Surrogates

MPE Bayes Forum

Garching, December 2012

Christoph Räth

Max-Planck-Institut für extraterrestrische Physik, Garching, Germany

Motivation

•To make known to Bayesians some key concepts of nonlinear data analysis (NLDA)

•To start another attempt to bring together the best from ‚the two worlds‘

‚The two worlds‘:

NLDA:
„The model is the data.“
(C. Grebogy)

Bayes:
"You always put prejudice in it. That's called the Bayesian method.“
(Dick Bond to George Efstathiou, Paris, Planck Meeting, 27.9.12)

I. Tools: Some Higher Order Statistics

Minkowski Functionals

two-dimensional image data/CMB data => three Minkowski functionals (MF):

Area : $M_0(\nu) = \int_{R(\nu)} dS$

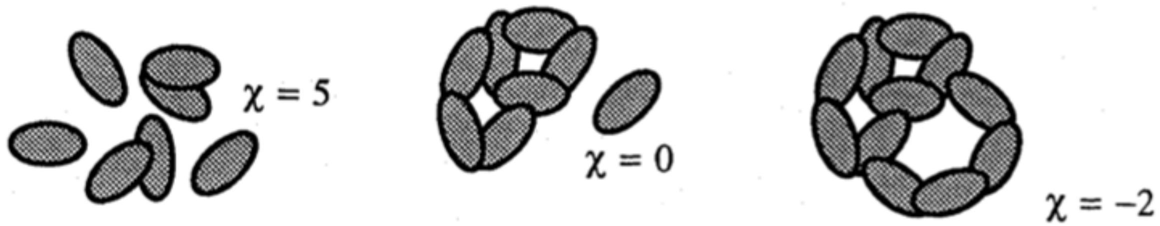
Circumference: $M_1(\nu) = \int_{\partial R(\nu)} dl$

Euler characteristic: $M_2(\nu) = \int_{\partial R(\nu)} \frac{dl}{r} dl$

of an excursion set $R(\nu)$

$M_2 = \chi =$
connected regions -
holes in the regions.

d=2 :



Information (of the sum) of *all* n-point correlation function is contained in the MF

Mecke et al., A & A, 1994
Schmalzing & Gorski, MNRAS, 1998

Scaling indices for spherical data

Idea: Assessing *local* scaling properties:

Consider a point distribution P:

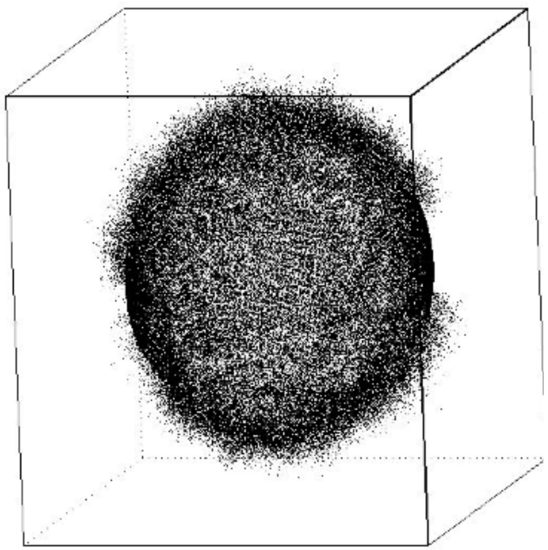
$$P = \{\vec{p}_i\}, i = 1, \ldots, N_{points},$$
$$\vec{p}_i = \{x_i, y_i, z_i\}$$

Local cumulative weighted density:

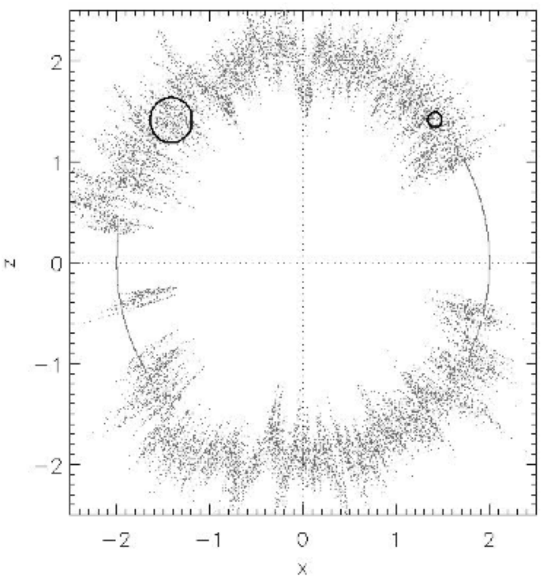
$$\rho(\vec{p}_i) = \sum_{j=1}^{N} e^{-\left(\frac{d_{ij}}{r}\right)^n}, d_{ij} = \left\|\vec{p}_i - \vec{p}_j\right\|$$

Scaling Indices:

$$\alpha(\vec{p}_i) \equiv \frac{\partial \log(\rho(\vec{p}_i))}{\partial \log(r)}$$



3D representation
of WMAP data



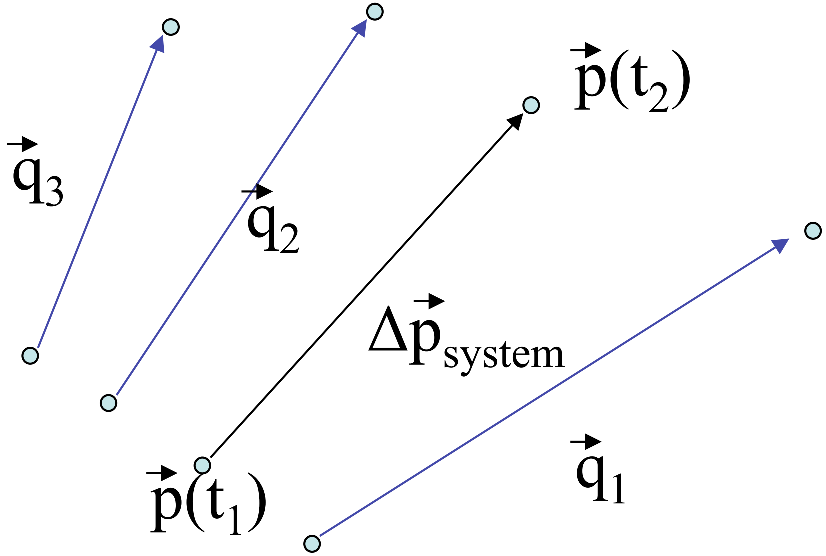
x-z-projection for
all points with |y|<0.1

$$\boxed{=> \alpha(\vec{p}_i) = \frac{\sum_{j=1}^{N} n \cdot \left(\frac{d_{ij}}{r}\right)^n \cdot e^{-\left(\frac{d_{ij}}{r}\right)^n}}{\sum_{j=1}^{N} e^{-\left(\frac{d_{ij}}{r}\right)^n}}}$$

See e.g. for a review:
G. Rossmanith et al., Adv. in Astron., 2011

Non-linear prediction error (NLPE)

Predicted vs. true flow in artificial phase space constructed with delay coordinates:



Global NLPE :

$$\Delta\vec{q}_{pred} = \frac{1}{N_p}\sum_{i=1}^{N_p}\vec{q}_i \qquad NLPE(i) = \left\|\Delta\vec{p}_{system}(i) - \Delta\vec{q}_{pred}(i)\right\| \qquad NLPE = \frac{1}{N}\sum_{i=1}^{N} NLPE(i)$$

$$\boxed{\psi(d,\tau,T,N) = \frac{1}{(M-T-(d-1)\tau)}\left(\sum_{n=(d-1)\tau}^{M-1-T}[\vec{x}_{n+T} - F(\vec{x}_n)]^2\right)^{1/2}}$$

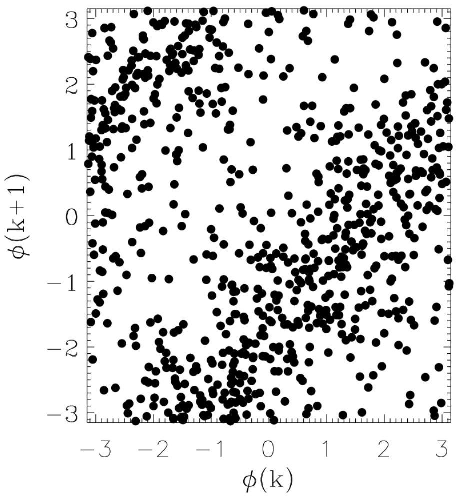
See e.g.: G. Sugihara and R. M. May, Nat., 344, 734(1990)

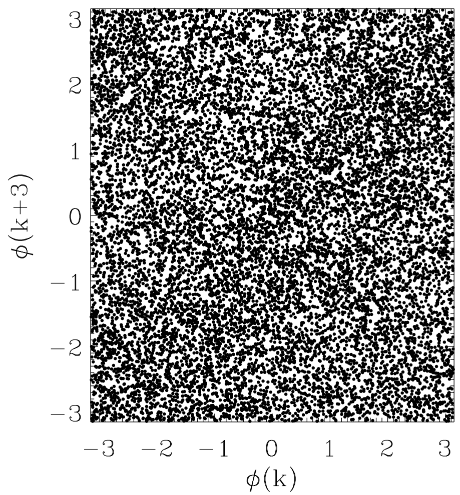
Phase Maps

Consider the Fourier Transform $FT(I(x)) = A(k)e^{i\varphi(k)}$ of a time series $I(x)$:

A phase map is a two-dimensional set of points $G = \{\varphi(k),\varphi(k+\Delta)\}$ where $\varphi(k)$ is the phase of the kth mode of the Fourier transform and Δ a mode delay.

Examples:





Note: If the phases are uniformly distributed and independent from each other, the phase maps are a random 2d distribution of points.

See e.g.: L.-Y. Chiang, et al., MNRAS 337, 488 (2002)

II. Surrogates

Surrogates

Definition:

‚Surrogates are data sets which have some properties with a given data set in common while all other properties are subject to randomisation‘

One of the key concepts of nonlinear data analysis

Background:

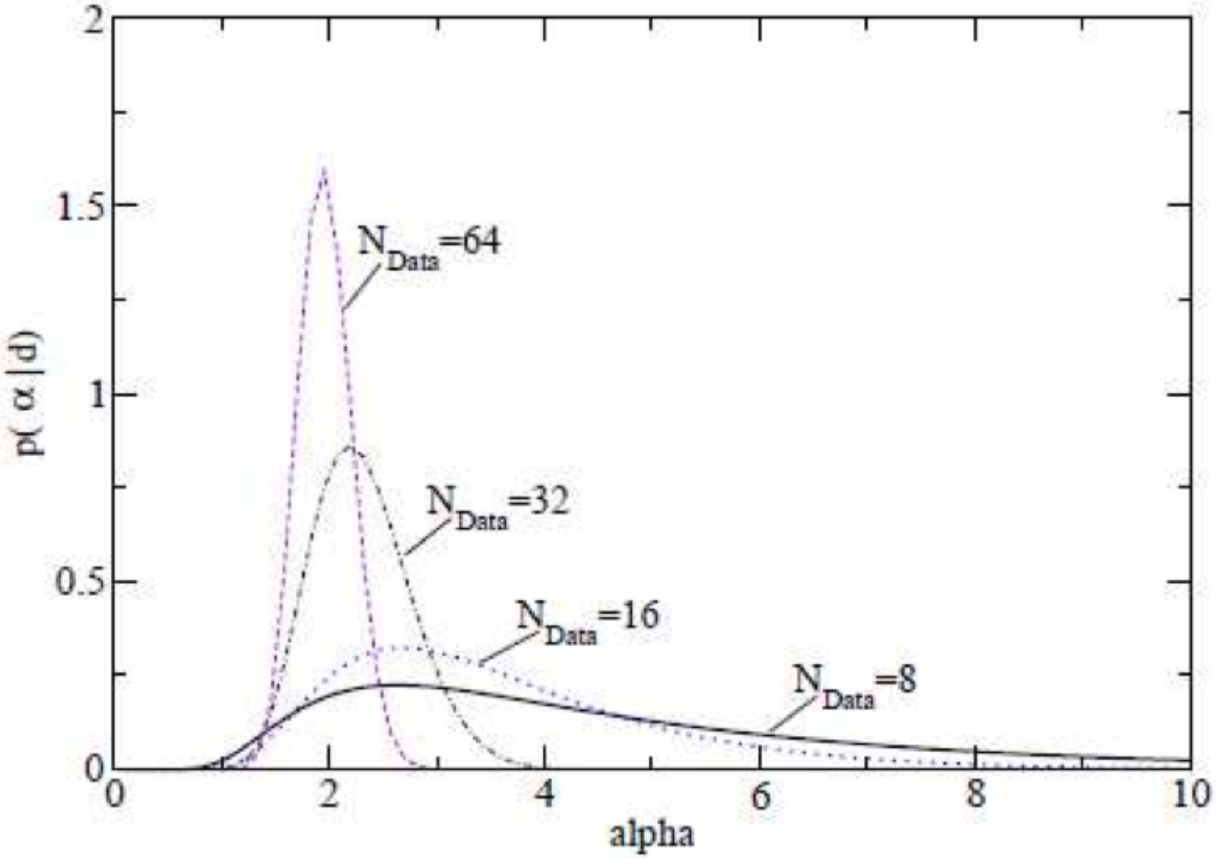
Resampling techniques: Jackknife, Bootstrapping, etc.

Most common surrogates:

Preserving linear properties, i.e. power spectrum, randomising all Higher Order Correlations <=> Fourier phases are random and correlation-free

Volker Dose's Talk:

alpha distributions, surrogate data



Scheme:

• A priori definition of a null hypothesis

• Generation of surrogates consistent with null hypothesis

• Computation of discrimination statistics being sensitive to the complement of the null hypothesis

• Comparison of the outcome of the discrimination statistics for original data and surrogates

•Accepting or rejecting null hypothesis

Note:
citations: 1630 (June 2012)
citations: 1740 (December 2012)
There‘s more than Bayes method...

We describe a statistical approach for identifying nonlinearity in time series. The method first specifies some linear process as a null hypothesis, then generates surrogate data sets which are consistent with this null hypothesis, and finally computes a discriminating statistic for the original and for each of the surrogate data sets. If the value computed for the original data is significantly different than the ensemble of values computed for the surrogate data, then the null hypothesis is rejected and nonlinearity is detected. We discuss various null hypotheses and discriminating statistics. The method is demonstrated for numerical data generated by known chaotic systems, and applied to a number of experimental time series which arise in the measurement of superfluids, brain waves, and sunspots; we evaluate the statistical significance of the evidence for nonlinear structure in each case, and illustrate aspects of the data which this approach identifies.

Probing Linearity / Gaussianity

Data Set

Constrained Randomisation

Surrogate data with the same power spectrum

Calculation of statistical measures M sensitiv to higher order correlations

M derived from e.g.:

- Phase maps
- NLPE
- Bispectrum
- Minkowski-functionals
- Wavelets
- Scaling indices
- Etc……

Statistical comparison in terms of e.g. significances, Confidence levels, etc.

III. Some Algorithms for Generating Surrogates

FT-algorithm

Original data:

$$I_0(x,y) \xrightarrow{\text{Rank ordered remapping onto Gaussian distribution}} I_1(x,y) \xrightarrow{\text{FT}} A_1(k_x,k_y)e^{i\phi_1(k_x,k_y)}$$

Random phases: $\phi(k_x,k_y)$

$$A_1(k_x,k_y)e^{i\phi(k_x,k_y)} \xrightarrow{\text{FT}^{-1}} I_{surro}(x,y)$$

Note: Phases are – by construction – random

Theiler et al., Physica D, 58, 77 (1992)

AAFT-algorithm

Original data:

$$I_0(x,y) \xrightarrow{\text{Rank ordered Remapping on to Gaussian}} I_1(x,y) \xrightarrow{\text{FT}} A_1(k_x,k_y)e^{i\phi_1(k_x,k_y)}$$

Random phases: $\phi\,(k_x,k_y)$

$$I_{surro}(x,y) \xleftarrow{\text{Rank ordered Remapping on to Original data}} \tilde{I}_{surro}(x,y) \xleftarrow{\text{FT}^{-1}} A_1(k_x,k_y)e^{i\phi\,(k_x,k_y)}$$

Note: Power spectrum is whitened by the remapping step.
Effect of remapping on the phases is not considered.

Theiler et al., Physica D, 58, 77 (1992)

IAAFT-algorithm

Original data:

$$\begin{tikzcd}
I_0(x,y) \arrow[r, "\text{FT}"] & A_0(k_x,k_y)e^{i\phi_0(k_x,k_y)} \arrow[r, "\text{FT}^{-1}"] & J(x,y) \arrow[r, "\text{Rank ordered remapping}"] & I_n(x,y) \arrow[d, "\text{FT}"] \\
\phi_{random}(k_x,k_y) \arrow[u] & & A_0(k_x,k_y)e^{i\phi_n(k_x,k_y)} \arrow[u, "\text{FT}^{-1}"] & A_n(k_x,k_y)e^{i\phi_n(k_x,k_y)} \arrow[l]
\end{tikzcd}$$

Replace φ_0 by a set of random phases φ_{random}:

Replacement of Fourier amplitudes

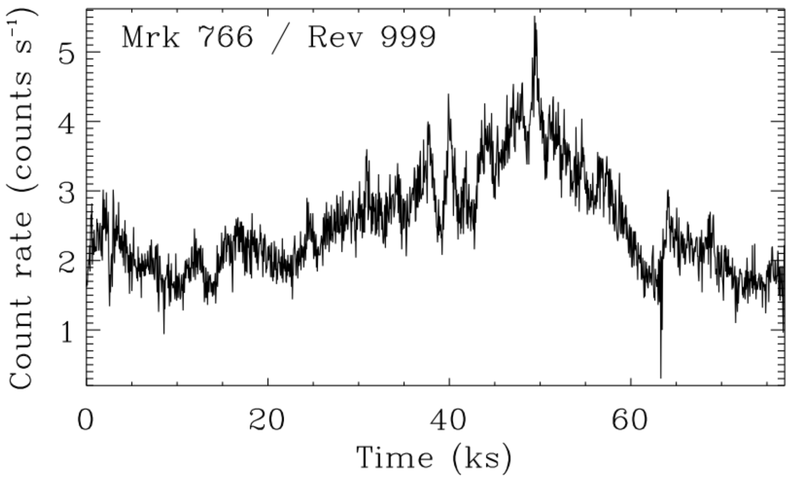
Note: Randomness of the phases is not controlled during iteration.

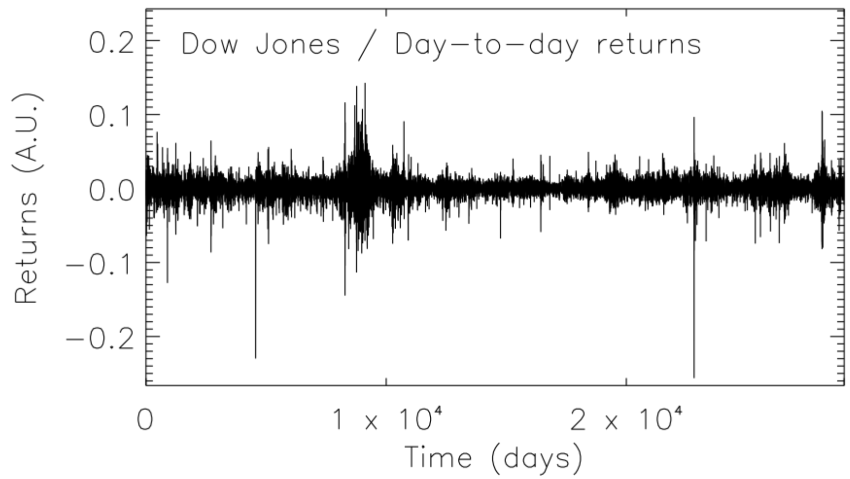
Schreiber & Schmitz, PRL, 77, 635 (1996)

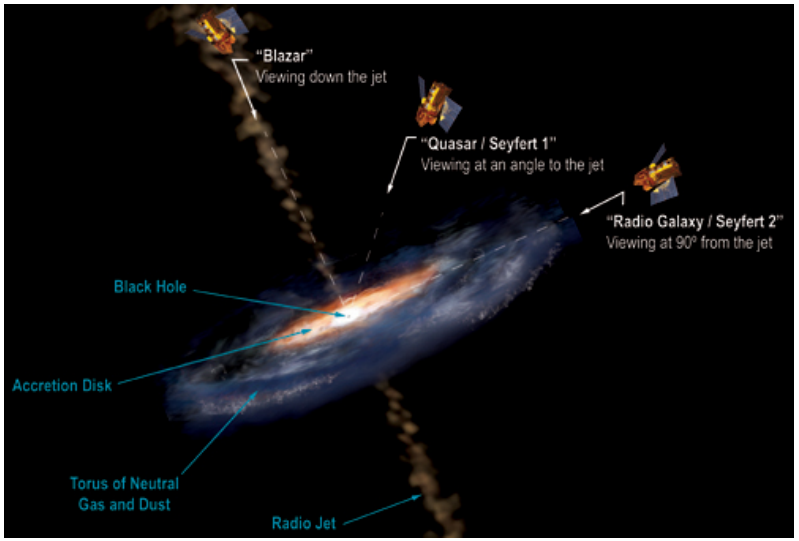
IV. Assessing FT, AAFT and IAAFT

Assessing FT, AAFT and IAAFT

Consider the following scalar time series from two – quite distinct - complex system, namely the X-ray observation of an AGN and a stock market index:

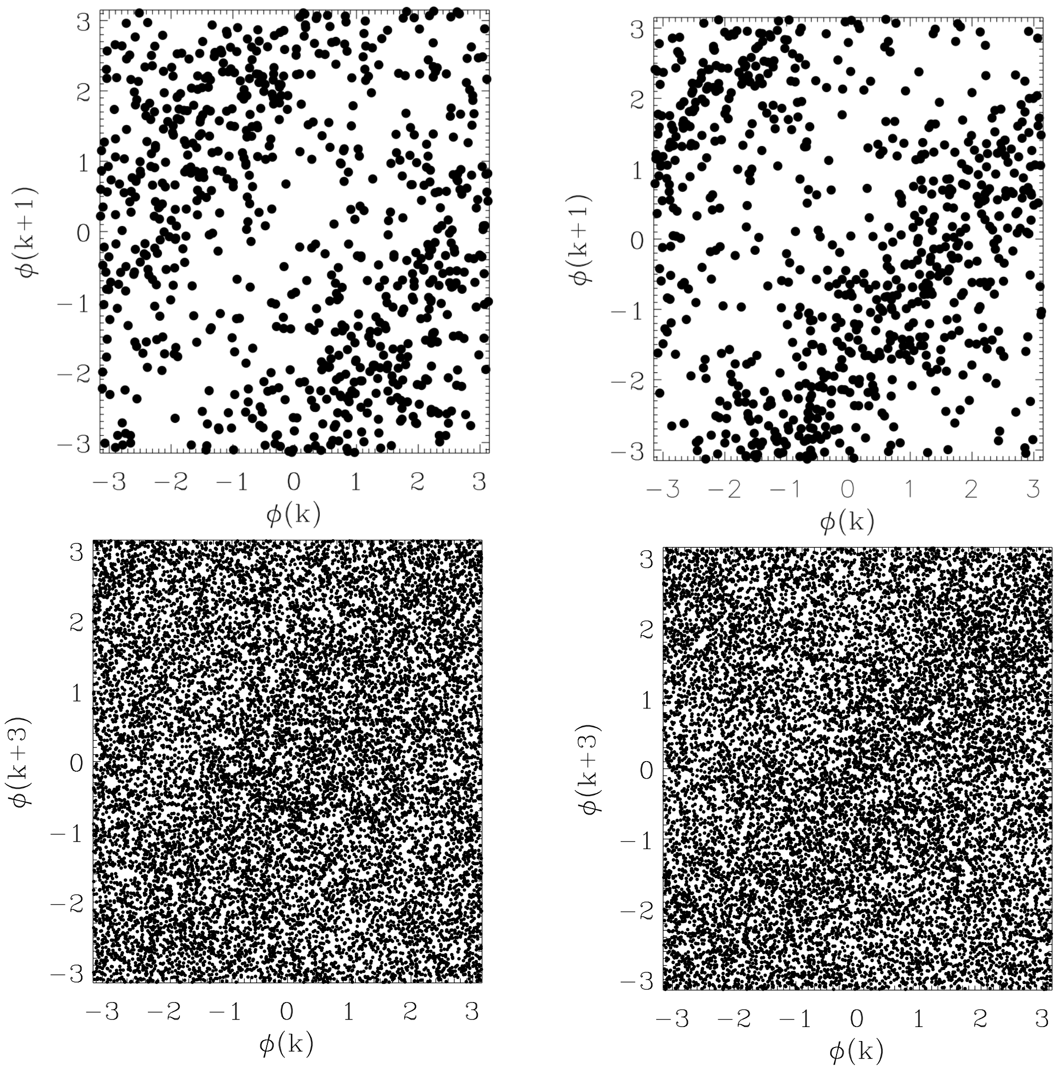


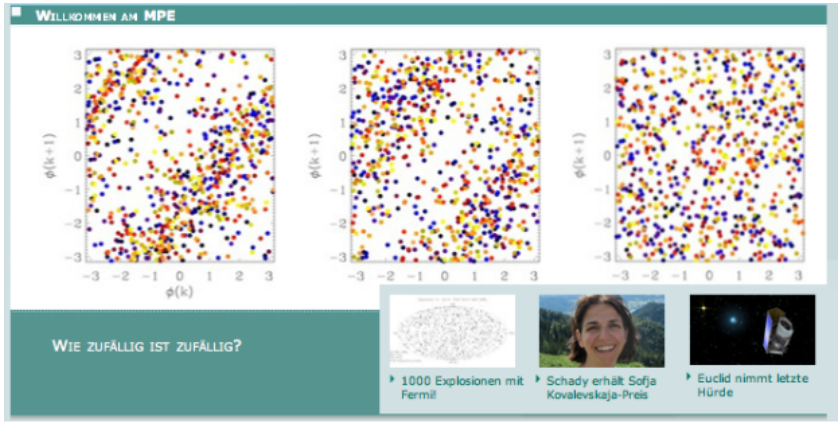




Assessing FT, AAFT and IAAFT

Phase maps for one realization of AAFT and IAAFT surrogates for Mrk and DJ:

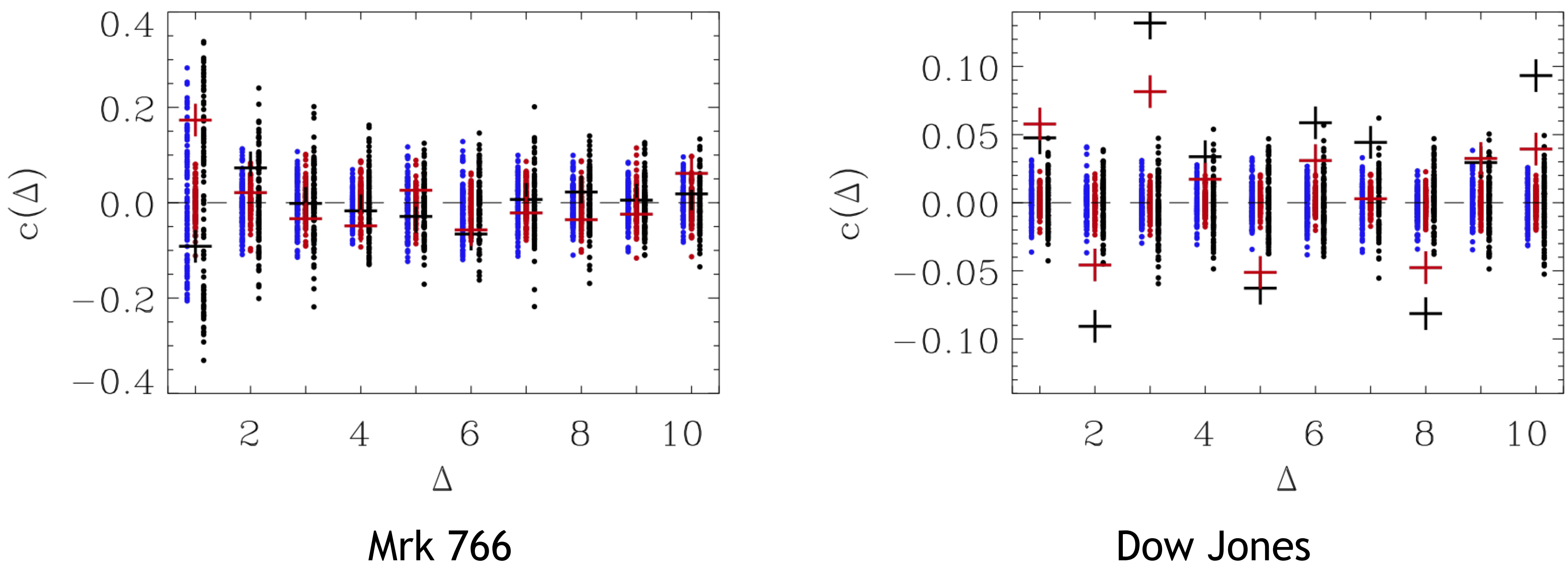




Phase maps show ‚features' !!!

Assessing FT, AAFT and IAAFT

(Linear) Cross-correlations of phases
for AAFT, FT and IAAFT surrogates for Mrk and DJ:

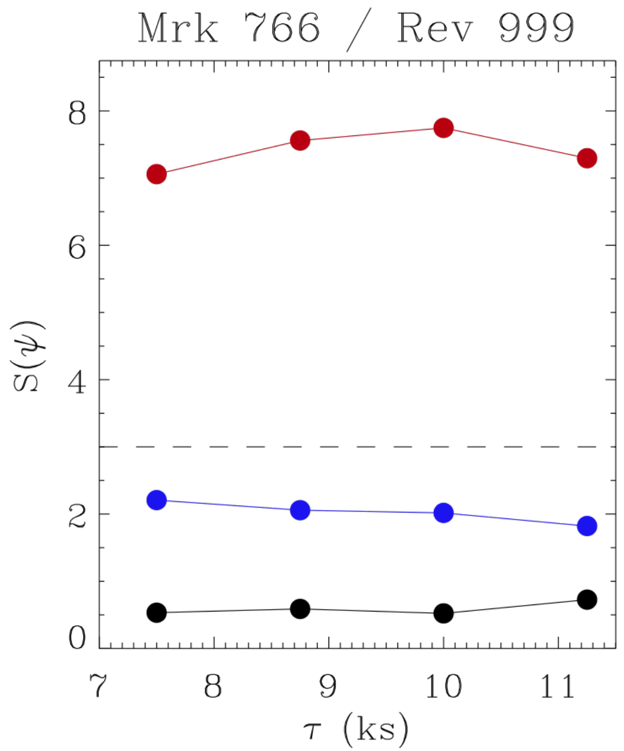


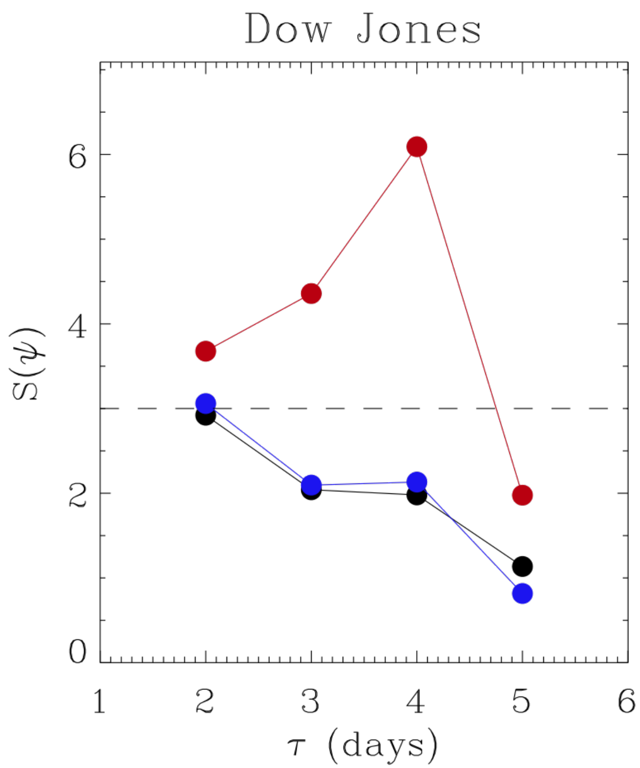
Systematic broadening for AAFT and IAAFT !!!

Räth et al., PRL, 2012

Assessing FT, AAFT and IAAFT

Significances based on the NLPE as derived from AAFT, FT and IAAFT surrogates:



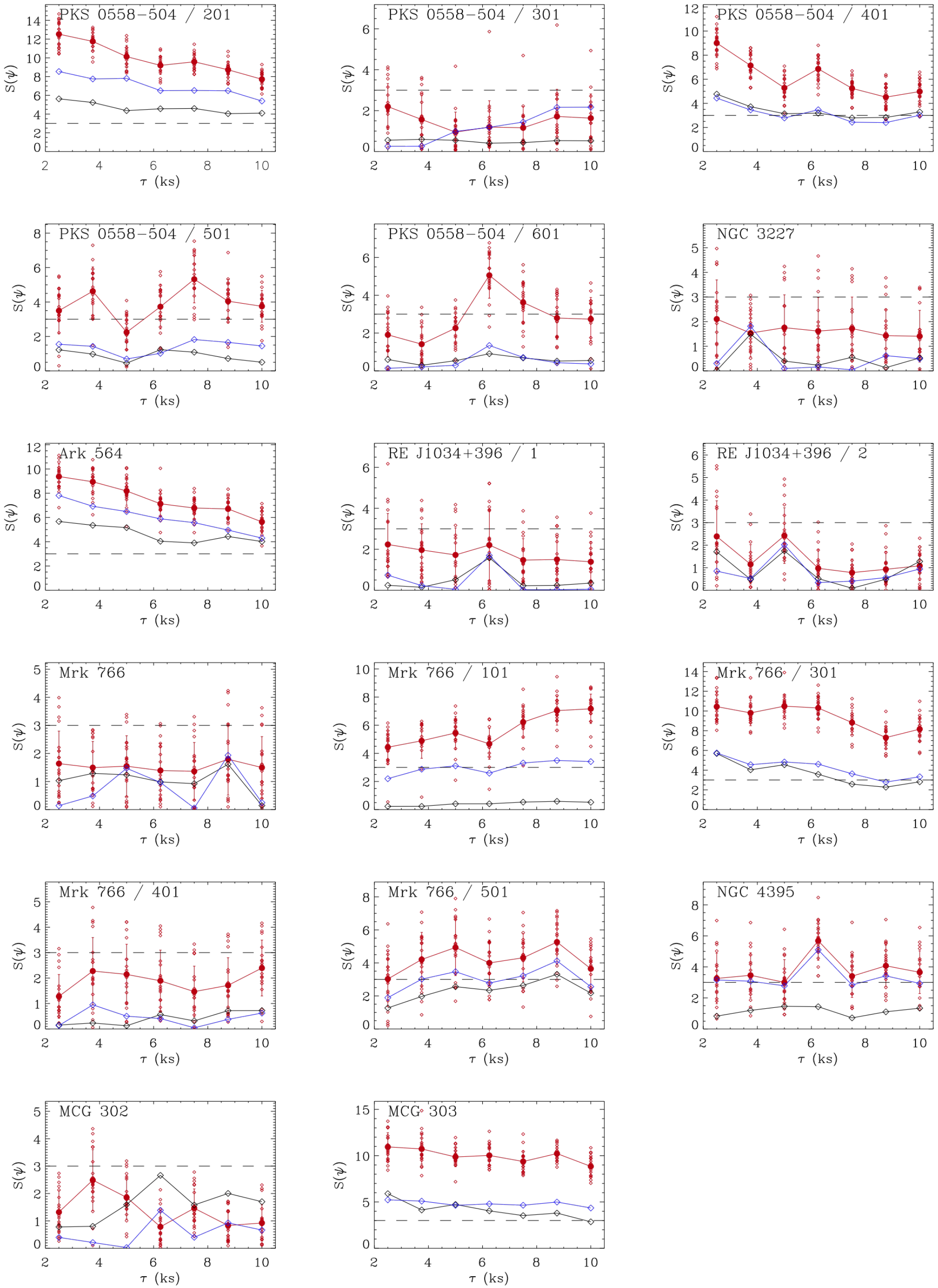


$$S(\psi) = \left| \frac{\psi_{original} - \langle \psi \rangle_{surro}}{\sigma_{\psi_{surro}}} \right|$$

Correlations in the phases propagate into the calculation of NLPE => non-detection of nonlinearities with AAFT and IAAFT !

Räth et al., PRL, 2012

Some more AGN time series:

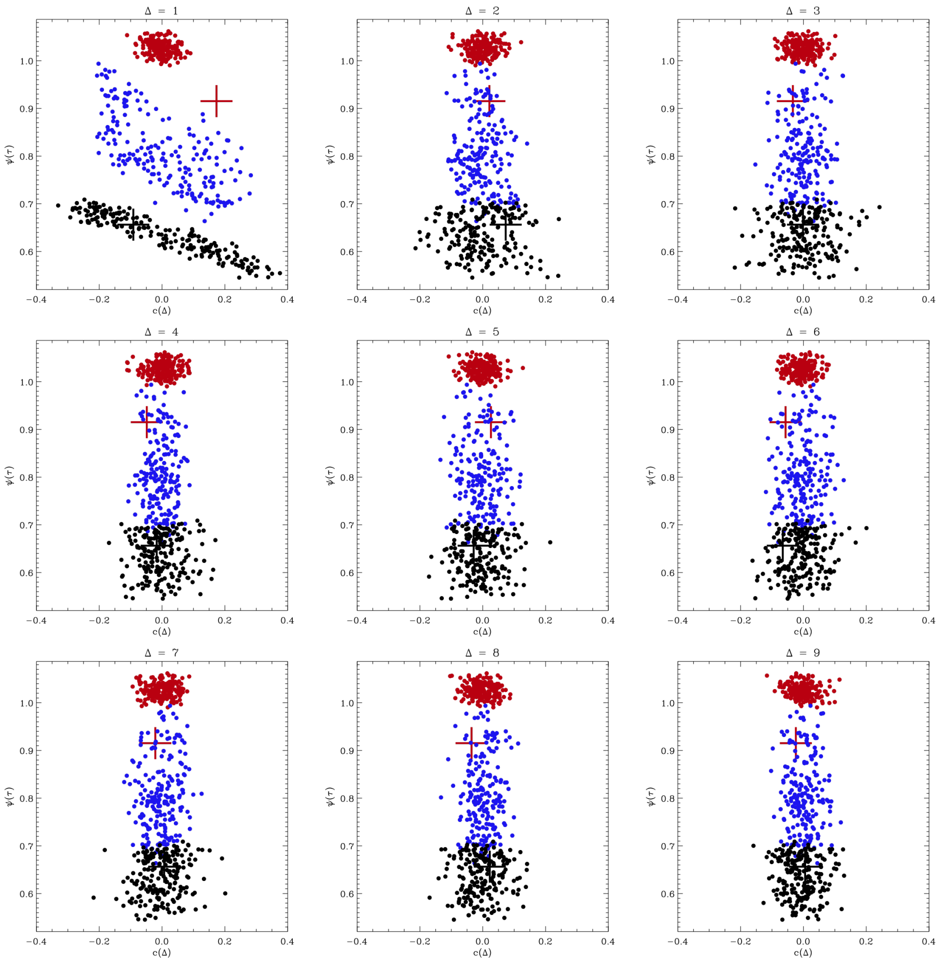


Significant differences for the outcome of surrogate tests depending on the class of surrogates being used.

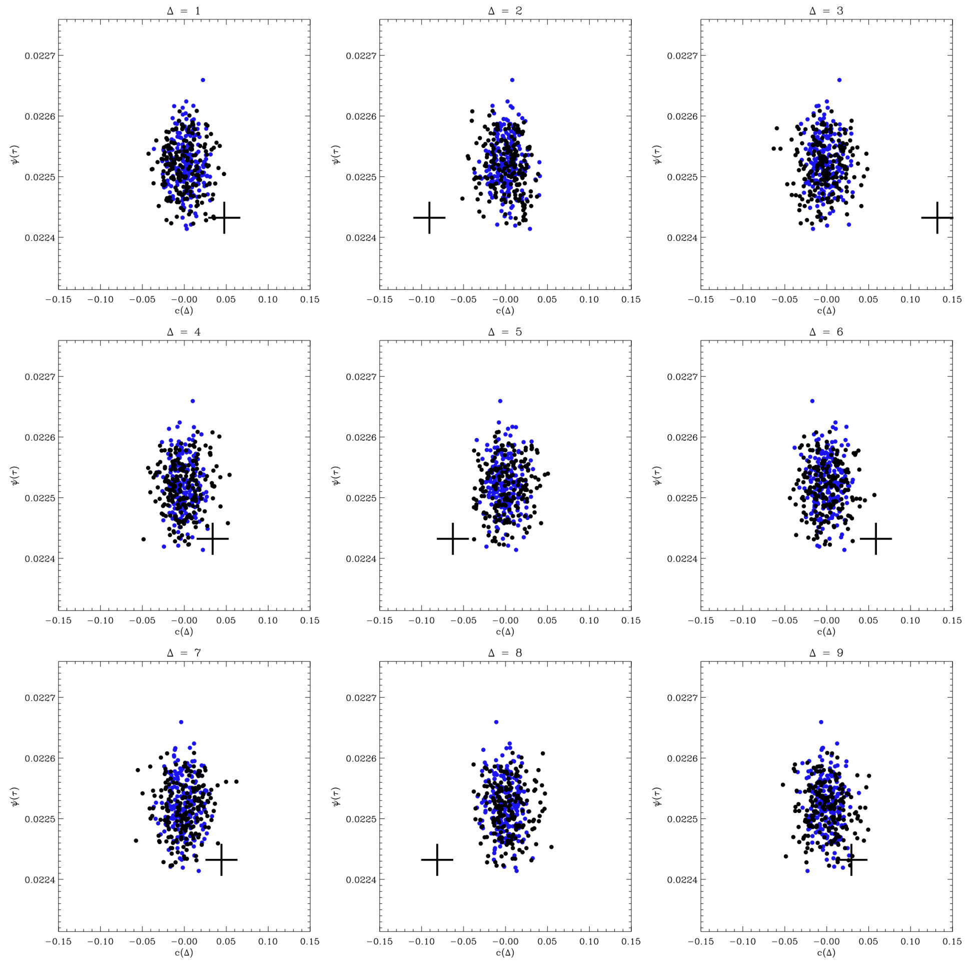
Less significant results for AAFT and IAAFT is a rule.

Phase information vs. HOS

Mrk 766

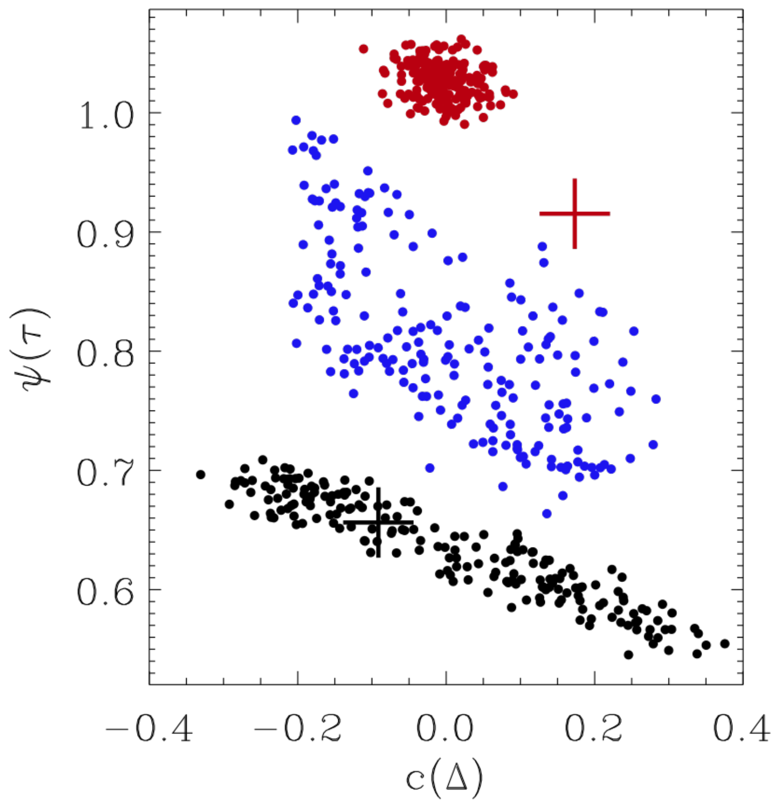


Dow Jones



No correlations found for the DJ,
High correlations detected for Mrk 766 (only) for Δ=1.

Phase information vs. HOS



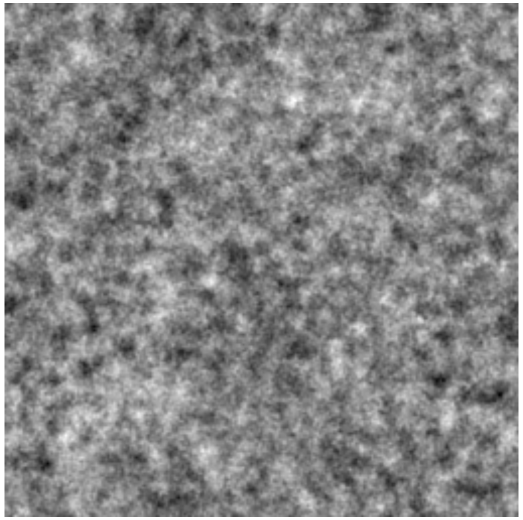
(Surrogate) time series can be constructed such that:

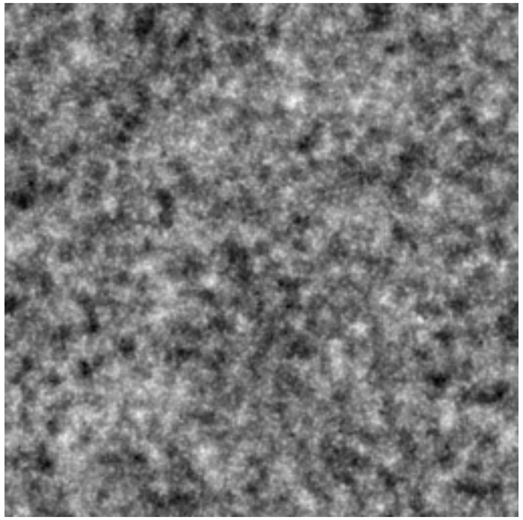
$$\boxed{\psi(\tau) \propto c(\Delta)}$$

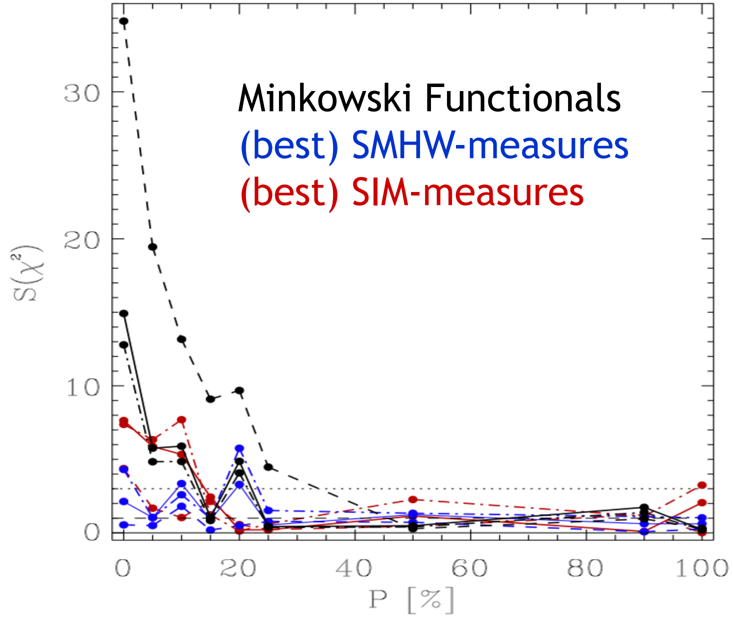
- Wiener-Chintschin-like relation between HOS and phases information detected
- Possibility, to ultimately get more insight into the meaning of Fourier phases for nonlinear data sets.

V. Assessing higher order statistics with Surrogates

Why SIM and MF ?







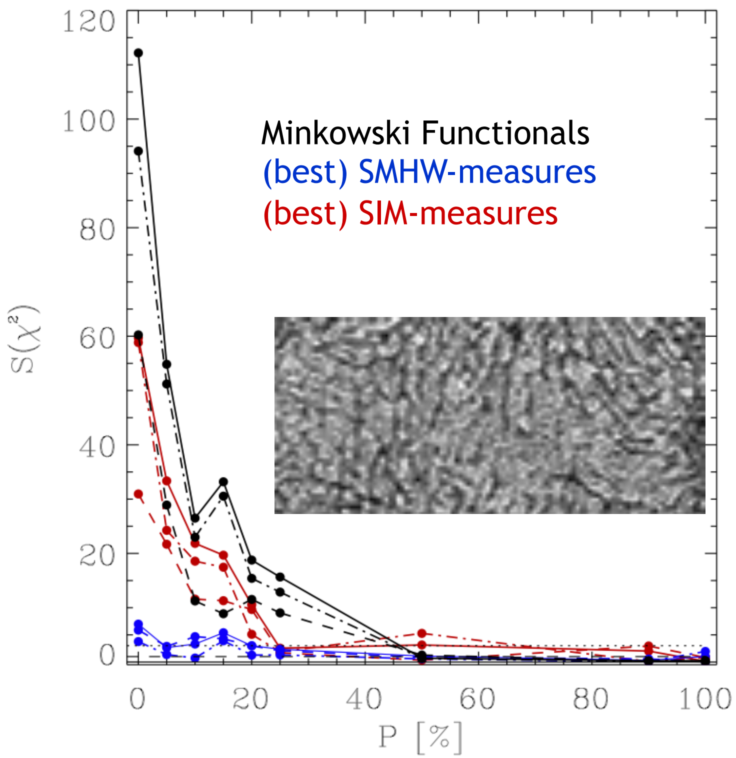
Simulated G and NG flat field, α_3=0.0, α_3=0.3

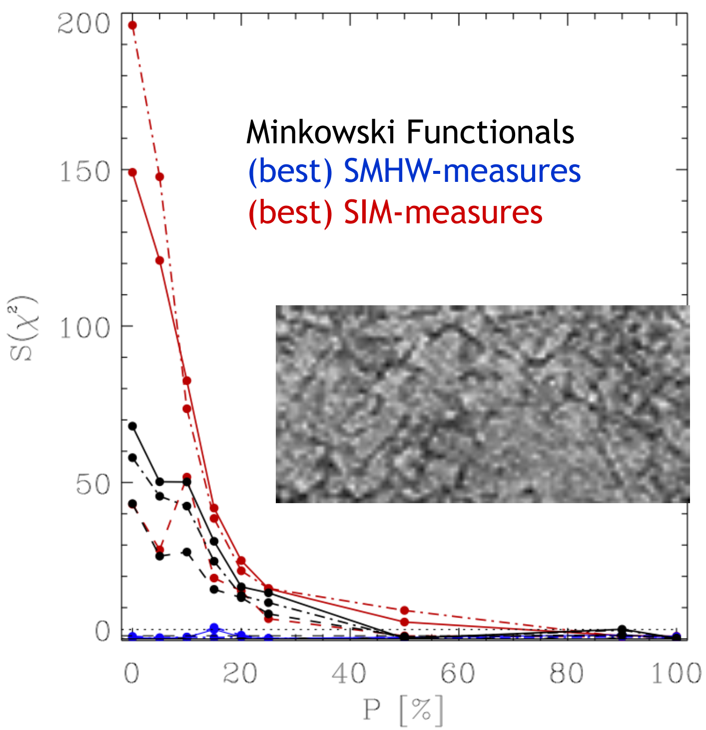
(Rocha et al., MNRAS, 2005)

Highly significant detection of HOCs in the original image with SIM and MF.

=> Assessing the performance of higher order statistics using surrogates.

Why SIM and MF (2nd example) ?





HRMRI images of a healthy (left) and osteoporotic (right) bone

(Müller et al., Osteop. Int., 2006, Räth et al., Proc SPIE, 2009)

Highly significant detection of HOCs in the original image with MF and SIM.

Only poor performance of wavelets.

VI. Surrogates and the CMB

Why (scale-dependent) non-Gaussianity?

- Non-Gaussianity for Inflation is like.....
  ...detection of the Higgs-particle for understanding mass
  ...direct detection of dark matter

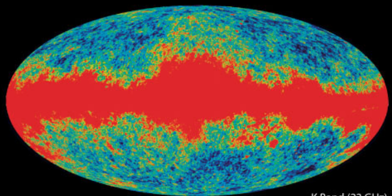
- Single-field inflation: density fluctuations are Gaussian

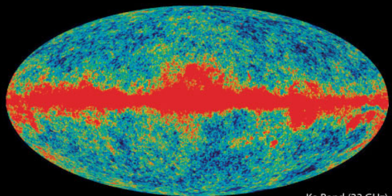
- Some non-standard inflationary models predict *scale-dependent* non-Gaussianities.

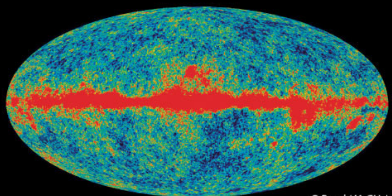
- Once one has found a signature using a model-independent test, one wants to explain its origin.

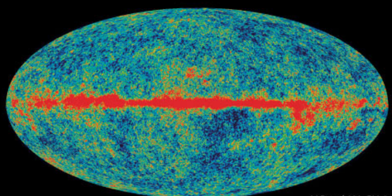
=> Testing whether existing models can account for the detected anomalies

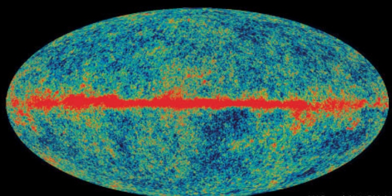
WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP)


K Band (23 GHz)

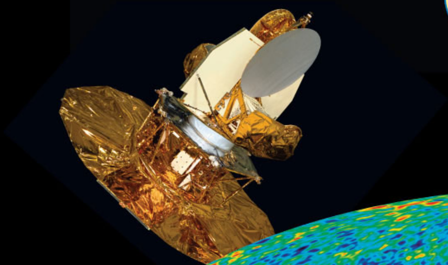

Ka Band (33 GHz)

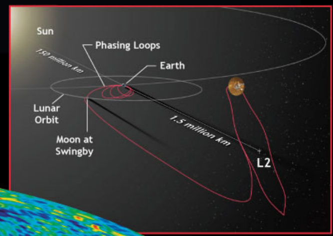

Q Band (41 GHz)

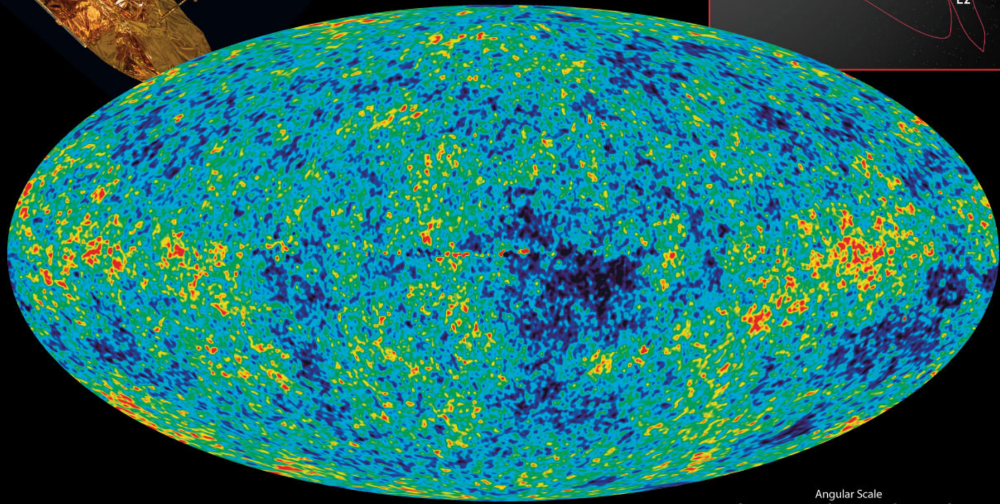

V Band (61 GHz)

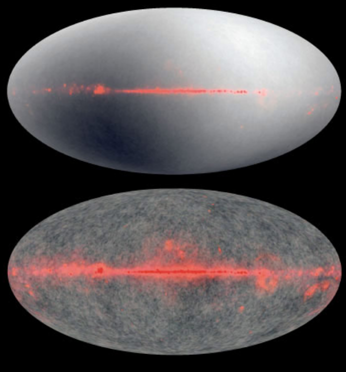

W Band (94 GHz)

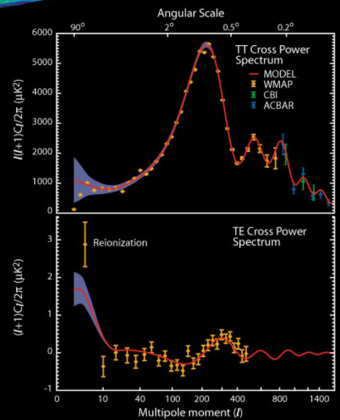
WMAP Full-sky Maps

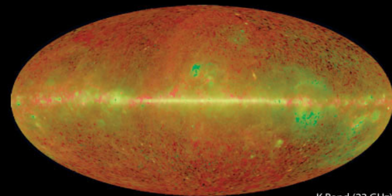


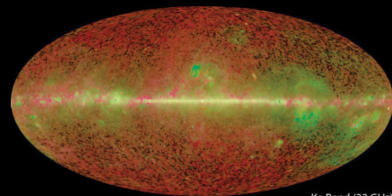


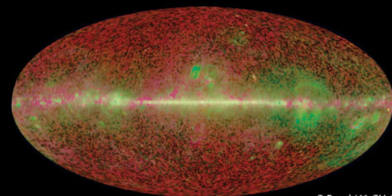


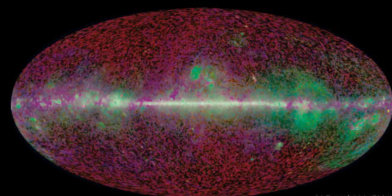

WMAP Foregrounds vs. Cosmic Microwave Background
Red=Q band Green=V band Blue=W band

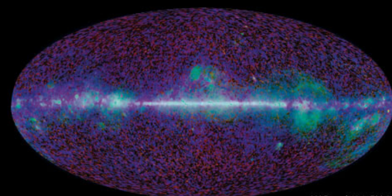



K Band (23 GHz)


Ka Band (33 GHz)


Q Band (41 GHz)


V Band (61 GHz)


W Band (94 GHz)

WMAP Foregrounds
Red=Synchrotron Green=Free-Free Blue=Thermal Dust

Goddard Space Flight Center • Princeton University • University of Chicago • UCLA • University of British Columbia • Brown University
http://map.gsfc.nasa.gov • http://lambda.gsfc.nasa.gov

NW-2002-5-032-GSFC

Generating Surrogates

Fourier Transform of the temperature map:

$$T(n) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(n) \qquad \text{with} \qquad a_{lm} = \int T(n) Y_{lm}^{*} d\Omega_n$$

One can write:

$$a_{lm} = |a_{lm}| e^{\boxed{i\phi_{lm}}} \qquad \text{with} \qquad \boxed{\phi_{lm}} = \arctan\left(\frac{\mathrm{Im}(a_{lm})}{\mathrm{Re}(a_{lm})}\right)$$

Non-Gaussian Field :

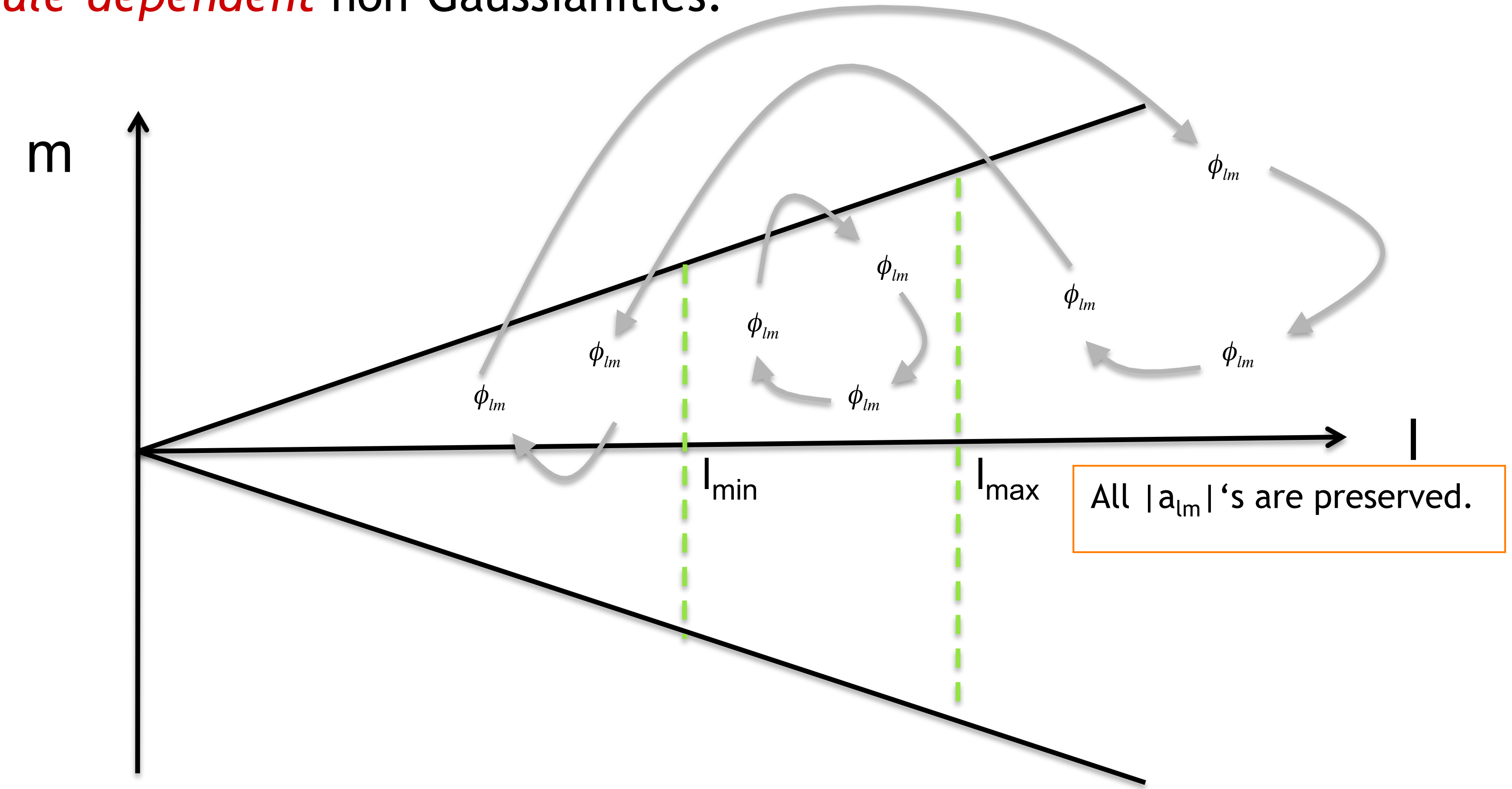
Fourier Phases are correlated and/or *not* uniformly distributed

How to test for possible phase correlations?

Destroy (only) them (by scale-dependent shuffling) and look what happens...

Generating Surrogates

Introducing a two-step shuffling/replacement scheme allows to test for *scale-dependent* non-Gaussianities:

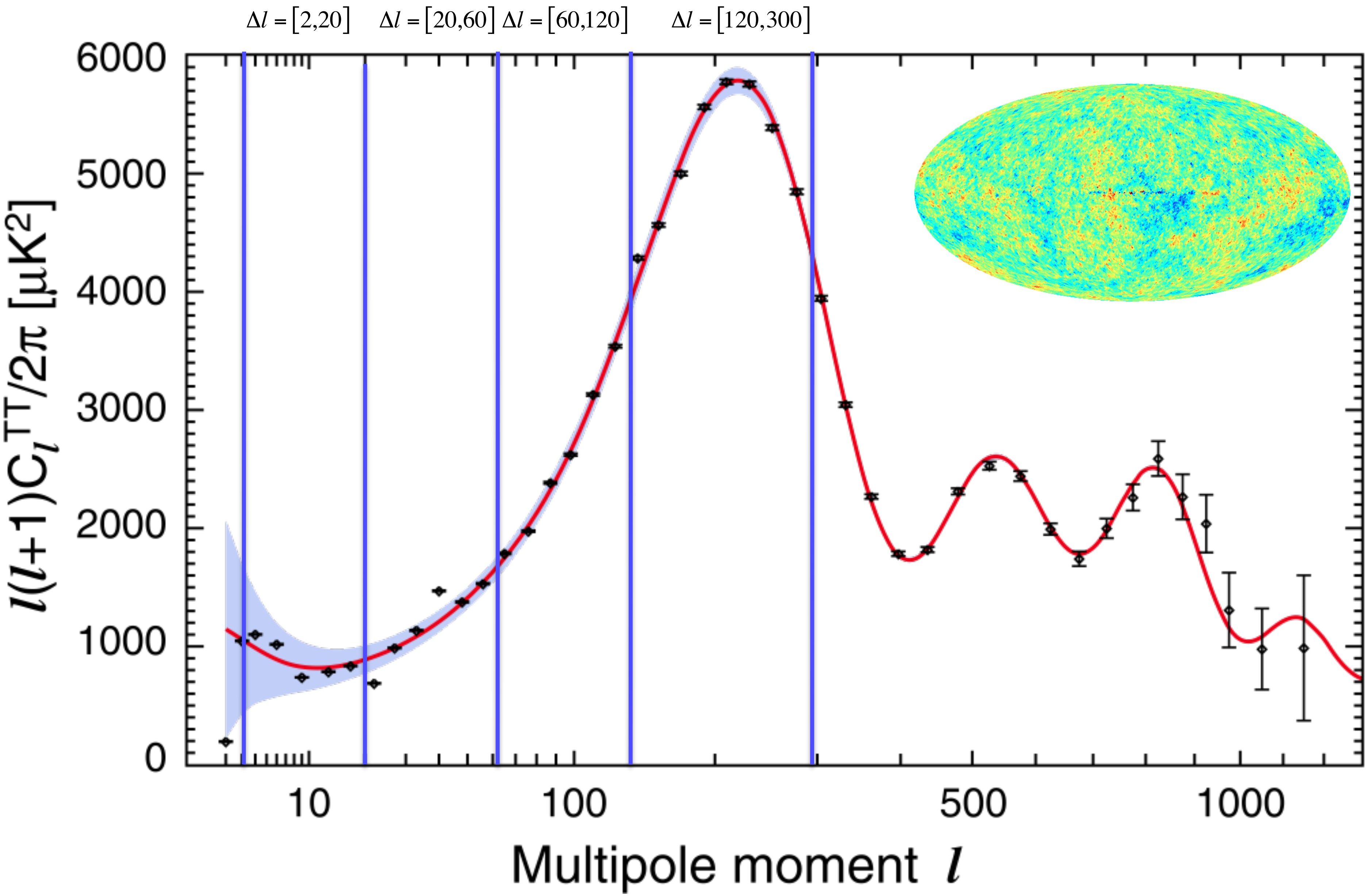


First order Surrogate: Shuffle outside (l_{min}, l_{max})

Second order Surrogates: Shuffle inside (l_{min}, l_{max})

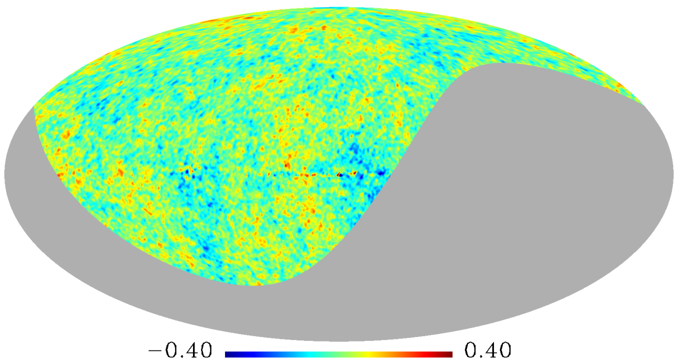
C. Räth et al., PRL, 2009

Generating Surrogates: Δl-intervals



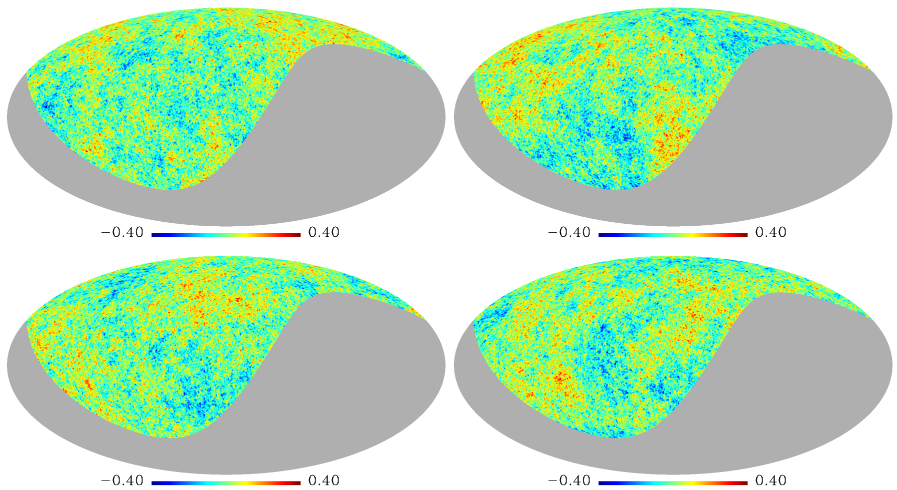
Deviation in rotated hemispheres

WMAP data / 1st order surrogate



compare with

Simulations / 1st or 2nd order Surrogates



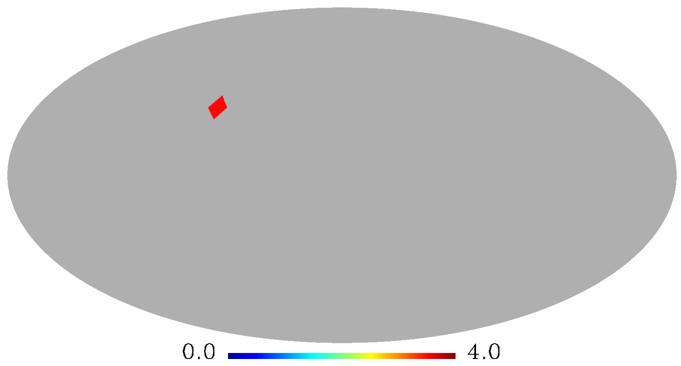
σ-normalised deviation S:

$$S(\vartheta,\phi) = \frac{X - \langle X \rangle}{\sigma_X},$$

$$X = <\alpha(r)>, \sigma_{\alpha(r)},$$

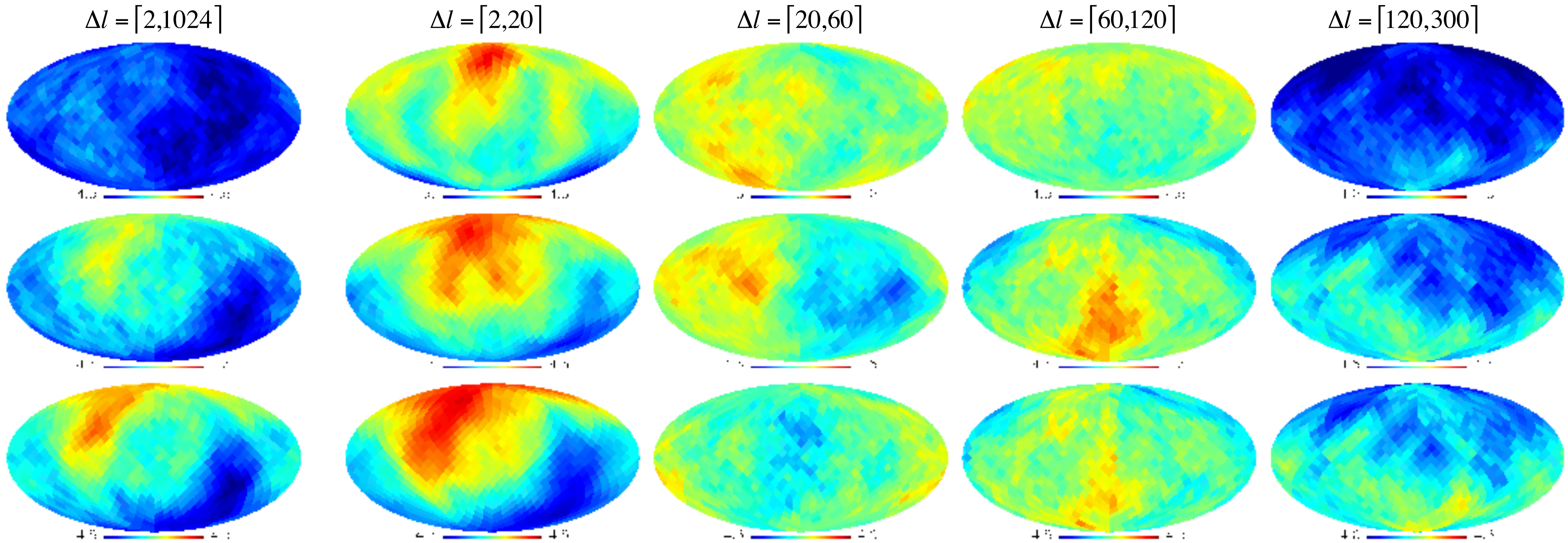
$$\chi^2(<\alpha(r)>, \sigma_{\alpha(r)}),$$

$$\chi^2(M_i), i = 0,1,2$$



Results for SIM

S(X) in rotated hemispheres for varying Δl and r:



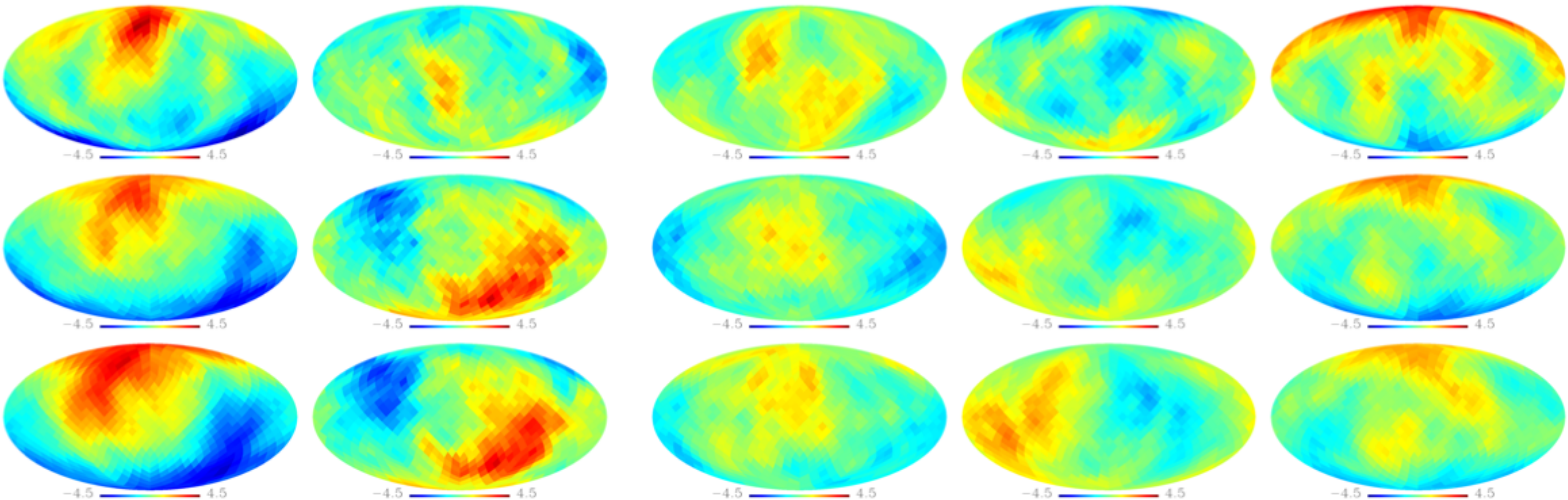
ILC 7yr map, X = $\langle \alpha_{r2} \rangle$, $\langle \alpha_{r6} \rangle$, $\langle \alpha_{r10} \rangle$ (from top to bottom)

- *Most significant deviations for Δl = [2,20] and Δl = [120,300]*
- *Signal in Δl = [2,1024] to be interpreted as superposition of the signals in Δl = [2,20] and Δl = [120,300]*

C. Räth et al., MNRAS, 2011

Results:

Checks on systematics ($\Delta l=[2,20]$):



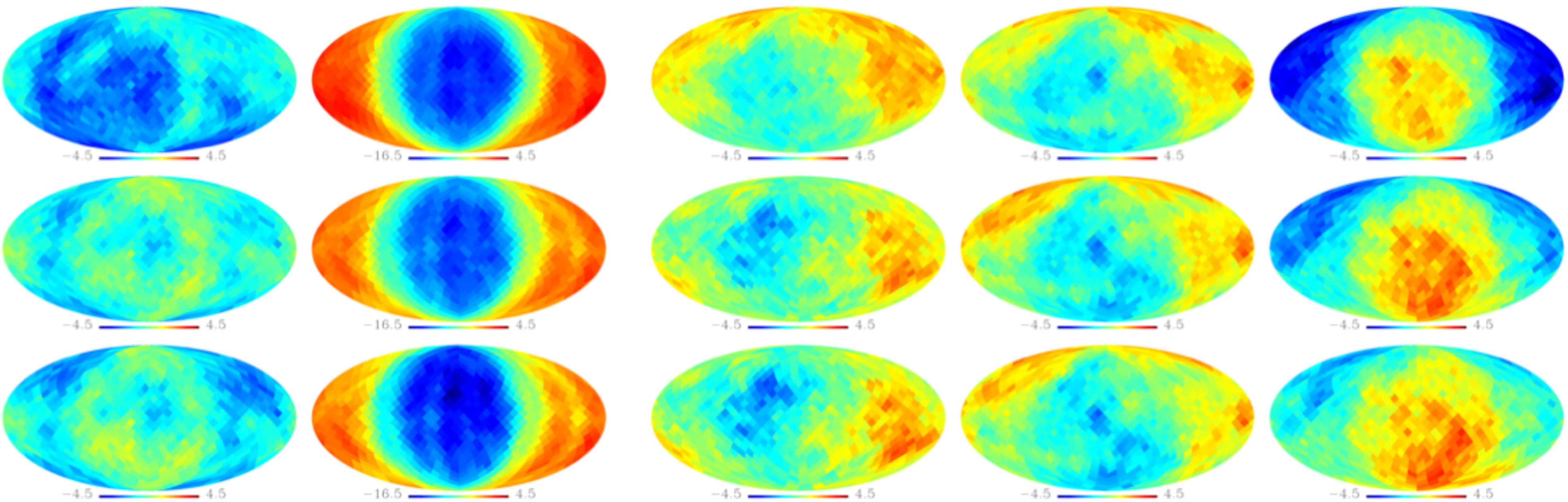
Uncorrected ILC map	Difference ILC map (year 7 – year6)	Asymmetric Beam map	Simulated Coadded VW-band map	Simulated ILC-like map

=> No test can so far explain the low-l anomalies!

C. Räth et al., MNRAS, 2011

Results:

Checks on systematics ($\Delta l=[120,300]$):



Uncorrected ILC map

Difference ILC map (year 7 – year6)

Asymmetric Beam map

Simulated Coadded VW-band map

Simulated ILC-like map

=> A number of ‚residuals' found for the high-l case

Results for MFs and SIM

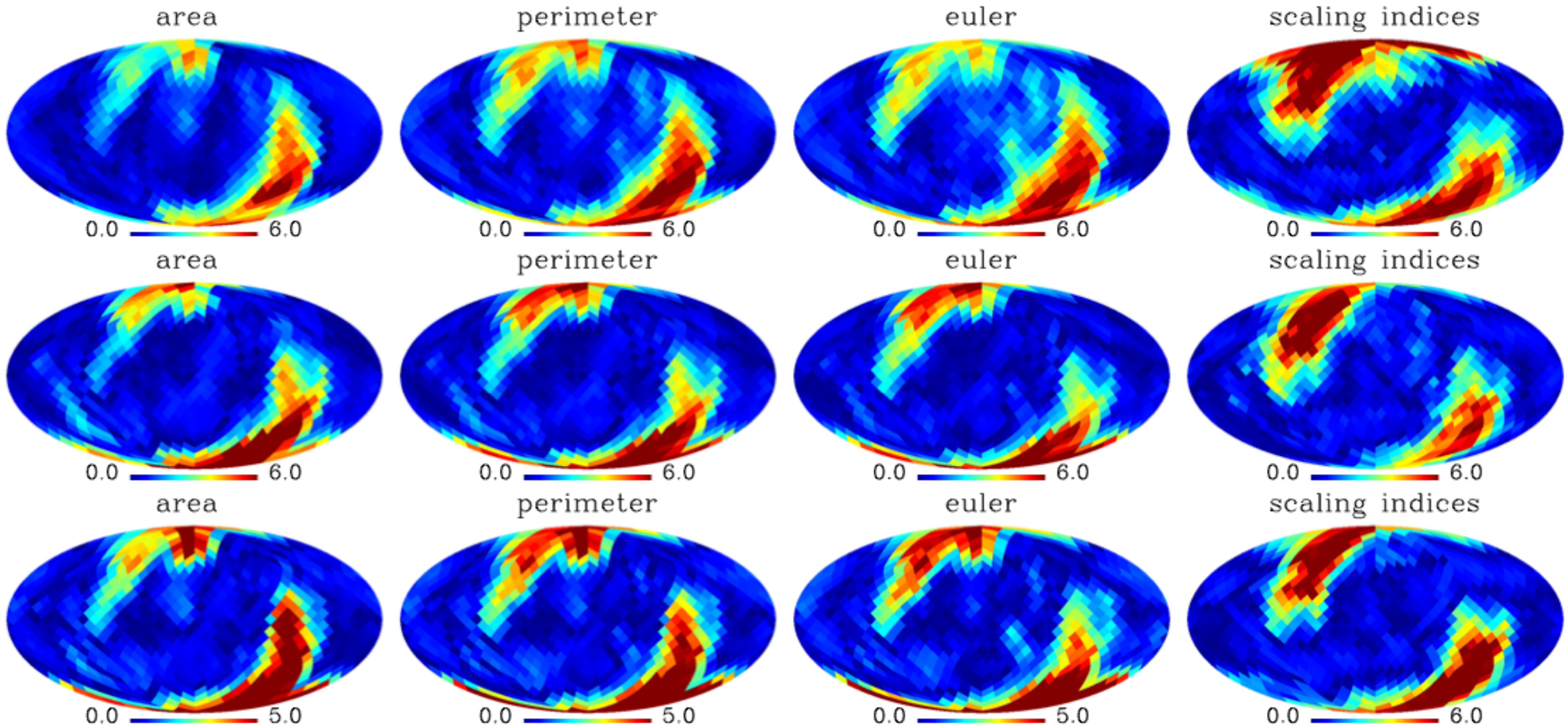
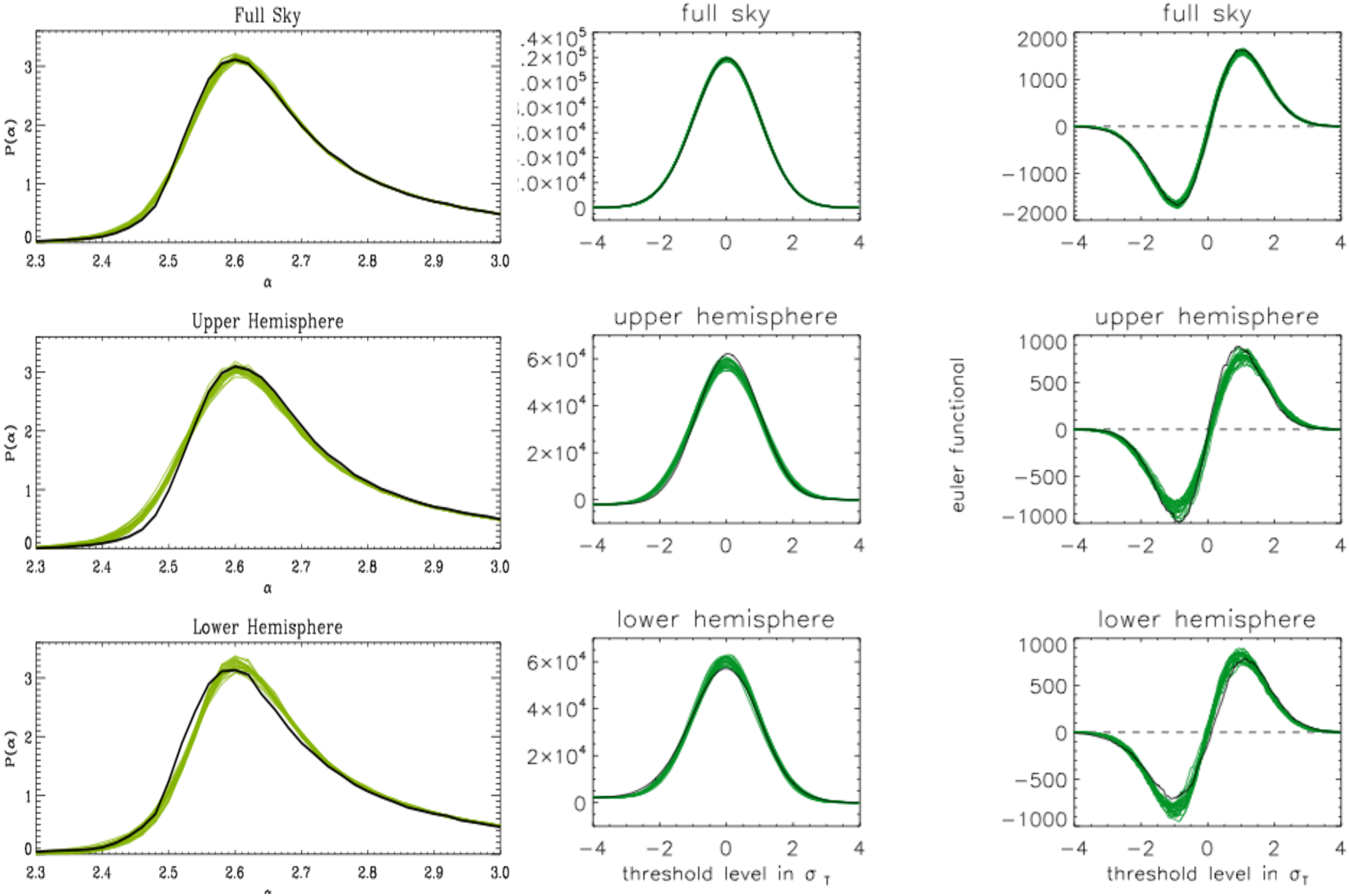


Figure 1. Deviations $S(\chi^2)$ of Minkowski Functionals M_0, M_1 and M_2 of the rotated hemispheres for the ILC7 (upper row, from left to right) and NILC7 map (middle row). In the lower row we show the results of the phase replacement method for NILC7. The l-range for the method of the surrogates is $\Delta l = [2, 20]$. The plots to the very right show the corresponding results $S(\chi^2)$ for the respective maps gained by the scaling index method.

H. Modest et al., MNRAS, 2012

Results for MFs and SIM



Results for MFs and SIM

	Full Sky	hemisphere S_{max}	hemisphere Opposite S
χ^2	$(S\|\%)$	$(S\|\%)$	$(S\|\%)$
Area	0.62 | 86.4	6.72 | 99.6	3.05 | 98.8
Perimeter	0.93 | 88.6	7.33 | >99.8	4.52 | 99.4
Euler	1.44 | 92.2	7.24 | >99.8	3.62 | 99.0
SIM	0.41 | 57.0	8.9 | >99.8	6.1 | 99.8

	Full Sky	hemisphere S_{max}	hemisphere Opposite S
χ^2	$(S\|\%)$	$(S\|\%)$	$(S\|\%)$
Area	1.03 | 88.2	9.51 | >99.8	5.98 | 99.8
Perimeter	0.89 | 86.4	9.97 | >99.8	7.31 | 99.8
Euler	0.77 | 84.4	9.50 | >99.8	7.22 | >99.8
SIM	0.29 | 51.4	7.53 | >99.8	6.23 | >99.8

Table 2. The same as Table 1, but for the NILC7 surrogate maps.

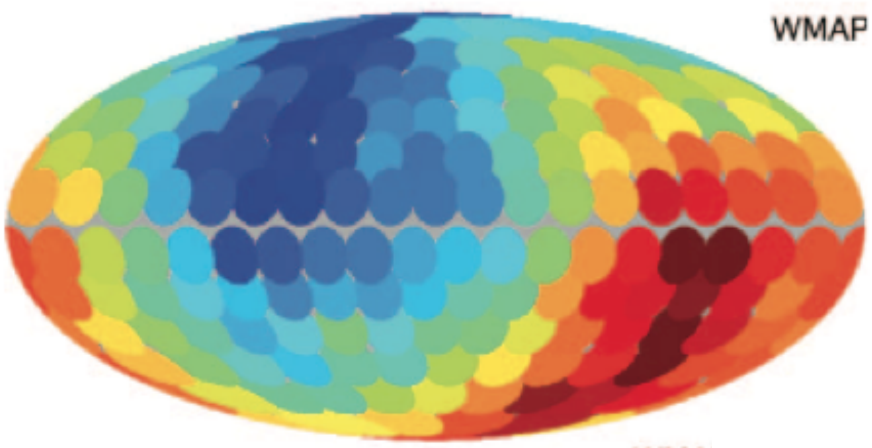
⇒Highly significant detection of NGs on large scales and of signatures anisotropies.

The signal is independent from:
- The input map
- The chosen higher order statistics

Thus, what about:
Single field slow roll inflation?
Copernican Principle?

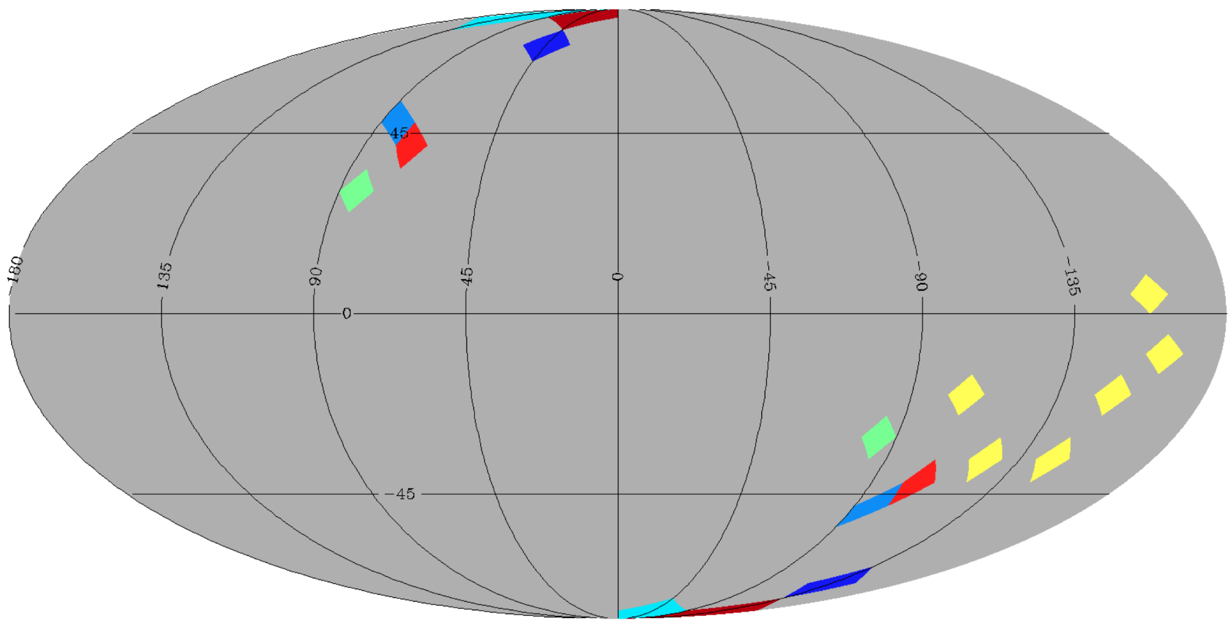
H. Modest et al., MNRAS, 2012

Results: Linear and nonlinear asymmetries



Hemispherical asymmetries of the Power spectrum
(e.g. Hansen et al., MNRAS, 2004
Hansen et al., ApJ, 2009)

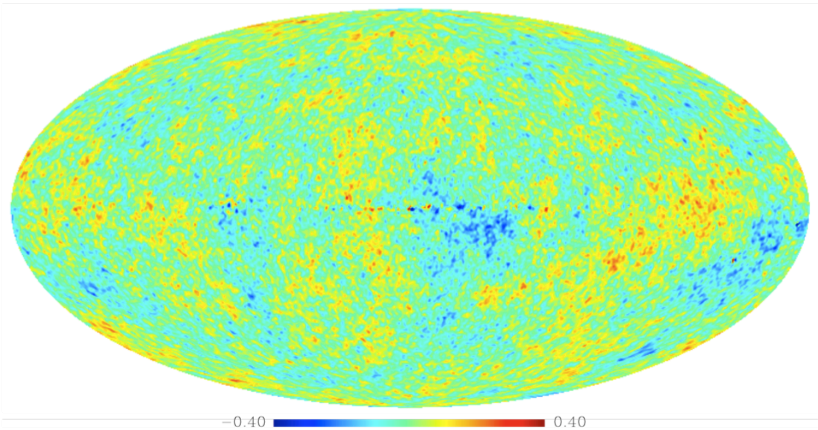
Directionality of the linear and nonlinear hemispherical asymmetries is not so different.

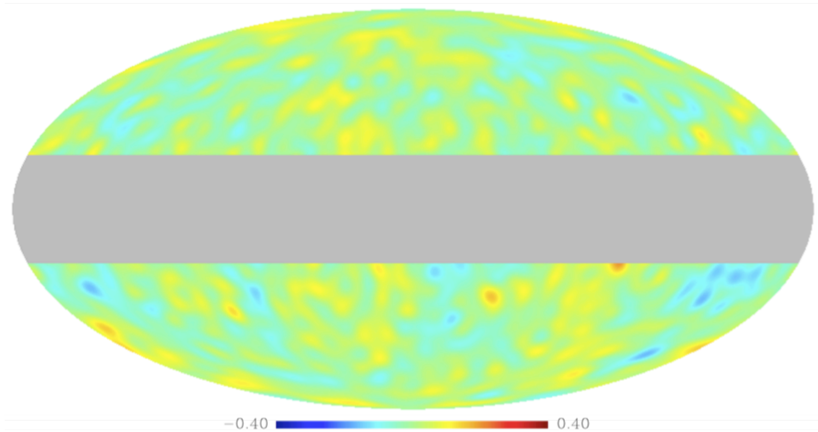


H. Modest et al., MNRAS, 2012

Surrogates for an incomplete sky

Possible foreground residuals in the galactic plane
⇒ Masking of the galactic plane
⇒ Basis functions Y_{lm} no longer orthogonal





$$f(x) = \sum_{\ell,m} a_{\ell m} Y_{\ell m}$$

$$f(x) = \sum_{\ell,m} a_{\ell m}^{cut} Y_{\ell m}^{cut}$$

Here: Cut is ± 20°, l_{max} = 40

Creating an orthonormal basis on an incomplete sky

How to obtain $a_{\ell m}^{cut}, Y_{\ell m}^{cut}$:

$$C = \int_{S^{cut}} Y(s) Y^*(s)\, d\Omega \quad \Rightarrow \quad C = A A^* \quad \Rightarrow \quad Y^{cut} = A^{-1} Y, \quad a^{cut} = A^T a$$

Construct the *Coupling Matrix* by integrating over the cut sky

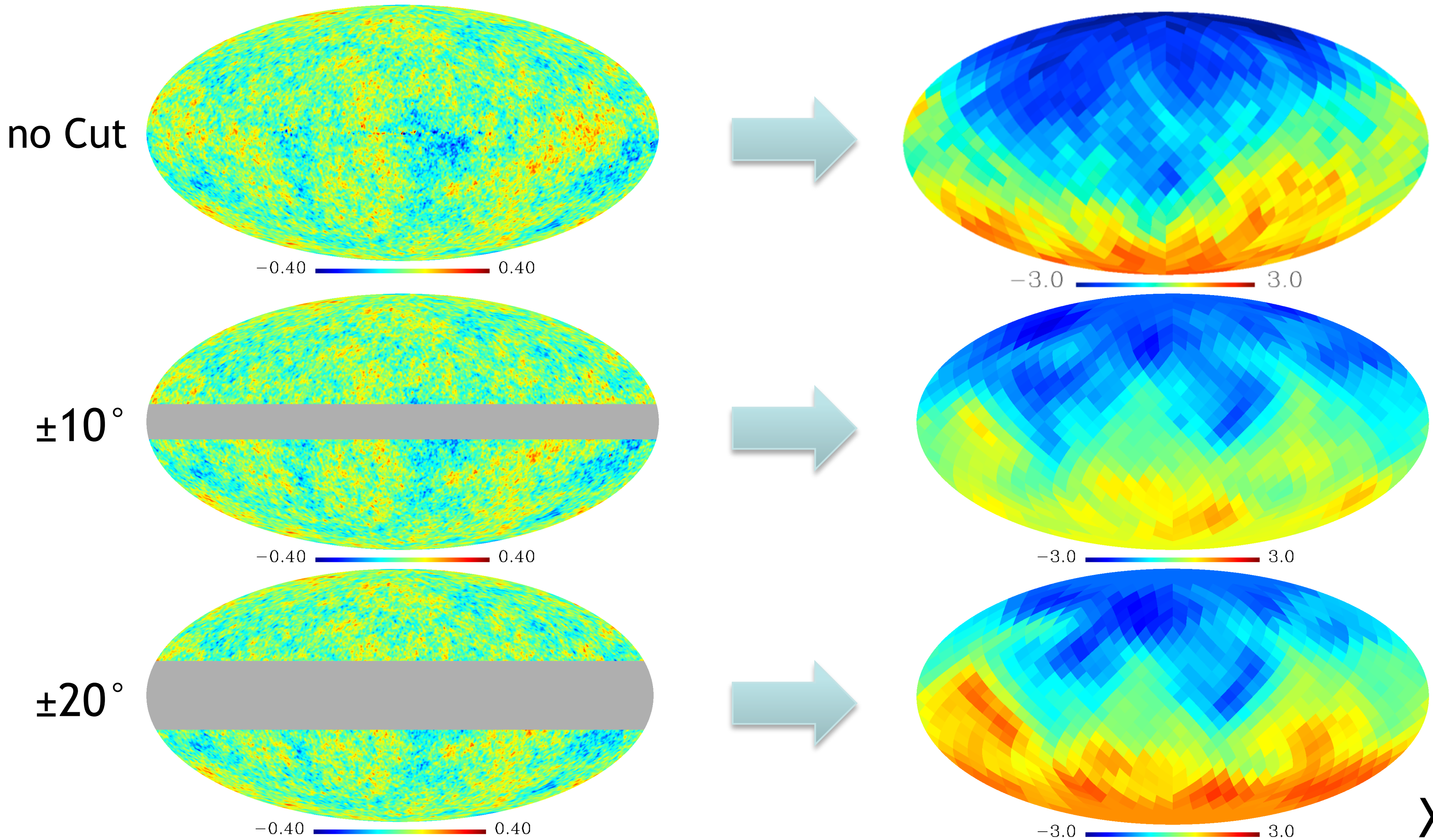
Decompose the Coupling Matrix with e.g. Cholesky Decomposition

Calculate the cut sky harmonics and its coefficients with the matrix A

Sounds straightforward, the implemention is, however, somewhat tedious....

See: Gorski et al. ApJ,1994a,b ,
Mortlock et al., MNRAS, 2002

Results for cut sky analysis



$X = \sigma(\alpha)$

Excluding the galactic plane doesn‘t change the results significantly.

Rossmanith et al., PRD, 2012

NGs of the local type

Perturbation of the curvature (NGs of the local type):

$$\psi(\vec{x}) = \psi_G(\vec{x}) + f_{NL}(\psi_G(\vec{x}) - < \psi_G(\vec{x}) >)^2$$

WMAP7 constraints on f_{NL}: $f_{NL} = 32 \pm 21$ (68% CL)
(Komatsu et al., ApJS, 2011)

Tests involving surrogates and f_{NL} realisations (Elsner & Wandelt, ApJ, 2010)

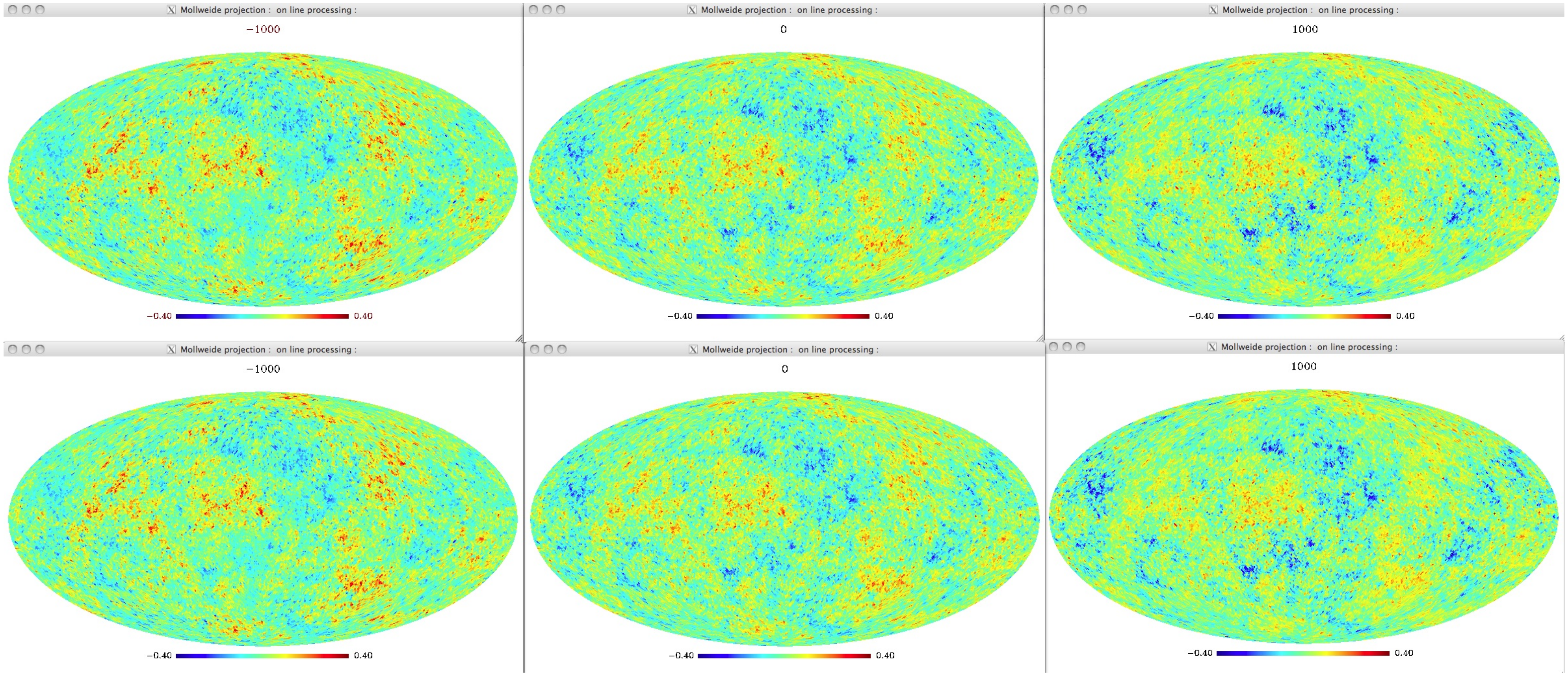
10 f_{NL} realisations, 5 f_{NL} values each (-1000, -100, 0, 100, 1000) + wmap data

1 1st order surro, 500 2nd order surros

$\Rightarrow$ 50 +1 * 501 maps = 25.551 maps

NGs of the local type

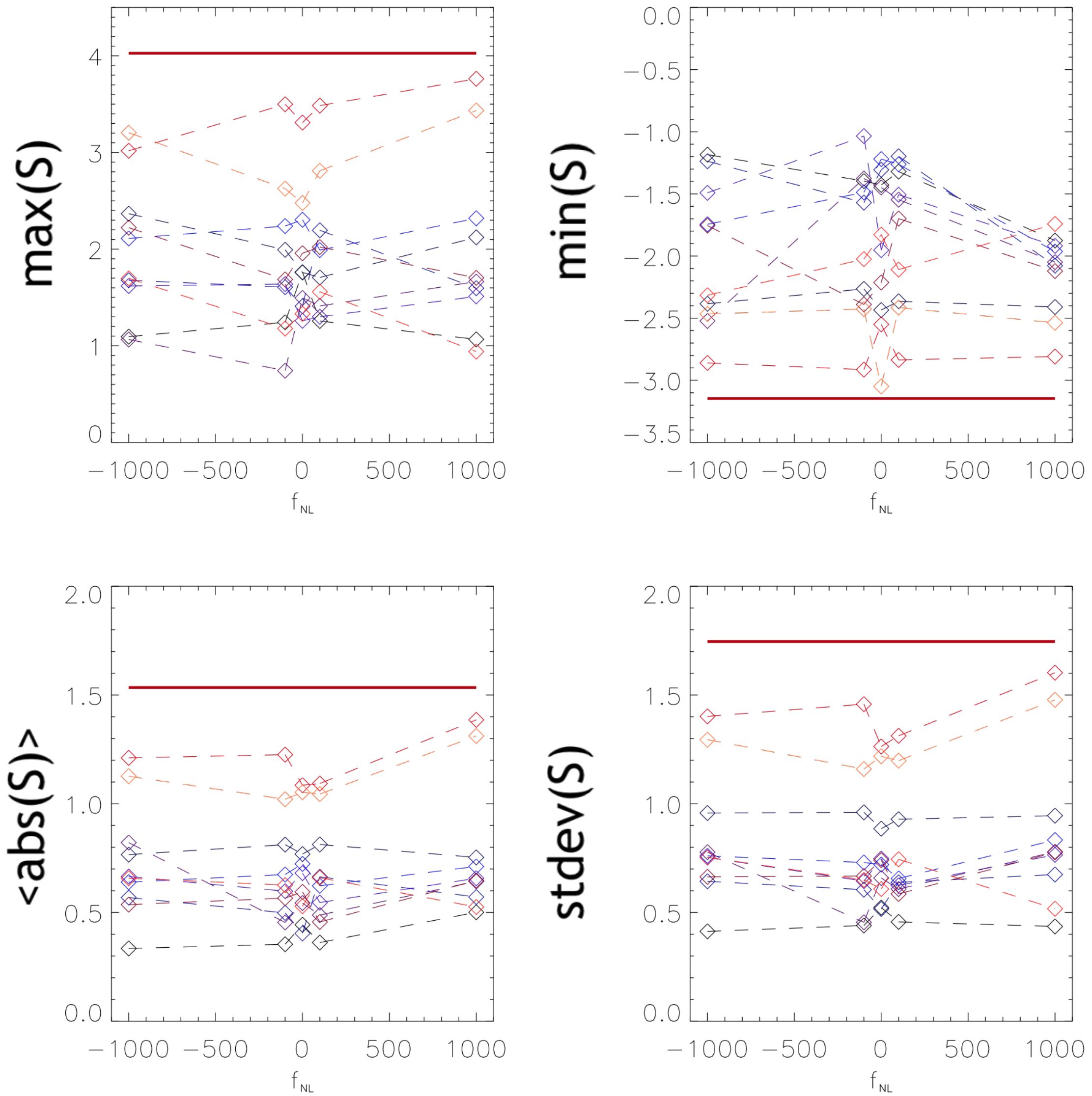
Simulation



Simulation + WMAP-like beam and noise properties

On the origin of low-l phase correlations

Statistics of S- maps based on scaling indices

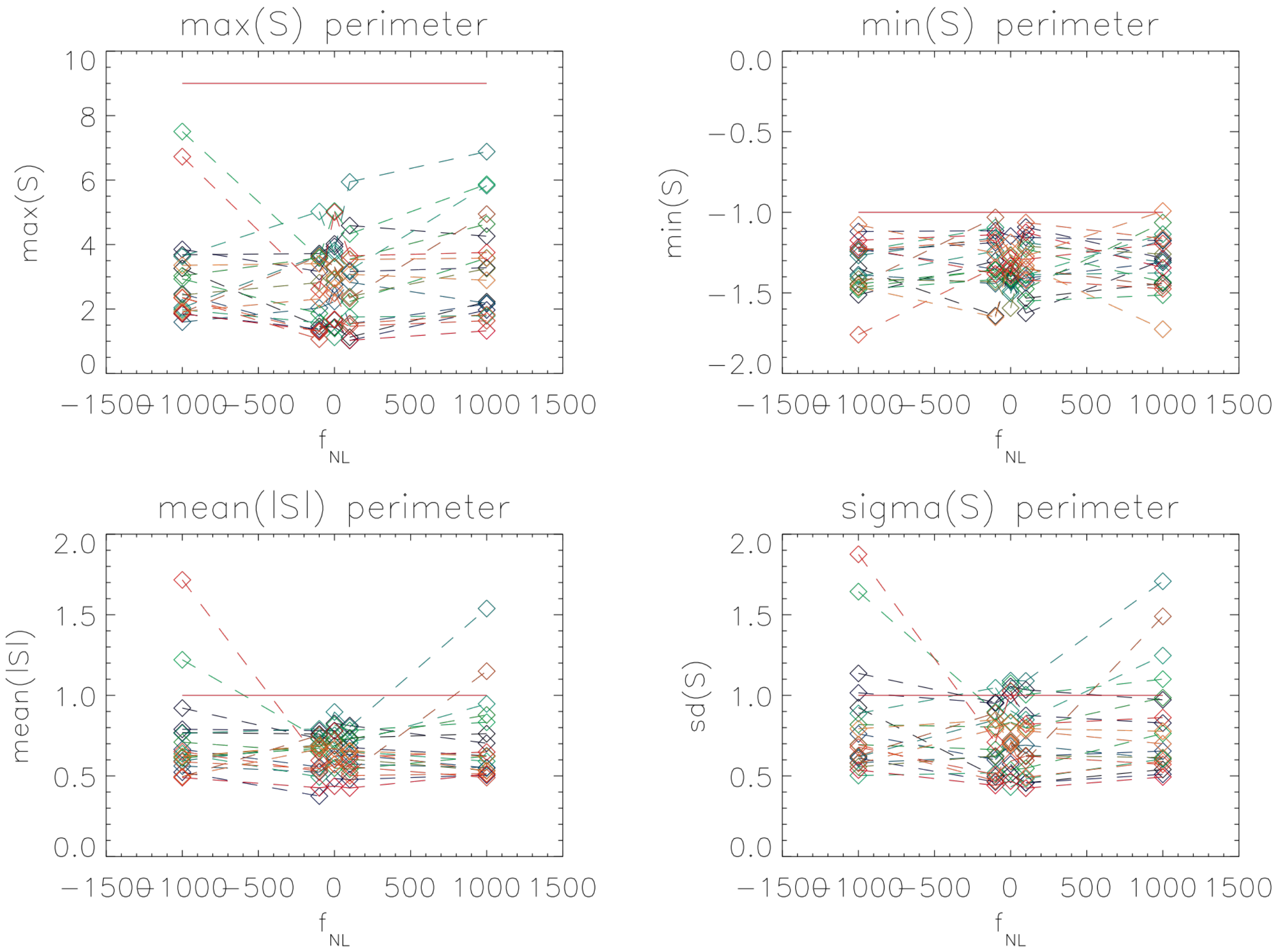


Nearly no variations with varying fnl

Neither extreme values nor first two moments can be reproduced by fnl maps

On the origin of low-l phase correlations

Statistics of S- maps based on Minkowski functionals



Only a few realisations show varaiations for very high abs(fnl)-values

For lower abs(fnl)-values the S-maps statistics of the CMB cannot be reproduced by the simulations

On the origin of low-l phase correlations

Another candidate: Bianchi-like template (see Jaffe et al., ApJ, 2006):

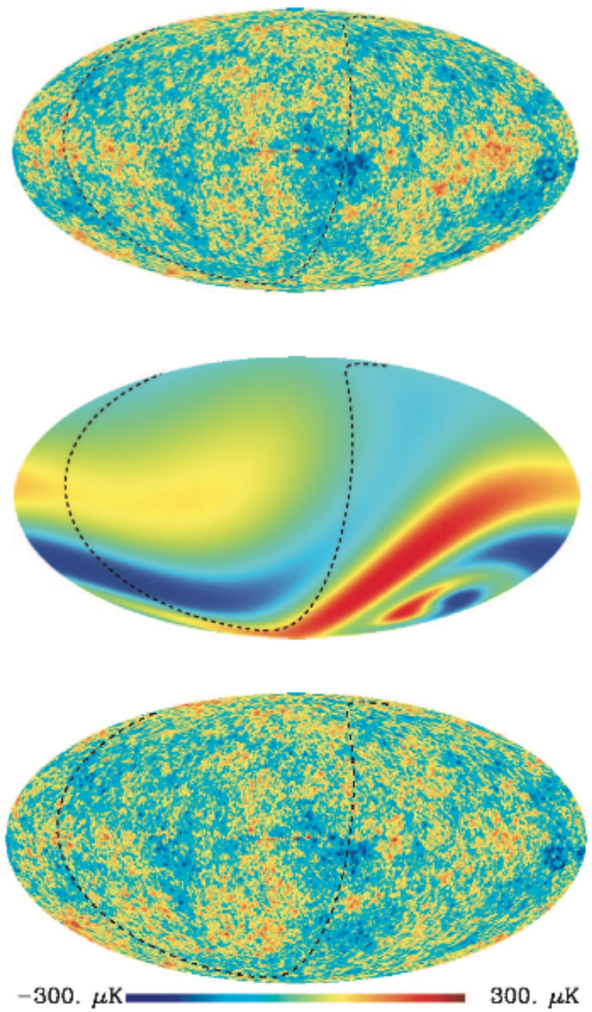


FIG. 4.—*Top*: *WMAP* Internal Linear Combination map. *Middle*: Best-fit Bianchi VII_h template (enhanced by a factor of 4 to bring out structure). *Bottom*: Difference between WILC and best-fit Bianchi template; the "Bianchi-corrected" ILC map. Overplotted on each as a dotted line is the equator in the reference frame that maximizes the power asymmetry as described in § 6.3.

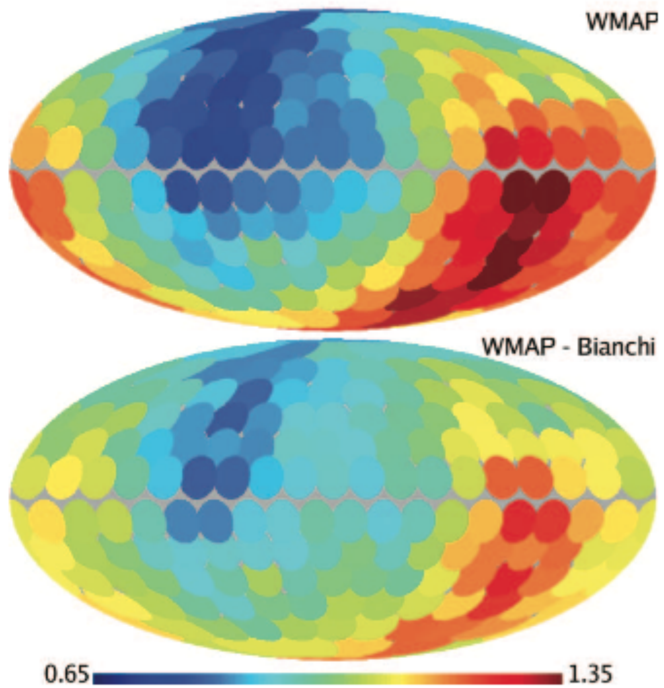


FIG. 10.—Power ratio between hemispheres in *WMAP* ILC, corrected (*bottom*) and uncorrected (*top*) for best-fit Bianchi component.

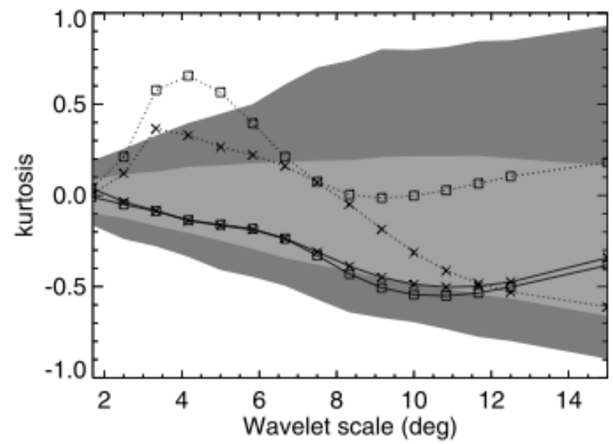
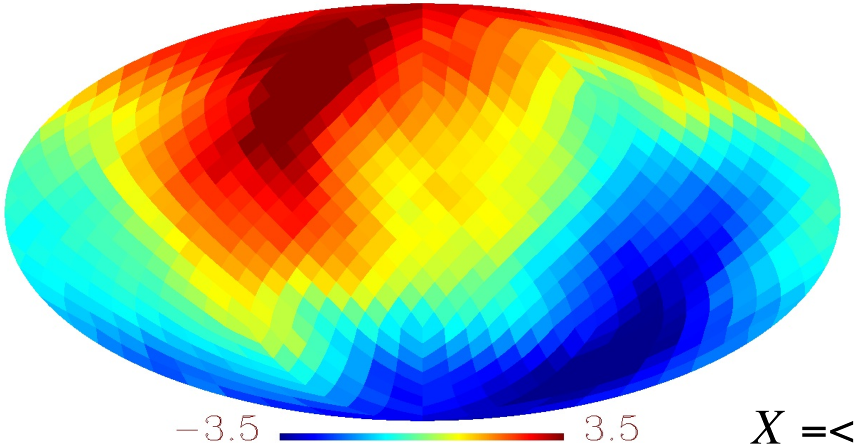


FIG. 11.—Kurtosis in wavelet coefficients. The boxes and crosses show the kurtosis before and after subtracting the Bianchi template, respectively, computed from the southern (*dotted line*) and northern (*solid line*) Galactic hemispheres.

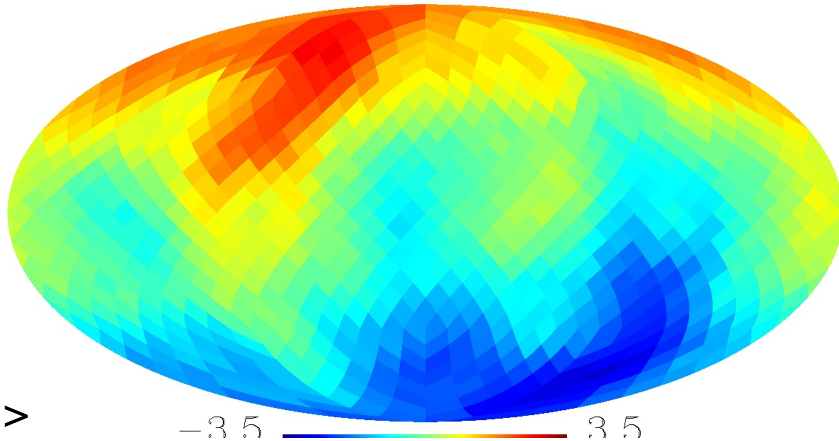
We consider the following two cases:
WMAP
WMAP - BIANCHI-Template

On the origin of low-l phase correlations

WMAP NILC7

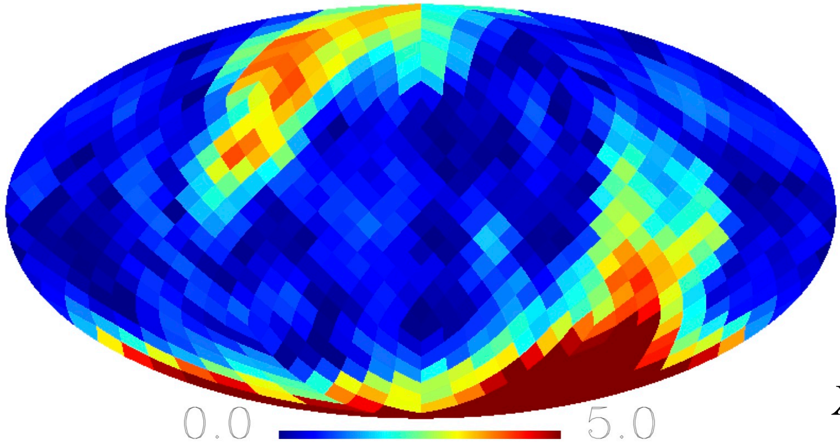


WMAP NILC7 – BIANCHI-Template

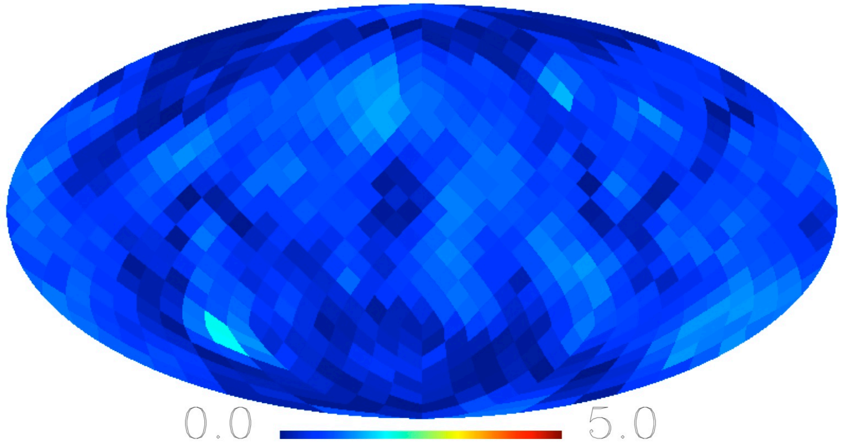


$X =< \alpha(r_{10}) >$

euler



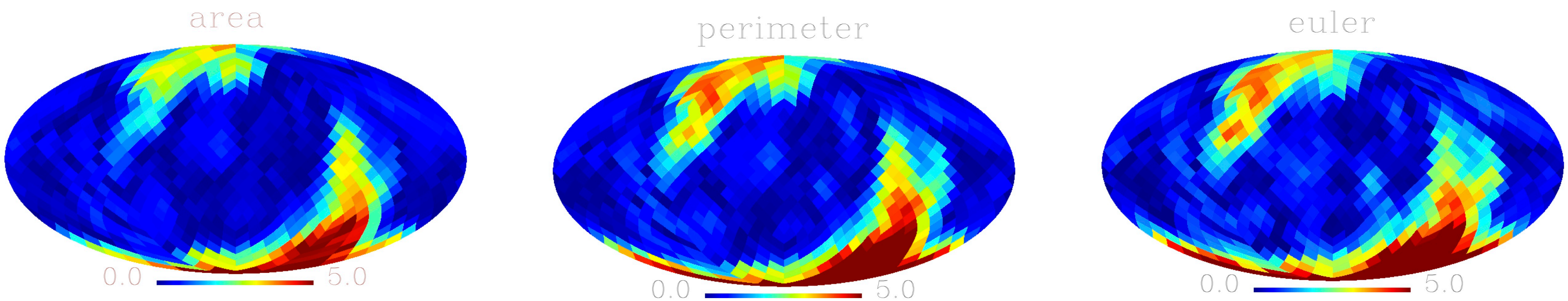
euler



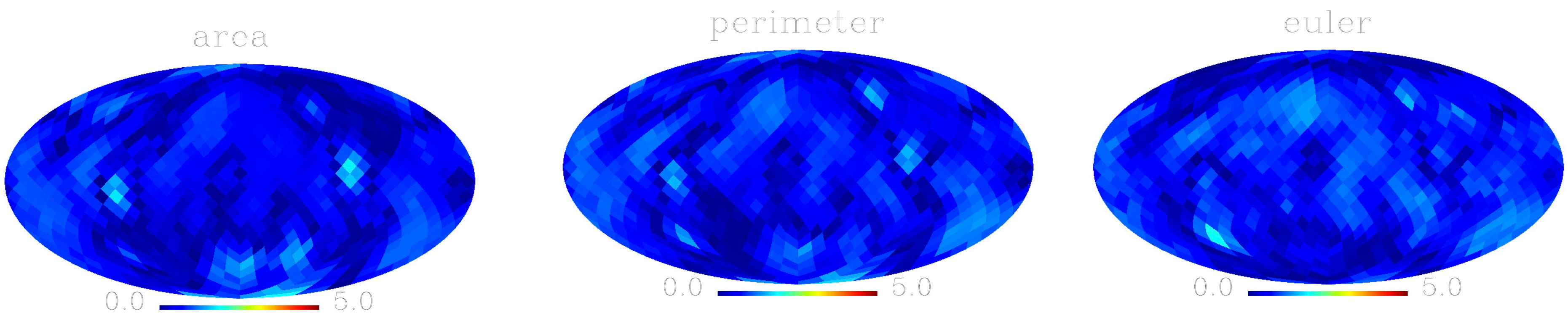
$X = \chi^2_{euler}$

On the origin of low-l phase correlations

WMAP NILC7



WMAP NILC7 – BIANCHI-Template



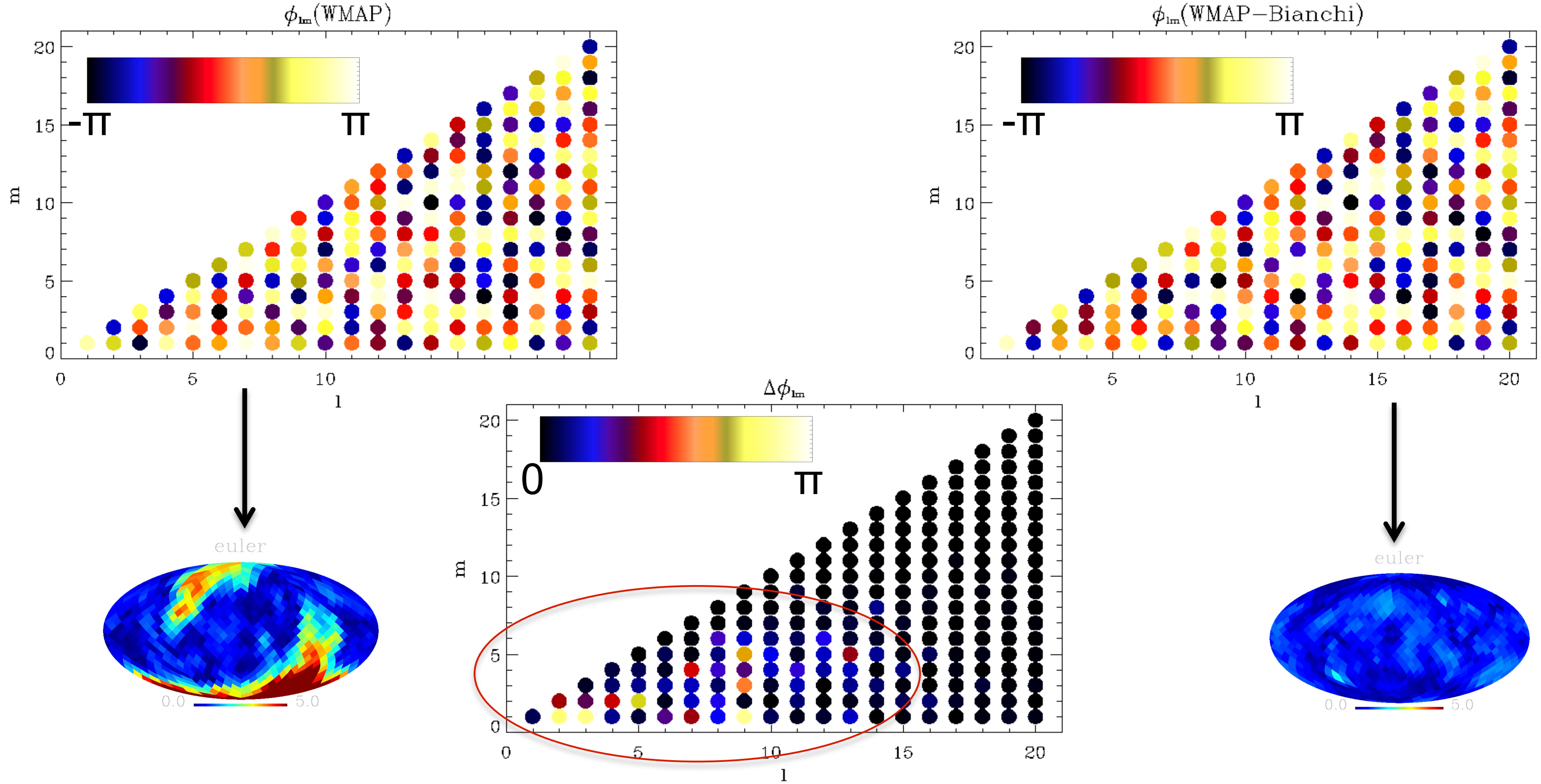
=>Interestingly enough, the anisotropic Bianchi template seems to be a viable model to (also) account for the low-l phase correlations

(H. Modest et al, in preparation)

A closer look at the low-l phase correlations

What makes the SIM-/MF-Signal appear/disappear ?

Low l-case (l<20)=> Number of basis functions Y_{lm} , and thus of phases φ_{lm}, is limited:



Only the variations in these modes make the difference.
Thus, the origin of the anomalies is considerably narrowed down.
More detailed parameter studies, more sophisticated surrogates
=> Relation between features of HOC in real space and phase information ?!

VI. Conclusions

- Surrogates are a versatile tool for (model-independent) data analysis, e.g. for detecting weak non-linearities in time series, non-Gaussianities in images etc.

- Not all surrogate generating algorithms are as good as they seemed to be. => Nonlinearities may remain undetected

However:

- Surrogates can help to shed (more) light on the meaning of Fourier phases and their relation to HOS
- Deeper understanding of the information coded in the phases may help in the development of nonlinear models

Thank you for your attention!