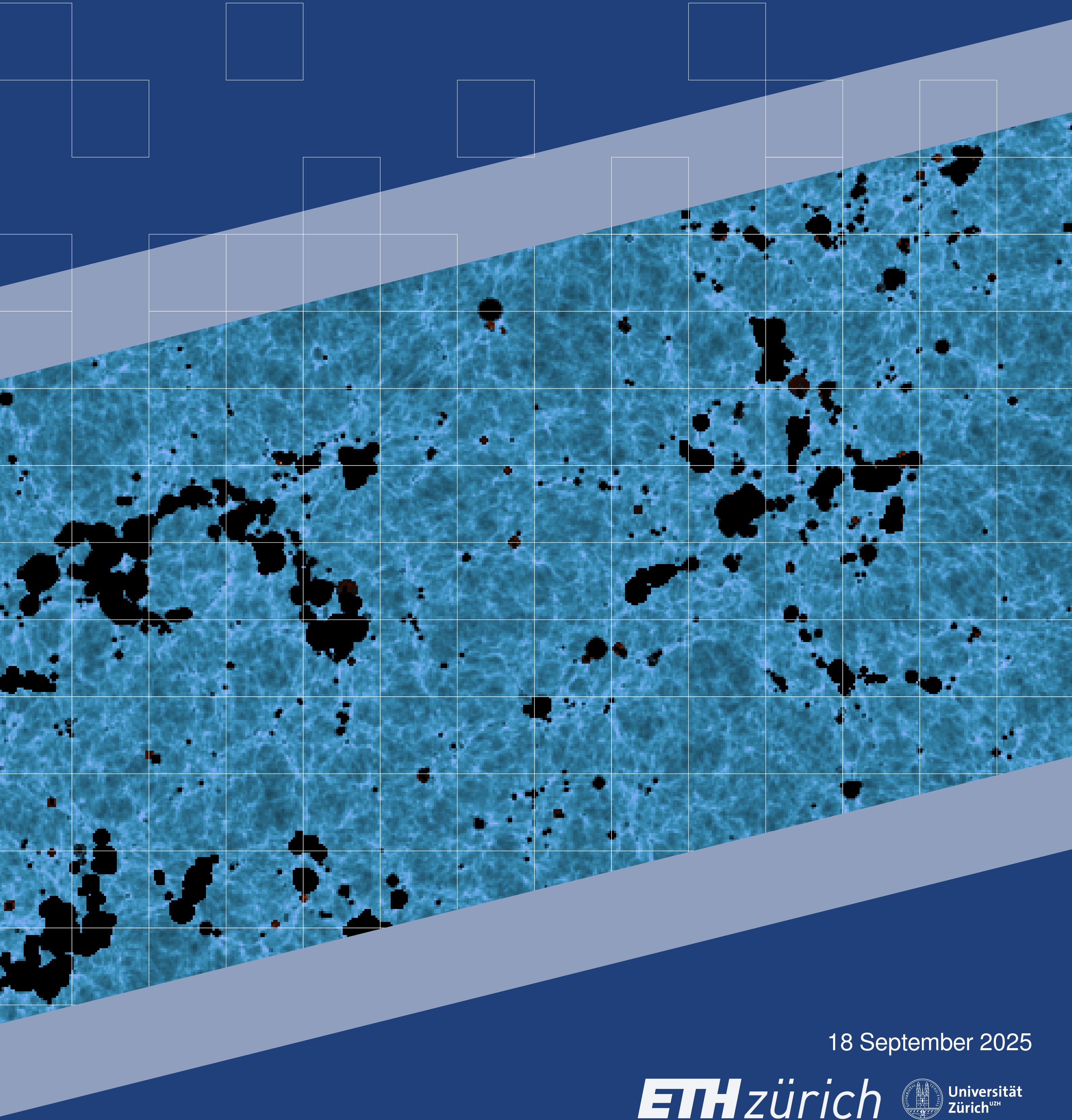
# Simulating the EOR with self-consistent growth of galaxies

## Master's Thesis

**Author:** Rémy Moll

**Supervisors:** Prof. Aurel Schneider
Prof. Alexandre Refregier



18 September 2025

ETH zürich

Universität Zürich

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

# Declaration of originality

The signed declaration of originality is a component of every written paper or thesis authored during the course of studies. **In consultation with the supervisor**, one of the following two options must be selected:

☐ I hereby declare that I authored the work in question independently, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used no generative artificial intelligence technologies[1].

☐ I hereby declare that I authored the work in question independently. In doing so I only used the authorised aids, which included suggestions from the supervisor regarding language and content and generative artificial intelligence technologies. The use of the latter and the respective source declarations proceeded in consultation with the supervisor.

**Title of paper or thesis**:

**Authored by**:
*If the work was compiled in a group, the names of all authors are required.*

| **Last name(s):** | **First name(s):** |
|---|---|
| | |
| | |
| | |
| | |

With my signature I confirm the following:

- I have adhered to the rules set out in the Citation Guidelines.
- I have documented all methods, data and processes truthfully and fully.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for originality.

| **Place, date** | **Signature(s)** |
|---|---|
| | |
| | |
| | |
| | |

*If the work was compiled in a group, the names of all authors are required. Through their signatures they vouch jointly for the entire content of the written work.*

---

[1] For further information please consult the ETH Zurich websites, e.g. https://ethz.ch/en/the-eth-zurich/education/ai-in-education.html and https://library.ethz.ch/en/researching-and-publishing/scientific-writing-at-eth-zurich.html (subject to change).

We present `BEoRN`, a semi-numerical simulation suite that uses a simplified description of radiation sources to efficiently generate maps of the cosmic dawn and the epoch of reionization. In its original formulation, the framework assumed exponential halo growth with a fixed accretion rate. We show how this assumption introduces systematic inaccuracies in the predicted observables of reionization. In this work, we implement a more realistic description of halo mass growth that follows the mass accretion history of individual halos. We achieve this consistent treatment of halo growth by incorporating individual accretion histories extracted from the merger trees provided by the `THESAN` simulations. This required an extensive refactoring of the `BEoRN` code base, redesigning the computation of the radiative source properties to accommodate variable growth while ensuring scalability and computational efficiency. We describe these modifications in detail and validate the improved suite against the original implementation. We demonstrate that the refined treatment alters the timing and morphology properties of the reionization. Finally, we discuss how this new version of `BEoRN` provides a more robust platform for exploring reionization scenarios, and we outline planned extensions and applications.

## Contents

# 1. Introduction

The formation of the earliest astrophysical objects, such as stars, galaxies, and black holes, have a profound influence on the evolution of the universe. Though poorly understood, these events have shaped the characteristics of our current universe, including the structure and distribution of matter itself. Despite the milestones achieved in observational cosmology, many aspects of the early universe and its dark ages remain difficult to probe. While traditional measurements provide insights into relatively recent epochs, and the cosmic microwave background (CMB) serves as an early snapshot of the universe, the dark ages are incompatible with direct observations. They represent a critical link between the late-time universe and the primordial conditions that has remained largely unexplored.

The epoch of reionization (EOR) spans the time period from the end of the dark ages until the time when the universe is fully ionized again. It is a period of complex interactions between matter and radiation but it is crucial to understand as it sets the stage for the subsequent evolution of the universe. Beyond its impact on the late universe, a detailed understanding of the reionization process has been shown to provide new and competitive constraints on the current cosmological model (e.g. Mao et al. (2008); McQuinn et al. (2006); Schneider et al. (2023) ). Understanding and being able to model the EOR is therefore essential for a comprehensive picture of cosmology.

The dark ages of the universe refer to the period after recombination where the primordial atoms remain neutral. They are characterized by the total lack of sources of radiation (beyond the radiation background). The dominant interactions during that period are either gravitational or due to the cooling of the primordial gas. During that time the formation of the first stars is obstructed by the lack of efficient cooling mechanisms in the absence of heavier nuclei. With the simplest cooling channel being the deexcitation of atomic hydrogen, the gas inside a virialized structure can only collapse if the enclosed mass is high enough. This so-called atomic cooling limit sets a minimum mass for the halos that can host star formation at around $10^8 M_\odot$. Other cooling channels such as the deexcitation of molecular hydrogen are suppressed by the emission of photons from the first stars.

The formation of the first stars marks the end of the dark ages. These so-called population III stars have zero metallicity and very distinct characteristics compared to later generations of stars. Their existence has not been confirmed observationally but they are thought to have shaped the subsequent formation of stars and galaxies and to have played a crucial role in the reionization of the universe (e.g. Mebane et al. (2020)).

Driven by the newly formed stars and galaxies, reionization is explained as an inside-out process (e.g. Iliev et al. (2006)) expanding from within the halos that host the first galaxies. The ionizing radiation emitted by these sources reaches the intergalactic medium (IGM) and creates ionized bubbles that grow and eventually overlap to fully ionize the universe again.

While reionization marks the gradual disappearance of neutral hydrogen, the preceeding abundance during the dark ages and cosmic dawn allows for an additional mode of observation: the 21-cm line. Due to the hyperfine transition of neutral hydrogen there is a characteristic emission or absorption of photons at a frequency of $1420\ \mathrm{MHz}$. The strength of this signal depends on the local conditions, encoded by the spin temperature. The redshifting of the photons allows to probe different epochs through the observed frequency.

The detection of the 21-cm signal of reionization is a major goal of current and upcoming radio telescopes, for instance the Square Kilometer Array (Koopmans et al. 2015, SKA) or the Hydrogen Epoch of Reionization Array (DeBoer et al. 2017, HERA) . These instruments are expected to detect the power spectrum of the 21-cm signal, providing further insights into the dynamics of the early universe. In particular the low-frequency component SKA-Low is expected to have the sensitivity to image the 21-cm signal directly and to produce maps of the ionization field during the EOR.

Beyond observations, an additional pillar of understanding the EOR is the modeling and simulation of the universe during that time. The main purpose of simulations is to constrain the EOR observables. Combined with the first observations, simulations will generate a wealth of information about the early universe at a range of redshifts that has previously been inaccessible. With the highest sensitivity and resolution forecasted for these observations, the simulations must be able to capture the full dynamic range of the interactions from the small-scale physics of star formation and feedback to the large-scale structure of the universe.

State-of-the-art simulations need to implement a range of physical processes, including gravitational interactions, hydrodynamics, radiative transfer, and feedback mechanisms. Prominent examples include for instance the THESAN simulations (Garaldi et al. 2022; Kannan et al. 2021; Smith et al. 2022) . Another approach is to use ray-tracing algorithms which give detailed descriptions of the radiative transfer (see e.g. Mellema et al. (2006)). These simulations are computationally expensive and cannot be used to to repeatedly explore the parameter space of reionization.

This work presents BEoRN, the *Bubbles during the Epoch of Reionization Numerical simulator* by Schaeffer et al. (2023). In its simplest description BEoRN is the implementation of the "halo model of reionization" by Schneider et al. (2023). In this model the radiative interactions are described through spherically symmetric profiles around sources embedded in dark matter halos. This effectively reduces the dimensionality of the radiative transfer problem. BEoRN uses the one-dimensional (1-d) profiles generated by this model to paint the 3-d space around halos which are obtained from a large-scale N-body simulation. A distinguishing feature of BEoRN is the self-consistent treatment of the expansion of the affected regions around the sources. This approach allows for simulations at the largest scales while still taking into account the core processes of reionization. The computational efficiency of BEoRN makes it suitable to explore the underlying parameters.

The first iteration of BEoRN focused on the impact of parameters related to the emission of photons whereas this work focuses on the effects of gravitational mass accretion. We show that the radiation profiles are sensitive to the growth rate and that the mass accretion history provdied by N-body simulations is too complex to be captured by simple parametrizations. Our improved version of BEoRN permits a more consistent treatment by considering the individual mass accretion history of each source. We demonstrate the resulting measurable effects on the 21-cm observables when compared to the simpler models.

This report is structured as follows: Section 2 describes the details of the simulation procedure, including the underlying model. Section 3 explains how mass evolution is modeled and its impact on the flux profiles used by BEoRN. In Section 4 we give an overview of the implementation and changes required by the refined modeling. Section 5 details the validation we perform on the refined procedure and in Section 6 we compare the resulting signals to quantify the impact of different models of mass accretion. Section 7 summarizes our findings and discusses potential future improvements.

Note that throughout this report, physical distance units are specified with the prefix "p", while comoving distance units are specified with the prefix "c".

# 2. Overview of the BEoRN framework

This section presents the model describing the sources of radiation that drive reionization. We explain how BEoRN implements this model to generate 3-d maps of the IGM during the epoch of reionization. The code of BEoRN as well as usage instructions are publicly available under https://github.com/cosmic-reionization/BEoRN[1].

## 2.1. The halo mass model of reionization

The distinguishing feature of BEoRN is the parametrization of sources of radiation through the properties of their host dark matter halo. This approach is based on the model presented by Schneider et al. (2023) and gives a description of the 21-cm signal through the treatment of flux profiles around sources. Using these profiles and allowing them to overlap enables efficient computations of the ionization state of the IGM without

---

[1] For an explicit overview of the changes referenced here, please refer to https://github.com/moll-re/BEoRN

the need for detailed radiative transfer simulations. The following description follows the derivation made by Schneider et al. (2021); Schneider et al. (2023) and we refer to these works for components that are not defined here.

The model describes the emission of ionizing radiation by galaxies. It assumes that the sources are hosted by dark matter halos and expresses the star formation and radiation properties as a function of the halo mass $M_\mathrm{h}$ and mass accretion rate $\dot{M}_\mathrm{h}$. The modeling of the halo mass evolution is discussed in Section 3. The model itself simply considers an arbitrary but known halo mass accretion history $M_\mathrm{h}(z)$.

The emission of radiation revolves around the star formation rate $\dot{M}_\star$ which is simply assumed to be proportional to the halo mass accretion rate via

$$\dot{M}_\star = f_\star(M_\mathrm{h}) \cdot \dot{M}_\mathrm{h} \tag{1}$$

where the star formation efficiency $f_\star$ introduces a mass dependence that enables the suppression of star formation in low mass halos and the implementation of a cooling limit. Motivated by abundance matching, Schneider et al. (2023) use the double power law

$$f_\star(M_\mathrm{h}) = f_{\star,0} \cdot \frac{2\left(\frac{\Omega_b}{\Omega_m}\right)}{\left(\frac{M_\mathrm{h}}{M_\mathrm{p}}\right)^{\gamma_1} + \left(\frac{M_\mathrm{h}}{M_\mathrm{p}}\right)^{\gamma_2}} \cdot S(M_\mathrm{h}) \tag{2}$$

with $M_\mathrm{p}$ the pivot mass where the efficiency peaks, and $\gamma_1$, $\gamma_2$ are the low and high mass slopes, and $f_{\star,0}$ is the normalization chosen at a value of $f_{\star,0} \sim 0.1$. An additional suppression factor $S(M_\mathrm{h})$ is introduced to account for reduced star formation in low mass halos whose effect is discussed by Schaeffer et al. (2023).

### 2.1.1. Expression of the profiles

Derived from the star formation rate the halo model predicts the production and distribution of photons in three distinct energy bands: Lyman-$\alpha$ photons, x-ray photons, and ionizing UV photons. Each of these bands has a different effect on the IGM and is treated separately.

#### *2.1.1.1. Lyman-$\alpha$ flux profile*

Lyman-$\alpha$ photons induce a coupling between the spin temperature and the kinetic temperature of the gas. This effect, known as the Wouthuysen-Field effect (Field 1958; Wouthuysen 1952), causes a characteristic absorption and emission feature in the 21-cm signal. Before the cosmic dawn the gas temperature is lower than the background temperature and the coupling leads to absorption. As the first stars heat the gas, the signal transitions to emission.

The Lyman-$\alpha$ flux profile around a halo of mass $M$ at redshift $z$ is given by

$$\rho_\alpha(r \mid M, z) = \frac{(1+z)^2}{4\pi r^2} \cdot \sum_{n=2}^{n_m} f_n \cdot \varepsilon_\alpha(\nu') \cdot f_\star \cdot \dot{M}(z' \mid M, z) \tag{3}$$

which depends on the lookback redshift $z'$ at which the photon was emitted, hence the expression in terms of $\nu' = \nu \cdot (1+z')/(1+z)$.

Finally the position dependent coefficient is expressed as

$$x_\alpha(r \mid M, z) = \frac{1.81 \cdot 10^{11}}{1+z} \cdot S_\alpha(z) \cdot \rho_\alpha(r \mid M, z) \tag{4}$$

using a suppression factor $S_\alpha(z)$.

#### 2.1.1.2. Temperature profile

The temperature of the IGM around the sources has a strong impact on the 21-cm signal. It is governed by two effects: the cooling temperature of the adiabatically expanding universe and the heating due to X-ray photons emitted by the newly formed stars. The heating profile $\rho_h$ follows

$$\frac{3}{2} \cdot \frac{\mathrm{d}\rho_h(r \mid M, z)}{\mathrm{d}z} = \frac{3\rho_h(r \mid M, z)}{1+z} - \frac{\rho_{\mathrm{xray}}(r \mid M, z)}{k_B(1+z)H(z)} \tag{5}$$

which is based on the flux profile of x-ray photons $\rho_{\mathrm{xray}}(r \mid M, z)$. The Boltzmann constant is given by $k_B$ and $H(z)$ is the Hubble parameter at redshift $z$.

#### 2.1.1.3. Reionization profile

The comoving ionized volume around a source of ionizing photons satisfies the differential equation

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\dot{N}_{\mathrm{ion}(t)}}{\overline{n}_H^0} - \alpha_B \cdot \frac{C}{a^3} \cdot \overline{n}_H^0 \cdot V \tag{6}$$

where $\alpha_B$ is the recombination coefficient, $C$ is the clumping factor, $a$ is the scale factor, and $\overline{n}_H^0$ is the mean density of hydrogen. We expressed this volume in terms of $\dot{N}_{\mathrm{ion}}$ the rate of change of the total number of ionizing photons per baryon. This description is not dependent on the frequency since we simply consider the contribution of all photons above the ionization threshold.

Ionizing photons, i.e. photons with energies above $13.6\ \mathrm{eV}$, experience a large optical depth, which justifies the sharp cutoff of the ionization profile at the bubble radius $R_b = \sqrt[3]{\frac{3}{4\pi} V(M, z)}$. The radial dependence of the ionized fraction is expressed through the Heaviside step function $\theta_{\mathrm{H}}$ and reads

$$x_{\mathrm{HII}}(r \mid M, z) = \theta_{\mathrm{H}}\left[R_b(M, z) - r\right] \tag{7}$$

### 2.1.2. Expression of the reionization signal

The above profiles express the local effect of radiation around a single halo as a 1-d simplification. A representation of a typical profile can be seen in Figure 3. Using an estimate of the spatial distribution of halos, these profiles can be applied to generate a full 3-d map if we assume spherical symmetry.

The observable signal of the 21-cm line is now a spatially dependent quantity obtained from a combination of the mapped quantities. It is expressed as the differential brightness temperature $dT_{\mathrm{b}}$ which describes the contrast between the foreground and the background radiation emitted by the CMB. Following e.g. Pritchard & Loeb (2012) an expression for $dT_{\mathrm{b}}$ is given by

$$dT_{\mathrm{b}}(\boldsymbol{x}, z) \simeq T_0(z) \cdot x_{\mathrm{HI}}(\boldsymbol{x}, z) \cdot (1 + \delta_b(\boldsymbol{x}, z)) \cdot \frac{x_\alpha(\boldsymbol{x}, z)}{1 + x_\alpha(\boldsymbol{x}, z)} \cdot \left(\frac{1 - T_{\mathrm{CMB}}(z)}{T_{\mathrm{gas}}(\boldsymbol{x}, z)}\right) \tag{8}$$

where the background radiation originates from the CMB. The previous considerations allow us to infer the values of $x_{\mathrm{HI}}$, $x_\alpha$, and $T_{\mathrm{gas}}$ from $x_{\mathrm{HII}}$, $\rho_\alpha$, and $\rho_h$ respectively. The baryonic overdensity $\delta_b$ is assumed to trace the dark matter overdensity $\delta_{\mathrm{dm}}$ which is obtained from the underlying N-body simulation. The amplitude of the signal is given by

$$T_0(z) = 27 \cdot \frac{\Omega_b h^2}{0.023} \cdot \sqrt{\frac{1+z}{10} \frac{0.15}{\Omega_m h^2}}\ \mathrm{mK} \tag{9}$$

where $\Omega_m$ and $\Omega_b$ are the matter and baryonic density parameters.

## 2.2. Simulation steps

The simulation procedure revolves around the implementation of the above spherical radiation profiles around halos. We give a brief overview of the main steps here. For a more detailed description of the implementation we refer to Schaeffer et al. (2023). We discuss our improvements and changes to the original implementation in Section 4.

### 2.2.1. Halo catalog - N-body simulations

As a prerequisite, the generation of map data requires a spatial distribution of dark matter halos as well as the underlying density field. Each snapshot can be used to generate a map at the corresponding redshift. `BEoRN` has been successfully validated against mock maps generated by `21cmFAST` (Mesinger et al. 2011) and produced meaningful signals by using the PKDGRAV3 (Potter et al. 2016) simulation suite.

### 2.2.2. Computation of radiation profiles

In accordance with the astrophysical parameters set by the user, radiation profiles are computed in order to be applied according to the halo catalog in a subsequent step. The computation of the profiles is deliberately separated from their application onto halos for a more efficient processing. The range of halo masses and redshifts covered by this precomputation is largely determined by the underlying halo catalog since it provides upper bounds on the halo masses[2].

### 2.2.3. Painting with the binned approach

The last step consists of applying the ionization and temperature distributions defined by the profiles onto a 3-d grid. This is done by iterating over the halos in the catalog and using their corresponding profile. For a given profile a 3-d kernel is generated and mapped onto the grid via convolution using the fast fourier transform implemented by the `astropy` (Astropy Collaboration et al. 2022) package. We refer to this procedure as "painting" since the addition of the contributions of each halo allows us to sequentially build up the final map. In general, contributions from multiple halos can overlap without any restrictions. The ionization map is treated specially: In order to conserve the overall number of ionizing photons, we ensure that each cell is only ionized once.

This final step ensures consistent painting: While the contributions to the temperature and coupling maps can be simply added, the ionization map requires a binary treatment. Cells cannot be ionized past a value of $x_{\mathrm{HII}} = 1$. Discarding the excesses would violate photon conservation. A redistribution of the excess photons to neighboring cells is performed to ensure that the total number of ionizations matches the total number of emitted photons.

The usage of precomputed profiles is crucial to the efficiency of the simulation but it introduces a discretization in halo mass since each halo is assigned the profile of the closest mass bin. The effect of this simplification has been shown to converge for a sufficient number of mass bins by Schaeffer et al. (2023).

### 2.2.4. Derivation of global quantities

The global signal as well as the power spectrum are derived from the map data to be compared to other models or observations. Being derived from a full 3-d N-body simulation, the results are sensitive to the underlying cosmology and the detailed profile modeling. This means that the results also depend on the underlying astrophysical model.

---

[2] The minimum halo mass that needs to be considered is already constrained by the atomic cooling limit. Depending on the mass resolution of the simulation it might not even be reached.

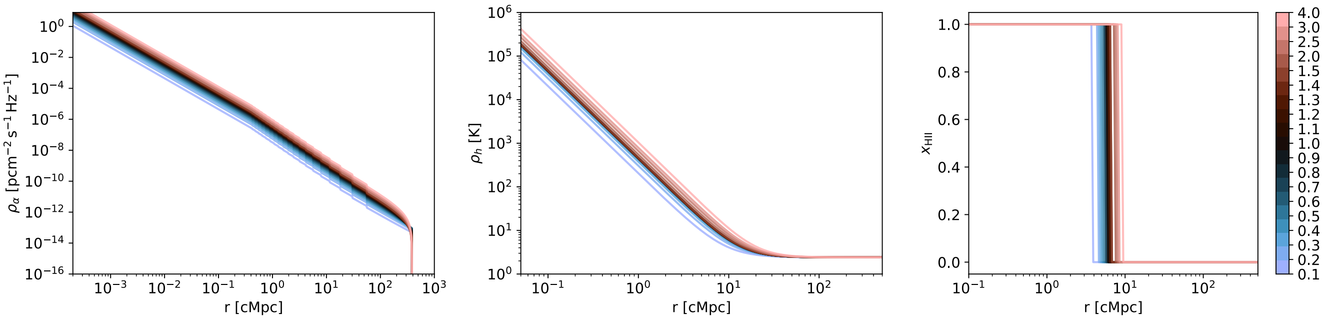


Figure 1: Flux profiles around halos with varying accretion rates. *Left:* Profile of the Lyman-$\alpha$ coupling coefficient $\rho_\alpha$. *Center:* Profile of the heating profile $\rho_h$. *Right:* Ionization fraction profile $x_{\rm HII}$. The effect of different mass accretion rates is visualized by the color gradient where bluer colors correspond to lower accretion rates and redder colors to higher accretion rates.

## 3. Halo mass history

This section shows the impact of the halo growth on the resulting radiation profiles and motivates the need for a more precise treatment of the halo mass history. We show how to leverage simulation data for an improved consistency during the simulation.

### 3.1. Modeling mass accretion

As described in Section 2.1 the fundamental assumption of `BEoRN` is the halo model of reionization by Schneider et al. (2023). It describes how observables of reionization can be parametrized in terms of the halo mass and more specifically its rate of change since they are derived from the star formation rate expressed in Eq. 1. In this simplified model, for a given star formation efficiency the halo mass history is the single most impactful property besides the mass itself.

`BEoRN`'s goal is to provide simulations of the map-level contributions to the 21-cm signal, meaning that we cannot rely on a distribution of halo masses and accretion rates alone. Instead, `BEoRN` leverages large scale N-body simulations to provide a spatial distribution of halos. Halo growth is considered to follow the exponential growth model

$$M_{\rm h}(z) = M_{\rm h}(z_0) \cdot \exp[-\alpha(z - z_0)] \tag{10}$$

where $\alpha = -\frac{\dot{M}_{\rm h}}{M_{\rm h}}$ is a free parameter describing the specific mass accretion rate. Following Dekel et al. (2013) a value of $\alpha = 0.79$ was assumed for all halos, independent of their mass or redshift in the initial version. This means that the requirements for the simulation data were minimal: Only a single halo catalog at a given redshift was required to generate a map at that redshift.

Using a simple exponential growth model is a significant simplification of the complex and time-sensitive process of halo growth (e.g. McBride et al. (2009)). Another limitation is the assumption of a constant accretion rate $\alpha$ for all halos, independently of their position, mass or redshift. In a realistic scenario we expect to observe significant stochasticity of the accretion process as well as systematic effects from the halo mass and the redshift. From a statistical perspective, this has been investigated by Schneider et al. (2021) who also consider a halo growth following the extended Press-Schechter formalism. This more detailed treatment shows that in particular small scales deviate from the simple exponential growth model. From a simulation perspective, an even more precise treatment is possible since the growth history of each halo is already encoded in the successive snapshots of the N-body simulation. Ignoring this information introduces inconsistencies by painting halos using profiles that might not reflect their actual growth history.

### 3.2. Effect on radiation profiles

To illustrate the necessity of a more precise treatment of the halo mass history, we first investigate the effect of different mass accretion rates on the resulting radiation profiles. To this end, we consider halos at fixed masses and vary their accretion rates around the fiducial value of $\alpha = 0.79$.

Figure 1 shows the three relevant profiles, computed for $M_\mathrm{h} = 6.08 \cdot 10^{11} M_\odot$. The variation of the accretion rate leads to noticeable differences in all three profiles, even at high radial distances. There is a clear and consistent trend for all three profiles: Higher accretion rates lead to higher fluxes, i.e. an effect that is more outreaching. This is expected as a higher accretion rate leads to a higher star formation rate and thus to the production of more photons. Note that the spread of the profiles seems symmetric. This is due to the logarithmic scaling of the plot and the higher range of alpha values above the fiducial value. In other words, the magnitude of the effect seems to be the same when shifting the accretion above or below the fiducial value.

This picture is more complex once we consider a distribution of accretion rates instead of a single value across all halos. We note that the dominating factor when considering a distribution is the contribution from the mean accretion rate. The scatter around the mean has a significantly smaller effect. We do not elaborate on the stochasticity of the accretion rate since the usage of N-body simulations allows for a more sophisticated investigation. Instead of assuming pure stochasticity we can extract the actual growth history of each halo and use it to assign a more meaningful accretion rate.

## 3.3. Merger trees

### 3.3.1. Using THESAN simulation data

In order to generate precise map-level predictions of the 21-cm signal, `BEoRN` combines the halo model of reionization with large-scale N-body simulations which provide realistic snapshots of the dark matter distribution. They give a spatial context to the generated profiles.

As described in Section 2 `BEoRN` was initially used to postprocess the PKDGRAV3 (Potter et al. 2016) simulation suite and to obtain a meaningful signal capable of constraining astrophysical parameters related to star formation. The aim of this thesis is not to merely increase the precision but to develop a proof of concept that leverages the mass history which can be extracted directly from the simulation to refine the underlying model.

To this end, we use the publicly available data from the THESAN simulation suite (Garaldi et al. 2022; Kannan et al. 2021; Smith et al. 2022) . The THESAN-DARK simulation in particular provides a dark-matter-only simulation that conveniently already includes halo catalogs and merger trees generated by the `LHaloTree` tree builder by Springel et al. (2005). This will allow us to extract the growth of each halo across different snapshots without significant preprocessing.

With a box length of $95.5\ \mathrm{cMpc}$ the simulation provides a sufficient volume to avoid box size effects (e.g. Iliev et al. (2014)) while still allowing us to iterate quickly and test the refined model without excessive computational cost. The simulation has two variants with different mass resolutions: THESAN-DARK 1 with $2100^3$ particles for a mass resolution of $3.70 \cdot 10^6 M_\odot$ per particle and THESAN-DARK 2 with $1050^3$ particles for a mass resolution of $2.96 \cdot 10^7 M_\odot$ per particle. Unless specified otherwise we use THESAN-DARK 2 since it provides a good compromise between good resolution and computational cost. We make use of THESAN-DARK 1 to perform convergence tests as described in Section 5.

### 3.3.2. Main progenitor branch

Growth of structure in $\Lambda\mathrm{CDM}$ is hierarchical: Small structures form first and merge to form larger structures. The growth of halos can be represented using merger trees. These tree-like structures describe the halo history in terms of the mergers of its smaller progenitors. A merger tree is constructed by linking halos in consecutive snapshots of the simulation where each halo as a single descendant but potentially multiple progenitors.

The main progenitor is the most massive progenitor and serves as a tracer of the halo mass history if we assume that the halo mass growth is dominated by mergers. Beyond that, we expect the main progenitor to be most representative of the baryonic conditions inside and outside the halo as the merger occurs. For the identification of accretion rates for `BEoRN` we therefore focus solely on the main progenitor branch of each halo.

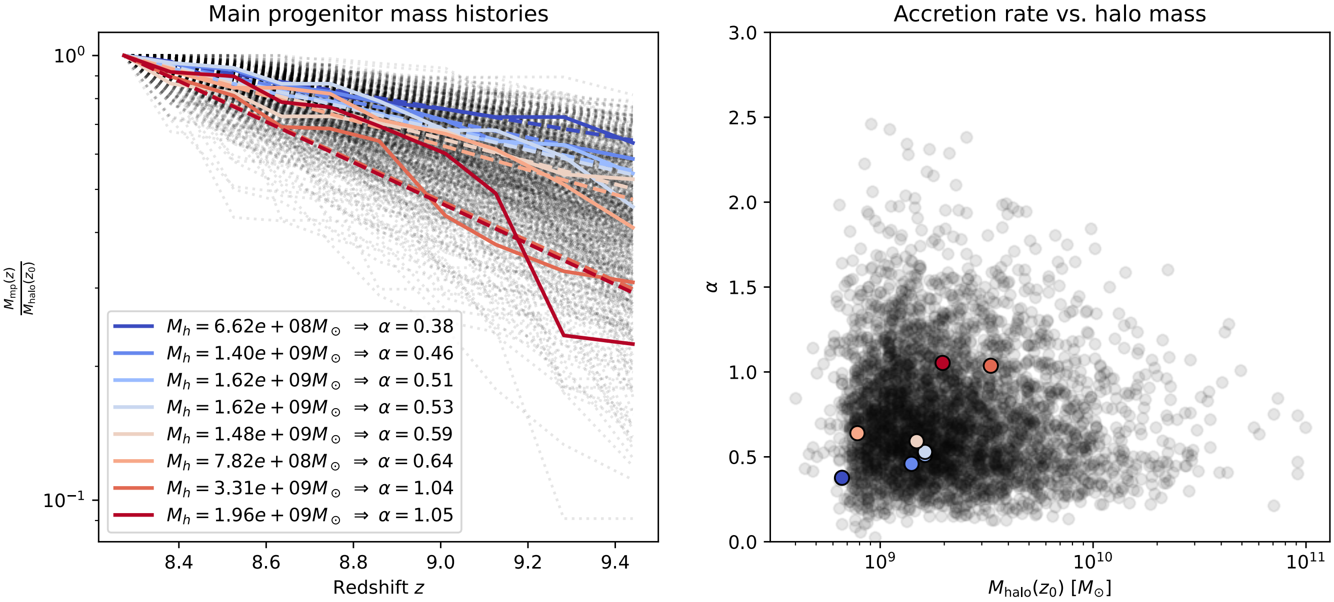


Figure 2: Usage of merger tree fitting to obtain accretion rate estimates. *Left:* Collection of normalized main progenitor branches with mass $M_{\rm mp}$ starting at $z = 10.3$ and looking back over $n = 10$ snapshots. Select histories and their corresponding exponential fits are highlighted. *Right:* Distribution of best-fit accretion rates $\alpha$ for all halos at $z = 8.29$.

Reducing the breadth of the merger tree reduces the data volume significantly and allows us to implement the tree handling in memory without excessive computational cost. For this purpose we provide a simple implementation of a tree walker that copies the simplified trees to a single file for easier access. Other preprocessing is not required, which allows `BEoRN` to keep all parameters related to the mass history as free parameters to be specified at runtime.

### 3.3.3. Fitting procedure

The restriction to the main progenitor corresponds to a reduction of the dimensionality of the mass history to a one-dimensional function of redshift compatible with the orginal assumption of an exponential growth model as in Eq. 10.

We use a linear regression in log-space to obtain estimates of the accretion rate $\alpha$ for each halo. This is implemented in a vectorized fashion to allow for efficient processing of the full dataset. For this fit we enforce the current halo mass as a boundary condition. This prevents inconsistent fits where the latest fitted mass deviates from the actual current halo mass. As a visualization of the fitting procedure Figure 2 shows a collection of normalized main progenitor branches starting at $z = 8.29$ and looking back over $n = 10$ snapshots. After fitting we overlay the estimated exponential growth history for a selection of halos. The right panel shows the distribution of best-fit accretion rates $\alpha$ for all halos at $z = 8.29$. Given the relative low mass of the halos we observe a strong clustering of accretion rates around a value of $\alpha \approx 0.5$. Outliers with significantly deviating values appear nevertheless and are not linked to a specific mass range.

Similarly to the halo mass itself, the accretion rate can then be taken into account during the painting procedure by selecting a profile corresponding to the halo mass and accretion rate of each halo. Consequently, the accretion rate is binned as well and the range that is covered during the painting is finite. We leave this as a free parameter to be specified at runtime.

## 3.4. Resulting accretion rates

In order to obtain a sensible range of $\alpha$ values to cover during the painting procedure, we investigate the global result of the fitting procedure. Our method of fitting trades speed and convenience for absolute precision: Not all halos are well represented in the merger tree and no further processing is done beyond the consideration of the tree. Additionally, we need to account for unphysical or incomplete histories due to limitations of the

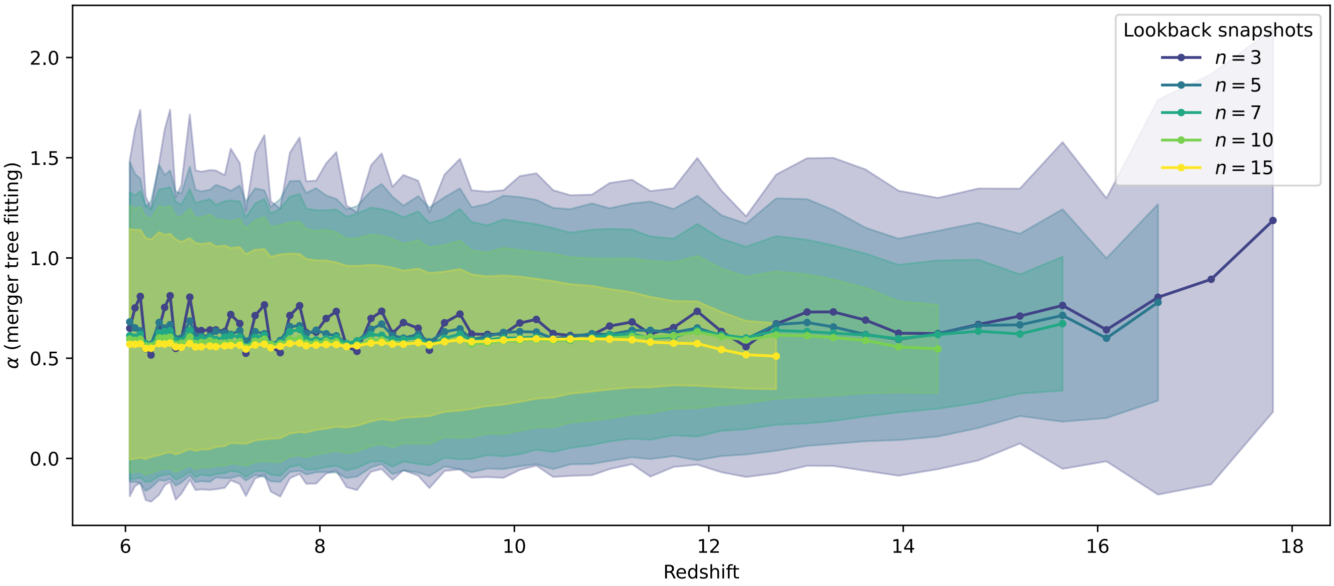


Figure 3: Evolution of the mean of the fitted accretion rates and the $1\sigma$ standard deviation (shaded area). For a given snapshot we consider different numbers of lookback snapshots $n$.

halo finder. We discuss this step in Section 4. For the current investigation we disregard these halos and only consider well-behaved, fully resolved trees. Figure 3 shows how the fitted accretion rate evolves when starting from the different snapshots. We plot the mean and $1\sigma$ standard deviation of the resulting distribution of $\alpha$ values. We consider different lookback lengths with the goal of assessing the stability of the fitting procedure.

We observe a clear stabilization of the mean accretion for longer lookbacks. Not only does it make sense to consider longer lookbacks because of their causal connection, but also because it helps to absorb short-term fluctuations most likely introduced by the halo finder. This is especially noticeable in the first few snapshots where mean and $1\sigma$ uncertainty are significantly higher. This is probably due to the overabundance of low mass halos whose mass history is more erratic and harder to reconstruct.

Numerically, the advantage of longer lookbacks is the stabilization of the fit leading to reduced scatter in the resulting distribution. We note that these behaviors stabilize once we consider around $n = 10$ snapshots of lookback. Both the mean and standard deviation follow a stable trend and the mean settles at $\alpha \approx 0.6$. We attribute the slight offset of the means to the fact that discarding incomplete trees favors more massive halos at higher lookbacks. These halos are more stable in terms of detection by the halo finder and are expected to have fewer fluctuations.

Physically, the lookback time is motivated from the flux profiles of the halos themselves. Due to their size in the $\mathrm{cMpc}$ range (see e.g. Figure 1) we attribute to each profile a timescale during which there is a causal effect on the region defined by the extent of that profile. For a profile of radius $\sim 100\ \mathrm{cMpc}$ this time is of the order of $\Delta t = 300\ \mathrm{Myr}$, corresponding to $\Delta z = 4$ (when looking back from a redshift of $z = 8$). Given the spacing of snapshots in THESAN using $n = 10$ snapshots still lies below the causal range. Since the fitted behavior seems to stabilize we suggest to not go beyond that since the consideration of additional snapshots slows down the simulation considerably.

## 4. Implementation of changes

This section describes the adaptations that were necessary in order to utilize the individual treatment of halo mass accretion histories in BEoRN. We distinguish between necessary changes that were required to implement the underlying model and secondary changes that affect the quality of the simulation outputs indirectly.

### 4.1. Profile generation taking into account halo mass history

For each halo we require a flux profile that matches the halo properties which now include the accretion rate additionally to the mass and the redshift. The profiles are generated in a preprocessing step following the redshifts of the snapshots and the mass and accretion bins defined in the configuration. Since the dynamic range of accretion rates is large, the resulting parameter space rapidly expands. The computation of the profiles therefore utilizes vectorized operations to achieve reasonable runtimes.

Note that the precomputation of the profiles introduces another “second degree” inconsistency: The flux profile attributes a radiative behavior to the halo that is motivated by its history. This is repeated for each snapshot creating possibly conflicting histories. In the case of stable halo growth this is not a problem but in the case of erratic growth (e.g. major mergers) this can lead to unphysical behavior. A more consistent approach would be to assume a more flexible mass growth model that distinguishes different regimes of growth. This would require a much more complex handling of the precomputed profiles and is beyond the scope of this work. The current approach remains a good approximation for the majority of halos.

### 4.2. Parallel painting of profile bins

Similarly to the computation of profiles, the painting step is affected by the increased parameter space. `BEoRN`’s fast simulation times revolve around the crucial simplification of the halo model: Halos with the same core properties are treated identically and can be mapped onto the grid in a single operation (see Section 2.2.3). Through the addition of the accretion rate as a parameter the degeneracy of identical halos is reduced. The number of halos that can be treated simultaneously decreases even though they have the same mass. To mitigate this effect we implement a parallelized version of the painting step that distributes the workload to multiple processes[3]. This implementation utilizes a shared memory approach and uses processes on a single node that share a common memory space to store the grid. This allows for a more efficient usage of node resources since the memory overhead of duplicating the grid for each process is avoided. The required pre- and postprocessing that ensure the correct execution of all processes are justified by the performance gain which is nearly linear with the number of processes used[4].

Part of the painting procedure remains inherently sequential. For instance the final ionization map requires conservation of the total photon count. This is achieved by distributing duplicate ionizations from overlapping bubbles to neighboring cells. A parallel spreading approach might create new overlaps that would require further iterations to resolve. We therefore perform this step in a single process. We aim to keep these inefficient computations to a minimum.

### 4.3. Merger tree processing

The central improvement of the simulation procedure is the consideration of the individual halo mass accretion histories during the painting and not just the assumption of a predefined value. As described in Section 3 we utilize the merger trees provided by the THESAN simulation. The inference of the accretion rate is performed at runtime. Further preprocessing of the simulation is not required, only a single step that merges the individual tree files into a single file.

The generated $\alpha$ values are binned as a result of the painting procedure and the permitted range is restricted as specified in the configuration. For our runs we find that an upper limit of $\alpha = 5$ only affects a sub-percent fraction of halos. Many of these halos exhibit erratic growth suggesting that allowing for very high accretion rates is not physical.

---

[3] A rudimentary parallel implementation using the message passing interface (`MPI`) already exists. It leverages the fact that each snapshot can be processed independently and distributes the snapshots to multiple processes.

[4] We test the scaling with a parallelization up to 70 processes and observe a continuous speedup. Part of this speedup is absorbed by the overhead of the much larger number of bins.

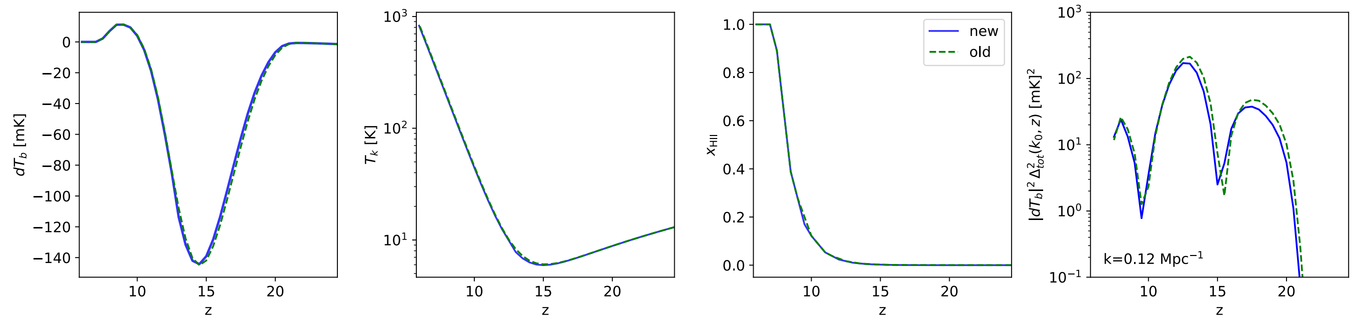


Figure 4: Validation signal comparison between the new implementation of `BEoRN` (blue) and the old version (green, dashed). From *left* to *right*: Global evolution of the differential brightness temperature $dT_{\mathrm{b}}$. Evolution of the mean kinetic temperature $T_k$. History of the mean ionization fraction $x_{\mathrm{HII}}$. 21-cm power spectrum as a function of redshift at $k = 0.12\mathrm{cMpc}^{-1}$.

The THESAN data provides a convenient way to iterate and refine the above procedure but is not without shortcomings. The merger trees are constructed in postprocessing and do not guarantee self-consistency of halo properties across multiple snapshots. This manifests itself through negative growth rates that cannot be represented in the current model. Furthermore the mass resolution of the THESAN-DARK simulations is apparently too coarse to accurately resolve halos down to the atomic cooling limit of $M_{\mathrm{h}} = 10^8 M_\odot$. This is an issue that becomes apparent in Section 5 where we compare the impact of the different mass resolutions. To account for this we follow the description of star formation efficiency employed by Schaeffer et al. (2023), picking a "boosted" model for the description of our halos. The resulting parameters for Eq. 2 are $f_{\star,0} = 0.1$, $M_p = 2.8 \times 10^{10} M_\odot$, $\gamma_1 = 0.49$ and $\gamma_2 = -0.61$.

## 4.4. Secondary changes

In addition to the changes directly linked to the new accretion model we implement several improvements that lead to better usability and reproducibility of the simulation outputs.

We improve the input/output handling by implementing proper `HDF5` support and caching of intermediate results. This allows for a more efficient usage of disk space and faster loading times. It also enables the resumption of interrupted simulations. The import of data from the original N-body simulation has been generalized to a reference class to ensure modularity and easy adaptation to other simulations. This has been part of a larger overhaul of the codebase to improve modularity and readability. `BEoRN` aims to be a flexible framework that produces fast results that the end user can customize to reflect their parameter choices. Usability is therefore a key aspect of the code design.

A general speedup from the cumulated effect of the above changes and code optimizations results in a faster painting procedure. A contribution to that speedup comes from the usage of the `pylians` package by Villaescusa-Navarro (2018). It provides efficient implementations of the grid mapping of the individual particles. This additionally allows for a rigorous implementation of redshift space distortions (RSD) by utilizing the exact velocity information of each dark matter particle individually. Previous implementations of RSD in `BEoRN` were based on approximations of the velocity field derived from the density field. The impact of RSD on the 21-cm signal has been discussed e.g. by Ross et al. (2021) but is not the focus of this work.

# 5. Validation

We perform several validation tests to ensure the accuracy and reliability of our simulation results. This section details how we verify that the new implementation remains robust and consistent with previous versions.

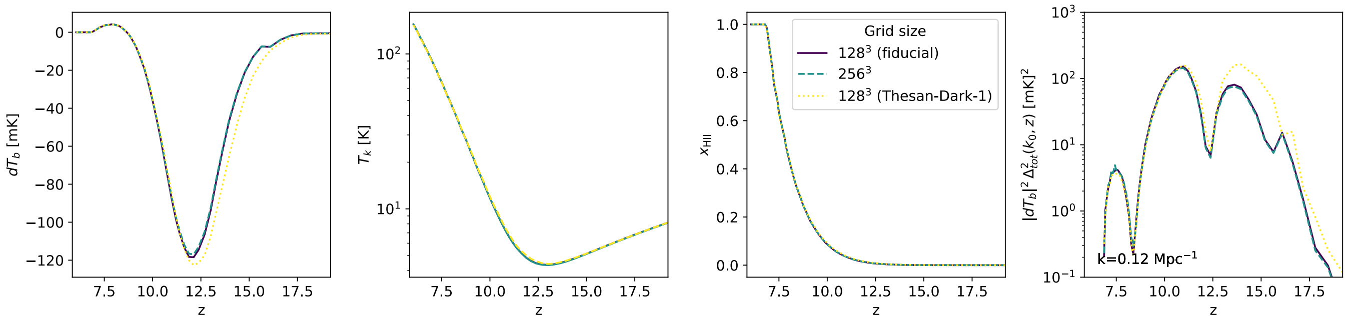


Figure 5: Validation signal comparison between runs using a grid resolution of $128^3$ (blue) and $256^3$ (green) as well as a run using the THESAN-DARK 1 simulation ($2100^3$ particles) with a grid resolution of $128^3$ (yellow). From *left* to *right*: Global evolution of the differential brightness temperature $dT_{\mathrm{b}}$. Evolution of the mean kinetic temperature $T_k$. History of the mean ionization fraction $x_{\mathrm{HII}}$. Dimensionless power spectrum of the $dT_{\mathrm{b}}$ field as a function of redshift at $k = 0.12\mathrm{cMpc}^{-1}$.

## 5.1. Code validation

We ensure the consistency of the updated BEoRN code with previous versions by running a series of simulations under identical conditions. We compare key outputs starting from the profiles of individual sources, to the ionization maps, and finally to the global reionization signals. This step-by-step comparison allows us to identify any discrepancy that may arise from the code changes. In Figure 4 we present a comparison of the central observables of reionization between the previous and current versions of BEoRN. The underlying simulation parameters are kept identical or updated to an equivalent setting where necessary. The results show an essentially unchanged evolution of the global brightness temperature, kinetic temperature, and ionization fraction. The small changes in timing are attributed to the slightly shifted mass bins of the updated procedure. Note that this run uses an artifical halo catalog and an unphysically high star formation efficiency to ensure a rapid reionization within the simulation range. As such this setup is not representative of a realistic reionization scenario.

Similarly we maintain backward compatibility with the input format used in previous BEoRN runs (i.e. snapshots generated by PKDGRAV3 or `21cmfast`). This allows us to reproduce the earlier runs and match the results as described by Schaeffer et al. (2023).

## 5.2. Convergence tests

To ensure that our results are not sensitive to the numerical resolution of the simulation, we perform convergence tests. We compare the following variations of resolution: Firstly, we scrutinize the impact of the grid resolution by running simulations with $128^3$ and $256^3$ cells. Secondly, we investigate how changing the mass resolution affects the results by comparing the outputs obtained from the Thesan-Dark 1 and 2 simulations, which have different particle masses, as mentioned in Section 2.

Figure 5 shows that there is no significant difference when using a finer grid resolution. This is expected since the ionized regions rapidly become larger than the cell size. The comparison between the different mass resolutions exhibits a deviation in the timing of reionization. This effect is expected and documented by Kannan et al. (2021): The lowest mass halos which are not resolved by THESAN-DARK 2 form small bubbles quickly (as early as $z = 10$) and contribute to the ionization budget at early times. We account for this by artificially increasing the lower mass cutoff during these validation runs but an imbalance between the two simulations remains. Additionally, the higher particle count per halo seems to enable a more consistent tree finding. The lack of good trees at the earliest snapshots is clearly visible as a dip in the $dT_b$ signal of the THESAN-DARK 2 runs. When the accretion rate cannot be inferred from the tree, a fallback value is used instead. This leads to a systematic shift in the signal that abruptly becomes apparent when the fallback value is no longer needed. After observing this effect we tune the fallback value to be in line with the expected accretion rate as shown in Figure 3. This largely removes the jump in the subsequent runs.

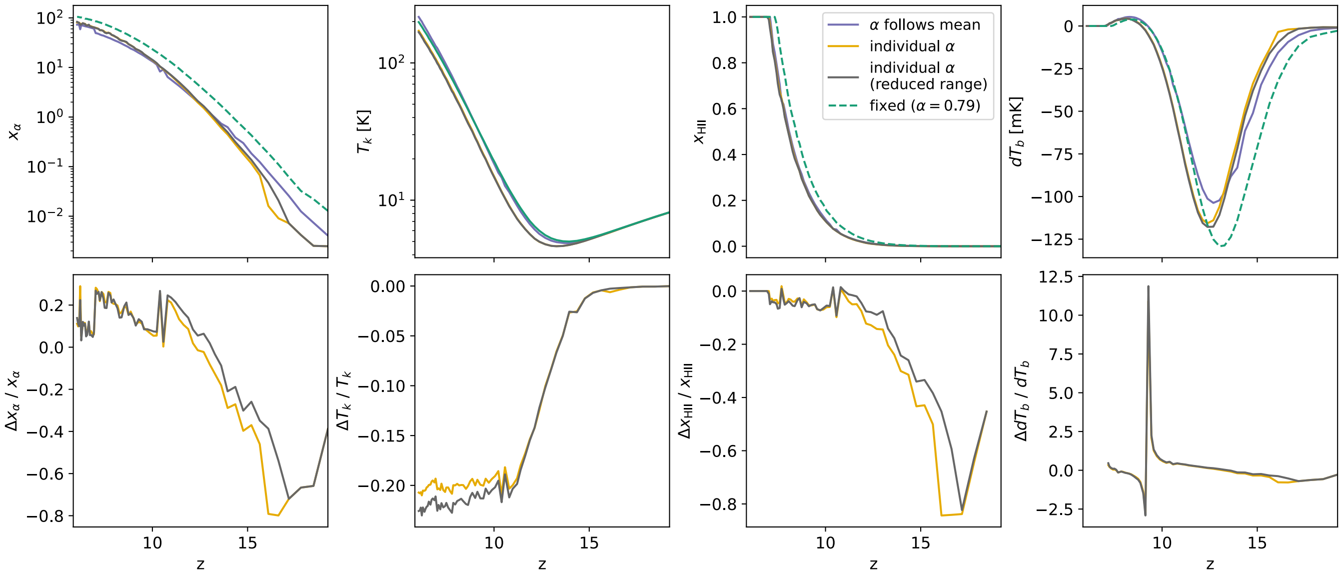


Figure 6: Signal comparison between full runs with the different accretion models: Single value of $\alpha$ for all halos according to the mean accretion rate (blue), individual accretion rates for each halo allowing a range from $\alpha = 0$ to $\alpha = 5$ (green), from $\alpha = 0$ to $\alpha = 3$ (yellow), and the previously model fixing $\alpha = 0.79$ (purple, dashed). From *left* to *right*: Evolution of the value of the coupling coefficient $x_\alpha$. Evolution of the mean kinetic temperature $T_k$. History of the mean ionization fraction $x_{\rm HII}$. Global evolution of the differential brightness temperature $dT_{\rm b}$. The bottom row shows the relative difference to the reference model, chosen to be the model following the mean accretion rate. The comparison of the original model with fixed $\alpha = 0.79$ is omitted for clarity.

Globally the shapes of the signals are very similar and, apart from the initially delayed cosmic dawn, the power spectra match closely after the discrepancy during the first few snapshots.

An additional convergence test is performed by varying the binning of the accretion rate $\alpha$ parameter. Keeping the overall range fixed between $\alpha = 0$ and $\alpha = 5.0$, we compare the signal generated when using $n = 5, 25, 50$ bins. A bulk of the halos where the accretion rate cannot be fitted is assigned a fallback value. This then falls into one of the bins, causing a global shift of the signal when changing the binning. However, the overall shape and features of the signal remain unchanged from $n = 25$ bins onwards.

## 6. Impact of individual mass history modeling

This section presents the outputs of the different simulation runs. We compare the effect of different accretion models on the global signal, map-level differences, and statistical properties of the 21-cm brightness temperature field. We focus on three different implementations:

- The fiducial model where the accretion rate is kept fixed, independently of the halo and the redshift. This corresponds to the original implementation of `BEoRN` where $\alpha = 0.79$.
- A model where the accretion rate is computed individually for each halo based on its mass growth history and is considered during the painting of each halo.
- A model where the accretion rate is computed individually for each halo but the considered value during the painting is set to the mean accretion rate of all halos at the respective redshift (effectively reducing the dynamic range of accretion rates).

### 6.1. Impact on the global signal

We first investigate the effect of the different accretion models on the global, i.e. the averaged quantities that consititute the 21-cm signal. Figure 6 shows the evolution of the coupling coefficient $x_\alpha$, the kinetic temperature $T_k$, the ionization fraction $x_{\rm HII}$, and their combined effect on the differential brightness temperature $dT_{\rm b}$. Moving away from the initial model where $\alpha = 0.79$ for all halos, we see a clear delay in the evolution of all quantities. This is expected since the accretion rates are overall lower when computed individually for each halo. The refined signals are noticeably less stable and show more fluctuations at early times. This is due to fluctuations in the tree fitting and is inherently linked to the difficulties of the THESAN simulation to resolve the earliest halos properly.

The more interesting comparison is then between the simulation using the moving mean accretion rate and the ones using the individual accretion rates. That is the difference illustrated in the bottom row of Figure 6. Using the mean accretion model as a reference, we compare the two remaining model that consider individiual mass histories over $n = 10$ snapshots. The first model allows a range of $0 \leq \alpha \leq 5.0$ and the "reduced" model allows $0 \leq \alpha \leq 3.0$.

For the individual models we see that heating is delayed by $\Delta z \approx 0.5$ and the coupling strength is initially lower but increases more rapidly at later times. This could be due to the apparition of more and more high accretion halos which contribute more to the signal while the mean remains low because of the overall increase in halo number. Halos with high accretion rates also have a significant impact the ionization fraction due to the formation of large bubbles, which explains the closely matched ionization histories.

Finally, a summary of these effects is seen in the differential brightness temperature $dT_\mathrm{b}$: The absorption trough is shifted to later times because the cosmic dawn is delayed. The delayed heating from late star formation results in a lower temperature. Even though the coupling is strong, the spin temperature remains closer to the CMB temperature, leading to a shallower absorption feature. The subsequent transition to emission is now less delayed already. In high accretion halos the star formation is increased and heating above the CMB temperature is faster. Finally the emission is shorter and drops to zero more rapidly, which is expected because the end of reionization occurs simultaneously for all models.

This comparison shows that even though the evolution of the ionization fraction is largely unaffected by our refined treatment, the global signal is nevertheless sensitive to the accretion model in ways that cannot be represented by only shifting the global accretion rate. An individual treatment of halos is the key to capture these effects.

## 6.2. Map-level investigation

Having established that the individual accretion model produces a distinct global signal, we now compare the map-level differences directly. When only considering a single fixed snapshot in time the original model and the model using the mean will create very similar maps since they use the same generalized trend. We therefore directly use a snapshot from the mean model as our reference so that the comparisons are not tainted by the timing differences to the original model.

Figure 7 shows slices through the simulation box for the different accretion models. We explicitly fix the ionization fraction of $x_\mathrm{HII} = 0.5$ which removes the effect of different timing of reionization. Thus we can focus on the spatial differences and compare the morphology of the ionized regions[5]. We omit the original model with $\alpha = 0.79$ and directly compare the two alternative accretion models. The maps resemble each other closely and we focus on the residual maps that highlight specific deviations produced when changing the accretion model. They show that fixing the mean accretion rate is not sufficient to fully represent the complex reionization behavior.

The coupling coefficient map sees an increase in all regions which is explained by the stronger emission of Lyman-$\alpha$ photons at these late stages. At the considered redshift, the mean model uses $\alpha \approx 0.51$. Regions where this reversal has not occured are in theory possible but don't seem to appear with this particular halo population.

As before, the kinetic temperature maps reflect the observation made for the signals. Most regions are colder than in the fiducial case due to the lower heating by fewer stars. The halos lag behind but some high accretion halos seem to catch up already, they practically vanish in the residual map. These differences are only visible because the mean model fails to capture this diversity of halo histories.

---

[5] Since the models compared here all have a similar ionization history, the redshifts are identical in this case.

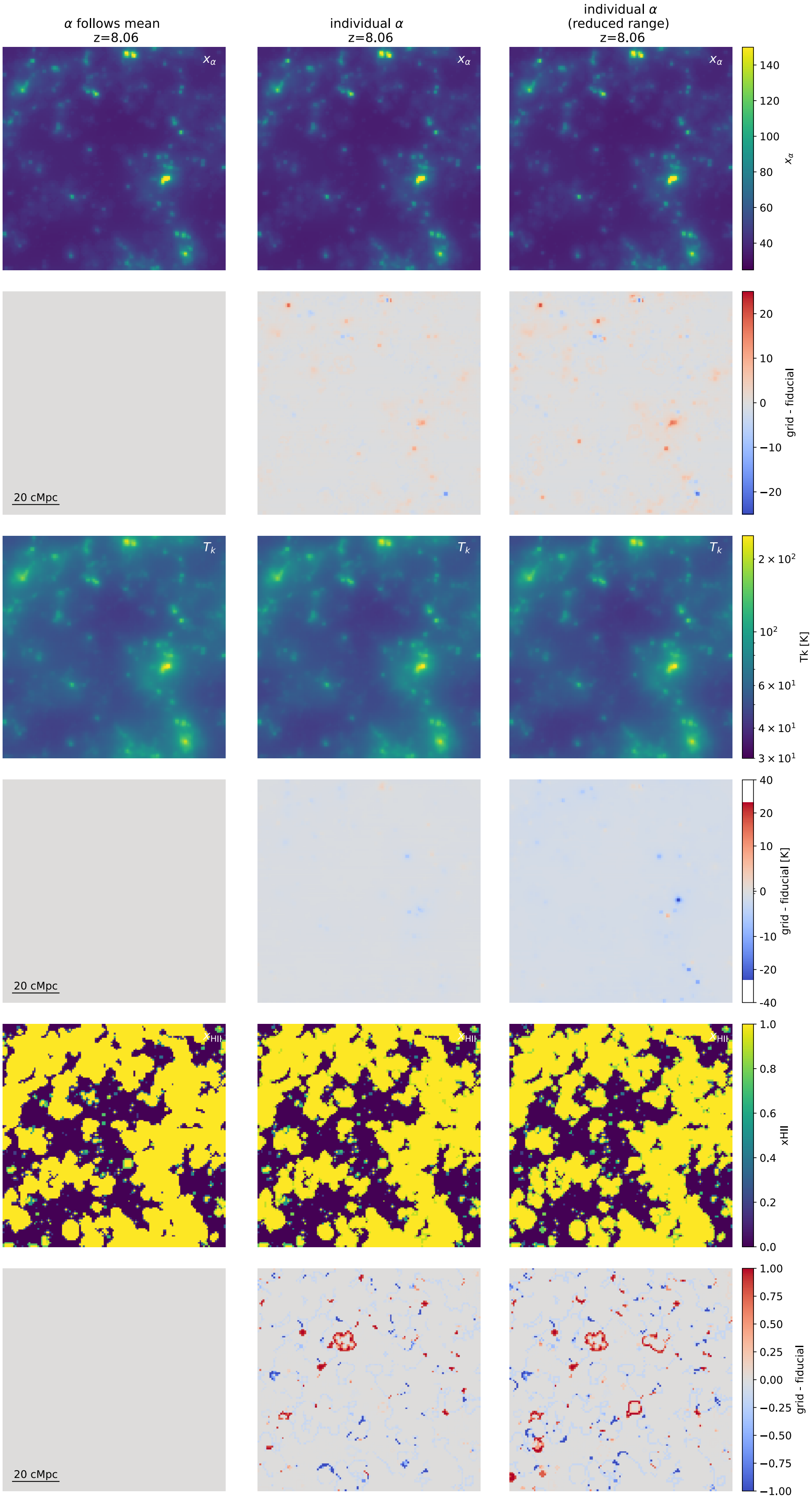


Figure 7: Map slices of the core profiles applied onto the simulation grid for the different accretion models plotted at a fixed ionization fraction of $x_{\mathrm{HII}} = 0.5$. We compare the model that uses a single value of $\alpha$ for all halos according to the mean accretion rate (left), individual accretion rates for each halo allowing a range from $\alpha = 0$ to $\alpha = 5$ (middle), and from $\alpha = 0$ to $\alpha = 3$ (right).

From *top* to *bottom*: Map of the $x_\alpha$ coupling coefficient and residual map when compared to the reference. Map of the kinetic temperature $T_k$ and residual map when compared to the reference. Map of the ionization fraction $x_{\mathrm{HII}}$ and residual map when compared to the reference. In the residual maps blue regions correspond to values lower than the reference model while red regions are higher than the reference model.

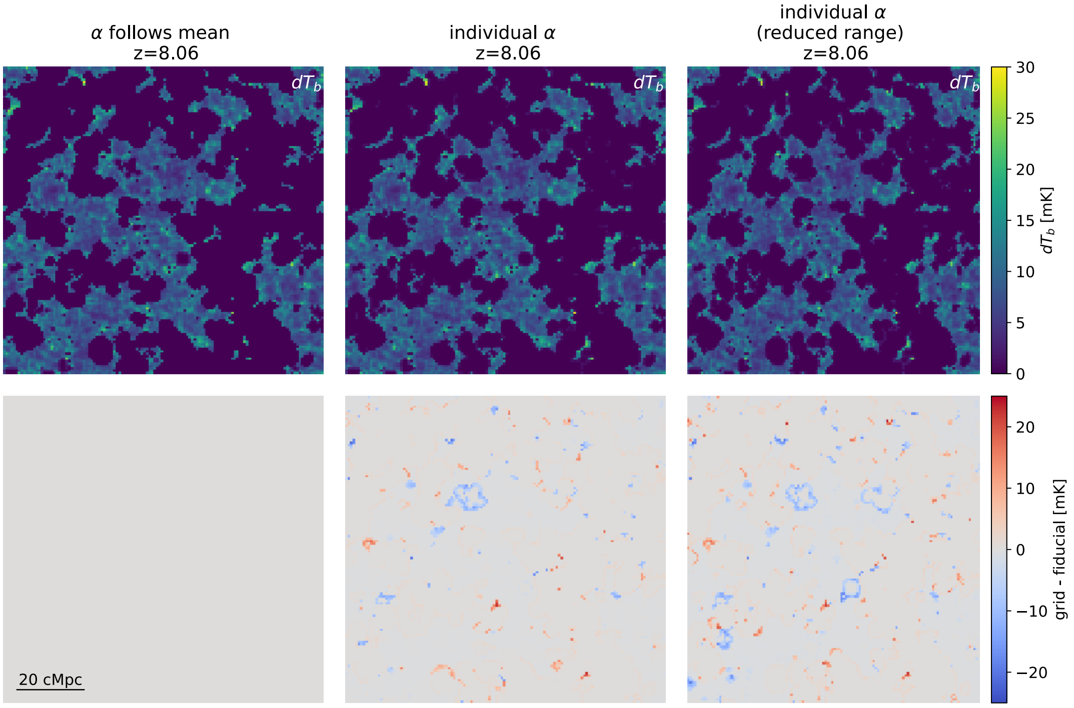


Figure 8: Map slices of the brightness temperature $dT_\mathrm{b}$ for the different accretion models plotted at a fixed ionization fraction of $x_\mathrm{HII} = 0.5$. The layout is the same as in Figure 7.

Finally, the ionization maps show the clearest differences due to the sharp bubble cutoff. There are multiple bubbles where the detailed mass accretion models generate a clear contrast, both positive and negative, compared to the mean model. The global picture remains largely unchanged, bubbles have formed and they are in the process of merging into larger contiguous structures. However, the individual bubbles show a nuanced morphology as a direct consequence of the individual halo histories. Capturing this diversity is therefore important to generate maps with the realistic range of existing structures.

So far we have treated the model with a high $\alpha$ range and the model with a lower range equally. Focusing on the differences between these two models throughout the previous maps, we see that they are mostly in agreement. This is expected since most halos are expected to have moderate accretion rates. The few very high accretion halos do however lead to small but visible differences that are most easily spotted in the ionization maps.

The derived brightness temperature maps are of particular interest. As a reminder, these are not a direct output of the simulation but the spatial distribution can be obtained from the local values of the simulated quantities via Eq. 8. They correspond to the actual observations that can be made by 21-cm surveys. We present map slices and their comparison to the mean model in Figure 8, as previously done for the individual fields.

The obtained differences are a combination of the previously discussed effects. The brightness temperature in unionized regions of the IGM remains lower due to the overall lower heating. The bubbles themselves, which appear as dark regions without signal, display a range of morphologies. Some bubbles have grown larger due to the presence of high accretion halos while others are smaller. Beyond the immediate boundary of the bubbles, the temperature seems to be affected as well, as noticeable from the faint ring-like structures around some bubbles. The detailed maps show the richness of structures that can be obtained when considering individual halo histories. The subtle differences induced by this better modeling are expected to eventually be resolved and should definitely be taken into account when interpreting future observations.

We conclude this section by commenting on the differences between the two individual accretion models. The local fluctuations are hard to quantify from the maps alone. An inspection of the resulting power spectra, which exceeds the scope of this report, reveals a small increase in power at small scales when allowing for a larger range of accretion rates. Fluctuations like these are expected to not be visible in the signal measurements but increasing the dynamic range has no detrimental effect on the performance, which is why we recommend the wider range as a default choice.

## 7. Conclusion

With the highly anticipated detection of the reionization signal and upcoming observations of the conditions of the IGM during the cosmic dawn, the interpretation of these observations requires accurate predictions from theoretical models and simulations. `BEoRN` by Schaeffer et al. (2023) is a semi-numerical simulation framework that implements the halo model of reionization (Schneider et al. 2021; 2023). It uses flux profiles to express the emission of radiation by sources in terms of their host halo to efficiently simulate the reionization on large volumes and to generate predictions for the 21-cm signal. It excels in its computational efficiency and flexibility, allowing for fast and flexible execution.

We have presented an extension to `BEoRN` that improves the physical accuracy by implementing a more consistent growth of galaxies based on the individual mass accretion histories of their host dark matter halo. We use the fact that the input data from the underlying N-body simulation already includes constraints on the growth from the halo properties at different snapshots. Disregarding this information and instead assuming a fixed accretion rate for all halos is an oversimplification. The proof-of-concept implementation presented here leverages the halo history encoded in the merger trees of the THESAN simulation. More broadly, the updated framework is now better suited to incorporate more detailed growth simulations and can be easily extended to other simulations. We also refactored the simulation procedure and achieved a faster execution time by a factor of two. This further enhances the usability of `BEoRN` for large parameter studies.

After validating the new procedure we have shown that the consistent modeling of halo growth produces simulation outputs which have distinct features compared to simpler models. We compared map outputs directly and also analyzed global quantities and their derived signal. The results are sensitive to the distribution of accretion rates, highlighting the importance of careful modeling of the halo growth.

Works going beyond this proof-of-concept implementation should utilize more sophisticated history tracking that ensures the consistency of halo properties across mutlitple timesteps (e.g. the `rockstar` halo finder by Behroozi et al. (2012)). We also highlighted the limitations incurred by the coarse mass resolution of the THESAN simulation, which is why subsequent research should be based on higher resolution simulations in order to benefit from accurate accretion rate matching down to the lowest halo masses.

Furthermore, our analysis of halo growth shows that a simple modeling relying on the halo mass alone is insufficient. Many of the radiative properties in the halo model of reionization have similarly been expressed as a function of halo mass. Given the increasing evidence against simple mass-only models such as the stellar-to-halo-mass relation, the refinement of these models using stochasticity or additional halo properties is a promising avenue for future research. The impact of revised models, including our refinement to the halo growth will be subject to a future publication.

## Acknowledgements

We would like to thank Sambit Giri and Yu-Hsiu Huang for their valuable input and helpful discussions during the development of this code. Their expertise and insights have significantly contributed to the robustness and accuracy of the simulation suite.

This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID `uzh27 - 32205` on Alps.

# Bibliography

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, The Astrophysical Journal, 935, 167, https://ui.adsabs.harvard.edu/abs/2022ApJ...935..167A

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2012, The Astrophysical Journal, 762 (American Astronomical Society), 109, http://dx.doi.org/10.1088/0004-637X/762/2/109

DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, Publications of the Astronomical Society of the Pacific, 129 (IOP Publishing), 45001, http://dx.doi.org/10.1088/1538-3873/129/974/045001

Dekel, A., Zolotov, A., Tweed, D., et al. 2013, Monthly Notices of the Royal Astronomical Society, 435, 999, https://doi.org/10.1093/mnras/stt1338

Field, G. B. 1958, Proceedings of the IRE, 46, 240

Garaldi, E., Kannan, R., Smith, A., et al. 2022, Monthly Notices of the Royal Astronomical Society, 512 (Oxford University Press (OUP), 4909, http://dx.doi.org/10.1093/mnras/stac257

Iliev, I. T., Mellema, G., Pen, U.-L., et al. 2006, Monthly Notices of the Royal Astronomical Society, 369, 1625, https://doi.org/10.1111/j.1365-2966.2006.10502.x

Iliev, I. T., Mellema, G., Ahn, K., et al. 2014, Monthly Notices of the Royal Astronomical Society, 439 (Oxford University Press (OUP), 725, http://dx.doi.org/10.1093/mnras/stt2497

Kannan, R., Garaldi, E., Smith, A., et al. 2021, Monthly Notices of the Royal Astronomical Society, 511 (Oxford University Press (OUP), 4005, http://dx.doi.org/10.1093/mnras/stab3710

Koopmans, L., Pritchard, J., Mellema, G., et al. 2015, in Proceedings of Advancing Astrophysics with the Square Kilometre Array — PoS(AASKA14) (Sissa Medialab), http://dx.doi.org/10.22323/1.215.0001

Mao, Y., Tegmark, M., McQuinn, M., Zaldarriaga, M., & Zahn, O. 2008, Physical Review D, 78 (American Physical Society (APS), http://dx.doi.org/10.1103/PhysRevD.78.023529

McBride, J., Fakhouri, O., & Ma, C.-P. 2009, Monthly Notices of the Royal Astronomical Society, 398 (Oxford University Press (OUP), 1858, http://dx.doi.org/10.1111/j.1365-2966.2009.15329.x

McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006, The Astrophysical Journal, 653 (American Astronomical Society), 815, http://dx.doi.org/10.1086/505167

Mebane, R. H., Mirocha, J., & Furlanetto, S. R. 2020, Monthly Notices of the Royal Astronomical Society, 493 (Oxford University Press (OUP), 1217, http://dx.doi.org/10.1093/mnras/staa280

Mellema, G., Iliev, I. T., Alvarez, M. A., & Shapiro, P. R. 2006, New Astronomy, 11, 374, https://www.sciencedirect.com/science/article/pii/S1384107605001405

Mesinger, A., Furlanetto, S., & Cen, R. 2011, Monthly Notices of the Royal Astronomical Society, 411, 955, https://doi.org/10.1111/j.1365-2966.2010.17731.x

Potter, D., Stadel, J., & Teyssier, R. 2016, https://arxiv.org/abs/1609.08621

Pritchard, J. R., & Loeb, A. 2012, Reports on Progress in Physics, 75, 86901

Ross, H. E., Giri, S. K., Mellema, G., et al. 2021, Monthly Notices of the Royal Astronomical Society, 506 (Oxford University Press (OUP), 3717, http://dx.doi.org/10.1093/mnras/stab1822

Schaeffer, T., Giri, S. K., & Schneider, A. 2023, Monthly Notices of the Royal Astronomical Society, 526 (Oxford University Press (OUP), 2942, http://dx.doi.org/10.1093/mnras/stad2937

Schneider, A., Giri, S. K., & Mirocha, J. 2021, Physical Review D, 103 (American Physical Society (APS), http://dx.doi.org/10.1103/PhysRevD.103.083025

Schneider, A., Schaeffer, T., & Giri, S. K. 2023, https://arxiv.org/abs/2302.06626

Smith, A., Kannan, R., Garaldi, E., et al. 2022, Monthly Notices of the Royal Astronomical Society, 512 (Oxford University Press (OUP), 3243, http://dx.doi.org/10.1093/mnras/stac713

Springel, V., White, S. D., Jenkins, A., et al. 2005, Nature, 435, 629, https://ui.adsabs.harvard.edu/abs/2005Natur.435..629S

Villaescusa-Navarro, F. 2018, ascl:1811.008, https://ui.adsabs.harvard.edu/abs/2018ascl.soft11008V

Wouthuysen, S. 1952, AJ, 57, 31, https://ui.adsabs.harvard.edu/abs/1952AJ.....57R..31W

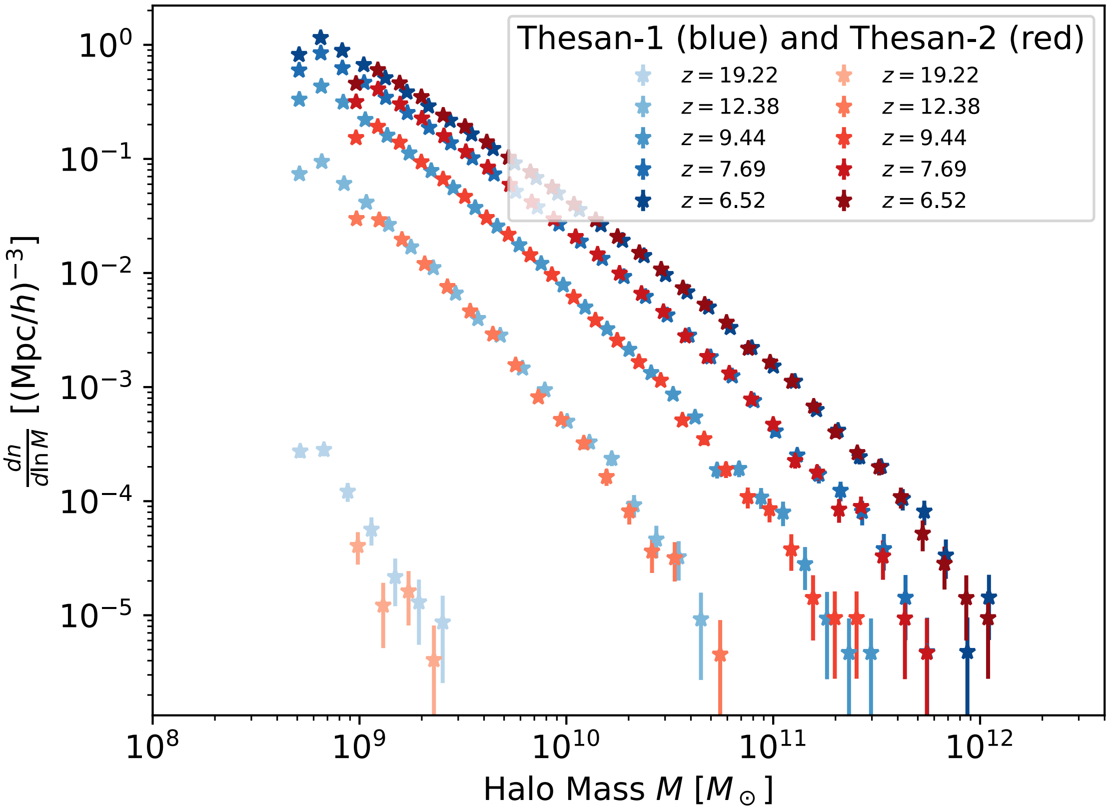


Figure 9: The halo mass functions of the THESAN-DARK 1 (blue) and THESAN-DARK 2 (red) plotted at different redshifts. The error bars indicate the Poisson error in each mass bin.

# Appendix

## A - Halo mass function of THESAN-DARK 1 and THESAN-DARK 2

We compare the halo mass functions of the two dark-matter-only simulations THESAN-DARK 1 and THESAN-DARK 2. As shown in Figure 9, the higher resolution of the former allows to resolve smaller halos more accurately. The two simulations only agree above a mass of about $5 \cdot 10^9 M_\odot$ and then yield the same halo mass function. This is consistent with our observations in the validation (Section 5) of both simulations and poses a limitation on the minimum halo mass that can be resolved in BEoRN when using merger trees from THESAN-DARK 2.

## B - Generation of the cover image

The cover image of this report has been generated using BEoRN. From a simulation run with a grid resolution of $512^3$ cells, a slice of the brightness temperature map has been extracted at $z = 8.07$. The slice shows an emission due to the spin temperature being higher than the CMB temperature. The first ionization bubbles appear as dark patches where the ionized hydrogen does not contribute to the 21-cm signal.