Draft version September 29, 2025
Typeset using LaTeX **twocolumn** style in AASTeX7.0.1

Detection of HI filament: Pair Stacking vs. Filament Stacking

Yuxi Meng ,[1,2] Jie Wang ,[1,2] Yingjie Jing ,[1,2] Hongxiang Chen ,[1,2] and Zerui Liu [1,2]

[1]*National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China*
[2]*School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China*

ABSTRACT

It is anticipated that the faint 21 cm signal emitted by neutral hydrogen within cosmic filaments can be detected. However, because of the signal's weakness, stacking techniques are necessary. We assessed two stacking methods—pair stacking and filament stacking—using the EAGLE and IllustrisTNG simulations. Pair stacking leverages the fact that cosmic filaments serve as the connectors between cosmic web nodes, while filament stacking directly aggregates cosmic filaments identified by galaxy distributions. Our analysis indicates that, although pair stacking is convenient, it faces contamination from massive structures, the signal from filament gets very weak after the contamination is removed. Conversely, HI detection via filament stacking appears more viable. The column density exceeds 10^{17} cm^{-2} even when all halos are masked, and it is nearly 10 times higher than what is achieved with pair stacking. The effectiveness of filament stacking could further increase with a high number density galaxy catalog and better spatial resolution in radio observation intensity mapping. With the advent of new optical and radio data, the future detection of HI filaments looks promising.

Keywords: Large-scale structure of the universe (902) — Cosmic web (330) — Intergalactic gas (812)

1. INTRODUCTION

Neutral hydrogen (HI), the most prevalent element in the universe, is crucial for galaxy star formation and serves as an effective tracer of baryon distribution. Cosmological simulations suggest the majority of HI is found within galaxies and the circum-galactic medium (CGM), with smaller amounts existing as clouds or pockets in the intergalactic medium (IGM). Large-scale cosmic filaments primarily consist of the photoionized warm-hot intergalactic medium (WHIM). Although HI is not the main component of the IGM, its presence is confirmed by Lyman-alpha absorption studies, indicating that HI may significantly influence star formation in galaxies. Specifically, the HI clouds located in cosmic filaments can move towards clusters, promoting rapid star formation in galaxies via cold accretion at high redshifts.

Detecting HI in cosmic filaments is challenging because of its low density, but the stacking technique offers a robust solution. By compiling measurements from numerous faint objects, the signal-to-noise ratio is enhanced, improving detectability. Filament detection is possible through the stacking of galaxy pairs, which is straightforward to apply and not sensitive to the choice of pairs. This technique has been employed in the detection of dark matter (J. Clampitt et al. 2016), magnetic fields (T. Vernstrom et al. 2021), mass-to-light ratio (T. Yang et al. 2020), WHIM (A. de Graaff et al. 2019), and shock fronts (T. Vernstrom et al. 2023). However, using pair stacking for neutral hydrogen detection (D. Tramonte et al. 2019) has not yielded high signals due to the very low HI column density in cosmic filaments. Thus, a more effective stacking technique is necessary for HI detection in cosmic filament.

The effectiveness of a stacking technique, as determined by the signal-to-noise ratio, largely hinges on the precision of stacking genuine cosmic filaments. Identifying filaments using galaxy distributions holds the potential for greater accuracy than merely pairing galaxies. Such stacking methods have been leveraged in discovering dark matter and the Warm-Hot Intergalactic Medium (WHIM) (H. Tanimura et al. 2020a,b, 2022). Yet, deriving accurate filament identifications from galaxy distributions can suffer due to incomplete galaxy catalogs. Various filament finders employing diverse algorithms have been developed, but their effectiveness is often debated. Consequently, we conduct a test to assess the effectiveness of both pair stacking and filament stacking.

Email: yxmeng@bao.ac.cn, jie.wang@nao.cas.cn

The latest cosmological hydrodynamical simulations are valuable for evaluating stacking techniques. They not only deliver galaxy populations with accurate properties but also represent the baryonic content in the CGM and IGM by integrating key physical processes affecting these regions. Consequently, they predict the HI content we focus on. The scale of these simulations is extensive, covering sufficient cosmic filaments spanning tens of megaparsecs, enabling more reliable predictions of filament signals.

This paper is organized as follows: Section 2 provides a concise introduction to the simulations utilized and outlines the stacking methodology. In Section 3, we explore the elements influencing the effectiveness of the stacking techniques and evaluate their capabilities. The discussion of results is presented in Section 4, followed by conclusions in Section 5.

2. METHODS

2.1. *Simulations*

We utilize the EAGLE (Evolution and Assembly of GaLaxies and their Environments) simulations alongside the IllustrisTNG (The Next Generation Illustris Simulations) to conduct our analysis. Each set employs distinct simulation algorithms and subgrid physics. This allows us to gauge the uncertainty stemming from varying galaxy formation models. Our focus on HI in the local universe prompts us to use the z=0 snapshots from both simulations.

The EAGLE simulations (The EAGLE team 2017; S. McAlpine et al. 2016) are a series of hydrodynamical simulations executed with a significantly enhanced version of the tree-SPH code GADGET-3. This simulation suite includes subgrid physics such as radiative cooling, photoheating, star formation, stellar evolution, metal enrichment, stellar feedback, black hole growth, and AGN feedback. It's calibrated against the stellar mass function, galaxy sizes, and the black hole–stellar mass relation of z=0 galaxies. In the EAGLE simulations, groups and subhalos are identified using the friends-of-friends (FOF) algorithm and the SUBFIND algorithm, respectively. We conducted postprocessing to derive the HI mass for each gas particle in accordance with A. Rahmati et al. (2013). EAGLE employs cosmological parameters derived from Planck 2014 ($\Omega_{m,0} = 0.307$, $\Omega_{\Lambda,0} = 0.693$, $\Omega_{b,o} = 0.04825$, $h = 0.6777$, $\sigma_8 = 0.8288$, $n_s = 0.9611$). We utilized the reference simulation REFL0100N1504, which features a 100 comoving Mpc box on each side, and contains 1504^3 dark matter particles with masses of $9.7 \times 10^6 \mathrm{M}_\odot$ and 1504^3 gas particles with masses of $1.81 \times 10^6 \mathrm{M}_\odot$.

The IllustrisTNG simulations (D. Nelson et al. 2019) utilize the moving-mesh code AREPO. Compared to the original Illustris simulation, the galaxy formation model has been improved, particularly in kinetic AGN feedback, galactic winds, and the incorporation of magnetic fields. Groups and subhalos are detected using the FOF and SUBFIND algorithms. Neutral hydrogen abundance is calculated on-the-fly as per A. Rahmati et al. (2013), and we subsequently conducted post-processing to determine the HI mass of the gas cells. The IllustrisTNG simulation adopts cosmological parameters based on Planck 2015 data ($\Omega_{m,0} = 0.3089$, $\Omega_{\Lambda,0} = 0.6911$, $\Omega_{b,o} = 0.0486$, $h = 0.6774$, $\sigma_8 = 0.8159$, $n_s = 0.9667$), which show slight deviations from Planck 2014. We applied the highest resolution run, TNG100-1, within the TNG100 series, encompassing a box size of 106.5 comoving Mpc per side, containing 1820^3 dark matter particles of $7.5\times10^6 \mathrm{M}_\odot$, alongside 1820^3 gas cells of $1.4 \times 10^6 \mathrm{M}_\odot$.

In our simulation, we classify the space into different types called knots, filaments, and voids. The knots are defined as areas within $5 \times R_{200,c}$ around halos with masses $M_h \geq 10^3$ $\mathrm{M}_\odot$/h. Filaments are designated as zones within 1 Mpc/h surrounding filaments identified by galaxies with stellar masses $M_* \geq 10^8$ $\mathrm{M}_\odot$/h using DisPerSE, while voids encompass the remaining portions. We observed that both the volume fraction and the HI mass fraction align consistently between the EAGLE and TNG simulations (see Fig. 1). Knots, being the densest structures, comprise approximately 20% of the HI content despite occupying just 2% of the volume. Filaments occupy about 7% of the total volume and account for nearly 50% of the HI. Meanwhile, the voids, which cover roughly 91% of the total volume, contain only around 30% of the HI.

2.2. *Stacking Techniques*

2.2.1. *Pair Stacking*

The pair stacking technique is developed to increase the signal emitted from the filaments which are identified by the straight lines between a pair of galaxies (J. Clampitt et al. 2016). It is supposed that galaxies residing in the most massive halos are usually connected by cosmic filaments. Therefore, by selecting pairs of galaxies with proper separations, the connection between pairs could be identified as the filaments in the cosmic webs. This is shown in the right panel of Fig. 2.

Other HI sources near the galaxy pairs might coincidentally align with the filaments(straight lines), potentially contaminating signals from the filaments. Assuming these structures are spherically distributed around the ends of the filaments in the stacked map, their con-

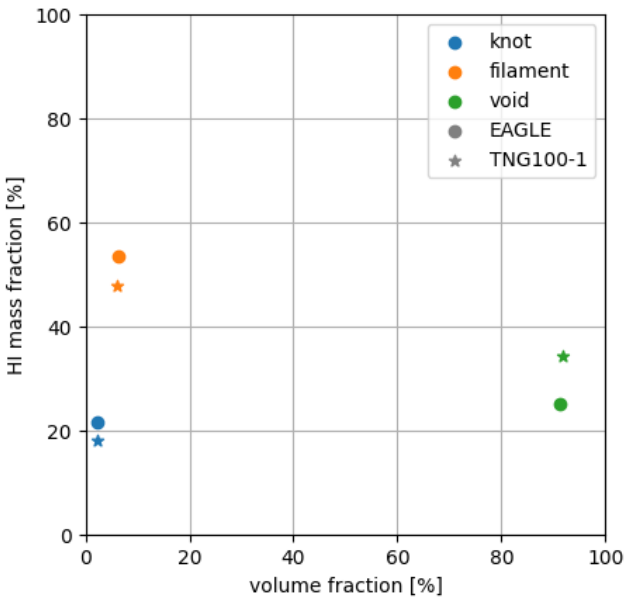

Figure 1. HI and volume fraction in knots, filaments, and voids in EAGLE (dots) and TNG100 (stars) simulations. The blue points represent knot regions, which are within 5 $R_{200,c}$ of haloes $M_h \geq 10^3$ M$_\odot$/h. The orange points represent filament regions, which are within 1 Mpc/h vicinity of the cosmic filaments that identified with galaxies of $M_* \geq 10^8$ M$_\odot$/h using DisPerSE. The green points represent void regions, which locate outside the knot and filament regions.

tribution can be subtracted from the total flux, resulting in a signal predominantly from cosmic filaments between the pairs, which peaks centrally in the stacked map.

To evaluate the detectability of filaments with pair stacking, we perform the stacking similarly to what the observations did. The procedures of stacking are as follows:

1. *Create a catalogue of galaxy pairs for the purpose of stacking.* The catalogue is generated based on galaxy database and specific criteria for selecting pairs. Central galaxies with a stellar mass above a certain threshold are chosen, analogous to Luminous Red Galaxies (LRG) defined by a limiting observational magnitude. Two galaxies are considered a pair if their projected separation lies between $6 - 14$ Mpc/h, and their line of sight distance is under 6 Mpc/h.

2. *Create maps of column density for each galaxy pair separately.* To create a column density map, we start by constructing a cube of HI density on a 0.1 Mpc/h grid. From this cube, we extract a box-shaped region around a pair, ensuring the region is 20 Mpc/h thick along the line of sight. We then rotate and project this box to obtain the density map.

3. *Produce stacked column density maps and evaluate the contamination from the end.* We stack maps for each galaxy pair to obtain a density map around the filaments, which includes the contribution from both the filaments and nearby contaminated galaxies. By assembling these maps and averaging each with its mirrored version, we produce a symmetrical map to depict contamination from adjacent galaxies at the ends. A least squares polynomial fit is conducted on sectors perpendicular to the paired galaxies, enabling us to determine the column density $\sigma(r)$ of the spherical profile, representing the contamination contribution.

4. *Estimate the intensity map around the filaments.* The stacked map is composed of filament signals and additional structures distributed spherically around the ends of the pairs. By subtracting this stacked map from the spherical profile, the residual map actually is the intensity map around the filaments.

5. *Identify and isolate the filament signal, then evaluate the associated uncertainties.* The residual map reveals a filament-like structure bridging both ends of the pair, with its peak intensity situated at the map's center. By isolating the central area of the residual map, we determine the filament signal's mean column density. We apply these steps across three projection directions within the simulation and assess the uncertainties in our findings using bootstrapping.

Following our procedures, we determined the column density of the filament signal along with an estimation of its uncertainties. The observed flux includes contributions from the filament signals, foreground contamination, and instrumental noise. Nonetheless, the filament signal strength measured reflects the intrinsic signal achievable through the pair stacking technique. A stronger filament signal increases the likelihood of detection using this technique.

2.2.2. *Filament Stacking*

Filament stacking leverages the pattern inherent in the distribution of galaxies themselves, rather than tracing cosmic filaments with pairs of galaxies. Various algorithms are designed to detect cosmic filaments within galaxy survey catalogues. Among these, DisPerSE (T. Sousbie 2011; T. Sousbie et al. 2011) excels due to its foundation in discrete Morse theory and persistence theory, enabling reliable identification of topological structures, such as cosmic filaments, from galaxy distributions (see Figure 2). DisPerSE begins by estimating a density field from the distribution of tracer particles like galaxies and then identifies critical points—maxima, saddle points, and minima—while tracing ridge-like integral lines between maxima and saddle points, which

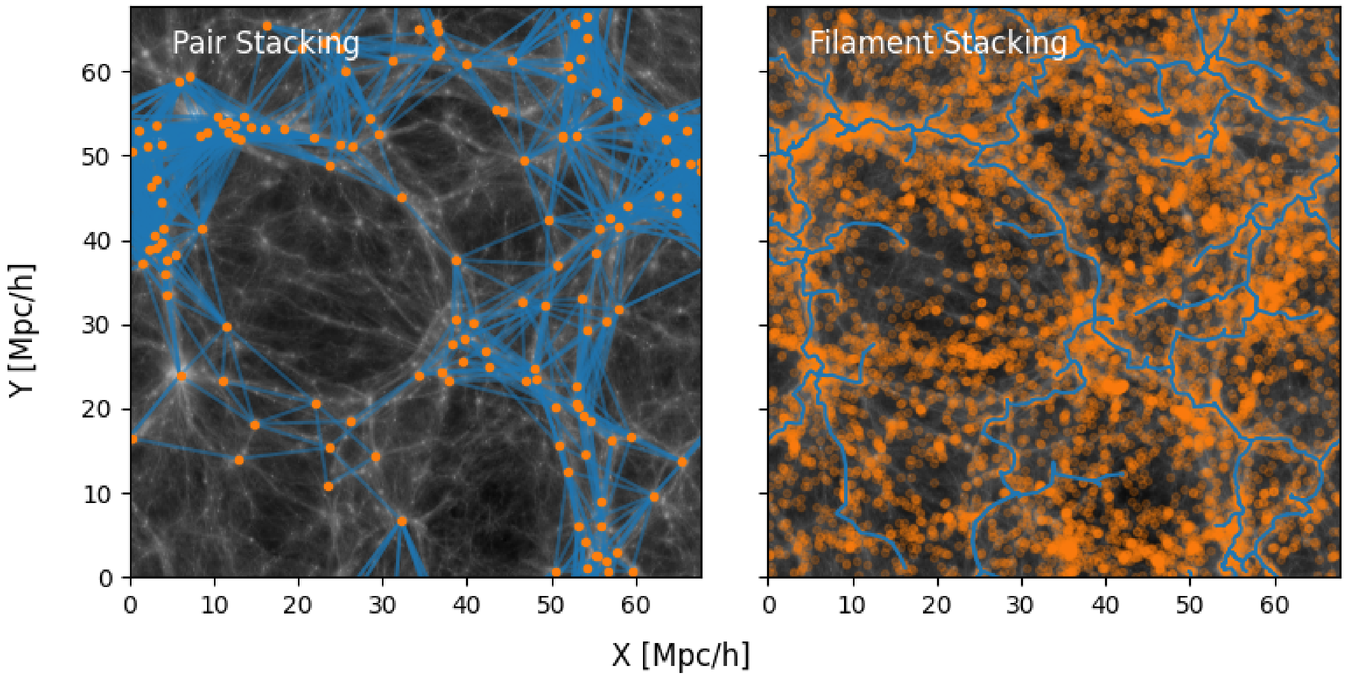

Figure 2. A demonstration of galaxy pairs (left panel) and filaments identified with galaxy distribution (right panel) in EAGLE simulation. Orange points represent galaxies in both panels, and blue lines represent galaxy pairs (left panel) or cosmic filaments (right panel). For clarity, the galaxy pairs are selected from the most massive galaxies of $M_* \geq 10^{11}$ $M_\odot$/h. The filaments are identified with galaxies with $M_* \geq 10^{10}$ $M_\odot$/h using DisPerSE.

serve as filament spines. Each filament is assigned a persistence ratio, and those with lower persistence (commonly measured in units of σ) are filtered out to retain only the strongest filaments. Filament stacking involves stacking the pixels (or voxels in 3D) directly according to their distance from the spines of cosmic filaments determined from the galaxy distribution.

We apply filament stacking to the simulation with the following prescriptions:

1. *Generate column density and galaxy maps from simulation.* We take a slice that is 20 Mpc/h thick from the simulation box, matching the thickness used in our pair stacking method. Then we project and create a two-dimensional distribution of galaxies and a HI column density map. Only galaxies above a specific stellar mass cutoff are included, which sets the corresponding galaxy number density, and the column density map is divided into pixels measuring 0.01 Mpc/h each.

2. *Identify the cosmic filaments with DisPerSE.* DisPerSE relies on several crucial inputs for identifying filaments, most notably the tracer distribution and the persistence threshold. In our research, we use the 2D distribution of galaxies above a certain stellar mass threshold (e.g., $M_* \geq 10^{10}h^{-1}M_\odot$), which corresponds to a given galaxy number density, as tracers. We select filaments that have persistence ratios greater than 3 for the stacking process, avoiding additional steps like smoothing. Figure 2 illustrates the tracers as orange dots and the detected filament spines as blue curves.

3. *Stack column densities radially around filament spines* Based on the column density map and filament catalogue, we sort pixels in the column density map by their distance to the filament spine, then get the radial profile of HI column density perpendicular to the filament spine. This method is implemented across all filament structures in the simulation box in three different projection directions. By taking the median of all the resulting column density profiles from these slices, we create a single curve, with the confidence interval determined by the 15th and 85th percentiles of all individual curves.

The column density profile we have determined demonstrates the fundamental capability of the filament stacking technique. Since the column density is measured in the same units as the filament signal obtained via pair stacking, we can evaluate the effectiveness of both stacking approaches by comparing their predicted column density values directly.

2.3. *Mask halo contribution*

HI in the universe resides in both the IGM and the galaxies, and it is necessary to evaluate the contribution of resolved galaxy HI sources in the stacked signals. HI in galaxies has column density as high as

$\sim 10^{20}\text{cm}^{-2}$, while HI packets in the IGM are only $\sim 10^{15}\text{cm}^{-2}$, orders of magnitude lower than the galaxies. The majority of galaxy HI resides in haloes of different masses, which can be masked. We found 70% of HI is within regions within twice $R_{200,c}$ of massive halos with $M_h \geq 10^{11}$ $\text{M}_\odot$/h, which only occupy $\sim 1\%$ of the volume of the simulation box.

For both pair and filament stacking, galaxies with HI in massive halos will be included. To clearly detect the IGM component, masking of resolved sources is necessary. This involves excluding masked pixels from column density maps and not including them in the average column density calculations for both stacking methods. The goal is to determine the mean density outside the vicinity of halos. We investigate the effectiveness of the two stacking methods, both with and without masking.

3. RESULTS

3.1. *Pair Stacking*

By stacking pairs selected from the galaxy catalogue, we obtain an average column density map on which a high-density bridge is connecting two peaks of density at both ends of the pair. The central region between the pair has a column density of $\sim 1.8 \times 10^{17}\text{cm}^{-2}$, including 75% contribution from background and 15% contribution from the spherical profile around both ends of the pair. The remaining 10% column density ($\sim 2 \times 10^{16}\text{cm}^{-2}$) is expected to represent cosmic filaments, and we can obtain its profile in the residual map by subtracting the stacked map from the fitted profile map (see the right-most part of the blue curve in Figure 3). We found this filament signal is generally smoothly distributed in a small region between the pair, and the maximum of the filament signal is located at the center of the pair.

Upon examining the column density maps of individual galaxy pairs, we notice that the stacked map is basically contributed by three types of pairs. First, there are pairs of galaxies connected by a cosmic filament, as anticipated. Through stacking these pairs, we can statistically determine the average HI column density within cosmic filaments. Secondly, we identify pairs where only a void exists between the galaxies. The absence of filaments in such pairs highlights the limitations of using galaxy pairs to trace cosmic filaments, potentially lowering the signal-to-noise ratio if too many of these pairs are included. The third case is when a high-density region, such as a galaxy group, located between the galaxies. These regions significantly contaminate the stacked signal because their HI column density is considerably higher than that found in the intergalactic medium (IGM). While subtracting the spherical profile can partially mitigate this contribution, a substantial impact still remains.

In pair stacking, there is significant contamination from the HI components within halos. Figure 3 shows the relationship between column density and the lower halo mass limit used for masking. The blue curve represents the column density stacked across all pairs, while the dotted and dash-dotted curves illustrate the contamination from halos and the residual filament signals, respectively. The total signal from all pairs reaches approximately $1.7 \times 10^{17}\text{cm}^{-2}$, with halo contamination accounting for around 80%, and the residual filament signal is roughly $2. \times 10^{16}\text{cm}^{-2}$.

The impact of halos with masses $M_h \geq 10^{15}$ $\text{M}_\odot$/h on the stacking results is minimal, likely due to their low number density. A decrease in stacked column density is observed when massive sparse distribution. Excluding halos with masses $M_h \geq 10^{14}$ $\text{M}_\odot$/h reduces the stacked column density, as these massive halos contribute about 1/8 of the total column density and 1/4 of the filament signal. Halos within the mass range of $M_h = 10^{11-14}$ $\text{M}_\odot$/h provide 3/4 of the total column density. When halos with masses equal to or greater than $M_h \geq 10^{13}$ $\text{M}_\odot$/h are masked, the filament signal decreases to approximately $\sim 10^{16}\text{cm}^{-2}$, which is less than half of the unmasked level. Further masking of halos with $M_h \geq 10^{12}$ $\text{M}_\odot$/h lowers the filament signal to under $2 \times 10^{15}\text{cm}^{-2}$, representing a tenfold reduction from its unmasked level. Finally, the filament signal disappears entirely when excluding halos with $M_h \geq 10^{11}$ $\text{M}_\odot$/h. Overall, detecting the HI signal within filaments in the stacked analyses of all pair samples is challenging.

Not all pairs in the stacking process improve the filament signal's signal-to-noise ratio. Pairs connected over voids, rather than over filaments, can actually diminish the signal. By selecting a subset of galaxy pairs carefully, one can reduce those linked to voids and thus enhance the stacked filament signal. Furthermore, it is also important to be cautious to avoid more contamination from halos, as choosing more pairs from dense regions can lead to this issue. Additionally, selecting fewer pairs might result in increased noise in actual observations. If the selection does not substantially improve the true filament signal, it could negatively affect the signal-to-noise ratio. Therefore, it is crucial to evaluate whether the filament signals can be reinforced to select certain subgroups from all pairs.

In Figure 3, we illustrate two distinct pair selection strategies, outlined as follows. The first strategy involves selecting pairs based on the stellar mass of galaxies at each end. Generally, more massive pairs are found

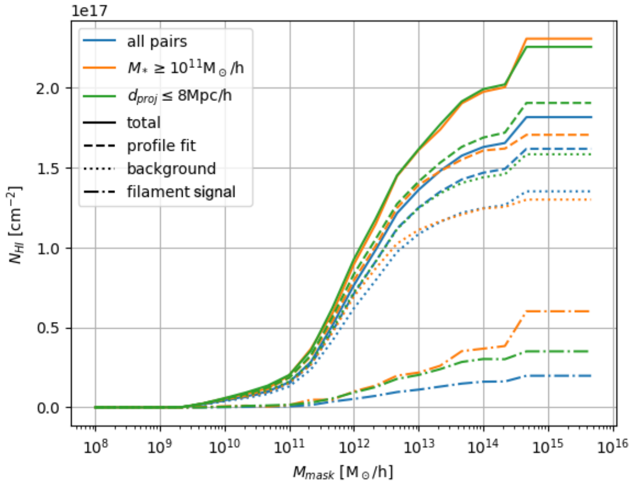

Figure 3. Total, halo contamination and background density and their dependence on the low limit of halo mass for masking. Curves of different colors represents different pair selections. Blue curves are from all pairs. Orange curves represents massive pairs with $M_* \geq 10^{11}$ M$_\odot$/h. Green curves are obtained with pairs of separation shorter than 8 Mpc/h. The solid, dashed, dotted, and dotted dashed curves represent total density, fitted spherical profile, background density level, and filament signal, respectively.

in denser regions and may align more closely with cosmic filaments. Pairs with galaxy masses $M_* \geq 3\times10^{10}$ M$_\odot$/h exhibit an overall density 5% greater than that of all pairs, though the filament signal column density shows only a slight difference. For the most massive pairs, with galaxies of $M_* \geq 10^{11}$ M$_\odot$/h, shown by the orange curves in Figure 3, the overall density is 25% higher without masking, and their filament signal density is approximately three times that for all pairs. However, most of this increased filament signal density is due to HI in massive halos, as the difference is negligible when halos of $M_h \geq 10^{12}$ M$_\odot$/h are masked.

The second strategy focuses on pair separation, indicating a different filament population. Short pairs signal shorter, more robust filaments in denser environments. By stacking pairs with separations less than 8 Mpc/h, we achieve an overall column density comparable to the most massive pairs, with filament signal density 75% higher than that of all pairs when unmasked. Yet, this filament signal advantage vanishes when masking halos of $M_h \geq 10^{12}$ M$_\odot$/h, suggesting the enhancement is mainly due to these halos. Longer pairs have column densities similar to or less than those of all pairs.

In a word, by carefully choosing the subsample, the signal can be significantly enhanced; however, separating it from halo contamination remains challenging.

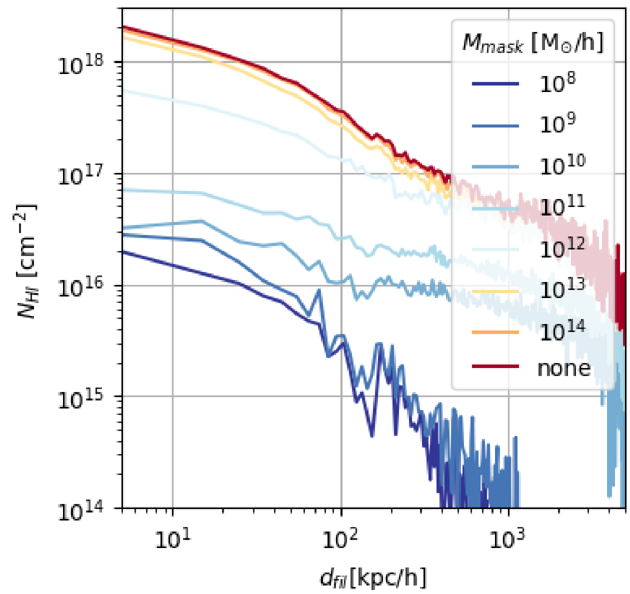

Figure 4. Density profile obtained with filament stacking, with different mask strength in EAGLE simulation. Filaments are identified with galaxies of $M_* \geq 10^8$ M$_\odot$/h, and curves from red to blue represent mask strength from none to $2 \times R_{200,c}$ region around halos of $M_h \geq 10^8$ M$_\odot$/h.

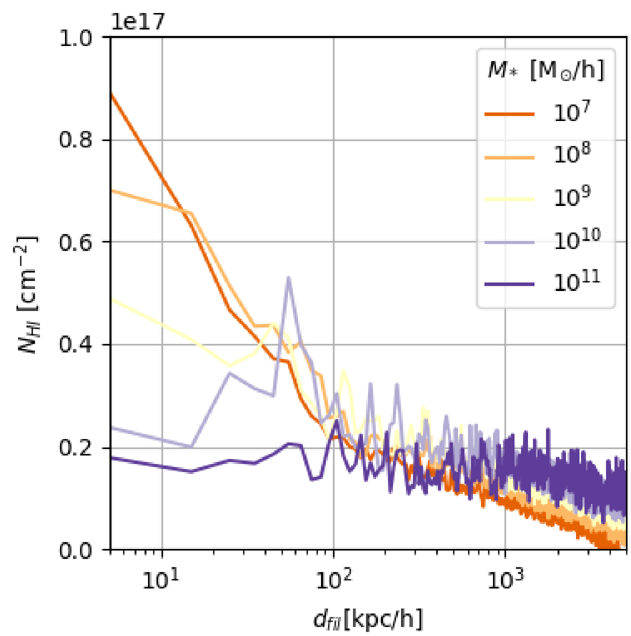

Figure 5. Density profile obtained with filament stacking, with different galaxy number densities (corresponding to different stellar mass thresholds) in EAGLE simulation. Regions within $2 \times R_{200,c}$ around halos of $M_h \geq 10^{11}$ M$_\odot$/h are masked. Curves from red to blue represent galaxy number densities increasing from the sparsest sample (all galaxies of $M_* \geq 10^{11}$ M$_\odot$/h) to the densest sample (all galaxies of $M_* \geq 10^7$ M$_\odot$/h).

3.2. *Filament Stacking*

The radial density profile obtained through filament stacking illustrates the density distribution perpendicular to the axis of cosmic filaments. Figure 4 shows this profile with various mask halo masses. These profiles are distinctly concentrated at the spine, unlike the average measurements observed in pair-stacked filaments. Within a 10 kpc/h radius, the stacked column density can be up to eight times higher than at 100 kpc/h. As a result, the effectiveness of filament stacking methods heavily depends on the resolution of observations, since only adequate resolution can reveal the peaks in density.

Masking HI in the most massive halos with $M_h \geq 10^{14}$ $\mathrm{M}_\odot$/h does not significantly alter the stacked density, since these halos occupy a small volume despite their high column density. In contrast, when HI is masked in halos with $M_h \geq 10^{11}$ $\mathrm{M}_\odot$/h, the average column density within 10 kpc/h decreases to approximately $5-10\%$ of the unmasked values, while the density at 100 kpc/h decreases to about $10-20\%$. HI in halos with $M_h = 10^{11-14}$ $\mathrm{M}_\odot$/h is highly concentrated at the center of the filament spine, so removing these halos results in a smoother stacked density profile. As for halos with $M_h = 10^{9-11}$ $\mathrm{M}_\odot$/h, the HI at the center of the spine is comparable to that in the intergalactic medium. These halos are more evenly spread perpendicular to the filament spine and substantially contribute to the outer regions of the stacked filaments. Meanwhile, the amount of HI in halos with $M_h < 10^9$ $\mathrm{M}_\odot$/h is negligible.

The success of stacking filaments hinges on how accurately filaments are identified; improved filament identification enhances stacking outcomes. The galaxy number density of the catalog (set by the stellar mass threshold) is the key determinant in filament identification efficacy. In Figure 5, we show how galaxy number density affects the density profile of stacked filaments. Identifying filaments using only the brightest galaxies with $M_* \geq 10^{11}$ $\mathrm{M}_\odot$/h often leads to ineffective identification. In these cases, the cosmic web seems to link galaxy clusters without genuine filaments. The hydrogen distribution, perpendicular to filament spines, appears nearly uniform. Masking halos drastically reduces column density, indicating the limitations of identification in an incomplete galaxy catalog that focuses mainly on halos. Higher galaxy number density (achieved by lowering the minimum stellar mass threshold) helps in correctly identifying filaments, pinpointing filament spines more precisely, and detecting smaller filaments missed by less complete catalogs. Avoiding false identifications by using catalogs with higher number density improves column density around filament spines in the stack. Specifically, within 10 kpc/h of the spine, higher galaxy number density results in increased column density due to precise spine location. However, at approximately 100 kpc/h, density declines when the number density rises for $M_* < 10^9$ $\mathrm{M}_\odot$/h, due partly to accurate spine localization. This trend is also influenced by including smaller, more diffuse filaments. Beyond 1 Mpc/h from the spine, higher number density lowers stacked column density, as this region includes voids and previously unidentified filaments. Higher number density reduces unidentified filaments, explaining the downward trend.

3.3. *Comparison of two methods*

Our study reveals that the filament spine's stacked density, derived through filament stacking, is ten times higher than the filament signal detected via pair stacking. Both techniques are affected by contamination from HI in halos, yet the signal from pair stacking diminishes more sharply as masking intensifies. The reliability of filament stacking is contingent upon accurate filament identification, and higher galaxy number density in the catalogue can enhance its performance. For pair stacking, restricting analysis to the most massive or shortest pairs amplifies the filament signal, but this is largely dominated by HI in halos with $M_h \geq 10^{12}$ $\mathrm{M}_\odot$/h, not the expected HI. The reliance on spatial resolution varies between the two stacking methods. Filament stacking generates a density profile with a central peak in the spine, and elevated resolution boosts the stacked density. In contrast, pair stacking remains unaffected by spatial resolution since the filament signal is uniformly distributed between the pair's centers. For an overview of the anticipated filament column density detection capabilities of these stacking methods, refer to Figure 6.

It is more practical to evaluate the efficacy of two stacking methods given existing observational capabilities. In state-of-the-art galaxy surveys like DESI BGS (C. Hahn et al. 2023) is expected to achieve comoving galaxy number density $n_{gal} \geq 10^{-2}$ h^3 Mpc^{-3} at low redshifts, comparible to minimum stellar mass $M_* \sim 10^9$ $M_\odot$ h^{-1}. Meanwhile, radio telescopes such as FAST can achieve a spatial resolution of approximately ~ 26 kpc/h at this redshift. With these levels of galaxy number density and spatial resolution, we anticipate that filament stacking (depicted by the orange solid curve in Figure 6) surpasses pair stacking (shown by the red solid curve in Figure 6) by about $\sim 1.5-2$ orders of magnitude across all masking strengths. This suggests that identifying filaments with the current galaxy catalogue is more effective than the basic assumption of cosmic filaments being in pairs, and the resolution

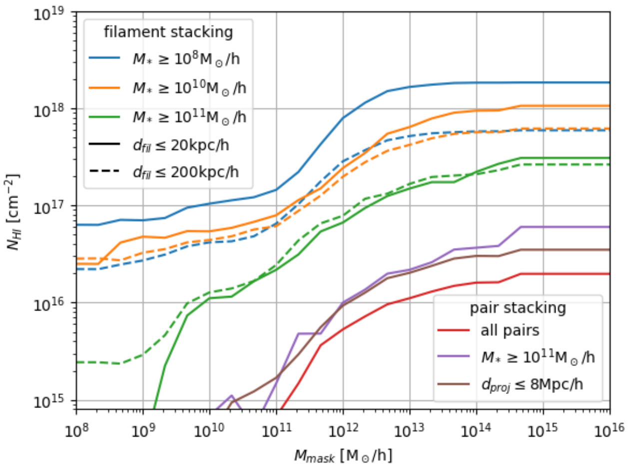

Figure 6. Comparison of two stacking methods. The blue, orange and green curves represents results of filament stacking, obtained with galaxy catalogues of different number densities (set by stellar mass thresholds). In these three curves, solid ones are the peak of the density profile within 20 kpc/h around the spine of filaments, and dashed ones are mean density within 200 kpc/h. The red, purple and brown curves represents filament signal of pair stacking. The red curve is obtained with all pairs, purple curve is obtain with pairs connecting the most massive galaxies of $M_* \geq 10^{11}$ $M_\odot$/h, and brown curve is obtained with the shortest pairs.

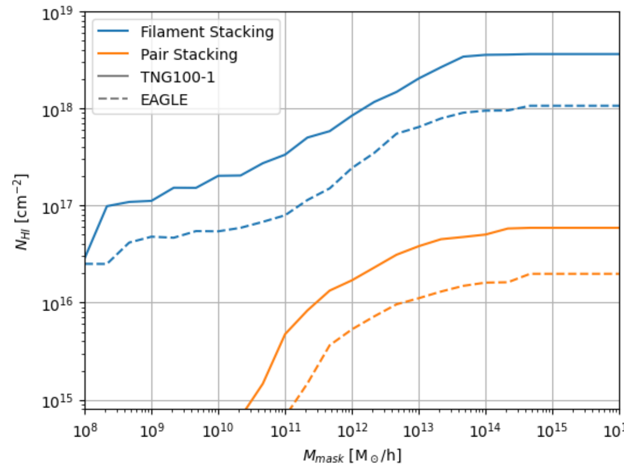

Figure 7. Comparison of two stacking methods in EAGLE and IllustrisTNG simulations. The blue curve is the mean HI density within 20 kpc/h around the filaments identified with galaxies of $M_* \geq 10^{10}$ $M_\odot$/h. The orange curve is the filament signal in all pairs with galaxies of $M_* \geq 10^{10}$ $M_\odot$/h. The solid curves represents results in TNG100-1, and dashed curves is from EAGLE simulations.

of single-dish telescopes is adequate to capitalize on the advantages of filament stacking.

4. DISCUSSIONS

When predicting HI content in simulations, caution is needed due to the substantial simplifications of physical processes compared to reality. Figure 7 illustrates the anticipated HI column density using two stacking methods applied in the EAGLE and IllustrisTNG simulations. Both simulations rely on post-processed data calibrated to the brightest sources, resulting in significant uncertainty regarding HI in the intergalactic medium and low-mass halos. The models predict different outcomes, with IllustrisTNG anticipating a HI column density approximately half an order of magnitude greater than EAGLE's prediction. Nevertheless, the relative differences between filament stacking and pair stacking methods are consistent across both simulations. This suggests that while the absolute HI density is contingent upon the simulation's subgrid physics, filament stacking demonstrates an effectiveness approximately $1.5-2$ orders of magnitude greater than pair stacking.

We adopted some approximations in our analysis as follows:

1. We conduct our analysis in real space, whereas observations occur in redshift space, which experiences redshift space distortion. This distortion affects both galaxies and HI in the large-scale structure, as well as the galaxy pairs and cosmic filaments identified with galaxies that trace the HI content in redshift space. However, we observe that redshift space distortion might influence the accuracy of filament identification, warranting careful consideration during observations. Despite this, we anticipate that the identification of filaments from projected galaxy distributions remains largely unaffected by redshift space distortion.

2.We consider that the observed signal, specifically the brightness temperature, correlates with the column density of HI when the HI object is optically thin. Yet, in galaxies, HI could be optically thick, causing 21cm line absorption. This would result in the actual 21cm emission being lower than what is predicted if assuming HI is optically thin. Nonetheless, this impact is likely minimal, as the majority of HI in the local universe is indeed optically thin.

5. CONCLUSIONS

This study contrasts two stacking techniques utilized in the EAGLE and IllutrisTNG simulations. The pair stacking method operates on the straightforward premise that cosmic filaments serve as structures linking massive halos, incorporating an additional spherical profile removal to assess filament contribution. Meanwhile, filament stacking leverages the cosmic web identified by galaxy distribution and is heavily dependent on effective filament identification. Both techniques share the objective of enhancing the detection of faint filaments by achieving a higher signal-to-noise ratio through the stacking process.

We have the following conclusions for these two stacking methods.

- Both stacking process is affected by contamination from HI present in halos. When the majority of halos are masked, filament stacking maintains roughly $\sim$ 10% of its original unmasked value. However, the filament signal observed in pair stacking diminishes significantly and loses its significance.

- Filament stacking results in a density profile that peaks at the spine's center, whereas in pair stacking, the filament signal is evenly distributed. Consequently, the outcomes of filament stacking are more influenced by the observation's spatial resolution.

- The selection of pairs in pair stacking can influence the values of the stacked column density, particularly for the most massive or shortest pairs. However, once the contribution from halos is excluded, this signal in the filament almost vanishes.

- The results of filament stacking are significantly influenced by the galaxy number density of the catalog (set by the stellar mass threshold). Catalogues with higher number density allow for the identification of more filaments and enhance the precision in determining the location of filament spines.

The two stacking methods stem from distinct principles. The pair stacking approach offers greater ease of use and shows less susceptibility to pair selection; however, it suffers from significant HI contamination in the most massive halos and struggles to identify HI within smaller halos or in cosmic filaments' diffuse phases. On the other hand, filament stacking, despite depending on the galaxy catalogue's thoroughness, surpasses pair stacking in efficacy. The advanced generation of galaxy surveys facilitates accurate cosmic filament identification, which supports HI detection through filament stacking, particularly after masking the most massive halos. Furthermore, higher-resolution observations with radio interferometers should reveal HI filaments, as the HI is primarily located along the cosmic filaments' spine. Overall, with the current galaxy catalogue, filament stacking shows enhanced effectiveness at low redshift.

ACKNOWLEDGMENTS

This work is supported by the National Key R&D Program of China (2022YFA1602901), the NSFC grant (Nos. 11988101, 11873051, 12125302, 11903043) and CAS Project for Young Scientists in Basic Research Grant (No. YSBR-062).

Software: Astropy (Astropy Collaboration et al. 2013, 2018), DisPerSE (T. Sousbie 2011; T. Sousbie et al. 2011), Matplotlib (J. D. Hunter 2007), Numpy (C. R. Harris et al. 2020), Scipy (P. Virtanen et al. 2020).

REFERENCES

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, doi: 10.3847/1538-3881/aabc4f

Clampitt, J., Miyatake, H., Jain, B., & Takada, M. 2016, MNRAS, 457, 2391, doi: 10.1093/mnras/stw142

de Graaff, A., Cai, Y.-C., Heymans, C., & Peacock, J. A. 2019, A&A, 624, A48, doi: 10.1051/0004-6361/201935159

Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2023, AJ, 165, 253, doi: 10.3847/1538-3881/accff8

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357, doi: 10.1038/s41586-020-2649-2

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90, doi: 10.1109/MCSE.2007.55

McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, Astronomy and Computing, 15, 72, doi: 10.1016/j.ascom.2016.02.004

Nelson, D., Springel, V., Pillepich, A., et al. 2019, Computational Astrophysics and Cosmology, 6, 2, doi: 10.1186/s40668-019-0028-x

Rahmati, A., Pawlik, A. H., Raičević, M., & Schaye, J. 2013, MNRAS, 430, 2427, doi: 10.1093/mnras/stt066

Sousbie, T. 2011, MNRAS, 414, 350, doi: 10.1111/j.1365-2966.2011.18394.x

Sousbie, T., Pichon, C., & Kawahara, H. 2011, MNRAS, 414, 384, doi: 10.1111/j.1365-2966.2011.18395.x

Tanimura, H., Aghanim, N., Bonjean, V., Malavasi, N., & Douspis, M. 2020a, A&A, 637, A41, doi: 10.1051/0004-6361/201937158

Tanimura, H., Aghanim, N., Douspis, M., & Malavasi, N. 2022, A&A, 667, A161, doi: 10.1051/0004-6361/202244158

Tanimura, H., Aghanim, N., Kolodzig, A., Douspis, M., & Malavasi, N. 2020b, A&A, 643, L2, doi: 10.1051/0004-6361/202038521

The EAGLE team. 2017, arXiv e-prints, arXiv:1706.09899, doi: 10.48550/arXiv.1706.09899

Tramonte, D., Ma, Y.-Z., Li, Y.-C., & Staveley-Smith, L. 2019, MNRAS, 489, 385, doi: 10.1093/mnras/stz2146

Vernstrom, T., Heald, G., Vazza, F., et al. 2021, MNRAS, 505, 4178, doi: 10.1093/mnras/stab1301

Vernstrom, T., West, J., Vazza, F., et al. 2023, Science Advances, 9, eade7233, doi: 10.1126/sciadv.ade7233

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Yang, T., Hudson, M. J., & Afshordi, N. 2020, MNRAS, 498, 3158, doi: 10.1093/mnras/staa2547