Cosmic flows, Bayesian Inference & the Local Universe

Yehuda Hoffman
(Racah Inst. of Physics, Hebrew University)

XIIth Rencontres du Vietnam
Large Scale Structure and Galaxy Flows
July, 3rd - 9th, 2016, Quy Nhon, Vietnam

Collaborators (partial list)

- Potsdam: Jenny Sorce, Stefan Gottloeber, Noam Libeskind, Matthias Steinmetz
- Madrid: Chris Brook, Alexander Knebe, Gustavo Yepes
- Lyon: Helene Courtois, Romain Graziani
- Saclay: Daniel Pomarede
- Hawaii: Brent Tully
- Copenhagen: Arianna di Cintio
- Moscow: Sergey Pilipenko
- Tartu Obser.: Elmo Tempel
- Jerusalem: Edoardo Carlesi

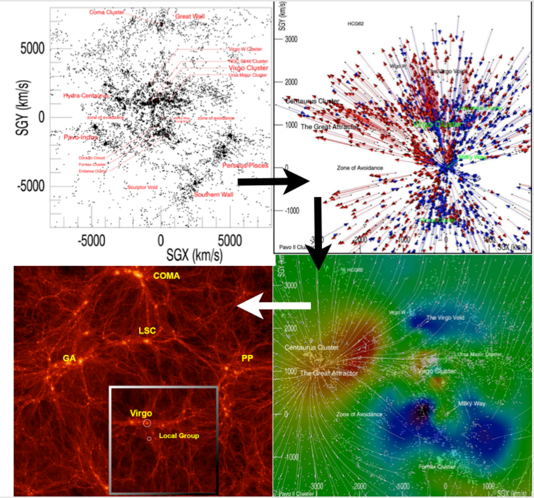
Outline

Cosmology with cosmic flows

- Cosmicflows
- Bayesian inference
- Near-field cosmology
- Constrained Local UniversE Simulations (CLUES)[a]

[a] http://www.clues-project.org

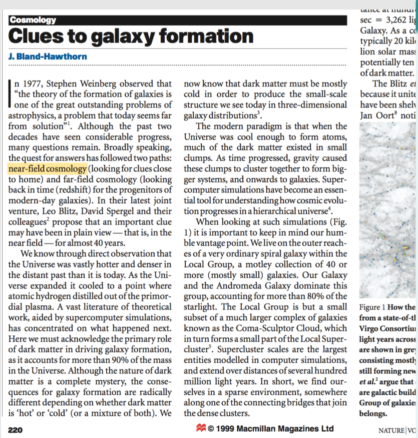
CLUES (in a nutshell)



The Three Pillars of CLUES

- Reconstruction of the Large Scale Structure (LSS) from noisy, sparse and incomplete data
- Time machine: from the (present epoch) reconstructed LSS to initial conditions (ICs)
- Constrained simulations: from ICs to the present epoch nearby universe

Near-field cosmology (Bland-Hawthorn, 1999)



Cosmology

Clues to galaxy formation

J. Bland-Hawthorn

In 1977, Stephen Weinberg observed that "the theory of the formation of galaxies is one of the great outstanding problems of astrophysics, a problem that today seems far from solution"[1]. Although the past two decades have seen considerable progress, many questions remain. Broadly speaking, the quest for answers has followed two paths: near-field cosmology (looking for clues close to home) and far-field cosmology (looking back in time (redshift) for the progenitors of modern-day galaxies). In their latest joint venture, Leo Blitz, David Spergel and their colleagues[2] propose that an important clue may have been in plain view — that is, in the near field — for almost 40 years.

We know through direct observation that the Universe was vastly hotter and denser in the distant past than it is today. As the Universe expanded it cooled to a point where atomic hydrogen distilled out of the primordial plasma. A vast literature of theoretical work, aided by supercomputer simulations, has concentrated on what happened next. Here we must acknowledge the primary role of dark matter in driving galaxy formation, as it accounts for more than 90% of the mass in the Universe. Although the nature of dark matter is a complete mystery, the consequences for galaxy formation are radically different depending on whether dark matter is 'hot' or 'cold' (or a mixture of both). We now know that dark matter must be mostly cold in order to produce the small-scale structure we see today in three-dimensional galaxy distributions[3].

The modern paradigm is that when the Universe was cool enough to form atoms, much of the dark matter existed in small clumps. As time progressed, gravity caused these clumps to cluster together to form bigger systems, and onwards to galaxies. Supercomputer simulations have become an essential tool for understanding how cosmic evolution progresses in a hierarchical universe[4].

When looking at such simulations (Fig. 1) it is important to keep in mind our humble vantage point. We live on the outer reaches of a very ordinary spiral galaxy within the Local Group, a motley collection of 40 or more (mostly small) galaxies. Our Galaxy and the Andromeda Galaxy dominate this group, accounting for more than 80% of the starlight. The Local Group is but a small subset of a much larger complex of galaxies known as the Coma-Sculptor Cloud, which in turn forms a small part of the Local Supercluster[5]. Supercluster scales are the largest entities modelled in computer simulations, and extend over distances of several hundred million light years. In short, we find ourselves in a sparse environment, somewhere along one of the connecting bridges that join the dense clusters.

220 © 1999 Macmillan Magazines Ltd NATURE | VO

Near Field Cosmology

- Our local neighbourhood is the part of the universe that we know the best - faint galaxies, satellites, tidal streams, …
- ΛCDM - is in an excellent agreement with the universe at large (scales), but tension exists on small scales: 'cusp vs. core', abundance of satellites
- The value of H_0: discrepancy between the near- and the far-field estimated values
- Nature of dark matter
- Galaxy formation

The notion of '**NEAR**-field' is stretched here to cover everything that is **LOCAL**.

Phases vs. Power

In the standard model of cosmology the universe emerges out of a Gaussian random perturbation field.

$$\delta(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3 k \delta_{\mathbf{k}} \exp\left(\imath\ \mathbf{k} \cdot \mathbf{r}\right)$$

$$\delta_{\mathbf{k}} = \left|\delta_{\mathbf{k}}\right| \exp\left(\imath\ \phi_{\mathbf{k}}\right)$$

Gaussian random perturbation field $\rightarrow$ RANDOM PHASE APPROXIMATION

Cosmologists are traditionally interested in the power (spectrum) of the perturbation field, $P(k)$.

Study of the near field involves the (reconstruction of the) phase structure.

Here we focus on the recovery of the power & phases (of the near-field).

The Copernican problem: near-field vs cosmology

How [a]typical is the Local Group?
How representative is the near field?

The physicist says: I want it to be typical so I can practice [near-field] cosmology on it.

The astronomer/astrophysicist says: I want it to be atypical. It is more interesting this way.

The LG and the near field can be typical with respect to some properties and unique with respect to others. And it is a matter of degree - not black or white.

Bayesian framework: WF/CRs

In the Bayesian approach one is interested in the posterior probability of a model given observational data:

$$\mathrm{P(model \mid \ data)} \propto \mathrm{P(data \mid \ model)\ \ P(model)}$$

- P(model) is the prior probability (knowledge) of the model
- P(data | model) is the likelihood of the data given the (prior) model
- P(model | data) is the posterior probability

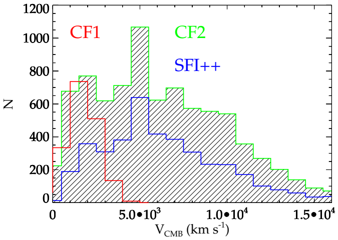
- Model: Gaussian random field, with the ΛCDM power spectrum
- Data: peculiar velocities of galaxies (Cosmicflows database)
- Reconstruction of the large scale structure by means of the Wiener filter (WF) and constrained realizations (CRs)

Data: Cosmicflows-2 (CF2)

The Astronomical Journal, 146:86 (25pp), 2013 October
doi:10.1088/0004-6256/146/4/86

COSMICFLOWS-2: THE DATA

R. Brent Tully[1], Hélène M. Courtois[1,2], Andrew E. Dolphin[3], J. Richard Fisher[4], Philippe Héraudeau[5], Bradley A. Jacobs[1], Igor D. Karachentsev[6], Dmitry Makarov[6], Lidia Makarova[6], Sofia Mitronova[6], Luca Rizzi[7], Edward J. Shaya[8], Jenny G. Sorce[2], and Po-Feng Wu[1]



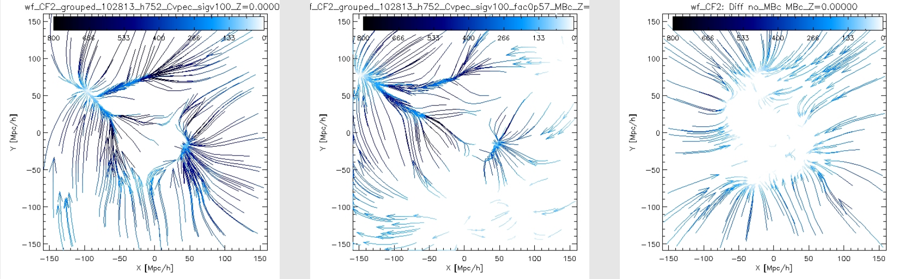
Number of data points: ≈8000 (4814 grouped)

Median redshift: $5895\,\mathrm{km\ s}^{-1}$

Malmquist bias correction (MBc)

With the (deeper) CF2 data the Malmquist bias cannot be neglected - it leads to a strong 'inhaling' (breathing) mode.

RAW MBc DIFF=MBc-RAW



Currently we have different MBc schemes: Sorce (2015), Graziani & Courtois (in preparation), YH. There are open issues with the methods and some conflicts. We are working on it!

Likelihood analysis of CF2 data w.r.t. ΛCDM

Data:

$$U_\mu = u_\mu + \epsilon_\mu = \mathbf{v}(\mathbf{r}_\mu) \cdot \hat{r}_\mu + \epsilon_\mu$$

Wiener filter:

$$v_\alpha^{\mathrm{WF}}(\mathbf{r}_i) = \xi_\mu^\alpha(\mathbf{r}_i)\xi_{\mu\nu}^{-1} U_\nu$$

Cross-correlation function:

$$\xi_\mu^\alpha(\mathbf{r}_i) = \left\langle v_\alpha(\mathbf{r}_i) U_\mu \right\rangle = \left\langle v_\alpha(\mathbf{r}_i)\mathbf{v}(\mathbf{r}_\mu) \cdot \hat{r}_\mu \right\rangle,$$

Auto-correlation function (covariance matrix):

$$\xi_{\mu\nu} = \left\langle U_\mu U_\nu \right\rangle = \left\langle u_\mu u_\nu \right\rangle + (\sigma_i^2 + \sigma_*^2)\delta_{\mu\nu}^K$$

Here σ_* is a free parameter that represents small scales non-linear velocities, $\sigma_* \approx (150 - 200)\,\mathrm{km\ s}^{-1}$.

Data likelihood:

Easy to calculate:

$$\mathcal{L}(\{U_\mu\}|\mathrm{model}) = \frac{1}{\sqrt{\det((\xi_{\mu\nu}))}} \exp\left[-\frac{U_\mu \xi^{-1}_{\mu\nu} U_\nu}{2}\right]$$

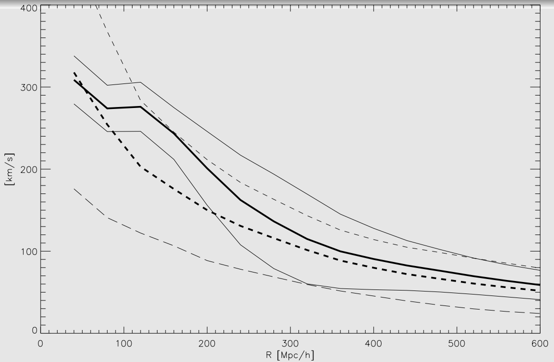
And

$$\chi^2/\mathrm{d.o.f.} = U_\mu \xi^{-1}_{\mu\nu} U_\nu / \ N_{\mathrm{d.o.f}}$$

With $N_{\mathrm{d.o.f}} = 4814$ the $\chi^2/\mathrm{d.o.f.}$ should be very close to 1.0, but it depends quite strongly on σ_*.

One needs to filter out the small scales so as to get a meaningful likelihood.

Vbulk of CF2 (Hoffman, Courtois & Tully, 2015)



Amplitude of the bulk velocity: The mean (thick lines) and the mean $\pm$ one standard deviation (thin lines) of an ensemble of 20 constrained (solid lines) and unconstrained (dashed lines) realizations. (In agreement with the CF2 analysis of Nusser & Davis, and the COMPOSITE database of Feldman, Hudson & Watkins.)

Likelihood analysis: bulk velocity (Hoffman, Nusser, Courtois & Tully 2016)

Bulk velocity of a sphere of radius R:

$$B_{\alpha}^{WF} = \frac{1}{V_R} \int_{r<R} \mathrm{d}^3\mathrm{r}\ v_{\alpha}^{WF}(\mathbf{r})$$

Bulk velocity likelihood function:

$$\mathcal{L}(B_{a,\alpha}^{WF}|\mathrm{model}) = \frac{1}{\sqrt{\det\left(\left\langle B_{a,\alpha}^{WF} B_{b,\beta}^{WF} \right\rangle\right)}}$$

$$\exp\left[-\frac{B_{a,\alpha}^{WF}\left\langle B_{a,\alpha}^{WF} B_{b,\beta}^{WF}\right\rangle^{-1} B_{b,\beta}^{WF}}{2}\right]$$

Bulk velocity auto-covariance matrix:

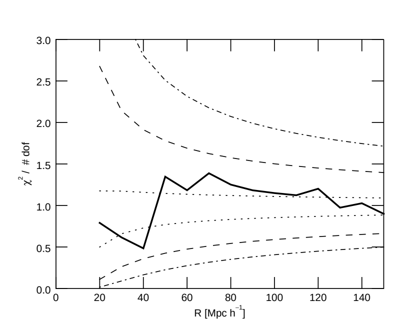
$$\left\langle B_{a,\alpha}^{WF} B_{b,\beta}^{WF} \right\rangle = \frac{1}{V_{R_a} V_{R_b}} \sum_{i\in R_a} \sum_{j\in R_b} \xi_{\mu}^{\alpha}(\mathbf{r}_i)\xi_{\mu\mu'}^{-1}\xi_{\mu'}^{\beta}(\mathbf{r}_j)$$

CF2 data and WMAP ΛCDM model: $\chi^2/\mathrm{d.o.f.}$

$$\chi^2/\mathrm{d.o.f.} = B^{WF}_{a,\alpha} \left\langle B^{WF}_{a,\alpha} B^{WF}_{b,\beta} \right\rangle^{-1} B^{WF}_{b,\beta} \; / \; N_{\mathrm{d.o.f}}$$

The $\chi^2/\mathrm{d.o.f.}$ of the (bulk velocity) of the CF2 data lies within the 2σ for $R = 20,\ 30,\ \ldots 150 h^{-1}\mathrm{Mpc}$.

The CF2 data is consistent with the (WMAP) ΛCDM model (within the framework discussed here).



Method can be extended to include monopole and quadruple moments and to do parameters estimation. It can gauge Malmquist bias corrections.

Cosmic web: V-web (Hoffman +, 2012)

Rescaling the shear tensor:

$$\Sigma_{\alpha\beta} = -\frac{1}{2H_0}\left(\frac{\partial v_\alpha}{\partial r_\beta} + \frac{\partial v_\beta}{\partial r_\alpha}\right)$$

Eigenvalues and eigenvectors: $\{\lambda_i\}$ & $\{\hat{\mathbf{e}}_i\}$ (i=1,2,3)

Constructing the V-web:

- Issue 1: The velocity field needs to be spanned on a (regular) grid. [Clouds in Cell (CIC) is used in simulations.]
- Issue 2: Spatial derivatives entail finite resolution. [Gaussian smoothing is used in simulations.]
- Issue 3: Assume a smoothing length (r_s) and a threshold value $\lambda_{\rm th}$.
- Web classification: number of eigenvalues above the threshold: knot (3), filament (2), sheet (1), void (0)
- Preferred directions: knots & voids - no, filaments - $\hat{\mathbf{e}}_3$, sheets - $\hat{\mathbf{e}}_1$

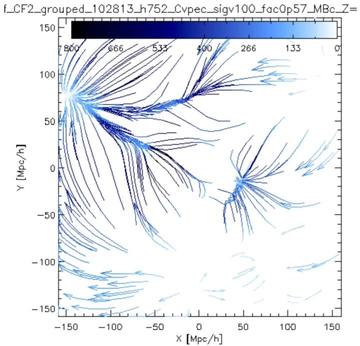
Peculiar velocities

Representing a velocity field by flow lines:

- Let s be the line parameter. The line element $d\vec{l}$ is defined by:

$$d\vec{l} = (\hat{x}v_x + \hat{y}v_y + \hat{z}v_z)ds$$

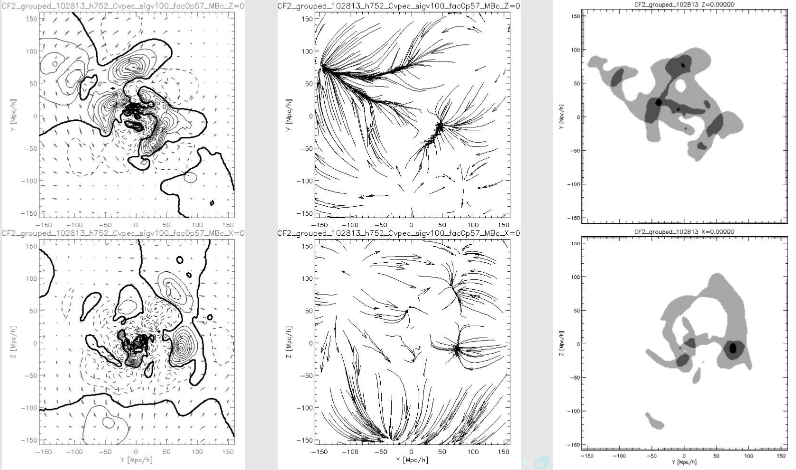
- The seed point of where the integration of a line starts needs to be determined (on a grid, at random, other options).
- Colour of the line represents the norm of the velocity vector.



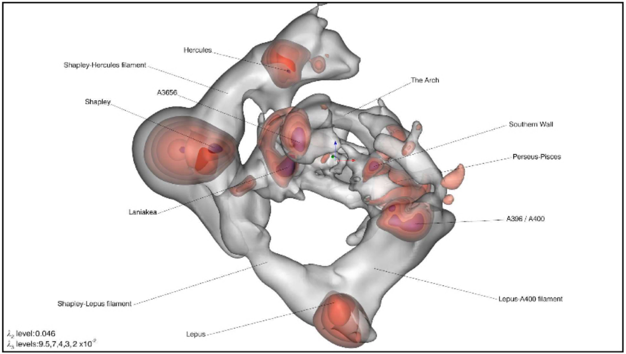
CF2: density, velocity & V-web:

$SGZ = 0$

$SGX = 0$



CF2 (V-web) and the large scale structure (LSS)

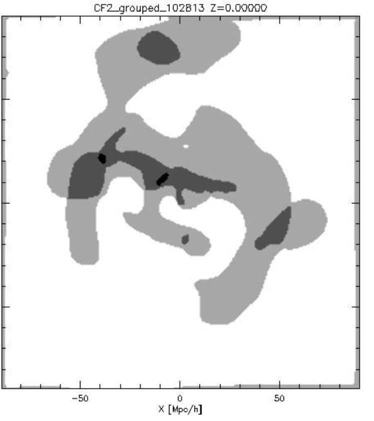


A 3D map of the Cosmic Web represented in terms of knots (red surfaces) and filaments (grey surface).

Orientation and dimensions are provided by the three-arrows signpost located at the origin of the supergalactic coordinate system with its 2000 km/s long arrows pointing to the three cardinal directions (red, green, blue for SGX, SGY, SGZ, respectively).

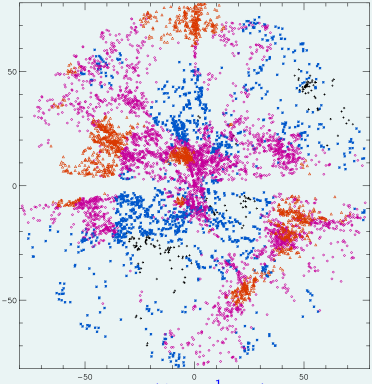
V-web and the 2MRS galaxies

CF2 V-web



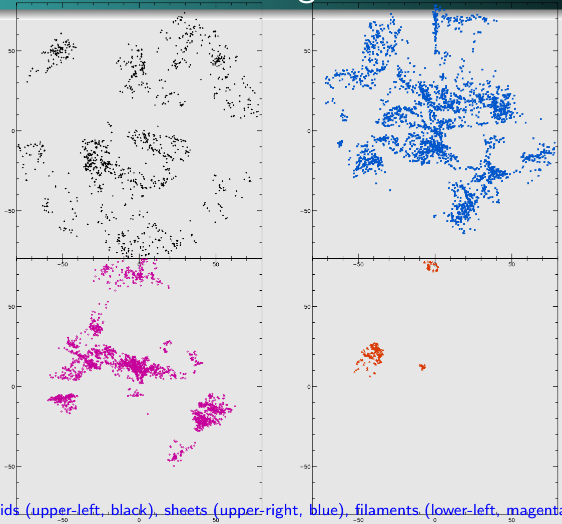
The Supergalactic Plane: the V-web derived from the CF2/WF reconstruction (threshold $\lambda_{th} = 0.03$).

2MRS galaxies & V-web



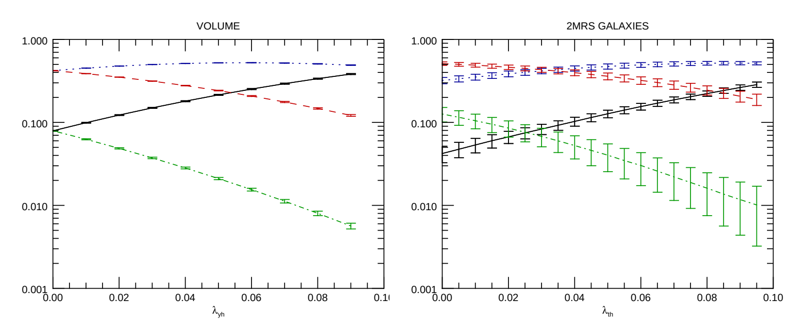
The Supergalactic Plane ($\pm 10 h^{-1}\mathrm{Mpc}$): 2MRS galaxies classified as voids (black), sheets (blue), filaments (magenta) and knots (oranges) (threshold $\lambda_{th} = 0.02$)

V-web classification of 2MRS galaxies



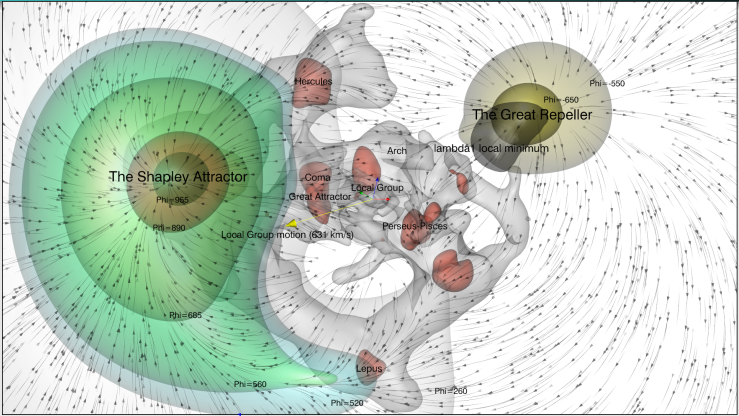
2MRS galaxies: voids (upper-left, black), sheets (upper-right, blue), filaments (lower-left, magenta), knots (lower-right, orange) (threshold $\lambda_{th} = 0.04$)

Volume and 2MRS galaxies web fraction

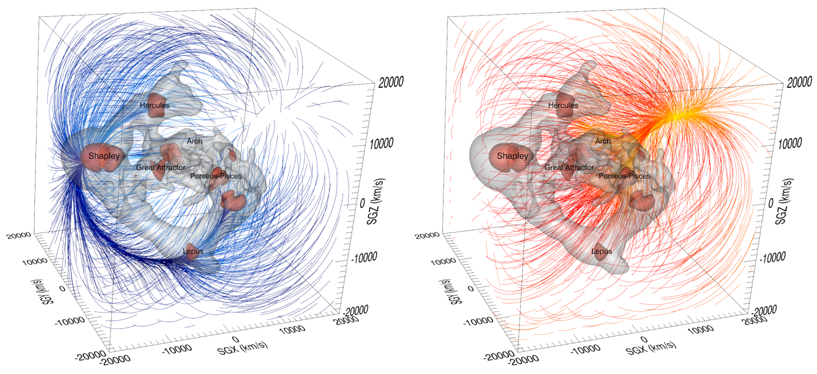


Volume (left) and number of (2MRS) galaxies (right) fractions in web elements voids (black), sheets (blue), filaments (red) and knots (green) as a function of threshold values (mean and scatter calculated over 20 CRs at $10h^{-1}\mathrm{Mpc}$ Gaussian smoothing).

Cosmography: Great Repeller and Shapley Attractor

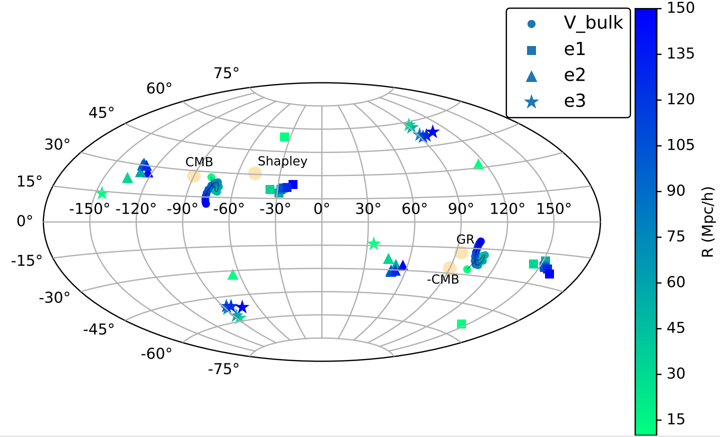


A face-on view of a slice $\pm 30h^{-1}\mathrm{Mpc}$ thick, normal to the direction of the pointing vector $\hat{r} = (0.604, 0.720, -0.342)$. Three different elements of the flow are presented: stream lines, red and grey surfaces present the knots and filaments of the V-web, and equipotential surfaces are shown in green and yellow. The yellow arrow indicates the direction of the CMB dipole.

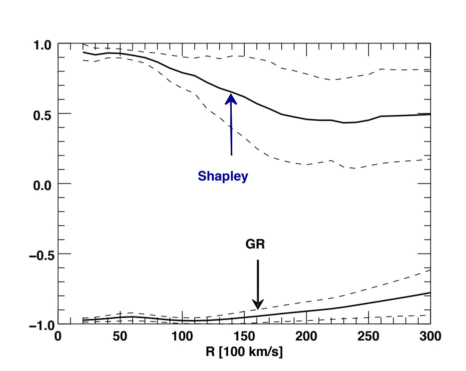


A 3D view of the stream lines of the flow field (left panel) and of the anti-flow (right panel). Stream lines are seeded on a regular grid and are coloured according to the magnitude of the velocity, The knots and filaments of the V-web are shown for reference.

What is the source of our motion w.r.t. the CMB?



Aitoff projection the Great Repeller, the Shapley Attractor, the CMB dipole, the bulk velocity and the three eigenvectors of the velocity shear tensor(evaluated across spheres of radius R.



Cosine of angles between a. Shapley Attractor and the 'expanding' eigenvector of the shear tensor; b. Great Repeller and the bulk velocity - of spheres of radius R (mean and scatter).

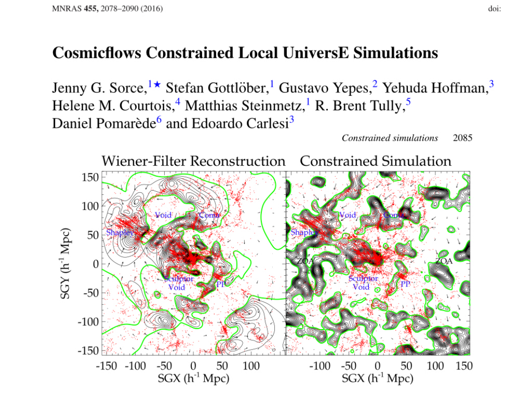
The Great Repeller (GR) and the Shapley Attractor

The Great Repeller is responsible for the CMB dipole - it 'pushes' the LG. The Shapley Supercluster acts as an attractor that drives the tidal component of the local flow field.

The GR is located at:

$$[SGX, SGY, SGZ] \approx [110, -60, 100] h^{-1}\mathrm{Mpc}$$

CF2 constrained simulations of the local Universe


MNRAS 455, 2078–2090 (2016) doi:

Cosmicflows Constrained Local UniversE Simulations

Jenny G. Sorce,[1★] Stefan Gottlöber,[1] Gustavo Yepes,[2] Yehuda Hoffman,[3] Helene M. Courtois,[4] Matthias Steinmetz,[1] R. Brent Tully,[5] Daniel Pomarède[6] and Edoardo Carlesi[3]

Constrained simulations 2085

- BOX $500h^{-1}\mathrm{Mpc}$
- N=512^3
- 15 constrained simulation and 10 random ones (for control)
- **The local universe (out to a few tens of Mpc) is robustly constrained.**
- Virgo: $M_{200} \sim (2.7-4.3) \times 10^{14} h^{-1}\mathrm{M}_{\odot}$, within $(3-4)h^{-1}\mathrm{Mpc}$ from actual position.

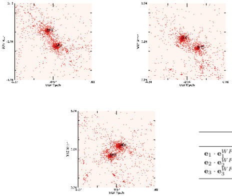
The Local Group factory (Carlesi et al 2016)

Constrained Local UniversE Simulations: A Local Group Factory

Edoardo Carlesi,[1] * Jenny G. Sorce,[2] Yehuda Hoffman,[1] Stefan Gottlöber,[2] Gustavo Yepes,[3,4]
Noam I. Libeskind,[2] Sergey V. Pilipenko,[5,6] Alexander Knebe,[3,4] Hélène Courtois,[7]
R. Brent Tully,[8] Matthias Steinmetz[2]

SIMULATIONS

Name	N_{simu}	L_{box}	R_{zoom}	m_p	z_{start}
SimuLN256	100	100	NO	5.26×10^9	60
SimuLN512	12	100	NO	6.57×10^8	80
SimuLGzoom	300	100	12	6.57×10^8	80



RECOVERY OF THE LOCAL NEIGHBORHOOD

	mean	σ
$\mathbf{e}_1 \cdot \mathbf{e}_1^{WF}$	0.95	0.11
$\mathbf{e}_2 \cdot \mathbf{e}_2^{WF}$	0.93	0.15
$\mathbf{e}_3 \cdot \mathbf{e}_3^{WF}$	0.97	0.08

(a)

	simu	*obs*
λ_1	0.174 ± 0.062	0.148 ± 0.038
λ_2	0.052 ± 0.075	0.051 ± 0.039
λ_3	-0.270 ± 0.074	-0.160 ± 0.033

(b)

Figure 4. Particle densities in a $5h^{-1}$ Mpc side slice of thickness $\pm 2h^{-1}$ Mpc showing the candidates $M31$ and MW in a single realisation, along the $X-Y$, $X-Z$ and $Y-Z$ planes.

Aim: To run a very large number of constrained simulations that 'mimic' the nearby universe.

Motivation: Statistics of look-alike simulations

Bayesian inference: posterior distribution

$$P\big(\mathrm{X} \mid \Lambda\mathrm{CDM}, \mathrm{Cosmicflows\ data},\ \mathrm{LG\ model}\big)$$

X = mass, tangential velocity, merging history, ...

The sampling of the posterior distribution function is done by looking for pairs of halos that obey the LG model at the LG position in constrained simulations.

Local Group model

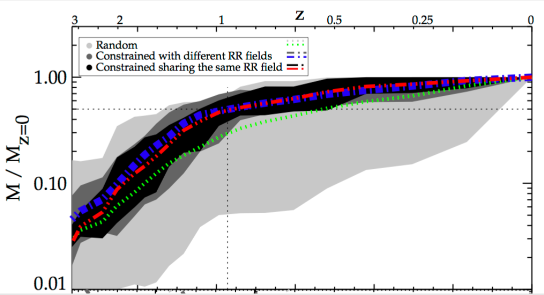
What is a LG?

- Simplest model: two halos, distance $d = (0.35 - 0.70)h^{-1}\mathrm{Mpc}$, physical radial velocity $v_r = (-135 \; - \; -80)\,\mathrm{km\ s^{-1}}$, isolation
- More advanced: add tangential velocity (v_{tan})
- Even more advanced and less observationally motivated: add mass
- More physical: (M, d, v_r, v_{tan}) $\longrightarrow$ (energy, angular momentum) i.e. an orbit.
- Observationally constrained physical model: fix the phase on the orbit (to get the correct d, v_r, v_{tan})
- Add galaxy formation considerations: e.g. quite recent merging history, disks

Bayesian inference: Virgo mass assembly history

“How did the Virgo cluster form?”

Sorce, Gottloeber, Hoffman & Yepes 2016)

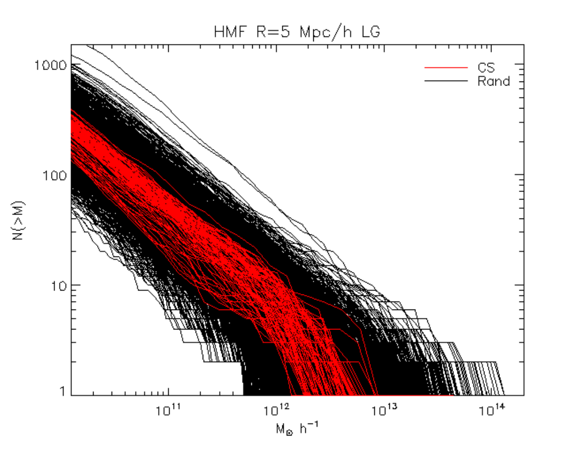


Virgo model

- How do we define a Virgo-like object?
- Mass - $\approx (1-5)h^{-1}\mathrm{M}_\odot$?
- possible prediction: mass assembly history (MMAH)

Bayesian inference: Local Neighbourhood mass function

(Carlesi, Hoffman, Sorce & Gottloeber, in preparation)



Local Neighbourhood model

- Sphere of $R = 5h^{-1}\mathrm{Mpc}$
- What else?
- possible prediction: mass function of DM halos

The Local Volume (Bland-Hawthorn & Freeman, 2013)

"A word of caution... to properly survey the effects of environment ..., we need to study resolved stellar populations out to at least 20 Mpc (Bland-Hawthorn & Freeman 2006). ... A physical scale of 20 Mpc ... cover[s] the full range of galaxy environments, from voids to massive groups and clusters. This is ... the Local Universe or Local Volume, now recognized by the International Astronomical Union (Division H). ... [T]his volume falls within the domain of the ... CLUES ... . All galaxies with masses equivalent to the LMC or larger can be imaged in most wavelength bands (e.g. x-rays, infrared, radio). ... In time, we fully expect near-field cosmological studies to extend to the Local Volume."

CLUES and the Local Volume

- The initial conditions of the Local Volume are very well constrained by the CLUES/CF2 machinery.
- In time, we fully expect near-field CONSTRAINED cosmological high resolution SIMULATIONS to extend to the Local Volume.
- The constrained simulations will test the Copernican hypothesis on the nature of the near-field.