

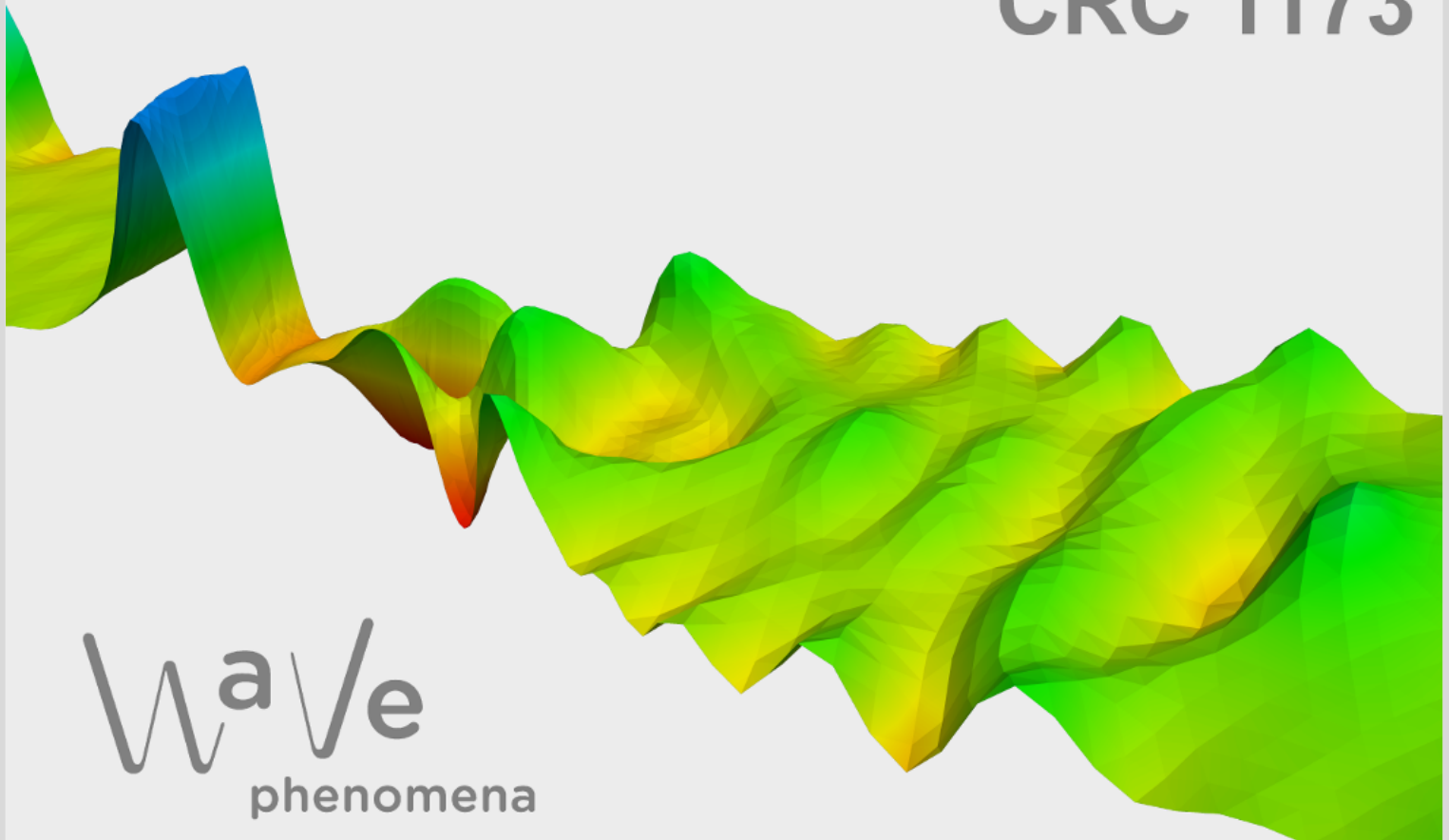
A dynamical low-rank multilevel Monte Carlo estimator for high-dimensional kinetic equations

Chinmay Patwardhan, Sebastian Krumscheid,
Jonas Kusch, Emil Løvbak, Pia Stammer

CRC Preprint 2026/26, July 2026

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Participating universities



Funded by



A Dynamical Low-rank Multilevel Monte Carlo Estimator for High-Dimensional Kinetic Equations

CHINMAY PATWARDHAN ¹
SEBASTIAN KRUMSCHEID ²
JONAS KUSCH ³
EMIL LØVBAK ⁴
PIA STAMMER ⁵

¹ Dept. of Mathematics, Karlsruhe Institute of Technology, Karlsruhe, Germany

Email address: chinmay.patwardhan@kit.edu

² Scientific Computing Center, Karlsruhe Institute of Technology, Karlsruhe, Germany

Email address: sebastian.krumscheid@kit.edu

³ Scientific Computing, Norwegian University of Life Sciences, Ås, Norway

Email address: jonas.kusch@nmbu.no

⁴ Scientific Computing Center, Karlsruhe Institute of Technology, Karlsruhe, Germany

Email address: emil.loevbak@kit.edu

⁵ Dept. of Radiation Science & Technology, Delft University of Technology (TU Delft), Delft, Netherlands

Email address: p.k.stammer@tudelft.nl.

Abstract. Kinetic equations are used to model a wide range of phenomena important for real-world applications. Their applications span astrophysics, nuclear physics, engineering, and social sciences. Due to their high-dimensional phase space, modelling and quantifying uncertainties, relevant for applications, poses a significant challenge even for modern computing infrastructure. In recent years, dynamical low-rank approximation (DLRA) has gained popularity for making fine grid simulations of high-dimensional problems feasible by evolving the solution of a time-dependent PDE as a low-rank factorization. This reduces the computational and memory requirements significantly.

In this work, we propose a low-rank multilevel Monte Carlo estimator for kinetic equations based on a probabilistic rank-adaptive DLRA time integrator. The level hierarchy of the low-rank multilevel estimator is constructed through spatial refinement and by ensuring that the low-rank error remains below the spatial discretization error. We demonstrate the efficacy of the estimator through several numerical experiments from radiation transport, radiation therapy, and shallow water flow.

Keywords. uncertainty quantification, multilevel, multifidelity, dynamical low-rank approximation, Monte Carlo.

2020 Mathematics Subject Classification. 65C05; 65Z05; 65C20; 35Q62; 35Q49.

Chinmay Patwardhan was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173.

Emil Løvbak was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 563450842.

Pia Stammer received funding from the German National Academy of Sciences Leopoldina for the project underlying this article, under grant number LPDS 2024-03.

1. Introduction

Kinetic equations model the microscopic effects of particle transport and their interactions, such as collisions and absorption, through a probability density function. By bridging the gap between atomistic particle models and macroscopic continuum models, kinetic equations play a crucial role in modelling multi-scale dynamics of transport phenomena. These equations describe a wide range of physical phenomena ranging from gas dynamics and radiation transport to shallow water flow and chemical reaction networks, with applications spanning astrophysics, nuclear physics, radiation therapy, engineering, and social sciences [52, 37, 1, 30]. Since real-world applications often introduce uncertainties, including these in models provides a more realistic description. These uncertainties stem from modeling errors, including the use of empirical collision kernels; measurement errors, such as incorrect measurements of the physical domain; and device errors, including errors in initial and boundary data. Quantifying these uncertainties is thus essential for robustness and reliability in real-world applications.

In this work, we consider uncertain kinetic equations of the type

$$\partial_t u(t, x, v; \omega) + \mathcal{A}(u) \partial_x u(t, x, v; \omega) = \mathcal{S}(u; \omega)(t, x, v; \omega) \quad (1.1a)$$

$$u(0, x, v; \omega) = u_0(x, v; \omega), \quad (1.1b)$$

$$u(t, x, v; \omega) = u_b(t, x, v; \omega), \quad x \in \Gamma_{\mathcal{D}}, \quad (1.1c)$$

where $u(t, x, v; \omega)$ denotes the phase space density at time $t \in \mathbb{R}_{\geq 0}$, position $x = (x_1, \dots, x_d) \in \mathcal{D} \subseteq \mathbb{R}^d$, moving with velocity $v = (v_1, \dots, v_d) \in \mathcal{V} \subseteq \mathbb{R}^d$. The random variable $\omega \in \mathbb{R}^k$ models the uncertainty and has the probability density function (pdf) $p(\omega)$. The left-hand side of (1.1a), known as the streaming term, models particle movement in the phase-space where $\mathcal{A}(u) \partial_x u$ is known as the transport or advection term. Changes in the velocity of particles due to inter-particle collisions or scattering and absorption by the background medium are modelled by the collision operator \mathcal{S} . The initial condition is given by u_0 and the boundary conditions are given by u_b on the boundary $\Gamma_{\mathcal{D}}$ of \mathcal{D} . The streaming and scattering operators are problem-dependent and determine the dynamics of the system along with the initial and boundary conditions.

We are often interested in quantifying uncertainties in the entire solution of a PDE or a quantity derived from it. Let $Q : \omega \mapsto Q(\omega) = Q(u(\cdot, \cdot, \cdot; \omega)) \in H$ be a bounded linear functional or a Lipschitz function of the weak solution $u(t, x, v; \omega)$ mapping into the Hilbert space H . Then Q denotes the quantity of interest (QoI) of the solution. For instance, if Q is a functional of the solution, then $H = \mathbb{R}$, while if it is the weak solution at some time $T \geq 0$, then $Q(\omega) := u(T, \cdot, \cdot; \omega) \in H = L^2(\mathcal{D} \times \mathcal{V})$. In this work, we are interested in estimating the expected value of the QoI under the random variable $\omega \sim p(\omega)$, i.e.

$$\mathbb{E}[Q] := \int_{\mathbb{R}^k} Q(u(\cdot, \cdot, \cdot; \omega)) p(\omega) d\omega, \quad (1.2)$$

with the variance

$$\mathbb{V}[Q] := \mathbb{E} \left[\|Q(\omega) - \mathbb{E}[Q]\|_H^2 \right]. \quad (1.3)$$

There are several ways of estimating $\mathbb{E}[Q] \in H$, broadly classified as *intrusive* and *non-intrusive* methods. As the name suggests, intrusive uncertainty quantification (UQ)

methods, such as the generalized polynomial chaos (gPC)-based stochastic Galerkin (sG) method, modify the underlying equations to obtain estimates of $\mathbb{E}[Q]$. In contrast, non-intrusive methods, such as sampling-based Monte Carlo (MC) methods and collocation-based methods, do so without altering the governing equations. A major challenge to quantifying uncertainties in kinetic equations is the high computational complexity and memory requirements for grid-based simulations. This is due to its high-dimensional phase space, which for each realization of ω , including time, is at least $2d+1$ dimensional, where d is the spatial dimension. In the literature, this phenomenon is referred to as the *curse of dimensionality*, wherein computational costs grow exponentially with the number of dimensions. Mitigation techniques, such as reduced-order models [38, 48], surrogate models [11, 25], and low-rank approximations [18, 27, 16, 32, 50], have been proposed to reduce the computational complexity for probabilistic and deterministic kinetic equations.

In recent years, dynamical low-rank approximation (DLRA) [27] has gained significant popularity in the kinetic theory community as an efficient numerical method for approximating deterministic high-dimensional time-dependent PDEs. DLRA is a hybrid approach in which time-dependent low-rank factors of the solution are updated according to the problem dynamics. This approach simultaneously reduces the computational costs and memory requirements of a simulation. A recent review paper highlights the use of DLRA in the context of kinetic equations [16]. We refer the reader to this work for a detailed comparison of the methods and [17, 15, 32, 49, 35] for further reading on DLRA in kinetic theory.

We are interested in developing DLRA-based UQ methods for kinetic equations. The first such approach using a slightly different representation was developed under the term dynamically orthogonal (DO) field expansion in [53] and applied to problems in ocean dynamics with random data [56]. Apart from the DO-type methods, there are primarily two approaches in the literature that address the combination of DLRA and uncertainties. The first approach treats the random variable ω as a phase-space variable and evolves an uncertain basis over time and has been explored in [45, 46, 32]. This approach can be viewed as a stochastic Galerkin method with a time-dependent, adaptive random basis. This approach can be viewed as a stochastic Galerkin method with a time-dependent, adaptive random basis. The second approach, outlined in [50], treats the random variable as an input parameter to the PDE and uses DLRA as a black-box approximation method for MC-based sampling. MC-based methods are widely adopted due to their ease of implementation and widespread use for propagating uncertainties in complex phenomena. Variance-reducing enhancements, such as the multilevel Monte Carlo (MLMC) [21, 22] method, have further accelerated the use of sampling-based methods. MLMC estimators exploit the hierarchical structures of numerical methods to construct estimators that balance stochastic and numerical approximation errors while minimizing computational costs. Extensive work on MLMC theory and estimators for elliptic equations, parabolic equations, and hyperbolic conservation laws can be found in [12, 13, 23, 2, 42, 43, 44]. Hence, we are interested in the second approach of combining DLRA with a sampling-based UQ methods, wherein we combine non-intrusive methods with DLRA.

To mitigate the curse of dimensionality, in our previous work [50], we constructed a DLRA-based low-rank Monte Carlo (DLR-MC) estimator for radiation transport in one-dimensional slab geometry. This estimator computed the low-rank approximation based on a *fixed-rank integrator*, i.e., one rank was chosen apriori and stayed constant over time. The variance of the estimator was further reduced by constructing a control variates estimator (DLR-CV) based on a “high” rank - “low” rank multi-fidelity strategy. The DLR-MC and DLR-CV estimators focus on reducing the statistical error of the QoI, with the implicit assumption that it dominates the numerical error. In higher dimensions ($d \geq 2$), this imposes a severe restriction on the extent of variance reduction since a highly resolved grid is required to satisfy this assumption. Moreover, these estimators utilize fixed-rank integrators, which necessitate prior knowledge of the rank (or maximal rank) of the solution over the entire simulation. This leads to an over-approximation of rank, resulting in higher computational costs.

In this work, we propose a rank-adaptive low-rank MC estimator based on the augmented BUG integrator [8], which improves on the fixed-rank augmented BUG estimator used in [50]. Further, we design a novel rank-adaptive dynamical low-rank MLMC framework for kinetic equations. The central challenge is efficiently balancing the low-rank and discretization errors, which we do by developing a spatially accurate augmented BUG integrator. The novelties of the work can be summarized as follows,

- *A rank-adaptive low-rank Monte Carlo (RaDLR-MC) estimator:* We propose a rank-adaptive RaDLR-MC estimator based on the augmented BUG integrator, in which the rank of the solution is selected automatically at each time step to satisfy a prescribed truncation tolerance. Unlike the fixed-rank estimator of [50], this allows the approximation to adapt to evolving dynamics of the problem with reduced computational overhead.
- *A novel rank-adaptive low-rank multilevel Monte Carlo (RaDLR-MLMC) framework:* We propose a low-rank MLMC framework based on geometric spatial refinement, a fixed discretization in velocity, and the augmented BUG integrator [8] for time-stepping on the low-rank manifold. The rank-adaptivity of the augmented BUG integrator is designed to preserve spatial discretization errors on each level, thereby ensuring that the MLMC theorem holds.
- *Numerical experiment for high-dimensional kinetic equations:* We validate the proposed RaDLR-MC and RaDLR-MLMC estimators on benchmark kinetic equations, including radiative transport, radiation therapy, and shallow water moment equations.

The rest of the paper is organized as follows: In section 2, we recap DLRA for deterministic PDEs and under Monte Carlo sampling. We present a rank-adaptive low-rank Monte Carlo (RaDLR-MC) estimator for kinetic equations and discuss the relation between the discretization of the phase space and the rank of the approximation. In section 3, we present a rank-adaptation strategy to obtain an augmented BUG integrator that has dominant spatial error. Based on this, we propose a new low-rank Multilevel Monte Carlo (RaDLR-MLMC) framework for kinetic equations. Finally, in section 4 we

present numerical results for radiation transport, radiation therapy, and shallow water moment equations under uncertainty.

2. Dynamical low-rank approximation and Monte Carlo sampling

2.1. Fundamentals of DLRA for deterministic problems

We begin this section by presenting the fundamental ideas of the continuous formulation of DLRA [27] for deterministic time-dependent problems then extend it to the uncertain PDE case.

Let $u(t, x, v)$ denote the solution to the deterministic version of (1.1):

$$\begin{aligned} \partial_t u(t, x, v) + \mathcal{A}(u) \partial_x u(t, x, v) &= \mathcal{S}(u)(t, x, v), \\ u(0, x, v) &= u_0(x, v), \\ u(t, x, v) &= u_b(t, x, v), \quad x \in \Gamma_{\mathcal{D}}. \end{aligned} \tag{2.1}$$

We restrict the solution of the deterministic kinetic equation to the low-rank manifold

$$\begin{aligned} \mathcal{M}_r := \{g \in L^2(\mathcal{D} \times \mathcal{V}) \mid g(x, v) &= \sum_{i,j=1}^r X_i(x) S_{ij} W_j(v) \\ &\text{where } X_i \in L^2(\mathcal{D}), W_j \in L^2(\mathcal{V}), S_{ij} \in \mathbb{R}\} \end{aligned}$$

such that u admits the ansatz:

$$u(t, x, v) \approx u_r(t, x, v) = \sum_{i,j=1}^r X_i(t, x) S_{ij}(t) W_j(t, v), \tag{2.2}$$

where $\{X_i\}_{i=1,\dots,r} \in L^2(\mathcal{D})$ and $\{W_j\}_{j=1,\dots,r} \in L^2(\mathcal{V})$ are orthonormal bases, while $S_{ij}(t)$ denotes the mixing coefficients. Let $\langle \cdot, \cdot \rangle_x$ and $\langle \cdot, \cdot \rangle_v$ denote the inner products defined on $L^2(\mathcal{D})$ and $L^2(\mathcal{V})$, respectively. Then, by definition, we have

$$\langle X_i, X_k \rangle_x = \delta_{ik}, \quad \langle W_j, W_l \rangle_v = \delta_{jl}.$$

The goal is to evolve X_i , W_j , and S_{ij} in time, such that the solution remains on the low-rank manifold \mathcal{M}_r and follows the dynamics of the problem. In DLRA [27], this is done by projecting the dynamics of the problem onto the tangent space to the low-rank manifold. That is, if \mathcal{P}_{u_r} denotes the orthogonal projection onto the tangent space to the low-rank manifold \mathcal{M}_r at u_r , then we solve

$$\partial_t u_r = \mathcal{P}_{u_r}(\mathcal{F}(t, u_r)), \tag{2.3}$$

where

$$\mathcal{F}(t, u_r) := -\mathcal{A}(u_r) \partial_x u_r + \mathcal{S}(t, u_r). \tag{2.4}$$

Using the short-hand notation $\mathcal{F} := \mathcal{F}(t, u_r)$ the orthogonal projection is given by

$$\mathcal{P}_{u_r}(\mathcal{F}) = \sum_{j=1}^r \langle W_j, \mathcal{F} \rangle_v W_j - \sum_{i,j=1}^r X_i \langle X_i W_j, \mathcal{F} \rangle_{x,v} W_j + \sum_{i=1}^r X_i \langle X_i, \mathcal{F} \rangle_x.$$

By imposing the gauge conditions $\langle \partial_t X_i, X_j \rangle_x = 0$, $\langle \partial_t W_i, W_j \rangle_v = 0$ we obtain unique evolution equations for X_i , S_{ij} , and W_j . Note that the specific gauge conditions have been chosen to obtain unique factorization of the low-rank factors in the tangent space [27].

Since the rank of the solution is not known beforehand, it is often over-approximated, resulting in evolution equations for the factors that are then severely ill-conditioned. This is due to the inversion of the coefficient matrix $\mathbf{S}(t) = (S_{ij}(t))$, which has near-zero singular values, that is required for determining the evolution of the bases [27]. Several robust integrators have been proposed in recent years that are robust to the presence of small singular values. The projector-splitting integrator (PSI), based on a Lie-Trotter or Strang splitting of the orthogonal projection, was introduced in [41] and proven to be robust to the presence of small singular values [26]. However, the PSI involves a backward in time step to update $S_{ij}(t)$, which leads to unstable dynamics for parabolic problems and even for hyperbolic problems [33]. The fixed-rank BUG integrator [10], also known as the unconventional integrator, was developed to mitigate this drawback. This development has facilitated the widespread use of DLRA integrators for kinetic problems [16]. Since its development, the BUG integrator has been modified to allow rank-adaptivity [8], completely parallel update steps [9], and provably high-order integrators [7, 31, 47]. In this work, we use the augmented BUG integrator [8], a first-order in time, rank-adaptive integrator for time-dependent problems.

We begin by deriving the evolution equations underlying all the BUG integrators. Let $K_j(t, x) := \sum_{i=1}^r X_i(t, x) S_{ij}(t)$ such that $u_r(t, x, v) = \sum_{j=1}^r K_j(t, x) W_j(t, v)$. Differentiating with respect to t yields

$$\partial_t u_r = \sum_{j=1}^r \partial_t (K_j W_j) = \sum_{j=1}^r (\partial_t K_j W_j + K_j \partial_t W_j).$$

Inserting the above expression into the projected equation (2.3) and projecting onto the space spanned by $\{W_j\}_{j=1, \dots, r}$. For $1 \leq l \leq r$, we obtain

$$\begin{aligned} \sum_{j=1}^r (\partial_t K_j \langle W_j, W_l \rangle_v + K_j \langle \partial_t W_j, W_l \rangle_v) &= \sum_{j=1}^r \langle W_j, \mathcal{F} \rangle_v \langle W_j, W_l \rangle_v \\ &\quad - \sum_{i,j=1}^r X_i \langle X_i W_j, \mathcal{F} \rangle_{x,v} \langle W_j, W_l \rangle + \sum_{i=1}^r X_i \langle X_i W_l, \mathcal{F} \rangle_{x,v}. \end{aligned}$$

Using the orthogonality of W_j and the gauge condition $\langle \partial_t W_i, W_j \rangle_v = 0$, we arrive at

$$\partial_t K_l(t, x) = \langle W_l, \mathcal{F} \rangle_v.$$

Note that since W_l on the right-hand side is also time-dependent, we need to close the system. The system is closed by assuming that W_l remains constant over the time step. Thus, to update the spatial basis from t_0 to $t_1 = t_0 + \Delta t$, we solve the following system of equations

$$\partial_t K_l(t, x) = \langle W_l, \mathcal{F} \rangle_v, \quad \partial_t W_l = 0. \quad (2.5)$$

Similarly, defining $L_i(t, v) := \sum_{j=1}^r S_{ij}(t) W_j(t, v)$ and following the same steps as for K_j , we obtain the following evolution equations for the remaining factors:

$$\partial_t L_i(t, v) = \langle X_i, \mathcal{F} \rangle_x, \quad \partial_t X_i(t, x) = 0, \quad (2.6)$$

$$\partial_t S_{ij}(t) = \langle X_i W_j, \mathcal{F} \rangle_{x,v} \quad \partial_t X_i(t, x) = 0, \quad \partial_t W_j(t, v) = 0. \quad (2.7)$$

The update for S_{ij} is obtained by testing (2.3) with $X_i W_j$ and integrating over the spatial and velocity domains. The choice of initial conditions for these closure equations

LOW-RANK MLMC FOR KINETIC EQUATIONS

Integrator	$S(t_0)$	rank-adaptive	parallel updates
fixed-rank BUG [10]	$X^{1,\top} X^0 S^0 W^{0,\top} W^1 \in \mathbb{R}^{r \times r}$	no	partial
fixed-rank augmented BUG [8]	$\tilde{X}^\top X^0 S^0 W^{0,\top} \tilde{W} \in \mathbb{R}^{2r \times 2r}$	no	partial
augmented BUG [8]	$\tilde{X}^\top X^0 S^0 W^{0,\top} \tilde{W} \in \mathbb{R}^{2r \times 2r}$	yes	partial
parallel BUG [9]	$S_0 \in \mathbb{R}^{r \times r}$	yes	full

TABLE 1. Comparison of BUG integrators for dynamical low-rank approximation, including the S-step initial condition, rank-adaptivity, and parallelism properties. \tilde{X} and \tilde{W} are as defined in the K- and L-steps of the augmented BUG integrator. “Partial” parallel updates are in the K- and L-steps, while “full” implies in all three sub-steps.

leads to various integrators. A summary of the integrators is provided in Table 1. The initial conditions for the K- and L-steps of all the integrators are the same as for the augmented BUG integrator described below.

In the augmented BUG integrator, (2.5) and (2.6) are first updated in parallel from initial data. The basis is expanded by augmenting the previous basis at time t_0 to the updated basis. The coefficients are subsequently updated by projecting onto the augmented basis and evolving time with (2.7). To make this more precise, we present one step of the augmented BUG integrator. For the given initial data $u_r(t_0, x, v) = \sum_{i,j=1}^{r_0} X_i^0(x) S_{ij}^0 W_j^0(v) = X^{0,\top}(x) \mathbf{S}^0 V^0(v)$ at time t_0 the augmented BUG integrator computes the solution at time $t_1 = t_0 + \Delta t$ as follows:

- (1) **K-step:** Update and augment the spatial basis from $\{X_i^0(x)\}_{i=1}^{r_0}$ to $\{\tilde{X}_i^1\}_{i=1}^{2r_0}$ by solving

$$\partial_t K_j = \langle W_j^0, \mathcal{F}_K \rangle_v, \quad K_j(t_0, x) = \sum_{i=1}^{r_0} X_i^0(x) S_{ij}^0,$$

where $\mathcal{F}_K(t, x, v) := \mathcal{F}(t, \sum_{j=1}^r K_j(t, x) W_j^0(v))$. Compute $\{\tilde{X}_i^1(x)\}_{i=1, \dots, 2r_0}$ as an orthonormal basis of $\{K_1^1(x), \dots, K_{r_0}^1(x), X_1^0(x), \dots, X_{r_0}^0(x)\}$ and store the matrix $\mathbf{M} \in \mathbb{R}^{2r_0 \times r_0}$ with entries $(\mathbf{M})_{\alpha\beta} = \langle \tilde{X}_\alpha^1, X_\beta^0 \rangle_x$.

- (2) **L-step:** Update and augment the velocity basis from $\{W_j^0(v)\}_{j=1}^{r_0}$ to $\{\tilde{W}_j^1\}_{j=1}^{2r_0}$ by solving

$$\partial_t L_i = \langle X_i^0, \mathcal{F}_L \rangle_x, \quad L_i^0(v) = \sum_{j=1}^{r_0} S_{ij}^0 W_j^0(v),$$

where $\mathcal{F}_L(t, x, v) := \mathcal{F}(t, \sum_{i=1}^r X_i^0(x) L_i(t, v))$. Compute $\{\tilde{W}_j^1\}_{j=1, \dots, 2r_0}$ as an orthonormal basis of $\{L_1^1, \dots, L_{r_0}^1, W_1^0, \dots, W_{r_0}^0\}$ and store the matrix $\mathbf{N} \in \mathbb{R}^{2r_0 \times r_0}$ with entries $(\mathbf{N})_{\alpha\beta} = \langle \tilde{W}_\alpha^1, W_\beta^0 \rangle_v$.

- (3) Define $\tilde{\mathbf{S}}^0 = \mathbf{M}\mathbf{S}^0\mathbf{N}^\top \in \mathbb{R}^{2r_0 \times 2r_0}$, where $\mathbf{S}^0 = (S_{ij}^0) \in \mathbb{R}^{r_0 \times r_0}$. Update the coefficients to $\tilde{\mathbf{S}}^1 = \tilde{\mathbf{S}}(t_1) \in \mathbb{R}^{2r_0 \times 2r_0}$ by solving

$$\frac{d}{dt}\tilde{S}_{ij}(t) = \langle \tilde{X}_i^1 \tilde{W}_j^1, \mathcal{F}_S \rangle_{x,v}, \quad \tilde{S}_{ij}(t_0) = (\tilde{\mathbf{S}}^0)_{ij},$$

where $\mathcal{F}_S(t, x, v) := \mathcal{F}(t, \sum_{i,j=1}^r \tilde{X}_i^1(x) \tilde{S}_{ij}(t) \tilde{W}_j^1(v))$.

- (4) **Truncation:** Compute the SVD $\tilde{\mathbf{S}} = \tilde{\mathbf{P}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{T}}^\top$, where $\tilde{\mathbf{\Sigma}} = \text{diag}(\sigma_k)$ contains the singular values of $\tilde{\mathbf{S}}$. Choose the new rank $r_1 \leq 2r_0$ as the smallest number satisfying:

$$\left(\sum_{k=r_1+1}^{2r_0} \sigma_k^2 \right)^{1/2} \leq \vartheta.$$

For $i, j = 1, \dots, r_1$, set $S_{ij}^1 = \sigma_i \delta_{ij}$, $X_i^1(x) = \sum_{k=1}^{2r_0} \tilde{X}_k^1(x) \tilde{P}_{ki}$, and $W_j^1(v) = \sum_{k=1}^{2r_0} \tilde{W}_k^1(v) \tilde{T}_{kj}$

Then the factorized solution at time t_1 with rank r_1 is given by $\sum_{i,j=1}^{r_1} X_i^1(x) S_{ij}^1 V_j^1(v)$.

Remark 2.1. Note that in practice, to update the low-rank factors, we need to define an appropriate discretization of the spatial and velocity domains and the corresponding inner products.

2.2. Rank-adaptive dynamical low-rank Monte Carlo (RaDLR-MC) estimator

In the probabilistic setting, we treat the random variable as a parameter of the kinetic equation and assume that the solution to (1.1) admits the following uncertainty-dependent low-rank ansatz:

$$u(t, x, v; \omega) \approx u_r(t, x, v; \omega) = \sum_{i,j=1}^r X_i(t, x; \omega) S_{ij}(t; \omega) W_j(t, v; \omega). \quad (2.8)$$

Thus, for each realization of the random variable, we can compute the low-rank approximation by repeatedly applying the augmented BUG integrator with a given truncation tolerance ϑ . This estimator can also be extended to any (rank-adaptive) BUG integrator with a pre-defined truncation tolerance ϑ in a straightforward way. Let Q_r denote the low-rank approximation to the QoI Q , computed with the augmented BUG integrator. To estimate $\mathbb{E}[Q]$ defined in (1.2), we construct and compute the estimator \hat{Q}_r to $\mathbb{E}[Q_r]$. For completeness, we state the RaDLR-MC estimator. Let $\omega_1, \dots, \omega_M$ be M independent and identically distributed (i.i.d.) random realizations of ω . Then, we define the RaDLR-MC estimator as

$$\hat{Q}_r^{MC} := \frac{1}{M} \sum_{i=1}^M Q_r^{(i)}$$

where $Q_r^{(i)} := Q_r(\omega_i)$ denotes the QoI approximated with the augmented BUG integrator with truncation tolerance ϑ for the i^{th} i.i.d. copy of ω . We quantify the accuracy of the

LOW-RANK MLMC FOR KINETIC EQUATIONS

estimator with the mean squared error (MSE) defined as

$$e\left(\widehat{Q}_r^{MC}\right)^2 := \mathbb{E}\left[\left\|\widehat{Q}_r - \mathbb{E}[Q]\right\|_H^2\right].$$

The estimator's dependence on the truncation tolerance ϑ can be made more explicit by replacing the subscript r by ϑ in u_r and Q_r . However, to avoid introducing new notation, we continue with the subscript r to denote the low-rank approximation. In the context of rank-adaptive integrators, r should be interpreted as a varying quantity that depends on ϑ .

Unlike the fixed-rank augmented BUG integrator in [50], the rank in the augmented BUG integrator evolves as a function of the truncation tolerance ϑ . Although this approach improves adaptability to the underlying problem dynamics and uncertainty, selecting an appropriate value for ϑ demands a thorough understanding of the integrator, the problem at hand, and the numerical method employed. For instance, the parallel BUG integrator requires a smaller tolerance compared to the augmented BUG integrator to get comparable results for the same test case [9]. Moreover, if $R(t; u_r, \vartheta)$ denotes the rank of the approximation u_r at time t for truncation tolerance ϑ , then in the probabilistic setting R is a random variable. Thus, the error in the low-rank approximation may vary from highly accurate to highly inaccurate across samples.

2.2.1. Evaluating the bias of augmented BUG

To gain more insight into the relation between the low-rank approximation and truncation tolerance, we evaluate the bias of the RaDLR-MC estimator. The MSE of the RaDLR-MC estimator can be decomposed as

$$e\left(\widehat{Q}_r^{MC}\right)^2 = \mathbb{E}\left[\left\|\frac{1}{M}\sum_{i=1}^M(Q_r^{(i)} - \mathbb{E}[Q_r])\right\|_H^2\right] + \|\mathbb{E}[Q_r - Q]\|_H^2,$$

where we have used the fact that H is a Hilbert space with the inner product norm $\|\cdot\|_H^2 = \langle \cdot, \cdot \rangle_H$ and

$$\mathbb{E}\left[\left\langle \widehat{Q}_r^{MC} - \mathbb{E}[Q_r], \mathbb{E}[Q_r] - \mathbb{E}[Q] \right\rangle_H\right] = \underbrace{\langle \mathbb{E}[\widehat{Q}_r^{MC}] - \mathbb{E}[Q_r], \mathbb{E}[Q_r] - \mathbb{E}[Q] \rangle_H}_{=\mathbb{E}[Q_r]} = 0$$

Thus, the MSE can be written in terms of the estimator variance V_r and bias B_r as

$$e\left(\widehat{Q}_r^{MC}\right)^2 = \frac{1}{M}V_r + B_r^2,$$

where

$$V_r := \mathbb{E}[\|Q_r^{(1)} - \mathbb{E}[Q_r]\|_H^2] \quad \text{and} \quad B_r := \|\mathbb{E}[Q_r - Q]\|_H.$$

We can further decompose the bias B_r into contributions from the spatio-velocity discretization and the low-rank approximation. Let u_{full} denote the approximation to (1.1) computed using the same numerical discretization method (without DLRA) as for the low-rank approximation u_r , and let Q_{full} denote the corresponding QoI. We assume that u_{full} and u_r have the exact same spatio-velocity discretization parameters, i.e., $\Delta x_{full} = \Delta x = \Delta x_r$ and $\Delta v_{full} = \Delta v = \Delta v_r$, where $\Delta x \in \mathbb{R}_{>0}^d$ and $\Delta v \in \mathbb{R}_{>0}^d$ are the

spatial and velocity discretization parameters. Adding and subtracting Q_{full} in the bias and using the triangle inequality, we obtain

$$B_r \leq \underbrace{\|\mathbb{E}[Q_r - Q_{full}]\|_H}_{\text{low-rank contr.}} + \underbrace{\|\mathbb{E}[Q_{full} - Q]\|_H}_{\text{spatio-velocity contr.}}.$$

To further simplify the inequality, we use Jensen's inequality and the Lipschitz continuity of Q , which yields

$$B_r \leq L \cdot \mathbb{E}[\|u_r - u_{full}\|] + L \cdot \mathbb{E}[\|u_{full} - u\|],$$

where $\|\cdot\|$ is an appropriate norm for the approximations u_r and u_{full} to (1.1), for instance the L^2 norm.

The error contribution of the spatio-velocity discretization can be estimated using standard theory on numerical methods. For an order $\alpha \geq 1$ method in space and order $q \geq 1$ in velocity domain, we assume that the error contribution of the spatio-velocity discretization has the form

$$\|u - u_{full}\| \leq K_1(\omega)h_x^\alpha + K_2(\omega)h_v^q,$$

where $h_x := \|\Delta x\|_\infty$, $h_v := \|\Delta v\|_\infty$, and $\|\cdot\|_\infty$ is the infinity norm.

In [8], the authors prove rigorous error bounds for the augmented BUG integrator that are robust to the presence of small singular values (over-approximation of the rank). However, these error bounds use the Lipschitz continuity of the right-hand side operator \mathcal{F} . For kinetic equations, the Lipschitz constant of the right-hand side scales as $\mathcal{O}(1/h_x)$, rendering the error bounds invalid.

While no rigorous bounds exist for the augmented BUG integrator without the Lipschitz constant, from experiments on deterministic problems, it is known that the error tends to respect the following bound

$$\|u_r(t_s, \cdot, \cdot) - u_{full}(t_s, \cdot, \cdot)\| \leq K_3\Delta t + K_4s\vartheta + K_5\epsilon,$$

where $t_s = t_0 + s\Delta t$ and $\Delta t = t_{\text{end}}/N_t$ for some $N_t \in \mathbb{N}$, ϵ is the low-rank projection error, and K_3, K_4, K_5 are independent of the singular values of u_r . Overall, the bias term has the following structure

$$B_r \leq L(\widehat{K}_1 h_x^\alpha + \widehat{K}_2 h_v^q + \widehat{K}_3 \Delta t + \widehat{K}_4 s\vartheta + \widehat{K}_5 \epsilon),$$

where L is the Lipschitz constant of Q and \widehat{K}_i , $i \in \{1, 2, 3, 4, 5\}$, are expectations of K_i over $p(\omega)$.

2.2.2. Cost analysis of the RaDLR-MC estimator

We evaluate the cost of achieving an MSE of ε^2 for the RaDLR-MC estimator, i.e.,

$$\begin{aligned} e\left(\widehat{Q}_r^{MC}\right)^2 &\leq \varepsilon^2, \\ \implies \frac{1}{M}V_r + B_r^2 &\leq \varepsilon^2. \end{aligned}$$

For $0 < \theta < 1$, the contribution of the statistical error and numerical error can be distributed as

$$\frac{1}{M}V_r \leq \theta \cdot \varepsilon^2, \quad B_r \leq \sqrt{1-\theta} \cdot \varepsilon.$$

This splitting, of the contributions to the MSE, leads to a trade-off between the accuracy of the approximation and the variance of the estimator. Usually, one is interested in computing a low-variance estimate of an accurate approximation of the QoI. In terms of the low-rank approximation error, this implies that the solution is computed on a fine grid in the phase space with a small truncation tolerance to accurately capture the solution. Moreover, $M \geq V_r/(\theta \cdot \varepsilon^2)$ samples must be performed to achieve the required low variance. Thus to construct such an estimator, we need to completely understand the relation between the bias and its various contributions. The behavior of the approximation error due to spatial and velocity discretizations is a well studied subject and can be found in any standard textbook on the topic. See [36, 6] for instance. A less understood relation, however, is that between the approximation error and the truncation tolerance ϑ . To study this for the augmented BUG integrator, we consider the linesource test case, described in [8].

In Figure 1, we plot the absolute error of the low-rank approximation of the QoI, Q_r , to the full rank approximation Q_{full} against the truncation tolerance of the augmented BUG integrator for various spatial grids and velocity discretizations.

From Figure 1 we see that, the error depends on both the spatio-velocity discretization and the truncation tolerance. Moreover, for a fixed spatial and velocity discretization, it is not possible to achieve a target MSE by varying the truncation tolerance alone (and hence the approximation rank), unless the spatio-velocity error is already bounded by $\varepsilon/\sqrt{2}$. This limitation extends to the low-rank control variates strategy proposed in [50]. Nevertheless, Figure 1 shows that refining the spatial grid at a fixed velocity resolution reduces the approximation error, and thus the bias and overall MSE of the estimator. We exploit this observation in the next section to construct a spatially accurate augmented BUG integrator and an adaptive estimator for kinetic equations.

3. A rank-adaptive framework for low-rank multilevel Monte Carlo

Exploiting spatial refinement as the primary mechanism for controlling approximation error, we construct a multilevel estimator over a hierarchy of spatial discretizations, each paired with a low-rank approximation of the kinetic equation. The intuition is straightforward: coarse levels are cheap to evaluate and capture the bulk of the solution structure at low rank, while successive refinements incrementally correct the bias introduced by both spatial discretization and rank truncation. The multilevel telescoping sum then aggregates these corrections, and the sample allocation across levels can be tuned to balance statistical and approximation errors. The outcome is an estimator that pairs the well-known variance reduction of MLMC with the computational savings of low-rank compression in phase space; a combination particularly well-suited to the high-dimensional setting of kinetic equations. We call this the *rank-adaptive dynamical low-rank multilevel Monte Carlo* (RaDLR-MLMC) estimator.

3.1. A spatially accurate low-rank approximation

The error of the low-rank approximation of the QoI is governed by four discretization parameters: spatial cell size Δx , velocity grid cell size Δv , time step size Δt , and truncation

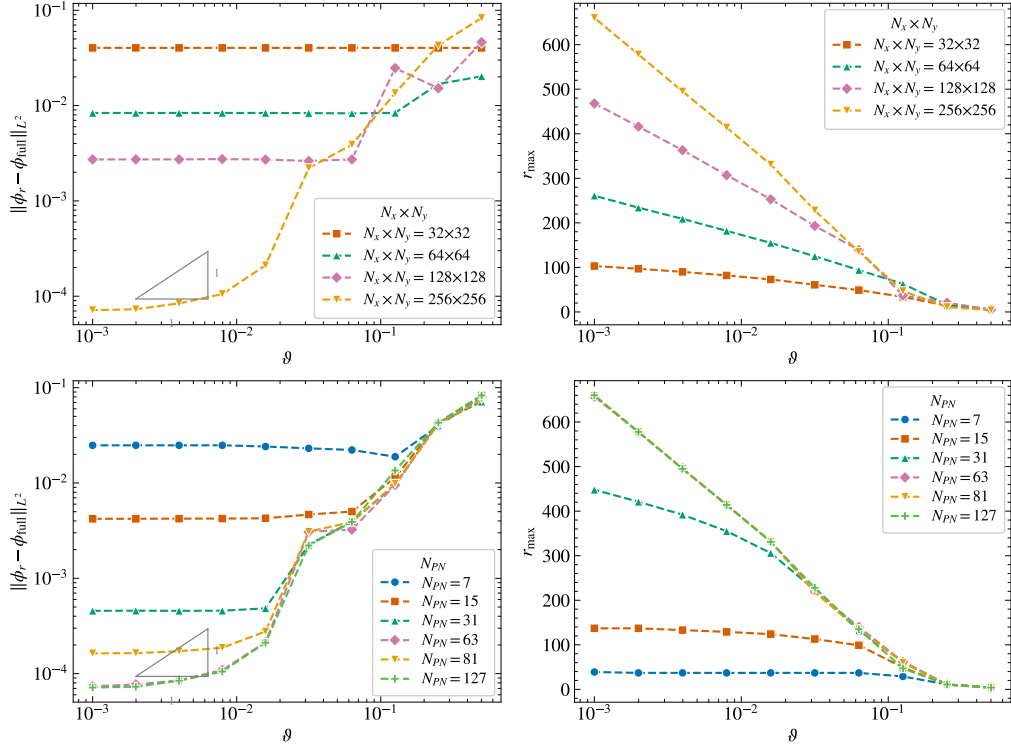


FIGURE 1. Dependence of the low-rank approximation error and maximum rank on the truncation tolerance ϑ for the line source test case [8]. Here $Q(x) := \phi(t = 1.0, x) = \int_{S^2} \psi(t = 1.0, x, \Omega) d\Omega$ denotes the scalar flux and $\Omega = v/|v|$ is the direction of flight. (Top) Spatial refinement study: discrete L^2 error of ϕ_r against a full-rank reference ϕ_{full} on a 256×256 grid (left) and maximum rank r_{max} over all time steps (right), for spatial grids ranging from 32×32 to 256×256 cells with fixed angular resolution. (Bottom) Angular refinement study: same quantities for angular resolutions ranging from $N = 8$ to $N = 128$ with fixed spatial grid.

tolerance ϑ (recall Figure 1). To lay the groundwork for the multilevel estimator, we first describe how spatial refinement can be incorporated into the low-rank approximation framework.

We use an explicit Euler time-stepping scheme and for stability considerations, the time step size is chosen to satisfy the following Courant-Friedrichs-Lewy (CFL) condition

$$\Delta t \leq \frac{\text{CFL}}{|v|\sigma} \left(\frac{\prod_{i=1}^d \Delta x_i}{\sum_{i=1}^d \prod_{i \neq j} \Delta x_j} \right), \quad (3.1)$$

where $|v| := \sup(\mathcal{V}) < \infty$. Note that the CFL condition may contain additional terms, like the scattering or absorption coefficient for radiation transport. These problem-specific terms are incorporated in σ . On a uniform spatial grid $\Delta x_i = h_x$, for $i = 1, \dots, d$, the product $\prod_{i=1}^d \Delta x_i = h_x^d$ and the sum of products is $\sum_{i=1}^d \prod_{i \neq j} \Delta x_j = d h_x^{d-1}$, so the

CFL condition reduces to

$$\Delta t \leq \frac{\text{CFL}}{d|v|\sigma} h_x. \quad (3.2)$$

For an order α method in space, choosing the truncation tolerance to satisfy

$$\vartheta \leq \frac{\widehat{K}_1}{\widehat{K}_4} \left(\frac{\text{CFL}}{d\sigma|v|t_{end}} \right) h_x^{\bar{\alpha}}, \quad (3.3)$$

where $\bar{\alpha} := \min\{\alpha + 1, 2\}$, ensures that the low-rank approximation error does not dominate the spatial discretization error. Since we use an asymptotically first order method in time, using a higher-order spatial discretization does not yield additional benefits. This could be mitigated by using a higher-order BUG integrator [7, 31, 47] in conjunction with a higher-order spatial discretization.

Substituting the CFL-constrained Δt and the bound on ϑ into the general error estimate, and assuming a uniform spatial grid, the dominant term in the bias reduces to

$$B_r \leq \widehat{K}_6 h_x + \widehat{K}_2 h_v^q + \widehat{K}_5 \epsilon, \quad (3.4)$$

where

$$\widehat{K}_6 = L \cdot \max \left\{ \widehat{K}_1, \widehat{K}_3, \widehat{K}_1 \left(\frac{\text{CFL}}{d\sigma|v|t_{end}} \right) \right\}.$$

3.2. The RaDLR-MLMC method and algorithm

Finally, we state the rank-adaptive dynamical low-rank multilevel Monte Carlo (RaDLR-MLMC) estimator for kinetic equations. We see from (3.4) that for a high-order method in velocity ($q > 1$) and small low-rank projection error $\epsilon \ll 1$, the bias is dominated by the spatial error. The assumption that the low-rank projection error is non-dominant remains valid for the problems considered in this work which require the solution to admit a low-rank structure. Thus, we use a high-order method in the velocity domain such that the spatial error remains the dominant contribution to the bias term and construct the multilevel hierarchy in space. Let $h_{x,\ell} := h_{x,0}/2^\ell$ denote the width of the uniform spatial grids, where $\ell = 0, 1, \dots, L$ are referred to as levels and $h_{x,0}$ is the size of a spatial cell on the coarsest level. Furthermore, we let

$$\vartheta_\ell := \frac{\widehat{K}_1}{\widehat{K}_4} \left(\frac{\text{CFL}}{d\sigma|v|t_{end}} \right) h_{x,\ell}^{\bar{\alpha}}, \quad \text{and} \quad \Delta t_\ell := \left(\frac{\text{CFL}}{d\sigma|v|} \right) h_{x,\ell}.$$

If we define the low-rank approximation on the grid parametrized by $(\Delta t_\ell, h_{x,\ell}, \Delta v, \vartheta_\ell)$ by Q_ℓ , $\ell = 0, 1, \dots, L$, then $\mathbb{E}[Q_L]$, on the finest level L , can be written as the telescoping sum

$$\mathbb{E}[Q_L] = \mathbb{E}[Q_0] + \sum_{\ell=1}^L (\mathbb{E}[Q_\ell] - \mathbb{E}[Q_{\ell-1}]) = \sum_{\ell=0}^L \mathbb{E}[\Delta Q_\ell], \quad (3.5)$$

where we define the random variable $\Delta Q_\ell(\omega) := Q_\ell(\omega) - Q_{\ell-1}(\omega)$, with $Q_{-1}(\omega) \equiv 0$. The basic principle of the MLMC method is to estimate the QoI on the finest level L by sampling from ΔQ_ℓ on level $\ell = 0, \dots, L$.

The RaDLR-MLMC estimator for $\mathbb{E}[Q_L]$ based on the telescoping sum (3.5) is defined as

$$\widehat{Q}_L^{ML} := \sum_{\ell=0}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \Delta Q_\ell^{(i)}, \quad (3.6)$$

where M_ℓ is the sample size on level ℓ and $\Delta Q_\ell^{(i)}$, $i = 1, \dots, M_\ell$, are M_ℓ i.i.d. copies of ΔQ_ℓ . The MSE of the RaDLR-MLMC estimator is given by

$$e\left(\widehat{Q}_L^{ML}\right)^2 = \mathbb{E}\left[\left\|\widehat{Q}_L^{ML} - \mathbb{E}[Q]\right\|_H^2\right] = \mathbb{E}\left[\left\|\widehat{Q}_L^{ML} - \mathbb{E}[Q_L]\right\|_H^2\right] + \|\mathbb{E}[Q_L - Q]\|_H^2.$$

Inserting the MLMC estimator (3.6) and the telescoping sum (3.5) in the MSE, we obtain

$$\begin{aligned} e\left(\widehat{Q}_L^{ML}\right)^2 &= \mathbb{E}\left[\left\|\sum_{\ell=0}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \left(\Delta Q_\ell^{(i)} - \mathbb{E}[\Delta Q_\ell]\right)\right\|_H^2\right] + \|\mathbb{E}[Q_L - Q]\|_H^2 \\ &= \sum_{\ell=0}^L \frac{1}{M_\ell} \mathbb{E}\left[\|\Delta Q_\ell - \mathbb{E}[\Delta Q_\ell]\|_H^2\right] + \|\mathbb{E}[Q_L - Q]\|_H^2 \\ &= \sum_{\ell=0}^L \frac{1}{M_\ell} V_\ell + \|\mathbb{E}[Q_L - Q]\|_H^2 \end{aligned}$$

where $V_\ell := \mathbb{E}\left[\|\Delta Q_\ell - \mathbb{E}[\Delta Q_\ell]\|_H^2\right] = \mathbb{V}[\Delta Q_\ell]$ has been defined in (1.3).

The first term of the MSE is called the total variance of the estimator, while the second term is the bias of the estimator. In the MLMC method, the statistical and numerical errors of the estimator are balanced to reach a certain predefined accuracy ε . To be specific, for a given $\varepsilon > 0$, we want to choose the finest discretisation level L and the number of samples M_ℓ on each level such that the MSE of the RaDLR-MLMC estimator satisfies $e\left(\widehat{Q}_L^{ML}\right)^2 \leq \varepsilon^2$. As discussed in Section 2.2.2, this leads to a trade-off between the error contributions which can be distributed using a parameter $0 < \theta \leq 1$. One way to resolve this is by choosing $\theta = 0.5$, i.e., equally splitting the ε -error between the two error contributions and thus requiring that the total variance and bias be at most $\varepsilon^2/2$. This splitting of the error is not necessarily the most optimal strategy [24], but rather a pragmatic compromise.

If C_ℓ is the cost of computing a single sample of $\Delta Q_\ell^{(i)}$, then the total cost of the RaDLR-MLMC estimator is given by

$$c\left(\widehat{Q}_L^{ML}\right) = \sum_{\ell=0}^L M_\ell C_\ell.$$

If we have an a priori estimate of the bias, then we can find an optimal L satisfying $\|\mathbb{E}[Q_L - Q]\|_H^2 \leq \varepsilon^2/2$. In general, the constants are not known a priori, and hence L is chosen on-the-fly by estimating the bias during computation as described in [12].

Assuming that we know the final level L , the number of samples on each level is chosen such that the total cost is minimized while the total variance is bounded by $\varepsilon^2/2$. That

Algorithm 1 RaDLR-MLMC algorithm

- 1: Initialize $L = 2$ with initial samples M_ℓ on levels $\ell = 0, 1, 2$
 - 2: Initialize $K_\ell = 0$ to be the count of already computed samples on level ℓ
 - 3: **while** $K_\ell < M_\ell$ **do**
 - 4: **for** $\ell = 1, \dots, L$ **do**
 - 5: On level ℓ , generate additional samples of $\Delta Q_\ell^{(i)}$ for $i = 1, \dots, M_\ell - K_\ell$ with parameters $(\Delta t_\ell, h_\ell, \Delta v, \vartheta_\ell)$ and the augmented BUG integrator
 - 6: Updates estimates of V_ℓ for $\ell = 0, \dots, L$
 - 7: Compute the optimal number of samples M_ℓ on each level $\ell = 0, 1, \dots, L$ according to (3.7)
 - 8: **if** Test for weak convergence of the QoI **then**
 - 9: If weak convergence fails, set $L := L + 1$, initialize M_L , and $K_L = 0$
 - 10: **end if**
 - 11: **end for**
 - 12: **end while**
-

is, the optimal number of samples of level ℓ is given by [12]

$$M_\ell = \left\lceil 2\varepsilon^{-2} \sqrt{V_\ell/C_\ell} \left(\sum_{\ell'=0}^L \sqrt{V_{\ell'} C_{\ell'}} \right) \right\rceil, \quad (3.7)$$

where $\lceil \cdot \rceil$ is the ceiling function defined as $\lceil a \rceil = \min\{m \in \mathbb{Z} \mid m \geq a\}$ given any $a \in \mathbb{R}$.

Generating addition samples $\Delta Q_\ell^{(i)}$ on level ℓ in the line 4 of algorithm 1, requires solving uncertain kinetic equation (1.1) on levels ℓ and $\ell - 1$, with the same realization of the random variable ω . Since, in most cases, the variance V_ℓ is not known a priori, it must be estimated as,

$$V_\ell = \mathbb{E} \left[\|\Delta Q_\ell - \mathbb{E}[\Delta Q_\ell]\|_H^2 \right] \approx \frac{1}{M_\ell - 1} \left(\sum_{i=1}^{M_\ell} \|\Delta Q_\ell^{(i)}\|_H^2 - \frac{1}{M_\ell} \left\| \sum_{i=1}^{M_\ell} \Delta Q_\ell^{(i)} \right\|_H^2 \right).$$

Due to the high dimensionality of the kinetic equations, accumulating samples to estimate the expected value and variance can be infeasible. To avoid this, we use Welford's online algorithm [57] to compute the expected value and variance on-the-fly on each level.

To test for convergence, in line 7 of algorithm 1, we verify if

$$\|\mathbb{E}[Q_L - Q]\|_H < \varepsilon/\sqrt{2}$$

is satisfied for the given tolerance ε . Since the exact solution $\mathbb{E}[Q]$ is unknown a priori, we estimate the error based on estimates of the last levels ΔQ_L . Assuming that the error of the differences behaves as $\|\mathbb{E}[Q_\ell - Q_{\ell-1}]\|_H = \mathcal{O}(2^{-\alpha\ell})$, as $\ell \rightarrow \infty$ for $\alpha \geq 1$, we get

$$\|\mathbb{E}[Q_L - Q]\|_H = \left\| \sum_{\ell=L+1}^{\infty} \mathbb{E}[Q_\ell - Q_{\ell-1}] \right\|_H \leq \sum_{\ell=L+1}^{\infty} \|\mathbb{E}[Q_\ell - Q_{\ell-1}]\|_H \simeq \frac{\|\mathbb{E}[Q_L - Q_{L-1}]\|_H}{2^\alpha - 1},$$

where

$$A \simeq B \iff cB \leq A \leq \hat{c}B, \quad c, \hat{c} > 0.$$

Remark 3.1. For robustness of the bias estimate, we extrapolate the error from the previous two levels $\|\mathbb{E}[Q_{L-1} - Q_{L-2}]\|_H$ and $\|\mathbb{E}[Q_{L-2} - Q_{L-3}]\|_H$, and check if the maximum of the three satisfies the convergence criterion. Note that since this is a heuristic approach that does not provide theoretical guarantees, we choose a very conservative strategy. Our later convergence tests in Figure 2 serve to motivate this and confirm that it is effective in keeping the error within the prescribed accuracy.

Note that the RaDLR-MLMC has been constructed to preserve the spatial accuracy of the underlying discretization method. Thus, we can use the following generic MLMC theorem to characterize the total cost of the RaDLR-MLMC estimate $\widehat{Q}_L^{\text{ML}}$ in a Hilbert space H .

Theorem 3.2 ([22]). *Assume that there exist constants $\alpha, \beta, \gamma > 0$, such that $\alpha \geq \frac{1}{2} \min\{\beta, \gamma\}$ and we have the asymptotic rates*

$$(i) \quad \|\mathbb{E}[Q_\ell - Q]\|_H \leq c_1 2^{-\alpha\ell},$$

$$(ii) \quad V_\ell \leq c_2 2^{-\beta\ell}, \text{ and}$$

$$(iii) \quad C_\ell \leq c_3 2^{\gamma\ell}.$$

Then for any given ε small enough, there exists a total number of level L and the number of samples M_ℓ , $\ell = 0, \dots, L$, such that the MSE satisfies

$$e\left(\widehat{Q}_L^{\text{ML}}\right)^2 < \varepsilon^2,$$

and

$$c\left(\widehat{Q}_L^{\text{ML}}\right) \leq c \begin{cases} \varepsilon^{-2}, & \text{if } \beta > \gamma, \\ \varepsilon^{-2}(\ln \varepsilon)^2, & \text{if } \beta = \gamma, \\ \varepsilon^{-2-\frac{\gamma-\beta}{\alpha}}, & \text{if } \beta < \gamma. \end{cases}$$

The computational complexity predicted by the MLMC theorem above is guaranteed, provided that the assumptions (i)–(iii) are satisfied. These are often difficult to verify theoretically for complex models, and we refer the reader to the numerical experiments section for numerical verifications for various kinetic equations considered in this paper.

Finally, we note that the classical complexity bound of Theorem 3.2 identifies three regimes, depending on the interplay between the variance decay rate β and the cost growth rate γ . The optimal regime $\beta > \gamma$ yields complexity $\mathcal{O}(\varepsilon^{-2})$, matching plain Monte Carlo in the idealized setting where sample cost is fixed, i.e., independent of the discretization level. The borderline case $\beta = \gamma$ incurs only a logarithmic overhead, $\mathcal{O}(\varepsilon^{-2} |\log \varepsilon|^2)$. While these favorable rates are often attainable through a judicious choice of discretization—for instance, when levels correspond solely to different step sizes in a time-stepping scheme—more involved discretizations combining spatial, temporal, and, as in this work, rank-based refinement often fall into the worst-case regime $\beta < \gamma$, particularly for computationally expensive models. It is important to stress, however, that even in this worst case, MLMC still reduces the complexity exponent by β/α relative to plain Monte Carlo, which can translate to savings of several orders of magnitude, as we demonstrate in the numerical experiments section.

LOW-RANK MLMC FOR KINETIC EQUATIONS

Dimensions	\mathcal{D}	\mathcal{V}	$\mathcal{A}(u)\partial_x u$
$d = 2$	$[-3, 3]^2$	$\mathbb{P}_{\mathbb{R}^2}(\mathbb{S}^2)$	$\Omega \cdot \partial_x u$
$d = 1$	$[-3, 3]$	$\mathbb{P}_{\mathbb{R}}(\mathbb{S}^2)$	$\mu \partial_x u$

TABLE 2. Domain of the spatial variable $x \in \mathcal{D}$ and the angular variable $\Omega \in \mathcal{V}$ for the radiative transport equation in one-, and two-dimensional slab geometry. The projection of \mathbb{S}^2 onto \mathbb{R}^2 is defined as $\mathbb{P}_{\mathbb{R}^2}(\mathbb{S}^2) := \left\{ \Omega = (\sqrt{1 - \mu^2} \sin \theta, \sqrt{1 - \mu^2} \cos \theta) \mid \theta \in [0, 2\pi), 0 \leq \mu \leq 1 \right\}$ and $\mathbb{P}_{\mathbb{R}}(\mathbb{S}^2) := \{ \mu \mid -1 \leq \mu \leq 1 \}$.

Remark 3.3. In case of multiple QoIs, say $\bar{Q} = (Q_1, \dots, Q_k)$, we estimate the variance and expected difference on level ℓ as the maximum of the individual variances and expected differences. That is,

$$V_\ell := \max_{i=1, \dots, k} \{V_{\ell,1}, \dots, V_{\ell,k}\},$$

$$\|\mathbb{E}[\bar{Q}_L - \bar{Q}_{L-1}]\|_H := \max_{i=1, \dots, k} \{\|\mathbb{E}[Q_{L,i} - Q_{L-1,i}]\|_H\},$$

where $V_{\ell,i}$ and $\|\mathbb{E}[Q_{L,i} - Q_{L-1,i}]\|_H$ are the variance and expected differences on level ℓ for the i^{th} QoI. This strategy is a conservative choice that doesn't strictly fit into the setting of Theorem 3.2; however, it preserves the algorithmic guarantees component-wise.

4. Numerical results

In the following, we present numerical experiments ranging from a one-dimensional test case, for which error convergence and computational costs can be assessed against a semi-analytical solution, to a challenging two-dimensional radiation transport benchmark, a high-dimensional radiation therapy dose calculation, and finally the hyperbolic shallow water moment equations, demonstrating the generality of our method beyond radiation transport. In all experiments, the velocity variable is discretized using the P_N method [6] with n basis functions and the MLMC methods use 10 warm-up samples except the shallow water equations and radiation therapy in which we use 20 and 50, respectively.¹

4.1. Radiation transport

We consider a one-speed radiation transport equation such that $\Omega \in \mathcal{V} = \mathbb{S}^{d-1}$ is the direction of flight of a particle traveling at speed v_0 . Without loss of generality we assume $v_0 = 1$. The spatial and angular domains \mathcal{D} and \mathcal{V} and the advection operator $\mathcal{A}(u)\partial_x u$ are specified in Table 2 for the one- and two-dimensional slab geometry. Note that in the one-dimensional slab geometry, we replace the angular variable Ω by μ , which corresponds to the projection of \mathbb{S}^2 onto \mathbb{R} under symmetry assumptions.

The collision operator is assumed to be isotropic. It is given by $\mathcal{S}(u; \omega) := \sigma_s \left(\int_{\mathcal{V}} u \, d\Omega - u \right) - \sigma_a u + s$, where s is the source term. In the collision operator, σ_s denotes the scattering

¹All numerical experiments presented in this section were conducted using our open-source Julia implementation, available at <https://github.com/chinsp/publication-RaDLR-MLMC-for-Kinetic-Equations.git>.

coefficient and σ_a the absorption coefficient specifying the probability of scattering and absorption of particles, respectively.

4.1.1. Gaussian pulse test case

In the first test case, we consider the uncertain Gaussian pulse test case with an uncertain scattering coefficient [3, 4, 20]. The QoI for this test case is the scalar flux at time $t = 1$ which we denote by $Q(x) := \int_{-1}^1 u(t = 1.0, x, \mu; \omega) d\mu$. The scattering coefficient has the form, $\sigma_s = \bar{\sigma}_s(1 + 0.1\omega)$, where $\omega \sim U(-1, 1)$ and $\bar{\sigma}_s = 1.0$ is the scattering coefficient. That is, we assume that there is 10% uncertainty in the probability of scattering of particles. The initial condition is a Gaussian pulse centered at $x = 0$

$$u(t = 0, x, \mu; \omega) = \exp(-x^2/\kappa^2),$$

where κ is a given constant and \mathcal{D} and \mathcal{V} correspond to the first row in Table 2. For this problem, a semi-analytic solution is available and is computed using the gPC-sG expansion [3, 4]. It serves as a reference for computing the errors of the Monte Carlo-based estimators. We denote the semi-analytic solution by $\mathbb{E}[Q_{\text{sa}}]$.

To study the convergence behavior and computational costs of the low-rank based estimators, we compare the RaDLR-MC, full-MLMC (without DLRA), and RaDLR-MLMC with varying tolerances ε to the semi-analytic solution on a reference grid of size $h_{x,ref} = 5.859375 \times 10^{-3}$. The coarsest discretization parameters for the MLMC estimators are $h_{x,0} = 0.375$ with $n = 501$ basis functions for angular discretization.

For a fair comparison, we discretize the RaDLR-MC to match the finest level of the full-rank MLMC and the RaDLR-MLMC for a specified tolerance ε . Notably, the finest levels for both estimators coincided across all experiments. The relative error of the estimators and their corresponding costs (minimum over 5 runs) for varying tolerances ε are shown in Figure 2.

From Figure 2, we see the bias of all three methods follows the expected curve as ε is varied. Moreover, by comparing the total computation cost of the estimator we see that the RaDLR-MC, at the finest level, is more expensive than the MLMC estimators at the same relative error. At large tolerances, the full-rank MLMC is computationally cheaper than RaDLR-MLMC which can be attributed to the additional costs of computing QR decompositions in the augmented BUG integrator [8]. At lower tolerances, the finest level of the MLMC estimators is large enough that the computational cost of computing a full-rank sample is much higher than the additional cost of QR decomposition. Thus, the RaDLR-MLMC allows for a computationally cheaper estimate of the expected value of the QoI. These effects are more pronounced in higher-dimensions wherein, for certain problems, it is infeasible to compute a single sample of the full-rank solution.

4.1.2. Lattice test case

Next, we consider an uncertain lattice test case for the radiation transport equation in two-dimensional slab geometry, which is an important benchmark in the radiative transfer community [34, 54]. The geometry of the lattice is given in Figure 3, where the colored blocks represent a strongly absorbing material, while the white regions are treated as weak scatterers of radiation. A constant source placed in the central block

LOW-RANK MLMC FOR KINETIC EQUATIONS

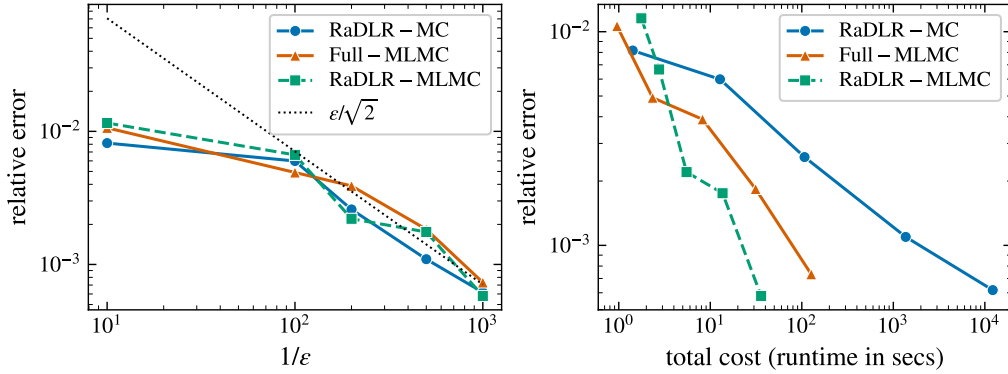


FIGURE 2. Results of the Gaussian pulse test case with $\bar{\sigma}_s = 1.0$ and $\kappa = 0.5$ for the RaDRL-MC (with augmented BUG), full-rank MLMC, and RaDRL-MLMC. The solution is compared to the semi-analytic solution $\mathbb{E}[Q_{sa}]$ computed using gPC-sG expansion [4, 3]. Left: Error of the different estimators with respect to the semi-analytic solution versus the tolerance of the adaptive estimators. Right: Error versus cost (minimum over 5 runs) for the three estimators.

emits particles in all directions isotopically. This radiation is absorbed by the blocks, forming radiation shadows away from the source. We provide the details of the random parameters and refer the reader to [54] for complete details of the test case.

	$\sigma_a^{\text{blue}}(\omega)$	$\sigma_a^{\text{orange}}(\omega)$	$s(t, x, \Omega; \omega)$
Experiment 1	U(9.5, 10.5)	U(9.0, 11.0)	U(0.9, 1.1)
Experiment 2	U(9.0, 11.0)	U(8.5, 11.5)	U(0.9, 1.1)

TABLE 3. Uncertain parameters for each experiment. The blue outer absorbers, orange inner absorbers, and isotropic source are modelled as independent uniform random variables.

We assume that the strength of the absorbing blocks in the lattice and the source is uncertain. We divide the absorbing blocks into two regions: the inner absorbers (σ_a^{orange} ; colored in orange in Figure 3) and the outer absorbers (σ_a^{blue} ; colored in blue in Figure 3). The quantity of interest is the scalar flux $Q(x) := \int_{\mathbb{P}(\mathbb{S}^2)} u(t, x, \Omega; \omega) d\Omega$. The distributions of the random parameters for the test case are given in Table 3. At the coarsest level, we set $h_{x,0} = 0.21875$ with 1600 basis functions for angular discretization.

In Figure 7, we plot $\mathbb{E}[Q_\ell]$ (in log-scale) and the differences $\mathbb{E}[\Delta Q_\ell]$ for the two test cases. We see that the majority of the statistical variation is captured at the coarsest level, while the higher levels correct for the bias in the estimator. We also see that for Experiment 2, with higher uncertainty in the absorbing blocks, the scalar flux shows more oscillations despite being estimated at the same MLMC tolerance ϵ . In Figure 3, we plot the variance and cost of evaluating a sample on level ℓ . We see that $\beta < \gamma$, which is not the most favorable case according to the MLMC theorem. However, based

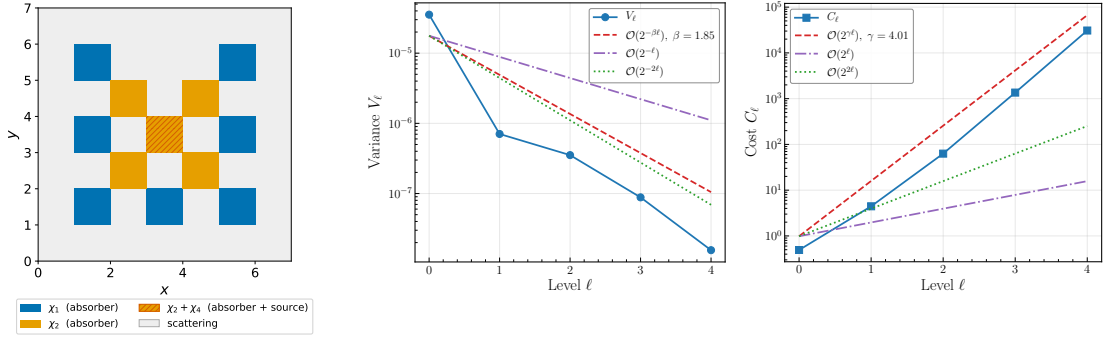


FIGURE 3. Left: geometry and setup of the lattice test case with uncertain absorption and source. Middle: variance of the RaDLR-MLMC estimator for the lattice test case across levels. Right: cost of the RaDLR-MLMC estimator for the lattice test case across levels.

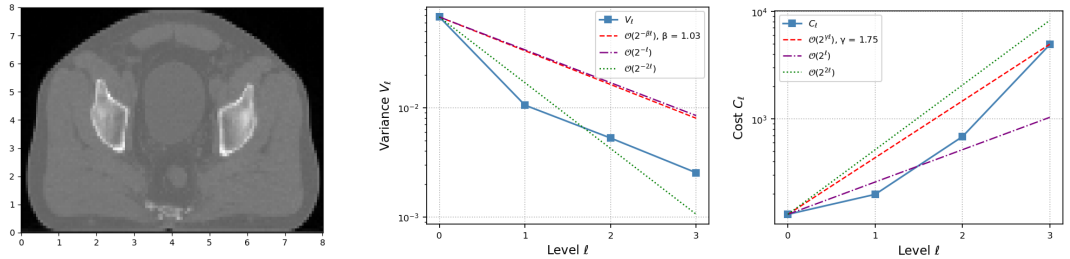


FIGURE 4. Left: CT image. Middle: variance of the RaDLR-MLMC estimator for the radiation therapy test case across levels. Right: cost of the RaDLR-MLMC estimator for the radiation therapy test case across levels.

on experiments, we see that this drop in efficiency across levels is still not as severe as using the full-rank integrators which have a much higher cost per sample.

4.2. Radiation therapy

Next, we consider a more computationally challenging, high-dimensional test case. In radiation therapy, treatment planning requires the computation of energy deposited in a patient's tissue for a specific treatment set up. Patient movements, set up errors, or inaccuracies of available physical and medical data can introduce uncertainties that need to be taken into account [39, 40]. However, already a single simulation of the full six-dimensional transport problem without physical simplifications or model order reduction can be extremely time-consuming. Thus, uncertainty quantification at a high enough resolution is often infeasible in clinical practice [19].

We now consider a simplified prostate patient CT from the CORT data set [14] (see Figure 4), irradiated with two 75 MeV proton beams which are placed in opposing lateral

LOW-RANK MLMC FOR KINETIC EQUATIONS

positions, i.e., at 90° and 270° gantry angle. Further, we assume a positioning error following a normal distribution with standard deviation 3mm. Note, that this choice of irradiation angles and uncertainty is similar to what would be used in application [51, 5], however we have scaled down the size of the CT scan and adapted the energy of the beams accordingly to limit computational costs.

Mathematically, we model proton transport using the continuous slowing down approximation to the steady linear Boltzmann transport equation:

$$\boldsymbol{\Omega} \cdot \nabla_{\mathbf{r}} \psi(E, \mathbf{r}, \boldsymbol{\Omega}) - \frac{\partial \mathcal{S}(E, \mathbf{r}) \psi(E, \mathbf{r}, \boldsymbol{\Omega})}{\partial E} - \frac{1}{2} \frac{\partial^2 T(E, \mathbf{r}) \psi(E, \mathbf{r}, \boldsymbol{\Omega})}{\partial E^2} = \Gamma \psi(E, \mathbf{r}, \boldsymbol{\Omega}), \quad (4.1)$$

where the phase space of the particle density ψ consists of energy $E \in [E_{\min}, E_{\max}] \subset \mathbb{R}_{\geq 0}$, space $\mathbf{r} \in \mathbb{R}^3$ and direction of flight $\boldsymbol{\Omega} \in \mathbb{S}^2$. The stopping power \mathcal{S} describes the average rate of energy loss and straggling T represents the statistical spread in energy loss due to stochastic interactions. The collision operator Γ on the right-hand side depends on the differential and total scattering cross sections Σ_s and Σ_t and is defined as

$$\Gamma \psi = \int_{\mathbb{S}^2} \Sigma_s(E, \mathbf{r}, \boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}) \psi(E, \mathbf{r}, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' - \Sigma_t(E, \mathbf{r}) \psi(E, \mathbf{r}, \boldsymbol{\Omega}). \quad (4.2)$$

We model the incoming physical beams with a standard deviation of 0.3 cm in space and 1% in energy. The random lateral beam shifts $\boldsymbol{\omega} \in \mathbb{R}^2$ are applied in beam's eye view (perpendicular to the beam directions) and follow a normal distribution $\mathcal{N}(\boldsymbol{\omega}; \boldsymbol{\mu}_{\omega}, \mathbf{C}_{\omega})$ with parameters

$$\boldsymbol{\mu}_{\omega} = (0, 0)^T, \mathbf{C}_{\omega} = \begin{bmatrix} 0.3^2 & 0 \\ 0 & 0.3^2 \end{bmatrix}.$$

The resolution on the coarsest level is chosen as 0.2 cm in beam direction (along the y-axis) and 0.4 cm in the other dimensions. More details on the implementation and numerical methods used can be found in [55].

Figure 4 shows the costs per sample and the variance for each level. Similarly to the previous test case here $\beta = 1.03 < 1.75 = \gamma$ which theoretically places us in the least favorable case of the MLMC theorem. However, it is clear from Figure 4 that computing a large number of samples for a Monte Carlo estimate of the expected value at highest resolution is very expensive despite already having greatly reduced the costs by use of the dynamical low-rank approximation. Using the low-rank MLMC approach, only a very small number of samples is required at highest resolution, thus making expected value computations much more feasible. Figure 5 shows the expected energy delivered in the described CT test case as well as the differences and means at each level. Here we can clearly see that while the basic structure of the solution can be represented at coarser levels, the finest levels are necessary to capture essential details in the solution such as the distinct peaks of the two beams.

4.3. Hyperbolic Shallow Water Moment Equations

The hyperbolic shallow water moment equations (HSWME) [30, 29] model free-surface flows arising in applications such as flood prediction, tsunami simulation, and weather forecasting. The HSWME are derived from the Navier–Stokes equations when assuming a polynomial velocity profile along the water depth [30]. The resulting kinetic equations

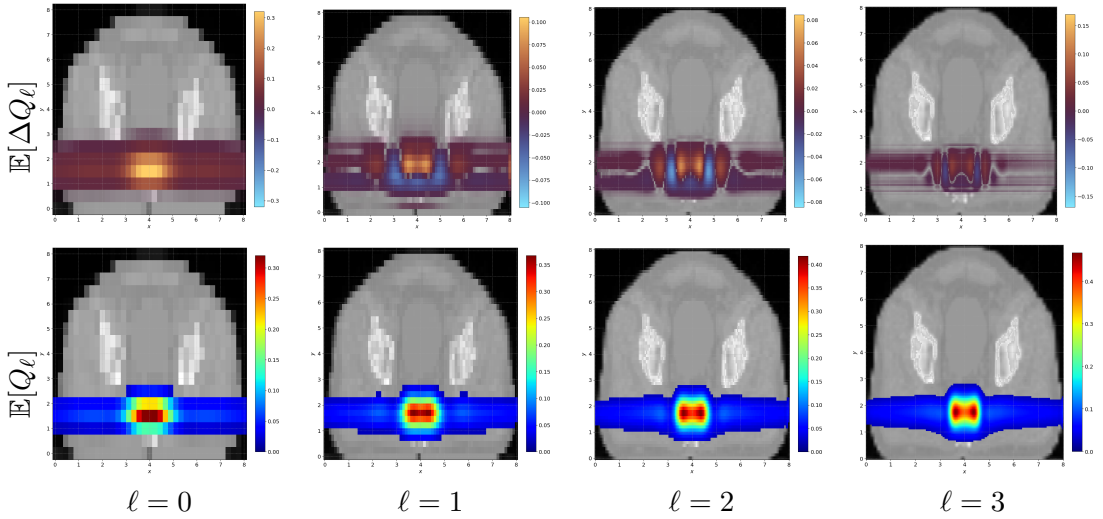


FIGURE 5. Two-dimensional slice along x-y-plane of the quantity of interest (deposited energy) for the radiation therapy test case with uncertain beam positions. Row 1: ΔQ_ℓ for $\ell = 0, 1, 2, 3$. Row 2: $Q_\ell = \sum_{i=0}^{\ell} \Delta Q_i$, $\ell = 0, 1, 2, 3$.

are then regularized to guarantee hyperbolicity as proposed in [29]. Despite being cheaper than simulating incompressible Navier-Stokes equations, the HSWME yet remains expensive to simulate at scale, motivating the use of model order reduction techniques such as DLRA. For a complete model description and details on DLRA for HSWME, we refer the reader to [28].

The macroscopic variables in HMSWE are, the height of the water column $h(t, x; \omega)$, the velocity profile $p(t, x, z)$, and the momentum of the water flow hp_m , where $p_m(t, x; \omega)$ denotes the mean velocity of the water flow. We consider an uncertain dam break test case [28] with independent uncertainty in the initial background height, shock position, and shock amplitude and set the QoI to be $Q := (h, hp_m)^\top$. To e specific, the initial condition is given by

$$h(t = 0, x; \omega) = 0.3\omega_1 + 0.35\omega_2 \cdot (\tanh(50x) - \tanh(50(x - 0.2\omega_3))),$$

where $\omega_i \in U(1 - \delta/100, 1 + \delta/100)$, for $i = 1, 2, 3$, represents $\pm\delta\%$ parametric uncertainty. The flow is assumed to be initially at rest, i.e., $p(t = 0, x, z) = 0$.

The coarsest level is discretized by $h_{x,0} \approx 0.02$ with 102 basis functions for the angular discretization using the method of moments [6]. In Figure 6, we plot the expected value of the height and momentum profiles for the dam break test case for three different MLMC tolerances $\varepsilon = 0.1$, $\varepsilon = 0.01$, and $\varepsilon = 0.005$. These plots show that the uncertainty in the initial condition of the dam break test case significantly effects the expected outcome of the model. These can be simulated and quantified efficiently with the RaDLR-MLMC estimators by choosing an appropriate tolerance. As expected, reducing ε improves the estimate of $\mathbb{E}[Q]$ while only minor effects are seen in the expected velocity profile $\mathbb{E}[hp_m]$.

LOW-RANK MLMC FOR KINETIC EQUATIONS

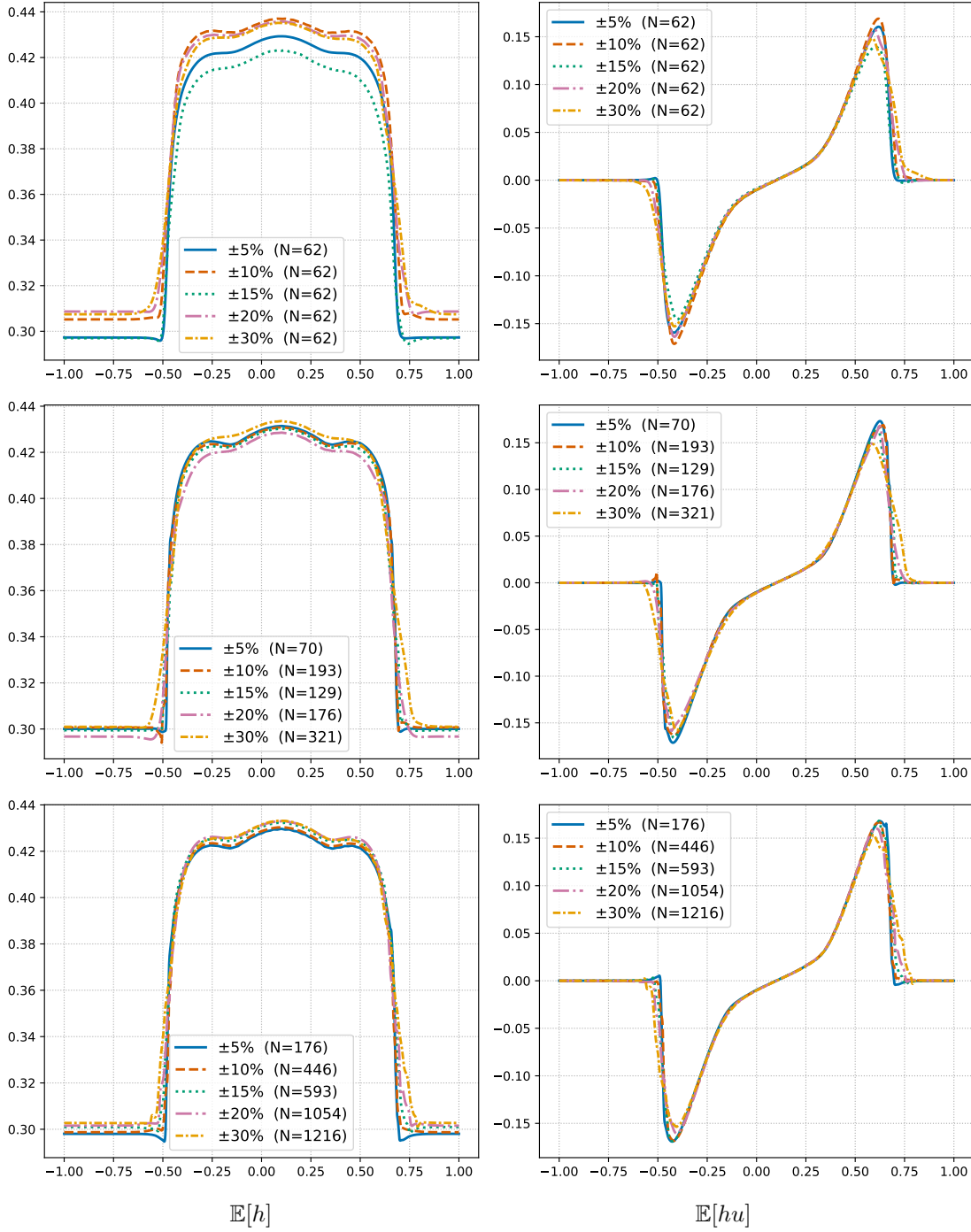


FIGURE 6. MLMC estimates of the water height $\mathbb{E}[h]$ (left) and momentum $\mathbb{E}[hp_m]$ (right) for the shallow water shock problem. Each curve corresponds to a different level of parametric uncertainty $\pm\delta\%$ of the nominal values, ranging from $\pm 5\%$ to $\pm 30\%$ ($\pm\delta\%$). Results are shown for MLMC tolerances $\varepsilon = 0.1$ (top), $\varepsilon = 0.01$ (middle), and $\varepsilon = 0.005$ (bottom). The number of solver calls N required to meet each tolerance is reported in the legend.

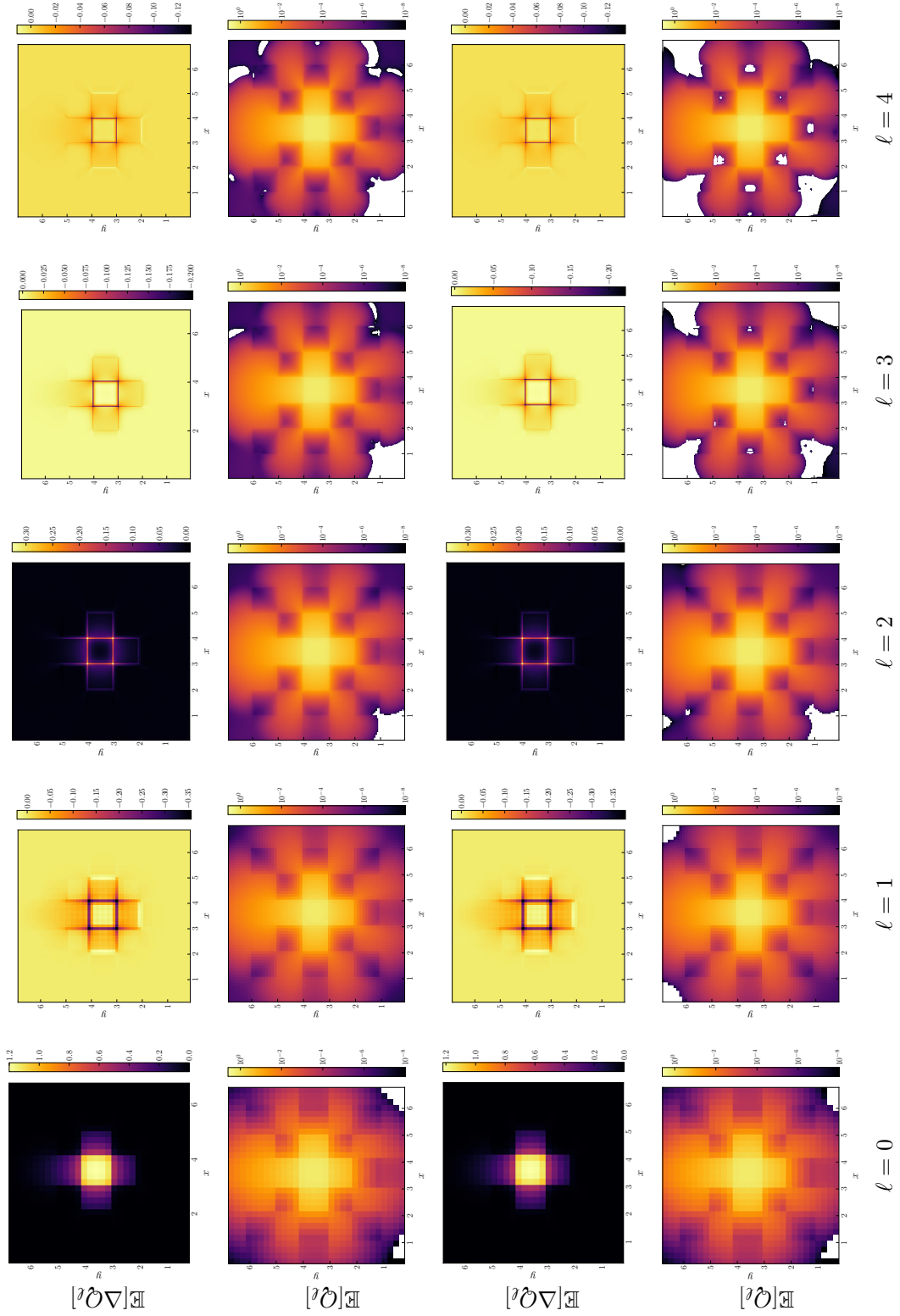


FIGURE 7. Quantity of interest (the scalar flux) for the lattice test case with uncertain absorption and source across levels. Row 1,3: ΔQ_ℓ for $\ell = 0, 1, 2, 3, 4$. Row 2,4: $Q_\ell = \sum_{i=0}^{\ell} \Delta Q_i$, $\ell = 0, 1, 2, 3, 4$, in a log scale. Row 1,2 correspond to experiment 1 in Table 3 while rows 3,4 are for experiment 2. The white regions correspond to values under the minimum cut-off 10^{-8} and negative values.

5. Conclusions and outlook

To overcome the computational bottleneck of sampling-based UQ methods for high-dimensional kinetic equations, we introduced a rank-adaptive dynamical low-rank Monte Carlo estimator that computes a low-rank approximation of the quantity of interest for each sample. A detailed bias analysis revealed a key computational challenge inherent to rank-adaptive integrators: the saturation of errors at smaller tolerances. To address this challenge, we developed a spatially accurate rank-adaptive integrator and leveraged it to construct a rank-adaptive dynamical low-rank multilevel Monte Carlo estimator, significantly reducing the overall computational cost. We demonstrated the efficacy of the proposed method on a range of challenging test cases. For higher-dimensional problems, the ability to obtain reliable estimates of the quantities of interest at all represents a significant achievement although the variance decay and cost growth remain far from ideal. To facilitate reproducibility and broader adoption, we provide an open-source Julia implementation of the method, compatible with any numerical solver, with or without DLRA.

Acknowledgements

The authors would like to thank Martin Frank for his helpful comments and suggestions.

Bibliography

- [1] D. Armbruster, D. Marthaler, and C. Ringhofer. Kinetic and Fluid Model Hierarchies for Supply Chains. *Multiscale Modeling & Simulation*, 2(1):43–61, January 2003.
- [2] Andrea Barth, Annika Lang, and Christoph Schwab. Multilevel Monte Carlo method for parabolic stochastic partial differential equations. *BIT Numerical Mathematics*, 53(1):3–27, March 2013.
- [3] W. Bennett and R. G. McClarren. Uncertainty Benchmarks for Time-Dependent Transport Problems. *Nuclear Science and Engineering*, 199(sup1):53, April 2025.
- [4] William Bennett and Ryan G. McClarren. Benchmarks for Infinite Medium, Time Dependent Transport Problems with Isotropic Scattering. *Journal of Computational and Theoretical Transport*, 51(4):205–221, June 2022.
- [5] Wenhua Cao, Gino J Lim, Yupeng Li, X Ronald Zhu, and Xiaodong Zhang. Improved beam angle arrangement in intensity modulated proton therapy treatment planning for localized prostate cancer. *Cancers*, 7(2):574–584, 2015.
- [6] Kenneth M. Case and Paul Frederick Zweifel. *Linear Transport Theory*. Addison-Wesley Publishing Company, 1967.

- [7] Gianluca Ceruti, Lukas Einkemmer, Jonas Kusch, and Christian Lubich. A robust second-order low-rank BUG integrator based on the midpoint rule. *BIT Numerical Mathematics*, 64(3):30, 2024.
- [8] Gianluca Ceruti, Jonas Kusch, and Christian Lubich. A rank-adaptive robust integrator for dynamical low-rank approximation. *BIT Numerical Mathematics*, 62(4):1149–1174, December 2022.
- [9] Gianluca Ceruti, Jonas Kusch, and Christian Lubich. A Parallel Rank-Adaptive Integrator for Dynamical Low-Rank Approximation. *SIAM Journal on Scientific Computing*, 46(3):B205–B228, June 2024.
- [10] Gianluca Ceruti and Christian Lubich. An unconventional robust integrator for dynamical low-rank approximation. *BIT Numerical Mathematics*, 62(1):23–44, March 2022.
- [11] Wei Chen, Giacomo Dimarco, and Lorenzo Pareschi. Structure and asymptotic preserving deep neural surrogates for uncertainty quantification in multiscale kinetic equations, June 2025. arXiv:2506.10636.
- [12] K A Cliffe, M B Giles, R Scheichl, and A L Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Computing and Visualization in Science*, 14(1):3–15, August 2011.
- [13] Nathan Collier, Abdul-Lateef Haji-Ali, Fabio Nobile, Erik Von Schwerin, and Raúl Tempone. A continuation multilevel Monte Carlo algorithm. *BIT Numerical Mathematics*, 55(2):399–432, June 2015.
- [14] David Craft, Mark Bangert, Troy Long, Dávid Papp, and Jan Unkelbach. Shared data for intensity modulated radiation therapy (imrt) optimization research: the cort dataset. *GigaScience*, 3(1):2047–217X, 2014.
- [15] Lukas Einkemmer, Jingwei Hu, and Jonas Kusch. Asymptotic-preserving and energy stable dynamical low-rank approximation. *SIAM Journal on Numerical Analysis*, 62(1):73–92, 2024.
- [16] Lukas Einkemmer, Katharina Kormann, Jonas Kusch, Ryan G. McClarren, and Jing-Mei Qiu. A review of low-rank methods for time-dependent kinetic simulations. *Journal of Computational Physics*, 538:114191, October 2025.
- [17] Lukas Einkemmer and Christian Lubich. A Low-Rank Projector-Splitting Integrator for the Vlasov–Poisson Equation. *SIAM Journal on Scientific Computing*, 40(5):B1330–B1360, January 2018.
- [18] Hillary R. Fairbanks, Alireza Doostan, Christian Ketelsen, and Gianluca Iaccarino. A low-rank control variate for multilevel Monte Carlo simulation of high-dimensional uncertain systems. *Journal of Computational Physics*, 341:121–139, July 2017.

- [19] Anqi Fu, Vicki T Taasti, and Masoud Zarepisheh. Distributed and scalable optimization for robust proton treatment planning. *Medical physics*, 50(1):633–642, 2023.
- [20] B D. Ganapol and OECD Nuclear Energy Agency. *Analytical benchmarks for nuclear engineering applications: case studies in neutron transport theory*. Data bank. Nuclear Energy Agency, France, 2008. OCLC: 430990895.
- [21] Michael B. Giles. Multilevel Monte Carlo Path Simulation. *Operations Research*, 56(3):607–617, June 2008.
- [22] Michael B. Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328, May 2015.
- [23] Marcus J. Grote, Simon Michel, and Fabio Nobile. Uncertainty Quantification by Multilevel Monte Carlo and Local Time-Stepping for Wave Propagation. *SIAM/ASA Journal on Uncertainty Quantification*, 10(4):1601–1628, December 2022.
- [24] Abdul-Lateef Haji-Ali, Fabio Nobile, Erik von Schwerin, and Raúl Tempone. Optimization of mesh hierarchies in multilevel Monte Carlo samplers. *Stochastics and Partial Differential Equations Analysis and Computations*, 4(1):76–112, March 2016.
- [25] Shi Jin, Zheng Ma, and Keke Wu. Asymptotic-Preserving Neural Networks for Multiscale Kinetic Equations. *Communications in Computational Physics*, 35(3):693–723, January 2024.
- [26] Emil Kieri, Christian Lubich, and Hanna Walach. Discretized Dynamical Low-Rank Approximation in the Presence of Small Singular Values. *SIAM Journal on Numerical Analysis*, 54(2):1020–1038, January 2016.
- [27] Othmar Koch and Christian Lubich. Dynamical Low-Rank Approximation. *SIAM Journal on Matrix Analysis and Applications*, 29(2):434–454, January 2007.
- [28] Julian Koellermeier, Philipp Krah, and Jonas Kusch. Macro-micro decomposition for consistent and conservative model order reduction of hyperbolic shallow water moment equations: a study using POD-Galerkin and dynamical low-rank approximation. *Advances in Computational Mathematics*, 50(4):76, August 2024.
- [29] Julian Koellermeier and Marvin Rominger. Analysis and numerical simulation of hyperbolic shallow water moment equations. *Communications in Computational Physics*, 28(3):1038–1084, 2020.
- [30] Julia Kowalski and Manuel Torrilhon. Moment Approximations and Model Cascades for Shallow Flow. *Communications in Computational Physics*, 25(3):669–702, January 2019.

- [31] Jonas Kusch. Second-order robust parallel integrators for dynamical low-rank approximation. *BIT Numerical Mathematics*, 65(3):31, September 2025.
- [32] Jonas Kusch, Gianluca Ceruti, Lukas Einkemmer, and Martin Frank. Dynamical Low-Rank Approximation for Burgers' Equation with Uncertainty. *International Journal for Uncertainty Quantification*, 12(5):1–21, 2022.
- [33] Jonas Kusch, Lukas Einkemmer, and Gianluca Ceruti. On the Stability of Robust Dynamical Low-Rank Approximations for Hyperbolic Problems. *SIAM Journal on Scientific Computing*, 45(1):A1–A24, February 2023.
- [34] Jonas Kusch, Steffen Schotthöfer, Pia Stammer, Jannick Wolters, and Tianbai Xiao. KiT-RT: An Extendable Framework for Radiative Transfer and Therapy. *ACM Trans. Math. Softw.*, 49(4):38:1–38:24, December 2023.
- [35] Jonas Kusch and Pia Stammer. A robust collision source method for rank adaptive dynamical low-rank approximation in radiation therapy. *ESAIM: Mathematical Modelling and Numerical Analysis*, 57(2):865–891, March 2023.
- [36] Randall J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 1 edition, August 2002.
- [37] Elmer Eugene Lewis and Warren F Miller. *Computational methods of neutron transport*. United States: John Wiley and Sons, Inc., 1984.
- [38] Liu Liu and Xueyu Zhu. A bi-fidelity method for the multiscale Boltzmann equation with random parameters. *Journal of Computational Physics*, 402:108914, February 2020.
- [39] AJ Lomax. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. *Physics in Medicine & Biology*, 53(4):1027–1042, 2008.
- [40] AJ Lomax. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions. *Physics in Medicine & Biology*, 53(4):1043–1056, 2008.
- [41] Christian Lubich and Ivan V. Oseledets. A projector-splitting integrator for dynamical low-rank approximation. *BIT Numerical Mathematics*, 54(1):171–188, March 2014.
- [42] Yan Luo and Zhu Wang. A Multilevel Monte Carlo Ensemble Scheme for Random Parabolic PDEs. *SIAM Journal on Scientific Computing*, 41(1):A622–A642, January 2019.
- [43] S. Mishra and Ch. Schwab. Sparse tensor multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random initial data. *Mathematics of Computation*, 81(280):1979–2018, April 2012.

- [44] S. Mishra, Ch. Schwab, and J. Šukys. Multi-level Monte Carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions. *Journal of Computational Physics*, 231(8):3365–3388, April 2012.
- [45] Eleonora Musharbash and Fabio Nobile. Dual Dynamically Orthogonal approximation of incompressible Navier Stokes equations with random boundary conditions. *Journal of Computational Physics*, 354:135–162, February 2018.
- [46] Eleonora Musharbash, Fabio Nobile, and Eva Vidličková. Symplectic dynamical low rank approximation of wave equations with random parameters. *BIT Numerical Mathematics*, 60(4):1153–1201, 2020.
- [47] Fabio Nobile and Sébastien Riffaud. High-Order BUG Dynamical Low-Rank Integrators Based on Explicit Runge–Kutta Methods. *Journal of Scientific Computing*, 107(3):102, June 2026.
- [48] Lorenzo Pareschi. An Introduction to Uncertainty Quantification for Kinetic Equations and Related Problems. In Giacomo Albi, Sara Merino-Aceituno, Alessia Nota, and Mattia Zanella, editors, *Trails in Kinetic Theory: Foundational Aspects and Numerical Methods*, pages 141–181. Springer International Publishing, Cham, 2021.
- [49] Chinmay Patwardhan and Jonas Kusch. A Parallel, Energy-Stable Low-Rank Integrator for Nonlinear Multi-Scale Thermal Radiative Transfer. *Journal of Computational and Theoretical Transport*, 55(2):145–189, February 2026.
- [50] Chinmay Patwardhan, Pia Stammer, Emil Løvnbak, Jonas Kusch, and Sebastian Krumscheid. Low-Rank Variance Reduction for Uncertain Radiative Transfer with Control Variates. In Christiane Lemieux and Ben Feng, editors, *Monte Carlo and Quasi-Monte Carlo 2024*, volume 522, pages 357–375. Springer Nature Switzerland, Cham, 2026. Series Title: Springer Proceedings in Mathematics & Statistics.
- [51] Zoltán Perkó, Sebastian R Van Der Voort, Steven Van De Water, Charlotte MH Hartman, Mischa Hoogeman, and Danny Lathouwers. Fast and accurate sensitivity analysis of impt treatment plans using polynomial chaos expansion. *Physics in Medicine & Biology*, 61(12):4646–4664, 2016.
- [52] Martina Prugger, Lukas Einkemmer, and Carlos F. Lopez. A dynamical low-rank approach to solve the chemical master equation for biological reaction networks. *Journal of Computational Physics*, 489:112250, September 2023.
- [53] Themistoklis P Sapsis and Pierre FJ Lermusiaux. Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Physica D: Nonlinear Phenomena*, 238(23-24):2347–2360, 2009.
- [54] Steffen Schotthöfer and Cory Hauck. Reference solutions for linear radiation transport: the Hohlraum and Lattice benchmarks, May 2025.

- [55] Pia Stammer, Niklas Wahl, Jonas Kusch, and Danny Lathouwers. A high-order deterministic dynamical low-rank method for proton transport in heterogeneous media. *Journal of Computational Physics*, page 114879, 2026.
- [56] M.P. Ueckermann, P.F.J. Lermusiaux, and T.P. Sapsis. Numerical schemes for dynamically orthogonal equations of stochastic fluid and ocean flows. *Journal of Computational Physics*, 233:272–294, 2013.
- [57] B. P. Welford. Note on a Method for Calculating Corrected Sums of Squares and Products. *Technometrics*, 4(3):419–420, August 1962.