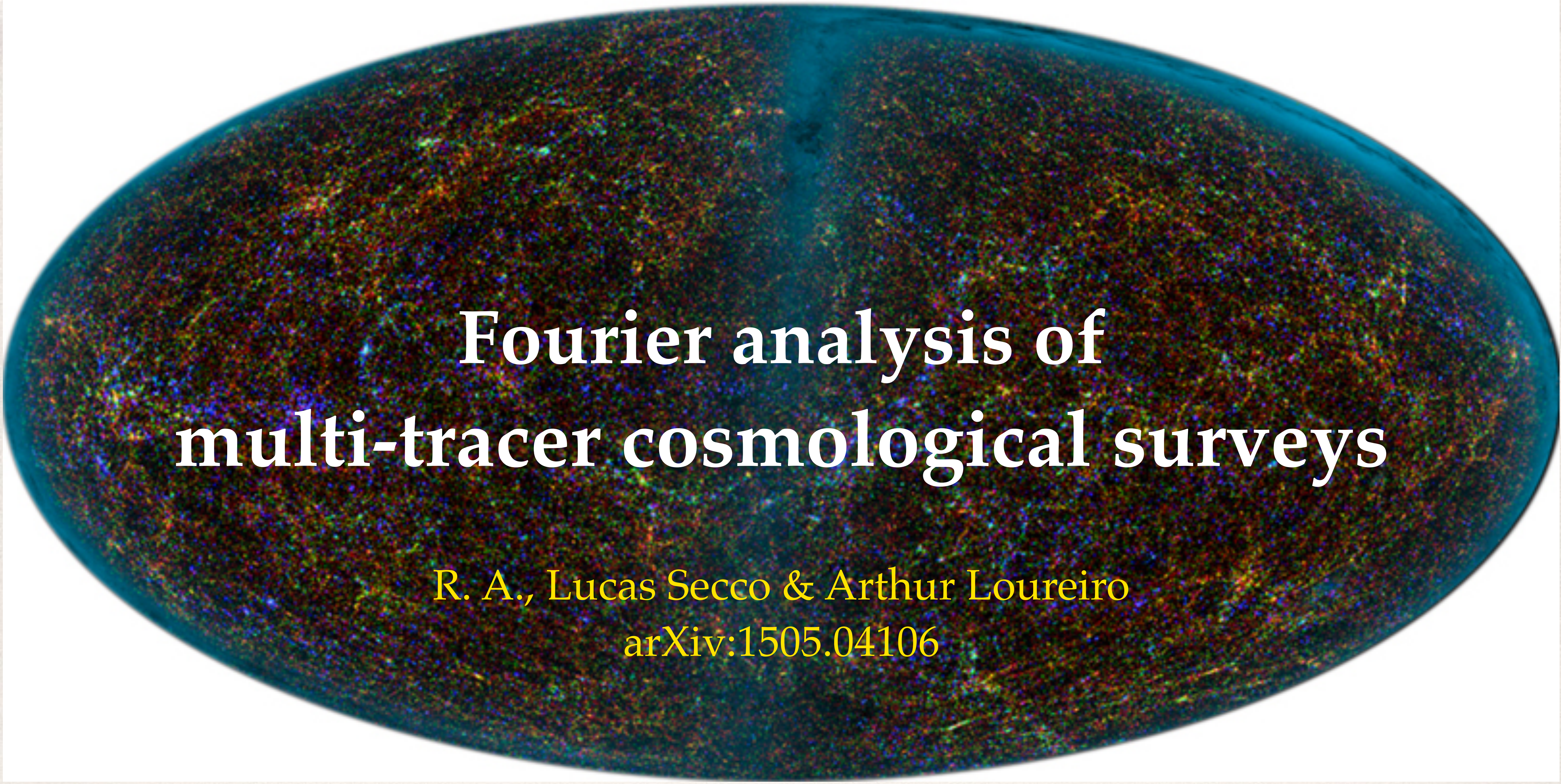


Fourier analysis of multi-tracer cosmological surveys

R. A., Lucas Secco & Arthur Loureiro

arXiv:1505.04106

Lucas

L. Raul Abramo

Physics Institute & LabCosmos

University Of Sao Paulo

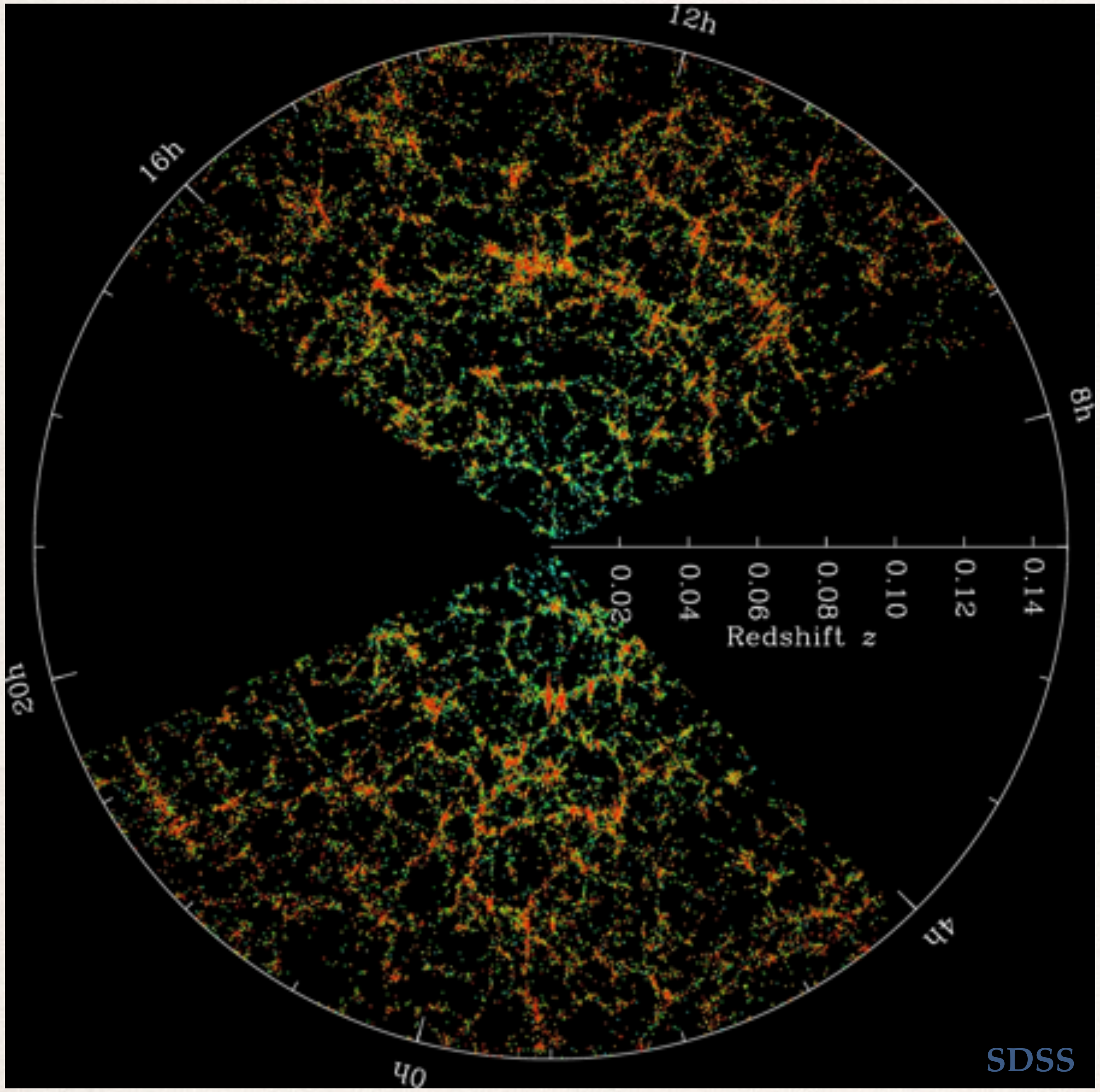
Arthur

LABCOSMOS

http://www.iag.usp.br/labcosmos/en/

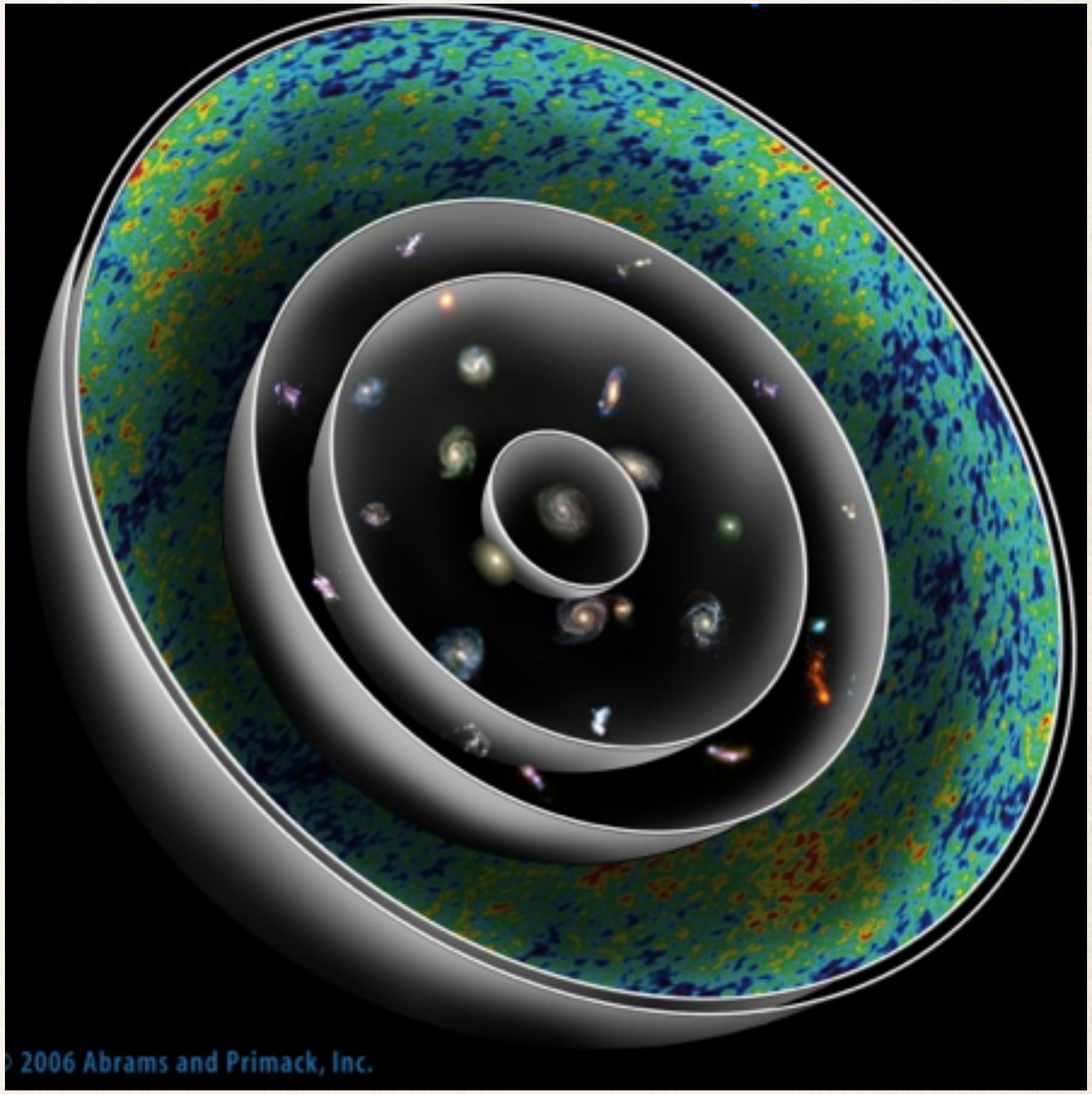
Galaxy surveys are evolving

We used to live in an
era of shot ("counts") noise



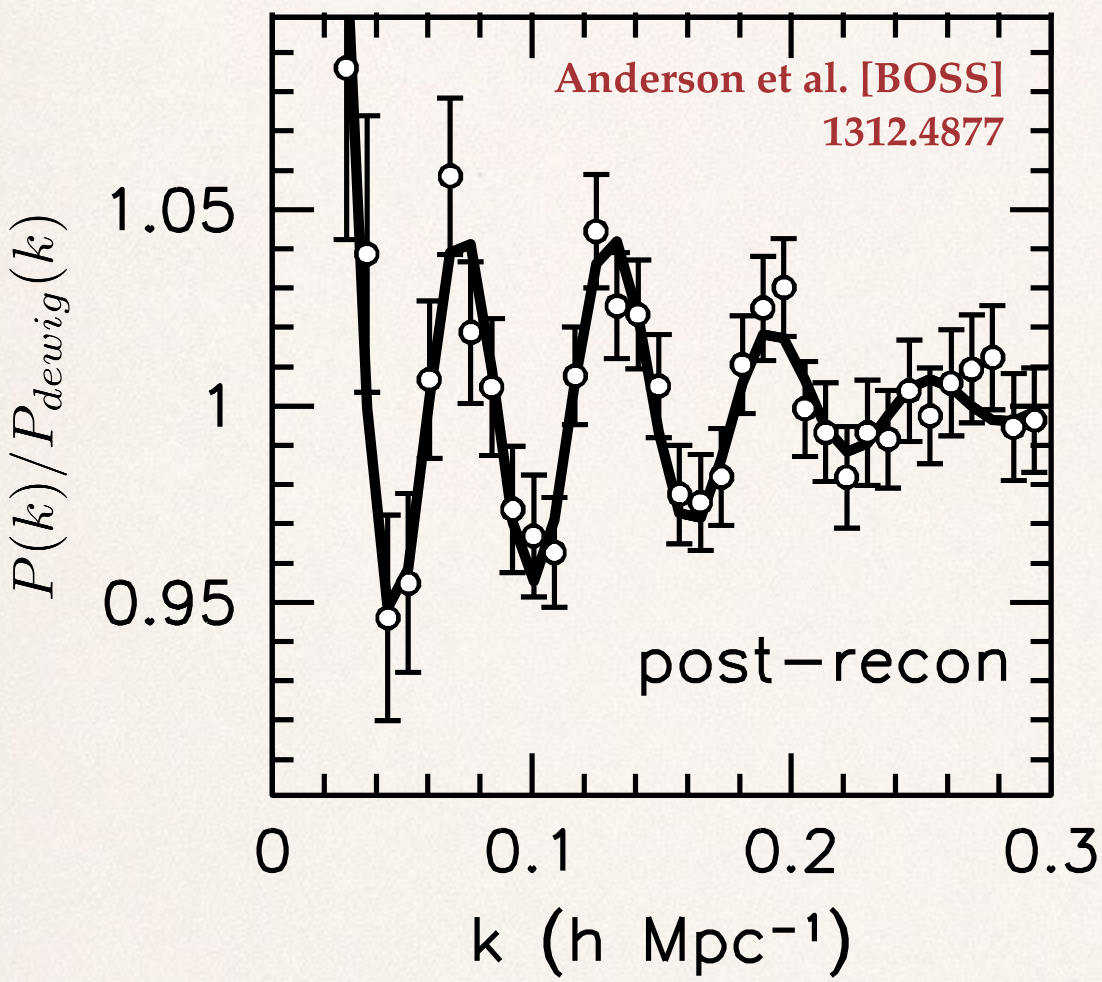
[**Finding galaxies** was the limiting factor]

We are now in the **age of cosmic variance and systematics**



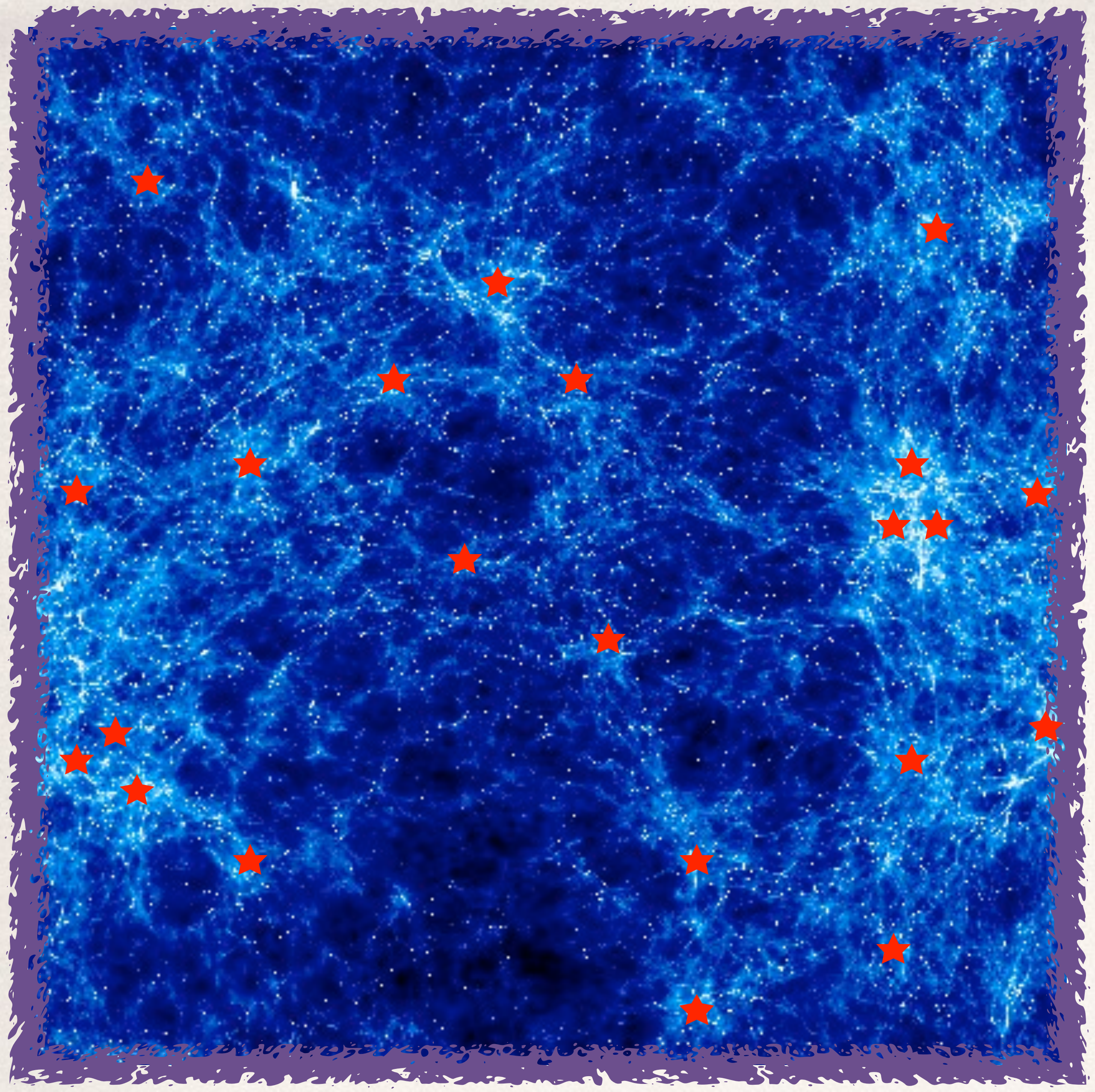
[**Gaining volume** is the limiting factor]

Surveys of large-scale structure are now **limited by cosmic variance** (and systematics)



These **error bars** are dominated by **cosmic variance** (and some systematics): in order to improve, we must increase the **volume** of our surveys

However, up to any given redshift there is only a finite volume. Moreover, we are reaching closer to the limits of the observable Universe!



Shot noise:

finite number of **counts** of the tracers of the underlying density field (Poisson statistics)

Cosmic variance:

finite **volume** inside which we can estimate the **amplitudes and phases** of the (Gaussian) random **modes** of the density field

$$\langle \frac{\delta n_g(\vec{x})}{\bar{n}_g} \frac{\delta n_g(\vec{y})}{\bar{n}_g} \rangle_V \simeq b_g^2 \, \langle \delta_m(\vec{x}) \delta_m(\vec{y}) \rangle_V + \frac{1}{\bar{n}_g} \delta_D(\vec{x} - \vec{y}) \quad \longleftrightarrow \quad P_g(\vec{k}) \simeq b_g^2 \, P_m(\vec{k}) + \frac{1}{\bar{n}_g}$$

Clustering in units of shot noise

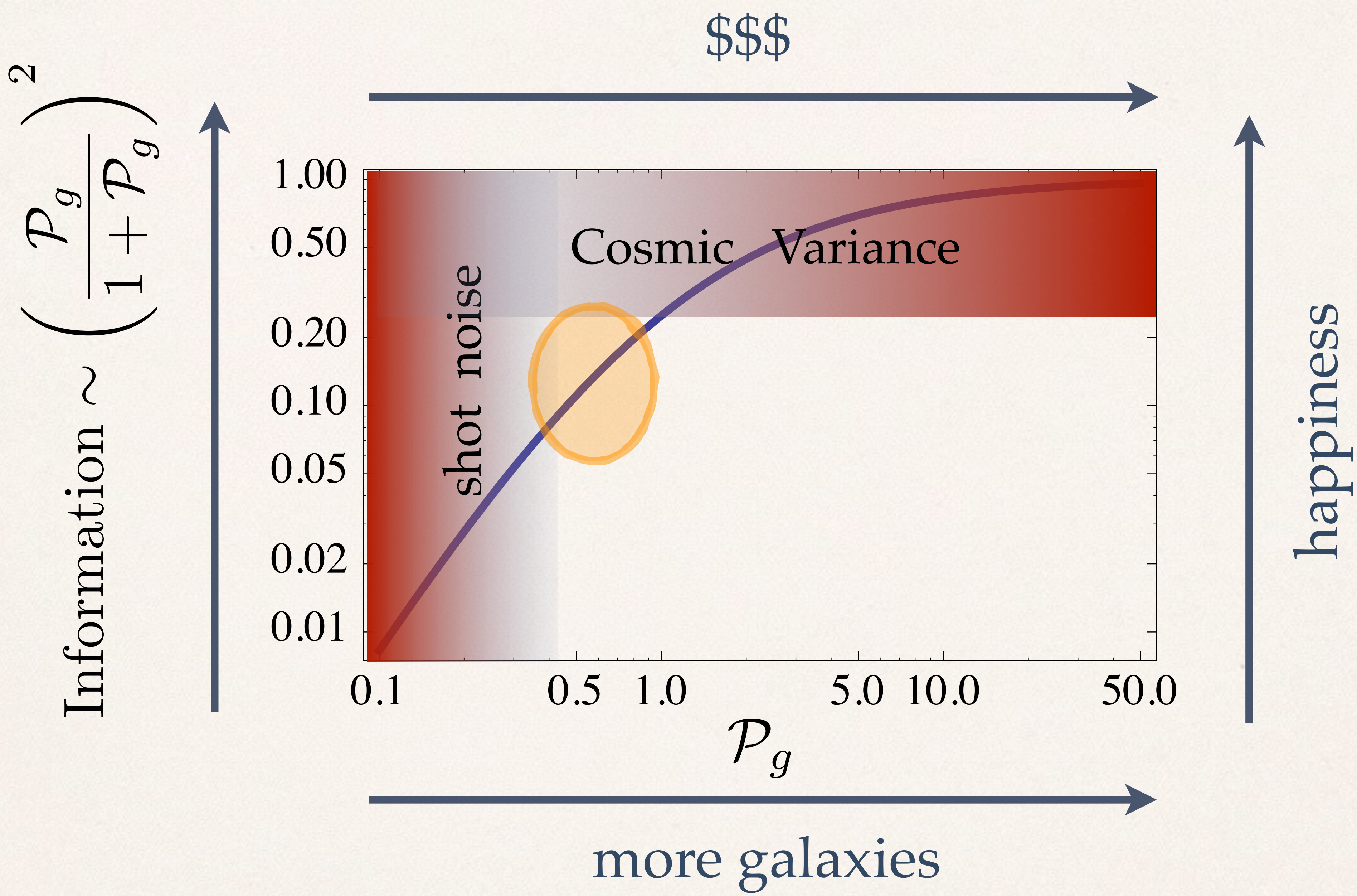
$$\boxed{\bar{n}_g \, P_g(\vec{k}) \simeq \bar{n}_g \, b_g^2 \, P_m(\vec{k}) + 1}$$

SNR

noise

Feldman, Kaiser & Peacock 1994 (**FKP**)
Tegmark et al. 1997, 1998

Fisher information of galaxy surveys (single type of galaxy)



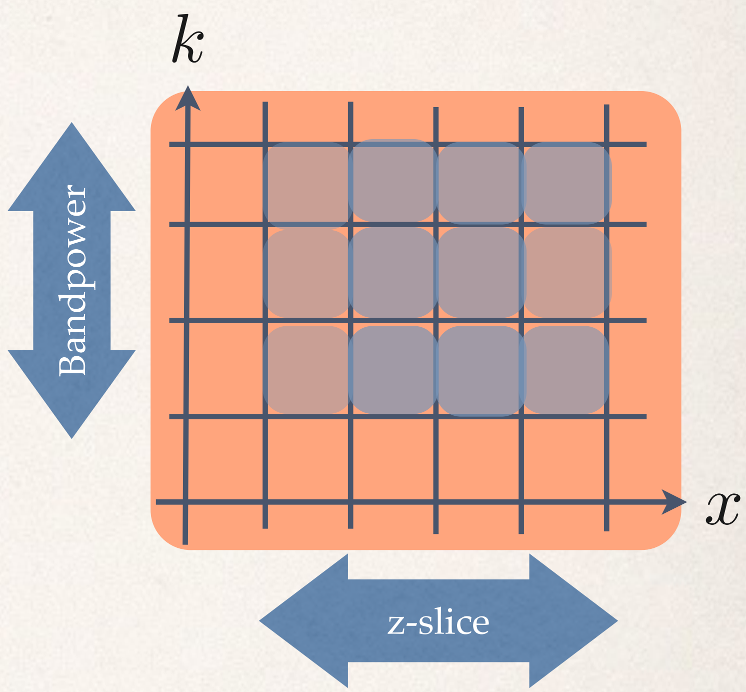
Signal/Noise: Clustering strength of galaxy type "*g*" in redshift space

$$\mathcal{P}_g(\vec{x},\vec{k}) \equiv \bar{n}_g(\vec{x}) \left[b_g(z,k) + f(z)\,\mu_k^2\right]^2 G^2(z)\, P_m(k)$$

FKP 1994
Hamilton 2005
R. A. 2012

Fisher information in phase space

On each **unit volume** of **phase space** there is a certain **amount of information** about the clustering strength:



$$F[\log \mathcal{P}_g] \times \frac{\Delta V_x\, \Delta V_k}{(2\pi)^3} = \frac{1}{2}\left(\frac{\mathcal{P}_g}{1+\mathcal{P}_g}\right)^2 \times \frac{\Delta V_x\, \Delta V_k}{(2\pi)^3}$$

Fisher information / (phase space volume)

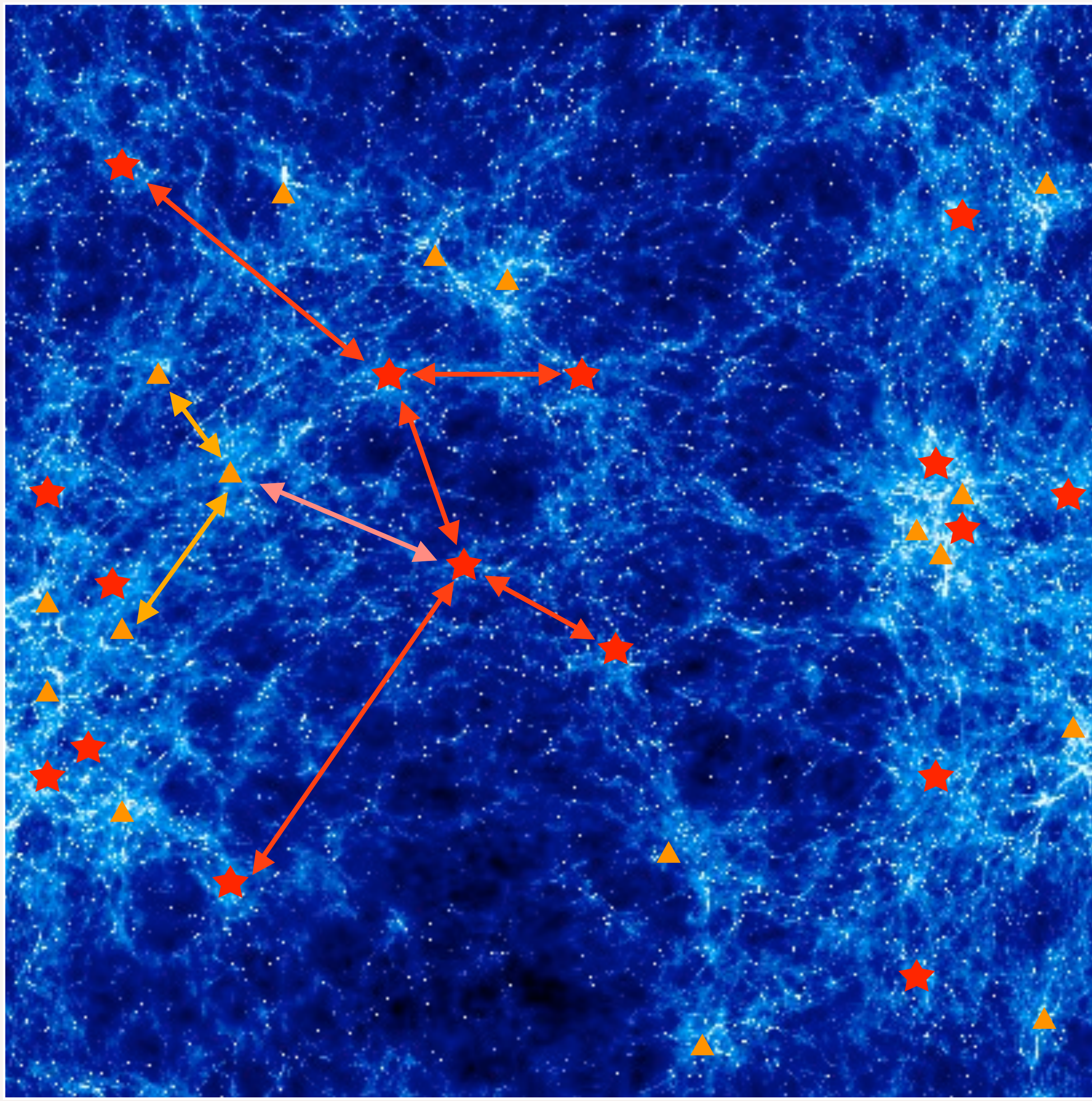
phase space **volume =** $\Delta\mathcal{V}$.

The **precision** (SNR) with which we can **estimate the clustering strength** is:

$$\frac{\mathcal{P}_g^2}{\sigma^2(\mathcal{P}_g)} = F[\log \mathcal{P}_g] \times \Delta\mathcal{V} = \frac{1}{2}\left(\frac{\mathcal{P}_g}{1+\mathcal{P}_g}\right)^2 \Delta\mathcal{V}$$

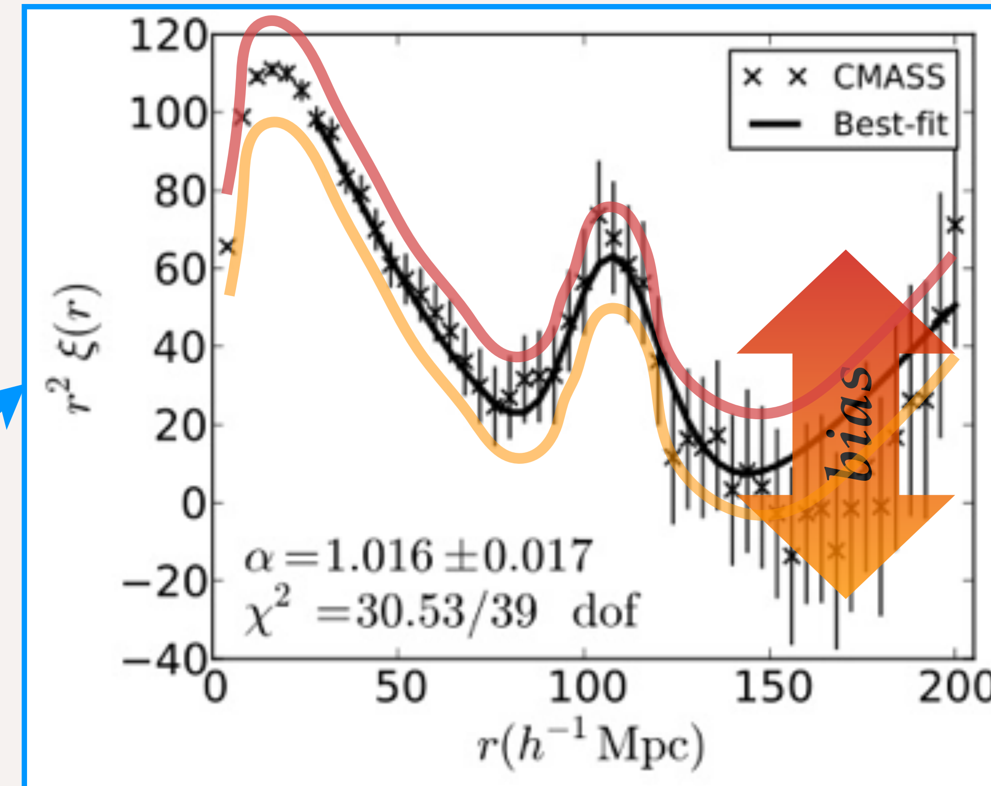
$$F = \frac{1}{2}\left(\frac{\text{signal}}{\text{signal } + \text{noise}}\right)^2 \leq \frac{1}{2}$$

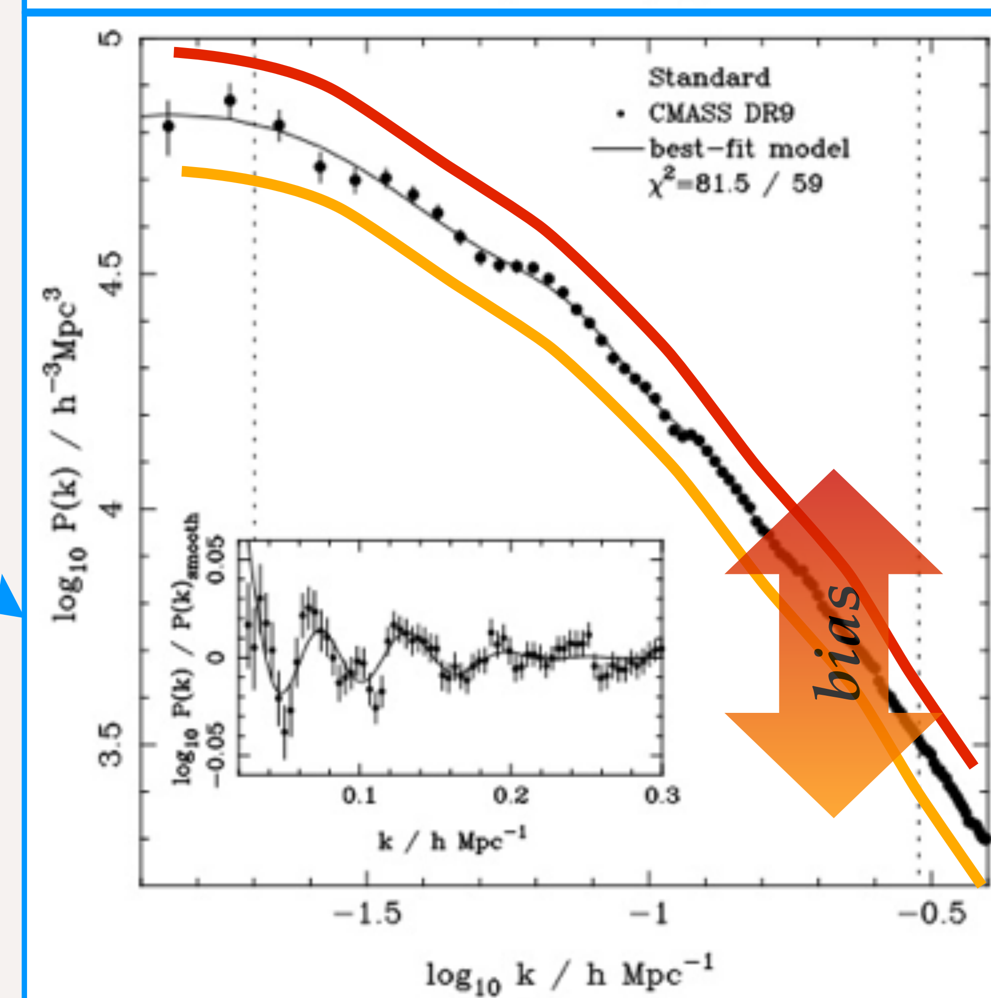
But what if there are **many different types** of galaxies?



Clustering in position space

Clustering in Fourier space





What is the **optimal** way of **combining** the clustering information from **several different types of tracers**?

Multi-tracer Fisher information matrix

<u>Multi-tracer technique:</u>
Seljak 2008
McDonald & Seljak 2008
Gil-Marín et al. 2011
Hamaus et al. 2011,2012
Cai & Bernstein 2011

<u>Fisher matrix:</u>
R.A. 2012
R.A. & K. Leonard 2013

$$\boxed{F[\log \mathcal{P}_\alpha, \log \mathcal{P}_\beta] = \frac{1}{4}\left[\frac{\mathcal{P}_\alpha \mathcal{P}_\beta (1-\mathcal{P})}{(1+\mathcal{P})^2} + \delta_{\alpha\beta}\frac{\mathcal{P}_\alpha \mathcal{P}}{1+\mathcal{P}}\right]}$$

Clustering strengths of **each** galaxy **type**: $\mathcal{P}_\alpha$

Total clustering strength of a survey: $\mathcal{P} \equiv \sum_\alpha \mathcal{P}_\alpha$

$$F = \frac{1}{2}\left(\frac{\mathcal{P}}{1+\mathcal{P}}\right)^2 \Longrightarrow \frac{1}{4}\begin{pmatrix} \frac{\mathcal{P}_1^2(1-\mathcal{P})}{(1+\mathcal{P})^2} + \frac{\mathcal{P}_1\mathcal{P}}{1+\mathcal{P}} & \frac{\mathcal{P}_1\mathcal{P}_2(1-\mathcal{P})}{(1+\mathcal{P})^2} \\ \frac{\mathcal{P}_1\mathcal{P}_2(1-\mathcal{P})}{(1+\mathcal{P})^2} & \frac{\mathcal{P}_2^2(1-\mathcal{P})}{(1+\mathcal{P})^2} + \frac{\mathcal{P}_2\mathcal{P}}{1+\mathcal{P}} \end{pmatrix} \Longrightarrow \frac{1}{4}\begin{pmatrix} \frac{\mathcal{P}_1^2(1-\mathcal{P})}{(1+\mathcal{P})^2} + \frac{\mathcal{P}_1\mathcal{P}}{1+\mathcal{P}} & \frac{\mathcal{P}_1\mathcal{P}_2(1-\mathcal{P})}{(1+\mathcal{P})^2} & \frac{\mathcal{P}_1\mathcal{P}_3(1-\mathcal{P})}{(1+\mathcal{P})^2} \\ \frac{\mathcal{P}_1\mathcal{P}_2(1-\mathcal{P})}{(1+\mathcal{P})^2} & \frac{\mathcal{P}_2^2(1-\mathcal{P})}{(1+\mathcal{P})^2} + \frac{\mathcal{P}_2\mathcal{P}}{1+\mathcal{P}} & \frac{\mathcal{P}_2\mathcal{P}_3(1-\mathcal{P})}{(1+\mathcal{P})^2} \\ \frac{\mathcal{P}_1\mathcal{P}_3(1-\mathcal{P})}{(1+\mathcal{P})^2} & \frac{\mathcal{P}_2\mathcal{P}_3(1-\mathcal{P})}{(1+\mathcal{P})^2} & \frac{\mathcal{P}_3^2(1-\mathcal{P})}{(1+\mathcal{P})^2} + \frac{\mathcal{P}_3\mathcal{P}}{1+\mathcal{P}} \end{pmatrix}$$

1 tracer **2 tracers** **3 tracers**

OK… but are we in fact **gaining** any information by **splitting galaxies** into **sub-types**, or are we just "shuffling around" the information?

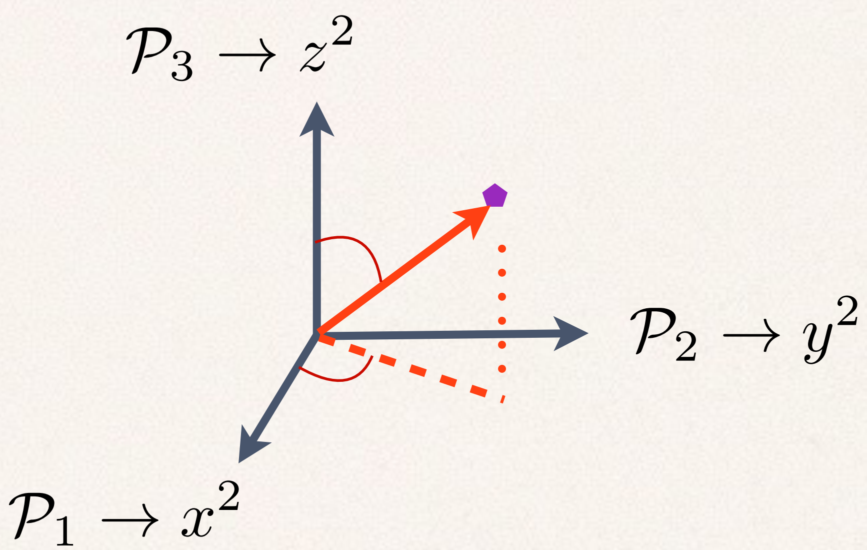
Multi-tracer Fisher information matrix

Yes, we **gain** information

> In fact, with **multiple tracers** the Fisher matrix is **unbounded**!

We can **diagonalize** the multi-tracer Fisher matrix by changing variables:

$\Rightarrow$ (hyper) **spherical coordinates!**



$$\left\{ \begin{array}{c} \mathcal{P}_1 \\ \\ \mathcal{P}_2 \\ \\ \mathcal{P}_3 \end{array} \right\} = \left\{ \begin{array}{c} x^2 \\ \\ y^2 \\ \\ z^2 \end{array} \right\} \iff \left\{ \begin{array}{c} r^2 \\ \\ \tan^2\theta \\ \\ \tan^2\phi \end{array} \right\} = \left\{ \begin{array}{c} \mathcal{P} \\ \\ \frac{\mathcal{P}_3}{\mathcal{P}_1+\mathcal{P}_2} \\ \\ \frac{\mathcal{P}_2}{\mathcal{P}_1} \end{array} \right\}$$

Total clustering strength

Relative clustering strengths

Multi-tracer Fisher information matrix

In "spherical" coordinates (i.e., using the total clustering strength and the relative clustering strengths) the **Fisher matrix becomes diagonal**!

E.g.: three species of tracers

Total clustering: < 1/2

$$F_{Sph} = \left\{ \begin{array}{ccc} \boxed{\frac{1}{2}\left(\frac{\mathcal{P}}{1+\mathcal{P}}\right)^2} & 0 & 0 \\ 0 & \boxed{\frac{1}{4}\frac{\mathcal{P}^2}{1+\mathcal{P}}\sin^2\theta\cos^2\theta} & 0 \\ 0 & 0 & \boxed{\frac{1}{4}\frac{\mathcal{P}^2}{1+\mathcal{P}}\sin^2\theta\sin^2\phi\cos^2\phi} \end{array} \right\}$$

Relative clusterings: information ~ $\mathcal{P} = \sum_\alpha \bar{n}_\alpha b_\alpha^2 P_m$

- ☛ **unbounded**
- ☛ **extra information!**

Seljak 2008
McDonald & Seljak 2008

Why?

Cosmic variance is only inherited through the **spectrum**

By **comparing** the clustering between **different tracers** of large-scale structure (e.g.: LRGs, ELGs, etc.), we can **measure with arbitrary accuracy*** the physical parameters that distinguish the different **clustering strengths**:

$$\mathcal{P}_1 = n_1\,(b_1 + f\,\mu_k^2)^2\,P(k;z)$$

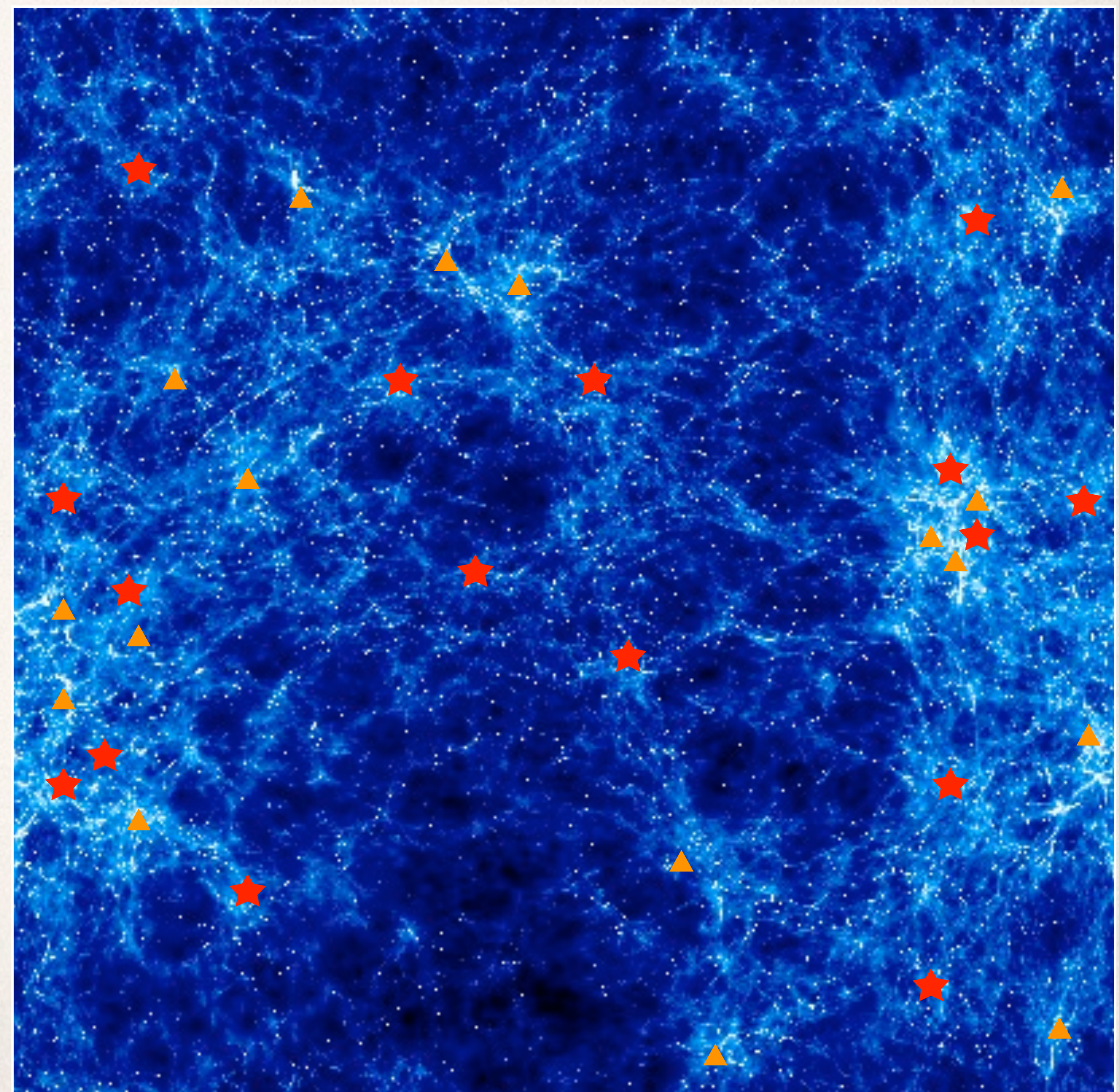
$$\mathcal{P}_2 = n_2\,(b_2 + f\,\mu_k^2)^2\,P(k;z)$$

$$\frac{\mathcal{P}_1}{\mathcal{P}_2} = \frac{n_1\,(b_1 + f\,\mu_k^2)^2}{n_2\,(b_2 + f\,\mu_k^2)^2}$$

Cosmic variance does **not** apply:

- *** bias * RSDs**
- *** NGs *HODs**

The key: **high number densities** of distinct types of tracers (red galaxies, blue galaxies, emission-line galaxies, quasars, etc.)



Simplest example: two types of tracers of large-scale structure

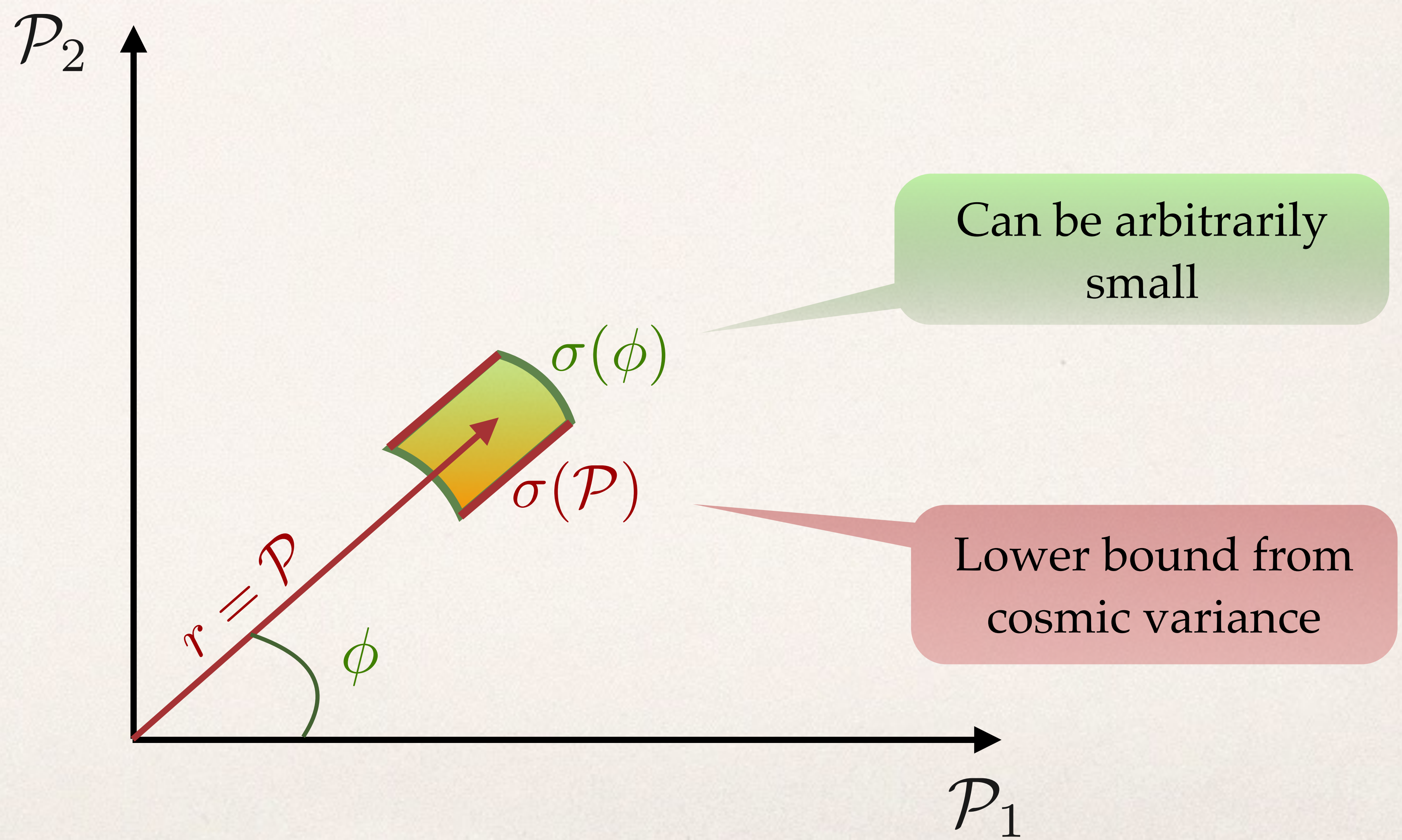
Total clustering

$$\mathcal{P} = \mathcal{P}_1 + \mathcal{P}_2$$

Relative clustering

$$\mathcal{R} = \frac{\mathcal{P}_2}{\mathcal{P}_1} = \tan^2\phi$$

$$F = \begin{pmatrix} F_{\mathcal{P}} & 0 \\ 0 & F_{\mathcal{R}} \end{pmatrix}$$



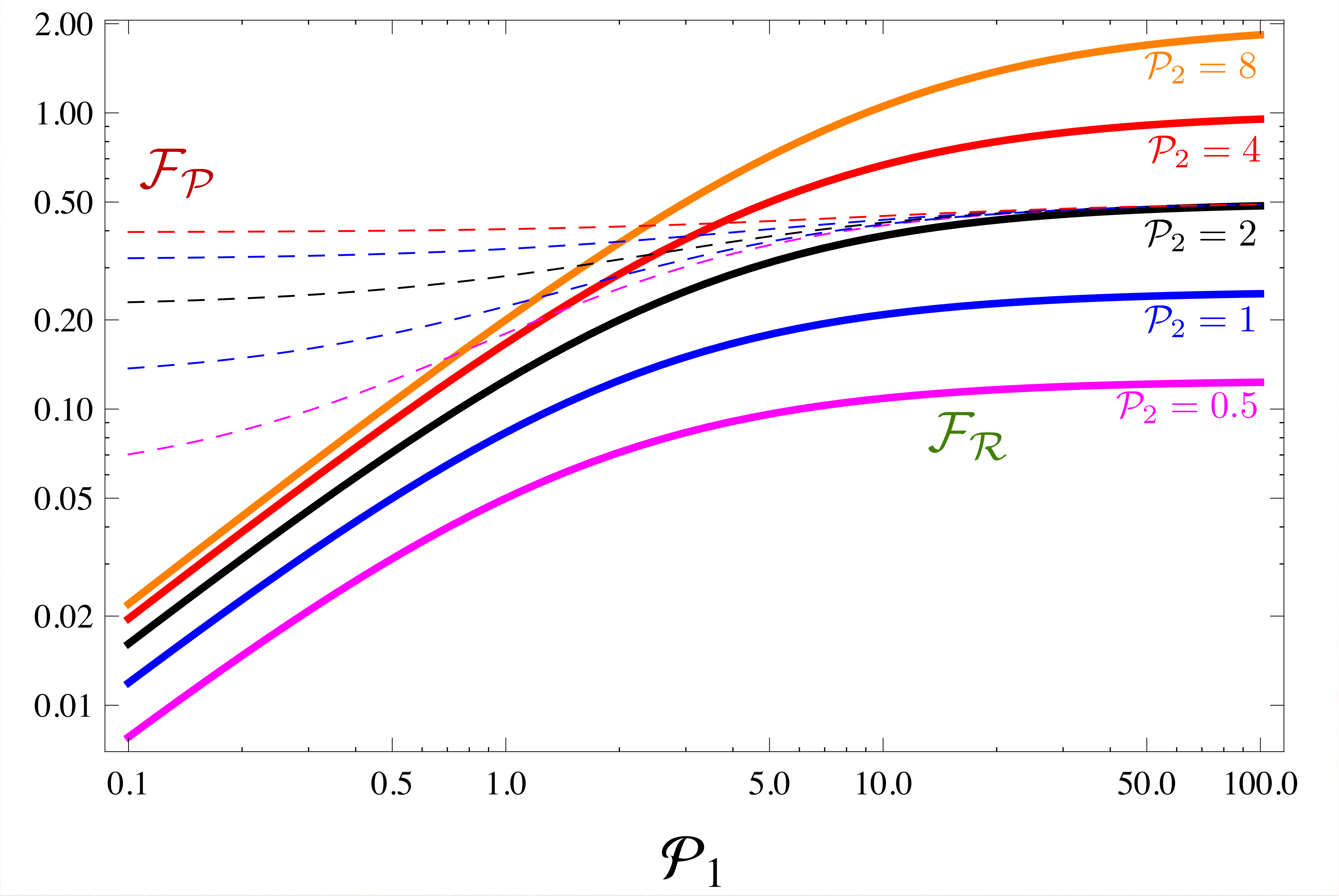
Simplest example:
two types of tracers of large-scale structure

Total clustering

$$\mathcal{F}_{\mathcal{P}} = \frac{\mathcal{P}^2}{\sigma^2(\mathcal{P})} = \frac{1}{2}\frac{\mathcal{P}^2}{(1+\mathcal{P})^2}$$

Relative clustering

$$\mathcal{F}_{\mathcal{R}} = \frac{(\mathcal{P}_2/\mathcal{P}_1)^2}{\sigma^2(\mathcal{P}_2/\mathcal{P}_1)} = \frac{1}{4}\frac{\mathcal{P}_1\,\mathcal{P}_2}{1+\mathcal{P}}$$



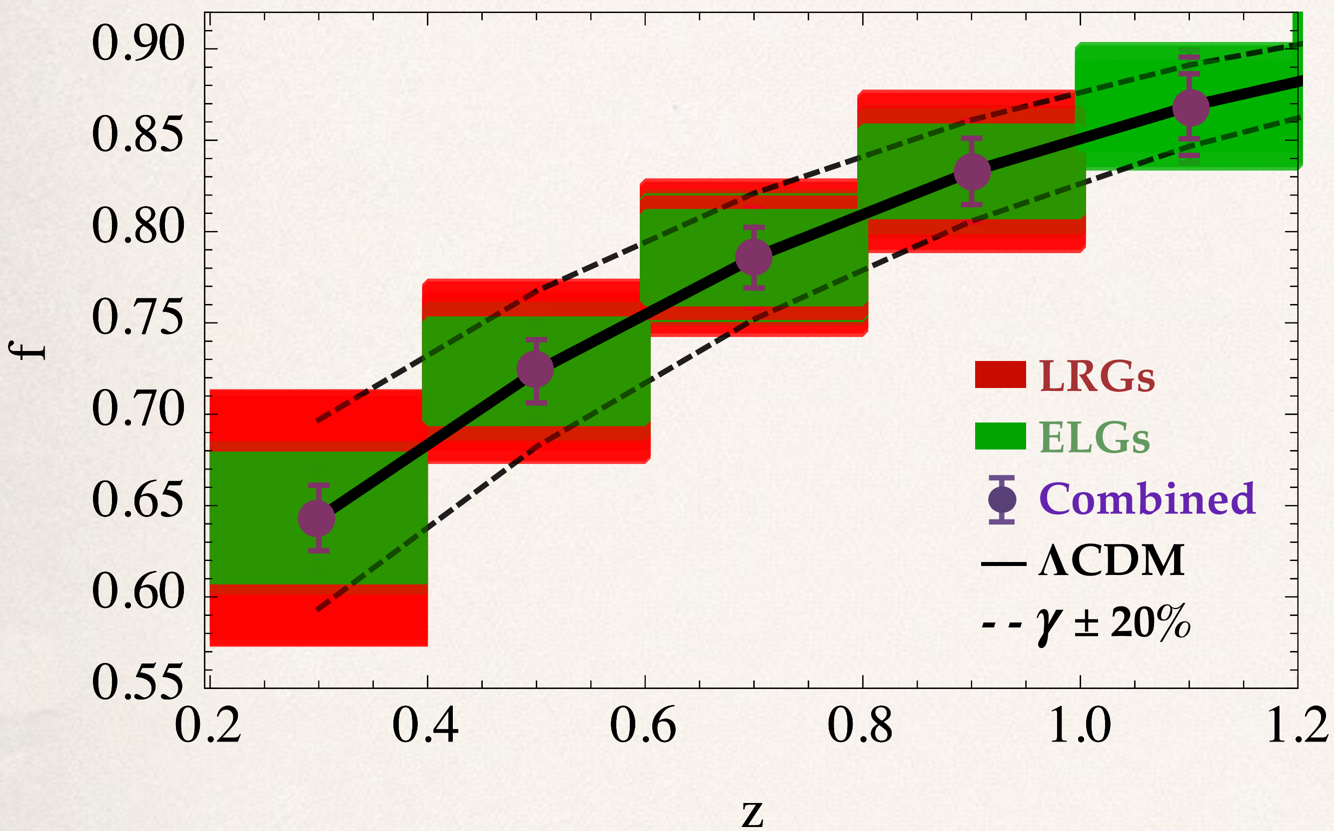
Example: RSDs in J-PAS

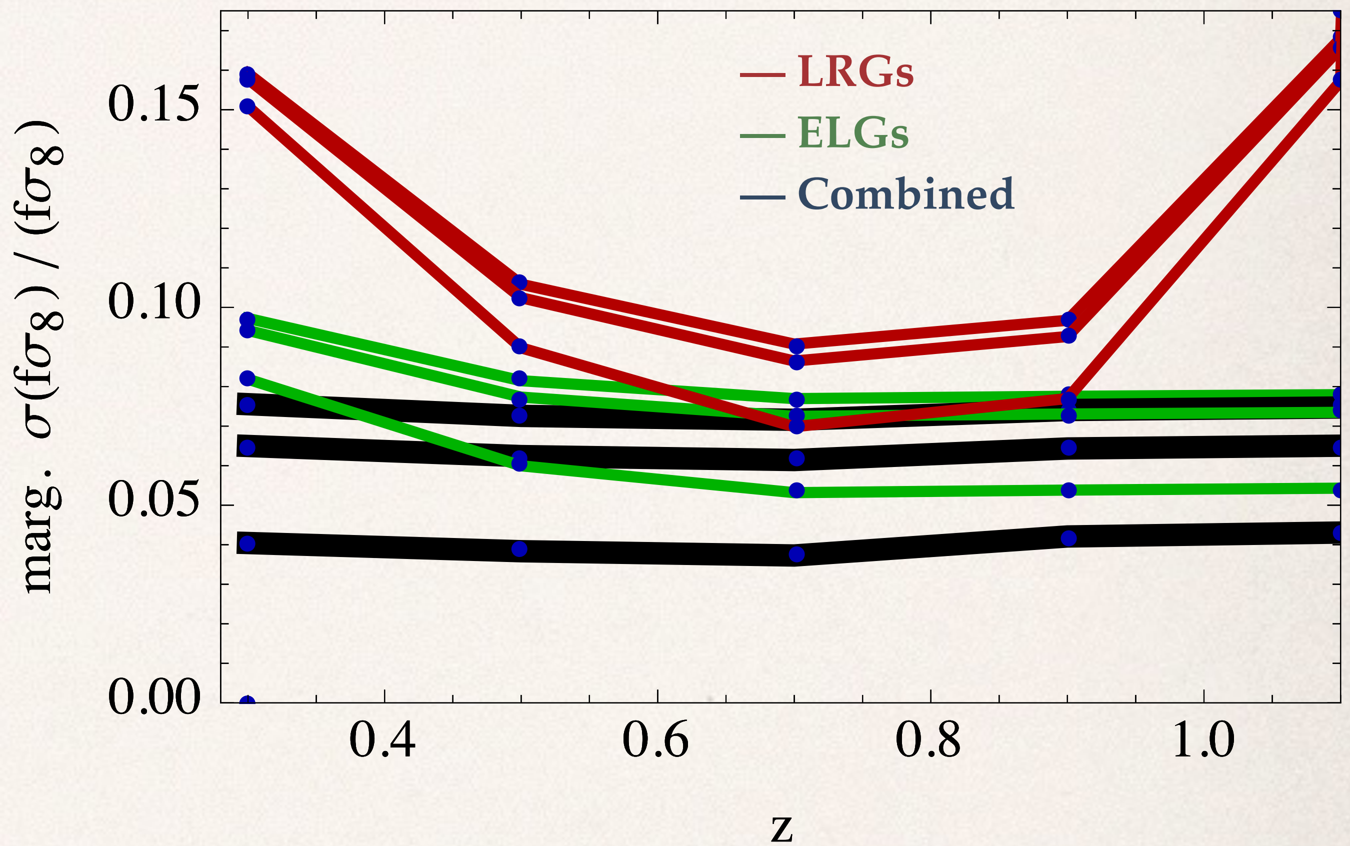
Marginalized* errors on matter growth rate

Modified gravity ($\gamma_{GR} \cong 0.55$)

$$\mathcal{P}_g = n_g \, (b_g + f \, \mu_k^2)^2 \, P(k;z)$$

$$f = \frac{d \ln G}{d \ln a} \simeq \Omega_m^{\gamma}$$





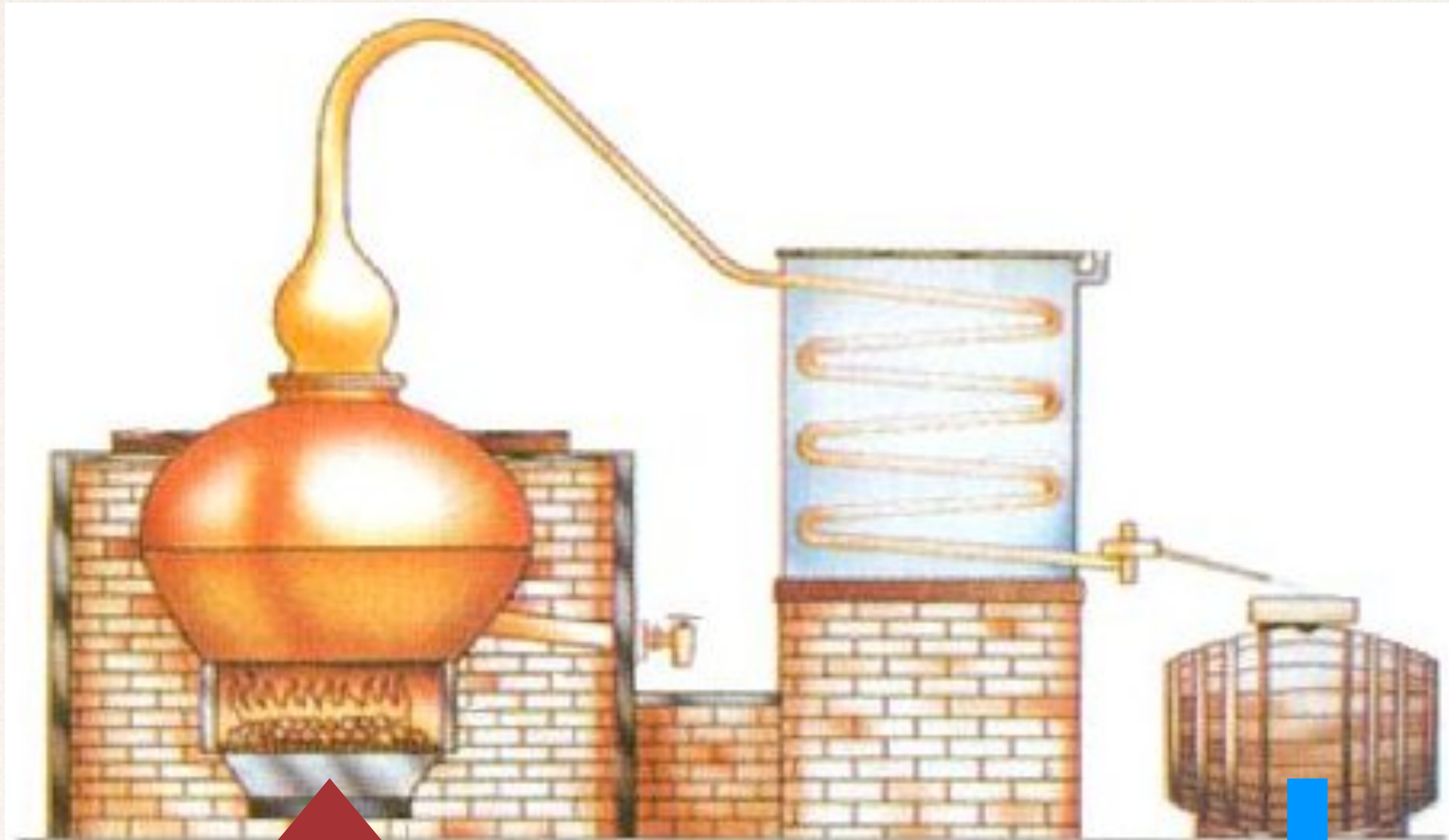
No prior on bias

Weak (~20%) prior on bias

Strong (~5%) prior on bias

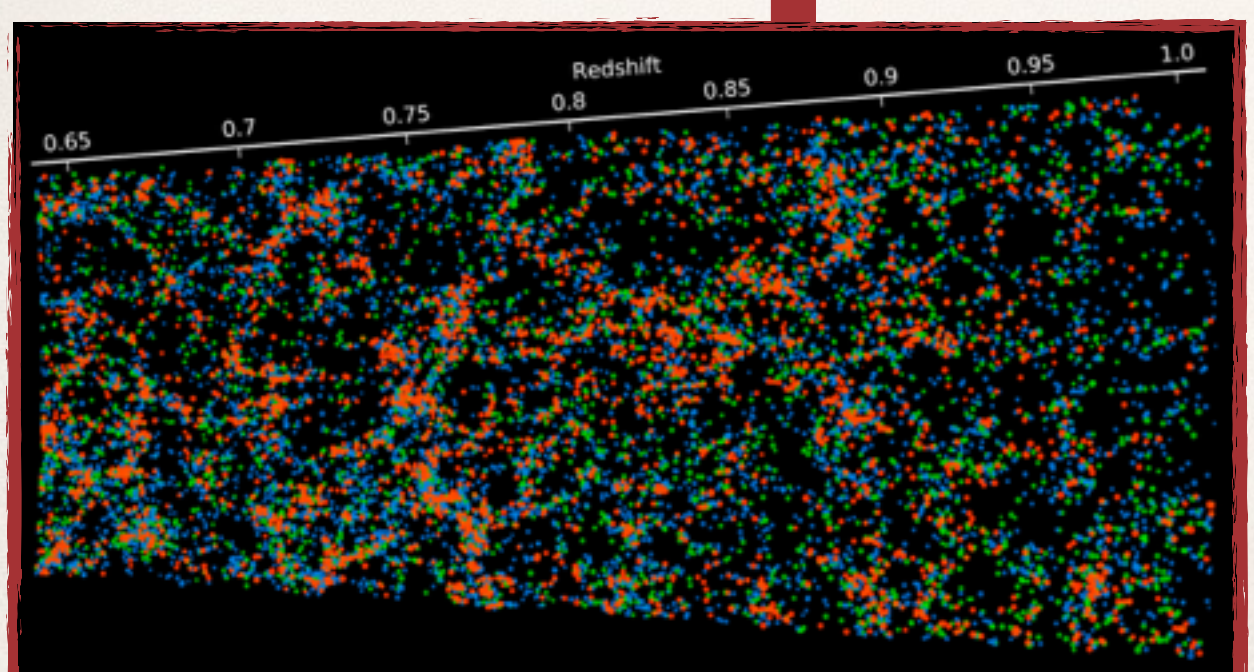
* Marginalized 7 "global" cosmological parameters ($\boldsymbol{\Omega}_m$, *h*, etc.) + 5 parameters on ***each*** redshift slice

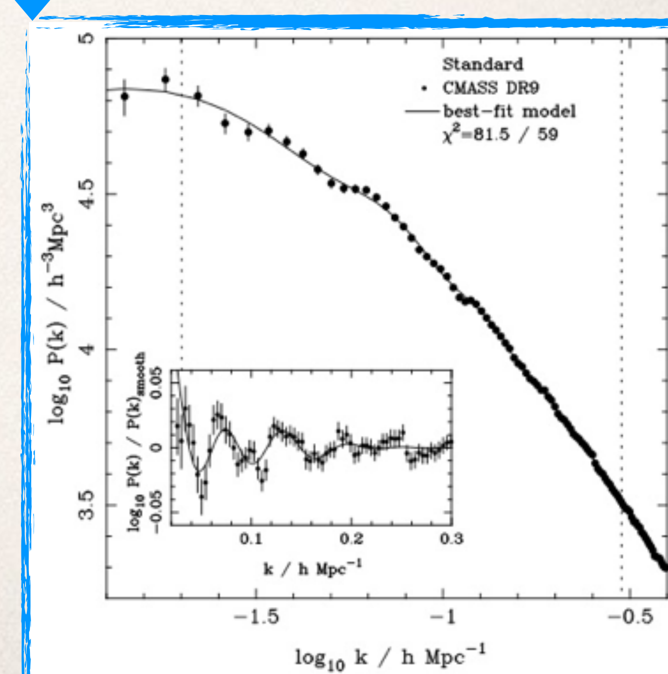
How to do it in practice?



Types and
positions
of galaxies

P(k), BAOs
RSDs
NGs, etc.





Feldman, Kaiser & Peacock, 1994 (**FKP**)

Fourier analysis of galaxy surveys

Given a **galaxy catalog $n_g(x)$**, the **optimal estimator** for the spectrum (FKP) is:

$$\frac{\delta n_g(x)}{\bar{n}_g} = \delta_g(x) \longrightarrow f_g(x) = w_{FKP}(x)\,\delta_g(x)$$

The **FKP weights** express the best **compromise** between **cosmic variance** and **shot noise**:

$$w_{FKP} = \frac{1}{1+\mathcal{P}_g}\,\bar{n}_g\,B_g \qquad\qquad B_g = b_g(z) + f(z)\mu_k^2 + \dots$$

The **estimated spectrum** for a Fourier bin k_i (the *bandpower*) is then:

$$\hat{P}_g(k_i) = \frac{1}{N}\,\langle|\tilde{f}_g|^2\rangle_{k_i} \qquad\qquad \tilde{f}_g(\mathbf{k}) = \mathrm{FFt}[f_g(\mathbf{x})]$$

The FKP estimator is **optimal** — it is **unbiased**, and it saturates the **Cramér-Rao bound**:

$$Cov[\hat{P}_g(k_i), \hat{P}_g(k_j)] \rightarrow [\mathrm{Fisher}]^{-1}$$

Fourier analysis of multi-tracer surveys

Given **any number of galaxy catalogs** $\boldsymbol{n_\mu(x)}$, the **optimal estimator** for the auto-spectra is:

$$\delta_\mu(x) \longrightarrow f_\mu(x) = \sum_\nu w_{\mu\nu}\, \delta_\nu(x)$$

The multi-tracer **weights** are:

$$w_{\mu\nu} = \left[\delta_{\mu\nu} - \frac{\mathcal{P}_\mu}{1+\mathcal{P}} \right] \bar{n}_\nu B_\nu$$

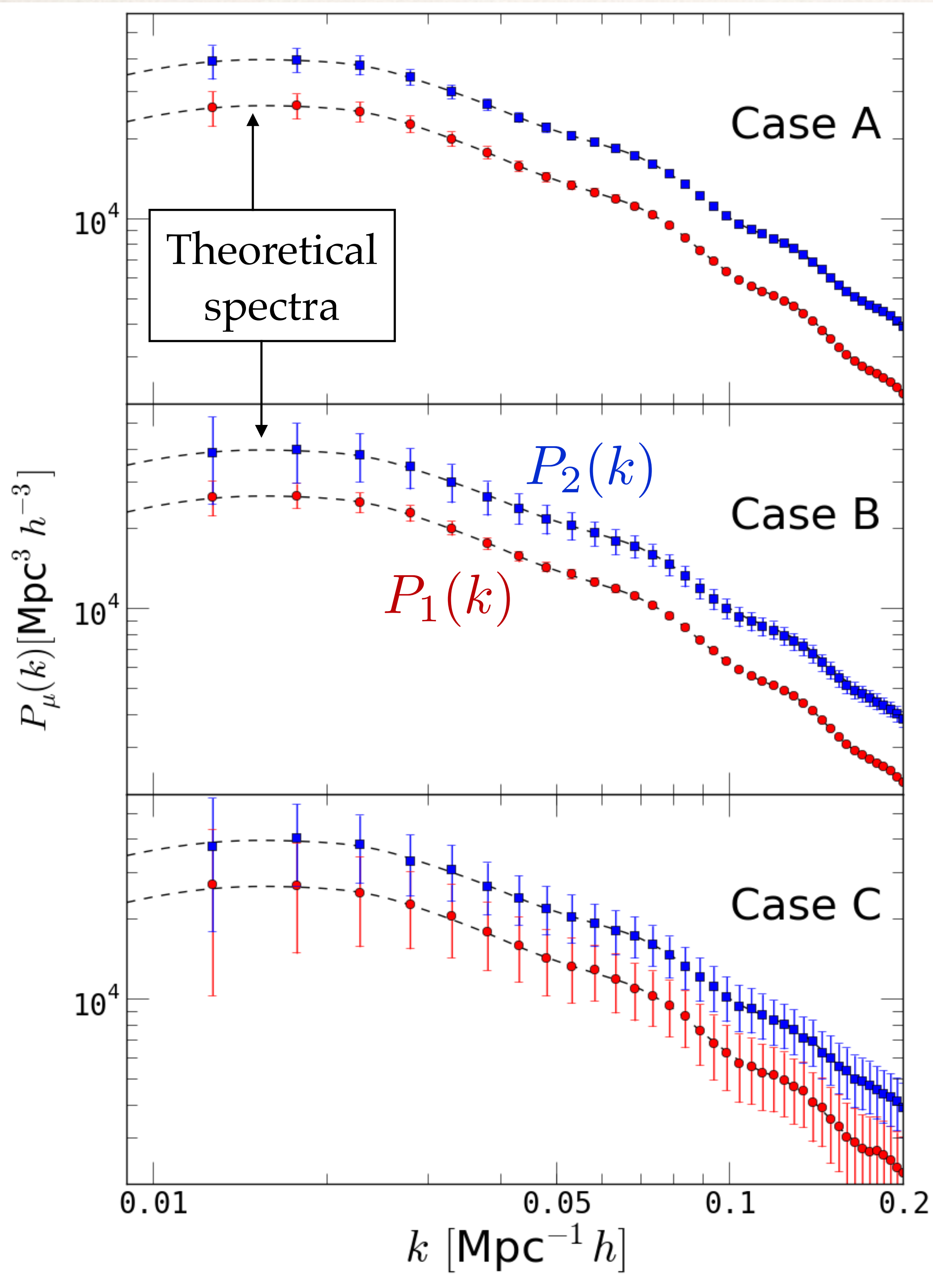
The **estimated auto-spectra** are:

$$\hat{P}_\mu(k_i) = \sum_\nu N^{-1}_{\mu\nu} \, \langle |\tilde{f}_\nu|^2 \rangle_{k_i}$$

The estimators are **optimal**: their **covariance** is the **inverse of the Fisher matrix**!

$$Cov[\hat{P}_\mu, \hat{P}_\nu] = [F_{\mu\nu}]^{-1}$$

Testing and validating the multi-tracer estimators



Case	$\bar{n}_1\ (h^3\ \mathrm{Mpc}^{-3})$	b_1	$\bar{n}_2\ (h^3\ \mathrm{Mpc}^{-3})$	b_2
A	$1.\,10^{-2}$	1.0	$1.\,10^{-2}$	1.2
B	$1.\,10^{-2}$	1.0	$1.\,10^{-5}$	1.2
C	$1.\,10^{-5}$	1.0	$1.\,10^{-5}$	1.2

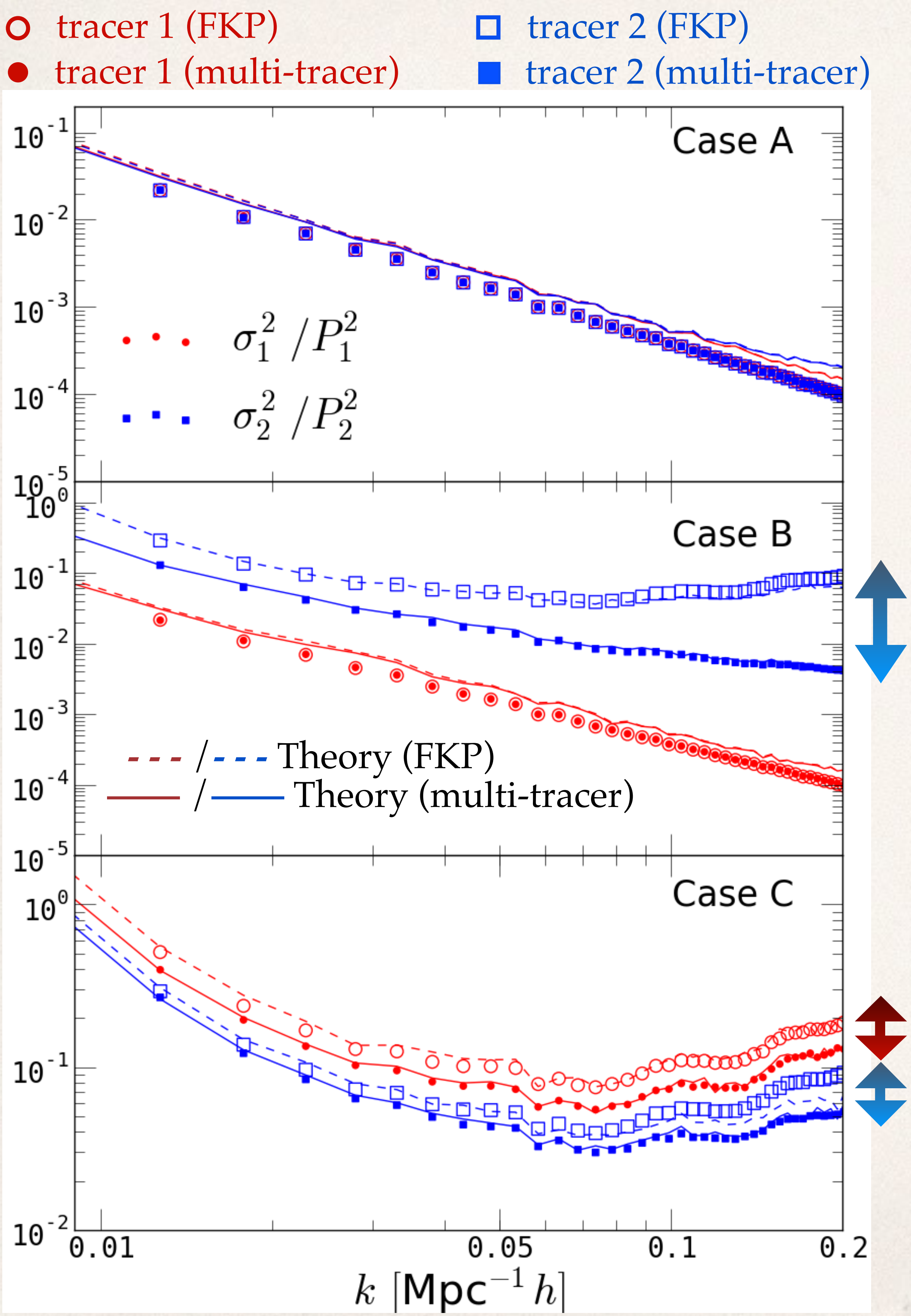
* Volume = $(1280\ h^{-1}\ \mathrm{Mpc})^3 = (128 \times 10\ h^{-1}\ \mathrm{Mpc})^3$

* 10^3 lognormal realizations

* Planck-vanilla fiducial parameters

✓ Unbiased

Testing and validating the multi-tracer estimators

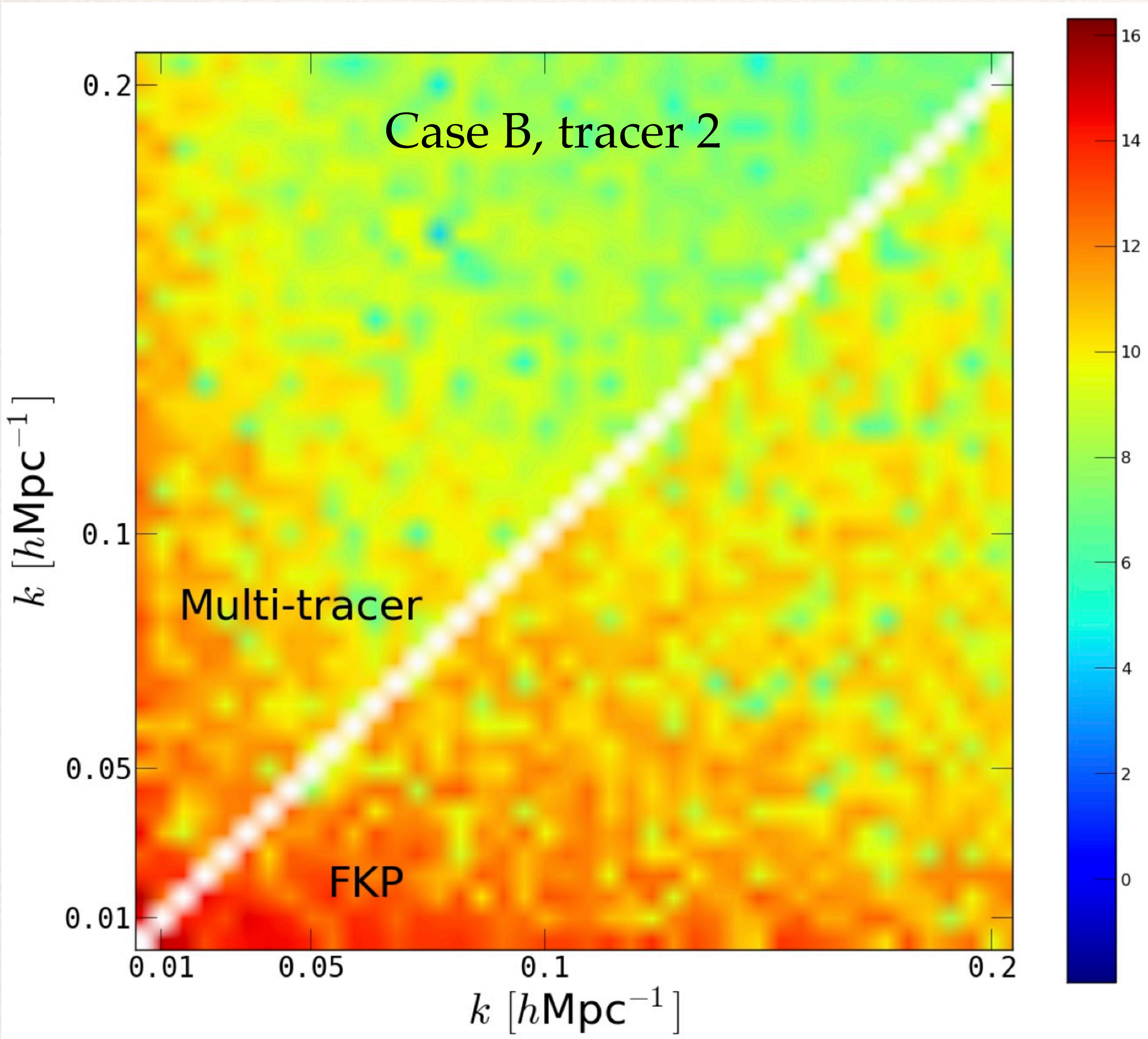


Case	$\bar{n}_1$ (h^3 Mpc^{-3})	b_1	$\bar{n}_2$ (h^3 Mpc^{-3})	b_2
A	$1.\,10^{-2}$	1.0	$1.\,10^{-2}$	1.2
B	$1.\,10^{-2}$	1.0	$1.\,10^{-5}$	1.2
C	$1.\,10^{-5}$	1.0	$1.\,10^{-5}$	1.2

✓ Unbiased
✓ Optimal

Testing and validating the multi-tracer estimators

Covariance matrix $Cov[P_2(k_i),P_2(k_j)]$



Case	$\bar{n}_1\ (h^3\ \mathrm{Mpc}^{-3})$	b_1	$\bar{n}_2\ (h^3\ \mathrm{Mpc}^{-3})$	b_2
A	$1.\,10^{-2}$	1.0	$1.\,10^{-2}$	1.2
B	$1.\,10^{-2}$	1.0	$1.\,10^{-5}$	1.2
C	$1.\,10^{-5}$	1.0	$1.\,10^{-5}$	1.2

✓ Unbiased

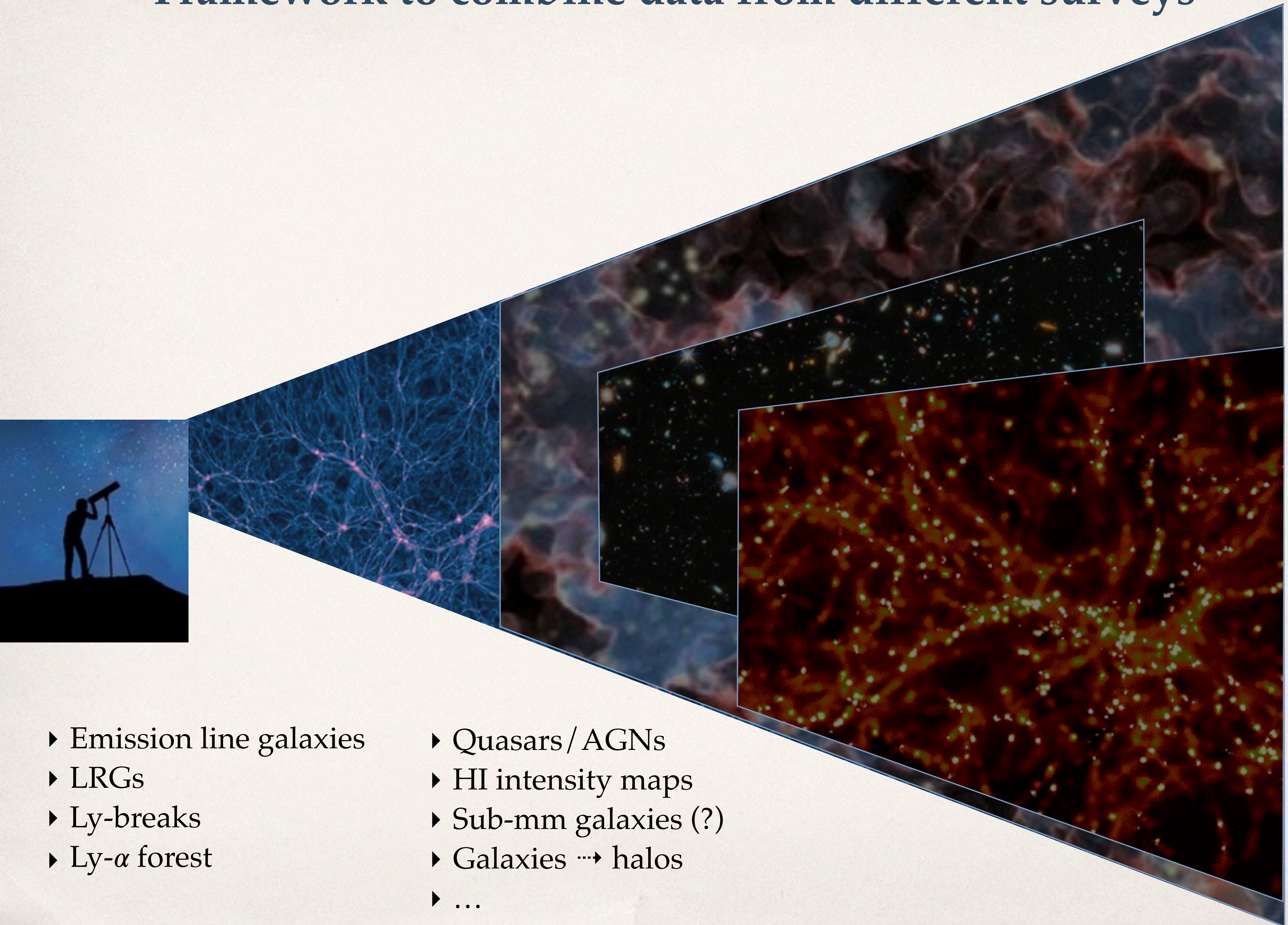
✓ Optimal

✓ Covariance matrix between bandpowers is well-behaved (better than FKP): diagonal terms dominate on all but largest scales

Summaryzing:

- The multi-tracer estimators achieved a **superior performance** compared to FKP, in **all the cases** we analyzed
- FKP methods, as well as generalizations such as that by Percival, Verde & Peacock (2004), are **only** able to estimate the **matter power spectrum**, $P_m(k)$ — **given fixed** (known) values for the bias, RSDs, etc.
- The multi-tracer estimator not only recovers the matter power spectrum, $P_m(k)$, but it also **estimates the biases of the tracers, RSDs, NGs — even the HODs** (see R. A., Secco and Loureiro 2015)
- The multi-tracer technique takes advantage of the fact that all tracers live on the **same underlying dark matter density field** — without assuming anything about how **differently** they trace that field

Framework to combine data from different surveys

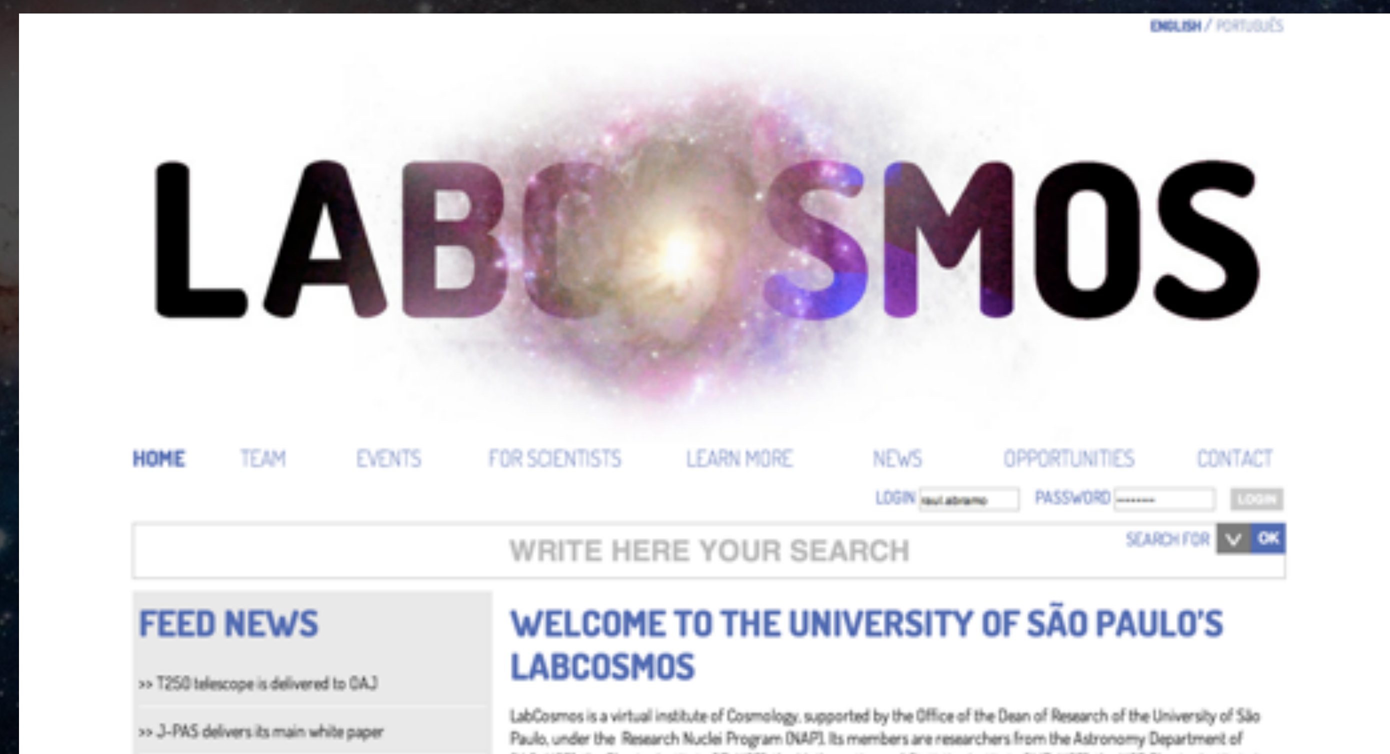


- Emission line galaxies
- LRGs
- Ly-breaks
- Ly-α forest

- Quasars / AGNs
- HI intensity maps
- Sub-mm galaxies (?)
- Galaxies ⇢ halos
- …

Thank you!

http://j-pas.org



http://www.iag.usp.br/labcosmos