The Quijote simulations is a suite of 45,500 full N-body simulations designed to:

- Quantify the information content on cosmological observables
- Provide enough statistics to train machine learning algorithms

Historically, Quijote was developed from the HADES simulations. Nowadays, it contains the full HADES data.
CHAPTER ONE

NEWS

**October 2022:** All the halo catalogues of the primordial non-Gaussianities simulations are now publicly available. Check *Primordial non-Gaussianities* for more details.

**September 2022:** All Quijote data located in the San Diego and New York clusters (almost 800 Terabytes) can now be accessed via Binder, a system that allows reading and manipulating the data without having to download it. Check *Data access* for further details.

**July 2022:** The snapshots of Quijote-PNG are now publicly available. Check *Primordial non-Gaussianities* for more details.

**June 2022:** The nwLH latin-hypercube, containing 2,000 simulations varying $\Omega_m$, $\Omega_b$, $h$, $n_s$, $\sigma_8$, $M_\nu$, $w$ is now publicly available! Check *Latin-hypercubes* for more details.
SCIENTIFIC GOALS

The Quijote simulations is a suite of more than 45,000 full N-body simulations that have been designed to accomplish two main goals:

- Quantify the information content on generic cosmological observables
- Provide enough data to train machine learning algorithms

For the first goal, Quijote provides a set of more than 35,000 simulations designed to calculate the information content on a generic cosmological observable by means of evaluating its Fisher matrix.

For the second goal, Quijote provides not only thousands of simulations on different latin-hypercubes, but the a total number of 44,100 N-body simulations, with billion of halos, galaxies, voids and millions of summary statistics such as power spectra, bispectra... et, to train machine learning algorithms, where having more data is always better.

The large number of simulations and data products available in Quijote allows many other scientific applications. See Publications for a list of different scientific usages of the data.
Chapter 2. Scientific goals
CHAPTER
THREE

PUBLICATIONS

1. SIMBIG: A Forward Modeling Approach To Analyzing Galaxy Clustering
   ChangHoon Hahn, Michael Eickenberg, Shirley Ho, Jiamin Hou, Pablo Lemos, Elena Massara, Chirag Modi,
   Azadeh Moradinezhad Dizgah, Bruno Régaldo-Saint Blancard, Muntazir M. Abidi
   2211.00723

2. SIMBIG: Mock Challenge for a Forward Modeling Approach to Galaxy Clustering
   ChangHoon Hahn, Michael Eickenberg, Shirley Ho, Jiamin Hou, Pablo Lemos, Elena Massara, Chirag Modi,
   Azadeh Moradinezhad Dizgah, Bruno Régaldo-Saint Blancard, Muntazir M. Abidi
   2211.00660

3. Cosmological Information in Skew Spectra of Biased Tracers in Redshift Space
   Jiamin Hou, Azadeh Moradinezhad Dizgah, ChangHoon Hahn, Elena Massara
   2210.12743

4. New applications of Graph Neural Networks in Cosmology
   Farida Farsian, Federico Marulli, Lauro Moscardini, Carlo Giocoli
   2210.11487

5. Tracer-Field Cross-Correlations with k-Nearest Neighbor Distributions
   Arka Banerjee, Tom Abel
   2210.05140

6. Squeezing $f_{NL}$ out of the matter bispectrum with consistency relations
   Samuel Goldstein, Angelo Esposito, Oliver H. E. Philcox, Lam Hui, J. Colin Hill, Roman Scoccimarro,
   Maximilian H. Abitbol
   2209.06228

7. Constraining CDM with density-split clustering
   Enrique Paillas, Carolina Cuesta-Lazaro, Pauline Zarrouk, Yan-Chuan Cai, Will J. Percival, Seshadri Nadathur,
   Mathilde Pinon, Arnaud de Mattia, Florian Beutler
   2209.04310

8. Bayesian evidence comparison for distance scale estimates
   Aseem Paranjape, Ravi K. Sheth
   2209.00668

9. Minkowski Tensors in Redshift Space – Beyond the Plane Parallel Approximation
   Stephen Appleby, Joby P. Kochappan, Pravabati Chingangbam, Changbom Park
   2208.10164

10. Correcting for small-displacement interlopers in BAO analyses
    Setareh Foroozan, Elena Massara, Will J. Percival
11. **Fast computation of non-linear power spectrum in cosmologies with massive neutrinos**
   Hernán E. Noriega, Alejandro Aviles, Sebastien Fromenteau, Mariana Vargas-Magaña
   2208.02791

12. **Estimating Cosmological Constraints from Galaxy Cluster Abundance using Simulation-Based Inference**
    Moonzarin Reza, Yuanyuan Zhang, Brian Nord, Jason Poh, Aleksandra Ciprijanovic, Louis Strigari
    2208.00134

13. **The Cosmic Graph: Optimal Information Extraction from Large-Scale Structure using Catalogues**
    T. Lucas Makinen, Tom Charnock, Pablo Lemos, Natalia Porqueres, Alan Heavens, Benjamin D. Wandelt
    2207.05202

14. **The Disordered Heterogeneous Universe: Galaxy Distribution and Clustering Across Length Scales**
    Oliver H. E. Philcox, Salvatore Torquato
    2207.00519

15. **Quijote PNG: The information content of the halo power spectrum and bispectrum**
    William R Coulton, Francisco Villaescusa-Navarro, Drew Jamieson, Marco Baldi, Gabriel Jung, Dionysios Karagiannis, Michele Liguori, Licia Verde, Benjamin D. Wandelt
    2206.15450

16. **Velocity profiles of matter and biased tracers around voids**
    Elena Massara, Will J. Percival, Neal Dalal, Seshadri Nadathur, Sladana Radinović, Hans A. Winther, Alex Woodfinden
    2206.14120

17. **Primordial non-Gaussianity and non-Gaussian Covariance**
    Thomas Floss, Matteo Biagetti, P. Daniel Meerburg
    2206.10458

18. **Field Level Neural Network Emulator for Cosmological N-body Simulations**
    Drew Jamieson, Yin Li, Renan Alves de Oliveira, Francisco Villaescusa-Navarro, Shirley Ho, David N. Spergel
    2206.04594

19. **Simple lessons from complex learning: what a neural network model learns about cosmic structure formation**
    Drew Jamieson, Yin Li, Siyu He, Francisco Villaescusa-Navarro, Shirley Ho, Renan Alves de Oliveira, David N. Spergel
    2206.04573

20. **Cosmological Information in the Marked Power Spectrum of the Galaxy Field**
    Elena Massara, Francisco Villaescusa-Navarro, ChangHoon Hahn, Muntazir M. Abidi, Michael Eickenberg, Shirley Ho, Pablo Lemos, Azadeh Moradinezhad Dizgah, Bruno Regaldo-Saint Blancard
    2206.01709

21. **Quijote-PNG: Quasi-maximum likelihood estimation of Primordial Non-Gaussianity in the non-linear dark matter density field**
    Gabriel Jung, Dionysios Karagiannis, Michele Liguori, Marco Baldi, William R Coulton, Drew Jamieson, Licia Verde, Francisco Villaescusa-Navarro, Benjamin D. Wandelt
    2206.01624

22. **Quijote-PNG: Simulations of primordial non-Gaussianity and the information content of the matter field power spectrum and bispectrum**
William R Coulton, Francisco Villaescusa-Navarro, Drew Jamieson, Marco Baldi, Gabriel Jung, Dionysios Karagiannis, Michele Liguori, Licia Verde, Benjamin D. Wandelt
2206.01619

23. Accurate predictions from small boxes: variance suppression via the Zel’dovich approximation
Nickolas Kokron, Shi-Fan Chen, Martin White, Joseph DeRose, Mark Maus
2205.15327

24. Robust Neural Network-Enhanced Estimation of Local Primordial Non-Gaussianity
Utkarsh Giri, Moritz Münchmeyer, Kendrick M. Smith
2205.12964

25. Two-loop power spectrum with full time- and scale-dependence and EFT corrections: impact of massive neutrinos and going beyond EdS
Mathias Garny, Petter Taule
2205.11533

26. Improving cosmological covariance matrices with machine learning
Natali S.M. de Santi, L. Raul Abramo
2205.10881

27. Fast and realistic large-scale structure from machine-learning-augmented random field simulations
Davide Piras, Benjamin Joachimi, Francisco Villaescusa-Navarro
2205.07898

28. Distinguishing Dirac vs. Majorana Neutrinos: a Cosmological Probe
Beatriz Hernandez-Molinero, Raul Jimenez, Carlos Pena-Garay
2205.00808

29. Accurate Model of the Projected Velocity Distribution of Galaxies in Dark Matter Halos
Han Aung, Daisuke Nagai, Eduardo Rozo, Brandon Wolfe, Susmita Adhikari
2204.13131

30. Wavelet Moments for Cosmological Parameter Estimation
Michael Eickenberg, Erwan Allys, Azadeh Moradinezhad Dizgah, Pablo Lemos, Elena Massara, Muntazir Abidi, ChangHoon Hahn, Sultan Hassan, Bruno Regaldo-Saint Blancard, Shirley Ho, Stephane Mallat, Joakim Andén, Francisco Villaescusa-Navarro
2204.07646

31. Quantification of high dimensional non-Gaussianities and its implication to Fisher analysis in cosmology
Core Francisco Park, Erwan Allys, Francisco Villaescusa-Navarro, Douglas P. Finkbeiner
2204.05435

32. Bayesian Control Variates for optimal covariance estimation with pairs of simulations and surrogates
Nicolas Chartier, Benjamin D. Wandelt
2204.03070

33. Probing massive neutrinos with the Minkowski functionals of large-scale structure
Wei Liu, Aoxiang Jiang, Wenjuan Fang
2204.02945

34. Perturbation Theory vs Simulation: Quasi-linear Scale, Binning Effect, and Visualization of Bispectrum
Joseph Tomlinson, Donghui Jeong
2204.00668

35. The effect of local universe constraints on halo abundance and clustering
Maxwell L. Hutt, Harry Desmond, Julien Devriendt, Adrianne Slyz
2203.14724

36. **Extracting high-order cosmological information in galaxy surveys with power spectra**
Yuting Wang, Gong-Bo Zhao, Kazuya Koyama, Will J. Percival, Ryuichi Takahashi, Chiaki Hikage, Héctor Gil-Marín, ChangHoon Hahn, Ruiyang Zhao, Weibing Zhang, Xiaoyong Mu, Yu Yu, Hong-Ming Zhu, Fei Ge
2202.05248

37. **Constraining cosmological parameters from N-body simulations with Bayesian Neural Networks**
Hector J. Hortua
2112.11865

38. **Detection of spatial clustering in the 1000 richest SDSS DR8 redMaPPer clusters with Nearest Neighbor distributions**
Yunchong Wang, Arka Banerjee, Tom Abel
2112.04502

39. **One-point statistics matter in extended cosmologies**
Alex Gough, Cora Uhlemann
2112.04428

40. **Cosmology with cosmic web environments I. Real-space power spectra**
Tony Bonnaire, Nabila Aghanim, Joseph Kuruvilla, Aurélien Decelle
2112.03926

41. **The Information Content of Projected Galaxy Fields**
Lucas Porth, Gary M. Bernstein, Robert E. Smith, Abigail J. Lee
2111.13702

42. **Cosmology and neutrino mass with the Minimum Spanning Tree**
Krishna Naidoo, Elena Massara, Ofer Lahav
2111.12088

43. **The Covariance of Squeezed Bispectrum Configurations**
Matteo Biagetti, Lina Castiblanco, Jorge Noreña, Emiliano Sefusatti
2111.05887

44. **NECOLA: Towards a Universal Field-level Cosmological Emulator**
Neerav Kaushal, Francisco Villaescusa-Navarro, Elena Giusarma, Yin Li, Conner Hawry, Mauricio Reyes
2111.02441

45. **The smearing scale in Laguerre reconstructions of the correlation function**
Farnik Nikakhtar, Ravi K. Sheth, Idit Zehavi
2110.03591

46. **Cosmology with the kinetic Sunyaev-Zeldovich effect: Independent of the optical depth and $\sigma_8$**
Joseph Kuruvilla
2109.13938

47. **Creating Jackknife and Bootstrap estimates of the covariance matrix for the two-point correlation function**
Faizan G. Mohammad, Will J. Percival
2109.07071

48. **The matter density PDF for modified gravity and dark energy with Large Deviations Theory**
Matteo Cataneo, Cora Uhlemann, Christian Arnold, Alex Gough, Baojiu Li, Catherine Heymans
49. **Towards an Optimal Estimation of Cosmological Parameters with the Wavelet Scattering Transform**  
   Georgios Valogiannis, Cora Dvorkin  
   2108.07821

50. **Beware of Fake $\nu_s$: The Effect of Massive Neutrinos on the Non-Linear Evolution of Cosmic Structure**  
   Adrian E. Bayer, Arka Banerjee, Uros Seljak  
   2108.04215

51. **The effects of peculiar velocities on the morphological properties of large scale structures**  
   Aoxiang Jiang, Wei Liu, Wenjuan Fang, Wen Zhao  
   2108.03851

52. **Analytic Gaussian Covariance Matrices for Galaxy N-Point Correlation Functions**  
   Jiamin Hou, Robert N. Cahn, Oliver H.E. Philcox, Zachary Slepian  
   2108.01714

53. **Modeling Nearest Neighbor distributions of biased tracers using Hybrid Effective Field Theory**  
   Arka Banerjee, Nickolas Kokron, Tom Abel  
   2107.10287

54. **The reach of next-to-leading-order perturbation theory for the matter bispectrum**  
   Davit Alkhanishvili, Cristiano Porciani, Emiliano Sefusatti, Matteo Biagetti, Andrei Lazanu, Andrea Oddo, and Victoria Yankelevich  
   2107.08054

55. **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

56. **The PDF perspective on the tracer-matter connection: Lagrangian bias and non-Poissonian shot noise**  
   Oliver Friedrich, Anik Halder, Aoife Boyle, Cora Uhlemann, Dylan Britt, Sandrine Codis, Daniel Gruen, ChangHoon Hahn  
   2107.02300

57. **Clustering in Massive Neutrino Cosmologies via Eulerian Perturbation Theory**  
   Alejandro Aviles, Arka Banerjee, Gustavo Niz, Zachary Slepian  
   2106.13771

58. **CARPool Covariance: Fast, unbiased covariance estimation for large-scale structure observables**  
   Nicolas Chartier, Benjamin D. Wandelt  
   2106.11718

59. **Extracting cosmological parameters from N-body simulations using machine learning techniques**  
   Andrei Lazanu  
   2106.11061

60. **Unsupervised Resource Allocation with Graph Neural Networks**  
   Miles Cranmer, Peter Melchior, Brian Nord  
   2106.09761

61. **Normalizing flows for random fields in cosmology**  
   Adam Rouhiainen, Utkarsh Giri, Moritz Münchmeyer  
   2105.12024
62. Joint analysis of anisotropic power spectrum, bispectrum and trispectrum: application to N-body simulations
   Davide Gualdi, Hector Gil-Marin, Licia Verde
   2104.03976

63. Clustering and halo abundances in early dark energy cosmological models
   Anatoly Klypin, Vivian Poulin, Francisco Prada, Joel Primack, Marc Kamionkowski, Vladimir Avila-Reese,
   Aldo Rodriguez-Puebla, Peter Behroozi, Doug Hellinger, Tristan L Smith
   MNRAS article

64. Detecting the radiative decay of the cosmic neutrino background with line-intensity mapping
   Jose Luis Bernal, Andrea Caputo, Francisco Villaescusa-Navarro, Marc Kamionkowski
   2103.12099

65. Information content in mean pairwise velocity and mean relative velocity between pairs in a triplet
   Joseph Kuruvilla, Nabila Aghanim
   2102.06709

66. Detecting neutrino mass by combining matter clustering, halos, and voids
   Adrian E. Bayer, Francisco Villaescusa-Navarro, Elena Massara, Jia Liu, David N. Spergel, Licia Verde,
   Benjamin Wandelt, Matteo Viel, Shirley Ho
   2102.05049

67. Information Content of Higher-Order Galaxy Correlation Functions
   Lado Samushia, Zachary Slepian, Francisco Villaescusa-Navarro
   2102.01696

68. Cosmological cross-correlations and nearest neighbor distributions
   Arka Banerjee, Tom Abel
   2102.01184

69. Learning the Evolution of the Universe in N-body Simulations
   Chang Chen, Yin Li, Francisco Villaescusa-Navarro, Shirley Ho, Anthony Pullen
   2012.05472

70. Constraining $M_\nu$ with the Bispectrum II: The Total Information Content of the Galaxy Bispectrum
   ChangHoon Hahn, Francisco Villaescusa-Navarro
   2012.02200

71. Fast and Accurate Non-Linear Predictions of Universes with Deep Learning
   Renan Alves de Oliveira, Yin Li, Francisco Villaescusa-Navarro, Shirley Ho, David N. Spergel
   2012.00240

72. Minkowski functionals and the nonlinear perturbation theory in the large-scale structure: second-order effects
   Takahiko Matsubara, Chiaki Hikage, Satoshi Kuriki
   2012.00203

73. The unequal-time matter power spectrum: impact on weak lensing observables
   Lucia F. de la Bella, Nicolas Tessore, Sarah Bridle
   2011.06185

74. Exploring KSZ velocity reconstruction with N-body simulations and the halo model
   Utkarsh Giri, Kendrick M. Smith
   2010.07193
75. **Modeling the Marked Spectrum of Matter and Biased Tracers in Real- and Redshift-Space**  
Oliver H.E. Philcox, Alejandro Aviles, Elena Massara  
2010.05914

76. **CARPool: fast, accurate computation of large-scale structure statistics by pairing costly and cheap cosmological simulations**  
Nicolas Chartier, Benjamin Wandelt, Yashar Akrami, Francisco Villaescusa-Navarro  
2009.08970

77. **Matter trispectrum: theoretical modelling and comparison to N-body simulations**  
Davide Gualdi, Sergi Novell, Héctor Gil-Marín, Licia Verde  
2009.02290

78. **The impact of massive neutrinos on halo assembly bias**  
Titouan Lazeyras, Francisco Villaescusa-Navarro, Matteo Viel  
2008.12265

79. **Capturing the Cosmic Web for Cosmology**  
Krishna Naidoo  
1829731

80. **Nearest Neighbor distributions: new statistical measures for cosmological clustering**  
Arka Banerjee, Tom Abel  
2007.13342

81. **The effects of massive neutrinos on the linear point of the correlation function**  
G. Parimbelli, S. Anselmi, M. Viel, C. Carbone, F. Villaescusa-Navarro, P.S. Corasaniti, Y. Rasera, R. Sheth,  
G.D. Starkman, I. Zehavi  
2007.10345

82. **A Lagrangian Perturbation Theory in the presence of massive neutrinos**  
Alejandro Aviles, Arka Banerjee  
2007.06508

83. **Discovering Symbolic Models from Deep Learning with Inductive Biases**  
Miles Cranmer, Alvaro Sanchez-Gonzalez, Peter Battaglia, Rui Xu, Kyle Cranmer, David Spergel, Shirley Ho  
2006.11287

84. **What does the marked power spectrum measure? Insights from perturbation theory**  
Oliver H.E. Philcox, Elena Massara, David N. Spergel  
2006.10055

85. **New Interpretable Statistics for Large Scale Structure Analysis and Generation**  
E. Allys, T. Marchand, J.-F. Cardoso, F. Villaescusa-Navarro, S. Ho, S. Mallat  
2006.06298

86. **A Faster Fourier Transform? Computing Small-Scale Power Spectra and Bispectra for Cosmological Simulations in $O(N^2)$ Time**  
Oliver H.E. Philcox  
2005.01739

87. **Effective halo model: Creating a physical and accurate model of the matter power spectrum and cluster counts**  
Oliver H.E. Philcox, David N. Spergel, Francisco Villaescusa-Navarro  
2004.09515
88. **What Can We Learn by Combining the Skew Spectrum and the Power Spectrum?**
   Ji-Ping Dai, Licia Verde, Jun-Qing Xia
   2002.09904

89. **Using the Marked Power Spectrum to Detect the Signature of Neutrinos in Large-Scale Structure**
   Elena Massara, Francisco Villaescusa-Navarro, Shirley Ho, Neal Dalal, David N. Spergel
   2001.11024

90. **Super-resolution emulator of cosmological simulations using deep physical models**
   Doogesh Kodi Ramanah, Tom Charnock, Francisco Villaescusa-Navarro, Benjamin D. Wandelt
   2001.05519

91. **Primordial non-Gaussianity without tails – how to measure fNL with the bulk of the density PDF**
   Oliver Friedrich, Cora Uhlemann, Francisco Villaescusa-Navarro, Tobias Baldauf, Marc Manera, Takahiro Nishimichi
   1912.06621

92. **Fisher for complements: Extracting cosmology and neutrino mass from the counts-in-cells PDF**
   Cora Uhlemann, Oliver Friedrich, Francisco Villaescusa-Navarro, Arka Banerjee, Sandrine Codis
   1911.11158

93. **Learning neutrino effects in Cosmology with Convolutional Neural Networks**
   Elena Giusarma, Mauricio Reyes Hurtado, Francisco Villaescusa-Navarro, Siyu He, Shirley Ho, ChangHoon Hahn
   1910.04255

94. **Constraining $M_\nu$ with the bispectrum. Part I. Breaking parameter degeneracies**
   ChangHoon Hahn, Francisco Villaescusa-Navarro, Emanuele Castorina, Roman Scoccimarro
   1909.11107

95. **Weighing neutrinos with the halo environment**
   Arka Banerjee, Emanuele Castorina, Francisco Villaescusa-Navarro, Travis Court, Matteo Viel
   1907.06598

96. **Anisotropic halo assembly bias and redshift-space distortions**
   Andrej Obuljen, Neal Dalal, Will J. Percival
   1906.11823

97. **The Quijote simulations**
   Francisco Villaescusa-Navarro, ChangHoon Hahn, Elena Massara, Arka Banerjee, Ana Maria Delgado,
   Doogesh Kodi Ramanah, Tom Charnock, Elena Giusarma, Yin Li, Erwan Allys, Antoine Brochard, Cora Uhlemann, Chi-Ting Chiang, Siyu He, Alice Pisani, Andrej Obuljen, Yu Feng, Emanuele Castorina, Gabriella Contardo, Christina D. Kreisch, Andrina Nicola, Justin Alsing, Roman Scoccimarro, Licia Verde, Matteo Viel, Shirley Ho, Stephane Mallat, Benjamin Wandelt, David N. Spergel
   1909.05273
CHAPTER
FOUR

FEATURES

• Simulations run with the TreePM code Gadget-III
• More than 40 Million CPU hours used
• Boxes of 1 Gpc/h. Combined total volume of more than 45,000 (Gpc/h)^3 at a single redshift
• 17,100 simulations for a fiducial Planck cosmology
• Between 500 and 1,000 simulations/cosmology for 27 different cosmologies
• 1,000 Separate Universe simulations
• 8,000 simulations in different latin-hypercubes
• More than 10 trillions of particles at a single redshift from all simulations
• Billions of halos and voids identified
• Full snapshots at redshifts 0, 0.5, 1, 2, 3 and 127 (initial conditions)
• More than 200,000 halo catalogues
• More than 200,000 void catalogues
• More than 1 million power spectra
• More than 1 million bispectra
• More than 1 million correlation functions
• More than 1 million marked power spectra
• More than 1 million probability distribution functions
• More than 1 Petabyte of data publicly available
The Quijote data is organized into different folders:

- **Snapshots.** This folder contains the snapshots of the simulations
- **Halos.** This folder contains the halo catalogues
- **voids.** This folder contains the void catalogues
- **Linear_Pk.** This folder contains the linear power spectra of each cosmological model
- **Pk.** This folder contains the non-linear power spectra
- **Marked_Pk.** This folder contains the marked power spectra
- **Bk.** This folder contains the bispectra
- **CF.** This folder contains the correlation functions
- **PDF.** This folder contains the pdfs
- **3D_cubes.** This folder contains the 3D density fields

Each of the above folders contain several subfolders, that represent the different cosmological models, e.g. \( h\_p \), \( \text{fiducial} \), and \( \text{Om}\_m \).

### 5.1 Cosmological models

A brief description of the different cosmologies is provided in the below table. The standard and paired fixed snapshots or data products will be located inside the same folder. The paired fixed (or fixed) will be located inside folders starting with NCV (from No Cosmic Variance). Further details can be found in the Quijote paper.

| Name            | \( \Omega_m \) | \( \Omega_b \) | \( h \)  | \( n_s \) | \( \sigma_8 \) | \( M_\nu \) | \( w \) | \( \delta_b \) | \( f_{\text{loc}}^{\text{NL}} \) |
|-----------------|----------------|----------------|---------|---------|-------------|-----------|------|----------------|------------------|
| fiducial        | 0.3175         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| fiducial        | 0.3175         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| fiducial_ZA     | 0.3175         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| fiducial_LR     | 0.3175         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| fiducial_HR     | 0.3175         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| Om_p            | 0.3275         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| Om_p            | 0.3275         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| Om_m            | 0.3075         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| Om_m            | 0.3075         | 0.049          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| Ob2_p           | 0.3175         | 0.051          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
| Ob2_p           | 0.3175         | 0.051          | 0.6711  | 0.9624  | 0.834       | 0         | -1   | 0              | 0                |
Quijote simulations, Release 0.1

| Name   | $\Omega_m$ | $\Omega_b$ | $h$    | $n_s$ | $\sigma_8$ | $M_\nu$ | $w$ | $\delta_b$ | $\beta_{loc}$ |
|--------|------------|------------|--------|-------|-----------|---------|----|------------|--------------|
| Ob2_m  | 0.3175     | 0.047      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| Ob2_m  | 0.3175     | 0.047      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| Ob_p   | 0.3175     | 0.050      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| Ob_m   | 0.3175     | 0.048      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| h_p    | 0.3175     | 0.049      | 0.6911 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| h_p    | 0.3175     | 0.049      | 0.6911 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| h_m    | 0.3175     | 0.049      | 0.6511 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| h_m    | 0.3175     | 0.049      | 0.6511 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| ns_p   | 0.3175     | 0.049      | 0.6711 | 0.9824| 0.834     | 0       | -1 | 0          | 0            |
| ns_p   | 0.3175     | 0.049      | 0.6711 | 0.9824| 0.834     | 0       | -1 | 0          | 0            |
| ns_m   | 0.3175     | 0.049      | 0.6711 | 0.9424| 0.834     | 0       | -1 | 0          | 0            |
| ns_m   | 0.3175     | 0.049      | 0.6711 | 0.9424| 0.834     | 0       | -1 | 0          | 0            |
| s8_p   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.849     | 0       | -1 | 0          | 0            |
| s8_p   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.849     | 0       | -1 | 0          | 0            |
| s8_m   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.819     | 0       | -1 | 0          | 0            |
| s8_m   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.819     | 0       | -1 | 0          | 0            |
| Mnu_p  | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0.1     | -1 | 0          | 0            |
| Mnu_p  | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0.1     | -1 | 0          | 0            |
| Mnu_pp | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0.2     | -1 | 0          | 0            |
| Mnu_pp | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0.2     | -1 | 0          | 0            |
| Mnu_ppp| 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0.4     | -1 | 0          | 0            |
| Mnu_ppp| 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0.4     | -1 | 0          | 0            |
| w_p    | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1.05 | 0          | 0            |
| w_m    | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -0.95 | 0          | 0            |
| DC_p   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0.035      | 0            |
| DC_m   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | -0.035     | 0            |
| LC_p   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0.100      | 0            |
| LC_m   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | -100       | 0            |
| EQ_p   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| EQ_m   | 0.3175     | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| OR_CMB_p| 0.3175    | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| OR_CMB_m| 0.3175   | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| OR_LSS_p| 0.3175   | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| OR_LSS_m| 0.3175   | 0.049      | 0.6711 | 0.9624| 0.834     | 0       | -1 | 0          | 0            |
| latin_hypercube | [0.1 - 0.5] | [0.03 - 0.07] | [0.5 - 0.9] | [0.8 - 1.2] | [0.6 - 1.0] | 0 | -1 | 0          | 0            |
| latin_hypercube | [0.1 - 0.5] | [0.03 - 0.07] | [0.5 - 0.9] | [0.8 - 1.2] | [0.6 - 1.0] | 0 | -1 | 0          | 0            |
| latin_hypercube | [0.1 - 0.5] | [0.03 - 0.07] | [0.5 - 0.9] | [0.8 - 1.2] | [0.6 - 1.0] | 0 | -1 | 0          | 0            |
| nwLH   | [0.1 - 0.5] | [0.03 - 0.07] | [0.5 - 0.9] | [0.8 - 1.2] | [0.6 - 1.0] | [0.01 - 1.0] | [-1.3 - -0.7] | 0 | 0            |

See *Latin-hypercubes* and *Primordial non-Gaussianities* for more information about the latin-hypercubes and the simulations with primordial non-Gaussianities.
Quijote contains over 1 petabyte of data. Given this large size, the data is currently distributed across three different clusters in New York (Rusty cluster), San Diego (GordonS cluster), and Princeton (Tiger cluster). The data can be accessed in two different ways:

- **Globus.** A system designed to easily transfer large amounts of data in a very efficient manner.
- **Binder.** A system that allows reading and manipulating the data online, without the need to download the data.

The table below describes the data each cluster contains and provides the links to the associated globus and binder systems.

| Cluster      | Content                                                                 | Access     |
|--------------|-------------------------------------------------------------------------|------------|
| New York     | - The snapshots of high-resolution latin-hypercube                      | globus     |
|              | - The snapshots of the nwLH latin-hypercube                             |            |
|              | - The PNG simulation snapshots and halo catalogues                       |            |
|              | - The 3D density fields                                                 |            |
|              | - The HADES data (if available)                                          |            |
|              | - 536 Terabytes                                                         |            |
| San Diego    | - The snapshots 8,000 - 14,999 of the fiducial cosmology                | globus     |
|              | - The snapshots of the standard & fixed LH latin hypercube              |            |
|              | - All halo catalogues                                                   |            |
|              | - All spherical overdensity void catalogues                             |            |
|              | - All power spectra                                                     |            |
|              | - All bispectra                                                         |            |
|              | - All correlation functions                                             |            |
|              | - All pdfs                                                              |            |
|              | - 235 Terabytes                                                         |            |
| Princeton    | - The snapshots of all other simulations                                 | globus     |
|              | - 620 Terabytes                                                         |            |
|              | Non available                                                           |            |
6.1 Globus

The data can be accessed through globus by clicking in the links from the above table. Note that to download the data to your local machine (e.g. laptop) you will need to install the globus connect personal. For further details see here. We now provide some simple instructions to use globus.

The simplest way to transfer data is to use the globus graphical environment. Just type the above names in collection (e.g. Quijote_simulations for the data in San Diego) 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.

In some cases, there are so many files in a given directory, that globus may not be able to list them all and will return an error. If this is the case, it is advisable to use the path line. For instance, if by clicking in Snapshots you get a time out error, you may want to just type in the path line: /Snapshots/ or ~/Snapshots/. This may show you the different content of the data and allow you to navigate it. You can also go to a given directory directly from there. E.g. to access the first realization of the fiducial cosmology, type in path: /Snapshots/fiducial/0/ or ~/Snapshots/fiducial/0/.

In some cases, the above option may not be desirable. For instance, imagine that you want to download all linear matter power spectra of the high-resolution latin-hypercube simulations. One of such files (realization 45) is located in /Snapshots/latin_hypercube HR/45/ICs/Pk_mm_z=0.000.txt, while the file for the realization 89 is located in /Snapshots/latin_hypercube HR/89/ICs/Pk_mm_z=0.000.txt.

Thus, to download all those files without involving downloading the full HR latin-hypercube folder, will require that you access each simulation folder, then the ICs folder and then transfer the file individually. For 2,000 files this is unpractical. For these situations, we recommend using the globus Command Line Interface (CLI). The first step is to install the CLI package, if you don’t have it. Next, login into globus by typing in a terminal:

globus login

Then, the following command allow you to determine the associated endpoint of the Quijote simulations:
You should do the same to know the endpoint of the machine where you are transferring the data to. You can then explore the filesystem of the Quijote simulations (or your machine) as:

```
ep1=c42757fe-d570-11e9-98e2-0a63aa6b37da
globus ls $ep1:/Snapshots/latin_hypercube_HR/45/ICs/
```

The above command will list the content in the /Snapshots/latin_hypercube_HR/45/ICs/ directory. A single file can be transferred as:

```
ep1=c42757fe-d570-11e9-98e2-0a63aa6b37da
ep2=ddb59af0-6d04-11e5-ba46-22000b92c6ec
globus transfer $ep1:/Snapshots/latin_hypercube_HR/45/ICs/Pk_mm_z=0.000.txt $ep2:/Quijote_simulations/linear_Pk/45/Pk_mm_z=0.000.txt --label "single file transfer"
```

Where ep2 should be the endpoint of the machine where you are transferring the data. Entire folders can be moved as follows:

```
ep1=c42757fe-d570-11e9-98e2-0a63aa6b37da
ep2=ddb59af0-6d04-11e5-ba46-22000b92c6ec
globus transfer $ep1:/Snapshots/latin_hypercube_HR/45/ICs $ep2:/Quijote_simulations/45/ICs --recursive --label "single folder transfer"
```

Many folders can be moved with a single command as

```
ep1=c42757fe-d570-11e9-98e2-0a63aa6b37da
ep2=ddb59af0-6d04-11e5-ba46-22000b92c6ec
globus transfer $ep1:/Snapshots/fiducial/ $ep2:/Quijote_simulations/fiducial/ --batch --recursive 0 0 --recursive 1 1 --recursive 2 2 --recursive 3 3 --recursive 4 4 --recursive 5 5 --recursive 6 6 --recursive 7 7 --recursive 8 8 --recursive 9 9 --label "CLI 10 folders" < folders.txt
```

where folders.txt is a text file containing

```
--recursive 0 0
--recursive 1 1
--recursive 2 2
--recursive 3 3
--recursive 4 4
--recursive 5 5
--recursive 6 6
--recursive 7 7
--recursive 8 8
--recursive 9 9
```

For more options and details see Command Line Interface (CLI).
6.2 Binder

Binder is a system that allows users to read and manipulate data that is hosted at the Flatiron Institute through either a Jupyter notebook or a unix shell. The user can find some basic documentation here. The links to the binder for the New York and San Diego cluster can be found in the table above. Note that the data in the Princeton cluster cannot be accessed through binder. Our binder environments contains the following packages:

- nbgitpuller
- sphinx-gallery
- pandas
- matplotlib
- astropy
- matplotlib
- scipy
- h5py
- corner
- future
- numba
- unyt
- Pylians
- pyfftw
- CAMELS-library

**Note:** The first time you log into binder it could take a while. This is because the system is downloading and installing all required packages. Clicking show you can see the progress.

**Warning:** Two important things need to be taken into account when using Binder. First, the Binder environment is ephemeral - after a few days of inactivity its contents are deleted, so one has to be vigilant about downloading any analysis results in time. Second, Binder is not designed to carry out long and heavy calculations. In this case we recommend the user to download the data and work with it locally.
Quijote provides several latin-hypercubes that can be classified into two main categories depending on whether they include massive neutrinos:

### 7.1 LH

The simulations in this category only consider massless neutrinos. There are three latin-hypercubes in this category, each containing 2,000 simulations that vary the value of $\Omega_m$, $\Omega_b$, $h$, $n_s$, $\sigma_8$. The limits of the latin-hypercubes are set by:

| Parameter | Limit |
|-----------|-------|
| $\Omega_m$ | $[0.1; 0.5]$ |
| $\Omega_b$ | $[0.03; 0.07]$ |
| $h$ | $[0.5; 0.9]$ |
| $n_s$ | $[0.8; 1.2]$ |
| $\sigma_8$ | $[0.6; 1.0]$ |

The value of the cosmological parameters for each simulation of a latin-hypercube of this category can be found here. Alternatively, inside each snapshot folder, there is a file called `Cosmo_params.dat` that contains the value of the cosmological parameters of that simulation. Each simulation of the latin-hypercube has a different value of the initial random seed. The value of the initial random seed of each simulation is written in the file `ICs/2LPT.param` inside each simulation folder.

The differences between the three latin-hypercubes are these:

- **standard**: This latin-hypercube contains 2,000 standard simulations with $512^3$ particles each. The snapshots, halo catalogues...etc of this latin-hypercube are located in a folder called `latin_hypercube`. The folder names are `X`, where `X` goes from 0 to 1999.

- **fixed**: This latin-hypercube contains 2,000 fixed simulations with $512^3$ particles each. The snapshots, halo catalogues...etc of this latin-hypercube are located in a folder called `latin_hypercube`. The folder names are `NCV_X` where `X` goes from 0 to 1999.

- **high-resolution**: This latin-hypercube contains 2,000 standard simulations with $1024^3$ particles each. The snapshots, halo catalogues...etc of this latin-hypercube are located in a folder called `latin_hypercube_HR`. The folder names are `X`, where `X` goes from 0 to 1999.

**Note:** The simulations in the standard and high-resolution latin-hypercubes share the same initial random seed. E.g. the simulation 723 of the standard latin-hypercube has the same initial random seed as the simulation 723 of the high-resolution latin-hypercube. The only difference is the maximum $k$ sampled in each.
7.2 nwLH

The simulations in this category include massive neutrinos. There is one single latin-hypercube in this category, and it contains 2,000 simulations that vary the value of $\Omega_m$, $\Omega_b$, $h$, $n_s$, $\sigma_8$, $M_\nu$, and $w$. The limits of this latin-hypercube are set by

$$\begin{align*}
\Omega_m &\in [0.1; 0.5] \\
\Omega_b &\in [0.03; 0.07] \\
h &\in [0.5; 0.9] \\
n_s &\in [0.8; 1.2] \\
\sigma_8 &\in [0.6; 1.0] \\
M_\nu &\in [0.01; 1.0] \text{ eV} \\
w &\in [−1.3; −0.7]
\end{align*}$$

The value of the cosmological parameters of each simulation of the latin-hypercube can be found here. Alternatively, inside each snapshot folder, there is a file called Cosmo_params.dat that contains the value of the cosmological parameters of that simulation. Each simulation of the latin-hypercube has a different value of the initial random seed. The value of the initial random seed of each simulation is written in the file ICs/NGenIC.param inside each simulation folder.

**Note:** Note that the initial conditions of these simulations have been generated using the Zel’dovich approximation, while the initial conditions of latin-hypercubes that do not include neutrinos were generated using 2LPT.

The snapshots, halo catalogues...etc of this latin-hypercube are located in a folder called latin_hypercube_nwLH. The folder names are X, where X goes from 0 to 1999.
Quijote contains 4,000 N-body simulations with primordial non-Gaussianities: **Quijote-PNG**. All these simulations contain $512^3$ dark matter particles in a periodic volume of $(1 \ h^{-1}\text{Gpc})^3$ and share the same cosmology as the fiducial model: $\Omega_m = 0.3175$, $\Omega_b = 0.049$, $h = 0.6711$, $n_s = 0.9624$, $\sigma_8 = 0.834$, $w = -1$, $M_{\nu} = 0.0$ eV. These are standard N-body simulations run with initial conditions generated in a particular way.

The simulations in Quijote-PNG can be classified into four different sets: 1) local, 2) equilateral, 3) orthogonal CMB, and 4) orthogonal LSS (see Bispectrum shapes). Each set contains 1,000 simulations: 500 with $f_{\text{NL}} = +100$ and 500 with $f_{\text{NL}} = -100$. Quijote-PNG is thus organized into eight different folders, depending on the non-Gaussianity shape and the value of $f_{\text{NL}}$:

- **LC_p**: contains data from 500 simulations with local type and $f_{\text{NL}} = +100$
- **LC_m**: contains data from 500 simulations with local type and $f_{\text{NL}} = -100$
- **EQ_p**: contains data from 500 simulations with equilateral type and $f_{\text{NL}} = +100$
- **EQ_m**: contains data from 500 simulations with equilateral type and $f_{\text{NL}} = -100$
- **OR_CMB_p**: contains data from 500 simulations with orthogonal CMB type and $f_{\text{NL}} = +100$
- **OR_CMB_m**: contains data from 500 simulations with orthogonal CMB type and $f_{\text{NL}} = -100$
- **OR_LSS_p**: contains data from 500 simulations with orthogonal LSS type and $f_{\text{NL}} = +100$
- **OR_LSS_m**: contains data from 500 simulations with orthogonal LSS type and $f_{\text{NL}} = -100$

Each of the above folders contains 500 sub-folders, each of them hosting the result of a different simulation. For instance, the folder **EQ_p/72/** contains the results of the 72th simulation run with $f_{\text{NL}} = +100$ for the equilateral shape. Depending on the location, these folder will contain the snapshots, halo catalogues, or other data products.

### 8.1 Bispectrum shapes

In Quijote-PNG we only consider models that have a primordial bispectrum, defined as

$$\langle \Phi(k_1)\Phi(k_2)\Phi(k_3) \rangle = (2\pi)^3\delta^{(3)}(k_1 + k_2 + k_3)B_\Phi(k_1, k_2, k_3),$$

where $\Phi(k)$ is the primordial potential. We consider four different shapes for the primordial bispectrum:

1) **Local**. The local shape can be characterized by

$$B^{\text{local}}_\Phi(k_1, k_2, k_3) = 2f_{\text{NL}}^\text{local}P_\Phi(k_1)P_\Phi(k_2) + \text{2 perm.}$$

2) **Equilateral**. The equilateral shape is described by

$$B^{\text{equil}}_\Phi(k_1, k_2, k_3) = 6f_{\text{NL}}^\text{equil}\left[-P_\Phi(k_1)P_\Phi(k_2) + \text{2 perm.}ight.$$

$$-2(P_\Phi(k_1)P_\Phi(k_2)P_\Phi(k_3))^\frac{2}{3} + P_\Phi(k_1)^\frac{2}{3}P_\Phi(k_2)^\frac{2}{3}P_\Phi(k_3)^\frac{2}{3} + \text{5 perm.}\right]$$
3) **Orthogonal CMB.** The orthogonal CMB template is given by

\[
B_{\Phi}^{\text{ortho-CMB}}(k_1, k_2, k_3) = 6f_{\text{NL}}^{\text{ortho-CMB}} \left[ -3P_\Phi(k_1)P_\Phi(k_2) + 2 \text{ perm.} - 8 \left( P_\Phi(k_1)P_\Phi(k_2)P_\Phi(k_3) \right)^{\frac{4}{3}} + 3P_\Phi(k_1)^{\frac{4}{3}}P_\Phi(k_2)^{\frac{4}{3}}P_\Phi(k_3) + 5 \text{ perm.} \right]
\]

4) **Orthogonal LSS.** The orthogonal LSS template is given by

\[
B_{\Phi}^{\text{ortho-LSS}}(k_1, k_2, k_3) = 6f_{\text{NL}}^{\text{ortho-CMB}} \left[ P_\Phi(k_1)P_\Phi(k_2)P_\Phi(k_3) \right]^{\frac{4}{3}} \left[ - \left( 1 + \frac{9p}{27} \right) \frac{k_1^2}{k_1k_2} + 2 \text{ perm.} + \left( 1 + \frac{15p}{27} \right) \frac{k_1}{k_3} + 5 \text{ perm.} - \left( 2 + \frac{60p}{27} \right) \right] + \frac{p}{27} \frac{k_1^4}{k_1^2k_2^2} + 2 \text{ perm.} - \frac{20p}{27} \frac{k_1k_2}{k_3^2} + 2 \text{ perm.} - \frac{6p}{27} \frac{k_1^2}{k_2k_3^2} + 5 \text{ perm.} + \frac{15p}{27} \frac{k_1^2}{k_2k_3^2} + 5 \text{ perm.} \right]
\]

### 8.2 Initial conditions

The initial conditions of the Quijote-PNG simulations have been generated using a modified version of the code described in Scoccimarro et al. 2012. Our modified version of the code is publicly available here.

The initial conditions of a given simulation can be found in a folder called ICs, that contains:

- ics.X. These are the initial conditions that contain the particle positions, velocities, and IDs. These are Gadget format-II snapshots and can be read as described in Snapshots. X can go from 0 to 127.
- 2LPT.params. This is the parameter file used to generate the initial conditions.
- logIC. The output of the initial conditions generator code.

The value of initial random seed for the simulation \(i\) is \(10 \times i + 5\) (this can be found in the 2LPT.params file) independently of the shape and \(f_{\text{NL}}\) value. For instance, the value of the initial random seed for OR_CMB_p/100 and OR_CMB_m/100 is 1005. This choice enables the calculation of partial derivatives, needed for Fisher matrix calculations.

For the details about the linear matter power spectrum used for these simulations see Linear power spectra.

### 8.3 Snapshots

We keep snapshots at redshifts 0, 0.5, 1, 2, and 3. The snapshots are saved as HDF5 files, and they can be read in the standard way (see Snapshots for details on this).
8.4 Halo catalogues

We store Friends-of-Friends (FoF) halo catalogues for each snapshot of each simulation in Quijote-PNG. We refer the user to Halo catalogues for details on how to read these files.

8.5 Team

Quijote-PNG was developed in 2022 by:

- William Coulton (CCA, USA)
- Gabriel Jung (Padova, Italy)
- Francisco Villaescusa-Navarro (CCA/Princeton, USA)
- Dionysios Karagiannis (Cape Town, South Africa)
- Drew Jamieson (MPA, Germany)
- Michele Liguori (Padova, Italy)
- Marco Baldi (Bologna, Italy)
- Licia Verde (Barcelona, Spain)
- Benjamin Wandelt (IAP, France)
The HADES simulations were the precursor of Quijote, and they contained around 1,000 N-body and hydrodynamic simulations run with different neutrino masses. HADES was designed to study neutrino effects of cosmological observables, while Quijote philosophy is more generic and not just focused on neutrinos.

Quijote now contains all HADES data. The data is however stored on tape, but can be retrieved back and placed in globus, url, and binder. If you need this please reach out to villaescusa.francisco@gmail.com
SNAPSHOTS

The snapshots are stored in either Gadget-II format or HDF5. They can be read using the readgadget.py and readsnap.py scripts. If you have Pylions installed you already have them.

The snapshots only contain 4 blocks:

- Header: This block contains general information about the snapshot such as redshift, number of particles, box size, particle masses... etc.
- Positions: This block contains the positions of all particles. Stored as 32-floats
- Velocities: This block contains the velocities of all particles. Stored as 32-floats
- IDs: This block contains the IDs of all particles. Stored as 32-integers. (This block may be removed in the future to reduce the size of the snapshots)

An example on how to read a snapshot is this:

```python
import numpy as np
import readgadget

# input files
snapshot = '/home/fvillaescusa/Quijote/Snapshots/h_p/snapdir_002/snap_002'
ptype = [1] # [1](CDM), [2](neutrinos) or [1,2](CDM+neutrinos)

# read header
header = readgadget.header(snapshot)
BoxSize = header.boxsize/1e3 # Mpc/h
Nall = header.nall  # Total number of particles
Masses = header.massarr*1e10  # Masses of the particles in Msun/h
Omega_m = header.omega_m  # value of Omega_m
Omega_l = header.omega_l  # value of Omega_l
h = header.hubble  # value of h
redshift = header.redshift  # redshift of the snapshot
Hubble = 100.0*np.sqrt(Omega_m*(1.0+redshift)**3+Omega_l)  # Value of H(z) in km/s/(Mpc/h)

# read positions, velocities and IDs of the particles
pos = readgadget.read_block(snapshot, "POS ", ptype)/1e3 # positions in Mpc/h
vel = readgadget.read_block(snapshot, "VEL ", ptype)  # peculiar velocities in km/s
ids = readgadget.read_block(snapshot, "ID ", ptype)-1  # IDs starting from 0
```

In the simulations with massive neutrinos it is possible to read the positions, velocities and IDs of the neutrino particles. Notice that the field should contain exactly 4 characters, that can be blank: "POS ", "VEL ", "ID ". The number in the name of the snapshot represents its redshift:

- 000 ——> z=3
• 001 ——> z=2
• 002 ——> z=1
• 003 ——> z=0.5
• 004 ——> z=0
The halo catalogues can be read through the readfof.py script. If you have Pylians installed you already have it. An example on how to read a halo catalogue is this:

```python
import readfof

# input files
snapdir = '/home/fvillaescusa/Quijote/Halos/s8_p/145/'  # folder hosting the catalogue
snapnum = 4  # redshift 0

# determine the redshift of the catalogue
z_dict = {4: 0.0, 3: 0.5, 2: 1.0, 1: 2.0, 0: 3.0}
redshift = z_dict[snapnum]

# read the halo catalogue
FoF = readfof.FoF_catalog(snapdir, snapnum, long_ids=False, swap=False, SFR=False, read_IDs=False)

# get the properties of the halos
pos_h = FoF.GroupPos/1e3  # Halo positions in Mpc/h
mass = FoF.GroupMass*1e10  # Halo masses in Msun/h
vel_h = FoF.GroupVel*(1.0+redshift)  # Halo peculiar velocities in km/s
Npart = FoF.GroupLen  # Number of CDM particles in the halo
```

The number in the name of the halo catalogue represents its redshift:

- 000 ——> z=3
- 001 ——> z=2
- 002 ——> z=1
- 003 ——> z=0.5
- 004 ——> z=0
ChangHoon Hahn has created a set of tens of thousands of galaxy catalogues from the Quijote simulations called the Molino catalogues.

You can find all the information and how to access this data here.
The void catalogues are stored as hdf5 files. They contain the following blocks:

- pos: the positions of the void centers in Mpc/h
- radius: the sizes of the voids in Mpc/h
- VSF: the void size function
- VSF_Rbins: the radii bins of the void size function
- parameters: the values of the void finder parameters used to generate the void catalogue

In python, the files can be read as

```python
import h5py

f = h5py.File('/home/fvillaescusa/Quijote/Voids/fiducial/0/void_catalogue_m_z=0.hdf5', 'r')

pos = f['pos'][:]  # void center positions in Mpc/h
radius = f['radius'][:]  # void radii in Mpc/h
VSF = f['VSF'][:]  # VSF (#voids/dR/Volume)
VSF_Rbins = f['VSF_Rbins'][:]  # VSF radii in Mpc/h
parameters = f['parameters'][:]  # parameters used to run the void finder

f.close()
```
14.1 Linear power spectra

The different folders contain both the CAMB parameter files and the matter power spectrum at z=0. In some cases transfer functions and power spectra for neutrinos, CDM, baryons, and CDM+baryons are also present. The format of the power spectrum files is

- \( k \mid P(k) \)

where the units of \( k \) and \( P(k) \) are comoving \( h/\text{Mpc} \) and \( (\text{Mpc}/h)^3 \), respectively. For the fiducial, \( \Omega_m \), \( \Omega_b \), \( h \), \( n_s \), \( s_8 \), \( \Lambda \), \( \Omega_\text{CDM} \), \( \Omega_\text{baryons} \), \( \Omega_\text{CDM+baryons} \), \( \Omega_\text{mass} \) the name of the matter power spectrum files at \( z=0 \) is CAMB_matterpow_0.dat. For \( M_{\text{nuc}_m} \), \( M_{\text{nuc}_mm} \), and \( M_{\text{nuc}_mmm} \) the files are called instead XeV_Pm_rescaled_z0.0000.txt, where \( X = 0.1(M_{\text{nuc}_m}), 0.2(M_{\text{nuc}_mm}) \) and \( 0.4(M_{\text{nuc}_mmm}) \). For the latin hypercube simulations, the files are named Pk_mm_z=0.000.txt.

Note that the matter power spectra at \( z = 0 \) are not normalized (this is because the normalization is performed in the code that generates the initial conditions). The normalization factor is stored in the file Normfac.txt. One example on how to obtain the correct normalized matter power spectrum for a given cosmology is this:

```python
import numpy as np

f_Pk = '/home/fvillaescusa/Quijote/Linear_Pk/ns_p/CAMB_TABLES/CAMB_matterpow_0.dat'
f_norm = '/home/fvillaescusa/Quijote/Linear_Pk/ns_p/Normfac.txt'

k, Pk = np.loadtxt(f_Pk, unpack=True)
Normfac = np.loadtxt(f_norm)

Pk_norm = Pk*Normfac
```

**Caution:** For the primordial non-Gaussian simulations, \( \Lambda \), \( \Omega_\text{mass}, \Omega_\text{CDM}, \Omega_\text{CDM+baryons} \), the linear power spectra files contain the Gaussian linear matter power spectrum from CAMB. The code that generates the initial conditions will take this Gaussian power spectrum and generate the non-Gaussian initial conditions.
### 14.2 Non-linear power spectra

The format of the power spectra are:

- $k \mid P(k)$ for power spectra in real-space
- $k \mid P_0(k) \mid P_2(k) \mid P_4(k)$ for power spectra in redshift-space

where $P_0(k)$, $P_2(k)$ and $P_4(k)$ are the monopole, quadrupole and hexadecapole, respectively. The units of $k$ are $h/\text{Mpc}$, while for the power spectra are $(\text{Mpc}/h)^3$.

In redshift-space there are three different files for each realization/redshift. These have been computed by placing the redshift-space distortions along the three different axes.

In python, the files can be read as:

```python
import numpy as np

k, Pk = np.loadtxt('$HOME/fvillaescusa/Quijote/Pk/matter/fiducial/3/Pk_m_z=0.txt', unpack=True)
k, Pk0, Pk2, Pk4 = np.loadtxt('$HOME/fvillaescusa/Quijote/Pk/matter/fiducial/3/Pk_m_RS1_z=0.txt', unpack=True)
```

### 14.3 Marked power spectra

The files whose name starts with:

- Mk_ contain marked power spectra $M(k)$ evaluated at wavenumber $k$
- Xk_ contain the cross spectra between marked and standard density field $X(k)$ evaluated at wavenumber $k$

The unit of $k$ is $h/\text{Mpc}$, while the one of $M(k)$ and $X(k)$ is $(\text{Mpc}/h)^3$.

Files with measurements performed in the fiducial cosmology have name:

- Mk_fiducial0-4999_....hdf5
- Xk_fiducial0-4999_....hdf5

where the first numbers (in the above case 0-4999) indicate the realizations saved in the file, and the dots specify the marked model considered.

The remaining files contain measurements performed in the other cosmologies and from 500 realization per cosmology. Their name is:

- Mk_fTH_....hdf5
- Xk_fTH_....hdf5

Also in this case the dots specify the marked model considered.

In python, the files can be read as:

```python
import numpy as np

f = h5py.File(FILENAME, 'r')
k = f['k'][:]
# Fiducial cosmology
Mk = f['i'][:]
```
# Massive neutrino cosmologies
\[ M_k = f[\text{cosmo/i\_suffix}][::] \]

# Other cosmologies
\[ M_k = f[\text{cosmo/i}][::] \]

where i is the number of the realization, cosmo is the wanted cosmology and suffix can be

- ‘m’ for the total matter field
- ‘cb’ for the cold dark matter plus baryons

In order to see the name of each cosmology type

```python
print(list(f.keys()))
```
The format of the individual bispectra files are:

- \( k_1/k_f \mid k_2/k_f \mid k_3/k_f \mid P_0(k_1) \mid P_0(k_2) \mid P_0(k_3) \mid B_0(k_1, k_2, k_3) \mid B_{SN}(k_1, k_2, k_3) \mid \text{counts} \)

where \( k_1, k_2, k_3 \) specify the length of the triangle sides, \( P_0(k) \) is the power spectrum monopole, \( B_0(k_1, k_2, k_3) \) is the bispectrum monopole, \( Q(k_1, k_2, k_3) \) is the reduced bispectrum, \( B_{SN} \) is the bispectrum shot noise correction, and \( \text{counts} \) is the number of triangles in the bin. \( B_0 \) is already shot-noise corrected. The header specifies \( k_f \), the fundamental mode, and \( N_{halo} \), the number of halos.

The individual bispectra files can be read in python as follows,

```python
import numpy as np

k1, k2, k3, p0k1, p0k2, p0k3, b123, q123, b_sn, cnts = np.loadtxt(FILENAME, skiprows=1, unpack=True, usecols=range(10))

# read header to get Nhalo
hdr = open(FILENAME).readline().rstrip()
Nhalo = int(hdr.split('Nhalo=')[-1])
```

Alternatively, sets of bispectra files for a specific redshift and cosmology can easily be accessed

```python
import h5py

fbk = h5py.File(FILENAME, 'r')
k1 = fbk['k1'][...]
k2 = fbk['k2'][...]
k3 = fbk['k3'][...]
p0k1 = fbk['p0k1'][...]
p0k2 = fbk['p0k2'][...]
p0k3 = fbk['p0k3'][...]
b123 = fbk['b123'][...]
q123 = fbk['q123'][...]
b_sn = fbk['b_sn'][...]
cnts = fbk['counts'][...]
# triangle counts
Nhalos = fbk['Nhalos'][...]
# number of halos
files = fbk['files'][...]
# names of individual files.
```
The format of the correlation functions are:

- \( R \mid x_i(R) \) for correlation functions in real-space
- \( R \mid x_i^0(R) \mid x_i^2(R) \mid x_i^4(R) \) for correlation functions in redshift-space

where \( x_i^0(R) \), \( x_i^2(R) \) and \( x_i^4(R) \) are the monopole, quadrupole and hexadecapole, respectively. The units of \( R \) are Mpc/h, while the different \( x_i \) are dimensionless.

In redshift-space there are three different files for each realization/redshift. These have been computed by placing the redshift-space distortions along the three different axes.

In python, the files can be read as

```python
import numpy as np

R, xi = np.loadtxt('/home/fvillaescusa/QUIJOTE/CF/matter/fiducial/0/CF_m_1024_z=0.txt', unpack=True)
R, xi0, xi2, xi4 = np.loadtxt('/home/fvillaescusa/QUIJOTE/CF/matter/fiducial/0/CF_m_RS0_1024_z=0.txt', unpack=True)
```
The format of the PDF files is:

- delta | pdf

where delta is the density contrast (\(\rho/\langle \rho \rangle - 1\)).

In python, the files can be read as

```python
import numpy as np

delta, pdf = np.loadtxt('/home/fvillaescusa/Quijote/PDF/matter/latin_hypercube/0/PDF_m_5..0_z=0.txt', unpack=True)
```
18.1 3D fields

The 3D density fields are located in the New York cluster (see Data access) under the 3D_cubes folder. There are different folders for the different cosmologies. Inside each cosmology folder there are the folder containing the data for the different realizations. Inside each of those folders the 3D density fields can be found with names as df_m_X_Y_z=Z.npy, where X can be 64, 128, 256, or 512, and it represents the grid size of the cube. Y represents the mass assignment scheme used to construct the density field, and can be something like CIC (cloud-in-cell) or PCS (piece-wise spline). Z represents the redshift of the density field. For instance, df_m_256_CIC_z=0.npy contains the 3D density field on a grid with $256^3$ voxels constructed using the CIC mass-assignment scheme at $z = 0$.

Note: These fields are constructed in real-space. Please reach us if you need them in redshift-space.

The files can be read simply as

```python
import numpy as np

df = np.load('/home/fvillaescusa/Quijote/3D_cubes/Om_p/df_m_128_PCS_z=0.npy')
```

18.2 2D fields

2D fields (say images) can be constructed from the above 3D fields by taking a slice and projected it into 2D. For instance:

```python
import numpy as np

# read the 3D density field
df_3D = np.load('/home/fvillaescusa/Quijote/3D_cubes/Om_p/df_m_128_PCS_z=0.npy')

# take a slice of 4 voxels width, i.e. 1000/128*4 = 31.25 Mpc/h
# along z-direction and project into 2D by computing the mean value
df_2D = np.mean(df_3D[:,:,0:4], axis=2)
```
TEAM

- Francisco Villaescusa-Navarro (Flatiron/Princeton)
- ChangHoon Hahn (Berkeley)
- Elena Massara (Flatiron/Waterloo)
- Arka Banerjee (Stanford)
- Ana Maria Delgado (CUNY/Flatiron)
- Doogesh K. Ramanah (Sorbonne, Paris)
- Tom Charnock (Sorbonne, Paris)
- Elena Giusarma (Flatiron/MTech)
- Yin Li (Berkeley/IPMU)
- Erwan Allys (ENS, Paris)
- Antoine Brochard (ENS, Paris)
- Cora Uhlemann (Cambridge)
- Chi-Ting Chiang (BNL)
- Siyu He (Flatiron)
- Alice Pisani (Princeton)
- Andrej Obuljen (Waterloo)
- Yu Feng (Berkeley)
- Emanuele Castorina (Berkeley)
- Gabriella Contardo (Flatiron)
- Christina D. Kreisch (Princeton)
- Andrina Nicola (Princeton)
- Justin Alsing (Oskar Klein/Flatiron)
- William Coulton (Flatiron)
- Drew Jamieson (MPA)
- Gabriel Jung (Padova, Italy)
- Dionysios Karagiannis (Cape Town, South Africa)
- Michele Liguori (Padova, Italy)
• Marco Baldi (Bologna)
• Roman Scoccimarro (NYU)
• Licia Verde (Barcelona)
• Matteo Viel (SISSA)
• Shirley Ho (Flatiron/Princeton)
• Stephane Mallat (ENS/College de France)
• Ben Wandelt (IAP, Paris)
• David Spergel (Flatiron/Princeton)
The below gif image has been created by linearly interpolating two images: 1) an image of the large-scale structure from one Quijote simulation, and 2) an image created by using the DeepDream software on top of 1). The goal of the animation is to emphasize and highlight regions of the cosmic web in a novel way.

The images below show Don Quijote riding his horse with the sky showing the large-scale structure of the Universe from its traditional version (top-right) to its machine learning version (bottom-left).
If you have used data from the Quijote simulations you may consider citing the Quijote paper.

```latex
@ARTICLE{Quijote_sims,  
  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,  
  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\}
}
```

If you use data from Molino, consider citing the Molino paper

```latex
@ARTICLE{Molino,  
  author = {{Hahn}, ChangHoon and {Villaescusa-Navarro}, Francisco},  
  title = "Constraining $M_{\nu}$ with the bispectrum. Part II. The information content of the galaxy bispectrum monopole",  
  journal = \jcap,  
  keywords = {cosmological parameters from LSS, cosmological simulations, neutrino masses from cosmology, redshift surveys, Astrophysics - Cosmology and Nongalactic Astrophysics},
}
```
If you use data from Gigantes, consider citing the Gigantes paper.

```latex
@ARTICLE{Gigantes,
    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 dataset: precision cosmology from voids in the machine learning era",
    journal = {arXiv e-prints},
    keywords = {Astrophysics - Cosmology and Nongalactic Astrophysics, Astrophysics - Instrumentation and Methods for Astrophysics},
    year = 2021,
    month = jul,
    eid = {arXiv:2107.02304},
    pages = {arXiv:2107.02304},
    archivePrefix = {arXiv},
    eprint = {2107.02304},
    primaryClass = {astro-ph.CO},
    adsurl = {https://ui.adsabs.harvard.edu/abs/2021arXiv210702304K},
    adsnote = {Provided by the SAO/NASA Astrophysics Data System}
}()
```

If you use data from Quijote-PNG, consider citing the Quijote-PNG paper.

```latex
@ARTICLE{Quijote-PNG,
    author = {{Coulton}, William R and {Villaescusa-Navarro}, Francisco and {Jamieson}, Drew and {Baldi}, Marco and {Jung}, Gabriel and {Karagiannis}, Dionysios and {Liguori}, Michele and {Verde}, Licia and {Wandelt}, Benjamin D.},
    title = "Quijote-PNG: Simulations of primordial non-Gaussianity and the information content of the matter field power spectrum and bispectrum",
    journal = {arXiv e-prints},
    keywords = {Astrophysics - Cosmology and Nongalactic Astrophysics},
    year = 2022,
    month = jun,
    eid = {arXiv:2206.01619},
    pages = {arXiv:2206.01619},
    archivePrefix = {arXiv},
    eprint = {2206.01619},
}()
```
primaryClass = {astro-ph.CO},
adsurl = {https://ui.adsabs.harvard.edu/abs/2022arXiv220601619C},
adsnote = {Provided by the SAO/NASA Astrophysics Data System}
MIT License
Copyright (c) 2019 Francisco Villaescusa-Navarro

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the “Software”), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED “AS IS”, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.
For problems, questions and general help you can reach us at villaescusa.francisco@gmail.com