Proceedings of the 4th International Conference on Computing Innovation and Applied Physics
DOI: 10.54254/2753-8818/87/2025.21031

Analysis of the Principle and Applications for the THESAN Project

Fanyue Zhou[1,a,*]

[1]*St. Margaret's Episcopal School, San Juan Capistrano, The United States of America*
a. 2025171@smes.org
**corresponding author*

Abstract: The THESAN Project combines the successful IllustrisTNG galaxy simulation framework with advanced radiation transport, non-equilibrium heating and cooling, realistic ionizing sources, and a cosmic dust model. Conducted on some of the world's most powerful supercomputers, THESAN consists of high-resolution Flagship simulation, which balances physical accuracy and box size to resolve hundreds of thousands of galaxies in a region of the universe spanning over 300 million light years. THESAN offers higher resolution than previous simulations of similar volume and employs galaxy formation simulations that match observations up to the present-day Universe. Its findings today align almost perfectly with recent observations. This study explains how THESAN accurately simulates different fields, e.g., having to consider concepts such as escape fractions and dark matter models. Moreover, to minimize errors, this simulation incorporated "softening", which reduces gravitational forces on scales of 2.2ckpc. The success of this project was also due to the assistance of other simulations, including the efficient radiation hydrodynamics solver AREPO-RT, which is explained later, that can capture the interaction between ionizing photons and gas and IllustrisTNG that can predict the properties of galaxies. With THESAN simulations, humanity will be able to understand interactions both within the IGM and the future Universe.

Keywords: Galaxy simulations, escape fractions, softening, AREPO-RT, THESAN project.

1. Introduction

The Universe contained hot dense plasma of matter and radiation after the Big Bang. As the gas expanded and cooled down, the protons were able to bond with free electrons and create hydrogen [1-3]. This led to less free election and made the Universe transparent to light. Later, the universe went into a period of darkness without sources of visible light. As stars and galaxies began to form, however, Lyman continuum radiation ionized and heated up the IGM, and this transition is known as the EoR [4-6]. The EoR refers to a period in the history of the universe when the IGM transitioned from being neutral to ionized. This event occurred around a billion years after the Big Bang, as the stars, galaxies, and black holes began emitting ultraviolet photons that stripped electrons from hydrogen atoms in the IGM [2, 7]. This ionization process created bubbles of ionized gas around galaxies, eventually allowing light to shine freely through space, marking a turning point of the Universe where baryonic matter first shaped the Universe. During the EoR, which occurred during the Cosmic Dawn, most of the hydrogen in the universe was not contained within galaxies (only about 5% was). At the same time, it influences the formation and evolution of later generations of galaxies

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/87/2025.21031

Figure 1 shows key phases, including the Age of Leptons, Nucleons, Nucleosynthesis, Ions, and Atoms, followed by the formation of the first stars and galaxies, and highlights important events, such as inflation, proton formation, and the cosmic microwave background [2, 8]. However, studying this period has been extremely challenging for scientists, as even the Hubble Space Telescope could only detect galaxies during the late stages of reionization [9]. While high-redshift quasars have provided some insights into the state of the large-scale IGM, they haven't fully enabled the construction of a comprehensive picture. Advancements such as the JWST helped scientists to get a more detailed and accurate understanding of cosmic reionization. The THESAN simulations are specifically designed to take full advantage of these observations, providing precise predictions of the large-scale reionization process. The THESAN project idea was first brought up in 2019 and the simulations were finished in 2021 [9, 10]. With the assistance of JWST, the accuracy and the researching scope expanded, since the JWST's infrared vision has allowed researchers to capture gas that accumulated to form these ancient galaxies and also measuring light from the first stars. The JWST has also used gravitational lensing to magnify distant light from the early universe, revealing faint galaxies that may have been responsible for most of the reionization. Furthermore, the 21 cm cosmology offers a new approach into the epoch of reionization, transforming humanity's understandings and views of the first stars and. In addition, 21 cm observations can help scientists distinguish between different reionization scenarios by analyzing the 21 cm bispectrum, which demonstrates differences in magnitude and sign across different scenarios, helping understand the nature of the sources of ionizing photons [11].

This paper will demonstrate how THESAN can provide humanity with a complete understanding of past, present and the future universe using its simulation abilities. Only by having a complete understanding of the universe and origins will scientists be able to achieve the next level of scientific accomplishments, such as moving on to the fourth dimension of research. Specifically, this paper will discuss the basic description and different stages that led to the present THESAN project. Other than the evolution of THESAN, numerous other important parts such as, applications, limitations, future goals, and newly updated data.

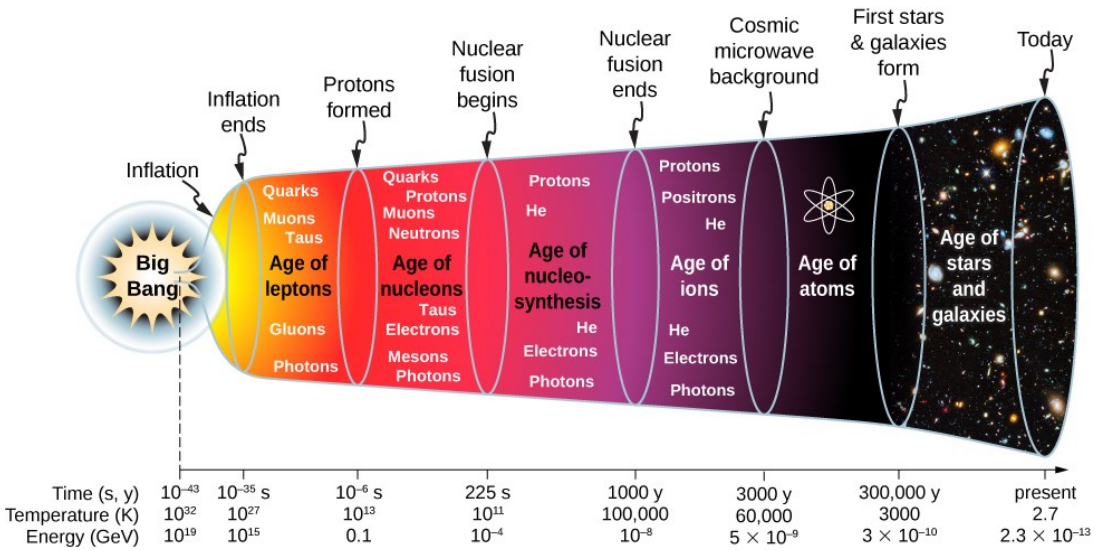


Figure 1: The history and evolution of the universe [2].

2. Principle and models for THESAN

As introduced above, scientists are able to solve hydrodynamical equations by using the AREPO code. To achieve higher-order accuracy, the method employs second order Runge-Kutta time integration

DOI: 10.54254/2753-8818/87/2025.21031

combined with a LSF gradient. This remains effective even on highly distorted meshes. The mesh is also regularly adjusted following the algorithm outlined by Vogelsberger. Magnetic fields are evolved using the ideal MHD equations [1], with an eight-wave formalism implemented to manage divergence errors. Gravity is computed using the HTPM approach, which uses a HOTA for short-range force estimation. The particle mesh technique is used for long-range forces, where particles are distributed into a grid of density values. Using the Fourier methods to solve the Poisson equation, scientists can calculate the gravitational potential. The Radiation Hydrodynamics Solver:

$$\frac{\partial Ni}{\partial t} + \frac{1}{a}\,\nabla \cdot \mathbf{Fi} = \mathbf{0}; \frac{\partial Fi}{\partial t} + \frac{c^2}{a}\,\nabla \cdot \mathbf{Fi} = \mathbf{0} \tag{1}$$

Some key ideas of cosmology need to be comprehended before one can understand this equation. Frequency Bins refers to the UV continuum (range of UV light frequencies) divided into three specific energy ranges that are particularly important for ionizing hydrogen and helium. These are: 13.6 eV, 24.6 eV, 54.4 eV, the energy that is required to ionize hydrogen, neutral helium, and singly ionized helium, respectively. Comoving Quantity is another concept. In cosmology, "comoving" refers to quantities that are adjusted to account for the expansion of the Universe. This adjustment ensures that researchers can compare quantities at different times in a meaningful way. In the Radiation Hydrodynamics Solver, the Comoving Photon Number Density (Ni) is the number of photons in a unit of volume, adjusted for the expansion of the Universe, while Comoving Photon Flux (Fi) is the rate at which photons pass through a unit area, also adjusted for the Universe's expansion. The first equation describes how the comoving photon number density changes over time, which is shown by partial derivative. The second term represents the divergence or spread of the photon flux in space, scaled by the factor to account for the expansion of the Universe. This is under the assumption that the Universe does not expand significantly before a photon is absorbed. Overall, the equation implies that any change in the photon number density over time is balanced by the flow of photons in space. The second equation describes how the comoving photon flux changes over time The second term involves the comoving pressure tensor P and the reduced speed of C~ (set to 0.2c in this context, where c is the speed of light in a vacuum). This term represents the spread of the photon flux, scaled by the expansion of the Universe. The reduced light speed approximation is used to simplify calculations without significantly affecting the accuracy of the results.

Escape fraction is another key idea in the THESAN Project [7]. The parameter of escape fraction refers to the measurement of HI ionizing photons that are not absorbed by the interstellar medium and are able to ionize the neutral hydrogen in the IGM:

$$\frac{dQ}{dt} = \frac{nion}{<nh>} - \frac{Q}{t} \tag{2}$$

Here, the first term represents the rate of change of the mass-weighted ionized hydrogen fraction $Q_{H\,II}$ over time. On the right-hand side, the first term is the rate at which ionizing photons contribute to the ionization of hydrogen, with n_{ion} being the total ionizing photon rate, which is determined by the sources that emit these photons (stars, black holes, while $<n_H>$ is the physical mean hydrogen density, which tells how efficiently the ionizing photons are ionizing hydrogen in the IGM. To elaborate on the former, it is calculated by multiplying the total number of ionized photons with the effective EF, which is the ones that escaped into the IGM. $Q_{H\,II}/\ t_{rec}$ accounts for the recombination of ionized hydrogen back into neutral hydrogen. Depending on different temperatures and density, t, which is the average recombination time, varies.

3. Applications

Since the inception of the THESAN project, numerous experiments and simulations have been conducted utilizing it (seen from Table 1) [2]. THESAN has generated a wealth of data under various

DOI: 10.54254/2753-8818/87/2025.21031

test conditions and with specific experimental parameters. Importantly, different THESAN models are tied to distinct research objectives, allowing scientists to select the most appropriate model for their specific field of study. This approach is analogous to the IllustrisTNG project, which offers models like TNG50, TNG100, and TNG300, each optimized for different applications. Table 1 is a table of the summary of different cosmological simulations of THESAN. The left column is the simulation or run of each model. Lbox in MPC is the size of the simulation box. N is the total number of particles used in the simulation. The mDM is the mass of dark matter particles, while Mgas is the mass of each gas particle. Regarding "softening", it refers to the softening length of gravitational forces, for instance, in order to prevent numerical divergences. Z end is the redshift at which the simulation ends. Fesc, is the EF of ionizing photons, which is how many photons escape into the IGM and can ionize the neutral hydrogen atoms.

Table 1: Data of THESAN simulations.

Name	Lbox[cMPC]	Nparticles	m dm[10Mo]	m gas [10Mo]
THESAN- 1	95.5	2 ×21003	3.12	0.58
THESAN- 2	95.5	2 ×10503	24.9	4.66

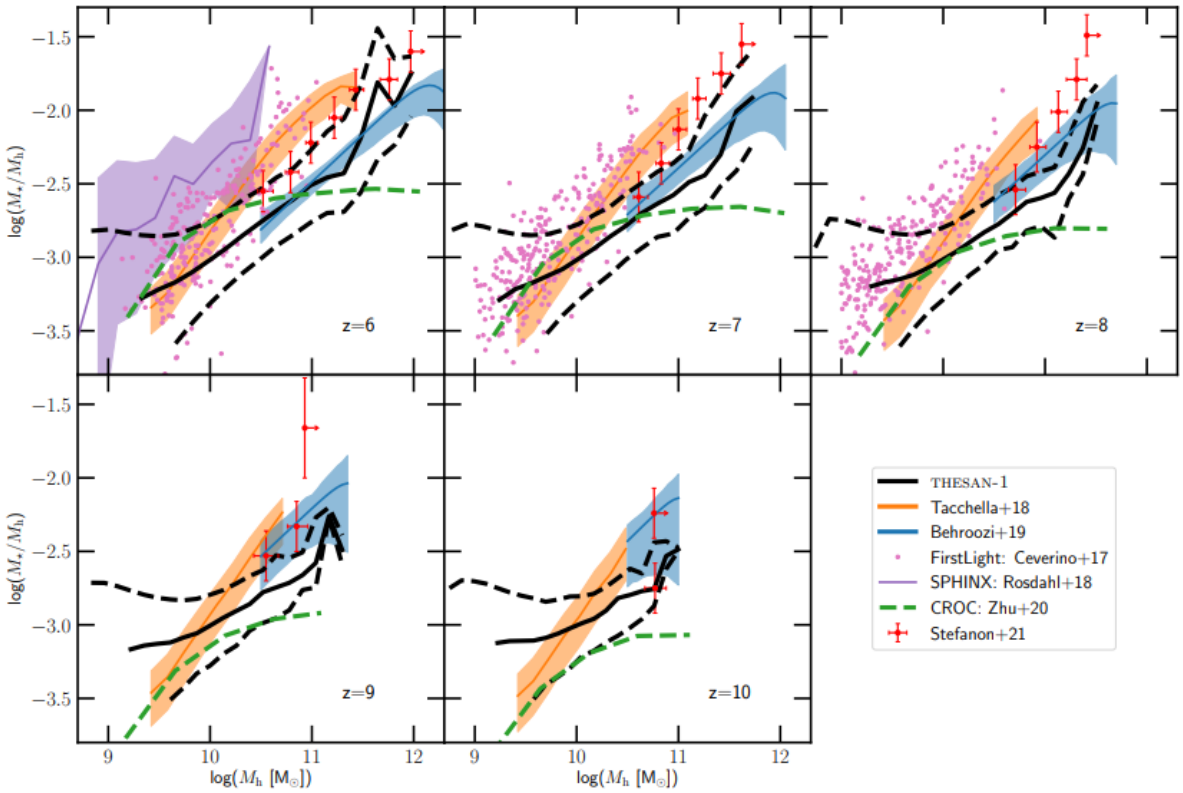


Figure 2: Simulations data from z = 6 - 10, in order. The black curve on each graph demonstrates the stellar to halo mass relation, while the dash curve indicates the range between 10 and 90 percentile of the galaxy distribution in the THESAN-1 simulation [4].

Figure 2 demonstrates how the abundance matching estimates from scientists align closely with the estimates from SHMR from Behroozi [4]. As mentioned in the background section, the James Webb Space Telescope (JWST) and the 21 cm cosmology contributed greatly to the THESAN project. JWST was essentially an upgrade to the Hubble Space telescope in many different perspectives, such as using infrared light detection instead of visible light, and therefore assisted the project undoubtedly.

DOI: 10.54254/2753-8818/87/2025.21031

As for the 21 cm cosmology, it is, to a certain extent, more complex. The behavior of the 21 cm power spectrum, a specific method in the 21 cm cosmology, is quite complex due to the potential influence caused by the ionization field and the gas density distribution, which could alter the outcomes. Using different redshifts as comparison can further complicate this situation, as the brightness temperature could evolve. Normally, at lower ionization fractions, which is below 0.3, and even at certain high ionization fractions, the mass density field is the main cause of the results, but as the bubble expands, a bump at K< 1 h/cMPC could occur. Since the bubble will eventually become so big that nearly all the power shifts to the large spatial scales, while the fluctuations in the ionization field diminishes. This position and bump is, therefore, important for insights into the reionization process [7].

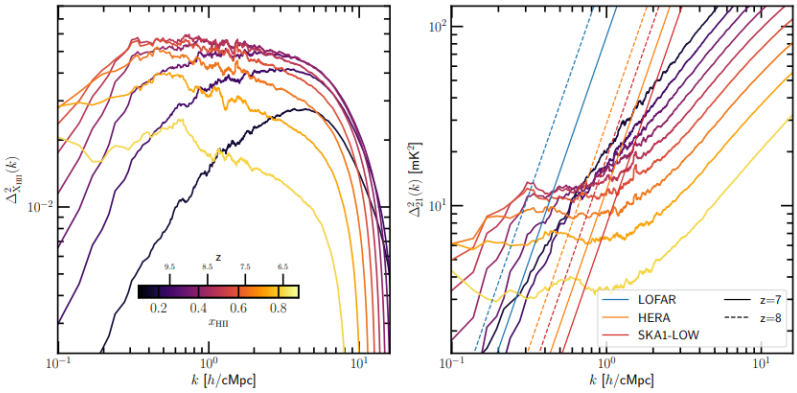


Figure 3: Simulation results collected from THESAN. The graph on the left side presents the power spectrum of IHs in THESAN 1 across various IFs with each increasing by 0.1. On the graph on the right, it demonstrates the data for a redshifted 21 cm signal [4].

A major challenge that had occurred in detecting the 21 cm signal is removing foregrounds. To illustrate, when diffusing supernova remnants, the ERS can be up to 10^6 times as the targeted signal. However, most of these sources are smooth, and there is a difference between foregrounds and the actual signal. Recent efforts have placed upper limits and got more accurate values, predicted by new simulations. Upcoming instruments will definitely improve in foreground avoidance and will lead to an accurate detection of the 21 cm signal during the EoR in the future. Seen from Figure 3, Ionized bubbles increased in size from the beginning of the ionization process to time afterwards. This is shown in the power spectrum, which peaks at higher wavenumbers and shifts to lower wavenumbers. Occurring in the 21cm signal as well, these changes are detectable by instruments including LOFAR, HERA, and SKA, particularly at redshifts z = 7 and 8 [4].

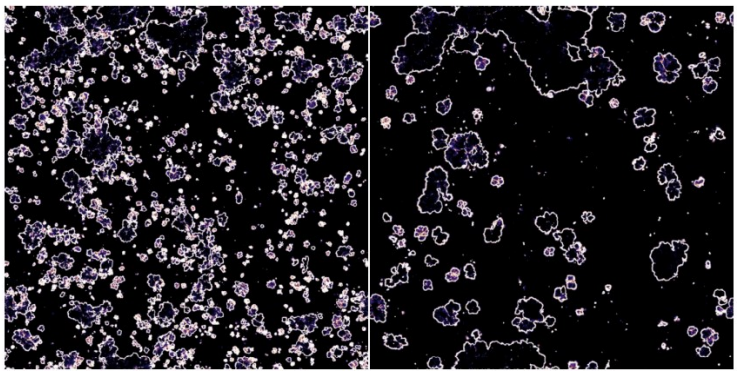


Figure 4: Simulation results collected from THESAN. The image on the left and right refers to a thesan-1 and thesan-2 simulation, respectively, both representing a bubble size distribution of ionization fraction of 0.3 [1].

DOI: 10.54254/2753-8818/87/2025.21031

To be more specific regarding the bubble sizes, the HII regions that are small and spread out evenly in the image of Figure 4 are those that are numerous but not luminous and more separated apart from each other [1]. The picture on the right demonstrates high mass halos, which are grouped more together, forming only in high density peaks. For this reason, the radiation output from one of the halos would give a really high and energetic one, causing the structure to be large bubble size with less numerous. The maps illustrate the boundaries of ionized bubbles by highlighting the regions with the largest ionization ingredients [8, 11]. This demonstrates how similar values of ionization could be represented differently, as thesan-1 exhibits many small bubbles, matching the bigger HI regions with a smaller number of bubbles in thesan-2 [4].

Currently, there is only one spin off project called THESAN HR, an extension of the THESAN project with higher resolution that is able to create smaller, more accurate, and more detailed simulations. The differences it created in patterns created significant different predictions for smaller objects, like low mass halos and low-metallicity star formation. As shown in Fig. 5, it is important to note that with the same initial conditions, physical model, output strategies, the two different models THESAN-HR and THESAN, which are top right, and bottom left respectively, exhibit slightly different simulation characteristics. The THESAN HR volume has much colder regions, where the temperature is less than 103 kelvins, and this is not possible in the Uniform UVB. The hot ionized regions are consistent between two models, but much denser cold filament is in THESAN. The zoomed in portion, as shown in Figure 5 for each section, refers to the HI gas in the massive galaxies, which is produced through self-shielding in Uniform UVB while in THESAN, scientists use the radiation fields to calculate self-shielding. Therefore, THESAN shows a factor of approximately 10 greater HI gas density in filaments, and a factor of 10 lower density in the surrounding CGM. Therefore, with a smaller volume and higher resolution, the THESAN-HR is able to examine detailed characteristics of the universe. In the future, there are definitely going to be more spin-off projects, such as the already planned THESAN-Zoom that hasn't been published yet. The THESAN simulation is above the curved white line while below it shows the UVB simulation. In the THESAN simulation, void regions exhibit cooler temperatures and remain neutral. The 80 kpc diameter circle surrounds the most massive galaxy in the volume, with close-up views provided for both simulations. In the visual, there is also THESAN simulation showing higher HI fraction than the UVB [1].

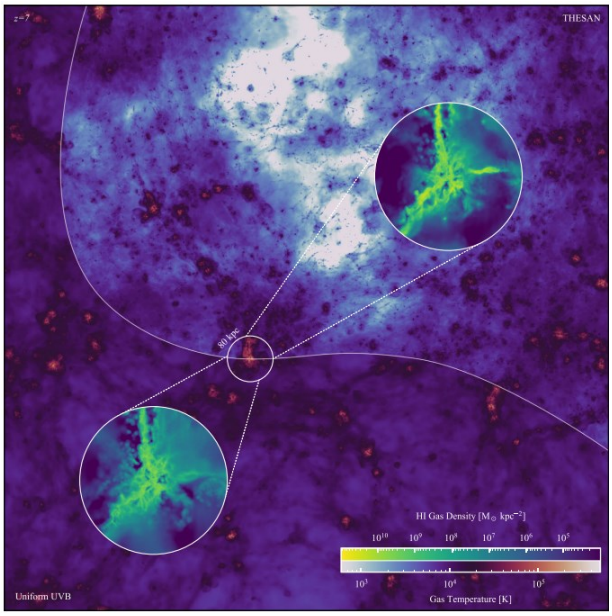


Figure 5: Simulation results collected from THESAN with a volume at a redshift of z = 7, and along the line, it shows the mass-weighted temperature [1].

DOI: 10.54254/2753-8818/87/2025.21031

4. Limitations and prospects

The foundation of the THESAN project has been set since 2022, and it has been improving ever since. From the initial IllustrisTNG with the assistance of Hubble Space Telescope to the new spin off projects and the complement of James Webb Space Telescope, the evolution of the THESAN Project has been extraordinary and impressive. The THESAN banner with two previous images: one illustrating the lightcone-like evolution and the other showing kaleidoscopic views. Zoomed- in region highlights a galaxy environment at z ~ 5.5, with additional insects depicting the galaxy as it would appear through the JWST and ALMA observatories [9-11].

Currently, there are still issues such as the constraints on the UVLF at faint magnitudes being unrealistic, but with more data and measurements of JWST, this discrepancy could be resolved. Furthermore, because high mass galaxies dominate the reionization process, which creates a graph of steady and consistent slope, it is hard to analyze the data. With future instruments, such as SKA, measurements could be more accurate and have higher sensitivity, and researchers will eventually have a detailed set of data, of which the slopes could be identified more clearly, and the graphs mentioned above would demonstrate a much more obvious trend. This can improve humanity's understanding of the bubble size distribution, and other ionizing photons at different redshifts. In the future, scientists have already planned on using the newly created instruments of JWST, ALMA, SKA etc., to specifically assign them to different areas and high-redshift structure formation and the Epoch of Reionization. For instance, the ALMA (Atacama Large Millimeter Array) is able to detect millimetric size wavelengths from the universe, which would help scientists with discovering the regions of Universe that are hidden from the range of visible light but provides images in the millimetric part of the spectrum. Lastly, THESAN also aims to contribute to educational purposes and making its findings more accessible to the public, including data and visuals. This project will eventually be used by educators, after it's completely perfected, for teaching the future generations and assisting them to have a better understanding of the processes that shaped the current universe.

5. Conclusion

The introduction of the THESAN project was presented in the first section, explaining how the THESAN project used the previous successful model IllustrisTNG simulation framework along with the world's most powerful supercomputers. Softening is another technique THESAN used, and it was helpful in reducing the gravitational forces that could alter the result, THESAN, therefore, provides a higher resolution of the simulation than any previous ones. The main purpose of THESAN is to model the IGM during the period of EoR, and with different assistance, it is able to provide a solid simulation. Furthermore, a background of the Universe during this time was introduced, as well as challenges that scientists faced that have the potential to be resolved in the future with better technologies and instruments. Moreover, the paper introduced two important equations used in the THESAN project: The Radiation Hydrodynamics Solver and the Effective Escape Fraction. The former is able to represent the interaction between ionizing H_{II} and gas, while the latter refers to the H_I photons not absorbed by the ISM and is able to ionize the neutral hydrogens. Then, several data and graphs were exhibited, along with explanations on the 21 cm cosmology and how it was able to assist the simulation. It is important for scientists to continue to put effort into this cutting-edge project of cosmology, and with the help of more data from instruments like JWST and ALMA, this project would definitely improve and contribute to the understanding of the universe.

DOI: 10.54254/2753-8818/87/2025.21031

References

[1] *Kannan, R., Garaldi, E., Smith, A. E., Pakmor, R., Springel, V., Vogelsberger, M., & Hernquist, L. (2022). Introducing the thesan project: radiation-magnetohydrodynamic simulations of the epoch of reionization. Monthly Notices of the Royal Astronomical Society, 511(3), 4005-4030.*

[2] *Garaldi, E., Kannan, R., Smith, A., et al. (2024). The thesan project: public data release of radiation-hydrodynamic simulations matching reionization-era JWST observations. Monthly Notices of the Royal Astronomical Society, 530(4), 3765-3786.*

[3] *Smith, A., Kannan, R., Garaldi, E., Vogelsberger, M., Pakmor, R., Springel, V., & Hernquist, L. (2022). The thesan project: Lyman-α emission and transmission during the Epoch of Reionization. Monthly Notices of the Royal Astronomical Society, 512(3), 3243-3265.*

[4] *Garaldi, E., Kannan, R., Smith, A., Springel, V., Pakmor, R., Vogelsberger, M., & Hernquist, L. (2022). The THESAN project: properties of the intergalactic medium and its connection to reionization-era galaxies. Monthly Notices of the Royal Astronomical Society, 512(4), 4909-4933.*

[5] *Borrow, J., Kannan, R., Garaldi, E., et al. (2023). thesan-hr: how does reionization impact early galaxy evolution?. Monthly Notices of the Royal Astronomical Society, 525(4), 5932-5950.*

[6] *Garaldi, E., Kannan, R., Smith, A., Springel, V., Pakmor, R., Vogelsberger, M., & Hernquist, L. (2020). Early structure formation in the THESAN radiation-magneto-hydrodynamics simulations. Proceedings of the International Astronomical Union, 16(S362), 1-7.*

[7] *Shen, X., Vogelsberger, M., Borrow, J., et al. (2024). The thesan project: galaxy sizes during the epoch of reionization. Monthly Notices of the Royal Astronomical Society, 534(2), 1433-1458.*

[8] *Moriwaki, K. (2022). Large-Scale Structure of the Universe: Cosmological Simulations and Machine Learning. Springer Nature.*

[9] *Yeh, J. Y., Smith, A., Kannan, R., et al. (2023). The thesan project: ionizing escape fractions of reionization-era galaxies. Monthly Notices of the Royal Astronomical Society, 520(2), 2757-2780.*

[10] *Kannan, R., Smith, A., Garaldi, E., et al. (2022). The THESAN project: predictions for multitracer line intensity mapping in the epoch of reionization. Monthly Notices of the Royal Astronomical Society, 514(3), 3857-3878.*

[11] *Neyer, M., Smith, A., Kannan, R., et al. (2024). The thesan project: connecting ionized bubble sizes to their local environments during the Epoch of Reionization. Monthly Notices of the Royal Astronomical Society, 531(3), 2943-2957.*