A New Fast Algorithm of Cosmic Statistics for Large Data Sets

Long-Long Feng
Purple Mountain Observatory

Outline

- Introduction
- The Filtered Cosmic Fields
- A Fast Algorithm in the MRA
- Counts-in-Cells: Numerical Tests
- The 2nd Cosmic Statistics
- Applications
- Summary

Cosmic Statistics

- Power Spectrum and Two-Point Correlation Function
- Bispectrum and 3-Point Correlation
- Covariance Matrix and Trispectrum
- N-Point Correlation Functions
- Counts in Cells, Moments and Conditional Cumulants

Filtered Density Fields

$$n_W(\mathbf{x}) = \sum_{i=1}^{N} w_i W(\mathbf{x} - \mathbf{x}_i)$$

Density Distribution:

$$n(\mathbf{x}) = \sum_{i=0}^{N} w_i \delta^3(\mathbf{x} - \mathbf{x_i})$$

$$n_W(\mathbf{x}) = \int W(\mathbf{x} - \mathbf{x}')n(\mathbf{x}')d^3\mathbf{x}'$$

Window Functions

- **Spherical Top Hat:**

$$W_{sphere}(r,R) = \frac{1}{(4\pi/3)R^3}\theta(R-r)$$

$$\hat{W}_{sphere}(k,R) = \frac{3}{k^3R^3}(\sin(kR) - kR\cos(kR))$$

- **Cubic Top Hat**

$$W_{cubic}(\mathbf{x},\mathbf{L}) = \frac{1}{L_xL_yL_z}\theta(L_x/2-|x|)\theta(L_y/2-|y|)\theta(L_z/2-|z|)$$

$$\hat{W}_{cubic}(\mathbf{k},\mathbf{L}) = \frac{\sin(k_xL_x)}{k_xL_x}\frac{\sin(k_yL_y)}{k_yL_y}\frac{\sin(k_zL_z)}{k_zL_z}$$

Window Functions

- **Spherical Shell Filter:**

$$W_{shell}(r, R) = \frac{1}{4\pi r^2}\delta_D(R - r)$$

$$W_{shell}(k, r) = \frac{\sin(kr)}{kr}$$

$$W_{shell}(\cdot, R) = \frac{1}{3R^2}\frac{d}{dR}(R^3 W_{sphere}(\cdot, R))$$

Window Functions

- **Cylinder Filter:**

$$W_G(\rho, z, R, h) = \frac{1}{\pi R^2 h}\theta(R-\rho)\theta(h/2-|z|)$$

$$\hat{W}_G(k_\perp, k_z, R, h) = \frac{\sin(k_z h/2)}{k_z h/2}\int_0^1 J_0(k_\perp R\sqrt{x})dx$$

- **Gaussian Filter:**

$$W_G(r, R) = \frac{1}{\sqrt{(2\pi)^3}R^3}\exp(-\frac{r^2}{2R^2})$$

$$\hat{W}_G(k, R) = \exp(-\frac{1}{2}k^2 R^2)$$

Multi-Resolution Analysis (MRA)

Express an arbitrary function at various levels of resolution, which forms a sequence of functional spaces

$$0 \subset \cdots \subset V_{-1} \subset V_0 \subset \cdots \subset L^2(\mathbf{R})$$

$$V_0 \quad \{\phi(x-k) | k \in \mathbf{Z}\}$$

$$V_j \quad \{\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k) \quad | k \in \mathbf{Z}\}$$

Dilation

Translation

Density Field in the MRA

In terms of scaling functions, we may make a decomposition of the density distribution in the MRA at a scale j

$$n(x) = \sum w_i \delta(x - x_i)$$

$$= \sum_l s_l^j \phi_{j,l}(x)$$

$$s_l^j = \int n(x)\phi_{j,l}(x)dx$$

$$= \sum_{i=1}^{N_p} w_i 2^{j/2} \phi(2^j x_i - l)$$

Operator in the MRA

for a kernel $W(x, y)$, projection onto V_j

yields a multiresolution representation

$$W \longrightarrow W_j(x, y) = \sum_{l,m} w^j_{l,m} \phi_{j,l}(x) \phi_{j,m}(y)$$

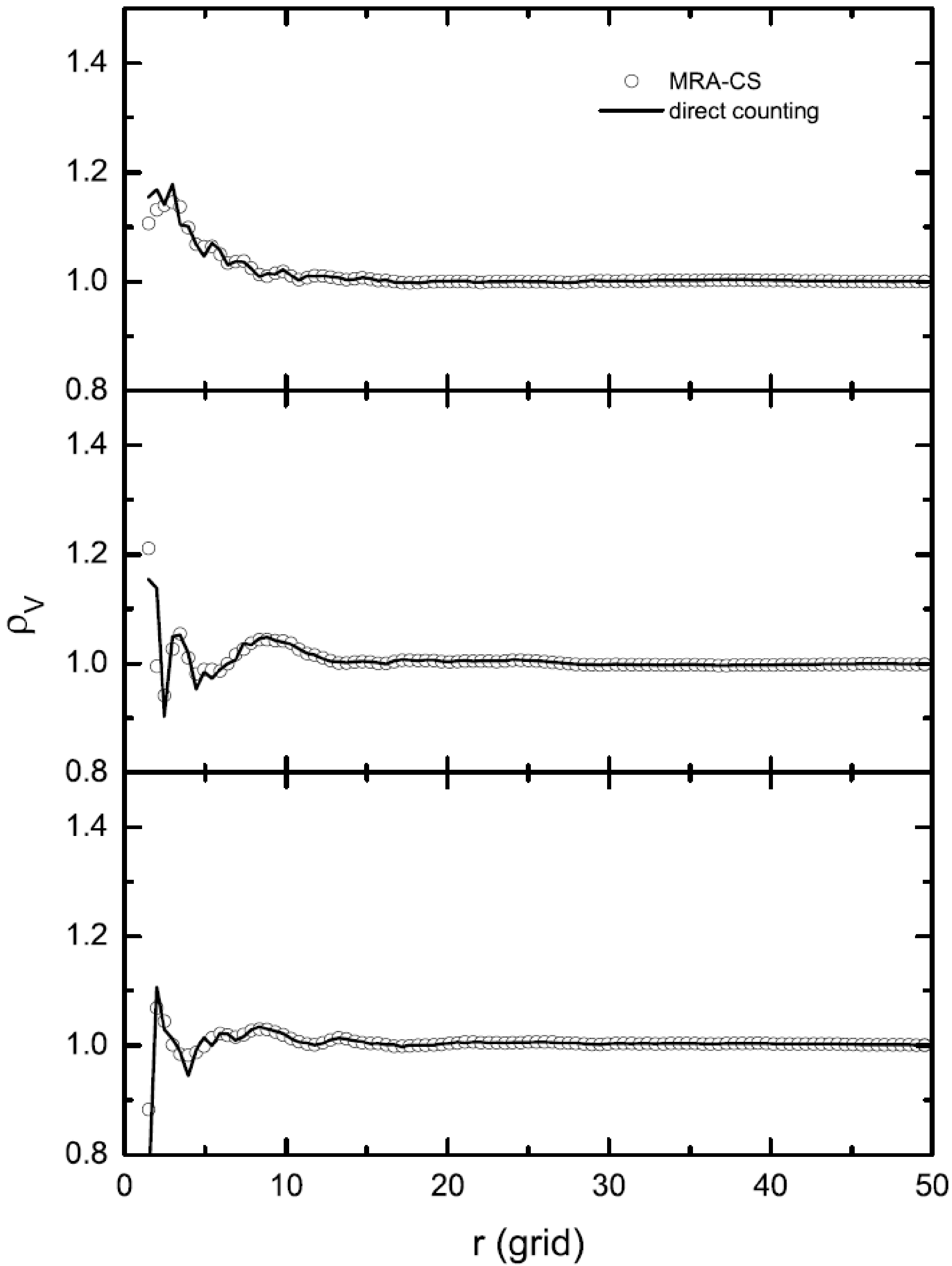
$$w^j_{l,m} = \int W(x, y) \phi_{j,l}(x) \phi_{j,m}(y) dx dy$$

Summation Rule

$$n_W(\mathbf{x}) = \sum_{i=1}^{N} w_i W(\mathbf{x} - \mathbf{x}_i)$$

$$n_W(x) \rightarrow n_W^j(x) = \sum_{l} \tilde{s}_l^j \phi_{j,l}(x)$$

$$\tilde{s}_l^j = \sum_{m} w_{l,m}^j s_m^j$$

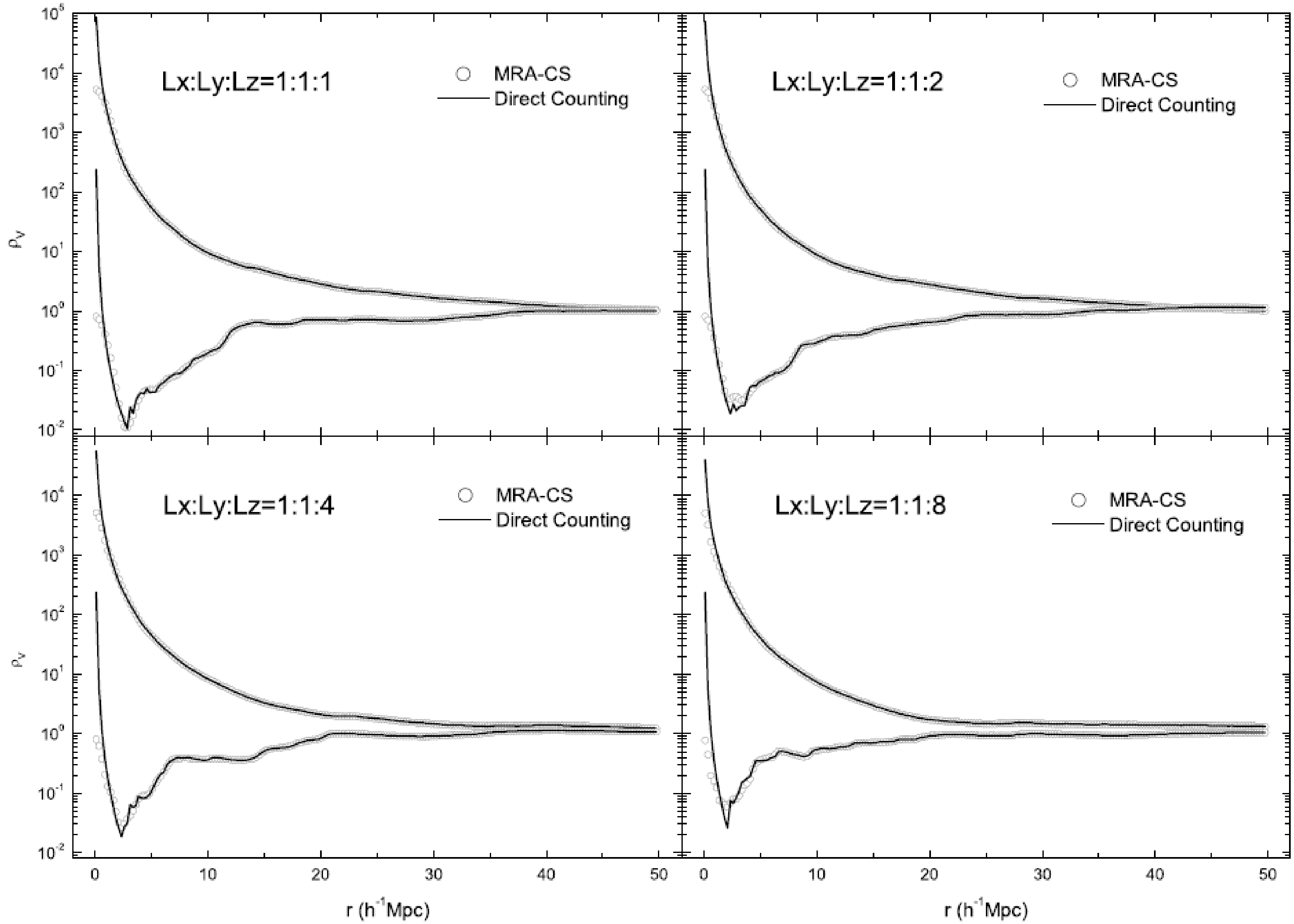


Random Sample

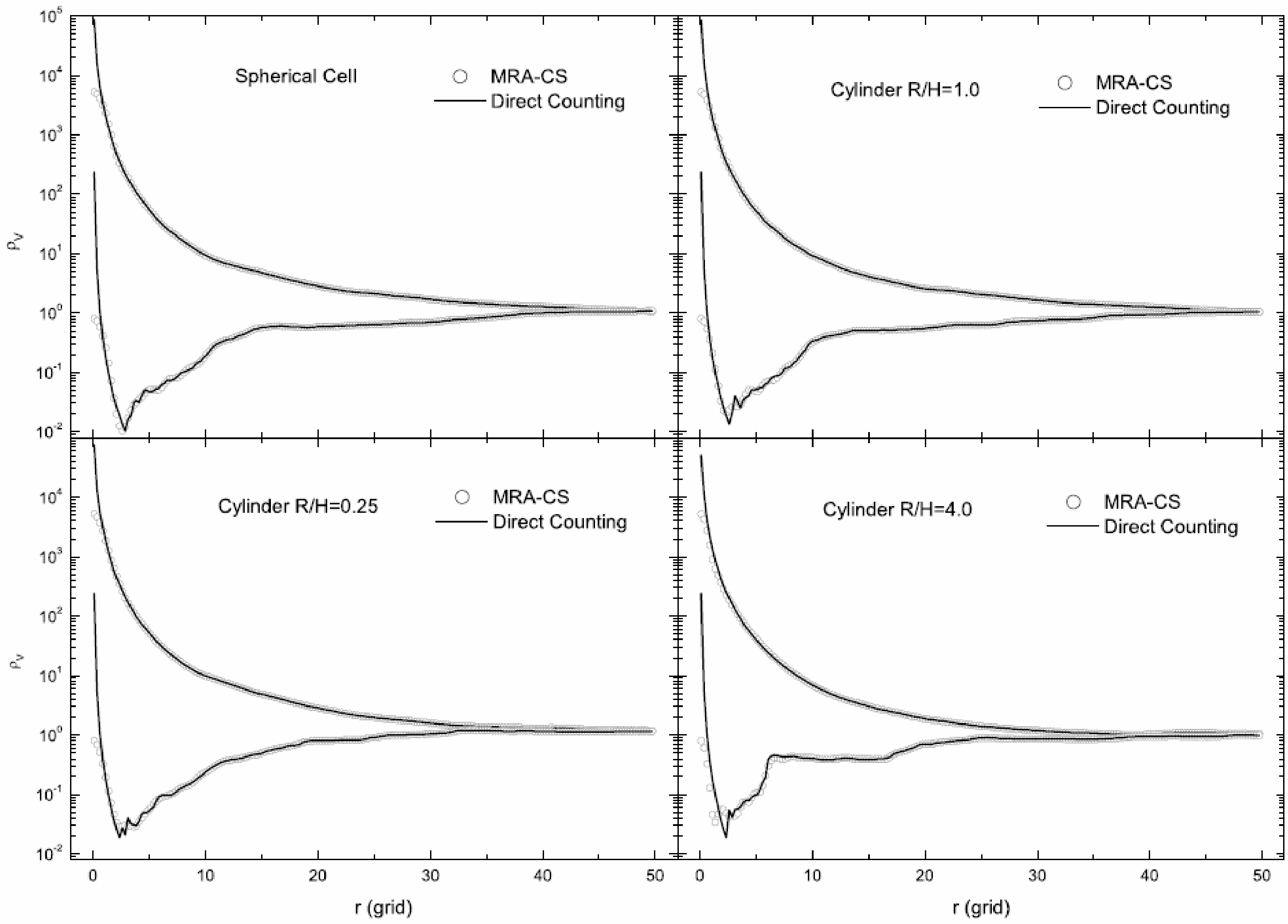
256^3 particles

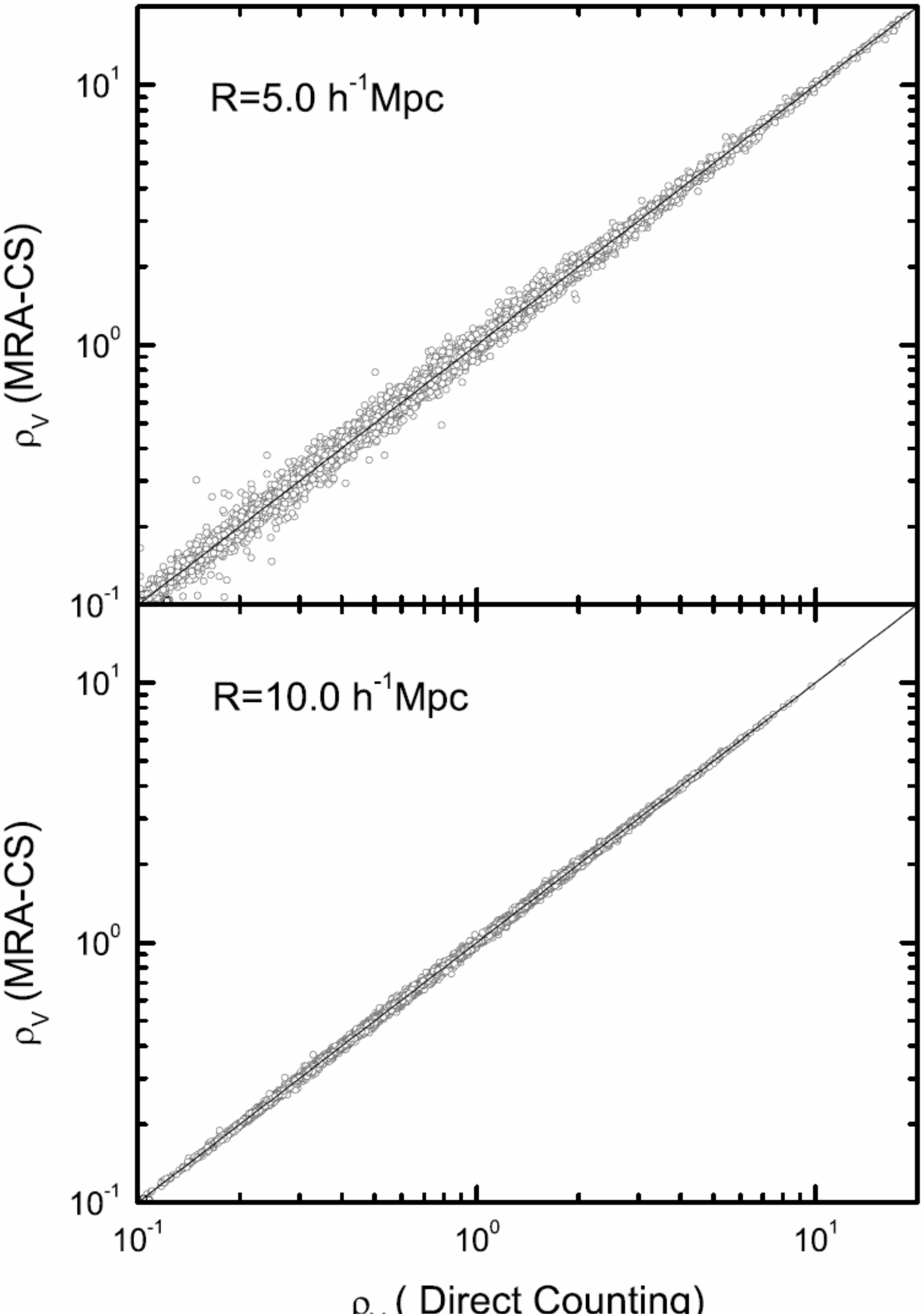
in a 256^3 grid

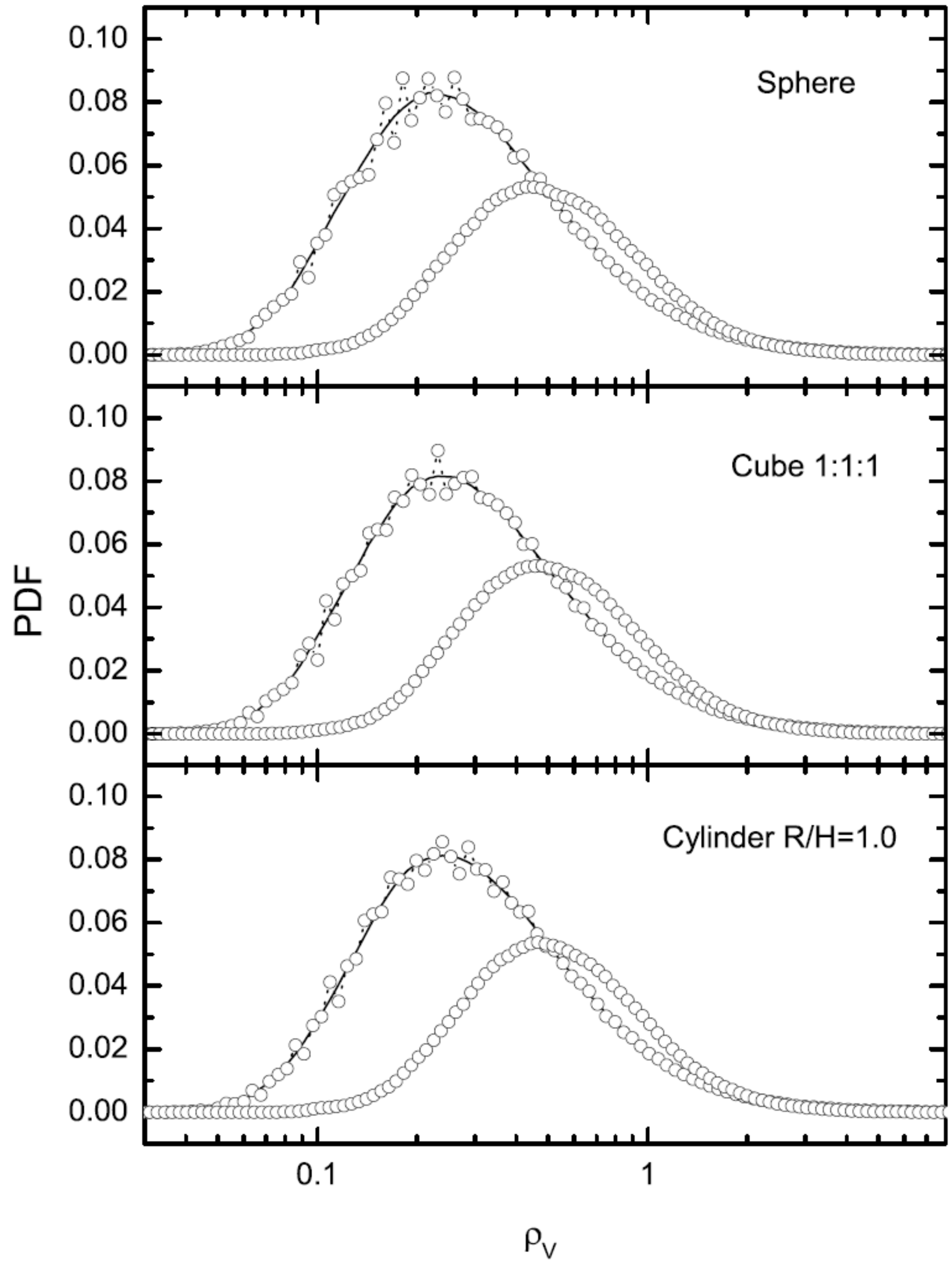
Cubic Counts-in-Cells in the Virgo Simulation Sample (LCDM Model)



Spherical and Cylinder Counts-in-Cells in the Virgo Simulation Sample (LCDM Model)




R=5.0 h^{-1}Mpc
R=10.0 h^{-1}Mpc
ρ_V (MRA-CS)
ρ_V (MRA-CS)
ρ_V (Direct Counting)
10^{-1}
10^{0}
10^{1}



Sphere
Cube 1:1:1
Cylinder R/H=1.0
PDF
0.00
0.02
0.04
0.06
0.08
0.10
0.1
1
ρV

Correlation Functions

- **Fourier Pair:**

$$\xi(r) = \frac{1}{2\pi^2}\int_0^\infty \hat{W}_{shell}(k,r)P(k)k^2dk$$

- **LS Estimator:**

$$\hat{\xi}_{LS}(r) = \frac{DD - 2DR + RR}{RR}$$

- **MRA-CS Scheme:**

$$DD = \sum_l \tilde{s}_l^j s_l^j = \tilde{\mathbf{s}}^j \cdot \mathbf{s}^j$$

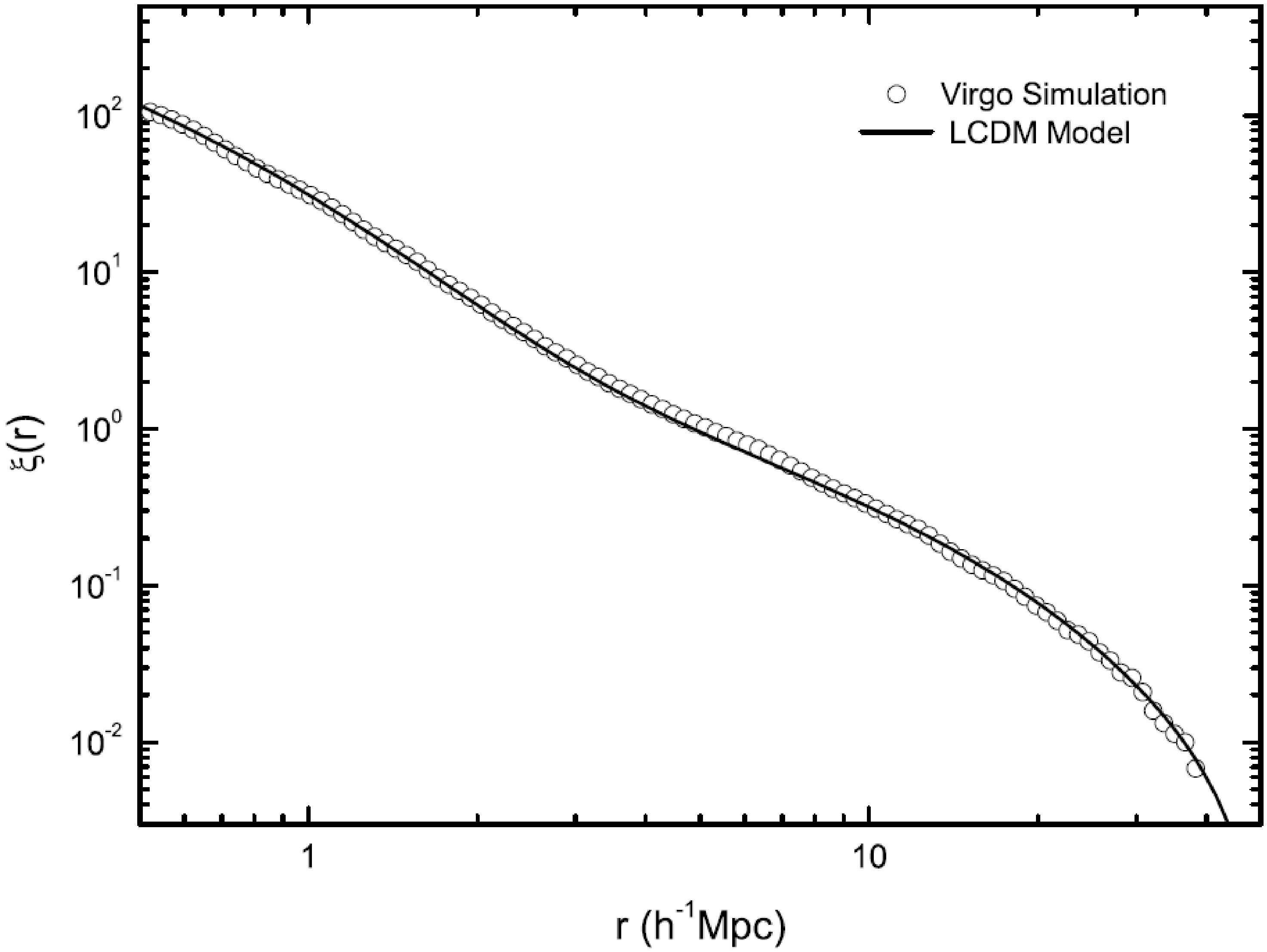
2nd Variances of the Filtered Field

$$\sigma^2(\cdot) = < \delta_W^2(\cdot) > = \frac{1}{(2\pi)^3} \int |W_{filter}(\mathbf{k}, \cdot)|^2 P(k) d^3\mathbf{k}$$

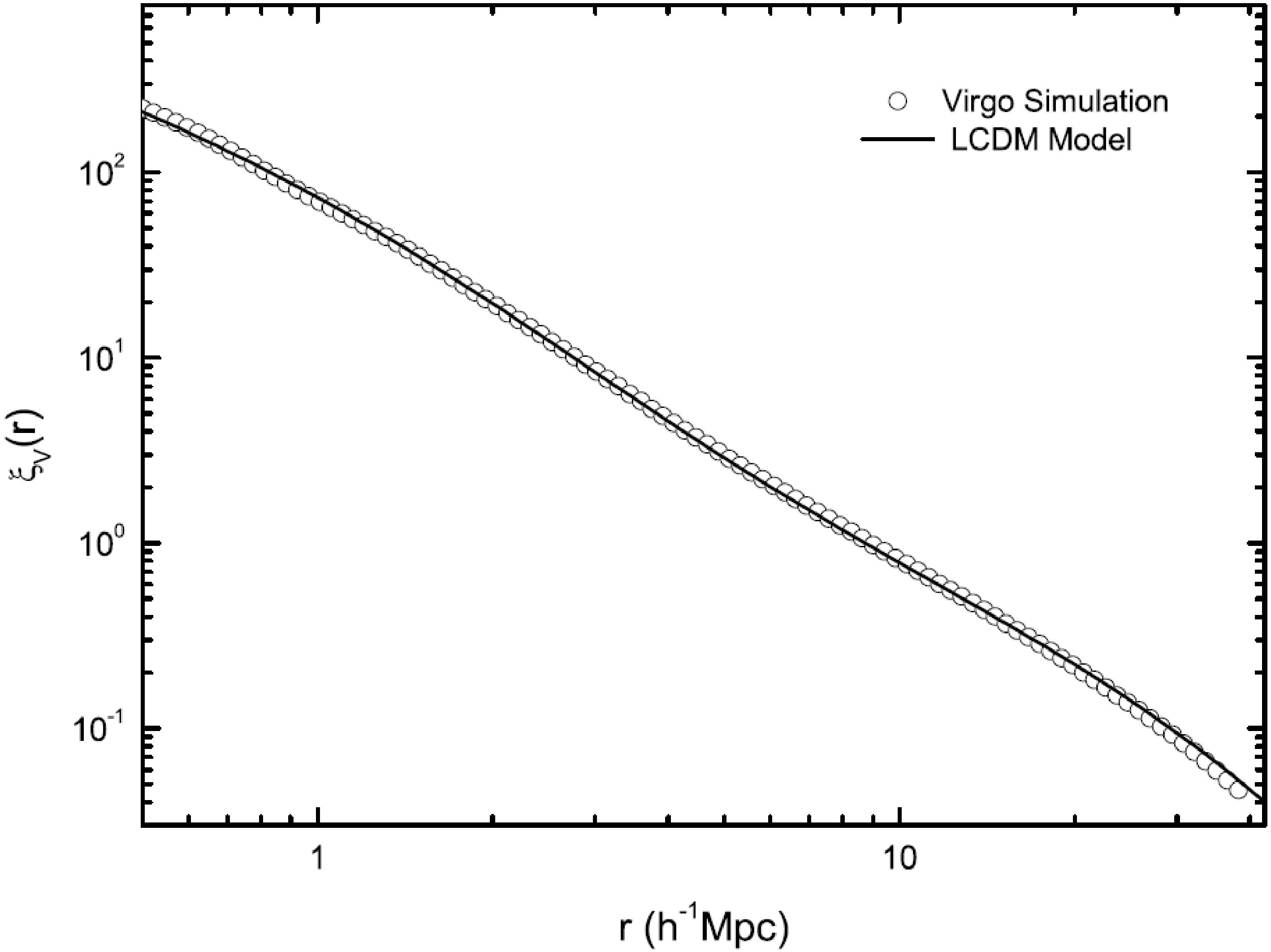
$$\sigma^2(\cdot) = \sum_l \tilde{s}_l^j \tilde{s}_l^j = |\tilde{\mathbf{s}}^j|^2 - 1$$

dot denotes for a set of parameters specifying the spatial geometry of window functions

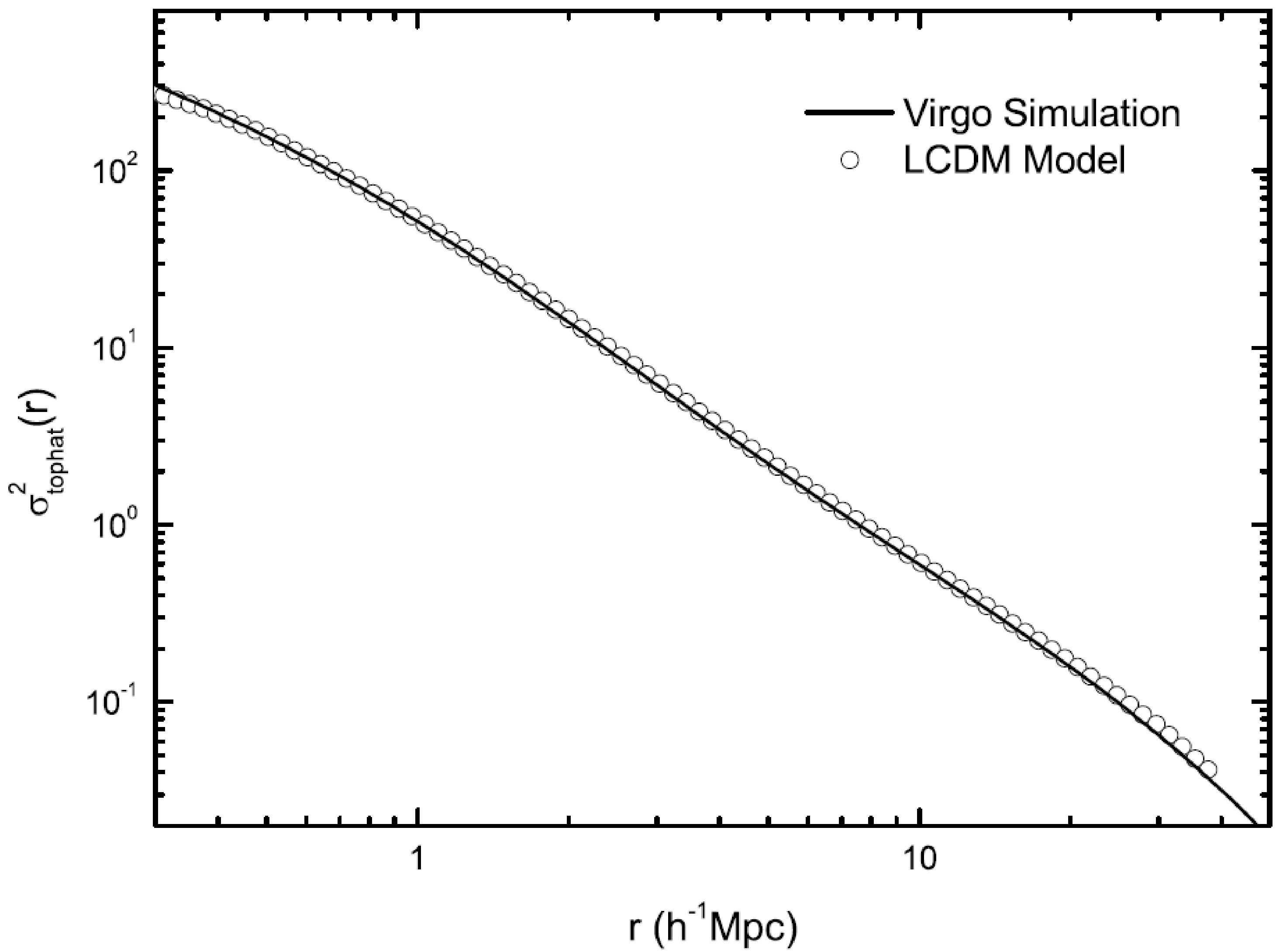
The Two-Point Correlation Function



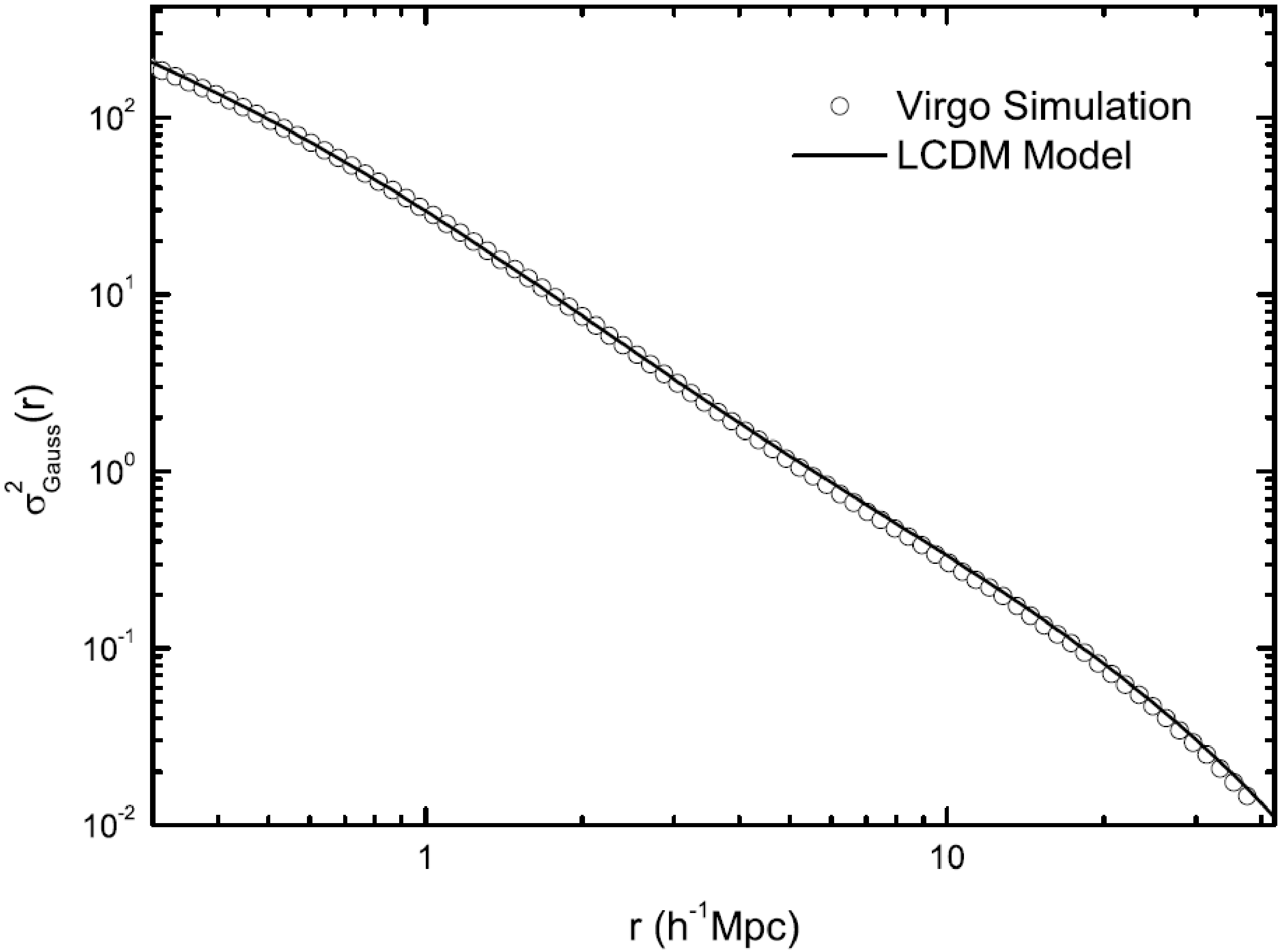
The Integral Two-Point Correlation Function



The 2nd Top-hat Filtered Variance



The 2nd Gaussian Filtered Variance



Multiresolution Analysis for Cosmic Statistics

- **Grid-based algorithm via FFT technique**

  scaled with O(*N*log*N*)
- **Computational cost: independent of size/shape of cells**

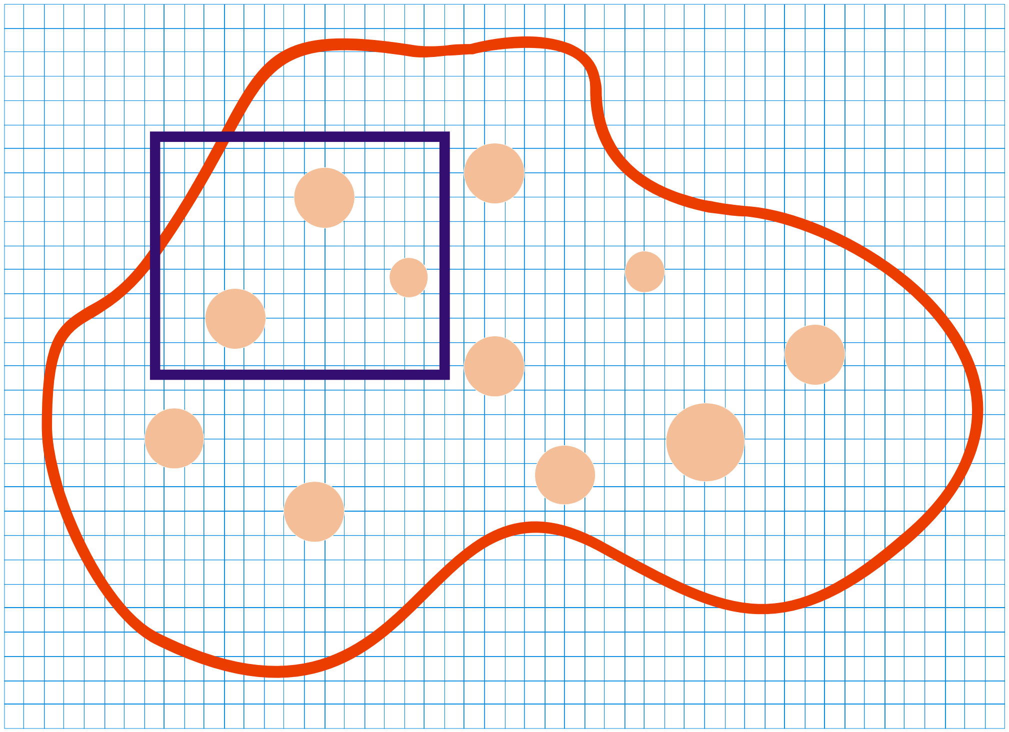
  shape dependence of high order statistics (topology, morphology and bias)
- **Easy parallelized**

  massive sampling

Application: Estimating Irregular Spatial Volume

Volume Incompleteness in real galaxy surveys

- bright star mask
- complicated survey geometry



Application: Gravity Solver

Gravitational Potential:

$$\Phi(\mathbf{r}) = \int \frac{G\rho(\mathbf{x})}{|\mathbf{r}-\mathbf{x}|} d^3\mathbf{x}$$

Decomposition in the MRA

$$\rho(\mathbf{x}) = \sum_{l_x,l_y,l_z} \epsilon_{j,l_x l_y l_z} \phi_{jl_x}(x)\phi_{jl_y}(y)\phi_{jl_z}(z)$$

$$\begin{aligned}\frac{1}{r} &= \frac{2}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-r^2 e^{2s}+s} ds \\ &= \sum_{i=1}^{M} w_i e^{-\alpha_i r^2} = \sum_{i}^{M} w_i e^{-\alpha_i x^2} e^{-\alpha_i y^2} e^{-\alpha_i z^2}\end{aligned}$$


Thank you !
天文臺