

# Inverse radiative transfer with goal-oriented $hp$ -adaptive mesh refinement: adaptive-mesh inversion

Shukai Du<sup>1</sup> and Samuel N. Stechmann<sup>1,2</sup>

<sup>1</sup>Department of Mathematics, University of Wisconsin–Madison

<sup>2</sup>Department of Atmospheric and Oceanic Sciences, University of Wisconsin–Madison

Submitted April 10, 2023; Revised August 18, 2023

## Abstract

The inverse problem for radiative transfer is important in many applications, such as optical tomography and remote sensing. Major challenges include large memory requirements and computational expense, which arise from high-dimensionality and the need for iterations in solving the inverse problem. Here, to alleviate these issues, we propose adaptive-mesh inversion: a goal-oriented  $hp$ -adaptive mesh refinement (AMR) method for solving inverse radiative transfer problems. One novel aspect here is that the two optimizations (one for inversion, and one for mesh adaptivity) are treated simultaneously and blended together. By exploiting the connection between duality-based mesh adaptivity and adjoint-based inversion techniques, we propose a goal-oriented error estimator, which is cheap to compute, and can efficiently guide the mesh-refinement to numerically solve the inverse problem. We use discontinuous Galerkin spectral element (DGSE) methods to discretize the forward and the adjoint problems. Then, based on the goal-oriented error estimator, we propose an  $hp$ -adaptive algorithm to refine the meshes. Numerical experiments are presented at the end and show convergence speed-up and reduced memory occupation by the goal-oriented mesh adaptive method.

## 1 Introduction

This paper concerns the *numerical solution of inverse radiative transfer equation*, which serves as the mathematical foundation for applications such as optical tomography [2, 4, 32], remote sensing [34, 13, 14, 29], and neutron transport [27, 28, 30]. Despite its wide applications, devising numerical methods for inverse radiative transfer is notoriously challenging because of the high-dimensionality of the forward problem, for which standard discretizations with sufficient accuracy would usually require large memory occupation, and can render the solver too slow to be applicable. Therefore, it is of great interest to devise numerical

methods which use fewer degrees of freedoms (DOFs) or less memory, while still maintain the accuracy requirement.

Among the various types of memory reduction methods, we are interested in *hp*-adaptive mesh refinement (*hp*-AMR) methods. This is for several reasons. Firstly, the method is versatile in the sense that it can efficiently represent the solution in regions where it is smooth, while also capturing local features. This advantage is especially beneficial for applications such as optical tomography and remote sensing, for which it is common to observe a pattern with a few local features embedded in a smooth background distribution (clouds in the sky for remote sensing; narrow, dirac-delta like inflow laser beam for optical tomography; etc.). Another reason concerns the well-developed theoretical understanding of adaptive mesh refinement (AMR), and its success in the field of solid and fluid mechanics. Indeed, the concept of *hp*-AMR traces back to at least the 1980s by the pioneering work of Babuška and his collaborators [18]; see also the review paper [5] for *hp*-finite element methods (FEM). It was shown that *hp*-adaptive FEM can achieve exponential convergence even when the solution presents singularities. This suggests the great potential of using *hp*-AMR to reduce the DOFs for radiative transfer. Our recent work [15] demonstrates this potential for the forward problem, while here we consider solving the inverse radiative transfer problem.

Before we proceed to the introduction of our proposed method, we first review some related existing work on inverse radiative transfer, or more generally on inverse transport problems. The well-posedness of inverse transport problems was studied in [12, 37, 6] by exploiting the singular decomposition of the albedo operators. Then, the stability of the inverse problem in different scaling was studied in [10, 11, 25, 26]. For the numerical discretization of the inverse problem, the time-independent inverse radiative transfer was considered in [24, 22, 23, 1], and the frequency domain problem was considered in [33]. Finally, [32] gives a review on the numerical techniques for the inverse transport problems in medical imaging. In [3], the effect of the numerical error on the quality of the reconstruction was discussed, where it was shown that the numerical error on the forward/adjoint solver can lead to significant errors in the reconstruction of the optical images.

Considering the efficiency of AMR in reducing numerical error, it is not surprising to see that AMR has also been applied to inverse problems. In [19, 20], the authors considered the diffuse optical imaging problem, where the connection between the forward/adjoint solver error and the reconstruction error was studied theoretically, and an AMR method was proposed based on their theoretical findings. In [36, 35], finite volume (FV) based AMR methods for fluorescence/luminescence imaging were studied numerically. In [7, 8], a FEM based AMR method was proposed for fluorescence-enhanced optical tomography, and a large savings in DOFs was observed in numerical tests.

In this paper, we propose a goal oriented *hp*-AMR method which is distinctive to the existing AMR approaches of solving the inverse radiative transfer equation in the following aspects. First, we devise a novel goal-oriented error estimator, which provides more efficient mesh-refinement strategies compared to energy or  $L^2$ -norm based error estimators; see [9, 39]. In addition, we use *hp*-AMR instead of only *h*-AMR with fixed  $p$ . The extra  $p$ -adaptivity allows more efficient representation of the solution where it is smooth. Finally, we consider the full radiative transfer equation instead of its diffusion-type approximation. The full radiative transfer equation is a more computationally challenging task because of the interplay of both the advection and the scattering terms. Since we aim for maximal

generality of our method for the different regimes of the radiative transfer equation, we use discontinuous Galerkin (DG) methods to approximate the forward and the adjoint equations. The DG methods can handle well both the advection-dominated and the diffusive regimes, by their flexible choices of the numerical traces. For instance, upwind-type fluxes allow the schemes to suppress unphysical oscillations in the advection-dominated regime, while in the diffusive regime, HDG-type fluxes enable the techniques of static condensation to save more DOFs. Finally, the DG setting allows a straightforward implementation of  $hp$ -adaptivity, which lies at the center of our memory reduction method.

The development of our method is based on two key observations. The first is that there exists a goal function, and by minimizing this goal function, the discretized error landscape  $\Phi_h(\sigma)$  will be sandwiched by the original landscape  $\Phi(\sigma)$  in the PDE setting (see (4) and (13) for their definitions). This observation motivates the use of this goal function for devising error estimators for AMR. Another observation is that when the numerical error is large, we can use the adjoint solution, which was originally calculated for updating the optical parameters, to also devise the goal-oriented error estimators for doing AMR. This observation suggests that the goal-oriented AMR is especially suitable for the inverse problem setting, since the adjoint solution only needs to be solved for one time but can serve two purposes, namely, the gradient calculation and the error estimator calculation.

The rest of the paper is organized as follows. In Section 2, we consider the forward and the inverse problems of the radiative transfer in the PDE setting, and propose an algorithm of reconstructing the optical coefficients. This algorithm will serve as the ‘‘PDE map’’ for the numerical discretization of the inverse equation. In Section 3, we first introduce the DG discretization for the forward and the adjoint problems, and the gradient calculation. Then, we propose the goal function and devise the goal-oriented error estimators. At the end, we combine the discretization and the error estimator and propose an algorithm of reconstructing the optical coefficient with  $hp$ -mesh adaptivity. Finally, in Section 4, we present numerical experiments to test the performance of our proposed method.

## 2 Forward and inverse problems

In this section, we introduce the forward and the inverse problems of the radiative transfer equation. Then, an algorithm (Algorithm 1) is proposed to reconstruct the optical coefficients in the PDE setting.

### 2.1 Forward problem

Let  $\Omega \subset \mathbb{R}^d$  be a bounded Lipschitz domain, and  $S$  represent the unit sphere in  $\mathbb{R}^d$ . Let  $\partial\Omega$  be the boundary of  $\Omega$ , and  $\mathbf{n}$  the outward-pointing unit normal vector on the boundary. We denote by

$$\Gamma^+ := \{(\mathbf{x}, \mathbf{s}) \in \partial\Omega \times S : \mathbf{n} \cdot \mathbf{s} \geq 0\}, \quad \Gamma^- := \{(\mathbf{x}, \mathbf{s}) \in \partial\Omega \times S : \mathbf{n} \cdot \mathbf{s} \leq 0\},$$

the outflow and the inflow boundary of  $\Omega \times S$ , respectively. We consider the following (time-independent) radiative transfer equation: find  $u \in V$  such that

$$\mathcal{D}(\sigma)[u] := \mathbf{s} \cdot \nabla u + (\sigma_a + \sigma_s)u - \sigma_s \int_S P(\mathbf{s}, \mathbf{s}')u(\mathbf{s}') = f \quad \text{in } \Omega \times S, \quad (1a)$$

$$u = g \quad \text{on } \Gamma^-, \quad (1b)$$

In the above equation,  $u(\mathbf{x}, \mathbf{s})$  is the radiative intensity at spatial location  $\mathbf{x}$  and along the direction  $\mathbf{s} \in S$ , while  $\sigma_a(\mathbf{x}) \geq 0$ ,  $\sigma_s(\mathbf{x}) \geq 0$ ,  $f(\mathbf{x}, \mathbf{s})$ , and  $g(\mathbf{x}, \mathbf{s})$  are the absorption coefficient, the scattering coefficient, the source term, and the inflow radiation, respectively. We assume the scattering phase function  $P(\mathbf{s}, \mathbf{s}')$  to have the form of the Henyey-Greenstein function:

$$P(\mathbf{s}, \mathbf{s}') = \frac{1 - g^2}{c(1 + g^2 - 2g \cos \theta)^{3/2}}, \quad (2)$$

where  $g$  is the asymmetric parameter, and  $c$  is chosen such that  $\int_S P(\mathbf{s}, \mathbf{s}')ds' \equiv 1$ . Here  $V$  is the solution space

$$V := \{u \in L^2(\Omega \times S) : \mathbf{s} \cdot \nabla u \in L^2(\Omega \times S)\}, \quad V^g := \{u \in V : u = g \text{ on } \Gamma^-\},$$

where  $V^g$  is the affine space of  $V$  satisfying the inflow condition (1b). Note that  $V^0$  is a subspace of  $V$ .

To render the presentation more concise, we rewrite (1) into a more compact form. Let  $F = (f, g)^T$ , and  $\mathcal{L}(\sigma) = (\mathcal{D}(\sigma), \gamma^-)^T$ , where  $\gamma^-$  is the trace operator to  $L^2(\Gamma^-)$ , and  $\sigma$  represents the collection of both  $\sigma_e$  and  $\sigma_s$ , namely  $\sigma(\mathbf{x}) := (\sigma_e(\mathbf{x}), \sigma_s(\mathbf{x}))$ . Then the equation (1) transforms into

$$\mathcal{L}(\sigma)[u] = F. \quad (3)$$

In the equation (1) (or (3)), we use the square bracket  $[\cdot]$  to indicate that  $\mathcal{D}(\sigma)$  (or  $\mathcal{L}(\sigma)$ ) depends linearly on  $u$ ; the round bracket suggests that  $\mathcal{D}$  (or  $\mathcal{L}$ ) can depend non-linearly on  $\sigma$ . In the setting of the forward problem, the optical property represented by  $\sigma$  and the source/boundary term  $F$  are given as input data. Then the task is to solve the equation (1) (or (3)) for the radiative intensity  $u$ . We refer to [16] and the references therein for the well-posedness of the forward problem.

## 2.2 Inverse problem

In the setting of the inverse problem, we are given certain measurements of the solution, here denoted as  $\mathcal{M}u$ , while the task is to infer the distribution of the optical property  $\sigma$  or the source term  $f$ . The former task fits into the field of optical tomography, while the latter one is usually referred to as inverse source problems. We refer to [2, 6] for a review on these topics.

In this paper, we focus on the optical tomography problem. To be more specific, we consider the task of reconstructing the optical property of the medium  $\sigma$ , based on a set of tests determined by the source terms  $F^i$  with  $i = 1, \dots, N_t$ , and a set of measurements on the outflow boundary  $\Gamma^+$ , here denoted as  $\mathcal{M}^j$  with  $j = 1, \dots, N_m$ . This problem can be

formulated as an optimization problem. Suppose  $\sigma^*$  is the target optical coefficient we would like to reconstruct. We define  $y^{ij} = \mathcal{M}^j u^i = \mathcal{M}^j(\mathcal{L}(\sigma^*))^{-1}[F^i]$  as the measurement data. For many applications, the data are polluted by noise. In such a case we define  $y^{ij} := \mathcal{M}^j u^i + \delta_{ij}$  where  $\delta_{ij}$  represent the noise. The optical tomography problem becomes

$$\min_{\sigma} \Phi(\sigma) := \sum_{i=1}^{N_t} \sum_{j=1}^{N_m} w_{ij} |\mathcal{M}^j(\mathcal{L}(\sigma))^{-1}[F^i] - y^{ij}|^2 + \alpha \mathcal{R}(\sigma), \quad (4)$$

where  $w_{ij}$  are the weights associated to each test-measurement signal  $y^{ij}$ . A typical choice is  $w^{ij} = \frac{1}{|y^{ij}|^2}$ ; see [32]. Here,  $\alpha \mathcal{R}(\sigma)$  is the regularization term and  $\alpha$  is the regularization parameter. Some widely used examples of regularization include the  $L^2$ -regularization, for which we choose  $\mathcal{R}(\sigma) = \|\sigma - \bar{\sigma}\|_{L^2(\Omega)}$  (where  $\bar{\sigma}$  is a spatial average), or  $H^1$ -regularization, for which we choose  $\mathcal{R}(\sigma) = \|\nabla \sigma\|_{L^2(\Omega)}$ . For more on regularization techniques and the choices of regularization parameters, we refer to the monograph [17] and the references therein.

Note that  $\Phi(\sigma)$  is nonlinear and there is no explicit expression for  $\Phi(\sigma)$  in general cases. To proceed, we transform (4) into a PDE-constrained optimization problem:

$$\min_{\sigma, u} \Phi(\sigma, u) := \sum_{i=1}^{N_t} \sum_{j=1}^{N_m} w_{ij} |\mathcal{M}^j u^i - y^{ij}|^2 + \alpha \mathcal{R}(\sigma), \quad (5a)$$

$$\text{subject to } \mathcal{L}(\sigma)[u^i] = F^i \quad \text{for } i = 1, \dots, N_t. \quad (5b)$$

To solve the problem, we introduce the Lagrangian

$$\Psi(\sigma, u, v) = \Phi(\sigma, u) + \sum_i \langle \mathcal{L}(\sigma)[u^i] - F^i, v^i \rangle, \quad (6)$$

where  $v^i$  is the Lagrange multiplier. By the Karush-Kuhn-Tucker (KKT) necessity condition [38], the solution of (5a), if exists, is also the first-order stationary point of  $\Psi$  in (6). By taking Fréchet derivatives of (6), we obtain

$$\frac{\delta \Psi}{\delta \sigma}[\delta \sigma] = \alpha \mathcal{R}'(\sigma)[\delta \sigma] + \sum_i \langle \mathcal{L}'(\sigma)[\delta \sigma][u^i], v^i \rangle, \quad (7a)$$

$$\frac{\delta \Psi}{\delta u^i}[\delta u^i] = 2 \sum_{j=1}^{N_m} w_{ij} (\mathcal{M}^j u^i - y^{ij}) \mathcal{M}^j \delta u^i + \langle \delta u^i, \mathcal{L}(\sigma)^*[v^i] \rangle \quad i = 1, \dots, N_t, \quad (7b)$$

$$\frac{\delta \Psi}{\delta v^i}[\delta v^i] = \langle \mathcal{L}(\sigma)[u^i] - F^i, \delta v^i \rangle \quad i = 1, \dots, N_t, \quad (7c)$$

where  $\langle \cdot, \cdot \rangle$  is the trial-test bracket, manifested as the summation of the  $L^2$ -inner product on  $\Omega \times S$  and  $\Gamma^-$ , and  $\mathcal{L}(\sigma)^*$  is the adjoint operator of  $\mathcal{L}(\sigma)$ .

Now, the optimization problem reduces to finding the stationary point  $(\sigma^*, u^*, v^*)$  such that equations (7) equal 0. Based on (7), we propose the following Algorithm 1 of finding  $\sigma^*$ . In the algorithm, the forward and the adjoint problems associated with (7c) and (7b) are first solved to obtain  $u^i$  and  $v^i$ , which are then used to update  $\sigma$  by an updating direction  $\delta \sigma^k$  and the step length  $\gamma > 0$ . The step length  $\gamma$  can be determined by a line search [31]. The updating direction can be simply chosen as the gradient given by (7a), or by Quasi-Newton method such as the BFGS method. We refer to [31] more details.

---

**Algorithm 1** Reconstructing optical coefficients - PDE setting

---

- 1: Set the maximal iteration count  $N_{\max\_iter}$  and the error tolerance  $\epsilon > 0$
  - 2: Initialize  $\sigma^0$
  - 3: Set  $k = 0$
  - 4: **while**  $k \leq N_{\max\_iter}$  **do**
  - 5:     Using  $\sigma^k$ , solve (7c) = 0 to obtain  $u^i$
  - 6:     Using  $\sigma^k$  and  $u^i$ , solve (7b) = 0 to obtain  $v^i$
  - 7:     Using  $\sigma^k$ ,  $u^i$ ,  $v^i$ , and (7a) to obtain the updating direction  $\delta\sigma^k$
  - 8:     Update the optical coefficients by  $\sigma^{k+1} = \sigma^k - \gamma \delta\sigma^k$ , where  $\gamma$  is determined by a line search
  - 9:     Update the iteration count:  $k = k + 1$
  - 10:    **if**  $|\Phi(\sigma^k) - \Phi(\sigma^{k-1})|/|\Phi(\sigma^k)| < \epsilon$  **then**
  - 11:        Break the loop
  - 12:    **end if**
  - 13: **end while**
- 

### 3 Numerical discretization

In this section, we introduce the numerical approximations to the inverse problem (4), manifested by using the discontinuous Galerkin spectral element (DGSE) methods to discretize the forward problem (7c)=0, the adjoint problem (7b)=0, and the gradient calculation (7a). Then, we propose a goal-oriented error estimator which can be easily calculated based on the forward and the adjoint numerical solution. Finally, based on the error estimator, we introduce an  $hp$ -adaptive mesh refinement algorithm to solve the inverse radiative transfer problem numerically.

#### 3.1 Discontinuous Galerkin

##### 3.1.1 Approximation spaces

We begin by considering the approximation spaces. For simplicity, we restrict ourselves to consider only a 2D domain  $\Omega$  discretized by rectangular meshes  $\mathcal{T}_h$ . The generalization to 3D and polyhedral meshes is straightforward under the DG setting. Since any  $K \in \mathcal{T}_h$  is rectangular, we can write  $K = K^x \times K^y$ . We require that the mesh  $\mathcal{T}_h$  satisfies the 1-irregular condition, i.e., for any element  $K \in \mathcal{T}_h$ , there are at most 2 neighbour elements connecting to  $K$  through each edge. Here  $h$  is both an index for the triangulations  $\mathcal{T}_h$  and the mesh-size, which is defined as  $h := \max_{K \in \mathcal{T}_h} h_K$ . For each  $K \in \mathcal{T}_h$ , we denote by  $\mathcal{T}_h^{a,K}$  the triangulation of the unit sphere  $S$ . We also let  $\mathcal{F}_K$  be the collection of the faces of  $K$ . Now we introduce the approximation space:

$$V_h := \left\{ \sum_{K \in \mathcal{T}_h} \sum_{K^a \in \mathcal{T}_h^{a,K}} \sum_{i,j,k} u_{i,j,k} \phi_i^{K^x}(x) \phi_j^{K^y}(y) \phi_k^{K^a}(\theta) : \phi_i^{K^*} \Big|_{i=1}^{p_{K^*}+1} \text{ are a set of polynomial bases on } K^* \right\},$$

where  $p_{K^*}$  represents the polynomial degree and  $\star \in \{x, y, a\}$ . Here we choose  $\phi_i^{[a,b]}(x) = \phi_i \circ \mathcal{F}_{[a,b]}^{-1}(x)$  where  $\phi_i$  supported on  $[-1, 1]$  is the Lagrange polynomial basis associated

to the  $i$ -th Gauss-Legendre-Lobatto quadrature points, and  $\mathcal{F}_{[a,b]}(\hat{x}) = \frac{b-a}{2}(\hat{x} + 1) + a$  is the push-forward map. This choice of basis gives us spectral element methods. Note that the space  $V_h$  is completely determined by the mesh discretization  $\mathcal{T}_h$  and the polynomial degrees  $p_{K^a}$ ,  $p_{K^y}$ , and  $p_{K^a}$  which are associated to each spatial-angular element  $K \times K^a$ . We will sometimes write  $\mathcal{T}_{hp}$  to denote the mesh  $\mathcal{T}_h$  encoded with also the polynomial degree information. Finally, to incorporate the boundary condition, we introduce

$$V_h^g := \{u_h|_{(\Omega^\circ \times S) \cup \Gamma^+} \in V_h : u_h|_{\Gamma^-} = g\}.$$

As a special case,  $V_h^0$  is a subspace of  $V_h$  with zero inflow radiation.

For notation conciseness, we define

$$(f, g)_{K \times K^a} := \int_K \int_{K^a} f(\mathbf{x}, \mathbf{s}) g(\mathbf{x}, \mathbf{s}) ds dx$$

as the  $L^2$  inner product on  $K \times K^a$ , and we denote the associated norm as  $\|f\|_{K \times K^a} := (f, f)_{K \times K^a}^{1/2}$ . Also, when there are two elements  $K^+$  and  $K^-$  which share a common face  $F$ , we introduce the jump notation  $[[u_h]] := u_h^+ \mathbf{n}^+ + u_h^- \mathbf{n}^-$ . If the face  $F$  is an exterior face, namely, if  $F \subset \partial\Omega$ , then  $[[u_h]] := 0$ .

### 3.1.2 DG for the forward and adjoint problems

Now we are ready to introduce the DG numerical discretization for the radiative transfer equations. We first define the following bilinear form:

$$\begin{aligned} a_h(u_h, v_h) := & \sum_{K \in \mathcal{T}_h} \sum_{K^a \in \mathcal{T}_h^{a,K}} \left( \int_{\partial K} \int_{K^a} (\mathbf{s} \cdot \mathbf{n}) \hat{u}_h v_h - \int_K \int_{K^a} u_h \mathbf{s} \cdot \nabla v_h \right. \\ & \left. + \int_K \int_{K^a} (\sigma_a + \sigma_s) u_h v_h - \int_K \sigma_s \int_{K^a} \int_S P(\mathbf{s}, \mathbf{s}') u_h(\mathbf{s}') ds' v_h(\mathbf{s}) ds \right), \end{aligned} \quad (8a)$$

where we take the upwind numerical trace  $\hat{u}_h$ :

$$\hat{u}_h|_{\partial K} := \begin{cases} u_h^{\text{nbr}} & \text{if } \mathbf{s} \cdot \mathbf{n}_F < 0 \text{ and } F \text{ is an interior face,} \\ u_h|_{\partial K} & \text{otherwise,} \end{cases} \quad (8b)$$

and  $u_h^{\text{nbr}}$  is the restriction of  $u_h$  from the neighbour elements of  $K$  on  $F$ .

The DG approximation for the  $i$ -th test of the forward problem (7c) reads as

$$\text{find } u_h^i \in V_h^{g^i} \quad \text{such that} \quad a_h(u_h^i, v_*) = \int_\Omega \int_S f^i v_* \quad \forall v_* \in V_h^0. \quad (9)$$

On the other hand, the DG approximation for the  $i$ -th test of the adjoint problem (7b) reads as

$$\text{find } v_h^i \in V_h^0 \quad \text{such that} \quad a_h(u_*, v_h^i) = M^i(u_h^i)[u_*^*] \quad \forall u_* \in V_h^0, \quad (10)$$

where

$$M^i(u_h^i)[u^*] := 2 \sum_j w_{ij} (y^{ij} - \mathcal{M}^j u_h^i) \mathcal{M}^j u^* \quad (11)$$

are the measurements. For different tests, we choose different underlying  $hp$ -meshes. We denote by  $\mathcal{T}_{hp}^i$  and  $\mathcal{T}_{hp}^{*-i}$  as the underlying  $hp$ -mesh for the  $i$ -th test of the forward and the adjoint problem, respectively. We shall assume that  $\mathcal{T}_{hp}^{*-i}$  is obtained by refining  $\mathcal{T}_{hp}^i$ . For instance,  $\mathcal{T}_{hp}^{*-i}$  can be a one-level  $h$  or  $p$ -refinement of  $\mathcal{T}_{hp}^i$ . This extra refinement for the adjoint meshes will be essential for devising the error estimators for the inverse problem, and will be discussed in more detail in Section 3.2.2.

### 3.1.3 Gradient calculation in the discrete setting

For the rest of this subsection, we consider how to use the forward and the adjoint numerical solutions, namely  $u_h^i$  and  $v_h^i$ , to calculate the gradient and update the optical parameter in the DG setting. Namely, we consider the DG discretization for the gradient calculation (7a).

We begin by introducing the approximation spaces for the optical parameter  $\sigma$ . Since we use the DG spectral element method as our radiative transfer equations solver, it is natural to choose the discretized optical parameter  $\sigma_h$  to live in a DG spectral element space. We shall denote by  $\mathcal{T}_{hp}^\sigma$  as the underlying mesh for this space, and by  $\Sigma_h$ , or  $\Sigma_h(\mathcal{T}_{hp}^\sigma)$  as the space.

Here we adopt the parametric reconstruction method [32] as a regularization to relieve the ill-posedness of the original inverse problem (4). This means that we shall set  $\mathcal{T}_{hp}^\sigma$  to live in a coarse mesh with large mesh-size. In the extreme case, we choose  $\mathcal{T}_{hp}^\sigma$  to be the one-element mesh covering the whole domain  $\Omega$ . Then, the polynomial degree  $p$  can be adjusted in the sense that a lower degree  $p$  represents a stronger regularization effect (restriction to the low-frequency modes) and vice versa.

Now, the discretized version of (7a) becomes

$$\frac{\delta \Psi_h}{\delta \sigma_h} [\delta \sigma_h] = \alpha \mathcal{R}'(\sigma_h) [\delta \sigma_h] + \sum_{i=1}^{N_{\text{tst}}} \sum_{K \in \mathcal{T}_h^i} \sum_{K^a \in \mathcal{T}_h^{a,K}} ((u_h^i v_h^i, \delta \sigma_{a,h})_{K \times K^a} + (u_h^i v_h^i - u_h^{i,s} v_h^i, \delta \sigma_{s,h})_{K \times K^a}), \quad (12)$$

for any test function  $\delta \sigma_h \in \Sigma_h$ , where  $u_h^{i,s} := \int_S P(\mathbf{s}, \mathbf{s}') u_h^i(\mathbf{s}') ds'$ , and  $u_h^i$  and  $v_h^i$  are the solution of (9) and (10), respectively. By letting the test function  $\delta \sigma_h$  go over all the basis polynomials of  $\Sigma_h$ , we obtain a vector as the updating gradient for  $\sigma_h$ .

We conclude Section 3.1 by the following Algorithm 2.

## 3.2 Goal-oriented error estimator

### 3.2.1 The goal function

The numerical solver (9) and (10) can be rewritten in a compact form as follows:

$$u_h^i = \text{DGSE}(\mathcal{T}_{hp}^i, \sigma)[F^i], \quad v_h^i = \text{DGSE}^*(\mathcal{T}_{hp}^i, \sigma)[M^i(u_h^i)],$$

---

**Algorithm 2** Reconstructing optical coefficients - DG discretization
 

---

- 1: Set the mesh for the optical parameter -  $\mathcal{T}_{hp}^\sigma$
  - 2: Set the mesh for each test  $i$  -  $\mathcal{T}_{hp}^i$  for the forward and  $\mathcal{T}_{hp}^{\star-i}$  for the adjoint problem
  - 3: Set the maximal iteration count  $N_{\max\_iter}$  and the error tolerance  $\epsilon > 0$
  - 4: Initialize  $\sigma_h^0$
  - 5: Set  $k = 0$
  - 6: **while**  $k \leq N_{\max\_iter}$  **do**
  - 7: Using  $\sigma_h^k$ , solve (9) to obtain  $u_h^i$
  - 8: Using  $\sigma_h^k$  and  $u_h^i$ , solve (10) to obtain  $v_h^i$
  - 9: Using  $\sigma_h^k$ ,  $u_h^i$ ,  $v_h^i$ , and (12) to obtain the updating direction  $\delta\sigma_h^k$
  - 10: Update the optical coefficients by  $\sigma_h^{k+1} = \sigma_h^k - \gamma \delta\sigma_h^k$ , where  $\gamma$  is determined by a line search
  - 11: Update the iteration count:  $k = k + 1$
  - 12: **if**  $|\Phi_h(\sigma_h^k) - \Phi_h(\sigma_h^{k-1})|/|\Phi_h(\sigma_h^k)| < \epsilon$  ( $\Phi_h$  defined in (13)) **then**
  - 13: Break the loop
  - 14: **end if**
  - 15: **end while**
- 

where DGSE and DGSE\* represent the forward and the adjoint solver, respectively. Using this notation, the discretized inverse problem (4) becomes

$$\min_{\sigma, \mathcal{T}_{hp}^I} \Phi_h(\sigma, \mathcal{T}_{hp}^I) := \sum_{i=1}^{N_t} \sum_{j=1}^{N_m} w_{ij} |\mathcal{M}^j \text{DGSE}(\mathcal{T}_{hp}^i, \sigma)[F^i] - y^{ij}|^2 + \alpha \mathcal{R}(\sigma), \quad (13)$$

where  $I := \{1, \dots, N_t\}$  is the index set for all tests. Note that the discretized error landscape functional  $\Phi_h$  depends both on the optical property  $\sigma$  and the set of the  $hp$ -meshes  $\mathcal{T}_{hp}^I$ . If we denote

$$\begin{aligned} \text{Err}_h(\sigma) &:= \sum_{i=1}^{N_t} \text{Err}_h^i(\sigma) := \sum_{i=1}^{N_t} \sum_{j=1}^{N_m} w_{ij} |\delta_h^{ij}(\sigma)|^2, \\ \delta_h^{ij}(\sigma) &:= \mathcal{M}^j (\mathcal{L}(\sigma))^{-1} [F^i] - \mathcal{M}^j \text{DGSE}(\mathcal{T}_{hp}^i, \sigma)[F^i], \end{aligned}$$

then the error landscape  $\Phi(\sigma)$  can be related to its discretized version  $\Phi_h(\sigma, \mathcal{T}_{hp})$  as follows:

$$\Phi_h(\sigma, \mathcal{T}_{hp}^I) \leq 2(\Phi(\sigma) + \text{Err}_h(\sigma)), \quad \Phi(\sigma) \leq 2(\Phi_h(\sigma, \mathcal{T}_{hp}^I) + \text{Err}_h(\sigma)), \quad \forall \sigma, \quad \forall \mathcal{T}_{hp}^I.$$

As a result, when  $\text{Err}_h(\sigma) \rightarrow 0$ , we have

$$\frac{1}{2}\Phi(\sigma) \leq \Phi_h(\sigma, \mathcal{T}_{hp}^I) \leq 2\Phi(\sigma).$$

The above estimate suggests that, when the error  $\text{Err}_h(\sigma)$  goes to zero, the minimizer of  $\Phi(\sigma)$  also minimizes  $\Phi(\sigma, \mathcal{T}_{hp}^I)$  and vice versa. This observation suggests that we should aim for devising numerical methods such that  $\text{Err}_h(\sigma)$  is reduced in the most efficient way. We next explain how this can be achieved in more detail.

Toward this aim, for simplicity, write  $u^i(\sigma) := (\mathcal{L}(\sigma))^{-1}[F^i]$  and  $u_h^i(\sigma) := \text{DGSE}(\mathcal{T}_{hp}^i, \sigma)[F^i]$ . Recall that we denote by  $\sigma^*$  the minimizer of  $\Phi(\sigma)$  so we have  $y^{ij} = \mathcal{M}^j u^i(\sigma^*)$  by definition. Now, for each test  $i$ , we introduce the functional  $J^i(\sigma)$ :

$$J^i(\sigma)[\varphi] := 2 \sum_j w_{ij} \mathcal{M}^j (u^i(\sigma) - u_h^i(\sigma)) \mathcal{M}^j \varphi.$$

Then, it holds that

$$J^i(\sigma)[u^i(\sigma) - u_h^i(\sigma)] = 2 \sum_{j=1}^{N_m} w_{ij} |\delta_h^{ij}(\sigma)|^2 = 2 \text{Err}_h^i(\sigma).$$

The above identity suggests that we can use the functional  $J^i(\sigma)$  to devise error estimators to guide the mesh-refinement of  $\mathcal{T}_{hp}^i$ . This approach fits into the well-known framework of goal-oriented error control and adaptive mesh refinement; see [9]. To be more specific, the error estimators are obtained by solving the dual equation

$$a_h(u_*, z_h^i) = J^i(\sigma)[u_*] \quad \forall u_* \in V_h^0,$$

which coincides with (10) except the right-hand-side term is different. Now note that

$$\begin{aligned} J^i(\sigma)[\varphi] &= 2 \sum_j w_{ij} \mathcal{M}^j (u^i(\sigma) - u_h^i(\sigma)) \mathcal{M}^j \varphi \\ &= M^i(u_h^i(\sigma))[\varphi] + 2 \sum_j w_{ij} (\mathcal{M}^j u^i(\sigma) - y^{ij}) \mathcal{M}^j \varphi. \end{aligned}$$

Therefore, when  $|y^{ij} - \mathcal{M}^j u^i(\sigma)| \ll |y^{ij} - \mathcal{M}^j u_h^i(\sigma)|$ , or namely, when the numerical error dominates the error caused by optical parameters, we have

$$J^i(\sigma)[\varphi] \approx M^i(u_h^i(\sigma))[\varphi].$$

Thus, in this case, we can use  $M^i(u_h^i(\sigma))$  as an alternative of  $J^i(\sigma)$  to devise error estimators to guide the mesh refinement of  $\mathcal{T}_{hp}^i$ . As a result, we have  $z_h^i \approx v_h^i$ . This suggests that we do not need to solve another adjoint equation for  $z_h^i$  but can simply use  $v_h^i$  to devise the error estimators.

### 3.2.2 The error estimators

For the rest of this subsection, we show how a goal-oriented error estimator can be devised based on a given functional  $J$  in the DG setting. Namely, we would like to control  $|J(u - u_h)|$  by a posteriori error estimators. To make the presentation concise, we shall consider the forward problem (1), and its DG discretization formulated as follows: find  $u_h \in V_h^g$  such that

$$a_h(u_h, v_*) = \int_{\Omega} \int_S f v_* \quad \forall v_* \in V_h^0. \quad (14)$$

Before we present the main result, we first introduce two lemmas on the consistency and the Galerkin orthogonality of the DG bilinear form (8a).

**Lemma 3.1** (Consistency of the DG bilinear form). *Suppose  $u$  solves (1). Then*

$$a_h(u, v_*) = \int_{\Omega} \int_S f v_* \quad \forall v_* \in V. \quad (15)$$

*Proof.* By multiplying (1a) with  $v_*$  and integrating on  $K \times K^a$ , we obtain

$$\begin{aligned} & \int_{K^a} \int_{\partial K} (\mathbf{s} \cdot \mathbf{n}) u v_* - \int_{K^a} \int_K \mathbf{s} u \cdot \nabla v_* + \int_{K^a} \int_K (\sigma_a + \sigma_s) u v_* \\ & - \int_{K^a} \int_K \sigma_s \int_S P(\mathbf{s}, \mathbf{s}') u(\mathbf{s}') d\mathbf{s}' v_*(\mathbf{s}) d\mathbf{s} = \int_{K^a} \int_K f v_*. \end{aligned}$$

Since  $u$  is the exact solution, we have  $(\mathbf{s} \cdot \mathbf{n})u = (\mathbf{s} \cdot \mathbf{n})\hat{u}$  by the consistency of the numerical trace, see (8b). Then we take summation for all  $K \in \mathcal{T}_h$  and  $K^a \in \mathcal{T}_h^{a,K}$ . This completes the proof.  $\square$

**Lemma 3.2** (Galerkin orthogonality). *Suppose  $u$  solves (1) and  $u_h$  solves (14). Then*

$$a_h(u_h - u, v_*) = 0 \quad \forall v_* \in V_h^0. \quad (16)$$

*Proof.* We simply choose  $v_* \in V_h^0$  in (15) and then take the difference between (14) and (15). This completes the proof.  $\square$

Consider the following dual equation: find  $z \in V^0$  such that

$$a_h(v, z) = J(v) \quad \forall v \in V^0, \quad (17)$$

Then we define

$$\eta_{K \times K^a} := \rho_{K \times K^a}^1 w_{K \times K^a}^1 + \rho_{K \times K^a}^2 w_{K \times K^a}^2, \quad \eta_K := \sum_{K^a \in \mathcal{T}_h^{a,K}} \eta_{K \times K^a}, \quad (18a)$$

where

$$\rho_{K \times K^a}^1 := \|f - \nabla u_h \cdot \mathbf{s} - (\sigma_a + \sigma_s) u_h - \sigma_s \int_S P(\mathbf{s}, \mathbf{s}') u_h(\mathbf{s}')\|_{K \times K^a}, \quad (18b)$$

$$w_{K \times K^a}^1 := \|z - \varphi_h\|_{K \times K^a}, \quad (18c)$$

$$\rho_{K \times K^a}^2 := \|(\mathbf{s} \cdot \mathbf{n})[[u_h]]\|_{\partial K \times K^a}, \quad (18d)$$

$$w_{K \times K^a}^2 := \|z - \varphi_h\|_{\partial K \times K^a}. \quad (18e)$$

In the above equations,  $\varphi_h$  can be any function in  $V_h^0$ .

**Proposition 3.1.** *Let  $u \in V^g$  be the solution of (1),  $u_h \in V_h^g$  solves (14), and  $z \in V^0$  be the solution of the dual problem (17). Then, we have*

$$|J(u - u_h)| \leq \sum_{K \in \mathcal{T}_h} \sum_{K^a \in \mathcal{T}_h^{a,K}} (\rho_K^1 w_K^1 + \rho_K^2 w_K^2).$$

*Proof.* Since  $u_h \in V_h^g$  and  $u \in V^g$ , we know  $(u - u_h) \in V^0$ . Thus, we can let  $v = u - u_h$  in (17), which gives us

$$J(u - u_h) = a_h(u - u_h, z) = a_h(u - u_h, z - \varphi_h) \quad \forall \varphi_h \in V_h^0,$$

where the Galerkin orthogonality identity (16) is used for the second equal sign. Then, by the definition of the bilinear form (8a) and its consistency property (15), we can proceed as follows:

$$\begin{aligned} |J(u - u_h)| &= |a_h(u, z - \varphi_h) - a_h(u_h, z - \varphi_h)| \\ &= \left| \sum_{K \in \mathcal{T}_h} \sum_{K^a \in \mathcal{T}_h^{a,K}} \int_K \int_{K^a} f(z - \varphi_h) \right. \\ &\quad - \sum_{K \in \mathcal{T}_h} \sum_{K^a \in \mathcal{T}_h^{a,K}} \left( \int_{\partial K} \int_{K^a} (\mathbf{s} \cdot \mathbf{n})(\hat{u}_h - u_h)(z - \varphi_h) + \int_K \int_{K^a} (\nabla u_h \cdot \mathbf{s})(z - \varphi_h) \right. \\ &\quad \left. \left. + \int_K \int_{K^a} (\sigma_a + \sigma_s)u_h(z - \varphi_h) - \int_K \sigma_s \int_{K^a} \int_S P(\mathbf{s}, \mathbf{s}')u(\mathbf{s}')ds'(z - \varphi_h)(\mathbf{s})ds \right) \right| \\ &\leq \sum_{K \in \mathcal{T}_h} \sum_{K^a \in \mathcal{T}_h^{a,K}} \left| \left( f - \nabla u_h \cdot \mathbf{s} - (\sigma_a + \sigma_s)u_h + \sigma_s \int_S P(\mathbf{s}, \mathbf{s}')u(\mathbf{s}')ds', z - \varphi_h \right)_{K \times K^a} \right| \\ &\quad + \sum_{K \in \mathcal{T}_h} \sum_{K^a \in \mathcal{T}_h^{a,K}} |((\mathbf{s} \cdot \mathbf{n})(u_h - \hat{u}_h), z - \varphi_h)_{\partial K \times K^a}|. \end{aligned}$$

Now, note that

$$\begin{aligned} |((\mathbf{s} \cdot \mathbf{n})(u_h - \hat{u}_h), z - \varphi_h)_{\partial K \times K^a}| &= \left| \sum_{F \in \mathcal{F}_K} ((\mathbf{s} \cdot \mathbf{n})(u_h - \hat{u}_h), z - \varphi_h)_{F \times K^a} \right| \\ &\leq \sum_{F \in \mathcal{F}_K} \|(\mathbf{s} \cdot \mathbf{n})(u_h - u_h^{\text{nbr}})\|_{F \times K^a} \|z - \varphi_h\|_{F \times K^a} \\ &\leq \sum_{F \in \mathcal{F}_K} \|(\mathbf{s} \cdot \mathbf{n})[[u_h]]\|_{F \times K^a} \|z - \varphi_h\|_{F \times K^a}. \end{aligned}$$

This completes the proof.  $\square$

Proposition 3.1 suggests that we need to solve a dual equation (17) to calculate the error estimators. Since a direct solve for  $z$  is not feasible, we calculate an approximate solution to  $z$  by  $z_h$ . Here  $z_h$  should live in a more refined space than the space for  $u_h$ . Otherwise, we can choose  $\varphi_h = z_h$  and this will render the error estimators vanishing. To be more specific, we seek  $z_h \in \tilde{V}_h^0$  such that

$$a_h(v_*, z_h) = J(v_*) \quad v_* \in \tilde{V}_h^0.$$

Here  $\tilde{V}_h$  is a more refined space than  $V_h$ . For instance, we can choose  $\tilde{V}_h$  to be an  $h$ -refined or  $p$ -refined version of  $V_h$ . In this paper, we shall use the  $p$ -refined version since it is

straightforward to implement. Once  $z_h$  is calculated, we choose  $\varphi_h = \Pi_V z_h$  where  $\Pi_V$  is the  $L^2$  projection onto the space  $V_h$ . Namely, the weighting terms become

$$w_K^1 = \|z_h - \Pi_V z_h\|_{K \times K^a}, \quad w_K^2 = \|z_h - \Pi_V z_h\|_{\partial K \times K^a}.$$

For the rest of this section, we use the error estimators derived here to design an  $hp$ -adaptive mesh refinement method to solve the inverse radiative transfer problem (4).

### 3.3 Goal-oriented mesh adaptation for the inverse problem

Up to this point we have introduced the forward solver (9), the adjoint solver (10), the gradient calculation (12), and the error estimators (18). Next, we show how they can be combined to solve the inverse problem (13) with mesh adaptation. We shall first go over the basic elements of the  $hp$ -adaptive mesh refinement and then propose the full algorithm.

#### 3.3.1 $hp$ -adaptive mesh refinement

Given a mesh  $\mathcal{T}_{hp}$ , we can refine it based on the error estimators introduced in (18). For  $h$ -refinement only, (18) is sufficient for an adaptive mesh refinement algorithm. For  $hp$ -adaptive mesh refinement ( $hp$ -AMR), we will need an additional smoothness estimators to complete the algorithm. Since we use a more refined mesh for the duality solution  $v_h$ , the dominating error comes from  $u_h$ . Hence, here we use  $u_h$  for devising the smoothness estimator.

We consider the smoothness estimators proposed in [21]. These estimators are obtained by examining the decaying pattern of the coefficients obtained from a Legendre expansion of the numerical solution. To be more specific, given the forward solution  $u_h$ , we first calculate its mean intensity  $u_h^{\text{ml}} := \int_S u_h$ . Then, for any element  $K \in \mathcal{T}_h$ , we calculate the Legendre expansion coefficients of  $u_h^{\text{ml}}$  on  $K$ , here denoted as  $a_{i,j}^K$  with  $i = 0, \dots, p_x$  and  $j = 0, \dots, p_y$ , where  $p_x$  and  $p_y$  are the polynomial degrees of the tensor-product basis function on  $K$  in the  $x$  and the  $y$  directions, respectively. Then, the smoothness estimator is defined as follows:

$$\eta_K^s := \frac{1}{2} \left( \frac{\log\left(\frac{2p_x+1}{2|a_{p_x}^x|^2}\right)}{2 \log(p_x)} + \frac{\log\left(\frac{2p_y+1}{2|a_{p_y}^y|^2}\right)}{2 \log(p_y)} \right), \quad \text{where} \quad (a_i^x)^2 := \sum_{j=0}^{p_y} |a_{i,j}^K|^2 \frac{2}{2j+1}, \quad (a_j^y)^2 := \sum_{i=0}^{p_x} |a_{i,j}^K|^2 \frac{2}{2i+1}. \quad (19)$$

With the error estimator  $\eta_K$  (defined in (18)) and the smoothness estimator  $\eta_K^s$  (defined in (19)), we propose the following Algorithm 3 to refine a given mesh  $\mathcal{T}_{hp}$ .

---

**Algorithm 3** *hp*-adaptive mesh refinement

---

- 1: Sort all element  $K \in \mathcal{T}_h$  in increasing order according to the error estimator  $\eta_K$  defined in (18).
  - 2: Mark the largest  $r_{\text{ref}}$  percentage of the elements for refinement –  $\mathcal{T}_h^{\text{mark}}$ .
  - 3: For all  $K \in \mathcal{T}_h^{\text{mark}}$ , calculate the smoothness estimators  $\eta_K^s$  defined in (19).
  - 4: For all  $K \in \mathcal{T}_h^{\text{mark}}$ , we perform the following refinement
  - 5: **if**  $p_K + 1 < \eta_K^s - \frac{1}{2}$  **then**
  - 6:     Perform  $p$ -refinement for  $K$ .
  - 7: **else**
  - 8:     Perform  $h$ -refinement for  $K$ .
  - 9: **end if**
- 

Note that for each test  $i$ , the error estimators  $\eta_K^i$  are computed based on the forward solution  $u_h^i$  and the adjoint solution  $v_h^i$ ; see equations (18), where  $z$  is replaced by  $v_h^i$ .

### 3.3.2 The full algorithm

We conclude this section by proposing the Algorithm 4 of solving the discretized inverse problem (13) with mesh adaptation.

---

**Algorithm 4** Reconstructing optical coefficients - DG discretization with *hp*-adaptive mesh refinement

---

- 1: Set  $\mathcal{T}_{hp}^\sigma$ , the mesh for the optical parameter
  - 2: Set the maximal iteration count  $N_{\text{max\_iter}}$ , the maximal DOFs count  $N_{\text{dofs}}$ , and the error tolerance  $\epsilon > 0$ .
  - 3: For each test  $i$ , initialize  $\mathcal{T}_{hp}^i$ , the mesh for solving the forward and the adjoint problems
  - 4: Initialize the optical parameter  $\sigma_h^0$ .
  - 5: Set  $k = 0$
  - 6: **while**  $k \leq N_{\text{max\_iter}}$  and DOFs  $\leq N_{\text{dofs}}$  **do**
  - 7:     Using  $\sigma_h^k$ , solve (9) to obtain  $u_h^i$
  - 8:     Using  $\sigma_h^k$  and  $u_h^i$ , solve (10) to obtain  $v_h^i$
  - 9:     Using  $\sigma_h^k$ ,  $u_h^i$ ,  $v_h^i$ , and (12) to obtain the updating direction  $\delta\sigma_h^k$
  - 10:    Update the optical coefficients by  $\sigma_h^{k+1} = \sigma_h^k - \gamma \delta\sigma_h^k$ , where  $\gamma$  is determined by a line search
  - 11:    Update the iteration counting:  $k = k + 1$
  - 12:    **if**  $|\Phi_h(\sigma_h^k) - \Phi_h(\sigma_h^{k-1})|/|\Phi_h(\sigma_h^k)| < \epsilon$  ( $\Phi_h$  defined in (13)) **then**
  - 13:       For each test  $i$ , refine the mesh  $\mathcal{T}_{hp}^i$  by Algorithm 3
  - 14:       Reinitialize the optical parameter  $\sigma_h^k = \lambda\sigma_h^k + (1 - \lambda)\sigma_h^0$
  - 15:    **end if**
  - 16: **end while**
- 

Note that previously in Algorithm 2, we stop the iteration when the relative difference  $|\Phi_h(\sigma_h^k) - \Phi_h(\sigma_h^{k-1})|/|\Phi_h(\sigma_h^k)| < \epsilon$ . Here in Algorithm 4, we apply mesh-refinement instead of stopping the iteration. This procedure is repeatedly applied until the maximal DOFs count  $N_{\text{dofs}}$  is reached. This can prevent the algorithm from refining indefinitely. The constant

$N_{\text{dofs}}$  can be chosen according to the memory-occupation restriction or computational time restriction. The parameter  $\lambda$  determines how much information will be inherited from the optical parameter recovered in the previous mesh setting. If the previous mesh setting can provide a relatively good approximation of  $\sigma$ , then we can increase  $\lambda$  for a better initial guess. Otherwise, we choose  $\lambda = 0$  to avoid potential bias from an inaccurate numerical solution. In the numerical tests we carried out in this paper, it only takes a few steps for  $\sigma_h^k$  to converge, so we always choose  $\lambda = 0$ .

## 4 Numerical experiments

In this section, we carry out numerical experiments to test out the performance of our proposed numerical methods. In the first subsection, we test the goal-oriented AMR method for the forward problem. Then, in the second subsection, we test the performance of the method for the inverse problem.

### 4.1 Goal-oriented AMR for the forward problem

In this subsection, we consider the numerical approximation to the forward problem (1).

Before testing the goal-oriented AMR methods, we first test the convergence of the discontinuous Galerkin spectral element (DGSE) method without adaptivity. We consider a rectangular spatial domain  $[0, L_x] \times [0, L_y]$  with the inflow radiation coming from the top and the left boundary:

$$u(\mathbf{x}, \mathbf{s}) = \frac{16}{\pi} \chi_{[\frac{7\pi}{4}, \frac{7\pi}{4} + \frac{\pi}{16}]}(\theta), \quad \text{where } \mathbf{s} = (\cos \theta, \sin \theta), \quad (x, y) \in [0, L_x] \times \{L_y\} \cup \{0\} \times [0, L_y].$$

For the other sides of the boundary, we assume zero inflow radiation and zero source term  $f$ . A scatterer is placed at the center of the domain, so that the extinction coefficient  $\sigma_e$  is

$$\sigma_e = \frac{11}{1 + \exp(-2k_0(0.1L_y - r(x, y)))}. \quad (20)$$

We choose the single scattering albedo to be  $\tilde{\omega} = \frac{10}{11}$ , so the scattering coefficient is  $\sigma_s = \tilde{\omega}\sigma_e$  and the absorption coefficient is  $\sigma_a = \sigma_e - \sigma_s$ . We choose  $k_0 = 2$ . The asymmetric parameter in the scattering phase function in (1) is chosen as  $g = 0.8$ , which is a typical value for water clouds in the atmosphere.

For this test we consider the DG method (9) with the polynomial degree  $p_x = p_y = 1$  and also with  $p_x = p_y = 2$  for the spatial discretization. To compare, we also consider a finite volume spatial discretization method, which is equivalent to the DG method with  $p_x = p_y = 0$ . For all these methods, we shall uniformly partition the domain  $[0, L_x] \times [0, L_y]$  into a collection of elements and slowly decrease the element size. For the angular discretization, we apply a uniform partition into 32 pieces by  $P_0$  element (finite volume angular discretization). Figure 1 shows that the DG method has errors that converge as  $\mathcal{O}(\text{DOFs}^{-(p+1)/d})$  where the dimension  $d = 2$ . This is the expected optimal convergence order for the DG method and the finite volume method (upwind).

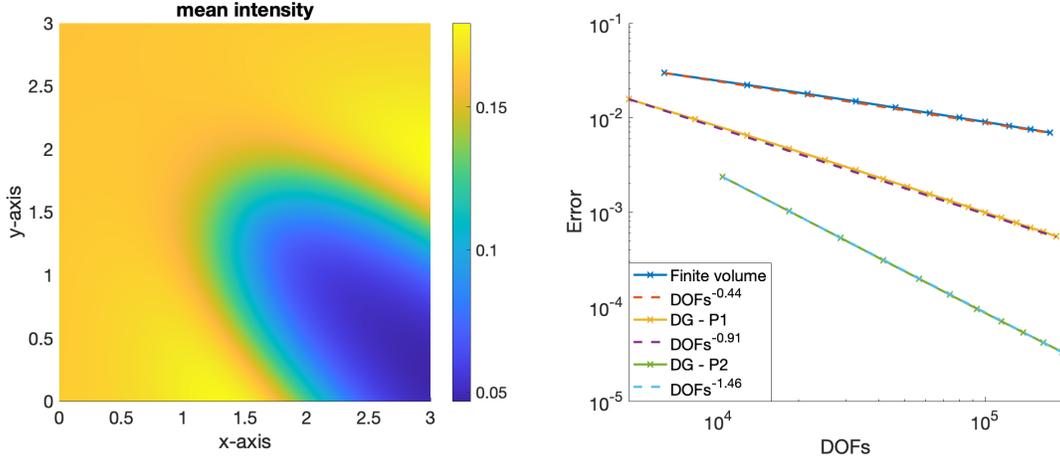


Figure 1: Convergence test for the forward problem, without AMR. Left: mean intensity of the radiative solution calculated by the DG- $P_2$  method. Right:  $L^2$  error of the mean intensity calculated by the finite volume method, DG with  $P_1$ , and DG with  $P_2$  polynomial approximation. All methods converge at the expected rates of  $\mathcal{O}(\text{DOFs}^{-(p+1)/d})$  with dimension  $d = 2$ , and the DG methods with larger  $p$  converge faster.

We next consider the forward problem with the goal-oriented AMR algorithm. We consider a rectangular spatial domain  $[0, L_x] \times [0, L_y]$  with the inflow radiation coming from the top boundary:

$$u(\mathbf{x}, \mathbf{s}) = \frac{16}{\pi} \chi_{[\frac{3\pi}{2}, \frac{3\pi}{2} + \frac{\pi}{16}]}(\theta), \quad \text{where } \mathbf{s} = (\cos \theta, \sin \theta), \quad x \in [0, L_x], \quad y = L_y.$$

For the other sides of the boundary, we assume zero inflow radiation and zero source term  $f$ . For this test, we also consider a scatterer placed at the center of the domain following the form of equation (20). To better test the AMR algorithm, here we consider the scatterer with a sharper boundary transition by taking  $k_0 = 10$ . The other settings are the same as the previous test; namely, we choose  $\tilde{\omega} = 10/11$  and  $g = 0.8$ . See Figure 2 for a visualization of  $\sigma_e$ .

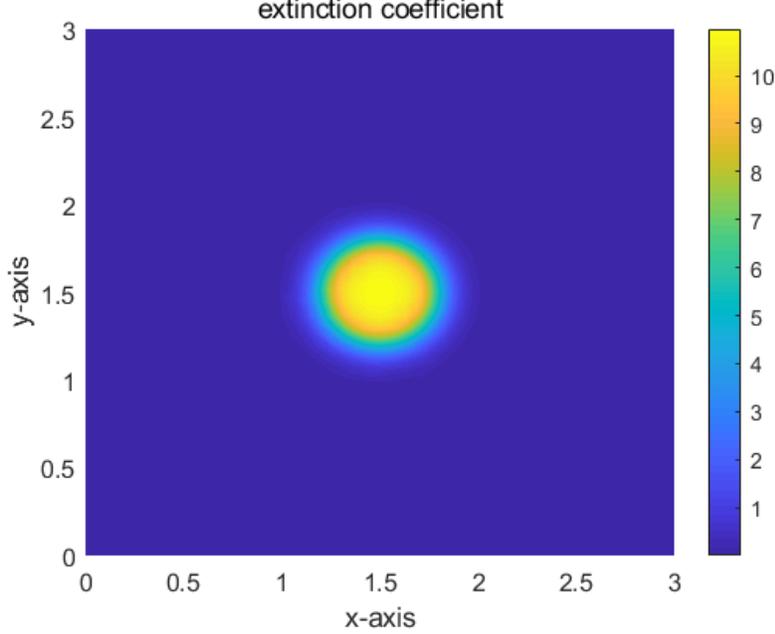


Figure 2: Extinction coefficient  $\sigma_e$ , to represent an idealized cloud located at the center of the domain.

Our aim is to find an efficient numerical approximation to  $J(u)$ , where  $u$  is the solution of (1). Here we set the goal function  $J$  to be (a smoothed version of) a point measurement on the outflow radiation, located at the right side of the domain:

$$J(u) = \int_0^{L_y} u(L_x, y)h(y)dy, \quad \text{where} \quad h(y) = \frac{1}{(L_y/20)\sqrt{2\pi}} \exp\left(-0.5\left(y - \frac{15}{21}L_y\right)^2 / (L_y/20)^2\right).$$

Specifically, the (smoothed) point measurement is put at  $(L_x, \frac{15}{21}L_y)$ . In the first test, we use the error estimator  $\eta_K$  defined in (18) to guide mesh refinement. The mesh refinement Algorithm 3 is repeatably applied with the refinement ratio  $r_{\text{ref}} = 0.2$ , until the DOFs for  $u_h$  reach the threshold  $N_{\text{max.dof}} = 2 \times 10^5$ . To better visualize where the meshes are refined, here we use a fixed polynomial degree  $p = 3$ .

To compare, we also consider a standard jump estimator

$$\eta_K^{\text{jump}} := \left( \frac{1}{|\partial K|} \int_{\partial K} \left| \llbracket \int_S u_h \rrbracket \right|^2 \right)^{1/2}. \quad (21)$$

Since the exact solution  $u$  is not known, we calculate  $|J(u_h) - J(\tilde{u})|$  as an approximation to the true functional error  $|J(u_h) - J(u)|$ , where  $\tilde{u}$  is calculated on a more refined mesh than  $u$ . Here  $\tilde{u}$  is calculated with an extra  $p$ -refinement for each element  $K \in \mathcal{T}_h$ .

Figure 3 shows the numerical solutions and the meshes calculated based on the goal-oriented estimator  $\eta_K$  and the standard jump estimator  $\eta_K^{\text{jump}}$ . We observe that the goal-oriented estimator can successfully guide the mesh to refine at the place where the solution has a sharp gradient, namely, top of the scatter. In addition, it refines along the path

connecting the place of the sharp gradient and the place of the point measurement (located at the right side boundary at  $y = \frac{15}{21}L_y$ ). In contrast, the standard estimator  $\eta^{\text{jump}}$  only refines where there are sharp gradients, without taking into consideration of the effect of the measurement  $J$ .

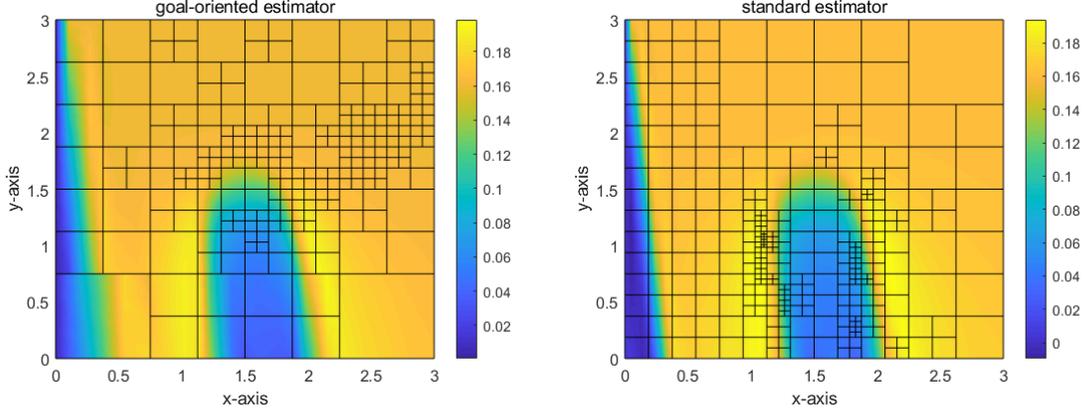


Figure 3: Goal-oriented AMR versus standard AMR. The plots show the mean intensity of the radiant intensity  $u_h$  and the meshes. Left: using the goal-oriented error estimator  $\eta_K$  introduced in (18). Right: using the standard jump estimator  $\eta_K^{\text{jump}}$  defined in (21). The standard jump estimator induces refinement near sharp gradients, whereas the goal-oriented error estimator also induces refinement along the path between the scatterer and the measurement location on the boundary at  $(x, y) = (L_x, (15/21)L_y)$ .

In Figure 4, we plot the error  $|J(u) - J(u_h)|$ . We observe that the goal-oriented estimator reduces the error efficiently while the standard estimator fails to decrease the error. This is consistent with our observation in Figure 3, where the goal-oriented estimator refines both the places of the sharp gradient and the path connecting to the measurement, while the standard estimator fails to refine those elements concerning the measurement  $J$ . The experiment shows that the goal-oriented estimator can be much more efficient than the standard one, when the goal function is not an  $L^2$  or energy-based norm, but instead a measurement on the domain boundary. Note that this is exactly the case for the applications such as optical tomography, which we shall consider in the next subsection.

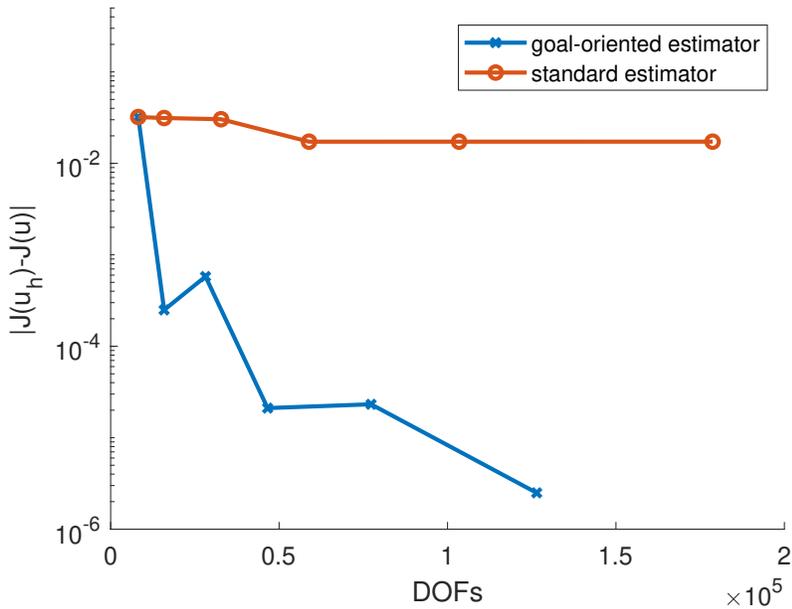


Figure 4: Evaluation of the error  $|J(u_h) - J(u)|$  by the goal-oriented and the standard error estimator. The goal-oriented estimator leads to a much smaller error.

## 4.2 Goal-oriented AMR for the inverse problem

In this subsection, we consider the numerical tests for the inverse problem (13). We aim to test the performance of Algorithm 4 and compare it with other mesh-refinement methods. We consider a rectangular domain  $[0, L_x] \times [0, L_y]$  where on each side, we put 2 inflow radiation laser beams and 20 measurements collecting the angularly-averaged outflow radiation. As a result, we have  $N_{\text{tst}} = 8$  and  $N_{\text{msm}} = 80$  in the formulation (4) or (13). The absorption is fixed as  $\sigma_a = 0.1$  and the asymmetric parameter is chosen as  $g = 0.1$ . The scattering coefficient  $\sigma_s$  is decomposed into a summation  $\sigma_s = \sigma_s^0 + \tilde{\sigma}_s$  where  $\sigma_s^0$  is a background state and  $\tilde{\sigma}$  represents a perturbation. We assume the background state  $\sigma_s^0$  and the absorption  $\sigma_a$  are known and we aim to recover the scattering coefficient  $\sigma_s$ . We shall consider two test cases with the same background state  $\sigma_s^0 = 1$  but with different perturbations  $\tilde{\sigma}_s$ . See the upper-left sub-figures of Figure 6 and Figure 9 for the true scattering coefficients  $\sigma_s$  for the first and the second test case, respectively.

To generate the data  $y^{ij} = \mathcal{M}^j \tilde{u}_h^i$ , we solve the forward problem (9) on a very refined mesh, namely, a 8 by 8 mesh with the polynomial degree chosen as  $p = 9$ , so we can regard  $\tilde{u}_h^i$  as the “true” solution. The optical parameter  $\sigma_h$  is set to live on the one-element mesh with  $p = 19$ .

To solve the inverse problem, we apply an  $H^1$ -regularization, namely  $\mathcal{R}(\sigma) = \|\nabla \sigma\|_{L^2(\Omega)}$ , with the regularization parameter chosen as  $\alpha = 10^{-1}$ ; see (4) where  $\alpha$  was first introduced. We implement Algorithm 4, with  $\epsilon = 10^{-3}$ , and  $r_{\text{ref}} = 0.2$  for the mesh refinement Algorithm 3. For the iteration method, we use the limited memory BFGS algorithm to calculate the updating direction  $\delta \sigma_h^k$ . The parameter  $\gamma$  is determined by a line search method, for which we use the backtracking algorithm which starts with  $\mu = 1$  and then followed by the candidate

step length  $\mu = 2^{-1}, 2^{-2}, 2^{-3}, \dots$ , until the Armijo condition

$$\Phi(\sigma_h^k - \mu \delta \sigma_h^k) \leq \Phi(\sigma_h^k) - c_1 \mu \nabla \Phi(\sigma_h^k) \delta \sigma_h^k$$

is satisfied, where  $c_1 = 10^{-4}$ . For more details on the BFGS method and the line search algorithm, we refer to [31].

We test the performance of Algorithm 4 by comparing it with other mesh-refinement methods. We set the initial mesh  $\mathcal{T}_{hp}^i$  ( $i = 1, \dots, N_{\text{tst}}$ ) by a 4 by 4 mesh with  $p = 2$ . To be more specific, we shall compare Algorithm 4 with other methods, for which, instead of applying the AMR Algorithm 3 in the Algorithm 4, we apply (1)  $h$ -refinement, for which we partition the domain into more elements (uniformly divided), (2)  $p$ -refinement, for which we increase the polynomial degree on each element, (3)  $h$ -AMR-goal, for which we only apply  $h$ -refinement in the Algorithm 3, and (4)  $hp$ -AMR-standard, for which we use the standard jump error estimator (18) to guide the mesh-refinement in the Algorithm 3. We remark that the  $h$ -refinement method with  $p = 2$  is similar to a finite volume method as mesh resolution is increased, while the  $p$ -refinement method with an increased polynomial order  $p$  is similar to a spectral method in the sense that an exponential (spectral) convergence can be achieved when the solution is sufficiently smooth.

Now we show the results for the first test case. Figure 5 shows that the  $hp$ -AMR-goal method is the most efficient method in reducing the error  $\|\sigma_h - \sigma_h^*\|_{L^2(\Omega)}$ . Namely, it uses the fewest DOFs to achieve the best recovery. The  $h$ -AMR-goal method is overall less efficient than the  $hp$ -AMR-goal method but their performances are similar. For the  $hp$ -AMR-standard,  $h$ -Ref, and  $p$ -Ref methods, we observe their errors first decrease and then increase, and in *all* DOFs levels, their recovered optical property is poorer than the one recovered by the  $hp$ -AMR-goal method. The error for the  $hp$ -AMR-standard method decreases fast at the beginning but then the error increases to back to the original value, but even in its decreasing phase its performance is less efficient than the  $hp$ -AMR-goal method. This consolidates our claim that the *goal-oriented* estimator is necessary for inverse problems. For  $h$ -Ref and  $p$ -Ref methods, we observe that they decrease the errors much more slowly than  $hp$ -AMR-goal. In summary, the  $hp$ -AMR-goal method behaves better than other methods in almost any measure, and the  $h$ -AMR-goal method behaves less well but similar to  $hp$ -AMR-goal.

To see this more clearly, we plot in Figure 6, the recovered scattering coefficients by the different refinement methods. To make a fair comparison, for all the refinement methods, we take the snapshot at the refinement step such that the DOFs are closest to  $10^6$ . The results agree with what we observe in Figure 5. Namely, the  $hp$ AMR-goal method gives the best approximation to the true scattering coefficient, while using the fewest number of DOFs.

As additional detail of what is seen in Figure 6, we observe that the  $h$ Ref method and  $hp$ AMR-standard method both fail to provide a satisfactory recovery to the true distribution (left-top sub-figure). All of the  $p$ Ref,  $h$ AMR-goal, and  $hp$ -AMR methods can successfully recover the correct location and the approximate shape of the perturbation  $\tilde{\sigma}_s$ . In further detail, though, the  $h$ -AMR-goal and  $hp$ -AMR-goal methods provide a more accurate recovery for the amplitude of perturbation in comparison to the  $p$ -Ref method. Note that here we chose a fixed regularization parameter  $\alpha$  for all refinement methods for a fair comparison. Despite this, we remark that if it is allowed for the regularization parameter  $\alpha$  to be adapted properly to the different refinement algorithms, the recovery for  $p$ Ref,  $h$ AMR-goal, and  $hp$ AMR-goal might further improve.

To visualize how the meshes are refined by the different methods, we plot the forward solution and the corresponding meshes in Figure 7. For goal-oriented  $hp$ -AMR method, we observe that the mesh is refined both at the place where the forward solution has a sharp gradient, and also at the left boundary where the measurement is located. In contrast, the standard  $hp$ -AMR method only refines the mesh where the sharp gradients are located, but fail to refine also at the measurements. The goal-oriented  $h$ -AMR method refines both at the sharp gradients and at the boundary where the measurements are located. However, it does not use a high order approximation as in the  $hp$ -AMR cases. This could explain why its performance is overall less great than  $hp$ -AMR-goal. Therefore, the goal-oriented  $hp$ -AMR is the only method that supports both (1) a balanced refinement strategy between resolving the forward and the adjoint solution, and (2) an automatic high-order approximation by  $p$ -adaptivity. Considering this, it is not surprising to see that the  $hp$ -AMR-goal method recovers the optical coefficient with the best quality in Figure 6.

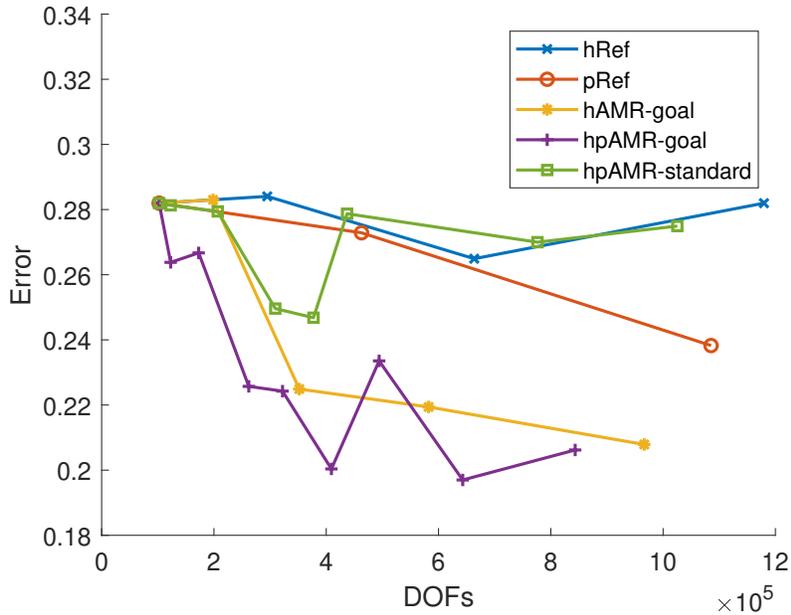


Figure 5: Total DOFs of the forward and the adjoint solutions (the summation of all tests) versus the  $L^2$  error of the optical property  $\|\sigma_h - \sigma_h^*\|_{L^2(\Omega)}$  for the first test case

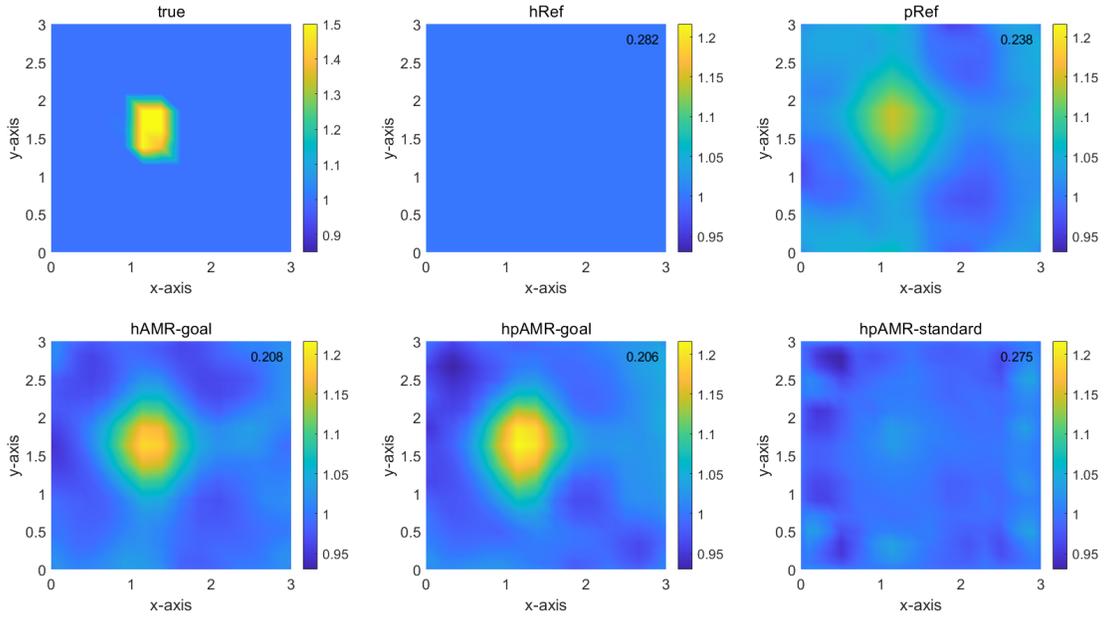


Figure 6: Recovered scattering coefficients  $\sigma_s$  for the different mesh refinement methods, with the total DOFs at around  $10^6$  for the first test case.  $L^2$  errors are printed at the upper-right corners for each sub-figure.

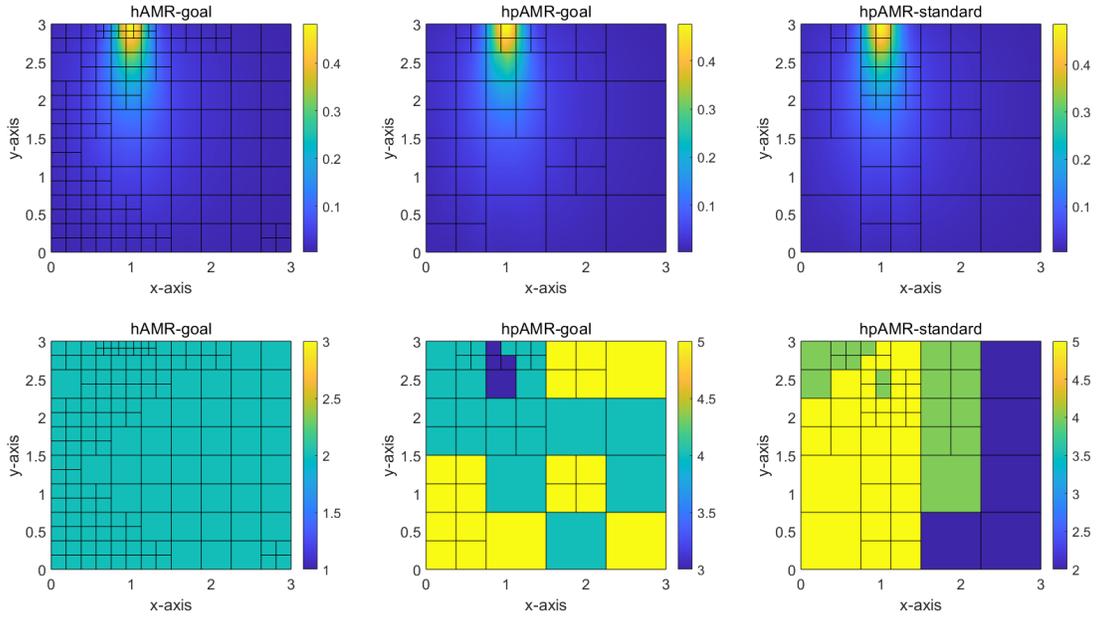


Figure 7: Top row: forward solution  $u_h^1$  for the first test case, by goal-oriented hpAMR (left), goal-oriented hAMR(middle), and standard hpAMR(right) at the last step of the refinement. Bottom row: the corresponding meshes  $\mathcal{T}_{hp}^1$ , where the colorbar represents the polynomial degrees.

To test the reliability of our proposed method, we conduct another experiment (the second test case) with the same setting used for Figure 5 but with a different scattering coefficient  $\sigma_s$  to be recovered; see the top-left figure of Figure 9 for a visualization of  $\sigma_s$ . From Figure 8 and Figure 9, we again observe that our proposed goal-oriented hpAMR method has the best performance compared to all the other refinement approaches. Namely, it uses the fewest DOFs but provides the best quality of the recovered scattering coefficient  $\sigma_s$ .

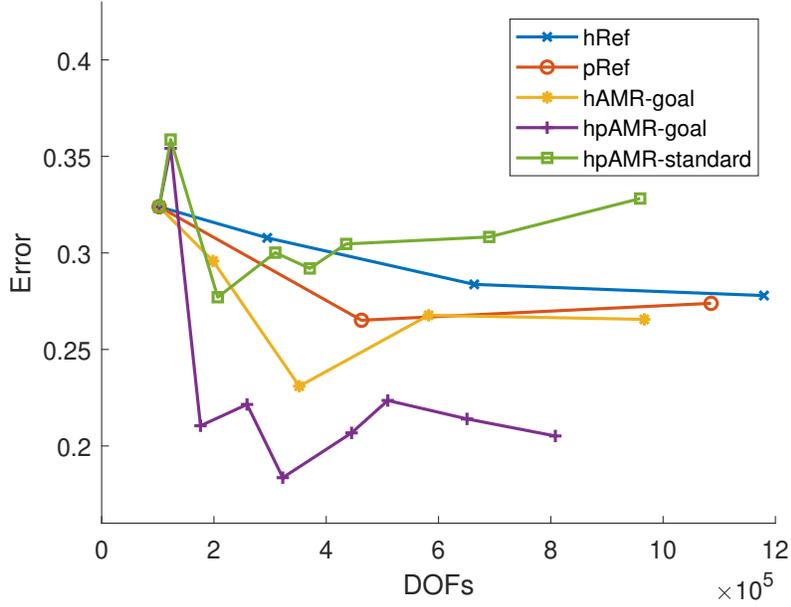


Figure 8: Total DOFs of the forward and the adjoint solutions (the summation of all tests) versus the  $L^2$  error of the optical property  $\|\sigma_h - \sigma_h^*\|_{L^2(\Omega)}$  for the second test case.

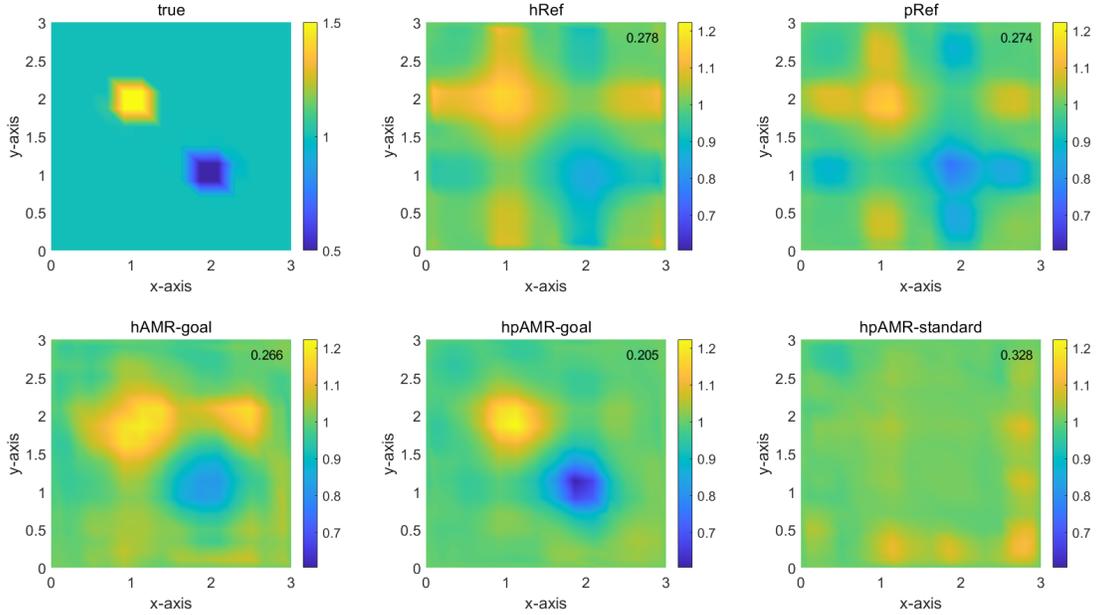


Figure 9: Recovered scattering coefficients  $\sigma_s$  for the different mesh refinement methods, with the total DOFs at around  $10^6$  for the second test case.  $L^2$  errors are printed at the upper-right corners for each sub-figures.

To test the robustness of the proposed goal-oriented  $hp$ -adaptive algorithm, we conduct

additional numerical tests with noise added to the signals  $y^{ij}$ . Namely, instead of using  $y^{ij}$ , we use  $\tilde{y}^{ij} := y^{ij}(1 + \delta X_{ij})$  where  $X_{ij}$  are independent, identically distributed random variables from a uniform distribution on  $[-1, 1]$ , and  $\delta$  represents the noise level. We consider two cases of  $\delta = 1\%$  and  $\delta = 10\%$ .

Figure 10 shows that when the noise level  $\delta = 1\%$ , the performance of the different refinement methods are similar to the noise-free case (Figure 5). We again observe that the  $hp$ -AMR-goal methods gives the smallest error while using the fewest DOFs. When the noise level  $\delta$  increases to  $10\%$ , we observe an increase of the errors for most refinement methods. Despite this, the  $hp$ -AMR-goal method still provides the overall smallest error compared to the other refinement approaches, especially in the last refinement level. In Figure 11 and Figure 12, we plot the recovered scattering coefficient by the different refinement methods with the noise level  $1\%$  and  $10\%$ , respectively. The results are consistent with what we observe in Figure 10.

We repeat the noise test for the second test case (with the true scatterer shown in the upper-left sub-figure of Figure 9). In Figure 13, we observe that in both cases of the noise level of  $\delta = 1\%$  and  $\delta = 10\%$ , the  $hp$ -AMR-goal method provides the overall smallest error. We plot in Figure 14 and 15 the recovered scattering coefficients. The results are consistent with what we observe in Figure 13 in the sense that the  $hp$ -AMR-goal methods provides the best quality of the recovery in the different noise levels.

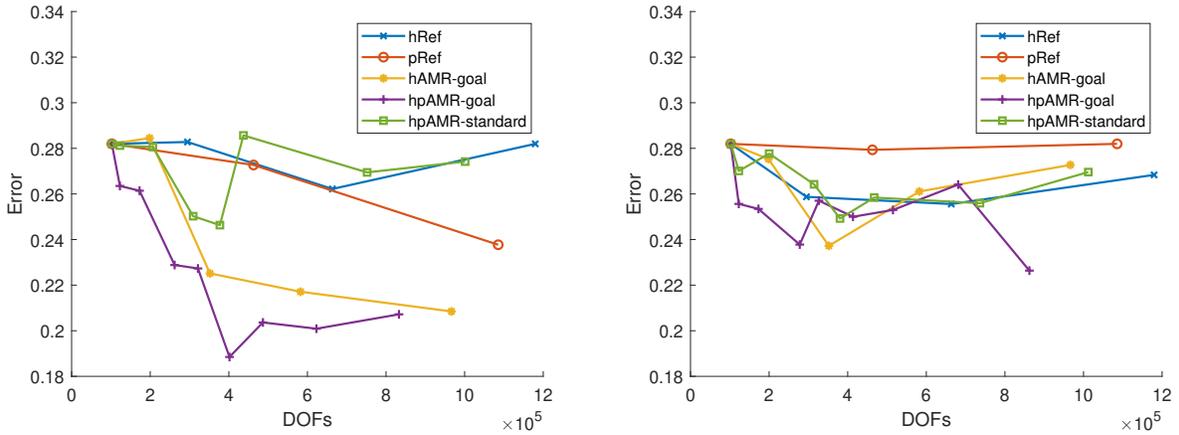


Figure 10: Total DOFs of the forward and the adjoint solutions (the summation of all tests) versus the  $L^2$  error of the optical property  $\|\sigma_h - \sigma_h^*\|_{L^2(\Omega)}$  for the first test case. Left:  $1\%$  noise is added to  $y^{ij}$ . Right:  $10\%$  noise is added to  $y^{ij}$ .

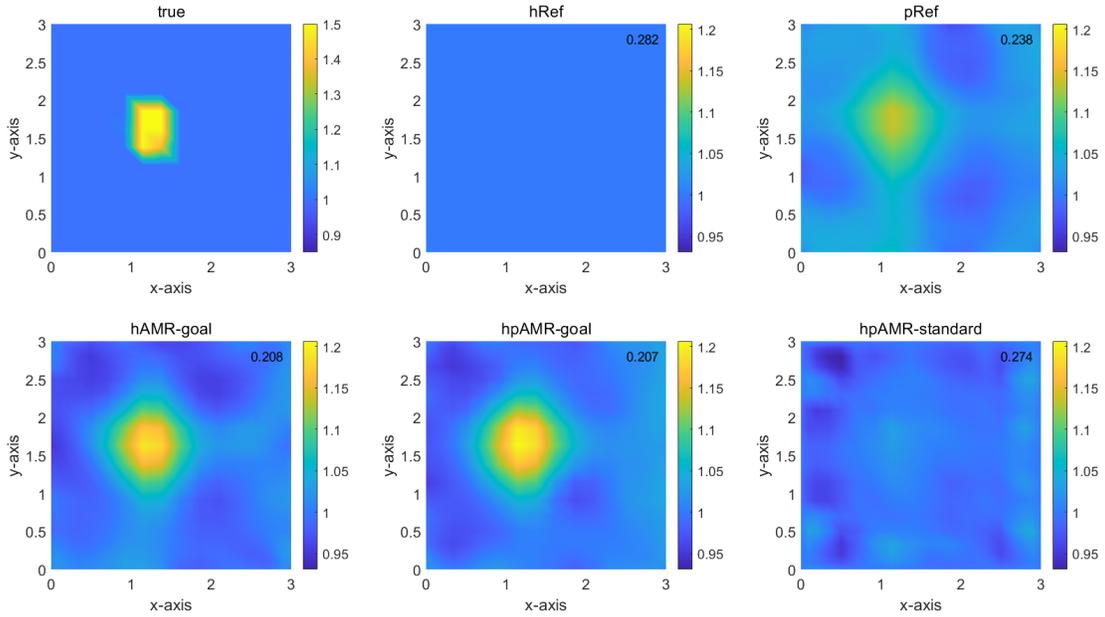


Figure 11: Recovered scattering coefficients  $\sigma_s$  for the different mesh refinement methods, with the total DOFs at around  $10^6$  for the first test case.  $L^2$  errors are printed at the upper-right corners for each sub-figures. 1% noise is added to  $y^{ij}$ .

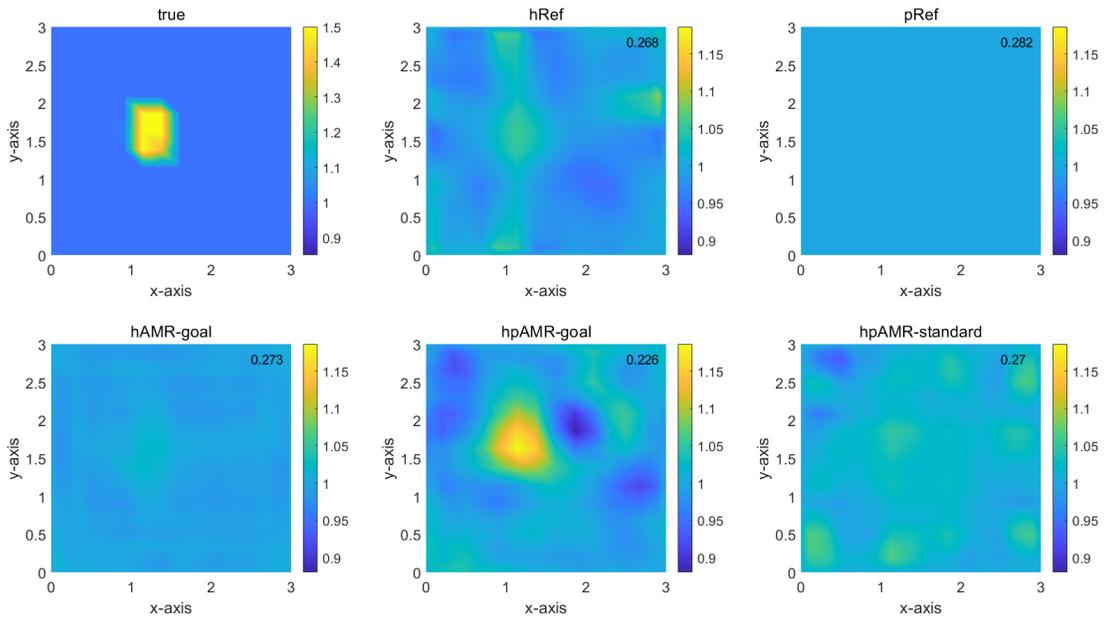


Figure 12: Recovered scattering coefficients  $\sigma_s$  for the different mesh refinement methods, with the total DOFs at around  $10^6$  for the first test case.  $L^2$  errors are printed at the upper-right corners for each sub-figures. 10% noise is added to  $y^{ij}$ .

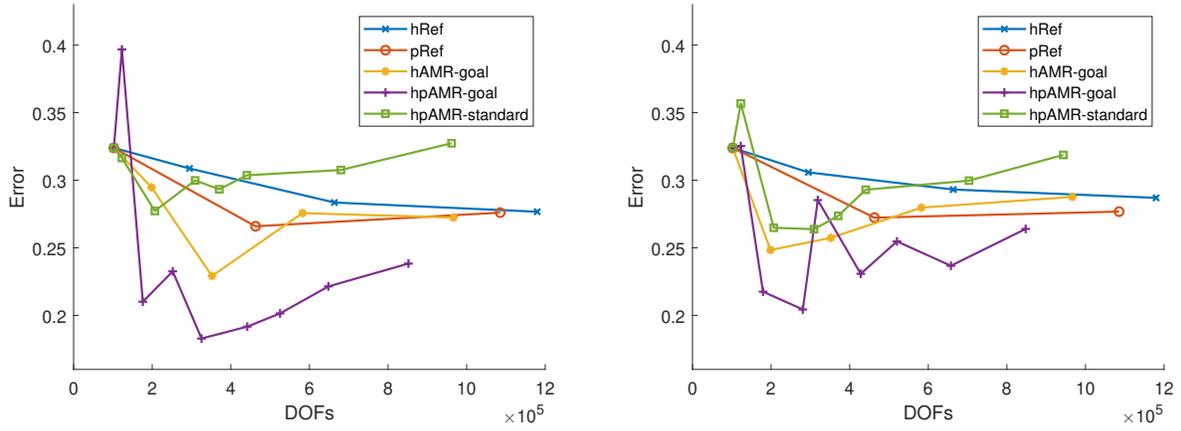


Figure 13: Total DOFs of the forward and the adjoint solutions (the summation of all tests) versus the  $L^2$  error of the optical property  $\|\sigma_h - \sigma_h^*\|_{L^2(\Omega)}$  for the second test case. Left: 1% noise is added to  $y^{ij}$ . Right: 10% noise is added to  $y^{ij}$ .

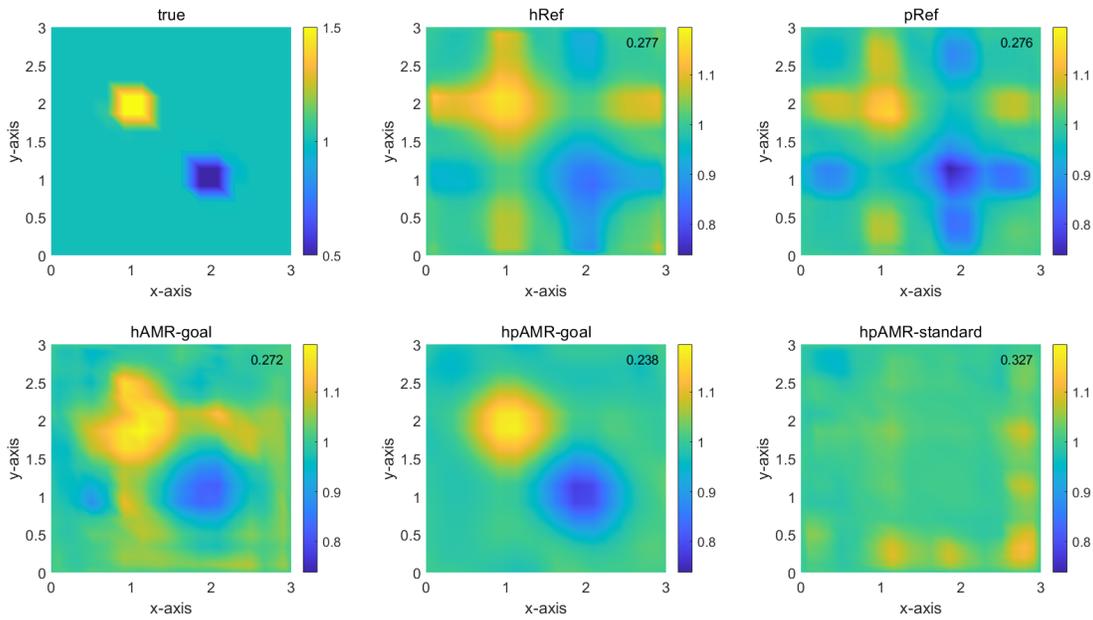


Figure 14: Recovered scattering coefficients  $\sigma_s$  for the different mesh refinement methods, with the total DOFs at around  $10^6$  for the second test case.  $L^2$  errors are printed at the upper-right corners for each sub-figures. 1% noise is added to  $y^{ij}$ .

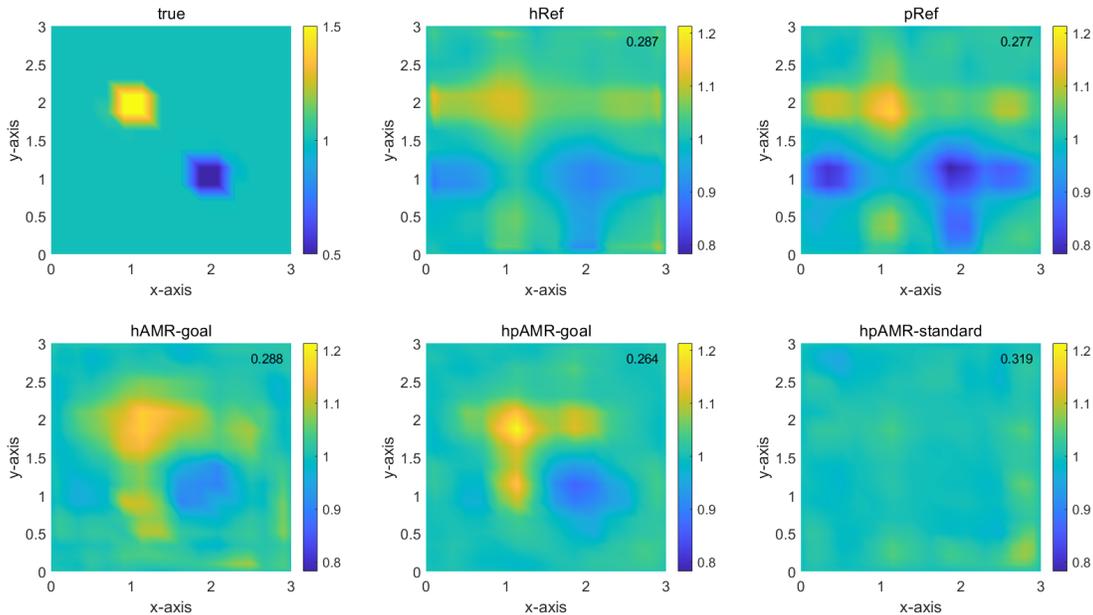


Figure 15: Recovered scattering coefficients  $\sigma_s$  for the different mesh refinement methods, with the total DOFs at around  $10^6$  for the second test case.  $L^2$  errors are printed at the upper-right corners for each sub-figures. 10% noise is added to  $y^{ij}$ .

## 5 Conclusion

In this paper we propose a goal-oriented  $hp$ -adaptive mesh refinement method to solve the inverse radiative transfer equation. The method is based on the development of a novel goal-oriented error estimator, which is achieved by connecting two kinds of duality arguments in different fields, namely, (1) the duality-based based mesh adaptivity for goal-oriented error minimization, and (2) the adjoint-based inversion techniques for solving inverse problems. The numerical tests suggest that the proposed method solves the inverse problem with the best quality of the recovered optical coefficients, while using the fewest DOFs. While our method is proposed here for inverse radiative transfer, the general principles of devising the error estimators and designing the refinement algorithms should be able to be naturally extended to enable adaptive-mesh inversion for other types of inverse problems, such as the Calderón problem or inverse scattering.

Potential future work includes further study of combinations of more involved regularization strategies, and applications to remote sensing problems.

## Acknowledgement

This research was partially supported by grant NSF DMS 2324368 and by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison with funding from the Wisconsin Alumni Research Foundation. The authors thank

two anonymous reviewers whose suggestions led to improvements of the paper.

## References

- [1] Gassan S Abdoulaev and Andreas H Hielscher. Three-dimensional optical tomography with the equation of radiative transfer. *Journal of Electronic Imaging*, 12(4):594–601, 2003.
- [2] S. R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15(2):R41–R93, 1999.
- [3] S. R. Arridge, J. P. Kaipio, V. Kolehmainen, M. Schweiger, E. Somersalo, T. Tarvainen, and M. Vauhkonen. Approximation errors and model reduction with an application in optical diffusion tomography. *Inverse Problems*, 22(1):175–195, 2006.
- [4] Simon R. Arridge and John C. Schotland. Optical tomography: forward and inverse problems. *Inverse Problems*, 25(12):123010, 59, 2009.
- [5] Ivo Babuška and Manil Suri. The  $p$  and  $h$ - $p$  versions of the finite element method, basic principles and properties. *SIAM Rev.*, 36(4):578–632, 1994.
- [6] Guillaume Bal. Inverse transport theory and applications. *Inverse Problems*, 25(5):053001, 48, 2009.
- [7] Wolfgang Bangerth. A framework for the adaptive finite element solution of large-scale inverse problems. *SIAM J. Sci. Comput.*, 30(6):2965–2989, 2008.
- [8] Wolfgang Bangerth and Amit Joshi. Adaptive finite element methods for the solution of inverse problems in optical tomography. *Inverse Problems*, 24(3):034011, 22, 2008.
- [9] Roland Becker and Rolf Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.
- [10] Ke Chen, Qin Li, and Jian-Guo Liu. Online learning in optical tomography: a stochastic approach. *Inverse Problems*, 34(7):075010, 26, 2018.
- [11] Ke Chen, Qin Li, and Li Wang. Stability of stationary inverse transport equation in diffusion scaling. *Inverse Problems*, 34(2):025004, 26, 2018.
- [12] Mourad Choulli and Plamen Stefanov. An inverse boundary value problem for the stationary transport equation. *Osaka J. Math.*, 36(1):87–104, 1999.
- [13] Adrian Doicu, Alexandru Doicu, Dmitry Efremenko, and Thomas Trautmann. Cloud tomographic retrieval algorithms. i: Surrogate minimization method. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 277:107954, 2022.
- [14] Adrian Doicu, Alexandru Doicu, Dmitry S Efremenko, and Thomas Trautmann. Cloud tomographic retrieval algorithms. ii: Adjoint method. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 285:108177, 2022.

- [15] Shukai Du and Samuel N Stechmann. Fast, low-memory numerical methods for radiative transfer via hp-adaptive mesh refinement. *Journal of Computational Physics*, 480:112021, 2023.
- [16] Herbert Egger and Matthias Schlottbom. An Lp theory for stationary radiative transfer. *Applicable Analysis*, 93(6):1283–1296, 2014.
- [17] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
- [18] B Guo and Ivo Babuška. The hp version of the finite element method. *Computational Mechanics*, 1(1):21–41, 1986.
- [19] Murat Guven, Birsen Yazici, Kiwoon Kwon, Eldar Giladi, and Xavier Intes. Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging. I. *Inverse Problems*, 23(3):1115–1133, 2007.
- [20] Murat Guven, Birsen Yazici, Kiwoon Kwon, Eldar Giladi, and Xavier Intes. Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging. II. *Inverse Problems*, 23(3):1135–1160, 2007.
- [21] Paul Houston and Endre Süli. A note on the design of hp-adaptive finite element methods for elliptic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):229–243, 2005.
- [22] Alexander D Klose and Andreas H Hielscher. Optical tomography using the time-independent equation of radiative transfer—part 2: inverse model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 72(5):715–732, 2002.
- [23] Alexander D. Klose and Andreas H. Hielscher. Quasi-Newton methods in optical tomographic image reconstruction. *Inverse Problems*, 19(2):387–409, 2003.
- [24] Alexander D Klose, Uwe Netz, Jürgen Beuthan, and Andreas H Hielscher. Optical tomography using the time-independent equation of radiative transfer—part 1: forward model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 72(5):691–713, 2002.
- [25] Ru-Yu Lai and Qin Li. Parameter reconstruction for general transport equation. *SIAM J. Math. Anal.*, 52(3):2734–2758, 2020.
- [26] Ru-Yu Lai, Qin Li, and Gunther Uhlmann. Inverse problems for the stationary transport equation in the diffusion scaling. *SIAM J. Appl. Math.*, 79(6):2340–2358, 2019.
- [27] Edward W. Larsen. Solution of multidimensional inverse transport problems. *J. Math. Phys.*, 25(1):131–135, 1984.
- [28] Edward W. Larsen. Solution of three-dimensional inverse transport problems. *Transport Theory Statist. Phys.*, 17(2-3):147–167, 1988.

- [29] Aviad Levis, Yoav Y Schechner, Anthony B Davis, and Jesse Loveridge. Multi-view polarimetric scattering cloud tomography and retrieval of droplet size. *Remote Sensing*, 12(17):2831, 2020.
- [30] NJ McCormick. Inverse radiative transfer problems: a review. *Nuclear science and Engineering*, 112(3):185–198, 1992.
- [31] Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.
- [32] Kui Ren. Recent developments in numerical techniques for transport-based medical imaging methods. *Commun. Comput. Phys.*, 8(1):1–50, 2010.
- [33] Kui Ren, Guillaume Bal, and Andreas H. Hielscher. Frequency domain optical tomography based on the equation of radiative transfer. *SIAM J. Sci. Comput.*, 28(4):1463–1489, 2006.
- [34] Roi Ronen, Yoav Y Schechner, and Eshkol Eytan. 4d cloud scattering tomography. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5520–5529, 2021.
- [35] Vadim Y Soloviev. Mesh adaptation technique for fourier-domain fluorescence lifetime imaging. *Medical physics*, 33(11):4176–4183, 2006.
- [36] Vadim Y Soloviev and Lada V Krasnosselskaia. Dynamically adaptive mesh refinement technique for image reconstruction in optical tomography. *Applied optics*, 45(12):2828–2837, 2006.
- [37] Plamen Stefanov. Inverse problems in transport theory. In *Inside out: inverse problems and applications*, volume 47 of *Math. Sci. Res. Inst. Publ.*, pages 111–131. Cambridge Univ. Press, Cambridge, 2003.
- [38] R. A. Tapia and M. W. Trosset. An extension of the Karush-Kuhn-Tucker necessity conditions to infinite programming. *SIAM Rev.*, 36(1):1–17, 1994.
- [39] Yaqi Wang and Jean C Ragusa. Standard and goal-oriented adaptive mesh refinement applied to radiation transport on 2D unstructured triangular meshes. *Journal of Computational Physics*, 230(3):763–788, 2011.