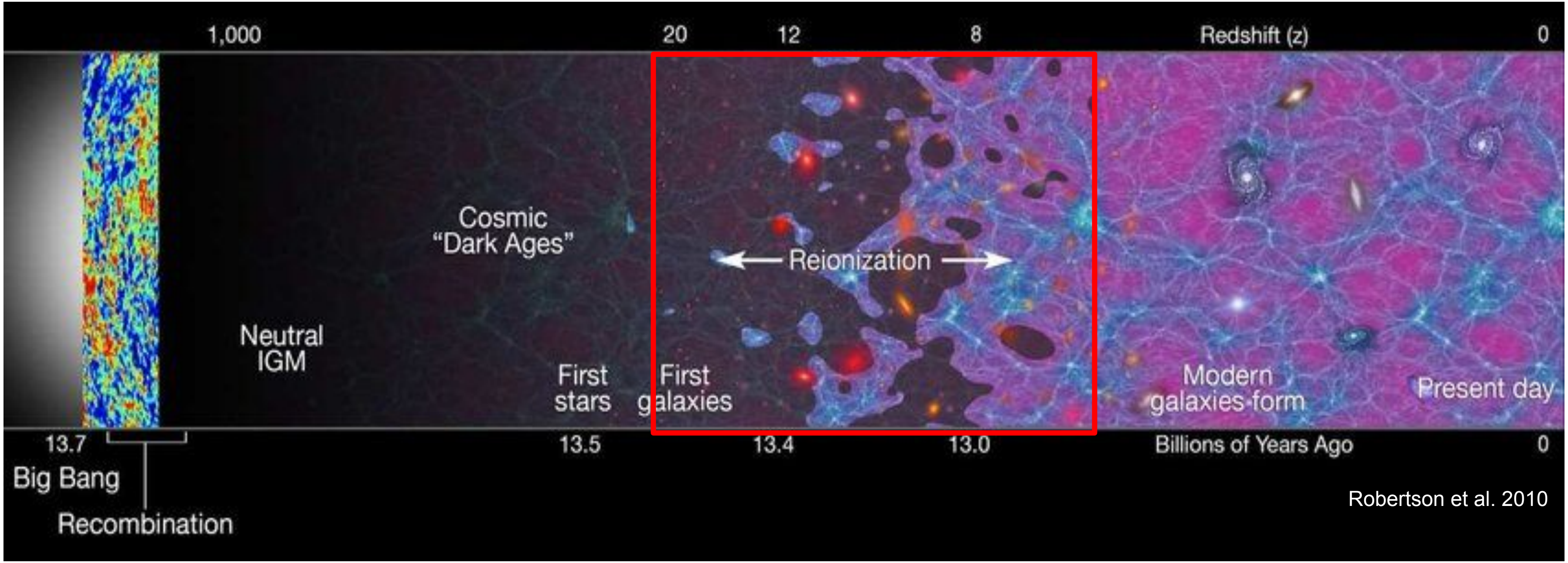
Constraining reionization with the $z > 5$ Lyman-α forest

2309.03341 - measurements
2309.05647, 2208.09013 - forecasting

Molly Wolfson (UCSB)

NASA, ESA and J. Olmsted (STScI)

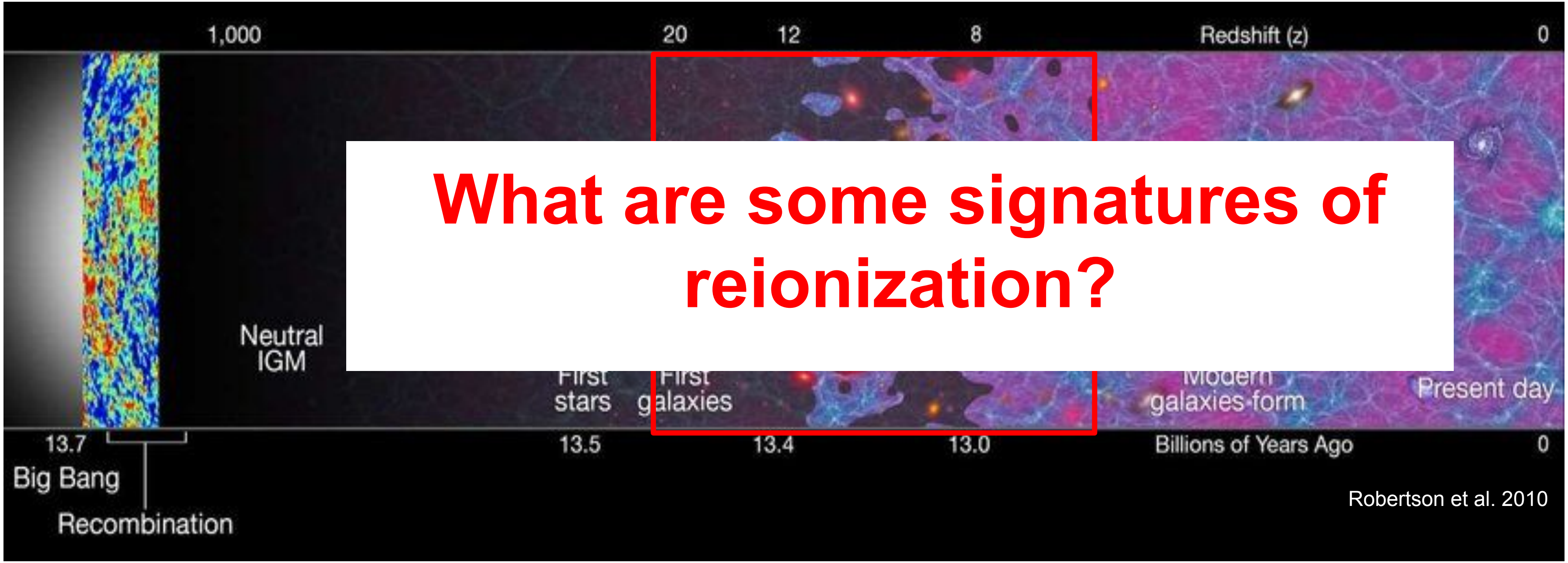
What is reionization?



Planck gives the midpoint: $z_{re} \sim 7.7$

Recent measurements suggest that it is not complete until $z < 6$

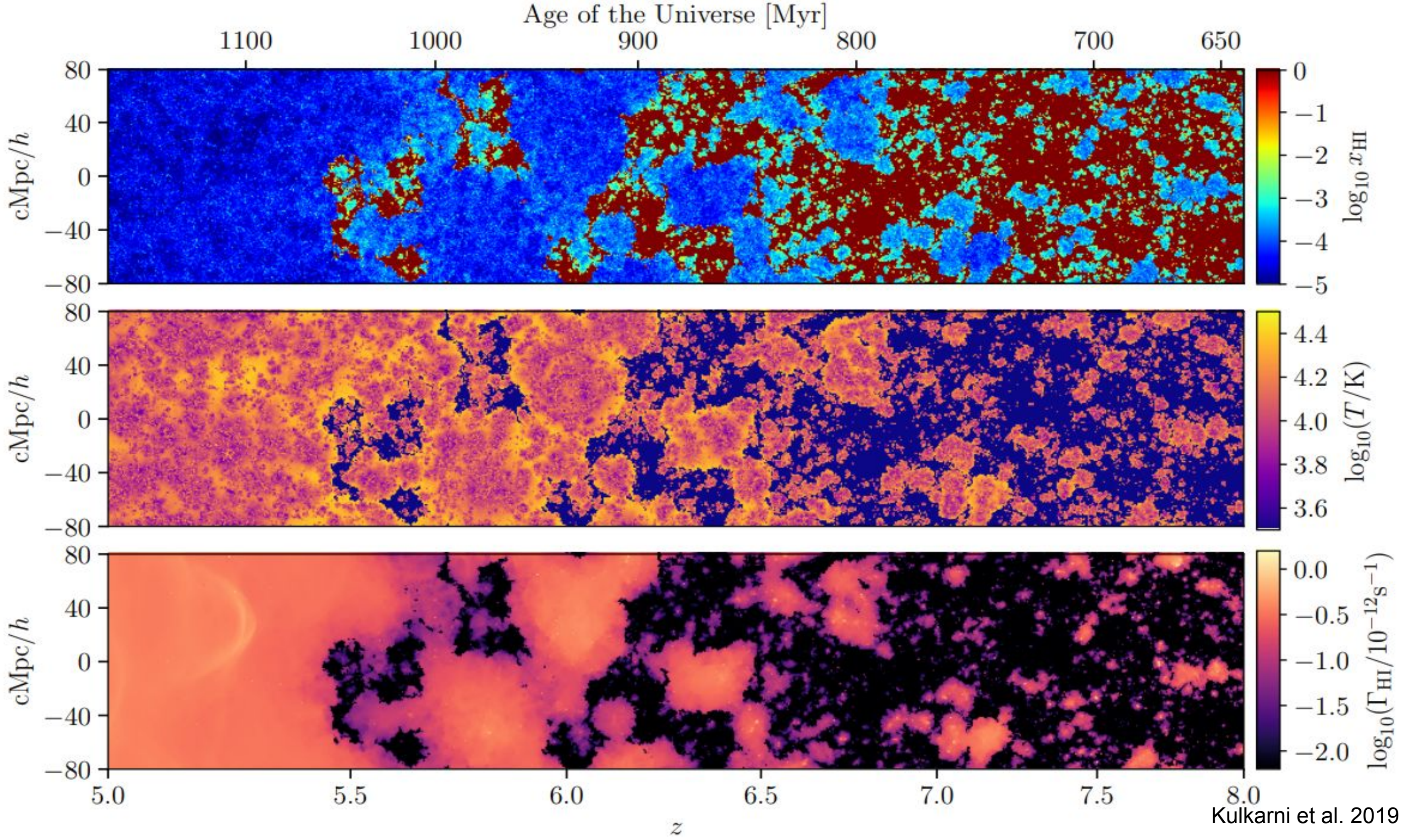
What is reionization?



What are some signatures of reionization?

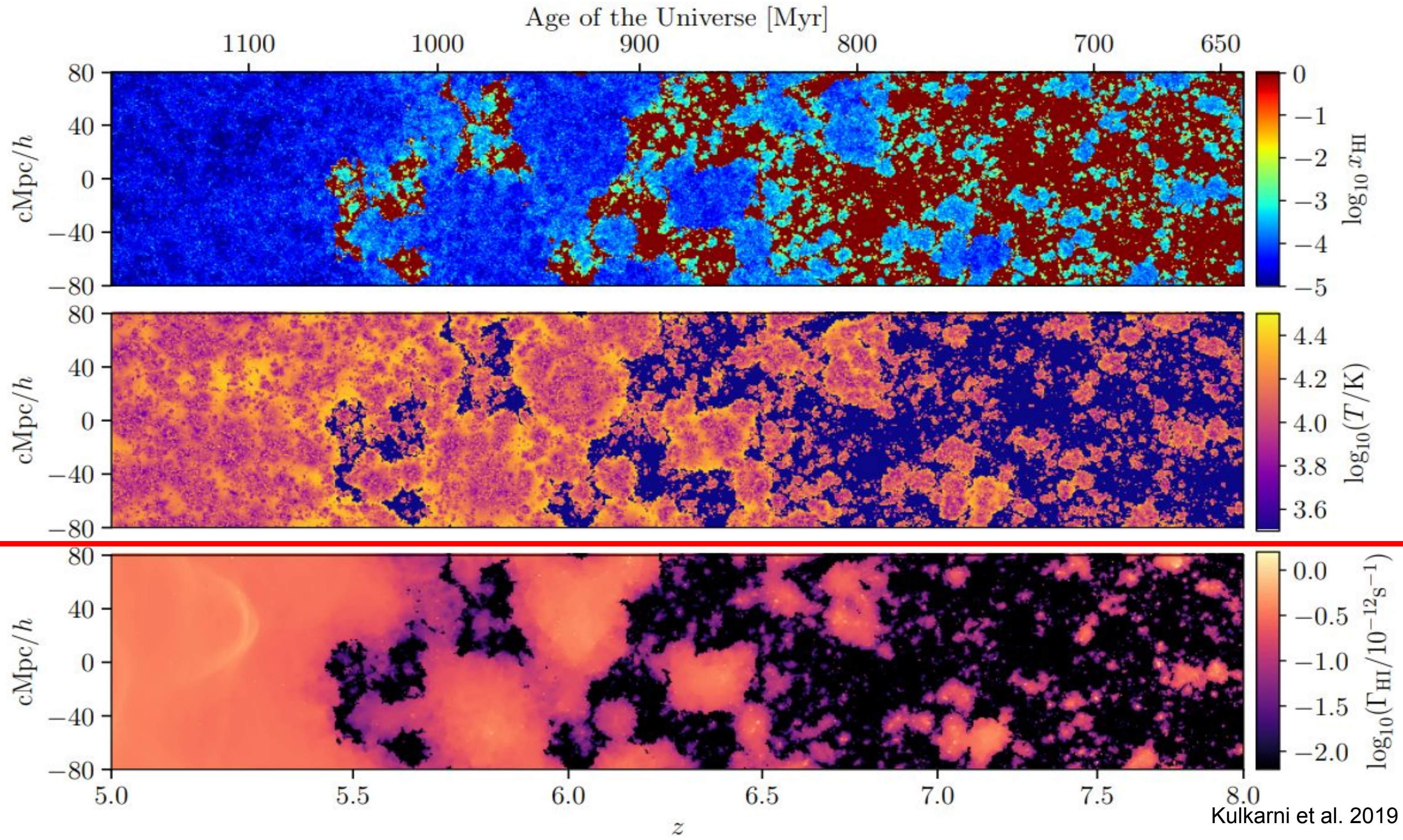
A lot still unknown (driving sources, number of photons required, impact on thermal state of IGM, and more)

Some signatures of reionization in the IGM



Kulkarni et al. 2019

Some signatures of reionization in the IGM



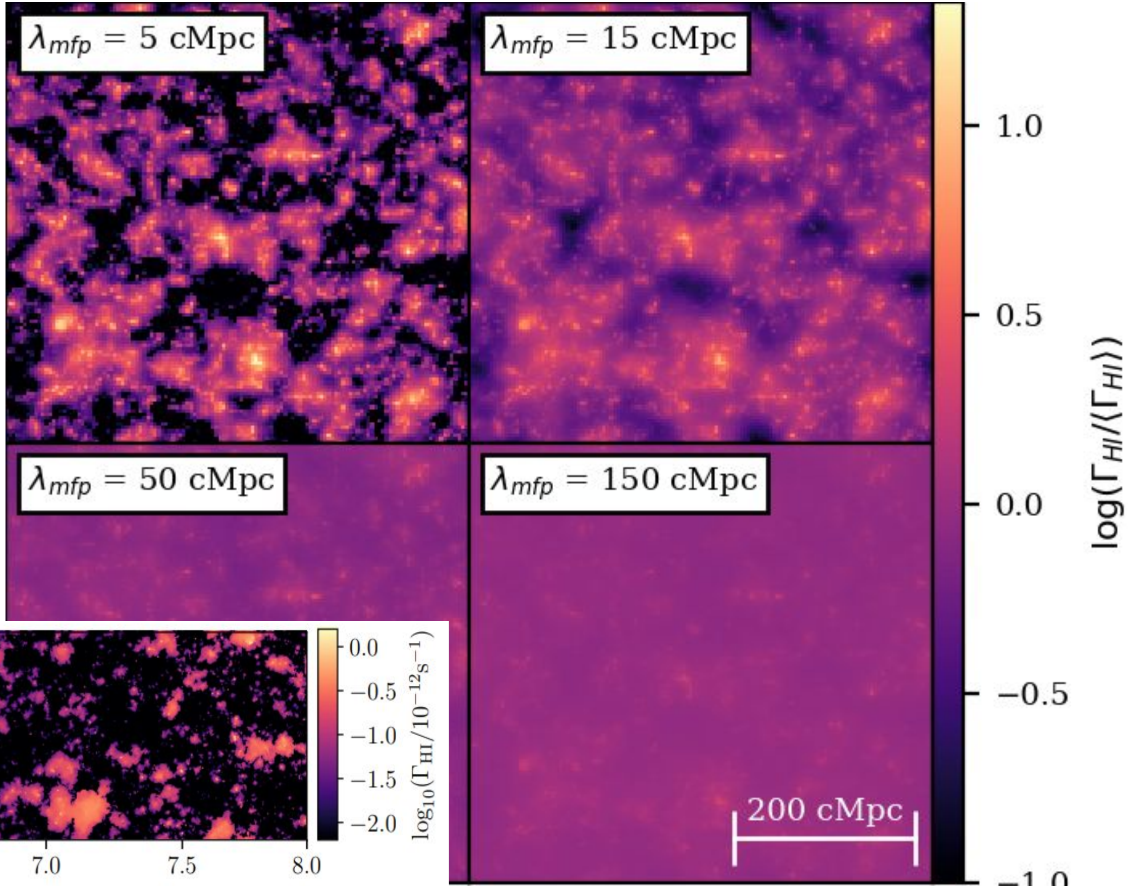
UVB

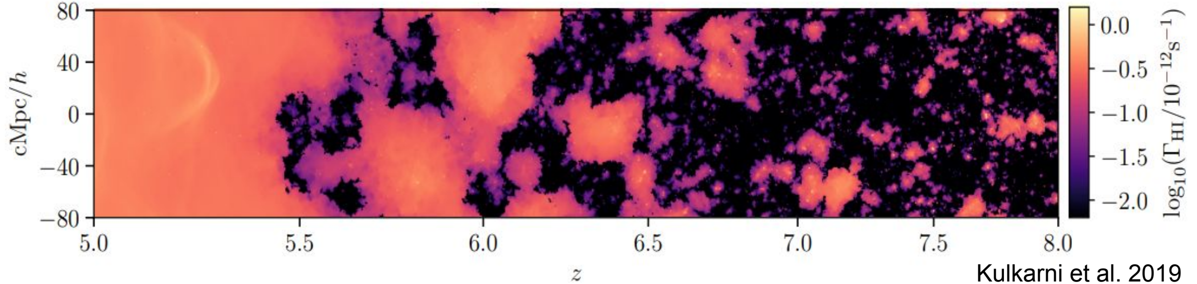
Kulkarni et al. 2019

Fluctuations in the UVB can be described by λ_{mfp}

λ_{mfp} - the average distance ionizing photons travel before interacting with neutral hydrogen

Rapid increase signals the end of reionization



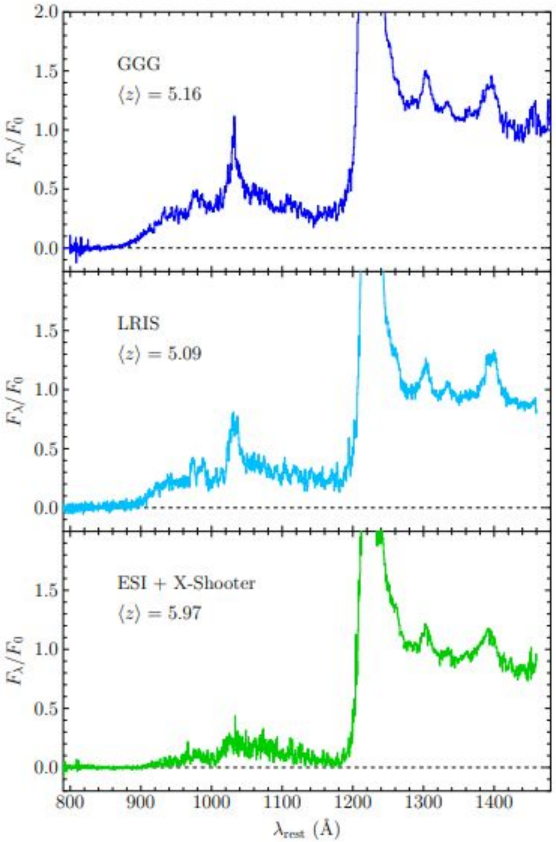


Kulkarni et al. 2019

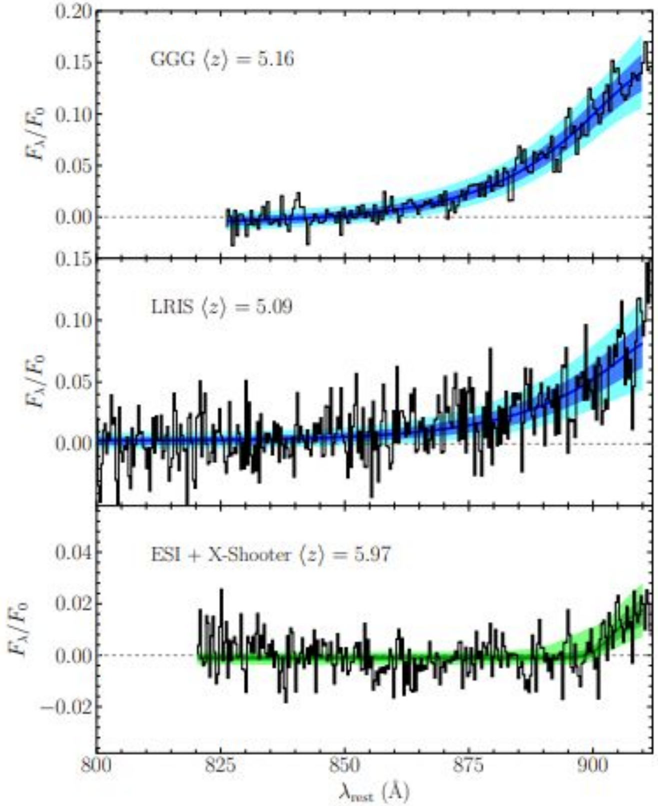
Wolfson et al. 2022

Existing measurements of λ_{mfp}

Most constraining method has been from flux beyond the Lyman limit in stacked quasar spectra



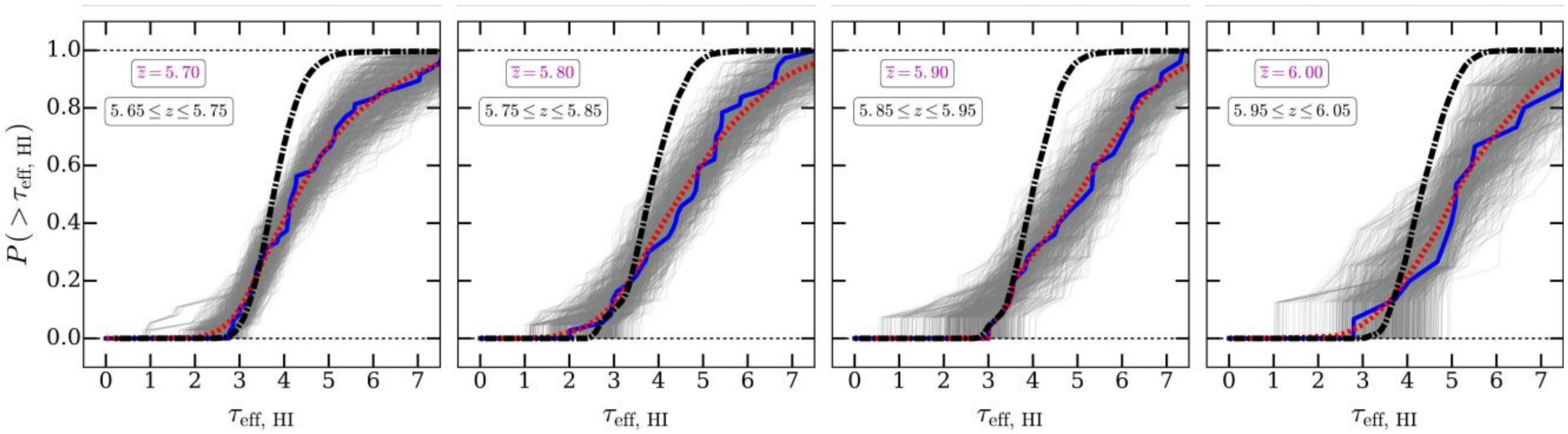
Becker et al. 2021



Becker et al. 2021

Existing measurements of λ_{mfp}

Additional recent measurements using the optical depth CDF

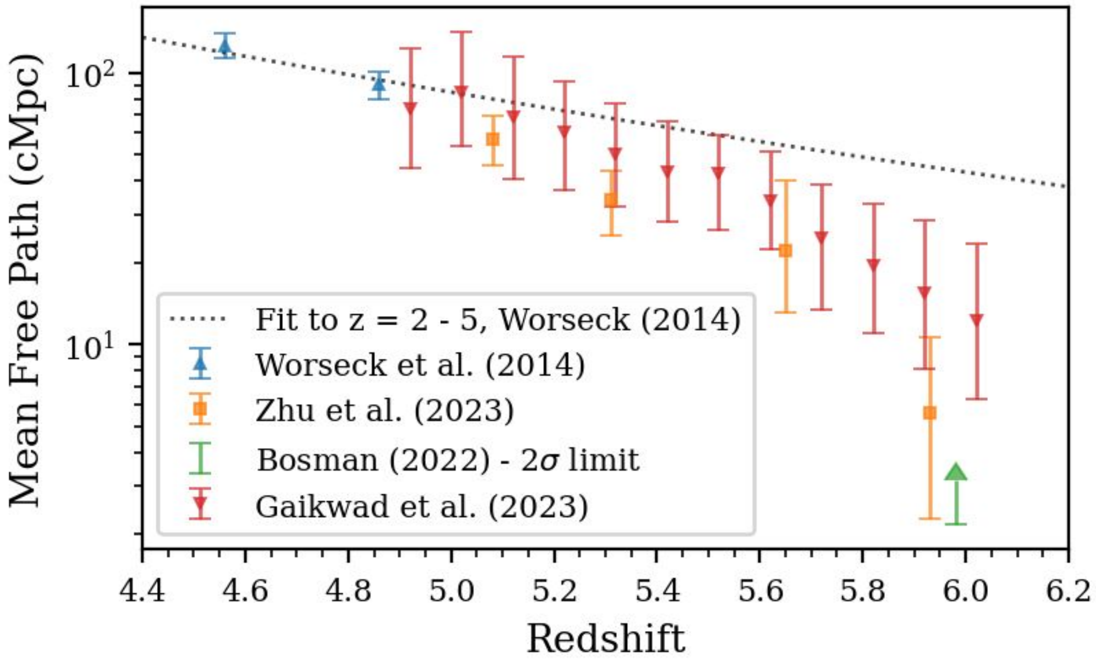


Gaikwad et al. 2023

Existing measurements of λ_{mfp}

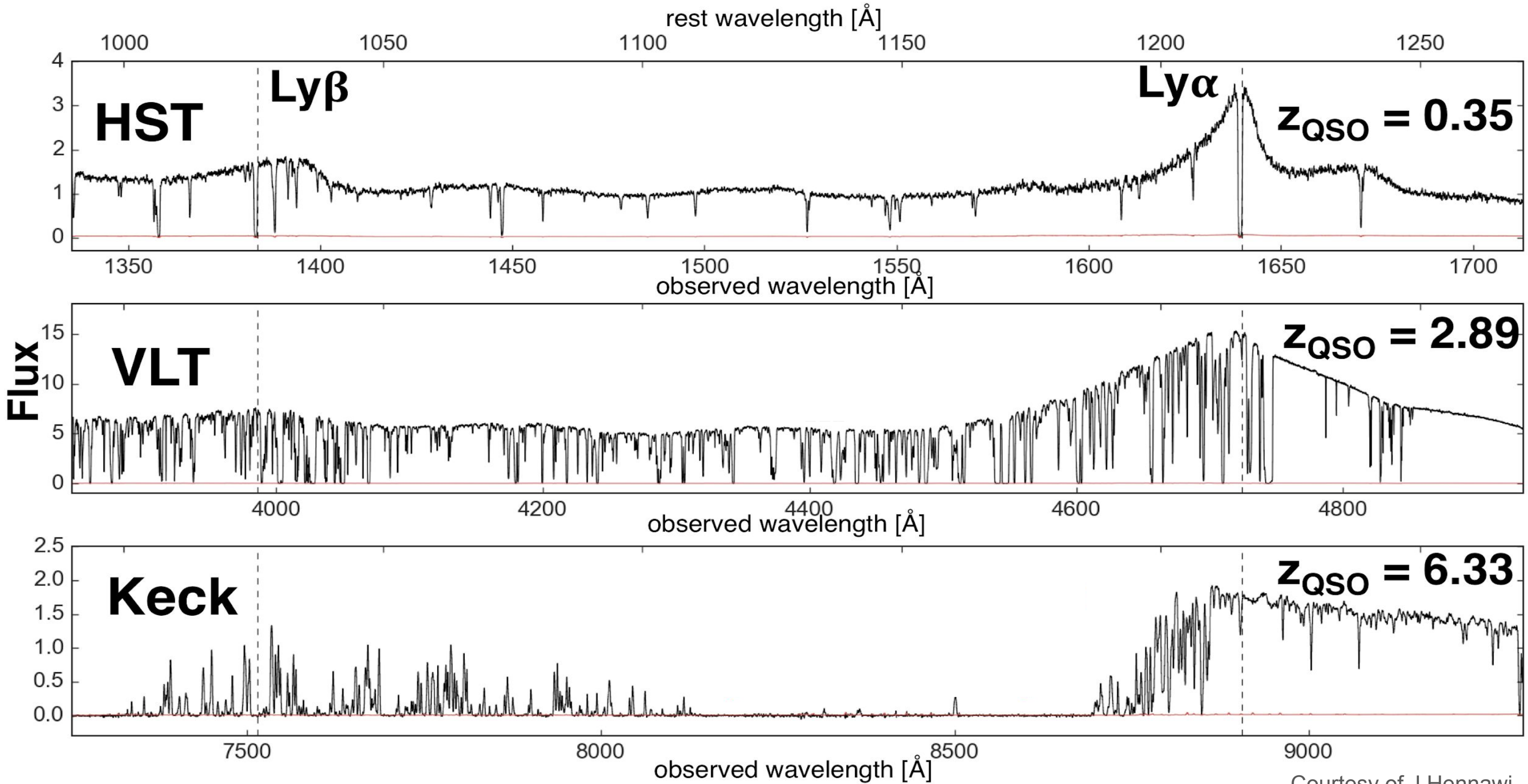
Blue and Orange points are from flux beyond the Lyman limit in stacked quasar spectra

Red points come from the the optical depth CDF



Lyman-α forest flux auto-correlation function

Probing the IGM with the Lyman-α forest:



Courtesy of J.Hennawi

Saturates for $\langle x_{\rm HI}(z)\rangle \gtrsim 10^{-4}$

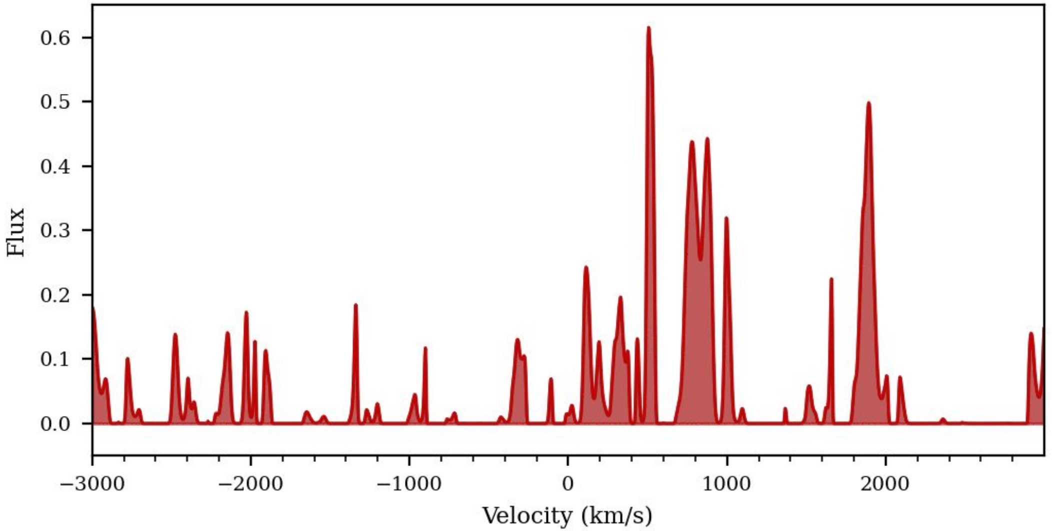
Auto-correlation function:

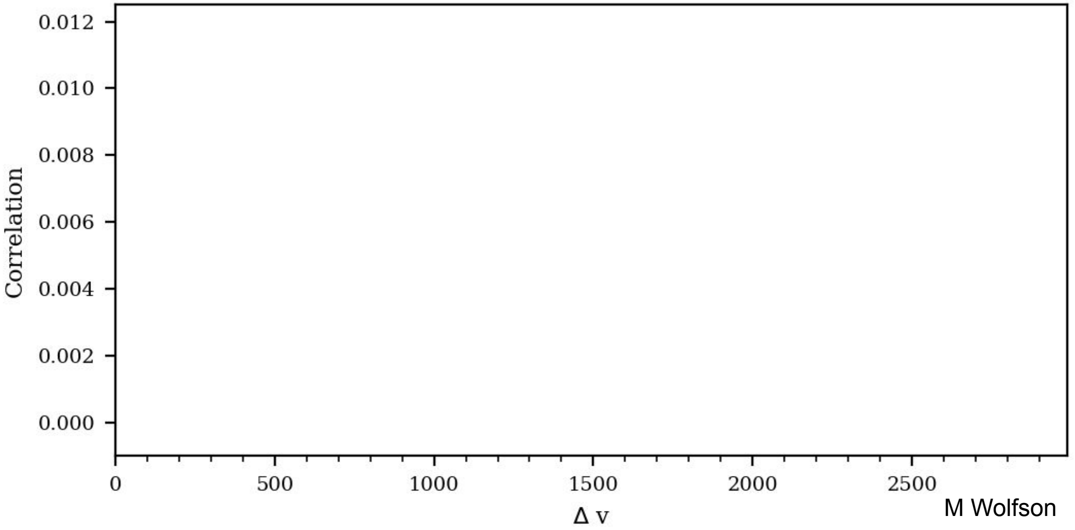
$$\xi_F(\Delta v) = \langle F(v) F(v + \Delta v) \rangle$$

The fourier transform of the power spectrum

Uncorrelated gaussian noise averages out

Easy to mask out DLAs etc





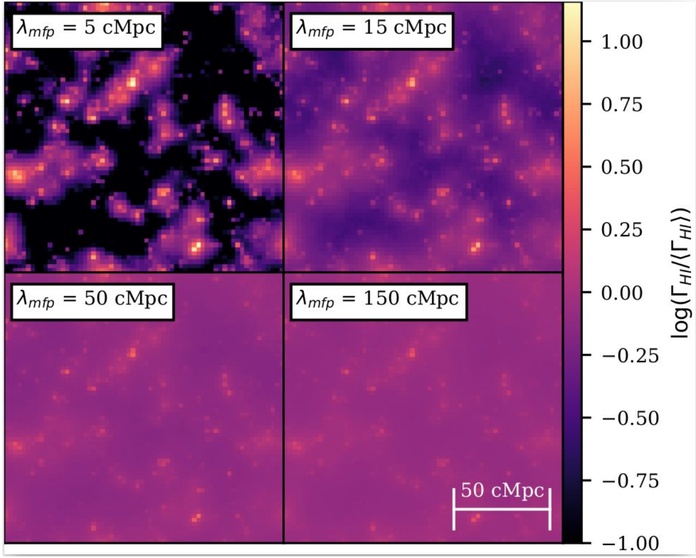
M Wolfson

Is the auto-correlation function of the Lyman-α forest sensitive to λ_{mfp}?

Simulation box properties

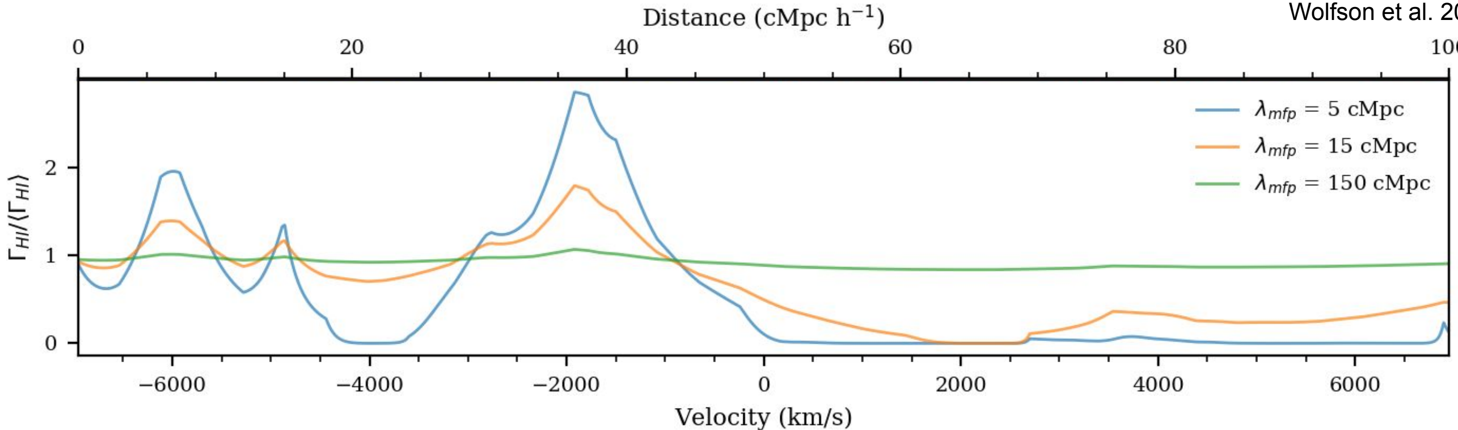
UVB boxes with L_{box} = 100 cMpc h^{-1}

- Method of Davies & Furlanetto 2016
- 128^3 pixels



Wolfson et al. 2023

Wolfson et al. 2023



Simulation box properties

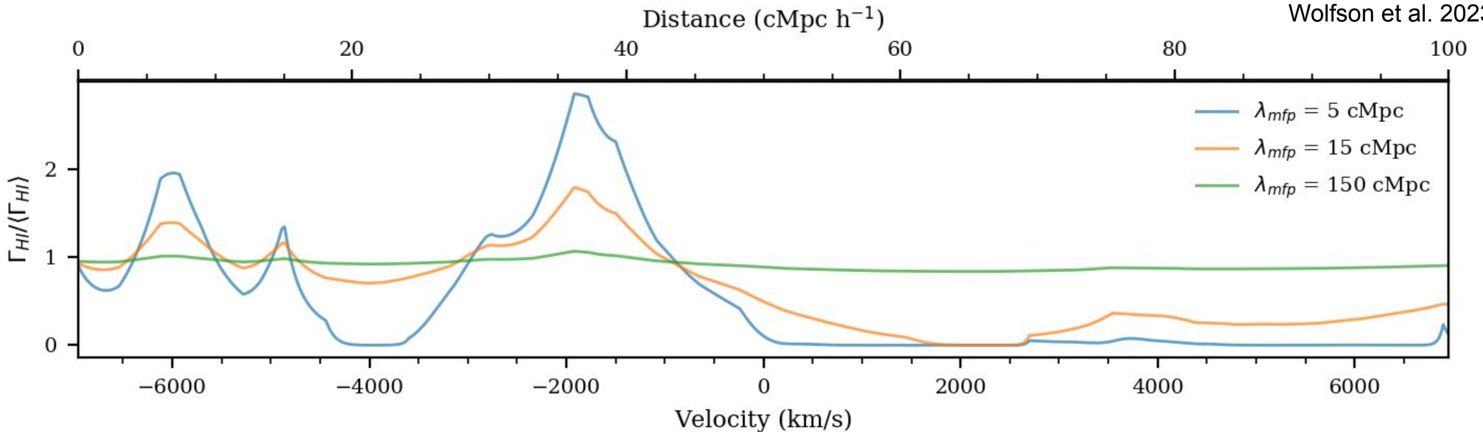
UVB boxes with L_{box} = 100 cMpc h^{-1}

- Method of Davies & Furlanetto 2016
- 128^3 pixels

Nyx box with L_{box} = 100 cMpc h^{-1}

- Hydrodynamical code designed for the Lyα forest
- 4096^3 dark matter particles, 4096^3 baryon grid cells

Wolfson et al. 2023



Simulation box properties

UVB boxes with L_{box} = 100 cMpc h^{-1}

- Method of Davies & Furlanetto 2016
- 128^3 pixels

Nyx box with L_{box} = 100 cMpc h^{-1}

- Hydrodynamical code designed for the Lyα forest
- 4096^3 dark matter particles, 4096^3 baryon grid cells

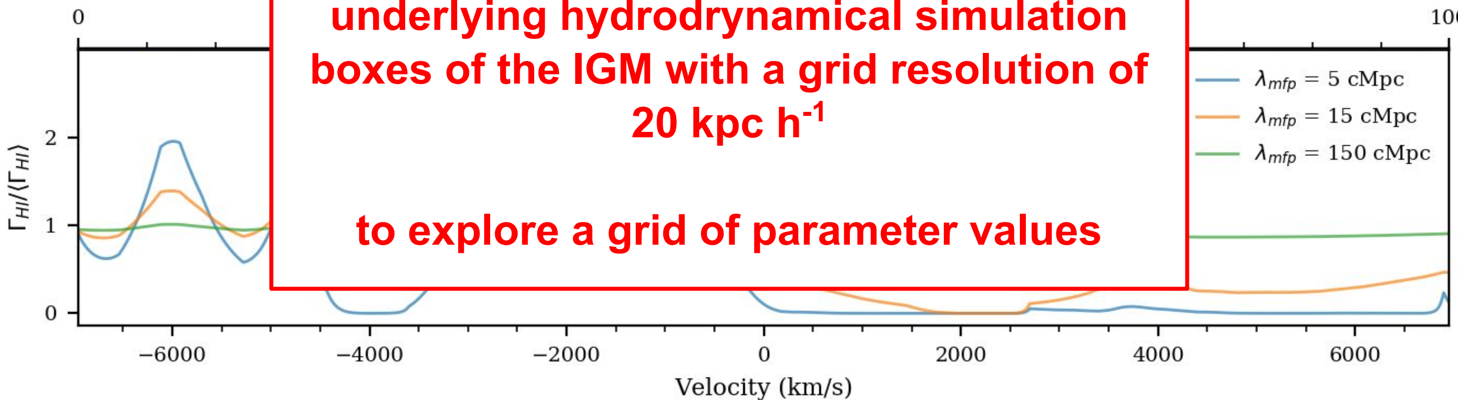
Need…

UVB box with L_{box} of multiple λ_{mfp}

underlying hydrodrynamical simulation boxes of the IGM with a grid resolution of 20 kpc h^{-1}

to explore a grid of parameter values

Wolfson et al. 2023



Simulation box properties

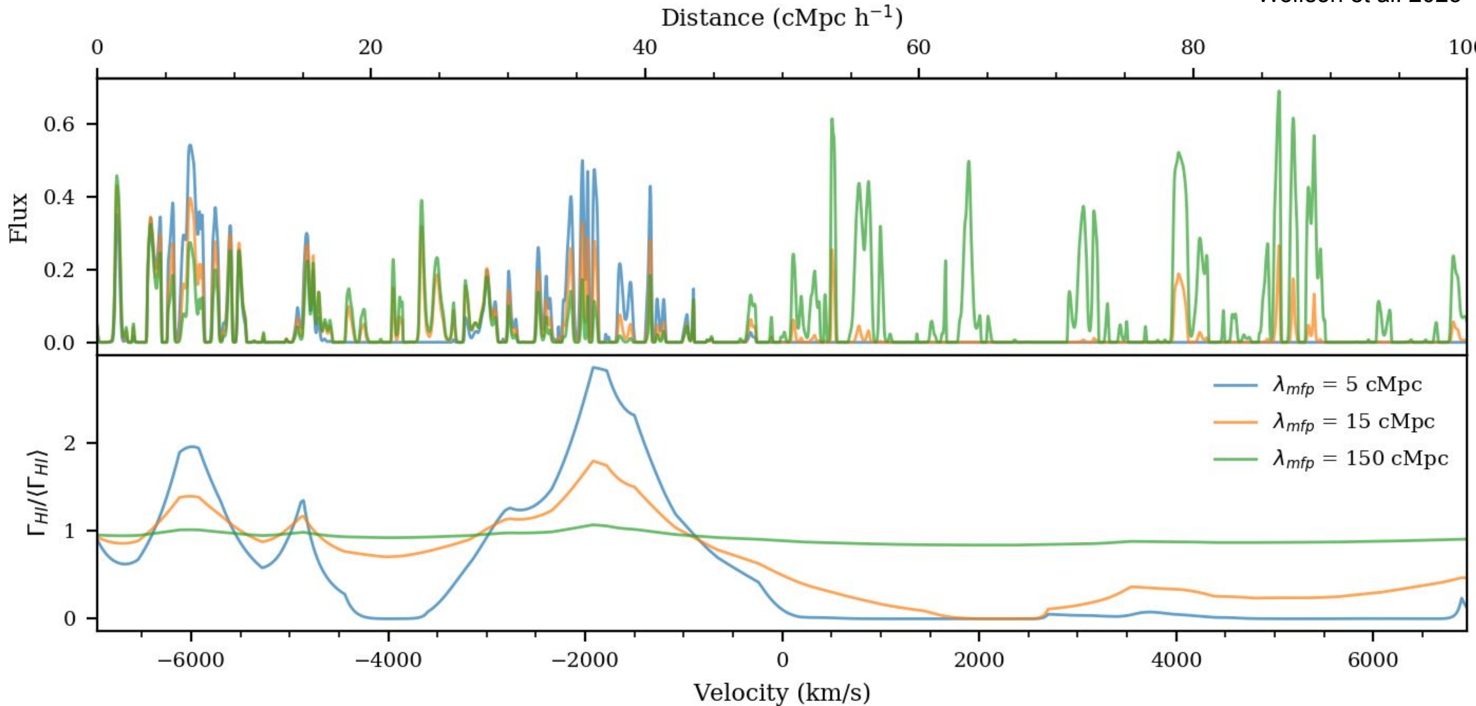
UVB boxes with L_{box} = 100 cMpc h^{-1}

- Method of Davies & Furlanetto 2016
- 128^3 pixels

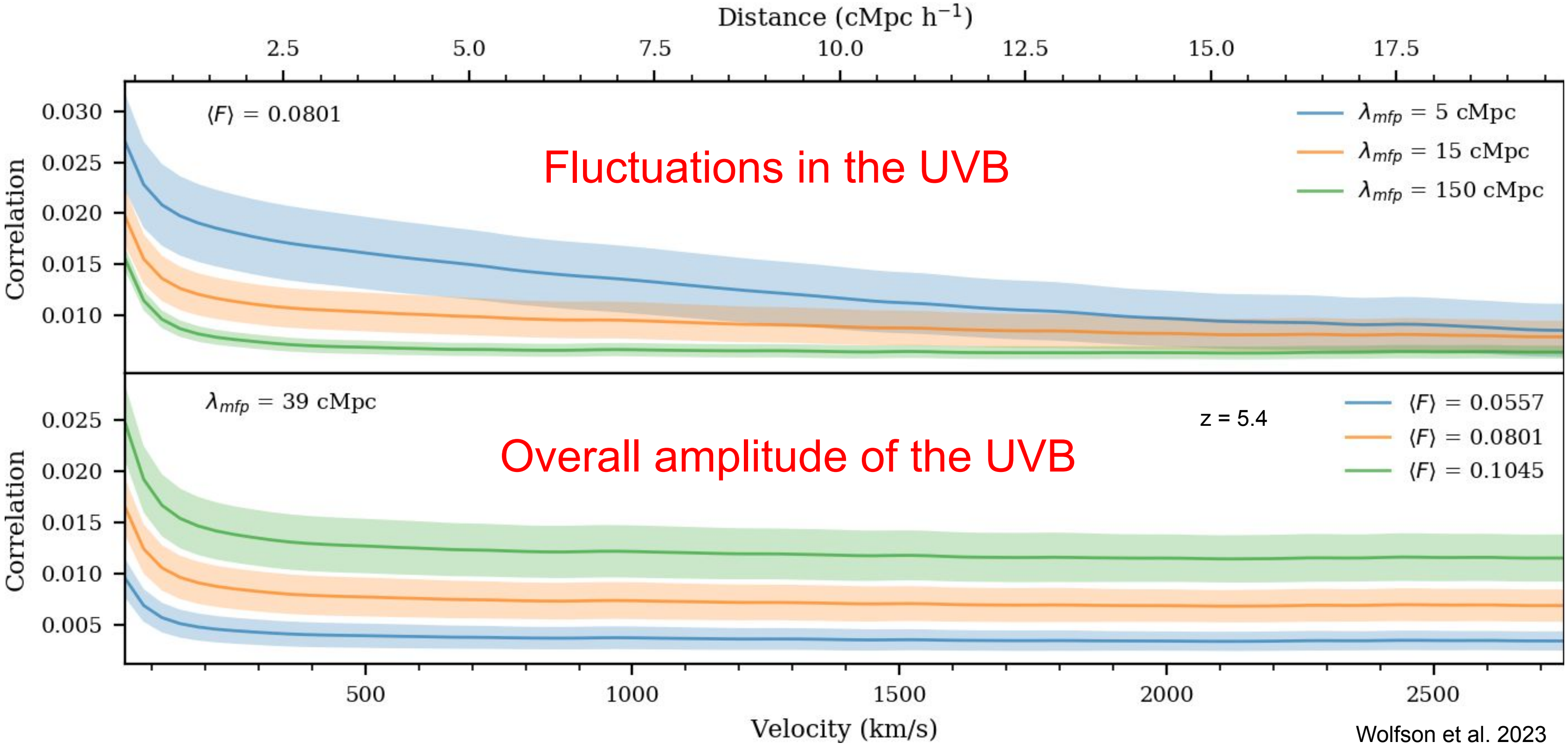
Nyx box with L_{box} = 100 cMpc h^{-1}

- Hydrodynamical code designed for the Lyα forest
- 4096^3 dark matter particles, 4096^3 baryon grid cells

Wolfson et al. 2023



Effect of λ_{mfp} on the auto-correlation function



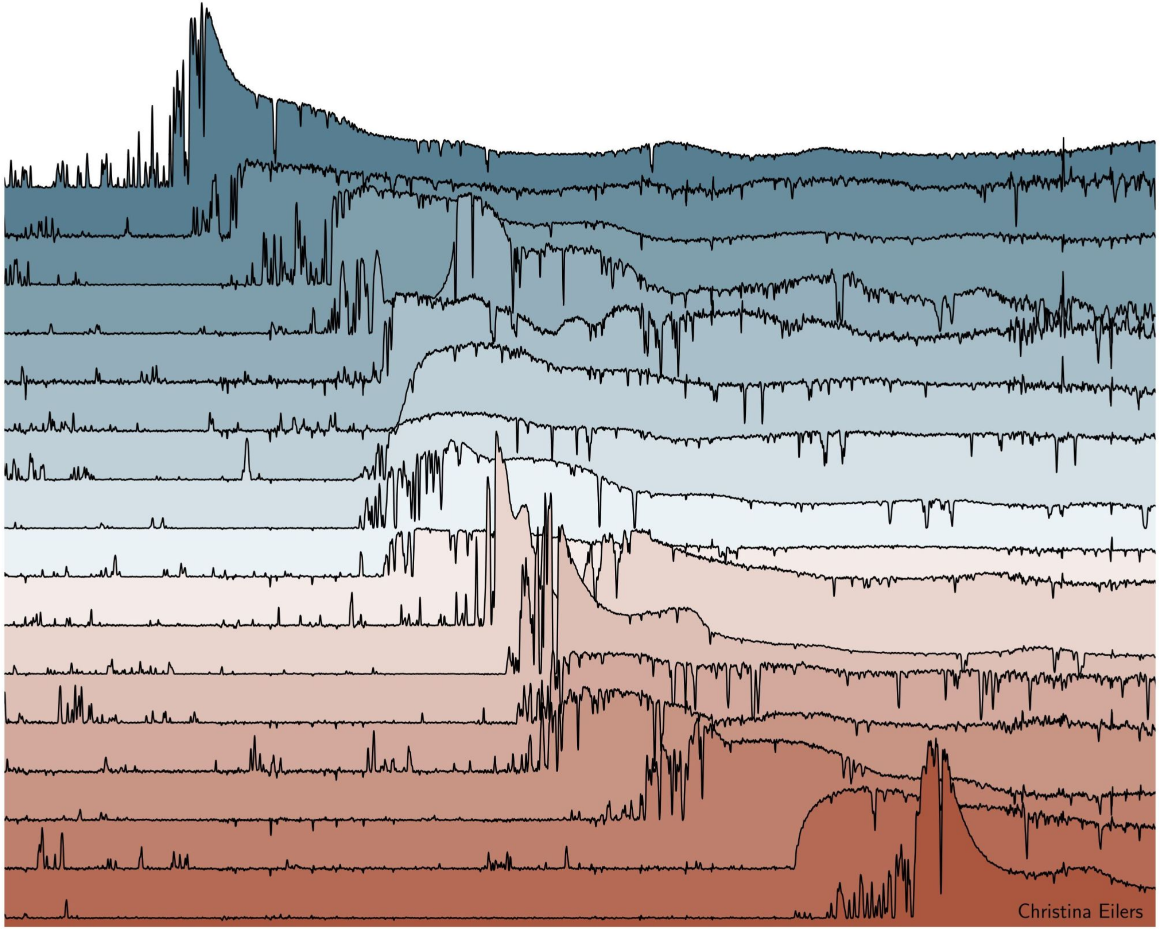
Wolfson et al. 2023

Data

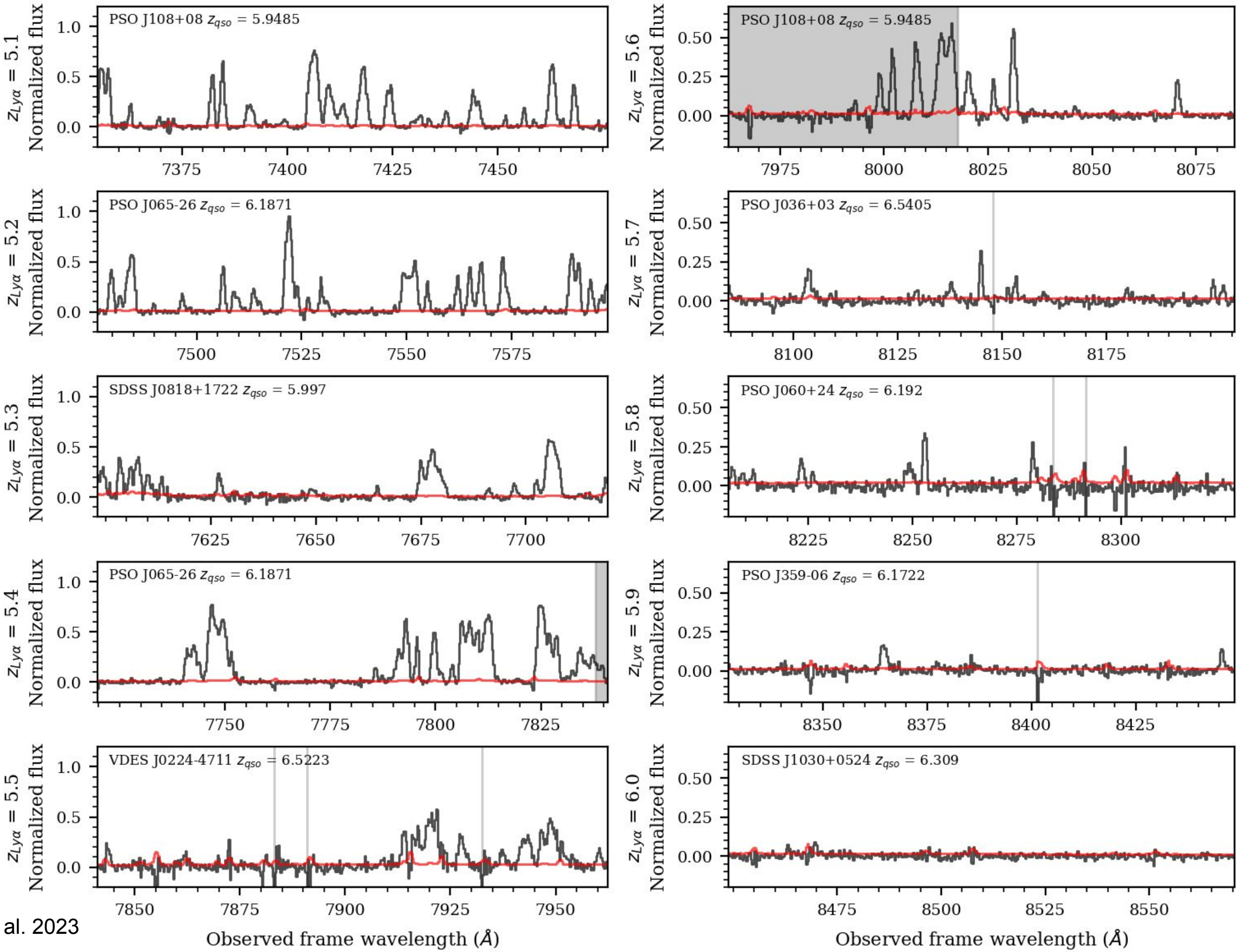
XQR-30 Quasars:

XQR-30 data (xqr30.inaf.it):

- dedicated ~250 hours of observations
- Uses VLT/X-Shooter (R ~ 8800 in the visible)
- 30 new observations of some of the most luminous z > 5.8 quasars observed
- Supplemented with 12 archival observations

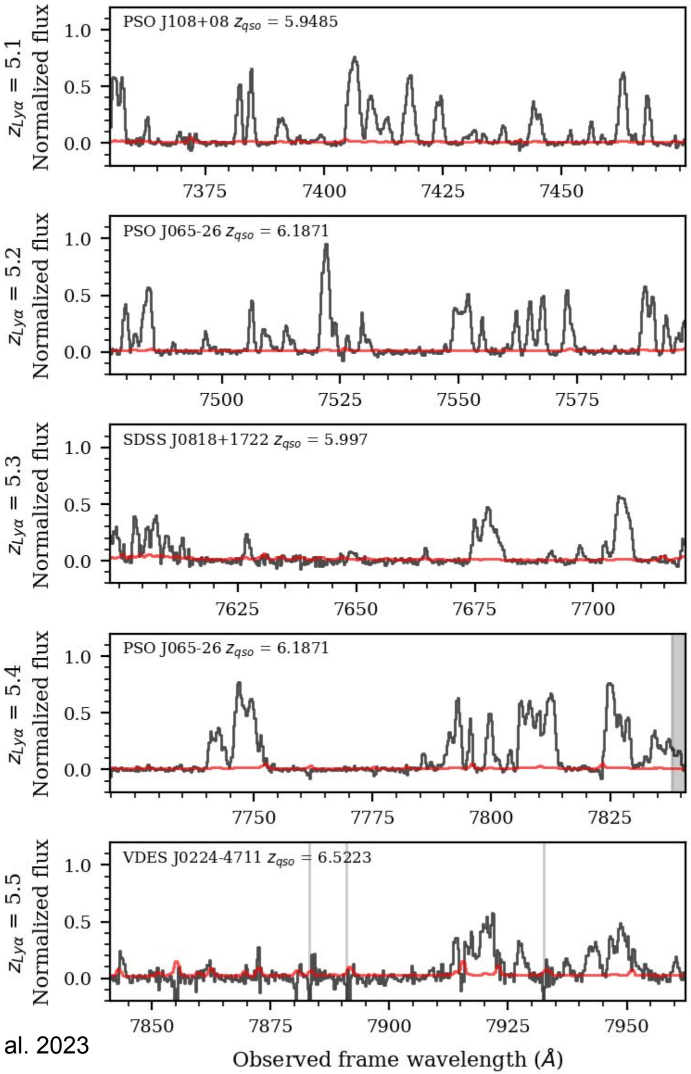


XQR-30 Data:



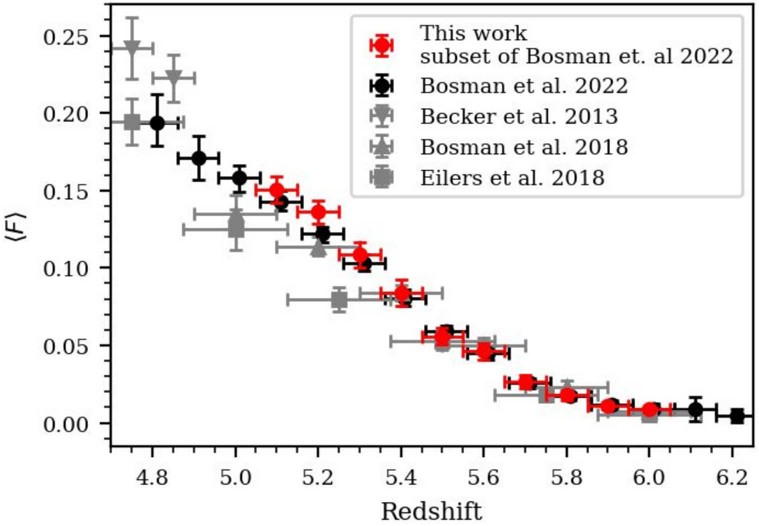
Wolfson et al. 2023

XQR-30 Data:

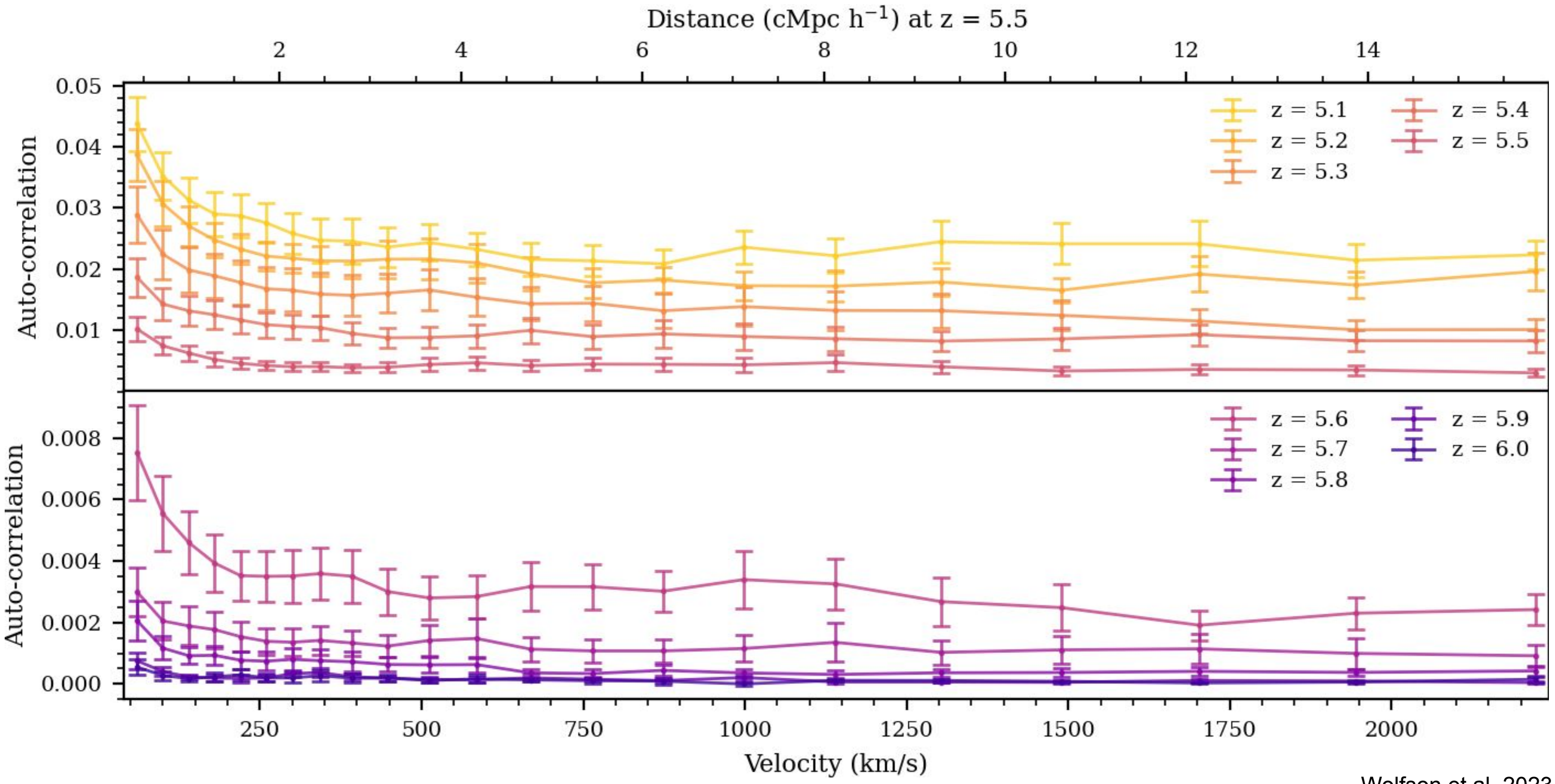


Wolfson et al. 2023

Wolfson et al. 2023

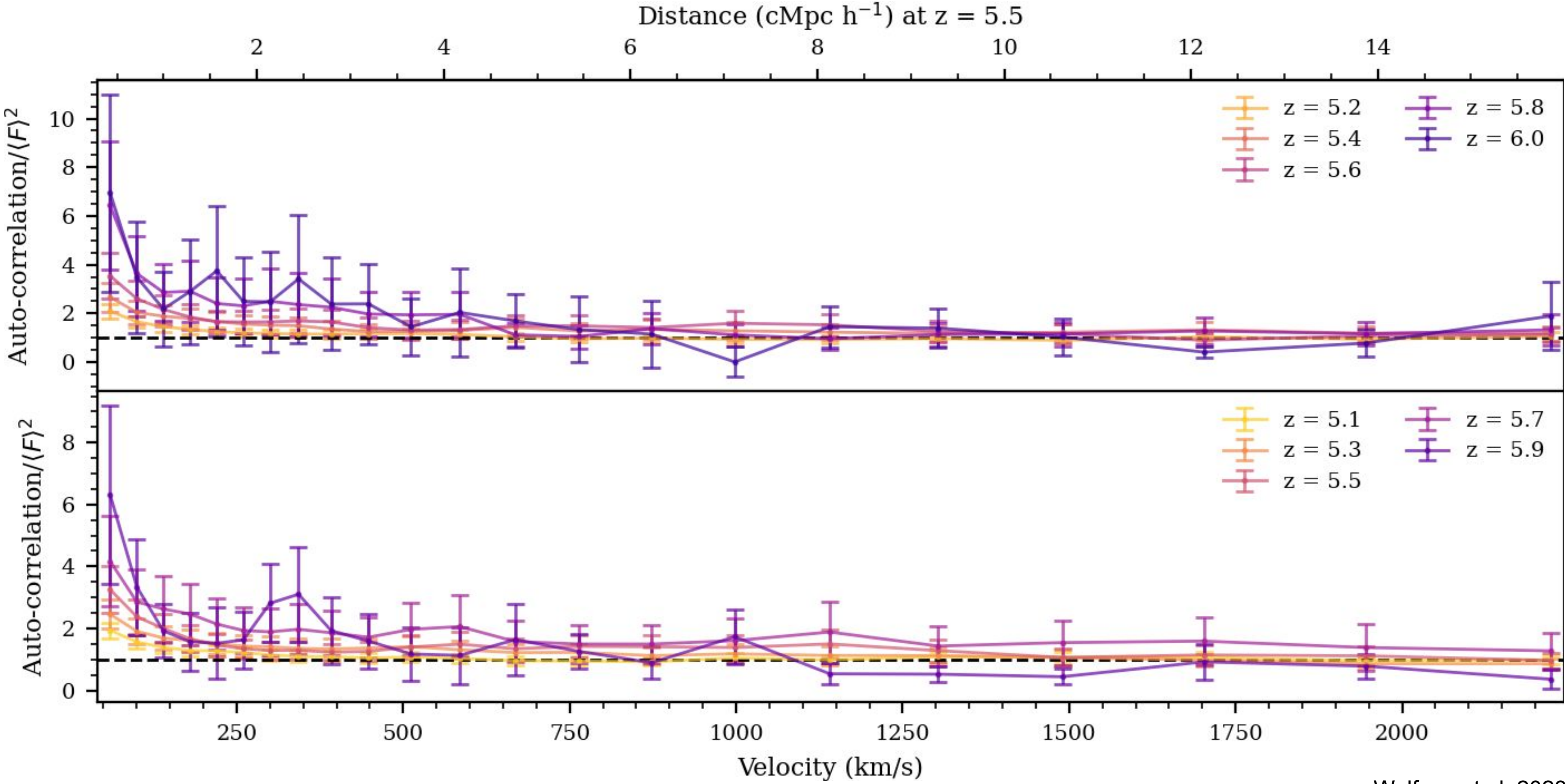


Auto-correlation function from XQR-30



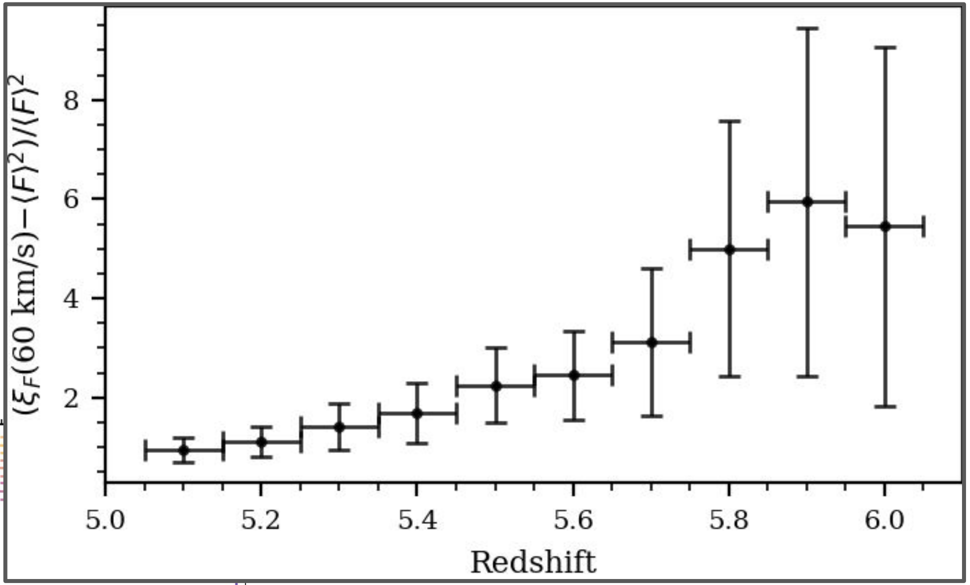
Wolfson et al. 2023

Auto-correlation function from XQR-30

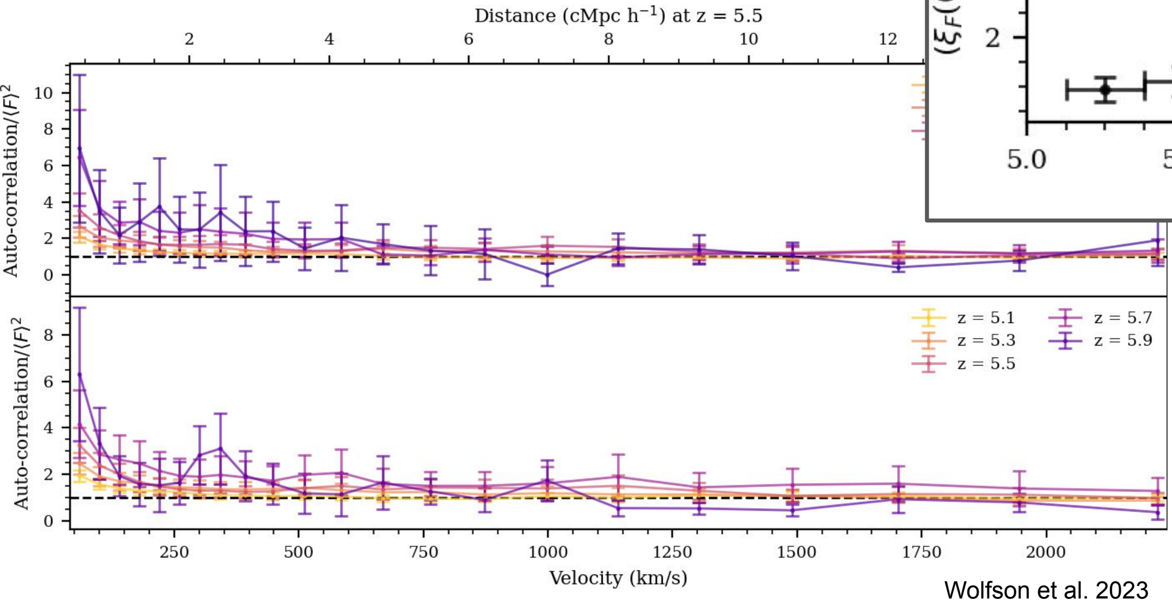


Wolfson et al. 2023

Smallest scale of the auto-correlation function from XQR-30:



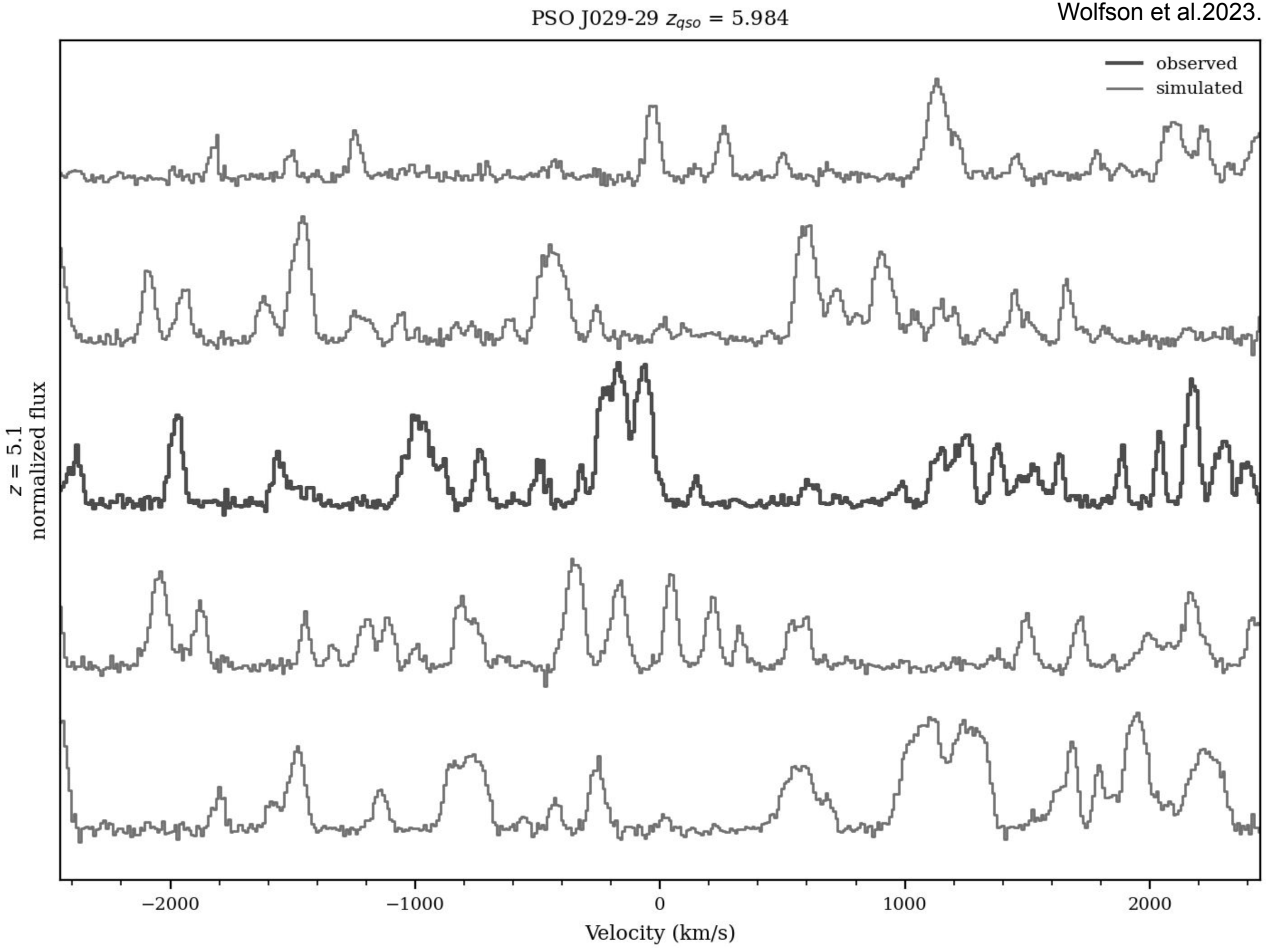
Wolfson et al. 2023



Wolfson et al. 2023

Compare with simulations:

Wolfson et al.2023.



Measuring λ_{mfp}

Gaussian likelihood:

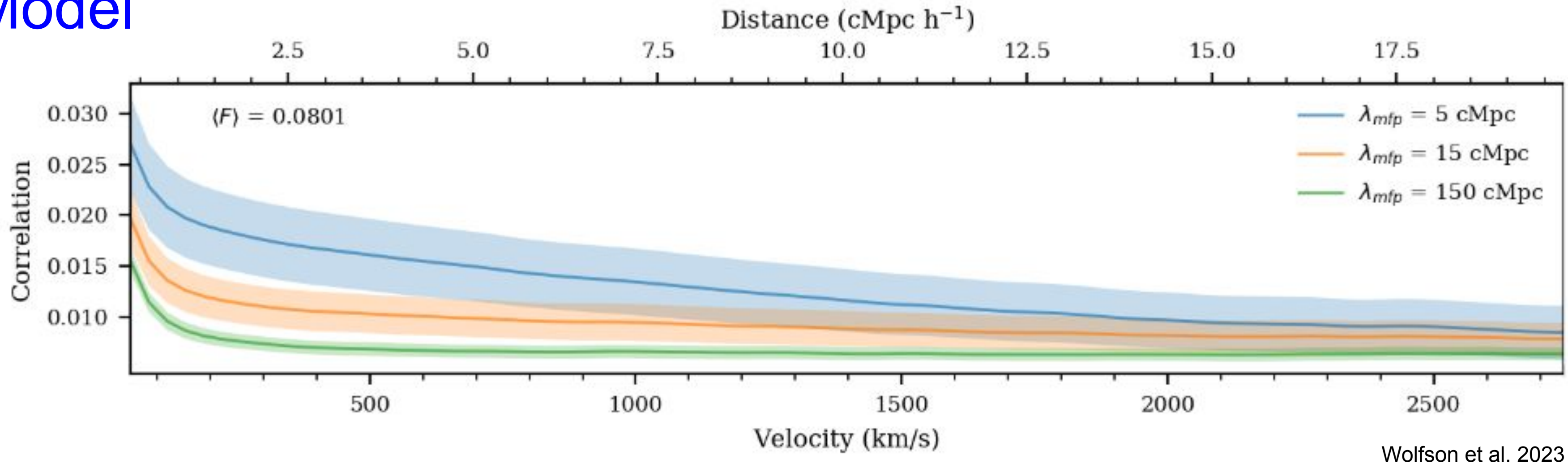
$$\mathcal{L} = \frac{1}{\sqrt{\det(\Sigma)(2\pi)^n}} \exp\left(-\frac{1}{2}\left(\boldsymbol{\xi} - \boldsymbol{\xi}_{\mathrm{model}}(\lambda_{\mathrm{mfp}}, \langle F \rangle)\right)^{\mathrm{T}} \Sigma^{-1} \left(\boldsymbol{\xi} - \boldsymbol{\xi}_{\mathrm{model}}(\lambda_{\mathrm{mfp}}, \langle F \rangle)\right)\right)$$

Measuring λ_{mfp}

Gaussian likelihood:

$$\mathcal{L} = \frac{1}{\sqrt{\det(\Sigma)(2\pi)^n}} \exp\left(-\frac{1}{2}\left(\boldsymbol{\xi} - \boxed{\boldsymbol{\xi}_{\mathrm{model}}(\lambda_{\mathrm{mfp}}, \langle F \rangle)}\right)^{\mathrm{T}} \Sigma^{-1} \left(\boldsymbol{\xi} - \boxed{\boldsymbol{\xi}_{\mathrm{model}}(\lambda_{\mathrm{mfp}}, \langle F \rangle)}\right)\right)$$

Model



Wolfson et al. 2023

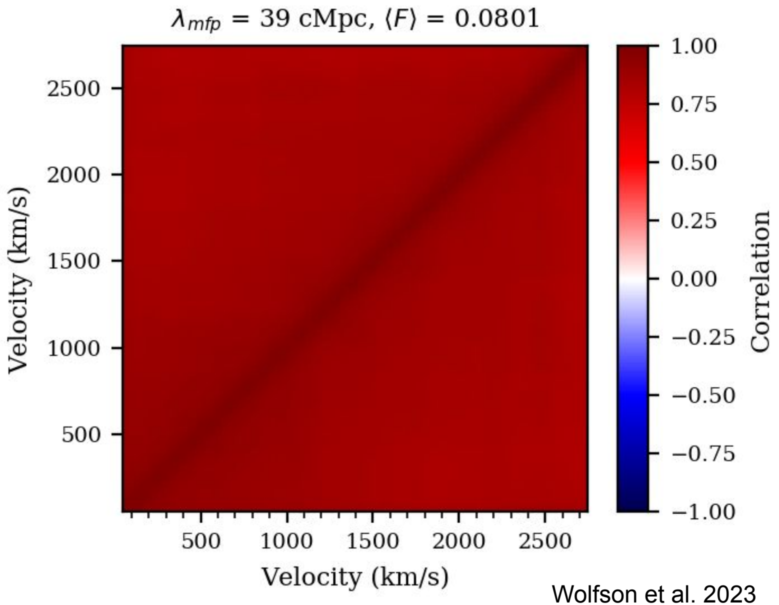
Measuring λ_{mfp}

Gaussian likelihood:

$$\mathcal{L} = \frac{1}{\sqrt{\boxed{\det(\Sigma)}(2\pi)^n}} \exp\left(-\frac{1}{2}\left(\boldsymbol{\xi} - \boxed{\boldsymbol{\xi}_{\mathrm{model}}(\lambda_{\mathrm{mfp}}, \langle F \rangle)}\right)^{\mathrm{T}} \boxed{\Sigma^{-1}} \left(\boldsymbol{\xi} - \boxed{\boldsymbol{\xi}_{\mathrm{model}}(\lambda_{\mathrm{mfp}}, \langle F \rangle)}\right)\right)$$

Model

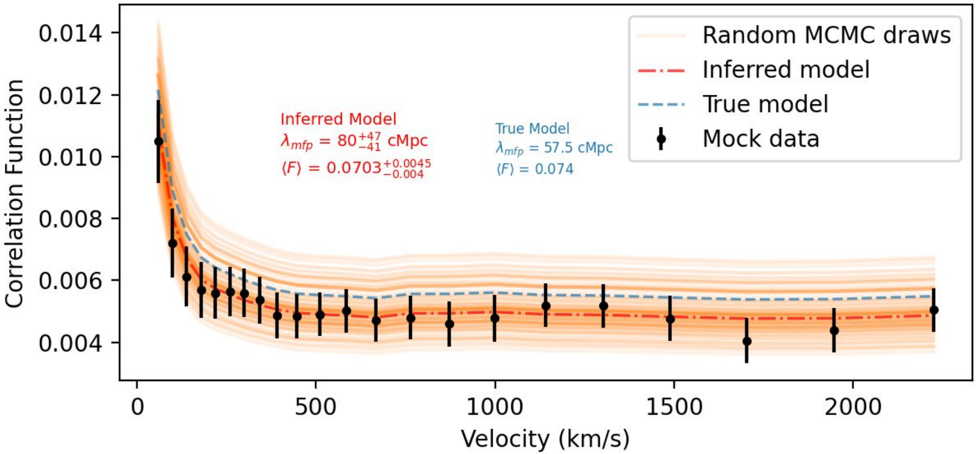
Covariance

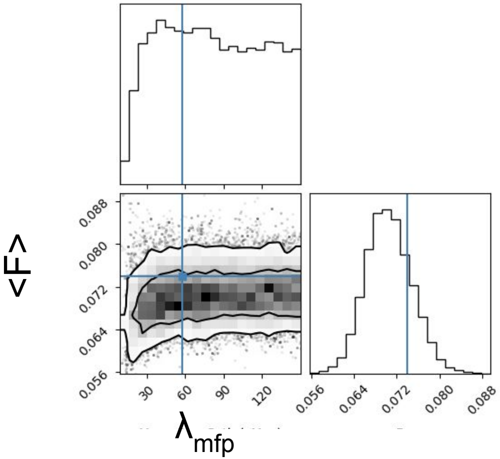


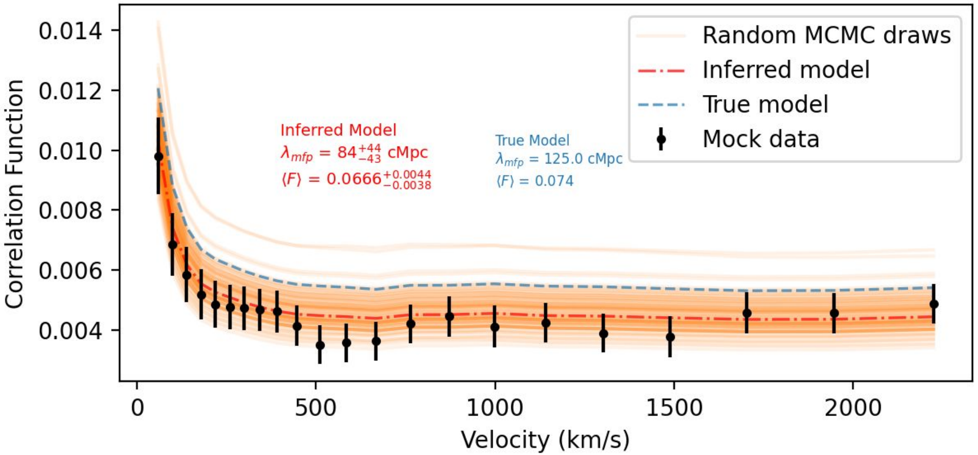
Wolfson et al. 2023

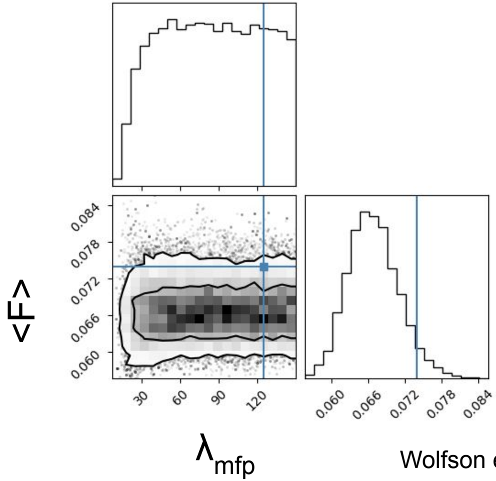
Fitting mock data at z = 5.4:

Wolfson et al. in prep.



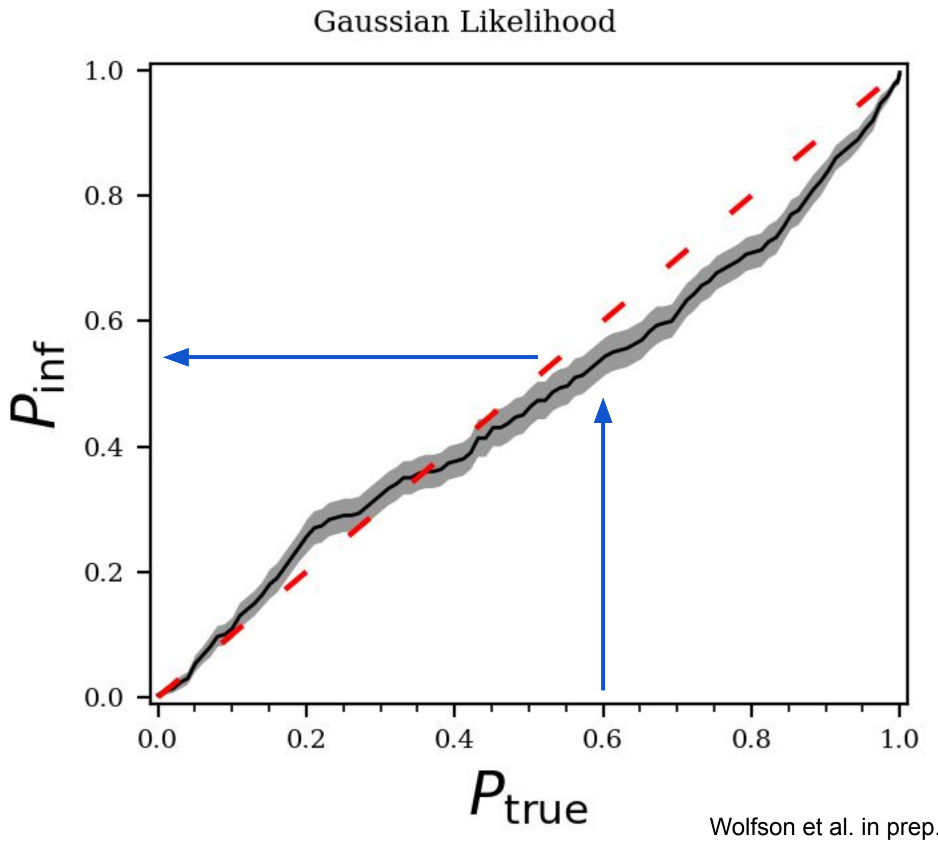






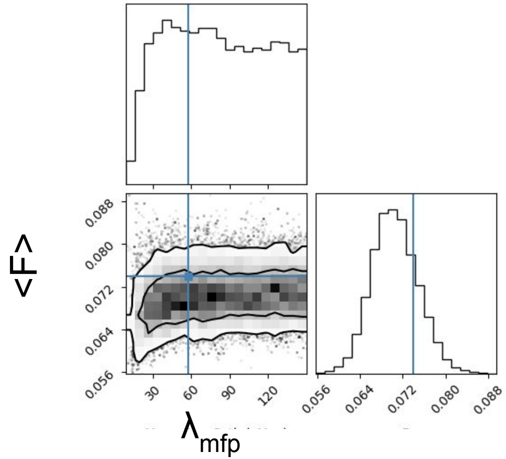
Wolfson et al. in prep.

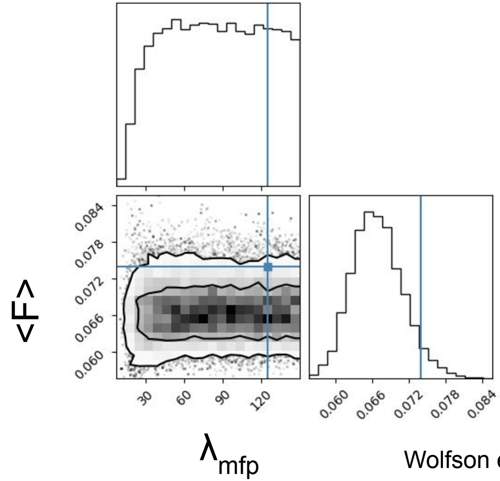
Fitting mock data at z = 5.4:
Is this behaving statistically correctly?



Wolfson et al. in prep.

Wolfson et al. in prep.

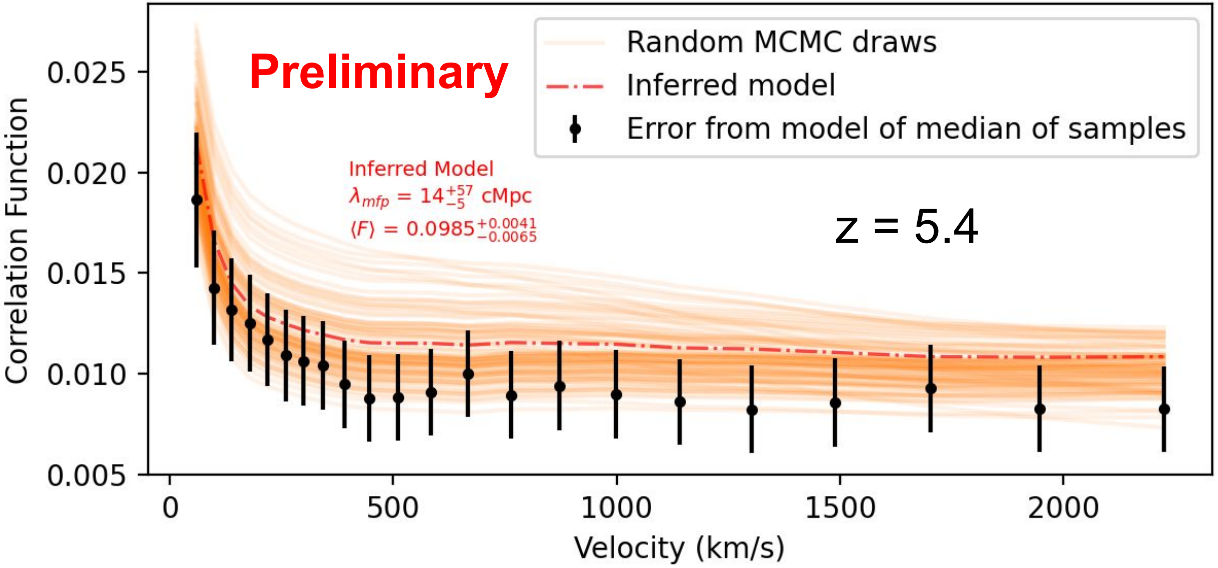


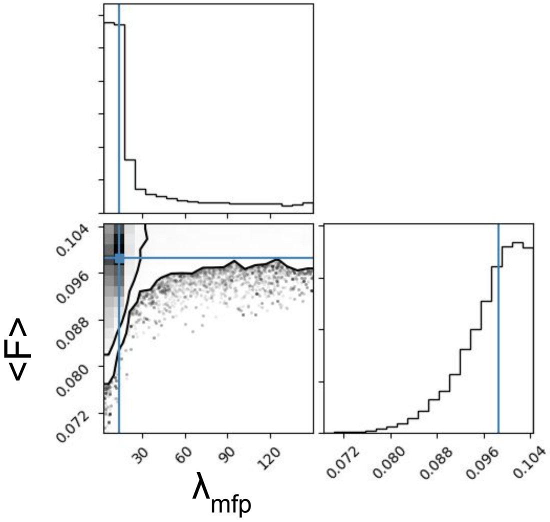


Wolfson et al. in prep.

Attempting a fit on XQR-30 data at z = 5.4

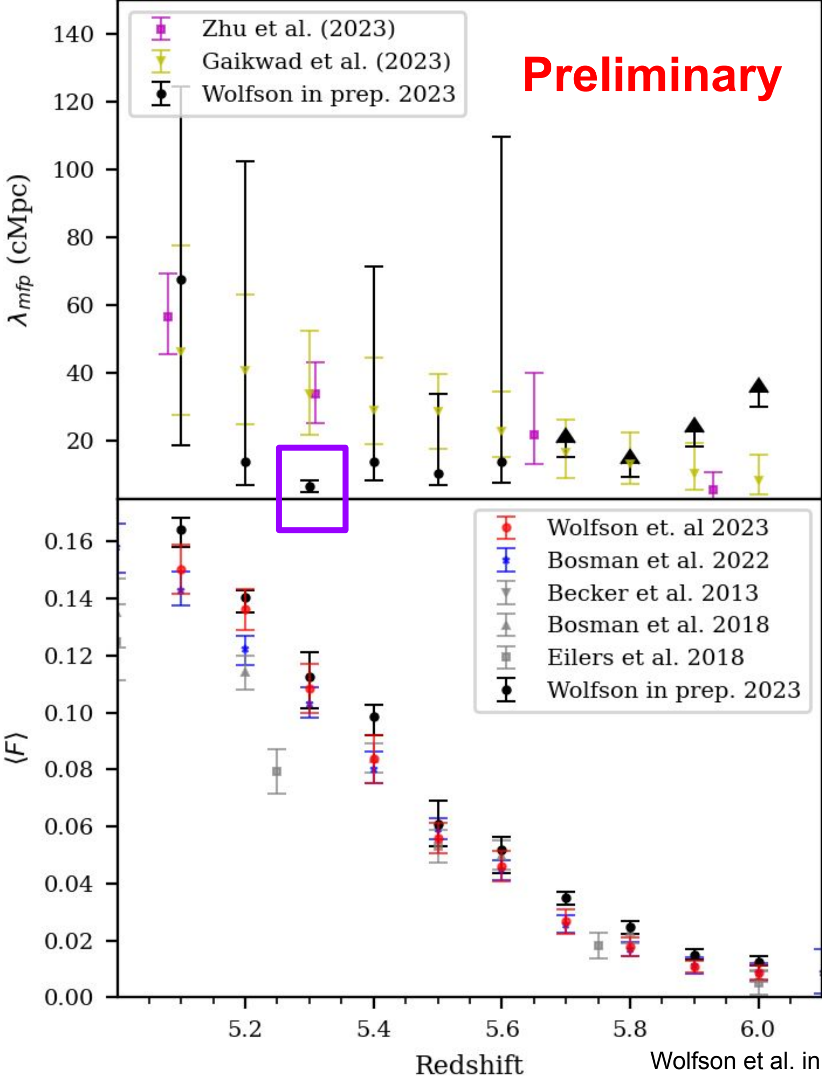
Wolfson et al. in prep.





What do the constraints on λ_{mfp} Look like for all z?

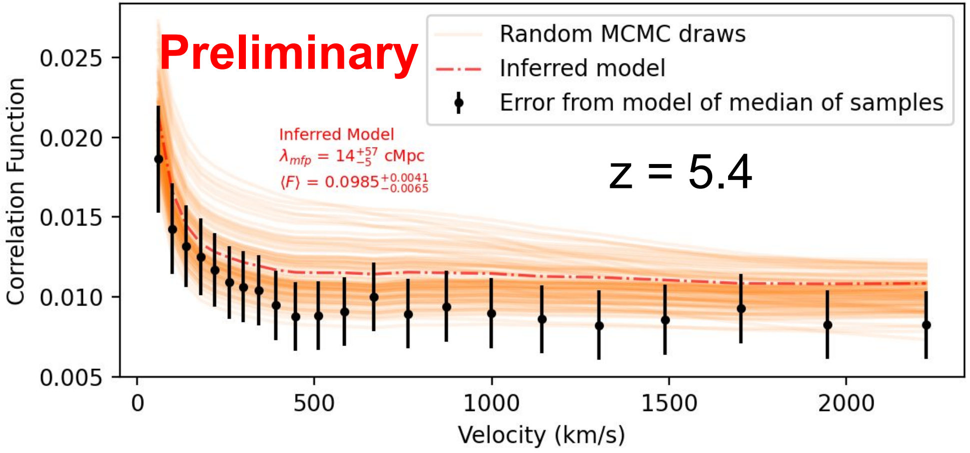
- Compare <F> from the auto-correlation with what was directly measured
- Compare λ_{mfp} from the auto-correlation with the literature values
    - Why do high-z favor large λ_{mfp}?
    - Why is z = 5.3 so precise?

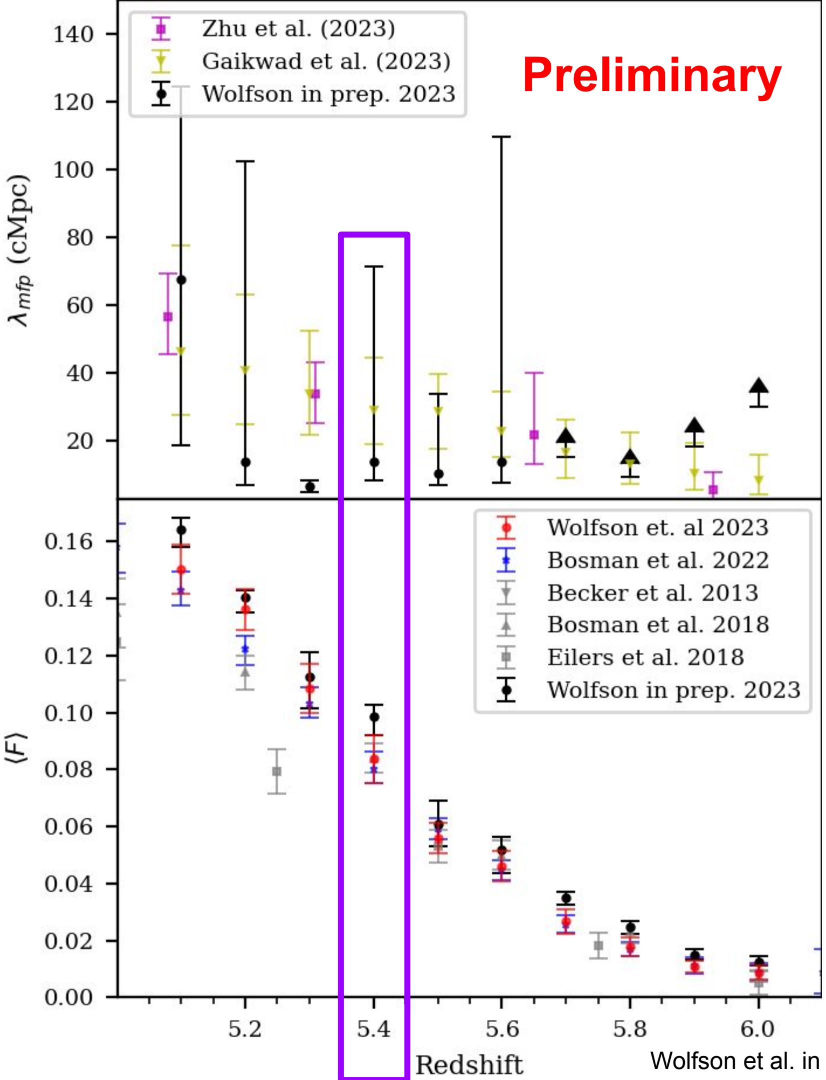


Wolfson et al. in prep.

What do the constraints on λ_{mfp} Look like for all z?

- How good are these fits by eye?

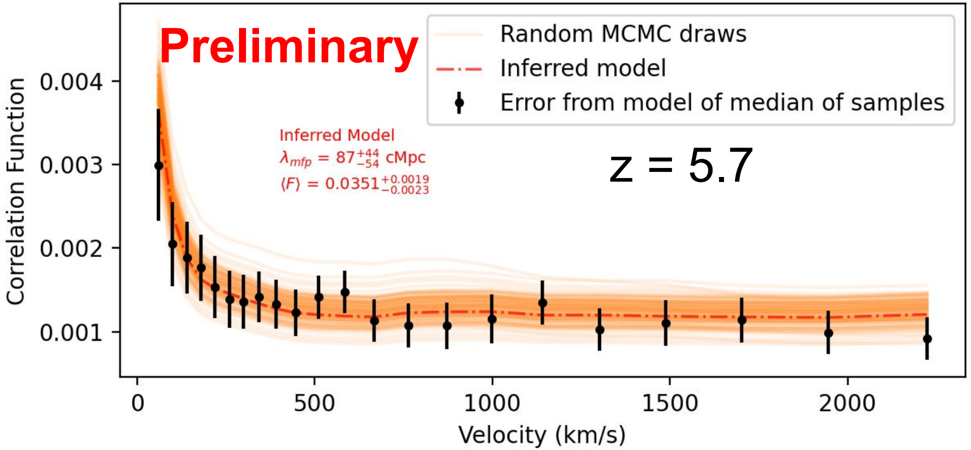


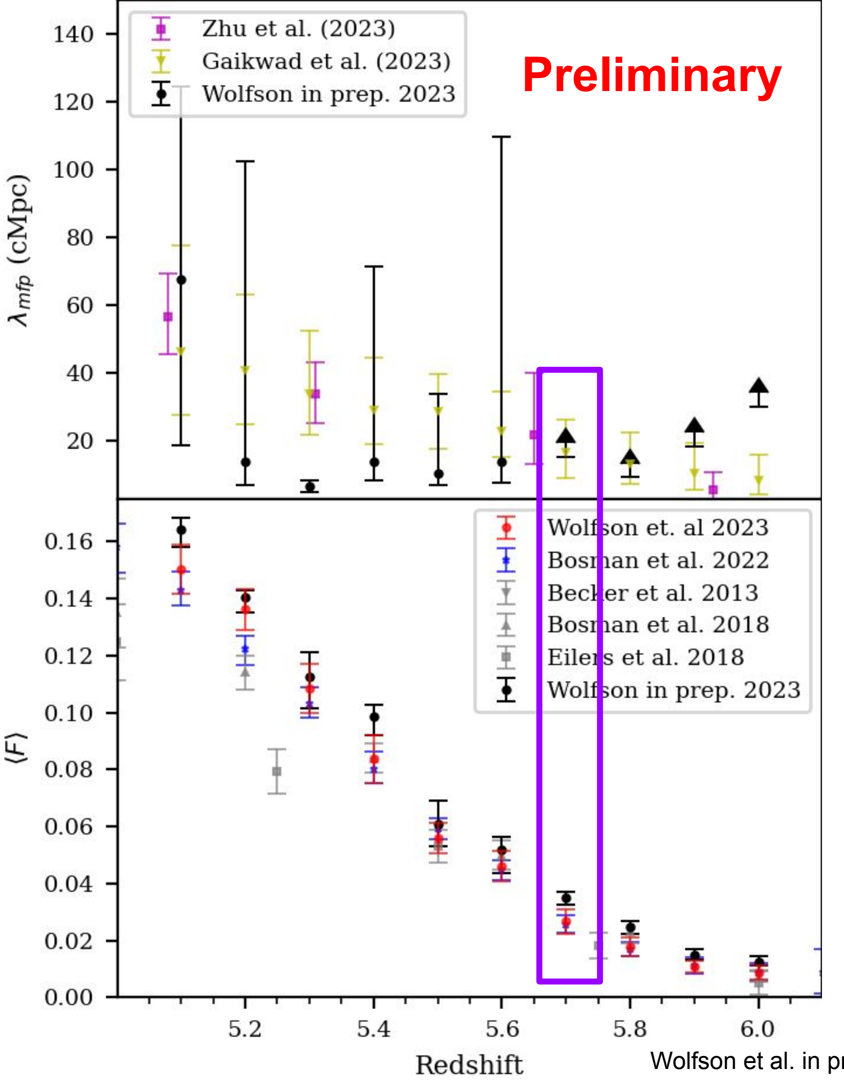


Wolfson et al. in prep.

What do the constraints on λ_{mfp} Look like for all z?

- How good are these fits by eye?

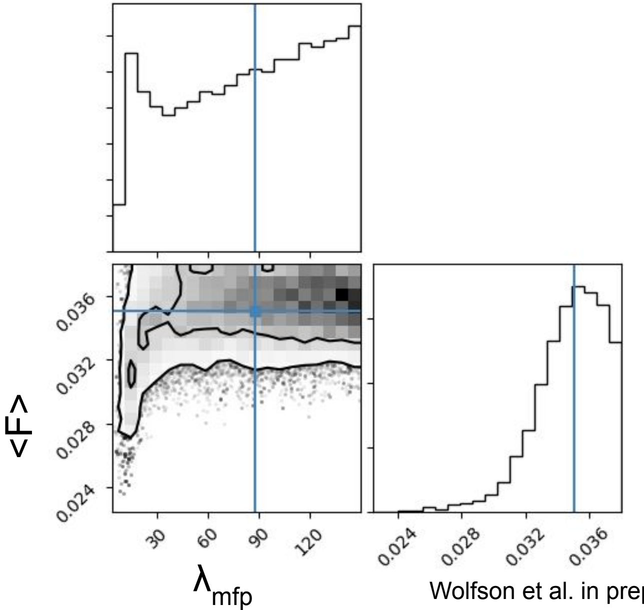




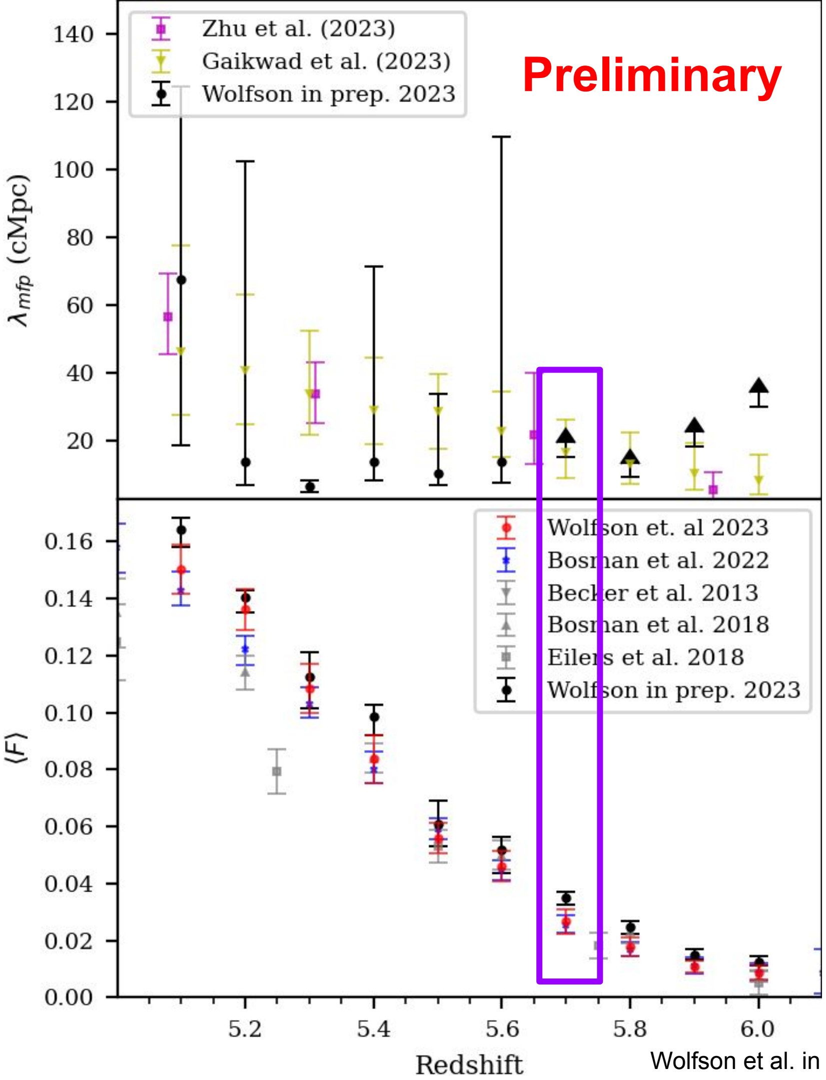
Wolfson et al. in prep.

What do the constraints on λ_{mfp} Look like for all z?

- How good are these fits by eye?



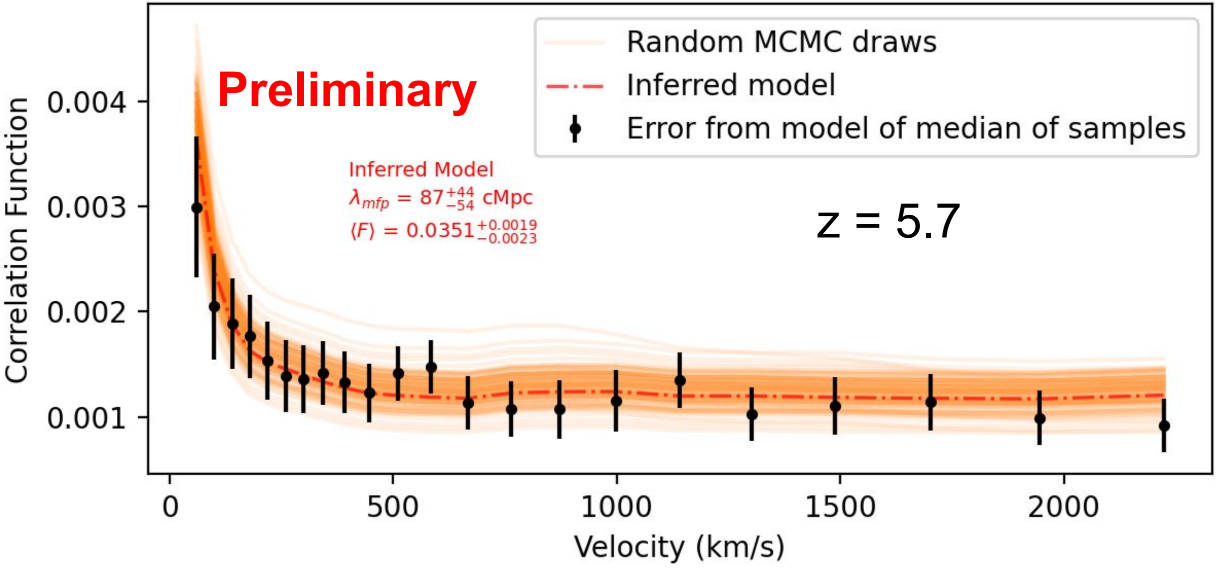
Wolfson et al. in prep.

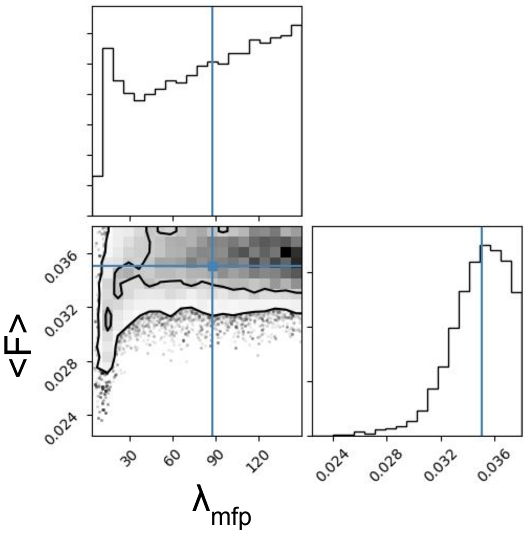


Wolfson et al. in prep.

Attempting a fit on XQR-30 data at z = 5.7

Wolfson et al. in prep.

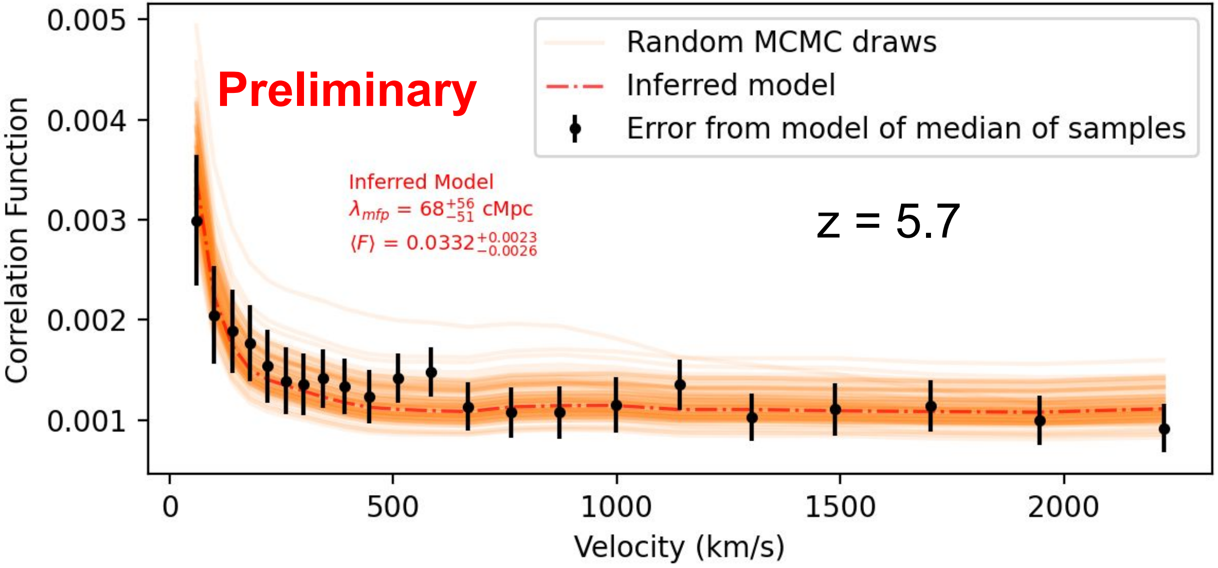


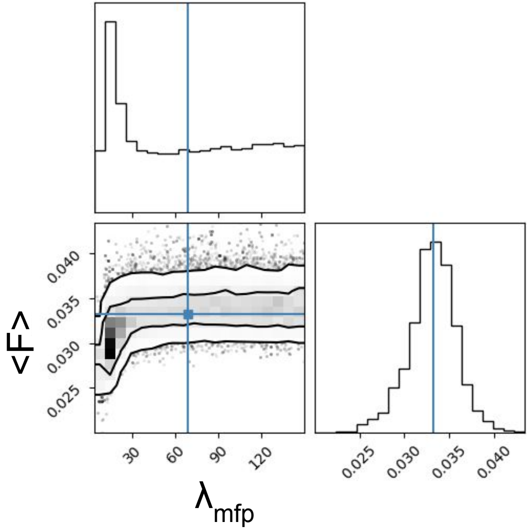


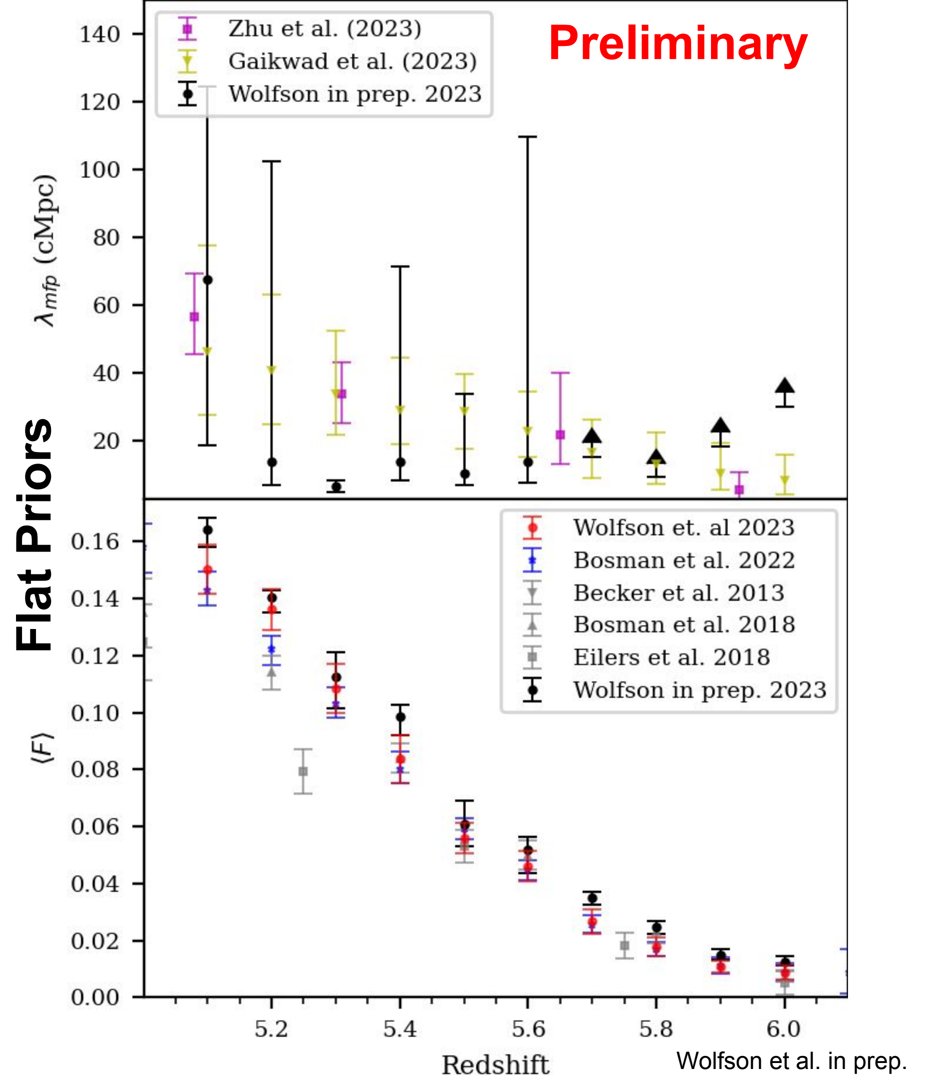
Attempting a fit on XQR-30 data at z = 5.7

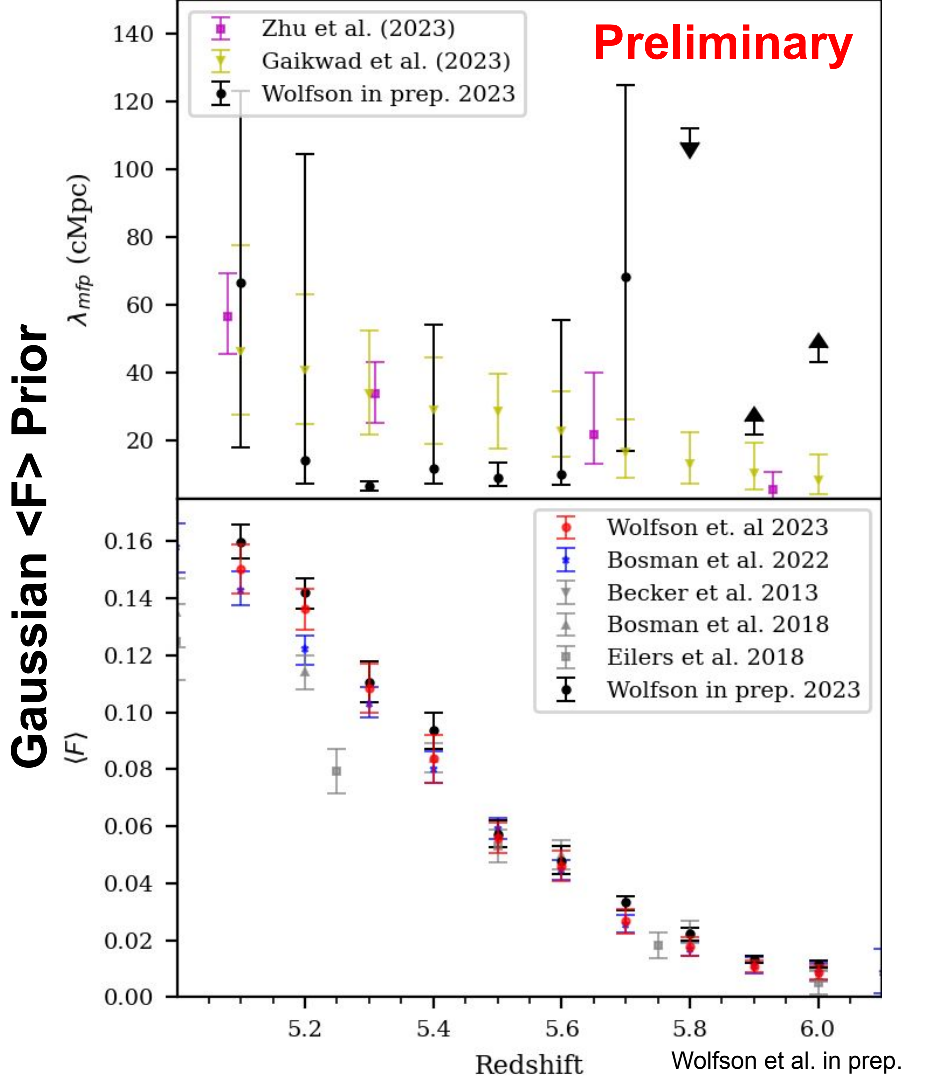
with a Gaussian prior on <F>

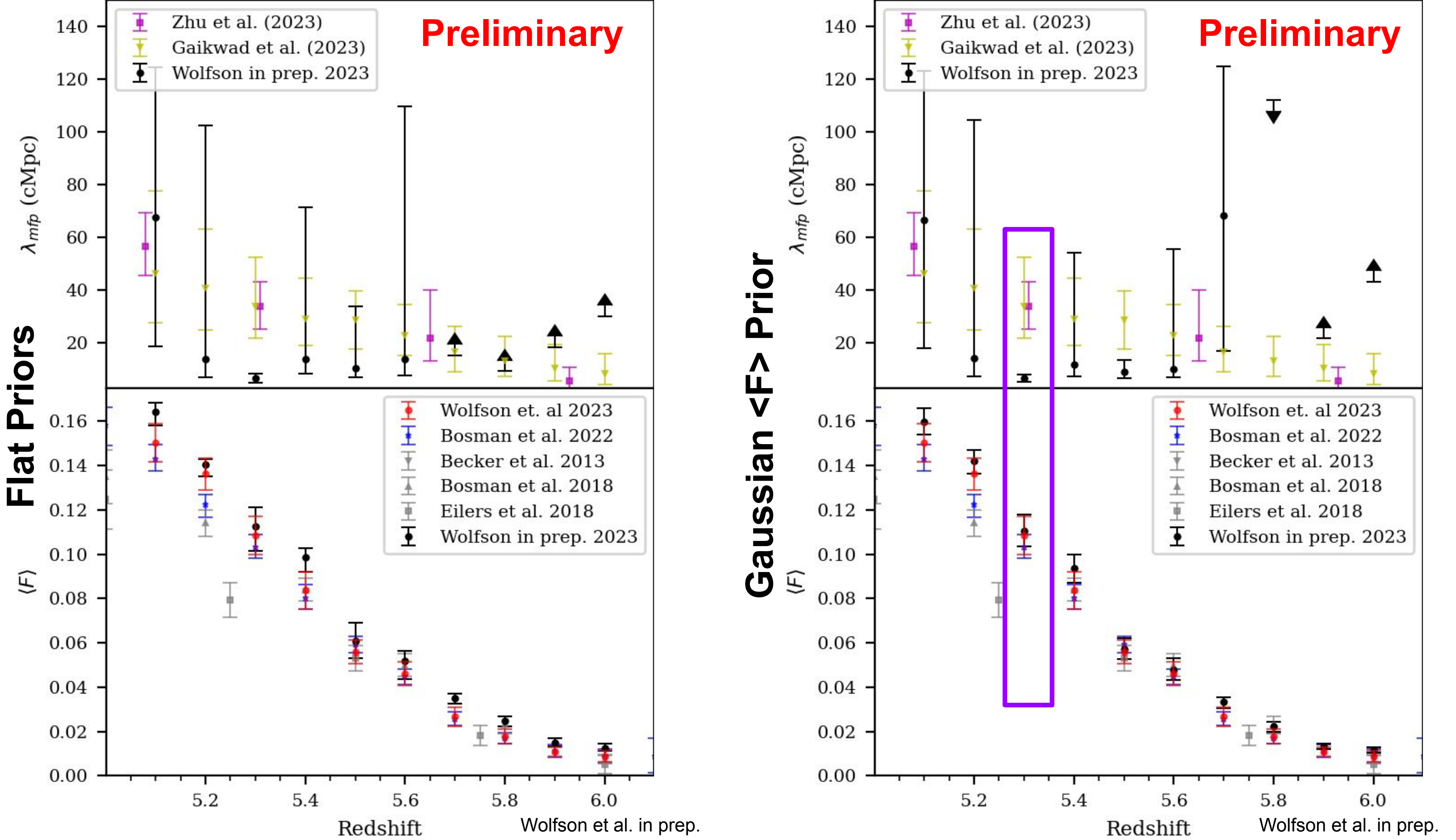
Wolfson et al. in prep.





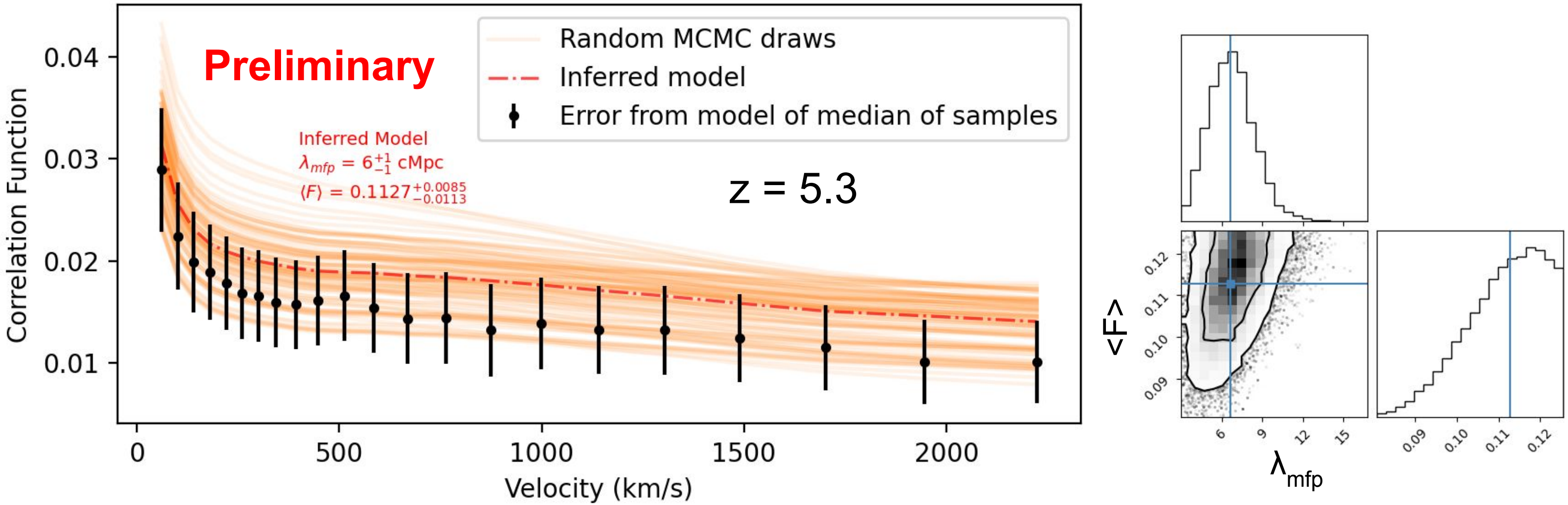

Flat Priors
Preliminary
Zhu et al. (2023)
Gaikwad et al. (2023)
Wolfson in prep. 2023
λ_{mfp} (cMpc)
Wolfson et. al 2023
Bosman et al. 2022
Becker et al. 2013
Bosman et al. 2018
Eilers et al. 2018
Wolfson in prep. 2023
$\langle F \rangle$
Redshift
Wolfson et al. in prep.


Gaussian <F> Prior
Preliminary
Zhu et al. (2023)
Gaikwad et al. (2023)
Wolfson in prep. 2023
λ_{mfp} (cMpc)
Wolfson et. al 2023
Bosman et al. 2022
Becker et al. 2013
Bosman et al. 2018
Eilers et al. 2018
Wolfson in prep. 2023
$\langle F \rangle$
Redshift
Wolfson et al. in prep.


Preliminary
Zhu et al. (2023)
Gaikwad et al. (2023)
Wolfson in prep. 2023
λ_{mfp} (cMpc)
Wolfson et. al 2023
Bosman et al. 2022
Becker et al. 2013
Bosman et al. 2018
Eilers et al. 2018
Wolfson in prep. 2023
$\langle F \rangle$
Redshift
Flat Priors
Gaussian <F> Prior
Wolfson et al. in prep.

Attempting a fit on XQR-30 data at z = 5.3

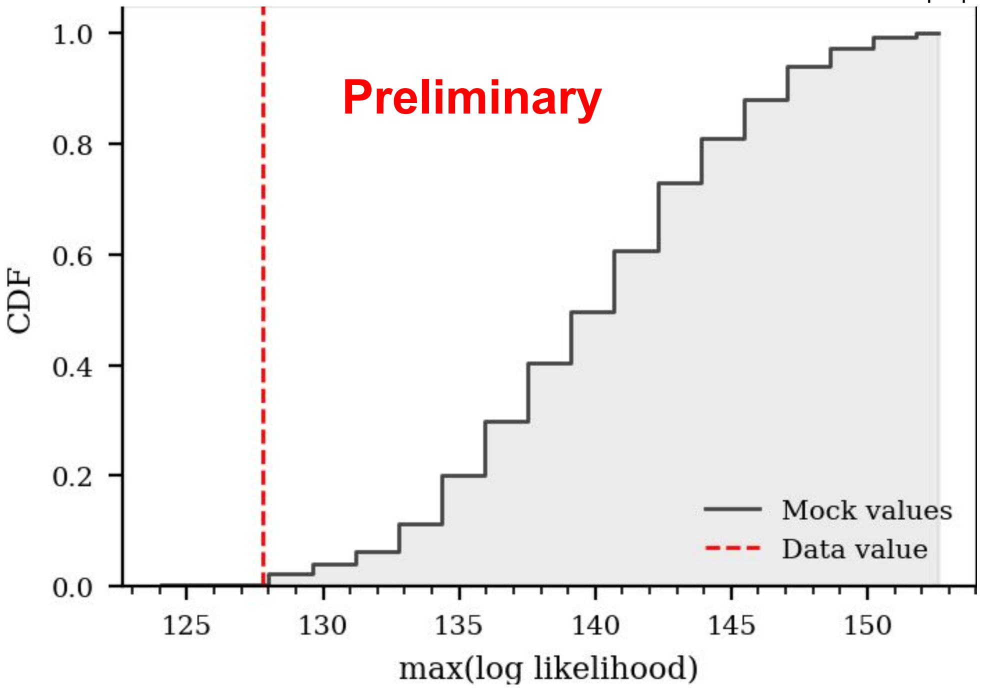
Wolfson et al. in prep.



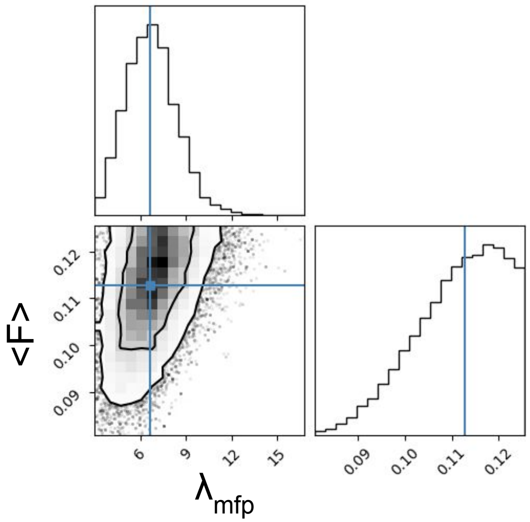
Can we quantify how ‘good’ this fit is compared to the mock data?

Attempting a fit on XQR-30 data at z = 5.3

Wolfson et al. in prep.



Wolfson et al. in prep.



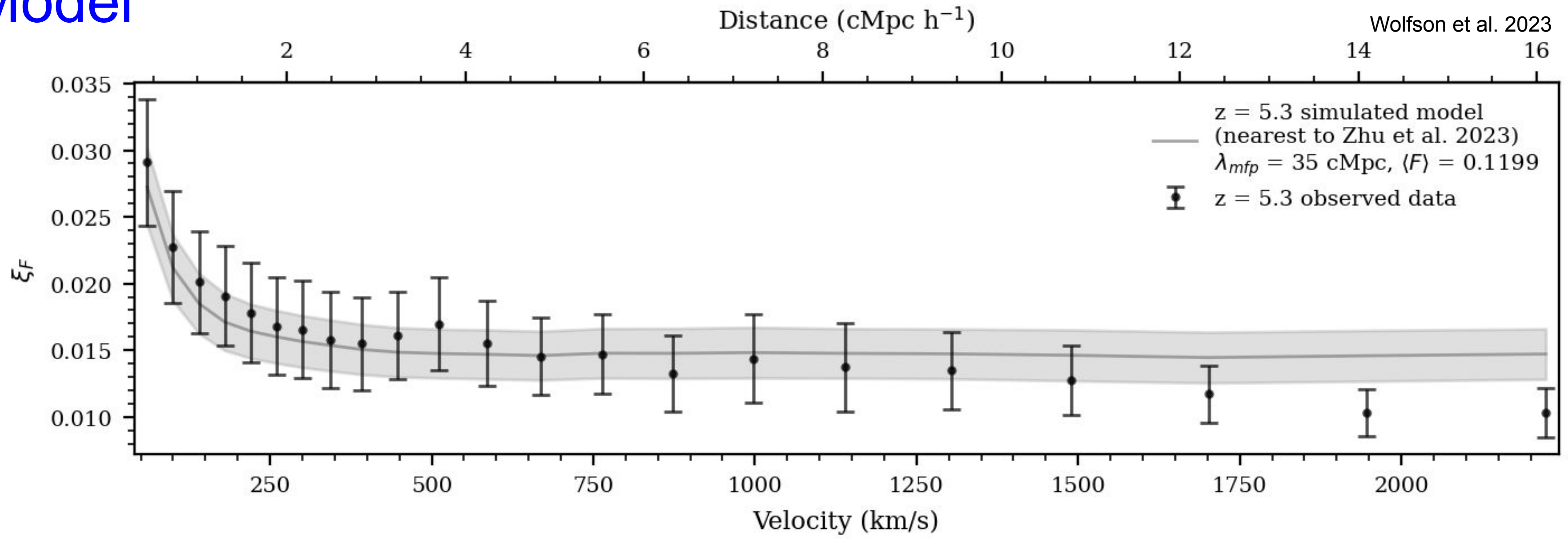
Can we quantify how 'good' this fit is compared to the mock data?

What is going wrong with fitting the data?

$$\mathcal{L} = \frac{1}{\sqrt{\det(\Sigma)(2\pi)^n}} \exp\left(-\frac{1}{2}\left(\boldsymbol{\xi} - \boxed{\boldsymbol{\xi}_{\rm model}(\lambda_{\rm mfp}, \langle F \rangle)}\right)^{\rm T} \Sigma^{-1} \left(\boldsymbol{\xi} - \boxed{\boldsymbol{\xi}_{\rm model}(\lambda_{\rm mfp}, \langle F \rangle)}\right)\right)$$

Model

Wolfson et al. 2023

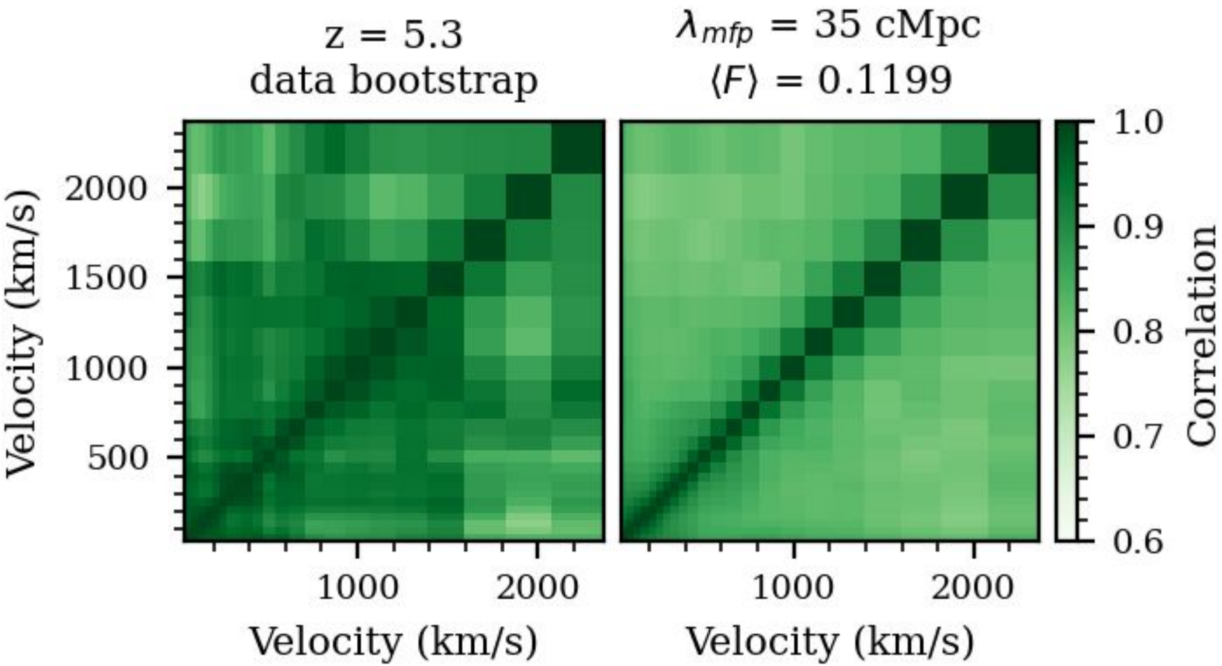


What is going wrong with fitting the data?

$$\mathcal{L} = \frac{1}{\sqrt{\boxed{\det(\Sigma)}(2\pi)^n}} \exp\left(-\frac{1}{2}\left(\boldsymbol{\xi} - \boldsymbol{\xi}_{\mathrm{model}}(\lambda_{\mathrm{mfp}}, \langle F\rangle)\right)^{\mathrm{T}} \boxed{\Sigma^{-1}} \left(\boldsymbol{\xi} - \boldsymbol{\xi}_{\mathrm{model}}(\lambda_{\mathrm{mfp}}, \langle F\rangle)\right)\right)$$

Wolfson et al. 2023

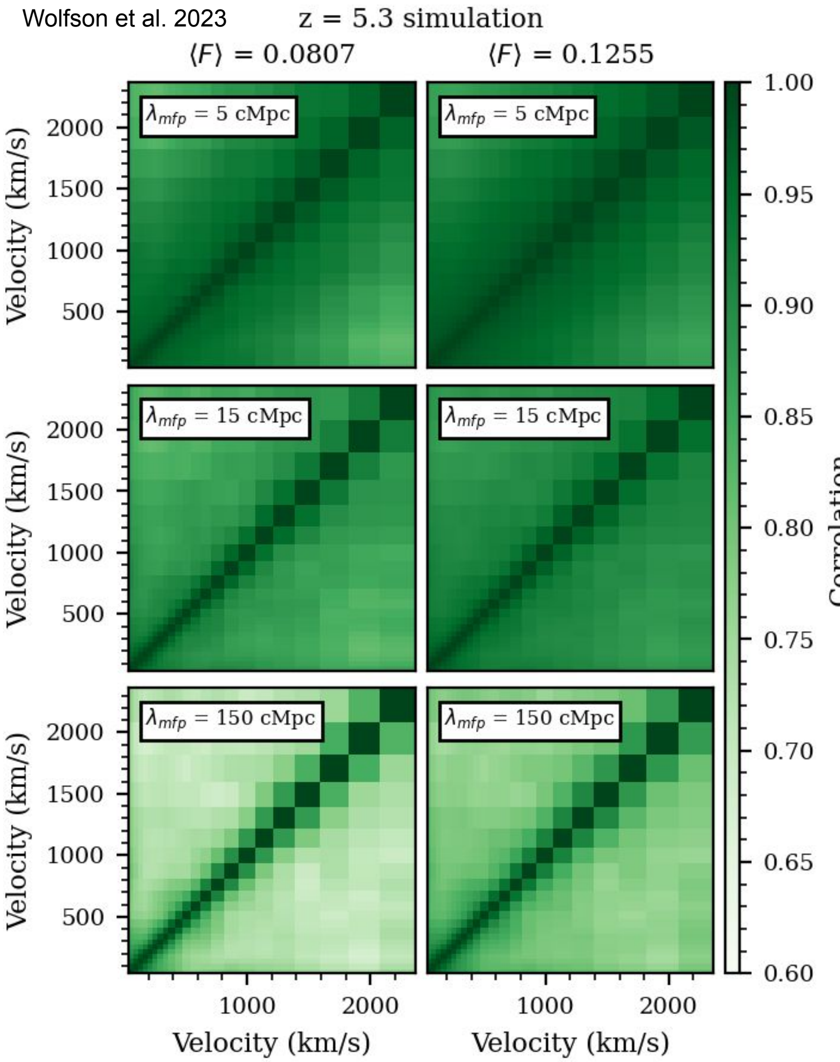
Covariance



Does the covariance matrix vary with λ_{mfp} or <F>?

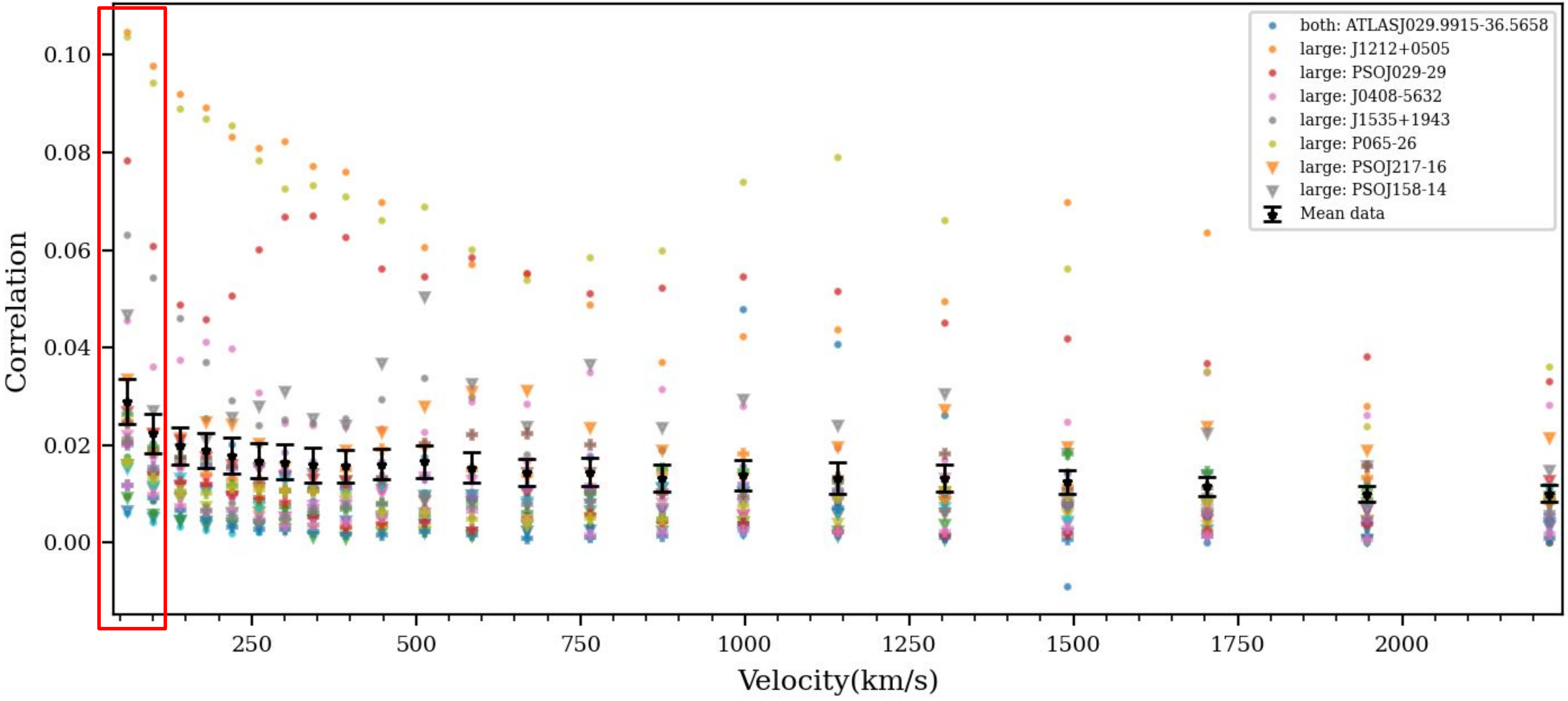
We are not only fitting the model line but also this covariance matrix structure

Wolfson et al. 2023



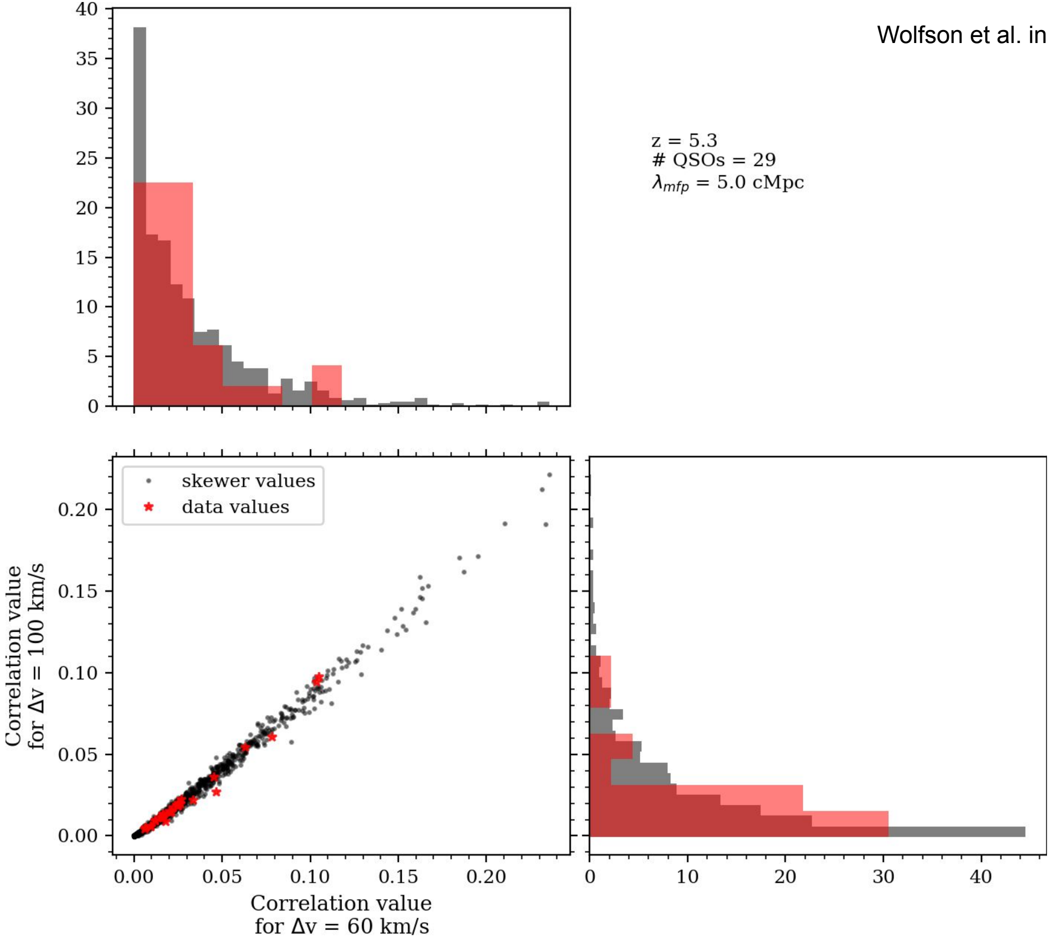
How do individual sightlines contribute to the bootstrap covariance matrix?

Wolfson et al. in prep.



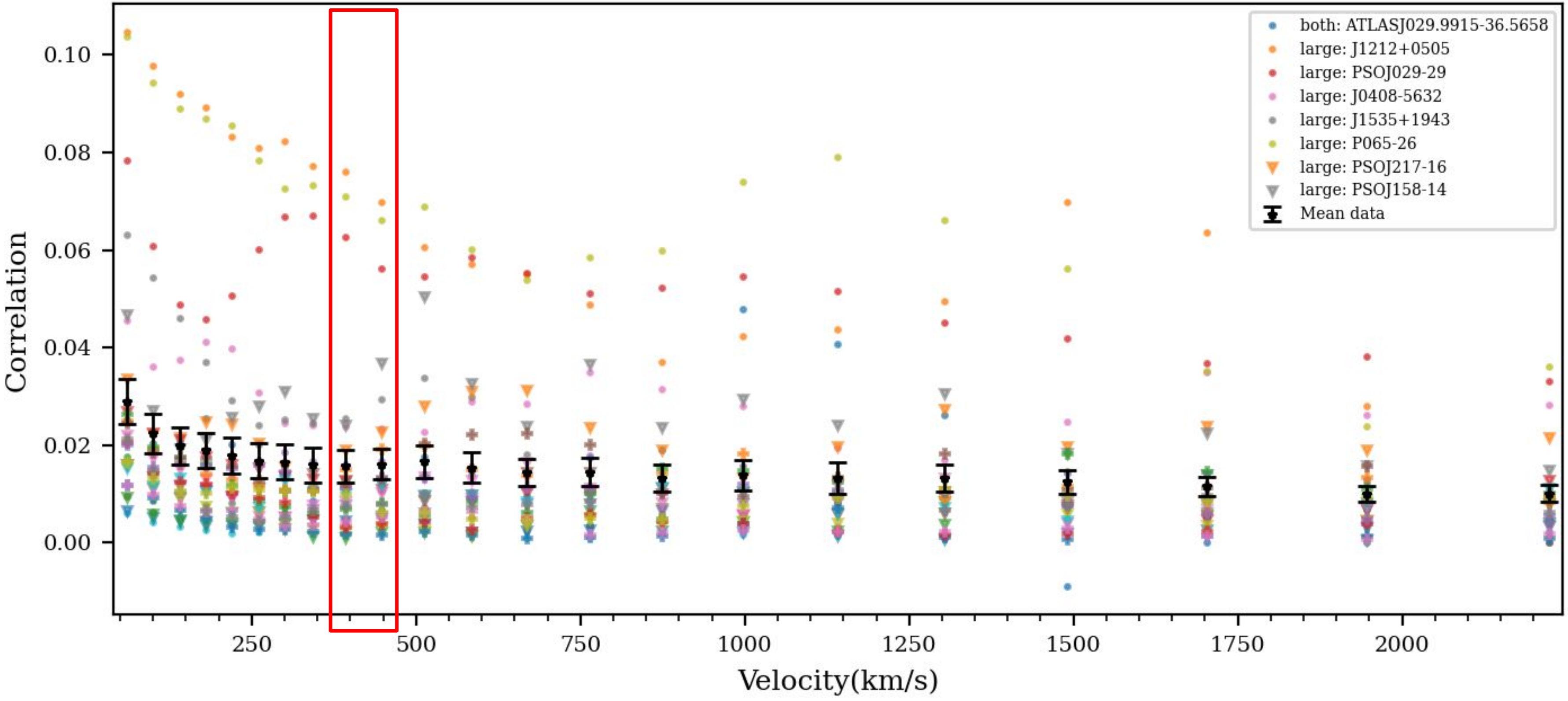
How do individual sightlines contribute to the bootstrap covariance matrix?

Wolfson et al. in prep.



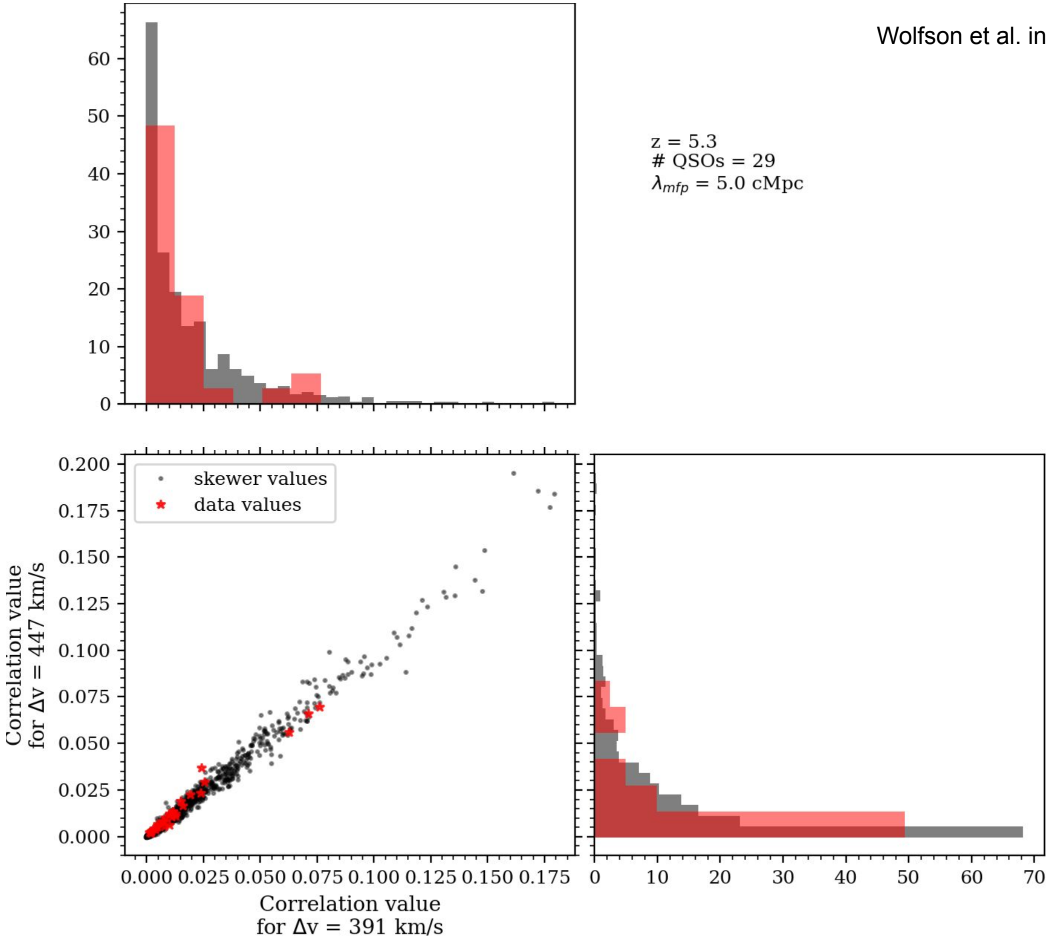
How do individual sightlines contribute to the bootstrap covariance matrix?

Wolfson et al. in prep.



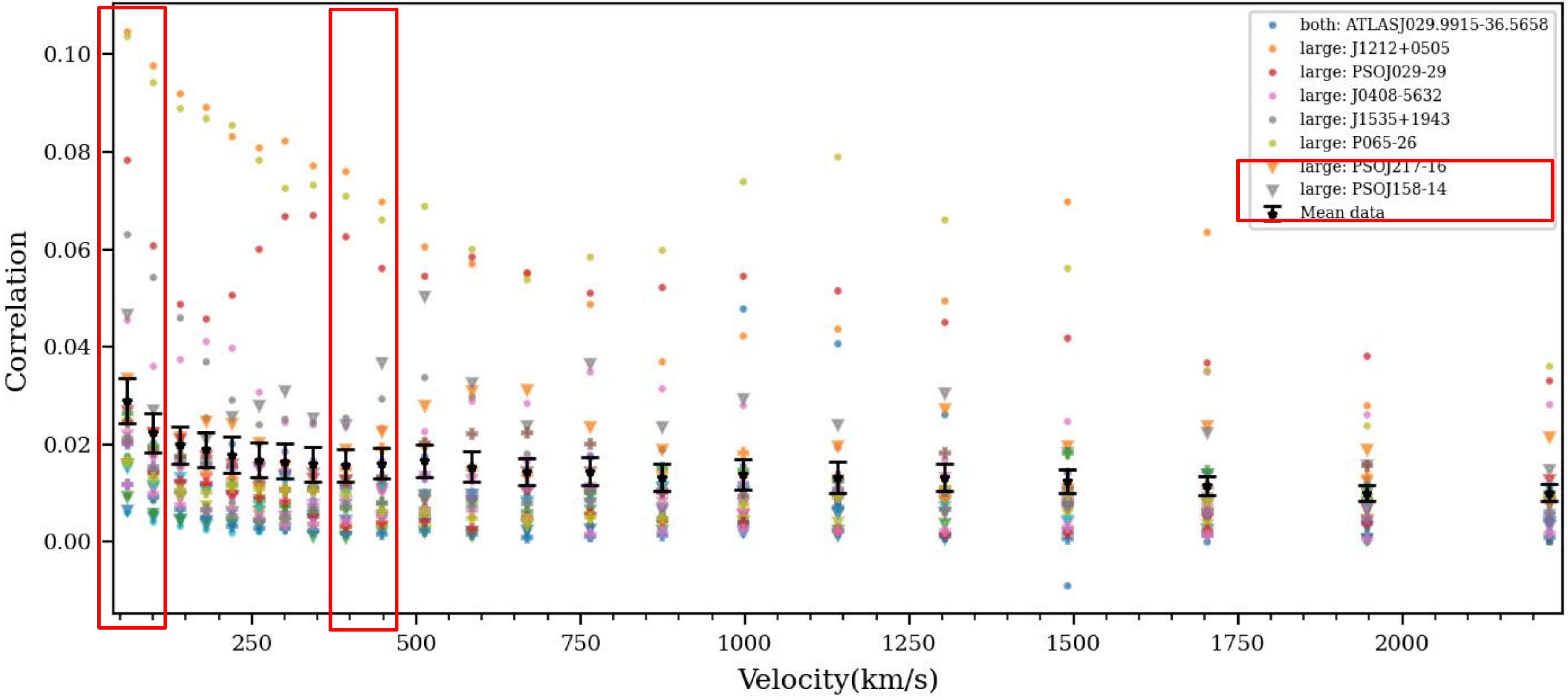
How do individual sightlines contribute to the bootstrap covariance matrix?

Wolfson et al. in prep.

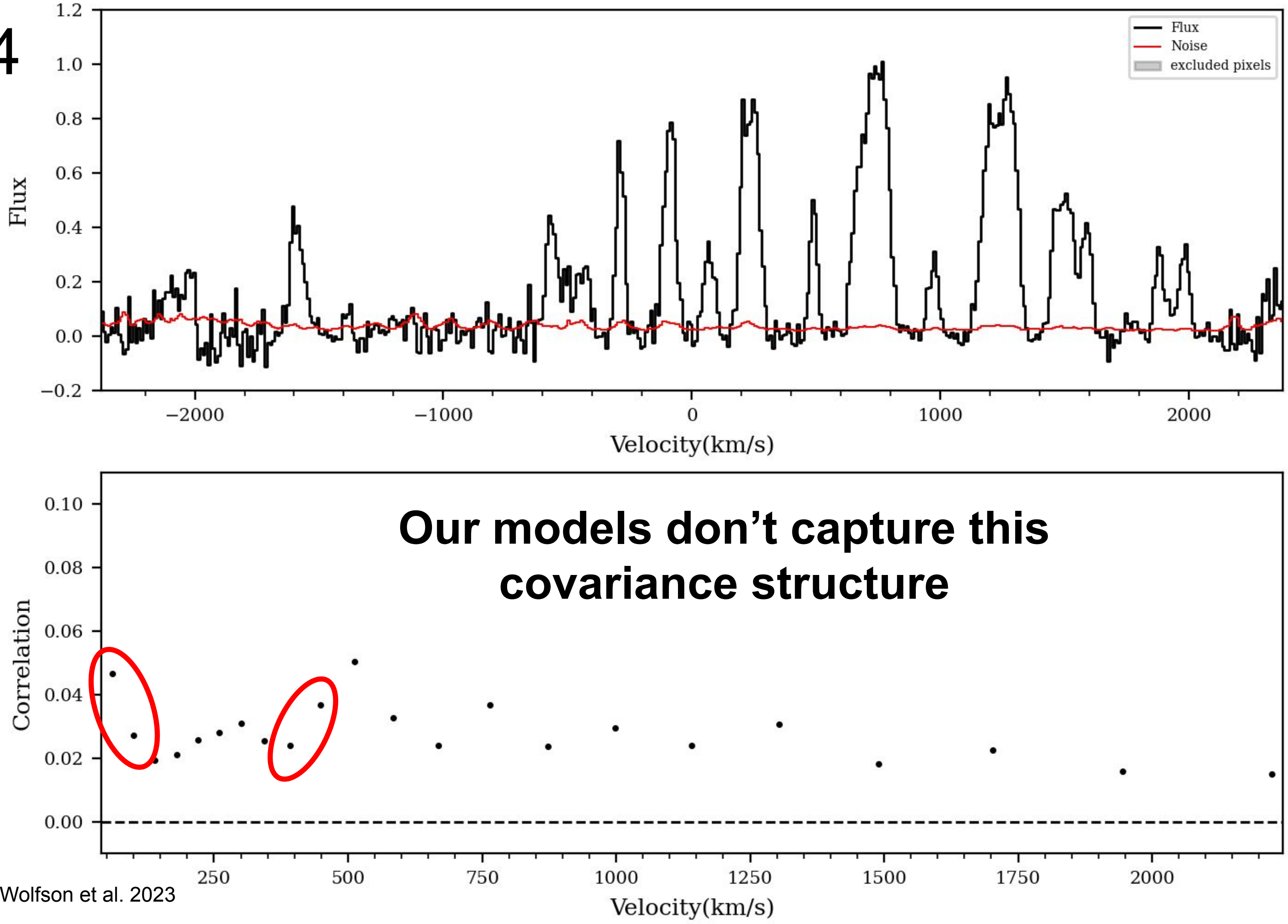


How do individual sightlines contribute to the bootstrap covariance matrix?

Wolfson et al. in prep.



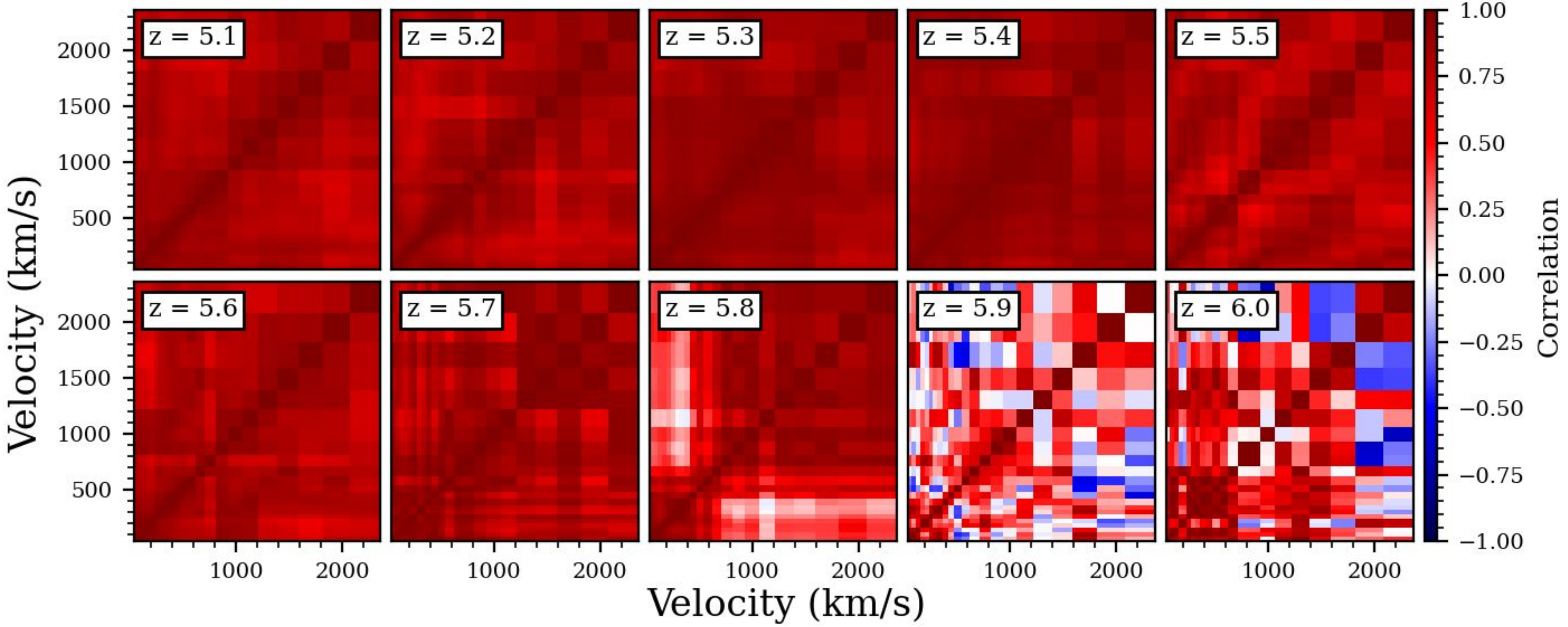
PSOJ158-14



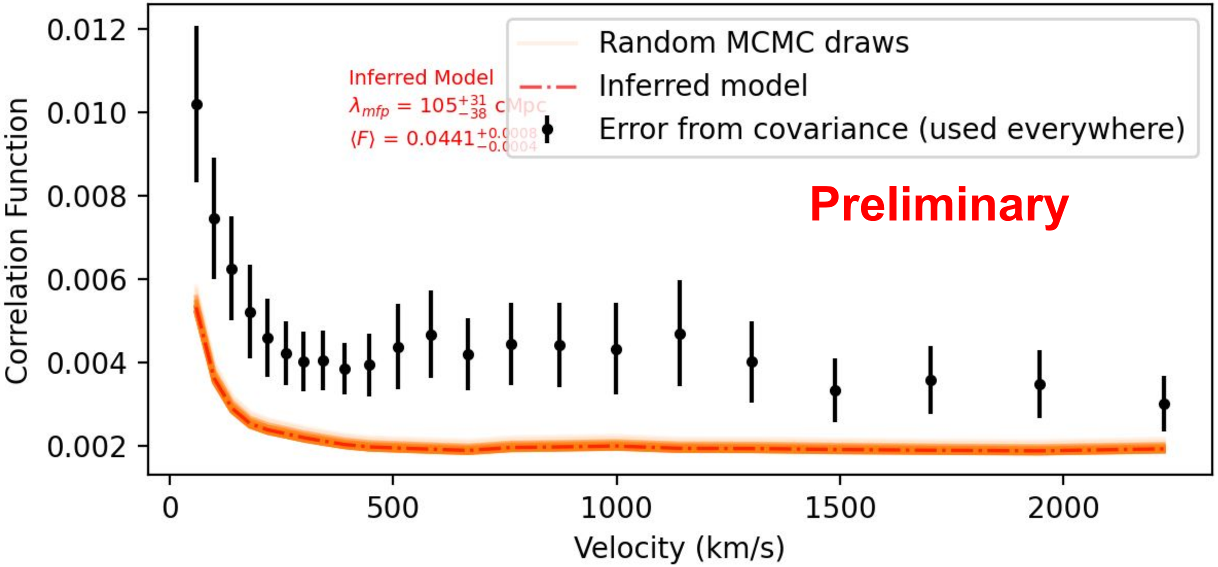
Wolfson et al. 2023

Can we measure the covariance matrix from the data?

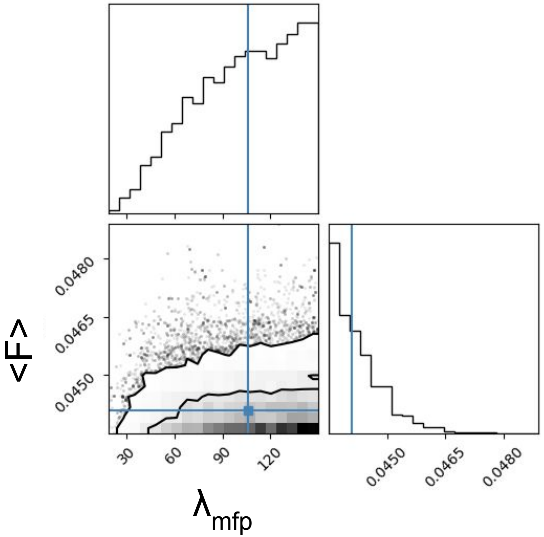
Wolfson et al 2023.



Attempting a fit on XQR-30 data at z = 5.6 with the covariance matrix estimated via bootstrapping:



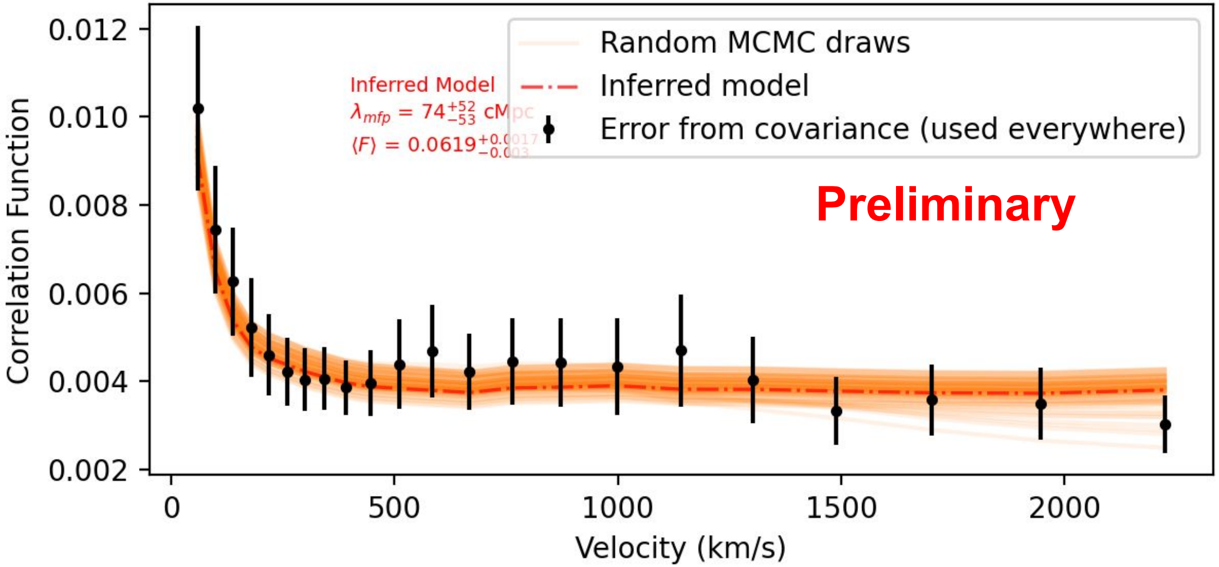
Wolfson et al. in prep.

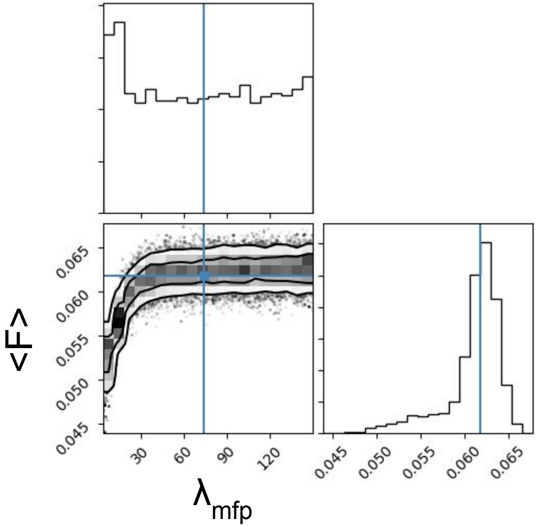


The bootstrap estimates are not good enough even where we have the most data

Attempting a fit on XQR-30 data at z = 5.6 with the covariance matrix estimated via bootstrapping:

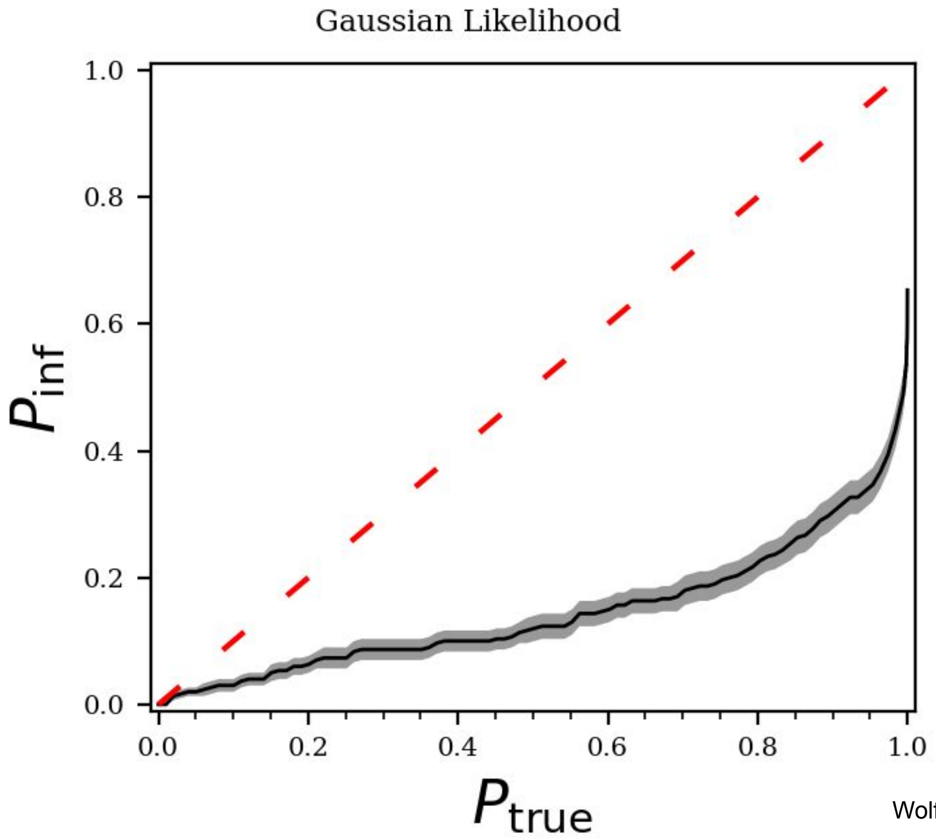
Wolfson et al. in prep.





Only using the diagonal of the bootstrap covariance matrix gives a good fit… but is this a valid thing to do?

Does using the diagonal of a covariance matrix estimated from bootstrap resampling the data behave statistically correctly?

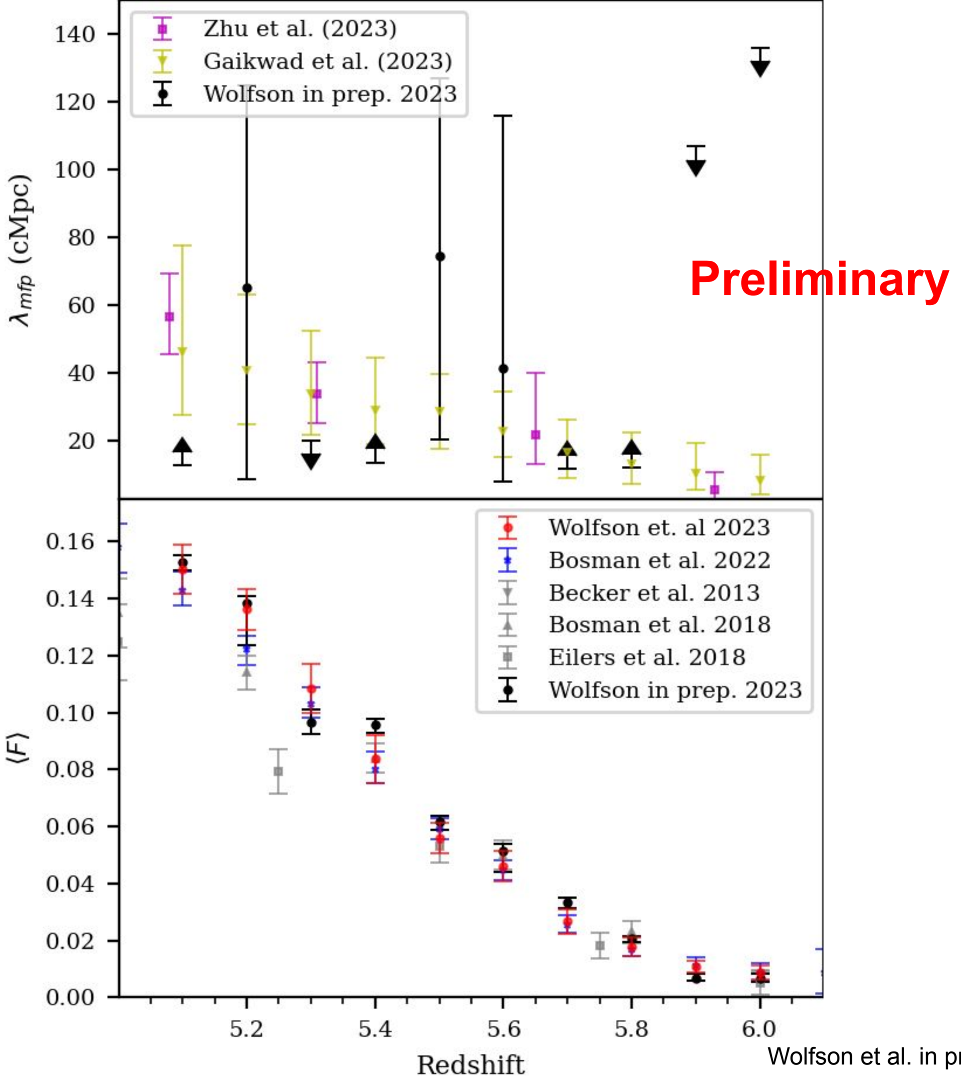


No

Wolfson et al. in prep.

What do the constraints on λ_{mfp} look like for all z for a diagonal covariance matrix from bootstrap resampling?

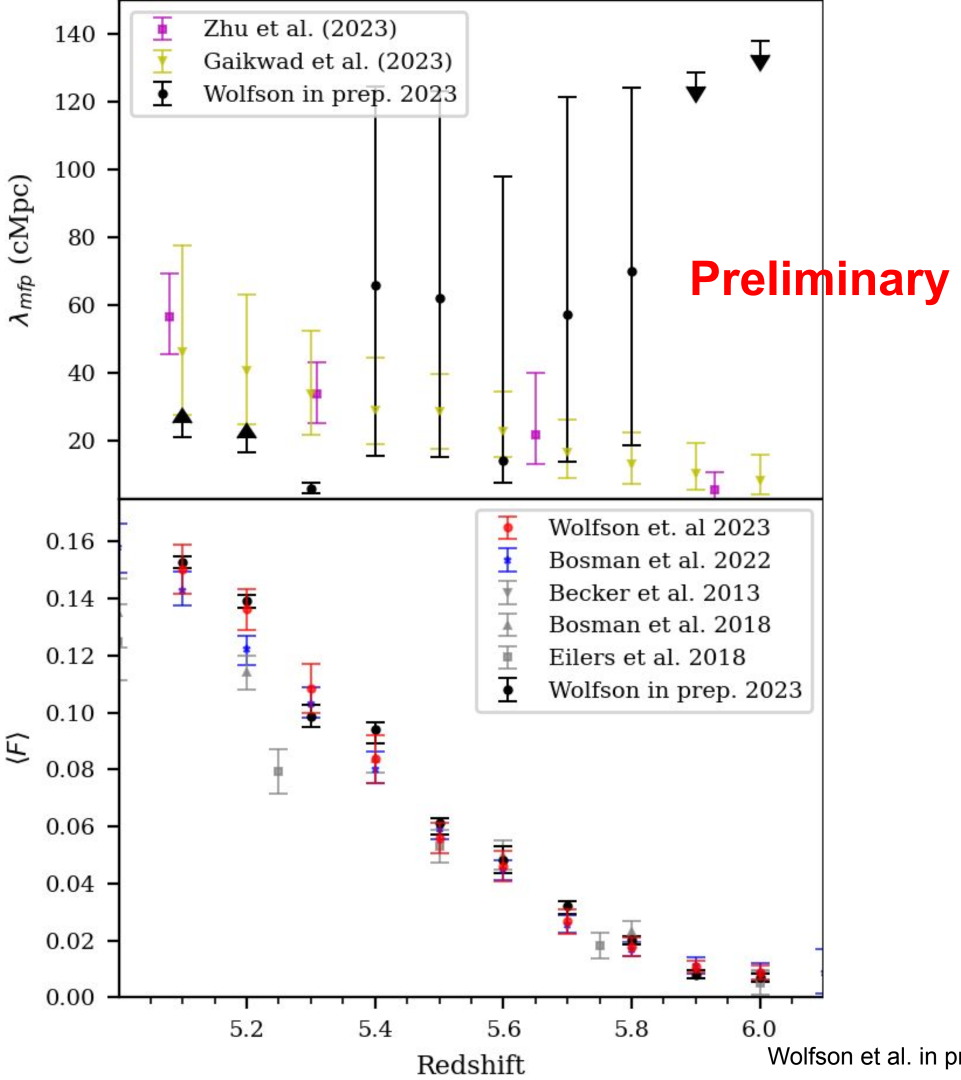
With flat priors and using all data



Wolfson et al. in prep.

What do the constraints on λ_{mfp} look like for all z for a diagonal covariance matrix from bootstrap resampling?

With a gaussian prior on <F> and using all data

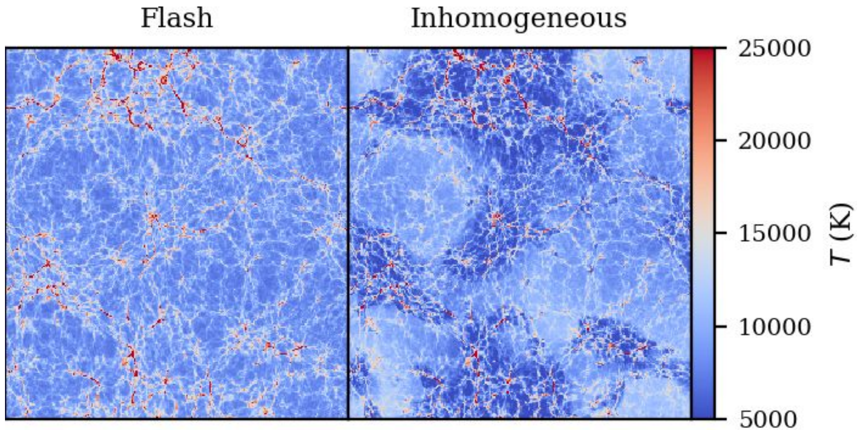


Wolfson et al. in prep.

What physics (or in our simulation) can effect these covariance matrices?

Adding temperature fluctuations effects the covariance structure

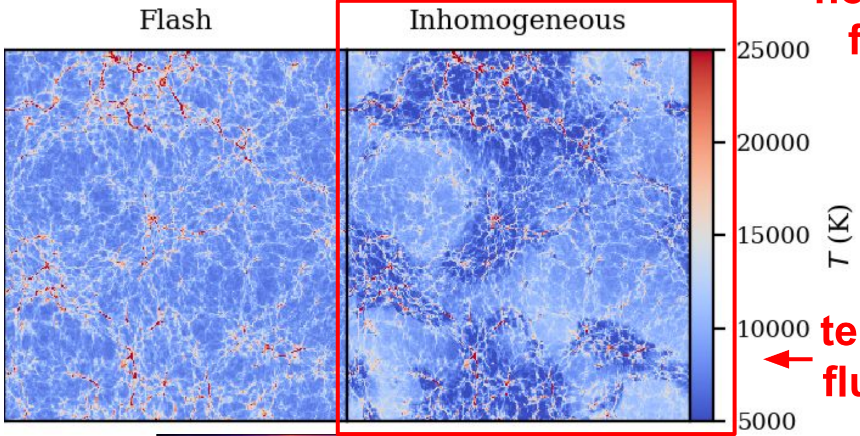
Wolfson et al. 2023

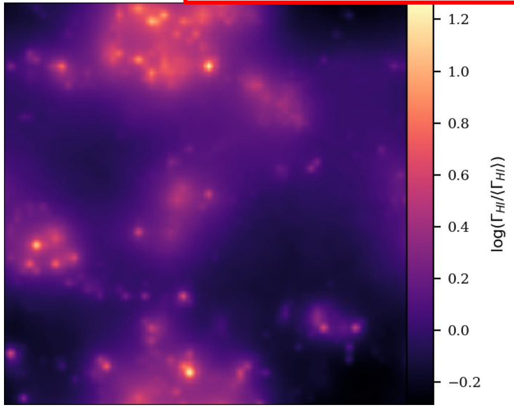


Reionization models from Oñorbe et al. 2019

Adding temperature fluctuations effects the covariance structure

Wolfson et al. 2023





λ_{mfp} = 15 cMpc

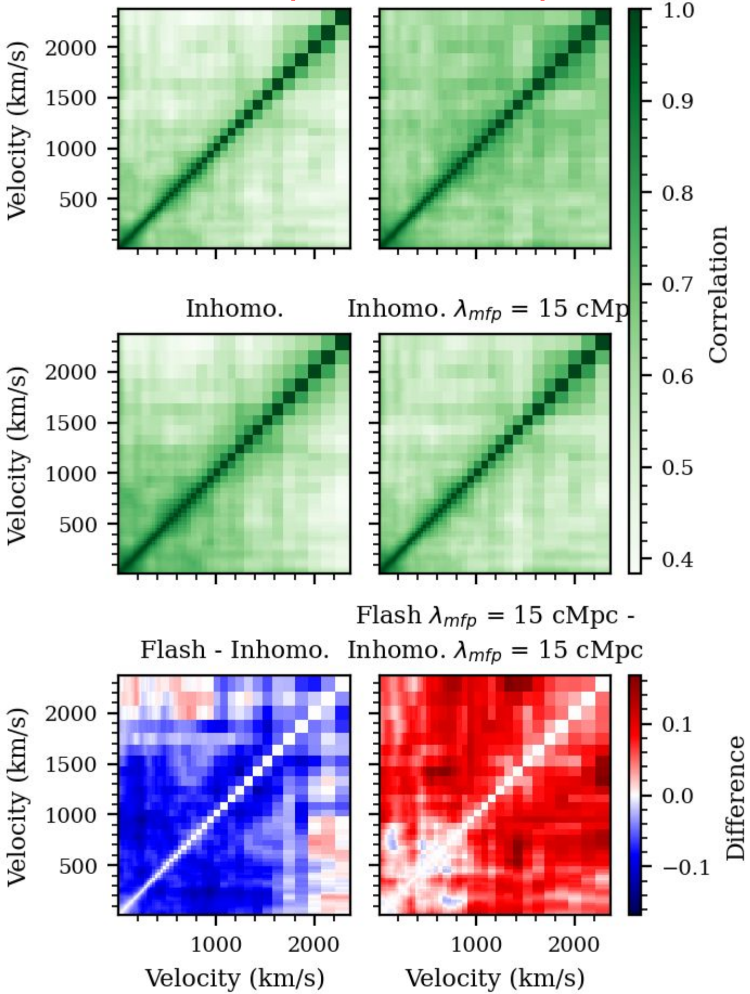
Wolfson et al. 2023

long λ_{mfp} **short λ_{mfp}** Wolfson et al. in prep.

no temperature fluctuations

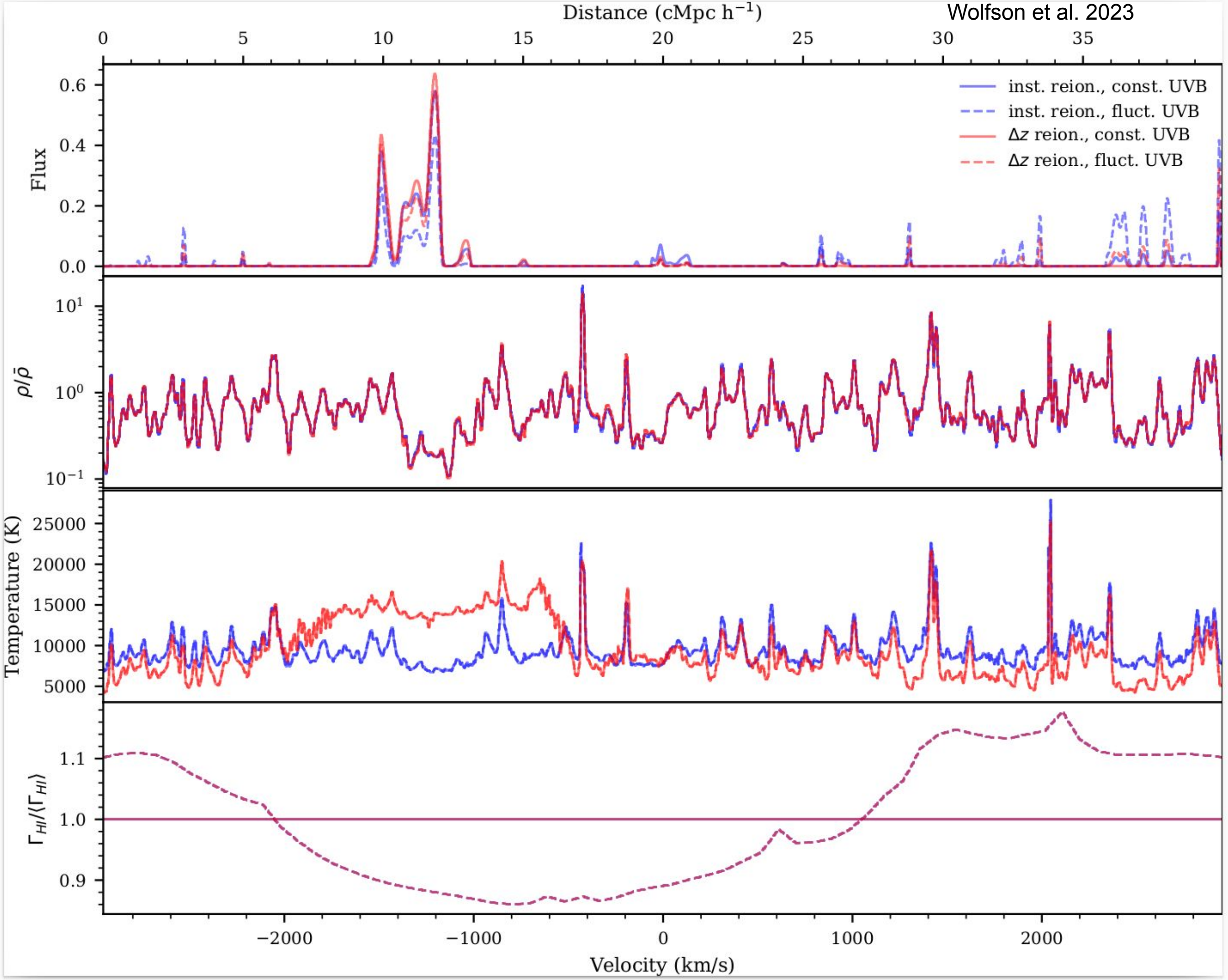
temperature fluctuations

difference



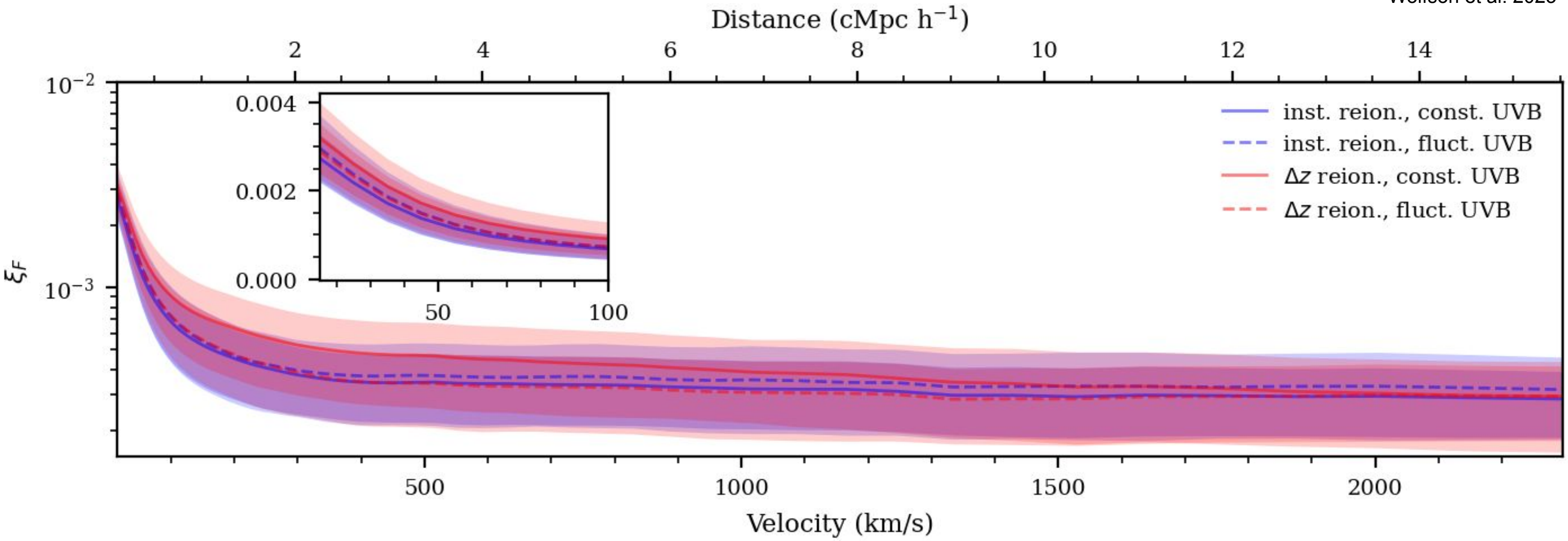
Will this significantly change measurements?

Temperature fluctuations effect on the flux:

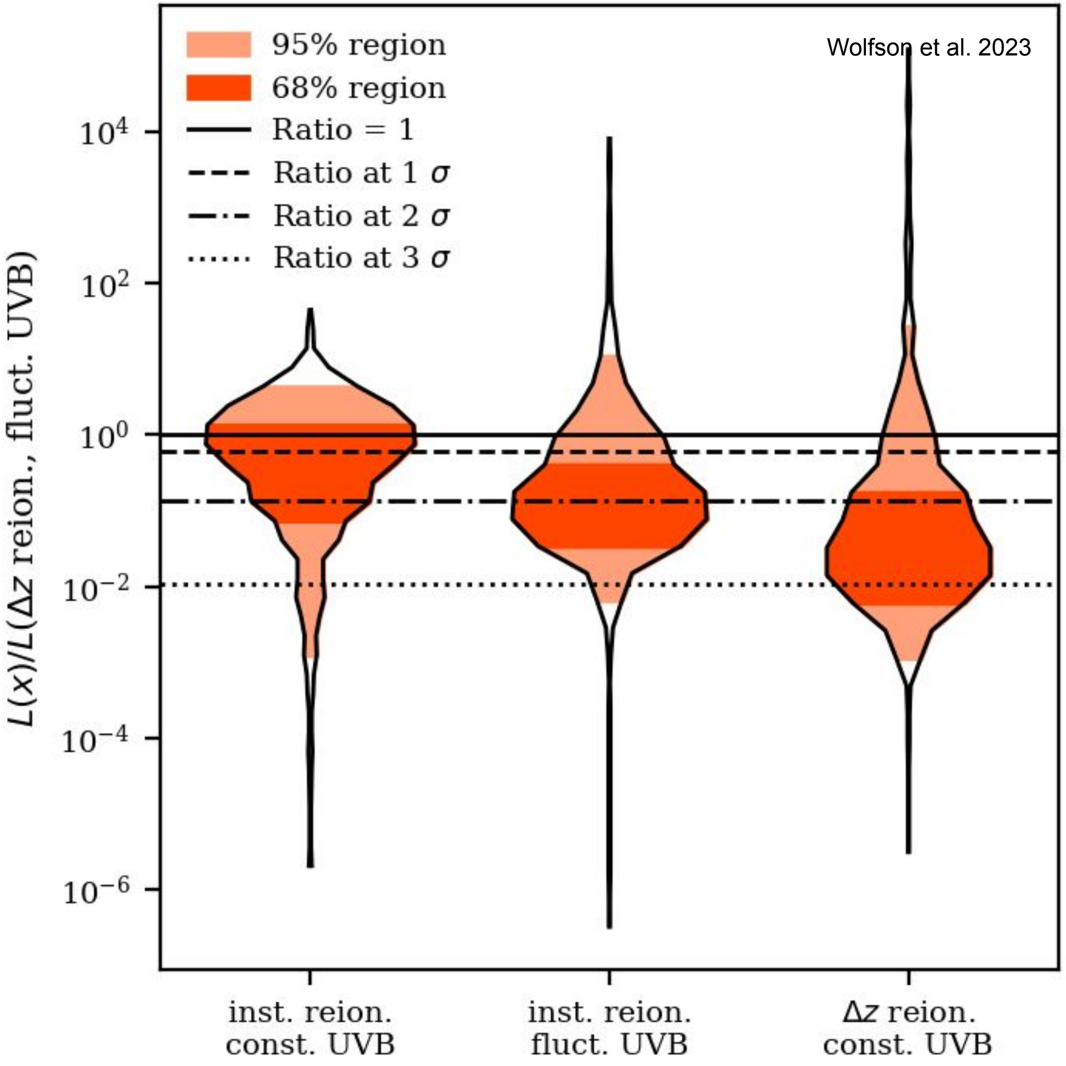


Temperature fluctuations effect on the correlation function:

Wolfson et al. 2023

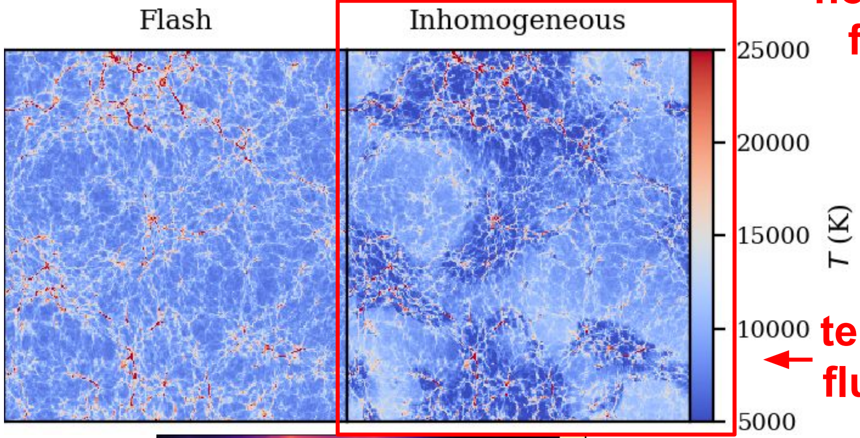


Temperature fluctuations effect on the likelihood:



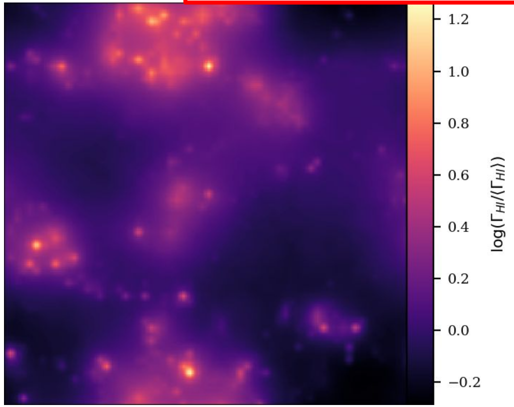
Adding temperature fluctuations effects the covariance structure

Wolfson et al. 2023



no temperature fluctuations

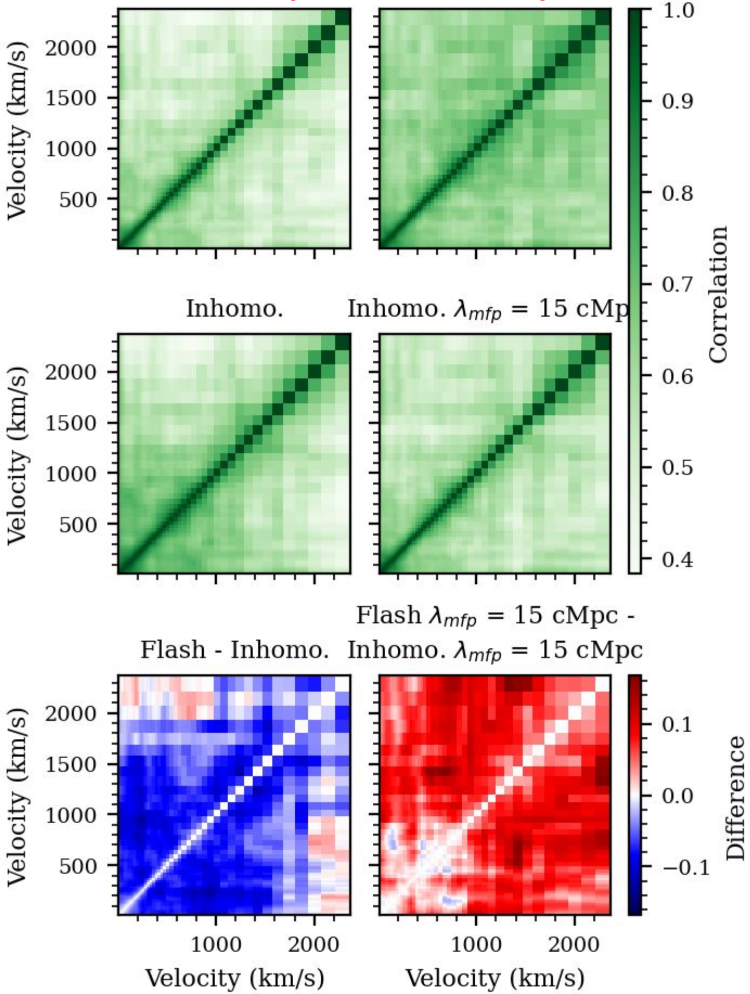
temperature fluctuations



λ_{mfp} = 15 cMpc

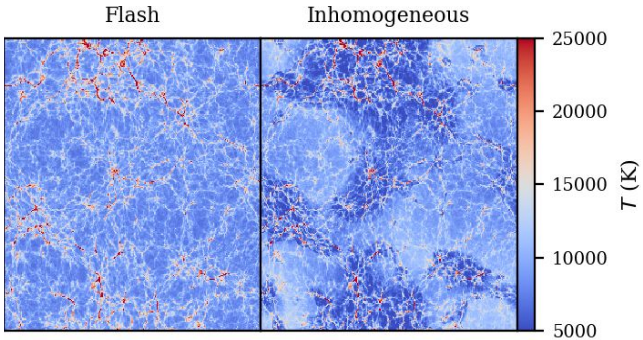
Wolfson et al. 2023

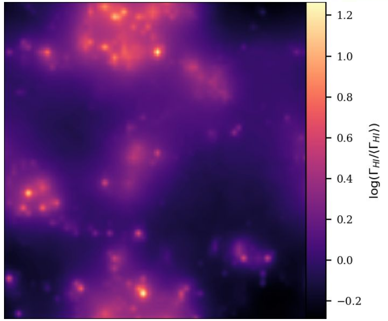
long λ_{mfp} short λ_{mfp} Wolfson et al. in prep.



difference

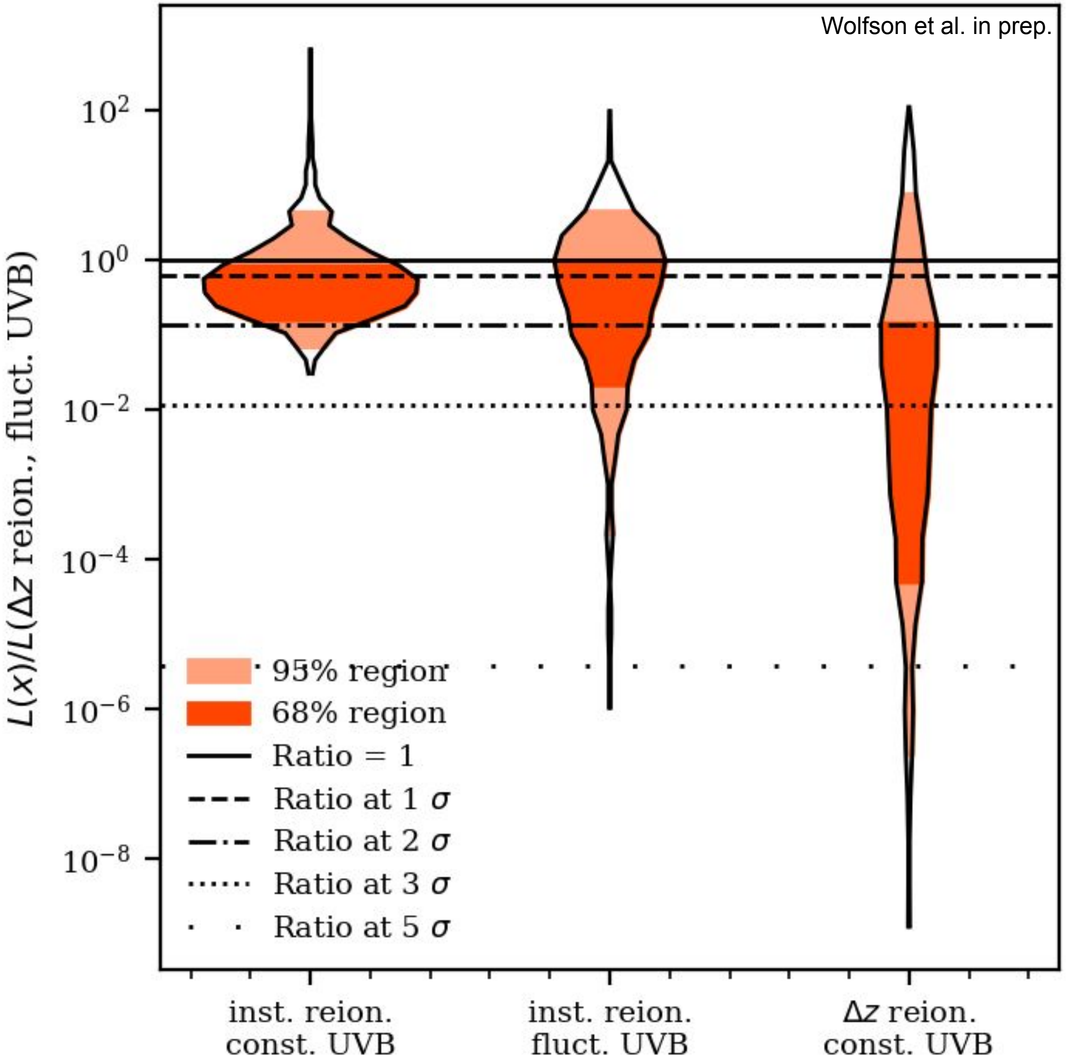
What if the models and the diagonals of the covariance are the same?



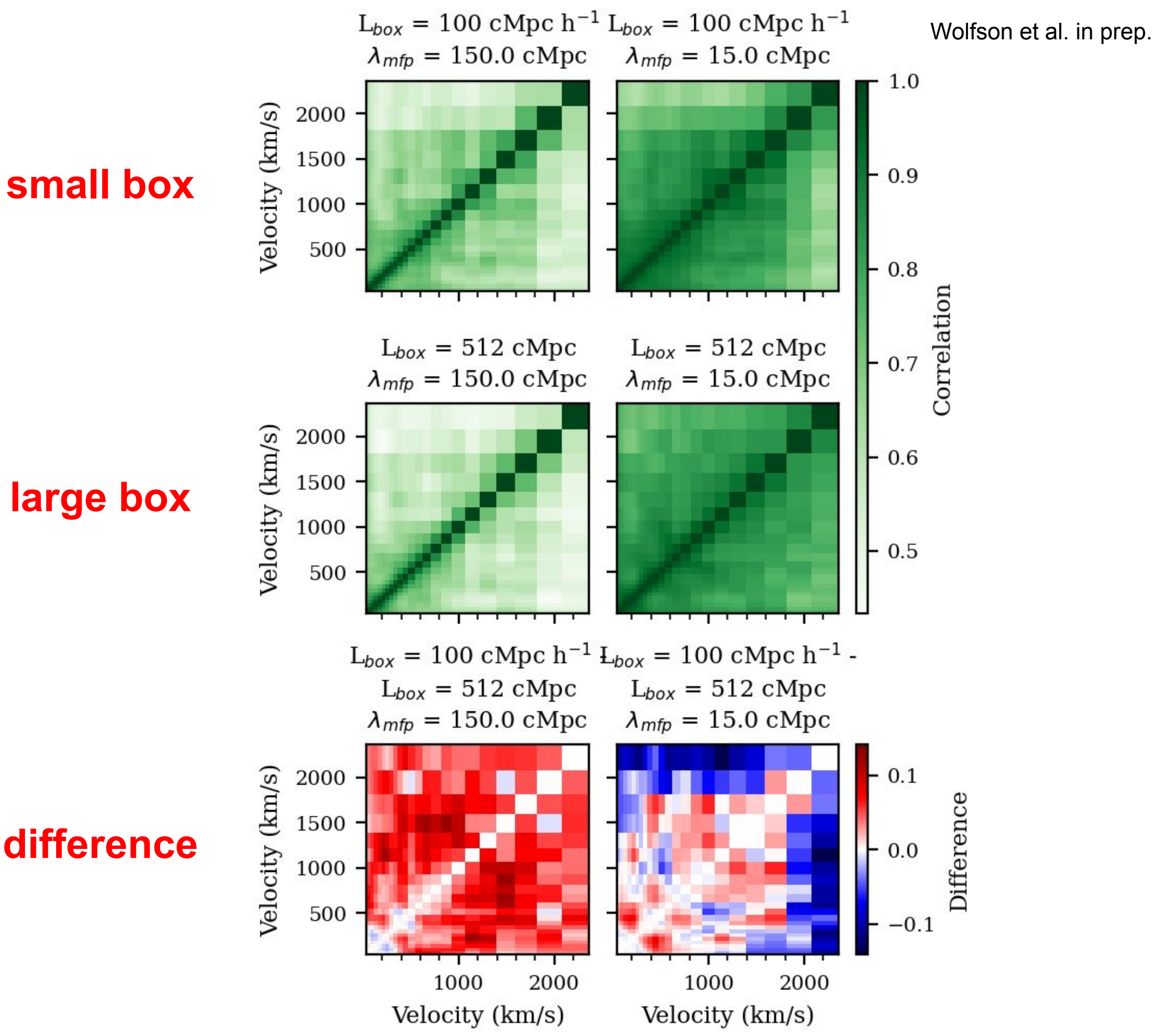


λ_{mfp} = 15 cMpc

Wolfson et al. 2023

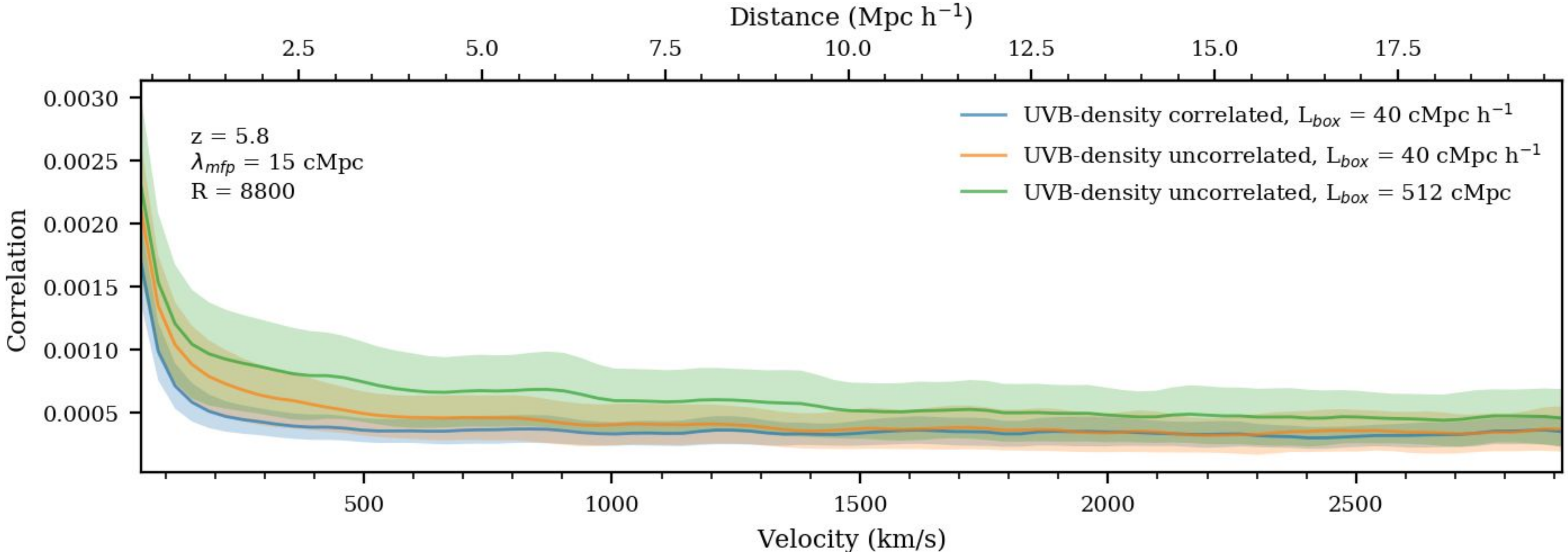


The size of the UVB box can cause similar effects on the covariance matrix:



Small UVB boxes also suppress the boost in the models :

Wolfson et al 2023.



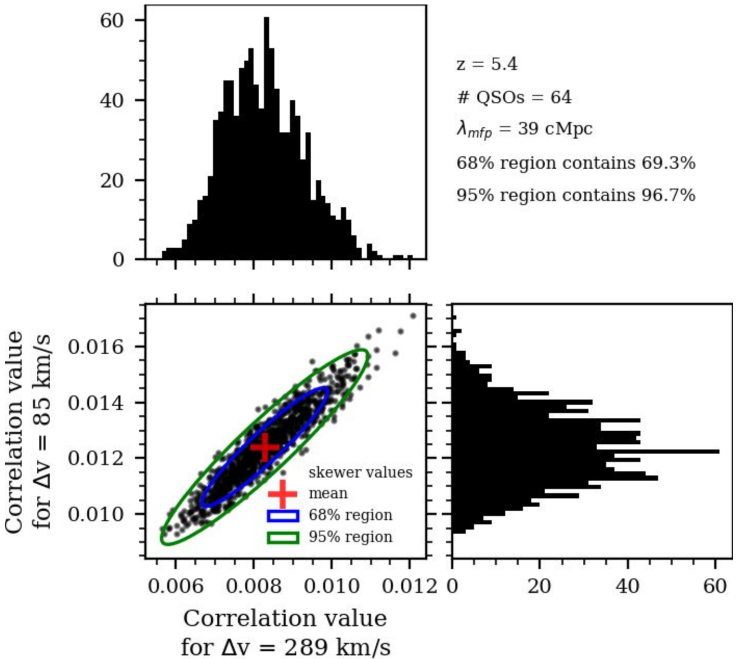
Future Directions:

- Improvements in the simulation modeling
    - Need coupled hydrodynamic cosmological simulations with radiative transfer of photons in large boxes
- Improvements in the methods
    - Mock data sets (especially at high-z) are not Gaussian distributed so using a Gaussian likelihood is wrong

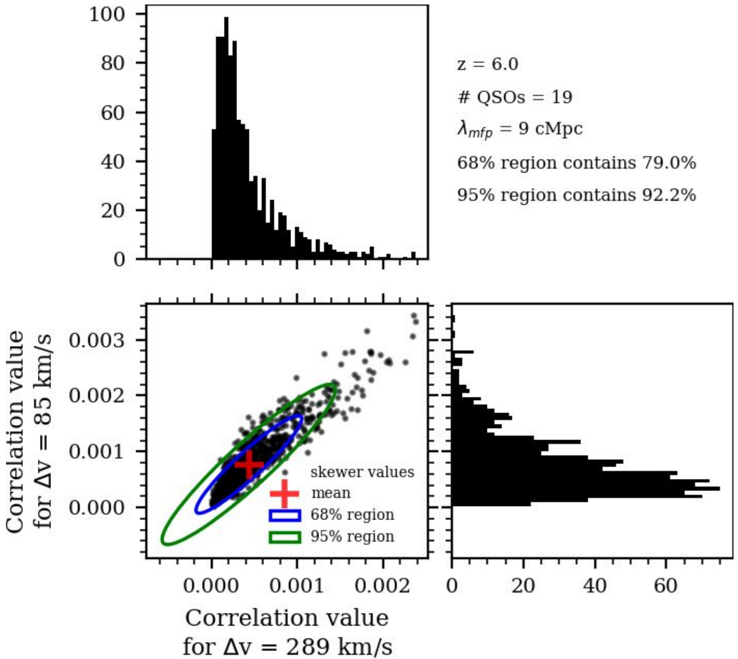
Future Directions:

- Improvements in the methods
  - Mock data sets (especially at high-z) are not Gaussian distributed so using a Gaussian likelihood is wrong

Wolfson et al. 2023



Wolfson et al. 2023



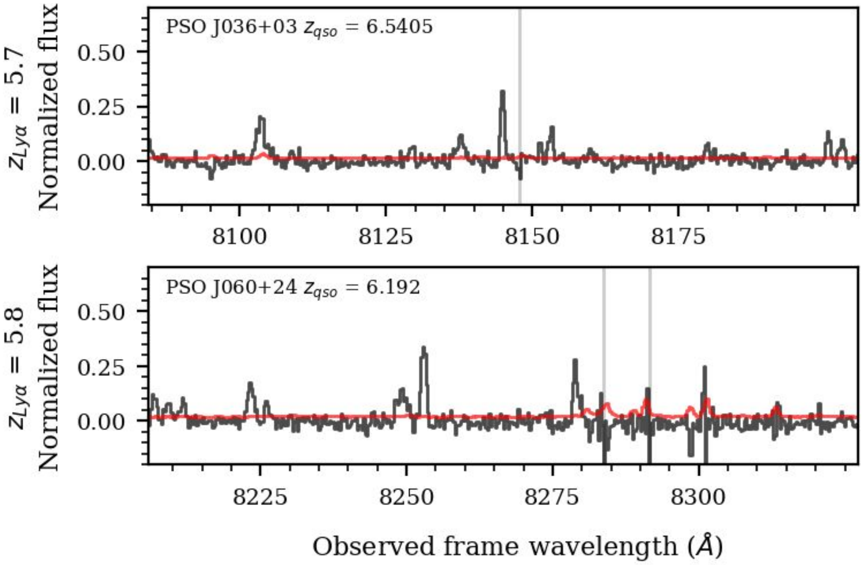
Future Directions:

- Improvements in the simulation modeling
    - Need coupled hydrodynamic cosmological simulations with radiative transfer of photons in large boxes
- Improvements in the methods
    - Mock data sets (especially at high-z) are not Gaussian distributed so using a Gaussian likelihood is wrong
    - Many sightlines at high-z are almost total absorption so computing a 2-pt statistic here may be sub-optimal (averaging with the rare transmission spikes)

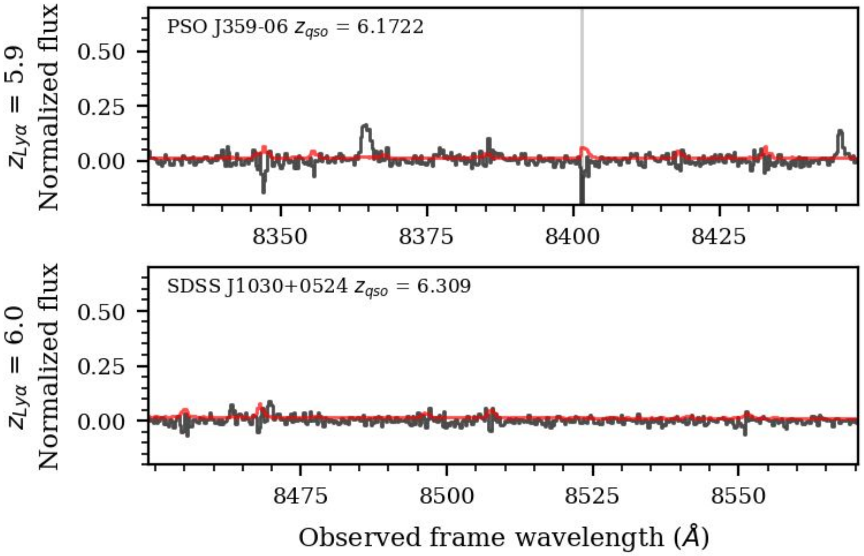
Future Directions:

- Improvements in the methods
  - Many sightlines at high-z are almost total absorption so computing a 2-pt statistic here may be sub-optimal (averaging with the rare transmission spikes)

Wolfson et al. 2023



Wolfson et al. 2023

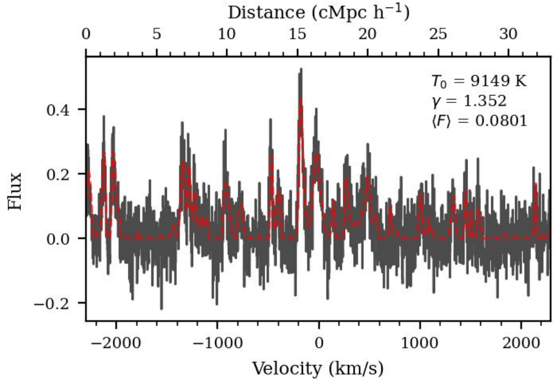


Future Directions:

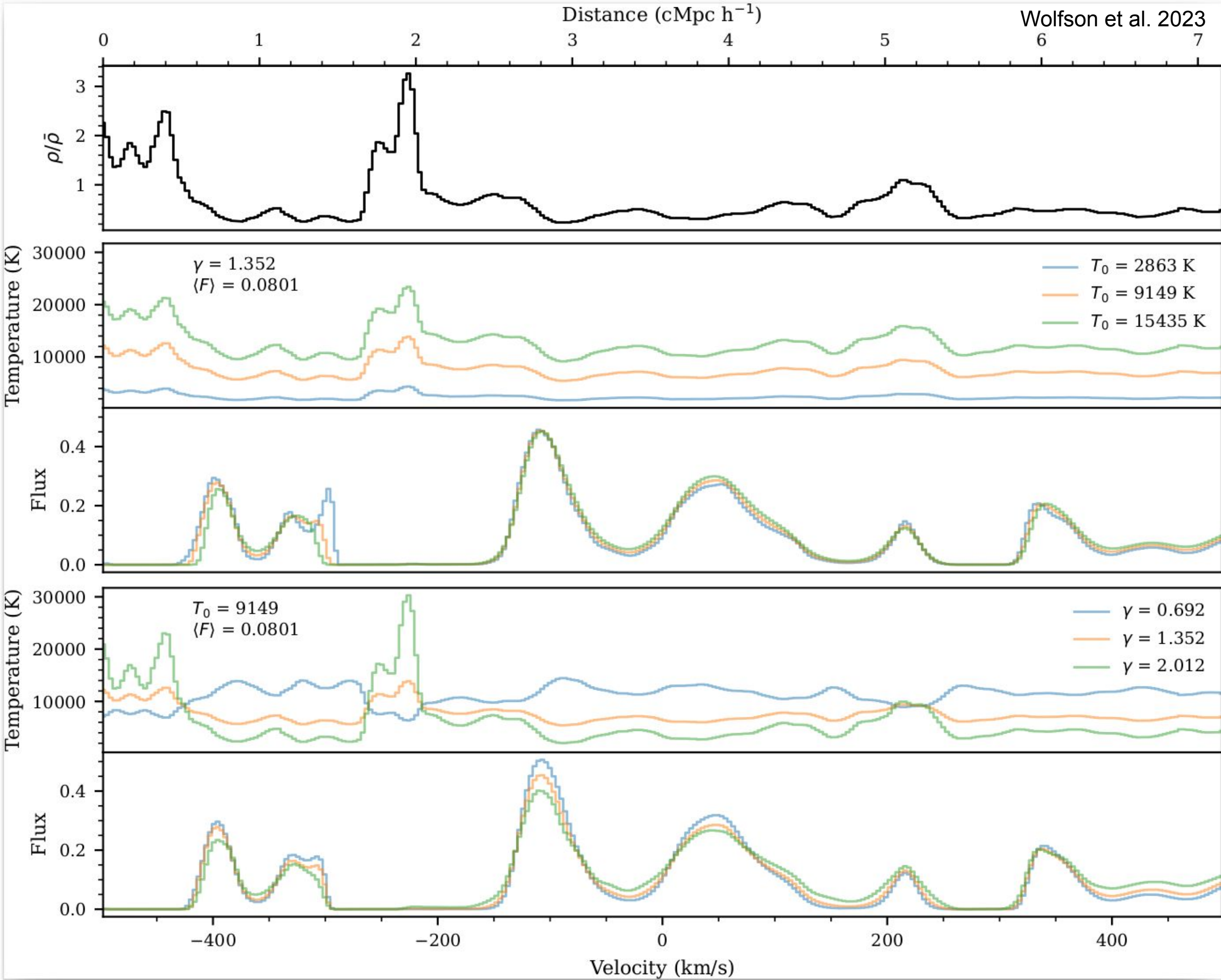
- Improvements in the simulation modeling
  - Need coupled hydrodynamic cosmological simulations with radiative transfer of photons in large boxes
- Improvements in the methods
  - Mock data sets (especially at high-z) are not Gaussian distributed so using a Gaussian likelihood is wrong
  - Many sightlines at high-z are almost total absorption so computing a 2-pt statistic here may be sub-optimal (averaging with the rare transmission spikes)
- Improvements in the data
  - High-resolution quasar spectra have access to smaller spatial scales which are sensitive to things like the thermal state of the IGM

Can you use high-resolution measurements to constrain the thermal state?

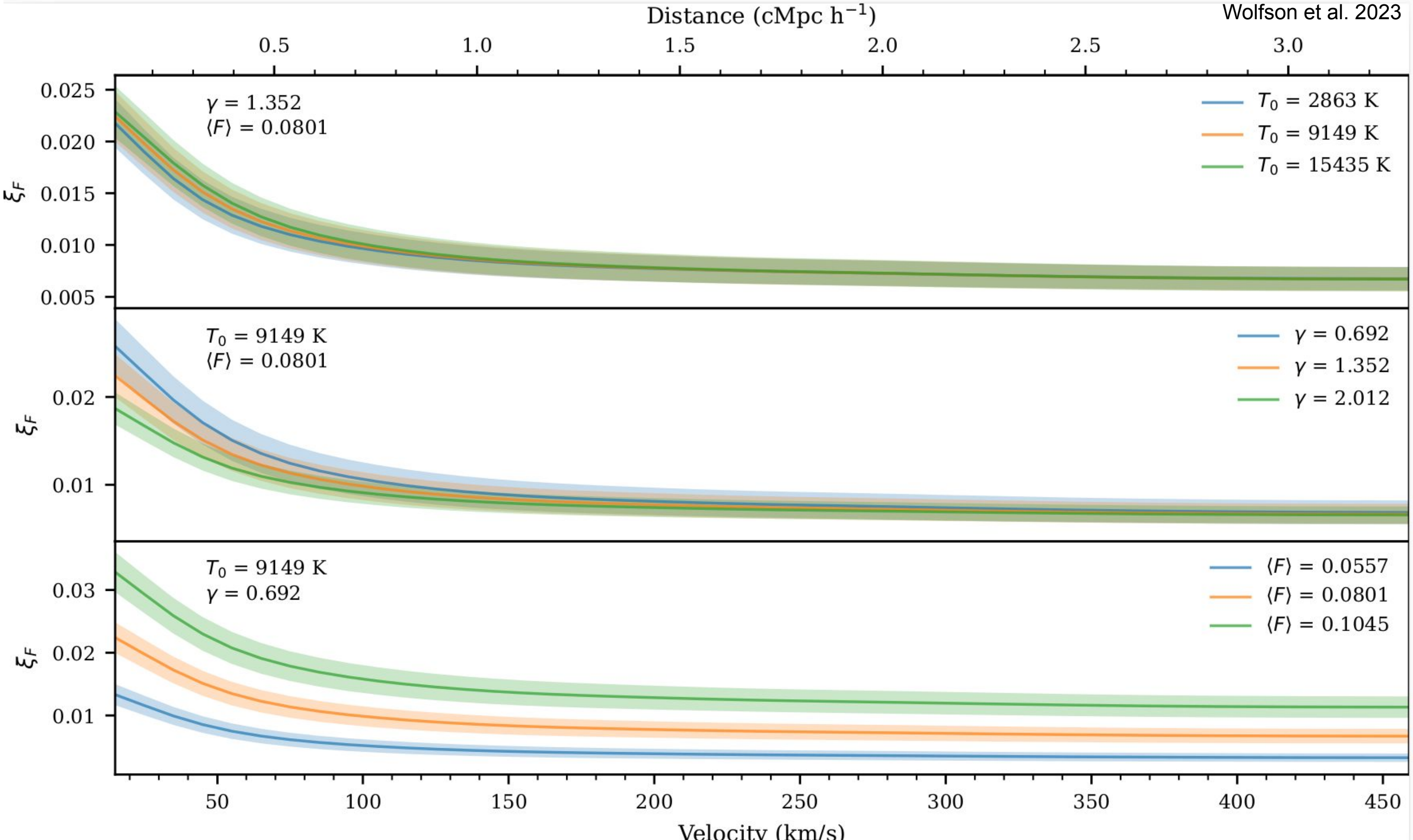
Wolfson et al. 2023



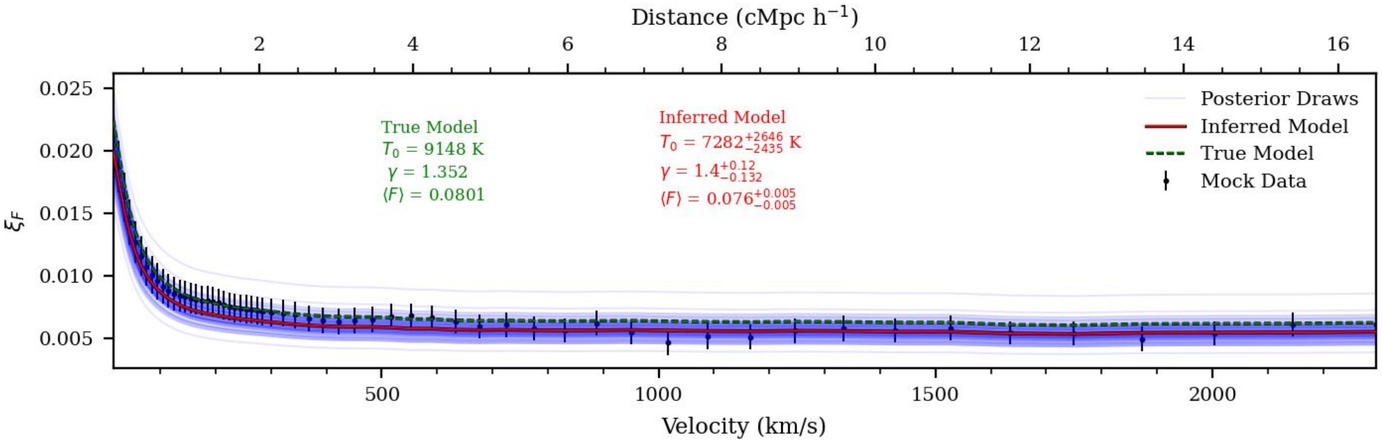
Wolfson et al. 2023

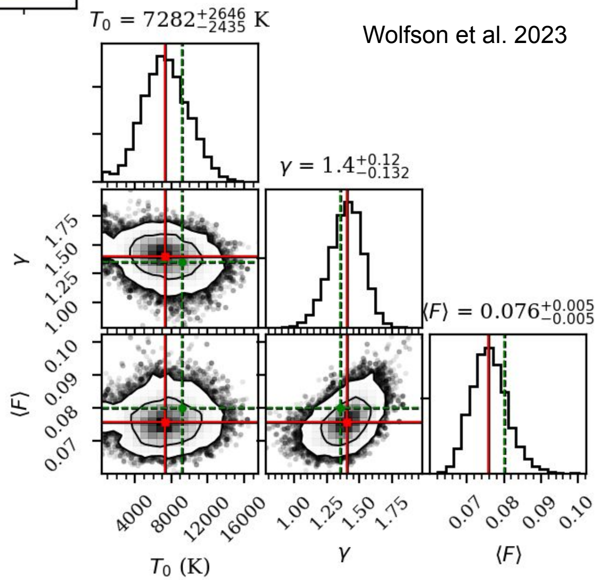


Thermal state effect on the correlation function:



Fitting a mock high-resolution data set:





Wolfson et al. 2023

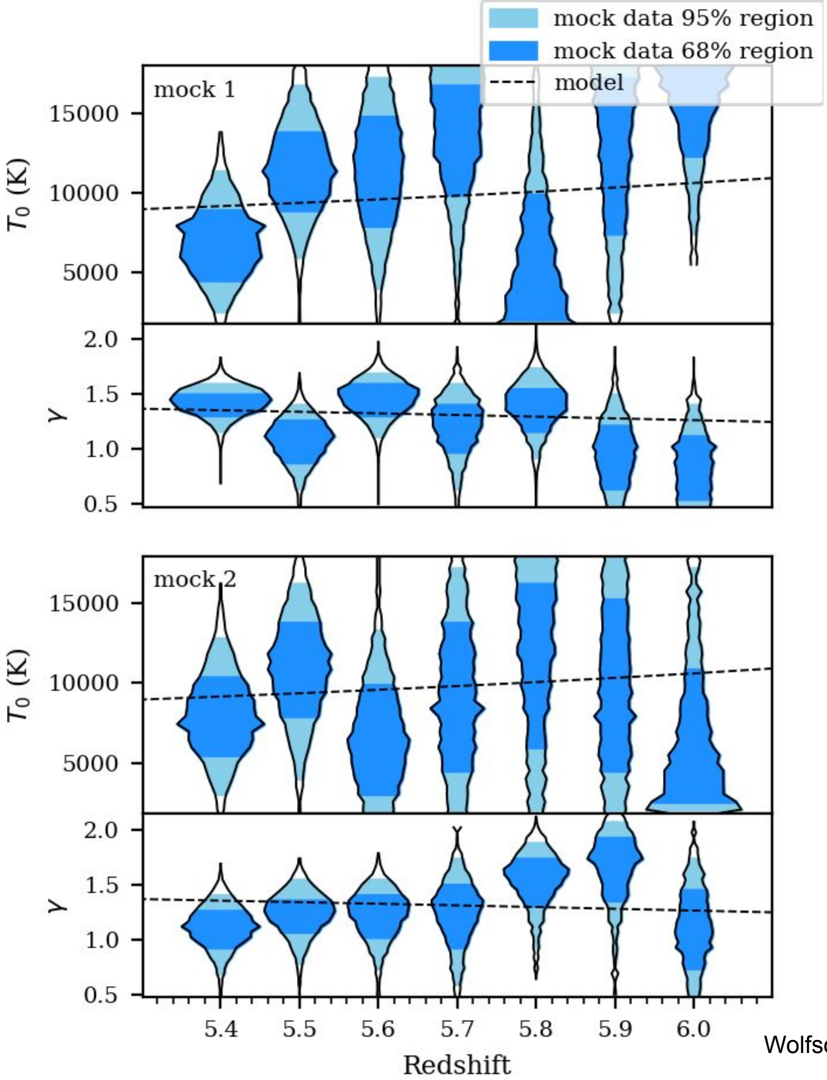
20 quasar observations

Keck Observatory archive has 56 unique z >4.5 quasars observed with HIRES

Multiple mock evolutions:

20 quasar observations

Keck Observatory archive has 56 unique z >4.5 quasars observed with HIRES



Wolfson et al. 202

Future Directions:

- Improvements in the simulation modeling
  - Need coupled hydrodynamic cosmological simulations with radiative transfer of photons in large boxes
- Improvements in the methods
  - Mock data sets (especially at high-z) are not Gaussian distributed so using a Gaussian likelihood is wrong
  - Many sightlines at high-z are almost total absorption so computing a 2-pt statistic here may be sub-optimal (averaging with the rare transmission spikes)
- Improvements in the data
  - High-resolution quasar spectra have access to smaller spatial scales which are sensitive to things like the thermal state of the IGM

Summary:

- The auto-correlation function provides a new way to competitively constrain the evolution with redshift of λ_{mfp} with existing data
- Practical considerations make the auto-correlation function a particularly useful statistic
  - Ex: not needing to model the noise or calculate the window function from DLA mask
- Measurement of λ_{mfp} from XQR-30 data is ongoing
- The covariance matrix of the auto-correlation function (4th order statistic) contains a lot of information on the high-z universe
  - **We can learn additional information about the IGM by improving our models to match the covariance fluctuations**
- Need more detailed physical simulations that explore this parameter space
  - Explore more reionization histories in sufficient box sizes