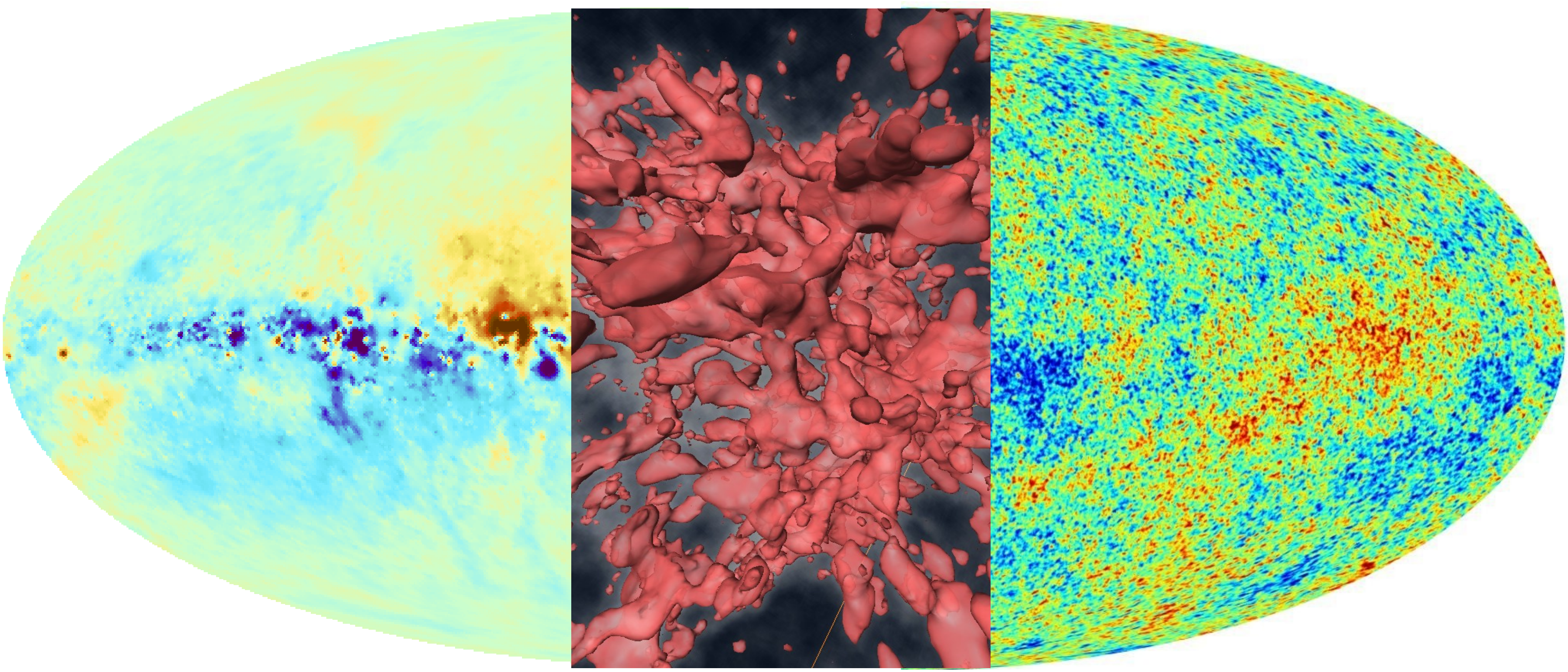
Information field theory

turning data into images

Max-Planck-Institut
für Astrophysik

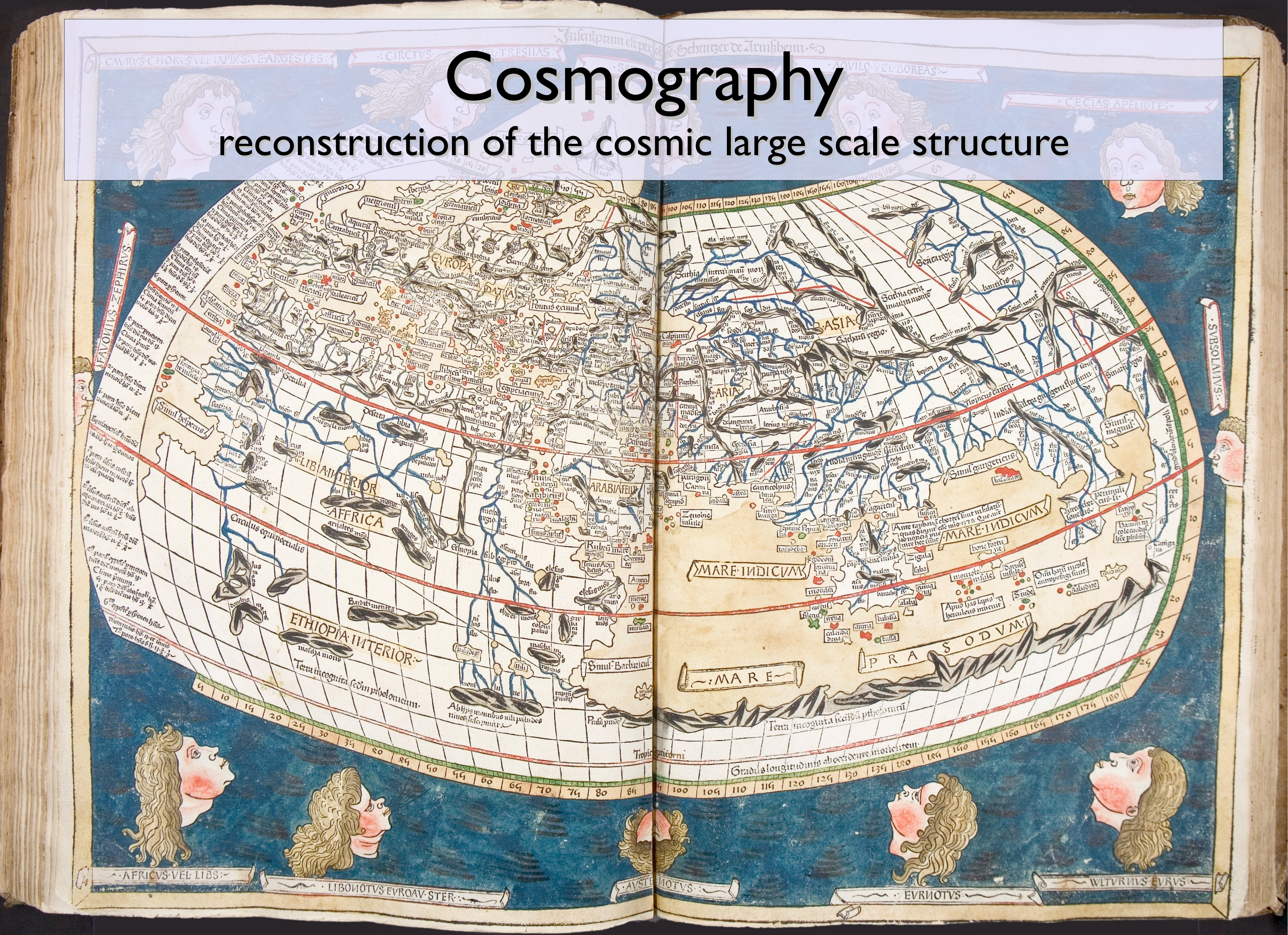
Torsten Enßlin

Mona Frommert, Jens Jasche, Henrik Junklewitz, Francisco Kitaura,
Niels Oppermann, Georg Robbers, Cornelius Weig, Marco Selig



Cosmography

reconstruction of the cosmic large scale structure



A signal field can only be reconstructed using prior information, since the infinite number of signal degrees of freedom are only constrained by a finite data set !

Corollary:
All imaging algorithms build on prior assumptions to regularize the reconstruction.

Why not to start directly with the assumptions?
space is continuous → information field theory

Cosmography

challenge:

how to translate ~10^6 galaxies uniquely into ~10^7 voxels?

→ ill-defined inverse problem, requires additional information

minimal additional information:

- statical homogeneity and isotropy, known power spectrum
- galaxy density traces dark matter density stochastically
- densities are non-negative, and vary logarithmically
- survey properties are understood (footprint, selection, ...)

information to be changed later:

- more realistic galaxy formation models
- including higher order n-point functions
- excluding cosmological prior

Information

"Information is what changes rational believes" (Caticha 2008)

"The actual science of logic is conversant at present only with things either certain, impossible, or entirely doubtful, none of which (fortunately) we have to reason on. Therefore the true logic for this world is the calculus of Probabilities, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind." (Maxwell 1850)

Cox's theorem (1946):
probability theory is the logic of uncertainties

Information Hamiltonian

$s = signal$
$d = data$

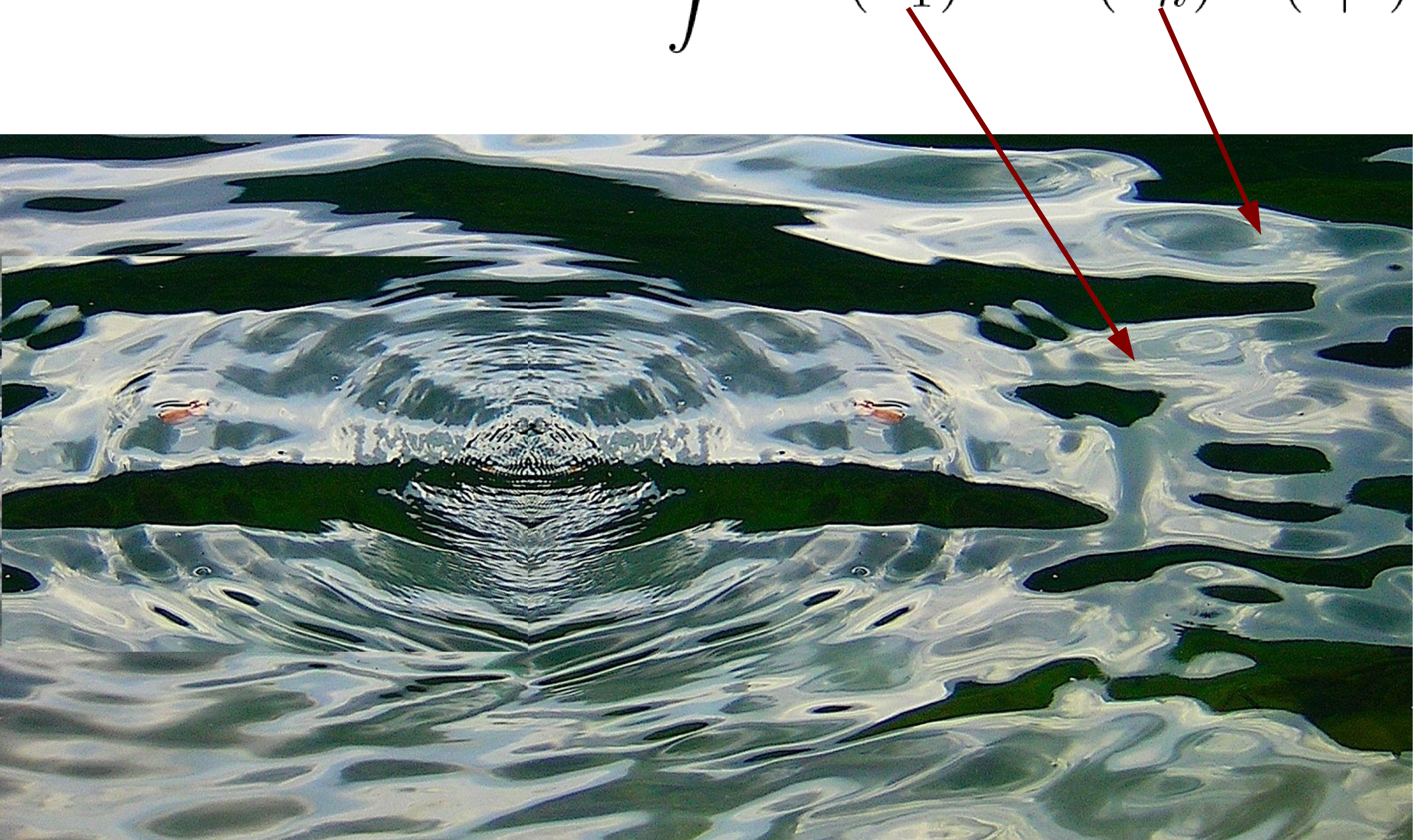
posterior — likelihood — prior

$$P(s|d) = \frac{P(d|s)\, P(s)}{P(d)}$$

evidence

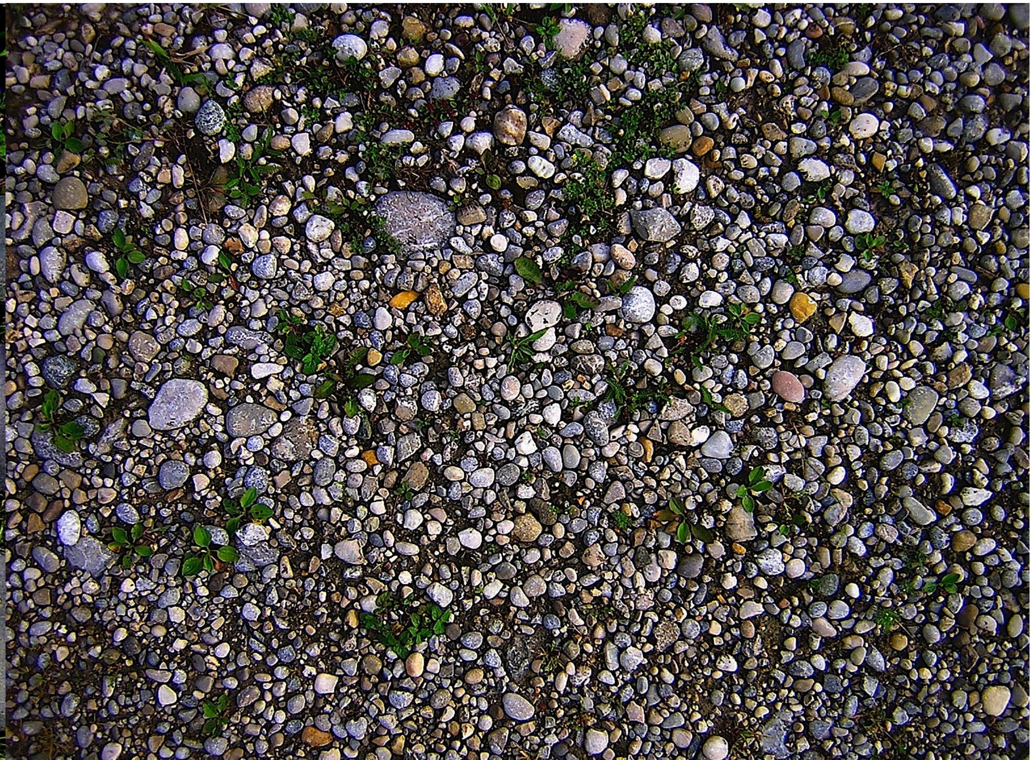
inference problem as information field theory

$$
\begin{aligned}
\langle s(x_1) \cdots s(x_n) \rangle_d &\equiv \langle s(x_1) \cdots s(x_n) \rangle_{(s|d)} \\
&\equiv \int \mathcal{D}s \, s(x_1) \cdots s(x_n) \, P(s|d)
\end{aligned}
$$



$$\begin{aligned}\langle s(x_1)\cdots s(x_n)\rangle_d &\equiv \langle s(x_1)\cdots s(x_n)\rangle_{(s|d)} \\ &\equiv \int \mathcal{D}s\, s(x_1)\cdots s(x_n)\, P(s|d)\end{aligned}$$

$$\int \mathcal{D}f\, F[f] \equiv \left(\prod_{i=1}^{N_{\mathrm{pix}}} \int df_i\right) F(f_1, \ldots, f_{N_{\mathrm{pix}}})$$



Free Theory

Gaussian signal & noise, linear response

signal :

$$
\begin{aligned}
P(s) &= \mathcal{G}(s,S) = \frac{1}{|2\pi\, S|^{1/2}} \exp\left(-\frac{1}{2}\, s^\dagger S^{-1} s\right) \\
j^\dagger s &= \int dx\, j^*(x) s(x) \\
S &= \langle s\, s^\dagger \rangle_{(s)}
\end{aligned}
$$

data :

$$
d = R\,s + n,\ P(d|s) = P(n = d - R\,s)
$$

noise :

$$
P(n) = \mathcal{G}(n,N),\ N = \langle n\, n^\dagger \rangle_{(n)}
$$

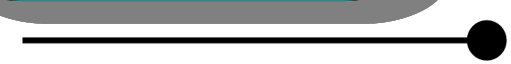
Wiener filter theory

known for 60 years

information source

$$
\begin{aligned}
H(s) &= -\log P(d,s) = -\log P(d|s) - \log \ldots \\
&= \frac{1}{2}(d - R\,s)^\dagger N^{-1}(d - R\,s) + \frac{1}{2}\,s^\dagger S^{-1} \ldots + \text{const} \\
&= \frac{1}{2}\,s^\dagger \underbrace{\left(S^{-1} + R^\dagger N^{-1} R\right)}_{\equiv D^{-1}} s + s^\dagger \underbrace{R^\dagger N^{-1} d}_{\equiv j} + \text{const} \\
&= \frac{1}{2}\,s^\dagger D^{-1}\,s + s^\dagger j + H_0
\end{aligned}
$$

information propagator

mean: $m = \langle s \rangle_{(s|d)} = D\,j =$ 

uncertainty: $\langle (s-m)\,(s-m)^\dagger \rangle_{(s|d)} = D$

Interacting Theory

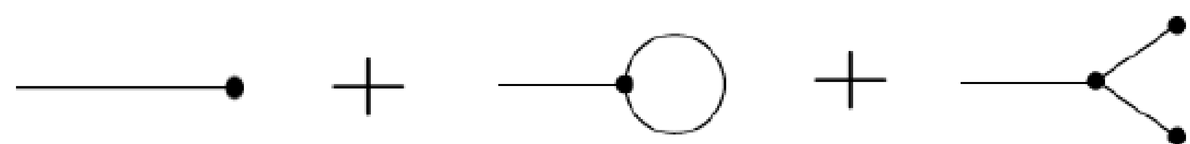
non-Gaussian signal, noise, or non-linear response

$$H[s] = \frac{1}{2}\, s^\dagger D^{-1}\, s - j^\dagger s + H_0 + \sum_{n=3}^{\infty} \frac{1}{n!}\, \Lambda^{(n)}_{x_1 \ldots x_n}\, s_{x_1} \cdots s_{x_n}$$

Taylor-Fréchet expansion of Hamiltonian

Use expansion into Feynman diagrams

$$\langle s \rangle_{(s|d)} =$$

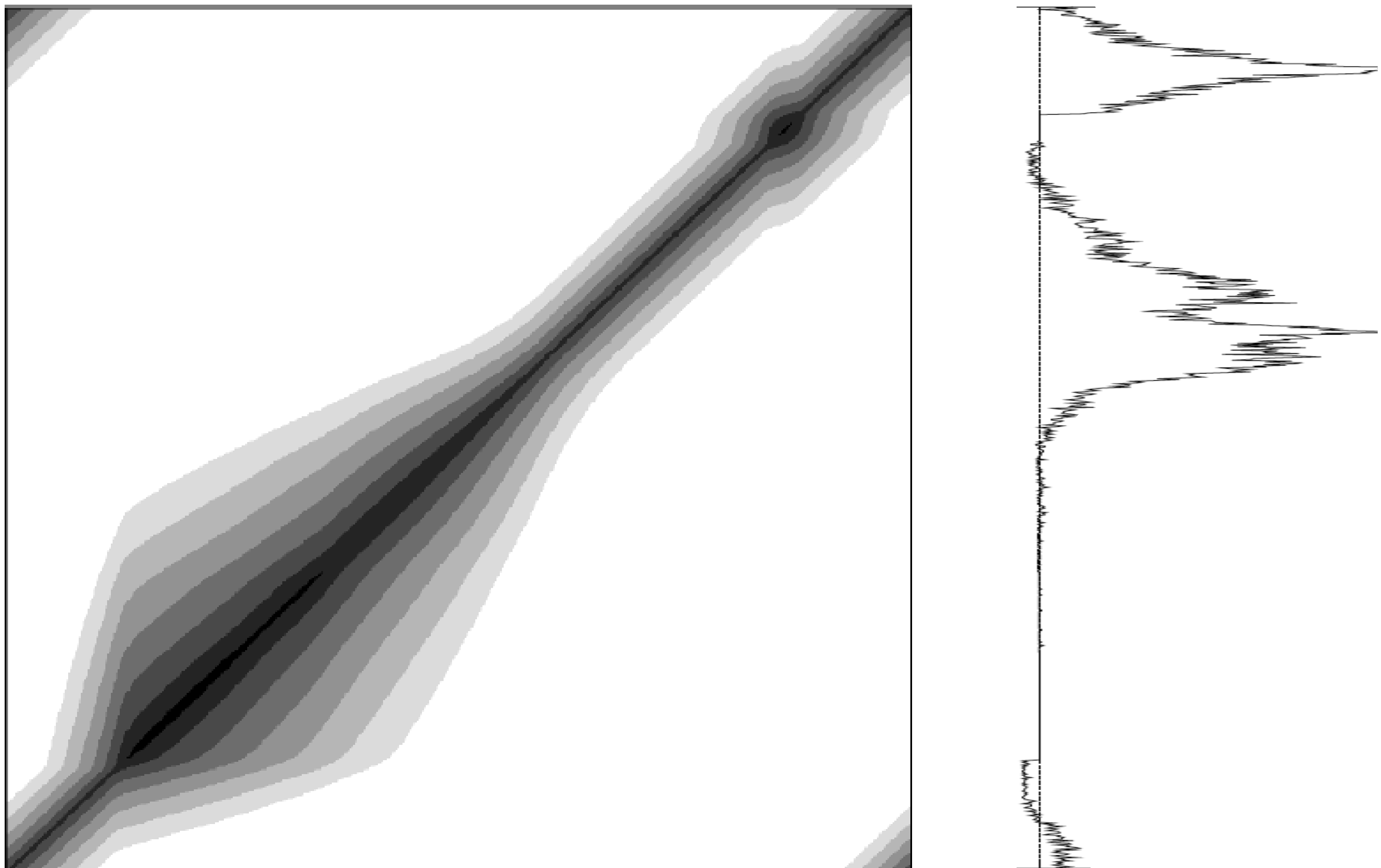


$+ \quad \ldots$

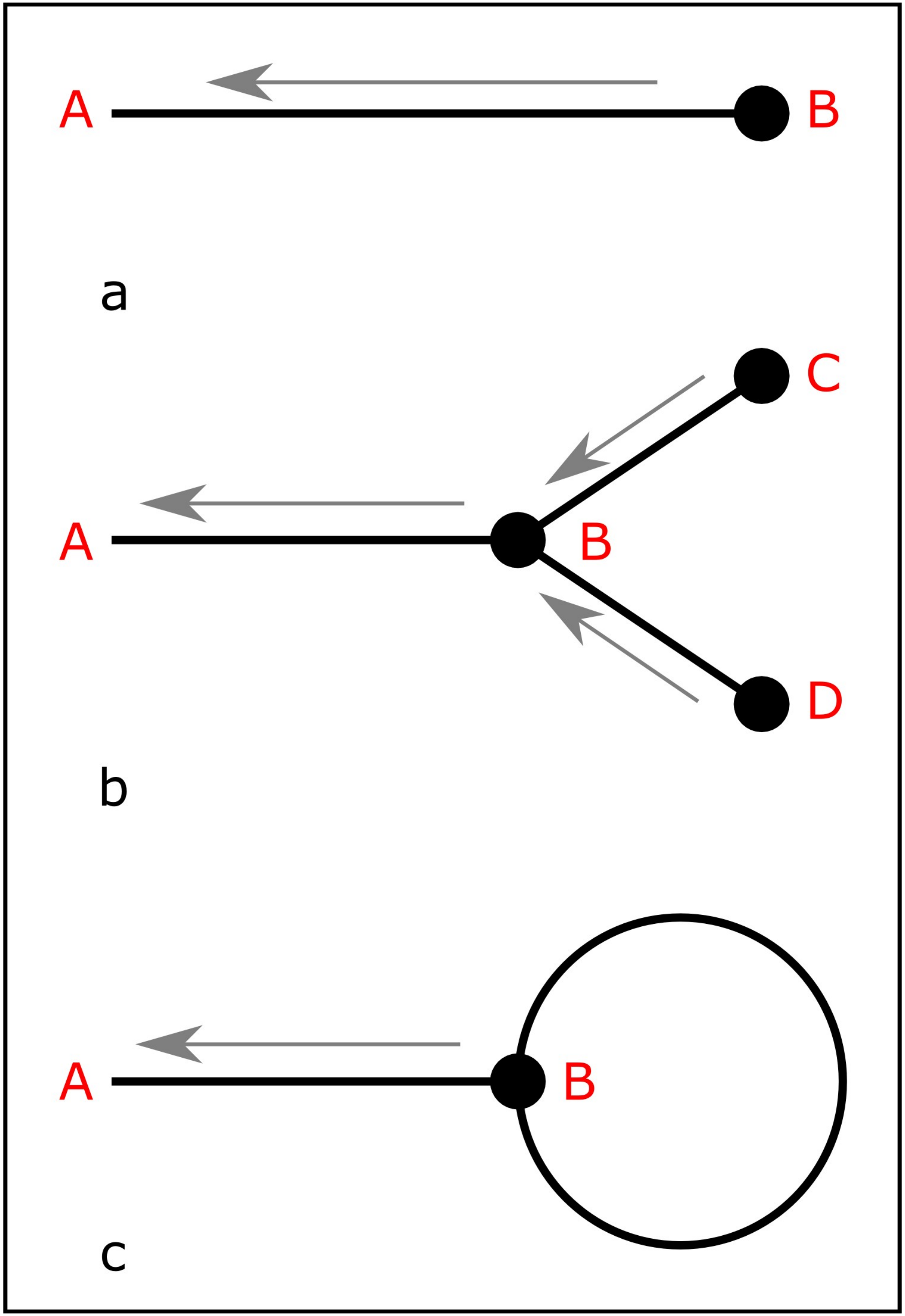
$$= D_{xy}\, j_y \;\; - \frac{1}{2} D_{xy}\, \Lambda^{(3)}_{yzu}\, D_{zu}$$

$$- \frac{1}{2} D_{xy}\, \Lambda^{(3)}_{yuz}\, D_{zz'}\, j_{z'}\, D_{uu'}\, j_u \quad + \quad \ldots$$

$$m = \langle s \rangle_{(s|d)} = D\, j = D_{xy}\, j_y$$



$$j = R^\dagger N^{-1} d \qquad D = \left[S^{-1} + R^\dagger N^{-1} R\right]^{-1}$$


A
B
a
C
A
B
D
b
A
B
c

dictionary

Translation:

inference problem $\rightarrow$ statistical field theory

$$P(s|d) = \frac{P(d|s)\,P(s)}{P(d)} \equiv \frac{1}{Z}\,e^{-H[s]}$$

Dictionary:

log-Posterior	=	negative Hamiltonian
Evidence	=	partition function Z
Wiener variance propagator	=	information
noise weighted data	$\rightarrow$	information source
inference algorithms	$\leftarrow$	Feynman diagrams
maximum a Posteriori	=	classical solution
uncertainty corrections	=	loop corrections
Shannon information	=	negative entropy

specifying information

Maximum entropy principle:

"In the problem of prediction, the maximization of entropy is not an application of a law of physics, but merely a method which ensures that no unconscious arbitrary assumptions have been introduced." (Jaynes 1957)

Cosmography:

- galaxy density traces dark matter density stochastically
  → Poissonian likelihood
- densities are non-negative, and vary logarithmically
  → log-density is ideal signal coordinate
- two point statistics is known
  → Gaussian prior for log-density signal

large-scale structure reconstruction

- galaxies trace dark matter density field
- log-normal density field: $\varrho = \varrho_0 \, e^{c\,s}$
- galaxy shot noise due to Poisson statistics
- inhomogeneous observation

Hamiltonian:

data = galaxy counts

galaxy bias

$$H[s] = \frac{1}{2}\, s^\dagger S^{-1} s - b\, d^\dagger s + \kappa_0\, e^{b\,s}$$

covariance of Gaussian log-density field *s*

expected mean galaxy counts

Enßlin, Frommert, Kitaura (2009)

classical & renormalized solution

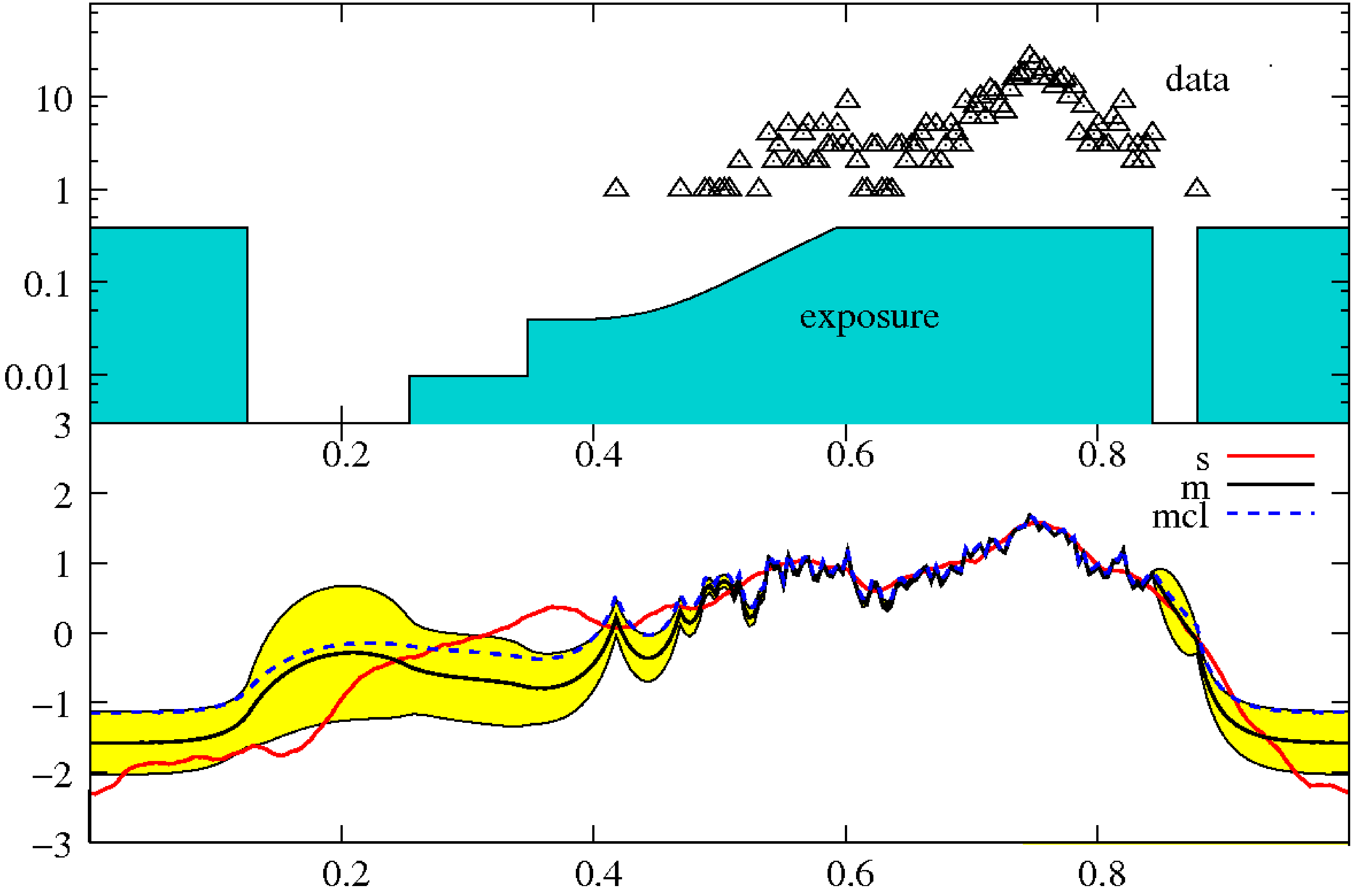
fix point solution for different temperatures T

T=0 classical solution, T=1: field theoretical solution

$$m = b\,S\,\left(d - \kappa_{b\,m + T\,b\,\widehat{D}/2}\right),$$

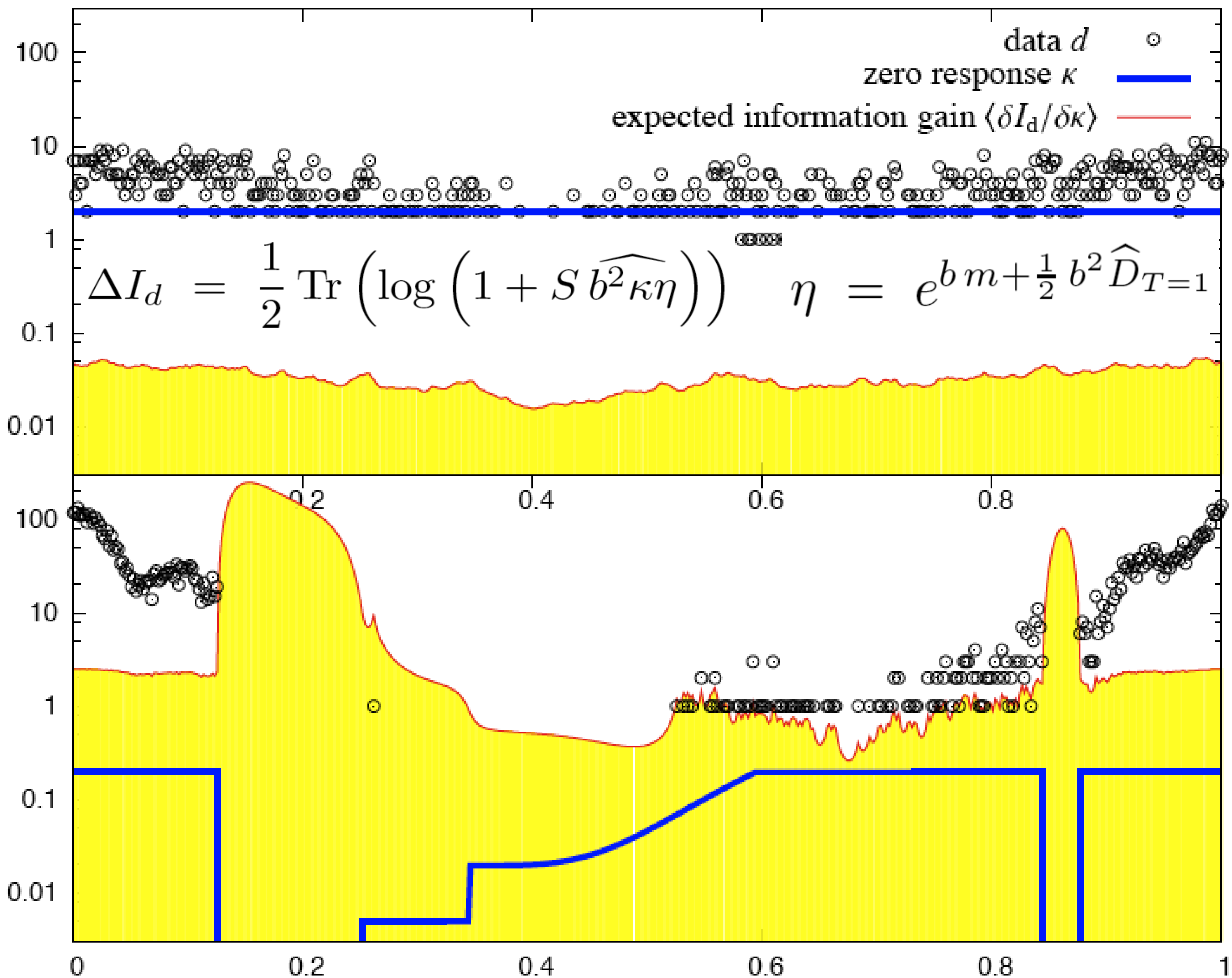
$$D = \left(S^{-1} + \widehat{\kappa}_{b\,m + T\,b\,\widehat{D}/2}\right)^{-1}$$

$$\kappa_s = \kappa\,e^{b\,s}$$



data
exposure
10
1
0.1
0.01
0.2
0.4
0.6
0.8
s
m
mcl
3
2
1
0
−1
−2
−3
0.2
0.4
0.6
0.8

expected information gain



non-linear reconstruction

nonlinear reconstruction of log-density from SDSS DR7 using Hamiltonian sampling

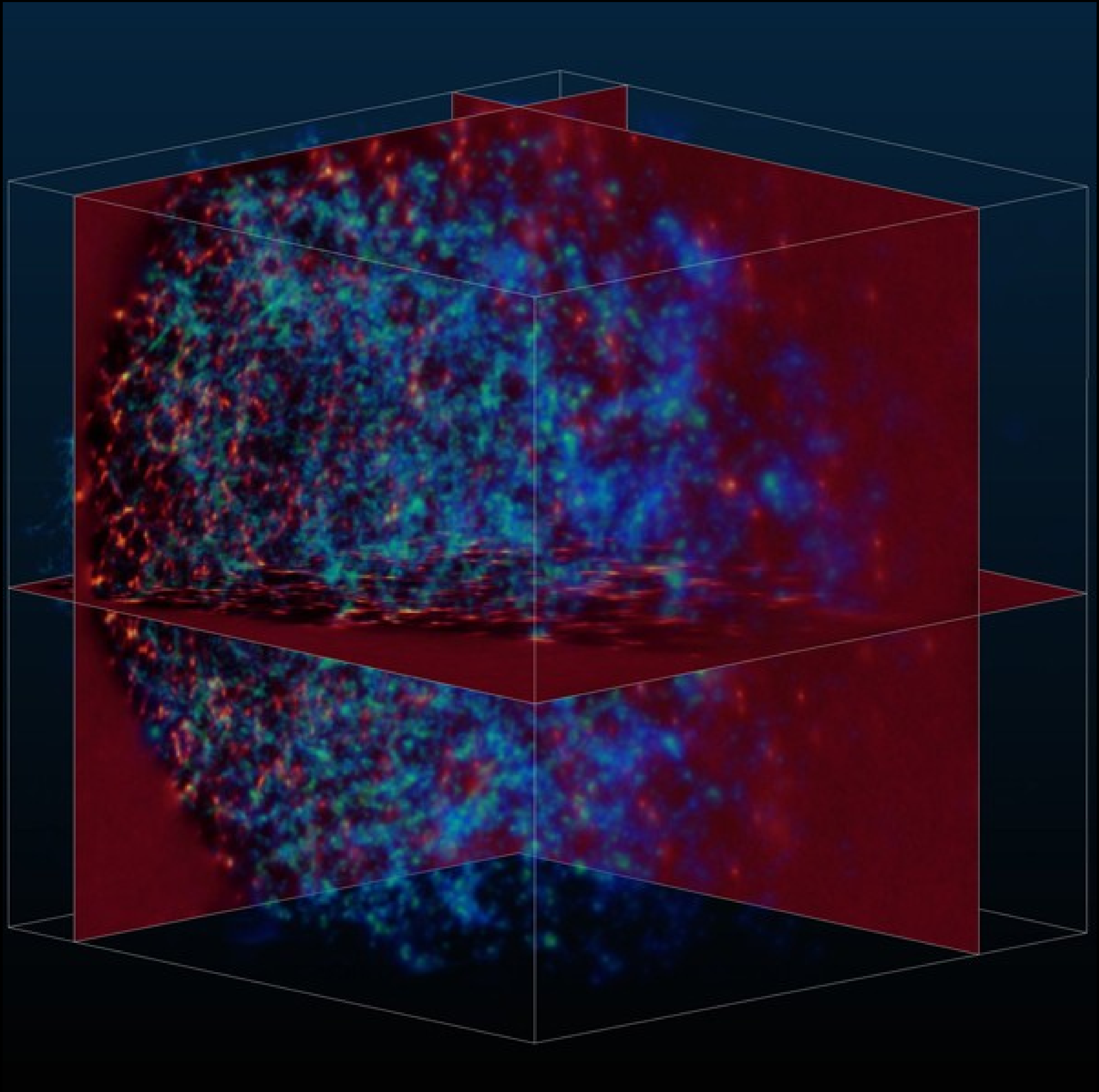
main sample galaxies $0.001 < z < 0.4$, covering a quarter of the sky

box : 750 Mpc/h per side, 256 voxels per side
→ resolution ~ 3 Mpc/h

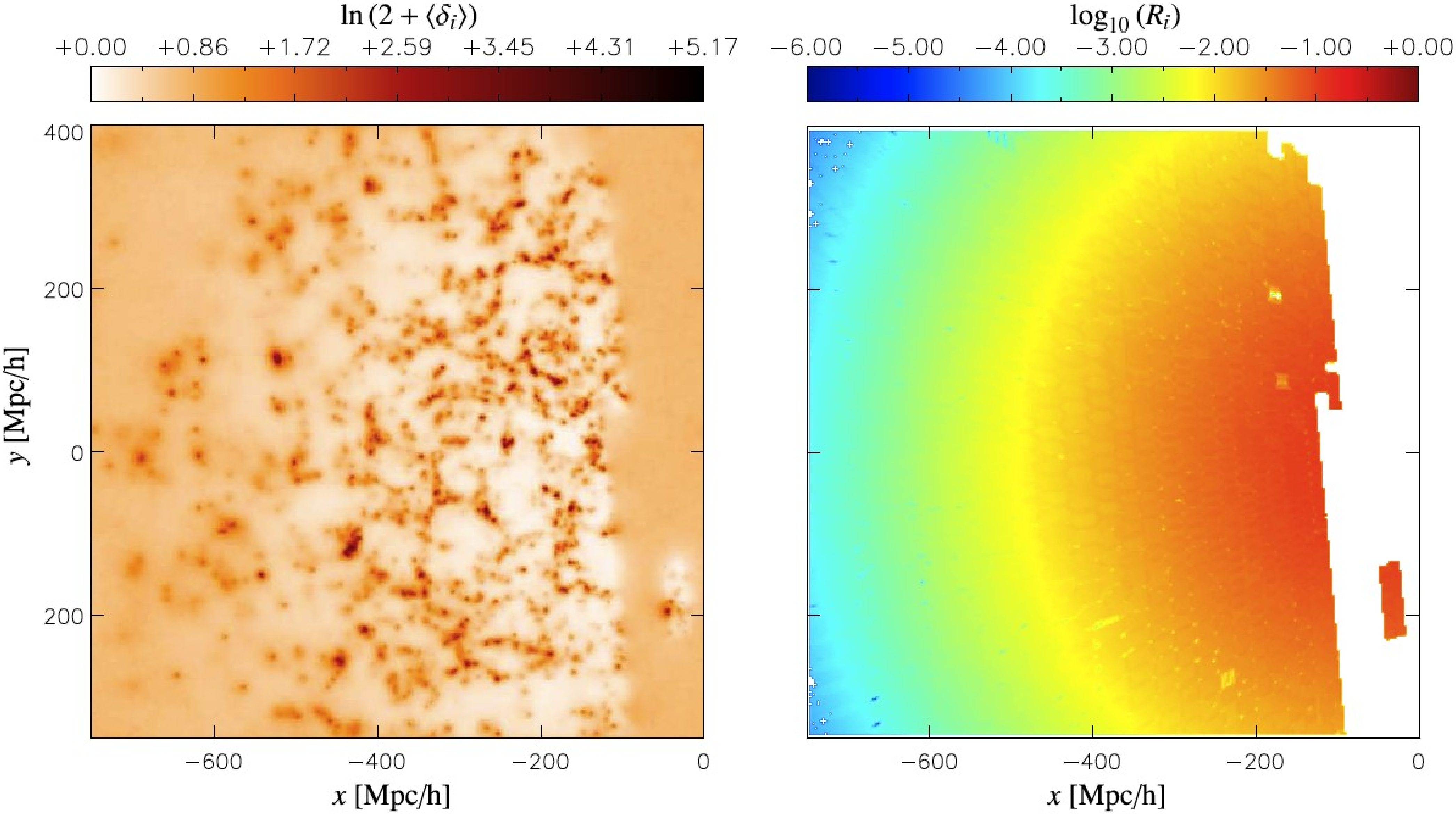
corrected for

- selection function
- completeness
- full shot noise

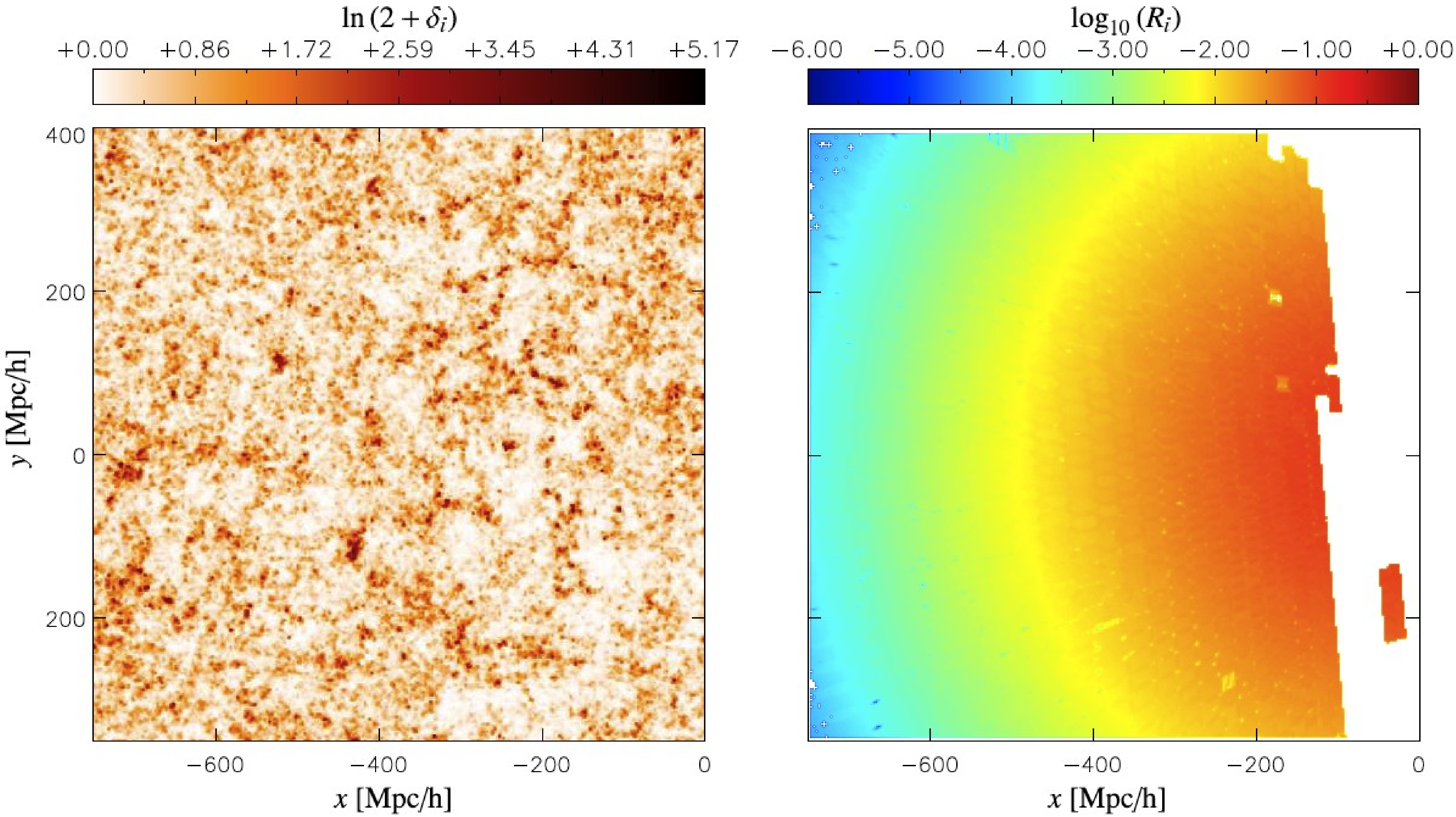
Jasche, Kitaura, Li, Enßlin (2010)



mean + mask


$\ln(2+\langle\delta_i\rangle)$
+0.00
+0.86
+1.72
+2.59
+3.45
+4.31
+5.17
$\log_{10}(R_i)$
−6.00
−5.00
−4.00
−3.00
−2.00
−1.00
+0.00
400
200
0
200
y [Mpc/h]
−600
−400
−200
0
x [Mpc/h]
−600
−400
−200
0
x [Mpc/h]

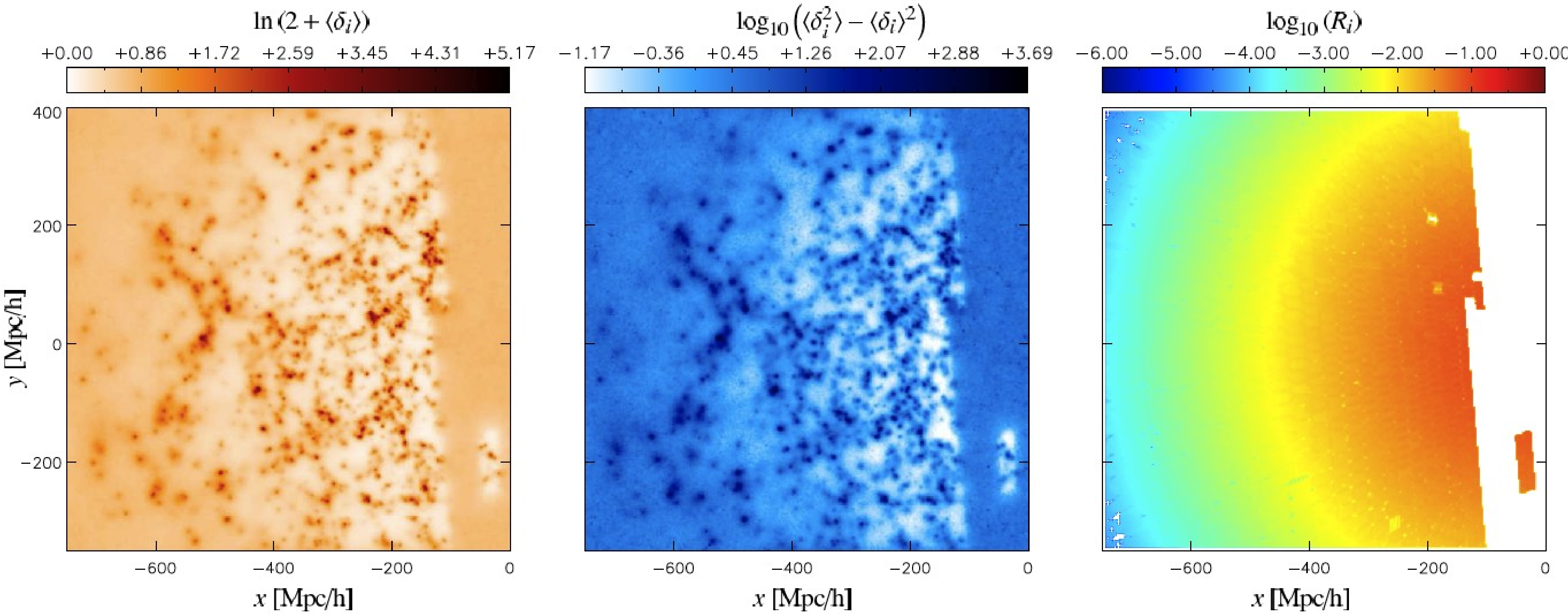
sample + mask



$\ln(2+\delta_i)$
+0.00 +0.86 +1.72 +2.59 +3.45 +4.31 +5.17
$\log_{10}(R_i)$
−6.00 −5.00 −4.00 −3.00 −2.00 −1.00 +0.00
400
200
0
200
y [Mpc/h]
−600 −400 −200 0
x [Mpc/h]
−600 −400 −200 0
x [Mpc/h]

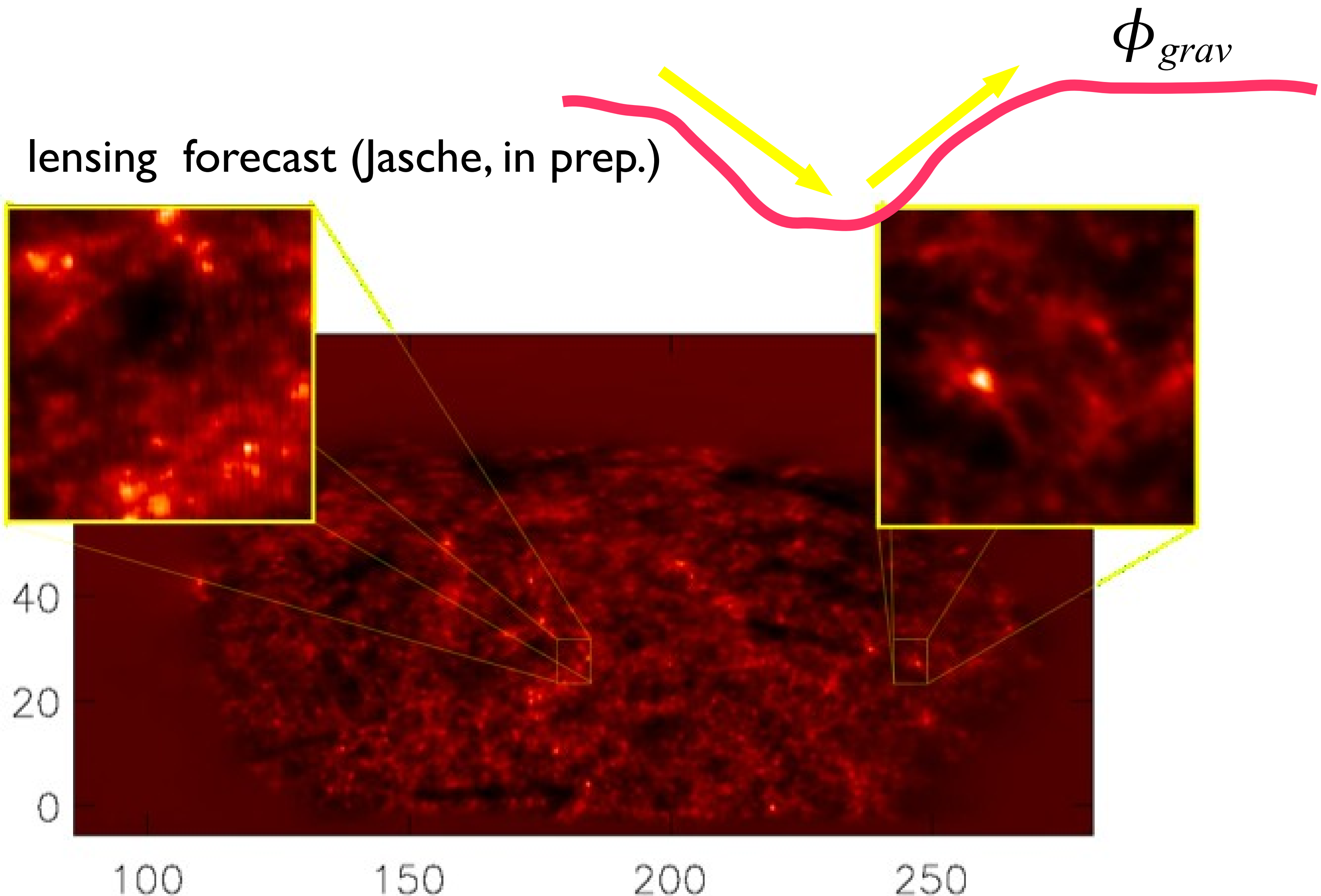
Skialbred.
Kluka
SVNDLEN DINGAFIOR DVNG.
Hekla
Medallan
Þorlaks hafn.

mean, dispersion, mask



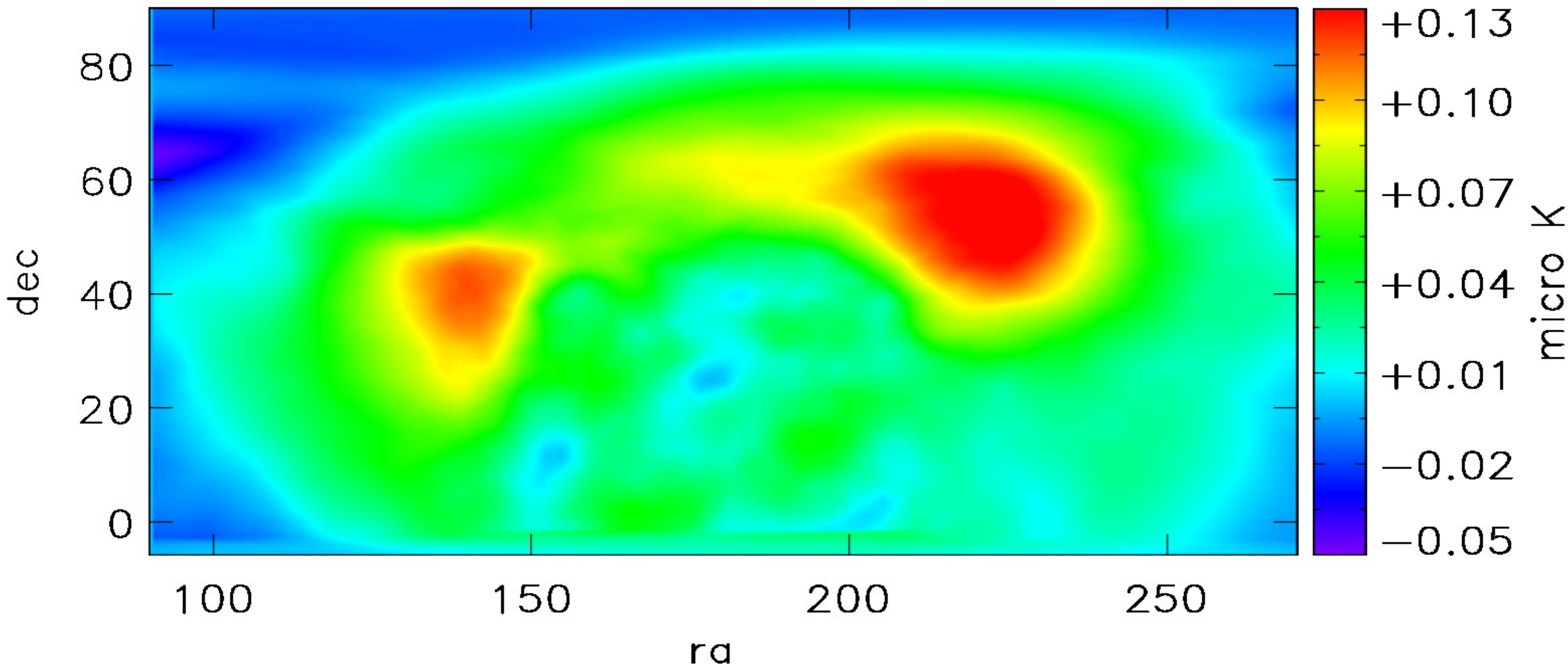
weak gravitational lensing

lensing forecast (Jasche, in prep.)

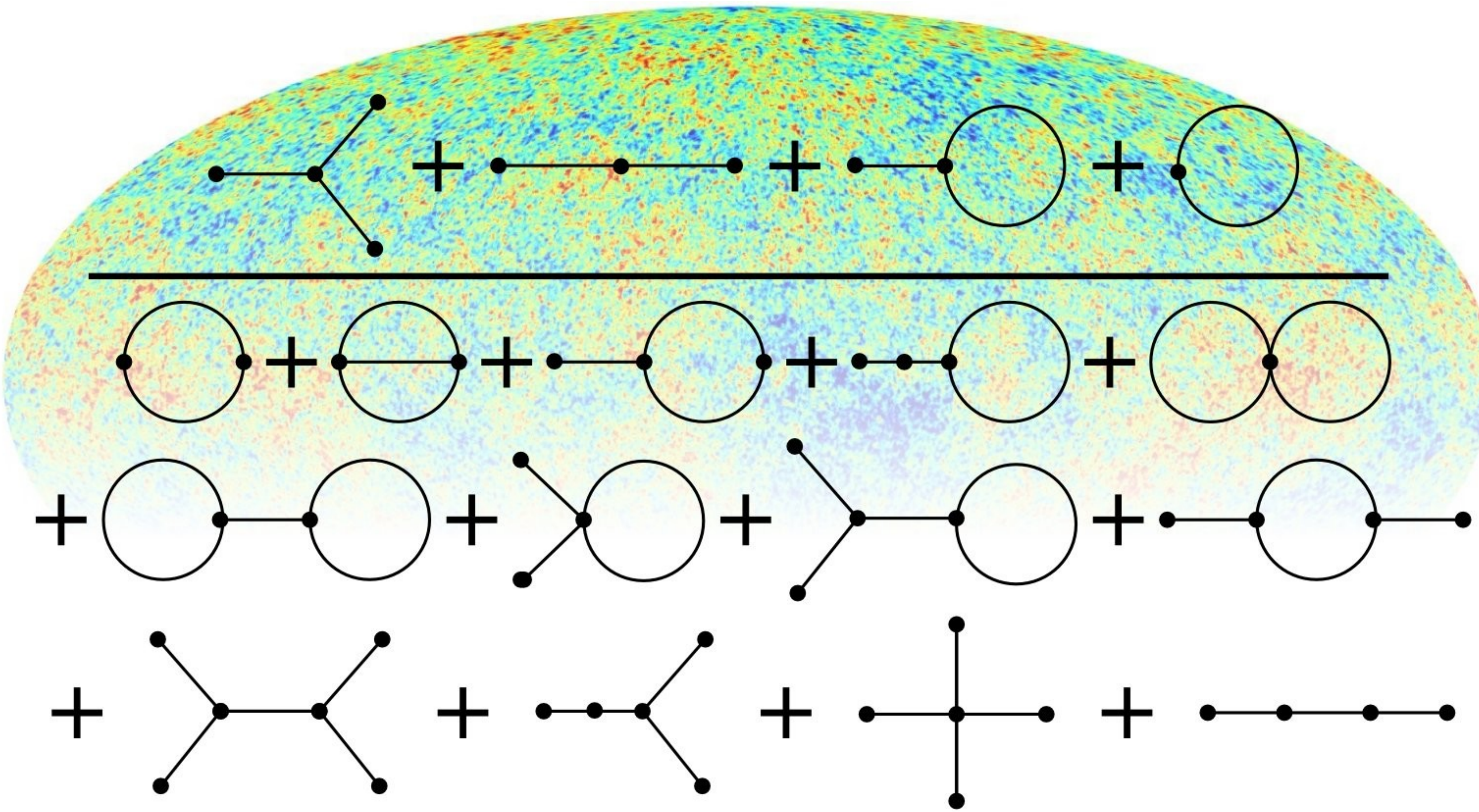


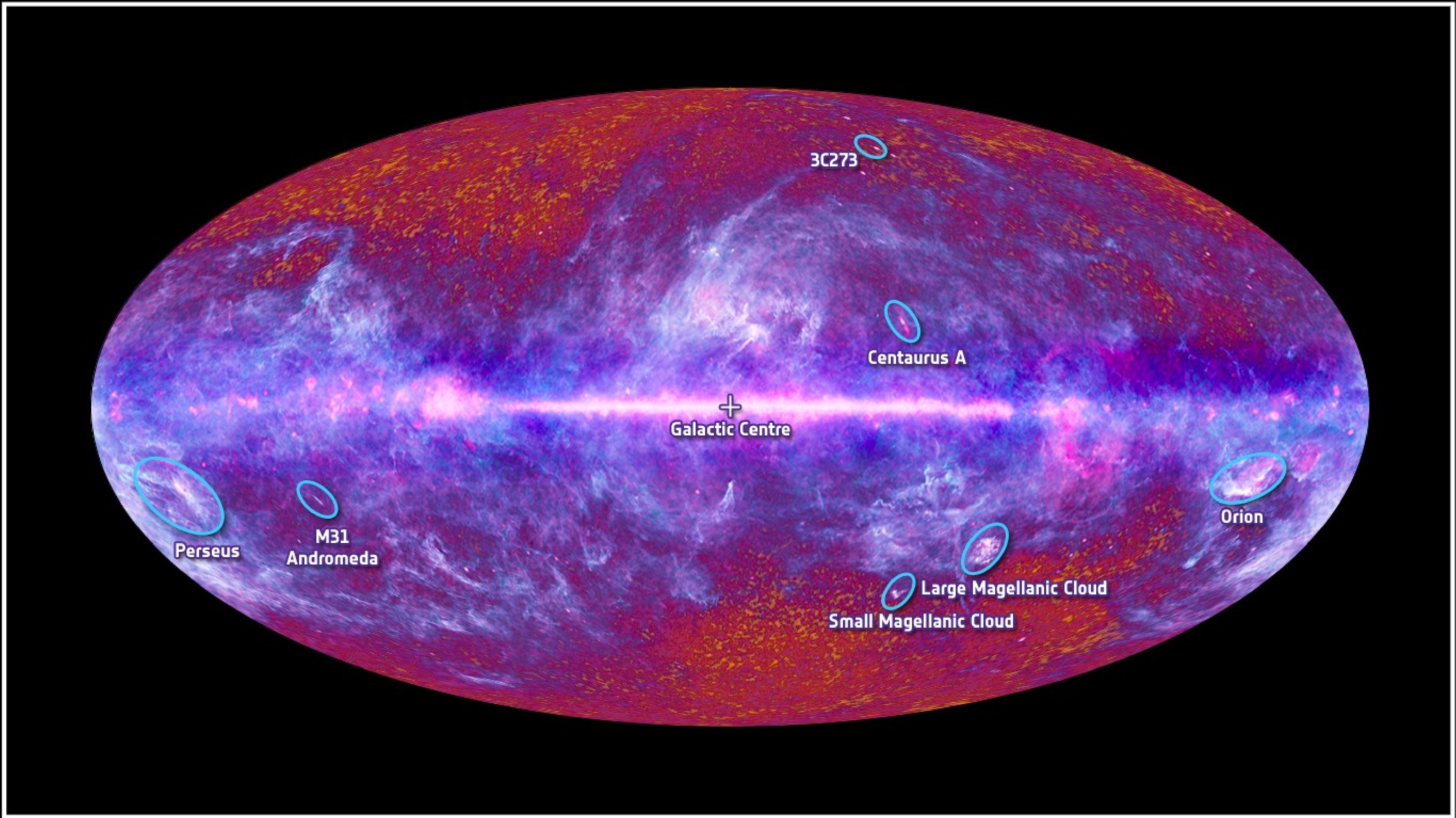
Integrated Sachs Wolfe Effect

ISW forecast (Jasche, in prep.)



Primordial non-Gaussianity detection



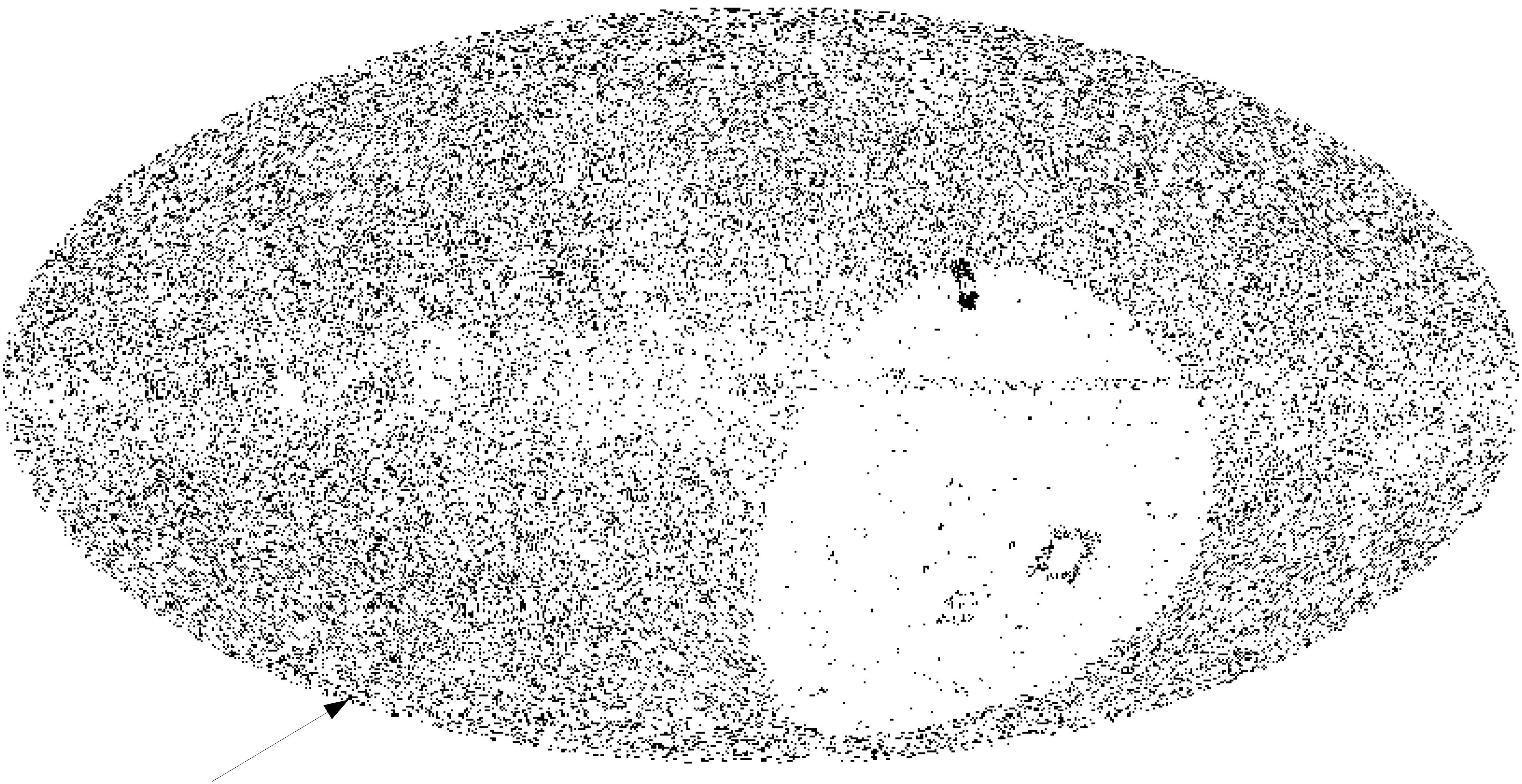


The Planck one-year all-sky survey

esa

Faraday grid

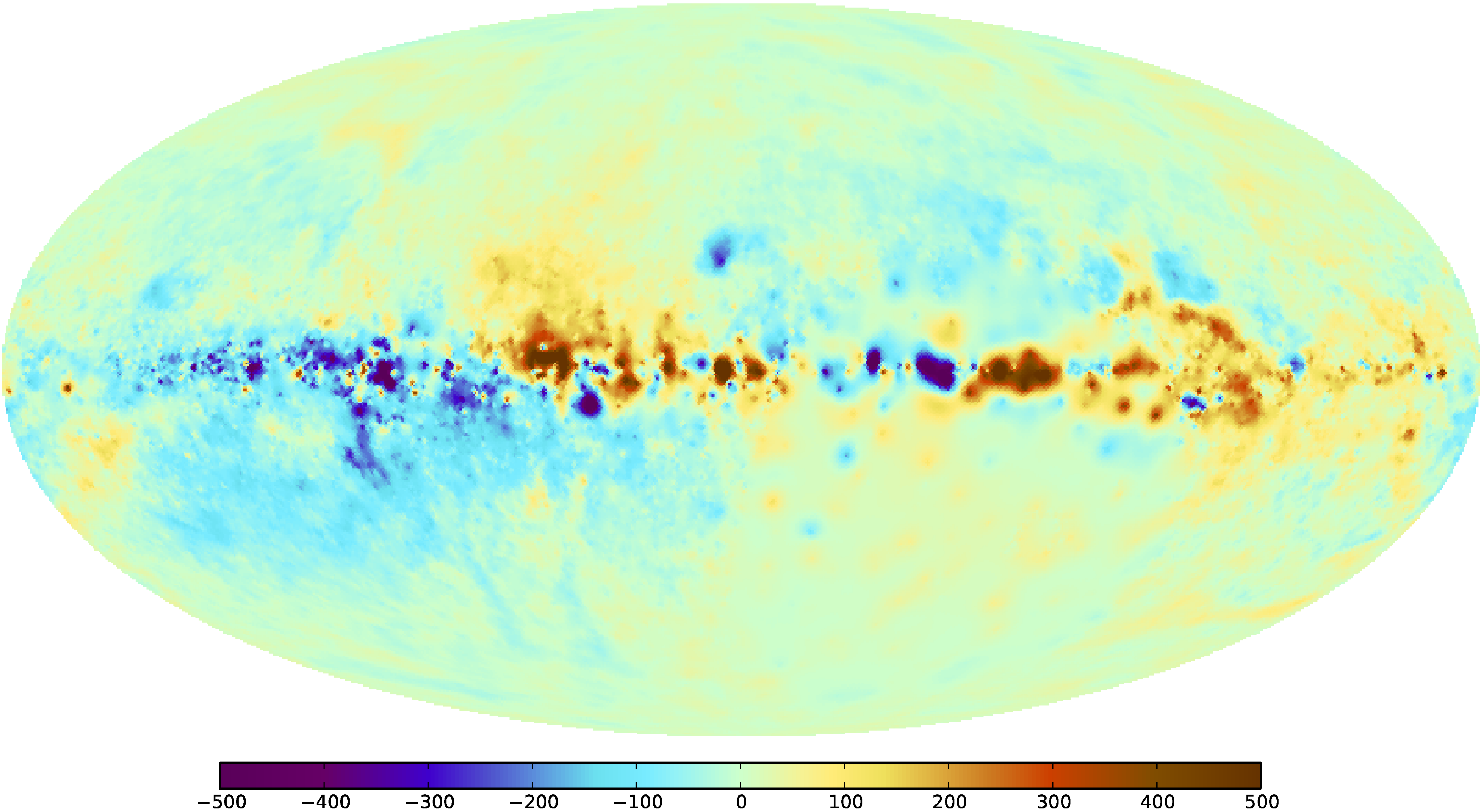
Oppermann et al. (arXiv:1111.6186)



41 330 extragalactic Faraday rotation measurements through the Galaxy

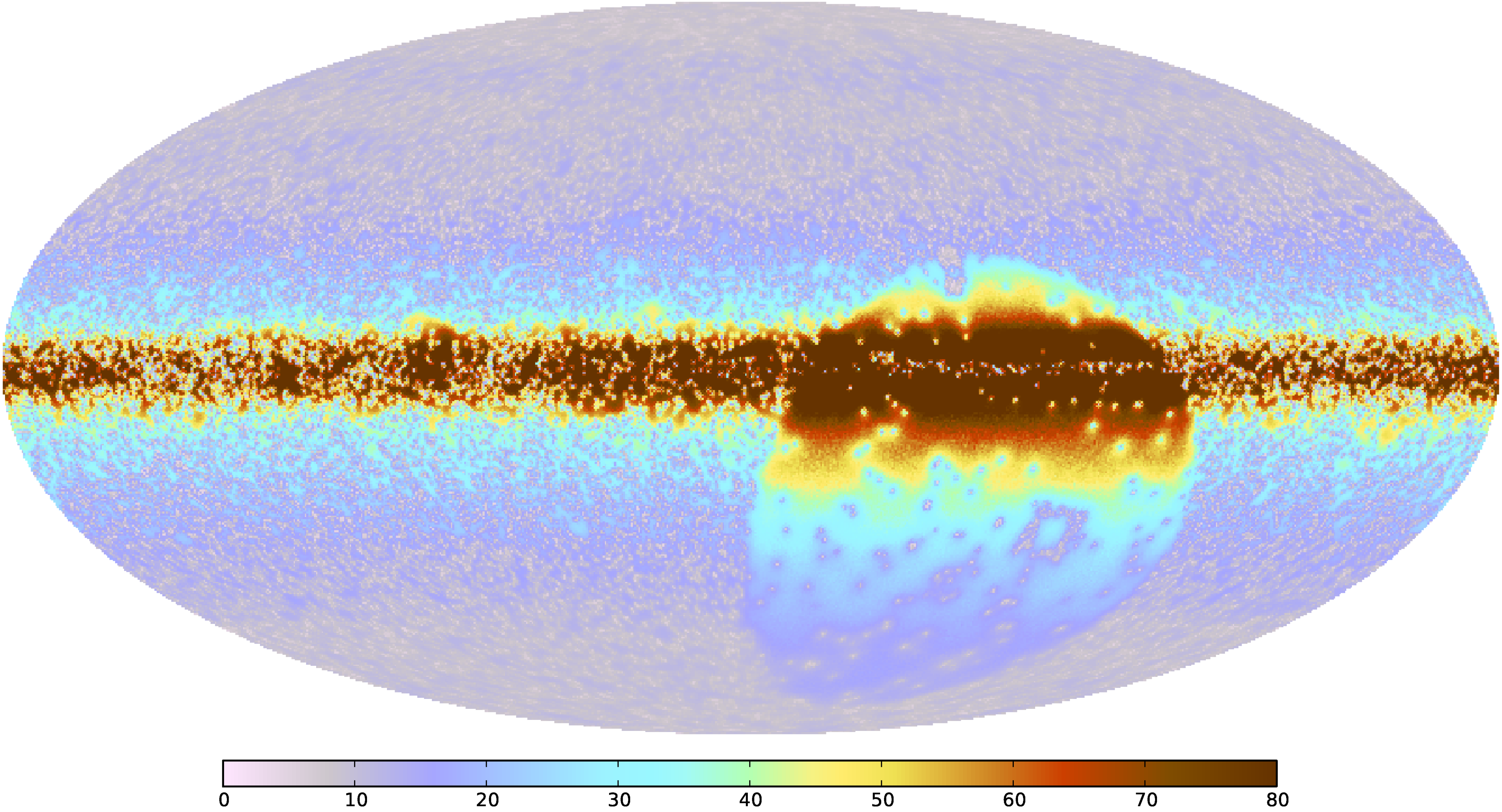
Faraday sky

Oppermann et al. (arXiv:1111.6186)



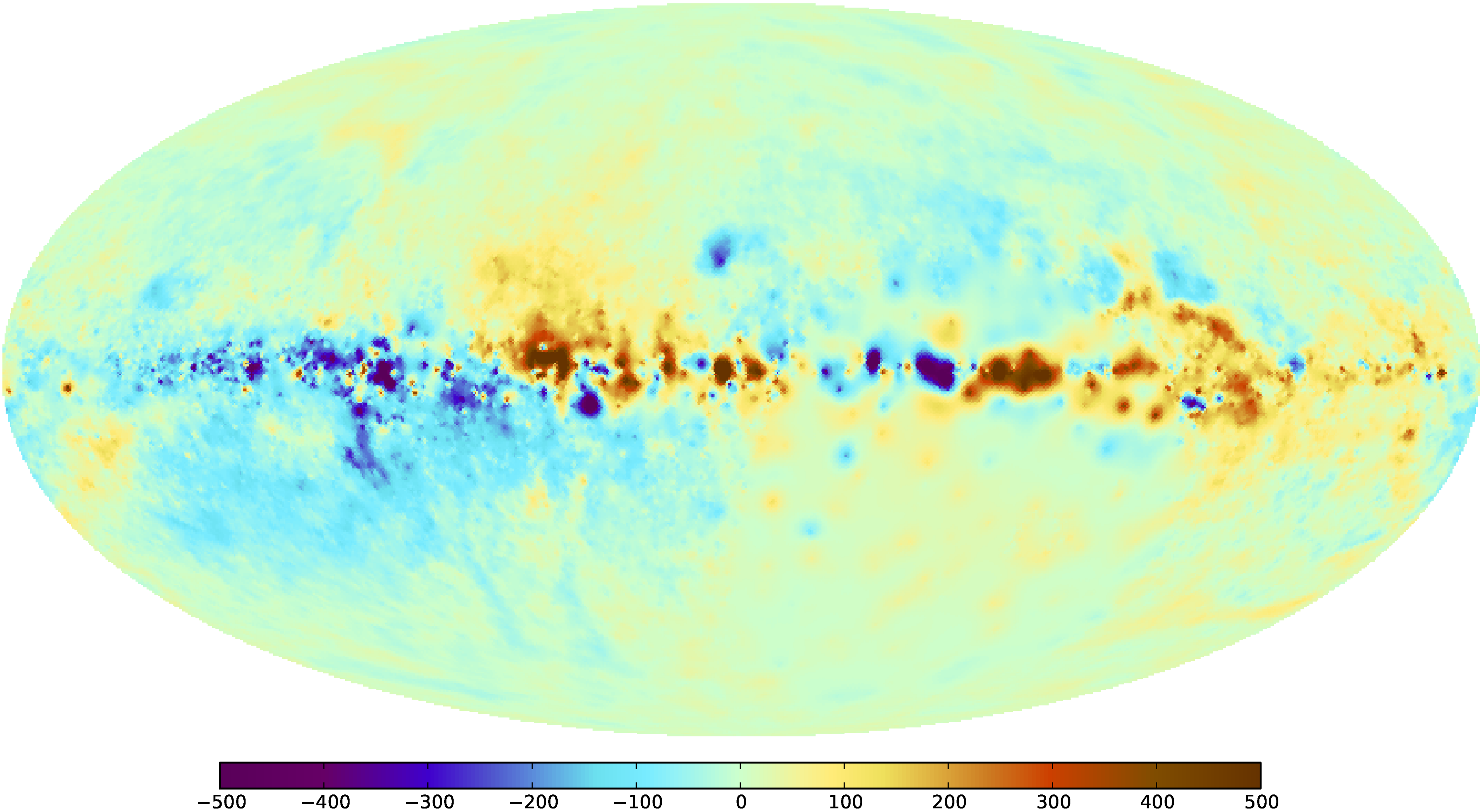
Faraday sky uncertainty

Oppermann et al. (arXiv:1111.6186)



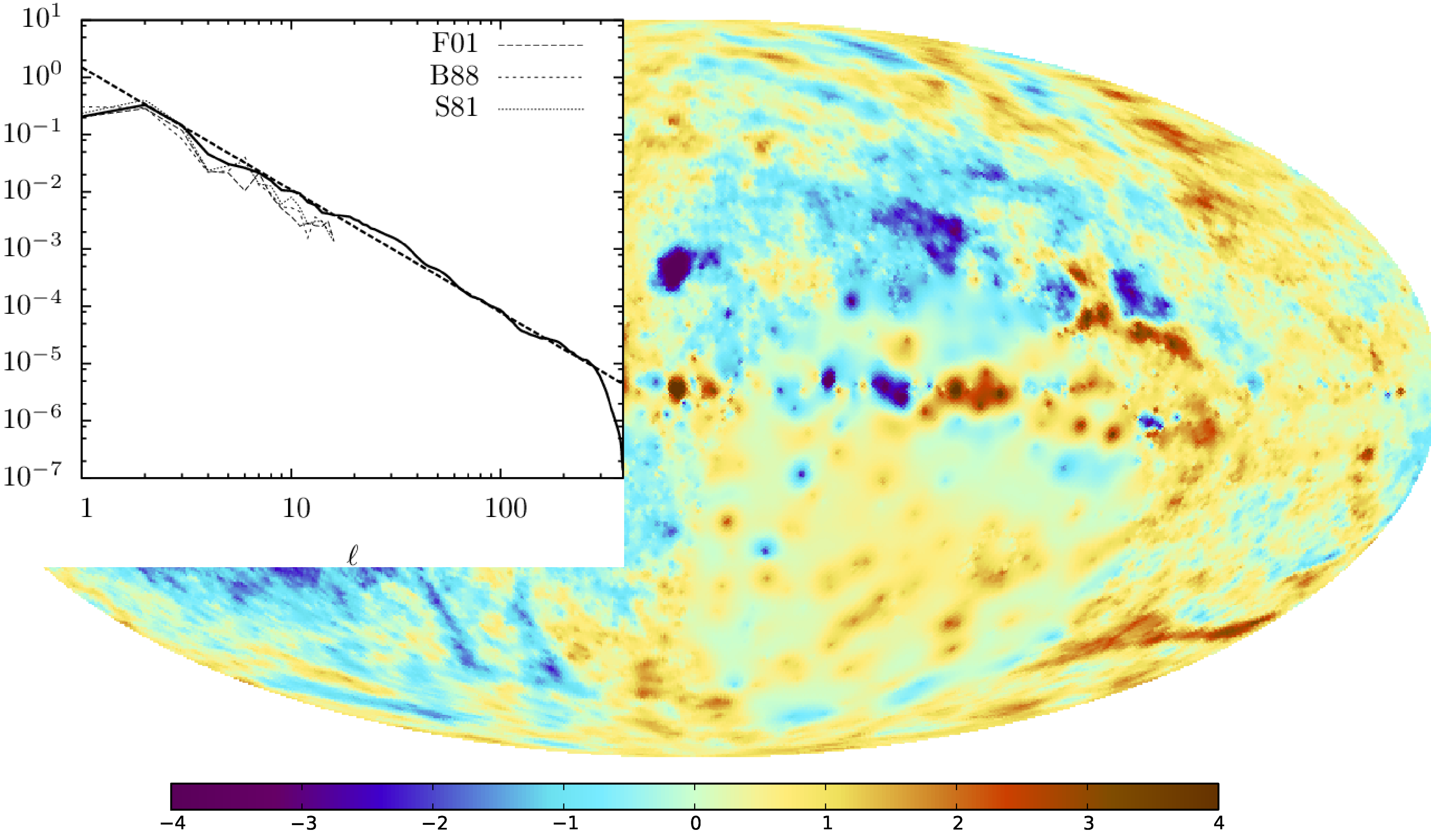
Faraday sky

Oppermann et al. (arXiv:1111.6186)



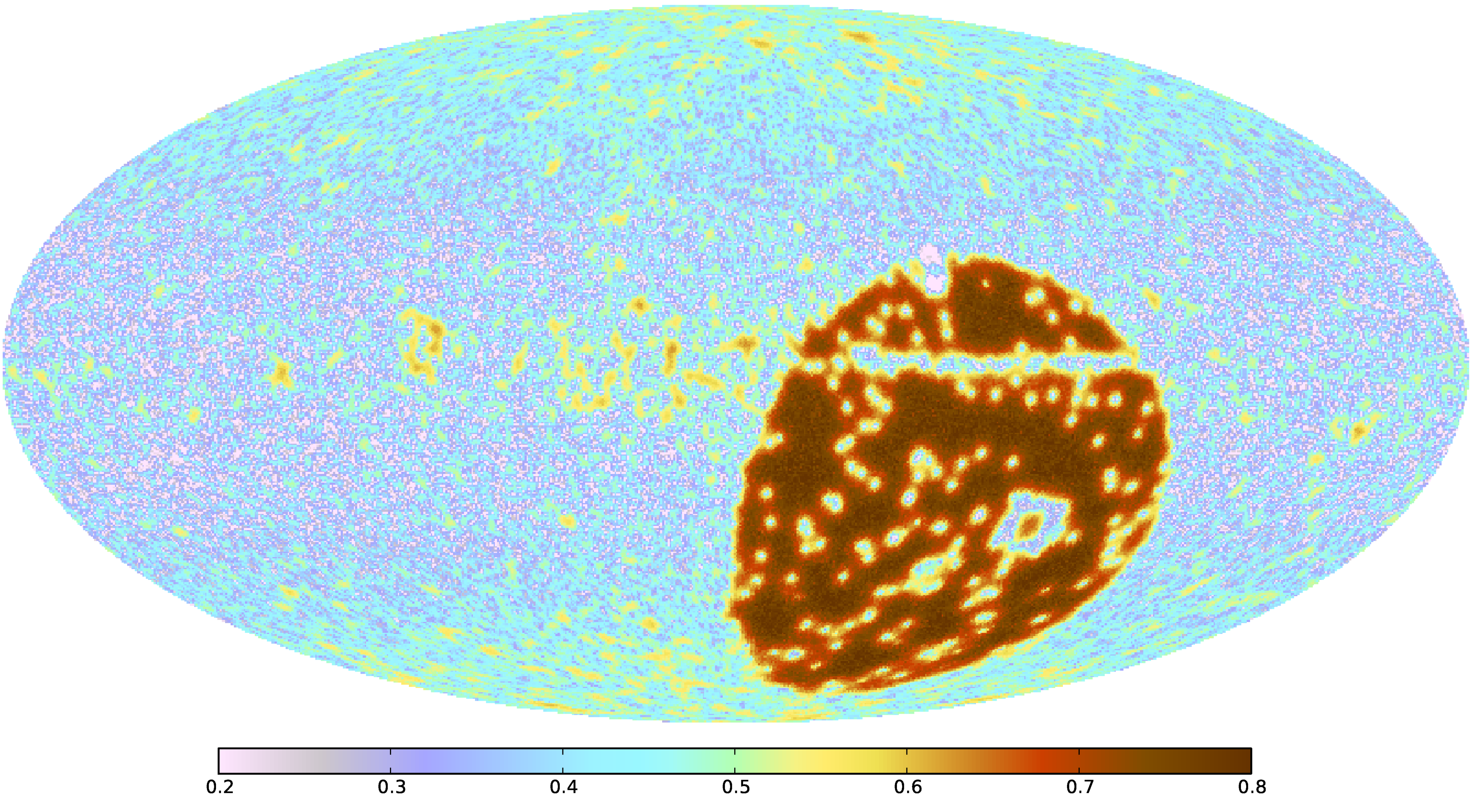
Faraday sky / galactic profile

Oppermann et al. (arXiv:1111.6186)



(Faraday sky / profile) uncertainty

Oppermann et al. (arXiv:1111.6186)



Conclusions

Information field theory (IFT):

- signal field reconstruction requires prior = regularization space is continuous → IFT, which is a statistical field theory
- field theoretical toolbox: free theory = Wiener filter theory, classical theory = max. a posteriori, Feynman diagrams, renormalisation flow, thermodynamical inference, ...

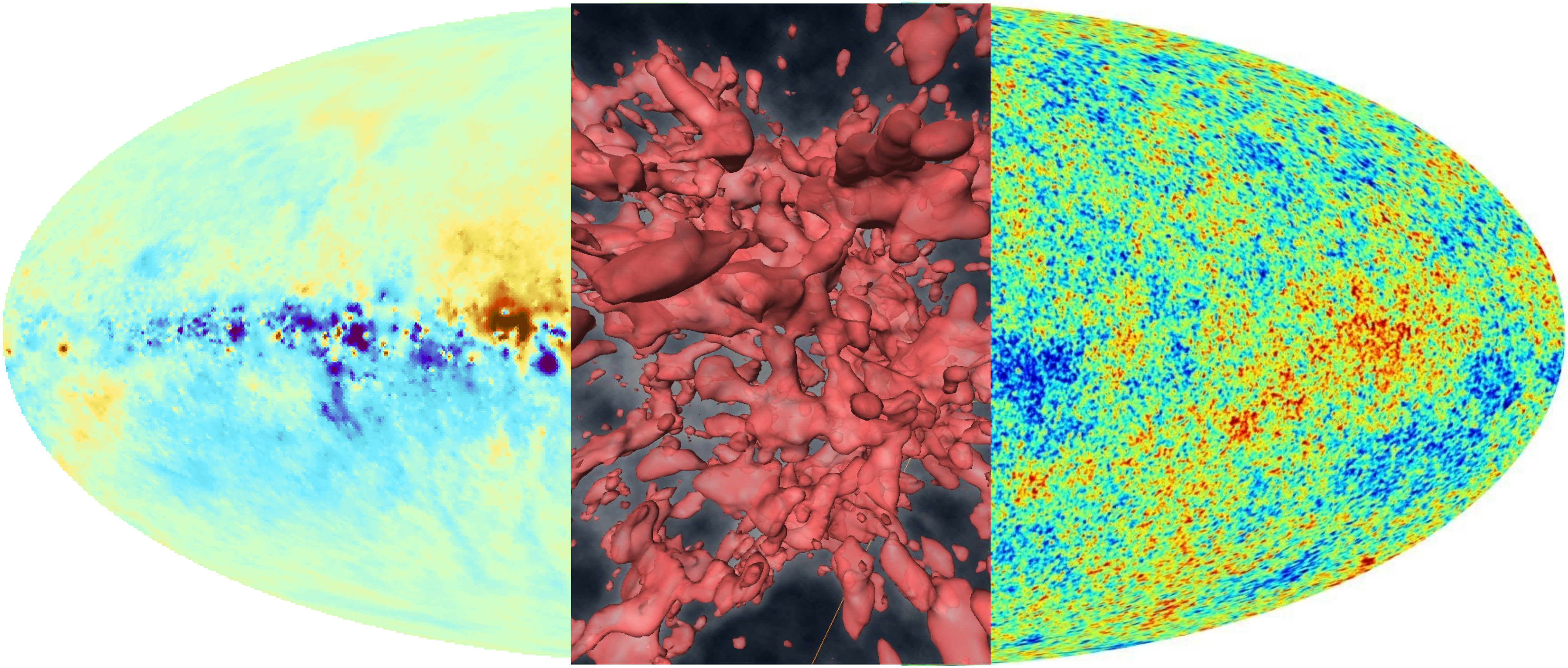
Cosmography & CMB science:

- clear assumptions (stat. homogeneity & isotropy, power, noise) → unique, high fidelity IFT algorithm
- LSS reconstruction, ISW detection, CMB non-Gaussianity

Critical filter methodology:

- reconstruction with unknown signal & noise spectra
- Faraday sky inference

Thank you !



www.mpa-garching.mpg.de/ift