Analytical covariance of velocity correlations

Chris Blake, DESI PV telecon update, 26 Oct 2023

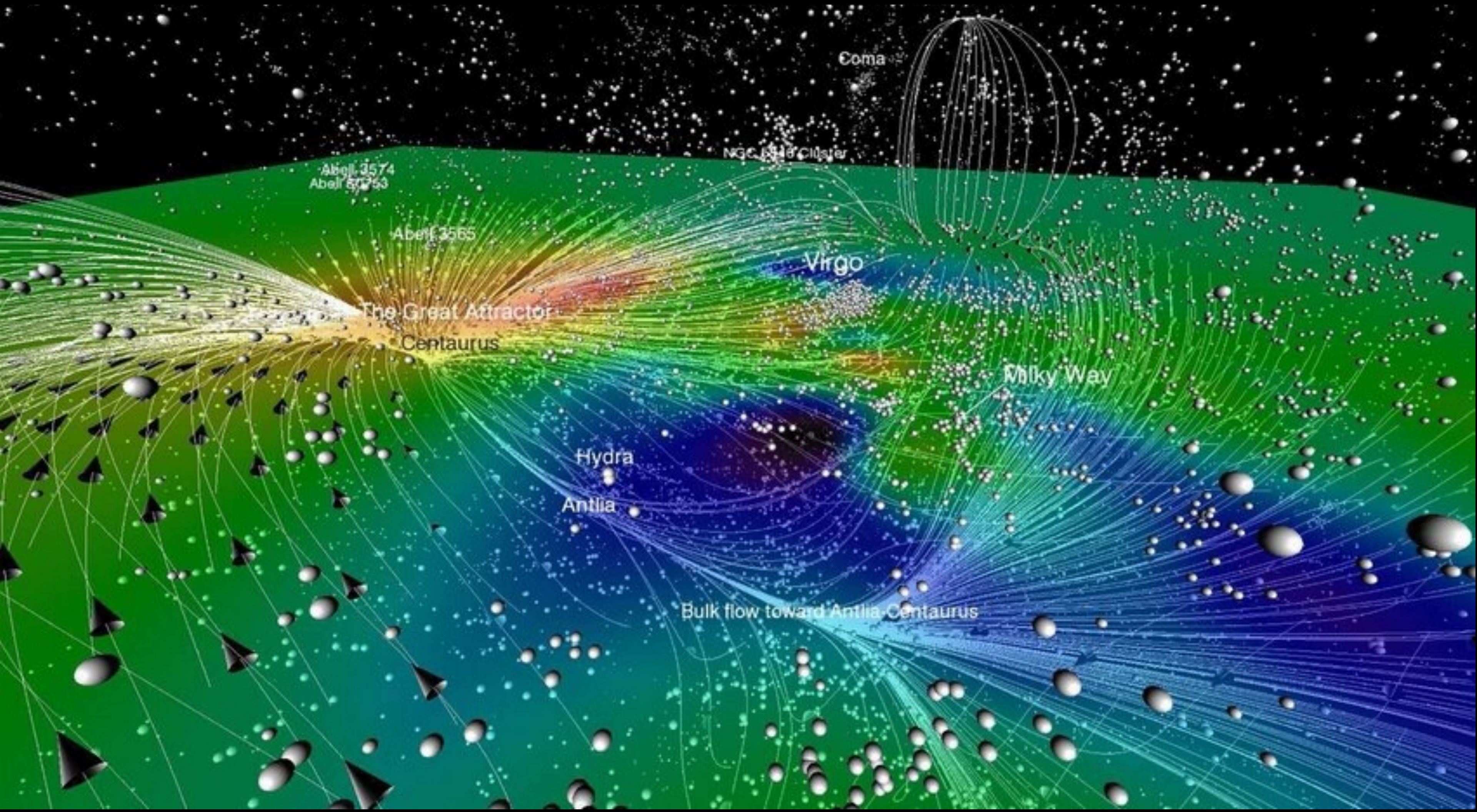


Image credit: Courtois, Pomarede, Tully et al., Cosmic Flows

Analytical covariance

- The covariance describes the fluctuations and correlations of our measurements between different statistics and scales:

$$\mathrm{Cov}[\xi_1(r), \xi_2(s)] = \langle \xi_1(r)\xi_2(s)\rangle - \langle \xi_1(r)\rangle\langle \xi_2(s)\rangle$$

- We need to know the data covariance for model comparison and Bayesian likelihood analysis

- There are different methods for obtaining a covariance:
  - Fluctuations across an ensemble of numerical simulations
  - Methods internal to the data (e.g. jackknife, bootstrap)
  - Analytical error propagation

- Analytical covariance can be useful if data vectors are large and/or sufficient numerical simulations are unavailable!

Analytical covariance

Paper available at: https://arxiv.org/abs/2308.15735

On the correlations of galaxy peculiar velocities and their covariance

Chris Blake,[1★] & Ryan J. Turner,[1]
[1] *Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
Measurements of the peculiar velocities of large samples of galaxies enable new tests of the standard cosmological model, including determination of the growth rate of cosmic structure that encodes gravitational physics. With the size of such samples now approaching hundreds of thousands of galaxies, complex statistical analysis techniques and models are required to extract cosmological information. In this paper we summarise how correlation functions between galaxy velocities, and with the surrounding large-scale structure, may be utilised to test cosmological models. We present new determinations of the analytical covariance between such correlation functions, which may be useful for cosmological likelihood analyses. The statistical model we use to determine these covariances includes the sample selection functions, observational noise, curved-sky effects and redshift-space distortions. By comparing these covariance determinations with corresponding estimates from large suites of cosmological simulations, we demonstrate that these analytical models recover the key features of the covariance between different statistics and separations, and produce similar measurements of the growth rate of structure.

Key words: cosmology: large-scale structure of Universe – cosmology: theory – methods: statistical

Velocity correlations

- What velocity correlation statistics are we interested in?

- Let's start from the galaxy position correlation function!

$$\xi_{gg}(\boldsymbol{r}) = \langle \delta_g(\boldsymbol{x})\, \delta_g(\boldsymbol{x}+\boldsymbol{r}) \rangle = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, P_{gg}(\boldsymbol{k})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \qquad \xi_{gg}(r) = \int \frac{dk\, k^2}{2\pi^2}\, P_{gg}(k)\, j_0(kr)$$

Neglecting RSD, this is an isotropic function, only depending on the magnitude $r = |\boldsymbol{r}|$, not the direction $\hat{\boldsymbol{r}}$. This is very useful when comparing correlation function measurements with theoretical predictions!

- Now let's consider 3D velocities in linear theory

$$\tilde{v}_i(\boldsymbol{k}) = -\frac{i\, k_i\, a\, H\, f}{k^2}\, \tilde{\delta}_m(\boldsymbol{k}) \qquad \psi_{ij}(\boldsymbol{r}) = \langle v_i(\boldsymbol{x})\, v_j(\boldsymbol{x}+\boldsymbol{r}) \rangle = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, \frac{k_i\, k_j}{k^2}\, P_{vv}(\boldsymbol{k})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}$$

The velocity correlation tensor $\psi_{ij}(\boldsymbol{r})$ is not an isotropic function!

Velocity correlations

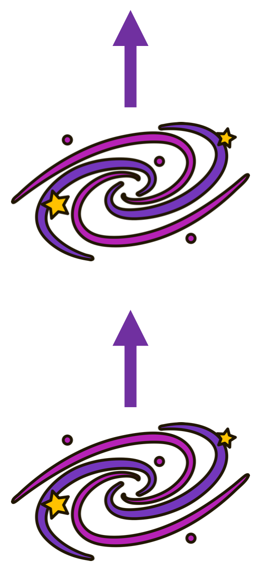
- Gorski (1988) showed that $\psi_{ij}(\boldsymbol{r})$ could be expressed in terms of two isotropic functions $\psi_\perp(r)$ and $\psi_\parallel(r)$:

$$\psi_{ij}(\boldsymbol{r}) = \psi_\perp(r)\,\delta^K_{ij} + [\psi_\parallel(r) - \psi_\perp(r)]\left(\frac{r_i\,r_j}{r^2}\right)$$

$$\psi_\perp(r) = \int \frac{dk\,k^2}{2\pi^2}\,P_{vv}(k)\,\frac{j_1(kr)}{kr}$$

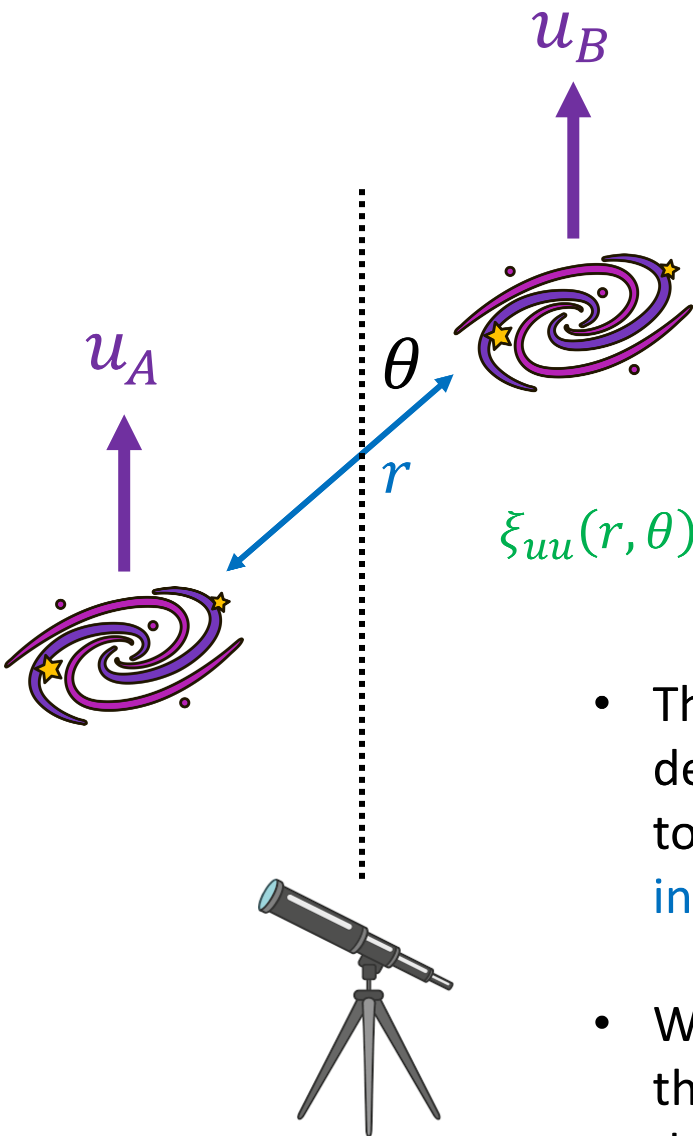
$$\psi_\parallel(r) = \int \frac{dk\,k^2}{2\pi^2}\,P_{vv}(k)\left[j_0(kr) - \frac{2j_1(kr)}{kr}\right]$$

$\psi_\perp(r)$: the extent to which galaxies are travelling together in the same bulk flow



$\psi_\parallel(r)$: the extent to which galaxies are approaching/receding along the line joining them

Velocity correlations



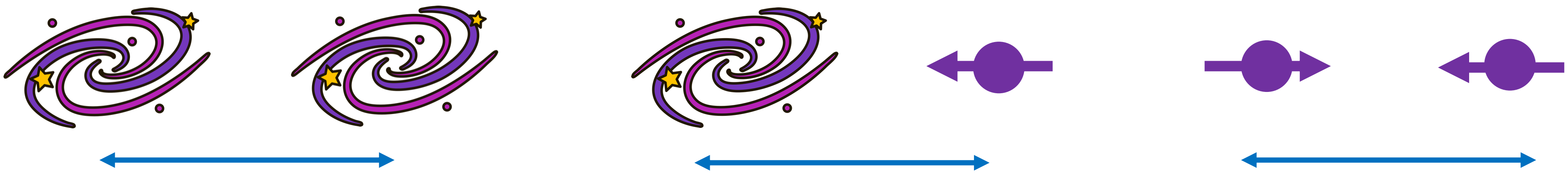
Consider the general correlation between the radial velocities u_A and u_B of two galaxies with separation vector $\boldsymbol{r}$ …

$$\xi_{uu}(r,\theta) = \langle u_A\, u_B \rangle = \psi_\perp(r) + [\psi_\parallel(r) - \psi_\perp(r)]\cos^2\theta$$

- The radial velocity correlation function is dependent on the angle of the separation vector to the line-of-sight → angle-averaging would lose information

- We can use the multipoles of $\xi_{uu}(r,\theta)$ to retain this information, which Gorski et al. (1989) defined as the $\psi_1(r)$ and $\psi_2(r)$ statistics

Velocity correlations



- In general, we will consider galaxy and velocity auto-correlations, and galaxy-velocity cross-correlations

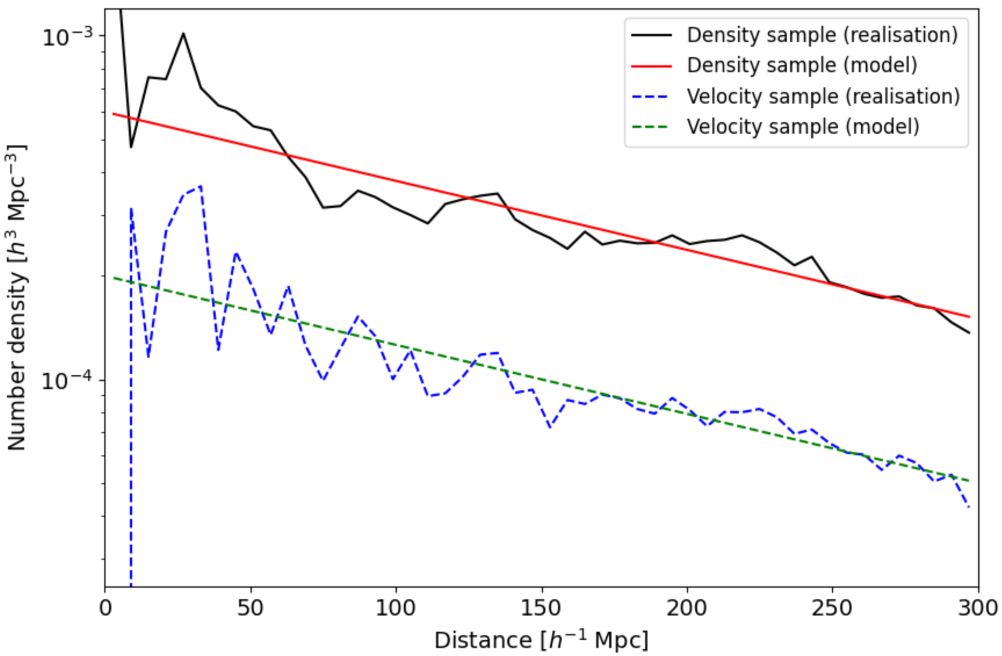
- For N-body simulations we know the full 3D velocities. We can analyse a set of 3 correlation functions: $(\xi_{gg}, \xi_{gv}, \xi_{vv}) \rightarrow$ simulation data vector

- For real observations we only know the radial velocities. Including RSD, we can analyse a set of 5 correlation functions: $(\xi^0_{gg}, \xi^2_{gg}, \xi^1_{gu}, \psi_1, \psi_2) \rightarrow$ observation data vector

- **Our goal: can we analytically predict the covariance of these groups of correlation functions?**

Mock catalogues for testing

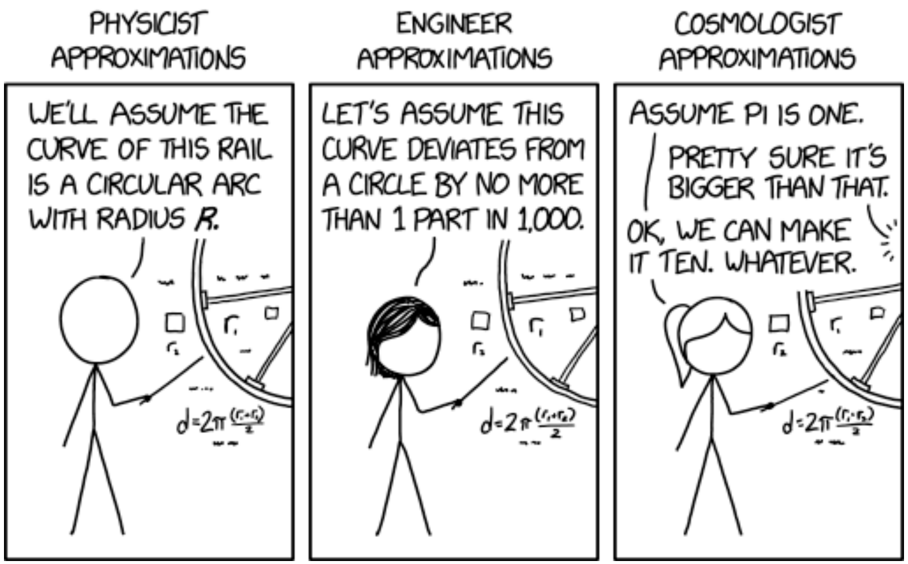
- We use $z < 0.1$ spheres drawn from 600 independent COLA mocks
- Density and velocity sub-samples with varying galaxy number density
- Apply velocity noise, which is 5% of distance
- Applying optimal weights to density and velocity tracers
- For observational statistics, we include curved-sky and RSD



1. Measure correlation functions *(see Ryan's talk)*
2. Compare analytical and simulation covariances
3. Fit growth rates

Approximations!

To determine the analytical covariances, we make some approximations!



https://xkcd.com/2205/

- Galaxy overdensity and velocity are Gaussian random fields
- Neglect higher-order correlations (e.g., the fact that velocity tracers preferentially sample overdense locations)
- Neglect the variation of the selection function and noise fields on the scale of the separation vector
- Use the "local plane-parallel approximation" (directions to locations on the scale of the separation vector are parallel)

Fun maths!

Integral over Fourier modes

Integral over survey volume

Selection function and weights

Model power spectra

Vector separations

Noise function and weights

$$\mathrm{Cov}\left[\hat{\xi}_{gg}(\mathbf{r}),\hat{\xi}_{gg}(\mathbf{s})\right] = N^2_{\xi_{gg}} \int \frac{V\,d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})} \int \frac{d^3\mathbf{x}}{V} \left[f_g^2(\mathbf{x})\,P_{gg}(\mathbf{k}) + \sigma_g^2(\mathbf{x})\right]^2$$

$$\mathrm{Cov}\left[\hat{\xi}_{vv}(\mathbf{r}),\hat{\xi}_{vv}(\mathbf{s})\right] = N^2_{\xi_{vv}} \int \frac{V\,d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})} \int \frac{d^3\mathbf{x}}{V} \left[f_v^2(\mathbf{x})\,P_{vv}(\mathbf{k})\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{k}})\,(\hat{\mathbf{s}}\cdot\hat{\mathbf{k}}) + \sigma_v^2(\mathbf{x})\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{s}})\right]^2$$

$$\mathrm{Cov}\left[\hat{\xi}_{gv}(\mathbf{r}),\hat{\xi}_{gv}(\mathbf{s})\right] = N^2_{\xi_{gv}} \int \frac{V\,d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})} \int \frac{d^3\mathbf{x}}{V} \frac{1}{2} \Big\{\left[f_g^2(\mathbf{x})\,P_{gg}(\mathbf{k}) + \sigma_g^2(\mathbf{x})\right]\left[f_v^2(\mathbf{x})\,P_{vv}(\mathbf{k})\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{k}})\,(\hat{\mathbf{s}}\cdot\hat{\mathbf{k}}) + \sigma_v^2(\mathbf{x})\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{s}})\right]$$
$$+ f_g^2(\mathbf{x})\,f_v^2(\mathbf{x})\,P_{gv}^2(\mathbf{k})\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{k}})\,(\hat{\mathbf{s}}\cdot\hat{\mathbf{k}})\Big\}$$

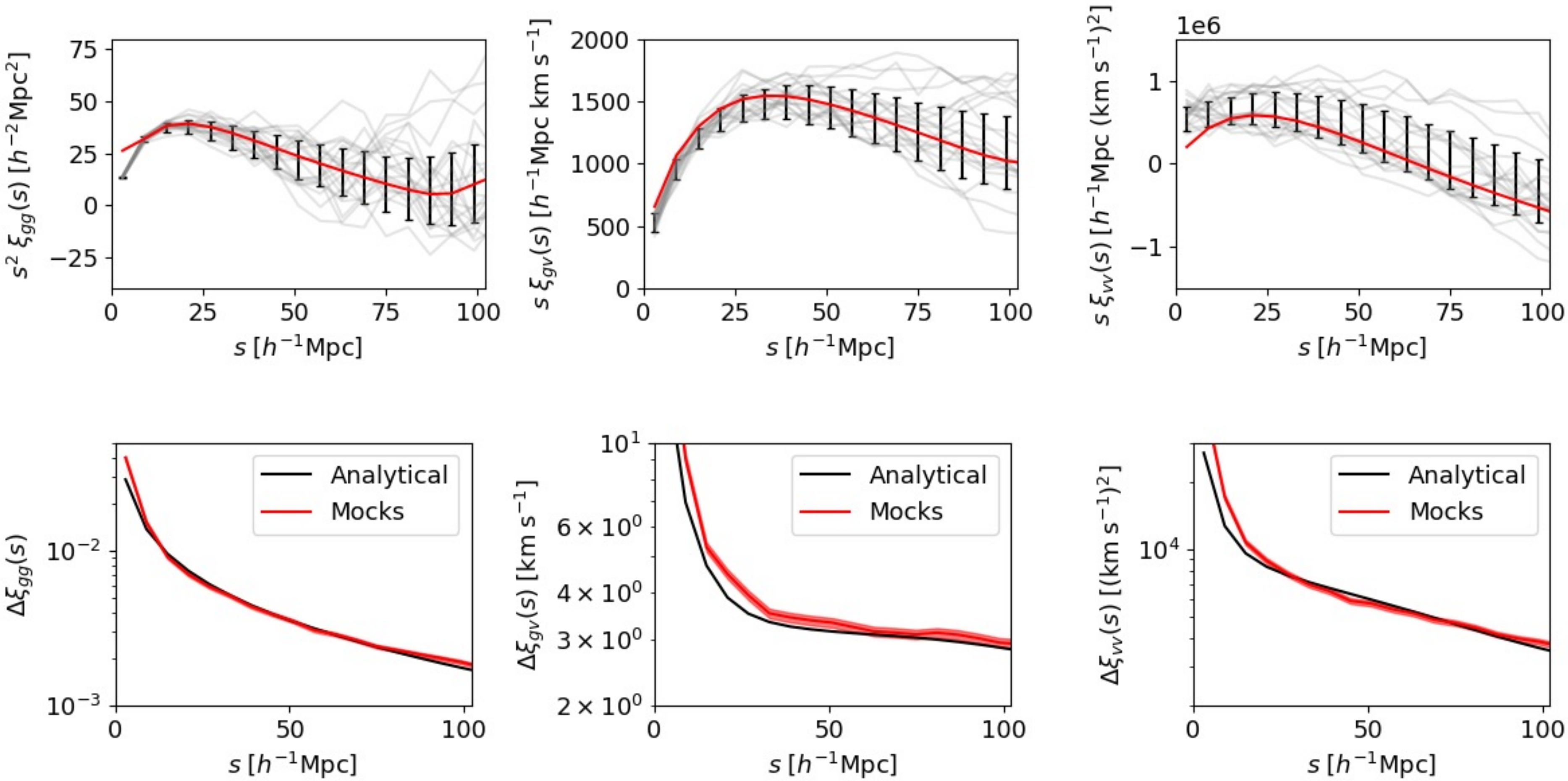
$$\mathrm{Cov}\left[\hat{\xi}_{gg}(\mathbf{r}),\hat{\xi}_{vv}(\mathbf{s})\right] = N_{\xi_{gg}} N_{\xi_{vv}} \int \frac{V\,d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})} \int \frac{d^3\mathbf{x}}{V} f_g^2(\mathbf{x})\,f_v^2(\mathbf{x})\,P_{gv}^2(\mathbf{k})\,(\hat{\mathbf{s}}\cdot\hat{\mathbf{k}})^2$$

$$\mathrm{Cov}\left[\hat{\xi}_{gg}(\mathbf{r}),\hat{\xi}_{gv}(\mathbf{s})\right] = N_{\xi_{gg}} N_{\xi_{gv}} \int \frac{V\,d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})} \int \frac{d^3\mathbf{x}}{V} \left[f_g^2(\mathbf{x})\,P_{gg}(\mathbf{k}) + \sigma_g^2(\mathbf{x})\right] f_g(\mathbf{x})\,f_v(\mathbf{x})\,P_{gv}(\mathbf{k})\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{k}})$$

$$\mathrm{Cov}\left[\hat{\xi}_{gv}(\mathbf{r}),\hat{\xi}_{vv}(\mathbf{s})\right] = N_{\xi_{gv}} N_{\xi_{vv}} \int \frac{V\,d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{s})} \int \frac{d^3\mathbf{x}}{V} \left[f_v^2(\mathbf{x})\,P_{vv}(\mathbf{k})\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{k}})\,(\hat{\mathbf{s}}\cdot\hat{\mathbf{k}}) + \sigma_v^2(\mathbf{x})\,(\hat{\mathbf{r}}\cdot\hat{\mathbf{s}})\right] f_g(\mathbf{x})\,f_v(\mathbf{x})\,P_{gv}(\mathbf{k})\,(\hat{\mathbf{s}}\cdot\hat{\mathbf{k}})$$

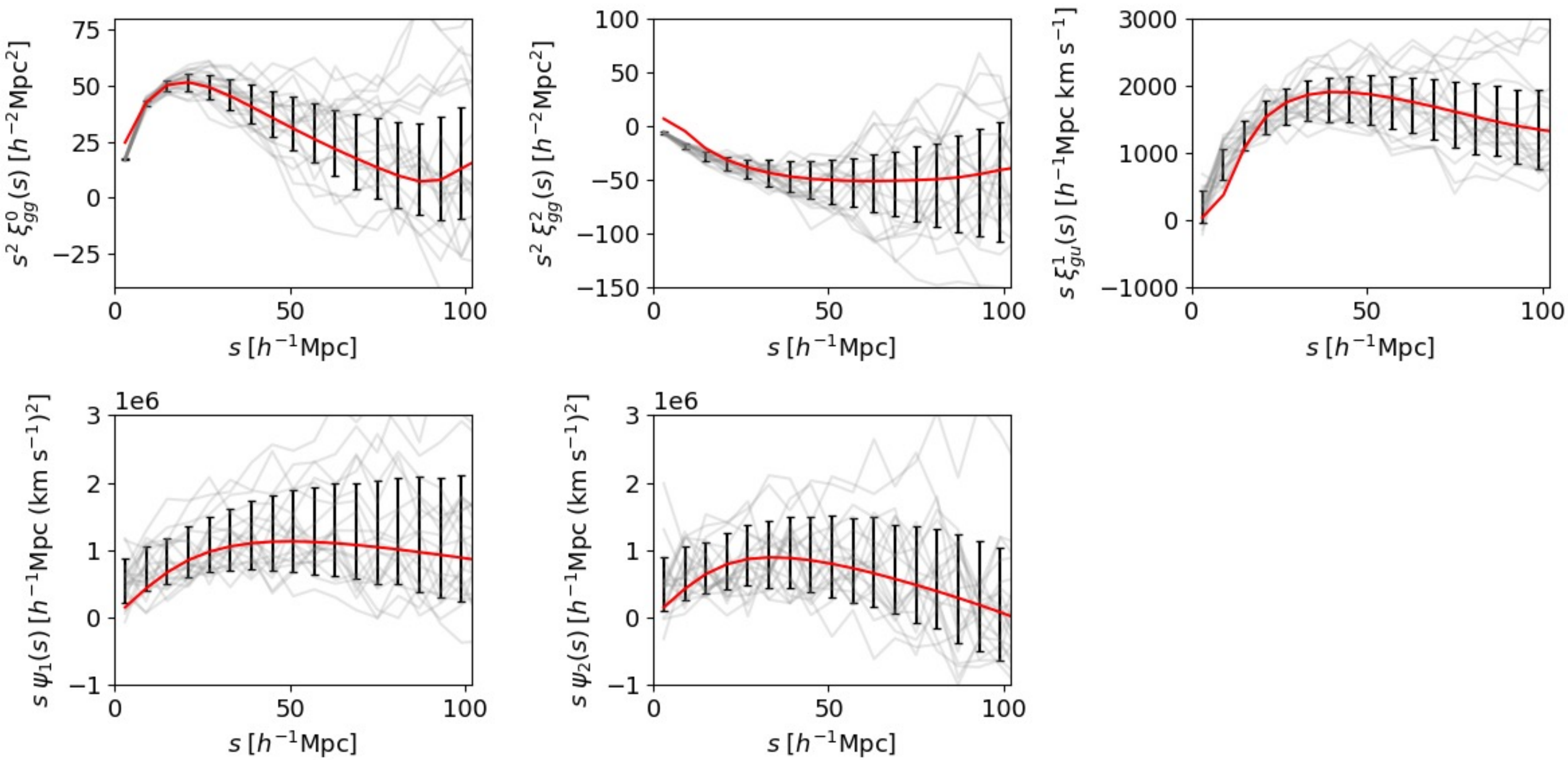
Correlation function measurements

- For the “simulation” 3D statistics: $(\xi_{gg}, \xi_{gv}, \xi_{vv})$



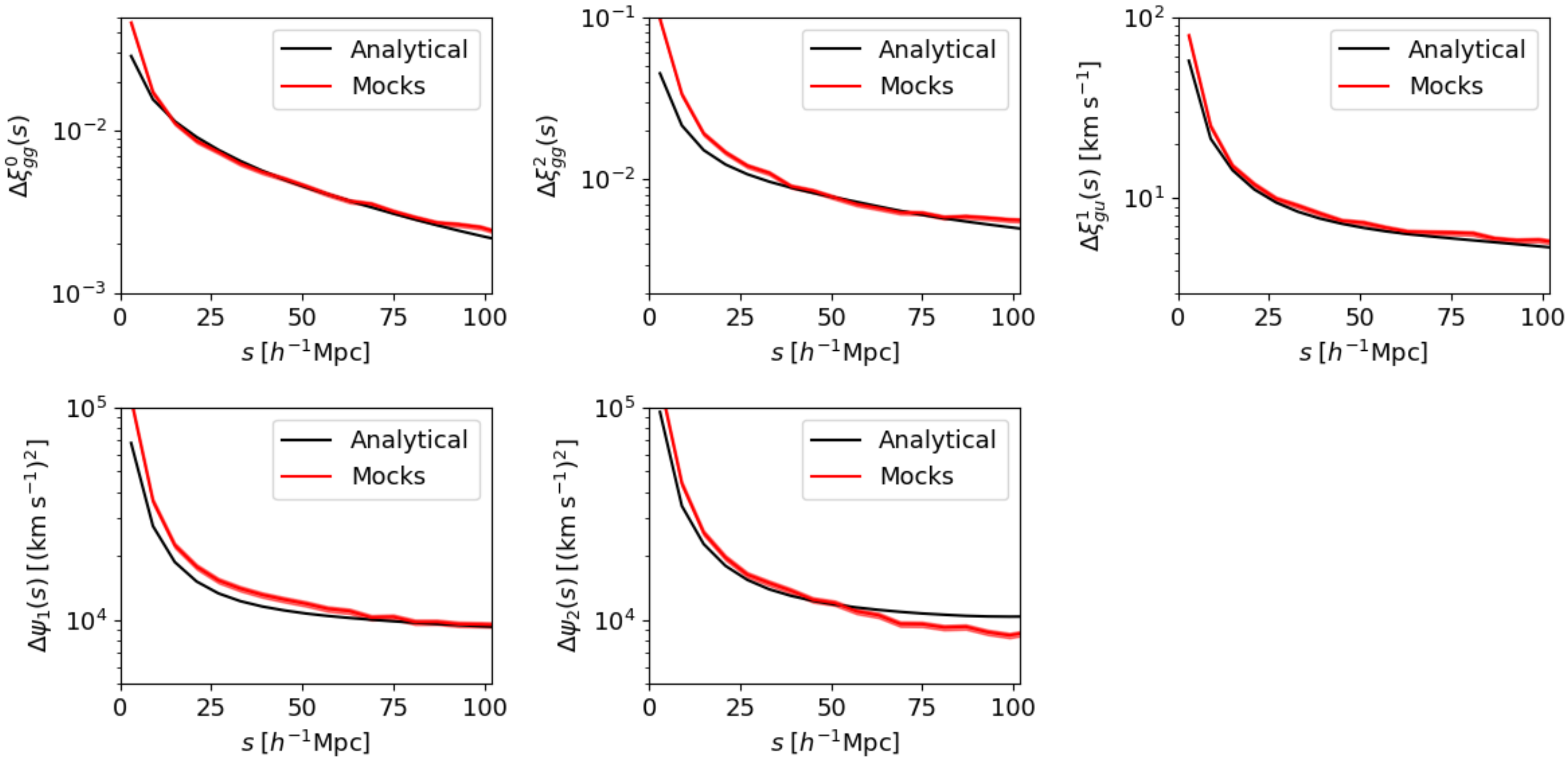
Correlation function measurements

- For the “observation” LOS statistics: $\left(\xi_{gg}^0, \xi_{gg}^2, \xi_{gu}^1, \psi_1, \psi_2\right)$



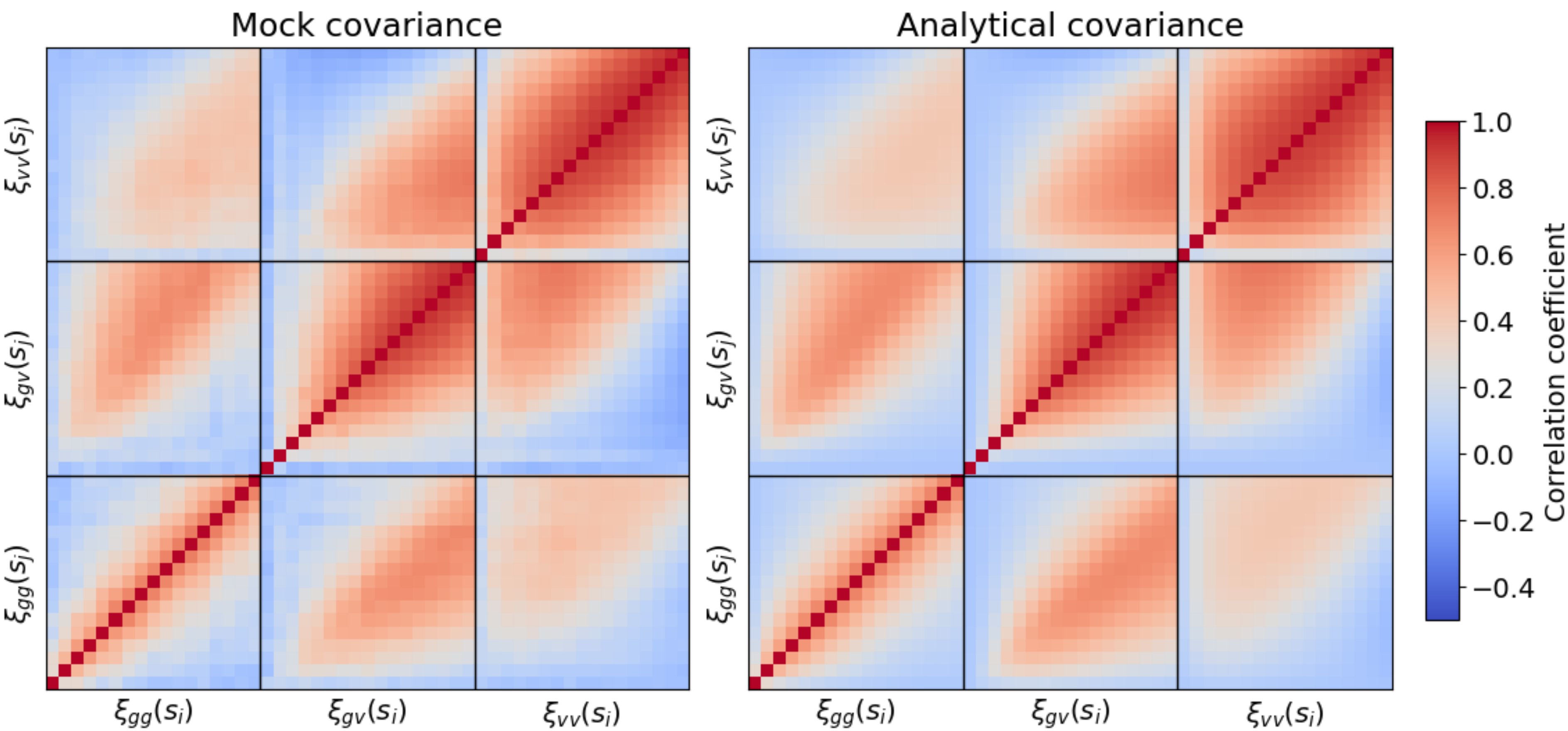
Correlation function measurements

- For the "observation" LOS statistics: $\left(\xi^0_{gg}, \xi^2_{gg}, \xi^1_{gu}, \psi_1, \psi_2\right)$



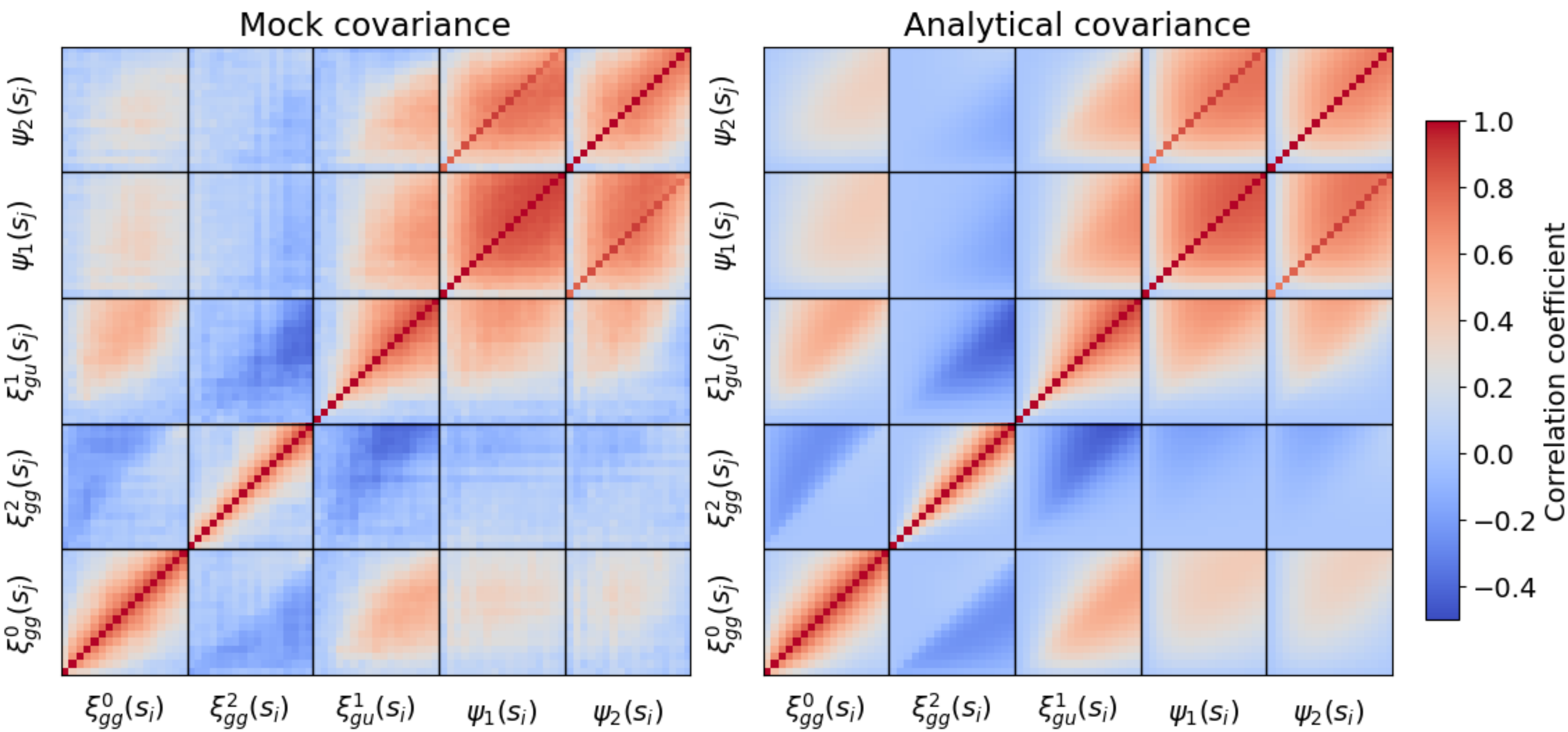
Analytical versus mock covariance

- For the "simulation" 3D statistics: $(\xi_{gg}, \xi_{gv}, \xi_{vv})$



Analytical versus mock covariance

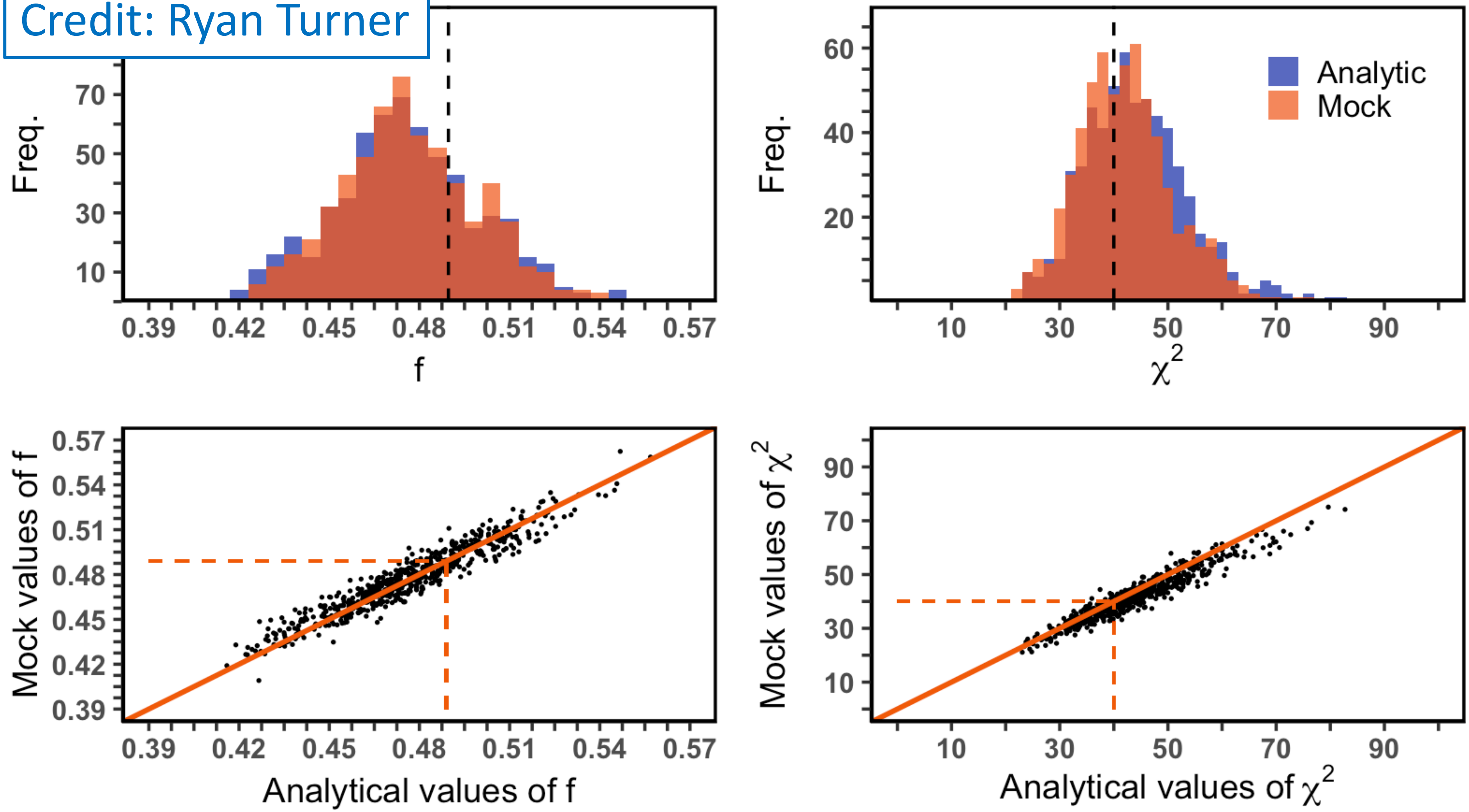
- For the "observation" LOS statistics: $\left(\xi_{gg}^{0}, \xi_{gg}^{2}, \xi_{gu}^{1}, \psi_1, \psi_2\right)$



Growth rate fits

- For the "simulation" 3D statistics: $(\xi_{gg}, \xi_{gv}, \xi_{vv})$

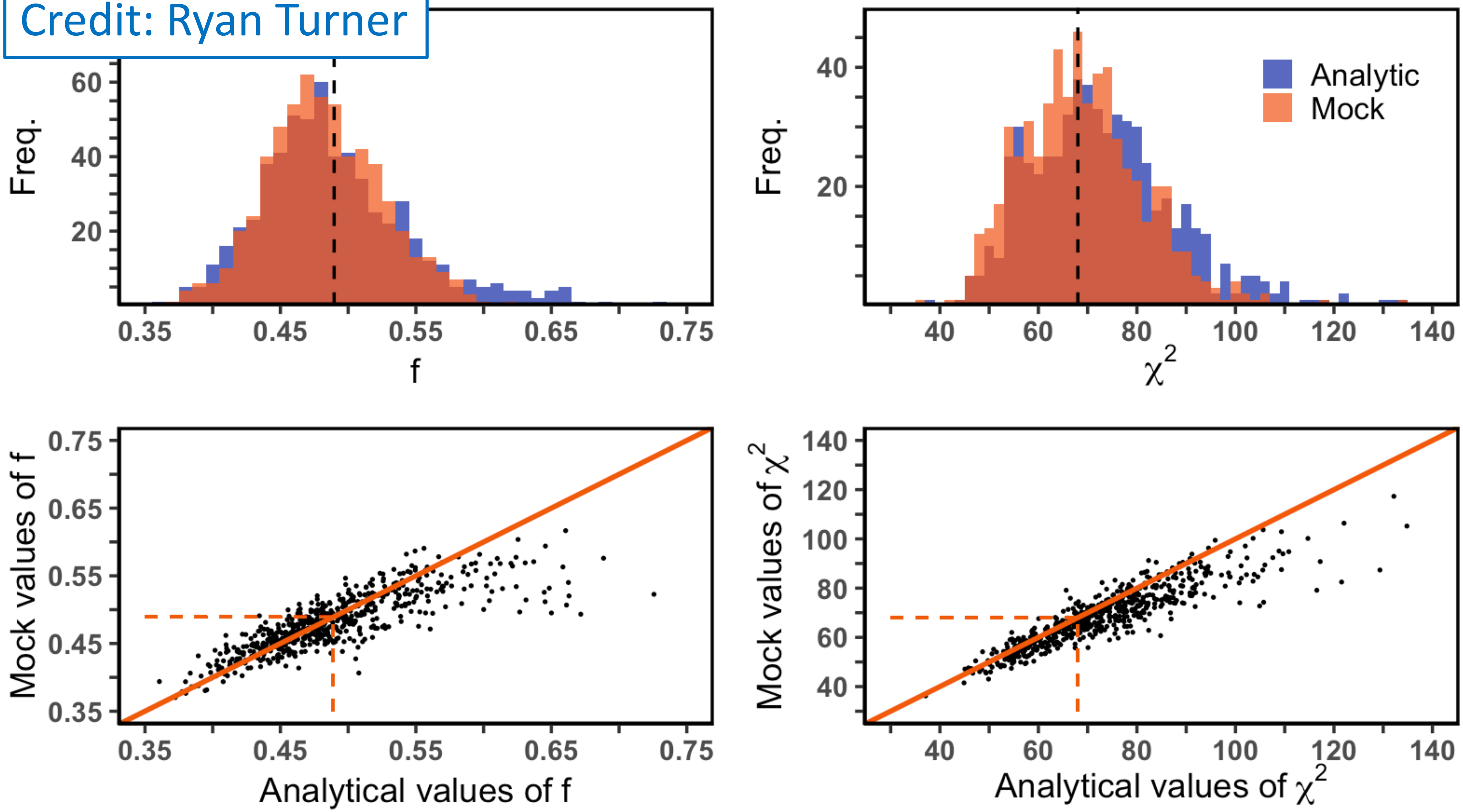
Credit: Ryan Turner



Growth rate fits

- For the “observation” LOS statistics: $\left(\xi_{gg}^{0}, \xi_{gg}^{2}, \xi_{gu}^{1}, \psi_1, \psi_2\right)$

Credit: Ryan Turner



Conclusions

- We have computed analytical covariances for velocity correlations including selection function, noise, curved-sky and RSD
- The analytical covariance is a good representation of the dispersion across simulations
- Growth rate fits using the analytical and simulation covariances agree well
- Analytical covariances could be a useful in cases where sufficient simulations are unavailable

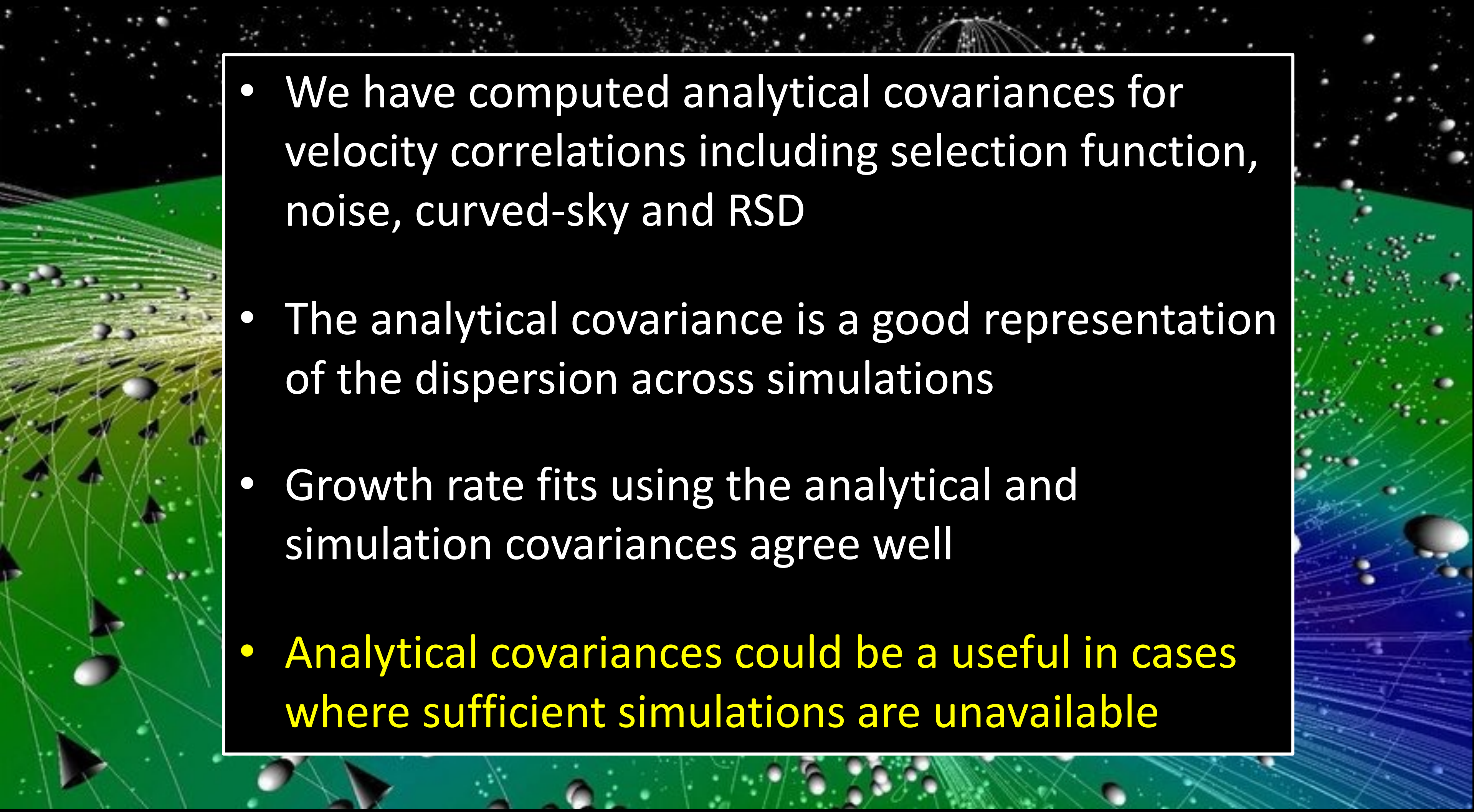


Image credit: Courtois, Pomarede, Tully et al, Cosmic Flows