Taming cosmological statistics: Logs, weightings, and Gaussianization

Mark Neyrinck
Johns Hopkins University

Main Collaborators:

- Nuala McCullagh, Bridget Falck, Xin Wang, Miguel Aragón-Calvo, Alex Szalay (JHU), István Szapudi, Julien Carron* (IfA, Hawaii)

*- see his talk, too

Mark Neyrinck, JHU

Outline

- Problem: two-point statistics poorly describe highly non-Gaussian fields
  - Solution: Gaussianization
  - Can be considered as a weighting more by volume than mass
  - Information gain: measurements from simulations
  - A new analytic model of the transformed-density correlation function
    - (detour: spherical-collapse local Lagrangian approximation)

Mark Neyrinck, JHU

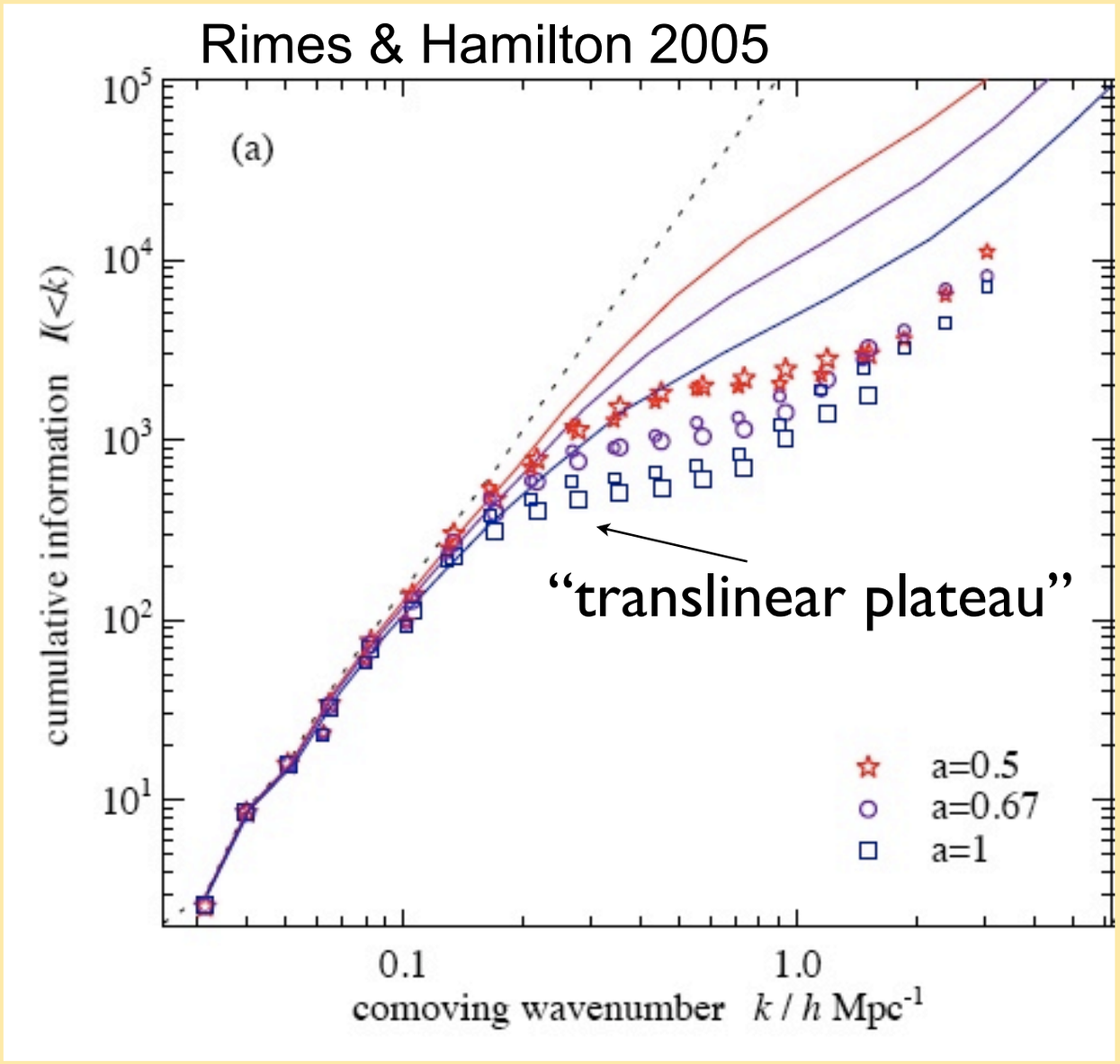
Scales		Stats
Hubble radius	Many Gpc	Cosmological parameters
Linear, Gaussian scales	x > ~100 Mpc/h k < ~0.1 h/Mpc	Power spectrum
Nonlinear scales	x < ~50 Mpc/h	Morphological statistics?

Why not use the power spectrum on small scales?

Fisher-information loss (small-scale covariance)
⇔ Fourier-mode coupling, loss in degrees of freedom

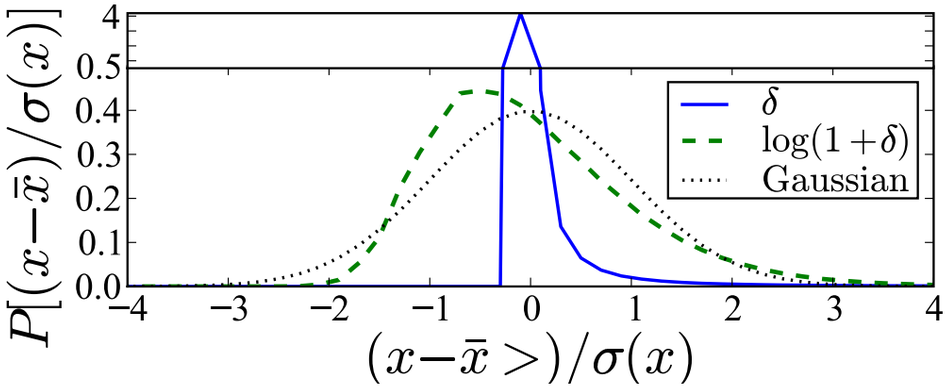
independent Fourier modes

large, linear scales



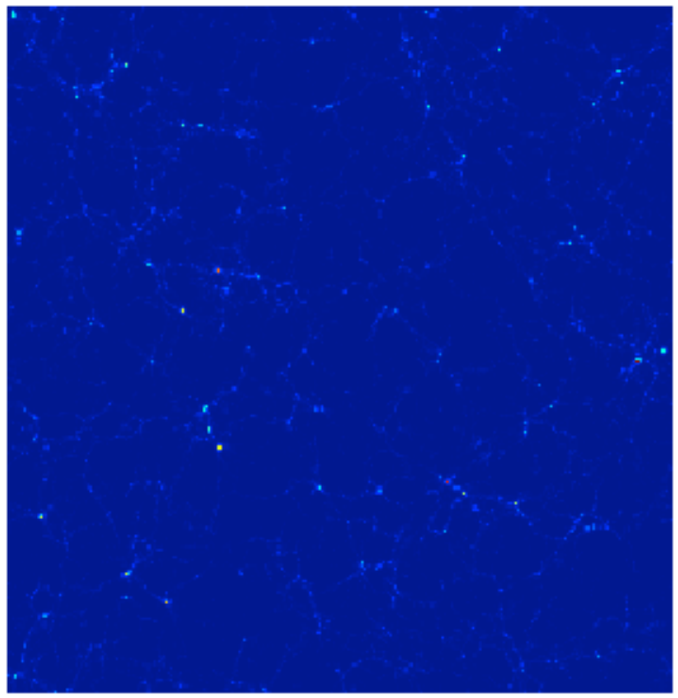
small, halo scales

Sharp peaks overweighted in mass-weighting -- try a ~volume-weighting.

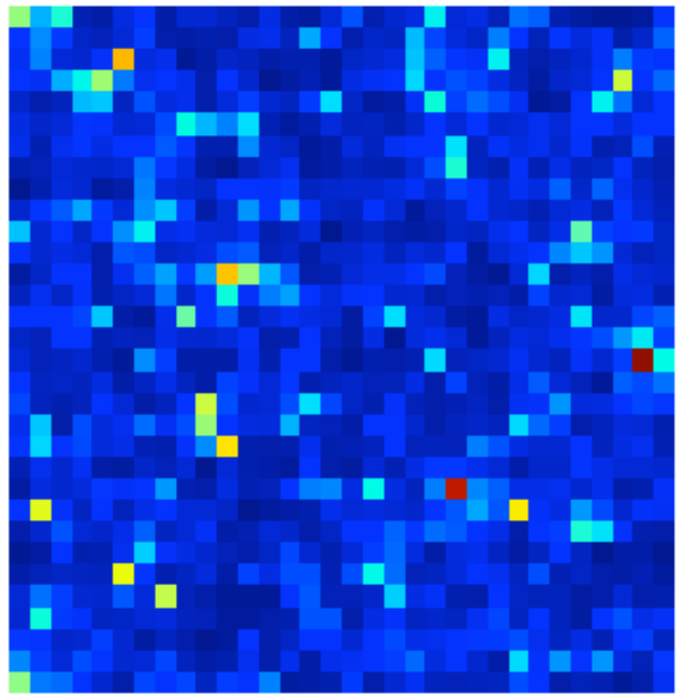


$$\delta(\mathbf{x}) \equiv \frac{\rho(\mathbf{x})}{\bar{\rho}} - 1$$

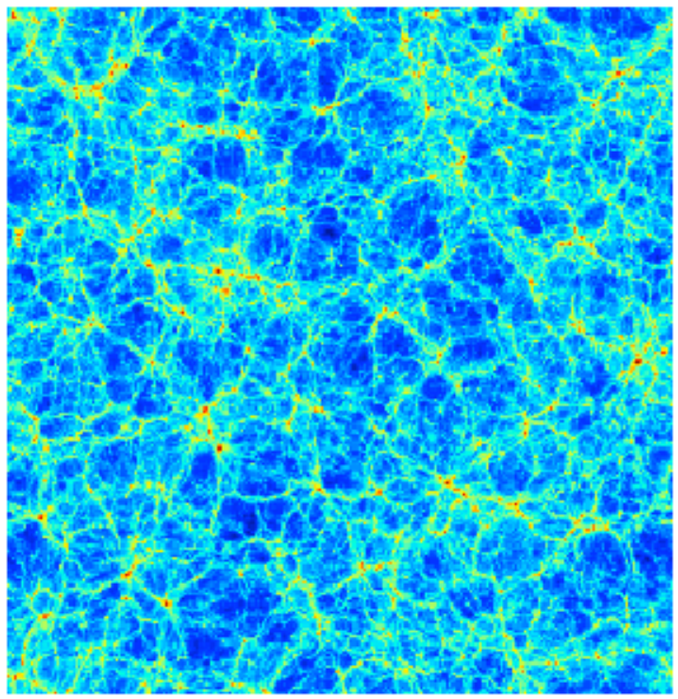
Overdensity δ



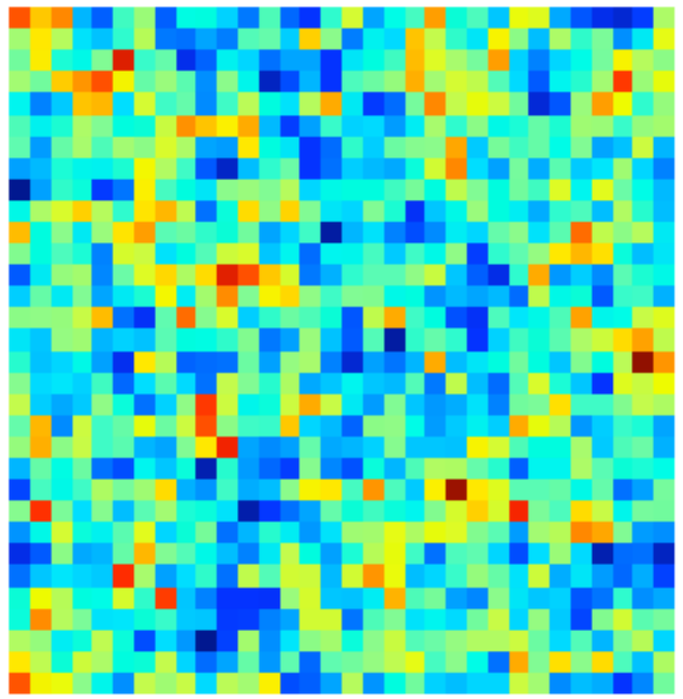
What we analyze



$\ln(1+\delta)$



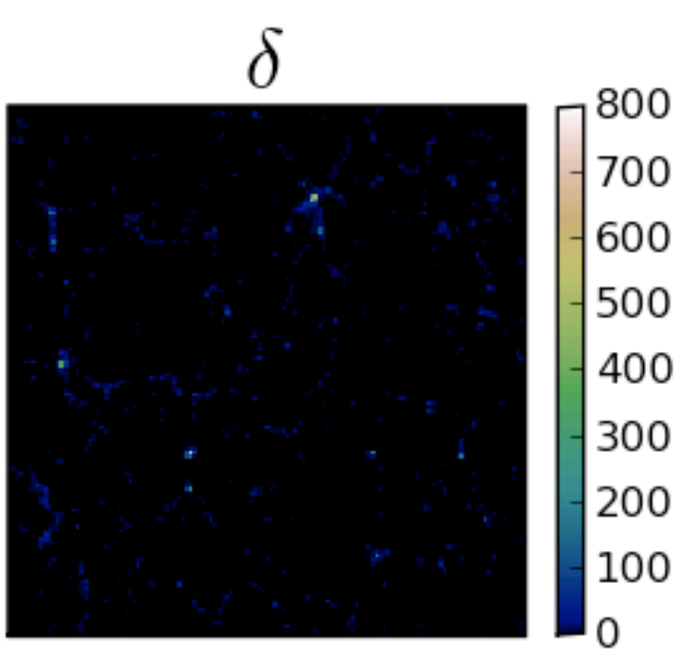
$\ln(1+\delta)$ (low−res)

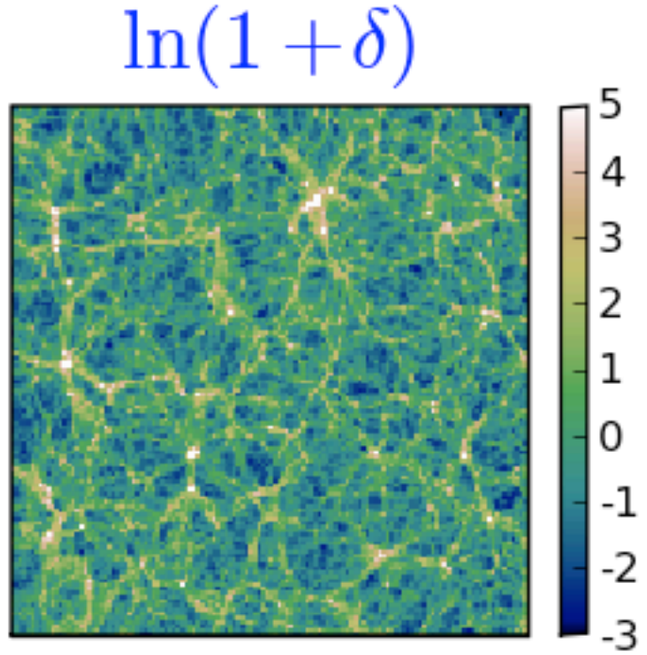


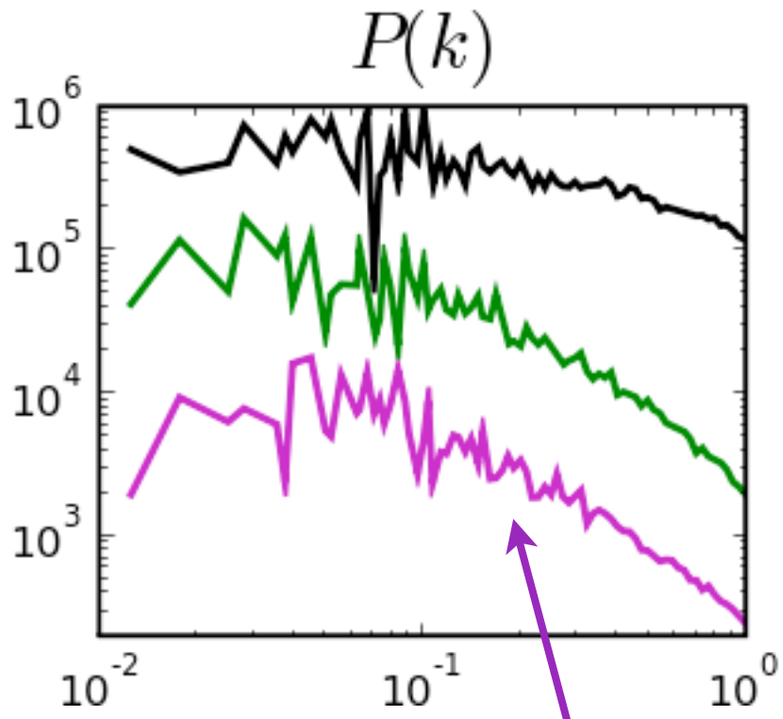
Millennium Sim. (Springel et al. 2005)

Mark Neyrinck, JHU

Sharp peaks overweighted in mass-weighting -- try a ~volume-weighting.

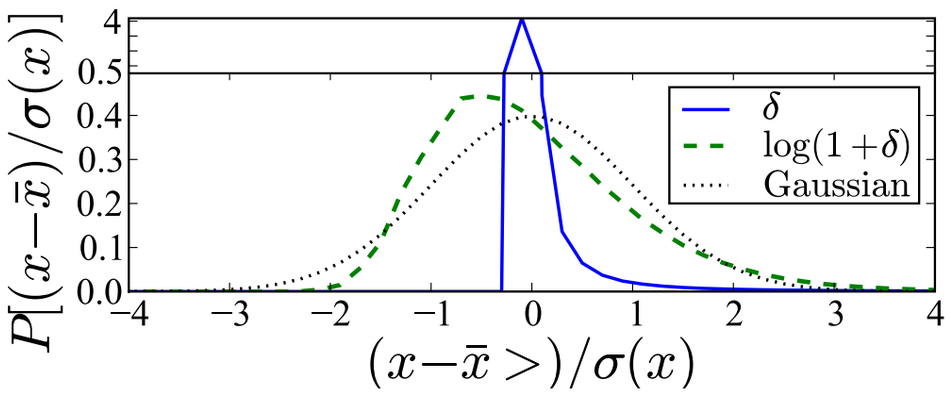






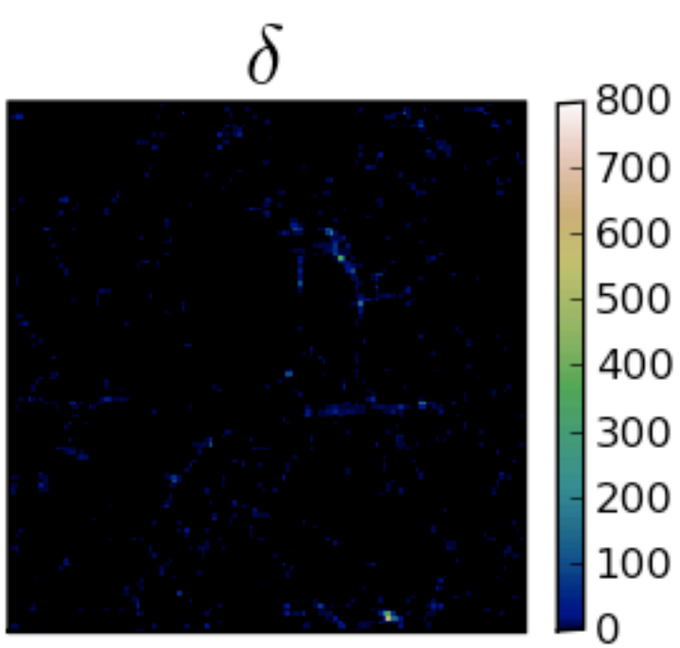
Initial power spectrum

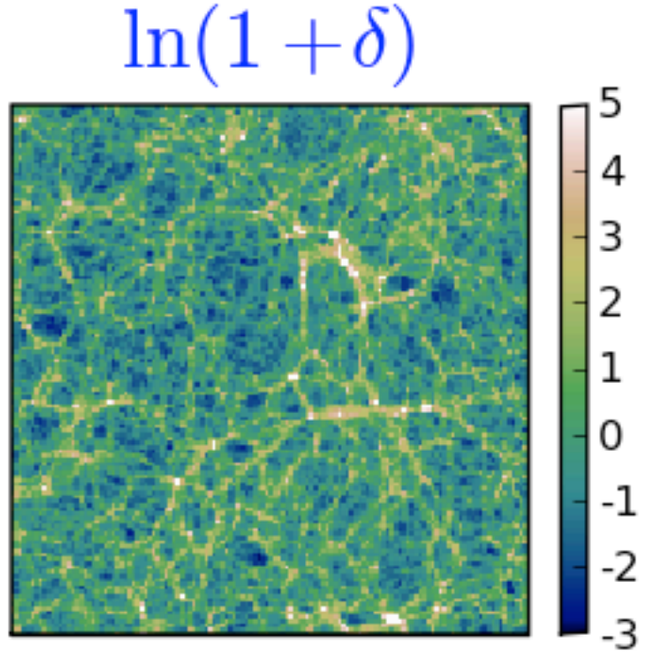
Substantial dependence of the power spectrum on the 1-pt PDF. Suggests decoupling the 1-pt from 2-pt statistics.

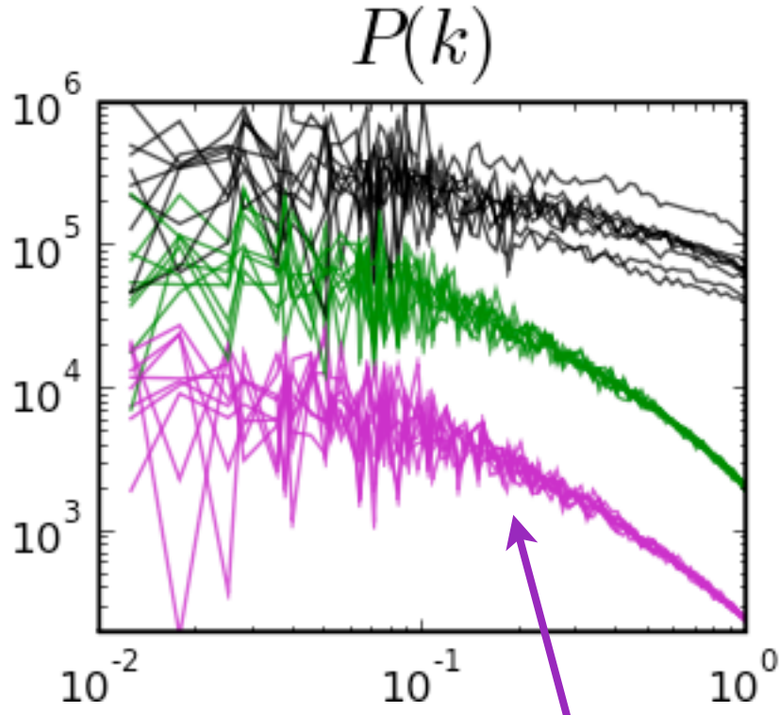


Mark Neyrinck, JHU

Sharp peaks overweighted in mass-weighting -- try a ~volume-weighting.

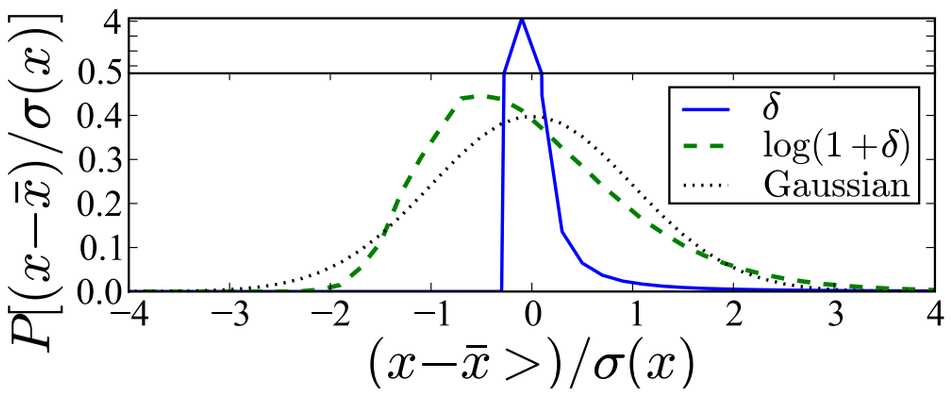






Initial power spectrum

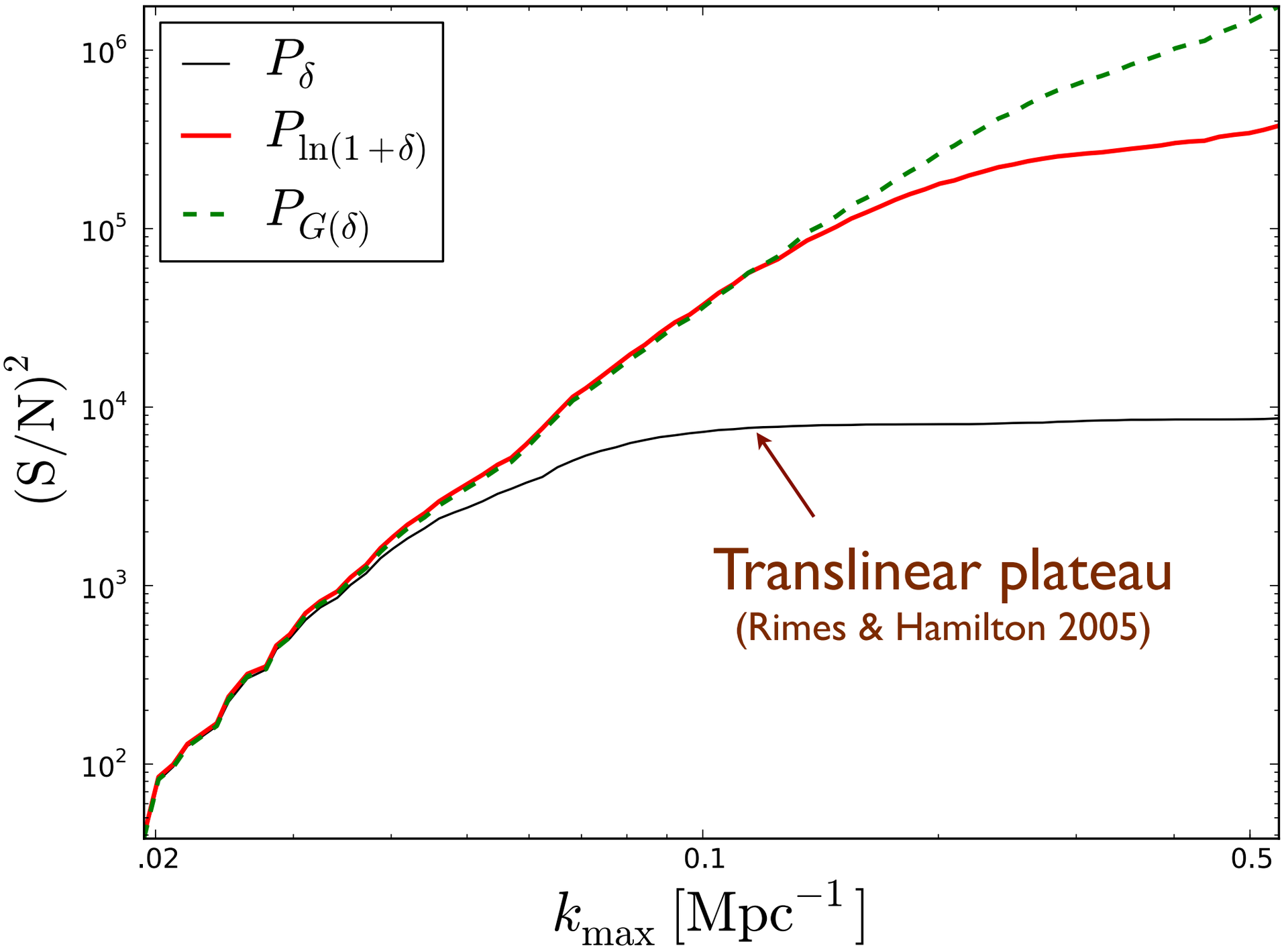
Substantial dependence of the power spectrum on the 1-pt PDF. Suggests decoupling the 1-pt from 2-pt statistics.



Mark Neyrinck, JHU

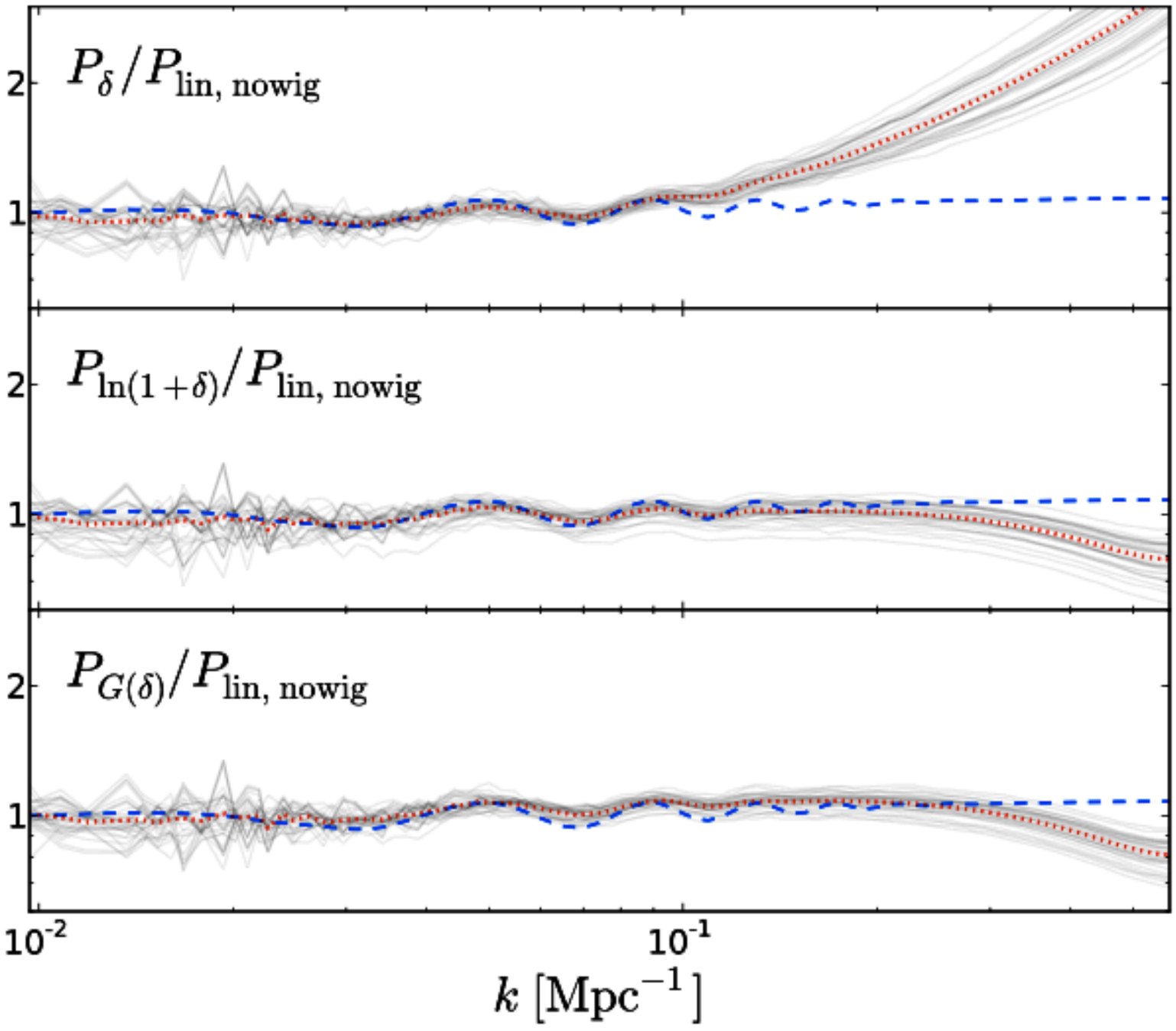
Intrinsic Fisher information greatly enhanced

statistically independent Fourier modes as a function of resolution (MN et al. 2009, 2011)

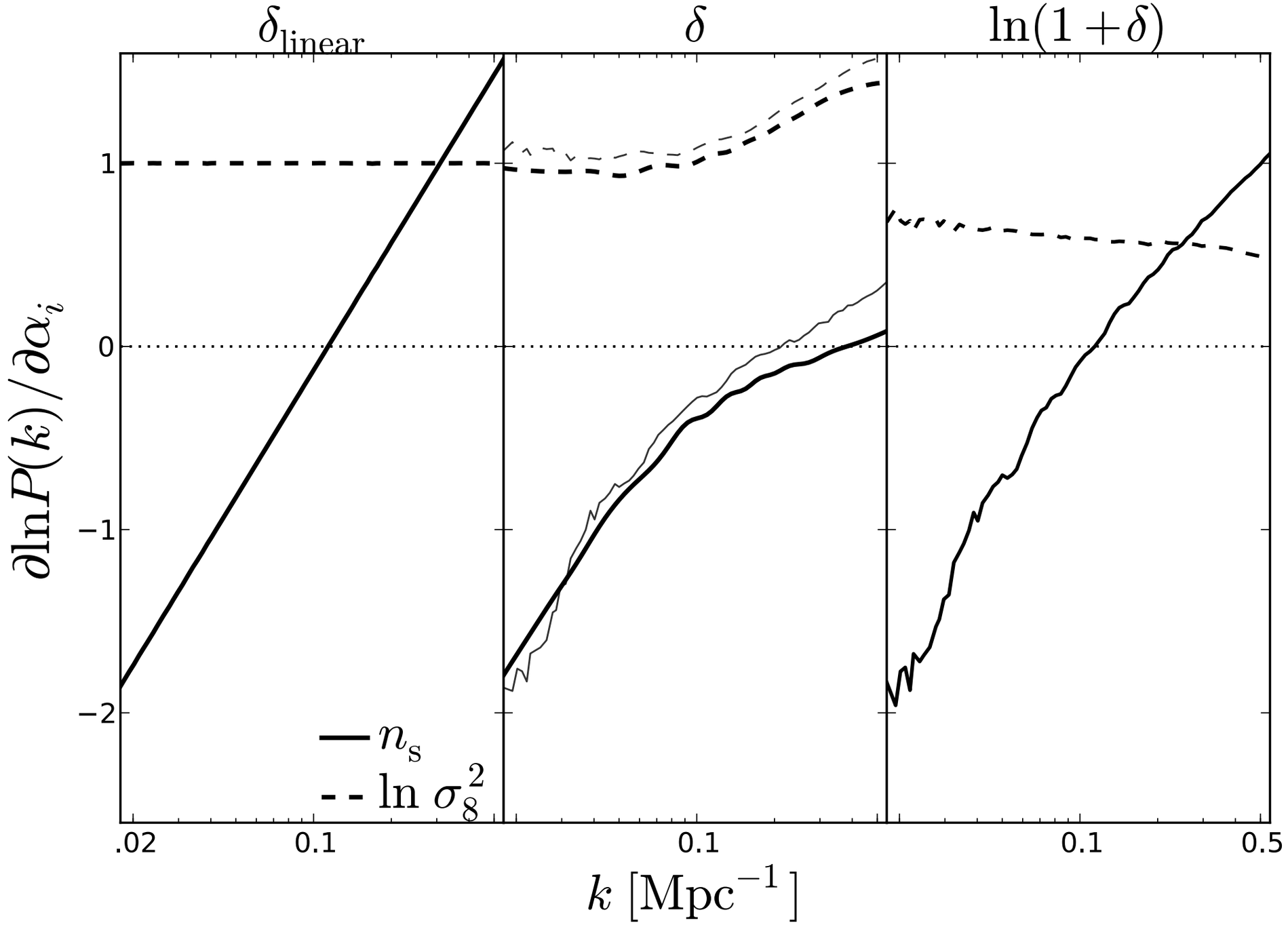


A fairer weighting also straightens the linear → nonlinear transfer function

37 Coyote Universe simulations shown, spanning a range of concordance cosmological params



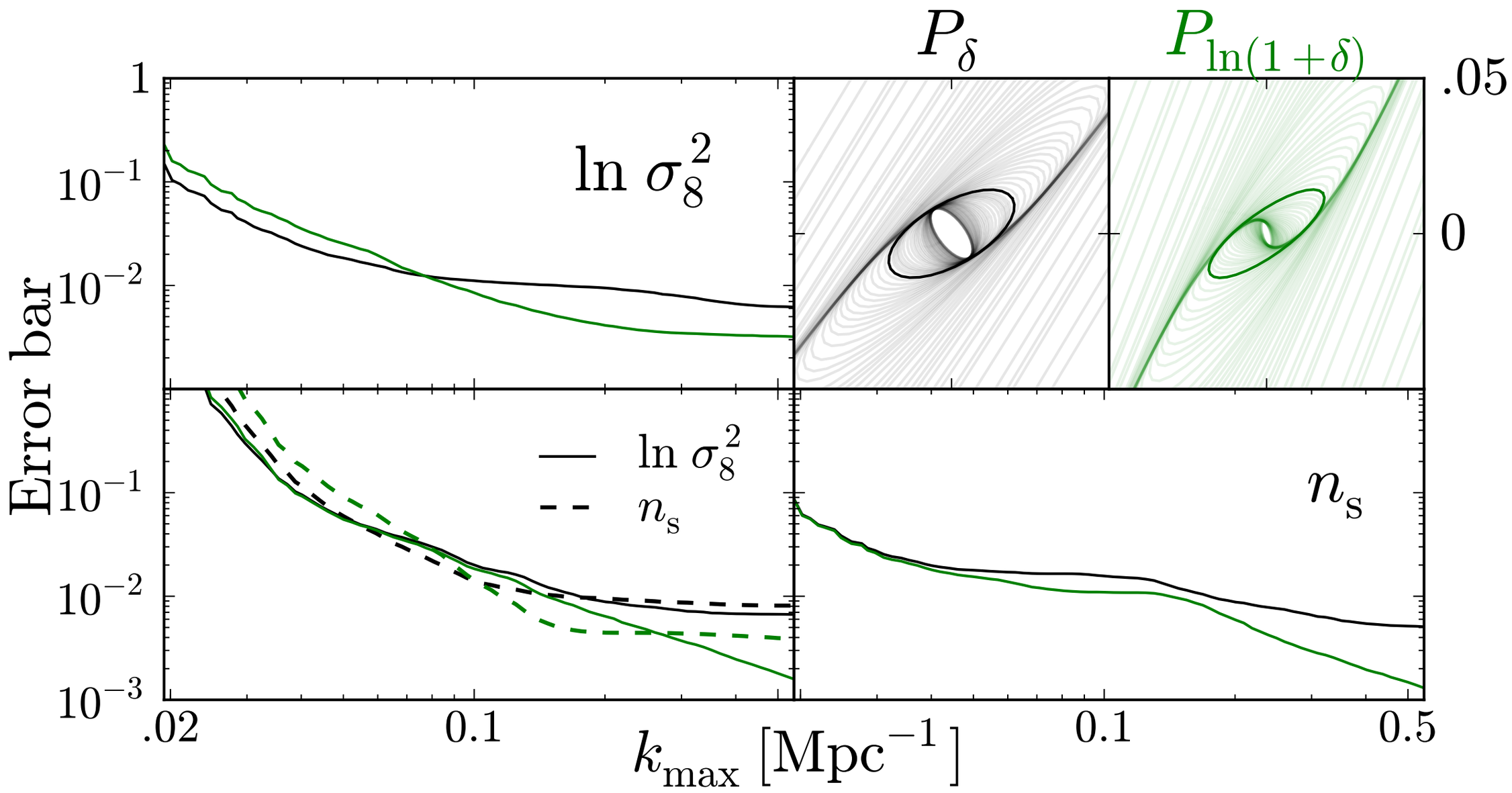
Dependence on parameters



Measured from the Coyote Universe simulations (Heitmann et al, 2008, 2009, 2010).

Mark Neyrinck, JHU

$P_{\ln(1+\delta)}$ significantly more cosmologically informative.



Mark Neyrinck, JHU

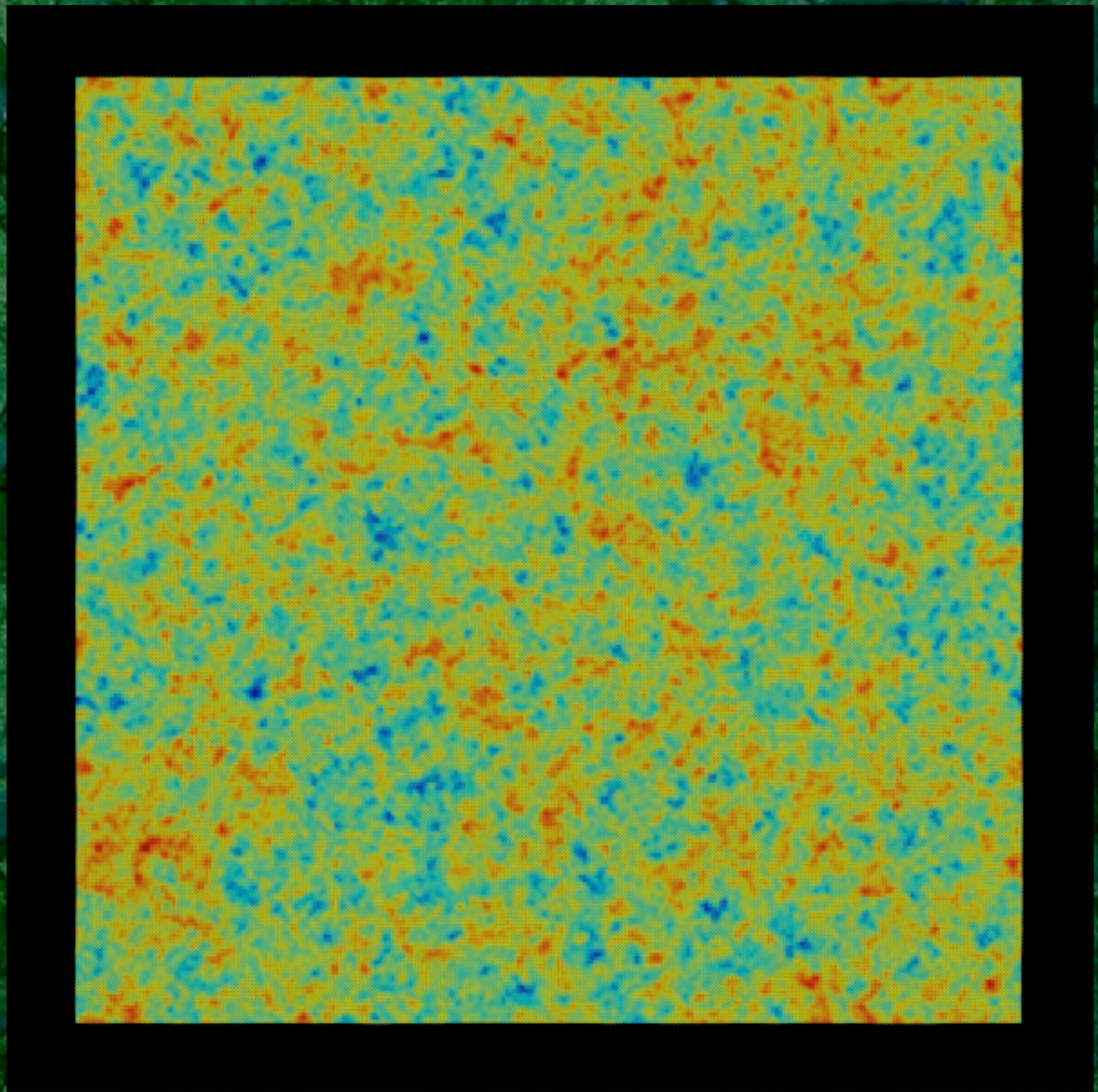
What if we move densities, not mass, around? Violates mass conservation, but should give a "better" weighting.
(i.e. statistically more friendly, and with more fidelity to initial conditions)

* Decouples mass-element bulk motion from gravitational enhancement

Mark Neyrinck, JHU

What if we move densities, not mass, around? Violates mass conservation, but should give a "better" weighting.
(i.e. statistically more friendly, and with more fidelity to initial conditions)

* Decouples mass-element bulk motion from gravitational enhancement



Mark Neyrinck, JHU

How to model analytically?
First, interlude about Bernardeau's (1992, 1994) spherical-collapse approximation, true in the limit $\Omega_\Lambda=0$, $\Omega_m \to 0$,

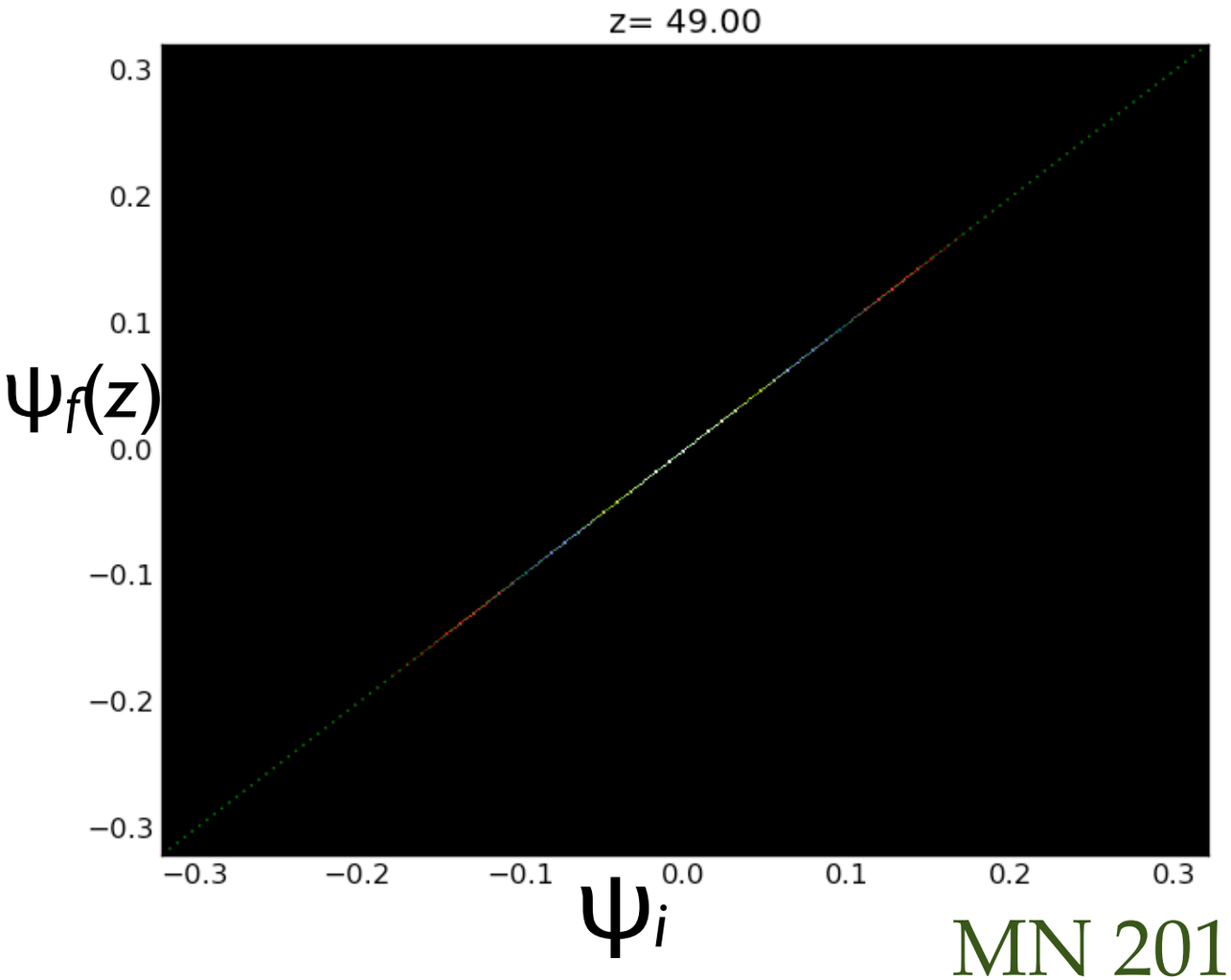
$$V(t) = V_0 \left(1 - \frac{2}{3}\delta_{\rm lin}\right)^{3/2}$$

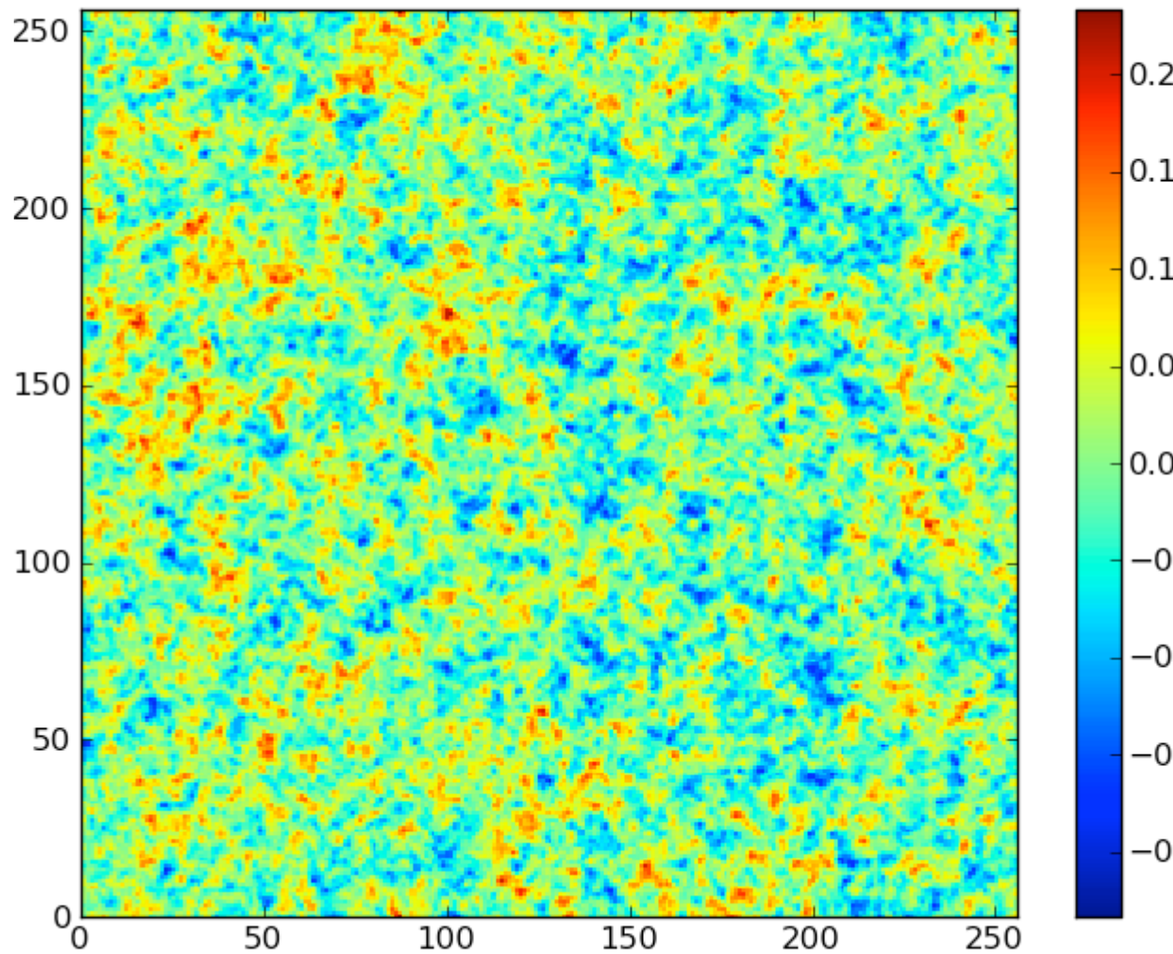
Allows a good model of the Lagrangian divergence of the displacement field, $\psi(\mathbf{q}) \equiv \boldsymbol{\nabla}_q \cdot \boldsymbol{\Psi}(\mathbf{q})$ (Mohayaee et al. 2006, Neyrinck 2012):

$$\psi_{\rm sc} = 3\left[\left(1 + \frac{2}{3}\psi_{\rm lin}\right)^{1/2} - 1\right]$$

$$\psi(\mathbf{q}) \equiv \nabla_q \cdot \mathbf{\Psi}(\mathbf{q})$$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



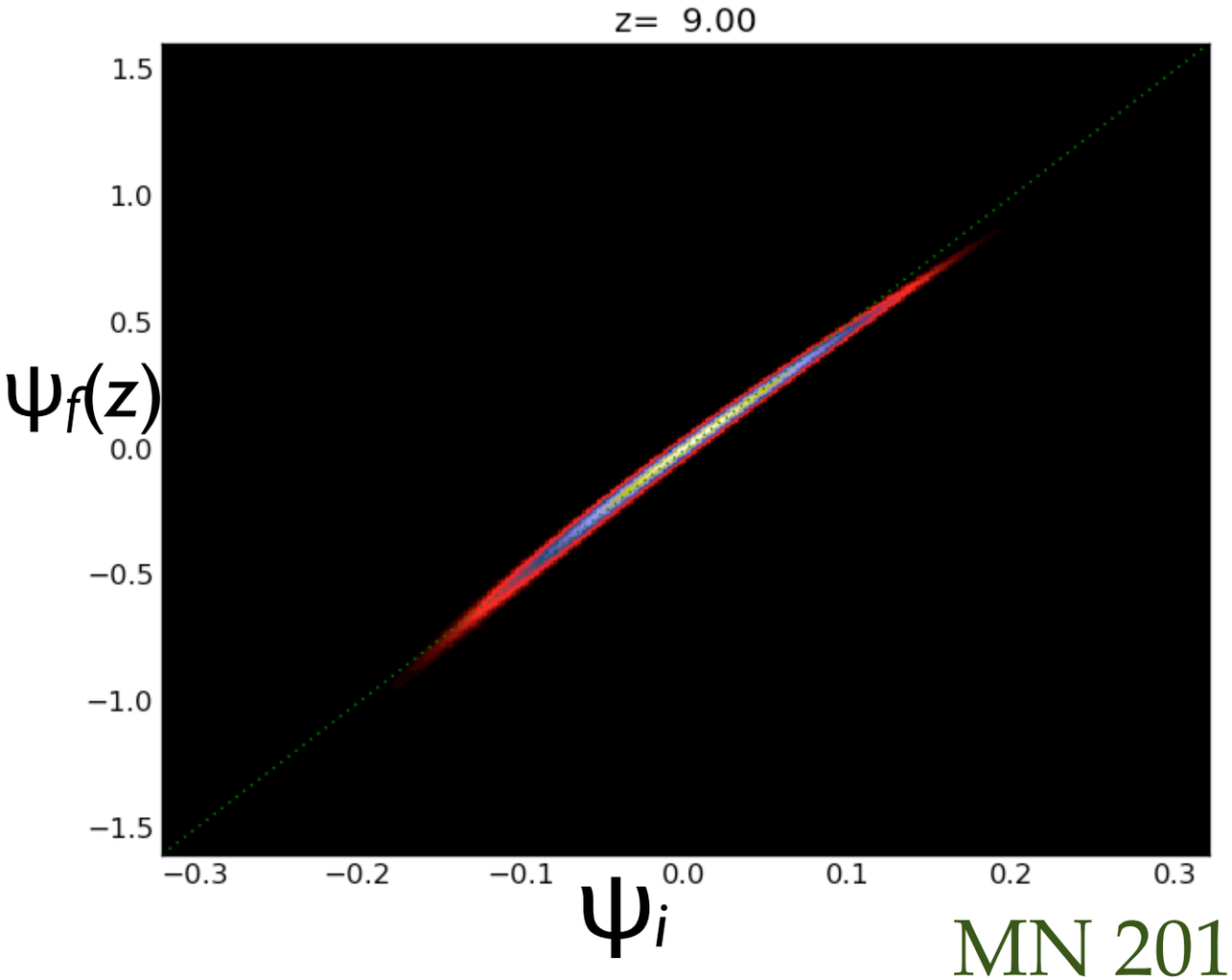


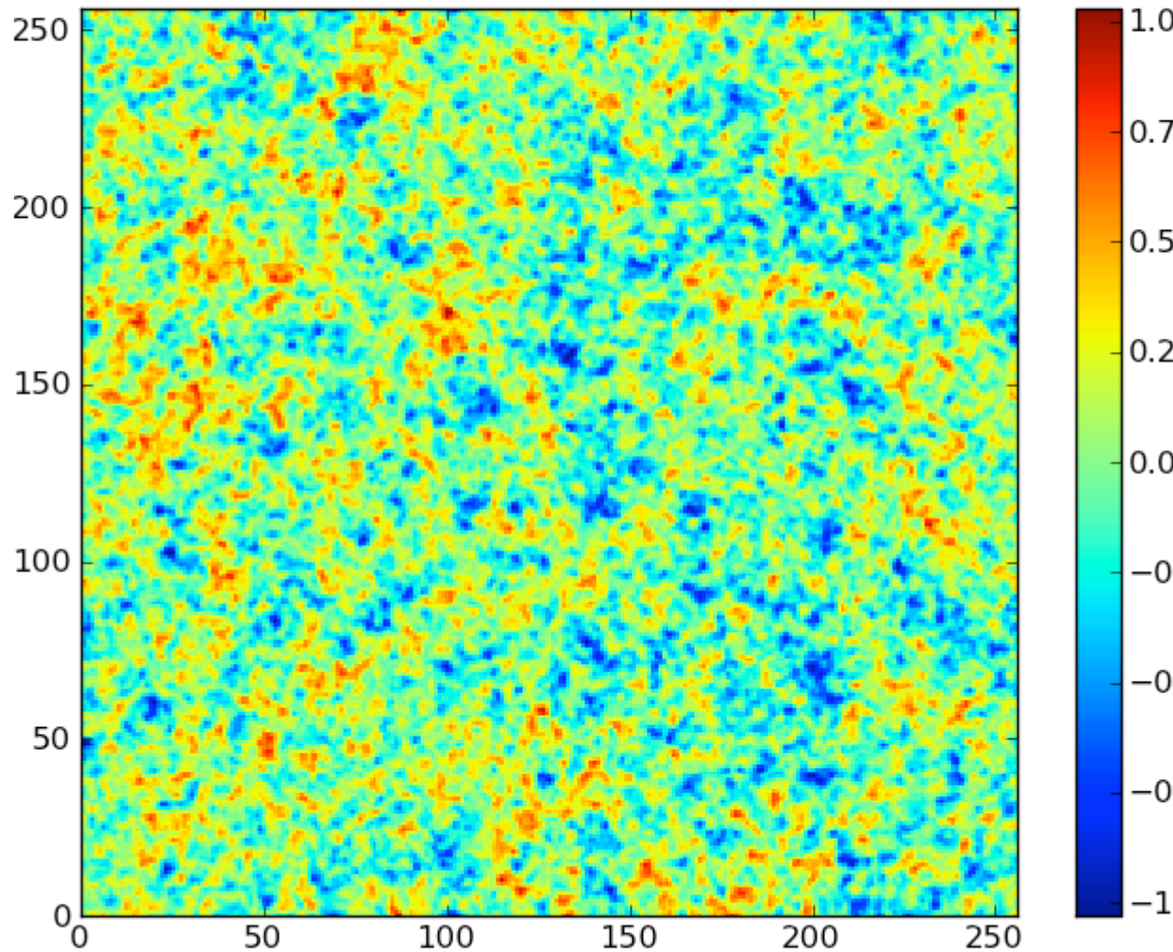
MN 2012, arXiv:1204.1326

Mark Neyrinck, JHU

$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



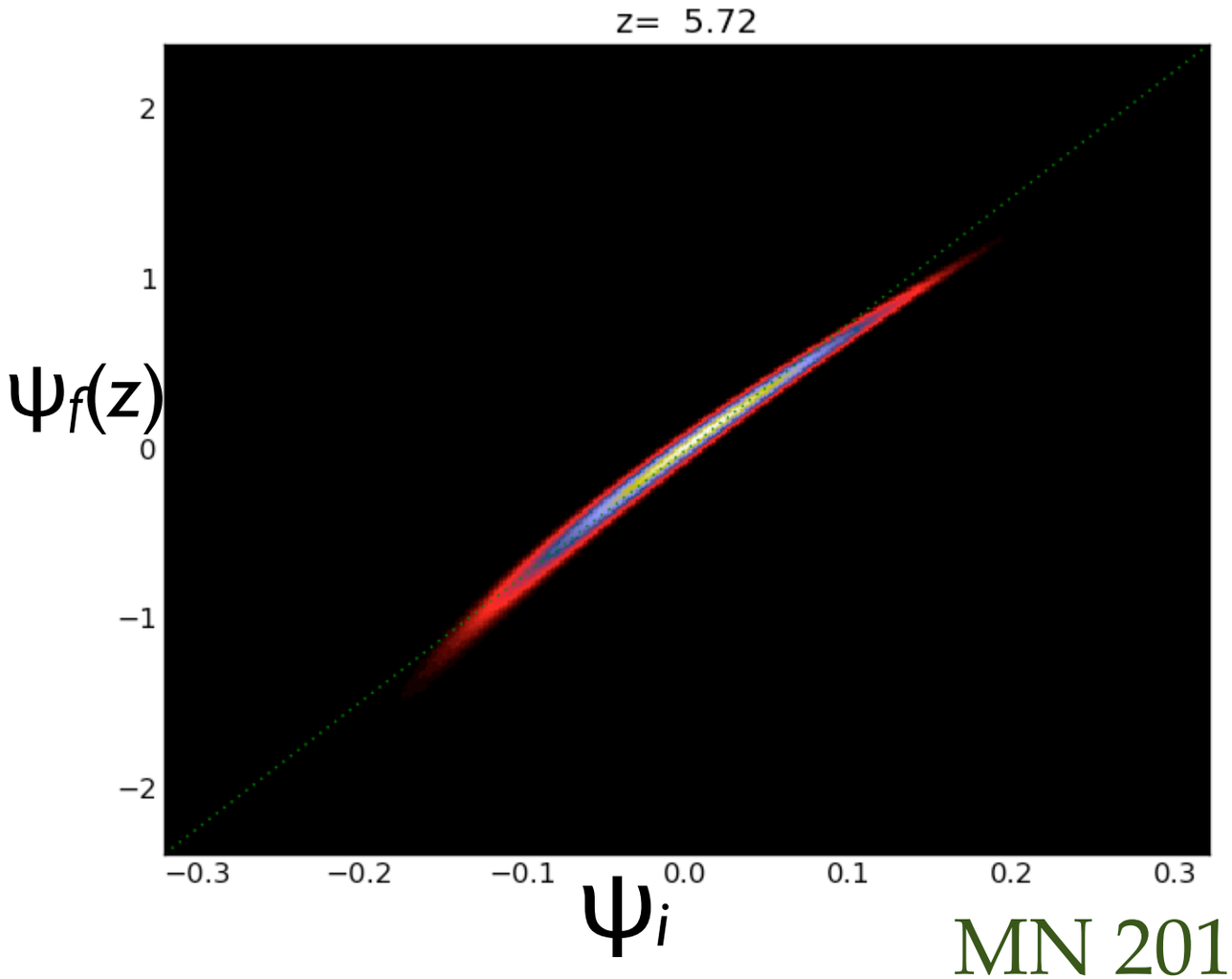


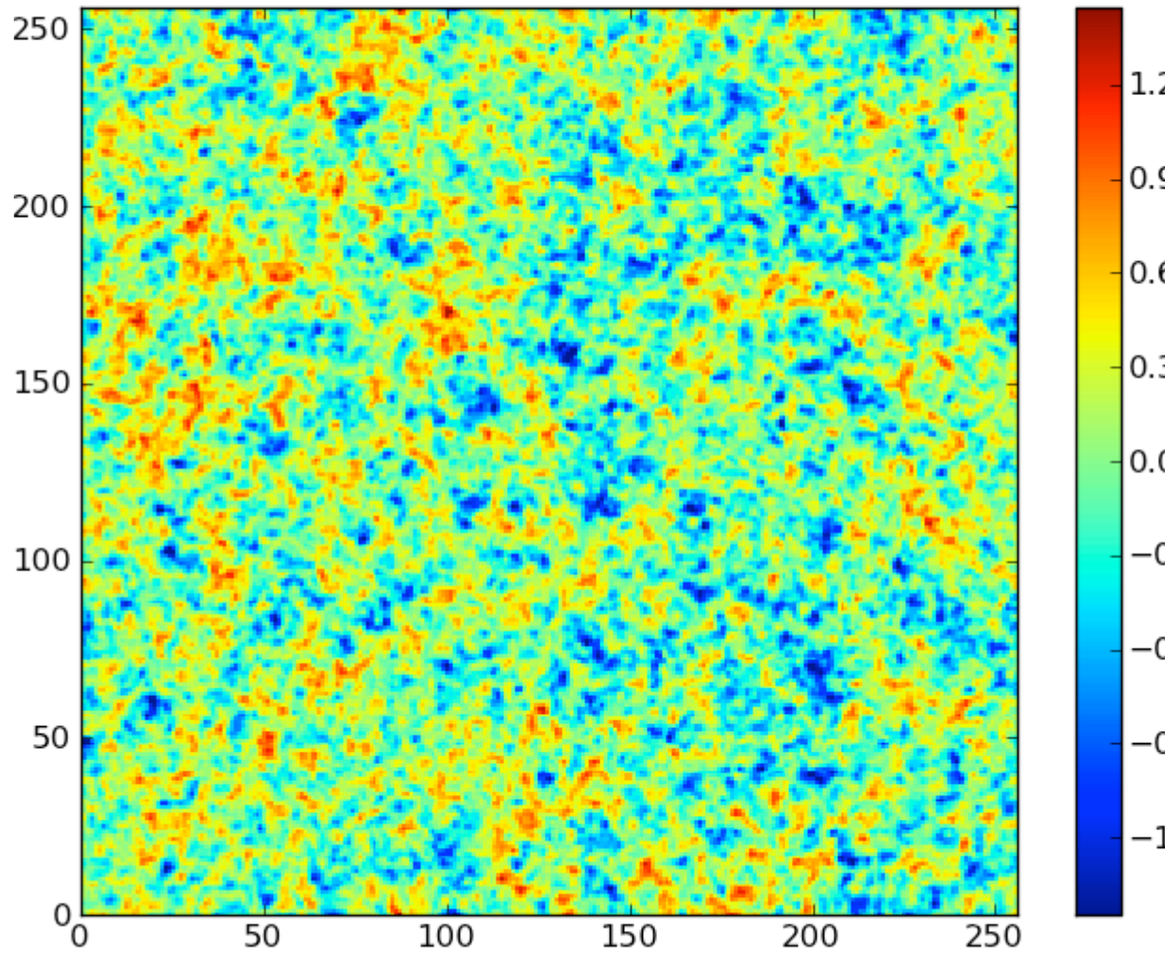
Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

$$\psi(\mathbf{q}) \equiv \boldsymbol{\nabla}_q \cdot \boldsymbol{\Psi}(\mathbf{q})$$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\boldsymbol{\nabla}_{\mathbf{L}} \cdot \boldsymbol{x}_f = 0$.



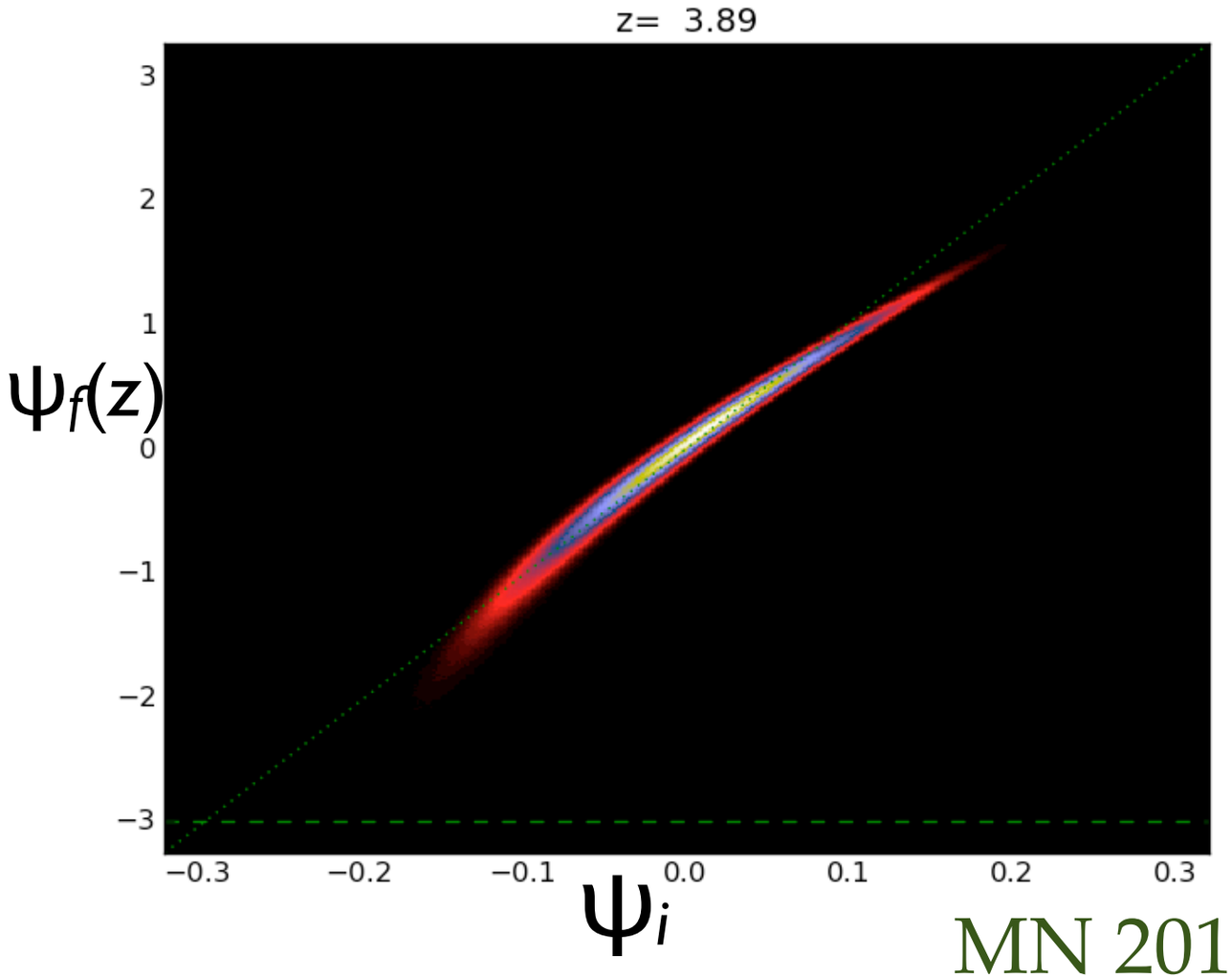


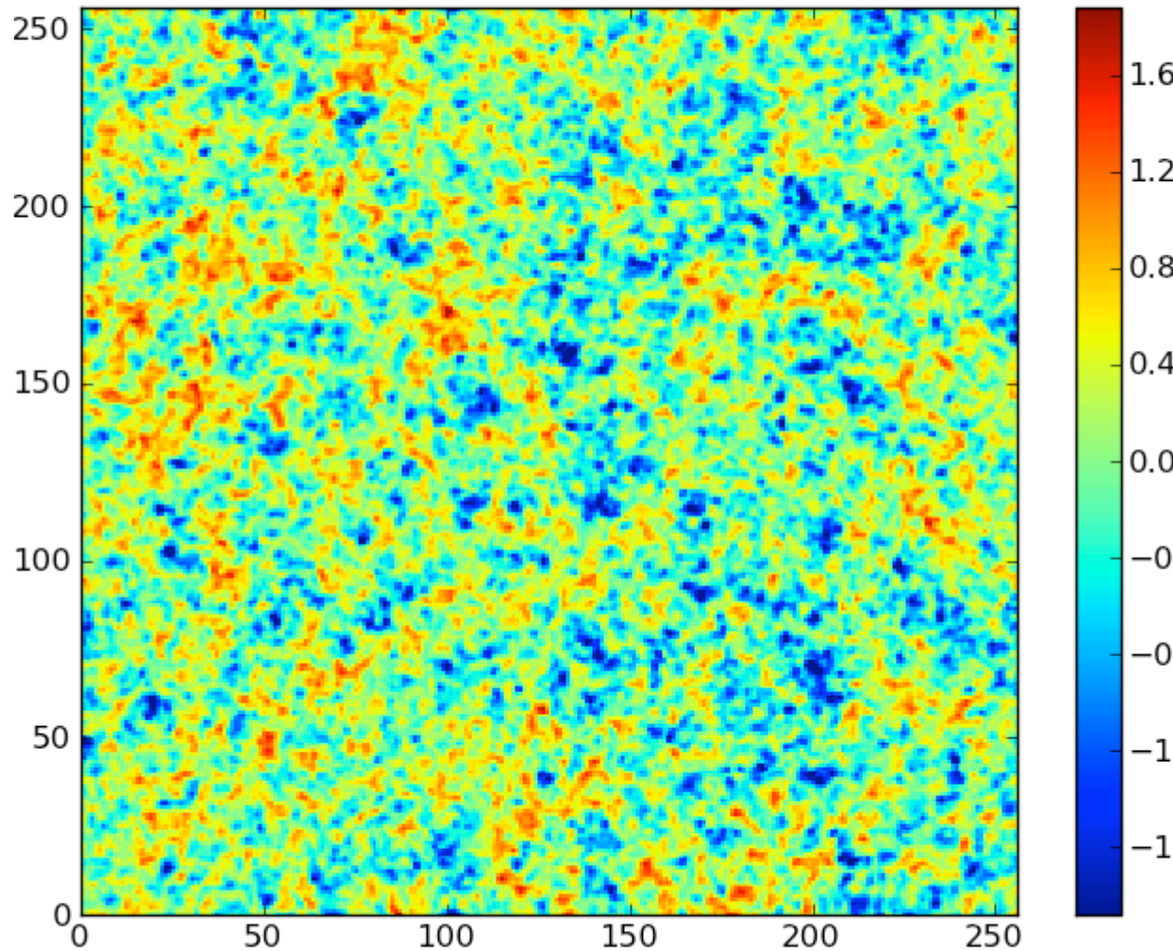
Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

$$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



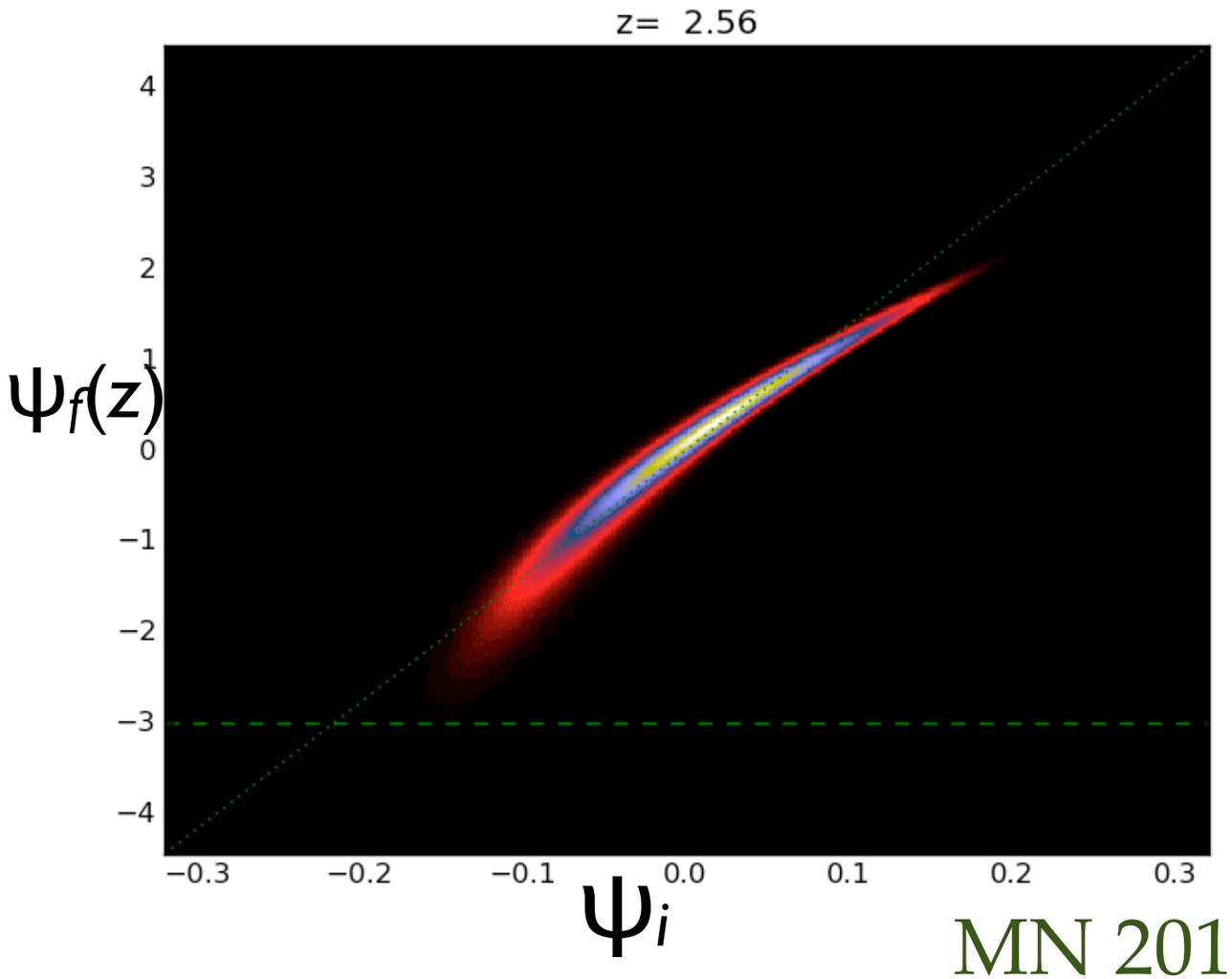


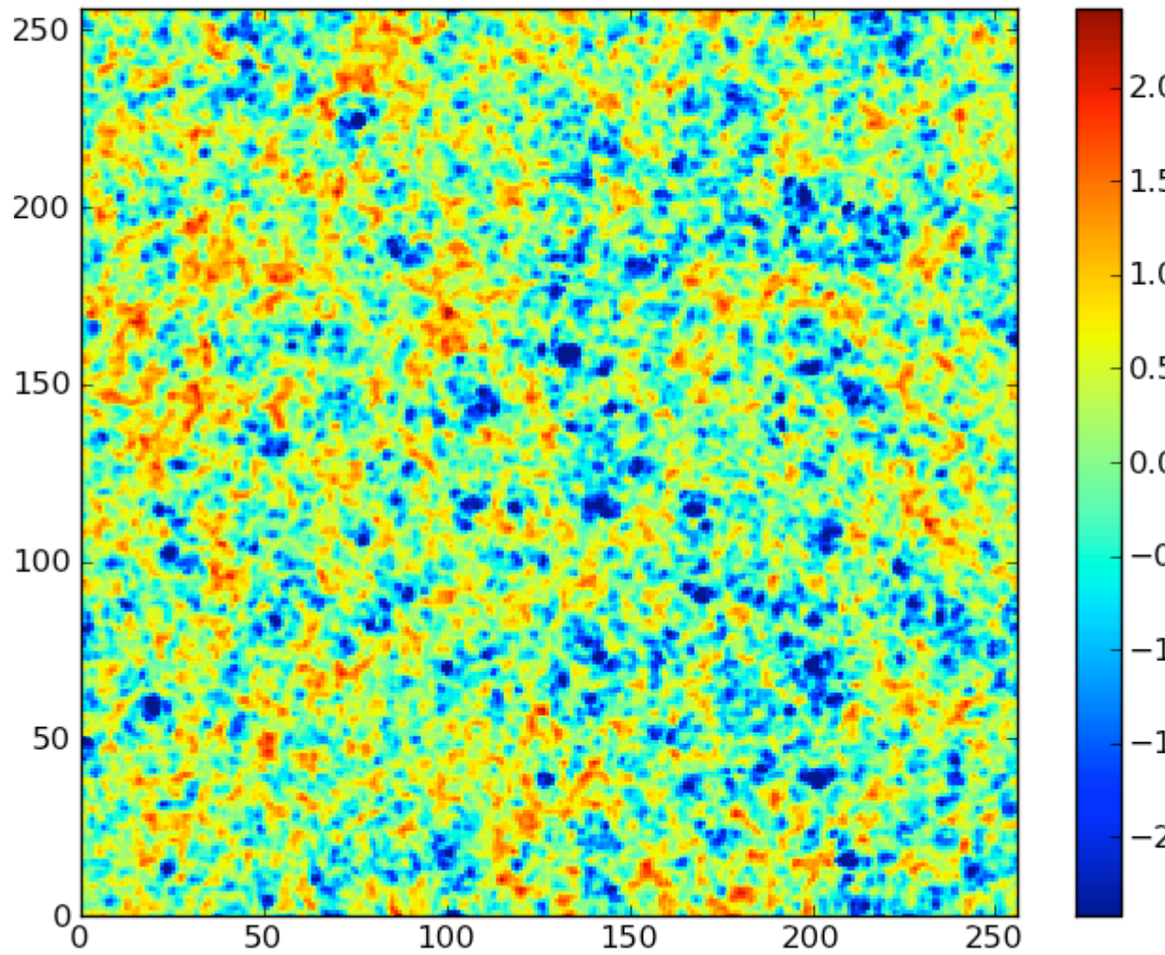
Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



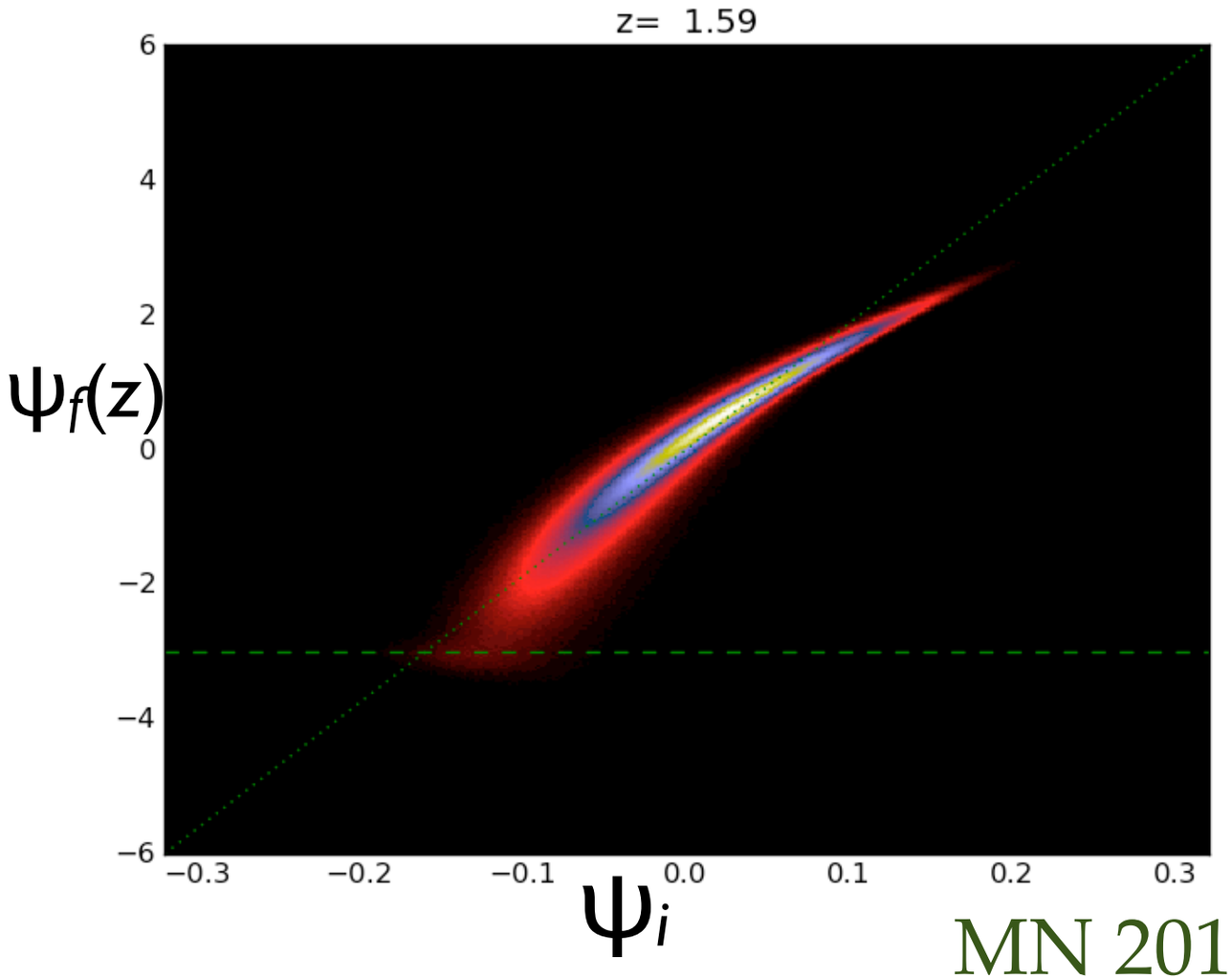


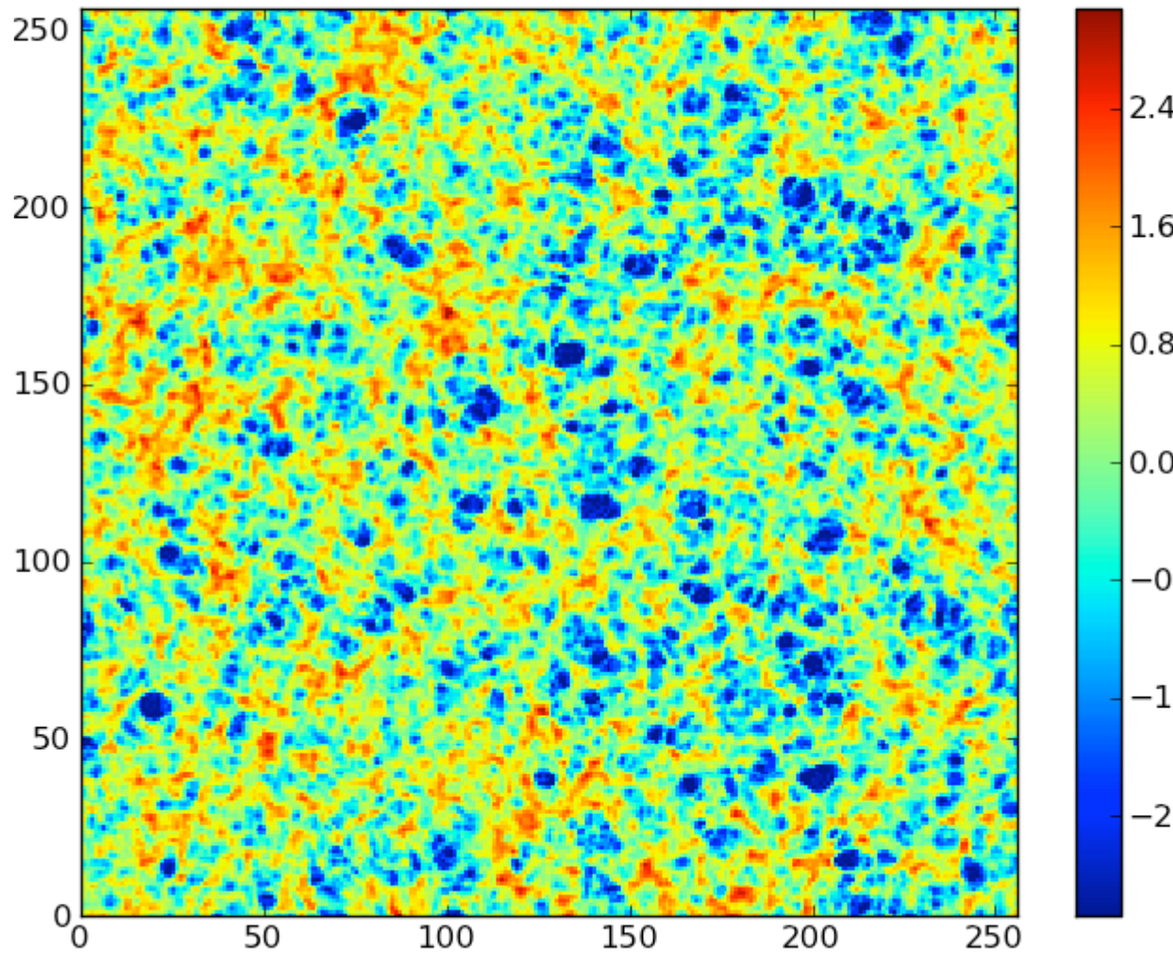
Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



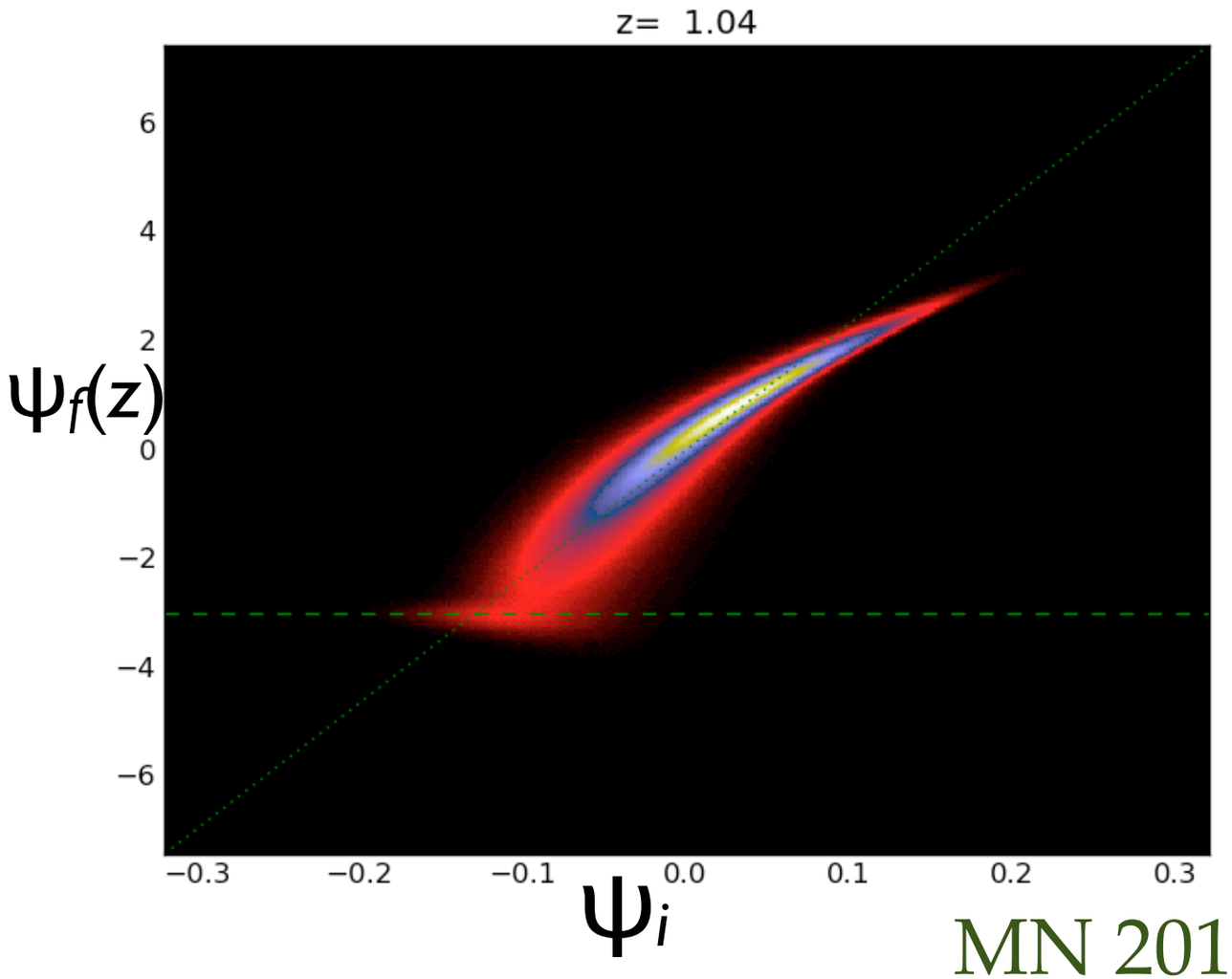


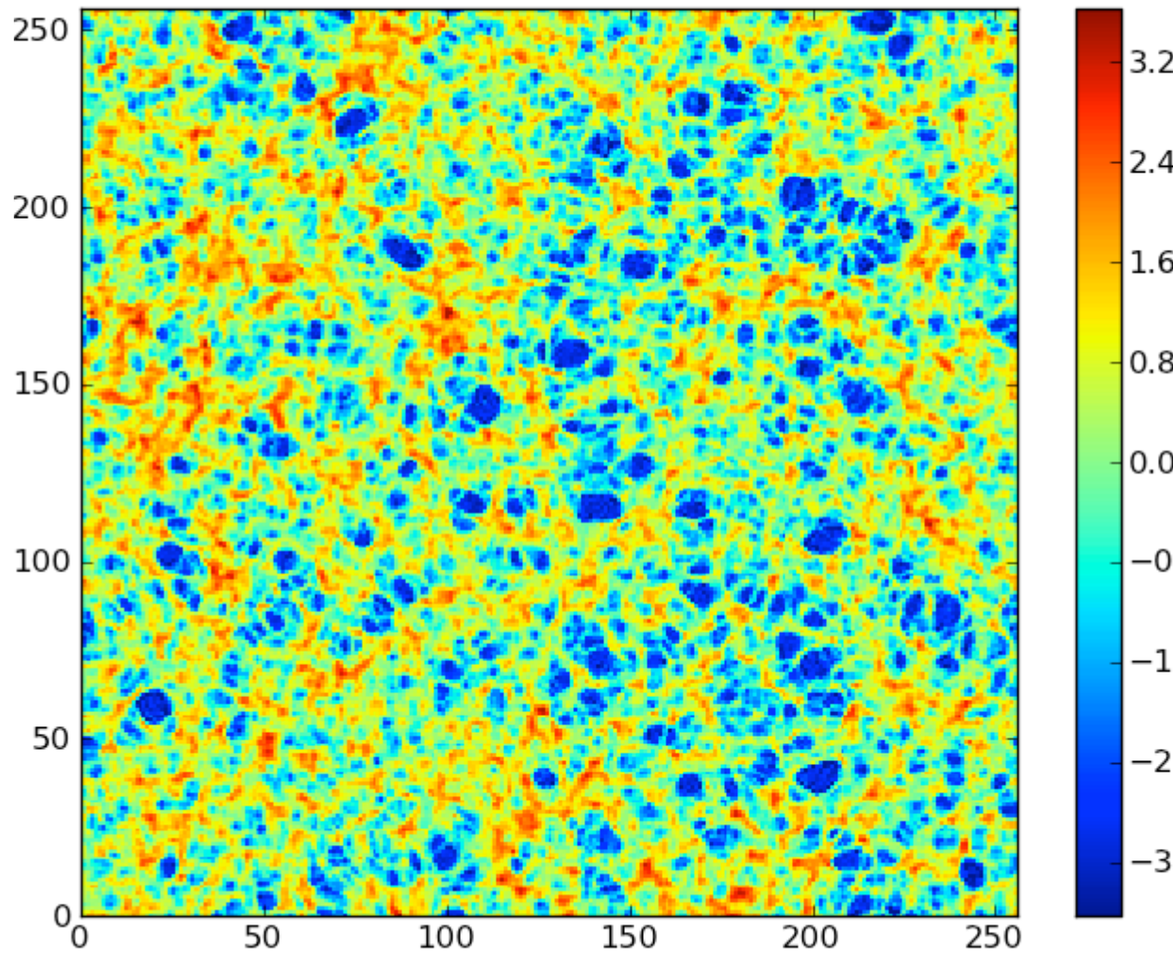
Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



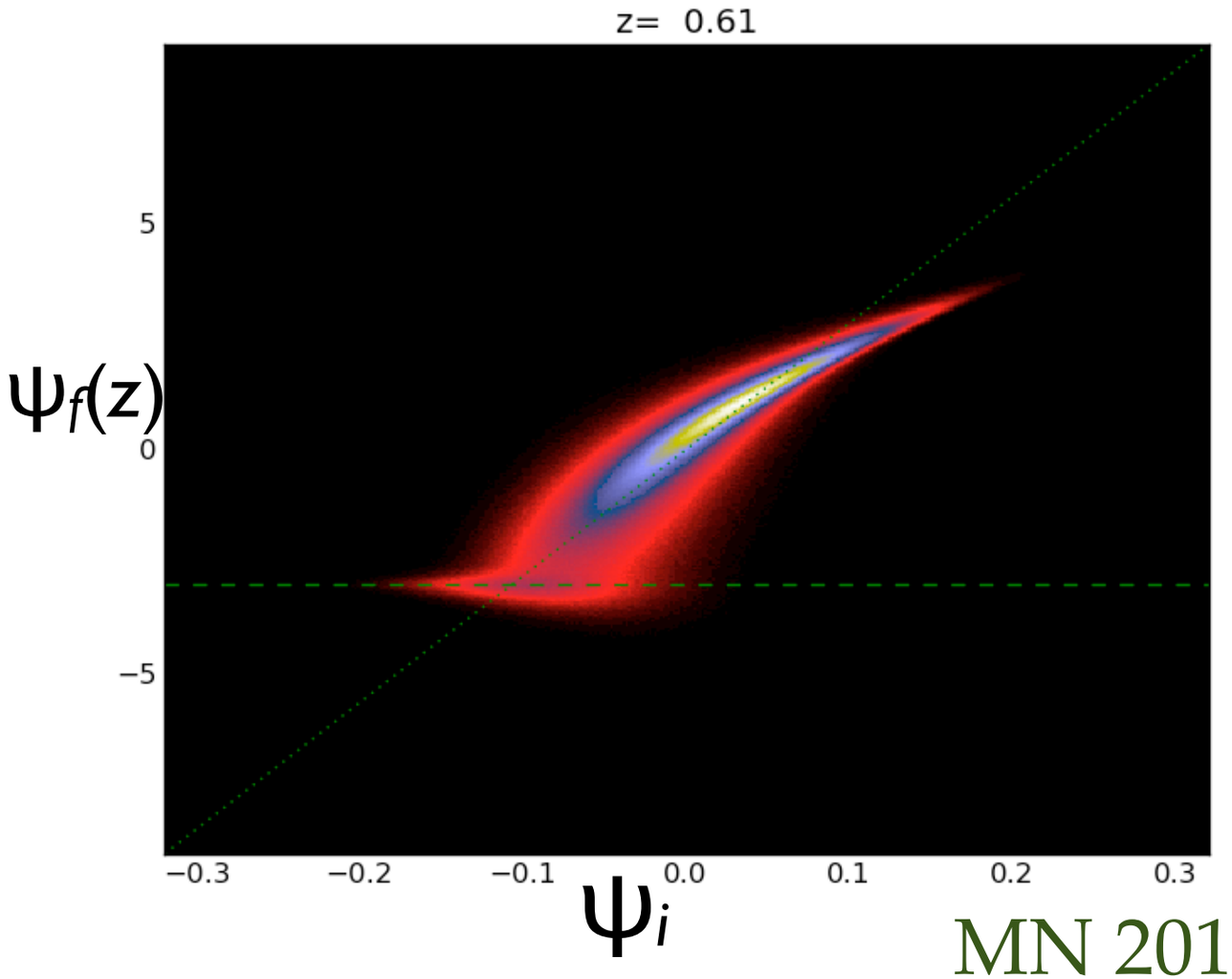


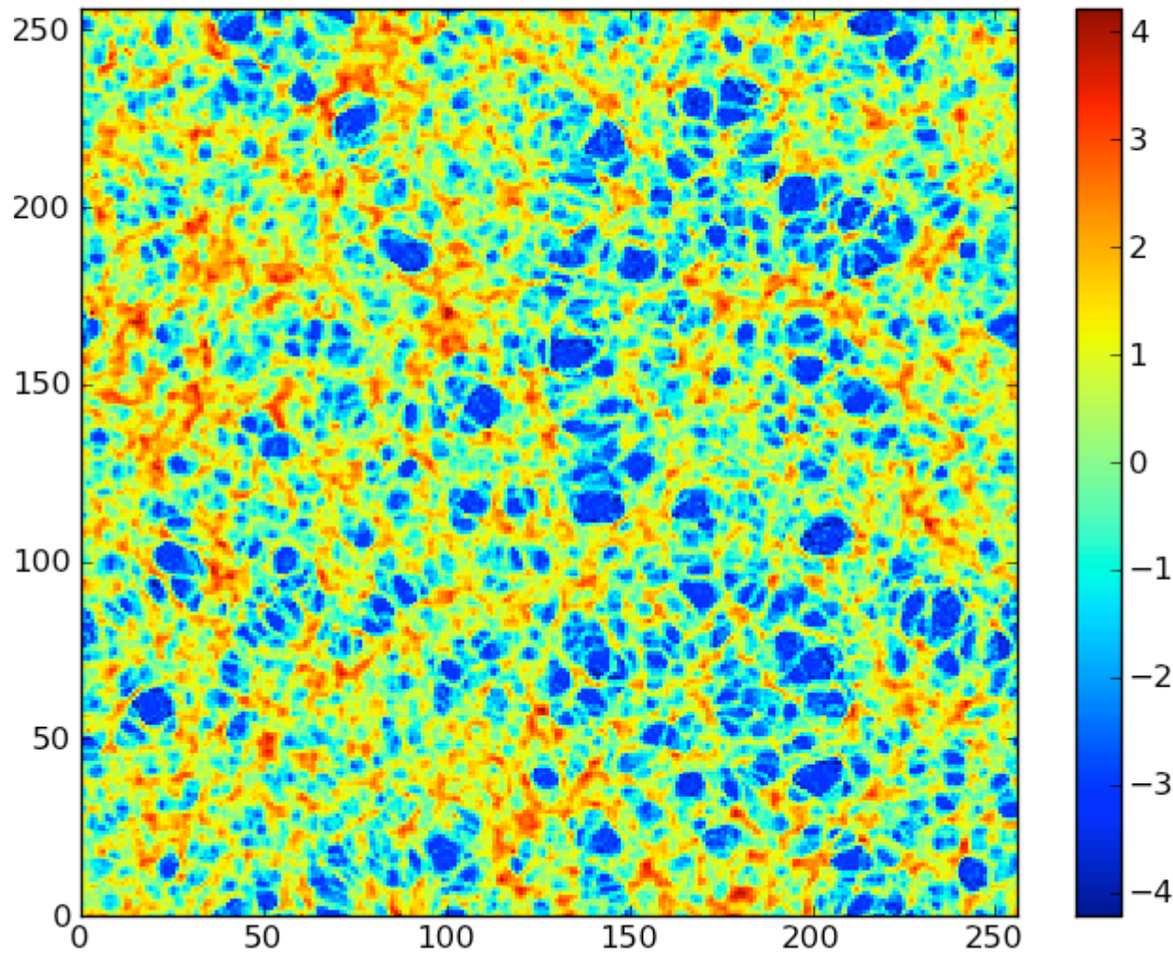
Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



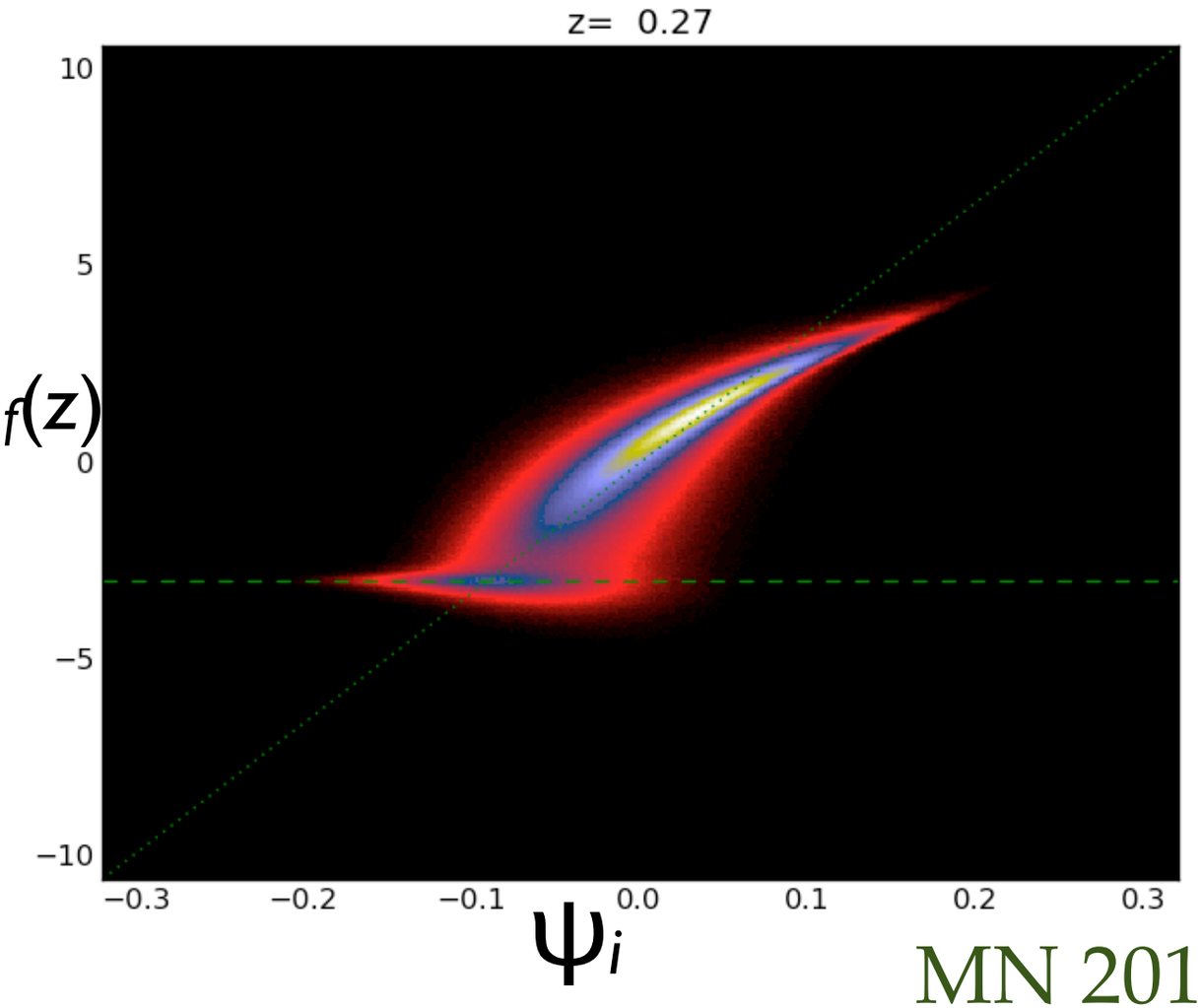


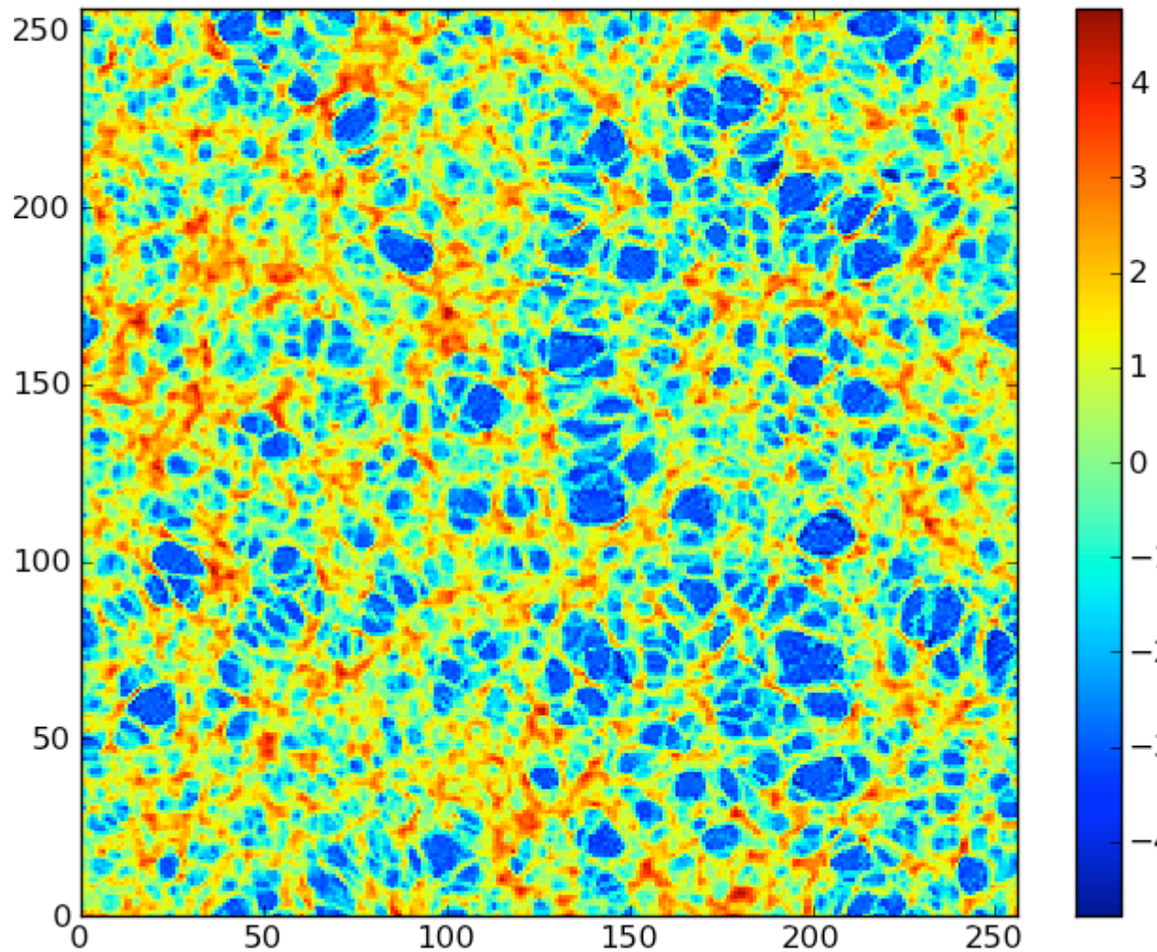
MN 2012, arXiv:1204.1326

Mark Neyrinck, JHU

$$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



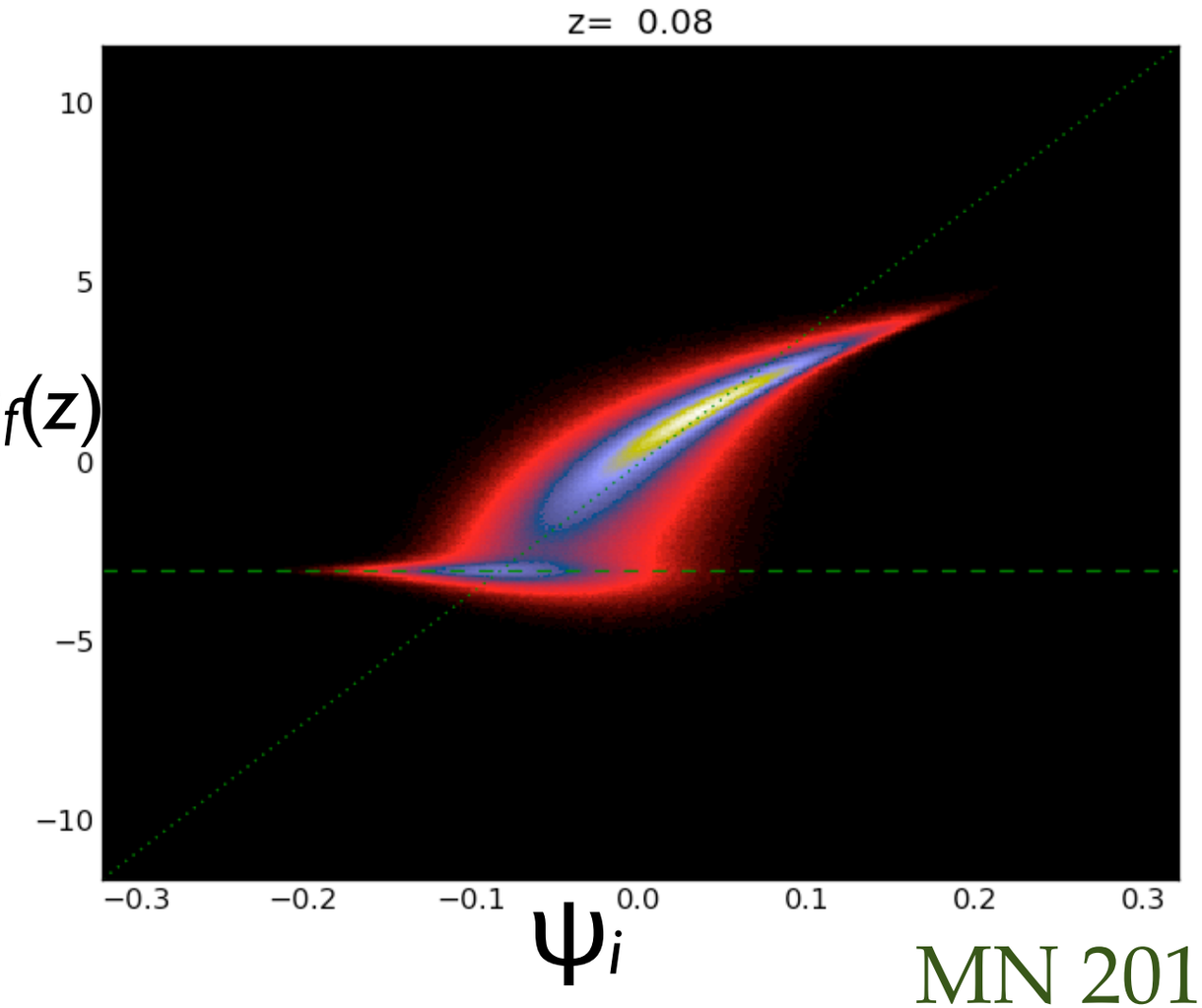


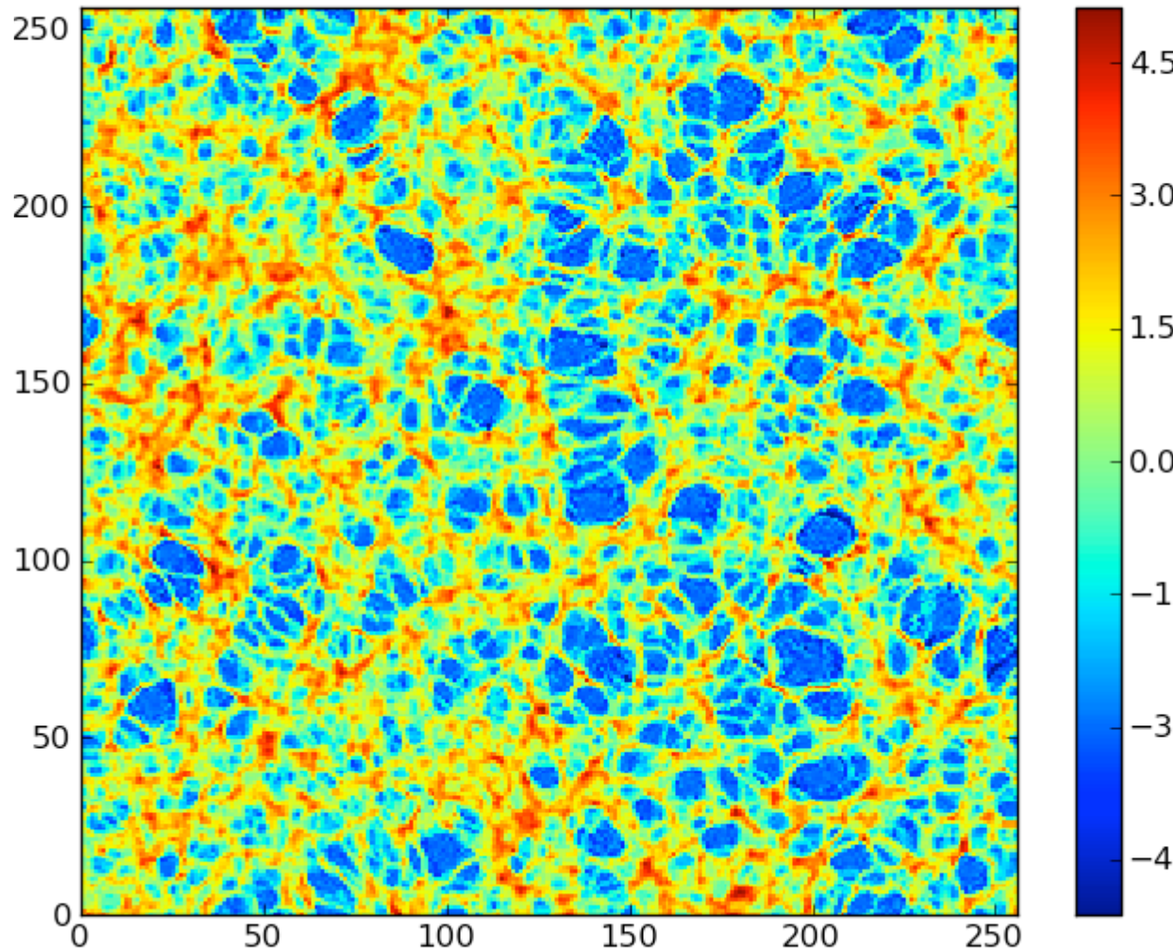
Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.



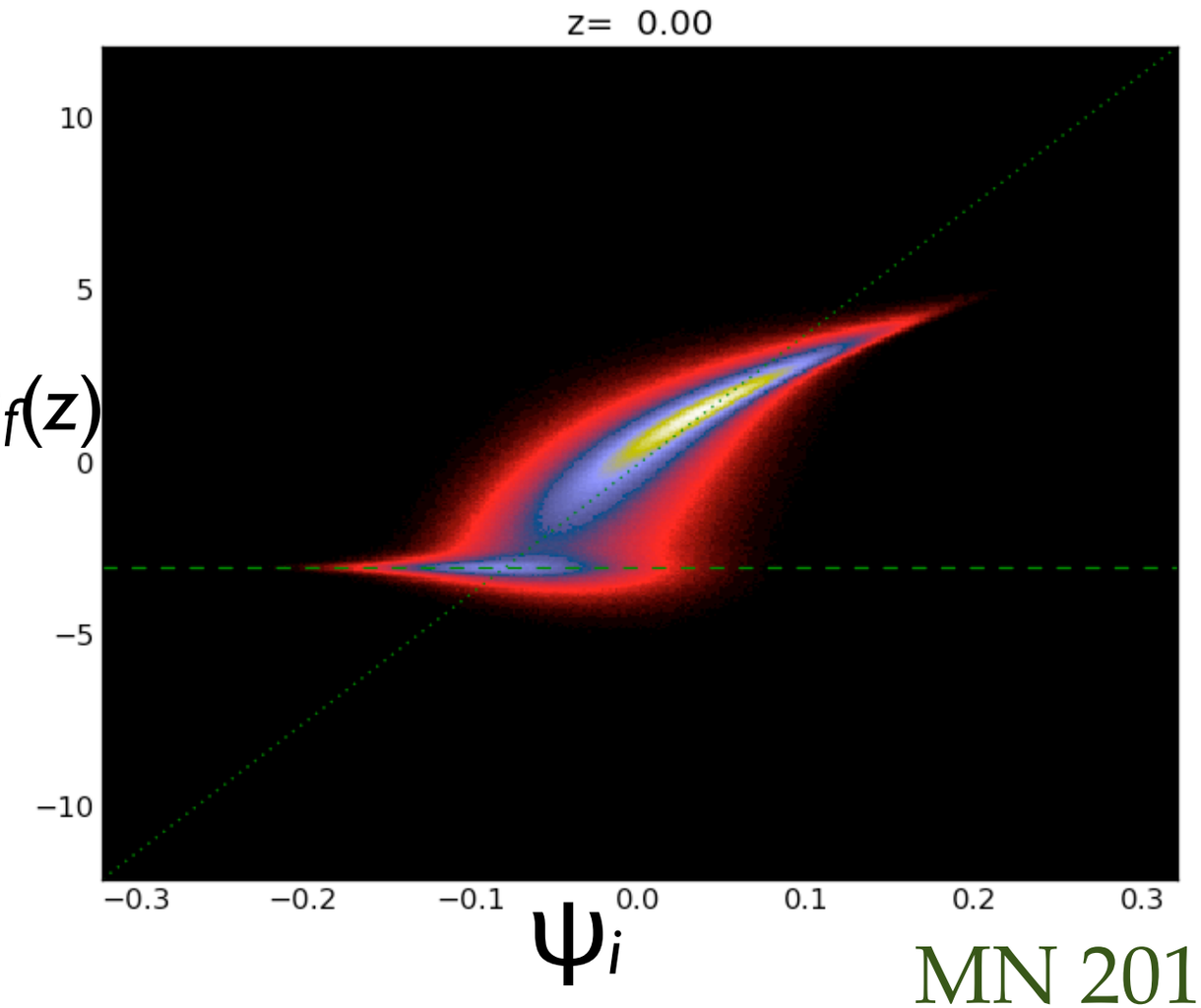


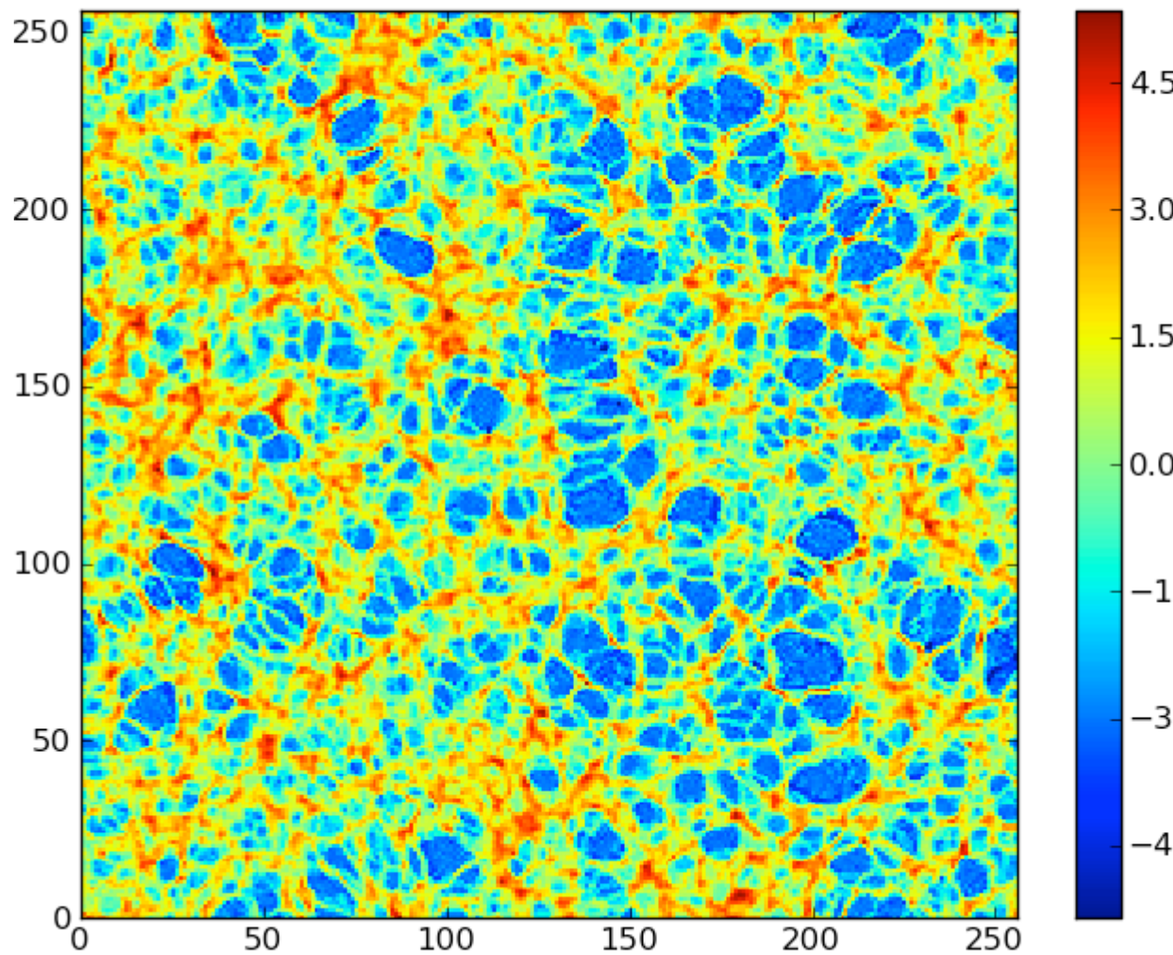
Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

$\psi(\mathbf{q}) \equiv \nabla_q \cdot \Psi(\mathbf{q})$ behavior

- Zel'dovich (1970): $\psi = -\delta_{linear}$.
- $\psi = -3$: halo formation, where $\nabla_L \cdot x_f = 0$.





Mark Neyrinck, JHU

MN 2012, arXiv:1204.1326

Can we get a good approximate particle realization from this?

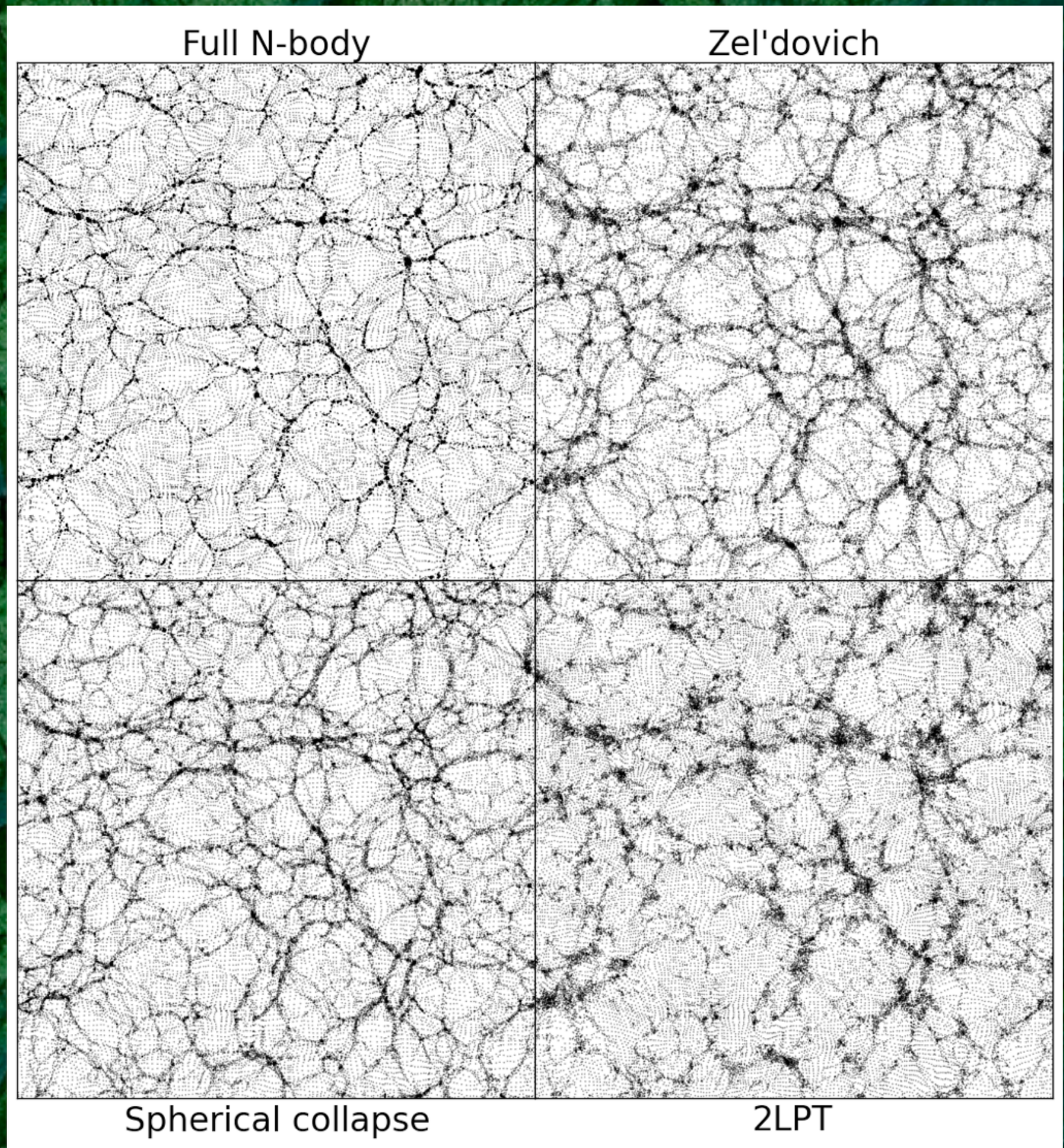
- Use this formula to predict $\nabla_L \cdot \psi$. Where collapse, set $\nabla_L \cdot \psi$=-3.

- Should give accuracy in voids, and little stream-crossing.

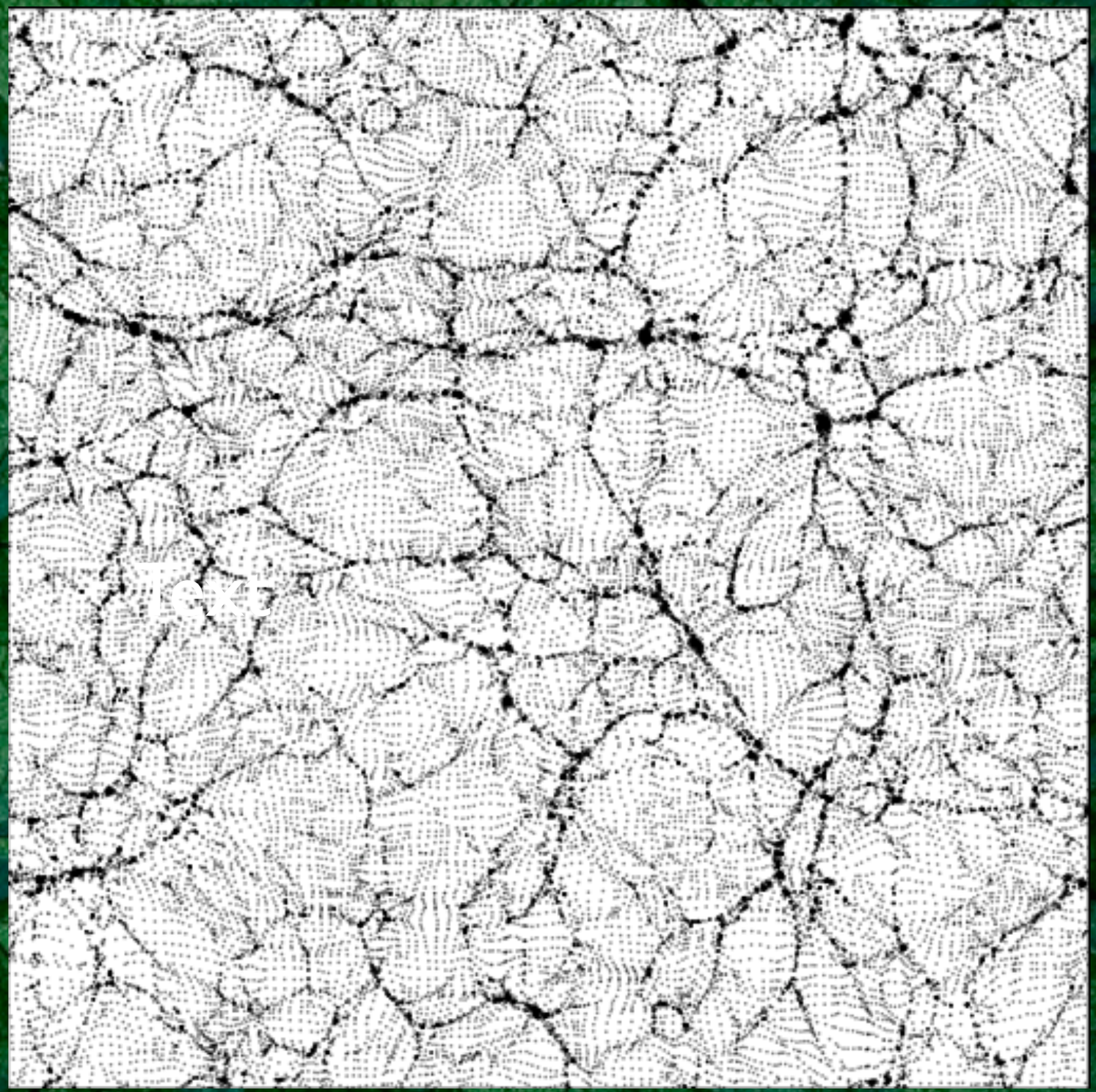
- Poor void behavior in voids for 2LPT (as in Sahni & Shandarin 1996)

MN 2012, arXiv: 1204.1326

Mark Neyrinck, JHU



N-body

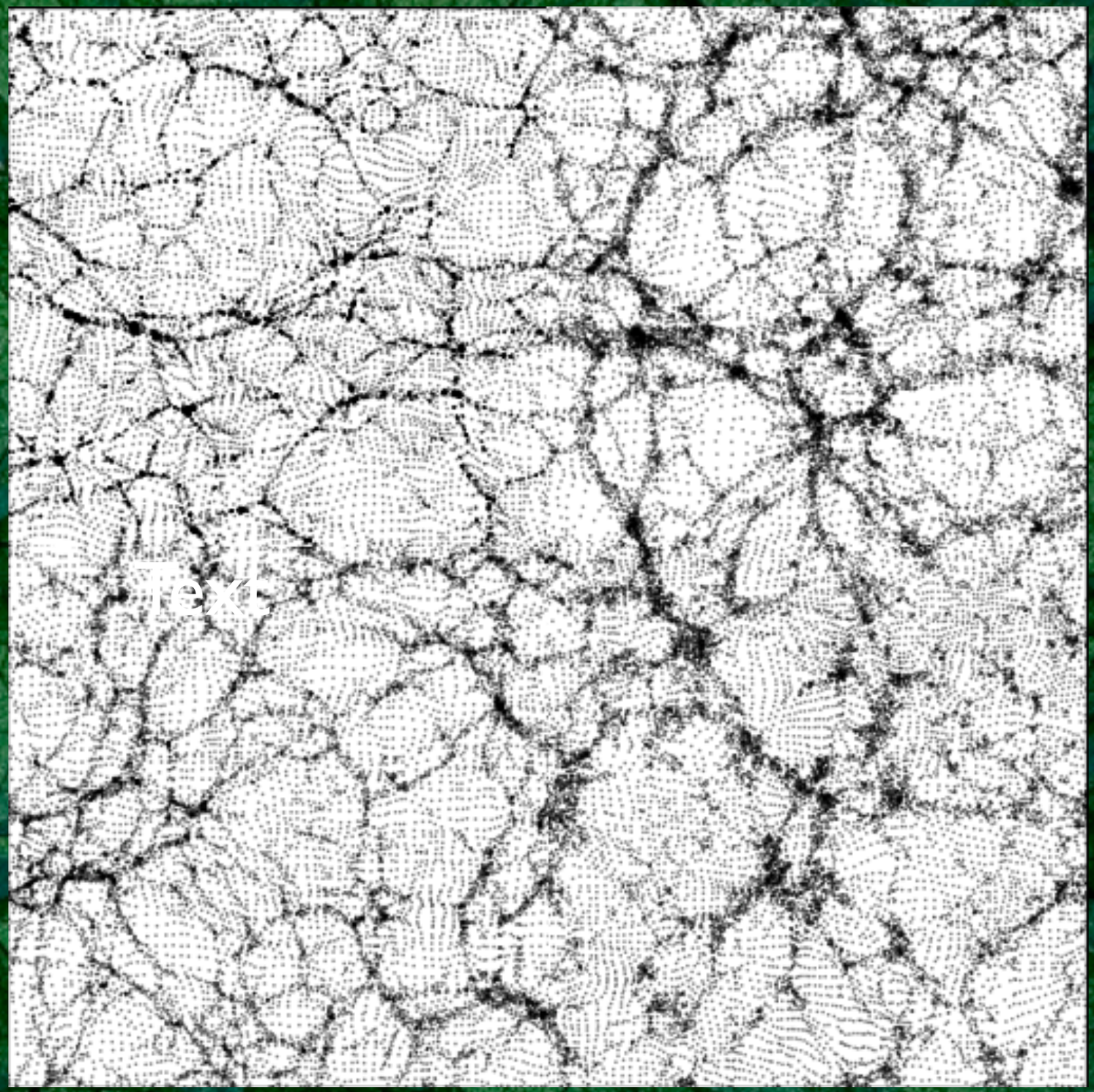


(suggests caution in using 2LPT at high resolution and low redshift -- obvious problem with using perturbation theory outside its range of validity)

MN 2012, arXiv: 1204.1326

Mark Neyrinck, JHU

(suggests caution in using 2LPT at high resolution and low redshift -- obvious problem with using perturbation theory outside its range of validity)



MN 2012, arXiv: 1204.1326

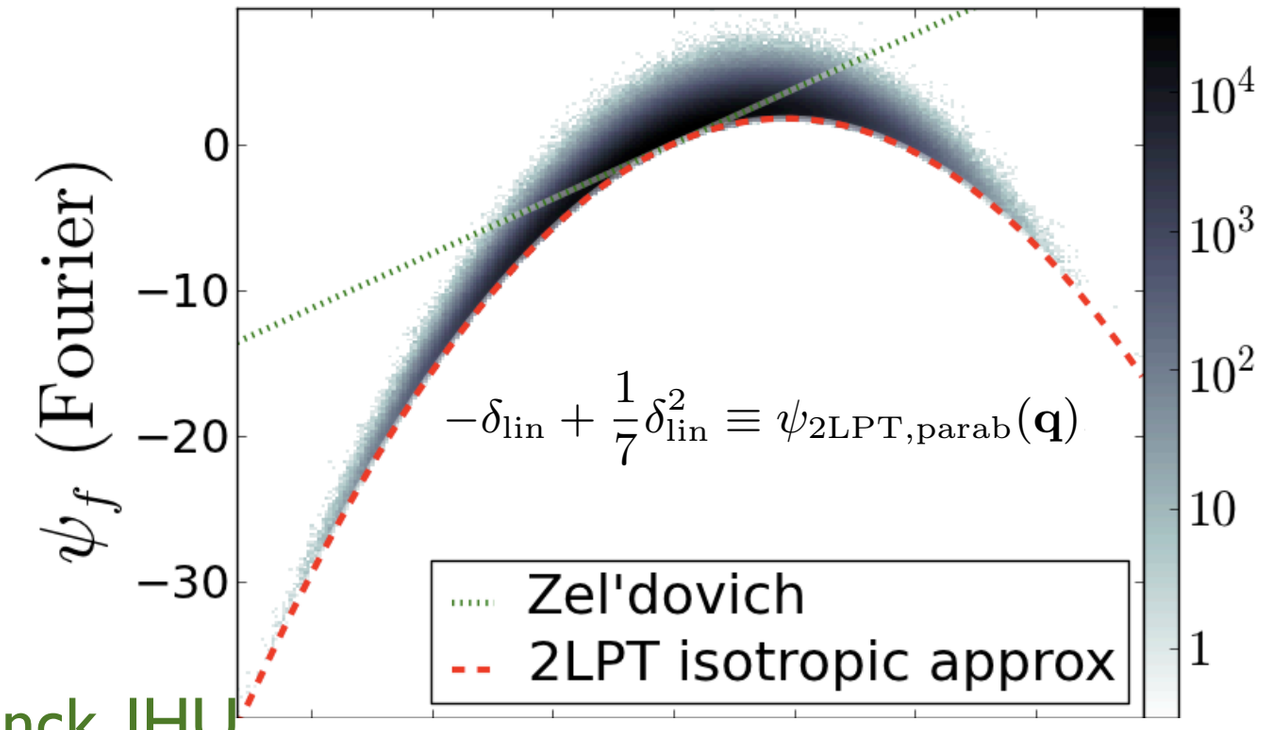
Mark Neyrinck, JHU

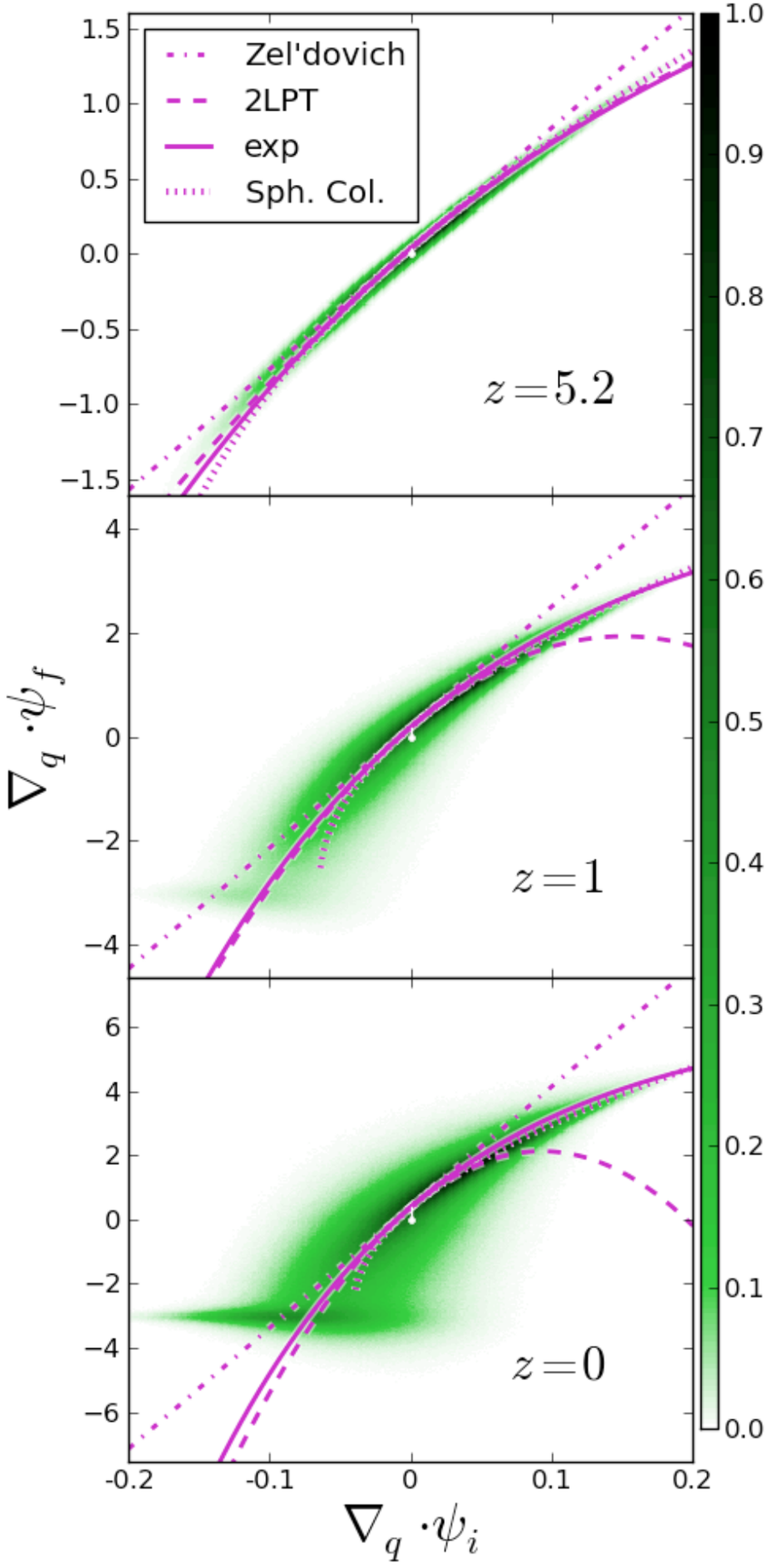
"Local Lagrangian" approximations work quite well (Protogeros & Scherrer 1997)

- Bernardeau's (1992, 1994) "spherical-collapse" fit (also see Mohayaee et al. 2006) works quite well:

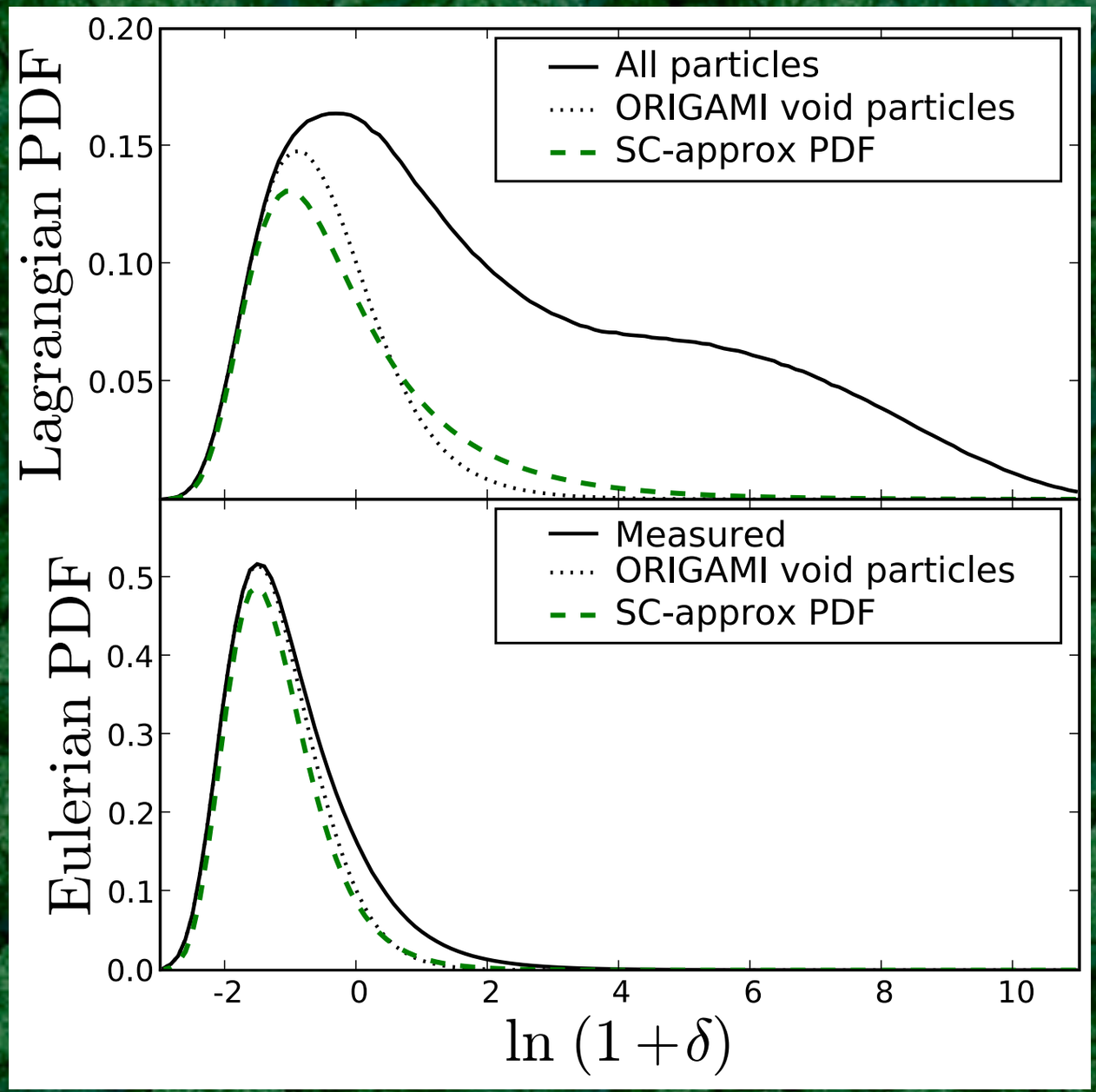
$$\psi_{\rm sc} \quad = \quad 3\left[\left(1+\frac{2}{3}\psi_{\rm lin}\right)^{1/2}-1\right]$$

- 2LPT on the other hand, behaves oddly for highly nonlinear density fields

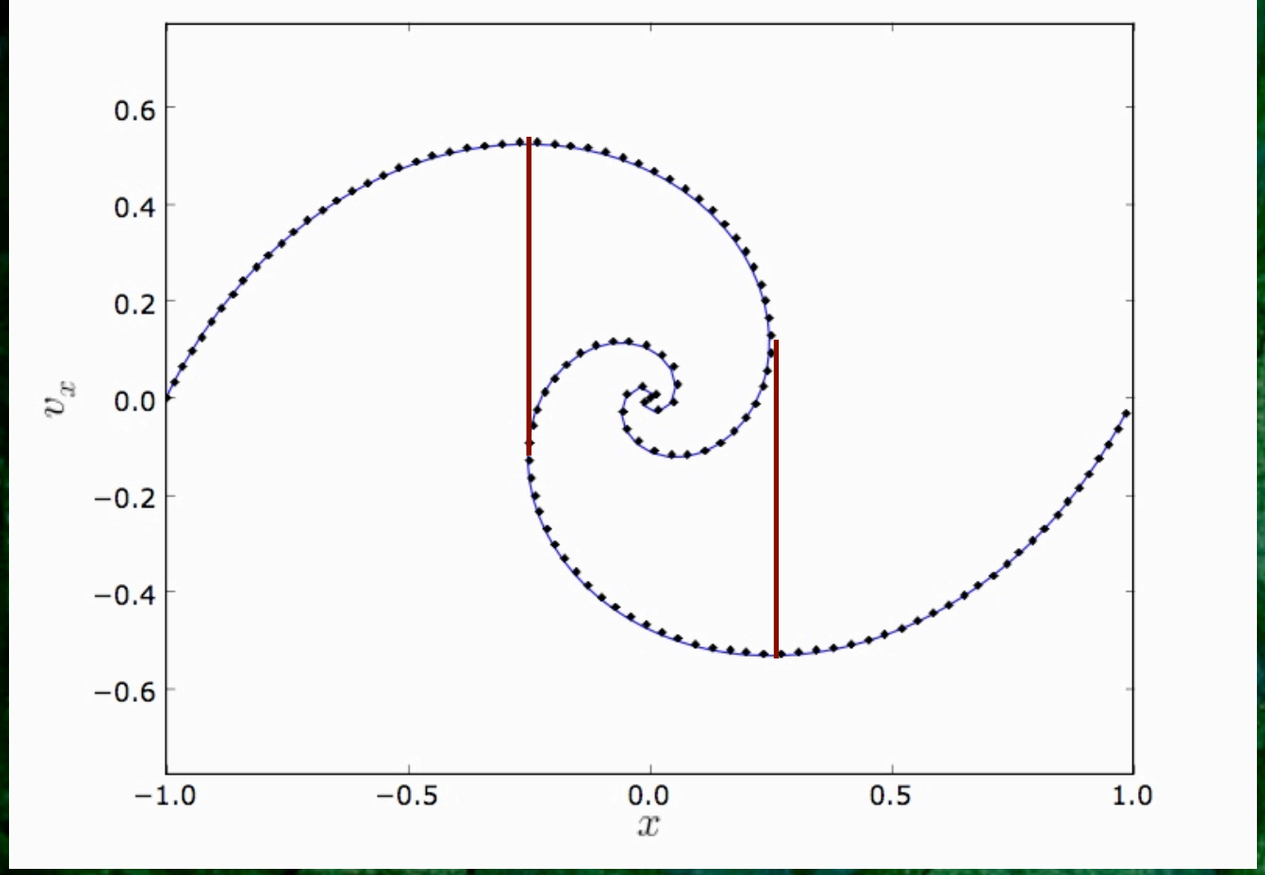


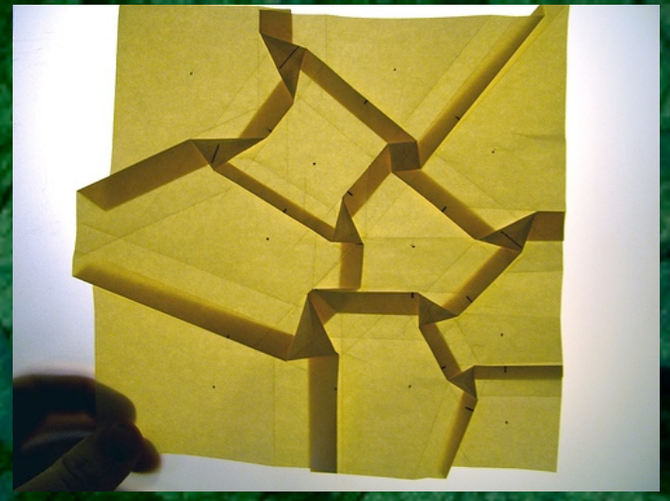


Allows a good description of the pdf in voids.



Mark Neyrinck, JHU





ORIGAMI

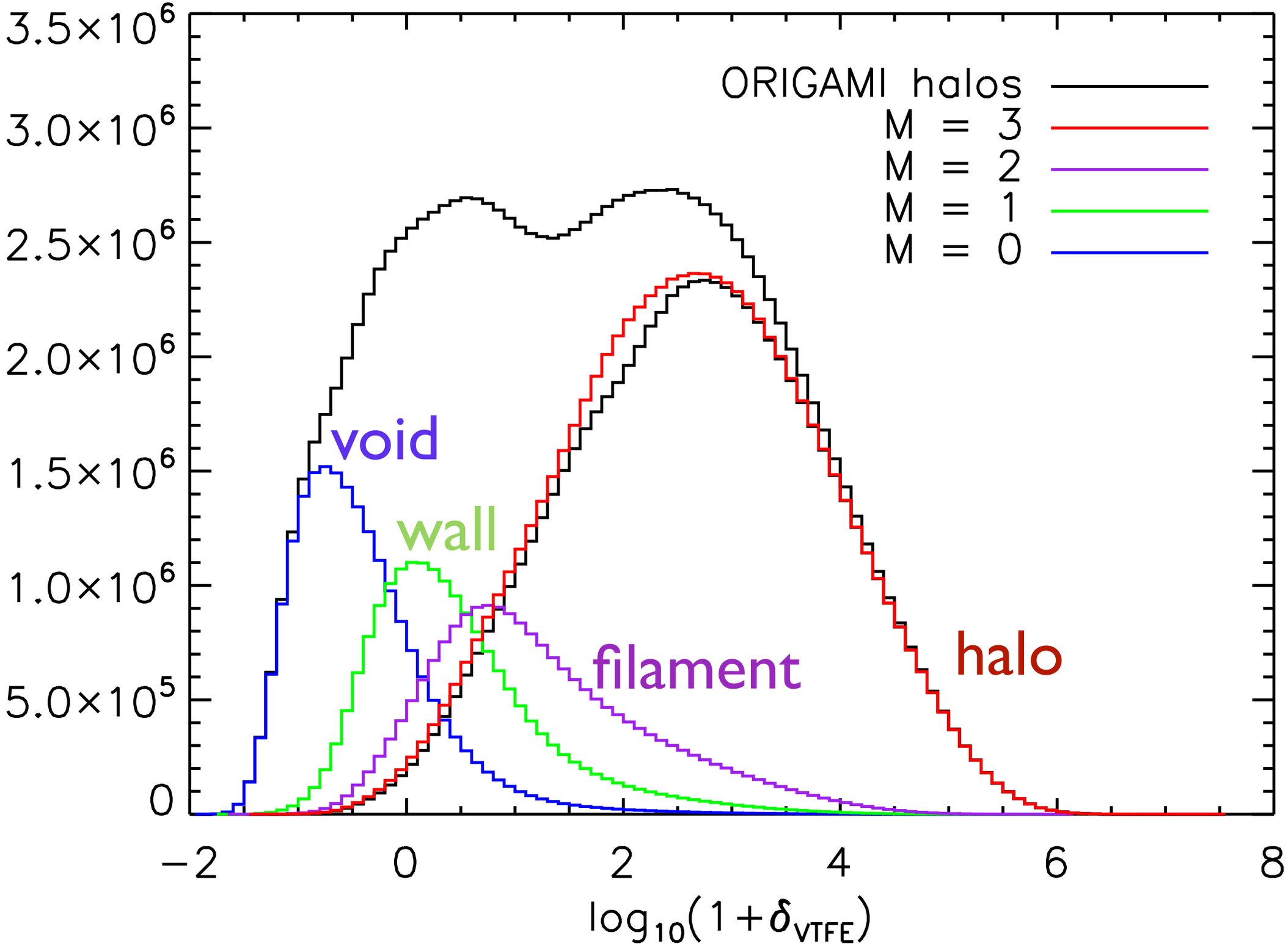
Order-ReversIng Gravity, Apprehended Mangling Indices

Bridget Falck, MN, & Alex Szalay,
arXiv:1201.2353

- *Objective, dynamic, geometric* definition of halo boundaries
- 1d: particle in a halo if its order wrt *any* another particle is swapped compared to the original Lagrangian ordering
- 3d: particle in a halo if this condition holds along 3 orthogonal axes (2 axes=filament, 1 axis=wall, 0 axes=void)
- Finds places where streams have crossed

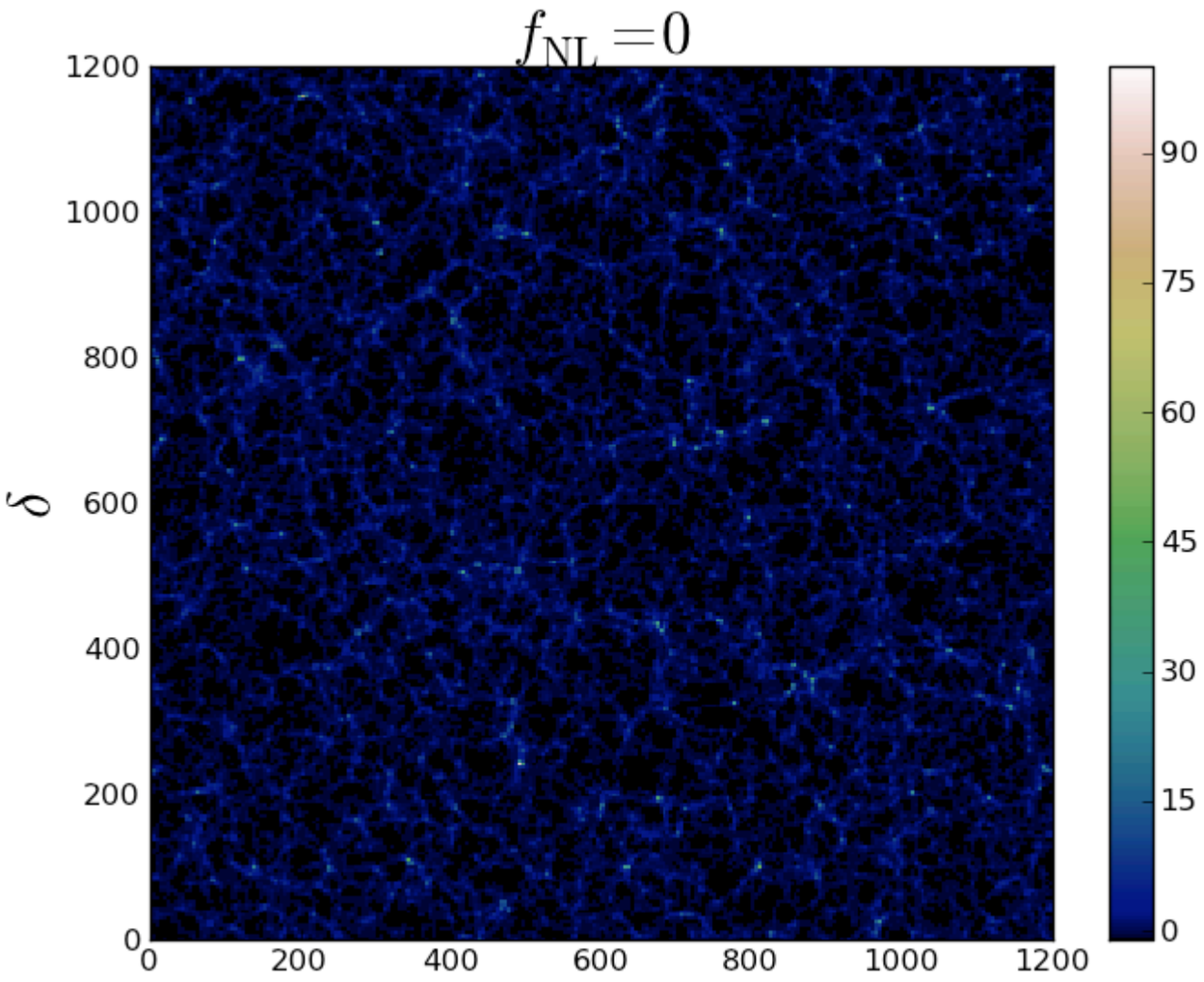
Mark Neyrinck, JHU

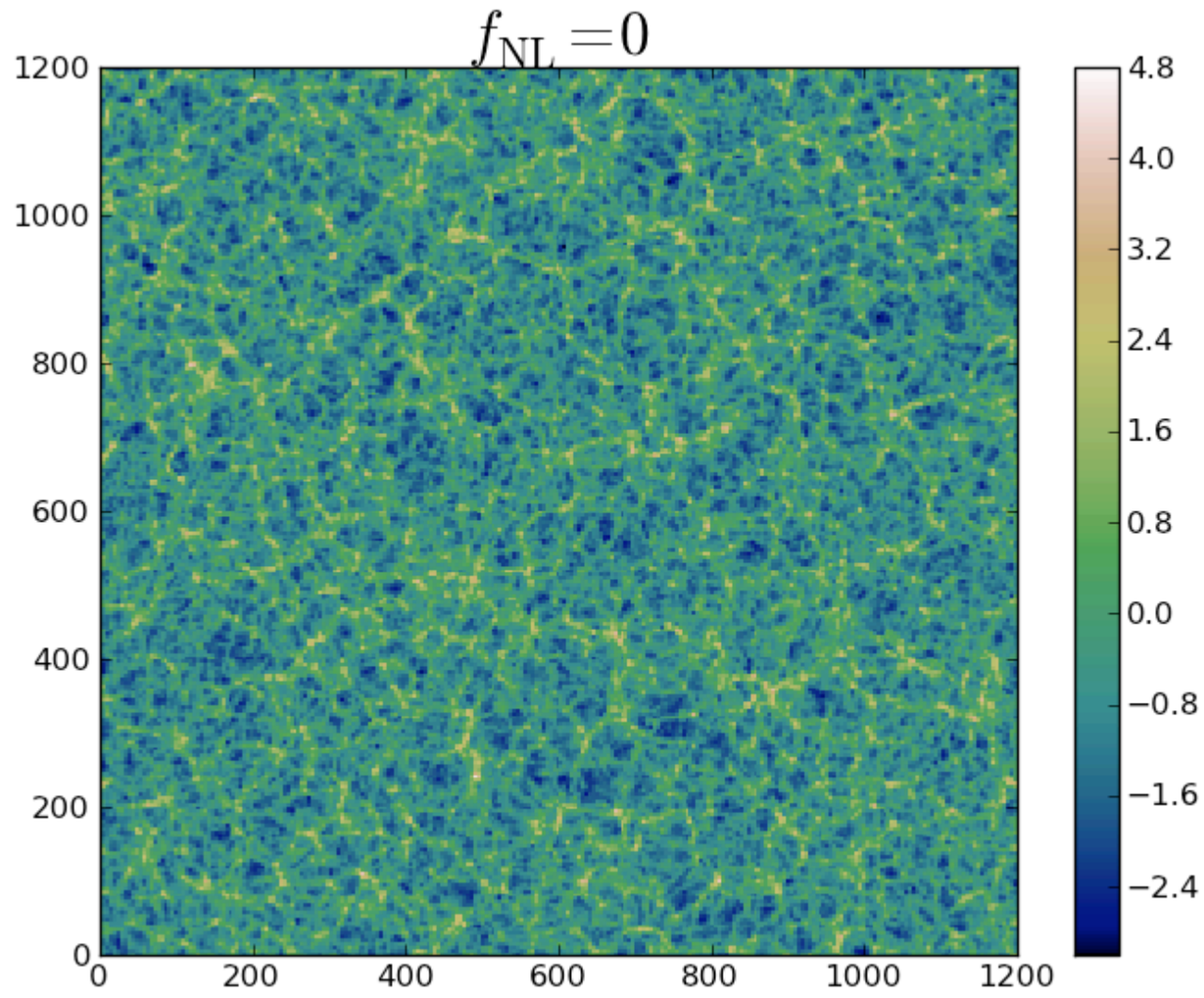
Mass-weighted histogram of particle density



Mark Neyrinck, JHU

Example of simplicity in single-stream regions: f_{NL}

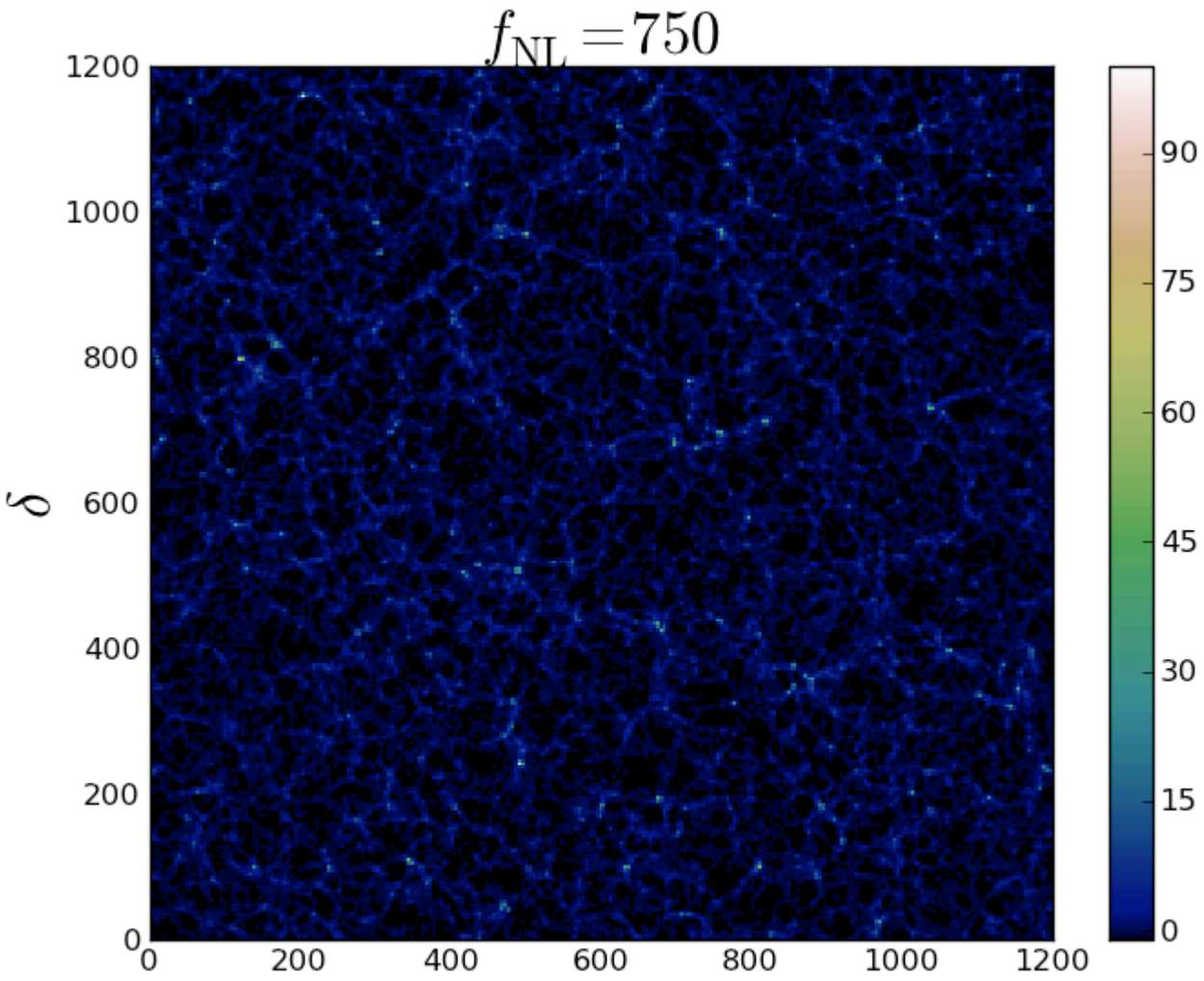


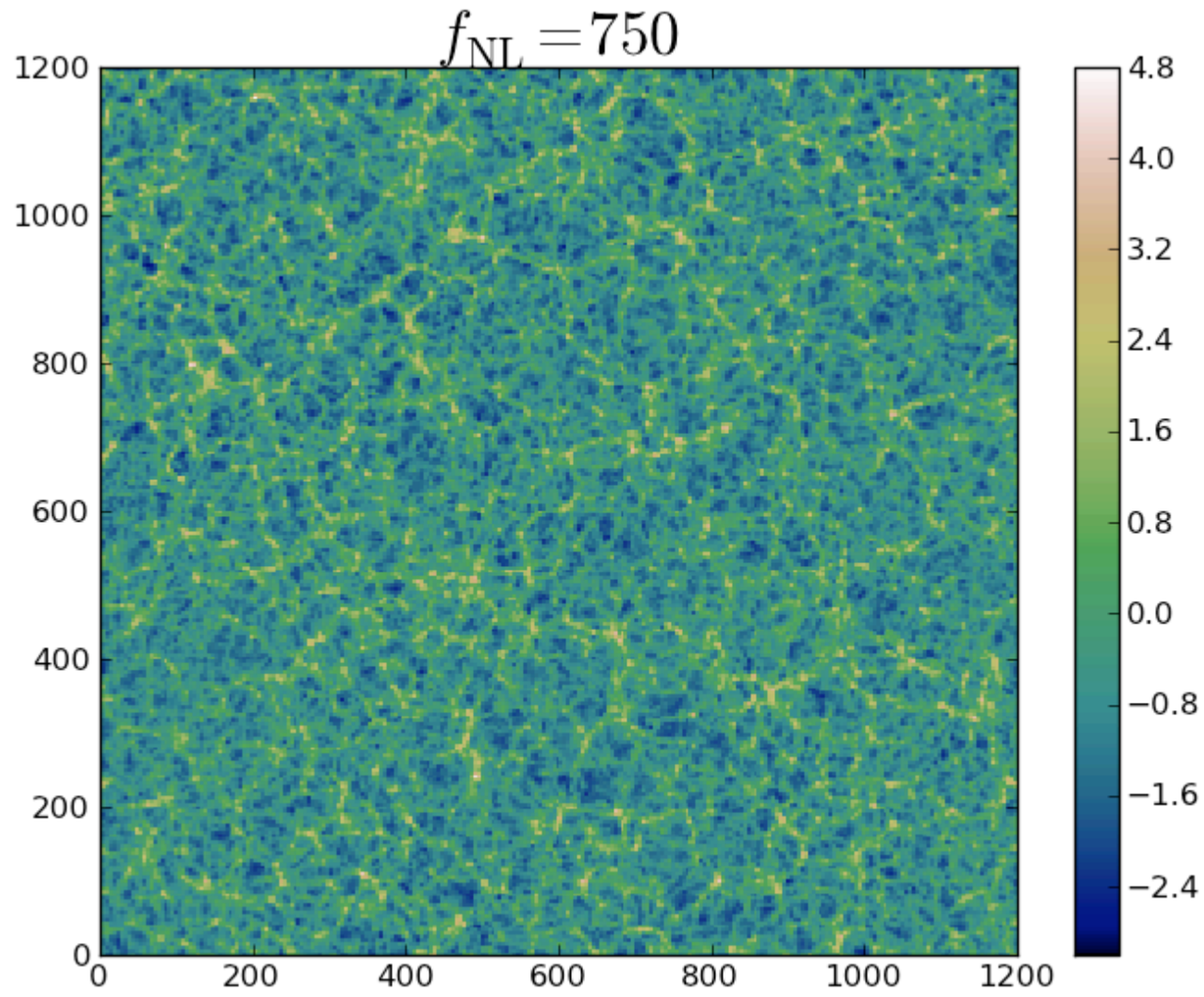


Simulations of Giannantonio & Porciani (2010), collaboration w/ many incl. Toni Riotto

Mark Neyrinck, JHU

Example of simplicity in single-stream regions: f_{NL}



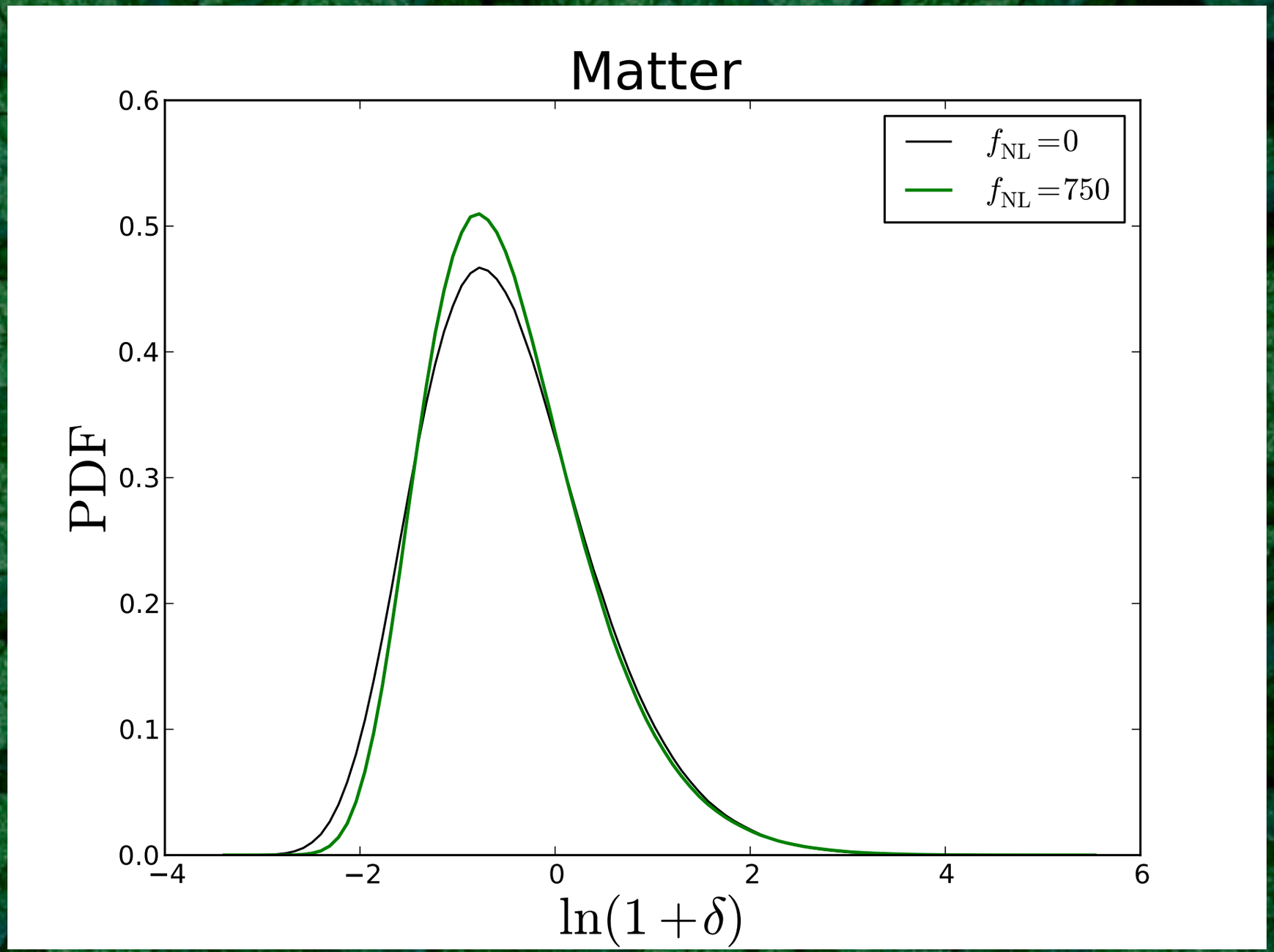


Simulations of Giannantonio & Porciani (2010), collaboration w/ many incl. Toni Riotto

Mark Neyrinck, JHU

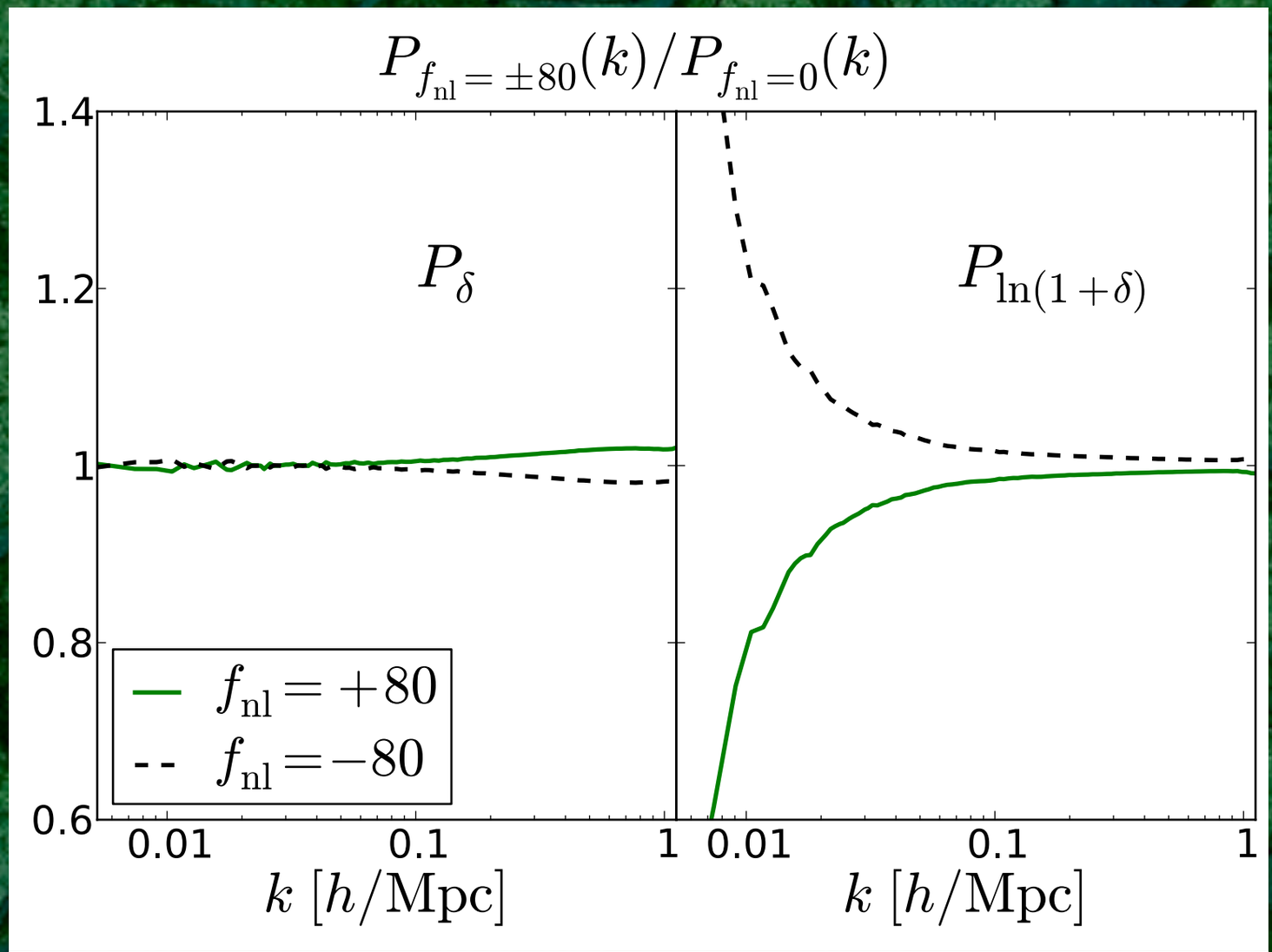
Example of simplicity in single-stream regions: f_{NL}

Difference mainly for low δ



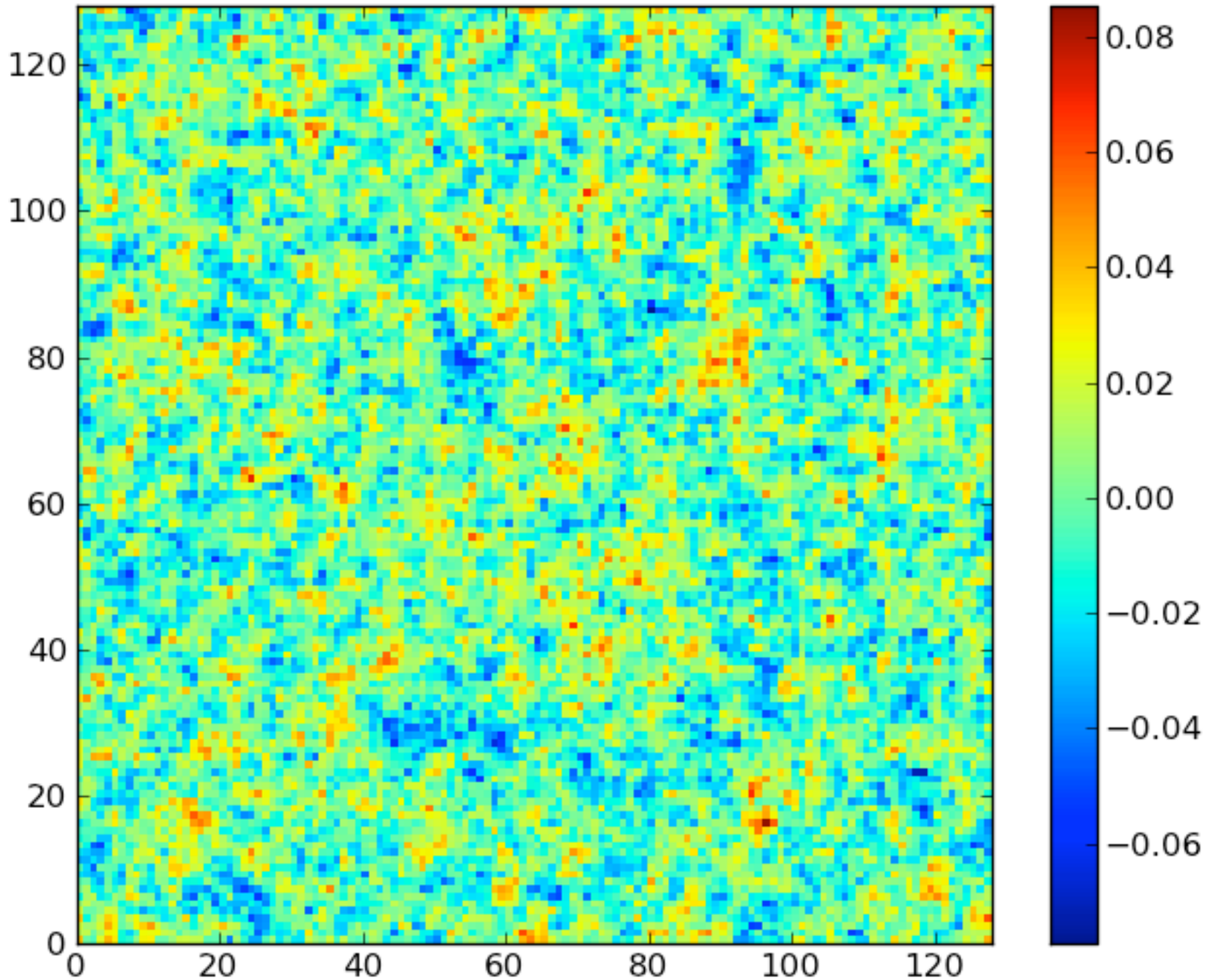
Mark Neyrinck, JHU

Power spectra: similar effect as large-scale halo biasing (Dalal et al. 2008)



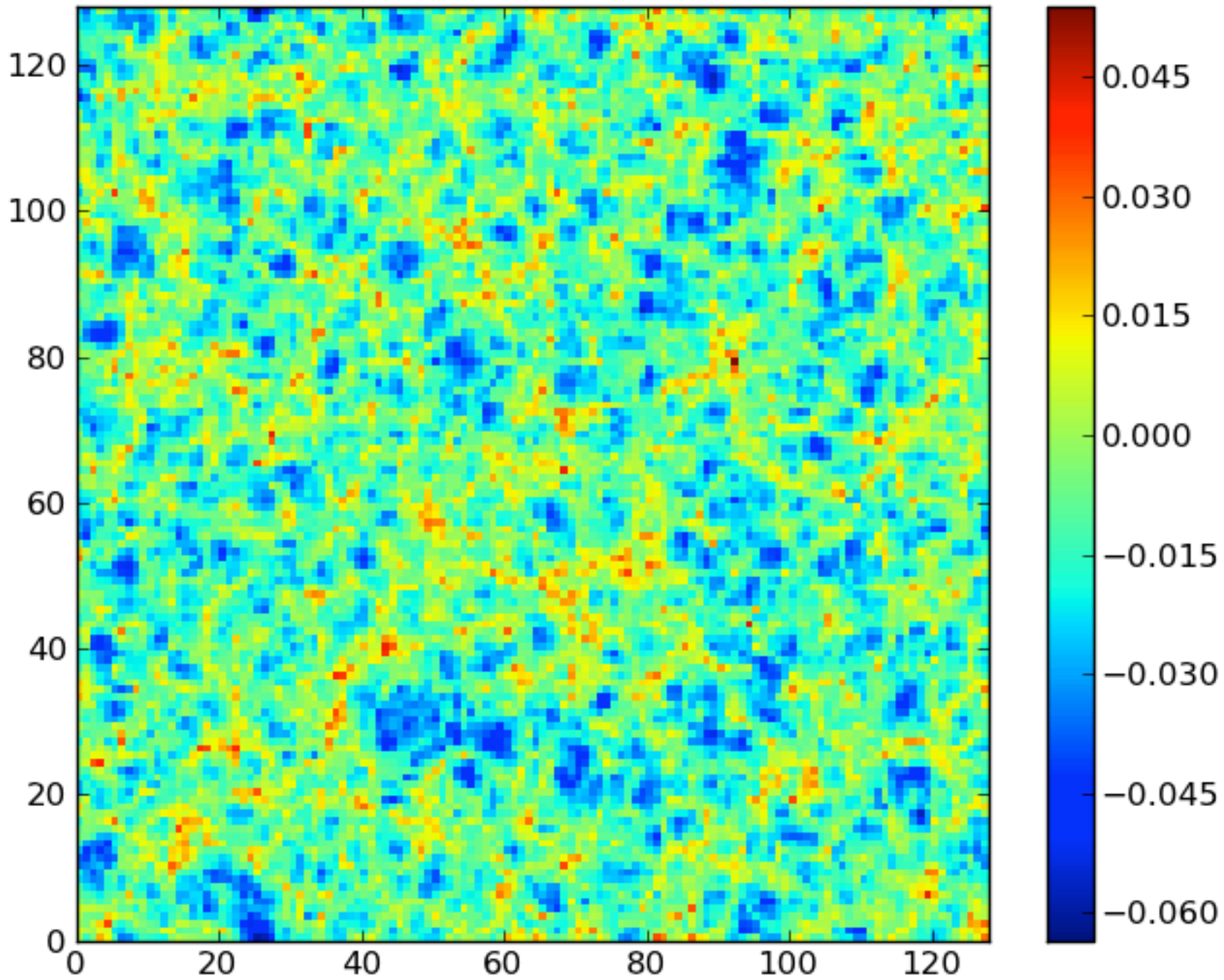
Mark Neyrinck, JHU

Now, back to analytic modeling, shifting around the density, not the mass ...



Mark Neyrinck, JHU

Now, back to analytic modeling, shifting around the density, not the mass ...



Mark Neyrinck, JHU

One motivation: BAO

- Sharp peaks move slightly together →
slight (-0.15%) bias in peak position from large-scale flows (e.g. Eisenstein et al. 2007, Matsubara 2008, Smith, Scoccimarro & Sheth 2008, Seo et al. 2008)

- Upweight smaller objects, even voids?
~volume-weighting, not ~mass weighting

- Hope: might this remove bias?, increase S/N?

Mark Neyrinck, JHU

(w/ Nuala McCullagh)

- Build on McCullagh & Szalay (2012)'s Zel'dovich-approx correlation function, where δ_{NL} is the Zel'dovich density:

$$\begin{aligned}\delta_x(\vec{x},t) &= \delta_{NL}(\vec{x} + D(t)\nabla_q\phi(\vec{q}(\vec{x}))) \\ &= (\delta_{NL}(\vec{q}) + D(t)\nabla_q\phi\cdot\nabla_q\delta_{NL}(\vec{q}) + ...)\,|_{\vec{q}=\vec{x}}\end{aligned}$$

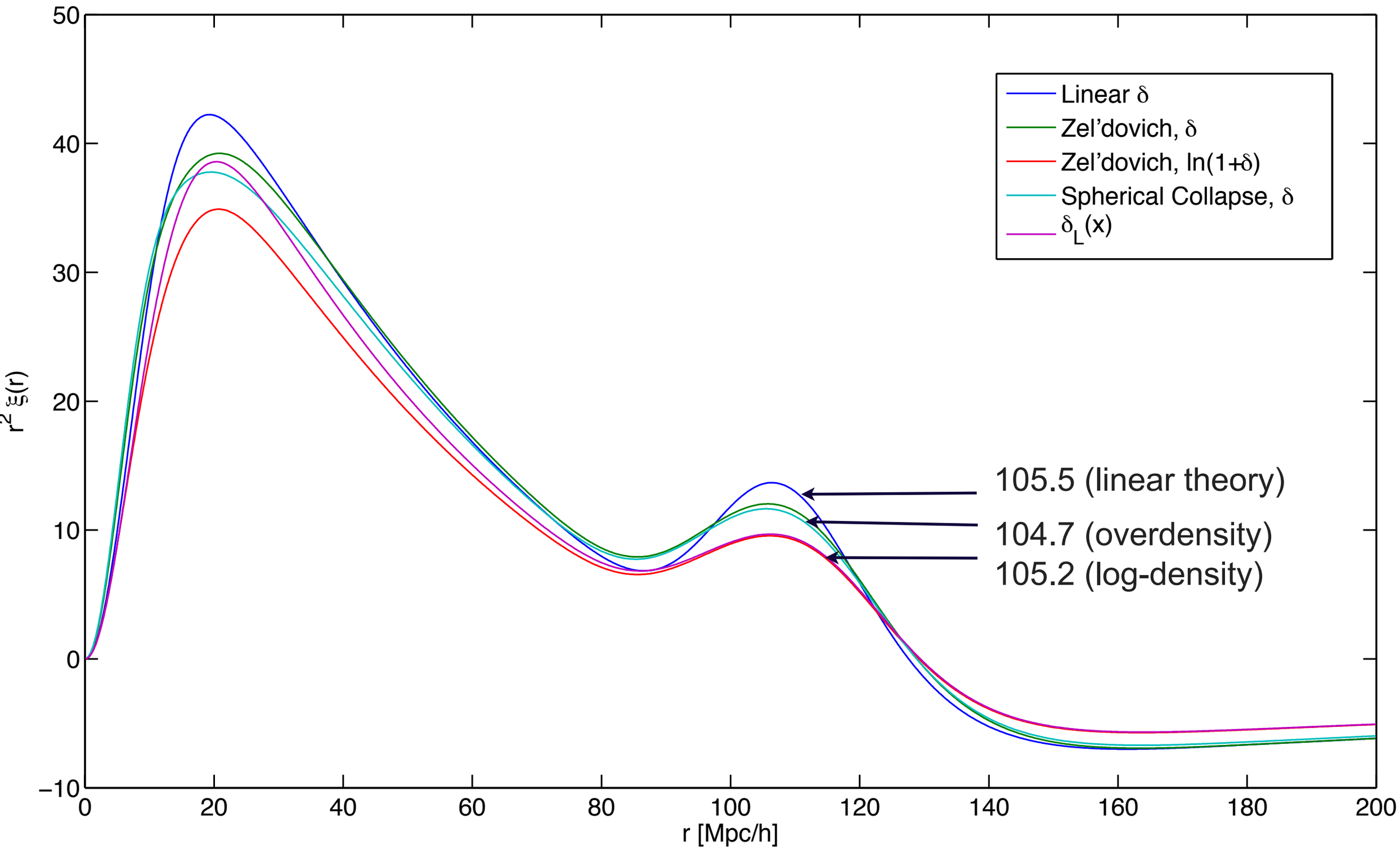
Try substituting δ_L for δ_{NL} instead. Could also try a (Taylor-expanded) nonlinear function of δ_L. Simple prescription for clustering of a nonlinearly-transformed field!
(difficult to model in general -- Wang et al. 2011). Try SC function:

$$1 + \delta_{NL}(\vec{q},t) = \frac{V_0}{V} = \left(1 - \frac{2}{3}\delta_L(\vec{q})D(t)\right)^{-3/2}$$

or the Zel'dovich log-density

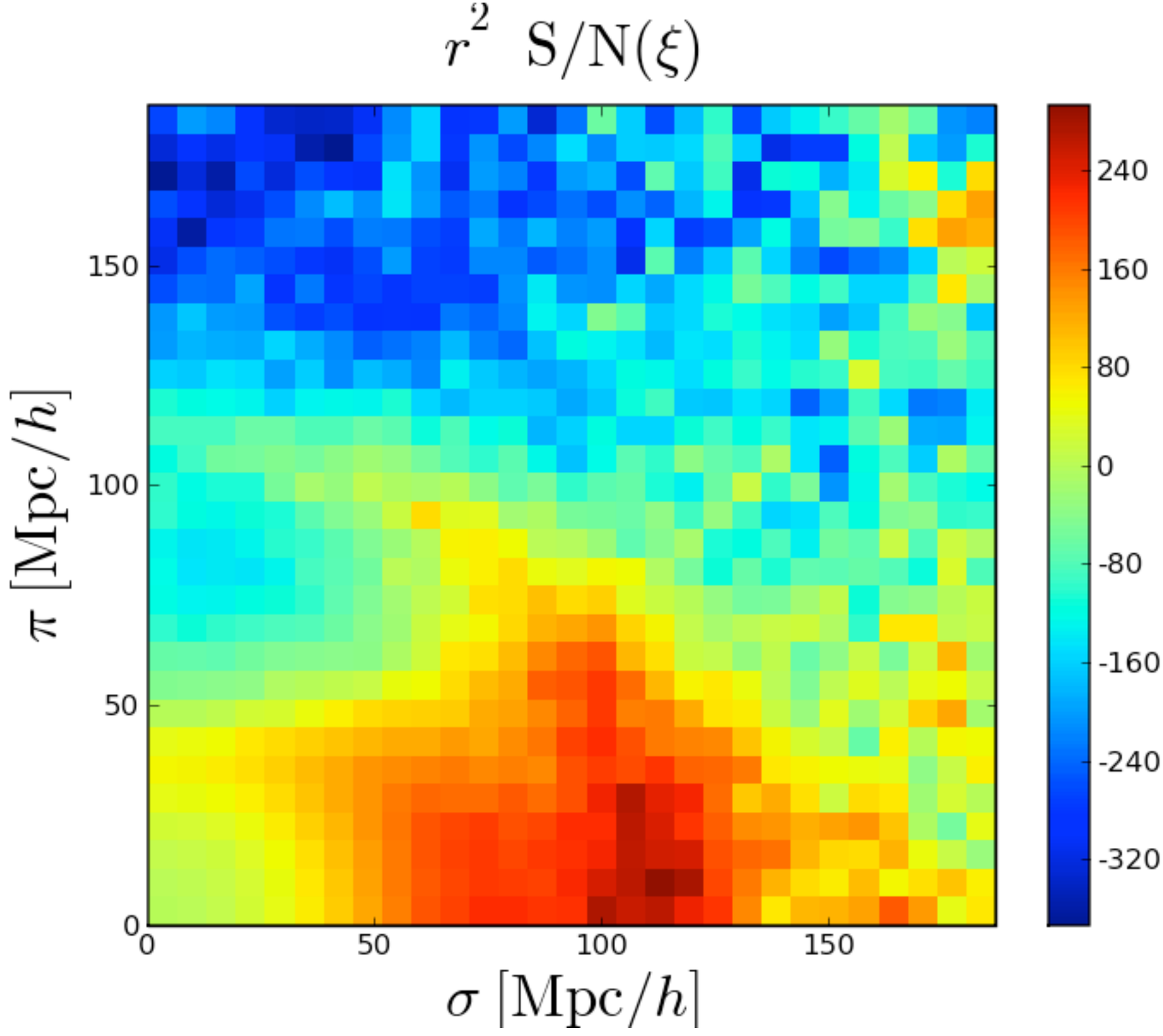
$$\begin{aligned}1 + \delta_q(\vec{q},t) &= \frac{1}{J(\vec{q},t)} \\ \ln\left(1+\delta_q(\vec{q},t)\right) &= -\ln J(\vec{q},t) \\ &= -\ln(1-D\lambda_1) - \ln(1-D\lambda_2) - \ln(1-D\lambda_3)\end{aligned}$$

“Volume-weighted” ξ’s largely debiased!



What about "reality?"

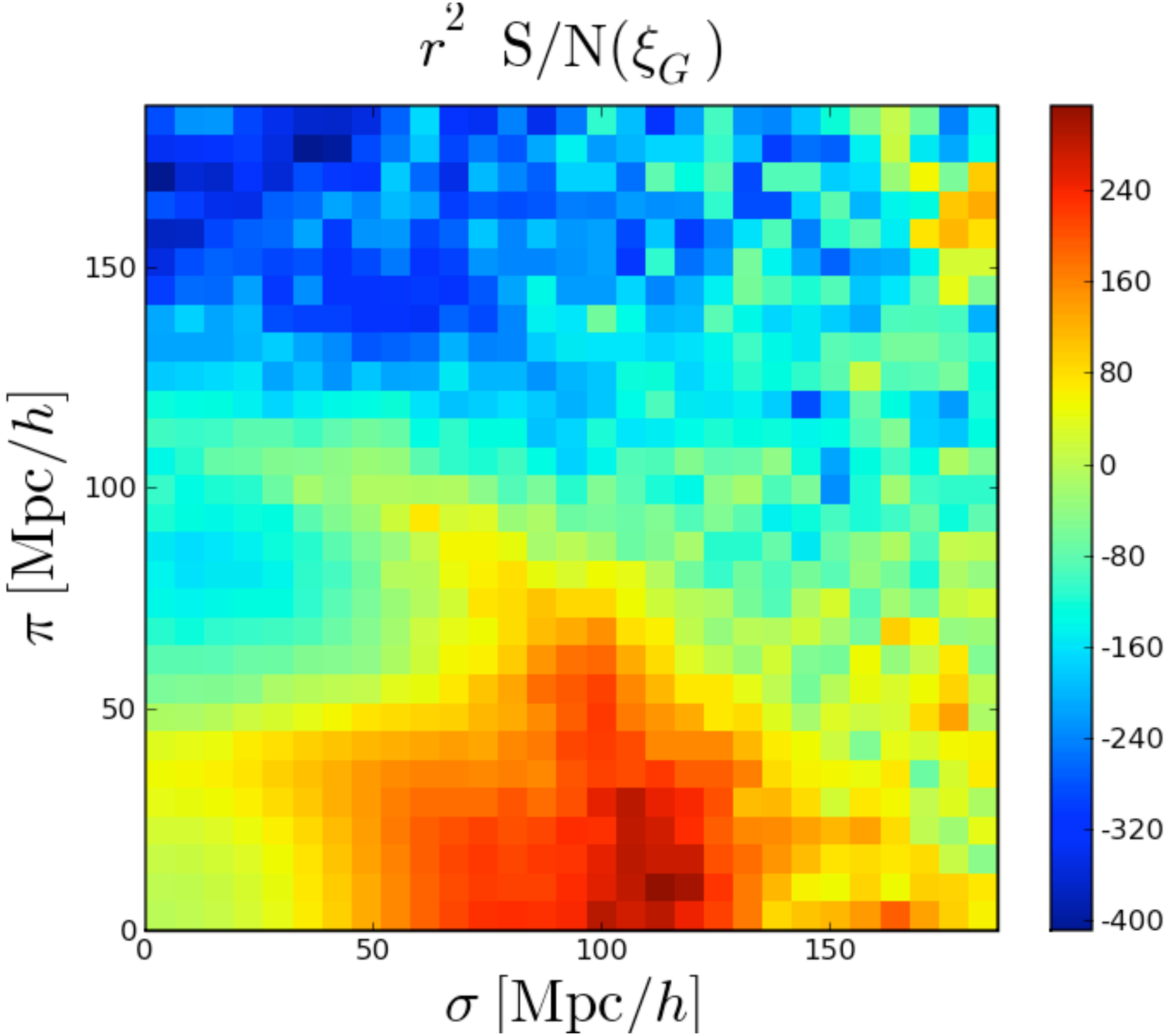
Horizon-1 LRG-like subhaloes, binned in 6.6 Mpc/h cells



Mark Neyrinck, JHU

What about "reality?"

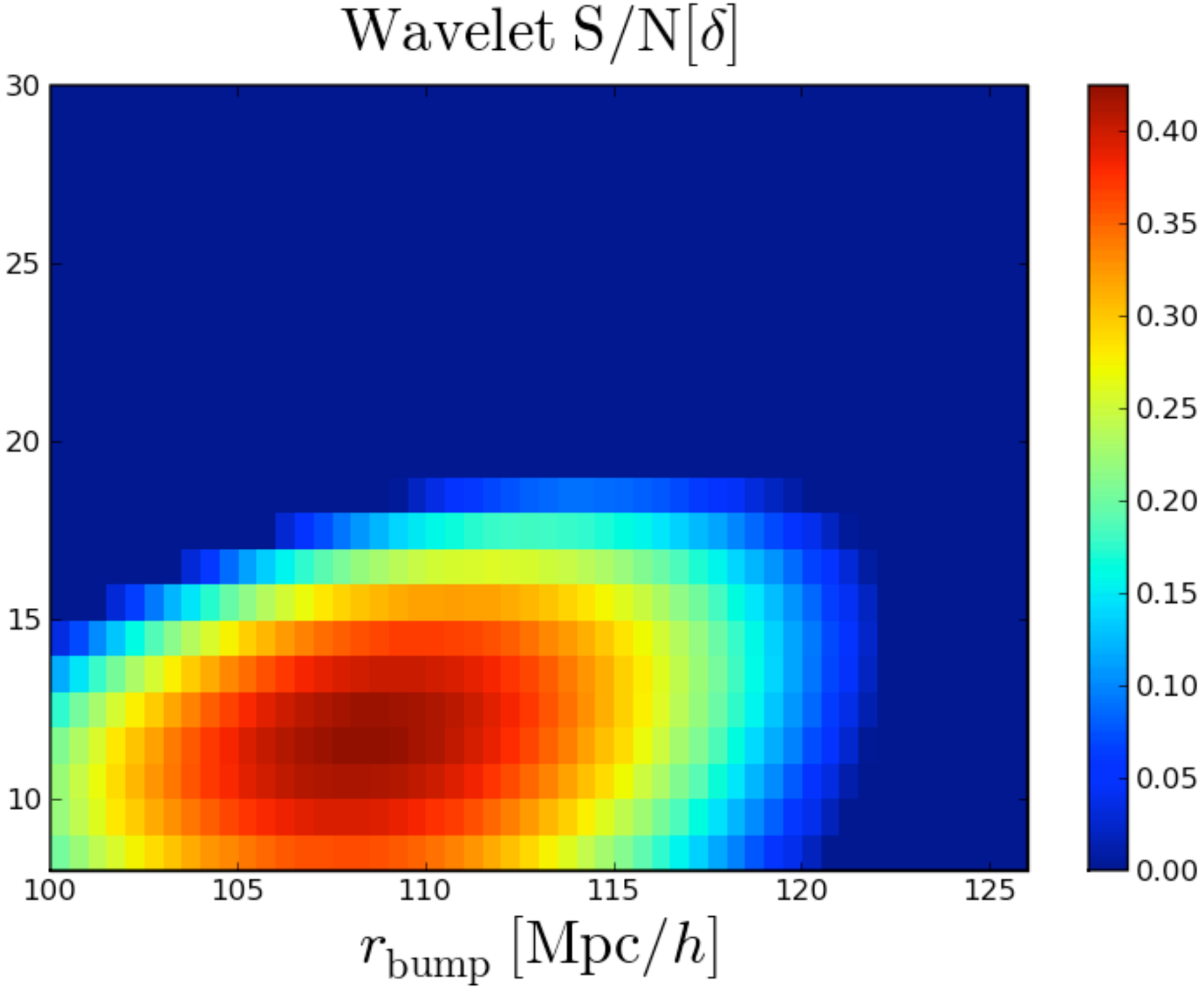
Horizon-1 LRG-like subhaloes, binned in 6.6 Mpc/h cells



Mark Neyrinck, JHU

- Small shift: +0.15% BAO bump shift in Gaussianized case. Close to negative bias in ξ_δ!

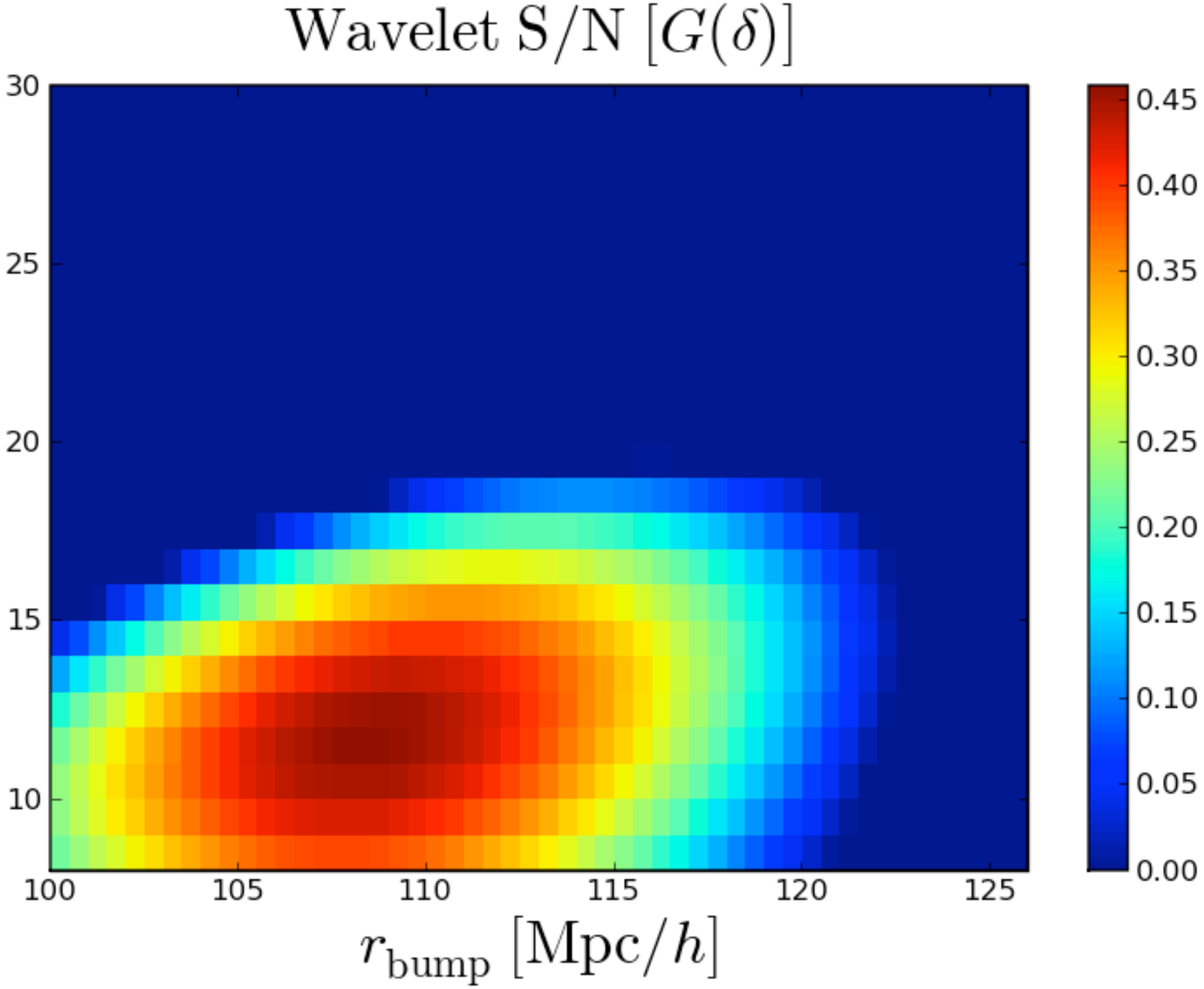
- The Gaussianized-density peak seems a bit broader



- $\Omega(r_{\rm bump}, s)$ -- Mexican-hat wavelet applied to correlation function

- Small shift: +0.15% BAO bump shift in Gaussianized case. Close to negative bias in ξ_δ!

- The Gaussianized-density peak seems a bit broader

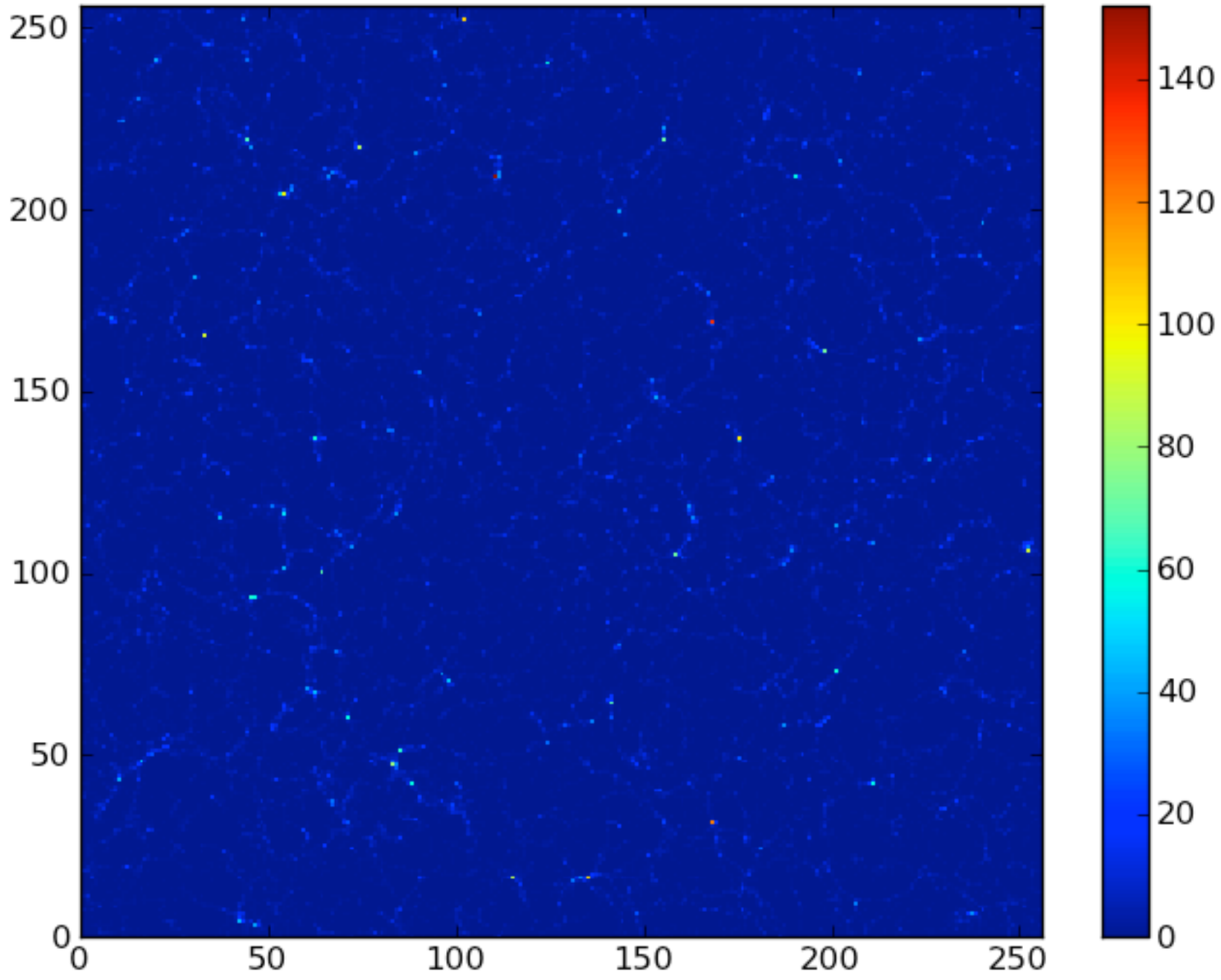


- $\Omega(r_{\rm bump}, s)$ -- Mexican-hat wavelet applied to correlation function

Benefits of Gaussianization in redshift space

Benefits of Gaussianization in redshift space

Real space



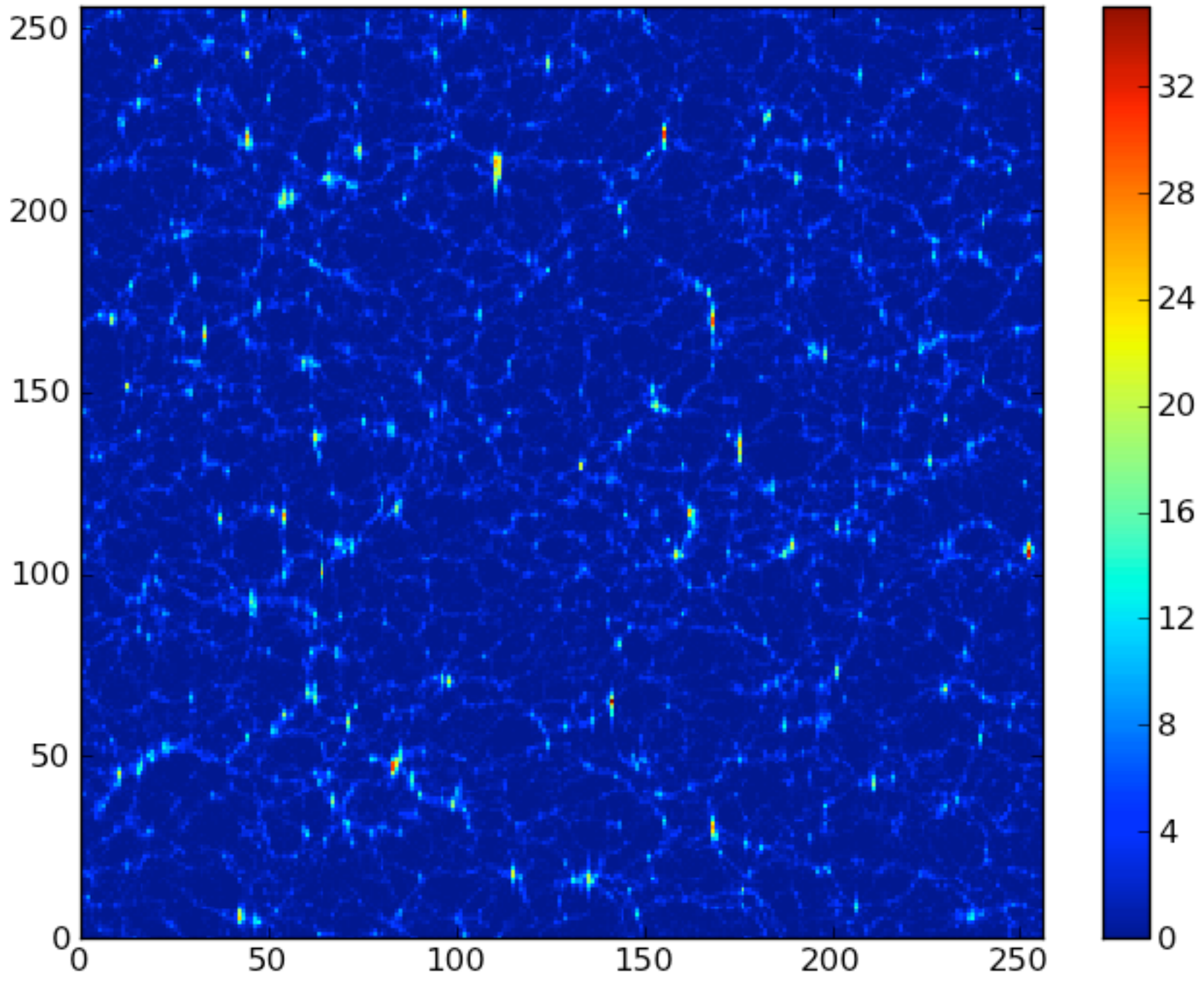
512^3 particles
1 h^{-1} Gpc box
4-h^{-1}Mpc cells

Mark Neyrinck, JHU

Redshift space

Sharp peaks become fingers of god -- tail of PDF contracts.

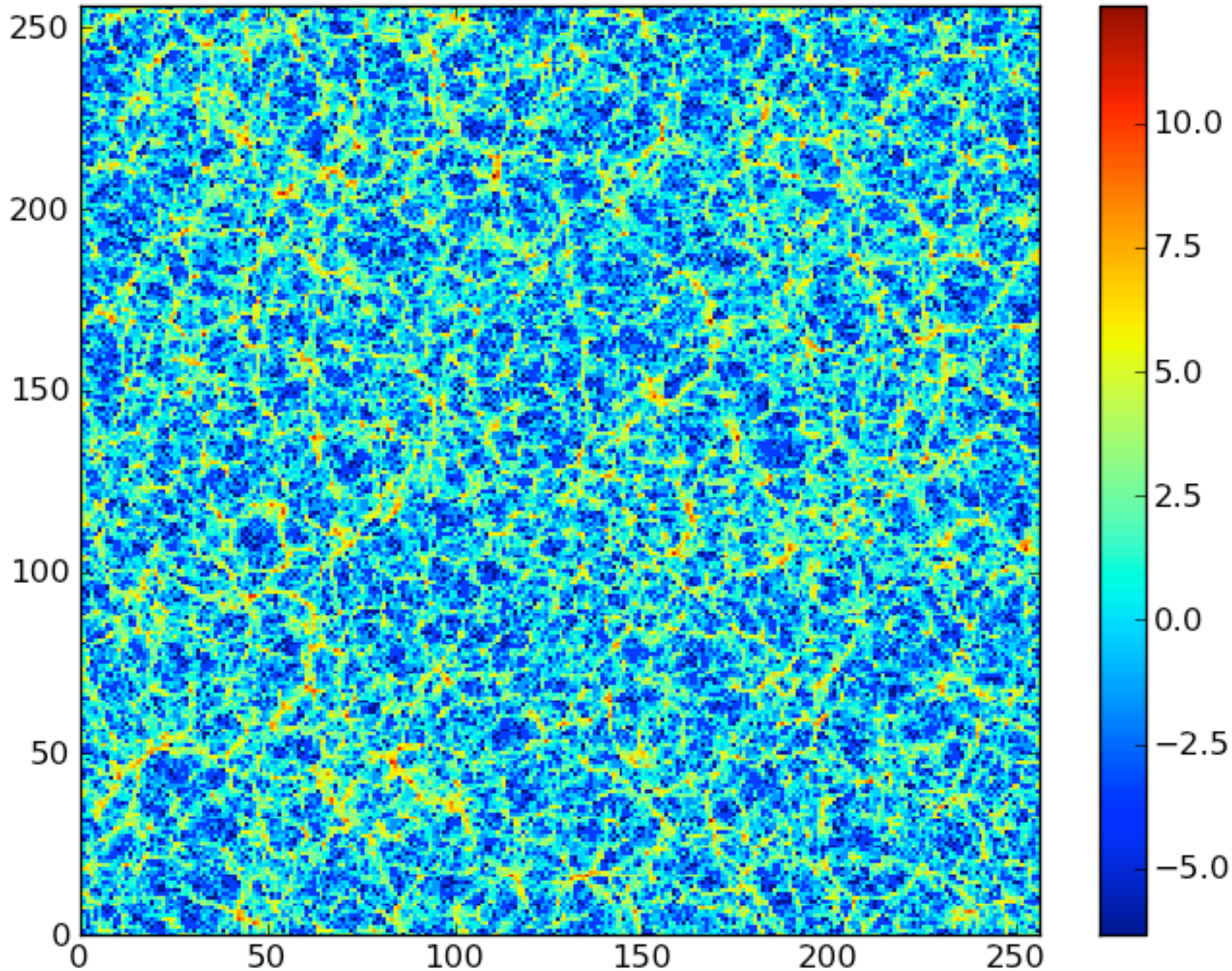
Does this mean Gaussianization will be useless?



1 h^{-1} Gpc box

Mark Neyrinck, JHU

Redshift space



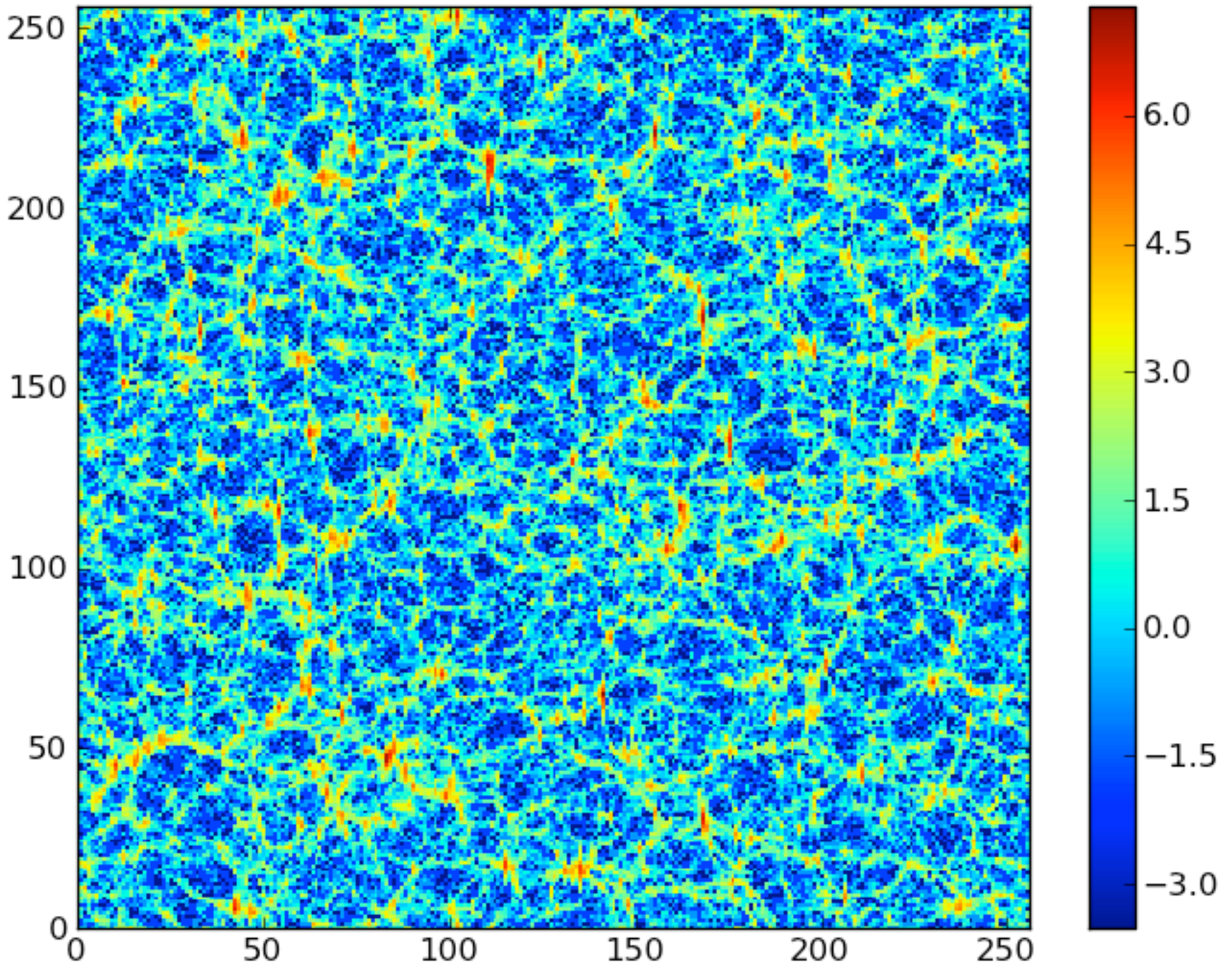
1 h^{-1} Gpc box

Mark Neyrinck, JHU

Gaussianized redshift space

- in δ, the most visible behavior is the development of FoG's.

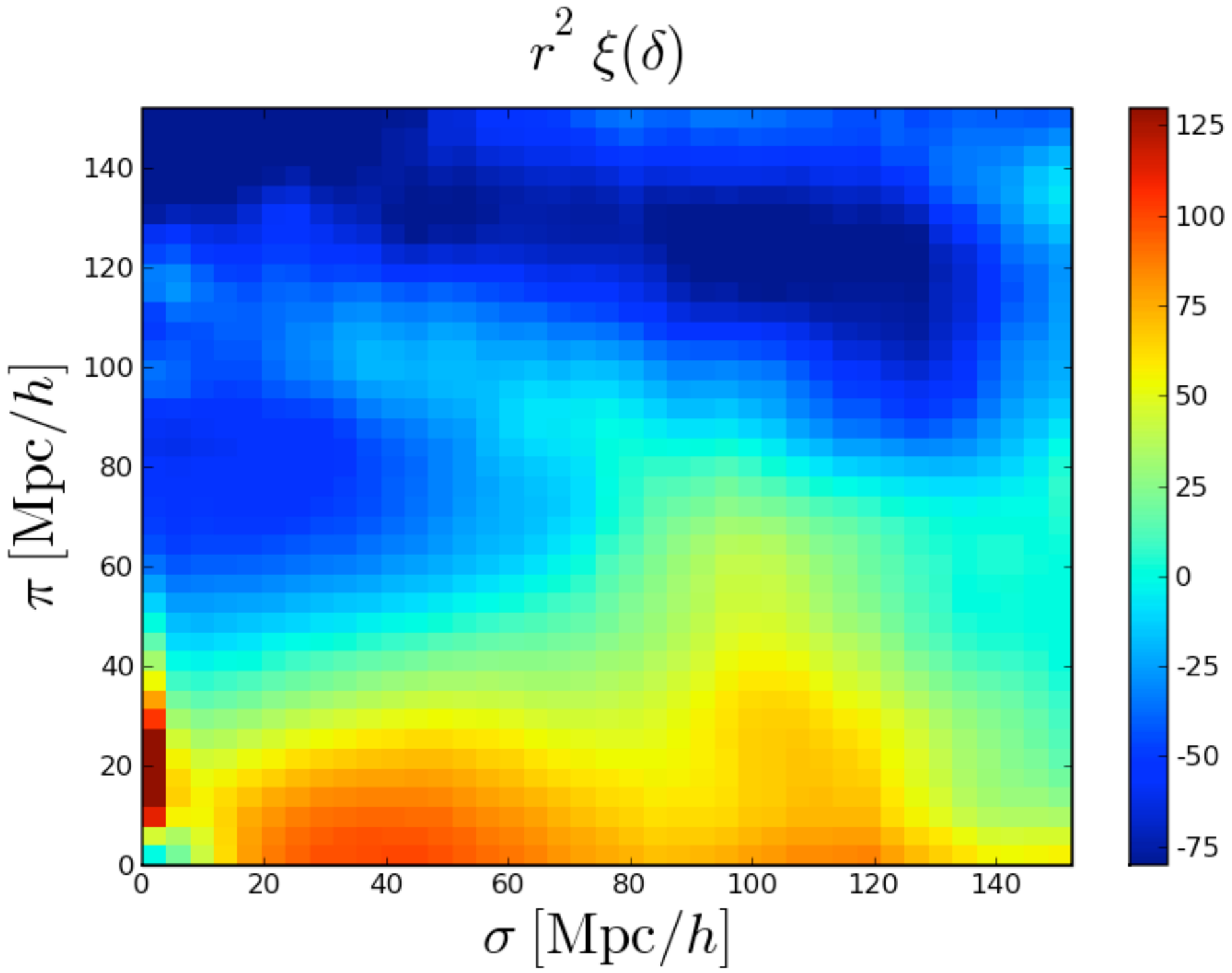
- in G(δ), on the other hand, *linear* squashing is the most visible effect.



1 h^{-1} Gpc box

Mark Neyrinck, JHU

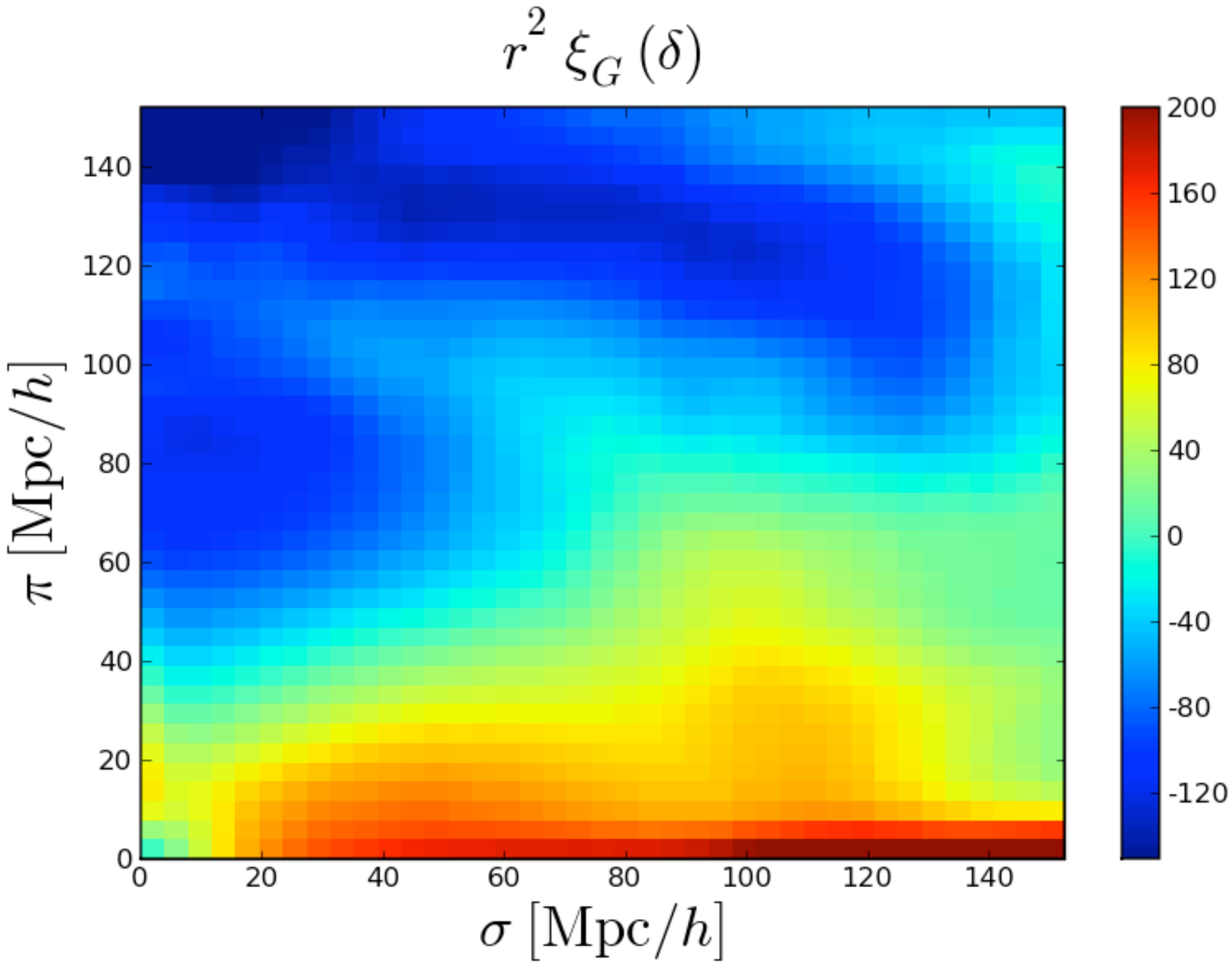
Correlation functions



Mark Neyrinck, JHU

Correlation functions

FoG's reduced compared to linear-type distortions in Gaussianized case.

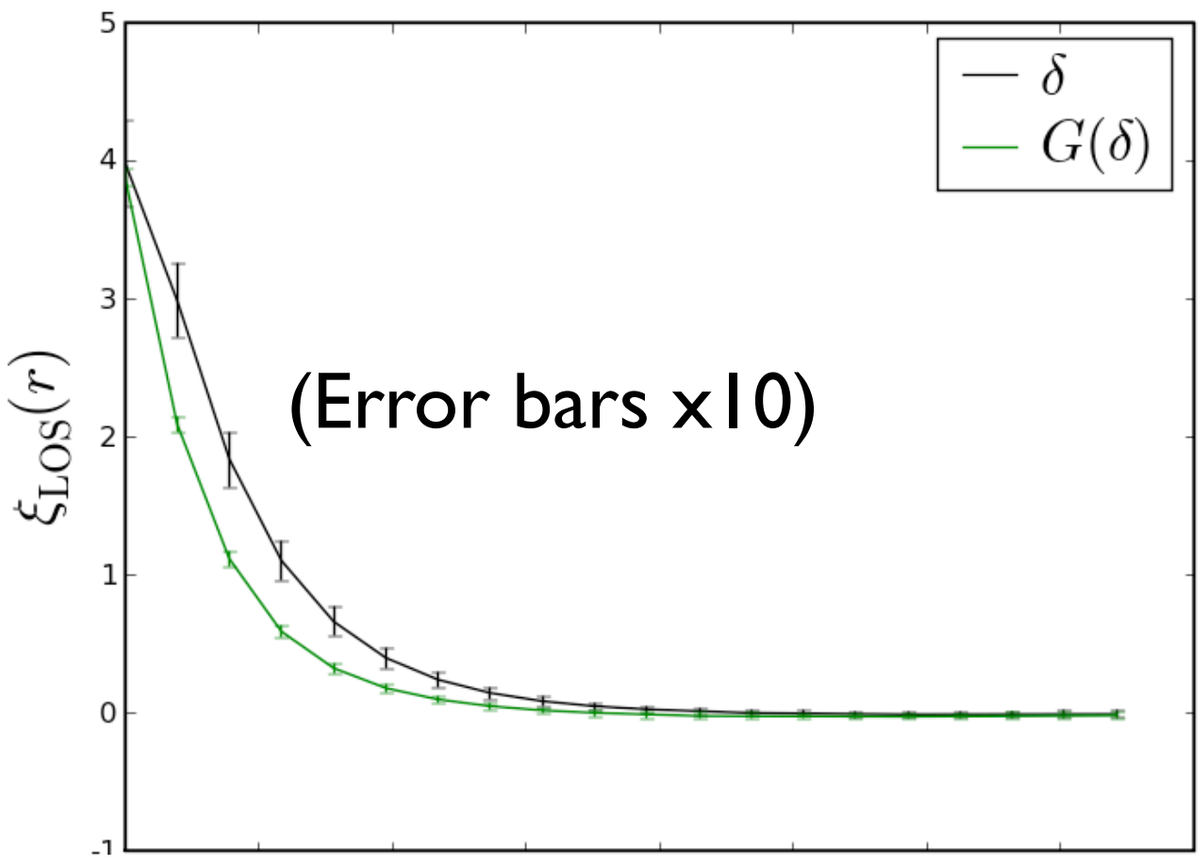


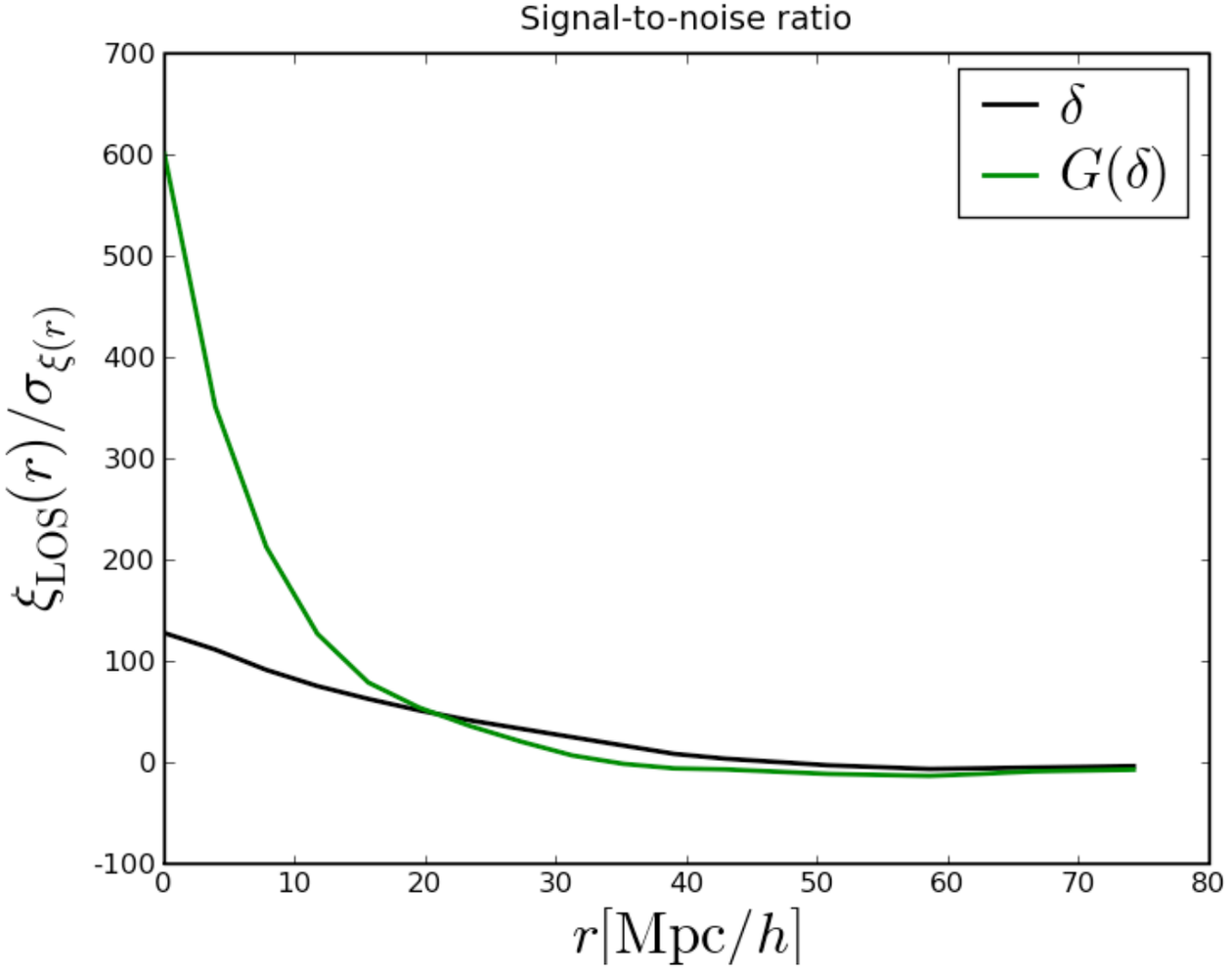
Mark Neyrinck, JHU

Line-of-sight correlation functions:

FoG's slightly less prominent in G(δ).

Error bars hugely reduced.

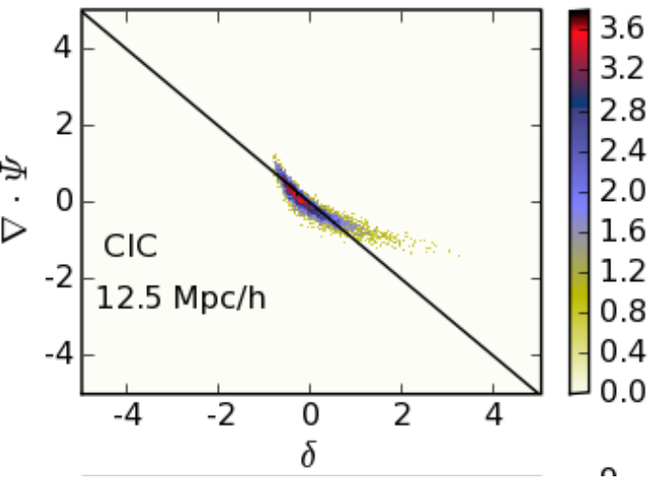


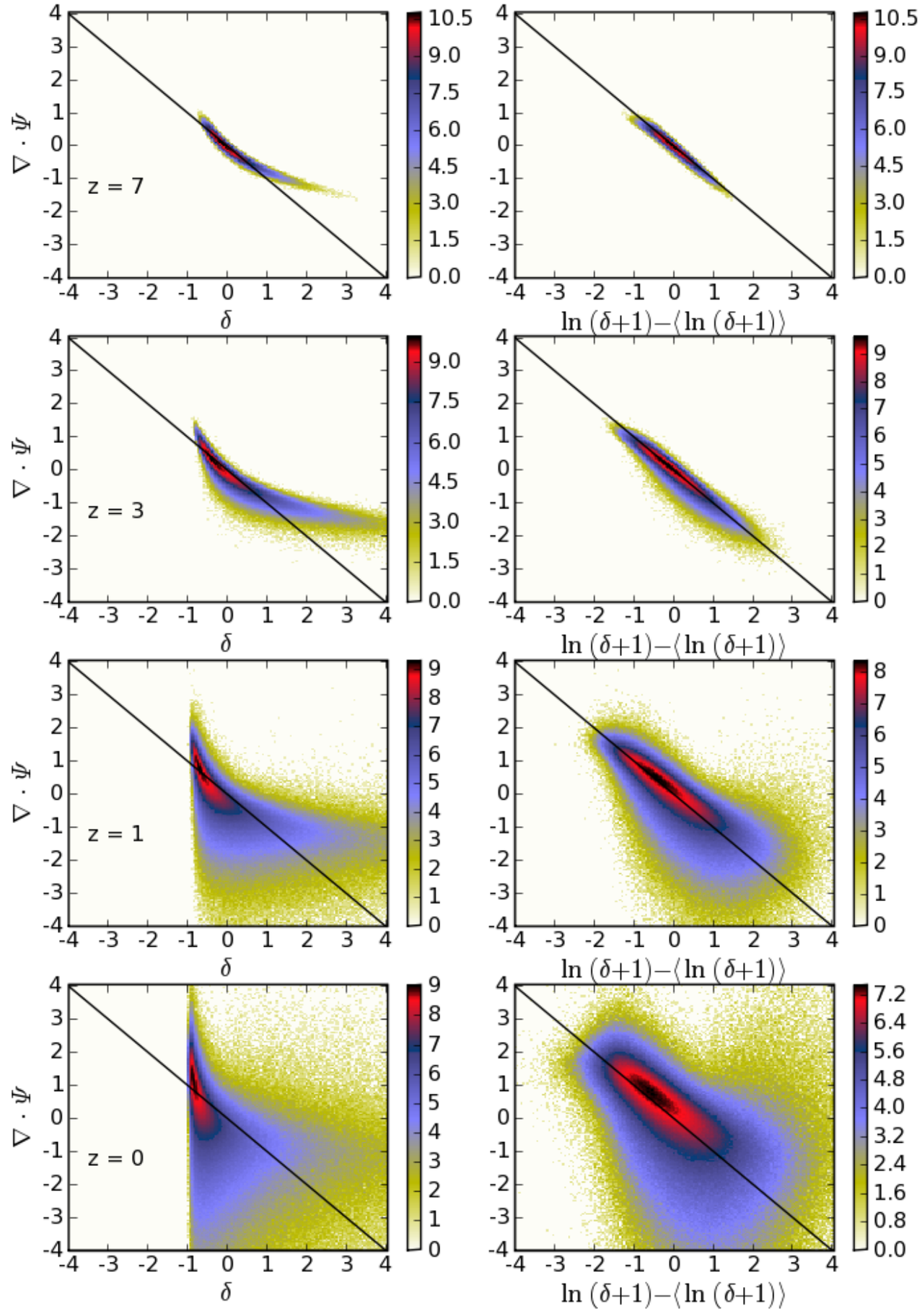


Mark Neyrinck, JHU

Displacement field better-estimated using log-density (Eulerian divergence here)

Highly nonlinear smoothing scale! (1.6 Mpc/h)





Falck, MN, et al 2012, ApJ, arXiv:1111.4466

Mark Neyrinck, JHU

Scales		Stats
Hubble radius	Many Gpc	Cosmological parameters
Linear, Gaussian scales	x > ~100 Mpc/h k < ~0.1 h/Mpc	Power spectrum
Nonlinear scales	x < ~50 Mpc/h	Morphological statistics?

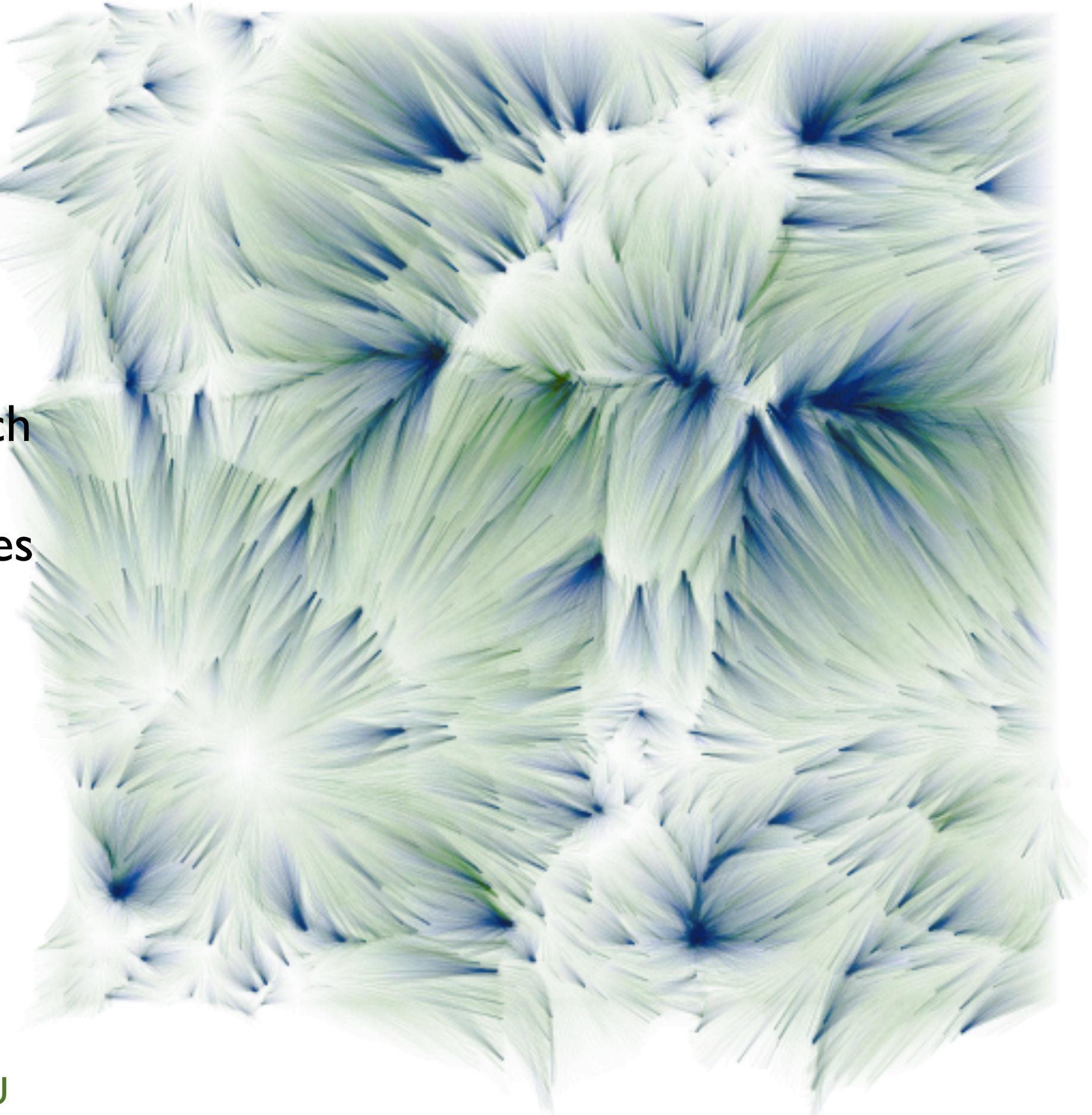
Scales		Stats
Hubble radius	Many Gpc	Cosmological parameters
Linear, Gaussian scales	x > ~100 Mpc/h k < ~0.1 h/Mpc	Power spectrum
Translinear scales	~ 0.1 h/Mpc < k < ~ 0.8 h/Mpc	Gaussianized power spectrum (my claim)
Fully nonlinear scales	x < few Mpc	Morphological statistics?

Mark Neyrinck, JHU

Conclusions

- Wrinkles more friendly in a Lagrangian approach

- When the wrinkles are not too bad, they can largely be Gaussianized away



Mark Neyrinck, JHU