Higher Order Statistics of the Large Scale Structure

Author: Paweł Drozda

Advisors: dr. hab. Wojciech A. Hellwing, prof. CFT PAN

dr. hab. Maciej A. Bilicki, prof. CFT PAN

Centrum Fizyki Teoretycznej Polskiej Akademii Nauk

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics

June 2025

Dedicated to my parents
Małgorzata and Piotr

Abstract

On the surface, our current understanding of the Universe appears to be accurate. General Relativity (GR) appropriately describes gravity, and the standard ΛCDM model provides reasonable explanations for most cosmological observations. However, the assumptions underlying ΛCDM reveal its two greatest mysteries: Λ (the cosmological constant) is a dark energy candidate to reproduce the observed accelerated expansion of the Universe, while CDM (cold dark matter) interacts only via gravity. Although adding these components yields agreement with data, their physical origin remains unknown.

In this context, testing gravity at cosmological scales is highly motivated - especially since gravitational interactions have not been robustly tested at scales larger than the Solar System.

A variety of extended gravity scenarios can drive cosmic acceleration without Λ. By comparing simulations in standard GR to those in extended models, we can begin to constrain the gravity. In this thesis, we consider two popular scenarios generalizing Einstein-Hilber action: the Hu–Sawicki f(R) model and the normal branch of the Dvali–Gabadadze–Porrati (nDGP) model. Although nDGP alone does not self-accelerate, studying both families provides complementary tests of two distinct extended gravity models.

Because gravity is nonlinear, it imprints non-Gaussian features into an evolving density field. Under the assumption of an initially Gaussian field, deviations from Gaussianity observed later can be interpreted as gravity signatures. These can be quantified via the central moments of number density counts. Since different gravity models leave distinct imprints, higher-order statistics, specifically, averaged correlation functions and cumulants based on central moments offer powerful discrimination. Working in a controlled suite of cosmological simulations, we establish a testbed for future observations and identify the regimes where the gravity scenarios can be distinguished (extended gravity signals, or EG signals). First, in a light-cone geometry that mimics real surveys, we measure angular statistics. We identify the optimal redshift range for detecting EG signals as $0.15 < z < 0.3$. We find that extended-gravity deviations from GR reach up to 20%. Analyzing dark-matter and mock galaxy catalogs separately, we detect these signals at $2-4\sigma$ significance, reaching even $\sim 3\sigma$ for sparse galaxy samples ($15\ deg^{-2}$), demonstrating the feasibility of observational tests with robust catalogs.

Encouraged by these results, we then study fully three-dimensional clustering in both real and redshift space. Focusing on skewness, we show that the Fingers-of-God (FoG) effect suppresses small-scale skewness in redshift space and EG signals become reduced compared to real-space measures. The z-space signal still reaches $\sim 4\%$ for galaxies. Notably, galaxy catalogs exhibit stronger deviations than halo catalogs, underscoring the potential of skewness as an observational probe.

Next, motivated by the richness of information contained in redshift space, we introduce ellipsoidal averaged correlation functions: an analogue of the anisotropic two-point func-

tion built from central moments. Employing dedicated simulations we find that developed ellipsoidal functions reveal features inaccessible to spherical counts: the optimal shape of ellipsoid at given volume corresponds with strength of FoG and Kaiser effects. We further show that, for skewness, redshift-space measurements outperform real-space ones in distinguishing models with different structure-growth histories.
Finally, we place our findings in the context of current and upcoming cosmological data, and discuss prospects for employing higher-order study in new observations.

Streszczenie

Nasze obecne zrozumienie Wszechświata z pozoru wydaje się być trafne. Teoria Względności (GR) opisuje grawitację wyjątkowo adekwatnie, podczas gdy standardowy model kosmologiczny ΛCDM rozsądnie wyjaśnia większość kosmologicznych obserwacji. Jednakże założenia modelu ΛCDM ujawniają dwie największe zagadki: Λ (stała kosmologiczna) jest kandydatką ciemnej energii mającą odtwarzać obserwowaną przyspieszoną ekspansję Wszechświata, podczas gdy CDM (zimna ciemna materia) oddziałuje wyłącznie grawitacyjnie. Uwzględnienie obydwu składników w opisie Wszechświata powoduje dużą zgodność z obserwacjami, jednak ich fizyczne pochodzenie pozostaje nieznane.

W tym kontekście szczególnie umotywowane jest przetestowanie i nałożenie ograniczeń na modele grawitacji oddziaływującej na skalach kosmologicznych, szczególnie biorąc pod uwagę, że nie ma obecnie solidnych testów siły grawitacji na skalach większych niż Układ Słoneczny.

Istnieją różne rozszerzone scenariusze grawitacji, które mogą powodować przyspieszoną ekspansję Wszechświata bez potrzeby wprowadzania Λ. Pierwsze kroki w kierunku ograniczania możliwych teorii grawitacji można podjąć, porównując wyniki symulacji uzyskane w standardowej teorii Einsteina (GR) z wynikami uzyskanymi z rozszerzonych modeli grawitacji. W tej pracy skupiamy się na dwóch popularnych scenariuszach uogólniających całkę działania Einsteina-Hilberta: na modelu $f(R)$ w formie Hu-Sawickiego oraz na normalnym wariancie modelu Dvali–Gabadadze–Porrati (nDGP). Choć sam model nDGP nie gwarantuje przyspieszonej ekspansji bez Λ, analiza obu rodzin rozszerzonych teorii grawitacji pozwala na cenne porównanie dwóch różnych podejść.

Ponieważ grawitacja jest nieliniowa, narzuca ona nie-Gaussowskie cechy w ewoluującym polu gęstości. Zakładając Gaussowskie warunki początkowe, obserwowane później odstępstwa od Gaussowskości mogą być interpretowane jako sygnatura grawitacji. Sygnatury te można ilościowo opisać za pomocą momentów centralnych rozkładu liczby obiektów pola gęstości. Ponieważ różne modele grawitacji pozostawiają odmienne ślady, statystyki wyższego rzędu – w szczególności uśrednione funkcje korelacji oraz kumulanty oparte na momentach centralnych – stają się pomocnym narzędziem do ich rozróżniania. Przeprowadzając badania w kontrolowanym środowisku symulowanej kosmicznej sieci dla różnych scenariuszy, tworzymy pole testowe pod przyszłe obserwacyjne ograniczenia modeli grawitacji i wskazujemy zakresy, w których użyte modele grawitacji są rozróżnialne (sygnały rozszerzonej grawitacji). W pierwszej części identyfikujemy sygnały w statystykach kątowych, korzystając z geometrii stożka świetlnego (lightcone), naśladującej rzeczywiste obserwacje. Określiliśmy optymalny przedział przesunięć ku czerwieni $0.15 < z < 0.30$ do wykrywania sygnałów i znaleźliśmy do 20% odchylenia rozszerzonych modeli grawitacji względem GR. Analizę przeprowadziliśmy oddzielnie dla katalogów ciemnej materii i symulowanych galaktyk. Wykryte sygnały wykazywały istotność na poziomie $2-4\sigma$, nawet w przypadku rzadkich katalogów galaktyk (gęstość rzutowana 15 deg^{-2}) dochodziła

do $\approx 3\sigma$. To stwarza realną szansę na przyszłe testy grawitacji z wykorzystaniem bardziej kompletnych katalogów. Zachęceni obiecującymi wynikami następnie zbadaliśmy w pełni trójwymiarowe uśrednione statystyki w przestrzeni rzeczywistej i w przestrzeni przesunięć ku czerwieni. Skupiając się na skośności, zauważyliśmy, że efekt „Fingers-of-God" tłumi skośność na małych skalach w przestrzeni przesunięć ku czerwieni i jednocześnie osłabia sygnał grawitacyjny. W przypadku galaktyk odchylenia od GR dochodziły do $\sim 4\%$. Sygnały w katalogach galaktyk były silniejsze niż w halach, co jest istotne w kontekście potencjalnych obserwacyjnych testów grawitacji przy użyciu skośności.
Następnie, zmotywowani bogactwem informacji zawartych w przestrzeni przesunięć ku czerwieni, opisaliśmy i zinterpretowaliśmy elipsoidalne uśrednione funkcje korelacyjne: oparty o momenty centralne analog anizotropowej dwu-punktowej funkcji korelacyjnej. Korzystając z dedykowanych symulacji zidentyfikowaliśmy dodatkowe cechy pola gęstości, które są niedostępne przy użyciu standardowych kulistych zliczeń: analiza optymalnego kształtu elipsoid przy stałej objętości ukazała jego zależność od siły efektów „Fingers-of-God" i Kaisera. Pokazaliśmy następnie, że w przypadku skośności pomiary w przestrzeni przesunięć ku czerwieni są lepsze od pomiarów z przestrzeni rzeczywistej w kontekście odróżniania modeli z różnymi scenariuszami wzrostu kosmicznej struktury.
Na koniec omawiamy wszystkie wyniki, odniesienia do aktualnych i przyszłych danych kosmologicznych oraz dyskutujemy o perspektywach mierzenia na nich statystyk przedstawionych w tej pracy.

Acknowledgements

The PhD studies have been a highly valuable time in my life, made possible by the support of many people.

Firstly, I would like to express my gratitude to my supervisors: Wojciech Hellwing and Maciej Bilicki. I am grateful for their invaluable advices and strong support of my research. They always encouraged me to explore every interesting side aspects of the main research, what not only led to unexpected insights, but also increased my confidence in conducting impactful research. I am especially thankful for their help with improving my scientific writing as well as giving talks. Also, I would like to thank them for inviting me to serve as the member of local organizing committee in conferences. This not only improved my soft skills, but also allowed to to apply my long-time hobby (digital art) in designing conference posters and badges. I am grateful to Prof. Enrique Gaztañaga for many helpful discussions, both at the beginning of my first clustering-related paper, and during my research visit. Thanks to your expertise and suggestions, I was able to find solutions to many challenging problems. I would like to thank Mariana Jaber for invaluable advices on giving talks and collaboration. Your help and humor is what I really appreciate.

I would like to thank colleagues in CFT: Oliver, Priyanka, Jorge, Krishna, Maria Luiza, Szymon, Julius, Tae Hun, Maria and Owidiusz, as well as participants of 5th Cosmology School. I owe special thanks to Suhani, Feven, Anjitha, Gursharanjit and Rasmus who I was sharing the office with. Your presence and friendly atmosphere you created made this PhD journey far more meaningful.

I am deeply grateful to my family, especially my parents Małgorzata and Piotr, who always supported me. Without them I would not have a chance to firstly develop my passion for astronomy and then to reach this point. I would like to thank my brother Karol, grandpa Wiesław, grandma Maria, aunts and uncles who supported me and with who I could have a great time. I am grateful to my friends: Michał, Asia, Tomek, Mateusz, Kasia, Rafał, Szymon, Magda, Robert, Kasia, Gosia, Czarek, Pratik, Agata, Teresa, Piotrek, Oliwia, Ula and others. I really appreciate staying in touch and meeting on various occasions, even many years after finishing studies together with some of you. I would like to thank Wojtek, Michał and Piotrek for wonderful explorations of wild places. Not every adventure was safe, but each of them created great memories. Also I thank senseis: Radek, Krzysztof, Kacper and participants of karate trainings. These definitely added a positive value into my life.

I am grateful to CFT PAN for opportunity to pursue my PhD studies. I would like to thank director Prof. Krzysztof Pawłowski and entire CFT administration for high achievability and support in navigating formal requirements. I would like to thank Prof. Remigiusz Augusiak for the help with procedures of submitting doctoral thesis.

I acknowledge financial support from National Science Center, Warsaw4PhD doctoral

school, CFT PAN and NAWA.

Declaration

I hereby declare that this thesis represents my work between 2021 and 2025 when I was a doctoral student in Center for Theoretical Physics PAS, under supervision of Prof. Wojciech A. Hellwing and Prof. Maciej Bilicki. Any part of this thesis has not been a subject for other degree application. The thesis includes following works:

- **P. Drozda**, W. A. Hellwing, M. Bilicki
  "Galaxy and halo angular clustering in ΛCDM and Modified Gravity cosmologies"
  DOI: 10.1103/PhysRevD.106.043513

- **P. Drozda**, W. A. Hellwing, M. Bilicki
  "Skewness as a Probe of Gravity: Real and Redshift Space Counts-In-Cells"
  arXiv:2504.12820 (in review in PRD)

- **P. Drozda**, W. A. Hellwing, M. Bilicki
  "Anisotropic Counts-in-Cells in Redshift Space: A New Route to Cosmological Constraints from Galaxy Surveys"
  arXiv:2506.01762 (submitted to PRD)

My contributions to the above mentioned scientific articles consist of:

- I made all the figures, and conducted the numerical analysis,

- together with coauthors (Prof. Wojciech A. Hellwing and Prof. Maciej Bilicki) I participated in scientific analysis, text writing and discussions of findings and conclusions.

Other contributions

During the doctoral studies I contributed in the work:

- J. Cisewski-Kehe, B. T. Fasy, W. A. Hellwing, M. R. Lovell, **P. Drozda**, M. Wu "Differentiating small-scale subhalo distributions in CDM and WDM models using persistent homology"
  DOI: 10.1103/PhysRevD.106.023521

The thesis does not contain nor discuss the aforementioned work.

Contents

List of Figures

Chapter 1

Introduction

Since ancient times, humanity has been trying to explain origin and nature of the Universe. Over the last hundred years cosmology has changed from the debate whether galaxies are parts of Milky Way or separate equivalent objects, to big-data driven science providing extensive studies on many of them. Human ingenuity led to formulation of theories describing objects which are yet physically inaccessible to us. Modern simulations and sky surveys allow to model and put tight constraints on the observables, leading to deepening our knowledge about the Universe.

At largest, cosmologic scales the dominating force is gravity. Its standard and yet most successful description is Einstein's General Relativity (GR) [1]. The GR assuming Universe to be filled just with matter and radiation predicts that its expansion slows down. However, the measurements reveal positive acceleration, suggesting the existence of so-called dark energy (DE). This led to the addition of term Λ to vanilla Einstein's Equations and in consequence modeling the Universe as filled with additional energy component causing the acceleration. The physical nature of dark energy is however unknown and recent observations may prefer even its equation of state to deviate from constant value [2]. On the top of background evolution, there is another enigmatic component: dark matter (DM), clumping into halos known to interact only gravitationally. While the DM has not been detected directly yet, it dominates the evolution of galaxies and in general, large-scale structure (LSS, or cosmic web - CW) - the complex web-shaped network into which the ingredients of Universe are distributed.

The unknown components of the Universe consist of $\approx 95\%$ of its energy budget [3]. Such situation gives hence rise for questioning our knowledge about force shaping the evolution of observable structures at largest scales. Robust General Relativity tests however did not yet cover cosmologic distance ranges of interactions. To remedy for this puzzle, one can employ so-called extended gravity models which generalize Einstein's GR equations. Careful choice of modification can provide models with different gravitational interactions while preserving its form at already tested regimes. To test the gravity at cosmological scales, one then needs to compare Large Structure simulated within GR and extended gravity, with the observations. The powerful tool for quantitative comparison is cluster-

ing statistics measuring how the objects group and describing the cosmic web density field. In practice, due to large complications related with modeling of galaxy catalogs, interpretability and possible biases, this approach firstly requires identifying differences in gravity signatures between simulated GR and extended gravity, before comparing with observations. The clustering and its differences between simulated gravity scenarios is the main subject which this thesis addresses.
Firstly we discuss the General Relativity and taking into account current unknowns and problems in cosmology we focus on extended gravity models.

1.1 Gravity and cosmology

Among four fundamental forces of nature, the one dominating at largest scales is gravity. According to current state of knowledge, it drives the formation and evolution of cosmic web and influences the expansion of Universe. The most successful description of gravity [4] is currently the General Relativity (GR). This is so far the most tested theory, from small distances [5, 6, 7], strong fields [8, 9, 10, 11, 12], to scales of Solar System [13, 14]. Also presently the GR tests are being suggested, e.g. [15]. Due to such strong agreement with reality, GR is usually adopted as a standard theory of gravity.

1.1.1 General Relativity and background evolution of the Universe

The major concept of General Relativity is the relation between spacetime and matter (energy). It is described by Einstein's equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{1.1}$$

where the left hand side describes the curvature of spacetime and is often collected into single Einstein tensor $G_{\mu\nu}$. The $R_{\mu\nu}$ is a Ricci curvature tensor, $g_{\mu\nu}$ stands for metric tensor and $R = g^{\mu\nu}R_{\mu\nu}$ is the Ricci scalar. The right side of Einstein Equations contains energy-momentum tensor $T_{\mu\nu}$ characterizing the distribution and dynamics of matter.
The cosmological constant Λ is a repulsive space-filing energy. First attempts were to tune it in such way to prevent the Universe from collapsing and make it static. It was however found that the Universe is not static, but rather expands and its rate increases in time [16]. The cosmological constant Λ allows for accelerated late-time expansion. Repulsive cosmological constant remains a necessary addition to General Relativity adopted in standard cosmological model (Sec. 1.1.3), however its physical nature is unknown.
The metric tensor $g_{\mu\nu}$ allow to define distances in a chosen spacetime. Given the metric, one can obtain the interval ds^2:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu, \tag{1.2}$$

where we used Einstein notation, i.e. repeating the indices means summing over them [1]. The x^η, where $\eta = \{0, 1, 2, 3\}$ are the time ($\eta = 0$) and spatial coordinates.

In cosmology, widely assumed is the Cosmological Principle. It states that at sufficiently large scales the Universe is homogeneous and isotropic. The isotropy is motivated with observations of Cosmic Microwave Background [17]. According to Copernican Principle, we are neither in a privileged nor a special position in a Universe. That inevitably implicates the homogeneity.
Given that, the global metric of the Universe is usually assumed to be Friedman-Lemaître-Robertson-Walker (FLRW) metric, which in spherical coordinates gives:

$$ds^2 = (c\ dt)^2 - a^2(t)\Big(\frac{dr^2}{1-kr^2} + r^2(d\theta + sin^2(\theta)d\phi^2)\Big), \tag{1.3}$$

where k is the curvature parameter and $a(t)$ is the so-called scale factor. The value of k defines whether the Universe is open ($k = -1$), flat ($k = 0$) or closed ($k = 1$). Scale factor is a quantity characterizing the size of Universe, normalized to unity at current time t_0, i.e. $a(t_0) = 1$. Multiplication of spatial terms by $a(t)$ in 1.3 enforces the distances to grow as the Universe expands and vice versa. It is then convenient to use so-called comoving coordinates which follow the scale factor evolution. The relation between position $\vec{r}$ and comoving coordinates $\vec{x}$ is:

$$\vec{r} = a(t) \cdot \vec{x}. \tag{1.4}$$

Differentiating this relation with respect to time allows to define Hubble parameter $H(t) \equiv \dot{a}/a$ which describes the expansion rate of the Universe. While considering only velocity coming from cosmological expansion, one end with Hubble law: $\vec{v} = H\vec{r}$.
By inserting the $FLRW$ metric 1.3 into Einstein equations 1.1, from (00) and (ii), $i = \{1, 2, 3\}$ components one obtains Friedman equations [18] which describe the background evolution of the Universe:

$$\Big(\frac{\dot{a}}{a}\Big)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \tag{1.5}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\Big(\rho + 3\frac{P}{c^2}\Big) + \frac{\Lambda c^2}{3}, \tag{1.6}$$

where ρ and P are energy density and pressure, respectively. Differentiating the first equation multiplied by a^2 and then inserting second equation gives:

$$3\frac{\dot{a}}{a}\Big(\rho + \frac{P}{c^2}\Big) + \dot{\rho} = 0. \tag{1.7}$$

Using first and second Friedman equations, one can compute so-called deceleration parameter $q \equiv -\ddot{a}a/\dot{a}^2$. When it was formulated, the Universe was believed to decelerate due to the gravity. The measured value estimated at current time $t = t_0$ due to accelerated expansion is negative $q_0 \sim -0.5$ [19].
Using first Friedman Equation one can define density $\rho_{all} = \rho + \Lambda c^2/(8\pi G)$ which includes

also the energy contribution coming from cosmological constant. The critical density ρ_c which is a reference density required to make Universe flat is:

$$\rho_c \equiv \frac{3H^2}{8\pi G}. \tag{1.8}$$

The density ρ_{all} includes all forms of energy in the Universe, of which the assumed are pressureless matter, radiation and cosmological constant. The general form of equation of state for these ingredients is $P = w\rho$, where w depends on the type of component. For pressureless matter $w = 0$, while radiation and cosmological constant have $w = 1/3$ and $w = -1$, respectively.
Using first law of thermodynamics $dE = -pdV$ and inserting $E = \rho a^2$ along with $V \propto a^3$ one gets the relation between density and scale factor:

$$\rho(t) \propto a(t)^{-3(1+w)}. \tag{1.9}$$

It is convenient to express the density as a fraction of critical density. For the component i, the density parameter Ω_i is:

$$\Omega_i \equiv \frac{\rho_i}{\rho_c}. \tag{1.10}$$

Inserting Hubble parameter at the current time $H_0 = H(t_0)$, one obtains density parameter $\Omega_{i,0}$ today. Also the common approach is to define the density associated with curvature k, what leads to $\Omega_k = -kc^2/(a^2H^2)$ while using 1.8 in First Friedman equation [20].
By inserting 1.8 with 1.9 into 1.10 for current moment t_0 and any given time t, one obtains time evolution of density parameter Ω_i:

$$\Omega_i(t) = \Omega_{i,0}a(t)^{-3(1+w)}\Big(\frac{H_0}{H(t)}\Big)^2. \tag{1.11}$$

Writing first Friedman Equation in terms of dimensionless density parameters gives then:

$$H^2 = H_0^2\Big(\Omega_{R,0}\ a^{-4} + \Omega_{M,0}\ a^{-3} + \Omega_{\Lambda,0} + \Omega_k\ a^{-2}\Big), \tag{1.12}$$

where subscripts $i = \{R, M, \Lambda, k\}$ correspond with radiation, matter, cosmological constant and curvature, respectively. This leads to a remarkable consequence: measuring current densities of the Universe along with Hubble parameter allows for estimating its past and future expansion rate. The observations favor flat ($k = 0$) Universe, with $\Omega_{M,0} \approx 0.315$, and $\Omega_{\Lambda,0} \approx 0.685$ [3]. The estimate of Hubble parameter for Cosmic Microwave Background (CMB) is $H_0 = 67.4 \pm 0.5\ km\ s^{-1}Mpc^{-1}$. This is however a subject of wide discussion, since depending on the method of estimation, one gets different value, we mention this so-called Hubble tension more in section 1.1.3.

1.1.2 Redshift and distances

The expression 1.12 is crucial in relating estimates of current cosmological density parameters with Hubble expansion. The scale factor a is however not obtained directly. Instead, from the observations one measures redshift z. The cosmological redshift is caused by change of wavelength of light traveling through expanding space. Observing the light red-shifted by z, one measures the change of its wavelength λ between observation time t_o and emission t_e:

$$1 + z \equiv \frac{\lambda_o}{\lambda_e} = \frac{a(t_o) = 1}{a(t_e)}, \tag{1.13}$$

where λ_e and λ_o is the wavelength at the moment of emission and observation, respectively. This relation can be obtained by comparing the intervals in $FLRW$ metric 1.3 between emission and observation times of adjacent light wave crests for $d\theta = d\psi = 0$.
The cosmological redshift of an object provides then information about the scale factor of Universe at the moment of emission. Since in expanding Universe the scale factor grows monotonically with time, the larger the redshift, the larger is the difference between $a(t_e)$ and scale factor now, hence the more time the light was traveling to the observer. Therefore objects with greater cosmological redshifts are more distant and simultaneously these are observed in larger time delay, i.e. the larger z, the younger the Universe was during the emission of light which now is reaching us. This is so-called lightcone geometry. It is then convenient to express that object is located at redshift z what informs both about the distance and elapsed time.
The relation between redshift and time can be obtained from $dz/dt = \dot{a}\ dz/da$ by combining the Equation 1.12 and 1.13:

$$\frac{dz}{dt} = -\frac{1}{a}H(z) = -(1+z)\ H_0\sqrt{\Omega_{R,0}(1+z)^4 + \Omega_{M,0}(1+z)^3 + \Omega_{\Lambda,0} + \Omega_k(1+z)^{-2}} \tag{1.14}$$

By inverting and integrating, one can obtain so-called lookback time describing how much we look in the past while observing given redshift:

$$t_0 - t(z) = \int_z^0 \frac{dt'}{dz'}dz', \tag{1.15}$$

where t_0 is present age of Universe and $t(z)$ is age of Universe at redshift z. By integrating from ∞ to 0 one gets age of Universe at redshift z, since $z \to \infty$ corresponds with Big Bang, where $a \to 0$.
Finding the distance depends on its definition. In following discussion we focus on flat ($k = 0$) Universe. Using comoving coordinates one can obtain comoving distance D_C which is a current ($t = t_0$) distance between objects and naturally remains constant regardless of the Universe expansion. At small distances ($z << 1$), the velocities caused by Universe expansion follow the Hubble law. Then, for small velocities $z \approx v/c$ and the relation between distance r and redshift can be approximated by $z = H_0 r/c$. However, $H(z)$ is not constant, thus at higher redshift one start observing the epochs where Hubble

parameter start differing significantly from H_0 and this simple linear relation is no more applicable.
For the flat spacetime ($k = 0$) assuming radial light travel in FLRW metric, one have $cdt = a(t)dr$. Writing dt in terms of redshift and Hubble parameter (1.14), one gets:

$$D_C(z) = \int_0^{r(z)} dr = c \int_0^z \frac{dz'}{H(z')}. \tag{1.16}$$

This relation describes comoving distance to the object at cosmological redshift z. The expression *cosmological* stands for the fact that in reality galaxies have also peculiar velocities resulting from their gravitational interactions. The radial component of peculiar velocity is a matter of Doppler effect which contributes to the observed redshift of an object. This effect is discussed in Sec. 1.2.6.

Another form is luminosity distance D_L defined as the distance at which the object with absolute luminosity L would be placed to provide flux F in standard Euclidean geometry, i.e.:

$$F = \frac{L}{4\pi D_L^2}. \tag{1.17}$$

The energy of photon approaching from redshift z is $1 + z$ times smaller. Simultaneously, photon arrival rate is decreased by the same factor, lowering the flux as it would be expected from placing the object at D_C, by $(1 + z)^2$. This fact combined with 1.17 gives the relation:

$$D_L = (1 + z)D_C. \tag{1.18}$$

The luminosity distance is a crucial measure connected with the observations. The third one is angular diameter distance D_A, obtained from the relation between physical and angular size of object placed at D_A. It is linked with comoving distance as $D_A = D_C/(1+z)$. The angular diameter distance is used for estimating distances to objects with known physical size such as peak scale of Baryonic Accoustic Oscillations (BAO). It can be used for comparing D_A computed by observed angular size, with the distance based on known redshift to estimate Universe curvature k.

The measurements of distances have a fundamental meaning in cosmology. By using redshift-independent distance indicators such as standard candles and standard rulers, the relations between distances and redshift allow for estimating Hubble factor and cosmological density parameters. These parameters at $t = t_0$ can be also extrapolated from Cosmic Microwave Background with acoustic peaks at temperature power spectrum. It gives opportunity to understand better the composition and the physics governing the evolution of the Universe and to predict its future shape. That can be achieved with robust and well verifiable cosmological model.

1.1.3 ΛCDM - standard cosmological model

Similarly to particle physics [21], currently cosmology has a standard commonly adopted model. The model is called Lambda-Cold Dark Matter (ΛCDM) [3]. It assumes that General Relativity is a proper theory of gravity at cosmological scales and adopts cosmological constant Λ which allows for accelerated expansion of the Universe. Along with the background evolution, the ΛCDM model considers the existence of cold dark matter which conjointly with the ordinary baryonic component consists of an ingredient of the Universe Ω_M. While the real nature of dark matter is not known yet, based on how the gravity of DM influence e.g. the shape of cosmic web one can put constraints on its properties. The name *cold* stands for the type of dark matter with relatively high particle mass. The CDM is then characterized by low velocities of particles what allows for creating halos at various scales where particles do not escape from gravitational wells of local overdensities. Besides CDM, the recognized is warm dark matter with particle masses of order keV [22] and hot dark matter with $\sim$ $1eV$ particles. While hot and thus relativistic dark matter has been falsified [23], the differences between CDM and WDM lie at small scales sub-galactic [24]. However, the CDM is usually used as a reference case for dark matter [25, 26].

The ΛCDM is a highly successful model in describing the observed Universe [27]. Starting from the Cosmic Microwave Background, it predicts the temperature fluctuation power spectra with high significance [3]. The resultant observed large-scale structure is in good agreement with cosmological simulations assuming ΛCDM model in terms of clustering measurements [28] or lensing [29]. The formation of galaxies as seen now was possible due to dark matter halos in which the baryonic matter can aggregate. The model assumes that dark matter is cold, what agrees well with Cosmic Microwave Background measurements [3], Lyman-alpha forest [30] or concentration-mass relation in galaxy clusters [31]. The ΛCDM model predicts many of observables and still new tests of its principles are being performed [32, 33].

Besides the great agreement between ΛCDM and many observations, there are also significant problems. One of main issues was so-called missing satellites problem accounting for the fact that ΛCDM simulations produce much more subhalos compared to number of observed satellite galaxies within the Local Group. The discrepancy was large: Local Group contain more than 40 galaxies of all types [34] of which dwarfs are the majority, while simulations can predict even ten times more [35]. One of common explanation was low efficiency of small halos in forming stars and thus galaxies [36]. However, when accounting properly with detectability of dwarf galaxies, the missing satellites problem can be resolved [37, 38].

A major issue is so-called the core-cusp problem [39] linked with the profile of dark matter in dwarf galaxies. The density profile obtained from galaxy rotation curves suggest constant density core at the center. Simulations with ΛCDM provide however cuspy

profiles, as well as fitted NFW profile (Section 1.2.4). There are many solutions of this problem, from different dark matter characteristics [40], to baryonic feedback[41], hence the core-cusp problem is an open issue. The other common disagreement is to-big-to-fail problem. It is similar to missing satellites issue, but points out that ΛCDM predicts sub-halos massive enough to form galaxies to be more abundant than real number of satellite galaxies [42]. Among the solutions, also baryonic feedback processes or alternative dark matter models [43] have been proposed.
An important category of disagreements is represented by tensions between local and early Universe. One of the most discussed one is Hubble tension which is a mismatch between Hubble parameter measured in local Universe ($H_0 = 73.04 \pm 1.04\ km\ s^{-1} Mpc^{-1}$ [44]) and H_0 extrapolated from Cosmological Microwave Background ($H_0 = 67.4 \pm 0.5\ km\ s^{-1} Mpc^{-1}$ [3]). There are many attempts to understand this $\sim 5\sigma$ tension, e.g. introducing time-variation of dark energy equation of state, models with additional interactions or just systematics in the data [45]. The Hubble tension becomes however an unresolved issue. Despite of great potential in predicting the evolution of Universe, ΛCDM is not fully describing all phenomena. The number of more or less significant disagreements [46] motivate for further investigations.

1.1.4 Why to test gravity?

Along with the addressed ΛCDM issues there are also fundamental problems of cosmology. First, a major dilemma is the existence of dark energy. While the cosmological constant Λ has been added to Einstein Equations and the Friedman equations can describe the accelerated expansion of Universe, the proper nature of dark energy is yet unknown. There are many other possible explanations, such as quintessence [47] where dark energy is related with a scalar field, but so far there are no strong claims which interpretation could be correct. Second unknown component is naturally dark matter, which together with dark energy constitute of $\Omega_\Lambda + \Omega_{DM} \approx 0.95$ of total Universe energy budget. As well as dark energy, there are no findings stating that dark matter exhibit any non-gravitational interaction with baryonic matter.
The existence of such abundant components that interact only gravitationally strongly motivates to question the validity of standard theory of gravity at largest scales. While the scales from micrometers up to Solar System have been investigated, there is still a room for testing the gravity at cosmological distances [48]. The ΛCDM paradigm stating that General Relativity is a valid theory of gravity at all scales does not necessarily have to be true. The extrapolation from scales of largest gravity tests to distances exceeding hundreds of megaparsecs in cosmological regime spans even by ~ 15 orders of magnitude. This is much more than the scale difference between classical physics ($< 1mm$) and quantum regime at nanometers scale. While General Relativity does not have a quantum limit, it may also be not fully applicable at largest scales. Yet another argument for testing the

gravity arises from the inflation. This involves e.g. primordial gravitational waves [49] or non-gaussianity [50]. Instead of validating the gravity at large scales, studying inflation gives opportunity for test gravity in extremely strong field regime.

1.1.5 Modified and extended gravity

The vast unknown about correct nature of gravity led to the formulations of its alternative theories and modified gravity models. While GR has so far passed all tests and become successful in predicting many phenomena, it is motivated to use it as a base for extensions and attempt for explaining large scale anomalies. The possibility of modifying GR is constrained by Lovelock theorem [51]. For the aim of this thesis we focus on $f(R)$ [52] and normal branch of Dvali-Gabadadze-Porrati ($nDGP$)[53] modified gravity. These models achieve accelerated expansion by modification of the Einstein-Hilbert action. The Einstein-Hilbert action in General Relativity with cosmological constant in natural units ($c = 1$) reads:

$$S_{GR} = \int d^4x \left[\frac{1}{2\kappa}(R - 2\Lambda) + \mathcal{L}_M\right]\sqrt{-det(g_{\mu\nu})}, \tag{1.19}$$

where $\kappa = 8\pi G$ and $\mathcal{L}_M$ is the matter Lagrangian density from which energy-momentum tensor $T_{\mu\nu}$ is derived. In the considered modified gravity scenarios this form is reshaped, simultaneously incorporating a constraint to recover GR in regimes where it has passed observational tests.

f(R) and chameleon screening

The $f(R)$ model assumes the generalization of Einstein-Hilbert action by dependence on Ricci scalar R and adding its functional $f(R)$ tuned to mimic the effect of cosmological constant Λ. It provides action of the form:

$$S_{f(R)} = \int d^4x \left[\frac{1}{2\kappa}(R + f(R)) + \mathcal{L}_M\right]\sqrt{-det(g_{\mu\nu})}. \tag{1.20}$$

For the aim of this work we focus on Hu-Sawicki f(R) [54]:

$$f(R) = -m^2\frac{c_1(-R/m^2)^n}{c_2(-R/m^2)^n + 1}, \tag{1.21}$$

where n, c_1 and c_2 are dimensionless parameters and m is a mass scale parameter. While varying the S_{GR} with respect to the metric provides Einstein Equation 1.1, performing the same with $S_{f(R)}$ gives modified Einstein equations for $f(R)$ scenario:

$$G_{\mu\nu} + \Big(R_{\mu\nu} - \nabla_\mu\nabla_\nu + g_{\mu\nu}\Box\Big)f_R - \frac{1}{2}f(R)g_{\mu\nu} = \kappa T_{\mu\nu}, \tag{1.22}$$

where $G_{\mu\nu}$ is the Einstein tensor introduced after Eq. 1.1, $\Box = \nabla^\mu \nabla_\mu$ is d'Alembertian and $f_R = df(R)/dR$ is scalaron being a scalar field specific for $f(R)$ gravity:

$$f_R = -n\frac{c_1}{c_2^2}\frac{(-R/m^2)^{n-1}}{\Big(1+(-R/m^2)^n\Big)^2}. \tag{1.23}$$

To match the observed Universe expansion history including acceleration as it would involve Λ, the parameters have to fulfill conditions $c_1/c_2 = 6\Omega_\Lambda/\Omega_M$ and $m^2 = H_0^2\Omega_M^2$. In other words, even if Λ does not appear explicitly in $f(R)$, the model is tuned to match ΛCDM expansion which is well described using $\Omega_{\Lambda,0}$ density parameter. Given also that $|R|/m^2 \gg 1$ and for $n = 1$, in current time $t = t_0$ the Hu-Sawicki $f(R)$ model can be characterized by one parameter:

$$f_{R0} = -\frac{1}{c_2}\frac{6\Omega_{\Lambda,0}}{\Omega_{M,0}}\Big(\frac{H_0^2\Omega_{M,0}}{R_0}\Big)^2, \tag{1.24}$$

where R_0 is Ricci scalar at $t = t_0$. Setting the value of c_2, or alternatively of f_{R0} defines then the Hu-Sawicki $f(R)$ model. In this work we focus on models with $f_{R0} = 10^{-6}$ and $f_{R0} = 10^{-5}$ denoted as $F6$ and $F5$, respectively.

The scalaron is an additional degree of freedom compared to standard GR and it is responsible for the enhancement of gravity in EG models, called fifth force. The value of fifth force can be computed with known scalaron field f_R, obtained by taking the trace of Eq. 1.22:

$$\Box f_R = \frac{1}{3}(R - Rf_R + 2f(R) + 8\pi G\rho). \tag{1.25}$$

One can neglect the time dependence, then $\Box \to \nabla^2$. To have a reliable gravity model, it has to recover General Relativity in regimes where it has been successfully validated. For modified gravity scenario it can be obtained with so-called screening mechanism. The common approach is screening mechanism suppressing the fifth force in high density or small scales and thus restoring GR there. The $f(R)$ gravity incorporates chameleon screening mechanism [55], where the extent of screening depends on local density. The scalaron field f_R is massive, with mass computed using relation 1.25 given by:

$$m_{f_R} = \frac{d^2}{df_R^2}\Big(V_{eff}\Big) \approx \frac{1}{3}{f_{RR}}^{-1}, \tag{1.26}$$

where V_{eff} is the effective potential such that $dV_{eff}/df_R = \Box f_R$ and $f_{RR} \equiv d^2 f(R)/dR^2$ [56]. By computing the fifth force:

$$F_5 = -\frac{1}{2}\vec{\nabla}\delta f_R, \tag{1.27}$$

where $\delta f_R = f_R - \bar{f_R}$ is the difference between local f_R and global scalaron $\bar{f_R}$ value, for a point mass source one obtains approximate solution of fifth force dependent on the distance r to be:

$$F_5(r) \propto (1 + m_{f_R} r)\frac{1}{r^2}e^{-m_{f_R} r}. \tag{1.28}$$

From 1.23, one can approximate $f_{RR} \propto R^{-(n+2)}$ and hence in high density regions the mass m_{f_R} scales as R^{n+2} (1.26) and the solution of Eq. 1.28 decays quickly resulting in suppression of the fifth force. In low density environments m_{f_R} becomes small, hence one obtains $F_5 \propto r^{-2}$ and the screening becomes weaker.

nDGP and Vainshtein screening

The $nDGP$ modified gravity scenario is a normal branch of DGP gravity. The idea is that the matter in Universe exists on a $3+1$-dimensional brane which is embedded within $4+1$ Minkowski spacetime. In this case the action is modified and contains both 4- and 5-dimensional term:

$$S_{nDGP} = \int_{brane} d^4x\sqrt{-det(g_{\mu\nu})}\Big[\frac{R}{16\pi G} + \mathcal{L}_M\Big] + \int d^5x\sqrt{-det(g^{(5)}_{\mu\nu})}\frac{R^{(5)}}{16\pi G^{(5)}}, \tag{1.29}$$

where $g^{(5)}_{\mu\nu}$ is five-dimensional metric, while $R^{(5)}$ and $G^{(5)}$ are Ricci scalar for $g^{(5)}_{\mu\nu}$ and five-dimensional gravitational constant, respectively. In $nDGP$ the gravity is assumed to propagate through entire five-dimensional spacetime, while other fundamental forces are limited to standard $3+1$ dimensional spacetime. This causes the gravity to leak into fifth dimension when the second term in 1.29 dominates. It happens at the so-called crossover scale r_C:

$$r_C = \frac{1}{2}\frac{G}{G^{(5)}}, \tag{1.30}$$

which defines the transition scale between standard 4-dimensional gravity playing a role at scales $r < r_C$ and 5-dimensional gravity, governing at $r > r_C$. In this work we consider two cases of $nDGP$: $N5$ and $N1$, parameterized with $H_0 r_C = 5$ and $H_0 r_C = 1$, respectively. In $nDGP$ models the large-scale gravity is however not weaker than in the case of GR, but rather enhanced due to extra scalar degree of freedom. The enhancement is limited at small scales due to Vainshtein screening mechanism [57] incorporated by $nDGP$ gravity model. Modified Poisson equation in $nDGP$ reads:

$$\nabla^2\Psi = 4\pi G a^2\rho + \frac{1}{2}\nabla^2\phi, \tag{1.31}$$

where Ψ is the full potential, first term on right hand-side refers to the standard gravity and ϕ is additional scalar field due to model degree of freedom. The equation of scalar field in $FLRW$ metric is given by:

$$\nabla^2\phi + \frac{r_C^2}{3\beta_n(a)a^2}\Big[(\nabla^2\phi)^2 - (\nabla_i\nabla_j\phi)(\nabla^i\nabla^j\phi)\Big] = \frac{8\pi G a^2}{3\beta_n(a)}\rho, \tag{1.32}$$

where the function $\beta_n(a)$ is:

$$\beta_n(a) = 1 + 2Hr_C\Big(1 + \frac{\dot{H}}{3H^2}\Big). \tag{1.33}$$

At large scales one have $|\nabla^2\phi| \ll r_C^{-2}$. In the relation 1.32 first term on left hand side dominates and:

$$\nabla^2\phi \approx \frac{8\pi G a^2}{3\beta_n(a)}\rho. \tag{1.34}$$

Inserting this to equation 1.31 one gets spatially independent enhancement of the field at large scales, leading to effective gravitational constant:

$$G_{eff} = G\Big(1 + \frac{1}{3\beta_n(a)}\Big). \tag{1.35}$$

At small scales one have $|\nabla^2\phi| \gg r_C^{-2}$, hence in 1.32 the second term on the left hand side dominate, and then one gets $\nabla^2\phi \ll 8\pi G a^2\rho/(3\beta_n)$. It recovers the standard Poisson equation when inserted to 1.31, making $G_{eff} \to G$. This is so-called Vainshtein screening mechanism which suppresses the fifth force at small scales. The $nDGP$ models are characterized by crossover scale r_C which separates the regimes of standard gravity and the one dominated by five-dimensional space time. The physical scale below which the fifth force originating from ϕ is screened, is given by Vainshtein radius r_V. Given the mass M of gravitational field source, it is: [56]:

$$r_V = \Big(\frac{16 r_C^2 GM}{9\beta_n^2}\Big)^{1/3}. \tag{1.36}$$

The $nDGP$ scenario is then an interplay of two effects. First one is leaking of gravity at scales $r > r_C$ into five-dimensional spacetime. The second is scalar field ϕ being an additional model degree of freedom. It enhances the gravity with fifth force leading to $G_{eff} > G$, but the Vainshtein screening mechanism suppresses it at distances $r < r_V$ from massive objects.

1.1.6 Constraints and testing

The screening mechanisms integrated with the $f(R)$ and $nDGP$ gravity scenarios allow them to pass stringent small-scale and high-density tests by recovering General Relativity in these regimes. To put constraints on these models, one need to investigate large-scale behavior of gravity in observations and identify potential differences in growth of structure compared to simulations. Therefore the growth rate and growth factor are the parameters to quantitatively discriminate between the gravity scenarios, given the observations and simulations outputs. The structure formation in GR and a mention in context of MG cases is discussed in Sec. 1.2. Constraining gravity models is widely performed with the clustering measurements and other statistics of large-scale structure, what is the subject of Sec. 2.

For $nDGP$ model, the promising test concerns also Integrated Sachs-Wolfe (ISW) effect [58]. This however plays a role at large scales where cosmic variance limits the measurement. Instead, the studies of redshift-space distortions and clustering allowed

to constrain the $nDGP$ gravity to be $(H_0 r_C)^{-1} < 0.97$ with 2σ confidence [59]. More recent research on weak lensing limited this parameter to $log_{10}(H_0 r_C) > 0.5$ with 25% precision [60]. The same study provided upper constraint for the $Hu - Sawicki$ model of $f(R)$ as $log_{10}(f_{R0}) < -6$ with 2% precision. The upper limit of $f(R)$ can be also found using peculiar velocities [61]. However, it is also worth investigating the $f(R)$ and $nDGP$ models with parameters in a range excluded by recent works. It helps to understand how the interplay between enhanced gravity and screening mechanisms shape the cosmic web. Knowing how the deviations from GR behave in more extreme MG models and comparing them with moderate scenarios allows for better constraints in the future.
Such research focuses on the formation and evolution of large-scale structure, which provides large laboratory for testing the models. Hence, understanding the physics behind that provides an opportunity to unveil a part of nature of the Universe.

1.2 Large-Scale Structure

Large-scale structure (LSS) is a greatest formation known to humanity so far. The so-called cosmic web is composed of matter arranged in nodes, connected with filaments and walls stretched around much emptier voids [62, 63]. It consists a complex system evolving on various scales, where smaller structures build up larger systems, creating hierarchical pattern.

According to current state of knowledge, LSS originates from tiny primordial density inhomogeneities being the product of pre-inflationary quantum fluctuations [64]. While Universe expanded and cooled, the physical size of the inhomogeneities increased up. Then, due to gravitational instability, the density field evolved towards the present shape.

Since young Universe was very smooth and density fluctuations at the era of recombination were of order 10^{-5} [65], the evolution of density field at these times can be well characterized by linear Perturbation Theory.

1.2.1 Gravitational instability and linear structure growth

To analytically describe the self-gravitating matter, it is convenient to work with density contrast $\delta(\vec{x}, t)$, defined by:

$$\delta(\vec{x}, t) \equiv \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \tag{1.37}$$

where $\rho(\vec{x}, t)$ and $\bar{\rho}(t)$ are the density at point $\vec{x}$ and time t and average density at moment t, respectively. Gravitational potentials at the scales relevant for structure growth are weak. Additionally, the scales themselves are significantly smaller than Hubble Radius. These facts allow for working in Newtonian limit [66], without applying full GR solutions. One of relations describing the evolution of density field is continuity equation which

characterizes the conservation of mass:

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla \cdot \Big((1+\delta)\vec{v}\Big) = 0, \tag{1.38}$$

where $\vec{v}$ is a peculiar velocity of matter element, motion relative to the expanding background. In addition to this, the evolution occurs with conservation of momentum, described by Euler equation:

$$\frac{\partial}{\partial t}(a\vec{v}) + (\vec{v}\cdot\nabla)\vec{v} = -\nabla\phi - \frac{1}{\bar{\rho}}\nabla(\delta p), \tag{1.39}$$

where ϕ is gravitational potential and p stands for pressure. On top of the expressions imposing conservation of mass and momentum, there is Poisson equation [67] specifying how density field and its own gravitational potential interact:

$$\nabla^2\phi = 4\pi G\bar{\rho}a^2\delta. \tag{1.40}$$

Assuming small perturbations with $|\delta| << 1$, the combination of equations 1.38, 1.39 and 1.40 in Fourier space brings:

$$\ddot{\delta}_k + 2\frac{\dot{a}}{a}\dot{\delta}_k + \Big(\frac{c_s^2k^2}{a^2} - 4\pi G\bar{\rho}\Big)\delta_k = 0, \tag{1.41}$$

where c_s and k are sound speed and wavenumber, respectively. For the case of presureless Einstein de-Sitter (EdS) Universe where $a \propto t^{2/3}$, the growing solution of Eq. 1.41 is $\delta \propto t^{2/3} \propto a$.

By considering the equilibrium and computing corresponding length scale $\lambda \equiv 2\pi/k$, one can obtain Jeans Mass, the border condition of density fluctuation to collapse:

$$M_J \propto \Big(\frac{c_s^2}{G}\Big)^{3/2}\frac{1}{\sqrt{\bar{\rho}}}. \tag{1.42}$$

For gravity-dominating case, i.e. when:

$$4\pi G\bar{\rho} > \frac{c_s^2k^2}{a^2}, \tag{1.43}$$

one obtains a solution of Eq. 1.41 being composed of growing and decaying terms. The decaying term vanishes in time and becomes irrelevant. While the Equation 1.41 contains only derivatives with respect to time and $\delta = \delta(\vec{x}, t)$, one can disentangle spatial and time-dependent part. The growth of structure can be then described with:

$$\delta(\vec{x}, t) = D(t)\delta_0(\vec{x}), \tag{1.44}$$

where $D(t)$ is the growth rate, $\delta_0(\vec{x}) \equiv \delta(\vec{x}, t = t_0)$ such that t_0 is a considered starting time. By construction, growth rate fulfills $D(t_0) = 1$. For any cosmological parameters and pressure-less case, the growth rate obtains a form:

$$D(t) = \frac{\dot{a}}{a}\int_0^a \frac{da'}{\dot{a}'^3}. \tag{1.45}$$

The solution 1.44 has a remarkable consequence: it shows that in linear regime the density fluctuations do not evolve spatially, but their amplitude increases. The growth of amplitude means that slightly overdense regions become denser in time, while underdense ones get emptier. This is a consequence of matter moving into gravitational potential wells.

By defining the growth factor f as:

$$f \equiv \frac{dlogD}{dloga}, \tag{1.46}$$

linearized continuity equation 1.38 with assumed $\delta << 1$ [66] becomes:

$$\nabla \vec{v} = -aHf\delta. \tag{1.47}$$

The growth factor relates with the density parameter at given redshift by $f \approx \Omega_M(z)^\gamma$, where $\gamma \approx 0.55$ [68]. The growth factor and its dependence on cosmological matter density parameter exhibits later on, while studying the statistics of LSS in redshift space what is discussed in Section 2.

1.2.2 Modified gravity case

While the considered modified gravity models: $f(R)$ and $nDGP$ are tuned to match ΛCDM cosmic expansion history, the structure formation differs of that from the standard model. The relation from Eq. 1.41 describing the evolution of density contrast instead of standard gravitational constant G, has to include effective gravitational constant G_{eff} (for nDGP 1.35) resulting from gravity enhanced by fifth force (for $f(R)$ 1.28). For $nDGP$ the linear growth is scale-independent, while $f(R)$ model reveals scale-dependent growth of structure. This later has remarkable implications on comparing clustering estimates as a function of scale, between these models and GR. For the case of $f(R)$, the enhancement of gravity called fifth force is screened in dense regions, hence making these areas to evolve as ruled by General Relativity. The sparse unscreened areas such as voids are then subjects of stronger gravitational interactions. These regions became a motivated source of searching for beyond-GR signals [69]. Also the abundance of voids can become a test for structure evolution [70]. Such selective screening mechanism may even partially restore the gaussianity of density field compared to standard GR scenario, where underdense regions become catastrophically empty due to mass transfer into large filaments and nodes. In $f(R)$, voids may not run out of such amount of mass due to locally enhanced gravity.

The fifth force accompanied with screening mechanism naturally affects also the creation of dark matter halos. The halo mass function i.e. distribution of dark matter halo masses is enhanced with respect to GR and depends strongly on gravity scenario [71]. It thus directly affects the population of galaxies, however the verification with observations

is nontrivial. In terms of representative cosmic volume mock galaxies within extended gravity, one is hence still limited to non-hydrodynamical schemes (see Sec. 1.2.7).
For velocity field, the deviations from GR are generally present at nonlinear scales $r \lesssim 10\ Mpc$ [72]. This might add potential extended gravity signal in redshift space at such separations. We discuss the beyond-standard models of gravity in Section 1.1.5.

1.2.3 Spherical collapse model

The analysis on linear structure growth concerns the young Universe where the only density fluctuations corresponding with structure formation were of $\delta << 1$. Nonetheless, at later epochs the density field evolved into more complex structure with density contrast significantly exceeding the unity at wide range of scales. The evolution of density field beyond the linear theory can be studied with Zeldovich approximation [73] which provides robust evolution of density field until the so-called shell-crossing occurs. The cosmic web evolution as a whole can not be efficiently described analytically after the relevant fluctuations entered highly non-linear regime.

The evolution of overdensities after the linear growth can described by nonlinear spherical collapse model. Since $\Omega_M \propto (1+z)^3$, the EdS model is a good approximation of our Universe at higher redshifts $z > 2$ when it was dominated by matter.

By assuming isolated symmetric spherical overdense region of mass M_s and radius r_s, its evolution can be described by Newtonian physics:

$$\frac{d^2 r_s}{dt^2} = -\frac{GM_s}{r_s^2}. \tag{1.48}$$

The parametric solution reads [74]:

$$r_s(\eta) = \alpha(1 - cos\eta) \tag{1.49}$$

$$t(\eta) = t_0 + \sqrt{\frac{\alpha^3}{GM_s}}(\eta - sin\eta), \tag{1.50}$$

where η is a parameter, $\alpha = const$ and $t(\eta = 0) = t_0$. The size of overdensity firstly increases following the Hubble flow, then it slows down due to its gravitation and stops at $\eta = \pi$ obtaining radius $r_{max} = 2\alpha$, called turnaround radius. With the density of fluctuation $\rho = M_s/(4/3\pi r_s^3)$, one can compute its density contrast. Inserting scale factor evolution $a \propto t^{2/3}$ in EdS Universe into first Friedman Equation, 1.6, one gets the relation:

$$\bar{\rho} = \frac{1}{6\pi G t^2} \tag{1.51}$$

describing average density evolution in EdS Universe.

Combining ρ with $\bar{\rho}$ in Eq. 1.37 and using the parametric solutions 1.50 with $t_0 = 0$, one obtains the density contrast of the overdensity:

$$\delta_s = \frac{9(\eta - sin\eta)^2}{2(1 - cos\eta)^3} - 1. \tag{1.52}$$

At the turnaround phase ($\eta_{turn} = \pi$), the contrast is $\delta_{turn} \approx 4.55$. This is well beyond the linear regime assuming $|\delta| << 1$. The spherical collapse model is a significant simplification of reality where e.g. the overdensities are not isolated and expansion of Universe also plays a role in this process. In more physically motivated scenario, the gravitational interactions with neighboring anisotropies make the collapse to be ellipsoidal [75]. However, the spherical case remains a useful ground description of nonlinear structure formation.

The solution 1.50 shows that after the turnaround phase the overdensity starts to collapse. The solution gives $r_s(\eta = 2\pi) = 0$ what would indicate collapsing into a singularity. However, it assumes isolated overdensity composed of pressureless matter with only radial velocities. In more physically accurate scenario, the overdensity collapses until it virializes, i.e. reaches equilibrium where potential and kinetic energy are balanced [76]. In such conditions the virialized dark matter halo is being created [77] in which eventually galaxies may form [78]. Gravitationally virialized system fulfills the condition:

$$2T + V = 0, \tag{1.53}$$

where T and V are averaged kinetic and potential energy, respectively. Taking into consideration two critical moments: turnaround phase where the entire energy of the system is in form of potential energy and virialization point, by assuming total energy conservation and using the relation 1.53, one obtains:

$$V_{vir} = 2V_{turn}, \tag{1.54}$$

where $V_{vir,turn}$ is average potential energy in virialization and turnaround respectively. Since $V \propto r$, the same relation connects overdensity radius at turnaround and virialization phase. Given that, the density relation at these moments is: $\rho_{vir} = 8\rho_{turn}$. Assuming that virialization occurs at the collapse time from solution 1.50 i.e. $t_{vir} = 2t_{turn}$, using the relation 1.51 one gets:

$$\bar{\rho}_{vir} = 4\bar{\rho}_{turn}, \tag{1.55}$$

where $\bar{\rho}_{vir,turn}$ is background density at virialization and turnaround, respectively. The relation of density and background density 1.55 at two phases allows for estimating the relative density of virialized dark matter halo:

$$\frac{\rho_{vir}}{\bar{\rho}_{vir}} = \frac{8\rho_{turn}}{\bar{\rho}_{turn}/4} = 32(\delta_{turn} + 1), \tag{1.56}$$

what gives halo density contrast of $\delta_{halo} \approx 177$.

1.2.4 Dark matter halos

According to the our state of knowledge, dark matter halos consist gravitational wells where baryonic matter may accumulate. While galaxies are not required to form a'priori within halos, the dark matter constituting $\approx 85\%$ [3] of matter in the Universe, dominates

the gravitational potential where these can evolve. According to hierarchical structure formation scenario where smaller structures form and build up larger structures, dark matter halos at certain scale contain smaller halos, called subhalos. While the main halo can host a galaxy, usually denoted as central galaxy, the embedded-in subhalo may contain a smaller companion galaxy, usually called satellite galaxy. The halo-subhalo hierarchy has been found in simulations [79, 80, 81, 82] thus supporting the hierarchical model.

Understanding the abundances and properties of dark matter halos is crucial in studying the large-scale structure due to existing galaxy-halo connections, described in more details in Sec. 1.2.5. The clumps of dark matter are not directly observable and the nature of this matter component is yet unknown. Therefore, the method of verifying the models of halos is based on constraining their effect on observed visible part - galaxies.

The most fundamental property of dark matter halo is its mass. Due to correlation between halo masses and abundance of galaxies, studying galaxy-halo connection provides strong cosmological constraints. The halo mass function (HMF), i.e. expected number of halos of a given mass in unit volume is described by the Press-Schechter theory [77].

Alongside with the distribution of halo masses, their internal structure also plays an important role in cosmology. The most commonly adopted density profile of dark matter halo is Navarro-Frenk-White (NFW) profile [83] which is a fit from N-body simulations:

$$\rho(R) = \frac{\rho_s}{\frac{R}{R_s}\left(1 + \frac{R}{R_s}\right)^2}, \tag{1.57}$$

where ρ_s and R_s are the parameters - characteristic density and scaling radius, respectively.

The density profile is a relevant quantity in terms of halo-galaxy connection. Estimating the profile provides better understanding of motion of galaxies hosted by the halo, what could influence redshift-space distortion signal within one-halo regime, i.e. scales comparable or smaller than size of halo. Additionally, halo density profile is often employed for statistically populating halos with galaxies in gravity-only simulations (see Sec. 1.2.7). Both issues have then further implications on galaxy clustering measurements.

1.2.5 Galaxies

Galaxies glowing in a vast space represent a visible tracers of large-scale structure in wide range of distances. These objects being composed of baryonic matter immersed within halos of dark matter, provide essential information about cosmology, interplay between gravitational potential of whole density field and visible matter, and physical processes ruling their evolution.

Galaxies started to form while the gas infall to dark matter-dominated gravitational wells. The gas shocks to the virial temperature being an equilibrium temperature of uniform and isothermal gas cloud [84]. The mass of cloud, i.e. the amount of baryonic

matter accumulated by halo depends on the gravitational potential of halo. The heavier the halo, generally the more baryons agglomerate within it what is the main aspect of halo-galaxy connection.

Properties

The observations show that the cosmic web is composed with galaxies of various types, such as disk galaxies, as well as e.g. elliptical or irregular. While the shapes of disk galaxies are due to angular momentum conservations, other types are assumed to evolve differently. Elliptical galaxies are believed to form by mergers of galaxies with comparable masses [85] or by the so-called monolith collapse scenario [86] where the formation of stars occurs rapidly compared to the time-scale of collapse after radiative cooling. While massive elliptical galaxies can form through both evolutionary channels, dwarf elliptical galaxies are thought to evolve through merging and other scenarios [87].
Due to direct observational data, the known picture of galaxies is far more complex than for dark matter halos. Besides the morphology, most significant properties of the galaxy are its mass, luminosity and colors. The mass can be estimated with different methods, such as measuring the rotation curve which provides the cumulative mass profile of baryonic and dark matter enclosed within corresponding radius. Similar approach is associated with measuring the motion of satellite galaxies. While there are various methods of measuring masses [88], the luminosity can be obtained e.g. with brightness-distance or Tully-Fisher relation [89, 84].The Tuly-Fischer relation can be used for obtaining so-called redshift-independent distances (discussed in more details in Sec. 1.2.6) for disk galaxies. With these one can study observationally radial peculiar velocities [90].

Galaxy-halo connection

The galaxies are not distributed randomly in space. Their positions are related with underlying density field. Since heavier halos can accumulate more baryonic mass, these objects are the hosts to statistically more massive galaxies.

However, not every dark matter halo contains a galaxy. Baryonic matter constitutes $\sim 15\%$ of matter. If the halo is not massive enough, its potential well will not accumulate sufficient amount of gas to form a galaxy. Such halo may be also stripped-out of gas due to interactions with nearby heavier objects. The minimal halo mass for hosting a galaxy is however not a constant value, but rather it depends on cosmic web environment and redshift [91, 92].

The galaxies are biased tracers of the density field, since their formation is more efficient in denser environment. The relation between galaxy and full matter density field can be described using high-peak bias model [93]. The relation between density contrast

of galaxies δ_g and matter density contrast δ_m can be expanded into Taylor series [94]:

$$\delta_g = \sum_{i=0}^{\infty} \frac{b_i}{i!} \delta_m^i, \tag{1.58}$$

where b_i are bias coefficients. The factor b_1 is a linear bias which gives robust description at linear scales where $|\delta_m| \ll 1$. Higher bias coefficients apply at nonlinear scales and play a role in more accurate modeling of cosmic web [95].

1.2.6 Velocity field and redshift space

The density field of cosmic web is directly linked with its velocity field. While studying the velocities of galaxies in clusters which provide insight into the clusters masses via virial theorem [96], velocity field allows for more precise study of LSS. One has then not only information about spatial distribution of objects, but also complementary insight into growth rate and direct link with so-called redshift-space distortions which we discuss below. The peculiar velocities can not be easily disentangled from motions related with the expansion of Universe. Differentiating 1.4 with respect to time yields:

$$\dot{\vec{r}} = \frac{\dot{a}}{a}\vec{r} + a\dot{\vec{x}}. \tag{1.59}$$

By ignoring the second term, one obtains Hubble law where the relative velocity between objects is caused only by expansion of space. However in reality on top of the background expansion each galaxy has peculiar velocity which is caused mostly by gravitational interactions. In observations disentangling between peculiar and cosmic motions is not straightforward, since both velocities add-up into measured redshift z. While mapping the objects in redshift space it leads to so-called redshift-space distortions (RSD) [97] which modify the distant-redshift relation due to nonzero peculiar velocities. The distorted position is:

$$\vec{r} = \vec{r_0} + \frac{1 + z_{cos}}{H(z_{cos})} v_{||,pec} \hat{e_r}, \tag{1.60}$$

where r_0 is undistorted position, z_{cos} stands for cosmological redshift of the object and $v_{||,pec}$ is its peculiar velocity projected onto observer's line-of-sight.

Disentangling between cosmic and peculiar velocity can be made with redshift-independent distance indicators, such as Tully-Fisher relation [89], standard candles such as Cepheids [98], Ia supernovae [99] or fundamental plane [100]. Redshift-independent distances allow not only for analysis of peculiar velocity field. After extracting the redshift counterpart responsible for cosmological expansion, these provide the estimate of local Hubble parameter. The typical bulk velocity i.e. average of local peculiar velocity flows, is $v_{pec} \sim 200 - 500\ km\ s^{-1}$ [90]. However the measurements of peculiar motions of single galaxies usually suffer from large errors of $\sim 20\%$. Furthermore, the errors of observed peculiar velocities increase with distance. Therefore these measurements are applicable mostly for galaxies at low redshift $z \lesssim 0.1$.

The velocity field and consequently redshift-space distortions depend on the scale of consideration. On small separations $r \lesssim 10Mpc$ [101, 102] the velocities are statistically random, dominated by virial motions. Due to the velocity dispersion, the spatially spherical overdensity in redshift space becomes stretched along observer's line of sight, while preserving the shape in tangential direction. The imprints on large-scale structure caused by this effect are known as Fingers of God. At larger scales $r \gtrsim 10Mpc$, the overdensities are smaller and Hubble flow becomes non-negligible. The areas of non-virialized overdensities expand, but slower than the background due to its gravity. The virial motions are not dominating the velocity field at such scales, hence in the surrounding of overdensity velocities generate coherent inflows towards the region of higher density. From observer's point of view, the objects between them and overdensity are moving away faster than it would result from the Hubble flow. Simultaneously, the galaxies behind the overdensity will be also moving towards it, decreasing the total radial velocity by v_{pec}. In redshift space, the overdense region become then squashed, creating the effect called Kaiser squashing [97]. Fig. 1.1 shows a simulation catalog slice with real space and redshift space mapping with distant observer approximation assuming line of sight parallel to axis x. The used simulation is described in Section 1.2.7. In real space no direction is distinguished, but z-space reveals clear distortions. At these scales the Kaiser squashing is dominating what is mainly visible in higher internal densities of filaments and nodes. However, in redshift space one can also see structures stretched along LOS (Fingers of God), e.g. at $\vec{r} \approx [50, 15]\ Mpc/h$.

The description of density contrast considering redshift-space distortions involves converting the coordinates into z-space. In linear regime the velocity field is curl-free, hence by considering the conservation of number of objects:

$$(1 + \delta(s))d^3s = (1 + \delta(r))d^3r, \tag{1.61}$$

where $\delta(r)$ and $\delta(s)$ is density contrast in real and redshift space respectively. In Fourier-space the relation between $\delta(r)$ and $\delta(s)$ in context of linear theory is [97]:

$$\delta_k^{(s)} = (1 + \beta\mu_k^2)\delta_k, \tag{1.62}$$

where $\mu_k \equiv \cos(\theta)$, $\beta = f/b_1$ and θ is angle between wave vector $\vec{k}$ and line of sight. This shows that redshift-space distortions which originate from peculiar velocity field, create a growth factor-dependent imprint onto the density field. The relation 1.62 is also relevant in measuring the statistics of the density field.

1.2.7 Simulations and catalogs

Studying the cosmic web at late times where highly nonlinear gravitational evolution and baryonic physics shape its formation requires numerical approach. The continuity 1.38, Euler 1.39 and Poisson 1.40 equations can not be solved fully analytically. Adopting

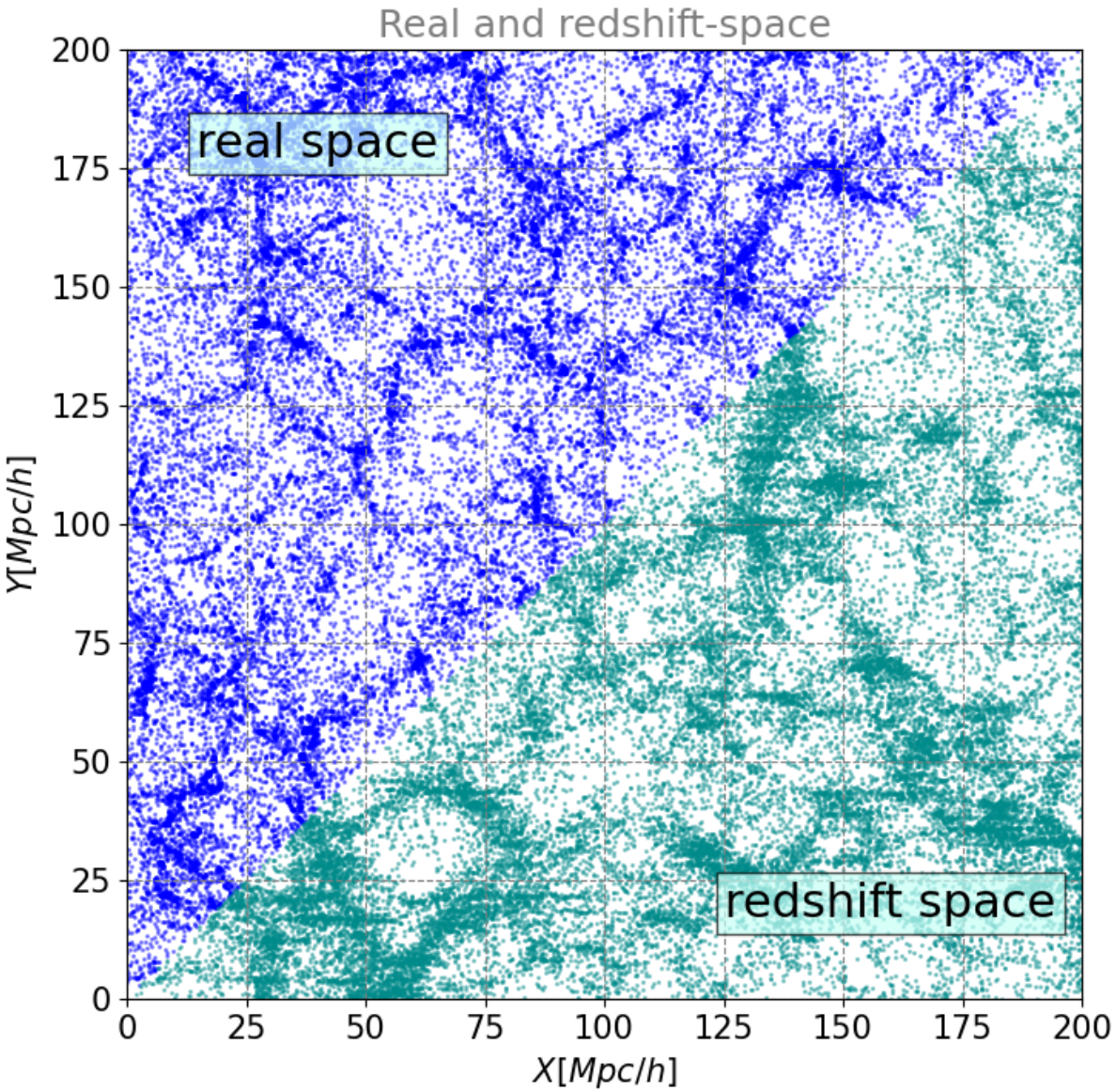

Figure 1.1: A 20 Mpc/h slice along $z-axis$ of $COLAVERSE$ simulation box at redshift $z = 0.5$, zoomed at $0 - 200$ Mpc/h range. The simulation is described in Section 1.2.7. The data points are dark matter pseudo-particles show in real space (left upper part) and redshift space (lower right). In z-space the distant observer approximation was assumed, i.e. the observer's lines of sight are parallel to each other. The assumed LOS direction is $x - axis$.

the Perturbation Theory which consider perturbative expansion up to i-th order gives reasonable approximations [103] as long as the terms of order bigger than i are negligible. However, the effort for computing these terms grows rapidly.

Cosmological simulations allow for tracking the large-scale structure formation in a controlled environment. This opens a wide range of possibilities for testing models by comparing the mock results with observations. From the verifications of various grav-

ity scenarios affecting the clustering [104], halo-galaxy connection [105], to cosmic web environment studies [63] or constraints on dark matter [106, 107], simulations provide multitude of applications in modern cosmology.

The standard method is to simulate the cosmic web in a box which expands according to the corresponding cosmology. To account for boundary effects, the simulation box is periodic. Since at larger scales Universe is homogeneous and isotropic, the influence of local density fluctuations on the Hubble expansion is expected to be negligible [108, 109].

In the simulations one assumes large number of pseudo-particles N_p interacting with each other, while the background cosmological expansion is not affected by the inhomogeneities. These are so-called N-body simulations. The aim is to evolve the positions and velocities of pseudo-particles under assumption of given cosmology, gravity and other interactions depending on type of the simulation. Due to non-relativistic velocities at cosmic distances, the N-body cosmological simulations generally use Newtonian approximations for inter-particle interactions rather than full General Relativity.
Since simulating the cosmic web with resolution of atoms is currently not possible, each pseudo-particle represents collective mass, corresponding with assumed resolution. In cosmology these masses can be of order 10^5 $M_\odot$ [82], up to $\sim 10^{12}$ $M_\odot$ [110]. The usual number of particles is in range of $N_p \sim 500^3$ up to thousands per simulation box side. The box length L_{BOX} depends on the purpose of the simulation. While estimating the statistics over cosmic web segmentation on filaments, wall, voids and nodes or studying clustering up to scales of tens of megaparsecs, one needs box of $L_{BOX} = 500 Mpc/h$, where $h = H/(100\ km\ s^{-1} Mpc^{-1})$[1] or more. However, the computational cost grows both with number of particles and the box size. Therefore large box simulations generally have lower resolution than these with small L_{BOX}. The simulations with box size of order $100 Mpc/h$ [80] do not allow for studying statistics over large scales, instead these have spatial an mass resolution suitable for e.g. investigating one-halo scales or satellite and central galaxy properties and their interaction with environment.

Besides of geometric properties such as box size and resolution, there are two major categories of cosmological simulations. The former ones, hydrodynamical simulations employ hydrodynamical schemes and account for gas thermodynamics or such processes as stellar formation and evolution physics, radiation heating or ionization. Such approach allows for more physically reliable recreation of galaxy populations. By comparing with existing cosmological survey such as e.g. GAMA [111], SDSS [112] or DESI [113], one can put constraints on galaxy evolution mechanisms. The hydrodynamical simulations are however computationally expensive and these usually are suited for relatively small volumes [82, 114].

The alternative category is N-body gravity-only cosmological simulations. These con-

[1] In cosmological simulations it is convenient to use units independent on Hubble parameter, e.g. Mpc/h for length or $M_\odot/h$ for mass, due to high uncertainty of $H(z)$ in real Universe

sider only gravitational forces what makes them much less expensive computationally when compared to hydrodynamical cases. This allows for constructing simulations with large L_{BOX}. Gravity-only approach is highly motivated while this is the dominating force at large scales and for $\sim 85\%$ of mass in the Universe there are no detections of non-gravitational interactions. However, neglecting the baryonic component prevent from studying the physics of galaxies directly. Instead, one can investigate populations of galaxies by linking them with dark matter halos, using methods as halo occupation distribution (HOD) [115], subhalo abundance matching (SHAM) [116] or conditional luminosity function [117] which are discussed in Sec. 1.2.7.

For the purpose of this thesis we focus on gravity-only simulations. In the study we have employed two non-hydrodynamical simulations. The former is ELEPHANT (Extended LEnsing PHysics using ANalytic ray Tracing)[118], N-body simulation of $N_p = 1024^3$ dark matter pseudo-particles in a box of $L_{BOX} = 1024\ Mpc/h$ assuming cosmology from $WMAP-9$ [119]. Being computed with $ECOSMOG$ [120] and $ECOMOG$ [121] codes which are based on publicly available $RAMSES$ code [122], the used $ELEPHANT$ simulation results consider the evolution in standard General Relativity, two $f(R)$ models: $F5$ and $F6$, and two $nDGP$ scenarios: $N1$ and $N5$ which have been discussed in 1.1.5 and 1.1.5.

The latter is $COLAVERSE$, dedicated simulation computed by Prof. Wojciech Hellwing considering $n_p = 1024^3$ particles inside $L_{BOX} = 500\ Mpc/h$ box assuming Planck-18 [123] cosmology. The $COLAVERSE$ simulation was computed with $COLA$ method which is an approximate method, described in section 1.2.7. This simulation provides models with varying two cosmological parameters: Ω_M and σ_8, while setting other parameters constant.

As for other various types of simulations, the resultant effect highly depends on state at the beginning. Simulating the cosmic web using identical physics and parameters may give quantitatively different outputs if their beginning setups were different. Therefore well-defined initial conditions are crucial for understanding the developed structure.

Initial conditions

The density perturbations at young Universe were small. The furthest so far [124] detection, Cosmic Microwave Background providing information about temperature fluctuations at redshift $z \sim 1100$ show that the relative density at that time was $\rho/\bar{\rho} \approx 10^{-5}$. The commonly used starting point in simulations is $t_{IC} \sim 10$ Myr after the Big Bang, where the density contrast was still $\delta \ll 1$. Therefore one can use Lagrangian Perturbation Theory to generate physically and observationally motivated initial conditions (IC) for simulations. The comoving positions x of particles at IC are related with so-called pre-initial positions q as:

$$x = q + \Psi, \tag{1.63}$$

where Ψ is the displacement field. The commonly used pre-initial positions are the regular grid in simulation box [125] or glass being a random field evolved with repulsive gravity with damping [126]. The displacement field Ψ has a perturbative solution. While the ICs are in highly linear regime, it is motivated to consider only first term, the Zeldovich approximation [73]:

$$|\Psi_1| = D(t)\ \nabla_q \phi, \tag{1.64}$$

where $D(t)$ is the growth rate and ∇_q is composed of partial derivatives with respect to positions q of the gravitational potential ϕ. By inserting Ψ_1 and differentiating 1.63 with respect to time, one can relate the IC velocities with scale factor f 1.46:

$$x = q - DHf\ \nabla_q \phi, \tag{1.65}$$

where $H = \dot{a}/a$ is Hubble factor. The gravitational potential is connected to density field with Poisson equation 1.40. Due to the definition 2.1, the displacement field is then also related with power spectrum. This allows for creating scale-dependent initial conditions which correspond with the appropriate cosmological parameters obtained from the observations of CMB and extrapolated to the redshift of IC, $z = z_{IC}$.
The common and convenient approach is to prepare for one model multiple realizations of initial conditions which are statistically identical, but differ with phases of displacement. The physics and cosmological parameters of simulations prepared on such ICs is equal. Such approach allows for estimating reliable uncertainties in further study, given that every realization is an independent representation of equivalent Universes with the same physics.

Computational methods in N-body simulations

Evolving the discretized field of particles is a nontrivial task even for collisionless dark matter-only simulations. Computing simply the force acting on each particle from the collective gravitational field by summing them and repeating that for all particles requires computational effort which scales as N_p^2. Therefore alternative methods and improvements have been developed.
Solving the Poisson equation 1.40 can be obtained by pixelizing the particles into a grid to get estimate of density field and convolving it with Green function [127]. The forces f from the definition $f = -\nabla\phi$ are then calculated with finite-differences. This so-called particle-mesh (PM) method is computationally efficient, however it can not resolve the structures at scales below the mesh resolution. The another method which addresses this problem is adaptive mesh refinement (AMR) in which the grid becomes finer in areas of higher density. This allows for resolving the structures which are inaccessible with PM method, but the performance becomes slower.

Simulating the evolution of particles field in extended gravity scenarios requires adjusted approach. For Hu-Sawicki variant of $f(R)$ and normal branch of Dvali-Gabadadze-

Porrati ($nDGP$) models and simulation which have been employed in this study, the extended-gravity version of AMR is used. Due to the specific density-dependent screening mechanisms in that models, the method resolving well the regions with higher density is crucial. Since the equations of motions in extended gravity have additional degrees of freedom compared to those for GR, these require a different approach. The codes $ECOMOG$ and $ECOSMOG$ which the $ELEPHANT$ simulation was computed with, solve the modified Poisson and motion equations using iterative Gauss-Seidel relaxation.

Among the methods for evolving the particle field, there is $COLA$ (COmoving Lagrangian Acceleration) method [128] which allows for time-effective simulation of cosmic web. The assumptions of this approach are to obtain positions and velocities of particles by computing corrections for Lagrangian Perturbation Theory instead of solving fully the equations of motions to find numerically the position in phase-space. In COLA method, the comoving position of particle is:

$$\vec{x} = \vec{x}_{LPT} + \vec{\delta x}, \tag{1.66}$$

where $\vec{x}_{LPT}$ is position obtained from Linear Perturbation Theory and $\vec{\delta x}$ is the correction. By computing second derivative with respect to time and inserting relation 1.63, for $\ddot{\vec{x}} = -\nabla\phi$ one gets:

$$\ddot{\vec{\delta x}} = -\nabla\phi - \ddot{\Psi}. \tag{1.67}$$

The displacement field Ψ is then expanded into perturbative series:

$$\Psi = \sum_i \Psi_i. \tag{1.68}$$

With Zeldovich approximation as the first-order result, the generalized solution for assuming standard ΛCDM model is:

$$\Psi_i = D_i(t)\Psi_{ini}, \tag{1.69}$$

where $D_i(t)$ is the growth rate corresponding with i-th order of perturbation. By computing the factor one can obtain corrected positions of particles. Similarly, the velocities are obtained by differentiating the relation 1.66. Due to the combination of analytical prescription, the COLA method provides fast convergence in small number of steps. Simulating up to a chosen redshift requires even thousand times less steps when compared to standard N-body algorithm. However the time-saving approach is occupied with low robustness at smaller scales, since COLA method can not properly evolve highly non-linear regimes. Nonetheless, this is an useful prescription for studying larger scale ranges with multiple simulations.

Identifying halos

The gravity-only simulations provide the distribution and velocities of dark matter pseudo-particles. In the real Universe dark matter clumps into halos of different sizes and masses.

To study dark matter halos from such simulations one needs to extract that information from DM pseudo-particle field being a discretized representation of density field.
The one of most commonly used methods of finding halos in simulations is Friends-of-Friends (FoF) algorithm [129]. It is based on a concept that two pseudo-particles belong to one halo if the distance between them is not greater than so-called linking length l_f. The frequently used linking length is 0.2 of mean inter-particle separation [84], i.e. $l_f = 0.2\sqrt[3]{L_{BOX}/N_{side}}$, where $N_{side} = N_p^{1/3}$ is the number of particles per simulation box side.
The method starts with a particle and searches for all neighbors (friends) which are situated not further than l_f. Then, for each of them one searches for another neighbors situated within $D \in [0, l_f]$ distance range. The procedure is repeated until there are no more particles fulfilling the distance criterion. By assumption of the method, the linked particles belong to one halo and other halos can be found by repeating this algorithm fo other yet-unconnected particles in the simulation.
The FoF is a reliable method for finding halos in the simulations. The pure algorithm however suffers from bridging problem which is artificially linking the separate halos which are connected with low density bridge of particles. To account for this, one can use spherical overdensity method and chose only symmetric overdense regions without bridge structures. For more accurate results one can search for halos using 6-dimensional phase-space, including both positions and velocities. Such extension is adopted in $ROCKSTAR$ code [130] which is adopted in ELEPHANT simulation. There are also other extensions and methods of halo finding, taking into account e.g. tidal fields [131] or mergers [132]. Algorithms for finding halos are extensively compared in [133].
The properties of identified halos such as mass can be defined in various forms. Due to virialized halos having overdensity of around 200 (1.56), the convenient mass estimators are M_{200b} and M_{200c}, the mass enclosed within a radius where the density is 200 times the background or critical density, respectively. What is estimated from cosmological sky surveys are the densities as a fraction of critical density $\rho_{crit} = 3H^2/8\pi G$. While M_{200b} would be preferred for simulations-only study due to its direct relation with mean density of pseudo-particle field, the usage of M_{200c} is hence more motivated for cosmological applications. The estimate of M_{200c} can be obtained by computing spherical overdensity around the halo center, as implemented in $ROCKSTAR$ code. Based on this quantity, one can define virial radius R_{200c}, distance from halo center at which the enclosed spherical overdensity mass reaches M_{200c}. Since halos do not have a boundary, the virial radius can be treated as a proxy for their size.

Non-hydrodynamic approaches for obtaining galaxies

Gravity-only N-body cosmological simulations due to their simplicity compared to fully hydrodynamical prescriptions allow for investigating much larger volumes of mock Uni-

verse. Although, with this approach one can not produce galaxies. Instead, there are statistical and observationally motivated methods for assigning galaxies to dark matter halos.

On of such schemes is SubHalo Abundance Matching (SHAM) [116]. The general concept is that galaxies residing in halos and subhalos have properties which are related with properties of their hosts. For halos the most substantial feature is their mass. The method then assumes that heavier halos contain more massive galaxies, what is well motivated by large-scale structure formation scenarios. By using the luminosity or stellar mass function of galaxies obtained from the observations, one match them to appropriate halos and subhalos to provide monotonic relation between galaxy stellar mass and halo mass. This ensures that the cumulative function for galaxies and halo mass function from simulation agree. The alternative property of halos which may be used for matching with galaxy is V_{max}, the maximum orbital velocity of dark matter particles within halo. Additionally, the employed property of halo can be also related with its history, e.g. maximal mass it had which correlates with stellar mass of residing galaxy [105].
The statistics of mock galaxies assigned with this method are found to agree well with observed data. Measurements of two- and three-point correlation functions for $SHAM$ galaxies are highly consistent with these measured on observed data [134, 135] what makes it a reliable method of obtaining galaxies in non-hydrodynamical simulations.

The $SHAM$ method is a deterministic prescription which assigns galaxies one-by-one to the halos, based on their correlating properties. An alternative approach for populating halos with galaxies is Halo Occupation Distribution (HOD). While the general assumption of positive halo mass - galaxy stellar mass correlation stays unchanged, HOD is a purely statistical prescription. For a halo with mass M_h it assumes the probability $P(N_g|M_h)$ that it hosts N_g galaxies. The probability usually discriminates between central and satellite galaxies. The common form of expression defining average number of central N_{cen} galaxies in a halo is [136]:

$$\Big\langle N_{cen}(M_h)\Big\rangle = \frac{1}{2}\Big[1 + \mathrm{erf}\Big(\frac{log(M_h) - log(M_{min})}{\sigma_{logM}}\Big)\Big], \tag{1.70}$$

where erf is error function and M_{min} and σ_{logM} are free parameters of HOD. Since halo can host at most one central galaxy, one draws them from Bernoulli distribution with mean from 1.70 [137], i.e. the value $\langle N_{cen}(M_h)\rangle$ becomes the probability that halo of mass M_h has central galaxy. By construction the parameter M_{min} is the halo mass for which the probability of hosting central galaxy is equal to 0.5, whereas the value of σ_{logM} controls the slope of probability function.
The average number of satellite galaxies in a halo is given by:

$$\Big\langle N_{sat}(M_h)\Big\rangle = \Big\langle N_{cen}\Big\rangle\Big(\frac{M_h - M_0}{M_1}\Big)^{\alpha}, \tag{1.71}$$

where M_0, M_1 and α are the parameters. The value of M_0 sets average minimal halo mass

required to host satellite galaxy, M_1 is normalization mass and α defines the slope. Multiplying by $\langle N_{cen} \rangle$ not only makes the number of satellites depend on host halo mass, but also it prevents from ambiguous situation in which halo hosts satellite galaxies while not having central galaxy. The number of satellites is often drawn from Poisson distribution with mean defined by 1.71.
Central galaxies are usually assumed to be placed in the center of host halo potential [136], while satellites are distributed within $[0, R_{200c}]$ distance range from the center, tracing NFW profile 1.57. The alternative approach is to attach satellite galaxies to positions of pseudo-particles sampled from the halo.
Assigning velocities depends on the status of the galaxy, similarly as the position. For centrals, it is usually given by velocity of the center of mass of the halo, what ensures that central galaxy remains at halo center. There exist also modifications which assume deviation from this, so-called velocity bias [138]. Velocities of satellites are often given by velocity of central galaxy with Gaussian scatter in each direction with dispersion matching root mean square dispersion of host halo [136].
Assigning the positions and velocities to galaxies is relevant also in the context of different scenarios of gravity. For $nDGP$ and $f(R)$ cases, the concentration of halos defined as [79]:

$$c_{200c} = \frac{R_{200c}}{R_S}, \tag{1.72}$$

where R_S is scaling radius from NFW profile (1.57), can be different than that for General Relativity [139, 140].

The Halo Occupation Distribution method is a powerful tool for populating halos from gravity-only simulations with galaxies. By tuning the parameters one can obtain mock galaxy catalogs which mimic the observations in context of chosen statistics, e.g. 2-point correlation function. This allows for studying different models with other statistics, testing systematics and understanding the halo-galaxy connection. Along with alternatives and extensions as $SHAM$ or Conditional Luminosity Function CLF [141] relating galaxy luminosity with M_h, one can create realistic galaxy sample within controlled environment.

Catalogs and lightcone projection

One can study mock cosmic web and its tracers in a geometry as they are simulated - in the box. This is the distant observer approximation mentioned in Sec. 1.2.6: the observer is assumed to be located far away from the simulation box and lines of sight are parallel to each other. It is a convenient approach for many applications which require statistical measurements over large volume at single redshift. The observed Universe is however not a simulation box with all the objects having the same redshift. One instead observes the large-scale structure in a geometry of lightcone: the farther the object is, the more time its light travels to the observer what combined with expansion of the Universe results in redshift-distance relation. Creating mock catalogs which reliably recover full geometry and

selection criteria of the observations provide opportunity to study the simulated Universe as it would be observed with sky survey. After careful accounting for systematics, it allows for confronting such mock data with existing survey result and verifying the validity of our predictions.

To create a lightcone, dependent on the required depth one needs to combine simulation outputs from different redshifts. For the case of flat ΛCDM cosmology, the relation 1.16 between redshift z and comoving distance D_C defines then the radial distance component of the object in lightcone representation while its exact redshift is given. In the observer frame of reference the $D_C(z)$ relation is continuous, while the state of simulated Universe is saved for discrete points in time called snapshots. Therefore placing the objects directly at comoving distances corresponding with simulation output redshifts would create the snapshots compressed into flat shells. For the simulations with dense snapshot grid, i.e. small time step between adjacent outputs, obtaining the positions and peculiar velocities at any redshift can be performed with interpolation [142, 143, 144]. It is convenient to start by physically centering the box of i-th simulation snapshot at distance $D_C(z_i)$ from the observer, where z_i is the redshift of the snapshot. Then the positions of objects placed closer or further from the observer than $D_C(z_i)$ need to be interpolated in time to match the relation:

$$D_C(z_{interp}) = |\vec{r}_{interp}|, \tag{1.73}$$

where $\vec{r}_{interp}$ is the interpolated position and z_{interp} is redshift of interpolated moment. This method requires additional caution to avoid e.g. double counting of the same object from different snapshots.

The size of simulation box is often smaller than desired lightcone depth, i.e. $L_{BOX} < D_C(z_{max})$, where z_{zmax} is maximum redshift considered. In such cases the common approach is to replicate the box and treat the combination as new box with larger size. This is highly motivated in simulations with periodic boundary conditions. Among the distance preserving mappings the rotation and reflection is also included, however these impose discontinuities in density and velocity fields.

However, not every simulation provides outputs with small time steps. For the cases of sparse snapshot sampling, the interpolation may be a subject of nonphysical artifacts in a lightcone. The convenient approach is to place replicated boxes from i-th snapshot by covering the comoving distance of $[0, D_C(z_{max})]$ around the observer and then cut out comoving distance ranges which correspond with other redshifts. The limits of distances $D_{C,i}$ for i-th snapshot can be defined as:

$$D_{C,i} \in \Big[\frac{D_C(z_{i+1}) + D_C(z_i)}{2}, \frac{D_C(z_i) + D_C(z_{i-1})}{2}\Big], \tag{1.74}$$

where $i-1$ and $i+1$ stand for earlier and later snapshot than i-th respectively, the boundaries for i-th snapshot are set to be in the middle of comoving distances to adjacent redshifts and naturally $D_C(z_{i+1}) < D_C(z_i) < D_C(z_{i-1})$. The resultant positions obtained

with this method are hence the same as one would get from the interpolation approach while assuming all velocities to be zero. With such constructed 3D distribution one then compute sky positions $[RA, DEC]$. From D_C the redshifts can be obtained by reversing the numerical quantities from the relation 1.16 using e.g. bisection or Newton method.

The lightcone generated with such method allows then for accessing the simulation output from observer-like point of view. Figure 1.2 shows a sample slice through redshifts up to $z \sim 0.5$ of a lightcone created from *ELEPHANT* simulation with standard *GR* gravity. For demonstration purposes and due to sparse galaxy sample in *Elephant*, the visualized data is dark matter pseudo particles.

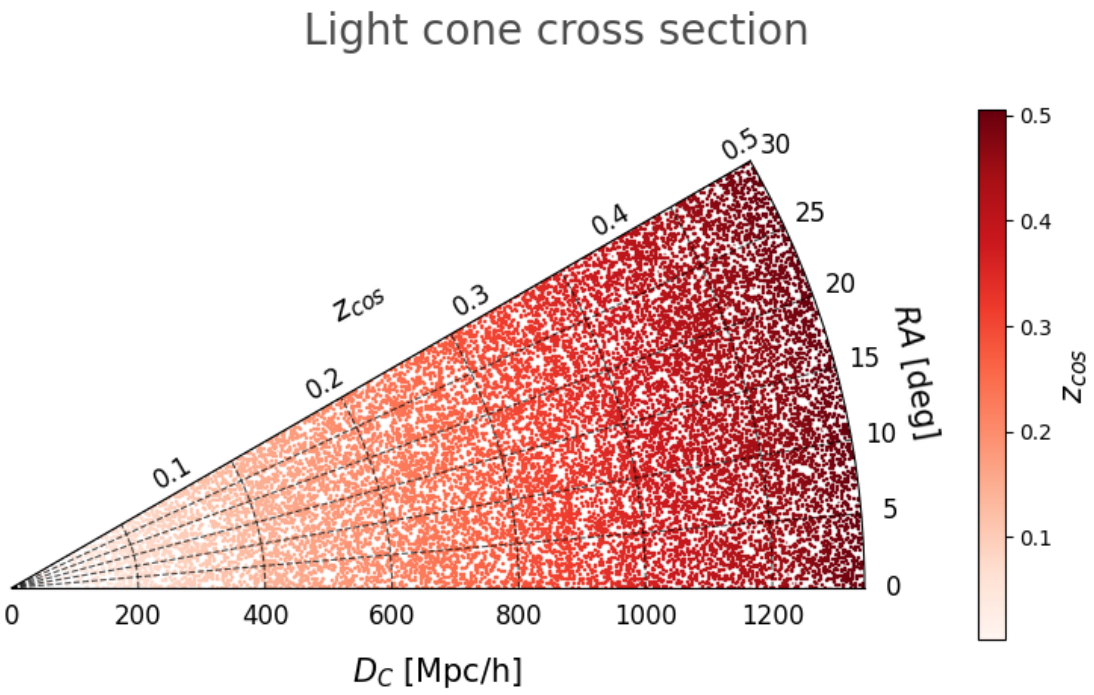

Figure 1.2: Radial cross section through light cone made from *ELEPHANT* dark matter pseudo-particles at $RA \in [0, 30]$ deg of 300 Mpc/h vertical thickness, i.e. $|D_C \sin(DEC)| < 200\ Mpc/h$.

The scatter plot shows the distribution with respect to comoving distance and cosmological redshift.

For the purpose of accurate mimicking the observations, besides creating galaxy lightcones one would need to incorporate additional effects, e.g. object selection. Reconstructing the survey geometry requires applying angular and radial masks. While the angular footprint is mostly a consequence of selective sky area observations, radial selection is mainly a result of limited sensitiveness of detector and exposure time. If mock catalog considers galaxy luminosities and colors, comparing lightcone with observations additionally requires either computing observed brightness of mock galaxies, or estimating

the absolute luminosities of observed galaxies with k-correction [145]. In such case the common approach is to then create so-called volume-limited or flux-limited galaxy samples, while the surveys have a magnitude limitation for galaxy brightness. For instance, volume-limited sample assumes choosing a defined redshift bin $[z_{min}, z_{max}]$ and selecting only the galaxies from this range whose absolute brightness is greater than absolute brightness of faintest observable galaxies at z_{max}. This allows for considering all galaxies within given distance ranges observable by a specific survey. The two latter samples are defined analogously by taking luminosity of flux as a first cut.
With lightcones, either these mimicking various properties of survey, or constructed to investigate projection and selection effects, one can consequently convert controlled-environment output with imposed physics into point of view of a hypothetical observer.

Chapter 2

Statistical properties of density field

Statistical descriptions of the cosmic web remain one of the key tools in cosmology. These have a wide spectrum of applications, from studying galaxy evolution to testing structure formation models and gravity constraints. In observational cosmology the clustering measurements of different galaxy populations tell us about abundance of galaxies or cosmological parameters [146, 147]. To achieve this, one needs to employ robust measure which can inform how the objects are statistically distributed within the large-scale structure.

2.1 Power spectrum and correlation function

The power spectrum P(k) and 2-point correlation function (2PCF) are the most commonly used statistical descriptions of the cosmic web. The former characterizes the amplitude of density field fluctuations as a function of wavenumber k related straightforwardly with scale r as $k = 2\pi/r$. Power spectrum is defined as squared modulus of Fourier space density contrast at given k:

$$P(k) \equiv \langle |\delta_k|^2 \rangle, \tag{2.1}$$

where $\langle ... \rangle$ indicate averaging by all orientations of vector $\vec{k}$ in Fourier space.

Since the inflation generated scale-invariant fluctuations, in post-inflationary Universe the $P(k)$ was in a shape of power-law [66]:

$$P(k) \propto k^{n_s}. \tag{2.2}$$

The n_s, factor describing slope of primordial power spectrum is an important cosmological parameter being derived e.g. from CMB observations [3]. Yet at the era of recombination, power spectrum of temperature anisotropies was different from power-law form. At later times due to growth of inhomogeneities and baryonic effects, the density field resulting from early-time anisotropies evolved, and so the matter power spectrum. This statistics is sensitive on structure formation physics. Therefore, except of estimating cosmological

parameters, $P(k)$ is being used e.g. in constraining dark matter properties [148], probing nonlinear-scale physics or studying direct gravity models [149]. By its construction based on δ_k, the power spectrum allows for investigation of redshift-space distortions. From the definition of power spectrum 2.1 and linear density contrast in redshift space 1.62, one gets [97]:

$$P^{(s)}(k) = (1 + \beta\mu_k)^2 P(k), \tag{2.3}$$

where $P^{(s)}(k)$ and $P(k)$ is redshift space and real space power spectrum, respectively. Such relation has remarkable consequences for analytical estimations of other statistics which can be linked with $P(k)$.

The second most frequently adopted statistics is 2-point correlation function. It defines the excess probability of finding a pair of objects separated by vector $\vec{r}$ relative to the one expected by random distribution. The excess probability dP can be written as [67]:

$$dP = \bar{\rho}^2(1 + \xi(\vec{r}))dV_1 dV_2, \tag{2.4}$$

where $\bar{n}$ is the mean density and $\xi(\vec{r})$ stands for $2PCF$. The dV_1 and dV_2 are volume elements occupied by two particles. Given that the density at point $\vec{x}$ implies finding dn objects in volume dV, i.e. $\rho(\vec{x}) = dn(\vec{x})/dV$, the expected number of particles at position $\vec{x} + \vec{r}$ occupying volume dV reads then:

$$dn(\vec{x} + \vec{r})|_{dV} = \Big(\delta(\vec{x} + \vec{r}) + 1\Big)\bar{\rho}dV, \tag{2.5}$$

where $\bar{\rho}$ is average density. The number of pairs dn_p between position $\vec{x}$ and $\vec{x} + \vec{r}$ occupying volumes dV_1 and dV_2 becomes

$$n_p = dn(\vec{x})|_{dV_1}\ dn(\vec{x} + \vec{r})|_{dV_2}. \tag{2.6}$$

While the average number of pairs is simply $n_{p,av} = \bar{\rho}dV_1\bar{\rho}dV_2$, one can compute the excess number of pairs between position $\vec{x}$ and $\vec{x} + \vec{r}$, over random distribution:

$$\frac{n_p}{n_{p,av}} - 1 = \Big(\delta(\vec{x}) + 1\Big)\Big(\delta(\vec{x} + \vec{r}) + 1\Big). \tag{2.7}$$

By computing the ensemble average of 2.7 over all positions $\vec{x}$, for given separation $\vec{r}$ one gets:

$$\left\langle \frac{n_p}{n_{p,av}} - 1 \right\rangle = \Big\langle \delta(\vec{x})\delta(\vec{x} + \vec{r}) + \delta(\vec{x}) + \delta(\vec{x} + \vec{r}) \Big\rangle. \tag{2.8}$$

By construction of the density contrast $\langle\delta(\vec{x})\rangle = 0$, therefore the definition of $2PCF$ is:

$$\xi(\vec{r}) \equiv \langle\delta(\vec{x} + \vec{r})\delta(\vec{x})\rangle. \tag{2.9}$$

If one assumes isotropy, the ensemble averaging occurs for all orientations of $\vec{r}$ while preserving the separation length $r = \sqrt{\vec{r}\cdot\vec{r}}$, resulting in $\xi(r)$ being dependent on separation length instead of generalized $\xi(\vec{r})$ being a function of separation vector $\vec{r}$.

The 2-point correlation function is related with power spectrum such that $P(k)$ is a Fourier counterpart of $2PCF$. With the assumption of isotropy in all directions, the connection becomes:

$$\xi(r) = \frac{1}{2\pi^2} \int_0^\infty k^2 \frac{sin(kr)}{kr} P(k) dk. \tag{2.10}$$

The isotropic 2-point correlation function being only a function of scale r is a convenient tool for clustering measurements in real space. However, in redshift space the used equivalent is anisotropic correlation function $\xi(r_p, \pi)$, which decomposes the separation r into a component r_p perpendicular to observer's line of sight, and π - parallel to it. Such two-dimensional representation gives direct insight into Fingers of God and Kaiser effects in the clustering measurements from simulations [150] and observations [151]. Without redshift-space distortions $\xi(r_p, \pi)$ would reveal purely radial dependence producing isovalue lines being concentric circles. In redshift space however, the FoG effect is amplifying $2PCF$ at small r_p what appears like smearing the signal along r_π. Simultaneously, Kaiser effect at larger scales causes increasing the signal along transverse direction r_p due to squashing the cosmic web elements along r_π.

However, not for every study the full 3D analysis involving radial measurements is applicable. For instance, investigating the clustering from observations may involve photometric data which are rich in terms of sky coverage and depth, however suffer from large redshift uncertainties. In such case it is convenient to work with projected statistics. For this aim the commonly used is projected correlation function $w_p(r_p)$, related with $\xi(r)$ and anisotropic 2PCF by [152]:

$$w_p(r_p) = 2 \int_0^\infty d\pi \xi(r_p, \pi) = 2 \int_0^\infty dy \; \xi\Big(\sqrt{r_p^2 + y^2}\Big). \tag{2.11}$$

The projected version of correlation function is used to examine real-space clustering of the cosmic web projected along LOS.
Another clustering measurement avoiding radial distances alongside with $w_p(r_p)$, is the angular correlation function $w(\theta)$. In that case, the counts of pairs are performed only for their angular separation θ. This is a simplified approach in which just the angular positions are used instead of $\xi(r_p, \pi)$ integrated along line of sight. Nevertheless, while applied to catalog, this method mixes physical scales due to nonzero redshift thickness.

The direct measurement of two-point correlation function from discrete data, being either a simulation or observation is achieved by pair counts between the data and so-called random catalog. The randoms are used to quantify the catalog counts with respect to random distribution and are usually denser than the data to account for Poisson shot noise effects. While for pure simulations the construction of random catalog is straightforward, in the case of observations these have to assume specific features resulting from the survey, e.g. angular footprint, completeness or fiber collisions [153].

The most commonly used estimator of $2PCF$ is Landy-Szalay estimator [154], which is robust in terms of variance:

$$\xi_{LS}(r) = \frac{DD - 2DR + RR}{RR}, \tag{2.12}$$

where DD is number of pairs of objects in data catalog separated by a chosen distance range $r \pm \Delta r$, RR is equivalent value for random catalog and DR is number of cross-pairs between data and random catalog. If randoms were constructed with different number of objects than data, the pair counts are scaled to account for random sample size.

The literature applying $2PCF$ in simulations and observed data is rich [151, 102, 56, 136, 150, 155]. Along with different variants of 2-point correlation function there exist also modifications such as marked correlation function [156] where galaxy pairs are weighted corresponding to their physical features. Unlike comparing the clustering among separate populations defined by binning over specific characteristics, this approach allows for investigating continuous relation between scale and properties of clustering galaxies.

2.2 Higher orders

Nonetheless, the 2-point analysis can not fully describe statistically the cosmic web. The large-scale structure is assumed to have Gaussian initial conditions [157], for which the entire clustering information is encoded in its power spectrum or $2PCF$. After the linear growth where different Fourier modes δ_k were evolving independently, the nonlinear gravitational evolution imposed mode coupling and thus higher-order dependencies emerged [158]. Since the gravity imprints higher-order statistics within the density field, these can be employed for gravity tests at cosmological scales. To characterize the hierarchy of cosmic web statistics more accurately, one has to investigate beyond 2-point statistics. The generalization of 2PCF are N-point correlation functions. Unlike $2PCF$, its higher order equivalents count sets of N objects with specific inter-distances. For instance, 3-point correlation function $\xi_3(r_1, r_2, r_3)$ counts sets of three objects separated by $r_1 \pm \Delta r$, $r_2 \pm \Delta r$ and $r_3 \pm \Delta r$ and confronting it with analogical counts in random distribution. However, the computational cost grows rapidly with the order N, the interpretability becomes challenging as well. There are developments of 3-point and 4-point correlation function [159, 160], however due to the above problems, the $2PCF$ still remains a more commonly used probe for studying the large-scale structure.

2.3 Smoothed density field and central moments

While studying the statistics of density field on certain scale R, it is also practical to employ smoothed density contrast $\delta_R(\vec{x})$, defined as:

$$\delta_R(\vec{x}) = \int d^3x' \delta(\vec{x'}) W_R(\vec{x} - \vec{x'}), \tag{2.13}$$

where $W_R(\vec{x} - \vec{x'})$ is a spherical top-hat normalized window function:

$$W_R(\vec{x} - \vec{x'}) = \begin{cases} \frac{3}{4\pi R^3} \text{ if } |\vec{x} - \vec{x'}| \leq R \\ 0 \text{ otherwise.} \end{cases} \tag{2.14}$$

Analyzing the probability distribution function (PDF) of the smoothed density contrast δ_R allows to measure the differences with respect to Gaussian field. Acquiring the information about the departure from the initial Gaussian field can be performed with n-point correlation functions, $PDFs$ or using moments of $PDFs$. Matching the shapes of PDFs one-by-one is however ambiguous e.g. due to sampling issues. Due to aforementioned problems with two former cases, we focus on the moments. These allow for pointing out relevant aspects of the distribution such as spread, asymmetry, tailedness or more convoluted characteristics. Simultaneously, according to the Wick's probability theorem, odd moments computed for variable X within a Gaussian distribution with mean equal to zero, vanish and even moments depend on variance σ [161], i.e.:

$$E[X^{2J}] = (2J - 1)!!\ \sigma^{2J}, \tag{2.15}$$

where E stands for expected value, $2J$ is the moment and !! is the double factorial. This fact thus makes the above method a powerful tool in detecting deviation from Gaussian distribution and in consequence gravity signatures 2.2.

By computing ensemble average of moment J of the smoothed field, one obtains so-called volume averaged correlation function $\bar{\xi}_J(R)$:

$$\bar{\xi}_J(R) \equiv \langle \delta_R(\vec{x})^J \rangle. \tag{2.16}$$

The name "averaged" comes from the construction of $\bar{\xi}$ and spatial smoothing. The $\delta_R(\vec{x})$ at each point is a density contrast δ averaged over the spherical volume surrounding that point and consequently the relation between averaged correlation function and N-point correlation function e.g. for case of $N = 3$ is [162, 163]:

$$\bar{\xi}_3(R) = \frac{1}{V} \int_V d^3r_1 d^3r_2 d^3r_3\ \xi_3(r_1, r_2, r_3), \tag{2.17}$$

where V is the probed volume such that $r_i \in [0, R]$, $i = 1, 2, 3$. While the volume-averaged correlation functions are only the functions of one scale parameter in contrast to multidimensional $NPCFs$, their interpretation is much more straightforward. Due to averaging, $\bar{\xi}_J$ provide less information than their N-point counterparts, however their computational cost is also lower what allows for tracing high orders J. That makes the volume averaged correlation functions a complementary method for statistically describing the cosmic web.

The density contrast can be expanded perturbatively around $\delta = 0$ as:

$$\delta = \sum_{i=1}^{\infty} \delta_i, \tag{2.18}$$

where δ_i is the i-th order Perturbation Theory solution [164]. By considering first two terms of 2.18, one can compute the variance:

$$\sigma^2 \equiv \langle \delta^2 \rangle = \langle \delta_1^2 \rangle + 2\langle \delta_1 \delta_2 \rangle + \langle \delta_2^2 \rangle. \tag{2.19}$$

The last two components are negligible, thus $\langle \delta^2 \rangle \approx \langle \delta_1^2 \rangle$. For the third order, by performing the same computations one obtains:

$$\langle \delta^3 \rangle = \langle \delta_1^3 \rangle + 3\langle \delta_1^2 \delta_2 \rangle. \tag{2.20}$$

The term $\langle \delta_1^3 \rangle$ vanishes, additionally $3\langle \delta_1^2 \delta_2 \rangle \propto \langle \delta_1^4 \rangle$ [165], what leads to the proportionality between $\langle \delta^3 \rangle$ and $\langle \delta^2 \rangle^2$. From the definition of averaged correlation functions 2.16, one gets $\bar{\xi_3} \propto \bar{\xi_2^2}$. Performing the analysis for higher orders allows for defining the proportionality factors S_J called cumulants or hierarchical amplitudes. By writing them in terms of averaged correlation functions [166, 94, 167], one gets:

$$S_J \equiv \frac{\bar{\xi_J}}{\bar{\xi_2^{J-1}}}. \tag{2.21}$$

By construction $S_J(R)$ measuring gravity-shaped cosmic web are thus weakly scale-dependent functions. The shape of cumulants inform about the hierarchy between developed orders of density field statistics.

To analytically describe the shape of averaged correlation functions and cumulants, it is convenient to work in Fourier space. According to convolution theorem, the convolution in Fourier space is a multiplication. Therefore the smoothed density contrast from 2.13 becomes:

$$\delta_R(\vec{x}) = \frac{1}{(2\pi)^3} \int d^3k \; \delta(\vec{k}) W_R(\vec{k}) e^{i\vec{k}\cdot\vec{x}}. \tag{2.22}$$

The expression for second-order averaged correlation function is then:

$$\langle \delta_R^2 \rangle = \frac{1}{(2\pi)^6} \int d^3k \; e^{i\vec{k}\cdot\vec{x}} \int d^3k' \; e^{-i\vec{k'}\cdot\vec{x}} W_R(\vec{k}) W_R(\vec{k'}) \langle \delta(\vec{k}) \delta(\vec{k'}) \rangle. \tag{2.23}$$

Employing the alternative definition of power spectrum $P(k)$ which is obtained from its connection with $2PCF$:

$$\langle \delta(\vec{k}) \delta^*(\vec{k'}) \rangle = (2\pi)^3 \delta_D(\vec{k} - \vec{k'}) P(k), \tag{2.24}$$

where δ_D is Dirac delta, from 2.23 one gets [165] the relation between second-order averaged correlation function and the power spectrum:

$$\bar{\xi}_2(R) \equiv \langle \delta_R^2 \rangle = \frac{1}{(2\pi)^3} \int d^3k \; \Big[W_R(\vec{k}) \Big]^2 P(k). \tag{2.25}$$

For the case of order $J = 3$, one can use the relation 2.20 and the solution of second order [165]:

$$\delta_2(\vec{x}) = \frac{1}{(2\pi)^3} \int d^3k e^{-i\vec{k}\cdot\vec{x}} \int d^3k e^{-i\vec{k'}\cdot\vec{x}} \delta_1(\vec{k}) \; \delta_1(\vec{k'}) \Big[\frac{5}{7} + \frac{\vec{k}\cdot\vec{k'}}{k'^2} + \frac{2}{7} \Big(\frac{\vec{k}\cdot\vec{k'}}{k \; k'} \Big)^2 \Big]. \tag{2.26}$$

One can obtain $\langle \delta_R^3 \rangle$ by inserting this to 2.20 and using the relation from 2.22. Performing this analysis for higher orders requires much more effort. These contain more contributions of coupled lower terms δ_i, hence applying next perturbative expansion terms becomes impractical due to computational reasons. The complication increases while studying additional effects, e.g. redshift space distortions where isotropy assumption fails. Quantitative studies of observed or simulated large-scale structure with volume averaged correlation functions of various orders require estimating them numerically [168, 169, 170].

2.3.1 Numerical approach

The definition 2.16 of $\bar{\xi}_J$ is constructed using continuous density field. However, the galaxies and dark matter halos, even dark matter pseudo-particles in the simulations form a discrete field. While considering the scale R, one can treat counts of these tracers occupying corresponding volume as a proxy for smoothed local density, i.e.:

$$N_{\vec{x},R} \approx \rho_R(\vec{x}), \tag{2.27}$$

where $N_{\vec{x},R}$ is number of objects within a sphere of radius R, centered at $\vec{x}$. Due to the linearity of convolution operation, the relation between smoothed density contrast $\delta_R(\vec{x})$ and smoothed density $\rho_R(\vec{x})$ is the same as for non-smoothed field, i.e. 1.37. Therefore one can expand the definition 2.16 into:

$$\bar{\xi}_J = \bar{\rho}_R{}^{-J} \Big\langle (\rho_R(\vec{x}) - \bar{\rho}_R)^J \Big\rangle, \tag{2.28}$$

where $\bar{\rho}_R$ is the average of smoothed density. The combination of relations 2.27 and 2.28 shows that the estimate of volume-averaged correlation function on discrete tracers can be obtained using central moments of counts:

$$m_{J,c} = E\Big[(N_i - \bar{N})^J\Big]. \tag{2.29}$$

Obtaining the set of counts N_i and average $\bar{N}$ is performed by counting the objects at various positions $\vec{x}_i$ allowing to cover statistically representative volume of catalog or simulation box. It is assumed that averaging the counts is equivalent to taking ensemble average if the volume is representative. This so-called counts-in-cells (CIC) method is a common approach for computing $\bar{\xi}_J$ [167, 171, 172, 173, 169].

Central moments of the counts contain contributions from lower orders. This is relevant especially in terms of Gaussian distribution where even orders of $J > 4$ are functions of variance. To subtract these contributions and, as a consequence, obtain moments that at $J > 2$ measure the departure from Gaussianity, one can use the so-called connected moments μ_J [167, 174]. The moment generating function $M(t)$ which enables to define moments m_J of given distribution is:

$$M(t) \equiv E[e^{tX}]. \tag{2.30}$$

Expanding the exponent into Taylor series around zero and inserting $m_J \equiv E[X^J]$ [1] gives the exact relation between $M(t)$ and moments:

$$M(t) = \sum_{J=0}^{\infty} \frac{m_J}{J!} t^J. \tag{2.31}$$

While the moments m_J can be then obtained from:

$$m_J = \frac{d^J}{dt^J} M(t) \Big|_{t=0}, \tag{2.32}$$

the connected moments μ_J are defined by:

$$\mu_J = \frac{d^J}{dt^J} ln\big(M(t)\big) \Big|_{t=0}. \tag{2.33}$$

By construction μ_J vanish for $J \geq 3$ if Gaussian distribution is assumed what makes it suitable for detecting non Gaussianity. Up to $J = 3$, μ_J and raw moments m_J match, however at bigger J the differences appear. Higher order connected moments up to $J = 5$ are:

$$\mu_4 = m_4 - 3m_2^2 \tag{2.34}$$

$$\mu_5 = m_5 - 10 m_2 m_3. \tag{2.35}$$

Since the computation of μ_J is performed on discrete objects, these are a subject of Poisson shot noise, especially at the scales where $\bar{N} \sim 1$ and lower. Subtracting this contribution requires assuming that the Poisson distribution has the same mean $\bar{N}$ as the resultant CIC measurements. The generating function for discrete field M_D reads [175]:

$$M_D(t) = M(e^t - 1). \tag{2.36}$$

Inserting the relation 2.36 into 2.31 calculated on connected moments provides the expression of connected moments with Poisson contributions. To obtain shot noise - corrected moments k_J, one needs to subtract the noise components. The corrections c_J are then:

$$c_J = \frac{d^J}{dt^J} \Big[\sum_{J=0}^{\infty} \frac{\mu_J}{J!} (e^t - 1)^J \Big] \Big|_{t=0} - \mu_J. \tag{2.37}$$

Finally, the corrected terms are provided by $k_J = \mu_J - c_J$. By computing the corrected moments up to order $J = 5$, one gets:

$$k_2 = \mu_2 - \bar{N} \tag{2.38}$$

$$k_3 = \mu_3 - 3k_2 - \bar{N} \tag{2.39}$$

$$k_4 = \mu_4 - 6k_3 - 7k_2 - \bar{N} \tag{2.40}$$

$$k_5 = \mu_5 - 10k_4 - 25k_3 - 15k_2 - \bar{N}. \tag{2.41}$$

[1] The objects m_J and $m_{J,c}$ from Eq. 2.29 are quantitatively the same. However, due to the qualitative differences: m_J is moment derived from distribution X while $m_{J,c}$ is centralized moment estimated from data, the subscript "c" will be kept for central moments.

The obtained connected and shot noise - corrected moments k_J quantify the departures from Gaussianity and are model-corrected to the shot noise. However, these are the central moments of the density while averaged correlation functions $\bar{\xi}_J$ are by definition 2.16 moments of the density contrast. Therefore to compute $\bar{\xi}_J$, one can use the expanded form 2.28. Then one have:

$$\bar{\xi}_J = \frac{k_J}{\bar{N}^J}. \tag{2.42}$$

Such computation is being performed for counts in spheres of radii R (Eq. 2.27), thus while repeating this method for different sizes of spheres, one obtains averaged correlation function as a function of scale, $\bar{\xi}_J(R)$.

2.3.2 Counts and pixelization

Depending on the size of simulation box or catalog of observed galaxies, one needs to compute $\bar{\xi}_J$ for different ranges of scales. The lower range is limited by the scales where Poisson correction is no longer applicable, what depends on small $\bar{N} < 1$ and sampling. The upper limit is restricted by the amount of information that one can acquire from n_C counts, since the larger is the scale R, the less spheres can cover the full catalog. At some scale the number of spheres containing full information becomes insufficient to provide statistically representative sample [176]. In practice the common upper range used is 1/10 or 1/20 of simulation box length.

To obtain relevant information from $\bar{\xi}_J(R)$ showing the trends with respect to scale, one needs to consider more than one scale. Additionally, at each scale the number of spheres could vary from thousands to even tens of millions. One usually needs to consider also different catalogs, models or realizations. Performing the counts for catalog with N_{all} objects in standard way, i.e. verifying for every object if it belongs to given sphere, would require $n_o = n_c \cdot N_{all}$ operations for just one scale. Instead, one can use so-called pixelization of catalog, i.e. grouping the objects in voxels with defined shapes and sizes. While verifying which objects are inside the sphere, one then rejects entire pixels which do not cross the sphere, rejecting simultaneously all objects within that pixels instead of inspecting them one-by-one. The commonly used pixelization concerns the spherical analysis on sky and is performed with *HEALPIX* software [177]. However in some specific cases and for three-dimensional problems the standard cubic pixelization is required.

While the aim is to minimize the number of operations, one needs to carefully adjust the pixel size L_{PIX}. For a basic case of spherical counts in simulation box with side L_{BOX}, n_o depends on L_{PIX}/R, N_{all} and L_{BOX}. Since counting the objects separately refers only

to pixels crossing partially the sphere, the number of operations[2] is:

$$n_o = \left[\left(\frac{L_{BOX}}{L_{PIX}}\right)^3 - \eta\left(\frac{R}{L_{PIX}}\right)\right] + \eta\left(\frac{R}{L_{PIX}}\right) \cdot \left(\frac{L_{BOX}}{L_{PIX}}\right)^3 \cdot N_{all}, \tag{2.43}$$

where $\eta(R/L_{PIX})$ is average number of pixels partially crossing randomly positioned sphere of radius R. The term in square bracket describes number of pixels being either fully inside or outside of the sphere. For very small L_{PIX} this contribution dominates since one has to either reject or add large number of small pixels what makes the pixelization ineffective. In the opposite situation where L_{PIX} becomes large, the second part dominates meaning that the number of objects needed to count one-by one becomes large, since pixels crossing the sphere occupy large volume. Therefore finding the optimal pixelization is crucial for computing averaged correlation functions. In practice it can be done separately for each given geometry of the problem.

2.3.3 Alternative averaged correlation functions

Similarly to N-point correlation functions, their averaged alternatives $\bar{\xi}_J$ can also be explored in different forms. In cases of angular catalogs it is convenient to work with angular area-averaged correlation functions $\bar{w}_J(\theta)$ [167]. The entire computational procedure is identical as for three-dimensional equivalent except of that the counts are performed within circles positioned on the sky. Additionally, the study is then performed for projected density field δ_p related with three-dimensional density ρ by:

$$\rho_p(\vec{\gamma})d\Omega = d\Omega \int_0^\infty dr\ r^2 F(r) \frac{\rho(r\vec{\gamma})}{\bar{\rho}}, \tag{2.44}$$

where γ is unit vector pointing from observer towards $\vec{\Omega}$ on the sky and $F(r)$ is the selection function which defines completeness of the data in catalog.

The angular averaged correlation function is useful if the data has large uncertainties of distance measurements, such as in photometric surveys. For mock catalogs or observational data with more precise distances one can perform three-dimensional study of redshift-space effects with two-point correlation of a form $\xi(r_p, \pi)$ by splitting inter-pairs separations into components perpendiculars and parallel to line of sight. For volume-averaged correlation functions such study can be made using non-spherical window function. The applicable modification has to acknowledge the asymmetry between both directions, i.e. parallel and perpendicular to line of sight. In the case of distant observer approximation, the simplest shape is either a cylinder with radius $r_\perp$ and length $r_\parallel$ elongated along LOS, or an ellipsoid with semi-axes of $\{r_\perp, r_\perp, r_\parallel\}$ where axis with length $r_\parallel$ is elongated along LOS and $r_\perp$ - perpendicular to it. While the former configuration

[2]Number of operations in this context is number of considerations whether single target (object or pixel) belongs to the sphere or not

seems slightly less complex, the ellipsoidal approach is more related to basic $\bar{\xi}_J$ which assumes spherical window function and does not impose sharp edges in kernel shape which are problematic in Fourier space. In contrast to spherical case, the ellipsoidal volume averaged correlation function depends on two scales which are the axes $r_\parallel$ and $r_\perp$. By orienting LOS along z axis, the ellipsoid window function is:

$$W_{r_\parallel,r_\perp}(\vec{r}-\vec{r'}) = \begin{cases} \frac{4}{3\pi r_\parallel r_\perp^2} \text{ if } \left(\frac{x-x'}{r_\perp}\right)^2 + \left(\frac{y-y'}{r_\perp}\right)^2 + \left(\frac{z-z'}{r_\parallel}\right)^2 \leq 1 \\ 0 \text{ otherwise,} \end{cases} \tag{2.45}$$

where $\vec{r} = [x, y, z]$ and $\vec{r'} = [x', y', z']$. This can be also related with spherical window function:

$$W_{r_\parallel,r_\perp}(\vec{r}-\vec{r'}) = \frac{1}{r_\parallel r_\perp^2} W_{R=1}\Big(\frac{x-x'}{r_\perp}, \frac{y-y'}{r_\perp}, \frac{z-z'}{r_\parallel}\Big). \tag{2.46}$$

Due to the scaling properties of Fourier transform, i.e. $F(g(x/a)) = a\ \tilde{g}(ka)$, the ellipsoid window function in Fourier space is:

$$W_{r_\parallel,r_\perp}(\vec{k}) = \frac{3}{k_r^3}(\sin(k_r) - k_r\cos(k_r)), \tag{2.47}$$

where $k_r = \sqrt{(r_\perp k_x)^2 + (r_\perp k_y)^2 + (r_\parallel k_z)^2}$. Since the only conceptual difference between spherical $\bar{\xi}_J(R)$ and ellipsoidal averaged correlation function is shape of the window kernel, one can implement the same analysis which led to the relation 2.25, but with discrimination between directions parallel and perpendicular to LOS. Taking advantage of the symmetry in azimuthal coordinate and expanding d^3k in cylindrical coordinates gives the expression for second order of ellipsoidal function:

$$\bar{\xi}_2(r_\parallel, r_\perp) = \frac{1}{(2\pi)^2}\int k_\rho dk_\rho dk_z\ P(k_\rho, k_z)\Big[W_{r_\parallel,r_\perp}(k_\rho, k_z)\Big]^2, \tag{2.48}$$

where $k_\rho^2 = k_x^2 + k_y^2$. Inserting the anisotropic power spectrum with redshift-space distortions from relation 2.3 or more nonlinear approximations [178] allow for studying the influences of bias and growth factor on ellipsoidal correlation function of the cosmic web.

While the analytical computations for higher orders of $\bar{\xi}_J(r_\parallel, r_\perp)$ become sophisticated, similarly to spherical case the numerical cost of estimating such function on simulated or observed data is relatively low. This allows for studying redshift-space effects in higher orders with additional degree of freedom. The richness of new [179] and upcoming cosmological data [180, 181, 182] combined with high variety of tools for statistical analysis of the cosmic web thus necessitate the development of cosmological simulations.

Chapter 3

Paper I: Angular clustering in ΛCDM and extended gravity

3.1 Introduction

Gravity is a dominant force shaping the structure formation on cosmological scales. The standard theory - General Relativity has not yet been robustly tested at cosmic distance ranges. The common approach is to extrapolate GR into largest scales assuming that it properly describes gravity there. To verify if such assumption is valid and better understand the physics governing the evolution of cosmic web, one need to reliably constrain the gravity at cosmological distances.

Gravity shapes the initial density field imposing non-Gaussianities. These can be quantified with moments-based averaged correlation functions $\bar{\xi}_J$ and cumulants s_J, which hence consist a powerful tool for testing gravity at cosmic scales. In this part we compare standard ΛCDM model assuming General Relativity, with two families of extended gravity: $f(R)$ and nDGP. The identification of potential extended gravity signals (EG signals), i.e. discrepancies between standard model and extended gravity can be performed by confronting clustering measurements from simulation, using $\bar{\xi}_J$ and s_J. To mimic the observations, instead of working with raw simulation boxes it is then motivated to construct light cones. In context of redshift measurements, the observations are either photometric or spectroscopic. The former ones are characterized by much larger sky coverage and depth, what comes with the price of larger redshift uncertainties than for spectroscopy. While we work with statistics, for the sake of future observational comparisons we focus on angular measurements. Despite of not containing valuable redshift-space signals, angular clustering still encode information valuable for discriminating between gravity scenarios. In this work, the angular averaged correlation function and angular cumulants are denoted by $W_J(\theta)$ and $S_J(\theta)$, respectively.

Galaxy and halo angular clustering in ΛCDM and modified gravity cosmologies

Paweł Drozda,[*] Wojciech A. Hellwing,[†] and Maciej Bilicki[‡]
Center for Theoretical Physics, Polish Academy of Sciences, Aleja Lotników 32/46, 02-668 Warsaw, Poland

(Received 13 April 2022; accepted 21 June 2022; published 5 August 2022; corrected 17 August 2022)

Using a suite of N-body simulations, we study the angular clustering of galaxies, halos, and dark matter in Λ cold dark matter and modified gravity (MG) scenarios. We consider two general categories of such MG models, one is the $f(R)$ gravity, and the other is the normal branch of the Dvali-Gabadadze-Porrati brane world. To measure angular clustering we construct a set of observer-frame light cones and resulting mock sky catalogs. We focus on the area-averaged angular correlation functions W_J, and the associated reduced cumulants $S_J \equiv W_J/W_2^{(J-1)}$, and robustly measure them up to the ninth order using counts in cells. We find that $0.15 < z < 0.3$ is the optimal redshift range to maximize the MG signal in our light cones. Analyzing various scales for the two types of statistics, we identify up to 20% relative departures in MG measurements from general relativity (GR), with varying signal significance. For the case of halos and galaxies, we find that third-order statistics offer the most sensitive probe of the different structure formation scenarios, with both W_3 and the reduced skewness S_3 reaching from 2σ to 4σ significance at angular scales $\theta \sim 0.13°$. The MG clustering of the smooth dark matter field is characterized by even stronger deviations ($\gtrsim 5\sigma$) from GR, albeit at a bit smaller scales of $\theta \sim 0.08°$, where baryonic physics is already important. Finally, we stress that our mock halo and galaxy catalogs are characterized by rather low surface number densities when compared to existing and forthcoming state-of-the-art photometric surveys. This opens up exciting potential for testing GR and MG using angular clustering in future applications, with even higher precision and significance than reported here.

DOI: 10.1103/PhysRevD.106.043513

I. INTRODUCTION

One of the greatest accomplishments of modern cosmology is the formulation of the concordance standard cosmological model, the so-called Lambda cold dark matter (ΛCDM). This is a phenomenological model assuming that around 30% of the present-day Universe energy density is in the form of nonrelativistic matter (baryonic and dark) and the remaining 70% is attributed to the "dark energy," an exotic phenomenon propelling the late-time accelerated expansion of the Universe. ΛCDM is a simple model with six free parameters that is able to pass successfully many stringent observational tests, e.g., [1–4].

One of the core predictions of ΛCDM is that the cosmic large-scale structure (LSS) originated from gravitational instability acting on early matter distribution [5]. The widely accepted scenario assumes adiabatic Gaussian initial conditions (as supported by the cosmic microwave background measurements [1]), which were then reshaped due to the nonlinear gravitational evolution. The resulting LSS is organized into the so-called cosmic web [6], as traced by its main building blocks—luminous galaxies. Among the most striking features of this cosmic-web arrangement are the high anisotropy of the underlying density distribution (i.e., volume dominance of voids, mass dominance of filaments) and a scale-dependent clustering amplitude, observed also in the spatial distribution of galaxies [7–11].

The scale-dependent hierarchical matter and galaxy clustering is one of the most striking manifestations of the gravitational instability paradigm [5].

These unique and characteristic features of the LSS have been extensively employed as powerful probes of the standard cosmological model and its core assumptions. In the past few decades, galaxy photometric and spectroscopic catalogs have been growing both in volume as well as in quality of the data, which allowed for more and more precise tests of the fundamental components of ΛCDM.

Thanks to these growing observational data, presently the spatial distribution of galaxies and their time evolution can be readily used for performing stringent tests of the gravitational instability scenario. The latter is rooted in two core assumptions: the adiabatic Gaussian initial conditions and general relativity (GR) as an adequate and valid description of gravitational clustering on all scales and at all times. In this paper we explore the possibility of

[*]pdrozda@cft.edu.pl
[†]hellwing@cft.edu.pl
[‡]bilicki@cft.edu.pl

employing the properties of late-time matter and galaxy angular clustering for testing and differentiating GR and beyond-GR structure formation scenarios.

One of the simplest characteristics of matter clustering is the two-point correlation function (2PCF). This statistics is relatively easy to measure, which makes it a fundamental object commonly used in cosmology for quantifying matter and galaxy clustering [8,12–16]. In this context, we can recall that Wick's theorem for a Gaussian random field states that the first two order statistics (i.e., the mean and the variance) are sufficient to provide a complete statistical description of the field and clustering. In other words, for normal distribution all the higher-order odd moments vanish, and all the higher-order even moments are proportional to the variance.

However, the distribution function of a developed late-time LSS on scales below a few hundred megaparsecs deviates from a pure Gaussian distribution. Because of gravitational evolution, depending on the scale and the epoch involved, non-Gaussian features emerge, leading to the highly anisotropic and complex large-scale structure. Therefore, the variance and related two-point statistics no longer provide a sufficient description of the late-time LSS on scales important for galaxy formation and clustering. To infer additional information, one can resort to higher-order or beyond-two-point clustering statistics. Analogically to 2PCF, we can define N-point equivalents. However, already starting from $N = 3$, such correlation functions (CFs) become very expensive computationally and their use in cosmology has been so far limited to some special cases [17–21]—but see Refs. [22,23] for some recent developments.

An approach that is complementary to N-point statistics involves Nth-order central moments. These are volume-averaged versions of their full N-point counterparts, but are significantly easier to compute and model. At the same time, higher-order central moments, and the associated cumulants, contain extra information on the shape and asymmetry of the matter and galaxy distribution. Thus, these statistics bear information that is complementary to that carried by two-point statistics, e.g., [24–26].

Both two-point and higher-order clustering statistics have proven to be insightful and rich cosmological probes. It is, however, worth mentioning that the higher central moments are especially well suited for testing departures from standard structure formation scenarios. This is thanks to their increased sensitivity to non-Gaussian features of the distribution functions [5,27,28]. In general, the nonstandard gravitational instability models would involve some level of modified gravity (MG). In such scenarios one usually deals with low-energy effective scalar-tensor modifications to the Einstein-Hilbert action integral [29]. In that sense, these models are not new fundamental theories of gravity in their own right, but rather phenomenological manifestations (and parametrizations) of deeper underlying theories.

In this work we consider two such MG models, which constitute a good representative sample of a whole family of effective phenomenological modifications to gravity. The first consists of the so-called $f(R)$ framework [30,31], where in the gravity action integral the classical Ricci's scalar R is generalized to a functional $f(R)$ form [32]. The second family is the so-called normal branch of the Dvali-Gabadadze-Porrati (nDGP) brane world model [33,34], where gravity can propagate in the full five-dimensional space-time, while the standard elementary particle forces are confined to a four-dimensional subset space-time of a brane [35]. Both of these MG scenarios admit the action of the so-called fifth force on cosmological scales. This extra force is a manifestation of the additional scalar degrees of freedom of these models that, when coupled to the usual matter fields, affect the action of the gravitational instability and structure formation on galactic and intergalactic scales [36]. The stringent tests of GR in the strong-field regime [37–39] and in the weak field for the Solar System and our own Galaxy [40,41] impose rigorous constraints on the scales and times on which such a MG-induced fifth force is allowed to manifest itself. In order to pass these fifth force tests, viable MG models need to suppress propagation of the extra degrees of freedom in environments such as the Solar System or the Milky Way. The physical phenomena that lead to the fifth force suppression are called the screening mechanisms. Both the $f(R)$ and the nDGP theories naturally admit for such effects. Tuning the related theoretical parameters of these theories allows for finding solutions that simultaneously pass the local gravity tests and match the global ΛCDM expansion histories.

We focus on *angular* correlations, i.e., those projected along the line of sight. While the full 3D CFs give direct access to such cosmologically important effects as redshift-space distortions, they can only be observationally studied with sufficient accuracy using spectroscopic redshift catalogs. These often suffer from small-area coverage and/or sparse sampling and, even in the era of the forthcoming Dark Energy Spectroscopic Instrument data, will include only a small fraction of all observable galaxies. Another problem connected to spectroscopic surveys is that their analysis requires theoretical input on redshift- to real-space mapping, which is a strongly model-dependent procedure [42–44]. On the other hand, photometric (imaging) surveys typically offer a much better combination of depth, sky coverage, and completeness than the spectroscopic ones and, if accompanied by photometric redshifts, give the possibility to perform tomographic analyses of the density field. In view of future multibillion galaxy catalogs from such campaigns as the Vera Rubin Observatory Legacy Survey of Space and Time [45] or the Euclid space mission [46], it is timely to investigate possible MG signals from higher-order angular clustering.

This paper is structured as follows: In Sec. II, we describe gravity models and simulations used. The broader

picture of clustering statistics is contained within Sec. III. Afterward, in Sec. IV, we show the method of clustering calculations and introduce corresponding analytical predictions. Later, in Sec. V, we motivate our choice of redshift ranges that we use in the search for modified gravity signals. Then we present all our results, which are summarized and concluded upon in Sec. VI.

II. MODIFIED GRAVITY MODELS AND SIMULATIONS

In our study we will examine and compare higher-order angular clustering in different growth-of-structure scenarios. For this purpose we invoke the extended lensing physics with analytical ray tracing (ELEPHANT) numerical N-body simulations [47] performed using the ECOSMOG code [48]. They assume the evolution of $N_{\rm part} = 1024^3$ particles within an $L_{\rm box} = 1024/h$ Mpc sized box. Our fiducial, or baseline, model of choice is the GR-based flat ΛCDM model with WMAP 9-yr cosmology [49], with matter and dark energy density parameters $\Omega_m = 0.281$ and $\Omega_\Lambda = 0.719$, and the Hubble constant $H_0 = 100\,h\,{\rm km\,s^{-1}\,Mpc^{-1}}$ where $h = 0.697$.

On top of this background ΛCDM model, we consider two beyond-GR scenarios. The first of them—the Hu-Sawicki variant of $f(R)$ MG—introduces the fifth force, which is suppressed in dense environments thanks to the virtue of the chameleon mechanism [50,51]. Adopting a standard choice of the free parameters for this model, e.g., [30,44,52], we are left with only one variable to be set in order to characterize the late-time modifications to GR: the amplitude of the background scalar field at the present times, usually denoted as f_{R0}. Following the previous works that employ the ELEPHANT simulation suite [47,53], we label the two $f(R)$ variants used as F6 and F5, which corresponds to $f_{R0} = \{-10^{-6}, -10^{-5}\}$, respectively.

The second MG family—nDGP models—incorporate the Vainshtein screening mechanism [33,54] to suppress the fifth force in the vicinity of massive bodies. In the parametrization adopted here, the nDGP models can be also fully characterized by a single choice of the model physical parameter. This is the so-called crossing-over scale r_c, which depicts a characteristic scale where the gravity propagation starts to leak out to the fifth spatial dimension. Taking $c = 1$, we can fix our nDGP variants to have $H_0 r_c = \{5, 1\}$ Gpc$/h$, which we label as N5 and N1, accordingly.

We calculate angular counts for the projected dark matter (DM) density field from subsampled data, using only 0.1% from the each initial 1024^3 particle load. Such subsampling severely limits the spatial and angular resolution of the density fields, but this is needed to facilitate numerical calculations. Dark matter halos were extracted using the ROCKSTAR halo finder [55] and mock galaxy catalogs were generated with the halo occupation distribution (HOD) method in Ref. [47] using parameters from [56]. Unlike the dark matter particles, dark matter halos and mock galaxies are not subsampled, and their relatively low number density (see Sec. V A) is related to the very nature of the ELEPHANT catalogs.

For our analysis we employ five independent random phase realizations of initial conditions and take snapshots saved at $z = 0.0$, 0.3, 0.5, and 1 for further analysis.

In order to work in a sky-projected observer frame, we need to construct proper observer light cones from our snapshots. The redshift range we consider, i.e., $0 < z < 0.5$ for galaxies and DM particles, and $0 < z < 1$ for halos, corresponds to comoving scales that by far exceed the ELEPHANT simulation box size. To cope with that we locate the observer at the $\vec{r}_0 = (0, 0, 0)$ corner of the box and use the box replications method (see [57–59]) to build the light cones. For each snapshot, we copy the adequate box within ranges defined as half the comoving distance between the redshift of the current and each adjacent snapshot.

From our light cones we generate two-dimensional sky catalogs consisting of a series of ∼1567 deg^2 chunks, which would correspond to sky patches of sizes $40° \times 40°$ if centered on the equator. Each of our 2D sky catalogs is a sum of several separate sky chunks.[1] We found that for 15 chunks we already attain the maximum spatially independent catalog information, as measured by the catalog effective volume $V_{\rm eff}$.

This effective volume needs to be defined because, due to box replications, the catalogs contain many copies of the same structures. Thus, the total amount of independent cosmological information is always smaller than it normally would result from the actual light cone comoving volume. To get a total measure of unique (i.e., not cloned) volume, we use the catalog effective volume $V_{\rm eff}$. We define it as a sum over all the simulation box texels that are used at least once, divided by their total number inside the box. For more details on how $V_{\rm eff}$ is measured, see the Appendix.

Following Ref. [44], for our halo samples we consider only objects with $M_{vir} \geq 10^{12}\ M_\odot/h$. In Ref. [44] it was found that the abundance of less-massive halos is already affected by the mass resolution limit of the simulations. After this initial mass cut, with the same universal threshold for all the simulations runs, we employ secondary individual sample mass cuts. These are administered in such a fashion to obtain the same object number density within a given initial condition realization suite among all different physical models (i.e., F5, F6, GR, N1, and N5). Here randomly selected least-massive halos are trailed off until a given sample is reduced to the target number density. The latter is set by the lowest number density sample within a given ensemble. This is done to mimic a volume-limited

[1]Note that some of the sky chunks can partially overlap over the mock sky. For details, see the Appendix.

sample selection effect. For the mock galaxy sample, we applied an analogous operation. However, since for galaxies we do not have masses nor luminosities, we trail off all galaxies picked up randomly.

III. HIGHER-ORDER CLUSTERING

Our goal is to study clustering properties of matter, halo, and galaxy distributions over a large range of scales and epochs. Here, we will define the basic objects and methods of our clustering analysis. We will work either with physical ($\vec{r}$) or comoving ($\vec{x}$) distance units, where the usual relation is $\vec{r} = \vec{x}/(1+z)$. Adopting a standard notation, we will refer to comoving units as either h^{-1} Mpc or h^{-1} Gpc.

We define the standard density contrast that measures a local (i.e., at point $\vec{x}$) fluctuation of the density around a uniform background as

$$\delta_{3D}(\vec{x}) = \frac{\rho(\vec{x})}{\bar{\rho}} - 1, \tag{1}$$

where $\bar{\rho}$ is the background density. Converting the coordinates into sky frame $\vec{x} \to [r, \vec{\Omega}]$, where r is radial distance and $\vec{\Omega}$ represents a sky-pointing angular vector, we can obtain the projected density contrast,

$$\delta_{2D}(\vec{\Omega}) = \int_{r_{\min}}^{r_{\max}} dr F(r) r^2 \delta_{3D}([r, \vec{\Omega}]). \tag{2}$$

Here, the $r_{\min/\max}$ values stand for the distance ranges covered by a given survey and $F(r)$ is its selection function. From now on, for simplicity we will drop the "2D" subindex whenever referring to δ_{2D}.

The global properties of density fluctuations are encapsulated in their 1D probability distribution function (PDF), which can be estimated by averaging the fluctuations over many sky directions. Gravitational clustering moves the shape of the density PDF away from its initial Gaussian form [60]. All the gravity-induced nontrivial PDF shape deviations at a certain scale can be characterized by a hierarchy of the central moments. We will use the standard definition of a central moment of Jth order,

$$\mu_J = E[(\delta - E(\delta))^J], \tag{3}$$

where E is the expected value, δ is our random variable, and all the variables intrinsically depend on the angular scale θ. In our work, θ is a radius of a circle centered on a particular sky direction $\vec{\Omega}$. We average over many such circles to obtain our PDFs.

The central moments μ_J estimated for a given sky area at some angular scale θ can be considered as area averages of the full J-point angular clustering functions. They are related with J-point correlation functions by [7]

$$W_J(\theta) \equiv \frac{\mu_J}{\langle\delta\rangle^J} = \frac{1}{A}\int_\Omega d\Omega_1 \ldots d\Omega_J w_J(\theta_1, \ldots, \theta_J), \tag{4}$$

where $A = 2\pi(1-\cos\theta)$ is the sky area enclosed by angle θ.

The Jth-order moments can be readily estimated using the counts-in-cells method [25,61]. The moments are the ensemble average over all the circular cells (of intrinsic angular scale θ) cast over the whole area of interest on the sky Ω,

$$W_J(\theta) \equiv \langle \delta_\theta^J \rangle, \tag{5}$$

where δ_θ is a projected angular density fluctuation estimated at scale θ from angular counts.

IV. CLUSTERING AND MOMENTS OF COUNTS IN CELLS

We estimate the moments of the angular clustering using the commonly adopted method of counts in cells (CIC) [25]. We randomly place N_C circles of angular radius θ within the investigated sky area, making sure they are fully within the considered region. Those extending outside the footprint are ignored and replaced by new randomly drawn ones. Then, we count the objects found inside each circle. The Jth central moment of the CIC distribution is

$$m_J(\theta) = \frac{1}{N_{\rm tot}} \sum_{i=0}^{N_{\rm tot}} (N_i - \langle N \rangle)^J, \tag{6}$$

where N_i stands for the object count in the ith cell, $\langle N \rangle$ is the mean count over all the circles with a given radius θ, and $N_{\rm tot}$ is the total number of circles used. We choose $N_{\rm tot} \propto A_{sky}/(2\pi(1-\cos(\theta)))$ to scale as the number of independent circles that we can place within the analyzed sky area. Since we are interested in the specific shape departures from a normal distribution, we will work with the connected moments μ_J. That is, we subtract from the central moments the parts expected for a Gaussian PDF. The first few connected moments are

$$\begin{aligned} \mu_2 &= m_2, \\ \mu_3 &= m_3, \\ \mu_4 &= m_4 - 3m_2^2, \\ \mu_5 &= m_5 - 10 m_3 m_2. \end{aligned} \tag{7}$$

Since we will work with relatively sparse samples, the mean counts, especially at small θ, can become small and the impact of the shot noise will become significant. To reduce it, we follow the procedure of Ref. [25] and subtract from the connected moments the contribution expected from a Poisson distribution for a given mean count $\langle N \rangle$ [see Eq. (A6) therein].

The shot-noise correction is obtained by considering the contribution to the moments from a Poisson distribution with the same mean number of counts $\langle N\rangle$ as the studied sample. For reference we only recall the first few shot-noise corrected moments k_J,

$$\begin{aligned} k_2 &= \mu_2 - \langle N\rangle, \\ k_3 &= \mu_3 - 3k_2 - \langle N\rangle, \\ k_4 &= \mu_4 - 7k_2 - 6k_3 - \langle N\rangle, \\ k_5 &= \mu_5 - 15k_2 - 25k_3 - 10k_4 - \langle N\rangle. \end{aligned} \tag{8}$$

Finally, the Jth-order corrected and averaged correlation function can be written as

$$W_J = \frac{k_J}{\langle N\rangle^J}, \tag{9}$$

and the rescaled cumulants, or more commonly dubbed in cosmology as "hierarchical amplitudes," will be

$$S_J = \frac{W_J}{W_2^{J-1}} \equiv \frac{W_J}{\sigma^{2J-2}}. \tag{10}$$

A. Signal significance

Following a standard approach, we estimate the error on the quantities given by Eqs. (9) and (10) as the variance around the mean, obtained as the ensemble average over all equivalent realizations of a given dataset (i.e., a light cone). In practice, we will be more interested in assessing the differences between each MG model and the fiducial ΛCDM case. This is measured by the relative difference of paired observables always taken with respect to the GR case. Both the fiducial GR and any given MG model sample will be characterized by their own individual variance. For that reason, comparing clustering moments of different models with different individual variances might be difficult and not intuitive. To foster a more natural and easy-to-interpret comparison, we will use the signal significance parameter ψ, defined as

$$\psi_J = \frac{X - Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}}, \tag{11}$$

where X and Y are measurements and $\sigma_{X,Y}$ are their respective uncertainties. Here, J indicates the order of the X and Y statistics used to calculate the significance. So, for example, ψ_3 can indicate either that W_3 or S_3 was used. In this work, we will always take Y as GR and any given MG model is taken as X. The significance ψ gives a simple notion of the direction of the difference with respect to the fiducial case, preserving the sign of the difference, and automatically traces the significance of this difference due to the normalization factor in the denominator.

B. Perturbation theory predictions

Our main results in this work are based on the analysis of N-body simulations, which by design can probe deeply into the nonlinear regime of structure formation. However, it is very informative and beneficial to provide also analytical predictions with which the numerical results can be gauged. For that purpose, we use calculations based on the weakly nonlinear perturbation theory (PT), which yield predictions for low-order moments. These are obtained by integrating over the matter power spectrum with appropriate window and selection functions (see [62,63]).

The second moment is given by

$$W_2(\theta) = \frac{1}{2\pi}\int_{R_{\rm min}}^{R_{\rm max}} r^4 F^2(r) dr \int_0^\infty k P(k) W_{2D}^2(k\theta r) dk, \tag{12}$$

where $P(k)$ is a given model power spectrum, and

$$W_{2D}(k) = 2\frac{J_1(k)}{k} \tag{13}$$

is the window function for which we take a circular top hat in the Fourier space with the first-order spherical Bessel function J_1. The $R_{\rm min\,/\,max}$ stand for catalog comoving distance ranges and $F(r)$ is the radial selection function normalized in a such way that

$$\int_{R_{\rm min}}^{R_{\rm max}} r^2 F(r) dr = 1. \tag{14}$$

In our catalogs, we do not use specific selections mimicking the observations, hence the selection function becomes

$$F(r) = \frac{3}{R_{\rm max}^3 - R_{\rm min}^3} = \text{const.} \tag{15}$$

For the third order, we have

$$\begin{aligned} W_3(\theta) = 6\frac{\theta^{-4}}{(2\pi)^2}\int_{R_{\rm min}}^{R^{\rm max}} r^2 F^3(r) dr \int_0^\infty q W_{2D}^2(q) P(k) dq \\ \times \left[\frac{5}{14}\int_0^\infty q W_{2D}^2(q) P(k) dq \right. \\ \left. - \frac{1}{4}\int_0^\infty q^2 W_{2D}^2(q)\frac{dP(k)}{dq} dq\right], \end{aligned} \tag{16}$$

where $q = k\theta r$.

The formulas for higher orders become longer and recurrently more involved (see, e.g., [7]). Thus, we opt to stop at the third order only, since detailed tests of PT are not our aim here, and we will use these predictions for approximate trend comparisons only. In yielding our PT predictions, we have used both linear and nonlinear (i.e., the Halofit [64]) power spectra models computed with

CAMB software [65] taken at the effective catalog redshift $z_{\rm eff} = 0.242$.

V. RESULTS

Now we are ready to present and investigate the results covering the angular clustering, distributions of the counts in cells, and the associated moments. All the presented results concern dark matter, halo, and galaxy samples extracted from the same depth ELEPHANT-based light cones, as described above.

A. Finding optimal light cone depth

Previous studies of clustering have indicated that, in the case of MG models considered here, the magnitude of the deviations from the GR-fiducial case is changing with the cosmic time, in a nonmonotonic way (see, e.g., [66–68]). In the case of single redshift snapshots, commonly used in the distant observer approximation, it is straightforward to depict a redshift with maximal deviation from GR. However, in our case, in the light cone projection, clustering information from the redshift range of the whole light cone is entangled. We want, therefore, to find an optimal redshift range for a light cone galaxy catalog that maximizes the relative deviation from GR of the MG clustering signal as far as the moments are concerned.

Considering the redshift ranges we have for our halo ($0 < z < 1$) and galaxy samples ($0 < z < 0.5$), we create a 4×4 grid of catalogs with varying minimum redshift $z_{\rm min}$ and a defined thickness $\Delta z = z_{\rm max} - z_{\rm min}$. For the optimization procedure, we define our merit parameter to be the ψ_3 estimator as defined later in Eq. (11), which measures the relative amplitude of the deviation between a given MG model and the GR case.

Optimization performed over all models and scales would be computationally very expensive. However, since our goal here is just to find an approximated optimal redshift light cone range, we opt to focus on only two MG models, N1 and F5, and only one angular scale of $\theta_{\rm opt} = 0.08°$. These two MG variants are characterized by the largest difference in the linear growth rate with respect to GR. The $\theta_{\rm opt}$ value was selected as a reasonable compromise between the nonlinear regime, where the clustering deviations usually are the largest, and at the same time a scale where the shot noise and simulation resolution effects are not too severe yet.

We have found that the redshift range[2] $0.15 < z < 0.3$ maximizes the MG signal for halos, while the range $0.15 < z < 0.25$ is optimal for the galaxies. Considering the fact that halo light cones provide wider redshift ranges in comparison with galaxies and the halos provide stronger signals than galaxies, we kept $0.15 < z < 0.3$ as the best redshift ranges for all the catalogs. The final data samples with the imposed redshift cuts have the following characteristic projected number densities:

(i) DM particles, ~ 51 deg^{-2};
(ii) halos, ~ 40 deg^{-2};
(iii) galaxies, ~ 15 deg^{-2}.

Given the effective depth of our light cones, the spatial resolution of the ELEPHANT suite, and taking into account that we consider only resolved halos, we can estimate that our catalogs will be spatially resolved down to ~ 0.5–$1\ h^{-1}$ Mpc [53]. Within the redshift range we use, this sets the minimum angular scales that we can consider as resolved to be $\theta_{res} \approx 0.05°$.

B. Probability density functions

We begin by showing in Fig. 1 an excerpt of angular counts-in-cells distributions for dark matter (top panel), halos (middle panel), and galaxies (bottom panel). These example PDFs are measured at the angular scale of $\theta = 0.3°$. Considering the median redshift of our light cones (i.e., $z \sim 0.242$), this scale corresponds to a projected comoving separation of $R = 3.6\ h^{-1}$ Mpc. We pick this scale since it constitutes a reasonable compromise between the scales where the influence of both the cosmic variance and the sparse sampling remains limited. For clarity we

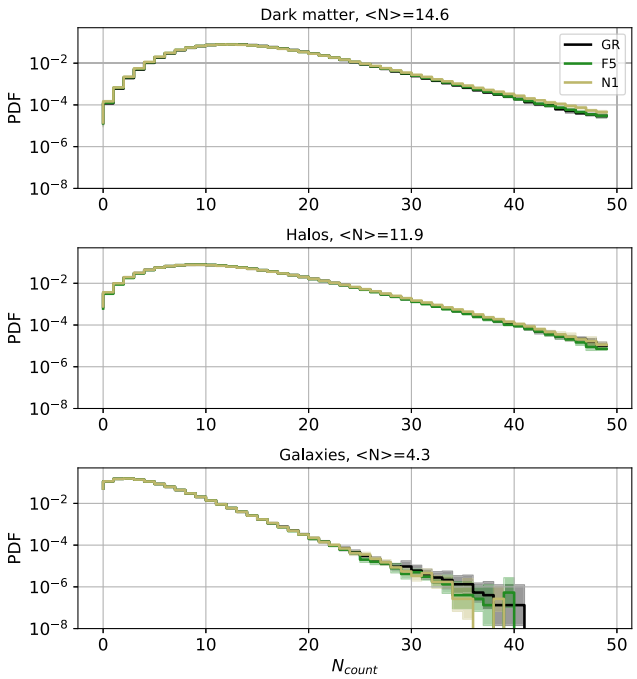

FIG. 1. Probability density functions of the counts in cells from circles of 0.3° radius, corresponding to $\sim 3.6\ h^{-1}$ Mpc physical scales at the effective redshift of our light cones ($z_{\rm eff} \sim 0.24$). Shown from top to bottom are the results for dark matter, halos, and galaxies, derived from the ELEPHANT suite for three gravity models indicated in the legend. The numbers in the headers indicate mean counts for each case.

[2]The exact range is $0.1525 < z < 0.3025$.

show only two modified gravity variants, N1 and F5, on top of the GR case.

This example already brings a few interesting observations. First, we can infer that dropping angular object number density drives the resulting PDF away from a Gaussian and more toward a Poissonian distribution, thus highlighting the importance of the shot-noise corrections for samples with small mean number counts $\langle N \rangle$. The second noticeable feature is that for all three samples (i.e., galaxies, halos, and dark matter) different models arrive at very similar mean number counts. Furthermore, the associated variances, or the distribution widths, are also comparable. Only, when we move away from the PDF's centers, toward the tails, the differences between GR and MG models become more and more appreciable. This is a clear illustration of how important it is to go beyond central and second-order moments, which are much more sensitive to information contained in the distributions tails (i.e., the PDF asymmetry and overall shape deviations).

C. Dark Matter

We start our analysis of angular clustering by looking at the projected dark matter density field. Although this is not directly observable, there is a strong connection between the underlying smooth projected dark matter distribution and quantities accessible via gravitational lensing effects. These, among others, include convergence and shear power spectra [62].

There are also tomographic techniques to obtain reconstructed 3D dark matter distribution on large scales [69–71]. Here, we will focus on simple sky-projected dark matter density fields measured from a subsampled dark matter N-body particle distributions. These distributions and their moments are not directly connected to observations, but provide a very good test bed. Moreover, analyzing the dark matter clustering will enable us to compare our CIC results with PT predictions as given in Sec. IV B and will provide additional physical insight about the higher-order angular clustering in MG in linear and nonlinear regimes.

We begin by comparing the N-body CIC moments with the PT predictions. This will serve both as a useful test of our estimators, as well as the indicator of scales where the transition between the nonlinear and weakly nonlinear angular clustering regimes occurs. In Fig. 2 we show the first two moments $W_2(\theta)$ and $W_3(\theta)$ (upper panel) and the reduced skewness $S_3(\theta)$ (bottom panel). The continuous lines indicate our N-body results; the dotted and dashed curves are the PT prediction obtained using the linear (dotted) and Halofit [64] (dashed) dark matter power spectra. The shaded regions indicate 1σ scatter from the simulation ensemble mean.

The PT predictions agree very well with the N-body results at large scales, $\theta \gtrsim 0.2°$. For smaller angles, the simulation results quickly surpass the values based on the

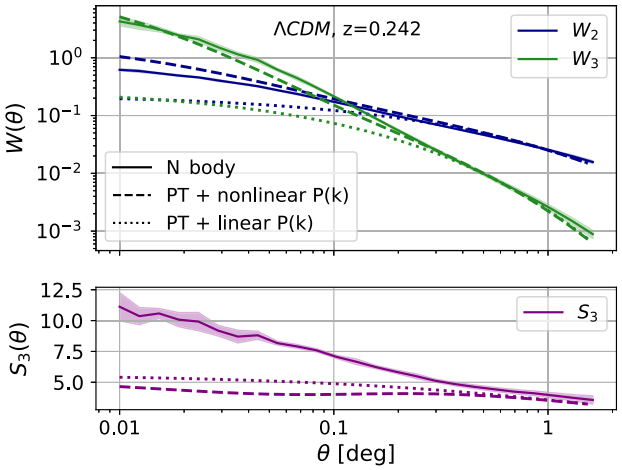

FIG. 2. First two reduced moments (top) and reduced skewness (bottom) of angular clustering calculated in the ΛCDM model at an effective redshift of $z = 0.242$. The solid lines with shaded error ranges show N-body simulation results, while the dotted (dashed) lines illustrate the perturbation theory predictions derived using linear (nonlinear) power spectra as detailed in Sec. IV B.

linear theory $P(k)$. Interestingly, using Halofit as the nonlinear power spectrum model readjusts the PT predictions, making them follow the N-body lines much more closely, extending good PT accuracy down to scales of $\theta \sim 0.04°$. However, even if the PT predictions for W_2 and W_3 separately look reasonable, their combination into the reduced skewness S_3 accumulates the deviation of each individual moment. This is clearly manifested in the bottom panel of Fig. 2, where both the PT-based forecasts fail and underpredict the skewness dramatically for $\theta \lesssim 0.3°$. To get a better prediction here, one would need to call for higher-order PT templates (see, e.g., [72–74]). This test indicates that our N-body results capture well both the linear and nonlinear regimes. In addition, all the significant differences that we might find in GR vs MG clustering above $\theta \simeq 0.3°$ could be highlighted in future analyses using weakly nonlinear PT predictions as detailed in Sec. IV B.

Before we move to the main part of the analysis, with the focus on clustering of halos and galaxies, we take a quick look at the angular variance and reduced skewness of the dark matter projected density field in our models, shown, respectively, in Figs. 3 and 4. To facilitate easier comparison, we group the MG models' families, keeping the $f(R)$ models in the left-hand side panels and nDGP in the right-hand side. The shaded areas for $\theta < 0.05°$ indicate the angles lower than the angular convergence scale of ELEPHANT.

Looking first at the angular variance (W_2), we can already make a number of very interesting observations. First, F6 seems to accommodate only minute differences from GR and is virtually indistinguishable from it for all

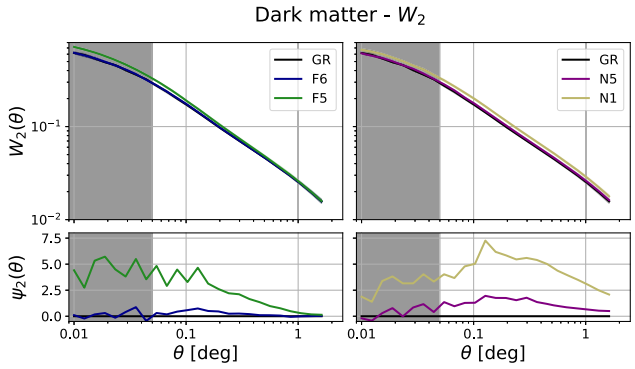

FIG. 3. Two-point area-averaged angular correlation function of dark matter particles for ΛCDM (black lines) as compared to two MG scenarios: $f(R)$ in the left-hand column and nDGP in the right-hand one. The MG models have two variants each, as indicated in the legends. Top: the correlation function. Bottom: illustrate the significance of departure in the MG models from the fiducial GR scenario, as defined in Eq. (11). The shaded regions cover the $\theta < 0.05°$ range, which we do not use for inferring the signal significance due to the limitations of the ELEPHANT simulations.

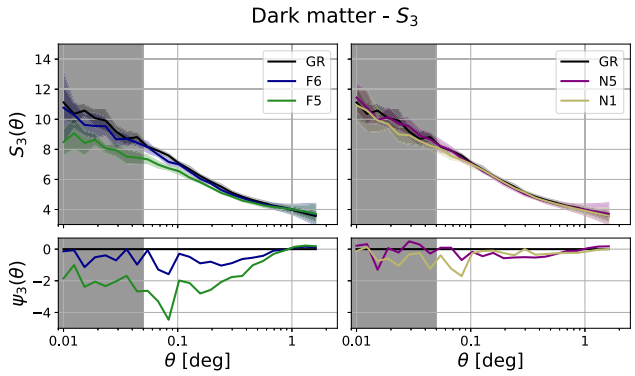

FIG. 4. Similar to Fig. 3, but for the reduced skewness.

scales. Second, all the remaining variants, i.e., F5, N1, and N5, exhibit in $W_2(\theta)$ some unique scale-dependent patterns of their significance signals. The departure of the F5 signal from GR saturates at around $\theta \lesssim 0.1°$ and then decreases to converge to GR at $\theta \simeq 1°$. For both the nDGP variants we notice a similar decrease in the signal, but now on two sides from the maximal deviation scale of $\theta \simeq 0.2°$. Thus, the departure of W_2 in nDGP from GR assumes a peaklike shape in angular scaling. Moreover, here also the N5 model, generally weakly departing from GR, fosters significant deviations from the GR case, in contrast to its $f(R)$ cousin, F6. This is a new result, as such features have not been found in the earlier 3D dark matter clustering studies (see, e.g., [67,68]).

The dark matter results for S_3 feature a bit different picture. Here, the noise and errors on both W_2 and W_3 moments are amplified, and the resulting signal becomes generally weaker and more erratic. While for N1 and N5 the significance is severely reduced, F5 interestingly still reaches $|\psi_3| \simeq 4$ at $\theta \sim 0.1°$. Interestingly enough, for all the MG models, their skewness takes values lower than in the fiducial GR, which indicates the known fact that the relative asymmetry of MG evolved density distributions is lower than in ΛCDM (see again [67,68]).

D. Halos and galaxies

The angular clustering patterns we observed for the dark matter, albeit very interesting, cannot be easily nor directly translated into expectations for any observables that we can extract from galaxy surveys. Some potential implications for weak gravitational lensing could be drawn, but we keep this discussion for later (see Sec. VI). On the other hand, the angular counts and related statistics of discrete objects, such as halos and galaxies, have much more direct and straightforward connection and interpretation in the context of existing observational catalogs. Thus, we now move to the main part of our analysis and take a look at the hierarchical clustering of halos and galaxies.

In Fig. 5 we summarize the third-order statistics (W_3, columns to the left; S_3, to the right) for halos (top blocks of panels) and galaxies (bottom blocks). We focus here on the third-order statistics, since a detailed analysis for all higher-order moments would be unfeasible, and additionally the moments higher than the fourth contain similar information to orders 3 and 4. We choose not to present here and discuss separately the case for the angular variance W_2, as its amplitude, in general, is fully degenerate with the first-order angular bias parameter b_θ,

$$W_2^{h,g}(\theta) = (b_\theta^{h,g})^2 W_2^{\mathrm{DM}}(\theta). \tag{17}$$

Here, (h, g) stands for halos and galaxies, respectively. In Ref. [44] it has been shown that the second-order clustering statistics in MG are affected by this bias degeneracy. For that reason, higher-order moments and their combinations (like skewness and kurtosis) may contain a more genuine MG signal. The reason is that, to the first order, the bias degeneracy is reduced for them (see also [53,75,76]).

Let us first discuss the MG signal for the halo population.

As we have verified for the case of DM density, the significance of the departure from the GR prediction is higher for W_3 alone, compared to the skewness. Again, this is expected given the standard error propagation properties. Focusing on the converged scales, i.e., $\theta \gtrsim 0.05°$, we can observe a number of interesting features. First, the weaker F6 variant is characterized by stronger deviations from the GR case than F5. This might appear as a surprise at first, but can be explained. Although, in terms of the background field value, the F6 variant should experience weaker scalar-field effects than F5, the former model is actually inherently more nonlinear than the latter, in terms of the chameleon screening behavior. This property manifests itself especially for the less

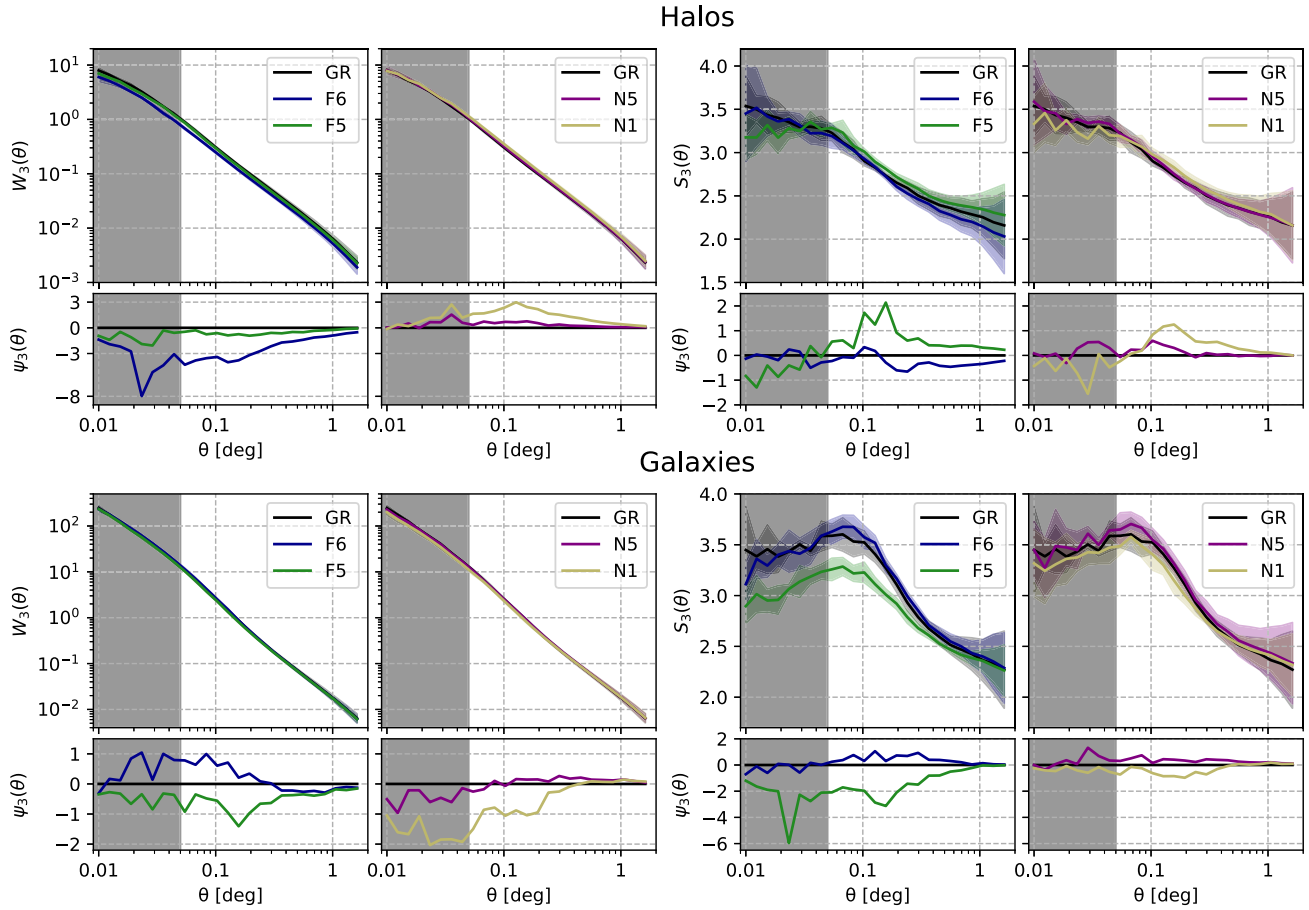

FIG. 5. Three-point-averaged angular correlation function (left columns to the left) and reduced skewness (columns to the right) for halos (top blocks) and galaxies (bottom blocks) extracted from ELEPHANT simulations for ΛCDM and two MG scenarios. See caption of Fig. 3 for further details.

massive, smaller halos (which actually dominate the sample). This phenomenon was to some extent already encountered and studied in Ref. [77]. This trend is reversed when we look at the skewness, where again F6 is marginally consistent with GR for all scales, while F5 also shows small deviations, albeit larger than F6.

In the case of the nDGP models, we observe behavior that qualitatively agrees with what we have seen already for the variance in the dark matter field. The N1 deviation again assumes a peaklike scale dependence, with the maximum significance attained at $\theta \simeq 0.1°$, both in W_3 and S_3. The departure signal of the N5 variant stays weak and insignificant for all scales. The differences in scale and magnitude dependence between the $f(R)$ and nDGP models clearly indicate that the different physics of their screening mechanisms manifests itself to some degree in different halo clustering.

For the galaxy population, we encounter an interestingly different picture. Here we need to recall that the ELEPHANT mock galaxy catalogs were constructed using HOD parameters tuned for each MG model separately, so the resulting distributions have the same (to within 1%–2%) projected 2PCFs. Since the real-space 3D clustering was anchored, the higher-order moments will carry here the genuine residual MG signals. This property of the mocks resulted in quite interesting trends we can single out in the galaxy statistics. First, we now see that signals for the F5 and F6 models have opposite signs of the effect on $W_3(\theta)$, but now also F5 reaches comparable deviations in the magnitude as F6. For nDGP, we no longer observe a peaklike shape of the scale dependence, but instead a saturation of significance at the level of ~ 1 for N1 at $\theta \lesssim 0.2°$. For N1, the effect is also opposite to the one we have just noted for halos. The disappearance of peaklike scale dependence in galaxies for the N1 scenario is related to the fact that our halo catalogs consist of only central halos, while for galaxies we also consider satellites. A related issue has been already discussed in Ref. [53].

In the case of the skewness, the effect is also flipped (when compared to W_3) for F5 and N1, with a clear difference that now for F5 the maximum signal is marginally stronger than for the halo population.

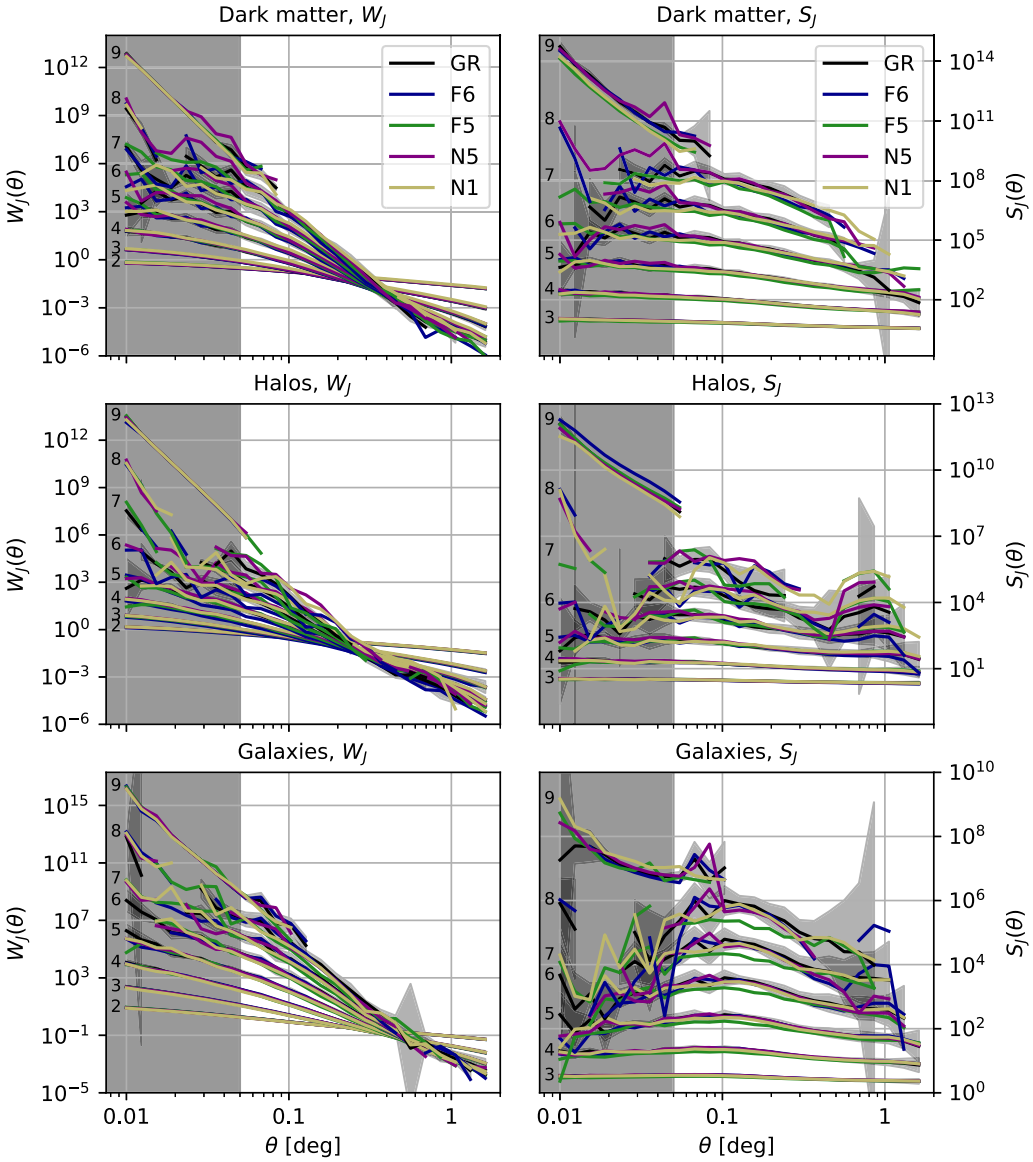

FIG. 6. Clustering of all orders considered in this work. From top to bottom, we present the results for dark matter, halos, and galaxies. The left column shows area-averaged correlation functions W_J for $2 \leq J \leq 9$; right column includes hierarchical amplitudes S_J for $3 \leq J \leq 9$. The particular orders are indicated with the numbered labels on the left-hand side of each panel. Lines of different colors correspond to the gravity models as in the legend. Light gray bands illustrate the errors in the GR case; those for MG are comparable. The dark gray vertical bands cover the angular scales, which we do not use for inferring the MG signal.

The remaining higher-order moments and reduced cumulants up to $J = 9$ reveal a qualitatively similar picture to what we have just shown and discussed for the third order. The general trends are continued, but the scatter in these quantities and associated errors grow quickly with the order. To obtain a general impression of the trends, we show all the collected W_J's and S_J's (for $J = 2, \ldots, 9$) for dark matter, halos, and galaxies in Fig. 6. Here the top row is for the smooth DM density, the central one for halos, and the bottom one for galaxies. Columnwise, the plots are organized as the area-averaged correlation functions W_J's (left-hand column) and reduced cumulants S_J's (right-hand column). The particular orders are organized as indicated by the labels, while the lines are colored according to gravity models in the same way as in Figs. 3–5.

For the cumulants, but also to some extent for the moments, we can observe that the relative errors explode at two regimes: for small and large angles. This can be easily attributed to the shot noise at $\theta \lesssim 0.01°$ scale and to finite catalog size (i.e., cosmic variance) influencing the $\theta \gtrsim 2°$ range. Higher moments are more sensitive to these effects, due to growing powers in Eq. (6), which multiplies

any initial error on the mean counts at a given scale $\langle N\rangle_\theta$. Such conditions introduce significant heteroscedasticity into the clustering measurements. The partly missing measurements in Fig. 6, especially for the eighth- and ninth-order statistics, are due to negative values that cannot be represented in the logarithmic scaling.

The precision of our measurements scales roughly as a square root of object number density (e.g., the relative errors are ∼1.5–2 times bigger in the case of galaxies than for DM pseudoparticles catalog; see Sec. V A). This fact makes searching for MG signals among highest orders inaccessible using these catalogs, since one would need to greatly increase the sampling or to consider much larger sky coverage. Overall, we have found that only orders of up to $J \leq 5$ are useful for MG signal searches in our catalogs. Although the deviations from the fiducial GR case are growing with the order, the associated scatter grows even faster. Thus, for the light cone samples used here, significant signals can be extracted only from the statistics up to fifth order.

A noteworthy feature is that galaxies are characterized by larger values of CIC moments than halos and DM particles. This is a natural consequence of both biased structure formation and the HOD catalog construction method used (i.e., introduction of satellites for the galaxy sample).

One final note about the hierarchical clustering ratios, shown in the right-hand panels of Fig. 6, is that their amplitudes generally do not depend strongly on the scale. This is in agreement with both theoretical predictions, as well as with the analyses of observational data (e.g., [25,27,67,78]).

E. Summary

With nine orders of moments, eight of cumulants, and with three different samples of five various gravity models, we have dealt with massive and complex data describing angular clustering in our light cones. Detailed analysis of all potential deviations from GR and associated signal significance would be very industrious and at best cumbersome. Therefore, we chose to present only the most interesting findings from this large body of statistics and we summarize them here in a survey manner.

The most promising features of the MG angular clustering we analyzed are given in Table I. There we quote the absolute values of the deviation from GR significance, $|\psi|$ [Eq. (11)], alongside the characteristic angular scales at which they are noted, the statistics for which the signal is found, and the MG model they are reported for.

As we have already indicated, the sampling density of ELEPHANT allows us to reliably use up to fifth-order statistics, as far as a ratio of MG to GR is concerned. The higher orders become too noisy, and the associated ratios with respect to the GR case become scatter dominated. When we are concerned with the MG signal significance, the general trends are that the lower-order statistics (i.e., $J = 2, 3$) are favored over the higher ones. However, in just one case (F5 for galaxies) the fourth and fifth orders reach higher significance than the lower moments. Here, the signal significance is reaching $|\psi_{3,4,5}| \simeq \{1.4, 2.4, 2.5\}$ for the third, fourth, and fifth order, respectively. Interestingly, we find cases where even if some W_J's reach small values of $|\psi|$, the associated reduced cumulants S_J's can arrive at much larger significance. This indicates that the reduced cumulants, due to their unique intrinsic length (or variance) scaling, contain extra constraining information about the underlying structure formation models.

TABLE I. The most significant MG deviations from GR in angular clustering in the light cones studied in this work. For a given tracer, we provide the model-statistics pair (columns 2 and 3) that give the largest signal significance as listed in the fourth column. The fifth column indicates the angular scales at which this maximal signal appears.

Tracer	Model	Statistics	$\lvert\psi_{\max}\rvert$	$\theta(^\circ)$
Dark matter	F5	W_2	4.8	0.05
	F5	S_3	4.5	0.08
	N1	W_2	7.3	0.13
Halos	F6	W_3	4.0	0.13
	F5	S_3	2.1	0.16
	N1	W_3	3.0	0.13
Galaxies	F5	W_5	2.5	0.16
	F5	S_3	3.1	0.16

The statistics we measured for the nDGP models reveal that this MG family is characterized by much less significant deviations from the fiducial GR case than $f(R)$. Especially if we focus on the hierarchical amplitudes S_J's. This would suggest that departures from GR at higher orders do not differ significantly from the departures at $J = 2$. Indeed, for instance, considering DM particles we obtain $|\psi_{2,3}| \simeq \{7.3, 6.4\}$ for $N1$, while $F5$ provides $|\psi_{2,3}| \simeq \{4.8, 1.8\}$. Generally, we find that the hierarchical amplitudes of the galaxy samples offer a slightly better sensitivity to deviations from GR than halos and DM. In contrast, for the halo sample, we find that the area-averaged correlation functions W_J's seem to perform marginally better in differentiating the models.

Finally, our results facilitate a general trend, where the MG models with larger theoretical growth-rate departure from the ΛCDM case are characterized by stronger angular clustering deviations as well. There is one notable exception to that trend: the F6 model, whose halo sample offers larger $|\psi|$ than that of the F5 variant, while the latter has theoretically larger growth rate. This is not a complete surprise, however, as previous studies already found evidence that the F6 model can exhibit more nonlinear behavior than its F5 cousin, owing to the intrinsic nonlinear nature of the chameleon screening [53,77].

VI. CONCLUDING REMARKS

In this work we have studied angular clustering by analyzing the moments of the counts in cells for two

modified gravity scenarios. The literature offers many studies of three-dimensional clustering, including redshift-space distortion analyses for such beyond-GR scenarios [44,52,79–81]. So far, however, very little (to our best knowledge) was known about the properties of angular clustering in the light cone sky-projected density in such scenarios. Our work here is the first approach to remedy this lack. To this end, based on the ELEPHANT suite, we have designed a number of light cones containing dark matter, halo, and galaxy samples. We then proceeded to build an ensemble of effective sky catalogs that we have used as the main object for our analysis.

Below, we summarize and recapture the most important results:

(i) The PT predictions for the GR W_2 and W_3 moments offer a reasonable agreement with the N-body dark matter results for $\theta \geq 0.2°$, if only the linear-theory power spectrum was used. A much better agreement is obtained, down to smaller scales, when using a nonlinear Halofit model for the power spectrum. For that case, the PT and simulations agree down to $\theta \sim 0.05°$. We note, however, that the PT prediction for the reduced skewness S_3 is grossly underestimated when compared with N-body data for $\theta \leq 0.5°$.

(ii) Our light cone analysis yielded an optimal catalog depth for maximizing the departures from GR in the angular clustering. For the case of our simulations and models, this turned out to be $0.15 \leq z \leq 0.3$.

(iii) On various scales and for various statistics, we found up to 20% relative departures from GR. Our data do not allow for probing robustly the scales both smaller than $\theta < 0.05°$ and larger than $\theta > 1°$, where the effects of the shot noise and a limited catalog size, respectively, dominate.

(iv) The reduced skewness S_3 of the galaxy sample has proven to be especially sensitive statistics for the $f(R)$ family models.

(v) We found significant signals even in the catalogs with as low object number density as 15 deg^{-2} (for galaxies); this indicates an optimistic outlook for measuring MG signals in real angular galaxy data.

(vi) Hierarchy between the second and higher moments is preserved in all the structure formation scenarios, with no clear or dramatic changes in weak scale dependence of the reduced cumulants for all scenarios.

(vii) In our data, the modified gravity signals can be extracted from higher-order statistics up to order $J = 5$. In practice, we can expect that catalogs with better sampling, in terms of both sky coverage and surface number density, should allow measurements that cover even higher orders and larger scales.

Our main findings agree with the picture where the angular correlations, due to their intrinsic spatial-scale mixing, offer a unique specific window for clustering analysis, especially in the context of scale-dependent GR modifications. In general, the deviations from GR might not get as large for angular correlations as in the case of redshift-space distortions and 3D clustering. However, the projected counts could gain much in signal significance, if one could tap the rich potential of much denser sampling stemming from usually many times bigger volumes of photometric galaxy catalogs than for often sparsely sampled redshift surveys.

Taking into account that, due to the limitations of the simulation suite used, our galaxy and halo catalogs are characterized by much smaller object number densities when compared to existing and forthcoming imaging sky surveys [45,82–85], one can expect that all the relevant shot-noise and even cosmic-variance effects should be strongly suppressed in future analysis of observational data. Our mock galaxy catalogs contain only ∼15 objects per square degree. For comparison, depending on the chosen galaxy sample, the Dark Energy Survey provides $\sim 3.42 \times 10^5$ galaxies at ∼4200 deg^2 for the redshift range similar to ours, $0.15 < z < 0.35$ [86], or even $\sim 1.7 \times 10^6$ galaxies at $0.2 < z < 0.35$. This gives from 5 up to 27 times higher galaxy surface density when confronted with our galaxy catalogs.

The fact that we have found significant, and therefore hopefully detectable, deviations from GR in our rather sparse mock galaxy catalogs, which have many times smaller object number density when compared to real galaxy samples, offers very promising prospects for testing GR and beyond-GR structure formation scenarios in the imaging data—an avenue that has not been exploited so far. However, to fully undertake such endeavor, one will need to account more robustly for the involved angular and redshift selection effects along with better galaxy population modeling in beyond GR. We leave this exciting undertaking for future work.

ACKNOWLEDGMENTS

The authors would like to acknowledge the fruitful discussions with Enrique Gaztañaga at early stages of this project. This work is supported by National Science Center, Poland under Agreements No. 2018/30/E/ST9/00698, No. 2018/31/G/ST9/03388, No. 2020/38/E/ST9/00395, and No. 2020/39/B/ST9/03494 and by the Polish Ministry of Science and Higher Education through Grant No. DIR/WK/2018/12. The research presented here also benefited from supercomputer calculations performed at the Interdisciplinary Center for Mathematical and Computational Modeling (University of Warsaw) under Projects No. GA67-17, No. GA67-16, and No. GB79-7.

APPENDIX: EFFECTIVE VOLUME

The so-called effective volume measures the amount of independent information in the catalog. As mentioned in

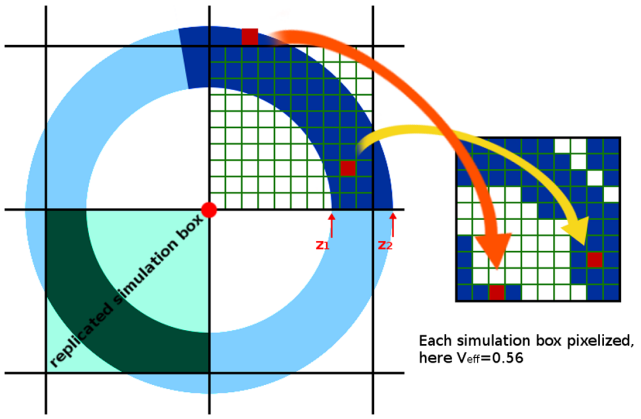

FIG. 7. Illustration of our method to calculate the effective volume. For clarity, we visualize only one light cone shell $[z_1, z_2]$, included within the catalog ranges $[z_{\rm min}, z_{\rm max}]$.

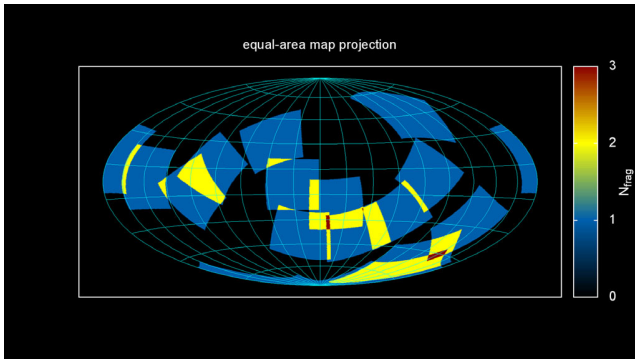

FIG. 8. Locations of sky fragments jointly maximizing the $V_{\rm eff}$. The $N_{\rm frag}$ value refers to the number of sky chunks overlapping at certain sky position.

Sec. II, calculating and maximizing the effective volume is critical for dealing with the issue of finite simulation box size.

We pixelize each replicated box (by dividing it into 30^3 cubic pixels) and then for each pixel we check whether its center belongs to the catalog. If this condition is fulfilled, the pixel contributes to $V_{\rm eff}$.

The partial effective volume $V_{{\rm eff},sh[i]}$ (i.e., corresponding to the ith shell) is then calculated from

$$V_{{\rm eff},sh[i]} = \frac{n_{\rm pix}^{\rm (good)}}{n_{\rm pix}}, \tag{A1}$$

where $n_{\rm pix}$ is the number of pixels that the simulation box was divided into, and $n_{\rm pix}^{\rm (good)}$ counts the pixels that fulfill the aforementioned condition. Next, we obtain complete effective volume by summing up the values from all the shells,

$$V_{\rm eff} = \sum_{i=0}^{n_{\rm shells}} V_{{\rm eff},sh[i]}. \tag{A2}$$

Figure 7 illustrates briefly our method of calculating $V_{\rm eff}$ for just one light cone shell considered ($[z_1, z_2]$ in the plot). For simplicity, the figure shows the two-dimensional case, but of course we perform the procedure in 3D space.

The left part of the figure symbolizes the light cone with the redshift cut (light blue) and also angular cut (dark blue). We marked the observer's position with the red dot in the center of the light cone. The right part visualizes the counts which would result from the situation presented in this figure. As an example, we distinguished two pixels. Note that we do not include two different pixels twice if they occupy the same position within the output box.

Focusing on the redshift range that we identified as optimal for our study ($0.15 \lesssim z \lesssim 0.30$), we maximize $V_{\rm eff}$ by drawing $n_{run} = 3000$ times a set of 15 randomly oriented sky chunks covering $\sim$1567 deg^2 each, as described in Sec. II. Such numerous collection of sky fragments enabled us to obtain a satisfying level of $V_{\rm eff}$ even from such thin redshift range in the catalog. From these we chose one set that provides the highest $V_{\rm eff}$ value. We noticed that $n_{run} = 3000$ is sufficient due to the saturation of the largest $V_{\rm eff}$ found.

Figure 8 shows the positions of sky chunks used in our catalogs within this work, in Hammer equal-area projection. Our procedure allowed us to obtain $V_{\rm eff} = 0.75$. Theoretically, it is possible to obtain $V_{\rm eff}$ as high as 0.94 with our redshift ranges. However, it requires a drastic increase in the number of sky chunks and becomes computationally ineffective, providing simultaneously a weak increase of independent information in the catalog. Note that the value of effective volume is not limited to 1 by definition. The value $V_{\rm eff} = 1$ would indicate that potentially constructed catalog contains all the information from one simulation box, i.e., one light cone shell. Getting $V_{\rm eff} > 1$ could be easily achieved, e.g., by selecting an entire sky and full redshift range of ELEPHANT light cone $z \in [0, 0.5]$—for that case, we would get $V_{\rm eff} \sim 2.23$.

[1] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, Bielewicz *et al.*, Astron. Astrophys. **641**, A6 (2020).

[2] S. Alam, M. Aubert, S. Avila, C. Balland, J. E. Bautista, M. A. Bershady, D. Bizyaev, M. R. Blanton, A. S. Bolton, J. Bovy, J. Brinkmann, J. R. Brownstein, E. Burtin, S. Chabanier, M. J. Chapman *et al.*, Phys. Rev. D **103**, 083533 (2021).

[3] T. M. C. Abbott, M. Aguena, A. Alarcon, S. Allam, O. Alves, A. Amon, F. Andrade-Oliveira, J. Annis, S. Avila, D. Bacon, E. Baxter, K. Bechtol, M. R. Becker, G. M. Bernstein, S. Bhargava (DES Collaboration), Phys. Rev. D **105**, 023520 (2022).

[4] D. Brout, D. Scolnic, B. Popovic, A. G. Riess, J. Zuntz, R. Kessler, A. Carr, T. M. Davis, S. Hinton, D. Jones, W. D. Kenworthy, E. R. Peterson, K. Said, G. Taylor, N. Ali *et al.*, arXiv:2202.04077.

[5] P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, 1980), Vol. 98.

[6] J. R. Bond, L. Kofman, and D. Pogosyan, Nature (London) **380**, 603 (1996).

[7] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002).

[8] E. Hawkins *et al.*, Mon. Not. R. Astron. Soc. **346**, 78 (2003).

[9] I. Zehavi, Z. Zheng, D. H. Weinberg, M. R. Blanton, N. A. Bahcall, A. A. Berlind, J. Brinkmann, J. A. Frieman, J. E. Gunn, R. H. Lupton, R. C. Nichol, W. J. Percival, D. P. Schneider, R. A. Skibba, Strauss *et al.*, Astrophys. J. Lett. **736**, 59 (2011).

[10] J. A. Piscionere, A. A. Berlind, C. K. McBride, and R. Scoccimarro, Astrophys. J. Lett. **806**, 125 (2015).

[11] F. Beutler, H.-J. Seo, S. Saito, C.-H. Chuang, A. J. Cuesta, D. J. Eisenstein, H. Gil-Marín, J. N. Grieb, N. Hand, F.-S. Kitaura, C. Modi, R. C. Nichol, M. D. Olmstead, W. J. Percival *et al.*, Mon. Not. R. Astron. Soc. **466**, 2242 (2017).

[12] S. J. Maddox, G. Efstathiou, and W. J. Sutherland, NASA STI/Recon Technical Report N **97** 27602 (1995).

[13] A. J. Connolly, R. Scranton, D. Johnston, S. Dodelson, D. J. Eisenstein, J. A. Frieman, J. E. Gunn, L. Hui, B. Jain, S. Kent *et al.*, Astrophys. J. Lett. **579**, 42 (2002).

[14] A. H. Maller, D. H. McIntosh, N. Katz, and M. D. Weinberg, Astrophys. J. Lett. **619**, 147 (2005).

[15] M. Crocce, A. Cabré, and E. Gaztañaga, Mon. Not. R. Astron. Soc. **414**, 329 (2011).

[16] Y. Wang, R. J. Brunner, and J. C. Dolence, Mon. Not. R. Astron. Soc. **432**, 1961 (2013).

[17] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **340**, 580 (2003).

[18] R. C. Nichol, R. K. Sheth, Y. Suto, A. J. Gray, I. Kayo, R. H. Wechsler, F. Marin, G. Kulkarni, M. Blanton, A. J. Connolly, J. P. Gardner, B. Jain, C. J. Miller, A. W. Moore, A. Pope *et al.*, Mon. Not. R. Astron. Soc. **368**, 1507 (2006).

[19] H. Guo, Z. Zheng, Y. P. Jing, I. Zehavi, C. Li, D. H. Weinberg, R. A. Skibba, R. C. Nichol, G. Rossi, C. G. Sabiu, D. P. Schneider, and C. K. McBride, Mon. Not. R. Astron. Soc. **449**, L95 (2015).

[20] Z. Slepian, D. J. Eisenstein, F. Beutler, C.-H. Chuang, A. J. Cuesta, J. Ge, H. Gil-Marín, S. Ho, F.-S. Kitaura, C. K. McBride, R. C. Nichol, W. J. Percival, S. Rodríguez-Torres, A. J. Ross, R. Scoccimarro *et al.*, Mon. Not. R. Astron. Soc. **468**, 1070 (2017).

[21] F. Sosa Nuñez and G. Niz, J. Cosmol. Astropart. Phys. 12 (2020) 021.

[22] Z. Slepian and D. J. Eisenstein, Mon. Not. R. Astron. Soc. **478**, 1468 (2018).

[23] O. Umeh, J. Cosmol. Astropart. Phys. 05 (2021) 035.

[24] S. D. M. White, Mon. Not. R. Astron. Soc. **186**, 145 (1979).

[25] E. Gaztanaga, Mon. Not. R. Astron. Soc. **268**, 913 (1994).

[26] D. J. Croton, E. Gaztanaga, C. M. Baugh, P. Norberg, M. Colless, I. K. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole *et al.*, Mon. Not. R. Astron. Soc. **352**, 1232 (2004).

[27] R. Juszkiewicz, F. R. Bouchet, and S. Colombi, Astrophys. J. **412**, L9 (1993).

[28] E. L. Lokas, R. Juszkiewicz, D. H. Weinberg, and F. R. Bouchet, Mon. Not. R. Astron. Soc. **274**, 730 (1995).

[29] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, Phys. Rep. **513**, 1 (2012).

[30] W. Hu and I. Sawicki, Phys. Rev. D **76**, 064004 (2007).

[31] T. P. Sotiriou and V. Faraoni, Rev. Mod. Phys. **82**, 451 (2010).

[32] H. A. Buchdahl, Mon. Not. R. Astron. Soc. **150**, 1 (1970).

[33] A. I. Vainshtein, Phys. Lett. **39B**, 393 (1972).

[34] G. Dvali, G. Gabadadze, and M. Porrati, Phys. Lett. B **485**, 208 (2000).

[35] V. Sahni and Y. Shtanov, J. Cosmol. Astropart. Phys. 11 (2003) 014.

[36] B. Li and H. Zhao, Phys. Rev. D **80**, 044027 (2009).

[37] K. Liu, R. P. Eatough, N. Wex, and M. Kramer, Mon. Not. R. Astron. Soc. **445**, 3115 (2014).

[38] J.-J. Wei, B.-B. Zhang, X.-F. Wu, H. Gao, P. Mészáros, B. Zhang, Z.-G. Dai, S.-N. Zhang, and Z.-H. Zhu, J. Cosmol. Astropart. Phys. 11 (2017) 035.

[39] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, K. Agatsuma, Aggarwal *et al.*, Astrophys. J. **896**, L44 (2020).

[40] I. H. Stairs, Living Rev. Relativity **6**, 5 (2003).

[41] F. De Marchi and G. Cascioli, Classical Quantum Gravity **37**, 095007 (2020).

[42] B. Bose, K. Koyama, W. A. Hellwing, G.-B. Zhao, and H. A. Winther, Phys. Rev. D **96**, 023519 (2017).

[43] B. Bose, K. Koyama, and H. A. Winther, J. Cosmol. Astropart. Phys. 10 (2019) 021.

[44] J. E. García-Farieta, W. A. Hellwing, S. Gupta, and M. Bilicki, Phys. Rev. D **103**, 103524 (2021).

[45] Z. Ivezic, J. Tyson, T. Axelrod, D. Burke, C. Claver, K. Cook, S. Kahn, R. Lupton, D. Monet, P. Pinto *et al.*, in *American Astronomical Society Meeting Abstracts No. 213* (Astronomical Observatory, 2009), Vol. 213, p. 460.03.

[46] R. Laureijs, J. Amiaux, S. Arduini, J.-L. Augueres, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet *et al.*, arXiv:1110.3193.

[47] M. Cautun, E. Paillas, Y.-C. Cai, S. Bose, J. Armijo, B. Li, and N. Padilla, Mon. Not. R. Astron. Soc. **476**, 3195 (2018).

[48] B. Li, G.-B. Zhao, R. Teyssier, and K. Koyama, J. Cosmol. Astropart. Phys. 01 (2012) 051.
[49] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, Jarosik *et al.*, Astrophys. J. Suppl. Ser. **208**, 19 (2013).
[50] J. Khoury and A. Weltman, Phys. Rev. D **69**, 044026 (2004).
[51] P. Brax, C. van de Bruck, A.-C. Davis, and D. J. Shaw, Phys. Rev. D **78**, 104021 (2008).
[52] C. Arnold, P. Fosalba, V. Springel, E. Puchwein, and L. Blot, Mon. Not. R. Astron. Soc. **483**, 790 (2019).
[53] S. Alam, C. Arnold, A. Aviles, R. Bean, Y.-C. Cai, M. Cautun, J. L. Cervantes-Cota, C. Cuesta-Lazaro, N. C. Devi, A. Eggemeier *et al.*, J. Cosmol. Astropart. Phys. 11 (2021) 050.
[54] E. Babichev and C. Deffayet, Classical Quantum Gravity **30**, 184001 (2013).
[55] P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu, Astrophys. J. Lett. **762**, 109 (2013).
[56] M. Manera, R. Scoccimarro, W. J. Percival, L. Samushia, C. K. McBride, A. J. Ross, R. K. Sheth, M. White, B. A. Reid, A. G. Sánchez *et al.*, Mon. Not. R. Astron. Soc. **428**, 1036 (2013).
[57] J. Blaizot, Y. Wadadekar, B. Guiderdoni, S. T. Colombi, E. Bertin, F. R. Bouchet, J. E. G. Devriendt, and S. Hatton, Mon. Not. R. Astron. Soc. **360**, 159 (2005).
[58] R. Overzier, G. Lemson, R. E. Angulo, E. Bertin, J. Blaizot, B. M. B. Henriques, G. D. Marleau, and S. D. M. White, Mon. Not. R. Astron. Soc. **428**, 778 (2013).
[59] A. Smith, S. Cole, C. Baugh, Z. Zheng, R. Angulo, P. Norberg, and I. Zehavi, Mon. Not. R. Astron. Soc. **470**, 4646 (2017).
[60] P. Coles and B. Jones, Mon. Not. R. Astron. Soc. **248**, 1 (1991).
[61] F. R. Bouchet and L. Hernquist, Astrophys. J. Lett. **400**, 25 (1992).
[62] E. Gaztanaga and F. Bernardeau, Astron. Astrophys. **331**, 829 (1998), https://core.ac.uk/download/pdf/25215052.pdf.
[63] A. Pollo, Acta Astron. **47**, 413 (1997), http://acta.astrouw.edu.pl/Vol47/n4/pap_47_4_2.pdf.
[64] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchman, Mon. Not. R. Astron. Soc. **341**, 1311 (2003).
[65] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. Lett. **538**, 473 (2000).
[66] B. Li, W. A. Hellwing, K. Koyama, G.-B. Zhao, E. Jennings, and C. M. Baugh, Mon. Not. R. Astron. Soc. **428**, 743 (2013).
[67] W. A. Hellwing, B. Li, C. S. Frenk, and S. Cole, Mon. Not. R. Astron. Soc. **435**, 2806 (2013).
[68] W. A. Hellwing, K. Koyama, B. Bose, and G.-B. Zhao, Phys. Rev. D **96**, 023515 (2017).
[69] D. J. Bacon and A. N. Taylor, Mon. Not. R. Astron. Soc. **344**, 1307 (2003).
[70] B. Jain and A. Taylor, Phys. Rev. Lett. **91**, 141302 (2003).
[71] R. Massey, J. Rhodes, R. Ellis, N. Scoville, A. Leauthaud, A. Finoguenov, P. Capak, D. Bacon, H. Aussel, J.-P. Kneib, A. Koekemoer, H. McCracken, B. Mobasher, S. Pires, Refregier *et al.*, Nature (London) **445**, 286 (2007).
[72] E. Hivon, F. R. Bouchet, S. Colombi, and R. Juszkiewicz, Astron. Astrophys. **298**, 643 (1995), https://www.osti.gov/biblio/1847347.
[73] T. Okamura, A. Taruya, and T. Matsubara, J. Cosmol. Astropart. Phys. 08 (2011) 012.
[74] F.-S. Kitaura and R. E. Angulo, Mon. Not. R. Astron. Soc. **425**, 2443 (2012).
[75] J. N. Fry and E. Gaztanaga, Astrophys. J. Lett. **413**, 447 (1993).
[76] W. A. Hellwing, XXXIX Polish Astronomical Society Meeting (2020), 10, 315, arXiv:1912.13026.
[77] D. Shi, B. Li, J. Han, L. Gao, and W. A. Hellwing, Mon. Not. R. Astron. Soc. **452**, 3179 (2015).
[78] A. J. Ross, R. J. Brunner, and A. D. Myers, Astrophys. J. Lett. **665**, 67 (2007).
[79] A. J. S. Hamilton, in *The Evolving Universe*, edited by D. Hamilton, Astrophysics and Space Science Library Vol. 231 (Springer, 1998), p. 185.
[80] P. Arnalte-Mur, W. A. Hellwing, and P. Norberg, Mon. Not. R. Astron. Soc. **467**, 1569 (2017).
[81] C. Hernández-Aguayo, J. Hou, B. Li, C. M. Baugh, and A. G. Sánchez, Mon. Not. R. Astron. Soc. **485**, 2194 (2019).
[82] J. T. De Jong, G. A. V. Kleijn, T. Erben, H. Hildebrandt, K. Kuijken, G. Sikkema, M. Brescia, M. Bilicki, N. R. Napolitano, V. Amaro *et al.*, Astron. Astrophys. **604**, A134 (2017).
[83] I. Tutusaus, M. Martinelli, V. F. Cardone, S. Camera, S. Yahia-Cherif, S. Casas, A. Blanchard, M. Kilbinger, F. Lacasa, Z. Sakr *et al.*, Astron. Astrophys. **643**, A70 (2020).
[84] T. Abbott, M. Adamow, M. Aguena, S. Allam, A. Amon, J. Annis, S. Avila, D. Bacon, M. Banerji, K. Bechtol *et al.*, Astrophys. J. Suppl. Ser. **255**, 20 (2021).
[85] M. Walmsley, C. Lintott, T. Géron, S. Kruk, C. Krawczyk, K. W. Willett, S. Bamford, L. S. Kelvin, L. Fortson, Y. Gal, W. Keel, K. L. Masters, V. Mehta, B. D. Simmons, Smethurst *et al.*, Mon. Not. R. Astron. Soc. **509**, 3966 (2021).
[86] A. Porredon, M. Crocce, P. Fosalba, J. Elvin-Poole, A. C. Rosell, R. Cawthon, T. Eifler, X. Fang, I. Ferrero, E. Krause *et al.*, Phys. Rev. D **103**, 043503 (2021).

Correction: The affiliation and the name in the first sentence of the Acknowledgments contained misspellings and have been set right.

3.2 Summary

The found EG signals reaching up to 20% consist a promising result in context of future observational constraints. The optimal redshift where the deviations from GR are maximized, is however linked with used geometry. In fully $3D$ representation the redshift favored in context of searching EG signals does not necessarily have to be identical. In lightcone geometry, selecting the objects within small redshift is related with much less data compared to larger-z probe with the same radial thickness. This increases the shot-noise uncertainties, consequently suppressing potential low-z EG signals. Additionally, the adopted angular counts-in-cells are accompanied with mixing physical scales at various distances. For the same radial thickness of the catalog, the mixing is naturally more severe at low redshifts. The combination of these effects leads to the optimal redshifts for EG signal as spotted in this work, $0.15 \leq z_{opt} \leq 0.3$.
Understanding how geometry effects and assumptions influence the findings is crucial for further studies on constraining the gravity.

Chapter 4

Paper II: Real and redshift-space clustering

4.1 Introduction

The previous study focused on identifying EG signals with angular clustering in real space. The observational data is however accessible in redshift space. Despite of the main argument for working with angular statistics which state that spectroscopic observations provide sparser data than photometry, the success in finding strong signals in angular statistics motivates to study the clustering fully in three dimensions. This involves investigating redshift space distortions which consist of Fingers of God and Kaiser effect. In order to better understand how the redshift-space representation affects cosmic web statistics, we prepare both real and z-space catalogs. While for angular case we worked with lightcone catalogs, in this study we focus on measuring the statistics on representative samples without mixing redshifts. To obtain that, we use the distant observer approximation.

This study depicts an approach for measuring 3D averaged correlation functions in real and z-space using dark matter pseudo-particles, halos and HOD galaxies. We focus on cumulants, especially skewness s_3, as it more clearly highlights the differences between two configurations: real and redshift space. The complete test of used gravity scenarios using these measures would require catalogs mimicking real observations, based on realistic and physically consistent cosmology simulations with GR, $f(R)$ and nDGP. This research however identify where to search for deviations from GR and how redshift space changes them.

Skewness as a Probe of Gravity: Real and Redshift Space Counts-In-Cells

Paweł Drozda,* Wojciech A. Hellwing, and Maciej Bilicki
Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland
(Dated: April 18, 2025)

We study the counts-in-cells reduced skewness s_3 for dark matter, halo, and galaxy distributions in both real and redshift space, using the ELEPHANT (*Extended LEnsing PHysics with ANalytical ray Tracing*) suite of N-body simulations. We compare General Relativity (GR) with two extended (EG) gravity models: $f(R)$ gravity with chameleon screening and the normal-branch Dvali–Gabadadze–Porrati (nDGP) model with Vainshtein screening. We quantify the suppression of s_3 by redshift-space distortions (RSD), finding that while small-scale skewness is strongly reduced, the $F5$ model retains a $\sim 4\%$ deviation from GR in galaxy samples, corresponding to a 2σ significance. We show that the ratio $s_3^{\rm RSD}/s_3^{\rm real}$ is approximately independent of the gravity model across tracers and redshifts. Our results demonstrate that real-space predictions can help reliably infer redshift-space skewness in both GR and extended gravity, providing a new tool for testing gravity with current and forthcoming galaxy redshift surveys.

I. INTRODUCTION

The large-scale structure of the Universe provides a rich testing ground for cosmological models, particularly through statistical analysis of density and velocity fields. Traditionally, two-point statistics such as the power spectrum and two-point correlation function have been instrumental in constraining cosmological parameters [*e.g.* 1–6]. Yet, these measures alone cannot fully capture the complexity of the Cosmic Web. Higher-order statistics, such as skewness and kurtosis, complement two-point measures by probing the non-linear gravitational clustering process, which can significantly differ between General Relativity (GR) and various Extended Gravity (EG) models (in the literature also dubbed as 'Modified Gravity')[7–9].

Recent studies highlight that these higher-order statistics are sensitive to subtle deviations from GR, making them powerful tools for testing structure formation scenarios in EG theories [10–13]. Measurements of skewness and kurtosis in particular can reveal modifications to gravity or dark energy beyond the standard ΛCDM model. Despite their promise, such analyses have predominantly focused on theoretical configuration space[7, 8, 14, 15]. However, actual observations readily access only galaxy positions in redshift space, where peculiar velocities introduce distortions that complicate the interpretation of clustering statistics [16–19]. These redshift-space distortions (RSD), while adding further complications to the galaxy clustering pattern, also carry valuable information about the growth rate of cosmic structure. Accurate modeling of RSD is thus crucial for interpreting observations and testing gravity theories [20, 21].

A systematic exploration of higher-order statistics in redshift space, particularly skewness, remains lacking but can become important for the next generation of robust cosmological tests. Constraints on cosmological parameters based on different statistics, scales, and phenomena are pivotal for internal consistency checks and for identifying potential unknown systematic effects that could limit the precision of these tests. Here, the potential of higher-order cumulants of the matter, halo, and galaxy density fields, remains largely unexploited and not fully tapped.

Addressing this gap, we perform a detailed study of skewness in both real and redshift space, comparing GR with two prominent EG scenarios: the $f(R)$ gravity model with chameleon screening and the normal-branch Dvali-Gabadadze-Porrati (nDGP) gravity model employing Vainshtein screening. We analyze how the transition from real to redshift space influences the detectability of gravity-induced differences, investigating separately the dark matter, halo, and galaxy distributions. Hence, this paper is a natural extension of our previous study focused on the angular clustering moments.[9]

Recent simulation-based studies have demonstrated that modifications to gravity can significantly alter higher-order clustering statistics, such as skewness and kurtosis, in both dark matter and halo distributions [7, 8, 22, 23]. In particular, deviations in the reduced skewness s_3 have been identified as a sensitive probe of screening mechanisms, with models like $f(R)$ and nDGP showing scale-dependent signatures in real-space clustering. However, translating these signatures into redshift space remains challenging. While RSD can suppress or mask EG signals at small scales [18, 19], new approaches incorporating realistic velocity distributions [20] and angular clustering statistics [9] offer promising avenues for observational tests. Furthermore, recent theoretical developments emphasize the need to move beyond two-point functions, using bispectra and parity-breaking correlations [10, 21], to robustly probe deviations from GR in upcoming surveys. Nonetheless, a comprehensive and systematic study of higher-order cumulants in redshift space across different tracer populations and gravity models remains largely unexplored. Our work aims to fill this gap.

* (pdrozda, hellwing, bilicki)@cft.edu.pl

This paper is structured as follows: Section II describes the simulations and catalogs; Section III introduces the statistical tools employed; Section IV presents our analysis and findings, emphasizing skewness across real and redshift spaces; and Section V discusses the broader implications of our results.

II. SIMULATIONS AND CATALOGS

In this work, we use the ELEPHANT (*Extended LEnsing PHysics with ANalytical ray Tracing*) N-body simulation suite [24], performed with the ECOSMOG code [25], a modified version of RAMSES adapted for EG models. The simulations evolve 1024^3 dark matter pseudo-particles within a $(1024/h\ \mathrm{Mpc})^3$ comoving box, generated from five independent realizations of the initial power spectrum. The assumed background cosmology corresponds to the WMAP9 results [26], with parameters $\Omega_M = 0.281$, $\Omega_\Lambda = 0.719$, and $h = 0.697$. Dark matter halos were identified using the ROCKSTAR halo finder [27], which employs a phase-space friends-of-friends algorithm. We also utilize ELEPHANT-dedicated galaxy mock catalogs, constructed using the Halo Occupation Distribution (HOD) method. The HOD parameters were calibrated to reproduce the projected two-point correlation function of BOSS CMASS DR9 galaxies [28] and to match the number density of a volume-limited sample [29]. The mocks are constructed to recover the assumed satellite-to-centrals ratio (varying from $10-15\%$ across redshifts see Hernández-Aguayo *et al.* [30] fro details), giving us a chance to study separately all-galaxy and centrals-only samples.

For biased tracers, we construct three halo samples, H1, H2, and H3, characterized by their mean number densities of 10^{-3}, 5×10^{-4}, and 10^{-4} $(h/\mathrm{Mpc})^3$, respectively, following the definitions in [31]. For each model, sample, and redshift, the desired number density is achieved by rejecting halos with masses smaller than an established threshold $M_{\rm min}$ that fulfills the target number density criterion. The exact mass cuts are listed in Table I, along with the corresponding number densities for both the halo and galaxy catalogs. The mass cuts are based on the halo virial mass, M_{200c}, defined as the mass enclosed within the volume where the mean density exceeds 200 times the critical density. For galaxies, the catalogs exhibit slight variations in number density; therefore, in Table I, we show the range of values rather than a single number.

A. Extended Gravity

In addition to the standard ΛCDM cosmology with GR-based structure formation, we consider two popular phenomenological extensions of General Relativity, implemented within the ELEPHANT suite. These are:

TABLE I. Mass cuts in halo catalogs for a given sample, model and redshift. The column corresponding to object number densities shows values both for halos (H1-H3 samples) and galaxies.

z	Sample	n $[10^{-4}\ (\mathrm{Mpc}/h)^{-3}]$	$M_{min}[10^{13}M_\odot/h]$ GR	F6	F5	N5	N1
0	H1	10	0.26	0.30	0.30	0.27	0.27
	H2	5	0.58	0.64	0.66	0.58	0.59
	H3	1	2.73	2.80	3.22	2.76	2.84
	Galaxies	3.8-3.9	-	-	-	-	-
0.3	H1	10	0.26	0.30	0.31	0.26	0.27
	H2	5	0.56	0.59	0.64	0.56	0.58
	H3	1	2.41	2.44	2.83	2.43	2.53
	Galaxies	3.1-3.3	-	-	-	-	-
0.5	H1	10	0.23	0.26	0.30	0.23	0.25
	H2	5	0.53	0.55	0.62	0.54	0.55
	H3	1	2.13	2.14	2.49	2.15	2.25
	Galaxies	3.1-3.3	-	-	-	-	-

- The $f(R)$ gravity model with the chameleon screening mechanism [32, 33],
- The normal branch of Dvali-Gabadadze-Porrati (nDGP) gravity with the Vainshtein screening mechanism [34, 35].

These two EG models provide a testbed for cosmological inference based on real and redshift-space reduced cumulants, and we will compare their predictions with those of GR.

B. Redshift-Space

In this study, we work both in configuration (real) space and in redshift space. The latter can be constructed using information about the positions and radial velocities of objects relative to the observer. To construct our redshift-space catalogs, we employ the distant observer approximation, that is, the galaxy/halo line-of-sight (LOS) directions are assumed to be parallel.

Geometrically, this can be understood as placing the simulation box at an effective comoving distance $D \gg \tilde{x}_{12}$, where $\tilde{x}_{12}$ is the mean separation between galaxies or halos. While at $z = 0.3$ and $z = 0.5$ the distant observer approximation is fully applicable, at $z = 0$ it should be interpreted as a redshift-space analysis without including projection effects. To obtain redshift-space positions, we modify the object positions along the LOS according to:

$$\vec{x}_{\rm RSD} = \vec{x} + \frac{1+z_{\rm snap}}{H(z_{\rm snap})}\left(\vec{v}\cdot\hat{e}_\parallel\right)\hat{e}_\parallel, \tag{1}$$

where $\vec{x}_{\rm RSD}$ and $\vec{x}$ are the redshift-space distorted and original comoving positions, respectively, $\vec{v}$ is the peculiar velocity, $\hat{e}_\parallel$ is the normalized LOS direction, and

$H(z_{\rm snap})$ is the Hubble parameter at the snapshot redshift $z_{\rm snap}$.

To maximize the information extracted from the 3D velocity field, we distort the simulation box separately along each of the three Cartesian directions $[x, y, z]$, assuming the LOS along each axis in turn. Computations are performed independently for each LOS orientation. Although the large-scale three-dimensional peculiar velocity field is correlated, meaning that the three projections are not fully independent, this procedure effectively yields approximately $\sqrt{3}$ times more independent redshift-space realizations compared to a single projection.

III. AVERAGED CORRELATION FUNCTIONS AND MOMENTS

We work with the three-dimensional volume-averaged correlation functions [36, 37] defined as

$$\bar{\xi}_J(R) \equiv \langle \delta_R^J \rangle, \tag{2}$$

where R is the smoothing scale.[1] Here, $\delta_R = \rho_R/\bar{\rho}_R - 1$ denotes the density contrast, defined as the excess of the local density ρ_R over the mean density $\bar{\rho}_R$. The lower index subscript R indicates that the density field is smoothed using a three-dimensional top-hat spherical window function of a co-moving radius R, and the angle brackets $\langle\cdot\rangle$ denote ensemble averaging over the entire catalog.

When working with discrete data, the smoothed density ρ_R at a given radius R can be approximated by counting the number of objects within a sphere of radius R. As long as the number of objects $N \gg 1$, the discrete distribution provides a good approximation to a continuous background fluid [1, 38]. We estimate $\bar{\xi}_J(R)$ by performing counts-in-cells (CIC) over random positions within the catalog and computing the central moments of the counts:

$$M_J(R) = \mathrm{E}\left[(N - \mathrm{E}[N])^J\right], \tag{3}$$

where $\mathrm{E}[N]$ denotes the expected value of the counts N, after applying normalization and shot-noise corrections (see e.g. [39]). Now, to get an estimate of $\bar{\xi}_J$ from Eqn. (2) we just need to dive the central moments by $E[N]^J$. In addition, we consider the reduced cumulants s_J defined as

$$s_J \equiv \frac{\bar{\xi}_J}{\bar{\xi}_2^{J-1}}. \tag{4}$$

Under the assumption of a power-law power spectrum, perturbation theory predicts the reduced cumulants to be weakly varying monotonic functions of scale [38, 40], making them useful diagnostics for cosmological tests.

The connection between volume-averaged correlations and the central moments shows that $\bar{\xi}_J(R)$ depends on the shape of the probability distribution function (PDF) of the counts. More precisely, these correlation functions are the connected parts of the moments, hence they quantify the departure from a Gaussian distribution. Simultaneously, the deviations of the PDF from an initially Gaussian form are driven by gravitational evolution of the density field (see [37, 41, 42]).

Thus, $\bar{\xi}_J(R)$ capture the growth of non-Gaussian features and quantify the gravitational structure formation and evolution of the large-scale Universe.

IV. REAL AND REDSHIFT SPACE SKEWNESS

We begin our analysis by examining the statistical properties of the density contrast. Figure 1 shows the probability distribution functions (PDFs) of the density contrast δ_R, estimated from CIC by treating the discrete particle counts as a proxy for the local density. We display results at three chosen smoothing scales, both in configuration (real) space and redshift space, for the GR model at redshift $z = 0.3$.

The histograms of counts converted into density contrast reveal clear RSD effects. At small scales ($R = 3\,h^{-1}$Mpc), the density contrast is suppressed in redshift space due to the Fingers-of-God (FoG) effect [17]. This is strikingly evident by comparing the long exponential tail of the real-space histogram, extending up to $\delta \sim 200$, with the much more compact redshift-space distribution, where few cells reach even moderate contrasts of $\delta \sim 100$.

At intermediate scales ($R = 15\,h^{-1}$Mpc), which mark the transition from the non-linear to the linear regime in classical gravitational instability theory, the real and redshift-space PDFs are remarkably similar.

Finally, in the large cell volume limit ($R = 43\,h^{-1}$Mpc), we observe that the density contrast PDF in redshift space is wider than that of real space. This is a clear manifestation of the so-called Kaiser squashing effect, where coherent large-scale infall motions enhance the redshift-space density around large-scale structures [17].

We can take a more quantitative view of the CIC distributions by examining the first two moments and the related $\bar{\xi}_J$ and s_J statistics. Figure 2 shows the second- and third-order volume-averaged correlation functions (left panel) and the corresponding reduced skewness s_3 (right panel) for both real and redshift space at $z = 0.3$. We limit our analysis to scales $R \lesssim 40\,h^{-1}$Mpc. For larger cell radii, the third moment of the CIC becomes significantly affected by finite-volume effects stemming from the combination of the simulation box size and the limited number of realizations [43].

The most striking feature is the significant reduction of s_3 in redshift space for scales $R \lesssim 30\,h^{-1}$Mpc com-

[1] For transparency, throughout this paper we denote the smoothing scale as R regardless of whether it refers to real or redshift space.

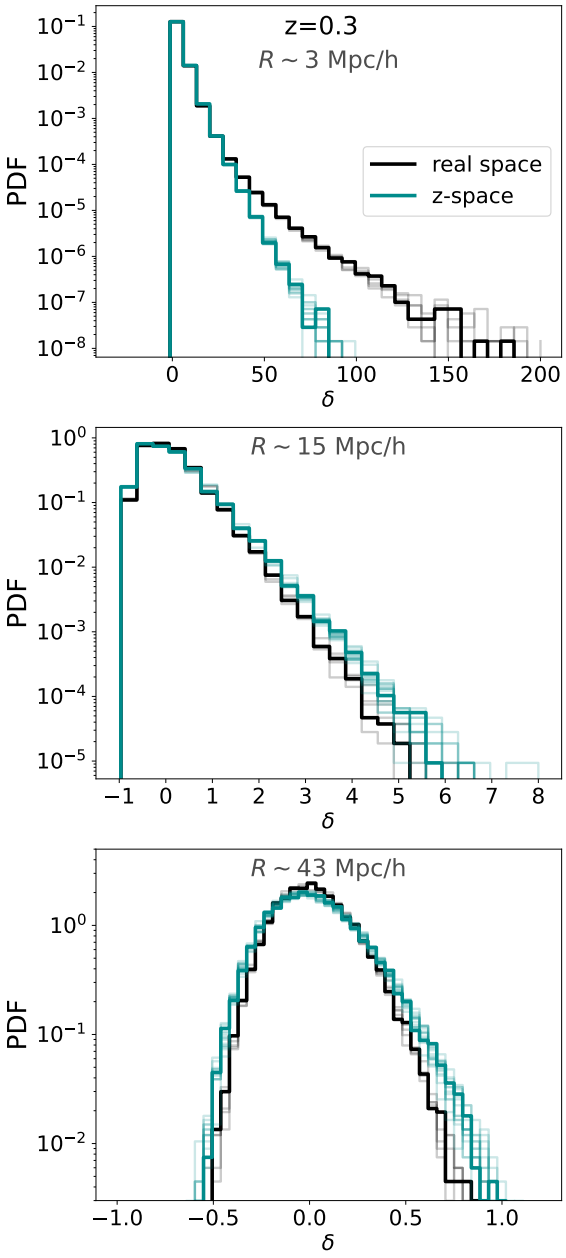

FIG. 1. Distribution of the density contrast of dark matter simulation pseudo-particles in real space (black) and redshift space (dark cyan) at $z = 0.3$ for three chosen smoothing scales. Thinner semi-transparent lines correspond to PDFs from individual realizations, while the thick solid line shows the average over all realizations, separately for each scale and space. Note the different horizontal and vertical axis ranges.

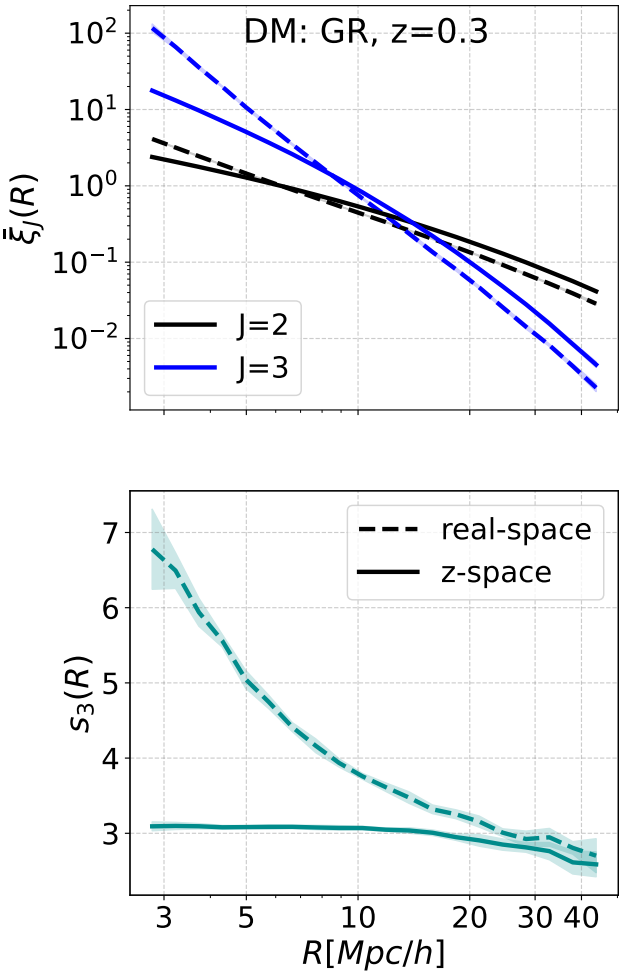

FIG. 2. The second- and third-order volume-averaged correlation functions (top panel) and the third-order reduced cumulant (bottom panel) at redshift $z = 0.3$. Dashed lines correspond to real space, while solid lines to redshift space. All results are for the GR scenario for dark matter particles.

pared to real space. This suppression is caused by statistically random peculiar motions of objects on small scales, leading to the FoG effect. This behavior of skewness was already reported by [44], who showed that the redshift-space s_3 is much less scale-dependent than its real-space counterpart. A similar suppression is also visible in the volume-averaged correlation functions $\bar{\xi}_J$ for $J = \{2, 3\}$, manifesting as a reduction of the signal at $R < 10\,h^{-1}$Mpc in redshift space compared to real space. These trends are consistent with the previous results obtained from the CIC analysis (Figure 1).

At larger scales, redshift space yields slightly higher $\bar{\xi}_J$ values compared to real space. However, due to the construction of the cumulants, the overall trend of the s_3 amplitude reduction in z-space is preserved. The interpretation of behavior of volume-averaged correlations is more convoluted, whereas cumulants show the real-to-redshift space transition that makes a more direct link to the predictions of perturbation theory (PT). Therefore, in this work, we focus primarily on the s_J statistics, especially the skewness s_3.

Next, we compare different gravity models in the context of real and redshift-space. Figure 3 shows the skew-

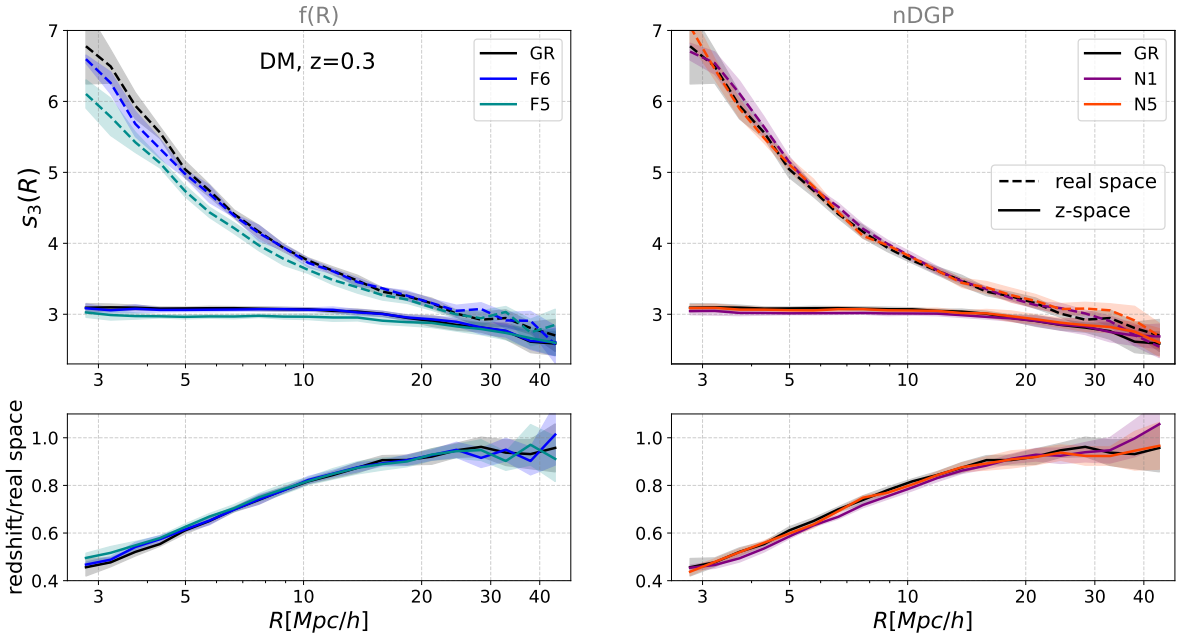

FIG. 3. Upper panels: third-order reduced cumulant s_3 of dark matter for GR (black), compared to $f(R)$ (left) and nDGP (right) gravity models at $z = 0.3$. Dashed lines represent real space, and solid lines indicate redshift space. Lower panels show the ratio of s_3 between redshift and real space for the particular gravity models.

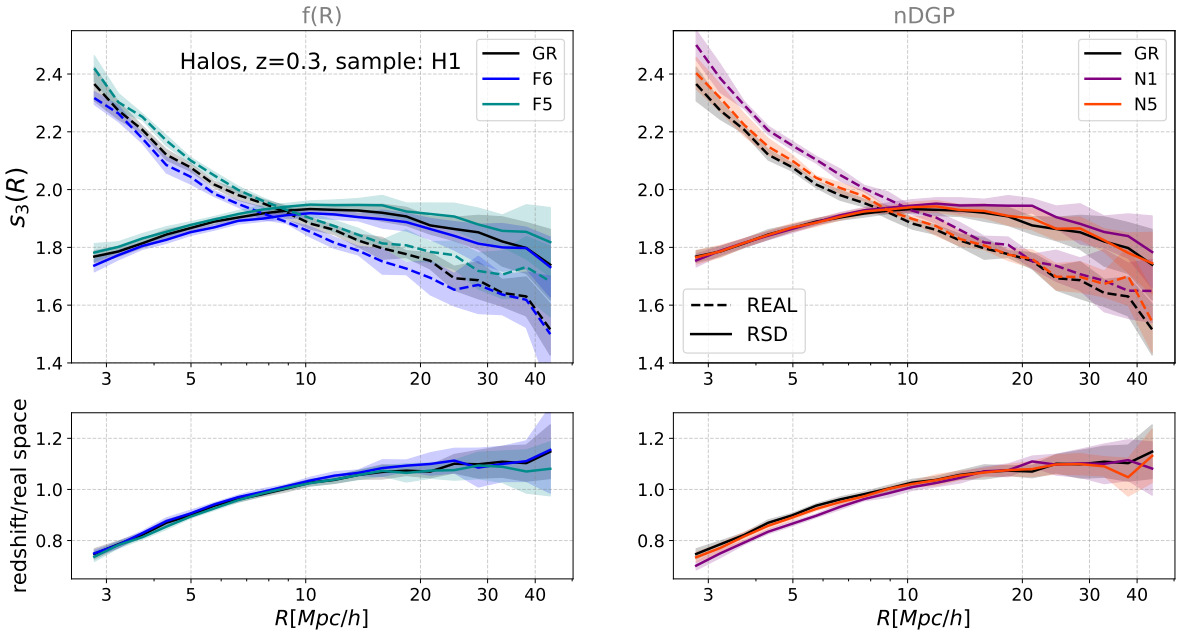

FIG. 4. Analogous plot as Fig. 3, but for the halos from sample H1.

ness s_3 for the GR, $f(R)$, and nDGP models. The trend of redshift-space suppression of the s_3 signal is quantitatively similar across all models. Even for the $F5$ model, which exhibits a significantly lower s_3 than the other models in real space, this difference relative to GR persists in redshift space.

On the other hand, the ratio of redshift- to real-space skewness, shown in the bottom panels, is in strong agreement among the different gravity scenarios. We discuss the potential universality of the real-to-redshift-space transition in Section IV C. Higher-order reduced cumulants s_J with $J > 3$ exhibit similar behavior and we do

not show them here.

As we checked, the skewness in redshift space also shows slight variations across redshifts. At $z = 0$, we observe that s_3 slowly decreases (by $\lesssim 10\%$) with scale between 3 and $30\,h^{-1}$Mpc, a feature that is not present at higher redshifts.

A. Dark matter halos

We now move from the counts of dark matter field tracers (particles) to the clustering analysis of halos. While the CIC of dark matter particles capture the properties of the smooth dark matter background field, it is the distribution of halos that is more directly related to observations. Halos are not only biased tracers of the underlying dark matter distribution but also host the luminous counterparts — galaxies. Thus, the analysis of the CIC distributions of the real and redshift-space halo fields provides information that has both stronger physical context and a more direct connection to observations.

The large-scale dark matter density field approximately follows a log-normal distribution, with positive skewness reflecting the continuous accumulation of matter into dense regions. Gravitational instability leads to the formation of collapsed halos, whose abundance — the halo mass function — depends primarily on the halo virial mass [45, 46]. Motivated by this, we divide our halo population into three samples based on different virial mass cuts (as listed in Table I) and analyze their hierarchical clustering separately.

We begin by examining the skewness of the H1 halo sample at $z = 0.3$, shown in Figure 4. This can be directly compared to the corresponding statistics for the smooth dark matter field presented in Figure 3. The first striking difference lies in the magnitude and scale dependence of RSD. For the smooth dark matter case, the s_3 amplitude was suppressed in redshift space up to $R \sim 30$–$40\,h^{-1}$Mpc; for the H1 halo sample, however, we identify a qualitatively different redshift-space behavior with two distinct regimes. At small scales ($R \lesssim 10\,h^{-1}$Mpc), the skewness is clearly suppressed in redshift space, similarly as was the case for the smooth density field. However, at larger separations ($R \gtrsim 10\,h^{-1}$Mpc), this trend is reversed: the amplitude of $s_3(R)$ is enhanced compared to real space. This behavior parallels the classical picture of the monopole of the two-point correlation function (2PCF): at small scales, random virial motions cause amplitude suppression (FoG effect), while at large scales, coherent infall enhances the clustering amplitude (Kaiser squashing effect [17]).

The Kaiser effect should be also present in the dark matter density field. For the case of our simulation scales and resolution it appears, however, that it is subdominant even at larger scales, where the overall suppression from virial motions seems to dominate the full s_3 amplitude. In contrast, for halos, where we only consider bulk motions of their centers of mass, the large-scale infall is dominating the velocity field, allowing the Kaiser enhancement to become significant at large-scales. Additionally, the overall amplitude of s_3 is consistently lower for halos than for the smooth dark matter field, reflecting the effect of halo biasing on hierarchical clustering [47].

From Figure 4, it is evident that the differences in skewness amplitude between GR and EG models identified in real space are strongly suppressed in redshift space. The fact that both the shapes and amplitudes of s_3 converge across GR and EG models in redshift space indicates that the overall effects of enhanced structure formation, as fostered by EG, somehow conspire to bring the redshift-space skewness back close to GR values. It seems that the more enhanced the real-space clustering becomes in a given EG model, the stronger the corresponding redshift-space suppression is, resulting in a near-cancellation of the two trends. This is particularly evident in the lower panels of Figure 4, where we show the ratio of skewness in redshift space to real space for all models. Both $f(R)$ and nDGP models exhibit ratios very close to the fiducial GR case. A weak exception appears for the N1 model, where H1 halos exhibit slightly stronger redshift-space suppression at $R \lesssim 10\,h^{-1}$Mpc compared to GR. However, given the size of the 1σ scatter, this effect is small and likely difficult to be detected observationally.

Interestingly, if this near-universality of the real-to-redshift-space skewness transition also holds for other halo samples and for galaxies, this implies that knowledge of the real-space clustering alone could be sufficient to predict the skewness amplitude in redshift space, regardless of the underlying gravity model. Qualitatively similar trends are found for higher-order reduced cumulants such as kurtosis ($J = 4$) and hyper-kurtosis ($J = 5$) but we do not display them here, focusing just on skewness.

So far, we have compared the skewness amplitudes between real and redshift space by considering all models together. We now investigate the actual relative differences between the EG models and the fiducial GR case taken as a reference. These results are shown in Figure 5. For this exercise, we selected the $z = 0.5$ and H2 halo sample, since for this case the deviations from GR appear to be maximized. Examining the top panel, we recover the well-known results for real-space clustering (see e.g., Fig. 11 in [48]) for a comparable halo sample. Here, the differences between various EG models and GR are typically contained within 5% at small scales ($R \lesssim 10\,h^{-1}$Mpc), while at larger scales the differences become insignificant. The situation is somewhat reversed when we move to redshift space, as illustrated in the bottom panel of Figure 5. At small scales ($R \lesssim 10\,h^{-1}$Mpc), we observe convergence of s_3 for all models, with no significant deviations. Only at larger scales do some EG models show hints of departure from GR. The most notable case is the F5 model, which around $R \sim 20$–$30\,h^{-1}$Mpc exhibits $\sim 2\%$ difference from GR, marginally reaching 1σ significance.

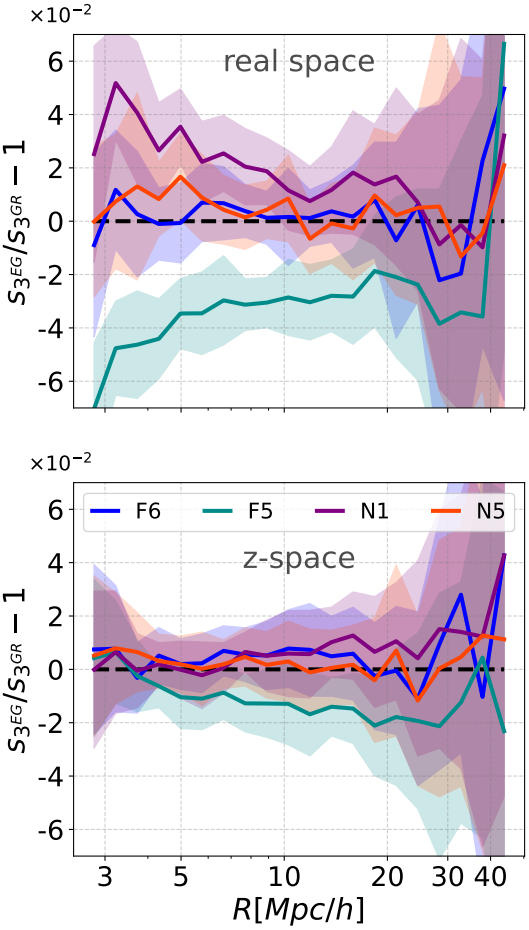

FIG. 5. Relative differences in the reduced third-order cumulant s_3 between EG models and GR for the halo sample H2 at $z = 0.5$ for real (top panel) and redshift-space (bottom panel).

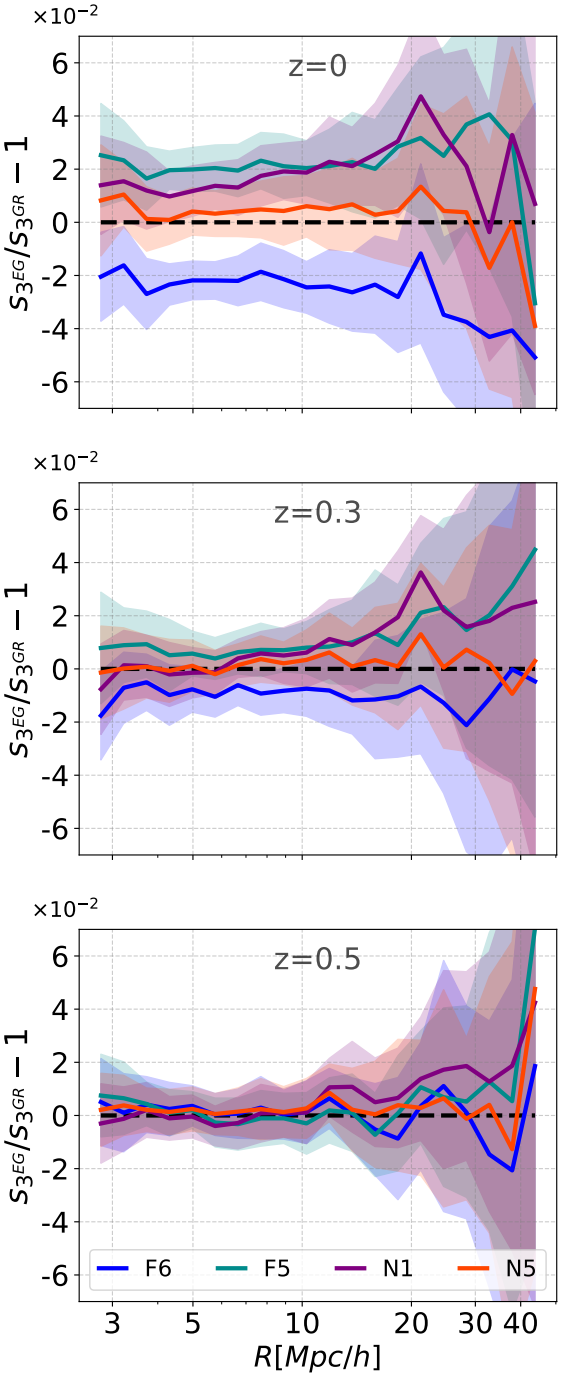

FIG. 6. Relative redshift-space differences in the reduced third-order cumulant s_3 between EG models and GR for the H1 halo sample at three redshifts: $z = 0$ (top), $z = 0.3$ (middle), and $z = 0.5$ (bottom).

While considering different halo samples at $z = 0.5$, we found that a statistically significant signal for the F5 model is also present for the H3 sample, both in real and redshift space (at the $\sim 5\%$ level), whereas for the H1 sample, the signal vanishes completely in redshift space. Interestingly, at $z = 0.3$, the only significant deviation from the GR case is observed for the F5 model and the H3 halo sample. The H1 sample, which has the highest spatial abundance, would in principle be expected to provide the strongest signal-to-noise ratio for CIC distribution functions. However, the suppression of the signal in redshift space persists.

For a more complete view, we present the results for the H1 sample collected across all three redshift epochs considered, as shown in Figure 6. Firstly, one can notice that the net differences between EG models and GR decrease with increasing redshift across all models considered. Interestingly, for this halo sample (i.e., H1), it is the F6 model that exhibits the strongest overall deviations from GR. Although initially surprising, this result is consistent with findings in the literature (see e.g., [49, 50]), which indicate strong non-linear behavior of F6 halos in the small-mass regime. As in the previous case of the smooth density field, we find qualitatively similar trends when examining higher-order moments, such

as kurtosis and hyper-kurtosis. The main takeaway from our halo sample CIC analysis is that, although significant differences in real-space skewness amplitudes exist between EG and GR cases, most — if not all — of the signal is suppressed when moving to counts performed in redshift space. This is a quite remarkable and important negative result. In the subsection below, we investigate how the situation changes when analyzing CIC for the mock galaxy sample.

B. Mock Galaxies

We now move to the observable tracers of the cosmic density field — galaxies. In this case, we work with a number density of $\sim 3.2 \times 10^{-4}\,(h/\mathrm{Mpc})^3$, which places our mock galaxy sample between the H2 and H3 halo datasets in terms of spatial abundance. We performed an analogous analysis to that presented above for halos.

We find that the nDGP models mostly match GR across all cases, except at $z = 0$ in real space, where 1σ deviations appear for the N1 model at some limited range of scales. However, these deviations disappear after moving to redshift space. For the $f(R)$ models, the situation is different. The F5 variant exhibits significant deviations from GR at every redshift in real space, and the signal partially survives in redshift space at $z = 0.5$ and, to a lesser extent, at $z = 0$ (at approximately the 1σ level). At $z = 0.3$, the F5 model shows strong deviations from GR, but only for the full galaxy sample in real space. For the F6 model, 1σ deviations are observed that appear to be largely independent of redshift.

Since signals in redshift space are the most relevant in the observational context, we focus our attention on the models at $z = 0.5$. Figure 7 shows the relative difference between EG and GR for the reduced skewness s_3, comparing the models in both real and redshift space at this redshift. One can notice a significant EG signal for the F5 model. Unlike F6, which drops below 1σ significance in redshift space, the F5 model maintains a $\sim 4\%$ deviation with $\sim 2\sigma$ significance in redshift space. This result is in agreement with Refs. [11]. For F6, the differences in redshift space are practically undetectable. Moreover, galaxies may in some cases provide a stronger EG signal in redshift space than halos. A comparison between Figure 7 and Figure 5 shows that the deviation of F5 from GR is approximately twice as large for galaxies compared to the H2 halo sample, which provides the most comparable number density (see Table I). However, at lower redshifts, this enhancement is much less pronounced. This result is encouraging, as galaxies are the only directly observable large-scale tracers of the cosmic web. Although the $\sim 5\%$ difference observed for the H3 halo sample at $z = 0.5$ is slightly stronger than that for galaxies, the galaxy signal remains promising for observational constraints.

The signal of F5 shown in Figure 7 is particularly promising in this context. Results from [51] indicate uncertainties on skewness measurements of $\sim$ 5–15% for selected galaxy samples at comparable redshifts (see Fig. 12 therein). Given that those measurements were based on a survey covering approximately 155 square degrees, surveys with larger coverage should provide much tighter constraints on skewness. Confronting such observational data with s_3 predictions for GR and F5 from mock galaxy catalogs mimicking the observations could potentially verify the validity of the models with high significance.

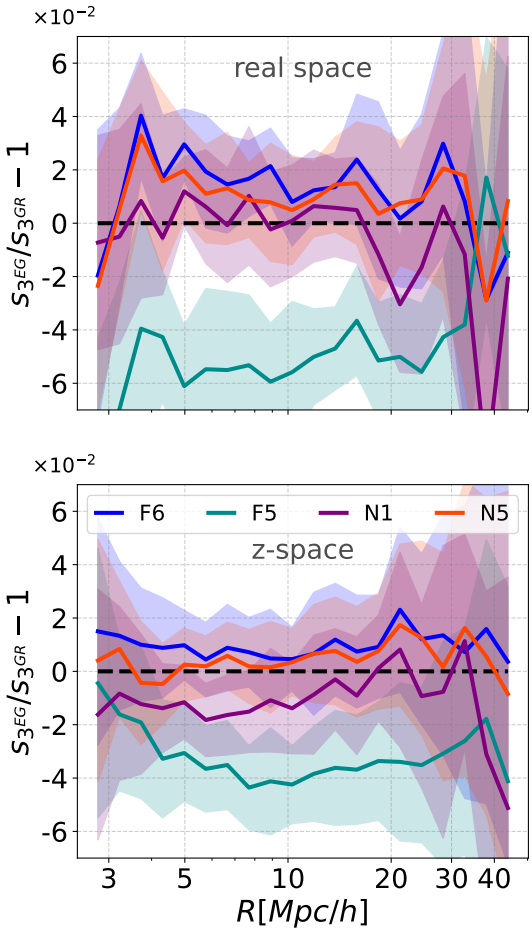

FIG. 7. Relative differences in the reduced third-order cumulant s_3 between EG models and GR for the full galaxy sample, comparing real (top) and redshift space (bottom) at $z = 0.5$.

We find qualitatively similar trends and signals for higher-order reduced cumulants $s_{4,5}$ for the mock galaxies. In fact, the differences between GR and extended gravity models grow larger for $J = 4$ and $J = 5$; however, the associated uncertainties also increase more rapidly reducing the overall significance. Therefore, we focus our analysis on s_3.

Up to this point, all results discussed refer to the full galaxy sample including both central and satellite galaxies. A notably different behavior of the skewness signal arises when considering only central galaxies in redshift space.

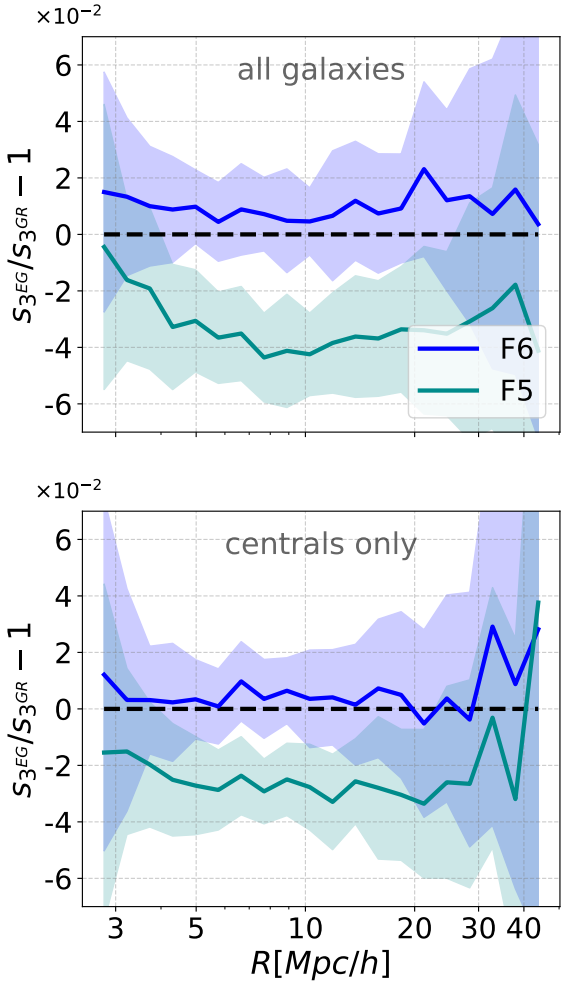

FIG. 8. Relative redshift-space differences in the reduced third-order cumulant s_3 between EG models and GR for mock galaxies, comparing the full sample (top) and centrals-only subsample (bottom), both at $z = 0.5$.

Figure 8 shows that the skewness signal for the full galaxy sample is stronger than that for the centrals-only at most scales. At first glance, this behavior might seem surprising, since satellite galaxies, which typically reside within the virial radius of their host halos, are expected to contribute additional signal primarily at scales smaller than the average halo size, roughly $\lesssim 2\,h^{-1}$Mpc [28, 52, 53].

The strengthening of deviations from GR for the full sample is likely related to the enhanced small-scale nonlinearities introduced by the satellite population, particularly within the one-halo regime.[2]

The observed all-scale increase of the F5 deviation in the full sample is a manifestation of the nature of the statistics used. Although the effects of satellites are most prominent at small distances, we use volume-averaged correlation functions in this work. Consequently, larger scales also incorporate contributions from smaller scales, allowing the satellite effect to propagate beyond the one-halo regime. We find that this effect persists at all redshifts considered, i.e $z = \{0, 0.3, 0.5\}$.

Another interesting feature we find is that for centrals-only galaxies, s_3 is smaller in real space than in redshift space at all scales, which is very distinct behavior than for both DM and halos. This inversion is indeed observed consistently across all gravity models we consider and contrasts with the behavior found for dark matter and halos, where real-space skewness typically exceeds redshift-space skewness. This can be interpreted in two ways: either adding RSD in centrals-only galaxies increases the nonlinearities, or ignoring satellite galaxies in real space decreases the skewness at all scales, but primarily around the one-halo regime, whereas in redshift space the effect is distributed more uniformly across scales. The first interpretation highlights that the centrals-only sample is much more selective compared to the full galaxy sample, leading to weaker spatial correlations on scales below the mean inter-galaxy separation. This is primarily due to the halo exclusion effect: in a centrals-only sample, no two galaxies can occupy the same halo, which suppresses the one-halo contribution to clustering. In contrast, including satellites reintroduces close galaxy pairs, reconstructing the missing small-scale clustering power associated with the one-halo term. The velocity field encodes additional clustering information — for example, from satellites that are absent in the centrals-only sample. Therefore, the redshift-space the amplitude is boosted relative to real space.

In contrast, for the full galaxy sample, small-scale motions primarily act to decrease the strong skewness signal originating from satellites, which are now included in the catalog. We find that the aforementioned effects are present also at other redshifts, except for $R < 6\,h^{-1}$Mpc at $z = 0$ for the centrals-only sample, where the real-space skewness becomes approximately 10% larger than in redshift space for the F5 model.

The second interpretation is more straightforward. Ignoring satellite galaxies in real space lowers the skewness primarily at the one-halo scale where satellites reside. However, due to volume averaging, this suppression also partially influences larger scales. In redshift space, much of the small-scale signal from satellites is suppressed by RSD. Hence, removing satellites does not induce as strong a scale-dependent change. Nonetheless, the additional nonlinearities encoded within the velocity field still enhance the skewness in the centrals-only

[2] In $f(R)$ gravity, satellites do not experience fundamentally different forces compared to centrals on small scales, due to efficient screening within dense halo environments. However, their presence enhances the small-scale clustering signal, particularly within the one-halo regime, and thus indirectly contributes to the observed skewness deviations.

redshift-space sample.

C. Universality of Real to Redshift-Space Transition

Our results, both for the dark matter density field and for halo clustering, indicate that the ratio between skewness in redshift and real space exhibits a very close agreement across all EG models and the fiducial GR case. This finding provides a promising opportunity for predicting higher-order redshift-space clustering cumulants in EG scenarios based solely on real-space predictions. We discuss this feature further in the Appendix, where we additionally compare one example of the s_3 ratios with the ratios of the volume-averaged correlation functions $\bar{\xi}_2$ and $\bar{\xi}_3$, from which the skewness s_3 is constructed. While these results hold robustly for the dark matter and halo populations the situation is different for galaxies. Here, the agreement between redshift-to-real-space skewness ratios across different gravity models is limited to specific cases and redshifts, largely due to the probabilistic nature of the halo occupation distribution (HOD) method used to construct the mock catalogs. In particular, the random assignment of satellite galaxies introduces additional small-scale peculiar velocities that are not fully correlated with the underlying matter density field. This synthetic velocity component partially decouples the galaxy clustering from the local density environment, making it more difficult to preserve the universality of the redshift-to-real-space ratio observed for dark matter haloes [see *e.g.* 54, 55].

V. SUMMARY AND CONCLUSIONS

In this work, we have investigated the skewness of the three-dimensional dark matter and galaxy density fields across different redshifts and gravity scenarios, focusing on the impact of redshift-space distortions. The comparison between real and redshift space revealed several remarkable trends, which we summarize below:

- For dark matter density, the skewness s_3 is suppressed in redshift space relative to real space at all scales and for all gravity models (Figure 3).
- For halos, skewness is not universally suppressed in redshift space (Figure 4). At large scales, s_3 is enhanced due to the Kaiser squashing effect, which is damped in dark matter particles by random motions.
- Differences between GR and EG models are suppressed in redshift space compared to real space, especially for halos. This significantly limits the utility of skewness as a gravity diagnostic for $f(R)$ and nDGP models.
- Up to $\sim$ 5% deviations between GR and EG are found for halos in real space (Figure 5), but redshift-space differences are generally smaller ($\sim$ 2%) and only marginally significant at $\sim 1\sigma$.
- For the H1 halo sample (number density of 10^{-3} $(\mathrm{Mpc}/h)^{-3}$), deviations in EG from GR decrease with redshift. Interestingly, the F6 model shows the strongest deviations for H1, whereas F5 dominates for the sparser H2 and H3 samples (Figure 6).
- For galaxies, nDGP models do not show significant deviations from GR in redshift space (Figure 7). In contrast, the F5 $f(R)$ model maintains a $\sim$ 2–4% deviation.
- Galaxies in redshift space can reveal stronger EG signals than halos. Comparing Figure 7 and Figure 5 shows that the F5 deviation is about twice as large for galaxies than for halos with similar number densities.
- Separating out central-only galaxies from the full sample reveals additional features (Figure 8). Samples containing both centrals and satellites show stronger skewness signals for F5, although for the F6 model the difference is negligible.
- The ratio of skewness between redshift and real space appears nearly universal across gravity models, especially for $f(R)$ (Figure 4). This glimpse of universality is discussed in the Appendix, where we find that while s_3 ratios match well across models, the underlying $\bar{\xi}_2$ and $\bar{\xi}_3$ ratios differ.

Our analysis demonstrates significant and complex effects of redshift-space distortions on higher-order clustering statistics. For the dark matter density field, skewness is universally suppressed in redshift space due to virial small-scale motions generating the Fingers-of-God effect. Halo populations exhibit a more nuanced behavior, with suppressed skewness at small scales transitioning into enhanced skewness at intermediate scales ($\sim 10\,h^{-1}$Mpc) due to coherent large-scale infall motions (Kaiser squashing).

Critically, the differences in skewness between GR and EG scenarios, prominent in real space, become much less pronounced in redshift space. Nevertheless, the F5 $f(R)$ model maintains a statistically significant deviation from GR at approximately 4% level for galaxy samples, suggesting potential observational detectability, particularly with the current new stage and large-scale galaxy surveys.

Interestingly, we uncover that the ratio of skewness between redshift and real spaces is remarkably stable across different gravity models. This approximate universality implies that real-space clustering predictions can be effectively used to estimate redshift-space skewness, greatly simplifying theoretical modeling for EG scenarios.

Furthermore, our study reveals enhanced sensitivity of galaxy skewness compared to halo skewness in redshift

space, especially when satellite galaxies are included. Satellites thus provide critical clustering information, reinforcing their importance in cosmological analyses aiming to test gravity models.

In summary, our findings highlight both challenges and opportunities in using higher-order clustering statistics to test gravity. While redshift-space distortions complicate direct interpretations of deviations from GR, our results point to viable pathways for utilizing observational data, emphasizing the role of satellite galaxies and the robustness of skewness ratios as powerful diagnostic tools.

VI. ACKNOWLEDGMENTS

The authors would like to thank Enrique Gaztañaga for valuable and engaging discussions at the various stages of this project. This work is supported via the research project 'COLAB' funded by the National Science Center, Poland, under agreement number UMO-2020/39/B/ST9/03494

Appendix A: Real to Redshift-Space Transition

In this section, we discuss the ratio between statistics computed in redshift and real space. We define the ratio as:

$$\eta_J = \frac{X_{J,\mathrm{RSD}}}{X_{J,\mathrm{REAL}}}, \tag{A1}$$

where X_J can refer either to the volume-averaged correlation functions $\bar{\xi}_J$ or to the reduced cumulants s_J, depending on whether ratios of correlations or cumulants are being considered. The subscripts RSD and REAL denote redshift and real space, respectively.

We find that the ratios for cumulants are, in many cases, universal with respect to the underlying gravity model. In this Appendix, we focus on halo clustering, as constructing halo catalogs comparable to ours requires relatively few additional assumptions. The situation for mock galaxy samples is more complicated due to the necessity of modeling specific selection functions, stellar masses, and luminosities. At the same time, the redshift-space clustering of the smooth dark matter density field cannot be directly observed or easily connected to measurable quantities. Thus, halos offer the most practical case for investigating potential universal relations between configuration and redshift space. We find that in many cases, the ratio η_J agrees between gravity models, particularly for s_J, but not necessarily for $\bar{\xi}_J$. Here, the nDGP model family however generally perform much worse than $f(R)$, indicating much lower degree of universality.

Figure 9 shows an example for the H1 halo sample at $z = 0.3$, for $f(R)$ and GR. For this example, we additionally fit the ratio $\eta_J(R)$ with an empirical function:

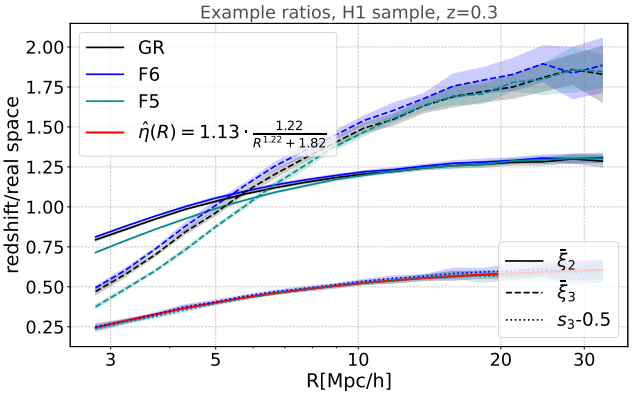

FIG. 9. Real-to-redshift space ratios at $z = 0.3$ for the H1 halo sample. The comparison is made in terms of $\bar{\xi}_2$ (solid lines), $\bar{\xi}_3$ (dashed), and s_3 (dotted) for the $f(R)$ gravity scenarios considered. The empirical fit for the s_3 case (Eq. A2) is marked in red. For clarity, the s_3 results have been offset downward by 0.5.

$$\hat{\eta}_J(R) = A \times \frac{R^\alpha}{R^\alpha + B}, \tag{A2}$$

where R is the smoothing scale in h^{-1}Mpc, and A, α, and B are free parameters. Due to finite-volume effects — i.e., cosmic variance limited by the simulation box size — the statistical uncertainties grow at larger scales. To ensure statistically robust fits, we restrict our analysis to scales $R < 50\,h^{-1}$Mpc. The ratios for other halo samples, redshifts, and higher orders are also well fitted by a function of the form given in Eq. A2. However, we do not provide the fitting parameters for all cases, as they strongly depend on the specific simulation details and do not carry physical meaning. Instead, we demonstrate the potential universality of η_J and provide an empirical fitting form that could be tested in future works.

The main findings regarding the real-to-redshift space transition can be summarized as follows:

- Differences in the η_J ratios between $f(R)$ models and GR are generally strongly suppressed when considering cumulants. We observe this effect consistently across all samples and redshifts.
- In most cases, one can reliably estimate the redshift-space skewness for both GR and extended gravity scenarios based on the universality of s_3.

The finding that the redshift-to-real-space ratio does not depend significantly on the assumed gravity model is promising for recovering real-space skewness from redshift-space observations. Furthermore, this effect can be tested against results from other simulations, helping to better constrain background fifth-force screening mechanisms in redshift space.

[1] P. J. E. Peebles, *The large-scale structure of the universe* (1980).

[2] S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, C. M. Baugh, C. S. Frenk, I. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, M. Colless, C. Collins, W. Couch, N. J. G. Cross, G. Dalton, V. R. Eke, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, K. Glazebrook, C. Jackson, A. Jenkins, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, B. A. Peterson, W. Sutherland, and K. Taylor, The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications, MNRAS **362**, 505 (2005), arXiv:astro-ph/0501174 [astro-ph].

[3] M. Tegmark, D. J. Eisenstein, M. A. Strauss, D. H. Weinberg, M. R. Blanton, J. A. Frieman, M. Fukugita, J. E. Gunn, A. J. S. Hamilton, G. R. Knapp, R. C. Nichol, J. P. Ostriker, N. Padmanabhan, W. J. Percival, D. J. Schlegel, D. P. Schneider, R. Scoccimarro, U. Seljak, H.-J. Seo, M. Swanson, A. S. Szalay, M. S. Vogeley, J. Yoo, I. Zehavi, K. Abazajian, S. F. Anderson, J. Annis, N. A. Bahcall, B. Bassett, A. Berlind, J. Brinkmann, T. Budavari, F. Castander, A. Connolly, I. Csabai, M. Doi, D. P. Finkbeiner, B. Gillespie, K. Glazebrook, G. S. Hennessy, D. W. Hogg, Ž. Ivezić, B. Jain, D. Johnston, S. Kent, D. Q. Lamb, B. C. Lee, H. Lin, J. Loveday, R. H. Lupton, J. A. Munn, K. Pan, C. Park, J. Peoples, J. R. Pier, A. Pope, M. Richmond, C. Rockosi, R. Scranton, R. K. Sheth, A. Stebbins, C. Stoughton, I. Szapudi, D. L. Tucker, D. E. vanden Berk, B. Yanny, and D. G. York, Cosmological constraints from the SDSS luminous red galaxies, Phys. Rev. D **74**, 123507 (2006), arXiv:astro-ph/0608632 [astro-ph].

[4] S. Alam, M. Ata, S. Bailey, F. Beutler, D. Bizyaev, J. A. Blazek, A. S. Bolton, J. R. Brownstein, A. Burden, C.-H. Chuang, J. Comparat, A. J. Cuesta, K. S. Dawson, D. J. Eisenstein, S. Escoffier, H. Gil-Marín, J. N. Grieb, N. Hand, S. Ho, K. Kinemuchi, D. Kirkby, F. Kitaura, E. Malanushenko, V. Malanushenko, C. Maraston, C. K. McBride, R. C. Nichol, M. D. Olmstead, D. Oravetz, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, M. Pellejero-Ibanez, W. J. Percival, P. Petitjean, F. Prada, A. M. Price-Whelan, B. A. Reid, S. A. Rodríguez-Torres, N. A. Roe, A. J. Ross, N. P. Ross, G. Rossi, J. A. Rubiño-Martín, S. Saito, S. Salazar-Albornoz, L. Samushia, A. G. Sánchez, S. Satpathy, D. J. Schlegel, D. P. Schneider, C. G. Scóccola, H.-J. Seo, E. S. Sheldon, A. Simmons, A. Slosar, M. A. Strauss, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, M. V. Magaña, J. A. Vazquez, L. Verde, D. A. Wake, Y. Wang, D. H. Weinberg, M. White, W. M. Wood-Vasey, C. Yèche, I. Zehavi, Z. Zhai, and G.-B. Zhao, The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample, MNRAS **470**, 2617 (2017), arXiv:1607.03155 [astro-ph.CO].

[5] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson, The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant, MNRAS **416**, 3017 (2011), arXiv:1106.3366 [astro-ph.CO].

[6] A. G. Adame, J. Aguilar, S. Ahlen, S. Alam, D. M. Alexander, M. Alvarez, O. Alves, A. Anand, U. Andrade, E. Armengaud, S. Avila, A. Aviles, H. Awan, B. Bahr-Kalus, S. Bailey, C. Baltay, A. Bault, J. Behera, S. BenZvi, A. Bera, F. Beutler, D. Bianchi, C. Blake, R. Blum, S. Brieden, A. Brodzeller, D. Brooks, E. Buckley-Geer, E. Burtin, R. Calderon, R. Canning, A. Carnero Rosell, R. Cereskaite, J. L. Cervantes-Cota, S. Chabanier, E. Chaussidon, J. Chaves-Montero, S. Chen, X. Chen, T. Claybaugh, S. Cole, A. Cuceu, T. M. Davis, K. Dawson, A. de la Macorra, A. de Mattia, N. Deiosso, A. Dey, B. Dey, Z. Ding, P. Doel, J. Edelstein, S. Eftekharzadeh, D. J. Eisenstein, A. Elliott, P. Fagrelius, K. Fanning, S. Ferraro, J. Ereza, N. Findlay, B. Flaugher, A. Font-Ribera, D. Forero-Sánchez, J. E. Forero-Romero, C. S. Frenk, C. Garcia-Quintero, E. Gaztañaga, H. Gil-Marín, S. Gontcho a Gontcho, A. X. Gonzalez-Morales, V. Gonzalez-Perez, C. Gordon, D. Green, D. Gruen, R. Gsponer, G. Gutierrez, J. Guy, B. Hadzhiyska, C. Hahn, M. M. S. Hanif, H. K. Herrera-Alcantar, K. Honscheid, C. Howlett, D. Huterer, V. Iršič, M. Ishak, S. Juneau, N. G. Karaçaylı, R. Kehoe, S. Kent, D. Kirkby, A. Kremin, A. Krolewski, Y. Lai, T. W. Lan, M. Landriau, D. Lang, J. Lasker, J. M. Le Goff, L. Le Guillou, A. Leauthaud, M. E. Levi, T. S. Li, E. Linder, K. Lodha, C. Magneville, M. Manera, D. Margala, P. Martini, M. Maus, P. McDonald, L. Medina-Varela, A. Meisner, J. Mena-Fernández, R. Miquel, J. Moon, S. Moore, J. Moustakas, E. Mueller, A. Muñoz-Gutiérrez, A. D. Myers, S. Nadathur, L. Napolitano, R. Neveux, J. A. Newman, N. M. Nguyen, J. Nie, G. Niz, H. E. Noriega, N. Padmanabhan, E. Paillas, N. Palanque-Delabrouille, J. Pan, S. Penmetsa, W. J. Percival, M. M. Pieri, M. Pinon, C. Poppett, A. Porredon, F. Prada, A. Pérez-Fernández, I. Pérez-Ràfols, D. Rabinowitz, A. Raichoor, C. Ramírez-Pérez, S. Ramirez-Solano, M. Rashkovetskyi, C. Ravoux, M. Rezaie, J. Rich, A. Rocher, C. Rockosi, N. A. Roe, A. Rosado-Marin, A. J. Ross, G. Rossi, R. Ruggeri, V. Ruhlmann-Kleider, L. Samushia, E. Sanchez, C. Saulder, E. F. Schlafly, D. Schlegel, M. Schubnell, H. Seo, A. Shafieloo, R. Sharples, J. Silber, A. Slosar, A. Smith, D. Sprayberry, T. Tan, G. Tarlé, P. Taylor, S. Trusov, L. A. Ureña-López, R. Vaisakh, D. Valcin, F. Valdes, M. Vargas-Magaña, L. Verde, M. Walther, B. Wang, M. S. Wang, B. A. Weaver, N. Weaverdyck, R. H. Wechsler, D. H. Weinberg, M. White, J. Yu, Y. Yu, S. Yuan, C. Yèche, E. A. Zaborowski, P. Zarrouk, H. Zhang, C. Zhao, R. Zhao, R. Zhou, and T. Zhuang, DESI 2024 VI: cosmological constraints from the measurements of baryon acoustic oscillations, J. Cosmology Astropart. Phys. **2025**, 021 (2025), arXiv:2404.03002 [astro-ph.CO].

[7] W. A. Hellwing, B. Li, C. S. Frenk, and S. Cole, Hierarchical clustering in chameleon f(R) gravity, MNRAS **435**, 2806 (2013), arXiv:1305.7486 [astro-ph.CO].

[8] W. A. Hellwing, K. Koyama, B. Bose, and G.-B. Zhao, Revealing modified gravity signals in matter and halo hierarchical clustering, Phys. Rev. D **96**, 023515 (2017), arXiv:1703.03395 [astro-ph.CO].

[9] P. Drozda, W. A. Hellwing, and M. Bilicki, Galaxy and halo angular clustering in Λ CDM and modified grav-

ity cosmologies, Phys. Rev. D **106**, 043513 (2022), arXiv:2204.05305 [astro-ph.CO].
[10] B. Bose and A. Taruya, The one-loop matter bispectrum as a probe of gravity and dark energy, J. Cosmology Astropart. Phys. **2018**, 019 (2018), arXiv:1808.01120 [astro-ph.CO].
[11] W. A. Hellwing, The skewness of z=0.5 redshift-space galaxy distribution in Modified Gravity, in *XXXIX Polish Astronomical Society Meeting*, Vol. 10, edited by K. Małek, M. Polińska, A. Majczyna, G. Stachowski, R. Poleski, Ł. Wyrzykowski, and A. Różańska (2020) pp. 315–322, arXiv:1912.13026 [astro-ph.CO].
[12] M. Cataneo, C. Uhlemann, C. Arnold, A. Gough, B. Li, and C. Heymans, The matter density PDF for modified gravity and dark energy with Large Deviations Theory, MNRAS **513**, 1623 (2022), arXiv:2109.02636 [astro-ph.CO].
[13] N. S. Sugiyama, D. Yamauchi, T. Kobayashi, T. Fujita, S. Arai, S. Hirano, S. Saito, F. Beutler, and H.-J. Seo, New constraints on cosmological modified gravity theories from anisotropic three-point correlation functions of BOSS DR12 galaxies, MNRAS **523**, 3133 (2023), arXiv:2302.06808 [astro-ph.CO].
[14] W. A. Hellwing, R. Juszkiewicz, and R. van de Weygaert, Hierarchy of N-point functions in the ΛCDM and ReBEL cosmologies, Phys. Rev. D **82**, 103536 (2010), arXiv:1008.3930 [astro-ph.CO].
[15] R. E. Fazolo, L. Amendola, and H. Velten, Skewness as a test of dark energy perturbations, Phys. Rev. D **105**, 103521 (2022).
[16] R. Scoccimarro, Redshift-space distortions, pairwise velocities, and nonlinearities, Phys. Rev. D **70**, 083007 (2004), arXiv:astro-ph/0407214 [astro-ph].
[17] N. Kaiser, Clustering in real space and in redshift space, MNRAS **227**, 1 (1987).
[18] C. Hernández-Aguayo, J. Hou, B. Li, C. M. Baugh, and A. G. Sánchez, Large-scale redshift space distortions in modified gravity theories, MNRAS **485**, 2194 (2019), arXiv:1811.09197 [astro-ph.CO].
[19] G. Valogiannis, R. Bean, and A. Aviles, An accurate perturbative approach to redshift space clustering of biased tracers in modified gravity, J. Cosmology Astropart. Phys. **2020**, 055 (2020), arXiv:1909.05261 [astro-ph.CO].
[20] J. Kuruvilla and C. Porciani, On the streaming model for redshift-space distortions, MNRAS **479**, 2256 (2018), arXiv:1710.09379 [astro-ph.CO].
[21] D. Kodwani and H. Desmond, Screened fifth forces in parity-breaking correlation functions, Phys. Rev. D **100**, 064030 (2019), arXiv:1904.12310 [astro-ph.CO].
[22] P. Arnalte-Mur, W. A. Hellwing, and P. Norberg, Real- and redshift-space halo clustering in f(R) cosmologies, MNRAS **467**, 1569 (2017), arXiv:1612.02355 [astro-ph.CO].
[23] J. Einasto, A. Klypin, G. Hütsi, L.-J. Liivamägi, and M. Einasto, Evolution of skewness and kurtosis of cosmic density fields, A&A **652**, A94 (2021), arXiv:2011.13292 [astro-ph.CO].
[24] M. Cautun, E. Paillas, Y.-C. Cai, S. Bose, J. Armijo, B. Li, and N. Padilla, The Santiago-Harvard-Edinburgh-Durham void comparison - I. SHEDding light on chameleon gravity tests, MNRAS **476**, 3195 (2018), arXiv:1710.01730 [astro-ph.CO].
[25] B. Li, G.-B. Zhao, R. Teyssier, and K. Koyama, ECOSMOG: an Efficient COde for Simulating MOdified Gravity, J. Cosmology Astropart. Phys. **2012**, 051 (2012), arXiv:1110.1379 [astro-ph.CO].
[26] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright, Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results, ApJS **208**, 19 (2013), arXiv:1212.5226 [astro-ph.CO].
[27] P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu, The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores, ApJL **762**, 109 (2013), arXiv:1110.4372 [astro-ph.CO].
[28] S. Alam, C. Arnold, A. Aviles, R. Bean, Y.-C. Cai, M. Cautun, J. L. Cervantes-Cota, C. Cuesta-Lazaro, N. C. Devi, A. Eggemeier, S. Fromenteau, A. X. Gonzalez-Morales, V. Halenka, J.-h. He, W. A. Hellwing, C. Hernández-Aguayo, M. Ishak, K. Koyama, B. Li, A. de la Macorra, J. Meneses Rizo, C. Miller, E.-M. Mueller, G. Niz, P. Ntelis, M. Rodríguez Otero, C. G. Sabiu, Z. Slepian, A. Stark, O. Valenzuela, G. Valogiannis, M. Vargas-Magaña, H. A. Winther, P. Zarrouk, G.-B. Zhao, and Y. Zheng, Towards testing the theory of gravity with DESI: summary statistics, model predictions and future simulation requirements, J. Cosmology Astropart. Phys. **2021**, 050 (2021), arXiv:2011.05771 [astro-ph.CO].
[29] M. White, M. Blanton, A. Bolton, D. Schlegel, J. Tinker, A. Berlind, L. da Costa, E. Kazin, Y. T. Lin, M. Maia, C. K. McBride, N. Padmanabhan, J. Parejko, W. Percival, F. Prada, B. Ramos, E. Sheldon, F. de Simoni, R. Skibba, D. Thomas, D. Wake, I. Zehavi, Z. Zheng, R. Nichol, D. P. Schneider, M. A. Strauss, B. A. Weaver, and D. H. Weinberg, The Clustering of Massive Galaxies at z ~0.5 from the First Semester of BOSS Data, ApJL **728**, 126 (2011), arXiv:1010.4915 [astro-ph.CO].
[30] C. Hernández-Aguayo, J. Hou, B. Li, C. M. Baugh, and A. G. Sánchez, Large-scale redshift space distortions in modified gravity theories, MNRAS **485**, 2194 (2019), arXiv:1811.09197 [astro-ph.CO].
[31] J. E. García-Farieta, W. A. Hellwing, S. Gupta, and M. Bilicki, Probing gravity with redshift-space distortions: Effects of tracer bias and sample selection, Phys. Rev. D **103**, 103524 (2021), arXiv:2103.14019 [astro-ph.CO].
[32] J. Khoury and A. Weltman, Chameleon cosmology, Phys. Rev. D **69**, 044026 (2004), arXiv:astro-ph/0309411 [astro-ph].
[33] P. Brax, C. van de Bruck, A.-C. Davis, and D. J. Shaw, f(R) gravity and chameleon theories, Phys. Rev. D **78**, 104021 (2008), arXiv:0806.3415 [astro-ph].
[34] A. I. Vainshtein, To the problem of nonvanishing gravitation mass, Physics Letters B **39**, 393 (1972).
[35] E. Babichev and C. Deffayet, An introduction to the Vainshtein mechanism, Classical and Quantum Gravity **30**, 184001 (2013), arXiv:1304.7240 [gr-qc].
[36] E. Gaztanaga and F. Bernardeau, The skewness and kurtosis of the projected density distribution function: validity of perturbation theory, A&A **331**, 829 (1998), arXiv:astro-ph/9707095 [astro-ph].
[37] A. Pollo, Gravitational Instability and 2-D Galaxy Surveys, Acta Astron. **47**, 413 (1997).

[38] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Large-scale structure of the Universe and cosmological perturbation theory, Phys. Rep. **367**, 1 (2002), arXiv:astro-ph/0112551 [astro-ph].

[39] E. Gaztanaga, High-Order Galaxy Correlation Functions in the APM Galaxy Survey, MNRAS **268**, 913 (1994), arXiv:astro-ph/9309019 [astro-ph].

[40] F. Bernardeau, The effects of smoothing on the statistical properties of large-scale cosmic fields., A&A **291**, 697 (1994), arXiv:astro-ph/9403020 [astro-ph].

[41] L. Amendola and S. Tsujikawa, *Dark Energy: Theory and Observations* (2010).

[42] P. Schneider, *Extragalactic Astronomy and Cosmology: An Introduction* (2015).

[43] C. M. Baugh, E. Gaztanaga, and G. Efstathiou, A comparison of the evolution of density fields in perturbation theory and numerical simulations - II. Counts-in-cells analysis, MNRAS **274**, 1049 (1995), arXiv:astro-ph/9408057 [astro-ph].

[44] E. Hivon, F. R. Bouchet, S. Colombi, and R. Juszkiewicz, Redshift distortions of clustering: a Lagrangian approach., A&A **298**, 643 (1995), arXiv:astro-ph/9407049 [astro-ph].

[45] W. H. Press and P. Schechter, Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation, ApJL **187**, 425 (1974).

[46] R. K. Sheth, H. J. Mo, and G. Tormen, Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes, MNRAS **323**, 1 (2001), arXiv:astro-ph/9907024 [astro-ph].

[47] J. N. Fry and E. Gaztanaga, Biasing and Hierarchical Statistics in Large-Scale Structure, ApJL **413**, 447 (1993), arXiv:astro-ph/9302009 [astro-ph].

[48] W. A. Hellwing, K. Koyama, B. Bose, and G.-B. Zhao, Revealing modified gravity signals in matter and halo hierarchical clustering, Phys. Rev. D **96**, 023515 (2017), arXiv:1703.03395 [astro-ph.CO].

[49] D. Shi, B. Li, J. Han, L. Gao, and W. A. Hellwing, Exploring the liminality: properties of haloes and subhaloes in borderline f(R) gravity, MNRAS **452**, 3179 (2015), arXiv:1503.01109 [astro-ph.CO].

[50] S. Gupta, W. A. Hellwing, M. Bilicki, and J. E. García-Farieta, Universality of the halo mass function in modified gravity cosmologies, Phys. Rev. D **105**, 043538 (2022), arXiv:2112.03699 [astro-ph.CO].

[51] M. Wolk, H. J. McCracken, S. Colombi, J. N. Fry, M. Kilbinger, P. Hudelot, Y. Mellier, and O. Ilbert, Evolution of hierarchical clustering in the CFHTLS-Wide since z $\sim$ 1, MNRAS **435**, 2 (2013), arXiv:1301.3301 [astro-ph.CO].

[52] Z. Zheng, A. A. Berlind, D. H. Weinberg, A. J. Benson, C. M. Baugh, S. Cole, R. Davé, C. S. Frenk, N. Katz, and C. G. Lacey, Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies, ApJL **633**, 791 (2005), arXiv:astro-ph/0408564 [astro-ph].

[53] M. T. Busha, R. H. Wechsler, P. S. Behroozi, B. F. Gerke, A. A. Klypin, and J. R. Primack, Statistics of Satellite Galaxies around Milky-Way-like Hosts, ApJL **743**, 117 (2011), arXiv:1011.6373 [astro-ph.CO].

[54] J. L. Tinker, D. H. Weinberg, and Z. Zheng, Redshift-space distortions with the halo occupation distribution - I. Numerical simulations, MNRAS **368**, 85 (2006), arXiv:astro-ph/0501029 [astro-ph].

[55] C. Hikage and K. Yamamoto, Impacts of satellite galaxies on the redshift-space distortions, J. Cosmology Astropart. Phys. **2013**, 019 (2013), arXiv:1303.3380 [astro-ph.CO].

4.2 Summary

While the differences between used extended gravity scenarios and GR are suppressed in redshift space, there are also promising results. Firstly, the deviations from GR are not reduced completely in z-space. It can be taken advantage of especially when employing larger catalogs, since the current uncertainties are still relatively large. In the era of large observational data, employing more detailed catalogs is highly motivated. Also as noticed in previous work, higher orders reveal bigger relative deviations from GR, but these also suffer from larger errors. Increasing the amount of information by using denser simulations should improve also high-order measures. Additionally, in this work about $3D$ clustering we found the signals coming from galaxy statistics to be twice as strong, compared to halos. This is promising, hence the decisive gravity tests using clustering will be naturally performed with galaxies rather than halos.

The aspect worth pointing out about studying the cosmic web with central moments-based statistics, is that even if these allow for low-cost investigation of higher orders, the information gained from J-point correlation functions is richer than from averaged correlation functions. Optimally one would use a method to obtain more information than from higher order $\bar{\xi}_J$, simultaneously keeping similar computational cost. Such a method requires generalization of counts-in-cells prescription and it is the subject of the next work.

Chapter 5

Paper II: Shape-dependent clustering with ellipsoid correlation function

5.1 Introduction

One of the solutions for better discriminating between cosmic web evolved through different gravity scenarios is to employ the statistics which allows for extracting more information. While the averaged correlation functions provide low-cost access to higher-order statistics of LSS, other statistics such as N-point correlation functions encode also the information about directional inter-object separations. Another example is anisotropic 2-point correlation function, splitting the inter-object separation on component parallel and perpendicular to line of sight. This is a convenient solution while working in redshift space. An equivalent approach can be used for volume averaged correlation functions by generalizing the counts-in-cells method. Instead of counts in spheres one can perform counts within ellipsoids with defined semi-axes perpendicular and parallel to line of sight. This enables for low-cost computation of any statistical order of LSS central moments and simultaneously gives access to additional information related with anisotropic counts. As for a new method, the results obtained from anisotropic averaged correlation functions firstly require careful interpretation and testing. In this part we discuss the developed anisotropic functions in terms of usage and link them with previously used statistics. Additionally we analyze how different growth of cosmic structure affects our statistics.

Anisotropic Counts-in-Cells in Redshift Space: A New Route to Cosmological Constraints from Galaxy Surveys

Paweł Drozda,* Wojciech A. Hellwing, and Maciej Bilicki
Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46,02-668 Warsaw, Poland
(Dated: June 3, 2025)

We introduce a novel extension of the volume-averaged correlation function (VACF) framework by replacing the traditional spherical smoothing kernels with anisotropic, ellipsoidal windows. This generalized approach enables the study of shape-dependent clustering statistics and captures directional information encoded in large-scale structure, particularly in redshift space where galaxy distribution is distorted by peculiar velocities. We define and compute ellipsoidal VACFs $\bar{\xi}_J(r_\parallel, r_\perp)$ and the corresponding reduced cumulants $s_J(r_\parallel, r_\perp)$, allowing for joint sensitivity to both scale and anisotropy across arbitrary statistical order J. Using a suite of COLA N-body simulations spanning a grid of cosmologies with varying Ω_M and σ_8, we analyze the behavior of ellipsoidal VACFs and cumulants in both real and redshift space. We find that the shape of the smoothing kernel that maximizes the clustering signal depends strongly on the redshift-space distortion regime: spherical in real space, prolate in the Fingers-of-God-dominated regime, and oblate in the Kaiser squashing-dominated regime. While the standard VACF amplitude is mainly sensitive to σ_8, the shape-dependence of redshift-space skewness shows a coherent response to the combined growth parameter $f\sigma_8$, with a typical sensitivity at the 1–3σ level between neighboring models. Our results demonstrate that ellipsoidal VACFs offer a computationally efficient and information-rich generalization of counts-in-cells analysis, with promising applications to galaxy survey data, halo catalogs, and cosmological tests of gravity beyond ΛCDM.

I. INTRODUCTION

The large-scale structure (LSS) of the Universe originates from the gravitational amplification of primordial density perturbations, modulated by baryonic processes in an expanding cosmological background. The current standard model of cosmology, ΛCDM, successfully describes a broad range of observations across cosmic time. Nonetheless, persistent tensions—such as those concerning the Hubble constant and the growth rate of structure—motivate the investigation of extensions to the standard model, including modifications to gravity and the nature of dark energy.

The two-point correlation function (2PCF) and its Fourier counterpart, the power spectrum, remain central tools in the statistical analysis of LSS. While powerful, these second-order statistics are inherently limited in their ability to capture non-Gaussian features and the full nonlinear complexity of cosmic structure formation. To address this, various higher-order statistics—such as three-point correlation function and bispectrum , marked correlation functions, and hierarchical cumulants—have been developed, albeit at considerable computational cost [*e.g.* 1–3].

The use of correlation functions in cosmology was pioneered by Peebles in the 1970s and 1980s [4]. Building on the fair-sample hypothesis, volume-averaged correlation functions (VACFs) were later introduced as the moments of the counts-in-cells (CiC) distribution, offering a more tractable alternative to high-order n-point functions. This framework was significantly advanced in the 1990s, mostly through the work of Juszkiewicz, Bernardeau, Colombi, and Gaztañaga, who demonstrated that VACFs and their reduced combinations could be used to test the gravitational instability paradigm and to constrain cosmological parameters [5–8]. In subsequent decades, CiC-based statistics have been widely employed to explore a range of alternative cosmological scenarios, including modified gravity models [*e.g.* 9–12].

In practice, VACFs are computed as the central moments of tracer counts within smoothing volumes of varying size—typically spherical. This traditional implementation, while computationally efficient and conceptually clear, suffers from a key limitation: the use of spherical kernels imposes isotropy and suppresses directional information. As a result, it fails to capture the anisotropic features of the Cosmic Web, such as the elongated morphology of filaments or the flattened structure of walls [13–17]. Furthermore, in spectroscopic surveys, peculiar velocities introduce anisotropies along the line-of-sight (LOS), giving rise to redshift-space distortions (RSD). These distortions encode additional cosmological information—especially about the growth rate of structure—but their directional signature is diluted by isotropic averaging.

This issue is especially pressing in the era of precision cosmology. Galaxy spectroscopic surveys provide three-dimensional maps of the cosmic density field with high fidelity, particularly along the LOS. Past and current datasets such as 2dFGRS, SDSS, or GAMA, along with next-generation surveys including Euclid, DESI, and 4MOST, offer unprecedented volumes and sizes of spectroscopic samples [*e.g.* 18–22]. These data present a

* (pdrozda, hellwing, bilicki)@cft.edu.pl

unique opportunity to perform high-precision CiC analyses that fully exploit their anisotropic information content. However, doing so requires extending the traditional isotropic framework to account for LOS-dependent effects introduced by RSD.

While RSD signals have been extensively utilized in two-point statistics to constrain the growth rate and galaxy bias, their incorporation into higher-order CiC statistics has remained underdeveloped. A natural and long-overdue extension of the CiC formalism is to introduce anisotropic smoothing volumes that are sensitive to redshift-space effects. In this work, we address this gap by generalizing the VACF methodology to include ellipsoidal smoothing kernels with independently controlled axes parallel and perpendicular to the LOS. This extension allows us to explicitly probe shape-dependent clustering, capture anisotropies induced by cosmic velocities, and gain access to information otherwise suppressed by spherical averaging.

We apply this new estimator to a suite of N-body simulations based on PiCOLA (*Parallel COmoving Lagrangian Accelerator*) dynamics[23], exploring a grid of cosmological models with varying matter density Ω_M and fluctuation amplitude σ_8. By analyzing anisotropic VACFs and their associated cumulants in both real and redshift space, we uncover distinctive morphological patterns and RSD signatures that trace the growth of structure. Our results show that ellipsoidal VACFs provide a powerful and flexible tool for extracting anisotropic clustering information from spectroscopic surveys.

This paper is organized as follows. In Section II, we define the ellipsoidal VACF estimator and describe its numerical implementation. Section III presents the simulation data and construction of redshift-space catalogs. In Section IV, we present our results for real- and redshift-space VACFs and cumulants across a range of cosmological models. Finally, in Section V, we summarize our findings and discuss their implications for future analyses of large-scale structure.

II. SHAPE-DEPENDENT DENSITY FIELD STATISTICS

The starting point of our analysis is the matter density contrast field, defined as

$$\delta(\vec{x}) = \frac{\rho(\vec{x})}{\bar{\rho}} - 1, \tag{1}$$

where $\rho(\vec{x})$ is the local matter density at position $\vec{x}$, and $\bar{\rho}$ is the mean cosmic density.

To probe the statistical properties of the density field beyond the two-point level, we consider the volume-averaged correlation function (VACF) of order J, defined as

$$\bar{\xi}_J(R) \equiv \langle \bar{\delta}^J \rangle, \tag{2}$$

where $\bar{\delta}(\vec{x})$ denotes the density contrast field smoothed over a finite volume using a three-dimensional window function W_{3D},

$$\bar{\delta}(\vec{x}) = \int d^3x'\, \delta(\vec{x}')\, W_{3D}(\vec{x} - \vec{x}'). \tag{3}$$

In the standard formulation, W_{3D} is taken to be a spherically symmetric top-hat kernel of radius R [2, 4]. Under the assumption that the smoothed local density can be approximated by counting tracers (e.g. halos or galaxies) within spheres of radius R, the VACF can be estimated as the central moments of the CiC distribution. As stated earlier, in this work we generalize the VACF formalism by introducing an anisotropic smoothing kernel with ellipsoidal geometry. We define the ellipsoidal window function as

$$W_{3D}(\vec{x}; r_\parallel, r_\perp) = \begin{cases} \dfrac{3}{4\pi r_\parallel r_\perp^2}, & \text{if } \dfrac{x_\parallel^2}{r_\parallel^2} + \dfrac{x_{\perp,1}^2 + x_{\perp,2}^2}{r_\perp^2} \leq 1 \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

where $r_\parallel$ and $r_\perp$ denote the semi-axes of the ellipsoid aligned parallel and perpendicular to the LOS, respectively. The position vector $\vec{x}$ is decomposed as $\vec{x} = (x_\parallel, x_{\perp,1}, x_{\perp,2})$, where we assume symmetry in the plane perpendicular to the LOS, i.e. $r_{\perp,1} = r_{\perp,2} = r_\perp$. This assumption preserves azimuthal symmetry around the LOS and ensures that rotational orientation within the perpendicular plane does not affect the clustering measurement.

The resulting generalized VACF is thus a function of both ellipsoid axes:

$$\bar{\xi}_J(r_\parallel, r_\perp) \equiv \langle \bar{\delta}^J \rangle \quad \text{with } W_{3D} \text{ as in Eq. (4).} \tag{5}$$

In practice, we estimate $\bar{\xi}_J(r_\parallel, r_\perp)$ by computing the central moments of the distribution of tracer counts within ellipsoidal volumes defined by $(r_\parallel, r_\perp)$. From these raw moments, we further derive the connected moments and apply Poisson noise corrections following the formalism of Gaztanaga [8].

A schematic illustration of the procedure is shown in Fig. 1. For each pair $(r_\parallel, r_\perp)$, we compute counts-in-ellipsoids across the simulation volume, derive the moments, and assemble the ellipsoidal VACFs.[1]

In addition to the raw and connected moments, we compute reduced cumulants—also known as hierarchical amplitudes—defined as

$$S_J(r_\parallel, r_\perp) = \frac{\bar{\xi}_J(r_\parallel, r_\perp)}{\bar{\xi}_2^{J-1}(r_\parallel, r_\perp)}, \quad \text{for } J \geq 3. \tag{6}$$

[1] The corresponding code is publicly available at github.com/Pawel-96/Avcorr.

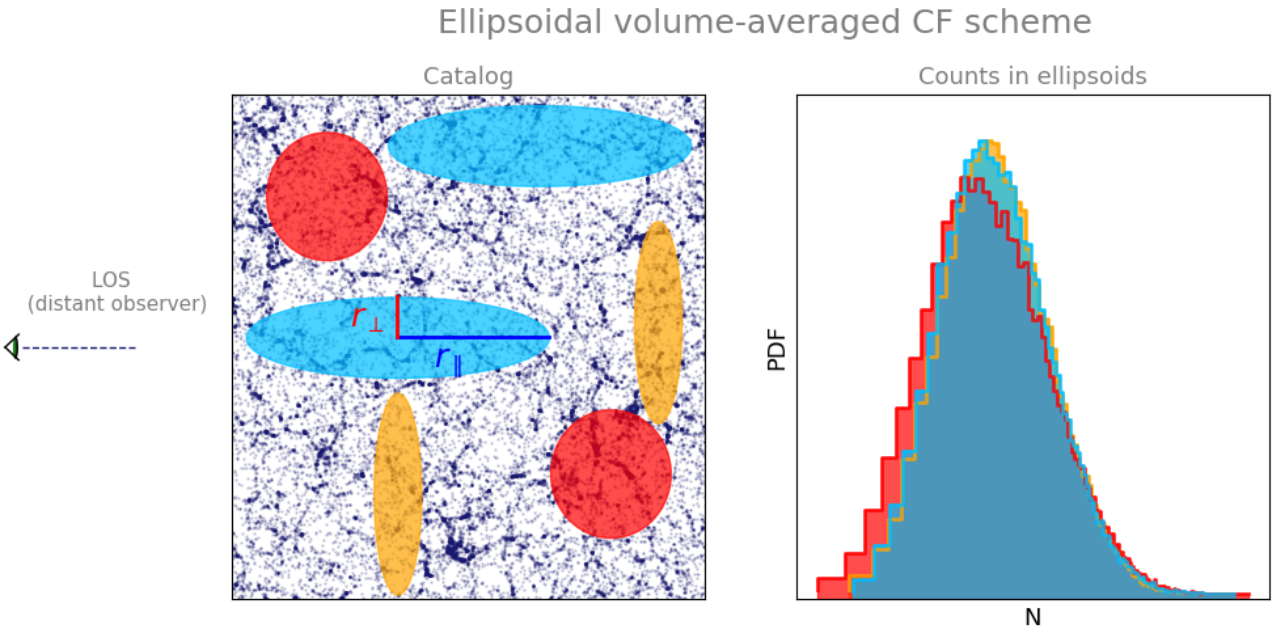

FIG. 1. Illustration of the ellipsoidal counts-in-cells procedure used to compute $\bar{\xi}_J(r_\parallel, r_\perp)$. For each pair of kernel axes $(r_\parallel, r_\perp)$, ellipsoidal shells are tiled across the simulation volume. The resulting count distribution is used to compute connected and Poisson-corrected central moments.

These normalized quantities remove the dominant scale dependence of $\bar{\xi}_2$ and highlight deviations from Gaussianity and linear growth. In particular, the reduced third-order cumulant or skewness, $S_3 = \bar{\xi}_3/\bar{\xi}_2^2$, is frequently used to test the predictions of gravitational instability and probe higher-order mode coupling,[5, 24]. In the weakly nonlinear regime, S_J are predicted to be nearly scale-independent [25] and are sensitive to both cosmological parameters and the shape of the initial power spectrum [26, 27].

III. DATA

This work is based on a new suite of gravity-only N-body simulations generated using the COLA (COmoving Lagrangian Acceleration) method [28], in particular we used the MG-COLA variant [29]. These simulations were specifically constructed to explore the impact of cosmological parameters on anisotropic clustering and higher-order statistics. They represent the first application of this COLA ensemble, which is optimized to strike a balance between computational efficiency and sufficient resolution for analyzing the matter density field.

The fiducial cosmology is based on the Planck 2018 results [30], corresponding to the TT,TE,EE+lowE+lensing+BAO solution from Table 2 of Planck Collaboration *et al.* [30]. The adopted cosmological parameters are:

- Hubble parameter: $h = 0.6766$,
- Baryon density: $\Omega_b = 0.049$,
- Total matter density: $\Omega_m = 0.3111$,
- Dark energy density: $\Omega_\Lambda = 0.6889$,
- Power spectrum normalization: $\sigma_8 = 0.8102$,
- Spectral index: $n_s = 0.9665$.

Each simulation evolves 1024^3 dark matter particles in a periodic box of size $L_{\rm box} = 500\,h^{-1}$Mpc, starting the time integration at redshift $z = 19$ and integrated using 52 global time steps. Forces are computed on a regular mesh with 1536^3 cells, yielding a spatial resolution of approximately 325 kpc/h. This level of resolution is well suited for studies of the continuous matter density field, though it is not sufficient for resolving halo substructure or detailed halo assembly histories. A dedicated analysis of halo and galaxy clustering using higher-resolution simulations will be presented in a follow-up work.

The full suite consists of 9 distinct background cosmologies, spanning $\pm 10\%$ variations in Ω_m and σ_8 around the fiducial model (see Table I). For each cosmology, we generate 5 independent realizations using different initial condition seeds, resulting in a total of 45 simulations. This design ensures robust statistical power for measuring redshift-space clustering and cosmological trends.

We analyze three redshift snapshots: $z = \{0.1, 0.2, 0.5\}$, using dark matter pseudo-particles as

TABLE I. Cosmological models used in this study, with specified values of σ_8, Ω_M, and the derived $f\sigma_8$. The fiducial model is denoted as `Cen`.

Model	σ_8	Ω_M	$f\sigma_8$
`Ls8COM`	0.7292	0.3111	0.3837
`Ls8HOM`	0.7292	0.3422	0.4043
`Cs8LOM`	0.8102	0.2800	0.4023
`Cen`	0.8102	0.3111	0.4263
`Cs8HOM`	0.8102	0.3422	0.4492
`Hs8LOM`	0.8912	0.2800	0.4425
`Hs8COM`	0.8912	0.3111	0.4689
`Hs8HOM`	0.8912	0.3422	0.4941

tracers. All statistics are computed in the distant observer approximation [31–33], without applying light-cone projection or relativistic corrections.

To facilitate interpretation and comparison with growth-based observables, we compute the derived combination

$$f\sigma_8 = f(\Omega_m)\,\sigma_8 \approx \Omega_m^{0.55}\sigma_8, \tag{7}$$

where $f \equiv d\ln D/d\ln a$ is the linear growth rate. This parameter is frequently used in redshift-space distortion analyses as a robust measure of structure growth [34–36].

In addition to real-space catalogs, we construct redshift-space versions of each snapshot using the distant observer approximation. For each tracer particle, the redshift-space position is computed as

$$\vec{x}_{\rm zspace} = \vec{x} + \frac{1+z_{\rm snap}}{H(z_{\rm snap})}\, v_\parallel\, \hat{e}_\parallel, \tag{8}$$

where $v_\parallel$ is the peculiar velocity component along the LOS and $\hat{e}_\parallel$ is the LOS unit vector. This transformation introduces anisotropies from both coherent infall (Kaiser effect) and small-scale dispersion (Finger-of-God effect) [37].

To average over orientation-dependent effects, we generate three redshift-space realizations for each snapshot by aligning the LOS direction with each Cartesian axis in turn. All redshift-space measurements are averaged over these three direction-projections and over 5 initial phase retaliations, following standard practice [38, 39].

IV. ANISOTROPIC COUNTS-IN-CELLS

We begin our analysis by examining the geometric and physical properties of ellipsoidal volume-averaged correlation functions (VACFs). As a starting point, we consider the second-order VACF, $\bar{\xi}_2(r_\parallel, r_\perp)$, measured in the real space for the fiducial cosmological model using dark matter particles. The result is shown in Figure 2.

To aid interpretation, we highlight two important geometric structures in the $(r_\parallel, r_\perp)$ plane:

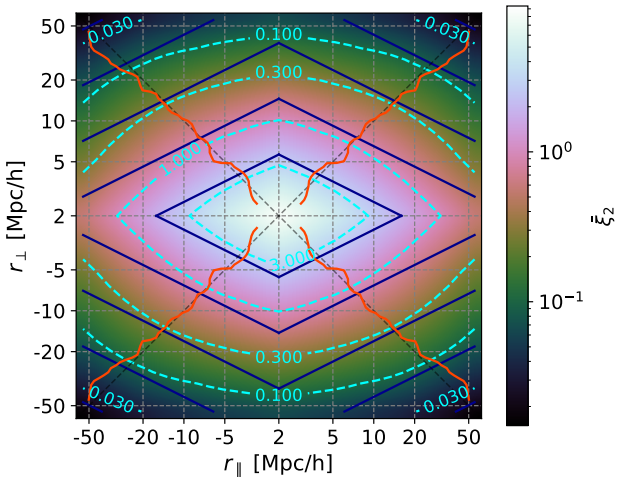

FIG. 2. Second-order ellipsoidal VACF $\bar{\xi}_2(r_\parallel, r_\perp)$ in real space at redshift $z = 0.1$. Dashed cyan lines denote iso-correlation contours; dark blue diamond-shaped curves indicate iso-volume contours ($4\pi/3 \cdot r_\parallel r_\perp^2 = \text{const}$); orange-red lines trace the ellipsoidal shapes $[r_\parallel, r_\perp]$ that maximize $\bar{\xi}_2$ at fixed volume.

- **Iso-volume contours** connect ellipsoidal kernels of constant volume, defined by $V = (4\pi/3) \cdot r_\parallel r_\perp^2$. These contours allow one to compare the clustering signal across different shapes while holding the enclosed volume fixed.

- **Iso-correlation contours** follow lines of constant VACF value, $\bar{\xi}_2(r_\parallel, r_\perp) = \text{const}$. Their geometry reveals how clustering strength depends jointly on the kernel size and anisotropy.

In an isotropic and homogeneous matter distribution, these two families of contours would be expected to align. Thus, any misalignment between them signals the presence of structure in the cosmic density field. Indeed, Figure 2 shows that the iso-correlation lines (dashed cyan) systematically deviate from the iso-volume contours (dark blue), reflecting the anisotropic nature of the Cosmic Web. As expected, the amplitude of $\bar{\xi}_2$ decreases with increasing scale. Along the diagonal $r_\parallel = r_\perp$, corresponding to spherical kernels, the VACF reduces to its standard isotropic form which we studied in [40]. However, for fixed volume, the clustering signal is consistently maximized for spherical kernels, as indicated by the orange-red curves. This behavior is in line with the statistical isotropy of over- and under-dense regions in real space: spherical kernels best isolate coherent structures, while elongated kernels tend to span across different environments, suppressing the variance in counts and hence reducing $\bar{\xi}_2$.

We also observe a symmetry in the VACF response to ellipsoidal shape: for fixed volume V, kernels with axis ratio $\eta = r_\parallel/r_\perp$ yield the same VACF as those with

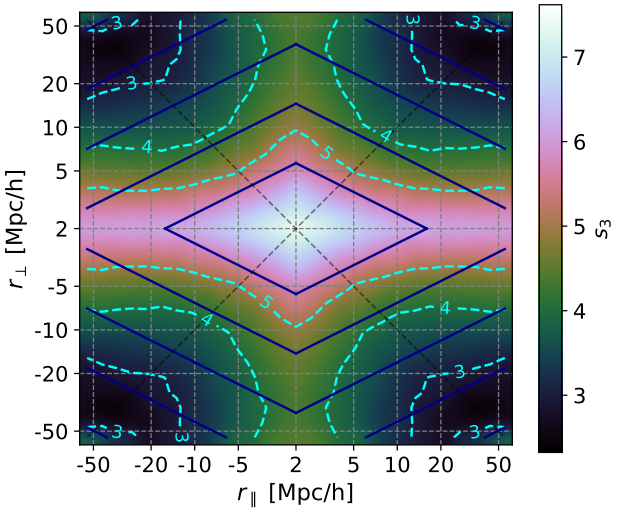

FIG. 3. Skewness S_3 of the density field in real space as a function of ellipsoidal kernel shape at redshift $z = 0.1$.

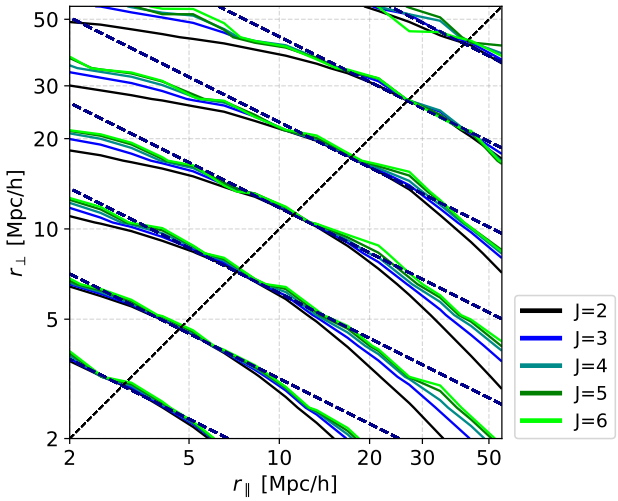

FIG. 4. Iso-value contours of $\bar{\xi}_J(r_\parallel, r_\perp)$ for orders $J = 2, ..., 6$, color-coded as indicated. Dashed dark blue lines mark iso-volume contours passing through $r_\parallel = r_\perp$. Only one quadrant is shown for clarity.

inverse ratio $1/\eta$,

$$\bar{\xi}_2(r_\parallel/r_\perp = \eta)\big|_V = \bar{\xi}_2(r_\parallel/r_\perp = 1/\eta)\big|_V. \quad (9)$$

This symmetry arises from the isotropy of real space, where no direction is privileged. Consequently, the VACF depends only on the kernel's shape and volume, not on its orientation with respect to the observer.

At higher orders $J > 2$, similar trends persist, though the statistical noise increases due to the growing sensitivity of higher-order moments to sample variance. Moreover, the location of maximum $\bar{\xi}_J$ at fixed volume begins to drift away from the diagonal as non-Gaussian features become more prominent.

To gain deeper insight, we now turn to the skewness statistic. Figure 3 shows the third-order reduced cumulant, S_3, for the same dark matter sample in real space.

In perturbation theory (PT), the reduced cumulants S_J are proportionality factors relating $\bar{\xi}_J$ to powers of $\bar{\xi}_2$, i.e., $S_J \propto \bar{\xi}_J/\bar{\xi}_2^{J-1}$ [2, 41]. In the linear regime, these ratios are scale-independent and determined by the gravitational growth dynamics [2]. Deviations from this behavior indicate nonlinear evolution or the breakdown of PT assumptions.

Several key features emerge in Figure 3. First, S_3 is the largest at small scales, where nonlinear gravitational clustering enhances higher-order correlations. Second, strong variations appear along the axes $r_\parallel = \text{const}$ and $r_\perp = \text{const}$, resulting in a cross-shaped structure. This behavior stems from the asymmetric response of $\bar{\xi}_3$ and $\bar{\xi}_2$ to changes in the ellipsoid's geometry.

In particular, we find that for fixed $r_\perp$ and increasing $r_\parallel \ll r_\perp$, the third-order moment $\bar{\xi}_3$ declines more rapidly than $\bar{\xi}_2^2$, leading to a local enhancement in S_3. A similar effect occurs for fixed $r_\parallel$ and small $r_\perp \ll r_\parallel$. This results in skewness enhancements along the vertical and horizontal axes. Interestingly, the effect is stronger along the horizontal axis ($r_\perp$), which reflects the geometry of the smoothing kernel: the ellipsoidal volume depends quadratically on $r_\perp$, making the statistic more sensitive to its variation.

This cross-like morphology in S_3 reflects a broader trend observed across higher-order VACFs. In Figure 4, we plot the iso-value contours of $\bar{\xi}_J(r_\parallel, r_\perp)$ for several orders $J = 2$ through 6.

As J increases, the iso-$\bar{\xi}_J$ contours converge toward the iso-volume lines (dashed blue diagonals in the Figure). This convergence explains why all reduced cumulants S_J exhibit similar cross-like features: the higher the order, the more closely $\bar{\xi}_J$ depends only on the total volume, with diminishing sensitivity to ellipsoid shape. For $J \to \infty$, one might expect $\bar{\xi}_J(r_\parallel, r_\perp)$ to become purely volume-dependent.

This behavior is somewhat counterintuitive. One might expect higher-order statistics—sensitive to rare events and extreme structures—to exhibit greater anisotropy. Yet our results suggest that the geometric dependence of $\bar{\xi}_J$ becomes increasingly isotropic with order. A possible explanation may lie in the slower evolution of higher-order nonlinearities, or in their reduced sensitivity to small-scale anisotropies, but further theoretical investigation is needed to clarify this effect.

A. Redshift-space effects

We now turn to the redshift space, the natural domain of galaxy spectroscopic surveys and the primary motivation for our anisotropic CiC estimator. In this context, the parameterization $[r_\parallel, r_\perp]$ acquires particu-

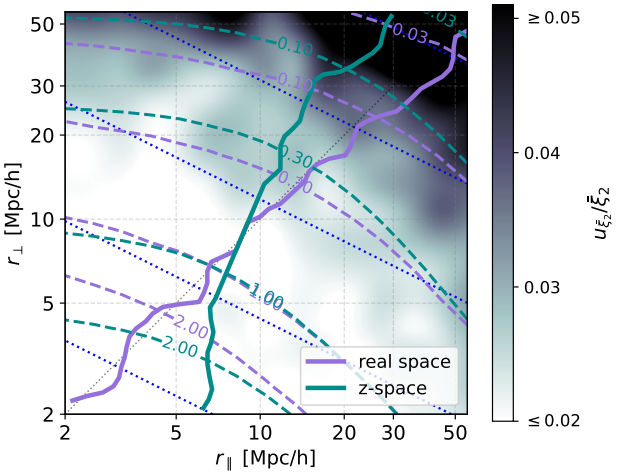

FIG. 5. Contour plot of dark matter VACF $\bar{\xi}_2(r_\parallel, r_\perp)$ in real (dashed violet) and redshift space (dashed cyan) at $z = 0.1$. Solid lines indicate kernel shapes that maximize $\bar{\xi}_2$ at constant ellipsoidal volume. Dotted blue lines denote iso-volume contours. The background color map shows the average relative uncertainty, $u_{\bar{\xi}_2}/\bar{\xi}_2$, combining real and redshift space errors.

lar relevance, as it allows us to separate the clustering signal along and across the LOS.

To facilitate comparison across different ellipsoidal volumes, we define an effective radius:

$$r_{\rm eff} \equiv \left(r_\parallel r_\perp^2\right)^{1/3}, \tag{10}$$

which corresponds to the radius of a sphere with the same volume as the ellipsoid.

Figure 5 shows iso-correlation contours of the second-order VACF, $\bar{\xi}_2(r_\parallel, r_\perp)$, in both real and redshift space at redshift $z = 0.1$. The solid lines indicate the ellipsoidal shapes that maximize $\bar{\xi}_2$ at fixed volume.

Several key trends emerge. In real space, the VACF is maximized by spherical kernels ($r_\parallel = r_\perp$), consistent with the statistical isotropy of density fluctuations. In redshift space, however, this symmetry is broken by redshift-space distortions (RSD), leading to a scale-dependent shift in the optimal kernel shape.

At small scales ($r_{\rm eff} \lesssim 10\ h^{-1}$Mpc), the maximal signal is achieved for prolate ellipsoids elongated along the LOS ($r_\parallel > r_\perp$). This reflects the dominance of the Fingers-of-God (FoG) effect [42, 43], where random virial motions elongate structures along the LOS, smearing the density field. At larger scales, the Kaiser squashing effect [32, 44] dominates: coherent infall velocities compress structures along the LOS, and the VACF is maximized for oblate ellipsoids with $r_\parallel < r_\perp$. The transition between these regimes occurs near $r_{\rm eff} \approx 12\ h^{-1}$Mpc, consistent with previous studies of redshift-space clustering [e.g., 31, 42, 45].

To build physical intuition, it is helpful to recall the standard configuration-space prediction for RSD in linear theory. In the Kaiser model[31], the redshift-space two-point correlation function $\xi^s(r, \mu)$ depends on the separation r and the cosine μ of the angle between the pair vector and the LOS:

$$\xi^s(r, \mu) = \sum_{\ell=0,2,4} \xi_\ell(r) L_\ell(\mu), \tag{11}$$

where $L_\ell(\mu)$ are Legendre polynomials and the multipoles $\xi_\ell(r)$ are given by:

$$\xi_0(r) = \left(1 + \frac{2\beta}{3} + \frac{\beta^2}{5}\right)\xi(r), \tag{12}$$

$$\xi_2(r) = \left(\frac{4\beta}{3} + \frac{4\beta^2}{7}\right)\left[\xi(r) - \bar{\xi}(r)\right], \tag{13}$$

$$\xi_4(r) = \frac{8\beta^2}{35}\left[\xi(r) + \frac{5}{2}\bar{\xi}(r) - \frac{7}{2}\bar{\bar{\xi}}(r)\right], \tag{14}$$

with $\beta = f/b$ the linear redshift-space distortion parameter. Here, $\xi(r)$ is the real-space two-point correlation function, while $\bar{\xi}(r)$ and $\bar{\bar{\xi}}(r)$ are its first and second volume-averaged integrals:

$$\bar{\xi}(r) = \frac{3}{r^3}\int_0^r \xi(s) s^2\, ds, \tag{15}$$

$$\bar{\bar{\xi}}(r) = \frac{5}{r^5}\int_0^r \xi(s) s^4\, ds. \tag{16}$$

Our ellipsoidal VACF estimator probes the same anisotropic effects encoded in $\xi^s(r, \mu)$, but from a complementary angle: by systematically varying the shape of the counting kernel and averaging over all directions, we recover the imprint of RSD without requiring explicit pair counts or angular binning.

Notably, the redshift-space VACF does not preserve the symmetry described in Eq. 9, as the LOS direction introduces a preferred axis. This breaks the equivalence between ellipsoids with axis ratios η and $1/\eta$, leading to asymmetric behavior in the $[r_\parallel, r_\perp]$ plane.

We also observe that, at small scales, the redshift-to-real-space ratio of $\bar{\xi}_2$ is suppressed (< 1), while at larger scales it becomes enhanced (> 1). This agrees with well-known results for the monopole component of the two-point function and its volume-averaged variants [e.g., 32, 33, 46].

Furthermore, the ellipsoidal formalism allows us to identify the shapes that yield the most extreme differences between real and redshift space. At small volumes, oblate ellipsoids ($r_\parallel < r_\perp$) provide the lowest signal in redshift space, amplifying the FoG suppression relative to real space. At large volumes, those same oblate shapes yield the strongest enhancement, aligning with the Kaiser effect. Interestingly, the maximal redshift-to-real-space ratio is not achieved at the most oblate shapes, suggesting that the mapping $\bar{\xi}_2(r_\parallel/r_\perp)$ is asymmetric and has a well-defined peak but no true minimum within the physical domain $r_\parallel > 0$.

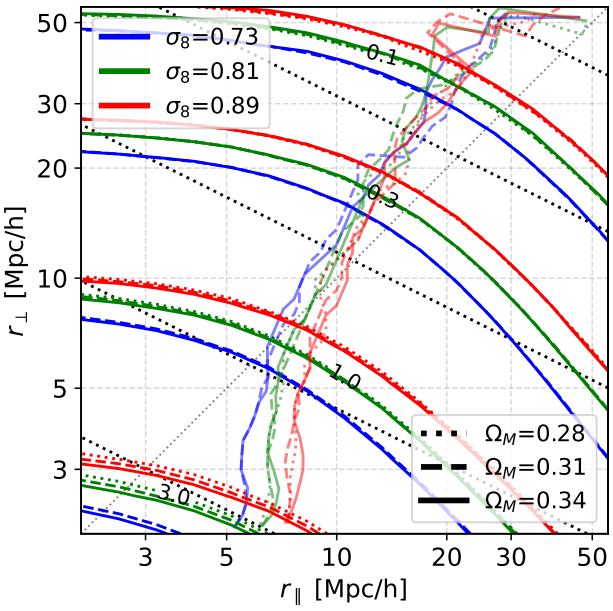

FIG. 6. Redshift-space $\bar{\xi}_2(r_\parallel, r_\perp)$ at $z = 0.1$. Colored iso-value contours correspond to different σ_8 values (blue, green, red), while line styles (dotted, dashed, solid) denote different values of Ω_M. Semi-transparent lines trace the ellipsoidal shapes that maximize $\bar{\xi}_2$ at fixed volume for each model.

These findings demonstrate that ellipsoidal VACFs encode rich anisotropic information beyond that available from spherical statistics. The additional degrees of freedom—smoothing scale and shape—enable a more nuanced analysis of large-scale structure morphology and redshift-space distortions.

B. Growth rate dependence

In the previous sections, we examined the behavior of ellipsoidal volume-averaged correlation functions and their cumulants in both real and redshift space. We now turn to investigating how the anisotropic VACFs respond to changes in the growth rate of cosmic structure, as encoded by variations in the cosmological parameters Ω_M and σ_8.

Figure 6 shows iso-correlation contours and lines of maximal $\bar{\xi}_2(r_\parallel, r_\perp)$ at constant effective volume for a set of cosmological models sampled at redshift $z = 0.1$, all computed in redshift space.

As expected, variations in σ_8 have a stronger effect on the amplitude of the correlation function than changes in Ω_M. This reflects the role of σ_8 in normalizing the matter power spectrum, to which $\bar{\xi}_2$ is directly related. In contrast, Ω_M influences the shape of the power spectrum and the growth rate f, resulting in more subtle changes to $\bar{\xi}_2$.

At scales smaller than the transition threshold $r_{\rm eff} \lesssim 10$ Mpc/h, increasing Ω_M leads to a suppression of the clustering signal. This is consistent with stronger redshift-space distortions—specifically the FoG effect—expected for models with faster growth rates. On larger scales, where the Kaiser effect dominates, the trend reverses slightly: models with larger Ω_M exhibit slightly enhanced $\bar{\xi}_2$, although the effect is weaker.

Interestingly, the location of the maximum $\bar{\xi}_2$ at fixed volume shifts with σ_8, especially on small scales. Higher σ_8 models tend to peak for more prolate ellipsoids ($r_\parallel > r_\perp$), consistent with stronger FoG elongations along the LOS. Conversely, for volumes beyond $r_{\rm eff} \sim 10$ Mpc/h, models with lower σ_8 exhibit maxima at more oblate shapes. This counterintuitive behavior arises because in high-σ_8 models, strong small-scale FoG suppression persists even when averaged over larger volumes.

At higher redshifts (not shown here), we observe a systematic shift of the maxima toward more oblate ellipsoids. This trend reflects a reduction in the FoG effect and an enhancement of the Kaiser effect at earlier cosmic times, when coherent infall motions were more prominent. This evolution is in agreement with theoretical expectations (e.g., [34]).

While the amplitude of $\bar{\xi}_2$ is clearly sensitive to σ_8, its dependence on the combined parameter $f\sigma_8 = \Omega_M^{0.55}\sigma_8$ is non-monotonic. For example, the model Cs8LOM ($f\sigma_8 = 0.4023$) yields a stronger $\bar{\xi}_2$ than Ls8HOM ($f\sigma_8 = 0.4043$), while in other cases, such as between Hs8COM and Hs8HOM, the ranking is reversed. This degeneracy makes $\bar{\xi}_2$ a poor discriminator of $f\sigma_8$.

The situation improves when we turn to higher-order statistics. In Figure 7, we show iso-contours of the reduced skewness s_3 in both real and redshift space, colored by the corresponding $f\sigma_8$ values.

In real space, all models yield nearly identical s_3 maps, making it difficult to distinguish between them. In contrast, in redshift space the contours are clearly separated and follow a consistent trend with increasing $f\sigma_8$. This enhanced sensitivity highlights the utility of anisotropic skewness as a growth-rate probe.

We also observe a significant difference in the shape of the s_3 map between real and redshift space. The redshift-space suppression of s_3 is strongest at small $r_\parallel$ and $r_\perp$, corresponding to small effective volumes. Moreover, the suppression is anisotropic: for fixed $r_{\rm eff}$, oblate ellipsoids ($r_\parallel < r_\perp$) experience greater suppression than prolate ones. This anisotropy reflects the directional nature of FoG distortions and confirms that shape-resolved cumulants offer a richer characterization of redshift-space clustering.

To explore these trends in more detail, we show cross-sections of $s_3(r_\parallel, r_\perp)$ at various kernel configurations in Figure 8. The diversity of shapes across panels underscores the richness of the signal. Differences between models reach the $1-3\sigma$ level, particularly for cross-sections that probe oblate ellipsoids. For instance, in the bottom right panel, the s_3 signal rises with increasing $r_{\rm eff}$ up to ~ 10 Mpc/h, before falling off at larger

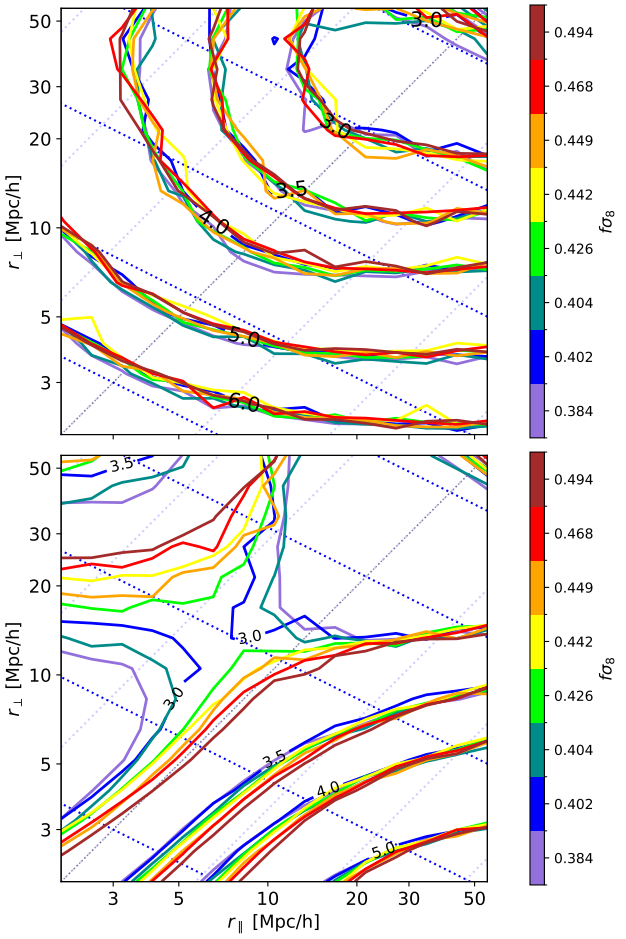

FIG. 7. Iso-contours of reduced skewness $s_3(r_\parallel, r_\perp)$ at $z = 0.1$, shown in real space (top panel) and redshift space (bottom panel). Colors correspond to different $f\sigma_8$ values. Blue dashed lines indicate iso-volume contours; the diagonal black dashed line marks spherical ellipsoids.

scales. This turnover scale is sensitive to $f\sigma_8$ and is promising for model discrimination.

We conclude that reduced skewness in redshift space, computed using anisotropic ellipsoidal kernels, is a sensitive and robust probe of cosmological growth. In contrast to isotropic VACFs, the directional information retained by ellipsoidal CiC provides access to subtle features of the Cosmic Web and redshift-space distortions, offering a promising avenue for future observational analyses.

V. CONCLUSIONS

In this work, we introduced a novel generalization of volume-averaged correlation functions (VACFs) and their associated cumulants by extending the standard counts-in-cells methodology to anisotropic, ellipsoidal kernels. These *ellipsoidal volume-averaged correlation functions*, denoted by $\bar{\xi}_J(r_\parallel, r_\perp)$, and the corresponding reduced cumulants $s_J(r_\parallel, r_\perp)$, represent a natural configuration-space analogue of the widely used anisotropic two-point function $\xi(r_p, \pi)$. Our approach distinguishes between directions parallel and perpendicular to the line of sight, allowing for explicit sensitivity to redshift-space distortions (RSD) in higher-order clustering statistics.

We explored the behavior of these estimators in both real and redshift space, using a suite of COLA-based N-body simulations with varying values of Ω_M and σ_8, and analyzed their ability to capture the imprint of cosmic structure growth and velocity-induced anisotropies. Our main findings can be summarized as follows:

- In real space, the VACF $\bar{\xi}_J(r_\parallel, r_\perp)$ at fixed ellipsoidal volume is maximized for spherical kernels ($r_\parallel = r_\perp$). In redshift space, however, this symmetry is broken by RSD. The preferred kernel shape becomes prolate ($r_\parallel > r_\perp$) on small scales dominated by Fingers-of-God (FoG) effects, and oblate ($r_\parallel < r_\perp$) on large scales where the Kaiser effect enhances LOS compression (Fig. 5).

- The shape dependence of higher-order statistics evolves systematically with order J. While lower-order VACFs show strong deviations from iso-volume contours, higher-order $\bar{\xi}_J$ functions tend to align with iso-volume shapes (Fig. 4). This suggests that at sufficiently high J, the statistics become effectively volume-dominated and less sensitive to ellipsoidal anisotropy.

- In redshift space, variations in σ_8 produce a significantly stronger effect on $\bar{\xi}_2$ than equivalent variations in Ω_M (Fig. 6). At small scales, models with higher σ_8 show stronger FoG suppression and favor more prolate kernel shapes. At larger scales, the transition to the Kaiser regime occurs at larger effective radii for these models.

- While $\bar{\xi}_J$ does not exhibit a monotonic dependence on the composite growth-rate parameter $f\sigma_8$, the redshift-space cumulants s_J, particularly the skewness s_3, do. In real space, models with different $f\sigma_8$ values are nearly indistinguishable, but in redshift space they exhibit clear, coherent differences (Fig. 7).

- Cross-sectional analyses of $s_3(r_\parallel, r_\perp)$ (Fig. 8) confirm that models with adjacent $f\sigma_8$ values can be separated at the $\sim 0.5{-}1\sigma$ level across a range of ellipsoidal configurations, especially in redshift space.

Ellipsoidal VACFs and their cumulants offer a powerful extension of the standard counts-in-cells framework,

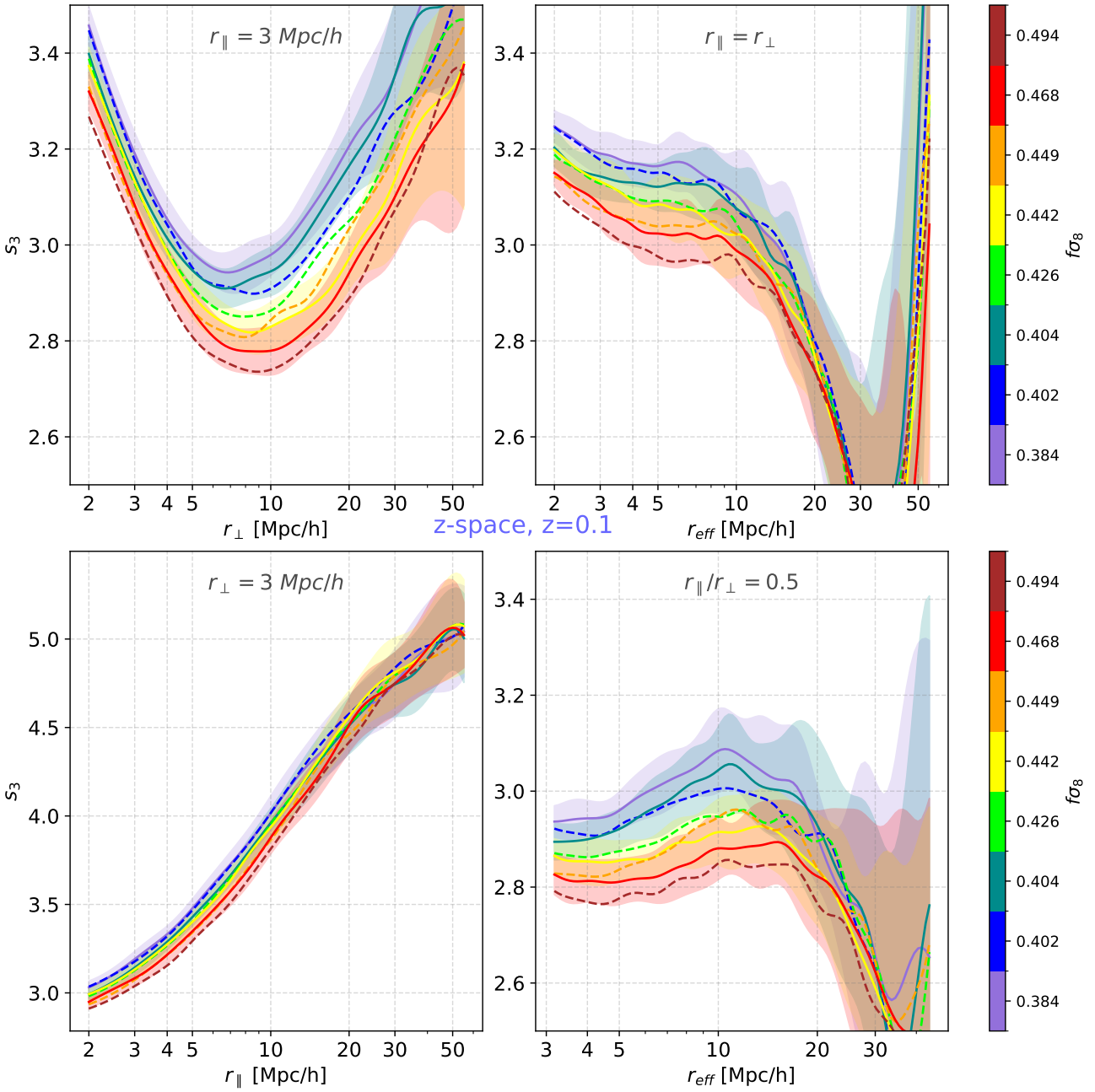

FIG. 8. Cross-sections of redshift-space skewness $s_3(r_\parallel, r_\perp)$ at $z = 0.1$. Panels show: (top left) varying $r_\perp$ at fixed $r_\parallel = 3\,\mathrm{Mpc}/h$; (top right) spherical case $r_\parallel = r_\perp$; (bottom left) varying $r_\parallel$ at fixed $r_\perp = 3\,\mathrm{Mpc}/h$; (bottom right) constant axis ratio $r_\parallel/r_\perp = 0.5$. Solid and dashed lines correspond to different $f\sigma_8$ values. For transparency every second model have errors shown (solid lines with shades), while for the rest we present only values (dashed lines).

with strong potential for extracting anisotropic clustering information from future redshift surveys. By providing access to directional and morphological dependencies in higher-order statistics, this approach opens a new window onto the physics of structure formation, redshift-space distortions, and the growth of cosmic large-scale structure. Its computational efficiency and general applicability to any cumulant order make it a promising addition to the toolbox of modern cosmological analysis.

Looking ahead, the application of ellipsoidal VACF statistics to galaxy and halo catalogs from simulations and observations represents a particularly promising avenue. Galaxy redshift surveys such as DESI, Euclid, and 4MOST will deliver high-fidelity three-dimensional

maps of the cosmic density field, ideally suited to this type of anisotropic analysis. Furthermore, since redshift-space distortions carry complementary information to real-space clustering, our method can serve as an independent probe of the growth of structure—enabling cross-validation of results from standard multipole-based RSD analyses. Finally, the sensitivity of ellipsoidal cumulants to directional anisotropies and nonlinear structure makes them a natural tool for testing extensions of the standard model, including modified gravity theories and alternative dark sector physics. In particular, their ability to access non-Gaussian signatures through higher-order cumulants while resolving geometric distortions makes them ideal for next-generation consistency tests of ΛCDM.

VI. ACKNOWLEDGMENTS

We are grateful to Enrique Gaztanga for helpful discussions at the early stages of this project. This work is supported by the Polish National Science Center through grants no. 2020/39/B/ST9/03494 and 2020/38/E/ST9/00395.

[1] R. Scoccimarro, ApJL **544**, 597 (2000), arXiv:astro-ph/0004086 [astro-ph].
[2] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002), arXiv:astro-ph/0112551 [astro-ph].
[3] R. Skibba, R. K. Sheth, A. J. Connolly, and R. Scranton, MNRAS **369**, 68 (2006), arXiv:astro-ph/0512463 [astro-ph].
[4] P. J. E. Peebles, *The large-scale structure of the universe* (1980).
[5] R. Juszkiewicz, F. R. Bouchet, and S. Colombi, ApJ **412**, L9 (1993), arXiv:astro-ph/9306003 [astro-ph].
[6] F. Bernardeau, ApJL **433**, 1 (1994), arXiv:astro-ph/9312026 [astro-ph].
[7] S. Colombi, F. R. Bouchet, and L. Hernquist, ApJL **465**, 14 (1996), arXiv:astro-ph/9508142 [astro-ph].
[8] E. Gaztanaga, MNRAS **268**, 913 (1994), arXiv:astro-ph/9309019 [astro-ph].
[9] W. A. Hellwing, B. Li, C. S. Frenk, and S. Cole, MNRAS **435**, 2806 (2013), arXiv:1305.7486 [astro-ph.CO].
[10] W. A. Hellwing, K. Koyama, B. Bose, and G.-B. Zhao, Phys. Rev. D **96**, 023515 (2017), arXiv:1703.03395 [astro-ph.CO].
[11] S. Alam, C. Arnold, A. Aviles, R. Bean, Y.-C. Cai, M. Cautun, J. L. Cervantes-Cota, C. Cuesta-Lazaro, N. C. Devi, A. Eggemeier, S. Fromenteau, A. X. Gonzalez-Morales, V. Halenka, J.-h. He, W. A. Hellwing, C. Hernández-Aguayo, M. Ishak, K. Koyama, B. Li, A. de la Macorra, J. Meneses Rizo, C. Miller, E.-M. Mueller, G. Niz, P. Ntelis, M. Rodríguez Otero, C. G. Sabiu, Z. Slepian, A. Stark, O. Valenzuela, G. Valogiannis, M. Vargas-Magaña, H. A. Winther, P. Zarrouk, G.-B. Zhao, and Y. Zheng, J. Cosmology Astropart. Phys. **2021**, 050 (2021), arXiv:2011.05771 [astro-ph.CO].
[12] M. Cataneo, C. Uhlemann, C. Arnold, A. Gough, B. Li, and C. Heymans, MNRAS **513**, 1623 (2022), arXiv:2109.02636 [astro-ph.CO].
[13] E. Hivon, F. R. Bouchet, S. Colombi, and R. Juszkiewicz, A&A **298**, 643 (1995), arXiv:astro-ph/9407049 [astro-ph].
[14] S. Codis, F. Bernardeau, and C. Pichon, MNRAS **460**, 1598 (2016), arXiv:1602.03562 [astro-ph.CO].
[15] A. Repp and I. Szapudi, MNRAS **498**, L125 (2020), arXiv:2006.01146 [astro-ph.CO].
[16] P. Drozda, W. A. Hellwing, and M. Bilicki, arXiv e-prints , arXiv:2504.12820 (2025), arXiv:2504.12820 [astro-ph.CO].
[17] B. McCarthy Gould, L. Castiblanco, C. Uhlemann, and O. Friedrich, The Open Journal of Astrophysics **8**, 1 (2025), arXiv:2409.18182 [astro-ph.CO].
[18] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo, Phys. Rep. **530**, 87 (2013), arXiv:1201.2434 [astro-ph.CO].
[19] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen, C. Allende Prieto, J. Annis, S. Bailey, C. Balland, O. Ballester, C. Baltay, L. Beaufore, C. Bebek, T. C. Beers, E. F. Bell, J. L. Bernal, R. Besuner, F. Beutler, C. Blake, H. Bleuler, M. Blomqvist, R. Blum, A. S. Bolton, C. Briceno, D. Brooks, J. R. Brownstein, E. Buckley-Geer, A. Burden, E. Burtin, N. G. Busca, R. N. Cahn, Y.-C. Cai, L. Cardiel-Sas, R. G. Carlberg, P.-H. Carton, R. Casas, F. J. Castander, J. L. Cervantes-Cota, T. M. Claybaugh, M. Close, C. T. Coker, S. Cole, J. Comparat, A. P. Cooper, M. C. Cousinou, M. Crocce, J.-G. Cuby, D. P. Cunningham, T. M. Davis, K. S. Dawson, A. de la Macorra, J. De Vicente, T. Delubac, M. Derwent, A. Dey, G. Dhungana, Z. Ding, P. Doel, Y. T. Duan, A. Ealet, J. Edelstein, S. Eftekharzadeh, D. J. Eisenstein, A. Elliott, S. Escoffier, M. Evatt, P. Fagrelius, X. Fan, K. Fanning, A. Farahi, J. Farihi, G. Favole, Y. Feng, E. Fernandez, J. R. Findlay, D. P. Finkbeiner, M. J. Fitzpatrick, B. Flaugher, S. Flender, A. Font-Ribera, J. E. Forero-Romero, P. Fosalba, C. S. Frenk, M. Fumagalli, B. T. Gaensicke, G. Gallo, J. Garcia-Bellido, E. Gaztanaga, N. Pietro Gentile Fusillo, T. Gerard, I. Gershkovich, T. Giannantonio, D. Gillet, G. Gonzalez-de-Rivera, V. Gonzalez-Perez, S. Gott, O. Graur, G. Gutierrez, J. Guy, S. Habib, H. Heetderks, I. Heetderks, K. Heitmann, W. A. Hellwing, D. A. Herrera, S. Ho, S. Holland, K. Honscheid, E. Huff, T. A. Hutchinson, D. Huterer, H. S. Hwang, J. M. Illa Laguna, Y. Ishikawa, D. Jacobs, N. Jeffrey, P. Jelinsky, E. Jennings, L. Jiang, J. Jimenez, J. Johnson, R. Joyce, E. Jullo, S. Juneau, S. Kama, A. Karcher, S. Karkar, R. Kehoe, N. Kennamer, S. Kent, M. Kilbinger, A. G. Kim, D. Kirkby, T. Kisner, E. Kitanidis, J.-P. Kneib, S. Koposov, E. Ko-

vacs, K. Koyama, A. Kremin, R. Kron, L. Kronig, A. Kueter-Young, C. G. Lacey, R. Lafever, O. Lahav, A. Lambert, M. Lampton, M. Landriau, D. Lang, T. R. Lauer, J.-M. Le Goff, L. Le Guillou, A. Le Van Suu, J. H. Lee, S.-J. Lee, D. Leitner, M. Lesser, M. E. Levi, B. L'Huillier, B. Li, M. Liang, H. Lin, E. Linder, S. R. Loebman, Z. Lukić, J. Ma, N. MacCrann, C. Magneville, L. Makarem, M. Manera, C. J. Manser, R. Marshall, P. Martini, R. Massey, T. Matheson, J. McCauley, P. McDonald, I. D. McGreer, A. Meisner, N. Metcalfe, T. N. Miller, R. Miquel, J. Moustakas, A. Myers, M. Naik, J. A. Newman, R. C. Nichol, A. Nicola, L. Nicolati da Costa, J. Nie, G. Niz, P. Norberg, B. Nord, D. Norman, P. Nugent, T. O'Brien, M. Oh, and K. A. G. Olsen, arXiv e-prints , arXiv:1611.00036 (2016), arXiv:1611.00036 [astro-ph.IM].

[20] E. N. Taylor, M. Cluver, E. Bell, J. Brinchmann, M. Colless, H. Courtois, H. Hoekstra, S. Kannappan, C. Lagos, J. Liske, E. Tempel, C. Howlett, S. McGee, K. Said, R. Skelton, M. Gunawardhana, S. Bellstedt, L. Hunt, T. Jarrett, C. Lidman, J. Lucey, S. Alam, M. Bilicki, A. de Graaff, W. Hellwing, S. Leslie, I. Loubser, L. Marchetti, M. Maseda, M. Mogotsi, P. Norberg, A. Sonnenfeld, J. G. Sorce, and 4HS Team, The Messenger **190**, 46 (2023).

[21] Euclid Collaboration, Y. Mellier, Abdurro'uf, J. A. Acevedo Barroso, A. Achúcarro, J. Adamek, R. Adam, G. E. Addison, N. Aghanim, M. Aguena, V. Ajani, Y. Akrami, A. Al-Bahlawan, A. Alavi, I. S. Albuquerque, G. Alestas, G. Alguero, A. Allaoui, S. W. Allen, V. Allevato, A. V. Alonso-Tetilla, B. Altieri, A. Alvarez-Candal, S. Alvi, A. Amara, L. Amendola, J. Amiaux, I. T. Andika, S. Andreon, A. Andrews, G. Angora, R. E. Angulo, F. Annibali, A. Anselmi, S. Anselmi, S. Arcari, M. Archidiacono, G. Aricò, M. Arnaud, S. Arnouts, M. Asgari, J. Asorey, L. Atayde, H. Atek, F. Atrio-Barandela, M. Aubert, E. Aubourg, T. Auphan, N. Auricchio, B. Aussel, H. Aussel, P. P. Avelino, A. Avgoustidis, S. Avila, S. Awan, R. Azzollini, C. Baccigalupi, E. Bachelet, D. Bacon, M. Baes, M. B. Bagley, B. Bahr-Kalus, A. Balaguera-Antolinez, E. Balbinot, M. Balcells, M. Baldi, I. Baldry, A. Balestra, M. Ballardini, O. Ballester, M. Balogh, E. Bañados, R. Barbier, S. Bardelli, M. Baron, T. Barreiro, R. Barrena, J. C. Barriere, B. J. Barros, A. Barthelemy, N. Bartolo, A. Basset, P. Battaglia, A. J. Battisti, C. M. Baugh, L. Baumont, L. Bazzanini, J. P. Beaulieu, V. Beckmann, A. N. Belikov, J. Bel, F. Bellagamba, M. Bella, E. Bellini, K. Benabed, R. Bender, G. Benevento, C. L. Bennett, K. Benson, P. Bergamini, J. R. Bermejo-Climent, F. Bernardeau, D. Bertacca, M. Berthe, J. Berthier, M. Bethermin, F. Beutler, C. Bevillon, S. Bhargava, R. Bhatawdekar, D. Bianchi, L. Bisigello, A. Biviano, R. P. Blake, A. Blanchard, J. Blazek, L. Blot, A. Bosco, C. Bodendorf, T. Boenke, H. Böhringer, P. Boldrini, M. Bolzonella, A. Bonchi, M. Bonici, D. Bonino, L. Bonino, C. Bonvin, W. Bon, J. T. Booth, S. Borgani, A. S. Borlaff, E. Borsato, B. Bose, M. T. Botticella, A. Boucaud, F. Bouche, J. S. Boucher, D. Boutigny, T. Bouvard, R. Bouwens, H. Bouy, R. A. A. Bowler, V. Bozza, E. Bozzo, E. Branchini, G. Brando, S. Brau-Nogue, P. Brekke, M. N. Bremer, M. Brescia, M. A. Breton, J. Brinchmann, T. Brinckmann, C. Brockley-Blatt, M. Brodwin, L. Brouard, M. L. Brown, S. Bruton, J. Bucko, H. Buddelmeijer, G. Buenadicha, F. Buitrago, P. Burger, C. Burigana, V. Busillo, D. Busonero, R. Cabanac, L. Cabayol-Garcia, M. S. Cagliari, A. Caillat, L. Caillat, M. Calabrese, A. Calabro, G. Calderone, F. Calura, B. Camacho Quevedo, S. Camera, L. Campos, G. Cañas-Herrera, G. P. Candini, M. Cantiello, V. Capobianco, E. Cappellaro, N. Cappelluti, A. Cappi, K. I. Caputi, C. Cara, C. Carbone, V. F. Cardone, E. Carella, R. G. Carlberg, M. Carle, L. Carminati, F. Caro, J. M. Carrasco, J. Carretero, P. Carrilho, J. Carron Duque, and B. Carry, A&A **697**, A1 (2025), arXiv:2405.13491 [astro-ph.CO].

[22] E. Di Valentino, J. Levi Said, A. Riess, A. Pollo, V. Poulin, A. Gómez-Valent, A. Weltman, A. Palmese, C. D. Huang, C. van de Bruck, C. Shekhar Saraf, C.-Y. Kuo, C. Uhlemann, D. Grandón, D. Paz, D. Eckert, E. M. Teixeira, E. N. Saridakis, E. Ó. Colgáin, F. Beutler, F. Niedermann, F. Bajardi, G. Barenboim, G. Gubitosi, I. Musella, I. Banik, I. Szapudi, J. Singal, J. Haro Cases, J. Chluba, J. Torrado, J. Mifsud, K. Jedamzik, K. Said, K. Dialektopoulos, L. Herold, L. Perivolaropoulos, L. Zu, L. Galbany, L. Breuval, L. Visinelli, L. A. Escamilla, L. A. Anchordoqui, M. M. Sheikh-Jabbari, M. Lembo, M. G. Dainotti, M. Vincenzi, M. Asgari, M. Gerbino, M. Forconi, M. Cantiello, M. Moresco, M. Benetti, N. Schöneberg, Ö. Akarsu, R. C. Nunes, R. C. Bernardo, R. Chávez, R. I. Anderson, R. Watkins, S. Capozziello, S. Li, S. Vagnozzi, S. Pan, T. Treu, V. Irsic, W. Handley, W. Giarè, Y. Murakami, A. Poudou, A. Heavens, A. Kogut, A. Domi, A. Łukasz Lenart, A. Melchiorri, A. Vadalà, A. Amon, A. Bonilla, A. Reeves, A. Zhuk, A. Bonanno, A. Övgün, A. Pisani, A. Talebian, A. Abebe, A. Aboubrahim, A. L. González Morán, A. Kovács, A. Papatriantafyllou, A. R. Liddle, A. Paliathanasis, A. Borowiec, A. K. Yadav, A. Yadav, A. A. Sen, A. J. W. Mini Latha, A. C. Davis, A. J. Shajib, A. Walters, A. Idicherian Lonappan, A. Chudaykin, A. Capodagli, A. da Silva, A. De Felice, A. Racioppi, A. Soler Oficial, A. Montiel, A. Favale, A. Bernui, A. C. Velasco, A. Heinesen, A. Bakopoulos, A. Chatzistavrakidis, B. Khanpour, B. S. Sathyaprakash, B. Zgirski, B. L'Huillier, B. Famaey, B. Jain, B. Marek, B. Zhang, B. Karmakar, B. Dragovich, B. Thomas, C. Correa, C. G. Boiza, C. Marques, C. Escamilla-Rivera, C. Tzerefos, C. Zhang, C. De Leo, C. Pfeifer, C. Lee, C. Venter, C. Gomes, C. Roque De bom, C. Moreno-Pulido, D. Iosifidis, D. Grin, D. Blixt, D. Scolnic, D. Oriti, D. Dobrycheva, D. Bettoni, D. Benisty, D. Fernández-Arenas, D. L. Wiltshire, D. Sanchez Cid, D. Tamayo, D. Valls-Gabaud, D. Pedrotti, D. Wang, D. Staicova, D. Totolou, D. Rubiera-Garcia, D. Milaković, D. Pesce, D. Sluse, D. Borka, E. Yusofi, E. Giusarma, E. Terlevich, E. Tomasetti, E. C. Vagenas, E. Fazzari, E. G. M. Ferreira, E. Barakovic, E. Dimastrogiovanni, E. Brinch Holm, E. Mottola, E. Özülker, E. Specogna, E. Brocato, E. Jensko, E. Antonette Enriquez, E. Bhatia, F. Bresolin, F. Avila, F. Bouchè, F. Bombacigno, F. K. Anagnostopoulos, F. Pace, F. Sorrenti, F. S. N. Lobo, F. Courbin, F. K. Hansen, G. Sloan, G. Farrugia, G. Lynch, G. Garcia-Arroyo, G. Raimondo, G. Lambiase, G. S. Anand, G. Poulot, G. Leon, G. Kou-

niatalis, G. Nardini, G. Csörnyei, G. Galloni, and G. Bargiacchi, arXiv e-prints , arXiv:2504.01669 (2025), arXiv:2504.01669 [astro-ph.CO].
[23] C. Howlett, M. Manera, and W. J. Percival, Astronomy and Computing **12**, 109 (2015), arXiv:1506.03737 [astro-ph.CO].
[24] R. Juszkiewicz, W. A. Hellwing, and R. van de Weygaert, MNRAS **429**, 1206 (2013), arXiv:1205.6163 [astro-ph.CO].
[25] F. R. Bouchet, R. Juszkiewicz, S. Colombi, and R. Pellat, ApJ **394**, L5 (1992).
[26] F. Bernardeau, ApJL **433**, 1 (1994), arXiv:astro-ph/9312026 [astro-ph].
[27] E. Gaztanaga and F. Bernardeau, A&A **331**, 829 (1998), arXiv:astro-ph/9707095 [astro-ph].
[28] S. Tassev, M. Zaldarriaga, and D. J. Eisenstein, J. Cosmology Astropart. Phys. **2013**, 036 (2013), arXiv:1301.0322 [astro-ph.CO].
[29] H. A. Winther, K. Koyama, M. Manera, B. S. Wright, and G.-B. Zhao, J. Cosmology Astropart. Phys. **2017**, 006 (2017), arXiv:1703.00879 [astro-ph.CO].
[30] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, *et al.*, A&A **641**, A6 (2020), arXiv:1807.06209 [astro-ph.CO].
[31] N. Kaiser, MNRAS **227**, 1 (1987).
[32] R. Scoccimarro, Phys. Rev. D **70**, 083007 (2004), arXiv:astro-ph/0407214 [astro-ph].
[33] A. Taruya, T. Nishimichi, and S. Saito, Phys. Rev. D **82**, 063522 (2010), arXiv:1006.0699 [astro-ph.CO].
[34] L. Guzzo, M. Pierleoni, B. Meneux, E. Branchini, O. Le Fèvre, C. Marinoni, B. Garilli, J. Blaizot, G. De Lucia, A. Pollo, H. J. McCracken, D. Bottini, V. Le Brun, D. Maccagni, J. P. Picat, R. Scaramella, M. Scodeggio, L. Tresse, G. Vettolani, A. Zanichelli, C. Adami, S. Arnouts, S. Bardelli, M. Bolzonella, A. Bongiorno, A. Cappi, S. Charlot, P. Ciliegi, T. Contini, O. Cucciati, S. de la Torre, K. Dolag, S. Foucaud, P. Franzetti, I. Gavignaud, O. Ilbert, A. Iovino, F. Lamareille, B. Marano, A. Mazure, P. Memeo, R. Merighi, L. Moscardini, S. Paltani, R. Pellò, E. Perez-Montero, L. Pozzetti, M. Radovich, D. Vergani, G. Zamorani, and E. Zucca, Nature **451**, 541 (2008), arXiv:0802.1944 [astro-ph].
[35] C. Blake, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, D. Gilbank, M. Gladders, K. Glazebrook, B. Jelliffe, R. J. Jurek, I. H. Li, B. Madore, D. C. Martin, K. Pimbblet, G. B. Poole, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, T. K. Wyder, and H. K. C. Yee, MNRAS **415**, 2876 (2011), arXiv:1104.2948 [astro-ph.CO].
[36] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, G. B. Poole, L. Campbell, Q. Parker, W. Saunders, and F. Watson, MNRAS **423**, 3430 (2012), arXiv:1204.4725 [astro-ph.CO].
[37] A. J. S. Hamilton, in *The Evolving Universe*, Astrophysics and Space Science Library, Vol. 231, edited by D. Hamilton (1998) p. 185, arXiv:astro-ph/9708102 [astro-ph].
[38] E. Jennings, *Simulations of dark energy cosmologies*, Ph.D. thesis, Durham University, UK (2011).
[39] W. A. Hellwing, in *XXXIX Polish Astronomical Society Meeting*, Vol. 10, edited by K. Małek, M. Polińska, A. Majczyna, G. Stachowski, R. Poleski, Ł. Wyrzykowski, and A. Różańska (2020) pp. 315–322, arXiv:1912.13026 [astro-ph.CO].
[40] P. Drozda, W. A. Hellwing, and M. Bilicki, arXiv e-prints , arXiv:2504.12820 (2025), arXiv:2504.12820 [astro-ph.CO].
[41] J. N. Fry and E. Gaztanaga, ApJL **413**, 447 (1993), arXiv:astro-ph/9302009 [astro-ph].
[42] J. C. Jackson, MNRAS **156**, 1P (1972), arXiv:0810.3908 [astro-ph].
[43] S. Dodelson, *Modern Cosmology* (2003).
[44] W. J. Percival and M. White, MNRAS **393**, 297 (2009), arXiv:0808.0003 [astro-ph].
[45] F. Shi, X. Yang, H. Wang, Y. Zhang, H. J. Mo, F. C. van den Bosch, S. Li, C. Liu, Y. Lu, D. Tweed, and L. Yang, ApJL **833**, 241 (2016), arXiv:1608.02313 [astro-ph.CO].
[46] J. A. Peacock and S. J. Dodds, MNRAS **267**, 1020 (1994), arXiv:astro-ph/9311057 [astro-ph].

5.2 Summary

While the developed anisotropic averaged correlation functions $\bar{\xi}_J(r_\parallel, r_\perp)$ along with constructed cumulants $s_J(r_\parallel, r_\perp)$ are less straightforward for interpretation than spherical functions, these consist an useful tool in cosmic web analysis. The appearance of anisotropic functions in $[r_\parallel, r_\perp]$ representation can be qualitatively explained by shapes of structures within cosmic web. An important feature raising especially in redshift space is the line of maximized $\bar{\xi}_J$ at given ellipsoid volume, which points out which shapes of structures are preferred to provide strongest correlation signal. This information can not be obtained with standard spherical approach and may consist an useful footprint for distinguishing cosmology models.

The anisotropic functions are well suited for comparing cosmic web with different structure growth. Especially anisotropic cumulants offer intriguing signals in $[r_\parallel, r_\perp]$ representation. Since the used method is new, there might be much more properties of anisotropic counts to take advantage of.

Chapter 6

Conclusions

In this thesis we have studied central moments-based averaged correlation functions and cumulants of simulated tracers of the cosmic web. The work is organized into two parts. In first part, we assessed the potential for testing gravity on cosmological scales by comparing extended gravity scenarios (Hu–Sawicki $f(R)$ and nDGP) to General Relativity (GR). In Chapter 3, we examined angular clustering in a light-cone geometry. By optimizing over the amount of information from catalogs, we identified the range in which deviations from GR (the EG signals) are strongest. We detected significant signals for dark matter particles, halos, and HOD galaxies, which motivated us to extend our analysis to fully three-dimensional clustering.

Because projection and scale mixing affect angular statistics, the redshifts corresponding to strongest EG signals in 3D differs from that found in the light-cone study. In Chapter 4, we therefore measured 3D clustering across all available simulation snapshots, in both real and redshift space. Rather than presenting every order of the correlation hierarchy, we focused on skewness, which most clearly highlights the differences between real and redshift space. We found that EG signals are weaker in redshift space - an important caveat, since observations are inherently redshift-space measurements. However,galaxy samples still exhibit stronger deviations than halos, offering a promising route for future tests. In both the angular and 3D analyses, we also observed that higher-order statistics produce larger EG deviations but suffer from rapidly increasing uncertainties, suggesting that very high orders may become useful only with larger, more complete catalogs.

The second part (Chapter 5) explores how to extract additional information from existing catalogs by replacing spherical counts with ellipsoidal ones. The used higher-order moments-based functions are averaged versions of J-point correlation functions. Despite of being much more achievable computationally, these contain less information than their multi-dimensional cousins. To account for that, we developed anisotropic, averaged correlation functions: central moment-based analogues of the standard anisotropic two-point function—computed along and across the line of sight, in both real and redshift space. Applying these new statistics to models with different structure-growth rates, we demonstrated that they reveal features inaccessible to spherical counts. This approach not only

remains far less computationally demanding than full J-point correlation functions but also captures direction-dependent clustering signatures, opening a novel window on large-scale structure.
Overall, this thesis identifies clear EG signals and illustrates the power of higher-order, averaged correlation functions to discriminate between cosmic webs evolved under different gravity scenarios. We quantify how detectability differs in real versus redshift space and introduce ellipsoidal averaged measures to harness the additional anisotropic information present in redshift-space data.

6.1 Future prospects

New and forthcoming photometric surveys such as Euclid [182] or Legacy Survey of Space and Time (LSST) [183] - will observe tens of billions of galaxies, vastly increasing the volume of available data. At the same time, new spectroscopic facilities like the Wide-field Spectroscopic Telescope (WST) [184] will furnish precise redshifts for hundreds of millions of galaxies. Together, these surveys will tighten constraints on cosmological parameters and deepen our understanding of dark energy and dark matter.

Such extensive datasets are especially valuable for clustering analyses. They will enable measurements of the higher-order averaged correlation functions developed in this work with unprecedented precision. Along with structure growth rate estimations, these statistics will allow a sensitive search for signatures of extended gravity. Moreover, by employing the anisotropic, ellipsoid-averaged functions introduced here, one can probe the cosmic web from a novel vantage point. Due to the richness of forthcoming observational data, one might identify features not detected so far.
In summary, the future will present both significant challenges and unique opportunities to refine our picture of the components of Universe, its evolution and to put tighter constraints on the nature of gravity itself.

Bibliography

[1] A. Einstein. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354(7):769–822, January 1916. doi: 10.1002/andp.19163540702.

[2] DESI Collaboration, M. Abdul-Karim, J. Aguilar, S. Ahlen, S. Alam, L. Allen, C. Allende Prieto, O. Alves, A. Anand, U. Andrade, E. Armengaud, A. Aviles, S. Bailey, C. Baltay, P. Bansal, A. Bault, J. Behera, S. BenZvi, D. Bianchi, C. Blake, S. Brieden, A. Brodzeller, D. Brooks, et al. DESI DR2 Results II: Measurements of Baryon Acoustic Oscillations and Cosmological Constraints. *arXiv e-prints*, art. arXiv:2503.14738, March 2025. doi: 10.48550/arXiv.2503.14738.

[3] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, et al. Planck 2018 results. VI. Cosmological parameters. A&A , 641: A6, September 2020. doi: 10.1051/0004-6361/201833910.

[4] Clifford M. Will. The Confrontation between General Relativity and Experiment. *Living Reviews in Relativity*, 17(1):4, December 2014. doi: 10.12942/lrr-2014-4.

[5] D. J. Kapner, T. S. Cook, E. G. Adelberger, J. H. Gundlach, B. R. Heckel, C. D. Hoyle, and H. E. Swanson. Tests of the Gravitational Inverse-Square Law below the Dark-Energy Length Scale. Phys. Rev. Lett. , 98(2):021101, January 2007. doi: 10.1103/PhysRevLett.98.021101.

[6] Andrew A. Geraci, Sylvia J. Smullin, David M. Weld, John Chiaverini, and Aharon Kapitulnik. Improved constraints on non-Newtonian forces at 10microns. Phys. Rev. D , 78(2):022002, July 2008. doi: 10.1103/PhysRevD.78.022002.

[7] Shan-Qing Yang, Bi-Fu Zhan, Qing-Lan Wang, Cheng-Gang Shao, Liang-Cheng Tu, Wen-Hai Tan, and Jun Luo. Test of the Gravitational Inverse Square Law at Millimeter Ranges. Phys. Rev. Lett. , 108(8):081101, February 2012. doi: 10.1103/PhysRevLett.108.081101.

[8] J. H. Taylor and J. M. Weisberg. A new test of general relativity - Gravitational radiation and the binary pulsar PSR 1913+16. ApJL , 253:908–920, February 1982. doi: 10.1086/159690.

[9] M. Kramer, I. H. Stairs, R. N. Manchester, M. A. McLaughlin, A. G. Lyne, R. D. Ferdman, M. Burgay, D. R. Lorimer, A. Possenti, N. D'Amico, J. M. Sarkissian, G. B. Hobbs, J. E. Reynolds, P. C. C. Freire, and F. Camilo. Tests of General Relativity from Timing the Double Pulsar. *Science*, 314(5796):97–102, October 2006. doi: 10.1126/science.1132305.

[10] Bryan A. Jacoby, P. B. Cameron, F. A. Jenet, S. B. Anderson, R. N. Murty, and S. R. Kulkarni. Measurement of Orbital Decay in the Double Neutron Star Binary PSR B2127+11C. In *American Astronomical Society Meeting Abstracts*, volume 209 of *American Astronomical Society Meeting Abstracts*, page 91.01, December 2006.

[11] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, K. Agatsuma, N. Aggarwal, O. D. Aguiar, L. Aiello, A. Ain, P. Ajith, B. Allen, A. Allocca, P. A. Altin, S. B. Anderson, W. G. Anderson, K. Arai, M. C. Araya, C. C. Arceneaux, J. S. Areeda, N. Arnaud, K. G. Arun, S. Ascenzi, G. Ashton, M. Ast, S. M. Aston, P. Astone, et al. Tests of General Relativity with GW150914. Phys. Rev. Lett. , 116(22):221101, May 2016. doi: 10.1103/PhysRevLett.116.221101.

[12] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt, B. Agarwal, M. Agathos, K. Agatsuma, N. Aggarwal, O. D. Aguiar, L. Aiello, A. Ain, P. Ajith, B. Allen, G. Allen, A. Allocca, M. A. Aloy, P. A. Altin, A. Amato, A. Ananyeva, S. B. Anderson, W. G. Anderson, S. V. Angelova, S. Antier, S. Appert, K. Arai, M. C. Araya, J. S. Areeda, M. Arène, N. Arnaud, et al. Tests of General Relativity with GW170817. Phys. Rev. Lett. , 123(1):011102, July 2019. doi: 10.1103/PhysRevLett.123.011102.

[13] B. Bertotti, L. Iess, and P. Tortora. A test of general relativity using radio links with the Cassini spacecraft. Nature , 425(6956):374–376, September 2003. doi: 10.1038/nature01997.

[14] A. K. Verma, A. Fienga, J. Laskar, H. Manche, and M. Gastineau. Use of MESSENGER radioscience data to improve planetary ephemeris and to test general relativity. A&A , 561:A115, January 2014. doi: 10.1051/0004-6361/201322124.

[15] Tonghua Liu, Marek Biesiada, Shuxun Tian, and Kai Liao. Robust test of general relativity at the galactic scales by combining strong lensing systems and gravi-

tational wave standard sirens. Phys. Rev. D , 109(8):084074, April 2024. doi: 10.1103/PhysRevD.109.084074.

[16] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of Ω and Λ from 42 High-Redshift Supernovae. ApJL , 517(2):565–586, June 1999. doi: 10.1086/307221.

[17] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE Differential Microwave Radiometer First-Year Maps. ApJ , 396:L1, September 1992. doi: 10.1086/186504.

[18] Aleksandr Aleksandrovich Friedmann. Über die kruümmung des raumes. *Uspekhi Fizicheskikh Nauk*, 93(2):280–287, 1967.

[19] M. Koussour, N. Myrzakulov, and M. K. M. Ali. Exploring Universe acceleration through observational constraints via Hubble parameter reconstruction. *Journal of High Energy Astrophysics*, 42:96–103, June 2024. doi: 10.1016/j.jheap.2024.04.003.

[20] Andrew Liddle. *An Introduction to Modern Cosmology, Second Edition*. 2003.

[21] Tom W. B. Kibble. The Standard Model of Particle Physics. *arXiv e-prints*, art. arXiv:1412.4094, December 2014. doi: 10.48550/arXiv.1412.4094.

[22] Kevork Abazajian, George M. Fuller, and Wallace H. Tucker. Direct Detection of Warm Dark Matter in the X-Ray. ApJL , 562(2):593–604, December 2001. doi: 10.1086/323867.

[23] S. D. M. White, M. Davis, and C. S. Frenk. The size of clusters in a neutrino-dominated universe. MNRAS , 209:27P–31P, July 1984. doi: 10.1093/mnras/209.1.27P.

[24] P. Bode, J. P. Ostriker, and N. Turok. Halo Formation in Warm Dark Matter Models. In *American Astronomical Society Meeting Abstracts*, volume 197 of *American Astronomical Society Meeting Abstracts*, page 72.04, December 2000.

[25] Simon D. M. White, Carlos S. Frenk, Marc Davis, and George Efstathiou. Clusters, Filaments, and Voids in a Universe Dominated by Cold Dark Matter. ApJL , 313: 505, February 1987. doi: 10.1086/164990.

[26] A. Schneider, D. Anderhalden, A. V. Maccio, and J. Diemand. Warm dark matter does not do better than cold dark matter in solving small-scale inconsistencies. MNRAS , 441:L6–L10, June 2014. doi: 10.1093/mnrasl/slu034.

[27] Phillip James E. Peebles. Status of the ΛCDM theory: supporting evidence and anomalies. *Philosophical Transactions of the Royal Society of London Series A*, 383 (2290):20240021, February 2025. doi: 10.1098/rsta.2024.0021.

[28] Daniel J. Eisenstein, Idit Zehavi, David W. Hogg, Roman Scoccimarro, Michael R. Blanton, Robert C. Nichol, Ryan Scranton, Hee-Jong Seo, Max Tegmark, Zheng Zheng, Scott F. Anderson, Jim Annis, Neta Bahcall, Jon Brinkmann, Scott Burles, Francisco J. Castander, Andrew Connolly, Istvan Csabai, Mamoru Doi, Masataka Fukugita, Joshua A. Frieman, Karl Glazebrook, James E. Gunn, John S. Hendry, Gregory Hennessy, Zeljko Ivezić, Stephen Kent, Gillian R. Knapp, Huan Lin, Yeong-Shang Loh, Robert H. Lupton, Bruce Margon, Timothy A. McKay, Avery Meiksin, Jeffery A. Munn, Adrian Pope, Michael W. Richmond, David Schlegel, Donald P. Schneider, Kazuhiro Shimasaku, Christopher Stoughton, Michael A. Strauss, Mark SubbaRao, Alexander S. Szalay, István Szapudi, Douglas L. Tucker, Brian Yanny, and Donald G. York. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. ApJL , 633(2):560–574, November 2005. doi: 10.1086/466512.

[29] T. M. C. Abbott, M. Aguena, A. Alarcon, O. Alves, A. Amon, F. Andrade-Oliveira, J. Annis, S. Avila, D. Bacon, E. Baxter, K. Bechtol, M. R. Becker, G. M. Bernstein, S. Birrer, J. Blazek, S. Bocquet, A. Brandao-Souza, S. L. Bridle, D. Brooks, D. L. Burke, H. Camacho, A. Campos, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, F. J. Castander, R. Cawthon, C. Chang, A. Chen, R. Chen, A. Choi, C. Conselice, J. Cordero, M. Costanzi, M. Crocce, L. N. da Costa, M. E. S. Pereira, C. Davis, T. M. Davis, J. DeRose, S. Desai, Di Valentino, et al. Dark Energy Survey Year 3 results: Constraints on extensions to Λ CDM with weak lensing and galaxy clustering. Phys. Rev. D , 107(8):083504, April 2023. doi: 10.1103/PhysRevD.107.083504.

[30] Lars Hernquist, Neal Katz, David H. Weinberg, and Jordi Miralda-Escudé. The Lyman-Alpha Forest in the Cold Dark Matter Model. ApJ , 457:L51, February 1996. doi: 10.1086/309899.

[31] Elise Darragh-Ford, Adam B. Mantz, Elena Rasia, Steven W. Allen, R. Glenn Morris, Jack Foster, Robert W. Schmidt, and Guillermo Wenrich. The Concentration-Mass relation of massive, dynamically relaxed galaxy clusters: agreement between

observations and ΛCDM simulations. MNRAS , 521(1):790–799, May 2023. doi: 10.1093/mnras/stad585.

[32] Guanlin Liu, Yu Wang, and Wen Zhao. Testing the consistency of early and late cosmological parameters with BAO and CMB data. *Physics Letters B*, 854:138717, July 2024. doi: 10.1016/j.physletb.2024.138717.

[33] Anna Chiara Alfano, Orlando Luongo, and Marco Muccino. Dark energy constraints using gamma-ray burst correlations with DESI 2024 data. *Journal of High Energy Astrophysics*, 46:100348, April 2025. doi: 10.1016/j.jheap.2025.100348.

[34] Yi Ren, Biwei Jiang, Yuxi Wang, Ming Yang, and Zhiqiang Yan. The Star Formation History in Local Group Galaxies. I. Ten Dwarf Galaxies. ApJL , 966(1):25, May 2024. doi: 10.3847/1538-4357/ad28c6.

[35] Erik J. Tollerud, James S. Bullock, Louis E. Strigari, and Beth Willman. Hundreds of Milky Way Satellites? Luminosity Bias in the Satellite Luminosity Function. ApJL , 688(1):277–289, November 2008. doi: 10.1086/592102.

[36] James S. Bullock. Notes on the Missing Satellites Problem. *arXiv e-prints*, art. arXiv:1009.4505, September 2010. doi: 10.48550/arXiv.1009.4505.

[37] Stacy Y. Kim, Annika H. G. Peter, and Jonathan R. Hargis. Missing Satellites Problem: Completeness Corrections to the Number of Satellite Galaxies in the Milky Way are Consistent with Cold Dark Matter Predictions. Phys. Rev. Lett. , 121(21):211302, November 2018. doi: 10.1103/PhysRevLett.121.211302.

[38] Amandine Doliva-Dolinsky, Michelle L. M. Collins, and Nicolas F. Martin. The satellite galaxies of the Milky Way and Andromeda. *arXiv e-prints*, art. arXiv:2502.06948, February 2025. doi: 10.48550/arXiv.2502.06948.

[39] Anna Genina, Alejandro Benítez-Llambay, Carlos S. Frenk, Shaun Cole, Azadeh Fattahi, Julio F. Navarro, Kyle A. Oman, Till Sawala, and Tom Theuns. The core-cusp problem: a matter of perspective. MNRAS , 474(1):1398–1411, February 2018. doi: 10.1093/mnras/stx2855.

[40] Heling Deng, Mark P. Hertzberg, Mohammad Hossein Namjoo, and Ali Masoumi. Can light dark matter solve the core-cusp problem? Phys. Rev. D , 98(2):023513, July 2018. doi: 10.1103/PhysRevD.98.023513.

[41] A. Del Popolo and F. Pace. The Cusp/Core problem: supernovae feedback versus the baryonic clumps and dynamical friction model. Ap&SS , 361(5):162, May 2016. doi: 10.1007/s10509-016-2742-z.

[42] R. Verbeke, E. Papastergis, A. A. Ponomareva, S. Rathi, and S. De Rijcke. A new astrophysical solution to the Too Big To Fail problem. Insights from the moria simulations. A&A , 607:A13, October 2017. doi: 10.1051/0004-6361/201730758.

[43] Mark R. Lovell, Violeta Gonzalez-Perez, Sownak Bose, Alexey Boyarsky, Shaun Cole, Carlos S. Frenk, and Oleg Ruchayskiy. Addressing the too big to fail problem with baryon physics and sterile neutrino dark matter. MNRAS , 468(3):2836–2849, July 2017. doi: 10.1093/mnras/stx621.

[44] Adam G. Riess, Wenlong Yuan, Lucas M. Macri, Dan Scolnic, Dillon Brout, Stefano Casertano, David O. Jones, Yukei Murakami, Gagandeep S. Anand, Louise Breuval, Thomas G. Brink, Alexei V. Filippenko, Samantha Hoffmann, Saurabh W. Jha, W. D'arcy Kenworthy, John Mackenty, Benjamin E. Stahl, and WeiKang Zheng. A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^{-1} Mpc^{-1} Uncertainty from the Hubble Space Telescope and the SH0ES Team. ApJ , 934(1):L7, July 2022. doi: 10.3847/2041-8213/ac5c5b.

[45] Eleonora Di Valentino, Olga Mena, Supriya Pan, Luca Visinelli, Weiqiang Yang, Alessandro Melchiorri, David F. Mota, Adam G. Riess, and Joseph Silk. In the realm of the Hubble tension-a review of solutions. *Classical and Quantum Gravity*, 38(15):153001, July 2021. doi: 10.1088/1361-6382/ac086d.

[46] L. Perivolaropoulos and F. Skara. Challenges for ΛCDM: An update. New A Rev. , 95:101659, December 2022. doi: 10.1016/j.newar.2022.101659.

[47] Shinji Tsujikawa. Quintessence: a review. *Classical and Quantum Gravity*, 30(21): 214003, November 2013. doi: 10.1088/0264-9381/30/21/214003.

[48] S. Shankaranarayanan and Joseph P. Johnson. Modified theories of gravity: Why, how and what? *General Relativity and Gravitation*, 54(5):44, May 2022. doi: 10.1007/s10714-022-02927-2.

[49] Lawrence M. Krauss, Scott Dodelson, and Stephan Meyer. Primordial Gravitational Waves and Cosmology. *Science*, 328(5981):989, May 2010. doi: 10.1126/science.1179541.

[50] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto. Non-Gaussianity from inflation: theory and observations. Phys. Rep. , 402(3-4):103–266, November 2004. doi: 10.1016/j.physrep.2004.08.022.

[51] David Lovelock. The Einstein Tensor and Its Generalizations. *Journal of Mathematical Physics*, 12(3):498–501, March 1971. doi: 10.1063/1.1665613.

[52] Thomas P. Sotiriou and Valerio Faraoni. f(R) theories of gravity. *Reviews of Modern Physics*, 82(1):451–497, January 2010. doi: 10.1103/RevModPhys.82.451.

[53] G. Dvali, G. Gabadadze, and M. Porrati. 4D gravity on a brane in 5D Minkowski space. *Physics Letters B*, 485(1-3):208–214, July 2000. doi: 10.1016/S0370-2693(00)00669-9.

[54] Wayne Hu and Ignacy Sawicki. Models of f(R) cosmic acceleration that evade solar system tests. Phys. Rev. D , 76(6):064004, September 2007. doi: 10.1103/PhysRevD.76.064004.

[55] Justin Khoury and Amanda Weltman. Chameleon cosmology. Phys. Rev. D , 69(4):044026, February 2004. doi: 10.1103/PhysRevD.69.044026.

[56] César Hernández-Aguayo, Jiamin Hou, Baojiu Li, Carlton M. Baugh, and Ariel G. Sánchez. Large-scale redshift space distortions in modified gravity theories. MNRAS , 485(2):2194–2213, May 2019. doi: 10.1093/mnras/stz516.

[57] A. I. Vainshtein. To the problem of nonvanishing gravitation mass. *Physics Letters B*, 39(3):393–394, May 1972. doi: 10.1016/0370-2693(72)90147-5.

[58] Tommaso Giannantonio, Yong-Seon Song, and Kazuya Koyama. Detectability of a phantom-like braneworld model with the integrated Sachs-Wolfe effect. Phys. Rev. D , 78(4):044017, August 2008. doi: 10.1103/PhysRevD.78.044017.

[59] Alexandre Barreira, Ariel G. Sánchez, and Fabian Schmidt. Validating estimates of the growth rate of structure with modified gravity simulations. Phys. Rev. D , 94(8):084022, October 2016. doi: 10.1103/PhysRevD.94.084022.

[60] Christopher T. Davies, Joachim Harnois-Déraps, Baojiu Li, Benjamin Giblin, César Hernández-Aguayo, and Enrique Paillas. Constraining modified gravity with weak-lensing peaks. MNRAS , 533(3):3546–3569, September 2024. doi: 10.1093/mnras/stae1966.

[61] Stuart Lyall, Chris Blake, Ryan Turner, Rossana Ruggeri, and Hans Winther. Testing modified gravity scenarios with direct peculiar velocities. MNRAS , 518(4):5929–5941, February 2023. doi: 10.1093/mnras/stac3323.

[62] Rien van de Weygaert and J. R. Bond. Observations and Morphology of the Cosmic Web. In M. Plionis, O. López-Cruz, and D. Hughes, editors, *A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure*, volume 740, page 24. 2008. doi: 10.1007/978-1-4020-6941-3_11.

[63] Wojciech A. Hellwing, Marius Cautun, Rien van de Weygaert, and Bernard T. Jones. Caught in the cosmic web: Environmental effect on halo concentrations, shape, and spin. Phys. Rev. D , 103(6):063517, March 2021. doi: 10.1103/PhysRevD.103.063517.

[64] Barbara Ryden. *Introduction to Cosmology*. 2016.

[65] C. L. Bennett, A. J. Banday, K. M. Gorski, G. Hinshaw, P. Jackson, P. Keegstra, A. Kogut, G. F. Smoot, D. T. Wilkinson, and E. L. Wright. Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results. ApJ , 464: L1, June 1996. doi: 10.1086/310075.

[66] Peter Schneider. *Extragalactic Astronomy and Cosmology: An Introduction*. 2015. doi: 10.1007/978-3-642-54083-7.

[67] P. J. E. Peebles. *The large-scale structure of the universe*. 1980.

[68] Eric V. Linder. Cosmic growth history and expansion history. Phys. Rev. D , 72 (4):043529, August 2005. doi: 10.1103/PhysRevD.72.043529.

[69] Yan-Chuan Cai, Nelson Padilla, and Baojiu Li. Testing gravity using cosmic voids. MNRAS , 451(1):1036–1055, July 2015. doi: 10.1093/mnras/stv777.

[70] Ixandra Achitouv, Marco Baldi, Ewald Puchwein, and Jochen Weller. Imprint of f (R) gravity on nonlinear structure formation. Phys. Rev. D , 93(10):103522, May 2016. doi: 10.1103/PhysRevD.93.103522.

[71] Suhani Gupta, Wojciech A. Hellwing, Maciej Bilicki, and Jorge Enrique García-Farieta. Universality of the halo mass function in modified gravity cosmologies. Phys. Rev. D , 105(4):043538, February 2022. doi: 10.1103/PhysRevD.105.043538.

[72] Wojciech A. Hellwing, Alexandre Barreira, Carlos S. Frenk, Baojiu Li, and Shaun Cole. Clear and Measurable Signature of Modified Gravity in the Galaxy Velocity Field. Phys. Rev. Lett. , 112(22):221102, June 2014. doi: 10.1103/PhysRevLett.112.221102.

[73] Ya. B. Zel'dovich. Gravitational instability: An approximate theory for large density perturbations. A&A , 5:84–89, March 1970.

[74] James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. 2008.

[75] Ravi K. Sheth, H. J. Mo, and Giuseppe Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. MNRAS , 323(1):1–12, May 2001. doi: 10.1046/j.1365-8711.2001.04006.x.

[76] Daniel Baumann. *Cosmology*. 2022. doi: 10.1017/9781108937092.

[77] William H. Press and Paul Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. ApJL , 187:425–438, February 1974. doi: 10.1086/152650.

[78] Malcolm S. Longair. *Galaxy Formation*. 2008.

[79] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White. The Aquarius Project: the subhaloes of galactic haloes. MNRAS , 391(4):1685–1711, December 2008. doi: 10.1111/j.1365-2966.2008.14066.x.

[80] Robert A. Crain, Joop Schaye, Richard G. Bower, Michelle Furlong, Matthieu Schaller, Tom Theuns, Claudio Dalla Vecchia, Carlos S. Frenk, Ian G. McCarthy, John C. Helly, Adrian Jenkins, Yetli M. Rosas-Guevara, Simon D. M. White, and James W. Trayford. The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations. MNRAS , 450(2):1937–1961, June 2015. doi: 10.1093/mnras/stv725.

[81] Anatoly Klypin, Gustavo Yepes, Stefan Gottlöber, Francisco Prada, and Steffen Heß. MultiDark simulations: the story of dark matter halo concentrations and density profiles. MNRAS , 457(4):4340–4359, April 2016. doi: 10.1093/mnras/stw248.

[82] Dylan Nelson, Volker Springel, Annalisa Pillepich, Vicente Rodriguez-Gomez, Paul Torrey, Shy Genel, Mark Vogelsberger, Ruediger Pakmor, Federico Marinacci, Rainer Weinberger, Luke Kelley, Mark Lovell, Benedikt Diemer, and Lars Hernquist. The IllustrisTNG simulations: public data release. *Computational Astrophysics and Cosmology*, 6(1):2, May 2019. doi: 10.1186/s40668-019-0028-x.

[83] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. A Universal Density Profile from Hierarchical Clustering. ApJL , 490(2):493–508, December 1997. doi: 10.1086/304888.

[84] Houjun Mo, Frank C. van den Bosch, and Simon White. *Galaxy Formation and Evolution*. 2010. doi: 10.1017/CBO9780511807244.

[85] Gabriella De Lucia, Volker Springel, Simon D. M. White, Darren Croton, and Guinevere Kauffmann. The formation history of elliptical galaxies. MNRAS , 366(2):499–509, February 2006. doi: 10.1111/j.1365-2966.2005.09879.x.

[86] R. B. Partridge and P. J. E. Peebles. Are Young Galaxies Visible? ApJL , 147:868, March 1967. doi: 10.1086/149079.

[87] S. de Rijcke, D. Michielsen, H. Dejonghe, W. W. Zeilinger, and G. K. T. Hau. Formation and evolution of dwarf elliptical galaxies. I. Structural and kinematical properties. A&A , 438(2):491–505, August 2005. doi: 10.1051/0004-6361:20042213.

[88] V. V. Bobylev and A. T. Bajkova. Review of current estimates of the Galaxy mass. *Publications of the Pulkovo Observatory*, 228:57–76, March 2023. doi: 10.31725/0367-7966-2023-228-3.

[89] R. B. Tully and J. R. Fisher. A new method of determining distances to galaxies. A&A , 54:661–673, February 1977.

[90] R. Brent Tully, Ehsan Kourkchi, Hélène M. Courtois, Gagandeep S. Anand, John P. Blakeslee, Dillon Brout, Thomas de Jaeger, Alexandra Dupuy, Daniel Guinet, Cullan Howlett, Joseph B. Jensen, Daniel Pomarède, Luca Rizzi, David Rubin, Khaled Said, Daniel Scolnic, and Benjamin E. Stahl. Cosmicflows-4. ApJL , 944(1):94, February 2023. doi: 10.3847/1538-4357/ac94d8.

[91] Alejandro Benitez-Llambay and Carlos Frenk. The detailed structure and the onset of galaxy formation in low-mass gaseous dark matter haloes. MNRAS , 498(4):4887–4900, November 2020. doi: 10.1093/mnras/staa2698.

[92] Mihir Kulkarni, Eli Visbal, and Greg L. Bryan. The Critical Dark Matter Halo Mass for Population III Star Formation: Dependence on Lyman-Werner Radiation, Baryon-dark Matter Streaming Velocity, and Redshift. ApJL , 917(1):40, August 2021. doi: 10.3847/1538-4357/ac08a3.

[93] J. A. Peacock. Large-scale surveys and cosmic structure. *arXiv e-prints*, art. astro-ph/0309240, September 2003. doi: 10.48550/arXiv.astro-ph/0309240.

[94] J. N. Fry and Enrique Gaztanaga. Biasing and Hierarchical Statistics in Large-Scale Structure. ApJL , 413:447, August 1993. doi: 10.1086/173015.

[95] Vincent Desjacques, Donghui Jeong, and Fabian Schmidt. Large-scale galaxy bias. Phys. Rep. , 733:1–193, February 2018. doi: 10.1016/j.physrep.2017.12.002.

[96] Marisa Girardi, Giuliano Giuricin, Fabio Mardirossian, Marino Mezzetti, and Walter Boschin. Optical Mass Estimates of Galaxy Clusters. ApJL , 505(1):74–95, September 1998. doi: 10.1086/306157.

[97] Nick Kaiser. Clustering in real space and in redshift space. MNRAS , 227:1–21, July 1987. doi: 10.1093/mnras/227.1.1.

[98] Henrietta S. Leavitt and Edward C. Pickering. Periods of 25 Variable Stars in the Small Magellanic Cloud. *Harvard College Observatory Circular*, 173:1–3, March 1912.

[99] M. M. Phillips. The Absolute Magnitudes of Type IA Supernovae. ApJ , 413:L105, August 1993. doi: 10.1086/186970.

[100] Alan Dressler, Donald Lynden-Bell, David Burstein, Roger L. Davies, S. M. Faber, Roberto Terlevich, and Gary Wegner. Spectroscopy and Photometry of Elliptical Galaxies. I. New Distance Estimator. ApJL , 313:42, February 1987. doi: 10.1086/164947.

[101] Idit Zehavi, Zheng Zheng, David H. Weinberg, Michael R. Blanton, Neta A. Bahcall, Andreas A. Berlind, Jon Brinkmann, Joshua A. Frieman, James E. Gunn, Robert H. Lupton, Robert C. Nichol, Will J. Percival, Donald P. Schneider, Ramin A. Skibba, Michael A. Strauss, Max Tegmark, and Donald G. York. Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity. ApJL , 736(1):59, July 2011. doi: 10.1088/0004-637X/736/1/59.

[102] Hong Guo, Idit Zehavi, Zheng Zheng, David H. Weinberg, Andreas A. Berlind, Michael Blanton, Yanmei Chen, Daniel J. Eisenstein, Shirley Ho, Eyal Kazin, Marc Manera, Claudia Maraston, Cameron K. McBride, Sebastián E. Nuza, Nikhil Padmanabhan, John K. Parejko, Will J. Percival, Ashley J. Ross, Nicholas P. Ross, Lado Samushia, Ariel G. Sánchez, David J. Schlegel, Donald P. Schneider, Ramin A. Skibba, Molly E. C. Swanson, Jeremy L. Tinker, Rita Tojeiro, David A. Wake, Martin White, Neta A. Bahcall, Dmitry Bizyaev, Howard Brewington, Kevin Bundy, Luiz N. A. da Costa, Garrett Ebelke, Elena Malanushenko, Viktor Malanushenko, Daniel Oravetz, Graziano Rossi, Audrey Simmons, Stephanie Snedden, Alina Streblyanska, and Daniel Thomas. The Clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Luminosity and Color Dependence and Redshift Evolution. ApJL , 767(2):122, April 2013. doi: 10.1088/0004-637X/767/2/122.

[103] F. S. Kitaura, F. Sinigaglia, A. Balaguera-Antolínez, and G. Favole. The cosmic web from perturbation theory. A&A , 683:A215, March 2024. doi: 10.1051/0004-6361/202345876.

[104] Rayne Liu, Georgios Valogiannis, Nicholas Battaglia, and Rachel Bean. Constraints on f (R) and normal-branch Dvali-Gabadadze-Porrati modified gravity model parameters with cluster abundances and galaxy clustering. Phys. Rev. D , 104(10): 103519, November 2021. doi: 10.1103/PhysRevD.104.103519.

[105] Risa H. Wechsler and Jeremy L. Tinker. The Connection Between Galaxies and Their Dark Matter Halos. ARA&A , 56:435–487, September 2018. doi: 10.1146/annurev-astro-081817-051756.

[106] S. D. M. White, C. S. Frenk, and M. Davis. Clustering in a neutrino-dominated universe. ApJ , 274:L1–L5, November 1983. doi: 10.1086/184139.

[107] Bodo Schwabe, Mateja Gosenca, Christoph Behrens, Jens C. Niemeyer, and Richard Easther. Simulating mixed fuzzy and cold dark matter. Phys. Rev. D , 102(8): 083518, October 2020. doi: 10.1103/PhysRevD.102.083518.

[108] A. Gruzinov, M. Kleban, M. Porrati, and M. Redi. Gravitational backreaction of matter inhomogeneities. J. Cosmology Astropart. Phys. , 2006(12):001, December 2006. doi: 10.1088/1475-7516/2006/12/001.

[109] Nan Li and Dominik J. Schwarz. Scale dependence of cosmological backreaction. Phys. Rev. D , 78(8):083531, October 2008. doi: 10.1103/PhysRevD.78.083531.

[110] 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, and David N. Spergel. The Quijote Simulations. ApJS , 250(1):2, September 2020. doi: 10.3847/1538-4365/ab9d82.

[111] J. Liske, I. K. Baldry, S. P. Driver, R. J. Tuffs, M. Alpaslan, E. Andrae, S. Brough, M. E. Cluver, M. W. Grootes, M. L. P. Gunawardhana, L. S. Kelvin, J. Loveday, A. S. G. Robotham, E. N. Taylor, S. P. Bamford, J. Bland-Hawthorn, M. J. I. Brown, M. J. Drinkwater, A. M. Hopkins, M. J. Meyer, P. Norberg, J. A. Peacock, N. K. Agius, S. K. Andrews, A. E. Bauer, J. H. Y. Ching, M. Colless, C. J. Conselice, S. M. Croom, L. J. M. Davies, R. De Propris, L. Dunne, E. M. Eardley, S. Ellis, C. Foster, C. S. Frenk, B. Häußler, B. W. Holwerda, C. Howlett, H. Ibarra, M. J. Jarvis, D. H. Jones, P. R. Kafle, C. G. Lacey, R. Lange, M. A. Lara-López, Á. R. López-Sánchez, S. Maddox, B. F. Madore, T. McNaught-Roberts, A. J. Moffett, R. C. Nichol, M. S. Owers, D. Palamara, S. J. Penny, S. Phillipps, K. A. Pimbblet, C. C. Popescu, M. Prescott, R. Proctor, E. M. Sadler, A. E. Sansom, M. Seibert, R. Sharp, W. Sutherland, J. A. Vázquez-Mata, E. van Kampen, S. M. Wilkins, R. Williams, and A. H. Wright. Galaxy And Mass Assembly (GAMA): end of survey report and data release 2. MNRAS , 452(2):2087–2126, September 2015. doi: 10.1093/mnras/stv1436.

[112] Abdurro'uf, Katherine Accetta, Conny Aerts, Víctor Silva Aguirre, Romina Ahumada, Nikhil Ajgaonkar, N. Filiz Ak, Shadab Alam, Carlos Allende Prieto, Andrés Almeida, Friedrich Anders, Scott F. Anderson, Brett H. Andrews, Borja Anguiano, Erik Aquino-Ortíz, Alfonso Aragón-Salamanca, Maria Argudo-Fernández, Metin Ata, Marie Aubert, Vladimir Avila-Reese, Carles Badenes, Rodolfo H. Barbá,

Barger, et al. The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data. ApJS , 259(2):35, April 2022. doi: 10.3847/1538-4365/ac4414.

[113] DESI Collaboration, A. G. Adame, J. Aguilar, S. Ahlen, S. Alam, G. Aldering, D. M. Alexander, R. Alfarsy, C. Allende Prieto, M. Alvarez, O. Alves, A. Anand, F. Andrade-Oliveira, E. Armengaud, J. Asorey, S. Avila, A. Aviles, S. Bailey, A. Balaguera-Antolínez, O. Ballester, C. Baltay, A. Bault, J. Bautista, J. Behera, S. F. Beltran, S. BenZvi, L. Beraldo e Silva, J. R. Bermejo-Climent, A. Berti, R. Besuner, F. Beutler, D. Bianchi, C. Blake, R. Blum, A. S. Bolton, S. Brieden, A. Brodzeller, D. Brooks, Brown, et al. The Early Data Release of the Dark Energy Spectroscopic Instrument. AJ , 168(2):58, August 2024. doi: 10.3847/1538-3881/ad3217.

[114] Roi Kugel, Joop Schaye, Matthieu Schaller, John C. Helly, Joey Braspenning, Willem Elbers, Carlos S. Frenk, Ian G. McCarthy, Juliana Kwan, Jaime Salcido, Marcel P. van Daalen, Bert Vandenbroucke, Yannick M. Bahé, Josh Borrow, Evgenii Chaikin, Filip Huško, Adrian Jenkins, Cedric G. Lacey, Folkert S. J. Nobels, and Ian Vernon. FLAMINGO: calibrating large cosmological hydrodynamical simulations with machine learning. MNRAS , 526(4):6103–6127, December 2023. doi: 10.1093/mnras/stad2540.

[115] Andreas A. Berlind, David H. Weinberg, Andrew J. Benson, Carlton M. Baugh, Shaun Cole, Romeel Davé, Carlos S. Frenk, Adrian Jenkins, Neal Katz, and Cedric G. Lacey. The Halo Occupation Distribution and the Physics of Galaxy Formation. ApJL , 593(1):1–25, August 2003. doi: 10.1086/376517.

[116] A. Vale and J. P. Ostriker. Linking halo mass to galaxy luminosity. MNRAS , 353 (1):189–200, September 2004. doi: 10.1111/j.1365-2966.2004.08059.x.

[117] Asantha Cooray. Halo model at its best: constraints on conditional luminosity functions from measured galaxy statistics. MNRAS , 365(3):842–866, January 2006. doi: 10.1111/j.1365-2966.2005.09747.x.

[118] Marius Cautun, Enrique Paillas, Yan-Chuan Cai, Sownak Bose, Joaquin Armijo, Baojiu Li, and Nelson Padilla. The Santiago-Harvard-Edinburgh-Durham void comparison - I. SHEDding light on chameleon gravity tests. MNRAS , 476(3):3195–3217, May 2018. doi: 10.1093/mnras/sty463.

[119] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP)

Observations: Cosmological Parameter Results. ApJS , 208(2):19, October 2013. doi: 10.1088/0067-0049/208/2/19.

[120] Baojiu Li, Gong-Bo Zhao, Romain Teyssier, and Kazuya Koyama. ECOSMOG: an Efficient COde for Simulating MOdified Gravity. J. Cosmology Astropart. Phys. , 2012(1):051, January 2012. doi: 10.1088/1475-7516/2012/01/051.

[121] Alexandre Barreira, Sownak Bose, and Baojiu Li. Speeding up N-body simulations of modified gravity: Vainshtein screening models. J. Cosmology Astropart. Phys. , 2015(12):059–059, December 2015. doi: 10.1088/1475-7516/2015/12/059.

[122] R. Teyssier. Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES. A&A , 385:337–364, April 2002. doi: 10.1051/0004-6361:20011817.

[123] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, et al. Planck 2018 results. VI. Cosmological parameters. A&A , 641:A6, September 2020. doi: 10.1051/0004-6361/201833910.

[124] Lawrence M. Krauss, Scott Dodelson, and Stephan Meyer. Primordial Gravitational Waves and Cosmology. *Science*, 328(5981):989, May 2010. doi: 10.1126/science.1179541.

[125] Edwin Sirko. Initial Conditions to Cosmological N-Body Simulations, or, How to Run an Ensemble of Simulations. ApJL , 634(2):728–743, November 2005. doi: 10.1086/497090.

[126] Simon D. M. White. Formation and Evolution of Galaxies: Les Houches Lectures. *arXiv e-prints*, art. astro-ph/9410043, October 1994. doi: 10.48550/arXiv.astro-ph/9410043.

[127] HMP Couchman. Cosmological simulations using particle-mesh methods. 1992.

[128] Svetlin Tassev, Matias Zaldarriaga, and Daniel J. Eisenstein. Solving large scale structure in ten easy steps with COLA. J. Cosmology Astropart. Phys. , 2013(6):036, June 2013. doi: 10.1088/1475-7516/2013/06/036.

[129] The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 292, May 15, 1985, p. 371-394. Research supported by the Science and Engineering Research Council of England and NASA.*, 292:371–394, 1985.

[130] Peter S. Behroozi, Risa H. Wechsler, and Hao-Yi Wu. The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores. ApJL , 762 (2):109, January 2013. doi: 10.1088/0004-637X/762/2/109.

[131] Pascal J. Elahi, Rodrigo Cañas, Rhys J. J. Poulton, Rodrigo J. Tobar, James S. Willis, Claudia del P. Lagos, Chris Power, and Aaron S. G. Robotham. Hunting for galaxies and halos in simulations with VELOCIraptor. PASA , 36:e021, May 2019. doi: 10.1017/pasa.2019.12.

[132] Jiaxin Han, Y. P. Jing, Huiyuan Wang, and Wenting Wang. Resolving subhaloes' lives with the Hierarchical Bound-Tracing algorithm. MNRAS , 427(3):2437–2449, December 2012. doi: 10.1111/j.1365-2966.2012.22111.x.

[133] Alexander Knebe, Steffen R. Knollmann, Stuart I. Muldrew, Frazer R. Pearce, Miguel Angel Aragon-Calvo, Yago Ascasibar, Peter S. Behroozi, Daniel Ceverino, Stephane Colombi, Juerg Diemand, Klaus Dolag, Bridget L. Falck, Patricia Fasel, Jeff Gardner, Stefan Gottlöber, Chung-Hsing Hsu, Francesca Iannuzzi, Anatoly Klypin, Zarija Lukić, Michal Maciejewski, Cameron McBride, Mark C. Neyrinck, Susana Planelles, Doug Potter, Vicent Quilis, Yann Rasera, Justin I. Read, Paul M. Ricker, Fabrice Roy, Volker Springel, Joachim Stadel, Greg Stinson, P. M. Sutter, Victor Turchaninov, Dylan Tweed, Gustavo Yepes, and Marcel Zemp. Haloes gone MAD: The Halo-Finder Comparison Project. MNRAS , 415(3):2293–2318, August 2011. doi: 10.1111/j.1365-2966.2011.18858.x.

[134] Mikito Yamamoto, Shogo Masaki, and Chiaki Hikage. Testing subhalo abundance matching from redshift-space clustering. *arXiv e-prints*, art. arXiv:1503.03973, March 2015. doi: 10.48550/arXiv.1503.03973.

[135] Hong Guo, Zheng Zheng, Peter S. Behroozi, Idit Zehavi, Johan Comparat, Ginevra Favole, Stefan Gottlöber, Anatoly Klypin, Francisco Prada, Sergio A. Rodríguez-Torres, David H. Weinberg, and Gustavo Yepes. Galaxy Three-point Correlation Functions and Halo/Subhalo Models. ApJL , 831(1):3, November 2016. doi: 10.3847/0004-637X/831/1/3.

[136] Shadab Alam, Christian Arnold, Alejandro Aviles, Rachel Bean, Yan-Chuan Cai, Marius Cautun, Jorge L. Cervantes-Cota, Carolina Cuesta-Lazaro, N. Chandrachani Devi, Alexander Eggemeier, Sebastien Fromenteau, Alma X. Gonzalez-Morales, Vitali Halenka, Jian-hua He, Wojciech A. Hellwing, César Hernández-Aguayo,

Mustapha Ishak, Kazuya Koyama, Baojiu Li, Axel de la Macorra, Jennifer Meneses Rizo, Christopher Miller, Eva-Maria Mueller, Gustavo Niz, Pierros Ntelis, Matia Rodríguez Otero, Cristiano G. Sabiu, Zachary Slepian, Alejo Stark, Octavio Valenzuela, Georgios Valogiannis, Mariana Vargas-Magaña, Hans A. Winther, Pauline Zarrouk, Gong-Bo Zhao, and Yi Zheng. Towards testing the theory of gravity with DESI: summary statistics, model predictions and future simulation requirements. J. Cosmology Astropart. Phys. , 2021(11):050, November 2021. doi: 10.1088/1475-7516/2021/11/050.

[137] Dylan Britt, Daniel Gruen, Oliver Friedrich, Sihan Yuan, and Bernardita Ried Guachalla. Bounds on galaxy stochasticity from halo occupation distribution modeling. A&A , 689:A253, September 2024. doi: 10.1051/0004-6361/202450266.

[138] Gillian D. Beltz-Mohrmann, Adam O. Szewciw, Andreas A. Berlind, and Manodeep Sinha. Toward Accurate Modeling of Galaxy Clustering on Small Scales: Halo Model Extensions and Lingering Tension. ApJL , 948(2):100, May 2023. doi: 10.3847/1538-4357/acc576.

[139] Myles A. Mitchell, Christian Arnold, Jian-hua He, and Baojiu Li. A general framework to test gravity using galaxy clusters II: A universal model for the halo concentration in f(R) gravity. MNRAS , 487(1):1410–1425, July 2019. doi: 10.1093/mnras/stz1389.

[140] Myles A. Mitchell, César Hernández-Aguayo, Christian Arnold, and Baojiu Li. A general framework to test gravity using galaxy clusters IV: cluster and halo properties in DGP gravity. MNRAS , 508(3):4140–4156, December 2021. doi: 10.1093/mnras/stab2817.

[141] Chun-Hao To, Rachel M. Reddick, Eduardo Rozo, Eli Rykoff, and Risa H. Wechsler. RedMaPPer: Evolution and Mass Dependence of the Conditional Luminosity Functions of Red Galaxies in Galaxy Clusters. ApJL , 897(1):15, July 2020. doi: 10.3847/1538-4357/ab9636.

[142] Joseph Hollowed. Lightcone Construction for HACC Cosmological Simulations with LANTERN. *arXiv e-prints*, art. arXiv:1906.08355, June 2019. doi: 10.48550/arXiv.1906.08355.

[143] Alex Smith, Shaun Cole, Cameron Grove, Peder Norberg, and Pauline Zarrouk. A light-cone catalogue from the Millennium-XXL simulation: improved spatial interpolation and colour distributions for the DESI BGS. MNRAS , 516(3):4529–4542, November 2022. doi: 10.1093/mnras/stac2519.

[144] Alexander I. Merson, Carlton M. Baugh, John C. Helly, Violeta Gonzalez-Perez, Shaun Cole, Richard Bielby, Peder Norberg, Carlos S. Frenk, Andrew J. Benson,

Richard G. Bower, Cedric G. Lacey, and Claudia del P. Lagos. Lightcone mock catalogues from semi-analytic models of galaxy formation - I. Construction and application to the BzK colour selection. MNRAS , 429(1):556–578, February 2013. doi: 10.1093/mnras/sts355.

[145] David W. Hogg, Ivan K. Baldry, Michael R. Blanton, and Daniel J. Eisenstein. The K correction. *arXiv e-prints*, art. astro-ph/0210394, October 2002. doi: 10.48550/arXiv.astro-ph/0210394.

[146] Ashley J. Ross, Robert J. Brunner, and Adam D. Myers. Higher Order Angular Galaxy Correlations in the SDSS: Redshift and Color Dependence of Nonlinear Bias. ApJL , 665(1):67–84, August 2007. doi: 10.1086/519020.

[147] D. J. Farrow, Shaun Cole, Peder Norberg, N. Metcalfe, I. Baldry, Joss Bland-Hawthorn, Michael J. I. Brown, A. M. Hopkins, Cedric G. Lacey, J. Liske, Jon Loveday, David P. Palamara, A. S. G. Robotham, and Srivatsan Sridhar. Galaxy and mass assembly (GAMA): projected galaxy clustering. MNRAS , 454(2):2120–2145, December 2015. doi: 10.1093/mnras/stv2075.

[148] Vincent R. Eke, Julio F. Navarro, and Matthias Steinmetz. The Power Spectrum Dependence of Dark Matter Halo Concentrations. ApJL , 554(1):114–125, June 2001. doi: 10.1086/321345.

[149] Philippe Brax and Patrick Valageas. Impact on the power spectrum of screening in modified gravity scenarios. Phys. Rev. D , 88(2):023527, July 2013. doi: 10.1103/PhysRevD.88.023527.

[150] Jorge Enrique García-Farieta, Wojciech A. Hellwing, Suhani Gupta, and Maciej Bilicki. Probing gravity with redshift-space distortions: Effects of tracer bias and sample selection. Phys. Rev. D , 103(10):103524, May 2021. doi: 10.1103/PhysRevD.103.103524.

[151] Ed Hawkins, Steve Maddox, Shaun Cole, Ofer Lahav, Darren S. Madgwick, Peder Norberg, John A. Peacock, Ivan K. Baldry, Carlton M. Baugh, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, Matthew Colless, Chris Collins, Warrick Couch, Gavin Dalton, Roberto De Propris, Simon P. Driver, George Efstathiou, Richard S. Ellis, Carlos S. Frenk, Karl Glazebrook, Carole Jackson, Bryn Jones, Ian Lewis, Stuart Lumsden, Will Percival, Bruce A. Peterson, Will Sutherland, and Keith Taylor. The 2dF Galaxy Redshift Survey: correlation functions, peculiar velocities and the matter density of the Universe. MNRAS , 346(1):78–96, November 2003. doi: 10.1046/j.1365-2966.2003.07063.x.

[152] Nick Mostek, Alison L. Coil, Michael Cooper, Marc Davis, Jeffrey A. Newman, and Benjamin J. Weiner. The DEEP2 Galaxy Redshift Survey: Clustering Dependence

on Galaxy Stellar Mass and Star Formation Rate at z ~1. ApJL , 767(1):89, April 2013. doi: 10.1088/0004-637X/767/1/89.

[153] ChangHoon Hahn, Roman Scoccimarro, Michael R. Blanton, Jeremy L. Tinker, and Sergio A. Rodríguez-Torres. The Effect of Fiber Collisions on the Galaxy Power Spectrum Multipoles. MNRAS , 467(2):1940–1956, May 2017. doi: 10.1093/mnras/stx185.

[154] Stephen D. Landy and Alexander S. Szalay. Bias and Variance of Angular Correlation Functions. ApJL , 412:64, July 1993. doi: 10.1086/172900.

[155] DESI Collaboration, A. G. Adame, J. Aguilar, S. Ahlen, S. Alam, D. M. Alexander, M. Alvarez, O. Alves, A. Anand, U. Andrade, E. Armengaud, S. Avila, A. Aviles, H. Awan, S. Bailey, C. Baltay, A. Bault, J. Behera, S. BenZvi, F. Beutler, D. Bianchi, C. Blake, R. Blum, S. Brieden, A. Brodzeller, D. Brooks, Z. Brown, E. Buckley-Geer, E. Burtin, R. Calderon, R. Canning, A. Carnero Rosell, R. Cereskaite, J. L. Cervantes-Cota, S. Chabanier, E. Chaussidon, J. Chaves-Montero, S. Chen, X. Chen, T. Claybaugh, S. Cole, et al. DESI 2024 II: Sample Definitions, Characteristics, and Two-point Clustering Statistics. *arXiv e-prints*, art. arXiv:2411.12020, November 2024. doi: 10.48550/arXiv.2411.12020.

[156] U. Sureshkumar, A. Durkalec, A. Pollo, M. Bilicki, J. Loveday, D. J. Farrow, B. W. Holwerda, A. M. Hopkins, J. Liske, K. A. Pimbblet, E. N. Taylor, and A. H. Wright. Galaxy and Mass Assembly (GAMA). Tracing galaxy environment using the marked correlation function. A&A , 653:A35, September 2021. doi: 10.1051/0004-6361/202140500.

[157] Peter Coles and Francesco Lucchin. *Cosmology: The Origin and Evolution of Cosmic Structure, Second Edition*. 2002.

[158] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro. Large-scale structure of the Universe and cosmological perturbation theory. Phys. Rep. , 367(1-3): 1–248, September 2002. doi: 10.1016/S0370-1573(02)00135-7.

[159] Zachary Slepian and Daniel J. Eisenstein. A practical computational method for the anisotropic redshift-space three-point correlation function. MNRAS , 478(2): 1468–1483, August 2018. doi: 10.1093/mnras/sty1063.

[160] Jiamin Hou, Zachary Slepian, and Robert N. Cahn. Measurement of parity-odd modes in the large-scale 4-point correlation function of Sloan Digital Sky Survey Baryon Oscillation Spectroscopic Survey twelfth data release CMASS and LOWZ galaxies. MNRAS , 522(4):5701–5739, May 2023. doi: 10.1093/mnras/stad1062.

[161] Xiaochun Luo and David N. Schramm. Kurtosis, Skewness, and Non-Gaussian Cosmological Density Perturbations. ApJL , 408:33, May 1993. doi: 10.1086/172567.

[162] E. Gaztanaga and F. Bernardeau. The skewness and kurtosis of the projected density distribution function: validity of perturbation theory. A&A , 331:829–837, March 1998. doi: 10.48550/arXiv.astro-ph/9707095.

[163] J. N. Fry and Enrique Gaztanaga. Redshift Distortions of Galaxy Correlation Functions. ApJL , 425:1, April 1994. doi: 10.1086/173956.

[164] M. H. Goroff, B. Grinstein, S. J. Rey, and M. B. Wise. Coupling of modes of cosmological mass density fluctuations. ApJL , 311:6–14, December 1986. doi: 10.1086/164749.

[165] Agnieszka Pollo. Gravitational Instability and 2-D Galaxy Surveys. Acta Astron. , 47:413–429, October 1997.

[166] Francis Bernardeau. The Gravity-induced Quasi-Gaussian Correlation Hierarchy. ApJL , 392:1, June 1992. doi: 10.1086/171398.

[167] E. Gaztanaga. High-Order Galaxy Correlation Functions in the APM Galaxy Survey. MNRAS , 268:913, June 1994. doi: 10.1093/mnras/268.4.913.

[168] C. M. Baugh, E. Gaztanaga, and G. Efstathiou. A comparison of the evolution of density fields in perturbation theory and numerical simulations - II. Counts-in-cells analysis. MNRAS , 274(4):1049–1070, June 1995. doi: 10.1093/mnras/274.4.1049.

[169] Wojciech A. Hellwing, Kazuya Koyama, Benjamin Bose, and Gong-Bo Zhao. Revealing modified gravity signals in matter and halo hierarchical clustering. Phys. Rev. D , 96(2):023515, July 2017. doi: 10.1103/PhysRevD.96.023515.

[170] Wojciech A. Hellwing. The skewness of z=0.5 redshift-space galaxy distribution in Modified Gravity. In Katarzyna Małek, Magdalena Polińska, Agnieszka Majczyna, Grzegorz Stachowski, Radosław Poleski, Łukasz Wyrzykowski, and Agata Różańska, editors, *XXXIX Polish Astronomical Society Meeting*, volume 10, pages 315–322, October 2020. doi: 10.48550/arXiv.1912.13026.

[171] C. M. Baugh, D. J. Croton, E. Gaztañaga, P. Norberg, M. Colless, I. K. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, C. Collins, W. Couch, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, J. A. Peacock, B. A. Peterson, W. Sutherland, K. Taylor, and 2dFGRS Team. The 2dF Galaxy Redshift Survey: hierarchical galaxy clustering. MNRAS , 351(2):L44–L48, June 2004. doi: 10.1111/j.1365-2966.2004.07962.x.

[172] D. J. Croton, E. Gaztañaga, C. M. Baugh, P. Norberg, M. Colless, I. K. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, C. Collins, W. Couch, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, J. A. Peacock, B. A. Peterson, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: higher-order galaxy correlation functions. MNRAS , 352(4):1232–1244, August 2004. doi: 10.1111/j.1365-2966.2004.08017.x.

[173] Wojciech A. Hellwing, Baojiu Li, Carlos S. Frenk, and Shaun Cole. Hierarchical clustering in chameleon f(R) gravity. MNRAS , 435(4):2806–2821, November 2013. doi: 10.1093/mnras/stt1430.

[174] F. Bernardeau. The effects of smoothing on the statistical properties of large-scale cosmic fields. A&A , 291:697–712, November 1994. doi: 10.48550/arXiv.astro-ph/9403020.

[175] Enrique Gaztanaga and Jun'ichi Yokoyama. Probing the Statistics of Primordial Fluctuations and Their Evolution. ApJL , 403:450, February 1993. doi: 10.1086/172216.

[176] J. S. Bagla, Jayanti Prasad, and Nishikanta Khandai. Effects of the size of cosmological N-body simulations on physical quantities - III. Skewness. MNRAS , 395(2):918–930, May 2009. doi: 10.1111/j.1365-2966.2009.14592.x.

[177] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. ApJL , 622(2):759–771, April 2005. doi: 10.1086/427976.

[178] Román Scoccimarro. Redshift-space distortions, pairwise velocities, and nonlinearities. Phys. Rev. D , 70(8):083007, October 2004. doi: 10.1103/PhysRevD.70.083007.

[179] ChangHoon Hahn, Michael J. Wilson, Omar Ruiz-Macias, Shaun Cole, David H. Weinberg, John Moustakas, Anthony Kremin, Jeremy L. Tinker, Alex Smith, Risa H. Wechsler, Steven Ahlen, Shadab Alam, Stephen Bailey, David Brooks, Andrew P. Cooper, Tamara M. Davis, Kyle Dawson, Arjun Dey, Biprateep Dey, et al. The DESI Bright Galaxy Survey: Final Target Selection, Design, and Validation. AJ , 165(6):253, June 2023. doi: 10.3847/1538-3881/accff8.

[180] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio. The Square Kilometre Array. *IEEE Proceedings*, 97(8):1482–1496, August 2009. doi: 10.1109/JPROC.2009.2021005.

[181] G. Riccio, K. Małek, A. Nanni, M. Boquien, V. Buat, D. Burgarella, D. Donevski, M. Hamed, P. Hurley, R. Shirley, and A. Pollo. Preparing for LSST data. Estimating the physical properties of z $<$ 2.5 main-sequence galaxies. A&A , 653:A107, September 2021. doi: 10.1051/0004-6361/202140854.

[182] Euclid Collaboration, R. Scaramella, J. Amiaux, Y. Mellier, C. Burigana, C. S. Carvalho, J. C. Cuillandre, A. Da Silva, A. Derosa, J. Dinis, E. Maiorano, M. Maris, I. Tereno, R. Laureijs, T. Boenke, G. Buenadicha, X. Dupac, L. M. Gaspar Venancio, P. Gómez-Álvarez, J. Hoar, J. Lorenzo Alvarez, G. D. Racca, G. Saavedra-Criado, J. Schwartz, R. Vavrek, M. Schirmer, H. Aussel, R. Azzollini, V. F. Cardone, et al. Euclid preparation. I. The Euclid Wide Survey. A&A , 662:A112, June 2022. doi: 10.1051/0004-6361/202141938.

[183] Željko Ivezić, Steven M. Kahn, J. Anthony Tyson, Bob Abel, Emily Acosta, Robyn Allsman, David Alonso, Yusra AlSayyad, Scott F. Anderson, John Andrew, James Roger P. Angel, George Z. Angeli, Reza Ansari, Pierre Antilogus, Constanza Araujo, Robert Armstrong, Kirk T. Arndt, Pierre Astier, Éric Aubourg, Nicole Auza, Tim S. Axelrod, Deborah J. Bard, Jeff D. Barr, Aurelian Barrau, James G. Bartlett, Amanda E. Bauer, Brian J. Bauman, Sylvain Baumont, Ellen Bechtol, Keith Bechtol, Andrew C. Becker, Jacek Becla, Cristina Beldica, Steve Bellavia, Federica B. Bianco, Rahul Biswas, Guillaume Blanc, Jonathan Blazek, Roger D. Blandford, Josh S. Bloom, Joanne Bogart, Tim W. Bond, Michael T. Booth, Anders W. Borgland, Kirk Borne, James F. Bosch, Dominique Boutigny, Craig A. Brackett, et al. LSST: From Science Drivers to Reference Design and Anticipated Data Products. ApJL , 873(2):111, March 2019. doi: 10.3847/1538-4357/ab042c.

[184] Roland Bacon, Vincenzo Maineiri, Sofia Randich, Andrea Cimatti, Jean-Paul Kneib, Jarle Brinchmann, Richard Ellis, Eline Tolstoi, Rodolfo Smiljanic, Vanessa Hill, Richard Anderson, Paula Sanchez Saez, Cyrielle Opitom, Ian Bryson, Philippe Dierickx, Bianca Garilli, Oscar Gonzalez, Roelof de Jong, David Lee, Steffen Mieske, Angel Otarola, Pietro Schipani, Tony Travouillon, Joel Vernet, Julia Bryant, Marc Casali, Matthew Colless, Warrick Couch, Simon Driver, Adriano Fontana, Matthew Lehnert, Laura Magrini, Ben Montet, Luca Pasquini, Martin Roth, Ruben Sanchez-Janssen, Matthias Steinmetz, Laurence Tresse, Christophe Yeche, and Bodo Ziegler. WST – Widefield Spectroscopic Telescope: Motivation, science drivers and top-level requirements for a new dedicated facility. *arXiv e-prints*, art. arXiv:2405.12518, May 2024. doi: 10.48550/arXiv.2405.12518.