

STABILITY OF INVERSE TRANSPORT EQUATION IN DIFFUSION SCALING AND FOKKER-PLANCK LIMIT

KE CHEN, QIN LI, AND LI WANG

ABSTRACT. We consider the inverse problem of reconstructing the scattering and absorption coefficients using boundary measurements for a time dependent radiative transfer equation (RTE). As the measurement is mostly polluted by errors, both experimental and computational, an important question is to quantify how the error is amplified in the process of reconstruction. In the forward setting, the solution to the RTE behaves differently in different regimes, and the stability of the inverse problem vary accordingly. In particular, we consider two scalings in this paper. The first one concerns with a diffusive scaling whose macroscopic limit is a diffusion equation. In this case, we showed, following the similar approach as in [Chen, Li and Wang, arXiv:1703.00097], that the stability degrades when the limit is taken. The second one considers a highly forward peaked scattering, wherein the scattering operator is approximated by a Fokker-Planck operator as a limit. In this case, we showed that a fully recover of the scattering coefficient is less possible in the limit, whereas obtaining a rescaled version of the scattering coefficient becomes more practice friendly.

1. INTRODUCTION

Radiative transfer equation (RTE) describes the dynamics of photon particles propagating in scattering and absorbing media [10]. A typical form reads as

$$\partial_t f + v \cdot \nabla_x f = \int k(x, v, v') f(x, v') dv' - \sigma(x, v) f, \quad (1)$$

equipped with a Dirichlet boundary condition

$$f|_{\Gamma_-} = \phi(t, x, v). \quad (2)$$

Here $f(t, x, v)$ is the distribution of particles at location $x \in \Omega \subset \mathbb{R}^d$ moving with velocity $v \in \mathbb{S}^{d-1}$. Since photons travel with a fixed speed, the velocity v is normalized to $|v| = 1$. $k(x, v, v')$ is the scattering cross section, representing the probability of particles that move in direction v' changing to direction v . σ is the total scattering coefficient that consists of the amount of photon particles being scattering and absorbed by the material. k and σ constitute the main optical property of the material.

The boundary condition (2) is a common choice for RTE, and Γ_- represents the “incoming” portion of the boundary, i.e.,

$$\Gamma_- = \{(x, v) : x \in \partial\Omega, v \cdot n_x < 0\}, \quad (3)$$

where n_x is the unit outer normal direction of the boundary. Similarly, one can define the “outgoing” portion of the boundary by

$$\Gamma_+ = \{(x, v) : x \in \partial\Omega, v \cdot n_x > 0\}. \quad (4)$$

The well-posedness of the forward problem (1) (2) is summarized in [15].

The work of K.C. and Q. L. is supported in part by a start-up fund of Q.L. from UW-Madison and National Science Foundation under the grant DMS-1619778 and DMS-1107291: RNMS KI-Net. The work of L.W. is supported in part by a start-up fund from SUNY Buffalo and the National Science Foundation under the grant DMS-1620135. Both Q.L. and L.W. are grateful to Prof. Kui Ren’s inspiring discussions.

RTE (1) is often incorporated with different scales that lead it to different equations. One typical scaling is the diffusive scaling, under which the RTE is well approximated by a diffusion equation

$$\partial_t \rho = C \nabla_x \left(\frac{1}{\sigma_s} \nabla_x \rho \right) + \sigma_a \rho,$$

where $\rho(t, x) = \int_{\mathbb{S}^{d-1}} f dv$, and σ_s, σ_a related to k and σ will be defined later. C is a generic constant depending on the dimension of the problem. This scaling is encountered in the long time limit with a strong scattering effect. Another is the Fokker-Planck scaling which emphasizes the highly forward peaked scattering. In this case, (1) reduces to

$$\partial_t f + v \cdot \nabla_x f = \mathcal{L}_{\text{FP}} f,$$

where

$$\mathcal{L}_{\text{FP}} = \left[\frac{\partial}{\partial v_3} (1 - v_3^2) \frac{\partial}{\partial v_3} + \frac{1}{1 - v_3^2} \frac{\partial^2}{\partial \psi^2} \right],$$

and $v = (\sqrt{1 - v_3^2} \cos \psi, \sqrt{1 - v_3^2} \sin \psi, v_3)$. In both scenarios, theory exists regarding the derivation, validity, and asymptotic error in the approximation, the reader can make references to [19, 9] for the former case and [20, 23] for the latter.

We study the inverse problem in this paper, with special attention paid to how its stability varies under the above two scalings. Unlike the forward setting wherein the optical properties k and σ are given, and one amounts to solve $f(t, x, v)$ for a specific boundary condition (2), in the inverse problem setting, one tries to recover the unknown optical properties from boundary measurements of $f(t, x, v)$. To be more precise, we define the albedo operator as a mapping from the boundary condition $\phi(t, x, v)$ to the outgoing data $f|_{\Gamma_+}$:

$$\mathcal{A}(k, \sigma) : \quad \phi \mapsto f|_{\Gamma_+},$$

then by adjusting the incoming data ϕ , and measuring the corresponding outgoing data $f|_{\Gamma_+}$, one gains a full knowledge of \mathcal{A} , which can be used to determine k and σ .

The inverse RTE problem benefits a broad application in optical tomography, atmospheric science and aerospace engineering. Optical tomography, with its major application in medical imaging, utilizes scattered light as a probe of structural variations in the optical properties of the tissue. Specifically, a narrow collimated beam of low energy visible or near infrared light is sent into biological tissues, and then collected by an array of detectors after it propagates through the media. The measurements collected are used to recover the optical properties of the media. In atmospheric science, or remote sensing, satellites cumulate hyperspectral light reflected from the earth and is used to infer mineral or plant distribution on the ground. In aerospace engineering, pictures taken by spacecrafts in the universe (Galileo's pictures from Jupiter, or Cassini's pictures from Saturn for example) are sent back to the earth for analyzing mineral/gas distribution on different planets. In all the applications, the forward solver for the light propagation is described by the RTE. One measures the reflected or propagated light intensity to reconstruct the optical properties, with which tissue/ground/gas components are inferred.

On the analytical side, there has been a vast literature on the wellposedness and stability of the inverse problem. In a pioneering paper [14], the authors showed that both k and σ can be uniquely determined by the incoming-to-outgoing map \mathcal{A} , assuming that σ is v -independent. With v dependence, the uniqueness up to the gauge-invariance was shown in [28]. The analysis is done through performing the singular decomposition: one separates the collected $f|_{\Gamma_+}$ data according to the singularities, and different parts are in charge of recovering different coefficients. Another approach is to linearize the equation before applying inverse Born series, and followed by showing the convergence of the series [22]. The results on the stability of the "inverse" dates back to [30] and was made systematic in [6, 7, 8]. Many papers concern the time-dependent case and the associated stability analysis has also been conducted [18, 26, 13], and also [5] for a review.

On the numerical side, special care is needed to address for the illposedness of the problems, both inherited from the continuous counterpart, and due to the incomplete corrupted data. Indeed, to uniquely determine k and σ_a , one needs a full knowledge of \mathcal{A} , which is impractical in real applications; and measurement error can easily propagate and get exaggerated. Typically Tikhonov type regularization is used to balance the pollution and the error tolerance, and the type of regularization embeds some prior knowledge. See [1] using the standard L_2 , [29] using TV regularization for the least variance, [24] summarizing H_1 regularization for some regularity, and L_1 regularization for sparsity. See also Tikhonov type regularization used on each element in the inverse Born series [21, 22]. Besides the illposedness, the size of the problem also brings extra difficulties, and Jacobian-type techniques [27] are introduced to advance the computation.

In the presence of different scales, however, the above mentioned theory or algorithms cannot be directly applied since the inverse problem may completely change its type. One example is the diffusive regime, while the inverse RTE with sufficient variation in measurement is shown to be well-posed, its diffusion limit is the Calderón type problem which is well acknowledged to be ill-posed [16, 17]. Our goal in this paper is to provide a rigorous connection between different scalings in the inverse setting, and show how stability varies with the scaling parameter. For the diffusive scaling, the connection is observed in [25, 2, 3] and addressed in [4, 11]. A similar problem on recovering the doping profile in the Boltzmann-Poisson system is presented [12], wherein numerical simulations also implies this relation. In the Fokker-Planck regime, the limit was briefly mentioned in [5] but the full discussion was rarely seen in the literature. The main contributions in this paper are

- 1) extend our previous analysis with diffusive scaling for steady problems [11] to time-dependent problems, and show that the stability degrades in the diffusion limit;
- 2) examine the well-posedness and stability in the Fokker-Planck scaling which has never been studied in detail before.

The rest of paper is organized as follows. Section 2 is devoted to diffusion regime and Section 3 is devoted to forward peaked regime, in which diffusion equation and Fokker-Planck equation are obtained as asymptotic limit respectively. In both cases, we utilize the linearization approach, study the well-posedness of the problem in both regimes, and examine the change of stability while passing the limit.

2. DIFFUSION REGIME

In this section, we study the wellposedness of the inverse RTE in the diffusion regime. First we briefly recapitulate the properties of RTE and its diffusion limit. For the ease of notation, we assume that the optical properties only have spatial dependence, and rewrite (1) as

$$\partial_t f + v \cdot \nabla_x f = \sigma_s(x) \int (f(t, x, v') - f(t, x, v)) dv' - \sigma_a(x) f(t, x, v), \quad (5)$$

where σ_s is termed the scattering coefficient, and

$$\sigma_a(x) = \sigma(x) - \sigma_s(x)$$

is the absorption coefficient.

The diffusion regime is achieved in the long time limit and when the scattering is much stronger than the absorption. To this end, we introduce a small parameter—Knudsen number Kn and rescale the RTE as follows:

$$\begin{cases} \text{Kn} \partial_t f + v \cdot \nabla_x f = \frac{1}{\text{Kn}} \sigma_s \mathcal{L} f - \text{Kn} \sigma_a f & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ f(0, x, v) = f^I(x, v), \\ f|_{\Gamma_-}(t, x, v) = \phi(t, x, v). \end{cases} \quad (6)$$

Here the collision operator \mathcal{L} is an abbreviation of

$$\mathcal{L} f(t, x, v) = \int (f(t, x, v') - f(t, x, v)) dv'. \quad (7)$$

There are two key features of the collision operator:

- Mass conservation: $\int \mathcal{L}f dv = 0$.
- One dimensional Null space: By setting $\mathcal{L}f = 0$, one gets $f = \int f dv$, meaning that f is a constant in velocity domain. We denote it as $\text{Null } \mathcal{L} = \text{span}\{\rho(t, x)\}$, the collection of functions that depend on t and x only.

2.1. Diffusion limit. When $\text{Kn} \ll 1$, the equation falls into the diffusion regime, and the RTE is asymptotically equivalent to a diffusion equation as $\text{Kn} \rightarrow 0$.

Theorem 1. *Suppose f solves (6) with initial data $f(0, x, v) = f^I(x)$ and boundary data $f|_{\Gamma_-} = \phi(t, x)$, both of which are independent of velocity v . Then as $\text{Kn} \rightarrow 0$, $f(t, x, v)$ converges to $\rho(t, x)$, which solves the heat equation:*

$$\begin{cases} \partial_t \rho - C \nabla_x \cdot \left(\frac{1}{\sigma_s} \nabla_x \rho \right) + \sigma_a \rho = 0, \\ \rho(0, x) = f^I(x), \\ \rho|_{\partial\Omega} = \phi(t, x). \end{cases} \quad (8)$$

Here C is a time dependent constant.

Proof. The proof follows standard asymptotic expansion. In the zero limit of Kn , the distribution converges to the local equilibrium, and by applying the standard asymptotic expansion technique, we write:

$$f_{\text{in}} = f_0 + \text{Kn} f_1 + \text{Kn}^2 f_2 + \dots \quad (9)$$

Inserting the expansion in the equation (6) and equate like powers of Kn :

- $\mathcal{O}(1)$ $\mathcal{L}f_0 = 0$. This immediately indicates that $f_0 \in \text{Null } \mathcal{L}$. With the form given in (7), $\text{Null } \mathcal{L}$ consists functions that are constants in v domain, and thus $f_0(t, x, v) = \rho(t, x)$.
- $\mathcal{O}(\text{Kn})$ $v \cdot \nabla_x f_0 = \sigma_s \mathcal{L}f_1$. This indicates that $f_1 = \mathcal{L}^{-1} \frac{v \cdot \nabla_x f_0}{\sigma_s}$. Notice that \mathcal{L} is invertible on $\text{Null } \mathcal{L}^\perp$, and consider the form of \mathcal{L} in (7), then $\text{Null } \mathcal{L}^\perp = \{f : \int f dv = 0\}$ and $\frac{v \cdot \nabla_x f_0}{\sigma_s} \in \text{Null } \mathcal{L}^\perp$, therefore $f_1 = \mathcal{L}^{-1} \frac{v \cdot \nabla_x f_0}{\sigma_s} = -\frac{v \cdot \nabla_x f_0}{\sigma_s}$.
- $\mathcal{O}(\text{Kn}^2)$ $\partial_t f_0 + v \cdot \nabla_x f_1 = \sigma_s \mathcal{L}f_2 - \sigma_a f_0$. Here we integrate the equation with respect to v . The second corrector f_2 will vanish and the left hand side becomes:

$$\partial_t \rho + \int v \cdot \nabla_x \left(-\frac{v}{\sigma_s} \cdot \nabla_x \rho \right) dv = -\sigma_a \rho \quad \Rightarrow \quad \partial_t \rho - C \nabla_x \cdot \left(\frac{1}{\sigma_s} \nabla_x \rho \right) = -\sigma_a \rho. \quad (10)$$

Integrating $\int v \cdot v dv$ out, we obtains the diffusion limit and conclude the theorem. The constant C depends on the dimension of the velocity space. \square

Remark 1. We comment that the initial and boundary conditions in the above theorem are relatively strict: they both need to be independent of v . If not, one needs to introduce the initial layer and the boundary layer to damp out the nonhomogeneities. More specifically, we write:

$$f(t, x, v) \approx f_{\text{il}}(t, x, v) + f_{\text{bl}}(t, x, v) + f_{\text{int}}(t, x, v), \quad (11)$$

where f_{int} stands for the interior solution and writes:

$$f_{\text{int}}(t, x, v) = \theta(t, x) - \text{Kn} \mathcal{L}^{-1}(v) \partial_x \theta(t, x),$$

with θ satisfying the diffusion equation (8). f_{il} is the initial layer and is governed by:

$$\partial_\tau f_{\text{il}} - \mathcal{L}f_{\text{il}} = 0,$$

where $\tau = t/\text{Kn}^2$ is the rescaled time. With appropriate initial data, f_{il} damps to 0 exponentially fast in τ and thus $f_{\text{il}} \sim e^{-t/\text{Kn}^2} \sim 0$ for finite t with $\text{Kn} \rightarrow 0$. f_{bl} is the boundary layer. At each point on the boundary, $x_0 \in \partial\Omega$, f_{bl} satisfies:

$$v\partial_z f_{\text{bl}} + \mathcal{L}f_{\text{bl}} = 0,$$

where z is the rescaled spatial coordinate around x_0 : $z = -\frac{(x-x_0)\cdot n_x}{\text{Kn}}$ with n_x being the normal direction pointing out of Ω at x_0 . It has been shown that f_{bl} exponentially decays to a constant in z . This constant is termed the extrapolation length, and is uniquely determined by the boundary data around x_0 . We denote it $\phi(t, x_0)$. This means that for x adjacent to x_0 , in the zero limit of Kn , $|f_{\text{bl}} - \phi(t, x_0)| \sim e^{-|x-x_0|/\text{Kn}} \sim 0$. We typically subtract this constant from f_{bl} and set it as the Dirichlet boundary condition for θ , and thus $f_{\text{bl}} \sim 0$ everywhere.

2.2. Recover absorption coefficient σ_a . In this section we assume that the scattering coefficient is known and aim to recover the absorption coefficient. Without loss of generality, we let $\sigma_s \equiv 1$.

2.2.1. Inverse problem setup. We first rewrite (6) into:

$$\begin{cases} \text{Kn}\partial_t f + v \cdot \nabla_x f = \frac{1}{\text{Kn}}\mathcal{L}f - \text{Kn}\sigma_a f & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ f(0, x, v) = 0 & \text{on } \{t=0\} \times \Omega \times \mathbb{S}^{n-1}, \\ f(t, x, v) = \phi(t, x, v) & \text{on } (0, T) \times \Gamma_-. \end{cases} \quad (12)$$

The solution to the above equation, denoted by $f(t, x, v; \phi)$, models the number density of photons with certain inflow ϕ . In experiment, time-dependent velocity-averaged data $m(t, x) := \int_{\mathbb{S}^{n-1}} v \cdot n(x) f(t, x, v)|_{\Gamma_+} dv$ is collected on the out-flow boundary Γ_+ . Therefore, we can define the albedo operator as:

$$\mathcal{A}(\sigma_a) : \quad \phi(t, x, v) \rightarrow m(t, x) = \int_{\mathbb{S}^{n-1}} v \cdot n(x) f(t, x, v)|_{\Gamma_+} dv.$$

Notice that \mathcal{A} nonlinearly depends on σ_a via the solution $f(t, x, v; \phi)$. In an effort to study the property of \mathcal{A} , we first derive a linearized version of it following the procedure outlined in [24].

Suppose a priori information about the absorption coefficient is known in the sense that σ_a can be considered as a small perturbation around a background state $\sigma_{a0}(x)$, i.e.,

$$\sigma_a(x) = \sigma_{a0}(x) + \tilde{\sigma}_a(x) \quad \text{with} \quad |\tilde{\sigma}_a| \ll |\sigma_a|, \quad \text{a.s.},$$

then a linearized problem with background state σ_{a0} and same initial and boundary data can be defined as

$$\begin{cases} \text{Kn}\partial_t f_0 + v \cdot \nabla_x f_0 = \frac{1}{\text{Kn}}\mathcal{L}f_0 - \text{Kn}\sigma_{a0}f_0 & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ f_0(0, x, v) = 0 & \text{on } \{t=0\} \times \Omega \times \mathbb{S}^{n-1}, \\ f_0(t, x, v) = \phi(t, x, v) & \text{on } (0, T) \times \Gamma_-. \end{cases} \quad (13)$$

Comparing (12) and (13), we define the residue $\tilde{f} = f - f_0$, then it solves, to the leading order:

$$\begin{cases} \text{Kn}\partial_t \tilde{f} + v \cdot \nabla_x \tilde{f} = \frac{1}{\text{Kn}}\mathcal{L}\tilde{f} - \text{Kn}\sigma_{a0}\tilde{f} - \text{Kn}\tilde{\sigma}_a f_0 & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ \tilde{f}(0, x, v) = 0 & \text{on } \{t=0\} \times \Omega \times \mathbb{S}^{n-1}, \\ \tilde{f}(t, x, v) = 0 & \text{on } (0, T) \times \Gamma_-, \end{cases} \quad (14)$$

which is obtained by subtracting (13) from (12) with the higher order term $\tilde{\sigma}_a \tilde{f}$ omitted. Notice here that both the linearized solution f_0 and the residue \tilde{f} are implicitly dependent on the incoming data ϕ . We then

introduced an adjoint problem of (13) and assign a Dirac delta function $\delta(\tau, y)$ at the boundary $(0, T) \times \Gamma_+$:

$$\begin{cases} -\text{Kn}\partial_t g - v \cdot \nabla_x g = \frac{1}{\text{Kn}} \mathcal{L}g - \text{Kn}\sigma_{a0}g & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ g(T, x, v) = 0 & \text{on } \{t = T\} \times \Omega \times \mathbb{S}^{n-1}, \\ g(t, x, v) = \delta(\tau, y) & \text{on } (0, T) \times \Gamma_+. \end{cases} \quad (15)$$

The solution is denoted by $g(t, x, v; \tau, y)$. Multiplying (14) with g , (15) with f_0 , integrating over $(0, T) \times \Omega \times \mathbb{S}^{n-1}$ and then subtracting them, we get:

$$\int_{\Gamma_+(y)} \tilde{f}(\tau, y, v) n(y) \cdot v dv = -\text{Kn} \int_{\Omega} \tilde{\sigma}_a(x) \int_{\mathbb{S}^{n-1}} \int_0^T f_0(t, x, v; \phi) g(t, x, v; \tau, y) dt dv dx. \quad (16)$$

We denote the LHS of (16) by $b(\tau, y, \phi)$, then according to the definition of \tilde{f} , it is simply

$$b(\tau, y, \phi) := \int_{\Gamma_+(y)} f(\tau, y, v) n(y) \cdot v dv - \int_{\Gamma_+(y)} f_0(\tau, y, v) n(y) \cdot v dv, \quad (17)$$

with the first term being the measurement from experiments, and the second term computed from equation (13). This term, therefore is known ahead of time. The RHS of (16) defines a linear mapping of $\tilde{\sigma}_a$. Let us denote

$$\gamma_{\text{Kn}}(x; \tau, y, \phi) := -\text{Kn} \int_{\mathbb{S}^{n-1}} \int_0^T f_0(t, x, v; \phi) g(t, x, v; \tau, y) dt dv, \quad (18)$$

then (16) defines a family of linear mapping from γ_{Kn} to the data on the LHS, parametrized by (τ, y, ϕ) :

$$\int_{\Omega} \tilde{\sigma}_a(x) \gamma_{\text{Kn}}(x; \tau, y, \phi) dx = b(\tau, y, \phi). \quad (19)$$

Therefore, (19) defines a linearized albedo operator, from which $\tilde{\sigma}_a$ can be obtained via solving a system of linear equations.

Remark 2. Equation (19) is a first type Fredholm operator, and it holds true for all parameter choices of τ , y and ϕ . The study on the wellposedness simply relies on the space expanded by $\{\gamma_{\text{Kn}}\}$. Suppose we look for $\tilde{\sigma}_a \in L_p(dx)$, then the uniqueness is guaranteed if $\{\gamma_{\text{Kn}}\}$ expands L_q space (with $\frac{1}{p} + \frac{1}{q} = 1$). There has been many studies on the topic and is not the main goal of the current paper. The wellposedness amounts to analyze the “conditioning” of γ_{Kn} . It is closely related to studying its “singular values”, as will be explained in better details below.

2.2.2. Ill-conditioning in the diffusion limit. Given the linearized albedo operator defined in (19), studying the stability of recovering $\tilde{\sigma}_a$ boils down to examining the property of the Fredholm operator of the first kind defined there. In this section, we intend to explore its conditioning with respect to Kn . More precisely, given a family of input-measurement pairs $(\phi(t, x, v), m(\tau, y, \phi))$, where $(t, x, v) \in (0, T) \times \Gamma_-$ and $m(\tau, y, \phi) = \int_{\mathbb{S}^{n-1}} f(\tau, y, v) n(y) \cdot v dv$, we can explicitly compute $\gamma_{\text{Kn}}(x)$ and $b(\tau, y, \phi)$ defined in (18) and (17), and study their dependence on Kn so as to get a sensitivity in recovering $\tilde{\sigma}_a$ with respect to Kn .

In this regard, we first introduce a distinguishability coefficient to quantify the perturbation of $\tilde{\sigma}_a$ when a δ -error is allowed for $b(\tau, y, \phi)$.

Definition 1. Consider linear equations (19) and γ_{Kn} defined in (18) and $b(\tau, y, \phi)$ defined in (17), we defined the distinguishability coefficient as

$$\kappa_a := \sup_{\sigma_a \in \Gamma_{\delta}} \frac{\|\sigma_a - \tilde{\sigma}_a\|_{L^{\infty}(dx)}}{\|\tilde{\sigma}_a\|_{L^{\infty}(dx)}}, \quad (20)$$

where

$$\Gamma_{\delta} = \left\{ \sigma_a : \sup_{\substack{\forall \|\phi\|_{L^{\infty}(\Gamma_-)} \leq 1, \\ \forall y \in \partial\Omega, \tau \in [0, T]}} |\langle \gamma_{\text{Kn}}, \sigma_a \rangle_{L^2(dx)} - b(\tau, y, \phi)| \leq \delta \right\},$$

and $\tilde{\sigma}_a$ is the exact solution to (19).

Here Γ_δ consists of all possible solutions to (19) within δ -tolerance, and the distinguishability coefficient κ quantifies supremum of relative error over Γ_δ . Therefore, in practice small κ is desired. However, this is not the case when Kn is small, as will be shown in the following theorem: small κ leads to very bad distinguishability.

Theorem 2. *For a family of linear equations defined in (19) and an error tolerance $\delta > 0$ on the measurement, the distinguishability coefficient satisfies*

$$\kappa_a = \mathcal{O}\left(\frac{\delta}{\text{Kn}}\right) \quad \text{when} \quad \text{Kn} \ll 1.$$

Proof. For any $\sigma_a \in \Gamma_\delta$, define $c = \sigma_a - \tilde{\sigma}_a$, then we have

$$\left| \int_{\Omega} \gamma_{\text{Kn}}(x) c(x) dx \right| \leq \delta. \quad (21)$$

When $\text{Kn} \ll 1$, from Theorem 1, $f_0(t, x, v)$ can be decomposed into two part

$$f_0(t, x, v) = f_{\text{int}}(t, x, v) + f_{\text{bl}}(t, x, v),$$

where $f_{\text{bl}}(t, x, v)$ encodes the boundary layer supported near the boundary with $\mathcal{O}(\text{Kn})$ width and f_{int} is the interior solution, and it approaches to its diffusion limit $\rho_f(t, x)$ which satisfies (8) with zero initial data and suitable boundary condition. Specifically, $f_{\text{int}}(t, x, v)$ can be expanded as:

$$f_{\text{int}}(t, x, v) = \rho_f(t, x) - \text{Kn} v \cdot \nabla_x \rho_f(t, x) + \mathcal{O}(\text{Kn}^2), \quad (22)$$

where ρ_f solves

$$\begin{cases} \partial_t \rho_f = C \Delta_x \rho_f - \sigma_a \rho_f & \text{in } (0, T) \times \Omega, \\ \rho_f(0, x) = 0 & \text{on } \{t = 0\} \times \Omega, \\ \rho_f(t, x) = \eta_\phi(t, x) & \text{on } (0, T) \times \partial\Omega. \end{cases} \quad (23)$$

Here the boundary value $\eta_\phi(x)$ is computed from $\phi(t, x, v)$ through the boundary layer analysis. (Details are provided in Remark 1).

Likewise, g admits the same decomposition that separates the interior part from the boundary part:

$$g_0(t, x, v) = g_{\text{int}}(t, x, v) + g_{\text{bl}}(t, x, v),$$

and $g_{\text{int}}(t, x, v)$ has the following expansion:

$$g_{\text{int}}(t, x, v) = \rho_g(t, x) - \text{Kn} v \cdot \nabla_x \rho_g(t, x) + \mathcal{O}(\text{Kn}^2),$$

with ρ_g satisfying

$$\begin{cases} -\partial_t \rho_g = C \Delta_x \rho_g - \sigma_a \rho_g & \text{in } (0, T) \times \Omega, \\ \rho_g(T, x) = 0 & \text{on } \{t = T\} \times \Omega, \\ \rho_g(t, x) = \eta_\delta(t, x) & \text{on } (0, T) \times \partial\Omega. \end{cases} \quad (24)$$

We plug the expansion of f_0 and g in the definition of γ_{Kn} , then in the interior away from the layer:

$$\begin{aligned} \gamma_{\text{Kn}}(x; \tau, y, \phi) &:= -\text{Kn} \int_{\mathbb{S}^{n-1}} \int_0^T f_0(t, x, v; \phi) g(t, x, v; \tau, y) dt dv, \\ &= -\text{Kn} \int_0^T \rho_f(t, x) \rho_g(t, x) dt + \mathcal{O}(\text{Kn}^3). \end{aligned}$$

Simplification is not available inside the layer. In the derivation, the $\mathcal{O}(\text{Kn}^2)$ terms are:

$$\int_0^T \rho_f \int_{\mathbb{S}} v \cdot \nabla_x \rho_g dv dt + \int_0^T \rho_g \int_{\mathbb{S}} v \cdot \nabla_x \rho_f dv dt,$$

and they disappear since the integrands are odd functions. Then inserting this $\gamma_{\mathbf{Kn}}$ back into (21) we have

$$\int_{\Omega} c(x) \gamma_{\mathbf{Kn}}(x) dx = -\mathbf{Kn} \int_{\text{int}} c(x) \int_0^T \rho_f(t, x) \rho_g(t, x) dt dx + \int_{\text{bl}} c(x) \int_0^T \gamma_{\mathbf{Kn}}(x) dt dx + \mathcal{O}(\mathbf{Kn}^3)$$

which readily implies that $\kappa = \mathcal{O}(\frac{\delta}{\mathbf{Kn}})$ by choosing $c = 0$ inside the layer. \square

We note that ρ_f and ρ_g are solutions to the heat equation, and have no dependence on \mathbf{Kn} . In the time-independent case [11], we compensate it by showing that the multiplication of ρ_f and ρ_g , the solutions to two elliptic equations are of low rank, which easily produces one more \mathbf{Kn} . It is no longer the case here: we cannot prove the term $\int_0^T \rho_f \rho_g dt$ being low rank in L_2 , and thus it is hard to obtain $\mathcal{O}(\delta/\mathbf{Kn}^2)$.

Remark 3. We would like to point out that the definition of distinguishability coefficient (20) is a “continuous” analog of the condition number in the discrete setting. In fact, if we discretize equation (19) in x and write it in a matrix form, we get:

$$\mathbf{A} \cdot \tilde{\sigma}_a^{\text{dis}} = \mathbf{b},$$

where each row of \mathbf{A} is $\gamma_{\mathbf{Kn}}$ evaluated at all discrete point with one particular τ , y and ϕ selected:

$$\mathbf{A}_{ij} = \gamma_{\mathbf{Kn}}(x_j; \tau_i, y_i, \phi_i), \quad \text{and} \quad \mathbf{b}_i = b(\tau_i, y_i, \phi_i).$$

Perform the singular value decomposition of \mathbf{A}

$$\mathbf{A} = \mathbf{U} \cdot \Sigma \cdot \mathbf{V}^T = \sum_{i=1}^N \lambda_i \mathbf{u}_i \mathbf{v}_i^T, \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N,$$

with λ_i being the singular values and \mathbf{u}_i and \mathbf{v}_i are column vectors, then

$$\tilde{\sigma}_a^{\text{dis}} = \mathbf{V} \cdot \Sigma^{-1} \cdot \mathbf{U}^T \mathbf{b}.$$

Similarly, a variation of $\tilde{\sigma}_a^{\text{dis}}$, denoted as σ_a^{dis} , satisfies

$$\sigma_a^{\text{dis}} = \mathbf{V} \cdot \Sigma^{-1} \cdot \mathbf{U}^T (\mathbf{b} + \mathbf{b}^\delta).$$

Then the equivalent definition of κ here is:

$$\kappa_{\mathbf{A}} = \max_{\mathbf{b}^\delta: \|\mathbf{b}^\delta\|_\infty < \delta} \frac{\|\sum \frac{1}{\lambda_i} \mathbf{v}_i \mathbf{u}_i^T \mathbf{b}^\delta\|_\infty}{\|\sum \frac{1}{\lambda_i} \mathbf{v}_i \mathbf{u}_i^T \mathbf{b}\|_\infty}.$$

Assuming \mathbf{b} is a fixed vector, and let the denominator being $\mathcal{O}(1)$, then the biggest number is achieved if \mathbf{b}^δ is aligned with \mathbf{u}_N so that $\kappa_{\mathbf{A}} = \frac{\delta}{\lambda_N}$. Or $\kappa_{\mathbf{A}}$ may implicitly depend on \mathbf{b} as well, and if the definition is replace by:

$$\kappa_{\mathbf{A}} = \max_{\mathbf{b}^\delta, \mathbf{b}: \|\mathbf{b}^\delta\|_\infty < \delta \|\mathbf{b}\|_\infty} \frac{\|\sum \frac{1}{\lambda_i} \mathbf{v}_i \mathbf{u}_i^T \mathbf{b}^\delta\|_\infty}{\|\sum \frac{1}{\lambda_i} \mathbf{v}_i \mathbf{u}_i^T \mathbf{b}\|_\infty},$$

then the maximum is achieved by the condition number: $\kappa_{\mathbf{A}} = \frac{\delta \lambda_1}{\lambda_N}$, by aligning \mathbf{b} with \mathbf{u}_1 and \mathbf{b}^δ with \mathbf{u}_N .

2.3. Recover scattering coefficient σ_s . To recover σ_s we follow the same route: first set up the inverse problem through a linearization and show that the stability degrades as $\mathbf{Kn} \rightarrow 0$. Without loss of generality, we set $\sigma_a = 1$.

2.3.1. *Inverse problem setup.* To set up the inverse problem we first recall the forward problem:

$$\begin{cases} \text{Kn} \partial_t f + v \cdot \nabla_x f = \frac{1}{\text{Kn}} \sigma_s \mathcal{L} f - \text{Kn} f & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ f(0, x, v) = 0 & \text{on } \{t = 0\} \times \Omega \times \mathbb{S}^{n-1}, \\ f(t, x, v) = \phi(t, x, v) & \text{on } (0, T) \times \Gamma_-, \end{cases} \quad (25)$$

then a similar linearization procedure can be conducted as follows. Assume that $\sigma_s(x)$ can be written as a superposition of a known background $\sigma_{s0}(x)$ and a perturbation $\tilde{\sigma}_s(x)$ from the background, i.e.,

$$\sigma_s(x) = \tilde{\sigma}_s(x) + \sigma_{s0}(x) \quad \text{with} \quad |\tilde{\sigma}_s| \ll |\sigma_s|, \quad \text{a.s.},$$

then the background solution f_0 satisfies equation:

$$\begin{cases} \text{Kn} \partial_t f_0 + v \cdot \nabla_x f_0 = \frac{1}{\text{Kn}} \sigma_{s0} \mathcal{L} f_0 - \text{Kn} f_0 & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ f_0(0, x, v) = 0 & \text{on } \{t = 0\} \times \Omega \times \mathbb{S}^{n-1}, \\ f_0(t, x, v) = \phi(t, x, v) & \text{on } (0, T) \times \Gamma_-. \end{cases} \quad (26)$$

The residue $\tilde{f} := f - f_0$ then solves:

$$\begin{cases} \text{Kn} \partial_t \tilde{f} + v \cdot \nabla_x \tilde{f} = \frac{1}{\text{Kn}} \sigma_{s0} \mathcal{L} \tilde{f} - \frac{1}{\text{Kn}} \tilde{\sigma}_s \mathcal{L} f_0 - \text{Kn} \tilde{f} & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ \tilde{f}(0, x, v) = 0 & \text{on } \{t = 0\} \times \Omega \times \mathbb{S}^{n-1}, \\ \tilde{f}(t, x, v) = 0 & \text{on } (0, T) \times \Gamma_-. \end{cases} \quad (27)$$

Write the adjoint problem of (26) as:

$$\begin{cases} -\text{Kn} \partial_t g - v \cdot \nabla_x g = \frac{1}{\text{Kn}} \sigma_{s0} \mathcal{L} g - \text{Kn} g & \text{in } (0, T) \times \Omega \times \mathbb{S}^{n-1}, \\ g(T, x, v) = 0 & \text{on } \{t = T\} \times \Omega \times \mathbb{S}^{n-1}, \\ g(t, x, v) = \delta(\tau, y) & \text{on } (0, T) \times \Gamma_+, \end{cases} \quad (28)$$

then multiply it with \tilde{f} and subtract the product of (27) with g , and integrate over $(0, T) \times \Omega \times \mathbb{S}^{n-1}$, we get

$$\int_{\Gamma_+(y)} \tilde{f}(\tau, y, v) n(y) \cdot v dv = \frac{1}{\text{Kn}} \int_{\Omega} \tilde{\sigma}_s(x) \int_{\mathbb{S}^{n-1}} \int_0^T g(t, x, v; \tau, y) \mathcal{L} f_0(t, x, v; \phi) dt dv dx. \quad (29)$$

This equation prompts a linear equation for $\tilde{\sigma}_a$, that is,

$$\int_{\Omega} \tilde{\sigma}_s(x) \gamma_{\text{Kn}}(x; \tau, y, \phi) dx = b(\tau, y, \phi), \quad (30)$$

where

$$\begin{aligned} \gamma_{\text{Kn}}(x; \tau, y, \phi) &:= \frac{1}{\text{Kn}} \int_{\mathbb{S}^{n-1}} \int_0^T g(t, x, v; \tau, y) \mathcal{L} f_0(t, x, v; \phi) dt dv, \\ &= \frac{1}{\text{Kn}} \int_0^T \langle g \rangle \langle f \rangle - \langle gf \rangle dt, \end{aligned}$$

and

$$b(\tau, y, \phi) := \int_{\Gamma_+(y)} (f(\tau, y, v) - f_0(\tau, y, v)) n(y) \cdot v dv.$$

Here $\langle f \rangle := \int_{\mathbb{S}^{n-1}} f(t, x, v) dv$. Again $b(\tau, y, \phi)$ is the data at our disposal—the difference between the measured data and computed data, and we end up with a Fredholm operator of first kind with kernel $\gamma_{\text{Kn}}(x)$. We study in the next section the ill-conditioning for this family of linear equations in the diffusion regime.

2.3.2. *Ill-conditioning in the diffusion limit.* Similar to the previous case, when Kn decreases, the transport equation approaches a diffusion equation and thus recovering the scattering coefficient σ_s is less stable. More precisely, we have the following theorem:

Theorem 3. *For a family of linear equations defined in (30) and an error tolerance $\delta > 0$ on the measurement, define the distinguishability coefficient as*

$$\kappa_s := \sup_{\sigma_s \in \Gamma_\delta} \frac{\|\sigma_s - \tilde{\sigma}_s\|_{L^\infty(dx)}}{\|\tilde{\sigma}_s\|_{L^\infty(dx)}}, \quad (31)$$

where

$$\Gamma_\delta = \left\{ \sigma_s : \sup_{\substack{\forall \|\phi\|_{L^\infty(\Gamma_-)} \leq 1, \\ \forall y \in \partial\Omega, \tau \in [0, T]}} |\langle \gamma_{\text{Kn}}, \sigma_s \rangle_{L^2(dx)} - b(\tau, y, \phi)| \leq \delta \right\},$$

and $\tilde{\sigma}_s$ is the exact solution to (30). Then we have

$$\kappa_s := \mathcal{O}\left(\frac{\delta}{\text{Kn}}\right) \quad \text{when } \text{Kn} \ll 1.$$

Proof. The proof again follows a boundary-interior decomposition and asymptotic expansion. First write g and f_0 as:

$$f_0 = f_{\text{bl}} + f_{\text{int}}, \quad g = g_{\text{bl}} + g_{\text{int}}$$

where f_{bl} and g_{bl} are the boundary layer part, and f_{int} and g_{int} are the interior part that admit the following expansion:

$$\begin{aligned} f_{\text{int}} &= \rho_f - \text{Kn} \frac{v \cdot \nabla_x \rho_f}{\sigma_{s0}} + \text{Kn}^2 f_2, \\ g_{\text{int}} &= \rho_g + \text{Kn} \frac{v \cdot \nabla_x \rho_g}{\sigma_{s0}} + \text{Kn}^2 g_2. \end{aligned} \quad (32)$$

Here ρ_f and ρ_g satisfy the diffusion equations:

$$\partial_t \rho_f + \rho_f - C \nabla_x \left(\frac{1}{\sigma_{s0}} \nabla_x \rho_f \right) = 0, \quad \partial_t \rho_g - \rho_g + C \nabla_x \left(\frac{1}{\sigma_{s0}} \nabla_x \rho_g \right) = 0$$

with suitable initial data and boundary condition.

Now decompose γ_{Kn} also into a layer and interior parts, i.e., $\gamma_{\text{Kn}} = (\gamma_{\text{Kn}})_{\text{bl}} + (\gamma_{\text{Kn}})_{\text{int}}$, then for the interior part, using (32), we have

$$\begin{aligned} (\gamma_{\text{Kn}})_{\text{int}} &= \frac{1}{\text{Kn}} \int_0^T (\langle f_{\text{int}} \rangle \langle g_{\text{int}} \rangle - \langle f_{\text{int}} g_{\text{int}} \rangle) dt \\ &= \frac{1}{\text{Kn}} \int_0^T (\rho_f + \text{Kn}^2 \langle f_2 \rangle) (\rho_g + \text{Kn}^2 \langle g \rangle) \\ &\quad - \left\langle \left(\rho_f - \frac{\text{Kn}}{\sigma_{s0}} v \cdot \nabla_x \rho_f + \text{Kn}^2 f_2 \right) \left(\rho_g + \frac{\text{Kn}}{\sigma_{s0}} v \cdot \nabla_x \rho_g + \text{Kn}^2 g_2 \right) \right\rangle dt \\ &= \frac{\text{Kn}}{\sigma_{s0}^2} \int_0^T \langle (v \cdot \nabla_x \rho_f)(v \cdot \nabla_x \rho_g) \rangle dt + \mathcal{O}(\text{Kn}^2) \\ &= \frac{C \text{Kn}}{\sigma_{s0}^2} \int_0^T \nabla_x \rho_f \cdot \nabla_x \rho_g dt + \mathcal{O}(\text{Kn}^2), \end{aligned} \quad (33)$$

where C again depends on the dimension of the velocity space.

Now denote $c(x) = \sigma_s(x) - \tilde{\sigma}_s(x)$, then one has

$$|\langle \gamma_{\text{Kn}}, c \rangle_{L^2(dx)}| \leq \delta.$$

then choose $\sigma_s(x)$ such that $c(x)$ vanishes in the layer,

$$\langle \gamma_{\text{Kn}}, c \rangle_{L^2(dx)} = \langle (\gamma_{\text{Kn}})_{\text{int}}, c_{\text{int}} \rangle_{L^2(dx)} = \left\langle -\frac{C\text{Kn}}{\sigma_{s0}^2} \int_0^T \nabla_x \rho_f \cdot \nabla_x \rho_g dt, c_{\text{int}} \right\rangle_{L^2(dx)} + \mathcal{O}(\text{Kn}^2),$$

we see that

$$c \sim \mathcal{O}\left(\frac{\delta}{\text{Kn}}\right).$$

□

3. HIGHLY FORWARD PEAKED REGIME

In this section, we consider the anisotropic scattering, and study the well/ill-posedness of the inverse RTE in the highly forward peaked regime, in which the time-dependent RTE is asymptotically equivalent to the Fokker-Planck equation. For simplicity, we study the critical case with zero absorption and x -independent scattering. The radiative transfer equation reads

$$\begin{cases} \partial_t f(t, x, v) + v \cdot \nabla_x f(t, x, v) = \mathcal{L}f(t, x, v), \\ f(0, x, v) = f^I(x, v) \quad \text{on } \Omega \times \mathbb{S}^{d-1}, \\ f(t, x, v) = \phi(t, x, v) \quad \text{on } (0, T) \times \Gamma_-, \end{cases} \quad (34)$$

where the collision operator takes the form:

$$\mathcal{L}f(t, x, v) = \frac{1}{\epsilon^2} \int_{\mathbb{S}^{d-1}} \sigma\left(\frac{1 - v \cdot v'}{\epsilon}\right) (f(t, x, v') - f(t, x, v)) dv'. \quad (35)$$

Without loss of generality, we assume that σ integrates to one, i.e.,

$$\frac{1}{\epsilon} \int_{\mathbb{S}^{d-1}} \sigma\left(\frac{1 - v \cdot v'}{\epsilon}\right) dv' = 1. \quad (36)$$

Considering v' is the incident direction and v is the scattering direction, then the small parameter ϵ reinforces “small-angle” scattering—the kernel is peaked in the forward direction of flight; it also plays a role of mean free path, which accounts for strong scattering effect. Here $v \in \mathbb{S}^{d-1}$ is a unit vector denoting the direction of flight. Hereafter, we will focus on dimension $d = 3$.

The existence of such a regime was longly exposed to the area [23, 20], but has received little attention in the inverse problem setting. It is not quite known how stabilities change according to ϵ despite some conjectures [5]. We address this issue in this section. We will first formally derive the Fokker-Planck limit in section 3.1, and set up the inverse problem in section 3.2. Stability with respect to ϵ will be discussed in section 3.3.

3.1. Fokker-Planck Limit. The equation, in the zero limit of ϵ , loses the large-angle scattering, and effectively is equivalent to the Fokker-Planck equation. The original derivation was seen in [23, 20]. Denote $\mu = v \cdot v'$ the cosine of the scattering angle, then the scattering cross-section has the Legendre polynomial expansion:

$$\frac{1}{\epsilon} \sigma\left(\frac{1 - \mu}{\epsilon}\right) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \sigma_n P_n(\mu),$$

where the projection coefficients on the n -th Legendre polynomial $P_n(\mu)$ is:

$$\sigma_n = \frac{2\pi}{\epsilon} \int_{-1}^1 \sigma\left(\frac{1 - \mu}{\epsilon}\right) P_n(\mu) d\mu. \quad (37)$$

It is immediate that $\sigma_0 = 1$ from (36).

To proceed, we write v using spherical coordinates: $v = (\sqrt{1-v_3^2} \cos \psi, \sqrt{1-v_3^2} \sin \psi, v_3)$ and introduce the spherical harmonic functions

$$Y_{n,m}(v) = \left[\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!} \right]^{1/2} \times (-1)^{(m+|m|)/2} P_{n,|m|}(v_3) e^{in\psi}, \quad n \geq 0, \quad -n \leq m \leq n,$$

where $P_{n,|m|}(v_3)$ are associated Legendre functions

$$P_{n,|m|}(v_3) = (1-v_3^2)^{m/2} \left(\frac{d}{dv_3} \right)^m P_n(v_3), \quad 0 \leq m \leq n.$$

The spherical harmonic functions form a complete set of orthonormal basis and thus any suitably smooth function $f(v)$ defined on the unit sphere can be expanded as

$$f(v) = \sum_{n=0}^{\infty} \sum_{m=-n}^n f_{n,m} Y_{n,m}(v), \quad f_{n,m} := \int_{\mathbb{S}^2} f(v) Y_{n,m}(v) dv. \quad (38)$$

Note also that $P_n(\mu)$ satisfy the addition formula:

$$P_n(v \cdot v') = \frac{4\pi}{2n+1} \sum_{m=-n}^n Y_{n,m}(v) Y_{n,m}(v').$$

Therefore, the collision (35) admits the following expansion

$$\begin{aligned} \mathcal{L}f &= \frac{1}{\epsilon} \sum_{n=0}^{\infty} \sum_{m=-n}^n \sigma_n \int_{\mathbb{S}^2} Y_{n,m}(v) Y_{n,m}(v') [f(v') - f(v)] dv' \\ &= \frac{1}{\epsilon} \left[\sum_{n=0}^{\infty} \sum_{m=-n}^n \sigma_n f_{n,m} Y_{n,m}(v) - \sigma_0 f(v) \right] \\ &= \frac{1}{\epsilon} \sum_{n=0}^{\infty} \sum_{m=-n}^n (\sigma_n - \sigma_0) f_{n,m} Y_{n,m}(v). \end{aligned} \quad (39)$$

The second equality holds because $\int Y_{n,m} dv = 0$ for all $n \geq 1$ and $\int Y_{n,m} dv = 1$ only if $n = m = 0$. Let $\alpha = \frac{1-\mu}{\epsilon}$, we rewrite σ_n in (37) as

$$\begin{aligned} \sigma_n &= 2\pi \int_0^{2/\epsilon} \sigma(\alpha) P_n(1-\epsilon\alpha) d\alpha \\ &= 2\pi \int_0^{2/\epsilon} \sigma(\alpha) \left[P_n(1) - P'_n(1)\epsilon\alpha + \frac{P''_n(1)}{2}(\epsilon\alpha)^2 + \dots \right] d\alpha. \end{aligned} \quad (40)$$

If we define

$$\begin{aligned} \xi_n &:= 2\pi \int_{-1}^1 \frac{1}{\epsilon^2} \sigma\left(\frac{1-\mu}{\epsilon}\right) (1-\mu)^n d\mu \\ &= \epsilon^{n-1} \left(2\pi \int_0^{2/\epsilon} t^n \sigma(t) dt \right) \\ &= \mathcal{O}(\epsilon^{n-1}), \end{aligned} \quad (41)$$

then σ_n can be rewritten as

$$\sigma_n = \epsilon \left[P_n(1)\xi_0 + P'_n(1)\xi_1 + \frac{1}{2}P''_n(1)\xi_2 + \frac{1}{3!}P'''_n(1)\xi_3 + \dots \right]. \quad (42)$$

Note that ξ_0 is fixed and has no dependence on σ due to (36):

$$\xi_0 = 2\pi \int_{-1}^1 \frac{1}{\epsilon^2} \sigma \left(\frac{1-\mu}{\epsilon} \right) d\mu = \frac{2\pi}{\epsilon}. \quad (43)$$

Since

$$P_n(1) = 1, \quad P'_n(1) = \frac{n(n+1)}{2}, \quad P_0(\mu) = 1,$$

we have from (40) that

$$\sigma_n - \sigma_0 = -\epsilon \frac{n(n+1)}{2} \xi_1 + \mathcal{O}(\epsilon^2),$$

and therefore plugging it into (39) we get

$$\mathcal{L}f = \sum_{n=0}^{\infty} \sum_{m=-n}^n -\frac{n(n+1)}{2} \xi_1 f_{n,m} Y_{n,m} + \mathcal{O}(\epsilon). \quad (44)$$

Recall that for the well-known Fokker-Planck operator in spherical coordinates:

$$\mathcal{L}_{\text{FP}} f(v) = \left[\frac{\partial}{\partial v_3} (1 - v_3^2) \frac{\partial}{\partial v_3} + \frac{1}{1 - v_3^2} \frac{\partial^2}{\partial \psi^2} \right] f(v),$$

we have

$$\mathcal{L}_{\text{FP}} Y_{n,m}(v) = -n(n+1) Y_{n,m}(v). \quad (45)$$

Comparing (44) and (45), we get the Fokker-Planck approximation:

$$\mathcal{L}f(v) = \sum_{n=0}^{\infty} \sum_{m=-n}^n -\left(\frac{n(n+1)}{2} \xi_1 + \mathcal{O}(\epsilon) \right) f_{n,m} Y_{n,m} = \frac{\xi_1}{2} \mathcal{L}_{\text{FP}} f(v) + \mathcal{O}(\epsilon),$$

with

$$\xi_1 = 2\pi \int_0^{\epsilon/2} t \sigma(t) dt \sim 2\pi \int_0^{\infty} t \sigma(t) dt. \quad (46)$$

In other words, when ϵ is small, the linear scattering operator \mathcal{L} converges to the Fokker-Planck operator with a scalar multiplication and the linear transport equation converges to the Fokker-Planck equation

$$\partial_t f + v \cdot \nabla_x f = \frac{\xi_1}{2} \mathcal{L}_{\text{FP}} f,$$

where \mathcal{L}_{FP} and ξ_1 are defined in (45) and (46), respectively.

Remark 4. We note that the unknown in the collision term defined in (35) is $\sigma(\mu)$. As a function of μ , it could be fully recovered only if all the coefficients σ_n in (37) are known. According to (42), this requires knowledge about ξ_n for all n . However, in the zero limit, the collision term converges to the Laplace operator, and there is only one scalar that is unknown: ξ_1 . As a result, the limiting Fokker-Planck equation is much easier to invert heuristically. This will be reflected in section 3.3.

3.2. Inverse problem setup. In the inverse problem setting, we are given inflow data and measure the outflow, with which we infer the scattering coefficient $\sigma(v, v')$. Here the albedo operator is given by:

$$\mathcal{A}(\sigma) : \quad \phi(t, x, v)|_{(0,T) \times \Gamma_-} \rightarrow \int_{\Gamma_+(y)} f(t, y, v) n(y) \cdot v dv.$$

We first linearize the albedo operator. Like always, we assume that a priori knowledge provides a background state σ_0 such that the residue $\tilde{\sigma} := \sigma - \sigma_0$ satisfies

$$|\tilde{\sigma}| \ll |\sigma|, \quad \text{a.s.},$$

then with background state σ_0 , one gets the solution f_0 that solves the following initial boundary value problem

$$\begin{cases} \partial_t f_0(t, x, v) + v \cdot \nabla_x f_0(t, x, v) = \mathcal{L}_0 f_0(t, x, v), \\ f_0(0, x, v) = 0 \quad \text{on } \Omega \times \mathbb{S}^2, \\ f_0(t, x, v) = \phi(t, x, v) \quad \text{on } (0, T) \times \Gamma_-, \end{cases} \quad (47)$$

where

$$\mathcal{L}_0 f_0(t, x, v) = \frac{1}{\epsilon^2} \int_{4\pi} \sigma_0 \left(\frac{1-\mu}{\epsilon} \right) (f_0(t, x, v') - f_0(t, x, v)) dv'.$$

The residue

$$\tilde{f}(t, x, v) := f(t, x, v) - f_0(t, x, v)$$

then satisfies

$$\partial_t \tilde{f}(t, x, v) + v \cdot \nabla_x \tilde{f}(t, x, v) = \mathcal{L}_0 \tilde{f}(t, x, v) + \tilde{\mathcal{L}} f_0(t, x, v) \quad (48)$$

with zero initial data and boundary data. Here

$$\tilde{\mathcal{L}} f_0(t, x, v) = \frac{1}{\epsilon^2} \int_{4\pi} \tilde{\sigma} \left(\frac{1-\mu}{\epsilon} \right) (f_0(t, x, v') - f_0(t, x, v)) dv'. \quad (49)$$

We also define an adjoint problem to (47):

$$\begin{cases} -\partial_t g(t, x, v) - v \cdot \nabla_x g(t, x, v) = \mathcal{L}_0 g(t, x, v), \\ g(T, x, v) = 0 \quad \text{on } \Omega \times \mathbb{S}^2, \\ g(t, x, v) = \delta(\tau, y) \quad \text{on } (0, T) \times \Gamma_+. \end{cases} \quad (50)$$

Multiply (48) and (50) by g and \tilde{f} respectively, and subtract them, we get, after integrating in x, v and t ,

$$\int_{\Gamma_+(y)} \tilde{f}(\tau, y, v) n(y) \cdot v dv = \int_{\Omega \times \mathbb{S}^2} \int_0^T g(t, x, v) \tilde{\mathcal{L}} f_0(t, x, v) dt dv dx, \quad (51)$$

where the LHS is the difference between measurement of $f(t, x, v)$ and the computed $f_0(t, x, v)$ at time τ and position $y \in \partial\Omega$, and we denote it by $b(\tau, y, \phi)$

$$b(\tau, y, \phi) = \int_{\Gamma_+(y)} \tilde{f}(\tau, y, v) n(y) \cdot v dv.$$

The RHS of (51) gives a linear function for $\tilde{\sigma}$. In particular, using (49) we have

$$b(\tau, y, \phi) = \frac{1}{\epsilon^2} \int_{\mathbb{S}^2 \times \mathbb{S}^2} \tilde{\sigma} \left(\frac{1-\mu}{\epsilon} \right) \gamma_\epsilon(v, v') dv' dv, \quad (52)$$

where

$$\gamma_\epsilon(v, v'; \tau, y, \phi) := \int_0^T \int_{\Omega} g(t, x, v) [f_0(t, x, v') - f_0(t, x, v)] dx dt. \quad (53)$$

By varying $\{\tau, y\}$ and ϕ , one obtains different g and f_0 , and thus γ_ϵ , making (52) a Fredholm operator of first kind with parameters $\{\tau, y, \phi\}$.

As mentioned in Remark 2, to have a unique recovery of $\tilde{\sigma}$ in L_p space, one needs γ_ϵ expanding the adjoint space L_q (with $\frac{1}{p} + \frac{1}{q} = 1$). The injectivity is beyond the scope of the current paper, and we only discuss the stability in the following section.

3.3. Stability in the highly forward peaked regime. In this section we study the stability in the recovery of σ in the forward peaked regime. There are two aspects of the problem:

1. To fully recover $\sigma(1 - v \cdot v')$, as mentioned in Remark 4, one needs all its moments σ_n , which in turn requires the information of ξ_n for all n . However, since ξ_n diminishes at the order of ϵ^{n-1} , obtaining ξ_n is very sensitive to the pollution in the data. Indeed, suppose the data has pollution of order δ , then there are at most $n_0 = \log_\epsilon \delta + 1$ terms that can be recovered. Now keeping δ fixed and sending ϵ to 0, the number n_0 decreases to 1, meaning that all the higher order information get lost. This is indeed consistent with the view of the singular decomposition [14, 5]. In that viewpoint, the reconstruction of σ relies on the separation of the ballistic component (pure transport) and the scattered components (mainly the single-scattering). In the forward peaked regime, however, the single-scattering concentrates on the original velocity and does not distinguish from the ballistic transport much, making the separation hard, and thus disables the reconstruction. This will be demonstrated in Theorem 5.
2. Nevertheless, in the Fokker-Planck regime, it's not the full information $\sigma(1 - v \cdot v')$ that matters, but the rescaled one $\frac{1}{\epsilon}\sigma\left(\frac{1-v \cdot v'}{\epsilon}\right)$. As written, when ϵ is small, the rescaled σ will concentrate around $v \cdot v' = 1$ and only a little information is needed to recover its shape. Indeed, according to (40), (41) and (42), ξ_n quickly decays to zero, and all of σ_n are dominated by the first few ξ_n for certain accuracy. For example, if ϵ accuracy is needed for σ_n , one only needs to recover one parameter ξ_1 . This significantly reduces the amount of measurements needed. In this sense, we find that the inverse problem with highly forward peaked scattering is actually more practice friendly. This is demonstrated in Theorem 4.

To recover $\frac{1}{\epsilon}\tilde{\sigma}_s\left(\frac{1-v \cdot v'}{\epsilon}\right)$, one simply needs to find all its Legendre coefficients $\tilde{\sigma}_n$ in the expansion

$$\frac{1}{\epsilon}\tilde{\sigma}\left(\frac{1-v \cdot v'}{\epsilon}\right) = \sum_{n=0}^{\infty} \tilde{\sigma}_n P_n(\mu), \quad \mu = v \cdot v'. \quad (54)$$

Using the same expression as in (42), one has:

$$\tilde{\sigma}_n = \varepsilon \left[P_n(1)\tilde{\xi}_0 + P'_n(1)\tilde{\xi}_1 + \frac{1}{2}P''_n(1)\tilde{\xi}_2 + \frac{1}{3!}P'''_n(1)\tilde{\xi}_3 + \cdots \right], \quad (55)$$

with

$$\tilde{\xi}_n = \epsilon^{n-1} 2\pi \int_0^{\epsilon/2} t^n \tilde{\sigma}_s(t) dt = \mathcal{O}(\epsilon^{n-1}). \quad (56)$$

Introducing the above relations into (52), we get

$$\begin{aligned} b(\tau, y, \phi) &= \frac{1}{\epsilon} \sum_{n=0}^{\infty} \tilde{\sigma}_{sn} \int_{\mathbb{S}^2} P_n(\mu) \gamma_\epsilon(v, v') dv dv' \\ &= \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} \xi_j P_n^{(j)}(1) \frac{1}{j!} \int_{\mathbb{S}^2} P_n(\mu) \gamma_\epsilon(v, v') dv dv' \\ &= \sum_{j=0}^{\infty} \xi_j \left(\frac{1}{j!} \sum_{n=0}^{\infty} P_n^{(j)}(1) \int_{\mathbb{S}^2} P_n(\mu) \gamma_\epsilon(v, v') dv dv' \right). \end{aligned}$$

Consequently, we obtain the following linear system for $\xi = (\xi_1, \xi_2, \dots)$

$$A\xi = \mathbf{b}, \quad (57)$$

where \mathbf{b} is a column vector whose size is equal to the number of experiments, and in matrix $\mathbf{A} = [a_{ij}]$, the component a_{ij} is determined by

$$a_{ij} = \frac{1}{j!} \sum_{n=0}^{\infty} P_n^{(j)}(1) \int_{\mathbb{S}^2} P_n(\mu) \gamma_{\epsilon}(v, v') dv dv',$$

where the subscript i represent experiment i with choice τ_i, y_i, ϕ_i . Recall the expression of γ_{ϵ} in (53), one has

$$\begin{aligned} \int_{\mathbb{S}^2} P_n(\mu) \gamma_{\epsilon}(v, v') dv dv' &= \frac{4\pi}{2n+1} \sum_{m=-n}^n \int Y_{n,m}(v) Y_{n,m}(v') \gamma_{\epsilon}(v, v') dv dv' \\ &= \frac{4\pi}{2n+1} \sum_{m=-n}^n \int Y_{n,m}(v) Y_{n,m}(v') \left[\int g(v) f_0(v') - g(v') f_0(v) dx dt \right] dv dv' \\ &= \frac{4\pi}{2n+1} \sum_{m=-n}^n \left[\bar{g}_{n,m}(\bar{f}_0)_{n,m} - (\bar{g} \bar{f}_0)_{n,m} \delta_{n,m} \right], \end{aligned}$$

where the over-line denotes integration in both x and t . Therefore,

$$a_{ij} = \sum_{n=0}^{\infty} \frac{P_n^{(j)}}{j!} \frac{4\pi}{2n+1} \sum_{m=-n}^n \left[\bar{g}_{n,m}(\bar{f}_0)_{n,m} - (\bar{g} \bar{f}_0)_{n,m} \delta_{n,m} \right]. \quad (58)$$

Theorem 4. *The recovery of $\frac{1}{\epsilon} \tilde{\sigma} \left(\frac{1-\mu}{\epsilon} \right)$ does not deteriorates as $\epsilon \rightarrow 0$. More precisely, if we define the distinguishability coefficient as*

$$\kappa_{\epsilon} = \sup_{\sigma \in \Gamma_{\delta}} \frac{\left\| \frac{1}{\epsilon} \sigma \left(\frac{1-\mu}{\epsilon} \right) - \frac{1}{\epsilon} \tilde{\sigma} \left(\frac{1-\mu}{\epsilon} \right) \right\|_{\infty}}{\left\| \frac{1}{\epsilon} \tilde{\sigma} \left(\frac{1-\mu}{\epsilon} \right) \right\|_{\infty}}, \quad (59)$$

where

$$\Gamma_{\delta} = \left\{ \frac{1}{\epsilon} \sigma \left(\frac{1-\mu}{\epsilon} \right) : \sup_{\substack{\forall \|\phi\|_{L^{\infty}(\Gamma_{-})} \leq 1, \\ \forall y \in \partial\Omega, \tau \in [0, T]}} \left| \frac{1}{\epsilon} \int \left[\frac{1}{\epsilon} \sigma \left(\frac{1-\mu}{\epsilon} \right) - \frac{1}{\epsilon} \tilde{\sigma} \left(\frac{1-\mu}{\epsilon} \right) \right] \gamma_{\epsilon}(v, v'; \tau, y, \phi) dv dv' \right| \leq \delta \right\},$$

then for $\epsilon \ll \delta$

$$\kappa_{\epsilon} \sim \mathcal{O}(\delta \epsilon).$$

Proof. From (54)–(56), we see that

$$\frac{1}{\epsilon} \tilde{\sigma} \left(\frac{1-\mu}{\epsilon} \right) = \epsilon \sum_{k=0}^{\infty} \left[\sum_{n=0}^{\infty} \frac{2n+1}{4\pi} P_n(\mu) \frac{P_n^{(k)}(1)}{k!} \right] \tilde{\xi}_k := \epsilon \sum_{k=0}^{\infty} h_k(\mu) \tilde{\xi}_k,$$

where $\tilde{\xi}_k$ is defined the same as in (56). Similarly, $\frac{1}{\epsilon} \sigma \left(\frac{1-\mu}{\epsilon} \right)$ has the expansion

$$\frac{1}{\epsilon} \sigma \left(\frac{1-\mu}{\epsilon} \right) = \epsilon \sum_{k=0}^{\infty} h_k(\mu) \xi_k.$$

Note firstly that $\tilde{\xi}_0 = \xi_0$, then for $\frac{1}{\epsilon} \sigma \left(\frac{1-\mu}{\epsilon} \right) \in \Gamma_{\delta}$, we have

$$\begin{aligned} \left| \frac{1}{\epsilon} \int \left[\frac{1}{\epsilon} \sigma \left(\frac{1-\mu}{\epsilon} \right) - \frac{1}{\epsilon} \tilde{\sigma} \left(\frac{1-\mu}{\epsilon} \right) \right] \gamma_{\epsilon}(v, v'; \tau, y, \phi) dv dv' \right| &= \left| \sum_{k=1}^{\infty} \int h_k(\mu) (\xi_k - \tilde{\xi}_k) \gamma_{\epsilon} dv dv' \right| \\ &= \left| \int h_1(\mu) (\xi_1 - \tilde{\xi}_1) \gamma_{\epsilon} dv dv' + \mathcal{O}(\epsilon) \right| \leq \delta, \end{aligned}$$

which implies that, for $\epsilon \ll \delta$, $|\xi_1 - \tilde{\xi}_1| \leq \mathcal{O}(\delta)$ since $h_1(\mu)$ and γ_ϵ are $\mathcal{O}(1)$. Plugging this result into the definition (59), one immediately sees that

$$\kappa_\epsilon = \sup_{\sigma \in \Gamma_\delta} \frac{(\xi_1 - \tilde{\xi}_1)h_1(\mu) + \mathcal{O}(\epsilon^2)}{\xi_0 h_0(\mu) + \xi_1 h_1(\mu) + \mathcal{O}(\epsilon^2)} \sim \mathcal{O}(\delta\epsilon),$$

where ϵ on the right comes from the fact that $\xi_0 \sim \mathcal{O}(\epsilon^{-1})$. \square

On the contrary of the above result, if we want to fully recover σ , then the presence of the small scale ϵ will make it impossible. Specifically, we have the following theorem.

Theorem 5. *Suppose $\sigma \in H_k(d\mu)$ (assuming k -th regularity in σ), the recovering σ becomes impossible in the limit of $\epsilon \rightarrow 0$ in the sense that if δ error is allowed in \mathbf{b} (e.g., measurement error, see (57)), then the error in σ will be $\left(\frac{\ln \epsilon}{\ln \delta}\right)^k$.*

Proof. Since $\sigma(\mu) \in L_2(d\mu)$, we expand it using Hermite functions:

$$\sigma = \sum_n \frac{1}{n!} \hat{\sigma}_n H_n(\mu) \quad (60)$$

with $\hat{\sigma}_n = \langle \sigma, H_n \rangle = \int \sigma H_n(\mu) d\mu$. Here H_n are weighted Hermite functions written as:

$$H_n(\mu) = p_n(\mu) e^{-\mu^2/2}, \quad \text{with} \quad \int p_m(\mu) p_n(\mu) e^{-\mu^2} d\mu = \delta_{mn} n!.$$

Note that other L_2 basis functions can be used. Hermite polynomial is only one possible choice.

Meanwhile we recall definition of ξ :

$$\xi_m = 2\pi \epsilon^{m-1} \int_0^{2/\epsilon} \sigma(\mu) \mu^m d\mu.$$

To prove the theorem, we allow Δ error in σ , and see how much it affects ξ , and then \mathbf{b} in the end. Here $\xi = [\xi_0, \xi_1, \dots]'$ and $\sigma = [\sigma_0, \sigma_1, \dots]'$. We first note that $\sigma \in H_k$, and thus its Hermite polynomial coefficients decay algebraically fast. The standard approximation theory from spectral accuracy indicates:

$$\sigma_n = \mathcal{O}(1/n^k).$$

Suppose we tolerant error up to Δ , then one needs to recover σ_n up to at least $n_0 = \Delta^{-1/k}$, and the allowed perturbation in σ_n is:

$$\Delta_{\sigma_n} \leq \Delta \quad \text{for} \quad n = 0, \dots, n_0.$$

We now look for explicit relation between ξ and σ_n . Considering the explicit relation between monomials and the Hermite polynomials

$$v^m = m! \sum_{k=0}^{m/2} \frac{1}{2^k k! (m-2k)!} p_{m-2k}(v),$$

plugging it back in the equation for ξ_m , we have

$$\begin{aligned} \xi_m &= \epsilon^{m-1} \sum_{k=0}^{m/2} \frac{2\pi m!}{2^k k! (m-2k)!} \int_0^{2/\epsilon} \sigma(v) p_{m-2k}(v) dv \\ &= \epsilon^{m-1} \sum_{k=0}^{m/2} \sum_{n=0}^{\infty} \frac{2\pi m!}{2^k k! (m-2k)! n!} D_{n,m,k} \sigma_n, \end{aligned}$$

where we have used expansion in (60) and defined $D_{n,m,k} = \int p_{m-2k} p_n e^{-v^2/2} dv$. In a matrix form one has:

$$\xi = \mathbf{C} \cdot \sigma,$$

where \mathbf{C} defined by:

$$C_{mn} = \epsilon^{m-1} \frac{2\pi m!}{n!} \sum_{k=0}^{m/2} \frac{D_{n,m,k}}{2^k k! (m-2k)!}.$$

Since we need to recover σ_n up to $n = n_0 = \Delta^{-1/k}$, and the recovered coefficients need to be within error tolerance δ , the tolerance for ξ_{n_0} then is:

$$\Delta_{\xi_{n_0}} \leq \Delta \sum_{k=0}^{n_0} C_{n_0,k} \sim \epsilon^{n_0-1} \Delta, \quad \text{with } n_0 = \Delta^{-1/k}.$$

Noting the relationship between ξ and \mathbf{b} in (57), we see that the error allowance on \mathbf{b} is $\epsilon^{n_0-1} \Delta$. With shrinking ϵ , this restriction becomes more and more severe, making the inverse problem less practical. More specifically if we have δ error in \mathbf{b} , then setting:

$$\epsilon^{\Delta^{-1/k}} \Delta \sim \delta$$

gives $\Delta > \left(\frac{\ln \epsilon}{\ln \delta}\right)^k \rightarrow \infty$ as $\epsilon \rightarrow 0$, meaning that the accuracy in recovering σ_n is lost and so it is with σ . \square

REFERENCES

- [1] G. ABDOULAEV, K. REN, AND A. HIELSCHER, *Optical tomography as a PDE-constrained optimization problem*, Inverse Problems, 21 (2005), pp. 1507–1530.
- [2] S. ARRIDGE, *Optical tomography in medical imaging*, Inverse Problems, 15 (1999), pp. R41–93.
- [3] S. ARRIDGE AND W. LIONHEART, *Nonuniqueness in diffusion-based optical tomography*, Opt. Lett., 23 (1998), pp. 882–884.
- [4] S. R. ARRIDGE AND J. C. SCHOTLAND, *Optical tomography: forward and inverse problems*, Inverse Problems, 25 (2009), p. 123010.
- [5] G. BAL, *Inverse transport theory and applications*, Inverse Problems, 25 (2009), p. 053001.
- [6] G. BAL AND A. JOLLIVET, *Time-dependent angularly averaged inverse transport*, Inverse Problems, 25 (2009), p. 075010.
- [7] ———, *Stability for time-dependent inverse transport*, SIAM J. Math. Anal., 42 (2010), pp. 679–700.
- [8] G. BAL, I. LANGMORE, AND F. MONARD, *Inverse transport with isotropic sources and angularly averaged measurement*, Inverse Probl. Imaging, 2 (2008), pp. 23–42.
- [9] C. BARDOS, S. SANTOS, AND R. SENTIS, *Diffusion approximation and computation of the critical size*, Trans. Amer. Math. Soc., 284 (1984), pp. 617–649.
- [10] K. M. CASE AND P. ZWEIFEL, *Linear Transport Theory*, Addison-Wesley, 1967.
- [11] K. CHEN, Q. LI, AND L. WANG, *Stability of stationary inverse transport equation in diffusion scaling*, arXiv: 1703.00097, (2017).
- [12] Y. CHENG, I. M. GAMBA, AND K. REN, *Recovering doping profiles in semiconductor devices with the Boltzmann-Poisson model*, J. Comput. Phys., 230 (2011), pp. 3391–3412.
- [13] M. CHOULLI AND P. STEFANOV, *Inverse scattering and inverse boundary value problems for the linear Boltzmann equation*, Comm. P.D.E, 21 (1996), pp. 763–785.
- [14] ———, *An inverse boundary value problem for the stationary transport equation*, Osaka J. Math., 36 (1998), pp. 87–104.
- [15] R. DAUTRAY AND J.-L. LIONS, *Mathematical analysis and numerical methods for science and technology*, Springer, Berlin, 1993.
- [16] A. GREENLEAF, M. LASSAS, AND G. UHLMANN, *On nonuniqueness for Calderón inverse problem*, Mathematical Research Letters, 10 (2003), pp. 685–693.
- [17] G. UHLMANN, *Electrical impedance tomography and Calderón’s problem*, Inverse Problems, 25 (2009), p. 123011.
- [18] E. LARSEN, *Solution of three-dimensional inverse transport problems*, Transport Theory and Stat. Phys., 17 (1988), pp. 147–167.
- [19] E. LARSEN AND J. KELLER, *Asymptotic solution of neutron transport problems for small mean free paths*, J. Math. Phys., 15 (1974), pp. 75–81.
- [20] C. L. LEAKEAS AND E. W. LARSEN, *Generalized Fokker-Planck approximations of particle transport with highly forward-peaked scattering*, Nuclear Science and Engineering, 137 (2001), pp. 236–250.
- [21] M. MACHIDA, G. Y. PANASYUK, Z.-M. WANG, V. A. MARKEL, AND J. C. SCHOTLAND, *Radiative transport and optical tomography with large datasets*, J. Opt. Soc. Am. A, 33 (2016), pp. 551–558.
- [22] M. MACHIDA AND J. C. SCHOTLAND, *Inverse Born series for the radiative transport equation*, Inverse Problems, 31 (2015), p. 095009.

- [23] G. POMRANING, *The Fokker-Planck operator as an asymptotic limit*, Mathematical Models and Methods in Applied Sciences, 02 (1992), pp. 21–36.
- [24] K. REN, *Recent developments in numerical techniques for transport-based medical imaging methods*, Comm. Comput. Phys, 8 (2010), pp. 1–50.
- [25] K. REN, G. BAL, AND A. H. HIELSCHER, *Transport- and diffusion-based optical tomography in small domains: a comparative study*, Appl. Opt., 46 (2007), pp. 6669–6679.
- [26] V. ROMANOV, *Stability estimates in problems of recovering the attenuation coefficient and the scattering indicatrix for the transport equation*, J. Inverse Ill-Posed Probl., 4 (1996), pp. 297–305.
- [27] T. SARATOON, T. TARVAINEN, B. T. COX, AND S. R. ARRIDGE, *A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation*, Inverse Problems, 29 (2013), p. 075006.
- [28] P. STEFANOV AND A. TAMASAN, *Uniqueness and non-uniqueness in inverse radiative transfer*, Proceedings of the American Mathematical Society, 137 (2009), pp. 2335–2344.
- [29] J. TANG, W. HAN, AND B. HAN, *A theoretical study for RTE-based parameter identification problems*, Inverse Problems, 29 (2013), p. 095002.
- [30] J. WANG, *Stability estimates of an inverse problem for the stationary transport equation*, Ann. Inst. Henri Poincaré, 70 (1999), pp. 473–495.

MATHEMATICS DEPARTMENT, UNIVERSITY OF WISCONSIN-MADISON, 480 LINCOLN DR., MADISON, WI 53705 USA.
E-mail address: ke@math.wisc.edu

MATHEMATICS DEPARTMENT AND WISCONSIN INSTITUTE OF DISCOVERY, UNIVERSITY OF WISCONSIN-MADISON, 480 LINCOLN DR., MADISON, WI 53705 USA.
E-mail address: qinli@math.wisc.edu

DEPARTMENT OF MATHEMATICS, COMPUTATIONAL AND DATA-ENABLED SCIENCE AND ENGINEERING PROGRAM, STATE UNIVERSITY OF NEW YORK AT BUFFALO, 244 MATHEMATICS BUILDING, BUFFALO, NY 14206 USA.
E-mail address: lwang46@buffalo.edu