Mean density estimation for ~~Accurate~~ power spectrum estimation

Jun Koda
Università Degli Studi Roma Tre

Sesto, 5 July 2018

FKP estimator

FKP: Feldmann, Kaiser and Peacock 1994, *ApJ* 426 23

FKP estimator

$$F(\boldsymbol{x}) = w(\boldsymbol{x})\left[n_{\rm data}(\boldsymbol{x}) - \alpha n_{\rm rand}(\boldsymbol{x})\right]$$

α = data/random ratio

Minimum variance weight

$$w_{\rm FKP}(\boldsymbol{x}) = \frac{1}{1 + \bar{n}(\boldsymbol{x})P_{est}}$$

which minimises the statistical error

$$\sigma_P(k)/P(k)$$

(monopole)

FKP estimator

FKP: Feldmann, Kaiser and Peacock 1994, ApJ 426 23

$$\tilde{P}(\boldsymbol{k}) = \frac{1}{\mathcal{N}} F(\boldsymbol{k}) F(\boldsymbol{k})^* - \mathcal{S}$$

The estimated power spectrum is related to the true power spectrum via window function convolution

$$\tilde{P}(\boldsymbol{k}) = \int \frac{d^3k}{(2\pi)^3} \, |\hat{W}(\boldsymbol{k} - \boldsymbol{k}')|^2 P(\boldsymbol{k}')$$

$$W(\boldsymbol{x}) = w(\boldsymbol{x}) \bar{n}(\boldsymbol{x})$$

Mean density for the FKP weight

FKP weight

$$w_{\mathrm{FKP}}(\boldsymbol{x}) = \frac{1}{1 + \bar{n}(\boldsymbol{x}) P_{est}}$$

ensemble average, or the density without clustering

$$\bar{n}(\boldsymbol{x}) = \langle n_g(\boldsymbol{x}) \rangle$$

usually,

$$\bar{n}(\boldsymbol{x}) = \bar{n}(z) \times f(RA, Dec)$$

In Euclid, we cannot make such decomposition
Need to estimate density in **3D** from random catalogue

Density estimation

<u>Kernel density estimation</u>

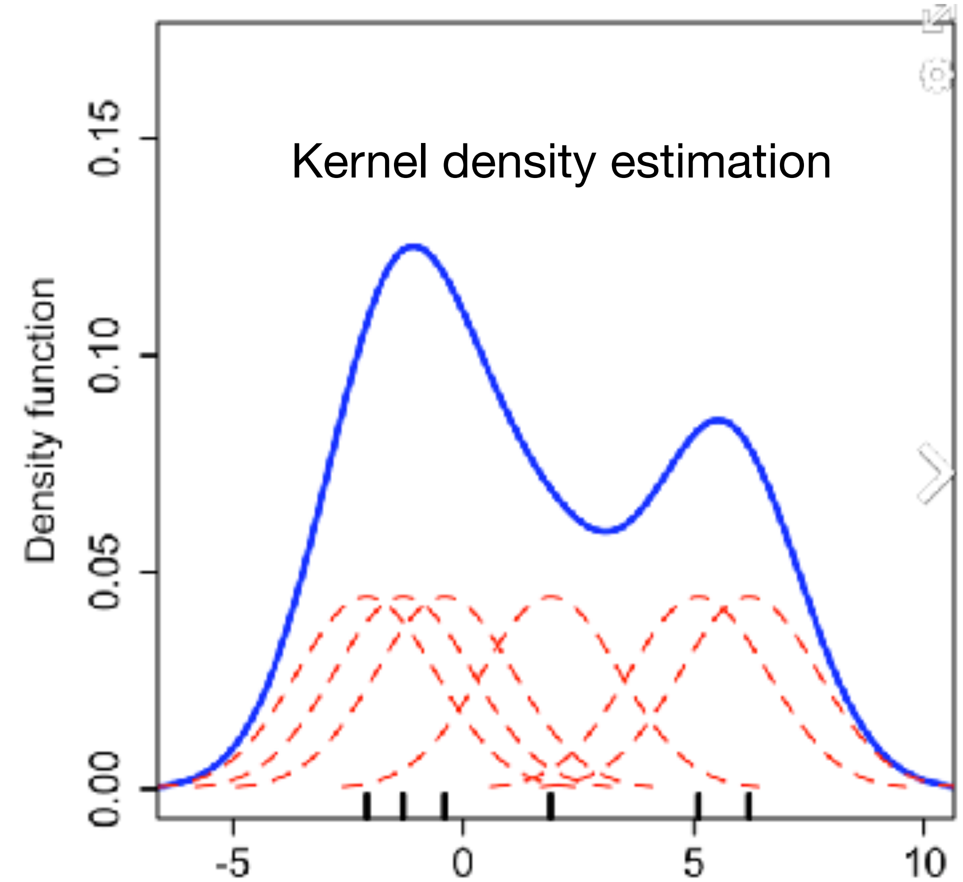
- Fixed kernel
  - PM simulation
  - P(k) estimation

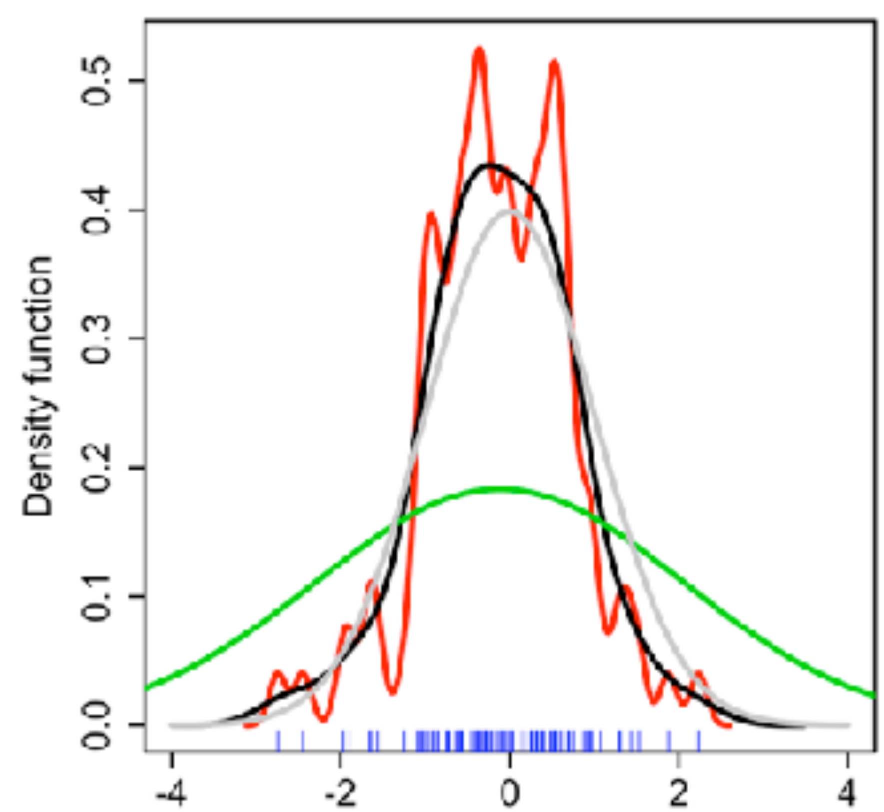
- Adaptive kernel
  - SPH hydro simulation

<u>Delaunay tessellation</u>

Statistical - systematic tradeoff

statistical fluctuation vs bias (over smoothing)





en.wikipedia.org/wiki/Kernel_density_estimation

How does the statistical/systematic error in mean density affect *P(k)* estimation?

$$w_{\mathrm{FKP}}(\boldsymbol{x}) = \frac{1}{1 + \bar{n}(\boldsymbol{x}) P_{est}}$$

1. Should not contribute to systematic error in *P(k)*

   everything is encoded in the window function

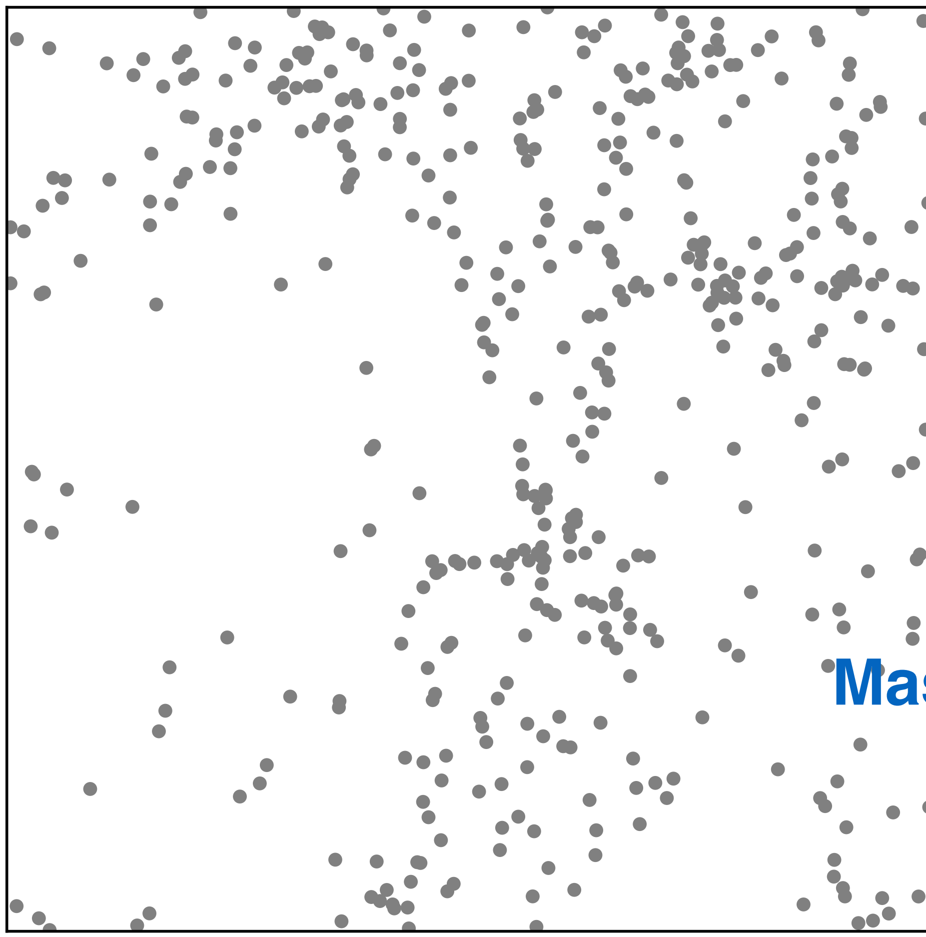
2. Increases the statistical error in P(k)

   because the weight is suboptimal

   but how significantly?

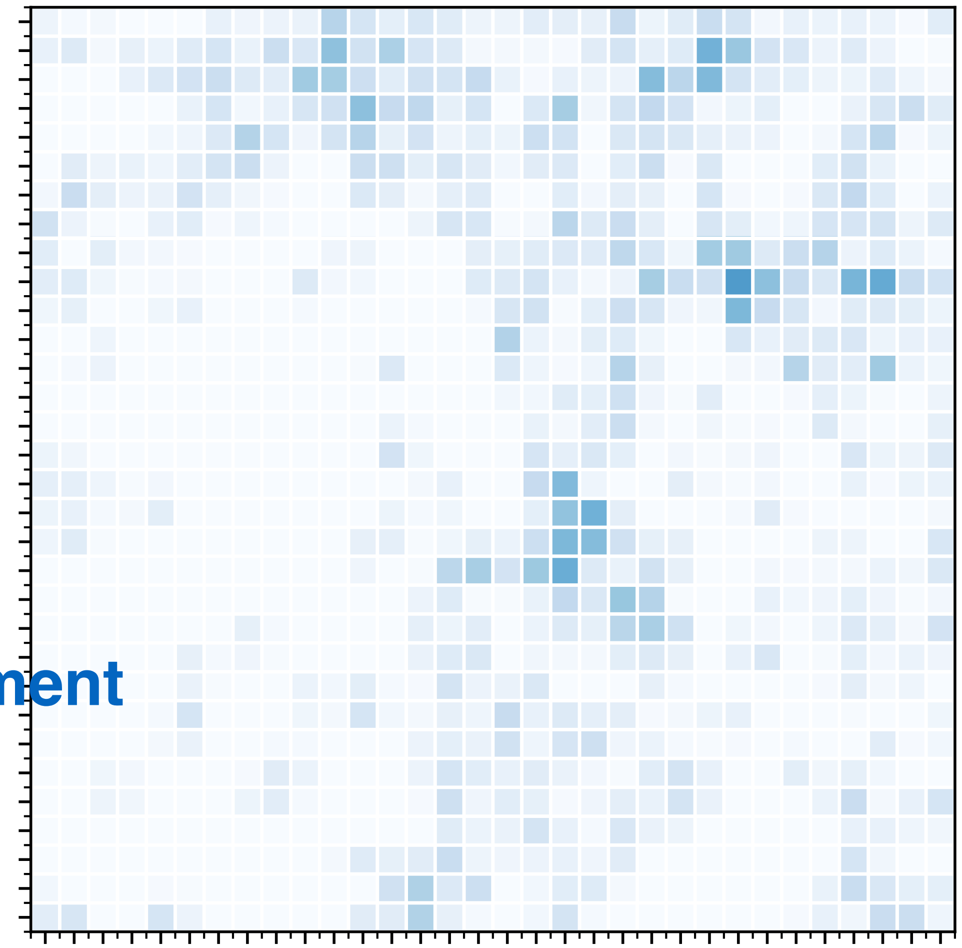
Using the FFT grid for nbar

Particles (galaxies/randoms)



Mass assignment

Density grid



FFT

$$F(\boldsymbol{k})$$

$$P_\ell(k)$$

Power spectrum multipoles

Using the FFT grid for nbar

Numbers

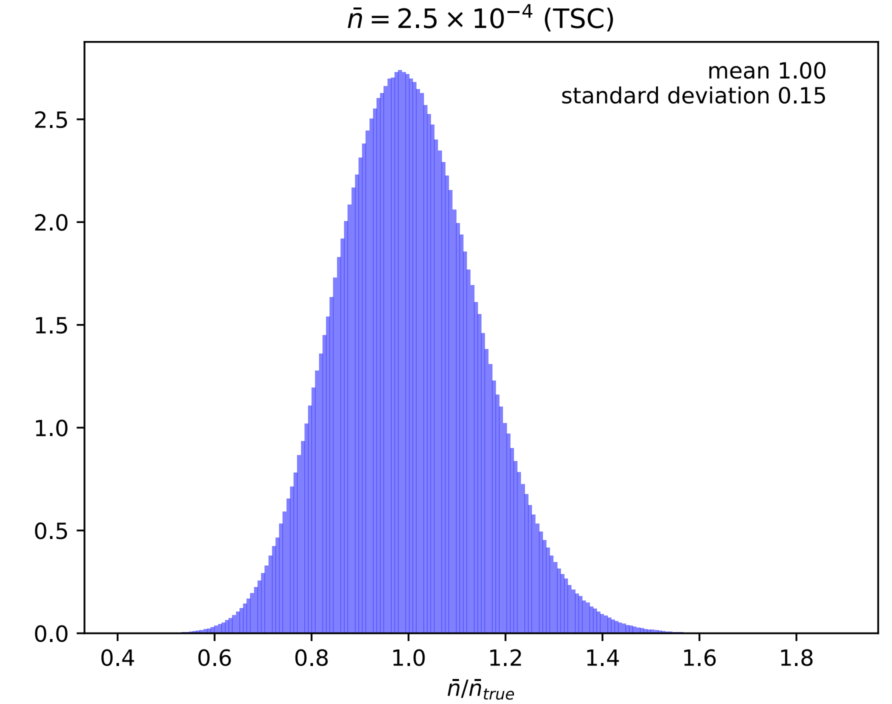
$$\Delta x \sim 2\pi \, h^{-1}\mathrm{Mpc}$$

$$k_{Nq} = \frac{\pi}{\Delta x} \sim 0.5 \, h\mathrm{Mpc}^{-1}$$

$$\bar{n} \sim 10^{-4} - 10^{-3} \, [h^{-1}\mathrm{Mpc}]^{-3}$$

$$\alpha^{-1} = 50 \qquad [\mathrm{random/data}]$$

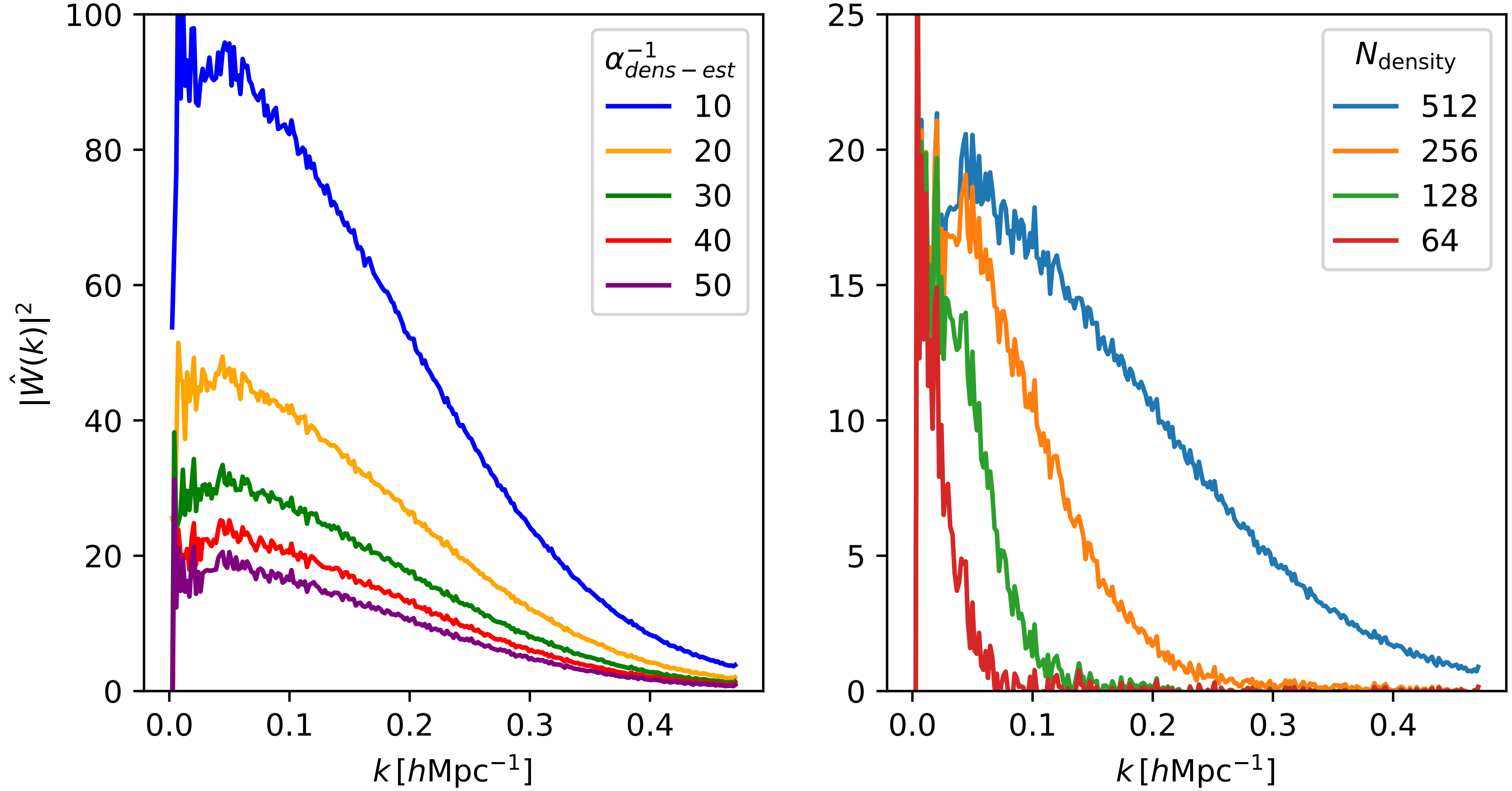
$$N_{rand} \text{ in cell} = 1 - 10 \text{ particles per cell}$$



nbar error effect on window function

Periodic box: true window function is δ_K at k = 0

change # random | change grid size



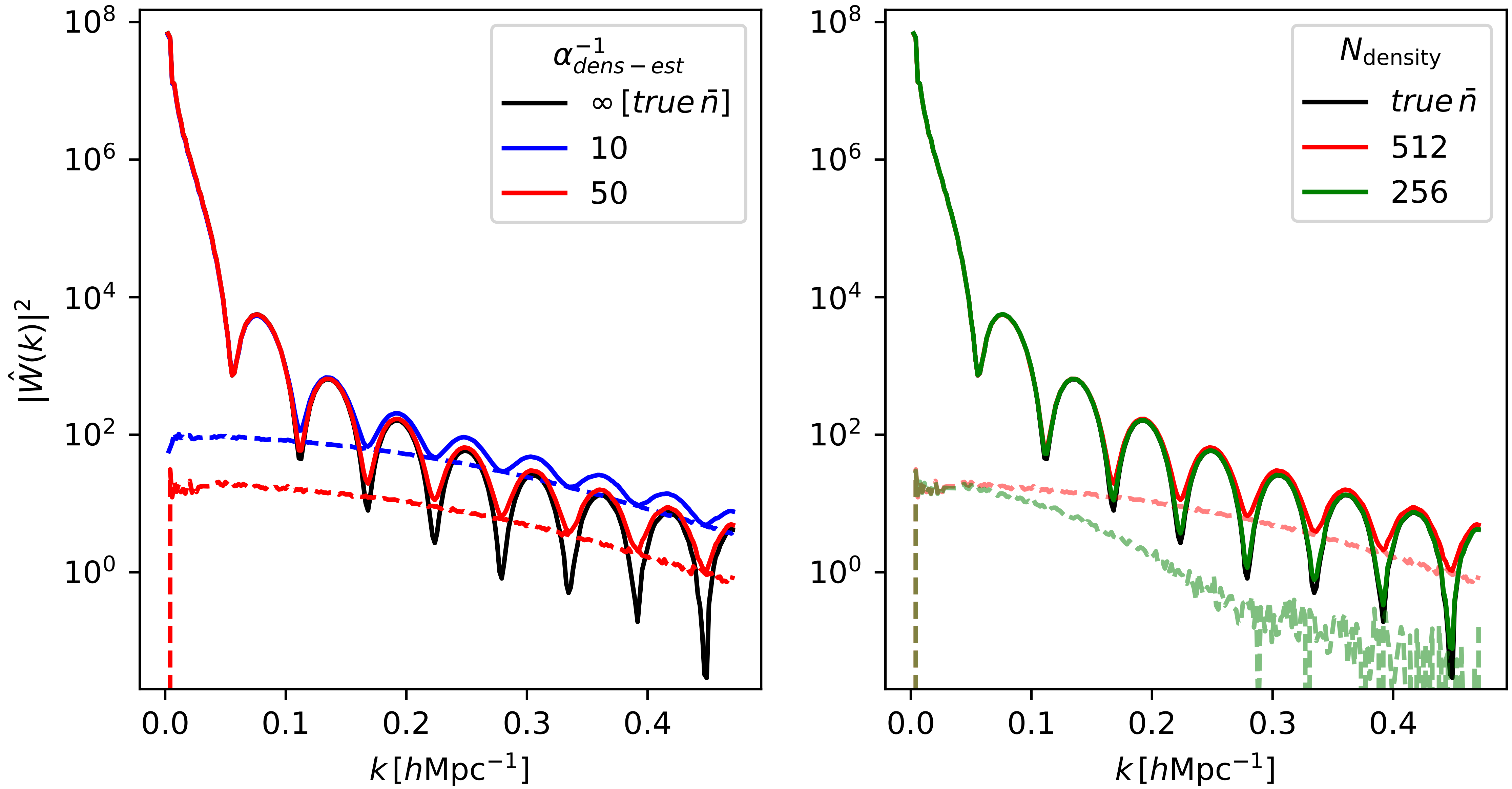
$$\langle |W(\boldsymbol{k})|^2 \rangle = \frac{\mathrm{Var}[\tilde{w}_a]}{\mathrm{Var}[\tilde{w}_a] + \langle \tilde{w}_a \rangle^2} V_{cell} \prod_{i=1}^{3} \hat{\chi}_i(k_i \epsilon), \quad \text{for } \boldsymbol{k} \neq 0. \tag{34}$$

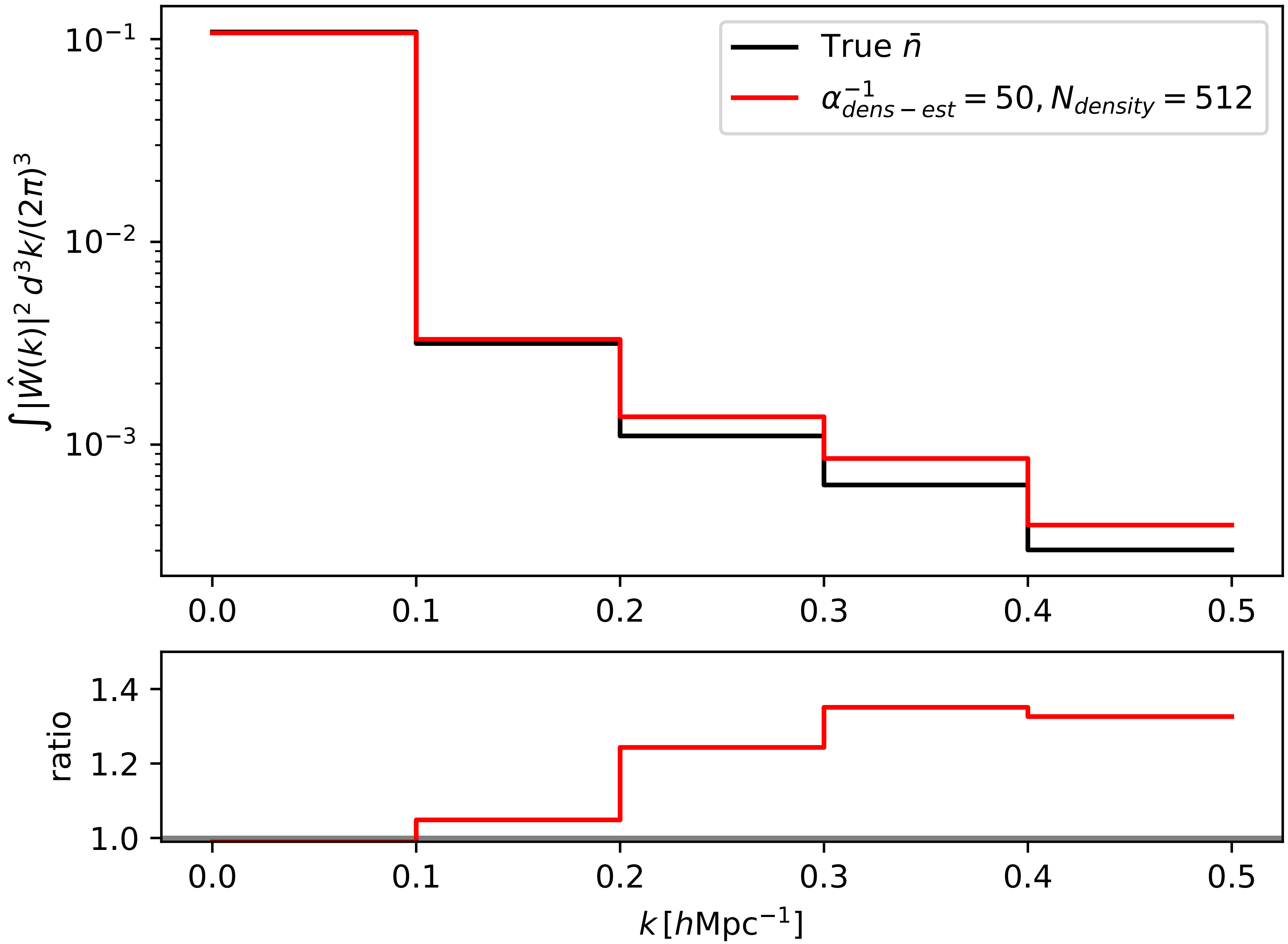
When the fluctuation is small, $\mathrm{Var}[\tilde{w}_a] \ll \langle w_a \rangle^2$, the amplitude of the window function is,

$$\frac{\mathrm{Var}[\tilde{w}_a]}{\langle \tilde{w}_a \rangle^2} V_{cell} = \left[\frac{\bar{n} P_{est}}{1 + \bar{n} P_{est}} \right] \bar{n}^{-1}_{dens-est}, \tag{35}$$

Statistical fluctuation term vs survey window function

Geometry: Flagship-like octant shell 1.7 < z < 1.8, n(z) = 4 × 10-4





$$\tilde{P}(\boldsymbol{k}) = \int \frac{d^3k}{(2\pi)^3} \, |\hat{W}(\boldsymbol{k} - \boldsymbol{k}')|^2 P(\boldsymbol{k}')$$

$$W(\boldsymbol{x}) = w(\boldsymbol{x})\bar{n}(\boldsymbol{x})$$

density error adds negligible mode mixing

nbar error effect on *P(k)* variance

$$\frac{\sigma_P^2(k)}{P^2(k)} = \frac{(2\pi)^3 \int d^3r\, \bar{n}^4 w^4 [1 + 1/\bar{n}P(k)]^2}{V_k [\int d^3r\, \bar{n}^2 w^2]^2} \,. \qquad (2.3.2)$$

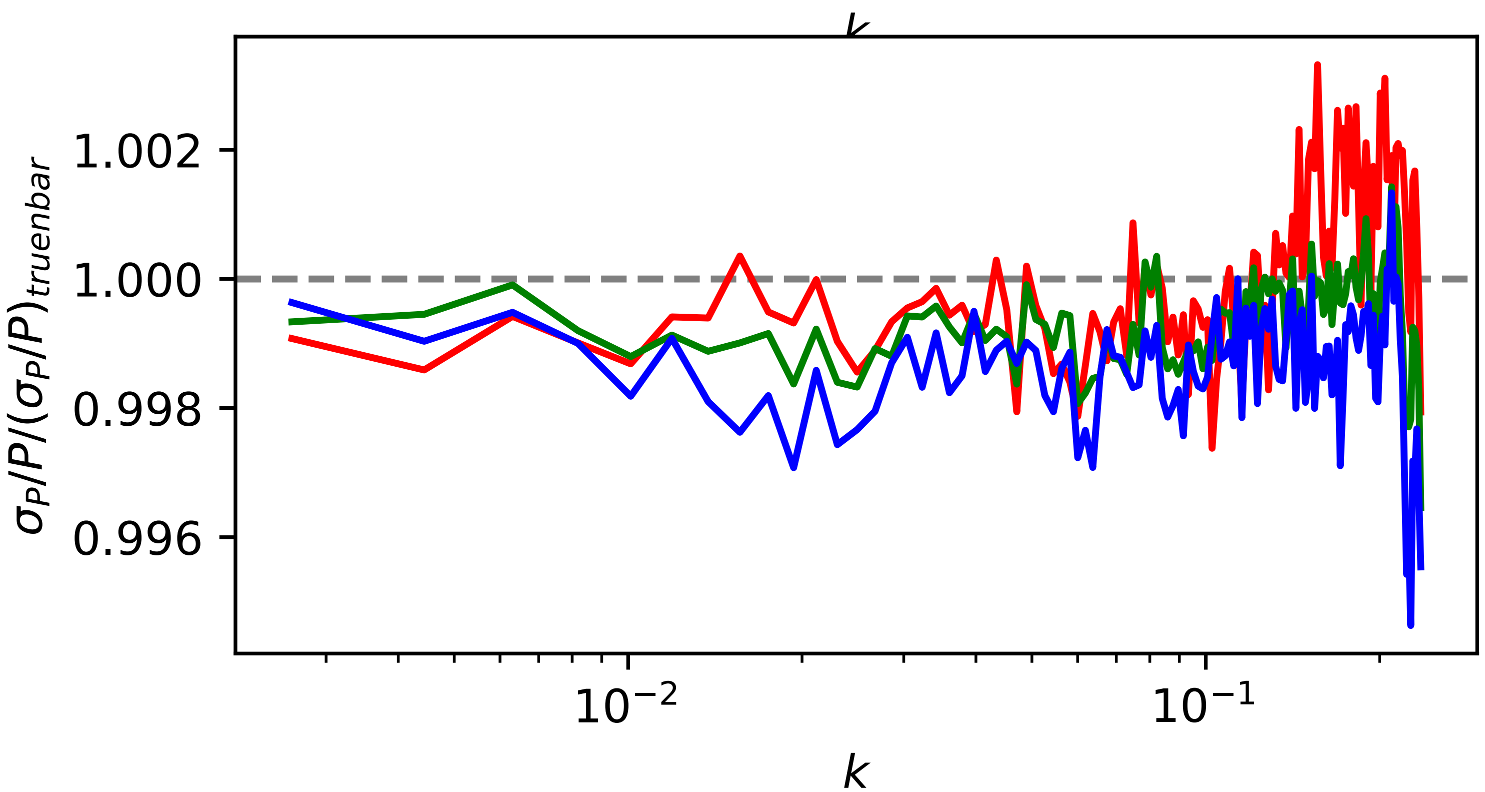
$$\frac{\sigma_P^2}{P^2} \propto \frac{\langle w^4 \rangle}{\langle w^2 \rangle^2}$$

only 2% increase in statistical error for fiducial setup

$\alpha_{dens-est}^{-1}$	precision ratio	$N_{density}$	precision ratio
10	1.12	512	1.022
20	1.057	256	1.0027
40	1.028	128	1.00034
50	1.022	64	1.000043

Lognormal mocks with Euclid Flagship-like octant

redshift slice 1.7 < z < 1.8, no angular mask, 1000 realisations



$N_{grid} = 256 \Longleftrightarrow \Delta x = 13.3\,h^{-1}\mathrm{Mpc}$

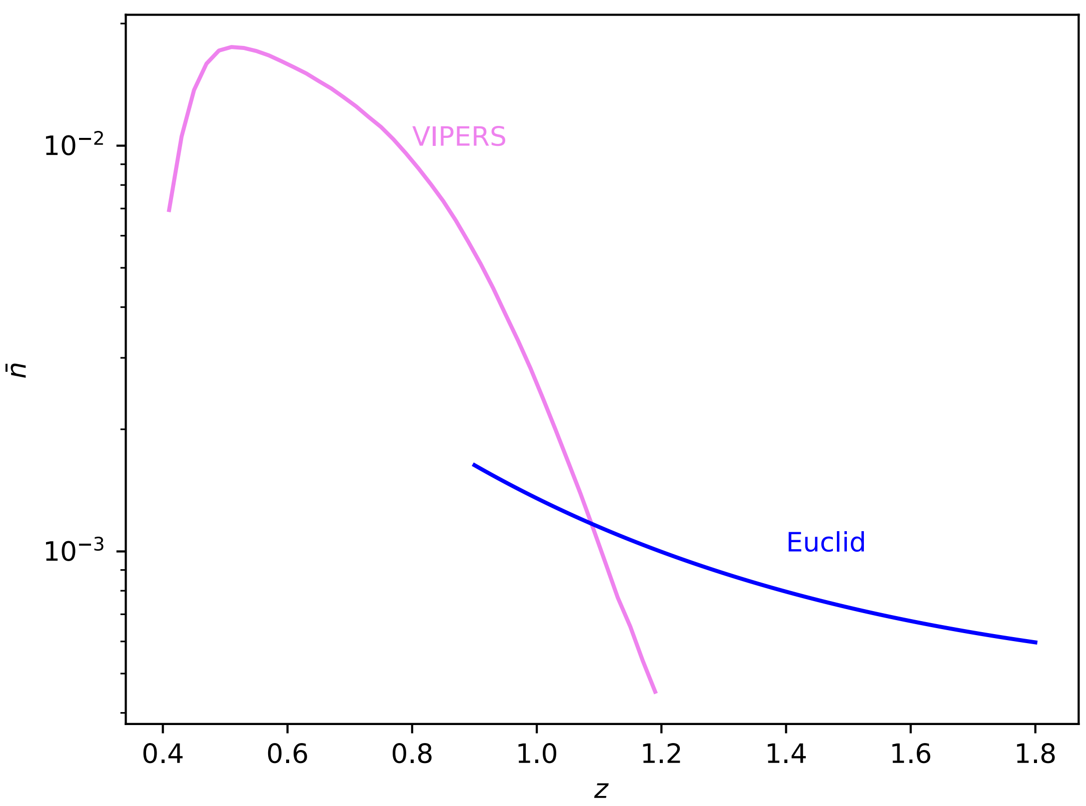
$P_{est} = 4000 \quad [(h^{-1}\mathrm{Mpc})^3]$

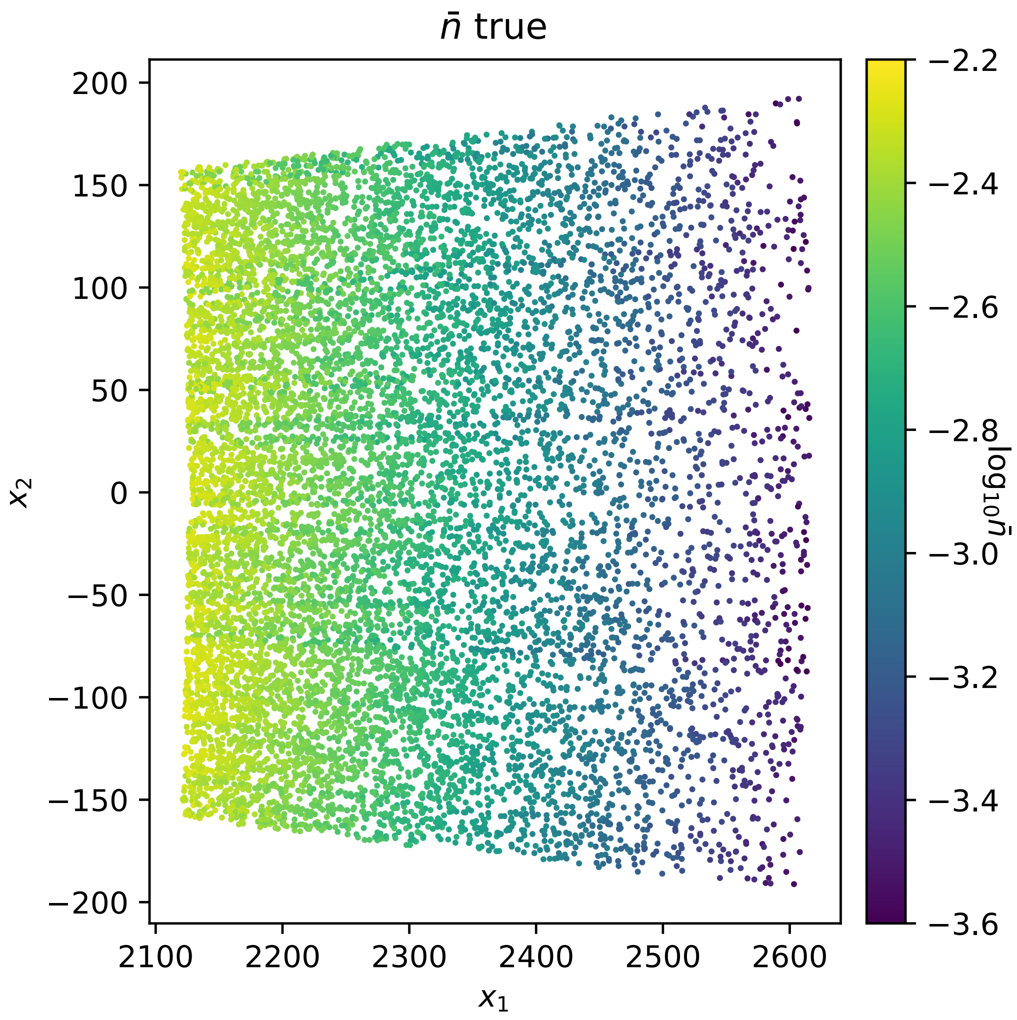
$$\frac{\sigma_P^2}{P^2} \propto \frac{\langle w^4 \rangle}{\langle w^2 \rangle^2}$$

$\alpha^{-1}_{dens-est}$	precision ratio	$N_{density}$	precision ratio
10	1.12	512	1.022
20	1.057	256	1.0027
40	1.028	128	1.00034
50	1.022	64	1.000043

Exercise with VIPERS COLA mocks

$$\bar{n}(z)$$



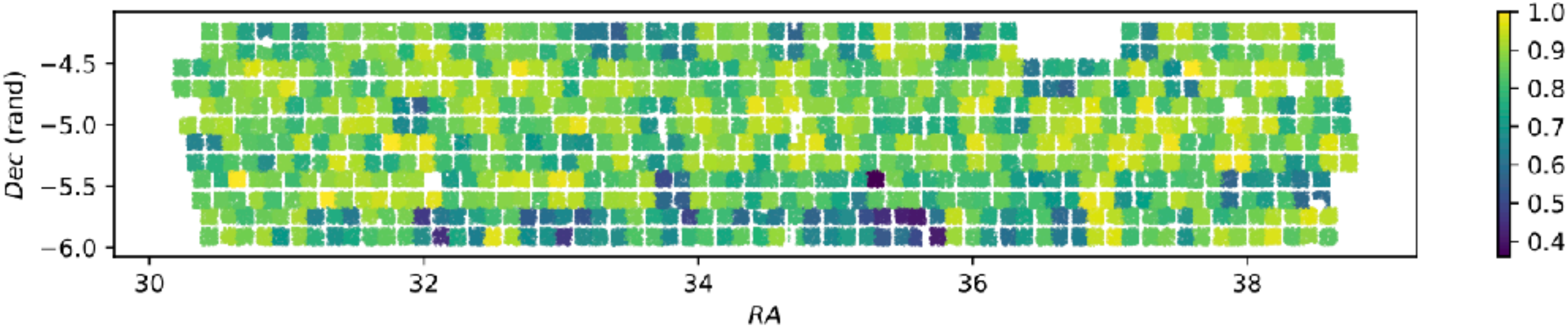


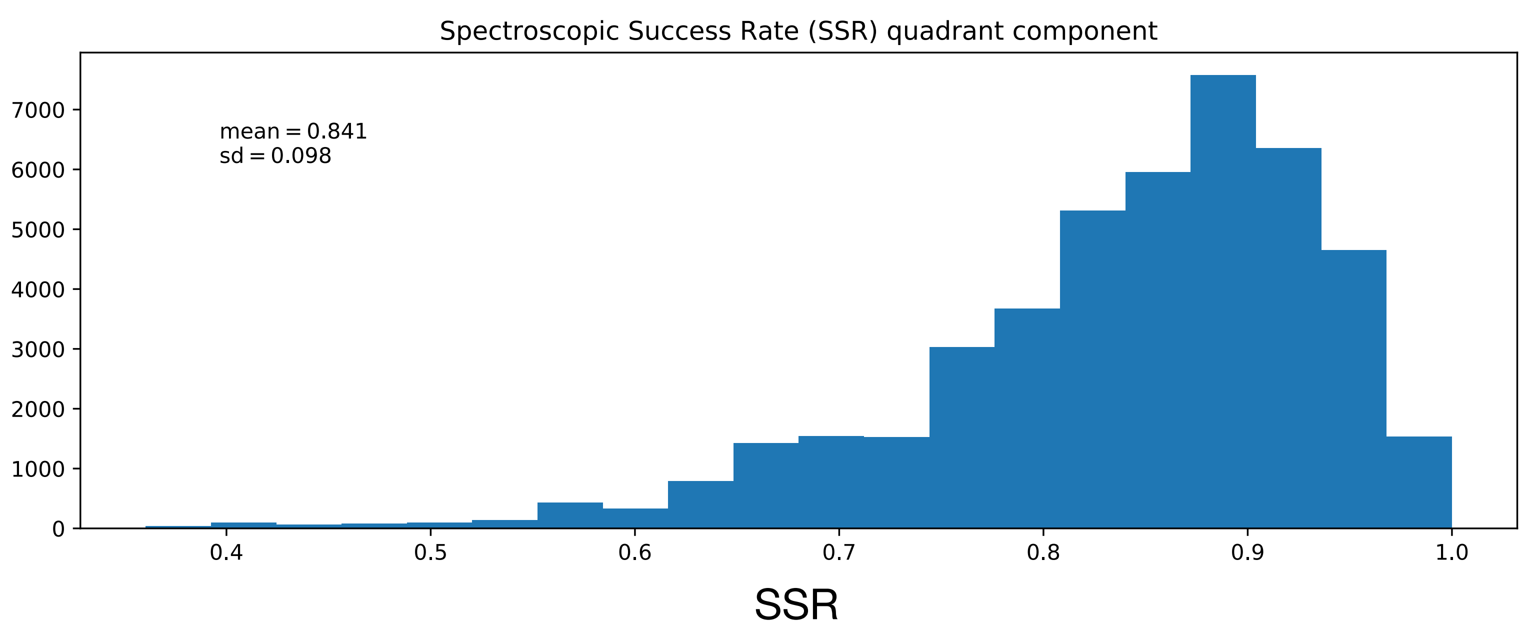
W1, high z bin (0.9 < z < 1.2)

Exercise with VIPERS COLA mocks

angular modulation of the mean number density

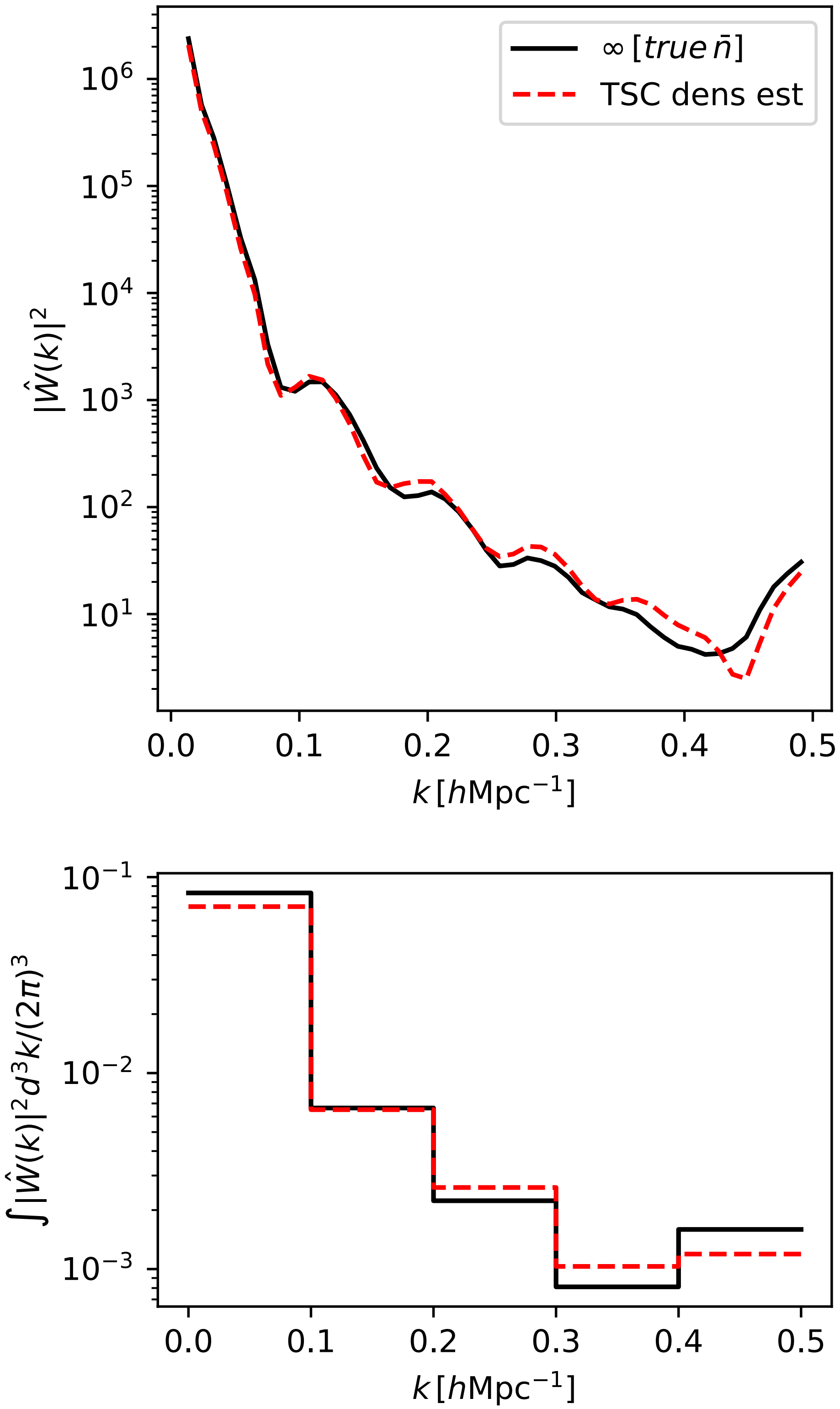
Spectroscopic Success Rate (SSR)



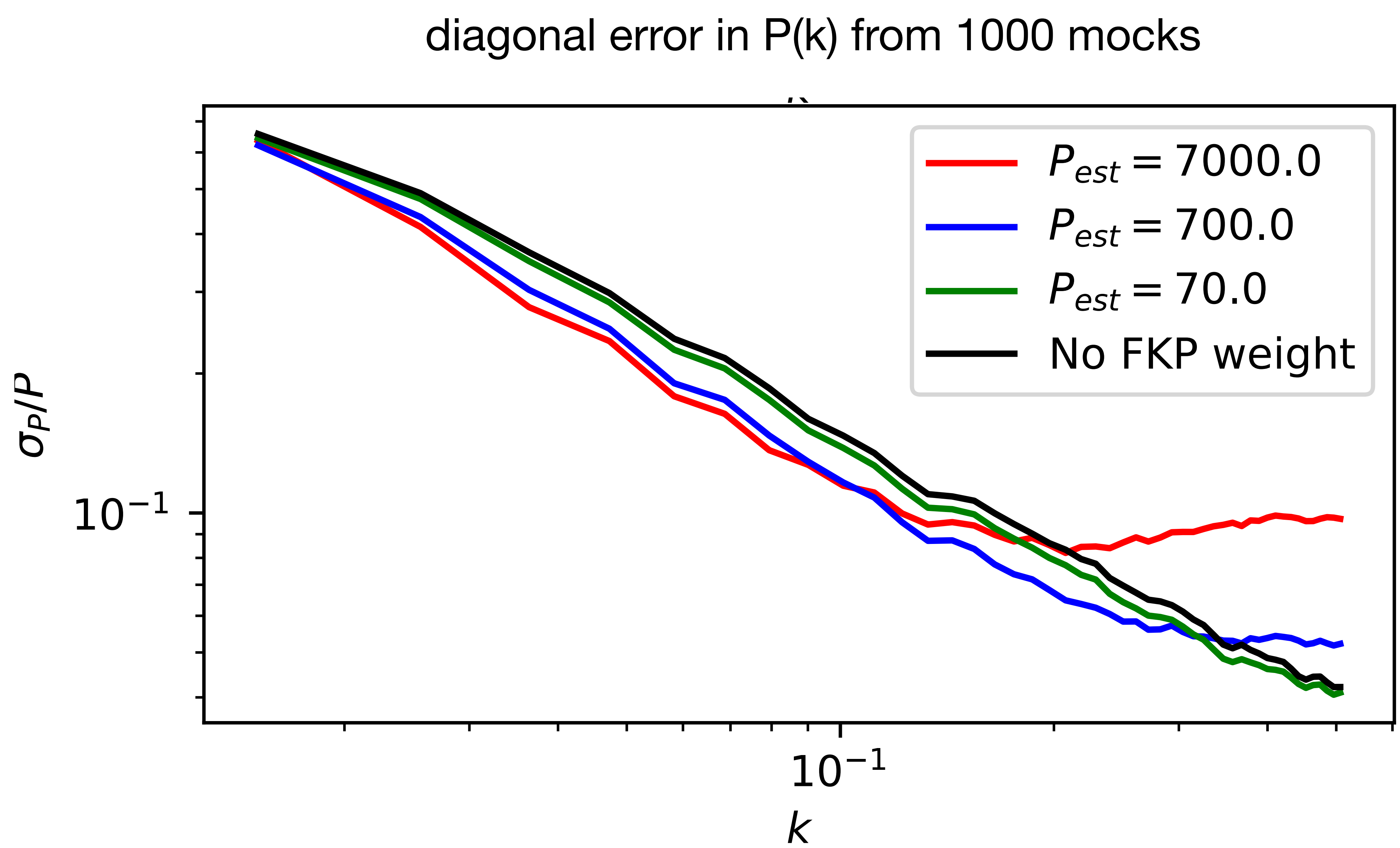


* quadrant dependence and redshift dependence is applied to the randoms

Window function with estimated nbar



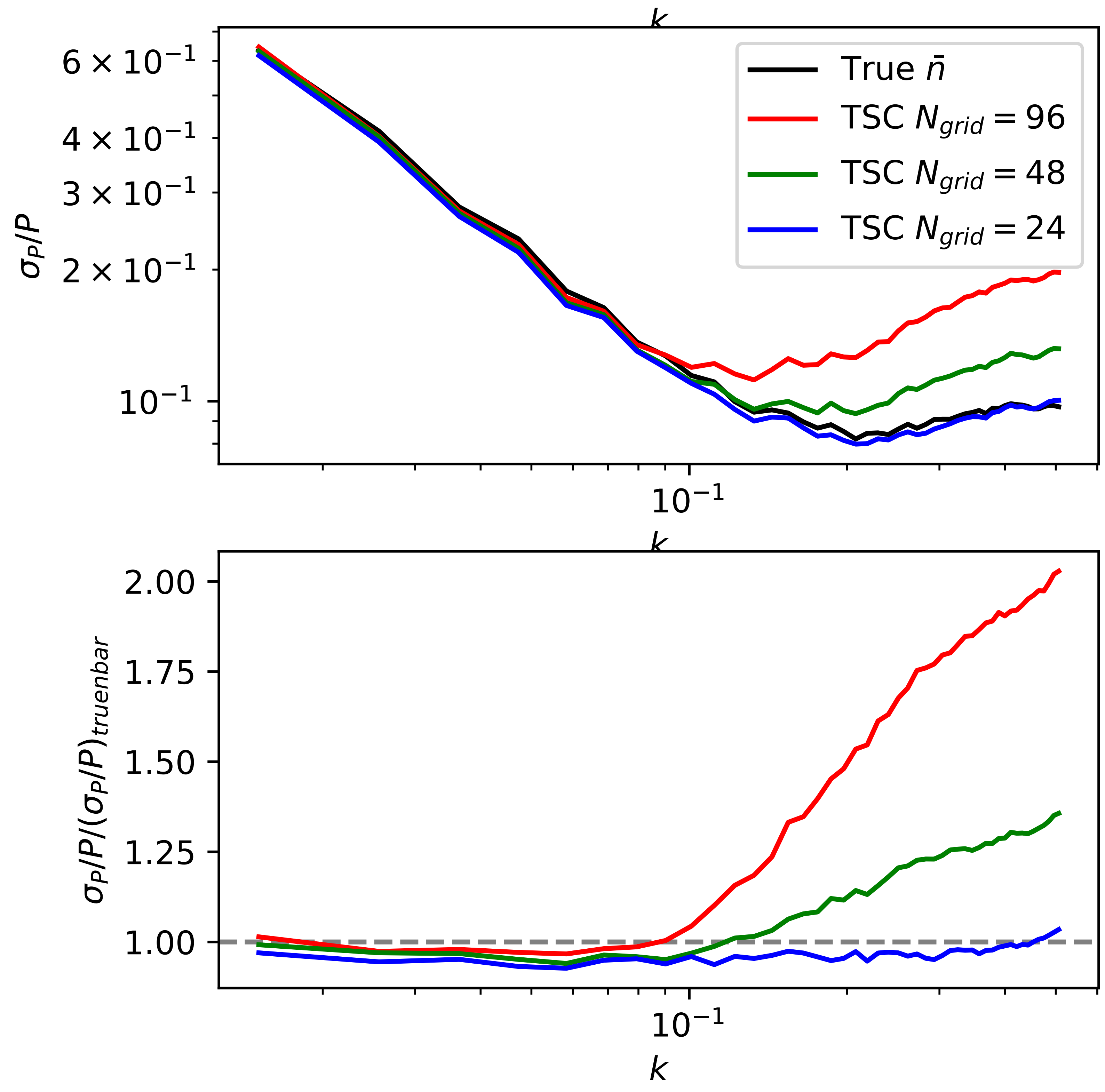
P(k) statistical error with / without FKP weights



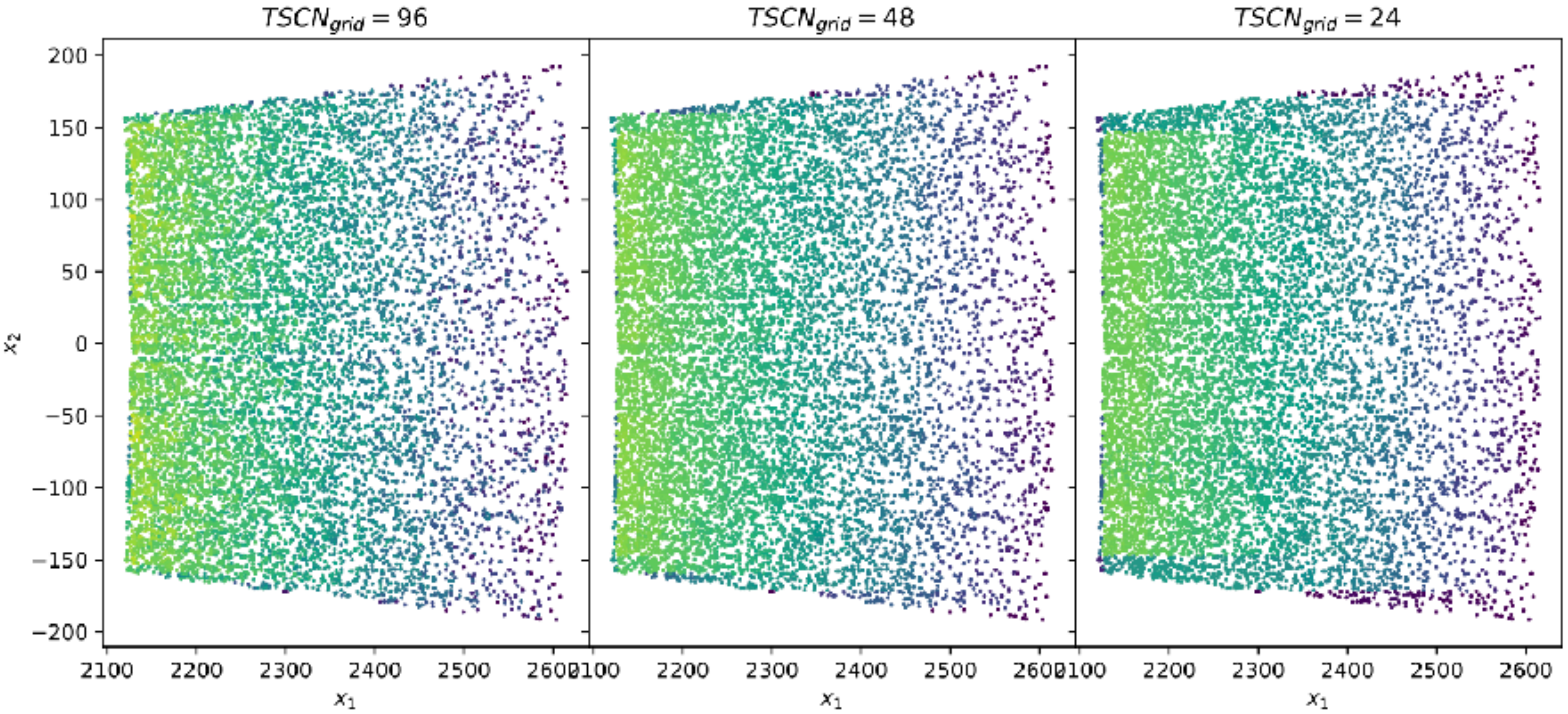
$$w_{\rm FKP}(\boldsymbol{x}) = \frac{1}{1 + \bar{n}(\boldsymbol{x})P_{est}}$$

True $\bar{n}(z)$ is used here

P(k) statistical error with estimated nbar



estimated mean density using TSC



Systematic error from coarse grid / large kernel is not affecting the P(k) error

Summary

- Euclid requires 3D density estimation for FKP estimator;
- scatter in density have negligible impact on window function; no significant additional mode mixing;
- scatter in density may or may not increase the *P(k)* error
  - could be a factor of 2 in *P(k)* error [VIPERS mock]
  - smooth the density field if necessary; systematic error (over smoothing) in mean density seems harmless for *P(k)* error
  - Using FFT grid seems for mean density is fine for Euclid, but worth checking with more realistic mocks