



Fisher information for the design of diffuse optical brain monitoring systems

ZHI-GUAN WANG,^{1,2}  JACK RADFORD,²  MING-JIE SUN,^{1,3,4} AND DANIELE FACCIO^{2,5} 

¹*School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China*

²*School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK*

³*Hangzhou International Innovation Institute, Beihang University, Hangzhou 311115, China*

⁴*mingjie.sun@buaa.edu.cn*

⁵*daniele.faccio@glasgow.ac.uk*

Abstract: Diffuse optical tomography is a non-invasive technique for assessing tissue structure and function by detecting scattered near-infrared light. A key question is how to select source–detector configurations and measurement modalities to maximize sensitivity, particularly for brain defect (e.g., tumor, stroke, etc.) monitoring. This study proposes Fisher information as a rapid, quantitative metric to optimize source-detector placement. We compare the performance of continuous-wave and time-resolved measurements and also explore the impact of cerebrospinal fluid on system design. This method confirms some known results but also highlights some counterintuitive results, including the finding that optimal source-detector distance depends on the type of detector and its related noise statistics and that the low-scattering cerebral spinal fluid layer in the brain can enhance measurement sensitivity. This is found to depend on how exactly the data is collected and processed, with fully time-resolved data always showing a net gain. Our results provide an efficient framework for analyzing and designing diffuse optical imaging systems for non-invasive monitoring and imaging via diffuse light.

Published by Optica Publishing Group under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

1. Introduction

Diffuse optical tomography (DOT) is a powerful non-invasive technology for probing the internal structure and functional dynamics of highly scattering media [1–3], such as biological tissue. By illuminating tissue with near-infrared light and measuring diffusely scattered photons at the surface, DOT enables the reconstruction of absorption and reduced scattering maps, which provide information about blood oxygenation, hemoglobin concentration, and other physiologically relevant parameters [4]. In medicine, DOT has been widely applied to monitor cerebral oxygenation and hemodynamics, particularly in neuroscience research and neonatal care. It has also been investigated for breast cancer detection by identifying abnormalities in tissue metabolism and vascularisation [5–8], as well as for functional brain imaging to study neural activity [9]. Recent advancements in high-density diffuse optical tomography (HD-DOT) have demonstrated that increasing the source-detector density significantly enhances imaging resolution and depth sensitivity [10,11]. However, regardless of the sensor count, determining the most effective spatial configuration remains critical, as the quality of data acquisition is intrinsically linked to the interrogation geometry. Consequently, a long-standing problem in DOT is where to place light sources and detectors to maximize measurement sensitivity. Although testing various source-detector (S-D) configurations and comparing the reconstructions with the ground truth is a viable approach, it can be quite time-consuming, especially for reconstructions that require solving large-scale, nonlinear inverse problems. Therefore, it is necessary to investigate configuration optimising methods that mainly exploit the forward model. For example, singular

value decomposition (SVD) can be applied to the reconstruction matrix, which can provide conceptually straightforward and quantitative metrics relating to the reconstruction quality that can be exploited as a design guide [12,13]. However, SVD is feasible when assuming a linear system model, which cannot be directly applied to nonlinear systems or iterative methods where the reconstruction matrix is continuously updated.

Aside from optimising the S-D placement, the choice of measurement type is also a crucial factor that affects the imaging quality. Three types of measuring systems are widely used in DOT: continuous-wave (CW), frequency-domain, or time-resolved (TR) systems. Time-resolved measurements yield the richest datasets, with various processing approaches such as time gating and temporal moment analysis [14–16]. The choice of imaging system is also typically related to the choice of sensor type, e.g., photodiodes for CW systems versus single photon avalanche diodes (SPADs) for TR measurements. The noise statistics will in turn depend on the chosen sensor, which will also influence the system performance and thus need to be considered in the optimisation. To compare the performance of different measurement types and data processing methods, we need a universally applicable metric that is independent of system type and noise model, which can then be used to guide system design.

A widely used tool for estimating and optimising measurement sensitivity is the Fisher information (FI). Fisher information quantifies the information that a random variable contains about an unknown parameter and has been used, e.g., to estimate the fundamental limits in localising fluorescent molecules [17] or for estimating the distance between incoherent light sources [18]. The application of FI spans various fields, including optical metrology [19], Lidar [20,21], and quantum systems [22,23]. Recent work has shown the importance of FI also for scattering media by demonstrating that scattering, rather surprisingly, does not degrade the average measurable information compared to the non-diffusive regime [24].

In this study, we combine FI and DOT for the optimisation of optical monitoring systems for the brain, and the FI is calculated by simulating photon propagation in a 2D circular scattering phantom as a simplified model of the brain and also in a 3D human atlas model. FI analysis provides a physically explainable quantity of the absolute uncertainty and different noise models can also be naturally introduced in the analysis. We identify the optimal S-D configurations under various conditions and compare the maximum achievable FI across different measurement modalities and data processing methods. Additionally, we investigate the influence of cerebrospinal fluid (CSF), a critical structure that affects photon propagation in the brain, on both the optimal S-D configuration and FI. Our results show that FI can be optimized by considering the details of each individual case, e.g., S-D separation depends on anomaly depth and noise statistics (Gaussian versus Poisson) of the detector and, cerebral spinal fluid has a non-trivial impact on sensitivity that manifests itself as an absolute increase for fully time-resolved measurements but has a more complex influence in other imaging modalities (CW or time-gated). Furthermore, the conclusion derived from the simplified brain model is further validated using a more realistic human head model, supporting the generalizability of our conclusions.

2. Method

2.1. Overview of Fisher information

Fisher information quantifies the amount of information that an observable random variable X provides about an unknown parameter θ . Specifically, under the context of DOT and brain monitoring of this study, θ denotes the absorption coefficient (μ_a) of the anomaly embedded in the brain, and X denotes the optical observable obtained from a measuring system. Let $f(x; \theta)$ represents the probability density function (PDF) for X conditioned on θ , then the $FI(\theta)$ is defined as:

$$FI(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f(X; \theta) \right)^2 \right]. \quad (1)$$

In real-world conditions, the randomness of X can arise from several mechanisms, including quantum indeterminacy, thermal fluctuations and instrumental noise. A distribution of X that is sharply peaked to changes in θ yields a higher value of $FI(\theta)$, indicating that a more precise estimation of θ can be obtained from X . Through the Cramér–Rao inequality, the FI is linked to a lower bound on the variance of any unbiased estimator $\hat{\theta}$. If $\hat{\theta}(X)$ is an unbiased estimator, i.e., $\mathbb{E}[\hat{\theta}(X)] = \theta$, then its variance satisfies:

$$\text{Var}(\hat{\theta}(X)) \geq \frac{1}{FI(\theta)}. \quad (2)$$

This inequality establishes that no unbiased estimator can have a variance smaller than the reciprocal of the FI. The computation of FI depends on the forward model from θ to X and the distribution of X , and both of which can be obtained analytically or numerically.

2.2. Fisher information for optical brain monitoring

In the near-infrared window (650–900 nm), light can penetrate several centimetres into brain tissue and scattered photons can contain information resulting from local changes in optical scattering and absorption properties, including blood oxygenation and also any anomalies, e.g., tumor or stroke-affected regions. For example, for a tumor, as the affected region grows, internal capillary density increases, resulting in a higher absorption coefficient [25–27]. Therefore, the absorption coefficient μ_a could serve as an indicator of the tumor progression, with changes in absorption reflected in variations in both the measured photon counts and their arrival-time distribution. A similar reasoning can be applied to other anomalies, such as stroke. In brain anomaly monitoring, the anomaly location and geometry can be determined using MRI or similar imaging methods, leaving the absorption coefficient μ_a as the only unknown parameter. Then, a question arises as to which S-D configuration and measurement modality (CW or TR) provides the most precise assessment of the anomalies via μ_a . To apply FI in this scenario, the unknown parameter θ is regarded as the anomaly absorption coefficient μ_a , and the observable random variable X represents either the total photon counts recorded by continuous-wave detectors or the counts in each time bin from time-resolved detectors.

Next, we need to determine the source of randomness in the measurement and build a distribution model for X . The potential sources of randomness in diffuse optics measurements are the noise originating from the source, from dynamic changes in the scattering media, and from the detectors. Generally speaking, fluctuations in the output power from the laser source due to the electronic circuit or temperature can be eliminated or assumed to be negligible to first approximation through careful control of the system and laser-power stabilisation techniques. *In vivo* biological tissue is a highly dynamic medium, where intrinsic physiological factors, such as cardiac pulsation and respiration, strongly modulate the detector signal. However, these signal components produced by these physiological factors typically exhibit pronounced periodic characteristics and can be regressed out by techniques like the General Linear Model (GLM), which is widely used in DOT and functional near infrared spectroscopy (fNIRS) [28,29]. Therefore, in this work, we assume the sufficient suppression of physiological perturbations, but with additional mathematical treatment (which is beyond the scope of this research), these components could still be included in the FI analysis [30]. From a microscopic perspective, the subcellular structures causing scattering in biological tissue, such as cells and organelles, are in constant random motion, which randomly modulates the measurement signal. Recent FI research focusing on such dynamic media is presented in [24,31–33]. Here, we focus on a situation where the scattering media are "stable" on the timescale of the measurements and the output light field can be predicted by μ_a and μ_s .

Based on the above analysis, in this work, we focus only on the randomness of X that results from the detector noise. Detector noise can be generally classified into shot noise and readout

noise. Shot noise follows a Poisson distribution, whereas readout noise is modelled as an additive Gaussian distribution. Most detectors, including CCD or CMOS cameras, photomultiplier tubes (PMTs), and photodiodes, are affected by both shot noise and readout noise. Some single-photon detectors, such as SPADs and superconducting nanowire single photon detectors (SNSPDs), operate in photon-counting mode, where a photon-induced electric pulse registers a photon-counting event. There is no analog signal amplifying process that introduces readout noise, and thus, the shot noise is dominant. Although under reasonable intensity, the high photon count will make the Poisson statistic similar to a Gaussian random variable, the analysis for Poisson noise remains relevant in photon-starved cases (e.g., in the case of a large source-detector separation). Consequently, depending on the detector, the observable random variable X is considered to follow a Gaussian or a Poisson distribution. Here, the random variable X refers to the detected photon count.

The distribution model for the observable random variable X is modelled as the photon counts N (total photon counts or the photon counts in a time bin) following a Poisson or Gaussian distribution with a mean value of $\bar{N}$ and standard deviation of σ . The FI of N with respect to the parameter $\theta = \mu_a$ is given by:

$$\begin{aligned}
 FI_N(\mu_a) &= \mathbb{E} \left[\left(\frac{\partial}{\partial \mu_a} \ln f(N; \mu_a) \right)^2 \right] \\
 &= \mathbb{E} \left[\left(\frac{\partial}{\partial \mu_a} \ln f(N; \bar{N}) \right)^2 \right] \\
 &= \mathbb{E} \left[\left(\frac{\partial}{\partial \bar{N}} \ln f(N; \bar{N}) \frac{\partial \bar{N}}{\partial \mu_a} \right)^2 \right] \\
 &= \mathbb{E} \left[\left(\frac{\partial \ln f(N; \bar{N})}{\partial \bar{N}} \right)^2 \right] \mathbb{E} \left[\left(\frac{\partial \bar{N}}{\partial \mu_a} \right)^2 \right] \\
 &= FI_N(\bar{N}) \left(\frac{\partial \bar{N}}{\partial \mu_a} \right)^2 = \frac{1}{\sigma^2} \left(\frac{\partial \bar{N}}{\partial \mu_a} \right)^2.
 \end{aligned} \tag{3}$$

Based on Eq. (1), we identify the target parameter θ as the absorption coefficient μ_a . By treating the MCX simulation as a forward operator $\mathbb{M}$, the expected measurement is expressed as a function of the optical properties, i.e., $\bar{N} = \mathbb{M}(\mu_a)$. Consequently, we reparameterize the log-likelihood function $\ln f(N; \mu_a)$ in terms of $\bar{N}$. The subsequent steps apply the chain rule, treating $\bar{N}$ as an intermediate variable dependent on μ_a . The last line uses a conclusion that for a Gaussian or Poisson variable, the FI of its mean value is the reciprocal of its variance. See [Supplement 1](#) for the derivation. The FI is inversely related to the variance of the underlying distribution. When N is a Poisson distribution, its variance σ^2 equals the mean value $\bar{N}$. The other part required for FI calculation is the derivative of $\bar{N}$ with regard to the μ_a of the anomaly region (the change in expected photon count, i.e., intensity, with respect to μ_a of the anomaly). In DOT, the spatial distribution of the derivative $\partial \bar{N} / \partial \mu_a$ is the Jacobian matrix. In our analysis, however, only the portion of the Jacobian overlapping with the anomaly region contributes to the calculation of FI. Both $\bar{N}$ and the Jacobian can be estimated by Monte Carlo simulations, e.g., using ‘Monte Carlo eXtreme’ (MCX) [34–36], a widely-used ray-tracing-based photon propagation simulator software optimized for scattering media. MCX (and in general the Monte Carlo method) is valid and accurate in the cerebral environment also when considering weak scattering media such as cerebrospinal fluid (CSF) where conversely, the diffusion equation is not strictly applicable [37,38].

In our Monte Carlo simulations, we obtain the simulated photon package weight of the detector, which is then normalized by the total launched package weight to yield an estimated reflectance. Next, using the Maximum Permissible Exposure (MPE) of the skin (3 mW/mm², from ANSI Z136.1 standard) and assuming 1 mm exposure area, we can use the reflectance to calculate

the detected photon count (for 800 nm light). The Jacobians are also obtained from each MCX simulation via the reply mode [35], providing $\partial \tilde{N} / \partial \mu_a$. The Gaussian noise is an additive noise with zero mean, and the variance is calculated by a typical Noise Equivalent Power (NEP) of a Photodiode ($1 \times 10^{-13} \text{ W/Hz}^{1/2}$). The mean and variance for the Poisson noise are equal to the simulated photon count ($\mu = \sigma^2 = \tilde{N}$). Finally, these values are substituted into Eq. (3) to compute the Fisher Information. This framework is fully compatible with experimental implementation. In a physical experiment, one can perform many measurements to numerically estimate the signal mean ($\tilde{N}$) and variance (σ^2). The sensitivity (Jacobian) can be determined via perturbation experiments or estimated from a reference model. These experimentally obtained values can then be directly substituted into Eq. (3) to calculate the Fisher information. We note that although the analysis is based on the simulated data, it is usable for experimental data if other sources of noise (such as the physiological components discussed before) are ideally suppressed and the data statistics can be modelled as Gaussian or Poisson variables. For cases where the experimental data contains remaining extrinsic perturbations, we speculate that the achievable Fisher information will decrease as additional perturbations worsen the estimation of the target parameter.

3. Fisher information of CW modality

3.1. Fisher information of CW measurements

The scope of this work is to provide a general framework for assessing the impact and optimising the positioning of S-Ds (as opposed to focusing on any specific detailed situation). Therefore, for simplicity and just an example, we model the head as a 2D uniform circular scattering medium of diameter 140 mm containing a 30 mm (matches the typical size of a brain tumor) diameter anomaly with higher absorption that we wish to detect or monitor. The optical parameters of the medium are set as the parameters of the grey matter $\mu_a = 0.033 \text{ mm}^{-1}$, $\mu_s = 7 \text{ mm}^{-1}$, $g = 0.9$, $n = 1.37$ provided in [39]. The absorption coefficient (μ_a) of the anomaly was set to 0.053 mm^{-1} , while other parameters remain the same as the background. This represents an approximate 50% contrast relative to the background, chosen to model the increased hemoglobin concentration associated with hyper-vascularized tissue (e.g., tumors). As an example and without loss of generality, seventeen detector positions are placed evenly on the boundary over the angular range $[-\pi/3, \pi/3]$, and nine source positions are uniformly distributed over $[-\pi/3, 0]$, as illustrated in Fig. 1(a). The sources are modelled as pencil beams with normal incidence, and the detectors are circular collection apertures. In this example, each light source is paired with a detector, yielding a total of 153 distinct S-D position combinations. The depth (distance from the circular phantom's external boundary) of the anomaly ranges from 14 mm to 50 mm with a step of 5 mm. Our objective is to determine the specific S-D configuration that yields the maximum Fisher Information for an anomaly at a given depth. To achieve this, for each specific anomaly depth, we conducted MCX simulations for all 153 possible S-D configurations. We then calculated the Fisher Information for every permutation to identify the optimal arrangement. The choice of a 2D phantom is also because of the wish to speed up the simulation, as we need to run extensive simulations. It takes 32 hours to run all 1224 simulations (8 anomaly depths $\times$ 153 S-D pairs) for the 2D phantom in a Nvidia GeForce RTX 4060 laptop GPU. The number of photon packets was dynamically scaled between 1×10^7 and 1×10^{11} depending on the source-detector separation. This strategy was employed to ensure that a sufficient number of photons were collected even at large distances, thereby effectively suppressing the intrinsic stochastic noise of the Monte Carlo method while maintaining computational efficiency.

Figure 1(b) illustrates how the FI varies with the S-D distance when the anomaly depth is 35 mm, based on the Poisson noise model (we will compare below to the case of Gaussian noise). As the distance between the source and the detector increases, the FI first rises and then falls, indicating that there is an optimal S-D distance that provides the maximum value of FI. It is well

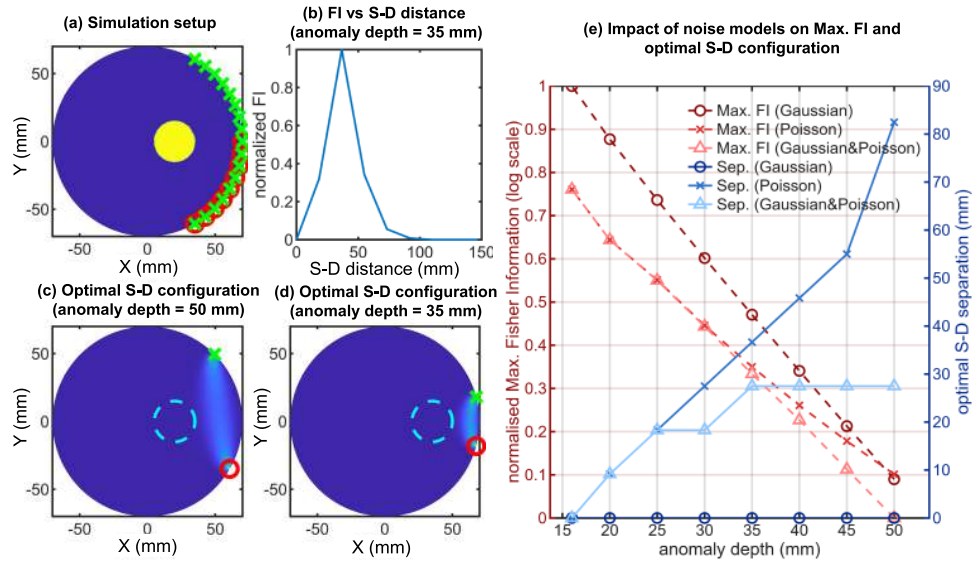


Fig. 1. Fisher information using CW illumination. (a) The 2D circular phantom with all possible positions of sources and detectors. The blue circle represents the background tissue, and the yellow circle is the anomaly. There is no skull-like structure around this phantom. Red circles represent sources, and green crosses indicate detectors. (b) An example of how FI changes with S-D distance at a fixed anomaly depth of 35 mm under Poisson noise. The corresponding Jacobian (the "banana" connecting the source and the detector) is also presented. (c) The optimal S-D configuration at an anomaly depth of 50 mm and detector Poisson noise. The dashed light-blue circle in this and the next panel indicates the shape of the anomaly. (d) The optimal S-D configuration at an anomaly depth of 35 mm and detector Poisson noise. (e) Comparison of maximum Fisher information results under different noise models (Gaussian, Poisson and mixed, as indicated in the legend). The left axis (red dashed curves) displays the maximum FI (normalized, log scale) achievable at each anomaly depth. The right axis (blue curves) indicates the corresponding optimal S-D separation, defined as the distance of the S-D pair yielding the highest FI.

understood in the field of fNIRS that there is a trade-off between sensitivity and signal-to-noise ratio (SNR) [40,41]. When the S-D distance is small, a higher SNR is expected. However, most of the detected photons only travel through the superficial region, resulting in limited sensitivity for deep targets [42]. It is widely agreed that a 30-35 mm S-D separation is required for adult brain activity monitoring due to the necessity to reach ~ 20 mm depths beyond the scalp to probe the grey matter [43]. There are also studies focusing on finding the optimal S-D distance that maximizes the sensitivity for the brain [44,45], but the impact of noise is overlooked.

The optimal S-D configurations (considering both sensitivity and noise in the FI approach) for anomaly depths of 50 mm and 35 mm are shown in Figs. 1(c) and 1(d). A more detailed mapping showing the highest FI and corresponding S-D distance at different anomaly depths is provided in Fig. 1(e). The results show that, under the Poisson noise model, if the anomaly is very close to the surface, a null S-D distance is preferred, and as the anomaly becomes deeper, the optimal S-D separation increases almost linearly with the anomaly depth. This result arises from both how photons propagate in scattering media and the noise model (we will demonstrate that the Gaussian noise model yields an entirely different S-D separation choice in the next subsection). As the S-D separation increases, the detected photon count decreases exponentially, but most of the detected photons will penetrate a deeper region (see the Jacobians given in Fig. 1). Although the Jacobian appears to extend deeper spatially, its absolute value (the change in detected photon

count per unit change in μ_a) still decreases. The Poisson noise variance, which is equal to the mean photon count, will also decrease with increasing S-D distance. Therefore, according to Eq. (3) (the FI is inversely proportional to the noise variance and scales with the square of the Jacobian, i.e., $FI \propto J^2/\sigma^2$), when the anomaly becomes deeper, increasing the S-D separation decreases the noise level and compensates for the loss of FI due to the reduced Jacobian, making this the optimal strategy.

3.2. Effects of the noise model on Fisher information

We also compared the relative effects of Gaussian noise and Poisson noise on the FI. Gaussian noise, which originates from the readout noise of the detector, is primarily correlated with the load resistance of the detector readout circuit and the temperature and is independent of the photon count. Thus, the resulting noise is a Gaussian distribution with a fixed variance across all the possible S-D configurations. For comparison, the Poisson noise model describes shot noise, whose variance equals its mean value, i.e., the mean photon count. Consequently, for different S-D placements, the Poisson noise variance varies while the Gaussian noise variance remains constant.

The effects of different noise models on the maximum FI and the optimal S-D placements are shown in Fig. 1(e). The results show that under the Gaussian noise model, zero S-D distance is always the best choice, while as discussed above, under the Poisson noise model, the optimal S-D separation increases with the anomaly depth. Combining the two types of noise results in an optimal S-D distance between these two cases. A key point here is that the FI for measurements affected by Gaussian noise indicates that no separation between source and detector is optimal for sensing at all depths. According to Eq. (3), when the noise variance is fixed, such as for Gaussian noise, the optimal strategy is to select the S-D configuration that maximizes the Jacobian in the anomaly region. Although increasing the S-D separation increases the fraction of detected photons that reach deep tissue, the Jacobian of the anomaly region is always the largest at zero S-D separation, regardless of the anomaly depth. The conclusion for Poisson noise is different as the noise variance increases with the mean expected photon counts $\bar{N}$. Although the null S-D distance yields the highest Jacobian, it also brings the highest Poisson noise. Therefore, the optimal distance is determined by achieving an appropriate balance between the Poisson noise level and maximizing the Jacobian at the anomaly position.

This example illustrates how different noise models influence FI. Gaussian and Poisson distributions are both fundamental noise models with well-defined analytical expressions. This framework can also be extended to other noise models with more complex distributions. In such cases, the calculation for FI may not be a closed-form expression but should instead be obtained through numerical methods.

4. Fisher information of TR modality

4.1. Fisher information of TR measurements

As opposed to the CW regime, time-domain measurements record the distribution of times for photons within a laser pulse to migrate through tissue from the source to the detector. The resulting histogram is known as the temporal point spread function (TPSF). A common approach in TR DOT is to calculate the moments and derive quantities of interest from these. The moments inherently incorporate contributions from all recorded photons, serving as compact yet comprehensive summaries of the TPSF shape. The moments of a TPSF $g(t)$ are defined by [46]:

$$M_0 = \int_{-\infty}^{\infty} g(t) dt, \quad (4)$$

$$M_k = \langle t^k \rangle = \frac{\int_{-\infty}^{\infty} t^k g(t) dt}{M_0}, \quad k = 1, 2, 3, \dots \quad (5)$$

The variance V , i.e., the second central moment follows $V = \frac{1}{M_0} \langle (t - M_1)^2 \rangle = M_2 - M_1^2$. In fNIRS, the most commonly used moments are M_0 , M_1 , and V , which correspond to the total photon count, the mean arrival time, and the temporal variance, respectively. Time-gating of the full TPSF is often used and offers straightforward depth discrimination by isolating early versus late photons, but suffers from poor SNR in late windows. In contrast, the moments leverage the entire TPSF for improved SNR.

As discussed in Section 2.2, Poisson noise is the dominant noise present in TR detectors, and the photon count in each time bin is an independent Poisson random variable. The FI of independent random variables is additive when considering the same parameter. Therefore, the FI of the full TPSF equals the sum of the FI of each time bin. For a time gate, the photon counts from the selected time bins are summed to create a new Poisson variable whose FI is then computed. The FI of M_0 is equivalent to that of the CW case, while FI for higher-order moments can be derived using the chain rule of differentiation:

$$\begin{aligned} FI_M(\mu_a) &= \mathbb{E} \left[\left(\frac{\partial}{\partial \mu_a} \ln f(M; \bar{M}) \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_i \frac{\partial \bar{N}_i}{\partial \mu_a} \frac{\partial \bar{M}}{\partial \bar{N}_i} \frac{\partial}{\partial \bar{M}} \ln f(M; \bar{M}) \right)^2 \right], \end{aligned} \quad (6)$$

where M can be M_1 , M_2 and V . $\bar{N}_i$ represents the expected photon count in the i -th time bin. Note that we apply a similar reparameterization to the log-likelihood function here, as that described in Eq. (3). The computation of FI for moments can be divided into three parts: $\partial \bar{N}_i / \partial \mu_a$ is the time-resolved Jacobian; $\partial \bar{M} / \partial \bar{N}_i$ describes how the moments will change with photon counts in a time bin, which depends on the shape of the TPSF and the type of the moment. The remaining part is the FI of M about its mean value $\bar{M}$. Higher-order moments are nonlinear functions of multiple independent Poisson random variables. As a result, their PDFs have no simple closed form and must be characterized numerically, for example, by a Gaussian approximation of their mean and variance for FI estimation.

The simulation setup for TR measurements is identical to that used in the CW case above. The TR detector in the simulation records photons over a 3-ns time range with a 50-ps temporal resolution, which is a typical value for most commercial time-resolved single-photon counting (TCSPC) detectors. We calculate and compare the FI of different TR data processing methods including M_0 (equivalent to a CW measurement), M_1 , M_2 , V , the full TPSF, the early time gate (ETG), and the late time gate (LTG). As different S-D configurations result in varying TPSFs, we employ a fixed time gating strategy that is independent of the TPSF shape. We calculate the ETG by accumulating the photon counts arriving in the first half of the time bins (i.e., from 0 ns to 1.5 ns), whereas the LTG selects photons arriving in the later half. For each metric and anomaly depth, we choose the highest FI value among all possible S-D configurations, and the results are shown in Fig. 2.

The full TPSF yields the highest FI across all different anomaly depths. For superficial anomalies, M_0 and M_1 provide high FI, but their performance is surpassed by M_2 and V as the anomaly depth increases. The reason is that $\partial \bar{M}_1 / \partial \bar{N}_i$ increases linearly with i , while $\partial \bar{M}_2 / \partial \bar{N}_i$ and $\partial \bar{V} / \partial \bar{N}_i$ increases quadratically with i , indicating that M_2 and V are more sensitive to late photons that contain more information about deep anomalies. Although V and M_2 are both second-order moments, V excludes mean-induced variability from M_1 , which reduces its variance and yields slightly higher FI than M_2 . ETG has the same information as M_0 for shallow anomalies

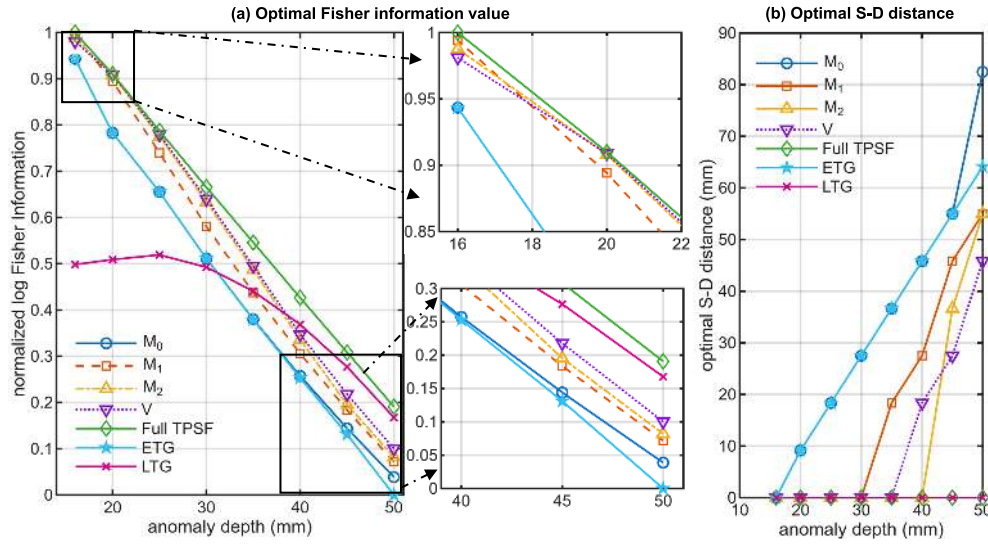


Fig. 2. Fisher information for different TR data processing methods. (a) The highest achievable FI at each anomaly depth. The full TPSF always gives the highest FI across all anomaly depths. M_1 is better than M_2 and V for shallow anomaly depths, but a contrary result is observed in larger depths. The LTG is the worst choice in shallow regions, but its FI shows the least attenuation with depth. (b) The S-D distance where the highest FI is obtained. An interesting observation is that the zero S-D separation is always the optimal choice for both full TPSF and LTG.

but significantly decays at greater depths. Conversely, although LTG is initially less informative than M_0 and ETG for shallow anomalies, it surpasses them at larger depths. Notably, for deep anomalies, M_0 , which includes additional early photons compared to LTG, yields lower FI, indicating that uninformative photons merely introduce noise and decrease the information.

Figure 2(b) shows the optimal S-D distance that maximizes FI at different anomaly depths. M_0 and ETG are obtained by summing the photon counts of all or the first half of the time bins, which is similar to a CW measurement, and the optimal choice is to increase S-D distance with the anomaly depth as explained in section 3.1. Since M_1 , M_2 , and V are intrinsically more sensitive to deeper anomalies, their optimal S-D separations remain zero for shallow anomalies. However, once the anomaly depth exceeds a specific range, it is still necessary to increase the S-D distance to extend the Jacobian deeper into the medium and suppress Poisson noise. An interesting finding is that the optimal S-D distance for the full TPSF case and also for LTG is always zero across all anomaly depths. This can be understood by considering that for shallow anomalies, one might expect a non-zero optimal S-D distance for deep anomalies, similar to the behavior observed in CW measurements under the Poisson noise model. For deep anomalies, those photons reaching the anomaly region arrive in late time bins and provide high FI. Similarly, since the LTG excludes the influence of early photons, zero S-D separation is also the optimal choice.

4.2. Correlation of TR moments

An additional question of interest is why it is conventional (in the fNIRS measurements) to use M_0 , M_1 , and V rather than M_0 , M_1 , and M_2 . Moreover, there is a more compact formula to retrieve μ_a and μ_s using M_0 , M_1 , and V [46].

Here, we investigate whether using M_2 or V affects the FI by comparing the joint FI $FI_{M_0, M_1, M_2}(\mu_a)$ and $FI_{M_0, M_1, V}(\mu_a)$. The joint FI refers to the total FI regarding the parameter μ_a

contained within multiple random variables, which also accounts for their correlations. The detailed steps to calculate the joint FI for moments can be found in [Supplement 1](#). The moments approximately follow a Gaussian distribution, and their joint FI can be calculated in the standard form of Gaussian variables $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}(\theta), \boldsymbol{\Sigma}(\theta))$ by:

$$FI_{\mathbf{X}}(\theta) = \left(\frac{\partial \boldsymbol{\mu}}{\partial \theta} \right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \boldsymbol{\mu}}{\partial \theta} \right), \quad (7)$$

where $\boldsymbol{\mu}$ is the mean value vector and $\boldsymbol{\Sigma}$ represents the covariance matrix. After determining the TPSF and Jacobian using MCX, we can calculate the mean values, variance, and correlation coefficient ρ to calculate the joint FI. The joint information is a function of ρ , and if two random variables are uncorrelated ($\rho = 0$), the joint FI is simply the sum of their individual FI. An example of the correlation coefficient between moments with a 35 mm anomaly depth is shown in Fig. 3. We found that M_0 is uncorrelated with the other moments, while M_1 and M_2 are highly correlated. Additionally, M_1 and V show a moderate correlation. The joint FI (calculated from Eq. (7)) indicates that any two-parameter combination of M_1 , M_2 , and V yields the same FI. This conclusion holds true across all other anomaly depths and S-D separations. From an intuitive explanation viewpoint, given $V = M_2 - M_1^2$, any two quantities uniquely determine the third, which explains why any pair yields identical FI. This is an intrinsic property of temporal moments and is independent of the setup. Furthermore, as M_0 is independent of the other moments, using M_0 , M_1 , and M_2 always yields the same amount of FI as M_0 , M_1 , and V .

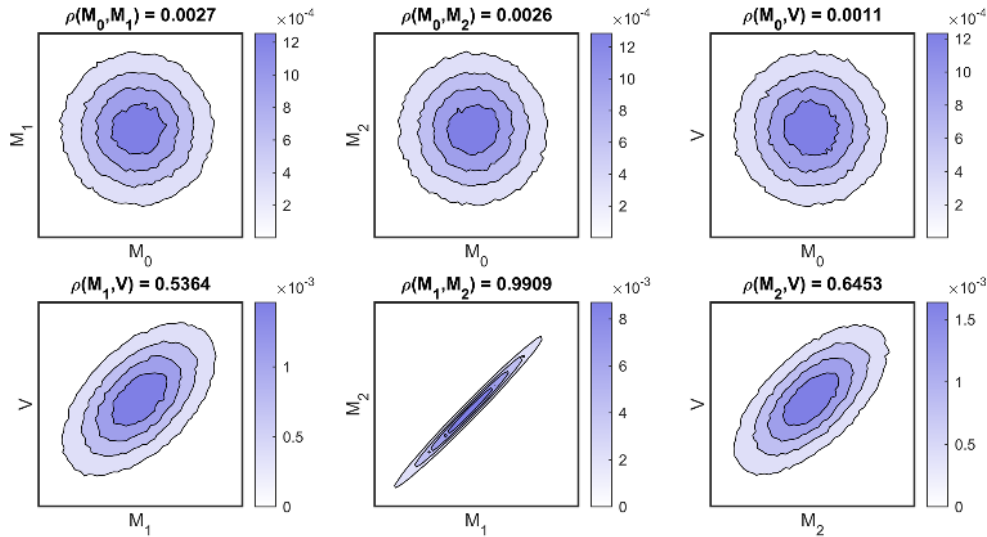


Fig. 3. Correlation between TR moments. The correlation between the TR moments when the photon counts in each time bin are independent Poisson variables. M_0 is uncorrelated with the other moments. M_1 and M_2 are highly correlated, while M_1 - V and M_2 - V exhibit moderate correlation. Correlation values, ρ are indicated in the title of each figure.

5. Effect of CSF on Fisher information

Cerebrospinal fluid (CSF) might be expected to have an impact on measurements as its scattering coefficient is significantly lower than that of surrounding tissues. A recent work [47] demonstrated that such structures can function as a scattering waveguide, significantly altering the propagation of photons over very long distances. Differing impacts of the CSF have been reported: the CSF

layer can confine photons to superficial regions, thereby reducing sensitivity to deep targets [48], while other studies report that it increases photon penetration into deeper tissue, thereby enhancing detection of deep targets [49,50]. Here, we compare the FI with and without a CSF-like layer, thereby providing a method for the quantitative assessment of its impact on the sensitivity of deep-region detection and on the choice of optimal S-D configuration.

In the simulation, the CSF-equivalent layer is 3 mm thick and is 12 mm below the surface, illustrated by a light blue ring in Fig. 4(a). The optical parameter of the CSF $\mu_a = 0.042 \text{ mm}^{-1}$, $\mu'_s = 0.01 \text{ mm}^{-1}$, $n = 1.37$ is obtained from [39], and all other parameters were kept consistent with those used in the CW and TR simulations. Figure 4(b) presents the logarithmic ratio of the optimal FI values with and without CSF across different anomaly depths. When the anomaly is located in the superficial region (15–25 mm), the CSF channel is totally blocked by the anomaly. Thus, the FI for most data processing methods remains unchanged. The reduction in the FI of LTG can be attributed to the loss of late anomaly-interacting photons that escape through the CSF. Within the anomaly depth range of 30–40 mm, the open CSF channel leads to an increase in FI for all methods, except for LTG. This implies that, overall, the presence of CSF facilitates detection within this depth range. The reduction in FI of LTG results from fewer scattering events each photon experiences in CSF, which shortens their average arrival time and causes many photons containing information to fall within the ETG window. As the anomaly depth further increases, the FI obtained by most methods remains higher than that without CSF. However, for deep anomalies, the FI of M_0 and ETG is considerably lower than when there is no CSF. This suggests that the presence of CSF generates more detected photons, but these photons arrive early and do not provide information about the target, thus acting as noise and decreasing the FI of M_0 and ETG. In summary, from the perspective of FI, CSF is generally beneficial for monitoring anomalies, but this benefit depends on the appropriate selection of data processing strategies. For example, if one uses a CW system (which is equivalent to M_0), the CSF acts as a hinge for monitoring deep anomalies, but the performance of using the full TPSF is either unaffected or even improved by CSF, regardless of the depth of the anomaly.

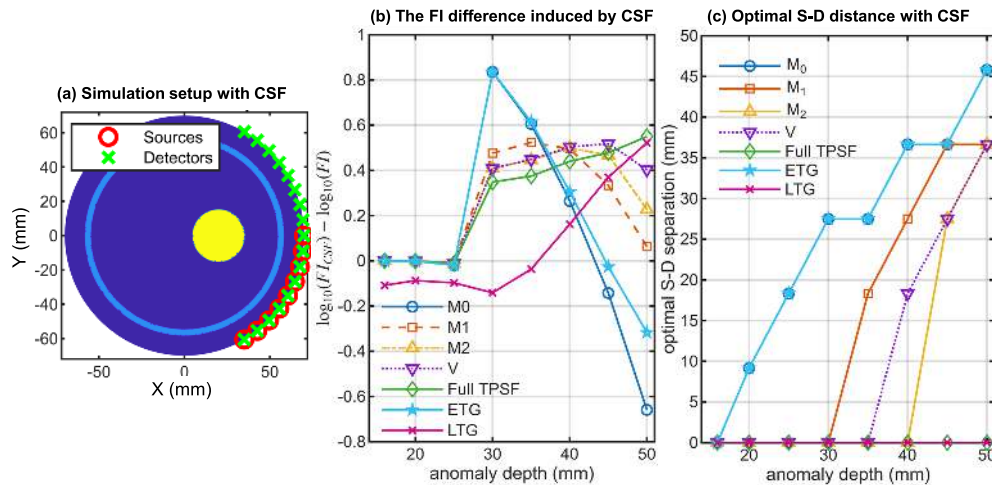


Fig. 4. The effect of CSF on the Fisher information. (a) The 2D circular phantom that contains CSF. The yellow circle and the light blue ring represent the anomaly and the CSF, respectively. The external scatter ring surrounding the CSF represents a skull-like structure. (b) The logarithmic ratio of the highest achievable FI values with and without the CSF. The presence of CSF generally increases the achievable FI, except for using ETG and M_0 for a deep anomaly. (c) The S-D distance where the highest FI is obtained with CSF. Compared with the case of no CSF, the optimal S-D separations are generally shortened.

The introduction of CSF also modifies the optimal S-D configuration, as shown in Fig. 4(b). Compared to the scenario without CSF, the optimal S-D separations for most methods decrease notably. However, the optimal S-D distance remains unchanged when using the full TPSF or LTG. Generally, increasing the S-D separation enhances the probability that detected photons reach deeper regions. The presence of CSF, however, allows deep regions to be probed without requiring a large S-D separation.

6. Fisher information analysis on a 3D human head model

Although the model used above is a simplified 2D circular model, we believe it still captures the essential features relevant to the conclusions presented in this paper. The spherical model itself was chosen as it enables significantly faster simulations, whereas full-scale MRI-based head mesh models, whilst absolutely feasible for a limited number of simulations despite the stronger attenuation and the consequently much larger number of rays required, the work presented here required scanning over a broad range of parameter settings and would have been computationally prohibitive. However, to demonstrate that the overarching conclusions do extend to the full brain model, we applied the FI analysis to a 3D human head atlas based on MRI data, which includes the main structures, i.e. the scalp, skull, CSF, gray matter and white matter. As exhaustive scanning of S-D positions in the 3D model is computationally prohibitive, we tested only a limited set of short S-D separations to determine whether zero separation remains optimal. We specifically chose this claim (i.e., that zero-separation S-D is optimal for time-domain sensing) because it is a known result from the literature [51–53], and therefore, one must expect our FI approach proposed here, as a minimum, to confirm this.

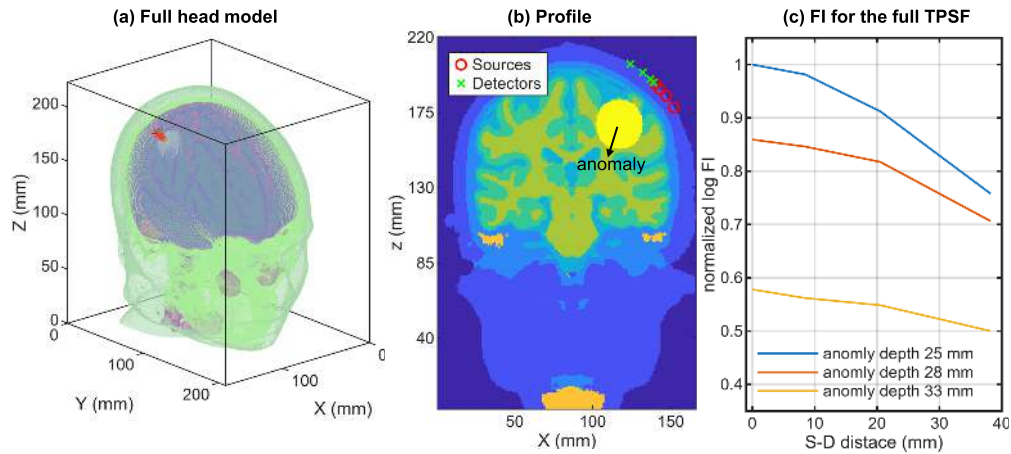


Fig. 5. Fisher information analysis on a 3D head model. (a) Schematic of the three-dimensional human head model. The 3D head model contains the scalp, skull, CSF, gray matter and white matter. (b) Cross-sectional view of the human head model showing the positions of the sources, detectors, and the anomaly. (c) Normalized log FI computed from the full TPSF of the TR system. The results indicate that zero S-D separation is always optimal, in agreement with the conclusions obtained from the simplified spherical model.

Figure 5 presents a schematic of the 3D head model, a cross-section profile (indicating the S-D positions and anomaly position), and the Fisher information results. We tested four S-D separations and three anomaly depths and computed the Fisher information using the full TPSF of the time-resolved system, as shown in Fig. 5(b). The results confirm that for time-domain measurements, zero S-D separation is indeed the optimal arrangement in the anatomically realistic

3D head model. This finding supports the generality of the conclusions obtained from the simplified head model and demonstrates the applicability of our FI analysis to complex scenarios.

7. Discussion and conclusion

The information theory or the use of FI and inverse biased retrievals have two completely different scopes and applications. Inverse retrieval, as the name says, is all about getting the best possible reconstruction of a ground truth for a given data set. FI optimisation is about trying to decide which measurement strategy optimizes the amount of information that can reach the detectors and be actually measured. This study focuses more on the FI metrics and is agnostic about the final image reconstruction method.

In general, Fisher information quantifies the ability to resolve changes in parameters from noisy measurements. As defined in Eq. (3), maximising FI requires optimising the ratio between the squared Jacobian and the noise variance ($FI \propto J^2/\sigma^2$). Consequently, the selection of the S-D configuration involves balancing the enhancement of the Jacobian of the target against the concomitant changes in signal-dependent noise levels. This principle can guide both the choice of S-D configuration and the data processing method. The choice of optimal S-D separation depends on whether the noise is signal-dependent. When the noise scales with the signal, increasing the S-D distance can reduce the noise, as in the case of Poisson-distributed noise. Compared with a CW system, a TR system is generally superior, as the additional temporal resolution provides a flexible approach to filter photons that carry relevant information, e.g., time-gating methods can intuitively select early or late photons that are sensitive to superficial or deep regions. Moments are statistical parameters defined over all the detected photons and cannot exclude specific subsets. Instead, their sensitivity to photons arriving at different times is tuned by moment order (e.g., M_2 and V are more sensitive to late-arriving photons than M_1 [54]). Using the full TPSF appears to be the optimal choice, but it requires higher data storage and processing capacity. A key finding is also that from the FI perspective, the optimal S-D position is always zero when measuring and analysing the full time-resolved histogram, regardless of anomaly depth and presence of cerebral spinal fluid layers. This conclusion is also verified in a 3D human head model with realistic anomaly structure. Consistent with our findings, previous studies have also reported the benefit of null S-D separation, demonstrating its ability to improve reconstruction contrast [51–53]. Although using a 2D or 3D model would affect the quantitative results of where to place the source and the detector, as that in a 3D model the photons will propagate in an additional dimension and their penetration depths could be different from that in a 2D model, the results of the general tendency of how FI changes with different system configuration could still provide insights for the DOT and fNIRS communities. Moreover, the proposed framework is extensible to other models with different optical coefficients or structures, enabling FI analysis in other scenarios.

In FI-based optimisation, two common criteria are A-optimality and D-optimality [55]. D-optimality aims to maximize the determinant of the FI matrix, which corresponds to maximising the overall information content. A-optimality, on the other hand, minimizes the trace of the inverse FI matrix and focuses on reducing the average estimation variance of the parameters. In our study, only a single parameter is considered, so the two criteria become equivalent. From the perspective of information theory, we naturally choose to maximize the FI, which follows the D-optimality. In future work, when multiple parameters (μ_s , the radius of the anomaly, etc.) are analysed using the proposed framework, the choice between A- and D-optimality becomes meaningful, as the two criteria have different characteristics, thereby leading to different optimisation results [56,57]. Besides, optimising the sensitivity of brain activity in the gray matter attracts many attentions [58], which is a typical scenario of multiparameter FI analysis. In such cases, the current framework is extensible by applying the D- and A-optimality criteria to the resulting FI Matrix.

CSF influences brain anomaly monitoring in multiple ways. Its low scattering coefficient increases the absolute number of photons reaching deep tissues and shortens their mean arrival time. However, the scattering waveguide effect also increases the total detected photon count and, consequently, the noise level. Thus, the impact of CSF on FI depends on the location of the anomaly and the data processing method. These findings underscore the need to account for CSF when designing diffuse optical monitoring systems for the brain.

Looking forward, from the perspective of FI, using (M_0, M_1, M_2) or (M_0, M_1, V) yields identical estimation accuracy for the target parameter. Since these quantities do not capture all the information contained in the full TPSF, future work could investigate whether introducing higher-order moments can yield additional gains in FI. Beyond temporal moments, alternative methods can extract information from the full TPSF's features, such as gamma-distribution fitting or the z -transform, and FI provides a theoretical framework to quantitatively compare which method captures the maximum useful information with the least data. Currently, the most commonly used cumulants in fNIRS and DOT are statistical moments of the counting distribution. However, the factorial cumulants [59], which fit the Poissonian nature of photon counting better, are less investigated in this field and could be an interesting point for future research. Also, the Shannon entropy could provide another perspective to determine the optimal way to perform measurement [60].

Finally, we note that the FI framework presented here is entirely model-based. Recently, deep learning techniques have emerged as powerful tools for diffusive imaging, including addressing inaccuracies in background parameter estimation [61], thereby enhancing the robustness of DOT across diverse subject populations. Furthermore, driven by the accumulation of system design data and experimental datasets within the DOT and fNIRS communities, machine learning strategies become a promising way for optimal experimental design [62]. While our physics-based Fisher Information approach provides theoretical bounds, future developments could benefit from a synergistic approach, combining analytical guidance with data-driven reconstruction methods to further guide system design and enhance real-world clinical applications.

Funding. Engineering and Physical Sciences Research Council (EP/T00097X/1, EP/Y029097/1); Royal Academy of Engineering Chairs in Emerging Technologies Scheme.

Acknowledgments. Z.-G. W. acknowledges the financial support from International Joint Doctoral Education Fund of Beihang University.

Disclosures. The authors declare that they have no conflicts of interest.

Data availability. Data for this work is available at [63].

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. 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).
2. M. B. Applegate, R. E. Istfan, S. Spink, *et al.*, "Recent advances in high speed diffuse optical imaging in biomedicine," *APL Photonics* **5**(4), 040802 (2020).
3. J. Bertolotti and O. Katz, "Imaging in complex media," *Nat. Phys.* **18**(9), 1008–1017 (2022).
4. S. G. Diamond, T. J. Huppert, V. Kolehmainen, *et al.*, "Dynamic physiological modeling for functional diffuse optical tomography," *NeuroImage* **30**(1), 88–101 (2006).
5. B. J. Tromberg, B. W. Pogue, K. D. Paulsen, *et al.*, "Assessing the future of diffuse optical imaging technologies for breast cancer management," *Med. Phys.* **35**(6Part1), 2443–2451 (2008).
6. M. L. Altøe, A. Marone, H. K. Kim, *et al.*, "Diffuse optical tomography of the breast: a potential modifiable biomarker of breast cancer risk with neoadjuvant chemotherapy," *Biomed. Opt. Express* **10**(8), 4305–4315 (2019).
7. K. M. S. Uddin, M. Zhang, M. Anastasio, *et al.*, "Optimal breast cancer diagnostic strategy using combined ultrasound and diffuse optical tomography," *Biomed. Opt. Express* **11**(5), 2722–2737 (2020).
8. M. Zhang, M. Xue, S. Li, *et al.*, "Fusion deep learning approach combining diffuse optical tomography and ultrasound for improving breast cancer classification," *Biomed. Opt. Express* **14**(4), 1636–1646 (2023).
9. A. T. Eggebrecht, S. L. Ferradal, A. Robichaux-Viehoever, *et al.*, "Mapping distributed brain function and networks with diffuse optical tomography," *Nat. Photonics* **8**(6), 448–454 (2014).

10. W. Fan, J. W. Trobaugh, C. Zhang, *et al.*, "Fundamental effects of array density and modulation frequency on image quality of diffuse optical tomography," *Med. Phys.* **52**(2), 1045–1057 (2025).
11. Z. E. Markow, J. W. Trobaugh, E. J. Richter, *et al.*, "Ultra high density imaging arrays in diffuse optical tomography for human brain mapping improve image quality and decoding performance," *Sci. Rep.* **15**(1), 3175 (2025).
12. F. Leblond, K. M. Tichauer, and B. W. Pogue, "Singular value decomposition metrics show limitations of detector design in diffuse fluorescence tomography," *Biomed. Opt. Express* **1**(5), 1514–1531 (2010).
13. A. Badon, V. Barolle, K. Irsch, *et al.*, "Distortion matrix concept for deep optical imaging in scattering media," *Sci. Adv.* **6**(30), eaay7170 (2020).
14. A. Pifferi, D. Contini, A. D. Mora, *et al.*, "New frontiers in time-domain diffuse optics, a review," *J. Biomed. Opt.* **21**(9), 091310 (2016).
15. H. Wabnitz, D. Contini, L. Spinelli, *et al.*, "Depth-selective data analysis for time-domain fNIRS: Moments vs. time windows," *Biomed. Opt. Express* **11**(8), 4224–4243 (2020).
16. A. Ortega-Martinez, D. Rogers, J. Anderson, *et al.*, "How much do time-domain functional near-infrared spectroscopy (fNIRS) moments improve estimation of brain activity over traditional fNIRS?" *Neurophotonics* **10**(01), 013504 (2022).
17. K. I. Mortensen, L. S. Churchman, J. A. Spudich, *et al.*, "Optimized localization analysis for single-molecule tracking and super-resolution microscopy," *Nat. Methods* **7**(5), 377–381 (2010).
18. M. Tsang, R. Nair, and X.-M. Lu, "Quantum theory of superresolution for two incoherent optical point sources," *Phys. Rev. X* **6**, 031033 (2016).
19. V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nat. Photonics* **5**(4), 222–229 (2011).
20. H. Li, K. Zheng, R. Ge, *et al.*, "Noise-tolerant LiDAR approaching the standard quantum-limited precision," *Light: Sci. Appl.* **14**(1), 138 (2025).
21. C.-H. Lee, Y. Kim, D.-G. Im, *et al.*, "Coherent Two-Photon LIDAR with Incoherent Light," *Phys. Rev. Lett.* **131**(22), 223602 (2023).
22. M. D. Lang and C. M. Caves, "Optimal Quantum-Enhanced Interferometry Using a Laser Power Source," *Phys. Rev. Lett.* **111**(17), 173601 (2013).
23. S. Gammelmark and K. Mølmer, "Fisher Information and the Quantum Cramér-Rao Sensitivity Limit of Continuous Measurements," *Phys. Rev. Lett.* **112**(17), 170401 (2014).
24. M. Horodyski, D. Bouchet, M. Kühmayer, *et al.*, "Invariance Property of the Fisher Information in Scattering Media," *Phys. Rev. Lett.* **127**(23), 233201 (2021).
25. R. Choe, M. E. Putt, P. M. Carlile, *et al.*, "Optically Measured Microvascular Blood Flow Contrast of Malignant Breast Tumors," *PLoS One* **9**(6), e99683 (2014).
26. Y. Shang, T. Li, and G. Yu, "Clinical applications of near-infrared diffuse correlation spectroscopy and tomography for tissue blood flow monitoring and imaging," *Physiol. Meas.* **38**(4), R1–R26 (2017).
27. D. Li, C. Qu, Q. Liu, *et al.*, "Monitoring the Real-Time Circulatory System-Related Physiological and Pathological Processes In Vivo Using a Multifunctional NIR-II Probe," *Adv. Funct. Mater.* **30**(6), 1906343 (2020).
28. H.-D. Nguyen, S.-H. Yoo, M. R. Bhutta, *et al.*, "Adaptive filtering of physiological noises in fNIRS data," *BioMed. Eng. OnLine* **17**(1), 180 (2018).
29. G. Bauernfeind, S. C. Wriessnegger, I. Daly, *et al.*, "Separating heart and brain: on the reduction of physiological noise from multichannel functional near-infrared spectroscopy (fNIRS) signals," *J. Neural Eng.* **11**(5), 056010 (2014).
30. F. Duan, F. Chapeau-Blondeau, and D. Abbott, "Exploring weak-periodic-signal stochastic resonance in locally optimal processors with a Fisher information metric," *Signal Processing* **92**(12), 3049–3055 (2012).
31. N. Byrnes and M. R. Foreman, "Universal bounds for imaging in scattering media*," *New J. Phys.* **22**(8), 083023 (2020).
32. J. Hüpfel, F. Russo, L. M. Rachbauer, *et al.*, "Continuity equation for the flow of Fisher information in wave scattering," *Nat. Phys.* **20**(8), 1294–1299 (2024).
33. I. Starshynov, M. Weimar, L. M. Rachbauer, *et al.*, "Model-free estimation of the Cramér-Rao bound for deep learning microscopy in complex media," *Nat. Photonics* **19**(6), 593–600 (2025).
34. Q. Fang and D. A. Boas, "Monte Carlo Simulation of Photon Migration in 3D Turbid Media Accelerated by Graphics Processing Units," *Opt. Express* **17**(22), 20178–20190 (2009).
35. R. Yao, X. Intes, and Q. Fang, "Direct approach to compute Jacobians for diffuse optical tomography using perturbation Monte Carlo-based photon "replay"," *Biomed. Opt. Express* **9**(10), 4588–4603 (2018).
36. S. Yan and Q. Fang, "Hybrid mesh and voxel based Monte Carlo algorithm for accurate and efficient photon transport modeling in complex bio-tissues," *Biomed. Opt. Express* **11**(11), 6262–6270 (2020).
37. J. B. Domínguez and Y. Bérubé-Lauzière, "Light propagation from fluorescent probes in biological tissues by coupled time-dependent parabolic simplified spherical harmonics equations," *Biomed. Opt. Express* **2**(4), 817–837 (2011).
38. A. Allali, A. D. Klose, and Y. Bérubé-Lauzière, "Efficient numerical modelling of time-domain light propagation in curved 3D absorbing and scattering media with finite differences," *Biomed. Opt. Express* **12**(3), 1422–1436 (2021).
39. P. Cassano, A. P. Tran, H. Katnani, *et al.*, "Selective photobiomodulation for emotion regulation: model-based dosimetry study," *Neurophotonics* **6**(01), 1 (2019).
40. G. Taga, F. Homae, and H. Watanabe, "Effects of source-detector distance of near infrared spectroscopy on the measurement of the cortical hemodynamic response in infants," *NeuroImage* **38**(3), 452–460 (2007).

41. F. Lange and I. Tachtsidis, "Clinical Brain Monitoring with Time Domain NIRS: A Review and Future Perspectives," *Appl. Sci.* **9**(8), 1612 (2019).
42. M. Calderon-Arnulphi, A. Alaraj, and K. V. Slavin, "Near infrared technology in neuroscience: past, present and future," *Neurol. Res.* **31**(6), 605–614 (2009).
43. P. Pinti, I. Tachtsidis, A. Hamilton, *et al.*, "The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience," *Ann. N. Y. Acad. Sci.* **1464**(1), 5–29 (2020).
44. G. E. Strangman, Z. Li, and Q. Zhang, "Depth Sensitivity and Source-Detector Separations for Near Infrared Spectroscopy Based on the Colin27 Brain Template," *PLoS One* **8**(8), e66319 (2013).
45. S. Brigadoi and R. J. Cooper, "How short is short? Optimum source-detector distance for short-separation channels in functional near-infrared spectroscopy," *Neurophotonics* **2**(2), 025005 (2015).
46. A. Liebert, H. Wabnitz, D. Grosenick, *et al.*, "Evaluation of optical properties of highly scattering media by moments of distributions of times of flight of photons," *Appl. Opt.* **42**(28), 5785–5792 (2003).
47. K. J. Mitchell, V. Gradauskas, J. Radford, *et al.*, "Energy transport in diffusive waveguides," *Nat. Phys.* **20**(12), 1955–1959 (2024).
48. E. Okada, M. Firbank, M. Schweiger, *et al.*, "Theoretical and experimental investigation of near-infrared light propagation in a model of the adult head," *Appl. Opt.* **36**(1), 21–31 (1997).
49. S. Wang, N. Shibahara, D. Kuramashi, *et al.*, "Effects of spatial variation of skull and cerebrospinal fluid layers on optical mapping of brain activities," *Opt. Rev.* **17**(4), 410–420 (2010).
50. J. Radford, V. Gradauskas, K. J. Mitchell, *et al.*, "Photon transport through the entire adult human head," *Neurophotonics* **12**(02), 025014 (2025).
51. A. Torricelli, A. Pifferi, L. Spinelli, *et al.*, "Time-Resolved Reflectance at Null Source-Detector Separation: Improving Contrast and Resolution in Diffuse Optical Imaging," *Phys. Rev. Lett.* **95**(7), 078101 (2005).
52. A. Puszka, L. Hervé, A. Planat-Chrétien, *et al.*, "Time-domain reflectance diffuse optical tomography with Mellin-Laplace transform for experimental detection and depth localization of a single absorbing inclusion," *Biomed. Opt. Express* **4**(4), 569 (2013).
53. A. D. Mora, D. Contini, S. Arridge, *et al.*, "Towards next-generation time-domain diffuse optics for extreme depth penetration and sensitivity," *Biomed. Opt. Express* **6**(5), 1749 (2015).
54. A. Liebert, H. Wabnitz, and C. Elster, "Determination of absorption changes from moments of distributions of times of flight of photons: Optimization of measurement conditions for a two-layered tissue model," *J. Biomed. Opt.* **17**(5), 057005 (2012).
55. N. Hyvönen, A. Seppänen, and S. Staboulis, "Optimizing Electrode Positions in Electrical Impedance Tomography," *SIAM J. Appl. Math.* **74**(6), 1831–1851 (2014).
56. B. Maus, G. J. P. van Breukelen, R. Goebel, *et al.*, "Robustness of optimal design of fMRI experiments with application of a genetic algorithm," *NeuroImage* **49**(3), 2433–2443 (2010).
57. C. Bossuyt, A. J. den Dekker, D. Iuso, *et al.*, "Framework for optimal design of a multi-source rectangular X-ray cargo scanning system," *Radiat. Phys. Chem.* **239**, 113332 (2026).
58. A. Machado, Z. Cai, G. Pellegrino, *et al.*, "Optimal positioning of optodes on the scalp for personalized functional near-infrared spectroscopy investigations," *J. Neurosci. Methods* **309**, 91–108 (2018).
59. D. Kambly, C. Flindt, and M. Büttiker, "Factorial cumulants reveal interactions in counting statistics," *Phys. Rev. B* **83**(7), 075432 (2011).
60. J. Radford and D. Faccio, "Information transport and limits of optical imaging in the highly diffusive regime," *Phys. Rev. Res.* **5**(2), L022008 (2023).
61. M. Mozumder, P. Hirvi, I. Nissilä, *et al.*, "Diffuse optical tomography of the brain: Effects of inaccurate baseline optical parameters and refinements using learned post-processing," *Biomed. Opt. Express* **15**(8), 4470–4485 (2024).
62. D. Smyl and D. Liu, "Optimizing Electrode Positions in 2-D Electrical Impedance Tomography Using Deep Learning," *IEEE Trans. Instrum. Meas.* **69**(9), 6030–6044 (2020).
63. Z.-G. Wang, J. Radford, M.-J. Sun, *et al.*, "Fisher information for the design of diffuse optical brain monitoring systems," University of Glasgow (2026), <https://doi.org/10.5525/gla.researchdata.2062>.