The Baryon Cycle through Cosmological Simulations: Taking COS to the Next Generation of Analysis

Scientific Category: Galaxies and the IGM

Scientific Keywords: Circumgalactic Medium, Galaxy Formation and Evolution, Metal Absorption Systems

Budget Size: Regular

UV Initiative: Yes Theory: Yes

Abstract

How do baryons cycle in and out of galaxies, and how does the baryon cycle drive galaxy evolution? Stellar feedback processes and accretion from the intergalactic medium give rise to extended metal-enriched gaseous structures surrounding galaxies: the circumgalactic medium (CGM). Arguably, the CGM is the key component of the baryonic cycle that regulates galaxy evolution and is therefore the most critical component of the cycle to pin down. COS was developed primarily for this purpose and 100s of HST orbits have been dedicated to CGM observations. However, high resolution hydrodynamic simulations show that standard observational analysis methods of absorption lines are fraught with assumptions that are no longer justified. We aim to address this, thereby increasing the intrinsic sceintific value of archived HST/COS spectra.

We propose to enhance the legacy of HST/COS by analyzing Eulerian hydrodynamic cosmological simulations incorporating physics based baryon cycle processes. We will (1) undertake absorption line analysis of simulated galaxies that accurately emulates observational data and analysis methods and compare the results to published COS programs, (2) compare the 3D simulations to the absorption line results, (3) assess commonly applied absorption line methods in order to provide insights into the interpretation of HST/COS data, and (4) develop new absorption line analysis methods that better reflect the CGM gas spatial and kinematic distributions.

Our goal is to provide deeper insights into current observational techniques for improving our observational knowledge of the baryon cycle from COS spectra.

Investigators:

Investigator	Institution	Country
C Churchill	New Mexico State University	USA/NM
G Kacprzak	Swinburne University of Technology	AUS
A Klypin	New Mexico State University	USA/NM
N Nielsen	Swinburne University of Technology	AUS
J Vander Vliet	New Mexico State University	USA/NM

Number of investigators: 5

Dataset Summary:

■ Scientific Justification

The last decadal report (2011, National Research Council) singled out the following questions as top priority for extragalactic science: **"How do baryons cycle in and out of galaxies, what do they do while they are there, and how does the cycle drive galaxy evolution?"** Elevating the "baryon cycle" to top priority comes with the realization that the cycling of baryons through dark matter halos is a dominant process responsible for the observed universe of stars and galaxies. Knowing the baryon cycle means a fundamental understanding of galaxies, star formation, and chemical evolution.

The interplay between stellar feedback processes in the interstellar medium (ISM) and filamentary accretion from the intergalactic medium (IGM) gives rise to extended (~ 150 kpc) metal-enriched gaseous structures surrounding galaxies, i.e., the circumgalactic medium (CGM). As the interface between the ISM and IGM, the CGM is a key regulating component of galaxies and therefore a critical component of the baryon cycle to understand.

From a theoretical standpoint, the key is to establish the physical processes by which baryonic matter responds to the full spectrum of gravity induced dark matter overdensities, thereby explaining many observed global galaxy relations (such as the stellar mass to halo mass ratio and its cosmic evolution, and the stellar mass metallicity relation; Tremonti+ 2004, ApJ, 613, 898; Andrews+ 2013, ApJ, 765, 140; Behroozi+ 2013, ApJ, 770, 57, Gu+ 2016, ApJ, 833, 2; Ly+ 2016, ApJ, 828, 67). Significant effort has been directed toward understanding the baryon cycle using hydrodynamic cosmological simulations via the tracking of gas cycles and the modeling of star formation and feedback physics (Kereš+ 2005, MNRAS, 363, 2; Oppenheimer+ 2008, MNRAS, 387, 577; Ceverino+ 2009, ApJ, 695, 292; Kereš+ 2009, MNRAS, 395, 160; Ceverino+ 2013, MNRAS, 442, 1545; Trujillo-Gomez+ 2015, MNRAS, 446, 1140; Christensen+ 2016, ApJ, 824, 57; Davé+ 2016, MNRAS, 426, 3265; Finlator 2016, arXiv:1612.00802). Davé et al. (2011a, MNRAS, 415, 11) were able to show that star-forming galaxies develop via a slowly evolving equilibrium balanced by inflows (driven by gravity/mass), wind recycling, star formation rates, and outflows, the latter regulating the fraction of inflow that gets converted into stars. The CGM gas content regulates the competition between IGM inflow and gas consumption within the ISM as governed by star formation (Birrer+ 2014, ApJ, 793, 12).

The best approach for observationally probing baryons in, around, and between galaxies is to analyze absorption lines in spectra of background sources (quasars) with sightlines passing near galaxies. The Cosmic Origins Spectrograph (COS) on *HST* was developed primarily for this purpose. From the observational standpoint, the key is to quantify and characterize the dynamics, spatial distributions, metallicities, densities, and temperatures of the gas flowing into, out of, and through galaxies. Absorption lines indirectly provide these properties; we say "indirectly" because chemical-ionization models of the gas are required to take incomplete information from various absorbing transitions of various ions in order to construct a complete visualization of the gas. A further difficulty is that the lines-of-sight to the luminous sources provide a pencil beam 1D probe of the gas as a function of line of sight velocity (3D velocity dotted into the line of sight). We are quite blind to reality.

• **UV Initiative and HST/COS/STIS Legacy**: Even with these well-known hurdles, large investments in *HST* time have been dedicated to absorption line observations of the CGM, such as COS-Halos (11598, 13033), COS-Dwarfs (12248), the COS Science Team (11541, 12025), our Large Program (12466, 12252, 13398), and others (e.g., 10151). These surveys, representing ~ 650 *HST* orbits, effectively provide the world's entire library of UV spectroscopic data on the CGM associated with identified galaxies (~ 200 galaxies). They are the collective UV legacy of *HST*/COS and STIS ultraviolet absorption line data that will remain available for study well beyond the lifetime of *HST*. *Our goals are to examine and "calibrate" current analysis methods of COS/STIS spectra using simulated data to gain deeper insight into the baryon cycle and spearhead the long-term collective process of increasing the value of archived HST UV absorption line spectra.*

The key to deeper insights into *HST*/COS+STIS absorption line data are to emulate observational survey parameters and analysis techniques employed by the observers to "mock" data from the cosmological simulations. Astonishingly, nobody is adopting this approach, which offers our only direct quantitative method for connecting the 1D probes provided by *HST*/COS+STIS to the 3D gas spatial distribution and velocity fields in simulations.

• **Historical precedent for dramatic success of absorption line analysis through simulations** is the revolution in understanding of the Lyα forest in the 1990s, which yielded the cosmic web *paradigm* of large scale structure. Armed with high-resolution absorption line data of the Lyα forest, our models of the H I gas were no more sophisticated than "spheres", "slabs", and "dark matter mini-halos" (Meylan 1995, "QSO Absorption Lines", Springer-Verlag). *The data alone had no potential to uncover the large scale coherent structure of* H I *in the IGM. Simulations forced a paradigm shift.* Currently, models of the CGM used to explain data are in their infancy; outflows are modeled as "bi-conical winds", accretion as planer inflows due to angular momentum loss, and incidence statistics by spherical "halo" distributions and halo occupation formalism (e.g., Tinker & Chen 2010, ApJ, 709, 1; Bouché+ 2013, Science, 341, 50; Bordoloi+ 2014 ApJ, 784, 108; Fox+ 2015, ApJ, 799, 7). The CGM gas itself is naively modeled as discrete isothermal "clouds" using Voigt profiles.

Modern cosmological simulations incorporate sophisticated stellar feedback processes into the baryon physics (supernovae, radiation pressure, radiative heating, etc.) These provide an open road for applying observational absorption line analysis methods to simulated galaxies. The coming change in the methods of analysis holds the potential for revolutionary paradigm shift in our interpretation of *HST*/COS+STIS absorption line data.

We propose to maximize the legacy of *HST*/COS by exploiting cutting-edge Eulerian adaptive mesh refinement hydrodynamic cosmological simulations incorporating physics based baryon cycle processes. We will (1) undertake absorption line analysis of simulated galaxies that accurately emulates observational data and analysis methods, (2) compare the 3D simulations to the 1D absorption line results, (3) quantitatively calibrate current methods in order to provide heightened insights into the interpretation of *HST*/COS data, and (4) work toward developing new absorption line analysis methods.

Our methods are founded on generating realistic spectra (pixel scales, resolution sampling, noise characteristics) and emulating the impact parameter, galaxy mass, redshift, and spectral detection sensitivity distributions of the observational surveys. Totally unique to our approach is that we have developed a method to isolate the *actual* gas cells in the simulations that created the absorption lines (Churchill+ 2015a, ApJ, 802, 10; Vander Vliet 2017, NMSU PhD). This is critically important because it means we can select the gas in the simulations that is being probed by the observations conducted by the *HST*/COS surveys, rather than generically comparing to the simulations.

In **Figure 1**, we see a simulated galaxy and CGM (Churchill+ 2015a) having such complexity that current observational analysis are rendered obsolete. For the shown line of sight (LOS), "mock" H I Lyβ, Mg II λ2796, C IV λ1548, and O VI λ1031 absorption lines are presented. The data are FUV and NUV COS spectra and, for Mg II, HIRES spectra. The LOS spatial locations, densities, temperatures, and metallicities of the gas cells responsible for the absorption (the "absorbing gas") are indicated with color coded "$\times$" data points. We find that H I and Mg II absorption arise in "cloud-like" structures that are confined to contiguous gas cells over small spatial scales and have a narrow temperature distribution, but with density and metallicity gradients. However, we find that the C IV and the O VI absorption is not cloud-like; *the absorption arises from gas spatially distributed over 100 kpc with a complex reversing velocity flow that results in multiple separated locations giving rise to absorption at a single value of LOS velocity.* (Dear reader, apologies for the figure's complexity- but so is the CGM!). Furthermore, note that the O VI absorption arises from gas without detectable H I absorption. How do we analyze that?

These insights run completely counter to currently applied analysis methods (Voigt profile modeling and chemical ionization modeling), which assume (1) discrete clouds with single valued densities, temperatures, and metallicities, and (2) discrete clouds at each LOS velocity. It follows that systematic errors will be the outcome of our current analysis methods. Determining these systematics is one of our goals (see Analysis Plans for all goals). To what degree do the gas properties we derive from commonly employed observational analysis techniques applied to absorption lines from simulations match the properties of the gas actually giving rise to the absorption? What signatures can be identified in the absorption line data that clearly indicate the complex velocity fields and spatially discrete structures that map to a given LOS velocity (such as we illustrated for O VI in Figure 1)? **What are plausible new approaches to absorption line analysis that incorporate this more realistic view of the CGM?** What are the implications for our understanding of the CGM and the baryon cycle once the signatures can be employed for observational data systematics are accounted?

Analysis of absorption properties in hydrodynamic cosmological simulation is a highly promising, yet undeveloped methodology with regards to the baryon cycle. We aim to develop novel analysis methods using simulations so that deeper insights can be gleaned from absorption line data, especially *HST*/COS+STIS spectra. Beyond *HST*, the COS+STIS archives can continue to deliver new insights into understanding galaxy evolution as analysis methods evolve; comparison to simulations will be key to the *HST* legacy.

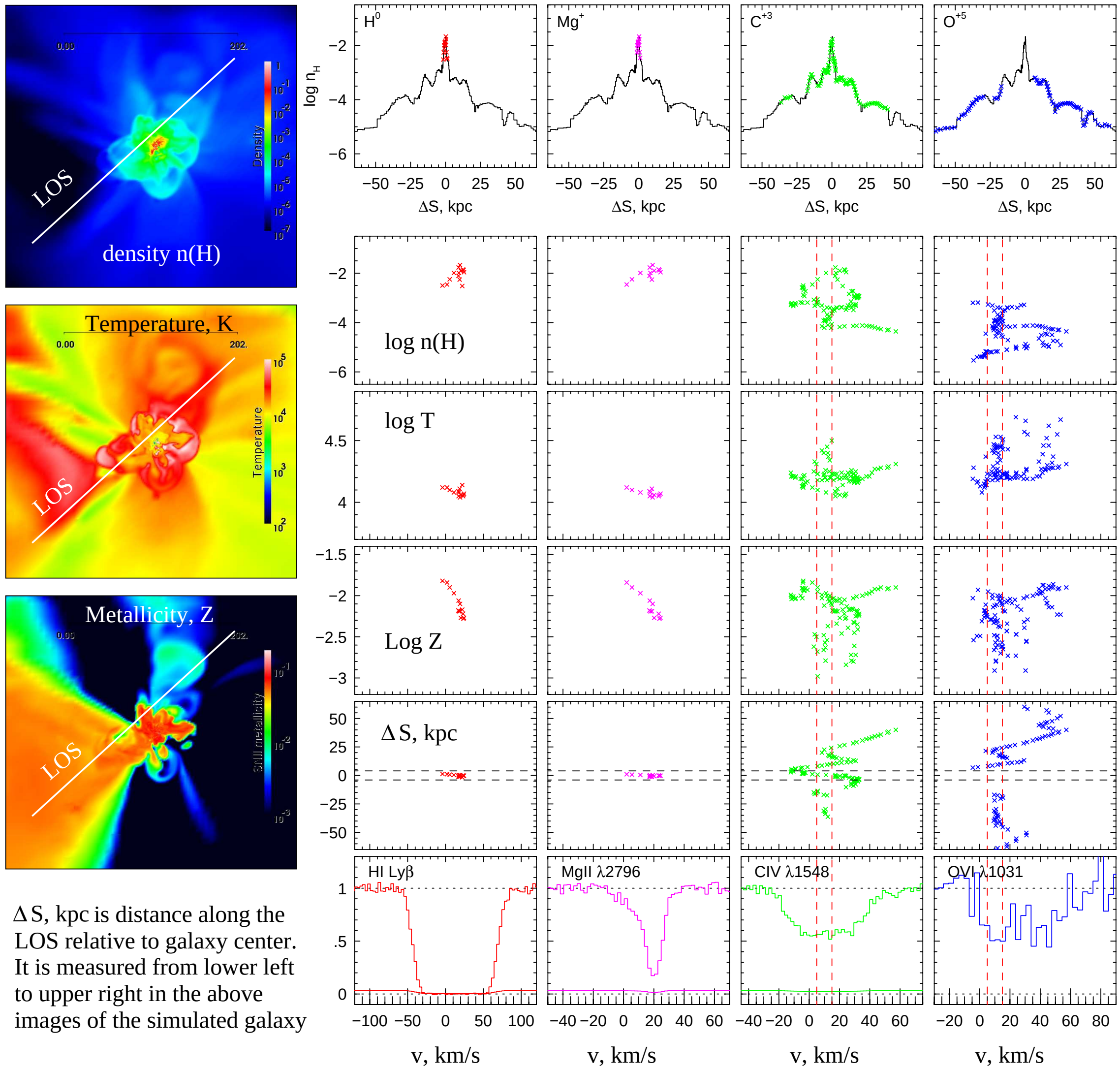

Figure 1: **Unique to our study** is that we focus on the gas cells that contribute to absorption and can therefore examine the physical properties of the "absorbing gas" for direct comparison of deduced properties from observational analysis methods applied to the simulated spectra. **This aspect of our study provides unprecedented insight into the accuracy of methods applied to *HST* COS and STIS spectra of the CGM.** — (left panels) The density, temperature, and metallicity images of the simulated galaxy, showing a "mock" line of sight (LOS). — (top right) the density of the gas as a function of LOS position, where $\Delta S = 0$ is the plane of the sky from the observer point of view. Absorbing gas cells are highlighted in red (H I), purple (Mg II), green (C IV), and blue (O IV). — (bottom right panels, top to bottom) The density, temperature, metallicity, and LOS position of absorbing gas cells, and the absorption profiles as a function of LOS velocity, v. Note that H I and Mg II absorbers are "cloud-like", with a tight range of properties, but that C IV and O VI absorbers comprise multiple spatially separated structures at any given LOS velocity (consider $v = 5$–15 km s^{-1} as denoted by the vertical dashed lines. *Conclusion: standard VP analysis of* C IV *and* O VI *is a failed analysis method.*

Analysis Plan

Our overall science goals for studying the baryon cycle are summarized by the following: **(1)** to examine how the absorbing gas properties of the CGM depend on different feedback models for different mass galaxies; **(2)** to examine how galaxies of different masses accrete gas from the IGM, and how different feedback models modify the accretion and transport of accreting gas through the CGM; **(3)** to investigate the mechanisms by which baryons are transported, mixed, and cycled back into the CGM following stellar feedback processes; **(4)** to determine the degree to which the gas properties derived from commonly employed observational analysis techniques, as applied to absorption lines from simulations, match the properties of the absorbing gas. Foremost of these goals are the objectives to understand the baryon cycle and its role in galaxy evolution while to provide improved insights into the analysis of COS spectra of the CGM.

• **The Simulations and "Mock" Absorption Lines**

We employ the adaptive mesh refinement N-body and hydrodynamics code hydroART (Kravtsov+ 1997, ApJS, 111, 73; Kravtsov & Klypin 1999, ApJ, 520, 437) as modified with cutting-edge stellar feedback physics (Ceverino & Klypin 2009, ApJ, 695, 292; Trujillo-Gomez 2015, MNRAS, 446, 114). Galaxies are formed and evolved in the full cosmological environment. Physical processes implemented in the code include star formation, stellar feedback, Type II and Ia metal enrichment, thermal and radiation pressure, photo-heating, and metallicity-dependent cooling and heating. Gas is self-shielded, advects metals, is heated by a homogeneous ultraviolet background, and can cool to 100 K due to metal and molecular line cooling. Gas flows, shock fronts, and metal disbursement follow self-consistently from this physics. Star formation occurs in the dense, cold molecular phase ($n_{\rm H} \sim 100$ cm^{-3}, $T \simeq 100$ K), which is disrupted by the combination of radiation pressure and photoionization by massive stars. The resolution ranges from 20–200 pc, depending on redshift and halo mass. We developed a full-treatment photo+collisional ionization code (fully tested against Ferland's Cloudy, Churchill+ 2015b, arXiv:1409.0916) that computes the ionization corrections for all ionic species in each cell of the simulation box. The treatment accounts for shielding by neighboring cells and for both UVB and stellar radiation (using Starburt99 SEDs accounting for the ages, metallicities, and distances of stellar particles).

We have built a library of 17 Milky Way mass galaxies run to $z = 1$, 40 dwarf galaxies run to $z = 0$, 80 sub-Milky Way mass galaxies to $z = 0$, and are currently running 60 dwarf and super-Milky Way mass galaxies to $z = 0$. We incorporate three components to the stellar feedback models: (1) supernovae and stellar winds, SNW; (2) radiation pressure, RP; (3) weak or strong photoheating, PH1/PH2. All galaxies include SNW, whereas variations of RP, and PH1 or PH2 are employed. Two star formation models are examined: (1) probabilistic, where 100% of the gas meeting the threshold for star formation is converted to stars based upon a fixed probability, and (2) deterministic, where 2% of the gas is converted to stars every time the gas meets the threshold for star formation. We thus, in the course of our study, will examine how the CGM and the absorption lines it gives rise to depend on various baryon cycle physics (i.e., feedback recipes).

- **Analysis of “Mock Absorption Lines”**

We have fully developed and tested a science-ready pipeline to (1) generate LOS vectors, (2) run the LOS through the simulations, (3) generate realistic simulated spectra with fully specified data characteristics, (4) automatically and objectively detect and quantitatively measure the absorption lines, (5) determine the gas cells in the simulations that give rise to the detected absorption lines for each ion for each LOS, and (6) catalog the results for science analysis. Details are presented in Churchill et al. (2015a). Typically, we run 1000 LOS through the simulated galaxies for two to four galaxy orientations. The LOS range from $0 \leq D/R_{\rm vir} \leq 2.0$. For comparison with observations, the synthetic spectra are generated using the exact characteristics of observed data, including the resolution (convolution with the instrumental spread function), pixelization, and noise characteristics.

- **Using HST UV Spectra to Constrain Baryon Cycle Physics**

In **Figure 2**, we present selected observations of the CGM (Nielsen+ 2013, ApJ, 776, 115; Churchill+ 2013, ApJ, 763, L42; Nielsen+ 2016, ApJ, 818, 171). These Mg II distributions serve to illustrate a small sample of observed characteristics of the CGM that can be examined. For H I, C IV, and O VI absorption, there exists a substantial body of similar observational work that can be exploited for direct comparison of “mock” absorption lines from the simulations having different baryon cycle physics (e.g., Fox+ 2013, ApJ, 778, 187; Stocke+ 2013, ApJ, 763, 149; Bordoloi+ 2014, ApJ, 796, 136; Werk+ 2014, ApJ, 792, 8; – 2016, ApJ, 833, 54; Kacprzak+ 2015, ApJ, 815, 22; Borthakur+ 2015, ApJ, 813, 46; – 2016, ApJ, 833, 259; Burchett+ 2016, ApJ, 832, 124; Keeney+ 2017, arXiv:1704:00235; Nielsen+ 2017, ApJ, 834, 148; Prochaska+ 2017, ApJ, 837, 169)

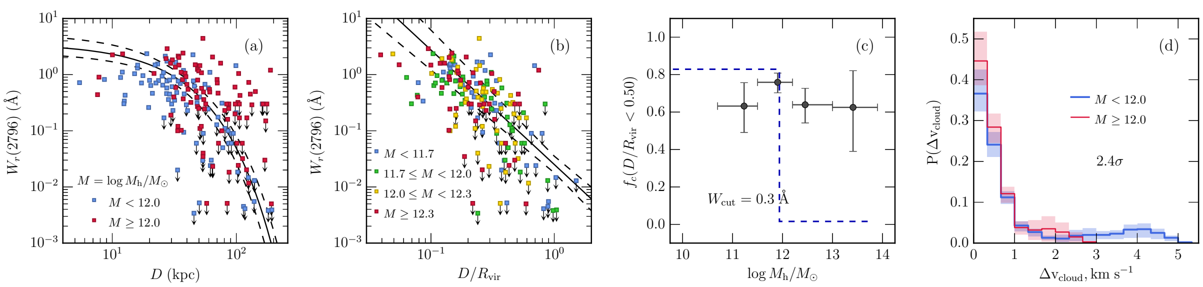

Figure 2: **Model Independent Testing of Baryon Cycle Physics** – (a) Mg II rest-frame equivalent width, $W_r(2796)$ vs impact parameter; higher mass halos have stronger absorption strengths in fixed impact parameter ranges. (b) $W_r(2796)$ versus $D/R_{\rm vir}$, where $R_{\rm vir}$ is the virial radius, showing $W_r(2796) \propto (D/R_{\rm vir})^{-2}$. (c) The covering fraction within $D/R_{\rm vir} = 0.5$ for $W_r(2796) \geq 0.3$ Å vs halo mass; covering fraction is independent of halo mass. The blue dotted line represents the theoretical prediction of cold-accretion. (d) The CGM velocity dispersion relative to the central galaxy (units of circular velocity) for low- and high-mass halos. Similar data for H I, C IV, and O VI absorption is abundant from studies of COS and STIS spectra.

- **Examining Model Dependent Absorption Line Analysis: New Approaches**

In addition to comparing the cause and effect response of the CGM to various baryon cycle physics in the simulations, we will focus on the comparison between the gas properties (density, temperature, ionization condition, metallicity, and kinematics) inferred from standard analysis techniques (Voigt profile, VP, fitting, chemical-ionization modeling) applied to "mock" absorption lines and the true underlying *distributions* of the gas properties responsible for the absorption in the CGM of the simulated galaxies.

Consider the following scenario: Some CGM absorption lines are observed in some spectra. VP fitting yields a set of column densities for various ions (kinematically complex absorption lines are modeled as multiple clouds with fixed column densities and isothermal temperatures). Photoionization models are run, constrained by the column densities, and metallicities are derived. Ultimately, gas properties are deduced that, for each VP component (cloud) a single metallicity, density, and perhaps an inferred abundance variation is quoted at each component velocity. This is what is published and employed for our understanding of the spatial distribution of the CGM as a function of galaxy property, projected distance, and redshift.

Now, simulate the above process with the absorption lines from the CGM of simulated galaxies in hydroART. Perform the above analysis just as an observer would and deduce the absorbing gas properties. Now, select out the gas cells from the simulation that give rise to the detected absorption lines (we have this capability), and examine the *distribution* of the "absorbing gas" metallicities, densities, temperatures, positions, and LOS velocities. How do they compare? Are there systematic offsets in the value (under- or over-estimates). Are the gas cells even physically contiguous (is the absorbing gas at a given velocity even come from the same physical location along the LOS)?

If Figure 1 is any indication of the nature of the true CGM, ions such as C IV and O VI are not "cloud-like" and thus the above described observational analysis methods would be failing at accurately describing the CGM. This is an example of an important result with far-reaching implications, but for which we may (or may not) be "fooled" by possible systematics lurking in our analysis methods. How can analysis be modified to better recover the underlying distributions and gradients of gas phase properties? How would doing so change our view of baryon cycle mixing in the CGM?

A more sophisticated modeling of absorption lines arising in gas with varying densities, velocity flows, and temperatures requires line formation with frequency redistribution, such as is employed for modeling expanding stellar winds. Examples of the mathematical formalism for such modeling is presented in Chapters 13 and 14 of Mihalas (1978, "Stellar Atmospheres", Freeman) and in Chapter 7 of Mihalas & Mihalas (1999, "Foundations of Radiation Hydrodynamics", Dover). We aim to adopt the formalism of Mihalas and apply it to absorption line formation in cosmological simulations as a way to extract more realistic insights into the density, velocity (flow and turbulence), temperature, and metallicity distributions of CGM gas, especially their LOS gradients. We have been exploring these methods for the last few years and are poised to demonstrate that they can be successfully applied to absorption line data such as COS and STIS spectra. The use of "mock" spectra through the simulation will be key for testing and calibrating the method. Preliminary studies indicate that transitions from multiple ions of a chemical species will be important. This is especially true for H I; the full Lyman series will be critical for estimating the metallicity distribution. Fortunately, each of the COS surveys includes spectra that have captured the full series.

■ Management Plan

All investigators will co-manage various components of the project and engage in interpreting results and writing papers. In alphabetical order:

Churchill will lead efforts to perform current Voigt profile fitting and chemical-ionization modeling methods with the goal of developing methods that more accurately recover the *distribution* of gas properties in the simulations from absorption line data.

Kacprzak is an expert on the absorption line galaxy connection and will engage in analyzing and comparing the observed and "mock" absorption line data to the results from *HST* programs. He will also lead efforts to quantify how the CGM varies as a function of baryon physics in the simulations.

Klypin is a primary architect of the hydroART code and an expert on feedback physics. He will be lead theorist, engaged in high-level interpretation of results and writing papers. He is a primary coder and will be responsible for running and managing the simulations (at NASA/Ames, NERSC, and SwinStar) and performing quality checks.

Nielsen will engage in analysis of the "mock" absorption line data, investigating the spatial-kinematic relationship within the absorption profiles and assisting with developing insights into current standard absorption line analysis methods.

Vander Vliet will run the absorption line pipeline and generate all "mock" spectra. He will engage in comparing simulated absorption line properties to both published observational data and the simulated gas properties. He is and expert at analyzing the simulations and will engage in quantifying the star formation rates and masses of the galaxies. He will analyze the baryon cycle for the various star formation and feedback models necessary for comparison with the CGM properties.

• **Two-Year Time-line**

We emphasize that we have all of the analysis software to generate the absorption line data from the simulations, extract the "absorbing gas" cells, and cull all these data into a convenient science ready format. We also have the analysis tools to perform the VP fitting and ionization modeling.

In Year 1, we will run the absorption line pipeline on galaxies of various masses and baryon cycle physics, generate absorption quantities, and measure simulated galaxy properties. All analysis will carefully reproduce realizations of HST survey characteristics (distribution of impact parameters, galaxy masses, star formation rates, redshifts, absorbing transitions/ions, spectral characteristics, detection sensitivities, etc.). In Year 2, we will compare the feedback recipes, engage in analysis of the baryon cycle physics and CGM properties, compare simulated "absorbing gas" properties to those inferred from VP fitting, and engage in exploring absorption line analysis methods that are not reliant on "cloud-based" assumptions to glean more accurate estimates of the distribution of CGM gas properties from *HST* spectra.

We anticipate three to four publications and that we will present our work at a minimum of two domestic and/or international conferences per year. The student and post-doc team members will be engaged in these conferences.