MATHEMATICS OF COMPUTATION
Volume 89, Number 324, July 2020, Pages 1605–1646
https://doi.org/10.1090/mcom/3505
Article electronically published on January 8, 2020

AN ADAPTIVE NESTED SOURCE TERM ITERATION FOR RADIATIVE TRANSFER EQUATIONS

WOLFGANG DAHMEN, FELIX GRUBER, AND OLGA MULA

Abstract. We propose a new approach to the numerical solution of radiative transfer equations with certified a posteriori error bounds for the L_2 norm. A key role is played by stable *Petrov–Galerkin*-type variational formulations of parametric transport equations and corresponding radiative transfer equations. This allows us to formulate an iteration in a suitable, infinite-dimensional function space that is guaranteed to converge with a fixed error reduction per step. The numerical scheme is then based on approximately realizing this iteration within dynamically updated accuracy tolerances that still ensure convergence to the exact solution. To advance this iteration two operations need to be performed within suitably tightened accuracy tolerances. First, the global scattering operator needs to be approximately applied to the current iterate within a tolerance comparable to the current accuracy level. Second, parameter dependent linear transport equations need to be solved, again at the required accuracy of the iteration. To ensure that the stage dependent error tolerances are met, one has to employ *rigorous a posteriori error bounds* which, in our case, rest on a *Discontinuous Petrov–Galerkin* (DPG) scheme. These a posteriori bounds are not only crucial for guaranteeing the convergence of the perturbed iteration but are also used to generate adapted parameter dependent spatial meshes. This turns out to significantly reduce overall computational complexity. Since the global operator is only applied, we avoid the need to solve linear systems with densely populated matrices. Moreover, the approximate application of the global scatterer is accelerated through low-rank approximation and matrix compression techniques. The theoretical findings are illustrated and complemented by numerical experiments with non-trivial scattering kernels.

1. Introduction

When dealing with problems giving rise to very complex discretizations, one often tacitly *assumes* that the numerical output represents the corresponding continuous object reasonably well, without being, however, able to actually quantify output quality in any rigorous sense. Often interest shifts then towards accurately solving the (fixed) discrete problem which by itself may indeed pose enormous challenges. Instead, the central objective of this article is to put forward a new algorithmic paradigm warranting *error controlled computation*. By this we mean the deviation

Received by the editor November 20, 2018, and, in revised form, October 3, 2019, and October 22, 2019.

2010 *Mathematics Subject Classification.* Primary 65N12, 65N15, 65N30; Secondary 65N50.

Key words and phrases. DPG transport solver, iteration in function space, fast application of scattering operator, Hilbert–Schmidt decomposition, matrix compression, a posteriori bounds, kinetic problems, linear Boltzmann, radiative transfer.

The research of the first author was supported by the NSF Grant DMS 1720297, and by the SmartState and Williams-Hedberg Foundation.

The third author is the corresponding author.

of the numerical result from the *exact continuous solution* is certifiably quantified and set to meet a given target accuracy with respect to a *problem relevant norm*. It goes without saying that the ability to quantify the accuracy of forward simulations is a necessary prerequisite of Uncertainty Quantification in general. In this article we develop such methods for a regime of kinetic models, described below, for which to the best of our knowledge error controlled schemes have so far not been available yet.

1.1. **Problem formulation.** We consider certain *kinetic models* describing the propagation of particles in a collisional medium modeling, e.g., heat transfer phenomena, neutron transport, or medical imaging processes. We confine the subsequent discussion to simple *monoenergetic radiative transfer models* which nevertheless exhibit the main obstructions to the design of efficient numerical methods for this problem class. Let $\mathrm{D} \subset \mathbb{R}^d$ be a bounded convex domain with piecewise C^1 boundary $\partial\mathrm{D}$, where $d \geqslant 1$. Hence, for almost all $x \in \partial\mathrm{D}$ the outward normal $\mathbf{n} = \mathbf{n}(x)$ is well defined. Furthermore, let $\mathrm{S} \subset \mathbb{R}^d$ denote the unit $(d-1)$-sphere representing the directions in which particles propagate. Since we focus on the monoenergetic case, the particles have all the same kinetic energy (which we assume to be equal to 1) but note that more general compact sets describing the admissible transport velocity field are possible and the subsequent developments generalize to a correspondingly wider scope of setups. In what follows, for $\vec{s} \in \mathrm{S}$

$$\Gamma_-(\vec{s}) := \{x \in \partial\mathrm{D} \mid \vec{s} \cdot \mathbf{n}(x) < 0\} \subset \partial\mathrm{D} \tag{1.1}$$

denotes the "inflow-boundary" for the given direction $\vec{s}$ while

$$\Gamma_- := \{(x,\vec{s}) \in \partial\mathrm{D} \times \mathrm{S} \mid \vec{s} \cdot \mathbf{n}(x) < 0\} \subset \partial\mathrm{D} \times \mathrm{S}$$

denotes the inflow portion of the corresponding space-direction cylinder. The corresponding outflow boundary portions $\Gamma_+(\vec{s})$, Γ_+ are defined analogously.

Given non-negative data $f\colon \mathrm{D} \times \mathrm{S} \to \mathbb{R}_+$, $g\colon \Gamma_- \to \mathbb{R}_+$, a cross-section function $\sigma\colon \mathrm{D} \times \mathrm{S} \to \mathbb{R}_+$, and a collision kernel $K\colon \mathrm{D} \times \mathrm{S} \times \mathrm{S} \to \mathbb{R}_+$, we want to find a function $u\colon \mathrm{D} \times \mathrm{S} \to \mathbb{R}_+$, satisfying

(1.2)

$$\begin{aligned} \vec{s} \cdot \nabla u(x,\vec{s}) + \sigma(x,\vec{s})u(x,\vec{s}) - \int_{\mathrm{S}} K(x,\vec{s}',\vec{s})u(x,\vec{s}')\,\mathrm{d}\vec{s}' &= f(x,\vec{s}) \qquad \forall (x,\vec{s}) \in \mathrm{D} \times \mathrm{S}, \\ u &= g \qquad \text{on } \Gamma_-. \end{aligned}$$

In the following, it will be useful to view the angular direction as a parameter and introduce the abbreviations

$$(\mathcal{T}_{\vec{s}}u)(x) := \vec{s} \cdot \nabla u(x,\vec{s}) + \sigma(x,\vec{s})u(x,\vec{s}), \quad (\mathcal{K}_{\vec{s}}u)(x) := \int_{\mathrm{S}} K(x,\vec{s}',\vec{s})u(x,\vec{s}')\,\mathrm{d}\vec{s}',$$

for the pure transport and collision operator, respectively. Splitting the transport part into

$$\mathcal{T}_{\vec{s}} = \mathcal{A}_{\vec{s}} + \sigma\mathrm{id}, \quad \mathcal{A}_{\vec{s}}v := \vec{s} \cdot \nabla v,$$

(1.2) can be written, for homogeneous boundary data $g \equiv 0$, as the operator equation

$$(\mathcal{B}u)(\cdot,\vec{s}) := \mathcal{T}_{\vec{s}}u - \mathcal{K}_{\vec{s}}u = \mathcal{A}_{\vec{s}}u + \sigma u - \mathcal{K}_{\vec{s}}u = f(\cdot,\vec{s}). \tag{1.3}$$

There is extensive literature addressing the solvability of (1.3) depending on the interrelation of the pair (σ, K) usually known as the *optical parameters*, see e.g., [5,14,16,25]. One may roughly distinguish two ends of the problem scope, namely

the case of *dominating scattering* near the diffusive limit (see, e.g., [20]), and the case of *dominating transport*. Here we restrict the subsequent considerations to the latter regime that is governed by at least weakly dominating transport and possibly anisotropic scattering. The precise conditions on corresponding pairs of optical parameters are discussed in a later section.

Note that when the kernel K vanishes the *pure transport problems*

$$\mathcal{T}_{\vec{s}} u = f, \quad u|_{\Gamma_-(\vec{s})} = g, \ \vec{s} \in \mathrm{S}, \tag{1.4}$$

may be viewed as a *parametric family* of PDEs giving rise to the corresponding family of *fiber solutions* $u_{\vec{s}}$, $\vec{s} \in \mathrm{S}$. Alternatively—and this is necessary for the full problem (1.3)—we can view solutions $u(x, \vec{s})$ as functions of the spatial variable $x \in \mathrm{D} \subset \mathbb{R}^d$ and the parametric variable $\vec{s} \in \mathrm{S} \subset \mathbb{R}^{d-1}$. It will therefore be important to identify a function space U consisting of functions over $\mathrm{D} \times \mathrm{S}$ for which (1.3) is well-posed in a sense to be made precise in Section 2.1.

1.2. **Common approaches and main obstructions.** There are at least two major groups of numerical strategies for approximately solving (1.3), namely the *method of moments* and the *discrete ordinates method* (DOM); see, e.g., [24] and [4,17,23,26], respectively. The method of moments builds on (low order) polynomial projections in the parameter domain and can be viewed as a model reduction. It seems to be rather difficult though to quantify the incurred model bias and develop rigorous error bounds for the deviation of the approximate solution from the exact one. Also, the accuracy of polynomial expansions suffers severely from low regularity. DOM hinges on transport solves for sufficiently many direction parameters. These can serve as quadrature nodes for the approximate application of the integral operator in combination with Jacobi-type iterations to approximately solve the very large densely populated linear systems. However, the convergence of this iteration in the discrete setting typically degrades with increasing dominance of the scatterer [23].

The common approach is to *first* discretize the (continuous) problem and then address the two—at first unrelated—issues: a) how to solve the (fixed) discrete problem efficiently; b) how to assess the accuracy attained by the solution of the discrete problem.

Modern strategies to face the complexity issues posed by a) concern the development of preconditioners or multigrid strategies or employ sparse tensor methods based on sparse grid or hyperbolic cross approximations. The former issue is impeded by the the fact that on a fixed discrete level it is hard to respect *intrinsic problem metrics* which play a central role in the current approach. Moreover, the distinct lack of sufficiently strong stability notions accounts, in particular, for increasing recent efforts to incorporate additional *structure preserving properties* into discrete concepts. Simple examples are non-negativity or mass conservation.

The viability and performance of sparse tensor methods, in turn, requires suitable a priori regularity assumptions such as the validity of a certain order of *mixed smoothness*, see, e.g., [2,3,17,22], which are then also invoked to address b).

In general, variational formulations for parametric transport problems like (1.4) or (1.3) are far less common than for elliptic problems. For instance, [17] considers least squares formulations minimizing residuals in $L_2(\mathrm{D} \times \mathrm{S})$. Corresponding trial

spaces require anisotropic regularity of the solution depending explicitly and sensitively on the transport direction. This may cause stability problems when the solution exhibits shear discontinuities. Alternatively, [15] proposes a mixed Galerkin formulation based on splitting the solution into symmetric and asymmetric parts. This still fails to tightly relate errors to residuals which is a key prerequisite for rigorous a posteriori error estimates.

We summarize now some of the intrinsic obstructions to an efficient and accuracy controlled numerical solution of such problems.

(1) The solution u of (1.3) is a function of $2d-1$ variables (or even more in non-stationary cases and realistic models involving energy levels). Hence, the problem is *high-dimensional* and standard schemes become possibly prohibitively inefficient.
(2) A non-trivial scattering kernel K would give rise to densely populated very large system matrices when using standard discretizations based on localization only.
(3) These obstructions are aggravated by the fact that solutions exhibit in general only a low degree of regularity, in particular, when dealing with highly concentrated and non-smooth boundary data. Standard a priori error estimates involving classical isotropic Sobolev regularity scales, often derived under unrealistic assumptions, are therefore not very useful for controlling accuracy.

The primary objective of this paper is to address the above issues and develop accuracy controlled schemes and corresponding stability notions. We confine the discussion to *stationary* problems but remark that the concepts carry over to time-dependent problems. In fact, unsteady problems become conceptually easier as it will become clear later (aside from having to deal with even more variables).

The numerical results in Section 7 indicate that the proposed stability concept, closely intertwining the continuous and discrete setting, produces meaningful physical results without explicitly imposing additional structure preserving measures.

1.3. **Conceptual roadmap.** The approach proposed in this paper is based on the following steps:

(I) Identify a pair of Hilbert spaces U, V over $\mathrm{D}\times\mathrm{S}$ for which (1.2) permits a *stable variational formulation* (see Section 2.1 for the precise meaning) where the (infinite-dimensional) trial space is to accommodate the solution of (1.2). Stability means that this variational formulation identifies the operator $\mathcal{B}$ in (1.3) as an *isomorphism* from U onto the dual V' of the (infinite-dimensional) test space V.

(II) Contrive an "ideal outer iteration"

$$u_{n+1} = u_n + \mathcal{P}(f - \mathcal{B}u_n), \quad n = 0, 1, 2, \dots, \tag{1.5}$$

that converges in U to the unique solution u of (1.2).

(III) Realize each iteration step approximately within dynamically updated error tolerances that are judiciously chosen so as to guarantee convergence of the perturbed iteration to the exact (infinite-dimensional) solution u of (1.2).

Steps (I) and (II) require analytic preparations which the numerical method is based upon while numerical aspects only enter in Step (III). The contributions of this paper culminate in Theorem 4.1, which we informally state here as follows.

Main contribution. *We contrive and theoretically justify a numerical algorithm that realizes Step* (III) *of the roadmap and prove that for any target accuracy* $\varepsilon > 0$ *it generates an approximate solution* u_ε *of* (1.3) *that deviates from the exact solution in* $L_2(\mathrm{D}\times\mathrm{S})$ *by at most* ε. *Since the algorithm progresses from coarse to successively finer accuracy levels termination at any stage comes with a current error certificate.*

This program relies on two points that guide the subsequent discussions. At *no stage* is there ever formulated beforehand any fixed discrete problem but discretizations are formed *adaptively* at each stage of the (perturbed) outer iteration (1.5). For this to work it is crucial that the accuracy of a current approximate solution can be rigorously quantified. The perhaps closest relative to the above roadmap are adaptive wavelet methods along the lines of [10]. However, these schemes rely essentially on *symmetric variational formulations* of *Galerkin-type* and *preconditioning* on the infinite-dimensional level results from finding a Riesz basis for the energy space. In the present context suitable variational formulations turn out to be intrinsically *unsymmetric*. In fact, obtaining suitable *a posteriori error bounds*, will be based on *unsymmetric stable variational formulations* of *Petrov–Galerkin-type* for (1.2) and corresponding pure transport problems (1.4); see also [12]. A central tool is the Banach–Nečas–Babuška Theorem that is briefly recalled in Section 2.1.

1.4. **Layout.** In the remainder of this section we describe the organization and layout of the paper following Steps (I)–(III).

ad (I) Since, depending on the optimal parameters, solutions to (1.3) may exhibit discontinuities we opt to choose $U := L_2(\mathrm{D}\times\mathrm{S}) = L_2(\mathrm{D})\otimes L_2(\mathrm{S})$ as *trial space*. For a variational formulation to be stable the (infinite-dimensional) test space V must then be different from U. As shown in Sections 2.2–2.4, for the regime of problems considered below a proper test space warranting stability is determined by the *graph norm* of the pure transport operator $\mathcal{T}$. Moreover, as a preparation for Step (II), we derive in Section 2.5 bounds for $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)}$ in terms of the optical parameters.

ad (II) With (I) at hand we identify in Section 3 (infinite-dimensional) *preconditioners* $\mathcal{P} \in \mathcal{L}(V',U)$ that warrant convergence of (1.5) in U and render Step (III) practically viable. In particular, we identify two problem regimes of *dominating transport* and *dominating scattering*, depending on whether $\mathcal{T}^{-1}\mathcal{K}$ is a contraction in $\mathcal{L}(U,U)$ or not; see Sections 3.1 and 3.2.

ad (III) The remainder of the paper is devoted to Step (III). In Section 4 we identify core routines needed for the approximate realization of (1.5) as well as error tolerances these routines need to meet in order to guarantee convergence of the perturbed outer iteration to the exact solution. Again we have to distinguish first the two regimes of dominating transport or scattering in Sections 4.1 and 4.2, respectively, in order to formulate then the main algorithm in Section 4.3 that covers both regimes.

We stress that one never has to invert a dense system involving a discretization of the *global operator* $\mathcal{K}$. Instead an error-controlled application of $\mathcal{K}$ is needed. While most numerical studies treat either local problems or simple kernels like constants we make a point on including non-trivial scatterers. In Section 5 we present a scheme based on *Alpert wavelet representation* of $\mathcal{K}$ and *low-rank* approximations; see Section 5.2.

As shown in Section 3, the application of the preconditioner $\mathcal{P}$ in (1.5) is ultimately reduced to the error-controlled approximate inversion of the "lifted" pure transport operator $\mathcal{T}$ (acting on functions on $\mathrm{D} \times \mathrm{S}$; see (2.11)), discussed in Section 6. This makes essential use of recent results from [7,13] where rigorous sharp a posteriori error bounds for linear transport equations are derived for Discontinuous Petrov–Galerkin (DPG) schemes.

Remark 1.1. When progressing with the (perturbed) outer iteration, target accuracies decrease step by step so that one starts initially with very coarse DPG discretizations. The only linear systems to be solved in the course of such a *nested iteration* are the symmetric positive definite sparse DPG systems for the spatial problems which are always kept as small as possible depending on the current target tolerances. The size of systems that need to be inverted is always significantly smaller than the number of overall generated degrees of freedom.

Finally, we present in Section 7 some first numerical experiments as a proof of concept. They demonstrate, in particular, the crucial role of adaptivity in the transport solver. In fact, the number of degrees of freedom shown in Figure 7 already for two spatial dimensions indicate that realizing the required error tolerances with uniform spatial grids would be infeasible.

When the specific value of a constant does not matter we frequently employ the notation $a \lesssim b$ to express that a is bounded by a fixed constant multiple of b independent of all parameters a and b may depend on, that are not explicitly mentioned.

2. Step (I)—Variational formulations and well-posedness

2.1. **Stability.** Our approach relies on appropriate *variational formulations* of (1.3) which allow us to interpret (1.3) as an operator equation

$$\mathcal{B}u = f, \tag{2.1}$$

where $\mathcal{B}$ is induced by this variational formulation as a linear mapping from an infinite-dimensional trial space U to the dual V' of some (infinite-dimensional) test space V (see Section 1.3 (I)). Here the spaces U, V host functions of both the spatial variables x and the parametric variables $\vec{s}$.

Denoting by $\mathcal{L}(X,Y)$ the space of all bounded linear operators from X to Y, the objective is then to establish well-posedness of (2.1) which means bounded invertibility of $\mathcal{B}$ or, more precisely, boundedness of the *condition number*

$$\kappa_{U,V'}(\mathcal{B}) := \|\mathcal{B}\|_{\mathcal{L}(U,V')}\|\mathcal{B}^{-1}\|_{\mathcal{L}(V',U)}.$$

Specifying the precise *mapping properties* is therefore the central objective of this section. The choice of the (Hilbert)spaces U, V tells us under which assumptions on the data a unique weak solution exists and in which norm the accuracy of approximate solutions is measured.

A well-known tool to be used in this context is the following result by Banach–Nečas–Babuška which we recall for the convenience of the reader.

Theorem 2.1. *Assume that $q(\cdot,\cdot)\colon X \times Y \to \mathbb{R}$ is a bilinear form on the Hilbert spaces X, Y (with norms $\|\cdot\|_X$, $\|\cdot\|_Y$). The validity of the following properties:*

(1) *$q(\cdot,\cdot)$ is continuous, i.e., there exists a $\bar{C} < \infty$ such that*

$$|q(w,z)| \leqslant \bar{C}\|w\|_X\|z\|_Y, \quad w \in X,\, z \in Y;$$

(2) *there exists a* $\underline{c} > 0$ *such that*

$$\inf_{w\in X}\sup_{z\in Y}\frac{q(w,z)}{\|w\|_X\|z\|_Y} \geqslant \underline{c}; \tag{2.2}$$

(3) *for each* $z \in Y\backslash\{0\}$ *there exists a* $w \in X$ *such that* $q(w,z) \neq 0$*;*

is equivalent to the solvability of the problem: given $f \in Y'$ *find* $u \in X$ *such that*

$$q(u,v) = \langle v, f\rangle, \quad v \in Y.$$

Moreover, one has the stability relation

$$\|u\|_X \leqslant \underline{c}^{-1}\|f\|_{Y'}.$$

Note that condition (3) can be replaced by a second inf-sup condition (2.2) with the roles of X and Y interchanged.

Denoting by $\mathcal{Q}$ the operator from X to Y induced by $q(\cdot,\cdot)$, the above theorem says in particular that

$$\kappa_{X,Y'}(\mathcal{Q}) \leqslant \frac{\bar{C}}{\underline{c}}.$$

2.2. **Variational formulations of the pure transport problem (2.3).** As indicated under **ad (I)** in Section 1.3, a crucial role is played by a suitable weak formulation for the pure transport equation

$$\vec{s}\cdot\nabla u(x,\vec{s}) + \sigma(x,\vec{s})u(x,\vec{s}) = f(x,\vec{s}) \quad \text{for almost all } (x,\vec{s}) \in \mathrm{D}\times\mathrm{S}, \tag{2.3}$$

defined on the phase space $\mathrm{D}\times\mathrm{S}$, where, in the following:

$$\sigma \geqslant 0, \quad \|\sigma\|_{L_\infty(\mathrm{D}\times\mathrm{S})} < \infty. \tag{2.4}$$

We consider first corresponding *fiber* problems obtained by freezing the transport direction $\vec{s} \in \mathrm{S}$. In favor of possibly low regularity requirements on the solution, we follow [12]. Formally applying integration by parts yields the variational problem

$$a(u,v;\vec{s}) := \int_{\mathrm{D}} u(\sigma(\cdot,\vec{s})v - \vec{s}\cdot\nabla v)\,\mathrm{d}x = -\int_{\partial\mathrm{D}} \mathbf{n}\cdot\vec{s}uv\,\mathrm{d}x + \int_{\mathrm{D}} fv\,\mathrm{d}x \tag{2.5}$$

for test functions v from a suitable space yet to be determined. In fact, the left hand side is now well-defined for $u \in L_2(\mathrm{D})$ and $v \in H(\vec{s};\mathrm{D})$, where

$$H(\vec{s};\mathrm{D}) := \{v \in L_2(\mathrm{D}) \mid \vec{s}\cdot\nabla v \in L_2(\mathrm{D})\} \tag{2.6}$$

is a Hilbert space endowed with the norm

$$\|v\|^2_{H(\vec{s};\mathrm{D})} := \|v\|^2_{L_2(\mathrm{D})} + \|\vec{s}\cdot\nabla v\|^2_{L_2(\mathrm{D})}.$$

However, for $u \in L_2(\mathrm{D})$ the trace on $\partial\mathrm{D}$ is not well-defined. Introducing the closed subspaces

$$H_{0,\Gamma_\pm(\vec{s})}(\vec{s};\mathrm{D}) := \mathrm{clos}_{\|\cdot\|_{H(\vec{s};\mathrm{D})}}\{v \in C^1(\bar{\mathrm{D}}) \mid v\big|_{\Gamma_\pm(\vec{s})} = 0\}, \tag{2.7}$$

and restricting the test functions to $H_{0,\Gamma_+(\vec{s})}(\vec{s};\mathrm{D})$, the boundary integral on the right hand side of (2.5) extends only over $\Gamma_-(\vec{s})$. Thus, prescribing inflow boundary data $g \in L_2(\Gamma_-(\vec{s}), \mathbf{n}\cdot\vec{s})$, the weighted L_2 space on $\Gamma_-(\vec{s})$ with weight $|\mathbf{n}\cdot\vec{s}|$, a weak formulation of (2.3) is to seek for

$$U(\vec{s}) = U := L_2(\mathrm{D}), \qquad V(\vec{s}) := H_{0,\Gamma_+(\vec{s})}(\vec{s};\mathrm{D}), \tag{2.8}$$

$u = u(\vec{s}) \in U(\vec{s})$ such that

$$a(u, v; \vec{s}) := \int_{\mathrm{D}} u(\sigma(\cdot, \vec{s})v - \vec{s} \cdot \nabla v)\, \mathrm{d}x$$

$$= \int_{\Gamma_-(\vec{s})} \mathbf{n} \cdot \vec{s} g v \, \mathrm{d}x + \langle v, f \rangle =: \langle v, F \rangle, \quad v \in V(\vec{s}). \tag{2.9}$$

Here $\langle v, f \rangle = \langle v, f \rangle_{V,V'}$ stands for the dual pairing between V and V'. In particular, Dirichlet boundary conditions become natural boundary conditions which is an advantage when the domain of the inflow boundary portion varies with $\vec{s}$ because they need not be incorporated in U. In this setting, at least formally, the trial space U is independent of $\vec{s}$ while the test space $V = V(\vec{s})$ depends essentially on $\vec{s}$.

The operator $\mathcal{T}_{\vec{s}}$ induced by $a(u, v; \vec{s})$ through

$$(\mathcal{T}_{\vec{s}} w)(v) = a(w, v; \vec{s}), \quad w \in L_2(\mathrm{D}),\ v \in H_{0,\Gamma_+(\vec{s})}(\vec{s}; \mathrm{D}),$$

defines a bounded linear operator from $L_2(\mathrm{D})$ to $(H_{0,\Gamma_+(\vec{s})}(\vec{s}; \mathrm{D}))'$. Accordingly, we have for its (exact) adjoint

$$\mathcal{T}_{\vec{s}}^* \in \mathcal{L}(H_{0,\Gamma_+(\vec{s})}(\vec{s}; \mathrm{D}), L_2(\mathrm{D})), \quad \langle w, \mathcal{T}_{\vec{s}}^* v \rangle = a(w, v; \vec{s}),$$
$$w \in L_2(\mathrm{D}),\ v \in H_{0,\Gamma_+(\vec{s})}(\vec{s}; \mathrm{D}).$$

Before addressing the invertibility of the operator $\mathcal{T}_{\vec{s}}$ we consider the "lifted" versions viewed as functions of x and $\vec{s}$; see [12]. The role of $H(\vec{s}; \mathrm{D})$ (see (2.6)) is now played by the space

$$H(\mathrm{D} \times \mathrm{S}) := \{v \in L_2(\mathrm{D} \times \mathrm{S}) \mid \vec{s} \cdot \nabla v \in L_2(\mathrm{D} \times \mathrm{S})\}.$$

The space $H(\mathrm{D} \times \mathrm{S})$ becomes a Hilbert space under the norm

$$\|v\|^2_{H(\mathrm{D}\times\mathrm{S})} := \int_{\mathrm{D}\times\mathrm{S}} \left(|\vec{s} \cdot \nabla v(x, \vec{s})|^2 + |v(x, \vec{s})|^2\right) \mathrm{d}x\, \mathrm{d}\vec{s}. \tag{2.10}$$

Likewise, the counterparts to the spaces (2.7) are given by the closed subspaces

$$H_{0,\pm}(\mathrm{D} \times \mathrm{S}) := \mathrm{clos}_{\|\cdot\|_{H(\mathrm{D}\times\mathrm{S})}} \{v \in C^1(\overline{\mathrm{D} \times \mathrm{S}}) \mid v|_{\Gamma_\pm} = 0\}.$$

The "lifted" bilinear form

$$a(w, v) := \int_{\mathrm{S}} a(w, v; \vec{s})\, \mathrm{d}\vec{s}$$

allows us to define, in analogy to the above fiber versions, $\mathcal{T}$ by

$$\langle \mathcal{T} w, v \rangle = a(w, v), \quad w \in U,\ v \in V, \tag{2.11}$$

where

$$U := L_2(\mathrm{D} \times \mathrm{S}), \qquad V := H_{0,+}(\mathrm{D} \times \mathrm{S}). \tag{2.12}$$

Thus, the variational problem: find $u \in U$ such that for any $f \in V'$

$$a(u, v) = \langle v, f \rangle, \quad v \in V, \tag{2.13}$$

is equivalent to the operator equation

$$\mathcal{T} u = f,$$

where $\mathcal{T}$ is viewed as a mapping from U into V'.

The invertibility of the fiber operators $\mathcal{T}_{\vec{s}}$ and the lifted version $\mathcal{T}$ will be seen to be an immediate consequence of the following norm-equivalences; see (2.10).

Theorem 2.2. *Under the assumption* (2.4) *one has*

$$\|\mathcal{T}_{\vec{s}}^* v\|_{L_2(\mathrm{D})} \sim \|v\|_{H(\vec{s};\mathrm{D})}, \qquad v \in V(\vec{s}) = H_{0,\Gamma_+(\vec{s})}(\vec{s};\mathrm{D}),\ \vec{s} \in \mathrm{S},$$

$$\|\mathcal{T}^* v\|_{L_2(\mathrm{D}\times\mathrm{S})} \sim \|v\|_{H(\mathrm{D}\times\mathrm{S})}, \qquad v \in V = H_{0,+}(\mathrm{D}\times\mathrm{S}),$$

as well as

$$\begin{aligned} \|\mathcal{T}_{\vec{s}} v\|_{L_2(\mathrm{D})} &\sim \|v\|_{H(\vec{s};\mathrm{D})}, \qquad v \in V(\vec{s}) = H_{0,\Gamma_-(\vec{s})}(\vec{s};\mathrm{D}),\ \vec{s} \in \mathrm{S}, \\ \|\mathcal{T} v\|_{L_2(\mathrm{D}\times\mathrm{S})} &\sim \|v\|_{H(\mathrm{D}\times\mathrm{S})}, \qquad v \in V = H_{0,-}(\mathrm{D}\times\mathrm{S}), \end{aligned} \tag{2.14}$$

where the constants in the first line are independent of $\vec{s} \in \mathrm{S}$ *and depend only on* $\sigma_{\min}$, $\sigma_{\max}$, *and* $\hat{\ell} = \operatorname{diam}(\mathrm{D})$.

In principle, these results have been already shown in [12]. We return to a proof in the next section in order to exhibit the dependence of involved constants from the optical parameters which will be needed for the numerical scheme.

As a consequence of Theorem 2.2 we obtain the following results.

Corollary 2.3. *Assume that* (2.4) *holds. Then there exist constants* $0 < \underline{c}, \bar{C} < \infty$ *such that for* $U(\vec{s})$, U, $V(\vec{s})$, V *defined by* (2.8), (2.12), *respectively,*

$$\|\mathcal{T}_{\vec{s}}\|_{\mathcal{L}(U(\vec{s}),V(\vec{s}))}, \|\mathcal{T}\|_{\mathcal{L}(U,V)} \leqslant \bar{C}, \qquad \|\mathcal{T}_{\vec{s}}^{-1}\|_{\mathcal{L}(V(\vec{s})',U(\vec{s}))}, \|\mathcal{T}^{-1}\|_{\mathcal{L}(V',U)} \leqslant \underline{c}^{-1}.$$

Hence, the variational problems (2.9), (2.13), *respectively, have unique solutions that depend continuously on the data.*

Proof. First note that Theorem 2.2 implies that

$$|\!|\!| v |\!|\!|_{\mathcal{T}_{\vec{s}}^*} := \|\mathcal{T}_{\vec{s}}^* v\|_{L_2(\mathrm{D})}, \qquad |\!|\!| v |\!|\!|_{\mathcal{T}^*} := \|\mathcal{T}^* v\|_{L_2(\mathrm{D}\times\mathrm{S})}, \tag{2.15}$$

are equivalent norms on $V(\vec{s}) = H_{0,\Gamma_+(\vec{s})}(\vec{s};\mathrm{D})$, $V = H_{0,+}(\mathrm{D}\times\mathrm{S})$, respectively. Endowing $V(\vec{s})$, V with these norms, observe that

$$\sup_{w\in U} \frac{a(w,v)}{\|w\|_U} = \sup_{w\in U} \frac{\langle w, \mathcal{T}^* v\rangle}{\|w\|_U} = \|\mathcal{T}^* v\|_{U'} = \|\mathcal{T}^* v\|_U = |\!|\!| v |\!|\!|_{\mathcal{T}^*}.$$

Since by (2.14), $\mathcal{T}$ is injective, and hence $\mathcal{T}^*$ is surjective, we obtain

$$\sup_{v\in V} \frac{a(w,v)}{|\!|\!| v |\!|\!|_{\mathcal{T}^*}} = \sup_{v\in V} \frac{\langle \mathcal{T} w, v\rangle}{|\!|\!| v |\!|\!|_{\mathcal{T}^*}} = \sup_{v\in V} \frac{\langle w, \mathcal{T}^* v\rangle}{|\!|\!| v |\!|\!|_{\mathcal{T}^*}} \geqslant \frac{\|w\|^2_{L_2(\mathrm{D}\times\mathrm{S})}}{\|w\|_{L_2(\mathrm{D}\times\mathrm{S})}} = \|w\|_{L_2(\mathrm{D}\times\mathrm{S})} = \|w\|_U,$$

which says that $\underline{c} = \bar{C} = 1$ and hence, by Theorem 2.1,

$$\kappa_{U,V'}(\mathcal{T}) = 1$$

for U, V as in (2.12). The treatment of the fiber operators $\mathcal{T}_{\vec{s}}$ is completely analogous. Hence, with the choice (2.15) of norms (2.9) and (2.13) are perfectly conditioned, i.e., the operators $\mathcal{T}_{\vec{s}}$, $\mathcal{T}$ are even *isometries* between the respective pairs of spaces. This completes the proof. □

Remark 2.4. Later, both the fact that the fiber operators $\mathcal{T}_{\vec{s}}$ as well as the lifted versions $\mathcal{T}$ have bounded condition numbers will be used in the envisaged numerical scheme.

It will be useful to clearly distinguish the two above variational formulations.

Variational formulation (F1): *determined by the combination of the bilinear form $a(\cdot,\cdot)$ from (2.9) with the pair of spaces U, V it is supposed to act on, namely*

$$\text{(F1)}\qquad a(u,v) = \int_{\mathrm{D}\times\mathrm{S}} u(x,\vec{s})(\sigma(x,\vec{s})v(x,\vec{s}) - \vec{s}\cdot\nabla v(x,\vec{s}))\,\mathrm{d}x\,\mathrm{d}\vec{s},$$
$$U = L_2(\mathrm{D}\times\mathrm{S}), \quad V = H_{0,+}(\mathrm{D}\times\mathrm{S}).$$

Variational formulation (F2): *determined by*

$$\text{(F2)}\qquad a(u,v;\vec{s}) := \int_{\mathrm{D}} (\vec{s}\cdot\nabla u + \sigma(\cdot,\vec{s})u)v\,\mathrm{d}x, \quad a(u,v) := \int_{\mathrm{S}} a(u(\cdot,\vec{s}),v(\cdot,\vec{s});\vec{s})\,\mathrm{d}\vec{s},$$
$$U(\vec{s}) = H_{0,\Gamma_-(\vec{s})}(\vec{s};\mathrm{D}), \qquad V(\vec{s}) = V = L_2(\mathrm{D}),$$
$$U = H_{0,-}(\mathrm{D}\times\mathrm{S}), \qquad V = L_2(\mathrm{D}\times\mathrm{S}).$$

Endowing $U = H_{0,-}(\mathrm{D}\times\mathrm{S})$ with the norm $|||w|||_{\mathcal{T}} := \|\mathcal{T}w\|_{L_2(\mathrm{D}\times\mathrm{S})}$, the same type of argument as in the proof of Theorem 2.3 again combined with Theorem 2.2 yields the following result; see also [12].

Proposition 2.5. *For data $f \in L_2(\mathrm{D}), L_2(\mathrm{D}\times\mathrm{S})$, respectively, the variational problems*

$$a(u(\vec{s}),v;\vec{s}) = \langle v, f\rangle, \quad v\in V(\vec{s}),\ \vec{s}\in\mathrm{S}, \qquad a(u,v) = \langle v,f\rangle, \quad v\in V, \tag{2.16}$$

have unique solutions in $U(\vec{s})$, U, defined by (F2), respectively, which depend continuously on the data.

Remark 2.6. The solutions in (2.16) are required to have more regularity than in the first version (F1), requiring, in particular, that $f\in L_2(\mathrm{D}\times\mathrm{S})$. Moreover, boundary conditions on $\Gamma_-(\vec{s})$, Γ_- are now essential boundary conditions that need to be built into the ansatz. Our interest in the formulation (F2) is a duality argument to be used later for the variational formulation of the full equation (1.3).

2.3. **Norm equivalences.** We establish next the norm equivalences in Theorem 2.2. As indicated earlier, a main reason for revisiting the proof is to prepare for Section 2.5 by determining the dependence of constants on the optical parameters. We use similar arguments as in [16] (see also [12] for related discussions).

Let the time of escape of free moving particles from D be

$$\ell_\pm(x,\vec{s}) := \inf\{t>0 \mid x\pm t\vec{s}\notin\mathrm{D}\}.$$

Then,

$$\ell(x,\vec{s}) := \ell_-(x,\vec{s}) + \ell_+(x,\vec{s})$$

is the length of the longest line segment through x in direction $\vec{s}$ completely contained in D and

$$\hat{\ell} := \sup_{(x,\vec{s})\in\mathrm{D}\times\mathrm{S}} \ell(x,\vec{s}) = \mathrm{diam}(\mathrm{D})$$

is the maximum time of escape. For a given $\vec{s}\in\mathrm{S}$, we can express any $x\in\mathrm{D}$ in terms of characteristic coordinates as follows. Denoting $x_-(x,\vec{s})\in\Gamma_-(\vec{s})$ the intersection of the line $x+t\vec{s}$, $t\in\mathbb{R}$, with $\Gamma_-(\vec{s})$, we can write

$$x = x_-(x,\vec{s}) + \ell_-(x,\vec{s})\vec{s}.$$

In these terms, define for $v \in L_2(\mathrm{D}\times\mathrm{S})$ and almost every $x = x_-(x,\vec{s}) + \ell_-(x,\vec{s})\vec{s} \in \mathrm{D}$, $x_-(x,\vec{s}) \in \Gamma_-(\vec{s})$

$$\begin{aligned} w(x,\vec{s}) &= w\left(x_-(x,\vec{s}) + \ell_-(x,\vec{s})\vec{s}, \vec{s}\right) \\ &:= \int_0^{\ell_-(x,\vec{s})} e^{-\int_r^{\ell_-(x,\vec{s})} \sigma(x_-(x,\vec{s})+\vec{s}\theta,\vec{s})\,\mathrm{d}\theta} v\left(x_-(x,\vec{s}) + r\vec{s}, \vec{s}\right)\,\mathrm{d}r. \end{aligned} \tag{2.17}$$

One readily verifies that w as well as $\mathcal{T}_{\vec{s}} w(\cdot,\vec{s}) = v(\cdot,\vec{s})$ belong to $L_2(\mathrm{D}\times\mathrm{S})$. Moreover,

$$\|\mathcal{T} w\|^2_{L_2(\mathrm{D}\times\mathrm{S})} = \int_{\mathrm{S}} \|\mathcal{T}_{\vec{s}} w(\cdot,\vec{s})\|^2_{L_2(\mathrm{D})}\,\mathrm{d}\vec{s} \leqslant C_1 \|w\|^2_{H(\mathrm{D}\times\mathrm{S})},$$

where C_1 depends on $\sigma_{\max}$, where we abbreviate

$$\sigma_{\min} := \inf_{(x,\vec{s})\in\mathrm{D}\times\mathrm{S}} \sigma(x,\vec{s}), \qquad \sigma_{\max} := \sup_{(x,\vec{s})\in\mathrm{D}\times\mathrm{S}} \sigma(x,\vec{s}).$$

We first derive a bound on $\mathcal{T}^{-1}$ as an operator mapping $L_2(\mathrm{D}\times\mathrm{S})$ into itself.

Lemma 2.7. *If $0 \leqslant \sigma \in L^\infty(\mathrm{D}\times\mathrm{S})$, then $\mathcal{T}^{-1}$ is a continuous operator from $L^2(\mathrm{D}\times\mathrm{S})$ to $L^2(\mathrm{D}\times\mathrm{S})$ and*

$$\|\mathcal{T}^{-1}\|_{\mathcal{L}(L^2(\mathrm{D}\times\mathrm{S}),L^2(\mathrm{D}\times\mathrm{S}))} \leqslant \sqrt{\hat{\ell}\frac{1-e^{-2\hat{\ell}\sigma_{\min}}}{2\sigma_{\min}}} \leqslant \min\left\{\hat{\ell}, \sqrt{\hat{\ell}/2\sigma_{\min}}\right\}. \tag{2.18}$$

Defining the formal adjoint of $\mathcal{T}$, by $\int_{\mathrm{D}\times\mathrm{S}}(\mathcal{T}^* v)w\,\mathrm{d}x\,\mathrm{d}\vec{s} = \int_{\mathrm{D}\times\mathrm{S}}(-\vec{s}\cdot\nabla v + \sigma v)w\,\mathrm{d}x\,\mathrm{d}\vec{s}$, the same bound holds for $\|\mathcal{T}^{-*}\|_{\mathcal{L}(L^2(\mathrm{D}\times\mathrm{S}),L^2(\mathrm{D}\times\mathrm{S}))}$.

Proof. For $v \in L_2(\mathrm{D}\times\mathrm{S})$, we consider w as defined in (2.17). One readily checks that w satisfies (2.13) for $v = f$. For $(x,\vec{s}) = (x_- + \ell_-(x,\vec{s})\vec{s}, \vec{s})$ and $0 \leqslant \ell_-(x,\vec{s}) \leqslant \ell(x_-,\vec{s})$, it follows from (2.17) and the Cauchy-Schwarz inequality

$$|w(x,\vec{s})|^2 \leqslant \left(\int_0^{\ell(x_-,\vec{s})} e^{-2\int_r^{\ell(x_-,\vec{s})}\sigma(x_-+\theta\vec{s},\vec{s})\,\mathrm{d}\theta}\,\mathrm{d}r\right)\left(\int_0^{\ell(x_-,\vec{s})} v(x_- + r\vec{s},\vec{s})^2\,\mathrm{d}r\right).$$

Since

$$\begin{aligned} \int_0^{\ell(x_-,\vec{s})} e^{-2\int_r^{\ell(x_-,\vec{s})}\sigma(x_-+\theta\vec{s},\vec{s})\,\mathrm{d}\theta}\,\mathrm{d}r &\leqslant \int_0^{\ell(x_-,\vec{s})} e^{-2\int_r^{\ell(x_-,\vec{s})}\sigma_{\min}\,\mathrm{d}\theta}\,\mathrm{d}r = \frac{1-e^{-2\ell(x_-,\vec{s})\sigma_{\min}}}{2\sigma_{\min}} \\ &\leqslant \frac{1-e^{-2\hat{\ell}\sigma_{\min}}}{2\sigma_{\min}}, \end{aligned}$$

we derive

$$|w(x,\vec{s})|^2 \leqslant \frac{1-e^{-2\hat{\ell}\sigma_{\min}}}{2\sigma_{\min}} \int_0^{\ell(x_-,\vec{s})} |v(x_- + r\vec{s},\vec{s})|^2\,\mathrm{d}r. \tag{2.19}$$

Integrating (2.19) over $\mathrm{D}\times\mathrm{S}$,

$$\begin{aligned} \|w\|^2_{L_2(\mathrm{D}\times\mathrm{S})} &= \int_{(x_-,\vec{s})\in\Gamma_-}\int_{t=0}^{\ell(x_-,\vec{s})} |w(x_- + t\vec{s},\vec{s})|^2\,|\vec{s}\cdot\mathbf{n}|\,\mathrm{d}t\,\mathrm{d}\Gamma_- \\ &\leqslant \frac{1-e^{-2\hat{\ell}\sigma_{\min}}}{2\sigma_{\min}} \int_{(x_-,\vec{s})\in\Gamma_-}\int_{t=0}^{\ell(x_-,\vec{s})}\int_{r=0}^{\ell(x_-,\vec{s})} |v(x_- + r\vec{s},\vec{s})|^2\,\mathrm{d}r|\vec{s}\cdot\mathbf{n}|\,\mathrm{d}t\,\mathrm{d}\Gamma_- \\ &\leqslant \hat{\ell}\frac{1-e^{-2\hat{\ell}\sigma_{\min}}}{2\sigma_{\min}} \|v\|^2_{L_2(\mathrm{D}\times\mathrm{S})}, \end{aligned}$$

where we have used that $\int_{t=0}^{\ell(x_-,\vec{s})} \mathrm{d}t \leqslant \hat{\ell}$ for all $(x_-,\vec{s}) \in \Gamma_-$ to derive the last bound. This yields the first bound for $\|\mathcal{T}^{-1}\|_{\mathcal{L}(L^2(\mathrm{D}\times\mathrm{S}),L^2(\mathrm{D}\times\mathrm{S}))}$ given in (2.18). The second bound follows directly from the fact that $\hat{\ell}(1-e^{-2\hat{\ell}\sigma_{\min}})/(2\sigma_{\min}) \leqslant \min\left\{\hat{\ell}^2, \hat{\ell}/(2\sigma_{\min})\right\}$ since $1-e^{-x} \leqslant \min\{1,x\}$ for any $x \geqslant 0$. The argument for $\mathcal{T}^*$ is the same. □

Proof of Theorem 2.2. The inequality (2.18) says that

$$\|v\|_{L_2(\mathrm{D}\times\mathrm{S})} \leqslant \min\left\{\hat{\ell}, \sqrt{\hat{\ell}/2\sigma_{\min}}\right\} \begin{cases} \|\mathcal{T}v\|_{L_2(\mathrm{D}\times\mathrm{S})}, & v \in H_{0,-}(\mathrm{D}\times\mathrm{S}), \\ \|\mathcal{T}^*v\|_{L_2(\mathrm{D}\times\mathrm{S})}, & v \in H_{0,+}(\mathrm{D}\times\mathrm{S}). \end{cases} \tag{2.20}$$

Integrating (2.19) only over $x \in \mathrm{D}$ leads to analogous statements for the fibers $\mathcal{T}_{\vec{s}}$, $\mathcal{T}_{\vec{s}}^*$, namely

$$\|v\|_{L_2(\mathrm{D})} \leqslant \min\left\{\hat{\ell}, \sqrt{\hat{\ell}/2\sigma_{\min}}\right\} \begin{cases} \|\mathcal{T}_{\vec{s}}v\|_{L_2(\mathrm{D})}, & v \in H_{0,\Gamma_-(\vec{s})}(\vec{s};\mathrm{D}), \\ \|\mathcal{T}_{\vec{s}}^*v\|_{L_2(\mathrm{D})}, & v \in H_{0,\Gamma_+(\vec{s})}(\vec{s};\mathrm{D}), \end{cases} \quad \vec{s} \in \mathrm{S}.$$

We infer from (2.20) that, for instance,

$$\|\mathcal{T}_{\vec{s}}v\|_{L_2(\mathrm{D})} \leqslant \|\vec{s}\cdot\nabla v\|_{L_2(\mathrm{D})} + \sigma_{\max}\|v\|_{L_2(\mathrm{D})} \leqslant (1+\sigma_{\max}^2)^{1/2}\|v\|_{H(\vec{s};\mathrm{D})}.$$

Conversely, one has

$$\begin{aligned} \|v\|_{H(\vec{s};\mathrm{D})} &\leqslant \|\vec{s}\cdot\nabla v\|_{L_2(\mathrm{D})} + \|v\|_{L_2(\mathrm{D})} \leqslant \|\mathcal{T}_{\vec{s}}v\|_{L_2(\mathrm{D})} + (1+\sigma_{\max})\|v\|_{L_2(\mathrm{D})} \\ &\leqslant \left(1 + (1+\sigma_{\max})\min\left\{\hat{\ell}, \sqrt{\hat{\ell}/2\sigma_{\min}}\right\}\right)\|\mathcal{T}_{\vec{s}}v\|_{L_2(\mathrm{D})}. \end{aligned}$$

The remaining assertions of Theorem 2.2 are derived analogously. □

Remark 2.8. $\|\mathcal{T}^{-1}\|_{\mathcal{L}(L^2(\mathrm{D}\times\mathrm{S}),L^2(\mathrm{D}\times\mathrm{S}))}$ is small when either $\mathrm{diam}(\mathrm{D})$ is small or when $\sigma_{\min}$ is large relative to $\hat{\ell}$.

2.4. **Variational formulation of the radiative transfer problem (1.3).**

Throughout this section we let $g \equiv 0$, i.e., we treat homogeneous inflow boundary conditions. Also, we assume that the kernel K satisfies

$$K(x,\vec{s},\vec{s}') \geqslant 0, \quad (x,\vec{s},\vec{s}') \in \mathrm{D}\times\mathrm{S}\times\mathrm{S}, \quad K \in L_\infty(\mathrm{D};L_2(\mathrm{S}\times\mathrm{S})) \subset L_2(\mathrm{D}\times\mathrm{S}\times\mathrm{S}), \tag{2.21}$$

so that we have

$$\mathcal{K}, \mathcal{K}^* \in \mathcal{L}(L_2(\mathrm{D}\times\mathrm{S}), L_2(\mathrm{D}\times\mathrm{S})). \tag{2.22}$$

Following the same lines as before for the pure transport operator $\mathcal{T}$ we can define the operator $\mathcal{B}$ by

$$b(w,v) = \langle \mathcal{B}w, v\rangle := \int_{\mathrm{S}} a(w(\cdot,\vec{s}), v(\cdot,\vec{s});\vec{s})\,\mathrm{d}\vec{s} - k(w,v) \quad \forall\, w \in U,\ v \in V, \tag{2.23}$$

where $k(w,v) = \langle \mathcal{K}w, v\rangle$, and the spaces U, V are chosen according to the formulations (F1), (F2), respectively.

A key property in what follows is *accretivity* of $\mathcal{B}$. In the present context this means that there exists some positive α such that

$$(\mathcal{B}v, v) \geqslant \alpha\|v\|^2_{L_2(\mathrm{D}\times\mathrm{S})}, \quad v \in H_{0,-}(\mathrm{D}\times\mathrm{S}). \tag{2.24}$$

We postpone for a moment listing conditions on the optical parameters which imply (2.24) but present first the central result in this section.

Theorem 2.9. *Assume that (2.21) and (2.24) hold. Then, for either one of the two formulations (F1), (F2) and any $f \in V'$ the problem: find $u \in U$ such that*

$$b(u,v) = \langle f, v\rangle, \quad v \in V, \tag{2.25}$$

has a unique solution satisfying

$$\|u\|_U \lesssim \|f\|_{V'},$$

with constants depending only on the optical parameters.

The operator $\mathcal{B}$, defined by (2.23) is in either setting a linear norm-isomorphism from U onto V', i.e., has a finite condition $\kappa_{U,V'}(\mathcal{B}) < \infty$.

The proof makes use of the following norm-equivalences.

Lemma 2.10. *Let $\mathcal{T}'$, $\mathcal{B}'$ denote the formal adjoints of $\mathcal{T}$, $\mathcal{B}$, respectively. Then, under the assumptions (2.28), (2.21) on σ and K one has*

$$\begin{aligned} \|w\|_{H(\mathrm{D}\times\mathrm{S})} &\sim \|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})} \sim \|\mathcal{T}w\|_{L_2(\mathrm{D}\times\mathrm{S})}, \quad w \in H_{0,-}(\mathrm{D}\times\mathrm{S}),\\ \|w\|_{H(\mathrm{D}\times\mathrm{S})} &\sim \|\mathcal{B}'w\|_{L_2(\mathrm{D}\times\mathrm{S})} \sim \|\mathcal{T}'w\|_{L_2(\mathrm{D}\times\mathrm{S})}, \quad w \in H_{0,+}(\mathrm{D}\times\mathrm{S}), \end{aligned} \tag{2.26}$$

where the constants depend on the optical parameters.

Proof of Lemma 2.10. By (2.22), we have for some constant C_1

$$\|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})} \leqslant \|\mathcal{T}w\|_{L_2(\mathrm{D}\times\mathrm{S})} + C_1\|w\|_{L_2(\mathrm{D}\times\mathrm{S})} \leqslant (1 + C_1C_2)\|\mathcal{T}w\|_{L_2(\mathrm{D}\times\mathrm{S})},$$

where we have used (2.20) in the last step. Conversely, again by (2.22), (2.24), and using Young's inequality yields

$$\begin{aligned} \|\mathcal{T}w\|_{L_2(\mathrm{D}\times\mathrm{S})} &\leqslant \|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})} + \|\mathcal{K}w\|_{L_2(\mathrm{D}\times\mathrm{S})}\\ &\leqslant \|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})} + C_1\|w\|_{L_2(\mathrm{D}\times\mathrm{S})}\\ &\leqslant \|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})} + \frac{C_1}{\sqrt{\alpha}}(\mathcal{B}w, w)^{1/2}\\ &\leqslant \|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})} + \frac{C_1}{\sqrt{\alpha}}\Big(\frac{\|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})}}{2\delta} + \delta\|w\|_{L_2(\mathrm{D}\times\mathrm{S})}\Big)\\ &\leqslant \|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})} + \frac{C_1}{\sqrt{\alpha}}\Big(\frac{\|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})}}{2\delta} + \delta C_2\|\mathcal{T}w\|_{L_2(\mathrm{D}\times\mathrm{S})}\Big), \end{aligned}$$

where $C_2 = \min\Big\{\hat{\ell}, \sqrt{\hat{\ell}/\sigma_{\min}}\Big\}$ is the constant from (2.20). Choosing δ small enough to ensure that $C_1C_2\delta/\alpha < 1$, the relation $\|\mathcal{B}w\|_{L_2(\mathrm{D}\times\mathrm{S})} \sim \|\mathcal{T}w\|_{L_2(\mathrm{D}\times\mathrm{S})}$ follows. The first line in (2.26) follows then from Theorem 2.2 proving the assertion for $\mathcal{B}$. The argument for $\mathcal{B}'$ is analogous. □

We are now in position to prove Theorem 2.9.

Proof of Theorem 2.9. First, under the given assumptions we clearly have for either formulation (F1) or (F2) with respective pairs U, V, that $\mathcal{B}$ is bounded

$$\mathcal{B} \in \mathcal{L}(U, V').$$

Then, it follows from Theorem 2.1 and (2.24) that under the above assumptions

$$\|\mathcal{B}^{-1}\|_{\mathcal{L}(L_2(\mathrm{D}\times\mathrm{S}),L_2(\mathrm{D}\times\mathrm{S}))} \leqslant \alpha^{-1}. \tag{2.27}$$

To prove the last statement of the theorem note that in view of (2.26), injectivity of $\mathcal{T}$ and $\mathcal{T}'$ implies injectivity of $\mathcal{B}$ and $\mathcal{B}'$. Suppose $\mathcal{B}$ were not surjective. Then there exists a $w_0 \neq 0$ in $L_2(\Omega)$ such that $\langle \mathcal{B}w, w_0\rangle = 0$ for all $w \in H_{0,-}(\Omega)$. By boundedness of $\mathcal{B}$ and denseness of $H_{0,-}(\Omega)$ in $L_2(\Omega)$, this leads to a contradiction to (2.24). We can argue in the same way for $\mathcal{B}'$ to conclude that $\mathcal{B}$ and $\mathcal{B}'$ are bijections for their respective pairs of spaces. This holds by duality, since $(\mathcal{B}')^*$ agrees with $\mathcal{B}$ as a mapping from $L_2(\Omega)$ to $(H_{0,-}(\Omega))'$. In view of Lemma 2.10, the proof of Theorem 2.9 can now be completed with the aid of Theorem 2.1 in exactly the same way as the proof of Corollary 2.3. □

When the specific choice of the settings (F1) or (F2) is clear from the context, we view (2.25) as an operator equation

$$\mathcal{B}u = f$$

with data f in the respective dual space V'.

We discuss next two general conditions on the optical parameters that entail (2.24). Defining the kernel averages

$$\bar\sigma(x,\vec s) := \int_{\mathrm{S}} K(x,\vec s,\vec s')\,\mathrm{d}\vec s' \quad \text{and} \quad \bar\sigma'(x,\vec s) := \int_{\mathrm{S}} K(x,\vec s',\vec s)\,\mathrm{d}\vec s',$$

a first frequently studied general class of optical parameters is signified by the fact that there exist $0 < \alpha, M_a < \infty$ such that for all $(x,\vec s) \in \mathrm{D}\times\mathrm{S}$,

(2.28)
$$\sigma(x,\vec s) - \bar\sigma(x,\vec s) \geqslant \alpha, \quad \sigma(x,\vec s) - \bar\sigma'(x,\vec s) \geqslant \alpha, \quad \bar\sigma(x,\vec s) \leqslant M_a, \quad \bar\sigma'(x,\vec s) \leqslant M_a'.$$

Note that this implies that the absorption coefficient σ is not allowed to vanish in D. For this class we recall the following well-known result (see, e.g., [14, Chapter XXI, §2, Theorem 4]).

Proposition 2.11. *If σ and K satisfy assumptions (2.21) and (2.28), then the operator $\mathcal{B}$ is* accretive, *i.e., for any $v \in H_{0,-}(\mathrm{D}\times\mathrm{S})$,*

$$(\mathcal{B}v, v) \geqslant \alpha\|v\|^2_{L_2(\mathrm{D}\times\mathrm{S})},$$

where the constant α is the one appearing in (2.28).

For the convenience of the reader we sketch the simple argument. It follows from conditions (2.28), (2.21) that $(\sigma v - \mathcal{K}v, v) \geqslant \alpha\|v\|^2_{L_2(\mathrm{D}\times\mathrm{S})}$ on $L_2(\mathrm{D}\times\mathrm{S})$, which, combined with the accretivity of $\mathcal{A}$ on $H_{0,-}(\mathrm{D}\times\mathrm{S})$, defined by $\langle \mathcal{A}w, v\rangle = \int_{\mathrm{D}\times\mathrm{S}} \vec s\cdot\nabla w(x,\vec s)v(x,\vec s)\,\mathrm{d}x\,\mathrm{d}\vec s$, i.e., $(\mathcal{A}v, v) \geqslant 0$ for all $v \in H_{0,-}(\mathrm{D}\times\mathrm{S})$, yields the conclusion.

We emphasize that condition (2.28) is not necessary for (2.24) to hold as can be seen from the following class of frequently used kernels with slightly more specified structure. Consider

(2.29)
$$K(x,\vec s,\vec s') = \kappa(x)G(\vec s,\vec s'), \quad G(\vec s,\vec s') = G(\vec s',\vec s), \quad G(\vec s,\vec s') \geqslant 0, \quad \vec s,\vec s' \in \mathrm{S}, \quad \kappa \geqslant \kappa_0 > 0,$$

with the normalization

$$\int_{\mathrm{S}} G(\vec s,\vec s')\,\mathrm{d}\vec s' = \int_{\mathrm{S}} G(\vec s,\vec s')\,\mathrm{d}\vec s = 1, \quad \vec s,\ \vec s' \in \mathrm{S}.$$

Once the integral over one argument is a constant, this latter relation can always be realized by rescaling κ. Assuming always that $\mathrm{d}\vec{s}$ is the Haar measure, it also follows that $\int_{\mathrm{S}\times\mathrm{S}} G(\vec{s},\vec{s}')\,\mathrm{d}\vec{s}\,\mathrm{d}\vec{s}' = 1$. Moreover, we split

$$\sigma = \sigma_a + \kappa,$$

where $\sigma_a \geqslant 0$ is the so-called absorption coefficient. Hence in this case $\sigma(x,\vec{s}) - \bar{\sigma}(x,\vec{s}) = \sigma(x,\vec{s}) - \bar{\sigma}'(x,\vec{s}) = \sigma_a(x,\vec{s})$ so that (2.28) does not hold whenever σ_a vanishes somewhere in D. On the other hand, let $\mathcal{C}_+ \subset L_2(\mathrm{D}\times\mathrm{S})$ be the cone of non-negative functions in $L_2(\mathrm{D}\times\mathrm{S})$ (in the weak sense) and define

$$\mathcal{K}_0 v := \int_{\mathrm{S}} G(\cdot,\vec{s}')v(\vec{s}')\,\mathrm{d}\vec{s}'.$$

Under the above conditions the largest eigenvalue of $\mathcal{K}_0$ is one, it is simple and has the constant as the corresponding eigenfunction. Therefore,

$$\sup\left\{(v,\mathcal{K}_0 v) \;\middle|\; v\in\mathcal{C}_+\cap H_{0,-}(\mathrm{D}\times\mathrm{S}),\ \|v\|_{L_2(\mathrm{D}\times\mathrm{S})}=1\right\} =: \beta < 1.$$

Thus, the accretivity condition (2.24) holds with

$$\alpha \geqslant (\sigma_a)_{\min} + \kappa_0(1-\beta),$$

which is strictly larger than zero even if the absorption coefficient vanishes in D.

In principle, one could base a numerical method on both formulations (F1), (F2), where the latter one would seek approximations in a stronger norm. However, in what follows we focus on the setting (F1) where the solution is sought in $U = L_2(\mathrm{D}\times\mathrm{S})$ and where boundary conditions are natural ones.

Remark 2.12. There is of course an alternate way of establishing bounded invertibility of $\mathcal{B}\in\mathcal{L}(U,V')$ whenever the condition

$$\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \leqslant \rho < 1 \tag{2.30}$$

holds. While continuity of $\mathcal{B}$ is immediate, a straightforward Neuman-series argument shows that then

$$\|\mathcal{B}^{-1}\|_{\mathcal{L}(V',U)} \leqslant (1-\rho)^{-1}\|\mathcal{T}^{-1}\|_{\mathcal{L}(V',U)}.$$

We refer to the regime of problems where (2.30) is valid as the *weakly transport dominated case*.

In addition, condition (2.30) will be seen to be crucial for the identification of preconditioners $\mathcal{P}$ in the idealized iteration (1.5). We therefore address the derivation of bounds for $\|\mathcal{B}^{-1}\|_{\mathcal{L}(V',U)}$ in the next section.

2.5. **Contractivity of $\mathcal{T}^{-1}\mathcal{K}$.** We begin with the following result taken from [14, Chapter XXI, §2, Lemma 1].

Proposition 2.13. *Assume that* (2.21) *and* (2.28) *hold. Then $\mathcal{K}$ maps $L_2 := L_2(\mathrm{D}\times\mathrm{S})$ boundedly into itself, with*

$$\|\mathcal{K}\|_{\mathcal{L}(L_2,L_2)} \leqslant \left(M_a M_a'\right)^{1/2}, \tag{2.31}$$

where M_a, M_a' are the constants from (2.28)*. Moreover, $\mathcal{K}$ maps $L_2^+(\mathrm{D}\times\mathrm{S})$, the cone of non-negative functions in $L_2(\mathrm{D}\times\mathrm{S})$, into itself.*

To specify bounds for the operator norm $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)}$ we introduce the quantities

$$\gamma := \sup_{(x,\vec{s})\in \mathrm{D}\times\mathrm{S}} \left\{ \frac{\bar{\sigma}(x,\vec{s})}{\sigma(x,\vec{s})}, \frac{\bar{\sigma}'(x,\vec{s})}{\sigma(x,\vec{s})} \right\}, \qquad \zeta := \frac{\gamma\sigma_{\max}}{\sigma_{\min}}. \tag{2.32}$$

Lemma 2.14. *Under assumptions (2.28) on the optical parameters,*

$$\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \leqslant \min\left\{ \zeta, (\sigma_{\max}-\alpha)/\sigma_{\min}, (M_a M_a')^{1/2} \min\left\{ \hat{\ell}, \sqrt{\hat{\ell}/2\sigma_{\min}} \right\} \right\}. \tag{2.33}$$

Proof. Combining (2.31) and (2.18) yields that

$$\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \leqslant (M_a M_a')^{1/2} \min\left\{ \hat{\ell}, \sqrt{\hat{\ell}/2\sigma_{\min}} \right\}.$$

To prove that $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \leqslant \min\{\zeta, (\sigma_{\max}-\alpha)/\sigma_{\min}\}$, we proceed as follows. For any $\varphi \in L_2(\mathrm{D}\times\mathrm{S})$ we have $\mathcal{K}\varphi \in L_2(\mathrm{D}\times\mathrm{S})$ so that there exists a unique $w \in H_{0,-}(\mathrm{D}\times\mathrm{S})$ such that $\mathcal{T}w = \mathcal{K}\varphi$. Thus, it suffices to prove that $\|w\|_{L_2(\mathrm{D}\times\mathrm{S})} \leqslant \min\{\zeta, (\sigma_{\max}-\alpha)/\sigma_{\min}\}\|\varphi\|_{L_2(\mathrm{D}\times\mathrm{S})}$. Since $\mathcal{A}$ is accretive on $H_{0,-}(\mathrm{D}\times\mathrm{S})$, we have

$$(\mathcal{K}\varphi, w) = (\mathcal{A}w, w) + (\sigma w, w) \geqslant (\sigma w, w) \geqslant \sigma_{\min}\|w\|^2_{L^2(\mathrm{D}\times\mathrm{S})}. \tag{2.34}$$

Furthermore,

$$\begin{aligned}
(\mathcal{K}\varphi, w) &\leqslant \int_{\mathrm{D}\times\mathrm{S}\times\mathrm{S}} |w(x,\vec{s})| K(x,\vec{s}',\vec{s}) |\varphi(x,\vec{s}')| \,\mathrm{d}x\,\mathrm{d}\vec{s}\,\mathrm{d}\vec{s}' \\
&\leqslant \left(\int_{\mathrm{D}\times\mathrm{S}} |w(x,\vec{s})|^2 \bar{\sigma}'(x,\vec{s}) \,\mathrm{d}x\,\mathrm{d}\vec{s} \right)^{1/2} \left(\int_{\mathrm{D}\times\mathrm{S}} |\varphi(x,\vec{s}')|^2 \bar{\sigma}(x,\vec{s}') \,\mathrm{d}x\,\mathrm{d}\vec{s}' \right)^{1/2} \\
&\leqslant \min\{\sigma_{\max}-\alpha, \gamma\sigma_{\max}\} \|w\|_{L^2(\mathrm{D}\times\mathrm{S})} \|\varphi\|_{L^2(\mathrm{D}\times\mathrm{S})},
\end{aligned}$$

where we have used Cauchy–Schwarz' inequality. Combining this with (2.34) yields the desired inequality $\|w\|_{L_2(\mathrm{D}\times\mathrm{S})} \leqslant \min\{\zeta, (\sigma_{\max}-\alpha)/\sigma_{\min}\}\|\varphi\|_{L_2(\mathrm{D}\times\mathrm{S})}$. □

It follows from (2.33) that having

$$\min\left\{ \zeta, (\sigma_{\max}-\alpha)/\sigma_{\min}, (M_a M_a')^{1/2} \min\left\{ \hat{\ell}, \sqrt{\hat{\ell}/2\sigma_{\min}} \right\} \right\} < 1$$

is a sufficient condition for $\mathcal{T}^{-1}\mathcal{K}$ to be a contraction. From this we can distinguish two different "physical regimes" that ensure contractivity:

- having $\zeta < 1$ or $(\sigma_{\max}-\alpha)/\sigma_{\min} < 1$ can be interpreted as quantifying the dominance of transport with respect to scattering with $\sigma(x,\vec{s})$ not varying too much in its arguments. This condition is a quantification of the well-known fact that DOM converges at a slower rate when collisions become more and more significant with respect to transport.
- having $(M_a M_a')^{1/2} \min\left\{ \hat{\ell}, \sqrt{\hat{\ell}/2\sigma_{\min}} \right\} < 1$ happens when $\hat{\ell} = \mathrm{diam}(\mathrm{D})$ is sufficiently small or $\sigma_{\min}/M_a M_a'$ sufficiently large, which is another expression to quantify how much transport effects dominate with respect to the scattering.

Of course, these conditions cannot be expected to hold in all relevant application scenarios. However, they are going to play a crucial role in what we call *preconditioning* on the continuous level, ensuring convergence in the infinite-dimensional continuous case.

3. Step (II)—Idealized iterations

We are now prepared to identify viable outer iterations of the form

$$u_{n+1} = u_n + \mathcal{P}(f - \mathcal{B}u_n), \quad n = 0, 1, 2, \dots, \tag{3.1}$$

(see Step (II) in Section 1.3). In the following, we will work with the pair of trial and test spaces U, V, given in (F1), that is,

$$U = L_2(\mathrm{D} \times \mathrm{S}), \qquad V = H_{0,+}(\mathrm{D} \times \mathrm{S}),$$

where we abbreviate in what follows $\|v\|_V := \|v\|_{H(\mathrm{D}\times\mathrm{S})}$. Of course, the *preconditioner* $\mathcal{P} \in \mathcal{L}(V', U)$ is to be chosen in such a way that

$$\exists\, \rho < 1 \text{ such that } \|u_{n+1} - u\|_U \leqslant \rho \|u_n - u\|_U, \quad n \in \mathbb{N}, \tag{3.2}$$

which holds if and only if $\|\mathrm{id} - \mathcal{P}\mathcal{B}\|_{\mathcal{L}(U,U)} \leqslant \rho < 1$. Note that for the variational formulation (F1) the residual $f - \mathcal{B}v$ is, by Theorem 2.9, well-defined in V' for any $v \in U$.

Recalling Remark 2.12, we consider two distinct problem regimes.

Remark 3.1. The operator equation $\mathcal{B}u = f$ implies homogeneous inflow-boundary conditions. Incorporating inhomogeneous boundary conditions could be treated by taking any function w in the domain of $\mathcal{B}$ that satisfies the required boundary conditions and subtract $f_b := \mathcal{B}w$ from f reducing the problem to homogeneous conditions.

3.1. Dominating transport: $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \leqslant \rho < 1$.

If we have the contraction

$$\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \leqslant \rho < 1,$$

then $\mathcal{P} := \mathcal{T}^{-1}$ is an admissible preconditioner. In fact, iteration (3.1) becomes

$$u_{n+1} = u_n + \mathcal{T}^{-1}(f - \mathcal{B}u_n) = \mathcal{T}^{-1}(\mathcal{K}u_n + f), \quad n \in \mathbb{N}_0, \tag{3.3}$$

and obviously satisfies (3.2), ensuring convergence in U to the solution u of the radiative transfer problem

$$\mathcal{B}u = (\mathcal{T} - \mathcal{K})u = f.$$

In particular, it follows that for any initial guess u_0

$$\|u - u_n\|_U \leqslant \rho^n \|u - u_0\|_U. \tag{3.4}$$

3.2. Dominating scattering: $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \geqslant 1$.

Throughout this section we continue to assume that (2.24) holds with some $\alpha > 0$.

To find a substitute for the preconditioner $\mathcal{P} = \mathcal{T}^{-1}$ of the transport dominated regime, consider for some fixed $a > 0$

$$\mathcal{T}_a := \mathcal{T} + a\,\mathrm{id}, \qquad \mathcal{B}_a := \mathcal{T}_a - \mathcal{K},$$

and take $\mathcal{P} := \mathcal{B}_a^{-1}$ in (3.1). This leads to the (ideal) iteration

$$u_{n+1} = u_n + (\mathcal{T}_a - \mathcal{K})^{-1}(f - (\mathcal{T} - \mathcal{K})u_n) = a\,\mathcal{B}_a^{-1}\big(u_n + a^{-1}f\big), \quad n \in \mathbb{N}_0, \tag{3.5}$$

where we have used that $(\mathcal{T}_a - \mathcal{K})^{-1}(\mathcal{T} - \mathcal{K}) = (\mathcal{T}_a - \mathcal{K})^{-1}(\mathcal{T}_a - \mathcal{K} - a\mathrm{id}) = -\mathrm{id} + a(\mathcal{T}_a - \mathcal{K})^{-1}$.

Thus, to ensure convergence we need that $\|a(\mathcal{T}_a - \mathcal{K})^{-1}\|_{\mathcal{L}(U,U)}$ is a contraction. Note that this is satisfied for any $a > 0$ since, by Proposition 2.11, we have that $(\mathcal{B}_a v, v) \geqslant \alpha + a$, which by Theorem 2.9 gives

$$\|a(\mathcal{T}_a - \mathcal{K})^{-1}\|_{\mathcal{L}(U,U)} \leqslant \frac{a}{a + \alpha} < 1. \tag{3.6}$$

So (3.5) converges in $U = L_2(\mathrm{D}\times\mathrm{S})$ to the true solution u with the error reduction rate $a/(a+\alpha)$ for any fixed $a > 0$.

Remark 3.2. Notice that $\mathcal{P} = \mathcal{B}_a^{-1}$ can be derived from a different perspective. Consider the time dependent initial-boundary value problem

$$\partial_t u + \mathcal{T}u - \mathcal{K}u = f, \qquad u(0,\cdot) = u^0 \text{ in } \mathrm{D}, \qquad u\big|_{\Gamma_-} = 0,$$

(where f, $\mathcal{T}$, $\mathcal{K}$ are still independent of t). Denoting by u_n the approximation of $u(t_n)$, $t_n = n\tau$, its backward-Euler semi-discretization in time reads

$$\frac{u_{n+1}-u_n}{\tau} + \mathcal{T}u_{n+1} - \mathcal{K}u_{n+1} = f, \quad n \in \mathbb{N}_0,$$

which gives

$$(\tau^{-1}\mathrm{id} + \mathcal{T} - \mathcal{K})u_{n+1} = \tau^{-1}u_n + f, \quad n \in \mathbb{N}_0.$$

This coincides with (3.5) for $a = \tau^{-1}$.

4. Step (III)—Perturbed iterations and the main algorithm

The practical realization of the scheme boils down to two tasks:

(T1) *Formulate a perturbed version of algorithms* (3.3) *and* (3.5) *with suitable error tolerances* η_n *that still guarantee convergence to the exact continuous solution.*

For this task, it will be convenient to use the following notational convention: Given an operator $\mathcal{G} \in \mathcal{L}(U,Y)$, we denote for any $\eta > 0$ by $[\mathcal{G}, w; \eta]$ an element in Y satisfying $\|\mathcal{G}w - [\mathcal{G}, w; \eta]\|_Y \leqslant \eta$. Specifically, for our purposes we require a routine to approximately apply the kernel, that is,

$$[\mathcal{K}, v; \eta] \to z_\eta \quad \text{such that} \quad \|\mathcal{K}v - z_\eta\|_{V'} \leqslant \eta. \tag{4.1}$$

Likewise the source is generally not given exactly and has to be approximated

$$[f; \eta] \to f_\eta \quad \text{such that} \quad \|f - f_\eta\|_{V'} \leqslant \eta. \tag{4.2}$$

The approximation $[f;\eta]$ of f depends on how the data are given. Finally, given a right hand side $g \in V'$, we have to provide a transport solver

$$[\mathcal{T}^{-1}, g; \eta] \to u_\eta \quad \text{such that} \quad \|u_\eta - \mathcal{T}^{-1}g\|_U \leqslant \eta, \tag{4.3}$$

where, as before, $\mathcal{T}$ is viewed as a mapping from U onto V' with $U = L_2(\mathrm{D}\times\mathrm{S})$, $V = H_{0,+}(\mathrm{D}\times\mathrm{S})$.

(T2) *Specify how to realize the above routines in* (4.1), (4.2), *and* (4.3).

In this section we concentrate only on (T1) and *assume* for the moment that the routines (4.1), (4.2), and (4.3) are available. These routines are detailed later on in Sections 5 and 6.

4.1. Dominating transport: $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \leqslant \rho < 1$.

An approximate realization of the ideal scheme (3.3) is

$$\bar{u}_{n+1} = [\mathcal{T}^{-1}, [\mathcal{K}, \bar{u}_n; \eta_{\mathcal{K}}] + [f; \eta_f]; \eta_{\mathcal{T}}], \quad n \geqslant 0. \tag{4.4}$$

In the following we take for simplicity $u_0 = 0$. Any other choice for u_0 that exploits additional information would, of course, be possible. We choose the individual tolerances proportional to

$$\eta_n = (1+n)^{-\beta}\rho^n \tag{4.5}$$

for some fixed $\beta > 1$ ($\beta = 1.5$ in later numerical experiments). Specifically, we set

$$\eta_{\mathcal{K}} := \kappa_1 \eta_n, \quad \eta_f := \kappa_2 \eta_n, \quad \eta_{\mathcal{T}} := \kappa_3 \eta_n,$$

where the parameters $\kappa_1, \kappa_2, \kappa_3 \geqslant 0$ satisfy

$$C_{\mathcal{T}}(\kappa_1 + \kappa_2) + \kappa_3 \leqslant 1, \tag{4.6}$$

with the upper bound $\|\mathcal{T}^{-1}\|_{\mathcal{L}(V',U)} \leqslant C_{\mathcal{T}}$ from (2.18).

In addition we need an upper bound for $\|u\|_U$. A first simple estimate that can be obtained from (2.24) or (2.27)

$$\|u\|_U \leqslant \|\mathcal{B}^{-1}\|_{\mathcal{L}(V',U)} \|f\|_{V'} \leqslant \alpha^{-1} \|f\|_{L_2(\mathrm{D})}. \tag{4.7}$$

Since this may be rather pessimistic when α is small we take

$$b_0(u) := \alpha^{-1} \|f\|_{L_2(\mathrm{D})}$$

only as an *initialization* which is refined during the course of the iteration based on a posteriori information. In the following, we will work with

$$b_{n+1}(u) := \min\left\{b_n(u), \|\bar{u}_{n+1}\|_U + (\rho\, b_n(u) + \zeta(\beta))\rho^{n-1}\right\}, \quad n \geqslant 0,$$

which is an upper bound that converges to $\|u\|_U$.

We are now prepared to present a detailed account of the perturbed iteration (4.4) in terms of the following Algorithm 1 called *Adaptive Source Term Iteration* (ASTI). We prove in Theorem 4.1 that for dominating transport ASTI$[\mathcal{T}, \mathcal{K}, f; \varepsilon]$ computes an approximate solution u_ε such that $\|u - u_\varepsilon\|_U \leqslant \varepsilon$.

Algorithm 1 ASTI$[\mathcal{T}, \mathcal{K}, f; \varepsilon] \to u_\varepsilon$

Fix κ_1, κ_2, κ_3 according to (4.6), fix $\beta > 1$, estimate ρ by (2.33), and choose $b_0(u)$, e.g., as in (4.7).
$n \leftarrow 0$
$\bar{u}_n \leftarrow 0$
$\mathrm{err} \leftarrow b_0(u)$
$b(u) \leftarrow b_0(u)$
while $\mathrm{err} > \varepsilon$ **do**
  $\eta_n \leftarrow (1+n)^{-\beta}\rho^n$
  $w \leftarrow [\mathcal{K}, \bar{u}_n; \kappa_1 \eta_n]$
  $g \leftarrow [f; \kappa_2 \eta_n]$
  $\bar{u}_{n+1} \leftarrow [\mathcal{T}^{-1}, w + g; \kappa_3 \eta_n]$
  $\mathrm{err} \leftarrow (\rho\, b(u) + \zeta(\beta))\rho^n$
  $b(u) \leftarrow \min\left\{b(u), \|\bar{u}_{n+1}\|_U + (\rho b(u) + \zeta(\beta))\rho^{n-1}\right\}$
  $n \leftarrow n + 1$
end while
$u_\varepsilon \leftarrow \bar{u}_n$

4.2. **Dominating scattering: $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \geqslant 1$.** For a given $a > 0$, the approximate realization of the scheme (3.5) takes the form

$$\bar{u}_{n+1} = [a\mathcal{B}_a^{-1}, \bar{u}_n + [a^{-1} f; \eta_n]; \eta_n], \quad n \in \mathbb{N}_0,$$

where the stage dependent tolerances η_n are chosen as in (4.5).

To render the approximate application of the preconditioner $a\mathcal{B}_a^{-1}$ practical, we choose the parameter a in such a way that the operator $\mathcal{B}_a$ is *transport dominated*, so that we can resort to the ASTI algorithm for its approximate inversion. To that end, recall from (2.33) that $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)}$ is estimated in terms of quantities ζ, γ from (2.32). When $\mathcal{T}$ is replaced by $\mathcal{T}_a$ these quantities depend on a and are therefore denoted for clarity by γ_a, ζ_a. Since the quantities $\bar{\sigma}$, $\bar{\sigma}'$ are not affected by the parameter a, we have

$$\gamma_a \leqslant \frac{\sigma_{\max} - \alpha}{\sigma_{\min} + a}, \qquad \zeta_a \leqslant \frac{(\sigma_{\max} - \alpha)(\sigma_{\max} + a)}{(\sigma_{\min} + a)(\sigma_{\min} + a)}.$$

In view of the bound (3.6) for $\|a\mathcal{B}_a^{-1}\|_{\mathcal{L}(U,U)}$, by choosing the parameter $a = a^*$ as the unique solution of

$$\frac{a}{a+\alpha} = \frac{(\sigma_{\max} - \alpha)(\sigma_{\max} + a)}{(\sigma_{\min} + a)(\sigma_{\min} + a)}, \tag{4.8}$$

one obtains simultaneously

$$\|a^*\mathcal{B}_{a^*}^{-1}\|_{\mathcal{L}(U,U)} \leqslant \rho^* \quad \text{and} \quad \|\mathcal{T}_{a^*}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} \leqslant \rho^* \quad \text{for some } \rho^* < 1. \tag{4.9}$$

Thus, an error controlled application of the preconditioner $a^*\mathcal{B}_{a^*}^{-1}$ is given for any right hand side g and accuracy η as

$$[\mathcal{B}_{a^*}^{-1}, g; \eta] = \mathrm{ASTI}[\mathcal{T}_{a^*}, \mathcal{K}, g; \eta].$$

Note that the algorithm consists now in nesting the outer iteration with an inner ASTI iteration for the application of the preconditioner. It is thus straighforward to formulate a general *Nested ASTI* scheme, where N-ASTI$[\mathcal{B}, f; \varepsilon]$ generates an approximate solution u_ε such that $\|u - u_\varepsilon\|_U \leqslant \varepsilon$ even when scattering dominates in $\mathcal{B}$ (see Algorithm 2).

4.3. Convergence of N-ASTI$[\mathcal{B}, f; \varepsilon]$.

Theorem 4.1. *For any target accuracy $\varepsilon > 0$, Algorithm 2 terminates and its output*

$$u_\varepsilon := \text{N-ASTI}[\mathcal{B}, f; \varepsilon]$$

satisfies

$$\|u - u_\varepsilon\|_U \leqslant \varepsilon,$$

where u is the exact solution of (1.3) with respect to the variational formulation (F1).

Proof. We first consider the transport dominated case where $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)} < 1$. The algorithm then reduces to ASTI, that is,

$$u_e = \mathrm{ASTI}[\mathcal{T}, \mathcal{K}, f; \varepsilon].$$

Let u_n denote the exact iterates of (3.3) and $\bar{u}_n$ the ones from the perturbed version (4.4). By the definition of the respective routines we have for given tolerances $\eta_{\mathcal{T}}$, $\eta_{\mathcal{K}}$, η_f

$$\begin{aligned} u_{n+1} - \bar{u}_{n+1} &= \mathcal{T}^{-1}(\mathcal{K}u_n + f) - [\mathcal{T}^{-1}, [\mathcal{K}, \bar{u}_n; \eta_{\mathcal{K}}] + [f; \eta_f]; \eta_{\mathcal{T}}] \\ &= \mathcal{T}^{-1}\big(\mathcal{K}(u_n - \bar{u}_n)\big) + \mathcal{T}^{-1}(\mathcal{K}\bar{u}_n - [\mathcal{K}, \bar{u}_n; \eta_{\mathcal{K}}]) + \mathcal{T}^{-1}(f - [f; \eta_f]) \\ &\quad + \mathcal{T}^{-1}([\mathcal{K}, \bar{u}_n; \eta_{\mathcal{K}}] + [f; \eta_f]) - [\mathcal{T}^{-1}, [\mathcal{K}, \bar{u}_n; \eta_{\mathcal{K}}] + [f; \eta_f]; \eta_{\mathcal{T}}]. \end{aligned}$$

Algorithm 2 N-ASTI$[\mathcal{B}, f; \varepsilon] \to u_\varepsilon$

$\rho \leftarrow$ Estimate $\|\mathcal{T}^{-1}\mathcal{K}\|_{\mathcal{L}(U,U)}$ using upper bound of (2.33).
if $\rho < 1$ **then** ▷ Dominating transport
  $u_\varepsilon \leftarrow \mathrm{ASTI}[\mathcal{T}, \mathcal{K}, f; \varepsilon]$
else ▷ Dominating scattering
  Estimate a^* from (4.8), estimate ρ^* from (4.9), fix $\beta > 1$.
  $n \leftarrow 0$
  $\bar{u}_n \leftarrow 0$
  $\mathrm{err} \leftarrow b_0(u)$
  $b(u) \leftarrow b_0(u)$
  while $\mathrm{err} > \varepsilon$ **do**
    $\eta_n \leftarrow (1+n)^{-\beta}(\rho^*)^n$
    $g \leftarrow \bar{u}_n + [(a^*)^{-1} f; \eta_n]$
    $\bar{u}_n = a^* \mathrm{ASTI}[\mathcal{T}_{a^*}, \mathcal{K}, g; \varepsilon]$
    $\mathrm{err} \leftarrow (\rho^* b(u) + (1+a^*)\zeta(\beta))\,(\rho^*)^n$
    $b(u) \leftarrow \min\{b(u), \|\bar{u}_{n+1}\|_U + ((\rho^*)b(u) + (1+a^*)\zeta(\beta))(\rho^*)^{n-1}\}$
    $n \leftarrow n+1$
  end while
  $u_\varepsilon \leftarrow \bar{u}_n$
end if
return u_ε

By the triangle inequality, bound (2.18) on $\|\mathcal{T}^{-1}\|_{\mathcal{L}(V',U)}$, and the properties of the routines, we obtain

$$\|u_{n+1} - \bar{u}_{n+1}\|_U \leqslant \rho\|u_n - \bar{u}_n\|_U + C_{\mathcal{T}}(\eta_{\mathcal{K}} + \eta_f) + \eta_{\mathcal{T}}.$$

For $\bar{u}_0 = u_0$ and with the choice $\eta_{\mathcal{K}} := \kappa_1\eta_n$, $\eta_f := \kappa_2\eta_n$ and $\eta_{\mathcal{T}} := \kappa_3\eta_n$ and (4.6), we get

$$\|u_{n+1} - \bar{u}_{n+1}\|_U \leqslant \rho\|u_n - \bar{u}_n\|_U + \eta_n,$$

which, by induction, yields

$$\|\bar{u}_{n+1} - u_{n+1}\|_U \leqslant \sum_{j=0}^{n} \rho^j \eta_{n-j}.$$

Specifically, taking the same η_n as in (4.5) for some fixed $\beta > 1$, we obtain

$$\|\bar{u}_{n+1} - u_{n+1}\|_U \leqslant \sum_{j=0}^{n} \rho^j \rho^{n-j}(1+(n-j))^{-\beta} = \rho^n \sum_{j=0}^{n} (1+j)^{-\beta} \leqslant \zeta(\beta)\rho^n, \tag{4.10}$$

where $\zeta(\beta) := \sum_{j\in\mathbb{N}} j^{-\beta}$ is the ζ-function. Hence, by triangle inequality

$$\|u - \bar{u}_{n+1}\|_U \leqslant \rho^{n+1}\|u\|_U + \zeta(\beta)\rho^n. \tag{4.11}$$

Thus, whenever at the nth stage of the algorithm $\|u\|_U \leqslant b_n(u)$, we conclude that

$$b_{n+1}(u) := \min\{b_n(u), \|\bar{u}_{n+1}\|_U + (\rho\, b_n(u) + \zeta(\beta))\rho^{n-1}\} \tag{4.12}$$

a bound for $\|u\|_U$ which converges to $\|u\|_U$. This yields the computable error bound

$$\|u - \bar{u}_{n+1}\|_U \leqslant (\rho\, b_{n+1}(u) + \zeta(\beta))\rho^n \tag{4.13}$$

which completes the proof for the transport dominated case.

For dominating scattering, denoting by u_n the exact iterates

$$u_{n+1} = a^*\mathcal{B}_{a^*}^{-1}(u_n + (a^*)^{-1}f), \quad n \in \mathbb{N}_0,$$

we readily obtain

$$\begin{aligned}\bar{u}_{n+1} - u_{n+1} = {} & [a^*\mathcal{B}_{a^*}^{-1}, \bar{u}_n + [(a^*)^{-1}f; \eta_n]; \eta_n] - a^*\mathcal{B}_{a^*}^{-1}(\bar{u}_n + [(a^*)^{-1}f; \eta_n]) \\ & + a^*\mathcal{B}_{a^*}^{-1}(\bar{u}_n + [(a^*)^{-1}f; \eta_n]) - a^*\mathcal{B}_{a^*}^{-1}(\bar{u}_n \\ & + (a^*)^{-1}f) + a^*\mathcal{B}_{a^*}^{-1}(\bar{u}_n - u_n).\end{aligned}$$

Hence,

$$\|\bar{u}_{n+1} - u_{n+1}\|_U \leqslant a^*\eta_n + \rho^*\eta_n + \rho^*\|\bar{u}_n - u_n\|_U.$$

We obtain as earlier with $\bar{u}_0 = u_0$

$$\|\bar{u}_{n+1} - u_{n+1}\|_U \leqslant (1 + a^*)\sum_{j=0}^{n}(\rho^*)^j\eta_{n-j}.$$

Specifically, taking η_n from (4.5) we get, on account of (3.4),

$$\|u - \bar{u}_n\|_U \leqslant \big(\rho^*\|u - u_0\|_U + (1 + a^*)\zeta(\beta)\big)(\rho^*)^{n-1}, \quad n \in \mathbb{N}, \tag{4.14}$$

and hence the same type of bound as in (4.11) for the transport dominated case. □

Remark 4.2. The recursion (4.12) successively mitigates a possibly over-pessimistic initial bound $b_0(u)$. It can be further improved by using the a posteriori bound $\|u - u_n\|_U \leqslant \frac{\rho}{1-\rho}\|u_n - u_{n-1}\|_U$. We also have (for $n \geqslant 2$)

$$\begin{aligned}\|u - u_n\|_U &\leqslant \frac{\rho}{1-\rho}\big\{\|\bar{u}_n - \bar{u}_{n-1}\|_U + \|u_n - \bar{u}_n\|_U + \|u_{n-1} - \bar{u}_{n-1}\|_U\big\} \\ &\leqslant \frac{\rho}{1-\rho}\big\{\|\bar{u}_n - \bar{u}_{n-1}\|_U + \zeta(\beta)(\rho^{n-1} + \rho^{n-2})\big\},\end{aligned}$$

which is a computable bound replacing $\|u - u_n\|_U$. However, the calculation of these a posteriori quantities would require storing two consecutive outer iterates.

4.4. **Complexity.** We conclude with some qualitative complexity estimates. Further quantifications depend on the realizations of the involved routines. The number $n(\varepsilon)$ of outer iteration steps required to realize $\|u - \bar{u}_{n(\varepsilon)}\|_U \leqslant \varepsilon$ is given by

$$n(\varepsilon) = \left\lceil \frac{|\ln\varepsilon| + \ln(\rho b(u) + a^*\zeta(\beta))}{|\ln\rho|} \right\rceil. \tag{4.15}$$

As detailed in the subsequent section the approximate application of the scatterer is typically dominated by the approximate inversion of the transport operator. As a consequence, in either version of the outer iteration the computational work per outer iteration step n is dominated by the computational cost $\mathbf{cost}_{\mathcal{P}}(\eta_n)$ of the preconditioner. Hence, the complexity $\mathbf{cost}_{\mathcal{B}^{-1}}(\varepsilon)$ of solving $\mathcal{B}u = f$ within accuracy ε can be bounded as

$$\mathbf{cost}_{\mathcal{B}^{-1}}(\varepsilon) \lesssim \sum_{j=1}^{n(\varepsilon)} \mathbf{cost}_{\mathcal{P}}(\eta_n).$$

Assuming that $\mathbf{cost}_{\mathcal{P}}(\eta) \lesssim \eta^{-\vartheta}$ holds for some positive ϑ (which is actually realistic as will be seen later), this yields

$$\begin{aligned}\mathbf{cost}_{\mathcal{B}^{-1}}(\varepsilon) &\lesssim \sum_{j=1}^{n(\varepsilon)} \rho^{-j\vartheta}(1+j)^{\beta\vartheta} \leqslant (1+n(\varepsilon))^{\beta\vartheta} \sum_{j=0}^{n(\varepsilon)} \rho^{-j\vartheta} \\ &\leqslant \frac{\rho^{-n(\varepsilon)\vartheta}}{1-\rho^{\vartheta}}(1+n(\varepsilon))^{\beta\vartheta} \leqslant C\varepsilon^{-\vartheta}|\ln\varepsilon|^{\beta\vartheta},\end{aligned}$$

where $C = C(\beta, \vartheta, \rho, u)$ is a constant depending on β, ϑ, ρ and a bound $b(u)$ for $\|u - u_0\|_U$. As a result, the cost of approximately inverting $\mathcal{B}$ is, up to a logarithmic factor, of the order of the one for the application of the preconditioner with the same accuracy, that is,

$$\mathbf{cost}_{\mathcal{B}^{-1}}(\varepsilon) \lesssim |\ln\varepsilon|^{\beta\vartheta}\mathbf{cost}_{\mathcal{P}}(\varepsilon).$$

The cost of the preconditioner, in turn, depends on the problem regime. For dominating transport $\mathbf{cost}_{\mathcal{P}}(\varepsilon) = \mathbf{cost}_{\mathcal{T}^{-1}}(\varepsilon)$, while for dominating scattering the approximate application of $a^*\mathcal{B}_{a^*}$ within accuracy ε requires (in the inner iteration) invoking $\mathcal{O}(|\ln\varepsilon|/|\ln\rho^*|)$ times an ε-accurate transport solve, i.e., $\mathbf{cost}_{\mathcal{P}}(\varepsilon) \lesssim \mathbf{cost}_{\mathcal{T}^{-1}}(\varepsilon)|\ln\varepsilon|/|\ln\rho^*|$.

In summary, the overall computational complexity for a given target accuracy is essentially determined by the cost of error-controlled transport solves (provided that a reasonably efficient approximate application scheme for the scatterer is at hand). A posteriori bounds for transport solvers are therefore pivotal. Moreover, since the target tolerances η_n are gradually tightened, early stages of the outer iteration (and its preconditioners) require only correspondingly cruder accuracy tolerances so that (up to logarithmic factors) the total complexity is dominated by the cost of the last outer iteration step.

The remainder of the paper is devoted to realizations of $[\mathcal{K}, v; \eta]$ and $[\mathcal{T}^{-1}, g; \eta]$.

5. The routine $[\mathcal{K}, v; \eta]$

5.1. **Introductory comments.** The scheme ASTI requires the application of the global operator $\mathcal{K}$ within dynamically updated accuracy tolerances. We present in this section an efficient error-controlled approximate application scheme that makes use of *wavelet-compression* and *low-rank* approximations. Fully non-linear versions with even better scaling are postponed to forthcoming work.

We confine the discussion to the class of kernels of the form (2.29), that is, $K(x, \vec{s}, \vec{s}') = \kappa(x)G(\vec{s}, \vec{s}')$, $G(\vec{s}, \vec{s}') = G(\vec{s}', \vec{s})$, with $G(\vec{s}, \vec{s}') \geqslant 0$, $\vec{s}, \vec{s}' \in \mathrm{S}$, $\kappa \geqslant \kappa_0 > 0$, and the normalization

$$\int_{\mathrm{S}} G(\vec{s}, \vec{s}')\,\mathrm{d}\vec{s}' = \int_{\mathrm{S}} G(\vec{s}, \vec{s}')\,\mathrm{d}\vec{s} = 1, \quad \vec{s}, \vec{s}' \in \mathrm{S}.$$

In the following, we adhere to the notation

$$\mathcal{K}_0 v := \int_{\mathrm{S}} G(\cdot, \vec{s}')v(\vec{s}')\,\mathrm{d}\vec{s}'.$$

The simplest examples are *isotropic* and *Rayleigh-type scattering* which are, respectively, of the form

$$G(\vec{s}, \vec{s}') := |\mathrm{S}|^{-1}, \quad G(\vec{s}, \vec{s}') = c\left(1 + (\vec{s}\cdot\vec{s}')^2\right).$$

Another variant of interest, used in [23], is given in terms of the similar expansion

$$G(\vec{s},\vec{s}') = \sum_{n=0}^{\infty} a_n T_n(\vec{s}\cdot\vec{s}'),$$

with $a_n \geqslant 0$ and T_n being the nth Chebyshev polynomial, $T_n(x) := \cos(n \arccos(x))$ for $|x| \leqslant 1$. It is shown in [23, Lemmata 2 and 3] that $\mathcal{K}$ is positive semi-definite with this type of kernel.

In our numerical scheme we focus on *Henyey–Greenstein-type* scattering represented by

$$G_\gamma(\vec{s},\vec{s}') := \begin{cases} \frac{1}{2\pi}\frac{1-\gamma^2}{1+\gamma^2-2\gamma\vec{s}\cdot\vec{s}'} & \text{if } d_{\mathrm{S}} = 1, \\ \frac{1}{4\pi}\frac{1-\gamma^2}{(1+\gamma^2-2\gamma\vec{s}\cdot\vec{s}')^{3/2}} & \text{if } d_{\mathrm{S}} = 2, \end{cases} \tag{5.1}$$

where $d_{\mathrm{S}} = d-1$ denotes the dimension of the parameter domain. This scattering model is widely used among physicists and was introduced in [21] to describe anisotropic effects via the parameter $-1 \leqslant \gamma \leqslant 1$. When $\gamma \geqslant 0$, the scattering is called *forward-peaked* and $\mathcal{K}_0$ is positive semi-definite. Moreover, for $d_{\mathrm{S}} = 2$ one has the expansion

$$\frac{1}{(1+\gamma^2-2\gamma\vec{s}\cdot\vec{s}')^{3/2}} = \sum_{n=0}^{\infty} \gamma^n P_n(\vec{s}\cdot\vec{s}'),$$

where P_n is the Legendre polynomial of degree n. Note that the closer γ comes to one, the slower is the decay and the larger is the model error when replacing G by a truncated expansion in favor of an efficient application of the scatterer to a given input.

Our focus on Henyey–Greenstein-type scattering is mainly motivated by the fact that varying the parameter γ allows us to quantitatively investigate different scattering regimes guiding the search for possibly different ways of exploiting sparsity.

The specification of $[\mathcal{K},\bar{u};\cdot]$ depends on the following *input format* of $\bar{u} \in L_2(\mathrm{D}\times\mathrm{S})$. As explained in Section 6, $\bar{u}$ is the output of a Discontinuous Petrov–Galerkin transport solver. It is a piecewise polynomial of degree m, subordinate to some current partition $\mathfrak{P}$ of the spatial domain D and whose coefficients are piecewise polynomials in the direction parameter $\vec{s} \in \mathrm{S}$. Thus, $\bar{u}$ has the form

$$\bar{u}(x,\vec{s}) = \sum_{T\in\mathfrak{P},i\in\mathcal{I}_T} v_{T,i}(\vec{s})\varphi_{T,i}(x), \tag{5.2}$$

where the spatial shape functions $\varphi_{T,i}$, $i \in \mathcal{I}_T$ are an orthonormal basis for $\mathbb{P}_m(T)$ and each parameter dependent coefficient $v_{T,i}$ is an element of $\mathbb{P}_M(\mathfrak{S})$ where $\mathfrak{S}$ is a partition of S. Hence,

$$(\mathcal{K}\bar{u})(x,\vec{s}) = \sum_{T\in\mathfrak{P}_h,i\in\mathcal{I}_T} (\mathcal{K}_0 v_{T,i})(\vec{s})\kappa(x)\varphi_{T,i}(x).$$

The simplest realization of $[\mathcal{K},\cdot;\cdot]$ rests on computing η-accurate approximations $w_{T,i} = [\mathcal{K}_0, v_{T,i};\eta]$ to $(\mathcal{K}_0 v_{T,i})$ so that (by orthonormality),

$$[\mathcal{K},\bar{u};\eta] := \sum_{T\in\mathfrak{P}_h,i\in\mathcal{I}_T} w_{T,i}\kappa\varphi_{T,i}, \qquad \|\mathcal{K}\bar{u} - [\mathcal{K},\bar{u};\eta]\|_U \leqslant \eta. \tag{5.3}$$

We focus therefore in what follows on the approximate application of $\mathcal{K}_0$ in the domain S.

5.2. **Matrix representations of $\mathcal{K}_0$, Alpert wavelets.** Suppose that $\Psi = \{\psi_\lambda \mid \lambda \in \Lambda\}$ is an *orthonormal* basis of $L_2(\mathrm{S})$ where Λ is a suitable infinite index set. Then, defining

$$G^\Psi_{\lambda,\lambda'} := (G, \psi_\lambda \otimes \psi_{\lambda'})_{\mathrm{S}\times\mathrm{S}} = (\psi_\lambda, \mathcal{K}_0\psi_{\lambda'})_{\mathrm{S}}, \qquad \mathbf{G}^\Psi := \left(G^\Psi_{\lambda,\lambda'}\right)_{\lambda,\lambda'\in\Lambda},$$

one has

$$G(\vec{s},\vec{s}') = \sum_{\lambda,\lambda'\in\Lambda} G^\Psi_{\lambda,\lambda'}\psi_\lambda(\vec{s})\psi_{\lambda'}(\vec{s}'),$$

i.e., $\mathbf{G}^\Psi$ is an *exact representation* of the kernel G and the associated operator in terms of an *infinite* matrix. By orthonormality of Ψ we have

$$\|\mathbf{G}^\Psi\| := \|\mathbf{G}^\Psi\|_{\mathcal{L}(\ell_2(\Lambda),\ell_2(\Lambda))} = \|\mathcal{K}_0\|_{\mathcal{L}(L_2(\mathrm{S}),L_2(\mathrm{S}))}.$$

An η-accurate application of $\mathcal{K}_0$ will be accomplished by identifying a "compressed" finite submatrix $\mathbf{G}^\Psi_\eta$ of $\mathbf{G}^\Psi$ that reduces the approximate application of $\mathcal{K}_0$ to an efficient matrix-vector multiplication.

As an appropriate choice for Ψ we advocate so-called *Alpert wavelet bases* of (at least) degree M from (5.2). For the convenience of the reader we briefly recapitulate some basic features of Alpert wavelets and refer to [1] for further details.

Starting from some initial partition $\mathfrak{S}_0$ of S (which could be the trivial one $\{\mathrm{S}\}$) and fixing a rule for splitting each cell C in a given partition into a fixed number of "children" forming the refinement $\mathcal{C}(C)$ of C, repeated refinements generate an infinite "master-tree" $\mathbb{T}$ whose nodes are cells and whose edges connect parents with children. We call a finite subtree of $\mathbb{T}$ complete if a child of a cell C belongs to the subtree if and only if all of $\mathcal{C}(C)$ is contained in the subtree. We consider only complete subtrees. Then the set of *leaves* of such a finite subtree forms a so-called "admissible" partition $\mathfrak{S}$ of S whose "refinement history" is determined by the subtree, i.e., there is a one-to-one correspondence between such (possibly very non-uniform) partitions $\mathfrak{S}$ and subtrees $\mathbb{T}_\mathfrak{S}$ of $\mathbb{T}$. The $\vec{s}$-dependent coefficients $v_{T,i}, w_{T,i}$ in (4.10), (5.3) will always be piecewise polynomials of degree M on such admissible partitions. We will make use of two different representations of such piecewise polynomials as described next.

Let $\mathbb{P}_M(C)$ denote the space of polynomials of (total) degree at most M over the cell C. Given an admissible partition $\mathfrak{S}$ of S, let $\mathbb{P}_M(\mathfrak{S})$ denote the space of piecewise polynomials of degree at most M, subordinate to the partition $\mathfrak{S}$. A canonical basis for $\mathbb{P}_M(\mathfrak{S})$ is obtained by associating with each cell $C \in \mathfrak{S}$ an orthonormal basis

$$\Phi_C = \{\phi_\nu := \chi_C P_{C,i} \mid \nu := (C,i),\ P_{C,i} \in \mathbb{P}_M(C),\ i \in \mathcal{I}_M := \{1,\dots,\dim \mathbb{P}_M\}\},$$

which gives rise to what is sometimes referred to as the orthonormal *scaling function* basis

$$\Phi_\mathfrak{S} := \bigcup_{C\in\mathfrak{S}} \Phi_C = \{\phi_\nu \mid \nu \in \Gamma_\mathfrak{S}\}, \quad \Gamma_\mathfrak{S} := \{(C,i) \mid C \in \mathfrak{S},\ i \in \mathcal{I}_M\}\},$$

to be always understood with respect to the uniform Haar measure on S induced by a convenient parametrization, i.e., $\int_\mathrm{S} \mathrm{d}\vec{s} = 1$ and $(v,w)_\mathrm{S} = \int_\mathrm{S} vw\,\mathrm{d}\vec{s}$.

Alpert wavelets provide alternative bases for such spaces of piecewise polynomials that encode "updates" obtained by passing to a refined partition. They are therefore better suited for meeting variable target accuracies. Since $\mathbb{P}_M(C) \subset \mathbb{P}_M(\mathcal{C}(C))$ one

can determine an *orthonormal* set of piecewise polynomials in $\mathbb{P}_M(\mathcal{C}(C))$. Setting $\mathcal{J}_M := \{1, \dots, \dim(\mathbb{P}_M(\mathcal{C}(C)) - \dim \mathbb{P}_M(C))\}$,

$$\Psi_C := \{\psi_\lambda \mid \lambda := (C, r),\, r \in \mathcal{J}_M\} \subset \mathbb{P}_m(\mathcal{C}(C))$$

spanning the orthogonal complement $\mathbb{W}(C) := \mathbb{P}_M(\mathcal{C}(C)) \ominus \mathbb{P}_M(C)$ between two successive levels of piecewise polynomials. Obviously,

$$\Psi := \{\psi_\lambda \mid \lambda \in \Lambda\}, \quad \Lambda := \{\lambda = (C, r) \mid r \in \mathcal{J}_M,\, C \in \mathbb{T}\},$$

is an orthonormal basis for $L_2(\mathrm{S})$. Clearly, for any admissible partition $\mathfrak{S}$ of S one easily identifies the subset $\Psi_{\mathfrak{S}} = \{\psi_\lambda : \lambda \in \Lambda_{\mathfrak{S}}\} \subset \Psi$ which forms a basis for $\mathbb{P}_M(\mathfrak{S})$, namely

$$\Lambda_{\mathfrak{S}} := \{\lambda = (C, r) \mid r \in \mathcal{J}_M,\ C \in \mathbb{T}_{\mathfrak{S}}\}.$$

Alpert bases are easy to construct, in particular, for domains like S. It is well known that changing from a scaling function representation of an element in $\mathbb{P}_M(\mathfrak{S})$ to its Alpert wavelet representation (and vice versa) can be done at $\mathcal{O}(\#\mathfrak{S})$ cost with the aid of the fast wavelet transform. Accordingly, one can efficiently pass from a scaling function representation of a compressed kernel to its wavelet representation and vice versa.

Moreover, ψ_λ, $|\lambda| > 0$, have *vanishing moments* of order $M + 1$, i.e.,

$$(P, \psi_\lambda)_{\mathrm{S}} = 0 \quad \forall P \in \mathbb{P}_M(\operatorname{supp} \psi_\lambda). \tag{5.4}$$

This has two important consequences. First, whenever a submatrix $\mathbf{G}_\eta^\Psi$ of $\mathbf{G}^\Psi$ is obtained by discarding entries $G^\Psi_{\lambda,\lambda'}$ with $|\lambda| + |\lambda'| > 0$ the corresponding kernel G_η still satisfies

$$\int_{\mathrm{S}\times\mathrm{S}} G_\eta(\vec{s}, \vec{s}')\, \mathrm{d}\vec{s}\, \mathrm{d}\vec{s}' = 1.$$

Second, (5.4) will be shown next to imply that $\mathbf{G}^\Psi$ is *nearly sparse* which provides the basis for an error-controlled efficient application of $\mathcal{K}_0$ through *matrix compression*.

5.3. **Compression of $\mathbf{G}^\Psi$.** As a guiding example, let us consider the case $d = 2$ (two spatial variables) such that S is the unit circle and has dimension $d_{\mathrm{S}} = d - 1 = 1$. Note that the Henyey–Greenstein kernel is then of the form

$$G_\gamma(\theta, \theta') = c(H_\alpha \circ \delta)(\theta, \theta'), \quad H_\alpha(\varphi) := \frac{1}{1 - \alpha\cos\varphi}, \qquad \text{and} \qquad \delta(\theta, \theta') = \theta - \theta',$$

where $c = \frac{1-\gamma^2}{2\pi(1+\gamma^2)}$ and $\alpha = \frac{2\gamma}{1+\gamma^2}$.

Proposition 5.1. *In the above terms one has*

$$\begin{aligned} &\left|(G_\gamma)_{\lambda,\lambda'}\right| \\ &\lesssim 2^{-\left(M+1+\frac{d_{\mathrm{S}}}{2}\right)\left||\lambda'|-|\lambda|\right|} 2^{-(M+1+d_{\mathrm{S}})\min\{|\lambda|,|\lambda'|\}} \max_{\ell \leqslant M+1} \left(\operatorname{dist}(S_\lambda, S_{\lambda'}) + 2^{-|\lambda|}\right)^{M+1-\ell} \\ &\times \sup_{\theta\in S_\lambda, \theta'\in S_{\lambda'}} |H_\alpha^{(2M+2-\ell)}(\theta - \theta')|. \end{aligned} \tag{5.5}$$

Proof. Recall that for $\lambda = (C, r)$ one has $S_\lambda := \operatorname{supp}\psi_\lambda = C$. Let us denote then by θ_λ the center of gravity of S_λ. Without loss of generality we can assume that $|\lambda| \leqslant |\lambda'|$. Taylor expansion of G_γ at θ_λ, using $(M+1)$st order vanishing moments of ψ_λ, yields for integration with respect to θ

$$\int_{-\pi}^{\pi} H_\alpha(\theta-\theta')\psi_\lambda(\theta)\,\mathrm{d}\theta = \int_{-\pi}^{\pi} (\theta-\theta')^{M+1} H_\alpha^{(M+1)}(\tilde\theta_\lambda-\theta')\psi_\lambda(\theta)\,\mathrm{d}\theta,$$

where $\tilde\theta_\lambda$ is some point in S_λ. Expanding $Y(\theta') := (\theta-\theta')^{M+1} H_\alpha^{(M+1)}(\tilde\theta_\lambda-\theta')$ at $\theta_{\lambda'} \in S_{\lambda'}$, yields upon integrating now first with respect to θ' and using again $(M+1)$st order vanishing moments,

$$\left|(G_\gamma)_{\lambda,\lambda'}\right| \lesssim \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\psi_\lambda(\theta)||\psi_{\lambda'}(\theta')||\theta'-\theta_{\lambda'}|^{M+1}|Y^{(M+1)}(\tilde\theta_{\lambda'})|\,\mathrm{d}\theta'\,\mathrm{d}\theta.$$

Since $|\theta'-\theta_{\lambda'}| \lesssim 2^{-|\lambda'|}$, $\|\psi_\lambda\|_{L_1(S_\lambda)} \lesssim 2^{-d_S|\lambda|/2}$ and since by Leibniz' rule

$$\begin{aligned}|Y^{(M+1)}(\tilde\theta_{\lambda'})| \leqslant C_M \max_{\ell\leqslant M+1} \left(\operatorname{dist}(S_\lambda, S_{\lambda'}) + 2^{-|\lambda|}\right)^{M+1-\ell} \\ \times \sup_{\theta\in S_\lambda, \theta'\in S_{\lambda'}} |H_\alpha^{(2M+2-\ell)}(\theta-\theta')|,\end{aligned}$$

the assertion follows. □

Of course, for $\alpha < 1$ the terms

$$C(M,\alpha,\lambda,\lambda') := \max_{\ell\leqslant M+1} \left(\operatorname{dist}(S_\lambda, S_{\lambda'}) + 2^{-|\lambda|}\right)^{M+1-\ell} \sup_{\theta\in S_\lambda, \theta'\in S_{\lambda'}} |H_\alpha^{(2M+2-\ell)}(\theta-\theta')|$$

are finite. The closer α (and hence γ) gets to one the larger one expects the second factor to become for small $\operatorname{dist}(S_\lambda, S_{\lambda'})$. On the other hand, for larger $\operatorname{dist}(S_\lambda, S_{\lambda'})$ the second factor turns out to be very small. In summary $C(M,\alpha,\lambda,\lambda')$ is bounded by a constant that possibly grows when γ tends to one but for fixed γ decreases when $|\lambda|$, $|\lambda'|$ grow regardless of the distance between the respective supports. $C(M,\alpha,\lambda,\lambda')$ in turn becomes very small when $\operatorname{dist}(S_\lambda, S_{\lambda'}) > c_\gamma$ where c_γ decreases when γ tends to one. This is illustrated in Figure 1 reflecting the strong near-sparsity of the representation. Moreover, defining

$$d(\lambda,\lambda') := 2^{\min\{|\lambda|,|\lambda'|\}} \operatorname{dist}(S_\lambda, S_{\lambda'}),$$

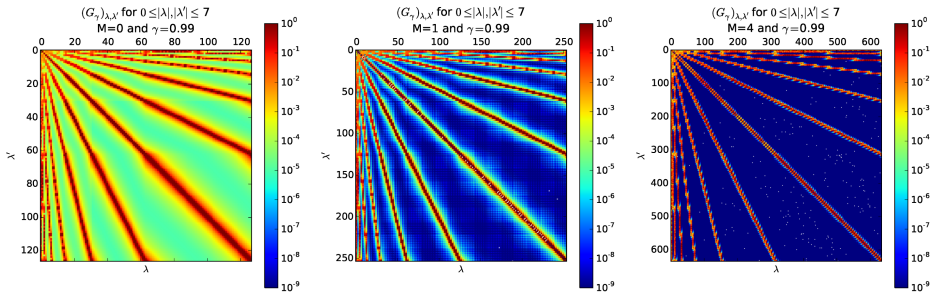


FIGURE 1. Alpert wavelet representation of $G_\gamma(\cdot,\cdot)$ of degrees $M = 0$, 1, and 4 and $\gamma = 0.99$.

and keeping in mind that $\operatorname{dist}(S_\lambda, S_{\lambda'})$ remains uniformly bounded because of the boundedness of S, one trivially has $d(\lambda,\lambda') \lesssim 2^{\min\{|\lambda|,|\lambda'|\}}$. Therefore, (5.5) yields the bounds

$$\left|(G_\gamma)_{\lambda,\lambda'}\right| \lesssim \frac{C(M,\alpha,\lambda,\lambda')2^{-\left(M+1+\frac{d_{\mathrm{S}}}{2}\right)\left||\lambda'|-|\lambda|\right|}}{(1+d(\lambda,\lambda'))^{M+1+d_{\mathrm{S}}}}. \tag{5.6}$$

Treating the term $C(M,\alpha,\lambda,\lambda')$ as a constant, this format allows us to directly invoke results on wavelet based matrix compression and corresponding *adaptive* approximate application tools; see, e.g., [9]. In particular, defining $s^* := (M+1)/d_{\mathrm{S}} = M+1$, (5.6) ensures that for each $s < s^*$ there exist positive summable sequences $(\alpha_j)_{j\in\mathbb{N}_0}$, $(\beta_j)_{j\in\mathbb{N}_0}$ and *compressed* versions $\mathbf{G}_J$ of $\mathbf{G}_\gamma = \mathbf{G}$, defined by concrete rules for replacing entries of $\mathbf{G}_\gamma$ by zero, such that

$$\|\mathbf{G}-\mathbf{G}_J\| \leqslant \beta_J 2^{-sJ}, \quad \#(\text{entries per row/column}) \leqslant \alpha_J 2^J, \quad J\in\mathbb{N}.$$

Here $\|\cdot\| = \|\cdot\|_{\mathcal{L}(\ell_2,\ell_2)}$ denotes the spectral norm.

5.4. **A linear compression scheme.** Since $\mathcal{K}_0$ is compact, (unlike the operators considered in [11]) the entries of $\mathbf{G}^\Psi$ decay uniformly with increasing level. Thus, given any "final" target accuracy ε, one can use (5.6) to find a level $L = L(\varepsilon)\in\mathbb{N}$ such that the finite matrix $\mathbf{G}^\Psi_L := \left(G_{\lambda,\lambda'}\right)_{|\lambda|,|\lambda'|\leqslant L}$ satisfies $\|\mathbf{G}^\Psi - \mathbf{G}^\Psi_L\| \leqslant \varepsilon$ and hence

$$\|\mathcal{K}_0 - \mathcal{K}_{0,L}\|_{\mathcal{L}(\mathrm{S},\mathrm{S})} \leqslant \varepsilon, \tag{5.7}$$

which, in turn, controls the accuracy of $\mathcal{K}$ as pointed out in (5.3). $\mathbf{G}^\Psi_L$ is computed in a preprocessing step but could later be updated due to the hierarchical nature of Ψ.

Then for any larger tolerance $\eta \geqslant \varepsilon$, arising in the outer iteration, one can combine the compression rules from [11, Section 4] with the decay estimates in Proposition 5.1 such that the resulting compressed matrix $\mathbf{G}^\Psi_\eta$ satisfies

$$\|\mathbf{G}^\Psi - \mathbf{G}^\Psi_\eta\| \leqslant \eta \quad \Leftrightarrow \quad \|\mathcal{K}_0 - \mathcal{K}_{0,\eta}\|_{\mathcal{L}(\mathrm{S},\mathrm{S})} \leqslant \eta.$$

Roughly speaking, the computational cost for applying $\mathcal{K}_0$ to an element in $\mathbb{P}_M(\mathfrak{S})$ scales like $\#\mathfrak{S}\cdot(\log\#\mathrm{S})^a$ for some $a>0$. A first variant of $[\mathcal{K}_0,\cdot;\cdot]$ is then given by

$$[\mathcal{K}_0, v; \eta] = \mathcal{K}_{0,\eta'}v, \qquad \eta' := \eta/\|v\|_{L_2(\mathrm{S})},$$

where the compressed operator $\mathcal{K}_{0,\eta'}$ is applied *exactly*. In fact, since the approximations $\bar u$ use the same piecewise polynomial degrees as the kernel representations, orthonormality yields for $\bar u(x,\vec s) = \sum_{T\in\mathfrak{P}_h, i\in\mathcal{I}_T}\Big(\sum_{\lambda\in\Lambda_{\mathfrak{S}}} v_\lambda^{T,i}\psi_\lambda(\vec s)\Big)\varphi_{T,i}(x)$ the scattering

$$(\mathcal{K}_{0,\eta'}\bar u)(x,\vec s) = \sum_{T\in\mathfrak{P}_h, i\in\mathcal{I}_T} w_{T,i}(\vec s)\varphi_{T,i}(x) \tag{5.8}$$

with

$$w_{T,i}(\vec s) = \sum_{\lambda\in\Lambda_{\eta'}}\Big(\sum_{\lambda'\in\Lambda_{\mathfrak{S}}}(G^\Psi_\eta)_{\lambda,\lambda'}v^{T,i}_{\lambda'}\Big)\psi_\lambda(\vec s),$$

where Λ_η contains the range of indices of $\mathbf{G}^\Psi_\eta$. Thus, the $\vec s$-dependent coefficients $w_{T,i}$ are obtained by compressed matrix-vector multiplication.

In summary, the computational cost of the resulting routine $[\mathcal{K}, \bar{u}; \eta]$ can be reduced to $\mathcal{O}(\#\mathfrak{P} \cdot \#\mathfrak{S} \cdot (\log(\#\mathfrak{S}))^a)$, where of course $\#\mathfrak{P}$ and $\#\mathfrak{S}$ depend on η, typically in an algebraic fashion. For the Henyey–Greenstein kernel such schemes are still effective when the parameter γ gets close to one; see Figure 1.

5.5. **Hilbert–Schmidt expansion of G.** There is an alternative way of efficiently applying the scattering operator when the parameter γ in the Henyey–Greenstein kernel stays bounded away from one. It uses the fact that, by our assumptions, the kernel G possesses a *Hilbert–Schmidt* decomposition of the form

$$G(\vec{s}, \vec{s}') = \sum_{k=1}^{\infty} \sigma_k g_k(\vec{s}) g_k(\vec{s}'), \quad \sigma_k \geqslant 0, \quad \sum_{k \in \mathbb{N}} \sigma_k^2 = \|G\|^2_{L_2(\mathrm{S}\times\mathrm{S})} \leqslant 1,$$

where

$$(g_k, g_l)_{\mathrm{S}} = \delta_{k,l}, \quad k, l \in \mathbb{N}.$$

An approximate Hilbert–Schmidt decomposition of G results from the singular value decomposition (SVD) of the matrix $\mathbf{G}_L^{\Psi}$ from (5.7) which we denote for simplicity again as $\mathbf{G}^{\Psi}$.

The singular value decomposition then yields vectors $\mathbf{g}_k$ such that

$$\mathbf{G}^{\Psi} = \sum_{k=1}^{N_\tau} \sigma_k' \mathbf{g}_k \otimes \mathbf{g}_k, \tag{5.9}$$

where N_τ is the rank of $\mathbf{G}_L^{\Psi}$ and $\mathbf{g}_k$ is the vector of expansion coefficients of g_k with respect to Ψ, i.e.,

$$\sigma_k = \sigma_k', \qquad g_k = \sum_{\mu \in \nabla} g_{k,\mu} \theta_\mu =: \mathbf{g}_k^T \Psi, \qquad k \in \mathbb{N}.$$

We can then consider *low-rank approximations* by further truncating (5.9)

$$G^r := \sum_{k \leqslant r} \sigma_k g_k \otimes g_k, \qquad \mathbf{G}^r := \sum_{k \leqslant r} \sigma_k \mathbf{g}_k \mathbf{g}_k^T.$$

This yields

$$\|\mathcal{K}_0 - \mathcal{K}_0^r\|_{\mathcal{L}(L_2(\mathrm{S}), L_2(\mathrm{S}))} = \|\mathbf{G}^r - \mathbf{G}\|_{\mathcal{L}(\ell_2, \ell_2)} = \sigma_{r+1}.$$

The application of the truncated operator $\mathcal{K}_0^r$ for coarser accuracy tolerances, however, requires further reduction compressing the arrays $\mathbf{g}_k$. The coefficient vectors $\mathbf{g}_k$, consisting of wavelet coefficients, can easily be compressed by thresholding providing best n-term approximations of desired accuracy. In particular, notice that $\|\mathbf{g} - \tilde{\mathbf{g}}\|_{\ell_2} \leqslant \delta$ implies that

$$\|\mathbf{g}\mathbf{g}^T - \tilde{\mathbf{g}}\tilde{\mathbf{g}}^T\| = \|\mathbf{g}\mathbf{g}^T - \tilde{\mathbf{g}}\tilde{\mathbf{g}}^T\|_{\mathcal{L}(\ell_2, \ell_2)} \leqslant 2\delta.$$

Thus, thresholding for a given tolerance η the basis vectors $\mathbf{g}_k$ so as to obtain approximations $\mathbf{g}_{k,\eta}$ satisfying

$$\|\mathbf{g}_k - \mathbf{g}_{k,\eta}\|_{\ell_2} \leqslant \frac{\gamma_k \eta}{2\sigma_k},$$

with positive weights $\sum_k \gamma_k \leqslant 1$, one can verify that for the truncated kernel $\mathbf{G}_\eta^{\Psi,r} := \sum_{k=1}^r \sigma_k \mathbf{g}_{k,\eta} \mathbf{g}_{k,\eta}^T$ one has

$$\|\mathbf{G}^{\Psi,r} - \mathbf{G}_\eta^{\Psi,r}\| \leqslant \eta.$$

(Updating the SVD for $\mathbf{G}_\eta^{\Psi,r}$ would improve stability.) As a consequence one obtains for the corresponding operator approximation $\mathcal{K}_0^{r,\eta}$ and a given $v(\vec{s}) = \sum_{\lambda\in\Lambda_\mathfrak{S}} v_\lambda\psi_\lambda(\vec{s})$

$$\|(\mathcal{K}_0-\mathcal{K}_0^{r,\eta})v\|_{L_2(\mathrm{S})} \leqslant \Big\{\sigma_{r+1}\|v\|_{L_2(\mathrm{S})} + \eta\Big(\sum_{\lambda\in\Lambda_\mathfrak{S}} |v_\lambda|^2_{L_2(\mathrm{D})}\Big)^{1/2}\Big\} = (\sigma_{r+1}+\eta)\|v\|_{L_2(\mathrm{S})}.$$

Hence, choosing r such that $\sigma_{r+1} \leqslant \frac{\eta}{\|v\|_{L_2(\mathrm{S})}}$, $\eta' \leqslant \frac{\eta}{\|v\|_{L_2(\mathrm{S})}}$, with this variant we take

$$[\mathcal{K}_0, v; \eta] := \mathcal{K}_0^{r,\eta'} v.$$

This strategy is particularly efficient when the singular values σ_k decay rapidly. For the Henyey–Greenstein kernel, as illustrated in Figure 4, the larger $1-\gamma$, the more this is the case.

6. The routine $[\mathcal{T}^{-1}, F; \eta]$

The numerical realization of the routine $[\mathcal{T}^{-1}, \cdot; \cdot]$ is based on solving *fiber problems*

$$\mathcal{T}_{\vec{s}} u := \vec{s}\cdot\nabla u + \sigma(\vec{s})u = \int_{\mathrm{S}} K(\cdot,\vec{s},\vec{s}')v(\cdot,\vec{s}')\,\mathrm{d}\vec{s}' + f =: F(\vec{s}), \quad \vec{s}\in\mathrm{S}, \tag{6.1}$$

for properly selected parameters $\vec{s}\in\mathrm{S}$ where $F\in L_2(\mathrm{D}\times\mathrm{S})$ is given. Achieving a given target accuracy depends on solving each fiber problem with sufficient accuracy and also on solving sufficiently many of them.

The approximate solution of (6.1) will be based on the Discontinuous Petrov–Galerkin (DPG) scheme developed and analyzed in [7, 13] whose main features we briefly recall for the convenience of the reader in Sections 6.1 and 6.2. In Section 6.3, we explain how to use the set of solutions to the fiber problems in order to adaptively build an approximation to u in $L_2(\mathrm{D}\times\mathrm{S})$ which will be the output of $[\mathcal{T}^{-1}, F; \eta]$.

6.1. A DPG transport solver for the fiber problems.

We outline the numerical transport solver that is the core constituent of the current realization of $[\mathcal{T}^{-1}, F; \eta]$. We denote by $\mathfrak{P}_h$, $h>0$ a family of uniformly shaped regular partitions of the spatial domain D. More specifically, in what follows we always assume that all spatial partitions $\mathfrak{P}_h$ are (possibly local) refinements of a hierarchy of dyadic partitions of D. These partitions therefore induce dyadic partitions of the boundary $\partial\mathrm{D}$ as well.

While typically h stands for a mesh size parameter in a quasi-uniform mesh, here h is a locally varying mesh size function covering local refinements of the above dyadic hierarchy. With a given $\mathfrak{P}_h$ we associate the skeleton $\partial\mathfrak{P}_h$, which however depends strictly speaking on an associated convective direction $\vec{s}\in\mathrm{S}$. In fact, in analogy to (1.1), for a given $\vec{s}\in\mathrm{S}$ we define $\partial T_\pm(\vec{s})$ for any given cell $T\in\mathfrak{P}_h$ and set

$$\partial\mathfrak{P}_h = \partial\mathfrak{P}_h(\vec{s}) := \bigcup\{\partial T_-(\vec{s}), T_+(\vec{s}) \mid T\in\mathfrak{P}_h\},$$

suppressing at times the dependence of $\partial\mathfrak{P}_h$ on $\vec{s}$. Note that for polyhedral domains $\Gamma_-(\vec{s})$ remains the same on certain neighborhoods in S.

Following [7], the DPG scheme is based on the *infinite-dimensional* mesh-dependent variational formulation over the trial and test space

$$U_{\vec{s}} := L_2(\mathrm{D}) \times H_{0,\Gamma_-(\vec{s})}(\vec{s};\partial\mathfrak{P}_h), \quad V_{\vec{s}} := H(\vec{s};\mathfrak{P}_h) = \prod_{T\in\mathfrak{P}_h} H(\vec{s};T),$$

endowed with the norms

$$\|\theta\|_{H_{0,\Gamma_-(\vec{s})}(\vec{s};\partial\mathfrak{P}_h)} := \inf_{w\in H_{0,\Gamma_-(\vec{s})}(\vec{s};\mathrm{D}):\ w\big|_{\partial\mathfrak{P}_h}=\theta} \|w\|_{H(\vec{s};\mathrm{D})}, \quad \|v\|^2_{H(\vec{s};\mathfrak{P}_h)} := \sum_{T\in\mathfrak{P}_h} \|v\|^2_{H(\vec{s};T)},$$

where as before $\|v\|^2_{H(\vec{s};T)} = \|v\|^2_{L_2(T)} + \|\vec{s}\cdot\nabla v\|^2_{L_2(T)}$. Recall from [7] that the introduction of the additional unknown field $\theta \in H_{0,\Gamma_-(\vec{s})}(\vec{s};\partial\mathfrak{P}_h)$, living on the skeleton $\partial\mathfrak{P}_h$, is necessary because the trace terms encountered in the usual derivation of DG bilinear forms may not exist for general elements in $L_2(\mathrm{D})$.

Remark 6.1. The spaces $U_{\vec{s}}$, $V_{\vec{s}}$ depend on the directions $\vec{s}$ and on $\mathfrak{P}_h$, and so will the solution $[u(\vec{s}),\theta(\vec{s})]$. However, when the solution component $u(\vec{s})$ is regular enough, i.e., $u(\vec{s}) \in H_{0,\Gamma_-(\vec{s})}(\vec{s};\mathrm{D})$, one can show that $u(\vec{s})$ is the solution of (6.1) and $\theta(\vec{s})$ is its trace on $\partial\mathfrak{P}_h$.

Defining

$$b_h(u,\theta,v;\vec{s}) = \sum_{T\in\mathfrak{P}_h} \underbrace{\int_T (\sigma(\vec{s})v - \vec{s}\cdot\nabla v)u\,\mathrm{d}x + \int_{\partial T} \mathbf{n}\cdot\vec{s}\theta v\,\mathrm{d}\Gamma}_{=:b_T(u,\theta,v;\vec{s})}, \tag{6.2}$$

and given $F(\vec{s}) \in L_2(\mathrm{D})$, we then wish to find $u(\vec{s}) \in L_2(\mathrm{D})$, $\theta \in H_{0,\Gamma_-(\vec{s})}(\vec{s};\partial\mathfrak{P}_h)$ such that

$$b_h(u(\vec{s}),\theta(\vec{s}),v;\vec{s}) = \int_{\mathrm{D}} F(\vec{s})v\,\mathrm{d}x, \quad v \in V_{\vec{s}} = H(\vec{s};\mathfrak{P}_h). \tag{6.3}$$

Remark 6.2. It immediately follows from [7, Theorem 3.1] that (6.2) is a uniformly stable variational formulation for the transport equation $\mathcal{T}_{\vec{s}}u_{\vec{s}} = F(\vec{s})$, i.e., continuity and inf-sup conditions according to Theorem 2.1 hold uniformly in $\vec{s} \in \mathrm{S}$ and in $\mathfrak{P}_h$.

A fully discrete scheme. The discretization of (6.3) requires two hierarchies of partitions $\mathfrak{P}_{\underline{h}}$, $\mathfrak{P}_h$ where the $\mathfrak{P}_h$ is a refinement of (locally) constant depth of $\mathfrak{P}_{\underline{h}}$, i.e., $\mathfrak{P}_{\underline{h}} \prec \mathfrak{P}_h$. (In fact, practical experiments usually indicate that depth-0 suffices, i.e., $h = \underline{h}$.) In that sense we can write $\underline{h} = \underline{h}(h)$ and $h = h(\underline{h})$. Given $\mathfrak{P}_{\underline{h}}$, $\mathfrak{P}_h$, we fix a polynomial degree $m \in \mathbb{N}$ and consider the finite-dimensional trial spaces

$$U^h_{\vec{s}} := \Big(\prod_{\underline{T}\in\mathfrak{P}_{\underline{h}}} \mathbb{P}_{m-1}(\underline{T})\Big) \times \Big(H_{0,\Gamma_-(\vec{s})}(\vec{s};\mathrm{D}) \cap \prod_{\underline{T}\in\mathfrak{P}_{\underline{h}}} \mathbb{P}_m(\underline{T})\Big)\Big|_{\partial\mathfrak{P}_h}.$$

Note that the second component consists of traces of globally continuous piecewise polynomials of one degree higher than for the discontinuous bulk-component but evaluated on the skeleton of the (possibly) *finer* mesh D_h.

Given the finite-dimensional trial space $U^h_{\vec{s}}$, it is critical to construct a suitable *test space* that renders also the finite-dimensional corresponding Petrov–Galerkin problem inf-sup stable, ideally with inf-sup constants independent of the trial and test space dimensions. We follow again [7] and fix the so-called *test search space* as

discontinuous piecewise polynomials of one degree higher on a *subgrid* $\mathfrak{P}_h$ of $\mathfrak{P}_{\underline{h}}$, namely

$$\hat{V}_{\vec{s}}^h := \prod_{T\in\mathfrak{P}_h} \mathbb{P}_{m+1}(T).$$

The actual *test space* $V_{\vec{s}}^h$ is then defined as the following $H(\vec{s};\mathfrak{P}_h)$-projection to the *test search* space $\hat{V}_{\vec{s}}^h$

$$V_{\vec{s}}^h := \big\{\breve{t}(u,\theta)\in\hat{V}_{\vec{s}}^h \mid (\breve{t}(u,\theta),v)_{V_{\vec{s}}} = b_h(u,\theta,v;\vec{s}),\ v\in\hat{V}_{\vec{s}}^h,\ [u,\theta]\in U_{\vec{s}}^h\big\}. \tag{6.4}$$

Since the local test search spaces over each cell $T\in\mathfrak{P}_{\underline{h}}$ have uniformly bounded finite dimension the overall computational work still remains proportional to the dimension of the trial spaces.

This gives rise to the *Petrov–Galerkin* formulation: find $[u_{\underline{h}}(\vec{s}),\theta_{\underline{h}}(\vec{s})]\in U_{\vec{s}}^{\underline{h}}$ such that

$$b_h(u_{\underline{h}}(\vec{s}),\theta_{\underline{h}}(\vec{s}),v_h;\vec{s}) = \int_{\mathrm{D}} F(\vec{s})v\,\mathrm{d}x =: F(\vec{s})(v),\quad v\in V_{\vec{s}}^h, \tag{6.5}$$

for $V_{\vec{s}}^h$ defined by (6.4). Here and below we sometimes use the shorthand notation $u_{\underline{h}} = u_{\mathfrak{P}_{\underline{h}}}, b_h = b_{\mathfrak{P}_h}, U^{\underline{h}} = U^{\mathfrak{P}_{\underline{h}}}$.

Before stating the corresponding stability result, we mention a variant where the skeleton component $\theta_{\underline{h}}(\vec{s})$ is replaced by the globally conforming piecewise polynomial $w_{\underline{h}}$ in $H_{0,\Gamma_-(\vec{s})}(\vec{s};\mathrm{D})\cap\prod_{T\in\mathfrak{P}_{\underline{h}}}\mathbb{P}_m(\underline{T})\big|_{\partial\mathfrak{P}_h}$. Then the local bilinear forms $b_T(u_{\underline{h}}(\vec{s}),\theta_{\underline{h}}(\vec{s}),v_h;\vec{s})$ from (6.2) can be rewritten as

$$\begin{aligned} &b_T(u_{\underline{h}},\theta_{\underline{h}},v_h;\vec{s})\\ &\quad= b_T(u_{\underline{h}},w_{\underline{h}},v_h;\vec{s}) = \int_T(\sigma(\vec{s})v_h-\vec{s}\cdot\nabla v_h)u_{\underline{h}}\,\mathrm{d}x + \int_{\partial T}\mathbf{n}\cdot\vec{s}w_{\underline{h}}v_h\,\mathrm{d}\Gamma\\ &\quad= \int_T \sigma(\vec{s})v_h(u_{\underline{h}}-w_{\underline{h}}) + \partial_{\vec{s}}v_h(w_{\underline{h}}-u_{\underline{h}}) + (\sigma w_{\underline{h}}+\partial_{\vec{s}}w_{\underline{h}})v_h\,\mathrm{d}x,\quad T\in\mathfrak{P}_h. \end{aligned}$$

Using $[u_{\underline{h}},w_{\underline{h}}]$ as unknowns one obviously has $\|w\|_{H_{0,\Gamma_-(\vec{s})}(\vec{s};\partial\mathfrak{P}_h)}\leqslant\|w\|_{H(\vec{s};\mathrm{D})}$. We will adopt this variant in what follows where it is now understood to use the norm

$$\|[u_{\underline{h}},w_{\underline{h}}]\|_{U_{\vec{s}}}^2 := \|u_{\underline{h}}\|_{L_2(\mathrm{D})}^2 + \|w_{\underline{h}}\|_{H(\vec{s};\mathrm{D})}^2.$$

The following facts are immediate consequences of the results in [7,13].

Theorem 6.3. *For a fixed but sufficiently large subgrid-depth $\underline{h}/h$, (depending on the shape parameters of the involved partitions) the scheme (6.5) is uniformly in $h\geqslant 0$, $\vec{s}\in\mathrm{S}$,* inf-sup *stable, i.e.,*

$$\inf_{[u_{\underline{h}},w_{\underline{h}}]\in U_{\vec{s}}^h}\sup_{v_h\in V_{\vec{s}}^h}\frac{b_h(u_{\underline{h}},w_{\underline{h}},v_h;\vec{s})}{\|[u_{\underline{h}},w_{\underline{h}}]\|_{U_{\vec{s}}}\|v_h\|_{V_{\vec{s}}}}\geqslant\bar{\beta}>0,\quad h\geqslant 0,\vec{s}\in\mathrm{S}, \tag{6.6}$$

where $\bar{\beta}$ depends on the shape parameters of the underlying partitions, on $\|\mathcal{T}_{\vec{s}}^{-1}\|_{\mathcal{L}(L_2(\mathrm{D}),H_{0,\Gamma_-(\vec{s})}(\vec{s};\mathrm{D}))}$, and on $\|\sigma\|_{L_\infty(\mathrm{S},W^1(L_\infty(\mathrm{D})))}$.

It is well known that the system matrices arising in (6.5) are always *symmetric positive definite* despite the asymmetric nature of transport equations.

While the conforming formulation (F1) does not require incorporating boundary conditions on Γ_- into the trial space, the skeleton component requires an adjustment in the DPG formulation. To that end, following [7, Remark 3.6], let

$w_0(\vec{s}) \in H(\vec{s};\mathrm{D})$ satisfy $w_0(\vec{s}) = g(\vec{s})$ on $\Gamma_-(\vec{s})$. Then, the (infinite-dimensional) DPG formulation of the problem $\mathcal{T}_{\vec{s}}\bar{u} = f - \mathcal{T}_{\vec{s}}w_0$, in D, $\bar{u} = 0$ in $\Gamma_-(\vec{s})$, is given by (6.7)

$$b_h(\bar{u}(\vec{s}), \bar{w}(\vec{s}), v; \vec{s}) = \langle f, v\rangle - b_h(\bar{w}_0(\vec{s}), \bar{w}_0(\vec{s}), v; \vec{s}) =: \langle f - F_b(w_0, \vec{s}), v\rangle, \quad v \in V.$$

Now one has $\bar{w}|_{\partial\mathfrak{P}_h} = \bar{u}|_{\partial\mathfrak{P}_h} = (u - w_0)|_{\partial\mathfrak{P}_h}$, i.e., it suffices to discretize (6.7).

6.2. **A posteriori error estimates.** As an immediate consequence of the fact that the DPG-induced transport operators $\mathcal{T}_{\vec{s},h}$ are norm isomorphisms, uniformly in $h \geqslant 0, \vec{s} \in \mathrm{S}$, errors in $\|\cdot\|_{U_{\vec{s}}}$ are equivalent to residuals in $\|\cdot\|_{V'_{\vec{s}}}$, i.e.,

$$\big\|[u(\vec{s}), u(\vec{s})] - [u_{\underline{h}}(\vec{s}), w_{\underline{h}}(\vec{s})]\big\|_{U_{\vec{s}}} \sim \big\|F(\vec{s}) - \mathcal{T}_{\vec{s},h}([u_{\underline{h}}(\vec{s}), w_{\underline{h}}(\vec{s})])\big\|_{V'_{\vec{s}}}, \quad h \geqslant 0,\ \vec{s} \in \mathrm{S},$$

holds with uniform constants. Thus, as soon as one can tightly estimate the dual norm $\|F(\vec{s}) - \mathcal{T}_{\vec{s},h}([u_{\underline{h}}(\vec{s}), w_{\underline{h}}(\vec{s})])\|_{V'_{\vec{s}}}$ of the residual, one obtains efficient and reliable a posteriori error bounds. Such tight bounds are established in [13] which we briefly recall. Define for $T \in \mathfrak{P}_h$ the Riesz lifts $\breve{R}_T(u_{\underline{h}}, w_{\underline{h}}, \bar{F}(\vec{s}))$ of the local residuals by

$$\big(\breve{R}_T(u_{\underline{h}}, w_{\underline{h}}, \bar{F}(\vec{s})), v_h\big)_{H(\vec{s};T)} = b_T(u_{\underline{h}}, w_{\underline{h}}, v_h; \vec{s}) - \bar{F}(\vec{s})(v_h), \quad v_h \in \hat{V}^h_{\vec{s}},$$

where $\bar{F}(\vec{s})|_T \in \mathbb{P}_m$ is a piecewise polynomial approximation to $F(\vec{s})$ and where $\hat{V}^h_{\vec{s}}$ is the same test search space as used before for the Petrov–Galerkin scheme. Thus, the computational cost per cell T is again uniformly bounded. Defining then

$$\begin{aligned}&\|\breve{R}_{\underline{h}}(u_{\underline{h}}, w_{\underline{h}}, \bar{F}(\vec{s}))\|^2_{H(\vec{s};\mathfrak{P}_{\underline{h}})}\\ &\quad = \|\breve{R}_{\mathfrak{P}_{\underline{h}}}(u_{\underline{h}}, w_{\underline{h}}, \bar{F}(\vec{s}))\|^2_{H(\vec{s};\mathfrak{P}_{\underline{h}})} := \sum_{T\in\mathfrak{P}_{\underline{h}}} \|\breve{R}_T(u_h, w_h, \bar{F}(\vec{s}))\|^2_{H(\vec{s};T)},\end{aligned}$$

the following holds; see [13, Theorem 4.1 and (4.4)].

Theorem 6.4. *If the operators $\mathcal{T}_{\vec{s},h}$ are norm isomorphisms uniformly in $h \geqslant 0$ and $\vec{s} \in \mathrm{S}$, then for a fixed maximal subgrid depth there exist constants $\underline{c}$, $\bar{C}$, depending on $\bar{\beta}$ from (6.6), but independent of $\vec{s}$, $\mathfrak{P}_h$, such that*

$$\begin{aligned}\underline{c}\|\breve{R}_h(u_{\underline{h}}, w_{\underline{h}}, \bar{F}(\vec{s}))\|_{H(\vec{s};\mathfrak{P}_h)} &\leqslant \|[u(\vec{s}), u(\vec{s})|_{\partial\mathfrak{P}_{\underline{h}}}] - [u_{\underline{h}}(\vec{s}), w_{\underline{h}}(\vec{s})]\|_{U_{\vec{s}}}\\ &\leqslant \bar{C}\|\breve{R}_h(u_{\underline{h}}, w_{\underline{h}}, \bar{F}(\vec{s}))\|_{H(\vec{s};\mathfrak{P}_h)}.\end{aligned} \tag{6.8}$$

In the present context it is particularly important to control the dependence of a posteriori bounds on the direction parameter $\vec{s} \in \mathrm{S}$. In this regard, the following further result from [13, Proposition 4.4] is relevant: there exists a constant $c_0 > 0$ such that the Petrov–Galerkin solution satisfies for each $T' \in \mathfrak{P}_{\underline{h}}$

$$\begin{aligned}&c_0\Big(\|u_{\underline{h}}(\vec{s}) - w_{\underline{h}}(\vec{s})\|^2_{L_2(T')} + \|\vec{s}\cdot\nabla w_{\underline{h}}(\vec{s}) + \sigma u_{\underline{h}}(\vec{s}) - \bar{F}(\vec{s})\|^2_{L_2(T')}\Big)\\ &\qquad\leqslant \sum_{T\in\mathfrak{P}_h, T\subset T'} \|\breve{R}_T(u_{\underline{h}}, w_{\underline{h}}, \bar{F}(\vec{s}))\|^2_{H(\vec{s};T)}\\ &\qquad\leqslant \|u_{\underline{h}}(\vec{s}) - w_{\underline{h}}(\vec{s})\|^2_{L_2(T')} + \|\vec{s}\cdot\nabla w_{\underline{h}}(\vec{s}) + \sigma u_{\underline{h}}(\vec{s}) - \bar{F}(\vec{s})\|^2_{L_2(T')}.\end{aligned} \tag{6.9}$$

For $d = 2$, i.e., S is the circle we can identify $\vec{s} = (\cos t, \sin t)^\top$ and the space $\mathbb{P}_M(\mathfrak{S})$ consists for a given admissible partition $\mathfrak{S}$ of S of 2π-periodic piecewise polynomials in $t \in (-\pi, \pi]$. Hence, the above error indicators are nearly piecewise polynomial in t when the components $u_{\underline{h}}$, $w_{\underline{h}}$ are of the form (5.2) with $\vec{s}$-dependent coefficients in $\mathbb{P}_M(\mathfrak{S})$; see Section 5.2.

The above DPG scheme and the associated a posteriori error bounds form the core constituent of the routine $[\mathcal{T}^{-1},\cdot;\cdot]$. We can use (6.8) to contrive adaptive mesh refinement strategies based on so-called *Dörfler marking* or *bulk chasing*. This means one marks those cells for subsequent refinement whose combined energy exceeds a fixed portion of the total lifted residual. It is shown in [13] that this entails a fixed error reduction for each refinement sweep and associated complexity estimates.

Remark 6.5. Convergence to zero of either one of the above residual error bounds guarantees convergence of errors in the spaces $U^{\vec{s}}$. The DPG output has two components, namely a piecewise polynomial u_h of degree m on the underlying mesh $\mathfrak{P}_h$ as well as a skeleton component which can be identified with the trace of a conforming piecewise polynomial of degree $m+1$. Therefore the a posteriori error bounds control in particular the convergence of the u-component in $L_2(\mathrm{D})$. For the realization of $[\mathcal{T}^{-1},F;\eta]$ below we always use only the u-component for the outer iteration.

6.3. **An adaptive solver in $U = L_2(\mathbf{D}\times\mathbf{S})$.** We describe next how $[\mathcal{T}^{-1},\cdot;\eta]$ is realized based on approximately solving, with the aid of the DPG scheme described above, fiber problems $\mathcal{T}_{\vec{s}}\bar{u} = F$ for the elements $\vec{s}$ from a stage-dependent discrete subset $\mathcal{Q}_\eta$ of the parameter domain S. Both $\mathcal{Q}_\eta$ as well as the meshes for each fiber solution are generated adaptively.

The data: The data $F = F(x,\vec{s})$ required by each call of $[\mathcal{T}^{-1},F;\eta]$ have a piecewise polynomial representation of the type (5.8). Specifically, they are of the form

$$F = w + g \in L_2(\mathrm{D}\times\mathrm{S}),$$

where w is the output of the routine $[\mathcal{K},\cdot;\cdot]$ and g is a stage-dependent approximation to the source term. More precisely, in the case of inhomogeneous boundary conditions g consists of two parts, namely $g = g_0 + g_1$ where g_0 stands for the "lifted boundary data" needed to correct the right hand side so as to reduce the problem to the homogeneous case; see (6.7). Both w and g need to be computed within the currently given accuracy tolerance. We omit the details concerning the computation of g.

Output format: The output of $[\mathcal{T}^{-1},\cdot;\eta]$ is a piecewise polynomial of degree m of the form (see (5.2))

$$\bar{u}_\eta(x,\vec{s}) = \sum_{T\in\mathfrak{P}_\eta}\sum_{i\in\mathcal{I}_T} v_{T,i}(\vec{s})\varphi_{T,i}(x),$$

where the $\varphi_{T,i}$ are polynomial basis functions of degree m supported in $T\in\mathfrak{P}_\eta$ and $\mathfrak{P}_\eta$ is a partition of the spatial domain D. The parameter dependent coefficients $v_{T,i}(\vec{s})$ are elements of a space $\mathbb{P}_M(\mathfrak{S}_\eta)$ of piecewise polynomials of degree M subordinate to a partition $\mathfrak{S}_\eta$ of S. We describe next how to compute the $v_{T,i}(\vec{s})$ as well as the partition $\mathfrak{P}_\eta$.

Computation of fiber solutions: The realization of $[\mathcal{T}^{-1},F;\eta]$ is based on approximately solving fiber transport problems $\mathcal{T}_{\vec{s}}u_{\vec{s}} = F(\cdot,\vec{s})$ for parameters $\vec{s}$ in a

suitable finite subset of S, Specifically, given a partition $\mathfrak{S}$ of the parameter domain S, we associate with each cell $C \in \mathfrak{S}$ a set of "quadrature points" $\mathcal{Q}_C$ whose union

$$\mathcal{Q}_{\mathfrak{S}} := \bigcup_{C\in\mathfrak{S}} \mathcal{Q}_C$$

is the discrete set of parameters for which we first compute error-controlled approximate fiber solutions. Before describing this in more detail, a few preparatory comments are in order. The realization of $[\mathcal{K}, \cdot; \cdot]$ is reduced to a frequent but efficient approximate application of a global operator acting in functions in $d-1$ variables. The bulk of computation therefore lies in $\#\mathcal{Q}_{\mathfrak{S}}$ approximate *inversions* of transport boundary value problems in d variables. It is therefore of primary importance to keep the size of each fiber transport problem as small as possible. In view of the inherently low regularity of the transport solutions (especially in the presence of rough boundary and source data) we opt for employing an adaptive DPG scheme for each fiber problem. The price to be paid is that then each fiber solution $\bar{u}_{\mathfrak{S}}(\cdot, \vec{s})$, $\vec{s} \in \mathcal{Q}_{\mathfrak{S}}$, comes with its own adaptive partition $\mathfrak{P}_{\vec{s}}$; see Figure 2. We refer to [7,13] for the details on an adaptive fiber transport solver

$$[\mathcal{T}_{\vec{s}}^{-1}, F; \eta] \to (\mathfrak{P}_{\vec{s}}, \bar{u}_{\vec{s}}), \qquad \bar{u}_{\vec{s}}(x) = \sum_{T\in\mathfrak{P}_{\vec{s}}, i\in\mathcal{I}_T} c_{T,i,\vec{s}}\varphi_{T,i}(x).$$

It consists in repeating the standard cycle

$$\textsc{Mark} \quad \to \quad \textsc{Refine} \quad \to \quad \textsc{Solve}$$

until the sum of squared indicators (in either (6.8) or (6.9)) is below the current threshold η^2. Here one needs for each $C \in \mathfrak{S}$ a good initial guess. If $C \in \mathfrak{S}$ was already obtained in the representation of the final DPG solution of the previous outer iteration we choose this one. Otherwise one can take the union of those fiber meshes associated with those parameter cells from the preceding outer iteration that intersect the current parameter cell.

For MARK we use a simple bulk criterion identifying for each selected quadrature point $\vec{s}$ a possibly small set of cells in the current partition such that the sum of the corresponding squared indicators exceeds a fixed portion of the full sum of squared indicators. Hence, the adaptively generated meshes depend on the directions $\vec{s}$. However, the approximate application of the scattering kernel in $[\mathcal{K}, \cdot; \cdot]$ requires an *aggregated* approximate solution $\bar{u}(x, \vec{s})$ as a function of the spatial and parametric variables which needs to be represented on a single mesh that is obtained by *merging* the parameter-dependent fiber meshes. Note that even the merged mesh involves a total number of degrees of freedom which is significantly smaller than that corresponding to a uniform mesh with the highest required resolution; see the rightmost picture in Figure 2.

A more detailed algorithmic description is beyond the present scope and can be found in [18, Section 6.3.2].

Aggregating fiber solutions: We discuss first how to generate an approximate solution $\bar{u}_{\mathfrak{S}} \in L_2(\mathrm{D} \times \mathrm{S})$ which is only based on approximate fiber solutions for $\vec{s} \in \mathcal{Q}_{\mathfrak{S}}$ where at this point $\mathfrak{S}$ is a *given* partition of $\mathfrak{S}$, e.g., generated by an error-controlled approximate application of $\mathcal{K}$. This can be formulated as a (preparatory) routine

$$[\mathcal{T}^{-1}, F, \mathfrak{S}; \eta] \to (\mathfrak{P}_{\mathfrak{S}}, \bar{u}_{\mathfrak{S}}), \qquad \bar{u}_{\mathfrak{S}}(x, \vec{s}) = \sum_{T\in\mathfrak{P}_{\mathfrak{S}}, i\in\mathcal{I}_T} v_{T,i}(\vec{s})\varphi_{T,i}(x), \tag{6.10}$$

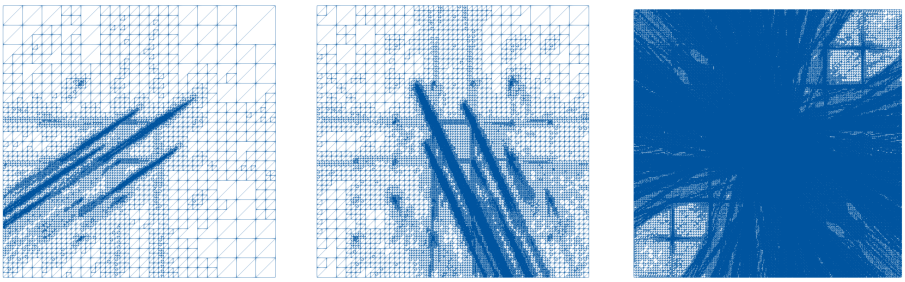


FIGURE 2. Adaptive meshes for fiber transport solutions with respect to two different directions as well as the merged mesh at iteration step 10.

that outputs a mesh $\mathfrak{P}_{\mathfrak{S}}$ and a piecewise polynomial $u_{\mathfrak{S}}(x,\vec{s})$ in x subordinate to $\mathfrak{P}_{\mathfrak{S}}$ with parameter-dependent coefficients $v_{T,i} \in \mathbb{P}_M(\mathfrak{S})$ and a spatial mesh $\mathfrak{P}_{\mathfrak{S}}$ such that

$$\|\breve{R}_{\mathrm{D}_{\mathfrak{S}}}(u_{\mathfrak{S}},\theta_{\mathfrak{S}},F(\vec{s}))\|_{H(\vec{s};\mathfrak{P}_h)} \leqslant \kappa_{\mathcal{T}}\eta, \quad \vec{s}\in\mathcal{Q}_{\mathfrak{S}}.$$

The workhorse called by $[\mathcal{T}^{-1},F,\mathfrak{S};\eta]$ is therefore the following subroutine providing a parameter-dependent approximate transport solution over a given cell C in the current parameter partition $\mathfrak{S}$:

$[C,F;\eta]\to(\mathfrak{P}_C,\bar{u}_C)$

C1: For $\vec{s}\in\mathcal{Q}_C$ invoke $[\mathcal{T}_{\vec{s}}^{-1},F;\eta]$;

C2: generate the mesh $\mathfrak{P}_C$ by merging the meshes $\mathfrak{P}_{\vec{s}}$, $\vec{s}\in\mathcal{Q}_C$ to obtain merged representations $\bar{u}_{\vec{s}}(x)=\sum_{T\in\mathfrak{P}_C,i\in\mathcal{I}_T}\tilde{c}_{T,i,\vec{s}}\varphi_{T,i}(x)$;

C3: Determine the polynomial $v_{C,T,i}(\vec{s})\in\mathbb{P}_M(C)$ that (quasi-)interpolates the values $\tilde{c}_{T,i,\vec{s}}$, $\vec{s}\in\mathcal{Q}_C$ and aggregate

$$\bar{u}_C(x,\vec{s}) := \sum_{T\in\mathfrak{P}_C}\tilde{v}_{C,T,i}(\vec{s})\varphi_{T,i}(x).$$

The output in (6.10) of $[\mathcal{T}^{-1},F,\mathfrak{S};\eta]$ is then given by

$$\bar{u}_{\mathfrak{S}}(x,\vec{s})=\sum_{C\in\mathfrak{S}}\bar{u}_{C,\eta}(x,\vec{s})=\sum_{T\in\mathfrak{P}_{\mathfrak{S}},i\in\mathcal{I}_T}v_{T,i}(\vec{s})\varphi_{T,i}(x),$$

where $\mathrm{D}_{\mathfrak{S}}$ is obtained by merging the cell-dependent meshes $\mathfrak{P}_C$, $C\in\mathfrak{S}$ produced by $[C,F;\eta]$.

Finding $\mathfrak{S}_\eta$: The accuracy requirement in $[\mathcal{T}^{-1},F;\eta]$ requires a mean square control over the parameter domain S. The output of the routine $[\mathcal{T}^{-1},F,\mathfrak{S};\eta]$ for a given parameter partition $\mathfrak{S}$ guarantees that the residual bounds satisfy the required accuracy η only at the quadrature points $\mathcal{Q}_{\mathfrak{S}}$ but a priori not necessarily for all parameter values in S. Our current approach is therefore to adaptively generate also a further refinement $\mathfrak{S}_\eta$ (if necessary) of some initial partition of S (dictated solely by the accuracy in the application of $\mathcal{K}$). We then apply quadrature with respect to $\mathcal{Q}_{\mathfrak{S}_\eta}$ to estimate the error in $L_2(\mathrm{D}\times\mathrm{S})$. Here we use that by (6.9), the true errors are rigorously sandwiched by error indicators that are piecewise defined as products of polynomials and trigonometric functions. Specifically, we apply the following steps:

S1: Take the partition $\mathfrak{S}=\mathfrak{S}_{\mathcal{K},\kappa_{\mathcal{K}}\eta}$ generated by $[\mathcal{K},\bar{u};\kappa_{\mathcal{K}}\eta]$ as initial guess.

S2: Given a partition $\mathfrak{S}$ of S compute $\bar{u}_{\mathfrak{S}} = [\mathcal{T}^{-1}, F, \mathfrak{S}; \eta]$.
S3: Subdivide each cell in $\mathfrak{S}$ to obtain a refined partition $\mathfrak{S}_r$.
S4: Evaluate the residual bounds (e.g., (6.9)) for the current approximation $\bar{u}_{\mathfrak{S}}(\cdot, \vec{s})$ at the new quadrature points $\vec{s} \in \mathcal{Q}_{\mathfrak{S}_r} \backslash \mathcal{Q}_{\mathfrak{S}}$ and mark all cells $C \in \mathfrak{S}_r$ containing a quadrature point for which a fixed threshold $\omega\eta$ ($\omega \leqslant 1$ fixed) is exceeded. If no cell is marked stop and set $\mathfrak{S} \to \mathfrak{S}_\eta$.
S5: The parents in $\mathfrak{S}$ of the marked cells are refined to generate a refined partition $\mathfrak{S}_{\text{new}}$ of $\mathfrak{S}$.
S6: Replace $\mathfrak{S}$ by $\mathfrak{S}_{\text{new}}$ and go to S2.

7. Numerical experiments

We consider the radiative transfer problem (1.2) on the unit square domain $\mathrm{D} = [0,1]^2$ with homogeneous boundary conditions. The structure of the source term f and absorption coefficient σ is illustrated by Figure 3. More precisely, we take $f = 0$ in the white and gray areas whereas $f = 1$ in the black area. Similarly, we set $\sigma = 10$ in the gray areas and $\sigma = 2$ everywhere else. Such checkerboard structure serves as a classical benchmark in the literature of radiative transfer and can be found in other works; see, e.g., [8].

The scattering is of Henyey–Greenstein-type (see formula (5.1))

$$K(x, \vec{s}, \vec{s}') = G(\vec{s}, \vec{s}') = \frac{1}{2\pi} \frac{1-\gamma^2}{1+\gamma^2-2\gamma\vec{s}\cdot\vec{s}'} \quad \forall x \in \mathrm{D}.$$

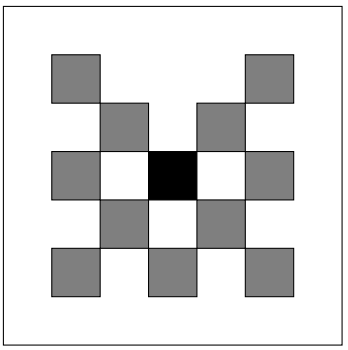


Figure 3. Geometry of the checkerboard benchmark.

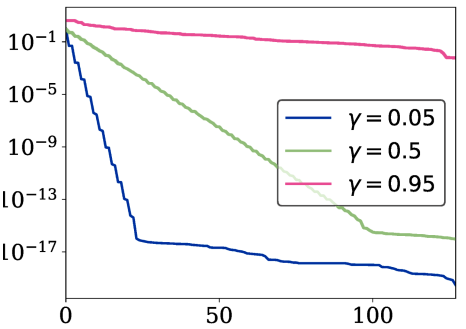


Figure 4. SVD of the matrix representation $\mathbf{G}^{\Xi}$, $\Xi \in \{\Psi, \Phi\}$ of G for different values of γ.

Figure 4 shows the decay of singular values of a highly accurate matrix representation $\mathbf{G}^{\Xi}$, $\Xi \in \{\Psi, \Phi\}$, of the scattering kernel G for different values of γ. For γ close to one this decay is very slow but Figure 1 reveals that the wavelet representation is nevertheless extremely sparse. Here we confine the subsequent discussion to moderately isotropic scattering $\gamma = 0.5$. The singular values still decay rapidly (see Figure 4) which allows us to apply the method outlined in Section 5.5 based on Hilbert–Schmidt decompositions. We present results with Alpert wavelets of degree 2.

We set $\varepsilon = 1.1 \cdot 5.10^{-3}$ as the final target accuracy. The problem is of transport-dominated nature ($\rho \leqslant 1$) so we can solve it with the ASTI algorithm. Table 1 gives the estimated values $C_{\mathcal{T}}, \rho, b_0(u)$ and $\kappa_1, \kappa_2, \kappa_3$. Note that $\kappa_2 = 0$ since we can evaluate the source term exactly. The remaining two parameters κ_1 and κ_3 balance the accuracy tolerances for the approximate application of the scattering operator and the approximate inversion of $\mathcal{T}$. Specifically, κ_1 determines on the one hand the number of quadrature points and hence the number of fiber transport problems to be solved and, on the other hand, κ_3 affects the spatial discretizations of these fiber problems.

TABLE 1. Values of the constants required to run the ASTI Algorithm 1.

$C_{\mathcal{T}}$	ρ	$b_0(u)$	κ_1	κ_2	κ_3
0.594604	0.594604	1/7	$0.2/C_{\mathcal{T}}$	0	0.8

Figure 5 displays the convergence history and degrees of freedom for the above choice of parameters. The left plot gives an approximation error of the scattering application $||\mathcal{K}(\bar{u}_n) - [\mathcal{K}, \bar{u}_n; \kappa_1\eta_n]||_{L_2(\mathrm{D}\times\mathrm{S})}$ (dark blue curve), the a posteriori error of the transport solves $||u_n - \bar{u}_n||_{L_2(\mathrm{D}\times\mathrm{S})}$ (light blue curve), and a bound for the global error $||u - \bar{u}_n||_{L_2(\mathrm{D}\times\mathrm{S})}$ (purple curve) based on (4.13). Recall that it is composed of the bounds for $\rho^n\|u\|_U$ and the the above two error tolerances. By the definition (4.5) of the tolerances η_n, the interior solution accuracies need to be somewhat finer which explains the gradual divergence between the global error bound and the interior error tolerances. To avoid this would require total a posteriori bounds based on the bilinear form $b(w, v) = ((\mathcal{T} - \mathcal{K})(w))(v)$ in combination with coarsening strategies, which is the subject of future work. The shaded blue regions in the right plot indicate statistics about the number of degrees of freedom that are associated for each selected angular direction.

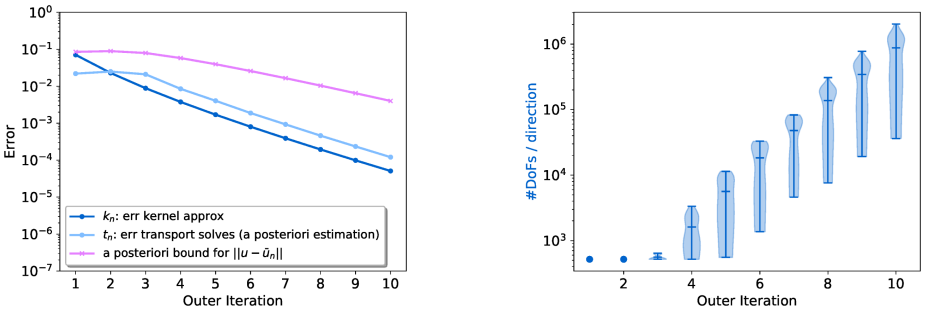
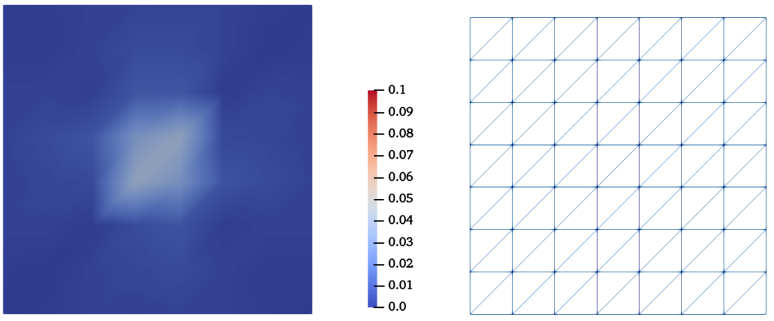
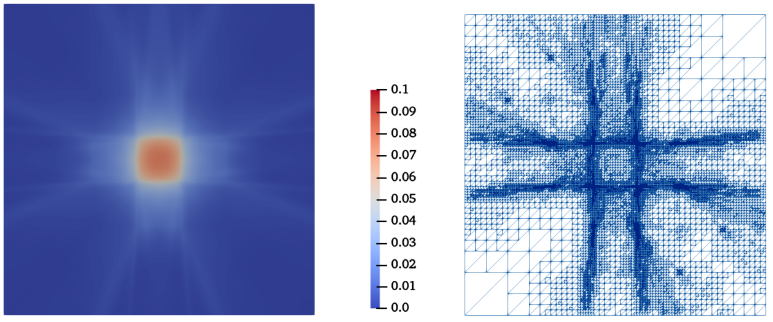


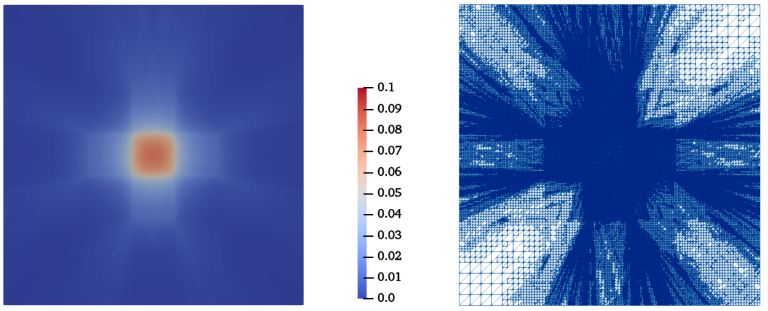
FIGURE 5. Convergence and number of DoFs for $\kappa_1 = \xi/C_{\mathcal{T}}$, $\kappa_2 = 0$, $\kappa_3 = (1 - \xi)/2$ with $\xi = 0.2$.



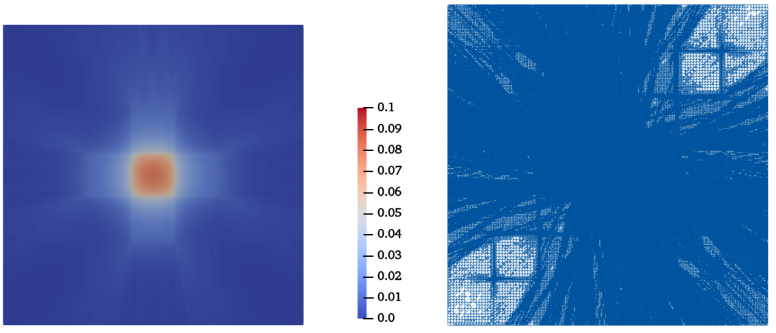
integrated solution and grid for iteration step 2.



integrated solution and grid for iteration step 6.



integrated solution and grid for iteration step 8.



integrated solution and grid for iteration step 10.

FIGURE 6. Integrated solutions $\int_S u_n(\cdot, \vec{s})\, d\vec{s}$ and corresponding merged grids.

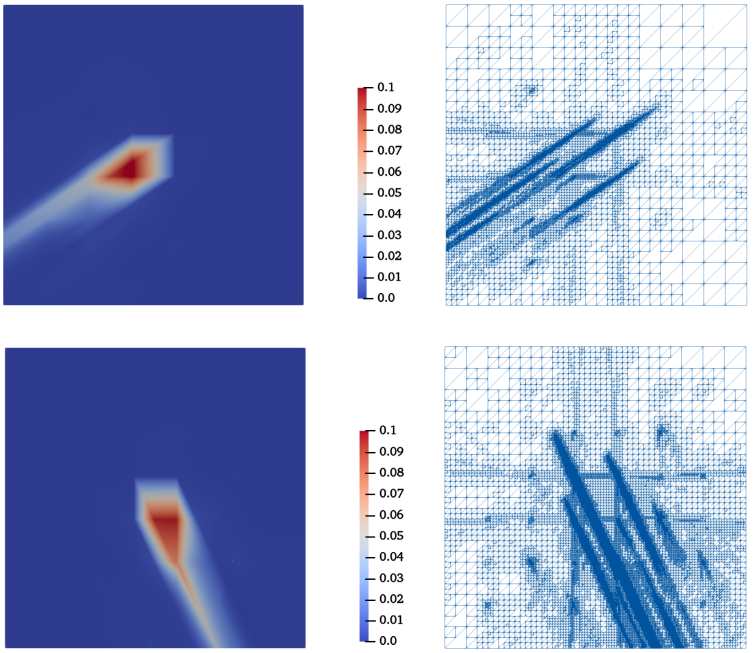


FIGURE 7. Solutions $\bar{u}_n(\cdot, \vec{s})$ for different directions $\vec{s}$ in final outer iterate.

The table below gives the precise values of the a posteriori error and the total degrees of freedom:

iteration	a posteriori error	#DoFs
1	0.0850598	6228
2	0.0891398	12456
3	0.079258	13392
4	0.0578653	38664
5	0.039463	135236
6	0.0258249	440648
7	0.0165168	1151102
8	0.010397	6586094
9	0.00647563	16570210
10	0.00400132	42179602

Figure 7 shows solutions $\bar{u}_n(\cdot, \vec{s})$ with their corresponding grids for the final iterate once the accuracy ε has been reached. Finally, Figure 6 shows the final averaged densities $\int_S \bar{u}_n(\cdot, \vec{s})\, \mathrm{d}\vec{s}$. They are computed on the merged grids.

We note that no special *structure preserving* measures had to be imposed on the numerical schemes to produce physically meaningful results.

Remark 7.1. The code to reproduce the numerical part of this article is available online at:

`https://gitlab.dune-project.org/felix.gruber/dune-dpg`

The implementation makes use of DUNE-DPG 0.4.2, a C++ based library which is built upon the multi-purpose finite element package DUNE [6]. Details of the DUNE-DPG library can be found in [18,19].

References

[1] B. K. Alpert, *A class of bases in L^2 for the sparse representation of integral operators*, SIAM J. Math. Anal. **24** (1993), no. 1, 246–262, DOI 10.1137/0524016. MR1199538

[2] M. Asadzadeh, *L_2-error estimates for the discrete ordinates method for three-dimensional neutron transport*, Transport Theory Statist. Phys. **17** (1988), no. 1, 1–24, DOI 10.1080/00411458808230852. MR950623

[3] M. Asadzadeh, *A finite element method for the neutron transport equation in an infinite cylindrical domain*, SIAM J. Numer. Anal. **35** (1998), no. 4, 1299–1314, DOI 10.1137/S0036142992238119. MR1620140

[4] M. Avila, R. Codina, and J. Principe, *Spatial approximation of the radiation transport equation using a subgrid-scale finite element method*, Comput. Methods Appl. Mech. Engrg. **200** (2011), no. 5-8, 425–438, DOI 10.1016/j.cma.2010.11.003. MR2749012

[5] G. Bal, *Inverse transport theory and applications*, Inverse Problems **25** (2009), no. 5, 053001, 48, DOI 10.1088/0266-5611/25/5/053001. MR2501018

[6] M. Blatt, A. Burchardt, A. Dedner, C. Engwer, J. Fahlke, B. Flemisch, C. Gersbacher, C. Gräser, F. Gruber, C. Grüninger, D. Kempf, R. Klöfkorn, T. Malkmus, S. Müthing, M. Nolte, M. Piatkowski, and O. Sander, *The Distributed and Unified Numerics Environment*, Version 2.4. Archive of Numerical Software, 4(100), pp. 13–29, May 2016. `http://dx.doi.org/10.11588/ans.2016.100.26526`.

[7] D. Broersen, W. Dahmen, and R. P. Stevenson, *On the stability of DPG formulations of transport equations*, Math. Comp. **87** (2018), no. 311, 1051–1082, DOI 10.1090/mcom/3242. MR3766381

[8] T. A. Brunner, *Forms of Approximate Radiation Transport*, Sandia Report SAND2002-1778, Sandia National Laboratories, July 2002. `http://dx.doi.org/10.2172/800993`.

[9] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods for elliptic operator equations: convergence rates*, Math. Comp. **70** (2001), no. 233, 27–75, DOI 10.1090/S0025-5718-00-01252-7. MR1803124

[10] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods. II. Beyond the elliptic case*, Found. Comput. Math. **2** (2002), no. 3, 203–245, DOI 10.1007/s102080010027. MR1907380

[11] W. Dahmen, H. Harbrecht, and R. Schneider, *Compression techniques for boundary integral equations—asymptotically optimal complexity estimates*, SIAM J. Numer. Anal. **43** (2006), no. 6, 2251–2271, DOI 10.1137/S0036142903428852. MR2206435

[12] W. Dahmen, C. Huang, C. Schwab, and G. Welper, *Adaptive Petrov-Galerkin methods for first order transport equations*, SIAM J. Numer. Anal. **50** (2012), no. 5, 2420–2445, DOI 10.1137/110823158. MR3022225

[13] W. Dahmen and R. P. Stevenson, *Adaptive strategies for transport equations*, Comput. Methods Appl. Math. **19** (2019), no. 3, 431–464, DOI 10.1515/cmam-2018-0230. MR3977482

[14] R. Dautray and J.-L. Lions, *Mathematical analysis and numerical methods for science and technology. Vol. 6*, Springer-Verlag, Berlin, 1993. Evolution problems II. MR1295030

[15] H. Egger and M. Schlottbom, *A mixed variational framework for the radiative transfer equation*, Math. Models Methods Appl. Sci. **22** (2012), no. 3, 1150014, 30, DOI 10.1142/S021820251150014X. MR2890452

[16] H. Egger and M. Schlottbom, *An L^p theory for stationary radiative transfer*, Appl. Anal. **93** (2014), no. 6, 1283–1296, DOI 10.1080/00036811.2013.826798. MR3195888

[17] K. Grella and C. Schwab, *Sparse discrete ordinates method in radiative transfer*, Comput. Methods Appl. Math. **11** (2011), no. 3, 305–326, DOI 10.2478/cmam-2011-0017. MR2844781

[18] F. Gruber, *Adaptive Source Term Iteration: A Stable Formulation for Radiative Transfer*, Ph.D. thesis, RWTH Aachen University, 2018. `http://dx.doi.org/10.18154/RWTH-2018-230893`.

[19] F. Gruber, A. Klewinghaus, and O. Mula, *The DUNE-DPG Library for Solving PDEs with Discontinuous Petrov–Galerkin Finite Elements*, Archive of Numerical Software, 5(1), pp. 111–128, 6 March 2017. `http://dx.doi.org/10.11588/ans.2017.1.27719`.

[20] J.-L. Guermond and G. Kanschat, *Asymptotic analysis of upwind discontinuous Galerkin approximation of the radiative transport equation in the diffusive limit*, SIAM J. Numer. Anal. **48** (2010), no. 1, 53–78, DOI 10.1137/090746938. MR2608358

[21] L. G. Henyey and J. L. Greenstein, *Diffuse Radiation in the Galaxy*, The Astrophysical Journal, 93, pp. 70–83, 1941.

[22] C. Johnson and J. Pitkäranta, *Convergence of a fully discrete scheme for two-dimensional neutron transport*, SIAM J. Numer. Anal. **20** (1983), no. 5, 951–966, DOI 10.1137/0720065. MR714690

[23] G. Kanschat, *Solution of radiative transfer problems with finite elements*, Numerical methods in multidimensional radiative transfer, Springer, Berlin, 2009, pp. 49–98, DOI 10.1007/978-3-540-85369-5_5. MR2484899

[24] M. F. Modest and J. Yang, *Elliptic PDE Formulation and Boundary Conditions of the Spherical Harmonics Method of Arbitrary Order for General Three-dimensional Geometries*, Journal of Quantitative Spectroscopy and Radiative Transfer, 109(9), pp. 1641–1666, 2008. http://dx.doi.org/10.1016/j.jqsrt.2007.12.018.

[25] M. Mokhtar-Kharroubi, *Mathematical Topics in Neutron Transport Theory*, Series on Advances in Mathematics for Applied Sciences, vol. 46, World Scientific Publishing Co., Inc., River Edge, NJ, 1997. New aspects; With a chapter by M. Choulli and P. Stefanov. MR1612403

[26] J. C. Ragusa, J.-L. Guermond, and G. Kanschat, *A robust S_N-DG-approximation for radiation transport in optically thick and diffusive regimes*, J. Comput. Phys. **231** (2012), no. 4, 1947–1962, DOI 10.1016/j.jcp.2011.11.017. MR2876596

Mathematics Department, University of South Carolina, 1523 Greene Street, Columbia, South Carolina 29208
Email address: dahmen@math.sc.edu

IGPM, RWTH Aachen, Templergraben 55, 52056 Aachen, Germany
Email address: felgru@gmx.de

CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine, PSL University, 75016 Paris, France
Email address: mula@ceremade.dauphine.fr