Volumetric Reconstruction for Interactive Analysis of the Cosmic Web

Joseph N. Burchett\textsuperscript{1,2}, David Abramov\textsuperscript{2}, Oskar Elek\textsuperscript{2}, Angus G. Forbes\textsuperscript{2}
\textsuperscript{1}New Mexico State University, \textsuperscript{2}University of California, Santa Cruz
burchett@ucolick.org, \{olek, dabramov, angus\}@ucsc.edu

\textbf{ABSTRACT}
We present a multi-faceted study of the Illustris TNG-100 dataset at redshift $z = 0$, utilizing two interactive visual analysis tools: \textit{Polyphorm}, which utilizes our Monte Carlo Physarum Machine (MCPM) algorithm to identify large scale structures of the Cosmic Web such as filaments and knots, and \textit{CosmoVis}, which provides an interactive 3D visualization of gas, dark matter and stars in the intergalactic and circumgalactic medium (IGM/CGM). We provide a complex view of salient cosmological structures within TNG-100 and their relationships to the galaxies within them. From our MCPM Cosmic Web reconstruction using the positions and masses of TNG-100 halos, we quantify the environmental density of each galaxy in a manner sensitive to the filamentary structure. We demonstrate that at fixed galaxy mass, star formation is increasingly quenched in progressively denser environments. With \textit{CosmoVis}, we reveal the temperature and density structure of the Cosmic Web, clearly demonstrating that more densely populated filaments and sheets are permeated with hotter IGM material, having $T > 10^{5.5}$ K, than more putative filaments ($T \sim 10^4$ K), bearing the imprint of feedback processes from the galaxies within. Additionally, we highlight a number of gas structures with interesting morphological or temperature/density characteristics for analysis with synthetic spectroscopy.

\section{1 \ INTRODUCTION}
The Cosmic Web represents the largest organizational scheme in the Universe and imprinted in its large-scale structure (LSS) is the cosmological history of the Universe. Embedded within the LSS, ecosystems of galaxies are actively forming and evolving, and in the process, accreting and expelling matter and channeling energy back into the system. Cosmological simulations are essential tools for expanding our theoretical understanding of the Universe. They universally predict networks of filaments, sheets, nodes, and voids, and modern simulations with hydrodynamics and galaxy formation physics also now yield realistic populations of galaxies that inhabit the Cosmic Web and the circumgalactic and intergalactic gas that permeates it. In the observed Universe, the LSS is readily apparent from the locations of spectroscopically measured galaxies. However, the underlying structure must be inferred from incomplete, partial tracers rather than mapping the LSS directly as it is seen in the simulations. Furthermore, as galaxies do not generally evolve in isolation but in ecosystems within the Cosmic Web, understanding the galaxy-Cosmic Web connection is paramount.

In this study, we draw from Illustris TNG, a state-of-the-art suite of cosmological simulations run in a variety of sizes and resolutions. The data products include the 3D distributions of physical properties including the density, temperature, metallicity, and magnetic fields of the gas throughout the volume, as well as various properties of the stars and dark matter distribution. We focus on Illustris TNG-100, a (100 Mpc)$^3$ simulation volume at redshift $z = 0$, i.e., the simulated equivalent of the present-day Universe. Due to the extraordinary data volume of the Illustris TNG simulation (billions of simulated mass quanta, TBs of data), as well as its highly multimodal nature, many different pathways can be taken towards understanding the data. Our approach uses interactive visualization methods and represents structures at multiple relevant scales.

Our solution is implemented in two separate software tools: \textit{Polyphorm} \cite{burchett2020,polyphorm_code}, designed to reconstruct and visualize large-scale structures, and \textit{CosmoVis}, designed for detailed inspection and analysis of intergalactic and galactic clusters and their surrounding medium. The specific tasks we investigate are as follows:

\begin{itemize}
\item \textbf{T1} Reconstruct and visualize the filaments, knots and voids of the Cosmic Web as a function of the positions and masses of the dark matter halos contained in the dataset.
\item \textbf{T2} Investigate the orientation statistics of the Cosmic Web filaments, which is a direct precursor to their clustering and classification.
\item \textbf{T3} Visualize the star formation rates for the halos superimposed on the Cosmic Web structures, and identify relations between quenching/star formation and the respective halo positions within the large-scale structure.
\item \textbf{T4} Quantify the local density field of galaxies in the Cosmic Web.\n\item \textbf{T5} Investigate the properties and formations of the gaseous medium surrounding the halos, especially in relation with their position within the large structure.
\item \textbf{T6} Study the impact of galaxy feedback on the circumgalactic/intergalactic medium (CGM/IGM).
\item \textbf{T7} Explore the formation and distribution of the warm-hot intergalactic medium (WHIM), purported to contain the unaccounted for fraction of gas in the Universe, i.e., the “missing baryons”.
\end{itemize}

\textit{Polyphorm (T1-4)} is described in Sec. 2 and \textit{CosmoVis (T5-7)} in Sec. 3. The Appendix provides a gallery showcasing outputs of our software tools, and we include two videos as additional material.

\section{2 \ Polyphorm: INFERRING COSMIC WEB STRUCTURES}
\textit{Polyphorm} \cite{polyphorm_code} is an interactive visualization and reconstruction tool designed to provide novel insights into the spatial LSS distribution \cite{burchett2020, polyphorm_code}. The main problem \textit{Polyphorm} addresses is reconstructing the Cosmic Web structure from the distribution of galaxies (in the real Universe) and dark matter halos (in simulations) by acting as a meaningful structural interpolator. \textit{Polyphorm} is based on Monte Carlo Physarum Machine (MCPM), an agent-based computational model that mimics the growth and foraging behavior of \textit{Physarum polycephalum} “slime mould”. MCPM is a non-linear model based on a feedback loop between a swarm of discrete particle-like agents (representing the virtual Physarum) and a continuous field representing the concentration of a marker emitted by the data. Further details about the model and its behavior can be found in Elek el al. \cite{polyphorm_code}; most importantly for this study, the model outputs a network-like 3D density field, which serves as a proxy of the Cosmic Web and the main vehicle for the visualizations presented in this section.

\textbf{Data preparation:} To reconstruct the LSS within TNG-100, we input dark matter halos identified with the Subfind algorithm and provided on the Illustris data access site. In addition to the 3D halo coordinates within the simulation volume (‘SubhaloCM’), \textit{Polyphorm} uses the ‘SubhaloMass’ attribute as halo weights to create a flow of agents proportional to their masses. Our star formation analysis also uses the star formation rates (SFR; ‘SubhaloSFR’) and stellar masses (M$^*’$; ‘SubhaloStellarPhotometricsMassInRad’). The resulting 4.37M halos were fitted by \textit{Polyphorm} in 1.5 minutes on our development machine (running Windows 10 and equipped with an NVIDIA TitanX GPU), using 10M model agents in a 768$^3$ simulation lattice, with a sensing distance parameter of 2.55 Mpc.

\textbf{Task 1 – Reconstruct and visualize the Cosmic Web:} To visualize
different aspects of the Cosmic Web, Polyphorm implements several rendering modalities showcased in Figs. 1 and 5. These modalities run interactively, and are based on direct volume rendering using ray marching with specific transfer functions designed for each task.

- **Trace density.** Provides an overview of the reconstructed MCPM scalar density field (the “trace”), obtained by spatio-temporal averaging of the agent trajectories as they navigate the LSS, guided by the halos. The transfer function maps the density to a user-defined color ramp, with a configurable optical density to tune the apparent opacity of the visualized structures. This quantity further serves as a proxy of the local matter density, as employed by Burchett et al. [3] and Simha et al. [16].

- **Trace segmentation.** Uses two user-defined thresholds (see Burchett et al. [3] for examples of how these can be calibrated) which separate the LSS into three regimes: knots plus filament cores, filament outskirts, and voids. This modality allows for a quick visual gauge of the relative proportions of these constituents within the Cosmic Web.

- **Galaxy color and filament orientation.** These two modalities are designed to visualize the galaxy ‘color’ derived from the specific star formation rate, and estimate the filament directional orientation. Tasks 2-4 below provide more details.

Both the galaxy color and filament orientation modalities are newly developed to support this data challenge. In addition, we have now extended the functionality of Polyphorm to include a global illumination solver based on Monte Carlo volumetric path tracing (Fig. 4). This is intended for rendering high-fidelity images of the Cosmic Web reconstruction, emphasizing spatial relationships between the important components of the LSS. This path tracing modality provides interactive navigation and parameter tweaking, with progressively converging rendering when the user interaction stops. Depending on the configuration, rendering takes seconds (small datasets, millions of agents, and higher grid resolutions) to fully converge.

**Task 2 – Investigate filament orientation:** The MCPM model at the core of Polyphorm readily identifies filamentary structure in the data, although it does not directly separate individual filaments from one another to provide data products such as filament catalog. However, the model does provide a path forward via estimating the orientation of reconstructed Cosmic Web filaments. Because the agents in the volume “flow” along filaments, the agents’ propagation directions are closely aligned with the filaments’ orientations. The average absolute values of the agents’ propagation directions are visualized in Figs. 4 and 5 by color-coding the volume (with the XYZ directions being mapped to each of the RGB color channels). Locations in the LSS with omnidirectional flow of agents (most notably knots) will have no preferential orientation and be colored white, whereas locations with a single dominant direction of flow (filaments) will have a distinguishable color. Here, we use this information to visually highlight individual coherent filament structures.

**Task 3 – Visualize the galactic star formation rates:** Using an empirically determined threshold for the star formation rate (sSFR, see Task 4), we segment the input halos into star-forming (blue) and quenched (red) whenever the information is provided in the TNG catalog. This yields 63k “blue” and 66k “red” halos. We then superimpose the reconstructed filament map (represented by the trace) on the halos, to elucidate the systematic relationship between the star formation rate and the respective halo’s position in the LSS. In Fig. 5, we see that this is indeed the case: large knots universally contain red, quenched galaxies, while the less dense filaments that typically surround voids tend to host blue, star-forming galaxies. In Task 4 we examine this relationship quantitatively, as a function of the Cosmic Web environmental density.

**Task 4 – Quantify the local galactic density field:** We extracted the resulting trace density field from the fit, and for each subhalo with nonzero stellar mass, found the MCPM density value ($\rho_{\text{phys}}$) at the galaxy’s position and calculated the specific star formation rate $\text{sSFR} = \text{SFR} / M_*$. We then binned the sample by $\rho_{\text{phys}}$ and $M_*$ and calculated the fraction of galaxies falling below our threshold of $\text{sSFR} = -10.75$ separating quenched and star-forming galaxies. We set this threshold based on visual inspection of the sSFR-mass distribution to approximate the well known bimodality in the galaxy population [1]. Because truly quenched galaxies have $\text{sSFR} = 0$ in the catalog, we set these galaxies’ sSFR $= -12.5$ for our analysis. Fig. 6 shows that at fixed galaxy mass, the quenched fraction increases with increasing environmental density for galaxies with $\log M_* / M_\odot < 10.2$. This result highlights the presence of “environmental quenching” in Illustris TNG as shown observationally by Peng et al. [13]. Intriguingly, the density at which this effect sets in increases with galaxy mass until $\log M_* / M_\odot \leq 10.2$, which suggests that the more massive galaxies below this mass are progressively more resilient to environmental quenching processes than less massive galaxies.

### 3 CosmoVis: Analyzing the Gas Physics within the Cosmic Web and Galaxy Halos

*CosmoVis* is a web-based tool for interactive volume rendering of large-scale hydrodynamic cosmological simulation data. This software not only visualizes extracted gas, dark matter, and stellar properties from the simulation, but it enables one to analyze simulated data using observational techniques, specifically absorption line

**Figure 1:** Overview of the modalities provided by Polyphorm (Section 2). Trace density: scalar density field representing the Cosmic Web estimate, mapped to a blue-yellow color ramp. Trace segmentation: same as before, manually segmented into knots plus filament cores (red), filament outskirts (green) and voids (blue). Filament orientation: dominant XYZ direction of the filaments mapped to RGB colors. Galaxy color: spatial distribution of galaxies that are star-forming (blue) vs. quenched (red) with respect to their position in the LSS (white).
sponding physical gas properties of galaxies within different neighborhoods of the Cosmic Web. We save multiple resolutions of the data (64^3, 128^3, 256^3, 384^3, 512^3) to accommodate different computing environments. Generating the voxelized grid for each field in TNG-100 at each of our resolutions takes ~2 hours on our development machine.

**Application workflow:** A 3D volume representing gas temperature, gas density, star distribution, and, for the EAGLE simulation, dark matter distribution is visualized a few seconds after the web interface is opened, once the dataset is loaded. A user can toggle on/off the available particle types, and, for each type, the color scale and the transfer function can be adjusted. Multiple gas attributes are available for selection via a dropdown menu. In this study, we focus on the Temperature and Metallicity gas fields. By default, we set “cool” gas temperatures, $T \sim 10^4$ K, as blue and “hot” gas, $T \sim 10^7$ K, as red, with green in an intermediate channel. These choices were entirely physically motivated, as astrophysicists typically adopt a convention where “cool”, “warm-hot”, and “hot” refer to temperature regimes $T \sim 10^4$, $T \sim 10^6$–$10^7$, and $T \gtrsim 10^9$, respectively. These temperature regimes typically connote distinct physical mechanisms, although they may all reside within the same physical structures, hence the term multiphase is used to describe the complex structures in the IGM/CGM. To elucidate a galaxy’s morphology, the rendered sizes of star particles can be interactively adjusted with a slider. Additional sliders are available to adjust the exposure of the scene, as well as opacity modulation based on gas density, distance from the camera, and attribute value, which can aid visibility of certain structures, particularly in large simulation volumes. Users can also slice in the X, Y, and Z planes to visually drill down on a specific region of interest. CosmoVis also enables users to place “virtual skewers” within the simulation visualization, which requests a spectrum and waits for it to be generated and retrieved from the server backend (which runs in the cloud as an Amazon Web Service). Lastly, different simulations and multiple grid resolutions can be loaded upon request via dropdown menus.

Star particles are rendered off-screen to a depth buffer. Voxel grids containing gas field information and dark matter density are loaded as 3D data textures, and combined in a volume rendering integral along with the stars stored in the depth buffer. Uniform values that feed into the shaders, ranging from color to simulation size, are updated in response to user interactions. For the synthetic IGM/CGM absorption spectra, skewer endpoints are sent to the python server backend, which utilizes Trident [9] and yt [20] to generate the synthetic quasar spectrum and send it back to the user. These spectra are convolved with real instrumental response functions to closely emulate actual observed data; we currently implement the Cosmic Origins Spectrograph (COS) on the Hubble Space Telescope, which has provided the lion’s share of spectroscopic information about the CGM/IGM in the low-redshift Universe. Therefore, astronomers can connect observationally inferred outflow models and empirical data to those produced by cosmological hydrodynamic simulations. Due to the much larger simulation volume than our previously tested EAGLE 25 Mpc box, we did not fully employ the interactive, real-time spectral generation functionality for TNG-100 within CosmoVis, but instead use CosmoVis to identify beginning and endpoints for skewers, such as piercing a galaxy halo, and send these to the backend to generate the spectra. This data challenge has highlighted a key performance optimization benchmark for our ongoing development.

**Task 5 - Explore morphological and thermodynamic structure:**

CIGN/IGM gas in simulations is typically visualized in 2D projections along some axis wherein the temperature, density, or metallicity...
Task 6 - Study the impact of galaxy feedback on the CGM/IGM:

CosmoVis

is color coded and weighted, e.g., by density along the line of sight. It is also a common practice for simulators to render movies wherein they can represent time evolution and/or rotate and move the scene in 3 dimensions. Both approaches limit exploratory analysis: all 3D structure is lost in the former and real-time exploration is lost in the latter. The CosmoVis framework overcomes both by enabling the user to interactively explore the 3D space while modifying color mapping, etc., on the fly to highlight salient features of the data.

Immediately striking in these data when one zooms in on individual galaxy halos is the diversity in morphology of the CGM. While some gaseous halos appear rather spherical, most have some irregular structure highlighting the effects of feedback from supernovae and/or active galactic nuclei (see the top panel of Fig. 9 and Task 6). Furthermore, as evidenced by the color contrast between halos in various environments, these media also have an array of thermodynamic properties, where even galaxies of similar size are sometimes dominated by hotter temperature gas than others (Fig. 9, top panel). This is likely a combination of effects ranging from the mass of the dark matter halo to feedback history to large-scale environmental interaction. Either way, this readily discernible phenomenon enabled by CosmoVis underscores the importance of analyzing circumgalactic media and its connection to galaxy evolution in context with the larger scale environment [5, 14, 22].

Task 7 - Explore the distribution and formation of the WHIM:

The diffuse warm-hot intergalactic medium (WHIM), at $T \sim 10^5 - 10^6$ K, is purported to comprise a substantial fraction (30-40%) of baryonic matter in the Universe [6], but the ionization state of this material renders it difficult to detect. Simulations, including Illustris TNG [10] predict the WHIM to reside primarily within Cosmic Web filaments. From Figs. 7-11, one clearly sees the ubiquity of the WHIM, showing up as a yellowish hue, within massive filaments. While such warm-hot gas can also arise from thermodynamic effect within galaxy halos, the top panel of Fig. 9 showcases regions in massive filaments far away from galaxies. These regions likely form through shock heating during structure formation [18] and are key to identify in cosmological simulations to enable comparison with potential signatures of the WHIM via absorption line spectra. With CosmoVis, we can quickly identify such regions, place skewers probing them, and produce synthetic spectra with Trident for direct comparison with observed putative WHIM absorption signatures [17].

4 Conclusion

We have demonstrated two novel software tools, Polyphorm and CosmoVis, which have been used in a multifaceted exploration and analysis of the emergent structures in the Illustris TNG-100 magnetohydrodynamic cosmological simulation dataset. Using Polyphorm, we have reconstructed the matter density field over the full simulation volume and provided initial visual and quantitative evidence that ‘red’ (quenched) galaxies systematically lie in the nodes of the Cosmic Web and other high-density regions. In addition, Polyphorm produces a robust estimate of the filament orientation, which we plan to leverage in the future to measure the directional alignment between galaxies and filaments. Further, the orientation statistics appear to be consistent enough to serve as a marker for clustering analysis and the eventual possibility of building a robust filament catalog.

Though CosmoVis is still under development, it already enables us to discern subtleties in Cosmic Web and galaxy halo IGM/CGM structure. Our framework enables straightforward implementation of additional gas attributes into the visualization, such as magnetic field strength. We are currently working on incorporating queryable galaxy catalog information using the Subfind data products to enable selecting samples of galaxies based on, e.g., stellar mass and star formation rate allowing for further insights into the nature of galaxies with respect to the large scale gas structures in which they reside.

We have demonstrated in myriad ways using our tools that galaxies know about their place in the Cosmic Web and vice versa: the Cosmic Web knows about the galaxies within, via energy injection and metal enrichment. Well beyond the simple observation that the more massive galaxies inhabit the nodes, this work demonstrates the interdependence of star formation rate on galaxy mass and large-scale environment. With the combination of Polyphorm and CosmoVis, we are poised to enable revolutionary analyses of galactic ecosystems in the Cosmic Web.
REFERENCES


APPENDIX: IMAGE GALLERY

In the pages below, we include a gallery of additional images created with the Polyphorm and CosmoVis software applications, along with figure captions that explain their relation to the tasks described in the main text above. Two videos are also included with our data challenge submission, one showcasing an analysis session using Polyphorm and the other an analysis session using CosmoVis.
Figure 4: Monte Carlo path tracing of the Cosmic Web structures, depicting the full TNG-100 dataset (left) and a slice about 20% of the full volume thickness (right). Here we interpret the input halos (red) as sources of illumination, and the reconstructed large-scale structures (blue-yellow) as a volumetric medium which not only emits, but also absorbs and scatters light. The combined effect of these interactions gives the dataset a better sense of depth and provides cues about the scale of the Cosmic Web features.
Figure 5: Details of the TNG-100 dataset reconstructed by Polyphorm (Section 2). The visualization modalities and their configuration are identical with Fig. 1. We focus on three slices along the X-axis, selected to showcase different configurations of LSS structures – knots, filaments and voids. The slices represent 20% of the total volume thickness, except for the ‘trace segmentation’ which uses 5% of the total thickness.
Figure 6: Distribution of the fraction of quenched galaxies as a function of galaxy mass and local environmental density as quantified by Polyphorm/MCPM. This figure shows that at fixed galaxy mass, the quenched fraction increases with increasing environmental density for galaxies with $\log M^*/M_\odot \leq 10.2$. This result, enabled by interactive analysis sessions using Polyphorm, highlights the presence of "environmental quenching" in Illustris TNG. Intriguingly, the density at which this effect sets in increases with galaxy mass until $\log M^*/M_\odot \leq 10.2$, suggesting that the more massive galaxies, but below this threshold, are progressively more resilient to environmental quenching processes than less massive galaxies. See Sec. 2, Task 4 for more details.
Figure 7: The full TNG-100 volume visualized in CosmoVis, with gas Temperature (left) and Metallicity (right). The hottest regions appear red, followed by yellow intermediate, and blue as the coldest (left). Star particles are yellow. The same color coding scheme is applied to metallicity. Galaxy clusters and connective filaments are clearly visible throughout the volume from this perspective. See Sec. 3, Tasks 5-7 for more details.
Figure 8: Two slices from TNG-100 visualized in CosmoVis with skewers (highlighted in green for increased visibility) probing two distinct regions. On the left is a supercluster of galaxies at the intersection of multiple filaments, on the right is a galaxy in a sparsely populated region of the volume. Gas Temperature is shown (top), as well as gas Metallicity (bottom). Star particles are color coded yellow and randomly sampled in the data. See Sec. 3 for details. Note the high temperature in the densely probed region (top left) corresponding to depleted metallicity (bottom left), as compared to the isolated galaxy with a relatively intermediate temperature (top right) and surrounding elevated metallicity (bottom right). See Sec. 3, Tasks 5-7 for more details.
Figure 9: TNG-100 slice showcasing multiple dense clusters and their connectivity in CosmoVis, with gas Temperature (top) and Metallicity (bottom). Star particles in yellow. See Sec. 3 for details. While some gaseous halos appear relatively spherical, most have irregular structures indicating that there are feedback effects from supernovae and/or active galactic nuclei. The diffuse warm-hot intergalactic medium (WHIM) is visible as well. See Sec. 3, Tasks 5-7 for more details.
Figure 10: TNG-100 slice in CosmoVis. Gas Temperature (top), Metallicity (bottom). See Sec. 3, Tasks 5-7 for more details.
Figure 11: TNG-100 slice in *CosmoVis*. Gas Temperature (left), Metallicity (right). See Sec. 3, Tasks 5-7 for more details.

Figure 12: TNG-100 slice in *CosmoVis*. Gas Temperature (left), Metallicity (right). Star particles in yellow. See Sec. 3 for details.
Figure 13: TNG-100 slice in *CosmoVis*. Gas Temperature (left), Metallicity (right). Star particles in yellow. Cosmic Web filaments can be seen within both the gas and metallicity, although appears more defined in the blue on the left. See Sec. 3 for details.

Figure 14: TNG-100 slice in *CosmoVis*. Gas Temperature (left), Metallicity (right). Star particles in yellow. See Sec. 3 for details.
Figure 15: Sample spectrum generated from TNG-100 using Trident [9] and yt [20]. Wavelength vs Flux. See Sec. 3 for more details.