

Information transport and limits of optical imaging in the highly diffusive regime

Jack Radford^{✉*} and Daniele Faccio^{✉†}

School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom



(Received 27 October 2022; accepted 22 December 2022; published 7 April 2023)

Imaging in the highly diffusive regime lies at the heart of various optical medical imaging technologies and other applications such as imaging through fog. However, due to the randomization of the photon propagation direction inside random, complex media, the spatial information of physiological features is degraded such that most diffuse optical imaging techniques are restricted to <10 transport mean free paths (ℓ^*). We present an information theoretical analysis of the limits of gathering information in the highly diffusive regime. Our results show that there is still information even when detection is performed at the single-photon level and beyond $200\ell^*$ (e.g., ~ 12 cm tissue). Image reconstructions are enhanced when resolving measurements in both space and time domains. These results provide a general framework for extremely deep diffuse imaging scenarios, such as imaging through the body, and a general context for developing optimized image retrieval strategies.

DOI: [10.1103/PhysRevResearch.5.L022008](https://doi.org/10.1103/PhysRevResearch.5.L022008)

Introduction. Overcoming the challenge of imaging through thick scattering materials will enable a paradigm shift in optical medical imaging devices for the human body and brain which could replace some of the functionality of magnetic resonance imaging, positron emission tomography, and computed tomography scanners with cheaper and nonharmful alternatives [1–12]. Highly scattering materials have a much greater probability of scattering photons than absorbing them, and the randomization of the directional information of each photon results in structural degradation of the transmitted image as well as attenuation of the number of photons transmitted in the forward direction through the material. The average distance traveled by a photon before it is absorbed or loses information about its initial propagation direction is quantified by the transport mean free path (TMFP) length $\ell^* = 1/(\mu'_s + \mu_a)$, where the reduced scattering coefficient μ'_s and the absorption coefficient μ_a of the medium describe the average number of scattering and absorption events experienced by a photon per unit length, respectively. Materials for which the total distance a photon has propagated through the material $L \gg \ell^*$ are said to be in the diffusive regime [13,14].

The task of reconstructing images in the highly diffusive regime is nontrivial since the problem is ill-posed [15–18], although computational strategies have successfully improved the condition of the inverse problem [19–29] and enabled imaging through $>80\ell^*$ [30]. In relatively thin materials $<10\ell^*$, there is a significant probability of detecting, at very early times, ballistic photons that are transmitted without any scattering events and which preserve spatial information

[31–37]. For thicker materials, time-gating is used to isolate early-arriving photons which have less interaction with the material and are more likely to preserve information about their initial propagation direction [18,38–44]. However, it has been shown that early-arriving photons give little advantage when imaging beyond $80\ell^*$ since all of the detected photons are highly scattered [45].

A pioneering study by Latimer [46] showed that concepts derived from information theory can be useful tools to analyze the statistical nature of light scattering for the determination of particulate size distributions. More recently, universal bounds for the limit of channel capacity through a scattering material have been derived [47] using ideal conditions. However, there still remains an open question regarding the actual distance over which it is ultimately possible to propagate image information in diffusive media in realistic conditions.

In this Rapid Communication, we use information theory to show that, when considering practical constraints on the problem such as shot noise and reasonable acquisition times, there remains information for imaging beyond $200\ell^*$. To put this in context, this implies that information is preserved even when the signal is effectively reduced to the single-photon level at the output of the diffusive medium. This information is not simply encoded in spatial degrees of freedom as one might expect based on more traditional imaging scenarios; there is an improved conditionality of the image retrieval inverse problem when measuring the full space-time-resolved photon counts. This indicates the possibility of imaging significantly beyond what has currently been achieved, effectively indicating a regime for diffuse imaging in future experiments.

Physics of the problem. The photon-diffusion approximation accurately describes photon statistics in the diffusive regime [14,48]:

$$\Phi(\vec{r}, t) = \frac{c}{(4\pi Dct)^{3/2}} \exp\left(-\frac{|\vec{r}|^2}{4Dct} - \mu_a ct\right), \quad (1)$$

where $\Phi(\vec{r}, t)$ is the photon fluence rate at the output (measured in W/cm^2), $D = 1/[3(\mu'_s + \mu_a)]$ is the diffusion

*j.radford.1@research.gla.ac.uk

†daniele.faccio@glasgow.ac.uk

coefficient, \bar{r} is the radial distance of the output surface of the material from a source/object pixel, t is time, and c is the speed of light in the medium.

We can therefore simulate experiments using a forward operator $\mathcal{A}(m)$ to model the spatiotemporal measurement of diffuse photons in the presence of an embedded absorbing object in transmission through a highly scattering material. This is achieved by approximating the intensity profile of a laser pulse $I(x, y, t)$ with a Kronecker delta in discretized time (55 ps time bins) applied uniformly over space and convolving with the photon fluence rate $\Phi(\bar{r}_m, t)$ corresponding to mid-way through the material. The result is then multiplied with a binary mask in the shape of the absorbing object we wish to image $m(x, y)$ and then convolving with the photon fluence rate at the output plane of the material $\Phi(\bar{r}_{\text{out}}, t)$:

$$\mathcal{A}(m) = \Phi(\bar{r}_{\text{out}}, t) * [[I(x, y, t) * \Phi(\bar{r}_m, t)] \odot m(x, y)_T], \quad (2)$$

where $*$ denotes a convolution and \odot denotes an element-wise multiplication. The radial vector $\bar{r}_m = \sqrt{x^2 + y^2 + L_m^2}$ is the three-dimensional (3D) Euclidean distance from the source position to the depth of the hidden absorbing mask L_m , and \bar{r}_{out} is the radial distance from the depth of the mask to the output plane of material with thickness L . The binary mask $m(x, y)_T$ is repeated T times (the total number of time bins) to multiply element-wise with the spatial profile of the laser pulse $I(x, y, t)$ at every time step.

The probability mass function is obtained from the normalized output photon fluence rate multiplied by the sum of photons expected at the detector N_p . The data are finally quantized into bins by rounding every element to the nearest integer.

The final simulation output is a 3D spatiotemporal data cube which can be integrated in space or time to generate time- or space-only data, respectively. In this case, space-only measurements are equivalent to using a continuous-wave laser.

The simplicity of this forward model is chosen since it must be computationally inexpensive to simulate enough examples to adequately populate statistical distributions required for analysis. Although the simulations are limited to a two-dimensional (2D) absorbing occlusion inside a homogeneous material, this is like simplified real-world scenarios such as imaging regions of blood absorption in layers of the human head with homogeneous properties or when imaging a single absorbing object through fog with uniform particle density.

Information measured in space and time domains. We use the Shannon entropy to measure the expected reduction in uncertainty, measured in bits, upon observing a discrete random variable $x_i \in X$, with probability $p(x_i)$ [49–51]:

$$H(X) = - \sum_i p(x_i) \log_2 p(x_i). \quad (3)$$

The aim is to quantify the information contained in the data that allows us to discriminate between N different absorbing objects (e.g., $N = 1024$ binary EMNIST images [52] in our study) hidden behind a homogeneous, thick scattering material when observing the number of photon counts $c_i \in C$ at given detector space and time coordinates (x, y, t) :

$$H(C) = - \sum_i^K p(c_i|x, y, t) \log_2 p(c_i|x, y, t), \quad (4)$$

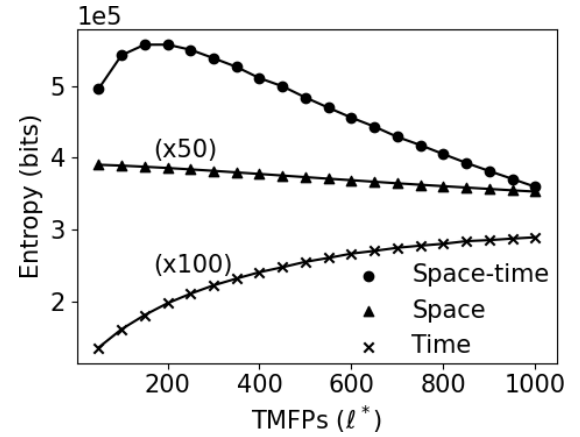


FIG. 1. The information content for a fixed number of measured photons $N_p = 1 \times 10^8$ for increasing number of transport mean free paths (TMFPs) when using both space and time information of the raw data (\bullet), integrating the time trace for every pixel to give only a number of photon counts over space (multiplied by 50; \blacktriangle) and integrating the spatial pixels to one time trace (multiplied by 100; \times). Using data resolved in both space and time gives orders of magnitude more information than only measuring in the space domain.

where K is the number of bins in the probability mass function. The total information content in the data is found by summing $H(C)$ for every pixel coordinate (x, y, t) .

Equation (4) is evaluated in each domain to compare information content for different measurement regimes. The results in Fig. 1 show, as may be expected, that information in space decreases for increasing number of TMFP lengths since the images become increasingly blurry and less diverse. Interestingly, the information in time increases with medium thickness. We interpret this counterintuitive finding to be due to the fact that the distribution of photon arrival times becomes spread out over an increasing number of time bins, and the signals are therefore increasingly diverse between different objects. However, there are orders of magnitude more information collected at the detector when using fully resolved space-time data. This is due to the much higher dimensionality of the data which enables more diversity in the measurements compared with only space or only time domain data. We also find that there is still information beyond $1000\ell^*$ (equivalent, e.g., to 60 cm for a medium with tissuelike properties $\mu'_s = 0.09 \text{ cm}^{-1}$ and $\mu_a = 16.5 \text{ cm}^{-1}$ [53]). However, these results do not account for photon attenuation (Fig. 1 considers a fixed number of 10^8 output photons for all TMFPs) or noise, which is addressed in the following section.

Practical considerations. The fluence rate (Φ) of the diffuse photons incident on the detector can be approximated by [54,55]

$$\Phi = \Phi_0 \frac{\delta\Omega}{4\pi} \exp[-\sqrt{3\mu_a(\mu'_s + \mu_a)}L], \quad (5)$$

where $\delta\Omega/4\pi$ is chosen, based on experimental considerations [45], to be of the order of 10^{-5} and accounts for the solid-angle fraction of light which reaches the lens of the detector. A modest value of 10 mW power was used for a $5 \times 5 \text{ cm}$ area, corresponding to $\Phi_0 = 4 \text{ W/m}^2$, which is ~ 3 orders of magnitude lower than the British standards maximum permissible exposure for human skin [56]. Using

Eq. (5), the number of transmitted photons N_p can be calculated (discretized by rounding to the nearest integer) in a $T = 60$ s total acquisition time. To simulate noise, a Poisson distribution is sampled for each bin where the mean is given by its number of photons.

Mutual information between input and output. To disambiguate redundant repeated information included when summing the entropy of every element of the measurement in Fig. 1, we consider the mutual information between the input pixel values and the measurements. Mutual information between input variables $x_i \in X$ and output variables $y_i \in Y$ is given by

$$MI(X; Y) = H(X) + H(Y) - H(X, Y), \quad (6)$$

where $H(X, Y) = -\sum_j \sum_i p(x_i, y_j) \log_2 p(x_i, y_j)$ is the joint entropy between the variables [49–51]. We consider the input variables to be the pixels at the plane at which the hidden absorbing objects are placed and the output variables are the measurement pixels at the detector plane.

Independent channels are identified by calculating the impulse response of the system $Q \in \mathbb{R}^{m \times n}$ when scanning the pixel-wise canonical basis at the plane of the absorbing hidden object, where m is the number of elements at the output (i.e., number of pixels/bins depending on the domain) and n is the number of elements at the input (i.e., pixels at absorption plane).

The inner product of the normalized measured data $R \in \mathbb{R}^{m \times 1}$ and the canonical impulse response basis calculates the projection coefficients for every independent channel output $Y = \langle R, Q \rangle$, where every column of Q is a unit vector. A probability mass function $P(Y)$ of projection coefficients is constructed by repeating for $i = 1024$ examples and binning the coefficients as binary high/low using the mean as a threshold. This reduces every input pixel (x_i) and projection coefficient (y_i) pair to be related by a binary symmetric channel [50] in which the probability of error when transmitting a bit is introduced by cross-talk noise and is dependent on the statistical fluctuation of every other channel by the process of photon diffusion. The mutual information between the distribution of input pixel values $P(X)$ and the projection coefficients $P(Y)$ for every channel is calculated independently using Eq. (6), where $H(X)$ and $H(Y)$ are found by substituting the probability mass functions into Eq. (3). Figure 2(a) shows the average mutual information per channel when using measurements resolved in space compared with space-time for noise-free and noisy data. There exists information beyond $200\ell^*$ when considering both attenuation and shot noise which, for typical tissue parameters such as $\mu'_s = 16.5 \text{ cm}^{-1}$ and $\mu_a = 0.09 \text{ cm}^{-1}$ [53], corresponds to ~ 12 cm of tissue. The limit of $\sim 200\ell^*$ in this regime is restricted by the lack of photons rather than the loss of information due to scattering.

We conducted an experiment to validate the assumptions about the number of transmitted photons in simulation using a phantom scattering material (see Supplemental Material [57] for experimental details). The experimental results confirm that the transmission of photons is possible at $226\ell^*$ using a 10 mW source and $266\ell^*$ for a 1 W source for the parameters used in simulation and can be extended to $293\ell^*$ and $333\ell^*$, respectively, for a system with greater solid angle of collection ($\delta\Omega/4\pi = 0.022$).

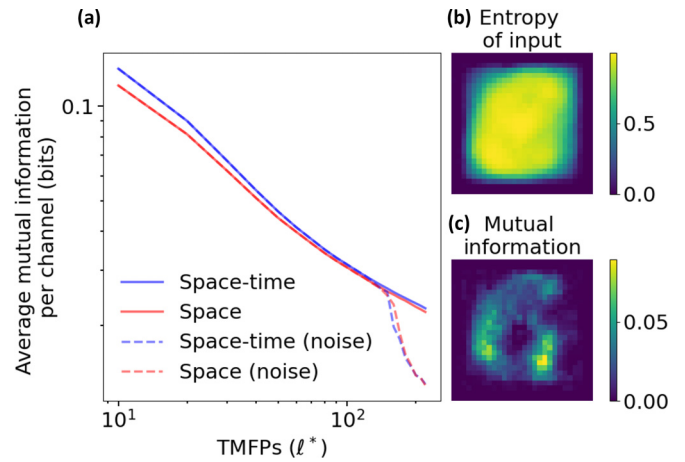


FIG. 2. (a) The average mutual information per channel for increasing transport mean free paths. Resolving measurements in both space and time reduces cross-talk between pixels and preserves more information through highly scattering materials. The same curves including also noise are shown in dashed lines: There is still information available beyond $200\ell^*$. (b) The entropy of the pixels and (c) the mutual information (in bits) between the input pixel values and projection coefficients for space-time-resolved measurements of 1024 EMNIST examples through $50\ell^*$.

Figure 2(a) also confirms that there is higher mutual information in the space-time domain, i.e., that there is increased diversity of the measurements compared with measuring in only the space domain.

Figure 2(b) shows the entropy of the input pixels for the 1024 binary EMNIST images, i.e., it shows the areas/pixels of high amounts of transmitted information. However, it is the distribution of mutual information, shown in Fig. 2(c) for the space-time domain data and a diffuser thickness of $50\ell^*$ that indicates the most information-rich regions of the measurements for a given set of images and physical configuration. The mutual information for each pixel/channel has a distinct distribution with a maximum at the outer edges and a minimum at the center where the highest number of photons is typically measured. We interpret this distribution to be due to the high cross-talk, i.e., photon blurring across pixels that occurs in the center channels compared with the lower cross-talk on the edge channels that are surrounded by areas of zero-photon counts.

Condition of the inverse problem. Ultimately, the goal of diffuse imaging is to retrieve images of objects. A typical inverse problem will be ill-conditioned if the ratio between the largest and smallest singular values of the forward operator A (σ_{\max} and σ_{\min} , respectively) is large. Therefore, we use the condition number $\kappa = \sigma_{\max}/\sigma_{\min}$ as a measure of the stability of the linear mapping and a quantification of the potential inaccuracy when reconstructing images in the absence of any priors [17,18,58,59]. The inverse problem of reconstructing images from the measured data is solved as the pseudo-inverse $A^{-1} = V\Sigma^{-1}U^T$, where V , Σ , and U are the left singular vector, singular values, and right singular vector, respectively. Figure 3(a) shows the condition numbers κ for the linear forward operators A using the space, time, and space-time resolved data. It is clear that using only time information is highly ill-posed, irrespective of the number of TMFPs. How-

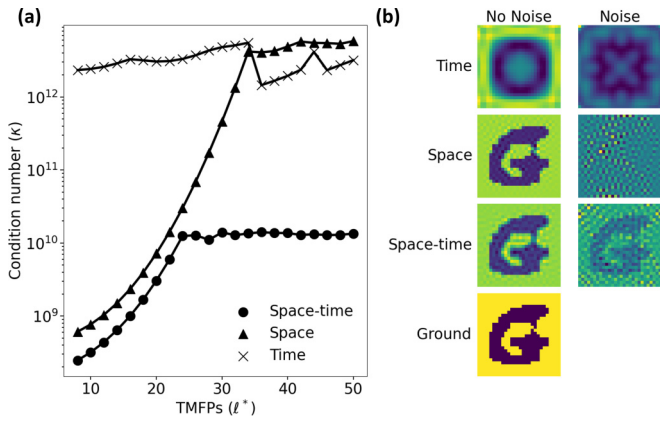


FIG. 3. (a) The condition number (κ) of the forward linear operator between when using measurements resolved in space-time (\bullet), space (\blacktriangle), and time (\times). As the thickness of a material increases, the problem becomes more ill-conditioned, but using the fully resolved space-time data can give orders of magnitude of improvement. (b) The reconstructed images of a target absorbing object through $50\ell^*$ when using the different measurement domains. When adding a small noise perturbation, space and time domain solutions are far from the ground truth, whereas there is a resemblance of the target when using space-time measurements.

ever, resolving the data in both space and time has advantages in reducing the condition number by increasing orders of magnitude (for increasing TMFPs) compared with using data resolved only in space or time.

The impact on the conditioning of the problem is illustrated in practical terms in Fig. 3(b) with image reconstructions. As a treatable example that does not require regularization terms nor any *a priori* information, we chose to reconstruct an image from photons propagated through $50\ell^*$. This is a nontrivial example which is, however, still solvable using a direct linear inversion and therefore does not require an in-depth analysis of different inversion techniques, which is beyond the scope of this Rapid Communication.

As can be seen, only the space-time data allow us to reconstruct an image that resembles the target object with both space- and time-only data failing due to the presence of noise. We underline that the purpose of these reconstructions is not to reproduce good images using regularization but to illustrate the improved conditioning of the problem at full rank when resolving in both space and time.

The target image used in Fig. 3 is arbitrary, and the reconstructed images are not an indication of the best possible image retrieval solutions—these would need to be studied separately in future work.

Conclusions. Using information theoretical analysis of a typical diffuse imaging experiment, we show that it is possible to measure information about a hidden absorption object through thicknesses $>200\ell^*$ using practical acquisition parameters, which corresponds to ~ 12 cm of human tissue. Remarkably, the limit of information transfer through diffusive material in this regime is defined by the presence of photons at the output plane, i.e., if there is a photon, there is information. This apparently simple conclusion is nontrivial in the presence of such strong diffusion, as one might expect the nondeterministic path of a photon through the material, where direction is randomized in many thousands of individual scattering events, could completely erase spatial information about interactions with an absorbing object inside the material.

We show that resolving photon counts in both space and time preserves more mutual information and that even a relatively small increase in information can directly improve the ill-conditioned inverse problem of image reconstruction, which results in more accurate solutions. Although tested in the absence of any regularization, the improvements in image retrieval will generally impact more complex inversion formulations. When combined with the observation of how the mutual information is distributed across the sensor at the output, one can start to formulate optimization strategies for data acquisition, by focusing only on those regions that contain the most information about the actual object. This is similar in spirit, for example, to recent work using machine learning to learn how to position sensors for compressive sensing [59] but instead, here, is guided by information theory rather than neural network extraction of the statistical properties of the measurement data.

It is beyond the scope of this Rapid Communication to determine how best to use the information in the measurements for image retrieval. Indeed, this Rapid Communication is not about deciding the best image retrieval policy but rather investigating the information limits for imaging which is prerequisite for implementing any given imaging protocol. This Rapid Communication therefore uncovers the potential for other bespoke studies to design information-rich source/detector arrangements and create algorithms that will enable the paradigm shift of imaging through diffusive media beyond $100\ell^*$.

Acknowledgments. The authors acknowledge funding from Engineering and Physical Sciences Research Council (EPSRC, UK, Grant No. EP/T00097X/1). D.F. is supported by the Royal Academy of Engineering under the Chairs in Emerging Technologies scheme. J.R. is supported by the EPSRC Centre for Doctoral Training in Intelligent Sensing and Measurement, Grant No. EP/L016753/1.

[1] A. Pifferi, D. Contini, A. D. Mora, A. Farina, L. Spinelli, and A. Torricelli, New frontiers in time-domain diffuse optics, a review, *J Biomed. Opt.* **21**, 091310 (2016).
 [2] R. J. Cooper, E. Magee, N. Everdell, S. Magazov, M. Varela, D. Airantzis, A. P. Gibson, and J. C. Hebden, MON-STIR II: A 32-channel, multispectral, time-resolved optical

tomography system for neonatal brain imaging, *Rev. Sci. Instrum.* **85**, 053105 (2014).
 [3] A. T. Eggebrecht, B. R. White, S. L. Ferradal, C. Chen, Y. Zhan, A. Z. Snyder, H. Deghani, and J. P. Culver, A quantitative spatial comparison of high-density diffuse optical tomography and fMRI cortical mapping, *NeuroImage* **61**, 1120 (2012).

- [4] A. D. Mora, D. Contini, S. Arridge, F. Martelli, A. Tosi, G. Boso, A. Farina, T. Durduran, E. Martinenghi, A. Torricelli, and A. Pifferi, Towards next-generation time-domain diffuse optics for extreme depth penetration and sensitivity, *Biomed. Opt. Express* **6**, 1749 (2015).
- [5] M. Alayed and M. Jamal Deen, Time-resolved diffuse optical spectroscopy and imaging using solid-state detectors: Characteristics, present status, and research challenges, *Sensors* **17**, 2115 (2017).
- [6] H. Zhao and R. J. Cooper, Review of recent progress toward a fiberless, whole-scalp diffuse optical tomography system, *Neurophotonics* **5**, 1 (2017).
- [7] Y. Yamada and S. Okawa, Diffuse optical tomography Present status and its future, *Opt. Rev.* **21**, 185 (2014).
- [8] C. Dunsby and P. M.W. French, Techniques for depth-resolved imaging through turbid media including coherence-gated imaging, *J. Phys. D* **36**, R207 (2003).
- [9] A. P. Gibson and H. Dehghani, Diffuse optical imaging, *Phil. Trans. R. Soc. A* **367**, 3055 (2009).
- [10] M. J. Niedre, G. M. Turner, and V. Ntziachristos, Time-resolved imaging of optical coefficients through murine chest cavities, *J Biomed. Opt.* **11**, 064017 (2006).
- [11] V. Ntziachristos, A. G. Yodh, M. Schnall, and B. Chance, Concurrent MRI and diffuse optical tomography of breast after indocyanine green enhancement, *Proc. Natl. Acad. Sci. USA* **97**, 2767 (2000).
- [12] 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 Process. Mag.* **18**, 57 (2001).
- [13] Z. Q. Zhang, I. P. Jones, H. P. Schriemer, J. H. Page, D. A. Weitz, and P. Sheng, Wave transport in random media: The ballistic to diffusive transition, *Phys. Rev. E* **60**, 4843 (1999).
- [14] K. M. Yoo, F. Liu, and R. R. Alfano, When Does the Diffusion Approximation Fail to Describe Photon Transport in Random Media? *Phys. Rev. Lett.* **64**, 2647 (1990).
- [15] H. Dehghani, S. Sri Nivasan, B. W. Pogue, and A. Gibson, Numerical modelling and image reconstruction in diffuse optical tomography, *Philos. Trans. R. Soc. A* **367**, 3073 (2009).
- [16] J. Ripoll, D. Yessayan, G. Zacharakis, and V. Ntziachristos, Experimental determination of photon propagation in highly absorbing and scattering media, *J. Opt. Soc. Am. A* **22**, 546 (2005).
- [17] P. C. Hansen, *Discrete Inverse Problems* (Society for Industrial and Applied Mathematics, Philadelphia, 2010).
- [18] F. Leblond, H. Dehghani, D. Kepshire, and B. W. Pogue, Early-photon fluorescence tomography: Spatial resolution improvements and noise stability considerations, *J. Opt. Soc. Am. A* **26**, 1444 (2009).
- [19] S. R. Arridge, Optical tomography in medical imaging, *Inverse Probl.* **15**, R41 (1999).
- [20] S. R. Arridge, Methods in diffuse optical imaging, *Trans. R. Soc. A* **369**, 4558 (2011).
- [21] J. C. Hebden and D. T. Delpy, Enhanced time-resolved imaging with a diffusion model of photon transport, *Opt. Lett.* **19**, 311 (1994).
- [22] W. Cai, B. B. Das, F. Liu, M. Zevallos, M. Lax, and R. R. Alfano, Time-resolved optical diffusion tomographic image reconstruction in highly scattering turbid media, *Proc. Natl. Acad. Sci. USA* **93**, 13561 (1996).
- [23] Y. Xu, N. Iftimia, H. Jiang, L. Lyndon Key, and M. B. Bolster, Three-dimensional diffuse optical tomography of bones and joints, *J Biomed. Opt.* **7**, 88 (2002).
- [24] S. D. Konecky, G. Y. Panasyuk, K. Lee, V. Markel, A. G. Yodh, and J. C. Schotland, Imaging complex structures with diffuse light, *Opt. Express* **16**, 5048 (2008).
- [25] G. Satat, B. Heshmat, D. Raviv, and R. Raskar, All photons imaging through volumetric scattering, *Sci. Rep.* **6**, 33946 (2016).
- [26] F. Tonolini, J. Radford, A. Turpin, D. Faccio, and R. Murray-Smith, Variational inference for computational imaging inverse problems, *J. Mach. Learn. Res.* **21**, 146 (2020).
- [27] F. Schiffers, L. Fiske, P. Ruiz, A. K. Katsaggelos, and O. Cossairt, Imaging through Scattering Media with a Learning Based Prior, *Proc. IS&T Int'l. Symp. on Electronic Imaging: Computational Imaging XVIII* **32**, 306-1 (2020).
- [28] J. Yoo, S. Sabir, D. Heo, K. H. Kim, A. Wahab, Y. Choi, S. I. Lee, E. Y. Chae, H. H. Kim, Y. M. Bae *et al.*, Deep learning diffuse optical tomography, *IEEE Trans. Med. Imaging* **39**, 877 (2020).
- [29] J. Han, S. Zhu, E. Guo, J. Gu, and L. Bai, Imaging through unknown scattering media based on physics-informed learning, *Photonics Res.* **9**, 210 (2021).
- [30] A. Lyons, F. Tonolini, A. Boccolini, A. Repetti, R. Henderson, Y. Wiaux, and D. Faccio, Computational time-of-flight diffuse optical tomography, *Nat. Photonics* **13**, 575 (2019).
- [31] L. Wang, P. P. Ho, C. Liu, G. Zhang, and R. R. Alfano, Ballistic 2-D imaging through scattering walls using an ultrafast optical Kerr gate, *Science* **253**, 769 (1991).
- [32] V. Gopal, S. Mujumdar, H. Ramachandran, and A. K. Sood, Imaging in turbid media using quasi-ballistic photons, *Opt. Commun.* **170**, 331 (1999).
- [33] S. Woo, M. Kang, C. Yoon, T. D. Yang, Y. Choi, and W. Choi, Three-dimensional imaging of macroscopic objects hidden behind scattering media using time-gated aperture synthesis, *Opt. Express* **25**, 32722 (2017).
- [34] A. V. Kanaev, A. T. Watnik, D. F. Gardner, C. Metzler, K. P. Judd, P. Lebow, K. M. Novak, and J. R. Lindle, Imaging through extreme scattering in extended dynamic media, *Opt. Lett.* **43**, 3088 (2018).
- [35] M. G. Tanner, T. R. Choudhary, T. H. Craven, B. Mills, M. Bradley, R. K. Henderson, K. Dhaliwal, and R. R. Thomson, Ballistic and snake photon imaging for locating optical endomicroscopy fibres, *Biomed. Opt. Express* **8**, 4077 (2017).
- [36] B. Brezner, S. Cahen, Z. Glasser, S. Sternklar, and E. Granot, Ballistic imaging of biological media with collimated illumination and focal plane detection, *J Biomed. Opt.* **20**, 076006 (2015).
- [37] S. Maruca, P. Rehain, Y. M. Sua, S. Zhu, and Y. Huang, Non-invasive single photon imaging through strongly scattering media, *Opt. Express* **29**, 9981 (2021).
- [38] K. Chen, L. T. Perelman, Q. Zhang, R. R. Dasari, and M. S. Feld, Optical computed tomography in a turbid medium using early arriving photons, *J. Biomed. Opt.* **5**, 144 (2000).
- [39] G. M. Turner, G. Zacharakis, A. Soubret, J. Ripoll, and V. Ntziachristos, Complete-angle projection diffuse optical tomography by use of early photons, *Opt. Lett.* **30**, 409 (2005).
- [40] G. M. Turner, A. Soubret, and V. Ntziachristos, Inversion with early photons, *Med. Phys.* **34**, 1405 (2007).

- [41] M. J. Niedre, R. H. De Kleine, E. Aikawa, D. G. Kirsch, R. Weissleder, and V. Ntziachristos, Early photon tomography allows fluorescence detection of lung carcinomas and disease progression in mice *in vivo*, *Proc. Natl. Acad. Sci. USA* **105**, 19126 (2008).
- [42] B. Zhang, X. Cao, F. Liu, X. Liu, X. Wang, and J. Bai, Early-photon fluorescence tomography of a heterogeneous mouse model with the telegraph equation, *Appl. Opt.* **50**, 5397 (2011).
- [43] J. Pichette, J. B. Domínguez, and Y. Bérubé-Lauzière, Time-domain geometrical localization of point-like fluorescence inclusions in turbid media with early photon arrival times, *Appl. Opt.* **52**, 5985 (2013).
- [44] E. P. McShane, H. K. Chandrasekharan, A. Kufcsák, N. Finlayson, A. T. Erdogan, R. K. Henderson, K. Dhaliwal, R. R. Thomson, and M. G. Tanner, High resolution TCSPC imaging of diffuse light with a one-dimensional SPAD array scanning system, *Opt. Express* **30**, 27926 (2022).
- [45] J. Radford, A. Lyons, F. Tonolini, and D. Faccio, Role of late photons in diffuse optical imaging, *Opt. Express* **28**, 29486 (2020).
- [46] P. Latimer, Light scattering, data inversion, and information theory, *J. Colloid Interface Sci.* **39**, 497 (1972).
- [47] N. Byrnes and M. R. Foreman, Universal bounds for imaging in scattering media, *New J. Phys.* **22**, 083023 (2020).
- [48] M. S. Patterson, B. Chance, and B. C. Wilson, Time resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties, *Appl. Opt.* **28**, 2331 (1989).
- [49] C. E. Shannon, A Mathematical Theory of Communication, *Bell Syst. Tech. J.* **27**, 623 (1948).
- [50] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge, 2003).
- [51] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley & Sons, Inc., Hoboken, 2006).
- [52] G. Cohen, S. Afshar, J. Tapson, and A. Van Schaik, EMNIST: Extending MNIST to handwritten letters, in *Proceedings of the International Joint Conference on Neural Networks*, Vol. 2017-May (IEEE, Anchorage, 2017) pp. 2921.
- [53] S. L. Jacques, Optical properties of biological tissues: A review, *Phys. Med. Biol.* **58**, R37 (2013).
- [54] D. C. Sordillo, L. A. Sordillo, P. P. Sordillo, L. Shi, and R. R. Alfano, Short wavelength infrared optical windows for evaluation of benign and malignant tissues, *J. Biomed. Opt.* **22**, 045002 (2017).
- [55] A. Yaroshevsky, Z. Glasser, E. Granot, and S. Sternklar, Transition from the ballistic to the diffusive regime in a turbid medium, *Opt. Lett.* **36**, 1395 (2011).
- [56] British Standards Institution, *Safety of Laser Products—Part 14: A User's Guide*, Tech. Rep., London (2004).
- [57] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.5.L022008> for details of experimental procedure and results showing the number of photons transmitted for increasing TMFP lengths, a discussion of determining the rank of forward operators used to produce Fig. 3, and additional image reconstructions using basic regularization.
- [58] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, Cambridge, 2007), p. 795.
- [59] K. Manohar, B. W. Brunton, J. N. Kutz, and S. L. Brunton, Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns, *IEEE Control Systems* **38**, 63 (2018).