Journal of Applied Physics

RESEARCH ARTICLE | DECEMBER 22 2025

Adagrad-optimized variational Bayesian reconstruction with sparsity-adaptive normal-generalized inverse Gaussian prior for fluorescence molecular tomography

Xin Cao ; Yiting He ; Xin Zhou ; Jiaxin Du ; Yi Chen ; Yangyang Liu ; Chengyi Gao;
Huangjian Yi

J. Appl. Phys. 138, 243103 (2025)
https://doi.org/10.1063/5.0305351

Journal of
Applied Physics

ARTICLE

pubs.aip.org/aip/jap

Adagrad-optimized variational Bayesian reconstruction with sparsity-adaptive normal-generalized inverse Gaussian prior for fluorescence molecular tomography

Cite as: J. Appl. Phys. **138**, 243103 (2025); doi: 10.1063/5.0305351
Submitted: 6 October 2025 · Accepted: 7 December 2025 ·
Published Online: 22 December 2025

Xin Cao,[1] Yiting He,[1] Xin Zhou,[1] Jiaxin Du,[1] Yi Chen,[1,2,a)] Yangyang Liu,[1,a)] Chengyi Gao,[3]
and Huangjian Yi[1]

AFFILIATIONS

[1] School of Information Science and Technology, Northwest University, Xi'an, Shaanxi 710127, China
[2] School of Electrical and Mechanical Engineering, The University of Adelaide, Adelaide, South Australia 5005, Australia
[3] Department of Oncology, The First Affiliated Hospital, Xi'an Jiaotong University, Xi'an, Shaanxi 710061, China

[a)] **Authors to whom correspondence should be addressed:** yichen.cgz@gmail.com and yyliu@nwu.edu.cn

ABSTRACT

Fluorescence molecular tomography (FMT) is a promising medical imaging technology with the ability to quantitatively reconstruct the three-dimensional distribution of fluorescently labeled probes *in vivo*. However, due to the strong scattering properties of biological tissues, conventional reconstruction methods encounter challenges such as low reconstruction accuracy and high computational complexity. Here, an adaptive online variational Bayesian method based on the normal-generalized inverse Gaussian (NGIG) prior is proposed. This method reduces computational complexity while ensuring that the globally optimal solution is maintained. Specifically, by utilizing variational inference, the optimization of the objective function is converted into a convex optimization problem that minimizes the variational lower bound, effectively reducing the function's complexity. Furthermore, to accurately capture the prior distribution, the NGIG prior is introduced. It imposes probabilistic constraints on the sparsity structure. This approach alleviates the adverse effects caused by overly strict sparsity constraints. In addition, the adaptive gradient algorithm (Adagrad) is employed to dynamically adjust the parameter learning rate, thereby preventing the algorithm from becoming trapped in local optima during the posterior inference process. The effectiveness of the proposed method is validated through numerical simulations and fluorescence source implantation experiments. The results show that the adaptive online variational Bayesian (AOVB)-NGIG method achieves superior performance in both fluorescence source localization and shape recovery. The minimum localization error is 0.243 mm, accompanied by a dice coefficient of 0.889. Meanwhile, the root mean square error and relative intensity error remain relatively low, indicating that the reconstructed results are the closest to the actual light source. These outcomes demonstrate that AOVB-NGIG can reliably reconstruct the spatial characteristics of the fluorescence source with high accuracy. This study is expected to advance the preclinical and clinical applications of FMT in early tumor detection.

I. INTRODUCTION

Fluorescence molecular imaging (FMI) is a non-invasive technique that enables real-time monitoring of structural and functional characteristics within biological tissues and has been extensively applied in both preclinical and clinical studies.[1] However, FMI is limited to acquiring only two-dimensional surface data, restricting the ability to obtain the spatial distribution of fluorescent targets. This limitation is crucial for the precise localization of fluorescent probes, quantifying biological processes, and improving diagnostic accuracy. Fluorescence molecular tomography (FMT) reconstructs the three-dimensional distribution of fluorescent probes *in vivo* at the molecular level,[2,3] exhibiting significant potential in tumor diagnosis and drug development.[4–6] Nevertheless, due to the significant scattering effects during photon propagation, the measured photon count is significantly lower than the actual number of emitted photons,[7] resulting in a highly ill-conditioned inverse problem for FMT.

Over the past two decades, significant efforts have been made by researchers to overcome these challenges. Initially, some scholars sought to improve the stability of the optimization process by incorporating constraints into the objective function,[8] such as sparse regularization,[9] which has been widely adopted. These include L_1 norm regularization (Lasso),[10] L_2 norm regularization (Tikhonov),[11] and L_p norm regularization (non-convex method) $(0 < p < 1)$.[12] Subsequently, De Mol and colleagues proposed elastic net (EN) regularization and total variation regularization,[13,14] which incorporate additional prior knowledge into the L_p regularization framework to better preserve the target's morphological features. Inspired by compressed sensing theory, L_1 norm-based regularization methods have been continuously optimized, including the fast iterative shrinkage-thresholding algorithm (FISTA), truncated conjugate gradient with incomplete variables (IVTCG), and the group orthogonal matching pursuit-based greedy algorithm (AGOMP).[15–17] Moreover, some researchers have incorporated regularization terms into the dictionary learning framework. These approaches construct redundant dictionaries to decompose complex signals into sparse linear combinations. By integrating prior knowledge, reconstruction accuracy is further enhanced.[18] Examples include K-dictionary learning methods such as K-SVD,[19] as well as the standard orthogonal matching pursuit (OMP) algorithm.[20] Although these approaches achieve satisfactory reconstruction to some extent, their optimization procedures remain intricate, and the imposed constraints are relatively weak.[13,19]

In addition, Bayesian-based methods have garnered widespread attention due to their reduced parameter selection complexity and improved stability.[21] These methods leverage Bayes' theorem to integrate prior information with observational data, iteratively updating the posterior probability distribution to optimize parameter estimation in sparse linear inverse problems, thereby enhancing reconstruction accuracy. Initially, Ando proposed the Bayesian predictive information criterion,[22] which improved the accuracy of the expected log-likelihood estimate by correcting the asymptotic bias of the posterior mean. Guided by this criterion, block-sparse Bayesian learning methods (SBL-LCGL),[23] the block-sparse Bayesian learning method with K-nearest neighbor strategy (KNN-GBSBL),[24] and adaptive group block-sparse Bayesian learning (AGBSBL)[25] were introduced to address sparse signal problems. Furthermore, the continuous development of hierarchical Bayesian[26] and variational Bayesian methods[27] has provided more flexible and efficient solutions for high-dimensional sparse reconstruction. However, Bayesian methods heavily rely on prior information, and their convergence speed is relatively slow in complex models or when applied to large-scale data sets.

Therefore, in this work, an adaptive online variational Bayesian (AOVB-NGIG) method for FMT reconstruction is proposed to overcome the above limitations. Specifically, to address the issue of computational complexity, variational inference is employed to reformulate the optimization process as a convex minimization problem of the variational lower bound. This reformulation significantly reduces computational costs. Additionally, the NGIG prior is utilized to impose constraints on the solution space, integrating the properties of both normal and generalized inverse Gaussian distributions to capture the sparsity and nonlinear characteristics inherent in the data. Within the context of Bayesian inference, this prior strikes a balance between model fitting accuracy and regularization, mitigating the risk of excessively sparse or overly smooth reconstructions. Consequently, the approach ensures the preservation of critical details and structural features within the reconstructed images. Simultaneously, the adaptive gradient algorithm (Adagrad) is applied to adjust the learning rate dynamically and to ensure optimized parameter updates while preventing convergence to local optima. The AOVB-NGIG method refines the learning process, accelerating convergence while maintaining reduced computational demands and preserving essential image features with high fidelity. A series of simulations and comparative experiments is conducted to validate the reconstruction performance of the proposed method.

The structure of this paper is organized as follows. In Sec. II, the FMT photon propagation model and the AOVB-NGIG method are introduced. In Sec. III, the experimental setup is described. In Sec. IV, the effectiveness of the proposed method is verified through numerical simulations and fluorescence source implantation experiments on a mouse. Finally, in Sec. V, the findings of this study are discussed and concluded.

II. METHODS

A. Photon propagation model

According to photon transport theory, the propagation of light in biological tissues can be described by the radiative transfer equation (RTE). However, the RTE equation is complex and involves numerous parameters. Directly solving, it is not only computationally intensive but also extremely challenging due to the heterogeneity of biological tissues.[28] Consequently, in steady-state FMT, the coupled diffusion equation (DE)[3] is frequently employed to model the propagation of photons in biological tissues within the near-infrared spectrum. The equation is expressed as follows:

$$\begin{cases} \nabla[D_x \nabla \Phi_x(r)] - \mu_{ax}\Phi_x(r) = -S(r), \; x \in \Omega, \\ \nabla[D_m \nabla \Phi_m(r)] - \mu_{am}\Phi_m(r) = -\Phi_x(r)x(r)', \; x \in \Omega, \end{cases} \tag{1}$$

where ∇ is the gradient operator, x and m represent the excitation and emission processes, D_x and D_m are the diffusion coefficients for the excitation light and emitted fluorescence, μ_{ax} and μ_{am} are the absorption coefficients for the excitation light and emitted fluorescence, r denotes the fluorescence source position, S is the light intensity distribution, $\Phi_x(r)$ and $\Phi_m(r)$ represent the light intensity distributions for the excitation light and emitted fluorescence, $x(r)$ is the fluorescence probe distribution to be determined, and Ω is the specified region.

Next, by applying the Robin boundary conditions,[29] the finite element method is used to discretize Eq. (1). This approach enables the efficient solution of the light propagation problem in tissues with complex geometries and boundary conditions. The transformation of Eq. (1) using the finite element method yields

$$\begin{cases} F_x\Phi_x = S_x, \; x \in \Omega, \\ F_m\Phi_m = FX', \; X \in \Omega, \end{cases} \tag{2}$$

where F_x and F_m are the system matrices, Φ_x and Φ_m are the photon densities during the excitation and emission processes, F is the discretized symmetric matrix, and X is the fluorescence target to be reconstructed.

By eliminating the non-measured values from the equation, the following linear relationship is obtained as

$$\Phi = AX. \tag{3}$$

Due to noise during data measurement and processing, certain deviations may occur. Therefore, the final linear relationship is derived as follows:

$$\Phi = AX + v, A \perp v \sim N(0, \gamma^2), \tag{4}$$

where Φ is the measurable fluorescence information on the surface of the organism, A is an $m \times n$ sparse matrix, where it represents the fluorescence intensity detected at each excitation point during the forward problem solution, X is the internal fluorescence source distribution, and v is the measured noise. $A \perp v$ denotes the independence between A and v, and γ is the scale factor of noise variance, where it characterizes the variance of noise in the direction orthogonal to matrix A.

B. FMT reconstruction based on the AOVB-NGIG method

Given the sparse distribution of fluorescent molecular probes in the reconstruction domain of FMT, compressed sensing reconstruction theory[30] is leveraged. By incorporating the AOVB framework, the optimization problem of the objective function is reformulated into a convex optimization model that integrates least squares loss with Bayesian inference. Considering the influence of noise, a new objective function is constructed as follows:

$$\mathcal{L}(X) = \min_{X}\left\{\|AX + v - \Phi\|_2^2 + \lambda D(q(X)||p(X))\right\}, \tag{5}$$

where $\|AX + \varepsilon - \Phi\|_2^2$ is the least squares fitting term, which quantifies the discrepancy between the model's predicted values and the observed measurements. $D(q(X)||p(X))$ is the Kullback–Leibler (KL) divergence of the variational distribution,[27] where $q(X)$ represents the variational distribution used to approximate the posterior distribution and $p(X)$ denotes the prior distribution. λ is the regularization parameter, which controls the balance between data fidelity and regularization. According to Eq. (5), it can be observed that in the process of variational Bayesian inference, the objective is to minimize $\mathcal{L}(X)$, where it corresponds to minimizing the KL divergence between the posterior distribution and the variational distribution.

To further refine the distribution of the fluorescence source X, the AOVB inference approach was employed to derive the posterior distribution from the observed data Φ. Within this framework, variational inference was initially utilized to estimate the posterior distribution $P(X|\Phi)$ given the observed measurements. Based on the Bayesian framework,[21] the posterior distribution can be expressed as

$$P(X|\Phi) = \frac{P(\Phi|X)P(X)}{P(\Phi)}, \tag{6}$$

where $P(X|\Phi)$ is the posterior distribution of X, $P(\Phi|X)$ is the likelihood function, $P(X)$ is the prior distribution, and $P(\Phi)$ is the marginal likelihood function, where it is used for model normalization.

Next, the NGIG prior is employed to accurately capture the sparsity of the fluorescence source. However, since the likelihood function follows a Gaussian distribution, which is not conjugate to the assumed prior, a hierarchical model[26] is adopted to reduce the computational complexity. In this context, the focus is placed on the marginal distribution prior of the hierarchical model,

$$P(X|\theta) = \prod\nolimits_{i=1}^{m} N(X_i|0, \theta_i), \tag{7}$$

$$P(\theta) = \prod\nolimits_{i=1}^{m} GIG(\theta_i; v, \delta, \lambda), \tag{8}$$

where $\theta = \{\theta_1, \theta_2, \ldots, \theta_p\}$ is a set of random variables, with each θ following a normal distribution, i.e., $\theta \sim N(0, \lambda)$. Π represents the product operation, X_i is the ith random variable, θ_i is the variance, and v, δ, and λ are the shape parameter, scale parameter, and location parameter. Among these, $GIG(\theta_i; v, \delta, \lambda) \propto \theta_i^{v-1}\exp[-1/2(\delta^2/\theta_i + \lambda^2\theta_i)]$.

The conditional distribution of the joint distribution is precisely known, as follows:

$$X|\theta, X_n, \Phi_n \sim N\left(X; K_n^{\theta}, C_n^{\theta}\right), \tag{9}$$

$$\theta_i|X, X_n, \Phi_n \sim GIG\left(\theta_i; v - 1/2, \sqrt{\delta^2 + A_i^2}, \lambda\right). \tag{10}$$

In Eq. (9), K_n^{θ} and C_n^{θ} are the parameters that need to be updated iteratively online, with the computation formula provided in Ref. 31,

$$\begin{aligned} K_n^{\theta} &= \left(X_n^T X_n + \gamma^2 D\left(\frac{1}{\theta}\right)\right)^{-1} X_n^T \Phi_n \\ &= D(\theta) X_n^T \left(\gamma^2 I_n + X_n D(\theta) X_n^T\right)^{-1} \Phi_n, \end{aligned} \tag{11}$$

$$\begin{aligned} C_n^{\theta} &= \left(\frac{1}{\gamma^2} X_n^T X_n + D\left(\frac{1}{\theta}\right)\right)^{-1} \\ &= \left(I_m - D(\theta) X_n^T (\gamma^2 I_n + X_n D(\theta) X_n^T)^{-1} X_n\right) D(\theta), \end{aligned} \tag{12}$$

where X_n is the design matrix containing the observed data, X_n^T is the transpose of the matrix X_n, and γ^2 is a coefficient for a penalty term used to prevent overfitting. $D(\theta)$ and $D(1/\theta)$ are diagonal matrices with diagonal elements θ and $1/\theta$, I_m and I_n are identity matrices of size $m \times m$ and $n \times n$, where the diagonal elements are all 1, and the off diagonal elements are 0.

In Bayesian inference, maximum a posteriori (MAP) estimation[27] can be employed to precisely solve for the parameters in Eqs. (11) and (12). The AOVB framework combines prior information with observed data, providing accurate MAP estimation and variational approximation of the posterior. This is achieved using the expectation-maximization (EM) algorithm.[32] The EM algorithm decomposes the original problem into two stages through an iterative optimization process: the expectation (E) step and the maximization (M) step.

In the E-step, the expectation of the posterior distribution $P(X|\Phi)$ needs to be calculated. During this process, the posterior probability distribution in the E-step is defined as $q^t(\theta) = P(\theta|X^t, \Phi)$. At the same time, $X^{t+1} = \mathrm{argmax} Q(X|X^t)$ is set, and by combining the likelihood function with the prior distribution, the objective function can be derived as

$$\begin{aligned} Q(X|X^t) &= \int \log\left(\frac{P(\Phi, X, \theta)}{P(\theta|X^t, \Phi)}\right) P(\theta|X^t, \Phi) d\theta \\ &= -\int \log(P(X, \Phi_n|\theta, X_n)) P(\theta|X^t, X_n, \Phi_n) d\theta + k(X^t, X_n, \Phi_n) \\ &= \frac{1}{2} X^T D\left(\frac{1}{\theta^t}\right) X + \frac{1}{2\gamma^2} |\Phi_n - X_n X|_2^2 + k(X^t, X_n, \Phi_n), \end{aligned} \tag{13}$$

where $1/\theta^t := \mathbb{E}[1/\theta^t | X^t, X_n, \Phi_n]$ and $Q(X|X^t)$ is the variational distribution, used to approximate the posterior distribution. $P(\Phi, X, \theta)$ is the joint probability distribution, and $P(\theta|X^t, \Phi)$ is the posterior distribution. X^t represents the variational parameters, k is a constant, and $d\theta$ denotes the integral with respect to the parameter θ.

In the M-step, the expectation function $Q(X|X^t)$ needs to be maximized to update the parameters. The equation is as follows:

$$X^{t+1} = \underset{X}{\arg\max}\, [Q(X|X^t)]. \tag{14}$$

During the solving process, to avoid local optima, the adaptive gradient algorithm (Adagrad)[33] is applied to adjust $Q(X|X^t)$. The Adagrad algorithm dynamically adjusts the learning rate for each parameter based on its historical gradients. This results in smaller learning rates for frequently updated parameters and larger learning rates for sparse updates. This adaptive mechanism effectively accommodates varying update frequencies across parameters, eliminating the need for manual learning rate adjustments and enhancing optimization efficiency. In Eq. (13), the optimization of the parameter θ is primarily addressed. Thus, during the maximization of the expectation function, the parameter θ undergoes continuous online dynamic adjustments, including both gradient accumulation and parameter updating, as illustrated in Eqs. (15) and (16),

$$s^t := s^{t-1} + \nabla_\theta Q(X|X^t) \odot \nabla_\theta Q(X|X^t), \tag{15}$$

$$\theta^t := \theta^{t-1} - \frac{\eta}{\sqrt{s^{(t)} + \epsilon}} \odot \nabla_\theta Q(X|X^t), \tag{16}$$

where s^t and θ^t are the accumulated squared gradients (used to adjust the learning rate) and the parameter values at the tth iteration. $\nabla_\theta Q(X|X^t)$ is the gradient of the objective function $Q(X|X^t)$ with respect to the parameter θ, and $\odot$ is the symbol for element-wise (Hadamard) multiplication, representing the square of the gradient. η is the learning rate, and ϵ is a small constant 10^{-6} used to prevent division by zero or to maintain numerical stability. By utilizing the Adagrad algorithm, more stable parameter updates can be achieved during the training process, thereby accelerating convergence.

Building upon this, the EM framework further minimizes the KL divergence of the posterior distribution by maximizing the evidence lower bound (ELBO),[23] thereby more effectively approximating the true posterior distribution. The relationship among the three is shown in Eq. (17),

$$\begin{aligned} &\log P(\Phi) - \int \log\left(\frac{P(\Phi, X, \theta)}{q(\theta)q(X)}\right) q(\theta)q(X) d\theta dX \\ &=: KL[q(\theta)q(X)||P(X, \theta|\Phi)]. \end{aligned} \tag{17}$$

Since the objective functions of $q(\theta)$ and $q(X)$ are convex and can be precisely minimized, iterative algorithms exist for $q^{t+1}(\theta)$ and $q^{t+1}(X)$,

$$\begin{aligned} q^{t+1}(\theta) &\propto \exp\left(\int \log P(\Phi, X, \theta) q^t(X) dX\right) \\ &\propto \exp\left(-\frac{1}{2}\sum\nolimits_{i=1}^{m} \frac{E_t\left[X_i^2\right]}{\theta_i}\right) P(\theta|\lambda), \end{aligned} \tag{18}$$

$$\begin{aligned} q^{t+1}(X) &\propto \exp\left(\int \log P(\Phi, X, \theta) q^{t+1}(\theta) d\theta\right) \\ &\propto \exp\left(-\frac{1}{2}\sum\nolimits_{i=1}^{m} X_i^2 E_{t+1}\left[\frac{1}{\theta_i}\right] - \frac{1}{2\gamma^2}|\Phi_n - X_n X|_2^2\right), \end{aligned} \tag{19}$$

where E_t is the expectation of the intermediate variational distribution at iteration t, which corresponds to coordinate ascent variational inference.[34]

For the posterior probability distribution, the model is enhanced through online updates and hyperparameter optimization. Variational inference is then used to iteratively sample from the conditional posterior distribution of each parameter, approximating the samples of the entire posterior distribution. By defining the observation matrix $\hat{\Phi}_n$ and X_n, both past and current observation data are utilized for estimation, thus incorporating historical information into the update process,

$$\hat{\Phi}_n = \begin{pmatrix} \hat{X}_{n-1} R_{n-1} \\ \tilde{\Phi}_n \end{pmatrix} \in R^{2M}, X_n = \begin{pmatrix} \hat{X}_{n-1} \\ \tilde{X}_n \end{pmatrix} \in R^{2M\times m}. \tag{20}$$

Therefore, the final parameter values after the update are

$$K_n^{t+1} = K_{n-1} + D(\theta_n^{t+1}) X_n^T \left(\gamma^2 I_{2M} + X_n D(\theta_n^{t+1}) X_n^T\right)^{-1} (\hat{\Phi}_n - X_n K_{n-1}), \tag{21}$$

$$C_n^{t+1} = \left(I_p - D(\theta_n^{t+1}) X_n^T (\gamma^2 I_{2M} + X_n D(\theta_n^{t+1}) X_n^T)^{-1} X_n\right) D(\theta_n^{t+1}), \tag{22}$$

where I is the identity matrix. During the iteration process, the latest parameters are provided by the previous step, with the parameters being continuously updated until the optimal values are obtained. The algorithm has a memory cost of only $O(M_p)$, demonstrating favorable performance.

In summary, the complete process for solving the parameters using the AOVB-NGIG method is outlined in Algorithm 1.

Algorithm 1. AOVB-NGIG

```
Input: vector $X$, parameters $v$, $\delta$, and $\lambda$, measured surface photon distribution $\Phi$.
Initialization: $T > 0$, maximum number of iterations is $k_{\max} = 1000$, convergence tolerance is $err = 1e - 6$.
While $t \leq T$ and $d((\theta_t, K_n^\theta, C_n^\theta), (\theta_{t-1}, K_n^{t-1}, C_n^{t-1})) > v$
    i. Compute $\theta_{AOVB}^{t+1}$ and $\theta_{EM}^{t+1}$;
    ii. Compute
        $G_{AOVB}^{t+1} = D(\theta_n^{t+1}) X_n^T (\gamma^2 I_{2M} + X_n D(\theta_n^{t+1}) X_n^T)^{-1}$
        $G_{EM}^{t+1} = D(\theta_n^{t+1}) X_n^T (\gamma^2 I_{2M} + X_n D(\theta_n^{t+1}) X_n^T)^{-1}$
    iii. Compute
        $K_n^{t+1} = K_{n-1} + G_{AOVB}^{t+1}(\hat{\Phi}_n - X_n K_{n-1})$
        $C_n^{t+1} = (I_p - G_{EM}^{t+1} X_n) D(\theta_n^{t+1})$
        $\theta^{t+1} = \theta^{t-1} - \frac{\eta}{\sqrt{s^{(t)} + \varepsilon}} \odot \nabla_\theta Q(X|X^t)$
    iv. $t = t + 1$.
End while
Output: At any given time, $(K_n^*, C_n^*) \in R^{2p+p^2}$
```

After obtaining the optimal parameter updates, the posterior distribution at the next time step is derived through MAP estimation. Specifically, the fluorescence source distribution X, given the current parameters and observation data, can be represented by a generalized inverse Gaussian mixture distribution. Based on the variational inference framework, the final posterior distribution is as follows:

$$q^{t+1}(X) = GIG\left(X_i; v - \frac{1}{2}, \sqrt{\delta^2 + C_{n,ii}^t + \left(K_{n,i}^t\right)^2}, \lambda\right). \tag{23}$$

By iterating repeatedly, the true fluorescence source distribution can be progressively approximated, thereby achieving high-precision reconstruction. In each iteration, the current parameter update θ_t and observation data Φ are utilized to continuously optimize the posterior distribution through variational Bayesian inference. Ultimately, the final estimate of X is obtained as follows:

$$X^* = \arg\max_X P(X|\Phi), \tag{24}$$

where X^* is the final estimate of X.

III. EXPERIMENT DESIGN

In this section, numerical simulations and fluorescence source implantation experiments were performed to validate the reconstruction performance of the AOVB-NGIG method. Furthermore, CG-Tikhonov,[35] IVTCG,[16] and OMP-L_1[20] methods were selected for comparison in terms of accuracy and shape recovery. All experiments and computations were executed on a personal computer equipped with Intel® Core™ i5-10210U CPU (1.60 GHz) and 8 GB RAM.

A. Evaluation indexes

To assess the reconstruction quality and accuracy of the AOVB-NGIG method more precisely, several quantitative metrics are employed. These include location error (LE),[36] dice coefficient (DICE),[23] root mean square error (RMSE),[37] and relative intensity error (RIE).[38]

LE represents the localization error. It measures the Euclidean distance between the reconstructed fluorescence source center and the actual fluorescence source position. The value of LE ranges from 0 to 1. A value closer to 0 indicates a higher localization precision of the algorithm,

$$LE = \|L_r - L_s\|_2^2, \tag{25}$$

where L_r and L_s are the centers of the reconstructed and actual targets.

DICE is used to measure the overlap between two sets. Its value ranges from 0 to 1. A DICE value closer to 1 indicates a higher similarity between the reconstructed and actual targets,

$$\text{DICE} = 2\frac{|R_r \cap R_s|}{|R_r| + |R_s|}, \tag{26}$$

where R_r and R_s are the regions of the reconstructed and actual targets.

RMSE is utilized to assess the difference between the reconstructed source center and the actual source center. The value ranges from 0 to infinity. A smaller RMSE indicates superior reconstruction quality,

$$\text{RMSE} = \sqrt{\frac{\sum_N (X_r - X_s)}{N}}, \tag{27}$$

where N is the number of nodes, X_r is the reconstructed bioluminescence source intensity, and X_s is the actual value.

RIE is utilized to quantify the informational disparity between the reconstructed and the true images. A lower RIE value, closer to 0, indicates a better restoration of the fluorescence source intensity,

$$\text{RIE} = \frac{|I_r - I_s|}{I_s}, \tag{28}$$

where I_r and I_s are the intensities of the reconstructed and actual sources.

B. Numerical simulation setup

In the numerical simulation, a cylindrical model with a radius of 10 mm and a height of 30 mm was designed to represent biological tissue. This model consists of five organs: muscle, lung, heart, bone, and liver, as shown in Fig. 1(a). The detailed optical parameters for each organ are provided in Ref. 37.

In the forward process, a tetrahedral element was used to simulate photon propagation. The intensity of each source is 1 nw/mm^3, and the surface fluorescence distribution was simulated using the molecular optical simulation environment MOSE 2.3[39] based on the Monte Carlo method, as shown in Fig. 1(e). In the inverse

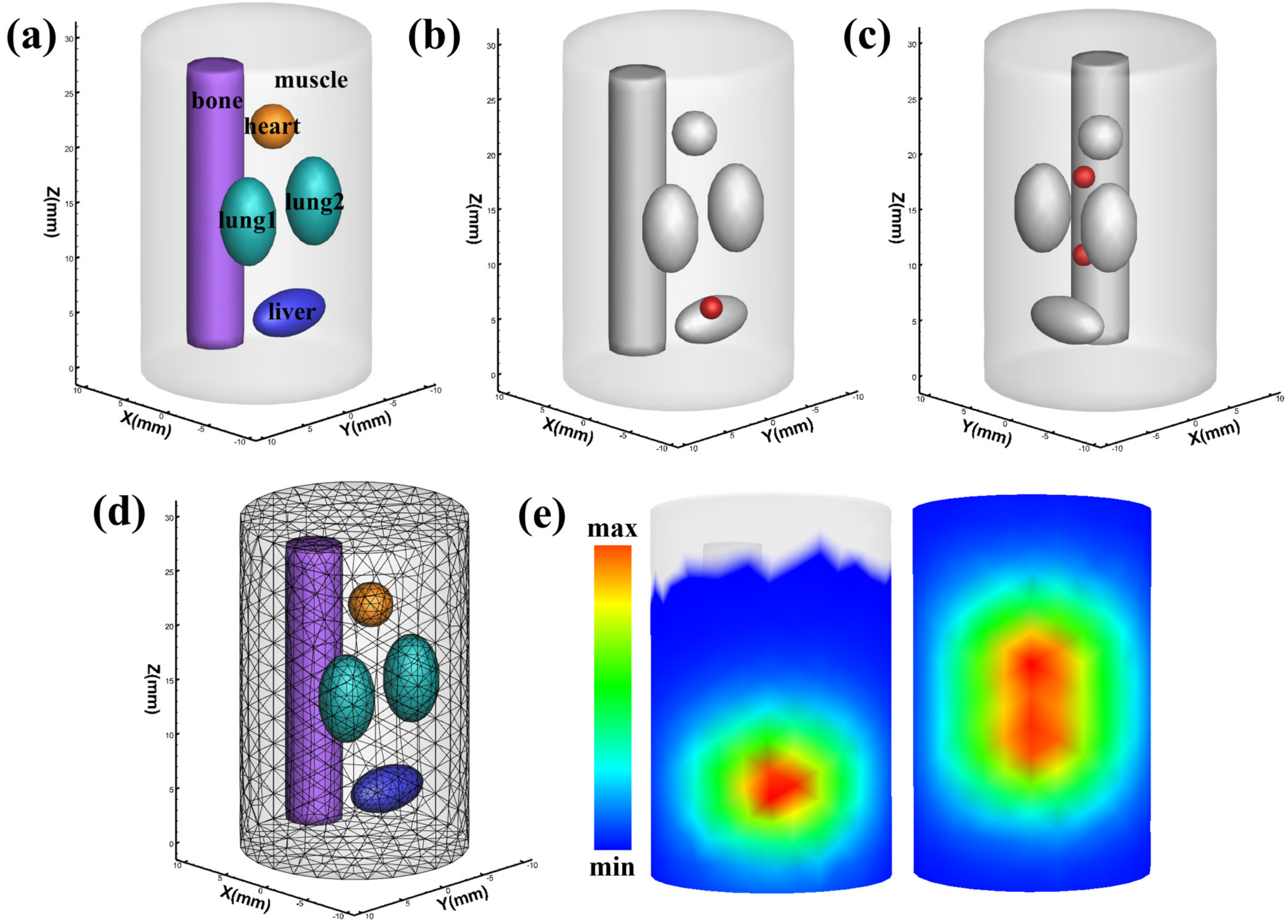

FIG. 1. (a) The 3D view of the cylinder model. (b) Cylindrical phantom with single-target. (c) Cylindrical phantom with dual-target. (d) The 3D view of meshed biological tissues. (c) Forward simulation of a cylindrical phantom with single-target. (e) Forward simulation of a cylindrical phantom with single-target and dual-target.

process, the geometry model was discretized into a uniform tetrahedral mesh using the COMSOL Multiphysics 6.1 platform (COMSOL, Inc., Burlington, Massachusetts), as shown in Fig. 1(d).

We designed single-source and dual-source simulation experiments to evaluate the performance of the AOVB-NGIG reconstruction method. In the single-source simulation experiment, one spherical fluorescence source with a radius of 1 mm was placed at (−6, 2, and 8) mm, as shown in Fig. 1(b). In the dual-source simulation experiment, two spherical fluorescence sources with a radius of 1 mm were placed at the center coordinates of (−1, 1, and 11) mm and (−1, 1, and 18) mm, as shown in Fig. 1(c). As noise is inevitable in FMT, the anti-noise experiment was also designed to evaluate the robustness of our method. Gaussian noise of 5%, 10%, 15%, and 20% was added to the measurement data based on the single-source numerical simulation, and then reconstruction was performed to assess the accuracy and robustness of the AOVB-NGIG method under different noise levels.

C. Implantation experiment setup

To further assess the practicality of the AOVB-NGIG method *in vivo* imaging, a fluorescence source implantation experiment was designed. All experimental animal protocols were approved by the Animal Ethics Committee of the Northwest University of China, and all procedures were strictly adhered to the approved guidelines. An adult BALB/C mouse was anesthetized for the fluorescence source implantation experiment. To minimize discomfort, all animal experiments were conducted under isoflurane gas anesthesia. The dual-modality FMT/CT imaging system used for data collection is shown in Fig. 2(a). This system consists of an optical acquisition module, a micro-CT system, and a control module. Specifically, the optical acquisition module includes an electron multiplying charge coupled device (EMCCD) camera, a continuous-wave optical laser, and optical filters. The micro-CT system comprises a high-resolution flat-panel x-ray detector and an

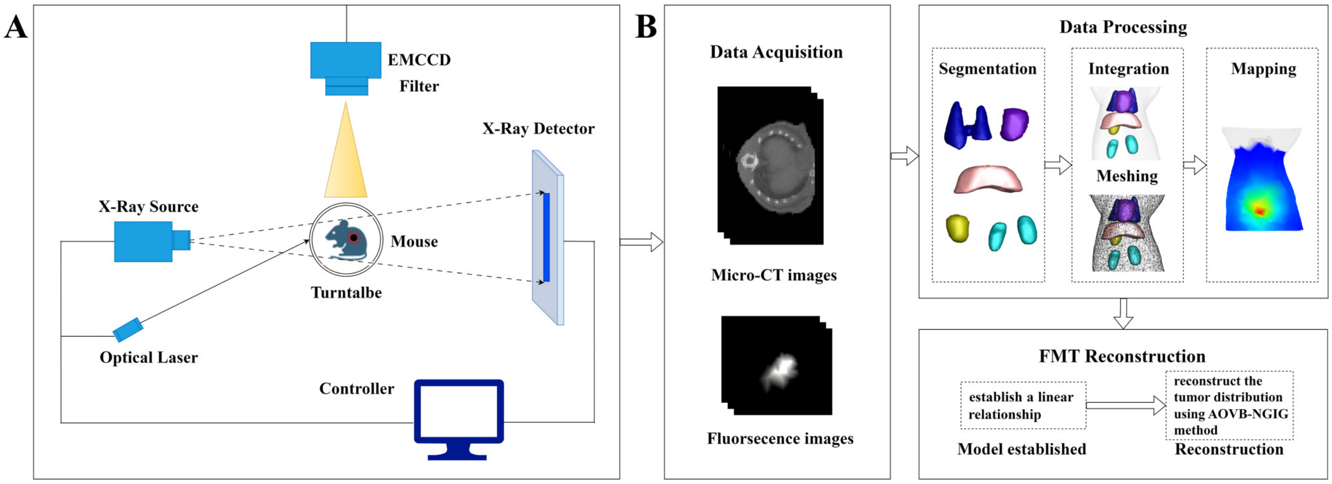

FIG. 2. (a) FMT/CT dual-modal imaging system. (b) Fluorescence source implantation experimental process [FMT/CT dual-modal imaging system data acquisition (top), data processing (bottom left), and FMT reconstruction (bottom right)].

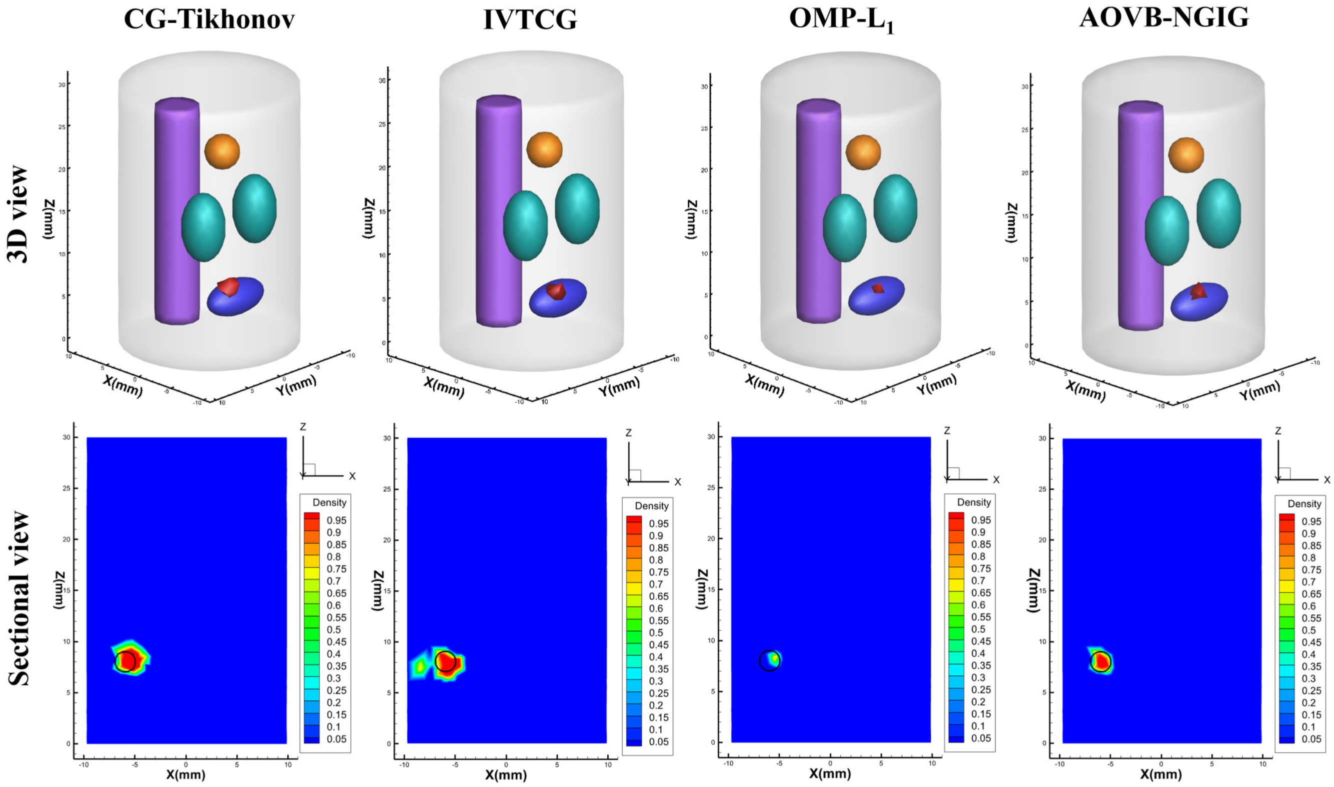

FIG. 3. Numerical simulation experiment single-source reconstruction result. The first row shows the 3D reconstruction results corresponding to the four methods, while the second row presents the cross-sectional view of the $Y = 2$ mm plane, where the black circle represents the true position of the fluorescent source.

TABLE I. Quantitative results of the single-source simulation experiment.

Method	True (mm)	Result (mm)	LE (mm)	DICE	RMSE	RIE
CG-Tikhonov	(−6, 2, 8)	(−5.159, 2.283, 8.026)	0.707	0.333	0.003 46	1.009
IVTCG	(−6, 2, 8)	(−5.585, 2.093, 7.687)	0.696	0.348	0.003 24	0.758
OMP-L_1	(−6, 2, 8)	(−5.483, 2.353, 8.259)	0.655	0.400	0.002 49	0.883
AOVB-NGIG	(−6, 2, 8)	(−5.827, 1.834, 7.994)	0.243	0.889	0.000 22	0.223

x-ray tube. The control module includes an electrically controlled rotating stage and an electrically controlled translation stage.

For data acquisition, a spherical fluorescent bead containing Cy5.5 solution with a radius of 1 mm was initially injected into the body of the mouse. After 6 h, laser light at 680 nm was emitted using a continuous-wave optical source. Surface fluorescence image with a 120° field of view was collected by a thermoelectric cooled electron multiplying charge coupled device (EMCCD) camera (−80 °C, iXonEM+888) with an exposure time of 1 s. Subsequently, the micro-CT system (tube voltage of 60 kVp, x-ray power of 40 W) was employed to perform CT imaging of the mouse and obtain 3D structural data.

After data collection was completed, the data processing stage began. Initially, we used Amira 2019.1 software (Amria, Visage Imaging, Australia) to segment the major organs, including the muscle, heart, lung, stomach, liver, and kidneys, and integrated them into the mouse body model. The fluorescent images were then mapped and registered onto the surface of the mouse body. The workflow for the fluorescence source implantation experiment is illustrated in Fig. 2(b). The fluorescence source was aligned and registered with the surface of the mouse trunk model. The true central position of the fluorescent target is (12.5, 11.5, and 19.5) mm, and the optical parameters for different organs were obtained from the literature.[40] It should be noted that due to the unknown intensity of the real fluorescence source, photon count calibration with the detected fluorescence source could not be performed. As a result, the evaluation metric RIE could not be obtained in the fluorescence source implantation mouse experiment.

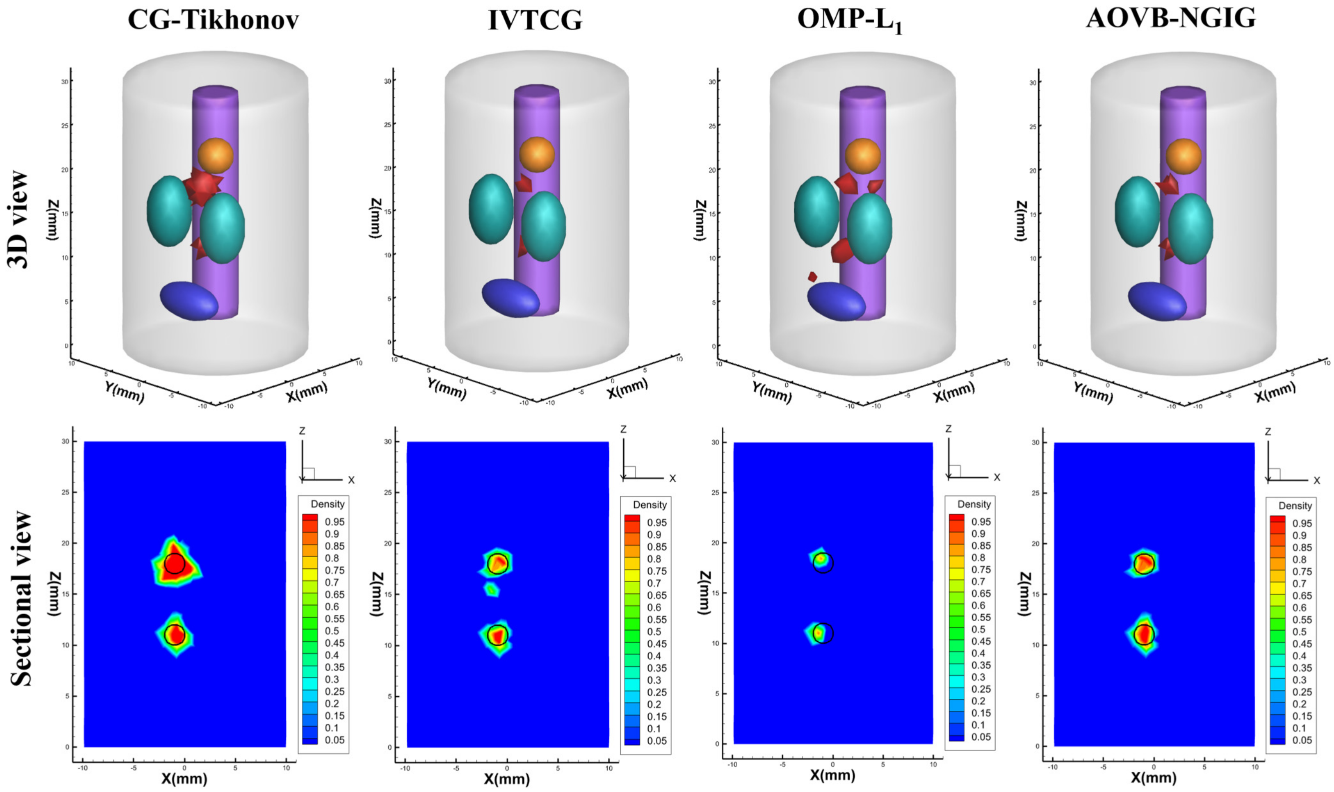

FIG. 4. Numerical simulation experiment dual-source reconstruction result. The first row shows the 3D reconstruction results corresponding to the four methods, while the second row presents the cross-sectional view of the $Y = 1$ mm plane, where the black circle represents the true position of the fluorescent source.

TABLE II. Quantitative results of the dual-source simulation experiment.

Method	True (mm)	Result (mm)	LE (mm)	DICE	RMSE	RIE
CG-Tikhonov	(−1, 1, 11)	(−0.887, 1.351, 11.330)	0.878	0.129	0.006 01	0.974
	(−1, 1, 18)	(−1.033, 1.290, 18.233)	0.868	0.140	0.005 92	1.084
IVTCG	(−1, 1, 11)	(−0.787, 1.211, 11.115)	0.728	0.308	0.004 89	0.705
	(−1, 1, 18)	(−0.861, 1.287, 18.246)	0.798	0.222	0.005 27	0.790
OMP-L_1	(−1, 1, 11)	(−1.451, 1.214, 10.974)	0.670	0.381	0.002 91	0.647
	(−1, 1, 18)	(−1.133, 1.054, 17.858)	0.620	0.444	0.003 16	0.553
AOVB-NGIG	(−1, 1, 11)	(−0.914, 1.198, 11.058)	0.397	0.727	0.001 65	0.360
	(−1, 1, 18)	(−1.056, 1.218, 17.927)	0.447	0.766	0.002 86	0.423

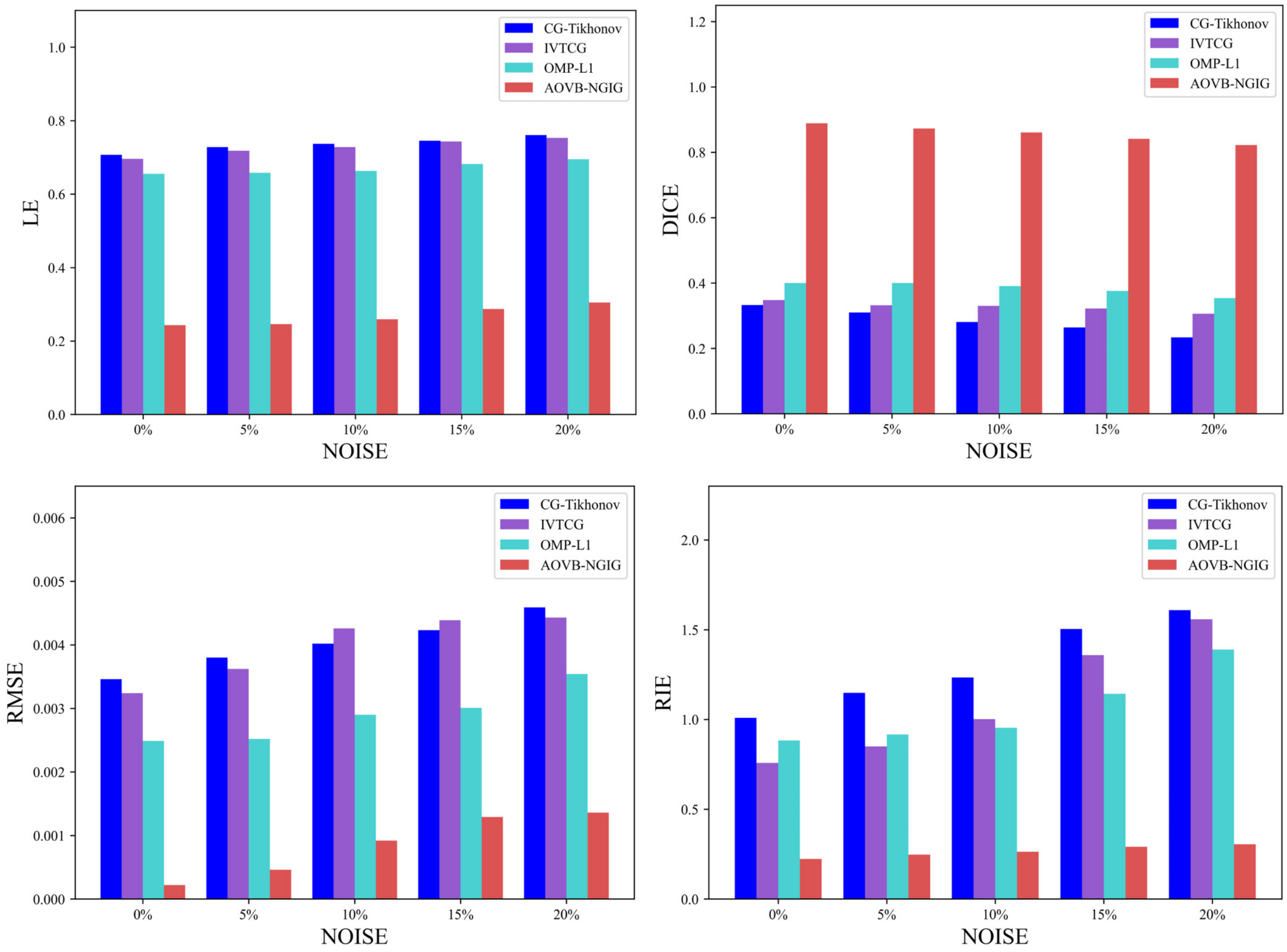

FIG. 5. The quantitative analysis results of the four methods under various Gaussian noise intensities.

IV. RESULTS

A. Numerical simulation results

1. Single-source simulation experiment

In the single-source simulation experiment, the 3D and sectional views of the reconstructed results with different methods were shown in Fig. 3. In the 3D view, the CG-Tikhonov and IVTCG methods generated overly smooth artifacts, while the OMP-L_1 method produced excessively convergent fluorescence source boundaries. In contrast, the AOVB-NGIG method yielded more accurate morphology and fewer image artifacts. The quantitative analysis of the reconstruction results of the four methods is listed in Table I. From these results, the LE, RMSE, and RIE for the AOVB-NGIG method were 0.223, 0.000 32, and 0.263. This indicates that the reconstructed fluorescence source position is closer to the actual fluorescence source, and the fluorescence yield is closer to the true value. Additionally, the DICE is relatively high at 0.889, which suggests a higher similarity between the reconstructed and true fluorescence sources. These results confirm the superior performance of the AOVB-NGIG method in fluorescence source reconstruction, morphological preservation, and spatial resolution.

2. Dual-source simulation experiment

In the dual-source simulation experiment, the 3D and sectional views of the reconstructed results with different methods are shown in Fig. 4, and the corresponding quantitative analysis results are summarized in Table II. Similar to the single-source case, the AOVB-NGIG method still outperforms the CG-Tikhonov, IVTCG, and OMP-L_1 methods in terms of multi-target resolution, with the reconstruction results being closer to the true sources and exhibiting fewer artifacts. The results in Table II indicate that the LE of the AOVB-NGIG method is significantly smaller, with an average of only 0.392, approximately half of that of the other methods. The DICE is also notably higher than the other three methods, reaching 0.001 65 and 0.002 86, indicating superior source localization and higher source overlap. At the same time, the RMSE and RIE are smaller, suggesting that the predicted values of the reconstructed sources have smaller spatial deviations. The results of the numerical simulation experiment demonstrate that the AOVB-NGIG method maintains strong source localization and reconstruction capabilities even with multiple sources, and it can accurately differentiate the specific locations of the sources.

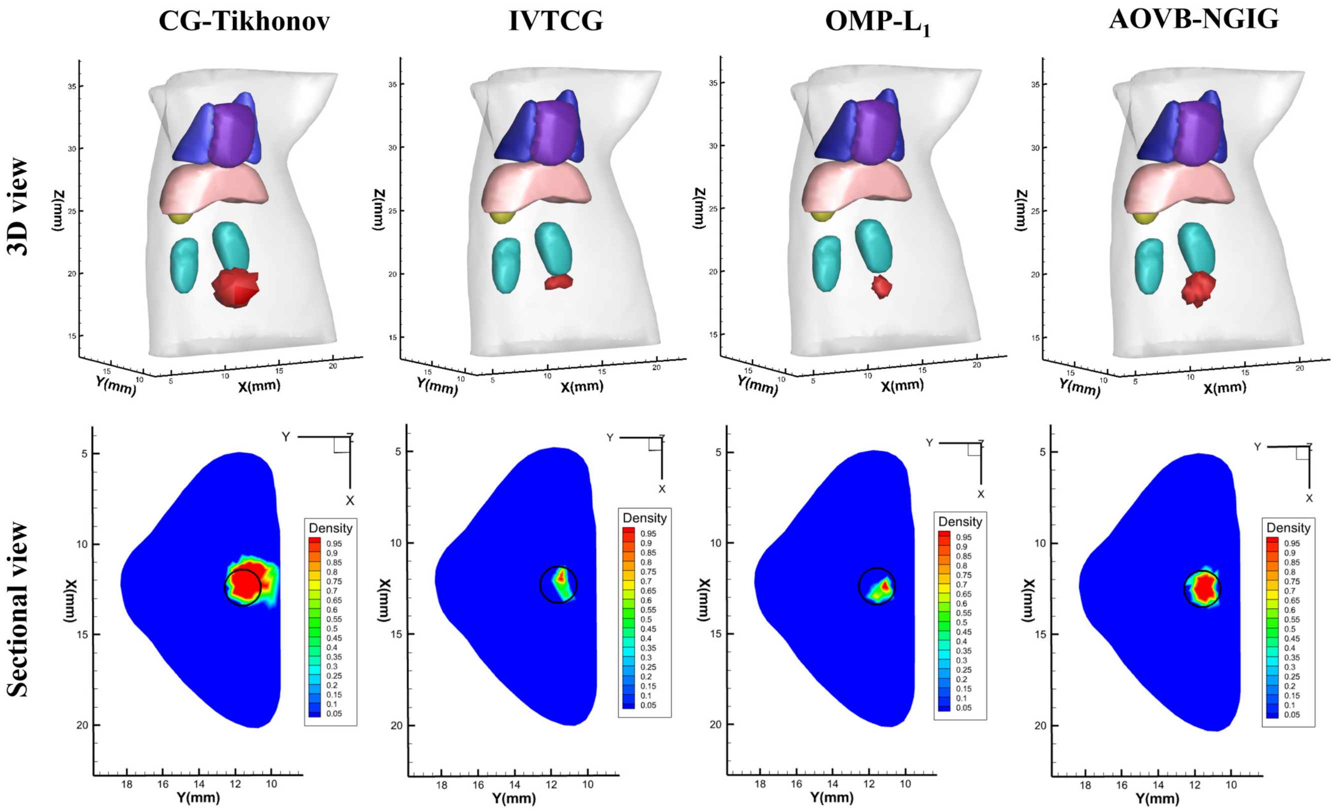

FIG. 6. Fluorescence source implantation experiment reconstruction results. The first row presents the 3D reconstruction results corresponding to the four methods. The second row displays the cross-sectional view at the $Z = 19.5$ mm plane, where the black circle indicates the actual location of the fluorescent source.

TABLE III. Quantitative results of the implantation experiment.

Method	True (mm)	Result (mm)	LE (mm)	DICE	RMSE
CG-Tikhonov	(12.5, 11.5, 19.5)	(12.368, 11.170, 19.275)	0.896	0.330	0.005 37
IVTCG	(12.5, 11.5, 19.5)	(12.841, 11.307, 19.802)	0.883	0.395	0.003 79
OMP-L_1	(2.5, 11.5, 19.5)	(12.845, 11.098, 19.262)	0.759	0.581	0.004 42
AOVB-NGIG	(12.5, 11.5, 19.5)	(12.500, 11.499, 19.173)	0.317	0.902	0.002 08

3. Anti-noise simulation experiment

As described in Eq. (4) of Sec. II, noise is inevitable in FMT. Therefore, a series of anti-noise simulation experiments was conducted to verify the noise resistance and robustness of the AOVB-NGIG method. The reconstruction performance of four methods was compared under Gaussian noise levels of 5%, 10%, 15%, and 20%. The results are shown in Fig. 5. Additionally, LE, DICE, RMSE, and RIE were selected as evaluation metrics to quantify the precision and reliability of the AOVB-NGIG method during the reconstruction process. Analysis of Fig. 5 shows that, under different noise levels, the LE values of all four methods remain relatively stable. However, the LE range for the AOVB-NGIG method is between 0.243 and 0.305 mm, demonstrating superior performance in comparison. Additionally, the AOVB-NGIG method achieves the highest DICE, with its mean remaining stable around 0.857 across various noise levels, indicating its ability to recover fluorescence distributions accurately. Furthermore, the variations in RMSE and RIE are minimal, with ranges of 0.000 22–0.001 36 and 0.223–0.305 mm, and both values remain below the optimal values for the other three methods under 0% noise. Quantitative results further confirm that, in the presence of noise in the fluorescence source, the AOVB-NGIG method outperforms the other methods in terms of reconstruction performance, with superior stability and robustness.

B. Implantation experiment results

The 3D view and sectional view are displayed in Fig. 6, and the quantitative evaluation results for the four methods are summarized in Table III. As illustrated in Fig. 6, the most favorable performance is achieved by the AOVB-NGIG method. The quantitative evaluation further corroborates this finding, as indicated in Table III. In the fluorescence source implantation experiment, the AOVB-NGIG method achieved the smallest LE value of 0.317, demonstrating a significantly higher localization accuracy. Additionally, a higher DICE of 0.902 and a lower RMSE of 0.002 08 were obtained. Compared to CG-Tikhonov, IVTCG, and OMP-L_1, the AOVB-NGIG method exhibited more than twice the accuracy, underscoring its superior reconstruction fidelity and enhanced fluorescence source recovery capability. Collectively, the fluorescence source implantation experiment further validated the exceptional performance of the AOVB-NGIG approach in FMT reconstruction.

V. DISCUSSION AND CONCLUSION

As a three-dimensional imaging modality, FMT enables the quantitative reconstruction of fluorescence distribution within biological tissues. However, its reconstruction process is constrained by ill-conditioned inverse problems, leading to insufficient reconstruction accuracy. To better recover the three-dimensional distribution of early-stage tumors, an AOVB-NGIG method is proposed in this study to achieve high-precision morphological reconstruction of tumor regions in FMT. This method employs the NGIG prior to capturing prior information and utilizes variational inference to decompose the original objective function into a more tractable convex problem. Finally, the Adagrad algorithm is incorporated, allowing the AOVB-NGIG method to dynamically adjust the learning rate throughout the iterative process, ensuring stable convergence. This guarantees the sparsity of the solution and the robustness of the reconstruction process.

To evaluate the performance of the proposed AOVB-NGIG method, both numerical simulations and fluorescence source implantation experiments with a live mouse were conducted, followed by qualitative and quantitative comparisons with the CG-Tikhonov, IVTCG, and OMP-L_1 methods. The results of the numerical simulation experiments demonstrate that under single-source, dual-source, and noise-resistant conditions, the AOVB-NGIG method achieves optimal results for LE, DICE, RMSE, and RIE while also providing stable reconstruction outcomes. In addition, the single-source implantation experimental results demonstrate that, compared to other methods, the tumor regions reconstructed by the AOVB-NGIG method more closely align with the true distribution. This method is capable of more accurately reconstructing the fluorescence source and achieving a higher fluorescence yield. Overall, the AOVB-NGIG method demonstrates significant advantages in terms of localization accuracy, spatial resolution, fluorescence yield recovery, and morphological restoration.

Although the AOVB-NGIG method has achieved outstanding results in terms of reconstruction accuracy, localization precision, relative sparsity, robustness, and morphological similarity, certain limitations remain. First, variational inference is sensitive to initial conditions, and the optimization process may become trapped in local minima. Second, the parameter optimization requires substantial memory. Finally, the validation was predominantly restricted to implantation models and did not include more complex *in vivo* conditions. To overcome these limitations, future studies will be directed toward the integration of additional prior knowledge and the refinement of optimization strategies. Moreover, experiments employing mouse tumor burden models will be performed to more accurately replicate physiological conditions. Overall, the AOVB-NGIG method has yielded promising results in FMT, and future investigations will further refine its shape reconstruction accuracy and explore its potential in preclinical applications.

In summary, an AOVB-NGIG method that integrates the NGIG prior with variational Bayesian inference is proposed for FMT reconstruction. This approach significantly enhances reconstruction accuracy while effectively reducing computational complexity through variational inference. The algorithm is validated by a series of numerical simulations and implantation experiment, and its performance is compared with three conventional reconstruction methods. The experimental results indicate that the AOVB-NGIG method has yielded promising results in FMT, and future investigations will further refine its shape reconstruction accuracy and explore its potential in preclinical applications.

ACKNOWLEDGMENTS

This work was supported in part by the National Major Scientific Research Instrument Development Projects of China (No. 82127805), the Key Research and Development Program of Shaanxi Province (No. 2024SF-YBXM-681), and the National Natural Science Foundation of China (NSFC) (Nos. 61701403 and 61806164).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Xin Cao: Methodology (equal); Software (equal); Writing – original draft (equal). **Yiting He:** Methodology (equal); Writing – original draft (equal). **Xin Zhou:** Validation (equal). **Jiaxin Du:** Resources (equal); Validation (equal). **Yi Chen:** Supervision (equal); Writing – review & editing (equal). **Yangyang Liu:** Supervision (equal); Writing – original draft (equal). **Chengyi Gao:** Conceptualization (equal). **Huangjian Yi:** Investigation (equal).

DATA AVAILABILITY

Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon request.

REFERENCES

[1] Z. Zhang, K. He, C. Chi, Z. Hu, and J. Tian, "Intraoperative fluorescence molecular imaging accelerates the coming of precision surgery in China," Eur. J. Nucl. Med. Mol. Imaging **49**(8), 2531–2543 (2022).

[2] C. Qin, J. Zhong, Z. Hu, X. Yang, and J. Tian, "Recent advances in Cerenkov luminescence and tomography imaging," IEEE J. Sel. Top. Quantum Electron. **18**(3), 1084–1093 (2012).

[3] Z. Hu, X. Chen, J. Liang, X. Qu, D. Chen, W. Yang, J. Wang, F. Cao, and J. Tian, "Single photon emission computed tomography-guided Cerenkov luminescence tomography," J. Appl. Phys. **112**(2), 024703 (2012).

[4] H. Chen, K. Shou, S. Chen, C. Qu, Z. Wang, L. Jiang, M. Zhu, B. Ding, K. Qian, A. Ji, H. Lou, L. Tong, A. Hsu, Y. Wang, D. W. Felsher, Z. Hu, J. Tian, and Z. Cheng, "Smart self-assembly amphiphilic cyclopeptide-dye for near-infrared window-II imaging," Adv. Mater. **33**(16), 2006902 (2021).

[5] Z. Hu, C. Fang, B. Li, Z. Zhang, C. Cao, M. Cai, S. Su, X. Sun, X. Shi, C. Li, T. Zhou, Y. Zhang, C. Chi, P. He, X. Xia, Y. Chen, S. S. Gambhir, Z. Cheng, and J. Tian, "First-in-human liver-tumour surgery guided by multispectral fluorescence imaging in the visible and near-infrared-I/II windows," Nat. Biomed. Eng. **4**(3), 259–271 (2020).

[6] C. Cao, S. Deng, B. Wang, X. Shi, L. Ge, M. Qiu, F. Zhang, M. Lu, L. Ma, C. Chi, Z. Hu, J. Tian, and S. Zhang, "Intraoperative near-infrared II window fluorescence imaging-assisted nephron-sparing surgery for complete resection of cystic renal masses," Clin. Transl. Med. **11**(10), e604 (2021).

[7] Z. Hu, M. Zhao, Y. Qu, X. Zhang, M. Zhang, M. Liu, H. Guo, Z. Zhang, J. Wang, W. Yang, and J. Tian, "*In vivo* 3-dimensional radiopharmaceutical-excited fluorescence tomography," J. Nucl. Med. **58**(1), 169–174 (2017).

[8] Y. Wang, S. Ma, H. Yang, J. Wang, and X. Li, "On the effective inversion by imposing a priori information for retrieval of land surface parameters," Sci. China Ser. D **52**(4), 540–549 (2009).

[9] C. Li, G. Wang, J. Qi, and S. R. Cherry, "Three-dimensional fluorescence optical tomography in small-animal imaging using simultaneous positron-emission-tomography priors," Opt. Lett. **34**(19), 2933–2935 (2009).

[10] Q. Zhang, H. Zhao, D. Chen, X. Qu, X. Chen, X. He, W. Li, Z. Hu, J. Liu, J. Liang, and J. Tian, "Source sparsity based primal-dual interior-point method for three-dimensional bioluminescence tomography," Opt. Commun. **284**(24), 5871–5876 (2011).

[11] C. Li, G. S. Mitchell, and S. R. Cherry, "Cerenkov luminescence tomography for small-animal imaging," Opt. Lett. **35**(7), 1109–1111 (2010).

[12] H. Guo, Z. Hu, X. He, X. Zhang, M. Liu, Z. Zhang, X. Shi, S. Zheng, and J. Tian, "Non-convex sparse regularization approach framework for high multiple-source resolution in Cerenkov luminescence tomography," Opt. Express **25**(23), 28068–28085 (2017).

[13] C. De Mol, E. De Vito, and L. Rosasco, "Elastic-net regularization in learning theory," J. Complexity **25**(2), 201–230 (2009).

[14] S. Niu, Y. Gao, Z. Bian, J. Huang, W. Chen, G. Yu, Z. Liang, and J. Ma, "Sparse-view x-ray CT reconstruction via total generalized variation regularization," Phys. Med. Biol. **59**(12), 2997 (2014).

[15] D. Kim and J. A. Fessler, "Another look at the fast iterative shrinkage/thresholding algorithm (FISTA)," SIAM J. Optim. **28**(1), 223–250 (2018).

[16] P. Jiao, H. Yi, Y. Hu, and X. He, "A permissible region strategy for fluorescence molecular tomography," Opt. Rev. **26**(6), 523–530 (2019).

[17] L. Kong, Y. An, Q. Liang, L. Yin, Y. Du, and J. Tian, "Reconstruction for fluorescence molecular tomography via adaptive group orthogonal matching pursuit," IEEE Trans. Biomed. Eng. **67**(9), 2518–2529 (2020).

[18] K. Skretting and K. Engan, "Recursive least squares dictionary learning algorithm," IEEE Trans. Signal Process. **58**(4), 2121–2130 (2010).

[19] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," IEEE Trans. Signal Process. **54**(11), 4311–4322 (2006).

[20] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," IEEE Trans. Inf. Theory **53**(12), 4655–4666 (2007).

[21] B. T. Knapik, A. W. Van Der Vaart, and J. H. van Zanten, "Bayesian inverse problems with Gaussian priors," Ann. Stat. **39**(5), 2626–2657 (2011).

[22] T. Ando, "Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models," Biometrika **94**(2), 443–458 (2007).

[23] Y. Wang, H. Wang, Q. Zhu, Y. Chen, L. Su, H. Yi, C. Gao, and X. Cao, "SBL-LCGL: Sparse Bayesian learning based on Laplace distribution for robust cone-beam x-ray luminescence computed tomography," Phys. Med. Biol. **69**(17), 175020 (2024).

[24] L. Yin, K. Wang, T. Tong, Y. An, H. Meng, X. Yang, and J. Tian, "Improved block sparse Bayesian learning method using K-nearest neighbor strategy for accurate tumor morphology reconstruction in bioluminescence tomography," IEEE Trans. Biomed. Eng. **67**(7), 2023–2032 (2019).

[25] L. Yin, K. Wang, T. Tong, Q. Wang, Y. An, X. Yang, and J. Tian, "Adaptive grouping block sparse Bayesian learning method for accurate and robust reconstruction in bioluminescence tomography," IEEE Trans. Biomed. Eng. **68**(11), 3388–3398 (2021).

[26] C. Faure, J. Antoni, F. Ablitzer, and C. Pézerat, "A fully Bayesian approach based on Bernoulli–Gaussian prior for the identification of sparse vibratory

sources from displacement measurements," J. Sound Vib. **524**, 116726 (2022).

[27] D. Urozayev, B. Ait-El-Fquih, and D. B. Peter, "Sparse variational Bayesian inversion for subsurface seismic imaging," IEEE Trans. Geosci. Remote Sens. **62**, 1–13 (2024).

[28] A. D. Klose, V. Ntziachristos, and A. H. Hielscher, "The inverse source problem based on the radiative transfer equation in optical molecular imaging," J. Comput. Phys. **202**(1), 323–345 (2005).

[29] H. Guo, Y. Hou, X. He, J. Yu, J. Cheng, and X. Pu, "Adaptive hp finite element method for fluorescence molecular tomography with simplified spherical harmonics approximation," J. Innovative Opt. Health Sci. **07**(02), 1350057 (2014).

[30] Y. Tsaig and D. L. Donoho, "Extensions of compressed sensing," Signal Process. **86**(3), 549–571 (2006).

[31] G. E. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis* (John Wiley & Sons, 2011).

[32] S. Zhang, Y. Liu, X. Li, and G. Bi, "Variational Bayesian sparse signal recovery with LSM prior," IEEE Access **5**, 26690–26702 (2017).

[33] R. Ward, X. Wu, and L. Bottou, "AdaGrad stepsizes: Sharp convergence over nonconvex landscapes," J. Mach. Learn. Res. **21**(219), 1–30 (2020), available at http://jmlr.org/papers/v21/18-352.html.

[34] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," J. Am. Stat. Assoc. **112**(518), 859–877 (2017).

[35] M. Cai, Z. Zhang, X. Shi, J. Yang, Z. Hu, and J. Tian, "Non-negative iterative convex refinement approach for accurate and robust reconstruction in Cerenkov luminescence tomography," IEEE Trans. Med. Imaging **39**(10), 3207–3217 (2020).

[36] H. Guo, L. Gao, J. Yu, X. He, H. Wang, J. Zheng, and X. Yang, "Sparse-graph manifold learning method for bioluminescence tomography," J. Biophotonics **13**(4), e201960218 (2020).

[37] Y. Chen, M. Du, W. Li, L. Su, H. Yi, F. Zhao, K. Li, L. Wang, and X. Cao, "ABPO-TVSCAD: Alternating Bregman proximity operators approach based on TVSCAD regularization for bioluminescence tomography," Phys. Med. Biol. **67**(21), 215013 (2022).

[38] Y. Chen, M. Du, J. Zhang, G. Zhang, L. Su, K. Li, F. Zhao, H. Yi, L. Wang, and X. Cao, "Generalized conditional gradient method with adaptive regularization parameters for fluorescence molecular tomography," Opt. Express **31**(11), 18128–18146 (2023).

[39] N. Ren, J. Liang, X. Qu, J. Li, B. Lu, and J. Tian, "GPU-based Monte Carlo simulation for light propagation in complex heterogeneous tissues," Opt. Express **18**(7), 6811–6823 (2010).

[40] H. Yi, D. Chen, W. Li, S. Zhu, X. Wang, J. Liang, and J. Tian, "Reconstruction algorithms based on l 1-norm and l 2-norm for two imaging models of fluorescence molecular tomography: A comparative study," J. Biomed. Opt. **18**(5), 056013 (2013).