CrossMark

# AREPO-MCRT: Monte Carlo Radiation Hydrodynamics on a Moving Mesh

Aaron Smith[1,5] , Rahul Kannan[2] , Benny T.-H. Tsang[3] , Mark Vogelsberger[1] , and Rüdiger Pakmor[4]
[1] Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; arsmith@mit.edu
[2] Center for Astrophysics, Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[3] Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA
[4] Max-Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany

## Abstract

We present AREPO-MCRT, a novel Monte Carlo radiative transfer radiation-hydrodynamics (RHD) solver for the unstructured moving-mesh code AREPO. Our method is designed for general multiple scattering problems in both optically thin and thick conditions. We incorporate numerous efficiency improvements and noise reduction schemes to help overcome efficiency barriers that typically inhibit convergence. These include continuous absorption and energy deposition, photon weighting and luminosity boosting, local packet merging and splitting, path-based statistical estimators, conservative (face-centered) momentum coupling, adaptive convergence between time steps, implicit Monte Carlo algorithms for thermal emission, and discrete-diffusion Monte Carlo techniques for unresolved scattering, including a novel advection scheme. We primarily focus on the unique aspects of our implementation and discussions of the advantages and drawbacks of our methods in various astrophysical contexts. Finally, we consider several test applications including the levitation of an optically thick layer of gas by trapped infrared radiation. We find that the initial acceleration phase and revitalized second wind are connected via self-regulation of the RHD coupling, such that the RHD method accuracy and simulation resolution each leave important imprints on the long-term behavior of the gas.

*Unified Astronomy Thesaurus concepts:* Radiative transfer (1335); Computational methods (1965); Radiative transfer simulations (1967); Hydrodynamical simulations (767)

## 1. Introduction

Observations of light at all wavelengths of the electromagnetic spectrum have been essential to our understanding of the cosmos and the wide range of astrophysical phenomena found therein. Intense radiation from stars and black holes can also affect the dynamical evolution of small- to large-scale systems as sources of thermal, mechanical, and chemical feedback. A substantial effort has been made to model the interplay between gas and radiation with robust theoretical models and numerical techniques. In many cases, the multiscale, multiphysics nature of such problems necessitates fully coupled radiation-hydrodynamics (RHD) simulations to precisely capture the underlying physics (Pomraning 1973; Mihalas & Mihalas 1984; Castor 2004).

The large variety of environments requiring the study of RHD and decoupled radiative transfer has led to the development of different algorithms with advantages for specialized applications. Part of the complexity is the high dimensionality of the radiation field, which can vary as a function of space, direction, frequency, and time. Most schemes employed by modern codes can broadly be classified as: (i) directly integrating the transport equation along characteristic paths; (ii) solving discrete-ordinate representations of the transport equation; and (iii) solving reduced-dimensionality angular-moment equations derived with an approximate closure relation. Each of these, or hybrid combinations thereof, have been highly successful in exploratory and precision studies in astrophysics. A few contemporary methods for multidimensional RHD include flux-limited diffusion (FLD; Levermore & Pomraning 1981; Turner & Stone 2001; Krumholz et al. 2007; Commerçon et al. 2011), first moment closure (M1; Levermore 1984; González et al. 2007; Rosdahl et al. 2013; Skinner & Ostriker 2013; Kannan et al. 2019), variable

Eddington tensor formulations (VET; Stone et al. 1992; Davis et al. 2012; Jiang et al. 2012, 2014), optically thin VET (OTVET; Gnedin & Abel 2001), adaptive ray-tracing (Abel et al. 1999; Whalen & Norman 2006; Trac & Cen 2007; Wise & Abel 2011; Rosen et al. 2017), transport in light cones (Pawlik & Schaye 2008), and the method of characteristics moment closure (MOCMC; Ryan & Dolence 2020). The community continues to benefit from numerous contributions enabling further radiative transfer studies.

In this paper, we present a Monte Carlo radiative transfer (MCRT) implementation for the moving-mesh hydrodynamics code AREPO (Springel 2010; Pakmor et al. 2011, 2016; Weinberger et al. 2020). Our efforts are complementary to other RHD methods already implemented in the code, including direct discretization (Petkova & Springel 2011), SIMPLEX triangulation (Jaura et al. 2018), and M1 closure (Kannan et al. 2019). The MCRT method provides an accurate approach to modeling radiation fields by sampling from physically motivated probability distribution functions (for a recent review see Noebauer & Sim 2019). Complex phenomena arising from microphysical processes, such as multiple scattering and frequency redistribution, can be accounted for from first principles and astrophysical observables can be assembled one photon packet at a time. This often leads to conceptually simple implementations and interpretations of the method. One of the main drawbacks of MCRT is the noise from photon packet discretization. However, with enough computational resources and efficiency-improving algorithms, convergence can always be obtained to arbitrarily high accuracy. Quantities derived from binned photon statistics typically suffer from a slow convergence rate with the number of samples as $1/\sqrt{N}$, although this is independent of the dimensionality of the physical setup.

---
[5] NHFP Einstein Fellow.

Indeed, MCRT constitutes a valuable alternative approach to RHD that can outperform other methods in terms of accuracy and emergent statistics in many computationally demanding situations. It is also clear that the MCRT RHD approach is maturing with an increasing number of codes and applications (Nayakshin et al. 2009; Noebauer et al. 2012; Cleveland & Gentile 2015; Harries 2015; Noebauer & Sim 2015; Roth & Kasen 2015; Ryan et al. 2015; Tsang & Milosavljević 2015, 2018; Harries et al. 2017; Smith et al. 2017b; Vandenbroucke & Wood 2018). Therefore, it is timely to incorporate on-the-fly and post-processing MCRT into AREPO for accurate native modeling of radiation fields and emergent observables. Especially given the demonstrated capabilities of AREPO for studying astrophysical problems over the past decade (for a brief but comprehensive list of applications see Weinberger et al. 2020). Additional access to accurate radiative transfer on a moving mesh can help facilitate important insights in astronomy spanning a wide range of scales, large and small.

Although many of the algorithms discussed in this work are new, we have adopted or adapted most of the concepts from the mature MCRT literature. For example, implicit Monte Carlo (IMC; Fleck & Cummings 1971; Brooks & Fleck 1986; Brooks 1989) algorithms provide an essential framework for stable and efficient time-dependent coupling of the radiative transfer and thermal energy equations by treating a portion of absorption followed by re-emission as effective scattering. However, there are tradeoffs with implicit schemes; for example, spurious behavior has been found in extreme cases which can only be mitigated with corrective IMC or iterative variants (Long et al. 2014; Cleveland & Wollaber 2018). More importantly, in astrophysical applications MCRT photon packets become diffusive and inefficient to track without accelerated transport. One option is to maintain continuous photon positions with movement in optically thick zones according to a modified random walk (Fleck & Canfield 1984). This technique is widely used and has undergone several recent improvements (Min et al. 2009; Robitaille 2010; Keady & Cleveland 2017).

In practice, random walks are not always optimal because they become inactive near cell interfaces. Still, such diffusion approximations are necessary to obtain accurate MCRT results at high optical depths (Camps & Baes 2018). Therefore, discrete-diffusion Monte Carlo (DDMC) techniques have been developed to increase the efficiency of MCRT calculations in opaque media (Gentile 2001; Densmore et al. 2007). The idea is to replace many unresolved scatterings with a single jump to a neighboring cell based on a discretized diffusion equation. Under the Fick's law closure relation of the radiative transfer equation (RTE), the diffusive term operates as spatial leakage across cell interfaces that is naturally incorporated into the Monte Carlo (MC) paradigm. A robust hybrid scheme allows the conversion between DDMC and MC particles for accurate propagation through optically thin cells as well. The DDMC method has also been extended to incorporate frequency-dependent transfer (Abdikamalov et al. 2012; Densmore et al. 2012; Wollaeger et al. 2013; Wollaeger & van Rossum 2014).

The paper is organized as follows. In Section 2, we discuss the main numerical methodology, emphasizing aspects that are unique to our implementation, hereafter known as AREPO-MCRT. In Section 3, we test our implementation against known analytic and numerical solutions from the literature. In Section 4, we apply the code to the problem of radiative forcing of a dusty atmosphere. Finally, in Section 5, we provide a summary of findings and discussion of future applications.

## 2. Numerical Methodology

In this section, we describe our Monte Carlo radiative transfer RHD implementation, which is fully integrated into the moving-mesh code AREPO (Springel 2010). Specifically, AREPO employs a second-order finite-volume method to solve the ideal magneto-hydrodynamical equations on an unstructured Voronoi tessellation, which is free to move with the local fluid velocity. Native ray-tracing through three-dimensional Voronoi tessellations is well understood in computational geometry and has been employed by several post-processing MCRT codes with great success (Camps et al. 2013; Camps & Baes 2015; Hubber et al. 2016; Smith et al. 2017a). We therefore focus on physics rather than geometry in this section.

### 2.1. Radiation Transport

The specific intensity $I_\nu(\boldsymbol{r}, \boldsymbol{n}, t)$ encodes all information about the radiation field taking into account the frequency $\nu$, spatial position $\boldsymbol{r}$, propagation direction unit vector $\boldsymbol{n}$, and time $t$. The general RTE is given in the lab frame by

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \boldsymbol{n} \cdot \boldsymbol{\nabla} I_\nu = j_\nu - k_\nu I_\nu, \qquad (1)$$

where $k_\nu$ is the absorption coefficient and $j_\nu$ is the emission coefficient. We define the mean intensity as $J_\nu \equiv \frac{1}{4\pi}\int d\Omega I_\nu$, which is related to the energy density by $u_\nu = \frac{4\pi}{c}J_\nu$. To avoid confusion, in this paper we denote the gas internal energy density by $u_{\rm g}$.

In this paper we restrict the discussion to monochromatic radiation, although this is not required in the code and will be relaxed in future applications. In this context a convenient parameterization for the absorption coefficient is through a constant opacity $\kappa \equiv k/\rho$ and scattering albedo $A \equiv k_{\rm s}/(k_{\rm s} + k_{\rm a})$, where $\rho$ denotes the gas density and $k_{\rm s}$ and $k_{\rm a}$ are the purely scattering and absorbing components. Furthermore, if we assume local thermodynamic equilibrium (LTE) then the right-hand side of Equation (1) simplifies to include (i) dust absorption and thermal emission as $k_{\rm a}(B - I)$ based on the Planck function $B$ at temperature $T$, (ii) isotropic elastic scattering as $\frac{k_{\rm s}}{4\pi}\int [I(\boldsymbol{n}') - I(\boldsymbol{n})]d\boldsymbol{n}' = k_{\rm s}(J - I)$, and (iii) external sources $j_{\rm ext}$, e.g., geometrically assigned sources. For completeness, the Planck (blackbody) distribution is defined as $B_\nu(T) \equiv \frac{2h\nu^3}{c^2}(e^{h\nu/k_{\rm B}T} - 1)^{-1}$ with a normalization such that the energy density that radiation would have if it were in thermodynamic equilibrium with gas is $u_{\rm r} = \frac{4\pi}{c}\int B_\nu(T)d\nu = a_{\rm B}T^4$. Finally, the Planck mean absorption coefficient is $k_{\rm P} = \int k_{\nu,{\rm a}}B_\nu(T)d\nu \big/ \int B_\nu(T)d\nu$, which is trivially $k_{\rm a}$ under the assumption of a frequency-independent (gray) opacity. Thus, we summarize the simplified version of Equation (1) as

$$\frac{1}{c}\frac{\partial I}{\partial t} + \boldsymbol{n} \cdot \boldsymbol{\nabla} I = k_{\rm a}(B - I) + k_{\rm s}(J - I) + j_{\rm ext}. \qquad (2)$$

The radiation flux and pressure are given by $\boldsymbol{F}_\nu = \int d\Omega I_\nu \boldsymbol{n}$ and $\mathsf{P}_\nu \equiv c^{-1}\int d\Omega I_\nu \boldsymbol{n} \otimes \boldsymbol{n}$, respectively. Therefore, the moment equations provide a general framework for radiation–gas coupling:

$$\frac{\partial u}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{F} = ck_{\rm a}(u_{\rm r} - u) \qquad (3)$$

and

$$\frac{1}{c^2}\frac{\partial \boldsymbol{F}}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{P} = -\frac{k\boldsymbol{F}}{c}. \qquad (4)$$

For presentation purposes we do not show moments of $j_{\text{ext}}$.

### 2.2. Radiation Hydrodynamics

The equations governing nonrelativistic hydrodynamics can be written in an Eulerian reference frame as a set of conservation laws for mass, momentum, and total energy (Castor 2004):

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) = 0, \qquad (5)$$

$$\frac{\partial \rho \boldsymbol{v}}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v} \otimes \boldsymbol{v}) + \boldsymbol{\nabla} P = \frac{k\boldsymbol{F}}{c}, \qquad (6)$$

and

$$\frac{\partial \rho e}{\partial t} + \boldsymbol{\nabla} \cdot [(\rho e + P)\boldsymbol{v}] = ck_{\text{a}}(u - u_{\text{r}}) + \boldsymbol{v} \cdot \frac{k\boldsymbol{F}}{c}. \qquad (7)$$

Here, $\rho$ is the density, $\boldsymbol{v}$ the velocity, $P$ the pressure, $e \equiv \epsilon + \frac{1}{2}|\boldsymbol{v}|^2$ the total specific energy. We typically assume an ideal gas equation of state so the pressure is isotropic and specified by $P = (\gamma_{\text{ad}} - 1)u_{\text{g}}$, where $u_{\text{g}} = \rho\epsilon$ and $\gamma_{\text{ad}} \equiv C_P/C_V$ is the adiabatic index, or ratio of specific heat at constant pressure to that at constant volume. These equations would also be modified in the presence of additional physics, such as magnetic fields or self-gravity, which is already present in the AREPO code (Springel 2010; Weinberger et al. 2020).

In the current work we primarily focus on the MCRT and RHD physics, noting that we expect to improve the implementation as needed for future applications. For now, the radiation is coupled to the hydrodynamics with standard operator splitting that is formally first order in time and space. While second-order schemes for spatial emission sampling and path integration utilizing the gradient information in AREPO are possible, we leave the exploration of potential benefits to future work. At this point we also require global time-stepping, but plan to allow compatibility with the individual time-stepping scheme of AREPO, i.e., the efficient factor of two hierarchy of time steps, which would then also benefit from a second-order time-integration scheme.

The energy and momentum exchanged between gas and radiation is performed at the end of each time step as follows. First, emission processes can remove internal energy from the gas, which is converted to radiation as MC photon packets. second, the MCRT transport calculations are performed, meanwhile tracking the cumulative interactions with gas due to photon absorption and scattering processes. Lastly, the conserved gas quantities are collectively updated to reflect the net exchange. In the next section we describe the IMC technique to treat the source and transport steps (semi-) implicitly. The mesh configuration remains unchanged during transport, however photon packets generally persist across time steps. This is achieved by saving the particle positions along with the host cell indices, which are validated before transport via a breadth-first face neighbor walk. The photon packets are also efficiently exchanged between tasks during a domain decomposition. Nonlocal transport is handled by collecting photons at domain boundaries and then sending batches via asynchronous point-to-point communication patterns (similar to Rosen et al. 2017). Additional details of the implementation are described in the remaining subsections.

### 2.3. Implicit Monte Carlo

Under LTE conditions, radiation and gas are tightly coupled through the strong temperature dependence of the thermal emission. The IMC algorithm employs a semi-implicit time discretization to transform the nonlinear thermal RTE into a system of linearized equations naturally incorporated into the MC method. Effectively, the method replaces a portion of absorption and re-emission with elastic scattering, thus reducing the amount of quasi-equilibrium energy exchange between gas and radiation. We refer the interested reader to Fleck & Cummings (1971) and Wollaber (2016) for detailed derivations and numerical discussions. For our purposes, we introduce the material equation as (simplified following Equation (2))

$$\frac{1}{c}\frac{\partial u_{\text{g}}}{\partial t} = \frac{1}{c\beta}\frac{\partial u_{\text{r}}}{\partial t} = k_{\text{a}}(u - u_{\text{r}}), \qquad (8)$$

where $\beta \equiv \partial u_{\text{r}}/\partial u_{\text{g}}$ quantifies the thermal coupling. We then expand $u_{\text{r}}$ to first order in time for $t \in [t^n, t^{n+1}]$

$$u_{\text{r}}(t) \simeq u_{\text{r}}^n + \alpha \Delta t (\partial u_{\text{r}}/\partial t)^n. \qquad (9)$$

The parameter $\alpha \in [0, 1]$ is a coefficient that interpolates between a fully explicit ($\alpha = 0$) and implicit ($\alpha = 1$) scheme for updating $u_{\text{r}}$ over the time step $\Delta t$. Substituting $u_{\text{r}}(t)$ from Equations (9) into (8) and eliminating $(\partial u_{\text{r}}/\partial t)^n$ gives the following:

$$u_{\text{r}}(t) = f u_{\text{r}}^n + (1 - f)u, \qquad (10)$$

in which we have introduced the so-called Fleck factor

$$f \equiv \frac{1}{1 + \alpha c \Delta t \beta k_{\text{P}}} = \frac{1}{1 + 4\alpha c \Delta t (1 - A)\kappa u_{\text{r}}/\epsilon}, \qquad (11)$$

where in the second equality we assume a constant scattering albedo $A$, opacity $\kappa$, and an ideal gas equation of state such that $\beta = 4u_{\text{r}}/u_{\text{g}}$. Finally, we substitute $u_{\text{r}}$ from Equation (10) into Equation (2) to obtain the implicit radiation transport equation:

$$\frac{1}{c}\frac{\partial I}{\partial t} + \boldsymbol{n} \cdot \boldsymbol{\nabla} I = k_{\text{ea}}(B - I) + k_{\text{es}}(J - I) + j_{\text{ext}}, \qquad (12)$$

where the effective absorption and scattering coefficients are

$$k_{\text{ea}} \equiv f k_{\text{a}} \qquad (13)$$

and

$$k_{\text{es}} \equiv k_{\text{s}} + (1 - f)k_{\text{a}}. \qquad (14)$$

Thus, after calculating the Fleck factor the numerical radiation transport coefficients are modified to reflect the replacement of a portion of thermal absorption with scattering. For notational simplicity, we drop the "effective" subscript with the understanding that the IMC scheme is implied throughout.

### 2.4. Monte Carlo Procedure

Under the MCRT paradigm we represent the specific intensity with an ensemble of photon packets, which are each characterized by an energy weight $\varepsilon_k$, position $\boldsymbol{r}_k$, normalized direction $\boldsymbol{n}_k$, and time $t_k$, where the index $k$ refers to an

individual MC packet. We emphasize that in our implementation different MC photon packets can each have varying energies, which can greatly improve the sampling and convergence statistics. AREPO employs a finite-volume method to solve the gas conservation equations so in the first-order scheme we assume constant gas properties within each cell. We treat the thermal emission term $k_a B$ in Equation (2) as a continuous process for the gas but stochastically for radiation. The total thermal radiation energy emitted by each cell assuming gray opacity is

$$\Delta E_{\mathrm{em},i} = c \Delta t V_i k_a a_B T^4, \qquad (15)$$

where $\Delta t$ denotes the time step and $V_i$ is the current volume of cell $i$. We note that with IMC the emission term is corrected by the Fleck factor as inherited from Equation (13). For numerical stability, we subtract this exact energy from each gas cell, however the insertion of MC packets is necessarily discretized. This is achieved by constructing the cumulative distribution function from individual cells, drawing a random number to find the emitting cell, and assigning the photon packet position uniformly within the Voronoi cell volume. The emission direction is isotropic in the comoving frame of the gas and the time is uniformly distributed over the duration of the time step, $t_k \in [t_n, t_{n+1}]$. Furthermore, we incorporate an emission weighting scheme to accelerate the convergence of MC sampling. This allows simulations to dynamically allocate photon packets to optimally sample the radiation field, which we currently implement as a power law luminosity boost $\propto L^\gamma$ properly normalized to conserve energy. Specifically, $\gamma \in (0, 1)$ alleviates the $1/N_{\mathrm{ph}}$ sensitivity limit inherent to uniform sampling, thereby reducing emissivity disparities that often lead to statistical noise in dim regions and unnecessary oversampling in bright regions. If the total radiation energy at a given time is $\mathcal{E}_n$ then the photon weight is $w_k \equiv \varepsilon_k / \mathcal{E}_n$, so by construction $\sum w_k = 1$.

The subsequent transport of photon packets follows the usual MC procedure for scattering and escape (e.g., Smith et al. 2015; Tsang & Milosavljević 2015). We determine the optical depth to scattering based on an exponential distribution, i.e., $\tau_{\mathrm{scat}} = -\ln \zeta$ where $\zeta$ is a random number uniformly distributed in [0, 1]. The location of the scattering interaction is calculated implicitly via piecewise constant integration, i.e., by continually moving the photon through each cell until $\tau_{\mathrm{scat}}$ is exhausted. After each scattering event the photon is assigned a new direction $\boldsymbol{n}_k$ and the ray-tracing procedure continues until the photon (i) reaches the end of the time step, (ii) arrives at a specified escape criterion, or (iii) is removed by an absorption process (see Section 2.6). Each time the photon moves a distance $\Delta \ell_k$ there is a corresponding change in position of $\Delta \boldsymbol{r}_k = \Delta \ell_k \boldsymbol{n}_k$, elapsed time of $\Delta t_k = \Delta \ell_k / c$, and traversed optical depth of $\Delta \tau_k = k_s \Delta \ell_k$.

### 2.5. Photon Splitting and Merging

The instantaneous representation of the MCRT radiation field is the result of nontrivial sourcing and transport. In addition, photons persist across time steps, which can lead to dense accumulations of packets in regions of $(\boldsymbol{r}, \boldsymbol{n}, \nu)$ phase space. We therefore implement photon splitting and merging. Specifically, at the beginning of each time step we split statistically overweight packets, e.g., down to a few standard deviations above the mean, $\varepsilon_k \gtrsim \langle \varepsilon \rangle + f_{\mathrm{split}} \langle \langle \varepsilon^2 \rangle - \langle \varepsilon \rangle^2 \rangle^{1/2}$. This simple prescription efficiently regulates the distribution of packet weights during transport calculations to reduce unnecessary variance. We note that in scattering media redundant photons quickly disperse to better sample alternative trajectories. In the future, particular applications may require additional splitting criteria to increase the signal-to-noise in optically thin regions or at large distances from the last scattering or emission event (e.g., similar to Harries 2015).

We also implement optional photon merging schemes to combine underweight photons, remove redundant information, or reduce the total number of photons in the simulation. As packets are already assigned to gas cells, at the end of each time step we perform a linear sort to count and identify merger candidates as groups of packets sharing a common host cell. This has the advantage of being computationally efficient and ensuring that the merged packets remain within the convex hull of the Voronoi cells. We currently consider all photons in the same cell as merger candidates, but one might include additional criteria such as the photon weight in the merge condition.

Once we have a list of photons in each cell we can further bin them into directional groups to preserve angular information after grouping. However, binning can lead to numerical artifacts by biasing paths toward certain directions. Therefore, our preferred implementation iteratively merges the pair of photons with closest angular separation, until all pairs exceed a threshold angle, e.g., $\theta > 30°$. This has the advantage of being unique and not favoring any particular directions. Although our algorithm is technically $\mathcal{O}(N^2)$, a matrix-like data structure means one particle's row and column is removed while the other's is reused for the merged packet. That is, we only recompute dot products for the row and column of the new packet, retaining angles for unaffected pairs and shrinking the effective matrix dimensions. In practice the merging calculations are a negligible fraction of the overall simulation runtime. However, when this is not the case, we optionally introduce a short-circuiting condition during the pair search to immediately merge sufficiently close photons, e.g., $\theta < 5°$, and limit the remaining candidates.

While merging can result in information loss, it is straightforward to characterize. Specifically, the merger only preserves the group momentum and center of energy position.[6] For a pair of photons the new energy is $\varepsilon' = (\varepsilon_1^2 + \varepsilon_2^2 + 2\varepsilon_1 \varepsilon_2 \cos \theta)^{1/2}$, which for equal weight photons results in a fractional loss of $f_{\mathrm{loss}} = 1 - \varepsilon'/(\varepsilon_1 + \varepsilon_2) = 1 - \frac{1}{2}\sqrt{2 + 2\cos\theta} \approx \frac{1}{8}\theta^2 + \mathcal{O}(\theta^4)$, such that the maximum is $f_{\mathrm{loss}} \lesssim 3.4\%(\theta/30°)^2$. The final outcome of multiple particle mergers is more complex but can be bounded by considering a ring configuration, i.e., equal weights and maximal separations, resulting in $f_{\mathrm{loss}} \leqslant 1 - \cos\theta \approx \frac{1}{2}\theta^2 = 13.7\%(\theta/30°)^2$. More realistic is the case of many particles isotropically distributed over a spherical cap, which can be derived via a surface rotation as half of the loss in the ring case, or $f_{\mathrm{loss}} \approx 6.9\%(\theta/30°)^2$. Unfortunately, these losses are unavoidable when employing splitting and merging to help regulate the weight distribution and avoid skewed or imbalanced statistics.

---

[6] Alternatively, one might prioritize conserving the total energy, which results in a fractional gain of momentum compared to the original radiation field. Beyond these simple choices, one might leverage the insight that particle data is often redundant to obtain a weighted subset (called a *coreset*) to replace direct merging with a consolidated representation of the intensity.

## 2.6. Continuous Absorption

We employ the continuous absorption method to minimize MC sampling noise. Treating the $-k_a I$ term in Equation (1) deterministically is a variance-reduction technique that accelerates the convergence of the radiation field by adjusting the weight of photon packets. Thus, after considering all other event-based processes, such as scattering or grid traversal, the photon energy weight is reduced by $e^{-\tau_a}$, where for notational simplicity we use $\tau_a \equiv k_a \Delta \ell_k$. Photon packets with negligible weight can be eliminated by applying a threshold condition, e.g., $w_k > w_{min} \sim 10^{-12}$. Furthermore, the internal energy deposited to the gas is

$$\Delta E_{abs,i} = \sum_{paths} \varepsilon_k (1 - e^{-\tau_a}), \qquad (16)$$

where the sum is over all photon paths within cell $i$. We also employ a path-based estimator for the radiation energy density specifically accounting for the decreasing energy contribution due to continuous absorption (Smith et al. 2018). Otherwise the energy and momentum deposition can be overestimated, especially in cases where the absorption optical depth is greater than unity. The correction factor is $\int_0^{\tau_a} e^{-\tau'_a} d\tau'_a / \tau_a = (1 - e^{-\tau_a})/\tau_a$, or approximately $1 - \tau_a/2$ in the optically thin limit for absorption, which is important for numerical stability although we use the exact expression in the equations that follow. The corresponding corrected energy density based on the total residence time of propagating photons is (Lucy 1999)

$$u_i = \sum_{paths} \frac{\varepsilon_k}{V_i} \frac{(1 - e^{-\tau_a})}{k_a c \Delta t}, \qquad (17)$$

with the sum again over all paths within cell $i$.

## 2.7. Momentum Coupling

The momentum deposition is also path-based and includes both absorption and scattering contributions as

$$\Delta \boldsymbol{p}_i = \sum_{paths} \frac{\tau \varepsilon_k}{c} \left( \frac{1 - e^{-\tau_a}}{\tau_a} \right) \boldsymbol{n}_k. \qquad (18)$$

Furthermore, as scattering occurs in the comoving frame the following kinetic energy is also transferred to the gas:

$$\Delta E_{kin,i} = \boldsymbol{v} \cdot \Delta \boldsymbol{p}_i. \qquad (19)$$

We note that momentum coupling in non-Cartesian coordinates is complicated by the fact that the reference unit vectors can change along traversed paths. In Appendix A, we provide additional discussion specific to calculations in spherical geometry, which is commonly used in astrophysical simulations.

We have implemented two methods for momentum coupling that incorporate different meanings to the summations in Equation (18). The first assigns the momentum directly to the host cell, i.e., the usual volume-integrated coupling. However, this can underestimate the radiation pressure force in the immediate vicinity around unresolved sources and therefore the impact on gas properties (Hopkins & Grudić 2019). It is often impractical to resolve photon mean free paths in hydrodynamics simulations so we provide a general Monte Carlo solution to conserve momentum in such cases. Our approach is a neighbor-based method, which also applies momentum to the cell that the photon packet would enter next. In the case of multiple scattering, momenta from each segment of the overall trajectory are imparted independently onto different neighbors according to individual path-based estimators. The schemes we propose are also particularly well suited for the unstructured meshes encountered in AREPO, and adhere to the MCRT philosophy of accurately capturing sub-grid physical interactions.[7]

We emphasize that both volume- and face-integrated methods can lead to unphysical results for isotropic sources embedded within optically thick cells. For example, a constant luminosity point source with negligible absorption produces an inverse square law flux $\boldsymbol{F} = L\hat{\boldsymbol{r}}/(4\pi r^2)$, such that the energy density for a pure scattering medium is given by $\boldsymbol{F} = -c\boldsymbol{\nabla}u/3k$ and $u = 3kL/(4\pi cr)$. By employing Gauss's theorem the integrated momentum rate is

$$\dot{\boldsymbol{p}}_{tot} = \int \frac{k\boldsymbol{F}}{c} dV = \oint \frac{u}{3} d\boldsymbol{A} = \oint \frac{kL}{4\pi cr} d\boldsymbol{A}. \qquad (20)$$

If the source is located near a cell center then the net momentum is zero even though the correct radially outward value, corresponding to a sphere with differential surface area $d\boldsymbol{A} = r^2 d\Omega \hat{\boldsymbol{r}}$, is

$$\dot{p}_{r,tot} = [\dot{\boldsymbol{p}}_{tot} \cdot \hat{\boldsymbol{r}}]_{r=const} = \frac{\tau(<r)L}{c}. \qquad (21)$$

The neighbor method conserves the total momentum because the integration is always directed toward an interface with positive vector orientation. However, as discussed by Hopkins & Grudić (2019), by representing cell interfaces as planes with finite area a fraction of momentum is lost due to residual vector cancellation in the transverse directions. To gain intuition, we approximate cells as having $N \gg 1$ faces subtending equal solid angles of $4\pi/N$. We further employ the small-angle approximation and approximate each face as a disk at a distance $z$ with relative differential surface area $d\boldsymbol{A}/z^2 = 2\pi\varphi d\varphi \hat{\boldsymbol{z}}$ and maximum conic opening angle $\varphi_{max} \approx 2/\sqrt{N}$. Thus, the total relative momentum rate is

$$\frac{\dot{p}_{N,tot}}{\dot{p}_{r,tot}} \approx \frac{N}{2} \int_0^{\varphi_{max}} \frac{\varphi \, d\varphi}{\sqrt{1+\varphi^2}} \approx 1 - \frac{1}{N} + \mathcal{O}(N^{-2}), \qquad (22)$$

so a Voronoi tessellation with a larger number of neighbors will have a smaller geometric loss than a regular Cartesian grid. We note that even when the point source is close to a cell boundary the geometric efficiency is still greater than $\dot{p}_{z,tot} / \dot{p}_{r,tot} \geqslant \int_0^1 \rho\sqrt{1-\rho^2} \, d\rho / \int_0^1 \rho \, d\rho = 2/3$ there.

On the other hand, if a similar source is located near a cell interface then a purely neighbor-based method encounters a scenario in which momentum is conserved but imparted to the wrong cells. This is because the cells hosting the photons are transparent to the momentum, even though this is where the unresolved absorption and scattering actually takes place. One solution is to construct a local tessellation around each source

---

[7] We recognize that alternative approaches could give qualitatively similar results. For example, the RTE might admit an exact or approximate solution including only local sources, which would provide the initial conditions for nonlocal MCRT with standard volume-integrated momentum coupling. On the other hand, one might treat self-canceling momentum as a source of turbulent energy or pressure in the hydrodynamics equations.

as is done by Hopkins & Grudić (2019), but this is not practical for MCRT. Instead we propose a scheme which interpolates between the volume- and face-integrated methods, such that the total momentum is split between the host and neighbor cells. Specifically, in highly diffusive regions we expect back-scattering of photons to cancel out most momentum depositions, consistent with a force multiplier of $\sim\tau$ despite having $\sim\tau^2$ scattering events. To reproduce this physical self-cancellation and ensure the momentum flux is conserved but small compared to the energy content, we apply half of the scattering momentum to the host and half to the neighbor. Note that momentum imparted from absorbed photons can safely be applied directly to the forward neighbor. We enforce positive orientation by comparing the path to the center of mass $r_{\rm host}$, corresponding to a distance of

$$\Delta\ell_{\rm host} \equiv \boldsymbol{n}_k \cdot (\boldsymbol{r}_{\rm host} - \boldsymbol{r}_k). \tag{23}$$

There are three cases: (i) if $\Delta\ell_{\rm host} \geqslant \Delta\ell_k$ then the momentum is shared with the backward neighbor; (ii) if $\Delta\ell_{\rm host} \leqslant 0$ then the momentum is shared with the forward neighbor; and (iii) otherwise the momentum from Equation (18) is split between the host and both neighbors according to the inward $\Delta p^\triangleleft$ and outward $\Delta p^\triangleright$ oriented portions.[8] For clarity, in Figure 1 we provide a diagram illustrating examples of each of these cases along with a concise implementation table. We also include a schematic diagram of the basic ideas motivating the design of the neighbor momentum coupling scheme.

### 2.8. Adaptive Convergence

Noise is an inherent feature of MCRT, which can compromise simulation accuracy when coupling to the hydrodynamics if convergence is not reached. Fortunately, the error is straightforward to quantify because the photon discretization is essentially a Poisson process and the relative signal-to-noise ratio increases with the number of photon packets as SNR $\propto \sqrt{N}$. For path-based Monte Carlo estimators the distribution for continuous energy deposition $\Delta E$ (from Equations (16) and (17)) can be highly unpredictable. However, if we estimate the discretized path integration as a weighted Poisson process then the relative error in cell $i$ is given by[9]

$$\delta_i \equiv \frac{\sqrt{\sum \Delta E^2}}{\sum \Delta E}. \tag{24}$$

In practice, there are also a fixed number of pre-existing packets carried over from previous time steps. These contribute to the overall path statistics but their variance contribution cannot be reduced. Our approach is to reweight Equation (24) by the appropriate relative energy, i.e., $\sum \Delta E / (\sum \Delta E + \sum \Delta E_{\rm pre})$. This is equivalent to ignoring the pre-existing variance in on-the-fly estimates, but properly accounts for adaptive convergence



**Figure 1.** Schematic diagram to motivate the optional neighbor momentum coupling scheme with brief pros and cons of each method. The accompanying diagram illustrates the implementation, which interpolates between volume- and face-integrated methods. First, if $\Delta\ell_{\rm host} \geqslant \Delta\ell_k$, then the momentum is shared with the backward neighbor. Second, if $\Delta\ell_{\rm host} \leqslant 0$, then the momentum is shared with the forward neighbor. Otherwise, the total momentum is split between the host and both neighbors according to the inward $\Delta p^\triangleleft$ and outward $\Delta p^\triangleright$ oriented portions as described in the text.

in time-dependent simulations. We note that the property of diminishing returns can necessitate a large number of photons for high-resolution three-dimensional simulations. However, with enough computational resources the noise can always be maintained below a specified tolerance level, e.g., $\delta_{\rm goal} = 0.1$ for smaller than ten per cent local error.

Additionally, the global number of photons required to achieve convergence strongly depends on the overall conditions of the gas and radiation. Therefore, we implement an adaptive convergence scheme such that the emission and propagation of new photons during each time step is performed iteratively in batches.[10] We employ a threshold-based metric to ensure the unconverged fraction remains low, e.g., $f_{\rm goal} = 0.1$ for over ninety per cent confidence of global convergence. Specifically, the effective fraction is

$$f \equiv \sum_{\delta_i > \delta_{\rm goal}} w_i, \tag{25}$$

where the sum is over all unconverged cells, i.e., $\delta_i > \delta_{\rm goal}$, and the cell weight is proportional to the new energy density from the time step $w_i \propto \sum \Delta E / V_i$ normalized such that $\sum w_i = 1$. The size of subsequent batches is based on the current level of convergence. In our current implementation we employ a rapid

---

[8] Equation (18) satisfies the additive property that the total momentum can be separated into disjoint parts: $\Delta p = \Delta p^\triangleleft + \Delta p^\triangleright$. In practice, we compute the inward oriented absorption portion as $\Delta p_a^\triangleleft = \frac{\varepsilon_k}{c}(1 - e^{-\tau_{\rm a,host}})$ where $\tau_{\rm a,host} = k_a \Delta\ell_{\rm host}$, the scattering portion as $\Delta p_s^\triangleleft = \Delta p_a^\triangleleft k_s / k_a$, and all momenta are applied in the $\boldsymbol{n}_k$ direction. Finally, we note that the backward neighbor corresponds to the shortest face distance in the $-\boldsymbol{n}_k$ direction and can be found when searching for the forward neighbor.

[9] In this subsection we abbreviate some notation to simplify the discussion of statistical moments, i.e., we do not explicitly include repetitive cell and path subscripts such as $(\sum_{\rm paths} \Delta E)_i$.

[10] Energy quantities depend on the cumulative number of photon packets. Thus, to ensure consistency a minor memory cost is associated with adaptive convergence to store both the previous and current deposition arrays.

growth mode to find the order of magnitude for the number of photons needed for convergence and a slower reduction mode until the threshold is attained, i.e., $f < f_{goal}$. If this is not satisfied then to avoid overcompensating during initial iterations we multiply the unconverged fraction by a factor of $1 - \Delta N_{ph}/2N_{ph}$, where $\Delta N_{ph}$ is the number of photons in the most recent batch and $N_{ph}$ is the cumulative total from all batches. If $f > 1/2$ the growth mode increases the number of photons in the next batch by a factor of $2^{2f-1}$, otherwise the reduction mode decreases the next batch by a factor of $1.1^{-\log(2f)/\log(2f_{goal})}$. For additional control we also limit the batch sizes between minimum and maximum values such that $\Delta N_{ph,min} < \Delta N_{ph} < \Delta N_{ph,max}$.

Finally, we note that the variance-to-mean ratio may lead to artificial convergence if the contributions are from highly skewed distributions. In this case it is possible to probe higher moments, such as the variance of the variance (VOV), which measures the relative statistical uncertainty in the estimated relative error and can be approximated as $VOV = \sum(\Delta E - \sum \Delta E)^4 / \sum(\Delta E - \sum \Delta E)^2$. Such statistics may still not provide the full picture but can be easily calculated on the fly and included in output files as a way to intelligently lower resolution to ensure sufficiently high signal to noise for internal and observed quantities. Still, MCRT convergence schemes may also benefit from tests that are not based on extrapolation. For example, one might consider the variance between chains of photon trajectories as is done with the Gelman–Rubin diagnostic widely employed in Bayesian inference (Gelman & Rubin 1992). We expect such tests to be more robust but also difficult to implement in practical applications so we leave this for future studies.

## 2.9. Discrete–Diffusion Monte Carlo

When the mean free paths of photons are unresolved in a simulation setup those cells are within the radiative diffusion regime. In this case the transport term can be approximated by an isotropic diffusion process. Specifically, we apply Fick's law as a closure relation to the zeroth order moment equation, such that the radiative flux is proportional to the energy density gradient $\boldsymbol{F} = -c\nabla u/3k_s$. The basic form of the RTE without source terms is now[11]

$$\frac{1}{c}\frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{\nabla u}{3k_s}\right) \equiv \mathcal{L}u. \qquad (26)$$

Since AREPO is a finite-volume code based on a Voronoi tessellation of mesh generating points, we recast the local linear operator on the right-hand side of Equation (26) into the form (for a similar discussion including anisotropic diffusion see Kannan et al. 2016)

$$\mathcal{L}u = \lim_{V \to 0} \frac{1}{V} \int \nabla \cdot \left(\frac{\nabla u}{3k_s}\right) dV, \qquad (27)$$

and apply Gauss's divergence theorem to get

$$\mathcal{L}u = \lim_{V \to 0} \frac{1}{V} \oint \frac{\nabla u}{3k_s} \cdot d\boldsymbol{A}. \qquad (28)$$

Therefore, the discretized radiation energy density in a finite-volume scheme on an unstructured mesh for each cell $i$ over all neighbor cells $\delta i$ is

$$\mathcal{L}u_i = \sum_{\delta i} \frac{A_{\delta i}}{V_i} \frac{(u_{\delta i} - u_i)}{3\Delta\tau_{s,\delta i}} \equiv \sum_{\delta i} k_{leak}^{\delta i}(u_{\delta i} - u_i), \qquad (29)$$

where $V_i$ is the current cell volume, $A_{\delta i}$ is the area of the shared face, and $\Delta\tau_{s,\delta i} \equiv (k_{s,i} + k_{s,\delta i})\Delta r_{\delta i}/2$ is the optical depth between the two mesh generating points, i.e., $\Delta r_{\delta i} \equiv \|\boldsymbol{r}_{\delta i} - \boldsymbol{r}_i\|$ with the cell interface halfway between.[12] We note that the general forms for the "leakage coefficients" $k_{leak}^{\delta i}$ in Equation (29) reduce to the expressions previously found for non-uniform Cartesian and spherical geometries (e.g., Densmore et al. 2007; Abdikamalov et al. 2012; Tsang & Milosavljević 2018). In the MCRT interpretation this discretization of the diffusion operator provides the mechanism for spatial transport of photon packets, with the final form of Equation (29) arranged to highlight photon flux conservation across cell interfaces.

### 2.9.1. Semi-deterministic DDMC Momentum Deposition

The MC procedure and RHD coupling are similar to those of continuous MCRT. However, the optical depth to scattering and distance to the neighboring cell are replaced by an effective distance to leakage drawn from an exponential distribution, i.e., $\Delta \ell_k = -\ln \zeta / k_{leak}^{\delta i}$. The traversed optical depth is then $\tau = k\Delta\ell_k$, such that the energy deposition and residence energy density are still given by Equations (16) and (17), respectively. However, the typical momentum imparted according to Equation (18) is too large by a factor of $\tau_{s,\delta i}$, which can be seen via direct substitution of the leakage coefficient: $\langle \Delta p_{tot} \rangle \approx \tau \varepsilon_k / c = k\varepsilon_k/ck_{leak}^{\delta i} = 3k\Delta\tau_{s,\delta i}V_i\varepsilon_k/cA_{\delta i} \approx \Delta\tau_{s,\delta i}\tau_i\varepsilon_k/c$, with the final approximation valid for a spherical cell with radius $r_i$ and optical depth $\tau_i = kr_i$. Following after Equation (20) but including both scattering and absorption gives

$$\Delta\dot{\boldsymbol{p}} = \int \frac{k\boldsymbol{F}}{c} dV = \oint \frac{u}{3} d\boldsymbol{A} \approx \sum_{\delta i} \frac{\boldsymbol{A}_{\delta i}}{3} \bar{u}, \qquad (30)$$

where the bar denotes the average value at the cell interface, which is approximately $\bar{u} \equiv (u_i + u_{\delta i}k_i/k_{\delta i})/2$ for an unstructured mesh and also accounts for a changing absorption coefficient across neighbors. To our best knowledge, this semi-deterministic DDMC momentum scheme has not appeared previously in the literature. It is variance reducing and efficiently applied at the end of the MCRT calculations.

### 2.9.2. Hybrid IMC–DDMC

The DDMC method is accurate as long as the diffusion approximation holds within the host cell, which can be violated when transitioning to optically thin regions. Therefore, following Densmore et al. (2007) we implemented a hybrid transport scheme in which DDMC packets can convert to spatially continuous MC packets and vice versa, depending on whether the cell optical depth is sufficiently large, i.e., how $\tau_i = k_{s,i} \min\{\Delta r_{\delta i}\}$ compares to $\tau_{DDMC}$. We briefly summarize the main ideas of the hybrid scheme adopting the "asymptotic

---

[11] Under the continuous absorption method only scattering is included as photons are reweighted according to the $e^{-\tau_a}$ correction factor.

[12] In steady-state radiative equilibrium the flux is $\boldsymbol{F}_\nu = -\frac{4\pi}{3k_\nu}\frac{dB_\nu}{dT}\nabla T$, which implies that the DDMC leakage coefficients should be constructed from the Rosseland mean: $k_R \equiv \int_0^\infty \frac{dB_\nu}{dT} d\nu / \int_0^\infty k_\nu^{-1}\frac{dB_\nu}{dT} d\nu$.

diffusion limit" as the interfacing boundary condition. If $\tau_i < \tau_{\text{DDMC}}$, then the leakage coefficient is redefined to be

$$k_{\text{DDMC}\to\text{MC}}^{\delta i} = \frac{1}{3\Delta r_{\delta i}} \frac{2}{k_{\text{s},i}\Delta r_{\delta i} + 2\lambda}, \qquad (31)$$

where $\lambda \approx 0.7104$ is the constant extrapolation distance (Habetler & Matkowsky [1975]). If this corresponds to the minimum distance then the DDMC packet becomes an MC packet with a random position on the cell interface and an isotropic outward direction. On the other hand, if an MC packet moves into a neighboring cell that is optically thick then it is converted into a DDMC packet in that cell with probability

$$P_{\text{MC}\to\text{DDMC}}^{\delta i} = \frac{2}{k_{\text{s},\delta i}\Delta r_{\delta i} + 2\lambda}\left(\frac{2}{3} + \mu\right), \qquad (32)$$

where $\mu$ is the directional cosine for the MC packet with respect to the cell interface. Otherwise, the packet scatters back into the original cell with a random isotropic inward direction. To have a valid probabilistic interpretation, we require $P_{\text{MC}\to\text{DDMC}}^{\delta i} \in [0, 1]$ for $\mu \in (0, 1]$, which imposes a condition that $\tau_{\text{DDMC}} \gtrsim 2$, although we suggest a more conservative choice of $\tau_{\text{DDMC}} = 5$. To further explore this hybrid scheme, in Appendix B we demonstrate the validity of our new semi-deterministic DDMC momentum scheme from Equation (30) across extreme DDMC–MC transitions.

## 2.10. DDMC with Advection

We now present a new unsplit DDMC scheme to incorporate an advection term into the leakage coefficients, which can be important in the dynamical diffusion regime where $\tau v/c \gtrsim 1$. Following Section 2.9, the basic form of the RTE without source terms is now

$$\frac{1}{c}\frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{\nabla u}{3k_{\text{s}}} - \frac{vu}{c}\right) \equiv \mathcal{L}u. \qquad (33)$$

In the finite-volume framework, we recast the local linear operator on the right-hand side of Equation (33) into the form

$$\mathcal{L}u = \lim_{V\to 0}\frac{1}{V}\int \nabla \cdot \left(\frac{\nabla u}{3k_{\text{s}}} - \frac{vu}{c}\right)dV, \qquad (34)$$

and apply Gauss's divergence theorem to get

$$\mathcal{L}u = \lim_{V\to 0}\frac{1}{V}\oint\left(\frac{\nabla u}{3k_{\text{s}}} - \frac{vu}{c}\right) \cdot dA. \qquad (35)$$

Therefore, the discretized radiation energy density in a finite-volume scheme for each cell $i$ over all neighbor cells $\delta i$ is

$$\begin{aligned}\mathcal{L}u_i &= \sum_{\delta i}\frac{A_{\delta i}}{V_i}\left[\frac{u_{\delta i} - u_i}{3\Delta\tau_{\text{s},\delta i}} - \frac{\bar{v}\bar{u}}{c}\right] \\ &= \sum_{\delta i}\frac{A_{\delta i}}{V_i}\left[\left(\frac{1}{3\Delta\tau_{\text{s},\delta i}} - \frac{\bar{v}}{2c}\right)u_{\delta i} - \left(\frac{1}{3\Delta\tau_{\text{s},\delta i}} + \frac{\bar{v}}{2c}\right)u_i\right],\end{aligned} \qquad (36)$$

where $V_i$ is the current cell volume, $A_{\delta i}$ is the area of the shared face, and $\Delta\tau_{\text{s},\delta i} \equiv (k_{\text{s},i} + k_{\text{s},\delta i})\Delta r_{\delta i}/2$ is the optical depth between the two mesh generating points, i.e., $\Delta r_{\delta i} \equiv \|r_{\delta i} - r_i\|$

with the cell interface halfway between. The bar denotes the average at the cell interface, which to first order is approximately $\bar{u} \equiv (u_i + u_{\delta i})/2$ and $\bar{v} \equiv (v_i + v_{\delta i}) \cdot (r_{\delta i} - r_i)/2\Delta r_{\delta i}$. In the MCRT interpretation this discretization provides the mechanism for spatial transport of photon packets, with the final form of Equation (36) arranged to highlight the asymmetric leakage due to the preferred direction of the gas motion. In fact, the inhibited- or enhanced-leakage coefficients are modified to reflect advective transport, which can be succinctly implemented by noticing that $k_{\text{leak,adv}}^{\delta i} = k_{\text{leak}}^{\delta i}(1 + 3\Delta\tau_{\text{s},\delta i}\bar{v}/2c)$.

We note that negative leakage coefficients are possible when the oncoming flow overwhelms the probability of upstream diffusion, and should be ignored. While this interpretation is physically meaningful, such a logical inconsistency is indicative of either insufficient spatial resolution or that a fully relativistic treatment of radiative transfer is necessary. Of course, even first-order comoving-frame DDMC is not needed for our present applications, but we hope this will serve as a useful tool in the DDMC community. The Doppler correction terms may also be treated in an analogous fashion. Finally, the momentum should also be modified:

$$\begin{aligned}\Delta\dot{p}_{\text{adv}} &= \int \nabla \cdot \left(\frac{v \otimes F}{c^2}\right)dV \\ &= -\oint\left(v \otimes \frac{\nabla u}{3ck}\right) \cdot dA \\ &\approx -\sum_{\delta i}\frac{A_{\delta i}\bar{v}}{3c\Delta\tau_{\delta i}}(u_{\delta i} - u_i),\end{aligned} \qquad (37)$$

such that the momentum correction over the time step is

$$\Delta p_{\text{adv}} \approx -\sum_{\delta i}k_{\text{leak}}^{\delta i}V_i\Delta t(u_{\delta i} - u_i)\frac{\bar{v}}{c}. \qquad (38)$$

## 3. Test Problems

### 3.1. Gray Diffusion

To test the spatial transport of the MC particles we consider pure scattering in optically thick media. In the case of a constant scattering coefficient this is equivalent to a random walk with a mean free path of $\lambda_{\text{mfp}} = k^{-1}$. For an arbitrary reference length scale $R \gg \lambda_{\text{mfp}}$, the light-crossing and diffusion times are $t_{\text{light}} = R/c$ and $t_{\text{diff}} = \frac{3}{2}\tau t_{\text{light}} = 3kR^2/2c$. Therefore, the evolution of the radiation energy density is governed by a diffusion equation $\partial u/\partial t = (c/3k)\nabla^2 u$ and the solution given an initial point source impulse of energy $\mathcal{E}_0$ is

$$\tilde{u} = \frac{e^{-\tilde{r}^2/2\tilde{t}}}{(2\pi\tilde{t})^{3/2}}, \qquad (39)$$

where we have rescaled into dimensionless units with radius $\tilde{r} = r/R$, time $\tilde{t} = t/t_{\text{diff}}$, and energy density $\tilde{u} = u\,R^3/\mathcal{E}_0$.

Our spatial transport test consists of a low-resolution three-dimensional Cartesian grid initialized with $10^6$ photon packets at the center at $t = 0$. The mesh configuration and resolution do not matter for this test because we output the photon packets and bin their positions in spherical shells for statistics that directly correspond to the analytic solution from Equation (39). We have verified that the cell-based energy density and momentum estimators give the same results, although path-based estimators represent averages over discrete time steps and cell volumes.
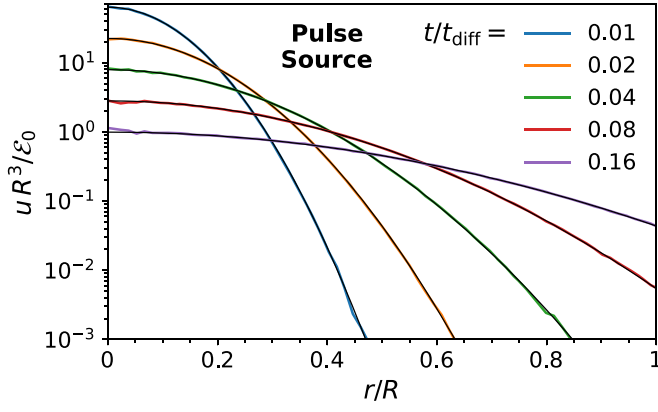
**Figure 2.** Radiation energy density $u(r)$ for a pulse source diffusing in a uniform medium over several doubling times, $t = \{1, 2, 4, 8, 16\} \times 10^{-2} \, t_{\mathrm{diff}}$, where the diffusion time is $t_{\mathrm{diff}} = \frac{3}{2}\tau t_{\mathrm{light}} = 3kR^2/2c$. The analytic solution from Equation (39) is shown by the black curves. For convenience the axes have been rescaled into dimensionless units.
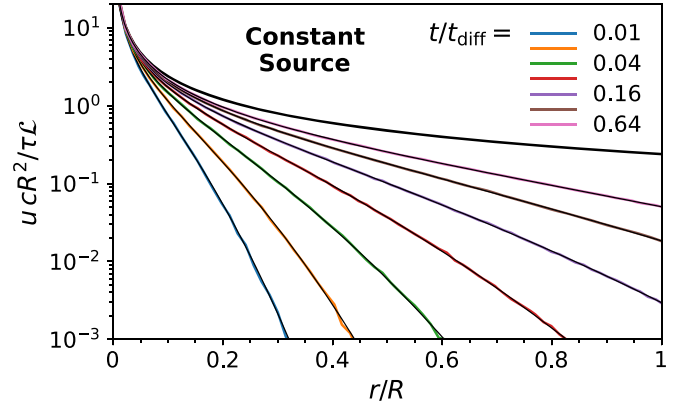


**Figure 3.** Radiation energy density $u(r)$ for a constant source diffusing in a uniform medium over several doubling times, $t = \{1, 2, 4, 8, 16, 32, 64\} \times 10^{-2} \, t_{\mathrm{diff}}$, where the diffusion time is $t_{\mathrm{diff}} = \frac{3}{2}\tau t_{\mathrm{light}} = 3kR^2/2c$. The analytic solution from Equation (40) is shown by the black curves with the final thick black curve showing the steady-state solution, $u = 3k\mathcal{L}/4\pi cr$.

Figure 2 shows the radiation energy density radial profile over several doubling times, $t = \{1, 2, 4, 8, 16\} \times 10^{-2} \, t_{\mathrm{diff}}$. The simulation provides excellent agreement with the exact analytical solution of Equation (39). To simulate an infinite domain, we employ periodic boundary conditions but keep track of the domain tiling to retain the absolute positions of each photon packet.

We also test the spatial diffusion of photon packets under a constant luminosity point source $\mathcal{L}$. This is the same setup as before but the evolution of the radiation energy density is given by

$$\tilde{u} = \frac{3}{4\pi\tilde{r}} \, \mathrm{erfc}\left(\frac{\tilde{r}}{\sqrt{2\tilde{t}}}\right), \qquad (40)$$

where we have again rescaled into dimensionless units with radius $\tilde{r} = r/R$, time $\tilde{t} = t/t_{\mathrm{diff}}$, energy density $\tilde{u} = u \, cR^2/\tau\mathcal{L}$, and the steady-state solution is $\tilde{u}|_{t\to\infty} = 3/4\pi\tilde{r}$. For this test we emit photon packets from the center at a constant rate of $5 \times 10^6/t_{\mathrm{diff}}$. Figure 3 shows the radiation energy density over several doubling times, $t = \{1, 2, 4, 8, 16, 32, 64\} \times 10^{-2} \, t_{\mathrm{diff}}$. The simulation provides excellent agreement with the analytical solution of Equation (40).

With exact solutions in hand we can directly quantify the numerical error of scattering-dominated transport. In Figure 4 we present the volume-weighted $L^1$ error of the radiation energy density, $\int |u - u_{\mathrm{exact}}| \, dV \big/ \int dV$, for the pulse and constant source diffusion tests as a function of the number of radial bins $N_{\mathrm{bins}}$ and photon packets $N_{\mathrm{ph}}$. The errors are normalized such that the total radiation energy is one, even for the constant source that would otherwise grow in time. We are also careful to integrate the analytic solutions from Equations (39) and (40) over the same volumes as the MCRT radial bins. The time-averaged comparisons therefore represent the precise numerical error due to the random walk process. In fact, we recover the expected $\propto \sqrt{N_{\mathrm{bins}}/N_{\mathrm{ph}}}$ noise relations for the number of bins and photon packets. We note that path-based estimators would allow photons to contribute to multiple cells thereby significantly reducing the error beyond what is shown. However, a fair comparison in general three-dimensional geometry would require a more complex treatment of
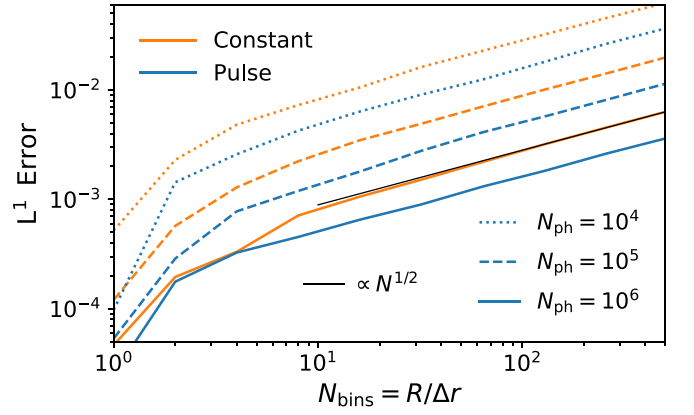


**Figure 4.** Volume-weighted $L^1$ error of the radiation energy density, $\int |u - u_{\mathrm{exact}}| \, dV \big/ \int dV$, for the pulse and constant source diffusion tests as a function of the number of radial bins $N_{\mathrm{bins}}$. We also show results with different numbers of photon packets $N_{\mathrm{ph}} \in \{10^4, 10^5, 10^6\}$ (per diffusion timescale for the constant source case). The errors are normalized to the total radiation energy and time averaged. The MCRT noise follows the expected $\propto \sqrt{N_{\mathrm{bins}}/N_{\mathrm{ph}}}$ relation for the number of bins and photon packets.

cell volumes and integrating over the lagged path time steps. These simple tests stress the need for on-the-fly convergence criteria in MCRT RHD simulations, where the resolution is largely determined by the gas dynamics.

Finally, we demonstrate the speedup that the DDMC scheme provides for scattering-dominated transport. In Figure 5 we present the relative runtime for the pulse and constant source diffusion tests as a function of the cell optical depth resolution $\Delta\tau_{\mathrm{cell}}$. The simulations have a total characteristic optical depth of $\tau = 512$ and are run for a full diffusion timescale on a periodic domain, so pure MCRT has $\gtrsim 10^5$ scattering events per photon. The runtimes are scaled to the fastest DDMC timings ($\approx 1$ second on a laptop computer), which employ a large number of photon packets ($\approx 10^7$). The MCRT timings are approximately constant while the DDMC speedup follows the expected $\propto \Delta\tau_{\mathrm{cell}}^{-2}$ scaling. Deviations are due to various overheads and nuances related to running this simple test while including redundant physics in an unstructured mesh code.
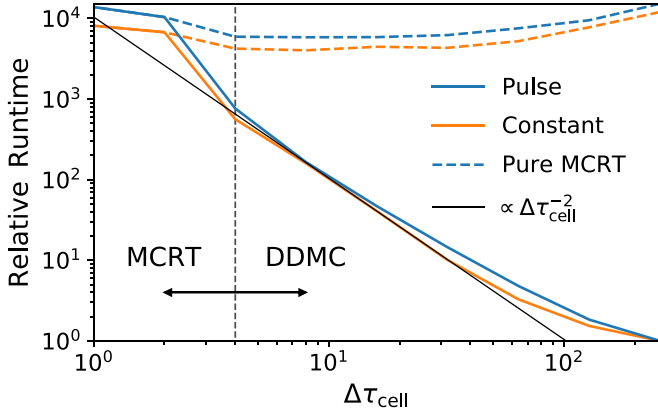
**Figure 5.** Relative simulation runtime for the pulse and constant source diffusion tests as a function of the cell optical depth resolution $\Delta\tau_{\mathrm{cell}}$. Aside from code overheads, these tests demonstrate that pure MCRT is approximately independent of the simulation resolution, while hybrid MCRT–DDMC exhibits the expected $\propto \Delta\tau_{\mathrm{cell}}^{-2}$ speedup from bypassing sub-grid scattering calculations. The simulations have the same characteristic optical depth of $\tau = 512$ and are run for a full diffusion timescale on a periodic domain.

### 3.2. Diffusion with Advection

We now provide two basic tests to demonstrate the accuracy of our DDMC advection scheme. In Figure 6 we validate the ability to capture flows from an Eulerian reference frame by considering the same setup as Section 3.1 but with a constant velocity. By symmetry the solution is the same as the one-dimensional version of Equation (39) but with a coordinate boost of $z \to z - v_0 t$. We set the advection crossing time to be equal to the diffusion time, i.e., $v_0 = R/t_{\mathrm{diff}} = 2c/\tau$. We find that this new unsplit approach is both highly efficient and accurate.

For the second test, we consider the homologous stretching of a one-dimensional infinite plane-parallel slab. In this case the velocity is given by $v(z) = v_0 z/R$, and the evolution of the radiation energy density is governed by the partial differential equation $\partial u/\partial t = (c/k)\partial^2 u/\partial z^2 - (v_0/R)(u + z\partial u/\partial z)$. The solution can be derived with the ansatz that diffusive stretching simply modifies the elapsed time on the global radiation clock. Upon substitution of $\tilde{t} \to \tilde{s}(\tilde{t})$ we find the ansatz reduces the problem to an ordinary differential equation $\tilde{s}'(\tilde{t}) = 1 + 2\tilde{v}_0\tilde{s}(\tilde{t})$ subject to the condition that $\tilde{s}(0) = 0$ (see dimensionless definitions below Equation (41)). Therefore, the full solution for an initial point source impulse of energy $\mathcal{E}_0$ undergoing homologous expansion is

$$\tilde{u} = \frac{e^{-\tilde{z}^2/2\tilde{s}}}{\sqrt{2\pi\tilde{s}}} \qquad \text{where} \qquad \tilde{s} = \frac{e^{2\tilde{v}_0\tilde{t}} - 1}{2\tilde{v}_0}. \qquad (41)$$

We have rescaled into dimensionless units with position $\tilde{z} = z/R$, velocity $\tilde{v}_0 = v_0 kR/2c$, time $\tilde{t} = t/t_{\mathrm{diff}}$, with $t_{\mathrm{diff}} = kR^2/2c$, and energy density $\tilde{u} = uR/\mathcal{E}_0$. In Figure 7 we validate the DDMC advection scheme under this more stringent test. Specifically, we set the characteristic expansion timescale to be equal to the diffusion time, i.e., $v_0 = R/t_{\mathrm{diff}} = 2c/\tau$.

For completeness, we also provide the analytic solution for diffusion under homologous expansion in spherical geometry. The derivation mirrors that of the one-dimensional slab with
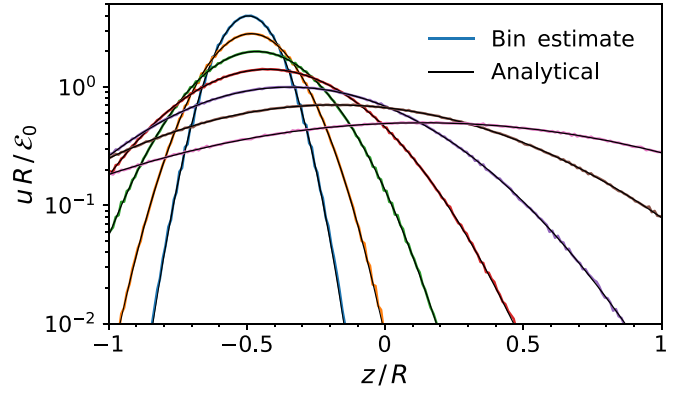


**Figure 6.** Radiation energy density $u(z)$ for gray diffusion undergoing constant relative motion in a uniform medium over several doubling times, $t = \{1, 2, 4, 8, 16, 32, 64\} \times 10^{-2} \, t_{\mathrm{diff}}$, where the diffusion time is $t_{\mathrm{diff}} = \frac{1}{2}\tau t_{\mathrm{light}} = kR^2/2c$. The numerical solution (colored curves) employ the DDMC advection scheme and the analytic solution (black curves) are from Equation (39) but shifted in time. For convenience the axes are in dimensionless units.
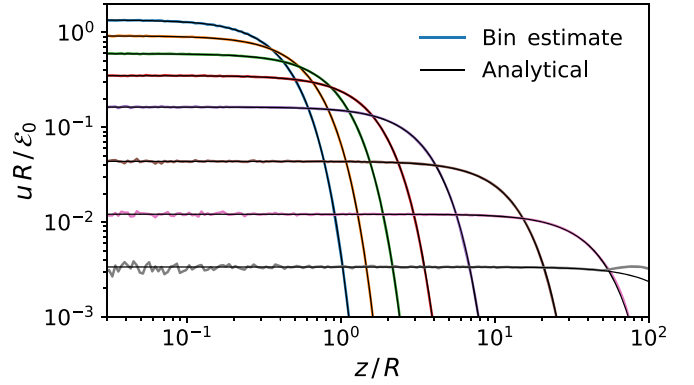


**Figure 7.** Radiation energy density $u(z)$ for gray diffusion in a uniform medium undergoing homologous expansion shown at times of $t = \{1, 2, 4, 8, 16, 32, 48, 64\} \times 10^{-2} \, t_{\mathrm{diff}}$, where the diffusion time is $t_{\mathrm{diff}} = \frac{1}{2}\tau t_{\mathrm{light}} = kR^2/2c$. The numerical solution (colored curves) employ the DDMC advection scheme and the analytic solution (black curves) are from Equation (41). The bump in the final curve is due to an unphysical boundary condition. For convenience the axes are in dimensionless units.

the full solution being

$$\tilde{u} = \frac{e^{-\tilde{r}^2/2\tilde{s}}}{(2\pi\tilde{s})^{3/2}} \qquad \text{where} \qquad \tilde{s} = \frac{e^{2\tilde{v}_0\tilde{t}} - 1}{2\tilde{v}_0}. \qquad (42)$$

We have rescaled into dimensionless units with radius $\tilde{r} = r/R$, velocity $\tilde{v}_0 = v_0 3kR/2c$, time $\tilde{t} = t/t_{\mathrm{diff}}$, with $t_{\mathrm{diff}} = 3kR^2/2c$, and energy density $\tilde{u} = uR^3/\mathcal{E}_0$. We can understand this general stretching behavior by considering that a test particle following the Lagrangian flow from homologous expansion has a physical coordinate of $r(t) = r_0 e^{v_0 t/R}$. This can be derived by induction from a Riemann integration of the velocity field with a starting radius $r_0$ and equal time segments $\Delta t = t/N$ such that $r_N = r_0(1 + v_0 t/RN)^N$, which reduces to the exponential function in the limit as $N \to \infty$. However, the equation of motion is modified by the random walk process in an invariant manner requiring that $\tilde{s} \approx \tilde{t}$ early on. To the best of our knowledge these homologous diffusion solutions are new and can serve as an
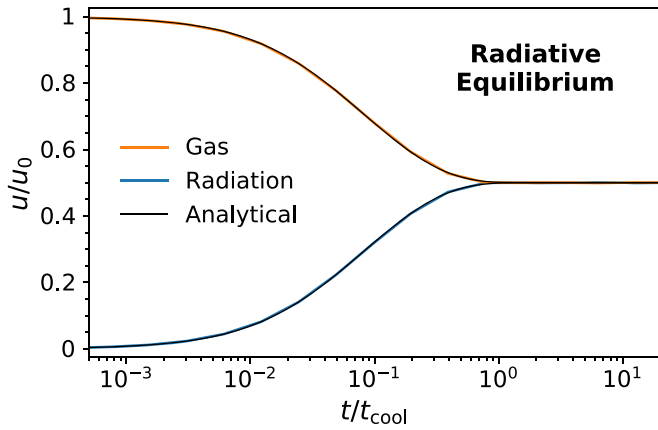
**Figure 8.** Gas and radiation energy densities $u(t)$ to demonstrate the evolution to radiative equilibrium in a uniform medium. Here the cooling time is defined as $t_{cool} \equiv 1/c\, k_a$. The numerical solution from Equation (43) is shown by the black curves. For convenience the axes have been rescaled into dimensionless units.

additional benchmark for codes wishing to accurately model radiation in the dynamical diffusion regime.

### 3.3. Radiative Equilibrium

To test the coupling of the gas internal energy and radiation fields we consider pure absorption in LTE over a uniform medium. Defining $u_r = a_B T^4$ as in Section 2.1, the stiff system of equations governing the gas and radiation energy density evolution is

$$\frac{d\tilde{u}_g}{d\tilde{t}} = -\frac{d\tilde{u}}{d\tilde{t}} = \tilde{u} - \zeta \tilde{u}_g^4. \tag{43}$$

Here we have rescaled the variables such that $\tilde{u}_g \equiv u_g/u_0$, $\tilde{u} \equiv u/u_0$, where the initial gas and radiation energy densities are $u_g|_{t=0} = u_0$ and $u|_{t=0} = 0$. We have also introduced a dimensionless coupling parameter that can be given in terms of the initial temperature and density as $\zeta \equiv a_B T_0^3 (\gamma - 1)\mu/k_B \rho_0$. For this test we set $\zeta = 8$ so that at equilibrium we have $u|_{t\to\infty} = u_0/2$.

Our radiative equilibrium test consists of a one-zone setup, which converges quickly regardless of the number of photon packets. Figure 8 shows the time evolution of the mean energy densities, with the MCRT result providing excellent agreement with the exact solution. For this test we start with a small time step and force subsequent time steps to be twice the previous one. We also enabled IMC with an implicitness parameter of $\alpha = 0.5$, which provides higher-order accuracy for this special test case.

### 3.4. Radiative Shock

We now test the hydrodynamical coupling by considering the formation of both sub- and super-critical radiative shocks following the initial conditions proposed by Ensman (1994), which have been reproduced in numerous RHD implementations (Hayes & Norman 2003; Whitehouse & Bate 2006; González et al. 2007; Commerçon et al. 2011; Noebauer et al. 2012; Tsang & Milosavljević 2015). We perform the test on a moving mesh starting with a box of radius $R = 7 \times 10^{10}$ cm with uniform density $\rho_0 = 7.78 \times 10^{-10}$ g cm$^{-3}$, absorption coefficient $k_a = 3.115 \times 10^{-10}$ cm$^{-1}$, mean molecular weight
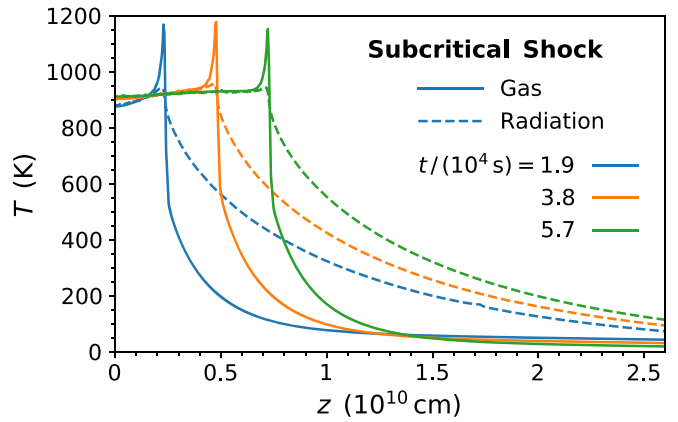


**Figure 9.** Gas and radiation temperatures $T(z)$ for the sub-critical radiative shock test, shown respectively as solid and dashed curves. The initial gas is colliding with a velocity of $v_0 = 6$ km s$^{-1}$. The profiles are given at the times $t = \{1.9, 3.8, 5.7\} \times 10^4$ s.
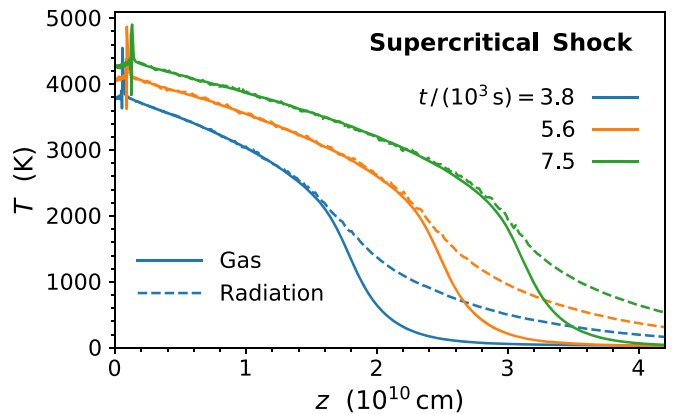


**Figure 10.** Gas and radiation temperatures $T(z)$ for the super-critical radiative shock test, shown respectively as solid and dashed curves. The initial gas is colliding with a velocity of $v_0 = 20$ km s$^{-1}$. The profiles are given at the times $t = \{3.75, 5.625, 7.5\} \times 10^4$ s.

$\mu = 1$, adiabatic index $\gamma = 7/5$, and a linear temperature profile decreasing from $T = 85$ K at the center to $T = 10$ K at the edges of the box. We employ IMC with $\alpha = 1$ and allow MC particles to escape at either boundary. The gases on the left and right are colliding toward the center at constant velocity of $v_0 = 6$ and 20 km s$^{-1}$ for the sub- and super-critical shocks, which generates an outwardly propagating shock wave. The thermal radiation diffuses upstream to pre-heat the pre-shock gas to the post-shock temperature (Zel'dovich & Raizer 1967). The temperature profiles for each test are shown in Figures 9 and 10 at several different times to illustrate the outward propagation. In both cases, our results are in agreement with previous simulations and analytical studies for the jump conditions across the shock (Mihalas & Mihalas 1984). Further verification of RHD codes could also include the semi-analytic radiative shock solutions of Lowrie & Edwards (2008), which are involved in their implementation but provide further insights within the context of gray nonequilibrium radiative diffusion.

## 4. Levitation of Optically Thick Gas

Radiative feedback can play an important role in galaxy formation and evolution by driving supersonic turbulence, reducing the star formation efficiency, and regulating galactic

winds (e.g., Thompson et al. 2005; Hopkins et al. 2011). In particular, systems with extreme star formation rate densities, including so-called ultraluminous infrared galaxies (ULIRGs), can experience efficient photon trapping as direct ultraviolet (UV) starlight is efficiently reprocessed by dust grains to multiscattered infrared (IR) radiation. In such environments the trapping effect boosts the momentum injection rate by a factor of the optical depth $\tau_{IR}$ relative to the intrinsic force budget of $\mathcal{L}/c$. However, in reality the interstellar medium has a hierarchical structure, which facilitates the escape of radiation and the associated momenta through low column density channels. Furthermore, the natural emergence of the Rayleigh–Taylor instability (RTI; Chandrasekhar 1961) in the presence of external forces, such as gravity, may limit the coupling of gas and radiation. Modeling these complexities requires multi-dimensional RHD simulations, which motivated Krumholz & Thompson (2012) to design a two-dimensional setup to investigate the efficiency of radiation pressure driving of a dusty atmosphere in a vertical gravitational field. Subsequently, several other groups have simulated this levitation setup with different codes and RHD methods, including FLD (Krumholz & Thompson 2012; Davis et al. 2014), VET (Davis et al. 2014), M1 (Rosdahl & Teyssier 2015; Kannan et al. 2019), and MCRT (Tsang & Milosavljević 2015). As the model and physics are described in detail by each of these authors, we only provide a brief summary along with the AREPO-MCRT results for comparison with the previous studies.

### 4.1. Simulation Setup

As discussed in Krumholz & Thompson (2012) and the subsequent studies, the goal is to simulate the evolution of molecular gas at temperatures where dust dominates the opacity and radiation is strong enough to trigger the RTI. For simplicity, we assume perfect thermal and dynamic coupling between gas and dust grains. The Planck $\kappa_P$ and Rosseland $\kappa_R$ mean opacities are given by

$$\{\kappa_P, \kappa_R\} = \{0.1, 0.0316\}\left(\frac{T}{10\,\mathrm{K}}\right)^2 \mathrm{cm}^2\,\mathrm{g}^{-1}, \qquad (44)$$

which approximately holds for dusty gas in LTE at $T \leqslant 150\,\mathrm{K}$ (Semenov et al. 2003). At higher temperatures we simply limit both mean opacities to their values at 150 K. We define a characteristic radiation temperature $T_r = (u/a_B)^{1/4}$, which is nearly identical to the gas temperature $T$ except in thin layers of shock-heated gas, e.g., at the front of the wind. For simplicity, the gravitational acceleration $g$ is assumed to be constant (downward) and the radiation field is sourced by a constant flux $F_*$ at the lower boundary (upward). This flux defines a characteristic temperature $T_* = (F_*/a_B c)^{1/4}$, which leads to the definition of a corresponding characteristic sound speed $c_* = \sqrt{k_B T_*/\mu m_H}$, scale height $h_* = c_*^2/g$, and sound crossing time $t_* = h_*/c_*$. As in previous studies we assume $\mu = 2.33$ and choose $T_* = 82\,\mathrm{K}$, corresponding to $F_* = 2.54 \times 10^{13}\,L_\odot\,\mathrm{kpc}^{-2}$ and $\kappa_{R,*} = 2.13\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. With these assumptions the levitation setup of radiation opposing gravity is characterized by two dimensionless numbers, the Eddington ratio

$$f_{E,*} = \frac{\kappa_{R,*} F_*}{g c} \qquad (45)$$

and the optical depth

$$\tau_* = \kappa_{R,*}\Sigma_*, \qquad (46)$$

where $\Sigma_*$ denotes the characteristic surface mass density. For the simulation presented in this paper we use $f_{E,*} = 1/2$ and $\tau_* = 3$, which is sufficient to probe dynamically unstable coupling between radiation and gas.

The initial setup corresponds to an isothermal atmosphere in hydrostatic equilibrium in the absence of radiation. Specifically, the temperature is $T = T_*$ and the vertical density profile is the exponential $\rho(y) = \rho_* \exp(-y/h_*)$, where $\rho_* = \Sigma_*/h_*$ is the characteristic density. We impose a density floor of $10^{-10}\rho_*$ and further induce a perturbation of the form

$$\frac{\delta\rho}{\rho} = \frac{1+\chi}{4}\sin\left(\frac{4\pi x}{L_x}\right), \qquad (47)$$

where $L_x = 512\,h_*$ is the simulation boxsize and $\chi$ is a random number uniformly distributed in $[-1/4, 1/4]$. Furthermore, no radiation is initially present. The boundary conditions are periodic in the $x$ direction, reflective at bottom ($y = 0$), and outflowing at the top ($L_y = 16L_x$), tall enough that no gas is lost during the simulation. The initial mesh consists of a high-resolution $(0.5h_*)^2$ Cartesian mesh at the bottom to resolve the high-density gas. The resolution is degraded slowly upwards until a minimum resolution of $(8h_*)^2$ is reached. As the simulation progresses the mesh moves according to the local fluid flow, is regularlized where needed, and undergoes adaptive refinement and derefinement to approximately maintain cell volumes between 0.25 and 64 $h_*^2$ with a target mass resolution of $\Sigma_* L_x/1024^2$. Finally, we run with an adaptive convergence criteria of $\delta_{goal} = f_{goal} = 0.05$, a luminosity boosting exponent of $1/2$, and the traditional volume-based momentum scheme which is accurate because geometric sources are always at the boundary interfaces and the majority of cells are optically thin.

### 4.2. Results

Figures 11–13 respectively show the evolution of the normalized gas density $\rho/\rho_*$, radiation temperature $T_r/T_*$, and the vertical Eddington ratio $f_{E,y}$, i.e., the ratio of radiative to gravitational accelerations for increasing times. The radiation field remains fairly smooth throughout the simulation and mirrors the density propagation as photons escape when they are able to reach the front of the wind. The trapping is high at early times ($t \approx 25\,t_*$), leading to efficient gas heating, which in turn increases the mean opacity to sustain an initial super-Eddington phase that quickly lifts the gas upwards. As Rayleigh–Taylor instabilities form, the vertical radiation coupling weakens and gravity slowly removes inertia ($t \approx 100\,t_*$). At later times, the photons regain sufficient trapping for levitation to continue ($t \approx 200\,t_*$) throughout the remainder of the simulation. Apparently, the radiation force continues to counterbalance gravity even as many of the dense filamentary structures stall out to what resembles a highly turbulent quasi-steady state.

Figure 14 provides complementary one-dimensional depictions of the evolving normalized gas density $\rho/\rho_*$, vertical (mass-weighted) gas velocity $v_y/c_*$, and radiation temperature $T_r/T_*$. These quantities are calculated in a conservative fashion
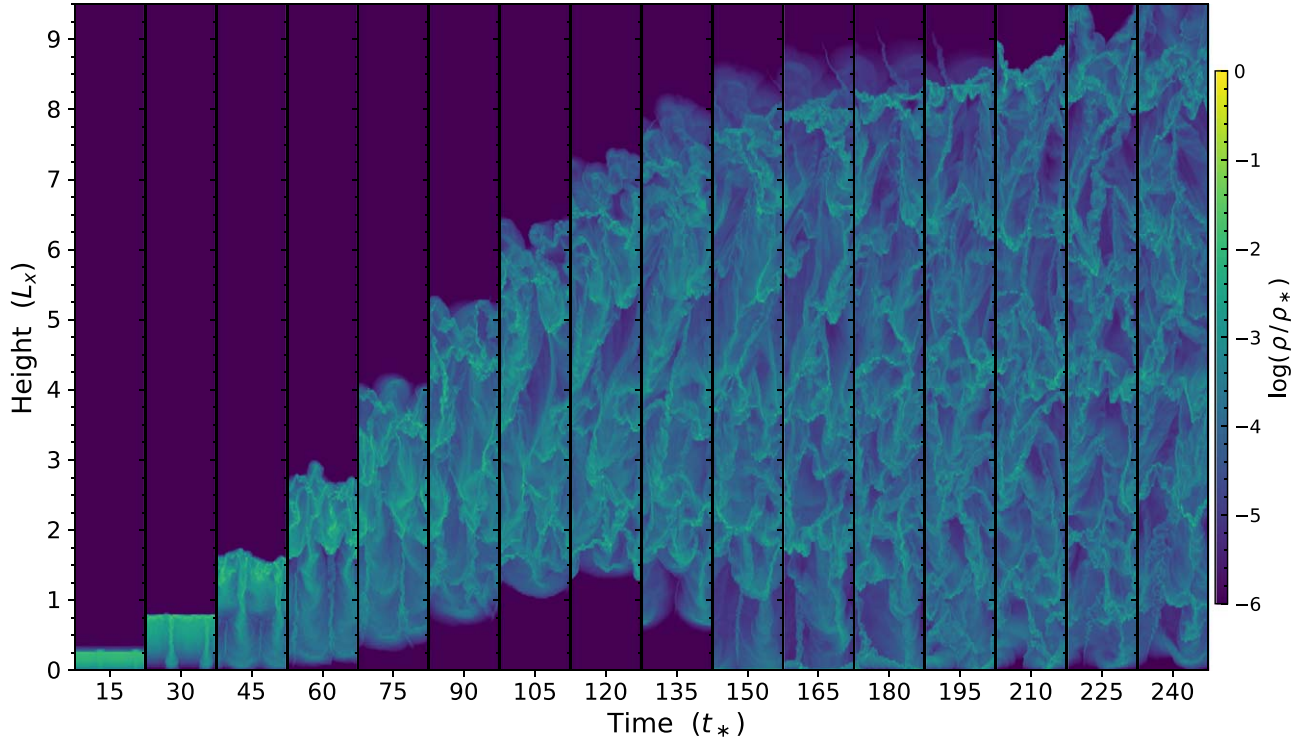
**Figure 11.** Evolution of the normalized gas density $\rho/\rho_*$ shown in intervals of $15\,t_*$. This view emphasizes the successful launch of the wind despite the gas becoming Rayleigh–Taylor unstable early on. Even after some cold filaments fall back down the gas structure remains highly elongated and turbulent in a quasi-steady-state configuration.
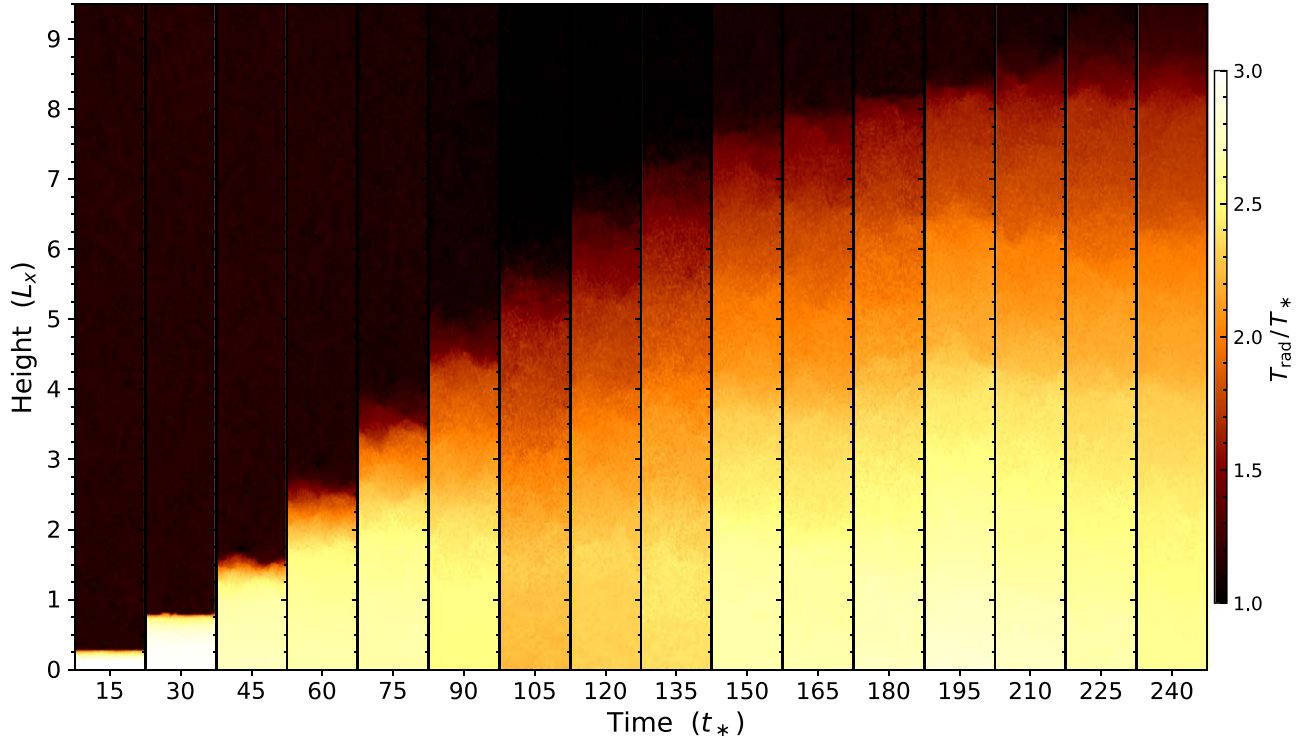


**Figure 12.** Evolution of the normalized radiation temperature $T_r/T_*$, which also closely mirrors the behavior of the gas temperature. The radiation energy is quite smooth due to photon trapping behind the wind front. By $t = 30\,t_*$ the radiation efficiently heats and pushes the gas. The rapid expansion, cooling, and escape channels lead to a noticeably reduced temperature by $t = 100\,t_*$, but this is built up again by $t = 150\,t_*$.

by binning the gas particles with a height resolution of $40\,h_*$. The density shows the gas propagation and remains fairly uniform throughout time despite the significant elongation. The velocity is fastest near the front of the wind but develops a clear wave-like pattern with a peak-to-peak distance of approximately $2\,L_x$. Finally, the radiation temperature derived from the
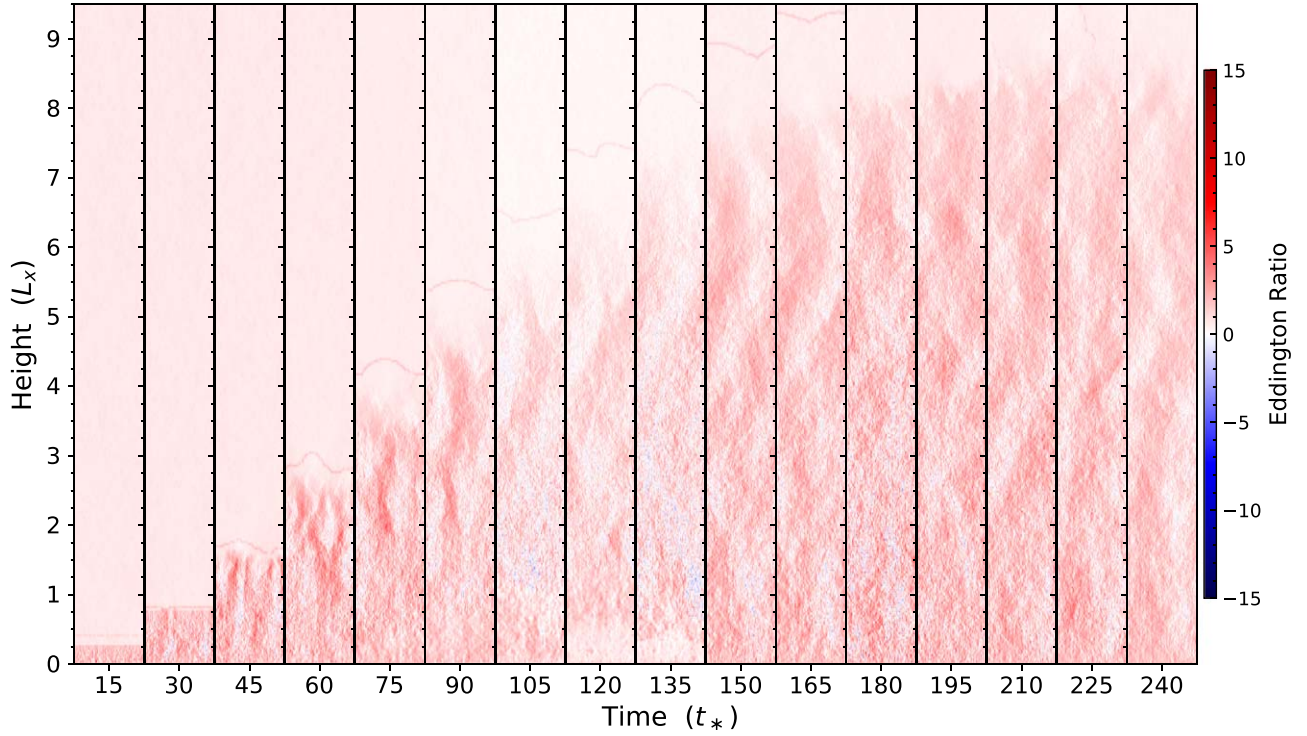
**Figure 13.** Evolution of the vertical Eddington ratio $f_{E,y}$, i.e., the ratio of radiative to gravitational accelerations. The volume is largely dominated by super-Eddington regions except for the dense filamentary clumps that descend until they are disrupted reminiscent of a recycling process.
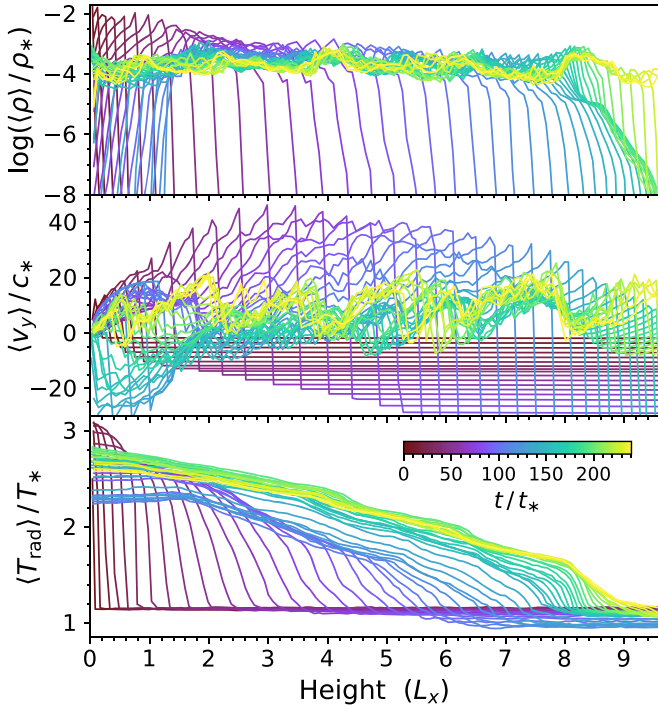


**Figure 14.** Evolution of the normalized gas density $\rho/\rho_*$ (top), vertical mass-weighted velocity $v_y/c_*$ (middle), and radiation temperature $T_r/T_*$ (bottom) as one-dimensional profiles binned with a height resolution of $40\,h_*$ given in time intervals of $5\,t_*$. These quantities demonstrate the early propagation of the wind front and the highly turbulent quasi-steady state at late times. The density becomes fairly uniform, the velocity develops wave-like behavior, and the temperature develops a characteristic gradient to support the continued levitation of the gas.

energy density allows us to better explore the cause of the apparent dimming in Figure 12. These profiles clearly show that the radiation pressure is significantly relieved as newly emitted radiation escapes rather than continuing to build up ($t \approx 100\,t_*$). While the development of the RTI is critical for opening up low column density pathways, this is also a result of self-regulation as the initial driving phase was efficient enough to promote cooling via expansion of both the gas and radiation fields. As the wind stalls the radiation field is eventually reassembled to suppress substantial fallback. Interestingly, a characteristic temperature gradient develops to continue to support the gas at late times.

Figure 15 illustrates the evolution of the mass-weighted height $\langle y \rangle$ (top), vertical velocity $\langle v_y \rangle$ (middle), and velocity dispersion $\sigma$ (bottom) normalized to the appropriate characteristic scales. The initial gas acceleration resembles the result from AREPO-RT, while at late times follows the behavior of the previous VET and MCRT studies. We find a terminal velocity of $>20\,c_*$, which is higher than all other results in the literature. The height reaches $\langle y \rangle \approx 2500h_*$ by the end of the simulation. The mass-weighted vertical velocity never falls significantly below zero, and remains highly turbulent for most of the run. To further explore the behavior of the radiation properties in Figure 16 we show additional globally averaged quantities. The net Eddington ratio (top) is defined as

$$f_{E,V} = \frac{\sum f_{y,\mathrm{rad},i}}{\sum g m_i}, \qquad (48)$$

where the sums are over all gas cells $i$. During the initial acceleration phase the Eddington ratio is high $\gtrsim 1.5$ for a prolonged period but is mostly $\lesssim 1$ thereafter. The mean
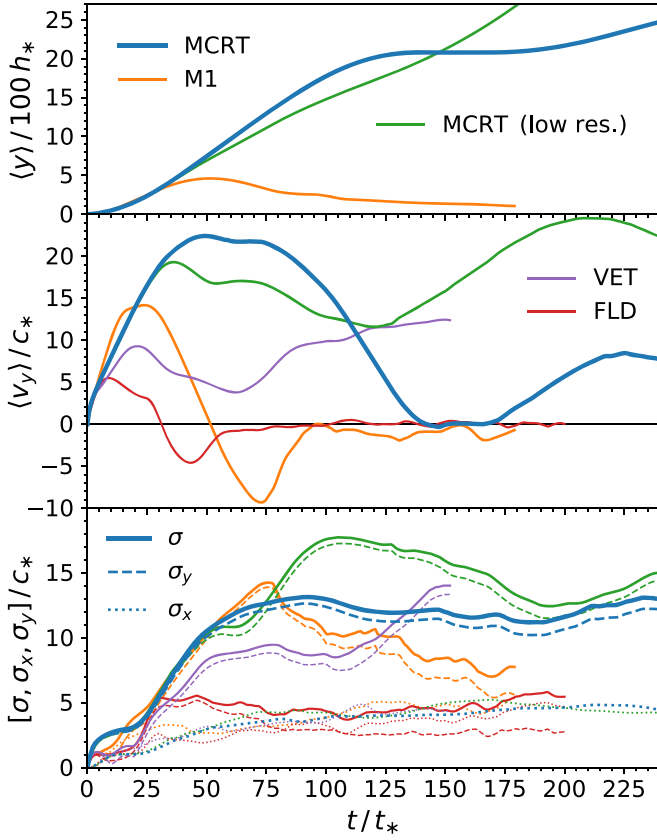
**Figure 15.** Evolution of the global mass-weighted height $\langle y \rangle$ (top), vertical velocity $\langle v_y \rangle$ (middle), and velocity dispersion $\sigma$ (bottom). We compare our MCRT method (blue curves) to the M1 scheme of AREPO-RT (orange curves; Kannan et al. 2019) along with VET and FLD results (purple and red curves; Davis et al. 2014), and find a significant difference in the ability to lift the gas upwards. The velocity reaches $\gtrsim 20\, c_*$ and despite the formation of instabilities remains positive even at late times. We also show results from a simulation with a target mass resolution that is four times lower than the main run (green curves).

vertical optical depth (middle) is

$$\tau_V = \frac{\sum k_{R,i} V_i}{L_x}, \qquad (49)$$

while the flux-weighted optical depth is

$$\tau_F = \frac{\sum V_i \sum f_{y,\mathrm{rad},i}}{L_x \sum f_{y,\mathrm{rad},i} / k_{R,i}}, \qquad (50)$$

given in terms of the radiation force and absorption coefficient $k_R$. From Figure 16 it is clear that the mean trapping optical depth remains high throughout the simulation around $\tau_V \approx 9$. The bottom panel shows the ratio $\tau_V / \tau_F$, which is close to unity at the beginning but drops to lower values ($\approx 0.6$–$0.8$) for the remaining time as a reflection of the net traversal through lower opacity channels between filaments.

These results are quite interesting when compared to previous results. In particular, in the figures we highlight the comparison to the results from AREPO-RT, as this uses the same hydrodynamics code but with a second-order M1 closure scheme (orange curves; Kannan et al. 2019). For additional comparison we also show results from the FLD (red curves) and VET (purple curves) methods (Davis et al. 2014). The mass-weighted gas height and velocity are significantly
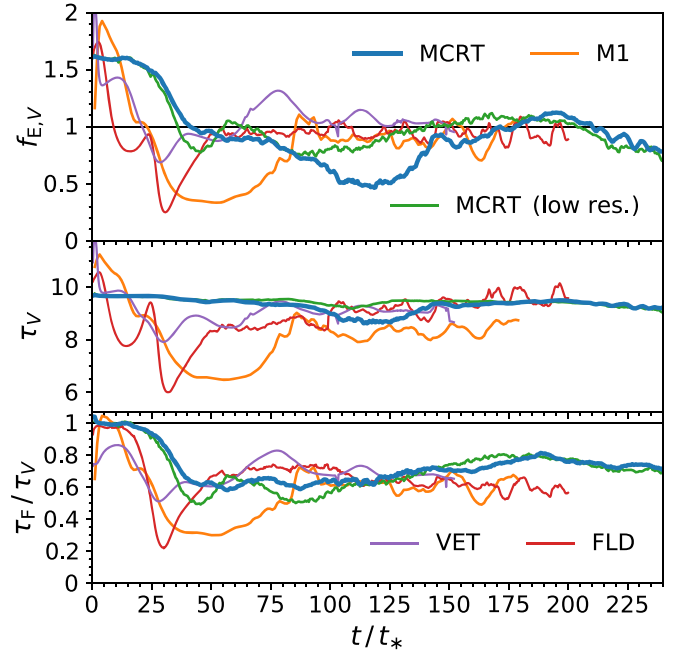


**Figure 16.** Evolution of the global volume-weighted Eddington ratio $f_{E,V}$ (top), vertical optical depth $\tau_V$ (middle), and ratio of vertical flux- to volume-weighted optical depths $\tau_F / \tau_V$ (bottom). There is strong upward driving at early times ($t \lesssim 25\, t_*$), which is reduced but still relatively effective throughout the remainder of the simulation. We again compare our MCRT results (blue and green curves) to those obtained from the M1 (orange curves), VET (purple curves), and FLD (red curves) methods.

different, with the MCRT method launching a successful wind while M1 falls back down after an initial strong liftoff. This emphasizes the known result that moment-based methods and ray-tracing short/long characteristic RT methods differ in the long-term evolution. Specifically, the MCRT and VET methods avoid the characteristic long-term fate of the M1 and FLD methods to end up with turbulent gas that is gravitationally confined at the bottom of the domain. The gas dynamics can be significantly different between different radiative transfer methods, but the long-term results may also be sensitive to what happens during the initial acceleration phase as there are even differences between the previous MCRT implementation by Tsang & Milosavljević (2015), although their results are similar to our lower resolution run. We also note that we followed their choice to cap the opacities $\kappa_{R,P}$ at their values at $T = 150$ K, whereas all other authors allow the $\kappa \propto T^2$ scaling to arbitrary temperatures. Despite the opacity ceiling we still see the strongest overall RHD response compared to all other studies.

Finally, we briefly comment on a few numerical considerations. To determine how simulation resolution affects our results we performed an identical run maintaining cell volumes between 1 and $64\, h_*^2$ with a target mass resolution of $\Sigma_* L_x / 512^2$, which is four times lower than the main run. The results are shown as green curves in Figures 15 and 16, which exhibit identical behavior until about $35\, t_*$ when the instabilities are fully developed. Interestingly, at this point the chaotic behavior results in qualitatively different velocity histories with the lower resolution run being slightly slower during the first peak inertial phase but only falling to about $12\, c_*$ at $120\, t_*$ and rising with a second wind thereafter. This is similar to the MCRT results by Tsang & Milosavljević (2015),

which had resolution comparable to our low-resolution run ($1\,h_*^2$). Davis et al. ([2014](#)) also discuss the impact of resolution on the Eddington ratio and development of density inhomogeneities between their fiducial and lower ($1\,h_*^2$) resolution simulations. Neither of our studies demonstrate resolution convergence but we agree that higher resolution enhances the impact of radiative feedback in this test setup. More importantly, there is a larger difference between the treatment of radiation and hydrodynamics. To explore this further, we also ran a numerical experiment in which all photon packets are merged in each cell at the end of each time step. This emulates a crude moment-based approximation with poor flux preservation designed to test the importance of accurately representing intersecting rays from nonlocal radiation sources. In such simulations the wind fails with similar long-term behavior as the FLD and M1 results.

## 5. Summary and Discussion

In this paper, we have presented AREPO-MCRT, a novel implementation of a highly accurate Monte Carlo radiative transfer RHD method in the moving-mesh code AREPO. The scheme uses a first-principles approach to sample the radiation field one photon trajectory at a time. The flow of a large but finite number of independently simulated photon packets provides a statistical representation of the collisionless radiation transport problem. The basic ideas are conceptually simple but the $\sqrt{N}$ rate of convergence requires variance reduction and importance sampling techniques to be competitive with other RHD methods in terms of accuracy to computational cost. We have incorporated many of the strategies employed throughout the community to overcome the efficiency barriers inherent to MCRT. Beyond this, we have invested significant effort to develop concepts, discussion, and algorithms to meet the unique needs of AREPO-MCRT. In the long term, we intend to target a variety of multiple scattering problems relevant to astrophysical applications.

We tested our implementation on a variety of standard problems, and accurately reproduced the time-dependent transport and coupling effects in each instance. Specifically, we demonstrated the accurate diffusion of pulse and constant sources of radiation, simulating an infinite domain by tracking the tiling within periodic boxes. We also explore the $L^1$ error and hybrid DDMC speedup for these tests. We then derive a new analytic solution combining diffusion with homologous expansion to test our DDMC advection scheme. We resolved the evolution of rapidly cooling gas via thermal emission to a final state of radiative equilibrium. We also tested the hydrodynamical coupling by considering the formation of both sub- and super-critical radiative shocks.

Finally, we explored the ability of a trapped IR radiation field to accelerate a layer of gas in the presence of an opposing external gravitational field. We found persistent radiation-driven levitation even after the formation of Rayleigh–Taylor instabilities that create dense filaments and chimneys promoting the escape of photons. Our results are in agreement with previous VET and MCRT studies, as the photon directions are well preserved in these methods. On the other hand, the long-term behavior of the FLD and M1 closure moment-based methods is that of a highly turbulent quasi-steady state concentrated at the bottom of the domain. We note that due to efficient trapping at early times we obtain higher initial

ballistic velocities than all previous studies. An important insight from our simulations is that the initial acceleration phase and the revitalized second wind are connected via self-regulation of the RHD coupling. As a consequence, the RHD implementation and simulation resolution are both crucially important when going beyond qualitative long-term effects.

We also emphasize that AREPO-MCRT can be used to post-process existing simulations to obtain accurate representations of radiation fields and emergent observables. In fact, MCRT is often the *de facto* method of choice for multi-frequency three-dimensional radiative transfer calculations when considering scattered and reprocessed light. While we do not anticipate that AREPO-MCRT will replace existing post-processing pipelines, there is a natural place for built-in and on-the-fly tools with a common codebase tightly coupled to the original simulation. Specifically, the code inherits the efficient mesh construction, domain decomposition, and other strategies that are increasingly necessary for state-of-the-art hydrodynamics simulations, while at the same time introducing new algorithms and data structures that address the unique challenges of the computationally demanding nonlocal photon transport.

In the current state, it is not yet feasible to perform full galaxy formation simulations. However, the RHD solver will undergo continued development to address the performance needs of various applications. Importantly, this includes allowing compatibility with the individual time-stepping scheme of AREPO and improving the strong scaling with the number of computational domains. Depending on the application, other strategies for MCRT physics algorithms and parallelization efficiency would also likely benefit our implementation (e.g., Harries et al. [2019](#); Michel-Dansac et al. [2020](#); Vandenbroucke & Camps [2020](#)). Still, even the high-resolution levitation setup can be run on a single compute node within a reasonable amount of time. This is partially due to the rapid adaptive convergence of the trapped radiation field and the ability to take longer time steps with the implicit Monte Carlo scheme, even when employing the full speed of light in transport calculations.

In future work, we plan to use this implementation to study timely problems in astrophysics. This includes the goal to study Lyman-$\alpha$ (Ly$\alpha$) radiation pressure with the first self-consistent three-dimensional Ly$\alpha$ RHD simulations. Previous studies have shown that resonant scattering by trapped Ly$\alpha$ photons can have a dynamical impact in dust-poor environments (Smith et al. [2017b](#), [2019](#); Kimm et al. [2018](#)). We plan to study these phenomena with the Ly$\alpha$ radiative transfer functionality that is already implemented within AREPO-MCRT. The full RHD interface will also take advantage of the new resonant DDMC scheme proposed by Smith et al. ([2018](#)) to break the efficiency barrier of frequency redistribution in this physical regime. Overall, it will be a valuable endeavor to push the limits of MCRT RHD schemes to provide an accurate and robust understanding of the role of radiation fields throughout the universe.

# Appendix A
## Spherical Coordinates

Ray tracing in spherical coordinates can be reduced to finding the intersection between a line and sphere. We take $\boldsymbol{r}$ and $\boldsymbol{n}$ to be the photon position and direction before traversal such that the radius is bounded by inner and outer radii, i.e., $r \equiv \|\boldsymbol{r}\| \in [r_-, r_+]$. Thus, the length is determined by $\|\boldsymbol{r} + \ell\boldsymbol{n}\|^2 = r_\pm^2$, which admits solutions of

$$\ell_\pm = -\mu r \pm \sqrt{r_\pm^2 - r_{\min}^2}, \qquad (A1)$$

where the unnormalized radial cosine is $\mu r \equiv \boldsymbol{n} \cdot \boldsymbol{r}$, the impact parameter is $r_{\min}^2 \equiv (1 - \mu^2)r^2$, and only positive lengths are intersections. In practice, if $\mu \geqslant 0$ or the inner discriminant $r_-^2 - r_{\min}^2 \leqslant 0$ (but $\mu < 0$), then the photon packet traverses toward the outer shell boundary. Otherwise, we use the inner solution. To calculate a radial momentum similar to Equation (18) we integrate the outward contribution along the path, which yields

$$\Delta p_r = \frac{k\varepsilon_k}{c} \int_0^\ell \hat{\boldsymbol{r}} \cdot (\boldsymbol{r} + \ell'\boldsymbol{n}) d\ell'$$
$$= \frac{k\varepsilon_k}{c} \int_0^\ell \frac{\ell' + \mu r}{\|\boldsymbol{r} + \ell'\boldsymbol{n}\|} d\ell'$$
$$= \frac{k\varepsilon_k}{c}(\|\boldsymbol{r} + \ell\boldsymbol{n}\| - r) \equiv \frac{\tau_r \varepsilon_k}{c}, \qquad (A2)$$

where $\tau_r \equiv k\Delta r$ is the net radial optical depth traversed. Unfortunately, there is no simple analytic formula when continuous absorption is taken into account by adding a factor of $e^{-\tau_a}$ inside the integral. However, in this case we can either (i) track the momentum flux through interfaces, (ii) switch to a probabilistic absorption scheme, or (iii) create an approximate numerical solution from

$$\Delta p_r = \frac{k\varepsilon_k}{c} \int_{\mu r}^{\mu r + \ell} \frac{\ell' e^{k_a(\mu r - \ell')}}{\sqrt{\ell'^2 + r_{\min}^2}} d\ell'$$
$$\approx \frac{k\varepsilon_k}{k_a c} e^{k_a(\mu r - r_{\min})}$$
$$\times \left[ \begin{array}{ll} \dfrac{e^{k_a(r_{\min} - \ell')}(1 + k_a\ell') - (1 + k_a r_{\min})}{k_a r_{\min}} & |\ell'| < r_{\min} \\[2ex] \mathrm{sgn}(\ell')(1 - e^{k_a(r_{\min} - \ell')}) & |\ell'| \geqslant r_{\min} \end{array} \right]_{\mu r}^{\mu r + \ell} \qquad (A3)$$

with the last line being an example of a simple but relatively accurate option. We emphasize that this is only consequential

when there is significant absorption along the path segment. In fact, the center of energy distance along the ray is

$$\langle \ell \rangle = \ell^{-1} \int_0^\ell \ell' e^{-k_a\ell'} d\ell' = \frac{\ell}{\tau_a^2}(1 - (1 + \tau_a)e^{-\tau_a}), \qquad (A4)$$

corresponding to a center of energy position of $\langle \boldsymbol{r} \rangle = \boldsymbol{r}_0 + \langle \ell \rangle \boldsymbol{n}$. If the cell optical depth is greater than unity this can be highly skewed toward the origin of the ray as opposed to the optically thin midpoint of $\langle \ell \rangle \approx \ell/2 - \tau_a\ell/3 + \mathcal{O}(\tau_a^2)$. For completeness, we provide the first-order correction to Equation (A2) as

$$\Delta p_r \approx \frac{\tau_r \varepsilon_k}{c} + \frac{k_a k\varepsilon_k}{2c}[\mu r\Delta r - \ell(r + \Delta r)$$
$$+ \frac{1}{2}r_{\min}^2 \log\left(\frac{(1 - \mu)(r + \Delta r + \ell + \mu r)}{(1 + \mu)(r + \Delta r - \ell - \mu r)}\right)]. \qquad (A5)$$

# Appendix B
## DDMC–MC Boundary Conditions

We now demonstrate the validity of our new semi-deterministic DDMC momentum coupling scheme across extreme transitions. Specifically, this refers to Equation (30) while also employing the hybrid IMC–DDMC boundary conditions described in Section 2.9.2. This is particularly important because opacity gradients often induce high radiative fluxes. Furthermore, transition regions are particularly sensitive to changes, so it is essential to accurately capture the local momentum coupling. We design a simple test to capture the relevant features of this numerical problem.

The setup is that of a one-dimensional uniform slab with a center-to-edge optical depth of $\tau = kR = 100$. Outside the central tophat region the density drops by a factor of 100, representing an extreme transition layer before the photons escape freely at a radius of $3R$. We choose to restrict the photon propagation to pure forward–backward scattering, which promotes sufficiently rapid escape for a convenient cadence of distinct curves in our demonstration. We have verified that three-dimensional transport gives similar results. At $t = 0$ we initialize $10^7$ photon packets as a pulse source uniformly distributed throughout the central region, in an effort to enhance the flux near the transition layer. Figure 17 shows the radiation energy and force density profiles over several representative times, $t = \{0.1, 0.2, 0.4, 0.8, 1.2, 1.6\} t_{\mathrm{diff}}$, where in this case the diffusion time is $t_{\mathrm{diff}} = kR^2/2c$. The course-grained DDMC results with central cell resolutions of $\Delta\tau = 20$ are in excellent agreement with the MC reference solutions with central cell resolutions of $\Delta\tau = 1$. The slight discrepancy at early times is due to time averaging effects related to the lagged path-based estimators and the limitation of ignoring the flux time-derivative term in the diffusion closure relation (see Equation (4)). We partially mitigate this by employing a slightly shorter time step ($\Delta t = 0.05 t_{\mathrm{diff}}$) than the near-equilibrium times of interest ($t \gtrsim 0.1 t_{\mathrm{diff}}$). We conclude that DDMC transport in this regime is highly accurate and significantly more efficient than traditional MCRT.
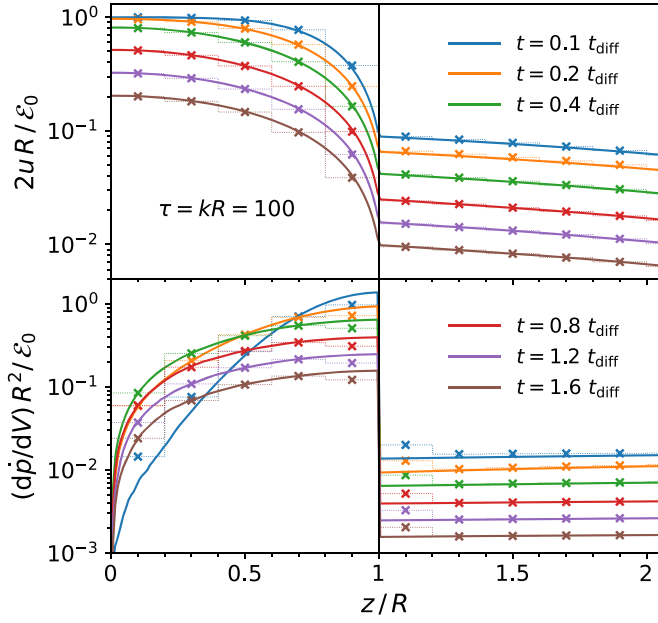
**Figure 17.** Radiation energy density $u(z)$ and force density $d\dot{p}/dV$ for a tophat configuration with center-to-edge optical depth of $\tau = 100$ but a $100\times$ reduced density outside. This test demonstrates the validity of the semi-deterministic DDMC momentum coupling scheme across extreme DDMC–MC transitions. The curves illustrate the evolution at times of $t = \{0.1, 0.2, 0.4, 0.8, 1.2, 1.6\}\, t_{\rm diff}$, where the diffusion time is $t_{\rm diff} = kR^2/2c$. The histograms employ the DDMC method a with resolution of $\Delta\tau = 20$, while the curves are the reference solutions with $\Delta\tau = 1$.

## Appendix C
## Neighbor Momentum Conservation

We now validate the efficacy of the neighbor momentum coupling scheme to conserve momentum compared to the cell-integrated method. We emphasize that problems arise only when individual cells are optically thick but that such scenarios are of common occurrence even in hydrodynamics simulations with state-of-the-art resolution. We refer to Hopkins & Grudić (2019) for a detailed discussion regarding non-scattering radiation pressure and to Section 2.7 above for our proposed general MCRT solution. The setup we explore is the one-dimensional slab version of the test from Section 3.1. The simulated domain is large enough so essentially no photons escape and although the system is scale free for concreteness we choose a characteristic optical depth of $\tau = kR = 500$. We choose to restrict the photon propagation to pure forward–backward scattering, which avoids transverse geometric vector cancellation and simplifies the derivations for reference solutions. Specifically, the evolution of the radiation energy density is governed by a diffusion equation $\partial u/\partial t = (c/k)\partial^2 u/\partial z^2$ and the solution given an initial point source impulse of energy $\mathcal{E}_0$ is $\tilde{u} = e^{-\tilde{z}^2/2\tilde{t}}/(2\pi\tilde{t})^{1/2}$, where we have rescaled into dimensionless units with radius $\tilde{r} = r/R$, time $\tilde{t} = t/t_{\rm diff}$ with diffusion time $t_{\rm diff} = kR^2/2c$, and energy density $\tilde{u} = u\,R/\mathcal{E}_0$. Therefore, the resulting force density is $d\dot{p}/dV = kF/c = -\nabla u = (\mathcal{E}_0/R^2)\tilde{z}\tilde{u}/\tilde{t}$. The total instantaneous force integrated over all space is $\dot{p}_{\rm tot} = 2\int_0^\infty (d\dot{p}/dz)dz = (\mathcal{E}_0/R)(2/\pi\tilde{t})^{1/2}$, such that the cumulative momentum up to a given point in time is $p_{\rm tot} = \int_0^t \dot{p}_{\rm tot}\, dt' = (\tau\mathcal{E}_0/c)(2\tilde{t}/\pi)^{1/2}$. We note that similar expressions may also be derived for spherical geometry.

We demonstrate the success of the three momentum implementations illustrated in Figure 1: (i) microphysical
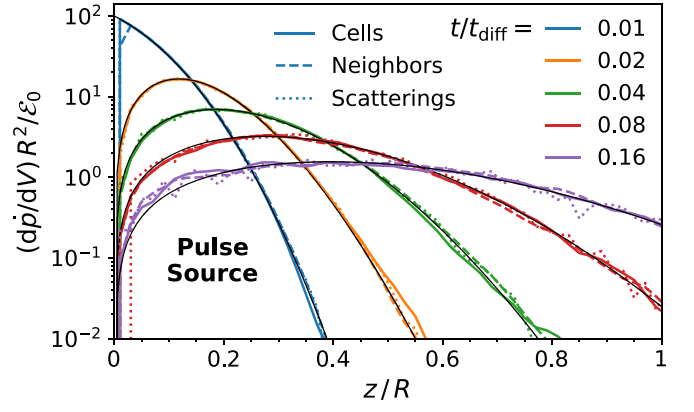


**Figure 18.** Radiation force density $d\dot{p}/dV$ for a pulse source diffusing in a one-dimensional uniform medium over several doubling times, $t = \{1, 2, 4, 8, 16\}\times 10^{-2}\, t_{\rm diff}$, where the diffusion time is $t_{\rm diff} = kR^2/2c$. Momentum schemes based on scattering events, path depositions, and neighbor integration all agree when the radiation field is well resolved. For direct comparison the analytic solution derived in Appendix C is averaged over the lagged time step.
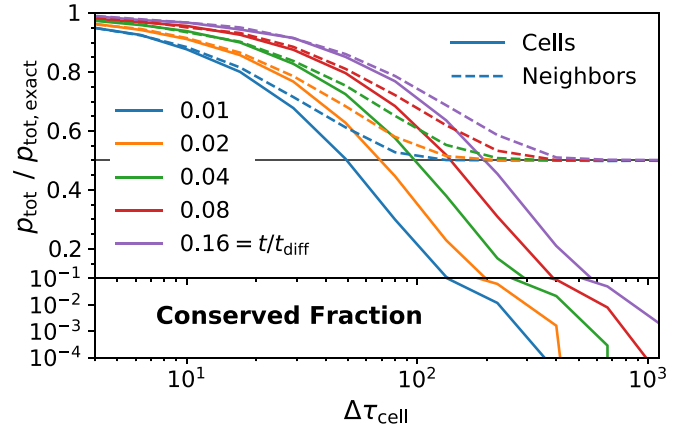


**Figure 19.** Momentum conservation as a function of cell optical depth resolution based on integrating the radiation force density over space and cumulative time and dividing by the exact analytic value of $p_{\rm tot} = (\tau\mathcal{E}_0/c)(2\tilde{t}/\pi)^{1/2}$. While standard volume averaged schemes can result in significant losses the neighbor method is guaranteed to conserve momentum within a factor of two.

scattering based exchange as described in Tsang & Milosavljević (2015), (ii) volume integration of path depositions, and (iii) neighbor-based coupling accounting for physical self-cancellation and outwardly oriented propagation. Figure 18 shows the radiation force density profile over several doubling times, $t = \{1, 2, 4, 8, 16\}\times 10^{-2}\, t_{\rm diff}$, employing $10^5$ photon packets so statistical variations are apparent. The simulation provides excellent agreement with the exact analytical solution, although we note that path-based estimators represent averages over discrete time steps and cell volumes so we plot the analytic solution including the appropriate time lag and bin integration. Finally, in Figure 19 we show the ratio of simulated to exact cumulative momenta, which demonstrates the failure of cell-integrated methods (of all varieties) to conserve momentum within optically thick cells. We place the point source at a cell center within the uniform grid and run simulations with varying resolutions. The main losses occur in the source cell at early times but persist to a noticeable degree until the diffusion is well resolved, i.e., $t \gg k\Delta z^2/c$. In contrast, the neighbor momentum scheme conserves at least half of the momentum even when the radiation field is confined within a single cell. We note that the

factor of two arises from systematic over cancellation, which can potentially be avoided by increasing the momentum imparted analogous to a closure relation. However, we prefer not to introduce this additional factor, which could erroneously apply too much momentum in other settings. By splitting the momentum between host and neighbor cells we successfully overcome the order of magnitude losses inherent to standard MCRT momentum coupling at poor optical depth resolutions.

## Appendix D
## Levitation Control Experiments

In this section we briefly explore the impact of various algorithm choices in the context of the levitation problem. We re-run the low-resolution setup under the following scenarios: (i) the luminosity exponent is 1, which samples equal weight photons in an unbiased fashion, (ii) with merging and splitting turned off, and (iii) with the neighbor-based momentum scheme activated. We note that these are the only changes to the simulation setup. The results indicate that luminosity boosting has a minor impact on the late time behavior after the development of large density fluctuations in the turbulent gas.
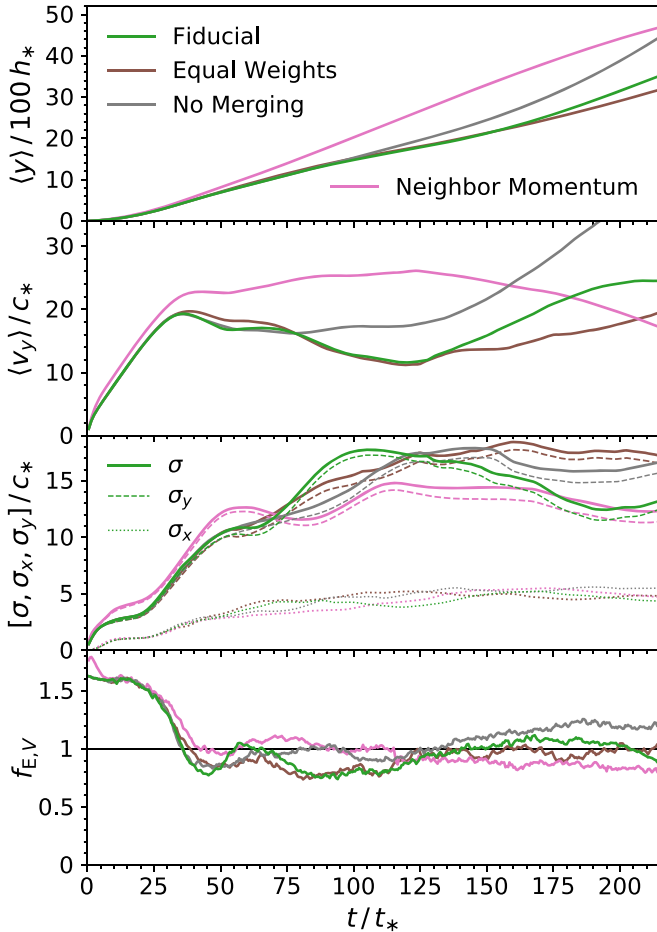


**Figure 20.** Evolution of the global mass-weighted height $\langle y \rangle$ (top), vertical velocity $\langle v_y \rangle$ (second), velocity dispersion $\sigma$ (third), and volume-weighted Eddington ratio $f_{E,V}$ (bottom). We compare the fiducial MCRT model with luminosity boosting, merging, and splitting as employed in Section 4 above (green curves) to identical simulations but with equal photon weights (brown curves), no merging or splitting (gray curves), and employing the neighbor momentum scheme (pink curves), all at lower resolution ($1\,h_*^2$). Although they are all within the expected agreement, this demonstrates that implementation details can also be important.

On the other hand, the effect of merging and splitting is more noticeable as these optimizations introduce artificial momentum and energy losses in the radiation field. Finally, the neighbor-based momentum method results in a significant boost in upward lift and gas expulsion, similar to the higher resolution fiducial run. We interpret this as an indication that the RHD physics is not converged at these low resolutions. We summarize these findings in Figure 20.

## ORCID iDs

Aaron Smith https://orcid.org/0000-0002-2838-9033
Rahul Kannan https://orcid.org/0000-0001-6092-2187
Benny T.-H. Tsang https://orcid.org/0000-0002-6543-2993
Mark Vogelsberger https://orcid.org/0000-0001-8593-7692
Rüdiger Pakmor https://orcid.org/0000-0003-3308-2420

## References

Abdikamalov, E., Burrows, A., Ott, C. D., et al. 2012, ApJ, 755, 111
Abel, T., Norman, M. L., & Madau, P. 1999, ApJ, 523, 66
Brooks, E. D. I. 1989, JCoPh, 83, 433
Brooks, E. D. I., & Fleck, J. A. J. 1986, JCoPh, 67, 59
Camps, P., & Baes, M. 2015, A&C, 9, 20
Camps, P., & Baes, M. 2018, ApJ, 861, 80
Camps, P., Baes, M., & Saftly, W. 2013, A&A, 560, A35
Castor, J. I. 2004, Radiation Hydrodynamics (Cambridge: Cambridge Univ. Press)
Chandrasekhar, S. 1961, Hydrodynamic and Hydromagnetic Stability (Oxford: Clarendon)
Cleveland, M. A., & Gentile, N. 2015, JCoPh, 291, 1
Cleveland, M. A., & Wollaber, A. B. 2018, JCoPh, 359, 20
Commerçon, B., Teyssier, R., Audit, E., Hennebelle, P., & Chabrier, G. 2011, A&A, 529, A35
Davis, S. W., Jiang, Y.-F., Stone, J. M., & Murray, N. 2014, ApJ, 796, 107
Davis, S. W., Stone, J. M., & Jiang, Y.-F. 2012, ApJS, 199, 9
Densmore, J. D., Thompson, K. G., & Urbatsch, T. J. 2012, JCoPh, 231, 6924
Densmore, J. D., Urbatsch, T. J., Evans, T. M., & Buksas, M. W. 2007, JCoPh, 222, 485
Ensman, L. 1994, ApJ, 424, 275
Fleck, J. A. J., & Canfield, E. H. 1984, JCoPh, 54, 508
Fleck, J. A. J., & Cummings, J. D. 1971, JCoPh, 8, 313
Gelman, A., & Rubin, D. B. 1992, StaSc, 7, 457
Gentile, N. A. 2001, JCoPh, 172, 543
Gnedin, N. Y., & Abel, T. 2001, NewA, 6, 437
González, M., Audit, E., & Huynh, P. 2007, A&A, 464, 429
Habetler, G. J., & Matkowsky, B. J. 1975, JMP, 16, 846
Harries, T. J. 2015, MNRAS, 448, 3156
Harries, T. J., Douglas, T. A., & Ali, A. 2017, MNRAS, 471, 4111
Harries, T. J., Haworth, T. J., Acreman, D., Ali, A., & Douglas, T. 2019, A&C, 27, 63
Hayes, J. C., & Norman, M. L. 2003, ApJS, 147, 197
Hopkins, P. F., & Grudić, M. Y. 2019, MNRAS, 483, 4187
Hopkins, P. F., Quataert, E., & Murray, N. 2011, MNRAS, 417, 950
Hubber, D. A., Ercolano, B., & Dale, J. 2016, MNRAS, 456, 756
Jaura, O., Glover, S. C. O., Klessen, R. S., & Paardekooper, J. P. 2018, MNRAS, 475, 2822
Jiang, Y.-F., Stone, J. M., & Davis, S. W. 2012, ApJS, 199, 14
Jiang, Y.-F., Stone, J. M., & Davis, S. W. 2014, ApJS, 213, 7
Kannan, R., Springel, V., Pakmor, R., Marinacci, F., & Vogelsberger, M. 2016, MNRAS, 458, 410
Kannan, R., Vogelsberger, M., Marinacci, F., et al. 2019, MNRAS, 485, 117
Keady, K. P., & Cleveland, M. A. 2017, JCoPh, 328, 160
Kimm, T., Haehnelt, M., Blaizot, J., et al. 2018, MNRAS, 475, 4617
Krumholz, M. R., Klein, R. I., McKee, C. F., & Bolstad, J. 2007, ApJ, 667, 626
Krumholz, M. R., & Thompson, T. A. 2012, ApJ, 760, 155
Levermore, C. D. 1984, JQSRT, 31, 149
Levermore, C. D., & Pomraning, G. C. 1981, ApJ, 248, 321
Long, A. R., Gentile, N. A., & Palmer, T. S. 2014, JCoPh, 277, 228
Lowrie, R. B., & Edwards, J. D. 2008, ShWav, 18, 129
Lucy, L. B. 1999, A&A, 344, 282
Michel-Dansac, L., Blaizot, J., Garel, T., et al. 2020, A&A, 635, A154

Mihalas, D., & Mihalas, B. W. 1984, Foundations of Radiation Hydrodynamics (New York: Oxford University Press)

Min, M., Dullemond, C. P., Dominik, C., de Koter, A., & Hovenier, J. W. 2009, A&A, 497, 155

Nayakshin, S., Cha, S.-H., & Hobbs, A. 2009, MNRAS, 397, 1314

Noebauer, U. M., & Sim, S. A. 2015, MNRAS, 453, 3120

Noebauer, U. M., & Sim, S. A. 2019, LRCA, 5, 1

Noebauer, U. M., Sim, S. A., Kromer, M., Röpke, F. K., & Hillebrandt, W. 2012, MNRAS, 425, 1430

Pakmor, R., Bauer, A., & Springel, V. 2011, MNRAS, 418, 1392

Pakmor, R., Springel, V., Bauer, A., et al. 2016, MNRAS, 455, 1134

Pawlik, A. H., & Schaye, J. 2008, MNRAS, 389, 651

Petkova, M., & Springel, V. 2011, MNRAS, 415, 3731

Pomraning, G. C. 1973, The Equations of Radiation Hydrodynamics (Oxford: Pergamon Press)

Robitaille, T. P. 2010, A&A, 520, A70

Rosdahl, J., Blaizot, J., Aubert, D., Stranex, T., & Teyssier, R. 2013, MNRAS, 436, 2188

Rosdahl, J., & Teyssier, R. 2015, MNRAS, 449, 4380

Rosen, A. L., Krumholz, M. R., Oishi, J. S., Lee, A. T., & Klein, R. I. 2017, JCoPh, 330, 924

Roth, N., & Kasen, D. 2015, ApJS, 217, 9

Ryan, B. R., & Dolence, J. C. 2020, ApJ, 891, 118

Ryan, B. R., Dolence, J. C., & Gammie, C. F. 2015, ApJ, 807, 31

Semenov, D., Henning, T., Helling, C., Ilgner, M., & Sedlmayr, E. 2003, A&A, 410, 611

Skinner, M. A., & Ostriker, E. C. 2013, ApJS, 206, 21

Smith, A., Becerra, F., Bromm, V., & Hernquist, L. 2017a, MNRAS, 472, 205

Smith, A., Bromm, V., & Loeb, A. 2017b, MNRAS, 464, 2963

Smith, A., Ma, X., Bromm, V., et al. 2019, MNRAS, 484, 39

Smith, A., Safranek-Shrader, C., Bromm, V., & Milosavljević, M. 2015, MNRAS, 449, 4336

Smith, A., Tsang, B. T. H., Bromm, V., & Milosavljević, M. 2018, MNRAS, 479, 2065

Springel, V. 2010, MNRAS, 401, 791

Stone, J. M., Mihalas, D., & Norman, M. L. 1992, ApJS, 80, 819

Thompson, T. A., Quataert, E., & Murray, N. 2005, ApJ, 630, 167

Trac, H., & Cen, R. 2007, ApJ, 671, 1

Tsang, B. T. H., & Milosavljević, M. 2015, MNRAS, 453, 1108

Tsang, B. T. H., & Milosavljević, M. 2018, MNRAS, 478, 4142

Turner, N. J., & Stone, J. M. 2001, ApJS, 135, 95

Vandenbroucke, B., & Camps, P. 2020, A&A, 641, A66

Vandenbroucke, B., & Wood, K. 2018, A&C, 23, 40

Weinberger, R., Springel, V., & Pakmor, R. 2020, ApJS, 248, 32

Whalen, D., & Norman, M. L. 2006, ApJS, 162, 281

Whitehouse, S. C., & Bate, M. R. 2006, MNRAS, 367, 32

Wise, J. H., & Abel, T. 2011, MNRAS, 414, 3458

Wollaber, A. B. 2016, J. Computational and Theoretical Transport, 45, 1

Wollaeger, R. T., & van Rossum, D. R. 2014, ApJS, 214, 28

Wollaeger, R. T., van Rossum, D. R., Graziani, C., et al. 2013, ApJS, 209, 36

Zel'dovich, Y. B., & Raizer, Y. P. 1967, Physics of Shock Waves and High-temperature Hydrodynamic Phenomena (New York: Academic Press)