Three-dimensional cosmography of the high redshift Universe using intergalactic absorption

Collin Politsch

@capolitsch

Carnegie Mellon University

October 23, 2020

Collaborators (PhD advisors)

Larry Wasserman (CMU)

Jessi Cisewski-Kehe (UW-Madison)

Rupert Croft (CMU)

Overview of Thesis Work

- **Extrasolar planets.**
  - Modeling phase-folded transit light curves with higher-order changepoint methods
- **Eclipsing binary stars.**
  - Nonparametric modeling of phase-folded light curves
- **Spectroscopic classification and redshift estimation.**
  - Efficient spectral template generation with observational spectra
- **Supernovae.**
  - Light-curve template generation
  - Nonparametric estimation of observable parameters
- **Intergalactic medium (via the Lyman-α forest).**
  - Nonparametric estimation of quasar continua
  - One-dimensional reconstruction of the intergalactic medium
  - Three-dimensional reconstruction of the intergalactic medium

Overview of Thesis Work

- **Extrasolar planets.**
  - Modeling phase-folded transit light curves with higher-order changepoint methods
- **Eclipsing binary stars.**
  - Nonparametric modeling of phase-folded light curves
- **Spectroscopic classification and redshift estimation.**
  - Efficient spectral template generation with observational spectra
- **Supernovae.**
  - Light-curve template generation
  - Nonparametric estimation of observable parameters

Chapters 2-3 of Thesis
Politsch et al. (2020a,b)
MNRAS

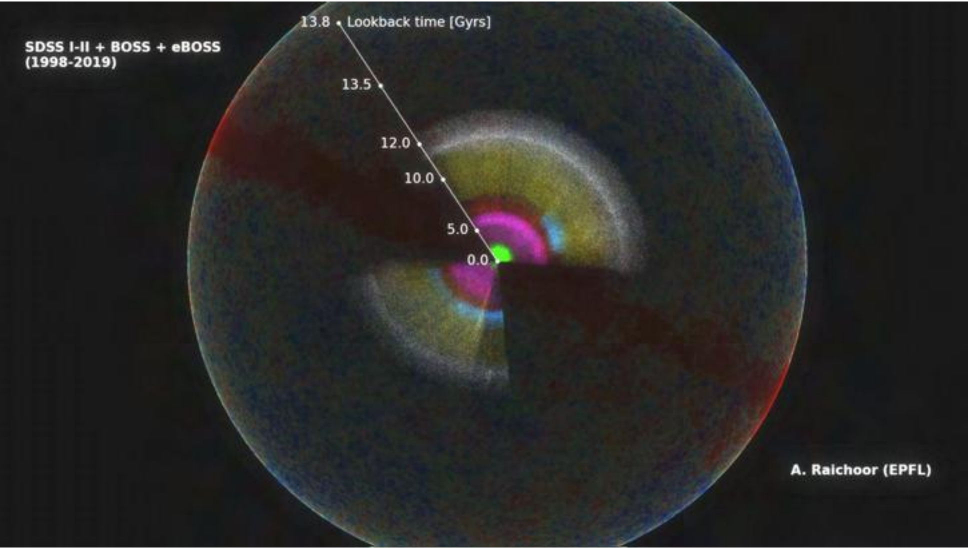
- **Intergalactic medium (via the Lyman-α forest).**
  - Nonparametric estimation of quasar continua
  - One-dimensional reconstruction of the intergalactic medium
  - Three-dimensional reconstruction of the intergalactic medium

Focus of this talk

Talk outline

1. The Lyman-α forest
   - Background
   - The Sloan Digital Sky Survey
   - Summary of work & results
2. One-dimensional mapping the intergalactic medium
   - Denoising observational spectra
   - Estimating the underlying density field
   - Uncertainty quantification
3. Three-dimensional mapping the intergalactic medium
   - Spatial model
   - Distributed computing
   - Optimized absorption field reconstruction
   - Uncertainty quantification
   - High-significance candidates for galaxy protoclusters and voids

The Sloan Digital Sky Survey



The cosmological matter distribution

The **intergalactic medium** is a highly diffuse gaseous medium that pervades the volume of intergalactic space and contains a majority of the baryonic matter in the Universe.

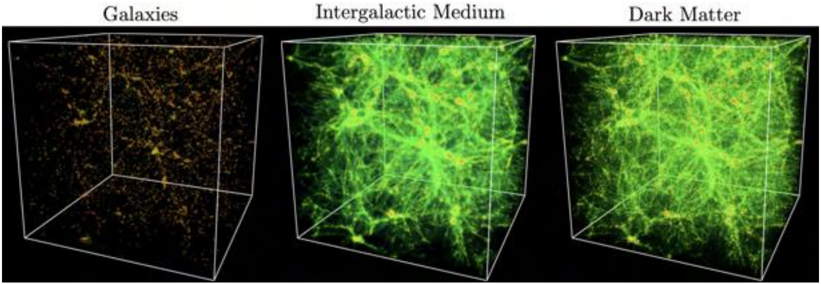
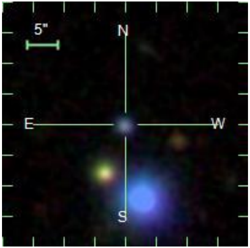


Image credit: R. Cen and J. P. Ostriker

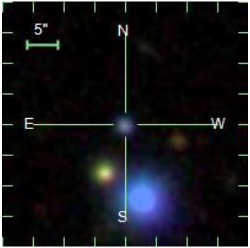
Absorption spectroscopy with quasar backlights

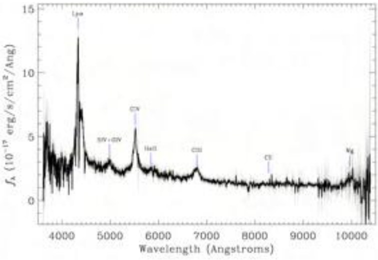
Absorption spectroscopy: Spectroscopic techniques that measure the absorption of light, as a function of wavelength, due to its interaction with a medium.



Absorption spectroscopy with quasar backlights

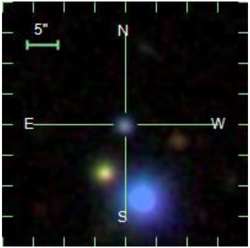
Absorption spectroscopy: Spectroscopic techniques that measure the absorption of light, as a function of wavelength, due to its interaction with a medium.

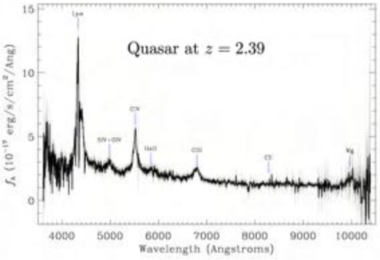




Absorption spectroscopy with quasar backlights

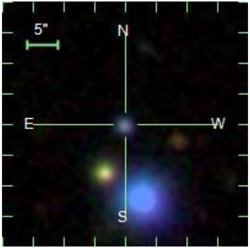
Absorption spectroscopy: Spectroscopic techniques that measure the absorption of light, as a function of wavelength, due to its interaction with a medium.

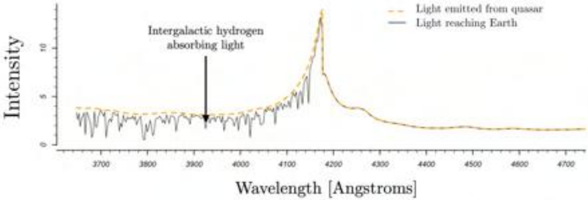




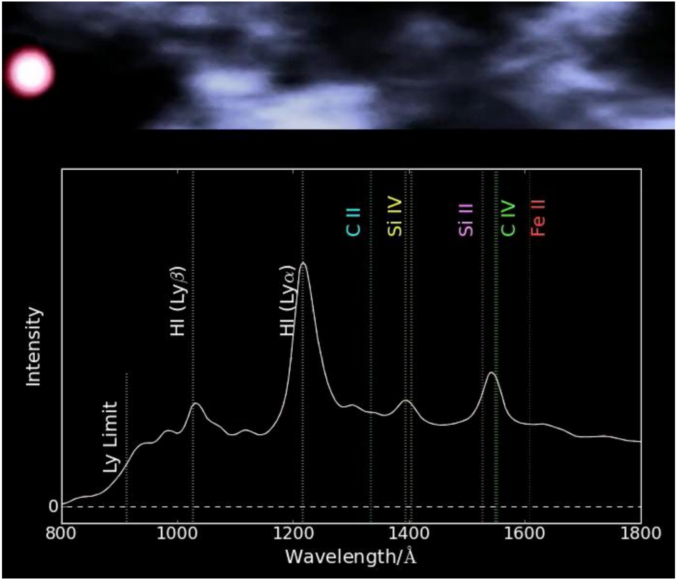
Absorption spectroscopy with quasar backlights

Absorption spectroscopy: Spectroscopic techniques that measure the absorption of light, as a function of wavelength, due to its interaction with a medium.



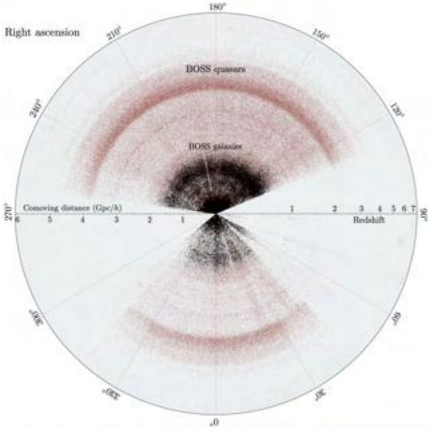


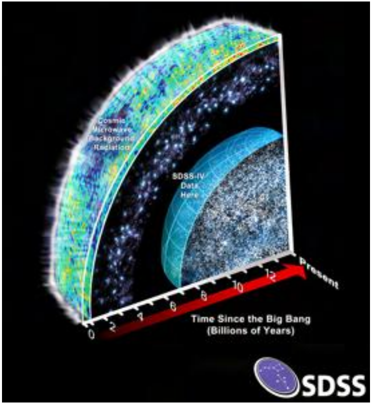
The Lyman-α forest absorption process



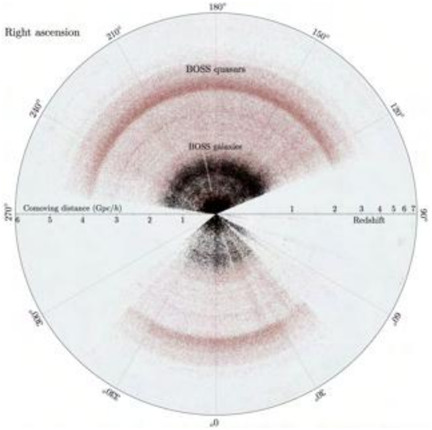
Video credit: Andrew Pontzen

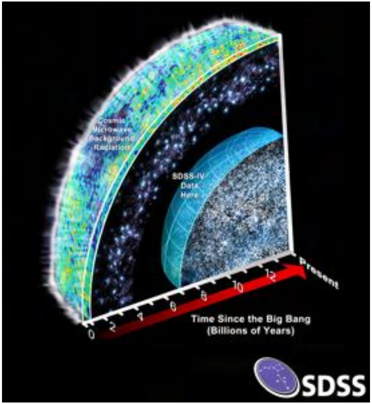
The Baryon Oscillation Spectroscopic Survey





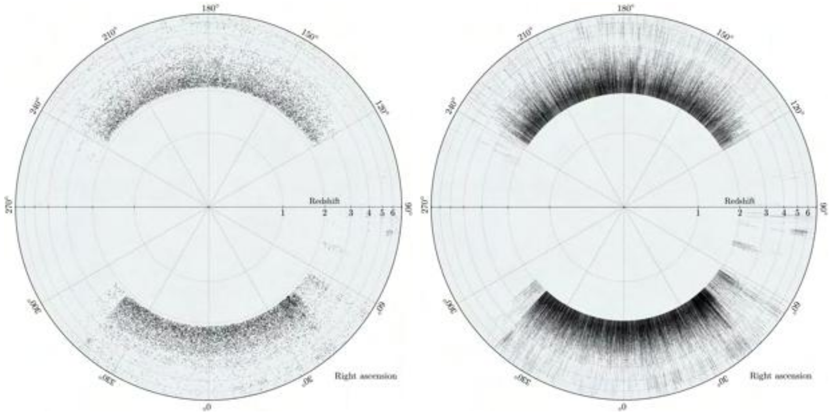
The Baryon Oscillation Spectroscopic Survey





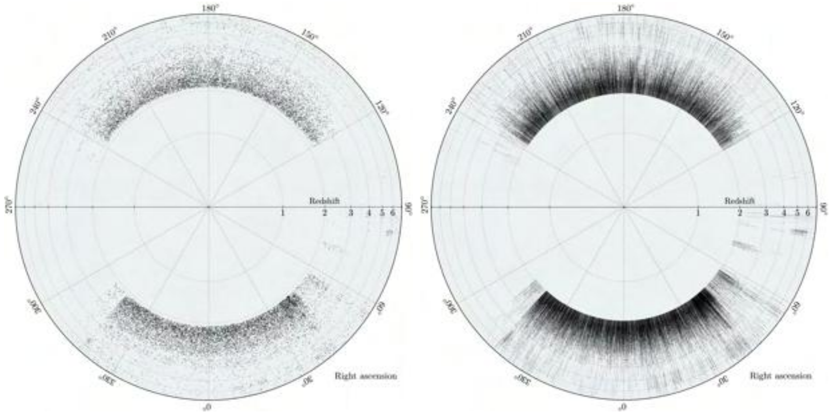
~160,000 usable
Lyman-α quasar spectra

High-level summary of work



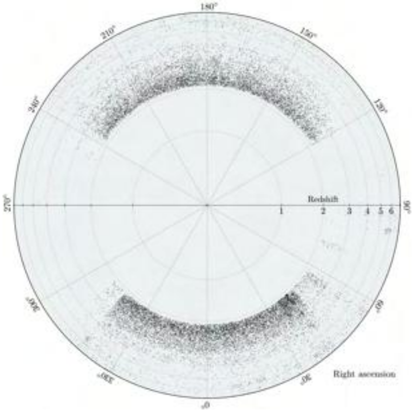
High-level summary of work

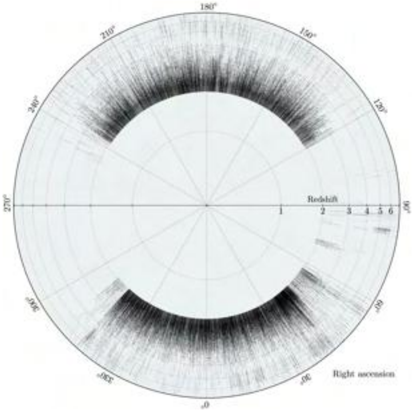
1. Infer the relative density of the IGM along each 1D quasar sightline



High-level summary of work

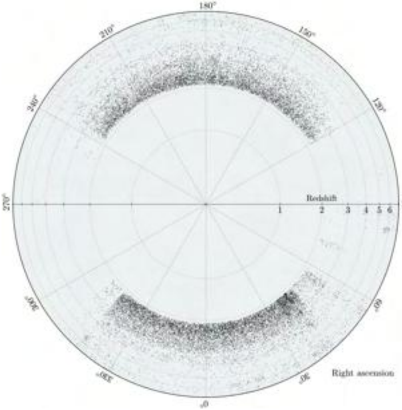
1. Infer the relative density of the IGM along each 1D quasar sightline
2. Pool the sightlines and reconstruct a full 3D large-scale structure map

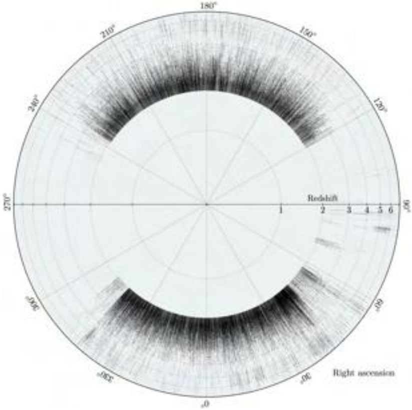




High-level summary of work

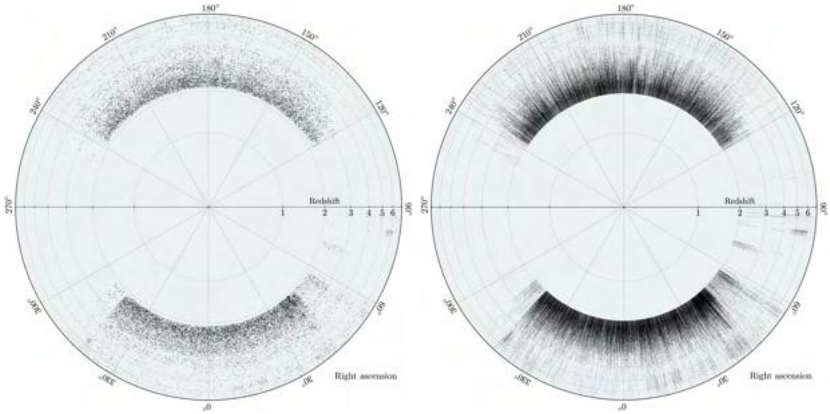
1. Infer the relative density of the IGM along each 1D quasar sightline
2. Pool the sightlines and reconstruct a full 3D large-scale structure map
3. Quantify the total statistical uncertainty in the full 3D map





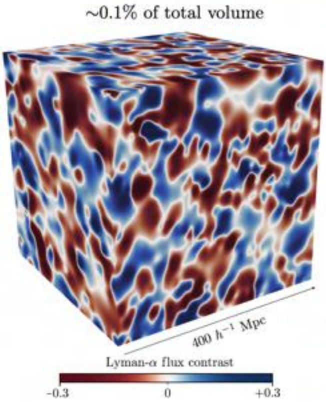
High-level summary of work

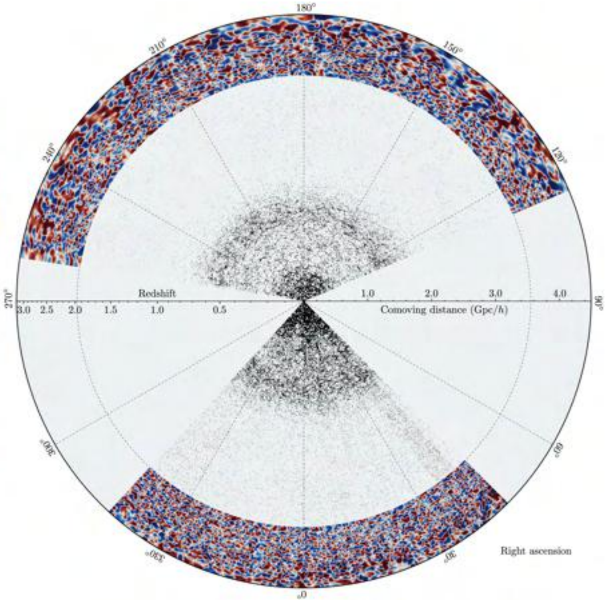
1. Infer the relative density of the IGM along each 1D quasar sightline
2. Pool the sightlines and reconstruct a full 3D large-scale structure map
3. Quantify the total statistical uncertainty in the full 3D map
4. Identify statistically significant candidates for galaxy protoclusters and cosmic voids



Summary of results: Largest volume LSS map of the Universe to date

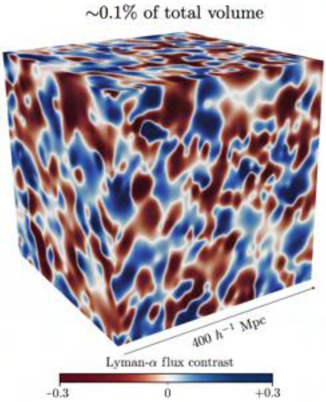
Summary of results: Largest volume LSS map of the Universe to date

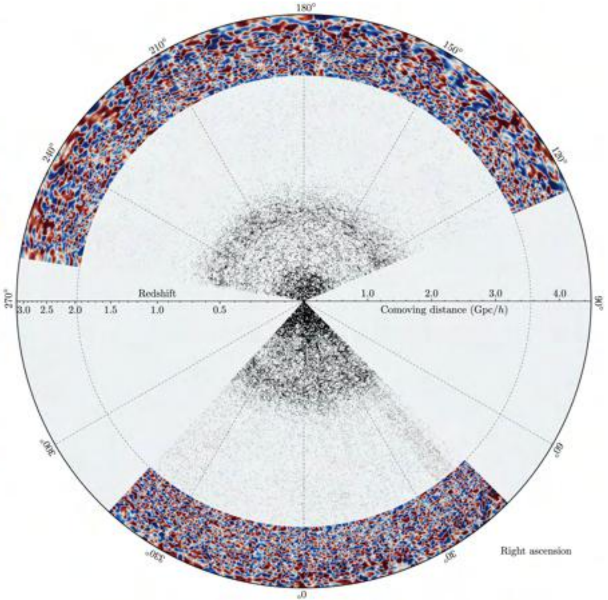




Summary of results: Largest volume LSS map of the Universe to date

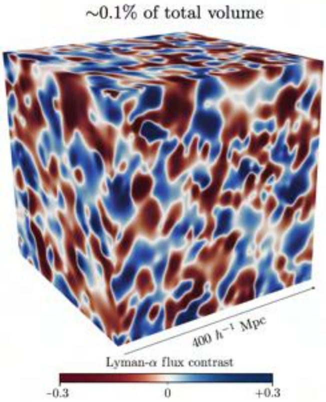
- 47 h^{-3} Gpc^3 = 154 Gpc^3 (Planck CMB)

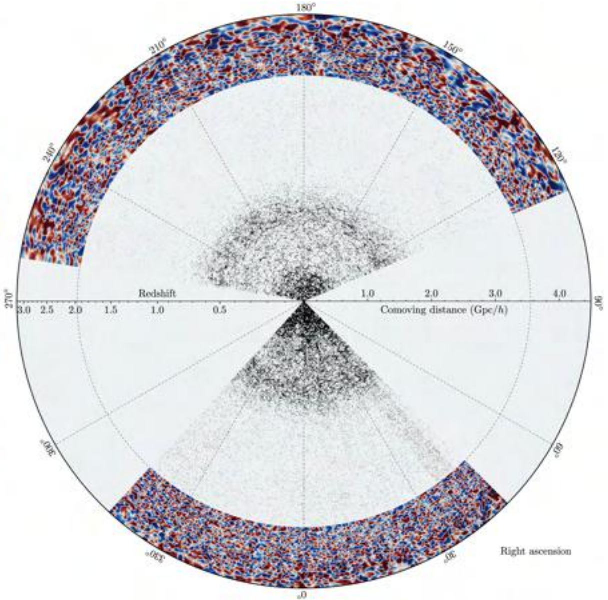




Summary of results: Largest volume LSS map of the Universe to date

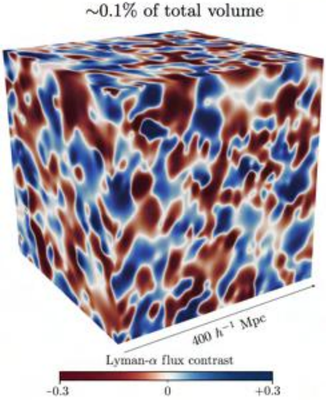
- 47 h^{-3} Gpc^3 = 154 Gpc^3 (Planck CMB)
- 10,332 deg^2 footprint (25% sky coverage)

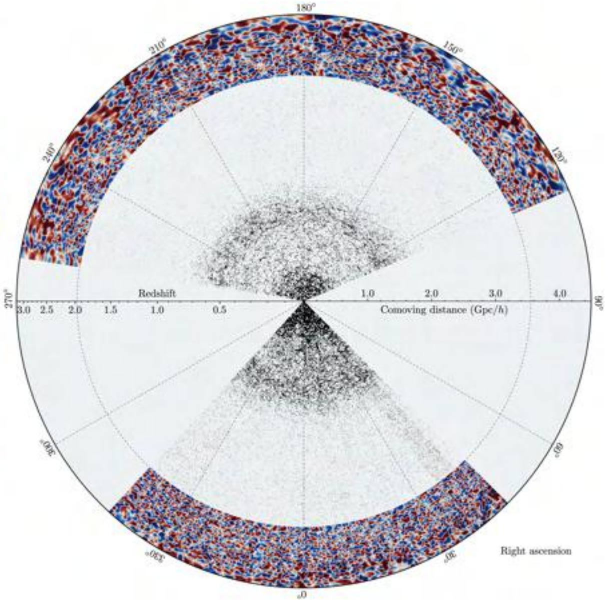




Summary of results: Largest volume LSS map of the Universe to date

- 47 h^{-3} Gpc^3 = 154 Gpc^3 (Planck CMB)
- 10,332 deg^2 footprint (25% sky coverage)
- Redshift range $1.98 \leq z \leq 3.15$

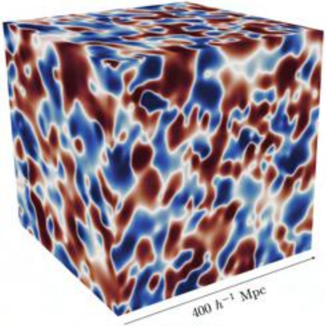


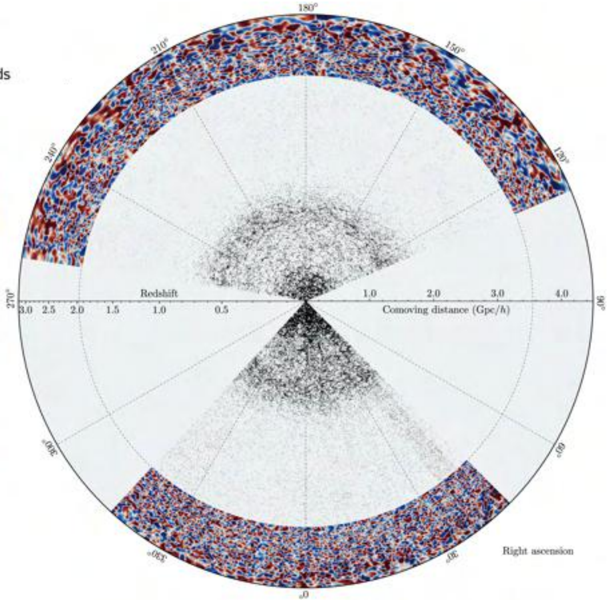


Summary of results: Largest volume LSS map of the Universe to date

- 47 h^{-3} Gpc^3 = 154 Gpc^3 (Planck CMB)
- 10,332 deg^2 footprint (25% sky coverage)
- Redshift range $1.98 \leq z \leq 3.15$
- ~180,000 candidates for protoclusters/voids

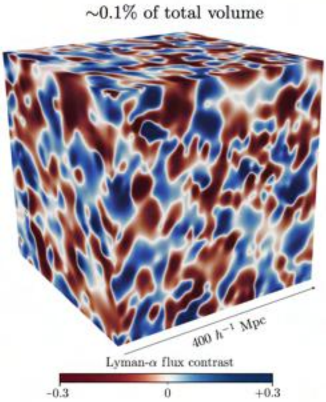
~0.1% of total volume

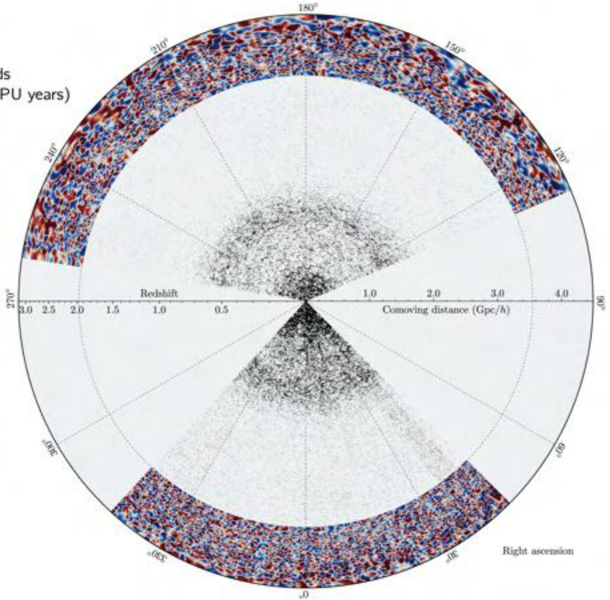




Summary of results: Largest volume LSS map of the Universe to date

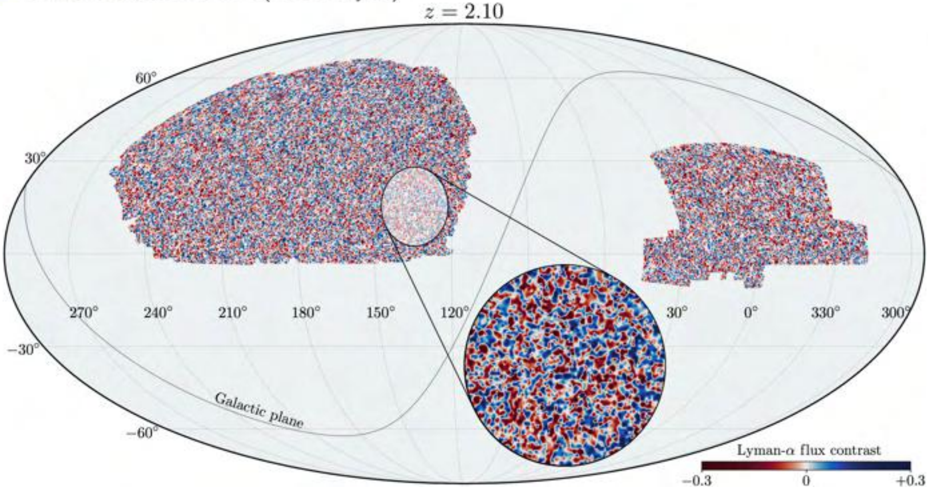
- 47 h^{-3} Gpc^3 = 154 Gpc^3 (Planck CMB)
- 10,332 deg^2 footprint (25% sky coverage)
- Redshift range $1.98 \leq z \leq 3.15$
- ~180,000 candidates for protoclusters/voids
- ~3 million distributed CPU hours (~342 CPU years)





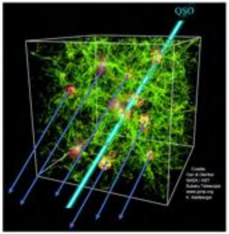
Summary of results: Largest volume LSS map of the Universe to date

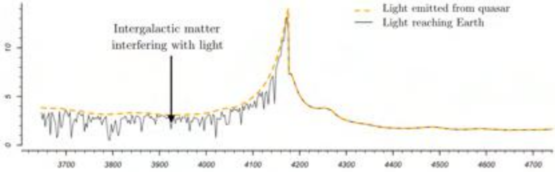
- 47 h^{-3} Gpc^3 = 154 Gpc^3 (Planck CMB)
- 10,332 deg^2 footprint (25% sky coverage)
- Redshift range $1.98 \lesssim z \lesssim 3.15$
- $\sim$180,000 candidates for protoclusters/voids
- $\sim$3 million distributed CPU hours ($\sim$342 CPU years)



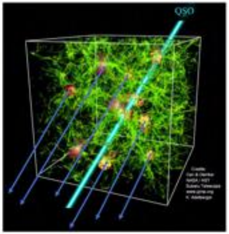
Outline

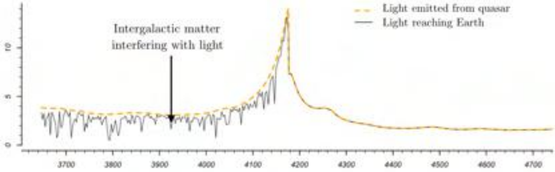
One-dimensional mapping the intergalactic medium





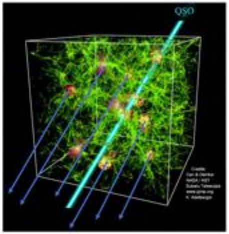
One-dimensional mapping the intergalactic medium

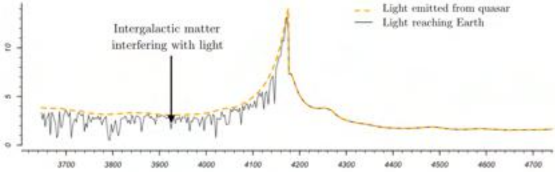




$$f(\lambda_i) = f_0(\lambda_i) + \epsilon_i$$

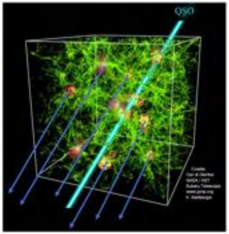
One-dimensional mapping the intergalactic medium

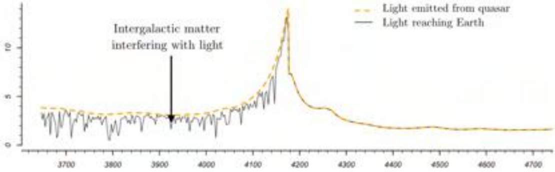




$$
\begin{aligned}
f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i
\end{aligned}
$$

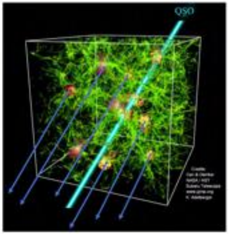
One-dimensional mapping the intergalactic medium

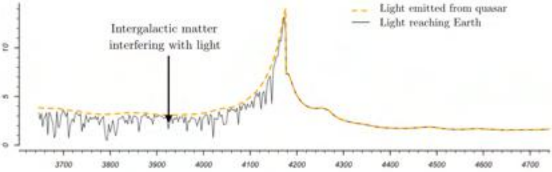




$$
\begin{aligned}
f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot \overline{F}(\lambda_i) \cdot (1 + \delta_F(\lambda_i)) + \epsilon_i
\end{aligned}
$$

One-dimensional mapping the intergalactic medium



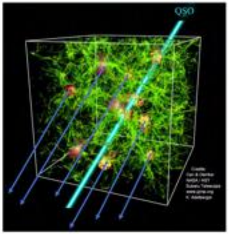


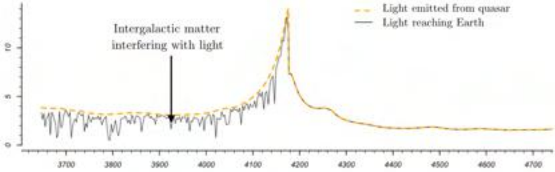
$$
\begin{aligned}
f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot \overline{F}(\lambda_i) \cdot (1 + \delta_F(\lambda_i)) + \epsilon_i
\end{aligned}
$$

Lyman-α flux contrast:

$$\delta_F(z) = \frac{F(z) - \overline{F}(z)}{\overline{F}(z)}$$

One-dimensional mapping the intergalactic medium





$$
\begin{aligned}
f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot \overline{F}(\lambda_i) \cdot (1 + \delta_F(\lambda_i)) + \epsilon_i
\end{aligned}
$$

Lyman-α flux contrast:

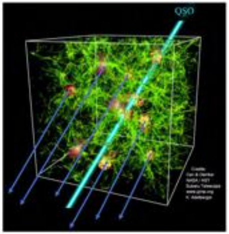
$$\delta_F(z) = \frac{F(z) - \overline{F}(z)}{\overline{F}(z)}$$

Estimator:

$$\widehat{\delta}_F(z) = \frac{\widehat{f}_0(z) - \widehat{m}(z)}{\widehat{m}(z)}$$

where $\quad m(z) = C(z) \cdot \overline{F}(z)$

One-dimensional mapping the intergalactic medium



$$
\begin{aligned}
f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot \overline{F}(\lambda_i) \cdot (1 + \delta_F(\lambda_i)) + \epsilon_i
\end{aligned}
$$

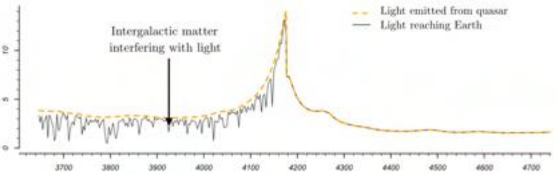
Lyman-α flux contrast:

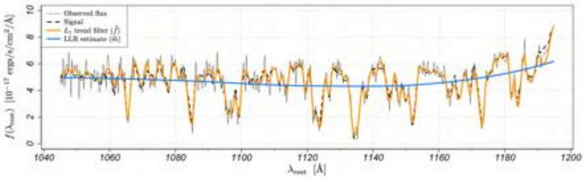
$$\delta_F(z) = \frac{F(z) - \overline{F}(z)}{\overline{F}(z)}$$

Estimator:

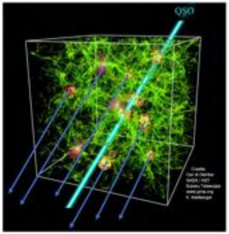
$$\widehat{\delta}_F(z) = \frac{\widehat{f}_0(z) - \widehat{m}(z)}{\widehat{m}(z)}$$

where $\quad m(z) = C(z) \cdot \overline{F}(z)$





One-dimensional mapping the intergalactic medium



$$
\begin{aligned}
f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot \overline{F}(\lambda_i) \cdot (1 + \delta_F(\lambda_i)) + \epsilon_i
\end{aligned}
$$

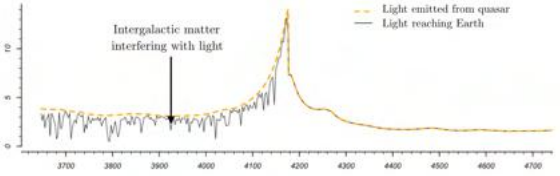
Lyman-α flux contrast:

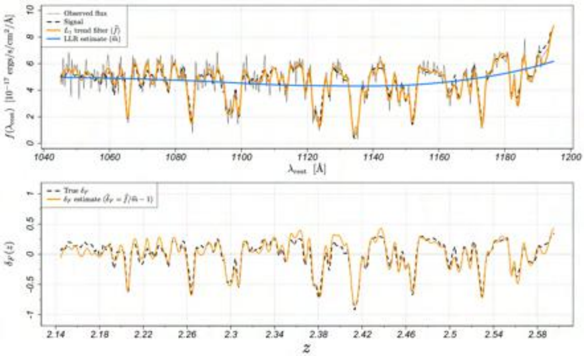
$$\delta_F(z) = \frac{F(z) - \overline{F}(z)}{\overline{F}(z)}$$

Estimator:

$$\widehat{\delta}_F(z) = \frac{\widehat{f_0}(z) - \widehat{m}(z)}{\widehat{m}(z)}$$

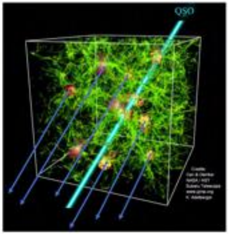
where $\quad m(z) = C(z) \cdot \overline{F}(z)$





Trend filtering: Tibshirani (2014)

One-dimensional mapping the intergalactic medium



$$\begin{aligned} f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\ &= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i \\ &= C(\lambda_i) \cdot \overline{F}(\lambda_i) \cdot (1 + \delta_F(\lambda_i)) + \epsilon_i \end{aligned}$$

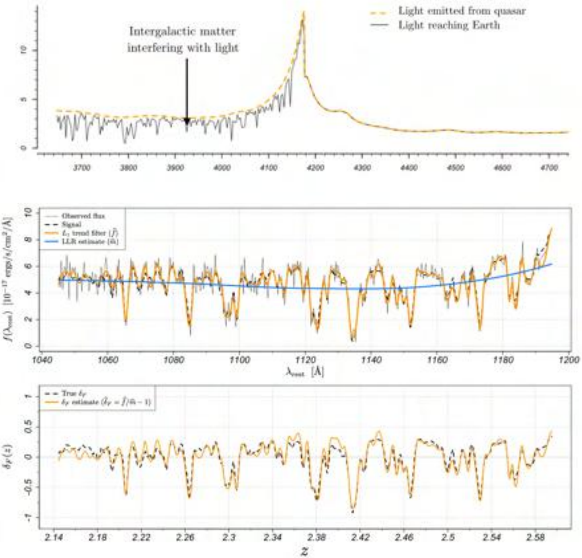
Lyman-α flux contrast:

$$\delta_F(z) = \frac{F(z) - \overline{F}(z)}{\overline{F}(z)}$$

Estimator:

$$\widehat{\delta}_F(z) = \frac{\widehat{f}_0(z) - \widehat{m}(z)}{\widehat{m}(z)}$$

where $\quad m(z) = C(z) \cdot \overline{F}(z)$



Trend filtering: Tibshirani (2014), Politsch et al. (2020 a,b)

Parametric bootstrapping the flux contrast

Uncertainty quantification?

Parametric bootstrapping the flux contrast

Uncertainty quantification?

1. Contstruct the bootstrap sample:

$$f_b^*(\lambda_i) = \widehat{f}_0(\lambda_i) + \epsilon_i^* \qquad \text{where } \epsilon_i^* \sim N(0, \widehat{\sigma}_i^2)$$

Parametric bootstrapping the flux contrast

Uncertainty quantification?

1. Contstruct the bootstrap sample:

$$f_b^*(\lambda_i) = \widehat{f}_0(\lambda_i) + \epsilon_i^* \qquad \text{where } \epsilon_i^* \sim N(0, \widehat{\sigma}_i^2)$$

2. Fit the trend filtering estimate $\widehat{f}_b^*$ for the flux signal and the LOESS estimate $\widehat{m}_b^*$ for the mean flux level

Parametric bootstrapping the flux contrast

Uncertainty quantification?

1. Contstruct the bootstrap sample:

$$f_b^*(\lambda_i) = \widehat{f}_0(\lambda_i) + \epsilon_i^* \qquad \text{where } \epsilon_i^* \sim N(0, \widehat{\sigma}_i^2)$$

2. Fit the trend filtering estimate $\widehat{f}_b^*$ for the flux signal and the LOESS estimate $\widehat{m}_b^*$ for the mean flux level
3. Define the flux contrast estimate:

$$\widehat{\delta}_{F,b}^*(z_i) = \frac{\widehat{f}_b^*(z_i) - \widehat{m}_b^*(z_i)}{\widehat{m}_b^*(z_i)}$$

Parametric bootstrapping the flux contrast

Uncertainty quantification?

1. Contstruct the bootstrap sample:

$$f_b^*(\lambda_i) = \widehat{f}_0(\lambda_i) + \epsilon_i^* \qquad \text{where } \epsilon_i^* \sim N(0, \widehat{\sigma}_i^2)$$

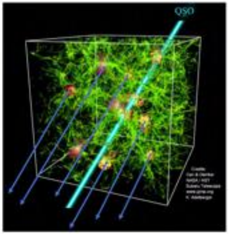
2. Fit the trend filtering estimate $\widehat{f}_b^*$ for the flux signal and the LOESS estimate $\widehat{m}_b^*$ for the mean flux level
3. Define the flux contrast estimate:

$$\widehat{\delta}_{F,b}^*(z_i) = \frac{\widehat{f}_b^*(z_i) - \widehat{m}_b^*(z_i)}{\widehat{m}_b^*(z_i)}$$

4. Contruct the pointwise $1-\alpha$ percentile band:

$$V_{1-\alpha}(z_i) = \left(\widehat{\delta}_{F,\alpha/2}^*(z_i),\ \ \widehat{\delta}_{F,1-\alpha/2}^*(z_i)\right)$$

One-dimensional absorption field reconstruction



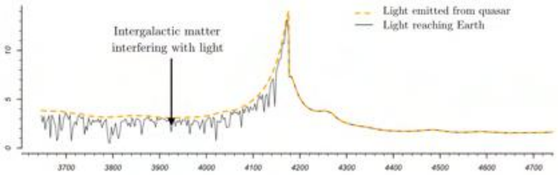
$$
\begin{aligned}
f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot \overline{F}(\lambda_i) \cdot (1 + \delta_F(\lambda_i)) + \epsilon_i
\end{aligned}
$$

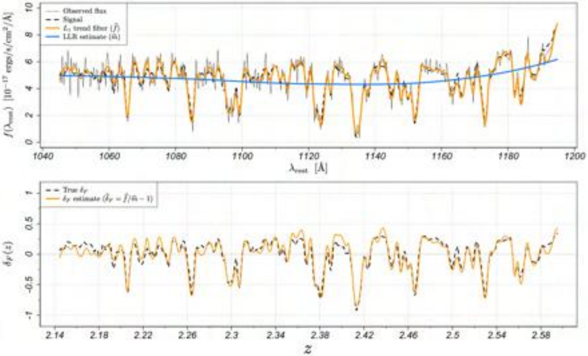
Lyman-α flux contrast:

$$\delta_F(z) = \frac{F(z) - \overline{F}(z)}{\overline{F}(z)}$$

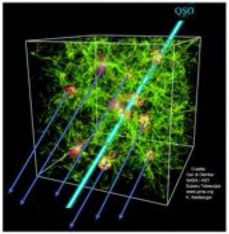
Estimator:

$$\widehat{\delta}_F(z) = \frac{\widehat{f}_0(z) - \widehat{m}(z)}{\widehat{m}(z)}$$





One-dimensional absorption field reconstruction



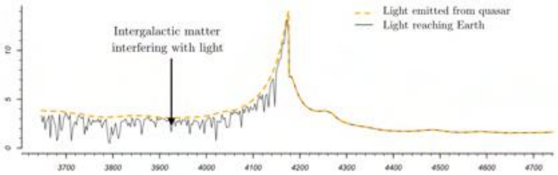
$$
\begin{aligned}
f(\lambda_i) &= f_0(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot F(\lambda_i) + \epsilon_i \\
&= C(\lambda_i) \cdot \overline{F}(\lambda_i) \cdot (1 + \delta_F(\lambda_i)) + \epsilon_i
\end{aligned}
$$

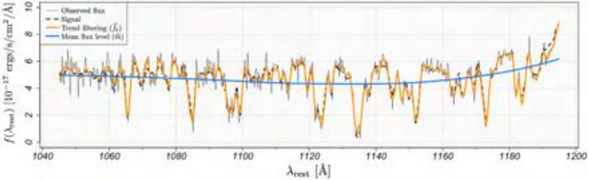
Lyman-α flux contrast:

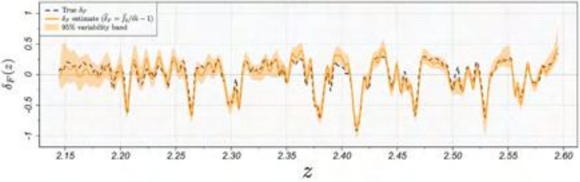
$$\delta_F(z) = \frac{F(z) - \overline{F}(z)}{\overline{F}(z)}$$

Estimator:

$$\widehat{\delta}_F(z) = \frac{\widehat{f}_0(z) - \widehat{m}(z)}{\widehat{m}(z)}$$





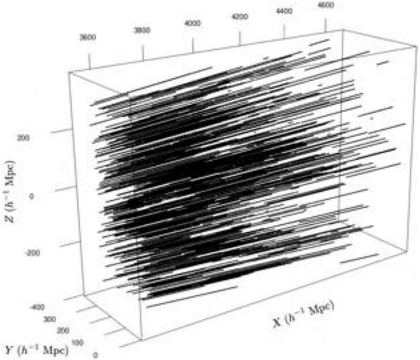


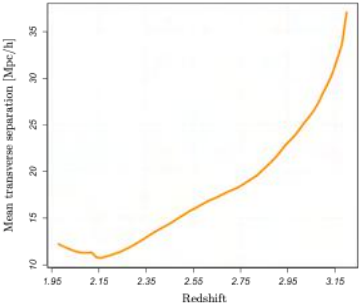
Outline

1. The Lyman-α forest
   - Background
   - The Sloan Digital Sky Survey
   - Summary of work & results
2. One-dimensional mapping the intergalactic medium
   - Denoising observational spectra
   - Estimating the underlying density field
   - Uncertainty quantification
3. Three-dimensional mapping the intergalactic medium
   - Spatial model
   - Distributed computing
   - Optimized absorption field reconstruction
   - Uncertainty quantification
   - High-significance candidates for galaxy protoclusters and voids

Lyman-α forest tomography

Objectives/Essentials:

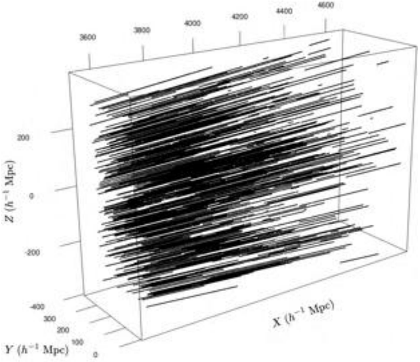


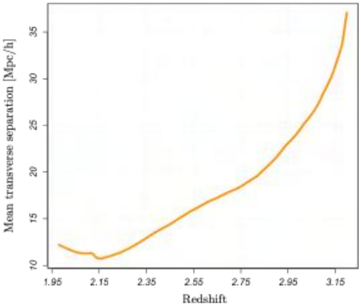


Lyman-α forest tomography

Objectives/Essentials:

1. Reconstruct a full 3D map large-scale structure map from the dense collection of 1D maps

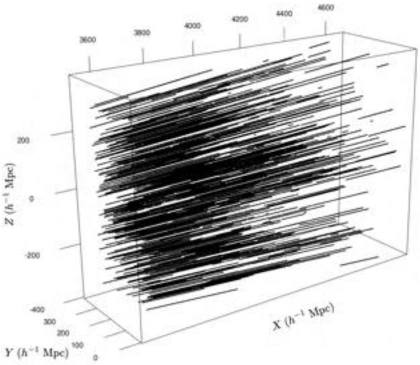


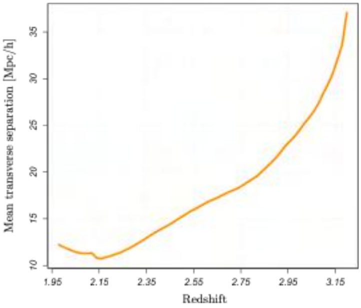


Lyman-α forest tomography

Objectives/Essentials:

1. Reconstruct a full 3D map large-scale structure map from the dense collection of 1D maps
2. **Nonparametric/data-driven modeling.** Optimize the map to be predictive of structure *between sightlines*

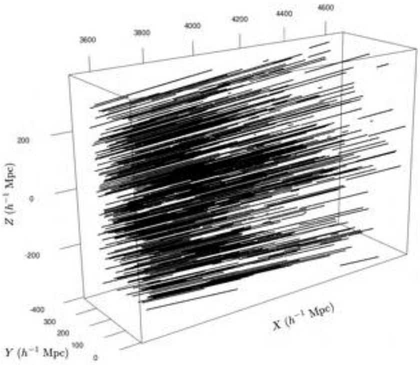


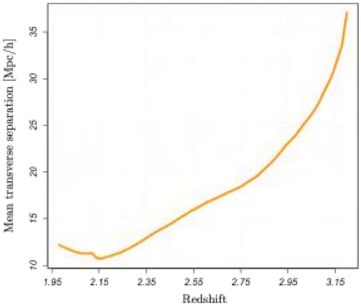


Lyman-α forest tomography

Objectives/Essentials:

1. Reconstruct a full 3D map large-scale structure map from the dense collection of 1D maps
2. **Nonparametric/data-driven modeling.** Optimize the map to be predictive of structure *between sightlines*
3. **Multi-resolution.** The Universe has structure of various scales, so too should our model

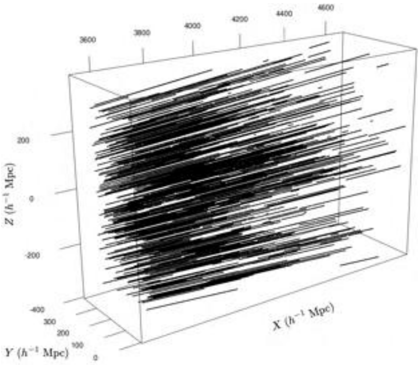


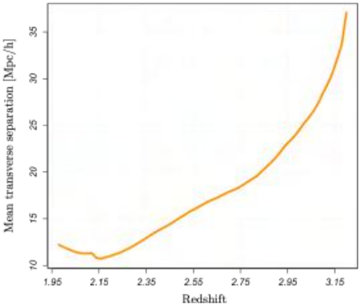


Lyman-α forest tomography

Objectives/Essentials:

1. Reconstruct a full 3D map large-scale structure map from the dense collection of 1D maps
2. **Nonparametric/data-driven modeling.** Optimize the map to be predictive of structure *between sightlines*
3. **Multi-resolution.** The Universe has structure of various scales, so too should our model
4. **Uncertainty quantification.** Track all statistical uncertainty in the reconstructed map

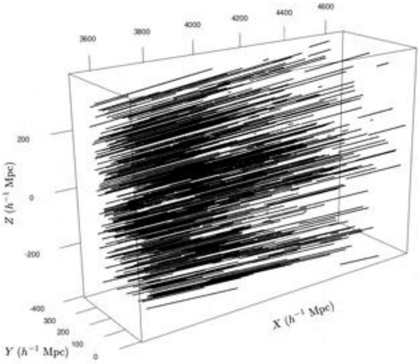


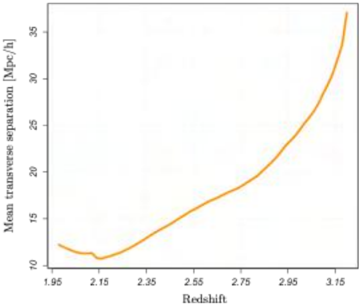


Lyman-α forest tomography

Objectives/Essentials:

1. Reconstruct a full 3D map large-scale structure map from the dense collection of 1D maps
2. **Nonparametric/data-driven modeling.** Optimize the map to be predictive of structure *between sightlines*
3. **Multi-resolution.** The Universe has structure of various scales, so too should our model
4. **Uncertainty quantification.** Track all statistical uncertainty in the reconstructed map
5. **Computational efficiency and scalability.** Sparsity and distributed computing





Multi-resolution spatial kernel ridge regression

Cost functional:

$$\min_{\beta \in \mathbb{R}^m} (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Multi-resolution spatial kernel ridge regression

Cost functional:

$$\min_{\beta \in \mathbb{R}^m} (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Three-dimensional estimator:

$$\widehat{\delta}_F^L(x) = \sum_{\ell,j} \widehat{\beta}_{\ell,j} \phi_{\ell,j}(x), \quad x \in \mathbb{R}^3$$

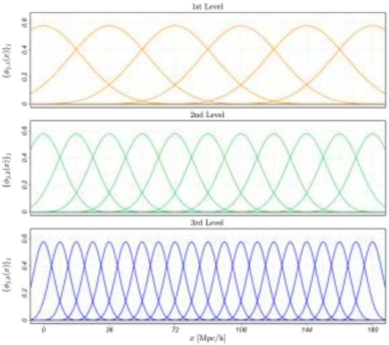
Multi-resolution spatial kernel ridge regression

Cost functional:

$$\min_{\beta \in \mathbb{R}^m} (\widehat{\delta}_F - \Phi\beta)^T W(\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Three-dimensional estimator:

$$\widehat{\delta}_F^L(x) = \sum_{\ell,j} \widehat{\beta}_{\ell,j}\, \phi_{\ell,j}(x), \quad x \in \mathbb{R}^3$$



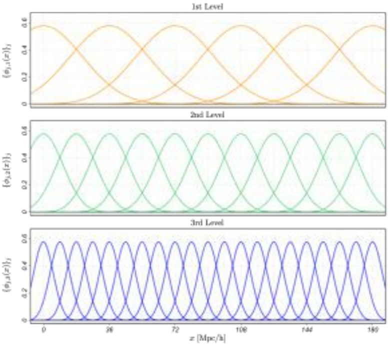
Multi-resolution spatial kernel ridge regression

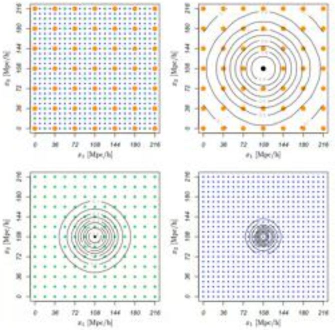
Cost functional:

$$\min_{\beta \in \mathbb{R}^m} (\widehat{\delta}_F - \Phi\beta)^T W(\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Three-dimensional estimator:

$$\widehat{\delta}_F^L(x) = \sum_{\ell,j} \widehat{\beta}_{\ell,j}\phi_{\ell,j}(x), \quad x \in \mathbb{R}^3$$





Multi-resolution generalized ridge penalty

Cost functional:

$$\widehat{\beta} = \underset{\beta \in \mathbb{R}^m}{\operatorname{argmin}} \, (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2,$$

Multi-resolution generalized ridge penalty

Cost functional:

$$\widehat{\beta} = \underset{\beta \in \mathbb{R}^m}{\operatorname{argmin}} \, (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2,$$

Block diagonal penalty:

$$D = \begin{bmatrix} \sqrt{\gamma_1}(\Delta_1 + \alpha_1 I) & 0 & \cdots & 0 \\ 0 & \sqrt{\gamma_2}(\Delta_2 + \alpha_2 I) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sqrt{\gamma_d}(\Delta_d + \alpha_d I) \end{bmatrix}.$$

Multi-resolution generalized ridge penalty

Cost functional:

$$\widehat{\beta} = \underset{\beta \in \mathbb{R}^m}{\operatorname{argmin}} \, (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2,$$

Block diagonal penalty:

$$D = \begin{bmatrix} \sqrt{\gamma_1}(\Delta_1 + \alpha_1 I) & 0 & \cdots & 0 \\ 0 & \sqrt{\gamma_2}(\Delta_2 + \alpha_2 I) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sqrt{\gamma_d}(\Delta_d + \alpha_d I) \end{bmatrix}.$$

First-order fusion:

$$\Delta_{\ell_{ij}} = \begin{cases} 6 & i = j, \\ -1 & j \in \mathcal{N}_i, \\ 0 & \text{otherwise,} \end{cases}$$

Multi-resolution generalized ridge penalty

Cost functional:

$$\widehat{\beta} = \underset{\beta \in \mathbb{R}^m}{\text{argmin}} \ (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2,$$

Block diagonal penalty:

$$D = \begin{bmatrix} \sqrt{\gamma_1}(\Delta_1 + \alpha_1 I) & 0 & \cdots & 0 \\ 0 & \sqrt{\gamma_2}(\Delta_2 + \alpha_2 I) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sqrt{\gamma_d}(\Delta_d + \alpha_d I) \end{bmatrix}.$$

First-order fusion:

$$\Delta_{\ell_{ij}} = \begin{cases} 6 & i = j, \\ -1 & j \in \mathcal{N}_i, \\ 0 & \text{otherwise}, \end{cases}$$

Expanded penalty:

$$\|D\beta\|_2^2 = \gamma_\ell \left(\sum_{\ell=1}^{d} \alpha_\ell^2 \beta_\ell^T \beta_\ell + 2\alpha_\ell \beta_\ell^T \Delta_\ell \beta_\ell + \beta_\ell^T \Delta_\ell^2 \beta_\ell \right), \quad \alpha_\ell > 0, \ \gamma_\ell \geq 0$$

Distributed computing

Cost functional:

$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W(\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Distributed computing

Cost functional:

$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Closed-form solution:

$$\widehat{\beta} = (\Phi^T W \Phi + D^T D)^{-1} \Phi^T W \widehat{\delta}_F$$

Distributed computing

Cost functional:

$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W(\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Closed-form solution:

$$\widehat{\beta} = (\Phi^T W \Phi + D^T D)^{-1} \Phi^T W \widehat{\delta}_F$$

$n \approx$ **72 million** **40 million** $\lesssim p \lesssim$ **50 million**

Distributed computing

Cost functional:

$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Closed-form solution:

$$\widehat{\beta} = (\Phi^T W \Phi + D^T D)^{-1} \Phi^T W \widehat{\delta}_F$$

$n \approx$ **72 million** 40 **million** $\lesssim p \lesssim$ **50 million**

Distributed approximation:

Distributed computing

Cost functional:

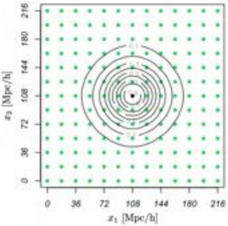
$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Closed-form solution:

$$\widehat{\beta} = (\Phi^T W \Phi + D^T D)^{-1} \Phi^T W \widehat{\delta}_F$$

$n \approx 72$ **million** 40 **million** $\lesssim p \lesssim$ 50 **million**

Distributed approximation:



Distributed computing

Cost functional:

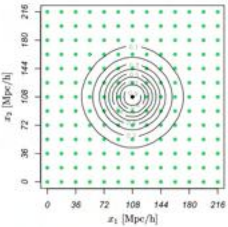
$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Closed-form solution:

$$\widehat{\beta} = (\Phi^T W \Phi + D^T D)^{-1} \Phi^T W \widehat{\delta}_F$$

$n \approx 72$ million **40 million $\lesssim p \lesssim$ 50 million**

Distributed approximation: Let $A_1, \ldots, A_r$ be a partition of the 47 h^{-3} Gpc^3 volume.



Distributed computing

Cost functional:

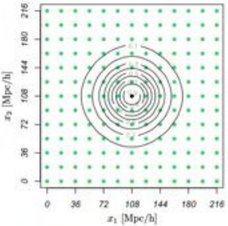
$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W(\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Closed-form solution:

$$\widehat{\beta} = (\Phi^T W \Phi + D^T D)^{-1} \Phi^T W \widehat{\delta}_F$$

$n \approx$ **72 million**   **40 million** $\lesssim p \lesssim$ **50 million**

Distributed approximation: Let $A_1, \ldots, A_r$ be a partition of the 47 h^{-3} Gpc^3 volume.



$$\widehat{\beta}_k = (\Phi_k^T W_k \Phi_k + D_k^T D_k)^{-1} \Phi_k^T W_k \widehat{\delta}_{F_k}$$

Distributed computing

Cost functional:

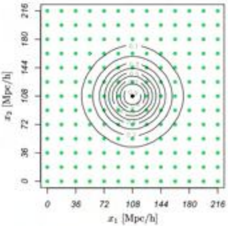
$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Closed-form solution:

$$\widehat{\beta} = (\Phi^T W \Phi + D^T D)^{-1} \Phi^T W \widehat{\delta}_F$$

$n \approx 72$ **million** 40 **million** $\lesssim p \lesssim$ 50 **million**

Distributed approximation: Let $A_1, \ldots, A_r$ be a partition of the 47 h^{-3} Gpc^3 volume.



$$\widehat{\beta}_k = (\Phi_k^T W_k \Phi_k + D_k^T D_k)^{-1} \Phi_k^T W_k \widehat{\delta}_{F_k}$$

$$\widehat{\delta}^L_{F_k}(x) = \sum_{\ell,j} \widehat{\beta}_{k_{\ell,j}} \phi_{k_{\ell,j}}(x), \quad x \in \widetilde{A}_k$$

Distributed computing

Cost functional:

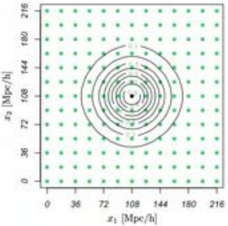
$$\min_{\beta \in \mathbb{R}^p} (\widehat{\delta}_F - \Phi\beta)^T W (\widehat{\delta}_F - \Phi\beta) + \|D\beta\|_2^2$$

Closed-form solution:

$$\widehat{\beta} = (\Phi^T W \Phi + D^T D)^{-1} \Phi^T W \widehat{\delta}_F$$

$n \approx$ **72 million** **40 million** $\lesssim p \lesssim$ **50 million**

Distributed approximation: Let $A_1, \ldots, A_r$ be a partition of the 47 h^{-3} Gpc^3 volume.



$$\widehat{\beta}_k = (\Phi_k^T W_k \Phi_k + D_k^T D_k)^{-1} \Phi_k^T W_k \widehat{\delta}_{F_k}$$

$$\widehat{\delta}^L_{F_k}(x) = \sum_{\ell, j} \widehat{\beta}_{k_{\ell,j}} \phi_{k_{\ell,j}}(x), \quad x \in \widetilde{A}_k$$

$$\widehat{\delta}^{L\|}_F(x) = \sum_{k=1}^{r} \mathbb{1}(x \in A_k) \cdot \widehat{\delta}^L_{F_k}(x)$$

Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters

Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters
- Construct an 85%/15% Train/Validation split on the set of background quasars

Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters
- Construct an 85%/15% Train/Validation split on the set of background quasars
- Estimate the large-scale density field along each validation sightline

Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters
- Construct an 85%/15% Train/Validation split on the set of background quasars
- Estimate the large-scale density field along each validation sightline

Kendall tau ranking distance:

$$L(\eta) = |\mathcal{D}_1| + |\mathcal{D}_2|$$

$$\mathcal{D}_1 = \{(i,j)_{i<j} : \rho(\widehat{\delta}_{F_i}^{L\|}) < \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) > \rho(\widetilde{\delta}_{F_j}^{L})\}$$

$$\mathcal{D}_2 = \{(i,j)_{i<j} : \rho(\widehat{\delta}_{F_i}^{L\|}) > \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) < \rho(\widetilde{\delta}_{F_j}^{L})\}$$

Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters
- Construct an 85%/15% Train/Validation split on the set of background quasars
- Estimate the large-scale density field along each validation sightline

Kendall tau ranking distance: (Bubble sort distance)

$$L(\eta) = |\mathcal{D}_1| + |\mathcal{D}_2|$$

$$\mathcal{D}_1 = \{(i,j)_{i<j} : \ \rho(\widehat{\delta}_{F_i}^{L\|}) < \rho(\widehat{\delta}_{F_j}^{L\|}) \ \cap \ \rho(\widetilde{\delta}_{F_i}^{L}) > \rho(\widetilde{\delta}_{F_j}^{L})\}$$

$$\mathcal{D}_2 = \{(i,j)_{i<j} : \ \rho(\widehat{\delta}_{F_i}^{L\|}) > \rho(\widehat{\delta}_{F_j}^{L\|}) \ \cap \ \rho(\widetilde{\delta}_{F_i}^{L}) < \rho(\widetilde{\delta}_{F_j}^{L})\}$$

Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters
- Construct an 85%/15% Train/Validation split on the set of background quasars
- Estimate the large-scale density field along each validation sightline

Kendall tau ranking distance: (Bubble sort distance)

$$L(\eta) = |\mathcal{D}_1| + |\mathcal{D}_2|$$

$$\mathcal{D}_1 = \{(i,j)_{i<j} :\ \rho(\widehat{\delta}_{F_i}^{L\|}) < \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) > \rho(\widetilde{\delta}_{F_j}^{L})\}$$

$$\mathcal{D}_2 = \{(i,j)_{i<j} :\ \rho(\widehat{\delta}_{F_i}^{L\|}) > \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) < \rho(\widetilde{\delta}_{F_j}^{L})\}$$

Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters
- Construct an 85%/15% Train/Validation split on the set of background quasars
- Estimate the large-scale density field along each validation sightline

Kendall tau ranking distance: (Bubble sort distance)

$$L(\eta) = |\mathcal{D}_1| + |\mathcal{D}_2|$$

$$\widehat{\eta} = \mathrm{argmin}_\eta \, L(\eta)$$

$$\mathcal{D}_1 = \{(i,j)_{i<j} : \rho(\widehat{\delta}_{F_i}^{L\|}) < \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) > \rho(\widetilde{\delta}_{F_j}^{L})\}$$

$$\mathcal{D}_2 = \{(i,j)_{i<j} : \rho(\widehat{\delta}_{F_i}^{L\|}) > \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) < \rho(\widetilde{\delta}_{F_j}^{L})\}$$

Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters
- Construct an 85%/15% Train/Validation split on the set of background quasars
- Estimate the large-scale density field along each validation sightline

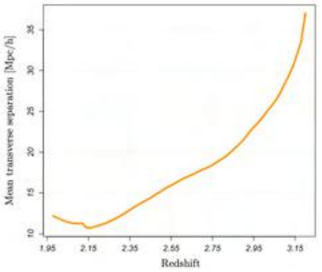
Kendall tau ranking distance: (Bubble sort distance)

$$L(\eta) = |\mathcal{D}_1| + |\mathcal{D}_2|$$

$$\widehat{\eta} = \text{argmin}_\eta \, L(\eta)$$

$$\mathcal{D}_1 = \{(i,j)_{i<j} : \ \rho(\widehat{\delta}_{F_i}^{L\|}) < \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) > \rho(\widetilde{\delta}_{F_j}^{L})\}$$

$$\mathcal{D}_2 = \{(i,j)_{i<j} : \ \rho(\widehat{\delta}_{F_i}^{L\|}) > \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) < \rho(\widetilde{\delta}_{F_j}^{L})\}$$



Model validation

- Let $\eta = (d, \alpha, \gamma_1, \ldots, \gamma_d)$ be the vector of model hyperparameters
- Construct an 85%/15% Train/Validation split on the set of background quasars
- Estimate the large-scale density field along each validation sightline

Kendall tau ranking distance: (Bubble sort distance)

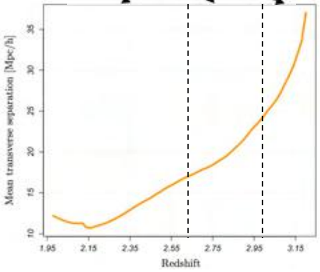
$$L(\eta) = |\mathcal{D}_1| + |\mathcal{D}_2|$$

$$\mathcal{D}_1 = \{(i,j)_{i<j} : \rho(\widehat{\delta}_{F_i}^{L\|}) < \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) > \rho(\widetilde{\delta}_{F_j}^{L})\}$$

$$\mathcal{D}_2 = \{(i,j)_{i<j} : \rho(\widehat{\delta}_{F_i}^{L\|}) > \rho(\widehat{\delta}_{F_j}^{L\|}) \cap \rho(\widetilde{\delta}_{F_i}^{L}) < \rho(\widetilde{\delta}_{F_j}^{L})\}$$

$$\widehat{\eta} = \text{argmin}_{\eta}\, L(\eta)$$

Optimize separately

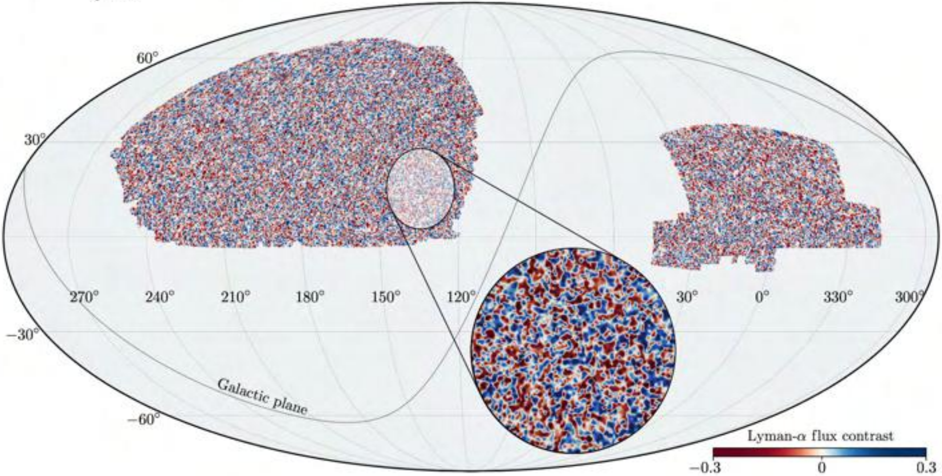


Cross-sectional sky map of the intergalactic medium

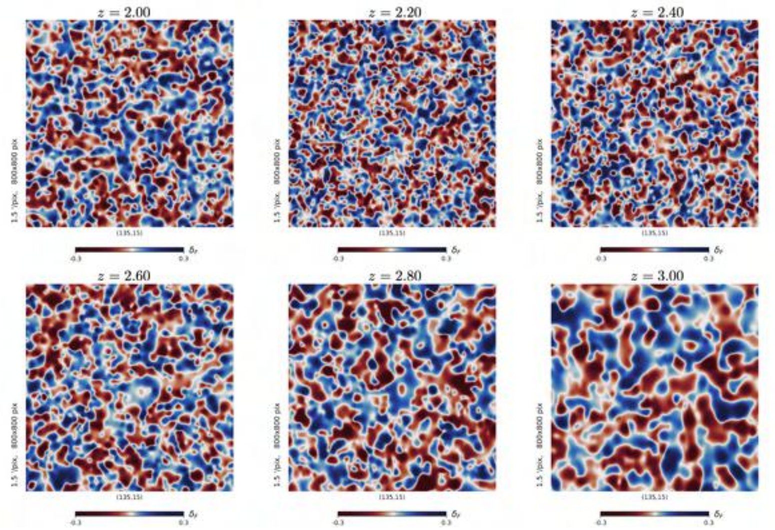
$$\widehat{\delta}_F^{L\|}(x) = \sum_{j=1}^{3}\sum_{k=1}^{r_j} \mathbb{1}(x \in A_{j,k}) \cdot \widehat{\delta}_F^{L,j,k}(x; \widehat{\eta}_j)$$

Cross-sectional sky map of the intergalactic medium

$$\widehat{\delta}_F^{L\|}(x) = \sum_{j=1}^{3}\sum_{k=1}^{r_j} \mathbb{1}(x \in A_{j,k}) \cdot \widehat{\delta}_F^{L,j,k}(x; \widehat{\eta}_j)$$



Slice of intergalactic medium at various redshifts



Uncertainty quantification

Uncertainty quantification

1. Construct a bootstrap sample of quasar spectra

$$f^*_{q,b}(\lambda_{q_i}) \sim N(\widehat{f}_{q,0}(\lambda_{q_i}), \widehat{\sigma}^2_{q_i}), \quad i = 1, \ldots, n, \quad q = 1, \ldots, 159,581$$

Uncertainty quantification

1. Construct a bootstrap sample of quasar spectra

$$f^*_{q,b}(\lambda_{q_i}) \sim N(\widehat{f}_{q,0}(\lambda_{q_i}), \widehat{\sigma}^2_{q_i}), \quad i = 1, \dots, n, \quad q = 1, \dots, 159,581$$

2. For each spectrum, fit the trend filtering estimate $\widehat{f}^*_{q,b}$ for the flux signal and the LOESS estimate $\widehat{m}^*_{q,b}$ for the mean flux level

Uncertainty quantification

1. Construct a bootstrap sample of quasar spectra

$$f^*_{q,b}(\lambda_{q_i}) \sim N(\widehat{f}_{q,0}(\lambda_{q_i}), \widehat{\sigma}^2_{q_i}), \quad i = 1, \dots, n, \quad q = 1, \dots, 159,581$$

2. For each spectrum, fit the trend filtering estimate $\widehat{f}^*_{q,b}$ for the flux signal and the LOESS estimate $\widehat{m}^*_{q,b}$ for the mean flux level

3. Define the flux contrast estimates:

$$\widehat{\delta}^*_{F,b}(z_{q_i}) = \frac{\widehat{f}^*_{q,b}(z_{q_i}) - \widehat{m}^*_{q,b}(z_{q_i})}{\widehat{m}^*_{q,b}(z_{q_i})}$$

Uncertainty quantification

1. Construct a bootstrap sample of quasar spectra

$$f^*_{q,b}(\lambda_{q_i}) \sim N(\widehat{f}_{q,0}(\lambda_{q_i}), \widehat{\sigma}^2_{q_i}), \quad i = 1, \dots, n, \quad q = 1, \dots, 159,581$$

2. For each spectrum, fit the trend filtering estimate $\widehat{f}^*_{q,b}$ for the flux signal and the LOESS estimate $\widehat{m}^*_{q,b}$ for the mean flux level

3. Define the flux contrast estimates:

$$\widehat{\delta}^*_{F,b}(z_{q_i}) = \frac{\widehat{f}^*_{q,b}(z_{q_i}) - \widehat{m}^*_{q,b}(z_{q_i})}{\widehat{m}^*_{q,b}(z_{q_i})}$$

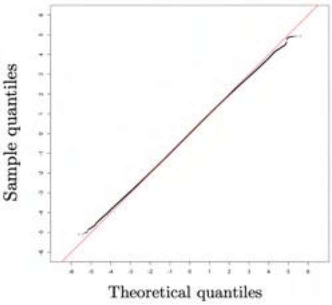
4. Compute the distributed-SKRR 3D reconstruction $\widehat{\delta}^{L\|*}_F(x)$ from the pooled sightline flux contrast estimates.

Uncertainty quantification

$\Longrightarrow$ Produce 50 bootstrap reconstructions of the full 47 h^{-3} Gpc^3 absorption field

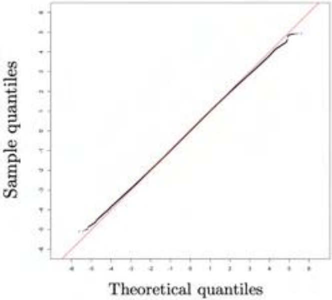
Uncertainty quantification

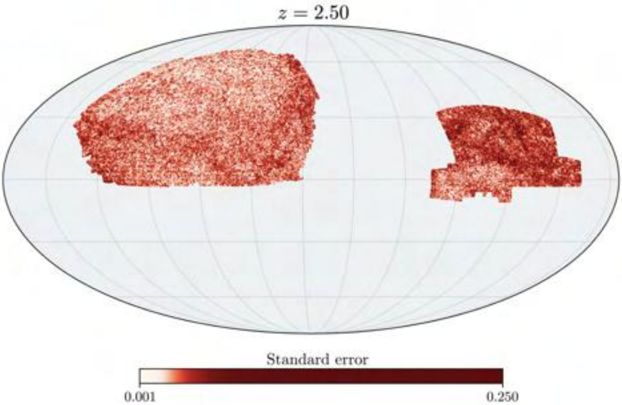
$\Longrightarrow$ Produce 50 bootstrap reconstructions of the full 47 h^{-3} Gpc^3 absorption field



Uncertainty quantification

$\Longrightarrow$ Produce 50 bootstrap reconstructions of the full 47 h^{-3} Gpc^3 absorption field





Candidates for galaxy protoclusters and voids

Candidates for galaxy protoclusters and voids

Statistically significant overdensities:

$$G_n = \{\widehat{\delta}_F^L(x) : \widehat{\delta}_F^L(x) < -n \cdot \widehat{\mathrm{se}}(\widehat{\delta}_F^L(x))\}$$

Candidates for galaxy protoclusters and voids

Statistically significant overdensities:

$$G_n = \{\widehat{\delta}_F^L(x) : \widehat{\delta}_F^L(x) < -n \cdot \widehat{\mathrm{se}}(\widehat{\delta}_F^L(x))\}$$

Statistically significant underdensities:

$$V_n = \{\widehat{\delta}_F^L(x) : \widehat{\delta}_F^L(x) > n \cdot \widehat{\mathrm{se}}(\widehat{\delta}_F^L(x))\}$$

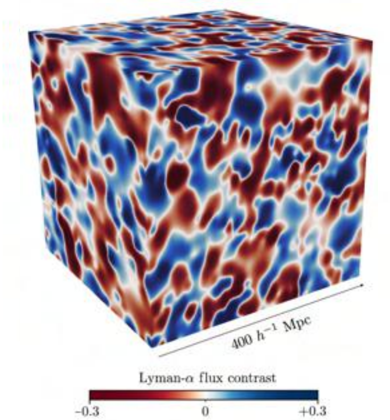
Candidates for galaxy protoclusters and voids

Statistically significant overdensities:

$$G_n = \{\widehat{\delta}_F^L(x) : \widehat{\delta}_F^L(x) < -n \cdot \widehat{\text{se}}(\widehat{\delta}_F^L(x))\}$$

Statistically significant underdensities:

$$V_n = \{\widehat{\delta}_F^L(x) : \widehat{\delta}_F^L(x) > n \cdot \widehat{\text{se}}(\widehat{\delta}_F^L(x))\}$$



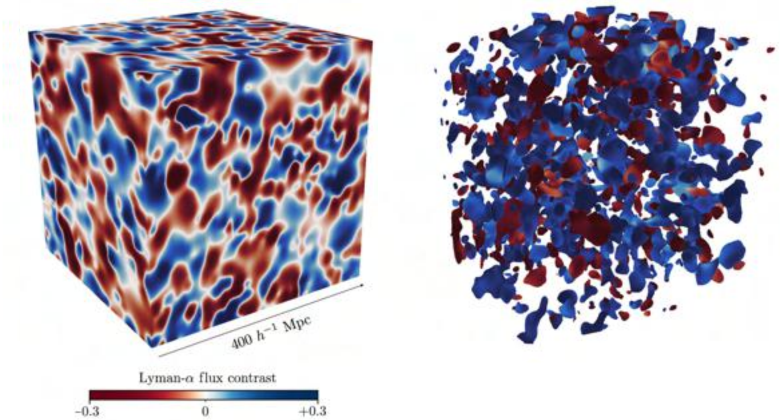
Candidates for galaxy protoclusters and voids

Statistically significant overdensities:

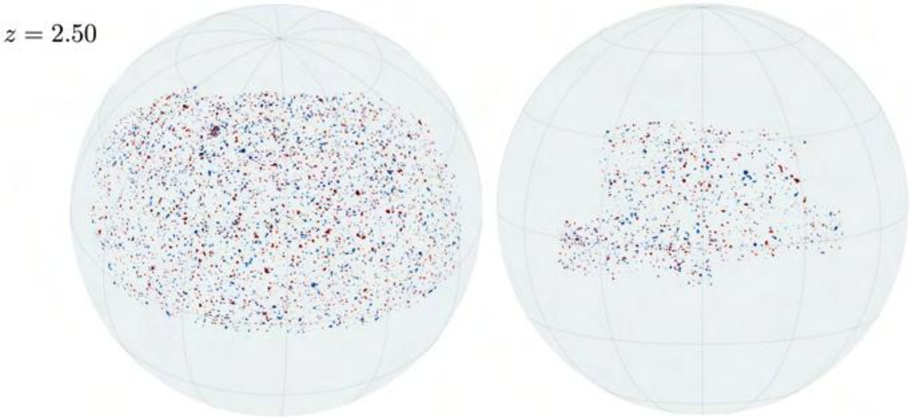
$$G_n = \{\widehat{\delta}_F^L(x) : \widehat{\delta}_F^L(x) < -n \cdot \widehat{\text{se}}(\widehat{\delta}_F^L(x))\}$$

Statistically significant underdensities:

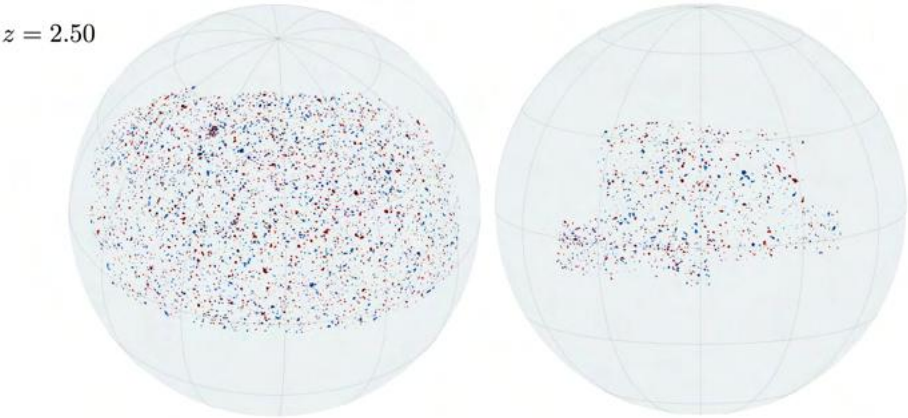
$$V_n = \{\widehat{\delta}_F^L(x) : \widehat{\delta}_F^L(x) > n \cdot \widehat{\text{se}}(\widehat{\delta}_F^L(x))\}$$



Candidates for galaxy protoclusters and cosmic voids



Candidates for galaxy protoclusters and cosmic voids



Census of candidates (rough estimates based on partial catalog)

	Significance level 3σ	4σ	5σ	6σ	7σ	8σ
Galaxy Protoclusters	184,250	130,535	82,209	50,029	29,493	18,272
Cosmic Voids	173,882	127,098	82,016	54,507	34,948	22,266

Dissemination of data products (in progress)

All data products will be disseminated at:

`http://stat.cmu.edu/Lyman-alpha-cosmos-map/` **(powered by PSC)**

Dissemination of data products (in progress)

All data products will be disseminated at:

`http://stat.cmu.edu/Lyman-alpha-cosmos-map/` **(powered by PSC)**

- SQL database ($\sim$30 TB) — 1 h^{-3} Mpc^3 voxel resolution

Dissemination of data products (in progress)

All data products will be disseminated at:

`http://stat.cmu.edu/Lyman-alpha-cosmos-map/` **(powered by PSC)**

- SQL database (~30 TB) — 1 h^{-3} Mpc^3 voxel resolution
- Downloadable HEALPix sky maps $z = 1.98, 1.99, \ldots, 3.15$

Dissemination of data products (in progress)

All data products will be disseminated at:

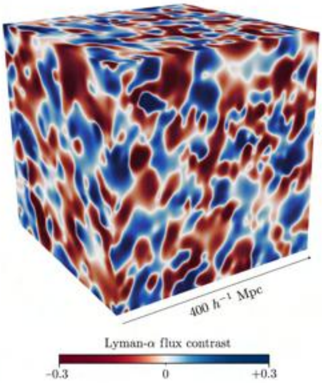
`http://stat.cmu.edu/Lyman-alpha-cosmos-map/` **(powered by PSC)**

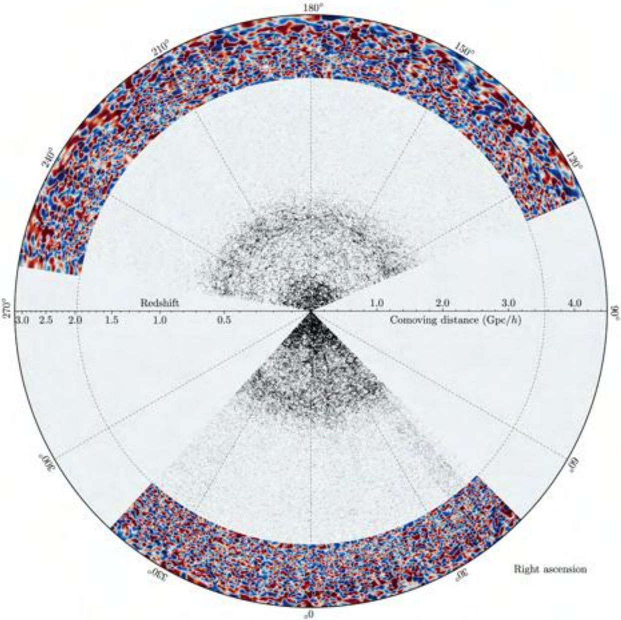
- SQL database (~30 TB) — 1 h^{-3} Mpc^3 voxel resolution
- Downloadable HEALPix sky maps $z = 1.98, 1.99, \ldots, 3.15$
- Downloadable catalog of candidates for galaxy protoclusters and cosmic voids

References

- **C. A. Politsch**, J. Cisewski-Kehe, R. A. C. Croft, L. Wasserman, Three-dimensional cosmography of the high redshift Universe using intergalactic absorption, In preparation (2020), Inquiry approved by *Nature*.
- **C. A. Politsch**, J. Cisewski-Kehe, R. A. C. Croft, L. Wasserman, Trend filtering – I. A modern statistical tool for time-domain astronomy and astronomical spectroscopy, Monthly Notices of the Royal Astronomical Society, Volume 492, Issue 3 (2020), 4005-4018.
- **C. A. Politsch**, J. Cisewski-Kehe, R. A. C. Croft, L. Wasserman, Trend filtering – II. Denoising astronomical signals with varying degrees of smoothness, Monthly Notices of the Royal Astronomical Society, Volume 492, Issue 3 (2020), 4019-4032.
- **C. A. Politsch** & R. A. C. Croft. Mapping the Large-Scale Universe through Intergalactic Silhouettes, CHANCE, Volume 32, Issue 3 (2019), 14-19.
- K. Dawson et al. The Baryon Oscillation Spectroscopic Survey of SDSS-III. The Astronomical Journal, Volume 145, Issue 1 (2013), 41 pp.
- S. Alam et al. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. The Astrophysical Journal Supplement Series, Volume 219, Issue 1 (2015), 27 pp.
- R. J. Tibshirani. Adaptive piecewise polynomial estimation via trend filtering. Annals of Statistics, Volume 42, Number 1 (2014), 285-323.
- R. J. Tibshirani & J. Taylor. The solution path of the generalized lasso. Annals of Statistics, Volume 39, Number 3 (2011), 1335-1371.
- A. Ramdas & R. J. Tibshirani. Fast and Flexible ADMM Algorithms for Trend Filtering. Journal of Computational and Graphical Statistics, Volume 25 (2014), 839-858.
- D. Nychka, S. Bandyopadhyay, D. Hammerling, F. Lindgren, S. Sain. A Multiresolution Gaussian Process Model for the Analysis of Large Spatial Datasets, Journal of Computational and Graphical Statistics, Volume 24, Number 2 (2015), 579-599.
- R. Cen & J. P. Ostriker. Where Are the Baryons? The Astrophysical Journal, Volume 514, Issue 1 (1999), 1-6.
- Planck Collaboration. Planck 2018 results. VI. Cosmological parameters. arXiv (2018).

Thank you!

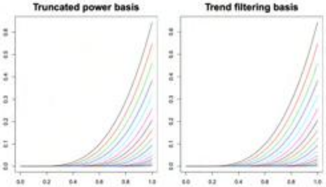

400 h^{-1} Mpc
Lyman-α flux contrast
−0.3
0
+0.3


180°
210°
150°
240°
120°
270°
90°
300°
60°
330°
30°
0°
Redshift
3.0 2.5 2.0 1.5 1.0 0.5
1.0 2.0 3.0 4.0
Comoving distance (Gpc/h)
Right ascension

Overview of trend filtering

$$f(t_i) = f_0(t_i) + \epsilon_i \qquad i = 1, \ldots, n$$

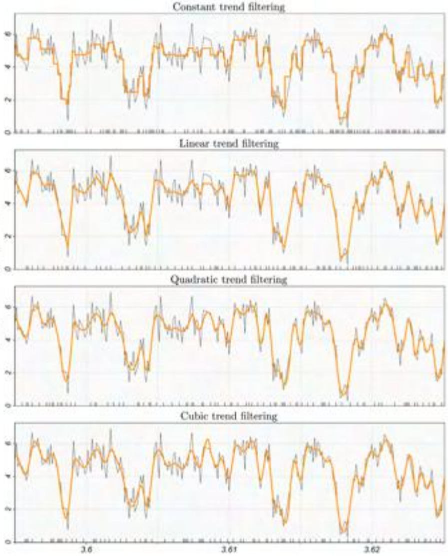
$$f_0(t) = \sum_j \beta_j h_j(t)$$



$$\min_{\{\beta_j\}} \sum_{i=1}^{n} \left(f(t_i) - \sum_{j=1}^{n} \beta_j h_j(t_i) \right)^2 + \gamma \cdot k! \cdot \Delta t^k \sum_{j=k+2}^{n} |\beta_j|$$

$$\widehat{f}_0(t;\gamma) = \sum_{j=1}^{n} \widehat{\beta}_j h_j(t)$$

References: Tibshirani & Taylor (2011), Tibshirani (2014), Ramdas & Tibshirani (2015), Politsch et al. (2020a,b)



ℓ_p-penalized linear models

	High-dimensional	1D splines*
$p = 0$	**Subset selection** $\min_{\{\beta_j\}} \sum_{i=1}^{n} \left(y_i - \sum_j \beta_j x_{ji} \right)^2$ s.t. $\sum_{j=1}^{d} \mathbb{1}\{\beta_j \neq 0\} = s$	**Variable-knot regression spline** $\min_{\{\beta_j\}} \sum_{i=1}^{n} \left(f(t_i) - \sum_j \beta_j \eta_j(t_i) \right)^2$ s.t. $\sum_{j \geq k+2} \mathbb{1}\{\beta_j \neq 0\} = s$
$p = 1$	**Lasso** $\min_{\{\beta_j\}} \sum_{i=1}^{n} \left(y_i - \sum_j \beta_j x_{ji} \right)^2$ s.t. $\sum_{j=1}^{d} \lvert\beta_j\rvert \leq \tau$	
$p = 2$	**Ridge regression** $\min_{\{\beta_j\}} \sum_{i=1}^{n} \left(y_i - \sum_j \beta_j x_{ji} \right)^2$ s.t. $\sum_{j=1}^{d} \beta_j^2 \leq \tau$	**Smoothing spline** $\min_{\{\beta_j\}} \sum_{i=1}^{n} \left(f(t_i) - \sum_j \beta_j \eta_j(t_i) \right)^2$ s.t. $\sum_{j,k=1}^{n} \beta_j \beta_k \omega_{jk} \leq \tau$

ℓ_p-penalized linear models

	High-dimensional	1D splines*
$p = 0$	**Subset selection** $$\min_{\{\beta_j\}} \sum_{i=1}^n \Big(y_i - \sum_j \beta_j x_{ji}\Big)^2 \quad \text{s.t.} \sum_{j=1}^d 1\{\beta_j \neq 0\} = s$$	**Variable-knot regression spline** $$\min_{\{\beta_j\}} \sum_{i=1}^n \Big(f(t_i) - \sum_j \beta_j \eta_j(t_i)\Big)^2 \quad \text{s.t.} \sum_{j \geq k+2} 1\{\beta_j \neq 0\} = s$$
$p = 1$	**Lasso** $$\min_{\{\beta_j\}} \sum_{i=1}^n \Big(y_i - \sum_j \beta_j x_{ji}\Big)^2 \quad \text{s.t.} \sum_{j=1}^d \lvert\beta_j\rvert \leq \tau$$	**Trend filtering** $$\min_{\{\beta_j\}} \sum_{i=1}^n \Big(f(t_i) - \sum_j \beta_j h_j(t_i)\Big)^2 \quad \text{s.t.} \sum_{j \geq k+2} \lvert\beta_j\rvert \leq \tau$$
$p = 2$	**Ridge regression** $$\min_{\{\beta_j\}} \sum_{i=1}^n \Big(y_i - \sum_j \beta_j x_{ji}\Big)^2 \quad \text{s.t.} \sum_{j=1}^d \beta_j^2 \leq \tau$$	**Smoothing spline** $$\min_{\{\beta_j\}} \sum_{i=1}^n \Big(f(t_i) - \sum_j \beta_j \eta_j(t_i)\Big)^2 \quad \text{s.t.} \sum_{j,k=1}^n \beta_j \beta_k \omega_{jk} \leq \tau$$