyal, and M. Unser, "Sparsity-driven reconstruction for FDOT with anatomical priors," IEEE Transactions on Medical Imaging **30**(5), 1143–1153 (2011).
27. D. L. Donoho, "For most large underdetermined systems of linear equations the minimal L_1-norm solution is also the sparsest solution," Communications on Pure and Applied Mathematics **59**(6), 797–829 (2006).
28. W. Lu, J. Duan, Z. Qiu, Z. Pan, R. W. Liu, and L. Bai, "Implementation of high-order variational models made easy for image processing," Mathematical Methods in the Applied Sciences **39**, 4208–4233 (2016).
29. S. Okawa, Y. Hoshi, and Y. Yamada, "Improvement of image quality of time-domain diffuse optical tomography with l_p sparsity regularization," Biomed. Opt. Express **2**(12), 3334–3348 (2011).
30. Q. Lyu, Z. Lin, Y. She, and C. Zhang, "A comparison of typical l_p minimization algorithms," Neurocomputing **119**, 413–424 (2013).
31. D. Kong and C. H. Q. Ding, "Non-convex feature learning via $L_{p,\infty}$ operator," In Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence (AAAI Press, 2014) pp. 1918–1924.
32. I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," Communications on Pure and Applied Mathematics **63**(1), 1–38 (2010).
33. M. V. Afonso, J. M. Bioucas-Dias and M. A. T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," IEEE Transactions on Image Processing **20**(3), 681–695 (2011).
34. M. V. Afonso, J. M. Bioucas-Dias and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," IEEE Transactions on Image Processing **19**(9), 2345–2356 (2010).
35. J. Duan, Z. Pan, X. Yin, W. Wei, and G. Wang, "Some fast projection methods based on Chan-Vese model for image segmentation," EURASIP Journal on Image and Video Processing **2014**(1), 1–16 (2014).
36. J. Duan, Z. Pan, B. Zhang, W. Liu, and X. Tai, "Fast algorithm for color texture image inpainting using the non-local CTV model," Journal of Global Optimization **62**(4), 853–876 (2015).

Biomedical Optics EXPRESS

37. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM Journal on Imaging Sciences **2**(1), 183–202 (2009).
38. T. Goldstein, C. Studer, and R. Baraniuk, "A field guide to forward-backward splitting with a FASTA implementation," https://arxiv.org/abs/1411.3406.
39. M. Schweiger, S. R. Arridge, M. Hiraoka, and D. T. Delpy, "The finite element method for the propagation of light in scattering media: boundary and source conditions," Medical Physics **22**(11), 1779–1792 (1995).
40. H. Dehghani, B. Brooksby, K. Vishwanath, B. W. Pogue, and K. D. Paulsen, "The effects of internal refractive index variation in near-infrared optical tomography: a finite element modelling approach," Physics in Medicine and Biology **48**(16), 2713 (2003).
41. M. Schweiger and S. Arridge, "The Toast++ software suite for forward and inverse modeling in optical tomography," J. Biomed. Opt. **19**(4), 040801 (2014).
42. S. R. Arridge and M. Schweiger, "Photon-measurement density functions. Part 2: Finite-element-method calculations," Appl. Opt. **34**(34), 8026–8037 (1995).
43. B. W. Pogue and M. S. Patterson, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," J. Biomed. Opt. **11**(4), 041102 (2006).
44. Y. Zhan, A. T. Eggebrecht, J. P. Culver, and H. Dehghani, "Singular value decomposition based regularization prior to spectral mixing improves crosstalk in dynamic imaging using spectral diffuse optical tomography," Biomed. Opt. Express **3**(9), 2036–2049 (2012).
45. V. A. Markel and J. C. Schotland, "Inverse problem in optical diffusion tomography. i. Fourier–Laplace inversion formulas," J. Opt. Soc. Am. A **18**(6), 1336–1347 (2001).
46. V. A. Markel, V. Mital, and J. C. Schotland, "Inverse problem in optical diffusion tomography. iii. inversion formulas and singular-value decomposition," J. Opt. Soc. Am. A **20**(5), 890–902 (2003).
47. X. Li, D. N. Pattanayak, T. Durduran, J. P. Culver, B. Chance, and A. G. Yodh, "Near-field diffraction tomography with diffuse photon density waves," Phys. Rev. E **61**(4), 4295 (2000).
48. C. B. Shaw and P. K. Yalavarthy, "Performance evaluation of typical approximation algorithms for nonconvex l_p-minimization in diffuse optical tomography," J. Opt. Soc. Am. A **31**(4), 852–862 (2014).
49. J. Duan, W. Lu, C. Tench, I. Gottlob, F. Proudlock, N. N. Samani and L. Bai, "Denoising optical coherence tomography using second order total generalized variation decomposition," Biomedical Signal Processing and Control **24**, 120–127 (2016).
50. J. Duan, Z. Qiu, W. Lu, G. Wang, Z. Pan and L. Bai, "An edge-weighted second order variational model for image decomposition," Digital Signal Processing **49**, 162–181 (2016).
51. B. Martinet, "Régularisation d'inéquations variationelles par approximations successives," RIRO **4**, 154–159 (1970).
52. A. Chambolle, R. A. De Vore, N-Y. Lee, and B. J. Lucier, "Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage," IEEE Transactions on Image Processing **7**(3), 319–335 (1998).
53. I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," Communications on Pure and Applied Mathematics **57**(11), 1413–1457 (2004).
54. M. A. T. Figueiredo and R. D. Nowak, "An Em algorithm for wavelet-based image restoration," IEEE Transactions on Image Processing **12**(8), 906–916 (2003).
55. E. T. Hale, W. Yin, and Y. Zhang, "A fixed-point continuation method for L_1-regularized minimization with applications to compressed sensing," CAAM TR07-07, Rice University (2007).
56. C. Vonesch and M. Unser, "A fast iterative thresholding algorithm for wavelet-regularized deconvolution," Proc. SPIE 6701, 1–5 (2007).
57. S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," IEEE Transactions on Signal Processing **57**(7), 2479–2493 (2009).
58. J. P. Culver, R. Choe, M. J. Holboke, L. Zubkov, T. Durduran, A. Slemp, V. Ntziachristos, B. Chance, and A. G. Yodh, "Three-dimensional diffuse optical tomography in the parallel plane transmission geometry: Evaluation of a hybrid frequency domain/continuous wave clinical system for breast imaging," Medical Physics **30**(2), 235–247 (2003).
59. J. Prakash and P. K. Yalavarthy, "A LSQR-type method provides a computationally efficient automated optimal choice of regularization parameter in diffuse optical tomography," Medical Physics **40**(3), 033101 (2013).
60. X. Wu, A. T. Eggebrecht, S. L. Ferradal, J. P. Culver, and H. Dehghani, "Quantitative evaluation of atlas-based high-density diffuse optical tomography for imaging of the human visual cortex," Biomedical Optics Express **5**(11), 3882–3900 (2014).
61. J. Ashburner and K. J. Friston, "Image segmentation," in *Human Brain Function* 2nd ed., R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, K.J. Friston, C.J. Price, S. Zeki, J. Ashburner, and W.D. Penny, eds. (Academic, 2003).
62. A. T. Eggebrecht, B. R. White, S. L. Ferradal, C. Chen, Y. Zhan, A. Z. Snyder, H. Dehghani, and J. P. Culver, "A quantitative spatial comparison of high-density diffuse optical tomography and fmri cortical mapping," Neuroimage **61**(4), 1120–1128 (2012).
63. H. Dehghani, B. R. White, B. W. Zeff, A. Tizzard, and J. P. Culver, "Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography," Appl. Opt. **48**(10), D137–D143 (2009).
64. H-Y. Wu, A. Filer, I. B. Styles, and H. Dehghani, "Development of a multi-wavelength diffuse optical tomography system for early diagnosis of rheumatoid arthritis: simulation, phantoms and healthy human studies," Biomedical

Biomedical Optics EXPRESS

Optics Express **7**(11), 4769–4786 (2016).
65. D. Lighter, A. Filer, and H. Dehghani, "Multispectral diffuse optical tomography of finger joints," Proc. SPIE **10412**, 104120N (2017).
66. J. A. Guggenheim, I. Bargigia, A. Farina, A. Pifferi, and H. Dehghani, "Time resolved diffuse optical spectroscopy with geometrically accurate models for bulk parameter recovery," Biomed. Opt. Express **7**(9), 3784–3794 (2016).

1. Introduction

Diffuse optical tomography (DOT) is a non-invasive, non-ionizing and low-cost imaging technology with applications in diagnosing breast cancer [1–3], analyzing brain function for functional neuroimaging [4–8], and imaging small animals for the study of disease detection, progression/regression and treatment [9, 10]. The imaging process typically involves the injection of near-infrared (NIR) light in the spectral range of 650−900nm into the imaging volume of interest (e.g. breast, head, mouse) through sources on the surface of the volume. The transmitted light is then measured at different locations using detectors that are also on the same volume surface. A number of measurements are acquired using different source-detector pairs, and the internal distribution of the tissue's optical properties is reconstructed using a transport-model-based image reconstruction algorithm [11].

When a single wavelength continuous-wave (CW) light source is used, only the amplitude of the fluence can be measured at the surface boundary. In this case, only tissue absorption at that wavelength can be estimated. However, since the quantities of interest in DOT experiments are typically chromophore concentrations (mainly oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb)) rather than the absorption itself, measurements have to be taken at multiple CW wavelengths in order to provide sufficient information to recover the distributions of these chromophores. There are two main approaches for reconstruction of chromophore images using multiple wavelength measurements: non-spectral methods and spectrally constrained methods. Non-spectral methods (traditional DOT) reconstruct the absorption coefficients at each wavelength independently and then calculate the chromophore concentrations using Beer's law [12]. Spectrally constrained DOT (SCDOT) directly reconstructs the chromophore distributions by using the known absorption spectra of the chromophores to constrain the reconstruction process. Compared with non-spectral methods, spectrally constrained reconstruction has been shown to be better at suppressing artefacts in the resulting reconstructed images and to reduce crosstalk between chromophores and scatter parameters in breast imaging [13–15]. It has also been shown that boundary measurements at two NIR wavelengths are sufficient to recover the concentrations of HbO_2 and Hb [16].

Reconstruction of images from DOT measurements is a difficult inverse problem. The limited availability of boundary measurements and the diffusive nature of light propagation in tissue [11, 17] make the problem non-linear and ill-posed, and iterative solutions with effective methods for regularization are necessary to obtain unique solutions. Many approaches have been used [18–20] and quadratic Tikhonov (L_2-norm) regularization is the most popular approach as the solutions to each iteration step can be computed analytically, simplifying the reconstruction process. This is known to suppress the high-frequency components of the reconstructed image (normally noise) leading to smooth reconstructions, but this has the drawback of being unable to preserve sharp features in the reconstructed images [19].

L_1-norm regularization has recently been adopted for single wavelength DOT image reconstruction. Features of interest in DOT, such as tumours in the breast or activations in the brain are typically spatially localized and in this case the vector corresponding to the difference in the optical properties relative to the background is sparse with only a few non-zero elements [21–24]. L_1-norm regularization is known to induce sparsity in the solution to inverse problems [24–26], and has been shown to give essentially the same sparsity as the true sparsity measure (L_0-norm) [27]. Compared to L_2 regularization, L_1 regularization is less sensitive to outliers, which correspond to sharp edges in image processing applications [28] and is thus

Biomedical Optics EXPRESS

able to preserve edge-like features. Both L_2 and L_1 regularization methods can be solved by convex optimization schemes where a unique solution can be guaranteed [24, 25]. The more general L_p-norm ($0 < p < 1$) regularization schemes have also been studied for DOT image reconstruction [23, 29], and are also known to introduce sparsity to the reconstructed image [30]; however, L_p regularization is known to be nonconvex meaning that local minima exist [31] and unique solutions cannot be guaranteed.

In traditional DOT, L_1 regularization can indeed be applied to each wavelength independently. However, there are no guarantees that the solutions will be consistent with Beer's law. It must be assumed that the regularizer will have the same sparsifying effect at all wavelengths. This may not necessarily be true, given that SNR, scattering etc are different. SCDOT can be used to constrain the solution space to those solutions that are physically plausible. Therefore, reconstruction with spatial and spectral regularization simultaneously applied will constrain the solution space much more reliably than their sequential application. To the best of our knowledge, L_1-norm has not yet been used in SCDOT image reconstruction. We introduce a novel algorithm, spectral-L_1, which combines the sparsity-preserving advantages of L_1 regularization with the spectral constraints imposed by coupling optical properties across multiple wavelengths, to solve the inverse problem for image reconstruction in SCDOT. The key advance is to adapt the DOT reconstruction process to incorporate efficient methods for solving each iterative step. These are necessary because the L_1-norm is non differentiable and the update terms in the reconstruction process cannot be computed analytically. We investigate three algorithms for solving the update term: iteratively reweighted least square algorithm (IRLS) [32], alternating directional method of multipliers (ADMM) [33–36] and fast iterative shrinkage-thresholding algorithm (FISTA) [37, 38]. All three methods have been widely used to obtain sparse solutions to linear systems. IRLS and ADMM are second-order algorithms that require explicit inversion of a large matrix; FISTA is a first-order algorithm that does not require explicit matrix inversion, but does require a gradient operator to be constructed.

We adapt the DOT reconstruction process to use these methods for the solution of the update terms. An automated method to automatically select the regularization parameters is developed which is based on the L-curve method but is modified for this use-case. Then we perform a systematic comparison of the different regularization methods (L_1 and L_2) and optimization algorithms (IRLS, ADMM and FISTA) on simulated data in two- and three-dimensions. The comparison evaluates the methods on the accuracy of image reconstruction; ability to preserve edges; robustness against noise; and computational efficiency. Comprehensive and robust qualitative and quantitative evaluations are performed to quantitatively compare the reconstruction results using average contrast (AC), Pearson correlation (PC) and peak signal-to-noise ratio (PSNR). To our knowledge, this is the first systematic study in the area of spectral DOT reconstruction to perform such a comprehensive evaluation. We then apply our methods to the reconstruction of functional activations in simulated human brain imaging data using realistic anatomical models and finally evaluate the proposed algorithms using experimentally acquired data, by imaging a tissue-mimicking, plastic phantom of known optical properties using a multispectral DOT system.

The paper is organized as follows: Section 2 introduces the theory for image reconstruction in SCDOT and proposes a new spectral-L_1 inverse model; Section 3 investigates the performance of the candidate three reconstruction algorithms for our proposed model; In section 4, a principled method for selecting the regularization parameter is described; Section 5 presents extensive comparative experiments in simulated models, and the results of experiments using tissue phantoms. In section 6, the conclusions that can be drawn from our results are discussed.

Biomedical Optics EXPRESS

2. Theory

Image reconstruction in SCDOT aims to find the tissue composition that best explains the boundary measurements. It typically requires the repeated evaluation of a forward model of light propagation in biological tissues as part of an inverse model that minimizes the difference between the measurements and the model's predicted measurements. In this section, the forward model for CW light propagation is introduced, followed by the spectrally constrained inverse model. The L_1 and L_2 regularization methods for the inverse problem are described at the end of the section.

2.1. The forward model

It is generally accepted that if the scattering coefficients dominate over absorption coefficients in tissues and the region of interest is far from the light sources, light propagation can be modelled by a diffusion equation (DE) [11]. The DE is able to generate isotropic fluence fields given a distribution of source fibres and the tissue optical properties. For a CW system, the DE can be written as

$$-\nabla \cdot \kappa(r,\lambda_i)\nabla\Phi(r,\lambda_i) + \mu_a(r,\lambda_i)\Phi(r,\lambda_i) = q_0(r,\lambda_i). \tag{1}$$

Here, $\mu_a(r,\lambda_i)$ is the absorption coefficient at position r for wavelength λ_i, $\kappa(r,\lambda_i) = 1/3[\mu_a(r,\lambda_i) + \mu_s'(r,\lambda_i)]$ in which $\mu_s'(r,\lambda_i)$ is the reduced scattering coefficient. $\Phi(r,\lambda_i)$ is the photon density at position r and wavelength λ_i, and the isotropic source term at wavelength λ_i is given by $q_0(r,\lambda_i)$. It should be noted that in CW imaging, the value of $\mu_s'(r,\lambda_i)$ is not updated by the reconstruction algorithm and is assumed to be a known constant. The air-tissue boundary is represented by an index-mismatched type III condition (Robin or mixed boundary condition) in which the fluence at the edge of the tissue exits and does not return [39, 40]. A finite element method (FEM) is used to numerically solve Eq. (1) on a discretized mesh, which has been implemented in several open-source software packages, notably TOAST++ [41] and NIRFAST [12]. In this work, the NIRFAST package is used for all computations.

In CW systems, the tissue absorption μ_a depends on the concentration of chromophores in the tissue. The relationship between the absorption coefficients at different wavelengths is therefore constrained by the intrinsic absorption properties of the chromophores via Beer's law. For a dual-wavelength imaging system, and for two chromophores, Beer's law is written in matrix-vector form as

$$\begin{pmatrix} \mu_{a,\lambda_1} \\ \mu_{a,\lambda_2} \end{pmatrix} = \begin{pmatrix} \varepsilon_{c_1,\lambda_1} & \varepsilon_{c_2,\lambda_1} \\ \varepsilon_{c_1,\lambda_2} & \varepsilon_{c_2,\lambda_2} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \tag{2}$$

where c_1 and c_2 are chromophore concentrations and λ_1 and λ_2 are two measurement wavelengths. In the remainder of the paper c_1 and c_2 correspond to oxyhemoglobin HbO_2 and deoxyhemoglobin Hb respectively, with $\lambda_1 = 750$nm and $\lambda_2 = 850$nm. $\varepsilon_{c_i,\lambda_i}$ $(i = 1, 2)$ are the extinction coefficients of the two chromophores at the corresponding wavelength λ_i. The values of $\varepsilon_{c_i,\lambda_i}$ have been documented by Zeff et al (2007) [5].

2.2. The inverse model for SCDOT image reconstruction

In SCDOT, chromophore concentrations c_1, and c_2 are directly estimated from the boundary measurements in preference to explicitly reconstructing optical properties at each wavelength. The following SCDOT inverse model allows direct estimation of chromophore parameters from two measurement wavelengths (i.e. 750 and 850nm) using some form of iterative procedure. Using a block notation, in which $(\vdots)$ represents the concatenation of two column vectors, we have:

$$c_1, c_2 = \underset{c_1,c_2}{\arg\min} \left\| \begin{pmatrix} \Phi^{M}_{\lambda_1} \\ \Phi^{M}_{\lambda_2} \end{pmatrix} - \begin{pmatrix} \Phi^{C(k)}_{\lambda_1} \\ \Phi^{C(k)}_{\lambda_2} \end{pmatrix} \right\|_2^2, \tag{3}$$

Biomedical Optics EXPRESS

where chromophores c_1 and c_2 are the model parameters to be recovered. $\Phi^{\mathrm{M}}_{\lambda_i}$ $(i = 1, 2)$ is the measured fluence at the tissue surface and $\Phi^{\mathrm{C}}_{\lambda_i}$ is the calculated data using the forward solver. The superscript k denotes the iteration number. Eq. (3) defines a non-linear least square problem which can be solved via the classical Gauss-Newton method in which the first order Taylor series is used to expand the forward solution $\Phi^{\mathrm{C}}_{\lambda_i}$ as

$$\begin{pmatrix} \Phi^{\mathrm{C}(k)}_{\lambda_1} \\ \Phi^{\mathrm{C}(k)}_{\lambda_2} \end{pmatrix} = \begin{pmatrix} \Phi^{\mathrm{C}(k-1)}_{\lambda_1} \\ \Phi^{\mathrm{C}(k-1)}_{\lambda_2} \end{pmatrix} + \mathrm{J}^{k-1} \begin{pmatrix} c_1^k - c_1^{k-1} \\ c_2^k - c_2^{k-1} \end{pmatrix}, \tag{4}$$

in which the spectral Jacobian J (also known as the sensitivity matrix) relates the changes in boundary data to changes in chromophore concentrations and can be constructed directly with the incorporation of spectral prior information using the adjoint method [42]. Note that when $k = 1$, an initial guess of the chromophore concentrations c_1^0 and c_2^0 is required which can be obtained by a data-calibration procedure explained elsewhere [43]. The spectral Jacobian J can be derived as [44]:

$$\mathrm{J} = \begin{pmatrix} \frac{\partial \Phi^{\mathrm{C}}_{\lambda_1}}{\partial c_1} & \frac{\partial \Phi^{\mathrm{C}}_{\lambda_1}}{\partial c_2} \\ \frac{\partial \Phi^{\mathrm{C}}_{\lambda_2}}{\partial c_1} & \frac{\partial \Phi^{\mathrm{C}}_{\lambda_2}}{\partial c_2} \end{pmatrix} = \begin{pmatrix} \frac{\partial \Phi^{\mathrm{C}}_{\lambda_1}}{\partial \mu_{a,\lambda_1}} \cdot \frac{\partial \mu_{a,\lambda_1}}{\partial c_1} & \frac{\partial \Phi^{\mathrm{C}}_{\lambda_1}}{\partial \mu_{a,\lambda_1}} \cdot \frac{\partial \mu_{a,\lambda_1}}{\partial c_2} \\ \frac{\partial \Phi^{\mathrm{C}}_{\lambda_2}}{\partial \mu_{a,\lambda_2}} \cdot \frac{\partial \mu_{a,\lambda_2}}{\partial c_1} & \frac{\partial \Phi^{\mathrm{C}}_{\lambda_2}}{\partial \mu_{a,\lambda_2}} \cdot \frac{\partial \mu_{a,\lambda_2}}{\partial c_2} \end{pmatrix} = \begin{pmatrix} \mathrm{J}_{\lambda_1} \cdot \varepsilon_{c_1,\lambda_1} & \mathrm{J}_{\lambda_1} \cdot \varepsilon_{c_2,\lambda_1} \\ \mathrm{J}_{\lambda_2} \cdot \varepsilon_{c_1,\lambda_2} & \mathrm{J}_{\lambda_2} \cdot \varepsilon_{c_2,\lambda_2} \end{pmatrix}, \tag{5}$$

where J_{λ_i} relates the changes in boundary data to changes in the absorption coefficient at wavelength λ_i. The size of J in this case is the number of wavelengths times the number of measurements per wavelength, by number of finite element nodes times number of chromophore parameters.

Substituting Eq. (4) into Eq. (3) leads to

$$\Delta c^k = \arg\min_{\Delta c} ||\Delta\Phi^{k-1} - \mathrm{J}^{k-1}\Delta c||_2^2, \tag{6}$$

where Δc^k is the change in the chromophore parameters at the k-th iteration and can be written as

$$\Delta c^k = \begin{pmatrix} \Delta c_1^k \\ \Delta c_2^k \end{pmatrix} = \begin{pmatrix} c_1^k - c_1^{k-1} \\ c_2^k - c_2^{k-1} \end{pmatrix}. \tag{7}$$

$\Delta\Phi$ in Eq. (6) is the data-model mismatch which is given by

$$\Delta\Phi^{k-1} = \begin{pmatrix} \Phi^{\mathrm{M}}_{\lambda_1} - \Phi^{\mathrm{C}(k-1)}_{\lambda_1} \\ \Phi^{\mathrm{M}}_{\lambda_2} - \Phi^{\mathrm{C}(k-1)}_{\lambda_2} \end{pmatrix}. \tag{8}$$

Minimizing Eq. (6) leads to the normal equations

$$\left(\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)}\right)\Delta c^k = \mathrm{J}^{(k-1)\mathrm{T}}\Delta\Phi^{k-1}. \tag{9}$$

which can be solved to find the update term Δc^k using the Gauss-Newton algorithm which is summarized in Algorithm 1.

It is however non-trivial to calculate the inverse of $\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)}$ in Eq. (9) (i.e. step 6 in Algorithm 1) because it is normally singular or close to singular. Furthermore, experimental noise in the measurements $\Phi^{\mathrm{M}}_{\lambda_i}$ tends to lead to reconstruction artefacts if this inversion is computed directly. Strategies that can be employed to invert such ill-posed matrices includes

Biomedical Optics EXPRESS

Algorithm 1: Gauss-Newton Algorithm for Minimizing Eq. (3).

INPUT: $\Phi^{\mathrm{M}}_{\lambda_i}$ $(i = 1, 2)$, $iter$, Tol
Initialize: c_1^0, c_2^0
  for $k = 1 : iter$
    1: Update μ_{a,λ_i} at each wavelength using Beer's law (Eq. (2))
    2: Update $\Phi^{\mathrm{C}(k-1)}_{\lambda_i}$ at each wavelength using the forward model (Eq. (1))
    3: Update $\Delta\Phi^{k-1}$ using Eq. (8)
    4: Stop if $k = iter$ or $||\Delta\Phi^{k-1} - \Delta\Phi^{k-2}||_1 \leqslant$ Tol, otherwise go to step 5
    5: Update J^{k-1} using Eq. (5)
    6: Update $\Delta c^k = (\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)})^{-1}\mathrm{J}^{(k-1)\mathrm{T}}\Delta\Phi^{k-1}$
    7: Update c_1^k and c_2^k using Eq. (7)
  end for
RETURN c_1^k and c_2^k

algebraic reconstruction technique (ART), truncated singular value decomposition (TSVD) or the simultaneous iterative reconstruction technique (SIRT) [1, 11, 45–47]. Regularization can also be employed to reduce model errors and artefacts caused by measurement noise. In the following section, we introduce an L_1-based regularization technique to solve this ill-posed inverse problem.

2.3. The proposed spectral-L_1 inverse model

To convert Eq. (6) into a more readily solvable problem, a Tikhonov (L_2) regularization term is usually introduced into the inverse problem:

$$\Delta c^k = \arg\min_{\Delta c} \left\{ ||\Delta\Phi^{k-1} - \mathrm{J}^{k-1}\Delta c||_2^2 + \lambda||\Delta c||_2^2 \right\}. \tag{10}$$

The regularization parameter λ determines the degree of regularization that will be imposed on the model. This can be solved analytically to give

$$\Delta c^k = (\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)} + \lambda\mathrm{I})^{-1}\mathrm{J}^{(k-1)\mathrm{T}}\Delta\Phi^{k-1}. \tag{11}$$

I is the identity matrix and its size is the same as that of $\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)}$. The introduction of $\lambda\mathrm{I}$ effectively reduces the condition number of the matrix, thus stabilizing the matrix inversion. This is widely known as the Levenberg-Marquardt algorithm and is in general more robust than the Gauss-Newton method. An analytical solution to this problem is possible because Eq. (10) is convex and quadratic which makes L_2-norm regularization an attractive choice for many inverse problems. However, in image reconstruction problems, it tends to over-smooth the result and sharp features such as object boundaries in the reconstructed images are often smeared. Moreover, L_2-norm discourages sparsity, and is not suitable for sparse image reconstruction. In SCDOT image recovery, the perturbation/change Δc^k is usually zero or close to zero when the region to be recovered is not in the vicinity of the region of interest. In this case, Δc^k is spatially sparse. Recently, L_1 regularization has been widely studied because of several useful properties: it is sparsity-promoting, convex, edge-preserving, and is more robust against noise. Therefore, we propose a new inverse model for SCDOT image recovery based on L_1 regularization that we refer to as spectral-L_1. This is formulated as

$$\Delta c^k = \arg\min_{\Delta c} \left\{ ||\Delta\Phi^{k-1} - \mathrm{J}^{k-1}\Delta c||_2^2 + \lambda||\Delta c||_1 \right\}. \tag{12}$$

Although L_1 regularization has many advantages over L_2 regularization, the L_1-norm is non-differentiable, which makes it difficult to solve Eq. (12). Three candidate algorithms for this task will be investigated in the next section.

Biomedical Optics EXPRESS

3. Candidate algorithms for solving the proposed spectral-L_1 method

We now consider three algorithms for the solution of Eq. (12): iteratively reweighted least squares (IRLS) [32]; alternating directional method of multipliers (ADMM) [33]; and fast iterative shrinkage-thresholding algorithm (FISTA) [37]. These algorithms will be incorporated into the image reconstruction process by substituting them into step 6 of Algorithm 1, which solves for the update term.

3.1. IRLS

Instead of solving the L_1-minimization problem directly, IRLS reformulates the problem as a sequence of weighted L_2 minimization problems. Specifically, by introducing a weight matrix W, the L_1 minimization can be converted into finding the optimal solution of the quadratic problem

$$\Delta c^k = \arg\min_{\Delta c} \left\{ ||\Delta\Phi^{k-1} - \mathrm{J}^{k-1}\Delta c||_2^2 + \lambda ||\mathrm{W}\Delta c||_2^2 \right\}. \tag{13}$$

W is a diagonal matrix with weights, w_s, along its diagonal that are given by

$$w_s = \begin{cases} |\Delta c_s^{i-1}|^{-0.5} & \text{if } |\Delta c_s^{i-1}| \geqslant \varepsilon \\ \varepsilon^{-1} & \text{if } |\Delta c_s^{i-1}| < \varepsilon \end{cases}. \tag{14}$$

The superscript i above denotes the i'th IRLS iteration (Algorithm 2), and it should be distinguished from the superscript k denoting the iterations of Algorithm 1. A small positive number $0 < \varepsilon \ll 1$ is used to avoid the possibility of division by zero. It has been suggested that ε should be a series of positive real numbers that decay to zero over iterations [48]. In practice, we have found that using a fixed value in the range $0.001 \leq \varepsilon \leq 0.01$ does not give significantly different results. Eq. (13) results in the normal equation

$$\left(\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)} + \lambda \mathrm{W}^{\mathrm{T}}\mathrm{W}\right)\Delta c = \mathrm{J}^{(k-1)\mathrm{T}}\Delta\Phi^{k-1}. \tag{15}$$

This is known as the weighted L_2-minimization scheme. We note that if the diagonal weights w_s are set to 1, the normal equation reduces to the conventional L_2-minimization scheme (Eq. (11)). The IRLS algorithm employing this method is summarized in Algorithm 2.

The calculation of the elements of W requires an initial guess for Δc for which we use the solution to the L_2-regularized problem, Eq. (11). One of the biggest advantages of IRLS is that Eq. (15) has an analytical solution which allows Eq. (13) to be solved exactly, making IRLS almost as easy to implement as the Levenberg-Marquadt scheme. In common with many sparsity-promoting optimization methods, the sparsity level in IRLS is controlled by the regularization parameter λ which must be chosen carefully for each specific problem.

Algorithm 2: Iteratively Reweighted Least Square Algorithm (IRLS)

INPUT: $\Delta\Phi^{k-1}$, J^{k-1}, λ, $iter$, Tol
Initialize: Set Δc^0 using Eq. (11)
  for $i = 1 : iter$
    1: Update W using Eq. (14)
    2: Update $\Delta c^i = (\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)} + \lambda \mathrm{W}^{\mathrm{T}}\mathrm{W})^{-1}\mathrm{J}^{(k-1)\mathrm{T}}\Delta\Phi^{k-1}$
    3: Stop if $i = iter$ or $||\Delta c^i - \Delta c^{i-1}||_1 \leqslant$ Tol, otherwise go to step 1
  end for
RETURN $\Delta c^k = \Delta c^i$

Biomedical Optics EXPRESS

3.2. ADMM

ADMM has been widely used to solve optimization problems in machine learning, signal processing, and standard image restoration and reconstruction. This method has become particularly important in the field of variational image processing, which frequently requires the minimization of non-differentiable objectives [33,34,49,50]. It has been shown to be able to solve constrained optimization problems effectively and efficiently. The basic idea is to decompose a complex optimization problem into several simpler subproblems, which usually have closed-form solutions [35]. Its simplicity, flexibility, and broad applicability have made it an important part of the modern optimization toolset. To apply ADMM to our spectral-L_1 problem, we first introduce an auxiliary splitting vector variable v, an augmented Lagrangian multiplier b, and a positive penalty parameter θ, reformulating Eq. (12) as the following unconstrained optimization problem

$$\Delta c, v, b = \underset{\Delta c, v, b}{\arg\min} \left\{ ||\Delta\Phi^{k-1} - \mathrm{J}^{k-1}\Delta c||_2^2 + \lambda ||v||_1 + \frac{\theta}{2}||v - \Delta c - b||_2^2 \right\}. \tag{16}$$

This multivariate optimization problem corresponds to a sub-minimization problem with respect to Δc, v and b, separately. When all the subproblems converge, the solution for Δc approximately represents that of Eq. (12). In order to find the minimizers for all of the subproblems, ADMM searches all the saddle points of Eq. (16) by first fixing the variables (v, b) and minimizing the subproblem with respect to Δc using the following normal equation

$$\left(\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)} + \theta\mathrm{I}\right)\Delta c = \mathrm{J}^{(k-1)\mathrm{T}}\Delta\Phi^{k-1} + \theta\left(v^{i-1} - b^{i-1}\right). \tag{17}$$

By inverting the matrix on the left-hand side of Eq. (17), a unique solution for Δc is found. We then fix variables Δc and b and set the first order derivative with respect to v to zero. This leads to

$$\lambda\frac{v}{|v|} + \theta\left(v - \Delta c^i - b^{i-1}\right) = 0, \tag{18}$$

which can be solved component-wise using an analytical shrinkage-thresholding method to give

$$v^i = \max\left(\left|\Delta c^i + b^{i-1}\right| - \frac{\lambda}{\theta}, 0\right) \circ sign\left(\Delta c^i + b^{i-1}\right), \tag{19}$$

where $\circ$ and $sign$ symbols denote component-wise multiplication and the signum function, respectively. The last step of ADMM is to update the augmented Lagrangian multiplier b, as $b^i = b^{i-1} + \Delta c^i - v^i$. The complete method is presented in Algorithm 3. The key advantage of ADMM is that Eqs. (17) and (18) have closed-form solutions. We note that the augmented Lagrangian multiplier means that different choices of the penalty parameter θ will provide similar results but with different rates of convergence. In all the experiments we have conducted, we used $\theta = 0.01$ to achieve fast convergence.

3.3. FISTA

FISTA is a very efficient optimization approach that uses the forward-backward splitting technique (FBS) [37,38]. It is an extension of the classical gradient descent method and therefore belongs to the class of first order methods that are a better choice for large-scale problems than second-order methods such as IRLS and ADMM because they do not require the explicit construction of very large matrices. Let us consider minimizing the L_1-regularized data fitting energy given by Eq. (12). We begin by analyzing the standard unregularized problem with $\lambda = 0$. Let $\mathrm{F}(\Delta c) = ||\Delta\Phi^{k-1} - \mathrm{J}^{k-1}\Delta c||_2^2$ and $\nabla\mathrm{F}(\Delta c) = \mathrm{J}^{(k-1)\mathrm{T}}(\mathrm{J}^{(k-1)}\Delta c - \Delta\Phi^{k-1})$ denote its gradient. We apply the gradient descent algorithm

$$\Delta c^i = \Delta c^{i-1} - t\nabla\mathrm{F}\left(\Delta c^{i-1}\right), \tag{20}$$

Biomedical Optics EXPRESS

Algorithm 3: Alternating Directional Method of Multipliers (ADMM)

INPUT: $\Delta\Phi^{k-1}$, J^{k-1}, λ, θ, $iter$, Tol
Initialize: $v^0 = 0$, $b^0 = 0$
for $i = 1 : iter$
1: Update $\Delta c^i = (\mathrm{J}^{(k-1)\mathrm{T}}\mathrm{J}^{(k-1)} + \theta\mathrm{I})^{-1}(\mathrm{J}^{(k-1)\mathrm{T}}\Delta\Phi^{k-1} + \theta(v^{i-1} - b^{i-1}))$
2: Update v^i using Eq. (19)
3: Update $b^i = b^{i-1} + \Delta c^i - v^i$
4: Stop if $i = iter$ or $||\Delta c^i - \Delta c^{i-1}||_1 \leqslant$ Tol, otherwise go to step 1
end for
RETURN $\Delta c^k = \Delta c^i$

where $t > 0$ is a suitable stepsize which controls how far the iteration moves along the gradient direction in the i'th iteration. The value of t is initialized by estimating the Lipschitz constant $\widetilde{L}$ of $\nabla\mathrm{F}$ as $\widetilde{L} = L(\nabla\mathrm{F})$ and then backtracking rules are adopted to guarantee that the objective has decreased sufficiently [38]. The gradient iteration given by Eq. (20) can be understood as a proximal regularization [51] of the linearized function $\mathrm{F}(\Delta c)$ at Δc^{i-1}, which corresponds to the following optimization problem:

$$\Delta c = \arg\min_{\Delta c}\left\{\mathrm{F}\left(\Delta c^{i-1}\right) + \nabla\mathrm{F}\left(\Delta c^{i-1}\right)\left(\Delta c - \Delta c^{i-1}\right) + \frac{1}{2t}||\Delta c - \Delta c^{i-1}||_2^2\right\}. \tag{21}$$

Analogously, adopting the same gradient descent idea to solve Eq. (12) with $\lambda \neq 0$ leads to the following minimization problem

$$\Delta c = \arg\min_{\Delta c}\left\{\frac{1}{2t}||\Delta c^{i-1} - t\nabla\mathrm{F}\left(\Delta c^{i-1}\right) - \Delta c||_2^2 + \lambda||\Delta c||_1\right\}. \tag{22}$$

Minimizing this results in a formulation similar to Eq. (18) and can be solved in the same way to give

$$\Delta c^i = \max\left(\left|\Delta c^{i-1} - t\nabla\mathrm{F}\left(\Delta c^{i-1}\right)\right| - t\lambda, 0\right) \circ sign\left(\Delta c^{i-1} - t\nabla\mathrm{F}\left(\Delta c^{i-1}\right)\right). \tag{23}$$

The minimizer of the Eq. (12) can then be found by iterating Δc^i in Eq. (23) to convergence. In isolation, Eq. (23) is known as the iterative shrinkage-thresholding algorithm (ISTA) [52–57], whose global convergence rate is $O(1/N)$, where N is the iteration counter. This is improved upon by using a Nesterov-type acceleration technique to obtain faster convergence. In the FISTA algorithm, the iterative shrinkage operator is not used on the value obtained from the previous iteration Δc^{i-1}, but rather on a combination of the values from the previous two iterations. Thus, in FISTA, Eq. (23) is replaced with

$$\Delta c^i = \max\left(\left|\Delta y^i - t\nabla\mathrm{F}\left(\Delta y^i\right)\right| - t\lambda, 0\right) \circ sign\left(\Delta y^i - t\nabla\mathrm{F}\left(\Delta y^i\right)\right), \tag{24}$$

where Δy^i comes from the prediction procedure given in step 4 of Algorithm 4. This step can help to push the solution to the current iteration further in the direction it moved during the previous iteration, which can significantly improve the computational efficiency. The complete FISTA method is presented in Algorithm 4, where steps 2 and 3 implement the acceleration strategy and can be viewed as an over-relaxation step that improves the global convergence rate from $O(1/N)$ to $O(1/N^2)$.

3.4. The spectral-L_1 algorithm

We have introduced three methods for solving the chromophore update terms in SCDOT with a sparsity enforcing constraint: IRLS, ADMM, and FISTA. These algorithms replace the single

Biomedical Optics EXPRESS

Algorithm 4: Fast Iterative Shrinkage-Thresholding Algorithm (FISTA)

INPUT: $\Delta\Phi^{k-1}$, J^{k-1}, λ, $iter$, Tol
Initialize: Set Δc^0 using Eq. (11), $\alpha^0 = 1$, $t < 1/\widetilde{L}$
  for $i = 1 : iter$
    1: Update Δc^i using Eq. (24)
    2: Stop if $i = iter$ or the relative residual $\leqslant$ Tol, otherwise go on to step 3
    3: Update $\alpha^i = (1 + \sqrt{1 + 4(\alpha^{i-1})^2})/2$
    4: Update $\Delta y^{i+1} = \Delta c^i + (\alpha^{i-1} - 1)/(\alpha^i)(\Delta c^i - \Delta c^{i-1})$
  end for
RETURN $\Delta c^k = \Delta c^i$

update term of the conventional reconstruction algorithm (step 6 of Algorithm 1). The flow-chart presented in Fig. 1 shows how these proposed methods are integrated into SCDOT reconstruction for CW imaging. Since the three proposed optimization schemes are themselves iterative, our method contains nested iterations. In IRLS and FISTA, an initial guess for Δc is required. We use the standard L_2-regularized solution (Eq. (11)). This is only required on the first iteration of the outer loop.

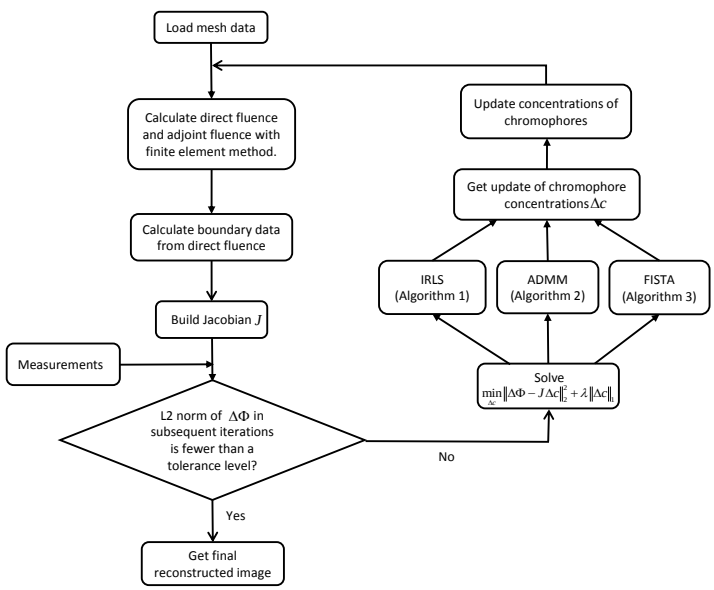


Fig. 1. Flow chart for SCDOT image reconstruction using the proposed spectral-L_1 model.

4. Parameter selection

The regularization parameter λ determines the trade-off between the goodness-of-fit of the model to the data, and the strict enforcement of the regularization criteria. An optimal value between the two quantities must therefore be found: if too much regularization is imposed on the model, then it will not fit the data properly; if the regularization parameter is too small, the fit will be good but the solution will be dominated by data errors and measurement noise (the overfitting regime). There are several methods to find an optimal compromise between these two considerations and the L-curve method is both simple and effective. By plotting the model-data mismatch $||\Delta\Phi - J\Delta c||_2^2$ against the model regularization $||\Delta c||_2^2$ or $||\Delta c||_1$ for a sequence of different λ,

Biomedical Optics EXPRESS

a curve which is typically L-shaped can be constructed. Figure 2 shows the L-curves obtained from each of the four candidate optimization schemes using the numerical experiments described by Zhan et al [44]. The optimal trade-off occurs at the "elbow" of the L-shaped curve and this can be located by determining the point of maximum curvature of the curve.

Since strong regularization can improve the conditioning of the linear system, we solve the formulations given by Eqs. (10) and (12) with a relatively large regularization parameter λ and then decrease it gradually by a fixed factor until the curvature of the L-curve starts to decrease. This corner point is considered to be at the optimum value of λ where both the model fit and the regularization function are simultaneously near to their minimum values. In principle, computing the L-curve requires the full image reconstruction process to be run multiple times which is computationally very expensive. We have found that it is sufficient to compute the L-curve for one iteration of the outer loop of Fig. 1, and then to use the resulting optimal value of λ for the remaining iterations. In addition, in order to avoid the special case where the L-curve does not allows an optimal value of λ to be found by purely numerical means [58, 59], we select a range around the parameter with the highest curvature value. We then adjust the values manually to get the final optimal parameter by visually inspecting the solutions and choosing the one that generates the sparsest solution with a well-defined compact localization. This approximate optimum is then used for subsequent iterations. We note that the choice of algorithm for L_1 regularized reconstruction significantly affects the shape of the L-curve and the optimal value of λ.

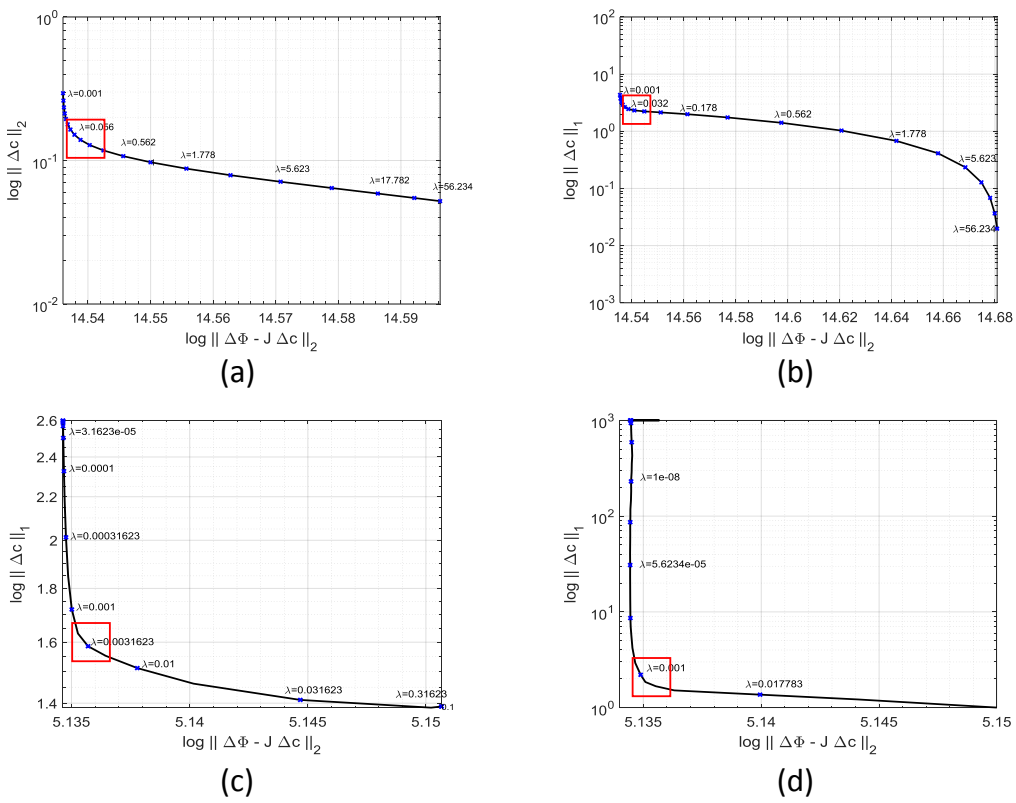


Fig. 2. L-curves (data fit against model regularization) derived from a synthetic example: a) Tikhonov regularization; b) L_1 regularization using the IRLS algorithm; c) L_1 regularization using the ADMM algorithm; d) L_1 regularization using the FISTA algorithm. The optimal regularization parameter is around the point of maximum curvature (within the red boxes).

Biomedical Optics EXPRESS

In addition to the regularization parameter λ that is common to all three L_1 algorithms, we have considered how to select the other parameters of each method to ensure that our comparison is fair and unbiased. IRLS has one parameter ε and we set this to $0.001 \leq \varepsilon \leq 0.01$ following the recommendations set out by Shaw and Yalavarthy [48]. ADMM has one parameter θ and the use of the augmented Lagrangian multiplier means that different choices of θ provide similar results but lead to different rates of convergence. In all the experiments, θ was set to 0.01 to achieve fast convergence. FISTA has two parameters t and α. t is initialized by estimating the Lipschitz constant and then backtracking rules are adopted to guarantee that the objective has decreased sufficiently [38]. t is therefore updated automatically. α is involved in an over-relaxation step (i.e. step 4 in Algorithm 3) and its update is also automatic (i.e. step 3 in Algorithm 3). The regularization parameter λ is therefore the only parameter that must be optimized for a specific problem.

5. Experiment setup

We have performed extensive experiments to evaluate the performance of different models and algorithms qualitatively and quantitatively. We first define three evaluation metrics to quantify the quality of the reconstructed images. We then describe simulated numerical experiments, and then real experiments performed on phantom samples. For experiments in which measurement noise was added, ten repeats were performed. In all cases, the forward model was implemented using the NIRFAST package [12] in Matlab R2013a (Mathworks, Natick, USA).

5.1. Quantitative evaluation metrics

Three quantitative evaluation metrics are considered: the average contrast (AC), Pearson correlation (PC) and peak signal-to-noise ratio (PSNR). Ideally, if the reconstructed image is exactly same as the ground truth image, AC is equal to 1. For PC and PSNR, the recovered image has higher quality if higher PC or PSNR values are obtained.

Average constrast (AC) is based on the mean value of the region of interest and is defined as:

$$\mathrm{AC} = \frac{\sum_{j=1}^{N} c_i^j / N}{\tilde{c}_i} \quad i = 1, 2 \tag{25}$$

where c_i^j denotes the recovered values of chromophore i on the finite element node j. N is the number of nodes in the activation region which is selected by thresholding the recovered changes based on 50% of the maximum recovered changes. $\tilde{c}_i$ are the ground truth values of the chromophores in the activation region.

The second evaluation metric PC is given by

$$\mathrm{PC} = \frac{\mathrm{COV}\left(c_i, \tilde{c}_i\right)}{\sigma\left(c_i\right)\sigma\left(\tilde{c}_i\right)} \quad i = 1, 2. \tag{26}$$

The numerator is the covariance (COV) of the recovered images with the ground truth and σ indicates standard deviation. The PC is thus a measure of the joint variability of the ground truth image with the reconstructed image.

Finally, PSNR evaluates the difference between the ground truth image and the recovered image. Larger PSNR values means less difference between the target and the recovered image. This measure is defined as follows

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}_{c_i}^2}{\mathrm{MSE}}\right) \quad i = 1, 2. \tag{27}$$

Here, MAX_{c_i} is the maximum pixel value of c_i and MSE is the mean squared error between the reconstructed and ground truth values.

$$\mathrm{MSE} = \frac{1}{N}\sum_{j=1}^{N}\left(c_i^j - \tilde{c}_i^j\right)^2 \tag{28}$$

For AC, values closer to 1 indicate better performance. For PC and PSNR, higher values are better.

5.2. Three-dimensional head numerical experiments

We first evaluate our proposed methods using a physically realistic three dimensional head model derived from T1-weighted MPRAGE scans originally acquired by Eggebrecht *et al* [8] that were kindly provided to us by the authors of that work. Following the process described by Wu et al [60], Statistical Parametric Mapping (SPM) software [61] was used to perform a parametric segmentation of the 5 tissue types (scalp, skull, cerebrospinal fluid (CSF), gray matter, white matter) based on the pixel intensity probability function distribution. These five different layers were then further processed in NIRFAST to create masks and layered volumetric FEM meshes.

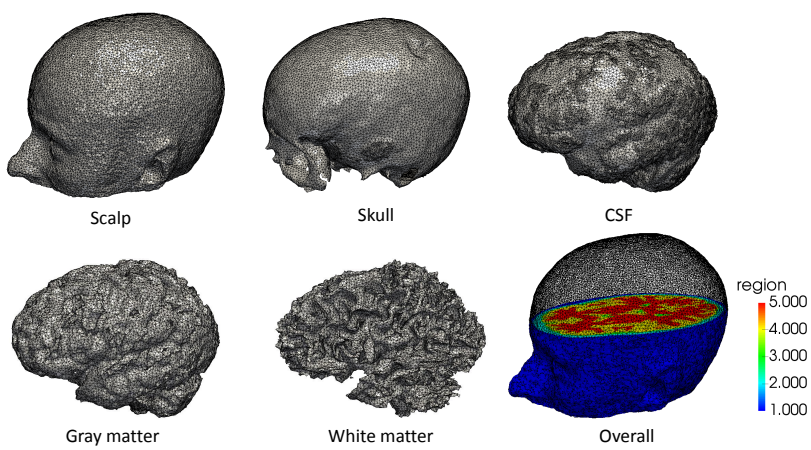


Fig. 3. Three-dimensional surface mesh for each of the five head layers.

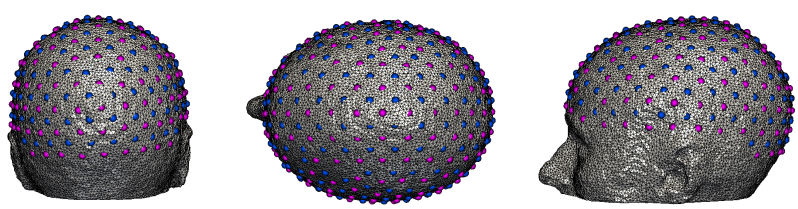


Fig. 4. Schematic view from three directions showing the distribution of the imaging array with 158 sources (blue circles) and 166 detectors (red circles).

The mesh is composed of 101046 nodes corresponding to 589658 tetrahedral elements. Each node is labelled by one of the five segmented head tissue types, as shown in Fig. 3. Chromophore concentrations assigned to each layer are computed from the tissue optical properties at 750nm and 850nm in a previous in vivo study [62] (Table 1) using Beer's law and Mie scattering formulae. A high-density (HD) imaging array containing 158 sources and 166 detectors (Fig. 4) [8] was

Biomedical Optics EXPRESS

placed over the whole head, with source-detector (SD) separation distances ranging from 1.3 to 4.8cm. In this study, 3478 differential measurements per wavelength were used to image the hemodynamic changes in the brain. Two individual activations were simulated in the visual cortex with chromophore concentrations of HbO_2 and Hb respectively increasing to double the background level in the gray matter (Fig. 5). Each simulated activation has a radius of 5mm. 0% to 2% distributed Gaussian noise at 0.5% intervals was added to the measurement vector.

Reconstructed chromophore concentrations of the simulated activation using the Tikhonov model (Eq. (10)) and the spectral-L_1 model (Eq. (12)) on noise-free data are displayed in Fig. 6, while those on data with 1% Gaussian noise are displayed in Fig. 7. We only show the area with changes in chromophore concentration greater than 0.0001mM. Compared to the results from the spectral-L_1 model, Tikhonov reconstructions have lower image contrast, which can be clearly seen from the first column of Fig. 6 and Fig. 7. Some artifacts (areas contained within green ellipses) can be easily observed around the source and detector areas. With increased levels of noise, larger artefacts are seen with Tikhonov regularization and the results are spatially smeared. In contrast, results from the spectral-L_1 model show fewer artifacts in the non-activation area. Higher noise does not noticeably affect the L_1-regularized reconstructions. IRLS produces visually more compact localizations than ADMM and FISTA, whilst ADMM appears to have better sparsity-inducing properties compared with IRLS and FISTA.

Table 1. Head tissue optical property for each of five layers.

	Scalp	Skull	Cerebrospinal Fluid(CSF)	Gray Matter	White Matter
c_1 (mM)	0.0575	0.0438	0.011	0.0548	0.0683
c_2 (mM)	0.0313	0.0209	0.0083	0.0354	0.0273
Scattering amplitude	0.53	0.7258	0.3	0.5040	0.8176
Scattering power	1.1599	0.8987	$0.9e^{-6}$	1.7757	1.3048
μ_a (mm^{-1})at 750nm	0.017	0.012	0.004	0.018	0.017
μ_s' (mm^{-1})at 750nm	0.74	0.94	0.3	0.84	1.19
μ_a (mm^{-1})at 850nm	0.019	0.014	0.004	0.019	0.021
μ_s' (mm^{-1})at 850nm	0.64	0.84	0.3	0.6726	1.0107

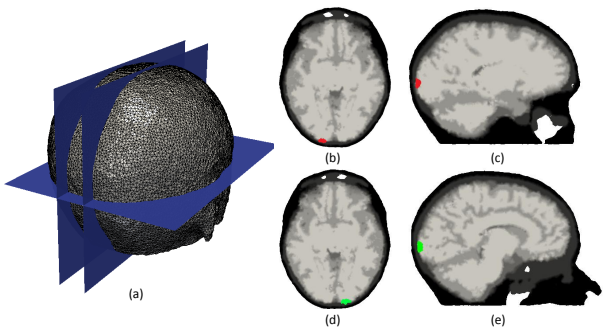


Fig. 5. Ground-truth image with the activation only exists in the gray matter and white matter. (a): Illustration of the overall distribution of slices. (b)-(c): Individual activation is color-coded in red and represents the individual simulation of HbO_2. (d)-(e): Individual activation is color-coded in green and represents the individual simulation of Hb.

Evaluation metrics from this experiments are shown in Fig. 8. It is clear that the spectral-L_1 model can achieve higher AC, PC and PSNR values than the Tikhonov model which means higher image contrast and accuracy can be achieved with L_1 regularization. Although the results of FISTA show more visual artifacts than other L_1-norm methods, it is still able to produce

Biomedical Optics EXPRESS

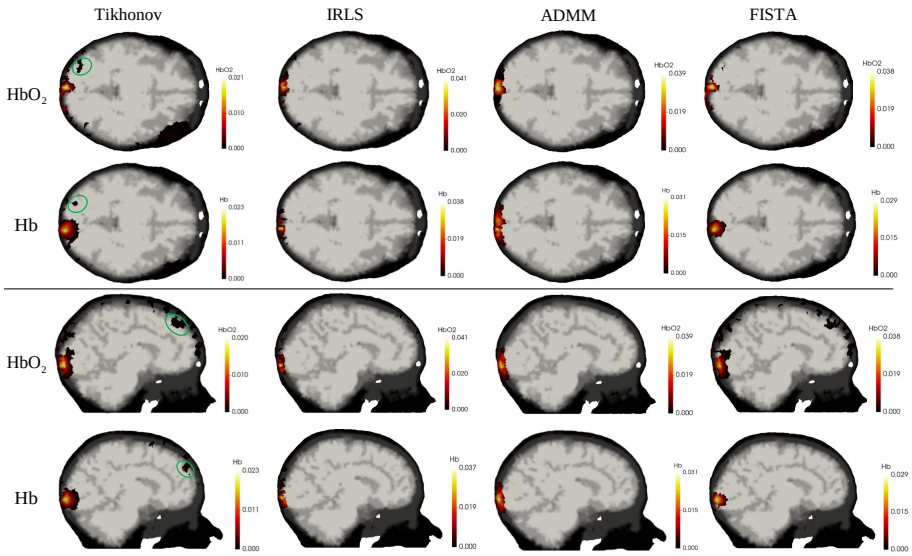


Fig. 6. The reconstructed image of the change of HbO_2 and Hb in mM with noise-free data. Some examples of reconstruction artefacts are highlighted in green ellipses.

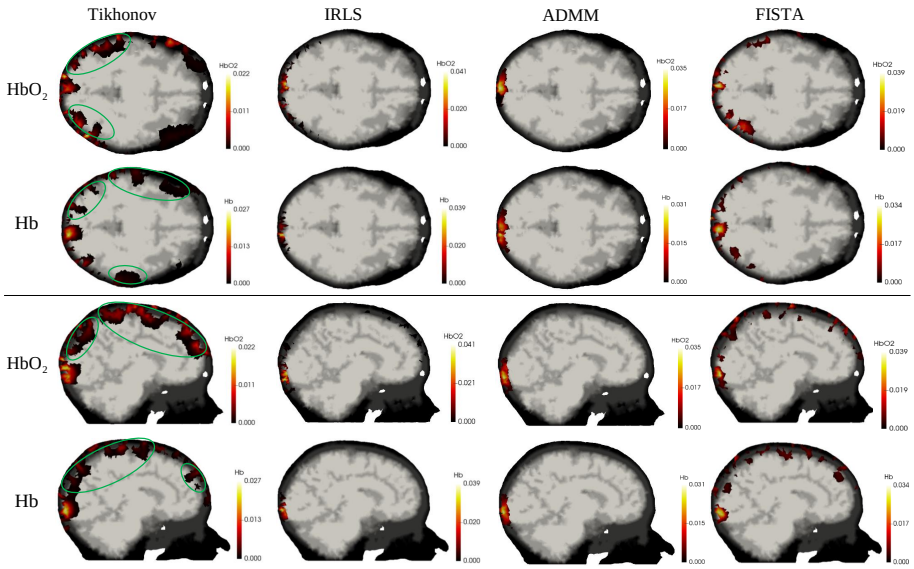


Fig. 7. The reconstructed image of the change of HbO_2 and Hb in mM with data contaminated by 1% Gaussian noise. Some examples of reconstruction artefacts are highlighted in green ellipses.

better performance based on the metrics. This is because (i) AC is defined on the activation region which is selected by thresholding the recovered changes based on 50% of the maximum recovered changes, artefacts away from this region do not influence this metric; (ii) By the other metrics (PC and PSNR), the improved ability of FISTA to localize the activation is sufficient to counteract the effect of the artefacts.

Biomedical Optics EXPRESS

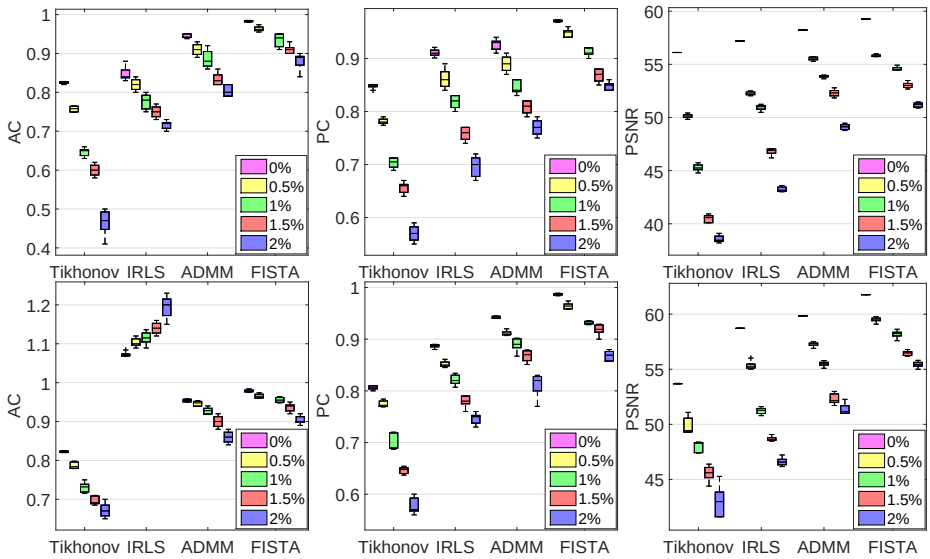


Fig. 8. Evaluation metrics comparing the performance of different methods on a simulated 3D head model at five different noise levels. Left to right column: AC index; PC index and PSNR index. The first row gives the results from HbO_2; the second row from Hb.

5.3. More realistic three dimensional head numerical experiments

Following the proof-of-concept experiments described in the previous section, we extended our analysis to a more realistic case with much smaller changes in chromophore concentrations. In the activation area we model a small region with changes in HbO_2 (c_1) of 5μM and Hb (c_2) of -5μM, relative to the background concentrations given in Table 1 (Fig. 9). The mesh is the same as that used in the previous section. In line with the expected in vivo performance of imaging systems, 0.12%, 0.15%, 0.41% and 1.42% Gaussian random noise was added to first (13mm), second (30mm), third (40mm) and fourth (48mm) nearest neighbour measurements to provide realistic data [63].

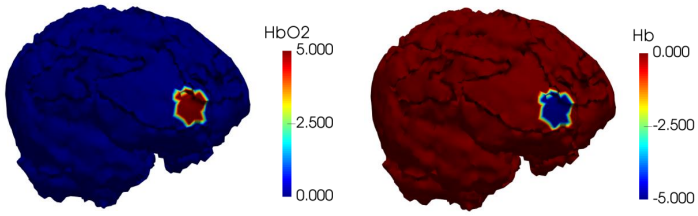


Fig. 9. Ground-truth image showing the change in chromophore concentration confined to the gray matter.

Reconstruction using the four methods considered here are shown in Fig. 10 with noise-free and noisy simulated data. With reference to results shown earlier in this paper, we make a similar observation that in comparison to the ground truth values, results using Tikhonov regularization are visually inferior to those from L_1 regularization. With increased noise, Tikhonov regularization performs progressively worse with more artefacts visible in the source-detector areas. L_1 regularization induces sparse results with fewer artefacts in non-activated areas. Visual inspection of the results from the three L_1 algorithms suggests that IRLS produces over-sparse

Biomedical Optics EXPRESS

reconstructions with strong activations confined to a small area. ADMM and FISTA results are much more visually realistic and they are seen to give higher quantitative accuracy. A quantitative evaluation using AC, PC and PSNR is given in Table 2 and Table 3 and these support the conclusion that even at small changes in chromophore concentration, the spectral-L_1 model can still guarantee higher image contrast and accuracy, with FISTA performing consistently better by all measures (AC closer to 1, higher PC, higher PSNR).

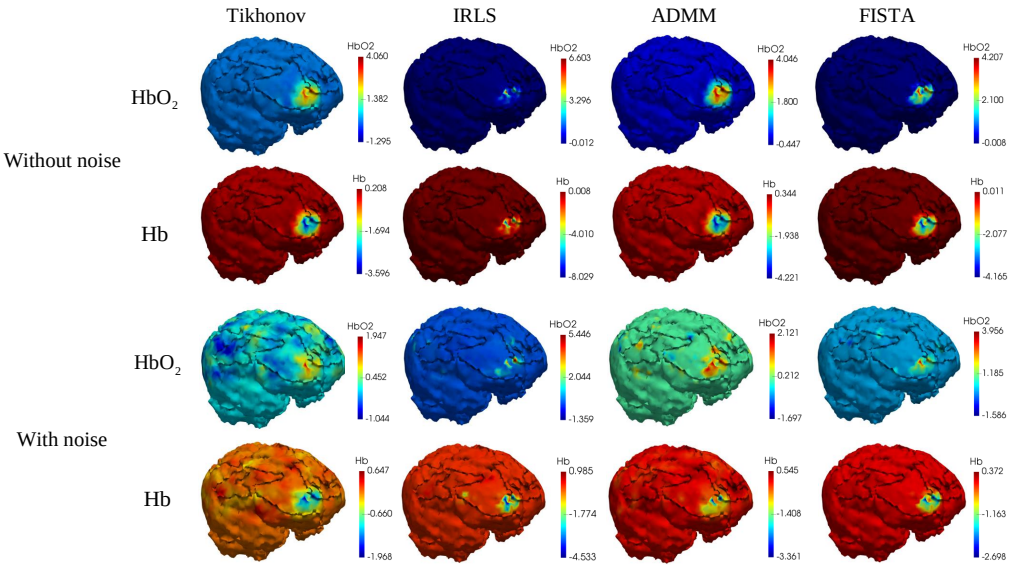


Fig. 10. Reconstruction of HbO_2 and Hb using (L-R): Tikhonov for L_2-norm regularization; IRLS, ADMM, FISTA algorithms for L_1-norm regularization with different noise levels. First two rows: results with clean simulated data; Last two rows: those with noisy data.

Table 2. Three evaluation metrics for HbO_2 on results by different methods

	Without noise				With noise			
	Tikhonov	IRLS	ADMM	FISTA	Tikhonov	IRLS	ADMM	FISTA
AC	0.7180	0.8132	0.8367	**0.9069**	0.6674	0.7188	0.7294	**0.7556**
PC	0.9997	0.9998	0.9998	**0.9999**	0.9973	0.9996	0.9996	**0.9998**
PSNR	73.8062	76.4843	77.3333	**78.5520**	38.082	63.7853	67.9154	**73.4270**

Table 3. Three evaluation metrics for Hb on results by different methods

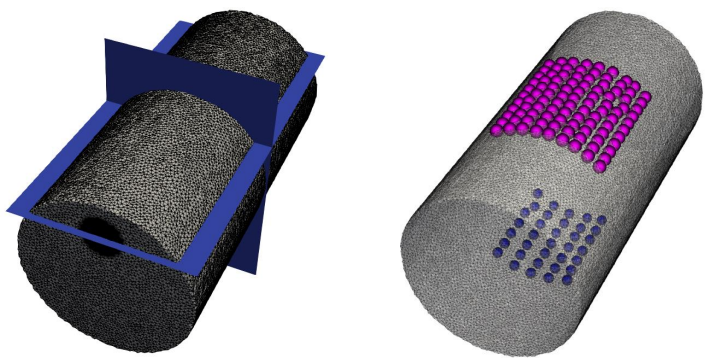
	Without noise				With noise			
	Tikhonov	IRLS	ADMM	FISTA	Tikhonov	IRLS	ADMM	FISTA
AC	0.8570	0.9394	0.9577	**0.9736**	0.7200	0.9086	0.9107	**0.9299**
PC	0.9997	0.9997	**0.9998**	**0.9998**	0.997	**0.9996**	**0.9996**	**0.9996**
PSNR	76.1785	76.9921	78.9364	**79.0713**	40.5900	66.3472	75.0008	**76.2553**

5.4. Experiments with phantom data

To evaluate the proposed algorithms on real experimental data, a multispectral, non-contact CW-DOT system designed for hand imaging [64, 65] was used to image a solid plastic cylindrical phantom (INO, Quebec, Canada) of radius 12.3mm and length 50mm. Boundary data was

Biomedical Optics EXPRESS

collected at five wavelengths (650nm, 710nm, 730nm 830nm and 930nm), in a transmission setup with a 7 x 5 grid of source positions on the underside of the phantom and a 11 x 9 grid of virtual detectors on top, displayed in Fig. 11(b). The spatially constant, but spectrally varying optical properties of the phantom were measured previously in time resolved experiments [66]. The absorbing dye within the phantom was treated as a chromophore that has unit concentration in the bulk of the phantom, the extinction coefficient of which was modelled by the measured absorption coefficient. A heterogeneous version of the phantom was imaged which contained a cylindrical rod of radius 3mm and length 50mm, at a depth of 5mm (Fig. 11(a)). The rod has twice the absorption coefficient of the bulk phantom which provides a 2:1 contrast in dye concentration compared to background (Fig. 12 left). A homogeneous version was also imaged, enabling calibration of the model/data mismatch and any source or detector coupling variation. The mesh, as shown in Fig. 11(a), consists of 85,205 nodes and 451,821 linear tetrahedral elements.



(*a*) (*b*)

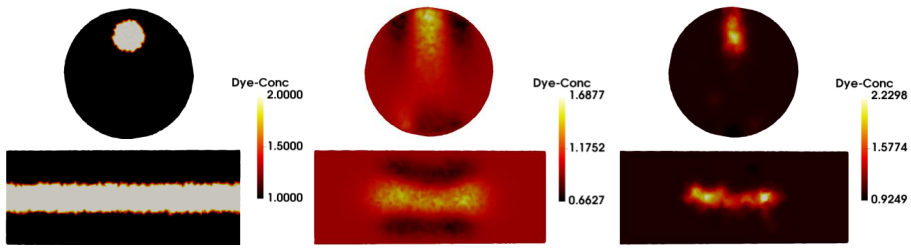
Fig. 11. (a): Illustration of the overall distribution of slices. (b): Distribution of sources and detectors.

Ground truth data and images reconstructed with L_2 and L_1 methods are shown in Fig. 12 respectively. The experiments described in the previous sections showed that the particular choice of L_1 method makes only a very small difference to the quality of the reconstruction, but there are very large differences in computational efficiency, with FISTA being far more efficient in this domain because of its superior ability to deal with large problems (as will be shown in section 5.5). Therefore in this experiment, we only use FISTA as the L_1 solver. It can be clearly observed from Fig. 12 that L_2 regularization over-smooths the reconstructed images which have much lower image contrast than the ground truth. Some artefacts can be seen in the source and detector areas. We note that only the central region can be reconstructed in both cases because the sources and detectors are confined to this region, with very low sensitivity away from the centre. The image contrast reconstructed by L_1 regularization is much closer to the ground truth but with more compact results. We calculate the three evaluation metrics in the volume of illumination (Table 4) and these support the same conclusions.

5.5. Comparison of CPU time consumed in the inverse model

We now compare the computational efficiency of the proposed methods. All experiments are performed using Matlab 2013a (Mathworks, Natick, USA) on a Windows 7 (Microsoft, Redmond, USA) platform with an Intel Core CPU i7-6700 at 3.40GHz and 64.0GB memory. The simulated experiments described in section 5.2 were used to perform this comparison. CPU times used in the inverse procedure only are measured. We run each method over ten realisations of noise at each of five noise levels to obtain reliable statistics. Figure 13 shows the CPU time consumed for the four different methods (Tikhonov, IRLS, ADMM, FISTA). In order to display the recorded times

Biomedical Optics EXPRESS



(*c*)

Fig. 12. Ground truth and reconstruction results with different regularizations. From left to right: ground truth; results with L_2 regularization; results with L_1 regularization using FISTA algorithm.

Table 4. Evaluation of L_1 and L_2 regularization methods for reconstruction of a single rod inclusion in a tissue-simulating phantom.

	L_2 Regularization	L_1 Regularization with FISTA
AC	0.7187	1.1387
PC	0.6617	0.7263
PSNR	13.7472	13.8821

from ten repetitions clearly, CPU times for one iteration of the outer loop of the reconstruction algorithm are shown in Fig. 13. Total CPU times for iteration to convergence are given in Table 5.

FISTA is clearly the fastest L_1 regularization method amongst those considered here, and it is faster even than Tikhonov regularization which does not use an inner iteration. FISTA only involves the computation of J^TJ which is much more computationally efficient than the computation of JJ^T. IRLS and ADMM are substantially slower because they require an inner iteration and inversion/multiplication of large matrices.

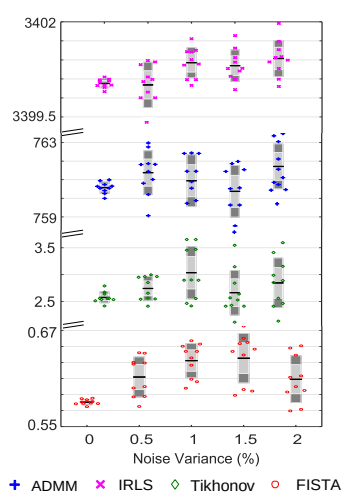


Fig. 13. Total CPU time consumed in the experiments described in section 5.2

Biomedical Optics EXPRESS

Table 5. Total CPU time(s) consumed in the inverse model for the experiments described in section 5.2

Evaluation metric	CPU time (Mean±SD)				
Chromophore	0%	0.5%	1%	1.5%	2%
Tikhonov	42.7902±0.211	33.4124±1.120	35.4658±1.399	33.4156±1.607	14.5622±1.463
IRLS	10171.96±0.019	10171.95±1.414	6823.03±0.896	6821.99±0.912	6822.01±1.274
ADMM	9100.34±0.508	7444.31±1.373	6195.01±1.583	4193.57±1.442	2892.30±1.200
FISTA	**2.532±0.092**	**1.964±0.193**	**1.844±0.154**	**1.850±0.218**	**1.927±0.307**

6. Conclusion

In this paper, a new model for spectrally constrained DOT reconstruction with L_1 regularization is proposed. Numerical experiments showed that compared to the L_2-norm, L_1 regularization can reduce crosstalk and maintain image contrast by inducing sparsity. These findings were tested on real experimental data using a tissue-simulating phantom and similar results were found. Although all L_1-based methods perform similarly in terms of reconstruction quality, FISTA performs marginally better than ADMM and IRLS by the measures of AC, PC, and PSNR, and is far more computationally efficient as it avoids direct matrix inversion and large matrix-matrix multiplications.

The contributions of this paper can be summarized as follows: 1) It is the first time that L_1 regularization methods and spectrally constrained DOT methods have been used together and it is their combination (i.e. spectral-L_1 model) that is original. We have given detailed descriptions of how this can be done, and performed systematic comparisons of the performance and efficiency of the different methods on both simulated and real data. We are not aware of any previously published work that proposes such a model and performs such a detailed analysis of its performance; 2) We have developed a method to automatically select the regularization parameters. This is based on the L-curve method but is modified for efficient application in this use-case; 3) We have conducted extensive numerical experiments on simulated data and on real experimental data, and have performed comprehensive and robust qualitative and quantitative evaluations. To the best of our knowledge, this is the first systematic study in the area of spectral DOT reconstruction to perform such a comprehensive evaluation.

Funding

European Union Horizon 2020 Marie Sklodowska-Curie BITMAP Innovative Training Network (675332). EPSRC Physical Sciences for Health Centre for Doctoral Training (EP/L016346/1) (studentship award to Daniel Lighter).

Acknowledgements

We are grateful to Joe Culver and Adam Eggebrecht from Washington University School of Medicine for supplying the MRI scans used in this work. Data supporting this research is openly available from the University of Birmingham data archive at http://findit.bham.ac.uk/.

Disclosures

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