Goethe University Frankfurt am Main
Department of Physics
Institute for Theoretical Physics

Master's Thesis

Modeling high-redshift X-ray sources and cosmological redshifting in the radiation hydrodynamics code AREPO-RT

Sonja Marie Köhler

Frankfurt am Main
31.05.2025

First examiner: Prof. Dr. Laura Sagunski
Second examiner: Asst.-Prof. Dr. Rahul Kannan

Abstract

The 21-cm line signal from neutral hydrogen is the subject of great interest as a unique window into the thermal and (re-)ionization history of the early Universe. Complementing extensive observational data expected in the coming decades from a theoretical standpoint, numerical radiation hydrodynamics simulations are an indispensable tool shedding light on the first galaxies and the Epoch of Reionization. In this context, accurately modeling high-redshift X-ray sources is crucial for predictions of the 21-cm signal, in particular due to the latter's sensitive dependence on the temperature of the intergalactic medium.

To this end, the first aim of this thesis is to implement the X-ray emissions of key astrophysical sources at high redshifts in the radiation hydrodynamics solver AREPO-RT within the moving-mesh code AREPO. This is accomplished by extending the modeled spectra of stars and active galactic nuclei into the soft X-ray band, described by added higher-energy frequency bins, as well as by newly incorporating the radiative contributions of the supernova shock-heated interstellar medium and high-mass X-ray binaries as additional sources.

Necessitated by the large mean free path of X-ray photons on cosmological scales, the other major aim is implementing a term accounting for the effects of cosmological redshifting. Two different methods for this are investigated and compared. The first modifies the moment equations of numerical radiative transfer based on an isotropic, piecewise power-law approximation of the specific radiative intensity across successive frequency bins. Additionally, a new numerical redshifting method is developed and introduced. This second method is based on an effective re-binning of radiative energy in frequency space, directly modeling the transfer of photons towards lower-energy bins at each timestep of a simulation.

The theoretical and numerical basis of the redshifting models, as well as the extended spectrum and newly included sources of X-ray radiation, are described in detail. Further, results of test runs verifying the physical validity and numerical stability of the implementation are presented and discussed.

Zusammenfassung

Das 21-cm-Signal von neutralem Wasserstoff ist von großem Forschungsinteresse, da es einzigartige Einblicke in die thermische Entwicklung und den (Re-)Ionisiationsprozess des frühen Universums ermöglicht. Neben umfangreichen Beobachtungsdaten, die in den kommenden Jahrzehnten zu erwarten sind, stellen numerische, strahlungshydrodynamische Simulationen ein unverzichtbares Mittel dar, um die ersten Galaxien und die Reionisationsepoche zu erforschen. In diesem Kontext ist es wichtig, Röntgenquellen im frühen Universum korrekt zu modellieren, um Vorhersagen über das 21-cm-Signal zu treffen – insbesondere da letzteres stark von der Temperatur des intergalaktischen Mediums abhängt.

Daher ist das erste Ziel dieser Masterarbeit, die Röntgenemissionen von wichtigen astrophysikalischen Quellen bei hoher Rotverschiebung in den strahlungshydrodynamischen Code AREPO-RT zu implementieren, der Teil des 'moving-mesh' ('mitbewegtes Gitter') AREPO-Codes ist. Dies wird erreicht, indem das modellierte Spektrum von Sternen und aktiven galaktischen Kernen auf die Energien weicher Röntgenstrahlen ausgeweitet wird, welche durch zusätzliche, energiereichere Frequenz-Bins beschrieben werden. Zudem werden als weitere Röntgenquellen die Strahlung des interstellaren Mediums, das von Supernovae schockerhitzt wurde, und die Strahlung massereicher Röntgendoppelsterne neu miteinbezogen.

Aufgrund der großen mittleren freien Weglänge von Röntgenphotonen im Universum ist das weitere Hauptziel dieser Arbeit, einen Term zu implementieren, der die Effekte kosmologischer Rotverschiebung berücksichtigt. Zwei verschiedene Methoden hierfür werden untersucht und verglichen. Die erste Methode modifiziert die Momentengleichungen des numerischen Strahlungstransports und basiert auf der Approximation einer isotropen, stückweise durch ein Potenzgesetz zwischen Frequenz-Bins beschriebenen spezifischen Strahlungsintensität. Zusätzlich wird eine neue Methode zur Modellierung numerischer Rotverschiebung entwickelt und beschrieben. Diese zweite Methode basiert auf einem effektiven 'Re-binning' von Strahlungsenergie in den Frequenz-Bins und modelliert den Übergang von Photonen in energieärmere Bins in jedem Zeitschritt einer Simulation.

Die theoretische und numerische Basis der Rotverschiebungsmodelle sowie der ausgeweitete Frequenzbereich und die neu implementierten Röntgenquellen werden detailliert beschrieben. Weiterhin werden Ergebnisse von Testsimulationen, welche die physikalische Validität und numerische Stabilität der Implementierung bestätigen, vorgestellt und diskutiert.

Contents

List of Figures

List of Tables

List of Abbreviations

2LPT ..... Second-order Lagrangian perturbation theory

AGN ..... Active galactic nuclei

AMR ..... Adaptive mesh refinement

BPASS ..... Binary Population and Spectral Synthesis (code)

CBE ..... Collisionless Boltzmann equation

CMB ..... Cosmic microwave background

CPU ..... Central processing unit

DM ..... Dark matter

EER ..... Effective energy-rebinning

EoR ..... Epoch of Reionization

EoS ..... Equation of State

EUV ..... Extreme UV (photon energy range)

FFT ..... Fast Fourier transform (algorithms)

FLRW ..... Friedmann-Lemaître-Robertson-Walker (metric)

FE ..... Fluid element

GFM ..... Galaxy formation model

GPU ..... Graphics processing unit

HMXB ..... High-mass X-ray binary

HPC ..... High-performance computing

IC ..... Initial condition

IGM ..... Intergalactic medium

TNG ..... 'The Next Generation' (Illustris, Millennium models)

ISM ..... Interstellar medium

ΛCDM ..... Λ cold dark matter

LMXB ..... Low-mass X-ray binary

LyC ..... Lyman continuum (radiation)

MMFV ..... Moving-mesh finite-volume

ODE ..... Ordinary differential equation

OTVET ..... Optically thin variable Eddington tensor

PDE ..... Partial differential equation

PPL ..... Piecewise power-law (approximation / method)

RHD ..... Radiation hydrodynamics

RSLA ..... Reduced speed of light approximation

RT ..... Radiative transfer

SED ..... Spectral energy distribution

SFR ..... Star formation rate

SMBH ..... Supermassive black hole

SPH ..... Smoothed particle hydrodynamics

TreePM ..... Tree particle-mesh

HERA ..... *Hydrogen Epoch of Reionization Array*

JWST ..... *James Webb Space Telescope*

LOFAR ..... *Low Frequency Array*

MWA ..... *Murchison Widefield Array*

PAPER ..... *Precision Array for Probing the Epoch of Reionization*

SKA ..... *Square Kilometre Array*

1 Introduction

The *Epoch of Reionization (EoR)* is a crucial, yet observationally and theoretically still rather uncertain, period in the evolution of the Universe (e.g. Furlanetto et al. 2006; Kannan et al. 2022a; Asthana et al. 2024). Following the formation of the first stars and (proto-) galaxies at cosmic dawn, the EoR is characterized by a cosmic-scale phase transition from mostly cold neutral gas to a hot ionized plasma, occurring within the first billion years of cosmic history (see Fig. 1). As such, it represents a key link in the Universe's evolution from the linear regime of structure formation constrained by the Cosmic Microwave Background (CMB) to the highly non-linear, clustered structures observed in the local Universe (e.g. Mellema et al. 2013; Dayal & Ferrara 2018).

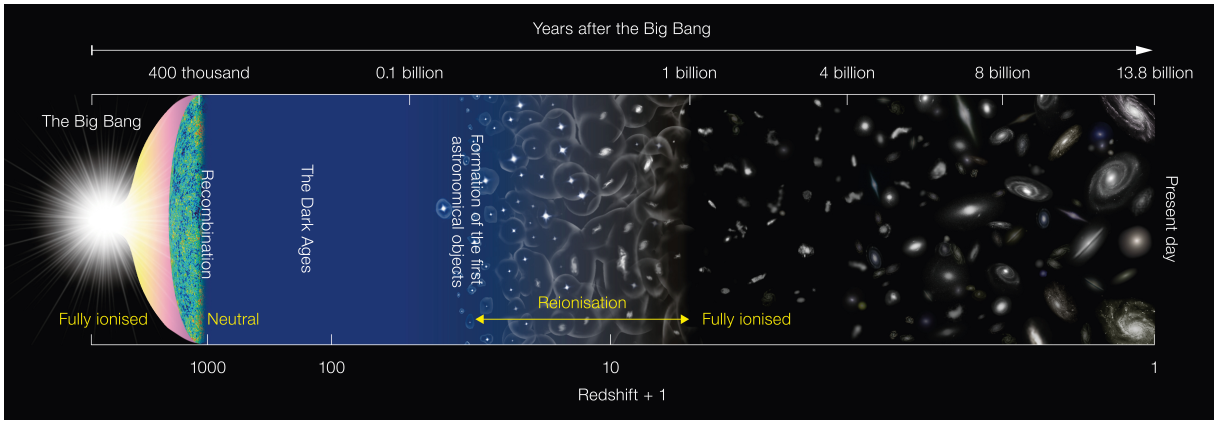

Figure 1: Schematic diagram of cosmic history. This figure was obtained from ESO / NAOJ (2016).

The process of reionization is primarily driven by stellar sources emitting UV-range Lyman continuum (LyC, $e \geq 13.6\,\mathrm{eV}$) radiation (e.g. Kannan et al. 2022a). Consequently, the phase transition progresses in growing 'bubbles' of hot, ionized gas around these sources, which eventually overlap as the entire Universe is reionized in the final stages of the EoR (see Fig. 2).
Notably, both cosmic dawn and the EoR remain highly active areas of investigation, with open questions concerning e.g. (i) the precise timing of the reionization of the Universe (e.g. Dayal & Ferrara 2018; Kannan et al. 2022a), (ii) the first appearance, radiative output, masses, and growth of early galaxies and active galactic nuclei (AGN; e.g. Silk et al. 2024; Harikane et al. 2024; Yuan et al. 2024; Greene et al. 2024), and (iii) the detailed thermal history of the intergalactic medium (IGM) during these periods (e.g. Pritchard & Loeb 2012).

Considerable observational effort is therefore dedicated to exploring these questions using both current and upcoming instruments. Firstly, the recently launched *James Webb Space Telescope* (*JWST*; Gardner et al. 2006) has already revolutionized our understanding of

the typical luminosities and masses of distant galaxies (e.g. Harikane et al. 2024; Weibel et al. 2024; Labbé et al. 2023) and AGN (e.g. Greene et al. 2024; Asthana et al. 2024), marking the beginning of a new era of precision observations of the early Universe.
Secondly, radio frequency interferometers such as the *Low Frequency Array* (*LOFAR*; Stappers et al. 2011), *Precision Array for Probing the Epoch of Reionization* (*PAPER*; Parsons et al. 2010), *Murchison Widefield Array* (*MWA*; Tingay et al. 2013), *Hydrogen Epoch of Reionization Array* (*HERA*; DeBoer et al. 2017), and the *Square Kilometre Array* (*SKA*; Mellema et al. 2013; Braun et al. 2019) currently under construction, are all striving to shed light on the EoR and gain a more complete picture of the physical processes governing it.
Specifically, the *21-cm line signal* is a key observational probe in this context, serving as a signature of neutral hydrogen (H I) in the Universe and encoding information about both its distribution and temperature. This emission line – with its characteristic rest-frame microwave wavelength $\lambda = 21\,\mathrm{cm}$ – originates from a spin-flip transition in H I atoms, in which the orientation of the electron spin reverses with respect to the proton spin (van de Hulst 1945; Ewen & Purcell 1951).
The effects of cosmological redshifting then allow observations to map out the 21-cm signal's evolution in cosmic time, thereby providing insights into the evolution of the thermal and ionization state of the Universe (e.g. Furlanetto et al. 2006; Pritchard & Loeb 2012).

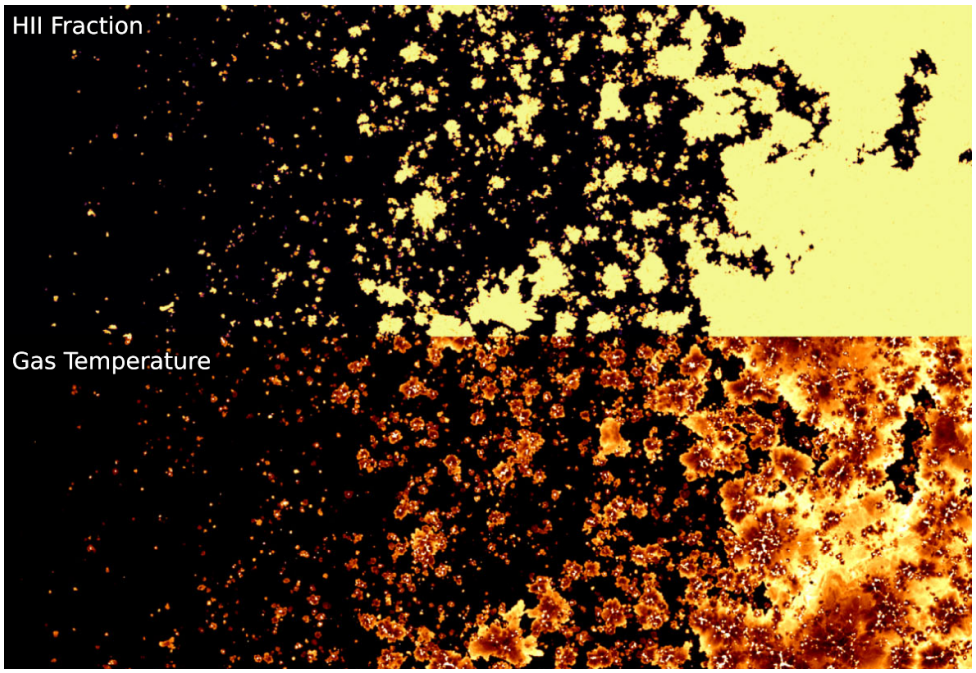

Figure 2: Mock lightcones showing the ionized hydrogen (H II) fraction and gas temperature from the THESAN reionization simulations (Kannan et al. 2022a; Smith et al. 2022; Garaldi et al. 2022) in relative units, illustrating the evolution of these properties in the intergalactic medium throughout the Epoch of Reionization. This figure was obtained and adapted from (Kannan et al. 2022a).

To complement this anticipated wealth of observational data in the coming years and decades, numerical *radiation hydrodynamics (RHD)* simulations are a crucial theoretical approach to testing and constraining current models of the EoR. RHD simulations are uniquely capable of directly, accurately, and self-consistently modeling the coupling of simulated cosmic structures to the radiation fields driving reionization (see Sec. 3.2). This, in turn, opens the door to a wide range of detailed predictions for key observables from a theoretical standpoint (see e.g. Gnedin 2014; Rosdahl & Teyssier 2015; Ocvirk et al. 2016, 2020; Kannan et al. 2022a; Garaldi et al. 2024; Kannan et al. 2025; Zier et al. 2025b; Shen et al. 2025). Examples include the evolution of the neutral hydrogen and helium fractions with redshift, i.e., the *reionization histories*, as well as the evolution of not only the average temperature of astrophysical gas, but also the spatial distribution of different gas temperatures in the simulation (see e.g. Sec. 4). Sufficiently large-volume ($V \gtrsim \mathcal{O}((100\,\mathrm{cMpc})^3)$) RHD simulations additionally facilitate numerical predictions for the 21-cm signal (e.g. Kannan et al. 2022a).

The latter is a key science goal of the currently ongoing THESAN-XL project (Zier et al. 2025a), forming the broader context of this thesis. At its core, THESAN-XL includes an extremely large-volume ($V \sim (400\,\mathrm{cMpc})^3$) high-resolution reionization simulation, which employs the RHD simulation code AREPO-RT (Kannan et al. 2019; Zier et al. 2024; see also Sec. 3.2) within the moving-mesh finite-volume hydrodynamics code AREPO (Springel 2010b; Pakmor et al. 2016; see also Sec. 3.1.3). It further builds on previous simulation efforts using AREPO-RT, most notably the recent large-volume ($V = (95.5\,\mathrm{cMpc})^3$) reionization simulation suite THESAN (Kannan et al. 2022a; Smith et al. 2022; Garaldi et al. 2022).
Due to the sensitive dependence of the 21-cm signal on the IGM temperature during the EoR, accurately modeling high-redshift X-ray sources – as well as the evolution of the radiative energy emitted by them – is crucial in this context. Specifically, the temperature of the IGM strongly depends not only on the presence and total luminosity of X-ray sources (e.g. Furlanetto 2006; Mesinger et al. 2013), but also on the details of their respective spectral energy distributions (SEDs; Pacucci et al. 2014).

Consequently, the first major aim of this thesis is to implement numerical models capturing the radiative emissions of key X-ray sources at high redshifts in AREPO-RT. To this end, the range of photon energies tracked by the code is increased from THESAN's *extreme UV (EUV)* range $[13.6\,\mathrm{eV}, 100\,\mathrm{eV}]$ to include the *soft X-ray band* ($[500\,\mathrm{eV}, 2\,000\,\mathrm{eV}]$), resulting in the significantly extended total range $[13.6\,\mathrm{eV}, 2\,000\,\mathrm{eV}]$ (see Sec. 3.3.1).
Further, new numerical models of *high-mass X-ray binaries (HMXB)* and the supernova shock-heated *interstellar medium (ISM)* are integrated into the code (Secs. 3.3.3 and 3.3.4). Finally, the SED of the radiative emissions of AGN is updated in order to more accurately model the extended photon energy range (Sec. 3.3.5).

Notably, X-ray photons have a significantly larger mean free path than the lower-energy UV photons due to the energy dependence of the photoionization cross-sections of neutral hydrogen H I, neutral helium He I, and singly ionized helium He II (e.g. Furlanetto et al. 2006; Eide et al. 2020; see also Sec. 2.1). Therefore, they have the potential to undergo significant *cosmological redshifting* before being absorbed. This decrease of photon energies, caused by the expansion of the Universe (see e.g. Baumann 2022 for further details), is commonly neglected in numerical radiative transfer (RT; e.g. Gnedin & Ostriker 1997; Gnedin & Abel 2001; Petkova & Springel 2009; Kannan et al. 2019, 2022a).
This is based on the assumption that the mean free path of EUV photons is sufficiently short – given their large photoionization cross-sections with the astrophysically abundant H I, He I, and He II species – for the impact of redshifting to be negligible on the scales of interest. However, the same approximation is no longer well-justified when modeling X-ray photons, particularly on the extensive spatial scales of a large-volume simulation.
Therefore, the second major aim of this thesis is to implement cosmological redshifting in AREPO-RT. Due to the complexity and fundamental limitations of numerical redshifting, two separate methods are integrated into the code, investigated, and compared. While the first of these approaches (*piecewise power-law (PPL)* method; Sec. 3.4) is primarily based on an existing method (Finlator et al. 2009, 2018) that is adapted to the requirements of the AREPO-RT code base, the second approach (*'effective energy-rebinning' (EER)*; Sec. 3.5) introduces a conceptually new method developed in the context of this thesis.

This work is organized as follows. First, the relevant theoretical background concerning X-ray heating and key high-redshift X-ray sources, as well as essentials of fluid dynamics and RHD, are introduced in Sec. 2. Sec. 3 forms the core of this thesis; it first presents an overview of the numerical methods used in cosmological (radiation-) hydrodynamic simulations, focusing on those employed within AREPO-RT, before discussing the newly developed implementations of X-ray sources and cosmological redshifting in detail. Results of test runs performed in the context of these new implementations are presented and discussed in Sec. 4. Finally, the summary and conclusions to this thesis are provided in Sec. 5.

2 Theoretical background

2.1 X-ray heating and key sources at high redshifts

During the EoR, soft-band X-ray radiation is expected to cause a more uniform heating of the IGM than UV photons alone, thereby leaving a characteristic imprint on the large-scale 21-cm power spectrum (Mesinger et al. 2013; cf. Sec. 1). In the following, we briefly describe the two key causes of this, before discussing the specific astrophysical sources most relevant to X-ray heating at high redshifts.

Firstly, we note that the photoionization cross-sections of H I, He I, and He II can be described in terms of analytic fitting functions (see Eq. (1) and Table 1 of Verner et al. 1996); these functional forms are shown in Fig. 3 for the EUV to soft X-ray energy range relevant to AREPO-RT ([13.6 eV, 2 000 eV]). Crucially, Fig. 3 shows that the photoionization of these species is most efficient for photon energies at and slightly above their respective ionization threshold energies in the EUV range – i.e., 13.6 eV for H I, 24.59 eV for He I, and 54.42 eV for He II– while the cross-sections rapidly decrease with increasing photon energies.

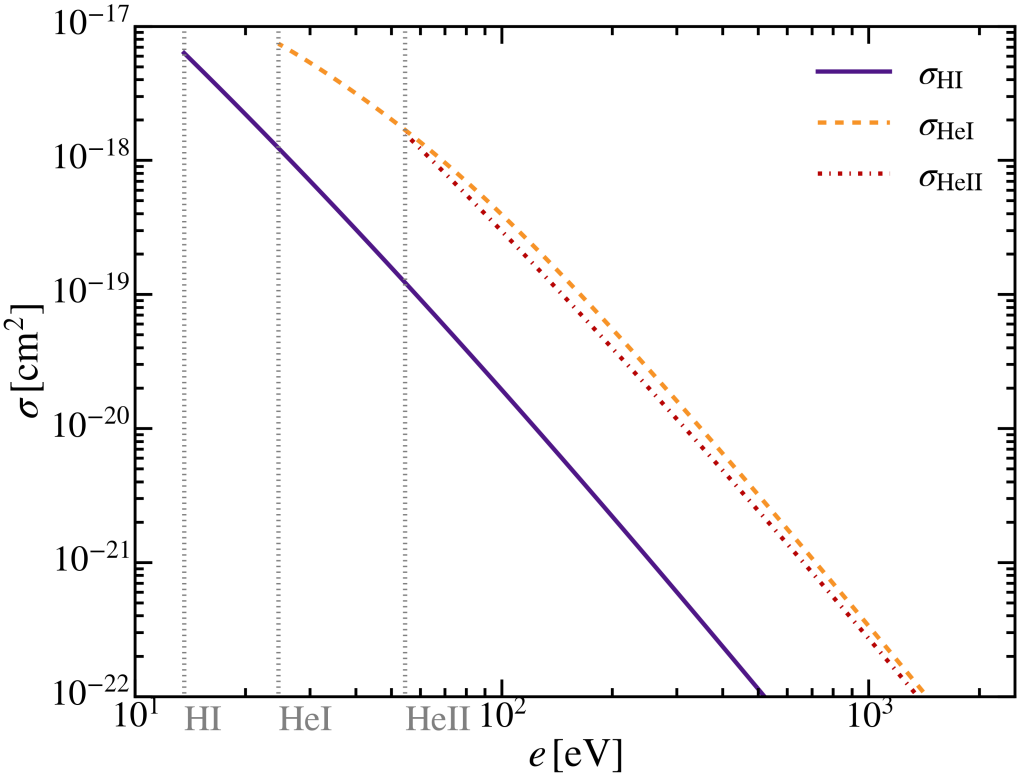

Figure 3: Photoionization cross-sections σ (Verner et al. 1996) of neutral hydrogen H I (purple solid curve), neutral helium He I (orange dashed curve), and singly ionized helium He II (red dash-dotted curve) as a function of photon energy e. Gray dotted vertical lines indicate the ionization threshold energies of these chemical species at 13.6 eV, 24.59 eV, and 54.42 eV, respectively.

This is the key reason for the long mean free path of X-ray photons mentioned in Sec. 1, which typically significantly exceeds the mean free path of EUV photons during the EoR due to the high abundances of H I, He I, and He II gas (cf. e.g. Eide et al. 2020, App. A). Consequently, X-rays are capable of propagating further into the IGM even while the neutral fractions of hydrogen and helium are still substantial, i.e., during the early stages of reionization, in which EUV photons still tend to be absorbed in the immediate surroundings of galaxies.

Secondly, X-ray photons contribute more energy to photoheating *per ionization event* than EUV photons. Specifically, photoheating due to the chemical species j is caused by photoionization events in which the photon energy $e = h\nu$ exceeds the ionization threshold energy $h\nu_{t,j}$, thereby transferring the 'leftover' energy,

$$\Delta e_j = h\nu - h\nu_{t,j} \ , \tag{1}$$

to heat. Here, h and ν refer to Planck's constant and the photon's frequency, respectively, while $\nu_{t,j}$ denotes the ionization threshold frequency for species j.

On the one hand, higher-energy photons therefore cause a larger injection of energy in the form of heat per ionization event; on the other hand, for photons with energies above $\sim 2\,\mathrm{keV}$, photoionization itself becomes so inefficient due to the small cross-sections that their contribution to both ionization and heating becomes negligible (Pacucci et al. 2014). For these reasons, the soft X-ray band is key to the uniform IGM heating mentioned above: here, the photon energies are large enough to allow propagation far into the IGM – but not so large that the photons essentially free-stream, never interacting with the vast majority of astrophysical gas – and each ionization event contributes a more substantial amount to photoheating than is the case for the lower-energy EUV photons.

Consequently, the most relevant sources of X-rays during the EoR are those which (i) emit a substantial amount of energy within the soft X-ray band, and (ii) already do so at the high redshifts ($z \gtrsim 5 - 6$) of this epoch in cosmic history.
The dominant sources fulfilling both of these criteria are those mentioned in Sec. 1: the hot ISM, HMXBs, and AGN (Pacucci et al. 2014; Eide et al. 2018, 2020).
Firstly, the hot ISM refers to the diffuse, volume-filling, hot ($T \approx \mathcal{O}(10^6 - 10^7\,\mathrm{K})$; Mineo et al. 2012b; Pacucci et al. 2014) phase of gas surrounding stars within a galaxy. This gas is primarily shock-heated by the effects of supernovae and stellar winds, resulting in an SED that is well-approximated by a thermal bremsstrahlung spectrum (Pacucci et al. 2014; see Sec. 3.3.3 for details).
Secondly, HMXBs are binary systems comprised of a massive star and a compact object,

i.e., a neutron star or stellar black hole, in which the compact object is accreting material from the star. The latter is therefore also referred to as the *donor* star of the system. This forms an accretion disk, which efficiently converts the gravitational energy released by the inspiraling material to highly energetic radiation, resulting in significant emissions in the X-ray range (e.g. Mineo et al. 2012a; Sec. 3.3.4). More generally speaking, X-ray binaries – also including intermediate- to low-mass X-ray binaries (LMXB) – are a broader class of these accreting binary systems, in which the donor can be a less massive star or even a white dwarf. However, due to the significantly longer lifespans of lower-mass stars, these systems operate on substantially longer timescales than HMXBs (e.g. Mineo et al. 2012a). Consequently, LMXBs only become significant contributors of X-ray radiation at redshifts $z \gtrsim 1.5$ (Madau & Fragos 2017), i.e., billions of years after the end of reionization at $z \sim 5 - 6$ (e.g. Kannan et al. 2022a; see also Sec. 4).
Thirdly, AGN are accreting supermassive black holes (SMBHs) at the centers of galaxies. The most relevant component of AGN in the context of this thesis are their highly luminous accretion disks, formed as the immense gravitational effects of the SMBH accelerate and heat surrounding matter. This results in radiative emissions across a broad spectrum of photon energies (e.g. Shen et al. 2020; see also Sec. 3.3.5).
Finally, we note that stars – while being the dominant source of H I-ionizing EUV photons and photoheating during the EoR (e.g. Madau et al. 1999; Haardt & Madau 2012; Eide et al. 2018, 2020) – are negligible sources of X-ray photons due to a steep drop-off in their luminosity at high photon energies $h\nu \gtrsim 54.4\,\mathrm{eV}$ (e.g. Haardt & Madau 2012).

2.2 Fluid dynamics

In essence, the term *fluid dynamics* refers to the theoretical description of the state and evolution of fluids, i.e., substances which (i) continuously deform under the influence of stress forces, (ii) adapt to the shape of a container, and (iii) are made up of a large number of constituent particles flowing freely past one another (e.g. van den Bosch 2020). Fluids can broadly be classified depending on the interactions between the constituent particles, such as those arising from gravitational, van der Waals, or Coulomb forces. In the following, dynamics and examples of *collisionless* and *collisional* astrophysical fluids are discussed.

2.2.1 Collisionless fluid dynamics

In a fully collisionless fluid, the fluid's constituent particles exclusively interact via the force of gravity. One of the most crucial astrophysical examples of this is cold dark matter, i.e., the dominant matter component of the universe within the standard Λ cold dark matter (ΛCDM; for details, see e.g. Baumann 2022) model. Additionally, the collisionless fluid description also serves as an excellent approximation for the large-scale dynamics of

stars within a galaxy, as individual stars are spatially separated on $\mathcal{O}(\mathrm{pc} - \mathrm{kpc})$ scales and can therefore be interpreted as an extremely dilute fluid of neutral, gravitationally interacting particles with a near-zero collision cross-section (e.g. van den Bosch 2020).
In the Newtonian continuum limit, the dynamics of such collisionless systems are described by the coupling of the *collisionless Boltzmann equation (CBE)*,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0 \,, \tag{2}$$

to the *Poisson equation*,

$$\nabla^2 \Phi = 4\pi G \int f(\mathbf{x}, \mathbf{v}, t) \; \mathrm{d}^3\mathbf{v} \tag{3}$$

(e.g. Springel et al. 2001; Vogelsberger et al. 2020; Baumann 2022, App. B.1.1, p. 417).
Here, Φ and G denote the gravitational potential and Newton's constant, respectively. The distribution function $f = f(\mathbf{x}, \mathbf{v}, t)$ describes the phase-space density of the fluid, defined as the number density of constituent particles in six-dimensional phase space $(\mathbf{x}, \mathbf{v})$ – or, equivalently, $(\mathbf{x}, \mathbf{p})$ – as a function of position $\mathbf{x}$, velocity $\mathbf{v}$ or momentum $\mathbf{p} = m\mathbf{v}$, and time t. Eq. (2) therefore expresses the conservation of the local phase-space density, i.e., Liouville's theorem (e.g. Nolting 2007, Ch. 1.2.3, pp. 21-23).

2.2.2 Collisional fluid dynamics and the Euler equations

Freely flowing, subsonic astrophysical gases can be described to good approximation as *incompressible fluids*, which further typically have low viscosity ν, i.e., resistance to shear forces, and low conductivity κ, i.e., capacity to conduct heat through microscopic collisions (e.g. van den Bosch 2020).
A highly useful and common theoretical description of astrophysical gas dynamics is therefore achieved by treating the gas as an *ideal fluid*, $\nu = \kappa = 0$, obeying the *Euler equations*[1] (e.g. van den Bosch 2020; Vogelsberger et al. 2020). The latter describe the fluid in a *macroscopic continuum approach*, which neglects the discreteness of individual constituent particles and instead approximates the fluid as consisting of smooth fluid elements (FEs), also referred to as fluid packets. These FEs are essentially volumes of fluid that are required to be small compared to the characteristic scale of the hydrodynamical problem of interest, yet large compared to the mean free path of the constituent particles[2]. Further, each FE is required to consist of a large number of individual particles in order for the continuum description to hold, as well as to allow the definition of the hydrodynamical properties density ρ, fluid velocity $\mathbf{v}$, pressure P, and specific energy e (i.e., energy per unit mass) in each FE (see e.g. van den Bosch 2020).

[1] The Euler equations are equivalent to the *Navier-Stokes* hydrodynamical equations in the special case of an ideal fluid; for more details, see e.g. van den Bosch (2020).

[2] Consequently, this fluid description is not applicable to the collisionless fluids discussed in Sec. 2.2.1, as these have an infinite mean free path by definition.

The Euler equations can be equivalently expressed in both the so-called *Eulerian* and *Lagrangian* formulations.
The former approaches the description of ideal fluid dynamics within a 'fixed' frame of reference, i.e., the equations explicitly describe the fluid flow with respect to a spatially static observer:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{4}$$

$$\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} + P\,\mathbf{I}) = 0 \tag{5}$$

$$\frac{\partial (\rho e)}{\partial t} + \nabla \cdot \Big((\rho e + P)\,\mathbf{v} \Big) = 0\ . \tag{6}$$

Here, the operator $\otimes$ denotes the outer tensor product, while $\mathbf{I}$ is the 3×3 unity matrix. Further, $\rho = \rho(\mathbf{x})$, $\mathbf{v} = \mathbf{v}(\mathbf{x})$, $P = P(\mathbf{x})$, and $e = e(\mathbf{x})$ represent a field description of the hydrodynamical variables, specifying them at (fixed) positions $\mathbf{x}$.

Alternatively, the Lagrangian formulation of the Euler equations approaches their solution in a 'comoving' manner: rather than solving the equations with respect to a spatially fixed frame of reference through which the FEs flow, the FEs themselves serve as the frame of reference in this case:

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = -\rho\, \nabla \cdot \mathbf{v} \tag{7}$$

$$\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = -\frac{1}{\rho}\, \nabla P \tag{8}$$

$$\frac{\mathrm{D}e}{\mathrm{D}t} = -\frac{1}{\rho}\, \nabla \cdot P\mathbf{v}\ . \tag{9}$$

Here, the *Lagrangian derivative* $\mathrm{D}/\mathrm{D}t$, also referred to as the *convective* or *material* derivative, describes the change of a quantity from the point of view of an observer flowing along with the fluid,

$$\frac{\mathrm{D}}{\mathrm{D}t} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\ . \tag{10}$$

We note that Eq. (10) shows the equivalence of the Eulerian and Lagrangian formulations given by Eqs. (4)-(6) and Eqs. (7)-(9), respectively.

The first of the hydrodynamical equations in either formulation (Eq. (4) or (7)) is also referred to as the *hydrodynamical continuity equation*, expressing the conservation of the mass density ρ with time.
Further, the set of three equations given equivalently by either Eq. (5) or Eq. (8) are the *momentum equations*, describing how a gradient in the pressure P results in a change to the fluid velocity $\mathbf{v}$ or momentum density $\rho\mathbf{v}$ in the absence of external forces. Essentially,

this describes the force on FEs exerted by pressure gradients analogously to the Newtonian equation of motion $\mathbf{F} = m\mathbf{a} = \dot{\mathbf{p}}$; the formal similarity to $\mathbf{a} = \mathbf{F}/m$ is particularly clear in the Lagrangian formulation, Eq. (8). We note that the momentum equations are also occasionally referred to as the Euler equations in a stricter definition of the latter term (e.g. van den Bosch 2020; for counter-examples, see e.g. Springel 2010b; Vogelsberger et al. 2020).
Finally, the *energy equation* (Eq. (6) or Eq. (9)) describes the change in the total specific energy e of a FE caused by adiabatic compression or expansion, expressed explicitly as a non-zero velocity field divergence $\nabla \cdot \mathbf{v} \neq 0$ (see e.g. van den Bosch 2020).

Either formulation consists of a set of five equations for six variables (ρ, P, e, and the three components of $\mathbf{v}$), which therefore require an additional constraint to close the system of equations. This necessary closure is typically provided by an Equation of State (EoS) of the form (e.g. Springel 2010b)

$$P = P(\rho, u) = (\gamma - 1)\,\rho u\ . \tag{11}$$

Here, γ denotes the adiabatic index and u the internal specific energy, which is related to the total specific energy as $e = u + \frac{\mathbf{v}^2}{2}$.
Astrophysical gases are typically sufficiently dilute to be well-approximated as ideal gases, for which the adiabatic index depends only on the number of internal degrees of freedom q and is given by

$$\gamma = \frac{q+5}{q+3}\ , \tag{12}$$

resulting in the constant $\gamma = \frac{5}{3}$ in the case of a monatomic ($q = 0$) ideal gas (e.g. van den Bosch 2020). Therefore, a commonly used approximate EoS (cf. e.g. Vogelsberger et al. 2020) is given by

$$P(\rho, u) = \frac{2}{3}\rho u\ . \tag{13}$$

Further, in an astrophysical context, the effects of gravity need to be considered in addition to the pure fluid dynamics discussed above. Specifically, to include gravitational forces $\mathbf{F}_{\text{grav}} = -\nabla\Phi$ acting on a fluid, the force vectors are added to the hydrodynamical momentum equations, i.e.,

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} + P\mathbf{I}) = -\nabla\Phi, \tag{14}$$

or

$$\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = -\frac{1}{\rho}\nabla P - \nabla\Phi \tag{15}$$

in the Eulerian or Lagrangian formulations, respectively (e.g. van den Bosch 2020).
In particular, a self-gravitating fluid is described by coupling the fluid equations discussed

above to Poisson's equation,

$$\nabla^2\Phi(\mathbf{x}) = 4\pi G\rho(\mathbf{x}) \ , \tag{16}$$

cf. Sec. 2.2.1. We note that the density ρ used in the macroscopic continuum description is related to the phase-space density $f(\mathbf{x}, \mathbf{v}, t)$ of constituent particles as

$$\rho(\mathbf{x}, t) = \int f(\mathbf{x}, \mathbf{v}, t) \ \mathrm{d}^3\mathbf{v} \ . \tag{17}$$

In the fluid description, the form of Poisson's equation given in Eq. (16) is therefore analogous to Eq. (3).

2.3 Radiation hydrodynamics and the radiative transfer equation

The *radiation hydrodynamic equations* in the Eulerian formulation (e.g. Vogelsberger et al. 2020),

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0 \tag{18}$$

$$\frac{\partial(\rho\mathbf{v})}{\partial t} + \nabla\cdot(\rho\mathbf{v}\otimes\mathbf{v} + P\,\mathbf{I}) = \Gamma_p \tag{19}$$

$$\frac{\partial(\rho e)}{\partial t} + \nabla\cdot\Big((\rho e + P)\,\mathbf{v}\Big) = \Gamma_e - \Lambda \ , \tag{20}$$

formally follow from the Euler equations of pure hydrodynamics (Eqs. (4)-(6); or, equivalently, Eqs. (7)-(9) in the Lagrangian formulation) through the addition of source and sink terms in the momentum and energy equations.
Specifically, including radiation introduces a new source term Γ_p into the momentum equations (Eq. 19), which captures the transfer of momentum from photons to the gas. In the energy equation (Eq. 20), Γ_e and Λ similarly describe the energy transferred into and out of the gas through its coupling to the radiation field; here, the sink term Λ is also referred to as the *cooling function* or *gas cooling rate* (e.g. Vogelsberger et al. 2020; Kannan et al. 2019).
In practice (e.g., in the context of numerical RHD simulations; see also Sec. 3.2), these source and sink terms are computed through detailed *thermochemistry* models, describing the mechanisms through which photons of specific energies couple to specific chemical species such as H I, H II, He I, He II, and He III (for details, see e.g. Kannan et al. 2019, 2022a).

Crucially, an accurate treatment of RHD requires an accurate theoretical description of the dynamics of the radiation field itself, i.e., *radiative transfer (RT)*. In the following, we therefore present a brief derivation of the most general RT equation in an expanding space-time.

First, the total differential of the photon distribution function $f_\gamma = f_\gamma(\mathbf{x}, \mathbf{p}, t)$ in comoving coordinates, characterizing the phase-space number density of photons, is given by

$$\mathrm{d}f_\gamma = \frac{\partial f_\gamma}{\partial t}\,\mathrm{d}t + \sum_{i=1}^{3} \frac{\partial f_\gamma}{\partial x_i}\,\mathrm{d}x_i + \sum_{i=1}^{3} \frac{\partial f_\gamma}{\partial p_i}\,\mathrm{d}p_i \;. \tag{21}$$

Here, t denotes the time, while $\mathbf{x}$ is the position in *comoving* coordinates, related to the physical position $\mathbf{x}_\mathrm{phys}$ as $\mathbf{x} = \mathbf{x}_\mathrm{phys}/a$, where $a = a(t)$ is the cosmological scale factor. Similarly, $\mathbf{p}$ denotes the comoving photon momentum, given by

$$\mathbf{p} = a\,\mathbf{p}_\mathrm{phys} = a\,\frac{h\nu}{c}\,\mathbf{n} \;, \tag{22}$$

where $\mathbf{p}_\mathrm{phys}$ is the physical momentum, h is Planck's constant, ν is the photon frequency, c is the speed of light, and $\mathbf{n}$ is the unit vector parallel to $\mathbf{p}$, i.e., the unit vector in the direction of photon propagation.

In comoving phase space $(\mathbf{x}, \mathbf{p})$, the distribution function f_γ of photons is relativistically invariant (Mihalas & Weibel-Mihalas 1984, pp. 152ff., 311f.), i.e., only changes with time in the presence of photon sources and sinks (absorbers). Consequently, the following continuity equation (e.g. Gnedin & Ostriker 1997; Petkova & Springel 2009) holds for f_γ:

$$\frac{\mathrm{d}f_\gamma}{\mathrm{d}t} = \frac{\partial f_\gamma}{\partial t} + \sum_{i=1}^{3} \dot{x}_i \frac{\partial f_\gamma}{\partial x_i} + \sum_{i=1}^{3} \dot{p}_i \frac{\partial f_\gamma}{\partial p_i} = \left.\frac{\partial f_\gamma}{\partial t}\right|_\mathrm{sources} - \left.\frac{\partial f_\gamma}{\partial t}\right|_\mathrm{sinks} \;. \tag{23}$$

While Eq. (23) describes the evolution of the radiation field from a *microscopic* point of view, i.e. in terms of the distribution function f_γ of individual photons, a complete *macroscopic* description can be provided by instead expressing it in terms of the *radiative specific intensity* I_ν (e.g. Mihalas & Weibel-Mihalas 1984, p. 311). Specifically, $I_\nu = I_\nu(\mathbf{x}, t, \mathbf{n}, \nu)$ describes the specific intensity of photons with frequencies in the infinitesimal range $[\nu, \nu + \mathrm{d}\nu]$ propagating in direction $\mathbf{n}$ at space-time coordinates $(\mathbf{x}, t)$. By definition,

$$\mathrm{d}E_\nu = I_\nu(\mathbf{x}, t, \mathbf{n}, \nu)\,(\mathbf{n} \cdot \mathrm{d}\mathbf{A})\,\mathrm{d}t\,\mathrm{d}\nu\,\mathrm{d}\Omega \tag{24}$$

then describes the rate of radiative energy from these photons flowing per area $\mathrm{d}\mathbf{A}$, in the time $\mathrm{d}t$, into the solid angle $\mathrm{d}\Omega$ centered around $\mathbf{n}$ (e.g. Mihalas & Weibel-Mihalas 1984, p. 311; Kannan et al. 2019). Consequently, the dimensions of I_ν in cgs units are $[I_\nu] = \mathrm{erg\,cm^{-2}\,s^{-1}\,Hz^{-1}\,sr^{-1}}$.

This macroscopic, or continuum, view in terms of I_ν can be translated to the microscopic view, and vice versa, using the relation (Mihalas & Weibel-Mihalas 1984, p. 312)

$$I_\nu(\mathbf{x}, t, \mathbf{n}, \nu) = \frac{h^4\nu^3}{c^2}\,f_\gamma(\mathbf{x}, t, \mathbf{p} = \mathbf{p}(\mathbf{n}, \nu)) \;. \tag{25}$$

Therefore, by inserting $f_\gamma = \frac{c^2}{h^4}\frac{I_\nu}{\nu^3}$ from Eq. (25) into Eq. (23) and substituting $\mathbf{p} = \mathbf{p}(\mathbf{n}, \nu)$ according to Eq. (22), we obtain

$$\begin{aligned}
\frac{\mathrm{d}f_\gamma}{\mathrm{d}t} &= \frac{\partial f_\gamma}{\partial t} + \sum_{i=1}^{3} \dot{x}_i \frac{\partial f_\gamma}{\partial x_i} + \dot{\nu}\frac{\partial f_\gamma}{\partial \nu} \\
&= \frac{c^2}{h^4}\left[\frac{1}{\nu^3}\frac{\partial I_\nu}{\partial t} + \sum_{i=1}^{3} \dot{x}_i \frac{1}{\nu^3}\frac{\partial I_\nu}{\partial x_i} + \dot{\nu}\Big(- 3\nu^{-4} I_\nu + \frac{1}{\nu^3}\frac{\partial I_\nu}{\partial \nu}\Big)\right] \\
&= \frac{c^2}{h^4\nu^3}\left[\frac{\partial I_\nu}{\partial t} + \sum_{i=1}^{3} \dot{x}_i \frac{\partial I_\nu}{\partial x_i} + \frac{\dot{\nu}}{\nu}\Big(- 3\frac{\partial I_\nu}{\partial \nu} + \nu\frac{\partial I_\nu}{\partial \nu}\Big)\right] \\
&= \frac{c^2}{h^4\nu^3}\left[\frac{\partial I_\nu}{\partial t} + \sum_{i=1}^{3} \dot{x}_i \frac{\partial I_\nu}{\partial x_i} - H\Big(\nu\frac{\partial I_\nu}{\partial \nu} - 3I_\nu\Big)\right] \\
&\stackrel{!}{=} \frac{c^2}{h^4\nu^3}\left(\left.\frac{\partial f_\gamma}{\partial t}\right|_{\text{sources}} - \left.\frac{\partial f_\gamma}{\partial t}\right|_{\text{sinks}}\right)
\end{aligned} \tag{26}$$

In the second to last step, the photon frequency's scaling with cosmological expansion,

$$\begin{aligned}
&\nu \propto a^{-1} \\
\Rightarrow\ &\frac{\dot{\nu}}{\nu} = -\frac{\dot{a}}{a} = -H\ ,
\end{aligned} \tag{27}$$

where H denotes the Hubble parameter, has been inserted (e.g. Mo et al. 2010, p. 108). Following conventional notation (see e.g. Petkova & Springel 2009; Kannan et al. 2019), we define

$$j_\nu = \frac{1}{c}\left.\frac{\partial I_\nu}{\partial t}\right|_{\text{sources}} \tag{28}$$

as the emission coefficient, and

$$\kappa_\nu \rho I_\nu = \frac{1}{c}\left.\frac{\partial f_\gamma}{\partial t}\right|_{\text{sources}} \tag{29}$$

as the absorption term proportional to I_ν, where κ_ν is the absorption coefficient and ρ the gas density.
Further, we note that

$$\sum_{i=1}^{3} \dot{x}_i \frac{\partial I_\nu}{\partial x_i} = \dot{\mathbf{x}} \cdot \nabla I_\nu = \frac{c\,\mathbf{n}}{a} \cdot \nabla I_\nu \tag{30}$$

for a photon traveling with speed c in the direction of the unit vector $\mathbf{n}$ in comoving coordinates.

By inserting Eqs. (28)-(30) into Eq. (26), the *general radiative transfer equation* (cf. e.g. Gnedin & Ostriker 1997; Petkova & Springel 2009) follows:

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \frac{\mathbf{n}}{a} \cdot \nabla I_\nu - \frac{H}{c}\left(\nu\frac{\partial I_\nu}{\partial \nu} - 3I_\nu\right) = j_\nu - \kappa_\nu \rho I_\nu\ . \tag{31}$$

The terms contained in Eq. (31) describe (i) source and sink terms of radiation, (ii) radiation transport, and (iii) cosmological effects on radiation.
Firstly, the *source and sink terms* – i.e., photon emission and absorption terms – are given by the right-hand side of Eq. (31), specified by the coefficients j_ν and κ_ν as described above.
Secondly, the first two terms on the left hand side of Eq. (31) describe pure photon transport, i.e., the evolution of the 'isolated' radiation field in the absence of photon emission and absorption, or cosmological effects. This is commonly expressed explicitly in the *photon transport equation* (e.g. Kannan et al. 2019),

$$\frac{\partial I_\nu}{\partial t} = -\dot{x}_i \frac{\partial I_\nu}{\partial x_i} \, . \tag{32}$$

Thirdly, the *cosmological terms* of the RT equation are the terms proportional to the Hubble parameter H in Eq. (31). They themselves consist of (a) the 'photon dilution term' $\propto 3I_\nu$, and (b) the cosmological redshifting term $\propto \nu \frac{\partial I_\nu}{\partial \nu}$. The former captures the reduction of the photon number density $\mathcal{N}$ caused by the expansion of space, scaling as $\mathcal{N} \propto a^{-3}$ due to all spatial volumes V expanding as $V \propto a^3$. The latter describes the decrease in energy *each* photon experiences through cosmological redshifting, in addition to the overall reduction in photon number density. In the context of the continuum RT description, this is essentially captured by 'shifting' the spectrum $I_\nu(\nu)$ towards lower frequencies ν in an expanding Universe ($H > 0$), consequently modifying the specific intensity I_ν at each fixed ν according to the local spectral slope $\frac{\partial I_\nu}{\partial \nu}$.

We note that in practice, Eq. (31) is frequently simplified by discarding the redshifting term due to its complex dependence on the radiation field's spectrum, particularly in the context of cosmological simulation codes including numerical radiative transfer (discussed further in Sec. 3.2; see e.g. Gnedin & Ostriker 1997; Gnedin & Abel 2001; Petkova & Springel 2009; Kannan et al. 2019). The limitations of this approximation, as well as numerical methods addressing cosmological redshifting, are discussed in detail below (Secs. 3.4 and 3.5).

3 Numerical methods and implementations

The following subsections describe the numerical methods utilized and developed in the context of this thesis. First, the relevant established methods of simulating galaxy and large-scale structure formation, as well as radiation fields and reionization, are summarized, focusing primarily on the specific methods and code used in this thesis. Subsequently, the newly implemented methods modeling high-redshift X-ray sources and cosmological redshifting, as well as further relevant code changes, are described in detail.

3.1 Cosmological simulations of structure formation and evolution

Cosmological simulations have been an indispensable tool in modern astrophysics research over the past decades, in particular to investigate the formation and evolution of galaxies and large-scale structures in the Universe (for recent reviews, see e.g. Naab & Ostriker 2017; Vogelsberger et al. 2020). Specifically, they are uniquely capable of modeling the non-linear evolution of structure in the universe – for which no generally applicable analytic models exist – while capturing physical processes on a wide range of time and length scales.

Modern simulations typically start out from high-redshift initial conditions (ICs) constrained by observations (e.g., Planck Collaboration et al. 2016, 2020a; see Sec. 3.1.1). These ICs are subsequently evolved by modeling the dynamics and interactions of dark matter (Sec. 3.1.2) and baryonic matter (Sec. 3.1.3) in an expanding Universe.

Particularly in the early days of the field, as well as in modern simulations capturing extremely large volumes ($\mathcal{O}(\mathrm{cGpc}^3)$ scales) of the Universe (e.g. Garrison et al. 2018; Maksimova et al. 2021; Hernández-Aguayo et al. 2023), dark matter-only simulations (so-called N-body simulations, see Sec. 3.1.2) with no baryonic component have been employed to save substantial amounts of computing resources (e.g. Springel et al. 2005a; Boylan-Kolchin et al. 2009; Angulo et al. 2012).

On the other hand, hydrodynamical simulations additionally model baryonic matter and the wealth of physical processes associated with this component. In cosmological simulations of structure formation, this is encompassed in a typically code-specific galaxy formation model (GFM), numerically capturing key processes such as gas cooling, star formation, or stellar and AGN feedback (see Sec. 3.1.3 for further details).

While incorporating the baryonic component and a wide range of physics models allows for substantially more accurate and detailed predictions, it also represents the most challenging and computationally expensive part of galaxy formation simulations due to the inherent multiphysics nature, large dynamical range, and frequently supersonic gas velocities in the processes involved (e.g. Vogelsberger et al. 2020).

However, rapid progress in the field has been driven by a number of factors, such as (i) the advancement of numerical methods and algorithms either applicable to, or even explicitly tailored for, galaxy formation simulations, (ii) continuous improvements in our understanding and models of an enormous range of physical processes at the heart of galaxy formation, such as those captured in specific GFMs, and (iii) technological advancements, particularly the substantial performance increases of high-performance computing (HPC) architectures in recent years (such as exascale supercomputers and powerful GPU accelerators; see e.g. Zier et al. 2024).

Leveraging these factors, modern simulations have been highly successful in reproducing a wide range of individual galaxy and galaxy population properties inferred from observations (see Sec. 3.1.3).
By using different theoretical models as inputs and comparing simulation results to the wealth of observational data provided by modern surveys, simulations can further serve as a powerful probe of theoretical models. In this way, they have played a crucial role in e.g. (i) constraining dark matter models, such as self-interacting dark matter (e.g. Vogelsberger et al. 2012; Rocha et al. 2013; Peter et al. 2013), atomic dark matter (e.g. Roy et al. 2025), or fuzzy dark matter (e.g. Nori & Baldi 2018); (ii) probing the nature of dark energy (e.g. Garrison et al. 2018; Maksimova et al. 2021; Grove et al. 2022; Ding et al. 2022; Yuan et al. 2022; Shen et al. 2024); and (iii) testing different hypotheses for alternative cosmological models, such as early dark energy (e.g. Shen et al. 2024) or cosmic string models (e.g. Jiao et al. 2024; Koehler et al. 2024).
Consequently, cosmological galaxy formation simulations have been established as key tools for providing both theoretical insights and interpretations of observational data.

3.1.1 Initial conditions

Serving as the starting point for numerical simulations, initial conditions (ICs) specify the distribution of dark matter and baryons at a high initial redshift (typically $z_{\rm initial} \sim 100$, e.g. Vogelsberger et al. 2014a; Hopkins et al. 2018; Naiman et al. 2018; Vogelsberger et al. 2020; Springel et al. 2021; Maksimova et al. 2021).
Broadly speaking, two key types of ICs are usually employed in the context of cosmological simulations: (i) ICs sampling a larger spatial volume *uniformly* and with periodic boundary conditions – capturing the large-scale homogeneity and isotropy of the Universe based on the cosmological principle – and (ii) so-called *zoom ICs* sampling a central high-resolution region (e.g., a single galaxy or dark matter halo) embedded in a lower-resolution background capturing surrounding density and temperature fields.
The former are particularly well-suited to modern large-volume ($\mathcal{O}(\mathrm{cMpc}^3) - \mathcal{O}(\mathrm{cGpc}^3)$) simulations studying large-scale structure formation and galaxy and dark matter halo population statistics (e.g. Springel et al. 2005a; Boylan-Kolchin et al. 2009; Angulo et al.

2012; Vogelsberger et al. 2014a; Springel et al. 2018; Hadzhiyska et al. 2021; Kannan et al. 2023), as well as cosmological models (e.g. Rocha et al. 2013; Maksimova et al. 2021) and cosmological-scale physical processes such as reionization (e.g. Kannan et al. 2022a,b; Smith et al. 2022; Garaldi et al. 2024).
Conversely, simulations based on zoom ICs can leverage the typically substantially smaller volume of the region of interest to numerically resolve it in significantly more detail (e.g. Springel et al. 2008; Hopkins et al. 2014, 2018, 2023; Kannan et al. 2025; Zier et al. 2025b; Shen et al. 2025). Notably, this allows highly accurate simulations of individual dark matter halos and galaxies, capturing their (sub-) structure through fine spatial and mass resolutions and frequently incorporating more detailed smaller-scale physics models, such as a multi-phase ISM (e.g. Kim & Ostriker 2017; Hopkins et al. 2018; Kannan et al. 2025) and more sophisticated AGN models (e.g. Bourne & Sijacki 2017; Fiacconi et al. 2018; Talbot et al. 2021, 2024).

ICs based on the concordance cosmological model, i.e., a ΛCDM cosmology with an inflationary phase after the Big Bang, are typically determined by superimposing Gaussian density fluctuations on a homogenous background described by a spatially flat (Ω_k=0) Friedmann-Lemaître-Robertson-Walker (FLRW) metric (see e.g. Baumann 2022 for details). The initially homogenous distribution of the simulation's dark matter and gas particles is practically achieved by initializing particle positions uniformly on a *Cartesian lattice* or by a so-called *gravitational glass* initialization. The latter is realized by setting up particle positions randomly and subsequently displacing them by simulating the effects of a repellent, effectively opposite gravitational force between particles until equilibrium is achieved, i.e., until each particle 'freezes' in comoving coordinates (White 1996).
The specific makeup of this homogenous background, described by the ratios between dark matter, dark energy, and baryonic matter, is generally set by observational constraints on the energy and matter content of the universe, i.e., the cosmological parameters Ω_m, Ω_Λ, and Ω_b (see e.g. Planck Collaboration et al. 2016, 2020b for further details). Similarly, the relative amplitudes and distribution of the Gaussian density fluctuations, described fully by the matter power spectrum $P(k)$ as a function of the wavenumber k, are typically constrained by precision observations of CMB anisotropies (e.g. Planck Collaboration et al. 2020a).
The gravitational effect of the resulting overdensities on the simulation's matter particles are captured by a corresponding shift in the initial particle positions and velocities. Specifically, these displacements are commonly calculated using the Zel'dovich approximation (Zel'dovich 1970; see also Klypin & Shandarin 1983a) or second-order Lagrangian perturbation theory (2LPT; see Scoccimarro 1998).

All simulations presented in this work are based on large-volume, uniform (non-zoom)

ICs sampling cubic boxes with periodic boundary conditions. To generate these ICs, the `NGEN-IC` code within GADGET-4 (Springel et al. 2021), a recent update to the GADGET cosmological simulation code base (Springel et al. 2001; Springel 2005), is employed. We further use a gravitational glass initialization for the larger-scale, more computationally expensive X-ray test runs presented in Sec. 4.1 and a simple uniform Cartesian lattice initialization for smaller test runs (Secs. 4.2 and 4.3). Additionally, all runs employ 2LPT to determine the initial particle positions and velocities. Finally, all ICs utilized in this thesis are based on the Planck Collaboration et al. (2016) or Planck Collaboration et al. (2020b) initial matter power spectrum and cosmological parameters, as specified in the context of each test run setup in Sec. 4.

3.1.2 Modeling dark matter

Analytically, the dynamics of collisionless dark matter are described in the continuum limit by the collisionless Boltzmann equation (CBE; Eq. (2)) coupled to the Poisson equation (Eq. (3)), cf. Sec. 2.2.1. However, solving the CBE numerically has presented a considerable challenge in the past, as its high dimensionality in the dark matter distribution function, $f = f(\mathbf{x}, \mathbf{v}, t)$, prevents a straightforward application of typical numerical discretization techniques aiming to solve partial differential equations (PDEs; e.g. Springel et al. 2001; Vogelsberger et al. 2020). Consequently, alternative methods have been developed in order to accurately model dark matter dynamics; of these, the most widely used example is the *N-body method*, consequently serving as the namesake of 'N-body', i.e. dark matter-only, simulations.

This method samples the continuous dark matter distribution function f by a discrete set of N tracer particles with masses m_i at phase-space coordinates $(\mathbf{x}_i, \mathbf{v}_i)$, $i = 1, \ldots, N$. The conservation of the local dark matter phase-space density expressed by the CBE (Sec. 2.2.1) is realized by conserving each of the tracer particle masses m_i along their respective trajectories.
Thereby, the N-body approach uses a *method of characteristics* technique to solve the CBE, i.e., the equation is solved along curves on which the original PDE (Eq. 2)) reduces to an ordinary differential equation (ODE; e.g. Vogelsberger et al. 2020).
Further, the N-body method essentially represents a *Monte Carlo* approach to solving the CBE, as the ensemble of discrete tracer particles sample the phase-space density N times and can therefore be interpreted as a coarse-grained phase-space density, $\langle f \rangle \approx \sum_i m_i \, f\left(\mathbf{x}_i(t), \mathbf{v}_i(t)\right)$. As such, a sufficiently large number of particles N is essential to reduce Poisson noise and obtain an accurate simulation of dark matter dynamics (Springel et al. 2001; Vogelsberger et al. 2020).

At small separations, a simple discretization of the originally smooth dark matter distri-

bution function can give rise to unphysical large-angle two-body scatterings between the individual sampling particles. To avoid this numerical artifact, a *gravitational softening* term is typically introduced, smoothing the particles' gravitational potential on small scales. This is commonly implemented numerically through kernel-based methods (for further details, see e.g. Hernquist & Katz 1989; Springel 2005; Price & Monaghan 2007; Springel 2010b).
The spatial scale of this smoothing is characterized by the *gravitational softening length*, which, in practice, is typically set as a code parameter (e.g. Springel 2010b; Springel et al. 2021) depending on the average particle density in the simulation volume (e.g. Springel et al. 2001). The latter is also commonly referred to as a simulation's *effective resolution*, describing the number of resolution elements – i.e., dark matter particles in the context of N-body simulations – per unit volume.
Due to the 'blurring' of the gravitational potential, larger gravitational softening lengths inherently reduce the spatial resolution of the calculated gravitational forces. To mitigate this effect, *adaptive softening schemes* have been introduced, which reduce the softening length in high-density regions within the simulation volume (Price & Monaghan 2007).

The computational problem at the heart of N-body simulations is calculating the gravitational forces exerted by and acting on each of the simulation particles, in order to determine their dynamical evolution through each time step of the simulation. Standard numerical methods addressing this problem can typically be classified as aiming to solve either the *integral* or *differential* form of the Poisson equation, or alternatively, as employing a combination of the latter two methods in a *hybrid* approach.
The first group of methods approaches the gravitational force calulation by solving the integral form of Poisson's equation,

$$\Phi(\mathbf{x}) = -G \int \mathrm{d}^3\mathbf{x}' \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \,. \tag{33}$$

Here, $\Phi(\mathbf{x})$ and $\rho(\mathbf{x})$ denote the gravitational potential and dark matter density at position $\mathbf{x}$, respectively, and G denotes Newton's constant. The *particle-particle scheme* solves Eq. (33) directly for the N tracer particles of the dark matter density, effectively resulting in a direct summation problem with computational complexity $\mathcal{O}(N^2)$ (e.g. Vogelsberger et al. 2020). This numerically expensive 'brute-force' approach of early N-body simulations (e.g. Peebles 1970; Press & Schechter 1974; White 1976; see also Holmberg 1941 for one of the earliest 'analog' uses of this method) can be substantially accelerated through the use of *tree algorithms* (Appel 1981, 1985; Barnes & Hut 1986). Rather than computing and summing up the gravitational force exerted by every individual particle to obtain the net force at a given point, this approach approximates the gravitational effects of more distant particles by treating them as a group. Specifically, particles are grouped in

a tree-like hierarchical structure, typically in a so-called *octree*. The latter structure is achieved by evenly dividing a cubic simulation volume – simultaneously serving as the root node of the tree – into eight cubic child cells of equal volume, and in turn iteratively subdividing each of the child cells into another eight cubic cells. This pattern is usually repeated until exactly one or zero simulation particles are left in the finest subdivision, i.e., the leaves of the tree; this is illustrated schematically in two dimensions in Fig. 4.
The above-mentioned particle groups are then defined as all simulation particles within a specific node of the resulting tree. Leveraging this structure, the computational speed-up is achieved by approximating the gravitational forces of distant particle groups on a given particle by using multipole expansions at a coarser level of the tree, resulting in an $\mathcal{O}(N \log N)$ complexity (Appel 1985; Barnes & Hut 1986; Springel 2010b).

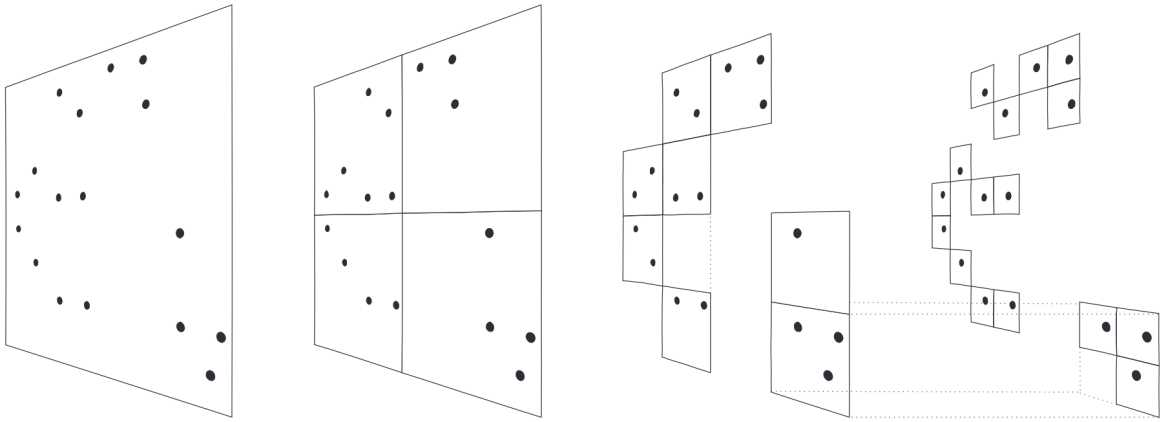

Figure 4: Schematic visualization of the Barnes & Hut (1986) octree subdivision of space, represented in two dimensions. The root node enclosing all particles, shown on the left, is iteratively subdivided until the final child nodes only contain exactly one particle each, as shown on the right. Empty nodes containing no particles are discarded and not further subdivided in the process. This figure was obtained from Springel et al. (2001).

Serving as a fundamentally different approach, the second group of methods to compute the gravitational forces in N-body simulations instead aims to solve the differential form of Poisson's equation (Eq. (16)).
Typically, this is achieved through the use of *mesh-based methods*. In the simplest case, the mesh is essentially a Cartesian grid, i.e., a lattice-like structure discretizing the simulated spatial domain into cubes of equal volume. In the *particle-mesh method* (Hockney & Eastwood 1981; Klypin & Shandarin 1983b), the gravitational potential acting on each of the simulation particles is then determined by solving Eq. (16) in Fourier space,

$$k^2 \tilde{\Phi}(\mathbf{k}) = -4\pi G \tilde{\rho}(\mathbf{k}) \; . \tag{34}$$

Here, $\mathbf{k}$ is the wavevector and $k = |\mathbf{k}|$ the wavenumber, while $\tilde{\Phi}$ and $\tilde{\rho}$ denote the Fourier-transformed gravitational potential and Fourier-transformed density, respectively.

Notably, a Cartesian mesh allows a numerically highly efficient solution of Eq. (34), as it enables the use of fast Fourier transform (FFT) algorithms. Provided this structure, the approach leads to a computational complexity of $\mathcal{O}(N \log N)$ of the particle-mesh method, where N represents the number of cells defined by the mesh (e.g. Vogelsberger et al. 2020). As an alternative to the FFT approach using a Cartesian grid, so-called *multi-grid* or *adaptive mesh refinement* (*AMR*; Kravtsov et al. 1997) schemes are also commonly employed (e.g. O'Shea et al. 2004; Bryan et al. 2014) in the context of the particle-mesh method. In this case, the mesh is constructed to have an octree-like structure, allowing a finer mesh spacing (i.e., increased numerical resolution) in key areas of interest, such as high-density regions within the simulation volume (e.g. Vogelsberger et al. 2020).

Finally, *hybrid* gravity solvers combine aspects of both the methods solving the integral form of Poisson's equation and those solving its differential form, i.e., mesh-based methods. Essentially, these hybrid approaches aim to strike an optimal balance between computational efficiency and accuracy of the force calculation by employing direct summation techniques (e.g., particle-particle schemes and tree-based algorithms) to model short-range gravitational interactions, while using mesh-based methods (e.g., FFT-based particle-mesh method) to model long-range interactions (Efstathiou et al. 1985).
One of the most widely used (see e.g. Bode & Ostriker 2003; Springel 2005; Hopkins 2015, 2017; Springel et al. 2021) hybrid methods is the *tree particle-mesh* (*TreePM*; Xu 1995; Bagla 2002) approach, which combines the tree and particle-mesh methods described above to compute short- and long-range gravitational forces, respectively. Notably, this is the approach taken by the AREPO code (Springel 2010b; Pakmor et al. 2016) for cosmological simulations and hence, the approach utilized to numerically model dark matter dynamics throughout this thesis.

3.1.3 Modeling baryonic matter

One of the key baryonic components of cosmological simulations is astrophysical gas, typically modeled as an ideal fluid governed by the Euler equations discussed in Sec. 2.2.2. In light of the two different formulations of these equations (see Eqs. (4)-(6) and Eqs. (7)-(9)), numerical schemes modeling astrophysical gas can also typically be classified as either *Eulerian* or *Lagrangian* approaches; additionally, so-called *arbitrary Lagrangian-Eulerian* numerical methods combine the latter two approaches.

As is the case for the analytical description, numerical Eulerian methods utilize a spatially fixed frame of reference to solve the Euler equations. In practice, this results in *grid-based* codes, which solve these hyperbolic PDEs on a Cartesian fixed or adaptive mesh (for a recent review, see Teyssier 2015). In the latter case, the resolution of the mesh can be refined in specific regions based on a set refinement criterion, such as the density of the

gas. This approach is therefore an example of an AMR scheme (Berger & Oliger 1984; Berger & Colella 1989), which – aside from its uses in the context of N-body simulations (see Sec. 3.1.2) – is commonly employed in Eulerian hydrodynamical codes to accurately capture the large dynamical range of cosmological simulations (e.g. Bryan & Norman 1997; Teyssier 2002; O'Shea et al. 2004; Bryan et al. 2014; Moseley et al. 2023; Andalman et al. 2024).

Conversely, numerical Lagrangian methods take the point of view of an observer 'comoving' with the fluid, again analogously to their analytical counterpart, i.e., the Lagrangian formulation of the Euler equations. To solve the latter, the typical method of choice is *smoothed particle hydrodynamics* (*SPH*; Lucy 1977; Gingold & Monaghan 1977; see also e.g. Hernquist & Katz 1989; Springel et al. 2001; Springel 2010a). This method models the fluid through discrete sampling particles obeying Eqs. (7)-(9). The hydrodynamical properties, such as the fluid density at a given position $\mathbf{x}$, are computed by smoothing over nearby particles (e.g. van den Bosch 2020), resulting in a mesh-free, *particle-based* approximate fluid description.
In this scheme, the number density of sampling particles therefore automatically follows the flow and clumping of gas mass throughout the simulation volume, effectively increasing the numerical resolution in areas of high density without the need for explicit refinement schemes. This represents a substantial advantage of Lagrangian methods for capturing large density contrasts, which commonly occur in the context of cosmological simulations (e.g. Springel 2010a).

Finally, an arbitrary Lagrangian-Eulerian method has been developed for astrophysical applications in the AREPO code used throughout this thesis (Springel 2010b). Rather than employing a static Cartesian mesh, as commonly used in Eulerian codes, or the mesh-free 'comoving' approach of SPH codes, AREPO solves the hydrodynamic equations on a *moving mesh* made up of unstructured cells defined by a *Voronoi tessellation*, each enclosing a finite volume of space. Formally, AREPO is therefore classified as a *moving-mesh finite-volume (MMFV)* code (Vogelsberger et al. 2020).
A Voronoi tessellation is a unique subdivision of space into convex, non-overlapping cells based on a specified set of *mesh-generating points*. Each Voronoi cell is constructed around a specific mesh-generating point, and defined as all points in space which are closer to this mesh-generating point than to any of the other mesh-generating points. The set of all cells defined by the Voronoi tessellation then form a generally unstructered, irregularly shaped mesh. An example of such a Voronoi tessellation in two dimensions is shown in Fig. 5.
The mesh-generating points are allowed to move freely in AREPO, and the Voronoi mesh is reconstructed according to their new positions in each hydrodynamical timestep of a simulation.

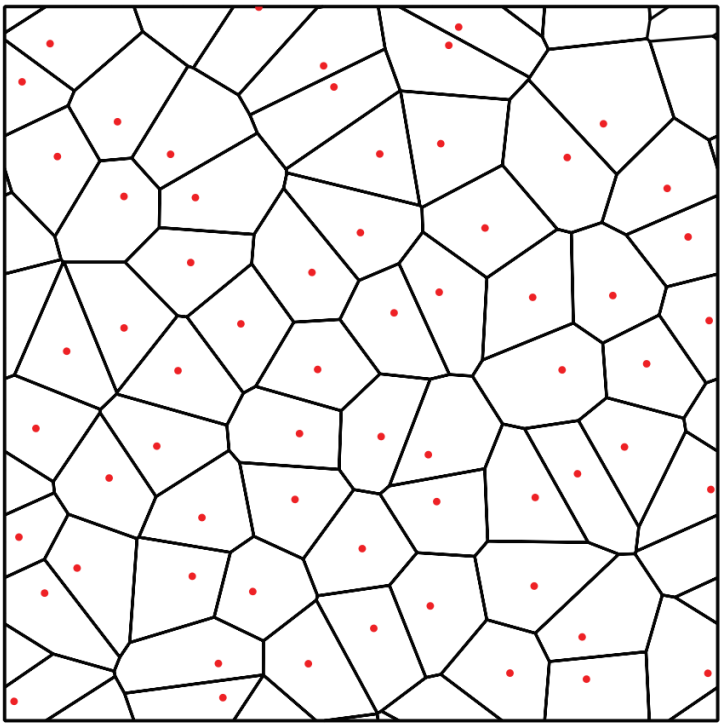

Figure 5: The two-dimensional Voronoi tessellation for an example set of $N = 64$ mesh-generating points (shown as red points), assuming periodic boundary conditions. This figure was obtained from Springel (2010b).

Formulated with respect to this arbitrarily moving mesh made up of finite-volume cells, the Euler equations then take the following form for each cell:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho\,\mathrm{d}V = -\int_S \rho\,(\mathbf{v}-\mathbf{w})\cdot\mathbf{n}\,\mathrm{d}S \tag{35}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho\mathbf{v}\,\mathrm{d}V = -\int_S \rho\mathbf{v}\,(\mathbf{v}-\mathbf{w})\cdot\mathbf{n}\,\mathrm{d}S - \int_S P\mathbf{n}\,\mathrm{d}S \tag{36}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho e\,\mathrm{d}V = -\int_S \rho e\,(\mathbf{v}-\mathbf{w})\cdot\mathbf{n}\,\mathrm{d}S - \int_S P\mathbf{v}\cdot\mathbf{n}\,\mathrm{d}S\,. \tag{37}$$

Here, V is the volume of the cell, S the area of the surface enclosing the volume V, $\mathbf{n}$ the normal vector on the surface S, and $\mathbf{w}$ the local velocity of the mesh.

The approach described above shares conceptual similarities with both Eulerian and Lagrangian codes, and can in fact be constrained to run in both Eulerian and Lagrangian modes.

Firstly, analogously to Eulerian codes, AREPO takes a *finite-volume discretization* approach, i.e., it numerically solves the Euler equations by discretizing space into cells, each with a finite volume and well-defined hydrodynamical properties averaged over the cell. Specifically, it utilizes a second-order unsplit Godunov scheme (Godunov 1959) and an exact Riemann solver (see e.g. Springel 2010b; Teyssier 2015 for further details) to solve the Euler equations – a widely used approach in Eulerian schemes. Notably, this implies

that for the special case of fixed ($\mathbf{w} = 0$) mesh-generating points regularly spaced on a Cartesian grid, the MMFV approach is equivalent to a typical Eulerian finite-volume method on a static grid.

However, a Lagrangian mode of operation can also be achieved by allowing the mesh-generating points to move along with the fluid, i.e., by setting their velocity equal to the fluid velocity in each cell ($\mathbf{w} = \mathbf{v}$) in the simplest case (Springel 2010b). Similarly to the ordinary Lagrangian SPH approach, this essentially results in discrete points carrying the hydrodynamic and thermodynamic information of the fluid along with its flow. In turn, it inherits the crucial advantage of SPH schemes to automatically increase numerical resolution in regions of converging flows and therefore, ultimately, in higher-density regions. This is because, by definition, the Voronoi tessellation of space will result in a smaller cell size (i.e., higher resolution) in regions with a higher number density of mesh-generating points, as illustrated in Fig. 6.

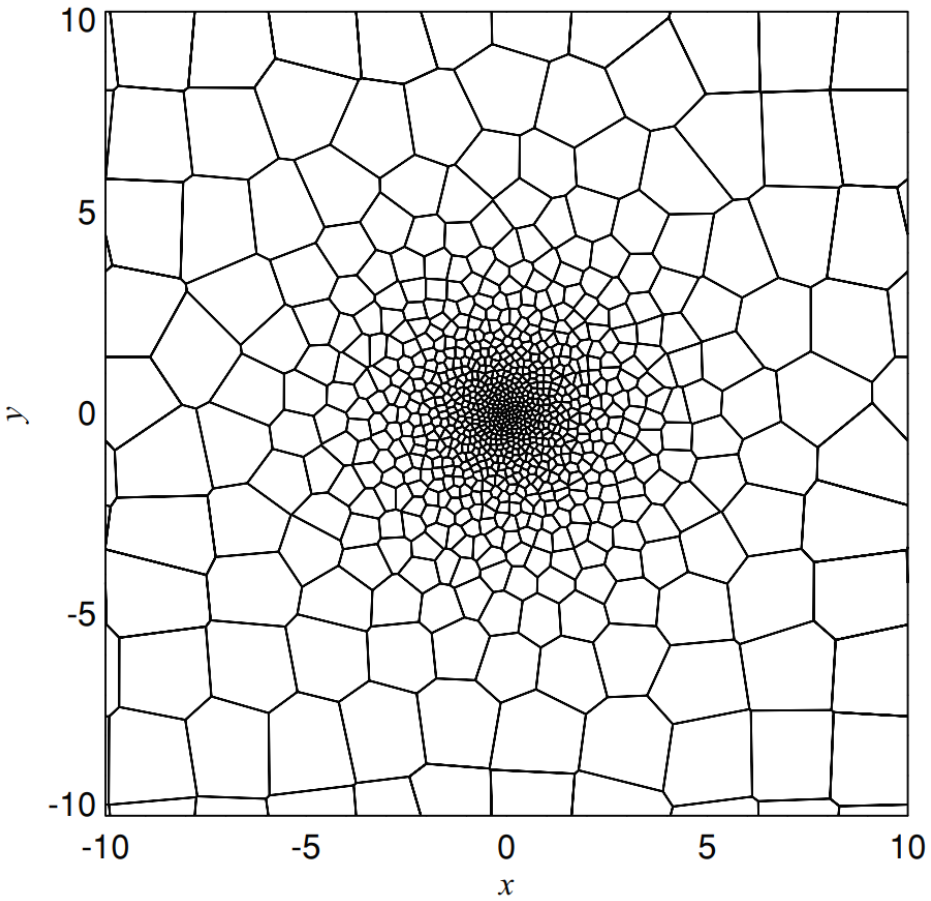

Figure 6: A two-dimensional, regularized Voronoi mesh for $N = 750$ particles distributed as an exponential disc centered on spatial coordinates $x = y = 0$ with scale radius $r_0 = 1.0$. The shown mesh structure illustrates the refinement of numerical resolution in regions with a high density of mesh-generating points in the moving-mesh finite-volume method of AREPO. This figure was obtained from Springel (2010b).

Besides sharing these numerical methods and advantages with both standard Eulerian and Lagrangian codes, the flexibility of the arbitrary Lagrangian-Eulerian MMFV approach allows the AREPO code to avoid well-known disadvantages of these standard methods. For instance, the MMFV approach is better equipped to capture hydrodynamic shocks and fluid instabilities, such as the Rayleigh-Taylor instability, than the Lagrangian, mesh-free SPH codes (see e.g. Agertz et al. 2007; Springel 2010b for further details).
Crucially, it also avoids Eulerian methods' lack of Galilean invariance, which introduces numerical uncertainties in the presence of large bulk velocities (e.g. Tasker et al. 2008; Wadsley et al. 2008). For example, in cosmological simulations of galaxy formation, it is common for galaxies to move with relative speeds which far exceed the sound speed of dense gas *within* each galaxy. This bulk velocity of a galaxy as a whole – with respect to the static Eulerian grid – can then introduce difficulties in resolving the fluid flow of gas within the galaxy, particularly if the latter are several orders of magnitude smaller than the former. These adverse effects of large bulk velocities can largely be eliminated by allowing the mesh-generating points and hence, the moving mesh itself, to follow the local fluid flow as described above (see also Springel 2010b).

Besides the hydrodynamical treatment of astrophysical gas, the GFM of a simulation crucially determines the numerical modeling of its baryonic component. It encompasses models of a broad range of physical processes relevant to the formation and evolution of galaxies, such as star formation and stellar feedback prescriptions, cooling processes of the astrophysical gas, SMBH seeding and growth, and AGN feedback, among many others (see e.g. Vogelsberger et al. 2020). In cosmological simulations encompassing large volumes of space (e.g., $\mathcal{O}(\mathrm{Mpc}^3)$), the majority of these processes physically occur on scales far below the resolution limit and therefore have to be approximated by effective models, referred to as *subresolution* or *subgrid* models. These empirically calibrated models are typically one of the greatest sources of uncertainty in modern cosmological simulations, and in turn one of the most significant points in which approaches and results diverge when comparing different simulation projects or codes (e.g. Vogelsberger et al. 2020).

Throughout this thesis, we use the *IllustrisTNG* (*Illustris 'The Next Generation'*; e.g. Springel et al. 2018; Nelson et al. 2019) GFM as a common basis for all runs.[3] As a state-of-the-art update to the previous Illustris model (Vogelsberger et al. 2013, 2014b,a; Torrey et al. 2014), the IllustrisTNG model has been extensively tested and successfully employed to reproduce numerous detailed properties of the observed galaxy population (e.g. Nelson et al. 2018, 2021; Springel et al. 2018; Marinacci et al. 2018; Naiman et al.

[3]We note, however, that we also include substantial extensions building on this core model, primarily through the inclusion of radiation hydrodynamics (see Kannan et al. 2019, 2022a), as well as through the code extensions developed in the context of this thesis. These modifications are discussed in detail in the sections below.

2018). For instance, it has been shown to produce galaxy sizes (Genel et al. 2018) and stellar mass functions (Pillepich et al. 2018) in excellent agreement with low-redshift observations, as well as a realisitic galactic color bimodality (Nelson et al. 2018) and accurate star-formation histories (Donnari et al. 2019). Some of its key numerical models are outlined in the following.

Firstly, the ISM is treated using a subresolution model describing a two-phase gas, which consists of a cold phase containing the majority of the ISM's gas mass in dense clumps, and a smooth, volume-filling hot phase made up of supernova-heated gas. In practice, the cold and dense phase is modeled through an effective, polytropic Equation of State (EoS) of the form $T \propto \rho^{\gamma}$, i.e., its temperature T is set by the gas density ρ and adiabatic index γ through the EoS (Springel & Hernquist 2003). This effective model is employed to overcome the technical challenge of modeling the cold gas directly, as its high density would otherwise require extremely small time steps to reliably capture its dynamics (e.g. Vogelsberger et al. 2020).
Further, star formation is treated through the stochastic conversion of cold ISM gas to stellar particles, constrained by the requirement that this cold gas meets or exceeds a set density threshold (e.g. Nelson et al. 2019). The resulting stellar particles are then treated as a collisionless fluid (cf. Sec. 2.2.1) and evolved dynamically through the use of a TreePM N-body approach (e.g. Springel 2010b; see also Sec. 3.1.2).
In turn, these stars also provide back-reactions to the surrounding gas, i.e., stellar feedback. In particular, the IllustrisTNG model includes the effects of supernovae and stellar winds, which are expected to inject both kinetic and thermal energy into the gas. Additionally, supernovae at the end of stars' life-cycles contribute to metal enrichment of the simulated gas (e.g. Nelson et al. 2019).
Specifically, the production and evolution of nine different chemical species (H, He, C, N, O, Ne, Mg, Si, and Fe) including seven metals, as well as the total metallicity of each gas cell, is tracked in the model. Metal line cooling through these heavier elements, as well as a range of other density-, temperature-, and redshift-dependent cooling mechanisms, are modeled accordingly (see e.g. Kannan et al. 2022a).
Finally, SMBHs at the centers of galaxies are also a key component of the GFM, as they can have a substantial impact on their host galaxy's properties and evolution – for instance, by playing a crucial role in regulating star formation in massive galaxies (e.g. Sijacki & Springel 2006; Hopkins et al. 2006; Madau & Dickinson 2014; Harrison et al. 2018). Due to their extremely compact nature, which is not feasible to spatially resolve directly in cosmological simulations, their effects are also included through effective subresolution models. Specifically, SMBH formation is modeled through a simple seeding prescription, which numerically creates a SMBH of a set black hole seed mass at the center of a halo, once the latter crosses a certain mass threshold. Subsequently, SMBH

growth is modeled through an idealized Bondi-Hoyle accretion model capped by the theoretical Eddington limit (see Weinberger et al. 2018). AGN feedback, i.e., injections of energy and momentum into the gas surrounding an accreting SMBH, is implemented through a subresolution model operating in either a quasar mode for AGN with a high mass accretion rate (Springel et al. 2005b), or in a radio mode for AGN with a low mass accretion rate (Weinberger et al. 2017). For a more detailed discussion of SMBH modeling and AGN feedback in the IllustrisTNG GFM, we refer to Weinberger et al. (2018).

To summarize, two key numerical tools underlying the cosmological simulations presented in this work are the MMFV code AREPO and the IllustrisTNG GFM. Both dark matter and stellar particles are treated as collisionless fluids and dynamically evolved using the TreePM approach to the N-body problem. Astrophysical gas is modeled as an ideal fluid and the Euler equations describing its dynamics are solved on a moving, unstructured Voronoi mesh in an Arbitrary Lagrangian-Eulerian approach. The GFM accounts for a wide range of physical processes relevant to galaxy formation and evolution, which are mostly implemented through the use of subresolution models.

3.2 Numerical radiative transfer and radiation hydrodynamic simulations

As mentioned in the previous section, the combination of the AREPO code and IllustrisTNG GFM as numerical ingredients has been extensively tested and successfully used in simulation projects such as the IllustrisTNG project and the more recent MillenniumTNG simulations (Kannan et al. 2023; Hernández-Aguayo et al. 2023; Pakmor et al. 2023). However, reliable investigations of key high-redshift epochs – in particular, the reionization of the Universe driven by high-energy radiation – require us to take these numerical models a step further. This next step, in the form of numerically treating *radiation hydrodynamics*, is described in detail in the following.

In the most general case, the dynamics of a radiation field embedded in an expanding space-time are theoretically described by the general RT equation (Eq. (31)) as discussed in Sec. 2.3. However, even the more straightforward *classical* RT equation,

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \mathbf{n}\cdot\nabla I_\nu = j_\nu - \kappa_\nu \rho I_\nu \ , \tag{38}$$

obtained by neglecting the cosmological terms proportional to the Hubble parameter H and setting $a = 1$ in Eq. (31), is quite challenging to solve numerically. This is in large part due to the problem's high dimensionality, requiring a discretization of not only the four space-time variables $(\mathbf{x}, t)$, but also the frequency variable ν and the two directional

angles specifying the unit vector **n** in the direction of photon propagation (e.g. Petkova & Springel 2009; Kannan et al. 2019; Zier et al. 2024).

Nevertheless, a number of numerical approaches to RT have been established. Firstly, *long characteristic ray-tracing* schemes (e.g. Abel et al. 1999; Abel & Wandelt 2002) cast and propagate rays from each source into the simulation volume, discretizing the radiation field in individual directions and solving the RT equation along each ray. For high angular resolutions, this method is remarkably accurate, but computationally expensive and challenging to parallelize efficiently (e.g. Kannan et al. 2019) in the context of HPC environments equipped with a large number of CPU cores.
Aiming to reduce this high computational demand by simplifying the problem and improving its numerical scalability, *short characteristic methods* (e.g. Whalen & Norman 2006; Petkova & Springel 2011; Jaura et al. 2020) have been developed. These schemes also use the ray-tracing approach described above; however, they only solve the RT equations along rays connecting nearby cells, rather than along rays cast through the entire simulation domain.

Serving as an alternative approach, *Monte Carlo methods* (e.g. Ciardi et al. 2001; Maselli et al. 2003; Semelin et al. 2007; Maselli et al. 2009) emit and propagate so-called *photon packets*, which are each representative of a large set of individual photons approximated to be traveling in the same direction (e.g. Dullemond 2017), and which sample the continuous radiation field. The evolution of each photon packet is determined by stochastically probing the photons' interaction lengths and scattering angles based on underlying probability density functions (Kannan et al. 2019). Similarly to the long characteristic ray-tracing approach, these methods tend to be highly accurate – provided a sufficient number N_γ of photon packets – on one hand, but numerically demanding on the other hand. In particular, numerical uncertainty is caused by the characteristic Poisson noise of Monte Carlo schemes, leading the signal-to-noise ratio to merely improve in proportion to $\sqrt{N_\gamma}$. This drives the need for a large N_γ to obtain an accurate statistical description of the radiation field at the cost of a high computational demand.
As a consequence, Monte Carlo methods are frequently only applied in RT *post-processing* (Vogelsberger et al. 2020), i.e., in RT calculations based on previously existing hydrodynamic simulations, rather than in *on-the-fly* RT. In the context of cosmological simulations, the latter refers to full RHD simulations that numerically account for both radiation and gas dynamics in each hydrodynamical time step, thereby capturing the co-evolution and coupling of the radiation fields with the gas according to the RHD equations (see Eqs. (18)-(20)) self-consistently.

Both ray-tracing and Monte Carlo schemes typically scale linearly with the number of

sources (e.g. Jaura et al. 2018; Vogelsberger et al. 2020), resulting in a prohibitively high computational cost of cosmological-scale RHD simulations with on-the-fly RT.
To still facilitate such simulations – which are crucial to e.g. accurately capture radiation fields and their co-evolution with galaxies during the EoR (e.g. Kannan et al. 2022a) – a third approach, *moment-based* RT, has been commonly employed in recent years (e.g. González et al. 2007; Petkova & Springel 2009; Rosdahl et al. 2013; Rosdahl & Teyssier 2015; Kannan et al. 2019, 2022a, 2025).
This class of methods solves only the first two angular moments of Eq. (38), i.e.,

$$\frac{\partial E_\nu}{\partial t} + \nabla \cdot \mathbf{F}_\nu = S_\nu - \kappa_\nu \, \rho \, E_\nu \ , \tag{39}$$

$$\frac{\partial \mathbf{F}_\nu}{\partial t} + c^2 \, \nabla \, \mathbb{P}_\nu = -\kappa_\nu \, \rho \, \mathbf{F}_\nu \ . \tag{40}$$

These *RT moment equations* are obtained by reducing the specific intensity I_ν to its first angular moments: specifically, the radiative *energy density* E_ν (zeroth angular moment) is defined as (e.g. Kannan et al. 2019)

$$cE_\nu = \int_{4\pi} I_\nu \, \mathrm{d}\Omega \ , \tag{41}$$

the *flux* $\mathbf{F}_\nu$ (first angular moment) as

$$\mathbf{F}_\nu = \int_{4\pi} \mathbf{n} I_\nu \, \mathrm{d}\Omega \ , \tag{42}$$

and the *radiation pressure tensor* $\mathbb{P}_\nu$ (second angular moment) as

$$c\mathbb{P}_\nu = \int_{4\pi} (\mathbf{n} \otimes \mathbf{n}) \, I_\nu \, \mathrm{d}\Omega \ . \tag{43}$$

Here, Ω denotes the solid angle; $\int_{4\pi} \mathrm{d}\Omega$ therefore represents an integration over all directions $\mathbf{n}$. Further, Eqs. (39) and (40) introduce the angular moments of the emission coefficient j_ν, where the zeroth moment defines the *source function* S_ν as

$$S_\nu = \int_{4\pi} j_\nu \, \mathrm{d}\Omega \ , \tag{44}$$

and the first moment vanishes for the assumption of an isotropic emission term (e.g. Gnedin & Abel 2001):

$$\int_{4\pi} \mathbf{n} j_\nu \, \mathrm{d}\Omega = j_\nu \int_{4\pi} \mathbf{n} \, \mathrm{d}\Omega = 0 \ . \tag{45}$$

The RT moment equations represent a *fluid* description of the radiation field (see Eqs. (39) and (40) in comparison to e.g. the hydrodynamical continuity and momentum equations, Eqs. (4) and (5)). Similarly to the Euler equations, they form an under-determined set of hyperbolic conservation equations, requiring an additional constraint to close the system

of equations (cf. the EoS in the hydrodynamical case; Sec. 2.2.2).
Providing this additional closure relation, the *Eddington tensor formalism* is a widely used approach in numerical RT (e.g. Vogelsberger et al. 2020). Here, the *Eddington tensor* $\mathbb{D}$ is defined as a proportionality tensor relating the energy density E_ν to the radiation pressure tensor $\mathbb{P}_\nu$ as (e.g. Kannan et al. 2019)

$$\mathbb{P}_\nu = E_\nu \, \mathbb{D} \; . \tag{46}$$

A wide range of methods to estimate $\mathbb{D}$ has been developed. For instance, the *flux-limited diffusion* (e.g. Alme & Wilson 1973; Krumholz et al. 2007) and *optically thin variable Eddington tensor* (*OTVET*; Gnedin & Abel 2001) approaches compute $\mathbb{D}$ by assuming that the material in between radiation sources and sinks is either completely optically thick (flux-limited diffusion) or thin (OTVET).
However, since large-volume RHD simulations aiming to realistically model the universe on cosmological scales are required to capture large gas density contrasts and, in turn, both optically thick and thin material, these approximations are commonly discarded in favor of the *M1 closure relation* (Levermore 1984; see also e.g. Aubert & Teyssier 2008; Rosdahl et al. 2013; Rosdahl & Teyssier 2015; Kannan et al. 2019, 2022a). This approach defines the Eddington tensor according to only local properties of the radiation field (E_ν and $\mathbf{F}_\nu$) within each gas cell, setting

$$\mathbb{D} = \frac{1-\chi}{2} \, \mathbb{I} + \frac{3\chi - 1}{2} \, \mathbf{n}_\mathbf{F} \otimes \mathbf{n}_\mathbf{F} \; . \tag{47}$$

Here,

$$\mathbf{n}_\mathbf{F} = \frac{\mathbf{F}_\nu}{|\mathbf{F}_\nu|} \tag{48}$$

describes the unit vector in the direction of the local flux $\mathbf{F}_\nu$. The parameter χ is defined as

$$\chi = \frac{3 + 4f^2}{5 + 2\sqrt{4 - 3f^2}} \; , \tag{49}$$

where

$$f = \frac{|\mathbf{F}_\nu|}{c \, E_\nu} \tag{50}$$

is the *reduced flux* f, i.e., the magnitude of the flux vector relative to its maximum magnitude given by the product of the signal speed c and local energy density E_ν (Kannan et al. 2019). We note that f is therefore required to obey the condition

$$0 \leq f \leq 1 \tag{51}$$

by definition.
Further, to illustrate how the definition of the Eddington tensor given by Eqs. (47)-(50)

encodes the direction of photon propagation, we note the special cases of (i) an isotropic radiation field with zero net flux of photons,

$$f = 0$$
$$\Rightarrow \ \chi = \frac{1}{3} \ \Rightarrow \ \mathbb{D} = \frac{2/3}{2}\,\mathbb{I} + 0 = \frac{1}{3}\,\mathbb{I}\,,$$

and (ii) a fully directional radiation field, with all net flux propagating in a single direction $\hat{\mathbf{z}}$,

$$\mathbf{n_F} = \frac{\hat{\mathbf{z}}}{|\hat{\mathbf{z}}|} = \hat{\mathbf{z}},\ f = 1$$
$$\Rightarrow \ \chi = 1 \ \Rightarrow \ \mathbb{D} = 0 + \mathbf{n_F} \otimes \mathbf{n_F} = \hat{\mathbf{z}} \otimes \hat{\mathbf{z}}\,.$$

Crucially, the computational cost of moment-based RT with the M1 closure is independent of the number and distribution of radiative sources, as the Eddington tensor only depends on local quantities as mentioned above. This makes the approach particularly well-suited to simulate galaxy- or even cosmological-scale volumes enclosing a large number of sources (e.g. Rosdahl & Teyssier 2015; Costa et al. 2018; Kannan et al. 2022a, 2025).

Therefore, this is the approach taken by the RHD code AREPO-RT (Kannan et al. 2019; Zier et al. 2024), the moment-based RT solver extending the MMFV code AREPO (Sec. 3.1.3), which forms the basis of the implementations developed in the context of this thesis.
AREPO-RT solves the RT equation using an operator split approach, i.e., it separately solves the components of Eq. (38) describing (i) source and sink terms and (ii) terms pertaining to pure radiation transport. The former are captured by the thermochemistry equations, determining the coupling of the radiation field to the gas of different chemical species as mentioned in Sec. 2.3. Crucially, this accounts for the processes of photoionization and photoheating, discussed in further detail in Sec. 3.3.6.
The latter are expressed by the moments of the photon transport equation (Eq. (32)),

$$\frac{\partial E_\nu}{\partial t} + \nabla \cdot \mathbf{F}_\nu = 0\,, \tag{52}$$
$$\frac{\partial \mathbf{F}_\nu}{\partial t} + c^2\,\nabla\,\mathbb{P}_\nu = 0\,, \tag{53}$$

which are solved using a finite-volume approach on AREPO's moving Voronoi mesh (Kannan et al. 2019; see also Sec. 3.1.3 and Springel 2010b).
Specifically, the operator splitting approach taken by AREPO-RT is a *Strang split scheme* (Strang 1968). Here, each RT timestep Δt_{RT} first updates the primitive RT variables, E_ν and $\mathbf{F}_\nu$, in a half-timestep $(\Delta t_{\mathrm{RT}}/2)$ update according to the radiative source and sink terms (the thermochemistry equations), then performs a full-timestep update based on the photon transport equations (Eqs. (52) and (53)), followed by another half-timestep update

based on the source and sink terms. This scheme achieves second-order convergence of the numerical solution to the RT equation (Kannan et al. 2019).

The extremely large speed of light c presents a fundamental challenge to numerical RT solvers. In particular, since c far exceeds the typical sound speeds of astrophysical gas, the transition from pure hydrodynamic to *radiation* hydrodynamic simulations increases the computational cost dramatically (e.g. Zier et al. 2024) due to the smaller RT timesteps required, $\Delta t_{\mathrm{RT}} \ll \Delta t_{\mathrm{hydro}}$. In order to maximize the efficiency of the code and facilitate large-volume simulations despite this challenge, AREPO-RT employs (i) *RT subcycling* and (ii) the *reduced speed of light approximation (RSLA)*.
Firstly, RT subcycling (Commerçon et al. 2014) refers to performing N_{sub} RT timesteps per hydrodynamic timestep, where $N_{\mathrm{sub}} > 1$. Thereby, the RT timestep $\Delta t_{\mathrm{RT}}{=}\Delta t_{\mathrm{hydro}}/N_{\mathrm{sub}}$ is set to be smaller than the hydrodynamic timestep $\Delta t_{\mathrm{hydro}}$. The advantage of this approach is that it allows a simulation to capture the relevant radiative processes on the comparatively short timescales involved, i.e. $\mathcal{O}(\Delta t_{\mathrm{RT}})$, while reducing the frequency with which the numerical routines associated with gravitational and hydrodynamic calculations are called. In particular, it avoids highly time-consuming frequent reconstructions of the Voronoi mesh in the case of AREPO-RT, substantially decreasing the total computing time (see Kannan et al. 2019, App. A). In this work, we employ a typical choice for the number of RT subcycles, $N_{\mathrm{sub}} = 64$, providing a good balance between numerical efficiency and accuracy (Kannan et al. 2022a; Zier et al. 2024).
Secondly, the RSLA replaces the physical speed of light c with an effective signal speed $\tilde{c} < c$ in the RT equations, aiming to essentially artificially increase the size of the timestep Δt_{RT}. This approximation is well-justified for problems in which the characteristic velocity of the system is sufficiently small, $v_{\mathrm{char}} \ll \tilde{c}$ (Kannan et al. 2019). Unfortunately, the latter does not always hold in cosmological-scale RHD simulations of reionization, in which the speed of ionization fronts in the dilute IGM can approach the physical speed of light, $\mathcal{O}(v_{\mathrm{char}}) \sim \mathcal{O}(c) > \tilde{c}$ (e.g. Rosdahl et al. 2013; Bauer et al. 2015). Nevertheless, it is frequently necessary to make use of the approximation in reionization simulations – though we note that in this case, it represents a compromise in physical accuracy – due to computational constraints. The simulations presented in this work therefore employ the RSLA by setting $\tilde{c} = 0.2\,c$ (following Kannan et al. 2022a).

The use of moment-based RT with the local M1 closure relation – as well as the application of the aforementioned methods to increase numerical efficiency – has facilitated large-volume RHD simulations of the EoR over recent years (e.g. Gnedin 2014; Ocvirk et al. 2016, 2020; Kannan et al. 2022a; Garaldi et al. 2024).
Notably, the THESAN simulations (Kannan et al. 2022a; Smith et al. 2022; Garaldi et al. 2022) use AREPO-RT in combination with the IllustrisTNG GFM (Sec. 3.1.3) to numerically

model reionization in a large-volume simulation box (length of the cubic box on each side: $L_{\mathrm{box}} = 95.5\,\mathrm{cMpc}$). THESAN further includes an additional dust model, following the production and destruction of cosmic dust based on empirical relations as described in McKinnon et al. (2016, 2017). This combination of numerical models, which we collectively refer to as the 'THESAN model' in the following, has been shown to successfully match a wide range of observational data at high redshifts (Kannan et al. 2022a; Garaldi et al. 2024; though see also Kannan et al. 2023; Koehler et al. 2024 for caveats regarding a similar numerical model in the context of recent *JWST* observations).
Crucially, the THESAN model forms the basis upon which the code extensions developed in this thesis build.[4]

Finally, we note that to further increase the efficiency of the code, AREPO-RT has been adapted and optimized for GPU-centric HPC architectures by Zier et al. (2024). For a more comprehensive description of the AREPO-RT code and THESAN model than summarized herein, we refer to Kannan et al. (2019, 2022a) for the original CPU-based implementation and to Zier et al. (2024) for the GPU-based acceleration.
Further details of the code and model, which are specifically relevant to one or more of the code extensions developed in this thesis, are additionally described in the following sections.

3.3 Implementation of X-ray sources

As mentioned in Sec. 1, high-redshift X-ray sources are crucial for obtaining accurate predictions of the thermal history of the early Universe. In particular, the most relevant contributors of soft X-ray photons (photon energy range $[0.5\,\mathrm{keV}, 2.0\,\mathrm{keV}]$) in this context are AGN, HMXBs, and the hot, diffuse ISM heated primarily by supernovae and stellar winds (Pacucci et al. 2014; see also Sec. 2.1).
We note that the contribution of stars is implemented in AREPO-RT using spectra generated with the *Binary Population and Spectral Synthesis* code (*BPASS*; Eldridge et al. 2017) as previously described in Kannan et al. (2022a). Although stellar sources are not particularly significant contributors of X-ray radiation (e.g. Haardt & Madau 2012; see Sec. 2.1), their simulated radiative output is also extended into the X-ray range according to the same *BPASS* tables.

In the following, we present the implementation of the X-ray sources mentioned above into the AREPO-RT code base.
We first describe the extension of the tracked range of photon energies into the soft X-ray band (Sec. 3.3.1). We then present an overview of how the X-ray sources' radiative energy

[4]Strictly speaking, we use an updated version of the core THESAN model, developed in the context of the THESAN-XL project (Zier et al. 2025a).

is injected into the simulated gas (Sec. 3.3.2), before describing the underlying spectral energy distributions (SEDs) and luminosities of each of the modeled sources in detail (Secs. 3.3.3-3.3.5). Finally, we discuss the computation of photoionization cross-sections, photoheating rates, and mean photon energies based on the different source spectra (Sec. 3.3.6).

In practice, including the new X-ray implementation is controlled by adding the compile-time flag `#define MRT_XRAY` to the configuration file (`Config.sh`) of the AREPO code. Additional related configuration options were also defined and implemented in this context. Firstly, we added the option to run with a total of five, rather than six, energy bins (see Sec. 3.3.1) by using the configuration flag `MRT_XRAY_FIVE_BINS`.
Secondly, using the flags `MRT_XRAY_NO_HMXB` and `MRT_XRAY_NO_ISM` causes the compiler to exclude the implementations associated with HMXBs and the hot ISM, respectively; thereby, we added the option of running the code with a custom set of included X-ray sources.
The source code and header files `RT_xray.cc`[5] and `RT_xray.h`, which contain the core parts of the X-ray implementation and source spectra, were added to the code base. Further, modifications specific to the X-ray models were made to the following existing files:

(i) `RT_init.cc`, containing the initialization routine called once at the beginning of each AREPO-RT run. It defines the energy bin edges in eV (see Sec. 3.3.1) and computes e.g. the photoionization cross-sections, photoheating rates, and mean photon energies (Sec. 3.3.6) in these energy bins based on the spectra of the included sources.

(ii) `RT.cc`, containing, among others, the function used to inject the radiative energies from all included sources (see Sec. 3.3.2) in each RT timestep of a run.

(iii) `RT_spectra.cc`, which defines the SEDs of the modeled radiative sources (Secs. 3.3.3-3.3.5).

(iv) `RT_proto.h`, a header file declaring functions used by the moment-based RT solver.

(v) `src/allvars.cc` and `src/allvars.h`, declaring (and, partially, defining) variables accessible to all other files of the code.

Additional modifications were made to the *Makefiles* controlling the compilation of the code, specifically to `makefiles/modules.make` to integrate the new files of the X-ray implementation into the larger code base, as well as to `makefiles/systypes.make` and

[5]Path in the AREPO code base directory: `/arepo/src/MRT/RT_xray.cc`. The files mentioned in the following share this base path, unless specified otherwise.

`Template-Makefile.systype` to add code compatibility with the GPU partition of the *Narval* supercomputer[6].

3.3.1 Extension of the photon energy range tracked in AREPO-RT

Photon frequencies ν – or, equivalently, photon energies $e = h\nu$ – are tracked in a set of discrete bins in AREPO-RT.[7] Using the energy density as an example, Fig. 7 shows an illustration of the way the primitive RT variables – the radiative energy density E and flux $\mathbf{F}$ – are therefore stored and processed in AREPO-RT, i.e., in a piecewise-constant description tracking only the total of each quantity in each bin.

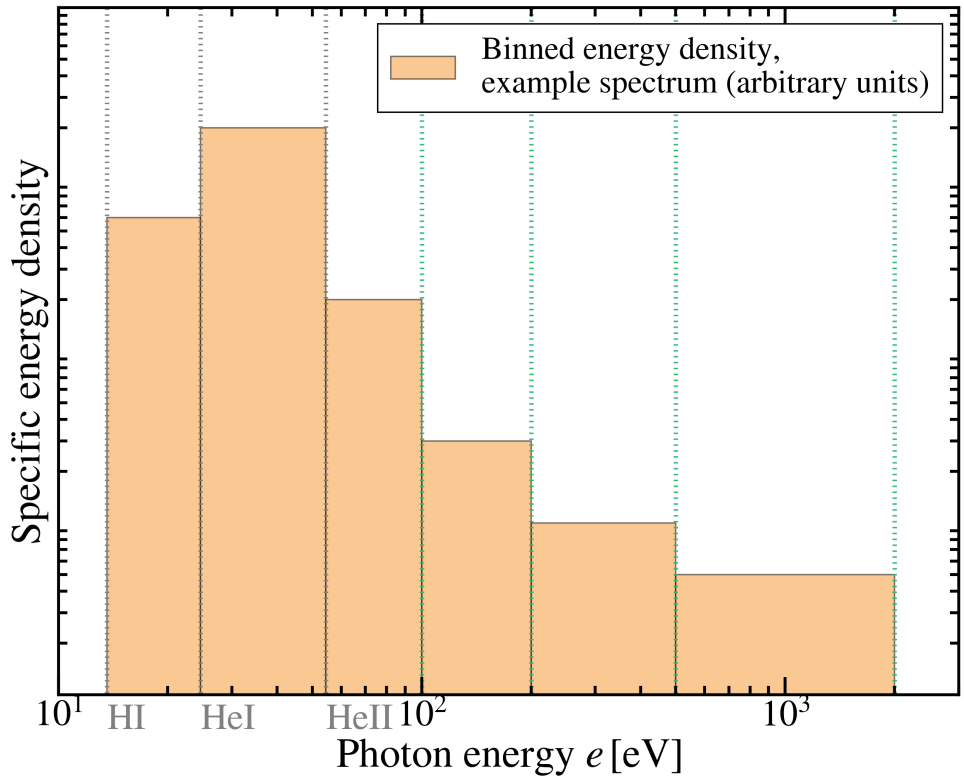

Figure 7: Schematic example spectrum in arbitrary units, illustrating how the radiative energy density as a function of the photon energy e is tracked within discrete energy bins in AREPO-RT. The radiative flux is handled analogously. The energy bin edges of AREPO-RT in the extreme UV range, as well as the ionization thresholds of neutral hydrogen H I, neutral helium He I, and singly ionized helium He II, are shown by the gray dotted vertical lines. Additionally, the edges of the newly added X-ray bins are indicated by the green dotted vertical lines.

Here, the binned primitive variables E_i and $\mathbf{F}_i$ in frequency bin i with edges $[\nu_i, \nu_{i+1}]$ are obtained from the monochromatic primitive variables E_ν and $\mathbf{F}_\nu$ (Sec. 3.2) by integrating

[6] `https://docs.alliancecan.ca/wiki/Narval/en`; see also the Acknowledgements for further details.

[7] We note that we refer to these bins as 'frequency bins' and 'energy bins' interchangeably in accordance with typical convention, as the definitions only differ by a constant factor, i.e., the Planck constant h.

over the specified frequency range, i.e.,

$$E_i = \int_{\nu_i}^{\nu_{i+1}} E_\nu \, \mathrm{d}\nu \ , \tag{54}$$

$$\mathbf{F}_i = \int_{\nu_i}^{\nu_{i+1}} \mathbf{F}_\nu \, \mathrm{d}\nu \ . \tag{55}$$

The RT moment equations for these binned quantities are then given by (cf. Eqs. (39) and (40))

$$\frac{\partial E}{\partial t} + \nabla \cdot \mathbf{F} = S - \kappa_E \, \rho \, E \tag{56}$$

$$\frac{\partial \mathbf{F}}{\partial t} + c^2 \, \nabla \, \mathbb{P} = -\kappa_F \, \rho \, \mathbf{F} \ . \tag{57}$$

Here, κ_E and κ_F refer to the mean absorption coefficient κ_ν in the range $[\nu_i, \nu_{i+1}]$ weighted by E_ν and $\mathbf{F}_\nu$, respectively. Further, the source function S_ν is binned analogously to the energy density and flux, i.e.,

$$S_i = \int_{\nu_i}^{\nu_{i+1}} S_\nu \, \mathrm{d}\nu \ . \tag{58}$$

Finally, as previously described for the monochromatic case (Eq. (46)), the binned radiation pressure tensor $\mathbb{P}_i$ is defined via the Eddington tensor $\mathbb{D}$ as $\mathbb{P}_i = E_i \, \mathbb{D}$, where $\mathbb{D}$ is computed using the M1 closure relation (Eq. (47)).

A typical choice for the energy bin edges in the EUV range is defining them based on the photoionization thresholds of H I, He I, and He II at the photon energies 13.6 eV, 24.59 eV, and 54.42 eV, respectively (e.g. Verner et al. 1996).
This choice is adopted in the original THESAN reionization simulations, which considered UV photons up to a cutoff energy of 100 eV, therefore resulting in the following three energy bins:

$$\begin{aligned} &\text{Bin 0: } [13.6\,\text{eV}, 24.59\,\text{eV}] \ , \\ &\text{Bin 1: } [24.59\,\text{eV}, 54.42\,\text{eV}] \ , \\ &\text{Bin 2: } [54.42\,\text{eV}, 100\,\text{eV}] \ . \end{aligned}$$

To cover the relevant photon energy range up to the soft X-ray band (Sec. 2.1), we introduce the three additional bins

$$\begin{aligned} &\text{Bin 3: } [100\,\text{eV}, 200\,\text{eV}] \ , \\ &\text{Bin 4: } [200\,\text{eV}, 500\,\text{eV}] \ , \\ &\text{Bin 5: } [500\,\text{eV}, 2\,000\,\text{eV}] \ . \end{aligned}$$

This definition of the photon energy bin edges is implemented in `RT_init.cc` in the AREPO-RT code and illustrated in Fig. 7. We note that while we refer to the newly added bins as the 'X-ray bins' in the following to differentiate them from the existing UV-range bins of the THESAN model, the standard definition of the soft X-ray band is $[500\,\mathrm{eV}, 2\,000\,\mathrm{eV}]$ (e.g. Mineo et al. 2012a,b; Pacucci et al. 2014; Madau & Fragos 2017); strictly speaking, only the highest-energy bin (bin 5) therefore covers this range. Additionally, we note that we implemented the configuration flag `MRT_XRAY_FIVE_BINS` mentioned above to combine bins 3 and 4 into a single bin covering the energy range $[100\,\mathrm{eV}, 500\,\mathrm{eV}]$.

The edges of the newly implemented bins were chosen to (i) cover the relevant photon energy range up to 2 keV (Pacucci et al. 2014), (ii) ensure that all bins are approximately logarithmically evenly spaced with respect to the photon energy (see e.g. Fig. 7), (iii) reflect substantial breaks in the X-ray source spectra (at $e \sim 200\,\mathrm{eV}$ for the hot ISM, HMXB, and AGN, and at $e \sim 500\,\mathrm{eV}$ for HMXB; cf. Sec. 3.3.2, Fig. 8), and (iv) ensure that the new bins each contribute a comparable amount of photon energy from the X-ray sources throughout a simulation (see Sec. 4.1.5, Fig. 22).

3.3.2 Injection of radiative energy from X-ray sources

For all considered sources, the injection of radiative energy into each of the frequency bins is determined by (i) the shape of the respective source SED $\frac{\mathrm{d}L}{\mathrm{d}e}$, illustrated in Fig. 8, and (ii) the total luminosity output $L^{13.6-2000\,\mathrm{eV}}$ expected from the source type in the complete energy range considered in AREPO-RT, i.e., $[13.6\,\mathrm{eV}, 2\,000\,\mathrm{eV}]$.
Specifically, in the context of the code's RT initialization (`RT_init.cc`) at the start of a run, each source SED is numerically integrated with respect to the photon energy within each of the intervals specified by the energy bins. Dividing the resulting value for bin i by the integral of the SED over the entire energy range of $[13.6\,\mathrm{eV}, 2\,000\,\mathrm{eV}]$ then determines the fraction $x_{e,i}$ of this source's radiative energy output injected into bin i (presented in Table 1 for each of the source types and energy bins).
Consequently, the luminosity in bin i is given by

$$L_i = x_{e,i}\, L^{13.6-2000\,\mathrm{eV}} \,, \tag{59}$$

i.e., scaled by the appropriate energy fraction based on the shape of the SED.
During an RT time step Δt_{RT}, this finally results in the following radiative energy injected into bin i:

$$E_i = x_{e,i}\, L^{13.6-2000\,\mathrm{eV}}\, \Delta t_{\mathrm{RT}} \,. \tag{60}$$

In practice, this injection of energies is carried out during an AREPO-RT run by calling the function `add_xray_fluxes()` (defined in `RT_xray.cc`) in `RT.cc`.

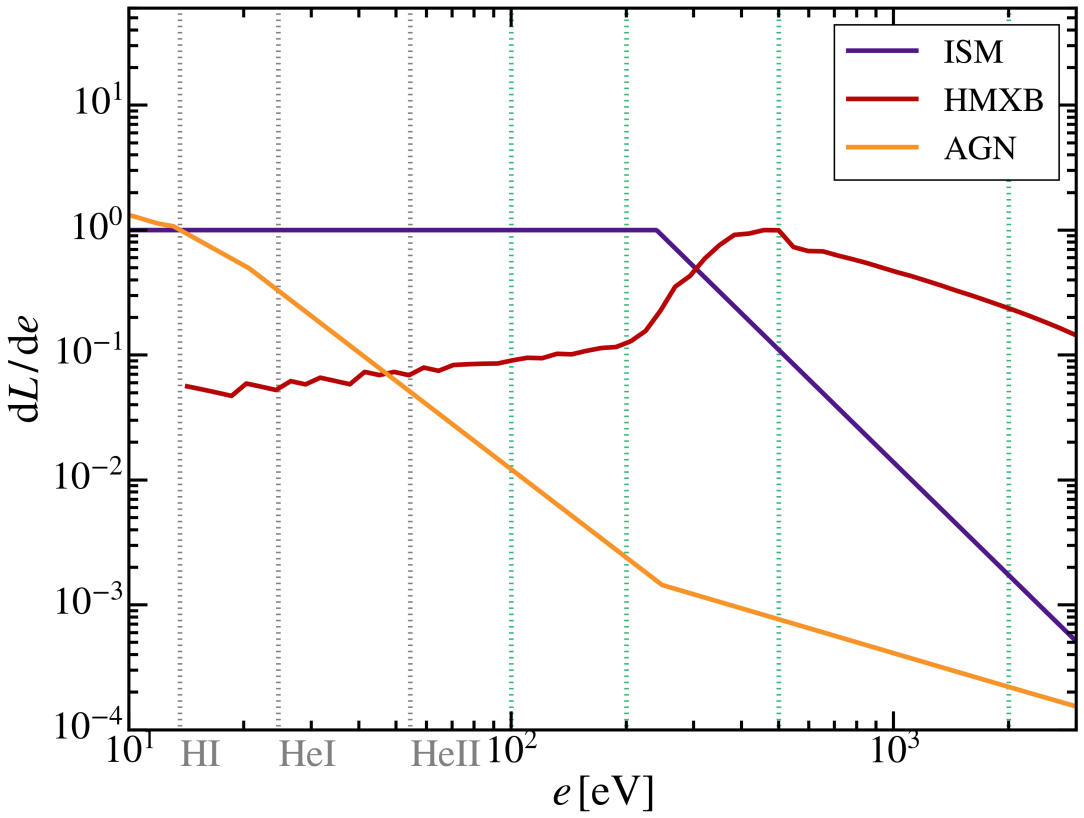

Figure 8: Spectral energy distributions of key high-redshift X-ray sources as a function of the photon energy e. This figure shows the spectral shapes in arbitrary units; the final normalizations are set by the luminosities of the corresponding sources (see text for details). The spectrum of the hot interstellar medium (ISM, purple curve) is modeled by a thermal bremsstrahlung profile with characteristic energy $k_B T_{\mathrm{ISM}} = 240\,\mathrm{eV}$ (Mineo et al. 2012b; Pacucci et al. 2014). For high-mass X-ray binaries (HMXB, red curve), we use the intrinsic $z \sim 10$ spectral energy distribution from Fragos et al. (2013b, 2016). Finally, we use the spectral energy distribution given in Shen et al. (2020) to model radiation from active galactic nuclei (AGN, orange curve).
Vertical dotted lines indicate the energy binning used in this work, cf. Sec. 3.3.1, as noted in the caption of Fig. 7.

The specifics of the underlying SEDs and luminosities for each of the X-ray source types are described in the following.

3.3.3 Hot interstellar medium

The emissions of the diffuse, hot interstellar medium, heated by supernovae and stellar winds, are modeled using a thermal bremsstrahlung SED (Pacucci et al. 2014; see also Eide et al. 2018) of the shape

$$\frac{\mathrm{d}L_{\mathrm{ISM}}(e)}{\mathrm{d}e} = \begin{cases} 1 & , e < k_B T_{\mathrm{ISM}}, \\ \left(\frac{e}{k_B T_{\mathrm{ISM}}}\right)^{-3} & , e \geq k_B T_{\mathrm{ISM}}\ . \end{cases} \tag{61}$$

Table 1: Fraction $x_{e,i}$ of the total radiative energy of each of the X-ray sources injected into each of the energy bins.

Energy bin [eV]	$x_{e,i}$, ISM [%]	$x_{e,i}$, HMXB [%]	$x_{e,i}$, AGN [%]
$13.6 - 24.59$	3.19	0.07	36.96
$24.59 - 54.42$	8.66	0.23	35.21
$54.42 - 100.0$	13.23	0.44	15.72
$100.0 - 200.0$	29.02	1.24	6.89
$200.0 - 500.0$	38.38	23.24	2.06
$500.0 - 2\,000.0$	7.52	74.79	3.15

Specifically, we assume a flat spectrum (i.e., spectral index $\alpha = 0$) for the hot ISM's SED at lower photon energies, describing the approximate spectral shape of thermal bremsstrahlung emissions below the characteristic thermal energy threshold $k_B T$ (e.g. Rybicki & Lightman 1979), where k_B denotes the Boltzmann constant. Above the characteristic thermal energy, we assume a drop-off of the SED with spectral index $\alpha = 3$, resulting in the relatively soft X-ray spectrum (Mineo et al. 2012b; Pacucci et al. 2014) illustrated in Fig. 8.
This thermal bremsstrahlung profile, which we implemented in `RT_spectra.cc` in AREPO-RT, has been shown to be a well-suited effective model for the SED of the hot ISM (e.g. Pacucci et al. 2014). Further, the latter's mean thermal energy for a typical galaxy – serving as the characteristic thermal energy threshold in Eq. (61) – has been determined as $\langle k_B T \rangle = 240\,\mathrm{eV}$ by Mineo et al. (2012b) using *Chandra* (Weisskopf et al. 2000) observations of nearby galaxies' diffuse X-ray emissions.

The normalization of this spectrum, described in relative units in Eq. (61), is set by an empirical estimate for the intrinsic luminosity of the hot ISM per star formation rate (SFR). In the photon energy band [0.3 keV, 10 keV], Mineo et al. (2012b) estimate this scaling of the intrinsic luminosity with the SFR as

$$L_{\mathrm{ISM}}^{0.3-10\,\mathrm{keV}}\,[\mathrm{erg\,s^{-1}}] = (7.3 \pm 1.3) \times 10^{39} \cdot \mathrm{SFR}\,[M_\odot\,\mathrm{yr}^{-1}]\,. \tag{62}$$

This value has been corrected for both Galactic and intrinsic absorptions, i.e., absorptions of the hot ISM's radiation within both the Milky Way and the respective observed galaxy itself (see Mineo et al. 2012b for further details).
We rescale this value to the photon energy band [13.6 eV, 2 keV] using the thermal bremsstrahlung SED given by Eq. (61). Specifically, we integrate the SED with respect to e in the energy ranges [0.3 keV, 10 keV] and [13.6 eV, 2 keV] to obtain $L_{\mathrm{ISM,rel}}^{0.3-10\,\mathrm{keV}}$ and $L_{\mathrm{ISM,rel}}^{13.6-2000\,\mathrm{eV}}$, respectively, in the relative units of Eq. (61). We then compute the factor converting the luminosity from the former energy range to the latter, based on the thermal

bremsstrahlung SED, as

$$\frac{L_{\mathrm{ISM,rel}}^{13.6-2000\,\mathrm{eV}}}{L_{\mathrm{ISM,rel}}^{0.3-10\,\mathrm{keV}}} = 4.4919\,. \tag{63}$$

As a cross-check, we performed this SED integration both numerically and analytically, verifying the value of the conversion factor given in Eq. (63).
Rescaling the mean value cited in Eq. (62) by this factor, we obtain the following scaling for the intrinsic ISM luminosity in our energy range:

$$L_{\mathrm{ISM}}^{13.6-2000\,\mathrm{eV}}[\mathrm{erg\,s^{-1}}] = 3.3 \times 10^{40} \cdot \mathrm{SFR}\,[\mathrm{M_\odot\,yr^{-1}}]\,. \tag{64}$$

Finally, the radiative energy input from the hot ISM in each bin, injected during the RT time step Δt_{RT}, is determined by Eq. (60) for each active cell, using Eq. (64) and the cell's current SFR as tracked by the code.

3.3.4 High-mass X-ray binaries

We employ the tabulated values from Fragos et al. (2013b) (*erratum:* Fragos et al. 2016) for the intrinsic HMXB SED $\frac{\mathrm{d}L_{\mathrm{HMXB}}(e)}{\mathrm{d}(\log e)}$, which are based on a large-scale population synthesis simulation (Fragos et al. 2013a) modeling the evolution of X-ray binaries using the `StarTrack` code (Belczynski et al. 2002, 2008). Specifically, we use the $z \sim 10$ intrinsic SED from Fragos et al. (2016)[8], as the shape of the SED is approximately redshift-independent (Fragos et al. 2013b). We linearly interpolate the tabulated values for $\frac{\mathrm{d}L_{\mathrm{HMXB}}(e)}{\mathrm{d}(\log e)}$ to obtain the spectrum as a function of energy. In case the photon energy for which we require the spectrum exceeds the energy range of the tabulated values, we use the minimum and maximum value, respectively, for energies below and above the given range. We note that in our implementation, this only affects the small range $[13.6\,\mathrm{eV}, 14.22\,\mathrm{eV}]$, as 14.22 eV is the minimum photon energy tabulated in Fragos et al. (2016).
Subsequently, we convert from $\frac{\mathrm{d}L_{\mathrm{HMXB}}(e)}{\mathrm{d}(\log e)}$ to $\frac{\mathrm{d}L_{\mathrm{HMXB}}(e)}{\mathrm{d}e}$ using $\mathrm{d}(\log e) \propto \mathrm{d}e/e$. The resulting spectral shape is plotted in Fig. 8.

Similarly to the hot ISM, the luminosity contributions from HMXBs are scaled linearly with the SFR; however, this scaling is additionally dependent on the metallicity Z of newly formed stars (Fragos et al. 2013b; Madau & Fragos 2017). We employ the Madau & Fragos (2017) parametrization for this metallicity dependence (see Fig. 9), which estimates the logarithm of the HMXB luminosity L_{HMXB} per SFR by an eighth-order polynomial in Z,

$$\log(L_{\mathrm{HMXB}}/\mathrm{SFR}) = \sum_{\mathrm{i}=0}^{8} \beta_{\mathrm{i}}\,\mathrm{Z^i}\,, \tag{65}$$

[8]The corresponding tabulated values were obtained from the machine-readable table available at `https://content.cld.iop.org/journals/2041-8205/827/1/L21/revision1/apjlaa2e15t1_mrt.txt`.

with the polynomial coefficients $\beta_0, \ldots, \beta_8$ for the intrinsic luminosity $L_{\mathrm{HMXB}}^{2-10\,\mathrm{keV}}$ in the energy range [2 keV, 10 keV] given in Table 1 of Madau & Fragos (2017).
To be representative of the metallicity of *newly formed* stars in this context, we use the gas metallicity, tracked as a property of each cell, rather than the stellar metallicity to evaluate $L_{\mathrm{HMXB}}^{2-10\,\mathrm{keV}}/\mathrm{SFR}$.
Further, in order to save computing time – since this calculation of $L_{\mathrm{HMXB}}^{2-10\,\mathrm{keV}}/\mathrm{SFR}$ is required for each active, star-forming cell in each RT time step – the computation of the eighth-order polynomial is replaced by an interpolation of $\log(L_{\mathrm{HMXB}}/\mathrm{SFR})$ from 100 precomputed evaluations of Eq. (65) covering metallicities within the range $0 \leq Z \leq 0.025$. For metallicities exceeding this range ($Z > 0.025$; see Fragos et al. 2013b), we set $\log(L_{\mathrm{HMXB}}/\mathrm{SFR}) = \log\left(L_{\mathrm{HMXB}}(Z = 0.025)/\mathrm{SFR}\right)$.

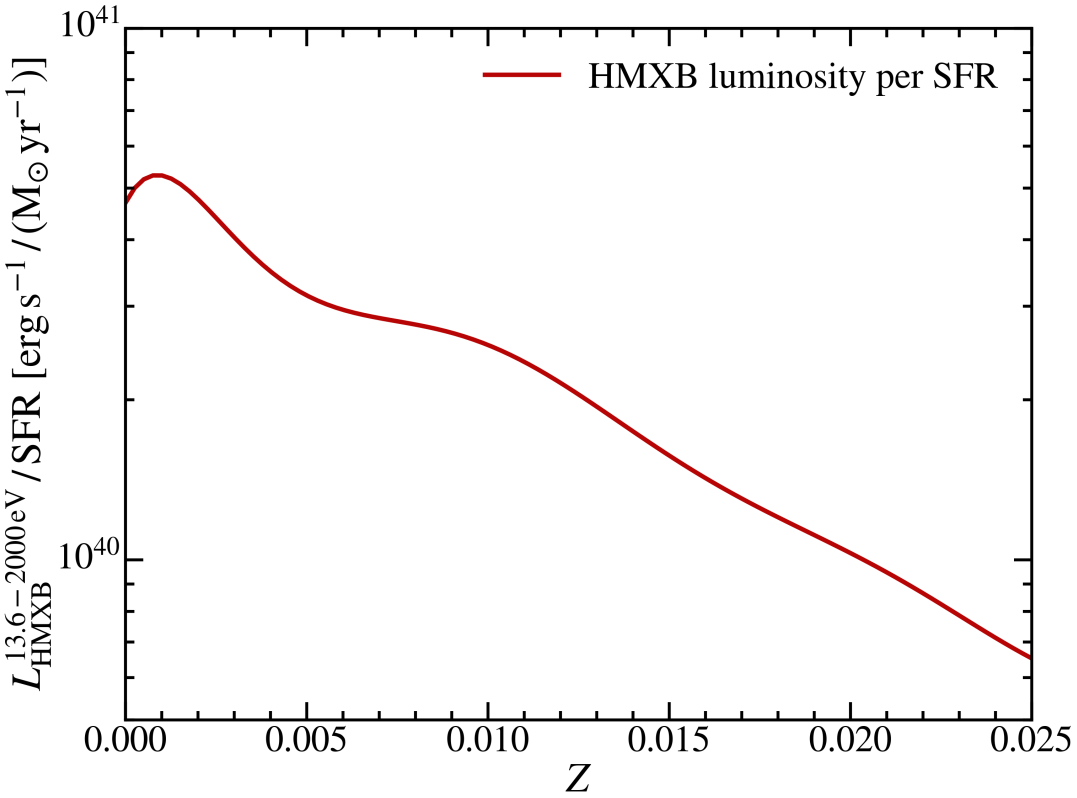

Figure 9: The high-mass X-ray binary (HMXB) luminosity per star formation rate (SFR) as a function of the metallicity Z of newly formed stars, based on the parametrization given in Madau & Fragos (2017) and rescaled to the energy range [13.6 eV, 2 keV] of AREPO-RT.

The result is then rescaled to the energy range of AREPO-RT, analogously to the case of the hot ISM described above. Here, we numerically integrate the Fragos et al. (2013b, 2016) SED with respect to $\log e$ to compute $L_{\mathrm{HMXB,rel}}^{2-10\,\mathrm{keV}}$ and $L_{\mathrm{HMXB,rel}}^{13.6-2000\,\mathrm{eV}}$. The resulting luminosity rescaling factor, converting from the [2 keV, 10 keV] energy range to our [13.6 eV, 2 keV] range, is given by

$$\frac{L_{\mathrm{HMXB,rel}}^{13.6-2000\,\mathrm{eV}}}{L_{\mathrm{HMXB,rel}}^{2-10\,\mathrm{keV}}} = 1.7369\,. \tag{66}$$

We verify the accuracy of our numerical $\mathrm{d}(\log e)$ integration by additionally integrating the $z \sim 10$ Fragos et al. (2013b) HMXB SED including interstellar absorption over the energy range [2 keV, 10 keV] and comparing the result ($8.33 \times 10^{37}\,\mathrm{erg\,s^{-1}Mpc^{-3}}$) to the left panel of Fig. 1 of Fragos et al. (2013b).
Using the rescaling factor computed in Eq. (66) to obtain $L_{\mathrm{HMXB}}^{13.6-2000\,\mathrm{eV}}/\mathrm{SFR}$ from the interpolation of Eq. (65), we calculate the HMXB luminosity in each active cell based on its respective SFR. The result is then used to compute the corresponding energy input in each bin during the RT time step Δt_{RT} as described by Eq. (60); this radiative energy is subsequently injected into the cell.

3.3.5 Active galactic nuclei

To describe the spectrum of AGN radiation in the EUV range, the original THESAN simulations use the Lusso et al. (2015) parametrization, providing an estimate for wavelengths $\lambda > 600\,\text{Å}$ (photon energies $e < 20.66\,\mathrm{eV}$).
In order to also reliably obtain the spectral shape in the range of the added X-ray energy bins, we instead use the AGN SED parametrization given by Shen et al. (2020). We linearly interpolate their tabulated values[9] for $\log(\nu L_\nu\,[\mathrm{erg\,s^{-1}}])$, where $L_\nu \equiv \frac{\mathrm{d}L}{\mathrm{d}\nu}$, analogously to the HMXB SED (Sec. 3.3.4) and subsequently convert this to $L_\nu\,[\mathrm{erg\,s^{-1}\,Hz^{-1}}]$.
Additionally, the SED is normalized using the condition $\int_0^\infty L_\nu\,\mathrm{d}\nu = 1$ by integrating the SED over the entire tabulated frequency range, [$3.16 \times 10^{12}\,\mathrm{Hz}$, $4.83 \times 10^{20}\,\mathrm{Hz}$] (corresponding to an energy range of [0.013 eV, 2.00 MeV]). This results in a normalization factor of $2.2156 \times 10^{46}\,\mathrm{erg\,s^{-1}}$, by which we divide the SED values L_ν.

We note that this procedure differs slightly from the manner in which the SED shape is set in Secs. 3.3.3 and 3.3.4 – i.e., as $\frac{\mathrm{d}L}{\mathrm{d}e}$ in relative units – to ensure compatibility with existing parts of the implementation of AGN radiation in AREPO-RT.
Further, the computation of the bolometric AGN luminosity, which is scaled with the SMBH mass accretion rate in the THESAN model (Weinberger et al. 2018; Kannan et al. 2022a), is not modified in the context of the X-ray implementation.

3.3.6 Photoionization and photoheating rates

Modeling the coupling of the energy injected by radiative sources, such as stars and the X-ray sources discussed above, to the simulated gas is a crucial component of RHD simulations (Sec. 3.2). In particular, photoionization of the chemical species H I, He I, and He II and the associated photoheating of the gas are key processes in the context of reionization simulations.

[9] Obtained from `https://bitbucket.org/ShenXuejian/quasarlf/src/master/data/MySED.dat`.

Table 2: Mean photon energies e in AREPO-RT's frequency bins based on the different source spectra of stars, the hot interstellar medium (ISM), high-mass X-ray binaries (HMXB), and active galactic nuclei (AGN).

Energy bin [eV]	e, stars [eV]	e, ISM [eV]	e, HMXB [eV]	e, AGN [eV]
$13.6 - 24.59$	18.16	18.56	18.58	17.70
$24.59 - 54.42$	32.04	37.56	38.16	34.68
$54.42 - 100.0$	57.15	74.93	75.64	70.34
$100.0 - 200.0$	103.05	144.31	146.80	128.49
$200.0 - 500.0$	203.56	276.46	374.01	294.71
$500.0 - 2\,000.0$	800.83	714.33	938.95	938.93

In the following, we therefore describe the computation and values of the mean photon energies e, photoionization cross-sections σ, and photoheating rates $\mathcal{E}$ per ionization event in the AREPO-RT frequency bins. These quantities are computed individually in the initialization routine `RT_init.cc` for each of the source types and vary based on the shape of the underlying source spectrum.

Firstly, the mean photon energy e_i emitted by a source in frequency bin i, defined by the edges $[\nu_i, \nu_{i+1}]$, is computed as (Kannan et al. 2019)

$$e_i = \frac{\int_{\nu_i}^{\nu_{i+1}} 4\pi J_\nu \, \mathrm{d}\nu}{\int_{\nu_i}^{\nu_{i+1}} \frac{4\pi J_\nu}{h\nu} \, \mathrm{d}\nu} \, , \tag{67}$$

where the *mean intensity* J_ν characterizes the spectrum of the source and is defined as the spatial average of the specific intensity I_ν (e.g. Mihalas & Weibel-Mihalas 1984, p. 312),

$$J_\nu = \frac{1}{4\pi} \int_{4\pi} I_\nu \, \mathrm{d}\Omega \, . \tag{68}$$

The mean photon energies computed according to Eq. (67) are presented in Table 2 for all of the newly implemented X-ray source spectra, as well as for the stellar spectrum.

Further, the mean photoionization cross-section $\bar{\sigma}_{ij}$ in bin i for chemical species j is calculated as

$$\bar{\sigma}_{ij} = \frac{\int_{\nu_i}^{\nu_{i+1}} \frac{4\pi J_\nu}{h\nu} \sigma_{j\nu} \, \mathrm{d}\nu}{\int_{\nu_i}^{\nu_{i+1}} \frac{4\pi J_\nu}{h\nu} \, \mathrm{d}\nu} \, . \tag{69}$$

Here, $\sigma_{j\nu}$ denotes the photoionization cross-section for species j as a function of the photon frequency, given by the Verner et al. (1996) parametrization in the current version of AREPO-RT and illustrated in Fig. 3.

The values of $\bar{\sigma}_{ij}$ for all source types and chemical species H I, He I, and He II are presented in Table 3. Additionally, we plot $\bar{\sigma}_{ij}$ as a function of the mean photon energies e_i for all sources in Fig. 10, directly comparing these values to the analytical Verner et al. (1996)

Table 3: Comparison between photoionization cross-sections σ from the different source spectra in AREPO-RT's frequency bins for the chemical species H I, He I, and He II.

Energy bin [eV]	$\sigma_{\mathrm{H\,I}}$, stars [cm^2]	$\sigma_{\mathrm{H\,I}}$, ISM [cm^2]	$\sigma_{\mathrm{H\,I}}$, HMXB [cm^2]	$\sigma_{\mathrm{H\,I}}$, AGN [cm^2]
$13.6 - 24.59$	3.32×10^{-18}	3.15×10^{-18}	3.15×10^{-18}	3.55×10^{-18}
$24.59 - 54.42$	7.00×10^{-19}	4.88×10^{-19}	4.67×10^{-19}	5.94×10^{-19}
$54.42 - 100.0$	1.08×10^{-19}	5.64×10^{-20}	5.48×10^{-20}	6.67×10^{-20}
$100.0 - 200.0$	1.78×10^{-20}	7.94×10^{-21}	7.56×10^{-21}	1.07×10^{-20}
$200.0 - 500.0$	2.09×10^{-21}	1.05×10^{-21}	4.20×10^{-22}	9.73×10^{-22}
$500.0 - 2\,000.0$	4.56×10^{-23}	5.49×10^{-23}	3.39×10^{-23}	3.40×10^{-23}
Energy bin [eV]	$\sigma_{\mathrm{He\,I}}$, stars [cm^2]	$\sigma_{\mathrm{He\,I}}$, ISM [cm^2]	$\sigma_{\mathrm{He\,I}}$, HMXB [cm^2]	$\sigma_{\mathrm{He\,I}}$, AGN [cm^2]
$13.6 - 24.59$	0	0	0	0
$24.59 - 54.42$	5.15×10^{-18}	4.02×10^{-18}	3.91×10^{-18}	4.59×10^{-18}
$54.42 - 100.0$	1.54×10^{-18}	9.12×10^{-19}	8.92×10^{-19}	1.04×10^{-18}
$100.0 - 200.0$	3.65×10^{-19}	1.76×10^{-19}	1.68×10^{-19}	2.31×10^{-19}
$200.0 - 500.0$	5.24×10^{-20}	2.70×10^{-20}	1.12×10^{-20}	2.51×10^{-20}
$500.0 - 2\,000.0$	1.30×10^{-21}	1.56×10^{-21}	9.68×10^{-22}	9.71×10^{-22}
Energy bin [eV]	$\sigma_{\mathrm{He\,II}}$, stars [cm^2]	$\sigma_{\mathrm{He\,II}}$, ISM [cm^2]	$\sigma_{\mathrm{He\,II}}$, HMXB [cm^2]	$\sigma_{\mathrm{He\,II}}$, AGN [cm^2]
$13.6 - 24.59$	0	0	0	0
$24.59 - 54.42$	0	0	0	0
$54.42 - 100.0$	1.41×10^{-18}	7.74×10^{-19}	7.54×10^{-19}	9.04×10^{-19}
$100.0 - 200.0$	2.73×10^{-19}	1.28×10^{-19}	1.23×10^{-19}	1.70×10^{-19}
$200.0 - 500.0$	3.77×10^{-20}	1.95×10^{-20}	8.23×10^{-21}	1.82×10^{-20}
$500.0 - 2\,000.0$	1.01×10^{-21}	1.21×10^{-21}	7.53×10^{-22}	7.55×10^{-22}

photoionization cross-sections used for the underlying computation (Eq. (69)).
To verify the accuracy of the numerical calculations of $\bar{\sigma}_{ij}$, we further perform a test using extremely narrow frequency bins, each covering a range of $\Delta(h\nu) = 0.1\,\mathrm{eV}$ and each centered around the AREPO-RT bin edges (e.g., resulting in the narrow bin $[13.55\,\mathrm{eV}, 13.65\,\mathrm{eV}]$ for the lower edge of bin 0 at 13.6 eV). Due to the small range of photon frequencies covered, the photoionization cross-sections computed in these bins are expected to match the Verner et al. (1996) parametrization exactly. We therefore compute $\bar{\sigma}_{ij}$ according to Eq. (69) in these narrow bins and plot the results in Fig. 10 accordingly, verifying the agreement with the Verner et al. (1996) cross-sections.

As mentioned in Sec. 2.1, the amount of photoheating caused by a photoionization event is determined by the energy each ionizing photon carries in excess of the species' ionization threshold energy $h\nu_{t,j}$. Specifically, the *total* photoheating rate Γ_j due to species j is therefore given by

$$\Gamma_j = \int_{\nu_{t,j}}^{\infty} \frac{4\pi J_\nu}{h\nu}\, \sigma_{j\nu}\, (h\nu - h\nu_{t,j})\, \mathrm{d}\nu \ . \tag{70}$$

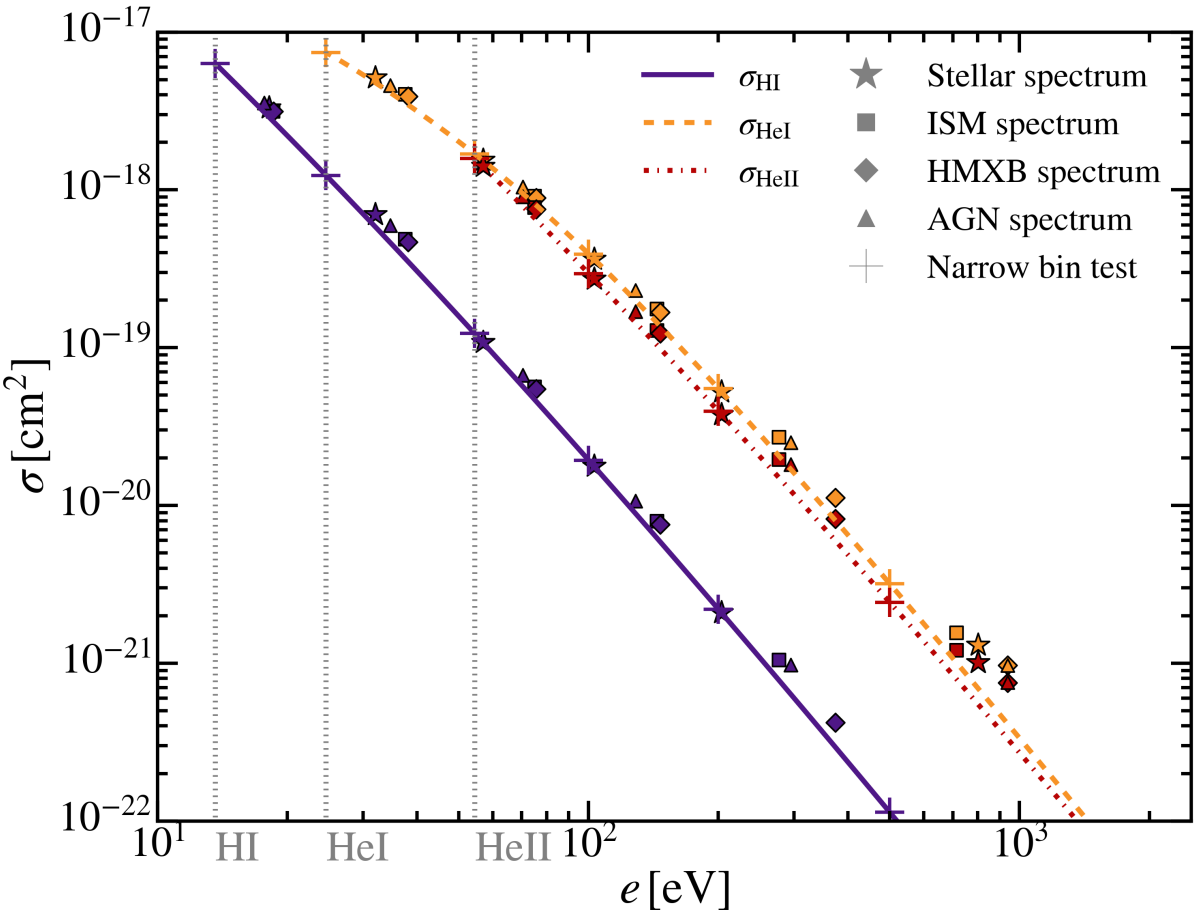

Figure 10: Mean photoionization cross-sections $\bar{\sigma}_{ij}$ (scatter markers) for all source types – stars, the hot interstellar medium (ISM), high-mass X-ray binaries (HMXB), and active galactic nuclei (AGN) – and for chemical species H I, He I, and He II as a function of the mean photon energy e_i in the AREPO-RT bins. For reference, we also show the analytic Verner et al. (1996) parametrization for these cross-sections σ as a function of the photon energy e (solid, dashed, and dash-dotted curves), as well as the ionization threshold energies for H I, He I, and He II (gray dotted vertical lines). Colors indicate the respective chemical species (H I: purple, He I: orange, He II: red), while the marker symbols indicate the underlying source spectrum (stellar: star markers, ISM: squares, HMXB: diamonds, AGN: triangles). Further, cross markers in the corresponding colors indicate the results obtained for $\bar{\sigma}_{ij}$ in a numerical test using extremely narrow frequency bins (see text for details).

Within frequency bin i, the energy $\mathcal{E}_{ij}$ injected *per photoionization event* by the photoheating of species j is consequently computed in AREPO-RT as (Kannan et al. 2019)

$$\mathcal{E}_{ij} = \frac{\int_{\nu_i}^{\nu_{i+1}} \frac{4\pi J_\nu}{h\nu} \sigma_{j\nu} (h\nu - h\nu_{t,j}) \, \mathrm{d}\nu}{\int_{\nu_i}^{\nu_{i+1}} \frac{4\pi J_\nu}{h\nu} \sigma_{j\nu} \, \mathrm{d}\nu} \, . \tag{71}$$

The resulting values for $\mathcal{E}_{ij}$ for all source types, and for photoionization events of the chemical species H I, He I, and He II, are presented in Table 4. We note that, as expected, this comparison of $\mathcal{E}_{ij}$ for the different source types shows one of the key points made

Table 4: Comparison between the mean photoheating energies $\mathcal{E}$ injected per ionization event, calculated using the different source spectra in AREPO-RT's frequency bins for H I, He I, and He II.

Energy bin [eV]	$\mathcal{E}_{\mathrm{H\,I}}$, stars [eV]	$\mathcal{E}_{\mathrm{H\,I}}$, ISM [eV]	$\mathcal{E}_{\mathrm{H\,I}}$, HMXB [eV]	$\mathcal{E}_{\mathrm{H\,I}}$, AGN [eV]
$13.6 - 24.59$	3.24	3.56	3.54	2.87
$24.59 - 54.42$	15.63	18.89	19.33	17.02
$54.42 - 100.0$	42.99	54.84	55.48	51.48
$100.0 - 200.0$	89.21	114.41	116.24	104.38
$200.0 - 500.0$	189.77	229.35	300.07	230.31
$500.0 - 2\,000.0$	601.15	580.22	635.72	635.85
Energy bin [eV]	$\mathcal{E}_{\mathrm{He\,I}}$, stars [eV]	$\mathcal{E}_{\mathrm{He\,I}}$, ISM [eV]	$\mathcal{E}_{\mathrm{He\,I}}$, HMXB [eV]	$\mathcal{E}_{\mathrm{He\,I}}$, AGN [eV]
$13.6 - 24.59$	0	0	0	0
$24.59 - 54.42$	5.51	9.54	10.06	7.28
$54.42 - 100.0$	32.13	45.13	45.79	41.49
$100.0 - 200.0$	78.25	104.81	106.72	94.25
$200.0 - 500.0$	178.79	219.33	291.41	220.68
$500.0 - 2\,000.0$	591.75	570.36	627.20	627.28
Energy bin [eV]	$\mathcal{E}_{\mathrm{He\,II}}$, stars [eV]	$\mathcal{E}_{\mathrm{He\,II}}$, ISM [eV]	$\mathcal{E}_{\mathrm{He\,II}}$, HMXB [eV]	$\mathcal{E}_{\mathrm{He\,II}}$, AGN [eV]
$13.6 - 24.59$	0	0	0	0
$24.59 - 54.42$	0	0	0	0
$54.42 - 100.0$	2.21	14.55	15.20	11.07
$100.0 - 200.0$	48.41	74.65	76.55	64.18
$200.0 - 500.0$	148.96	189.90	262.83	191.45
$500.0 - 2\,000.0$	564.07	542.01	600.81	600.82

in Sec. 2.1 regarding X-ray heating: the sources with a harder spectrum, i.e., a larger relative contribution at high photon energies ($[500\,\mathrm{eV}, 2\,000\,\mathrm{eV}]$), contribute significantly more heating energy per photoionization event. Specifically, this is evident here for the harder HMXB and AGN spectra (cf. Fig. 8).

3.4 Piecewise power-law implementation of cosmological redshifting

Numerical approaches to RT commonly neglect the effects of cosmological redshifting (Sec. 1), described theoretically by a complex, spectrum-dependent term in the general RT equation (Eq. (31)); the original THESAN model is no exception to this trend (Kannan et al. 2019, 2022a). The underlying assumption is that all photons only propagate sufficiently small distances – on the time and length scales considered – for the Universe's expansion to be negligible in the time between each photon's emission and absorption (e.g. Gnedin

& Ostriker 1997). This is typically a fair approximation for EUV-range photons (e.g., $13.6\,\mathrm{eV} \leq h\nu \leq 100\,\mathrm{eV}$ in the THESAN model; Kannan et al. 2022a) due to their large ionization cross-sections with the astrophysically abundant H I, He I, and He II gases (see Fig. 3). However, as noted in Sec. 2.1, X-ray photons have substantially longer mean free paths; consequently, this assumption breaks down at higher photon energies. Here, the effects of cosmological redshifting become more pronounced, motivating the development of numerical methods capturing this process.

We present two distinct methods addressing this, which have been newly implemented into AREPO-RT in the context of this thesis. Specifically, the key functions modeling redshifting are contained in the newly added file `RT_redshift.cc`, and the inclusion of redshifting in the code is controlled by the compile-time flag `MRT_REDSHIFT` in AREPO's configuration file `Config.sh`. Additional modifications were made to the following previously existing files:

(i) `RT_init.cc` (see Sec. 3.3), to compute the average photon frequency in each bin (Eq. (87)) once in the RT initialization routine.

(ii) `RT_run.cc`, the main file containing the steps taken in each RT timestep during a run (e.g., thermochemistry calculations) in the CPU version of AREPO-RT (Kannan et al. 2019). The main redshifting function is called in this file at each *hydrodynamical* timestep, provided `MRT_REDSHIFT` is active. We note that we implement redshifting on the hydrodynamical timesteps, rather than the RT timesteps, as redshifting only has a meaningful impact on photon energies on large timescales ($\gg \mathcal{O}(\Delta t_{\mathrm{RT}})$).

(iii) `gpu/RT_gpu_run.cu`, written in the CUDA programming language for the GPU-accelerated version of AREPO-RT (Zier et al. 2024); roughly speaking, the GPU equivalent of `RT_run.cc`. This file calls the main redshifting function in runs performing the RT calculations on GPUs. We note, however, that the computations specific to the redshifting implementations are nevertheless carried out on CPUs in such runs; they are merely integrated into the larger RT GPU implementation in this file.

(iv) `RT.h`, a header file modified to include the declaration of the main redshifting function.

(v) `src/allvars.cc` and `src/allvars.h` (see Sec. 3.3), declaring redshifting variables to ensure they are accessible across all files of the code.

(vi) `src/timer.h`, creating a timer (`CPU_RT_REDSHIFT`) implemented to measure the computational cost of the redshifting methods. This primarily served as a numerical check confirming that neither redshifting method significantly increases the computational cost of simulations with AREPO-RT.

Finally, the *Makefiles* controlling the compilation of the code were modified to integrate the new file `RT_redshift.cc` into the rest of the code base in `makefiles/modules.make` (cf. Sec. 3.3).

In this section, we first discuss a numerical redshifting method based on a piecewise power-law (PPL) approximation of the photon spectrum, as well as our implementation of the method into AREPO-RT.
This PPL approach is primarily based on an existing method proposed by Finlator et al. (2009). It has previously been successfully used to capture the redshifting evolution of an inhomogenous UV background, resolved by 24 frequency bins in the EUV range $[13.6\,\mathrm{eV}, 136\,\mathrm{eV}]$, in the *Technicolor Dawn* reionization simulations (Finlator et al. 2018). We caution, however, that the typical THESAN setup using AREPO-RT has a substantially lower frequency resolution, covering a comparable photon energy range in the EUV with only 3 frequency bins (Sec. 3.3.1).

3.4.1 Comoving rescaling of the cosmological RT equations

As derived and described in detail in Sec. 2.3, the most general RT equation including cosmological terms is given by Eq. (31), where the terms proportional to the Hubble parameter H are of a cosmological nature: the term proportional to $3I_\nu$ accounts for the dilution of the number density of photons, while the term proportional to $\nu\frac{\partial I_\nu}{\partial \nu}$ describes the effect of ongoing cosmological redshifting.
Implementing the latter in numerical radiative transfer is challenging, since it requires information about the photon spectrum (i.e., the radiative intensity as a function of frequency), while numerical RT solvers necessarily only track a discretized version of the spectrum in a finite number of frequency bins (see e.g. Fig. 7; Sec. 3.3.1).

The cosmological dilution term, on the other hand, can be implemented more simply by a rescaling of the specific intensity and emission coefficient (Wu et al. 2019; Kannan et al. 2022a). By defining the *rescaled*, or *comoving*, specific intensity as $I'_\nu \equiv a^3 I_\nu$, and the rescaled emission coefficient as $j'_\nu \equiv a^3 j_\nu$, the partial derivative of I_ν with respect to time t in Eq. (31) can be rewritten as

$$\frac{\partial I_\nu}{\partial t} = \frac{\partial}{\partial t}\left(a^{-3} I'_\nu\right) = -3a^{-3} H I'_\nu + a^{-3}\,\frac{\partial I'_\nu}{\partial t}\,. \tag{72}$$

After inserting Eq. (72) into Eq. (31) and expressing the result in terms of the rescaled quantities, the modified general RT equation follows,

$$\frac{1}{c}\frac{\partial I'_\nu}{\partial t} + \frac{\mathbf{n}}{a}\cdot\nabla I'_\nu - \frac{H}{c}\cdot\nu\,\frac{\partial I'_\nu}{\partial \nu} = j'_\nu - \kappa_\nu \rho I'_\nu\,. \tag{73}$$

A direct comparison of Eqs. (31) and (73) shows that this rescaling automatically accounts for the cosmological dilution of the specific intensity by substituting I_ν and j_ν with the comoving quantities I'_ν and j'_ν; only the classical terms and the cosmological redshifting term, now $\propto \nu \frac{\partial I'_\nu}{\partial \nu}$, remain in the equation.

Using the definition of the rescaled specific intensity given above and the fact that the cosmological scale factor $a = a(t)$ is independent of the direction $\mathbf{n}$, it follows that the rescaled angular moments of I_ν (see Sec. 3.2) are simply given by $E'_\nu = a^3 E_\nu$, $\mathbf{F}'_\nu = a^3 \mathbf{F}_\nu$, and $\mathbb{P}'_\nu = a^3 \mathbb{P}_\nu$. Similarly, the form of the rescaled source function $S'_\nu = a^3 S_\nu$ follows from the definition of the rescaled emission coefficient $j'_\nu = a^3 j_\nu$.
With this, the zeroth and first angular moments of the RT equation with cosmological terms and comoving rescaling can be computed analogously to the classical case (Eqs. (39)-(40)),

$$\frac{\partial E'_\nu}{\partial t} + \frac{1}{a} \nabla \cdot \mathbf{F}'_\nu - H\nu \frac{\partial E'_\nu}{\partial \nu} = S'_\nu - \kappa_\nu \rho E'_\nu \tag{74}$$

$$\frac{\partial \mathbf{F}'_\nu}{\partial t} + \frac{c^2}{a} \nabla \mathbb{P}'_\nu - H\nu \frac{\partial \mathbf{F}'_\nu}{\partial \nu} = -\kappa_\nu \rho \mathbf{F}'_\nu \, . \tag{75}$$

Finally, in the frequency bin convering the range $[\nu_i, \nu_{i+1}]$, the rescaled cosmological RT moment equations are then given by (cf. Eqs. (56) and (57))

$$\frac{\partial E'}{\partial t} + \frac{1}{a} \nabla \cdot \mathbf{F}' - H \int_{\nu_i}^{\nu_{i+1}} \nu \frac{\partial E'_\nu}{\partial \nu} \, \mathrm{d}\nu = S' - \kappa_E \rho E' \tag{76}$$

$$\frac{\partial \mathbf{F}'}{\partial t} + \frac{c^2}{a} \nabla \mathbb{P}' - H \int_{\nu_i}^{\nu_{i+1}} \nu \frac{\partial \mathbf{F}'_\nu}{\partial \nu} \, \mathrm{d}\nu = -\kappa_F \rho \mathbf{F}' \, , \tag{77}$$

where E', $\mathbf{F}'$, and $\mathbb{P}'$ are defined analogously to their non-rescaled counterparts (Sec. 3.3.1) as the frequency integrals of E'_ν, $\mathbf{F}'_\nu$, and $\mathbb{P}'_\nu$. Eqs. (76) and (77) represent the form of the equations we aim to numerically solve in AREPO-RT in the context of this redshifting method.

3.4.2 Piecewise power-law spectrum approximation

To compute the redshifting terms in Eqs. (76) and (77) given only the values for the total energy density, flux, and radiation pressure in each bin, an additional assumption about the intensity spectrum is necessary, as the partial derivative with respect to ν cannot be directly computed from these discrete values alone.
For this purpose, we assume a PPL spectrum for the intensity, i.e. $I_\nu \propto \nu^{-\alpha}$, where the spectral index α is computed separately in each frequency bin and at each point in discretized space-time, i.e. in each cell, at each hydrodynamical timestep. Generally, a PPL spectrum is expected to be a good approximation for realistic photon spectra for an appropriate choice of the frequency bin edges at the typical sharp transition regions of the

spectrum for UV photons (e.g. Baumschlager et al. 2023), i.e., at the ionization thresholds of H I, He I, and He II.

The PPL approximation substantially simplifies the redshifting term (Finlator et al. 2009, 2018). Assuming an isotropic spectral index α, it follows that the spectra of the comoving energy density and flux are also described by a power law in each frequency bin, $E'_\nu \propto \nu^{-\alpha}$ and $\mathbf{F}'_\nu \propto \nu^{-\alpha}$. Setting $E'_\nu \equiv C\nu^{-\alpha}$, where C is an arbitrary proportionality constant, the redshifting term of Eq. (76) becomes proportional to

$$\begin{aligned}\int_{\nu_i}^{\nu_{i+1}} \nu \frac{\partial E'_\nu}{\partial \nu} \,\mathrm{d}\nu &= \int_{\nu_i}^{\nu_{i+1}} \nu \frac{\partial}{\partial \nu} \left(C\nu^{-\alpha}\right) \,\mathrm{d}\nu \\ &= \int_{\nu_i}^{\nu_{i+1}} \nu \cdot \left(-\alpha C \nu^{-\alpha-1}\right) \,\mathrm{d}\nu \\ &= -\alpha \int_{\nu_i}^{\nu_{i+1}} C\nu^{-\alpha} \,\mathrm{d}\nu \\ &= -\alpha \int_{\nu_i}^{\nu_{i+1}} E'_\nu \,\mathrm{d}\nu \\ &= -\alpha E' \,. \end{aligned} \tag{78}$$

Using $\mathbf{F}'_\nu \propto \nu^{-\alpha}$, it follows analogously that

$$\int_{\nu_i}^{\nu_{i+1}} \nu \frac{\partial \mathbf{F}'_\nu}{\partial \nu} \,\mathrm{d}\nu = -\alpha \mathbf{F}'_\nu \,. \tag{79}$$

Therefore, the PPL approximation reduces Eqs. (76) and (77) to

$$\frac{\partial E'}{\partial t} + \frac{1}{a} \nabla \cdot \mathbf{F}' = S' - \left(\kappa_E \,\rho\, c + H\alpha\right) E' \tag{80}$$

$$\frac{\partial \mathbf{F}'}{\partial t} + \frac{c^2}{a} \nabla \,\mathbb{P}' = -\left(\kappa_F \,\rho\, c + H\alpha\right) \mathbf{F}' \,. \tag{81}$$

The redshifting term is fully described by the spectral index α in this approximation. The numerical implementation of Eqs. (80) and (81) is straightforward and analogous to isotropic absorption (note the formal equivalence of the cosmological terms $\propto H$ to the absorption terms $\propto \kappa_{E/F}$).

Specifically, we implemented this approach in the function `mrt_redshift_photons_PPL()` in AREPO-RT, defined within the newly added file `RT_redshift.cc`. To ensure numerical stability, we employ an implicit scheme (e.g. Press & Teukolsky 2007, p. 931f.): we discretize

$$\left.\frac{\mathrm{d}E'}{\mathrm{d}t}\right|_{\text{redshift}} = -H\alpha E' \tag{82}$$

from Eq. (80) as

$$\Delta E' \equiv E'_{t+\Delta t} - E'_t = -\Delta t \cdot H\alpha E'_{t+\Delta t} \,, \tag{83}$$

where Δt denotes the numerical timestep and $\Delta E'$ the associated update of E' during this timestep. The implicit formulation is obtained by evaluating the right hand side of Eq. (83) in terms of the *updated* value $E'_{t+\Delta t}$. This updated value is therefore computed numerically as

$$\begin{aligned} E'_{t+\Delta t}\,(1 + \Delta t\, H\alpha) &= E'_t \\ \Rightarrow E'_{t+\Delta t} &= \frac{E'_t}{1 + \Delta t\, H\alpha}\,. \end{aligned} \tag{84}$$

We note that this ensures that $E'_{t+\Delta t} \to 0$ for $\Delta t \to \infty$, rather than approaching unphysical negative energy densities.
Finally, each component F'_j $(j = 1, \ldots, 3)$ of the flux $\mathbf{F}'$ is updated analogously,

$$F'_{j,\,t+\Delta t} = \frac{F'_{j,\,t}}{1 + \Delta t\, H\alpha}\,, \tag{85}$$

resulting in an implementation formally equivalent to *isotropic* absorption.

The key remaining challenge is therefore to obtain an estimate for the power-law spectral index α from the discrete, low-resolution spectrum tracked in AREPO-RT (see e.g. the example spectrum plotted in Fig. 7).
We approach this for each frequency bin i with associated energy density E'_i by considering the change towards the energy density E'_{i+1} of the next-higher frequency bin $i + 1$. Specifically, the power-law spectral index α_i for bin i is computed as the negative slope in log-space:

$$\alpha_i = -\frac{\log E'_{i+1} - \log E'_i}{\log \overline{\nu_{i+1}} - \log \overline{\nu_i}}\,. \tag{86}$$

Here, $\overline{\nu}$ refers to the average of the frequencies in each bin, i.e. for bin i defined by the frequency range $[\nu_i, \nu_{i+1}]$,

$$\overline{\nu_i} = \frac{\nu_i + \nu_{i+1}}{2}\,. \tag{87}$$

This approach is sketched schematically in Fig. 11.
The reasoning behind considering only the slope towards the next-higher frequency bin, rather than e.g. the average value of the slopes towards the next-lower *and* next-higher frequency bin, is that the impact of redshifting at any given frequency is physically independent of the spectrum below this frequency, since redshifting exclusively *decreases* photon energies.

However, we caution in this context that the above approach to estimating the spectral index α as the power-law slope *between* frequency bins does not necessarily capture the slope of the spectrum *within* the bins accurately. The latter would be closer to the canonical meaning of a 'spectral index' in the context of a power-law spectrum, but is

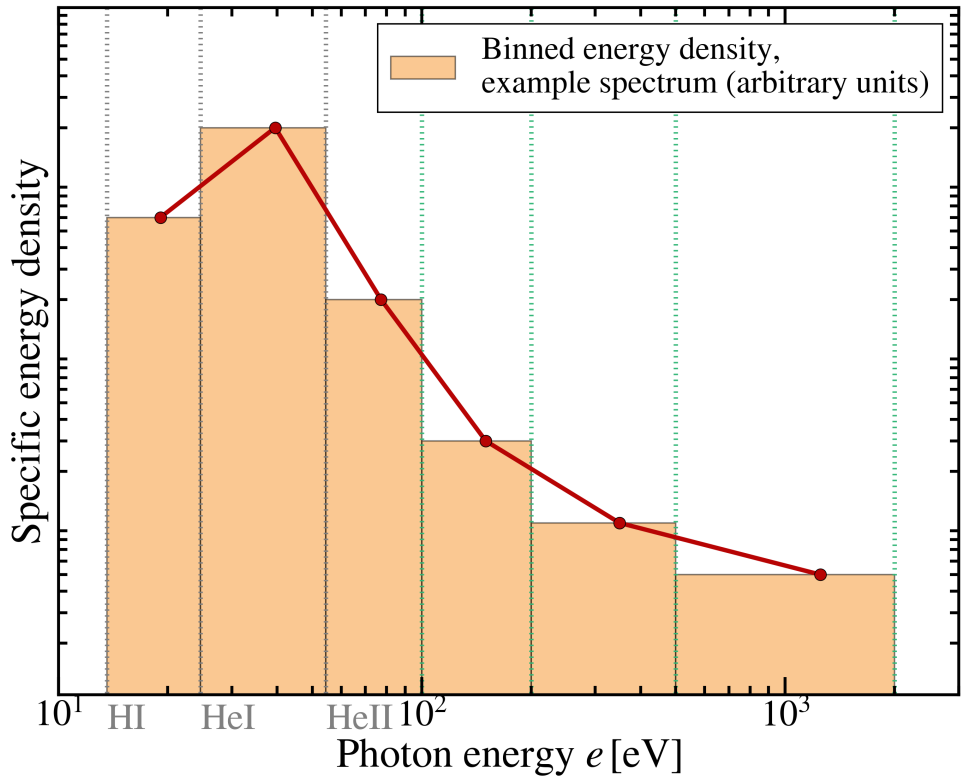

Figure 11: Illustration of the computation of the spectral index α from the piecewise-constant description of the energy density in AREPO-RT, using the same example spectrum in arbitrary units as presented in Fig. 7.

unfortunately fundamentally inaccessible from a piecewise-constant, binned description of the spectrum; in other words, no spectral information is tracked by the code within the bins.

An exception to the above approach of computing α based on the energy density of the next-higher bin (Eq. (86)) is necessarily the highest-energy bin. For this case, we instead use a modified description that assumes a linear decline in photon energy density from the low-energy edge of the bin towards zero at the high-energy edge. This approach is motivated by (i) the implicit assumption that the spectral intensity approaches zero for photon energies above those tracked within the highest-energy bin, and (ii) the fact that the latter assumption cannot be described by the above power-law approximation, as $\lim_{E' \to 0^+} \log E' = -\infty$; the linear decline therefore serves as a simple alternative model of a gradual decline towards zero. We note that, of course, the validity of the assumption made in point (i) heavily depends on the modeled sources and frequency range in a practical numerical setup.

The specific form of this linear description is set by the constraint that the total energy density in the highest-energy bin, computed by integrating over the linearly declining spectrum, must be equal to the value for E' tracked by the code, in combination with the condition that $E'_\nu = 0$ at the high-energy end. This again results in the moment equations

of the same form as Eqs. (80) and (81), however with a value of α_{lin} now given by

$$\alpha_{\mathrm{lin}} = \frac{\nu_{j_2} + \nu_{j_1}}{\nu_{j_2} - \nu_{j_1}} \,. \tag{88}$$

Here, j_1 and j_2 refer to the bin edges of the highest-energy bin j.
We stress that this is only one possible option of handling redshifting in the highest-energy bin; in a practical application of the PPL redshifting method, other methods may be better-suited, depending on the spectra and abundances of the sources of radiation considered, as well as the definition of the frequency bin ranges. Other potential options include, for instance, (i) simply neglecting redshifting in the highest bin (equivalent to assuming a flat spectrum in this range), (ii) setting the spectral index equal to that of the second-highest bin (equivalent to assuming a single power-law spectrum across the second-to-highest and highest frequency bins), or (iii) determining α based on an approximate spectral index of the SED of a dominant source at high energies (e.g., inferring a spectral index from the HMXB SED in the soft X-ray range).

In light of the latter option, we note that it is unfortunately not *generally* a good approximation to simply use the spectral slope of source spectra as an estimate for α. The reason for this is illustrated in Fig. 12: as EUV radiation passes through (partially) neutral gas, it typically undergoes significant spectral hardening (e.g. Maselli et al. 2009) due to the strong energy dependence of the photoionization cross-sections of H I, He I and He II (see Fig. 3), as photons with energies closer to the respective ionization thresholds get absorbed significantly faster.
This spectral hardening can substantially alter the shape of the SED and, in turn, the spectral index. For example, it is clear from Fig. 12 that a constant α based on the source spectrum (top curve) would in this case misrepresent the actual spectral shape within a distance of only a few cells from the source (see Maselli et al. 2009 for details on the underlying RT implementation). In fact, this approach would even result in the wrong sign of α within the first few bins as defined in the THESAN model (Sec. 3.3.1).
However, we note that it may still be a well-justified approximation for bins at higher photon energies, such as those of the soft X-ray band and beyond, due to the significantly lower photoionization cross-sections of hydrogen and helium in this range (see Sec. 2.1; Table 3). In the scope of this thesis, we leave a more detailed investigation of this potential approach to future work.[10]

[10]The primary reasons we did not explore this further in the context of this thesis are unrelated fundamental limitations of PPL-based redshifting in the THESAN model, which are discussed in detail in Secs. 3.4.3 and 4.2.

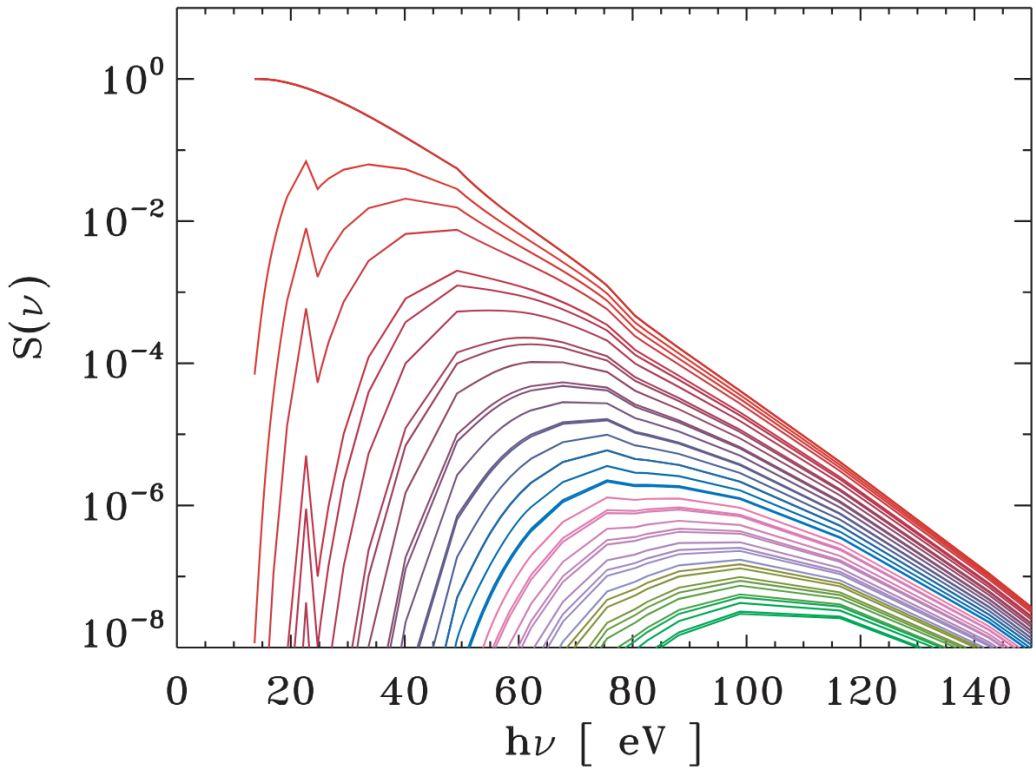

Figure 12: Evolution of the spectral shape of radiation passing through neutral hydrogen and helium gas, illustrating spectral hardening near the ionization thresholds of H I and He I. Specifically, the curves show the spectral energy distribution of a photon packet (cf. Sec. 3.2) passing through gas cells simulated with the CRASH2 code (see Maselli et al. 2003, 2009), with the curves ordered from top to bottom in the cell crossing order. This figure was obtained from Maselli et al. (2009).

3.4.3 Limitations of the approach and slope limiter for the spectral index

There are several significant numerical challenges and limitations associated with the PPL approach described in the sections above, most of which fundamentally boil down to the low frequency resolution of the AREPO-RT code in the THESAN model. We note that the latter is unfortunately a difficult limitation to resolve in the current setup, as the computational cost of RHD simulations with AREPO-RT scales approximately linearly with the number of frequency bins; i.e., it is not possible to significantly improve the frequency resolution without dramatically increasing the computational cost.

The first limitation of the approach is the lack of information about the spectral index *within* each bin, as discussed in more detail above.
Secondly, perhaps the most fundamental drawback is that an attempted reconstruction of the photon spectrum based on only a few values of the energy density in broad bins, as illustrated in Fig. 11, is unlikely to capture the true shape and details of the spectrum accurately. This necessarily leads to a potentially large uncertainty in the value of α.

Finally, the power-law approximation for α_i breaks down for very low energy densities in either bin i or the next-higher bin $i+1$. This can be made clear from Eq. (86), as $E'_i \to 0$ and/or $E'_{i+1} \to 0$ lead to a divergent α_i. AREPO-RT does set a lower limit for the minimum photon energy density, `MINDENSPHOT` $\equiv 10^{-25}$ in internal energy density units[11], to avoid floating point exceptions caused by energies approaching zero and by numerical uncertainty in the RT solver. However, this low value does not prevent α_i from reaching high values extremely sporadically (as also observed by Finlator et al. 2018 in the context of this method), which are fundamentally associated with a large numerical uncertainty due to the low energy densities involved.

In an attempt to mitigate the impact of the latter problem, we follow the approach of Finlator et al. (2018) and introduce a *slope-limiter* for the spectral index by enforcing $|\alpha| \leq 10$ in the code.[12]
The impact of this slope limiter is investigated and discussed in further detail in the context of test runs presented in Sec. 4.2.

3.5 Effective energy-rebinning approach to cosmological redshifting

Due to the aforementioned limitations of the PPL method, we explore a fundamentally different approach to numerical redshifting in the following. This alternative approach is, to the best of our knowledge, a conceptually new method developed in the context of this thesis, which we refer to here as 'effective energy-rebinning' ('EER').
This method has been implemented in the same file as the PPL redshifting approach in AREPO-RT, i.e., `RT_redshift.cc`. The integration into the larger code base and the additional modified files are therefore identical for the two redshifting methods (see the description provided in Sec. 3.4).

The central idea of the method is that the gradual 'shift' of the photon spectrum, progressing towards lower energies due to cosmological redshifting, can equivalently be modeled by effectively shifting the energy *bins* towards *higher* energies. In other words, the approach is to imagine keeping the photon spectrum – which, physically, would be shifted by the energy $\Delta e_{\mathrm{rs},\gamma} = -h\,|\Delta\nu_{\mathrm{rs}}(\nu)|$ – fixed, but shifting the bin edges by the opposite amount, $\Delta e_{\mathrm{rs,bins}} = h\,|\Delta\nu_{\mathrm{rs}}(\nu)|$, effectively rebinning the spectrum. We note that we refer to this as only an 'effective' rebinning, since the *physical* values of the energy bin edges (e.g., 13.6 eV, 24.59 eV, etc.) remain fixed in the implementation.

[11] The internal energy density units correspond to a physical value of $[E] = 6.18 \times 10^{52}\,\mathrm{erg\,ckpc^{-3}}$ in all test runs discussed below (Sec. 4.2).

[12] We note that this is a slightly more strict condition than employed in Finlator et al. (2018), who set only the lower limit $\alpha \geq -10$.

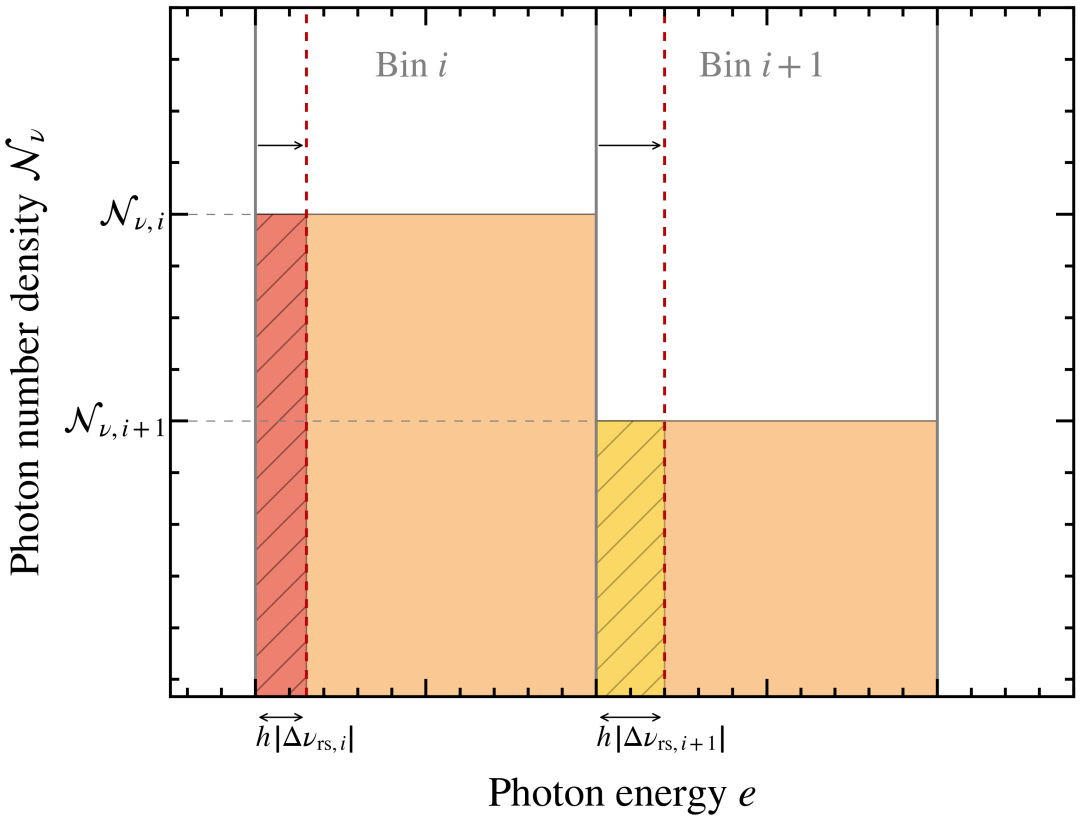

Figure 13: Schematic illustration of the effective energy-rebinning approach to numerical redshifting, showing a binned sample spectrum (orange bars) in arbitrary units for two adjacent frequency bins, i.e., bin i and bin $i+1$ with associated specific photon number densities $\mathcal{N}_{\nu,i}$ and $\mathcal{N}_{\nu,i+1}$. The solid gray vertical lines indicate the bin edges, while the dashed red vertical lines illustrate the effective 'rebinning' for bin i, shifting the bin edges by the photon energies $h\,|\Delta\nu_{\mathrm{rs},i}|$ and $h\,|\Delta\nu_{\mathrm{rs},i+1}|$ (see text for details). Finally, the hatched red and yellow shaded areas represent the total photon number density redshifted out of and into bin i, respectively.

This idea is sketched schematically in Fig. 13, where the gray vertical lines show the original bin edges, and the red vertical lines represent the imagined 'shifted' bins.
Here, the size of the energy shift Δe_{rs} of the bins is simply determined by the size of the timestep Δt, since the relative change in photon frequency due to redshifting is equal and opposite to the relative change in the scale factor a (cf. Eq. (27)), i.e.,

$$\frac{\Delta\nu_{\mathrm{rs}}}{\nu} = -\frac{\Delta a(\Delta t)}{a(t)}\,. \tag{89}$$

Noting that $\Delta\nu < 0$ and $\Delta a > 0$ in an expanding Universe, it follows that the energy shift of the bins is given by

$$\Delta e_{\mathrm{rs,bins}}(\nu) = h\,|\Delta\nu_{\mathrm{rs}}(\nu)| = h\,\frac{\nu\,\Delta a}{a}\,. \tag{90}$$

To estimate the amount of energy that should therefore be transferred into and out of bin i, we first switch from the previously used description of the radiation field in terms of the energy density E_i to a description in terms of the *photon number density* $\mathcal{N}_i$. This step is taken in order to account for the fact that redshifting is a photon number-conserving process, rather than an energy-conserving process.
We obtain the approximate (comoving) photon number density by dividing the comoving energy density by the mean photon energy e_i in bin i,

$$\mathcal{N}_i \approx \frac{E'_i}{e_i} \ . \tag{91}$$

Analytically, $\mathcal{N}_i$ is given by the integral of the monochromatic photon number density $\mathcal{N}_\nu$ over the bin's frequency range $[\nu_i, \nu_{i+1}]$ (cf. Eq. (54)), i.e., by the area under the curve of the spectrum $\mathcal{N}_\nu = \mathcal{N}_\nu(\nu)$. However, only $\mathcal{N}_i$ can be obtained by E'_i, which is tracked by AREPO-RT; neither the exact shape of the spectrum, nor the values of $\mathcal{N}_\nu$ at the frequency bin edges (labeled as $\mathcal{N}_{\nu,i}$ and $\mathcal{N}_{\nu,i+1}$ in Fig. 13), are known during a run.
To still estimate $\mathcal{N}_{\nu,i}$ and $\mathcal{N}_{\nu,i+1}$ for the sake of the redshifting calculation, we therefore approximate the spectrum as piecewise-constant, as illustrated by the rectangular, orange-shaded bars of Fig 13. $\mathcal{N}_i$ is then represented by the areas of these orange-shaded bars, covering the frequency range $[\nu_i, \nu_{i+1}]$; consequently, $\mathcal{N}_{\nu,i}$ (represented by the height of the bars) is estimated in our implementation as

$$\mathcal{N}_{\nu,i} = \frac{\mathcal{N}_i}{\nu_{i+1} - \nu_i} \ . \tag{92}$$

With this, we can then compute the amount of photons *lost* from bin i as the area under the spectral curve – estimated as $\mathcal{N}_{\nu,i} = \text{const.}$ – in the frequency interval given by $[\nu_i, \nu_i + |\Delta\nu_{\mathrm{rs},i}|\,]$, i.e.,

$$\Delta\mathcal{N}_{i,\mathrm{lost}} = \mathcal{N}_{\nu,i} \cdot h \ |\Delta\nu_{\mathrm{rs},i}| \ . \tag{93}$$

This is illustrated by the red hatched rectangle in Fig. 13.
Further, we set the amount of photons *gained* by bin i as equal to the amount of photons *lost* by bin $i+1$ (yellow hatched rectangle in Fig. 13),

$$\Delta\mathcal{N}_{i,\mathrm{gained}} = \Delta\mathcal{N}_{i+1,\mathrm{lost}} = \mathcal{N}_{\nu,i+1} \cdot h \ |\Delta\nu_{\mathrm{rs},i+1}| \ . \tag{94}$$

Crucially, this ensures that our approach is fundamentally photon number-conserving.
Converting the photon number densities back to energy densities in a final step, the net effect of redshifting on E'_i in our implementation can be summarized as follows for a timestep $t \to t + \Delta t$:

$$E'_{i,\,t+\Delta t} = E'_{i,\,t} - \frac{\Delta\mathcal{N}_{i,\mathrm{lost}} - \Delta\mathcal{N}_{i,\mathrm{gained}}}{e_i} \ . \tag{95}$$

We note that we enforce all updated values of E'_i to remain at or above the set value for the numerical minimum energy density `MINDENSPHOT` (Sec. 3.4.3) by setting $E' =$ `MINDENSPHOT` if Eq. (95) results in $E' <$ `MINDENSPHOT`.
Finally, the comoving flux $\mathbf{F'_i}$ in bin i is updated isotropically by the same ratio as the energy density,

$$\mathbf{F'}_{\mathbf{i},\,t+\Delta t} = \mathbf{F'}_{\mathbf{i},\,t} \cdot \frac{E'_{i,\,t+\Delta t}}{E'_{i,\,t}} \,. \tag{96}$$

During this step, we ensure that the reduced flux f (Eq. (50)) still obeys the condition $0 \leq f \leq 1$ (Eq. (51); see Sec. 3.2 for details) in each bin after the update of the flux. This is done by first computing the value f_{updated} that would follow from updating $\mathbf{F'_i}$ as specified in Eq. (96); we then proceed with the flux update if $f_{\text{updated}} \leq 1$. Otherwise, we instead modify the ratio given in Eq. (96) by dividing by f_{updated}, i.e.,

$$\mathbf{F'}_{\mathbf{i},\,t+\Delta t} = \mathbf{F'}_{\mathbf{i},\,t} \cdot \frac{E'_{i,\,t+\Delta t}}{E'_{i,\,t}} \cdot \frac{1}{f_{\text{updated}}} \quad , \text{ if } f_{\text{updated}} > 1 \,. \tag{97}$$

Similarly to the PPL method, there are different options for handling the highest-energy bin in this treatment of redshifting. In the test runs presented in Sec. 4.3, we chose to neglect any radiative energy shifted *into* the highest bin, instead only modeling its energy *loss* to the second-to-highest bin due to redshifting. This is conceptually roughly equivalent to the approach taken for the PPL method, in the sense that it also neglects all photons that carry energies exceeding the range of the bins, and that are therefore not tracked by the code.
One alternative option is to instead assume that an equal amount of energy is shifted into and out of the highest bin, i.e., effectively assuming a flat spectrum in the corresponding energy range. This is the most straightforward approach to implement, as it simply results in no net gain or loss of energy in the highest bin.
More generally, the best way to approach the highest-energy bin depends crucially on the setup of the simulation or problem under consideration – in particular, on the source SEDs and modeled frequency range – as also argued for the PPL method in Sec. 3.4.2.

4 Test run results

4.1 Effects of high-redshift X-ray sources

4.1.1 Test run setup

In the following, we present results from RHD simulations of the EoR, testing the X-ray implementation developed in the context of this thesis and investigating the impact of high-redshift X-ray sources on the thermal and ionization histories in detail. To this end, we ran two separate test runs with AREPO-RT up to a final redshift of $z = 5$.

The baseline run, also referred to as the *'No X-rays'* run in the following, uses only the original AREPO-RT code in the THESAN setup with three EUV-range energy bins, covering photon energies up to 100 eV. It only includes stars and AGN as radiative sources, further using the Lusso et al. (2015) parametrization for the AGN spectrum at lower photon energies (cf. Sec. 3.3.5). We note that in this EUV range, the latter is very similar to the Shen et al. (2020) parametrization employed in the run below (see Fig. 1 of Shen et al. 2020).
The run testing the new X-ray implementation – also referred to as the *'X-ray run'* or simply as *'With X-rays'* in figure headers in this section – further uses the methods described in Sec. 3.3. While it is based on the THESAN model and ran with the AREPO-RT code like the *No X-rays* run, it extends the covered spectrum as described in Sec. 3.3.1, modeling photon energies up to 2 keV in six energy bins. In addition to stars and AGN, it includes the hot, diffuse ISM and HMXBs as X-ray sources as discussed in Secs. 3.3.3 and 3.3.4, respectively. Further, the AGN source SED is updated to the Shen et al. (2020) parametrization (Sec. 3.3.5).

The ICs of both runs are identical and based on the Planck Collaboration et al. (2020b) cosmological parameters, set up with GADGET-4 using a gravitational glass initialization as described in Sec. 3.1.1. The simulation models a cubic volume with periodic boundary conditions; the box size L_{box}, describing the length of the cubic volume on a side, is $L_{\mathrm{box}} = 25\,\mathrm{cMpc}$. This is resolved by $N_{\mathrm{DM}} = 300^3 = 27\,000\,000$ dark matter particles, resulting in a mass resolution of $m_{\mathrm{DM}} = 1.92 \times 10^7\,\mathrm{M}_\odot$.

4.1.2 Global reionization history and temperature evolution

To provide an overview of the progression of reionization, as well as the mean gas temperature in the simulation volume, we first present the ionization and thermal histories for both the *X-rays* and *No X-rays* runs in Fig. 14. Specifically, we show the evolution of the volume-averaged neutral hydrogen fraction and gas temperature throughout the EoR. The former shows excellent agreement between both the *X-rays* and *No X-rays* runs. The

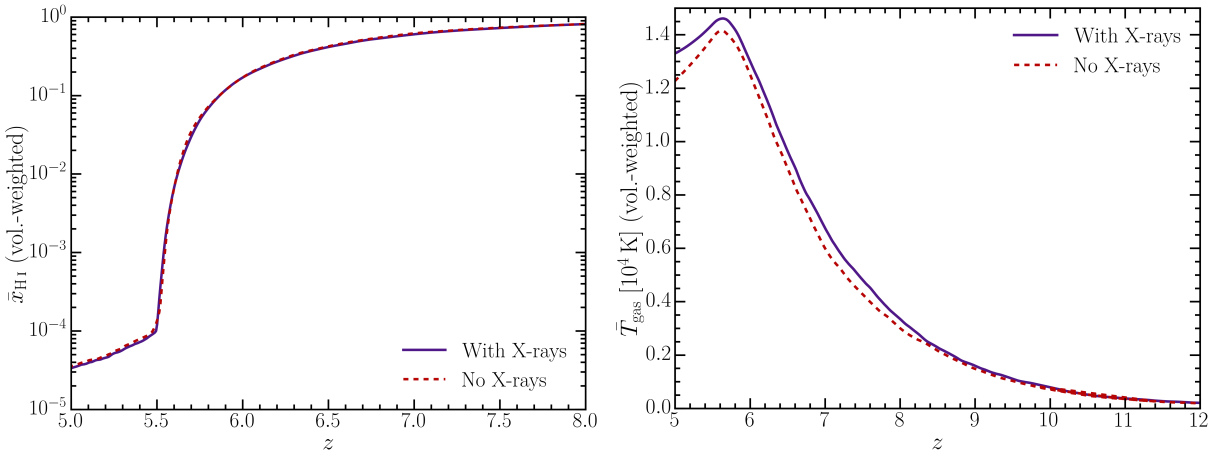

Figure 14: Volume-averaged fraction of neutral hydrogen $\bar{x}_{\mathrm{H\,I}}$ (left panel) and gas temperature $\bar{T}_{\mathrm{gas}}$ (right panel) as a function of redshift for both the *X-rays* (purple solid curves) and *No X-rays* (red dashed curves) simulations.

mean gas temperature shows a tendency towards slightly higher values in the *X-rays* run (see also Sec. 4.1.3); nevertheless, its overall evolution agrees fairly well between both runs. We note that both of these findings are expected physically, as the global progression of reionization is dominated by the EUV emissions of stars (e.g. Kannan et al. 2019, 2022a). Therefore, the neutral fraction and temperature evolution for *both* runs shown in Fig. 14 are broadly in agreement with typical simulated reionization histories (see e.g. Figs. 4 and 6 of Kannan et al. 2022a for the analogous figures based on the original THESAN simulation suite). These quantities first show a gradual progression towards a lower neutral fraction and higher temperatures, respectively, with both of these trends accelerating towards lower redshifts. The steepest drop-off in neutral fraction occurs approximately between redshifts $z \sim 6$ and $z \sim 5.5$. This corresponds to a *late reionization scenario* (e.g. Kannan et al. 2022a), i.e., there is still a significant fraction of neutral gas left below redshifts $z \lesssim 6$ in our simulation volume. Further, during this period, the global amount of photoheating reaches its peak due to the frequent ionization events (cf. Sec. 2.1); consequently, the maximum mean gas temperatures of $\bar{T}_{\mathrm{gas}} \approx 1.45 \times 10^4\,\mathrm{K}$ (*X-rays* run) and $\bar{T}_{\mathrm{gas}} \approx 1.4 \times 10^4\,\mathrm{K}$ (*No X-rays* run) are reached at $z \sim 5.6$.

Finally, the end of reionization is reached approximately at redshift $z \sim 5.5$ to $z \sim 5$, resulting in a final neutral hydrogen fraction of $\bar{x}_{\mathrm{H\,I,final}} \approx 3 \times 10^{-5}$. The final mean temperature shows a mild discrepancy between the *X-rays* and *No X-rays* runs, with $\bar{T}_{\mathrm{gas}} \approx 1.33 \times 10^4\,\mathrm{K}$ for the *X-rays* run and $\bar{T}_{\mathrm{gas}} \approx 1.23 \times 10^4\,\mathrm{K}$ for the *No X-rays* run at $z = 5$. This temperature difference and profile is discussed further in the following section.

4.1.3 X-ray heating of the intergalactic medium

A key signature of X-ray sources is a fairly uniform heating of the IGM during the EoR, due to the ability of X-ray photons to propagate further outside of the ionized bubbles

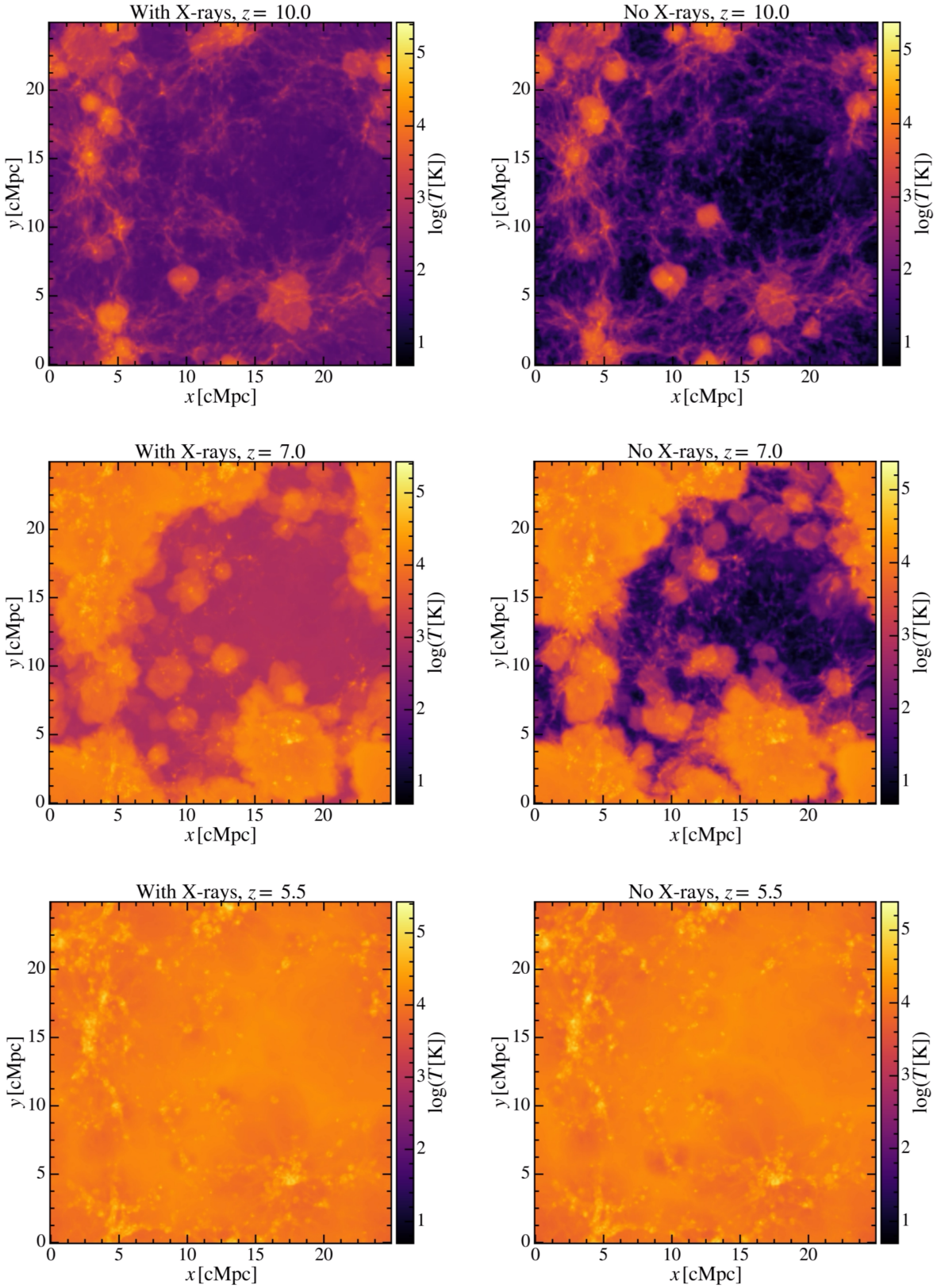

Figure 15: X-ray heating of the intergalactic medium throughout the Epoch of Reionization, illustrated by the simulated gas temperature distribution projected onto a slice through the simulation volume of the *X-rays* (left panels) and *No X-rays* (right panels) runs at redshifts $z = 10$, $z = 7$, and $z = 5.5$. Comparing the left and right panels illustrates qualitatively that our X-ray model successfully captures the characteristic uniform heating of the intergalactic medium.

than EUV-range photons, as well as their potential to cause more significant heating per ionization event (cf. Sec. 2.1; Table 4). As also noted in Sec. 2.1, this characteristic X-ray heating in particular is predicted to be discernible in the large-scale 21-cm power spectrum (Mesinger et al. 2013).

We first show qualitatively how our X-ray implementation captures this expected effect in Fig. 15, showing projections of the gas temperature onto a slice of the simulation volume of both the *X-rays* and *No X-rays* runs. The projections are created with a depth of 14 %, i.e., within a cuboid of the simulated box with side lengths $(\Delta x \times \Delta y \times \Delta z) = (L_{\mathrm{box}} \times L_{\mathrm{box}} \times 0.14\, L_{\mathrm{box}})$; this slice shows a representative sample of the spatial distribution of the simulated gas temperature.
It is clear from these temperature distributions that the inclusion of the X-ray sources has a substantial effect during the EoR. Firstly, at a redshift of approximately $z \sim 10$, the comparison to the *No X-rays* run shows that the simulation volume begins to heat up uniformly in the *X-rays* run.
This effect increases as reionization progresses; in particular, the IGM (visible in the $z \sim 10$ and $z \sim 7$ *No X-rays* panels of Fig. 15 as the black and dark purple regions; see also Fig. 17) is heated to temperatures $T \approx \mathcal{O}(10^2 - 10^3\,\mathrm{K})$ in the *X-rays* run. On the other hand, significant regions of the IGM stay cold ($T \lesssim \mathcal{O}(10\,\mathrm{K})$), even until $z \sim 7$, in the run not including the X-ray implementation.
Finally, towards the end of reionization at $z \sim 5.5$, the spatial distribution of heated gas roughly converges between the *X-rays* and *No X-rays* runs again. By this redshift, the simulation volumes of both runs are heated nearly uniformly to a final mean temperature $\bar{T} \approx \mathcal{O}(10^4\,\mathrm{K})$ (cf. Sec. 4.1.2).

We present the effects of the X-ray implementation in further quantitative detail in Fig. 16. At each point, these histograms show the fraction of cells simulated with arepo-rt as a function of the associated gas temperature for the *X-rays* and *No X-rays* runs. Consequently, this allows direct insights into the impact of X-rays on cells with specific temperatures.
At high redshifts, $z \sim 12$ to $z \sim 10$, we find a mild, but significant shift of the coldest gas cells towards higher temperatures. As also illustrated in Fig. 15, these coldest gas cells are located within the IGM, far from the major sources of heating (visible as the orange-yellow regions in e.g. the $z \sim 10$ panel of Fig. 15).
Throughout the middle stages of reionization, this heating of the coldest cells increases substantially, as shown in the $z \sim 8$ to $z \sim 6$ panels of Fig. 16. At the same time, the expansion of heated bubbles causes more and more cells to shift towards higher temperatures, independently of the presence of X-ray sources.

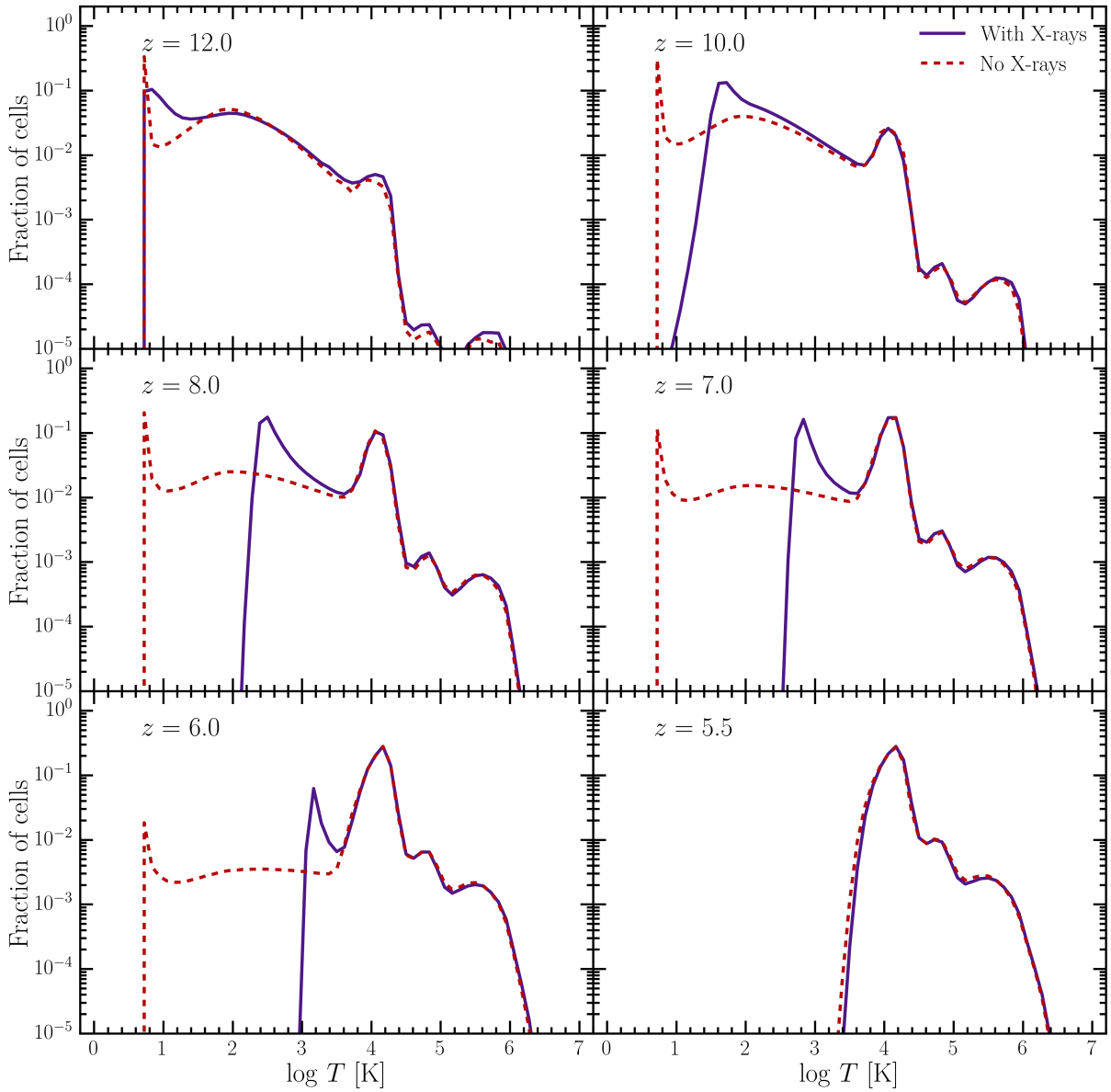

Figure 16: Histograms of the fraction of simulated cells as a function of their gas temperature for both the *X-rays* (solid purple curves) and *No X-rays* (dashed red curves) runs at redshifts $z = 12$ to $z = 5.5$. The implementation of the X-ray sources causes significant heating of the coldest gas cells throughout the Epoch of Reionization, before converging towards the end of reionization at $z \sim 5.5$.

This can be seen clearly by the first two sharp peaks of the *No X-rays* curves, located at $T \approx \mathcal{O}(10\,\mathrm{K})$ and $T \approx \mathcal{O}(10^4\,\mathrm{K})$ between redshifts $z \sim 10$ to $z \sim 6$. As reionization progresses with decreasing redshift, the initially large fraction of cells around the first peak gradually shifts towards the second peak.

This compounds with the effect of the more uniform heating of the IGM in the *X-rays* run. As a result, the hottest regions of the simulation volume are in good agreement with the *No X-rays* run, while the coldest cells progressively heat due to the presence of X-ray sources as mentioned above. The latter causes a growing discrepancy between the *X-rays*

and *No X-rays* runs between $z \sim 12$ and $z \sim 6$.
However, the two runs converge again at $z \sim 5.5$, as the heated bubbles finally encompass the whole simulation volume (see also the $z \sim 5.5$ panels of Fig. 15). At this point, the high temperatures ($T \gtrsim \mathcal{O}(10^4\,\mathrm{K})$ *within* the heated bubbles fully dominate the temperature profile.

Finally, we present the phase-space diagrams of the gas in both the *X-rays* and *No X-rays* runs at redshifts $z = 9$, $z = 7$, and $z = 5.5$ in Fig. 17. They show the distribution of gas mass in the phase space defined by (T, n_H), where T denotes the gas temperature and n_H the number density of hydrogen atoms.
At high temperatures $T \gtrsim \mathcal{O}(10^4\,\mathrm{K})$, both runs show a similar phase-space distribution of the gas at all redshifts, further supporting the observation made above that X-ray heating is subdominant within the hot expanding bubbles during the EoR. This is the case across the entire density range. We note in this context that the very narrow distribution of gas in the phase-space diagrams at high hydrogen number densities, $n_\mathrm{H} \geq 10^{-1}\,\mathrm{cm}^{-3}$, is a reflection of the Springel & Hernquist (2003) effective EoS model employed in the simulations (see Sec. 3.1.3 for further details).
However, the phase-space diagrams clearly show the effect of the X-ray implementation on the simulated *cold* gas at redshifts $z = 9$ and $z = 7$. In addition to the difference in temperature profiles – which reflect the points mentioned in the context of Fig. 16 well – it is clear from the phase-space diagrams that the X-ray radiation affects the cold, *low-density* gas in particular, i.e., the IGM gas. Specifically, at $z \sim 7$, the dilute gas ($n_\mathrm{H} \lesssim \mathcal{O}(10^{-2.5}\,\mathrm{cm}^{-3})$) is completely heated to temperatures $T \gtrsim 10^{2.5}\,\mathrm{K}$, while gas of an intermediate density ($10^{-2}\,\mathrm{cm}^{-3} \lesssim n_\mathrm{H} \lesssim 10^{-1}\,\mathrm{cm}^{-3}$) remains nearly unaffected in its phase-space distribution by the inclusion of X-ray sources. This quantitatively confirms that the X-ray radiation in our implementation propagates into and heats the cold, dilute IGM gas specifically – as expected due to the long mean free path of X-ray photons (Sec. 2.1).

4.1.4 Impact on the gas ionization state

The expansion of the main ionized bubbles during the EoR is largely independent of the uniform X-ray heating discussed in Sec. 4.1.3. This is illustrated in Fig. 18, which shows a direct comparison of the spatial distributions of the H II fractions and gas temperatures at redshift $z = 7$ for the run including X-rays. The distribution of cells with highly ionized hydrogen ($x_\mathrm{H\,II} \gtrsim 0.9$) is primarily contained within individual bubbles, leaving the rest of the IGM gas in a low ionization state ($x_\mathrm{H\,II} \lesssim 0.1$), as expected throughout the middle stages of reionization. On the other hand, the comparison to the temperature distribution clearly shows that a significant amount of heating progresses into the regions *outside* of these ionized bubbles. As shown in the previous section (see the

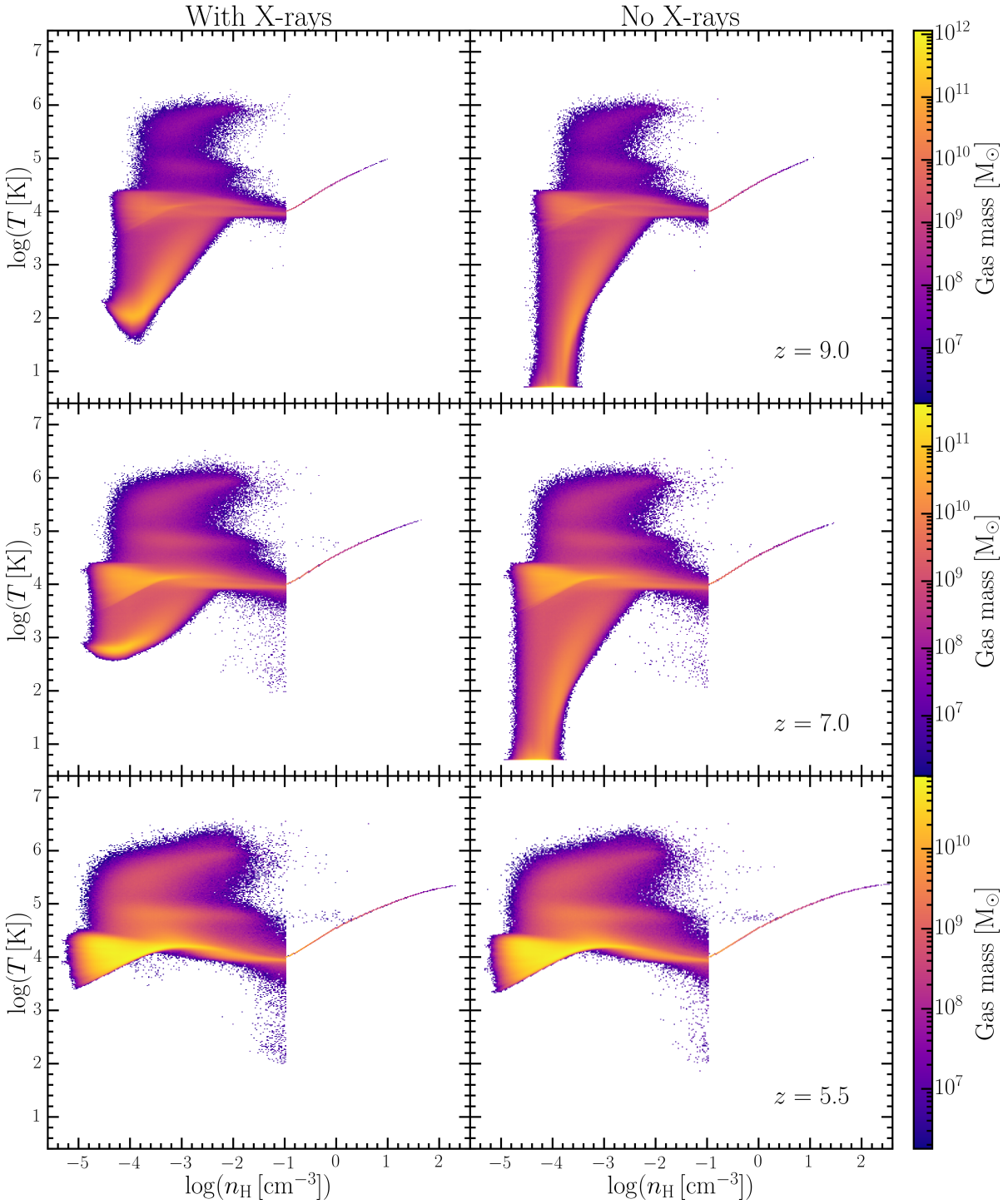

Figure 17: Phase-space diagrams of the *X-rays* (left column) and *No X-rays* (right column) runs at redshifts $z = 9$, $z = 7$, and $z = 5.5$ (rows; as indicated in the figure). T and $n_{\rm H}$ denote the gas temperature and hydrogen number density, respectively. The characteristic X-ray heating of the cold, dilute intergalactic medium gas during the Epoch of Reionization is clearly reflected in the $z = 9$ and $z = 7$ panels, before the phase-space diagrams with and without X-rays converge towards the end of reionization at $z \sim 5.5$. We note that the narrow phase-space distribution of the dense gas ($n_{\rm H} \geq 10^{-1}\,{\rm cm}^{-3}$) is caused by employing the effective Equation of State approximation of Springel & Hernquist (2003) in our model (cf. Sec. 3.1.3).

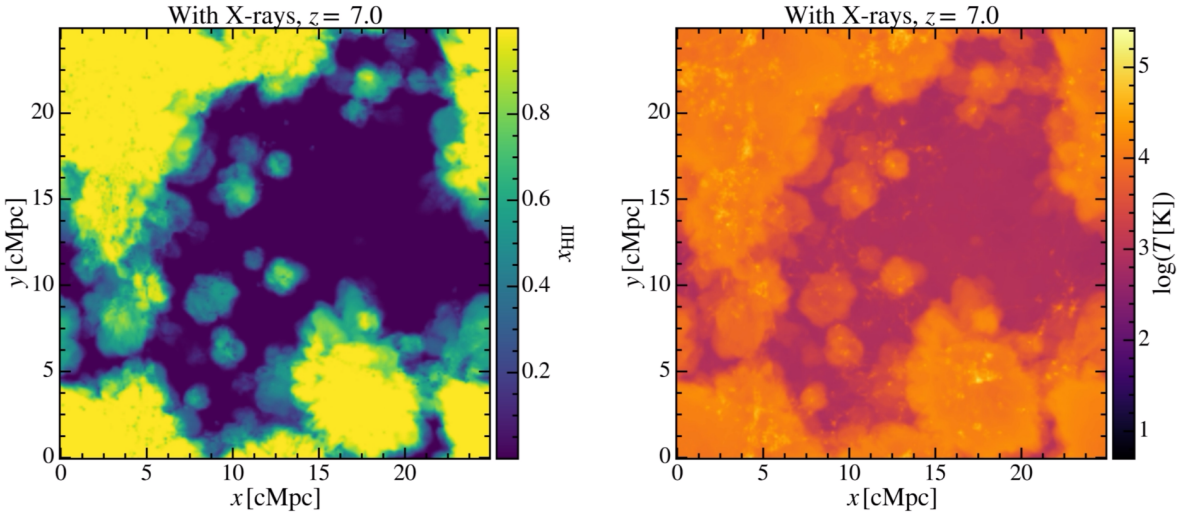

Figure 18: Comparison of the projections of the ionized hydrogen fraction $x_{\mathrm{H\,II}}$ (left panel) and gas temperature T (right panel) distributions in a slice through the simulation volume of the *X-rays* run at redshift $z = 7$. We find that X-ray photons propagate into regions of the intergalactic medium past the ionized bubbles, visible by the X-ray heating of these regions (cf. Fig. 15, $z = 7$ panels).

$z = 7$ panels of Fig. 15), this more uniform heating occurs only when including X-rays, i.e., it is indicative of characteristic X-ray heating of the IGM. Consequently, Fig. 18 illustrates clearly that X-ray photons are able to propagate past the ionization fronts – i.e., the edges of the highly ionized bubbles – in our implementation, as expected physically.

A more detailed look into the ionization state of hydrogen during the EoR in our simulations is presented in Fig. 19, in which we show the fractions of cells throughout the simulation volume as a function of their respective ionized hydrogen fraction $x_{\mathrm{H\,II}}$ for both the *X-rays* and *No X-rays* runs in a redshift range of $z = 12$ to $z = 5.5$.
The resulting $x_{\mathrm{H\,II}}$ distributions progress in a similar manner to the temperature profile's evolution presented in Fig. 16. Firstly, the cells most strongly affected by X-rays at redshifts $z \sim 12$ to $z \sim 6$ are those in the lowest ionization state, i.e., those with the lowest H II fraction ($x_{\mathrm{H\,II}} \approx 10^{-6}$ in the *No X-rays* run). This is analogous to the coldest cells being most substantially affected in the temperature profiles; in fact, the latter are largely the *same* cells as those in a low ionization state. Both are located within the IGM, far away from sources of ionizing and heating radiation (cf. e.g. Fig. 18).
Further, the impact of the X-ray implementation on these cells grows progressively larger until $z \sim 6$, at which point the lowest H II fractions reached are increased to approximately $x_{\mathrm{H\,II}} \approx 10^{-3.8}$ in the *X-rays* run. Nevertheless, the *X-rays* and *No X-rays* runs converge towards the end of reionization at $z \sim 5.5$ – again similarly to the evolution of the temperature distributions – due to the ionized bubbles having expanded throughout the entire simulated volume.

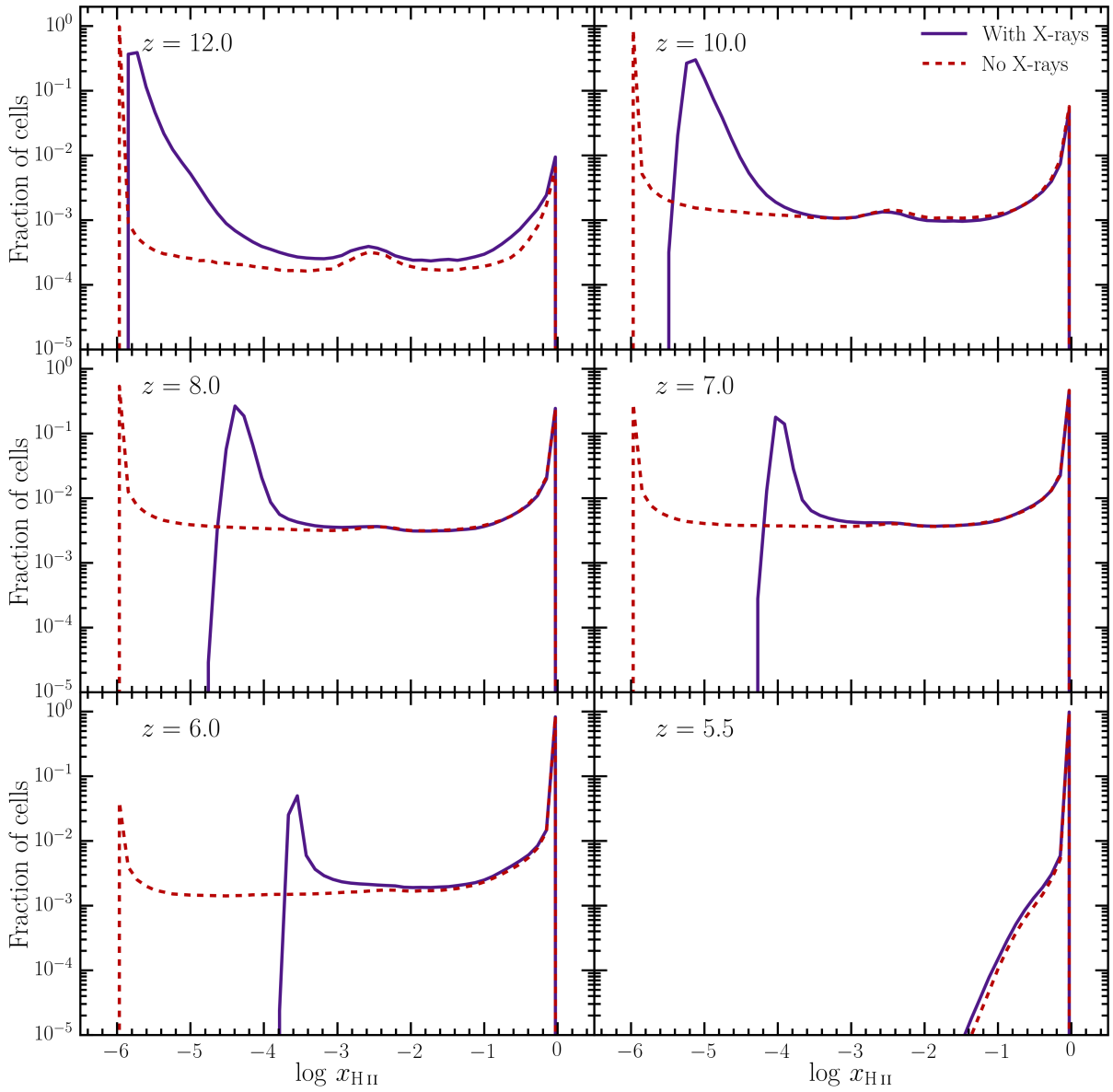

Figure 19: Histograms of the distribution of simulated cells' ionized hydrogen fractions $x_{\mathrm{H\,II}}$, shown for both the *X-rays* (solid purple curves) and *No X-rays* (dashed red curves) runs at redshifts $z = 12$ to $z = 5.5$. Including X-ray sources causes a substantial shift towards a higher ionization state in the cells with the lowest H II fractions throughout the Epoch of Reionization, before the H II fractions approximately converge for the *X-rays* and *No X-rays* runs at $z \sim 5.5$.

At this stage, the fraction of ionized hydrogen is consistently above $x_{\mathrm{H\,II}} \gtrsim 10^{-1.5}$ and nears $x_{\mathrm{H\,II}} \approx 1$ for the majority of cells, marking the end of the EoR by definition.

We additionally show the effect of our implementation of X-ray sources on the fraction of doubly ionized helium He III in Figs. 20 and 21. Due to the higher ionization threshold of He II compared to H I or He I, and the resulting larger ionization cross-sections at higher photon energies, the fraction of He III is expected to be more significantly impacted by the modeling of high-energy X-ray photons (see Fig. 3).

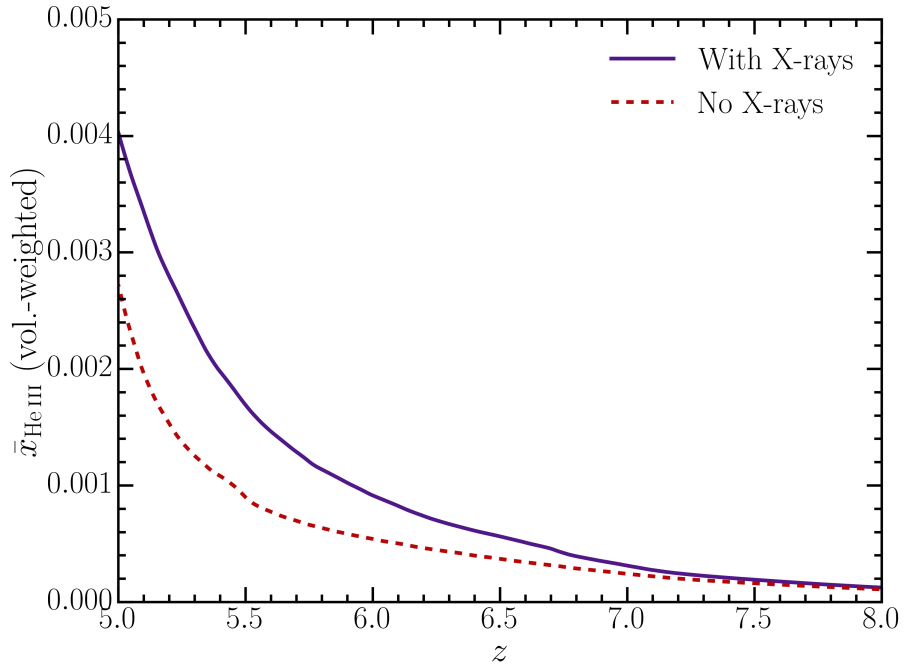

Figure 20: Volume-averaged fraction of doubly ionized helium $\bar{x}_{\mathrm{He\,III}}$ as a function of redshift in the *X-rays* (solid purple curve) and *No X-rays* (dashed red curve) runs. Including X-rays results in a globally higher fraction of He III in the later stages of reionization ($z \gtrsim 7$), a trend that persists beyond the end of the Epoch of Reionization at $z \sim 5.5$.

Indeed, we find this to be the case in our test runs. Firstly, Fig. 20 shows the evolution of the global, volume-averaged He III fraction $\bar{x}_{\mathrm{He\,III}}$ from redshifts $z = 8$ to $z = 5$. We find that in the *X-rays* run, $\bar{x}_{\mathrm{He\,III}}$ is significantly higher throughout the later stages of the EoR ($z \gtrsim 7$) than in the *No X-rays* run. This discrepancy increases until the end of reionization at $z \sim 5.5$, before remaining approximately stable until $z = 5$.

Providing more detailed insights into the full ionization of helium in our simulations, Fig. 21 shows the distribution of individual cells' He III fractions $x_{\mathrm{He\,III}}$ between redshifts $z = 12$ and $z = 5.5$. Here, it is clear that the X-ray sources already begin to mildly affect a significant number of cells at $z \sim 12$, before this subsequently becomes visible in the global volume-averaged He III fraction discussed above. Besides shifting the minimum He III fraction in cells from approximately $x_{\mathrm{He\,III}} \approx 10^{-9}$ to $x_{\mathrm{He\,III}} \approx 10^{-7.5}$ at high redshifts $z \sim 10 - 12$, the X-ray implementation further causes a second, flatter peak at higher He III fractions (centered around $x_{\mathrm{He\,III}} \approx 10^{-6}$ at $z = 10$) in the $x_{\mathrm{He\,III}}$ distributions. This spread-out peak indicates the presence of a larger number of cells with a broadly 'intermediate' He III fraction between the sharper peaks of the distribution at the very low (e.g., $x_{\mathrm{He\,III}} \approx 10^{-7}$ at $z = 10$) and high (e.g., $x_{\mathrm{He\,III}} \approx 10^{-1}$ at $z = 10$) ends; i.e., it indicates that the X-ray photons in our implementation are able to partially ionize He II in additional cells.

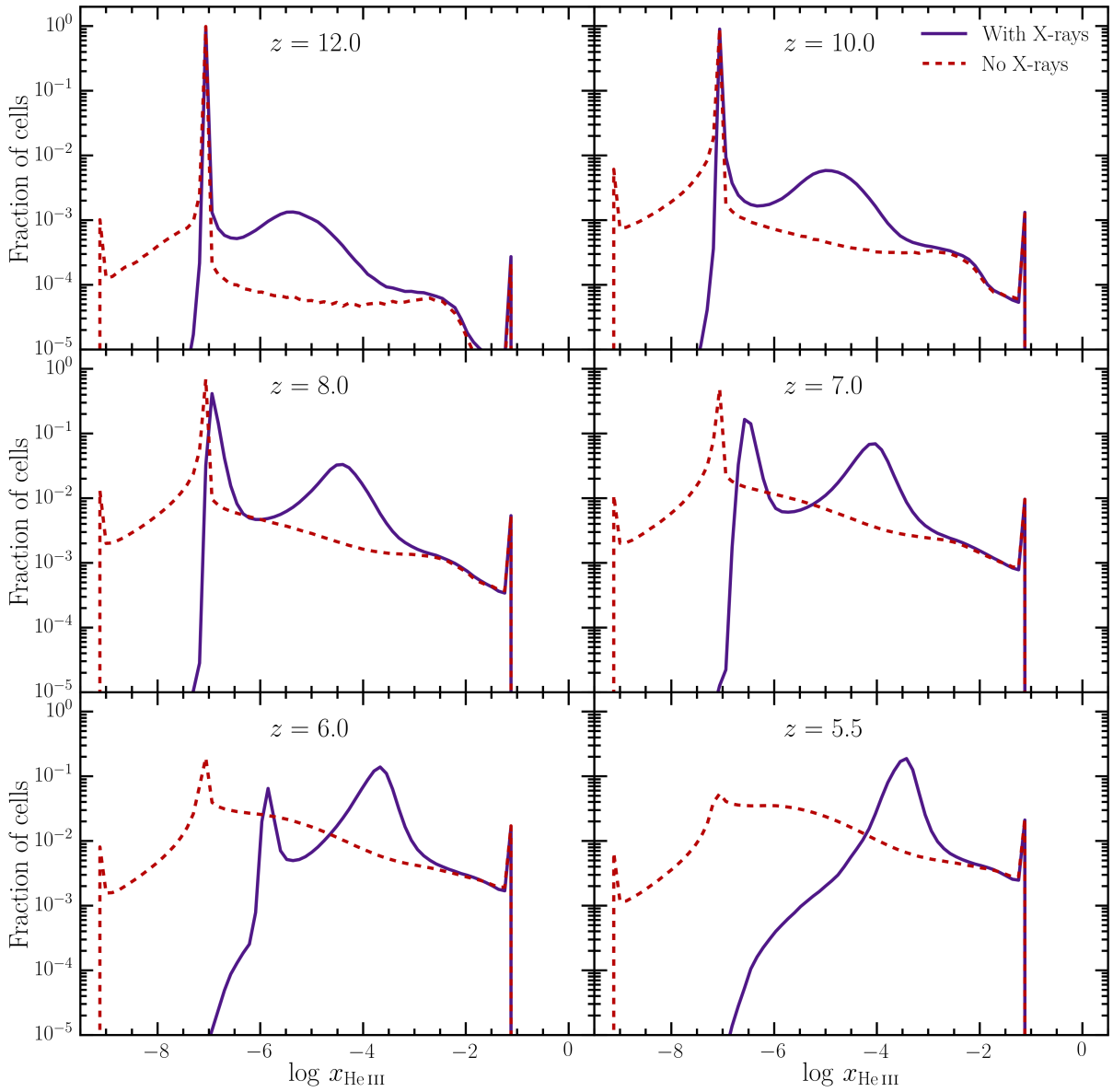

Figure 21: Histograms of the distribution of cells' fractions of doubly ionized helium $x_{\mathrm{He\,III}}$ for the *X-rays* (solid purple curves) and *No X-rays* (dashed red curves) runs at redshifts $z = 12$ to $z = 5.5$. The implementation of X-ray sources results in a significant shift towards higher He III fractions; this trend gradually increases until the end of reionization at $z \sim 5.5$.

Further, this partial ionization appears to be an ongoing process in the period between $z = 12$ and $z = 6$ in the *X-rays* run, as the intermediate, spread-out peak gradually shifts towards higher He III fractions throughout, ending at $x_{\mathrm{He\,III}} \approx 10^{-3.5}$ at $z = 6$. At the same time, the peak at the minimal He III fractions of cells continues to increase to $x_{\mathrm{He\,III}} \approx 10^{-6}$ at $z = 6$ in the *X-rays* run, while remaining approximately constant in the *No X-rays* run. This indicates that as time progresses, *all* cells in the simulation volume are impacted to some degree by the inclusion of X-ray sources. We note that this is consistent with the long mean free path of X-ray photons theoretically expected (Sec. 2.1) and practically observed in our test runs (Sec. 4.1.3).

Finally, unlike the distributions of cells with respect to the temperature or H II fraction – which both converge to excellent agreement between the *X-rays* and *No X-rays* runs by $z \sim 5.5$ (Figs. 16 and 19) – the distribution of He III fractions remains substantially altered by the X-ray implementation past the end of reionization. At $z = 5.5$, the inclusion of X-ray sources has shifted the majority of cells towards the partially ionized, intermediate He III fractions of the broader peak described above ($x_{\mathrm{He\,III}} \approx 10^{-3.5}$), with only few cells remaining at the fraction of the previous minimum peak of the *X-rays* run at $x_{\mathrm{He\,III}} \approx 10^{-6}$, and none at the minimum peak of the *No X-rays* run at $x_{\mathrm{He\,III}} \approx 10^{-9}$. This shows that the X-ray sources have a significant impact on the He III fraction even *within* the ionized, heated bubbles encompassing the entire domain by $z \sim 5.5$, whereas the impact on the temperature and H II fractions is minimal due to the dominant effects of stellar radiation. Additionally, this is consistent with the discrepancy in the global He III fraction between the *X-rays* and *No X-rays* runs as discussed in the context of Fig. 20.

4.1.5 Contribution of individual energy bins

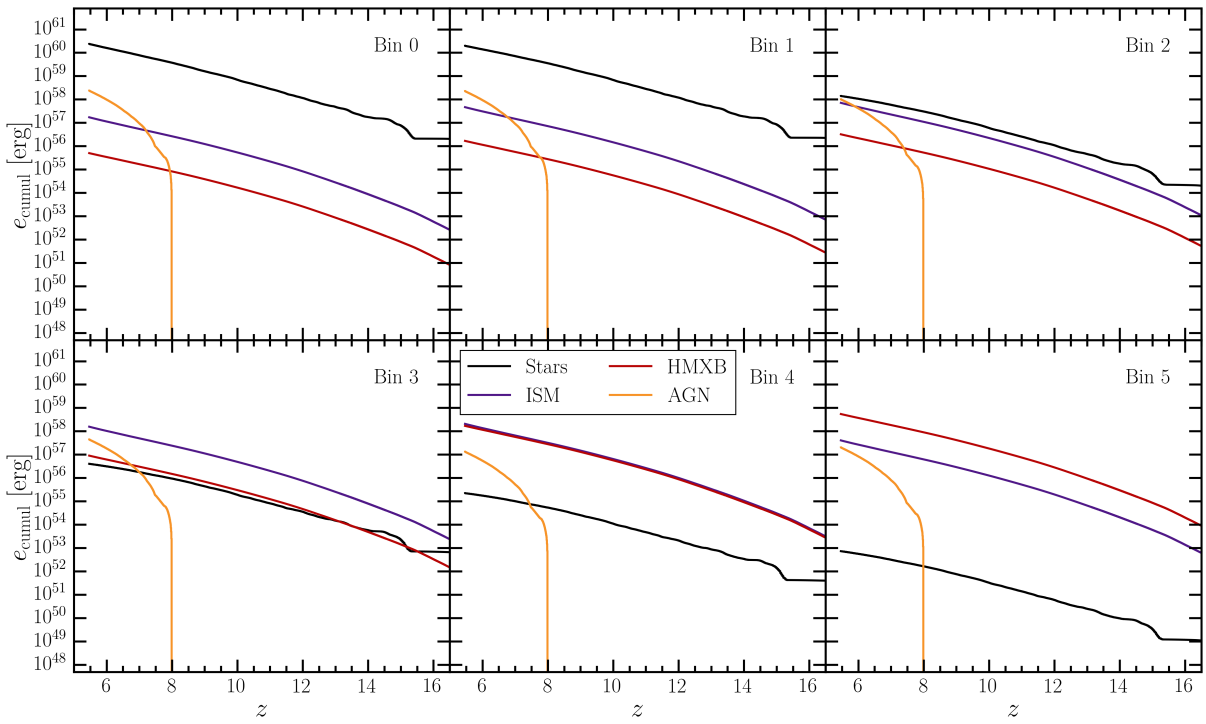

Figure 22: Cumulative photon energy e_{cumul} emitted in the *X-rays* run by source type – stars (black curves), the hot interstellar medium (ISM, purple curves), high-mass X-ray binaries (HMXB, red curves), and active galactic nuclei (AGN, orange curves) – within each of the six energy bins (panels, as labeled) as a function of redshift. The bins are defined by the photon energy ranges [13.6 eV, 24.59 eV] (Bin 0), [24.59 eV, 54.42 eV] (Bin 1), [54.42 eV, 100 eV] (Bin 2), [100 eV, 200 eV] (Bin 3), [200 eV, 500 eV] (Bin 4), and [500 eV, 2 000 eV] (Bin 5); see also Sec. 3.3.1.

The contribution of each of the energy bins used in our X-ray implementation (see Sec. 3.3.1) to the total photon output is presented in Fig. 22. Specifically, we show the total cumulative photon energy emitted by each of the source types – i.e., stars and the three X-ray source types (hot ISM, HMXB, and AGN) – within each of the six energy bins as a function of redshift.
Within the first three UV-range bins, chosen to be identical to the bins used in the original THESAN model and covering the photon energy range $[13.6\,\mathrm{eV}, 100\,\mathrm{eV}]$, stars are the dominant contributors to the total radiative energy output throughout the entire redshift range as expected (e.g. Kannan et al. 2022a). In fact, the cumulative energy they emit within the first two bins exceeds that emitted by the next-highest contributing source – i.e., the hot ISM for $z \gtrsim 8 - 8.5$ and AGN for $z \lesssim 8 - 8.5$ – by several orders of magnitude. However, the relative contribution of the hot ISM increases from bin 0 to bin 4, leading it to exceed the strongly decreasing stellar emissions within bins 3 to 5.
The energy emitted by HMXBs, on the other hand, remains steadily below the ISM emissions from bins 0 to 3, before the two become approximately equal in bin 4. As the HMXB energy contributions increase significantly at high energies (see also the SEDs presented in Fig. 8), HMXBs eventually dominate over the hot ISM with its softer X-ray spectrum – as well as over all other sources – in the highest-energy bin.
Finally, while AGN only begin to contribute radiative energy at lower redshifts $z \lesssim 8$ in our numerical seeding model (cf. Sec. 3.1.3), they rapidly become not only significant sources of radiation, but even the second-to-most dominant sources within bins 0 to 3 by $z \sim 6 - 8.5$. Specifically, only stars (bins 0 to 2) and the hot ISM (bin 3) exceed their contributions in this range by the final simulated redshift, $z = 5$. However, in the highest-energy bins, they remain outshone by both the hot ISM and HMXBs, i.e., the newly implemented X-ray sources in the context of this thesis.

4.2 Impact of the piecewise power-law approach to redshifting

4.2.1 Test run setup

We test our implementation of PPL-based numerical redshifting and investigate its effects in RHD simulations of the EoR in the following. Specifically, we run three simulations using AREPO-RT in its original configuration with three UV-range frequency bins (i.e., using the original THESAN model) up to the final redshift $z = 5$. The first test run, referred to as the *'No redshifting'* run in the following, does not include the new implementation of numerical redshifting and provides a baseline for comparison. The second and third test runs (*'With redshifting'* runs) both include the PPL redshifting implementation; additionally, the third test run employs the slope-limiter $|\alpha| < 10$ (Sec. 3.4.3).

All test runs share the same initial conditions, which were initialized uniformly on a

Cartesian lattice (Sec. 3.1.1), and are based on the Planck Collaboration et al. (2016) cosmological parameters. The box size of the simulations is $L_{\mathrm{box}} = 23.3\,\mathrm{cMpc}$, which is resolved by a total of $N_{\mathrm{DM}} = 128^3 = 2\,097\,152$ dark matter particles. This sets the mass resolution of the test runs to $m_{\mathrm{DM}} = 2.0 \times 10^8\,\mathrm{M}_\odot$.

4.2.2 Reionization history and impact of the slope limiter

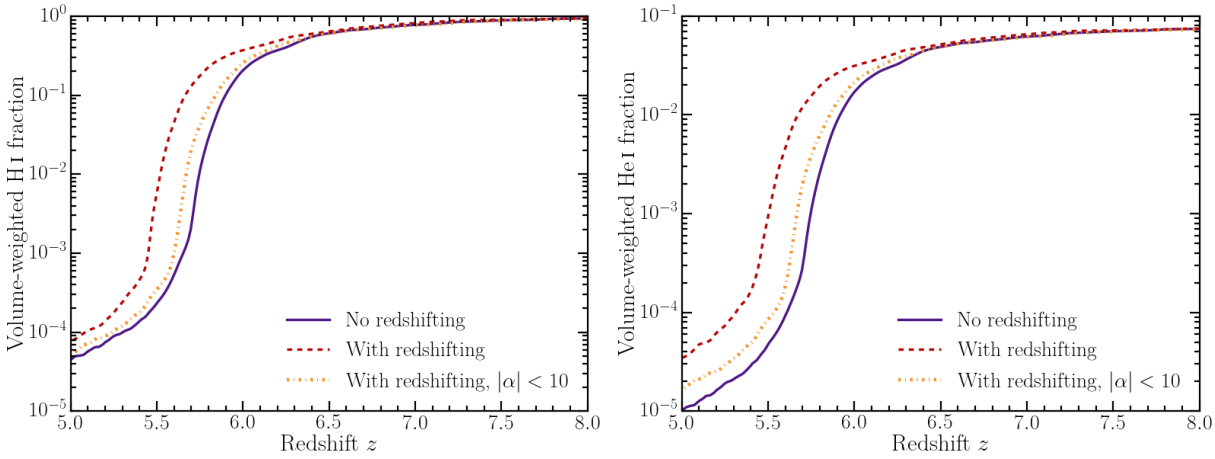

Figure 23: Volume-weighted neutral hydrogen (left panel) and neutral helium (right panel) fractions as a function of redshift from test runs without numerical redshifting (purple curve), with piecewise power-law-based redshifting (red dashed curve), and with slope-limited ($|\alpha| < 10$) piecewise power-law-based redshifting (orange dash-dotted curve). The discrepancy between the latter two curves shows an unphysical impact of the numerical slope limiter for our frequency resolution.

We present the global, volume-averaged fractions of neutral hydrogen H I and neutral helium He I in Fig. 23, comparing the reionization histories of the three test runs described above. The evolution of both neutral fractions show that the inclusion of the numerical PPL-based redshifting method significantly delays the progression of reionization. This is particularly true for the run not employing the slope-limiter $|\alpha| < 10$; here, the steepest drop in the neutral fractions is visibly delayed from a redshift of approximately $z \sim 6$ to $z \sim 5.7$. On the other hand, the differences between the run without redshifting and the run with slope-limited PPL redshifting are less pronounced, with the latter showing only a mild effect on the global reionization history of both H I and He I.
Additionally, we find that both runs with PPL redshifting end with a slightly higher neutral fraction of hydrogen and helium in the simulation volume in the final stages of reionization ($z \sim 5$). Again, this difference is more pronounced when no slope-limiter is employed, raising the final volume-averaged H I and He I fractions from $\bar{x}_{\mathrm{H\,I,final}} \approx 5 \times 10^{-5}$ to $\bar{x}_{\mathrm{H\,I,final}} \approx 8 \times 10^{-5}$ and from $\bar{x}_{\mathrm{He\,I,final}} \approx 10^{-5}$ to $\bar{x}_{\mathrm{He\,I,final}} \approx 3.5 \times 10^{-5}$, respectively.

Physically, both the delay in the reionization history and the final higher neutral fractions are caused by an overall reduction in radiative energy due to the redshifting implementation. We note that in reionization simulations, the 'overionization' of hydrogen gas – i.e., the simulations' tendency towards a lower neutral hydrogen fraction throughout the later stages of reionization than inferred from observations – presents a well-known potential problem, particularly when not employing the RSLA (Ocvirk et al. 2019). In this context, the slightly higher final H I fraction in the redshifting run seemingly suggests that numerically accounting for the effects of redshifting may help reduce this problem. However, we caution that the discrepancy between the slope-limited and non-slope-limited redshifting runs calls the reliability of this interpretation into question.

More generally, the discrepancy between the reionization histories inferred from the latter two runs indicates a serious issue of the PPL redshifting method as used in our numerical setup based on the THESAN model. It is clear from Fig. 23 that the slope limiter $|\alpha| < 10$, introduced as a purely numerical parameter (Sec. 3.4.3), has a significant impact on the physical results of our simulations.
There are several issues associated with this. Firstly, the choice $|\alpha| < 10$ is not physically motivated, but rather only introduced to limit numerical noise; as such, any impact of the parameter on the ionization state of the simulated gas is unphysical by nature.
Secondly, simply setting a global slope limiter does not distinguish between cases in which the spectral index α happened to be exceptionally high due to physical reasons – i.e., the actual photon spectrum – and cases in which this is due to numerical uncertainty, in line with the original intention of the slope limiter.
Finally, further testing of the implementation, in which we computed statistics of the values of α at each timestep, revealed that the calculated spectral index' typical spread is frequently beyond the range $\alpha \in [-10, 10]$.

This is illustrated in Fig. 24, showing the percentage of cells with each value for α at redshifts $z \sim 6.7$ and $z \sim 5.0$ for bins 0 and 1 with energy ranges of $[13.6\,\mathrm{eV}, 24.59\,\mathrm{eV}]$ and $[24.59\,\mathrm{eV}, 54.42\,\mathrm{eV}]$, respectively. Specifically, we show this for (i) all cells, (ii) only the cells with a minimum energy density of $E_{i,\,\mathrm{cell}} > 10^{-20}$ in internal units (see Sec. 3.4.3), and (iii) only the cells with $E_{i,\,\mathrm{cell}} > 10^{-10}$ in internal units. We compute and show these cases separately to differentiate between cells with a higher energy density – i.e., $E_{i,\,\mathrm{cell}}$ well above the numerical minimum value `MINDENSPHOT` $= 10^{-25}$ in internal units – and those with a lower energy density, which are likely to be affected by relatively higher numerical errors.
Indeed, we find a significantly smaller spread of α values when limiting the statistics to cells with higher energy densities. This holds for both bin 0 and 1 and provides evidence that the sporadic high values of the spectral index are, in fact, primarily due to the cells

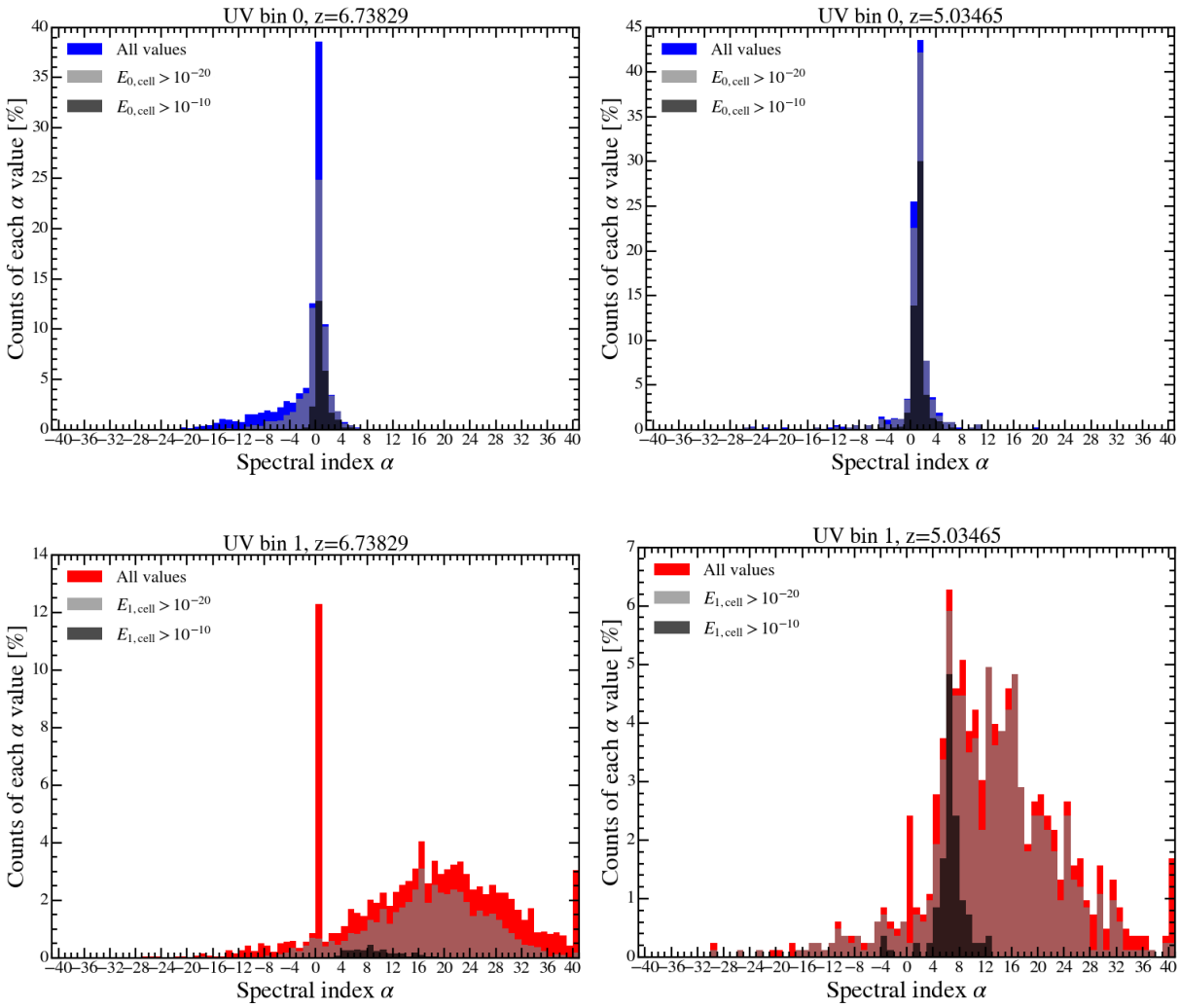

Figure 24: Distribution of values for the piecewise power-law spectral index α in the UV-range energy bins 0 ($[13.6\,\mathrm{eV}, 24.59\,\mathrm{eV}]$; upper panels) and 1 ($[24.59\,\mathrm{eV}, 54.42\,\mathrm{eV}]$; lower panels) of AREPO-RT in the THESAN setup (cf. Sec. 3.3.1), shown at redshifts $z \approx 6.7$ (left panels) and $z \approx 5.0$ (right panels). The blue and red bars show the fractions of cells with each α value for bins 0 and 1, respectively. Values of $|\alpha| > 40$, if and when they occurred, were counted into the highest ($\alpha > 40$) or lowest ($\alpha < -40$) shown bins. The light and dark gray shaded bars show this same fraction for only the cells with a minimum energy density of $E_{i,\mathrm{cell}} > 10^{-20}$ and $E_{i,\mathrm{cell}} > 10^{-10}$, respectively; here, $E_{i,\mathrm{cell}}$ in bins $i = 0, 1$ are specified in internal energy density units, corresponding to physical units of $[E] = 6.18 \times 10^{52}\,\mathrm{erg\,ckpc^{-3}}$. Setting these lower limits is intended to exclude the values of α that may be disproportionately affected by numerical uncertainty due to their low values in internal units. We find that particularly in bin 1, the spread of the spectral index frequently far exceeds the range $[-10, 10]$ enforced by the slope limiter (see text for details), even for cells with a comparatively high energy density.

with lower energy densities. Further, the spread of spectral indices is substantially higher within bin 1 than bin 0, likely for a similar reason: as stars are the dominant sources of radiation in the simulation volume, but only emit a small fraction of their luminosity at energies $h\nu \gtrsim 54.4\,\mathrm{eV}$ (e.g. Haardt & Madau 2012; see also Sec. 3.3), the energy density in bin 2 ($[54.42\,\mathrm{eV}, 100\,\mathrm{eV}]$) is expected to be particularly low in the setup considered here. This is likely to introduce numerical uncertainty in the computation of the spectral index α_1 in bin 1, which is based on the energy densities within both bin 1 and bin 2 (Eq. (86)).

Most crucially, Fig. 24 shows that it is far from uncommon for the spectral index to exceed the range $[-10, 10]$, if these limits are not enforced by a slope limiter. In frequency bin 1, this is frequently the case even for cells with a relatively high energy density.
This finding is consistent with the observation made above that enforcing a non-physical cutoff $|\alpha| < 10$ can significantly change the physical effects of the redshifting implementation. It further suggests that even a slightly different choice of slope limiter, e.g. $|\alpha| \leq 15$, would affect the amount of energy lost or gained by PPL-based redshifting in a substantial number of cells, potentially arbitrarily impacting simulation results.

As mentioned in Sec. 3.4.3, we expect that the low frequency resolution employed in our numerical setup is at the core of the limitations we find for the PPL method of redshifting. In particular, a higher resolution – such as frequently employed in the context of RT postprocessing works – is likely to significantly reduce errors in the method caused by a poor reconstruction of the radiation spectrum (see e.g. Fig. 11). However, given that the methods described here are implemented in light of their intended use in large-volume, on-the-fly RHD simulations, substantially increasing the frequency resolution is unfortunately not feasible at present (Sec. 3.4.3). We therefore chose to additionally develop and pursue the alternative EER approach to redshifting, discussed further in Sec. 4.3 below.

Nevertheless, we present a brief outlook on potential future work on PPL redshifting in the following.
The first and most straightforward way to potentially improve on the method presented above may be restricting the use of the slope limiter to bins with low energy densities, i.e., with E_i' below a certain threshold set in internal code units. In this case, the slope would no longer be artificially limited in cells and bins where the energy densities are high enough for numerical uncertainties to play a negligible role. It can be reasonably expected that this would reduce the (unphysical) impact of the choice of the slope limiter on simulation results, particularly since cells with high energy densities represent, by definition, more dominant parts of the radiation field.
As a more involved, but likely preferable alternative, we suggest using the maximum spectral index of the input source SEDs as a physically motivated slope limiter. Since

the bin edges for EUV photons are chosen at the ionization thresholds of H I, He I and He II, interactions with gas are expected to harden the spectrum in each of these bins (as illustrated in Fig. 12), *decreasing* α_i with respect to the original source spectrum. Further, gas ionization becomes increasingly inefficient for higher-energy photons (Fig. 3) – such as those modeled in the X-ray implementation discussed in Sec. 3.3 – causing this same trend to weaken, but persist across the entire modeled photon energy spectrum.
Therefore, we conjecture that this approach may provide a reasonable choice for an *upper limit* to α_i. An additional potential advantage is the ability to separately define the slope limiter within each bin, potentially helping to capture the spectrum more accurately than a fixed global maximum value such as $|\alpha| < 10$.

Finally, we note that a numerical RT approach designed to accurately track the radiation spectrum using a PPL approximation – albeit in a different context than the redshifting method discussed above – has been developed by Baumschlager et al. (2023). This method defines two very narrow frequency bins around each of the *edges* of the typically used bins (see Sec. 3.3.1) and infers the photon spectrum in between the narrow bins by using a PPL approximation, in lieu of tracking the spectrum in broad frequency bins.
Although this also requires implementing more bins than used in the original THESAN model, thereby driving up the computational cost (Sec. 3.4.3), it may be a promising approach to use in conjunction with the PPL-based redshifting method. This is because it appears to effectively avoid the most uncertain and frequency resolution-dependent step of the redshifting method – i.e., estimating the spectral index from a coarse, piecewise-constant description of the radiation spectrum – by instead tracking the spectral index in each bin directly by design. We plan to investigate this possibility further in future work.

4.3 Impact and convergence of the effective energy-rebinning approach to redshifting

4.3.1 AREPO-RT test run: setup and reionization history

To test the EER method of numerical redshifting and its implementation within the larger code base, we first present results from a set of test runs with AREPO-RT using the updated THESAN model including X-ray sources and six energy bins (Sec. 3.3).
These test simulations were run in a box of the size $L_{\mathrm{box}} = 11.6\,\mathrm{cMpc}$, resolved by $N_{\mathrm{DM}} = 128^3 = 2\,097\,152$ dark matter particles, resulting in a mass resolution of $m_{\mathrm{DM}} = 3.0 \times 10^7\,\mathrm{M}_\odot$. Further, they are each based on the same initial conditions, are initialized uniformly on a Cartesian lattice (see Sec. 3.1.1), and use the Planck Collaboration et al. (2016) values for the cosmological parameters.
Specifically, we ran two test simulations with this setup: similarly to the approach in previous sections, we first ran a baseline simulation that does not include numerical redshifting

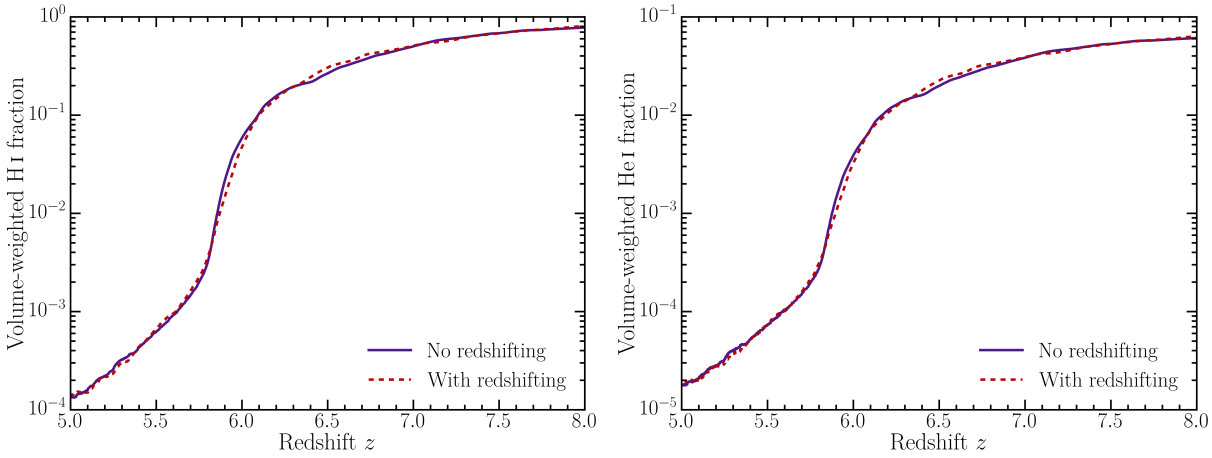

Figure 25: Volume-weighted fractions of neutral hydrogen H I and neutral helium He I as a function of redshift from test runs with (red dashed curve) and without (purple solid curve) numerical redshifting based on the effective energy-rebinning approach. We find no significant impact of the redshifting implementation on the global reionization history.

(*'No redshifting'* run). The second test run (*'With redshifting'* run) includes numerical redshifting based on the AREPO-RT implementation of the EER method described in Sec. 3.5.

In Fig. 25, we present the evolution of the neutral hydrogen and neutral helium fractions from redshift $z = 8$ to $z = 5$ in these test simulations with AREPO-RT. As is clear from the figure, the *With redshifting* and *No redshifting* test runs are in excellent agreement with each other; including the EER redshifting method does not significantly impact either the hydrogen or helium reionization history.
Fig. 25 therefore indicates a lack of the unexpected numerical 'runaway' effects we observed and discussed in the context of the PPL-based redshifting method (see Sec. 4.2.2; Fig. 23) when using the EER redshifting implementation. Rather, it appears to correctly capture the fact that redshifting is expected to have a negligible effect on the reionization history: we stress again in this context that the ionization state of the gas during the EoR is primarily driven by the EUV photons emitted by stars (e.g. Kannan et al. 2022a), which have a comparatively short mean free path during this epoch (Sec. 2.1; see also e.g. Eide et al. 2020, App. A). Consequently, the dominant portion of ionizing photons gets absorbed quickly after being emitted, compared to the timescale on which redshifting has a significant effect – in fact, this is the core argument justifying neglecting redshifting in simulations (e.g. Gnedin & Abel 2001; Kannan et al. 2022a). This is therefore well-reflected in our implementation of EER redshifting by its lack of impact on the ionization states of H I and He I gas.

4.3.2 Accuracy of the method and convergence with the number of energy bins

In the following, we present results from a numerical test evaluating the accuracy of the EER redshifting method with respect to the analytically expected redshifting of a set of source spectra, as well as the method's performance based on the number of energy bins used to discretize the radiation spectrum. We perform this test in a separate code from AREPO-RT, which essentially isolates the redshifting implementation from the rest of the code base and applies it to a well-defined, single-source spectrum.
Specifically, we first initialize one of two analytically well-defined source spectra[13] in this separate code, i.e., either the thermal bremsstrahlung spectrum of the hot ISM – given by Eq. (61) in relative units – or a thermal blackbody spectrum, described by the SED

$$\frac{\mathrm{d}L_{\mathrm{BB}}(e)}{\mathrm{d}e} = \frac{2e^3}{(hc)^2} \frac{1}{\exp\left(\frac{e}{k_B T_{\mathrm{eff}}} - 1\right)} \,. \tag{98}$$

Here, h, c, and k_B denote Planck's constant, the speed of light, and Boltzmann's constant, respectively. T_{eff} is the effective temperature of the blackbody spectrum, set to $T_{\mathrm{eff}} = 10^6\,\mathrm{K}$ in the following tests to obtain significant emissions in both the EUV and soft X-ray range covered by the AREPO-RT bins (total range $[13.6\,\mathrm{eV}, 2\,000\,\mathrm{eV}]$; note that $k_B T_{\mathrm{eff}} = 86\,\mathrm{eV}$ for $T_{\mathrm{eff}} = 10^6\,\mathrm{K}$).

To test our redshifting implementation, we then use the EER method to numerically redshift this initialized spectrum in N_{b} bins covering the energy range $[13.6\,\mathrm{eV}, 2\,000\,\mathrm{eV}]$, using a total of $N_t = 10^6$ timesteps of the size $\Delta(\log a) = 10^{-6}$. For an initial redshift $z_{\mathrm{init}} = 10$, this results in a final redshift of $z_{\mathrm{fin}} \approx 3.7$. Additionally, we compute the analytically expected redshifted spectrum at z_{fin} and bin this in the same N_{b} energy bins, serving as the 'ideal' result; in other words, this represents the expected result, if the numerical redshifting method were perfectly accurate.
Subsequently, we estimate the error of the redshifting implementation as the sum of the absolute values of the difference between the numerically and analytically redshifted final energy density in each bin. Additionally, we compute the error resulting from neglecting redshifting in the same way, in this case by considering the differences between the *initial* (non-redshifted) spectrum and the analytically redshifted spectrum.
We perform both of these caluculations for a range of different numbers of bins N_{b}; the results are presented in Fig. 26 for both the ISM and blackbody spectra.

[13] We note that we limited this test to spectra available in functional form to ensure that the analytical redshifted spectrum can be computed accurately. This excludes e.g. the HMXB and AGN spectra discussed in Secs. 3.3.4 and 3.3.5, as these are given in tabulated form and numerically interpolated to obtain the spectrum at intermediate values of the photon energy e in AREPO-RT.

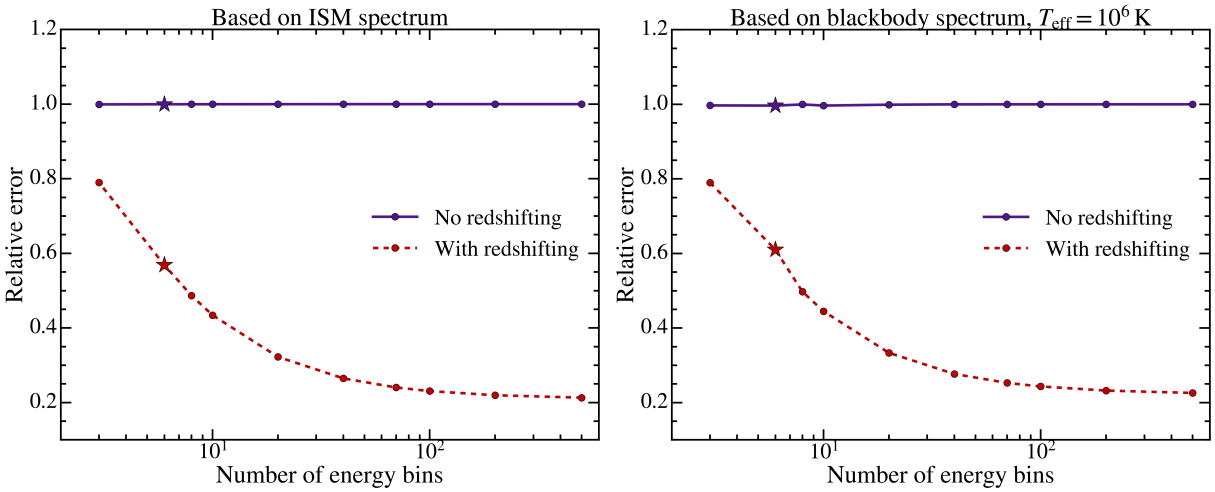

Figure 26: Accuracy of the effective energy-rebinning redshifting method and its convergence with respect to the number of energy bins used. This test (see text for details) was performed based on the spectral energy distribution of the hot interstellar medium (left panel; see Sec. 3.3.3) and based on a blackbody spectrum with an effective temperature of $T_{\mathrm{eff}} = 10^6$ K (right panel). We initialized each spectrum at $z = 10$ and redshifted it both analytically and numerically to $z \approx 3.7$. Curves show the relative error of the final binned spectrum, obtained by either using the effective energy-rebinning redshifting method (red dashed curves) or by neglecting redshifting (purple solid curves), as a function of the number of bins used in the computation. The absolute error is calculated with respect to the analytically redshifted spectrum; this is subsequently normalized by the maximum error obtained either with or without redshifting. Star markers indicate the error computed using the energy discretization of AREPO-RT (six bins; see Sec. 3.3.1).

We note that these errors are computed in relative units, as the underlying SEDs are inevitably in relative units to begin with (cf. Sec. 3.3: the energy output in absolute units, i.e. the normalization of the SED, is only set by the luminosity). Therefore, we normalize these errors for each source spectrum by dividing by the maximum error obtained for any of the tested N_{b} values, either with or without using redshifting, to obtain the relative errors shown in Fig. 26.
Further, the energy bins employed in this test are specified by dividing the total energy range of AREPO-RT into N_{b} intervals that are logarithmically evenly spaced. The only exception to this is $N_{\mathrm{b}} = 6$ (highlighted by the star markers in Fig. 26), which instead uses the same energy discretization as the X-ray version of AREPO-RT (see Sec. 3.3.1).

For both the computations based on the hot ISM spectrum and those based on the blackbody spectrum, we find that the numerical EER redshifting method becomes less accurate for a lower number of bins and more accurate for a larger number of bins as expected. We note that the method's error does not appear to approach zero even for very large numbers of energy bins ($N_{\mathrm{b}} = 200 - 500$), instead showing a tendency towards a leftover error of approximately 20 % of the maximum error for both spectra. One factor

affecting this may be the method's rather approximate handling of the highest-energy bin (see Sec. 3.5); even for a large value of N_b, for which each single bin is comparatively very narrow, this potential error in the highest-energy bin may progressively propagate to lower-energy bins due to the nature of the method.
Nevertheless, crucially, the error made in the final spectrum is consistently lower when using the EER-based redshifting method than when neglecting redshifting – which is the current standard approach of typical on-the-fly RHD solvers (cf. Secs. 2.3 and 3.2) – even for a number of bins as low as $N_\mathrm{b} = 3$, or AREPO-RT's $N_\mathrm{b} = 6$.
Specifically, the error is reduced by approximately 40 % by EER redshifting in the latter case; this is an imperfect, but significant improvement that further adds only negligible computational expense in the larger context of an AREPO-RT simulation. The method additionally does not appear to suffer from unexpected and difficult to predict numerical issues, as found for the PPL-based redshifting method – despite being applied in the context of the same low energy resolution. Consequently, we believe this method has the potential to serve as a viable approximate treatment of numerical redshifting in future large-volume RHD simulation projects.

4.3.3 Convergence of the method with the number of timesteps

Finally, we investigate the convergence of the EER redshifting method with the number of timesteps used over a constant total period of time. For this, we use the same basic testing setup as described in the previous section to numerically redshift the hot ISM spectrum from an initial redshift of $z_\mathrm{init} = 15$ to a final redshift of $z_\mathrm{fin} \approx 5.5$, using the standard X-ray setup of the AREPO-RT bins (Sec. 3.3.1).
This redshifting evolution is computed with a varying number of timesteps N_t, ranging from only $N_t = 10$ to $N_t = 10^6$. Each timestep changes the logarithm of the scale factor a by a constant amount ($\Delta(\log a) = \text{const.}$, equivalent to $\Delta a/a = \text{const.}$). This definition is chosen so that running e.g. $N_t = 10^1$ timesteps with $\Delta(\log a) = 10^{-1}$, or $N_t = 10^2$ timesteps with $\Delta(\log a) = 10^{-2}$, and so on, all result in the same total period of time covered, i.e., the same final redshift.

We record the final energy density in each of the six bins after the initial spectrum has been redshifted in this manner; the result is presented in Fig. 27 for each N_t. Similarly to the previous test, this final energy density is again given in relative, or arbitrary, units (see Sec. 4.3.2).
With the exception of $N_t = 10^1$ – an extremely low number of timesteps to cover the redshift evolution from $z_\mathrm{init} = 15$ to $z_\mathrm{fin} \approx 5.5$ – we find generally good agreement between the final energy densities. In fact, the discrepancy between the results using $N_t \geq 10^3$ timesteps and those only using $N_t = 10^2$ timesteps is very mild; in Fig. 27, it can only be discerned by 'zooming into' the overall curves (inset panel).

For a moderate number of timesteps and above, $N_t \geq 10^3$, we find no significant difference between the final energy densities; rather, we find excellent convergence of the EER-based redshifting method.

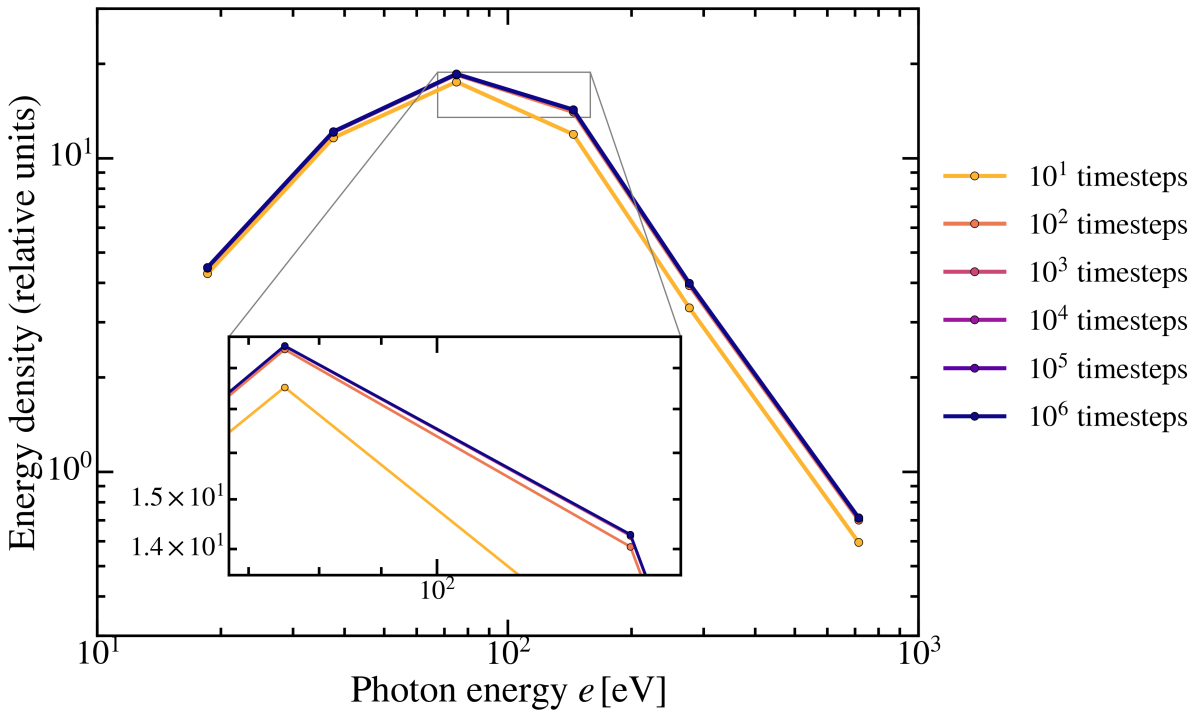

Figure 27: Convergence of the effective energy-rebinning redshifting method with respect to the number of timesteps. Curves show the final energy density in relative units, obtained by numerically redshifting the source spectrum of the hot interstellar medium from redshift $z_{\text{init}} = 15$ to $z_{\text{fin}} \approx 5.5$ in 10^1 (yellow curve), 10^2 (orange curve), 10^3 (red curve), 10^4 (magenta-purple curve), 10^5 (dark purple curve), and 10^6 (dark blue curve) timesteps. The size of the timesteps is determined by imposing the condition that $\Delta(\log a) = \text{const.}$ for each timestep. The inset panel shows a zoomed-in region of the curves to highlight the small difference between the runs using $N_t = 10^2$ and $N_t \geq 10^3$ timesteps. Beyond this, we find excellent convergence of the redshifting method for $N_t \geq 10^3$ timesteps covering the redshift evolution specified above.

5 Summary and conclusion

In this thesis, we numerically explored the Epoch of Reionization, a particularly compelling field of study in light of the wide range of remaining open questions and the extensive observational efforts currently dedicated to investigating them (cf. Sec. 1).
We described in detail how numerical RHD simulations leverage state-of-the-art algorithms and HPC environments to provide complementary theoretical predictions. Specifically, large-volume reionization simulations such as the THESAN suite are capable of providing predictions for numerous cosmological-scale, reionization-era observables, including the global 21-cm line signal. These predictions can not only serve as comparisons to observational data – such as those gathered by current and next-generation radio interferometers – but also facilitate theoretical interpretations of these data and provide insights into the underlying physical processes at play.

However, to achieve this, accurately modeling X-ray sources at high redshifts – i.e., the hot ISM, HMXBs, and AGN – is essential. The soft X-ray photons these sources emit are capable of propagating far into the IGM, causing a characteristic, relatively uniform X-ray heating, which is expected to be reflected in the large-scale 21-cm power spectrum.
In this work, we have demonstrated that our new implementation of these X-ray sources into the moment-based RT solver AREPO-RT successfully captures this characteristic X-ray heating of the IGM, along with other key physical signatures. We showed directly that the simulated X-ray radiation is capable of propagating past the ionization fronts of the heated, ionized bubbles during the EoR, reflecting the expected long mean free path of X-ray photons. Additionally, we investigated the latter's effects on the fraction of doubly ionized helium, finding a significantly higher mean He III fraction that persists past the end of reionization at $z \sim 5.5$.

Motivated by X-ray photons' long mean free path, we additionally explored two different methods of implementing cosmological redshifting into AREPO-RT. In particular, we presented the challenges associated with a piecewise power-law-based approach to redshifting in the context of the low energy resolution that is typical of numerically demanding on-the-fly RHD simulations. We consequently developed a new method of handling numerical redshifting, referred to as 'effective energy-rebinning' in this thesis.
We verified that this EER approach does not suffer from the numerical artifacts observed for the PPL method in the context of the reionization histories, despite employing the same energy discretization of the radiation field. Further, we demonstrated that the use of this method consistently improves agreement with analytically redshifted spectra, both for the AREPO-RT energy binning and for alternative discretizations using a total of $N_\mathrm{b} = 3$ to $N_\mathrm{b} = 500$ energy bins. Finally, we tested the approach using a varying number of

timesteps to evolve a spectrum over a set period of time and found excellent convergence of the method in this context.

The THESAN reionization simulation suite exclusively considered the impact of photons with energies up to 100 eV, focusing on UV photons as the dominant contributors to reionization. However, as the THESAN-XL project is poised to numerically investigate the periods of cosmic dawn and reionization in even larger spatial volumes, it sets the stage for more detailed, accurate predictions for the global 21-cm line signal. For the reasons discussed throughout this thesis, the extensions to the physical model described here are expected to form an important building block towards this objective.

Acknowledgements

I would first like to thank Rahul Kannan and Laura Sagunski for being excellent main thesis advisors. Thank you for welcoming me into your research groups, for the great collaboration in the EXPLORE-IV program, for agreeing to co-supervise my Master's thesis, and for providing a ton of invaluable guidance as well as practical advice and support along the way. Needless to say, this thesis would not have been possible without you.

I also want to thank both Rahul Kannan and Oliver Zier for consistently helping me with both technical details in my research and big-picture advice. I am incredibly grateful for the expertise, suggestions, and ideas you shared with me, and for the time you took to discuss this research in countless helpful conversations. Additionally, I would like to thank Oliver for helping me proofread a key part of this thesis.

I am also immensely grateful to Mark Vogelsberger for inviting and welcoming me as a visiting student at the MIT Kavli Institute for Astrophysics and Space Research. Thank you for your excellent advice – both in terms of my research and my long-term career plans – and thank you for always taking the time to meet for many helpful discussions.

Further, I would like to thank Volker Springel, Aaron Smith, and Rüdiger Pakmor for sharing their expertise and providing invaluable insights in many productive meetings about the THESAN-XL project and THESAN as a whole. Thank you for the great collaboration, and for sharing ideas that substantially shaped this research.

Last but absolutely not least, I am extremely grateful to my family, my friends, and my fellow students and research group members. I would like to say a special thank you to Tamara Caldas-Cifuentes, Yannik Schaper, and Edwin Genoud-Prachex for the warm welcome into the DMGW group here in Frankfurt, and to Teo Bulichi, Meredith Neyer, Jacob Shen, Cian Roche, and Mouza Almualla for all of the great times during my stays in the US. Thank you so much for all your support, and for all the things I got to learn from you not only in formal meetings or journal clubs, but also while just chatting and having a ton of fun.

I gratefully acknowledge funding from a Goethe University International Lab Visits scholarship and MIT Visiting Student scholarship in support of this work.

This research was enabled in part by support provided by the Shared Hierarchical Academic Research Computing Network (SHARCNET; `www.sharcnet.ca`) and Digital Research Alliance of Canada (`alliancecan.ca`). Computations were performed on the *Narval* supercomputer and the *Niagara* supercomputer at the SciNet HPC Consortium (Loken et al. 2010; Ponce et al. 2019). SciNet is funded by Innovation, Science and Economic Development Canada; the Digital Research Alliance of Canada; the Ontario Research Fund: Research Excellence; and the University of Toronto.

In the larger context of the THESAN-XL project, further simulations relevant to this project

were carried out on the *Frontier* supercomputer. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

I acknowledge the Texas Advanced Computing Center (TACC; `http://www.tacc.utexas.edu`) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this thesis.

Data storage resources in support of this research were provided by the MIT *Engaging* HPC cluster.

References

Abel T., Wandelt B. D., 2002, MNRAS, 330, L53

Abel T., Norman M. L., Madau P., 1999, ApJ, 523, 66

Agertz O., et al., 2007, MNRAS, 380, 963

Alme M. L., Wilson J. R., 1973, ApJ, 186, 1015

Andalman Z. L., Teyssier R., Dekel A., 2024, arXiv e-prints, p. arXiv:2410.20530

Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046

Appel A. W., 1981, An Investigation of Galaxy Clustering Using an Asymptotically Fast N-Body Algorithm. Undergraduate Thesis, Princeton University

Appel A. W., 1985, SIAM Journal on Scientific and Statistical Computing, 6, 85

Asthana S., Haehnelt M. G., Kulkarni G., Bolton J. S., Gaikwad P., Keating L. C., Puchwein E., 2024, arXiv e-prints, p. arXiv:2409.15453

Aubert D., Teyssier R., 2008, MNRAS, 387, 295

Bagla J. S., 2002, Journal of Astrophysics and Astronomy, 23, 185

Barnes J., Hut P., 1986, Nature, 324, 446

Bauer A., Springel V., Vogelsberger M., Genel S., Torrey P., Sijacki D., Nelson D., Hernquist L., 2015, MNRAS, 453, 3593

Baumann D., 2022, Cosmology. Cambridge University Press, doi:10.1017/9781108937092

Baumschlager B., Shen S., Wadsley J. W., 2023, arXiv e-prints, p. arXiv:2310.16902

Belczynski K., Kalogera V., Bulik T., 2002, ApJ, 572, 407

Belczynski K., Kalogera V., Rasio F. A., Taam R. E., Zezas A., Bulik T., Maccarone T. J., Ivanova N., 2008, ApJS, 174, 223

Berger M. J., Colella P., 1989, Journal of Computational Physics, 82, 64

Berger M. J., Oliger J., 1984, Journal of Computational Physics, 53, 484

Bode P., Ostriker J. P., 2003, ApJS, 145, 1

van den Bosch F. C., 2020, Astrophysical Flows. Lecture Notes, Yale University, `http://www.astro.yale.edu/vdbosch/Astrophysical_Flows.pdf`

Bourne M. A., Sijacki D., 2017, MNRAS, 472, 4707

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150

Braun R., Bonaldi A., Bourke T., Keane E., Wagg J., 2019, arXiv e-prints, p. arXiv:1912.12699

Bryan G. L., Norman M. L., 1997, in Clarke D. A., West M. J., eds, Astronomical Society of the Pacific Conference Series Vol. 12, Computational Astrophysics; 12th Kingston Meeting on Theoretical Astrophysics. p. 363 (arXiv:astro-ph/9710186), doi:10.48550/arXiv.astro-ph/9710186

Bryan G. L., et al., 2014, ApJS, 211, 19

Ciardi B., Ferrara A., Marri S., Raimondo G., 2001, MNRAS, 324, 381

Commerçon B., Debout V., Teyssier R., 2014, A&A, 563, A11

Costa T., Rosdahl J., Sijacki D., Haehnelt M. G., 2018, MNRAS, 479, 2079

Dayal P., Ferrara A., 2018, Phys. Rep., 780, 1

DeBoer D. R., et al., 2017, PASP, 129, 045001

Ding Z., et al., 2022, MNRAS, 514, 3308

Donnari M., et al., 2019, MNRAS, 485, 4817

Dullemond C. P., 2017, Radiative transfer in astrophysics. Lecture Notes, University of Heidelberg

ESO / NAOJ 2016, Image: Schematic diagram of the history of the Universe. Obtained from ESO (European Southern Oberservatory) website, credited there to NAOJ (National Astronomical Observatory of Japan). Accessed May 7, 2025, https://www.eso.org/public/images/eso1620a/

Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, ApJS, 57, 241

Eide M. B., Graziani L., Ciardi B., Feng Y., Kakiichi K., Di Matteo T., 2018, MNRAS, 476, 1174

Eide M. B., Ciardi B., Graziani L., Busch P., Feng Y., Di Matteo T., 2020, MNRAS, 498, 6083

Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, PASA, 34, e058

Ewen H. I., Purcell E. M., 1951, Nature, 168, 356

Fiacconi D., Sijacki D., Pringle J. E., 2018, MNRAS, 477, 3807

Finlator K., Özel F., Davé R., 2009, MNRAS, 393, 1090

Finlator K., Keating L., Oppenheimer B. D., Davé R., Zackrisson E., 2018, MNRAS, 480, 2628

Fragos T., et al., 2013a, ApJ, 764, 41

Fragos T., Lehmer B. D., Naoz S., Zezas A., Basu-Zych A., 2013b, ApJ, 776, L31

Fragos T., Lehmer B. D., Naoz S., Zezas A., Basu-Zych A., 2016, ApJ, 827, L21

Furlanetto S. R., 2006, MNRAS, 371, 867

Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181

Garaldi E., Kannan R., Smith A., Springel V., Pakmor R., Vogelsberger M., Hernquist L., 2022, MNRAS, 512, 4909

Garaldi E., et al., 2024, MNRAS, 530, 3765

Gardner J. P., et al., 2006, Space Sci. Rev., 123, 485

Garrison L. H., Eisenstein D. J., Ferrer D., Tinker J. L., Pinto P. A., Weinberg D. H., 2018, ApJS, 236, 43

Genel S., et al., 2018, MNRAS, 474, 3976

Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

Gnedin N. Y., 2014, ApJ, 793, 29

Gnedin N. Y., Abel T., 2001, New A, 6, 437

Gnedin N. Y., Ostriker J. P., 1997, ApJ, 486, 581

Godunov S. K., 1959, Matematicheskii Sbornik, 47, 271

González M., Audit E., Huynh P., 2007, A&A, 464, 429

Greene J. E., et al., 2024, ApJ, 964, 39

Grove C., et al., 2022, MNRAS, 515, 1854

Haardt F., Madau P., 2012, ApJ, 746, 125

Hadzhiyska B., Liu S., Somerville R. S., Gabrielpillai A., Bose S., Eisenstein D., Hernquist L., 2021, MNRAS, 508, 698

Harikane Y., Nakajima K., Ouchi M., Umeda H., Isobe Y., Ono Y., Xu Y., Zhang Y., 2024, ApJ, 960, 56

Harrison C. M., Costa T., Tadhunter C. N., Flütsch A., Kakkad D., Perna M., Vietri G., 2018, Nature Astronomy, 2, 198

Hernández-Aguayo C., et al., 2023, MNRAS, 524, 2556

Hernquist L., Katz N., 1989, ApJS, 70, 419

Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles. McGraw-Hill

Holmberg E., 1941, ApJ, 94, 385

Hopkins P. F., 2015, MNRAS, 450, 53

Hopkins P. F., 2017, arXiv e-prints, p. arXiv:1712.01294

Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Hopkins P. F., et al., 2018, MNRAS, 480, 800

Hopkins P. F., et al., 2023, MNRAS, 519, 3154

van de Hulst H. C., 1945, Nederlandsch Tijdschrift voor Natuurkunde, 11, 210

Jaura O., Glover S. C. O., Klessen R. S., Paardekooper J. P., 2018, MNRAS, 475, 2822

Jaura O., Magg M., Glover S. C. O., Klessen R. S., 2020, MNRAS, 499, 3594

Jiao H., Brandenberger R., Refregier A., 2024, Phys. Rev. D, 109, 123524

Kannan R., Vogelsberger M., Marinacci F., McKinnon R., Pakmor R., Springel V., 2019, MNRAS, 485, 117

Kannan R., Garaldi E., Smith A., Pakmor R., Springel V., Vogelsberger M., Hernquist L., 2022a, MNRAS, 511, 4005

Kannan R., Smith A., Garaldi E., Shen X., Vogelsberger M., Pakmor R., Springel V., Hernquist L., 2022b, MNRAS, 514, 3857

Kannan R., et al., 2023, MNRAS, 524, 2594

Kannan R., et al., 2025, arXiv e-prints, p. arXiv:2502.20437

Kim C.-G., Ostriker E. C., 2017, ApJ, 846, 133

Klypin A. A., Shandarin S. F., 1983a, MNRAS, 204, 891

Klypin A. A., Shandarin S. F., 1983b, MNRAS, 204, 891

Koehler S. M., Jiao H., Kannan R., 2024, arXiv e-prints, p. arXiv:2412.00182

Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73

Krumholz M. R., Klein R. I., McKee C. F., Bolstad J., 2007, ApJ, 667, 626

Labbé I., et al., 2023, Nature, 616, 266

Levermore C. D., 1984, J. Quant. Spectr. Rad. Transf., 31, 149

Loken C., et al., 2010, in Journal of Physics Conference Series. IOP, p. 012026, doi:10.1088/1742-6596/256/1/012026

Lucy L. B., 1977, AJ, 82, 1013

Lusso E., Worseck G., Hennawi J. F., Prochaska J. X., Vignali C., Stern J., O'Meara J. M., 2015, MNRAS, 449, 4204

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Madau P., Fragos T., 2017, ApJ, 840, 39

Madau P., Haardt F., Rees M. J., 1999, ApJ, 514, 648

Maksimova N. A., Garrison L. H., Eisenstein D. J., Hadzhiyska B., Bose S., Satterthwaite T. P., 2021, MNRAS, 508, 4017

Marinacci F., et al., 2018, MNRAS, 480, 5113

Maselli A., Ferrara A., Ciardi B., 2003, MNRAS, 345, 379

Maselli A., Ciardi B., Kanekar A., 2009, MNRAS, 393, 171

McKinnon R., Torrey P., Vogelsberger M., 2016, MNRAS, 457, 3775

McKinnon R., Torrey P., Vogelsberger M., Hayward C. C., Marinacci F., 2017, MNRAS, 468, 1505

Mellema G., et al., 2013, Experimental Astronomy, 36, 235

Mesinger A., Ferrara A., Spiegel D. S., 2013, MNRAS, 431, 621

Mihalas D., Weibel-Mihalas B., 1984, Foundations of radiation hydrodynamics. Oxford University Press

Mineo S., Gilfanov M., Sunyaev R., 2012a, MNRAS, 419, 2095

Mineo S., Gilfanov M., Sunyaev R., 2012b, MNRAS, 426, 1870

Mo H., van den Bosch F., White S., 2010, Galaxy formation and evolution. Cambridge University Press, doi:10.1017/CBO9780511807244.004

Moseley E. R., Teyssier R., Draine B. T., 2023, MNRAS, 518, 2825

Naab T., Ostriker J. P., 2017, ARA&A, 55, 59

Naiman J. P., et al., 2018, MNRAS, 477, 1206

Nelson D., et al., 2018, MNRAS, 475, 624

Nelson D., et al., 2019, Computational Astrophysics and Cosmology, 6, 2

Nelson E. J., et al., 2021, MNRAS, 508, 219

Nolting W., 2007, Grundkurs Theoretische Physik 6: Statistische Physik. Springer, doi:10.1007/978-3-540-68871-6

Nori M., Baldi M., 2018, MNRAS, 478, 3935

O'Shea B. W., Bryan G., Bordner J., Norman M. L., Abel T., Harkness R., Kritsuk A., 2004, arXiv e-prints, pp astro–ph/0403044

Ocvirk P., et al., 2016, MNRAS, 463, 1462

Ocvirk P., Aubert D., Chardin J., Deparis N., Lewis J., 2019, A&A, 626, A77

Ocvirk P., et al., 2020, MNRAS, 496, 4087

Pacucci F., Mesinger A., Mineo S., Ferrara A., 2014, MNRAS, 443, 678

Pakmor R., Springel V., Bauer A., Mocz P., Munoz D. J., Ohlmann S. T., Schaal K., Zhu C., 2016, MNRAS, 455, 1134

Pakmor R., et al., 2023, MNRAS, 524, 2539

Parsons A. R., et al., 2010, AJ, 139, 1468

Peebles P. J. E., 1970, AJ, 75, 13

Peter A. H. G., Rocha M., Bullock J. S., Kaplinghat M., 2013, MNRAS, 430, 105

Petkova M., Springel V., 2009, MNRAS, 396, 1383

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

Pillepich A., et al., 2018, MNRAS, 475, 648

Planck Collaboration et al., 2016, A&A, 594, A13

Planck Collaboration et al., 2020a, A&A, 641, A1

Planck Collaboration et al., 2020b, A&A, 641, A6

Ponce M., et al., 2019, in Practice and Experience in Advanced Research Computing 2019: Rise of the Machines (Learning). PEARC '19. Association for Computing Machinery, New York, NY, USA, doi:10.1145/3332186.3332195

Press W. H., Schechter P., 1974, ApJ, 187, 425

Press W. H., Teukolsky S. A., 2007, Numerical Recipes. The Art of Scientific Computing, 3 edn. Cambridge University Press

Price D. J., Monaghan J. J., 2007, MNRAS, 374, 1347

Pritchard J. R., Loeb A., 2012, Reports on Progress in Physics, 75, 086901

Rocha M., Peter A. H. G., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Oñorbe J., Moustakas L. A., 2013, MNRAS, 430, 81

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

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

Roy S., Shen X., Barron J., Lisanti M., Curtin D., Murray N., Hopkins P. F., 2025, ApJ, 982, 175

Rybicki G. B., Lightman A. P., 1979, Radiative Processes in Astrophysics. Wiley, doi:10.1002/9783527618170

Scoccimarro R., 1998, MNRAS, 299, 1097

Semelin B., Combes F., Baek S., 2007, A&A, 474, 365

Shen X., Hopkins P. F., Faucher-Giguère C.-A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020, MNRAS, 495, 3252

Shen X., Vogelsberger M., Boylan-Kolchin M., Tacchella S., Naidu R. P., 2024, MNRAS, 533, 3923

Shen X., et al., 2025, arXiv e-prints, p. arXiv:2503.01949

Sijacki D., Springel V., 2006, MNRAS, 366, 397

Silk J., Begelman M. C., Norman C., Nusser A., Wyse R. F. G., 2024, ApJ, 961, L39

Smith A., Kannan R., Garaldi E., Vogelsberger M., Pakmor R., Springel V., Hernquist L., 2022, MNRAS, 512, 3243

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010a, ARA&A, 48, 391

Springel V., 2010b, MNRAS, 401, 791

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V., Yoshida N., White S. D. M., 2001, New A, 6, 79

Springel V., et al., 2005a, Nature, 435, 629

Springel V., Di Matteo T., Hernquist L., 2005b, ApJ, 620, L79

Springel V., et al., 2008, MNRAS, 391, 1685

Springel V., et al., 2018, MNRAS, 475, 676

Springel V., Pakmor R., Zier O., Reinecke M., 2021, MNRAS, 506, 2871

Stappers B. W., et al., 2011, A&A, 530, A80

Strang G., 1968, SIAM Journal on Numerical Analysis, 5, 506

Talbot R. Y., Bourne M. A., Sijacki D., 2021, MNRAS, 504, 3619

Talbot R. Y., Sijacki D., Bourne M. A., 2024, MNRAS, 528, 5432

Tasker E. J., Brunino R., Mitchell N. L., Michielsen D., Hopton S., Pearce F. R., Bryan G. L., Theuns T., 2008, MNRAS, 390, 1267

Teyssier R., 2002, A&A, 385, 337

Teyssier R., 2015, ARA&A, 53, 325

Tingay S. J., et al., 2013, PASA, 30, e007

Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, 438, 1985

Verner D. A., Ferland G. J., Korista K. T., Yakovlev D. G., 1996, ApJ, 465, 487

Vogelsberger M., Zavala J., Loeb A., 2012, MNRAS, 423, 3740

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Vogelsberger M., et al., 2014a, MNRAS, 444, 1518

Vogelsberger M., et al., 2014b, Nature, 509, 177

Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nat. Rev. Phys., 2, 42

Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, MNRAS, 387, 427

Weibel A., et al., 2024, MNRAS,

Weinberger R., et al., 2017, MNRAS, 465, 3291

Weinberger R., et al., 2018, MNRAS, 479, 4056

Weisskopf M. C., Tananbaum H. D., Van Speybroeck L. P., O'Dell S. L., 2000, in Truemper J. E., Aschenbach B., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4012, X-Ray Optics, Instruments, and Missions III. pp 2–16 (arXiv:astro-ph/0004127), doi:10.1117/12.391545

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

White S. D. M., 1976, MNRAS, 177, 717

White S. D. M., 1996, in Schaeffer R., Silk J., Spiro M., Zinn-Justin J., eds, Cosmology and Large Scale Structure. p. 349

Wu X., Kannan R., Marinacci F., Vogelsberger M., Hernquist L., 2019, MNRAS, 488, 419

Xu G., 1995, ApJS, 98, 355

Yuan S., Hadzhiyska B., Bose S., Eisenstein D. J., 2022, MNRAS, 512, 5793

Yuan Y., Martin-Alvarez S., Haehnelt M. G., Garel T., Sijacki D., 2024, MNRAS, 532, 3643

Zel'dovich Y. B., 1970, A&A, 5, 84

Zier O., Kannan R., Smith A., Vogelsberger M., Verbeek E., 2024, MNRAS, 533, 268

Zier O., Kannan R., Smith A., Vogelsberger M., Springel V., Pakmor R., Hernquist L., Koehler S. M., 2025a, THESAN-XL, Unpublished. For reference, see https://doeleadershipcomputing.org/wp-content/uploads/sites/123/2024/11/2025INCITEFactSheets.pdf, p. 76.

Zier O., et al., 2025b, arXiv e-prints, p. arXiv:2503.03806

Erklärung nach § 39 (15) Ordnung 2020 für den Masterstudiengang Physik

Hiermit erkläre ich, dass ich die Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel verfasst habe. Alle Stellen der Arbeit, die wörtlich oder sinngemäß aus Veröffentlichungen oder aus anderen fremden Texten entnommen wurden, sind von mir als solche kenntlich gemacht worden. Ferner erkläre ich, dass die Arbeit nicht – auch nicht auszugsweise – für eine andere Prüfung oder Studienleistung verwendet wurde.

Frankfurt am Main, den 31.05.2025

Sonja Marie Köhler