UC Berkeley
UC Berkeley Previously Published Works

Title
OBSERVATIONAL REQUIREMENTS FOR Lyα FOREST TOMOGRAPHIC MAPPING OF LARGE-SCALE STRUCTURE AT z ~ 2

Authors
Lee, Khee-Gan
Hennawi, Joseph F
White, Martin
et al.

OBSERVATIONAL REQUIREMENTS FOR LYMAN-α FOREST TOMOGRAPHIC MAPPING OF LARGE-SCALE STRUCTURE AT z ∼ 2

Khee-Gan Lee[1], Joseph F. Hennawi[1], Martin White[2,3], Rupert A.C. Croft[4], Melih Ozbek[4]
Accepted by ApJ

ABSTRACT

The $z \gtrsim 2$ Lyα forest traces the underlying dark-matter distribution on large scales and, given sufficient sightlines, can be used to create 3D maps of large-scale structure. We examine the observational requirements to construct such maps and estimate the signal-to-noise as a function of exposure time and sightline density. Sightline densities at $z = 2.25$ are $n_{\rm los} \approx [360, 1200, 3300] \, \mathrm{deg}^{-2}$ at limiting magnitudes of $g = [24.0, 24.5, 25.0]$, resulting in transverse sightline separations of $\langle d_\perp \rangle \approx [3.6, 1.9, 1.2] \, h^{-1} \, \mathrm{Mpc}$, which roughly sets the reconstruction scale. We simulate these reconstructions using mock spectra with realistic noise properties, and find that spectra with S/N ≈ 4 per angstrom can be used to generate maps that clearly trace the underlying dark-matter at overdensities of $\rho/\langle\rho\rangle \sim 1$. For the VLT/VIMOS spectrograph, exposure times $t_{\rm exp} = [4, 6, 10] \, \mathrm{hrs}$ are sufficient for maps with spatial resolution $\epsilon_{\rm 3D} = [5.0, 3.2, 2.3] \, h^{-1} \, \mathrm{Mpc}$. Assuming $\sim 250 \, h^{-1} \, \mathrm{Mpc}$ is probed along the line-of-sight, $1 \, \mathrm{deg}^2$ of survey area would cover a comoving volume of $\approx 10^6 \, h^{-3} \, \mathrm{Mpc}^3$ at $\langle z \rangle \sim 2.3$, enabling efficient mapping of large volumes with 8-10m telescopes. These maps could be used to study galaxy environments, detect proto-clusters, and study the topology of large-scale structure at high-z.

Subject headings: cosmology:observations — galaxies:high-redshift — intergalactic medium — quasars: absorption lines — surveys — techniques:spectroscopic

1. INTRODUCTION

Cosmography – the cartography of the Universe – has been a primary goal of astronomical surveys ever since the first telescopic attempts to map the structure of the Milky Way (Herschel 1785). These efforts were naturally limited to our galaxy until the nearly contemporaneous discoveries of the extragalactic Universe and cosmological redshifts (Hubble 1926, 1929), which expanded the milieu of interest towards the large-scale structure of the galaxy distribution. However, the cartography of 3D large-scale structure beyond the $d \lesssim 10 \, \mathrm{Mpc}$ Local Universe remained challenging due to the inefficiency of obtaining galaxy redshifts with photographic plates: for example, the 920 redshifts listed in Humason et al. (1956) represented a heroic effort, spanning two decades, to obtain $5 - 10\,\mathrm{hr}$ exposures on $V \sim 12$ galaxies. The advent of photoelectric detectors in astronomical spectrographs precipitated a quantum leap in cosmography starting in the early 1980s, with progressively larger galaxy redshift surveys ranging from CfA Redshift Survey (Davis et al. 1982; Geller & Huchra 1989), Las Campanas Redshift Survey (Shectman et al. 1996), through to the 2dFGRS (Colless et al. 2001), SDSS-I and -II (Abazajian et al. 2009), and GAMA (Driver et al. 2011) surveys that have now comprehensively mapped out the galaxy distribution out to $z \sim 0.3$. The resulting 3D galaxy distri-

butions from these surveys have revealed the beautiful filamentary 'cosmic web' structure agreeing with theoretical predictions for a cold dark matter (CDM) universe, and forms one of the pillars of the current cosmological paradigm. Apart from the statistical clustering measurements (e.g., Tegmark et al. 2004; Cole et al. 2005; Reid et al. 2012), the maps from galaxy redshift surveys have enabled the study of galaxy properties in the context of large-scale environment (e.g., Lewis et al. 2002; Gómez et al. 2003; Kauffmann et al. 2004; Blanton et al. 2005).

However, since galaxy surface brightness scales with redshift as $\propto (1 + z)^{-4}$, it becomes increasingly expensive to obtain redshifts for higher-redshift galaxies. Various redshift surveys are underway to measure samples of tracer galaxies (e.g. luminous red galaxies or emission-line galaxies, see Hill et al. 2008; Drinkwater et al. 2010; Ahn et al. 2012; Comparat et al. 2013) to measure large-scale ($\gtrsim 20 \, h^{-1} \, \mathrm{Mpc}$) clustering out to $z \sim 1$, but it is considerably more difficult to obtain volume-limited samples of galaxies at sufficient space densities to directly map out large-scale structure. Even with 8-10m-class telescopes, redshift surveys capable of creating maps with resolutions of $\sim$ several Mpc out to $z \sim 1$ have been limited to elongated 'pencil-beam' geometries over $\sim 1 \, \mathrm{deg}^2$ areas (e.g., Davis et al. 2003; Lilly et al. 2007; Le Fèvre et al. 2013). Similar maps out to $z \sim 2$ would require volume-limited samples of $\mathcal{R} > 25$ galaxy redshifts requiring very long ($t_{\rm exp} \sim 10\,\mathrm{hrs}$) exposures on 8-10m telescopes in order to obtain secure redshifts from intrinsic absorption lines[5]. Such surveys over cosmolog-

lee@mpia.de
[1] Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany
[2] E.O. Lawrence Berkeley National Lab, 1 Cyclotron Rd., Berkeley, CA, 94720, USA
[3] Department of Astronomy, University of California at Berkeley, B-20 Hearst Field Annex # 3411, Berkeley, CA 94720
[4] Department of Physics, Carnegie-Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213

[5] Galaxy redshifts up to $\mathcal{R} \leq 25.5$ have been obtained with $t_{\rm exp} \sim 2 - 4\,\mathrm{hrs}$ (e.g., Steidel et al. 2004; Lilly et al. 2007), but at the faint-end most of the successful redshifts were from objects that
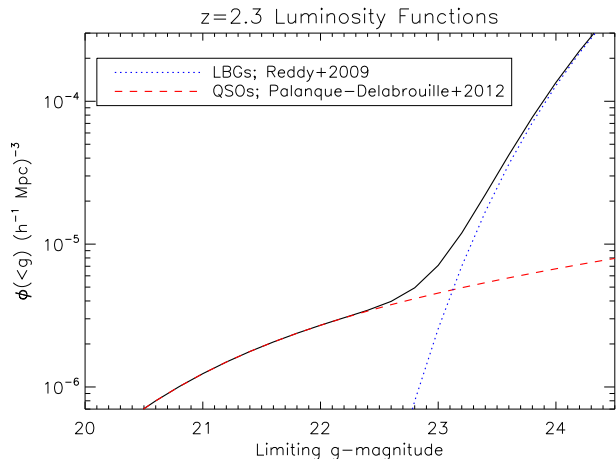
FIG. 1.— Luminosity functions used in this paper, shown here as a function of limiting g-magnitude. The red dashed curve shows the QSO luminosity function from Palanque-Delabrouille et al. (2013), while the dotted blue curve is the LBG luminosity function from Reddy et al. (2008). The solid curve shows the sum of both luminosity functions.

ically interesting volumes would be extremely expensive on existing instrumentation, and will likely be only feasible with future 30m-class telescopes.

At $z \gtrsim 2$, an alternative probe of large-scale structure is the Lyman-α (Lyα) forest absorption in the sightline of distant background sources (Lynds 1971). With the insight that the Lyα absorption in the photoionized intergalactic medium (IGM) comes from residual H I that directly traces the underlying dark matter overdensity (Cen et al. 1994; Bi et al. 1995; Zhang et al. 1995; Hernquist et al. 1996; Miralda-Escudé et al. 1996; Bi & Davidsen 1997), the Lyα forest has been established as the premier probe of the IGM and structure $z \gtrsim 2$ universe at density ranges near the cosmic mean. While the Lyα forest has traditionally been treated as 1-dimensional probes of the IGM (Croft et al. 1998, 2002; McDonald et al. 2000, 2005, 2006; Zaldarriaga et al. 2003; Viel et al. 2004), the BOSS Lyα Forest Survey (Lee et al. 2013) has recently measured Lyα forest correlations across multiple quasar lines-of-sight to constrain the large-scale baryon acoustic oscillations (BAO) scale at $\langle z \rangle \approx 2.3$ (Busca et al. 2013; Slosar et al. 2013; Kirkby et al. 2013).

The projected sightline density of the BOSS Lyα forest quasars is $\approx 17 \, \mathrm{deg}^{-2}$ at a limiting apparent magnitude of $g_{\mathrm{lim}} \approx 21.5$ (Bovy et al. 2011; Ross et al. 2012; Pâris et al. 2012). While the number of quasars per square degree increases with limiting magnitude, the slope of the quasar luminosity function is relatively flat (e.g., Palanque-Delabrouille et al. 2013) — this is illustrated in Figure 1. The corresponding area density of quasars at limiting magnitudes of $g_{\mathrm{lim}} \approx 23$ is $\sim 50 \, \mathrm{deg}^{-2}$ at $2.2 \lesssim z \lesssim 3$. At around $g \sim 23$, however, star-forming Lyman-break galaxies (LBGs[6]) begin to dominate the overall UV luminosity function (Figure 1). Furthermore,

the LBG luminosity function (Reddy et al. 2008) is considerably steeper than its quasar equivalent — at the $g_{\mathrm{lim}} \sim 24.5$ magnitude limit corresponding to L_* galaxies at $z \sim 2.5$, the projected area density of LBGs is $\gtrsim 1000 \, \mathrm{deg}^{-2}$ over $2.2 \gtrsim z \gtrsim 3$.

At sightline densities of $\sim 1000 \, \mathrm{deg}^{-2}$, the typical projected separation between the sightlines is $\sim 2'$; at $z \sim 2$, this corresponds to transverse comoving distances of $\sim 2 \, h^{-1} \, \mathrm{Mpc}$. It then becomes possible to combine the transverse sampling with the line-of-sight absorption to generate a 3D tomographic map of the Lyα absorption field with a spatial resolution corresponding roughly to the sightline separation[7]. Since each individual spectrum typically samples the Lyα forest over $400-500 \, h^{-1} \, \mathrm{Mpc}$ along the line-of-sight, this method has the potential to efficiently survey large volumes at $z \gtrsim 2$. Lyα forest tomography is not a new idea: Pichon et al. (2001) presented a Wiener reconstruction formalism for this problem, while Caucci et al. (2008) applied this formalism to numerical simulations to show that the technique can recover the large-scale topology of the dark matter field at $z \sim 2$.

It is often assumed that the spectra of $g \gtrsim 24$ LBGs required for Lyα forest tomography cannot be be obtained with the current generation of telescopes, and will require 30m mirror apertures. In the literature, we have encountered rather formidable observational requirements for IGM tomography, e.g. Steidel et al. (2009) specifies S/N ~ 30 per pixel at $R = 5000$ for $R = 24.5$ sources, while Evans et al. (2012) calls for S/N ≥ 8 per resolution element for $r = 24.8$ sources. Such spectra would require exposure times of the order $t_{\mathrm{exp}} \sim 10$hrs even on 30m-class telescopes. However, we could not find detailed justification for such steep requirements, neither for the resolution nor the signal-to-noise.

However, according to the fluctuating Gunn-Peterson picture of the IGM (e.g., Croft et al. 1997; Rauch et al. 1997; Croft et al. 1998), Lyα absorption is a non-linear tracer of the smoothly-varying dark matter distribution. Therefore, large-scale structure measurements on comoving scales of $\gtrsim 1 \, h^{-1} \, \mathrm{Mpc}$ do not have to resolve individual Lyα forest lines, and the individual spectra can be noisy if appropriate noise weighting is implemented in the analysis (McDonald et al. 2006; McDonald & Eisenstein 2007; McQuinn & White 2011). This is borne out by the success of the BOSS Lyα Forest Survey, in which the typical spectrum has S/N ~ 2 per angstrom within the Lyα forest (Lee et al. 2013).

It is therefore the purpose of this paper to examine in detail the actual requirements necessary to carry out Lyα forest tomographic mapping at spatial resolutions of $\sim 1-5 \, h^{-1} \, \mathrm{Mpc}$. We will first evaluate the availability of absorption sightlines at various magnitude limits and redshifts (§ 2.1). In § 3, we directly carry out tomographic reconstructions on mock spectra derived from numerical simulations of the Lyα forest, in which we have included instrumental effects reflecting various choices of sightline density and exposure times. As we shall show, Lyα forest tomography on scales of $\sim 2-5 \, h^{-1} \, \mathrm{Mpc}$ is already feasible on the current generation of 8-10m

---

had emission lines. Obtaining redshifts from *complete* samples that include non-emission line galaxies, require much longer exposures.

[6] Technically, 'LBGs' refers to star-forming galaxies selected at $z \gtrsim 3$ by their $\lambda < 912 \, \text{Å}$ Lyman-limit absorption, but we use the term to cover all $z \gtrsim 2$ galaxies with a star-formation far-UV continuum.

[7] Although this can also be done with smaller sightline densities, e.g. M. Ozbek et al (in prep) is currently working on a tomographic map using the BOSS data.

telescopes with reasonable exposure times. We also discuss an analytic model for the reconstruction signal-to-noise (§ 3.3), that allows us to quickly explore the effect of various exposure times and sightline densities on the reconstructed maps. In § 4 we will discuss possible science applications for Lyα tomography and potential survey strategies on various existing and proposed spectrographs.

We assume a flat ΛCDM cosmology with $\Omega_m = 0.274$, $\Omega_\Lambda = 0.726$ and $H_0 = 70 \text{ km s}^{-1}\text{Mpc}^{-1}$.

2. SOURCE AVAILABILITY

In this section, we first examine the availability of background sightlines for Lyα forest tomography using the published luminosity functions for LBGs and QSOs, followed by an analytic estimate of the observational requirements for Lyα forest tomography.

2.1. *Estimating Sightline Densities*

The critical observational parameter for 3D Lyα forest tomographic mapping is the area density of effective background sightlines, $n_{\rm los}$. In the conclusions of their paper, Caucci et al. (2008) tabulated the projected area density of $z \gtrsim 2$ background quasars and LBGs at several observed magnitude limits. However, this quantity is *not* identical to $n_{\rm los}$, since the finite Lyα forest path length in each background sightline (typically $1040 \text{ Å} \lesssim \lambda_{\rm rest} \lesssim 1180 \text{ Å}$) means that only a subset of these sightlines will pierce the volume at a given foreground redshift.

The effective number of sightlines per unit area that actually probe any given absorption redshift z and limiting magnitude $m_{\rm lim}$ is given by the differential sightline density

$$n_{\rm los}(z, m_{\rm lim}) = \int_{z_1}^{z_2} dz_{\rm bg} \int_{\infty}^{m_{\rm lim}} dm \frac{dl_c}{dz_{\rm bg}} \phi(z_{\rm bg}, m) \ (1)$$

where $\phi(z_{\rm bg}, m)$ is the luminosity function of background sources at redshift $z_{\rm bg}$ and apparent magnitude m, dl_c is the comoving line-element along the line-of-sight. The integration limits on $z_{\rm bg}$ take into account the finite lengths of Lyα forest in each sightline that could intersect the absorption redshift z: $1 + z_2 = (\lambda_\alpha/\lambda_{\rm rest,min})(1 + z)$ and $1 + z_1 = (\lambda_\alpha/\lambda_{\rm rest,max})(1 + z)$ where $\lambda_\alpha = 1215.7 \text{ Å}$ is the Lyα absorption wavelength and $[\lambda_{\rm rest,min}, \lambda_{\rm rest,max}] \approx [1040 \text{ Å}, 1180 \text{ Å}]$ is the range of useful Lyα forest wavelengths in each sightline. This $1040 \text{ Å} \leq \lambda_{\rm rest} \leq 1180 \text{ Å}$ range is chosen to avoid the quasar proximity zone as well as to avoid having to predict the shape of the Lyβ emission line (McDonald et al. 2006; Lee et al. 2013); these criteria might not apply for LBGs, but we use it as a conservative value.

We estimate $n_{\rm los}$ by integrating over the published luminosity functions for both LBGs and QSOs, illustrated in Figure 1. For the LBG contribution, we use the Schechter function fits from Table 14 in Reddy et al. (2008), evaluated in the redshift range $1.9 < z < 2.7$. This luminosity function was evaluated in $m_{\mathcal{RG}}$, a composite bandpass between $\mathcal{G}$ and $\mathcal{R}$ (in the custom filter set of Steidel & Hamilton 1993) that is roughly equivalent to Johnson V-band. In this paper we will work mostly in the SDSS g-band, so we make color corrections through
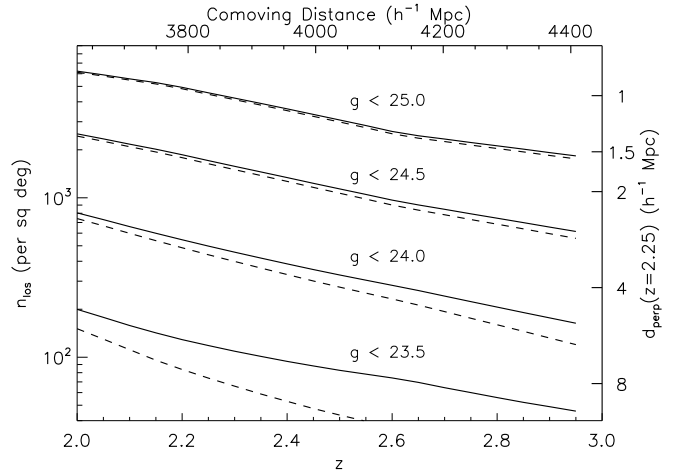

FIG. 2.— The number of background sightlines per sq deg, $n_{\rm los}$, available to probe redshift z at several limiting g-magnitudes. The dashed lines indicate the contribution from galaxies, while the solid line shows the total contribution from both galaxies and quasars. The top axis indicates the comoving distance to each redshift, while the right axis labels the typical sightline separation corresponding to $n_{\rm los}$, evaluated at $z = 2.25$.

the following: we first generate a toy model of the mean LBG spectrum at $z = 2.4$, represented by a power-law $f_\lambda \propto \lambda^\beta$ with Lyα forest attenuation at $\lambda < 1216(1 + z_{\rm bg})$ given by Becker et al. (2013), where $\beta = -1.1$ is the UV continuum slope estimated for $R \approx 24.5$ galaxies at $z = 2.5$ from Bouwens et al. (2009). By comparing the convolution of this model spectrum by the various filter bandpasses and assuming $m_{\mathcal{RG}} = \sqrt{\mathcal{RG}}$, we find $g - m_{\mathcal{RG}} \approx 0.2$. This appears consistent with the LBG color-color diagrams shown in Steidel et al. (2004). The resulting $n_{\rm los}$ distribution is shown as a function of z by the dashed lines in Fig. 2. At $z = 2.45$, we find $n_{\rm los} = [48, 301, 1162, 3243]$ per sq deg at limiting magnitudes of $g_{\rm lim} = [23.5, 24.0, 24.5, 25.0]$, respectively, from galaxies.

For the contribution of quasars to the background sightlines, we use the quasar luminosity function published by Palanque-Delabrouille et al. (2013) in Equation 1. The quasars provide $n_{\rm los} = [41, 54, 71, 91]$ per sq deg at $z = 2.45$ at $g_{\rm lim} = [23.5, 24.0, 24.5, 25.0]$. Quasars clearly provide only a small fraction of background sightlines at $g_{\rm lim} \geq 24$, but they represent the brightest sources at $g \lesssim 23$ and cannot be neglected in an actual survey.

The total $n_{\rm los}$ from galaxies and quasars, at various limiting magnitudes, is shown by the solid curves in Fig 2. This illustrates the exponential increase in $n_{\rm los}$ with limiting magnitude at $g \gtrsim 23$: every magnitude increase in depth yields an order-of-magnitude more sightlines. Since $n_{\rm los}$ quantifies the number of sightlines probing an infinitesimal redshift slice, one must observe a projected source density greater than $n_{\rm los}$ to cover a finite distance along the line-of-sight at the same density, For example, to ensure roughly uniform $n_{\rm los}$ over the range $2.15 \leq z \leq 2.45$, a survey would have to target a projected source density of $\sim 1.8 \, n_{\rm los}(z = 2.45)$, with sources spanning $2.3 \lesssim z_{\rm bg} \lesssim 3.1$. In this paper, we will adopt this scenario as our fiducial survey, i.e. a $\langle z \rangle = 2.25$ survey that assumes a constant differential sightline density of $n_{\rm los}(z = 2.45)$ over the entire

volume, but all other parameters (e.g. angular diameter distance) are evaluated at the lower redshift. Since the comoving distance between $z = 2.15$ and $z = 2.45$ is $\approx 250 \ h^{-1}$ Mpc, such a survey would cover a volume of roughly $1.1 \times 10^6 \ h^{-3}$ Mpc^3 per square degree of sky observed.

For a given n_{los}, we can calculate the typical inter-sightline separation, $\langle d_\perp \rangle \approx \sqrt{1/n_{los}}$. This can be written in units of comoving transverse Mpc:

$$\langle d_\perp \rangle \approx \left[\frac{n_{los}}{4200 \deg^{-2}} \right]^{-1/2} \left[\frac{1+z}{3.25} \right]^{-3/2} h^{-1} \text{Mpc}, \quad (2)$$

where the redshift dependence comes from changes in angular diameter distance. Intuitively, we do not expect to be able to map out scales much smaller than $\langle d_\perp \rangle$, so $\langle d_\perp \rangle$ naïvely sets the spatial resolution ϵ_{3D} of the to-mographic reconstructions feasible from a given set of sightlines. However, as we shall see later, the scale factor relating ϵ_{3D} and $\langle d_\perp \rangle$ is of order unity, but can be varied depending on the desired map properties. The right-axis of Figure 2 labels $\langle d_\perp \rangle$ corresponding to several values of n_{los}, assuming the angular diameter distance at $z = 2.25$. In terms of limiting magnitudes, a survey needs to reach $g_{lim} = [23.5, 24.0, 24.5, 25.0]$ in order to achieve source densities such that $\langle d_\perp \rangle = [7.1, 3.6, 1.9, 1.2] \ h^{-1}$ Mpc.

We find that a simple analytic approximation,

$$\log_{10} n_{los} \approx g_{lim} - 21.5 - 0.75(z - 2.45) \quad (3)$$

fits the source counts to within $\sim 15\%$ over the redshift and limiting magnitudes considered here. This also allows us to write down the median g-magnitude corresponding to n_{los}:

$$\bar{g} \approx \log_{10} n_{los} + 21.0 + 0.75(z - 2.45), \quad (4)$$

by simply setting $n_{los} \to 0.5 \, n_{los}$ in Equation 3.

Alternatively, we insert the approximation of Equation 2 to find

$$\log_{10} \langle d_\perp \rangle \approx -\frac{1}{2}[g_{lim} - 25.1 - 0.75(z - 2.45)]$$
$$\approx -\frac{1}{2}[\bar{g} - 24.6 - 0.75(z - 2.45)]. \quad (5)$$

We will use these approximations in the simple estimates of the following sub-section.

2.2. *Order-of-Magnitude Exposure Time Estimates*

We will now make some simple analytic estimates of the observational requirements required to carry out IGM tomography at various scales, ϵ_{3D}. While we will carry out definitive tests with simulations in the next section, these simple calculations argue that Lyα tomography should already be feasible on the current generation of 8-10m telescopes.

The first parameter we consider is the spectral resolving power, $R \equiv \Delta\lambda/\lambda$, required for IGM tomography. We conjecture that not only is it unnecessary to resolve individual Lyα forest absorbers (requiring echelle spectrographs with $R \gtrsim 10^4$), it is necessary only to resolve the desired 3D reconstruction scale of the map, which we define as ϵ_{3D}. By converting ϵ_{3D} to its equivalent span

in observed wavelength[8], we find

$$R > 1300 \left(\frac{1 \ h^{-1} \text{Mpc}}{\epsilon_{3D}} \right) \left[\frac{(1+z)}{3.25} \right]^{-1/2}. \quad (6)$$

This implies that only moderate-resolution spectra are required for IGM tomography, with up to $R \approx 1300$ required for $\epsilon_{3D} = 1 \ h^{-1}$ Mpc reconstructions. Alternatively, if one desires $\epsilon_{3D} \approx 3 \ h^{-1}$ Mpc maps then $R \approx 400$ is sufficient — this is well within the resolving power of the main moderate-resolution blue grisms on VLT-VIMOS (HR-Blue; Le Fèvre et al. 2003) and Keck-LRIS (B600/4000; Oke et al. 1995).

We now investigate the telescope exposure time, t_{exp}, required to reach a given sightline density and transverse separation. In the background-limited regime, the relationship between t_{exp} and source magnitude, g, is given by

$$t_{exp} = [\text{S/N}]^2 10^{0.8(g - g_0)} \text{ hrs}, \quad (7)$$

where S/N is the spectral signal-to-noise per angstrom, while g_0 is the magnitude at which a telescope and instrument combination gives S/N $= 1$ per angstrom with an integration time of $t_{exp} = 1$ hr. We estimate $g_0 = 24.75$ for the VIMOS spectrograph on the VLT using the ESO exposure time calculator[9]. This assumed $0''.8$ seeing, $1''$ slit size, 1.3 airmass and 3 days from new moon on the $R = 1150$ HR-Blue grating. We also assume a point source, since the typical $z \sim 2 - 3$ LBGs are compact objects with half-light radii of $\sim 0''.3$ (e.g., Shapley 2011); furthermore, Giavalisco et al. (1996) found that for 70% of the LBGs in their sample, the magnitudes within an $0''.7$ aperture differed from the isophotal magnitude by less than 0.1 mag.

Using Equation 5, we can relate the median source magnitude to the expected separation between sightlines, and find

$$t_{exp} = 49 \left(\frac{\text{S/N}}{8 \text{ per Å}} \right)^2 \left(\frac{1 \ h^{-1} \text{Mpc}}{\langle d_\perp \rangle} \right)^{1.6}$$
$$\times \ 10^{0.8(24.75 - g_0)} \text{ hrs.} \quad (8)$$

This indicates that $t_{exp} = 49$ hrs would be required to obtain absorption spectra with S/N $= 8$ per angstrom (roughly the value quoted by Evans et al. 2012, as necessary for IGM tomography) from the $\bar{g} \approx 24.6$ sources that are separated by $\langle d_\perp \rangle \approx 1 \ h^{-1}$ Mpc between lines-of-sight.

If we assume that the spatial resolution of a tomographic map, ϵ_{3D}, is given by the sightline separation $\epsilon_{3D} \approx \langle d_\perp \rangle$, then $\epsilon_{3D} = 1 \ h^{-1}$ Mpc tomography is clearly not feasible with current 8-10m telescopes, and would require the $\sim 15 - 20\times$ greater collecting areas of 30m-class telescope to achieve the necessary depths. However, due to the $t_{exp} \propto \langle d_\perp \rangle^{-1.6}$ scaling, the requisite exposure times drop quickly for maps with coarser resolution: to reach $\langle d_\perp \rangle = 4 \ h^{-1}$ Mpc at fixed signal-to-noise, we find $t_{exp} = 5.3$ hrs. Exposure times of this order are regularly carried out in galaxy redshift surveys on 8m telescopes, e.g. the zCOSMOS-Deep survey

---

[8] For reference, $1 \ h^{-1}$ Mpc comoving distance spans 1.25 Å or $\Delta v = 95$ km s^{-1} along the line-of-sight at $z = 2.25$
[9] http://www.eso.org/observing/etc/

(Lilly et al. 2007). However, the data from these surveys are not useful for Lyα tomography since they are usually low-resolution ($R \sim 200$) spectra that do not have sufficient resolution to resolve structure along the line-of-sight (Equation 6).

However, it is still unclear what spectral signal-to-noise ratio is required for Lyα tomography — in the previous paragraph we have picked a somewhat large value assumed by Evans et al. (2012). This is an important question to address, as the exposure time depends sensitively on the minimum signal-to-noise requirement of the survey. For example, if S/N = 4 per angstrom were the requirement, then only $t_{exp} \approx 12$ hrs would be required to reach a source separation of $\langle d_\perp \rangle = 1\ h^{-1}$ Mpc. Moreover, we have assumed that the spatial resolution of the map is directly set by the typical sightline separation, $\epsilon_{3D} = \langle d_\perp \rangle$. This relationship makes sense at a rough intuitive level since we do not expect to resolve features on scales smaller than the sightline separation, but we need to investigate the exact relationship between ϵ_{3D} and $\langle d_\perp \rangle$, which is possibly also scale-dependent. To address these questions, we will, in the next section, directly carry out tomographic reconstructions on mock Lyα forest spectra derived from numerical simulations.

3. TOMOGRAPHIC RECONSTRUCTION ON SIMULATIONS

In this section, we use mock skewers derived from numerical simulations to explicitly test the signal-to-noise and sightline densities required to map out the IGM at various scales. We first describe the simulations and reconstruction technique, and then discuss the quality of the reconstructions as a function of telescope exposure time and reconstruction scale. Finally, we will derive an analytic expression that allows us to understand some of the trends we see in the simulated reconstructions.

3.1. Simulations and Mock Spectra

We use the dark matter-only numerical simulation used in White et al. (2010). Briefly, this simulation was run using a TreePM code (White 2002) that evolved 2048^3 dark matter particles over a $(250\ h^{-1}$ Mpc$)^3$ periodic cube. The resulting particle masses are $1.4 \times 10^8\ h^{-1} M_\odot$ and have a Plummer equivalent smoothing of $2.5\ h^{-1}$ kpc. The assumed cosmological parameters were $\Omega_m = 0.274$, $\Omega_\Lambda = 0.726$, $h = 0.7$, and $\sigma_8 = 0.8$. Our particular epoch of interest is $z \approx 2.3$, so we used the simulated dark matter distribution output at $z = 2.57$ but subsequently treat it as if it were at $z = 2.25$, so that values for e.g. the mean forest transmission, sightline densities, and angular diameter distance are computed for the latter redshift — for the purposes of this paper the deviations from the true simulation redshift are irrelevant. From the simulated particle distribution, we generate simulated Lyα forest spectra following the procedure described in Rorai et al. (2013); but see also Le Goff et al. (2011); Greig et al. (2011), and Peirani et al. (2014) for alternative methods of generating simulated Lyα forest skewers in the context of large-scale structure studies.

In brief, the dark matter field is first smoothed by an assumed Jeans scale, λ_J, to mimic the Jeans pressure smoothing experienced by baryons, followed by the use of the fluctuating Gunn-Peterson approximation (e.g., Croft et al. 1998; Gnedin & Hui 1998; Weinberg et al.

2003) to link the underlying dark matter overdensity $\Delta_{dm} \equiv \rho_{dm}/\langle \rho_{dm} \rangle$ to the Lyα optical depth, $\tau \propto \Delta_{dm}^{2-0.7(\gamma-1)}$, where γ governs the (assumed) power-law temperature-density relationship of the IGM, $T(\Delta) \propto \Delta^{\gamma-1}$. The Jeans scale is not well-constrained by observations but we choose $\lambda = 100\ h^{-1}$ kpc which conforms to theoretical expectations (Gnedin & Hui 1998). Meanwhile, the value of the temperature-density slope has been the subject of multiple discrepant measurements (Bolton et al. 2008; Viel et al. 2009; Garzilli et al. 2012; Calura et al. 2012; Rudie et al. 2012; Lee et al. 2014) so we pick $\gamma = 1.0$ as a meta-average from these papers. The uncertainties in these (and other) IGM parameters are unimportant for us in this paper since we will be making quantitative comparisons only between the 'true' and tomographically reconstructed absorption fields, but future efforts to directly the infer the dark-matter distribution will have to marginalize over uncertainties in the astrophysics of the IGM, which would otherwise lead to errors in the contrast of the resulting maps.

The optical depths are next convolved with the peculiar velocity field to give the skewers in redshift space. The Lyα forest transmission is then given by $F = \exp(-\tau)$; the full set of skewers is normalized to give a mean transmission, $\langle F \rangle$, consistent with the measurements of Faucher-Giguère et al. (2008). In total, there were 62500 skewers covering the full $250^3\ h^{-3}$ Mpc^3 simulation box at transverse separations of $1\ h^{-1}$ Mpc, with each skewer binned into 2048 line-of-sight pixels with a size of 12 km s^{-1}.

For a given tomographic reconstruction, we randomly select a number of sightlines from our full $1\ h^{-1}$ Mpc grid until we reach the desired area density n_{los}. Note that due to the quantized nature of our sightlines in the transverse plane, we limit ourselves to reconstructions where the inter-sightline separation is $\langle d_\perp \rangle > 1.4\ h^{-1}$ Mpc. We then smooth the spectra to the assumed spectrograph resolution using a 1-dimensional Gaussian kernel. In practice, we find no difference in the reconstruction quality at different resolutions, so long as map resolution, ϵ_{3D}, is resolved by the spectrograph (i.e. Equation 6 is obeyed).

Assuming a complete target selection, the luminosity functions allows us to randomly assign an object magnitude g to each sightline, giving us a distribution of source magnitudes down to the survey limit g_{lim} that corresponds to the chosen sightline density n_{los}. The telescope exposure time, t_{exp}, then relates the assigned magnitude, g, of each source to the signal-to-noise in the Lyα forest via Equation 7. The quantity g_0 in that equation is determined by the specific telescope and spectrograph; we use as a fiducial value $g_0 = 24.75$ corresponding to the HR-Blue mode of the VIMOS spectrograph on the VLT, although it is easy to rescale t_{exp} to a different telescope collecting area and instrument throughput.

For each mock spectrum we then add mock noise to the native simulation pixels by adding a vector of Gaussian deviates with a standard deviation equal to $(\sqrt{95/12})/$S/N, where S/N is determined by t_{exp} and individual source magnitude, and the factor in square-roots are to rescale the noise from the 95 km s^{-1} velocity separation corresponding to 1Å to the 12 km s^{-1} simulation pixels. Figure 3 shows several examples of our
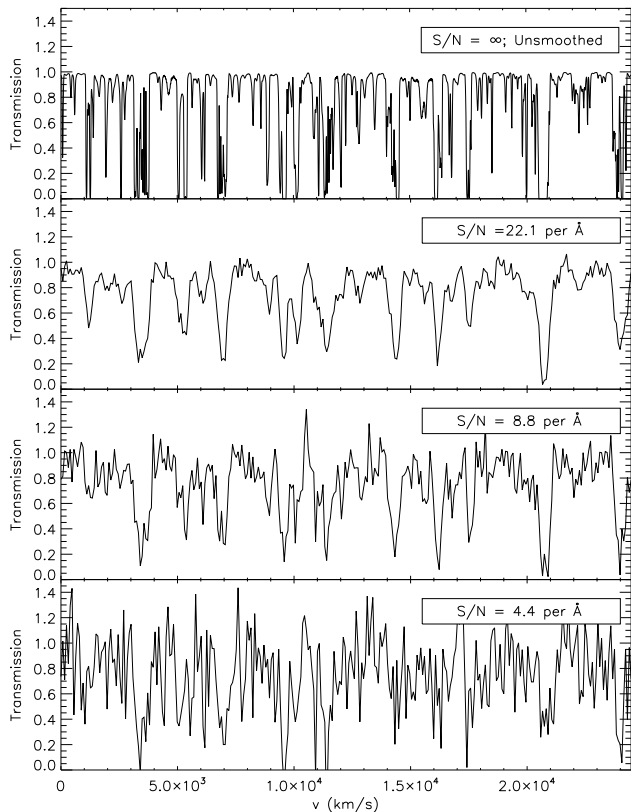
FIG. 3.— Mock spectrum from our simulations, shown at several signal-to-noise ratios. The top panel shows the spectrum without noise or resolution effects, and binned in 11 km s^{-1} pixels. The lower 3 panels show the same spectrum after it has been smoothed to mimic a resolving power of $R = 1150$, rebinned to 1 Å pixels, and had noise added corresponding to the signal-to-noise shown in the labels. The lower panel corresponds roughly to the signal-to-noise limit set by the necessity to make a redshift identification. Assuming $t_{\rm exp} = 5$ hrs on the VIMOS, the noisy spectra correspond to source magnitudes of $g = [22.3, 23.3, 24.0]$ from top to bottom.

mock spectra that have been smoothed to the assumed spectrograph resolution, rebinned to 1 Å pixels, and have had pixel noise added, assuming source magnitudes of $g = [22.3, 23.3, 24.0]$ and an exposure time of $t_{\rm exp} = 5$ hrs on the VIMOS.

3.1.1. *Wiener Reconstruction*

To carry out the tomographic reconstructions, we have written a parallel Fortran90 code that closely follows the Wiener interpolation algorithm described in Caucci et al. (2008). This code, which will be described in more detail in M. Ozbek et al (in prep), was written to create large-scale ($\epsilon_{\rm 3D} \sim 20\ h^{-1}\,{\rm Mpc}$) tomographic maps from the BOSS Lyα forest data. In this approach, the reconstructed map m is estimated from the data d by evaluating

$$m = \mathbf{K} \cdot d$$
$$= \mathbf{C}_{\rm MD} \cdot (\mathbf{C}_{\rm DD} + \mathbf{N})^{-1} \cdot d, \qquad (9)$$

where $\mathbf{K} \equiv \mathbf{C}_{\rm MD} \cdot (\mathbf{C}_{\rm DD} + \mathbf{N})^{-1}$ is a form of Wiener filter (Wiener 1942; Press et al. 1992; Zaroubi et al. 1995) and $\mathbf{N}$ is the noise covariance matrix (assumed to be diagonal), while $\mathbf{C}_{\rm MD}$ and $(\mathbf{C}_{\rm DD} + \mathbf{N})$ describe the map-data and data-data covariances, respectively. In principle, $\mathbf{C}_{\rm MD}$ and $\mathbf{C}_{\rm DD}$ are set by the observed flux power

spectrum or correlation function, but it is well-known that for the purposes of constructing the Wiener filter these only need to be approximately correct (see, e.g, Press et al. 1992). In this paper we therefore adopt the *ad hoc* approach of Caucci et al. (2008), which assumes $\mathbf{C}_{\rm DD} = \mathbf{C}_{\rm MD} = \mathbf{C}(\mathbf{r_1}, \mathbf{r_2})$ and

$$\mathbf{C}(\mathbf{r_1}, \mathbf{r_2}) = \sigma_F^2 \exp\left[-\frac{(\Delta r_\parallel)^2}{L_\parallel^2}\right] \exp\left[-\frac{(\Delta r_\perp)^2}{L_\perp^2}\right], (10)$$

where $\Delta r_\parallel$ and $\Delta r_\perp$ are the distance between $\vec{r_1}$ and $\vec{r_2}$ along, and transverse, to the line-of-sight, respectively. In the case of $\mathbf{C}_{\rm MD}$, $\vec{\Delta r}$ corresponds to the separation between a grid point in the final map grid and a pixel in one of the sightline skewers, while for $\mathbf{C}_{\rm DD}$ the $\vec{\Delta r}$ is between two separate pixels in the absorption skewers, whether in different skewers or along the same skewer. The variance of the flux fluctuations, σ_F^2, is, for the reconstructions, measured directly from the 3D Lyα forest field of the simulation itself. The parameters $L_\parallel$ and $L_\perp$ set the correlation length in the line-of-sight and transverse directions. We set $L_\parallel$ to the line-of-sight distance corresponding to the smoothing FWHM of the assumed instrumental resolution (Equation 6), while $L_\perp = \langle d_\perp \rangle = \sqrt{1/n_{\rm los}}$ as suggested by Caucci et al. (2008).

In practice, we first convert the Lyα flux transmission $F = \exp(-\tau)$ skewers to fluctuations about the mean-flux, $\delta_F = F/\langle F \rangle - 1$. The δ_F from the smoothed and noisy mock spectra are then binned into a 50^3 grid. Since the algorithm is computationally expensive, we carry out reconstructions on subvolumes of $150^3\ h^{-3}\,{\rm Mpc}^3$, $100^3\ h^{-3}\,{\rm Mpc}^3$, or $50^3\ h^{-3}\,{\rm Mpc}^3$ extracted from the full simulation box and binned to 50^3 grid cells. The exact choice of subvolume depends on the tomographic reconstruction scale $\epsilon_{\rm 3D}$, but we ensure that the grid cell sizes are always $\leq \epsilon_{\rm 3D}$. We have checked that the choice of binning does not significantly affect our subsequent results. The binned fluctuations δ_F constitute the 'data' matrix d in Equation 9. Similarly, the noise matrix $\mathbf{N}$ is estimated for each cell by summing in quadrature the inverse of the noise variances, σ_N^2, of the contributing sightlines. This step ensures correct weighting based on the signal-to-noise, and is crucial for a reasonable reconstruction since our sightlines are always dominated by the faint-end of the luminosity function.

To speed up the reconstruction, the matrix inversion in Equation 9 is carried out separately on chunks of 5^3 gridcells. Buffer regions of $\sim 2\epsilon_{\rm 3D}$ are added to each face of the chunks to mitigate edge artifacts; we discard these buffer zones from the final reconstructed volume. As a final step, the reconstructed volume is smoothed by a 3D Gaussian kernel with standard deviation $\epsilon_{\rm 3D}$. Caucci et al. (2008) set $\epsilon_{\rm 3D} = 1.4\langle d_\perp \rangle$, but we shall treat it as a free parameter that allows us to vary the mapping scale and reconstruction fidelity of the map, as we shall see in the next section.

3.2. *Results*

In this section, we carry out tomographic reconstructions on the simulated Lyα forest absorption skewers described in § 3.1, in order to study the quality of the resulting maps, at various spatial resolutions $\epsilon_{\rm 3D}$, as a function of observational parameters.
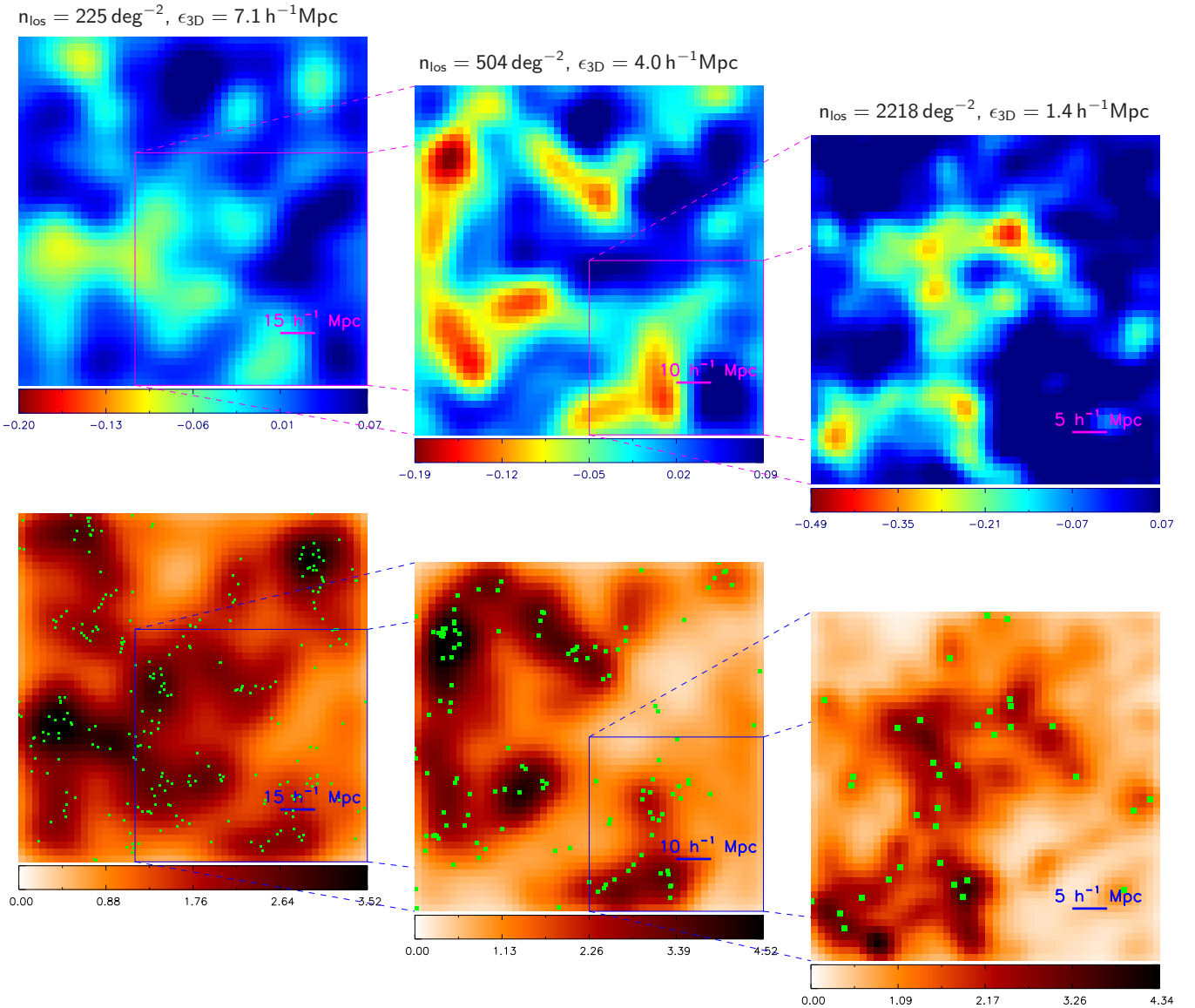
FIG. 4.— Simulated Lyα forest tomographic reconstructions at various scales ϵ_{3D}, shown in the plane of the sky. The top-row shows the Lyα absorption field reconstructed from randomly-selected sets of mock spectra with various areal densities n_{los} and smoothed to ϵ_{3D} (both labeled on top), while the bottom row shows the corresponding dark matter overdensity smoothed to the same ϵ_{3D}. From left to right, the dimensions of the simulation slices are $(150 \ h^{-1} \mathrm{Mpc})^2$, $(100 \ h^{-1} \mathrm{Mpc})^2$, and $(50 \ h^{-1} \mathrm{Mpc})^2$ respectively with thickness of $3 \ h^{-1}$ Mpc along the line-of-sight; the smaller slices are subvolumes of the larger slices. These maps correspond to transverse areas of $[5.09 \deg^{-2}, 2.26 \deg^{-2}, 0.56 \deg^{-2}]$ on the sky (from left to right), and we have assumed telescope exposure times of $t_{exp} = [3, 5, 16]$ hrs to generate the noisy mock spectra. The green dots overlaid on the DM maps show a complete sample of $\mathcal{R} \leq 25.5$ galaxies co-eval with the Lyα forest field, obtained from abundance matching to dark-matter halos in the simulation volume.

The primary observational parameters we can vary are the differential sightline density, $n_{los}(z)$, that is tied to the limiting magnitude[10], g_{lim}, and ultimately the telescope exposure time, t_{exp}. In principle, n_{los} and t_{exp} can be varied independently but we begin by assuming that t_{exp} is tied to g_{lim} and n_{los}. In other words, longer exposure times yield more sightlines because more sources become available at fainter magnitudes. While it is in principle possible to target increasingly noisy faint sources below a given magnitude limit, there is, in practice, a minimum spectral signal-to-noise threshold below which a redshift identification becomes difficult. For LBGs with

no intrinsic Lyα emission line, S/N = 4 per angstrom seems to be a conservative threshold for measuring the redshift of a non-emission line LBG from its intrinsic absorption lines (Steidel et al. 2010). Combining Equations 3 and 7, we therefore have

$$t_{exp} \approx 9 \left(\frac{n_{los}}{1000 \deg^{-2}} \right)^{0.8} \mathrm{hrs}$$

$$\approx 9 \left(\frac{2.1 \ h^{-1} \mathrm{Mpc}}{\langle d_\perp \rangle} \right)^{1.6} \mathrm{hrs} \qquad (11)$$

where we have assumed the performance of the HR-Blue mode of the VIMOS spectrograph on the VLT. This equation is only an approximation for illustrative purposes, because for the simulated spectra we directly use

---

[10] We ignore, for now, inefficiencies in target selection and assume that we can observe all background sources down to the magnitude limit.

the source luminosity functions (§ 2.1) to determine the relationship between $g_{\rm lim}$ and $n_{\rm los}$.

For a given $n_{\rm los}$, we randomly select simulated sightlines and assign to each sightline a source magnitude (assuming the luminosity functions described in § 2.1) and seed the pixels with Gaussian random noise assuming a signal-to-noise given by $t_{\rm exp}$ (Equation 7). We then carry out the Wiener filtering procedure described in the previous section. The final step involves smoothing the reconstructed volume with a 3D Gaussian with a standard deviation $\epsilon_{\rm 3D}$ — we treat this as a free parameter that allows us to choose the final map resolution.

In Figure 4 we present transverse-plane slices of simulated Lyα forest tomographic reconstructions with smoothing scales of $\epsilon_{\rm 3D} = [7.1, 4.0, 1.4]\ h^{-1}\,{\rm Mpc}$, generated from mock surveys with $n_{\rm los} = [225, 504, 2218]\,{\rm deg}^{-2}$ and $t_{\rm exp} = [3, 5, 16]\,{\rm hrs}$ (we will explain the choices of $\epsilon_{\rm 3D}$ later). The reconstructions were carried out on sub-volumes of the overall simulation box, with dimensions $(150\ h^{-1}\,{\rm Mpc})^3$, $(100\ h^{-1}\,{\rm Mpc})^3$, and $(50\ h^{-1}\,{\rm Mpc})^3$ for the $\epsilon_{\rm 3D} = [7.1, 4.0, 1.4]\ h^{-1}\,{\rm Mpc}$ reconstructions, respectively; the smaller boxes are sub-volumes of the larger boxes. This is reflected in the slices shown in Figure 4; for example, one can see how the overall anvil-like structure in the $\epsilon_{\rm 3D} = 1.4\ h^{-1}\,{\rm Mpc}$ map (at right) varies as we move to coarser map resolutions.

For a qualitative comparison, we juxtapose the underlying dark matter overdensity field $\Delta_{\rm dm} \equiv \rho_{\rm dm}/\langle\rho_{\rm dm}\rangle$, clipped to $\Delta_{\rm dm} \leq 20$ and smoothed to the same $\epsilon_{\rm 3D}$ as the corresponding IGM maps. The tomographic reconstructions agree very well with the underlying DM overdensity: features on scales larger than several $\epsilon_{\rm 3D}$ are well recovered, although the contrast is somewhat different due to the different dynamic range of the exponential $F \equiv \exp(-\tau_{\rm Ly\alpha})$ transformation in the Lyα forest. It is also clear that, even at the coarsest reconstruction scales, IGM tomography is exceptionally good at detecting voids in large-scale structure, a non-trivial endeavour with galaxy redshift surveys even at low-redshifts (e.g., Tinker et al. 2008; Kreckel et al. 2011; Sutter et al. 2012). The efficient detection of voids will facilitate the study of the topology of large-scale structure out to high-redshifts. However, in this paper we do not move beyond this qualitative comparison between the tomographic absorption maps and the dark-matter field; we will address the problem of inverting the dark-matter field to a future paper.

It is important to emphasize that these maps were reconstructed from mock spectra that *include realistic pixel noise assuming feasible exposure times on existing 8-10m telescopes.* The 5-hour exposure times, which enables $\epsilon_{\rm 3D} = 4\ h^{-1}\,{\rm Mpc}$ maps (middle column of Figure 4), is already regularly carried out for galaxy redshift surveys on 8m telescopes — indeed, we have set $t_{\rm exp}$ by the necessity of measuring the source redshifts. The longer $t_{\rm exp} \gtrsim 10\,{\rm hrs}$ exposure times required for $\epsilon_{\rm 3D} \lesssim 2\ h^{-1}\,{\rm Mpc}$ tomography is at the margins of integration times that have been attempted for individual pointings in a galaxy redshift survey. However, the stunning detail revealed by these maps is arguably unparalleled by any other cosmographical technique outside of the $z \lesssim 0.1$ Local Universe, and should motivate at-
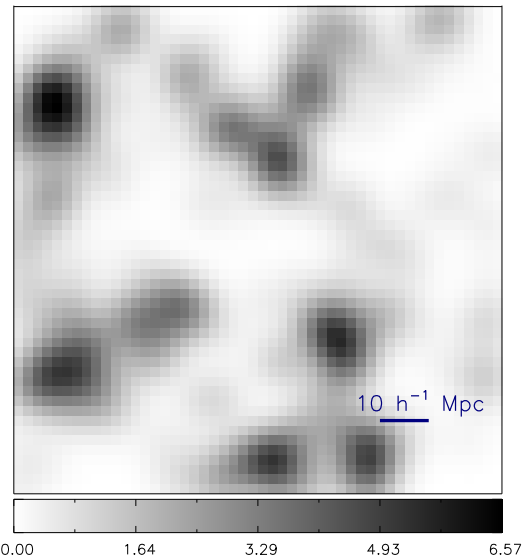

FIG. 5.— Galaxy overdensity field in the same simulated slice shown in the middle panels of Fig. 4. This used the distribution of $\mathcal{R} \leq 25.5$ galaxies obtained through halo abundance matching with the dark-matter halos in the simulation, which was then smoothed with a Gaussian filter with $4\ h^{-1}\,{\rm Mpc}$ standard deviation to match the $\epsilon_{\rm 3D}$ of the tomographic reconstructions in Fig. 4. The dimensions of this map is $100\ h^{-1}\,{\rm Mpc} \times 100\ h^{-1}\,{\rm Mpc}$ in the plane of the page, and $2\ h^{-1}\,{\rm Mpc}$ into the plane of the page.

tempts to carry this out on existing instrumentation.

To compare with the tomographic maps, we have also generated a mock galaxy catalog through halo abundance-matching (Kravtsov et al. 2004; Tasitsiomi et al. 2004; Vale & Ostriker 2004) between the DM halo catalog from the simulation and the Reddy et al. (2008) LBG luminosity function. On the DM overdensity maps we have over-plotted the positions of $\mathcal{R} \leq 25.5$ galaxies that fall within the same simulated volume. These are foreground galaxies that would not be targeted as background sources within a tomographic survey, but sub-samples of such galaxies could be obtained through other surveys for comparison with the tomographic maps (or vice-versa: the tomographic mapping survey could target well-studied galaxy fields such as CANDELS, Grogin et al. 2011; Koekemoer et al. 2011, or 3D-HST, Brammer et al. 2012). It is clear that the LBGs are clustered around the same features traced out by the dark matter distribution and the Lyα forest absorption, suggesting that Lyα tomography will be a powerful tool in the study of galaxy environments at $z \sim 2$.

In Figure 5, we plot the overdensity of $\mathcal{R} \leq 25.5$ galaxies, $\Delta_{\rm galaxy} = \rho_{\rm galaxy}/\langle\rho_{\rm galaxy}\rangle$ occupying the same simulation slice as the $\epsilon_{\rm 3D} = 4\ h^{-1}\,{\rm Mpc}$ tomographic reconstruction shown in the middle panel of Figure 4. This has been smoothed with a $\sigma = 4\ h^{-1}\,{\rm Mpc}$ Gaussian kernel to match the Lyα forest reconstructions. The strongest galaxy overdensities form quasi-spherical groups that trace out highly overdense regions of the simulation volume, but do not trace well the filamentary structures closer to mean-density. This is further illustrated by Figure 6, which show scatter plots relating both tomographic reconstruction fluxes and galaxy overdensity to the underlying DM overdensity. At regions of under- and average-density ($\Delta_{\rm dm} \lesssim 1$), the reconstructed Ly$\alpha$
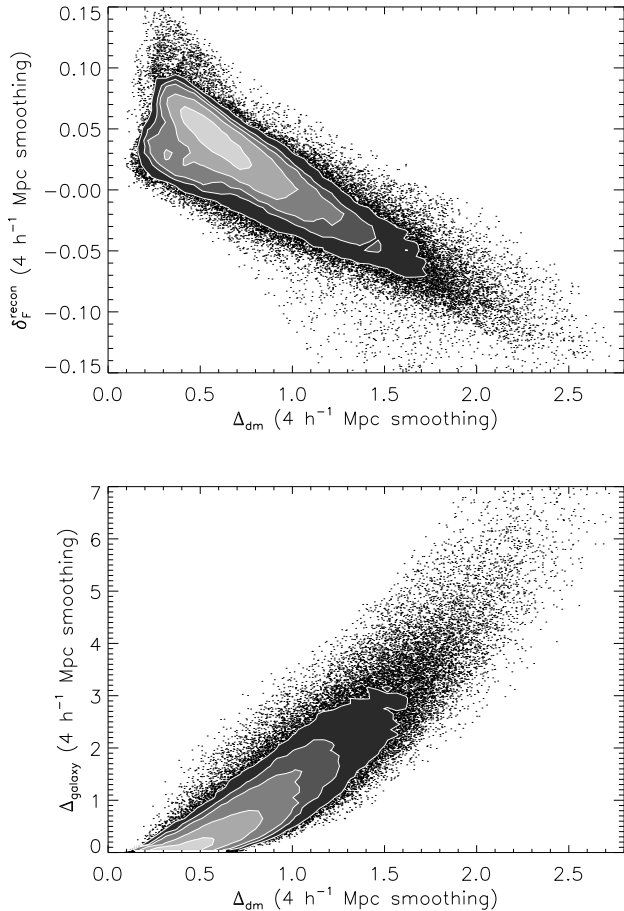
FIG. 6.— Reconstructed Lyα forest flux (top) and galaxy over-density (bottom) from the simulation volume plotted as a function of the corresponding dark matter overdensity, within $(2\ h^{-1}\ \mathrm{Mpc})^3$ voxels. The corresponding visualizations are in the middle panels of Fig. 4 and Fig. 5, respectively. The contours denote the 10th, 20th, 30th, 50th and 80th percentiles of the distributions. The Lyα forest tomographic reconstruction clearly gives a better mapping of the smoothed DM distribution at $\Delta_{\mathrm{dm}} \lesssim 1$, while the relationship between galaxy overdensity and DM overdensity becomes quasi-linear at $\Delta_{\mathrm{dm}} \gtrsim 1$.

fluxes are roughly linear with respect to Δ_{dm}, but the Δ_{galaxy} distribution is nearly flat with respect to Δ_{dm} in this regime. The smoothed galaxy distribution only becomes linear with respect to Δ_{dm} at denser regimes, and at $\Delta_{\mathrm{dm}} \gtrsim 2$ it exhibits less scatter than the Lyα forest reconstruction. We will pursue these comparisons in more detail in a separate paper (although see, e.g., Kitaura et al. 2012).

Apart from visual inspection, we quantify the fidelity of the tomographic maps by plotting the scatter of the reconstructed field δ_F^{rec} against the full 3D Lyα forest field δ_F^{orig}, in which both fields have been binned and smoothed as described in § 3.1.1. This is shown in Figure 7: the distribution of points in the scatter plot appears to be quite linear. However, we see a bias in the sense that the best-fit linear regression differs from the $\delta_F^{\mathrm{rec}} = \delta_F^{\mathrm{orig}}$ relation that would be obeyed in the case of a perfect reconstruction. This is likely due to the *ad hoc* nature of our Wiener reconstruction algorithm: for example, in the final step we smooth the full field with an
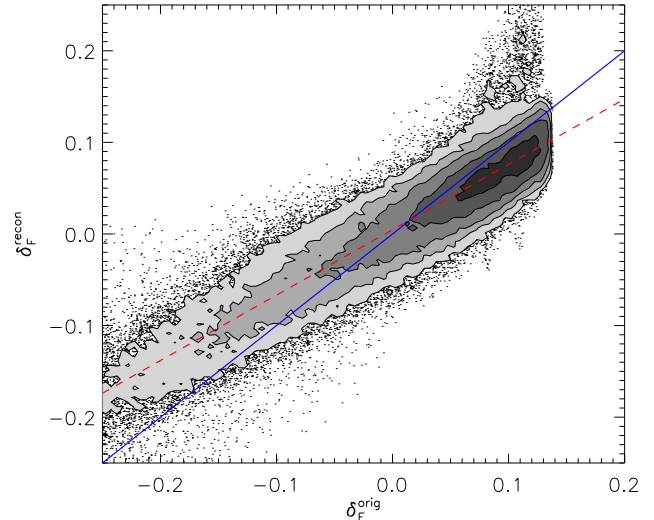

FIG. 7.— Scatter plot of the reconstructed absorption field fluctuations, δ_F^{rec}, against the true forest fluctuations, δ_F^{orig}, in the $n_{\mathrm{los}} = 2218\ \mathrm{deg}^{-2}$ and $\epsilon_{\mathrm{3D}} = 1.4\ h^{-1}\ \mathrm{Mpc}$ reconstruction as shown in the top-right of Figure 4. Both fields have been smoothed to $1.4\ h^{-1}\ \mathrm{Mpc}$ with a spherical Gaussian filter. The contours denote the 5th, 15th, 30th, 50th, and 80th percentiles of the distribution. The solid blue line is the $\delta_F^{\mathrm{orig}} = \delta_F^{\mathrm{rec}}$ relation, while the red dashed line shows the best-fit regression linear fit for the points. The relationship between δ_F^{rec} and δ_F^{orig} is quite linear with a cross-correlation coefficient of $r = 0.93$, allowing a straightforward correction for the overall bias. The distribution of δ_F^{rec} shown here corresponds to $\mathrm{SNR}_{\epsilon} = 2.49$.

isotropic 3D Gaussian kernel, even though the geometry of the problem is intrinsically anisotropic between the line-of-sight and transverse directions (c.f. Equation 10). In this paper we are more concerned with the observational requirements of IGM tomographic mapping, so we defer improvements on the algorithm to future work. For now, we simply correct δ_F^{rec} by the best-fit regression linear function to remove the bias; the reconstructed maps shown in Figure 4 have already been corrected for this bias.

The δ_F^{rec}-δ_F^{orig} comparison also allows us to define a reconstruction signal-to-noise ratio, SNR_{ϵ}, through the signal variance and residual variance:

$$\mathrm{SNR}_{\epsilon}^2 = \frac{\mathrm{Var}(\delta_F^{\mathrm{orig}})}{\mathrm{Var}(\delta_F^{\mathrm{rec}} - \delta_F^{\mathrm{orig}})}, \qquad (12)$$

where δ_F^{rec} has been corrected for the bias as described in the previous paragraph. For a fixed set of sightlines, SNR_{ϵ} can be varied by changing the smoothing scale ϵ_{3D}: by smoothing to larger scales on both the reconstruction and true map, the agreement between the two is improved and vice-versa.

For a qualitative comparison of how reconstructions at various SNR_{ϵ} appear visually, in Figure 8 we show several tomographic maps with different n_{los}(and t_{exp}, related through Equation 11), but smoothed to the same ϵ_{3D}. Compared to the true underlying field, the mock data set with $[t_{\mathrm{exp}}, n_{\mathrm{los}}] = [8\,\mathrm{hrs}, 971\,\mathrm{deg}^{-2}]$ gives a very good reconstruction of the main structures with $\mathrm{SNR}_{\epsilon} = 3.0$, although there are some discrepancies at flux levels close to the $\delta_F = -0.05$, which approximately traces the mean-density of the dark-matter distribution.
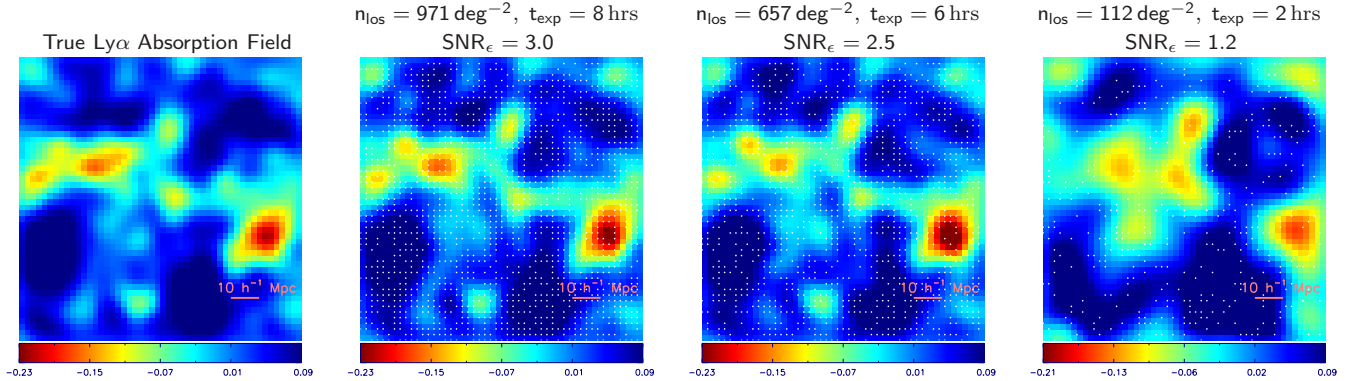
FIG. 8.— Simulated tomographic reconstructions of the same field generated from different survey parameters $[t_{\rm exp}, n_{\rm los}]$ and compared with the true underlying 3D Lyα forest absorption field (at left). All the maps have been smoothed with a Gaussian kernel of $\epsilon_{\rm 3D} = 3.5 \ h^{-1}$ Mpc to facilitate a direct comparison. The simulation slices have dimensions $(100 \ h^{-1} {\rm Mpc})^2 \times 2 \ h^{-1}$ Mpc, the line-of-sight direction is into the plane of the page. The positions of the Lyα forest mock skewers that contributed to each reconstruction are indicated by the white dots in each map. The reconstruction signal-to-noise, SNR_ϵ, is also labeled and the deterioration of reconstruction fidelity with lower $n_{\rm los}$ and $t_{\rm exp}$ is apparent.

The quality of this map is unsurprising, as the sight-line density corresponds to an average sightline separation of $\langle d_\perp \rangle = 2.1 \ h^{-1}$ Mpc (c.f. Equation 2), which oversamples the field at this $\epsilon_{\rm 3D} = 3.5 \ h^{-1}$ Mpc mapping resolution, i.e. within a $\epsilon_{\rm 3D}^2$ patch on the sky, there are ~ 2.7 sightlines that sample the area. Moving down to the survey parameters $[t_{\rm exp}, n_{\rm los}] = [6\,{\rm hrs}, 657\,{\rm deg}^{-2}]$, the resulting value of SNR_ϵ is somewhat decreased to 2.5, but visually the true absorption is still well reconstructed apart from some deterioration to delicate features, again at $\delta_F \sim -0.05$. This survey's sightline density of $n_{\rm los} = 657 \, {\rm deg}^{-2}$ corresponds to a sightline separation of $\langle d_\perp \rangle = 2.6 \ h^{-1}$ Mpc, which is again well-sampled with respect to $\epsilon_{\rm 3D} = 3.5 \ h^{-1}$ Mpc. We move on to the right-most panel in Figure 8 and see that the $[t_{\rm exp}, n_{\rm los}] = [2\,{\rm hrs}, 112\,{\rm deg}^{-2}]$ now gives an inferior tomographic reconstruction, with large distortions and none of the delicate filamentary structures reproduced. This is because the sightline separation, $\langle d_\perp \rangle = 6.3 \ h^{-1}$ Mpc, corresponding to this survey depth is too coarse compared to $\epsilon_{\rm 3D} = 3.5 \ h^{-1}$ Mpc, although the agreement can be improved by smoothing to larger $\epsilon_{\rm 3D}$. Qualitatively, $\mathrm{SNR}_\epsilon \approx 2.0 - 2.5$ seems to be a reasonable threshold for a 'good' reconstruction, but more quantitative criteria will have to depend on the science goals of the survey — these will be explored in future papers where we investigate the utility of IGM tomography for galaxy environment studies and hunting for the progenitors of galaxy clusters.

Intuitively, one expects a reasonable reconstruction when $\epsilon_{\rm 3D} \gtrsim \langle d_\perp \rangle$, but there is some leeway in selecting the smoothing scale $\epsilon_{\rm 3D}$ depending on the desired reconstruction fidelity SNR_ϵ. We can relate the survey parameters $[t_{\rm exp}, n_{\rm los}]$ as a function of mapping scale $\epsilon_{\rm 3D}$ at fixed SNR_ϵ, by simply varying $\epsilon_{\rm 3D}$ at fixed $[t_{\rm exp}, n_{\rm los}]$ until the desired SNR_ϵ is achieved. This is shown by the symbols in Figure 9, and also tabulated in Table 1. At fixed reconstruction fidelity, the $t_{\rm exp}$ vs $\epsilon_{\rm 3D}$ curve remains relatively flat at $\epsilon_{\rm 3D} \gtrsim 4 \ h^{-1}$ Mpc, i.e. small increases in survey parameters enable large improvements in the mapping resolution. For example, going from $t_{\rm exp} = 3\,{\rm hrs}$ to $t_{\rm exp} = 5\,{\rm hrs}$ leads to nearly a dou-

bling of the map resolution from $\epsilon_{\rm 3D} = 7.2 \ h^{-1}$ Mpc to $\epsilon_{\rm 3D} = 3.8 \ h^{-1}$ Mpc, at fixed $\mathrm{SNR}_\epsilon \approx 2.5$. With the current generation of 8-10m telescopes, it therefore makes sense for the first generation of IGM tomographic surveys to target the regime where the curves in Figure 9 begin to rise steeply, i.e. $t_{\rm exp} \approx 5 - 6\,{\rm hrs}$ and $n_{\rm los} \approx 500 - 600 \, {\rm deg}^{-2}$, that would enable maps with $\epsilon_{\rm 3D} = 3 - 4 \ h^{-1}$ Mpc at $\mathrm{SNR}_\epsilon \sim 2.5$. To achieve map resolutions of $\epsilon_{\rm 3D} \lesssim 2 \ h^{-1}$ Mpc, the survey requirements become increasingly stringent, with $t_{\rm exp} \gtrsim 10\,{\rm hrs}$ and rising rapidly. Tomographic maps with $\epsilon_{\rm 3D} \approx 1 \ h^{-1}$ Mpc resolution would require $t_{\rm exp} \gtrsim 20\,{\rm hrs}$ to obtain sufficient data from the $g_{\rm lim} \approx 25$ background sources, rendering it very challenging for the current generation of telescopes. However, with future 30m-class telescopes it would require only of order 1-2 hours of integration to obtain the spectra. Note that this is considerably less than the $\sim 10\,{\rm hrs}$ integrations previously assumed to be necessary for $\epsilon_{\rm 3D} \sim 1 \ h^{-1}$ Mpc IGM tomography on 30m telescopes. (Steidel et al. 2009; Evans et al. 2012).

3.3. Analytic Estimates of Reconstruction Noise

In this section, we attempt to calculate analytically the relationship between the survey parameters ($t_{\rm exp}$, $n_{\rm los}$, $\langle d_\perp \rangle$) and the characteristics of the resulting tomographic map ($\epsilon_{\rm 3D}$, SNR_ϵ). The purpose of this calculation is two-fold: (a) to develop intuition for how the various survey parameters affect the final map, and (b) to provide a 'quick-and-dirty' way to determine the effect of survey complications such as target-selection inefficiencies, uneven exposure times etc. As such, we do not attempt to make a rigorous derivation, and will defer such work to a future paper where we work on improvements to the Wiener reconstruction algorithm.

We begin by defining[11] the data as $d = s + n$, i.e. the sum of the true underlying signal s and a noise contribution, n. The signal and noise covariances are denoted as $\mathbf{S} \equiv \langle s \cdot s^T \rangle$ and $\mathbf{N} \equiv \langle n \cdot n^T \rangle$ respectively, giving the

---

[11] For clarity, in this section we have changed the notation of some variables. The signal s and map m here correspond, respectively, to $\delta_F^{\rm orig}$ and $\delta_F^{\rm rec}$ in the previous section.

$t_{\exp}$[a] (hrs)	$g_{\lim}$[b]	$n_{\los}$[b] (deg^{-2})	$\langle d_\perp \rangle$[d] (h^{-1} Mpc)	ϵ_{3D}[d] (h^{-1} Mpc)		
				$\mathrm{SNR}_\epsilon = 2.0$	$\mathrm{SNR}_\epsilon = 2.5$	$\mathrm{SNR}_\epsilon = 3.0$
2	23.6	110	6.3	8.1	13.9	20.5
3	23.8	230	4.4	4.6	7.2	10.5
4	24.0	360	3.5	3.2	4.8	6.9
5	24.1	500	3.0	2.6	3.8	5.3
6	24.2	660	2.6	2.3	3.2	4.5
8	24.4	970	2.1	1.9	2.7	3.7
10	24.5	1300	1.9	1.6	2.3	3.2
12	24.6	1600	1.7	1.4	2.0	2.7
16	24.8	2200	1.4	1.0	1.3	1.7

[a] Telescope exposure time, assuming the HR-Blue mode on VLT-VIMOS.
[b] Limiting g-magnitude of survey, defined by requirement to achieve S/N ≥ 4 per angstrom at survey limit.
[c] Areal density of sightlines used in reconstruction. Related to $t_{\exp}$ by Equation 11.
[d] Typical comoving separation between sightlines in the transverse direction.
[e] Map resolution that would yield the specified reconstruction fidelity SNR_ϵ for the choice of $t_{\exp}$ and $n_{\los}$.
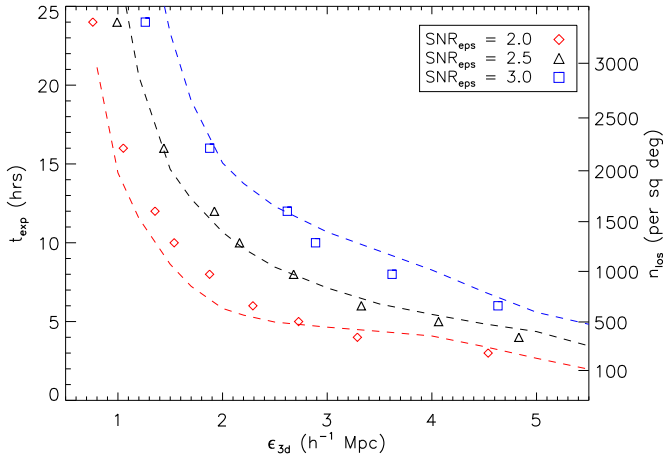

FIG. 9.— The exposure time, on the VLT-VIMOS spectrograph, required to carry out tomographic reconstructions at a given spatial resolution, ϵ_{3D} at various reconstruction signal-to-noise, SNR_ϵ (denoted by different colors). The symbols are derived from our simulated reconstructions, while the dashed-lines show the same quantity estimated with the analytic estimates of § 3.3. The right-hand axis labels the background source densities achievable given the exposure time, assuming that the spectra must have at least S/N ≥ 4 per angstrom.

corresponding data covariance

$$\mathbf{D} \equiv \langle d \cdot d^T \rangle = \mathbf{S} + \mathbf{N} \qquad (13)$$

if we assume that the noise and signal are uncorrelated, $\langle s \cdot n^T \rangle = 0$. Our reconstructed map is the data convolved by a filter, $\hat{m} = \mathbf{K} \cdot d$, with the covariance

$$\hat{\mathbf{M}} \equiv \langle \hat{m} \cdot \hat{m}^T \rangle = \mathbf{K} \langle d \cdot d^T \rangle \mathbf{K}^T = \mathbf{K} \cdot \mathbf{D} \cdot \mathbf{K}^T, \qquad (14)$$

where the hat indicates that the reconstructed map is defined, in principle, on a different set of coordinates from the data. In the case of Wiener filtering,

$$\mathbf{K} = \hat{\mathbf{S}} \cdot (\mathbf{S} + \mathbf{N})^{-1}. \qquad (15)$$

In our tomographic reconstructions, both the reconstructed map and true map have been binned on to the same grid, so we drop the hat in the notation, e.g. $\hat{\mathbf{S}} = \mathbf{S}$.

We can thus write down the covariance of the residual between the reconstructed map and the true underlying field as:

$$\mathbf{R} \equiv \langle (m - s)(m - s)^T \rangle$$
$$= (\mathbf{K} - \mathbf{I}) \cdot \mathbf{S} \cdot (\mathbf{K} - \mathbf{I})^T + \mathbf{K} \cdot \mathbf{N} \cdot \mathbf{K}^T, \qquad (16)$$

where $\mathbf{I}$ is the identity matrix and we have used the fact that

$$m - s = \mathbf{K} \cdot (s + n) - s$$
$$= (\mathbf{K} - \mathbf{I}) \cdot s + \mathbf{K} \cdot n, \qquad (17)$$

and again assumed that $\langle s \cdot n^T \rangle = 0$.

We are now in the position to estimate analytically SNR_ϵ at some map resolution ϵ_{3D} and telescope exposure time $t_{\exp}$, which sets the noise of the map. In the notation of this section, this is

$$\mathrm{SNR}_\epsilon^2(\epsilon_{3D}, t_{\exp}) = \frac{\sigma_S^2(\epsilon_{3D})}{\sigma_R^2(\epsilon_{3D}, t_{\exp})}, \qquad (18)$$

i.e. the ratio of the signal and residual variances from the reconstructions. In the present case, the numerator, σ_S^2, is the variance of the true Lyα forest absorption field[12] smoothed over standard deviation ϵ_{3D}. Meanwhile, σ_R^2 is the residual variance from the reconstruction smoothed over the same scale. The latter has a dependence on $t_{\exp}$ because this determines the pixel noise in the absorption spectra that go into the reconstructions.

In our case, $\sigma_S^2(\epsilon_{3D})$ is the signal s smoothed over a Gaussian window with standard deviation ϵ_{3D}, which is easily evaluated by carrying out a volume integral over the flux power-spectrum:

$$\sigma_S^2(\epsilon_{3D}) = \frac{1}{2\pi} \int_0^\infty k^2 dk \int_{-1}^1 d\mu$$
$$\times P_F(k, \mu) W^2(k\epsilon_{3D}), \qquad (19)$$

where $P_F(k, \mu)$ is the anisotropic 3D Lyα forest flux power spectrum, $\mu = \cos\theta$ is the ratio between the parallel component of the wave-vector, $k_\parallel$, and its modulus k, and $W(kR)$ is the Fourier space Gaussian filter:

$$W(kR) = \exp\left(-\frac{1}{2}k^2 R^2\right). \qquad (20)$$

[12] This quantity is more usually denoted σ_F^2 in the literature

For the Lyα forest power spectrum $P_F(k,\mu)$, we use the analytic model (e.g., McDonald 2003; McQuinn et al. 2011; McQuinn & White 2011)

$$P_F(k,\mu) = b^2(1+\beta\mu^2)^2 P_L(k)\exp(-k_\parallel^2/k_D^2), \quad (21)$$

where b and β are the bias and anisotropy parameters respectively (Kaiser 1987), $P_L(k)$ is the linear-theory dark-matter power spectrum, and k_D parametrizes the small-scale cut-off in the line-of-sight power spectrum from Jeans and thermal smoothing. We set $k_D = 0.08\,\mathrm{s\,km^{-1}}$ as in McQuinn & White (2011), but its exact value is unimportant to us as the cutoff is at much smaller scales than we are concerned with.

By analogy with Equation 19, we can evaluate σ_R^2 as the integral over the residual power spectrum multiplied by a Gaussian filter. If we assume translational invariance then all of the matrices in the expression for the residual covariance (Equation 16) are diagonal in k-space, and we can just replace the matrices with power spectra to obtain an expression for the residual power spectrum, $P_R(k,\mu)$. In addition, the linear algebra becomes regular algebra. We then get

$$\begin{aligned} P_R &= (K-1)P_F(K-1) + K(\bar{n}_{\mathrm{eff}}^{-1}P_{\mathrm{los}})K \\ &= \frac{P_F(\bar{n}_{\mathrm{eff}}^{-1}P_{\mathrm{los}})^2 + P_F^2\,\bar{n}_{\mathrm{eff}}^{-1}P_{\mathrm{los}}}{(P_F + \bar{n}_{\mathrm{eff}}^{-1}P_{\mathrm{los}})^2}, \end{aligned} \quad (22)$$

where we have substituted in the Fourier-space Wiener filter $K = P_F/(P_F + \bar{n}_{\mathrm{eff}}^{-1}P_{\mathrm{los}})$ in the second line, and do not explicitly write out the dependencies on k and μ. The term $\bar{n}_{\mathrm{eff}}^{-1}P_{\mathrm{los}}$ is an approximation for the Lyα forest noise power from Equation 12 in McQuinn & White (2011), where P_{los} is the line-of-sight Lyα forest power,

$$\bar{n}_{\mathrm{eff}} \equiv \frac{1}{A}\sum_n \nu_n, \qquad \nu_n = \frac{P_{\mathrm{los}}}{P_{\mathrm{los}} + P_{N,n}}, \quad (23)$$

A is the transverse area of the survey, and

$$P_{N,n} = 0.8\langle F\rangle^{-2}[(\mathrm{S/N})_n]^{-2}\left(\frac{1+z}{4}\right)^{-3/2} \quad (24)$$

is the line-of-sight power arising from pixel noise in an individual spectrum n that has signal-to-noise $(\mathrm{S/N})_n$ (per angstrom), with a redshift dependence that arises from the conversion between observed wavelength and line-of-sight comoving distance. For each spectrum n in a mock survey, we can evaluate ν_n as a function of its spectral signal-to-noise, which is in turn set by the source magnitude and exposure time. The sum of all the ν_n divided by the survey area gives $\bar{n}_{\mathrm{eff}}$, which sets the noise in the reconstructions. McQuinn & White (2011) assumed a fixed value of P_{los} since the line-of-sight power spectrum is approximately white at the scales they were concerned with, but we carry out the full evaluation of P_{los} as a function of $k_\parallel = k\,\mu$.

With the residual power spectrum in hand (Equation 22), we then compute the residual variance within a smoothing window ϵ_{3D} in an analogous integral to Equation 19. This now allows us to compute, through Equation 18, the SNR_ϵ expected for a given combination of ϵ_{3D} and t_{exp}. The analytic curves for t_{exp} against ϵ_{3D}, at fixed SNR_ϵ, is plotted as the dashed-lines in Figure 9.
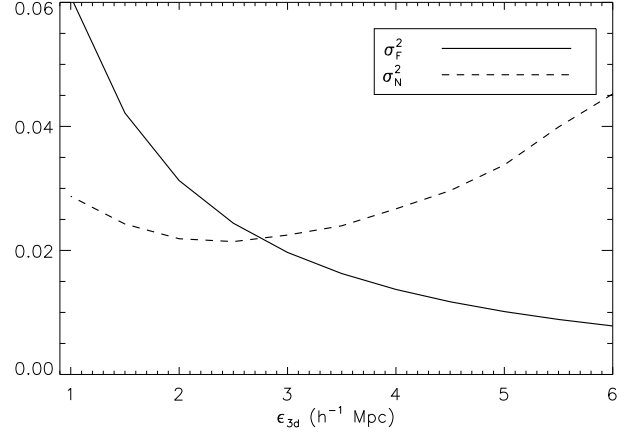
FIG. 10.— The Lyα forest flux variance σ_{F}^2 and shot-noise variance σ_{N}^2, as a function of reconstruction scale ϵ_{3D}. In this particular case we have assumed that $\epsilon_{\mathrm{3D}} = \langle d_\perp\rangle$, and set $\bar{n}_{\mathrm{eff}} = n_{\mathrm{los}}$ to remove the contribution from pixel noise in the spectra and show only the shot-noise from the random sampling of sightlines.

The only free parameters we have in this model is b and β in the 3D Lyα forest power spectrum (Equation 21) — the curves in Figure 9 assume $b = 0.21$ and $\beta = 0.5$, which are both consistent with the fits of Slosar et al. (2011). The analytic curves have the same qualitative behavior as the points estimated from the simulated reconstructions, although the simulation points also have an error from the Poisson sampling of background sightlines. Note that in this analytic calculation, we have assumed that the Wiener filter uses the correct Lyα forest correlation function $\mathbf{S}$, whereas in the simulations we have used the *ad hoc* form in Equation 10 as suggested by Caucci et al. (2008). This could also explain part of the discrepancy seen in Figure 9. Nevertheless, this analytic approach works surprising well and provides us with some intuition as to how the tomographic reconstructions scale with the quality and quantity of survey data.

An interesting insight from this analysis is that in the hypothetical case of infinite spectral signal-to-noise, $\bar{n}_{\mathrm{eff}} \to n_{\mathrm{los}}$ such that the noise term in the residual power spectrum does not go to zero. In other words, for a finite set of skewers there is always a shot-noise contribution from the random sampling of sightlines even in the absence of pixel noise in the spectra. We find that as t_{exp} goes to infinity, SNR_ϵ asymptotes to a value that increases with n_{los} or, equivalently, decreases with $\langle d_\perp\rangle$. For example, if we set $\epsilon_{\mathrm{3D}} = \langle d_\perp\rangle$, then we find $\mathrm{SNR}_{\epsilon,\infty} = [2.56, 2.17, 1.88]$ for $\epsilon_{\mathrm{3D}} = [2,3,5]\ h^{-1}\,\mathrm{Mpc}$. The reason for this is that the Lyα forest variance, $\sigma_F^2(\epsilon_{\mathrm{3D}})$ declines with ϵ_{3D} (dashed line in Figure 10), but the variance from the shot noise term,

$$\begin{aligned} \sigma_N^2(\epsilon_{\mathrm{3D}}) = \frac{1}{2\pi}\int_0^\infty k^2\mathrm{d}k \int_{-1}^1 \mathrm{d}\mu \\ \times\, n_{\mathrm{los}}^{-1}P_{\mathrm{los}}(k,\mu)\,W^2(k\epsilon_{\mathrm{3D}}), \quad (25) \end{aligned}$$

remains roughly constant if we assume $\epsilon_{\mathrm{3D}} = \langle d_\perp\rangle$. This is because as n_{los} is decreased, ϵ_{3D} increases and hence P_{los} is also smoothed over a larger window which reduces its value and keeps the ratio of both quantities roughly constant. This shot-noise variance σ_N^2 is shown by the

solid curve in Figure 10. At $\epsilon_{3D} \gtrsim 3\ h^{-1}$ Mpc, the shot-noise variance dominates over the flux variance. This behaviour is reflected by the fact that to achieve reconstructions at fixed SNR$_\epsilon$, the ratio $\epsilon_{3D}/\langle d_\perp \rangle$ decreases with $\langle d_\perp \rangle$ as can be seen in Table 1.

So far in this paper, we have tied the availability of background sightlines, $n_{\rm los}$, to the exposure time, $t_{\rm exp}$, based on the assumption that a minimum signal-to-noise must be achieved on the source continuum to obtain a source redshift. This is a conservative assumption, as $\sim 50\%$ of LBGs have Lyα emission lines (Shapley et al. 2003) that enable a redshift measurement at fainter magnitudes. Although these fainter spectra will have noisy Lyα forest pixels, they will reduce the shot-noise in the survey and allow significant improvements in the tomographic reconstruction, whether in terms of ϵ_{3D} or SNR$_\epsilon$. With our analytic derivation in hand, we are in a position to rapidly explore more complicated survey scenarios. For example, we find that a survey with $[t_{\rm exp}, n_{\rm los}, g_{\rm lim}] = [4\,{\rm hrs}, 360\,{\rm deg}^{-2}, 24.0]$, which obeys the S/N ≥ 4 per angstrom requirement, enables a $\epsilon_{3D} = 5\ h^{-1}$ Mpc reconstructed map with SNR$_\epsilon = 2.38$. If we keep $t_{\rm exp}$ fixed but observe ≈ 1000 additional faint targets per square degree down to $g_{\rm lim} = 24.5$, and assume a 50% success rate in measuring the redshift, this gives ~ 500 additional sightlines that contribute to the reconstruction. This improves the reconstruction fidelity, which is now increased to SNR$_\epsilon = 2.73$ at the same $\epsilon_{3D} = 5\ h^{-1}$ Mpc. Alternatively, one might choose to smooth the map to a smaller scale, $\epsilon_{3D} = 4.2\ h^{-1}$ Mpc, to retain the same reconstruction fidelity SNR$_\epsilon = 2.38$ as the fiducial survey.

This strategy only works at $\epsilon_{3D} \gtrsim 3\ h^{-1}$ Mpc where the shot-noise contribution is dominant over the intrinsic forest power. At smaller scales ($\epsilon_{3D} \lesssim 2\ h^{-1}$ Mpc), there is little gained from adding additional sightlines, and the only way to improve SNR$_\epsilon$ is to increase the integration times.

4. DISCUSSION

Beyond what we have presented in this paper, there are questions regarding Lyα forest tomography that we have not discussed.

An important issue that needs to be addressed is that of continuum-fitting in the LBG spectra. At the low signal-to-noise ratios that comprise the typical spectra in a tomographic survey sample (Figure 3), it would not be possible to directly estimate the continuum from regions in the Lyα forest deemed free from absorption (e.g., Savaglio et al. 2002). However, while there are likely to be inhomogeneities in the intrinsic LBG spectrum between restframe Lyα and Lyβ wavelengths, these occur mostly from stellar/interstellar medium absorption lines that can be masked once the LBG redshift is known — Shapley et al. (2003) have shown that these are relatively limited in number within the Lyα forest wavelength range of the average LBG at those redshifts. The remaining spectral undulations appear to be of order $\lesssim 20\%$ RMS in star-forming galaxies with sub-solar metallicities (c.f. Leitherer et al. 2011; Heckman et al. 2011). These can be treated using the various PCA techniques that have been developed to estimate quasar continua (Suzuki et al. 2005; Pâris et al. 2011; Lee et al. 2012) from $\lambda > 1216$ Å. Indeed, these techniques are

likely to work better on LBGs than quasars: the spectra on either side of the intrinsic galaxy Lyα wavelength are physically correlated, whereas in the case of quasars there is little correlation between the continuum-slopes either side of $\lambda \sim 1200$ Å (Telfer et al. 2002), the physical origin of which is still a mystery. In any case, we have shown in this paper that the signal-to-noise requirements for a tomographic survey are quite modest, so the continuum estimation do not have to perform much better than the $\sim 15 - 25\%$ pixel noise RMS at the survey limiting magnitude. However, the need for continuum estimation means that a reasonable portion of the LBG spectra at $\lambda > 1216$ Å needs to be observed in order to identify the intrinsic absorption lines in a region free of Lyα forest absorption.

The tomographic reconstruction algorithm also could benefit from further optimization. Although the Wiener interpolation-based scheme we have used here seems to work reasonably well, there are various *ad hoc* aspects that could be improved upon. For example, the correlation function (Equation 10) that relates the data cells to the mapping volume should in principle be derived from recent measurements of the 3D Lyα forest correlation function (e.g., Slosar et al. 2011). The algorithm could also be improved by incorporating adaptive and/or anisotropic smoothing as the final step, instead of the isotropic Gaussian smoothing we are currently implementing. Moreover, techniques need to be refined to deconvolve the dark-matter overdensity field and peculiar velocity fields from the 3D absorption maps, e.g. along the vein of Nusser & Haehnelt (1999) and Pichon et al. (2001). These techniques will need to take into account the uncertainties in the IGM thermal parameters such as the temperature at mean density, temperature-density relationship and Jeans' scale.

We have also not touched on the practical issue of target selection: the background LBGs and quasars need to be identified as such from imaging data in order to be targeted for spectroscopy. The selection of $z \sim 2$ LBGs through the BX criterion is relatively efficient ($\sim 85\%$ at $\mathcal{R} \gtrsim 23.5$), but this is a relatively broad-brush selection that captures LBGs in the range $1.7 < z < 2.6$ (Steidel et al. 2004). For a tomographic survey it is desirable to have a fine-tuned target selection to identify the sources that maximize Lyα forest coverage at the redshifts of interest. It would therefore be useful to apply sophisticated density estimation techniques (e.g., Richards et al. 2004; Bovy et al. 2011) on training sets of known LBGs to reduce the contamination rate. Quasar target selection has evolved considerably in recent years, and with color-color selection alone the success rate is $\sim 60\%$ (e.g. Ross et al. 2012; Bovy et al. 2011), although this increases significantly if UV or IR data is included (Bovy et al. 2012). Since the first generation of tomographic surveys are likely to be in well-studied fields such as COSMOS (Scoville et al. 2007) or GOODS (Dickinson et al. 2003) that have ample multi-wavelength data, we expect target selection to be straightforward, and in any case the spectrographs (see § 4.2) will have target densities of order several thousand per square degree, affording a relatively inefficient ($\lesssim 50 - 60\%$) target selection assuming $n_{\rm los} \lesssim 1000\,{\rm deg}^{-2}$ — this is the success rate of the Steidel et al. (2004) sam-

ple, which used only 3-color selection.

4.1. *Potential Science Applications*

Due to the hitherto distant prospects for mapping large-scale structure at $z \sim 2$, there has been a dearth of literature on its scientific possibilities. Here, we outline some possible applications for such maps. Some of these themes will be explored in more detail in subsequent papers, and we also invite others to contribute their ideas.

Galaxy Environments: The $z \sim 2 - 3$ epoch is a particularly interesting time for the study of galaxy formation, as the star-formation rate density of the Universe peaked at this epoch, whilst the present-day Hubble sequence of galaxies were still in the process of being assembled. However, while the galaxies in this epoch have been intensively studied both through imaging and spectroscopy, it has been extremely challenging to study them as a function of environment (although see Diener et al. 2013). Lyα forest tomographic maps on scales of $\epsilon_{3D} \sim 1 - 4\ h^{-1}\,\mathrm{Mpc}$ would reveal the impact of environment on the various galaxy properties at $z \sim 2$, e.g. color, morphology, star-formation rate, gas properties. Indeed, since the Lyα forest absorption is a continuous tracer of the underlying dark matter field at overdensities of $\delta\rho/\langle\rho\rangle \sim 1$, it would be a less biased tracer of the density field than even the most comprehensive galaxy redshift surveys (Figure 4). However, there are few papers in the literature that directly study galaxy properties as a function of the underlying DM field (as opposed to using groups or clusters), so some work needs to be done to tie theories of galaxy formation directly to the Lyα forest field. Moreover, the $\epsilon_{3D} \sim 1\ h^{-1}\,\mathrm{Mpc}$ scale of the most ambitious tomographic reconstructions (e.g. right panel of Figure 4) corresponds to $\sim 400\,\mathrm{kpc}$ proper distance at those redshifts. This approaches the 'circumgalactic medium' (CGM) scale for a typical LBG, and would provide valuable insights into the gaseous H I environment of such galaxies. However, since the size of the CGM scales with virial radius, coarser-resolution tomographic maps with ϵ_{3D} of several Mpc could already probe the CGM of galaxy groups and protoclusters.

Galaxy Protoclusters: Even though massive galaxy clusters have been discovered out to $z \gtrsim 1.5$ (e.g., Nastasi et al. 2011; Gobat et al. 2011; Brodwin et al. 2012; Muzzin et al. 2013), the mechanisms for their formation is not well-understood due to the difficulty in mapping large-scale structure at $z \gtrsim 1$. In the hierarchical view of large-scale structure, galaxy clusters are believed to form through the mergers of smaller protoclusters at the intersections of the filaments in the cosmic web. With large-scale structure maps from Lyα forest tomography, it should thus be possible to directly identify cluster progenitors. It is likely that this should be feasible through coarser maps with resolutions of several $h^{-1}\,\mathrm{Mpc}$; in an upcoming paper we will investigate methods to identify protoclusters through IGM tomographic maps. Selecting galaxy protoclusters directly through through their signature in the large-scale structure would be more systematic and thorough than either serendipitous discoveries (e.g., Steidel et al. 2005) or by targeting galaxy overdensities around high-redshift radio galaxies (Overzier et al. 2006; Venemans et al. 2007; Hatch et al. 2011).

Topology: The topology of large-scale structure has yet to be studied beyond the Local Universe. Caucci et al. (2008) have already shown that Lyα forest tomographic maps can effectively recover the large-scale topology of the Universe at $z \sim 2$. The measurement of cosmic topology at $z \sim 2$ will directly test the idea that large overdensities such as galaxy groups and clusters form at the intersection of filaments in the cosmic web (see, e.g., Bond et al. 1996; Kravtsov & Borgani 2012). With the higher-resolution maps $\epsilon_{3D} \approx 1\ h^{-1}\,\mathrm{Mpc}$, it could also be possible to directly test the cold-flow accretion picture of galaxy formation (Kereš et al. 2005, 2009; Dekel et al. 2009), in which star formation in galaxies is fed by cold gas and dwarf galaxies that stream in along the filaments in the cosmic web. While this mode of accretion is challenging to constrain through traditional 1-dimensional absorption-line studies due to the modest covering-factor of cold streams (Faucher-Giguère & Kereš 2011; Fumagalli et al. 2011, 2014; Hennawi & Prochaska 2013), and the challenge of finding background quasar sightlines at small impact parameter to foreground galaxies (Crighton et al. 2011; Rudie et al. 2012) a tomographic map of the IGM with $\sim 1\ h^{-1}\,\mathrm{Mpc}$ resolution will be a powerful test of this scenario. Another interesting possibility is to use the characteristic topology of inflationary cold dark matter cosmology as a standard ruler to measure the expansion rate of the Universe (Park & Kim 2010; Zunckel et al. 2011). This will require mapping out large volumes, but the scales necessary for the measurement are so large ($\epsilon_{3D} \gtrsim 15\ h^{-1}\,\mathrm{Mpc}$) that this could be feasible through the main surveys of MS-DESI (see below) or even BOSS.

Power Spectrum and Cross-Correlations: Beyond tomographic maps, the highly dense sets of Lyα forest spectra obtained for such surveys will also be powerful for Lyα forest auto-correlation and cross-correlation studies. The BOSS Lyα Forest Survey (Dawson et al. 2013; Lee et al. 2013) has pioneered the measurement of the 3D Lyα forest auto-correlation on scales of $\Delta r \gtrsim 10\ h^{-1}\,\mathrm{Mpc}$ (Slosar et al. 2011; Busca et al. 2013; Slosar et al. 2013), but on smaller ($\Delta r \lesssim 10\ h^{-1}\,\mathrm{Mpc}$) scales the number of available pixel pairs is limited due to the comparatively large transverse separations ($\sim 15'$) between BOSS quasars. A tomographic survey that targets ~ 2000 spectra over 1 square degree would give twice the number of $\Delta r \lesssim 10\ h^{-1}\,\mathrm{Mpc}$ pixel pairs compared with the final ~ 160000 quasar BOSS Lyα forest sample. The measurement of the small-scale Lyα forest 3D auto-correlation (or, equivalently, the 3D power spectrum) would place constraints on the underlying dark-matter power spectrum at that epoch, as well as the gas temperature of the IGM (A. Ariño, in prep). Beyond the auto-correlation, the Lyα forest has proven its utility in cross-correlation with other large-scale structure tracers, such as with damped Lyα systems (DLAs, Font-Ribera et al. 2012), and quasars (Font-Ribera et al. 2013). A dense grid of absorption sightlines would enable cross-correlation studies with e.g., weak-lensing maps (Massey et al. 2007), cosmic microwave background lensing convergence maps (e.g., Sherwin et al. 2012), and the cosmic near-infrared background (e.g., Fernandez et al. 2010).

4.2. *Spectrographs and Telescopes for Lyα Tomography*

We have established in this paper that surprisingly noisy moderate resolution spectra are sufficient to carry out IGM tomographic reconstructions from Lyα forest absorption spectra, in many cases requiring exposure times that are already feasible on the current generation of 8-10m telescopes. Here, we outline several regimes in which Lyα forest tomography can be carried out on existing and near-future instruments that vary in aperture, field-of-view (FOV) and multiplexing.

Medium-FOV Spectrographs on 8m Telescopes: Throughout this paper, we have used the exposure time calculator for the VIMOS spectrograph (Le Fèvre et al. 2003) on the 8m VLT to predict exposure times. As we have seen in Table 1, $t_{\rm exp} = 5$ hrs will yield sufficient signal-to-noise to identify complete samples of $g = 24.1$ LBGs at $z_{\rm bg} \sim 2-3$. This gives sufficient source density to create tomographic maps with $\epsilon_{\rm 3D} \approx 3-4~h^{-1}$ Mpc. With the $4 \times 7' \times 8'$ field-of-view on VIMOS, it would take about 18 pointings to cover an area of $1\,{\rm deg}^2$, or about 140 hrs including overheads. Assuming a line-of-sight distance of $\sim 250~h^{-1}$ Mpc is covered at the same minimum $n_{\rm los}$ (requiring a total projected source density of $\approx 1.8\,n_{\rm los}$; see § 2.1), this corresponds to a comoving volume of $\approx 1.1 \times 10^6~h^{-3}$ Mpc^3. This is equivalent to the volumes covered by the VVDS (Marinoni et al. 2008) and zCOSMOS (Kovač et al. 2010) galaxy redshift maps out to $z \sim 1$, but with a considerably less elongated geometry that will allow the characterization of structures spanning the transverse direction to the line-of-sight. The LRIS spectrograph (Oke et al. 1995) on the 10m Keck telescope has a field-of-view $\frac{1}{4}$ that of VIMOS, but its larger aperture, better blue throughput and lower observing overheads mean that it would take the same overall time to carry out the $1\,{\rm deg}^2$ survey.

Wide-FOV Spectrographs on 8m Telescopes: The Subaru Prime-Focus Spectrograph (Takada et al. 2014) is a $1\,{\rm deg}^2$ field-of-view spectrograph with 2400 fibers currently planned for the 8.2m Subaru telescope. While the spectrograph throughput and telescope aperture is similar to VIMOS, it has the advantage of a much larger ($\sim 18\times$) field-of-view. The $16\,{\rm deg}^2$ galaxy evolution survey described in Takada et al. (2014) involves obtaining redshifts for ~ 1800 LBGs per square degree at $2 < z < 3$, with $t_{\rm exp} = 3$ hrs exposures. This can be used to generate an IGM tomographic map over a comoving volume of $\approx 1.8 \times 10^7~h^{-3}$ Mpc^3 covering $2.1 \leq z \leq 2.4$ — our analytic estimates indicate that a map with $\epsilon_{\rm 3D} = 3.5~h^{-1}$ Mpc (similar to the VIMOS/LRIS map described above) would have a reconstruction fidelity of SNR$_\epsilon \approx 2.4$, i.e. with a quality similar to the 2nd panel from the right in Figure 8. Another possibility is to use the spectrograph for dedicated deep observations with $t_{\rm exp} \sim 10$ hrs of $g_{\rm lim} \approx 24.5$ sources that will reach the sightline densities sufficient for $\epsilon_{\rm 3D} = 1.5~h^{-1}$ Mpc tomographic maps. These scales, corresponding to ≈ 400 kpc physical at $z \sim 2.3$, is close the regime of interest for circumgalactic medium (CGM) studies, and will enable great insights into the process of galaxy formation.

Wide-FOV Spectrographs on 4m Telescopes: The Mid-Scale Dark Energy Spectroscopic Instrument (MS-DESI, Levi et al. 2013) is a 5000-fiber spectrograph

with a $7\,{\rm deg}^2$ field-of-view that is intended to be mounted on the 4m Mayall telescope at the Kitt Peak National Observatory. The main survey will target LRGs and ELGs out to $z \sim 1.4$, and Lyα forest quasars at $z \sim 2.3$ over 14,000 deg^2 to measure the BAO signal. The main Lyα forest survey will target $g \lesssim 23$ quasars at a projected area density of $\sim 50\,{\rm deg}^{-2}$ that is too low for tomographic reconstruction on scales we are interested in. However, the instrument may be made available for community use, so we can envision a dedicated survey for Lyα tomography. Such a survey could aim for $n_{\rm los} \approx 200\,{\rm deg}^{-2}$ to a magnitude limit of $g_{\rm lim} \approx 23.8$ to obtain maps with $\epsilon_{\rm 3D} \approx 7~h^{-1}$ Mpc. This would require $t_{\rm exp} \approx 12$ hrs integrations per pointing (we have simply rescaled the VIMOS $t_{\rm exp}$ values to a 4m aperture). Since each pointing would cover a volume of $\approx 8\times10^6~h^{-1}$ Mpc, despite the long exposure times it will be possible to build up large volumes rapidly to, e.g., identify large numbers of galaxy protoclusters or precisely measure the topology of large-scale structure.

5. CONCLUSIONS

In this paper, we have conducted a detailed study of the observational requirements necessary to carry out IGM tomography, at various map resolutions, by Wiener interpolation over dense grids of Lyα forest absorption spectra. Using empirical luminosity functions for background quasars and LBGs, the transverse separation between sightlines is $\langle d_\perp \rangle = [1, 2, 3, 5]~h^{-1}$ Mpc at limiting magnitudes of $g_{\rm lim} \approx [25.0, 24.5, 24.1, 23.7]$. With the *ansatz* that the source separation roughly sets the resolution $\epsilon_{\rm 3D}$ of the reconstructed map, this argues that tomographic reconstructions on scales of $\epsilon_{\rm 3D} = 3-5~h^{-1}$ Mpc is feasible with moderate-resolution ($R \sim 1000$) spectra of $g_{\rm lim} \approx 24$ sources, which is accessible to the current generation of 8-10m telescopes and instrumentation.

We directly tested this with Wiener reconstructions of mock Lyα forest absorption spectra generated from numerical simulations, in which we have added realistic pixel noise based on the source luminosity functions and assumed telescope exposure times, $t_{\rm exp}$. Assuming a conservative signal-to-noise requirement (S/N ≥ 4 per angstrom), necessary to measure the source redshifts, exposure times of $t_{\rm exp} = [3, 5, 8, 16]$ hrs on an 8m telescope can obtain background sightlines at sufficient source densities to create good-quality (SNR$_\epsilon = 2.5$) tomographic maps with spatial resolutions of $\epsilon_{\rm 3D} = [7.2, 3.8, 2.7, 1.4]~h^{-1}$ Mpc (e.g., Figure 4). We also show that for given set of sightlines, one can choose between a finer map resolution, $\epsilon_{\rm 3D}$, or a better reconstruction fidelity, SNR$_\epsilon$.

We also derived an analytic expression that allows us to compute SNR$_\epsilon$ as a function of $\epsilon_{\rm 3D}$ and signal-to-noise properties of a given set of sightlines. This shows the dominance, on scales $\epsilon_{\rm 3D} \gtrsim 3~h^{-1}$ Mpc, of shot-noise in the sampling of the sightlines — this argues, on these coarser scales, for obtaining additional, fainter, sightlines over increasing $t_{\rm exp}$ to improve SNR$_\epsilon$ or reduce $\epsilon_{\rm 3D}$.

These findings motivate a survey targeted at obtaining Lyα forest spectra from faint $g_{\rm lim} \approx 24$ LBGs over ~ 1 square degree on existing 8-10m telescopes and spectrographs, which would enable tomographic reconstructions with resolutions of $\epsilon_{\rm 3D} \sim 3-4~h^{-1}$ Mpc over $\sim$

$10^6 \ h^{-3} \, \mathrm{Mpc}^3$ of comoving volume. Such a survey would require exposure times of $4 - 5 \, \mathrm{hrs}$ on VLT-VIMOS or $2 - 2.5 \, \mathrm{hrs}$ on Keck LRIS, with a total time requirement of $\sim 130 \, \mathrm{hrs}$ (including overheads) to cover 1 square degree. These exposure times and magnitude limits are comparable to contemporary high-z galaxy redshift surveys (e.g., Steidel et al. 2004; Lilly et al. 2007; Le Fèvre et al. 2013), but the data from these surveys are not suitable for tomographic reconstruction since they are of low resolution ($R \lesssim 400$), which does not provide adequate line-of-sight sampling for our purposes.

IGM tomography is far more efficient than galaxy redshift surveys at mapping out $z \sim 2$ large-scale structure, particularly at low overdensities ($\rho/\langle\rho\rangle \lesssim 10$), due to the $\sim 400 - 500 \ h^{-1} \, \mathrm{Mpc}$ probed along the line-of-sight of Lyα forest spectrum; a tomographic survey requires only a dense *areal* sampling rather than *volume* sampling that is required for comparable galaxy surveys. Our fiducial 1 square degree, $\epsilon_{3\mathrm{D}} \sim 3 \ h^{-1} \, \mathrm{Mpc}$ tomographic map is already a factor of ~ 2 better, resolution-wise, compared to the Kitaura et al. (2009) map of $\sim 250000 \, \mathrm{SDSS}$ galaxies at $z \sim 0.1$, which had an effective resolution of $\sim 7 \ h^{-1} \, \mathrm{Mpc}$ over a $350^3 \ h^{-3} \, \mathrm{Mpc}^3$ volume. To achieve a similar galaxy number density ($n_{\mathrm{gal}} \approx 0.005 \, h^3 \, \mathrm{Mpc}^{-3}$) at $z \sim 2$ would require obtaining redshifts for a *volume-limited sample of $\mathcal{R} \approx 26$ galaxies*. Even with 30m-class telescopes, it would be very expensive to obtain such samples over cosmologically interesting volumes.

With the wide-field spectrographs slated to become available in the near-future, IGM tomography will be improved in two distinct regimes: (a) the Subaru PFS spectrograph on the 8m Subaru telescope will allow suffi- ciently deep ($t_{\mathrm{exp}} \gtrsim 10 \, \mathrm{hrs}$) exposures over ~ 1 square degree fields to enable cosmography at $\epsilon_{3\mathrm{D}} \approx 2 \ h^{-1} \, \mathrm{Mpc}$, a mapping scale achievable through galaxy redshift surveys only within the immediate ($z \lesssim 0.03$) Local Universe (e.g., Courtois et al. 2013). (b) the MS-DESI spectrograph on the 4m KPNO Mayall telescope, which would enable Lyα tomography with map resolutions of $\epsilon_{3\mathrm{D}} \sim 7 \ h^{-1} \, \mathrm{Mpc}$ over comoving volumes of $\approx 8 \times 10^6 \ h^{-3} \, \mathrm{Mpc}^3$ per $t_{\mathrm{exp}} \approx 10 \, \mathrm{hrs}$ pointing, enabling huge volumes to be efficiently mapped.

The IGM tomographic maps will open up new science possibilities at $z \sim 2$. In this paper we have outlined a few applications, such as studying galaxy environments at $z \sim 2$, searching for galaxy protoclusters, and measuring the topology of large-scale structure. However, much observational and theoretical work needs to be carried out before Lyα forest tomography becomes a useful scientific tool. We therefore invite the community to contribute their intellectual energies to this exciting and potentially very fruitful new observational technique.

REFERENCES

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, ApJS, 203, 21

Becker, G. D., Hewett, P. C., Worseck, G., & Prochaska, J. X. 2013, MNRAS, 430, 2067

Bi, H., & Davidsen, A. F. 1997, ApJ, 479, 523

Bi, H., Ge, J., & Fang, L.-Z. 1995, ApJ, 452, 90

Blanton, M. R., Eisenstein, D., Hogg, D. W., Schlegel, D. J., & Brinkmann, J. 2005, ApJ, 629, 143

Bolton, J. S., Viel, M., Kim, T., Haehnelt, M. G., & Carswell, R. F. 2008, MNRAS, 386, 1131

Bond, J. R., Kofman, L., & Pogosyan, D. 1996, Nature, 380, 603

Bouwens, R. J., Illingworth, G. D., Franx, M., et al. 2009, ApJ, 705, 936

Bovy, J., Hennawi, J. F., Hogg, D. W., et al. 2011, ApJ, 729, 141

Bovy, J., Myers, A. D., Hennawi, J. F., et al. 2012, ApJ, 749, 41

Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13

Brodwin, M., Gonzalez, A. H., Stanford, S. A., et al. 2012, ApJ, 753, 162

Busca, N. G., Delubac, T., Rich, J., et al. 2013, A&A, 552, A96

Calura, F., Tescari, E., D'Odorico, V., et al. 2012, MNRAS, 422, 3019

Caucci, S., Colombi, S., Pichon, C., et al. 2008, MNRAS, 386, 211

Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Comparat, J., Kneib, J.-P., Escoffier, S., et al. 2013, MNRAS, 428, 1498

Courtois, H. M., Pomarède, D., Tully, R. B., Hoffman, Y., & Courtois, D. 2013, AJ, 146, 69

Crighton, N. H. M., Bielby, R., Shanks, T., et al. 2011, MNRAS, 414, 28

Croft, R. A. C., Weinberg, D. H., Bolte, M., et al. 2002, ApJ, 581, 20

Croft, R. A. C., Weinberg, D. H., Katz, N., & Hernquist, L. 1997, ApJ, 488, 532

—. 1998, ApJ, 495, 44

Davis, M., Huchra, J., Latham, D. W., & Tonry, J. 1982, ApJ, 253, 423

Davis, M., Faber, S. M., Newman, J., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4834, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. P. Guhathakurta, 161–172

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10

Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Nature, 457, 451

Dickinson, M., Giavalisco, M., & GOODS Team. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini, 324

Diener, C., Lilly, S. J., Knobel, C., et al. 2013, ApJ, 765, 109

Drinkwater, M. J., Jurek, R. J., Blake, C., et al. 2010, MNRAS, 401, 1429

Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971

Evans, C. J., Barbuy, B., Bonifacio, P., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Faucher-Giguère, C., Lidz, A., Hernquist, L., & Zaldarriaga, M. 2008, ApJ, 682, L9

Faucher-Giguère, C.-A., & Kereš, D. 2011, MNRAS, 412, L118

Fernandez, E. R., Komatsu, E., Iliev, I. T., & Shapiro, P. R. 2010, ApJ, 710, 1089

Font-Ribera, A., Miralda-Escudé, J., Arnau, E., et al. 2012, JCAP, 11, 59

Font-Ribera, A., Arnau, E., Miralda-Escudé, J., et al. 2013, JCAP, 5, 18

Fumagalli, M., Hennawi, J. F., Prochaska, J. X., et al. 2014, ApJ, 780, 74

Fumagalli, M., Prochaska, J. X., Kasen, D., et al. 2011, MNRAS, 418, 1796

Garzilli, A., Bolton, J. S., Kim, T.-S., Leach, S., & Viel, M. 2012, MNRAS, 424, 1723

Geller, M. J., & Huchra, J. P. 1989, Science, 246, 897

Giavalisco, M., Steidel, C. C., & Macchetto, F. D. 1996, ApJ, 470, 189

Gnedin, N. Y., & Hui, L. 1998, MNRAS, 296, 44

Gobat, R., Daddi, E., Onodera, M., et al. 2011, A&A, 526, A133

Gómez, P. L., Nichol, R. C., Miller, C. J., et al. 2003, ApJ, 584, 210

Greig, B., Bolton, J. S., & Wyithe, J. S. B. 2011, MNRAS, 418, 1980

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35

Hatch, N. A., De Breuck, C., Galametz, A., et al. 2011, MNRAS, 410, 1537

Heckman, T. M., Borthakur, S., Overzier, R., et al. 2011, ApJ, 730, 5

Hennawi, J. F., & Prochaska, J. X. 2013, ApJ, 766, 58

Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, ApJ, 457, L51+

Herschel, W. 1785, Royal Society of London Philosophical Transactions Series I, 75, 213

Hill, G. J., Gebhardt, K., Komatsu, E., et al. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 399, Panoramic Views of Galaxy Formation and Evolution, ed. T. Kodama, T. Yamada, & K. Aoki, 115

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Hubble, E. P. 1926, ApJ, 64, 321

Humason, M. L., Mayall, N. U., & Sandage, A. R. 1956, AJ, 61, 97

Kaiser, N. 1987, MNRAS, 227, 1

Kauffmann, G., White, S. D. M., Heckman, T. M., et al. 2004, MNRAS, 353, 713

Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2

Kirkby, D., Margala, D., Slosar, A., et al. 2013, JCAP, 3, 24

Kitaura, F.-S., Gallerani, S., & Ferrara, A. 2012, MNRAS, 420, 61

Kitaura, F. S., Jasche, J., Li, C., et al. 2009, MNRAS, 400, 183

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36

Kovač, K., Lilly, S. J., Cucciati, O., et al. 2010, ApJ, 708, 505

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, ApJ, 609, 35

Kravtsov, A. V., & Borgani, S. 2012, ARA&A, 50, 353

Kreckel, K., Platen, E., Aragón-Calvo, M. A., et al. 2011, AJ, 141, 4

Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4841, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. M. Iye & A. F. M. Moorwood, 1670–1681

Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, A&A, 559, A14

Le Goff, J. M., Magneville, C., Rollinde, E., et al. 2011, A&A, 534, A135

Lee, K.-G., Suzuki, N., & Spergel, D. N. 2012, AJ, 143, 51

Lee, K.-G., Bailey, S., Bartsch, L. E., et al. 2013, AJ, 145, 69

Lee, K.-G., Hennawi, J. P., Spergel, D. N., et al. 2014, arXiv:1405.1072

Leitherer, C., Tremonti, C. A., Heckman, T. M., & Calzetti, D. 2011, AJ, 141, 37

Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv:1308.0847

Lewis, I., Balogh, M., De Propris, R., et al. 2002, MNRAS, 334, 673

Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70

Lynds, R. 1971, ApJ, 164, L73

Marinoni, C., Guzzo, L., Cappi, A., et al. 2008, A&A, 487, 7

Massey, R., Rhodes, J., Ellis, R., et al. 2007, Nature, 445, 286

McDonald, P. 2003, ApJ, 585, 34

McDonald, P., & Eisenstein, D. J. 2007, Phys. Rev. D, 76, 063009

McDonald, P., Miralda-Escudé, J., Rauch, M., et al. 2000, ApJ, 543, 1

McDonald, P., Seljak, U., Cen, R., Bode, P., & Ostriker, J. P. 2005, MNRAS, 360, 1471

McDonald, P., Seljak, U., Burles, S., et al. 2006, ApJS, 163, 80

McQuinn, M., Hernquist, L., Lidz, A., & Zaldarriaga, M. 2011, MNRAS, 415, 977

McQuinn, M., & White, M. 2011, MNRAS, 415, 2257

Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M. 1996, ApJ, 471, 582

Muzzin, A., Wilson, G., Demarco, R., et al. 2013, ApJ, 767, 39

Nastasi, A., Fassbender, R., Böhringer, H., et al. 2011, A&A, 532, L6

Nusser, A., & Haehnelt, M. 1999, MNRAS, 303, 179

Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, PASP, 107, 375

Overzier, R. A., Miley, G. K., Bouwens, R. J., et al. 2006, ApJ, 637, 58

Palanque-Delabrouille, N., Magneville, C., Yèche, C., et al. 2013, A&A, 551, A29

Pâris, I., Petitjean, P., Rollinde, E., et al. 2011, A&A, 530, A50

Pâris, I., Petitjean, P., Aubourg, É., et al. 2012, A&A, 548, A66

Park, C., & Kim, Y.-R. 2010, ApJ, 715, L185

Peirani, S., Weinberg, D. H., Colombi, S., et al. 2014, ApJ, 784, 11

Pichon, C., Vergely, J. L., Rollinde, E., Colombi, S., & Petitjean, P. 2001, MNRAS, 326, 597

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in FORTRAN. The art of scientific computing (Cambridge University Press)

Rauch, M., Miralda-Escude, J., Sargent, W. L. W., et al. 1997, ApJ, 489, 7

Reddy, N. A., Steidel, C. C., Pettini, M., et al. 2008, ApJS, 175, 48

Reid, B. A., Samushia, L., White, M., et al. 2012, MNRAS, 426, 2719

Richards, G. T., Nichol, R. C., Gray, A. G., et al. 2004, ApJS, 155, 257

Rorai, A., Hennawi, J. F., & White, M. 2013, ApJ, 775, 81

Ross, N. P., Myers, A. D., Sheldon, E. S., et al. 2012, ApJS, 199, 3

Rudie, G. C., Steidel, C. C., & Pettini, M. 2012, ApJ, 757, L30

Savaglio, S., Panagia, N., & Padovani, P. 2002, ApJ, 567, 702

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1

Shapley, A. E. 2011, ARA&A, 49, 525

Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65

Shectman, S. A., Landy, S. D., Oemler, A., et al. 1996, ApJ, 470, 172

Sherwin, B. D., Das, S., Hajian, A., et al. 2012, Phys. Rev. D, 86, 083006

Slosar, A., Font-Ribera, A., Pieri, M. M., et al. 2011, JCAP, 9, 1

Slosar, A., Iršič, V., Kirkby, D., et al. 2013, JCAP, 4, 26

Steidel, C., Martin, C., Prochaska, J. X., et al. 2009, in Astronomy, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey, 286

Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2005, ApJ, 626, 44

Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289

Steidel, C. C., & Hamilton, D. 1993, AJ, 105, 2017

Steidel, C. C., Shapley, A. E., Pettini, M., et al. 2004, ApJ, 604, 534

Sutter, P. M., Lavaux, G., Wandelt, B. D., & Weinberg, D. H. 2012, ApJ, 761, 44

Suzuki, N., Tytler, D., Kirkman, D., O'Meara, J. M., & Lubin, D. 2005, ApJ, 618, 592

Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, 1

Tasitsiomi, A., Kravtsov, A. V., Wechsler, R. H., & Primack, J. R. 2004, ApJ, 614, 533

Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, ApJ, 606, 702

Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, ApJ, 565, 773

Tinker, J. L., Conroy, C., Norberg, P., et al. 2008, ApJ, 686, 53

Vale, A., & Ostriker, J. P. 2004, MNRAS, 353, 189

Venemans, B. P., Röttgering, H. J. A., Miley, G. K., et al. 2007, A&A, 461, 823

Viel, M., Bolton, J. S., & Haehnelt, M. G. 2009, MNRAS, 399, L39

Viel, M., Haehnelt, M. G., & Springel, V. 2004, MNRAS, 354, 684

Weinberg, D. H., Davé, R., Katz, N., & Kollmeier, J. A. 2003, in American Institute of Physics Conference Series, Vol. 666, The Emergence of Cosmic Structure, ed. S. H. Holt & C. S. Reynolds, 157–169

White, M. 2002, ApJS, 143, 241

White, M., Cohn, J. D., & Smit, R. 2010, MNRAS, 408, 1818

Wiener, N. 1942, 'The interpolation, extrapolation and smoothing of stationary time series' (MIT)

Zaldarriaga, M., Scoccimarro, R., & Hui, L. 2003, ApJ, 590, 1

Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995, ApJ, 449, 446

Zhang, Y., Anninos, P., & Norman, M. L. 1995, ApJ, 453, L57+

Zunckel, C., Gott, J. R., & Lunnan, R. 2011, MNRAS, 412, 1401