Primordial non-Gaussianity without tails - how to measure $f_{\text{NL}}$ with the bulk of the density PDF

Oliver Friedrich$^{1,2}$, Cora Uhlemann$^{3,4}$, Francisco Villaescusa-Navarro$^{5,6}$, Tobias Baldauf$^{3,7}$, Marc Manera$^{3,8}$, Takahiro Nishimichi$^{9,10}$

1 Kavli Institute for Cosmology, University of Cambridge, CB3 0HA Cambridge, United Kingdom
2 Churchill College, University of Cambridge, CB3 0DS Cambridge, United Kingdom
3 Centre for Theoretical Cosmology, DAMTP, University of Cambridge, CB3 0WA Cambridge, United Kingdom
4 Fitzwilliam College, University of Cambridge, CB3 0DG Cambridge, United Kingdom
5 Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton NJ 08544-0010, USA
6 Center for Gravitational Physics, Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan
7 Institute de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain
8 Instituto de Ciencias Físicas de Compostela (ICFC), University of Santiago de Compostela, 15782 Santiago de Compostela, Spain
9 Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, 10010, New York, NY, USA
10 Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, Chiba 277-8583, Japan

ABSTRACT

We investigate the possibility to detect primordial non-Gaussianity by analysing the bulk of the probability distribution function (PDF) of late-time cosmic density fluctuations. For this purpose we devise a new method to predict the impact of general non-Gaussian initial conditions on the late-time density PDF. At redshift $z = 1$ and for a smoothing scale of $30\,\text{Mpc}/h$ our predictions agree with the high-resolution Quijote N-body simulations to $\sim 0.2\%$ precision. This is within cosmic variance of a $\sim 100\,\text{(Gpc}/h)^{3}$ survey volume. When restricting to this $30\,\text{Mpc}/h$ smoothing scale and to mildly non-linear densities ($\delta[30\,\text{Mpc}/h] \in [-0.3,0.4]$) and also marginalizing over potential ignorance of the amplitude of the non-linear power spectrum an analysis of the PDF for such a survey volume can still measure the amplitude of different primordial bispectrum shapes to an accuracy of $\Delta f_{\text{NL}}^\text{equi} = \pm 7.4$, $\Delta f_{\text{NL}}^\text{ortho} = \pm 22.0$, $\Delta f_{\text{NL}}^\text{loc} = \pm 46.0$. When pushing to smaller scales and assuming a joint analysis of the PDF with smoothing radii of $30\,\text{Mpc}/h$ and $15\,\text{Mpc}/h$ ($\delta[15\,\text{Mpc}/h] \in [-0.4,0.5]$) this improves to $\Delta f_{\text{NL}}^\text{equi} = \pm 3.3$, $\Delta f_{\text{NL}}^\text{ortho} = \pm 11.0$, $\Delta f_{\text{NL}}^\text{loc} = \pm 17.0$ - even when marginalizing over the non-linear variances at both scales as two free parameters. Especially, such an analysis could simultaneously measure $f_{\text{NL}}$ and the amplitude and slope of the non-linear power spectrum. However, at $15\,\text{Mpc}/h$ our predictions are only accurate to $\lesssim 0.8\%$ for the considered density range. We discuss how this has to be improved in order to push to these small scales and make full use of upcoming surveys with a PDF-based analysis.

Key words: cosmology: theory – large-scale structure of Universe – inflation – methods: analytical – numerical

1 INTRODUCTION

Data of the large scale structure of the Universe can be successfully analysed on the basis of the 1-point probability distribution function (PDF) of the matter density field - even in the presence of tracer bias and redshift uncertainties. This has been demonstrated e.g. by Gruen et al. (2016); Friedrich et al. (2018); Gruen et al. (2018); Brouwer et al. (2018). In particular, Gruen et al. (2018) and Friedrich et al. (2018) measured the PDF of galaxy density and then used measurements of gravitational lensing to relate that to the PDF of the underlying matter density field quantile-by-quantile. This way they could simultaneously...
ness of matter density fluctuations are related on mildly non-linear scales
b) constrain a two parameter galaxy bias model, that accounts for both linear bias and density dependent shot-noise
c) measure the late-time matter density and the amplitude of late-time density fluctuations as encoded by the parameters $\Omega_m$ and $\sigma_8$ of the $\Lambda$CDM model.

Given the rich amount of information that can be harvested from the PDF (see also Uhlemann et al. 2020), it is time to explore its potential for constraining fundamental physics and to compare it to other cosmological probes. In this paper we showcase one specific application: we study how primordial non-Gaussianity (see e.g. Komatsu & Spergel 2001; Fergusson & Shellard 2009; Scoccimarro et al. 2012; Meerbeg et al. 2019; Biagetti 2019, and references therein) are imprinted in the late time density PDF and how constraints from such an analysis compare to the ones obtained from direct measurements of moments of the density field. The impact of primordial non-Gaussianity on the matter density PDF has previously been discussed e.g. by Valageas (2002b); Uhlemann et al. (2018b). We extend on their results in two ways: First, we present a new method to model the impact of general non-Gaussian initial conditions on the PDF of the late-time density field. This method directly models the cumulant generating function (CGF) of the late-time density field from the CGF of the early-time density field. As shown in Section 4 such an approach requires fewer approximating steps than existing modelling approaches and is close to what would be called modelling 'from first principles'. Secondly, we take into account the full covariance matrix of measured density PDFs across different density contrasts and for two different smoothing scales to determine how well measurements of the density PDF can determine the amplitude of different primordial bispectrum templates. In the context of this task, we also compare the statistical power of the density PDF to that of direct measurements of the cumulants of the density field. The latter have recently been pushed towards applicability in real large scale structure analyses by Gatti et al. (2019) and the impact of primordial non-Gaussianity on higher-order weak lensing statistics has e.g. been investigated by Pace et al. (2011).

In general, scale-dependent tracer bias is believed to be the most promising signature of local primordial non-Gaussianity in the large-scale structure (e.g. Dalal et al. 2008; Desjacques et al. 2009; Jeong & Komatsu 2009; Scoccimarro et al. 2012; Biagetti et al. 2017), especially when combined with cosmic variance cancellation techniques (Seljak 2009). Recently, it was pointed out that a similar scale-dependent bias effect from primordial non-Gaussianity can be observed with voids (Chan et al. 2019), although massive neutrinos produce a similar effect on scales smaller than their maximal comoving free streaming scale (Banerjee & Dalal 2016). Detecting primordial non-Gaussianity in scale-dependent tracer bias requires analyses of clustering power spectra at very large scales. This poses a challenge in terms of cosmic variance, systematic effects (Laurent et al. 2017) as well as modelling of large scale relativistic effects (Bartolo et al. 2011; Camera et al. 2015; Contreras et al. 2019). A way to measure primordial non-Gaussianity that does not suffer from these challenges (but instead from other ones) is to probe the PDF of densities in spheres and their density-dependent clustering on intermediate scales (Uhlemann et al. 2018b; Codis et al. 2018b). In fact, the density PDF is sensitive to all primordial bispectrum shapes and can hence probe equilateral or orthogonal templates for which scale-dependent bias is less pronounced. In addition, density-dependent clustering allows to disentangle local $f_{NL}$ (causing primordial skewness) and $g_{NL}$ (generating primordial kurtosis) by scanning different density environments. Studying the 1-point PDF parallels a number of efforts to understand the cosmic structures beyond their N-point statistics - both for the purpose of detecting primordial non-Gaussianity (e.g. Chiang et al. 2015; Nusser et al. 2018; Karagiannis et al. 2019; Moradinezhad Dizgah et al. 2019) and to test the theory of structure formation in general (e.g. Jain & Van Waerbeke 2000; Simpson et al. 2013; Codis et al. 2016a; Kapczynski et al. 2016; Coulton et al. 2019).

Eventhough there are numerous ways how primordial non-Gaussianity can emerge from inflation, one can categorise them according to the primordial bispectrum shape they generate (see e.g. Babich et al. 2004; Chen et al. 2007; Liguori et al. 2010, for a discussion of concrete models). As suggested by its name, the local shape is typically generated by local interactions, such as in multi-field inflation (Bernardeau & Uzan 2002) or curvaton models (Bartolo et al. 2004), with a small amplitude also being produced in single-field slow-roll inflation (Acquaviva et al. 2003). The equilateral shape requires an amplification of non-linear effects around horizon exit and hence modifications to single-field inflation (Chen et al. 2007). Particular examples are non-canonical kinetic terms as in the Dirac-Born-Infeld model (Alishahiha et al. 2004) or higher-derivative terms such as in K-inflation (Armendariz-Picón et al. 1999), ghost inflation (Arkani-Hamed et al. 2004), effective field theories of inflation (Cheung et al. 2008) or Galileon inflation (Burrage et al. 2011). The orthogonal shape (Senatore et al. 2010) is able to distinguish between variants of non-canonical kinetic terms and higher-derivative interactions.

The late-time matter density PDF at a given smoothing scale is mostly sensitive to the skewness of the primordial density field at that scale and to the running of that skewness around the smoothing scale. As such - unless the PDF is measured on a wide range of smoothing scales - it can only poorly distinguish between different primordial bispectrum shapes. Any model that produces mainly one of the possible bispectrum template can however be successfully tested with PDF measurements. In this paper we consider an analysis of the PDF at redshift $z = 1$ in a survey volume of $V = 100 \ (\text{Gpc}/h)^3$, which is smaller than the effective volume of upcoming surveys such as Spherex with $V_{\text{eff}} \approx 150 \ (\text{Gpc}/h)^3$ and somewhat larger than existing surveys such as BOSS with $V_{\text{eff}} \approx 55 \ (\text{Gpc}/h)^3$ (Doré et al. 2014; Alam et al. 2017). At a smoothing scale of $30 \text{Mpc}/h$ we find our PDF model to agree with the high-resolution run of the Quijote N-body simulations (Villasecusa-Navarro et al. 2019) to $\lesssim 0.2\%$ accuracy over a range of $d [30 \text{Mpc}/h] \in [-0.3, 0.4]$. This is within cosmic variance of the considered volume of $100 \ (\text{Gpc}/h)^3$ (which is also the combined volume of the Quijote high-resolution boxes). Restricting to this smoothing scale and to these mildly non-linear densities we find that a PDF based analysis can measure the amplitude of different primordial bispectrum shapes to an accuracy of $\Delta f_{NL}^{\text{emp}} = \pm 7.4$, $\Delta f_{NL}^{\text{emp}} = \pm 22.0$, $\Delta f_{NL}^{\text{th}} = \pm 46.0$. 

MNras 000, 000-000 (0000)
- even when marginalising over the non-linear variance of the PDF with smoothing radii of 30Mpc/h and 15Mpc/h (δ15Mpc/h ∈ [−0.4, 0.5]) this improves to ∆f_{NL} = ±3.3, ∆f_{NL}^{\text{ortho}} = ±11.0, ∆f_{NL}^{\text{meas}} = ±17.0 - even when marginalising over the non-linear variances at both scales as two free parameters. Especially, such an analysis can simultaneously measure f_{NL} and the amplitude and slope of the non-linear power spectrum. Note that any dependence of these forecasts on σ_8 is completely mitigated by this marginalisation. We do not consider the impact of Ω_m on our signals (see Uhlemann et al. 2020, for an investigation of the general cosmology dependence of the PDF) though Friedrich et al. (2018) and Gruen et al. (2018) have demonstrated that parameters of the ΛCDM model and higher order moments of the density field can be measured simultaneously from what they call lensing-around-cells. Ultimately, we are working towards a combination of a late-time PDF analysis with the early-universe results of Planck Collaboration (2019). These two analyses have the potential to complement each other: the CMB providing information about the background ΛCDM spacetime, the late-time density PDF providing information about non-linear structure growth and both of them containing independent information about the imprint of primordial non-Gaussianities on the large scale structure.

Our paper is outlined as follows: Section 2 summarizes our procedure of modelling the matter density PDF and its moments and compares their statistical power for measuring primordial non-Gaussianity. Section 3 provides intuitive explanations for the impact of primordial non-Gaussianity on the density PDF and its moments, while Section 4 presents a detailed derivation of our actual modelling approach. Section 5 describes the different simulations we used and how we estimate the covariance matrix of the density PDF and moments at the scales under consideration. We conclude and discuss our results in Section 6. Python and C++ tools to carry out the calculations presented in this paper are publicly available\footnote{https://github.com/OliverFHD/CosMomentum}.

## 2 SUMMARY OF PROCEDURE AND FORECAST OF STATISTICAL POWER

### 2.1 Modelling the matter PDF

We start by summarizing the main technical result of this paper. Assume that we know the cumulant generating function (CGF) of the linear density contrast

\[ \varphi_{L,R}(j) \equiv \sum_{n=2}^{\infty} \frac{\langle \delta_{L,R}^{n} \rangle}{n!}, \tag{1} \]

where \( \langle \delta_{L,R}^{n} \rangle \) are the local connected moments (or cumulants) of the linear density contrast field \( \delta_L(x) \) today, averaged over spheres of radius \( R \)

\[ \delta_{L,R} = \int d^3x \, W_R(x) \delta_L(x), \quad W_R(x) = \frac{3\Theta(R - |x|)}{4\pi (R)^3}. \tag{2} \]

For Gaussian initial conditions the linear CGF is simply

\[ \varphi_{L,R}^{\text{Gaus}}(j) = \frac{\langle \delta_{L,R}^{n} \rangle}{2} j^2, \tag{3} \]

while for small primordial non-Gaussianity it can be approximated as

\[ \varphi_{L,R}^{\text{PNG}}(j) \approx \frac{\langle \delta_{L,R}^{n} \rangle}{2} j^2 + \frac{\langle \delta_{L,R}^{3} \rangle}{6} j^3, \tag{4} \]

where the skewness \( \langle \delta_{L,R}^{3} \rangle \) of the linear density contrast can be calculated from the primordial bispectrum, as described in Section 4.4 (see also Uhlemann et al. 2018b).

We then derive in this paper that the cumulant generating function \( \varphi_R(\lambda, z) \) of the non-linear density contrast at redshift \( z \) and smoothing scale \( R \) can be approximated as

\[ \varphi_R(\lambda, z) \approx -s_\lambda(\delta^*, j^*) \tag{5} \]

where \( \delta^* \) and \( j^* \) minimize the function

\[ s_\lambda(\delta, j) = -\lambda F(\delta) + j \delta - \varphi_{L,R}(1 + F(\delta))/\lambda(j). \tag{6} \]

\( F(\delta) = F(\delta, z) \) being the function that describes the spherical collapse of a density fluctuation that has linear density contrast \( \delta_L = \delta \) today (see Appendix A). Hence, we derive an approximation for computing the cumulant generating function of the evolved density field directly from the cumulant generating function of the linear density field.

This result extends the path integral approach of Valageas (2002a) and Valageas (2002b) for Gaussian initial conditions and limited types of primordial non-Gaussianity to general non-Gaussian initial conditions. As for Gaussian initial conditions, the above procedure yields the cumulant generating function at leading order in standard perturbation theory. The accuracy of Equation 5 can be significantly improved with the re-scaling (Bernardeau et al. 2015; Uhlemann et al. 2018a,b,c; Friedrich et al. 2018)

\[ \varphi_R(\lambda, z) = \frac{\sigma_{L,R}^2(z)}{\sigma_{NL,R}^2(z)} \varphi_R^{\lambda_0} \left( \lambda \frac{\sigma_{NL,R}^2(z)}{\sigma_{L,R}^2(z)} z \right), \]

\[ = \sigma_{L,R}^2(z) \sum_{n=2}^{\infty} S_n^{\lambda_0}(z) \sigma_{NL,R}^{2(n-1)}(z) \frac{\Lambda^n}{n!}. \tag{7} \]

Here, \( \varphi_R^{\lambda_0} \) is the leading order CGF from Equation 5, \( S_n \) are the reduced cumulants defined as

\[ S_n \equiv \frac{\langle \delta_{NL,R}^{n} \rangle}{\sigma_{NL,R}^{2(n-1)}}. \tag{8} \]

and \( \sigma_{L,R}^2(z) \) and \( \sigma_{NL,R}^2(z) \) are the variances of the linear and non-linear density field at smoothing radius \( R \). These variances can be calculated from the linear power spectrum \( P_L \) and nonlinear power spectrum \( P_{NL} \) as

\[ \sigma_{L,R}^2(z) = \int \frac{dk}{2\pi^2} P_L(k,z) k^2 W_R^2(k), \tag{9} \]

where \( W_R(k) \) is the Fourier transform of the spherical top-hat kernel from Equation 2, given by (cf. Appendix B)

\[ W_R(k) = 3 \left( \frac{\sin(Rk)}{(kR)^3} - \frac{\cos(Rk)}{(kR)^2} \right). \tag{10} \]

The non-linear power spectrum required for the re-scaling in Equation 7 can be obtained from N-body simulations, from fitting formulae such as halofit (Smith et al. 2003;
Figure 1. Comparing the matter density PDF measured in the Quijote N-body simulations (Villaescusa-Navarro et al. 2019) to our analytic model for Gaussian initial conditions. In this paper we compare cosmological information obtained from the bulk of the PDF (grey area, ≈ 87% of probability) to that obtained from moments of the density field. The latter can strongly depend on the tails of the PDF which are impacted more severely by the non-linear evolution of the density field or baryonic physics. Furthermore, methods to recover the matter density PDF from the galaxy density field (Friedrich et al. 2018; Gruen et al. 2018) require modelling of non-linear tracer bias which is also more difficult in the tails of the PDF.

Figure 2. We test the accuracy of our PDF model for exaggerated amplitudes of primordial bispectrum templates at a smoothing scale of \( R = 15 \text{Mpc}/h \) and redshift \( z = 1 \). Left: difference between PDFs obtained from non-Gaussian and Gaussian initial conditions in the Oriana simulations (points with errorbars, see Scoccimarro et al. 2012; using \( f_{NL}^{loc} = 100 \ f_{NL}^{qu} = -400 \ f_{NL}^{th} = -400 \)). Our model predictions for these differences are displayed by the solid lines. Right: simulations run by Nishimichi (2012); Valageas & Nishimichi (2011) for local primordial non-Gaussianity with \( f_{NL}^{loc} = \pm 100 \). Note that in all simulations the primordial non-Gaussianity also changes the late-time non-linear variance. We absorb this by fitting different values for this variance to each simulation. The errorbars in each panel are for individual simulations, corresponding to a volume of \( 14 \, (\text{Gpc}/h)^3 \) for Oriana and \( 70 \, (\text{Gpc}/h)^3 \) for Nishimichi. This actually overestimates the uncertainty since Gaussian and pNG versions of each simulation have strongly correlated initial conditions. The remaining mismatch between model and simulations may seem small, but it is not negligible wrt. the precision of future surveys. In Section 6 we discuss possible causes of these discrepancies and how to address them in the future.

Takahashi et al. (2012) or from response-function based approaches (e.g. respresso, Nishimichi et al. 2017), see Uhlemann et al. (2020) for a comparison. In this work, we treat the nonlinear variance as a free parameter in order to mitigate potential theoretical uncertainty in the modelling of late-time structure growth. Note that marginalising over the amplitude of non-linear density fluctuations makes our \( f_{NL} \) constraints also independent of \( \sigma_8 \) (the amplitude of the linear density contrast field on an 8Mpc/h smoothing scale; see
also Uhlemann et al. 2020 for a discussion of the dependence of the density PDF on \(\Lambda\)CDM parameters).

Once the cumulant generating function \(\varphi_R(\lambda, z)\) has been calculated, the PDF of \(\delta_R\) can be obtained from an inverse Laplace transform, i.e.

\[
p(\delta_R, z) = \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} e^{-i\lambda\delta_R + \varphi_R(\lambda, z)} .
\]  

(11)

A description of how to efficiently solve this integral is provided in Valageas (2002a); Bernardeau et al. (2015); Friedrich et al. (2018).

In Figure 1 we show a comparison of this model to the density PDF measured in the Quijote N-body simulations (see Villascusa-Navarro et al. 2019, details are also given in our Section 5.1). The figure shows the PDFs at redshift \(z = 1\) and for a smoothing radius of \(R = 15\) Mpc/h. Figure 2 displays the difference between PDFs evolved from Gaussian initial conditions and PDFs evolved from primordial non-Gaussianity with different primordial Bispectrum shapes. The left panel compares our model to simulations run by Scoccimarro et al. (2012), while the right panel uses simulated data based on methods developed in Nishimichi (2012); Valageas & Nishimichi (2011) (see also Uhlemann et al. 2018b or our Section 5.1 for details). It can be seen there that our analytical model realistically captures the shape of the PDF as found in N-body simulations. Note that the approximations made in Section 4 become more precise in the limit of \(f_{NL} \to 0\) and that also the finite resolution of the simulations influences the comparison of Figure 2, as we discuss in Section 6.

Since our model captures the impact of primordial non-Gaussianities on the late-time density PDF realistically, we now discuss the impact of values of \(f_{NL}\) that are compatible with current experimental bounds. In Figure 3 we show the theoretically predicted response of the PDF and its 2nd, 3rd and 4th cumulants (the variance, skewness and kurtosis) to a primordial bispectrum of equilateral shape and with amplitude \(f_{NL}^{\text{equi}} = 47\), corresponding to the 1σ uncertainty of Planck Collaboration (2019)\(^2\). For these figures, the non-linear variance of the late-time density field was calculated with the halofit power spectrum (Smith et al. 2002b) and orthogonal type non-Gaussianity, the PDF is only sensitive to their combination, as we discuss in Appendix D.

\(\text{Figure 3. Comparing the predicted response of the PDF (left) and the cumulants (right) of the density field to changes in } f_{NL}^{\text{eq}} \text{ (the amplitude of an equilateral shape of the primordial Bispectrum) and in the late-time non-linear variance. In the left panels we show the absolute difference between modified and fiducial PDFs, while on the right panels we show relative differences between modified and fiducial cumulants. The value of } f_{NL}^{\text{eq}} = 47 \text{ corresponds to the 1σ uncertainty of Planck Collaboration (2019) (though the latter simultaneously vary both } f_{NL}^{\text{ortho}} \text{ and } f_{NL}^{\text{eq}}, \text{ see our discussion in Appendix D). Error bars assume a survey volume of } V = 100(\text{Gpc}/h)^3. \text{ The grey regions indicate the bulk of the PDFs that is used for the forecasts in this work. In total it excludes about 13% of the probability in the tails for } R = 15\text{ Mpc}/h \text{ and about 5% of the probability for } R = 30\text{ Mpc}/h.\)
Correlation matrix of \((\hat{p}(\delta_{15}), \hat{p}(\delta_{30}), \langle \delta_{n15} \rangle_c, \langle \delta_{n30} \rangle_c)\)

upper right: N-body; lower left: log-normal

-0.75 -0.50 -0.25 0.00 0.25 0.50 0.75 1.00

Figure 4. Correlation matrix of the combined data vector of the PDF and cumulants measured at two smoothing scales. The first block (1-10) is the PDF measured on a 15Mpc/h smoothing scale and for density contrasts \(\delta \in [-0.4, 0.5]\). The second block (11-17) is the PDF measured on a 30Mpc/h smoothing scale and for density contrasts \(\delta \in [-0.3, 0.4]\). The binning of the PDF within these ranges was chosen to match the binning of the fiducial Quijote data products (Villaescusa-Navarro et al. 2019). The last two blocks (18-20 and 21-23) are the variance, skewness and kurtosis measured on smoothing scales of 15Mpc/h and 30Mpc/h respectively. The upper right triangle uses the Quijote N-body simulations. This is the covariance that we are using in our forecasts. To investigate a cheap way of producing covariances for future analyses we also investigate log-normal simulations that are tuned to produce the correct variance and skewness on a 15Mpc/h scale. The correlation matrix obtained from these is shown in the lower left triangle (cf. Section 5.2).

The correlation matrix obtained from these is shown in the lower left triangle (cf. Section 5.2).

2.2 Constraining \(f_{\text{NL}}\) with the PDF and cumulants

Based on our theoretical model and the covariance matrix estimated from the Quijote simulations we investigate how well analyses of the matter density PDF can measure the amplitude of the different primordial bispectrum templates as well as the late-time amplitude of density fluctuations. To do so, we consider the 3 parameters \(\theta = [\text{Var}_{15}, \text{Var}_{30}, f_{\text{NL}}]\), where \(\text{Var}_R\) denotes the non-linear variance of the density contrast at smoothing scale \(R\) and \(z = 1\) and \(f_{\text{NL}}\) denotes the amplitude of different primordial bispectrum templates. Given a model for a data vector, \(x = x(\theta)\), and an expected covariance matrix \(\mathbf{C}\) for that data vector, one can estimate the covariance matrix \(\mathbf{C}^{-1}_{\text{param}}\) of the statistical uncertainties in the parameters as (Krause et al. 2017)

\[
(\mathbf{C}^{-1}_{\text{param}})_{ij} = \frac{dx}{d\theta_i} \cdot \mathbf{C}^{-1} \cdot \frac{dx}{d\theta_j}.
\]

This assumes that the noise in measurements \(\hat{x}\) of \(x\) has a multivariate Gaussian distribution and that the dependence of \(\hat{x}\) on the parameters \(\theta\) is close to linear. In our case, \(x\) is either of the following data vectors:

- The PDF measured for \(\delta_{15\text{Mpc}/h} \in [-0.4, 0.5]\) (\(\approx 87\%\) of probability) or \(\delta_{30\text{Mpc}/h} \in [-0.3, 0.4]\) (\(\approx 95\%\) of probability), see the blue contours in Figure 5.
- Measurements of the first two non-vanishing cumulants (variance and skewness) or the first three non-zero cumulants (variance, skewness and kurtosis) of the density field at these two smoothing scales, see the red and green contours in Figure 5.
- The combined data vector of either the PDF or the cumulants measured at both smoothing scales, see Figure 6.

The above cuts in the PDFs where chosen such that they remove approximately the same amount of probability in both the underdense and overdense tails. The range \(\delta_{15\text{Mpc}/h} \in [-0.4, 0.5]\) is motivated by demanding that our PDF model be in \(\sim 1\%\) agreement with the high-resolution runs of the Quijote simulations. The motivation for choosing the range \(\delta_{30\text{Mpc}/h} \in [-0.3, 0.4]\) was to cut enough of the tail probabilities in order to assume multivariate Gaussian noise on the PDF measurements (see explanations below).

We estimate the covariance of each of these data vectors from the Quijote simulations and choose our binning of the PDFs to match that of Villaescusa-Navarro et al. (2019). Using the ensemble of PDF and cumulant measurements from Quijote we also test that individual data points have a close to Gaussian distribution. If we were to analyse the PDFs over the entire range \(\delta_R \in [-1, \infty]\), then the noise of PDF measurements could not have a multivariate Gaussian distribution, because of the normalization condition \(\int p(\delta_R)d\delta_R = 1\). Also, PDF measurements will always be positive which necessarily slow their distribution. This is especially noticeable in the tails of the PDF, where sampling noise is expected to
Measuring $f_{NL}$ with the bulk of the density PDF

Figure 5 shows the constraints on the amplitude of an equilateral primordial bispectrum, $f_{NL}^{\text{equi}}$, that can be obtained when analysing the PDF at either of the smoothing radii $R = 15\,\text{Mpc}/h$ and $R = 30\,\text{Mpc}/h$ - again assuming a survey volume of $V = (100\,\text{Gpc}/h)^3$ at $z = 1$ (blue solid contours). A comparison of the upper and lower panel of the figure indicates that the smaller scale is more powerful in constraining $f_{NL}$. We want to stress again that our model is only accurate enough to analyse the PDF at the 30Mpc/h smoothing scale. At $R = 15\,\text{Mpc}/h$ the residual $\lesssim 0.8\%$ modelling error is significantly larger than the cosmic variance of the considered survey volume, cf. the discussion in Section 6.1. As explained in that section, our remaining inaccuracy likely result from next-to-leading order corrections derived in Ivanov et al. (2019). These corrections should be included in any real data analysis but we do not expect them to significantly impact the Fisher analysis presented here.

In both panels of Figure 5 we also show the constraints on $f_{NL}^{\text{equi}}$ that can be obtained from direct measurements of the variance and skewness of density fluctuations at these scales (red dash-dotted contours) and from the variance, skewness and kurtosis combined (green dashed contours). In can be seen there, that the PDF indeed contains more information than just the 2nd and 3rd moment of fluctuations combined. The moment-based analysis only catches up with the PDF-based one once the kurtosis is also considered. At the same time, smoothing scale including the kurtosis in the moment based analysis doesn’t add as much information for $R = 30\,\text{Mpc}/h$ as it does for the 15Mpc/h smoothing scale. This may be expected because the density field becomes increasingly Gaussian at larger smoothing scales.

In Figure 6 we show the statistical power achievable with a joint analysis of the PDF at smoothing radii $R = 15\,\text{Mpc}/h$ and $R = 30\,\text{Mpc}/h$ - again assuming a survey volume of $V = (100\,\text{Gpc}/h)^3$ at $z = 1$ (blue solid contours). Also, we show the constraints obtainable from analyses of the variance and skewness (red dash-dotted contour) and the variance, skewness and kurtosis (green dashed contour). Note again that in each of these cases we consider the non-linear variance of density fluctuations at both smoothing scales as two free parameters. This means that the analyses tested here can simultaneously measure primordial non-Gaussianity and the amplitude and slope of the non-linear power spectrum.

In Table 1 we summarize the constraints of the considered analyses for different bispectrum types (Fergusson & Shellard 2009; Scolnicario et al. 2012) - the equilateral template ($f_{NL}^{\text{equi}}$), the orthogonal template ($f_{NL}^{\text{orth}}$) and a bispectrum from local non-Gaussianity ($f_{NL}^{\text{local}}$). For the smoothing scale of $30\,\text{Mpc}/h$ these constraints are compatible with current measurements of $f_{NL}$ from the CMB Planck Collaboration (2019). If our model accuracy can be improved along the lines of Ivanov et al. (2019) to also encompass the $15\,\text{Mpc}/h$ scale then the PDF-based analysis could even significantly improve over the CMB measurements. (As could the moment based analysis. However, we have not rigorously evaluated the accuracy of our moment predictions. Especially, we expect them to suffer more from modelling uncertainties in the high-density tail of the PDF.) We want to stress that the PDF alone can not distinguish between different bispectrum shapes, as we discuss in Appendix D. Hence, we have to consider one bispectrum template at a time in Figure 6.

lead to a Poisson-like rather than a Gaussian noise. Both of these problems are alleviated in our analysis because we only consider the bulk of the PDF. We investigate multivariate statistical behavior of the PDF measurements in Appendix E and find that it is indeed well described by a multivariate Gaussian distribution.
Eventually, we are working towards combining an analysis of the late-time PDF with the early-universe results of Planck Collaboration (2019). In such a combined analysis, the CMB would provide information about the background ΛCDM parameters, which change the linear skewness and ΛCDM parameters like Ωm and ns which drive the scale-dependence of the late-time PDF. We explain the details of these results in Section 4 and Section 5.

3 GAINING INTUITION FOR THE IMPACT OF PRIMORDIAL NON-GAUSSIANITY

Before presenting the details of our full modelling approach in Section 4 we want to provide approximations that make the imprints of non-Gaussian initial conditions, encoded in that modelling approach, more transparent.

3.1 Impact of primordial skewness on cumulants

For Gaussian initial conditions, the reduced cumulants \( \langle s^3 \rangle / \langle s^2 \rangle^{3/2} \) of the unsmoothed density contrast are given by simple constants and a smoothing scale \( R \) induces only a mild running due to the changing slope in the linear variance (Peebles 1980; Bernardeau 1994). At leading order this gives

\[
S_3^G = \frac{34}{7} + \gamma, \quad \gamma = \frac{d \log \sigma_L^2(R)}{d \log R},
\]

\[
S_4^G = \frac{60712}{1323} + \frac{62}{3} + \frac{\gamma}{3} - \frac{7}{3} \gamma^2,
\]

for the 3rd and 4th order reduced cumulants. For primordial non-Gaussianity, analytical predictions for the reduced skewness without smoothing by Fry & Scherrer (1994) demonstrate that the Gaussian result is offset by terms depending on the initial 3-point correlation function (and higher-point correlations that we neglect here). Chodorowski & Bouchet (1996) generalised this result to the unsmoothed kurtosis, which is also coupled to the initial skewness by non-linear evolution. Gaztanaga & Fosalba (1998) used spherical collapse to relate the reduced cumulants \( S_N^G \) in the presence of primordial non-Gaussianity to their Gaussian counterparts from equation (13). Denoting the 3rd central moment of the linear density field as \( \kappa_L^3 \), they obtain

\[
S_3^{NG} = S_3^G + \frac{1}{\sigma_L^2} \kappa_L^3 - 2 \left( \frac{S_3^G}{3} - 1 \right) \left( \frac{\kappa_L^3}{\sigma_L^2} \right)^2
\]

\[
S_4^{NG} = S_4^G + 4 \frac{S_3^G}{\sigma_L^2} \kappa_L^3 + \left( 3 + 7 S_3^G - \frac{14}{3} \left( S_3^G \right)^2 + \frac{3}{2} S_4^G \right) \frac{\kappa_L^3}{\sigma_L^2}.
\]

The most important contribution to \( S_3 \) is given in terms of the linearly evolved reduced skewness (induced by primordial non-Gaussianity), \( \kappa_L^3/\sigma_L^2 \), which decays linearly with the growth function, because \( \delta_L(z) \propto D(z) \). On the other hand, nonlinear evolution leads to almost constant reduced cumulants \( S_N^G \), such that the primordial skewness offsets the nonlinear reduced skewness and kurtosis by an amount that is inversely proportional to the growth function. Note that this signature is qualitatively different from the impact of ΛCDM parameters, which change the linear variance in Equation 13, but leave the hierarchical ratios \( S_N \) close to constant in time (for a discussion, see Section 2.1 in Uhlemann et al. 2020). In a multi-redshift analysis, we expect that this property would allow to jointly constrain \( f_{NL} \), as amplitude of the linear skewness, and ΛCDM parameters like \( \Omega_m \) and \( n_s \) which drive the scale-dependence of the linear variance.

Let us now quantitatively estimate the effect of different primordial bispectrum shapes discussed in Section 4.4. The amplitudes of the initial rescaled skewness for a radius of \( R = 15 \text{ Mpc}/h \) for different templates are

\[
\frac{\kappa_L^3}{\sigma_L^2} (15 \text{ Mpc}/h) \simeq \left( 28 S_3^L, 8.5 S_3^F, -5 S_3^\text{ortho} \right) \times 10^{-5}.
\]

The predicted ratios of reduced cumulants from Equations 13-14 agree well with the measurements in the Oriana simulations for radii \( R = 20 - 40 \text{ Mpc}/h \) at redshift \( z = 0.34 \) (see Figure 1 in Mao et al. 2014). Since the nonlinear variance depends very weakly on the amplitude of primordial non-Gaussianity \( f_{NL} \) (Uhlemann et al. 2018b), the ratio of connected moments is close to the ratio of reduced cumulants from Equation 44. When considering the equilateral model with \( f_{NL} = 47 \), we obtain a 0.4% increase for the skewness and a 1% increase in the kurtosis, in good agreement with the result from the full shape of the matter PDF shown in the upper right panel of Figure 3.

3.2 Heuristic approximation for the matter PDF

While we use the recipe described in Section 4.5 to model the density PDF, one can gain intuition for its sensitivity to primordial non-Gaussianity from a simplistic saddle-point approximation for the PDF from Equation 11 (see Uhlemann et al. 2018b). According to the large-deviation principle (Bernardeau & Reimberg 2016), the exponential decay of the late-time density PDF with increasing density contrast, \( \psi(\delta_L) \), can be predicted from the cumulant generating function and the reduced cumulants entering in Equation 7 are determined by spherical collapse. The matter PDF can then
Measuring $f_{NL}$ with the bulk of the density PDF

V = 100(Gpc/h)$^3$, using $R = 15$ & 30 Mpc/h : 

- 2nd & 3rd moments
- 2nd - 4th moments
- PDF, $\delta_{15} \in [-0.4, 0.5]$ and $\delta_{30} \in [-0.3, 0.4]$

Figure 6. Forecast for measurements of $f_{equi}$ (the amplitude of a primordial Bispectrum with equilateral shape, Fergusson & Shellard 2009) and the late-time variance of the density field when analysing PDF or cumulants at both $R = 15$ Mpc/h and $R = 30$ Mpc/h. Note that we are able to simultaneously measure the amplitude of the primordial Bispectrum and both the amplitude and slope of the late-time power spectrum. Marginalisation over the late-time variances also makes our constraints on primordial non-Gaussianity insensitive to $\sigma_8$. Ultimately we are preparing a joint analysis of the late-time density PDF and early-universe constraints from Planck Collaboration (2019), which is why we keep all other cosmological parameters fixed in this forecast (see also Uhlemann et al. 2020, for an investigation of the general cosmology dependence of the PDF).

be expressed in terms of this decay-rate function as

$$p(\delta_R, z) \propto \sqrt{\psi_{R,z}'(\delta_R) + \frac{\psi_{R,z}''(\delta_R)}{1 + \delta_R} \exp[-\psi_{R,z}(\delta_R)]},$$

(16)

where $\delta_L(\delta_R)$ is the mapping between linear and nonlinear density contrasts, $R_L(R, \delta_R) = R(1 + \delta_R)^{1/3}$ accounts for mass conservation, and $\kappa_3^L/\sigma_3^L \propto f_{NL}$ is the linear rescaled skewness caused by the presence of a primordial bispectrum. This approximation reflects that the leading-order corrections to the reduced cumulants are controlled by the primordial skewness, as we have seen explicitly for the skewness and kurtosis in Equations 14. The exponent dominates the behaviour in the tails, where a positive primordial skewness leads to an enhancement of high densities and a suppression of low densities. Around the peak of the PDF, the prefactor becomes relevant and leads to an enhancement in the PDF for moderately underdense spheres and a reduction for moderately overdense spheres, in agreement with the left panels of Figure 3.

4 SADDLE POINT METHOD FOR GENERAL NON-GAUSSIAN INITIAL CONDITIONS

We now present a derivation of our main technical result - the approximation in Equation 5 for the cumulant generating function (CGF) of the late time density contrast $\delta_R$ for arbitrary non-Gaussian initial conditions and our model for the PDF $p(\delta_R)$ that follows from it.
4.1 Path integral approach for the cumulant generating function

The cumulant generating function \( \varphi_R(\lambda, z) \) of the non-linear density contrast \( \delta_R \) is defined as

\[
e^{\varphi_R(\lambda, z)} = \exp \left( \sum_{n=1}^{\infty} \frac{\langle \delta_R(x,z)^n \rangle}{n!} \lambda^n \right)
\]

\[
\equiv \langle e^{\lambda \delta_R} \rangle = \int d\delta_R \ p(\delta_R|x,z) \ e^{\lambda \delta_R} . \tag{17}
\]

Here \( \langle \delta_R(x,z)^n \rangle \), refers to the \( n \)th connected moment (or cumulant) of the smoothed nonlinear density contrast \( \delta_R(x,z) \) and the second line is the cumulant expansion theorem (Bernardeau et al. 2002). Because of homogeneity and isotropy we will only consider \( \delta_R(x,z) \) and hence drop the label \( x \). For simplicity, we will also suppress the dependence of \( \delta_R \) on redshift \( z \).

In a procedure similar to that of Valageas (2002a), we can write the above expectation value as a functional integral over all possible configurations of today’s linear density contrast,

\[
e^{\varphi_R(\lambda)} = \int \mathcal{D}\delta_L \ \mathcal{P}[\delta_L] e^{\lambda \mathcal{R}[\delta_L]}, \tag{18}
\]

where \( \mathcal{P}[\delta_L] \) is the probability density functional of the linear density contrast field \( \delta_L \) and the non-linear density contrast \( \delta_R \) has been expressed as a functional of \( \delta_L \). (Note again that we dropped the dependence on \( z \) from our notation - this dependence is entirely carried by the functional \( \delta_R[\cdot] = \delta_R[\cdot,z] \) since we always consider \( \delta_L \) at \( z = 0 \).

For Gaussian initial conditions \( \delta_L \) will be a Gaussian random field and the probability density functional \( \mathcal{P}[\delta_L] \) can be directly expressed through the linear power spectrum \( P_L(k) \). This was done by Valageas (2002a) to derive an approximation for the late-time CGF from non-Gaussian initial conditions. Valageas (2002b) also studied a limited set of non-Gaussian initial conditions for which explicit expressions of \( \mathcal{P}[\delta_L] \) are available.

Here we extend these studies to general non-Gaussian initial conditions. Defining the linear cumulant generating functional associated with \( \mathcal{P}[\delta_L] \) as

\[
\Phi[J_L] = \sum_{n=1}^{\infty} \frac{1}{n!} \int \prod_{i=1}^{n} \mathrm{d}^3 x_i \ J_L(x_i) \ \xi_{L,n}(x_1, \ldots, x_n) \tag{19}
\]

we can express \( \varphi_R \) as

\[
e^{\varphi_R(\lambda)} = \frac{1}{N} \int \mathcal{D}\delta_L \ \mathcal{D}J_L \ e^{\lambda \mathcal{R}[\delta_L] - i J_L \cdot \Phi[i J_L]} . \tag{20}
\]

The linear cumulant generating functional \( \Phi[J_L] \) encodes the initial conditions. For Gaussian initial fluctuations it can be entirely expressed through the linear two-point correlation function and our results would reduce to those of Valageas (2002a). For the primordial non-Gaussianity models we consider here, it will additionally depend on the primordial three-point function \( \xi_{L,3} \) or the corresponding bispectrum, for which we discuss concrete templates in Section 4.4. Comparison with Equation 11 shows that the normalisation constant \( \mathcal{N} \) is formally given by \( |2\pi| \), i.e. the determinant of \( 2\pi \) times the unit operator in the space of \( J_L \)'s and hence infinite. But this normalisation will eventually drop in our calculations. Note also that we introduced the abbreviation

\[
J_L \cdot \delta_L \equiv \int \mathrm{d}^3 x \ J_L(x) \ \delta_L(x) . \tag{21}
\]

We will calculate the above functional integral with Laplace’s method, i.e. by approximating the integrand around its maximum with a Gaussian functional, which is also called steepest descent or saddle-point method (e.g. Valageas 2002a; Bernardeau et al. 2015; Uhlemann et al. 2018a). This means we define an action \( S_\lambda \) as

\[
S_\lambda[\delta_L, J_L] = -\lambda \delta_R[\delta_L] + i J_L \cdot \delta_L - \Phi[i J_L] \tag{22}
\]

and find the saddle point configurations \( \delta_L^* \) and \( J_L^* \) that minimise this action in order to approximate

\[
e^{\varphi_R(\lambda)} = \frac{1}{N} \int \mathcal{D}\delta_L \ \mathcal{D}J_L \ e^{-S_\lambda[\delta_L, J_L]}
\]

\[
\approx \frac{1}{A^{1/2}} e^{-S_\lambda[\delta_L^*, J_L^*]} . \tag{23}
\]

Here \( A \) is the determinant of the Hessian matrix of the functional \( S_\lambda \) (now considered as a matrix in the combined space of \( \delta_L^* \)’s and \( J_L^* \)’s) evaluated at the saddle point configurations \( \delta_L^* \) and \( J_L^* \). The cumulant generating function of the late-time smoothed density contrast is then given by

\[
\varphi_R(\lambda) \approx -S_\lambda[\delta_L^*, J_L^*] - \frac{1}{2} \ln A . \tag{24}
\]

The second term in this approximation is a next-to-leading order correction to the first term (for Gaussian initial conditions it is equivalent to 1-loop corrections to the CGF, cf. Valageas 2002c; Ivanov et al. 2019) and will be neglected in this paper. Instead, we apply a re-scaling of the cumulant generating function that accounts for strong late-time non-linearity by leaving the variance of non-linear density fluctuations as a free parameter. We describe this re-scaling in Section 4.3 after deriving an explicit expression for the first, leading order term of Equation 24 in the following section. We would like to stress here: omitting the second term in Equation 24 and using the re-scaling of Section 4.3 instead is the biggest conceptual weakness of our analysis. We discuss implications of this for the interpretation of our results and future work in Section 6.1.

4.2 Minimizing the action \( S_\lambda \)

In the following, let \( dF/df(x) \) denote functional derivation of a functional \( F \) wrt. the function \( f \). To calculate the leading order term in Equation 24 we have to find configurations \( \delta_L^* \) and \( J_L^* \) such that

\[
dS_\lambda \bigg|_{\delta_L^*, J_L^*} = 0 = \left. dS_\lambda \right|_{\delta_L^*, J_L^*} \tag{25}
\]

\[
\Rightarrow i J_L^*(x) = \lambda \left. \frac{d\delta_R}{dJ_L(x)} \right|_{\delta_L^*} \tag{26}
\]

\[
\delta_L^*(x) = \left. \frac{d\Phi}{dJ_L(x)} \right|_{i J_L^*} . \tag{27}
\]

Because we allow for arbitrary shapes of primordial \( N \)-point functions in the definition of \( \Phi \), it is not obvious that these equations have spherically symmetric solutions. We nevertheless show that a spherically symmetric ansatz \( \delta_L^*(x) = \delta_L^*(x), J_L^*(x) = J_L^*(x) \) solves Equations 26 and 27.
and argue in the end that the spherically symmetric solution is indeed a global minimum of the action - at least for small deviations from primordial Gaussianity.

With a spherically symmetric ansatz Equation 26 can be solved (along the lines of Valageas 2002a) by noting that

$$\delta_R(\delta^*_R) = F(\delta^*_R, L, R_L)$$

(28)

where the function $F$ describes spherical collapse of the density fluctuation (cf. Appendix A) from the initial, linear radius

$$R_L = R (1 + \delta_R(\delta^*_R))^{1/3}$$

(29)

to the final radius $R$. Hence one can see that

$$\frac{d\delta_R}{d\delta_R(\delta^*_R)} \bigg|_{\delta^*_R} = F'(\delta^*_R, L, R_L)$$

$$= F'(\delta^*_R) \left( W_{RL}(x) + \frac{\delta L_R}{d\delta_R(x)} \bigg|_{RL} \frac{\delta R}{\delta \lambda} \right)$$

$$= F'(\delta^*_R) \left( W_{RL}(x) + \frac{\delta L_R}{d\delta_R(x)} \bigg|_{RL} \frac{R^3}{3 \delta \lambda x} \right) \bigg|_{\delta^*_R}$$

(30)

$$\Rightarrow \lambda \frac{d\delta_R}{d\delta_R(\delta^*_R)} \bigg|_{\delta^*_R} = \lambda \frac{F'(\delta^*_R, L, R_L) W_{RL}(x)}{1 - F'(\delta^*_R) \frac{\delta L_R}{d\delta_R(x)} \bigg|_{RL} \frac{R^3}{3 \delta \lambda x}}$$

$$\Rightarrow i J^L_L(x) = A_L(\delta^*_R) W_{RL}(x)$$

(31)

(32)

Inserting Equation 32 into 27 we also get a simplified expression for $\delta^*_R$ as

$$\delta^*_R(x) = \frac{d \Phi}{d J^L_L(x)} \bigg|_{i J^L_L}$$

$$= \sum_{n=2}^{\infty} \frac{1}{(n-1)!} \left( \prod_{i=1}^{n} (\delta^*_R, L, R_L) \frac{\delta L_R}{d\delta_R(x)} \bigg|_{RL} \right)$$

$$\Rightarrow \delta^*_R(x) = \sum_{n=2}^{\infty} \frac{A_n}{(n-1)!} \left( \delta^*_R, L, R_L \right)$$

(33)

From this we can further deduce the useful relations

$$\delta^*_R(x) = \sum_{n=2}^{\infty} \frac{A_n}{(n-1)!} \left( \delta^*_R, L, R_L \right)$$

$$\Rightarrow \delta^*_R(x) = \sum_{n=2}^{\infty} \frac{A_n}{(n-1)!} \left( \delta^*_R, L, R_L \right)$$

(34)

$$\Rightarrow \delta^*_R(x) = \sum_{n=2}^{\infty} \frac{A_n}{(n-1)!} \left( \delta^*_R, L, R_L \right)$$

(35)

$$\Rightarrow \delta^*_R(x) = \sum_{n=2}^{\infty} \frac{A_n}{(n-1)!} \left( \delta^*_R, L, R_L \right)$$

(36)

where we have defined the cumulant generating function of the linear, spherically averaged density contrast $\delta_{L,R}$ as $\varphi_L,R$. Equations 32, 35 and 36 allow us to obtain $S_A(\delta^*_R, J^L_L)$ by solving a simple 2-dimensional optimization problem. To see this, we first use 32 in 22 to get

$$S_A(\delta^*_R, J^L_L) = -\lambda F(\delta^*_R, R_L) + A_L \delta^*_R L, R_L - \varphi^L, R_L (A_L) \bigg|_{RL}$$

(37)

This is in fact equal to the minimum of the 2-dimensional function

$$s_A(\delta, j) = -\lambda F(\delta) + j \delta - \varphi^L, R_L(1 + F(\delta))^{1/3}(j)$$

(38)

since solving

$$\frac{\delta s_A}{\delta j} \bigg|_{\delta^*_R, J^L_L} = 0 = \frac{\delta s_A}{\delta j} \bigg|_{\delta^*_R, J^L_L}$$

(39)

leads to

$$j^* = -\lambda F'(\delta^*_R) + \frac{d \varphi^L, R_L(1 + F(\delta))}{d F(R_L)} \bigg|_{R_L}$$

(40)

$$\delta^*_R = \frac{d \varphi^L, R_L(1 + F(\delta))}{d F(R_L)} \bigg|_{R_L}$$

(41)

which - as can be seen from 35 and 36 - has the solutions

$$\delta^*_R = \delta^*_R, L, j^* = A_L$$

(42)

We have shown that a spherically symmetric ansatz for the configurations $\delta^*_R$ and $J^L_L$ can extremize the action $S_A$ and that the corresponding extreme values $S_A(\delta^*_R, J^L_L)$ can be obtained by a simple 2-dimensional optimisation problem. We have not yet shown that these $\delta^*_R$ and $J^L_L$ are indeed global minima of the action. We however note that Valageas (2002a) has shown that for Gaussian initial conditions there exists a range of values of $\lambda$ for which the spherically symmetric saddle point is a global minimum. Especially, the Hessian of the action $S_A$ is positive definite for these values of $\lambda$. Small deviations from Gaussian initial conditions will change the location of this saddle point, but because of continuity there will be a range of deviations for which the Hessian at this point is still positive definite, such that it is still a minimum of the action. In fact, because of continuity this will also stay a global minimum for sufficiently small deviations from Gaussian initial conditions.

Note that for large enough $\lambda$ the action $S_A(\delta^*_R, J^L_L)$ can indeed have two extrema - even for Gaussian initial conditions (Valageas 2002a; Bernard et al. 2014). This leads to a 2-branch structure in the cumulant generating function where the 2nd branch governs the extreme rare-events tail of the PDF. This is an additional reason for why we avoid the high-density tail here.

4.3 Rescaling to the nonlinear variance

It has been shown (Fosalba & Gaztanaga 1998; Valageas 2002a,c) that the saddle-point approximation gives the cumulants of the non-linear density contrast at leading order in perturbation theory. This especially means that the above approximation for the CGF yields the wrong variance of density fluctuations even in the mildly non-linear regime. It is hence much better to consider the re-scaled cumulant generating function (Bernardeau & Reimberg 2016)

$$\varphi_R(\lambda) \equiv \frac{\lambda^2}{\sigma^2} \varphi_R \left( \frac{\lambda^2}{\sigma^2} \right) = \sum_n S_n \frac{\lambda^n}{n!}$$

(43)
where $\sigma_R^2$ is the variance of density fluctuations and

$$S_n \equiv \frac{(\delta^2 R)_{\text{c}}}{\sigma_R^{2(n-1)}} .$$

The reduced cumulants (or hierarchical coefficients) $S_n$ are significantly less sensitive to non-linear evolution than the raw cumulants (Bernardeau et al. 2002, 2014, 2015). This is because tidal terms that are not captured by the leading order term in Equation 24 are largely erased by smoothing effects in the reduced cumulants, both for Gaussian and non-Gaussian initial conditions (Fosalba & Gaztanaga 1998; Gaztanaga & Fosalba 1998).

Hence, our modelling strategy in this paper is to compute the rescaled CGF at leading order in perturbation theory by calculating the first term in Equation 24 and using the linear variance $\sigma_R^2$ in Equation 43 to compute the leading order, re-scaled cumulant generating function $\tilde{\phi}_R(\lambda)$. The final non-linear CGF $\phi_R(\lambda)$ is then obtained by inverting Equation 43 with $\sigma_R^2$ as a free parameter, i.e.

$$\phi_R(\lambda) = \frac{1}{\sigma_R^2} \tilde{\phi}_R(\sigma_R^2 \lambda) .$$

As can be seen in Figure 1, this procedure reproduces not only the variance but also the overall shape of the linear density PDF observed in N-body simulations. This re-scaling to the non-linear variance has also been successfully applied to the non-linear variance itself is also affected by primordial non-Gaussianity through mode-coupling. In particular, it generates an additional odd-order term in the 1-loop power spectrum,

$$P(k, z) = D^3(z) P_L(k) + P^{(12)}(k, z) + [P^{(22)} + P^{(13)}](k, z) ,$$

which is absent for Gaussian initial conditions and reads (Taruya et al. 2008)

$$P^{(12)}(k, z) = 2D^3(z) \int \frac{d^3q}{(2\pi)^3} F^{(2)}_{\text{sym}}(q, k-q) B_0(-k, q, k-q) ,$$

where $F^{(2)}_{\text{sym}}$ is the symmetrized kernel for the second order perturbative solution of the fluid equations (see e.g. equation (45) in Bernardaeu et al. 2002). By marginalising over the non-linear variance we are hence ignoring part of the information about primordial non-Gaussianity that is contained in the density PDF. At the smoothing scales considered here this effect is largest for the orthogonal bispectrum template.

### 4.4 Application: primordial bispectrum shapes

If the initial density field was not drawn from a Gaussian distribution, then the linear density contrast today will have a non-zero 3-point function

$$\xi_{3,L}(x_1, x_2, x_3) = \langle \delta_L(x_1) \delta_L(x_2) \delta_L(x_3) \rangle .$$

The Bispectrum $B_3(k_1, k_2, k_3)$ is defined through the Fourier transform of $\xi_{3,L}$ in each of its arguments (see Appendix B for our Fourier conventions) as

$$\langle \tilde{\delta}_L(k_1) \tilde{\delta}_L(k_2) \tilde{\delta}_L(k_3) \rangle \equiv (2\pi)^3 \delta_D(k_1 + k_2 + k_3) B_3(k_1, k_2, k_3) .$$

The skewness of the linear density field averaged over a spherical top-hat filter of radius $R$ is then given in either real space or Fourier space by

$$\langle \tilde{\delta}_L(k, R) \rangle = \frac{3}{2} \int \frac{d^3k}{(2\pi)^3} \tilde{W}_R(k) \tilde{\delta}_L(k) ,$$

where $W_R(x)$ and $\tilde{W}_R(x)$ are given in Equation 2 and Equation 10 respectively. The Fourier space expression can be simplified to

$$\langle \tilde{\delta}_L(k, R) \rangle = \frac{3}{2} \int \frac{d^3k}{(2\pi)^3} \tilde{W}_R(k) \tilde{\delta}_L(k) B_3(k_1, k_2, k_3) ,$$

where in the last line we set $k_3 \equiv |k_1 + k_2| = \sqrt{k_1^2 + k_2^2 + 2\mu_k k_1 k_2}$ and $k_1 \cdot k_2 \equiv \mu_k k_1 k_2$. The above integral only depends on the angle between the two remaining wave vectors. Hence we can perform the angular integral for one of these to get

$$\langle \tilde{\delta}_L(k) \rangle = \frac{2}{(2\pi)^2} \int d\ln k_1 k_1^2 \tilde{W}_R(k_1) \int d\ln k_2 k_2^2 \tilde{W}_R(k_2) \times$$

$$\times \int_{|k_1 - k_2|}^{k_1 + k_2} d\ln k_3 k_3^2 \tilde{W}_R(k_3) B_3(k_1, k_2, k_3) .$$

We would now like to express $B_3(k_1, k_2, k_3)$ in terms of a Bispectrum of primordial potential fluctuations, $B_\phi(k_1, k_2, k_3)$. The initial potential fluctuations $\tilde{\phi}_c(k)$ are related to the linear density contrast today via the Poisson equation

$$\tilde{\phi}_c(k) = -2\tilde{T}(k) k^2 D(z) \alpha_z(\zeta) \tilde{\phi}_{\text{c}}(k) ,$$

where $\tilde{T}(k)$ is the transfer function (defined such that $\tilde{T}(k \to 0) = 1$, $D(z)$ and $\alpha(z)$ are the linear growth factor and scale factor (with $D = 1 = a$ today) and $\zeta$, is some redshift during matter domination. Now common templates for $B_\phi$ contain terms of the form

$$B_\phi(k_1, k_2, k_3) \sim P_\phi(k_1)^{2\gamma_1} P_\phi(k_2)^{2\gamma_2} P_\phi(k_3)^{2\gamma_3}$$

MNRAS 000, 000–000 (0000)
such that \(\alpha_1 + \alpha_2 + \alpha_3 = 2\) (Fergusson & Shellard 2009; Scocciarino et al. 2012; Uhlmann et al. 2018c). In terms of the linear matter power spectrum today these terms read

\[
P_c(k_1) \sim P_c(k_2) \sim P_L(k_3) \times \left(\frac{3\Omega_m H_0^2}{2} \frac{a(z_1)}{D(z_1)}\right)^4 P_L(k_1)^{\alpha_1} \frac{P_L(k_2)^{\alpha_2}}{P_L(k_3)^{\alpha_3}} \times T(k_1)^{-2\alpha_1} T(k_2)^{-2\alpha_2} T(k_3)^{-2\alpha_3} k_1^{-4\alpha_1} k_2^{-4\alpha_2} k_3^{-4\alpha_3}.
\]

Hence, the linear density Bispectrum will contain terms like

\[
B_L(k_1, k_2, k_3) \geq \frac{-3\Omega_m H_0^2 a(z_1)}{2D(z_1)} P_L(k_1)^{\alpha_1} P_L(k_2)^{\alpha_2} P_L(k_3)^{\alpha_3} \times T(k_1)^{-2\alpha_1} T(k_2)^{-2\alpha_2} T(k_3)^{-2\alpha_3} k_1^{-4\alpha_1} k_2^{-4\alpha_2} k_3^{-4\alpha_3}.
\]

Inserting this into our above equations for the skewness we arrive at

\[
\langle \delta_{L,R}^3 \rangle \geq \frac{-3\Omega_m H_0^2 a(z_1)}{(2\pi)^3} \frac{D(z_1)}{D(z_1)} \times \int d\ln k_1 k_1^{-4\alpha_1} T(k_1)^{-2\alpha_1} \tilde{W}_R(k_1) P_L(k_1)^{\alpha_1} \times \int d\ln k_2 k_2^{-4\alpha_2} T(k_2)^{-2\alpha_2} \tilde{W}_R(k_2) P_L(k_2)^{\alpha_2} \times \int_{\min(k_1,k_2)}^{\max(k_1,k_2)} d\ln k_3 k_3^{-4\alpha_3} T(k_3)^{-2\alpha_3} \tilde{W}_R(k_3) P_L(k_3)^{\alpha_3}.
\]

In order to calculate the skewness of the linear density field as a function of \(R\) we now only have to specify what terms of the form of (54) are present in the primordial bispectrum. Following e.g. Uhlmann et al. (2018c) (see also the other references given in this section) we consider local, orthogonal and equiangular bases for the shape which are given by

\[
B^{\text{loc}}_L(k_1, k_2, k_3) = -2\Omega_{NL}^L (P_\phi(k_1) P_\phi(k_2) + 2 \text{ permutations})
\]

\[
B^{\text{uni}}_L(k_1, k_2, k_3) = 6\Omega_{NL}^L \left[ (P_\phi(k_1) P_\phi(k_2) + 2 \text{ perm.}) + 2 P_\phi^{2/3}(k_1) P_\phi^{2/3}(k_2) P_\phi^{2/3}(k_3) - (P_\phi(k_1) P_\phi^{2/3}(k_2) P_\phi^{2/3}(k_3) + 5 \text{ perm.}) \right]
\]

\[
B^{\text{ortho}}_L(k_1, k_2, k_3) = 6\Omega_{NL}^L \left[ 3(P_\phi(k_1) P_\phi(k_2) + 2 \text{ perm.}) + 8 P_\phi^{2/3}(k_1) P_\phi^{2/3}(k_2) P_\phi^{2/3}(k_3) - 3(P_\phi(k_1) P_\phi^{2/3}(k_2) P_\phi^{2/3}(k_3) + 5 \text{ perm.}) \right].
\]

Note that these expression differ in sign from equations given in other publications (e.g. Mao et al. 2014; Uhlmann et al. 2018b). This is because here \(\phi\) denotes the Newtonian potential and not Bardeen’s potential (Salopek & Bond 1990).

4.5 Summary of our recipe for the matter PDF

To summarize, the model predictions in Figures 1-3 and the forecasts presented in Figures 6-5 are calculated as follows:

- computing skewness of the linear density contrast \(\langle \delta_{L,R}^3 \rangle_c \) from Equation 52, using either of the primordial bispectrum templates given in Equations 58-60
- approximating the cumulant generating function of the linear density field in terms of the linear variance and skewness as
- using this in Equation 38 to calculate the late-time cumulant generating function (applying also the variance rescaling described in Section 4.3)
- numerically evaluating the inverse Laplace transform of Equation 11 to obtain the late-time density PDF.

4.6 Numerical calculation of the CGF and extraction of individual cumulants

In Section 4.2 we demonstrated that the leading order approximation of the cumulant generating function

\[
\varphi_R(\lambda) = \sum_{n=2}^{\infty} \langle \delta_R^n \rangle_c \frac{\lambda^n}{n!}
\]

is given by the minimum of the 2-dimensional function

\[
s_\lambda(\delta, j) = -\lambda \mathcal{F}(\delta) + j \delta - \varphi_{L,R}(1+\mathcal{F}(\delta)^{1/3})^3(j)\]

Minimizing this function amounts to solving the equations

\[
\lambda = \mathcal{F}^{-1}(\delta^*) + \frac{d\varphi_R(j^*)}{dj}
\]

\[
\delta^* = \frac{d\varphi_{L,R}(1+\mathcal{F}(\delta^*)^{1/3})^3(j^*)}{dj}.
\]
However, this is not what we do in practice. Instead, it is much easier to proceed as follows:

(i) For an array of values for $\delta^*$, calculate the corresponding arrays $F(\delta^*)$ and $F^\prime(\delta^*)$ using Appendix A. Also compute the variance and skewness of the linear density field at the radii $R_L(\delta^*) = (R + F(\delta^*))^{1/3}$.

(ii) For each value of $\delta^*$ invert Equation 65 to obtain $J^* = J^*(\delta^*)$. Since we approximate $\varphi_{L,R}(\lambda)$ as a cubic function in Equation 66, this just corresponds to solving a quadratic equation.

(iii) Now invert Equation 64 to obtain $\lambda = \lambda(\delta^*)$.

(iv) Finally, obtain the CGF from $\varphi_R(\lambda) = s(\delta^*, J^*)$.

This is a constructive procedure to obtain the cumulant generating function which does not actually require one to solve any optimization problem.

The above steps yield the cumulant generating function $\varphi_R(\lambda)$ on a grid of values of $\lambda$. In order to extract individual cumulants from that (i.e. the Taylor coefficients in Equation 62) one might attempt to simply fit a polynomial of finite degree in $\lambda$ to the CGF $\varphi_R(\lambda)$. This is however highly unstable, mainly because positive values of $\delta^*$ are mapped to a very small interval of $\lambda(\delta^*)$ compared to the interval that negative values of $\delta^*$ are mapped to (see e.g. Figure 3 of Valageas 2002a). A more robust way to extract individual moments is to first define the new variable (Bernaldeau & Valageas 2000; Valageas 2002a; Bernardeau et al. 2015)

$$\tau(\delta^*) = \frac{\delta^*}{\langle (d_L R_L(\delta^*))^3 \rangle} \cdot (66)$$

Then one proceeds as follows:

(i) Fit a polynomial of finite order $N$ in $\tau$ to both $\lambda(\tau)$ and $\varphi_R(\lambda(\tau))$.

(ii) Determine the $N \times N$ Bell-Jabotinsky matrices (Jabotinsky 1963) of $\lambda(\tau)$ and $\varphi_R(\lambda(\tau))$ wrt. $\tau$.

$$\left( B^{\lambda(\tau)} \right)_{k,l} = \frac{1}{k!} \frac{d^k \lambda}{d \tau^k} \bigg|_{\tau = 0} ; \quad \left( B^{\varphi_R(\lambda)} \right)_{k,l} = \frac{1}{k!} \frac{d^k \varphi_R(\lambda)}{d \tau^k} \bigg|_{\tau = 0} \cdot (67)$$

(iii) The Bell-Jabotinsky matrix of $\phi$ wrt. $\lambda$ (whose column $\ell = 1$ contains the cumulants of order $k$ divided by $k!$) is then given by

$$B^{\phi(\lambda)} = \left( B^{\lambda(\tau)} \right)^{-1} \cdot (68)$$

The cumulant ratios shown in Figure 3 and Figure 7 were computed along these lines. Figure 7 shows that this procedure converges up to the 7th cumulant for a polynomial degree of about $N = 16$. However, this statement is highly dependent on the redshift and smoothing scale at which the density field is considered and the above steps should not be carried out as a black box! Instead, we recommend comparisons such as Figure 7 as tests for robustness.

5 SIMULATED DATA AND COVARIANCE MATRIX

5.1 Data from N-body simulations

In the following paragraphs we summarise the different N-body simulations used in our work. Their cosmological parameters and available primordial non-Gaussianity model amplitudes are compared in Table 2.

5.1.1 Quijote simulations

The Quijote simulations (described in Villaescusa-Navarro et al. 2019) are a large suite of N-body simulations developed for quantifying the cosmological information content of large scale structure observables. They contain 43100 simulations spanning over 7000 cosmological models with variation in $\nu$CDM parameters. At a single redshift, the combined number of particles in the simulation suite is $> 8.5$ Trillion, with a combined volume of 43100 (Gpc$/h^3$) starting from $z = 127$. Here $N$ takes the values: $N = 256$ (low-resolution), $N = 512$ (fiducial-resolution) and $N = 1024$ (high-resolution). Overall, 15000 simulations are provided for the fiducial resolution, allowing to accurately estimate the covariance matrices of cosmological summary statistics. We refer the reader to Villaescusa-Navarro et al. (2019) for further details. In this paper we investigate to the probability distribution function (PDF) of the cosmic matter density field. The PDFs have been computed from the Quijote simulations as follows. First, particle positions and masses are assigned to a regular grid with 512$^3$ cells using the Cloud-in-Cell (CIC) mass assignment scheme (1024$^3$ for the high-resolution simulations). Next, the value of the overdensity field in each grid cell is computed by dividing the mass of each cell by the average mass. Finally, the PDF is estimated by calculating the fraction of cells that lie in a given overdensity bin, over the width of the overdensity bin itself.

The covariance matrix used for our forecasts is entirely obtained from the Quijote simulations. We combine simulated measurements of the PDF and the cumulants at smoothing scales $R = 15$, 30Mpc/$h$ into data vectors $d_i$ ($i = 1, \ldots, 15000$) to estimate the covariance as

$$\hat{C} = \frac{1}{N_{\text{sim}} - 1} \sum_i (d_i - \bar{d})(d_i - \bar{d})^T$$

where $d$ is the mean of all measurements and $N_{\text{sim}} = 15000$ for Quijote. In Figure 4 we show the corresponding correlation matrix of measured PDF histograms as well as of the measured variance, skewness and kurtosis. It can be seen there, that at low and high densities the PDFs are largely anti-correlated with the height of the PDF peak. Related to that, also measurements of the variance are anti-correlated with the peak height of the density PDF. In contrast, all cumulants are positively correlated to the probabilities of low and high densities in the PDFs. However, the onset of the positive correlation moves to higher densities when going to higher cumulant orders.

5.1.2 Simulations by Nishimichi et al. for local primordial non-Gaussianity

For Figure 2 we also investigated a set of simulations that have previously been studied by Uhlemann et al. (2018b). These simulations are based in a non-Gaussian initial condition generator developed by Nishimichi (2012) as well as...
Table 2. Cosmological parameters of the considered simulations, all run for a flat ΛCDM cosmology with $\Omega_m = 1 - \Omega_b$. The values quoted for the amplitude of primordial non-Gaussianity $f_{\text{NL}}$ are $\{f_{\text{NL, eqi}}/f_{\text{NL, orth}}\}$. \\

| Simulation | $\Omega_m$ | $h$ | $n_s$ | $\sigma_8$ | $f_{\text{NL}}$ |
|-----------|------------|-----|-------|------------|----------------|
| Quijote   | 0.3175     | 0.6811 | 0.9624 | 0.834      | -              |
| Nishimichi| 0.279      | 0.701  | 0.96   | 0.8137     | $\{\pm 100, 0\}$ |
| Oriana    | 0.25       | 0.7    | 1.0    | 0.8        | $\{100, -400\}$ |

Measuring $f_{\text{NL}}$ with the bulk of the density PDF

We also explore a cheap way of estimating the covariance matrix for the purpose of future data analyses by generating zero-mean shifted log-normal density fields (Hilbert et al. 2011; Xavier et al. 2016) with $\mu_g = -\sigma_g^2/2$. The 2-point correlation functions of $\delta(x)$ and $g(x)$ are then related through (Hilbert et al. 2011)

$$\xi_g(x) = \ln \left( 1 + \frac{\xi(x)}{\sigma_g^2} \right)$$

For a given choice of $\delta_0$ and a desired power spectrum of $\delta$, this enables us to calculate the corresponding power spectrum of $g$ and hence to draw it from the appropriate Gaussian distribution.

We follow Friedrich et al. (2018) and Gruen et al. (2018) and choose $\delta_0$ such that the resulting log-normal field $\delta$ has the same skewness as that calculated from our fiducial PDF model on a scale of $R = 15\,\text{Mpc}/h$ (and at $z = 1$). In practice, to perform these steps, we make use of the python tool nbodykit\(^3\) and generate the random fields $\delta(x)$ and $g(x)$ on a grid of volume $L^3 = 1/(\text{Gpc}/h)^3$ and with a number of $512^3$ grid points. Note that an important ingredient in our procedure is a theorem by Szykszkowicz & Yanikomeroglu (2009). They show that averaging a log-normal random field of the form given in (70) on a scale $R$ yields a random field that in certain limits is also well described by a log-normal random field, and what is more: by a log-normal random field with the same value of $\delta_0$. This allows us to impose our desired value of $\delta_0$ on the grid scale and obtain the same value also on larger smoothing scales.

6 CONCLUSIONS & DISCUSSION

In this study we have have quantified the impact of primordial non-Gaussianity on the late-time density PDF and extended existing work (Valageas 2002b; Uhlemann et al. 2018c) in two ways: First, we presented a new method of modelling the impact of general non-Gaussian initial conditions on the PDF of the late-time density field. This method is an extension of the steepest descent approach by Valageas (2002b) for Gaussian initial conditions and requires fewer approximations than existing approaches to model the impact primordial non-Gaussianity on the PDF.

Secondly, we considered the full covariance matrix of the log-normal density field to obey the relation

$$\langle \delta(x) \delta(y) \rangle = \sigma_\delta^2 \delta_0^2 \frac{e^{\delta_0}}{\delta_0}.$$
measured density PDFs to forecast the statistical power of such measurements in determining the amplitude of different primordial bispectrum shapes. We considered a combined analysis of the PDF on scales of 15 and 30 Mpc/h at redshift $z = 1$ in a survey volume of $V = 100$ (Gpc/h)$^3$. This is smaller than the effective volume of upcoming surveys such as Spherex with $V_{\text{eff}} \approx 150$ (Gpc/h)$^3$ and larger than existing surveys such as BOSS with $V_{\text{eff}} \approx 55$ (Gpc/h)$^3$ (Doré et al. 2014; Alam et al. 2017). We found that such an analysis can measure the amplitudes of different primordial bispectrum templates with statistical uncertainties of $\Delta f_{NL}^{\text{eqi}} = \pm 3.1$, $\Delta f_{NL}^{\text{orth}} = \pm 10.0$, $\Delta f_{NL}^{\text{th}} = \pm 17.0$, even when treating the non-linear variance of the density field at both smoothing scales as two independent, free parameters. This marginalisation makes our results independent of the amplitude of linear density fluctuations as parametrised by sigma $\sigma_8$. Other cosmological parameters such as $\Omega_m$ have been kept fixed in our analysis (see Uhlemann et al. 2020, for an investigation of the general cosmology dependence of the PDF). But we note that Friedrich et al. (2018) and Gruen et al. (2018) have demonstrated how lensing-around-structures can be used to simultaneously obtain information about parameters of a background $\Lambda$CDM model and higher order moments of the density field in a PDF-based analysis. Our work was done in preparation of a combined analysis of the late-time PDF and the early-universe results of Planck Collaboration (2019). These two cosmological probes have the potential to powerfully complement each other: the CMB can provide information about the background $\Lambda$CDM spacetime, the late-time density PDF contains information about non-linear structure growth and both of them contain independent information about the imprint of primordial non-Gaussianity on the large scale structure.

We want to stress again, that observational data of the large scale structure have already been successfully analysed based on a modelling framework that is closely related to the one presented here (Friedrich et al. 2018; Gruen et al. 2018). Still, to harvest the statistical power of the density PDF demonstrated here, a number of problems have to be addressed which we will discuss in the following sections.

### 6.1 Precision of our model

An important question is whether the analytic modelling and/or the simulated data presented here are accurate enough to analyse future large scale structure data. In Figure 8 we show the relative deviation of PDFs measured in N-body simulations wrt. our theoretical model. The red and blue points use measurements from the fiducial Quijote run (red, spacing of initial particle grid $\approx 2$ Mpc/h) and from a high-resolution version of Quijote (blue, spacing of initial particle grid $\approx 1$ Mpc/h). The green points use an additional set of simulations at an intermediate resolution ($\approx 1.5$ Mpc/h, cf. Baldauf et al. 2016 as well as our Appendix C for details). Our model for the Quijote simulations was computed for a non-linear variance that gives the best fit to the high-resolution data. For the intermediate-resolution data we had to fit the non-linear variance separately, since those simulations were run on a slightly different cosmology.

The upper panel of Figure 8 shows these residuals for a smoothing scale of $R = 30$ Mpc/h and the lower panel shows them for a smoothing scale of $R = 15$ Mpc/h. At both scales, going to higher resolution in the simulations improves the fit between model and data. At the $R = 30$ Mpc/h scale, our model is consistent with the high-resolution simulations at about $1 \sigma$, with $\chi^2 = 9.9$ (expected: $\approx 6 \pm 3.46$). Overall we have 100 high-resolution boxes of size (1 Gpc/h)$^3$, i.e. the error bars on the blue points in Figure 8 correspond to the survey volume of $V \approx 100$ (Gpc/h)$^3$ that we considered throughout this paper. We hence conclude that at $R = 30$ Mpc/h our model captures the non-linear evolution of the PDF accurately enough for analysis on these kind of volumes.
For $R = 15 \, \text{Mpc}/h$ our model agrees with the high-resolution data within $1\%$. But this is significantly larger than the statistical uncertainties expected in future large scale structure data! We ran a limited number of simulations with even higher resolution (cf. Appendix C) to test whether this may be due to inaccuracy of the simulations. Within the limited statistical power of this comparison this does not seem to be the case (cf. Figure C1).

On the modelling side, the most critical approximation we made is the assumption that the reduced cumulants

$$S_n = \frac{\langle \delta_R^n \rangle_c}{\langle \delta_R^2 \rangle_c^{n/2}}$$  \hspace{1cm} (73)

are well modelled by the leading order term of Equation 24 (i.e. by tree level perturbation theory) even in the regime where the cumulants $\langle \delta_R^n \rangle_c$ themselves have significant next-to-leading order contribution. This assumption allows us to apply the variance re-scaling described in Section 4.3 and it is justified by the observation that tidal terms which are not captured by the leading order term are largely erased by smoothing effects in the reduced cumulants (Fosalba & Gaztanaga 1998; Gaztanaga & Fosalba 1998). A way to improve the accuracy of our variance re-scaling (and potentially even make it unnecessary) is to take into account the next-to-leading order term in Equation 24. Ivanov et al. (2019) have demonstrated how to calculate this term for Gaussian initial conditions.

Finally, Figure 2 demonstrates that there is also residual disagreement between how the PDF responds to primordial non-Gaussianity in N-body simulations and in our model. We note that the simulations investigated in this figure were only available to us in resolutions similar to the fiducial Quijote run. And Figure 8 indicates that at our smoothing scales this is insufficient to obtain accurate density PDFs. Also, the simulations considered in Figure 2 exhibit exaggerated amounts of primordial non-Gaussianity. This may impact the performance of the saddle point approximation derived in Section 4 as well as the assumption that the primordial CGF is well approximated by a cubic function (Equation 61).

In summary, we conclude the following regarding the precision of our analytic modelling:

- At $R = 30 \, \text{Mpc}/h$ and $z = 1$ our model is consistent with high-resolution simulated data that corresponds to an overall survey volume of $V \approx 100(\, \text{Gpc}/h)^3$.
- At $R = 15 \, \text{Mpc}/h$, there is a residual disagreement of $\lesssim 1\%$ between our model and the Quijote high-resolution run. We will investigate this in future work using even higher resolution N-body simulations.
- There is also residual disagreement between simulations and our model regarding the response of the PDF to primordial non-Gaussianity (Figure 2). This is likely because of the limited resolution of the considered simulations as well as their exaggerated values of $f_{\text{NL}}$.
- A promising route to improve our analytic modelling is to extend the formalism of Ivanov et al. (2019) for non-Gaussian initial conditions. A major obstacle in applying such an approach in real data analysis is however that it is computationally very expensive.

6.2 Redshift uncertainties

In this paper we have only considered the 3D density field at one redshift. In real situations we cannot access this kind of information for a number of reasons. Firstly, any cosmological observations take place along our past light-cone. Secondly, peculiar velocities of tracers introduce distortions in the mapping between observed redshift and actual radial positions of tracers (redshift-space distortions; see e.g. Mao et al. 2014 for the impact of this on moments of the density field and Uhlemann et al. 2018a for the impact in the density PDF). Finally, both photometric and spectroscopic redshift measurements will only have a finite precision (see Hoyle et al. 2018, for recent work on photometric redshift estimation).

A way to circumvent these problems is to consider 2-dimensional projections of the density field instead - e.g. the convergence of gravitational lensing which is a line-of-sight projection of the matter density field whose PDF can be modelled with approaches similar to the one developed here (Bernaldeau 1995; Bernaldeau & Valageas 2000; Barthelemy et al. 2019). But even the PDF of tracer galaxies from photometric galaxy surveys can be successfully analysed with a formalism related to the one presented here (Friedrich et al. 2018; Gruen et al. 2018). To modify our 3D formalism accordingly, one has to consider cylindrical (as opposed to spherical) collapse in our derivations of Section 4. We defer this adjustment to future work.

6.3 Baryonic feedback, tracer bias and stochasticity

In this paper we have considered dark-matter-only simulations and neglected the impact of baryonic effects on the late-time evolution of structures (see e.g. Schneider et al. 2019). Foreman et al. (2019) find that baryonic physics leaves the Fourier space kernel that relates the late-time matter bispectrum and the late-time power spectrum mostly untouched, suggesting that baryonic feedback can be propagated into higher order correlators through its impact on the 2-point function. The latter can e.g. be calibrated with hydrodynamic simulations, see Eifler et al. (2015); Huang et al. (2019) and references therein. For the modelling of the density PDF this would mean that baryonic physics can be largely incorporated through the variance re-scaling described in Section 4.3. Alternatively, Ivanov et al. (2019) have investigated effective field theory corrections to their model of the matter density PDF. This may also be suitable to incorporate baryonic effects (see e.g. Lewandowski et al. 2015, who investigate this for the power spectrum). This would again require calibration with simulations. The safest strategy to avoid the challenges of modelling baryonic feedback (and in fact any strongly non-linear evolution) is to analyse the PDF at high redshifts, e.g. with Quasars as tracer samples.

When inferring the matter density PDF from a tracer sample (e.g. galaxies) another major obstacle is the fact that there is in general only a biased and stochastic relationship between the tracer density contrast and the total matter density contrast (tracer bias and tracer stochasticity; see Friedrich et al. 2018 and Uhlemann et al. 2018a for their impact of the galaxy density PDF and Mao et al. 2014 for...
their impact on moments of the galaxy density field as well as Dekel & Lahav 1999; Desjacques et al. 2018 for general introductions).

This problem can be circumvented entirely by directly studying the PDF of the convergence of gravitational lensing (Barthelemy et al. 2019) since it directly measures the matter density field. An alternative ansatz pursued by Friedrich et al. (2018) and Gruen et al. (2018) is to jointly study counts-in-cells (i.e. the galaxy density PDF) and lensing-around-cells. In early data of the Dark Energy Survey, they measured the parameters of a stochastic galaxy bias model and could still infer information about both variance and skewness of the matter density field.

ACKNOWLEDGEMENTS

OF gratefully acknowledges support by the Kavli Foundation and the International Newton Trust through a Newton-Kavli-Junior Fellowship and by Churchill College Cambridge through a postdoctoral By-Fellowship. CU kindly acknowledges funding by the STFC grant RG84196 ‘Revealing the Structure of the Universe’. Part of the work of FVN has been supported by the Simons Foundation. TB acknowledges support from the Cambridge Center for Theoretical Cosmology through a Stephen Hawking Advanced Fellowship. MM acknowledges support from the European Union’s Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement No 6655919Y. TN was supported by the Simons Foundation. TB acknowledges support from the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement No 6655919Y. TN was supported in part by JSPS KAKENHI Grant Numbers JP17K14273 and 19H00677, and by the Japanese Science and Technology Agency CREST JPMHCR1414. We would like to thank Kacper Kornet for HPC support as well the JST and the International Newton Trust through a Newton-Kavli Fellowship and by Churchill College Cambridge through a postdoctoral By-Fellowship. CU kindly acknowledges support by the Kavli Foundation and the International Newton Trust through a Newton-Kavli-Junior Fellowship and by Churchill College Cambridge through a postdoctoral By-Fellowship. CU kindly acknowledges support by the Kavli Foundation and the International Newton Trust through a Newton-Kavli-Junior Fellowship and by Churchill College Cambridge through a postdoctoral By-Fellowship. CU kindly acknowledges support by the Kavli Foundation and the International Newton Trust through a Newton-Kavli-Junior Fellowship and by Churchill College Cambridge through a postdoctoral By-Fellowship.

DATA AVAILABILITY

C++ and python tools to compute our model predictions are publicly available at https://github.com/OliverFHD/CosMomentum . Summary statistics measured in the Quijote N-body simulations are publicly available at https://github.com/franciscovillaescusa/Quijote-simulations .

REFERENCES

Arcaivi V., Bartolo N., Matarrese S., Riotto A., 2003, Nuclear Physics B, 667, 119
Alam S., et al., 2017, MNRAS, 470, 2617
Alisahahi M., Silverstein E., Dong T., 2004, Phys. Rev. D, 70, 123505
Arkani-Hamed N., Creminelli P., Mukohyama S., Zaldarriaga M., 2004, JCAP, 2004, 001
Armendariz-Picon C., Damour T., Mukhanov V., 1999, Physics Letters B, 458, 290
Bahbich D., Creminelli P., Zaldarriaga M., 2004, JCAP, 2004, 009
Baldauf T., Schaan E., Zaldarriaga M., 2016, JCAP, 2016, 017
Banerjee A., Dalal N., 2016, Journal of Cosmology and Astroparticle Physics, 2016, 015–015
Barthelemy A., Codis S., Uhlemann C., Bernardeau F., Gavazzi R., 2019, arXiv e-prints, p. arXiv:1909.02615
Bartolo N., Matarrese S., Riotto A., 2004, Phys. Rev. D, 69, 043503
Bartolo N., Matarrese S., Riotto A., 2011, JCAP, 2011, 011
Bernardeau F., 1994, A&A, 291, 697
Bernardeau F., 1995, A&A, 301, 309
Bernardeau F., Reinecke M., 2016, Phys. Rev. D, 94, 063520
Bernardeau F., Uzan J.-P., 2002, Phys. Rev. D, 66, 103506
Bernardeau F., Valageas P., 2000, A&A, 364, 1
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Physics Reports, 367, 1
Bernardeau F., Pichon C., Codis S., 2014, Phys. Rev. D, 90, 103519
Bernardeau F., Codis S., Pichon C., 2015, MNRAS, 449, L105
Biagetti M., 2019, Galaxies, 7, 71
Biagetti M., Lazeyras T., Baldauf T., Desjacques V., Schmidt F., 2017, MNRAS, 468, 3277
Brouwer M. M., et al., 2018, MNRAS, 49, 5189
Burrage C., de Rham C., Seery D., Tolley A. J., 2011, JCAP, 2011, 014
Camera S., Santos M. G., Maartens R., 2019, Journal of Cosmology and Astroparticle Physics, 2019, arXiv e-prints, p. arXiv:1910.01417
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Dalal N., Doré O., Huterer D., Shirokov A., 2008, Phys. Rev. D, 77, 123514
Dekel A., Lahav O., 1999, ApJ, 520, 24
Desjacques V., Seljak U., Iliev I. T., 2009, MNRAS, 396, 85
Desjacques V., Jeong D., Schmidt F., 2018, Phys. Rep., 733, 1
Doré O., et al., 2014, arXiv e-prints, p. arXiv:1412.4872
Eifler T., Krause E., Dodelson S., Zentner A. R., Hearin A. P., Gnedin N. Y., 2015, MNRAS, 454, 2451
Gausmann J. R., Shellard E. P. S., 2009, Phys. Rev. D, 80, 043510
Foreman S., Coulton W., Villaescusa-Navarro F., Barreira A., 2019, arXiv e-prints, p. arXiv:1910.03597
Fosalba P., Gaztanaga E., 1998, MNRAS, 301, 503
Friedrich O., Eifler T., 2018, MNRAS, 473, 4150

MNRAS 000, 000–000 (0000)
APPENDIX A: SPHERICAL COLLAPSE IN $\Lambda\text{CDM}$

In the Newtonian approximation and setting $G = 1 = c$ the evolution of spherical, cylindrical or planar perturbations $\delta$ is described by

$$\ddot{\delta} + \mathcal{H} \dot{\delta} - \frac{N + 1}{N} \frac{\delta^2}{1 + \delta} = 4\pi \rho_m a^2 \delta (1 + \delta),$$

(A1)

where $\tau$ is conformal time, $\mathcal{H} = d \ln a / d \tau$ is the conformal expansion rate and $N = 3$ for a spherical perturbation, $N = 2$ for a cylindrical perturbation and $N = 1$ for a planar perturbation (see Mukhanov (2005) who demonstrates this for $N = 1$ and $N = 3$). To evaluate Equation 28 and related expressions we choose $N = 3$ and solve Equation A1 with the initial conditions

$$\delta_{i} = \delta_{i,RL} \left(D(z_i)\right), \quad \dot{\delta}_{i} = \delta_{i} \mathcal{H}(z_i),$$

(A2)

where $z_i$ is a redshift chosen during matter domination. (In fact, in our calculation of $D(z)$ we set the radiation density $\Omega_r$ to zero and then choose $z_i = 4000$.)

APPENDIX B: FOURIER SPACE CONVENTIONS

We define the Fourier transform of a function $f(x)$ in real space as

$$\tilde{f}(k) := \int d^3 x \ f(x) \ e^{-i x k}. $$

(B1)

In this convention, the convolution theorem takes the form

$$\widetilde{f \ast g}(k) = \tilde{f}(k) \tilde{g}(k), \quad \widetilde{f g}(k) = \frac{1}{(2\pi)^3} \left( \tilde{f} \ast \tilde{g} \right)(k).$$

(B2)

And the power spectrum of the density contrast field $\delta(k)$ is given by

$$\langle \delta(k) \tilde{\delta}(q) \rangle = (2\pi)^3 \delta_D(k + q) \ P(k).$$

(B3)
For the 2-point correlation function of the real space density contrast $\delta(x)$ this means that
\[
\xi(r) = \langle \delta(x+r) \delta(x) \rangle = \int \frac{d^3 k}{(2\pi)^3} \langle \delta(k) \rangle e^{i k(x+r)} e^{i q x} \equiv \int \frac{d^3 k}{(2\pi)^3} P(k) e^{i k r}, \tag{B4}
\]
i.e. it is the Fourier transform of the power spectrum. The variance of the density field when averaged over a top-hat filter of radius $R$ is given by
\[
\langle \delta_L^2 R \rangle = \int \frac{d^3 k_1 d^3 k_2}{(2\pi)^6} \langle \delta(k_1) \delta(k_2) \rangle \tilde{W}_R(k_1) \tilde{W}_R(k_2) \equiv \frac{1}{2\pi^2} \int d\ln k k^3 P_L(k) \tilde{W}_R(k)^2, \tag{B5}
\]
where $\tilde{W}_R(k)$ is given in Equation 10.

APPENDIX C: TESTING THE MODEL ACCURACY WITH SIMULATIONS OF EVEN HIGHER RESOLUTION

In Figure 8 we had complemented the N-body simulations described in Section 5.1 with a set of simulations run by Baldauf et al. (2016). These dark-matter-only simulations consist of 14 boxes of volume $V = (1.5\text{Gpc}/h)^3$ sampled by 1024\(^4\) matter particles. The spacing between particles in the initial density field of these simulations is hence $\approx 1.5\text{Mpc}/h$, which places them between the resolutions of the fiducial Quijote runs and the high-resolution Quijote runs.

To push our test of the impact of simulations resolution to even higher resolutions we ran another 5 N-body boxes at the same cosmology and with the same number of particles as in Baldauf et al. (2016) but with a volume of $V = (0.5\text{Gpc}/h)^3$ instead. In Figure C1 we show the relative deviation between the PDF measured at $R = 15\text{Mpc}/h$ and $z = 1$ and our analytic prediction - again computed for the best-fit non-linear variance. We also show the lower resolution version of these simulations in the figure. It seems that going to resolutions that are even higher than that of the Quijote high-resolution runs (blue points in Figure 8) does not further improve the agreement between our model and the PDF of the simulated density field. However, the statistical uncertainty in the PDF measurements obtained from these 5 small boxes is substantial and we are going to investigate this further with larger sets of high-resolution simulations in the future.

APPENDIX D: SIMULTANEOUSLY VARYING DIFFERENT BISPECTRUM SHAPES

In general, the primordial bispectrum could be a super position of different templates. E.g. Planck Collaboration (2019) simultaneously varied $f_{\text{NL}}^{\text{equil}}$ and $f_{\text{NL}}^{\text{ortho}}$ in their analysis. In Figure D1 we forecast the same measurements as displayed in Figure 6 but this time simultaneously vary $f_{\text{NL}}^{\text{equil}}$ and $f_{\text{NL}}^{\text{ortho}}$. The strong degeneracy between these two parameters in a PDF analysis demonstrates that circular apertures are not well suited to distinguish between different primordial Bispectrum shapes. However, note that the analysis of

APPENDIX E: ASSUMPTION OF GAUSSIAN LIKELIHOOD OF PDF MEASUREMENTS

Throughout this paper we employed the assumption that statistical uncertainties in PDF measurements follow a multivariate Gaussian distribution. This cannot be completely true for two reasons: First, PDF measurements will always be positive which necessarily skews their distribution. This is especially noticeable in the tails of the PDF, where sampling noise is expected to lead to a Poisson-like rather than a Gaussian distribution. Secondly, histograms of density fluctuations that measure the density PDF are normalised such that integration over these histograms gives 1. This means that the bins in which the PDF has been measured cannot be perfectly independent of each other, which seems to further exclude a multivariate Gaussian distribution.

However, in our analysis we have cut the tails of the density PDF which reduces both of these problems. We demonstrate this in Figure E1 where we show histograms of PDF measurements in the Quijote simulations for the lowest density bins that went into our analysis of either the $R = 30\text{Mpc}/h$ and $R = 15\text{Mpc}/h$ smoothing scale. The figure also shows analytic Gaussian distributions with the same mean and variance. There is a close agreement of these analytic distributions with the actual distributions in Quijote. This in
\[ V = 100 (\text{Gpc}/h)^3, \text{ using } R = 15 \& 30 \text{ Mpc}/h \]
and PDFs on the ranges
\[ \delta_{15} \in [-0.4, 0.5], \delta_{30} \in [-0.3, 0.4] \]

**Figure D1.** Same as figure Figure 6 but simultaneously varying \( f_{\text{equi}}^{\text{NL}} \) and \( f_{\text{ortho}}^{\text{NL}} \). The strong degeneracy between these two parameters for both the moment and PDF analysis demonstrates that circular apertures are not well suited to distinguish between different primordial Bispectrum shapes.

\[ \Delta \text{Var}_{15} \times 10^{-4} \quad \Delta \text{Var}_{30} \times 10^{-4} \quad f_{\text{equi}}^{\text{NL}} \quad f_{\text{ortho}}^{\text{NL}} \]

Indicates that even for the bins in our analysis that reach farthest into the tails, the measurement uncertainties are close to Gaussian. We have checked that this also holds for all other density bins in our analysis.

To test for the assumption of multivariate Gaussianity we also look at the \( \chi^2 \) of our entire PDF measurements. To do so we use 10,000 of the Quijote simulations to estimate the covariance of the PDF measurements and then use that covariance to compute the \( \chi^2 \) of the remaining 5,000 Quijote PDF measurements wrt. their mean measurement. The histograms of these 5,000 \( \chi^2 \) values is shown in Figure E2 and again there is a good agreement with the analytic expectation of a (multivariate) Gaussian distribution. Note that with 10,000 simulations any noise in the covariance estimate (and its inverse) will be negligible in this test (Taylor et al. 2013; Friedrich & Eifler 2018).
Figure E1. The blue histograms show the distribution of individual PDF measurements in the Quijote simulations at smoothing scales of $R = 30\,\text{Mpc}/h$ (left panel) and $R = 15\,\text{Mpc}/h$ (right panel) and at the lowest densities that went into our analyses. These density bins are the ones that show the strongest deviation from a Gaussian distribution. The orange lines display analytic Gaussian distributions with the same mean and variance.

Figure E2. The orange histograms show the distribution of $\chi^2$ between the mean PDF measured in 5000 Quijote simulations and the individual PDF measurements in these simulations at smoothing scales of $R = 30\,\text{Mpc}/h$ (left panel) and $R = 15\,\text{Mpc}/h$ (right panel). The blue lines display the analytic $\chi^2$ distributions that would be expected if statistical uncertainties in the PDF measurements were drawn from a multi-variate Gaussian distribution.