From Pixels to Parameters

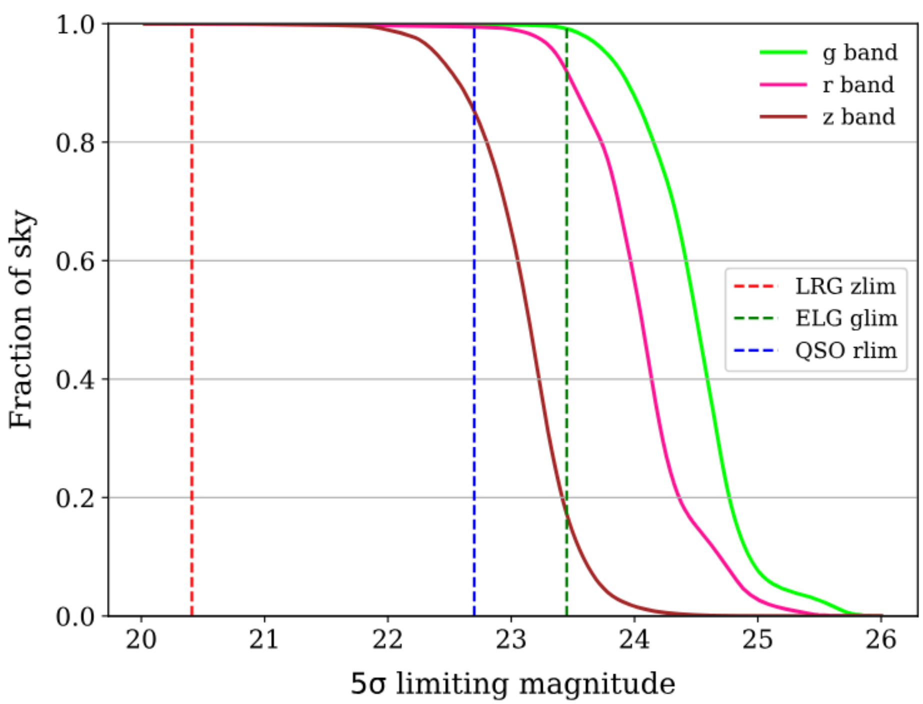
- Part 1: Galaxy Surveys
  - Galaxies beyond the point-particle picture
- **Part 2: From Pixels to Power Spectra**
  - **Systematics, estimators & covariances**
- Part 3: From Power Spectra to Parameters
  - Inference & error bars you can trust
- Part 4: Weak Lensing
  - Galaxies beyond the spin-2 field picture

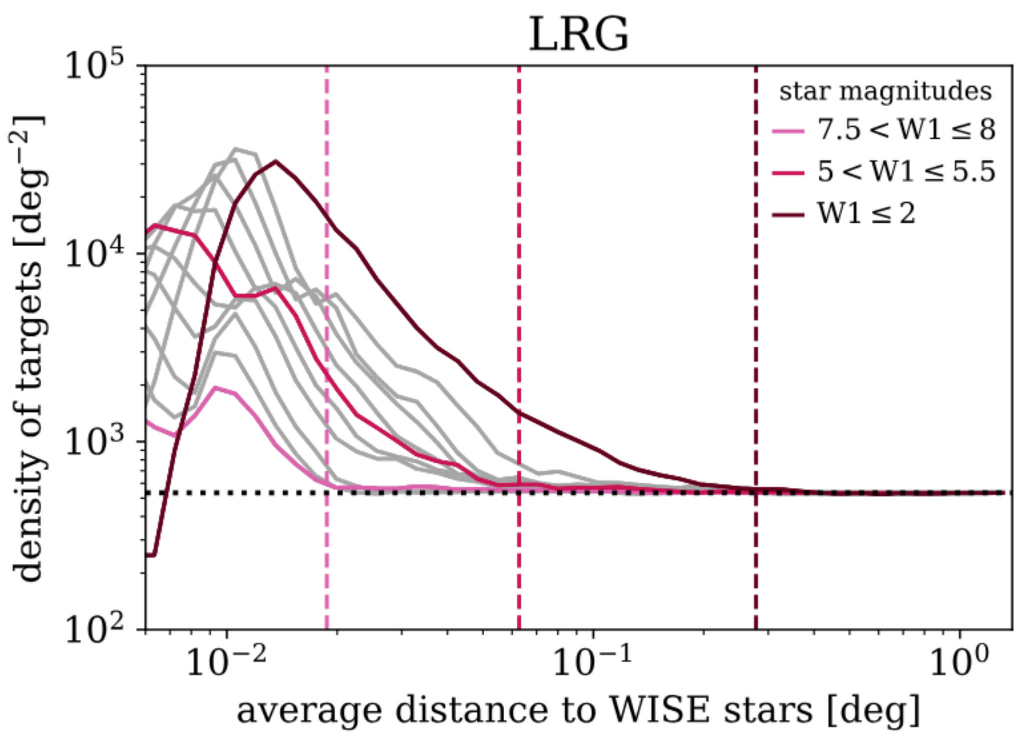
Catalog Systematics: Foregrounds

Target selection from imaging data (color+magnitude cuts)

→ selection probability of spectroscopic sample modulated by systematics of the imaging survey(s)

$$N_{\rm obs}(\hat{\mathbf{n}}) = (1 + f_{\rm sys}(\hat{\mathbf{n}}))N_{\rm true}(\hat{\mathbf{n}})$$





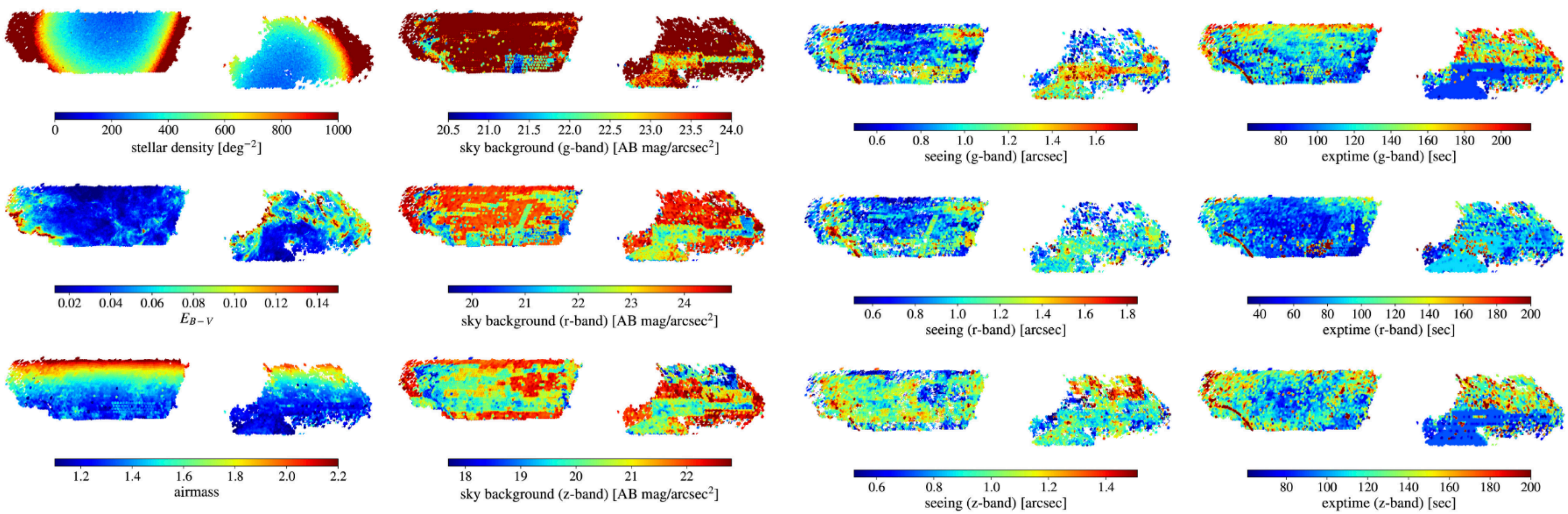
Kitanidis+2020

Catalog Systematics: Foregrounds

Target selection from imaging data (color+magnitude cuts)

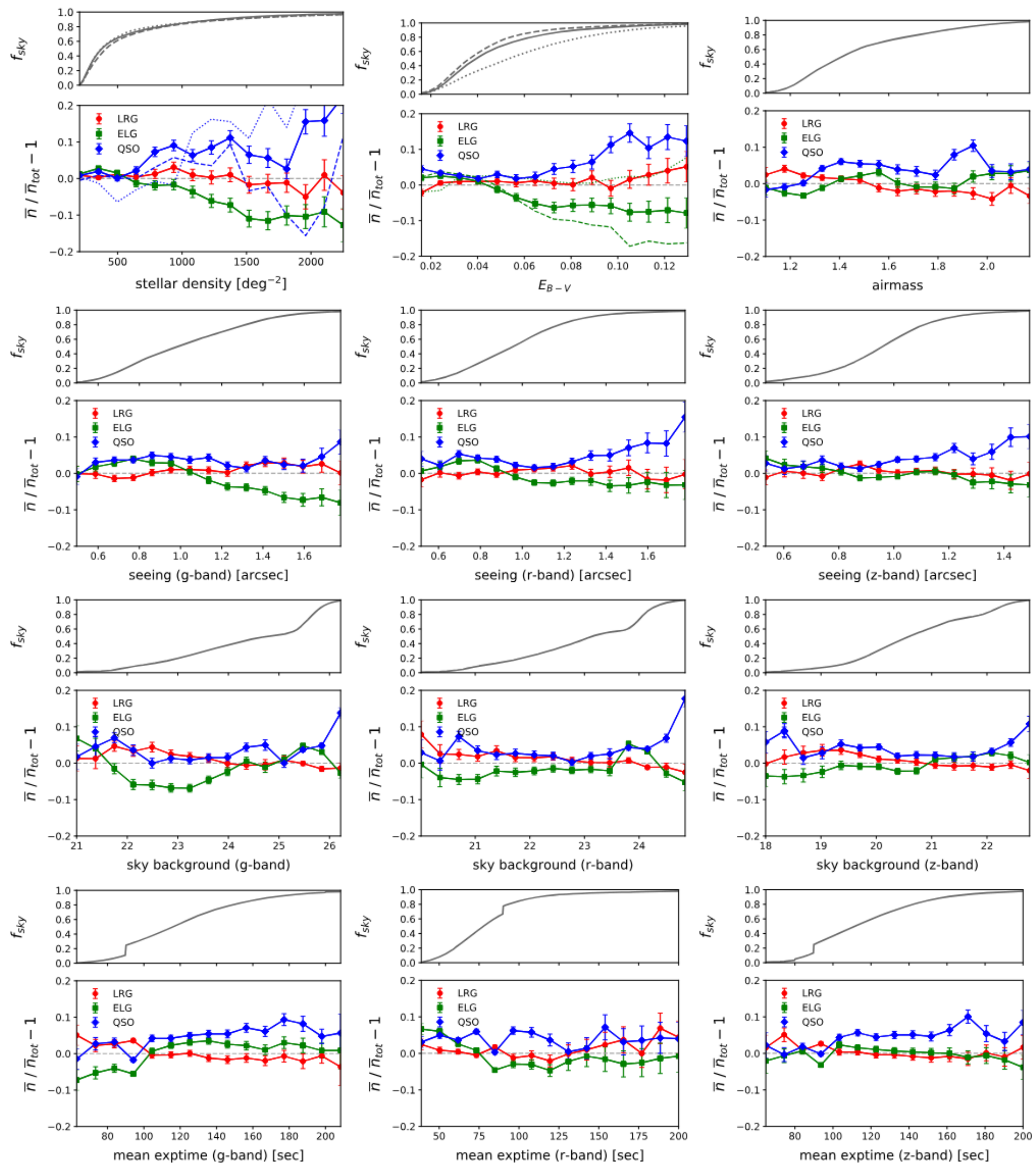
→ selection probability of spectroscopic sample modulated by systematics of the imaging survey(s)

$$N_{\rm obs}(\hat{\mathbf{n}}) = (1 + f_{\rm sys}(\hat{\mathbf{n}}))N_{\rm true}(\hat{\mathbf{n}})$$



potential foreground systematics affecting DESI target selecting (Kitanidis+2020)

Catalog Systematics: Foregrounds



Kitanidis+2020

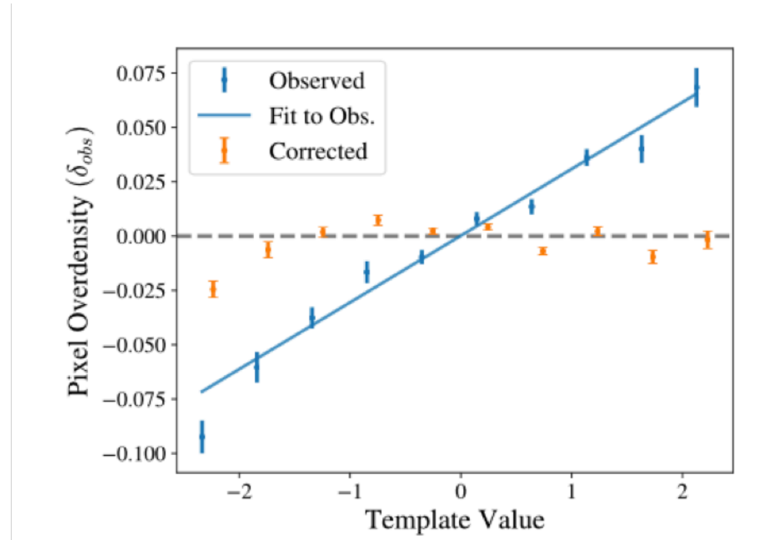
Catalog Systematics: Foregrounds

•List of systematics depends on ground/space, imaging-based target selection or slit-less survey

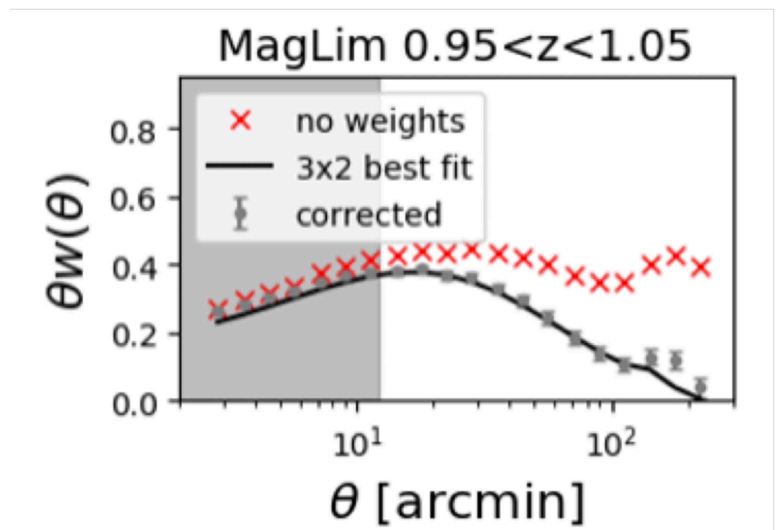
•Mitigation generally template based

- multi-linear regression at field level (Ross+2016, ff)
- template subtraction (Ho+12, ff), mode projection (Leistedt+2013, ff) at summary statistics level
- methods are equivalent for certain prior choices (Weaverdyck & Huterer 2021)

beware of overfitting!



Weaverdyck & Huterer 21



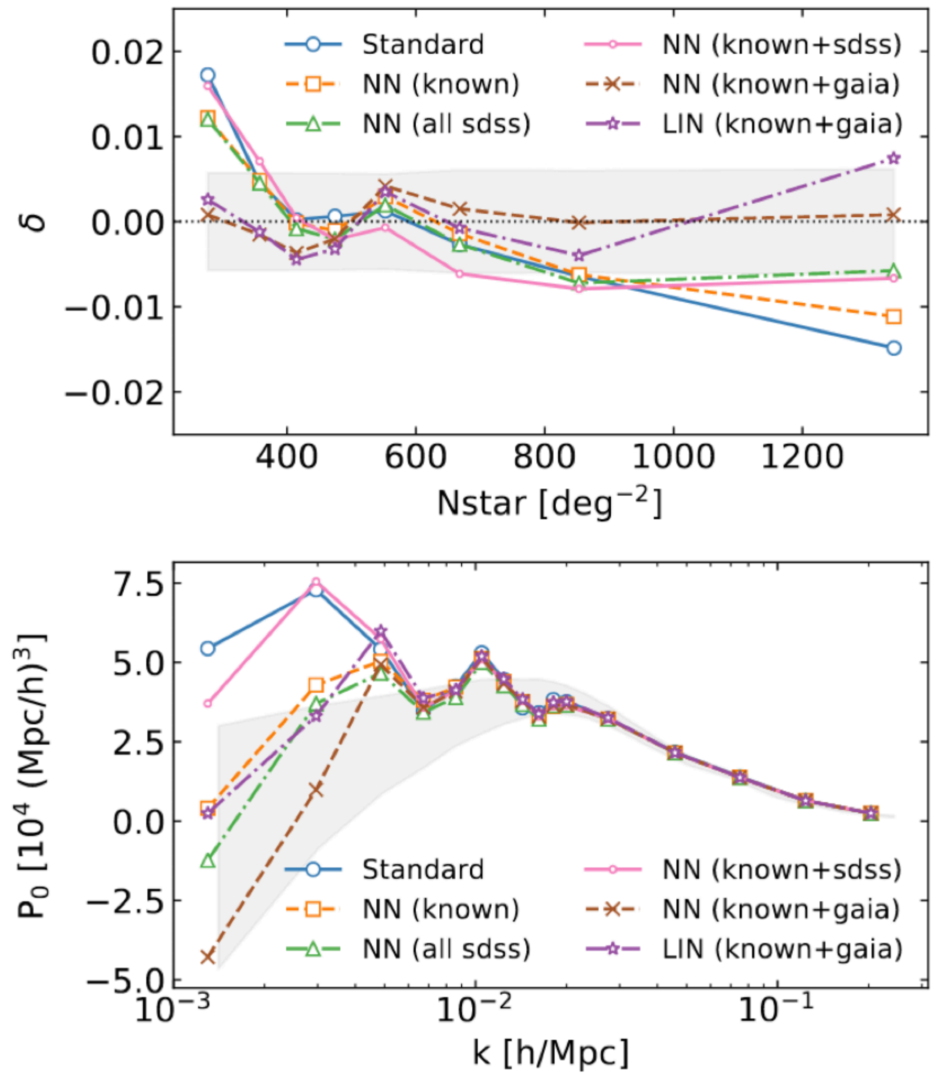
DES: Rodriguez-Monroy + 22

Catalog Systematics: Foregrounds

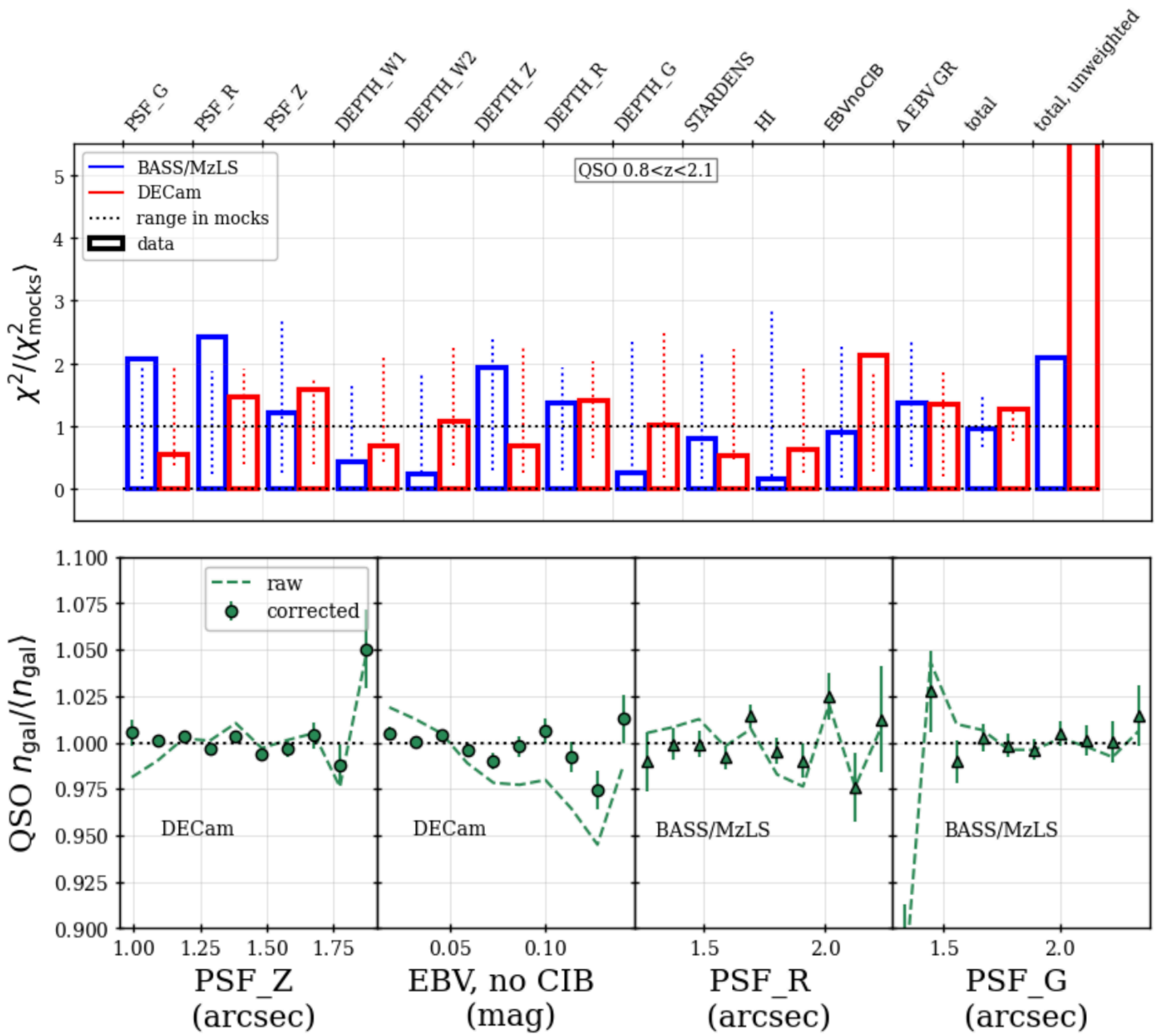
Rezaie+2020,2021 validated FCNN for non-linear weights of imaging sys on eBOSS ELGs/quasars.

Example: mean density as a function of stellar density against expected variance from EZmocks without systematics.

Notable effect of lin/NN and choice of stellar density template on low-k power spectrum monopole.



Catalog Systematics: Foregrounds

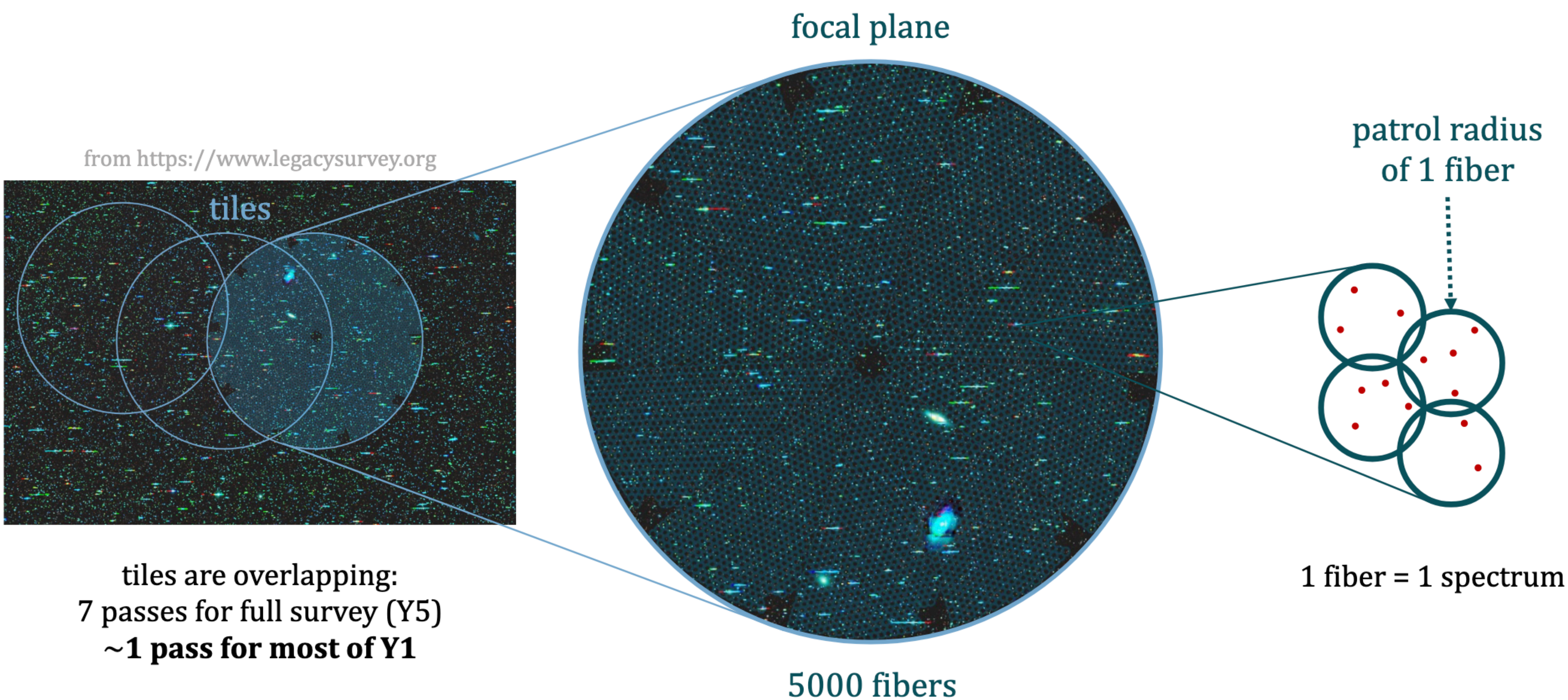


Identify most significant foregrounds for each sample (ELG, LRG, QSO)

Sample density after multi-dimensional foreground regression: assign weight w_{imsys} to each galaxy

DESI 2025 (2411.12020)

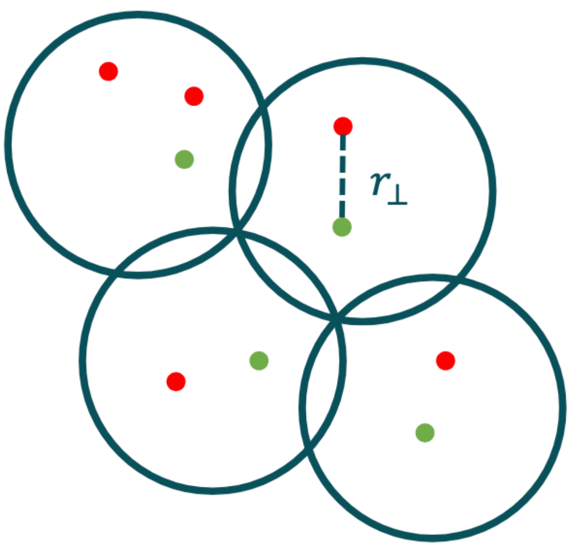
Catalog Systematics: Fiber Collisions

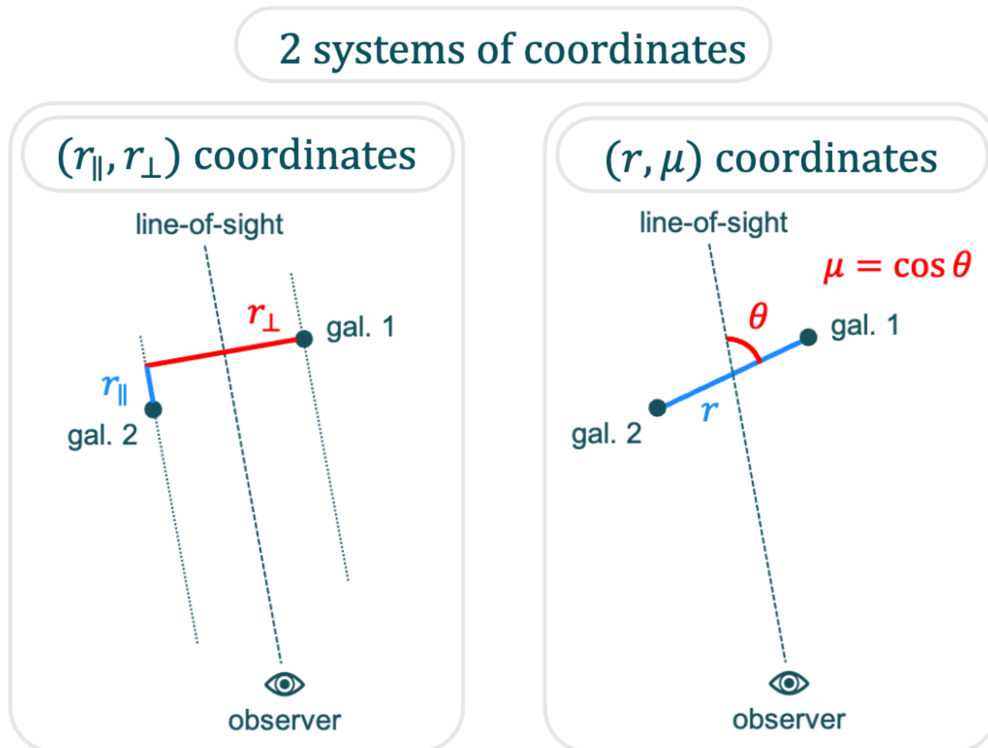


Credit: Mathilde Pinon

Catalog Systematics: Fiber Collisions

observed galaxies
missed galaxies

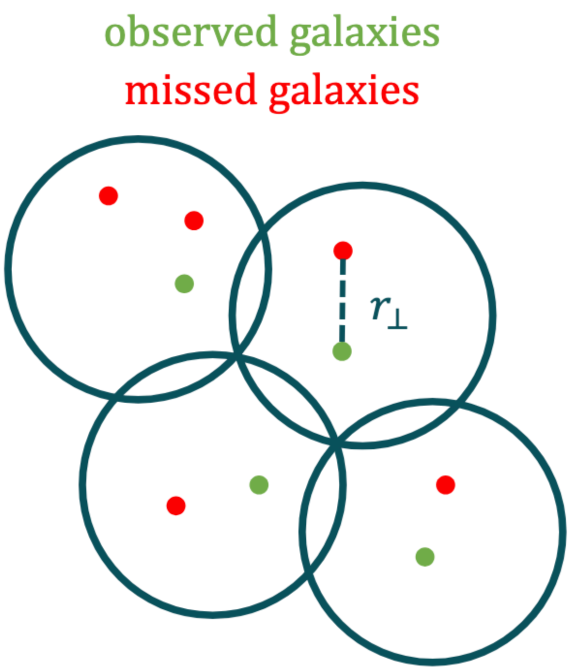




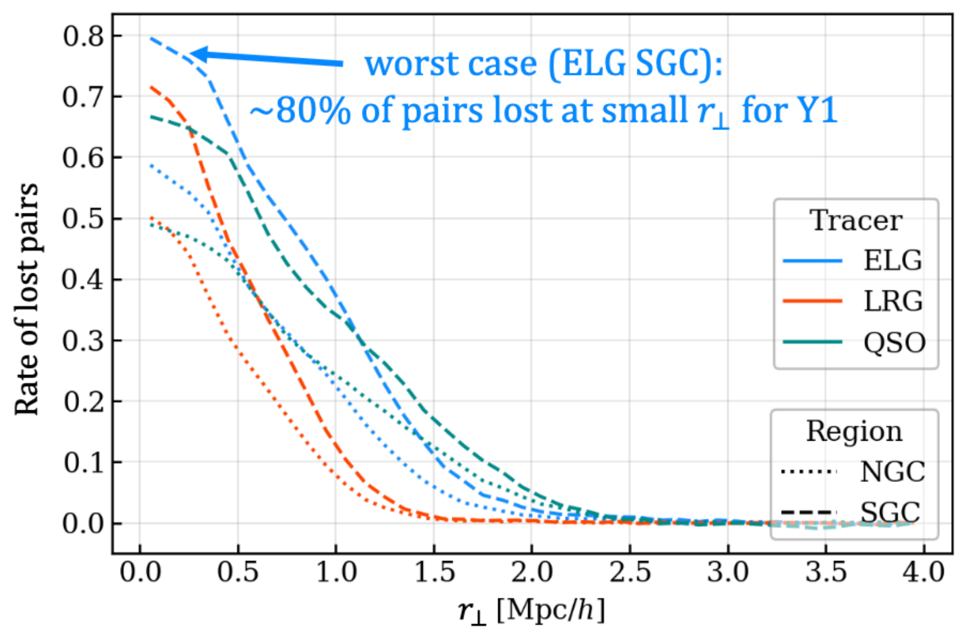
usually used
(Legendre decomposition)

Credit: Mathilde Pinon

Catalog Systematics: Fiber Collisions

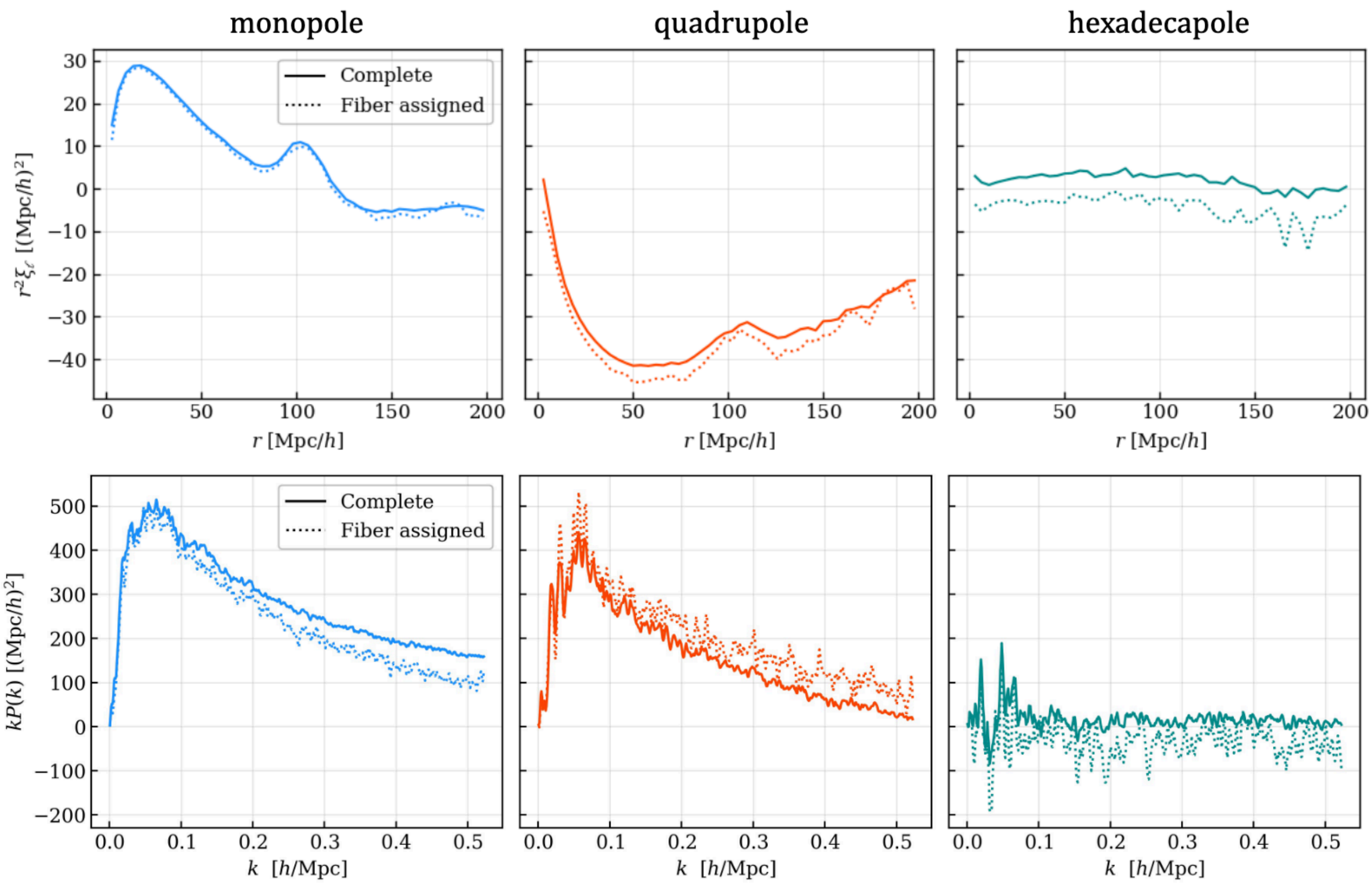


missing galaxy pairs
at small transverse separation $r_\perp$



Credit: Mathilde Pinon

Catalog Systematics: Fiber Collisions



correlation function multipoles

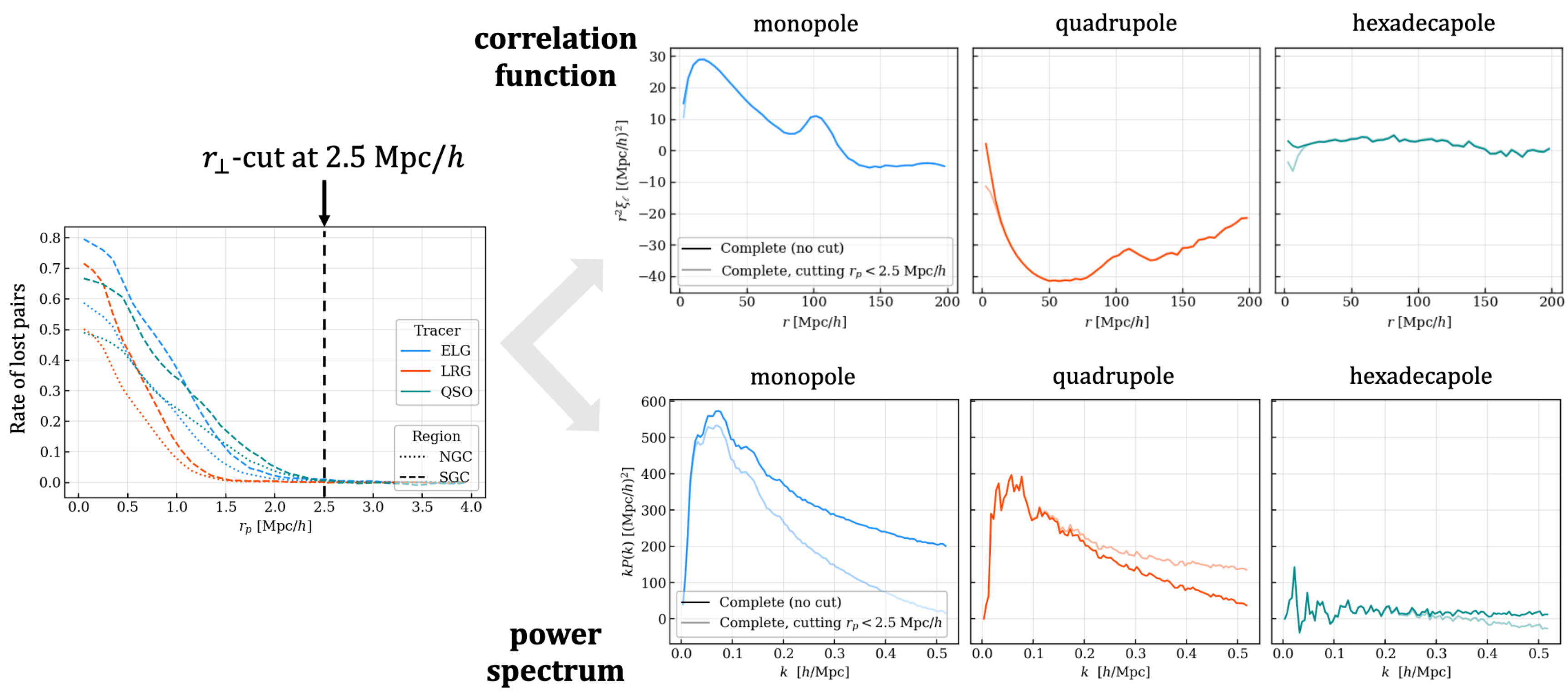
power spectrum multipoles

Pinon+2025

Catalog Systematics: Fiber Collisions

DESI2025: modify 2-pt estimators by removing all galaxy pairs at small transverse separation

→ must also be accounted for in theoretical modeling



Note: does not generalize to beyond-2pt summary statistics

Pinon+2025

Catalog Systematics

Imaging systematics and fiber collisions are just two prominent examples for catalog systematics – there are many more, known unknowns and unknown unknowns!

Quantifying the significance of a systematic effect and mitigating its impact may each amount to many person years of effort.

Increasing statistical precision of data sets puts more and more stringent requirements on systematic mitigation algorithms.

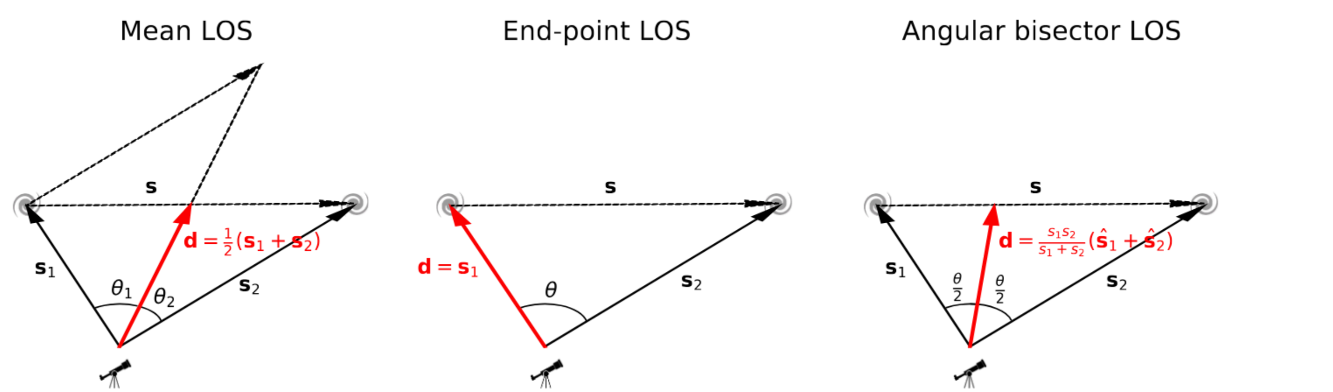
Note that requirements on systematics may differ for different summary statistics, and different science cases!

Power Spectrum Estimators

Power spectrum estimation in redshift space (Yamamoto+06)

On large scales, need to pay close attention to geometry + wide angle effects…

$$P_\ell(k) = \frac{(2\ell+1)}{I} \int \frac{d\Omega_k}{4\pi} \left[\int d^3r_1 \int d^3r_2\, F(\mathbf{r}_1) F(\mathbf{r}_2) e^{i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)} \mathcal{L}_\ell(\hat{\mathbf{k}}\cdot\hat{\eta}) - S_\ell(\mathbf{k}) \right]$$



Alternatives to FKP & related estimators:

- window-less quadratic estimators (Philcox 20ff)
- spherical Fourier-Bessel decomposition (Binney&Quinn 91; Leistedt+13,…)

Covariance Details

Redshift-space galaxy power spectrum

$$\mathbf{C}^{\mathrm{G}}_{\ell_1\ell_2}(k_1,k_2) = \frac{(2\ell_1+1)(2\ell_2+1)}{\mathrm{I}_{22}^2}\int_{\hat{\mathbf{k}}_1,\hat{\mathbf{k}}_2,\mathbf{x}_1,\mathbf{x}_2} P_{\mathrm{local}}(\mathbf{k}_2;\mathbf{x}_1)P_{\mathrm{local}}(\mathbf{k}_1;\mathbf{x}_2)$$
$$\times W_{22}(\mathbf{x}_1)W_{22}(\mathbf{x}_2)\, e^{-i(\mathbf{x}_1-\mathbf{x}_2)\cdot(\mathbf{k}_1-\mathbf{k}_2)}\mathcal{L}_{\ell_1}(\hat{\mathbf{x}}_1\cdot\hat{\mathbf{k}}_1)\Big[\mathcal{L}_{\ell_2}(\hat{\mathbf{x}}_2\cdot\hat{\mathbf{k}}_2)+\mathcal{L}_{\ell_2}$$
$$\simeq \sum_{\ell_1',\ell_2'} P_{\ell_1'}(k_1)P_{\ell_2'}(k_2)\Big\{\frac{(2\ell_1+1)(2\ell_2+1)}{\mathrm{I}_{22}^2}\int_{\hat{\mathbf{k}}_1,\hat{\mathbf{k}}_2,\mathbf{x}_1,\mathbf{x}_2} W_{22}(\mathbf{x}_1)W_{22}(\mathbf{x}_2)\, e^{-i(\mathbf{x}_1-\mathbf{x}_2)\cdot(\mathbf{k}_1-\mathbf{k}_2}$$
$$\times \mathcal{L}_{\ell_1}(\hat{\mathbf{x}}_1\cdot\hat{\mathbf{k}}_1)\mathcal{L}_{\ell_1'}(\hat{\mathbf{x}}_2\cdot\hat{\mathbf{k}}_1)\mathcal{L}_{\ell_2'}(\hat{\mathbf{x}}_1\cdot\hat{\mathbf{k}}_2)\Big[\mathcal{L}_{\ell_2}(\hat{\mathbf{x}}_2\cdot\hat{\mathbf{k}}_2)+\mathcal{L}_{\ell_2}(\hat{\mathbf{x}}_1\cdot\hat{\mathbf{k}}_2)$$
$$\equiv \sum_{\ell_1',\ell_2'} P_{\ell_1'}(k_1)\, P_{\ell_2'}(k_2)\; \mathcal{W}^{(1)}_{\ell_1,\ell_2,\ell_1',\ell_2'}(k_1,k_2)$$

$$\mathbf{C}^{\mathrm{T}_0}_{\ell_1\ell_2}(k_1,k_2) = \frac{1}{\mathrm{I}_{22}^2}\int_{\hat{\mathbf{k}}_{\ell_1},\hat{\mathbf{k}}_{\ell_2},\boldsymbol{\epsilon}} |W_{22}(\boldsymbol{\epsilon})|^2\Big\{\Big[8P_{\mathrm{L}}^2(\mathbf{k}_1)Z_1^2(\mathbf{k}_1)P_{\mathrm{L}}(\mathbf{k}_1+\mathbf{k}_2)Z_2^2(-\mathbf{k}_1,\mathbf{k}_1+\mathbf{k}_2)+(\mathbf{k}_1\leftrightarrow\mathbf{k}_2)\Big]$$
$$+16P_{\mathrm{L}}(\mathbf{k}_1)Z_1(\mathbf{k}_1)P_{\mathrm{L}}(\mathbf{k}_2)Z_1(\mathbf{k}_2)P_{\mathrm{L}}(\mathbf{k}_1+\mathbf{k}_2)Z_2(-\mathbf{k}_1,\mathbf{k}_1+\mathbf{k}_2)Z_2(-\mathbf{k}_2,\mathbf{k}_1+\mathbf{k}_2)$$
$$+\Big[12\,Z_1^2(\mathbf{k}_1)P_{\mathrm{L}}^2(\mathbf{k}_1)Z_1(\mathbf{k}_2)P_{\mathrm{L}}(\mathbf{k}_2)Z_3(\mathbf{k}_1,-\mathbf{k}_1,\mathbf{k}_2)+(\mathbf{k}_1\leftrightarrow\mathbf{k}_2)\Big]\Big\}$$
$$= \frac{\mathrm{I}_{44}}{\mathrm{I}_{22}^2}\int_{\hat{\mathbf{k}}_{\ell_1},\hat{\mathbf{k}}_{\ell_2}} P_{\mathrm{L}}(\mathbf{k}_1+\mathbf{k}_2)\Big[\Big(8P_{\mathrm{L}}^2(\mathbf{k}_1)Z_1^2(\mathbf{k}_1)Z_2^2(-\mathbf{k}_1,\mathbf{k}_1+\mathbf{k}_2)+(\mathbf{k}_1\leftrightarrow\mathbf{k}_2)\Big)$$
$$+16P_{\mathrm{L}}(\mathbf{k}_1)Z_1(\mathbf{k}_1)P_{\mathrm{L}}(\mathbf{k}_2)Z_1(\mathbf{k}_2)Z_2(-\mathbf{k}_1,\mathbf{k}_1+\mathbf{k}_2)Z_2(-\mathbf{k}_2,\mathbf{k}_1+\mathbf{k}_2)\Big]$$
$$+\frac{\mathrm{I}_{44}}{\mathrm{I}_{22}^2}\int_{\hat{\mathbf{k}}_{\ell_1},\hat{\mathbf{k}}_{\ell_2}}\Big[12\,Z_1^2(\mathbf{k}_1)P_{\mathrm{L}}^2(\mathbf{k}_1)Z_1(\mathbf{k}_2)P_{\mathrm{L}}(\mathbf{k}_2)Z_3(\mathbf{k}_1,-\mathbf{k}_1,\mathbf{k}_2)+(\mathbf{k}_1\leftrightarrow\mathbf{k}_2)\Big]$$

Wadekar & Scoccimarro 2020

Projected 2pt statistics + clusters

$$\mathrm{Cov}^{\mathrm{G}}\left(C^{ij}_{AB}(l_1),C^{kl}_{CD}(l_2)\right) = \frac{4\pi\delta_{l_1l_2}}{\Omega_{\mathrm{s}}(2l_1+1)\Delta l_1}\left[\left(C^{ik}_{AC}(l_1)+\delta_{ik}\delta_{AC}N^i_A\right)\left(C^{jl}_{BD}(l_2)+\delta_{jl}\delta_{BD}N^j_B\right)+\left(C^{il}_{AD}(l_1)+\delta_{il}\delta_{AD}N^i_A\right)\left(C^{jk}_{BC}(l_2)+\delta_{jk}\delta_{BC}N^j_B\right)\right],$$

$$\mathrm{Cov}^{\mathrm{NG},0}\left(C^{ij}_{AB}(l_1),C^{kl}_{CD}(l_2)\right) = \frac{1}{\Omega_{\mathrm{s}}}\int_{|\mathbf{l}|\in l_1}\frac{d^2\mathbf{l}}{A(l_1)}\int_{|\mathbf{l}'|\in l_2}\frac{d^2\mathbf{l}'}{A(l_2)}\int \mathrm{d}\chi\,\frac{q^i_A(\chi)q^j_B(\chi)q^k_C(\chi)q^l_D(\chi)}{\chi^6}\; T^{ijkl}_{ABCD}(\mathbf{l}/\chi,-\mathbf{l}/\chi,\mathbf{l}'/\chi,-\mathbf{l}'/\chi;z(\chi))$$

$$\mathrm{Cov}\left(N^i_{\lambda_\alpha},N^j_{\lambda_\beta}\right) = \delta_{i,j}\delta_{\alpha,\beta}N^i_{\lambda_\alpha}+\Omega_{\mathrm{s}}^2\int d\chi\, q^i_{\lambda_\alpha}(\chi)q^j_{\lambda_\beta}(\chi)\left[\int dM\,\frac{dn}{dM}b_{\mathrm{h}}(M,z)\int_{\lambda_{\alpha,\mathrm{min}}}^{\lambda_{\alpha,\mathrm{max}}} d\lambda\, p(M|\lambda,z)\right]\left[\int dM'\,\frac{dn}{dM'}b_{\mathrm{h}}(M',z)\int_{\lambda_{\beta,\mathrm{min}}}^{\lambda_{\beta,\mathrm{max}}} d\lambda'\, p(M'|\lambda',z)\right],$$

$$\mathrm{Cov}\left(N^i_{\lambda_\alpha},C^{jk}_{AB}(l)\right) = \Omega_{\mathrm{s}}\int d\chi\,\frac{q^i_{\lambda_\alpha}(\chi)q^j_A(\chi)q^k_B(\chi)}{\chi^2}\left[\int dM\,\frac{dn}{dM}b_{\mathrm{h}}(M,z)\int_{\lambda_{\alpha,\mathrm{min}}}^{\lambda_{\alpha,\mathrm{max}}} d\lambda\, p(M|\lambda,z)\right]\frac{\partial P_{AB}(k,z(\chi))}{\partial\delta_b}\sigma_b(\Omega_{\mathrm{s}};z(\chi))$$

Krause & Eifler 2017