Measuring large-scale structure with the Lyα forest

Martin White
UCB/LBNL

Santa Fe, 2011

LSS and the IGM

The IGM can be used to constrain both small-scale power and large-scale power at high redshift.

- For BAO, distance constraints become tighter as one moves to higher *z*
  - More volume per sky area.
  - Less non-linearity.
- Expensive if use galaxies as tracers.
- Any tracer will do: HI
  - 21cm from HI in galaxies: SKA or custom expt.
  - Lyα from IGM as probed by QSOs.
    - Absorption traces mass in a calculable way.
    - A dense grid of QSO sightlines could probe BAO
      - (White 2003, McDonald & Eisenstein 2007, Slosar++09, White++10, McQuinn & White 2011)

The future: 3D imaging

- In the past, Lyα forest surveys were analyzed as a large number of 1D “skewers” through the density field.
- Next generation surveys (and BOSS) achieve densities of 10-100 QSOs/deg^2.
- Cross-correlations between sightlines become relevant.
  - Have a poorly sampled 3D field, with an odd window function, not a large number of well sampled 1D fields.
- This allows S/N>1 on tens of Mpc modes in 3D.
- On scales <1Mpc, auto-correlations win.
- On scales >1Mpc, cross-correlations can offer huge statistical and systematic wins.

Impacts

- Ability to measure large-scale structure at z=2-3 allows:
  - Better constrains on distance-redshift relation.
    - Measure of $H(z)$ at z~2.5.
    - Constraint on Ω_K at 10^{-3} level.
  - Measurement of temperature fluctuations from He reionization or intensity fluctuations from UVbg sources.
    - If could push to z~4, may be able to detect T fluctuations from H reionization too! (McQuinn++10)
- Amazingly a single number characterizes the sensitivity of a survey to the 3D flux power spectrum (c.f. FKP).

FGPA

- Physics of the forest is “straightforward”.
  - Gas making up the IGM is in photo-ionization equilibrium with a (uniform?) ionization field which results in a tight ρ-T relation for the absorbing material
  - The HI density is proportional to a power of the baryon density.
  - Since pressure forces are sub-dominant on “large” scales, the gas traces the dark matter (0.1-10Mpc/h).
  - The structure in the QSO spectrum thus traces, in a calculable way, the fluctuations in the matter density along the line-of-sight to the QSO.

On large scales

- Differences with the galaxies
  - Signal is $e^{-\tau}$, so downweights high-δ.
  - Need to be slightly careful about redshift space distortions (τ conserved, not *n*, except in line-dominated regime).
  - Projection/finite sampling.
- Additional physics
  - Absorption could be affected by non-gravitational physics
    - Fluctuations in the UV background
    - Temperature fluctuations due to HeII reionization
    - Your favorite astrophysical phenomenon here.

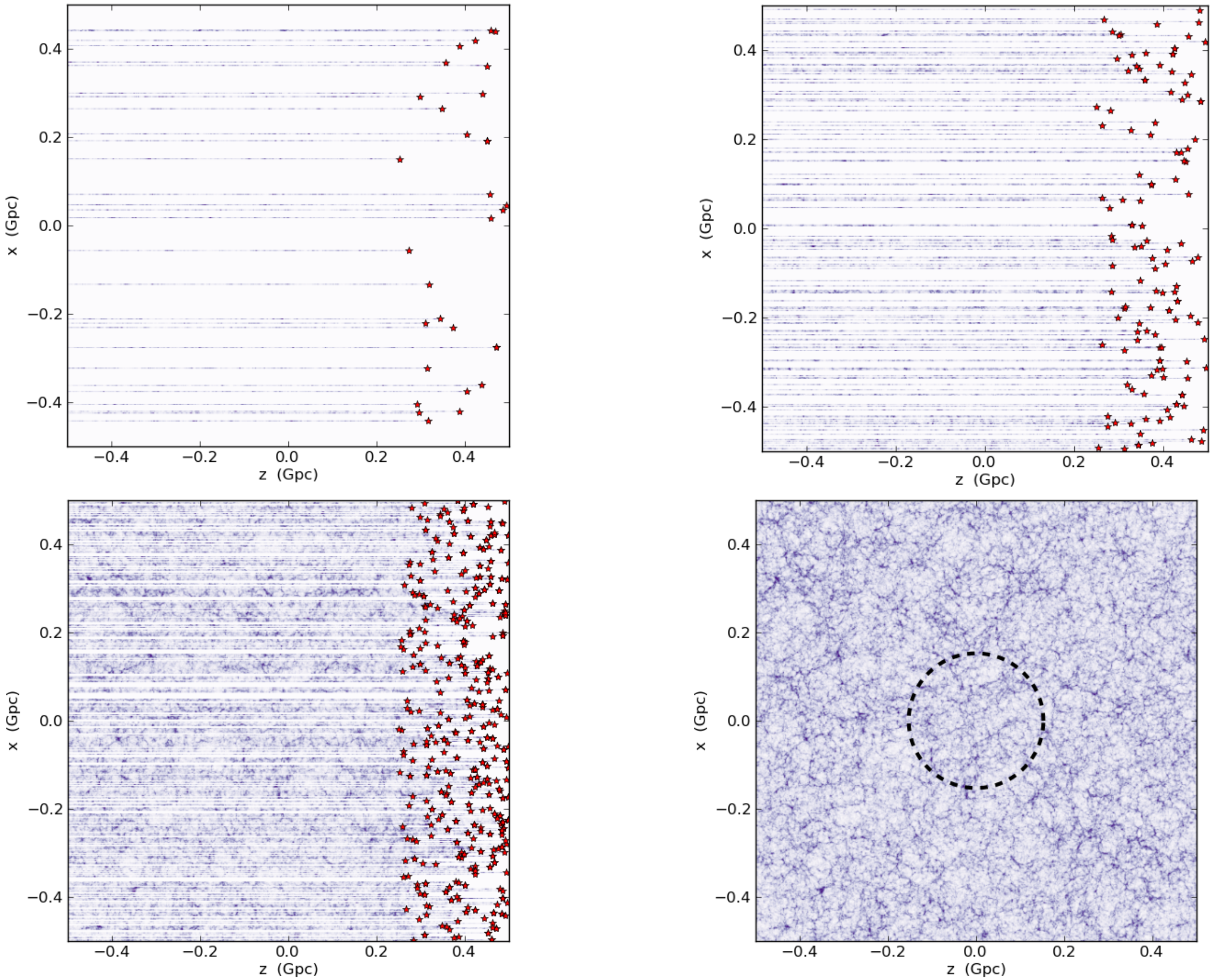
Orientation: distances & redshifts

z	λ_α	<F>	$\Delta\chi$	$d\lambda/d\chi$	$dv/d\chi$	b
2.0	3647	0.88	575	1.11	91	0.12
2.5	4255	0.80	546	1.37	97	0.18
3.0	4863	0.70	518	1.66	102	0.27

$$P_N = \frac{0.8}{\langle F\rangle^2}\ [S/N]^{-2}_{\Delta\lambda}\left(\frac{\Delta\lambda}{1\mathring{A}}\right)\left(\frac{1+z}{4}\right)^{-3/2}$$

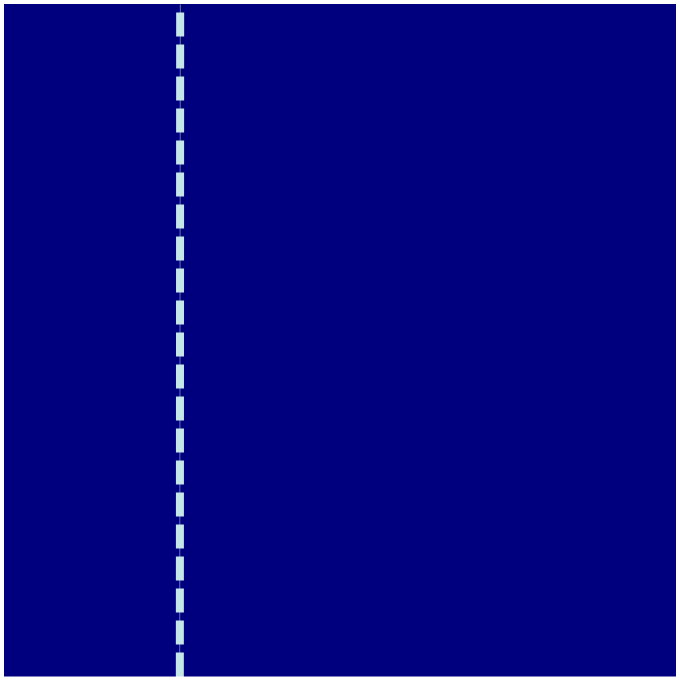
Distances are (comoving) Mpc/h, wavelengths in Å and velocities in km/s.

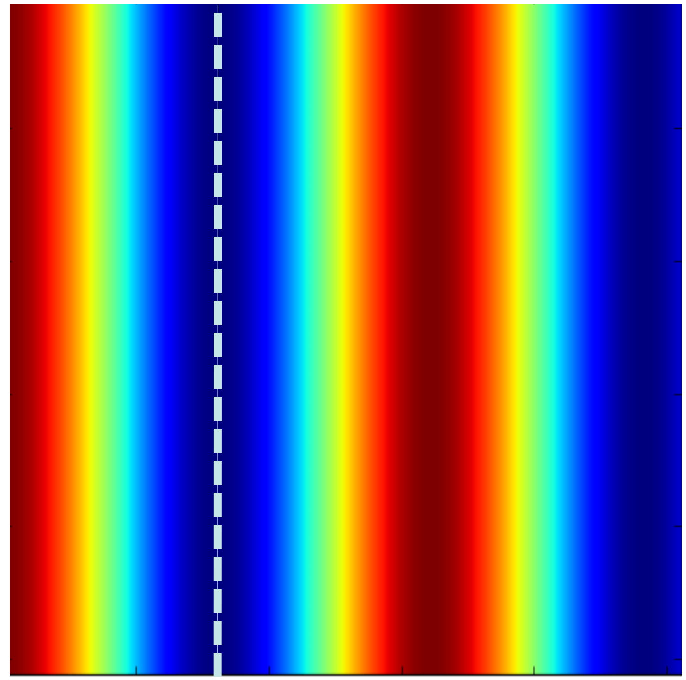
Skewer density



Aliasing

$$P_{\rm los}(k_{||}) = \int \frac{d^2\mathbf{k}_\perp}{(2\pi)^2} P_F(k_{||}, \mathbf{k}_\perp)$$





Can't tell the difference between a constant field (k_x=k_y=k_z=0) and one varying transverse to the line-of-sight (k_x>0 or k_y>0)

Skewer density

- Looking along a finite number of sightlines leads to power aliasing.
  - As the number of sightlines increases this aliasing is tamed, eventually we reach sample variance.
- Aliasing = sample variance at a critical number density of sightlines (about 50 quasars/sq. deg.)
  - Set by the ratio of P_F to P_{los}
  - The more small-scale power the larger P_{los} at fixed P_F, and the more skewers you need.

The flux over"density"

- Define $\delta_F(\boldsymbol{x})=F(\boldsymbol{x})/\langle F\rangle-1$ with FT $\delta(\boldsymbol{k})$.
- $P_F(\boldsymbol{k})=b^2(z)[1+g\mu^2]^2P_{lin}(|\boldsymbol{k}|)\mathrm{Exp}[-k^2/k_D^2]$
  - $g\sim1$ and $b\sim0.2$, $k_D\sim0.1$km/s.
- Window function $W(\boldsymbol{x})\sim\Sigma\ \delta^{(D)}(\boldsymbol{x}_p-\boldsymbol{x}_{pn})$
- And the survey measures $\delta_F(\boldsymbol{x})W(\boldsymbol{x})$.
- Take the FT:
  - The line-of-sight is straightforward.
  - Transverse the quasars provide a Monte-Carlo sampling of the FT integral.
  - Allow weights, *w*, per skewer.

Large-N limit

if omit self-pairs

- In the limit N➔∞

$$P_{\rm obs} = P_F(\mathbf{k}) + \bar{n}^{-1}\left[\cancelto{0}{P_N} + \bar{w}^2 P_{\rm los}(k_{||})\right]$$

- where

$$P_{\rm los} \equiv \int \frac{d^2\mathbf{k}_\perp}{(2\pi)^2}\; P_F(\mathbf{k}_\perp, k_{||})$$

$k_{\|\|}$	$z = 2.2$	$z = 2.6$	$z = 3.0$	$z = 3.6$	$z = 4.0$
0.15	0.27(2)	0.43(2)	0.58(4)	1.05(10)	1.47(22)
0.20	0.26(1)	0.38(2)	0.58(3)	0.86(6)	1.08(13)
0.30	0.18(1)	0.30(1)	0.44(2)	0.81(5)	0.85(10)
0.50	0.15(1)	0.24(1)	0.35(1)	0.59(3)	0.81(07)

Covariance

- Can also calculate Cov[P,P]
  - Assume Gaussian (large scales).
  - Off-diagonal terms are small, even when shell averaged, if omit self-pairs.
  - Important to cap total information content.
  - Quasar clustering is a small correction.

$$\mathrm{cov}[\widehat{P}_{\mathrm{F}}(k), \widehat{P}_{\mathrm{F}}(k')]_{k_{\parallel}} = 2P_{\mathrm{tot}}^2 \delta_k^{k'} + \tfrac{4}{N}\, \overline{\tilde{w}^2}\, P_{\mathrm{F}}(k) P_{\mathrm{F}}(k')$$
$$+ \tfrac{1}{N}\, \overline{\tilde{w}^2}^2\, \bar{n}^{-1} \int \tfrac{d^2 k_{\perp}^1}{(2\pi)^2}\, P_{\mathrm{F}}(k_{\mathrm{ll}}, k_{\perp}^1) P_{\mathrm{F}}(k_{\mathrm{ll}}, k_{\perp}^1 - k_{\perp} - k'_{\perp})$$
$$+ \tfrac{1}{N}\, \overline{\tilde{w}^2}^2\, \bar{n}^{-1} \int \tfrac{d^2 k_{\perp}^1}{(2\pi)^2}\, P_{\mathrm{F}}(k_{\mathrm{ll}}, k_{\perp}^1) P_{\mathrm{F}}(k_{\mathrm{ll}}, k_{\perp}^1 + k_{\perp} - k'_{\perp})$$

Optimal weights

- Can choose weights, *w*, to minimize error.
- Find $w(k_{||}) = B/[P_N+P_{los}]$
  - Remember P_{los} is almost constant at low $k_{||}$.

$$\mathrm{Var}\left[P_{\rm obs}\right] = 2P_{\rm tot}^2 \qquad \text{with} \qquad P_{\rm tot} = P_F + \bar{n}_{\rm eff}^{-1} P_{\rm los}$$

$$\bar{n}_{\rm eff} \equiv \frac{1}{\mathcal{A}} \sum_{n=1}^{N} \nu_n \quad , \quad \nu_n \equiv \frac{P_{\rm los}}{P_{\rm los} + P_{N,n}}$$

One number!

Weight per quasar/sightline

Analytic model

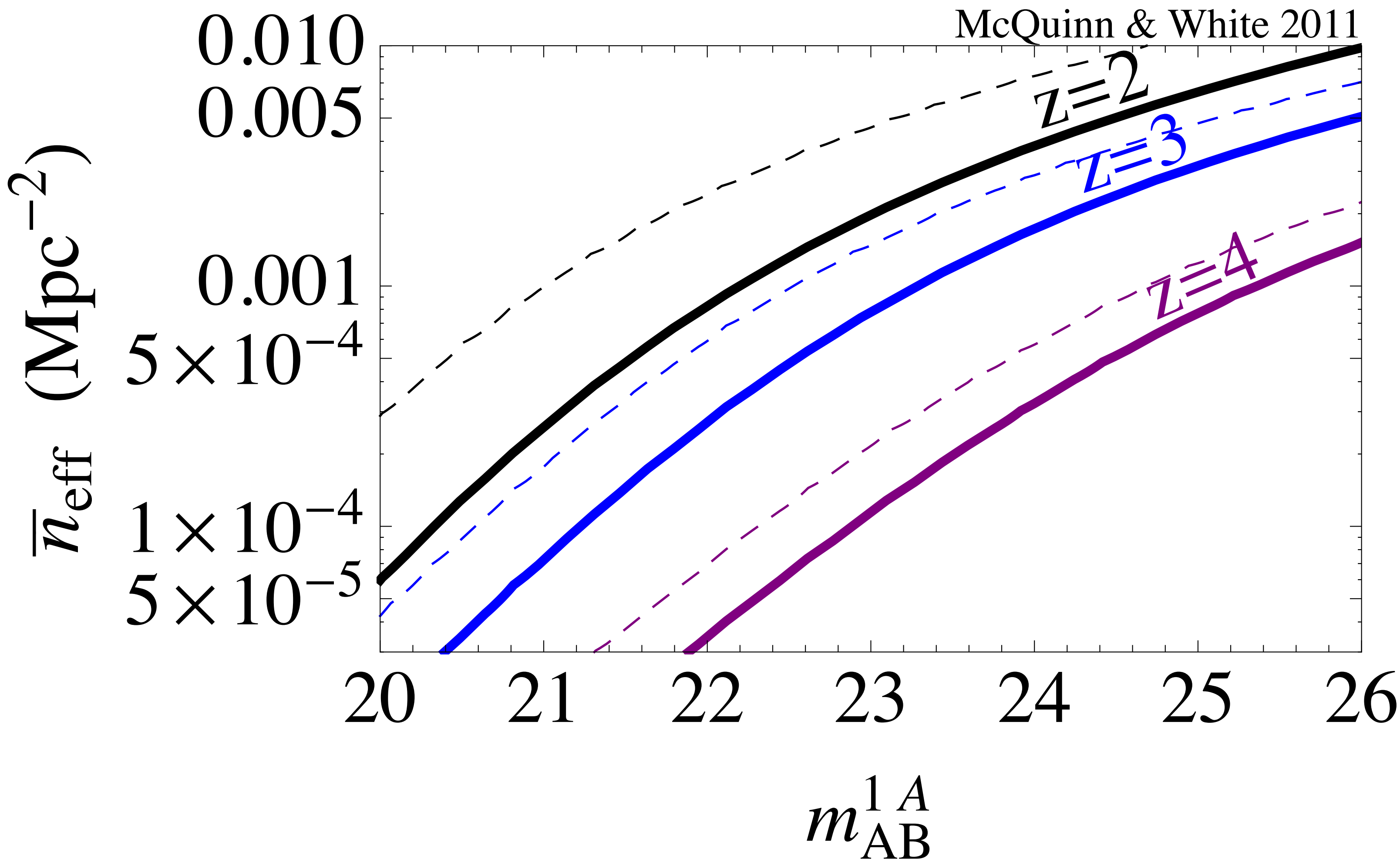
- Suppose $N(f)=(N_0/f_0)\ (f/f_0)^{-\alpha}$
  - $N(>f_{min})=N_0/(\alpha-1)\ \ (f_{min}/f_0)^{1-\alpha}$

- Choose f_0 s.t. $P_N=P_{los}$ at f_0, scaling as f^{-2}
- Define n_0 as N_0/Area.

- For $\alpha=2$: $n_{eff}=n_0 \tan^{-1}(f_0/f_{min})$
- For $\alpha=3$: $n_{eff}=½n_0 \log(1+[f_0/f_{min}]^2)$

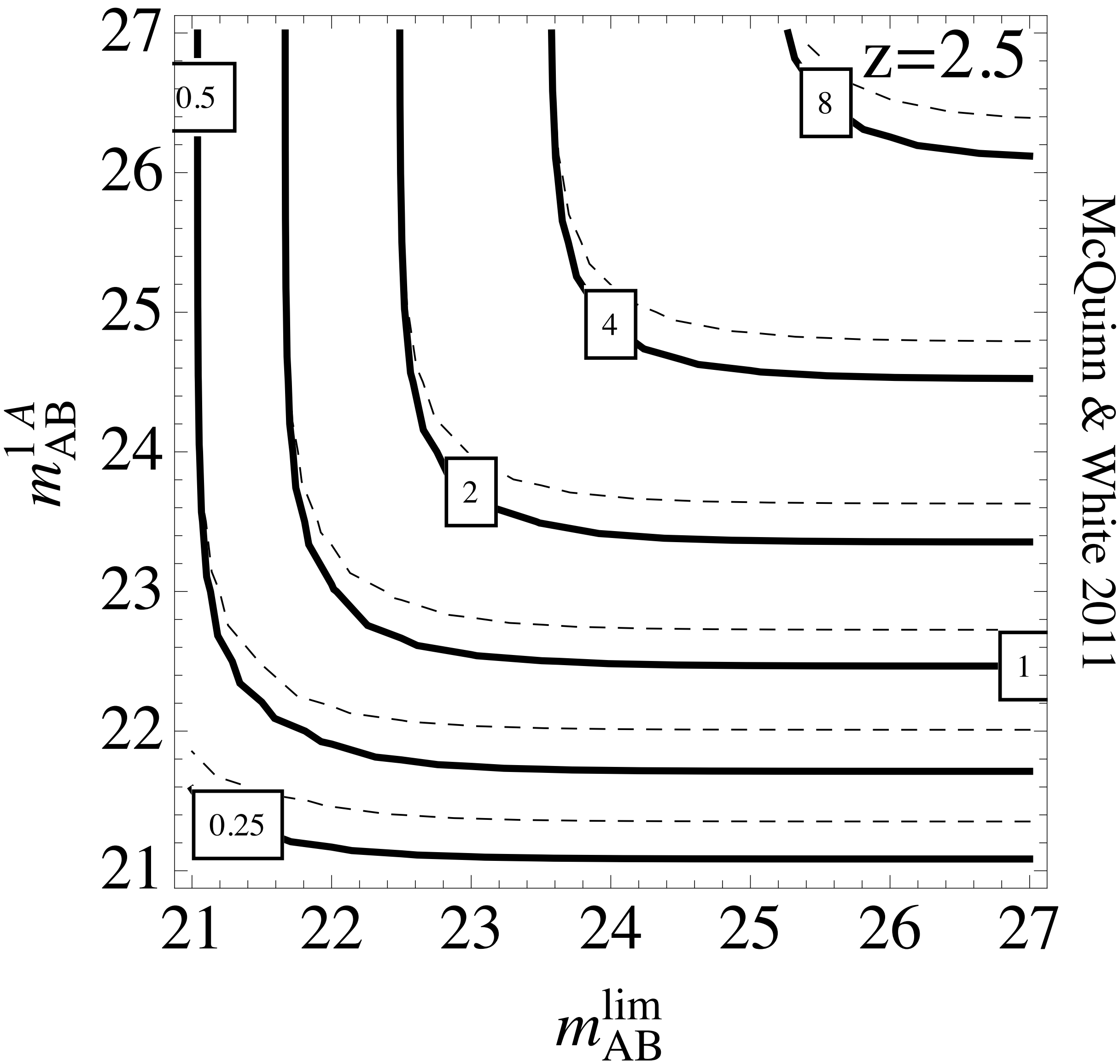
- Saturates at $f_{min}\sim f_0$.

The "effective" number density

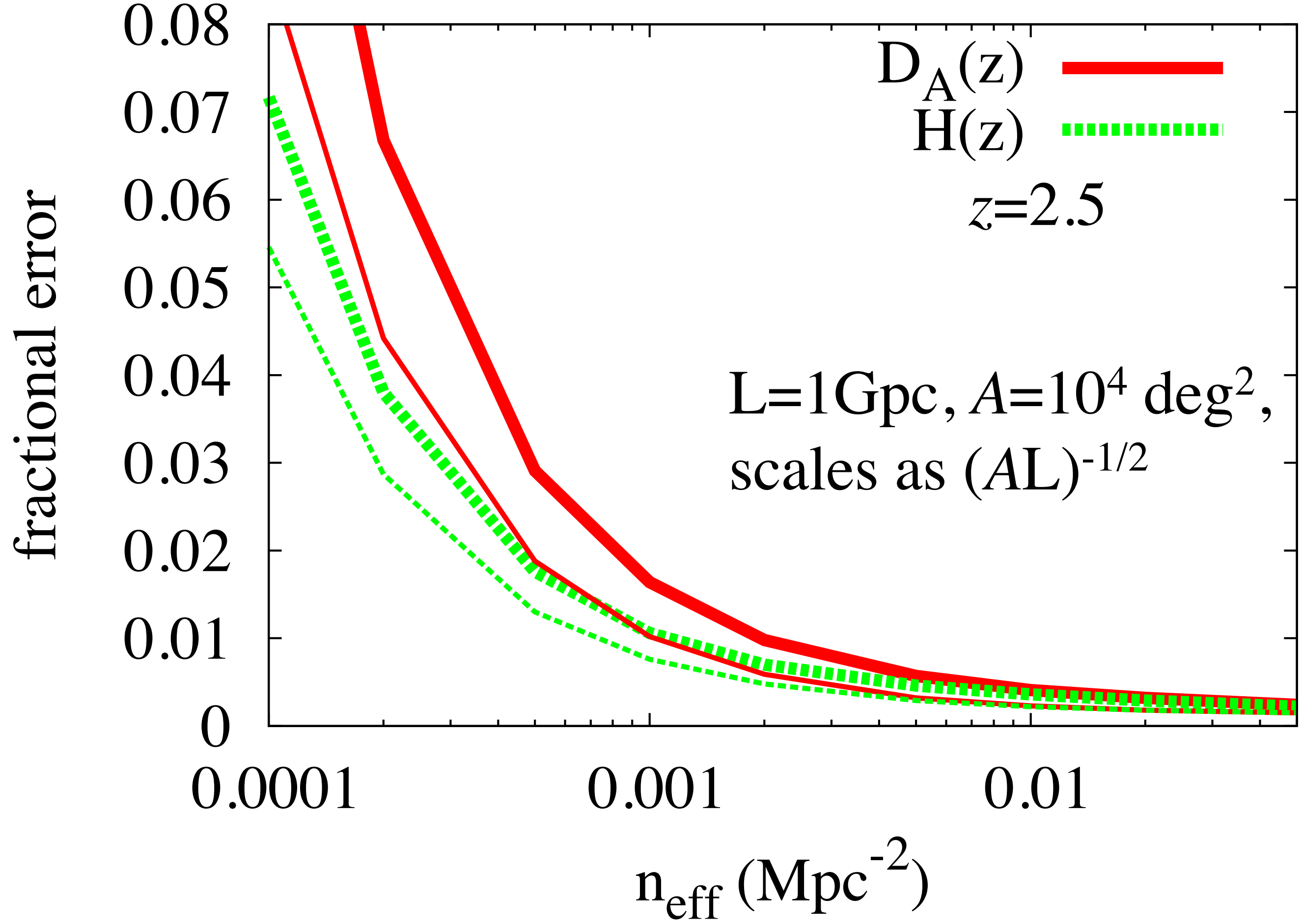


B-band, AB magnitude at which S/N=1 per Å.

Faint quasars don't help much

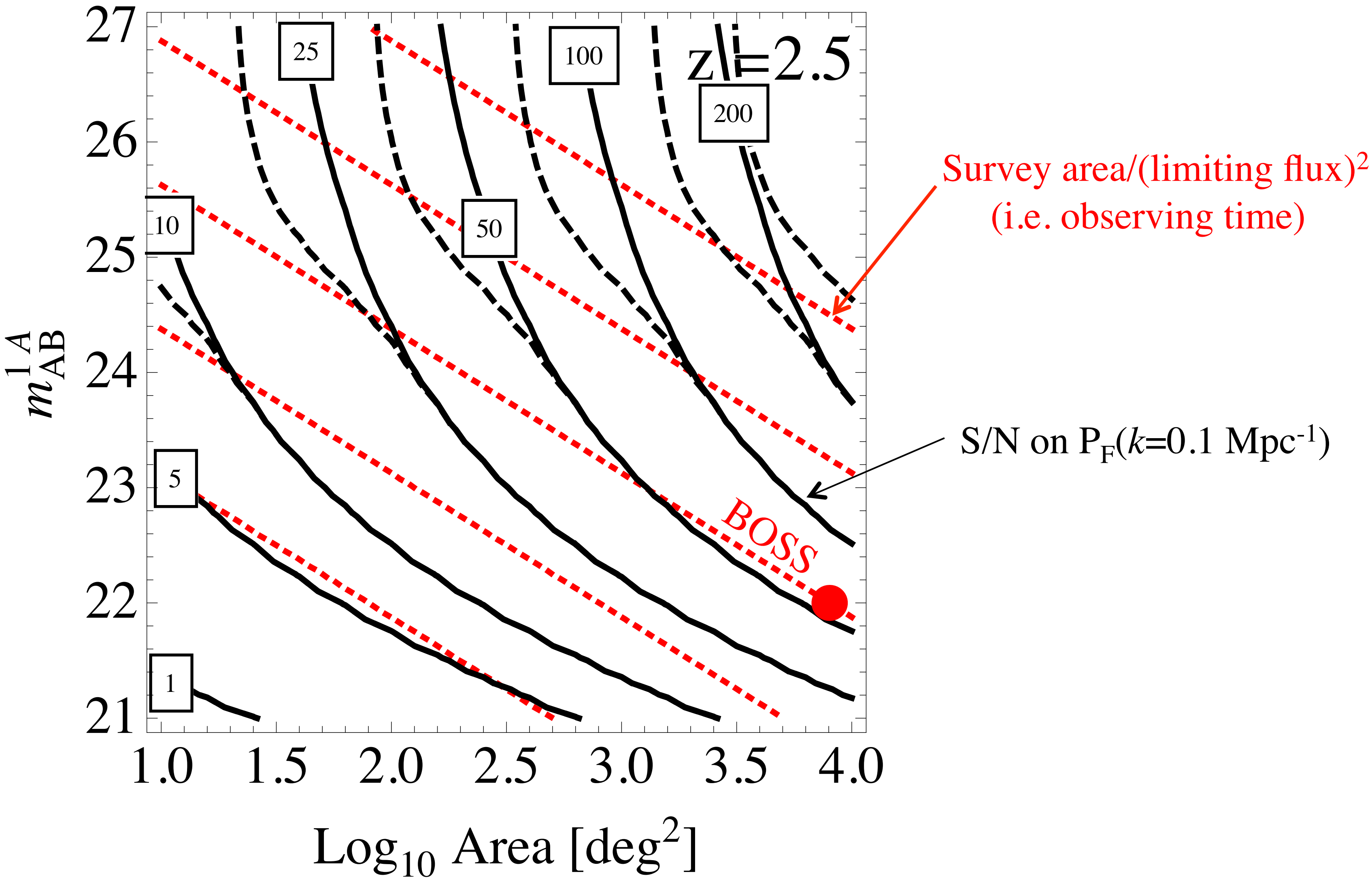


Parameter forecasts



Can constrain k^{-1} to 10%, k^{-2} to 3% and k^{0} to 0.03%

Constraining power



Comparison with FKP

- Have an "effective volume" for a quasar survey, just like FKP derived for galaxies.
  - Feldman, Kaiser & Peacock (1994; ApJ 426, 23)
- Derivation is similar, but for Lyα shot-noise is in plane of sky and is modulated by line-of-sight power.

$$V_{\mathrm{eff,gal}} = V_{\mathrm{gal}} \left(\frac{P(k)}{P(k) + \bar{n}_{3D}^{-1}} \right)^2$$

$$V_{\mathrm{eff,Ly}\alpha} = \mathcal{A}L \left(\frac{P_F(k)}{P_F(k) + \bar{n}_{\mathrm{eff}}^{-1} P_{\mathrm{los}}} \right)^2$$

Quadratic estimator

- Can show that these minimum variance weights are the lowest order term in a series approximation to the optimal quadratic estimator.
  - Estimator weights all pairs.
  - $P \sim \delta^T C^{-1} X C^{-1} \delta$ + bias
  - Iteratively invert C.
- For reasonable quasar densities the lowest order term is almost as good as the full estimator.
- Next order term suppresses contribution from quasars overabundant within $r_{perp} \leq k_{perp}^{-1}$

Configuration space

- Some advantages to working in configuration space and computing ξ_F by pair counts.
- In configuration space our weights, *w(k)*, become convolutions along the line-of-sight.
- Since *w(k)* is so flat, convolution is "small".
- Weights can be well approximated by a single number, w_n, per sightline.

- Optimal weighting: $w_n \sim (P_{los}+P_{N,n})^{-1} \sim (1+\sigma^2_N/\sigma^2_{los})^{-1}$
  – Variance smoothed on ~10Å
- The same weights work for cross-correlation with e.g. galaxy density field.

Advantages of 3D analysis

- There are several other advantages to 3D (c.f. 1D) analyses.
  - Continuum fluctuations less important.
    - Doesn't bias power in cross-spectra.
    - Increase in noise is small at low *k*.
  - Mean flux evolution only affects μ=1.
    - Marginalize low k_{perp} modes.
    - Power bleeding from *W(k)* not too important if reasonable guess for F(z) is known.
  - Damping power less important.
    - A lot of the los power from DLA systems is shot-noise.
    - Excluding self-pairs drastically reduces this power.
    - Power depends only on $k_{||}$

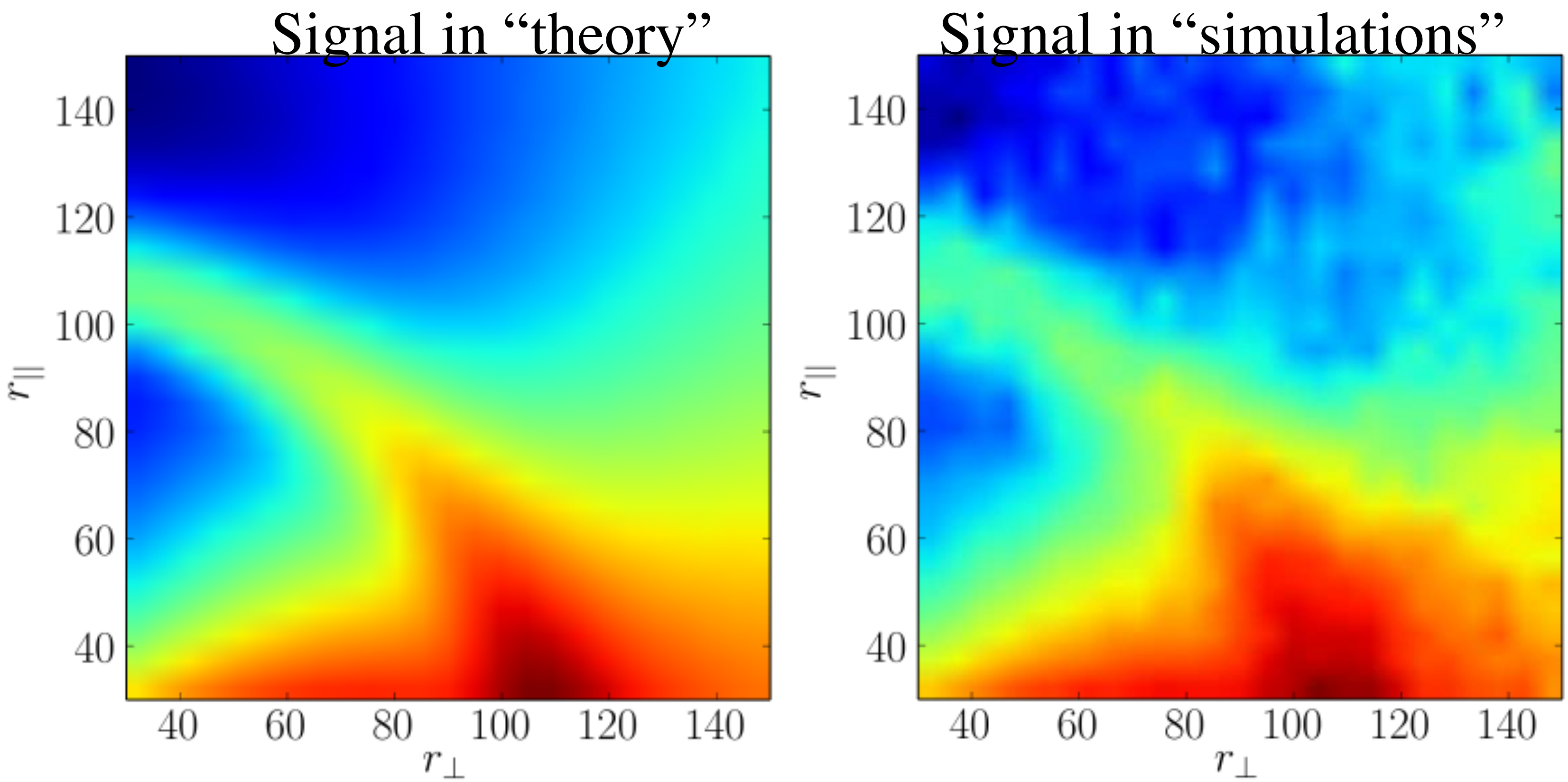
Summary

- The future of Lyα studies is 3D.
- We derived a simple formula for weighting sightlines in 2 -point analyses.
  - $w_n \sim (1+\sigma^2_N/\sigma^2_{los})^{-1}$
  - A good approx. to OQE.
- Survey sensitivity characterized by n_{eff}.
- Optimal strategy: get S/N=2 per Å for an L_* quasar.
- Surveys can provide (very) strong constraints on cosmology and astrophysics.

The End

BAO at high *z*

Slosar, Ho, White & Louis (2009)



BAO feature survives in the LyA flux correlation function, because on large scales flux traces density. Relatively insensitive to astrophysical effects*.