

A robust collision source method for rank adaptive dynamical low-rank approximation in radiation therapy

Jonas Kusch, Pia Stammer

CRC Preprint 2022/5, January 2022

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Participating universities



Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



Funded by

DFG

A robust collision source method for rank adaptive dynamical low-rank approximation in radiation therapy

Jonas Kusch^a, Pia Stammer^{b,c,d}

^aKarlsruhe Institute of Technology, Mathematical Science and Computational Methods, Karlsruhe, Germany, jonas.kusch@kit.edu

^bKarlsruhe Institute of Technology, Steinbuch Centre for Computing, Karlsruhe, Germany, pia.stammer@kit.edu

^cGerman Cancer Research Center - DKFZ, Department of Medical Physics in Radiation Oncology, Heidelberg, Germany

^dHIDSS4Health - Helmholtz Information and Data Science School for Health, Karlsruhe/Heidelberg, Germany

Abstract

Deterministic models for radiation transport describe the density of radiation particles moving through a background material. In radiation therapy applications, the phase space of this density is composed of energy, spatial position and direction of flight. The resulting six-dimensional phase space prohibits fine numerical discretizations, which are essential for the construction of accurate and reliable treatment plans. In this work, we tackle the high dimensional phase space through a dynamical low-rank approximation of the particle density. Dynamical low-rank approximation (DLRA) evolves the solution on a low-rank manifold in time. Interpreting the energy variable as a pseudo-time lets us employ the DLRA framework to represent the solution of the radiation transport equation on a low-rank manifold for every energy. Stiff scattering terms are treated through an efficient implicit energy discretization and a rank adaptive integrator is chosen to dynamically adapt the rank in energy. To facilitate the use of boundary conditions and reduce the overall rank, the radiation transport equation is split into collided and uncollided particles through a collision source method. Uncollided particles are described by a directed quadrature set guaranteeing low computational costs, whereas collided particles are represented by a low-rank solution. It can be shown that the presented method is L^2 -stable under a time step restriction which does not depend on stiff scattering terms. Moreover, the implicit treatment of scattering does not require numerical inversions of matrices. Numerical results for radiation therapy configurations as well as the line source benchmark underline the efficiency of the proposed method.

Keywords: Dynamical low-rank approximation, radiation therapy, kinetic equations, rank adaptivity, model order reduction

1. Introduction

Radiation therapy is one of the main tools in cancer treatment. In treatment planning, the set-up of radiation beams is optimized such that the tumor receives the prescribed dose, while minimizing the damage of surrounding risk organs or healthy tissue. Here, exact and fast dose calculation is of basic importance. Exact dose calculations however require the solution of a high dimensional coupled system of linear transport equations. Therefore, clinical dose calculation algorithms often rely on simplified *pencil beam* models [2], which are based on the Fermi-Eyges theory of radiative transfer [17]. Although these models are computationally efficient, they can only describe layered heterogeneities [25]. Thus simulation results are inaccurate, especially in cases including air cavities or other inhomogeneities [25, 29].

More exact dose calculation can be achieved by an appropriate Monte Carlo (MC) algorithm, where individual interacting particles are directly simulated [4]. However, while recent performance-tuned MC

implementations have achieved large run time improvements, the high computational costs and associated statistical noise still limit their clinical usage [18, 26]. The application of deterministic Boltzmann equations for dose calculation can achieve similar accuracy as MC simulations, but also exhibit the same computational complexity for grid-based solutions [6]. Moment models for the Boltzmann equation in radiation dose calculation have for example been considered in [19, 35]

In this work, we tackle the challenges arising from high-dimensional phase spaces in radiation therapy applications through dynamical low-rank approximation (DLRA) [28]. DLRA represents the solution by a low-rank ansatz. When the solution is an n by m time-dependent matrix with huge n and m , such a low-rank representation can be defined by a singular value decomposition (SVD) truncated to a (small) rank r .¹ Time evolution equations for each of the low-rank factors of the SVD are then derived by minimizing the residual while maintaining the solution’s low-rank structure. When the original method requires $O(n \cdot m)$ operations per time step, the updates of the low-rank factors only require $O(r^2 \cdot (n + m))$ operations. Robust integrators for the time evolution of these factors are the matrix projector–splitting integrator (PSI) [34] and the recently developed “unconventional” basis-update & Galerkin low-rank matrix integrator of [9]. Unlike the unconventional integrator, the projector–splitting integrator can be extended to high order. However, since it includes a backward step in time, the projector–splitting integrator can yield instabilities for parabolic equations. Moreover, the backward time step can result in unstable schemes for hyperbolic problems, which is not an issue for the unconventional integrator [30]. Following [30], stable discretizations of the PSI can however be constructed by using the continuous DLRA formulation proposed in [14].

The efficiency of dynamical low-rank approximation has been demonstrated in several applications, including kinetic theory [14, 15, 38, 37, 16, 12, 13, 8, 30, 11]. Two main challenges of DLRA in the context of kinetic theory and radiation transport specifically are the preservation of mass as well as capturing the asymptotic limit. Methods to guarantee mass conservation include re-scaling strategies [38], a high-order low-order (HOLO) decomposition [37] and the incorporation of certain basis functions in the tangent space of low-rank functions [13]. An asymptotic preserving scheme for dynamical low-rank approximation has been proposed in [11]. The key ingredients of this scheme are the ordering of low-rank updates and the choice of an implicit time discretization. A further method uses a HOLO scheme to guarantee preserving the asymptotic limit [12]. Even though radiation therapy applications do not exhibit sufficiently strong scattering to fall into the diffusive regime, stiff scattering terms remain a challenge. While implicit time discretizations guarantee a stable treatment of such terms, their significantly increased computational costs pose difficulties.

The efficiency of DLRA highly depends on the rank required to capture important solution characteristics. Choosing this rank sufficiently high to guarantee a satisfactory solution quality while at best maintaining low computational costs requires a great amount of intuition. Furthermore, a fixed choice of the rank does not capture the time evolution of the solution complexity. Rank adaptive DLRA integrators which pick the rank in an automated fashion during run time have for example been proposed in [10, 39, 8].

This work presents a dynamical low-rank approximation for radiation therapy. To employ the DLRA framework in this setting, we formulate the energy dependency of the continuous slowing down equation as a pseudo-time. Dynamical low-rank approximation is then used to update the low-rank factors of the solution in energy. An efficient choice of the rank for every energy is provided through the rank adaptive integrator of [8]. Further novelties of this work are:

1. *A stable and efficient time discretization of stiff scattering terms.* Following [36, 30], stiff scattering terms are split from the radiation transport equation. The unconventional integrator is used to time update the streaming part and the matrix projector-splitting integrator updates the scattering part. According to [30], the scattering part only requires an update of a single low-rank factor. By using an implicit time discretization on this part of the integrator and explicit updates on the remainder, we significantly reduce computational costs while allowing for a less restrictive CFL condition.

¹Note that the low-rank solution does not require a diagonal r by r singular value matrix. In fact, DLRA uses dense coefficient matrices.

2. *A first-collision source method to reduce the rank and impose boundary conditions.* Boundary conditions in radiation therapy are often Dirichlet conditions of uncollided particles traveling into a single direction. To efficiently incorporate this information into our solution ansatz, we perform a collided-uncollided split, see e.g. [1, 24]. Here, collided and uncollided particles are treated in two separate equations. For the uncollided particles, an S_N method with a directed quadrature set resolving only the small number of relevant directions can be used. The collided part of the solution is represented through dynamical low-rank approximation. This strips away highly peaked particle distributions in the low-rank approximation, thereby potentially reducing the rank. Furthermore, since the density of uncollided particles is zero at the boundary, imposing boundary conditions becomes straightforward.
3. *A multilevel dynamical low-rank approximation.* The collided-uncollided split can be extended to L -collided splits. This can be interpreted as writing the solution as a telescoping sum. Expecting a reduced rank in every of these telescoping updates, individual dynamical low-rank approximations are derived for every term.

This paper is structured as follows: Sections 2 and 3 give an overview on methods used in this work to point to existing work and to fix notation. A general background to the used radiation transport methods, especially the continuous slowing down equation and numerical methods to solve it is given in Section 2. Section 3 provides a recap of dynamical low-rank approximation as well as robust integrators for the DLRA evolution equations. The main method of this work is presented in Section 4. In Section 5, we derive a CFL condition which ensures stability and extend the results to the rank adaptive unconventional integrator in Section 6.

2. Recap: Mesoscopic transport models in radiation therapy

2.1. Continuous slowing down equation

In the following, we discuss kinetic transport in the field of radiation treatment planning as well as related numerical methods used in this work. The task of computational mathematics in radiation therapy is to predict the transport of radiation particles in cancer patients. An accurate model is provided by the linear Boltzmann equation, which describes the dynamics of the particle density ψ on a mesoscopic level [19]. The continuous slowing down (CSD) approximation [32] to the linear Boltzmann equation reads

$$\boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \psi(E, \mathbf{x}, \boldsymbol{\Omega}) + \rho(\mathbf{x}) \Sigma_t(E) \psi(E, \mathbf{x}, \boldsymbol{\Omega}) = \int_{\mathbb{S}^2} \rho(\mathbf{x}) \Sigma_s(E, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}') \psi(E, \mathbf{x}, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' + \partial_E (\rho(\mathbf{x}) S_t(E) \psi(E, \mathbf{x}, \boldsymbol{\Omega})) , \quad (1a)$$

$$\psi(E, \mathbf{x}, \boldsymbol{\Omega}) = \psi_{\text{BC}}(E, \mathbf{x}, \boldsymbol{\Omega}) \quad \text{for } \mathbf{x} \in \partial D , \quad (1b)$$

$$\psi(E_{\text{max}}, \mathbf{x}, \boldsymbol{\Omega}) = \psi_{\text{max}}(\mathbf{x}, \boldsymbol{\Omega}) . \quad (1c)$$

The phase space of the particle density ψ consists of energy $E \in [0, E_{\text{max}}] \subset \mathbb{R}_+$, space $\mathbf{x} \in D \subset \mathbb{R}^3$ and direction of flight $\boldsymbol{\Omega} \in \mathbb{S}^2$. We use $S_t : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ to denote the stopping power, which describes the rate at which particles lose energy. The tissue density of the patient is $\rho : D \rightarrow \mathbb{R}_+$. Material cross sections Σ_t and Σ_s describe scattering and absorption interactions of particles with tissue. Stopping power and material cross-sections are given from physical databases [23]. The quantity of interest is the dose absorbed by the tissue, which can be determined from

$$D(\mathbf{x}) = \frac{1}{\rho(\mathbf{x})} \int_0^\infty \int_{\mathbb{S}^2} S_t(E) \psi(E, \mathbf{x}, \boldsymbol{\Omega}) d\boldsymbol{\Omega} dE . \quad (2)$$

Since particles are assumed to lose energy continuously, the energy can be interpreted as a pseudo-time t . By performing a transformation of cross-sections and the particle density, cf. [5], the CSD equation (1) becomes

$$\partial_t \tilde{\psi} = -\boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \frac{\tilde{\psi}}{\rho} - \tilde{\Sigma}_t \tilde{\psi} + \int_{\mathbb{S}^2} \tilde{\Sigma}_s(t, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}') \tilde{\psi}(t, \mathbf{x}, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' := R(t, \tilde{\psi}) . \quad (3)$$

We omit phase space dependencies for ease of presentation. The pseudo-time t is defined as

$$t(E) := \int_0^{E_{\max}} \frac{1}{S(E')} dE' - \int_0^E \frac{1}{S(E')} dE' . \quad (4)$$

A tilde is used to denote the transformation of particle densities and cross sections, given by $\tilde{\Sigma}_t(t) = \Sigma_t(E(t))$, $\tilde{\Sigma}_s(t, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}') = \Sigma_s(E(t), \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}')$ and

$$\tilde{\psi}(t, \mathbf{x}, \boldsymbol{\Omega}) := S(E(t))\rho(\mathbf{x})\psi(E(t), \mathbf{x}, \boldsymbol{\Omega}) . \quad (5)$$

To simplify our presentation, we will from now on omit the tilde and always refer to the transformed quantities by ψ , Σ_t and Σ_s . Furthermore, the in-scattering operator will be denoted as \mathcal{S} , i.e.,

$$(\mathcal{S}\psi)(t, \mathbf{x}, \boldsymbol{\Omega}) := \int_{\mathbb{S}^2} \Sigma_s(t, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}')\psi(t, \mathbf{x}, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' . \quad (6)$$

2.2. Collision source method

First collision source methods split the solution to the radiation transport equation (3) into a collided and uncollided part, see e.g. [3, 24]. Let us write $\psi(t, \mathbf{x}, \boldsymbol{\Omega}) = \psi_u(t, \mathbf{x}, \boldsymbol{\Omega}) + \psi_c(t, \mathbf{x}, \boldsymbol{\Omega})$, where ψ_u represents uncollided and ψ_c represents collided particles. Then, the radiation transport equation (3) can be split into

$$\partial_t \psi_u = -\boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \frac{\psi_u}{\rho} - \Sigma_t \psi_u := R_u(t, \psi_u) , \quad (7a)$$

$$\partial_t \psi_c = -\boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \frac{\psi_c}{\rho} - \Sigma_t \psi_c + \mathcal{S}(\psi_u + \psi_c) := R_c(t, \psi_u, \psi_c) . \quad (7b)$$

The first equation describes the dynamics of uncollided particles. Since collided particles cannot generate or deplete uncollided particles, (7a) solely depends on ψ_u . Furthermore, as uncollided particles do not generate in-scattering, only the outscattering term $-\Sigma_t \psi_u$ describes interactions with the background material. The dynamics of collided particles are described by (7b) and include in-scattering from uncollided particles. This methodology of representing the solution in terms of collided and uncollided particles can be developed further: Denoting the particles that have collided $\ell = 0, \dots, L$ times as ψ_ℓ and particles that have collided more than L times as ψ_c , we have $\psi = \psi_0 + \psi_1 + \dots + \psi_L + \psi_c$. Then, we obtain the equations

$$\partial_t \psi_0 = -\boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \frac{\psi_0}{\rho} - \Sigma_t \psi_0 := R_0(t, \psi_0) , \quad (8a)$$

$$\partial_t \psi_\ell = -\boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \frac{\psi_\ell}{\rho} - \Sigma_t \psi_\ell + \mathcal{S}\psi_{\ell-1} := R_1(t, \psi_{\ell-1}, \psi_\ell) , \quad \text{for } \ell = 1, \dots, L, \quad (8b)$$

$$\partial_t \psi_c = -\boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \frac{\psi_c}{\rho} - \Sigma_t \psi_c + \mathcal{S}(\psi_L + \psi_c) := R_c(t, \psi_L, \psi_c) . \quad (8c)$$

2.3. Angular discretization

To allow for numerical approximations of solutions to the presented equations, time, space and angle need to be discretized. Numerical artifacts that numerical solutions exhibit crucially depend on the angular discretization. Deterministic discretizations can be classified into nodal and modal methods. A conventional nodal method is the *discrete ordinates* (S_N) [33] method, which evolves the solution on a chosen angular quadrature set. A conventional modal discretization technique is the *spherical harmonics* (P_N) method [7] which spans the solution in terms of spherical harmonics basis functions. For degree ℓ and order k , the spherical harmonics for an angle $\boldsymbol{\Omega} = (\sqrt{1 - \mu^2} \cos \varphi, \sqrt{1 - \mu^2} \sin \varphi, \mu)^T$ are defined as

$$Y_\ell^k(\boldsymbol{\Omega}) = \sqrt{\frac{2\ell + 1}{4\pi} \frac{(\ell - k)!}{(\ell + k)!}} e^{ik\varphi} P_\ell^k(\mu) ,$$

where P_ℓ^k are the associated Legendre polynomials. In this work, we use the real spherical harmonics

$$m_\ell^k = \begin{cases} \frac{(-1)^k}{\sqrt{2}} (Y_\ell^k + (-1)^k Y_\ell^{-k}), & k > 0, \\ Y_\ell^0 & k = 0, \\ -\frac{(-1)^k}{\sqrt{2}} (Y_\ell^{-k} - (-1)^k Y_\ell^k), & k < 0. \end{cases}$$

Let us collect all basis functions up to degree N in a vector

$$\mathbf{m} = (m_0^0, m_1^{-1}, m_1^0, m_1^1, \dots, m_N^N)^T \in \mathbb{R}^{(N+1)^2}$$

and choose the modal approximation $\psi(t, \mathbf{x}, \boldsymbol{\Omega}) \approx \mathbf{u}(t, \mathbf{x})^T \mathbf{m}(\boldsymbol{\Omega})$. The coefficient (or moment) vector \mathbf{u} has $m := (N+1)^2$ entries. Then, the P_N equations for $\mathbf{x} = (x, y, z)^T$ read

$$\partial_t \mathbf{u}(t, \mathbf{x}) = -\mathbf{A} \cdot \nabla_{\mathbf{x}} \frac{\mathbf{u}(t, \mathbf{x})}{\rho(\mathbf{x})} - \Sigma_t(t) \mathbf{u}(t, \mathbf{x}) + \boldsymbol{\Sigma} \mathbf{u}(t, \mathbf{x}),$$

where $\mathbf{A} \cdot \nabla_{\mathbf{x}} := \mathbf{A}_1 \partial_x + \mathbf{A}_2 \partial_y + \mathbf{A}_3 \partial_z$ with $\mathbf{A}_i := \int_{\mathbb{S}^2} \mathbf{m} \mathbf{m}^T \Omega_i d\boldsymbol{\Omega}$. The diagonal in-scattering matrix $\boldsymbol{\Sigma}$ has entries $\Sigma_{kk}(t) = 2\pi \int_{[-1,1]} P_k(\mu) \Sigma_s(t, \mu) d\mu$. Note that $\Sigma_t(t) = \Sigma_{11}(t) > 0$ and

$$|\Sigma_{kk}(t)| \leq 2\pi \int_{[-1,1]} |P_k(\mu)| \cdot |\Sigma_s(t, \mu)| d\mu \leq 2\pi \int_{[-1,1]} |\Sigma_s(t, \mu)| d\mu = \Sigma_{11}(t). \quad (9)$$

2.4. Spatial discretization

Several benchmarks in radiation transport assume a two-dimensional spatial domain, i.e., the spatial variable becomes $\mathbf{x} = (x, y)^T$. In this setting, we focus on the discretization of the P_N equations, since these will become relevant in the proposed DLRA method. Then, we have

$$\partial_t \mathbf{u}(t, \mathbf{x}) = -\mathbf{A}_x \partial_x \mathbf{u}(t, \mathbf{x}) - \mathbf{A}_y \partial_y \mathbf{u}(t, \mathbf{x}) + \mathbf{G} \mathbf{u}(t, \mathbf{x}). \quad (10)$$

A finite volume discretization splits the spatial domain D into $N_x \cdot N_y$ cells. In case of a structured quadrangle grid, the x and y domains are discretized into uniform one-dimensional grids $x_1 \leq x_2 \leq \dots \leq x_{N_x+1}$ and $y_1 \leq y_2 \leq \dots \leq y_{N_y+1}$ with grid size Δx and Δy respectively. Then, the cell of index (i, j) is defined on $I_{ij} := [x_i, x_{i+1}] \times [y_j, y_{j+1}]$ on which the numerical solution is chosen as

$$\mathbf{u}_{ij}(t) \simeq \frac{1}{\Delta x \Delta y} \int_{I_{ij}} \mathbf{u}(t, \mathbf{x}) d\mathbf{x}.$$

In the same manner we compute the patient density ρ_{ij} . To simplify the presentation of the spatial discretization, let us collect $\mathbf{u}_{ij}(t) = (u_{ijk}(t))_{k=1}^m \in \mathbb{R}^m$ for $i = 1, \dots, N_x$ and $j = 1, \dots, N_y$ into a matrix $\mathbf{u}(t) \in \mathbb{R}^{n_x \times m}$, where $n_x := N_x \cdot N_y$. For this, we define the function $\text{idx} : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ as $\text{idx}(i, j) = (i-1) \cdot N_x + j$. Then, the entries of $\mathbf{u}(t)$ are defined as $u_{\text{idx}(i,j), k}(t) = u_{ijk}(t)$.

With this notation at hand, we can define a finite volume scheme for (10) in compact notation. Let us define the sparse diffusion stencil matrices $\mathbf{L}_{x,y}^{(1)} \in \mathbb{R}^{n_x \times n_x}$ as

$$\begin{aligned} L_{x, \text{idx}(i,j), \text{idx}(i,j)}^{(1)} &= \frac{1}{\rho_{ij} \Delta x}, & L_{x, \text{idx}(i,j), \text{idx}(i\pm 1, j)}^{(1)} &= -\frac{1}{2\rho_{i\pm 1, j} \Delta x}, \\ L_{y, \text{idx}(i,j), \text{idx}(i,j)}^{(1)} &= \frac{1}{\rho_{ij} \Delta y}, & L_{y, \text{idx}(i,j), \text{idx}(i, j\pm 1)}^{(1)} &= -\frac{1}{2\rho_{i, j\pm 1} \Delta y}, \end{aligned}$$

as well as the sparse advection stencil matrices $\mathbf{L}_{x,y}^{(2)} \in \mathbb{R}^{n_x \times n_x}$ as

$$L_{x, \text{idx}(i,j), \text{idx}(i\pm 1, j)}^{(2)} = \frac{\pm 1}{2\rho_{i\pm 1, j} \Delta x}, \quad L_{y, \text{idx}(i,j), \text{idx}(i, j\pm 1)}^{(2)} = \frac{\pm 1}{2\rho_{i, j\pm 1} \Delta y}.$$

Furthermore with $\mathbf{A}_{x,y} = \mathbf{V}_{x,y} \mathbf{\Lambda}_{x,y} \mathbf{V}_{x,y}^T$, the Roe matrices $|\mathbf{A}_x|, |\mathbf{A}_y| \in \mathbb{R}^{m \times m}$ are defined as

$$|\mathbf{A}_x| := \mathbf{V}_x |\mathbf{\Lambda}_x| \mathbf{V}_x^T, \text{ and } |\mathbf{V}_y| := \mathbf{V}_y |\mathbf{\Lambda}_y| \mathbf{V}_y^T.$$

Here, $\mathbf{V}_{x,y}$ collects the orthonormal eigenvectors of the symmetric matrices $\mathbf{A}_{x,y}$ and $\mathbf{\Lambda}_{x,y} = \text{diag}(\lambda_1^{x,y}, \dots, \lambda_m^{x,y})$ are the corresponding real eigenvalues. Then, the semi-discrete finite volume update becomes a huge matrix differential equation of the form

$$\dot{\mathbf{u}}(t) = \mathbf{F}(\mathbf{u}(t)) + \mathbf{G}(t, \mathbf{u}(t), \mathbf{u}(t)),$$

where

$$\begin{aligned} \mathbf{F}(\mathbf{u}) &:= \mathbf{L}_x^{(2)} \mathbf{u} \mathbf{A}_x^T + \mathbf{L}_y^{(2)} \mathbf{u} \mathbf{A}_y^T + \mathbf{L}_x^{(1)} \mathbf{u} |\mathbf{A}_x|^T + \mathbf{L}_y^{(1)} \mathbf{u} |\mathbf{A}_y|^T, \\ \mathbf{G}(t, \mathbf{v}, \mathbf{u}) &:= -\Sigma_t(t) \mathbf{u} + \mathbf{v} \Sigma(t). \end{aligned}$$

The costs of evaluating the right-hand side are $C_{P_N} \lesssim n_x \cdot m$, when accounting for the sparsity of all stencil, flux and Roe matrices which leads to linear costs in n_x and m of matrix products.

3. Recap: Dynamical low-rank approximation

3.1. Main framework

This section gives a brief overview on dynamical low-rank approximation [28] for matrix differential equations $\dot{\mathbf{u}}(t) = \mathbf{F}(t, \mathbf{u}(t))$. Dynamical low-rank approximation represents and evolves the solution on a manifold of rank r matrices, which we denote by \mathcal{M}_r . A low-rank representation is given by the SVD-like factorization

$$\mathbf{u}(t) \approx \mathbf{X}(t) \mathbf{S}(t) \mathbf{W}(t)^T, \quad (11)$$

where $\mathbf{X} \in \mathbb{R}^{n_x \times r}$ and $\mathbf{W} \in \mathbb{R}^{m \times r}$ are basis matrices with orthonormal column vectors and $\mathbf{S} \in \mathbb{R}^{r \times r}$ is a dense coefficient matrix. Time evolution equations for the factors can be defined by imposing

$$\dot{\mathbf{u}}(t) \in T_{\mathbf{u}(t)} \mathcal{M}_r \quad \text{such that} \quad \|\dot{\mathbf{u}}(t) - \mathbf{F}(t, \mathbf{u}(t))\| = \min. \quad (12)$$

We use $T_{\mathbf{u}(t)} \mathcal{M}_r$ to denote the tangent space of \mathcal{M}_r at $\mathbf{u}(t)$, i.e., the solution should remain of rank r over time while minimizing the defect. These conditions yield a time evolution equation for the low-rank solution [28, Lemma 4.1], which reads

$$\dot{\mathbf{u}}(t) = \mathbf{P}(\mathbf{u}(t)) \mathbf{F}(t, \mathbf{u}(t)). \quad (13)$$

The operator \mathbf{P} is the orthogonal projection onto the tangent space, given by

$$\mathbf{P} \mathbf{g} = \mathbf{X} \mathbf{X}^T \mathbf{g} - \mathbf{X} \mathbf{X}^T \mathbf{g} \mathbf{W} \mathbf{W}^T + \mathbf{g} \mathbf{W} \mathbf{W}^T.$$

Evolution equations can then be derived for the factors \mathbf{X} , \mathbf{S} and \mathbf{W} , see [28]. However, the resulting equations depend on the inverse of the commonly ill conditioned coefficient matrix, which substantially limits the permitted step size [27].

3.2. Robust fixed rank integrators

Two robust integrators have been proposed for the evolution equation (13). First, the matrix projector-splitting integrator [34] splits (13) by a Lie-Trotter splitting technique, yielding

$$\dot{\mathbf{u}}_I(t) = \mathbf{F}(\mathbf{u}_I(t)) \mathbf{W} \mathbf{W}^T, \quad \mathbf{u}_I(t_0) = \mathbf{u}(t_0), \quad (14a)$$

$$\dot{\mathbf{u}}_{II}(t) = -\mathbf{X} \mathbf{X}^T \mathbf{F}(\mathbf{u}_{II}(t)) \mathbf{W} \mathbf{W}^T, \quad \mathbf{u}_{II}(t_0) = \mathbf{u}_I(t_1), \quad (14b)$$

$$\dot{\mathbf{u}}_{III}(t) = \mathbf{X} \mathbf{X}^T \mathbf{F}(\mathbf{u}_{III}(t)), \quad \mathbf{u}_{III}(t_0) = \mathbf{u}_{II}(t_1). \quad (14c)$$

The resulting consecutive movement in the low-rank manifold ensures robustness irrespective of singular values and thereby allows for increased step sizes [34]. Defining the decompositions $\mathbf{u}_I = \mathbf{K} \mathbf{W}^T$ as well as $\mathbf{u}_{III} = \mathbf{X} \mathbf{L}$ gives the *matrix projector-splitting integrator*, which updates the low-rank factors $\mathbf{X}^0 = \mathbf{X}(t_0)$, $\mathbf{W}^0 = \mathbf{W}(t_0)$ and $\mathbf{S}^0 = \mathbf{S}(t_0)$ to time $t_1 = t_0 + \Delta t$:

1. *K-step*: Update \mathbf{X}^0 to \mathbf{X}^1 and \mathbf{S}^0 to $\tilde{\mathbf{S}}^0$ via

$$\dot{\mathbf{K}}(t) = \mathbf{F}(\mathbf{K}(t)\mathbf{W}^{0,T})\mathbf{W}^0, \quad \mathbf{K}(t_0) = \mathbf{X}^0\mathbf{S}^0. \quad (15)$$

Determine \mathbf{X}^1 and $\tilde{\mathbf{S}}^0$ with $\mathbf{K}(t_1) = \mathbf{X}^1\tilde{\mathbf{S}}^0$ by performing a QR decomposition.

2. *S-step*: Update $\tilde{\mathbf{S}}^0$ to $\tilde{\mathbf{S}}^1$ via

$$\dot{\tilde{\mathbf{S}}}(t) = -\mathbf{X}^{1,T}\mathbf{F}(\mathbf{X}^1\tilde{\mathbf{S}}(t)\mathbf{W}^{0,T})\mathbf{W}^0, \quad \tilde{\mathbf{S}}(t_0) = \tilde{\mathbf{S}}^0 \quad (16)$$

and set $\tilde{\mathbf{S}}^1 = \tilde{\mathbf{S}}(t_1)$.

3. *L-step*: Update \mathbf{W}^0 to \mathbf{W}^1 and $\tilde{\mathbf{S}}^1$ to \mathbf{S}^1 via

$$\dot{\mathbf{L}}(t) = \mathbf{X}^{1,T}\mathbf{F}(\mathbf{X}^1\mathbf{L}(t)), \quad \mathbf{L}(t_0) = \tilde{\mathbf{S}}^1\mathbf{W}^{0,T}. \quad (17)$$

Determine \mathbf{W}^1 and \mathbf{S}^1 with $\mathbf{L}(t_1) = \mathbf{S}^1\mathbf{W}^{1,T}$ by performing a QR decomposition.

Then, the time updated solution is $\mathbf{u}(t_1) = \mathbf{X}^1\mathbf{S}^1\mathbf{W}^{1,T}$. It has been noted in [30] that when the flux function takes the form $\mathbf{F}(t, \mathbf{u}(t)) = \mathbf{u}(t)\mathbf{G}(t)$, the *K* and *S*-steps cancel out and only the *L*-step determines the dynamics.

The second robust integrator is the *unconventional integrator*, which has recently been introduced in [9]. This integrator first performs basis updates of \mathbf{X} and \mathbf{W} in parallel and then updates the coefficient matrix \mathbf{S} by a Galerkin step. This integrator shares the robustness properties of the matrix projector-splitting integrator [9]. It takes the form

1. *K-step*: Update \mathbf{X}^0 to \mathbf{X}^1 via

$$\dot{\mathbf{K}}(t) = \mathbf{F}(\mathbf{K}(t)\mathbf{W}^{0,T})\mathbf{W}^0, \quad \mathbf{K}(t_0) = \mathbf{X}^0\mathbf{S}^0. \quad (18)$$

Determine \mathbf{X}^1 with $\mathbf{K}(t_1) = \mathbf{X}^1\mathbf{R}$ and store $\mathbf{M} = \mathbf{X}^{1,T}\mathbf{X}^0$.

2. *L-step*: Update \mathbf{W}^0 to \mathbf{W}^1 via

$$\dot{\mathbf{L}}(t) = \mathbf{X}^{0,T}\mathbf{F}(\mathbf{X}^0\mathbf{L}(t)), \quad \mathbf{L}(t_0) = \mathbf{S}^0\mathbf{W}^{0,T}. \quad (19)$$

Determine \mathbf{W}^1 with $\mathbf{L}(t_1) = \mathbf{W}^1\tilde{\mathbf{R}}$ and store $\mathbf{N} = \mathbf{W}^{1,T}\mathbf{W}^0$.

3. *S-step*: Update \mathbf{S}^0 to \mathbf{S}^1 via

$$\dot{\mathbf{S}}(t) = \mathbf{X}^{1,T}\mathbf{F}(\mathbf{X}^1\mathbf{S}(t)\mathbf{W}^{1,T})\mathbf{W}^1, \quad \mathbf{S}(t_0) = \mathbf{M}\mathbf{S}^0\mathbf{N}^T \quad (20)$$

and set $\mathbf{S}^1 = \mathbf{S}(t_1)$.

3.3. Rank adaptive unconventional integrator

The unconventional integrator has recently been extended to allow for rank adaptivity [8]. That is, given a tolerance parameter ϑ , the integrator adapts the rank in time.

Starting from time t_0 where the solution has rank r_0 , the integrator gives the factored solution at time t_1 with rank $r_1 \leq 2r_0$. In the following, we use $r = r_0$ and use hats to denote matrices of rank $2r$. Then the rank adaptive integrator reads

1. *K-step*: Update $\mathbf{X}^0 \in \mathbb{R}^{n_x \times r}$ to $\hat{\mathbf{X}}^1 \in \mathbb{R}^{n_x \times 2r}$ via

$$\dot{\mathbf{K}}(t) = \mathbf{F}(\mathbf{K}(t)\mathbf{W}^{0,T})\mathbf{W}^0, \quad \mathbf{K}(t_0) = \mathbf{X}^0\mathbf{S}^0. \quad (21)$$

Determine $\hat{\mathbf{X}}^1$ with $[\mathbf{K}(t_1), \mathbf{X}^0] = \hat{\mathbf{X}}^1\mathbf{R}$ and store $\hat{\mathbf{M}} = \hat{\mathbf{X}}^{1,T}\mathbf{X}^0 \in \mathbb{R}^{2r \times r}$.

2. **L-step:** Update $\mathbf{W}^0 \in \mathbb{R}^{m \times r}$ to $\widehat{\mathbf{W}}^1 \in \mathbb{R}^{m \times 2r}$ via

$$\dot{\mathbf{L}}(t) = \mathbf{X}^{0,T} \mathbf{F}(\mathbf{X}^0 \mathbf{L}(t)), \quad \mathbf{L}(t_0) = \mathbf{S}^0 \mathbf{W}^{0,T}. \quad (22)$$

Determine $\widehat{\mathbf{W}}^1$ with $[\mathbf{L}(t_1), \mathbf{W}^0] = \widehat{\mathbf{W}}^1 \widetilde{\mathbf{R}}$ and store $\widehat{\mathbf{N}} = \widehat{\mathbf{W}}^{1,T} \mathbf{W}^0$.

3. **S-step:** Update $\mathbf{S}^0 \in \mathbb{R}^{r \times r}$ to $\widehat{\mathbf{S}}^1 \in \mathbb{R}^{2r \times 2r}$ via

$$\dot{\widehat{\mathbf{S}}}(t) = \widehat{\mathbf{X}}^{1,T} \mathbf{F}(\widehat{\mathbf{X}}^1 \widehat{\mathbf{S}}(t) \widehat{\mathbf{W}}^{1,T}) \widehat{\mathbf{W}}^1, \quad \widehat{\mathbf{S}}(t_0) = \widehat{\mathbf{M}} \mathbf{S}^0 \widehat{\mathbf{N}}^T \quad (23)$$

and set $\widehat{\mathbf{S}}^1 = \widehat{\mathbf{S}}(t_1)$.

4. **Truncation:** Determine the SVD $\widehat{\mathbf{S}}^1 = \widehat{\mathbf{P}} \widehat{\mathbf{\Sigma}} \widehat{\mathbf{Q}}^T$ where $\widehat{\mathbf{\Sigma}} = \text{diag}(\sigma_j)$. For a given tolerance ϑ , choose the new rank $r_1 \leq 2r$ such that

$$\left(\sum_{j=r_1+1}^{2r} \sigma_j^2 \right)^{1/2} \leq \vartheta.$$

Compute the new factors as follows: Let \mathbf{S}^1 be the $r_1 \times r_1$ diagonal matrix with the r_1 largest singular values and let $\mathbf{P}_1 \in \mathbb{R}^{2r \times r_1}$ and $\mathbf{Q}_1 \in \mathbb{R}^{2r \times r_1}$ contain the first r_1 columns of $\widehat{\mathbf{P}}$ and $\widehat{\mathbf{Q}}$, respectively. Finally, set $\mathbf{X}^1 = \widehat{\mathbf{X}}^1 \mathbf{P}_1 \in \mathbb{R}^{m \times r_1}$ and $\mathbf{W}^1 = \widehat{\mathbf{W}}^1 \mathbf{Q}_1 \in \mathbb{R}^{n \times r_1}$.

4. A robust collision source method for dynamical low-rank approximation

In this section we present the main method, which aims at providing an efficient and robust alternative to conventional strategies. Key ingredients for the construction are 1) a collision source method to define a splitting of the original equation, 2) a further splitting of collision terms which are treated implicitly, 3) computing individual DLRA updates by using the unconventional integrator for collided particles.

4.1. Collided-uncollided split

We start by deriving moment equations for the collided particles in the collision source method and use an S_N method for uncollided particles. Without going into detail, we denote the S_N solution of the uncollided particles as $\boldsymbol{\psi}$ and the right-hand side of a time continuous S_N method for streaming (i.e., particles move without interacting with tissue) as $\mathbf{F}_S(\boldsymbol{\psi})$. Then the evolution equations for the semi-discrete solution become

$$\dot{\boldsymbol{\psi}}(t) = \mathbf{F}_S(\boldsymbol{\psi}(t)) - \Sigma_t(t) \boldsymbol{\psi}(t), \quad (24a)$$

$$\dot{\mathbf{u}}_1(t) = \mathbf{F}(\mathbf{u}_1(t)) - \Sigma_t(t) \mathbf{u}_1(t) + \boldsymbol{\psi}(t) \mathbf{T}_M \boldsymbol{\Sigma}(t), \quad (24b)$$

$$\dot{\mathbf{u}}_\ell(t) = \mathbf{F}(\mathbf{u}_\ell(t)) - \Sigma_t(t) \mathbf{u}_\ell(t) + \mathbf{u}_{\ell-1}(t) \boldsymbol{\Sigma}(t), \quad \text{for } \ell = 2, \dots, L, \quad (24c)$$

$$\dot{\mathbf{u}}_c(t) = \mathbf{F}(\mathbf{u}_c(t)) - \Sigma_t(t) \mathbf{u}_c + (\mathbf{u}_L(t) + \mathbf{u}_c(t)) \boldsymbol{\Sigma}(t), \quad (24d)$$

where the matrix \mathbf{T}_M maps the nodal solution onto its moments. These $L + 2$ equations can be solved consecutively. Since radiation therapy commonly investigates the effects of particle beams which enter the computational domain from the boundary, the S_N equations from the uncollided particles can be solved efficiently by using a biased quadrature rule. I.e., the quadrature only encodes the small number of possible flight directions. After the first collision, particles move into all directions. To account for the increased complexity, we describe the collided solution through dynamical low-rank approximation. Following [30], we split streaming and scattering parts and use the matrix projector-splitting integrator to update in-scattering in equation (24d) as well as the unconventional integrator for the remainder.

To simplify our presentation, we start by discussing this strategy for (24b) and then extend it to the remaining equations. Let us first split *streaming* and *scattering* in (24b). Omitting the subscript 1 gives

$$\dot{\mathbf{u}}_I(t) = \mathbf{F}(\mathbf{u}_I(t)), \quad \mathbf{u}_I(t_0) = \mathbf{u}_1(t_0), \quad (25a)$$

$$\dot{\mathbf{u}}_{II}(t) = -\Sigma_t(t) \mathbf{u}_{II}(t) + \boldsymbol{\psi}(t_1) \mathbf{T}_M \boldsymbol{\Sigma}(t), \quad \mathbf{u}_{II}(t_0) = \mathbf{u}_I(t_1). \quad (25b)$$

The updated solution at time $t_1 = t_0 + \Delta t$ is then given as $\mathbf{u}_1(t_1) = \mathbf{u}_{II}(t_1)$. Note that the splitting method introduces an error of $O(\Delta t)$, which can be reduced by high order splitting methods. We start with the derivation of the basis update and Galerkin step equations of the unconventional integrator for the streaming part (25a). That is, we derive evolution equations for $\mathbf{X}_I(t), \mathbf{S}_I(t), \mathbf{W}_I(t)$ such that $\mathbf{u}_I(t) \approx \mathbf{X}_I(t)\mathbf{S}_I(t)\mathbf{W}_I(t)^T$. To simplify notation let us omit Roman indices in the following. The K -step equations (18) read

$$\begin{aligned}\dot{\mathbf{K}}(t) &= \mathbf{F}(\mathbf{K}(t)\mathbf{W}^{0,T})\mathbf{W}^0 \\ &= \mathbf{L}_x^{(2)}\mathbf{K}(t)\mathbf{W}^{0,T}\mathbf{A}_x^T\mathbf{W}^0 + \mathbf{L}_y^{(2)}\mathbf{K}(t)\mathbf{W}^{0,T}\mathbf{A}_y^T\mathbf{W}^0 \\ &\quad + \mathbf{L}_x^{(1)}\mathbf{K}(t)\mathbf{W}^{0,T}|\mathbf{A}_x|^T\mathbf{W}^0 + \mathbf{L}_y^{(1)}\mathbf{K}(t)\mathbf{W}^{0,T}|\mathbf{A}_y|^T\mathbf{W}^0 \\ &= \mathbf{L}_x^{(2)}\mathbf{K}(t)\widehat{\mathbf{A}}_x^0 + \mathbf{L}_y^{(2)}\mathbf{K}(t)\widehat{\mathbf{A}}_y^0 + \mathbf{L}_x^{(1)}\mathbf{K}(t)|\widehat{\mathbf{A}}_x|^0 + \mathbf{L}_y^{(1)}\mathbf{K}(t)|\widehat{\mathbf{A}}_y|^0,\end{aligned}\quad (26)$$

where we use $\widehat{\mathbf{A}}_{x,y}^0 := \mathbf{W}^{0,T}\mathbf{A}_{x,y}^T\mathbf{W}^0$ and $|\widehat{\mathbf{A}}_{x,y}|^0 := \mathbf{W}^{0,T}|\mathbf{A}_{x,y}|^T\mathbf{W}^0$. The numerical costs to compute these matrices are of $O(r^2 \cdot m)$ and evaluating the right-hand side of the K -step equations has costs of $O(r^2 \cdot n_x)$. To point out that the spatial basis is not yet updated by the scattering step (25b), we define the superscript $1/2$, i.e., the solution is given as $\mathbf{X}^{1/2} := \mathbf{X}_I(t_1)$.

The L -step equations (19) read

$$\begin{aligned}\dot{\mathbf{L}}(t) &= \mathbf{X}^{0,T}\mathbf{F}(\mathbf{X}^0\mathbf{L}(t)) \\ &= \mathbf{X}^{0,T}\mathbf{L}_x^{(2)}\mathbf{X}^0\mathbf{L}(t)\mathbf{A}_x^T + \mathbf{X}^{0,T}\mathbf{L}_y^{(2)}\mathbf{X}^0\mathbf{L}(t)\mathbf{A}_y^T \\ &\quad + \mathbf{X}^{0,T}\mathbf{L}_x^{(1)}\mathbf{X}^0\mathbf{L}(t)|\mathbf{A}_x|^T + \mathbf{X}^{0,T}\mathbf{L}_y^{(1)}\mathbf{X}^0\mathbf{L}(t)|\mathbf{A}_y|^T \\ &= \widehat{\mathbf{L}}_x^{(2),0}\mathbf{L}(t)\mathbf{A}_x^T + \widehat{\mathbf{L}}_y^{(2),0}\mathbf{L}(t)\mathbf{A}_y^T + \widehat{\mathbf{L}}_x^{(1),0}\mathbf{L}(t)|\mathbf{A}_x|^T + \widehat{\mathbf{L}}_y^{(1),0}\mathbf{L}(t)|\mathbf{A}_y|^T,\end{aligned}\quad (27)$$

where we use $\widehat{\mathbf{L}}_{x,y}^{(2),0} := \mathbf{X}^{0,T}\mathbf{L}_{x,y}^{(2)}\mathbf{X}^0$ and $\widehat{\mathbf{L}}_{x,y}^{(1),0} := \mathbf{X}^{0,T}\mathbf{L}_{x,y}^{(1)}\mathbf{X}^0$. The numerical costs to compute these matrices are of $O(r^2 \cdot n_x)$ and evaluating the right-hand side of the L -step equations has costs of $O(r^2 \cdot m)$. To point out that the directional basis is not yet updated by the scattering step (25b), we define $\mathbf{W}^{1/2} := \mathbf{W}_I(t_1)$.

The S -step equations (20) read

$$\begin{aligned}\dot{\mathbf{S}}(t) &= \mathbf{X}^{1/2,T}\mathbf{F}(\mathbf{X}^{1/2}\mathbf{S}(t)\mathbf{W}^{1/2,T})\mathbf{W}^{1/2} \\ &= \mathbf{X}^{1/2,T}\mathbf{L}_x^{(2)}\mathbf{X}^{1/2}\mathbf{S}(t)\mathbf{W}^{1/2,T}\mathbf{A}_x^T\mathbf{W}^{1/2} + \mathbf{X}^{1/2,T}\mathbf{L}_y^{(2)}\mathbf{X}^{1/2}\mathbf{S}(t)\mathbf{W}^{1/2,T}\mathbf{A}_y^T\mathbf{W}^{1/2} \\ &\quad + \mathbf{X}^{1/2,T}\mathbf{L}_x^{(1)}\mathbf{X}^{1/2}\mathbf{S}(t)\mathbf{W}^{1/2,T}|\mathbf{A}_x|^T\mathbf{W}^{1/2} + \mathbf{X}^{1/2,T}\mathbf{L}_y^{(1)}\mathbf{X}^{1/2}\mathbf{S}(t)\mathbf{W}^{1/2,T}|\mathbf{A}_y|^T\mathbf{W}^{1/2} \\ &= \widehat{\mathbf{L}}_x^{(2),1/2}\mathbf{S}(t)\widehat{\mathbf{A}}_x^{1/2} + \widehat{\mathbf{L}}_y^{(2),1/2}\mathbf{S}(t)\widehat{\mathbf{A}}_y^{1/2} + \widehat{\mathbf{L}}_x^{(1),1/2}\mathbf{S}(t)|\widehat{\mathbf{A}}_x|^{1/2} + \widehat{\mathbf{L}}_y^{(1),1/2}\mathbf{S}(t)|\widehat{\mathbf{A}}_y|^{1/2}.\end{aligned}\quad (28)$$

The numerical costs to compute update matrices and to evaluate the right-hand side of the S -step equations are of $O(r^2 \cdot (n_x + m))$. To point out that the coefficient matrix is not yet updated by the scattering step (25b), we define $\mathbf{S}^{1/2} := \mathbf{S}_I(t_1)$.

In a second step, we use the updated factors as initial condition, i.e., $\mathbf{X}_{II}(t_0) = \mathbf{X}^{1/2}$, $\mathbf{S}_{II}(t_0) = \mathbf{S}^{1/2}$ and $\mathbf{W}_{II}(t_0) = \mathbf{W}^{1/2}$. Again Roman numbers are omitted in the following. However, we do include a subscript 1 to denote that we are solving for the factors of \mathbf{u}_1 . For the scattering step (25b), determining the K , L and S -steps is straightforward and leads to

$$\dot{\mathbf{K}}_1(t) = -\Sigma_t(t)\mathbf{K}_1(t) + \psi(t_1)\mathbf{T}_M\Sigma(t)\mathbf{W}_1^{1/2}, \quad (29a)$$

$$\dot{\mathbf{L}}_1(t) = -\Sigma_t(t)\mathbf{L}_1(t) + \mathbf{X}_1^{1/2,T}\psi(t_1)\mathbf{T}_M\Sigma(t), \quad (29b)$$

$$\dot{\mathbf{S}}_1(t) = -\Sigma_t(t)\mathbf{S}_1(t) + \mathbf{X}_1^{1,T}\psi(t_1)\mathbf{T}_M\Sigma(t)\mathbf{W}_1^1. \quad (29c)$$

The time updated solution after streaming and scattering is given as $\mathbf{u}_1(t_1) = \mathbf{X}_1(t_1)\mathbf{S}_1(t_1)\mathbf{W}_1(t_1)^T$. When the directed S_N quadrature set has n_q nodes, computational costs are of $O(r \cdot n_q \cdot (n_x + m))$. Note that

since radiation therapy uses highly peaked particle beams as boundary conditions or source terms, only a limited number of directions needs to be resolved by the quadrature, i.e., n_q is expected to be small.

In the same manner, evolution equations for the factors of the solutions to the remaining moment equations (24c) and (24d) are derived. Here, the streaming update can be determined with (26), (27) and (28) (except for (24a), since for ψ we use a directed S_N method instead of a dynamical low-rank approximation). The scattering update for a general $\ell = 2, \dots, L$ which we denote by a subscript reads

$$\begin{aligned}\dot{\mathbf{K}}_\ell(t) &= -\Sigma_t(t)\mathbf{K}_\ell(t) + \mathbf{X}_{\ell-1}^1 \mathbf{S}_{\ell-1}^1 \mathbf{W}_{\ell-1}^{1,T} \Sigma(t) \mathbf{W}_\ell^{1/2}, \\ \dot{\mathbf{L}}_\ell(t) &= -\Sigma_t(t)\mathbf{L}_\ell(t) + \mathbf{X}_\ell^{1/2,T} \mathbf{X}_{\ell-1}^1 \mathbf{S}_{\ell-1}^1 \mathbf{W}_{\ell-1}^{1,T} \Sigma(t), \\ \dot{\mathbf{S}}_\ell(t) &= -\Sigma_t(t)\mathbf{S}_\ell(t) + \mathbf{X}_\ell^{1,T} \mathbf{X}_{\ell-1}^1 \mathbf{S}_{\ell-1}^1 \mathbf{W}_{\ell-1}^{1,T} \Sigma(t) \mathbf{W}_\ell^1.\end{aligned}$$

Computational costs are of $O(r^2 \cdot (n_x + m))$. Lastly, for the collided solution we perform a further splitting step. Omitting the subscript c , we have

$$\dot{\mathbf{u}}_I(t) = \mathbf{F}(t, \mathbf{u}_I(t)), \quad \mathbf{u}_I(t_0) = \mathbf{u}_1(t_0), \quad (31a)$$

$$\dot{\mathbf{u}}_{II}(t) = \mathbf{u}_L(t_1)\Sigma(t), \quad \mathbf{u}_{II}(t_0) = \mathbf{u}_I(t_1), \quad (31b)$$

$$\dot{\mathbf{u}}_{III}(t) = -\Sigma_t(t)\mathbf{u}_{III}(t) + \mathbf{u}_{III}(t)\Sigma(t), \quad \mathbf{u}_{III}(t_0) = \mathbf{u}_{II}(t_1). \quad (31c)$$

In this case, the K , S , L -equations for inscattering from \mathbf{u}_L , i.e., equation (31b) read (omitting Roman indices)

$$\dot{\mathbf{K}}_c(t) = \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \Sigma(t) \mathbf{W}_c^{1/2}, \quad (32a)$$

$$\dot{\mathbf{L}}_c(t) = \mathbf{X}_c^{1/2,T} \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \Sigma(t), \quad (32b)$$

$$\dot{\mathbf{S}}_c(t) = \mathbf{X}_c^{1,T} \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \Sigma(t) \mathbf{W}_c^1. \quad (32c)$$

For the in-scattering and out-scattering of the collided flux, i.e., equation (31c) we use the matrix projector-splitting integrator. Following [30], only the L -step needs to be computed and we are left with

$$\dot{\mathbf{L}}_c(t) = -\Sigma_t(t)\mathbf{L}_c(t) + \mathbf{L}_c(t)\Sigma(t). \quad (33)$$

The costs for the collided particles are again $O(r^2 \cdot (n_x + m))$. According to the derived steps, the scheme then consecutively updates the uncollided particles ψ , the factors of $\mathbf{u}_1, \dots, \mathbf{u}_L$ and lastly the factors of \mathbf{u}_c . In every step, the factors are first updated by a streaming step, followed by a scattering step. Lastly, when updating the factors of the collided flux, the additional L -step is performed to account for self-scattering. This procedure is repeated until a final time (or minimal energy) is reached. A flow chart to visualize the presented method is given in Figure 1.

4.2. Time (or energy) discretization

The presented equations still continuously depend on the pseudo-time (or energy) t . To treat stiff scattering terms, we use an implicit time update method for the scattering equations. The remainder uses explicit time discretizations. More specifically, we use implicit and explicit Euler time-discretizations in this work. Let us start with $\mathbf{X}_\ell^n = \mathbf{X}_\ell^0$, $\mathbf{S}_\ell^n = \mathbf{S}_\ell^0$ and $\mathbf{W}_\ell^n = \mathbf{W}_\ell^0$, where ℓ denotes the individual collision steps, i.e., $\ell \in \{1, \dots, L-1, L, c\}$. The streaming update is the same for all collision steps. Hence, when omitting a specific collision index, we obtain

$$\mathbf{K}^{1/2} = \mathbf{K}^0 + \Delta t \left(\widehat{\mathbf{L}}_x^{(2)} \mathbf{K}^0 \widehat{\mathbf{A}}_x^0 + \widehat{\mathbf{L}}_y^{(2)} \mathbf{K}^0 \widehat{\mathbf{A}}_y^0 + \widehat{\mathbf{L}}_x^{(1)} \mathbf{K}^0 |\widehat{\mathbf{A}}_x|^0 + \widehat{\mathbf{L}}_y^{(1)} \mathbf{K}^0 |\widehat{\mathbf{A}}_y|^0 \right), \quad \mathbf{X}^{1/2} \mathbf{R}_1 = \mathbf{K}^1, \quad (34a)$$

$$\mathbf{L}^{1/2} = \mathbf{L}^0 + \Delta t \left(\widehat{\mathbf{L}}_x^{(2),0} \mathbf{L}^0 \mathbf{A}_x^T + \widehat{\mathbf{L}}_y^{(2),0} \mathbf{L}^0 \mathbf{A}_y^T + \widehat{\mathbf{L}}_x^{(1),0} \mathbf{L}^0 |\mathbf{A}_x|^T + \widehat{\mathbf{L}}_y^{(1),0} \mathbf{L}^0 |\mathbf{A}_y|^T \right), \quad \mathbf{W}^{1/2} \mathbf{R}_2 = \mathbf{L}^{1,T}, \quad (34b)$$

$$\mathbf{S}^{1/2} = \widetilde{\mathbf{S}}^0 + \Delta t \left(\widehat{\mathbf{L}}_x^{(2),1/2} \widetilde{\mathbf{S}}^0 \widehat{\mathbf{A}}_x^{1/2} + \widehat{\mathbf{L}}_y^{(2),1/2} \widetilde{\mathbf{S}}^0 \widehat{\mathbf{A}}_y^{1/2} + \widehat{\mathbf{L}}_x^{(1),1/2} \widetilde{\mathbf{S}}^0 |\widehat{\mathbf{A}}_x|^{1/2} + \widehat{\mathbf{L}}_y^{(1),1/2} \widetilde{\mathbf{S}}^0 |\widehat{\mathbf{A}}_y|^{1/2} \right), \quad (34c)$$

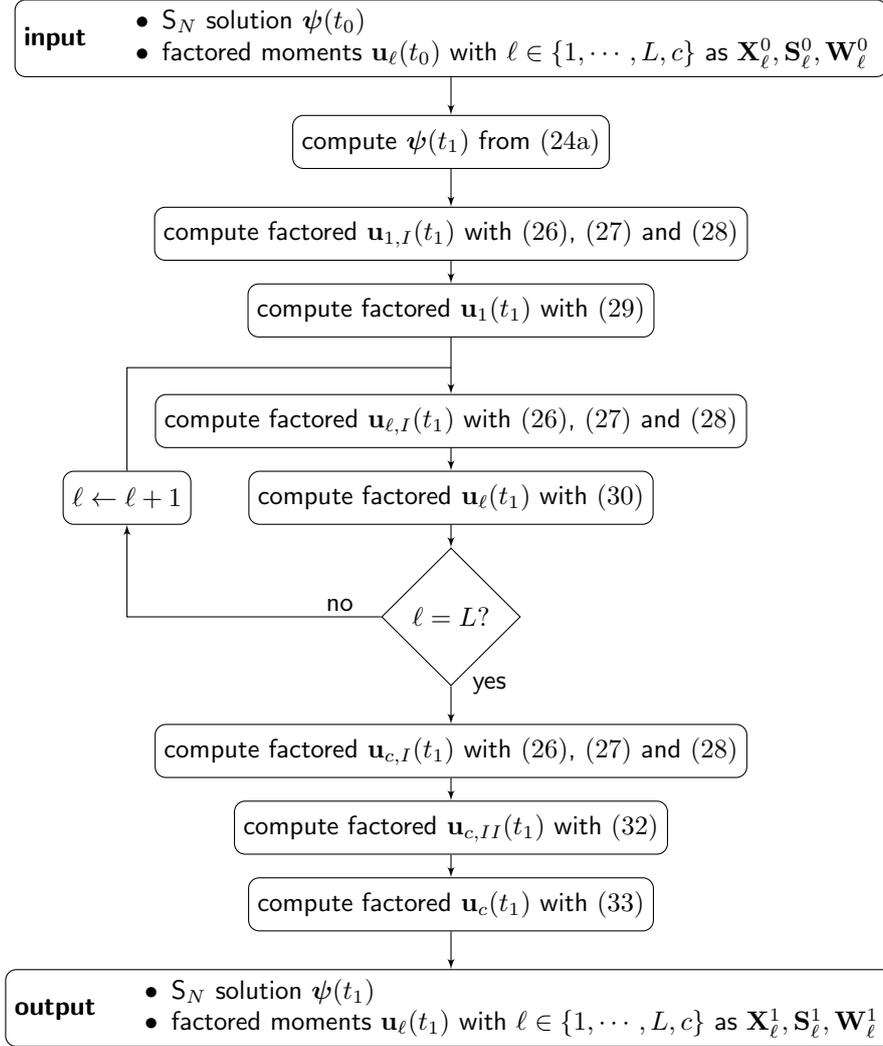


Figure 1: Flowchart of the presented method.

where $\tilde{\mathbf{S}}^0 = \mathbf{X}^{1/2,T} \mathbf{X}^0 \mathbf{S}^0 \mathbf{W}^{0,T} \mathbf{W}^{1/2}$ and flux matrices are computed before evaluating the right-hand side. The collision equations differ for the $\ell = 1$ and $\ell \in \{2, \dots, L\}$ collided fluxes as well as the collided flux. For $\ell = 1$, we have

$$\mathbf{K}_1^1 = \frac{1}{1 + \Delta t \Sigma(t_1)} \left(\mathbf{K}_1^{1/2} + \Delta t \boldsymbol{\psi}(t_1) \mathbf{T}_M \Sigma(t_1) \mathbf{W}_1^{1/2} \right), \quad \mathbf{X}_1^1 \mathbf{R}_1 = \mathbf{K}_1^1, \quad (35a)$$

$$\mathbf{L}_1^1 = \frac{1}{1 + \Delta t \Sigma(t_1)} \left(\mathbf{L}_1^{1/2} + \Delta t \mathbf{X}_1^{1/2,T} \boldsymbol{\psi}(t_1) \mathbf{T}_M \Sigma(t_1) \right), \quad \mathbf{W}_1^1 \mathbf{R}_2 = \mathbf{L}_1^{1,T}, \quad (35b)$$

$$\mathbf{S}_1^1 = \frac{1}{1 + \Delta t \Sigma(t_1)} \left(\tilde{\mathbf{S}}_1^{1/2} + \Delta t \mathbf{X}_1^{1,T} \boldsymbol{\psi}(t_1) \mathbf{T}_M \Sigma(t_1) \mathbf{W}_1^1 \right). \quad (35c)$$

For $\ell \in \{2, \dots, L\}$ we have

$$\mathbf{K}_\ell^1 = \frac{1}{1 + \Delta t \Sigma(t_1)} \left(\mathbf{K}_\ell^{1/2} + \Delta t \mathbf{X}_{\ell-1}^1 \mathbf{S}_{\ell-1}^1 \mathbf{W}_{\ell-1}^{1,T} \Sigma(t_1) \mathbf{W}_\ell^{1/2} \right), \quad \mathbf{X}_\ell^1 \mathbf{R}_1 = \mathbf{K}_\ell^1, \quad (36a)$$

$$\mathbf{L}_\ell^1 = \frac{1}{1 + \Delta t \Sigma(t_1)} \left(\mathbf{L}_\ell^{1/2} + \Delta t \mathbf{X}_\ell^{1/2,T} \mathbf{X}_{\ell-1}^1 \mathbf{S}_{\ell-1}^1 \mathbf{W}_{\ell-1}^{1,T} \Sigma(t_1) \right), \quad \mathbf{W}_\ell^1 \mathbf{R}_2 = \mathbf{L}_\ell^{1,T}, \quad (36b)$$

$$\mathbf{S}_\ell^1 = \frac{1}{1 + \Delta t \Sigma(t_1)} \left(\tilde{\mathbf{S}}_\ell^{1/2} + \Delta t \mathbf{X}_\ell^{1,T} \mathbf{X}_{\ell-1}^1 \mathbf{S}_{\ell-1}^1 \mathbf{W}_{\ell-1}^{1,T} \Sigma(t_1) \mathbf{W}_\ell^1 \right). \quad (36c)$$

The collided flux is then updated through

$$\mathbf{K}_c^1 = \mathbf{K}_c^{1/2} + \Delta t \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \Sigma(t_1) \mathbf{W}_c^{1/2}, \quad \mathbf{X}_c^1 \mathbf{R}_1 = \mathbf{K}_c^1, \quad (37a)$$

$$\mathbf{L}_c^1 = \mathbf{L}_c^{1/2} + \Delta t \mathbf{X}_c^{1/2,T} \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \Sigma(t_1), \quad \tilde{\mathbf{W}}_c^1 \mathbf{R}_2 = \mathbf{L}_c^{1,T}, \quad (37b)$$

$$\tilde{\mathbf{S}}_c^1 = \tilde{\mathbf{S}}_c^{1/2} + \Delta t \mathbf{X}_c^{1,T} \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \Sigma(t_1) \tilde{\mathbf{W}}_c^1, \quad (37c)$$

$$\tilde{\mathbf{L}}_c^1 = \tilde{\mathbf{S}}_c^1 \tilde{\mathbf{W}}_c^{1,T} (\mathbf{I} + \Sigma_t(t_1) \Delta t \mathbf{I} - \Delta t \Sigma(t_1))^{-1}, \quad \mathbf{W}_c^1 \mathbf{S}_c^{1,T} = \tilde{\mathbf{L}}_c^{1,T}. \quad (37d)$$

Note that since Σ is a diagonal matrix, the inversion in (37d) is given explicitly without having to solve a linear system of equations. The time updated solution is then given by $\mathbf{X}_\ell^{n+1} = \mathbf{X}_\ell^1$, $\mathbf{S}_\ell^{n+1} = \mathbf{S}_\ell^1$ and $\mathbf{W}_\ell^{n+1} = \mathbf{W}_\ell^1$, where $\ell \in \{1, \dots, L-1, L, c\}$.

Remark 1. *The proposed idea of multilevel DLRA can be applied in various settings with various strategies. The core ingredient is to write the solution as a sum of different contributions*

$$\mathbf{u}(t) = \mathbf{u}_1(t) + \mathbf{u}_2(t) + \dots + \mathbf{u}_L(t).$$

Strategies to write the solution as a sum of different components can be the use of telescoping identities, a split into symmetric and anti-symmetric solution contributions, a splitting of the original phase space, e.g. particles that move forward and backward and many more. In a second step, evolution equations for every component need to be derived. Third, every component \mathbf{u}_i is represented through a low-rank factorization and evolution equations for every factor are derived with DLRA.

5. L²-stability of the proposed scheme

The derived method is robust in that its time step restriction (or CFL number) does not depend on small singular values of the coefficient matrix or stiff terms arising in the scattering step. By the choice of the splitting steps, we ensure that this stability is achieved without having to invert matrices or solve a nonlinear problem, which is commonly the case for implicit time integration methods. To determine a suitable CFL condition, let us investigate the L²-stability of the proposed scheme, which follows the approach taken in [30]. In contrast to [30], the update equations include the inverse density, which will pose a challenge. We

first note that the inverse density can be pulled out of the stencil matrices. For this, we define the sparse diffusion stencil matrices without density dependence $\mathbf{T}_{x,y}^{(1)} \in \mathbb{R}^{n_x \times n_x}$ as

$$\begin{aligned} T_{x,\text{idx}(i,j),\text{idx}(i,j)}^{(1)} &= \frac{1}{\Delta x}, & T_{x,\text{idx}(i,j),\text{idx}(i\pm 1,j)}^{(1)} &= -\frac{1}{2\Delta x}, \\ T_{y,\text{idx}(i,j),\text{idx}(i,j)}^{(1)} &= \frac{1}{\Delta y}, & T_{y,\text{idx}(i,j),\text{idx}(i,j\pm 1)}^{(1)} &= -\frac{1}{2\Delta y}, \end{aligned}$$

as well as the sparse advection stencil matrices without density dependence $\mathbf{T}_{x,y}^{(2)} \in \mathbb{R}^{n_x \times n_x}$ as

$$T_{x,\text{idx}(i,j),\text{idx}(i\pm 1,j)}^{(2)} = \pm \frac{1}{2\Delta x}, \quad T_{y,\text{idx}(i,j),\text{idx}(i,j\pm 1)}^{(2)} = \pm \frac{1}{2\Delta y}.$$

With $\boldsymbol{\rho}^{-1} = \text{diag}\left(\left(\rho_{\text{idx}(i,j)}^{-1}\right)_{i,j=1}^{N_x, N_y}\right)$, we have that $\mathbf{L}_{x,y}^{(1,2)} = \mathbf{T}_{x,y}^{(1,2)} \boldsymbol{\rho}^{-1}$. Following [30], we pursue a von Neumann-like approach and investigate how the scheme amplifies and dampens certain Fourier modes. Let us store these modes in a matrix $\mathbf{E} \in \mathbb{C}^{n_x \times n_x}$ with entries

$$E_{\text{idx}(\ell,k),\text{idx}(\alpha,\beta)} = \sqrt{\Delta x \Delta y} \exp(i\alpha\pi x_\ell) \exp(i\beta\pi y_k),$$

where $i \in \mathbb{C}$ denotes the imaginary unit. This matrix has several properties. First, it is orthonormal, i.e., $\mathbf{E}\mathbf{E}^H = \mathbf{E}^H\mathbf{E} = \mathbf{I}$, where an uppercase H denotes the complex transpose. Second, the matrix \mathbf{E} applied to the spatial flux matrices diagonalizes the scheme:

$$\mathbf{T}_{x,y}^{(1,2)} \mathbf{E} = \mathbf{E} \mathbf{D}_{x,y}^{(1,2)}. \quad (38)$$

The diagonal matrices $\mathbf{D}_{x,y}^{(1,2)} \in \mathbb{R}^{n_x \times n_x}$ have entries

$$\begin{aligned} D_{x,\text{idx}(\alpha,\beta),\text{idx}(\alpha',\beta')}^{(1)} &= \frac{1}{2\Delta x} \left(e^{i\alpha\pi\Delta x} - 2 + e^{-i\alpha\pi\Delta x} \right) \delta_{\alpha\alpha'} \delta_{\beta\beta'} = \frac{1}{\Delta x} (\cos(\alpha\pi\Delta x) - 1) \delta_{\alpha\alpha'} \delta_{\beta\beta'}, \\ D_{y,\text{idx}(\alpha,\beta),\text{idx}(\alpha',\beta')}^{(1)} &= \frac{1}{2\Delta y} \left(e^{i\alpha\pi\Delta y} - 2 + e^{-i\alpha\pi\Delta y} \right) \delta_{\alpha\alpha'} \delta_{\beta\beta'} = \frac{1}{\Delta y} (\cos(\alpha\pi\Delta y) - 1) \delta_{\alpha\alpha'} \delta_{\beta\beta'}, \\ D_{x,\text{idx}(\alpha,\beta),\text{idx}(\alpha',\beta')}^{(2)} &= \frac{1}{2\Delta x} \left(e^{i\alpha\pi\Delta x} - e^{-i\alpha\pi\Delta x} \right) \delta_{\alpha\alpha'} \delta_{\beta\beta'} = -\frac{i}{\Delta x} \sin(\alpha\pi\Delta x) \delta_{\alpha\alpha'} \delta_{\beta\beta'}, \\ D_{y,\text{idx}(\alpha,\beta),\text{idx}(\alpha',\beta')}^{(2)} &= \frac{1}{2\Delta y} \left(e^{i\alpha\pi\Delta y} - e^{-i\alpha\pi\Delta y} \right) \delta_{\alpha\alpha'} \delta_{\beta\beta'} = -\frac{i}{\Delta y} \sin(\alpha\pi\Delta y) \delta_{\alpha\alpha'} \delta_{\beta\beta'}. \end{aligned}$$

With these tools at hand, we can prove stability. Let us start with the streaming steps:

Lemma 1. *Assume that the CFL condition*

$$\frac{\lambda_{\max}(\mathbf{A}_{x,y}) \Delta t}{\rho_{\min}} \frac{\Delta t}{\Delta x} \leq \frac{1}{2} \quad (39)$$

holds true. Then, the streaming scheme (34) is L^2 -stable, i.e.,

$$\|\mathbf{X}^{1/2} \mathbf{S}^{1/2} \mathbf{W}^{1/2,T}\|_F \leq \|\mathbf{X}^0 \mathbf{S}^0 \mathbf{W}^{0,T}\|_F.$$

Proof. First, we include the identity $\mathbf{E}\mathbf{E}^H$ inside the spatial flux matrices of the S -step (34c):

$$\begin{aligned} \widehat{\mathbf{L}}_{x,y}^{(1),1/2} &= \mathbf{X}^{1/2,T} \mathbf{T}_{x,y}^{(1)} \mathbf{E}\mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{X}^{1/2} \stackrel{(38)}{=} \mathbf{X}^{1/2,T} \mathbf{E} \mathbf{D}_{x,y}^{(1)} \mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{X}^{1/2}, \\ \widehat{\mathbf{L}}_{x,y}^{(2),1/2} &= \mathbf{X}^{1/2,T} \mathbf{T}_{x,y}^{(2)} \mathbf{E}\mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{X}^{1/2} \stackrel{(38)}{=} \mathbf{X}^{1/2,T} \mathbf{E} \mathbf{D}_{x,y}^{(2)} \mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{X}^{1/2}. \end{aligned}$$

Then, with $\tilde{\rho}^{-1} := \mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{E}$ and $\mathbf{u} = \mathbf{X}^{1/2} \tilde{\mathbf{S}}^0 \mathbf{W}^{1/2, T}$, we have

$$\begin{aligned} \mathbf{S}^{1/2} &= \tilde{\mathbf{S}}^0 + \Delta t \left(\widehat{\mathbf{L}}_x^{(2), 1/2} \tilde{\mathbf{S}}^0 \widehat{\mathbf{A}}_x^{1/2} + \widehat{\mathbf{L}}_y^{(2), 1/2} \tilde{\mathbf{S}}^0 \widehat{\mathbf{A}}_y^{1/2} + \widehat{\mathbf{L}}_x^{(1), 1/2} \tilde{\mathbf{S}}^0 |\widehat{\mathbf{A}}_x|^{1/2} + \widehat{\mathbf{L}}_y^{(1), 1/2} \tilde{\mathbf{S}}^0 |\widehat{\mathbf{A}}_y|^{1/2} \right) \\ &= \mathbf{X}^{1/2, T} \mathbf{E} \mathbf{E}^H \mathbf{u} \mathbf{W}^{1/2} + \Delta t \mathbf{X}^{1/2, T} \mathbf{E} \left(\mathbf{D}_x^{(2)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} \mathbf{A}_x + \mathbf{D}_x^{(1)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} |\mathbf{A}_x| \right) \mathbf{W}^{1/2} \\ &\quad + \Delta t \mathbf{X}^{1/2, T} \mathbf{E} \left(\mathbf{D}_y^{(2)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} \mathbf{A}_y + \mathbf{D}_y^{(1)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} |\mathbf{A}_y| \right) \mathbf{W}^{1/2}. \end{aligned}$$

Taking the norm and noting that for the spectral norm we have $\|\mathbf{X}^{1/2, T} \mathbf{E}\| = \|\mathbf{W}^{1/2}\| = 1$ yields

$$\begin{aligned} \|\mathbf{S}^{1/2}\|_F &\leq \left\| \mathbf{E}^H \mathbf{u} + \Delta t \left(\mathbf{D}_x^{(2)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} \mathbf{A}_x + \mathbf{D}_x^{(1)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} |\mathbf{A}_x| + \mathbf{D}_y^{(2)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} \mathbf{A}_y + \mathbf{D}_y^{(1)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} |\mathbf{A}_y| \right) \right\|_F \\ &\leq \left\| \frac{1}{2} \mathbf{E}^H \mathbf{u} + \Delta t \left(\mathbf{D}_x^{(2)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} \mathbf{A}_x + \mathbf{D}_x^{(1)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} |\mathbf{A}_x| \right) \right\|_F \end{aligned} \quad (40a)$$

$$+ \left\| \frac{1}{2} \mathbf{E}^H \mathbf{u} + \Delta t \left(\mathbf{D}_y^{(2)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} \mathbf{A}_y + \mathbf{D}_y^{(1)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u} |\mathbf{A}_y| \right) \right\|_F. \quad (40b)$$

Now we investigate the terms (40a) and (40b) individually. Recall that we have $\mathbf{A}_{x,y} = \mathbf{V}_{x,y} \boldsymbol{\Lambda}_{x,y} \mathbf{V}_{x,y}^T$. Then for (40a), which we denote by $\|\mathbf{e}\|_F$, we define $\mathbf{w} := \mathbf{E}^H \mathbf{u} \mathbf{V}_x$ which gives

$$\begin{aligned} \|\mathbf{e}\|_F &= \left\| \frac{1}{2} \mathbf{w} \mathbf{V}_x^T + \Delta t \left(\mathbf{D}_x^{(2)} \tilde{\rho}^{-1} \mathbf{w} \boldsymbol{\Lambda}_x + \mathbf{D}_x^{(1)} \tilde{\rho}^{-1} \mathbf{w} |\boldsymbol{\Lambda}_x| \right) \mathbf{V}_x^T \right\|_F \\ &\leq \left\| \frac{1}{2} \mathbf{w} + \Delta t \left(\mathbf{D}_x^{(2)} \tilde{\rho}^{-1} \mathbf{w} \boldsymbol{\Lambda}_x + \mathbf{D}_x^{(1)} \tilde{\rho}^{-1} \mathbf{w} |\boldsymbol{\Lambda}_x| \right) \right\|_F. \end{aligned}$$

Note that with $\mathbf{e}_k := (e_{jk})_{j=1}^{n_x}$ and $\mathbf{w}_k := (w_{jk})_{j=1}^{n_x}$, we have

$$\begin{aligned} \|\mathbf{e}\|_F^2 &= \sum_{k=1}^m \|\mathbf{e}_k\|_2^2 = \sum_{k=1}^m \left\| \frac{1}{2} \mathbf{w}_k + \Delta t \left(\mathbf{D}_x^{(2)} \tilde{\rho}^{-1} \mathbf{w}_k \lambda_k^x + \mathbf{D}_x^{(1)} \tilde{\rho}^{-1} \mathbf{w}_k |\lambda_k^x| \right) \right\|_2^2 \\ &\leq \sum_{k=1}^m \left\| \frac{1}{2} \mathbf{I} + \Delta t \left(\lambda_k^x \mathbf{D}_x^{(2)} + |\lambda_k^x| \cdot \mathbf{D}_x^{(1)} \right) \tilde{\rho}^{-1} \right\|_2^2 \cdot \|\mathbf{w}_k\|_2^2, \end{aligned}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. Now, for the spectral norm in the above expression we have the upper bound

$$\left\| \frac{1}{2} \mathbf{I} + \Delta t \left(\lambda_k^x \mathbf{D}_x^{(2)} + |\lambda_k^x| \cdot \mathbf{D}_x^{(1)} \right) \tilde{\rho}^{-1} \right\| \leq \left| \frac{1}{2} + \Delta t \lambda_{\max} \left(\lambda_k^x \mathbf{D}_x^{(2)} + |\lambda_k^x| \cdot \mathbf{D}_x^{(1)} \right) \cdot \lambda_{\max}(\tilde{\rho}^{-1}) \right|.$$

We note that $\lambda_{\max}(\tilde{\rho}^{-1}) = \min_j \rho_j^{-1} =: \rho_{\min}^{-1}$. With $\nu := \frac{\max_k |\lambda_k^x| \Delta t}{\rho_{\min} \Delta x}$ we have

$$\begin{aligned} \|\mathbf{e}\|_F &\leq \max_{\alpha} \left| \frac{1}{2} + \nu (\cos(\alpha \pi \Delta x) - 1) - i \nu \sin(\alpha \pi \Delta x) \right| \cdot \|\mathbf{w}\| \\ &= \max_{\alpha} \sqrt{\frac{1}{4} + \nu (\cos(\alpha \pi \Delta x) - 1) + \nu^2 (\cos(\alpha \pi \Delta x) - 1)^2 + \nu^2 \sin^2(\alpha \pi \Delta x)} \cdot \|\mathbf{w}\| \\ &= \max_{\alpha} \sqrt{\frac{1}{4} + \nu (1 - 2\nu) (\cos(\alpha \pi \Delta x) - 1)} \cdot \|\mathbf{w}\|. \end{aligned}$$

To obtain stability, we need $\|\mathbf{e}\|_F \leq \|\mathbf{w}\|/2$, i.e., $\nu \leq \frac{1}{2}$. In the same way, we can get an estimate for (40b), which yields $\frac{\max_k |\lambda_k^y| \Delta t}{\rho_{\min} \Delta y} \leq \frac{1}{2}$. With

$$\|\mathbf{w}\|_F = \|\mathbf{E}^H \mathbf{u} \mathbf{V}_{x,y}\|_F \leq \|\mathbf{u}\|_F = \|\mathbf{X}^{1/2} \tilde{\mathbf{S}}^0 \mathbf{W}^{1/2, T}\|_F \leq \|\mathbf{S}^0\|_F,$$

we know that $\|\mathbf{S}^{1/2}\|_F \leq \|\mathbf{S}^0\|_F$, which with Parseval's identity proves the theorem. \square

Theorem 1. Assume that the CFL conditions (39) and

$$\max_k \frac{1}{1 + \Delta t \Sigma_t(t) - \Delta t \Sigma_{kk}(t)} \leq 1 \quad (41)$$

hold true for all pseudo-times $t \in [0, T]$. Then, the scheme is L^2 -stable, in the sense that with $\mathbf{u}_c^1 := \mathbf{X}_c^1 \mathbf{S}_c^1 \mathbf{W}_c^{1,T}$ and $\mathbf{u}_\ell^1 := \mathbf{X}_\ell^1 \mathbf{S}_\ell^1 \mathbf{W}_\ell^{1,T}$ we have

$$\|\mathbf{u}_c^1\|_F + \sum_{\ell=1}^L \|\mathbf{u}_\ell^1\|_F + \|\boldsymbol{\psi}(t_1)\|_F \leq \|\mathbf{u}_c^0\|_F + \sum_{\ell=1}^L \|\mathbf{u}_\ell^0\|_F + \|\boldsymbol{\psi}(t_0)\|_F .$$

Proof. We start with the collided equations. The last two steps of (37) read

$$\begin{aligned} \bar{\mathbf{S}}_c^1 &= \tilde{\mathbf{S}}_c^{1/2} + \Delta t \mathbf{X}_c^{1,T} \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \boldsymbol{\Sigma}(t_1) \tilde{\mathbf{W}}_c^1 , \\ \tilde{\mathbf{L}}_c^1 &= \bar{\mathbf{S}}_c^1 \tilde{\mathbf{W}}_c^{1,T} (\mathbf{I} + \Sigma_t \Delta t \mathbf{I} - \Delta t \boldsymbol{\Sigma})^{-1} . \end{aligned}$$

Writing this as a single expression gives

$$\tilde{\mathbf{L}}_c^1 = \left(\tilde{\mathbf{S}}_c^{1/2} + \Delta t \mathbf{X}_c^{1,T} \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \boldsymbol{\Sigma}(t_1) \tilde{\mathbf{W}}_c^1 \right) \tilde{\mathbf{W}}_c^{1,T} (\mathbf{I} + \Sigma_t \Delta t \mathbf{I} - \Delta t \boldsymbol{\Sigma})^{-1} .$$

We take the norm of the above expression and note that since $\tilde{\mathbf{L}}_c^1 = \mathbf{S}_c^1 \mathbf{W}_c^{1,T}$

$$\|\mathbf{S}_c^1\|_F \leq \max_k \frac{1}{1 + \Delta t \Sigma_t - \Delta t \Sigma_{kk}} \|\tilde{\mathbf{S}}_c^{1/2}\|_F + \max_k \frac{\Delta t \Sigma_{kk}}{1 + \Delta t \Sigma_t - \Delta t \Sigma_{kk}} \|\mathbf{S}_L^1\|_F .$$

Note that we used

$$\|\boldsymbol{\Sigma}\| \cdot \|(\mathbf{I} + \Sigma_t \Delta t \mathbf{I} - \Delta t \boldsymbol{\Sigma})^{-1}\| = \max_k \Sigma_{kk} \cdot \max_k \frac{1}{1 + \Delta t \Sigma_t - \Delta t \Sigma_{kk}} = \max_k \frac{\Sigma_{kk}}{1 + \Delta t \Sigma_t - \Delta t \Sigma_{kk}} .$$

Adding $\|\mathbf{S}_L^1\|_F$ to both sides and noting that by Lemma 1, we have $\|\tilde{\mathbf{S}}_c^{1/2}\|_F \leq \|\mathbf{S}_c^{1/2}\|_F \leq \|\mathbf{S}_c^0\|_F$ leads to

$$\|\mathbf{S}_c^1\|_F + \|\mathbf{S}_L^1\|_F \leq \max_k \frac{1}{1 + \Delta t \Sigma_t - \Delta t \Sigma_{kk}} \|\mathbf{S}_c^0\|_F + \max_k \frac{1 + \Delta t \Sigma_t}{1 + \Delta t \Sigma_t - \Delta t \Sigma_{kk}} \|\mathbf{S}_L^1\|_F .$$

By (36c) we have

$$\begin{aligned} \|\mathbf{S}_L^1\|_F &\leq \frac{1}{1 + \Delta t \Sigma_t} \left(\|\tilde{\mathbf{S}}_L^{1/2}\|_F + \Delta t \max_k \Sigma_{kk} \|\mathbf{S}_{L-1}^1\|_F \right) \\ &\leq \frac{1}{1 + \Delta t \Sigma_t} \left(\|\mathbf{S}_L^0\|_F + \Delta t \max_k \Sigma_{kk} \|\mathbf{S}_{L-1}^1\|_F \right) . \end{aligned}$$

Using the chosen CFL condition gives

$$\|\mathbf{S}_c^1\|_F + \|\mathbf{S}_L^1\|_F \leq \|\mathbf{S}_c^0\|_F + \|\mathbf{S}_L^0\|_F + \max_k \frac{\Delta t \Sigma_{kk}}{1 + \Delta t \Sigma_t - \Delta t \Sigma_{kk}} \|\mathbf{S}_{L-1}^1\|_F .$$

Recursively continuing this process until $\ell = 1$ gives

$$\|\mathbf{S}_c^1\|_F + \sum_{\ell=1}^L \|\mathbf{S}_\ell^1\|_F \leq \|\mathbf{S}_c^0\|_F + \sum_{\ell=1}^L \|\mathbf{S}_\ell^0\|_F + \max_k \frac{\Delta t \Sigma_{kk}}{1 + \Delta t \Sigma_t - \Delta t \Sigma_{kk}} \|\boldsymbol{\psi}(t_1)\|_F .$$

Adding $\|\boldsymbol{\psi}(t_1)\|_F$ on both sides, noting that (when choosing an implicit time discretization for scattering of uncollided particles)

$$\|\boldsymbol{\psi}(t_1)\|_F = \frac{1}{1 + \Delta t \Sigma_t} \|\boldsymbol{\psi}(t_0)\|_F$$

and the use of Parseval's identity proves the theorem. \square

Remark 2. Commonly, the CFL condition of the scattering step is fulfilled automatically, since $\Sigma_{kk} \leq \Sigma_t$ for all k . Therefore, a time step restriction is only imposed by the streaming update through (39).

6. Extension to rank adaptivity

In a last step, we discuss the extension of the proposed scheme to the rank adaptive integrator of [8] and how it simplifies for a forward Euler time discretization. The core ingredient of this method is to extend the time updated basis with the basis at time t_0 . Hence, for the streaming step, the updated basis becomes $\widehat{\mathbf{X}}^{1/2} = [\mathbf{X}^0, \overline{\mathbf{X}}^{1/2}]$, where $\overline{\mathbf{X}}^{1/2}$ is chosen such that the column range of $\widehat{\mathbf{X}}^{1/2}$ contains $\mathbf{K}(t_1)$ from the streaming K -step and the basis is orthonormal, i.e., $\mathbf{X}^{0,T} \overline{\mathbf{X}}^{1/2} = \mathbf{0}$ and $\overline{\mathbf{X}}^{1/2,T} \overline{\mathbf{X}}^{1/2} = \mathbf{I}$. Hence the matrix to compute the initial condition of the S -step reads $\widehat{\mathbf{M}} = \widehat{\mathbf{X}}^{1/2,T} \mathbf{X}^0 = [\mathbf{I}, \mathbf{0}]^T \in \mathbb{R}^{2r_0 \times r_0}$, where r_0 is the rank at time t_0 . In the same manner, we have $\widehat{\mathbf{N}} = [\mathbf{I}, \mathbf{0}]^T \in \mathbb{R}^{2r_0 \times r_0}$. Hence, as pointed out in [8], we have $\widehat{\mathbf{X}}^{1/2} \widehat{\mathbf{S}}(t_0) \widehat{\mathbf{W}}^{1/2,T} = \mathbf{X}^0 \mathbf{S}(t_0) \mathbf{W}^{0,T} \in \mathcal{M}_{r_0}$. The fact that the initial condition of the S -step is of rank r_0 can be used to reduce computational costs when using an explicit Euler step. The S -step of the rank adaptive integrator for the streaming step then reads

$$\widehat{\mathbf{S}}^{1/2} = [\mathbf{I}, \mathbf{0}]^T \mathbf{S}^0 [\mathbf{I}, \mathbf{0}] + \Delta t \left(\overline{\mathbf{L}}_x^{(2),1/2} \mathbf{S}^0 \overline{\mathbf{A}}_x^{-1/2} + \overline{\mathbf{L}}_y^{(2),1/2} \mathbf{S}^0 \overline{\mathbf{A}}_y^{-1/2} + \overline{\mathbf{L}}_x^{(1),1/2} \mathbf{S}^0 |\overline{\mathbf{A}}_x|^{1/2} + \overline{\mathbf{L}}_y^{(1),1/2} \mathbf{S}^0 |\overline{\mathbf{A}}_y|^{1/2} \right),$$

where the flux matrices are given by

$$\begin{aligned} \overline{\mathbf{A}}_{x,y}^{-1/2} &:= \mathbf{W}^{0,T} \mathbf{A}_{x,y}^T \widehat{\mathbf{W}}^{1/2} \in \mathbb{R}^{r_0 \times 2r_0}, & |\overline{\mathbf{A}}_{x,y}|^{1/2} &:= \mathbf{W}^{0,T} |\mathbf{A}_{x,y}|^T \widehat{\mathbf{W}}^{1/2} \in \mathbb{R}^{r_0 \times 2r_0}, \\ \overline{\mathbf{L}}_{x,y}^{(2),1/2} &:= \widehat{\mathbf{X}}^{1/2,T} \mathbf{L}_{x,y}^{(2)} \mathbf{X}^0 \in \mathbb{R}^{2r_0 \times r_0}, & \overline{\mathbf{L}}_{x,y}^{(1),1/2} &:= \widehat{\mathbf{X}}^{1/2,T} \mathbf{L}_{x,y}^{(1)} \mathbf{X}^0 \in \mathbb{R}^{2r_0 \times r_0}. \end{aligned}$$

Hence, the flux matrices that need to be computed have $2r_0^2$ entries. When using more general time integration schemes to solve the S -step for the rank adaptive integrator, the flux matrices have $4r_0^2$ entries. We determine the solution factors after the streaming update $\mathbf{X}^{1/2}$, $\mathbf{S}^{1/2}$ and $\mathbf{W}^{1/2}$ as well as the rank $r_{1/2}$ through the truncation step of the rank adaptive integrator. For components $\ell \in \{1, 2, \dots, L, c\}$, the scattering steps have modified S -step equations

$$\begin{aligned} \widehat{\mathbf{S}}_1^1 &= \frac{1}{1 + \Delta t \Sigma(t_1)} \left([\mathbf{I}, \mathbf{0}]^T \mathbf{S}_1^{1/2} [\mathbf{I}, \mathbf{0}] + \Delta t \widehat{\mathbf{X}}_1^{1,T} \boldsymbol{\psi}(t_1) \mathbf{T}_M \Sigma(t_1) \widehat{\mathbf{W}}_1^1 \right), \\ \widehat{\mathbf{S}}_\ell^1 &= \frac{1}{1 + \Delta t \Sigma(t_1)} \left([\mathbf{I}, \mathbf{0}]^T \mathbf{S}_\ell^{1/2} [\mathbf{I}, \mathbf{0}] + \Delta t \widehat{\mathbf{X}}_\ell^{1,T} \mathbf{X}_{\ell-1}^1 \mathbf{S}_{\ell-1}^1 \mathbf{W}_{\ell-1}^{1,T} \Sigma(t_1) \widehat{\mathbf{W}}_\ell^1 \right), \quad \text{for } \ell = 2, \dots, L, \\ \widehat{\mathbf{S}}_c^1 &= [\mathbf{I}, \mathbf{0}]^T \mathbf{S}_c^{1/2} [\mathbf{I}, \mathbf{0}] + \Delta t \widehat{\mathbf{X}}_c^{1,T} \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1,T} \Sigma(t_1) \widehat{\mathbf{W}}_c^1. \end{aligned}$$

For $\ell \in \{1, \dots, L, c\}$ we use $\widehat{\mathbf{X}}_\ell^1 = [\mathbf{X}_\ell^{1/2}, \overline{\mathbf{X}}_\ell^1]$, where $\overline{\mathbf{X}}_\ell^1$ is chosen such that the column range of $\widehat{\mathbf{X}}_\ell^1$ contains $\mathbf{K}_\ell(t_1)$ from the scattering K -step and the basis is orthonormal, i.e., $\mathbf{X}_\ell^{1/2,T} \overline{\mathbf{X}}_\ell^1 = \mathbf{0}$ and $\overline{\mathbf{X}}_\ell^{1,T} \overline{\mathbf{X}}_\ell^1 = \mathbf{I}$. The directional basis $\widehat{\mathbf{W}}_\ell^1$ is defined analogously. Note that for the scattered particles, we need to do a final L -step (37d) after having updated the coefficient. In this case, the truncation step of the rank adaptive integrator yields $\widehat{\mathbf{W}}_c^1$ and $\widehat{\mathbf{S}}_c^1$. Since (37d) is constructed through the fixed-rank projector–splitting integrator, it will not modify the rank.

Besides allowing for a dynamic choice of the rank, the rank adaptive integrator remains L^2 stable.

Lemma 2. *Assume that the CFL condition (39) holds true. Then, the streaming scheme of the rank adaptive integrator is L^2 -stable, i.e.,*

$$\|\mathbf{X}^{1/2} \mathbf{S}^{1/2} \mathbf{W}^{1/2,T}\|_F \leq \|\mathbf{X}^0 \mathbf{S}^0 \mathbf{W}^{0,T}\|_F. \quad (42)$$

Proof. The proof follows the steps taken in Lemmas 1 and Theorem 1. We begin with the streaming part. Using the Fourier matrix \mathbf{E} , the spatial flux matrices can be written as

$$\begin{aligned} \overline{\mathbf{L}}_{x,y}^{(1),1/2} &= \widehat{\mathbf{X}}^{1/2,T} \mathbf{T}_{x,y}^{(1)} \mathbf{E} \mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{X}^0 = \widehat{\mathbf{X}}^{1/2,T} \mathbf{E} \mathbf{D}_{x,y}^{(1)} \mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{X}^0, \\ \overline{\mathbf{L}}_{x,y}^{(2),1/2} &= \widehat{\mathbf{X}}^{1/2,T} \mathbf{T}_{x,y}^{(2)} \mathbf{E} \mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{X}^0 = \widehat{\mathbf{X}}^{1/2,T} \mathbf{E} \mathbf{D}_{x,y}^{(2)} \mathbf{E}^H \boldsymbol{\rho}^{-1} \mathbf{X}^0. \end{aligned}$$

With $\tilde{\rho}^{-1} := \mathbf{E}^H \rho^{-1} \mathbf{E}$ and $\mathbf{u}_a = \widehat{\mathbf{X}}^{1/2} \widetilde{\mathbf{S}}^0 \widehat{\mathbf{W}}^{1/2, T} = \mathbf{X}^0 \mathbf{S}^0 \mathbf{W}^{0, T}$, we have

$$\begin{aligned} \widehat{\mathbf{S}}^{1/2} &= \widehat{\mathbf{X}}^{1/2, T} \mathbf{u}_a \widehat{\mathbf{W}}^{1/2} + \Delta t \widehat{\mathbf{X}}^{1/2, T} \mathbf{E} \left(\mathbf{D}_x^{(2)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u}_a \mathbf{A}_x + \mathbf{D}_x^{(1)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u}_a | \mathbf{A}_x \right) \widehat{\mathbf{W}}^{1/2} \\ &\quad + \Delta t \widehat{\mathbf{X}}^{1/2, T} \mathbf{E} \left(\mathbf{D}_y^{(2)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u}_a \mathbf{A}_y + \mathbf{D}_y^{(1)} \tilde{\rho}^{-1} \mathbf{E}^H \mathbf{u}_a | \mathbf{A}_y \right) \widehat{\mathbf{W}}^{1/2}. \end{aligned}$$

Since for the spectral norm we have $\|\widehat{\mathbf{X}}^{1/2, T} \mathbf{E}\| = \|\widehat{\mathbf{W}}^{1/2}\| = 1$, taking norms yields the inequality (40) and the remainder of the proof follows as in Theorem 1. Note that instead of $\mathbf{u} = \mathbf{X}^{1/2} \widetilde{\mathbf{S}}^0 \mathbf{W}^{1/2, T}$ we now use \mathbf{u}_a . However, this does not impact the derivation, since $\|\mathbf{u}_a\|_F = \|\widehat{\mathbf{X}}^{1/2} \widetilde{\mathbf{S}}^0 \widehat{\mathbf{W}}^{1/2, T}\|_F = \|\mathbf{S}^0\|_F$. Finally, we note that the truncation step of the rank adaptive integrator does not increase the norm of the solution, in fact

$$\|\mathbf{S}^{1/2}\| \leq \|\widehat{\mathbf{S}}^{1/2}\| \leq \|\mathbf{S}^0\|.$$

□

Lastly, we include scattering contributions.

Theorem 2. *Assume that the CFL conditions (39) and (41) hold true for all pseudo-times $t \in [0, T]$. Then, the scheme is L^2 -stable, in the sense that with $\mathbf{u}_c^1 := \mathbf{X}_c^1 \mathbf{S}_c^1 \mathbf{W}_c^{1, T}$ and $\mathbf{u}_\ell^1 := \mathbf{X}_\ell^1 \mathbf{S}_\ell^1 \mathbf{W}_\ell^{1, T}$ we have*

$$\|\mathbf{u}_c^1\|_F + \sum_{\ell=1}^L \|\mathbf{u}_\ell^1\|_F + \|\boldsymbol{\psi}(t_1)\|_F \leq \|\mathbf{u}_c^0\|_F + \sum_{\ell=1}^L \|\mathbf{u}_\ell^0\|_F + \|\boldsymbol{\psi}(t_0)\|_F.$$

Proof. The proof is essentially that of Theorem 1. The only difference is that we extend and truncate the basis and coefficient matrices. However, since both of these operations do not increase the Frobenius norm of the solution, the stability property is not violated by rank adaptivity. To make this observation more rigorous, we start with the collided equations. The last two steps of (37) read

$$\begin{aligned} \widehat{\mathbf{S}}_c^1 &= [\mathbf{I}, \mathbf{0}]^T \mathbf{S}_c^{1/2} [\mathbf{I}, \mathbf{0}] + \Delta t \widehat{\mathbf{X}}_c^{1, T} \mathbf{X}_L^1 \mathbf{S}_L^1 \mathbf{W}_L^{1, T} \boldsymbol{\Sigma}(t_1) \widehat{\mathbf{W}}_c^1, \\ \widetilde{\mathbf{L}}_c^1 &= \widetilde{\mathbf{S}}_c^1 \widetilde{\mathbf{W}}_c^{1, T} (\mathbf{I} + \boldsymbol{\Sigma}_t \Delta t \mathbf{I} - \Delta t \boldsymbol{\Sigma})^{-1}. \end{aligned}$$

Note that $\|\widetilde{\mathbf{L}}_c^1\|_F = \|\mathbf{S}_c^1\|_F$ and $\|\widetilde{\mathbf{S}}_c^1\|_F \leq \|\widehat{\mathbf{S}}_c^1\|_F$. Hence, taking norms of the above equations yields

$$\begin{aligned} \|\widehat{\mathbf{S}}_c^1\|_F &\leq \|\mathbf{S}_c^{1/2}\|_F + \Delta t \max_k \boldsymbol{\Sigma}_{kk} \|\mathbf{S}_L^1\|_F, \\ \|\mathbf{S}_c^1\|_F &\leq \max_k \frac{1}{1 + \boldsymbol{\Sigma}_t \Delta t - \Delta t \boldsymbol{\Sigma}_{kk}} \|\widehat{\mathbf{S}}_c^1\|_F = \frac{1}{1 + \boldsymbol{\Sigma}_t \Delta t - \Delta t \max_k \boldsymbol{\Sigma}_{kk}} \|\widehat{\mathbf{S}}_c^1\|_F, \end{aligned}$$

where we used that $\max_k |\boldsymbol{\Sigma}_{kk}| = \max_k \boldsymbol{\Sigma}_{kk}$. Written as a single expression, we have

$$\|\mathbf{S}_c^1\|_F \leq \max_k \frac{1}{1 + \boldsymbol{\Sigma}_t \Delta t - \Delta t \boldsymbol{\Sigma}_{kk}} \left(\|\mathbf{S}_c^{1/2}\|_F + \Delta t \max_k \boldsymbol{\Sigma}_{kk} \|\mathbf{S}_L^1\|_F \right) \quad (43)$$

$$\leq \|\mathbf{S}_c^{1/2}\|_F + \max_k \frac{\Delta t \boldsymbol{\Sigma}_{kk}}{1 + \boldsymbol{\Sigma}_t \Delta t - \Delta t \boldsymbol{\Sigma}_{kk}} \|\mathbf{S}_L^1\|_F. \quad (44)$$

For the L -collided particles, the coefficient vector reads

$$\|\mathbf{S}_L^1\|_F \leq \|\widehat{\mathbf{S}}_L^1\|_F \leq \frac{1}{1 + \Delta t \boldsymbol{\Sigma}_t} \left(\|\mathbf{S}_L^{1/2}\|_F + \Delta t \max_k \boldsymbol{\Sigma}_{kk} \|\mathbf{S}_{L-1}^1\|_F \right).$$

Adding $\|\mathbf{S}_L^1\|_F$ to both sides of (43) gives

$$\begin{aligned} \|\mathbf{S}_c^1\|_F + \|\mathbf{S}_L^1\|_F &\leq \|\mathbf{S}_c^{1/2}\|_F + \max_k \frac{1 + \boldsymbol{\Sigma}_t \Delta t}{1 + \boldsymbol{\Sigma}_t \Delta t - \Delta t \boldsymbol{\Sigma}_{kk}} \|\mathbf{S}_L^1\|_F \\ &\leq \|\mathbf{S}_c^{1/2}\|_F + \|\mathbf{S}_L^{1/2}\|_F + \max_k \frac{\Delta t \boldsymbol{\Sigma}_{kk}}{1 + \boldsymbol{\Sigma}_t \Delta t - \Delta t \boldsymbol{\Sigma}_{kk}} \|\mathbf{S}_{L-1}^1\|_F. \end{aligned}$$

By Theorem 2, we have that

$$\|\mathbf{S}_c^1\|_F + \|\mathbf{S}_L^1\|_F \leq \|\mathbf{S}_c^0\|_F + \|\mathbf{S}_L^0\|_F + \max_k \frac{\Delta t \Sigma_{kk}}{1 + \Sigma_t \Delta t - \Delta t \Sigma_{kk}} \|\mathbf{S}_{L-1}^1\|_F.$$

Continuing recursively and following the last steps of the proof for Theorem 1 proves the statement. \square

7. Numerical results

In the following, we demonstrate numerical experiments to compare conventional and the proposed methods. All results can be reproduced with the openly available code framework [31].

7.1. Line source benchmark

To demonstrate the applicability of the proposed collision source method for dynamical low-rank approximation in general radiation transport applications, we first take a look at the time dependent radiation transport equation for the *line source* benchmark [20, 21]

$$\begin{aligned} \partial_t \psi + \boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \psi + \Sigma_s \psi &= \frac{\Sigma_s}{4\pi} \int_{\mathbb{S}^2} \psi d\boldsymbol{\Omega}, \quad (\mathbf{x}, \boldsymbol{\Omega}) \in [-1.5, 1.5]^2 \times \mathbb{S}^2, \\ \psi(t_0) &= \frac{1}{4\pi\sigma^2} \exp\left(-\frac{\|\mathbf{x}\|^2}{4\sigma^2}\right), \end{aligned} \quad (45)$$

where $\Sigma_s = 1$ and $\sigma = 0.03$. This equation can be recovered from the continuous slowing down approximation when choosing $\rho \equiv 1$ and treating the energy variable as time. The line source benchmark is a common test case for radiation transport problems, exposing deficiencies of different methods. A comparison of conventional methods for this benchmark can for example be found in [22]. Common methods require high computational costs or parameter tuning to yield a satisfactory approximation. Uses of dynamical low-rank approximation for this benchmark are [38, 37, 8], where it is observed that high ranks are needed to achieve a desired level of accuracy. Nevertheless, in comparison to classical methods, DLRA yields reduced run times and memory requirements. We use the following computational parameters for our calculations:

$n_x = N_x \cdot N_y = 40000$	number of spatial cells
$m = (N + 1)^2 = 484$	number of spherical basis functions
$n_q = 968$	number of quadrature points for uncollided flux
$t_{\text{end}} = 1$	end time

We use a CFL number of 0.5 according to (39). The scalar flux $\Phi(t = 1, \mathbf{x}) = \int_{\mathbb{S}^2} \psi(t = 1, \mathbf{x}, \boldsymbol{\Omega}) d\boldsymbol{\Omega}$ computed with different methods can be found in Figure 2. We observe a significant increase in the solution quality when using $L = 4$ instead of $L = 1$ levels. The level 4 approximation with a tolerance parameter of $\vartheta = 0.3$ agrees well with the P_N solution. Here, we use the term P_N to indicate the use of an S_N solver for uncollided particles as well as a P_N solver for the remainder. While P_N takes 5408 seconds to compute the scalar flux at time $t = 1$, the DLRA methods with $L = 1$ and $L = 4$ levels only require 1009 and 1278 seconds respectively. Since particles move into all directions, a main factor in this run time is the S_N solution. Taking a look at the rank evolution in time, depicted in Figure 3, we see that most information is carried by the uncollided flux as well as solution components with little collisions.

7.2. Treatment planning for lung patient

In the following we examine the application of the proposed method to a realistic 2D CT scan of a lung patient. The patient is radiated with an electron beam of $E_{\text{max}} = 21$ MeV. We model this beam as

$$\begin{aligned} \psi_{\text{in}}(E, \mathbf{x}, \boldsymbol{\Omega}) &= 10^5 \cdot \exp(-(\Omega_{1,*} - \Omega_1)^2 / \sigma_{\Omega_1}) \cdot \exp(-(E_{\text{max}} - E)^2 / \sigma_E) \\ &\quad \cdot \exp(-(x_* - x)^2 / \sigma_x) \cdot \exp(-(y_* - y)^2 / \sigma_y), \end{aligned}$$

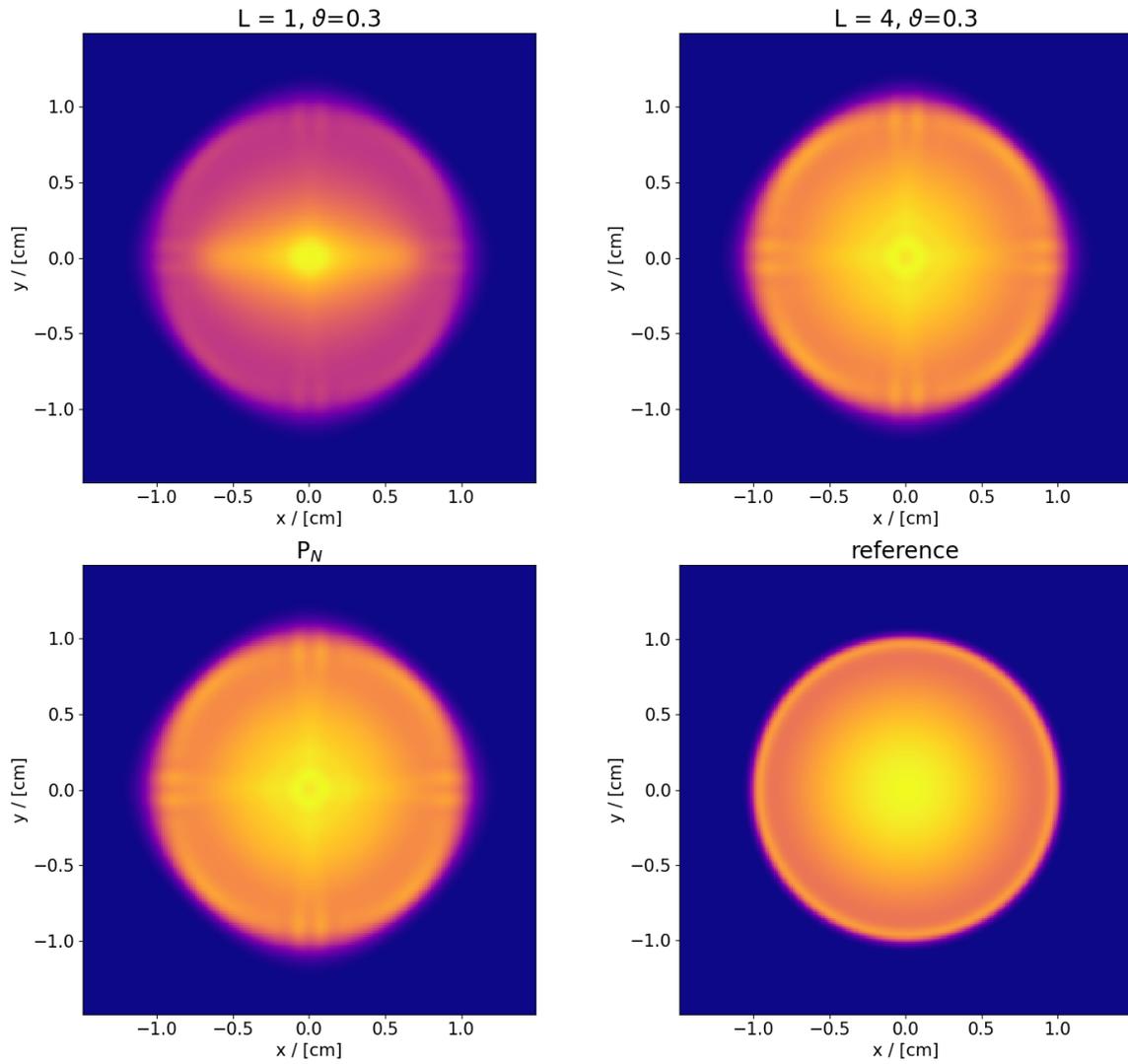


Figure 2: Scalar flux $\Phi(t = 1, \mathbf{x}) = \int_{\mathbb{S}^2} \psi(t = 1, \mathbf{x}, \mathbf{\Omega}) d\mathbf{\Omega}$ with different methods and analytic reference solution.

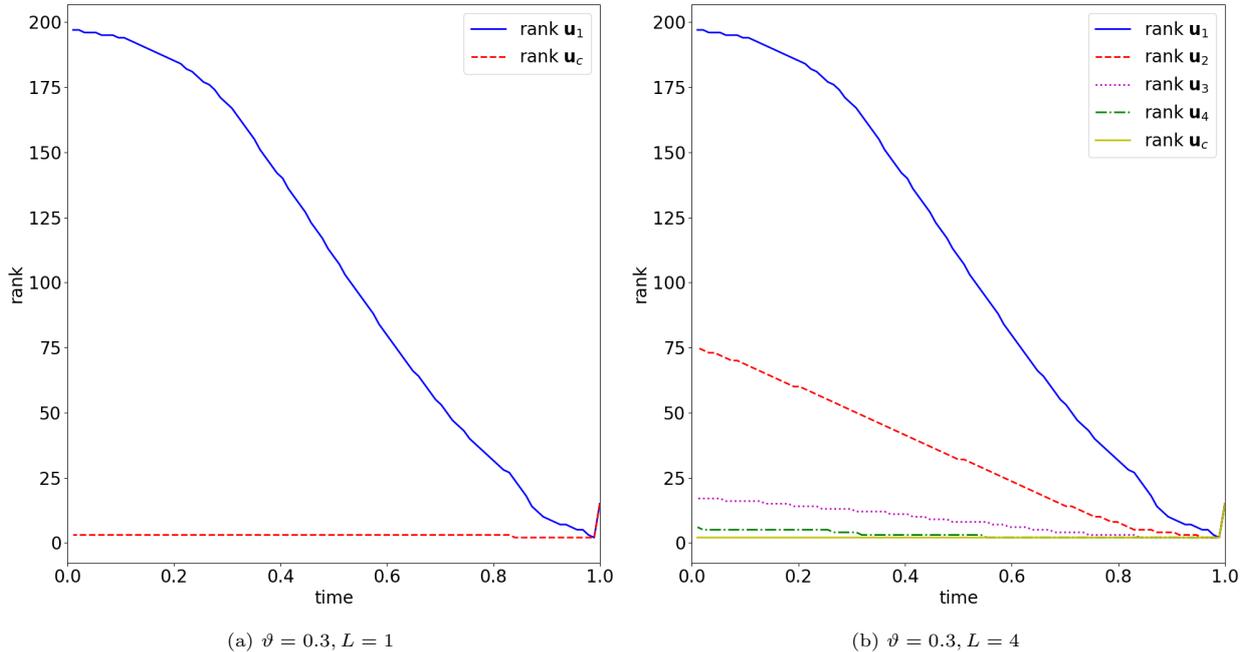


Figure 3: Rank evolution in energy for different tolerance parameters ϑ .

which is used as boundary condition for the uncollided particle flux. To determine a tissue density ρ for given gray-scale values of the CT image, we assume a value of one, i.e., a white pixel to consist of bone material with density $\rho_{\text{bone}} = 1.85 \text{ g/cm}^3$. The remaining tissue is scaled such that a pixel value of zero corresponds to a minimal density of $\rho_{\text{min}} = 0.05 \text{ g/cm}^3$. Air around the patient is filled with material, since this region does not impact the dose distribution. The chosen settings are the same as in Section 7.1. Since we are using a directed particle beam as boundary condition for the uncollided particles, the number of quadrature points n_q reduces by over 59 percent. The remaining parameters are:

$n_q = 396$	number of quadrature points for uncollided flux
$E_{\text{max}} = 21$	energy of beam in MeV
$x_* = 7.25, y_* = 14.5$	spatial mean of particle beam in cm
$\Omega_{1,*} = 1$	directional mean of particle beam
$\sigma_{\Omega_1}^{-1} = 75$	inverse directional beam variance
$\sigma_x^{-1} = \sigma_y^{-1} = 20$	inverse spatial beam variance
$\sigma_E^{-1} = 100$	inverse energy variance

We use a CFL number of 0.5 according to (39). For this setting we compute the full P_N solution, the proposed dynamical low-rank method with a fixed rank of 50 consisting only of collided and uncollided particles as well as the rank adaptive version with $L = 1$ intermediate levels. Due to its reduced computational costs, the DLRA methods show a significantly reduced run time. While the full P_N method runs for 47329 seconds, the DLRA methods have a run time of 4373, 3917 and 5392 seconds respectively. The resulting dose distribution can be found in Figure 4. All considered variations of the proposed method are able to capture the effect of heterogeneities in the patient density and agree very well with the P_N solution in the relevant dose areas. The efficiency of the method concerning both time and memory makes it feasible for practical applications. This includes the generation of optimal treatment plans with gradient-based optimization methods. It is observed that choosing a low refinement tolerance $\vartheta = \bar{\vartheta} \cdot \|\mathbf{S}\|_F$ with $\bar{\vartheta} = 0.01$ leads to a slight difference to the full solution for the smallest isoline. Note that isolines appearing below and above the main beam on the right are artifacts by dose values close to zero and are not of interest. The corresponding ranks at

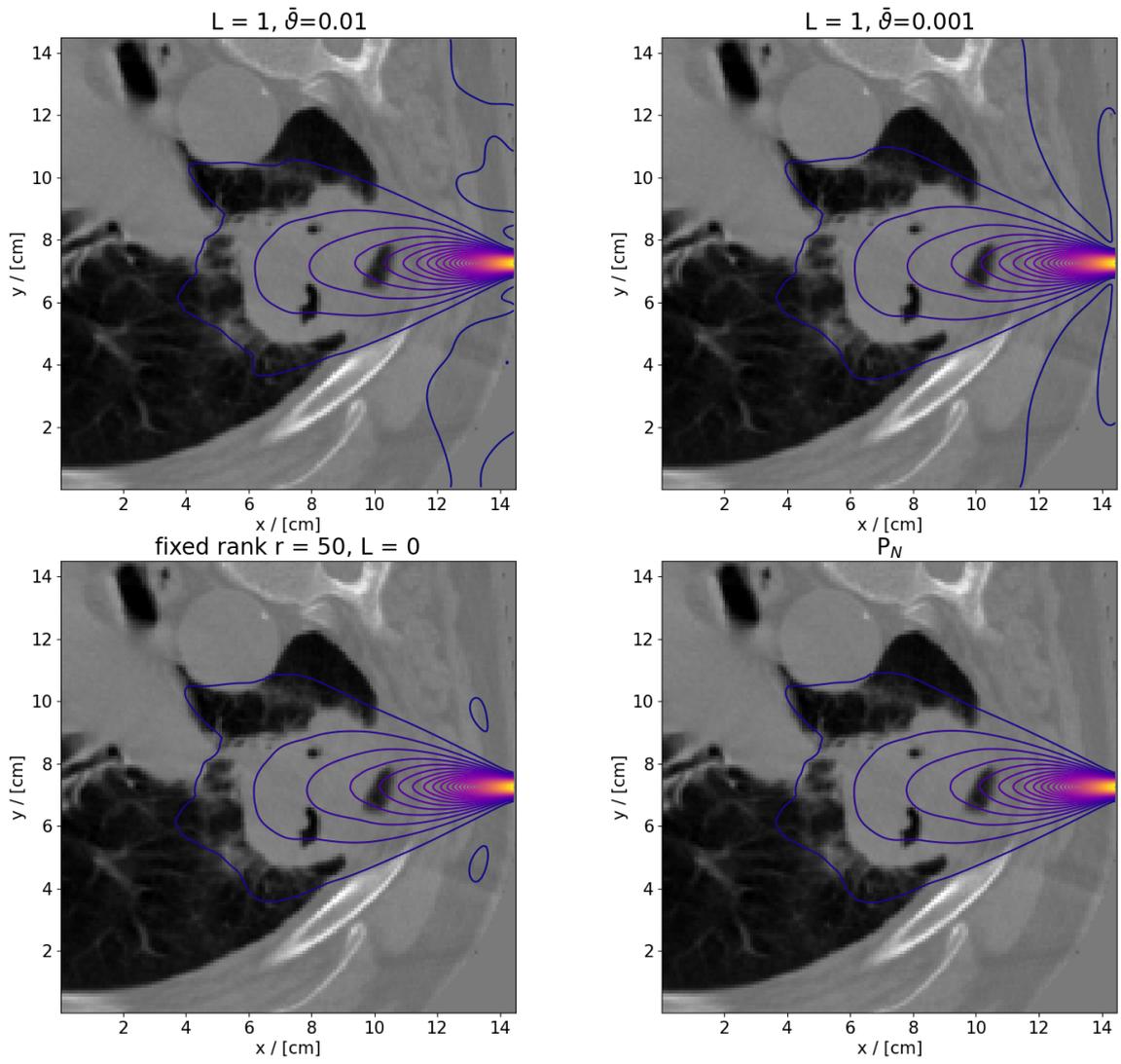


Figure 4: Dose distribution with different methods.

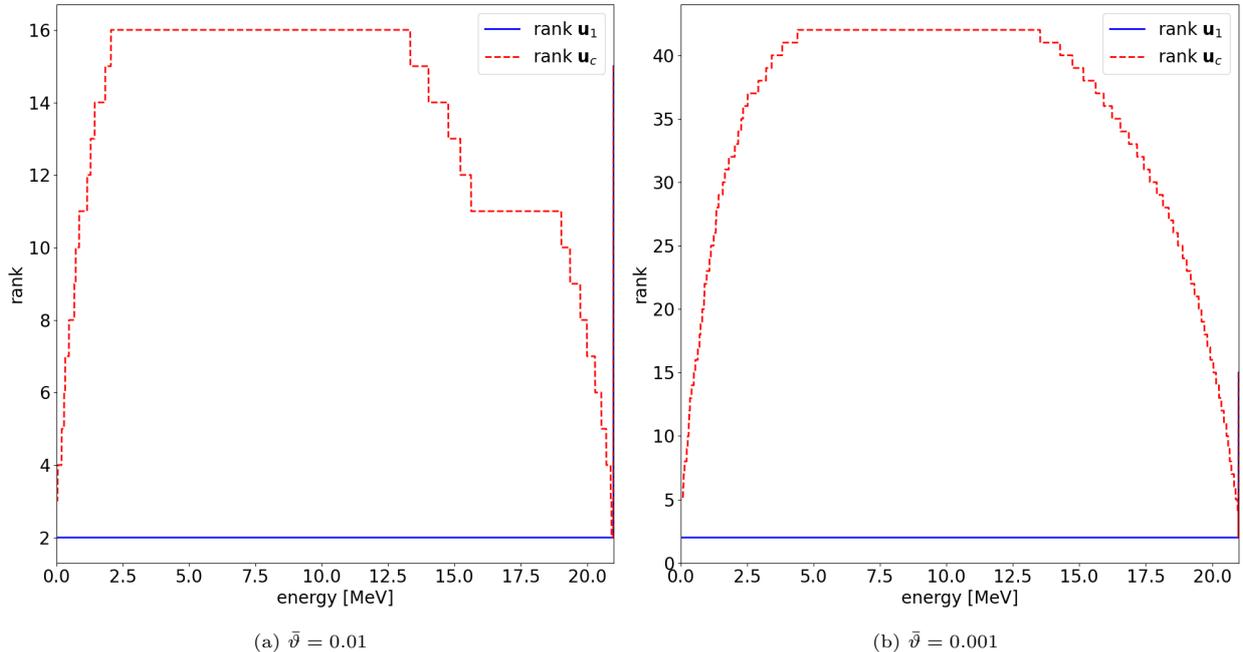


Figure 5: Rank evolution in energy for different $\vartheta = \bar{\vartheta} \cdot \|\mathbf{S}\|_F$.

different energies for $\bar{\vartheta} \in \{0.001, 0.01\}$ are depicted in Figure 5. It is observed that the rank of collided particles remains small at low and high energies. At intermediate energies, the rank reaches its maximum. Particles that have collided once can be described with a small rank throughout the simulation. Note that particles which have collided once are not present for energies below 15 MeV. The reason for this is that particles directly enter the patient tissue and are therefore directly subject to scattering.

The first four dominant spatial modes at the lowest energy are depicted in Figure 6 and the first four dominant directional modes are shown in Figure 7. These modes have been computed by an SVD of the coefficient matrix $\mathbf{S} = \mathbf{U}\mathbf{D}\mathbf{V}^T$. We then plot the first four columns of $\mathbf{X}\mathbf{U}$ and $\mathbf{W}\mathbf{V}$. It is observed that the directional basis carries the information that particles are predominantly travelling into the x -direction, i.e. into the direction of the particle beam. The spatial basis encodes that particles with low energies are situated at the left of the CT scan and can mostly be found in high-density tissue. Note that due to the low energy of these particles, no significant contribution to the overall dose distribution is observed.

8. Summary and outlook

In this work, we proposed a dynamical low-rank approximation for the use in radiation therapy. Instead of computing the full solution, dynamical low-rank approximation evolves a low-rank factorization of the solution in time, thereby significantly reducing computational times and memory requirements. The method can be understood as a Galerkin method, which automatically and dynamically picks basis functions to capture relevant information of the solution. Through rank adaptivity, the rank of the solution approximation (that is, the number of basis functions) decreases or increases in time according to the solution complexity. To use DLRA for treatment planning applications, we formulate the energy dependence of the governing steady state equations as a pseudo-time according to the continuous slowing down approximation. A rank adaptive unconventional integrator is then chosen to update the factorized solution in energy.

Furthermore, a collided-uncollided split is used to evolve only collided particles with a dynamical low-rank approximation. Thereby, we potentially reduce the required rank while facilitating the implementation of boundary conditions. This approach can be extended to L -collided splits, which can be interpreted as

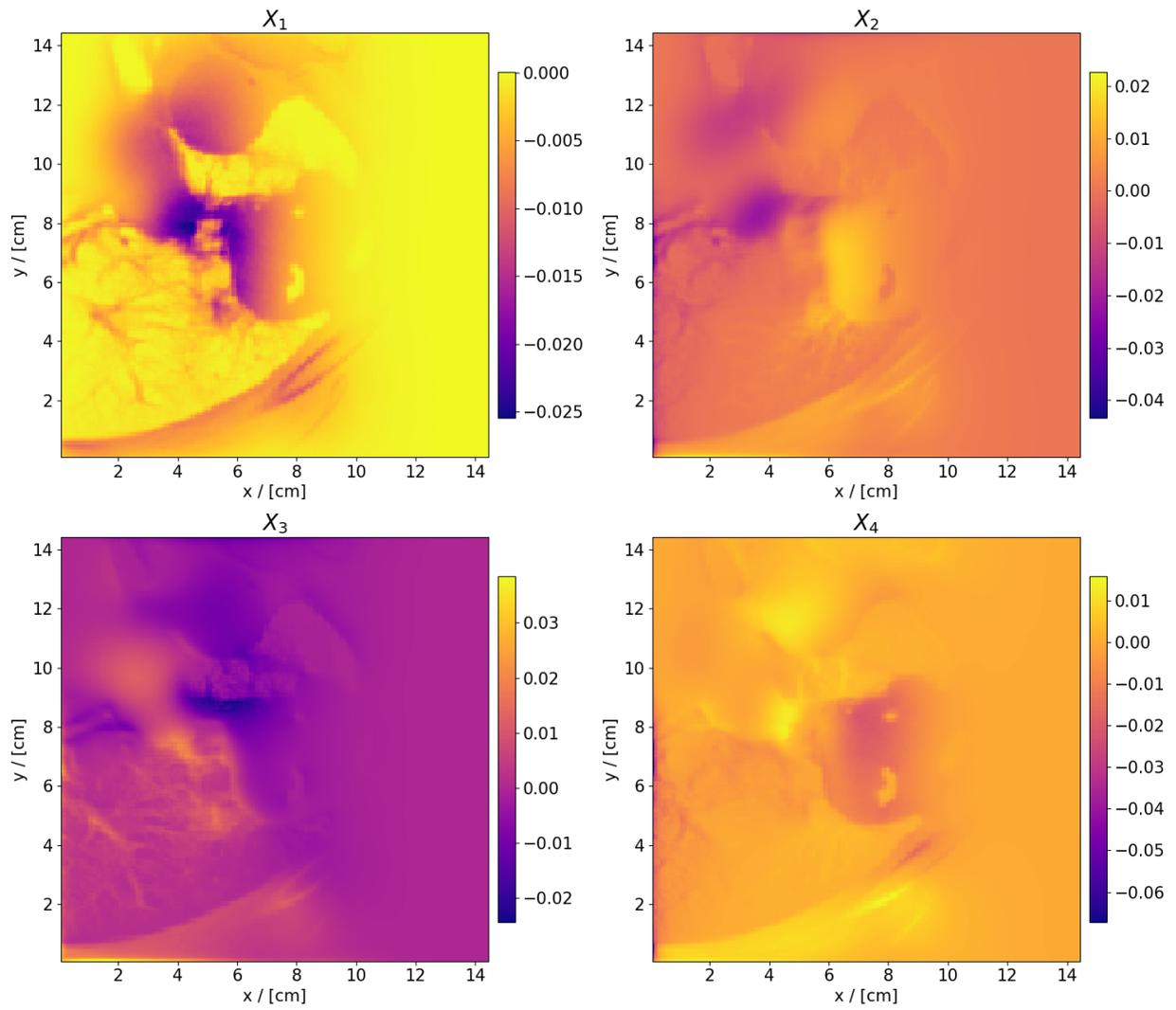


Figure 6: First four dominant spatial modes with fixed rank integrator.

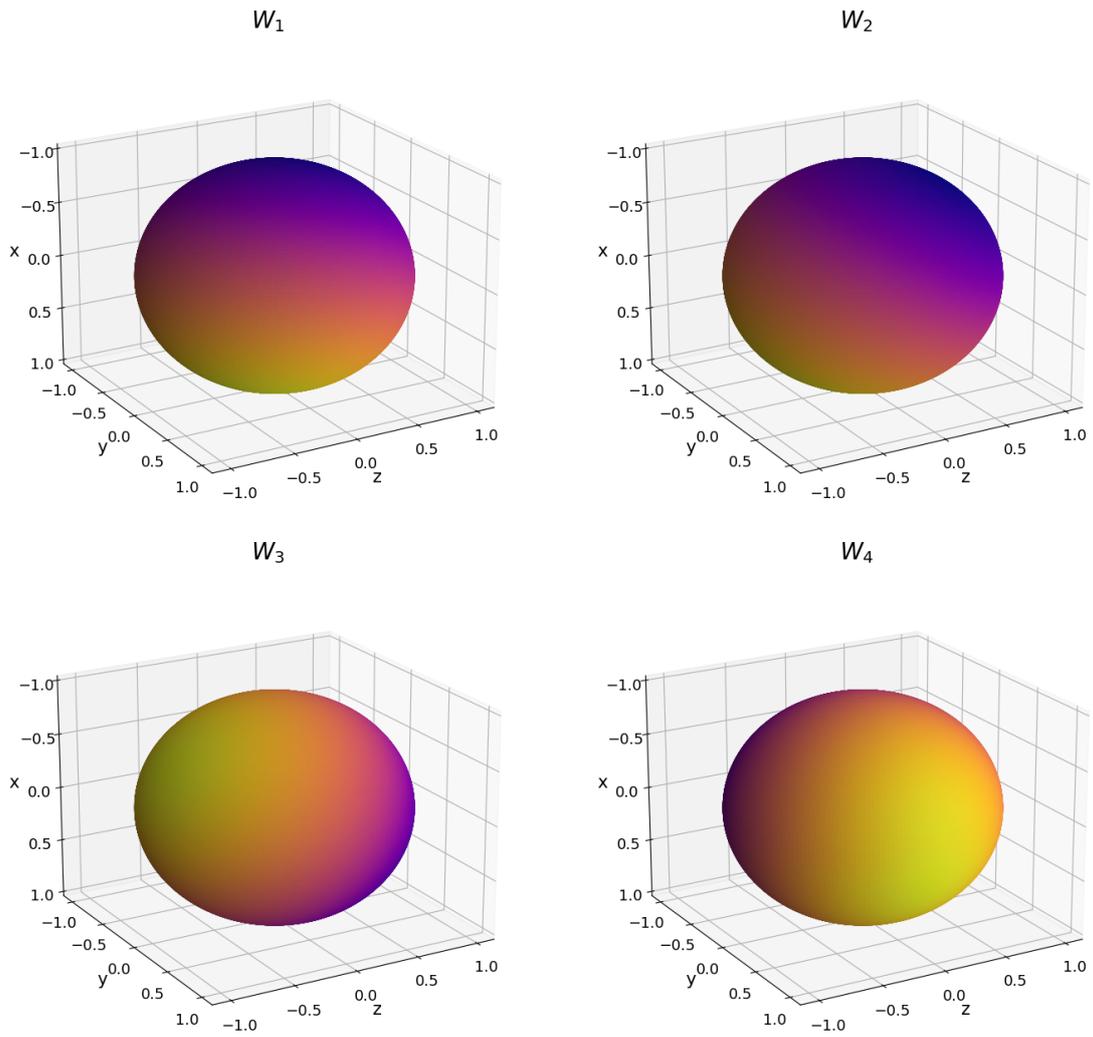


Figure 7: First four dominant directional modes with fixed rank integrator.

writing the solution as a telescoping sum. By applying DLRA to each update in this sum, we can decrease the solution complexity in a multilevel fashion. Additionally, we propose an efficient implicit time discretization for scattering terms, which allows for increased time step sizes while not requiring the inversion of matrices. The proposed method is shown to be L^2 -stable under a CFL restriction which only depends on the streaming term. In our numerical experiments, we observe that the solution can be captured with a reduction in run times by a factor of over 12.

In future work, we aim at using the proposed forward method to facilitate optimization and uncertainty quantification in radiation therapy. To improve the understanding of dynamical low-rank approximation in this new field of application, further validations, especially in 3D geometries and against the current standard pencil beam or MC algorithms are necessary. A further advantage of DLRA that needs to be investigated could be the use of GPU parallelizations which is limited for conventional particle based methods. Our results promise an approach which is efficient enough for a use in dose optimization, while still taking into account all relevant physical interactions.

Acknowledgments

Jonas Kusch has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — Project-ID 258734477 — SFB 1173. Pia Stammer is supported by the Helmholtz Association under the joint research school HIDSS4Health – Helmholtz Information and Data Science School for Health.

References

- [1] M. L. Adams and E. W. Larsen. Fast iterative methods for discrete-ordinates particle transport calculations. *Progress in nuclear energy*, 40(1):3–159, 2002.
- [2] A. Ahnesjö and M. M. Aspradakis. Dose calculations for external photon beams in radiotherapy. *Physics in Medicine & Biology*, 44(11):R99, 1999.
- [3] R. Alcouffe, R. O’Dell, and F. Brinkley Jr. A first-collision source method that satisfies discrete S_n transport balance. *Nuclear Science and Engineering*, 105(2):198–203, 1990.
- [4] P. Andreo. Monte Carlo techniques in medical radiation physics. 36(7):861–920, jul 1991.
- [5] C. Berthon, M. Frank, C. Sarazin, and R. Turpault. Numerical methods for balance laws with space dependent flux: application to radiotherapy dose calculation. *Communications in Computational Physics*, 10(5):1184–1210, 2011.
- [6] C. Börgers. Complexity of Monte Carlo and deterministic dose-calculation methods. *Physics in Medicine & Biology*, 43(3):517, 1998.
- [7] K. M. Case and P. F. Zweifel. Linear transport theory. 1967.
- [8] G. Ceruti, J. Kusch, and C. Lubich. A rank-adaptive robust integrator for dynamical low-rank approximation. *arXiv preprint arXiv:2104.05247*, 2021.
- [9] G. Ceruti and C. Lubich. An unconventional robust integrator for dynamical low-rank approximation. *BIT Numerical Mathematics*, pages 1–22, 2021.
- [10] A. Dektor, A. Rodgers, and D. Venturi. Rank-adaptive tensor methods for high-dimensional nonlinear PDEs. *arXiv preprint arXiv:2012.05962*, 2020.
- [11] Z. Ding, L. Einkemmer, and Q. Li. Dynamical low-rank integrator for the linear Boltzmann equation: error analysis in the diffusion limit. *arXiv:1907.04247*, 2019.
- [12] L. Einkemmer, J. Hu, and L. Ying. An efficient dynamical low-rank algorithm for the Boltzmann-BGK equation close to the compressible viscous flow regime. *arXiv preprint arXiv:2101.07104*, 2021.
- [13] L. Einkemmer and I. Joseph. A mass, momentum, and energy conservative dynamical low-rank scheme for the Vlasov equation. *arXiv preprint arXiv:2101.12571*, 2021.
- [14] L. Einkemmer and C. Lubich. A low-rank projector-splitting integrator for the Vlasov–Poisson equation. *SIAM J. Sci. Comput.*, 40(5):B1330–B1360, 2018.
- [15] L. Einkemmer and C. Lubich. A quasi-conservative dynamical low-rank algorithm for the Vlasov equation. *SIAM Journal on Scientific Computing*, 41(5):B1061–B1081, 2019.
- [16] L. Einkemmer, A. Ostermann, and C. Piazzola. A low-rank projector-splitting integrator for the Vlasov–Maxwell equations with divergence correction. *Journal of Computational Physics*, 403:109063, 2020.
- [17] L. Eyges. Multiple scattering with energy loss. *Physical Review*, 74(10):1534, 1948.
- [18] M. Fippel and M. Soukup. A Monte Carlo dose calculation algorithm for proton therapy. *Medical physics*, 31(8):2263–2273, 2004.
- [19] M. Frank, H. Hensel, and A. Klar. A fast and accurate moment method for the Fokker–Planck equation and applications to electron radiotherapy. *SIAM Journal on Applied Mathematics*, 67(2):582–603, 2007.
- [20] B. Ganapol. Homogeneous infinite media time-dependent analytic benchmarks for X-TM transport methods development. *Los Alamos National Laboratory*, 1999.

- [21] B. D. Ganapol. Analytical benchmarks for nuclear engineering applications. *Case Studies in Neutron Transport Theory*, 2008.
- [22] C. K. Garrett and C. D. Hauck. A comparison of moment closures for linear kinetic transport equations: The line source benchmark. *Transport Theory and Statistical Physics*, 42(6-7):203–235, 2013.
- [23] V. Grégoire and T. Mackie. State of the art on dose prescription, reporting and recording in Intensity-Modulated Radiation Therapy (ICRU report No. 83). *Cancer/Radiothérapie*, 15(6-7):555–559, 2011.
- [24] C. D. Hauck and R. G. McClarren. A collision-based hybrid method for time-dependent, linear, kinetic transport equations. *Multiscale Modeling & Simulation*, 11(4):1197–1227, 2013.
- [25] K. R. Hogstrom, M. D. Mills, and P. R. Almond. Electron beam dose calculations. *Physics in Medicine & Biology*, 26(3):445, 1981.
- [26] X. Jia, J. Schümann, H. Paganetti, and S. B. Jiang. GPU-based fast Monte Carlo dose calculation for proton therapy. *Physics in Medicine & Biology*, 57(23):7783, 2012.
- [27] E. Kieri, C. Lubich, and H. Walach. Discretized dynamical low-rank approximation in the presence of small singular values. *SIAM J. Numer. Anal.*, 54(2):1020–1038, 2016.
- [28] O. Koch and C. Lubich. Dynamical low-rank approximation. *SIAM J. Matrix Anal. Appl.*, 29(2):434–454, 2007.
- [29] T. Krieger and O. A. Sauer. Monte Carlo- versus pencil-beam-/collapsed-cone-dose calculation in a heterogeneous multi-layer phantom. 50(5):859–868, feb 2005.
- [30] J. Kusch, L. Einkemmer, and G. Ceruti. On the stability of robust dynamical low-rank approximations for hyperbolic problems. *arXiv preprint arXiv:2107.07282*, 2021.
- [31] J. Kusch and P. Stammer. *Numerical testcases for "A robust collision source method for rank adaptive dynamical low-rank approximation in radiation therapy"*, 2021. <https://github.com/JonasKu/publication-A-robust-collision-source-method-for-rank-adaptive-dynamical-low-rankapproximation.git>.
- [32] E. W. Larsen, M. M. Miften, B. A. Fraass, and I. A. Bruinvis. Electron dose calculations using the method of moments. *Medical physics*, 24(1):111–125, 1997.
- [33] E. E. Lewis and W. F. Miller. *Computational methods of neutron transport*. 1984.
- [34] C. Lubich and I. V. Oseledets. A projector-splitting integrator for dynamical low-rank approximation. *BIT*, 54(1):171–188, 2014.
- [35] E. Olbrant and M. Frank. Generalized Fokker–Planck theory for electron and photon transport in biological tissues: application to radiotherapy. *Computational and mathematical methods in medicine*, 11(4):313–339, 2010.
- [36] A. Ostermann, C. Piazzola, and H. Walach. Convergence of a low-rank Lie–Trotter splitting for stiff matrix differential equations. *arXiv:1803.10473*, 2018.
- [37] Z. Peng and R. G. McClarren. A high-order/low-order (HOLO) algorithm for preserving conservation in time-dependent low-rank transport calculations. *arXiv preprint arXiv:2011.06072*, 2020.
- [38] Z. Peng, R. G. McClarren, and M. Frank. A low-rank method for two-dimensional time-dependent radiation transport calculations. *J. Comput. Phys.*, 421:109735, 2020.
- [39] M. Yang and S. R. White. Time-dependent variational principle with ancillary Krylov subspace. *Phys. Rev. B*, 102(9):094315, 2020.