ISSN 2959-6157

Cosmological Radiation Simulation Analysis: Taking Arepo-RT as an Example

Yifan He[1, *]

[1]Chongqing Liangjiang New Area High School Affiliated to Southwest University, Chongqing, China
*30178074@alu.cqu.edu.cn

Abstract:

With the fast-growing progress in computational physics, astrophysical simulations are made possible and continually upgraded with novel physical models. Contemporarily, many modern simulations feature the capability of cosmological radiation. In order to collate the recent improvements in this field, the numerical solver Arepo-RT is discussed as an important example. This study reorganizes and interprets the main physical models and numerical techniques used in Arepo-RT and similar solvers, with some details simplified to suit common readers. This paper describes simply and briefly the finite volume method, the radiative transfer model, the radiation source model, the Godunov-type schemes and the time integration, all of which are frequently used in modern radiation simulations. Apart from the realizations of the solvers, the results and limitations are also outlined in this paper, including the justification of reduced speed of light approximation, the results compared with JWST, and the issue of missing photons. These results shed light on guiding further improvements for Arepo-RT.

Keywords: Astrophysical simulation; cosmological radiation; Arepo-RT; THESAN.

1. Introduction

In the past decades, the development in computer science has triggered a series of renovations in astrophysics, where numerical simulations have become an essential part. Early numerical models mostly focused on the predominant dark-matter interactions, considering only gravitational fields. Later, with the advent of Gadget and Arepo [1], cosmic gas and magnetohydrodynamics, as relatively complicated physics, were introduced to the simulations. These leaders have been boosting the understanding of cosmological evolution in a wide range of redshifts. However, about a higher-redshifted universe, including the CMB, the Epoch of Reionization and the first galaxies, there is still a lack of knowledge. To perceive these crucial stages of universe evolution requires a comprehensive model for the radiation transport (radiation hydrodynamics) and cosmic dust. Unfortunately, being a six-dimensional non-linear problem, the radiative transfer (RT) equations in the models have no general solutions and are also challenging for numerical methodologies [2, 3].

Recently, a breakthrough has been made by taking the moment-based solutions. Two examples are Ramses-RT [4] and Arepo-RT [5], two solvers for RT problems. The latter one is an extension of Arepo and is currently a major part of the THESAN project. The THESAN project features an all-around model for gravity, radiation hydrodynamics, cosmic dust, etc., despite some still uncertain parameters [6]. The results of THESAN are generally concordant with observations and have thus proved the utility of Arepo-RT. Moreover, the launch of James Webb Space Telescope (JWST) has enabled direct observations of the early universe. According to some results [7], the power of numerical simulations is further justified.

Apero-RT incorporates typical physical models and widely used numerical schemes. Hence, if one wishes to comprehend the principles used in these simulations, Apero-RT is an example worth reviewing. The paper is organized as follows: some basic concepts are discussed in Sec. 2. The physical modeling and approximations are discussed in Sec. 3. Section 4 is dedicated to the numerical schemes used in, e.g., Arepo, Arepo-RT and Ramses-RT. In Secs. 5 and 6, the validity of those methods is examined, including the advantages and drawbacks. Finally, the whole paper is concluded in Sec. 7.

2. Descriptions

In modern physics, the cosmological principle claims that the universe is homogeneous and isotropic at large scales, which means that the large-scale property of the universe is the same for all observers at any location and

any direction. According to the Friedmann Equation, the expansion of the universe can be described with the scale factor, $a(t)$, which is a function of time. Many solvers, including Arepo [1, 8], based on the spatial positions on co-moving coordinates, namely, position vector, velocity and density defined as:

(Tex translation failed) \[\]where the unscripted quantities denote those in proper coordinates. The co-moving coordinates extract the scale factor from the proper coordinates, so that the change in these quantities will no longer be influenced by the expansion of the universe. The co-moving coordinates can better describe and compare the evolution of the local sections of the universe through time. The physics taken into account usually includes the ΛCDM model, the gravitational fields and the magnetohydrodynamics. Later, photon injection and radiation transport are added to Arepo-RT. These physical equations are rewritten into the form of co-moving quantities. The numerical methods for radiation are analogous to those used for hydrodynamics. Before Arepo, the mainstream methods were smoothed particle hydrodynamics (SPH) and adaptive mesh refinement (AMR). Springel suggested a new method called “moving, unconstructed mesh”, which is a kind of finite volume approaches [1]. This method was later inherited by Arepo-RT [5].

Arepo-RT, like Arepo, discretizes the 3D space into cells to simplify calculations. The cells are constructed by Voronoi tessellation, forming a mesh [1, 8]. Voronoi tessellation produces cells from mesh-generating points. For each mesh-generating point, its cell is the region that is closer to it than to any other point. The connections between neighboring mesh-generating points form an ensemble of tetrahedra (i.e., Delaunay triangulation). To construct a mesh, the mesh-generating points are produced in the procedures: (i) start with four points (corners of a tetrahedron); (ii) insert a mesh-generating point into the tetrahedron; (iii) split the tetrahedron into four smaller ones and (iv) repeat for each tetrahedron. The interfaces between cells are then constructed according to the mesh-generating points. In the simulation, each mesh-generating point can move, with its velocity and acceleration affected by the overall motion of gas (or something else) in its cell. The motion is updated during each time step. However, the mesh and the gas move independently. Additionally, the size, position, and volume of a cell can change, but each cell should keep its size and mass within a desired interval, which are controlled by several methods: velocity corrections, (de-) refinement and regularization. In practical simulations, the algorithm uses a hybrid of particle-mesh and oct-tree, or a particle-particle particle-mesh, for diverse types of interaction.

3. Physical Models

3.1 Radiation Transport

To describe the radiation of a certain gas/plasma, one can write down the Euler Continuity Equation:

(Tex translation failed) where the specific intensity I_ν in the direction n is described with a six-dimensional non-linear equation, which is quite challenging. Some algorithms including Arepo-RT and RAMSES-RT simplify the equation by taking the moments of each term. By taking the zeroth moment, one can obtain [4, 5]:

(Tex translation failed) which connects energy density $\mathcal{E}_r$ with flux Φ_r, as can be interpreted as changeindensity+divergenceofflux=emission¨Cabsorption, where the emission (S) is due to the emissivity of the plasma, while the absorption ($\kappa_\varepsilon \rho c \mathcal{E}_r$) is due to the opacity of the gas. (c is a reduced approximation for the speed of light) This equation is intuitive that the energy is conserved, just like the mass is conserved in a fluid, unless there is a source or a sink. Next, by taking the first moment,

(Tex translation failed) connects flux with radiation pressure tensor $\mathbb{P}_r$. These equations are similar to those hydrodynamic equations used in Arepo, so one can continue to use the existing numerical methods in which Arepo deals with hydrodynamics.

3.2 M1 Closure

However, this derivation has introduced new variables. Now there are three variables ($\mathcal{E}_r$, Φ_r, $\mathbb{P}_r$) but only two equations. Besides, adding higher-order moment equations will introduce even more variables. Therefore, to make the system solvable, it is necessary to assume an extra connection between these variables other than the Euler equation. A plausible mathematical approach for this purpose is the M1 Closure. Analogous to the hydrodynamic closure, $p = \rho(\gamma - 1)e$ (with pressure p, mass density ρ, energy density e and adiabatic index γ), Levermore et al. suggested

(Tex translation failed) where $\mathbb{D}$ and χ are Eddington tensor and factor, respectively; they describe the diffusion and the propagation speed of the radiation field. The relation is the consequence of a Lorentz-boosted Planckian approximation, which assumes that the radiation diffusion reaches the asymptotic limit (where photons propagate in random walk), where $\mathbb{D} \approx \mathbb{I}/3$. It is also suitable for simulating the radiation when $\mathbb{D} \approx nn$. However, between

these two limits or in the presence of multiple sources, this approximation may face some challenges [4].

With the approximation valid, the M1 closure results in a PDE system with hyperbolic conservation. Consequently, one can approximate the solution by using an *operator split scheme*. As discussed by Rosdahl & Teyssier [4], the scheme splits the equations into two approximately independent processes: the update of the state vector $(\text{density}, \text{flux})$ and the update of the source term $(\text{emission} - \text{absorption})$. In free transport of photons, the $(\text{emission} - \text{absorption})$ term equals zero. Notably, photons are relativistic, and therefore the model must be discreet when dealing with the Galilean Invariance. On the other hand, the source term is dependent on the conditions: different types of sources may vary in frequency ranges, intensities, opacities, etc. They are thus applied with different approximations. To model these scenarios, one can split the emitted photons into several frequency bins. For each frequency bin, one can employ the empirical values of ion species, ionization cross-section, photon energies etc. [6].

3.3 Local Thermodynamic Equilibrium

Since not all the details of emission and absorption are known, one can assume a local thermodynamic equilibrium (LTE), to simplify the model of radiation source's, especially when radiation-matter interactions are significant [2]. LTE differs from complete thermodynamic equilibrium, since the former does not require an exact Planckian radiation field but rather assumes it. If a system is in or close to LTE, the emission of radiation can be simply modelled as blackbody radiation. This is valid especially for the infrared radiation from cosmic dust, since dust is so dense that its thermodynamic equilibria can hardly be affected by the radiation field. However, this approximation becomes inaccurate for some ultraviolet (UV) sources, e.g., hydrogen/helium recombination.

3.4 Cosmic Dust and Infrared (IR) Radiation

Cosmic dust is also included in Arepo-RT. In this model three main processes are considered: the generation of new cosmic dust, its interstellar growth, and its destruction. Hence, a semi-empirical model is developed by McKinnon et al. [9]. The dust usually originates from the evolution of Asymptotic Giant Branch (AGB) stars and supernovae (SN). The mass of dust generated from stellar evolution is proportional to the change in mass of a component in the star. For AGB stars, the ratio between masses of carbon and oxygen (C/O) determines which species of dust can be generated. After the generation, the dust will grow, when gas-phase elements collide with dust and form more dust, satisfying:

$$\left(\frac{\mathrm{d}M_{i,\mathrm{dust}}}{\mathrm{d}t}\right)_{\mathrm{growth}} = \left(1 - \frac{M_{i,\mathrm{dust}}}{M_{i,\mathrm{metal}}}\right)\left(\frac{M_{i,\mathrm{dust}}}{\tau_{\mathrm{g}}}\right) \tag{6}$$

where τ_{g} is a characteristic growth timescale affected by density and temperature.

Dust grains could also be destroyed when they go through SN shocks, sputtering, grain-grain collisions, etc. The overall destruction rate is

$$\left(\frac{\mathrm{d}M_{i,\mathrm{dust}}}{\mathrm{d}t}\right)_{\mathrm{destruction}} = -\frac{M_{i,\mathrm{dust}}}{\tau_{\mathrm{d}}} \tag{7}$$

where τ_{d} is a characteristic destruction timescale, determined by the shock-destruction efficiency (≈ 0.3), the local Type II SN rate, and the mass of gas that is shocked to $\geqslant 100\mathrm{km/s}$. Similar to the ionization fraction, the amount of dust is treated as a passive scalar, whose motion tracked within individual gas cells. Dust is assumed to be advected between gas cells in each time step where the net growth rate is calculated before updating the local dust mass [4]. In Arepo-RT and similar solvers, gas (coupled with dust) emits as a black body, so the emission term for infrared (IR) should be

$$S_{\mathrm{IR}} = \kappa_{\mathrm{Planck}} \rho c a T^4 \propto \mathrm{Temperature}^4 \tag{8}$$

provided the LTE approximation is valid here. Notably, those non-IR photons are not included in the model.

By assuming the opacities κ_{Planck} and κ_{ε} are equal, the resultant continuity equation can be written as

$$\frac{\partial \mathcal{E}_{\mathrm{IR}}}{\partial t} + \nabla \cdot \boldsymbol{\Phi}_{\mathrm{IR}} = \kappa_{\mathrm{Planck}} \rho \left(c a T^4 - \tilde{c} \mathcal{E}_{IR}\right) \tag{9}$$

To avoid the 4th power of temperature, one can apply a linear symmetrical approximation, and finally obtain [4]:

$$\Delta \mathcal{E}_{\mathrm{IR}} = \frac{c a T^4 - \tilde{c} \mathcal{E}_{IR}}{\left(\kappa_{\mathrm{Planck}} \rho \Delta t\right)^{-1} + \tilde{c} + 4 c a T^3 C_V^{-1}} \tag{10}$$

3.5 Photon Injection

LTE cannot be applied to UV radiation; thus, detailed physics should be modeled. In this case, photons participate in multiple interactions, including thermochemistry, heating and cooling, injection form stars and AGNs. Thermochemistry considers the change in ionization states of different species, including $\mathrm{H_I}$, $\mathrm{H_{II}}$, $\mathrm{He_I}$, $\mathrm{He_{II}}$ and $\mathrm{He_{III}}$. The species are reflected by the ionization fractions, which are passive scalar coupled with the gas. The photons are also divided into several frequency bins. For example, the dividing pattern used in THESAN [6] is $[13.6, 24.6, 54.4, \infty)$ eV. The source term of photons in bin i reacting with ion species j is [10]

$$S_{ij} = -\tilde{c} n_j n_\gamma^i \bar{\sigma}_{ij} + s_{ij} \tag{11}$$

where n_j and n_γ^i are the number densities of species j and photons in bin i; $\bar{\sigma}_{ij}$ is the average cross-section of the ion species, which indicates the probability of ionization; s_{ij} is the source due to recombination, which is approximated to 0 in optically thick media. The equations are closed with the conservation of the ion number densities, analogous to M1 Closure in free transport.

The heating and cooling are contributed by gas, metal, photoheating and dust-gas-radiation field reactions, according to Kannan et al. [11]. As discussed by Kannan et al. [6], the metal cooling is assumed to be an ionization equilibrium, with a spatially uniform UV background. Stars and active galactic nuclei (AGNs) also emit radiation, which is described by the BPASS spectra. The radiation from a star is influence by its age and metallicity. While for AGNs, the radiation is proportional to the mass accretion rate. Besides, the radiation from black holes is neglected [6, 12].

4. Numerical Techniques

4.1 Moving Mesh

Arepo-RT defines two state vectors:

$$\boldsymbol{U} = \begin{pmatrix} \mathcal{E}_r \\ \boldsymbol{\Phi}_r \end{pmatrix}, \boldsymbol{F} = \begin{pmatrix} \boldsymbol{\Phi}_r \\ \mathbb{P}_r \end{pmatrix} \tag{12}$$

which is analogous to the hydrodynamic definition in Arepo [1]:

$$U_{\text{hydro}} = \begin{pmatrix} \rho \\ \rho v \\ \rho e \end{pmatrix}, F_{\text{hydro}} = \begin{pmatrix} \rho v \\ \rho v v^T + P \\ \rho e v + P v \end{pmatrix} \tag{13}$$

For the free transport case, where $emission - absorption = 0$, the continuity equation can be rewritten as

$$\frac{\partial \boldsymbol{U}}{\partial t} + \nabla \cdot \boldsymbol{F} = 0. \tag{14}$$

One can employ the finite-volume method, describing the fluid's state by the cell averages of the state vectors for these cells [1]. For cell i, its state vector U:

$$\boldsymbol{Q}_i = \int_{V_i} \boldsymbol{U} \mathrm{d}V, \tag{15}$$

assuming that the state vector does not vary within a volume element, and that the position corresponding to the average state vector is at the cell's centroid. The changing rate of Q_i will be used in next steps. It is calculated using Gauss's divergence theorem which converts the volume integral into a surface integral,

$$\frac{\mathrm{d}\boldsymbol{Q}_i}{\mathrm{d}t} = -\int_{S_i} \left(\boldsymbol{F} - \boldsymbol{U}\boldsymbol{w}^{\mathrm{T}} \right) \boldsymbol{n} \mathrm{d}S \tag{16}$$

where w is the velocity of each point of the cell boundary. The changing rate (flux exchange) is calculated for each of the cell interfaces adjacent to cell i, then the total changing rate is their sum. Due to the failure of Galilean Invariance, Arepo-RT has to calculate the flux exchanges in the frame that is moving with the cell interface. Later, it returns to the lab frame to apply these values to e.g. the source terms [5]. Here, w indicates the relative speed between the two frames, and thus has to be taken into account in the calculation. Here, w is geometrically dependent on the velocities of the adjacent cells across the interface. As discussed by Springel [1], primarily, the velocity of the center of the interface between cells L and R is

$$\boldsymbol{w} \approx \frac{\boldsymbol{v}_{\mathrm{R}} + \boldsymbol{v}_{\mathrm{L}}}{2} \tag{17}$$

where v_R and v_L are the velocities of the mesh-generating points of the cells, respectively. These velocities will also affect the cell volume V_i.

4.2 Riemann Problem

A Riemann Problem is a set of partial differential equations expressing a certain conservation law with piecewise initial values. Solving the radiation transfer is in fact solving a Riemann Problem. The M1 closure ensures the system to have a hyperbolic conservation, while the finite volume method causes the state vectors to change suddenly at the cell interfaces. A classic solver to this problem is suggested by Godunov, called the Godunov's Scheme, which uses the intercell fluxes to calculate U at every cell interface, and linearly extrapolates for $U_{L,R}$ at the centroids of its neighboring cells (left and right). One can carry out the extrapolation in the interface frame, with the coordinates set normal to the interface. To make it easier to understand, here is a one-dimensional (normal to the interface) example [10]:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{F_{i+1/2}^n - F_{i-1/2}^n}{\Delta x} = 0 \tag{18}$$

where n means the nth time step; $i-1/2$ and $i+1/2$ denote the cell interfaces of the cell i.

The continuity can also be expressed with a Jacobian matrix $\partial F / \partial U$, whose eigenvalues indicate the signal speed (or propagations speed, wave speed) of the system. For radiation transfer problems, they are limited within the speed of light. The Harten-Lax-van Leer (HLL) scheme is an approximated flux function using such property, specifically, the intercell flux

$$F^n_{i+1/2} = \frac{\lambda^+ F^n_i - \lambda^- F^n_{i+1} + \lambda^+ \lambda^- (U^n_{i+1} - U^n_i)}{\lambda^+ - \lambda^-} \quad (19)$$

where λ^+ and λ^- are maximum and minimum signal speeds, respectively. The method is proposed by van Leer et al. Moreover, it can be shown that the (λ^+, λ^-) of a static system and those of a moving system (with a normal speed w) have a linear relationship in terms of w. This relationship enables a simple transformation between the lab frame and interface frame [5].

If the λ^+ and λ^- are set to the speed of light ($\pm c$), the function can be simplified to:

$$F^n_{i+1/2} = \frac{F^n_i + F^n_{i+1}}{2} + \frac{(U^n_{i+1} - U^n_i)c}{2}. \quad (20)$$

This is the global Lax–Friedrich (GLF) function, proposed by Rusanov. However, GLF is more diffusive than HLL, and therefore better simulates, e.g., the beams and shadows. By contrast, HLL has an inherent directionality, due to the choice of the axis, resulting in asymmetric simulations for originally isotropic sources [13].

4.3 Extrapolation and Time Integration

As Godunov's scheme suggests, tracing the change in U involves a temporal extrapolation $\partial U / \partial t$ as well as a spatial extrapolation $\partial U / \partial r$. Both extrapolations are made independently. For temporal extrapolation, the unit of the integration is the length of the timestep. Godunov's scheme assumes that the flux is unchanged within a whole timestep (i.e. remains in its initial value) and uses the flux to calculate the change in the state vector. Then the fluxes are updated, which will be used in the next timestep. This method obtains a first-order convergence. However, for better accuracy, Arepo-RT uses a hybrid of Heun's method and MUSCL [5, 14], which use the average between the initial and final fluxes, differing from Godunov's using the initial flux only. Analogous to trapezium integration, this new scheme obtains a second order accuracy.

Spatial extrapolation is to link the state vectors at the cell centroids with those at the interfaces. The unit of the integration is the displacement between the interface and the cell centroids. For each cell, the local gradient and geometry should be known. The gradient is obtained by local least square fit, that is, determining a cell's gradient by the quantities of its neighboring cells. The gradient is estimated such to fit the differences in quantities between the cell and its neighbors. Arepo-RT, similar to Arepo [1], bounds the gradient of a cell such that the intercell differences reproduced by the gradient are smaller than the maximum actual difference. The geometry is obtained by mesh regeneration based on the movements of mesh-generating points:

$$\boldsymbol{r}_{n+1} = \boldsymbol{r}_n + \boldsymbol{v}_n \Delta t \quad (21)$$

where v_n is the speed of the mesh-generating point and assumed to be constant during the nth time step Δt. To improve accuracy, two different mesh geometries are used in one time step. However, only one mesh per timestep is created, since the final mesh of the last step can be reused for the first half of the current step. In the middle of the step, a new mesh is updated, which will be reused for the next step.

4.4 Slope Limiter

One constraint of the physical system is that the propagation speed of radiation cannot exceed the speed of light. The property can be expressed by a factor (called the reduced flux) f,

$$f = \frac{|\boldsymbol{\Phi}_r|}{\tilde{c}\mathcal{E}_r} \in [0,1]. \quad (22)$$

This property is not included in most extrapolation schemes, so some solvers use slope limiting functions (i.e. slope limiters) to restrict the gradients. The limited gradients are used to construct piecewise linear functions within each cell and extrapolate for Φ_r and $\mathcal{E}_r$. Finally, the value of f is restricted. Nonetheless, some limiters may produce noises, such non-physical oscillations [1, 5]. As Kanan et al. mentioned [5], Arepo-RT makes the time and spatial extrapolations dependent on each other, by considering the gradings not only of U and F but of f. Since the gradients are bounded in extrapolation, the f can be naturally limited within the interval $[0,1]$. Therefore, this so-called *slope-limited scheme* can avoid using extra slope limiters that cause noises. Consequently, a second order accuracy is secured.

4.5 Operator Splitting

The radiative transfer equations involve the variables in cross connections. To solve the change in one variable in a timestep, another variable should also be known. However, both variables can change simultaneously but only the initial values are known. Therefore, a simple approach is to assert one variable to be unchanged during a timestep and solve the variables one by one. Similarly, Ramses-RT and Arepo-RT use an operator splitting scheme [4, 5]. The split scheme can be employed only when M1 closure is valid. The scheme involves three operations in a time-step. The first one is a half-step update ($\Delta t / 2$) of energy and flux due to the source term, where photons are emitted/absorbed into the neighboring cells. The second operation is a full-step update (Δt) that solves the RT equations, obtaining the intercell fluxes. The final one is a half-step

update $\Delta t/2$ of source terms, including the ionization states and heating/cooling processes [4, 15]. The operator splitting scheme is depicted in Fig. 1. Notably, operator splitting does not mean a dimensionally split scheme, as described in Springel [1]. Since the mesh is not Cartesian, it is impractical to decompose all operations into the same combination of independent spatial directions. Therefore, Arepo-RT uses an unsplit scheme, constructing a coordinate system for each cell face, with the axes orthogonal to the face.

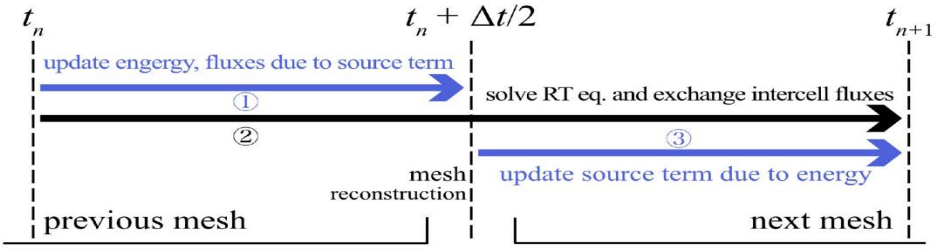

Fig. 1 A diagram describing the split scheme and mesh construction.

4.6 Timestep Constraints

A small timestep can increase the accuracy of the simulation. However, it could also be expensive. For radiative transfer (RT), the timestep of each cell is constrained by the von Neumann stability condition (in the lab frame), which is [5]:

$$\Delta t_{\mathrm{RT}} \leqslant \eta \frac{\text{cell width}}{c + \text{ speed of cell}}. \tag{23}$$

As c is a large value, the timestep is too small and too hard to compute. In most cases, a reduced speed of light approximation (RSLA), c, can be used, which will be discussed in section 5. There are also constraints due to hydrodynamics and magnetics, namely $\Delta t_{\mathrm{hydro}}$ and Δt_{mag}. If all the regimes have the same length of the timestep, the timestep should be restricted by all the constraints (i.e. the minimum of the constraints). However, RT usually needs a higher updating rate than does hydrodynamics. Therefore, if the length of the timestep can be made different for different regimes, the demands can be reached with better efficiency, and such method is called sub-cycling.

In addition, there are some special constraints for thermochemistry and IR radiation. A slight change in the photon density can lead to rapid changes in the ionization state and temperature of the gas. Similarly, IR radiation, dependent on the 4th power of temperature, may also change dramatically. Thus, a small timestep is required. Arepo-RT employs a semi-implicit approach, which first solves the number density implicitly before updating other variables. During each time step, if one of variables changes by over 10%, SUNDIALS CVODE (another algorithm) should be used instead to check it [4, 5].

4.7 Subcycling

Subcycling is to give the radiative transfer (RT) a different length of timestep from that of hydrodynamics. It is used by many solvers, so as to keep the convergence with only a small number of time steps [5]. Specifically, $\Delta t_{\mathrm{hydro}}$ is divided into a number (N_{sub}) of RT sub-steps, which are arranged in a routinized process. Before the operation, the length of the timestep is set to $\Delta t_{\mathrm{hydro}} / N_{\mathrm{sub}}$. The first operation is thermochemistry and momentum injection, followed by the construction of Active interfaces (those having at least one adjacent cell). The timestep is then reset to the $1/N_{\mathrm{sub}}$ of the minimum timestep of adjacent cells. Finally, then each cell's flux is exchanged over those interfaces and the variables are updated. Fig. 2 shows that the RT sub-steps are based on the original Arepo hydrodynamics [16].

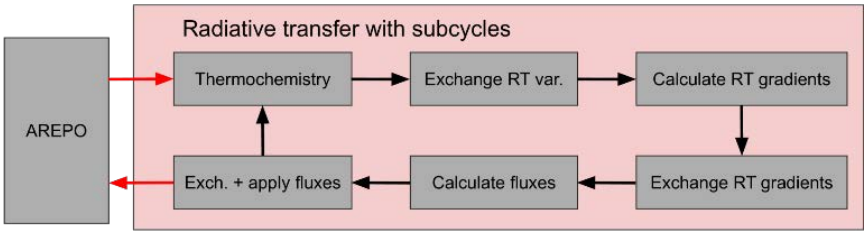

Fig. 2 The subcycling process of Arepo-RT as an extension Arepo [16].

5. Results

5.1 The Performance of HLL and GLF Solvers

Kannan et al. examined the propagation of free-streaming radiation waves in a completely absorbing, homogeneous medium with low optical depth [5]. To test the convergences of the algorithms, HLL and GLF schemes were used, along with their piecewise-constant (PC) versions (PC-HLL and PC-GLF). The latter two were to mimic the piecewise constant approximation used in obsolete solvers, so as to verify the improvements of the new linear extrapolation schemes. The tests were carried out using a regular Cartesian (0% deviation) and three irregular ones (1%, 10% and 20% deviations). As the results shown in Fig. 3, the PC-GLF and PC-HLL only obtained a first-order convergence, whereas the GLF and HLL schemes showed a faster convergence with an order of ~ 2.0.

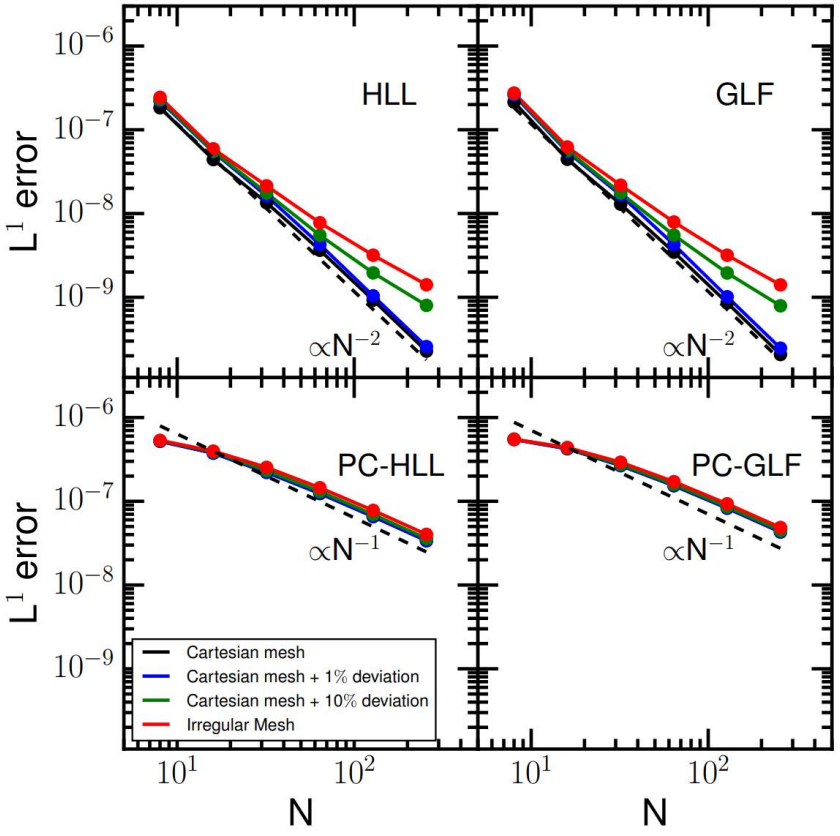

Fig. 3 The convergences of different schemes with different mesh distortions [5].

It is notable that the convergence was hardly affected by the choice of HLL or GLF but heavily affected by piecewise approximation. In other words, the difference in diffusion caused by the PC scheme is much more significant than that between HLL and GLF. Therefore, the new schemes are far less diffusive than the PC ones. However, the convergences of all schemes decreased as meshes became more irregular, which highlights the necessity for mesh regularization.

5.2 Effect of Reduced Speed of Light

To avoid a small time-step and boost computing efficiency, a reduced speed of light approximation (RSLA) may be employed. The approximation is valid as long as the reduced speed of light is still faster than any other radiation transport. According to Wu et al. [13], which used Arepo-RT to simulate the interplay between the photo-heating and the star formation rate (SFR), the RSLA has no significant impact on this simulation. In their experiment, it was shown that a $c = 0.1c$ results a 10 times smaller photoionization rate (post-reionization). To further assess the impact, Wu et al. [13] carried out simulations of $25\mathrm{cMpc}/h$ box size with 2×256^3 resolution elements (L25n256) with $0.1c$, $0.3c$, and $1.0c$. Fig. 4 illustrates the gas temperature T against redshift z, with varying overdensities of $1\bar{\rho}$, $10\bar{\rho}$, $100\bar{\rho}$, $1000\bar{\rho}$. Generally,

$T\left(\bar{\rho}\right)$ and $T\left(1000\,\bar{\rho}\right)$. showed an acceptable convergence at all redshifts. When $z<6.5$, however, the $0.1c$ simulations with $10\,\bar{\rho}$ and $100\,\bar{\rho}$ showed an underestimation of $5{,}000-10{,}000\mathrm{K}$.

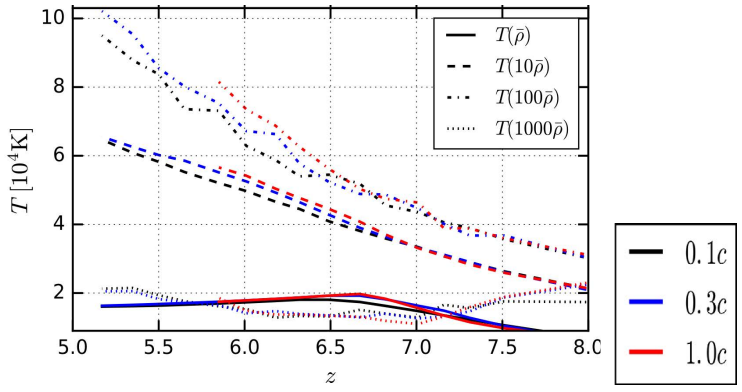

Fig. 4 The simulated gas temperature against redshift with different overdensities [13].

The effect of such differences can be evaluated through the comparison between FG09 simulation (mimicking $0.1c$ simulation) and FG09x10 simulation (mimicking $1.0c$ simulation), the former one with 10 times smaller photoionization rates. It can be shown that, at $z=6$, the $\sim 10{,}000\mathrm{K}$ difference in the halo gas temperature only caused an enlargement of $0.1\sim 0.2\mathrm{dex}$ in the halo mass needed for SFR suppression. Furthermore, the same simulation also showed good accuracy in intergalactic medium temperatures. Briefly, in this situation, the accuracy of the simulation can be kept even with such a low value of $0.1c$, suggesting that the RSLA is a credible and robust approach.

5.3 Comparisons with JWST

Arepo-RT is a decisive component of the THESAN project. Hence, the results obtained by the THESAN project can evidently reveal the performance of Arepo-RT. According to Garaldi et al. [7], the THESAN simulations show excellent agreement with the JSWT observation results. On the one hand, the galaxy main sequence at high redshift ($z<6$) simulated by THESAN matches the JWST observations well. In the low-redshift context, the main sequence is the majority of galaxies with their total stellar mass (M_*) correlated with their star formation rate (SFR). However, such correlation at high redshifts ($z>6$) is still under debate. The left panel of Fig. 5 compares the simulated results with the observations, from which one can conclude that the two sources of data are generally consistent with each other, despite some over-predicted SFRs when $M_*\lesssim 10^8 M_\odot$ and under-predicted SFRs partly when $M_*\gtrsim 10^9 M_\odot$. This consistency offers strong evidence to understand the high-redshift correlation of galaxy main sequence.

On the other hand, THESAN also relatively accurately predicts the UV slope distribution at $z\geqslant 5.5$. The UV slope (β) is defined as the power-law index of the rest-frame galaxy UV continuum, which is related to the stellar ages and the metal compositions. A lower β means a *bluer* radiation, which indicates a young galaxy with little dust and metal, while $\beta\approx -3$ is considered as a limit value. The right panel shows the UV slope–UV magnitude relation $\beta\left(M_{\mathrm{UV}}\right)$ in five redshift intervals. Apparently, the galaxies simulated by THESAN lie in a U-shaped distribution: (i) at $\mathrm{M_{UV}}\approx -18$ is the lowest (*bluest*) UV slopes, typically implying some young galaxies; (ii) For brighter or fainter ones ($\mathrm{M_{UV}}\neq -18$), the UV slope increases towards both poles, possibly due to some older galaxies, where dust gradually escalates and makes the radiation redder. Compared with the observed data, the bright galaxies depict an overall alignment with observed UV slopes, but some faint stars show significant deviations due to the selection biases (limitations on photometric measurement). Despite a larger scatter compared to the simulations, the simulation is considered successful even with only a small volume [7].

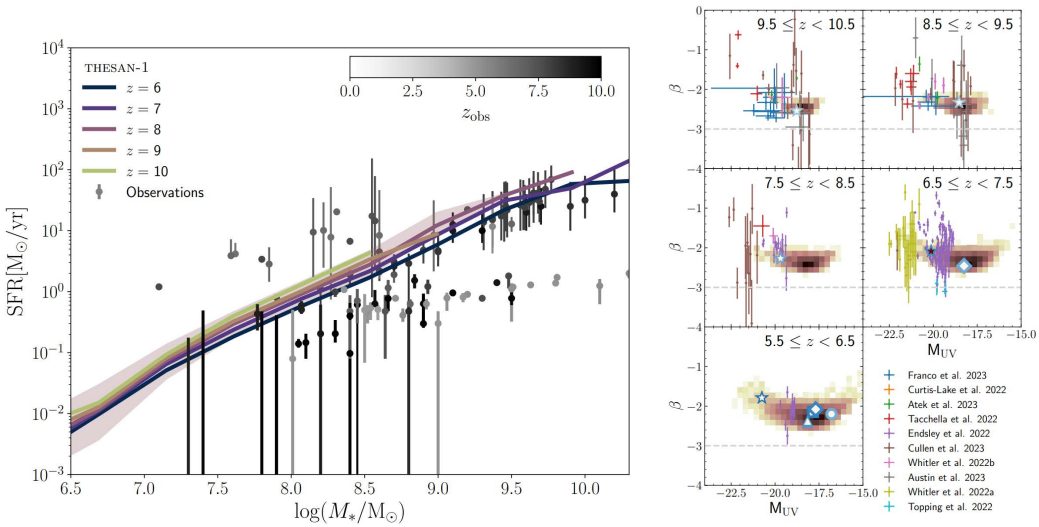

Fig. 5 The star formation rate plotted with observation data (left) and the UV slope distribution where darker squares mean smaller UV slope (right) [7].

6. Limitations

6.1 M1 Closure

Many numerical solvers, including Arepo-RT, employ the M1 Closure relation, which performs well for most problems. However, the M1 closure relation has some inherent disadvantages, such as the failure of asymptotic limit and the imperfection of modeling multiple-source geometry. According to Kannan et al. [5], modeling convergent rays can highlight the shortcomings of M1 Closure. A simulation is carried out for the free-streaming radiation distribution from a thin disk in 2D box, where the disk was surrounded by a torus of optically thick medium, outside which was optically thin and tenuous gas (The disc kept constant photon energy density).

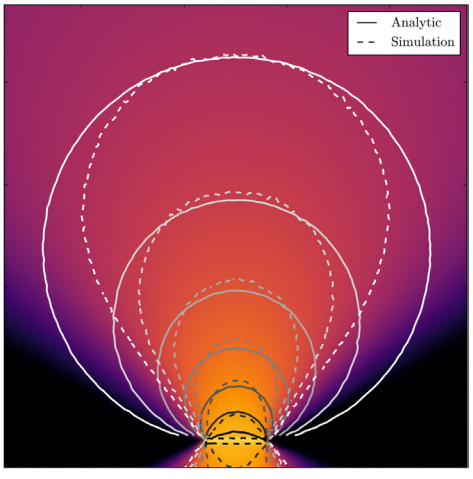

Fig. 6 The radiation field contours of analytic solution and numerical simulation [5].

Fig. 6 depicts the histogram of the radiation field intensity and the radiation field contours. As can be seen, the simulation matches the analytic result only qualitatively, where the simulation slightly overshoots the analytic solution in the y−direction but undershoots in the x−direction. In this scenario, the photons from the left side and the right side could have been passing through each other. However, the M1 closure usually overestimated the interaction between the photons, blocking the x–direction movement of photons and producing spurious fluxes at y–direction. Consequently, the M1 Closure caused deviations from the analytic solution [5]. Simulations using M1 Closure should be careful about these issues. If possible, an analytic solution should be attempted and added to the algorithm, so as to compensate for the numerical schemes.

6.2 Missing Photons

Deng et al. considered the evolution of H_{II} regions around massive stars, which spans large ranges both spatially and temporally [15]. The small-scale physics inside the Strömgren spheres is decisive to understanding the evolution but simulating the process requires high resolutions. (Strömgren sphere: a very small and fully ionized regions with a temperature $\approx 10^4$K from which the massive star's radiative feedback originates). They points out another issue of the Arepo-RT or similar uncoupled solvers: using an insufficient temporal resolution can result in a serious error, the missing photons.

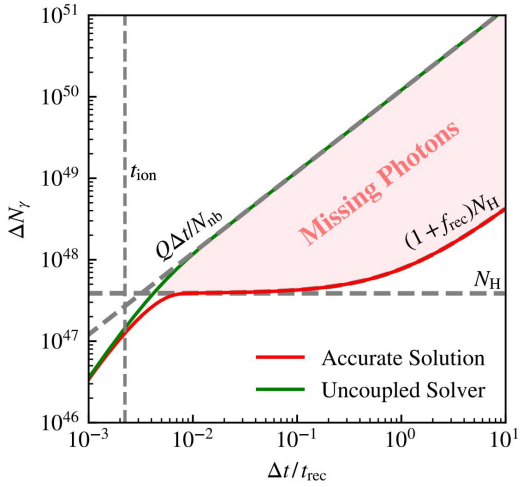

Fig. 7 The change in photon number density obtained by an accurate solution and an uncoupled solver, with the missing photons shown [15].

The missing photons are those absorbed in a medium but without any effect on the ionization state of the medium. This error is mainly due to the operator splitting scheme used in Arepo-RT and other uncoupled solvers, which calculate the photon absorption independently of solving thermochemistry equations. If a too long timestep is used, the error emerges where there is not a uniform cadence in the photon thermochemistry. It can be illustrated by simulating the photon absorption in neutral gas with and without the split scheme, respectively. As shown in Fig. 7, as $\Delta t / t_{\rm rec}$ (ratio of the timestep to the recombination timescale) increases, the uncoupled solver dramatically deviates from the accurate solution, leaving a missing photon region. As discussed in Deng et al. [15], in a timestep, the total amount of absorbed photons ΔN_γ and that of missing photons δN_γ can be modeled as:

$$\Delta N_\gamma = \frac{Q}{N_{\rm nb}}\Delta t, \delta N_\gamma = \frac{Q}{N_{\rm nb}}\Delta t - \frac{Q}{\alpha_{\rm B} n_{\rm H} \mathcal{R}_i}(1+ f_{\rm rec}). \quad (24)$$

when Δt is large enough, $\delta N_\gamma \approx \Delta N_\gamma$, which means almost all photons are absorbed without any effect on the ionization of the gas.

More seriously, the missing photons issue will propagate in the next timesteps, as it will repeat after each injection. However, there is no mechanism to compensate for these missing photons. To settle this problem, one can simply abandon the split scheme but instead couple the photon densities with the thermochemical solver. Still, this appeasement cannot solve the problem of the high demand of computational resources, especially when many ion species are included. Furthermore, Jaura et al. suggested an approximate uncoupled solver to correct the issue [17]. The new solver forces the number of absorbed photons not to surpass the number of ionizable atoms, provided that there is no sudden change in recombination rate. This approach is inexpensive but suitable for most cases.

6.3 Others

As mentioned by Borrow et al. [18], the assumption of a uniform UV background in calculating metal cooling rates can result in potential problems, if reionization is simulated to proceed differently. Nonetheless, fully calculating with the non-equilibrium metal cooling model is still impractical, since the computational cost is huge. Besides, as Garaldi et al. pointed out [7], the Arepo model used in THESAN only employs the two-phase interstellar medium (ISM) model. Yet, a multi-phase ISM model is necessary for more self-consistent simulations, but the compatibility for multi-phase model does not inhere in the algorithm. Thus, the current simulations are unable to cover cold, dense and molecular phases of the ISM, giant molecular clouds, etc.

7. Conclusion

To sum up, this study focused on the main features of Arepo-RT, introducing the physical models and numerical schemes. Specifically, the radiative transfer model is a moment-based form of Euler Continuity Equation, where the source term is set to zero. The radiative transfer possesses a similar algebraic structure to that of hydrodynamics. Technically, it is conducted on a moving mesh, with the intercell fluxes calculated through HLL and GLF schemes. The time integration is a hybrid of Heun's method and MUSCL, where the timesteps are sub-cycled within hydrodynamics steps. The results and limitations of Arepo-RT are also included in this paper, as a crucial

proof for the capabilities of the solver. These methodologies incorporated by Arepo-RT are representative of the radiation simulations. This reserach aims to provide an overview of cosmological radiation simulations, so that the readers can understand those principles with ease and thus learn the insights of those simulations. This paper is not only educational, but helps contribute to the knowledge system, making this field more open to the public. The author hopes more people can be inspired by this paper and get involved in astrophysical simulation research. However, the field is highly interdisciplinary and rapidly evolving, when new algorithms and computational forces are continually emerging. A further step in catching up on the new developments is still necessary.

References

[1] Springel V. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. Monthly Notices of the Royal Astronomical Society, 2010, 401(2): 791–851.

[2] Graziani F. Radiation Diffusion: An Overview of Physical and Numerical Concepts. Open Issues in Core Collapse Supernova Theory, 2005: 29–66.

[3] Vogelsberger M, Marinacci F, Torrey P, et al. Cosmological simulations of galaxy formation. Nature Reviews Physics, 2020, 2: 42–66.

[4] Rosdahl J and Teyssier R. A scheme for radiation pressure and photon diffusion with the M1 closure in RAMSES-RT. Monthly Notices of the Royal Astronomical Society, 2015, 449(4): 4380–4403.

[5] Kannan R, Vogelsberger M, Marinacci F, McKinnon R, Pakmor R and Springel V. AREPO-RT: radiation hydrodynamics on a moving mesh. Monthly Notices of the Royal Astronomical Society, 2019, 485(1): 117–149.

[6] Kannan R. Introducing the THESAN project: radiation-magnetohydrodynamic simulations of the epoch of reionization. Monthly Notices of the Royal Astronomical Society, 2022, 511(3): 4005–4030.

[7] Garaldi E. The THESAN project: public data release of radiation-hydrodynamic simulations matching reionization-era JWST observations. Monthly Notices of the Royal Astronomical Society, 2024, 530(4): 3765–3786.

[8] Weinberger R, Springel V and Pakmor R. The AREPO Public Code Release. The Astrophysical Journal Supplement Series, 2020, 248(2).

[9] McKinnon R, Torrey P and Vogelsberger M. Dust formation in Milky Way-like galaxies. Monthly Notices of the Royal Astronomical Society, 2016, 457(4): 3775–3800.

[10] Rosdahl J, Blaizot J, Aubert D, Stranex T and Teyssier R. RAMSES-RT: radiation hydrodynamics in the cosmological context. Monthly Notices of the Royal Astronomical Society, 2013, 436(3): 2188–2231.

[11] Kannan R. Simulating the interstellar medium of galaxies with radiative transfer, non-equilibrium thermochemistry, and dust. Monthly Notices of the Royal Astronomical Society, 2020, 499(4): 5732–5748.

[12] Weinberger R. Supermassive black holes and their feedback effects in the IllustrisTNG simulation. Monthly Notices of the Royal Astronomical Society, 2018, 479(3): 4056–4072.

[13] Wu X, Kannan R, Marinacci F, Vogelsberger M and Hernquist L. Simulating the effect of photoheating feedback during reionization. Monthly Notices of the Royal Astronomical Society, 2019, 488(1): 419–437.

[14] Pakmor R. Improving the convergence properties of the moving-mesh code AREPO. Monthly Notices of the Royal Astronomical Society, 2016, 455(1): 1134–1143.

[15] Deng Y, Li H, Kannan R, Smith A, Vogelsberger M and Bryan G L. Simulating ionization feedback from young massive stars: impact of numerical resolution. Monthly Notices of the Royal Astronomical Society, 2024, 527(1): 478–500.

[16] Zier O, Kannan R, Smith A, Vogelsberger M and Verbeek E. Adapting AREPO-RT for exascale computing: GPU acceleration and efficient communication. Monthly Notices of the Royal Astronomical Society, 2024, 533(1): 268–286.

[17] Jaura O, Glover S C O, Klessen R S and Paardekooper J P. SPRAI: coupling of radiative feedback and primordial chemistry in moving mesh hydrodynamics. Monthly Notices of the Royal Astronomical Society, 2018, 475(2): 2822–2834.

[18] Borrow J. THESAN-HR: how does reionization impact early galaxy evolution. Monthly Notices of the Royal Astronomical Society, 2023, 525(4): 5932–5950.