# gigantes

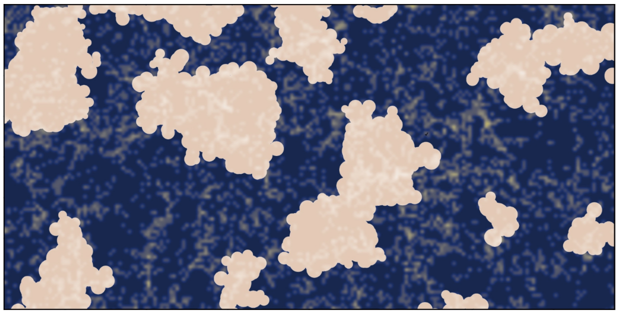
*Release 0.1*

Christina Kreisch, Alice Pisani, Francisco Villaescusa-Navarro

**Jul 21, 2023**

# SCIENCE

Gigantes is the most extensive and realistic void catalog suite ever released, containing over 1 billion cosmic voids covering a volume larger than the observable Universe, more than 20 TB of data, and created by running the void finder VIDE on the halos from the Quijote simulations. Gigantes provides the tools for intensive machine learning exploration of void statistics.



CHAPTER

ONE

# GOALS

Cosmic voids are a recent tool for the extraction of cosmological information from large scale-surveys. Theory development to prepare void analysis requires large void numbers. Voids are arguably the largest regions in the Universe, therefore huge simulation volumes are necessary.

Gigantes, built on halos from the massive Quijote simulation suite, aims to provide enough voids to study different void statistics in a realistic set-up. Examples of void statistics sensitive to cosmology are:

- the void-galaxy cross-correlation function,
- the void size function,
- the void auto-correlation function.

Gigantes will allow to:

- foster research on void evolution by providing a large number of voids at different redshifts and in different cosmologies;
- explore the information content of the different void statistics (see Applications);
- push the exploration for cosmology of a number of void properties (examples include ellipticities, density constrasts).

Finally, Gigantes is a massive dataset for intensive machine learning exploration of void statistics. The Gigantes void catalogs built from simulations with different cosmological models (the Quijote latin-hypercube) are instrumental to find a connection between void statistics and the value of cosmological parameters (see applications).

CHAPTER

# TWO

# PUBLICATIONS

1. **The Gigantes dataset: precision cosmology from voids in the machine learning era**
   Christina D. Kreisch, Alice Pisani, Francisco Villaescusa-Navarro, David N. Spergel, Benjamin D. Wandelt, Nico Hamaus, Adrian E. Bayer
   2107.02304
2. **Machine learning cosmology from void properties**
   Bonny Y. Wang, Alice Pisani, Francisco Villaescusa-Navarro, Benjamin D. Wandelt
   2212.06860

CHAPTER

THREE

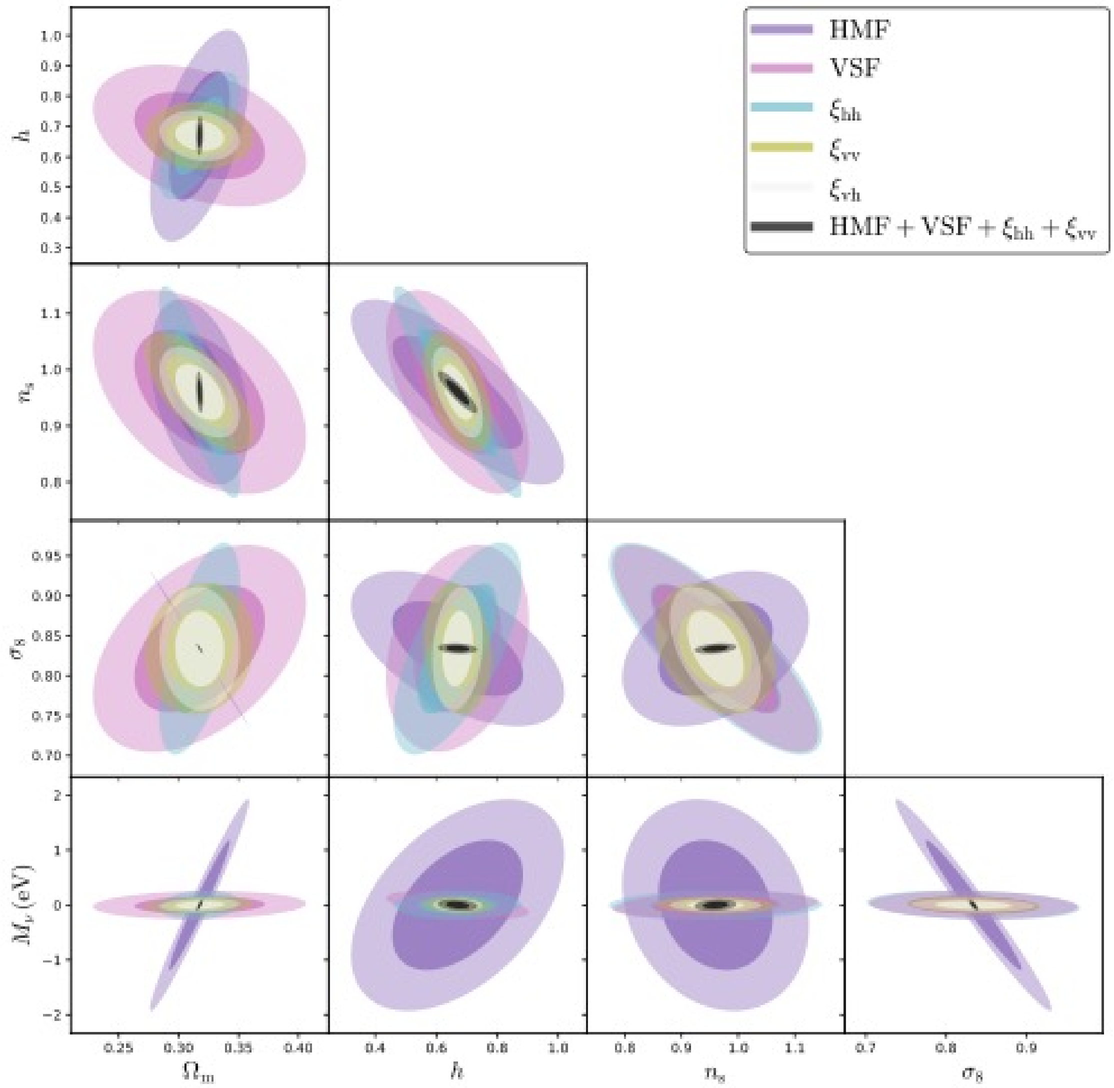
# APPLICATIONS

With the release of the Gigantes void catalogs suite we provide some example applications (see Kreisch et al. 2022). These are straightforward applications showcasing Gigantes' power and opening the way to exciting projects to perform with Gigantes.

## 3.1 Application 1: cosmological information in void statistics

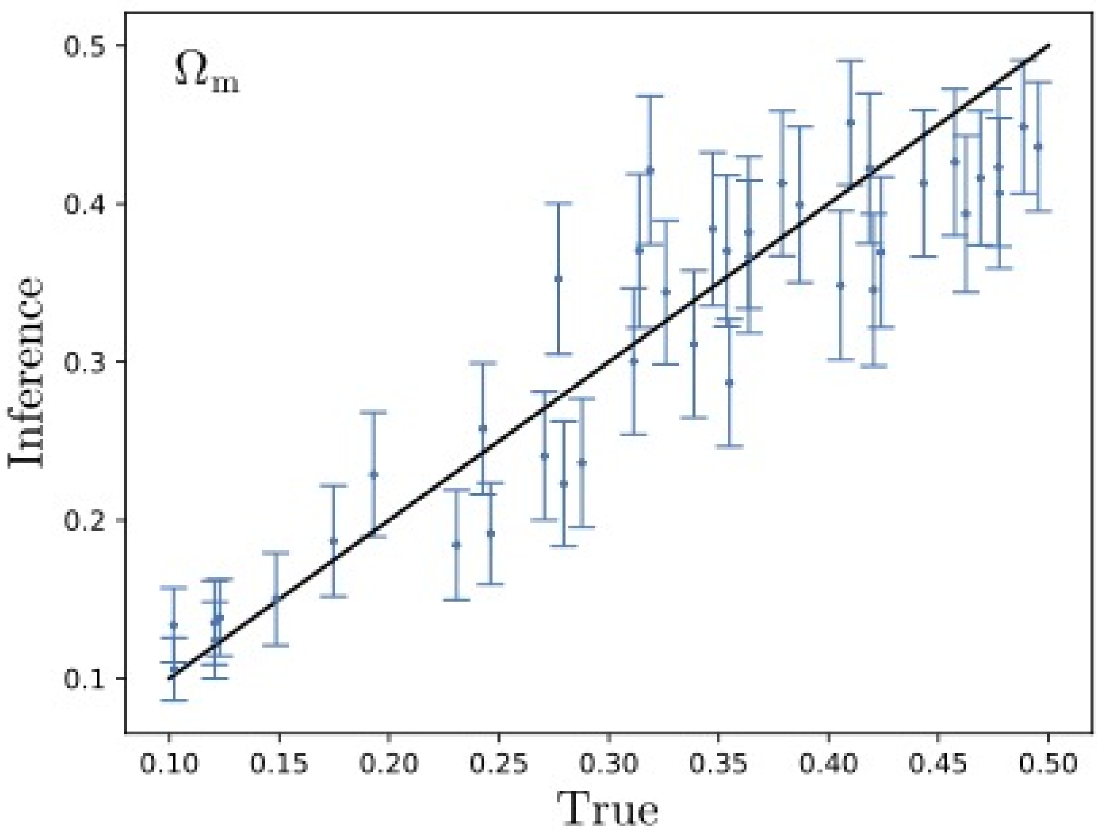
We aim to answer the following question: Do voids carry additional information? The question of whether voids, defined by relying on the halo (or galaxy) distribution, carry additional cosmological information with respect to traditional tools based on the halo (or galaxy) distribution, such as the two-point correlation function or the halo-mass function, has been a debated topic. Relying on a Fisher forecast from the Gigantes dataset we show that void statistics allow to constrain regions of the parameter space that would otherwise remain unconstrained.

The figure below shows constraints on cosmological parameters from the voids (void size function, void-halo, and void-void correlation functions) and halos (halo mass function, halo auto-correlation function), and the combined power of voids and halos.




## 3.2 Application 2: likelihood-free inference on the void size function

This application is an example of a machine learning application to the Gigantes dataset. Our goal is to perform likelihood-free inference from one of the most important summary statistics associated to cosmic voids: the void size function. In order to carry out this task, we need many examples from different cosmological models in order to be able to extract unique patterns that allow us to find a connection between the void size function and the value of the cosmological parameters. The goal of this application is to predict the mean and standard deviation of the posterior from the void size function.

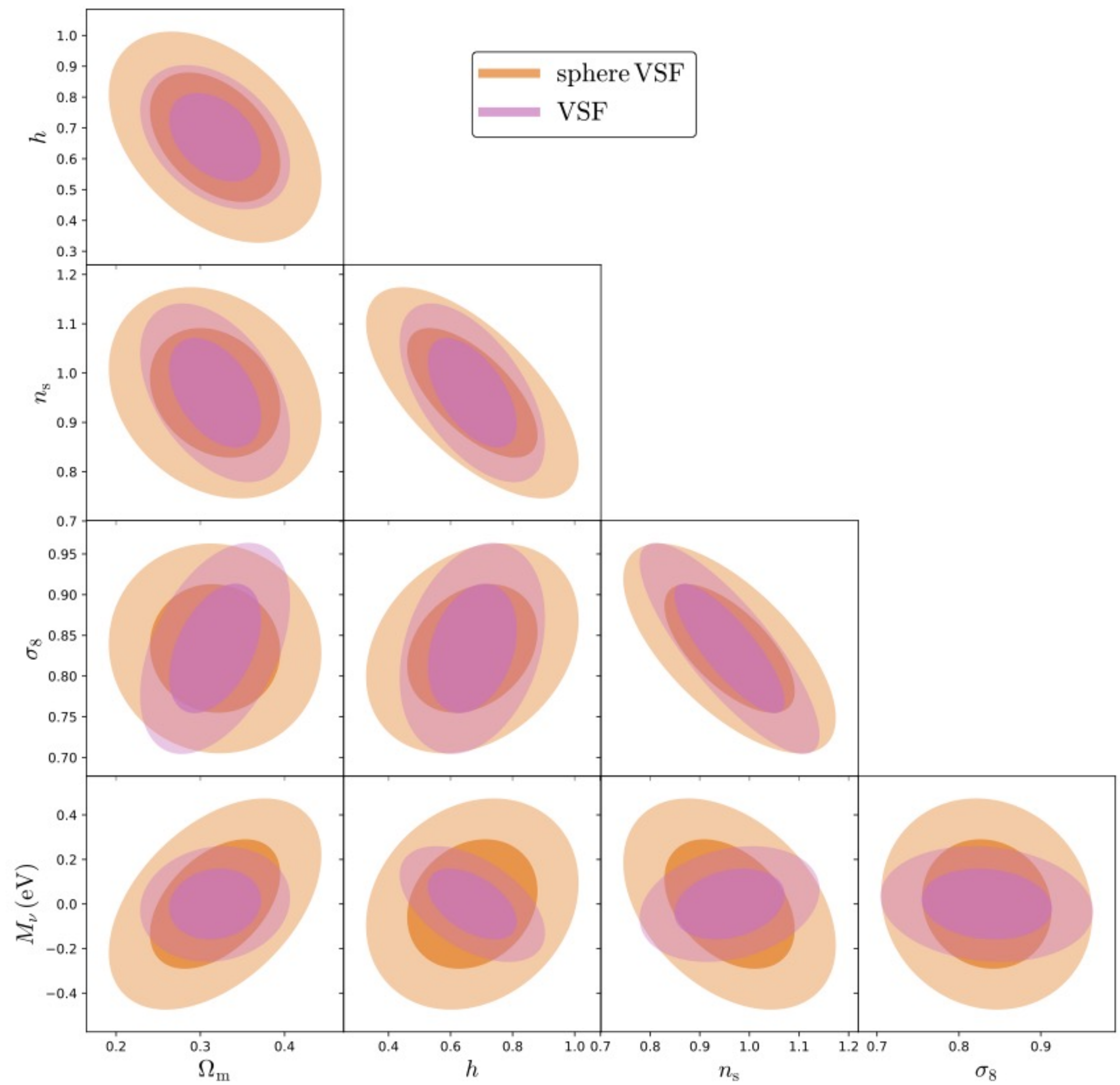



## 3.3 Application 3: void finder sensitivity to shape adds information

For this example we compare the information content captured when the void shape is measured in detail with the case in which a spherical assumption is made by the void finder. In other words we compare constraints obtained when selecting voids with VIDE, a void finder with no prior on void shape, and a more simplistic spherical-assumption based void finder.

For most of the cosmological parameters considered in this paper the void size function measured by VIDE provides more stringent constraints than the void size function measured by the spherical void finder.

These results showcase for the first time that even for non-shape based applications, such as the void size function, shape plays a strong role in determining the quality of constraints.




The large number of void catalogs in Gigantes allows many other scientific applications.

# CHAPTER FOUR

# DESCRIPTION

Void catalogs: Gigantes includes 15,000 VIDE fiducial cosmology void catalogs, as well as over 9,000 catalogs in non-fiducial cosmologies, spanning various values of the following cosmological parameters $\Omega_\mathrm{m}$, $\Omega_\mathrm{b}$, $h$, $n_s$, $\sigma_8$, $M_\nu$, and fully leveraging the Quijote simulation suite, which covers redshifts z = 0.0, 0.5, 1.0, and 2.0 in real and redshift space.

Void finder: The void finding relies on the popular public void finder VIDE (Sutter et al. 2015b), arguably the most used void finder, as testified by its use in a plethora of papers performing both simulation-based theoretical modelling and data analysis from modern surveys (Papers using VIDE).

Information about VIDE, used to build Gigantes can be found here : VIDE Wiki .

Provided statistics: Void center position (x,y,z—this is the volume-weighted barycenter, a.k.a. macrocenter), Void effective radius (Mpc/h), Void ID, void volume, redshift, ellipticity, density contrast, number of children (sub-voids), central density …

CHAPTER

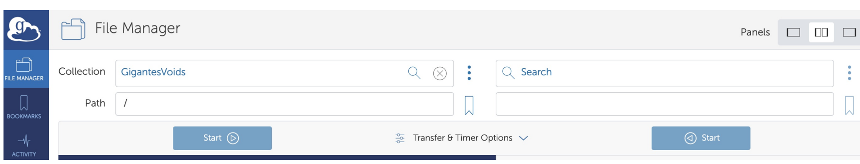
FIVE

# DOWNLOAD

The data can be accessed through globus:

- Log in into globus (create an account if you dont have one).
- To **access** the data, type: GigantesVoids (or with this link).


- To **transfer** the data, the simplest way is to use the globus graphical environment. Just type the above names in collection (GigantesVoids) or click the associated link. You will need to choose where the data is being moved in the other collection (e.g. your laptop or another supercomputer). Once the collection points are set, select the data you want to transfer and destiny folder and click in Start.




Note that to download the data to your local machine (e.g. laptop) you will need to install the globus connect personal.

CHAPTER

SIX

# ORGANIZATION

The Gigantes data is organized in different folders, following the structure of the Quijote simulations. The main folders are: fiducial, fiducial_ZA, fisher and latin_hypercube. In each folder it is possible to find a subfolder for each redshift considered, and void catalogs can be found there, labelled sample_[considered_simulation].

Each void catalog has the following structure (VIDE void catalog structure):

- sample_info.txt: Small text file which lists some basic info of the sample
- zobov_slice_[sample_name].dat: Binary file with the galaxy positions, RA, Dec, redshift, and unique catalog ID. Note that this file contains all particles in the particular sample, not just void particles.
- voidDesc_[all,central]_[sample_name].out: ASCII file with void indices and basic void properties in ZOBOV-normalized units.
- voidZone_[sample_name].out: Binary file which links voids to zones.
- voidPart_[sample_name].out: Binary file which links zones to particles.
- vol_[sample_name].out: Binary file with particle local volumes

In addition, VIDE provides the following derived void information:

- centers_[all, central]_[sample_name].txt: ASCII file with void centers, volume, effective radius, void ID, density contrast, etc.
- sky_positions_[all, central]_[sample_name].txt: ASCII file with void RA, dec, effective radius, and void ID
- shapes_[all, central]_[sample_name].txt: ASCII file with best-fit eigenvalues and eigenvectors for each void, in same coordinate system as x,y,z values in centers file

In addition to the [all,central] version of each file above, there are four versions of each catalog:

- no prefix: only parent voids, density cut applied (default catalog)
- untrimmed: all voids in hierarchy, density cut not applied
- untrimmed_dencut: all voids in hierarchy, density cut applied [deprecated]
- trimmed_nodencut: only parent voids, density cut not applied [deprecated]

# CHAPTER SEVEN

# TEAM

- Christina D. Kreisch (Citadel/Princeton University)
- Alice Pisani (Flatiron CCA/The Cooper Union/Princeton University)
- Francisco Villaescusa-Navarro (Flatiron/Princeton)
- David N. Spergel (Flatiron/Princeton)
- Benjamin D. Wandelt (IAP, Paris/Flatiron CCA)
- Nico Hamaus (USM Munich)
- Adrian E. Bayer (Princeton University/Flatiron CCA)

CHAPTER

# EIGHT

# CITATION

If you use data from Gigantes in your work, please consider citing the Gigantes paper and the Quijote paper. Thank you!

```
@ARTICLE{2022ApJ...935..100K,
     author = {{Kreisch}, Christina D. and {Pisani}, Alice and {Villaescusa-Navarro},␣
↪Francisco and {Spergel}, David N. and {Wandelt}, Benjamin D. and {Hamaus}, Nico and
↪{Bayer}, Adrian E.},
      title = "{The GIGANTES Data Set: Precision Cosmology from Voids in the Machine-
↪learning Era}",
    journal = {\apj},
   keywords = {N-body simulations, Cosmology, Large-scale structure of the universe,␣
↪Voids, Cosmic web, Computational astronomy, Cosmological parameters, Cosmological␣
↪parameters from large-scale structure, Cosmological neutrinos, Astrostatistics, 1083,␣
↪343, 902, 1779, 330, 293, 339, 340, 338, 1882, Astrophysics - Cosmology and␣
↪Nongalactic Astrophysics, Astrophysics - Instrumentation and Methods for Astrophysics},
       year = 2022,
      month = aug,
     volume = {935},
     number = {2},
        eid = {100},
      pages = {100},
        doi = {10.3847/1538-4357/ac7d4b},
        archivePrefix = {arXiv},
     eprint = {2107.02304},
primaryClass = {astro-ph.CO},
     adsurl = {https://ui.adsabs.harvard.edu/abs/2022ApJ...935..100K},
    adsnote = {Provided by the SAO/NASA Astrophysics Data System}
  }
```

```
@ARTICLE{Quijote_sims2020,
         author = {{Villaescusa-Navarro}, Francisco and {Hahn}, ChangHoon and {Massara},␣
↪Elena and {Banerjee}, Arka and {Delgado}, Ana Maria and {Ramanah}, Doogesh Kodi and
↪{Charnock}, Tom and {Giusarma}, Elena and {Li}, Yin and {Allys}, Erwan and {Brochard},␣
↪Antoine and {Uhlemann}, Cora and {Chiang}, Chi-Ting and {He}, Siyu and {Pisani}, Alice␣
↪and {Obuljen}, Andrej and {Feng}, Yu and {Castorina}, Emanuele and {Contardo},␣
↪Gabriella and {Kreisch}, Christina D. and {Nicola}, Andrina and {Alsing}, Justin and
↪{Scoccimarro}, Roman and {Verde}, Licia and {Viel}, Matteo and {Ho}, Shirley and
↪{Mallat}, Stephane and {Wandelt}, Benjamin and {Spergel}, David N.},
         title = "{The Quijote Simulations}",
         journal = {\apjs},
```

(continues on next page)

(continued from previous page)

```
        keywords = {N-body simulations, Cosmological parameters, Astrostatistics, Large-
↪scale structure of the universe, Cosmological neutrinos, 1083, 339, 1882, 902, 338,␣
↪Astrophysics - Cosmology and Nongalactic Astrophysics, Astrophysics - Instrumentation␣
↪and Methods for Astrophysics},
        year = 2020,
        month = sep,
        volume = {250},
        number = {1},
        eid = {2},
        pages = {2},
        doi = {10.3847/1538-4365/ab9d82},
        archivePrefix = {arXiv},
        eprint = {1909.05273},
        primaryClass = {astro-ph.CO},
        adsurl = {https://ui.adsabs.harvard.edu/abs/2020ApJS..250....2V},
        adsnote = {Provided by the SAO/NASA Astrophysics Data System}
}
```

CHAPTER

# NINE

# LICENSE

MIT License

CHAPTER

TEN

# HELP

For problems, questions and general help you can reach us at apisani@flatironinstitute.org .