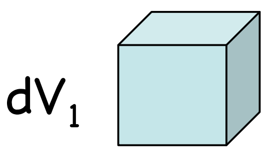
Lectures III and IV

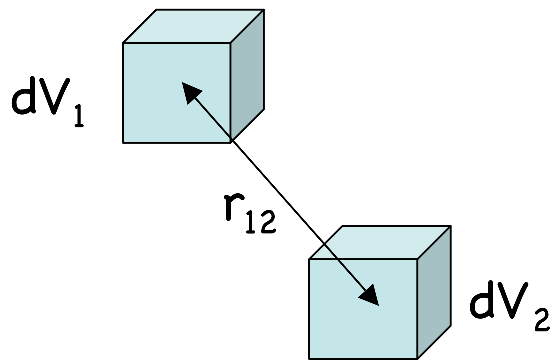
Cristiano Porciani

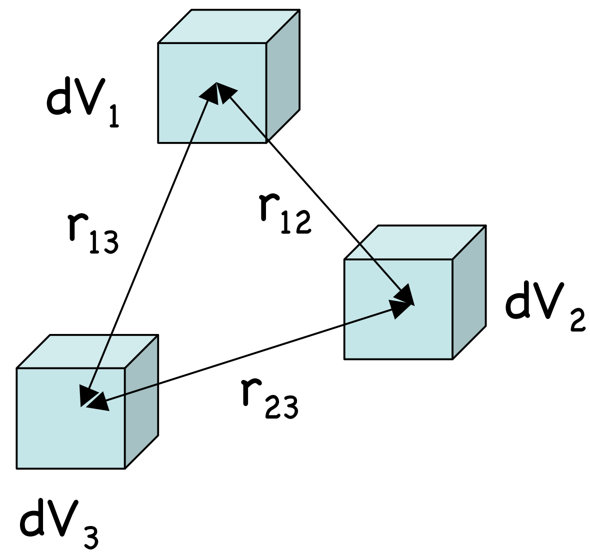
AIfA, Bonn

Correlation functions

Consider a stationary point process with mean density n and write the probability of finding N points within N infinitesimal volume elements



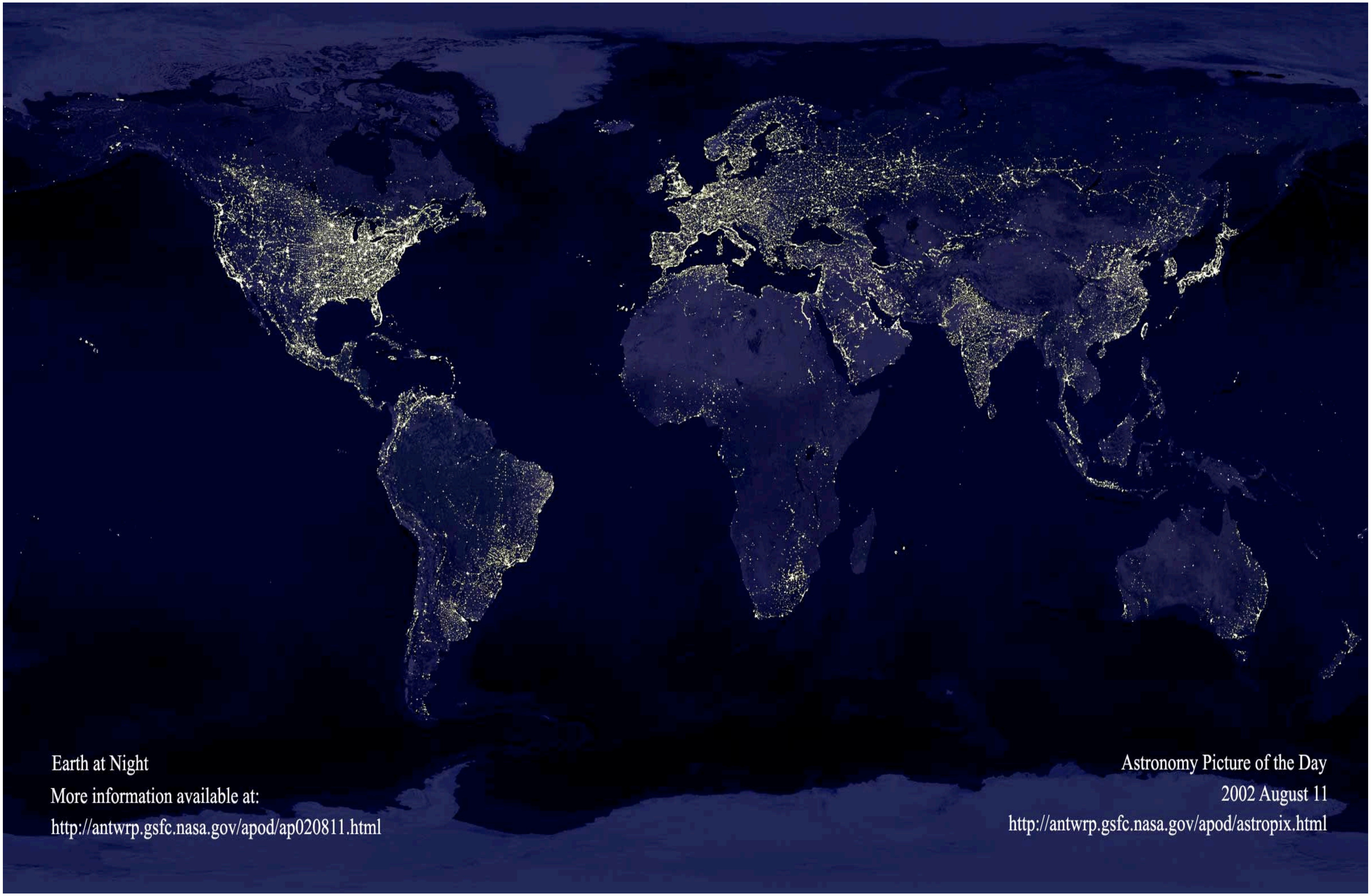




$dP_1 = n\, dV_1$

$dP_{12} = n^2\, [1+\xi(r_{12})]\, dV_1\, dV_2$

$dP_{123} = n^3\, [1+\xi(r_{12})+ \xi(r_{13})+ \xi(r_{23})+\zeta(r_{12}, r_{13}, r_{23})]\, dV_1\, dV_2\, dV_3$


Earth at Night
More information available at:
http://antwrp.gsfc.nasa.gov/apod/ap020811.html
Astronomy Picture of the Day
2002 August 11
http://antwrp.gsfc.nasa.gov/apod/astropix.html

Different galaxies cluster differently

96791 2dF galaxies with redshifts between 0.01 and 0.15

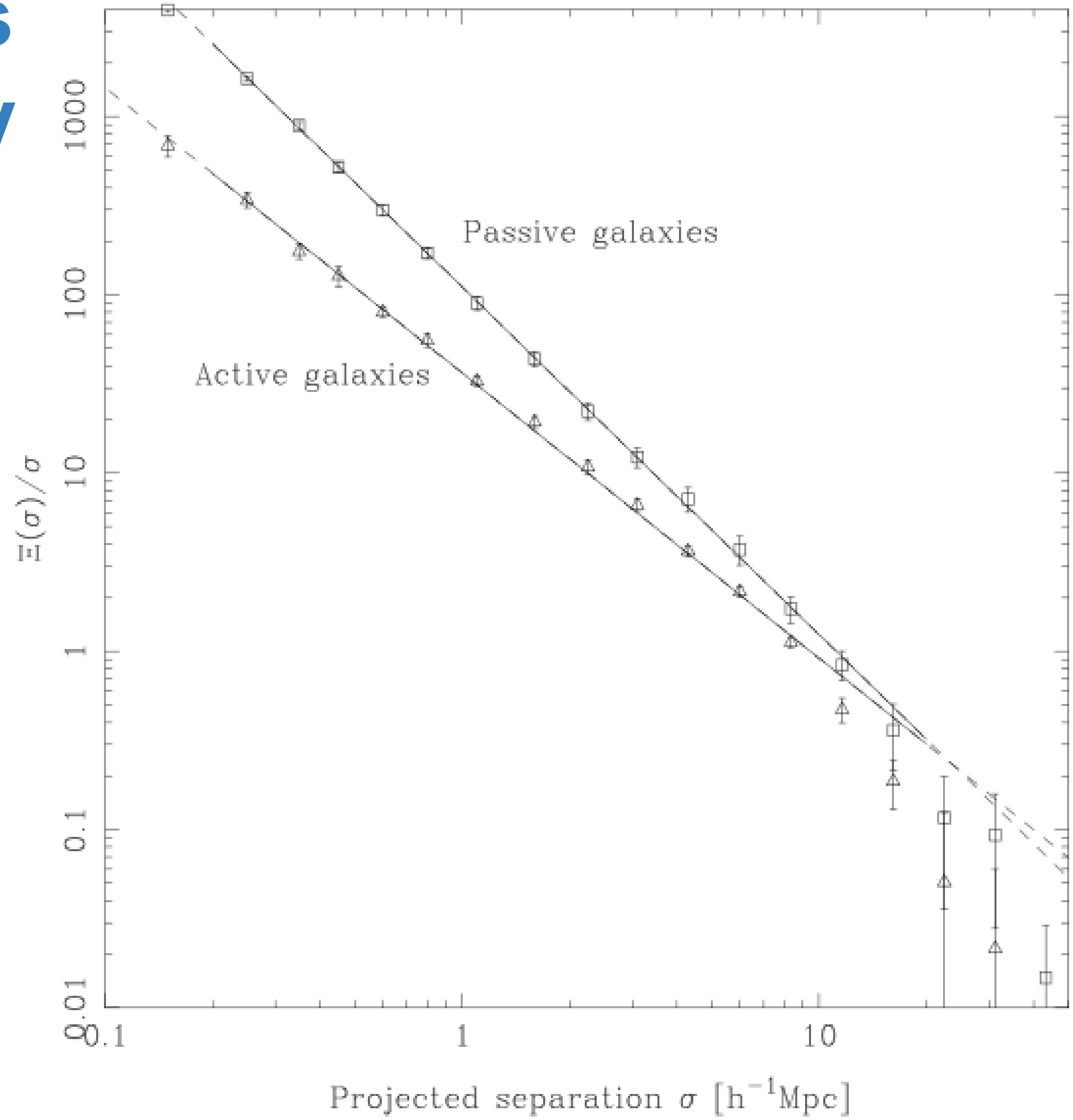
Spectral classification based on principal component analysis

Power-law fit, $\xi(r) = (r_0/r)^\gamma$

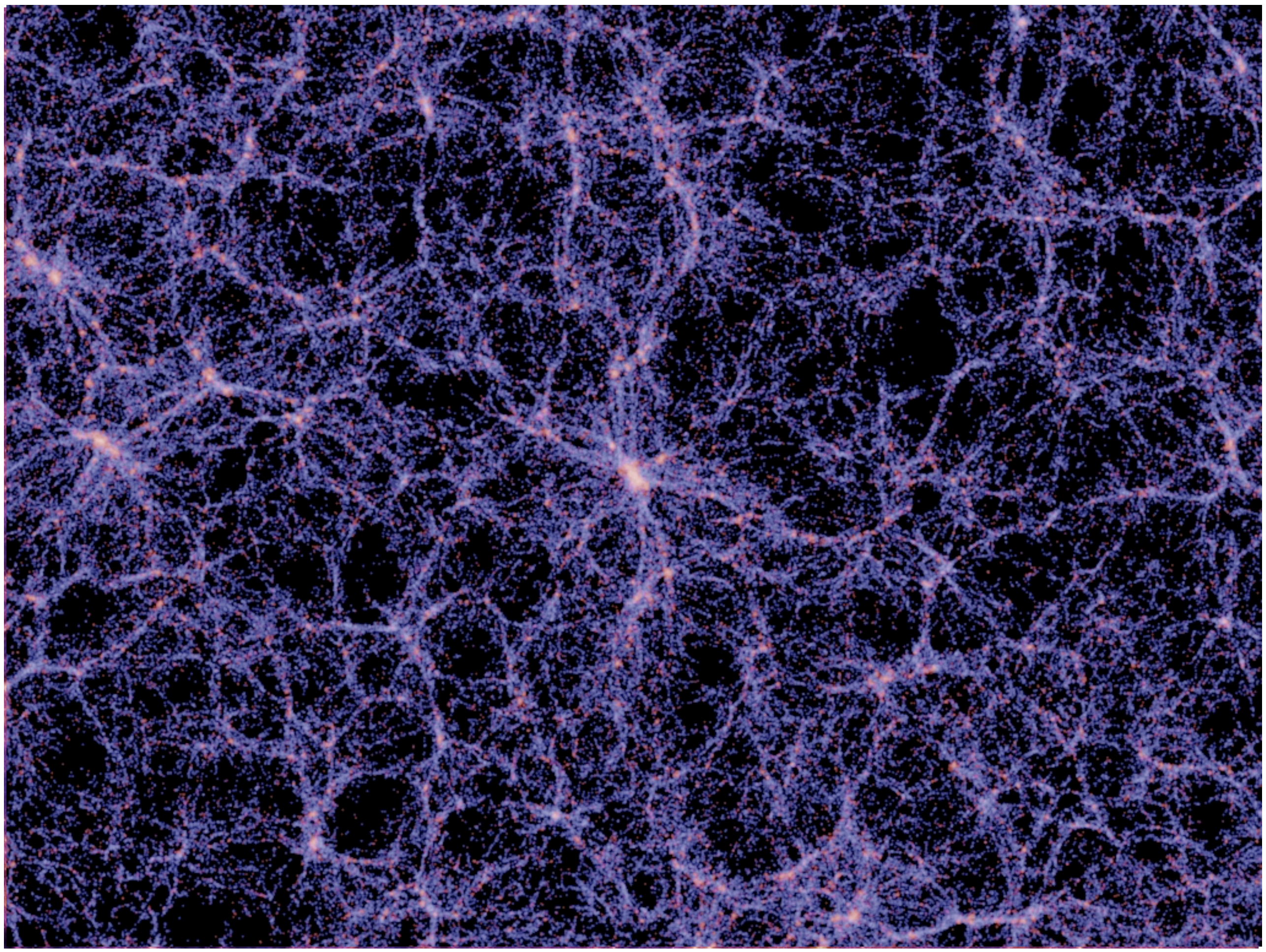
Early-type (passively evolving):
$\gamma = 1.95 \pm 0.03$,
$r_0 = 6.10 \pm 0.34$ Mpc/h

Late-type (star-forming):
$\gamma = 1.60 \pm 0.04$,
$r_0 = 3.67 \pm 0.30$ Mpc/h

Madgwick et al. (2003)



Galaxies and DM Halos as Tracers of the Cosmic Web



Courtesy V. Springel

At the 2-point level, biasing can be represented by the ratio between the correlation of the tracers and of the underlying mass distribution

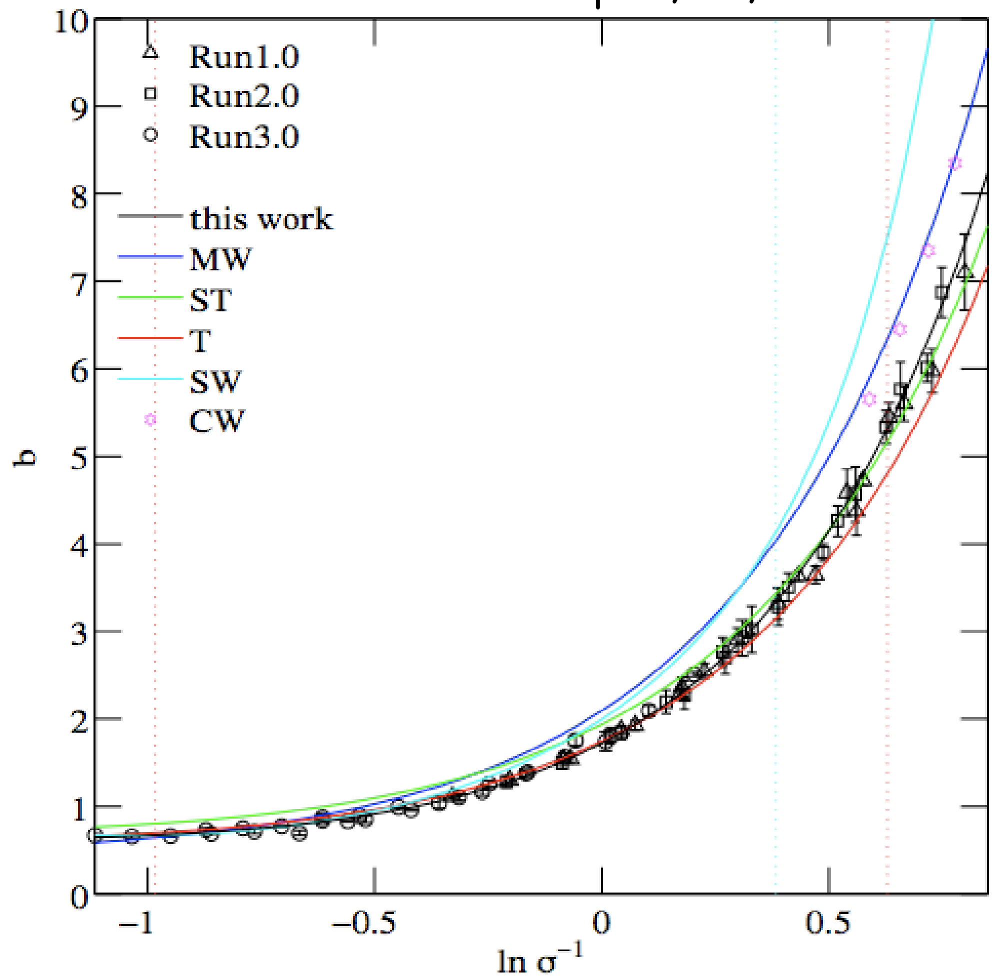
$\xi_{tt}(r)=b^2(r)\ \xi_{mm}(r)$

For large spatial separations (r>> 10Mpc),

$b \rightarrow$ constant

with a numerical value which depends on the mass of the tracers

Pillepich, CP, Hahn 2009



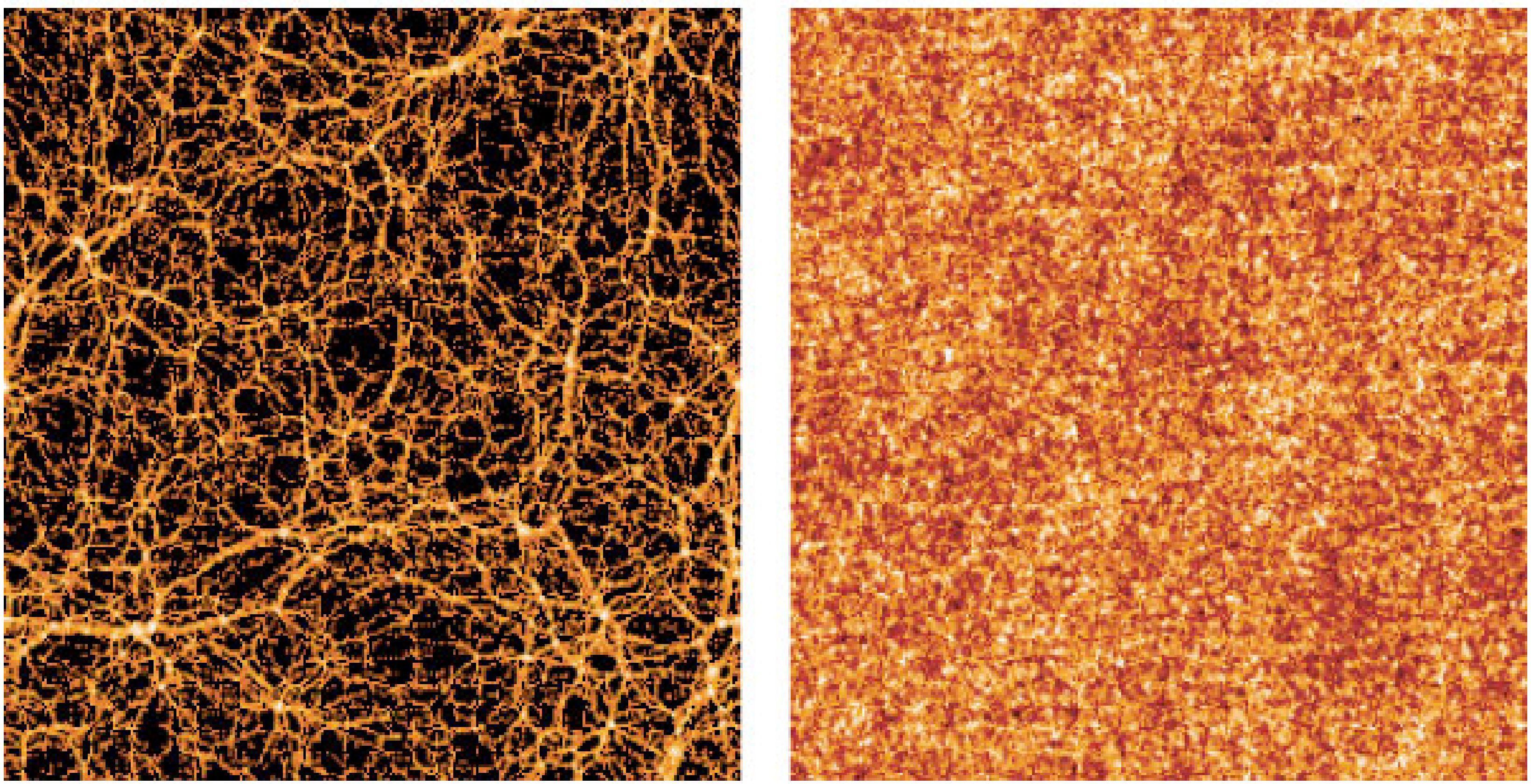
Lauro is desperate! He spent all the conference money on the social dinner. He now has only one possibility to keep his job…

The game show has new rules: 12 questions, you can call a friend 3 times but it must be for three consecutive questions, you can withdraw any time but if you answer wrong you loose everything.

Lauro, however, does not know much beyond the SZ effect and so he will answer at random.

For this reason he asked for help to Alan and promised to answer what Alan suggests for the 3 questions.
Knowing that Lauro lost after giving a wrong answer at the 6^{th} question, what can you say about Alan's success rate? (the population value…not the sample one)

Beyond 2-point statistics: preserving phase information



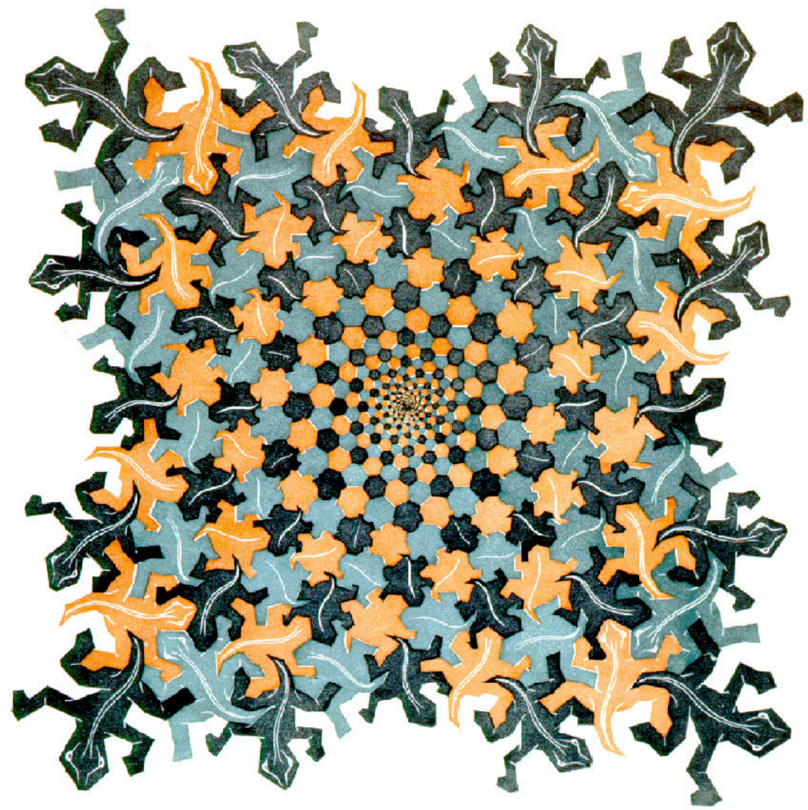
Coles & Chiang 2000

Primordial non-Gaussianity

The origin of structure

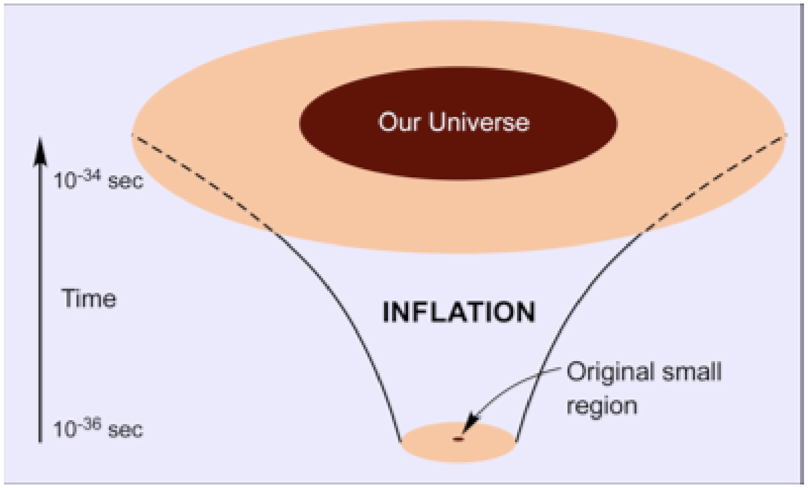
- One of the major unsolved mysteries in cosmology

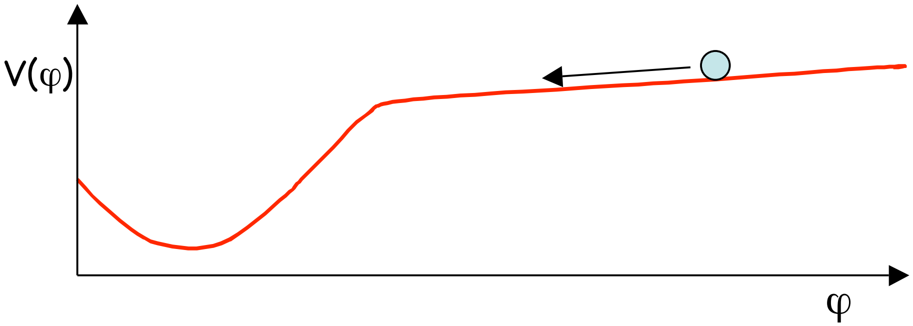
- Many competing theories that differ in their predictions some of which are accessible to astronomical observations today



Canonical inflation

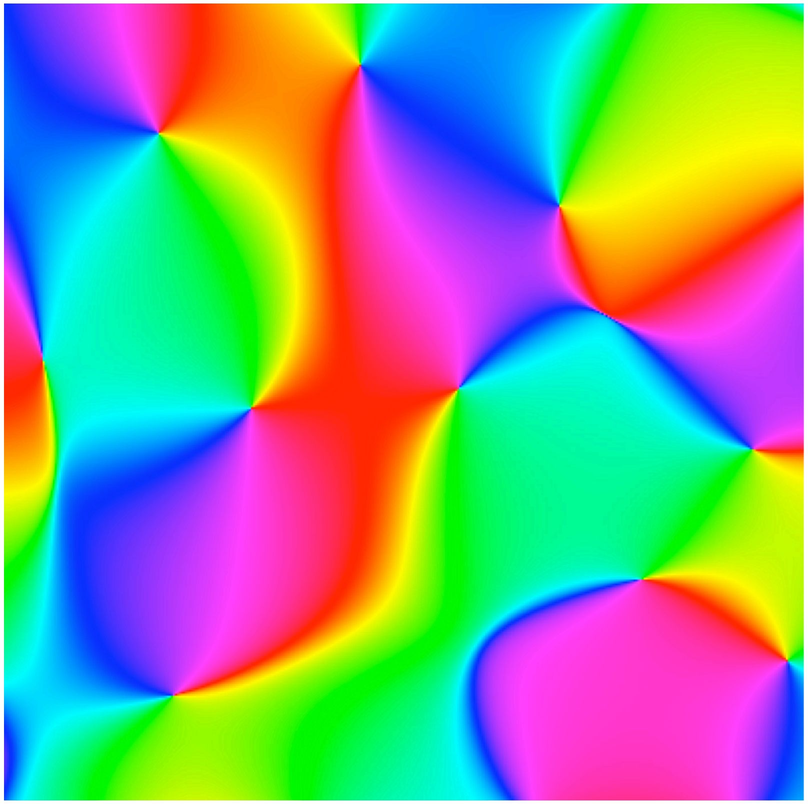
- Our universe derives from the exponential expansion of a tiny causally connected patch
- The accelerated expansion is driven by a spatially homogeneous scalar field (the inflaton) which is slowly rolling down its potential
- Introduced to address problems that were eating away at the foundations of the hot big-bang model (flatness, horizon, relic density, expansion)





Perturbations and inflation

- Small quantum fluctuations of the inflaton are naturally produced
- The exponential expansion rapidly stretches them to longer and longer wavelengths
- On super-Hubble scales, fluctuations become overdamped, leading to curvature perturbations that can be described as classical



Courtesy of Prof. Sir M. Berry

Inflationary zoo

D-term

curvaton

R^2

eternal

new

brane

ghost

oscillating

curvaton scenario

double

hybrid

Two scalar fields: one drives the accelerated expansion (the inflaton) and the second one is responsible for the formation of structure with its quantum fluctuations (the curvaton)

extended

tachyon

F-term

chaotic

multi-field

old

natural

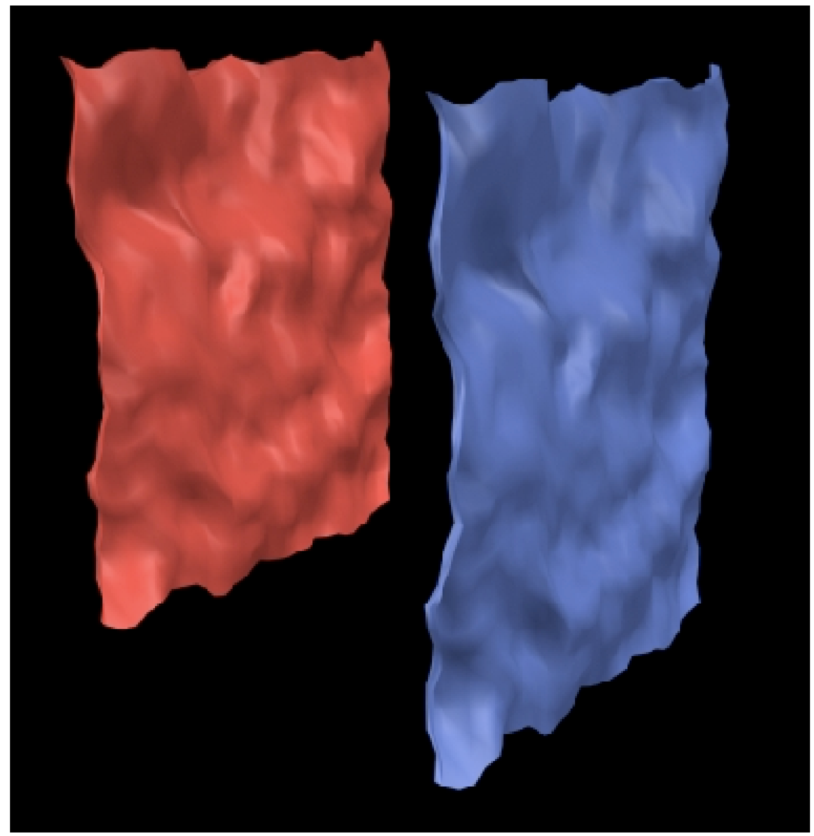
super-natural

extra-natural

Ekpyrotic/cyclic models

- Inspired by heterotic M-theory
- The collision of 2 parallel branes embedded in an extradimensional bulk marks the beginning of the hot, expanding phase of the universe
- Prior to the collision the universe is contracting, perturbations are generated during this phase

Khoury et al. 2001, 2002, Steinhardt & Turok 2002



Model predictions

	Canonical inflation	Curvaton scenario	Ekpyrotic universe
Geometry	Flat √ WMAP	Flat √ WMAP	Flat √ WMAP
Spectrum of perturbations	Nearly scale-invariant √ WMAP	Nearly scale-invariant √ WMAP	Nearly scale-invariant √ WMAP
Statistics of perturbations	Nearly Gaussian	Non-Gaussian	Non-Gaussian

A quote from S.F. Shandarin

Speaking of non-Gaussianity is like speaking of non-dogs!

Spectra and multi-spectra

Power spectrum

$$\left\langle \phi(\vec{k}_1)\phi(\vec{k}_2) \right\rangle = (2\pi)^3 \delta(\vec{k}_1 + \vec{k}_2) P(k_1)$$

$$\left\langle \phi(\vec{k}_1)\phi(\vec{k}_2)\phi(\vec{k}_3) \right\rangle = (2\pi)^3 \delta(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) B(k_1, k_2, k_3)$$

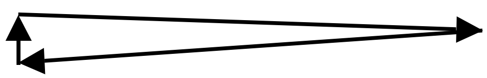
Bispectrum

For a Gaussian field, B=0

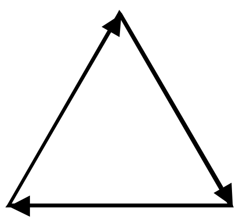
Shape of non-Gaussianity

Depending on the properties of the bispectrum of the gravitational potential, non-Gaussianity can be broadly classified into two classes (Babich et al. 2004):

- NG of the local (squeezed) form, where $B(k_1, k_2, k_3)$ is dominated by the configurations with $k_1 \ll k_2 \approx k_3$
  (curvaton models, Ekpyrotic/cyclic models)



- NG of the equilateral form, where $B(k_1, k_2, k_3)$ is dominated by the configurations with $k_1 \approx k_2 \approx k_3$
  (ghost inflation, DBI inflation)



A particularly simple model

Most of the local models can be reduced to the simple form (Salopek & Bond 1990; Falk et al. 1993; Gangui et al. 1994):

$$\Phi(\vec{x}) = \phi(\vec{x}) + f_{NL}\left[\phi^2(\vec{x}) - \langle\phi^2\rangle\right]$$

Bardeen's potential

auxiliary Gaussian field

non-linearity parameter (a real number)

Fractional non-Gaussian corrections are $\approx 10^{-5} f_{NL}$

Predicted non-Gaussianity

f_{NL}		Model	References
$(n_s-1)/4$	≈ 0.01	canonical inflation	Maldacena 2003 Acquaviva et al. 2003
$-5/(4r)$	>10	curvaton	Lyth et al. 2003
≈ 100		multi-field inflation	Linde & Mukhanov 1997
20-100		new ekpyrotic models	Buchbinder et al. 2007, Creminelli & Senatore 2007, Koyama et al. 2007
≈ 100 (equilateral)		DBI inflation	Alishahiha et al. 2004
≈ 100 (equilateral)		ghost inflation	Arkani-Hamed et al. 2004

The state of the art

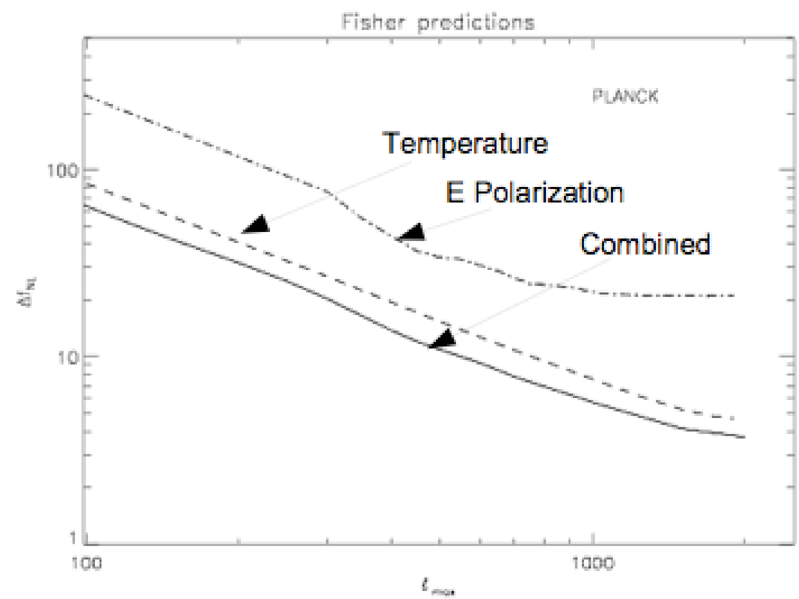
f_{NL} (95% CL)	Data	Method	Estimator	Reference
$-58<f_{NL}<134$	WMAP 1yr	"bispectrum"	KSW	Komatsu et al. 2003
$-27<f_{NL}<121$	WMAP 1yr	"bispectrum"	C+	Creminelli et al. 2006
$-54<f_{NL}<114$	WMAP 3yr	"bispectrum"	KSW	Spergel et al. 2007
$-36<f_{NL}<100$	WMAP 3yr	"bispectrum"	C+	Creminelli et al. 2006
$27<f_{NL}<147$	WMAP 3yr	"bispectrum"	YW	Yadav & Wandelt 2008
$-9<f_{NL}<111$	WMAP 5yr	"bispectrum"	YW	Komatsu et al. 2008
$-4<f_{NL}<80$	WMAP 5yr	"bispectrum"	SZ	Smith et al. 2009
$-70<f_{NL}<91$	WMAP 3yr	Minkowski functionals	-	Hikage et al. 2008
$23<f_{NL}<75$	WMAP 3yr	1-point PDF	-	Jeong & Smoot 2007
$-178<f_{NL}<64$	WMAP 5yr	Minkowski functionals	-	Komatsu et al. 2008
$-8<f_{NL}<111$	WMAP 5yr	wavelet decomposition	-	Curto et al. 2008

Summary of CMB studies

- Current 95% limits from bispectrum: $-9<f_{NL}<111$ and $-4<f_{NL}<80$
- Most-likely value $f_{NL} \approx 40 - 60$
- 7 per cent chance that a Gaussian map gives a higher most-likely value
- Two studies claim detection of non-Gaussianity at high confidence
- Topological methods give somewhat different results

The near future: Planck

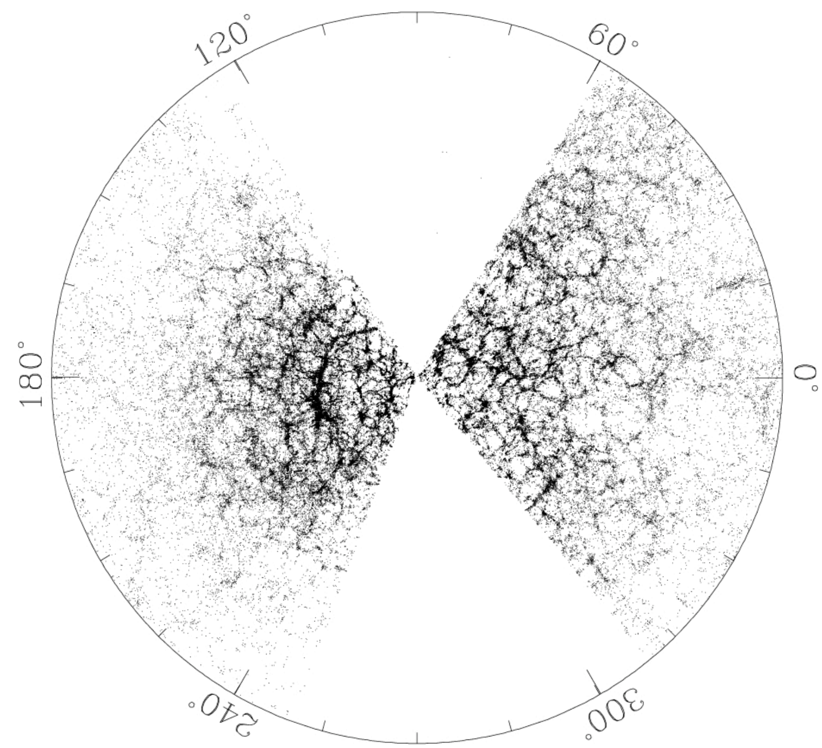
- The Planck satellite (a new European CMB probe) is on his way to the L2 point
- The expected Cramer-Rao limit is $\Delta f_{NL} \approx 5$ from temperature anisotropies (Komatsu & Spergel 2001)
- This reduces to $\Delta f_{NL} \approx 3$ from the joint analysis of temperature and polarization maps (Yadav, Komatsu & Wandelt 2007)



Smaller angular scales

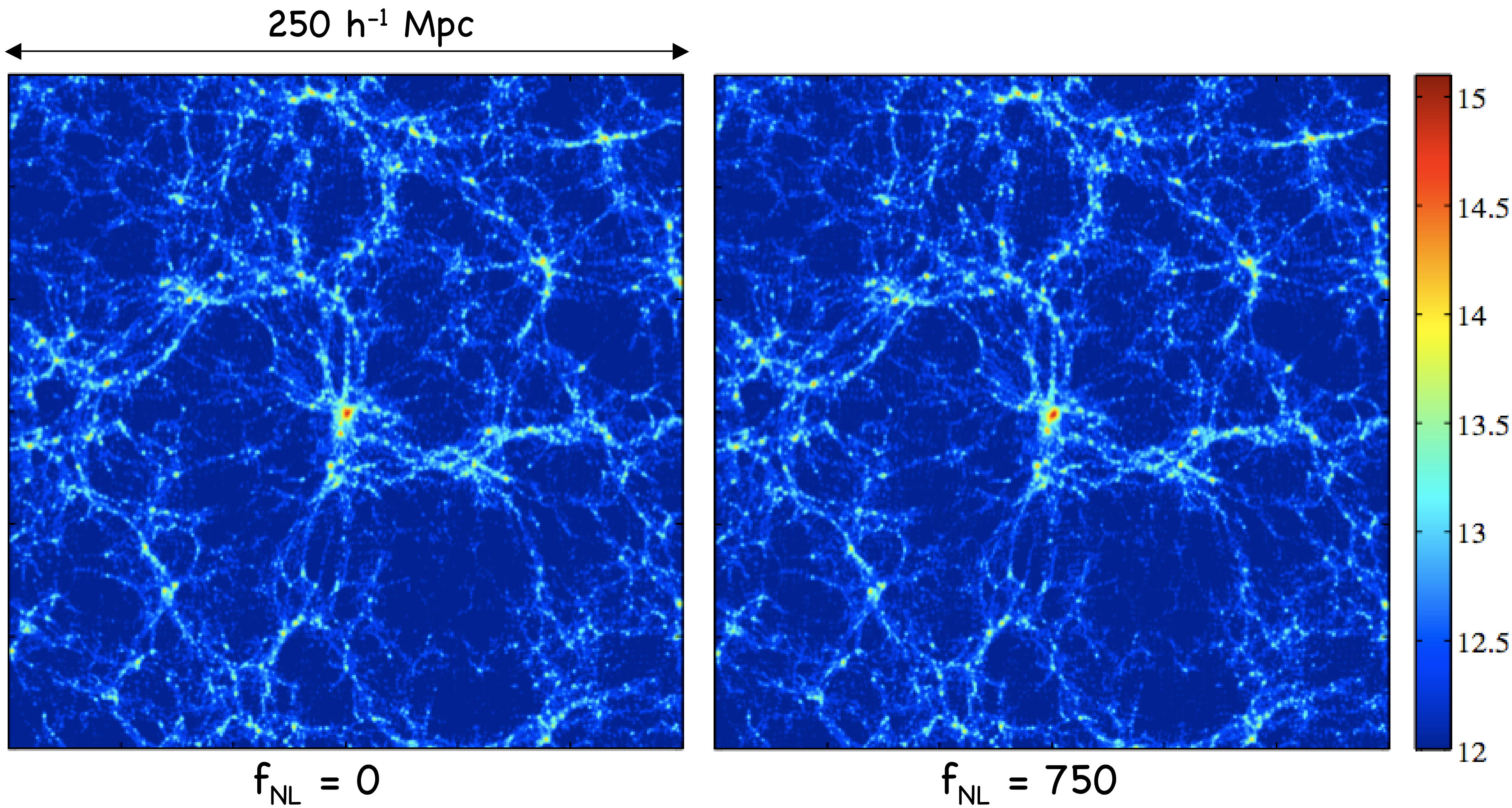
Other probes of non-Gaussianity?

- It would be important to cross-check results against probes with different systematics
- What about the large-scale structure? E.g. the galaxy bispectrum?
- Unfortunately the non-linear growth of perturbations superimposes a much stronger non-Gaussian signal onto the primordial one that is then difficult to recover



(Verde et al. 2000, Scoccimarro et al. 2004, Sefusatti & Komatsu 2007).

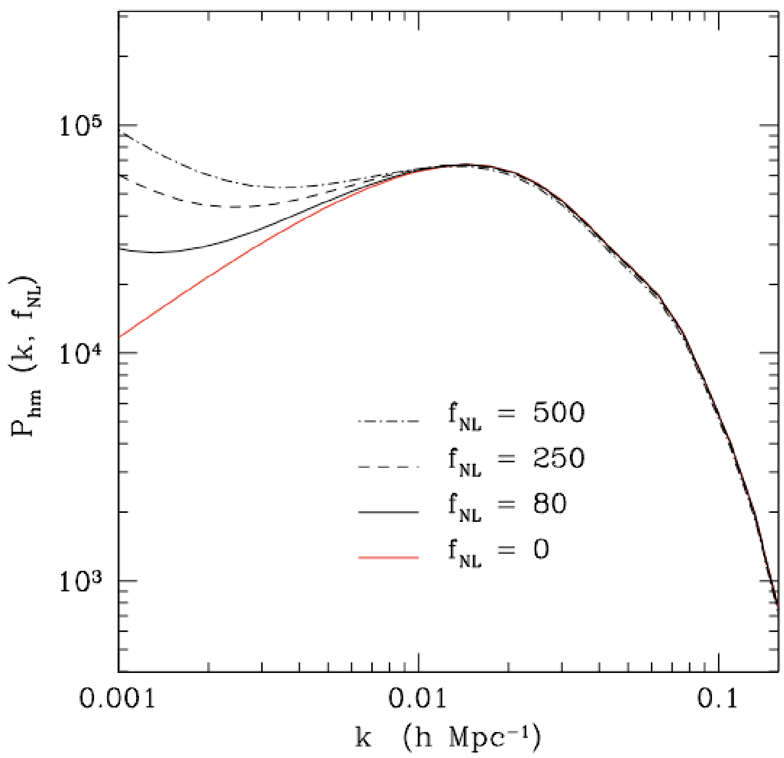
The large-scale structure



Back to life in 2008

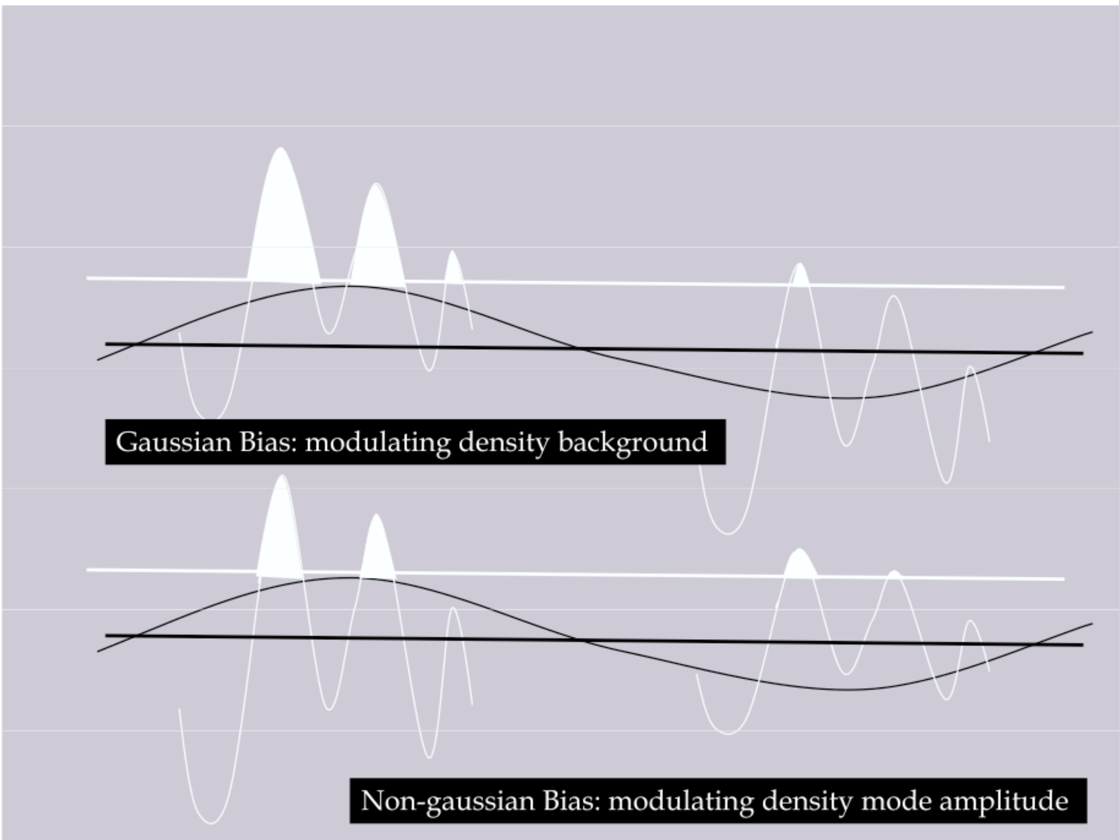
- The large-scale bias of collapsed objects (galaxies, galaxy clusters) as measured by the power spectrum has a strong k-dependence and scales linearly on f_{NL}!!!

- An approximated model based on linear theory captures all the relevant physics (Dalal et al. 2008, Matarrese & Verde 2008, Slosar et al. 2008, Afshordi & Tolley 2008, McDonald 2008)



How does it work?

Afshordi & Tolley 2008



- Gaussian case: long and short wavelength modes of the density field are independent

- Non-Gaussian case: the long modes modulate the amplitude of the short ones via the gravitational potential

$$\delta_s = \delta_G[1 + 2 f_{NL} \phi_l / g(z)]$$

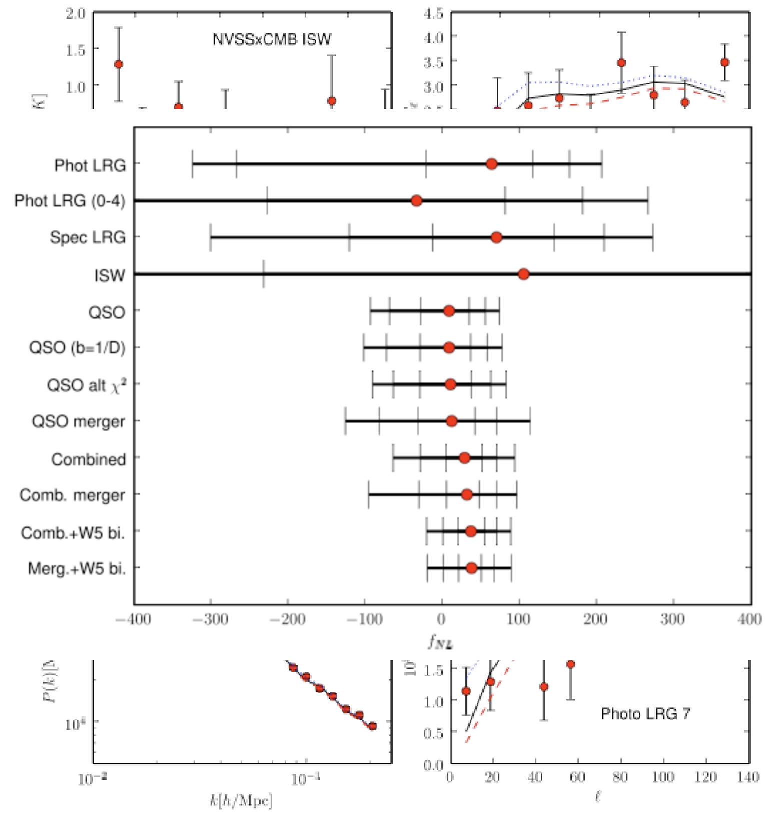
Competitive with CMB!

- Fitting the best datasets for galaxy clustering with the <u>approximated model</u>, gives:
  $-29<f_{NL}<70$ at 95% CL

- This is competitive with the WMAP 5yr results!
  Combined: $0<f_{NL}<69$

- Caveat: how accurate is the first-order model?

Slosar et al. 2008



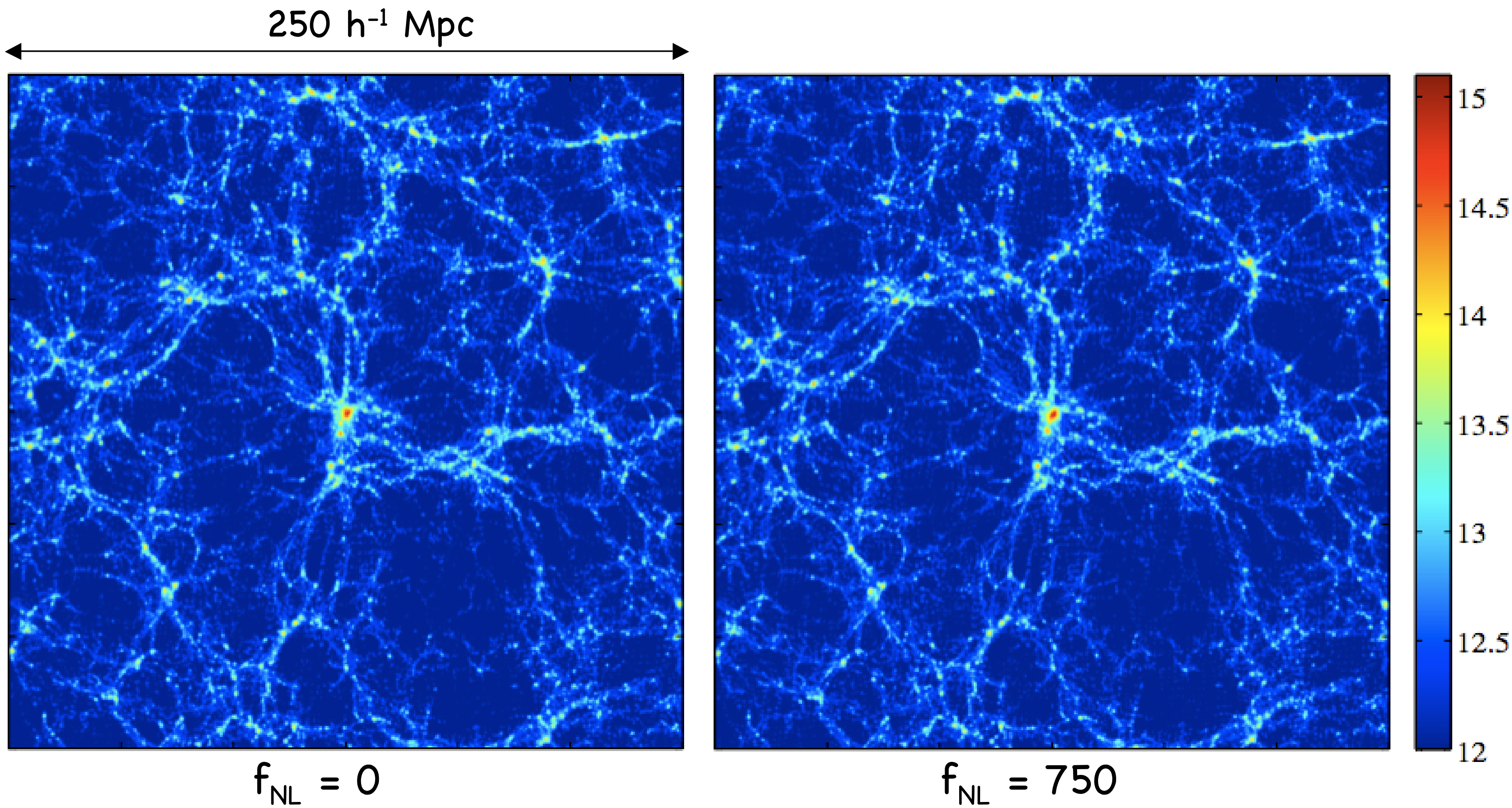
Future prospects

Predicted uncertainties for detection (not for measurement)

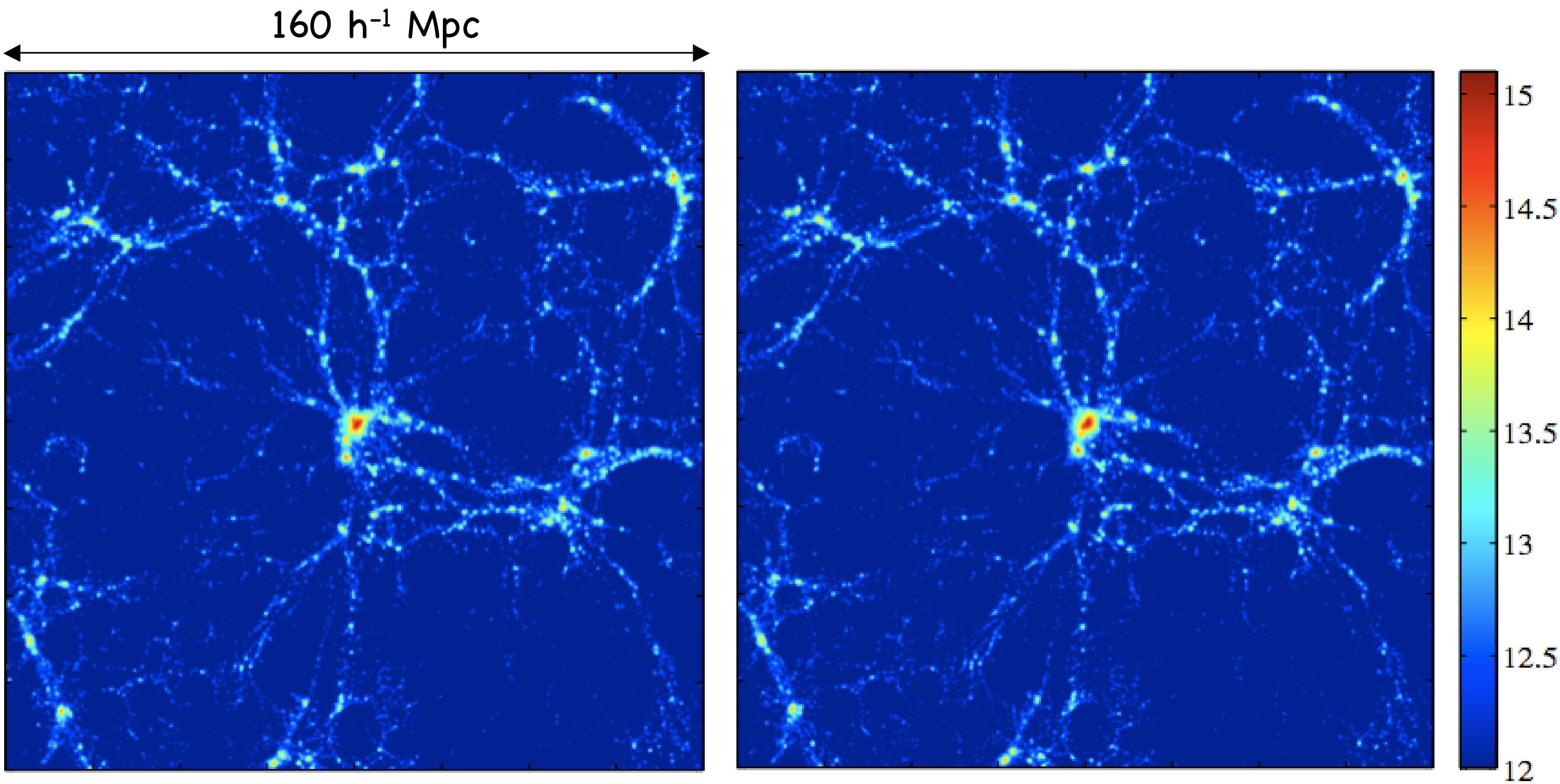
Survey	z Range	Square Degrees	Mean Galaxy Density $(h\ \mathrm{Mpc}^{-1})^3$	$q' \approx 0.8$ $\Delta f_{\mathrm{NL}}/q'$ LSS
SDSS LRGs	$0.16 < z < 0.47$	7.6×10^3	1.36×10^{-4}	37
BOSS	$0 < z < 0.7$	10^4	2.66×10^{-4}	18.8
WFMOS low z	$0.5 < z < 1.3$	2×10^3	4.88×10^{-4}	15
WFMOS high z	$2.3 < z < 3.3$	3×10^2	4.55×10^{-4}	17.5
ADEPT	$1 < z < 2$	2.8×10^4	9.37×10^{-4}	1.7
EUCLID	$0 < z < 2$	2×10^4	1.56×10^{-3}	1.9
DES	$0.2 < z < 1.3$	5×10^3	1.85×10^{-3}	8.4
PanSTARRS	$0 < z < 1.2$	3×10^4	1.72×10^{-3}	3.7
LSST	$0.3 < z < 3.6$	3×10^4	2.77×10^{-3}	0.7

Dalal et al. 2008, Carbone et al. 2008, McDonald 2008, Afshordi & Tolley 2008

The large-scale structure



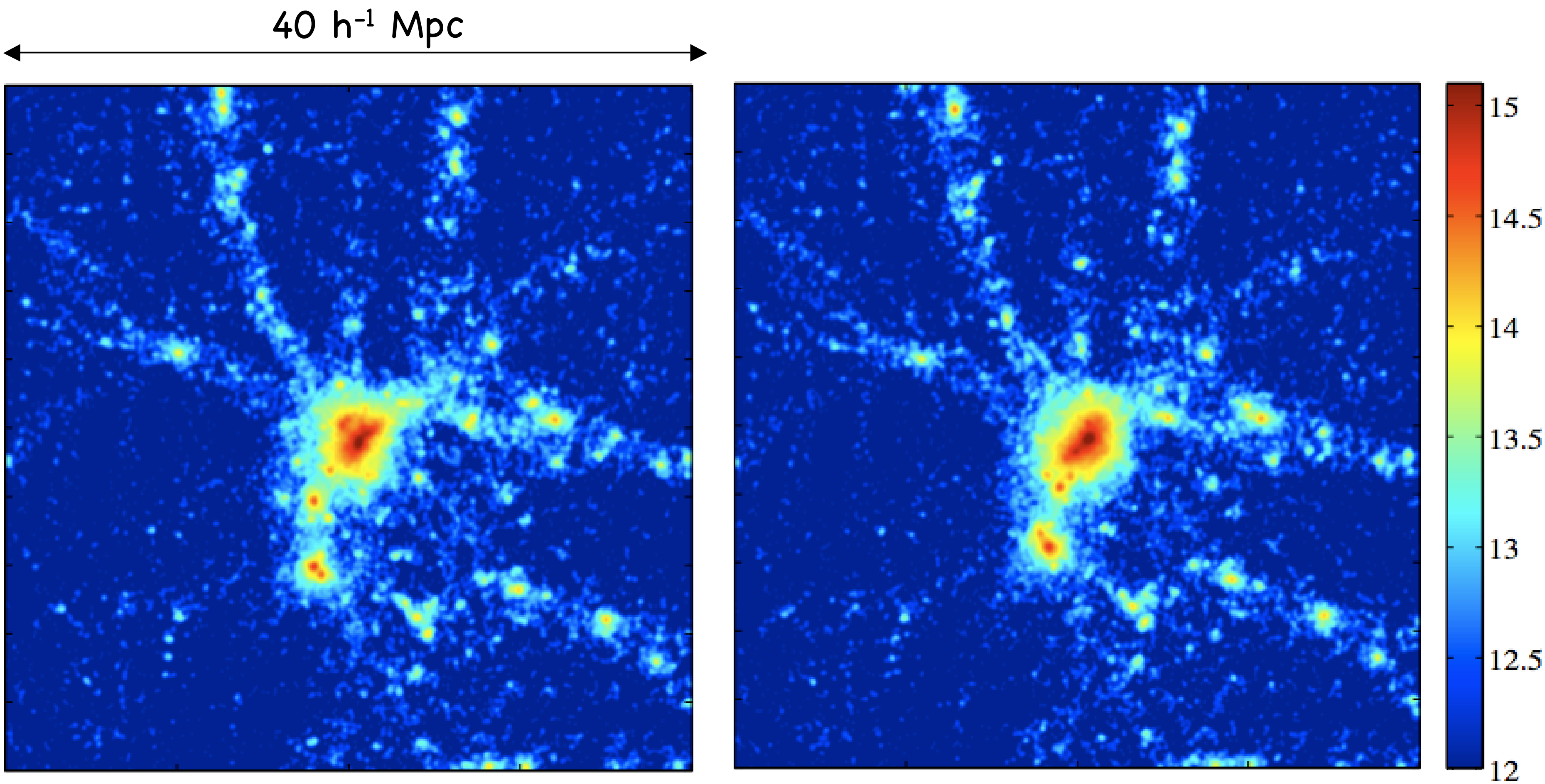
Zooming in



f_{NL} = 0

f_{NL} = 750

A massive galaxy cluster



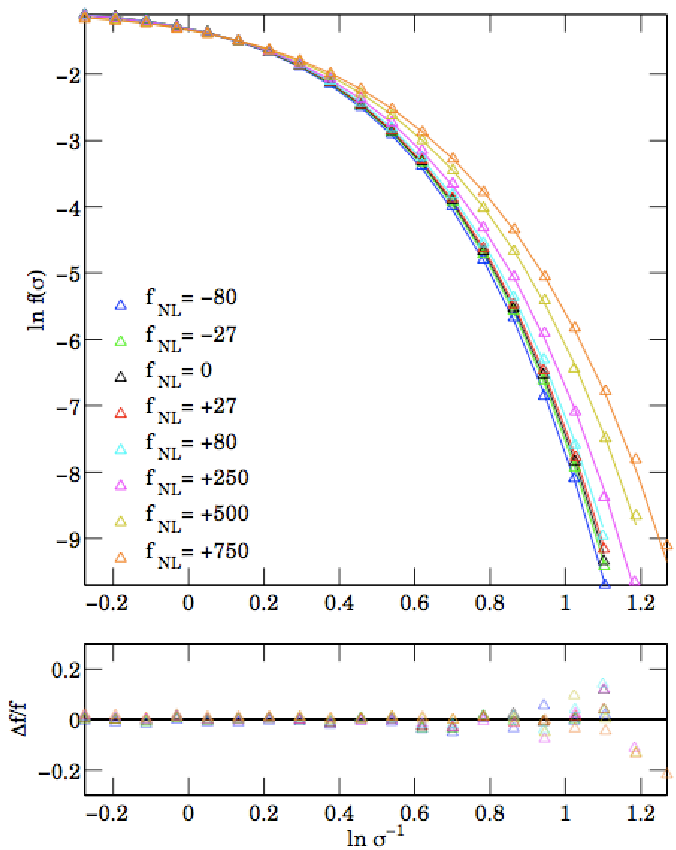
f_{NL} = 0 f_{NL} = 750

f_{NL} from rare events

- Rare events have the advantage that they maximize deviations from the Gaussian case
- However, they have the obvious disadvantage of being ... rare!
- Early attempts of using the cluster mass function to measure f_{NL} have been inconclusive due to small number statistics but the future looks promising

Mass function

Pillepich, CP & Hahn 2009



- A number of groups (Desjacques et al. 2009, Pillepich et al. 2009, Grossi et al. 2009) measured the mass distribution of galaxy clusters and groups as a function of f_{NL} in numerical simulations (cf. analytic approach by Matarrese et al. 2000, Lo Verde et al. 2007, Maggiore & Riotto 2009, Yam Lam & Sheth 2009)

- The halo mass function extracted from the simulations provides a benchmark for future determinations of f_{NL} with galaxy-cluster surveys

The Wiener filter

Norbert Wiener

1894-1964

The grand challenge

- d=s+n
- s is a Gaussian variable with zero mean and known variance σ^2_s
- n is a Gaussian variable with zero mean and known variance σ^2_n
- You measure d, what is your best Bayesian point estimate for s?
- What is the posterior PDF for s?

Solution

- We know that noise is Gaussian, then

$$L(s) = P(d \mid s) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left[-\frac{(d-s)^2}{2\sigma_n^2}\right]$$

- We know that the signal is Gaussian, then we can use the prior

$$P(s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left[-\frac{s^2}{2\sigma_s^2}\right]$$

- We finally write down Bayes theorem and compute the mean value of s over the posterior, finding

$$\langle s \rangle = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2}\, d$$

Wiener filter

$d = Hs + n$

Power spectrum of the signal is known

Power spectrum of the noise is known

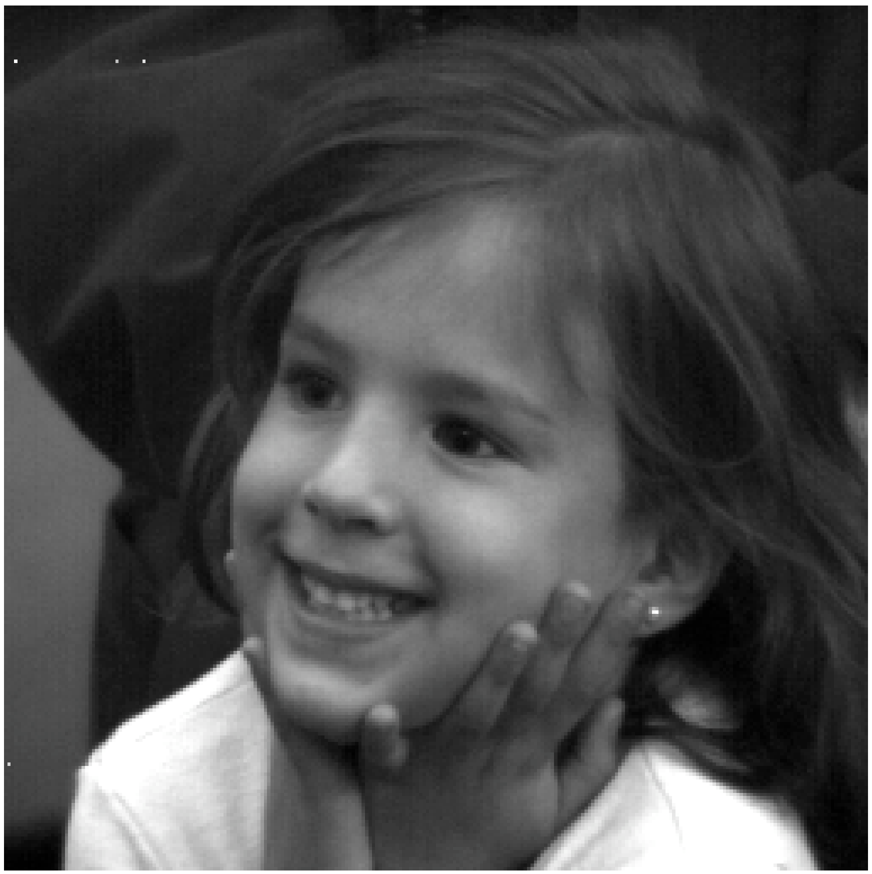
Point-spread function is known

Then, the optimal linear filter (*minimum mean-square error*) *is*

$$\hat{s}(\vec{k}) = G(\vec{k})d(\vec{k})$$

$$G(\vec{k}) = \frac{H^*(\vec{k})P_s(\vec{k})}{|H|^2(\vec{k})P_s(\vec{k}) + P_n(\vec{k})}$$

Example: Lenina

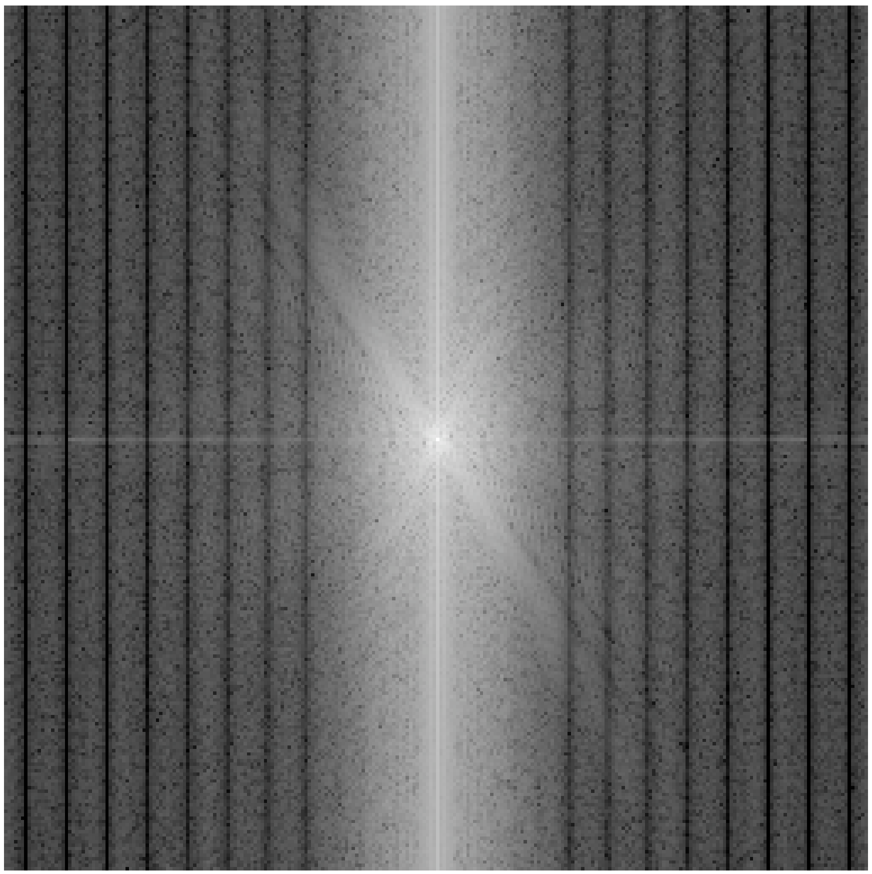


Original

Blurred

1-D Gaussian smoothing along x axis

Example

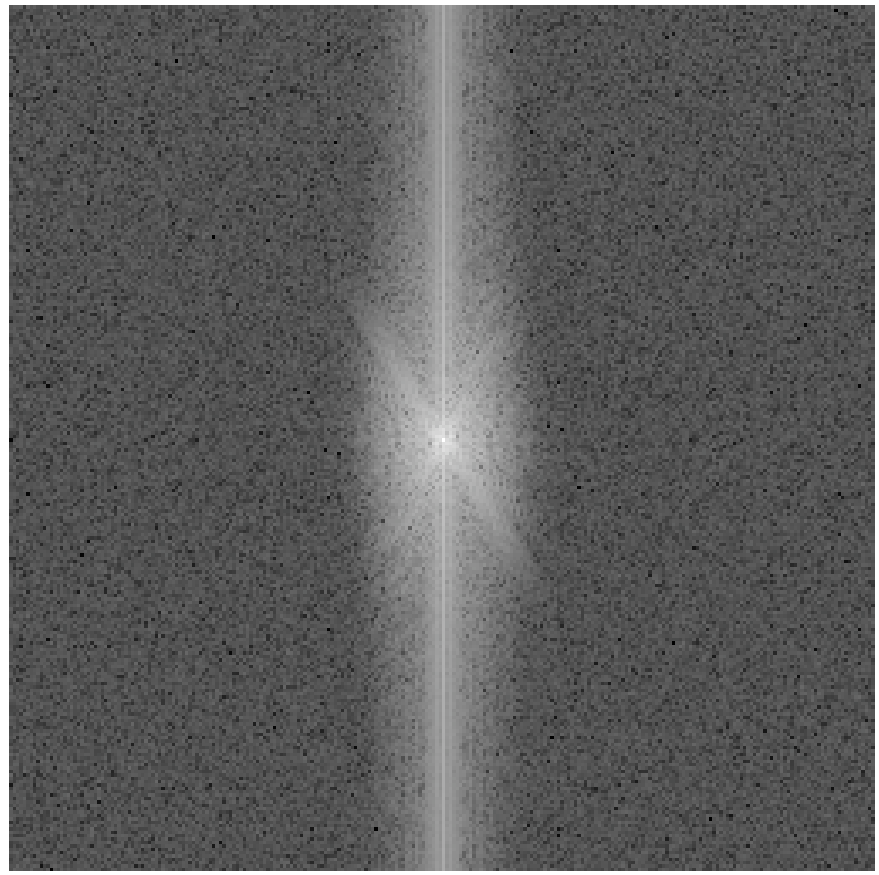


FT of blurred image
Where is the distortion?

Reconstructed
Inverse filter

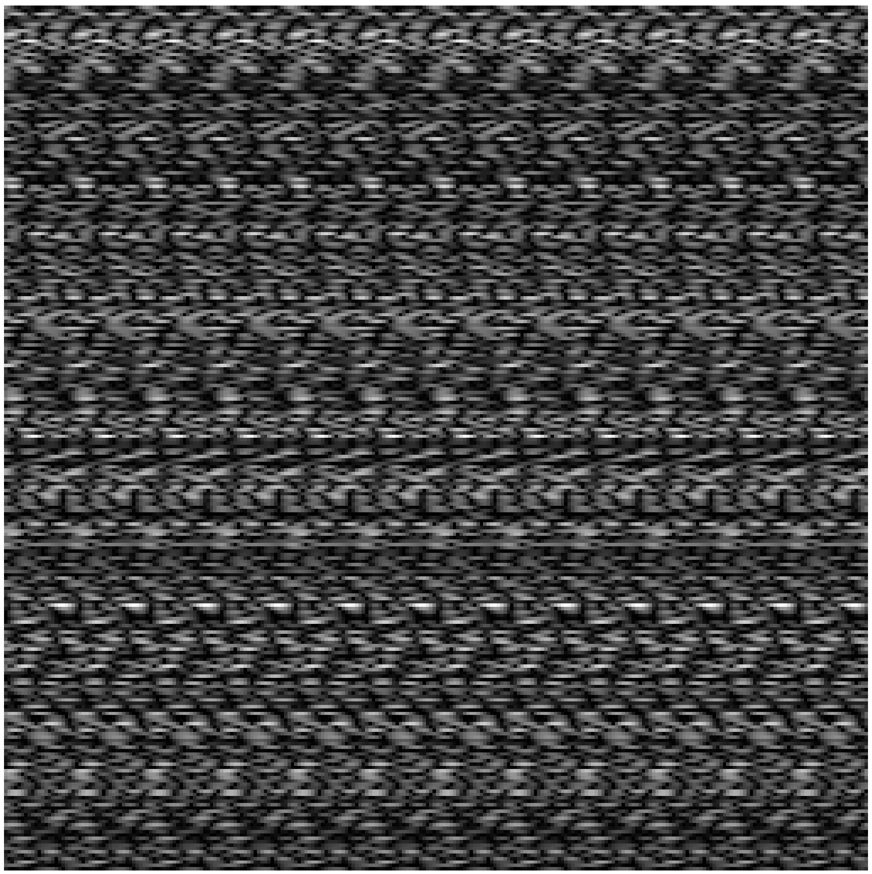
Example

Blurred + noise added



FT

Example

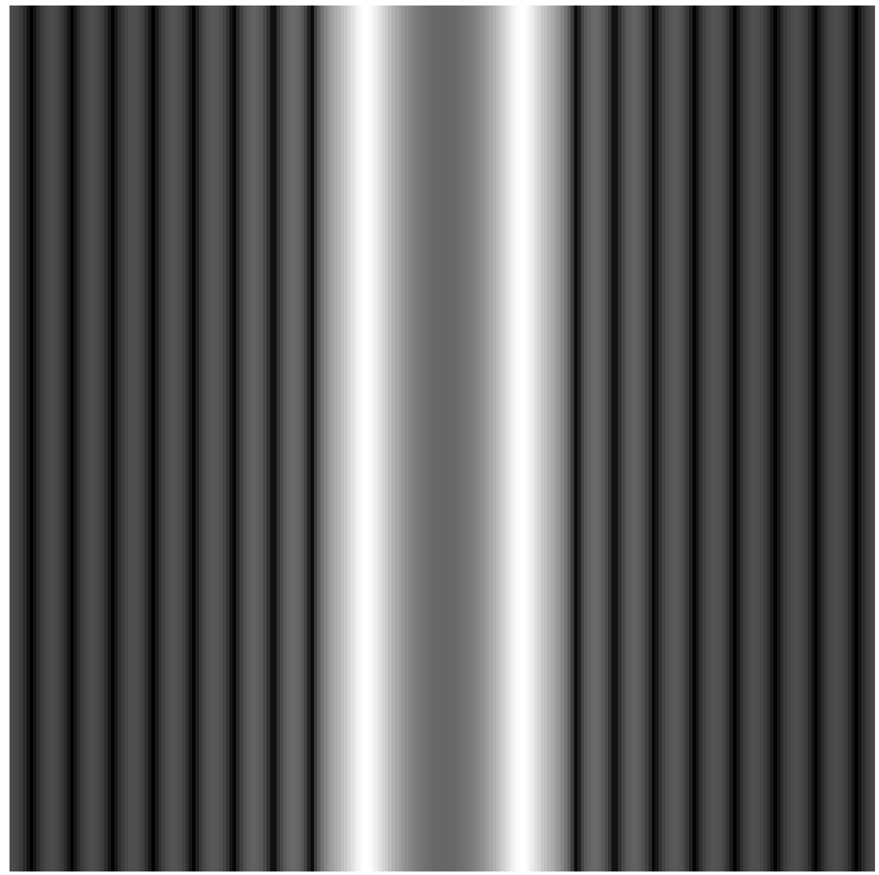


Reconstructed
Inverse filter

Reconstructed
Pseudo-inverse filter

Example

Reconstructed
Wiener filter



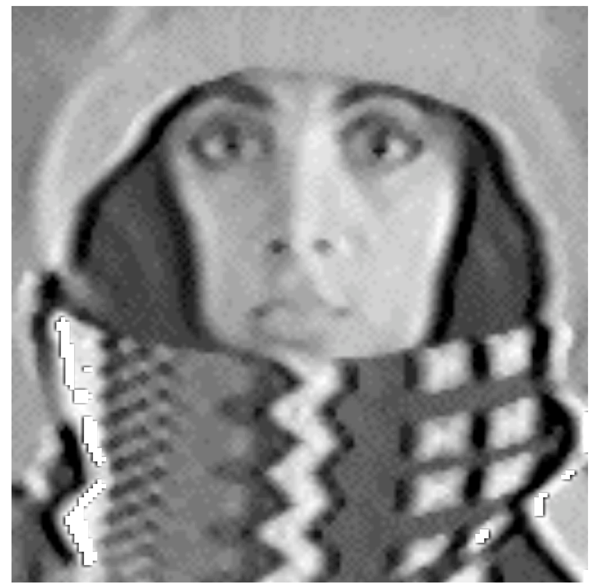
FT of Wiener filter

Wiener filtering

Wiener filtering

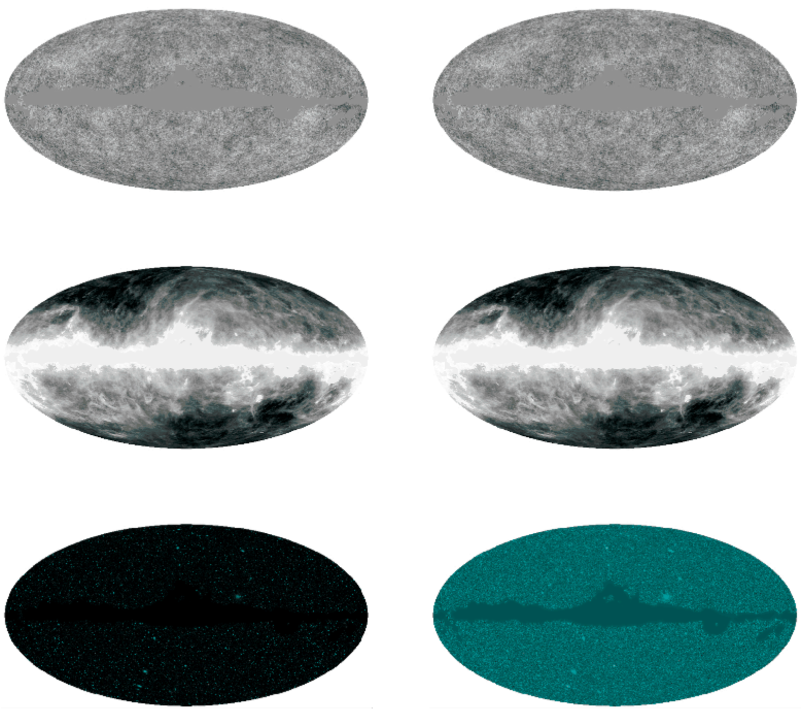
Noisy

Wiener filtered



Median (3x3) filtered

Wiener filter in astronomy



Starck et al. 2006

- Astronomical images: clean up
- CMB: each pixel contains a measurement of the beam-smoothed temperature plus noise
- LSS: to reconstruct the cosmic density field from the galaxy distribution

Lauro is desperate! He spent all the conference money on the social dinner. He now has only one possibility to keep his job…

The game show has new rules: 12 questions, you can call a friend 3 times but it must be for three consecutive questions, you can withdraw any time but if you answer wrong you loose everything.

Lauro, however, does not know much beyond the SZ effect and so he will answer at random.

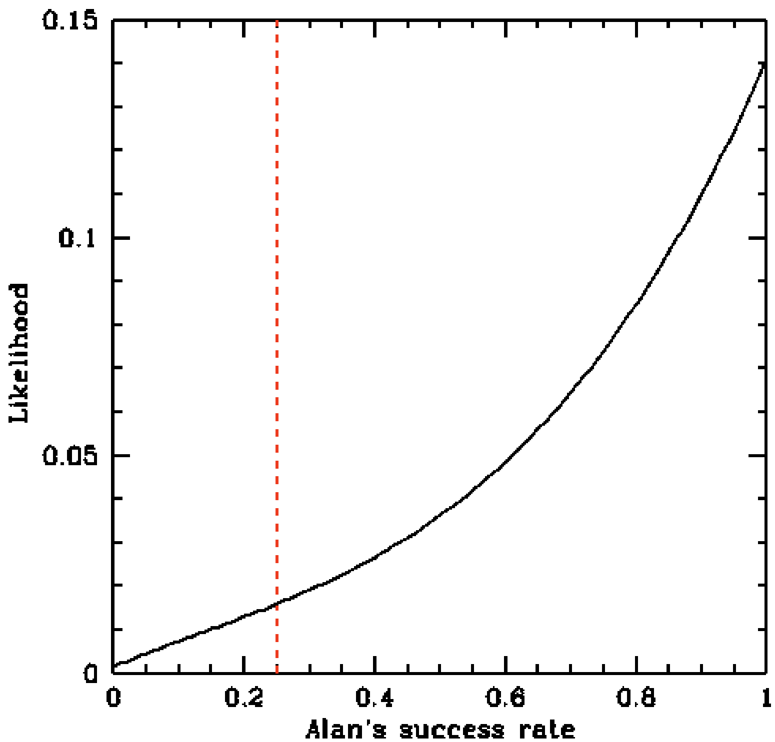
For this reason he asked for help to Alan and promised to answer what Alan suggests for the 3 questions.
Knowing that Lauro lost after giving a wrong answer at the 6th question, what can you say about Alan's success rate? (the population value…not the sample one)

Solution

- AAALLL
- LAAALL
- LLAAAL
- LLLAAA
- LLLLAA
- LLLLLA
- LLLLLL
- All possibilities are equally likely
- Likelihood:

$$L(p) = P(data \mid p) = 3\left[\left(\frac{1}{4}\right)^2 \frac{3}{4} p^3 + \left(\frac{1}{4}\right)^3 p^2(1-p) + \left(\frac{1}{4}\right)^4 p(1-p) + \left(\frac{1}{4}\right)^5 (1-p) + \left(\frac{1}{4}\right)^5 \frac{3}{4}\right]$$

Solution



- Maximum likelihood: p=1
- Assume flat prior (unless you already know something about Alan's success rate)
- Mean posterior: p≈0.72
- Median posterior: p≈0.77
- Note that the likelihood is not peaked, it is then difficult to give a credibility interval. Best, 1-tailed CI: p>0.26 at 95% significance
- Prob(p>0.25)≈0.953
- The likelihood function (multiplied by a prior) contains all the information