New methods in the analysis of galaxy distance catalogues

Guilhem Lavaux

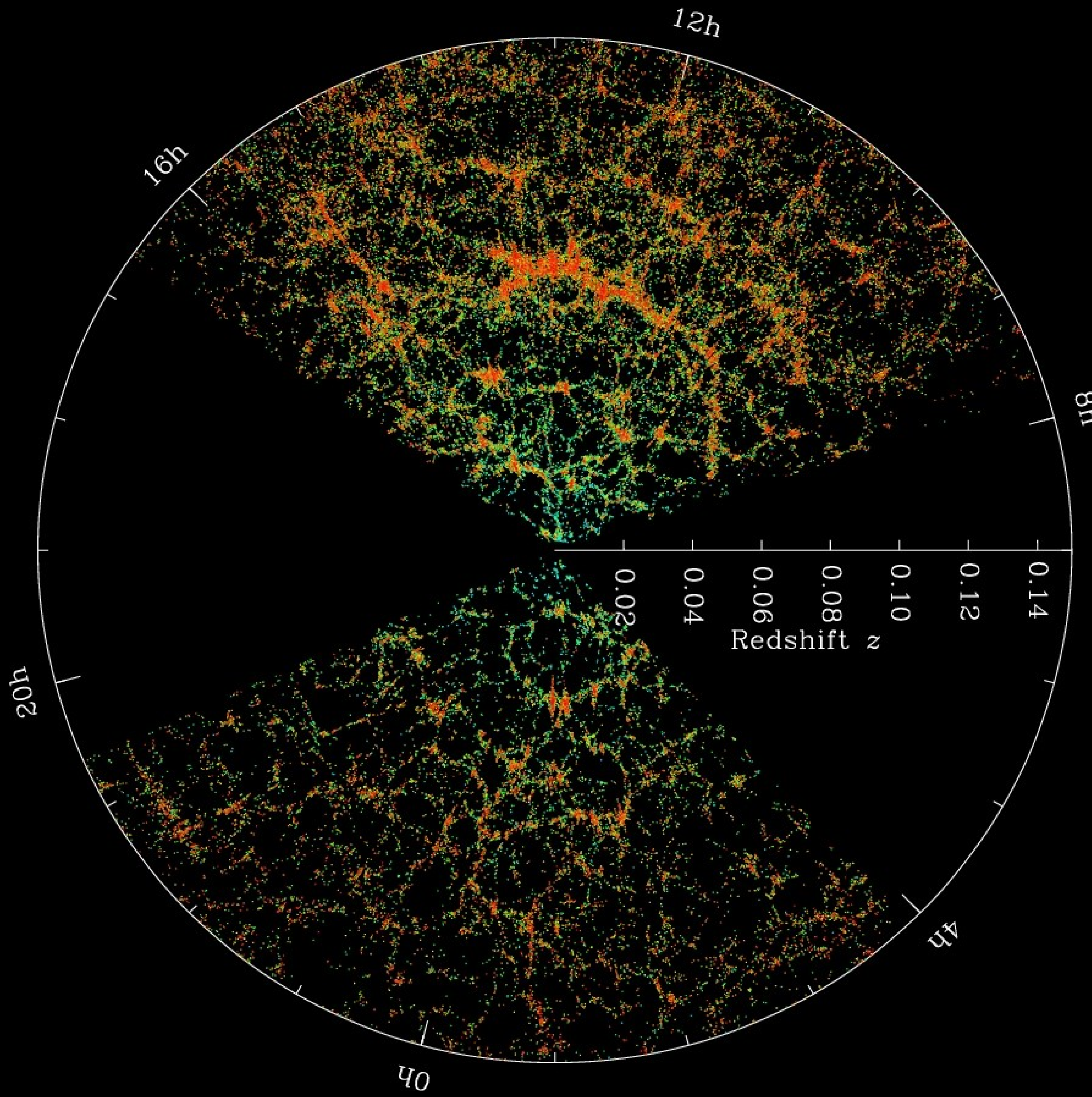
Institut d'Astrophysique de Paris

+ collaborators/students: Michael J. Hudson, Jonathan Carrick, Stephen Turnbull

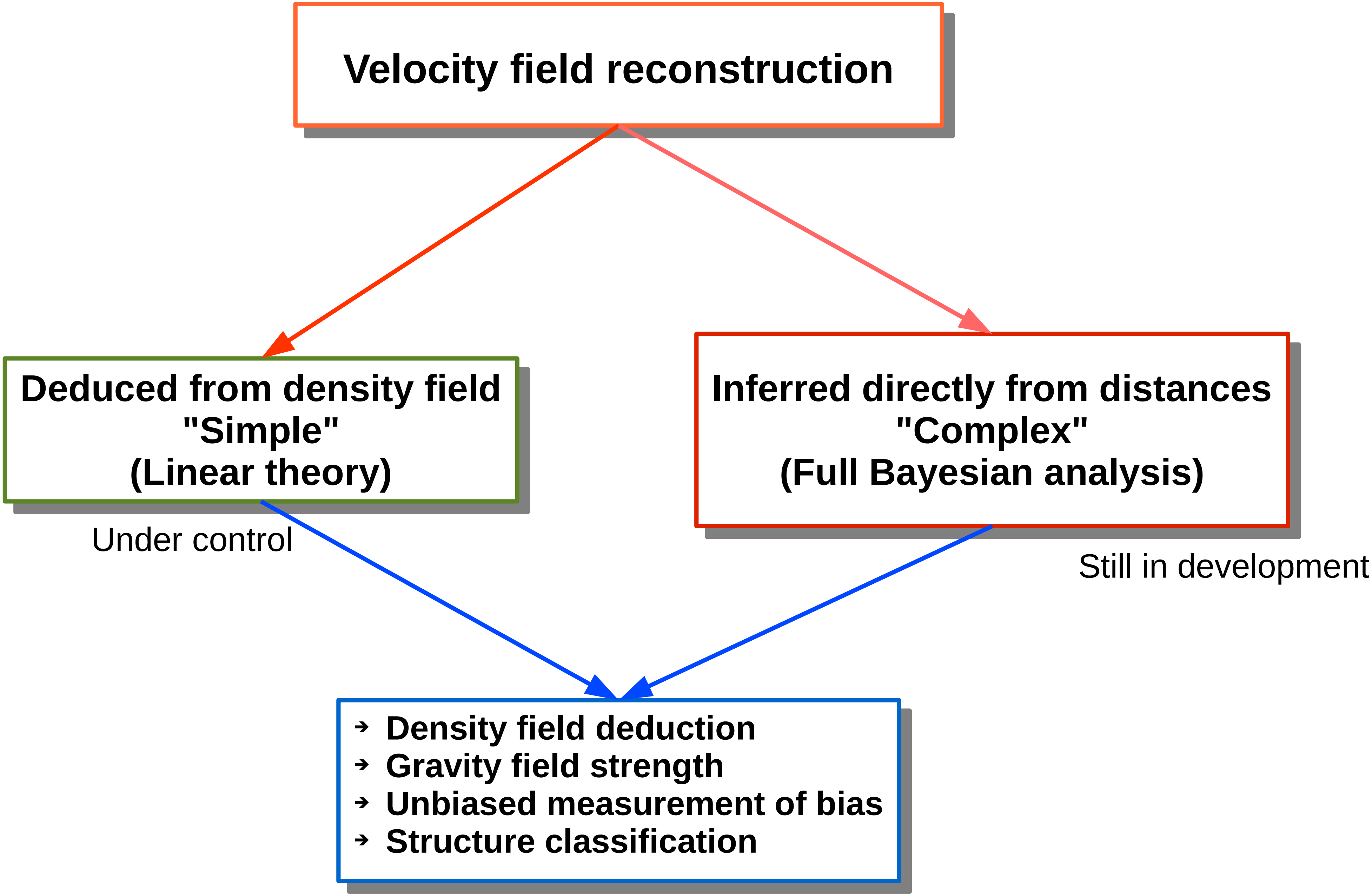
LSS conference, Garching/ESO

July 21st 2015

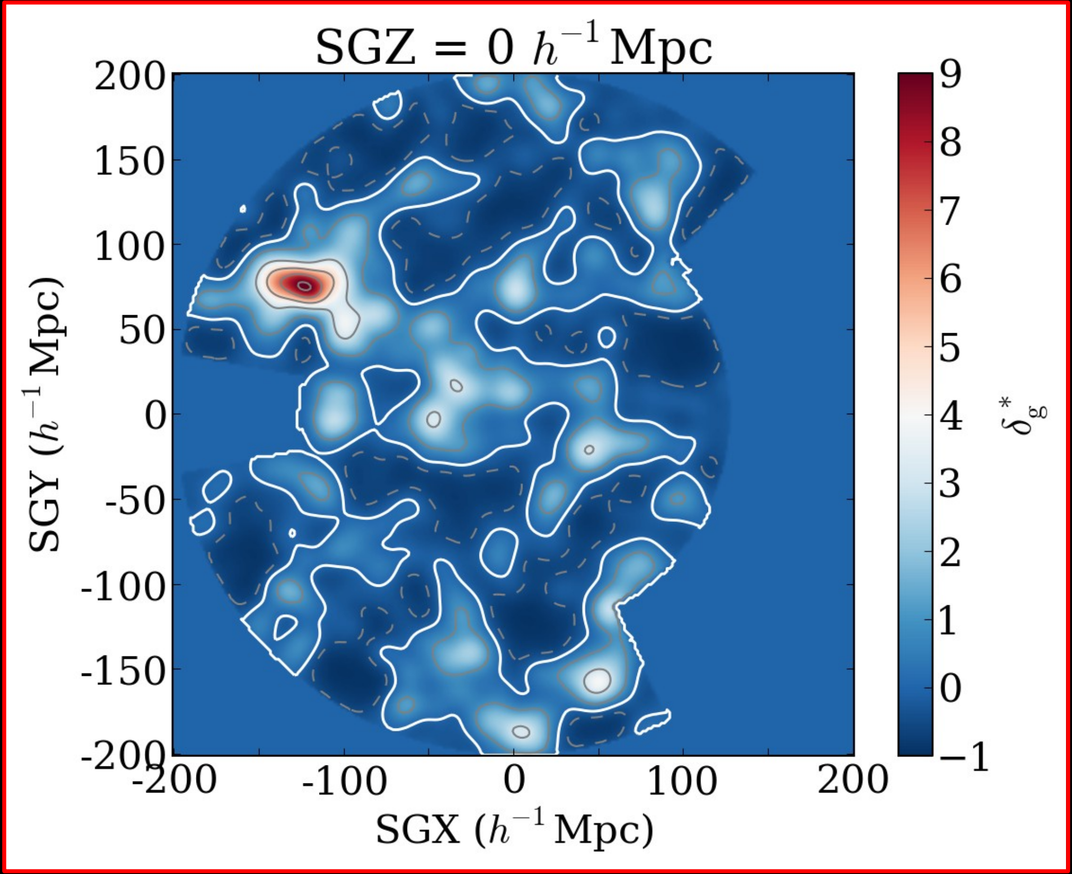
Galaxies in Large scale structures



Some new development in cosmic flows



Large scale statistical analysis (I)



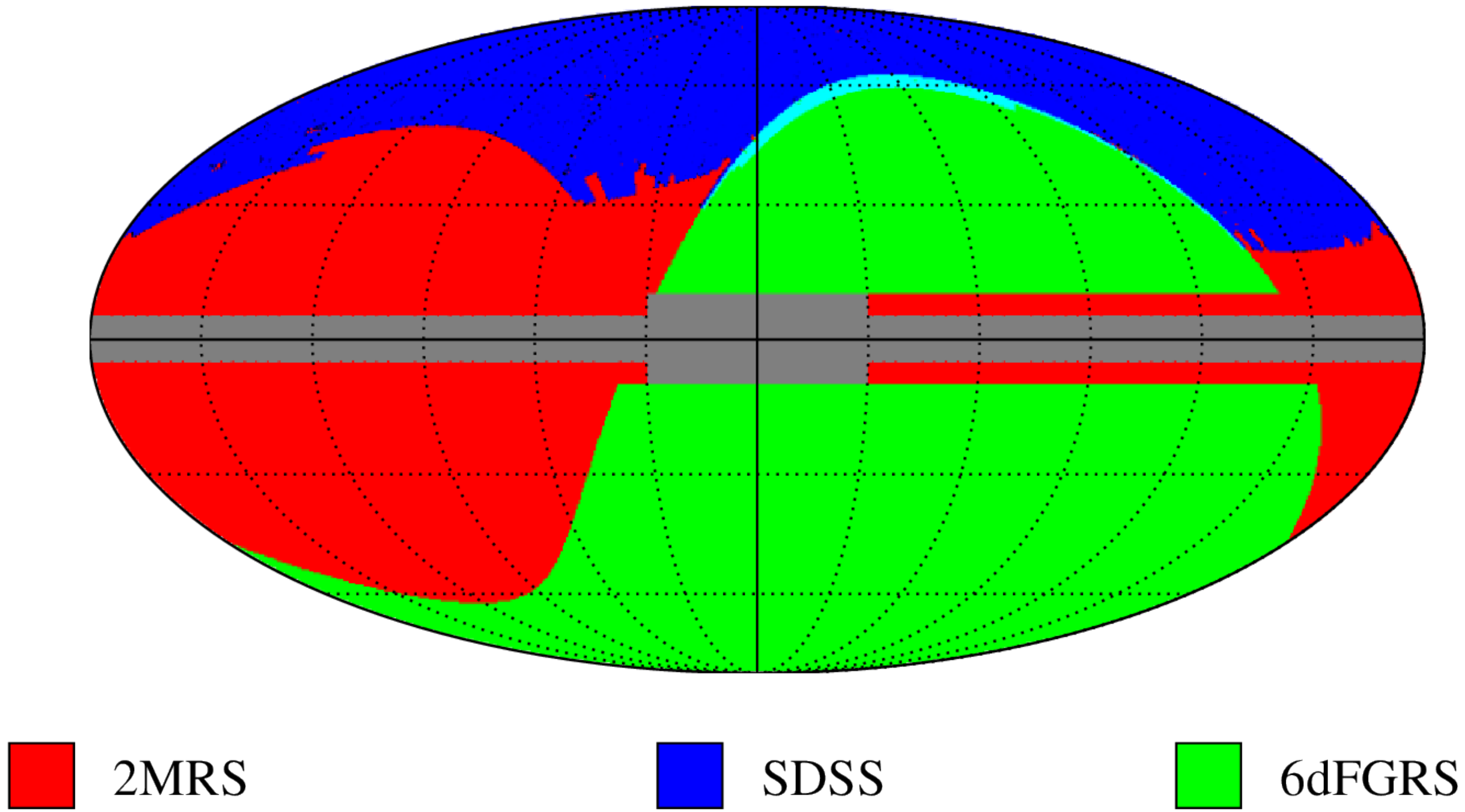
Lavaux & Hudson (2011, MNRAS)

Carrick, Turnbull, Lavaux & Hudson (2015, MNRAS)

http://cosmicflows.iap.fr/

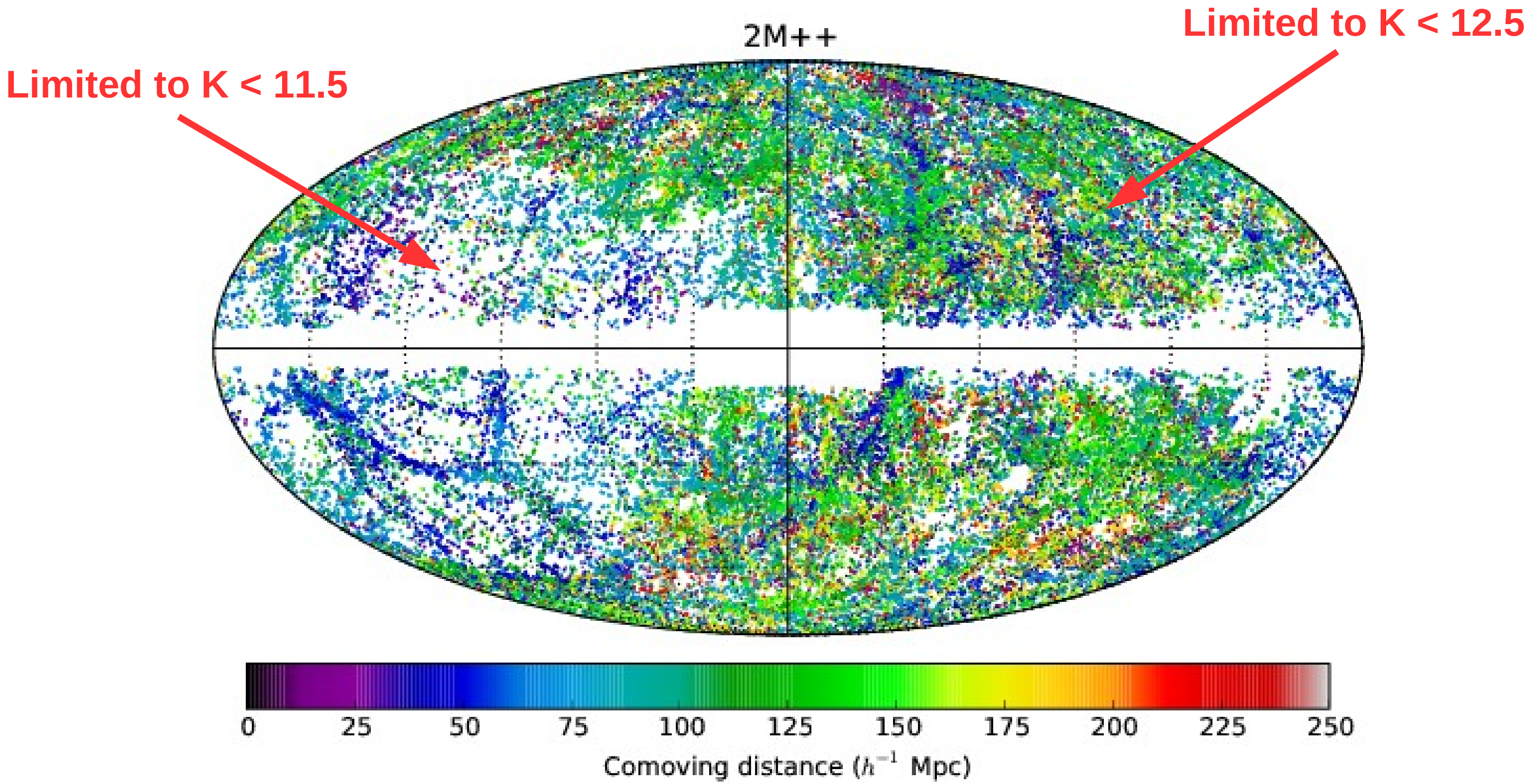
http://cosmicflows.uwaterloo.ca/

2M++ galaxy compilation: sources



Lavaux & Hudson (2011, MNRAS)

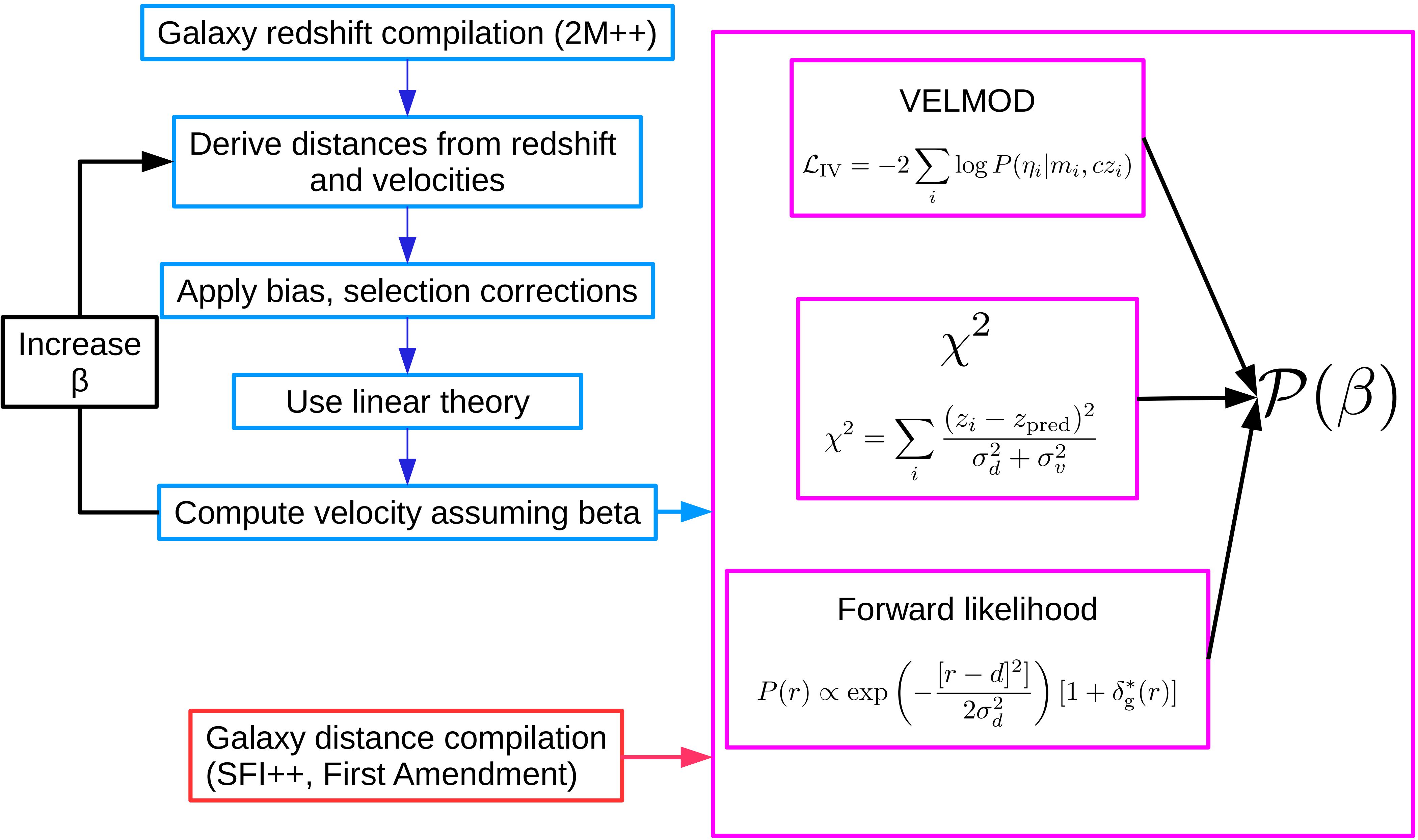
2M++ galaxy compilation: galaxies distribution



Lavaux & Hudson (2011, MNRAS)

Analysis framework

velocity field reconstruction from density



Carrick, Turnbull, Lavaux & Hudson (2015, MNRAS)

Analysis framework

velocity field reconstruction from density

Galaxy redshift compilation (2M++)

Derive distances from redshift and velocities

Apply bias, selection corrections

Use linear theory

Compute velocity assuming beta

Increase β

+ Free extra bulk-flow!

+ Some calibration on N-body

Galaxy distance compilation (SFI++, First Amendment)

VELMOD

$$\mathcal{L}_{\rm IV} = -2 \sum_i \log P(\eta_i | m_i, cz_i)$$

χ^2

$$\chi^2 = \sum_i \frac{(z_i - z_{\rm pred})^2}{\sigma_d^2 + \sigma_v^2}$$

Forward likelihood

$$P(r) \propto \exp\left(-\frac{[r-d]^2]}{2\sigma_d^2} \right) [1 + \delta_{\rm g}^*(r)]$$

$$\mathcal{P}(\beta)$$

Carrick, Turnbull, Lavaux & Hudson (2015, MNRAS)

Results

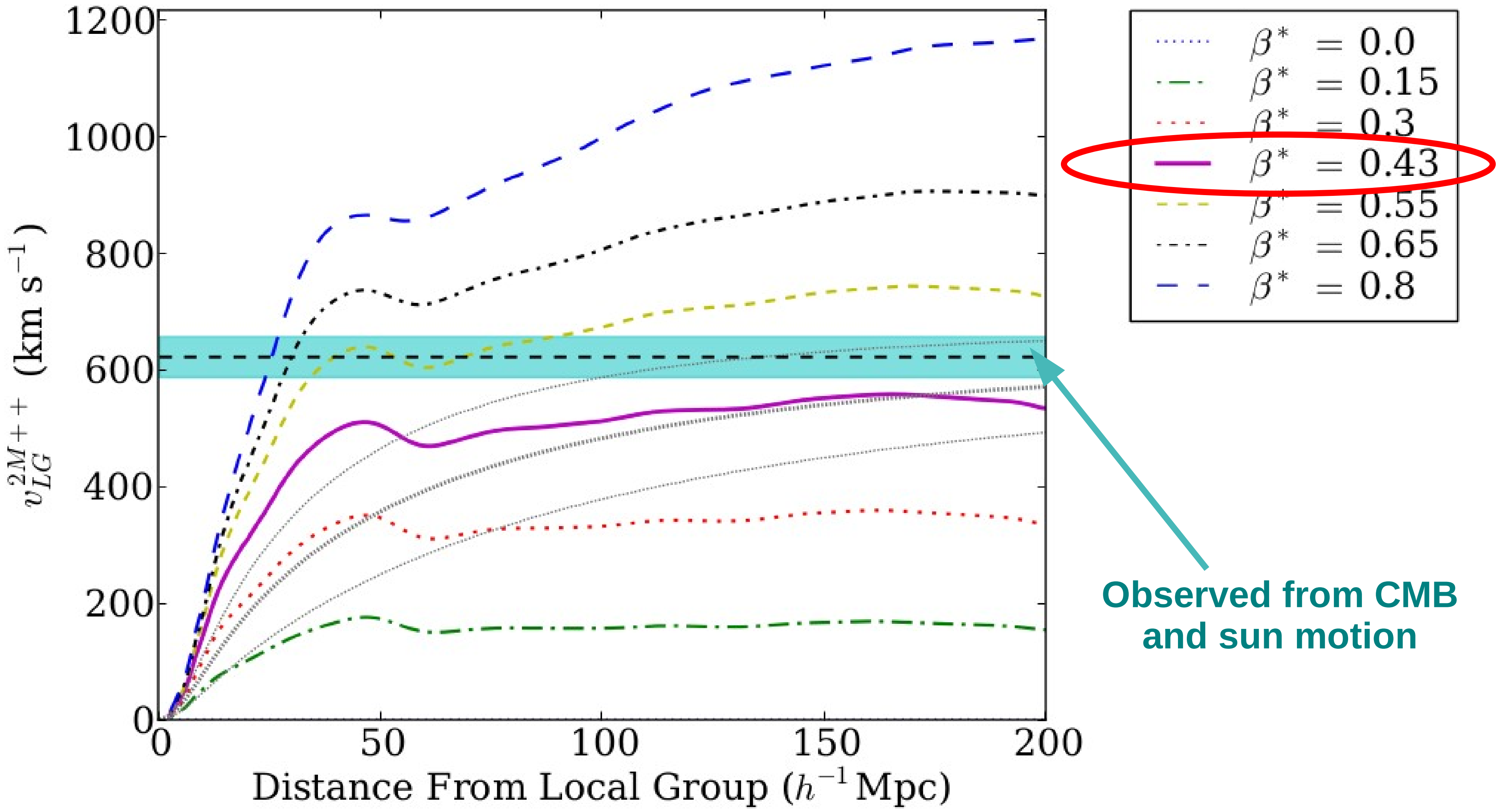
Growth of structure

	β^*	χ^2/(D.O.F.)
Forward Likelihood (LW)		
A1	0.440 ± 0.023	-
SFI++ Galaxy Groups	0.429 ± 0.022	-
SFI++ Field Galaxies	0.423 ± 0.045	-
All	$\mathbf{0.431 \pm 0.021}$	-
Forward Likelihood (NW)	0.439 ± 0.020	-
Inverse VELMOD (LW)	0.387 ± 0.048	-
χ^2 (LW)	0.444 ± 0.026	2194/2899
χ^2 (NW)	0.442 ± 0.028	2200/2899

Bulk flow

	v_x (km s^{-1})	v_y (km s^{-1})	v_z (km s^{-1})
BF_{2M++}	-3 ± 8	-72 ± 11	38 ± 11
LG_{2M++}	-18 ± 27	-422 ± 41	328 ± 37
$\mathbf{V}_{\mathrm{ext}}$	89 ± 21	-131 ± 23	17 ± 26
$\mathrm{BF}_{2M++} + V_{\mathrm{ext}}$	86 ± 22	-203 ± 26	55 ± 28
$\mathrm{LG}_{2M++} + V_{\mathrm{ext}}$	71 ± 34	-553 ± 47	345 ± 46

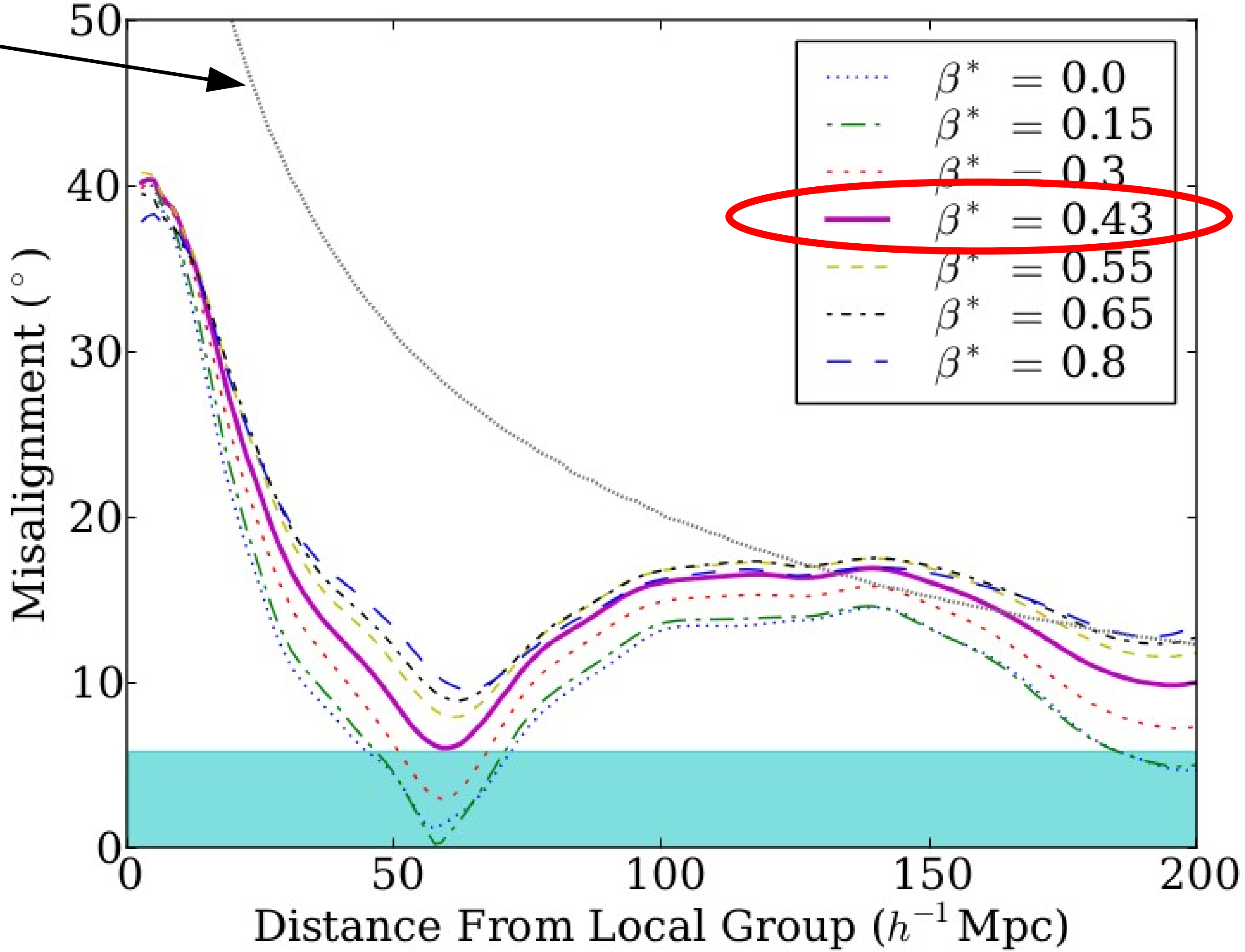
Results: convergence of the velocity of LG



AMPLITUDE

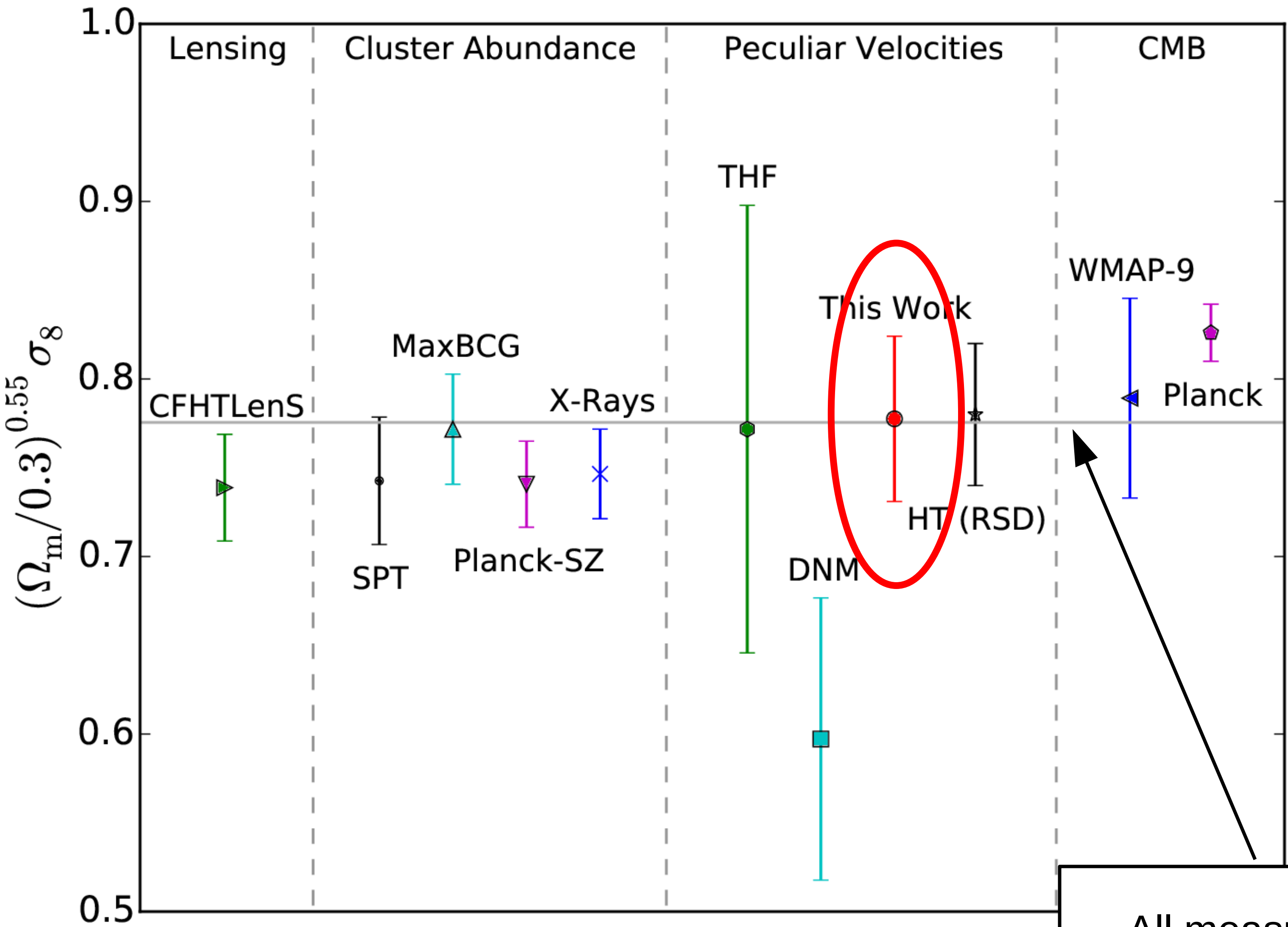
Results: convergence of the velocity of LG

95% limits from ΛCDM



ALIGNMENT

Results: gravity

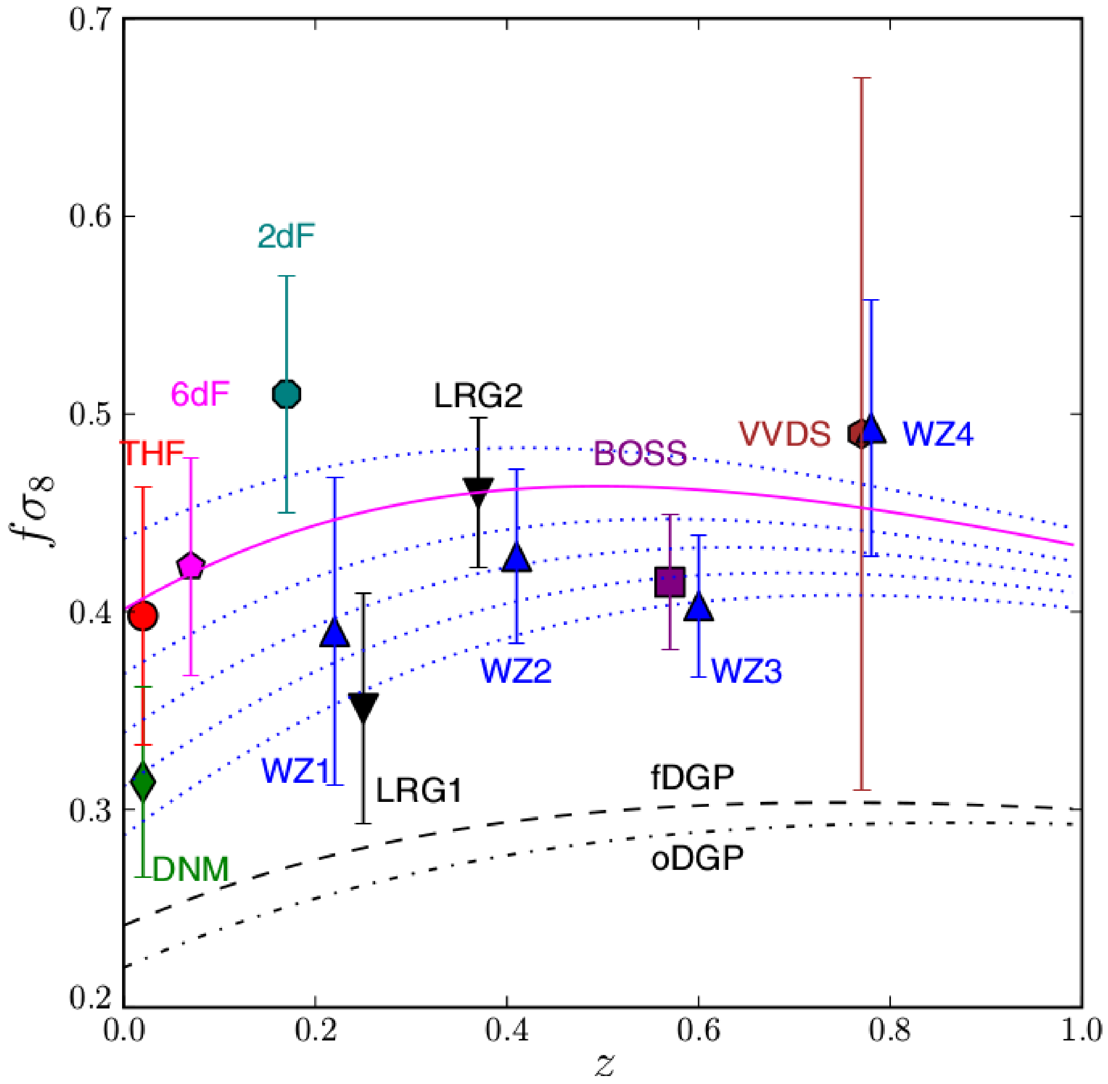


All measurements mean value

$f\sigma_8 = 0.400 \pm 0.005$

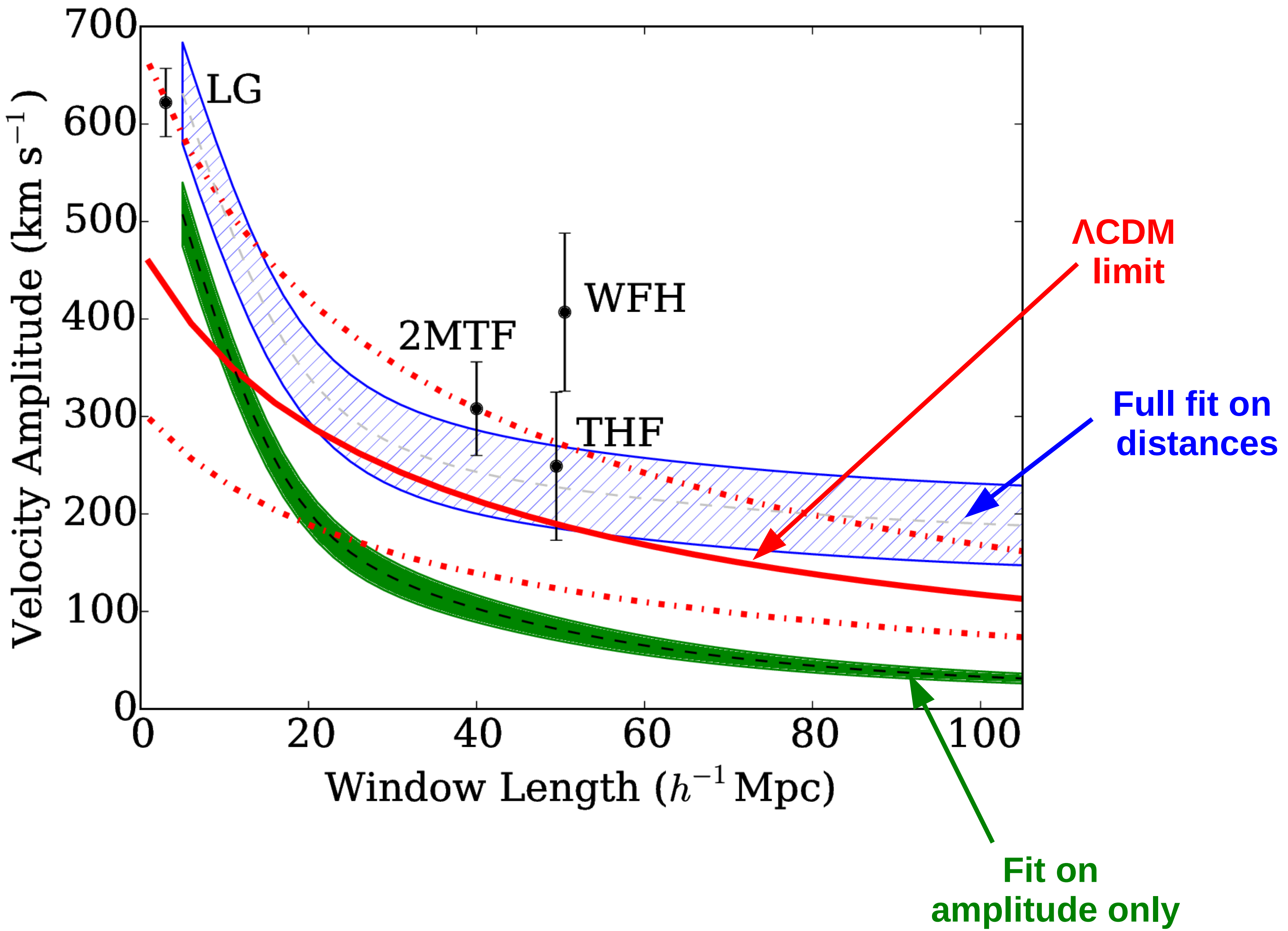
Carrick, Turnbull, Lavaux & Hudson (2015, MNRAS)

Insisting on low-z peculiar velocities



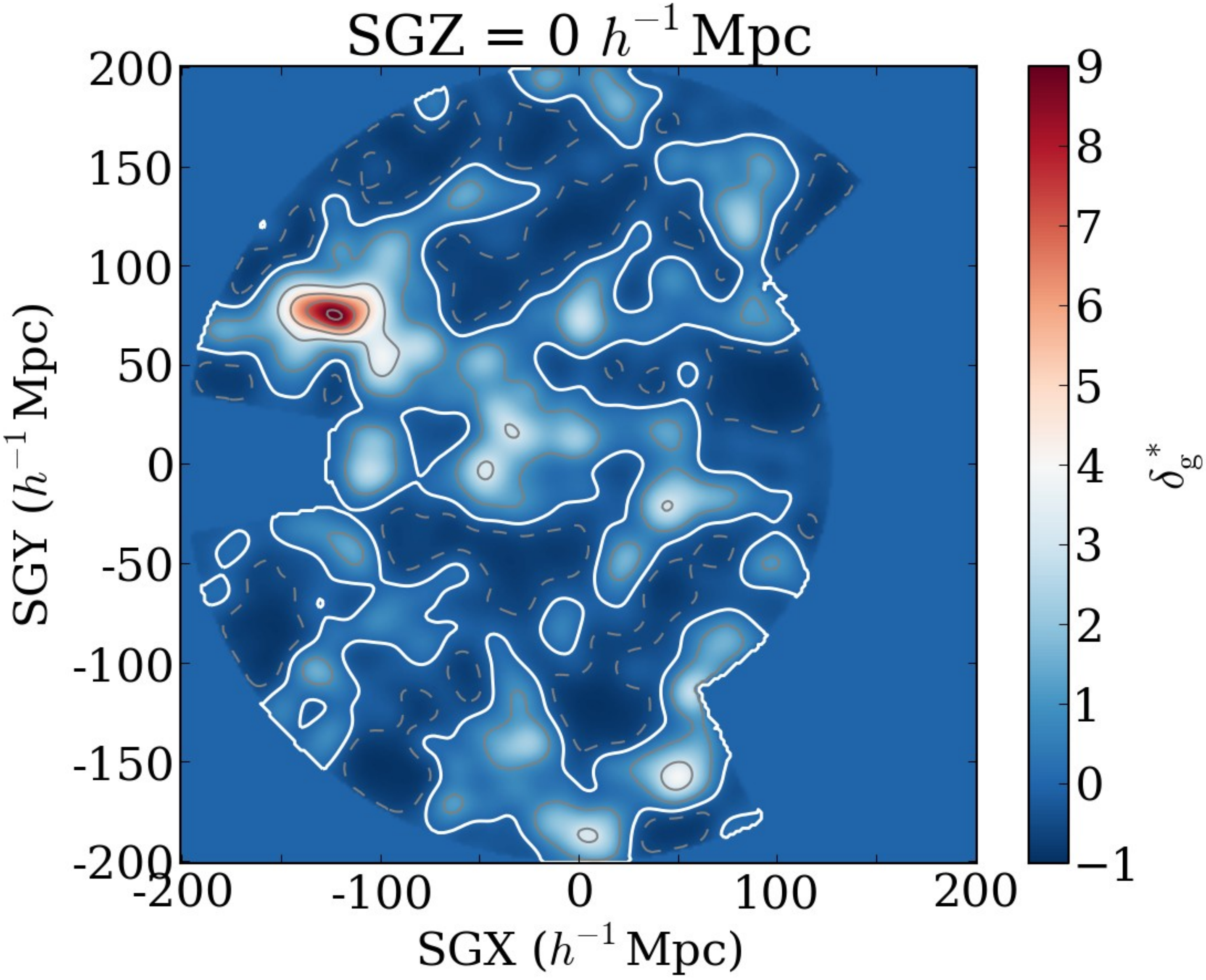
Hudson & Turnbull (2012)

Results: bulk flow



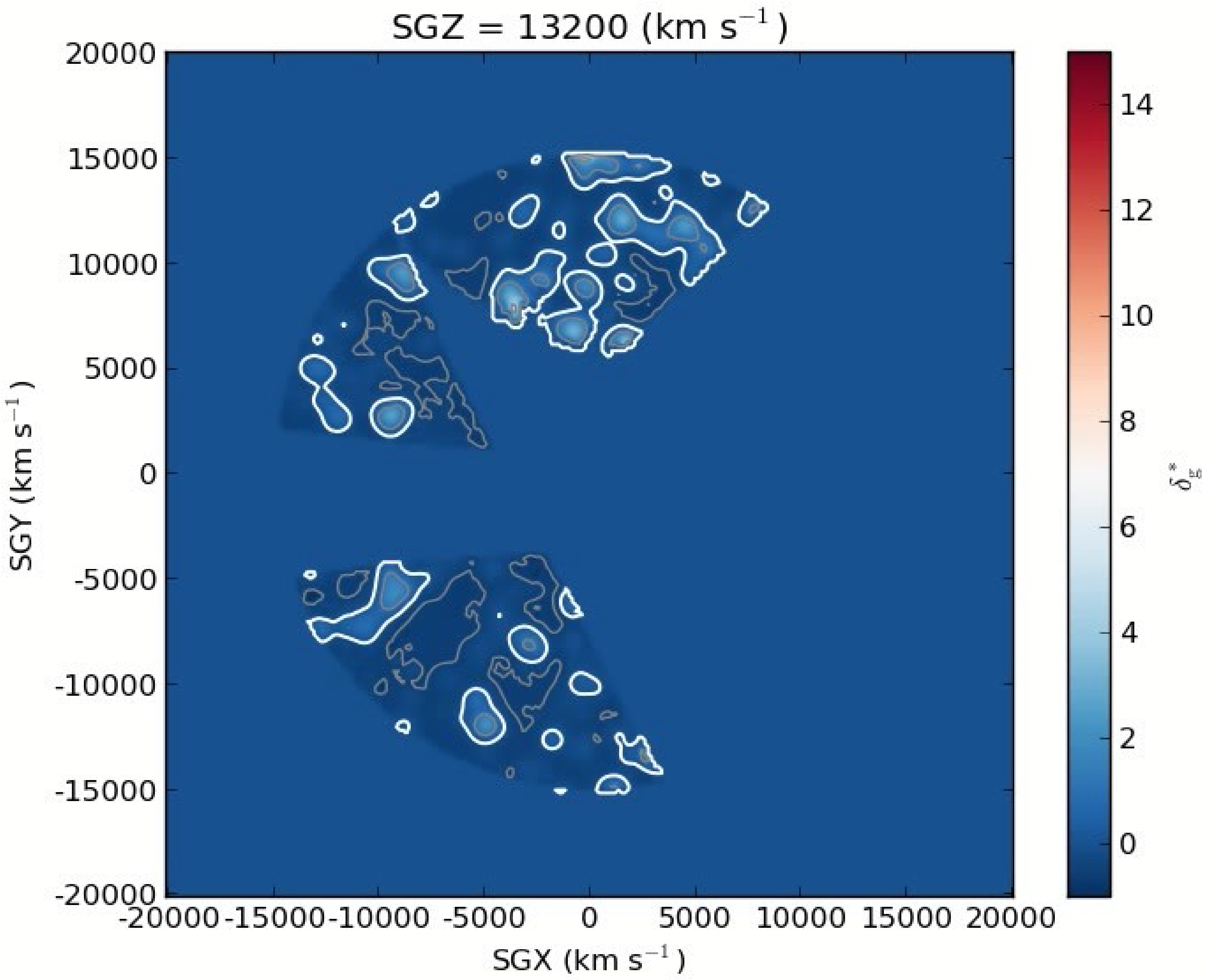
Carrick, Turnbull, Lavaux & Hudson (2015, MNRAS)

Large scale structure calibrated map



Carrick, Turnbull, Lavaux & Hudson (2015, MNRAS)

Large scale structure calibrated map



Carrick, Turnbull, Lavaux & Hudson (2015, MNRAS)

Summaries of results

- LG convergence as **expected** by LCDM
- Misalignment fluctuations **within** LCDM predictions
- $f\sigma_8$ in **agreement** with results from other probes
- **Bulk flow** still **high** but in **good agreement** with both observations and expectations

Calibrated velocities and maps at http://cosmicflows.iap.fr and http://cosmicflows.uwaterloo.ca .

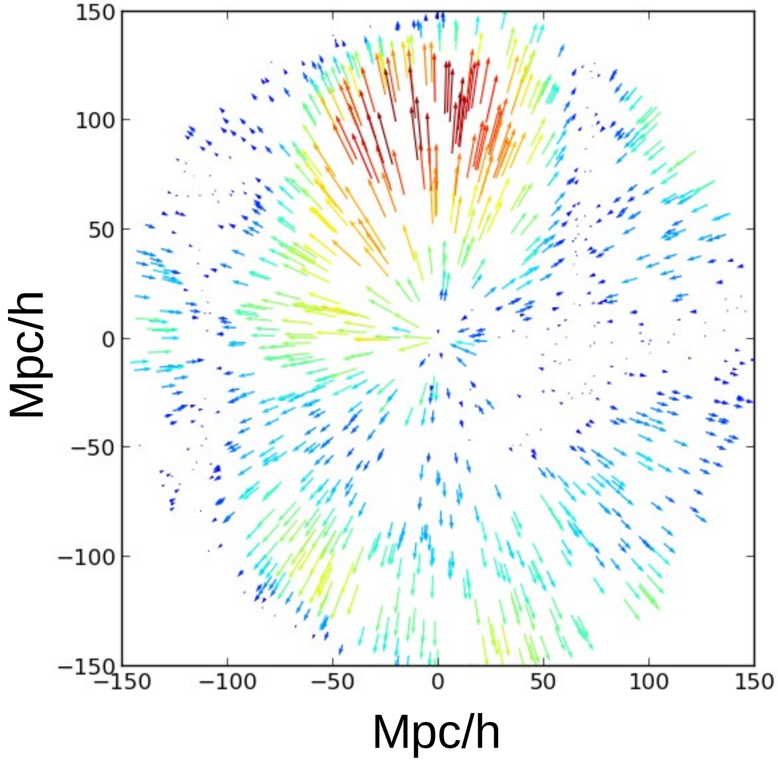
Large scale statistical analysis (II)

Red-figured volute-krater showing Hippolytus. Apulia c. 340 BC. Photo © Maicar Förlag - GML

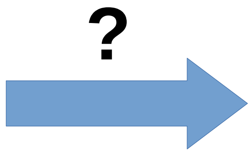
Velocity Reconstruction using Bayesian Inference Scheme **VIRBIuS**

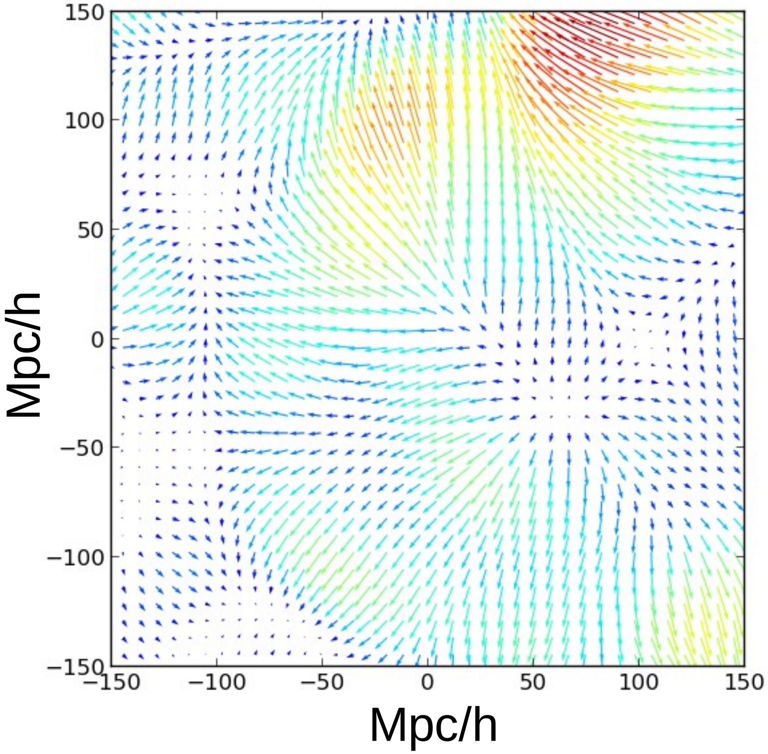
Lavaux (2015, submitted, in review)

The reconstruction problem: extrapolation



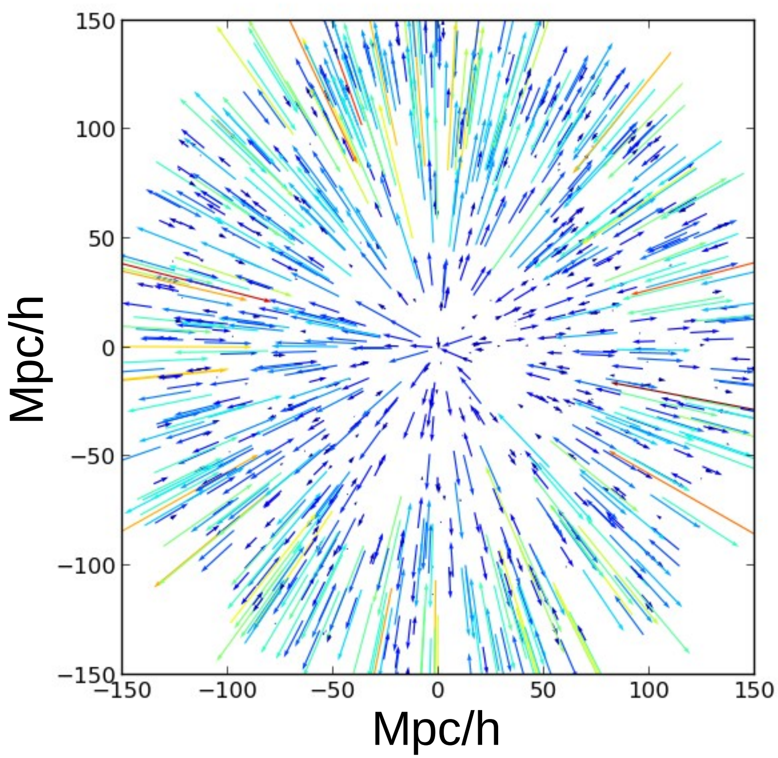
Ideal observations





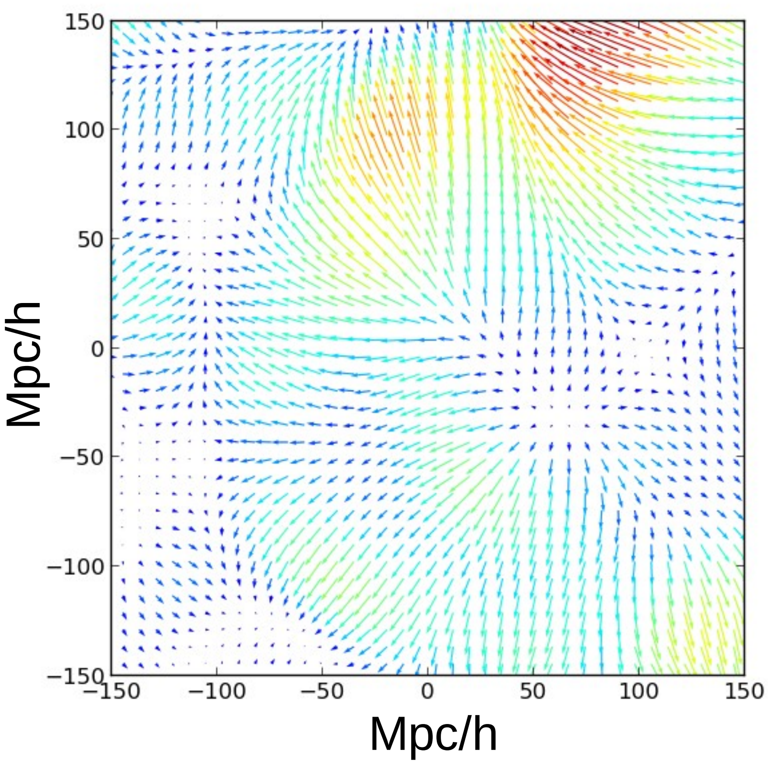
Full velocity field

The reconstruction problem: extrapolation



With full observational errors

?

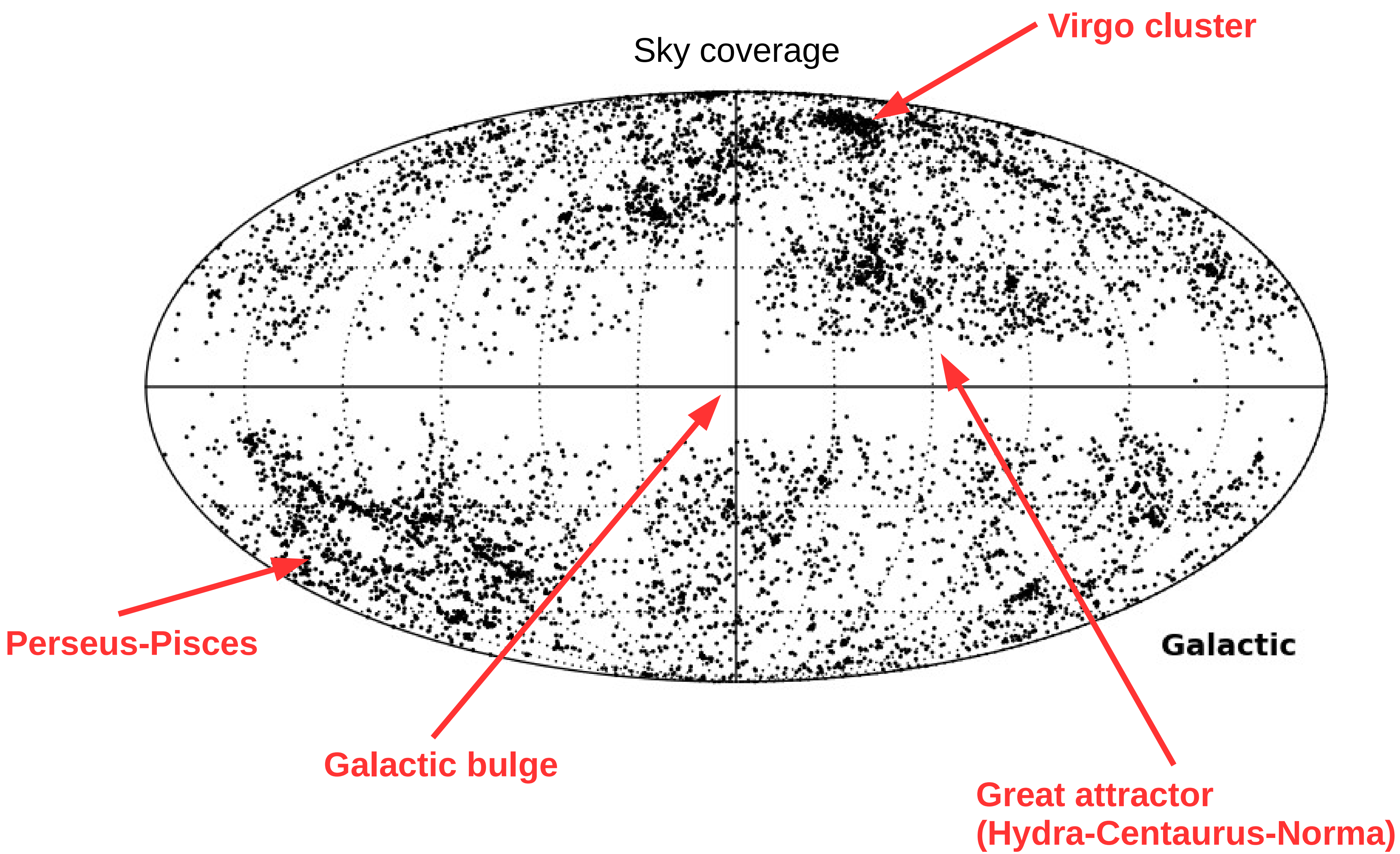


Full velocity field

Problems

- Inhomogeneous distance errors
- Unknown locations of tracers
- Hard to model non-linear dynamics
- Velocity/density bias

An example: the Cosmic Flows-2 catalog

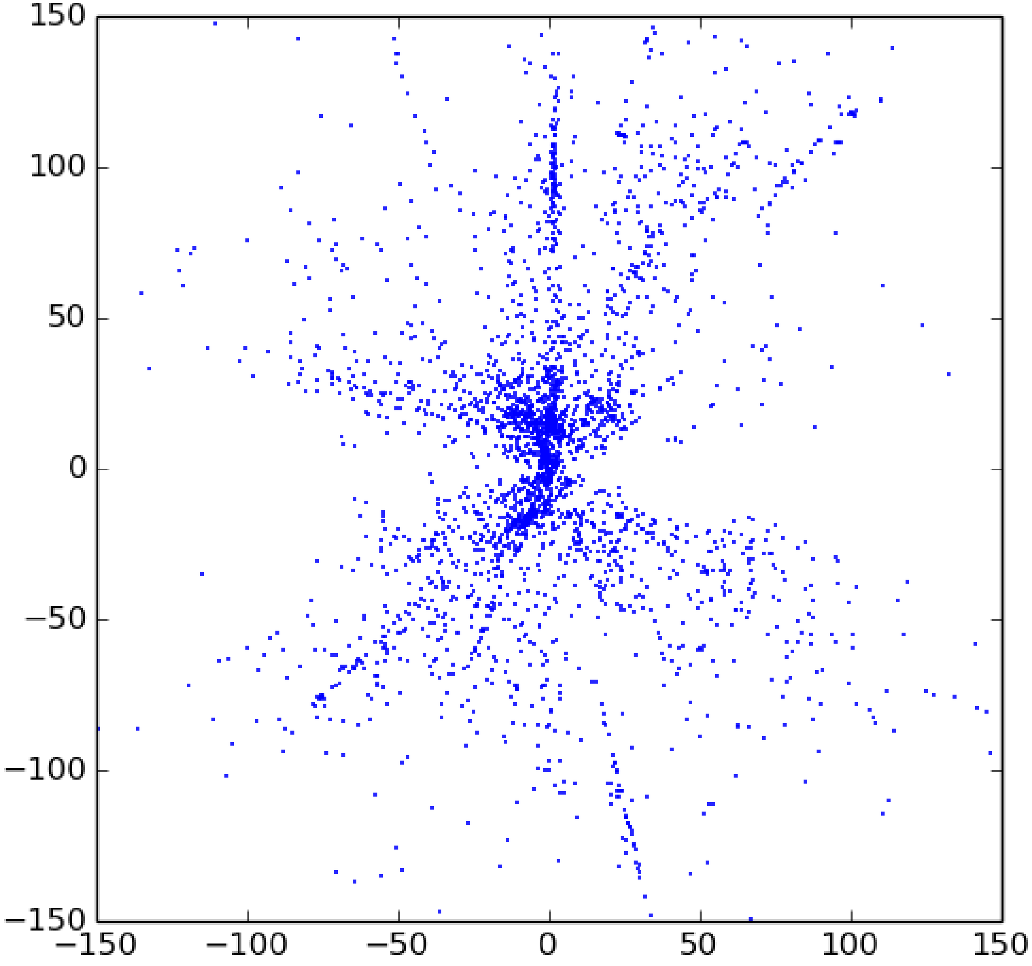


8139 galaxies (TRGB, TF, SNe, FP)

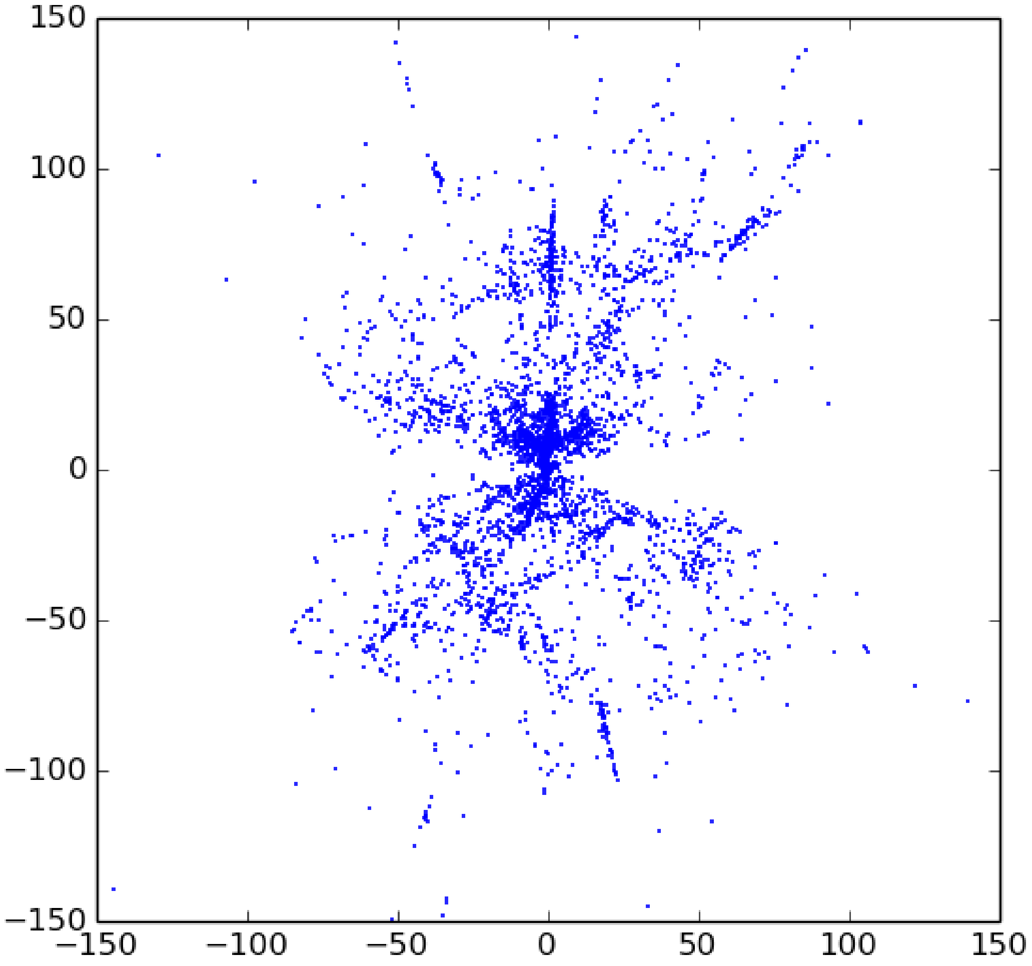
Tully et al (2013, AJ)

An example: the Cosmic Flows-2 catalog

Using measured distances



Using measured redshifts



The VIRBIuS model

Two observational constraints:

Redshift → $v_r = \widetilde{H} d + \mathbf{v}^W \left(\frac{\widetilde{H}}{H} d\mathbf{u} \right) .\mathbf{u} + \varepsilon$

Distance modulus → $\mu_{\mathrm{obs}} = 5 \log_{10} \left(\frac{\widetilde{H} d}{H \times 1\mathrm{pc}} \right) + \varepsilon_\mu$

Simplifying assumptions:

Curl-free velocity field $\theta(\mathbf{x}) = \nabla.\mathbf{v}(\mathbf{x})$

Velocity tracers are not biased

Isotropic radial selection effect for distances

Residual ε uncorrelated

Priors:

Velocity field θ Gaussian Random field $\langle \hat{\theta}(\mathbf{k}) \hat{\theta}(\mathbf{k}') \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') P(k)$

Distances $\pi(d) \propto d^p \exp \left[- \left(\frac{d}{d_{\mathrm{cut}}} \right)^n \right]$ Extra free parameters

Lavaux (submitted, 2015)

Numerical issues: The problem

Huge posterior: > 10^7 parameters

$$P\left(\{d_i\},\{\hat{\theta}_{i,j,k}\},\sigma_{\mathrm{NL}},A_S,H,\tilde{H},n,p,d_{\mathrm{cut}}\,\middle|\,\{z_i\},\{\mu_i\},\{\sigma_{z,i}\},\{\sigma_{\mu,i}\}\right)$$

Numerical issues: The problem

Huge posterior: > 10^7 parameters

$$P\left(\{d_i\},\{\hat{\theta}_{i,j,k}\},\sigma_{\mathrm{NL}},A_S,H,\tilde{H},n,p,d_{\mathrm{cut}}\,|\,\{z_i\},\{\mu_i\},\{\sigma_{z,i}\},\{\sigma_{\mu,i}\}\right)$$

+

Totally non-gaussian

Cannot evaluate a gridded posterior

Cannot run a classical Metropolis-Hasting Markov-Chain algorithm

BUT if we have a good proposal for a step of the chain it is doable

Likelihood and posterior

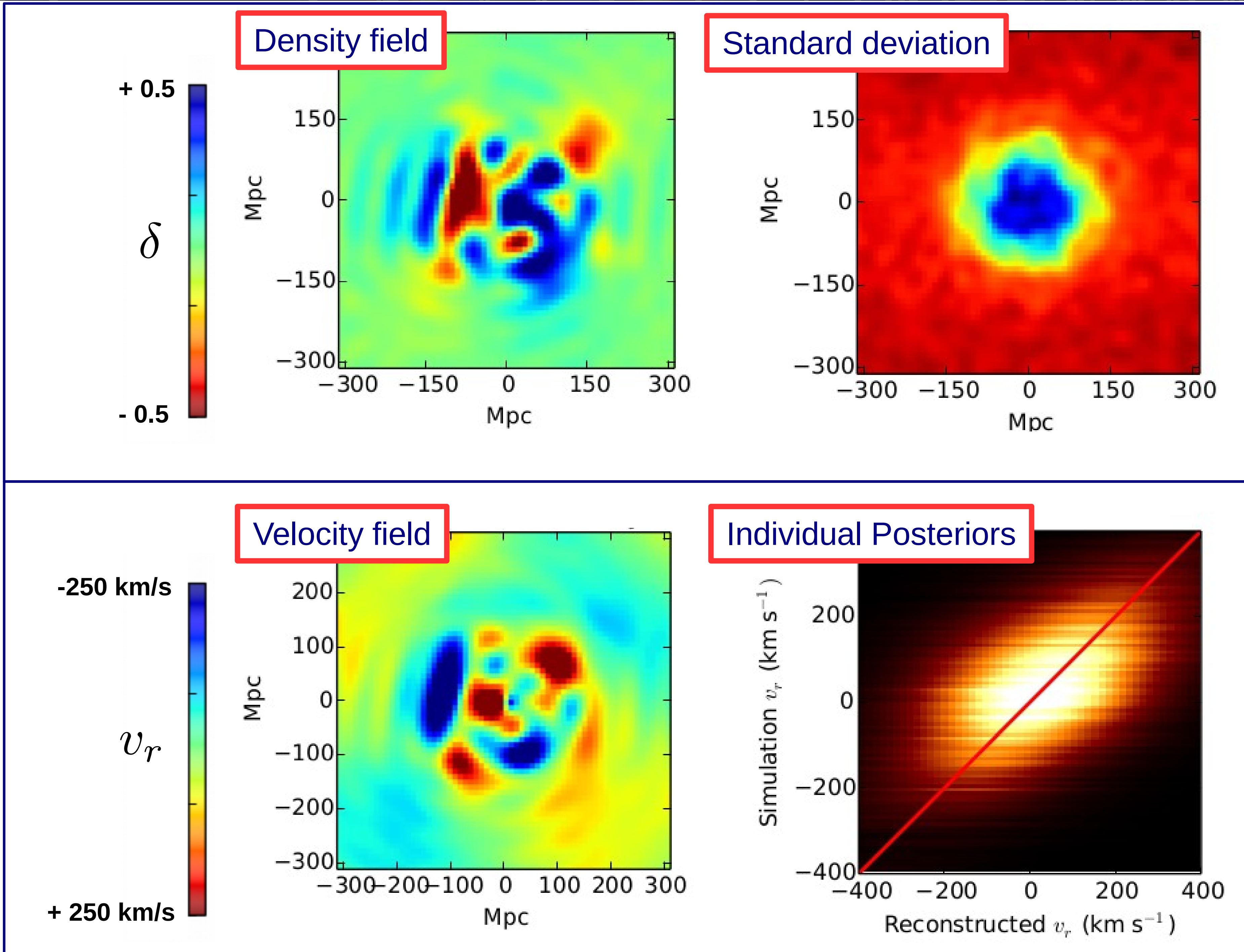
$$\mathcal{L} = P(\{\mu_i, z_i\}|\{d_i\}, \{\sigma_{z,i}, \sigma_{\mu,i}\}, \{\hat{\Theta}(k_q)\}, H, \widetilde{H}, \Sigma_{\rm NL}, \mathcal{T}, \{p_q^{type}\})$$

$$\propto \prod_{i=1}^{N_d} \left(\sigma_{z,i}^2(1+\bar{z}_i)^2 + \sigma_{{\rm NL},type(i)}^2\right)^{-1/2} \times$$

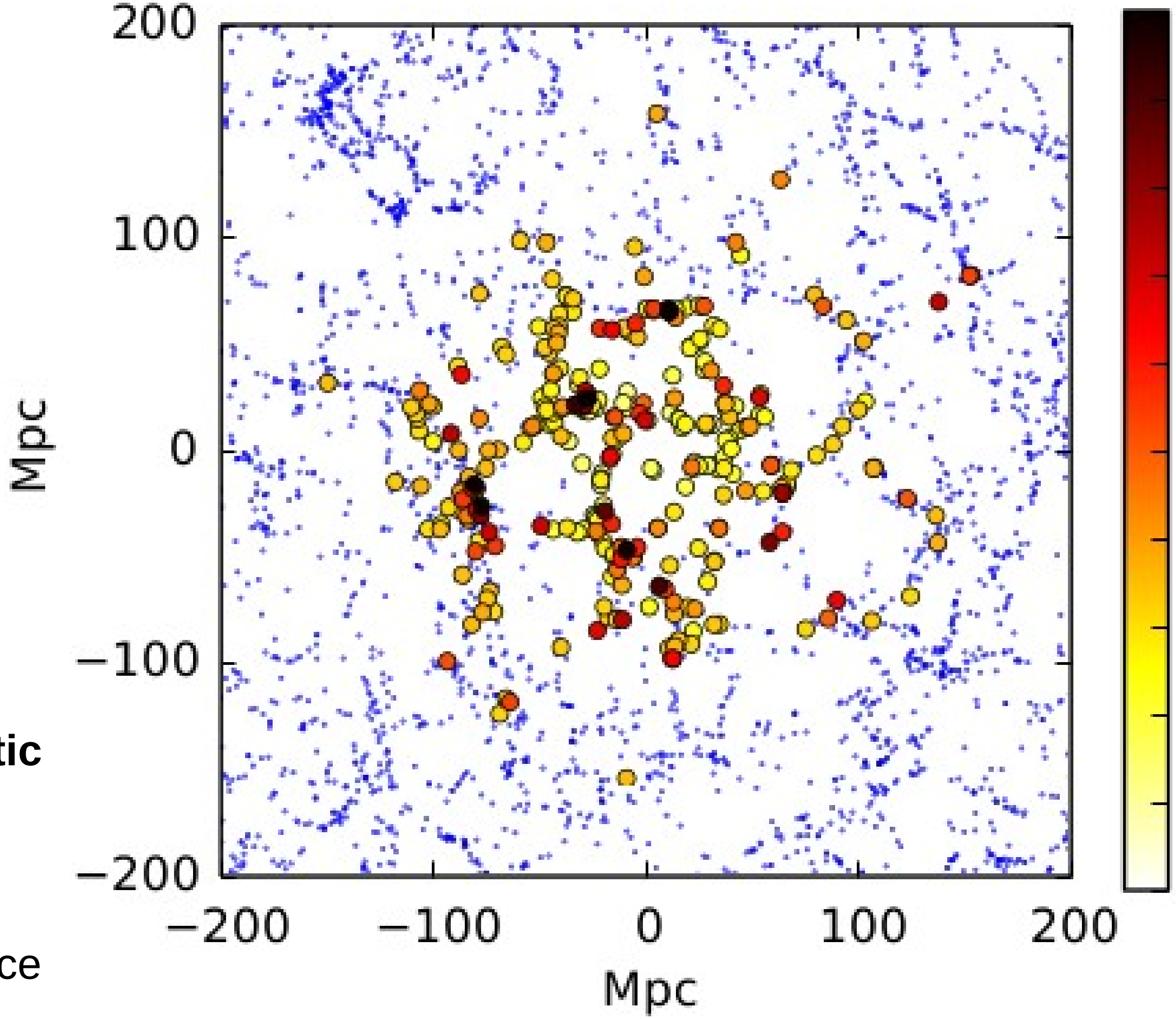
$$\exp\left\{-\frac{1}{2}\sum_{i=1}^{N_d} \frac{\left[v_i^r(z_i, d_i) - Hf\Psi_{r,i}(q^h)\right]^2}{\left[(\sigma_{z,i}^2(1+\bar{z}_i(d_i))^2 + \sigma_{{\rm NL},type(i)}^2\right]} - \frac{(\mu_i - 5\log_{10}(d_i^L/10\ {\rm pc}))^2}{\sigma_{\mu,i}^2}\right\}$$

$$P(\mathcal{D}^L = \{d_i\}, \hat{\Theta} = \{\hat{\Theta}(k_q)\}, \widetilde{H}, H, \Delta_{\mathcal{R}}^2, \{\sigma_{{\rm NL},q}\}, \mathcal{T}|$$
$$\mathcal{M} = \{\mu_i\}, \mathcal{Z} = \{z_i\}, \Sigma_z = \{\sigma_{z,i}\}, \Sigma_\mu = \{\sigma_{\mu,i}\}) =$$
$$\frac{\mathcal{L} \times \pi(\mathcal{D}^L)\pi(\hat{\Theta})\pi(\Sigma_{\rm NL})\pi(H)\pi(\{type(q)\})\pi(\Delta_{\mathcal{R}}^2)}{\int {\rm d}H{\rm d}\hat{\Theta}{\rm d}\mathcal{D}^L{\rm d}\Sigma_{\rm NL}\ \mathcal{L} \times \pi(\mathcal{D}^L)\pi(\{type(q)\})\pi(\hat{\Theta})\pi(\{\sigma_{{\rm NL},q}\})\pi(H)}.$$

Reconstructed « density » and velocity field



Application: halo mock catalog



Probably **high** velocity dispersion

Threshold is 100 km/s

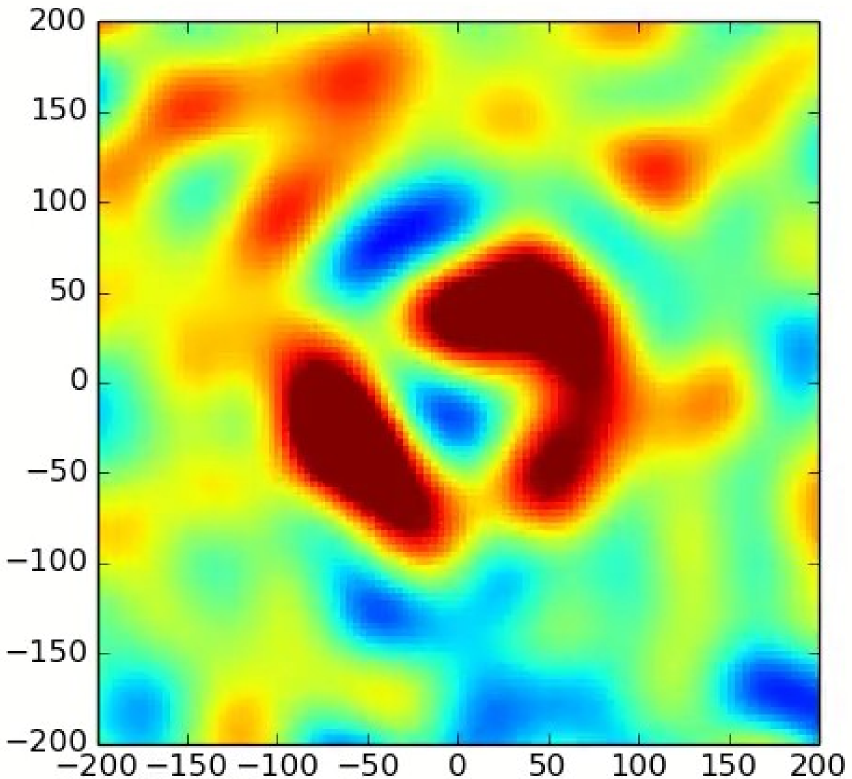
Probably **low** velocity dispersion

3000 tracers

Complex realistic radial selection

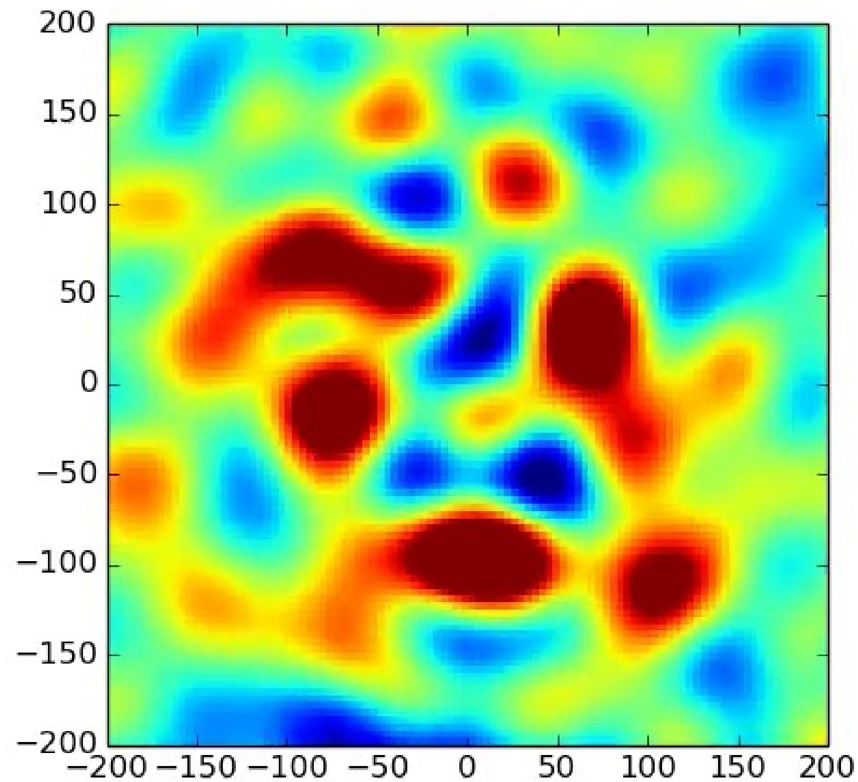
Typical observational **errors** on distance

Some early results on Cosmic Flows-2



Mpc

Galactic X plane



Mpc

Galactic Y plane

- 1  **+ 1**

Fourier scale cut at k=0.1 h/Mpc (~60 Mpc/h real scale)

Conclusion on VIRBIuS

- **Promising** results on mock catalogs
- Applicable to all distance data (SN, TF, FP, ...)
- **Cosmology is built-in** in the model, includes all potential known/unknown errors in tracers.
- Velocity bias model starts being important
- New tool, later available publicly with documentation

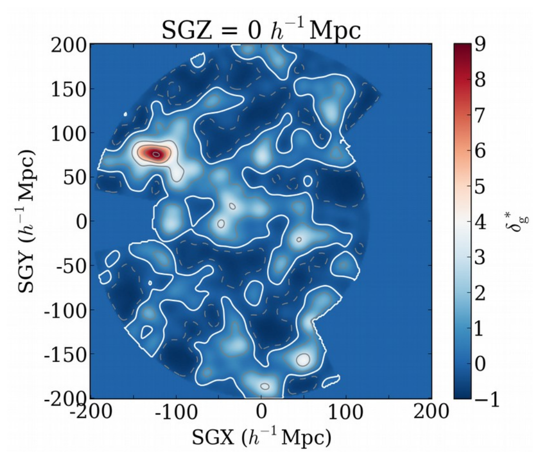
Conclusion

Conclusion

Making a better use of distance catalogs

Bayesian velocity field reconstruction

Iterative velocity field reconstruction



- cosmology
- galaxy environmental classification
- full 3d density field unbiased inference

- cosmology
- calibrated density field

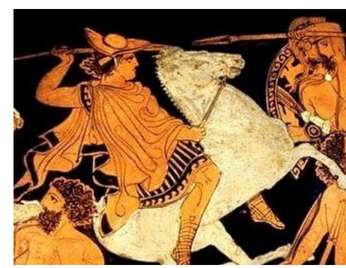
Conclusion

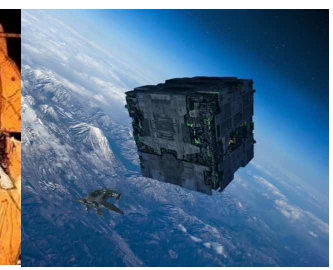
Going beyond the static analysis of LSS

1. Velocity field reconstruction

2. Chronocosmography (Jens Jasche' talk)

Challenging! But getting there





Understand the data collection
Proper statistical modeling
Develop analytical description of dynamics
Numerical integration, scaling to large computing farms
Posterior interpretation