



The short ionizing photon mean free path at $z=6$ in Cosmic Dawn III, a new fully-coupled radiation-hydrodynamical simulation of the Epoch of Reionization

Joseph S. W. Lewis, Pierre Ocvirk, Jenny G. Sorce, Yohan Dubois, Dominique Aubert, Luke Conaboy, Paul R. Shapiro, Taha Dawoodbhoy, Romain Teyssier, Gustavo Yepes, et al.

► To cite this version:

Joseph S. W. Lewis, Pierre Ocvirk, Jenny G. Sorce, Yohan Dubois, Dominique Aubert, et al.. The short ionizing photon mean free path at $z=6$ in Cosmic Dawn III, a new fully-coupled radiation-hydrodynamical simulation of the Epoch of Reionization. Monthly Notices of the Royal Astronomical Society, 2022, 10.1093/mnras/stac2383 . insu-03763116

HAL Id: insu-03763116

<https://insu.hal.science/insu-03763116>

Submitted on 11 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The short ionizing photon mean free path at $z = 6$ in Cosmic Dawn III, a new fully coupled radiation-hydrodynamical simulation of the Epoch of Reionization

Joseph S. W. Lewis^{1,2,★}, Pierre Ocvirk³, Jenny G. Sorce^{4,5}, Yohan Dubois⁶, Dominique Aubert³, Luke Conaboy⁷, Paul R. Shapiro⁸, Taha Dawoodbhoy⁸, Romain Teyssier⁹, Gustavo Yepes^{10,11}, Stefan Gottlöber⁴, Yann Rasera^{2,12}, Kyungjin Ahn¹³, Ilian T. Iliev¹⁴, Hyunbae Park^{14,15,16} and Émilie Thélie³

¹*Institut für Theoretische Astrophysik, Zentrum für Astronomie der Universität Heidelberg, Albert-Ueberle-Straße 2, D-69120 Heidelberg, Germany*

²*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*

³*Observatoire Astronomique de Strasbourg, Université de Strasbourg, CNRS UMR 7550, 11 Rue de l'Université, F-67000 Strasbourg, France*

⁴*Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*

⁵*Université Paris-Saclay, CNRS, Institut d'Astrophysique Spatiale, F-91405 Orsay, France*

⁶*Institut d'Astrophysique de Paris, UMR 7095, CNRS, UPMC Univ. Paris VI, 98 bis Boulevard Arago, F-75014 Paris, France*

⁷*Astronomy Center, Department of Physics and Astronomy, University of Sussex, Pevensey II Building, Falmer, Brighton BN1 9QH, UK*

⁸*Department of Astronomy, University Texas, Austin, TX 78712-1083, USA*

⁹*Institute for Theoretical Physics, University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

¹⁰*Departamento de Física Teórica M-8, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*

¹¹*Centro de Investigación Avanzada en Física Fundamental (CIAFF), Universidad Autónoma de Madrid, E-28049 Madrid, Spain*

¹²*Laboratoire Univers et Théories, Université de Paris, Observatoire de Paris, Université PSL, CNRS, F-92190 Meudon, France*

¹³*Chosun University, 375 Seosuk-Dong, Dong-Gu, Gwangju 501-759, Korea*

¹⁴*Lawrence Berkeley National Laboratory, Berkeley, CA 94720-8139, USA*

¹⁵*Berkeley Center for Cosmological Physics, UC Berkeley, Berkeley, CA 94720, USA*

¹⁶*Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*

Accepted 2022 August 8. Received 2022 August 5; in original form 2022 February 11

ABSTRACT

Recent determinations of the mean free path of ionizing photons (mfp) in the intergalactic medium (IGM) at $z = 6$ are lower than many theoretical predictions. In order to gain insight, we investigate the evolution of the mfp in our new massive fully coupled radiation-hydrodynamics cosmological simulation of reionization: Cosmic Dawn III (CoDa III). CoDa III's scale (94^3 cMpc^3) and resolution (8192^3 grid) make it particularly suitable to study the IGM during reionization. The simulation was performed with RAMSES-CUDATON on Summit, and used 13 1072 processors coupled to 24 576 GPUs, making it the largest reionization simulation, and largest ever RAMSES simulation. A superior agreement with global constraints on reionization is obtained in CoDa III over Cosmic Dawn II (CoDa II), especially for the evolution of the neutral hydrogen fraction and the cosmic photoionization rate, thanks to an improved calibration, later end of reionization ($z = 5.6$), and higher spatial resolution. Analysing the mfp, we find that CoDa III reproduces the most recent observations very well, from $z = 6$ to $z = 4.6$. We show that the distribution of the mfp in CoDa III is bimodal, with short (neutral) and long (ionized) mfp modes, due to the patchiness of reionization and the coexistence of neutral versus ionized regions during reionization. The neutral mode peaks at sub-kpc to kpc scales of mfp, while the ionized mode peak evolves from $0.1 \text{ Mpc } h^{-1}$ at $z = 7$ to $\sim 10 \text{ Mpc } h^{-1}$ at $z = 5.2$. Computing the mfp as the average of the ionized mode provides the best match to the recent observational determinations. The distribution reduces to a single neutral (ionized) mode at $z > 13$ ($z < 5$).

Key words: galaxies: high-redshift – dark ages, reionization, first stars – intergalactic medium – large-scale structure of Universe.

1 INTRODUCTION

The Epoch of Reionization (EoR) began as the first ionizing galactic sources started to ionize their immediate surroundings in the intergalactic medium (IGM) some few hundred million years after the big bang. This epoch lasted approximately 600 Myr, until the reionization

* E-mail: Joe.Lewis@uni-heidelberg.de

of the IGM was complete at some point between $z \approx 6$ and $z \approx 5.3$ (Bosman et al. 2022, hereafter [Bo22](#)). Constraining this distant epoch offers unique insight into the formation of the first galaxies and stars, and is an important step in bridging the gap between our knowledge of ‘modern’ galaxies, and the first moments after the big bang. Indeed, the rising and expanding ultraviolet (UV) background during the EoR may suppress star formation in low-mass galaxies (Ocvirk et al. 2016, 2020, hereafter [Oc16](#) and [Oc20](#), respectively; Dawoodbhoy et al. 2018). This in turn motivated efforts to investigate and predict possible signatures of the EoR in populations of galaxies of the lowest mass, which are best observed in the Local Group (Iliev et al. 2011; Ocvirk & Aubert 2011; Ocvirk et al. 2013, 2014; Dixon et al. 2018).

Because of the very nature of the EoR and the extreme distances involved, direct observations of the reionizing sources as they drive reionization are extremely challenging. Despite this, over the past decades innovative indirect constraints on both the sources of reionization and on the IGM have emerged. Further, the future is bright as upcoming space- and ground-based observatories (e.g. the *James Webb Space Telescope*, the *Nancy Grace Roman Space Telescope*, or the Extremely Large Telescope) will improve our capability to target the drivers of reionization. At the same time, current and planned 21-cm radio observations [e.g. Experiment to Detect the Global EoR Signature (EDGES), Hydrogen Epoch of Reionization Array (HERA), LOw-Frequency ARray (LOFAR), NenuFAR, and Square Kilometre Array (SKA)] may bring new constraints on the ionization of the IGM during the EoR and at Cosmic Dawn. Thus, the study of the EoR appears to be set for rapid progress over the coming years. In preparation of this new era of EoR observations, the simulation community is hard at work making predictions for these new observational campaigns, and investigating the physical processes that drive the sources of reionization. However, the predictive power of these simulations relies on the quality of their agreement with existing observational constraints, some of which have proved difficult to match.

For instance, recent determinations of the mean free path of ionizing photons (mfp), at a higher redshift than ever before, have shown that models and simulations of reionization overestimate the mfp when $z \approx 6$ (and possibly when $5.24 \lesssim z \lesssim 6$; see Rahmati & Schaye 2018; D’Aloisio et al. 2020; Keating et al. 2020), casting doubt on the realism of their portrayal of the ionization state of the IGM and thus cosmic reionization as a whole (see Becker et al. 2021, hereafter [Be21](#)). Already, new work has investigated the implications of these new data, finding that they require late and rapid reionization scenarios (Cain et al. 2021, hereafter [Ca21](#); Davies et al. 2021, hereafter [Da21](#)), which are also favoured by constraints on the Lyman α forest (Kulkarni et al. 2019; [Bo22](#)). In fact, [Ca21](#) have been able to obtain a reasonable agreement with both the new and old mfp constraints. We consider below these two recent works attempting to match the new very low mfp measurement of [Be21](#) and show how using Cosmic Dawn III (CoDa III)¹ may improve upon them. Both studies above use a fairly low-resolution simulation of the IGM (~ 1 cMpc cell size), and must recourse to complex subgrid models to the following.

(i) *Model the source population.* [Da21](#)’s description of sources involves a range of parameters that fold in a number of complex and coupled aspects, such as the sensitivity of sources to supernova and active galactic nucleus (AGN) feedback, their star formation

efficiency, their escape fractions, and the global dependence of these processes with respect to mass. However, as we show in Ocvirk et al. (2021, hereafter [Oc21](#)), radiative suppression of star formation in low-mass haloes is crucial to obtaining a realistic Lyman α forest after overlap. This regulation, for instance, is not included in [Da21](#), and CoDa III improves on their approach by allowing this process to occur self-consistently thanks to the fully coupled radiation-hydrodynamics (RHD) framework employed.

(ii) *Set the transmission properties of the IGM subgrid.* [Ca21](#) use a subgrid parametrization of the mean free path, which relies on simulations of irradiated patches (D’Aloisio et al. 2020). Such simulations do not include star formation, and one can therefore estimate that their densest patch structures would at some point form stars, and therefore both decrease their own mean free path and the mean free path around them.

Our approach circumvents the need for using such a subgrid model of the mean free path, as the whole volume of CoDa III is eligible to star formation provided the relevant criteria in density and temperature are met. Moreover, the mfp is self-consistently set in CoDa III by the physical state of the IGM, and is self-consistently affected by radiation received by sources near and far and the high-resolution hydrodynamical description of the IGM.

Beyond these two studies, it is desirable, in all cases, to deploy a fully coupled RHD formalism as we do here, in order to break free from the limitations imposed by the parametrization of the physical processes at hand as in [Da21](#), [Ca21](#), and check that the results they claim can be confirmed in more self-consistent simulations where more physics are resolved and therefore do not require such a level of subgrid modelling. In this direction, the THESAN project (Garaldi et al. 2022; Kannan et al. 2022) has recently performed a series of AREPO-RT (Kannan et al. 2019) simulations of the EoR. While they reach high spatial resolution inside of galaxies, CoDa III improves on their approach in at least two aspects: their fiducial run, THESAN-1, stops at $z = 5.5$, whereas we were able to go beyond this redshift, down to $z = 4.6$, and as a result compare CoDa III mfp with observations in a larger redshift range, which, most importantly, includes post-reionization measurements as well. Also, THESAN-1 uses a typical speed of light reduction of 0.2, which can have a strong impact on the post-reionization residual neutral hydrogen fraction (Ocvirk et al. 2019), and to some extent change the timing of reionization (Deparis et al. 2019), complicating the interpretation of their simulation results. To avoid such difficulties, CoDa III uses the full speed of light thanks to the GPU-optimized radiative transfer module CUDATON ([Oc16](#); [Oc20](#)).

First, we present our methodology, including the particularities of our simulation code RAMSES-CUDATON, the simulation set-up, and the computational steps we take to derive the mfp. Then, we present the global properties of reionization in the CoDa III simulation, and the evolution of the mfp in CoDa III. Finally, we conclude our work.

2 METHODS

2.1 Simulation

2.1.1 Overview of RAMSES-CUDATON

The CoDa III simulation was performed using the fully coupled radiation and hydrodynamics code RAMSES-CUDATON (presented in [Oc16](#); [Oc20](#)). Here we briefly recall its main properties and details concerning its deployment.

¹More information can be found about the whole simulation suite and the CoDa project here: <https://codasimulation.github.io/>

RAMSES-CUDATON results from the coupling of the RAMSES code for N -body dark matter dynamics, hydrodynamics, star formation, and stellar feedback (Teyssier 2002), with the radiative transfer module ATON (Aubert & Teyssier 2008), which handles the radiative transfer of ionizing photons via the M1 method (Levermore 1984), hydrogen photoionization, and thermochemistry. Motivated by the important computational cost of radiative transfer, we accelerate the ATON module using GPUs (Aubert & Teyssier 2010), whilst the remaining RAMSES physics run on CPUs. The performance boost thus obtained allows us to run fully coupled simulations with a full speed of light, circumventing pitfalls encountered by certain reduced speed of light approaches (Gnedin 2016; Deparis et al. 2019), in particular on the residual neutral fraction after overlap (Ocvirk et al. 2019).

2.1.2 Improvements over our previous simulation CoDa II

For the most part, the code and simulation set-up are as described in Oc20, but with higher resolution. Other features of CoDa III that differ from past CoDa simulations were fully tested and calibrated by a large suite of smaller simulations of the same numerical resolution as CoDa III to ensure a good match to observational constraints on the EoR. While the chief purpose of these calibration runs was to prepare for CoDa III, the calibration runs themselves were novel enough to warrant new science as described in Oc21. Since the new features of CoDa III were already described there, we refer the reader to that paper for more detail. We quickly enumerate the novelties of CoDa III with respect to Cosmic Dawn II (CoDa II) hereafter.

(i) CoDa III uses an 8192^3 grid while CoDa II had a 4096^3 grid, for the same volume, meaning its physical resolution is a factor of 2 finer, and its mass resolution is eight times higher.

(ii) We take a more detailed approach with respect to our source model than in the previous CoDa simulations. In CoDa III, each stellar particle receives an ionizing emissivity value based on its mass, age, and metallicity. We pre-compute emissivities using BPASS v2.2.1 (Stanway & Eldridge 2018), considering that each stellar particle is a star cluster of about $10^4 M_\odot$ forming over a few Myr as in Oc21. As a result, in CoDa III individual stellar particles' ionizing luminosities decrease more gradually and realistically than in CoDa II.²

(iii) One of the main model differences with respect to CoDa II lies within the star formation subgrid model. In CoDa III, for a gas cell to form stars, we not only require that it be dense enough, but also that it be cooler than 2×10^4 K. This additional temperature threshold for star formation has been widely used in similar simulations before CoDa III including CoDa I (Oc16) and others (e.g. Dubois & Teyssier 2008). Its physical interpretation is as follows: gas cells hotter than the threshold are so because of shocks, usually caused by supernovae, which are not favoured sites for star formation until they manage to cool back down to below the 2×10^4 K threshold.

We also found that including the temperature threshold allowed our simulations to achieve better agreement with the residual neutral fraction at the end of the EoR and with the ionization rate of ionized regions, thus advocating its use in CoDa III. For more information and a comparison between star formation with, and without the temperature threshold, we refer the reader to Oc21.

(iv) We use a subgrid stellar escape fraction of $f_{\text{esc}}^{\text{sub}} = 1$, whereas we used $f_{\text{esc}}^{\text{sub}} = 0.42$ in CoDa II, i.e. all photons produced by the stellar particles in CoDa III are released in the host cell. We shall fully investigate the escape of ionizing photons in an upcoming paper, however we can already tentatively explain why we require a higher $f_{\text{esc}}^{\text{sub}}$ in CoDa III. First, and due to the changes to the star formation model, star formation in CoDa III is not permitted in very hot ionized cells. For the same amount of star formation it is then simple to imagine that one requires more photons to complete reionization in CoDa III with respect to CoDa II. Second, CoDa III's higher resolution likely gives rises to more and denser clumps of gas in the IGM, potentially further raising the required number of photons to complete reionization. Third, the ionizing emissivity of stellar particles evolves in a smoother fashion in CoDa III, plausibly altering the balance between ionization and recombination (although we found this effect to be relatively moderate in smaller test volumes).

(v) Just as in CoDa I and CoDa II (the two previous Cosmic Dawn simulation projects), upon reaching 10 Myr of age, a mass fraction $\eta_{\text{SN}} = 0.2$ of stellar particles detonates as supernovae. Each supernova event injects 51 erg of energy for every $10 M_\odot$ of progenitor mass into its host cell (using the kinetic feedback of Dubois & Teyssier 2008). After the supernova event, a long-lived particle of mass $(1 - \eta_{\text{SN}})M_{\text{birth}}$ remains. Unlike CoDa I and CoDa II, CoDa III features the standard RAMSES chemical enrichment.³ Metals are produced by supernovae. A fraction $y = 0.05$ of the total supernova mass is reinjected as metals. Metallicity is treated as a passive scalar, and metals are thus advected along with the gas.

(vi) CoDa III features a physically motivated model for the formation and destruction of dust in galaxies on the fly, which we coupled to the ionizing radiative transfer of ATON (see Lewis et al. 2022, for details). An independent paper on the dust model is also in preparation (Dubois et al., in preparation). The dust aspects of CoDa III will be analysed in a forthcoming paper, and can be left aside for this study.

For quick reference, Table 1 recaps the essential physical and numerical parameters of CoDa III.

2.1.3 Initial conditions and cosmology

As in previous CoDa simulations, CoDa III's initial conditions were constrained so as to produce facsimiles of the progenitors of the Local Group (see Sorce et al. 2016; Sorce & Tempel 2018) within the Constrained Local Universe Simulations (CLUES)⁴ project. We will use this unique opportunity to investigate the reionization of the Local Group in a forthcoming paper. For this study and for the purpose of this paper, we can forgo the Local Universe aspect and treat CoDa III as a random realization of the Universe, as its mass function, for instance, is compatible with that of a random patch of Universe. The cosmology used ($\Omega_\Lambda = 0.693$, $\Omega_m = 0.307$, $\Omega_b = 0.0482$, $H_0 = 67.77 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\sigma_8 = 0.829$, and $n = 0.963$) here is taken from the Planck Collaboration XVI (2014) constraints (Planck14),⁵ as summarized in Table 1. The initial conditions were generated at $z = 150$.

³For clarification: this is not a new code feature, but was not activated at run time in the previous simulations, for simplicity.

⁴<https://www.clues-project.org/cms/>

⁵These constraints are compatible with subsequent *Planck* publications.

²In the previous CoDa simulations, stellar particles had a fixed stellar luminosity until their massive stars went supernova, at which point their ionizing luminosities were set to 0.

Table 1. Overview of the CoDa III simulation.

Set-up	
Grid size and dark matter particle number	8192 ³
Box size	94.43 cMpc
Force resolution	
Comoving	11.53 ckpc
Physical ($z = 6$)	1.65 pkpc
Dark matter particle mass	$5.09 \times 10^4 M_\odot$
Average cell gas mass	$9.375 \times 10^3 M_\odot$
Stellar particle mass	11 732 $M_\odot$
Star formation and feedback	
Density threshold for star formation	$50(\rho_{\text{gas}})$
Temperature threshold for star formation	$2 \times 10^4 \text{ K}$
Star formation efficiency, ϵ_*	0.03
Massive star lifetime	10 Myr
Supernova energy	10^{51} erg
Supernova mass fraction, η_{SN}	0.2
Supernova ejecta metal mass fraction	0.05
Radiation	
Stellar ionizing emissivity model	BPASS v2.2.1 binary
(from Stanway & Eldridge 2018)	
Stellar particle subgrid escape fraction, $f_{\text{esc}}^{\text{sub}}$	1.0
Effective photon energy	20.28 eV
Effective H I cross-section (at 20.28 eV)	$2.493 \times 10^{-22} \text{ m}^2$
Reionization	
CMB electron scattering optical depth, $\tau_{\text{es}}(z = 14)$	0.0497
Reionization mid-point, $z_{\text{re}}, 50 \text{ per cent}$	6.81
Reionization complete, $z_{\text{re}}, 99.99 \text{ per cent}$	5.53

2.1.4 Deployment

CoDa III was run on Summit at the Oak Ridge National Laboratory/Oak Ridge Leadership Computing Facility. It was deployed on 13 1072 processors coupled with 24 576 GPUs, distributed on 4096 nodes, and ran for ~ 10 d, allowing CoDa III to reach a final redshift of $z = 4.6$, i.e. sufficiently late after overlap that we can consider reionization complete and sample well into the post-overlap phase. Snapshots were written every ~ 10 Myr, we use a subset of them to derive the properties of the IGM in the following section. Subsequent processing of the more than 20 PB of data outputs took place at the same facility on Andes and Summit.

2.2 Computing the mean free path of ionizing photons

To compute the mean free path of ionizing photons (mfp), we employ four approaches.

(i) First, we employ a method (called average transmission method from hereon) inspired by Kulkarni et al. (2016). For a given simulation snapshot, we draw 1024 lines of sight (LoS) across the whole simulation box in one direction. Our LoS are evenly separated so as to evenly sample the box.

For each LoS, we then compute the transmitted flux at $912 \text{ \AA}$ as a function of position along the LoS x . These can then be averaged into one average transmitted flux $F(x)$, from which we obtain λ_{AT} by fitting it with equation (1) for the parameter λ :

$$F(x) = F_0 \exp\left(-\frac{x}{\lambda}\right), \quad (1)$$

where F_0 is a normalization parameter, equivalent to a flux incident where $x = 0$, i.e. free during the fit. For key redshifts, we check that our results are converged with respect to the number of LoS, and find

no significant difference in our result with $8\times$ or $64\times$ more LoS. In order to mimic the small number statistics regime of the observations, we draw samples of only 13 LoS (the same LoS sample size as in Be21, at $z = 6$) from our full sample of 1024. We repeat this 10^5 times, and our final λ_{AT} result is then the median of these 10^5 realizations. Also of interest, this approach yields the standard deviation of this set of mfp.

(ii) Second, in order to access the distribution of mfp belonging to individual LoS (λ_{los}), we fit equation (1) to the transmitted flux of every LoS individually. We can then also compute the mean, median, and spread of the resulting distribution at every redshift. The results from bad fits (reduced $\chi^2 > 1.0$) are discarded. This occurs only five times for the 1024×14 LoS that we process, and correspond to LoS that cross a large overdensity. In Appendix A, we compare this method to that of Rahmati & Schaye (2018) (also found in Garaldi et al. 2022), and show that both methods for computing the mean free path along individual LoS are remarkably consistent.

(iii) Third, as we will see in Section 3.2, our λ_{los} are divided into two families that reside in neutral, and ionized areas of the IGM. We can easily isolate the evolution of the λ_{los} from the ionized regions so as to more closely mimic the observational context of the measurement of mfp, since such measurements cannot be performed in still neutral medium. To this end, for every redshift we divide the distribution of λ_{los} in two modes, at the approximate position of the minimum of the distributions and therefore the transition between the modes, and consider only the long mode of the distribution, corresponding to ionized LoS. We refer to this approach as the ionized method. We use the notation λ_i to refer to mfps computed in this manner.

3 RESULTS

3.1 Global properties of the IGM during the EoR

The top left-hand panel of Fig. 1 shows the evolution of the average volume weighted neutral fraction of hydrogen ($\langle x_{\text{H I}} \rangle$) in CoDa III and CoDa II, along with a range of constraints. In both simulations, $\langle x_{\text{H I}} \rangle$ starts to drop noticeably between $z = 10$ and $z = 9$. However, in CoDa III, reionization ends later and less abruptly than in CoDa II, only finishing by $z \approx 5.6$ (≈ 6 in CoDa II). Further, in CoDa II the final average value of the neutral fraction is far lower than in CoDa III. This latter difference is due to the combination of the new star formation subgrid model and improved resolution, as explained in Oc21. Considering this, along with the rather late reionization obtained (compared to the canonical value of $z = 6$ for the end of reionization in use since Fan et al. 2006), CoDa III is in much better agreement with the constraints from the transmission in the Lyman α forest (and modelling) on $\langle x_{\text{H I}} \rangle$ after reionization. In particular, the agreement with the data from Bo22 is excellent. The match with constraints from Becker et al. (2015) is also remarkable in CoDa III after the ankle of the $\langle x_{\text{H I}} \rangle$ curve, although Becker et al. (2015) favour slightly earlier reionization than in CoDa III. Finally, we find that the older constraints from Fan et al. (2006) are disfavoured both by the more recent determinations of Bo22 and Becker et al. (2015), and our own simulations.

The bottom left-hand panel of Fig. 1 shows the evolution of the average ionized fraction of hydrogen ($x_{\text{H II}}$) on a linear scale, along with the same constraints. CoDa III is in better agreement with a greater fraction of the collection of constraints for $x_{\text{H II}} \gtrsim 0.1$ than CoDa II. Overall the agreement of CoDa III with modern constraints on the evolution of the progress of reionization is very good.

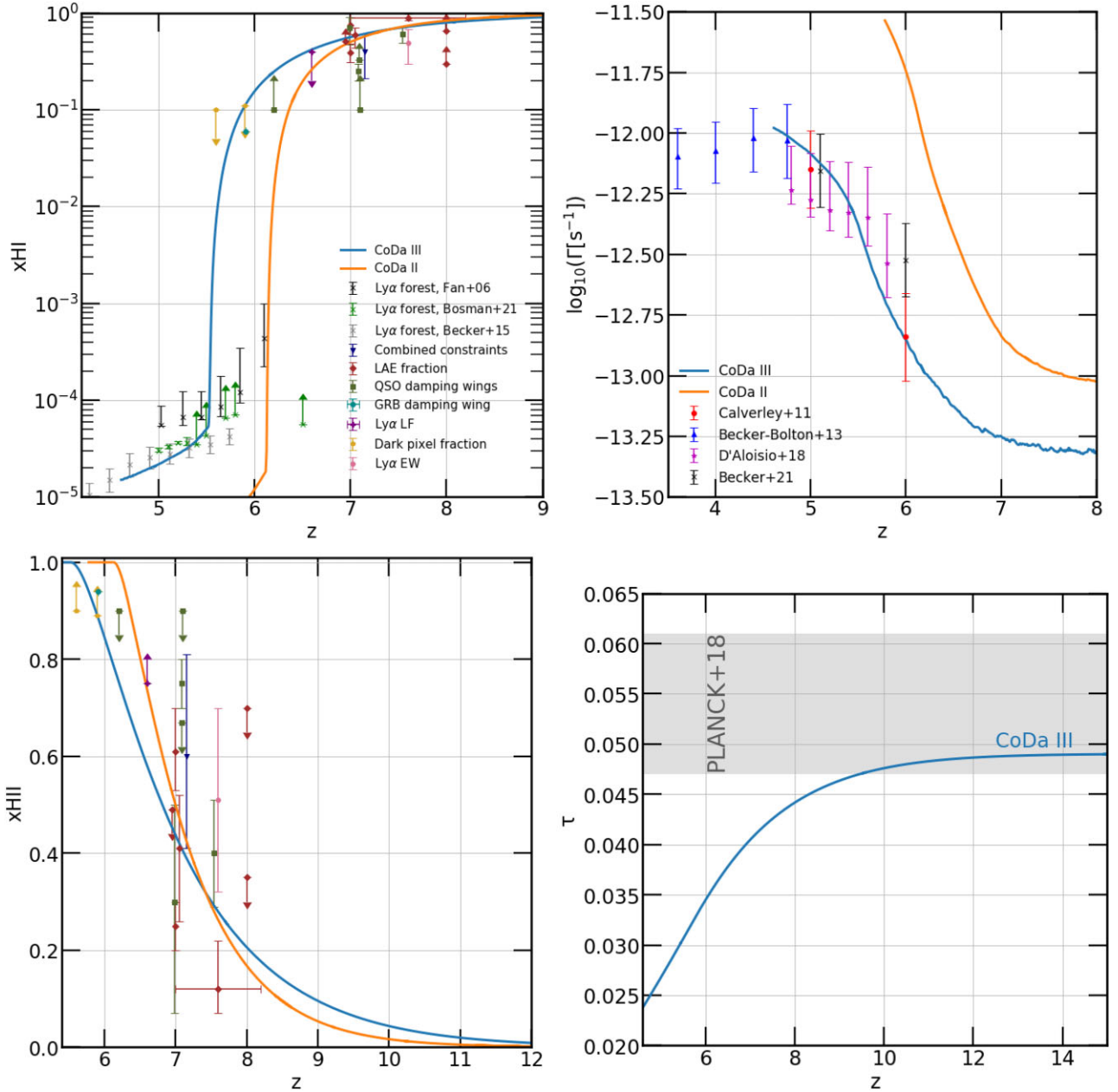


Figure 1. Top left: volume-weighted average neutral fraction along with a collection of constraints including Fan et al. (2006), Becker et al. (2015), Bosman et al. (2022), Ouchi et al. (2010), McGreer, Mesinger & D’Odorico (2015), Totani et al. (2016), Wang et al. (2021), Āurovūřková et al. (2020), Bañados et al. (2018), Schroeder, Mesinger & Haiman (2013), Mortlock et al. (2011), Tilvi et al. (2014), Pentericci et al. (2014), Schenker et al. (2014), Ono et al. (2012), Mason et al. (2018), Hoag et al. (2019), Greig & Mesinger (2017), and Jung et al. (2020). Top right: ionization rate in ionized regions when compared to constraints from Calverley et al. (2011), Becker & Bolton (2013), D’Aloisio et al. (2018), and Becker et al. (2021). Bottom left: volume-weighted ionized fraction. Bottom right: optical depth to CMB photons due to scattering by free electrons, compared to the Planck Collaboration LVII (2020) results (Planck18). CoDa III is in very good agreement with constraints on the history of ionization of the IGM and on the ionization rate, whilst remaining compatible with the optical depth due to electron scattering measured by Planck, despite a rather late end of reionization. Overall the improvement over CoDa II is very significant.

The top right-hand panel of Fig. 1 shows the evolution of the ionization rate in ionized regions ($\langle\Gamma_{\text{H II}}\rangle$) in both CoDa III and CoDa II. In the former, $\langle\Gamma_{\text{H II}}\rangle$ rises rapidly by roughly a factor of 10 between $z = 7$ and $z = 5.5$, after which the slope becomes shallower. CoDa III’s $\langle\Gamma_{\text{H II}}\rangle$ is a very good fit to the mix of observational constraints available, and a great improvement on the $\langle\Gamma_{\text{H II}}\rangle$ from CoDa II, which overshoots all constraints for the presented redshift range. This excess is symptomatic of CoDa II’s excessive ionization of the IGM (see Oc20; Oc21, for discussions about this). As with

$\langle x_{\text{H I}}\rangle$, the origin of this improvement lies both in the switch to a late reionization calibration and in the improved star formation subgrid model, as well as increased resolution (as discussed in Oc21).

The bottom right-hand panel of Fig. 1 shows that although reionization ends later in CoDa III, the optical depth due to scattering on free electrons seen by the cosmic microwave background (CMB) remains compatible with the constraints given by the latest Planck results (Planck18), demonstrating that a late reionization

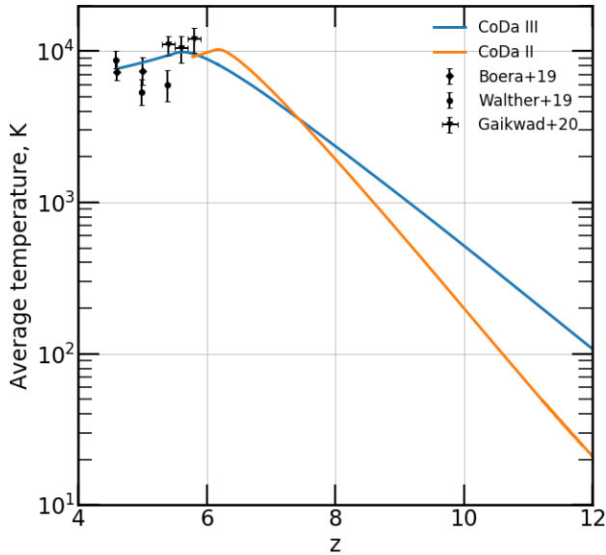


Figure 2. Average evolution of the gas temperature in CoDa III (blue) and CoDa II (orange). For comparison, constraints from Boera et al. (2019), Walther et al. (2019), and Gaikwad et al. (2020) are also shown.

scenario may still yield a high enough electron scattering optical depth.⁶

Fig. 2 shows the average evolution of the gas temperature in CoDa III (blue) and CoDa II (orange). In both simulations, the average temperature rises as reionization progresses and a larger and larger fraction of the IGM is photoheated. Eventually, the temperatures reach a characteristic peak as hydrogen reionization completes, before cooling slightly. We find that the average temperature in CoDa III (and to some extent in CoDa II) is consistent with recent constraints when available ($6 \geq z \geq 4.6$). CoDa II heats up much more rapidly, which is coherent with the earlier reionization and higher photoionization rates shown in Fig. 1.

Overall reionization in CoDa III reaches a better agreement with observational constraints on the IGM and other works from the literature than CoDa II did, signalling a marked improvement in our description of reionization.

3.2 Results: the mean free path of ionizing photons

We now move to our titular subject: the mean free path of ionizing photons (mfp). Fig. 3 shows the evolution of mfp over time in CoDa III. Observations from Worseck et al. (2014) and Be21 show that the mfp increases over time for $z < 6$. The full blue line with error bars shows that in CoDa III λ_{AT} increases from ≈ 1 pMpc at $z \approx 6$ to ≈ 20 pMpc at $z \approx 4.6$. In fact, λ_{AT} grows faster between $z \approx 5.75$ and $z \approx 5.2$, allowing CoDa III to match the newest low mfp from observations at $z = 6$ (Be21) and the high observed mfp at $z \lesssim 5.25$, whereas most other theoretical work has tended to overshoot the observed mfp at $z = 6$. Because of the stronger evolution of the mfp in CoDa III, we report values that are significantly lower than in other theoretical work for $z \gtrsim 6$, where observations are very challenging. This points to substantially different progressions of reionization in

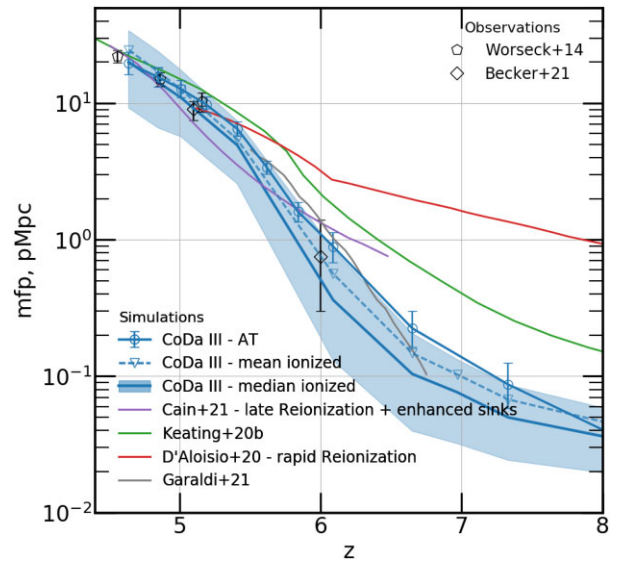


Figure 3. Evolution of the mean free path with redshift in CoDa III, and comparison to observational constraints from Worseck et al. (2014) and Becker et al. (2021) and previous simulation results from the literature (D'Aloisio et al. 2020; Keating et al. 2020; Cain et al. 2021; Garaldi et al. 2022). The blue graphics show results from CoDa III. The full blue line with circles and error bars shows the median and 25th–75th percentile region for λ_{AT} . The full blue line with triangles shows the mean λ_i , the blue dashed line and blue region show the median and 25th–75th percentile region λ_i . The median mean free path in CoDa III is in good agreement with existing observational constraints from $z = 4.5$ till $z = 6$.

the IGM of these other simulations, feasibly implying differences in ionizing emissivities and source populations.

Already, previous work (Ca21; Da21) has highlighted that the latest determinations of mfp by Be21 point to a late and rapid reionization, which CoDa III matches. However, unlike in Ca21, we do not need to invoke unresolved photon sinks in order to obtain our agreement with the observed mfp.

The THESAN simulation (Garaldi et al. 2022) also finds a good agreement with the $z = 6$ data point from Be21, however, unfortunately, the simulation ended too soon to comment on their agreement with lower redshift constraints, and it is therefore difficult to gauge the overall success of their simulation in reproducing the evolution of the properties of the IGM over the whole EoR, down to the post-overlap regime. In any case, it is possible that fully coupled RHD simulations are better able to match the mfp constraint from Be21 at $z = 6$ than post-processing or modelling methods, perhaps because of the more consistent treatment of the radiative–hydrodynamical interactions in the IGM and possible radiative feedback in galaxies, and/or the superior resolution.

In order to gain further insight into what drives the mfp measured in CoDa III, we show the distribution of individual LoS mfp (i.e. λ_{los}) at several redshifts in Fig. 4. At $z = 12.9$, before reionization is fully underway, the distribution of λ_{los} has a single mode near 10^{-4} pMpc. Although this is shorter than the simulation's cell size, it is physical; indeed, the general formula for the mfp reads $1/\sigma\rho_{\text{HI}}$, where σ is the interaction cross-section and ρ_{HI} the number density of neutral hydrogen atoms. In the dense, neutral high-redshift universe before reionization, mean free paths shorter than our cell size naturally occur and are fully consistent with the measurements presented here.

By $z = 7.3$, reionization is in full swing, and the distribution of λ_{los} is now bimodal, displaying a short mode and a long mode. The

⁶An even higher optical depth may still be achieved by adding the contribution from minihaloes at $z \geq 15$, adding an extra $\Delta\tau_{z \geq 15} \approx 0.008$, that are not treated in this work but suggested by Ahn & Shapiro (2021).

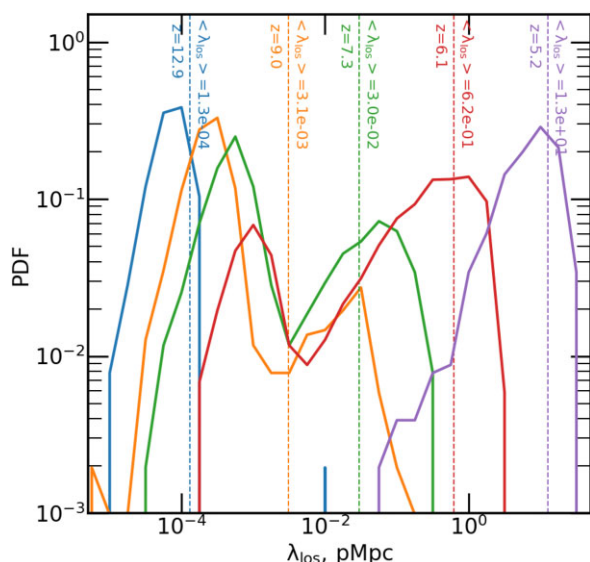


Figure 4. Distribution of λ_{los} . Dashed vertical lines show the distribution mean for each redshift. For readability, only a subset of redshifts roughly 200 Myr apart between $z = 12.9$ and $z = 5.2$ is shown. The λ_{los} are distributed in a neutral opaque mode of short mean free paths, and an ionized transparent mode of long mean free paths. The relative weight of these modes changes as reionization progresses.

short mode peaks around 8×10^{-4} pMpc and seems to descend from the $z = 12.9$ mode. However, approximately 30 per cent of LoS are now scattered around a secondary, longer mode peaking near $\lambda_{\text{los}} = 10^{-1}$ pMpc. This bimodality reflects the coexistence of large fully ionized regions already by $z = 7.3$, when reionization is almost complete in half of CoDa III’s volume (about 40 per cent ionized reading from the bottom left-hand panel of Fig. 1). At $z = 6.1$, both modes have moved to slightly higher λ_{los} values (10^{-3} pMpc for the short λ_{los} peak, and 0.64 pMpc for the long λ_{los} peak), and the relative weight between modes has swapped: now most LoS belong to the long λ_{los} peak.

By $z = 5.2$, reionization is complete, all of the LoS have been ionized and the short mode has disappeared in favour of the long mode. The evolution of the λ_{los} distribution reveals that the CoDa III λ_{los} evolves due to both gradual changes along every LoS (due to cosmic expansion driving the average physical density to decrease over time) and to rapid changes along short λ_{los} LoS due to reionization. The rapid increase of λ_{AT} in Fig. 3 between $z = 6.1$ and $z = 5.2$ could be driven by the disappearance of the shortest λ_{los} values as more and more LoS experience reionization.

Fig. 4 shows that the transition between the ionized mode of the λ_{los} distribution and the neutral mode occurs near 10^{-2} pMpc. Accordingly we select only those λ_{los} longer than this threshold for the ionized method. The blue solid curve with triangles in Fig. 3 shows the evolution of λ_i in CoDa III. For $z > 5$, both the mean and the median λ_i are lower than λ_{AT} . This makes our prediction based on the mean λ_i a better match to the $z = 6$ constraint from Be21, and also shows that the average transmission method (measuring the mean free path from an average simulated transmission spectrum) is biased towards high transmissions and overestimates the median λ_{los} at the redshifts considered in Fig. 3. For $z \leq 5$, the mean λ_i is slightly above the constraints from Worseck et al. (2014), however the median remains in very good

agreement and there is a broad scatter, comparable to the error bar of Be21.

4 CONCLUSIONS

In this paper, we have presented first results from our latest large-scale simulation of the EoR, CoDa III, the largest fully coupled RHD simulation of reionization ever performed. It improves on our previous CoDa II simulation in many aspects, such as using a two times higher spatial resolution and eight times higher mass resolution, as well as updates of the physical modules and subgrid models, in particular star formation and stellar source modelling. We show that, thanks to these improvements, CoDa III compares much better than CoDa II with observational constraints of the high-redshift Universe, in particular the evolution of the neutral hydrogen fraction that sees a dramatic improvement, as well as the ionizing rate in ionized regions, while the evolution of the ionized fraction and the electron scattering optical depth are also in good agreement with the constraints available. Moreover, we showed that CoDa III is the first fully coupled radiative hydrodynamics simulation to reproduce the latest constraints on the ionizing photon mean free path all the way from $z = 6$ down to $z = 4.6$ thanks to a late reionization ($z_{\text{rei}} = 5.6$) and a rapid evolution of the ionizing mean free path. In particular, CoDa III shows that the surprisingly low mean free path measured by Be21 can naturally occur in a well-calibrated simulation of the EoR such as ours. The distribution of mean free paths during reionization is generally bimodal, displaying a short, neutral mode and a long, ionized mode. Both modes shift to higher mfp as reionization progresses: the neutral mode peak shifts from sub-kpc to kpc scales of mfp, and the ionized mode peak from $0.1 \text{ Mpc } h^{-1}$ (at $z = 7$) to $\sim 10 \text{ Mpc } h^{-1}$ (at $z = 5.2$). The balance of these modes reflects the degree of ionization of the simulated volume. The bimodal distribution reduces to a single short (long) mode before (after) reionization, respectively. Finally, by using three different methods for computing the mean free path from CoDa III, we show that the choice of the method is important. In particular, measuring the mean free path from an average transmitted flux stacking a number of LoS tends to overestimate the mean free path otherwise obtained as the average or the median of the mfp of individual LoS. Overall, CoDa III’s description of reionization in the IGM appears to be a marked improvement over CoDa II, providing an adequate basis with which to perform a broad range of studies of the EoR, such as Lyman α transmission in the IGM, but also of star formation suppression, and the ionizing photons budget of galaxies.

ACKNOWLEDGEMENTS

The authors thank F. Davies and H. Katz for fruitful discussions. Cosmic Dawn III was performed on Summit at Oak Ridge National Laboratory/Oak Ridge Leadership Computing Facility (Project AST031) using an INCITE 2020 allocation. JSWL acknowledges support from the DFG via the Heidelberg Cluster of Excellence STRUCTURES in the framework of Germany’s Excellence Strategy (grant EXC-2181/1 – 390900948). JGS thanks Sergey Pilipenko for sharing the latest version of the ginnungagap code and for his guidance with using it. JGS acknowledges support from the ANR LOCALIZATION project, grant ANR-21-CE31-0019 of the French Agence Nationale de la Recherche. JGS also acknowledges the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for providing computing time to build the initial conditions of the simulations for this project on the GCS Supercomputer SuperMUC-NG at Leibniz Supercomputing

Centre (www.lrz.de). GY acknowledges financial support from MICIU/FEDER under project grant PGC2018-094975-C21. KA was supported by NRF-2016R1D1A1B04935414, 2021R1A2C1095136, and 2016R1A5A1013277. ITI was supported by the Science and Technology Facilities Council (grant numbers ST/I000976/1 and ST/T000473/1) and the South East Physics Network (SEP-Net). HP was supported by the World Premier International Research Center Initiative (WPI), MEXT, Japan and JSPS KAKENHI grant no. 19K23455.

DATA AVAILABILITY

The authors will share the data used for this paper upon reasonable request.

REFERENCES

- Ahn K., Shapiro P. R., 2021, *ApJ*, 914, 44
 Aubert D., Teyssier R., 2008, *MNRAS*, 387, 295
 Aubert D., Teyssier R., 2010, *ApJ*, 724, 244
 Bañados E. et al., 2018, *Nature*, 553, 473
 Becker G. D., Bolton J. S., 2013, *MNRAS*, 436, 1023
 Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, *MNRAS*, 447, 3402
 Becker G. D., D'Aloisio A., Christenson H. M., Zhu Y., Worseck G., Bolton J. S., 2021, *MNRAS*, 508, 1853 (Be21)
 Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, *ApJ*, 872, 101
 Bosman S. E. I. et al., 2022, *MNRAS*, 514, 55 (Bo22)
 Cain C., D'Aloisio A., Gangolli N., Becker G. D., 2021, *ApJ*, 917, L37 (Ca21)
 Calverley A. P., Becker G. D., Haehnelt M. G., Bolton J. S., 2011, *MNRAS*, 412, 2543
 D'Aloisio A., McQuinn M., Davies F. B., Furlanetto S. R., 2018, *MNRAS*, 473, 560
 D'Aloisio A., McQuinn M., Trac H., Cain C., Mesinger A., 2020, *ApJ*, 898, 149
 Davies F. B., Bosman S. E. I., Furlanetto S. R., Becker G. D., D'Aloisio A., 2021, *ApJ*, 918, L35 (Da21)
 Dawoodbhoy T. et al., 2018, *MNRAS*, 480, 1740
 Deparis N., Aubert D., Ocvirk P., Chardin J., Lewis J., 2019, *A&A*, 622, A142
 Dixon K. L., Iliev I. T., Gottlöber S., Yepes G., Knebe A., Libeskind N., Hoffman Y., 2018, *MNRAS*, 477, 867
 Dubois Y., Teyssier R., 2008, *A&A*, 477, 79
 Ďurovčíková D., Katz H., Bosman S. E. I., Davies F. B., Devriendt J., Slyz A., 2020, *MNRAS*, 493, 4256
 Fan X. et al., 2006, *AJ*, 132, 117
 Gaikwad P. et al., 2020, *MNRAS*, 494, 5091
 Garaldi E., Kannan R., Smith A., Springel V., Pakmor R., Vogelsberger M., Hernquist L., 2022, *MNRAS*, 512, 4909
 Gnedin N. Y., 2016, *ApJ*, 833, 66
 Greig B., Mesinger A., 2017, *MNRAS*, 465, 4838
 Hoag A. et al., 2019, *ApJ*, 878, 12
 Iliev I. T., Moore B., Gottlöber S., Yepes G., Hoffman Y., Mellema G., 2011, *MNRAS*, 413, 2093
 Jung I. et al., 2020, *ApJ*, 904, 144
 Kannan R., Garaldi E., Smith A., Pakmor R., Springel V., Vogelsberger M., Hernquist L., 2022, *MNRAS*, 511, 4005
 Kannan R., Vogelsberger M., Marinacci F., McKinnon R., Pakmor R., Springel V., 2019, *MNRAS*, 485, 117
 Keating L. C., Weinberger L. H., Kulkarni G., Haehnelt M. G., Chardin J., Aubert D., 2020, *MNRAS*, 491, 1736
 Kulkarni G., Choudhury T. R., Puchwein E., Haehnelt M. G., 2016, *MNRAS*, 463, 2583
 Kulkarni G., Keating L. C., Haehnelt M. G., Bosman S. E. I., Puchwein E., Chardin J., Aubert D., 2019, *MNRAS*, 485, L24

- Levermore C. D., 1984, *J. Quant. Spectrosc. Radiat. Transf.*, 31, 149
 Lewis J. S. W., Ocvirk P., Dubois Y., Aubert D., Chardin J., Gillet N., Thélie E., 2022, preprint ([arXiv:2204.03949](https://arxiv.org/abs/2204.03949))
 Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, *ApJ*, 856, 2
 McGreer I. D., Mesinger A., D'Odorico V., 2015, *MNRAS*, 447, 499
 Mortlock D. J. et al., 2011, *Nature*, 474, 616
 Ocvirk P. et al., 2014, *ApJ*, 794, 20
 Ocvirk P. et al., 2016, *MNRAS*, 463, 1462 (Oc16)
 Ocvirk P. et al., 2020, *MNRAS*, 496, 4087 (Oc20)
 Ocvirk P., Aubert D., 2011, *MNRAS*, 417, L93
 Ocvirk P., Aubert D., Chardin J., Deparis N., Lewis J., 2019, *A&A*, 626, A77
 Ocvirk P., Aubert D., Chardin J., Knebe A., Libeskind N., Gottlöber S., Yepes G., Hoffman Y., 2013, *ApJ*, 777, 51
 Ocvirk P., Lewis J. S. W., Gillet N., Chardin J., Aubert D., Deparis N., Thelie E., 2021, *MNRAS*, 507, 6108 (Oc21)
 Ono Y. et al., 2012, *ApJ*, 744, 83
 Ouchi M. et al., 2010, *ApJ*, 723, 869
 Pentericci L. et al., 2014, *ApJ*, 793, 113
 Planck Collaboration LVII, 2020, *A&A*, 643, A42
 Planck Collaboration XVI, 2014, *A&A*, 571, A16
 Rahmati A., Schaye J., 2018, *MNRAS*, 478, 5123
 Schenker M. A., Ellis R. S., Konidaris N. P., Stark D. P., 2014, *ApJ*, 795, 20
 Schroeder J., Mesinger A., Haiman Z., 2013, *MNRAS*, 428, 3058
 Sorce J. G. et al., 2016, *MNRAS*, 455, 2078
 Sorce J. G., Tempel E., 2018, *MNRAS*, 476, 4362
 Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
 Teyssier R., 2002, *A&A*, 385, 337
 Tilvi V. et al., 2014, *ApJ*, 794, 5
 Totani T., Aoki K., Hattori T., Kawai N., 2016, *PASJ*, 68, 15
 Walther M., Oñorbe J., Hennawi J. F., Lukic Z., 2019, *ApJ*, 872, 13
 Wang B. et al., 2021, *ApJ*, 916, 3
 Worseck G. et al., 2014, *MNRAS*, 445, 1745

APPENDIX A: CONSISTENCY CHECK OF MEAN FREE PATHS FOR INDIVIDUAL LINES OF SIGHT

In order to verify the λ_{los} and λ_i values for individual lines of sight (LoS), we also compute these using another method found in the literature (e.g. Rahmati & Schaye 2018). Using equation (1), one finds that

$$\frac{F(x)}{F_0} = \exp(-1), \text{ if } x = \lambda. \quad (\text{A1})$$

Applying this to each LoS, we search for the cells whose transmission $T_a(x_a) = F_a(x_a)/F_0$ and $T_b(x_b)$ bracket $\exp(-1)$. We then interpolate between the cells to find x so that $T(x) = \exp(-1)$. There are two limit cases that we must deal with. First, we must treat the case of very transparent LoS for which λ is longer than CoDa III's box length. In order to estimate the mean free path along these LoS, we equate their high transmissions with the absence of large overdensities and assume that the density of neutral hydrogen is constant and small along the LoS. Thus, we can obtain an estimate of the full mean free path by extrapolating the additional distance necessary for the integrated optical depth to reach one. Second, there are the cases of very opaque LoS (or at least starting in an opaque cell) where λ is smaller than the cell size of CoDa III. Here, assuming that this cell does not contain an unresolved absorber, we use the cell's optical depth to deduce λ .

The left-hand panel of Fig. A1 shows the distributions for both the fitting method presented in Section 2.2, and the method described above (that we call intercept method from hereon). Both methods are very consistent for short, long, and intermediate mean free paths, and throughout reionization in CoDa III. However, there are some discrepancies, most notably for the shortest and longest mean free

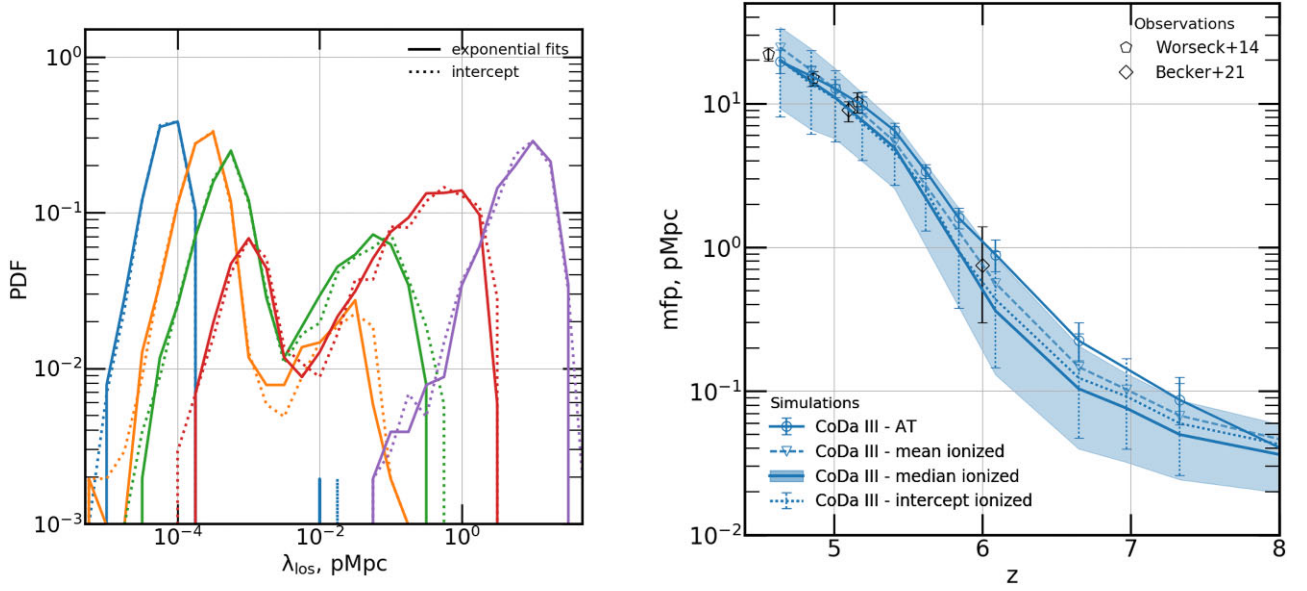


Figure A1. Left: distribution of λ_{los} . Dashed vertical lines show the distribution mean for each redshift. For readability, only a subset of redshifts roughly 200 Myr apart between $z = 12.9$ and $z = 5.2$ is shown. The λ_{los} are distributed in a neutral opaque mode of short mean free paths, and an ionized transparent mode of long mean free paths. The relative weight of these modes changes as reionization progresses. Right: we compute the evolution of the mean free path along ionized LoS, by applying the same methodology to the intercept distribution as used to compute the λ_{los} . The dotted lines show the result when compared to the other techniques we employ.

paths. This is not too surprising as the exponential fit is harder in these cases, and at the same time we must make some strong assumptions for the intercept method. The right-hand panel of Fig. A1 shows the average evolution of the ionized LoS as computed with all tested techniques, including the intercept method. As expected based on the distributions, the result is very similar to the evolution of the mean free path as derived using fitting. In fact, for $z \leq 6$, the two methods are almost indistinguishable. Though at higher redshifts, the

intercept method seems to predict a slightly higher mean free path. That these two methods independently yield such similar mean free paths, shows the robustness of our technique and results.

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.