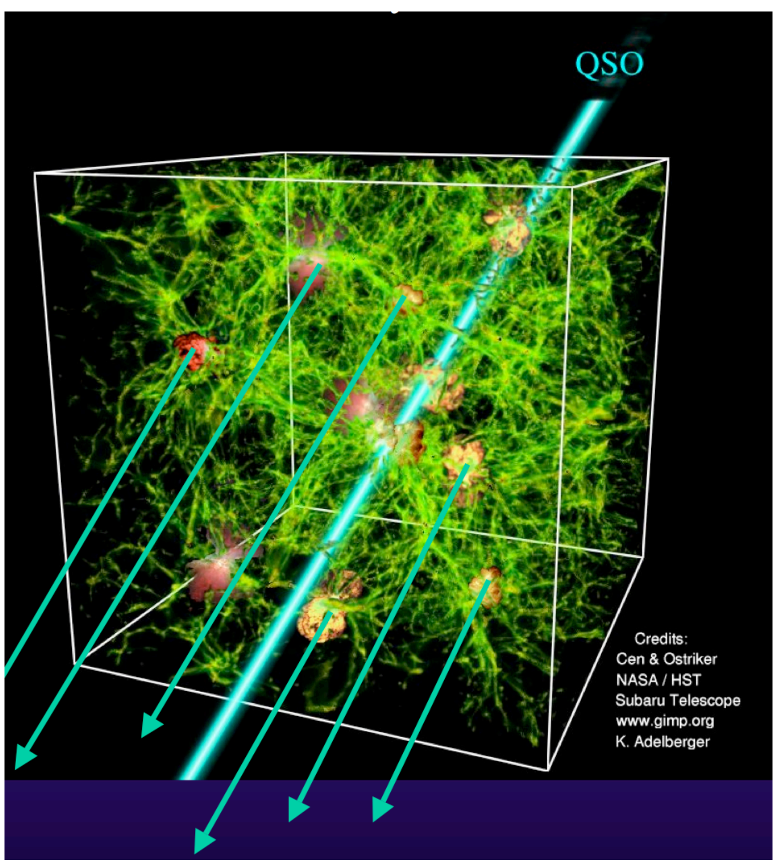
Mapping the “Cosmic Web” During the Peak Epoch of Galaxy Formation

Science White Paper submitted to the Astro2010 Survey
Science Frontier Panels: Galaxies Across Cosmic Time, Cosmology and Fundamental Physics

Authors: Charles Steidel[1] and Christopher Martin (Caltech), J. Xavier Prochaska (UCSC), Max Pettini (IoA), Joop Schaye and Olivera Rakic (Leiden)

Abstract: We outline the scientific opportunities enabled by the simultaneous study of galaxies and the intergalactic medium in the same survey volumes at high redshift ($1.6 \lesssim z \lesssim 3.5$). Combining two powerful techniques will provide complementary information on the state of baryons, both within and outside galaxies. The IGM presents a laboratory in which the effects of galaxy formation and AGN activity– radiative and hydrodynamical “feedback”, and gas flows into and out of forming galaxies– can be measured on scales not accessible using direct observations of the galaxies; similarly, galaxy observations in the same cosmic volume inform our interpretation of IGM absorption line information. The proposed opportunity will be within reach using the next-generation 30m-class O/IR telescopes; the gain in sensitivity over current capabilities enables a very high surface density of background IGM probes, and dense sampling of the galaxy distribution by reaching > 2 magnitudes fainter than L* throughout the cosmic epoch when galaxy formation peaked.



[1] e-mail: ccs@astro.caltech.edu

Background: Over the last decade, the statistical study of high redshift galaxies has flourished, with many different techniques contributing to establishing an overall census while revealing a new set of questions. Among the fruits of this enterprise has been a schematic picture of the global star formation history of the universe, at least over the last $\sim 90\%$ of cosmic time. It now appears relatively uncontroversial that the most active period of galaxy growth in the universe's history occurred in the redshift range $1.5 \lesssim z \lesssim 3.5$ (e.g., Reddy et al 2008), during which as much as 50% of the present-day stellar mass was formed. With the cosmological world model now a much less contentious issue from the standpoint of structure formation, the most significant remaining problems in understanding galaxy formation pertain to the baryon physics– not the distribution of dark matter– and how the energy produced by star formation, supernova explosions, and super-massive black hole accretion affect the subsequent development of structure on galaxy scales. Thus, there is a developing broad-brush picture of *when* galaxy formation is happening, and techniques exist for finding and collecting samples of most of the denizens of the high redshift universe, but we are lacking a detailed physical picture of *how* galaxies form. Further insight will require greater sensitivity and synthesis of techniques, such as we those we are suggesting below.

With the advent of high resolution spectrographs on 8-10m telescopes, QSO absorption line spectroscopy, coupled with increasingly sophisticated numerical simulations of structure formation) has revolutionized our view of the intergalactic medium (IGM). The beauty of this approach is that galaxies and other gaseous structures in the IGM are detected independently of their luminosity. Moreover, the information accessible in the absorption line spectra is of extremely high quality, with many orders of magnitude greater sensitivity than would be possible if the gas were required to be "self-luminous". For example, the Lyα line provides such a sensitive probe of neutral hydrogen gas that densities much less than the cosmic mean can be easily detected at redshifts $1.8 \lesssim z \lesssim 3.5$. We have learned that diffuse gas in the intergalactic medium may be a reliable tracer of the initial dark matter power spectrum on small scales that are not measurable using self-luminous objects (because sites of galaxy formation have become non-linear, thereby erasing the primordial signal.) We have also learned that much of the intergalactic gas, if not all of it, has been enriched by an early generation of star formation in galaxies; at $z \sim 3$, $\sim 90\%$ of the baryons in the universe reside in the diffuse intergalactic medium. Thus, the IGM acts as both a reservoir for fueling the galaxy formation process, and as a sink for receiving its products (e.g., metals, energy). The main limitation of the IGM studies is that QSOs bright enough to serve as background beacons are rare, and the information content along a single line of sight, while of extremely high quality, does not provide three-dimensional information nor show directly how the absorption signatures relate to luminous objects inhabiting the same (or nearby) regions of space.

A simultaneous study of galaxies and the IGM, in the same cosmic volumes during the extremely eventful $3.5 \geq z \geq 1.8$ era, offers the possibility of combining two powerful lines of investigation that provide complementary information on the state of baryons, both those collapsed into galaxies and those residing outside of galaxies. The IGM and "circum-galactic medium" (CGM–loosely defined as regions within $\lesssim 300$ kpc [physical] of galaxies) together present a laboratory in which the effects of galaxy formation and AGN accretion (e.g., radiative and hydrodynamical "feedback" and its recent history) can be measured on scales that are not accessible using direct observations of galaxies; similarly, the galaxy distribution relative to the lines of sight to background objects tells us more about how the physical information garnered from the absorption line studies should be interpreted.

The Big Questions: Over the last several years, there have been a number of surprising observations that would not have been predicted based solely on an understanding of the relevant cosmology and the distribution of dark matter. In hierarchical structure formation, gravitationally bound structures begin on small scales, and the largest bound objects are expected to be the last to form. A universal peak in overall star- formation and black hole accretion, as observed, is not necessarily expected. Neither is the observation that the most massive known galaxies appear to finish forming their stars earliest, or that typical galaxies at redshifts z$\sim 2-4$ appear to be forming stars at rates that are 10 - 100 times higher, within regions physically much smaller, than for typical galaxies in the local universe. The symbiotic relationship between the formation of galaxy spheroids and central black holes also does not follow directly from a cosmological framework. The complexity of galaxy formation requires following the behavior of dark matter and the "gastrophysics" of star formation, feedback of energy, accretion and expulsion of gas, and effects of environment. Among the remaining puzzles:

- Why does galaxy formation apparently proceed in an "anti-hierarchical" fashion, in spite of broad agreement that cosmological large scale structure develops hierarchically?
- How do galaxies acquire their gas ("cold flows", "hot mode accretion", mergers?), and how much cross-talk and exchange is there between galaxies and the IGM?
- What are the primary drivers of galaxy-scale star formation during the epoch of galaxy formation? What shuts it off? Is the mode of star formation analogous to anything in the local universe, or is it inherently different ?
- How do feedback processes (star formation, AGN activity, supernovae) influence galaxy formation and the subsequent evolution of both galaxies and the IGM?
- What is the root cause of apparent environmental dependencies of the galaxy formation process (large-scale over-densities, availability of gas supply, etc.)?
- To what extent can one use observable stars and gas to understand the underlying distribution of dark matter?
- Do super-massive black holes play a leading or subsidiary role in regulating the galaxy formation process?

Answers to these questions will depend on detailed studies of *both galaxies and gas* at high redshift. Key measurements include the distribution and metallicity of gas inside and outside of galaxies, the dynamics and physical conditions (e.g. temperature, density) of the gas, the star formation rates, internal structure, stellar and gaseous content, and dynamical states of galaxies, all as a function of environment and of cosmic time. Complete understanding will require continued development of state-of-the-art hydrodynamic simulations with all relevant physics incorporated in realistic ways.

Methodology: Somewhat ironically, a large fraction of what we currently know about galaxies at redshifts $z > 1.5$ comes from observations made in their rest-frame UV light. The main reason for this is practical - it is easy to identify high redshift (star forming) galaxies based on their distinctive far-UV colors as the Lyman limit of hydrogen passes through the observed passbands, and once flagged, the naturally dark terrestrial background in the optical (defined for the present purposes as $0.31 - 1\mu$m) makes spectroscopy of faint objects feasible for at least the brightest examples with 8 - 10m telescopes. Practicalities aside, the rest-UV spectra of galaxies contain a vast amount of information because of the very large number of ground-state transitions of astrophysically abundant ions at these wavelengths (see Fig. 1) At the redshifts of interest, $z \sim 1.5 - 3.5$, essentially all galaxies spectroscopically observed to date have extremely strong, broad absorption lines which arise in interstellar gas associated with the galaxy itself, stellar wind P-Cygni lines of CIV, NV, HeII, and SiIV from O-stars, as well as O and B star photospheric features. With sufficiently high quality data (as illustrated below), one can extract information on the stellar initial mass function (IMF) shape at the high mass end, stellar chemical abundances, the chemistry of the interstellar gas, and the kinematics of gas motions within the galaxy. For redshifts $z > 2.5$, the hydrogen Lyman limit is redshifted above the atmospheric cutoff so that a direct measurement of the ionizing radiation escaping from the galaxy is possible (highly relevant for the physics of the IGM, including the process of re-ionization), and for redshifts $z > 1.6$, all galaxy spectra will contain information on the "Lyman α forest" and associated lines of heavier elements from the IGM along the line of sight (see Fig.1).

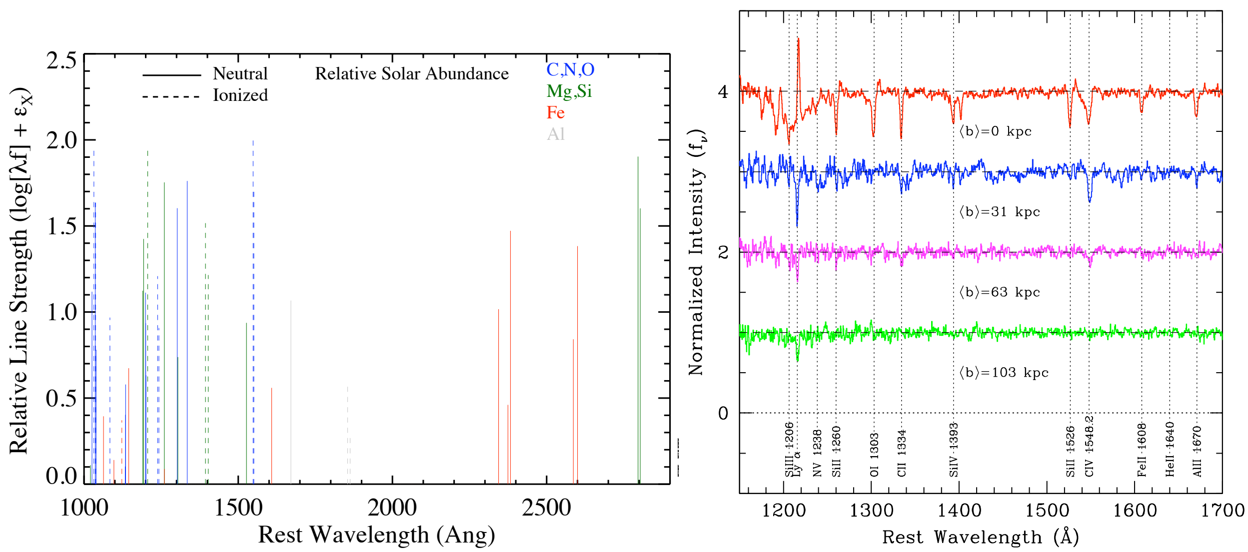


Figure 1: (Left) The product of cosmic abundance and oscillator strength for transitions of the most abundant elements in the universe (excluding H and He). Note the concentration of ground-state transitions between 1000 and ~ 1700 Å. **(Right)** Crude measures of "CGM" absorption as a function of distance from star-forming galaxies, based on many pairs of $z = 1.8 - 3.5$ galaxies with small angular separations but different redshifts. Each spectrum is the median of all background galaxy spectra after being shifted to the foreground galaxy redshift, in bins of angular separation, corresponding to median values of the projected impact parameters as indicated below each composite. The spectra were all obtained with a 10m telescope at $\mathcal{R} \simeq 800$, and individually have low S/N, but serve to illustrate the potential of using galaxies as IGM/CGM probes.

A relevant example of what is possible is shown in Fig. 2 - the z$= 2.723$ gravitationally-lensed galaxy MS1512-cB58 is one of a handful of high redshift galaxies bright enough to obtain spectral resolution $\mathcal{R} = 5000$ spectra using an 8m-class telescope. From this spectrum alone, one obtains a measure of the stellar metallicity (~ 0.4 solar); constraints on the slope of the high mass end of the IMF (top-heavy mass functions appear ruled out); the interstellar abundances of 8 different elements, showing clear enhancement of alpha elements, and a N abundance indicating that cB58 is very chemically young; velocity profiles showing the clear presence of a galaxy-scale outflow, with blue-shifted gas-phase velocities of up to ~ 800 km s^{-1} (relative to the systemic redshift) , and an inferred mass outflow rate of the same order as the star formation rate ($\gtrsim 50$ solar masses per year). Unfortunately, the typical spectra of even bright (but un-lensed) star forming galaxies at similar and higher redshifts do not approach the quality of the cB58 spectrum, and generally many spectra (of separate objects) must be co-added in order to measure any spectral feature with reasonable S/N (e.g., Shapley et al. 2003; see Fig. 1). The problem is that the characteristic apparent magnitude (i.e. L*) of galaxies at z$= 2$, 3, 4, and 5 are R$= 23.9$, 24.5, 25.0, and 25.6, respectively - whereas cB58

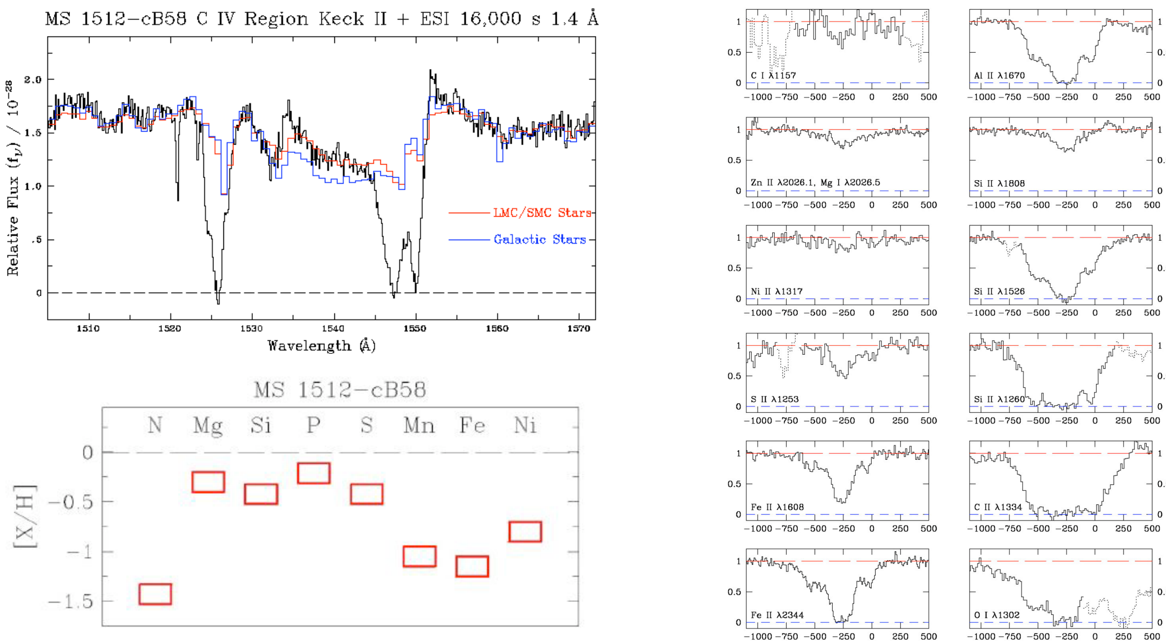


Figure 2: High-resolution ($\mathcal{R} \simeq 5000$) spectroscopy of MS1512-cB58. **(Left, top)** The CIV (1549 Å) P-Cygni profile in the Keck/ESI spectrum of cB58. The black line indicates the observed spectrum, while the red and blue lines indicate Starburst99 constant star-formation model fits to the P-Cygni profile using LMC/SMC and solar metallicity stars, respectively. The large absorption doublet within the P-Cygni profile is from interstellar CIV. The P-Cygni profile rules out instantaneous burst models for the star-formation history, and favors a Salpeter IMF at the high-mass end (Pettini et al 2000). **(Left, bottom)** The interstellar abundance pattern in cB58. This pattern reveals the enhancement of alpha elements (Mg, Si, P, S) relative to both Fe-peak elements (Mn, Fe, Ni) and nitrogen, indicative of preferential enrichment by Type II supernovae in a young stellar population (from Pettini et al. 2002). **(Right)** Velocity profiles of low-ionization interstellar absorption lines tracing predominantly neutral H I gas. Velocities are plotted relative to the stellar systemic redshift. Outflowing neutral gas extends to velocities of -750 km s^{-1}, indicating that it will most likely escape the potential well of cB58 (Pettini et al. 2002).

has an apparent magnitude of R$= 20.5$, ~ 4 magnitudes brighter than its unlensed ($\sim$L*) magnitude of R$= 24.4$.

The Epoch of Galaxy Formation in 3-D: Because the physics of diffuse gas can be very simple, theory has told us that the neutral hydrogen (HI) optical depth, as well as the gas temperature, is modulated by the gas density and the intensity and spectral shape of the metagalactic ionizing radiation field. This (assumed) simplicity makes it possible to turn a Lyman α forest spectrum into a one-dimensional map of the density along the line of sight, as a function of redshift. Because high resolution spectra of QSOs have the remarkable capability to detect HI column densities as low as $\lesssim 10^{12}$ cm^{-2} (associated with regions of the Universe that are *below* the mean density), the Ly-α forest may offer the best means of measuring the spectrum of these perturbations on small scales $1 - 10$h^{-1} Mpc (co-moving) where other techniques cannot (e.g., Croft et al 1998; McDonald et al 2006).

The highest quality data on the physical properties of the IGM and CGM have been recorded for the redshift range $z = 1.6 - 3.5$, for several reasons: first, Ly-α must be redshifted above the atmospheric UV cutoff of 0.31μm, setting the minimum redshift; second, the Ly-α forest evolves extremely rapidly with redshift, so that by $z > 3.5$, it is so dense with absorption that it loses dynamic range for measuring weak spectral features; third, QSOs bright enough for echelle observations on 8m-class telescopes ($m < 19$) are very rare even at their peak near $z \simeq 2.5$, and by $z = 6$ there is only a handful currently known in the entire sky. Thus, in general, our current view of the

IGM remains "one-dimensional", and the connection between the "3-D" galaxy distribution and the CGM/IGM remains relatively unexplored.

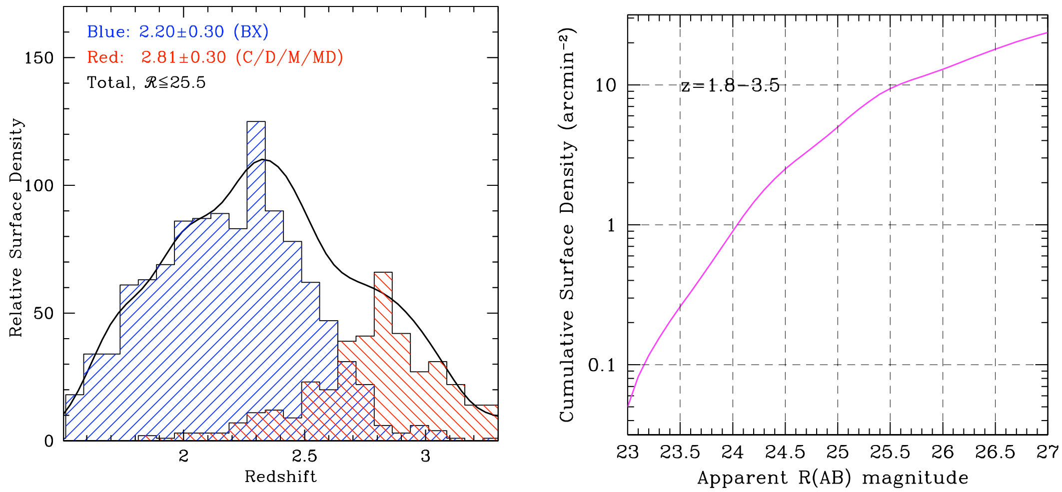


Figure 3: (Left) The relative surface density versus redshift for rest-UV selected galaxy samples to a fixed apparent magnitude limit (based on spectroscopic samples observed to R(AB)= 25.5). The relative distribution of objects to R(AB)= 26.5 is expected to be similar. The heavy black curve is the sum of the two sub-samples. **(Right)** The cumulative surface density of galaxies in the targeted redshift range $1.8 \lesssim z \lesssim 3.5$ as a function of apparent magnitude, based on measured rest-frame far-UV luminosity functions obtained from existing large spectroscopic samples (Steidel et al 1999, 2004; Reddy et al 2008).

One of the most exciting possibilities within reach in the foreseeable future is to combine the precision measurements of the astrophysics of the IGM using relatively bright background sources with direct observations of the luminous material - galaxies and AGN - in the same volumes of space, providing for the first time a densely-sampled three-dimensional view of the distribution of baryonic material (whether or not it is "self-luminous") in the high redshift universe. The scarcity of suitably bright background sources has prevented anything more than initial forays into learning about the IGM/CGM structure transverse to the line of sight. The situation is very different when the light-gathering power of a 30m-class telescope is brought to bear on the problem. While the surface density of QSOs with apparent magnitudes suitable for IGM studies increases significantly toward fainter apparent magnitudes, the space density of compact, UV-bright star forming galaxies sufficiently bright to yield very high quality spectra easily overtakes QSOs (by number) as background sources. For galaxies with $R \simeq 24.5$ ($\sim 0.6\mu$Jy), approximately the limiting apparent magnitude (for a 30m aperture telescope) to which one can obtain a spectrum at $R \sim 5000$ with $S/N \gtrsim 30$, galaxies in the appropriate redshift range $(z = 1.8 - 3.5)$ outnumber QSOs by more than a factor of 50. More importantly, the surface density of usable background sources increases by more than 2 orders of magnitude over the current situation for 8m-class telescopes, exceeding 2 arcmin^{-2} (see Fig. 3). This means that the IGM properties can be densely sampled on physical scales of $\lesssim 300$ kpc ($\simeq 1$ Mpc co-moving), approximately the maximum sphere of influence of individual galaxies on the IGM/CGM and comparable to the expected coherence length of the undisturbed IGM (e.g., Adelberger et al 2003, 2005; see also Fig. 4). Thus, the 3-D IGM can be effectively reconstructed "tomographically" using a dense grid of background sources over the range $1.8 < z < 3.5$ where the Lyman-α forest can be observed from the ground with good dynamic range.

Most importantly, the observations will reveal the inner workings of many of the most poorly–understood physical processes that must be incorporated in order to understand how galaxies form

- feedback (both hydrodynamic and radiative) from AGN and galaxies, gas accretion, details of the relationship between structure in diffuse gas versus that traced by galaxies. Once these "complications" are understood, the same data could be used successfully for measurement of the small-scale matter power spectrum with the Ly-α forest– with the ability to remove the systematic uncertainties from line-of-sight peculiar velocities which currently limit the precision of such measures.

Addressing the Questions: What is Needed?– As shown in Fig. 1 Access to a very high density of diagnostic absorption features in the IGM and CGM requires excellent spectroscopic sensitivity from the atmospheric cutoff at 3100 Å to $\simeq 7400$ Å if one were to concentrate on the $1.8 \lesssim z \lesssim 3.5$ era. As shown in Fig. 3, it is straightforward to select galaxies throughout this redshift range using well-established photometric criteria (of course, other criteria may be used in order to make the galaxy sample more "complete"), and the surface density of galaxies is extremely high ($>$ 20 arcmin^{-2}) to R(AB)$= 26.5$, which would be an easily achievable limit for identification-quality ($\mathcal{R} \sim 1000$) spectra for a 30m-class telescope equipped with a high-efficiency optical spectrometer. Higher quality spectra ($\mathcal{R} \simeq 5000$), with the full diagnostic power shown in the cB58 example above, would be possible for AGN and galaxies to R(AB)$= 24.5$, which have a surface density of at least 2.5 arcmin^{-2}. It is worth emphasizing that this dense grid of background rest-UV-bright sources could be used to map the IGM and CGM in *any region of the sky*, to complement observations of dense molecular gas (e.g., with ALMA), starlight, potential high redshift galaxy clusters, X-ray sources, IGM emission, etc. The absorption line probes are sensitive to gas in the range $T \simeq 10^4 - 10^6$ K– this would include all of the dominant photo-ionized IGM, and all but the very hottest shock-heated gas in the CGM.

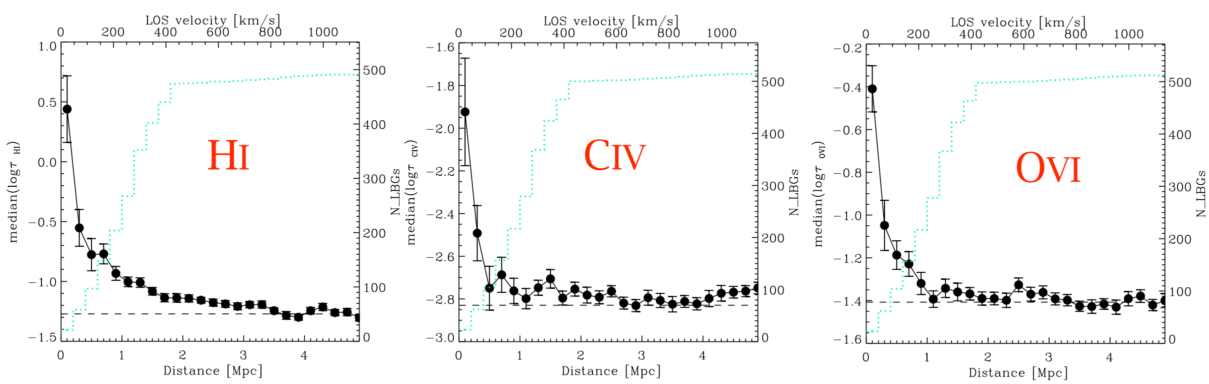


Figure 4: The median optical depth τ in the Lyα, CIV, and OVI transitions as a function of physical distance from star-forming galaxies, using the "pixel optical depth" method. These results show the statistical median CGM/IGM from a "galaxy-centric" point of view, for a sample of $z \sim 2.5$ galaxies. The light green histogram shows the number of galaxies contributing to the $\tau(d)$ curves, with the axis label on the right; the dashed lines show the median optical depth of all pixels in the bright QSO spectra over the redshift range of the galaxy survey.

While some progress in making quasi-tomographic observations of the IGM can be made with current-generation 8-10m telescopes (e.g., see Figs. 1, 4), it is clear from experience that larger telescope apertures equipped with very efficient multi-object spectrographs will be necessary to reach the high surface density "grid" with spectral resolution $R \simeq 5000$ required for quantitative analysis with adequate spatial resolution.

The same capabilities, i.e., 30m aperture with high-throughput spectrograph, would make possible very ambitious surveys of the high redshift universe that would be unprecedented in the quantity of information delivered. As an example, consider a hypothetical survey designed to cover a volume of the $z = 1.8 - 3.5$ universe that is as statistically representative as SDSS has been for the local universe. The relationship between angular scale on the sky and co-moving scale at the targeted cosmic epoch is vastly different between $z \sim 0.1$ (SDSS) and $z \simeq 2.5$. The SDSS was carried out with a telescope+instrument combination capable of observing over a field of view with a diameter of 2.5 degrees, or a transverse scale of $\simeq 18$ Mpc (co- moving) at the median redshift of the survey. At $z \simeq 2.5$, the same 18 Mpc co-moving scale is subtended by an angle of only 10.6 arcmin. Thus, a "wide-field" capability would be achieved with angular coverage feasible for imaging spectrometers on next-generation 30m-class telescopes. A survey of a representative volume of the Universe ($\sim 10^8$ co-moving Mpc^3) covers a solid angle of $\simeq \pi$ steradians at $z \sim 0.1$; for the proposed $z \sim 2.5$ baryonic structure survey, the same volume is covered by only $\simeq 4$ deg^2 degrees on the sky. Within a total survey area of 4 deg^2 there would be $\sim 650,000$ star-forming galaxies with R(AB) $\lesssim 26.5$ (0.08 μJy) in the redshift range $1.8 < z < 3.5$ that could be selected for spectroscopy using simple photometric criteria such as those used in Fig. 3. This apparent magnitude limit reaches $L < 0.12L^*$ at all redshifts in the range $1.8 \leq z \leq 3.5$. . The total number of targets with R(AB)< 24.5 and $1.8 < z < 3.5$ in the same volume would be $\simeq 30,000$.

The products of such a survey would include: 1) Identification in redshift space of 1000 over-dense regions that will become clusters by the present- day. The physical state of potential hot gas in the proto-intracluster media can be matched against Sunyaev-Zeldovich signatures in future high resolution CMB maps, providing a complete census of baryons in all phases within the densest regions in the Universe. 2) Exquisite far-UV spectra of a large number of galaxies in the same redshift range for which near-IR spectrographs can obtain rest-frame optical spectra with additional diagnostic capabilities. The far-UV spectra will provide measures of outflow kinematics, chemistry, stellar IMF, and in some cases mass outflow rate (e.g., Rix et al., 2004; Steidel et al., 2004). 3) The 30,000 high quality sightlines through the IGM will map intergalactic HI and metals in 3-D, to be compared with the galaxy distribution in the same cosmic volumes (e.g., Adelberger et al., 2003, 2005). The kinematics, chemistry, and physical state of CGM gas will track inflow and outflow throughout the survey volume; comparison of the properties of the widespread gas with respect to survey galaxies can be traced as a function of large-scale over-density and galaxy properties, and the influence of individual galaxies and AGN on the nearby IGM/CGM will directly measure where metals have been deposited by galaxy super-winds or AGN-driven outflows. The metallicity of galaxies can be compared with the metal content of gas in their environs. 4) The extremely densely sampled fainter galaxies will provide still finer spatial resolution for mapping gaseous material outside of galaxies (25 times better spatial sampling on all scales than possible with current facilities), and may allow the mapping of inhomogeneities in the UV ionizing radiation field and measurement of the lifetime of bright UV sources via the transverse proximity effect (e.g., Adelberger, 2004).

If carried out, the survey volume would be comparable to any spectroscopic survey to date (at any redshift), and would provide information far beyond mere "galaxy mapping"– the detailed census of IGM and CGM gas throughout the survey volume would be absolutely unique compared to *any other survey*. This census is an essential ingredient to understanding galaxy formation.

References:

Adelberger, K.L., & Steidel, C.C 2000, ApJ,544,218
Adelberger, K.L. et al. 2003, ApJ, 584, 45
Adelberger, K.L. 2004, ApJ, 612, 706
Adelberger, K.L. et al. 2005, ApJ, 619, 697
Croft, R.A.C., et al 2006, ApJ, 495, 44
McDonald, P., et al 2006, ApJS, 163, 80
Pettini, M., et al. 2002, ApJ, 569, 742
Pettini et al. 2000, ApJ, 528, 96.
Pettini et al. 2002, ApJ, 569, 742.
Reddy, N.A. et al 2008, ApJS, 175, 48
Rix, S.M. et al. 2004, ApJ, 615, 98.
Shapley, A. E., et al 2003, ApJ, 588, 65
Steidel et al. 1999, ApJ, 519, 1
Steidel et al. 2003, ApJ, 592, 728