CosmoVis: Visualizing Hydrodynamic Cosmological Simulations at Galactic and Intergalactic Scales

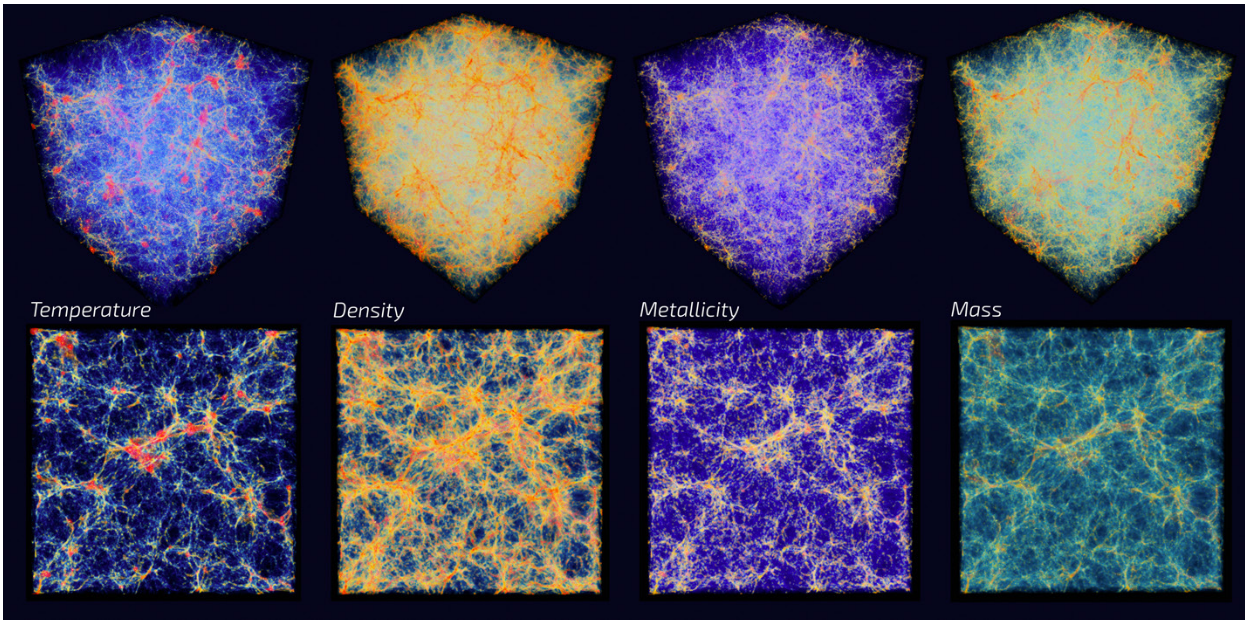

Fig. 1. This figure shows a series of interactive volume renders of gas fields created with the *CosmoVis* visualization tool. The data is extracted from the TNG100-1 z = 2.32 cosmological simulation snapshot. The top row depicts the entire 100 Mpc volume from an angle, and the bottom row shows a head-on perspective of a 15 Mpc thick slice of the volume.

Abstract—We introduce *CosmoVis*, an open-source web-based astrophysics visualization tool that facilitates the interactive analysis of large-scale hydrodynamic cosmological simulation datasets. *CosmoVis* enables astrophysicists as well as citizen scientists to share and explore these datasets, which are often comprised of complex, unwieldy data structures greater that 1 TB in size. Our tool visualizes a range of salient gas, dark matter, and stellar attributes extracted from the source simulations, and enables further analysis of the data using observational analogues, specifically absorption line spectroscopy. *CosmoVis* introduces novel analysis functionality through the use of "virtual skewers" that define a sightline through the volume to quickly obtain detailed diagnostics about the gaseous medium along the path of the skewer, including synthetic spectra that can be used to make direct comparisons with observational datasets. We identify the main analysis tasks that *CosmoVis* enables, and evaluate the software by presenting a series of contemporary scientific use cases that utilize *CosmoVis*. Additionally, we conduct a series of task-based interviews with astrophysicists indicating the usefulness of *CosmoVis* for a range of data analysis tasks.

Index Terms—AstroVis, cosmological simulations, astrophsyics, spectrography.

1 Introduction

Hydrodynamical simulations of galaxy formation have become essential tools in modern astrophysics over the past two decades. Thanks to these simulations, astronomers have gained unprecedented understanding of the structural and chemical composition of the Universe as well as the key processes driving galaxy formation and evolution. While the many variables in these highly complex simulations are tuned to match salient observable characteristics of the galaxy population in the Universe, many aspects of the simulated Universe result from these initial conditions and serve as pure theoretical predictions. In turn, simulations are critical tools for interpreting observations, both of the stars within galaxies and the gaseous environments they are bathed in.

The computational advances in generating these large-scale simulations over recent years has resulted in a double edged sword: on one side the high resolution and dimensionality are powerful and informative, and on the other, their size and complexity are prohibitive for most users to access these data. Even researchers with access to a powerful enough machine capable of loading and processing these simulations are nonetheless limited in the availability of tools for interactively exploring, inspecting, and analyzing these simulations. General purpose visualization tools, such as *ParaView* or *Blender* can render these data, but require data wrangling or scripting expertise, and their steep learning curve can overwhelm less computationally savvy users. Members of the general public are further isolated from the scientific inquiry enabled by these simulations.

Working in close collaboration with astrophysicists, we introduce *CosmoVis*, a web-based interactive visual analysis tool targeted toward

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

both simulation experts and the greater astronomical community, as well interested members of the general public. The *CosmoVis* software follows a client-server model where the visualization interface is presented to the user in the web browser, and analytics are handled by a *Flask* Python server that stores and manages the simulation data. *CosmoVis* simplifies the analysis workflow by providing access to multiple simulation databases, as well as providing a means to toggle on and off different simulation data layers, such as gas, dark matter and stars/subhalos. These fields are rendered using an optically based volume raymarching integral, with star particles embedded within the gas and dark matter media. Furthermore, users can interactively place "virtual skewers", or sightlines, within the simulation volume. These skewers provide 1D samples of the simulated datasets that are analogous to how observational astronomers probe the Universe, with spectra taken along sightlines to bright, distant quasars.

CosmoVis accomplishes three core requirements that were identified in collaboration with domain experts: (1) host existing large-scale hydrodynamical simulation datasets, (2) enable observers and theorists to easily produce synthetic spectra probing the diffuse circum- and intergalactic media within simulation volumes, and (3) enable on-the-fly analysis of these datasets. Through our iterative design process, we uncovered an additional practical use of the virtual skewers: superimposing elemental ion column densities and key physical properties such as density, entropy, metallicity, and temperature in different regions of the simulated data (see Section 3.5). These measurements can help answer open questions in astronomy such as how baryonic matter, particularly the heavy elements, are distributed through the universe, or how galaxies shape the gaseous environments around them.

This ability to interactively drop "virtual skewer" sightlines using straightforward point-and-click mouse interactions and retrieve analytic results is a main innovation of *CosmoVis*. These skewers enable users to sample the simulation data in ways that are both readily interpretable as well as directly comparable to empirical absorption line spectroscopy observations of the physical universe. Absorption line spectroscopy is an essential tool in observational astronomy for deriving physical properties of material within the line of site of a telescope and sensor. Instrumentation on the Hubble Space Telescope (HST) such as the Cosmic Origin Spectrograph (COS) can collect observational spectra, which are used to study the origin, formation, and evolution of solar systems, galaxies, and large scale structures of the Universe.

Since we present synthetic spectral data alongside the physical properties that give rise to these spectra, we provide a new avenue for utilizing simulations to address a range of open questions in astrophysics. What drives the evolution of galaxies? What quenches the cool gas supply off from massive galaxies leading to their decreased star formation and relatively quiescent state? How are galaxies fed star-formation fuel through cool and/or hot flows of gas from between galaxies? *CosmoVis* can help scientists address these questions, providing easy, intuitive methods for identifying and measuring spectral and spatial features across multiple simulations. Along synthetic skewers, physical column densities and quantities can be produced as well as synthetic spectra, which can be analyzed with the same tools that observational astronomers use. From here, astronomers can develop and test new hypotheses about the nature of the cosmos by interacting with the simulations and extrapolating their synthetic observations to the known universe.

2 Background & Related Work

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.

2.1 Hydrodynamical Simulations

Two main types of simulations are employed in studying the evolution of a galaxy in its larger cosmological context. *Large-volume cosmological simulations*, as the name suggests, spread their computational power out over a large computational domain, usually 50-500 megaparsecs (Mpcs) in size. These models are able to resolve hundreds of galaxies simultaneously, but at relatively coarse resolution. *Zoom-in simulations* focus on a smaller region, often a single galaxy, and can thus achieve significantly finer resolution in modeling its behavior, while still coarsely sampling the cosmological environment around it. *EAGLE* [36, 48] and *Illustris TNG* [43] are large-volume simulations, whereas *FIRE* [26] and *Tempest* [28] are zoom-in simulations. For the most part, state-of-the-art hydrodynamic simulations include as many "resolution elements" as the supercomputing infrastructure will allow for, about 20 billion resolution elements. The difference in simulations is primarily the volume over which these resolution elements are applied to: big boxes will have a coarser resolution, and smaller boxes will have a finer resolution. But also, big boxes have thousands of galaxies present in the simulation, whereas smaller boxes may only include one or two galaxies resolved in higher detail.

Different simulations suites employ different codes, each with their own distinct implementation of the physics present (e.g., hydrodynamics, gravity). Lagrangian codes use a technique called Smoothed Particle Hydrodynamics (SPH) [37] to represent gas parcels as zero-dimensional particles for the purposes of transport, then apply a 3D smoothing kernel to "smear" them into a finite space (e.g., *EAGLE*). Eulerian codes represent the simulation domain as a series of nested cartesian grid cells [4], and allow gas to travel between resolution elements (e.g., *Tempest*). Finally, hybrid models use tracer particles to flow with the fluid, but define non-Cartesian grid cells at each timestep generated by a voronoi tesselation [49] based on the particle locations (e.g., *TNG* and *FIRE*). At infinitely high resolution, all of these hydrodynamical methods converge to the same solution, but at finite resolution, they will have slight differences in how the hydrodynamic solution is approximated.

Aside from the hydrodynamics, the other primary difference in these simulations is the treatment of energetic feedback from supernovae and active galactic nuclei (AGN, i.e., supermassive blackholes). These two non-linear energy sources can have a profound effect on how the galaxy and its environment evolve with time. The finest spatial resolutions found in these simulations are parsecs to hundreds of parsecs, whereas the scales at which stars and black holes form and evolve are many orders of magnitude below this, thus stars and black holes cannot be modeled self-consistently. The solution is a "sub-grid" model, which provides parameterizations of both how stars and black holes form, age, and interact with their environments through exchange of mass, energy, radiation, and gas composition. Sub-grid models are based on analytic models for how stars and black holes behave as well as external higher-resolution computational simulations executed previously. Small discrepancies in the parameterization of these complex non-linear processes can result in significant differences in the outcomes of the galactic cosmological simulations. By comparing the behavior of galaxies in different simulations using different implementations and sub-grid models, theorists can converge on which galactic behaviors are real, and which are artifacts of a particular numerical implementation.

The largest of these simulations span a length of ~ 100 Mpc along each side and contain physical information about tens of thousands of galaxies. These models take years to develop and are executed on the most powerful supercomputers in the world. The initial conditions of the simulations describe the distribution of matter in the universe shortly after the Big Bang, when gravitational perturbations are still

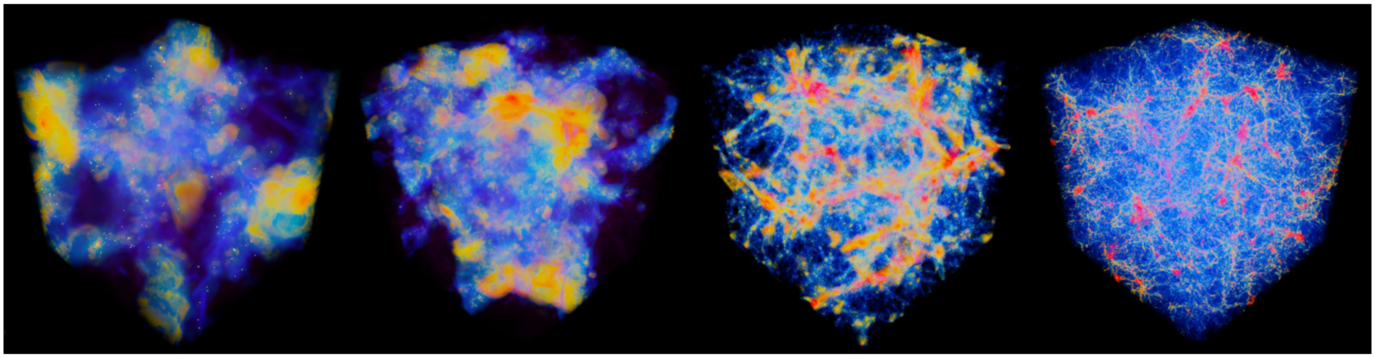

Fig. 2. 3D gas temperature distribution in several simulations. From left to right: 12 Mpc EAGLE z=0.0; 25 Mpc EAGLE z=0.0; 100 Mpc EAGLE z=0.0; 100 Mpc* TNG100-1 z=2.3.

linear. Time is represented in both linear terms of years and gigayears but also as an equivalent cosmological redshift z. This cosmological redshift [27] encodes the Doppler effect resulting from the cosmological expansion of the space between light emitted by a distant object at a particular epoch and absorbed by an observer at present [10].

2.2 Cosmological Absorption Line Spectroscopy

Cosmological absorption line spectroscopy is a powerful observational technique that is instrumental in uncovering the secrets of our Universe. Observational astronomers have identified a myriad of absorption profiles for specific element ions, where ions heavier than Helium are traditionally referred to as metals by the astrophysics community. Each of these ions has a characteristic absorption signature, which when measured can give an indication of how far away an emission source is by measuring its redshift, and the presence of intermediate material between the source and observe. These spectra can also provide information about the motions of gas or rotations of galaxies.

The spectral analysis component of *CosmoVis* is directly inspired by techniques in observational astronomy with an aim to bridge the gap between our knowledge of our known Universe and a simulation we are examining. The juxtaposition of "skewers", i.e., quasar sightlines, within a volume of galaxies was introduced by Burchett et al. [12] in a visualization tool called *IGM-Vis*, which facilitated visual comparison of measurements within observed cosmic data collected by the Hubble Space Telescope (HST) during sky surveys, such as the Sloan Digital Sky Survey (SDSS). This tool inspires the concept of the "virtual skewer" as a means to visually represent linear cosmological measurements embedded within a 3D volume showing the spatial relationship between galaxies using empirically observed data. *CosmoVis* extends this visualization approach to be applicable in creating new inquiries within simulations. This is made possible by *Trident*, a Python package that generates synthetic measurements using simulation data.

2.3 Simulation Visualization Tools

A variety of techniques have been employed to visualize simulations. For example, Woodring et al. [55] introduce new analysis features implemented within *ParaView* [1,2] to analyze simulation data. The Open Space "astrographics" system [6] facilitates the interactive display of integrated data from multiple sources, including simulation datasets, supporting research and science communication. Python package *yt* includes both a script-based and interactive volume renderer for displaying grid-based simulations (rather than particle-based or moving mesh datasets). An interactive tool called *FIREFly* [21] supports the rendering of particles from *FIRE* datasets generated from the *GIZMO* code [25]. The software *AstroBlend* [38] makes use of both *Blender* and *yt*, and can be used to perform analysis and visualization through scripting. *Polyphorm* [19] uses an unconventional nature-inspired approach to interactively reconstruct and visualize large-scale cosmic web density fields from either simulated halos or observed galaxy data, but does not handle additional attributes such as gas temperature.

The *Hardware/Hybrid Accelerated Cosmology Code* (or *HACC*) framework [23] includes models of baryonic matter as well as active galactic nuclei associated with violent bursts of energy from supermassive black holes, and has been used to illustrate how this energy is imparted to the surrounding gas and affects subsequent structure formation. Several submissions to the 2019 *IEEE Scientific Visualization Contest*[1] explored a 64 Mpc box *HACC* simulation. Nguyen et al. [40] used *ParaView* to visualize data from these simulations as point cloud renderings, and observed the changes over multiple time steps. Fritschi et al. [20] also visualized the *HACC* data using a combination of the open source tools *Visualization Tool Kit (VTK)* and *Qt* to run their graphic pipeline. They also chose to render the particles in the data as discrete points. Hesse-Edenfeld et al. [24] used a linked-multiview approach when building their custom visualization application, whereby interactive filtering on a plot automatically updates the rendering. Instead of rendering the particles as points, they implemented a direct volume rendering approach, and found that 256^3 Cartesian-grid based voxels provided sufficient spatial resolution while maintaining high frame rates. Schatz et al. [47] used a molecular visualization tool called *MegaMol* to create a series of linked visualizations, including parallel coordinate plots, 3D volumes, and isosurfaces. Despite these existing tools, our astrophysicist collaborators identified a clear need for interactive visualization software to effectively render large simulation datasets and to support a range of simulation analysis tasks.

3 Analysis Tasks

As we describe below, the analysis tasks facilitated by *CosmoVis* enable astronomers to investigate open questions about our Universe. What drives the evolution of galaxies? What quenches the cool gas supply off from massive galaxies leading to their decreased star formation and relatively quiescent state? How are galaxies fed star-formation fuel through cool and/or hot flows of gas from between galaxies? There are some salient features that are more readily measurable in a simulation than in the real universe, and there is still much to be uncovered about the nature of our Universe.

Through studying the results of different simulations, scientists can make more informed decisions when planning where to collect new observations with their telescopes. For example, astronomers can identify systems with clear signatures of outflows, or "bubbles", blown out from energized galaxies; they can find walls or sheet-like structures, which are complex regions where multiple filaments coming together, or they can find hot regions of the intergalactic medium within the cosmic web. That is, by consulting simulations, where it may be easier to identify and isolate a particular structure as compared to part of the real universe, new identifiable characteristics can be tested and confirmed with observational data. *CosmoVis* enables users to establish connections between the different gas and dark matter properties, to discover relationships between local galactic regions and large scale filamentary structures, and to more easily identify, extract, and interpret synthetic results that can inform our observations of the Universe.

[1] https://wordpress.cels.anl.gov/2019-scivis-contest/

3.1 Task 1: Access diverse simulation datasets

CosmoVis enables multiple cosmological simulation datasets to be viewed on demand, effectively homogenizing differences in the underlying data structures in simulations produced by different teams of collaborators. Modern cosmological simulations are large enough to contain tens of thousands of galaxies comparable to our own Milky Way. Our preprocessing pipeline (see Sec. 4.1) leverages existing data wrangling tools and introduces a custom visualization architecture so that physical quantities across simulations are accessible and comparable. Thus far, we have tested *CosmoVis* on *EAGLE*, *Illustris TNG*, and *FIRE* datasets, and users can ingest additional datasets as needed. Four simulations we have visualized are highlighted in Fig. 2.

Specifically, we make use of functionality available in the *yt* [53] and *Trident* [29] software packages to standardize the data outputs for visualization and analysis. *yt* enables relevant simulation data to be identified, gathered, and stored using a range different formats (e.g., *Enzo*, *GIZMO*, *Gadget*, and *FITS*), each of which can be loaded and processed using similar scripts. *Gas* and *dark matter* fields within these simulations are extracted as voxelized grids, and star particles are saved as a series of discrete points. A list of example particle types and fields found in cosmological simulations can be found in Fig 1. This backend preprocessing step only needs to occur once per simulation dataset, making the data available for the user to access via the frontend client application.

Multiple resolutions of volume data are available for visualization in order to accommodate end user network bandwidth and graphical capabilities (implementation details are discussed below). The interactive visualization software runs in the web browser, retrieving data from the backend AWS server. *CosmoVis* does not require any installation on the end user's machine, although expert users are welcome to set up a local webserver if desired.

3.2 Task 2: Identify structures in the IGM, CGM, & ICM

Cosmological simulations contain features at the circum- and intragalactic medium scales (CGM & IGM) as well as within the intracluster medium (ICM). In *CosmoVis*, galaxy morphologies are visually discernible as they span across a few kiloparsec (kpc) diameter each. Some examples include: Elliptical, spiral, spherical (S0), and irregular galaxies. This is accomplished by embedding star particles as individual points within a gas and dark matter volume, which becomes optically occluded based on the physical hydrogen number density. The relative spatial positions of the star particles is sufficient to describe the shape of the galaxy, and the density-modulated optical thickness provides effective depth cues.

Galaxy clusters, where galaxies collide and coalesce into larger scale structures, can also be easily identified. These denser regions may exhibit different properties when compared to galaxies located in sparse areas. On an even larger scale, *CosmoVis* enables users to see the emergent filamentary structure of the cosmic web. This network of matter and dark matter interweaves the galactic material and traverses across the entire volume. These structures are made apparent through attenuating the optical thickness in the volume raycasting integral based on the local physical density. In addition, we use different color maps to differentiate between gas and dark matter fields in the volume. By default, the gas color scheme ranges from blue (low), green (medium), and red (high) to correspond to different attribute data. Dark matter is represented by a dark purple to magenta color gradient to visually indicate low to high density regions. Each of the color maps used in *CosmoVis* can be customized by the user.

3.3 Task 3: Home in and focus upon regions of interest

CosmoVis enables users to navigate through and filter the simulated volume in order to focus on smaller subregions on demand. By rotating, zooming, and/or slicing the volume, a range of different cosmological structures can become more apparent. Low-density bubbles, accretion shocks, temperature variations, and remnants of galactic superwinds (big bubbles or jets) can be observed by toggling on various gas properties on or off. When the 'Temperature' field is enabled, hot regions of the intergalactic medium within the cosmic web can also be identified.

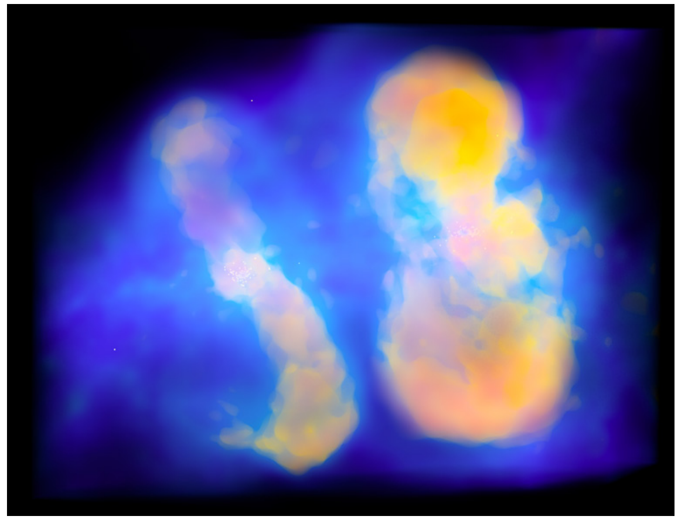

Fig. 3. Zoom in on a region containing two distinct circumgalactic environments in the 12 Mpc EAGLE simulation, with color representing temperature. Star particles are represented here as faint yellow disks at the center of these two adjacent systems. There appear to be visible galactic outflows of hot gas above and below the nuclei, indicated by the yellow-orange coloration.

Walls or sheet-like structures can also be discovered using *CosmoVis*. These complexes exist where multiple dark matter filaments come together and contain a range of hydrodynamic processes. To analyze these structures, simulations are sometimes re-run to focus on smaller subsections to achieve higher resolutions. However, it can be a time-consuming and cumbersome process to find and isolate a region of interest in this manner. *CosmoVis* simplifies the task of isolating regions of interest by allowing real-time interactive filtering of the volume. We present more details about using *CosmoVis* to investigate sheet-like structures in our first science use case (Sec. 5.2).

3.4 Task 4: Compare physical properties among matter components

CosmoVis facilitates a holistic overview of the primary components of the simulation, namely the *gas*, *dark matter*, and *stars*. Users can toggle the visibility of each of these while our rendering pipeline enables each of these types of fields to be visualized simultaneously (stars as particles, gas and dark matter as volumetric fields; Sec. 4.2). In this way, relationships between galaxies and the properties of their surrounding environments can be explored.

In addition, *CosmoVis* enables users to switch between different *gas* attributes to display in the volume rendering and customize the visualization quantitatively. We have incorporated several one-dimensional scalar fields into this framework, including temperature, density, metallicity, and element abundances such as hydrogen, carbon and oxygen. A science use case comparing metallicity in the CGM and IGM can be found in Sec. 5.4.

The transfer function controlling the color map for each gas field can be interactively adjusted, and a range of values can be specified to filter the data. This functionality enables myriad possibilities of exploratory analysis and hypothesis testing directly from the visualization alone. For example, assuming one is visualizing the temperature and wishes to focus on the so-called warm-hot regime, which is defined as $10^5 - 10^6$ K and is of particular interest in a number of contexts, the user can input these temperatures as the clipped minimum and maximum values for the color mapping and only the regions with these temperatures will appear in the visualization. Then, the user can identify interesting regions and switch the visualized quantity to density to assess whether this warm-hot region of the gas volume is also density-enhanced. Because we render the star particles, one can immediately check whether galaxies are nearby and therefore could be responsible for heating the gas to these temperatures.

In principle, the attributes available for analysis is limited only

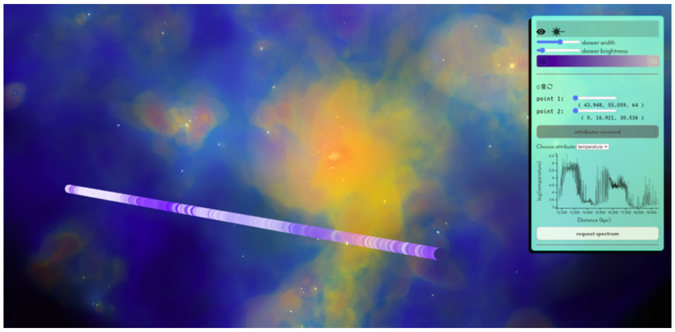

Fig. 4. *Temperature* along a virtual skewer. Using the default this color map, purple indicates a low temperature and gray indicates a high temperature. These temperature values along the skewer are also displayed in the line plot in the *Observe* panel on the right, where the user can also adjust the skewer position and define a custom color map.

by those provided by the simulation; *CosmoVis* is designed with the flexibility such that any attribute provided in a simulation dataset or derivable from those provided can be imported and visualized in such a customizable manner. A summary of fields that are included in the *EAGLE* and *TNG* simulation suites can be found in Table 1.

3.5 Task 5: Place virtual skewers

A main innovation in *CosmoVis* is providing researchers with the ability to interactively probe the simulation volume using virtual skewers in order quantitatively characterize its gas properties. With the "drop skewer" mode activated, a user can place these sightlines anywhere within the volume on demand. By default, the skewers' orientation is determined by the current camera view, pointing into the screen. They automatically clip to either the full volume or to the boundaries of the current volume subset, if the user has zoomed in to a region of interest (T3), and their endpoints can be interactively fine tuned after initial placement. Using these skewers, a researcher can request multiple types of data products from the *CosmoVis* backend server. These are retrieved using the *Trident* package, and include: 1) physical quantities along the length of the skewer line of site such as temperature, density, entropy, and metallicity; 2) column densities of several ion species (e.g., H I, O VI, and many others) along the length of the skewer; and 3) synthetic spectra associated with the skewer.

Column densities are a primary measurable quantity from observational datasets. The column densities represent the abundances of different elemental ions (elements heavier than helium are referred to as 'metals' in the astronomy community) along the length skewer, and are mapped using physical distance units in kiloparsecs (kpc). The physical quantities (temperature, metallicity, etc.) are also mapped the same way. Each column density or physical property can be plotted on a graph within the skewer panel, which updates the color banding along the skewer in the 3D volume. These different fields can be selected using a dropdown menu in the skewer panel once the data has been generated. An example skewer sightline with the physical property temperature is shown in Fig. 4. As gas temperature is generally not directly observable, cosmological simulations are useful for diagnosing the physical conditions an observed parcel of gas may possess. The plots of the physical attributes and column densities along the skewer are generated quickly, are easily interpretable, and provide a type of measurement that would be difficult if not impossible to measure in our physical universe using traditional instrumentation.

CosmoVis can also generate synthetic absorption line spectroscopy plots, or spectra, on demand via the user placed virtual skewers (see Sec. 2.2). These spectra are analogous to the data generated by instruments on the Hubble Space Telescope, and make it possible for users to compare the simulated universe to observations of the actual universe. These spectra encode the gas motions as well as the amount and temperature of the gas, but these need to be extracted with further analysis. This is why it is useful to include the computed column densities separately, which can provide extra information that is not immediately observable in the spectra.

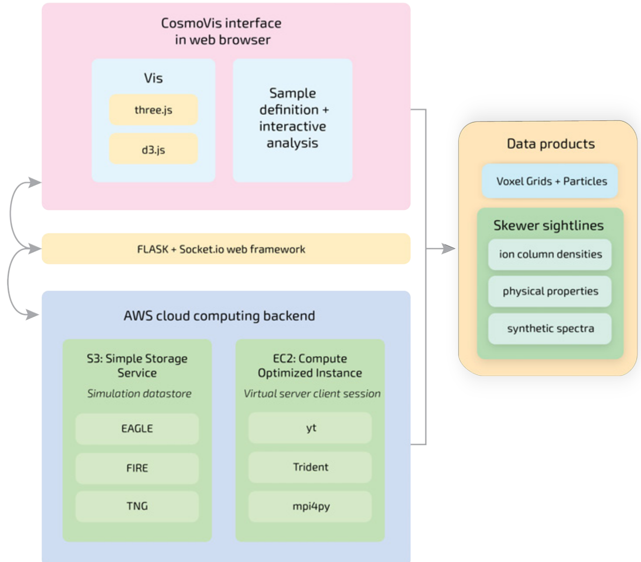

Fig. 5. Architecture of the *CosmoVis* software. The *CosmoVis* interface is accessible via modern web browsers, where the visualization is powered by the libraries D3.js for loading and displaying graphs, and THREE.js for rendering. The interface allows for interactive exploration and on-demand analysis via the skewers. Requests are sent via Socket.IO between the client frontend and the server backend hosted on an EC2 instance on Amazon Web Server. Metadata about the simulation, skewer placement and desired data product are sent to the backend. On the Python server running Flask, the requested simulation snapshot (e.g., EAGLE or TNG) is loaded using the *yt* package. Ion column densities, physical properties, and synthetic spectra are processed by *Trident*. These are sent back to the frontend where plots are dynamically generated by D3.js.

The synthetic data results, including both column densities and spectra, generated with the skewers are presented as two-dimensional line plots within the *CosmoVis* interface. There are a number of fields available to visualize as column densities, which update a plot as they are sifted through (also updating the banding of the skewer itself within the 3D volume). The synthetic spectra are also examined and compared within a separate panel. The spectra are aligned by either wavelength or velocity along the x-axis, with a subdomain being selectable using a slider. Users can also localize on different wavelength regions, such as those around the spectral lines of neutral hydrogen (H I), to see how their absorption compares across skewers. The synthetic spectra can also be downloaded as a FITS file format from the *CosmoVis* interface, which can be further analysed using traditional observational analysis scripts. Further, the 3D volume renders themselves illustrate interesting features of cosmological structures, and can be used to share results or for inclusion in publications.

4 Application Design of *CosmoVis*

CosmoVis has been developed through an iterative design process informed by continuous feedback from researchers in both the observational astronomy and simulation communities. Our initial design was based on two main goals: 1) create an interactive volume rendering solution to visualize a range of cosmological simulation datasets, and 2) enable the placement of virtual skewer sightlines through the simulated volumes and use them to generate synthetic spectra. As we became more familiar with the needs of these research communities and the simulation datasets themselves, our design goals coalesced into supporting each of the analysis tasks (Sec. 3), some of which also arose from specific scientific use-cases (Sec. 5). In this section we detail the key components of the *CosmoVis* system. The client-server architecture is summarized in Fig. 5.

Particle Types	Example Fields
Gas	Cooling Rate, Coordinates, Density, Electron Abundance, Element Abundance, Energy Dissipation, Entropy, Gravitational Potential Energy, Halo Mass, Internal Energy, Mass, Magnetic Field, Metallicity, Radiation, Star Formation Rate, Temperature, Velocity
Dark Matter	Coordinates, Density, Gravitational Potential Energy, Velocity
Stars	Birth Position, Birth Velocity, Coordinates, Density, Gravitational Potential Energy, Initial Mass, Mass, Metallicity, Metals, Star Formation Time, Velocity
Black Holes	Coordinates, Cumulative Mass, Density, Gravitational Potential, Magnetic Pressure, Mass, Mass Accretion Rate, Mergers, Pressure, Sound Speed Thermal Energy, Velocities

Table 1. Example particle types and fields in cosmological simulation snapshots. The datasets are usually broken down into different particle types (gas, dark matter, stars, and black holes), each expressing a variety of physical quantity fields.

4.1 Data Preprocessing

The data from most cosmological simulations (e.g., EAGLE or TNG) are distributed in the form of 'tracer' particle clouds. This particle representation mirrors the prevalent simulation methodology: smoothed particle hydrodynamics. However, these tracers in fact represent discretized field data (gas or dark matter density, gas temperature and pressure) or agglomerations of discrete macroscopic objects (stars or black holes). Even though such representation trivially lends itself to particle-based visualization, the sizes of these massive datasets (typically containing billions of tracers) quickly becomes unmanageable, especially when it comes to their transfer and rendering. We therefore opted for a hybrid approach: converting those tracers that inherently represent field quantities (gas and dark matter attributes) into uniform voxel grids, while retaining the discrete agglomerates in their original particle form.

To perform the conversion we rely on the *yt* Python package. The optimal sampling rates for the voxelized grids depend on the size of the simulated domain, number of contained tracers, as well as the actual details resolved by the simulation. Rather than attempting to determine this automatically, we preprocess several grid resolutions (between 64^3 and 512^3) and let the user choose the most appropriate one in runtime. To further optimize the storage, transfer, and rendering, we compress the original Float64 voxels to UInt8 and unpack them during the rendering stage.

4.2 Interactive Rendering

To visualize the hybrid field+particle data (Sec. 4.1), *CosmoVis* uses three rendering passes: a star particle pass, a skewer sightline pass (see Sec. 4.3), and a volume ray marching pass using the standard emission-absorption model [35]. In the first two passes, the star particles and skewers are rendered separately and saved to off-screen position and color buffers. In the third pass, a three-channel 3D texture containing the values of each gas, dark matter, and density voxel is integrated through by a physically based ray-marching shader. Here, the optical thickness of the gas and dark matter media are attenuated by the local hydrogen number density, making denser regions appear optically thicker. The depth buffers from the particle and skewer passes are used as early stopping criteria for the integration, which allows for consistent compositing of the volume, particle, and surface colors (Fig. 7).

In this design, the volume is effective in visualizing the field quantities: dark matter density, gas density, temperature, metallicity and other attributes, all with potentially custom transfer functions. Notably, we use two distinct color transfer functions to differentiate between the baryonic matter (gas) and dark matter filaments. The discrete star macro-particles then give sufficient cues about the shape and orientation of galaxies. On top of volumes up to 512^3 (Sec. 4.1), we can render in the order of 10^5 particles and dozens of active probing skewers.

Another design premise of *CosmoVis* is to provide the user with an unhindered, real-time analysis of the visualized simulation dataset. To this end we implemented standard modes of interaction: zooming, rotating, and panning directly in the visualization canvas. Users can also filter and slice through the volume using sliders in the *Data Selection* panel, interactively fine-tune the volume transfer functions within the *Layers* panel in the user interface, and place virtual skewers throughout the volume in the *Observe* panel (Sec. 4.3). This additional functionality helps to better localize and probe regions of interest in the data (Fig. 7).

4.3 User Interface

The design of the user interface minimizes visual clutter while maintaining easy accessibility for core functionality. The full width and height of the web browser window is used as the interactive 3D visualization canvas. A fully adjustable color transfer function is applied to the selected gas attribute (blue, green, red by default), dark matter (purple), and star particles (yellow). Hovering over star particles provides tabular information about the corresponding sub-halo.

The UI is controlled through a set of labeled pop-up panels on the right side of the rendering canvas. The *Data Selection* panel provides dropdown lists of available simulations and their preprocessed volume resolutions for gas attributes and dark matter, (enabling Task 1, Sec. 3.1) as well as sliders for interactively slicing the volume along the X, Y, and Z axes (enabling Task 3, Sec. 3.3). A grid overlay with 1-Mpc spacing can also be toggled via this panel.

An array of toggles in the *Layers* panel show or hide the various data fields in the main rendering window. Selecting *Gas*, *Dark Matter* or *Stars* fields expands a menu for fine-tuning each field. For example, for *Gas* the user can select the available attributes such as temperature, entropy, carbon, oxygen content, or metallicity. The user can also tune the maximum and minimum boundaries for controlling the 3-channel color transfer function, which can be customized with interactive color pickers and sliders controlling the density-modulated volume optical thickness. Similarly, the user can tune the *Dark Matter* density range and color transfer function. The *Star* menu allows for adjusting the size of the stars to accommodate different screen resolutions. Under *Visualization Options*, the user can adjust the strength of the hydrogen gas number density modulation (which controls the optical thickness of the gas and dark matter volumes), value modulation (which adjusts the optical thickness based on the magnitude of the active attribute), as well as overall exposure of the scene.

The *Observe* panel provides functionality for placing skewers, requesting synthetic data products, and viewing data plots associated with the skewers. When clicking on the simulated universe in the canvas, skewers are placed along the camera axis, with endpoints automatically clipped to the active volume boundaries. After placement, the *Observe* panel provides controls for the skewer's spatial extent, as well as options for generating synthetic data products: collecting column densities along the skewer for a wide variety of lyman-alpha absorption elements, as well as fields such as temperature, metallicity, and entropy. Once received by the client, the user can switch between the different synthesized data fields, updating the graph in the *Observe* panel, as well as the banding rendered along the skewer within the visualization. The second data product that can be generated are synthetic spectra, which are analogous to physical observations captured by sensors on the Hubble Space Telescope.

The *Spectra* panel is populated with processed synthetic data generated through user interactions via the *Observe* panel, which can be cross-compared by centering on specific wavelengths and linked brushing. Once a spectrum is computed by the server, a corresponding new plot is displayed. Each spectrum is aligned such that when selecting a specific spectral feature from the dropdown menu (such as the hydrogen Lyman-alpha line), switching between unit spaces (either wavelength or velocity space), or using the brush slider, each spectrum can then be visually compared against one another vertically. The panel finally allows to export all generated spectra as a FITS file to be used for further analysis.

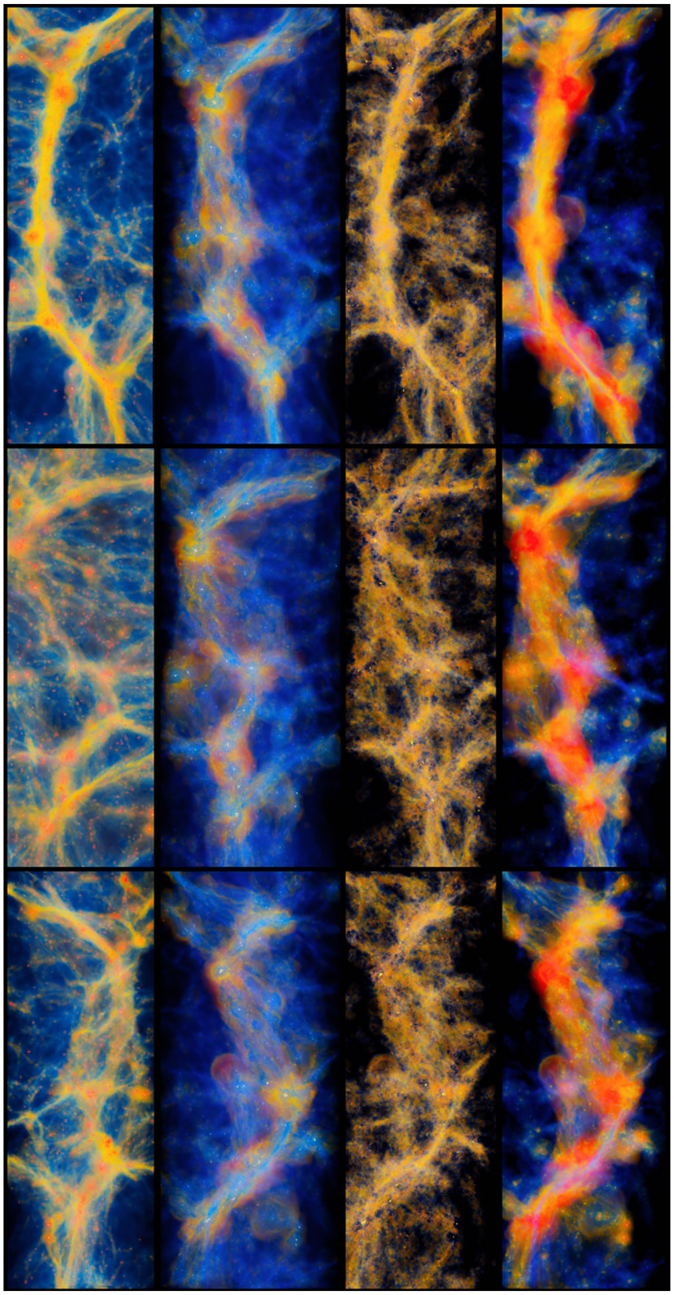

Fig. 6. Example of a cosmic sheet found in the 100 Mpc EAGLE simulation shown from three different angles arranged top to bottom. Gas fields from left to right: density, entropy, metallicity, and temperature. See Use Case 1, Sec. 5.1 for more information about these cosmic structures.

5 Scientific Use Cases

5.1 Case 1: Identifying sheet regions

A fundamental design motivation for *CosmoVis* was to provide an ability to easily explore the large volumes of cosmological hydrodynamical simulations to identify interesting cosmic structures [32]. These could be structures that are noted as interesting from a general, free-form exploration of the simulated Universe or a targeted search for the structures with predetermined characteristics that could then be analyzed in greater detail. Here, we use *CosmoVis* to identify some cosmic sheets that can be followed up with simulations at zoom-in resolution. We will also conduct some initial quantitative analysis on the sheet regions.

When one initially loads *CosmoVis*, the cosmic web structure of the Universe is readily apparent from the filamentary structures that contain the large majority of galaxies and intergalactic gas [8, 13, 16]. A somewhat closer inspection reveals that the not all of these structures are equal in their size and shape. Certain filaments are rather sparse with only a few galaxies and may appear to be merely offshoots of larger, more robust filaments that have much richer galaxy contents [18]. Furthermore, the initial visualization modality wherein the gas temperature is color coded reveals that the gas physical conditions can vary widely from depending on one's location in the cosmic web, such as at the intersections of filaments (nodes).

Certain regions appear as plane-like structures where multiple filaments converge or appear to have formed from a larger coherent structure. Such a 'cosmic sheet' is shown in Fig. 6, and these regions have attracted heightened recent attention. Cosmic sheets are laboratories for studying complex hydrodynamical processes that may produce nearly metal-free gas clouds in the intergalactic medium [34]. The relevant hydrodynamical processes, however, require resimulating the smaller regions of interest (a subset of a larger simulation volume) at increased resolutions not afforded for a large-volume cosmological simulation. Such workflow is common practice in the simulator community where a 'large box' simulation is run first at a fiducial (lower) resolution and certain regions, e.g., the immediate surroundings a particular galaxy, are rerun in a 'zoom-in' simulation at higher resolution where the initial conditions are provided by the original simulation at the fiducial resolution [22]. However, identifying regions of interest within the large boxes is generally a difficult, inefficient process. *CosmoVis* offers a transformative improvement to this workflow.

For this study, we employ the EAGLE 100 Mpc box visualized at 256^3 resolution (both options found in the 'data selection' menu) [48]. We kept the gas visualization modality with the temperature attribute. Larger simulated volumes by their nature contain greater numbers of the rarer, more massive structures, hence our choice of the largest volume. Sheets are plane-like structures with two primary dimensions, and are oriented in arbitrary direction. We proceed by using the slicing feature within the 'data selection' interface and narrowing the thickness of the volume visualized in the x-direction to $\sim$0.2 times the full volume (physically corresponding to 20 Mpc widths). The 3D rendering and interactivity of *CosmoVis* is critical for identifying sheets, as this slice of the simulated volume is inspected for sheet candidates by rotating it in several directions. Sheets are confirmed by finding (interactively) a camera orientation parallel to the structure. From this angle, the structure appears to collapse to one dimension. We log identified sheets by hovering over galaxies within the sheet and recording the approximate coordinates of the structure from the 'star particle details' window. We then view the next 20 Mpc slice of the simulation volume, inspect for sheet candidates, and proceed as above through the remaining volume in the x-direction. We then expand the slice to include the entire x-dimension and narrow the range to 20 Mpc in the y-direction, continuing the search as we did in the x-direction.

In all, we identified > 20 sheet candidates in the EAGLE 100 Mpc volume. Through the process of identifying sheets, we observed (from the colorization) that certain sheets seemed to contain much more high temperature gas $T > 10^5$ K than others. As a preliminary investigation of the temperature variation from sheet to sheet within individual sheets, we selected two sheets: one predominately filled with high temperature gas (orange-red in color) and another with predominately cool gas (blue-green in color). We oriented each sheet with the camera angle perpendicular its plane and placed three skewers through each using the functionality under the *Observe* menu (see Task 5, Sec. 3.5). Using the 'request skewer attributes', we inspected the temperature, density, metallicity, and N(H I) (neutral hydrogen column density) attributes. For the cooler-gas sheet, we found highly variable physical conditions along each skewer within the sheet. The cumulative N(H I) ranged from log N(H I)/cm^{-2} = 11.77 - 15.25. The gas probed by the highest column density skewer could be easily detected by routine Hubble Space Telescope observations, where the others would not. The temperatures along the skewers varied from log T/K = 3.0 - 6.0, with the higher temperature regions coinciding with density peaks along the skewer. We conclude that the intergalactic medium in such a sheet might be detectable only in small regions of enhanced density, where as the entire structure is largely undetectable with current observational capabilities.

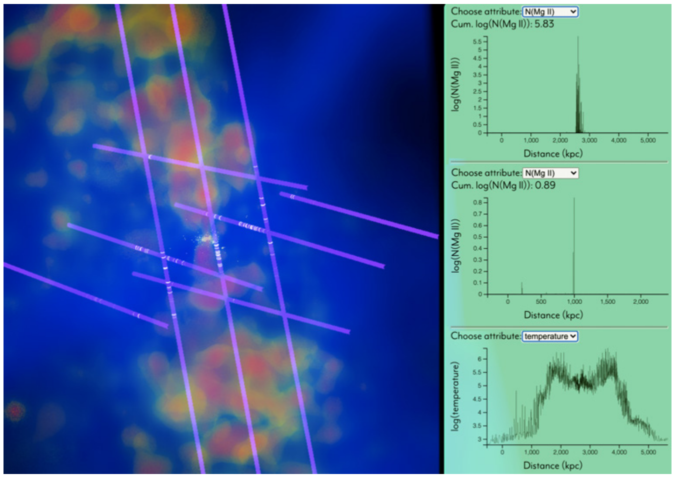

Fig. 7. The CGM environment analyzed in science Case 2. Here we have placed 10 skewers through an apparent biconical galactic superwind to analyze its thermodynamic structure.

5.2 Case 2: Simulating a QSO absorption line study of the circumgalactic medium

The circumgalactic environments of galaxies mediate the critical inflow and outflow processes that drive galaxy evolution [50]. Because the circumgalactic medium (CGM) is so diffuse, the gas does not emit light at levels bright enough to observe by imaging with current telescopes and instruments except for a few isolated cases [14, 46]. Therefore, the most reliable method for detecting this diffuse medium is via quasar absorption line spectroscopy, wherein a bright background source such as a quasar, or quasi-stellar object (QSO), is observed through spectroscopy, and the foreground material leaves its imprint on the QSO spectrum in the form of absorption lines [3]. This technique is highly sensitive to diffuse media and has the power to detect at densities several orders of magnitude below gas that emits light and can be imaged. However, this technique does have a major drawback: the sources that are bright enough to feasibly observe and measure their spectra are somewhat rare, and one must generally compile a statistical picture of the contents and gas motions of the CGM by compiling samples of multiple galaxies that have suitable probes of the CGM [5, 15, 30, 51]. Individual galaxies that contain multiple QSO probes of their CGM are exceedingly rare [9, 33].

Therefore, a fundamental question haunts the interpretation of these studies: are hidden variables contaminating the statistical composite picture of the CGM owing to the selection of multiple galaxies with underlying different properties? In this science Use Case, we attempt to to address this question by performing an experiment that simply cannot be conducted in the real Universe: by generating a suite of synthetic spectra probing individual galaxies to test the intrinsic variance of absorption properties within a single gaseous halo.

For our experiment, we identify a galaxy in a relatively isolated environment in the EAGLE 12 Mpc volume. The 12 Mpc box is ideal for this study because we seek a typical (not high-mass) galaxy that is fairly common in the Universe. Large statistical CGM studies are dominated by low-to-intermediate mass galaxies due to their prevalence among the galaxy population as a whole and it is in turn much more feasible to find bright QSOs (which rare) to probe them. Many suitable candidates exist within the 12 Mpc volume. Our galaxy of interest is shown in Figure 7. Note the plumes of warm/hot gas extending to either sides of galaxies. These result of gas outflows driven by supernovae and supermassive black hole activity [39, 54]. Clearly, the CGM of this galaxy contains inhomogenities in all of these quantities. We first examine the temperature, density, entropy, and metallicity gas modalities within the 'layers' menu, noting that the hotter inner regions inside these 'super-bubbles' appear to be enhanced in metallicity, indicating that the outflow is clearly carrying enriched material. However, these same regions are *suppressed* in density. We then placed ten skewers through this region to investigate how the CGM substructure manifests in the sightlines. The skewers were placed to probe directly through both lobes of the plume structure, through the central regions the galaxy CGM, and its periphery. The skewers reveal quantitatively the temperature distribution through the gas plumes, varying from $T \sim 10^6$ in one plume to 10^5 K in the central regions of the galaxy and back to to 10^6 K in the other plume. The column densities of H I and other ions in these skewers are quite low, with log N(H I)/cm^{-1} = 11.8-13.3 and log N(O VI)/cm^{-1} = 11.3-12.5. Typical values from the CGM literature are log N(H I)/cm^{-1} > 14.5 and log N(O VI)/cm^{-1} > 14.0 [31, 51, 52]. These low column densities might indicate that the hot winds from the galaxy are simultaneously sweeping out the material and ionizing it to states beyond those we are measuring, e.g., to O VII or O VIII. Indeed, checking the column densities of those species in *CosmoVis* show that well exceed N(O VI).

5.3 Case 3: Metal enrichment of the intergalactic medium

According to the accepted Big Bang Cosmology, elements with atomic number greater than two (helium) were only present in very small trace amounts at the beginning of the Universe [42]. Thus, the large amounts of carbon, nitrogen, iron, etc. out of which humans and our planet are constructed must have been forged in the nuclear reactions within stars. These heavy elements (or 'metals' in the common parlance of astronomers) therefore are ejected as stars shed their gaseous envelopes and explode in supernovae. These elements are then transported vast distances from their points of origin into the CGM and IGM [54]. Absorption signatures of metals have been long detected in spectra of quasars, and absorption line surveys of the circumgalactic medium have revealed unambiguously that the halos of galaxies are 'enriched' with metals [7, 17]. In fact, authors have posited that nearly all of the metal line systems detected to date arise from circumgalactic environments, even if the host galaxy (that presumably lies near the QSO line of sight) is unknown [45]. Posed differently, the question is "Have we yet detected truly IGM metal absorbers?"

In this Use Case, we turn *CosmoVis* to this question by using a combination of the 3D large-scale visualization and the virtual skewers to obtain column density measurements through various structures. Two main factors play into whether an metal absorption line system will be imprinted on a spectrum: 1) the medium must be metal enriched and 2) the medium must be dense enough to contain enough of a given species to leave a detectable absorption line.

Our general procedure as follows: We visualize the metallicity attribute of the gas layer and search for regions of the simulation volume that have relatively high metallicity values but that are well separated from the stars in galaxies. Then, visualizing the gas density layer of the same regions, we identify the highest-density subregions of the high-metallicity intergalactic environments. Using the skewer tool, we place skewers through this region, minding the extent of the skewer so that it only probes the region of question. Lastly, we retrieve the skewer attributes and scan the total column densities for a ions commonly detected in the literature, including Mg II, O VI, and C IV.

Upon examining all simulation volumes, metal-enriched regions of intergalactic filaments are ubiquitous. Contrary to expectation, we find many regions immediately surrounding galaxies devoid of enriched gas. Similar to the scenario observed in Science Case 2, it is likely that these environments have their gas by strong galactic outflows. In the EAGLE simulations, this phenomena is largely driven by black hole feedback and can clear a galaxy's CGM, eventually causing it to stop forming new stars [41]. Fig. 8 shows a large-scale view of the metallicity distribution in a slice of the EAGLE 25 Mpc volume, where several of these metal-poor cavities are evident. Identifying several candidate regions that might produce observable intergalactic metal absorption, we place skewers and inspect their ion column densities. We find no sightlines where, according the simulation, we should detect metal line absorbers in truly intergalactic space (i.e., not in the CGM).

6 Evaluation

In addition to evaluating *CosmoVis* through the scientific use cases presented above, we gathered extensive feedback when presenting the software and describing its capabilities at a number of conferences and

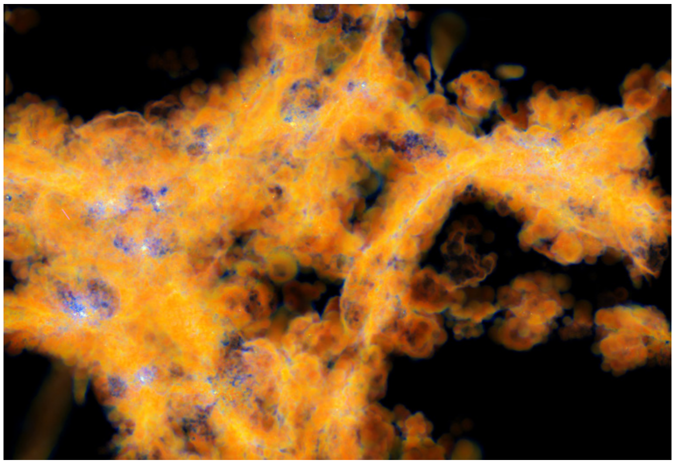

Fig. 8. The large-scale metallicity distribution in the EAGLE 25 Mpc volume. This is a ~ 5 Mpc slice of the cosmic web, revealing that much of the intergalactic medium in cosmic web filaments is highly enriched with heavy elements (low to high metallicity colored blue to red). Also seen in here are a number of regions in the immediate vicinities of galaxies likely evacuated of their metal-rich gas due to galactic winds [41].

workshops. In October 2020, we presented a visual analysis based on a prototype of *CosmoVis* to the *IEEE VisAstro Workshop* as part of their Data Challenge, which sought novel visualizations of relationships between galaxies and the cosmic web in order to gain new insight into the physical processes that shape galaxies across cosmic time [11]. In December 2020, we introduced *CosmoVis* at the *RHytHM: ResearcH using yt Highlights* Meeting. In February 2021, we showcased the *CosmoVis* software and discussed selected use cases as part of the *Workshop on the Fundamentals of Gaseous Halos*, organized by the Kavli Institute for Theoretical Physics (KITP). At the KITP workshop, 172 astrophysicists attended our software demonstration, and 47 of them accepted our invitation to attend a more in-depth exploration of *CosmoVis*. Feedback from these presentations was very positive, with many researchers interested in learning how they could use *CosmoVis* to investigate research questions using their own datasets.

Additionally, six researchers recruited from the KITP workshop provided us with detailed qualitative feedback to an open-ended study. We invited participants to use *CosmoVis* to complete specific tasks, soliciting information about the effectiveness of our visualization and interaction techniques for a range of analysis tasks. All participants were members of astrophysics labs engaged in research involving simulation datasets, and included one professor, three PhD students, one Master's student, and one undergraduate student. Surprisingly, three of the participants indicated that they normally did not use *any* visualization software to conduct their research. Two of the remaining participants mentioned using custom python and/or Matlab scripts to plot data. One researcher uses a range of tools, including Jupyter notebooks, yt, and Pynbody [44]. Prior to attempting the tasks, participants had watched a live software demonstration that illustrated the various elements of the *CosmoVis* user interface, and were given access to a recording of this. The tasks were chosen to represent a range of analysis activities that we expected to be familiar to workshop attendees, but difficult to carry out using existing software tools. They included the following: 1) Identify a region with 'warm-hot' 105K-106K gas; 2) Identify a region of high metallicity; 3) Find a signature of galactic winds; and 4) Use the virtual skewer tool to measure a distribution of temperature and ion species.

All participants were able to carry out the first two tasks, but two participants explained that they were not sufficiently comfortable with galactic winds research to carry out the third task, and two participants had trouble controlling the skewers when attempting to complete the fourth task. After finishing the tasks, we asked users to provide feedback on their experience and to rate their interest in using *CosmoVis* for different activities. Four users indicated that they were very likely to incorporate *CosmoVis* in their own research workflow as well to use it as an exploratory tool for investigating simulation datasets, and all six users indicated that they believed *CosmoVis* would be useful as an education tool in pedagogical contexts.

The participants commented positively on their experience using *CosmoVis*. One told us that they "really enjoyed the overall experience of looking at the data from different positions and perspectives." Another wrote "I like that you can start with the whole dataset and then it's responsive enough that you can focus-in on an interesting feature quickly." Another was excited about the possibility of using *CosmoVis* for public outreach: "I think this is a fantastic tool to introduce non-scientists to our field of research which is so typically walled off by hurdles such as spectroscopy, scaling relations, and calculations that are difficult to convey to a typical outreach/classroom audience".

We were especially interested in users' reactions to using skewers within the simulation volumes. One participant told us that "the skewers were a little confusing." Another found CosmoVis to be "so slick! [...] an amazing tool to quickly plot physical properties and ion content", while also warning that "it's very easy to accidentally add additional skewers without meaning to." Another participant also initially found it difficult to use of the skewers, but "Once I got the skewers to work the ability to generate spectra on-the-fly was great! On-the-whole there's a good range of features but not so many that it's overwhelming." Three of the participants mentioned the importance of more extensive documentation (i.e., beyond on the video tutorial), with one commenting that "a tutorial or user guide would go a long way toward making me feel like I'm getting the most out of *CosmoVis*."

Users also reported having new insights into their own research, even after working with *CosmoVis* for only a few hours. One participant expressed surprise that "IGM filaments have higher metallicity than I expected!". Another was intrigued by what they learned about gas properties: "Why are the hottest gases in the cube primarily at the edges? Fascinating!" Yet another told us "It's cool to see how stars are so neatly embedded in the clouds of metals." One participant told us that they looked forward to using *ComsoVis* in a current research project: "I could envision aggregating multiple sightlines and then comparing to the ionization-modelled quantities from observed spectra." Yet another explained why they would find *CosmoVis* useful for their own data explorations: "I can quickly get a overall visual of my snapshot *and* look at various parts of data structs – e.g. temperature."

Participants also expressed interest in additional features, including the following requests: "I wanted a velocity field for finding a super bubble"; "It would be great to be able to incorporate custom fields"; "A second skewer type that were instead radial shells might be useful"; "As well as having the sliders (for masking regions and choosing the skewer width), being able to input a numerical value would be useful". Based on this feedback, we are planning to incorporate some of these ideas into future updates of *CosmoVis*.

7 Conclusion

In this paper, we have introduced *CosmoVis*, a novel visualization software application for rendering and analyzing cosmological simulation datasets. We presented a series of scientific use cases demonstrating that *CosmoVis* is a useful tool for supporting a range of analysis tasks that enable new approaches to investigate a range of astrophysical phenomena, including identifying sheet regions, simulating a QSO absorption line to analyze the circumgalactic medium, and exploring metal enrichment within the intergalactic medium. We described the positive reception of *CosmoVis* by different communities of astrophysicists, and provided initial detailed feedback from six domain experts.

Future work will incorporate additional simulation datasets, and simplify the pipeline for ingesting custom datasets. Additionally, while *CosmoVis* enables graphing of skewer column densities and synthetic spectra, we plan to incorporate more analysis functionality, such as generating equivalent width measurements. We also are currently in the process of generalizing our representation of star particles to provide more information about the specific subhalos or galaxies of which they are a part of. Another area for future exploration is extending *CosmoVis* to animate the evolution of simulations across a range of redshift snapshots.

REFERENCES

[1] J. Ahrens, B. Geveci, and C. Law. ParaView: An end-user tool for large data visualization. In *The Visualization Handbook*, chap. 36, pp. 717–732. Elsevier, 2005.

[2] U. Ayachit. *The ParaView Guide: A Parallel Visualization Application*. Kitware, Inc., 2015.

[3] J. N. Bahcall and L. Spitzer, Jr. Absorption Lines Produced by Galactic Halos. *ApJL*, 156:L63, May 1969.

[4] M. J. Berger and J. Oliger. Adaptive mesh refinement for hyperbolic partial differential equations. *Journal of Computational Physics*, 53(3):484–512, 1984.

[5] J. Bergeron and P. Boissé. A sample of galaxies giving rise to mg II quasar absorption systems. *A&A*, 243:344–366, Mar. 1991.

[6] A. Bock, E. Axelsson, J. Costa, G. Payne, M. Acinapura, V. Trakinski, et al. Openspace: A system for astrographics. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):633–642, 2019.

[7] A. Boksenberg and W. L. W. Sargent. The existence of CA II absorption lines in the spectrum of the quasar 3c 232 due to the Galaxy NGC 3067. *The Astrophysical Journal*, 220:42–46, 1978.

[8] J. R. Bond, L. Kofman, and D. Pogosyan. How filaments of galaxies are woven into the cosmic web. *Nature*, 380(6575):603–606, Apr 1996.

[9] D. V. Bowen, D. Chelouche, E. B. Jenkins, T. M. Tripp, M. Pettini, D. G. York, and B. L. Frye. The Structure of the Circumgalactic Medium of Galaxies: Cool Accretion Inflow Around NGC 1097. *ApJ*, 826:50, July 2016.

[10] E. F. Bunn and D. W. Hogg. The kinematic origin of the cosmological redshift. *American Journal of Physics*, 77(8):688–694, 2009.

[11] J. N. Burchett, D. Abramov, O. Elek, and A. G. Forbes. Volumetric reconstruction for interactive analysis of the cosmic web. In *IEEE VIS Workshop on Visualization in Astrophysics (VisAstro)*, pp. 1–15, 2020.

[12] J. N. Burchett, D. Abramov, J. Otto, C. Artanegara, J. X. Prochaska, and A. G. Forbes. IGM-Vis: Analyzing intergalactic and circumgalactic medium absorption using quasar sightlines in a cosmic web context. In *Computer Graphics Forum*, vol. 38, pp. 491–504, 2019.

[13] J. N. Burchett, O. Elek, N. Tejos, J. X. Prochaska, T. M. Tripp, R. Bordoloi, and A. G. Forbes. Revealing the dark threads of the Cosmic Web. *The Astrophysical Journal Letters*, 891(2):L35, 2020.

[14] J. N. Burchett, K. H. R. Rubin, J. X. Prochaska, A. L. Coil, R. R. Vaught, and J. F. Hennawi. Circumgalactic Mg II Emission from an Isotropic Starburst Galaxy Outflow Mapped by KCWI. , 909(2):151, Mar. 2021.

[15] J. N. Burchett, T. M. Tripp, R. Bordoloi, J. K. Werk, J. X. Prochaska, J. Tumlinson, C. N. A. Willmer, J. O'Meara, and N. Katz. A Deep Search For Faint Galaxies Associated With Very Low-redshift C IV Absorbers: III. The Mass- and Environment-dependent Circumgalactic Medium. *ApJ*, 832(124):124, Dec. 2016.

[16] R. Cen, J. Miralda-Escudé, J. P. Ostriker, and M. Rauch. Gravitational collapse of small-scale structure as the origin of the Lyman-alpha forest. *The Astrophysical Journal Letters*, 437:L9–L12, Dec. 1994.

[17] H.-W. Chen, K. M. Lanzetta, and J. K. Webb. The origin of c IV absorption systems at redshifts $z < 1$: Discovery of extended c IV envelopes around galaxies. *ApJ*, 556:158–163, July 2001.

[18] M. Crone Odekon, G. Hallenbeck, M. P. Haynes, R. A. Koopmann, A. Phi, and P.-F. Wolfe. The Effect of Filaments and Tendrils on the H I Content of Galaxies. *The Astrophysical Journal*, 852:142, Jan. 2018.

[19] O. Elek, J. N. Burchett, J. X. Prochaska, and A. G. Forbes. Polyphorm: Structural analysis of cosmological datasets via interactive Physarum polycephalum visualization. *Transactions of Visualization and Computer Graphics*, 27(2), 2021.

[20] L. Fritschi, I. B. Rojo, and T. Günther. Visualizing the temporal evolution of the universe from cosmology simulations. In *Proc. IEEE Scientific Visualization Conference (SciVis)*, pp. 1–11, 2019.

[21] A. Geller and A. Gurvich. *Firefly: A WebGL* tool to explore particle-based data, 2021. `https://agurvich.github.io/Firefly/`. Accessed March 30th, 2021.

[22] J. Guedes, S. Callegari, P. Madau, and L. Mayer. Forming Realistic Late-type Spirals in a ΛCDM Universe: The Eris Simulation. , 742(2):76, Dec. 2011.

[23] S. Habib, A. Pope, H. Finkel, N. Frontiere, K. Heitmann, D. Daniel, et al. HACC: Simulating sky surveys on state-of-the-art supercomputing architectures. *New Astronomy*, 42:49–65, 2016.

[24] C. Hesse-Edenfeld, M. J. Steinke, N. Santalidis, K. Huesmann, S. Leistikow, and L. Linsen. Interactive multi-level visualization of combined particle and multi-field cosmology data. In *Proc. IEEE Scientific Visualization Conference (SciVis)*, pp. 12–30, 2019.

[25] P. F. Hopkins. A new class of accurate, mesh-free hydrodynamic simulation methods. *Monthly Notices of the Royal Astronomical Society*, 450(1):53–110, 2015.

[26] P. F. Hopkins, A. Wetzel, D. Kereš, C.-A. Faucher-Giguère, E. Quataert, M. Boylan-Kolchin, et al. FIRE-2 simulations: Physics versus numerics in galaxy formation. *Monthly Notices of the Royal Astronomical Society*, 480(1):800–863, 2018.

[27] E. Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Science*, 15:168–173, 1929.

[28] C. B. Hummels, B. D. Smith, P. F. Hopkins, B. W. O'Shea, D. W. Silvia, J. K. Werk, et al. The impact of enhanced halo resolution on the simulated circumgalactic medium. *The Astrophysical Journal*, 882(2):156, 2019.

[29] C. B. Hummels, B. D. Smith, and D. W. Silvia. Trident: A Universal Tool for Generating Synthetic Absorption Spectra from Astrophysical Simulations. *The Astrophysical Journal*, 847:59, 2017.

[30] S. D. Johnson, H.-W. Chen, and J. S. Mulchaey. On the possible environmental effect in distributing heavy elements beyond individual gaseous haloes. *Monthly Notices of the Royal Astronomical Society*, 449:3263–3273, 2015.

[31] G. G. Kacprzak, J. R. Vander Vliet, N. M. Nielsen, S. Muzahid, S. K. Pointon, C. W. Churchill, D. Ceverino, K. S. Arraki, A. Klypin, J. C. Charlton, and J. Lewis. The Relation between Galaxy ISM and Circumgalactic O VI Gas Kinematics Derived from Observations and ΛCDM Simulations. , 870(2):137, Jan. 2019.

[32] N. I. Libeskind, R. van de Weygaert, M. Cautun, B. Falck, E. Tempel, T. Abel, et al. Tracing the cosmic web. *Monthly Notices of the Royal Astronomical Society*, 473:1195–1217, 2018.

[33] S. Lopez, N. Tejos, C. Ledoux, L. F. Barrientos, K. Sharon, J. R. Rigby, M. D. Gladders, M. B. Bayliss, and I. Pessa. A clumpy and anisotropic galaxy halo at redshift 1 from gravitational-arc tomography. *Nature*, 554:493–496, 2018.

[34] N. Mandelker, F. C. van den Bosch, V. Springel, and F. van de Voort. Shattering of Cosmic Sheets due to Thermal Instabilities: A Formation Channel for Metal-free Lyman Limit Systems. , 881(1):L20, Aug. 2019.

[35] N. Max. Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2):99–108, 1995.

[36] S. McAlpine, J. C. Helly, M. Schaller, J. W. Trayford, Y. Qu, M. Furlong, R. G. Bower, R. A. Crain, J. Schaye, T. Theuns, et al. The EAGLE simulations of galaxy formation: Public release of halo and galaxy catalogues. *Astronomy and Computing*, 15:72–89, 2016.

[37] J. J. Monaghan. Smoothed particle hydrodynamics. *Annual review of astronomy and astrophysics*, 30(1):543–574, 1992.

[38] J. Naiman. Astroblend: An astrophysical visualization package for blender. *Astronomy and Computing*, 15:50–60, 2016.

[39] D. Nelson, A. Pillepich, V. Springel, R. Pakmor, R. Weinberger, S. Genel, P. Torrey, M. Vogelsberger, F. Marinacci, and L. Hernquist. First results from the TNG50 simulation: galactic outflows driven by supernovae and black hole feedback. , 490(3):3234–3261, Dec 2019.

[40] B. D. Nguyen, N. V. Nguyen, V. Pham, and T. Dang. Visualization of data from HACC simulations by ParaView. In *Proc. IEEE Scientific Visualization Conference (SciVis)*, pp. 31–32, 2019.

[41] B. D. Oppenheimer, J. J. Davies, R. A. Crain, N. A. Wijers, J. Schaye, J. K. Werk, J. N. Burchett, J. W. Trayford, and R. Horton. Feedback from supermassive black holes transforms centrals into passive galaxies by ejecting circumgalactic gas. *arXiv e-prints*, Apr. 2019.

[42] P. J. E. Peebles. *Principles of Physical Cosmology*. Number 978-0-691-01933-8. Princeton University Press, 1993.

[43] A. Pillepich, V. Springel, D. Nelson, S. Genel, J. Naiman, R. Pakmor, et al. Simulating galaxy formation with the IllustrisTNG model. *Monthly Notices of the Royal Astronomical Society*, 473(3):4077–4106, 2018.

[44] A. Pontzen, R. Roškar, G. S. Stinson, R. Woods, D. M. Reed, J. Coles, and T. R. Quinn. pynbody: Astrophysics Simulation Analysis for Python, 2013. Astrophysics Source Code Library, ascl:1305.002.

[45] J. X. Prochaska, B. Weiner, H.-W. Chen, J. Mulchaey, and K. Cooksey. Probing the intergalactic medium/galaxy connection. V. On the origin of and O VI absorption at $z < 0.2$. *The Astrophysical Journal*, 740:91, 2011.

[46] D. S. N. Rupke, A. Coil, J. E. Geach, C. Tremonti, A. M. Diamond-Stanic, E. R. George, R. C. Hickox, A. A. Kepley, G. Leung, J. Moustakas, G. Rudnick, and P. H. Sell. A 100-kiloparsec wind feeding the circumgalactic medium of a massive compact galaxy. , 574(7780):643–646, Oct. 2019.

[47] K. Schatz, C. Müller, P. Gralka, M. Heinemann, A. Straub, C. Schulz, et al. Visual analysis of structure formation in cosmic evolution. In *Proc. IEEE Scientific Visualization Conference (SciVis)*, pp. 33–41, 2019.

[48] J. Schaye, R. A. Crain, R. G. Bower, M. Furlong, M. Schaller, T. Theuns, et al. The EAGLE project: Simulating the evolution and assembly of galaxies and their environments. *Monthly Notices of the Royal Astronomical Society*, 446(1):521–554, 2015.

[49] N. Sukumar. Voronoi cell finite difference method for the diffusion operator on arbitrary unstructured grids. *International Journal for Numerical Methods in Engineering*, 57(1):1–34, 2003.

[50] J. Tumlinson, M. S. Peeples, and J. K. Werk. The circumgalactic medium. *Annual Review of Astronomy and Astrophysics*, 55:389–432, 2017.

[51] J. Tumlinson, C. Thom, J. K. Werk, J. X. Prochaska, T. M. Tripp, N. Katz, R. Davé, B. D. Oppenheimer, J. D. Meiring, A. B. Ford, J. M. O'Meara, M. S. Peeples, K. R. Sembach, and D. H. Weinberg. The COS-Halos survey: Rationale, design, and a census of circumgalactic neutral hydrogen. *The Astrophysical Journal*, 777:59, 2013.

[52] J. Tumlinson, C. Thom, J. K. Werk, J. X. Prochaska, T. M. Tripp, D. H. Weinberg, M. S. Peeples, J. M. O'Meara, B. D. Oppenheimer, J. D. Meiring, N. S. Katz, R. Davé, A. B. Ford, and K. R. Sembach. The large, oxygen-rich halos of star-forming galaxies are a major reservoir of galactic metals. *Science*, 334:948, 2011.

[53] M. J. Turk, B. D. Smith, J. S. Oishi, S. Skory, S. W. Skillman, T. Abel, and M. L. Norman. yt: A multi-code analysis toolkit for astrophysical simulation data. *The Astrophysical Journal Supplement Series*, 192(9):1–16, 2010.

[54] S. Veilleux, G. Cecil, and J. Bland-Hawthorn. Galactic winds. *Annual Review of Astronomy and Astrophysics*, 43:769–826, 2005.

[55] J. Woodring, K. Heitmann, J. Ahrens, P. Fasel, C.-H. Hsu, S. Habib, and A. Pope. Analyzing and visualizing cosmological simulations with ParaView. *The Astrophysical Journal Supplement Series*, 195(1):11, 2011.