Tomographic local 2D analyses of the WISExSuperCOSMOS all-sky galaxy catalogue

Camila P. Novaes[⋆], Armando Bernui[⋆],
Henrique S. Xavier[‡], Gabriela A. Marques[⋆]

[*MNRAS, 478, 3253 (2018)*]

[⋆]Observatório Nacional
[‡]IAG, USP

December 18, 2018

Motivation:

Structures like large voids, walls, filaments, and superclusters are expected to be present in the sky, but to disclose their features (space location, size, morphology, etc.) in the data sets is not an easy task.

Our aim here:

- To perform detailed analyses of the LSS density fields as traced by the distribution of galaxies.
- More specifically, to reveal regions with unexpected excess-of or lack-of luminous matter, suggestive of the presence of galaxies or giant voids, respectively.

$\rightarrow$ *Local analyses using the Minkowski Functionals.*

Data and simulations

The data set: WISExSuperCOSMOS (WSC) catalogue[1] (Bilicki et al. 2016) - constructed by cross-matching the currently largest all-sky photometric samples, namely, the WISE in the mid-infrared, and the SuperCOSMOS in the optical.

Why:

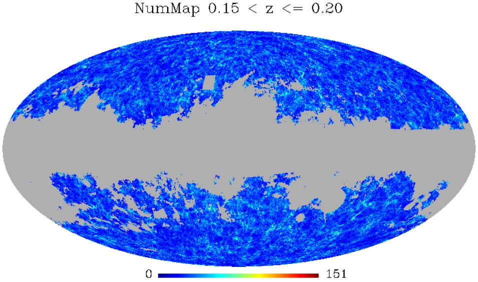
- Sky coverage: $f_{sky} \sim 0.55$,
- Number density of galaxies and angular resolution good enough for our local analyses.
- There are two versions of the WSC catalogue.

[1] http://ssa.roe.ac.uk/WISExSCOS

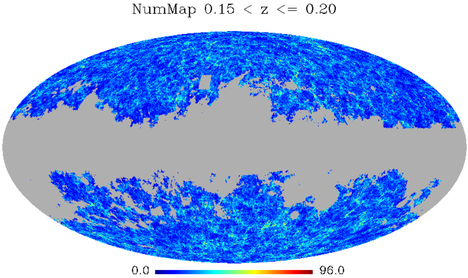
Data and simulations

The WSC galaxy number count maps:

WSC-*clean*



WSC-*svm*



Cleaning processes: color cuts and support vector machine (*svm*).

Data and simulations

The WSC data sets:

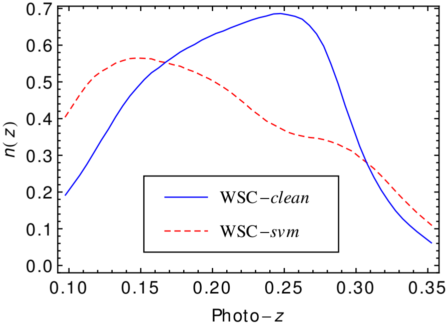
Bin #	photo-z range	WSC-*clean*		WSC-*svm*	
		counts	z_0	counts	z_0
1	0.10 - 0.15	1435113	0.130	2108545	0.127
2	0.15 - 0.20	2307948	0.176	2190296	0.174
3	0.20 - 0.25	2682398	0.226	1754763	0.223
4	0.25 - 0.30	2336109	0.272	1379258	0.274
5	0.30 - 0.35	719750	0.316	859436	0.320

Median redshits: z_0^{clean} = 0.22 and z_0^{svm} = 0.20.

Data and simulations

The simulated data: 5000 ΛCDM mock realizations of each photo-z bin generated using the FLASK code[2] (Xavier et al. 2016).

Using: a set of *auto- and cross-angular power spectra* $C_\ell^{i,j}$ for the number counts in all the photo-z bin i, j and the *radial selection function* of the survey.



[2] http://www.astro.iag.usp.br/∼flask

Minkowski Functionals

- The morphological properties of a given field $\mathcal{F}$ in a d-dimensional space can be described using d + 1 Minkowski Functionals (MF; Minkowski 1903).

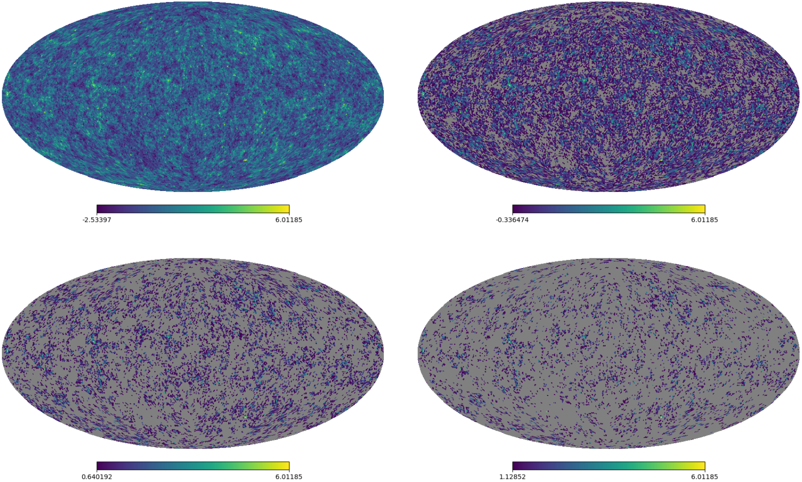
- Widely used to investigate the statistical properties of the 2D CMB temperature field and the 3D distribution of galaxies in the universe.

- Efficiently applied in masked skies or still in small regions.

- For the 2D CMB field ($\mathcal{F} = \Delta T$): given a connected region such that $\nu(\theta, \phi) = \Delta T/\sigma_0 > \nu_t$, the MF are
  - Area $\Rightarrow$ $V_0 = A(\nu) = \sum a_i$,
  - Perimeter $\Rightarrow$ $V_1 = P(\nu) = \sum l_i$,
  - Genus $\Rightarrow$ $V_2 = G(\nu) = \sum g_i = N_{hot} - N_{cold}$.

Minkowski Functionals

2D maps for different ν thresholds



Minkowski Functionals: our approach

3D galaxy catalogue

$\Downarrow$

2D galaxy number count maps
(in photo-z bins)

Using the HEALPix pixelization grid

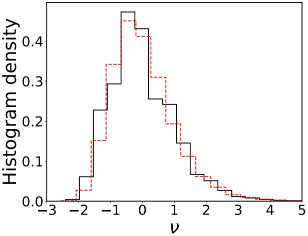
Resolution:

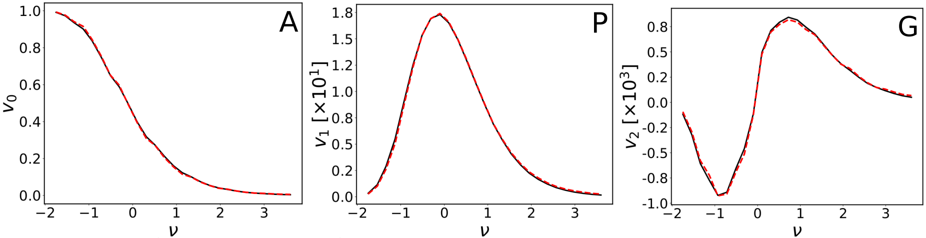
$N_{side} = 128 \rightsquigarrow \sim 2 \times 10^5$ pixels (27.5')

Minkowski Functionals: our approach

Example:

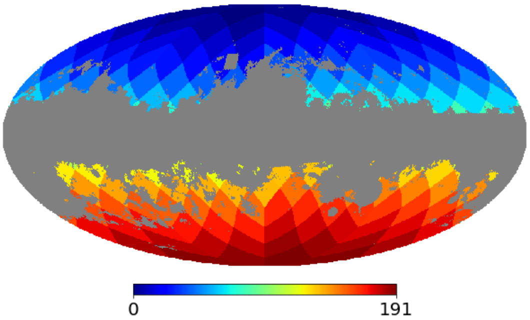
Masked WSC at $0.15 < z < 0.20$ (red line) and one of the corresponding mock realization (black line).

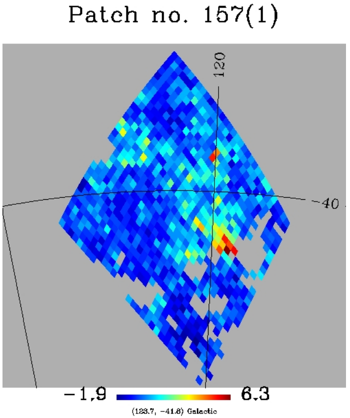




The local analyses

Considering the HEALPix resolution of $N_{side} = 4 \rightarrow$ 192 (126 analysed) patches ($\simeq 14.7^{\circ 2}$).





Identification and analyses of extreme regions

For the p-th patch and k-th MF:

$$\mathrm{v}_k^p \equiv (V_k(\nu_1), V_k(\nu_2), ...V_k(\nu_n))|_{\texttt{for the } p\texttt{-th patch}}$$

for $\{V_k, k = 0, 1, 2\} \equiv (V_0, V_1, V_2)$, $[\nu_{min}, \nu_{max}] = [-1.75, 3.6]$ and $n = 27$.

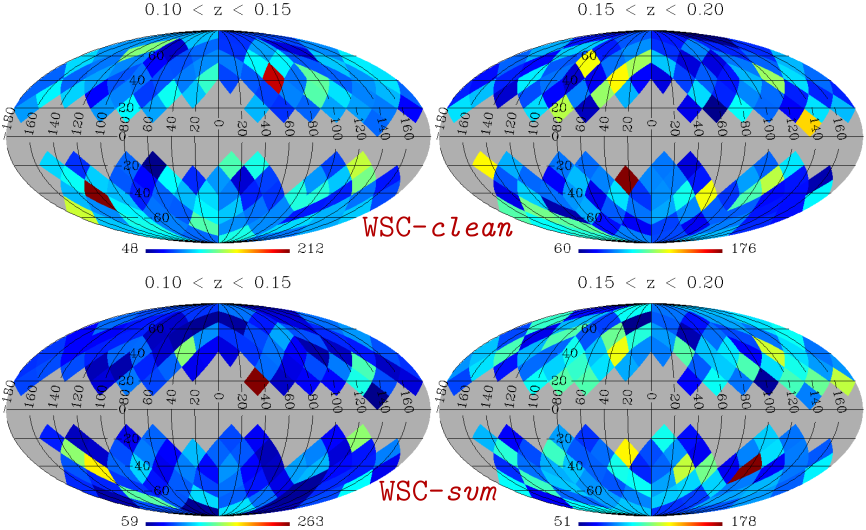
χ^2 analysis (Perimeter and Genus):

$$\begin{aligned}\mathcal{V}^p &\equiv (\mathrm{v}_0^p, \mathrm{v}_1^p, \mathrm{v}_2^p) \\ &= (V_0^p(\nu_1), ..., V_0^p(\nu_{27}), V_1^p(\nu_1), ..., V_1^p(\nu_{27}), V_2^p(\nu_1), ..., V_2^p(\nu_{27})),\end{aligned}$$

$$\chi^2 \equiv \sum_{i=1}^{81} \sum_{j=1}^{81} [\mathcal{V}_i^{\mathrm{WSC}} - \langle \mathcal{V}_i^{\mathrm{Mock}} \rangle] \, \mathsf{C}_{i,j}^{-1} \, [\mathcal{V}_j^{\mathrm{WSC}} - \langle \mathcal{V}_j^{\mathrm{Mock}} \rangle],$$

Identification and analyses of extreme regions

χ^2 value from each patch and z-bin $\rightarrow$ χ^2-map.



Identification and analyses of extreme regions

How extreme (anomalous?) are the patches with χ^2 far above the mean?

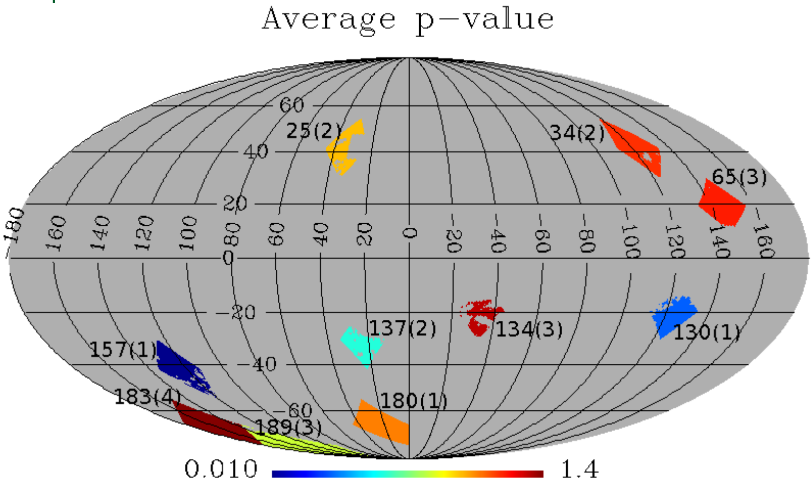
The significance: data *versus* simulations.

- χ^2-maps were constructed for each of the 5000 mock realizations of each sample (WSC-*clean* and WSC-*svm*) and photo-z bin.

- p-values were estimated by calculating the frequency of occurrences of the observed χ^2 amplitude in the mock realizations.

Identification and analyses of extreme regions

Selected extreme regions: p-value $< 1.4\%$ in both samples.



A total of 10 patches in the first 4 bins.

Identification and analyses of extreme regions

How extreme are the patches with χ^2 far above the mean?

Notice that:

- We found 10 patches with p-values < 1.4% among a total of 630 patches (126 in each of the 5 photo-z bins).
  ↘ ~38.6% of the simulations have at least 10 patches satisfying p-values < 1.4%.
- Patch no. 157(1): p-values = 0.01%.
  ↘ ~11.8% of the simulations have at least 1 patch (among the 630) satisfying p-values < 0.01%.

Still, these are extreme patches, very rare regions, what motivates their scrutiny in order to look for the possible reasons for their uncommon behaviour.

Detailed analyses of the selected regions

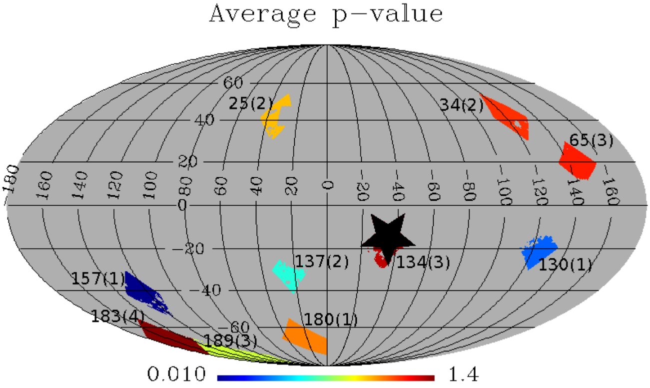
Aiming to investigate the amplitude and signature of the extreme regions of the WSC projected maps, we compare their MF vectors to the expected from ΛCDM mock realizations:

$$\Delta \mathrm{v}_k \equiv \frac{\mathrm{v}_k^{\mathrm{WSC}} - \langle \mathrm{v}_k \rangle}{\langle \mathrm{v}_k \rangle^{\mathrm{MAX}}},$$

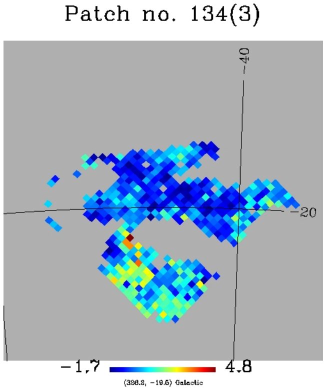
for $k = 0, 1, 2$.

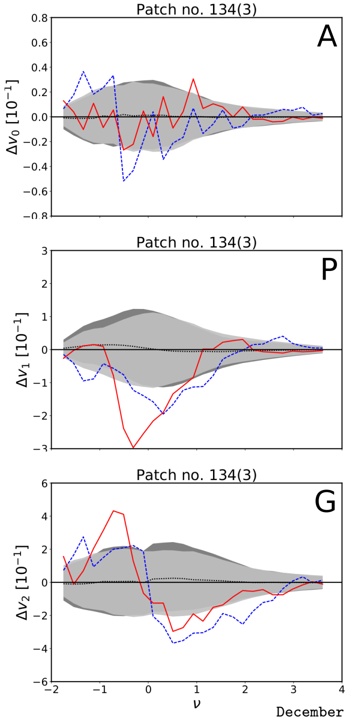
Near the

Galactic plain:
e.g. no. 134(3)

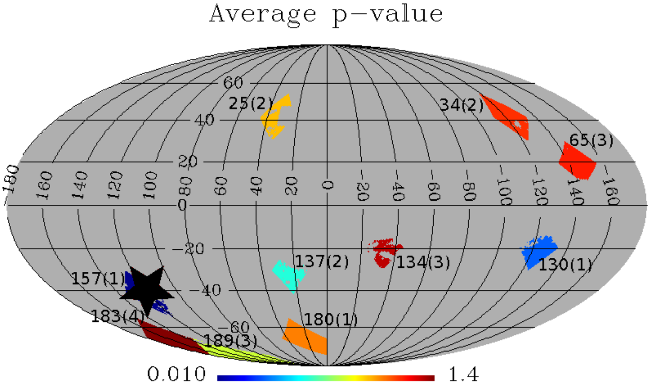


Near the Galactic plain: e.g. no. 134(3)

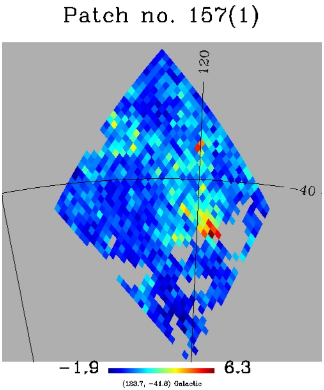




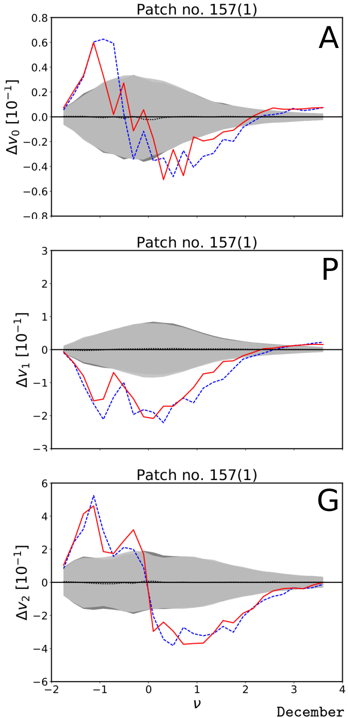
Far from the Galactic plain: e.g. no. 157(1)



Far from the Galactic plain: e.g. no. 157(1)



p-value = 0.01%

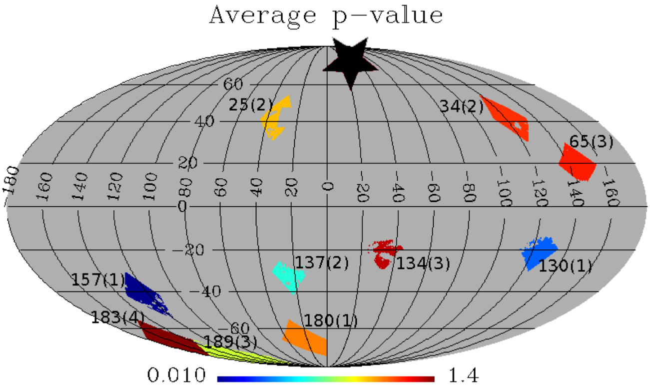


Identification and analyses of extreme regions

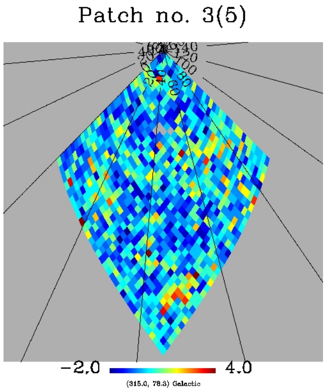
What can we infer from the behaviour of the MF curves?

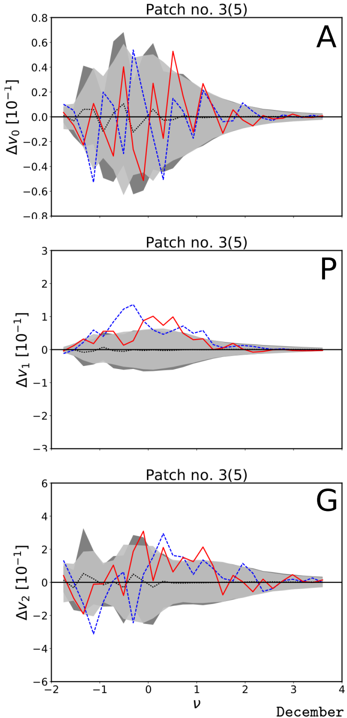
- $\nu < 0$ and $\nu > 0$ are associated to under- and over-densities, respectively.
- The curve of 2 or 3 MF extrapolating the 2σ region indicates the presence of some structure.
- The amplitude of the MF curves wrt the mean from simulations furnishes information about the structure responsible for the extremeness of a patch:
  - number of over-densities (under-dens.) respect to under-densities (over-dens.).
  - size and density.
- It is important to analyse the 3 MF features altogether.

The (noisy) last bin:
e.g. no. 3(5)



The (noisy) last bin: e.g. no. 3(5)





Conclusions

- We identified a total of 10 extreme patches, in 4 photo-z bins ($0.10 < z < 0.30$), in disagreement with the mocks, with p-value $<$ 1.4%.
- But, not in fact discrepant: our results indicate that the observed Universe as given by the WSC data is in agreement not only with the fiducial cosmological model, ΛCDM, but also with the contamination model, selection function, lognormal distribution of the objects, and other astrophysical features assumed to generate the mock realizations.

Not discrepant, but still extreme patches ...

Conclusions

From the detailed analyses we can conclude:

- The MF are highly efficient in obtaining a topological description of the distribution of galaxies in small regions of the sky: clustering of galaxies.
- The MF signature ($\Delta \mathrm{v}_k$) and the proximity of a patch to the galactic region allow to identify the presence of contamination as the reason for its extremeness.
- No extreme patches the last photo-z bin: shot-noise or the Universe at higher redshift, being less evolved, is better modelled by the mocks.

Conclusions

This confirms the capacity of the MF to discriminate features associated with the distributions of galaxies and those coming from contamination.

Our approach for mapping the galaxy distribution can be applied to a variety of catalogues, specially due to the advantage of the MF in not depending on the size of the analysed region.

Thanks!