Smearing scale in Laguerre reconstructions of the correlation function

Farnik Nikakhtar[1,*] Ravi K. Sheth[1,2] and Idit Zehavi[3]

[1]*Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street,*
Philadelphia, Pennsylvania 19104, USA
[2]*Center for Particle Cosmology, University of Pennsylvania, 209 South 33rd Street,*
Philadelphia, Pennsylvania 19104, USA
[3]*Department of Physics, Case Western Reserve University, Cleveland, Ohio 44106-7079, USA*

To a good approximation, on large scales, the evolved two-point correlation function of biased tracers is related to the initial one by a convolution with a smearing kernel. For Gaussian initial conditions, the smearing kernel is Gaussian, so if the initial correlation function is parametrized using simple polynomials, then the evolved correlation function is a sum of generalized Laguerre functions of half-integer order. This motivates an analytic "Laguerre reconstruction" algorithm which previous work has shown is fast and accurate. This reconstruction requires as input the width of the smearing kernel. We show that the method can be extended to estimate the width of the smearing kernel from the same dataset. This estimate, and associated uncertainties, can then be used to marginalize over the distribution of reconstructed shapes and hence provide error estimates on the value of the distance scale. This procedure is not tied to a particular cosmological model. We also show that if, instead, we parametrize the evolved correlation function using simple polynomials, then the initial one is a sum of Hermite polynomials, again enabling fast and accurate deconvolution. If one is willing to use constraints on the smearing scale from other datasets, then marginalizing over its value is simpler for this latter, "Hermite" reconstruction, potentially providing further speed-ups in cosmological analyses.

I. INTRODUCTION

On large ($\sim$100 h^{-1} Mpc) cosmological scales, the evolved two-point correlation function, even of unbiased tracers of the density field, differs in shape from the unbiased linear theory correlation function [1,2]. To leading order, this change in shape is due to a (three-dimensional) convolution [3]. For Gaussian initial conditions, the convolution kernel is very well approximated by a Gaussian. Convolving a polynomial of order n with a Gaussian yields a generalized Laguerre function of order $n/2$, and [4] used this to motivate an algorithm for estimating the initial shape from measurements of the evolved one. Essentially, fitting a series of half-integer Laguerre functions to the two-point correlation function of biased tracers of the cosmological density field allows one to perform the deconvolution analytically,[1] even in the presence of redshift space distortions [7].

This "Laguerre" reconstruction of the initial shape requires as input a guess for the width of the Gaussian

smearing kernel Σ. In cold dark matter (CDM) models, the physics that gives rise to the convolution relates the smearing scale to the background cosmological model,

$$\Sigma^2 = \frac{1}{3\pi^2} \int dk \, P_{\mathrm{L}}(k, z), \qquad (1)$$

where $P_{\mathrm{L}}(k, z)$ is the linear theory power spectrum [1] at redshift z. So, if the model parameters are sufficiently well known, then they can be used to provide an estimate of the smearing scale. However, it is interesting to ask if the data constrain this scale independently, i.e., without having to assume a fiducial cosmological model. This is particularly interesting because the uncertainties on this estimate can be propagated into uncertainties on the shape of the reconstructed correlation function. In turn, this allows one to marginalize over the value of the smearing scale when estimating cosmological parameters, such as the cosmological distance scale at the redshift of the survey, that is not tied to a fiducial model. This is in contrast to almost all other reconstruction algorithms [8–10], for which a fiducial model must be specified [11,12], although there has been recent progress in accelerating the process of marginalizing over this fiducial choice [13].

In Sec. II, we show that this is indeed possible, in principle, but whether or not the constraint that results is

---

*farnik@sas.upenn.edu
[1]In this respect, the philosophy of the approach is similar to that which motivates the use of Gaussian mixture models to approximate distributions that have been broadened by Gaussian measurement errors [5] and the sinc function representation for evaluating the propagator in Feynman diagrams [6].

sufficiently tight to be interesting (better than $\sim 10\%$ precision) will depend on the dataset (volume and tracer number density). We illustrate the method using the real-space dark matter correlation function, before applying it to the monopole of the redshift-space distorted correlation function measured in mock galaxy catalogs.

Section III shows how this fiducial model-free determination of the smearing scale can be used to estimate the cosmological distance scale and, in particular, to provide realistic estimates of the precision of this determination. If the smearing scale is not well constrained, then this decreases the precision of the distance scale constraint. Therefore, especially for small surveys, it may be that other datasets provide more useful determinations of the smearing. We argue that this motivates consideration of other parametrizations of the reconstruction problem. For example, even if useful constraints on the smearing scale must come from independent datasets, some parametrizations of the reconstruction problem may allow one to marginalize over the value of the smearing scale more easily than others. This is the subject of Sec. IV, where we develop what we call "Hermite reconstruction". A final section summarizes our results.

We illustrate our arguments using the dark matter correlation functions measured in the $z = 1$ outputs of the Quijote simulations [14], as well as mock galaxy catalogs in the $z = 0$ outputs of these same simulations. Each of these 15000 simulations followed the evolution of 512^3 particles in a periodic box of side $L = 1h^{-1}$ Gpc (comoving). The fiducial cosmology of this set is flat ΛCDM with $(\Omega_m, \Omega_b, h, n_s, \sigma_8) = (0.3175, 0.049, 0.6711, 0.9624, 0.834)$, for which Σ of Eq. (1) equals 5.1 h^{-1} Mpc at $z = 1$ and 8.5 h^{-1} Mpc at $z = 0$.

To explore how our results scale with effective volume, we average together the correlation functions measured in 10, 50, and 100 boxes at a time to crudely mimic effective volumes of 10, 50, and 100 $(h^{-1}\,\mathrm{Gpc})^3$ each. These correspond approximately to survey volumes that are between that of BOSS and Euclid, DESI, and a larger futuristic survey. By never reusing a box, this gives us 1500, 300, and 150 independent realizations with which to study cosmic variance associated with our three effective volume choices. In all cases, the correlation functions in these boxes are measured in bins of width 1 h^{-1} Mpc, and all our analyses only make use of the scales between 70 and 110 h^{-1} Mpc. On these scales, the covariance between the different bins is well described by a simple "smeared linear theory plus Poisson shot-noise model" [4,15,16].

II. SMEARING SCALE FROM DATA

Our starting point is that the evolved pair correlation function ξ_{NL} is related to that predicted by linear theory (i.e., the initial one multiplied by a growth factor) ξ_L, by a (three-dimensional) convolution,
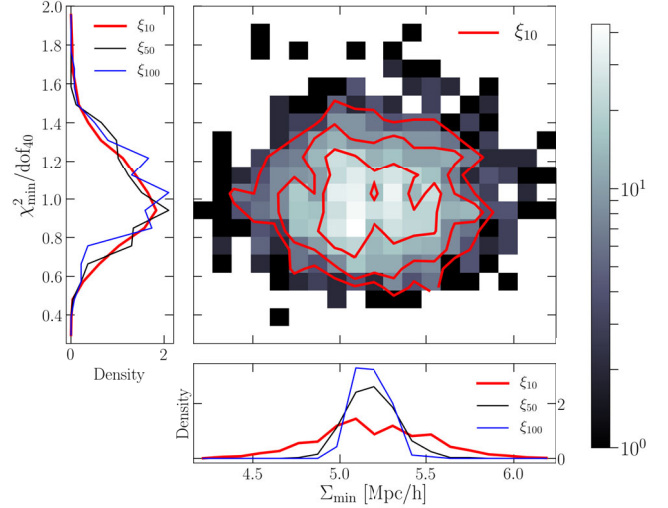

FIG. 1. Estimated Σ when we use exactly the correct shape and amplitude of ξ_L when fitting the right-hand side of Eq. (2) to the measured ξ_{NL}, but Σ is determined from the fit, as well as a measure of the goodness of fit, $\chi^2_{\min}/\mathrm{d.o.f.}$ Main panel shows this joint distribution from 1500 realizations, each of an effective volume 10 h^{-3} Gpc^3. Histograms in the bottom panel show the distribution of Σ for three different effective volumes as indicated (10, 50 and 100 h^{-3} Gpc): as expected, larger volumes return a narrower distribution, hence a tighter constraint on the smearing scale. Histograms in the panel on the left show the associated distributions of $\chi^2_{\min}/\mathrm{d.o.f.}$; as expected, the goodness of fit does not depend on the effective volume.

$$\xi_{\mathrm{NL}}(\mathbf{s}) = \xi_L \otimes G + \xi_{\mathrm{MC}}(\mathbf{s}) \approx \int d\mathbf{r}\, \xi_L(\mathbf{r}) G(\mathbf{s} - \mathbf{r}|\Sigma), \quad (2)$$

e.g., [1], where G is a three-dimensional isotropic Gaussian smoothing kernel of width Σ in each direction, and the final expression assumes that the "mode-coupling term" ξ_{MC} of Ref. [1] can be ignored. We ignore ξ_{MC} in all of the analysis which follows, since none of the main points we make are changed if we include it.

A. "Optimal" estimate

We begin with an exploration of the precision with which the smearing scale can be estimated if the shape and amplitude of the linear theory power spectrum are known. This means that we simply fit the right-hand side of Eq. (2) to the measured ξ_{NL} to determine the value of Σ. The fitting uses measurements in bins of width 1 h^{-1} Mpc over the range 70–110 h^{-1} Mpc and uses the covariance matrix described in [4] to account for the fact that bins in ξ_{NL} are correlated.

The two-dimensional histogram in Fig. 1 shows the joint distribution of the best-fitting Σ and the associated value of $\chi^2_{\min}/\mathrm{degrees}$ of freedom (d.o.f.) constructed from the 1500 realizations of an effective volume of 10 h^{-3} Gpc^3. The red curve in the left-hand panel shows the distribution of

$\chi^2_{\min}$/d.o.f. values obtained by projecting out the Σ values. It peaks close to unity, indicating that the fits are generally acceptable. The red curve in the bottom panel shows the distribution of best-fit Σ values, obtained by projecting out the $\chi^2_{\min}$ values. It peaks at about 5.2 h^{-1} Mpc, which is slightly higher than the linear theory value of 5.1 h^{-1} Mpc, consistent with previous work [1]. The rms of this distribution is 0.33 h^{-1} Mpc, which is about a 7% precision estimate of the smearing scale.

The other curves show similar projections of the histograms (not shown) made from 300 and 150 realizations of effective volumes of size 50 and 100 h^{-3} Gpc^3. Clearly, larger volumes are more constraining: the rms on Σ is 0.14 and 0.1 h^{-1} Mpc, scaling approximately with the inverse of the square root of the effective volume (Appendix A discusses why). Thus, Fig. 1 suggests that if the shape and amplitude of ξ_L are known then the expected precision on the estimated Σ is about $3\sqrt{50 \ h^{-3} \, \mathrm{Gpc}^3/V_{\mathrm{eff}}}$ percent. Allowing for the amplitude of the power spectrum to be a free parameter when fitting only slightly degrades the precision of the Σ estimate. (There is a slight degeneracy between the amplitude and the smearing scale. A higher initial peak must be smeared more to produce the same observed amplitude, but because the smearing matters little at scales $\sim 60 \ h^{-1}$ Mpc, the degeneracy is reduced by including scales that are far from the peak when fitting.)

That said, in practice, the biased tracers will be less abundant than the dark matter, so although the shape of the correlation function of biased tracers should not be too different from that of the dark matter, the measurement errors will be larger. This will degrade the precision of the constraint on Σ, which will propagate to other analyses which use its value. For example, in the context of reconstructing the shape of ξ_L, the distribution shown in Fig. 1 could be used as a prior on the value of the smearing scale when averaging over distributions such as those shown in Fig. 8 of Ref. [4]. Nevertheless, we think further analysis of this estimator of Σ is potentially interesting. In particular, the next subsection explores the accuracy and precision of estimated Σ if neither the shape nor the amplitude of ξ_L is known *a priori*.

B. Laguerre-based estimate

If ξ_L is a function of $|\mathbf{r}|$ and the three-dimensional Gaussian smearing kernel is isotropic with rms Σ in each direction, then the angular integral can be done analytically, making

$$\xi_{\mathrm{NL}}(s) = \int_0^\infty \frac{dr \, r^2}{\Sigma^3} \frac{e^{-(r^2+s^2)/(2\Sigma^2)}}{\sqrt{2\pi}} 2\frac{\sinh(rs/\Sigma^2)}{rs/\Sigma^2} \xi_L(r). \quad (3)$$

The terms other than ξ_L in the integral define a noncentral-Chi distribution in r/Σ with 3 degrees of freedom and

noncentrality parameter s/Σ. So, if ξ_L is written as a sum of polynomials, then ξ_{NL} is a sum over moments of the χ_3 distribution: generalized Laguerre functions. That is, if

$$\xi_L(r) = \sum_{k=0}^{n} a_k (r/\mathcal{R})^k \quad (4)$$

for some set of coefficients a_k ($\mathcal{R}$ merely serves to make the a_k dimensionless, it plays no fundamental role), then

$$\xi_{\mathrm{NL}}(s) = \sum_{k=0}^{n} a_k (s/\mathcal{R})^k (\Sigma/s)^k \mu_k(s/\Sigma), \quad (5)$$

where

$$\mu_{2n} = 2n!! L_n^{(1/2)}(-x^2/2)$$
$$\mu_{2n-1} = (2n-1)!! \sqrt{\frac{\pi}{2}} L_{n-1/2}^{(1/2)}(-x^2/2), \quad (6)$$

and the $L_\beta^{(\alpha)}(x)$ are generalized Laguerre functions. This shows that the shape of ξ_{NL} differs from ξ_L because $\mu_k(x)/x^k \neq 1$.

Reference [4] made the point that if the a_k are determined by fitting Eq. (5) to the observed $\xi_{\mathrm{NL}}(s)$ then the "Laguerre reconstructed/deconvolved" shape $\xi_{\mathrm{Lag}}(r)$ is given by inserting the fitted a_k in Eq. (4). If Σ is known, then the fitting reduces to a simple linear least squares problem. In Eq. (5), the terms that involve Σ multiply the a_k. So if Σ must also be determined from the fitting process, then the problem to be solved is nonlinear, but there is no other complication.

To illustrate, we have fit Eq. (5), with $n = 9$, to the same ξ_{NL} measurements we show in Fig. 1 for a number if choices of Σ, i.e., for each Σ, we solve a linear least squares problem. We then compare the $\chi^2_{\min}$ values and choose that Σ for which $\chi^2_{\min}$ is smallest. (We have checked that the values of $\chi^2_{\min}$/d.o.f. are consistent with unity, indicating the fits are acceptable.) The histograms show the distribution of estimated Σ values for effective volumes of 10, 50, and 100 $(h^{-1} \, \mathrm{Gpc})^3$ (red, black, and blue, respectively). In all three cases, the mean values are slightly larger than the linear theory value of 5.1 h^{-1} Mpc, by about the same amount as in the bottom panel of Fig. 1. The rms values of the distributions are 0.46, 0.21 and 0.15 h^{-1} Mpc; they are about 50% larger than in the bottom panel of Fig. 1 when the shape and amplitude were fixed to their correct values. This is not surprising; the Laguerre-based approach must determine the amplitude *and* shape of ξ_L as well as the value Σ. Viewed from this perspective, the Laguerre-based approach does rather well.

Since the Laguerre-based estimate of Σ could be artificially broadened if there are no parameter choices for which Eq. (4) can provide a good description of the linear theory shape, it is interesting to search for other parametrizations
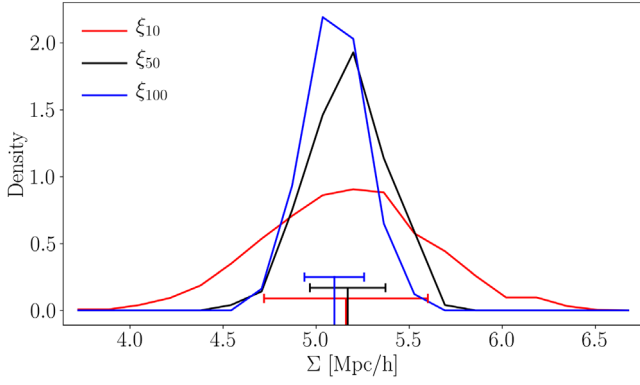
FIG. 2. Distribution of Σ (histograms) estimated from our Laguerre-based analysis [fit Eq. (5) to the measured ξ_{NL}] of the 1500, 300, and 150 realizations of the three different effective volumes indicated (10, 50 and 100 h^{-3} Gpc^3); larger volumes return a narrower distribution, hence a tighter constraint on the smearing scale.

of ξ_L which may constrain Σ better. We study this in Sec. IV.

C. Illustration using biased tracers

So far, we have focused on the dark matter correlation function in configuration space. Real data of rare, biased tracers will include redshift-space distortions and shot noise. To study how our constraints degrade in a more realistic setting, we have explored the following extreme scenario. We work with ξ_0, the monopole of the redshift-space distorted correlation function of mock galaxies from the same Quijote simulation set (the Molino suite of mock galaxy catalogs [17]) but now at $z = 0$. These mock catalogs use the standard [18] halo occupation distribution model. The number density of the mock galaxies is 1.63×10^{-4} h^{-3} Mpc^3, so the shot noise is significantly larger than for the dark matter, and the bias factor is $b = 2.4$. In addition, the lower redshift means that the smearing scale is larger: $\Sigma = 8.5$ h^{-1} Mpc. Moreover, the fact that we are working in redshift space means that the smearing scale and the bias factor are modified to

$$b_{eff}^2 \approx b^2[1 + 2\beta/3 + \beta^2/5] \tag{7}$$

and

$$\Sigma_{eff}^2 = \Sigma^2\left[1 + \frac{f(2+f)}{3}\frac{1 + 6\beta/5 + 3\beta^2/7}{1 + 2\beta/3 + \beta^2/5}\right], \tag{8}$$

where $f \equiv d\ln D/d\ln a = 0.53$ and $\beta \equiv f/b = 0.22$ (see Ref. [7] for a derivation). In our mocks, $b_{eff} = 2.6$ and $\Sigma_{eff} = 10.34$ h^{-1} Mpc.

In practice, the Laguerre method assumes that Eq. (3) still applies with $\xi_L \rightarrow b_{eff}^2\xi_L$ and $\Sigma \rightarrow \Sigma_{eff}$. So the question
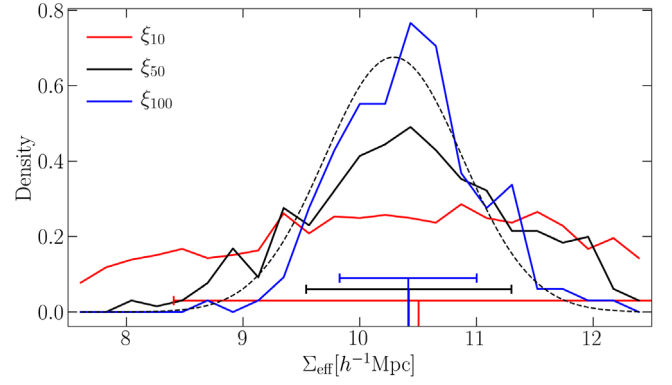

FIG. 3. Distribution of estimated Σ_{eff} from our Laguerre-based analysis [fit Eq. (5) to the measured redshift-space correlation function]. Three histograms show results from 1500, 300, and 150 realizations of the three different effective volumes as indicated (10, 50, and 100 h^{-3} Gpc^3). Vertical bars near the bottom mark the mean values of the distributions; horizontal error bars show the mean plus and minus the rms. Dashed curve shows the constraint on Σ_{eff} from propagating the error on a single realization of the ξ_{50} ensemble (the one shown in Fig. 4); it is slightly narrower than the actual distribution of measurements in the ensemble (black histogram).

is this: how well does the method recover Σ_{eff}, while also recovering the shape and amplitude of $b_{eff}^2\xi_L$?[2]

Figure 3 shows the result of estimating Σ_{eff} by fitting ninth-order Laguerres to the 1500, 300, and 150 realizations of ξ_0 (each realization has an effective volume of 10, 50, or 100 h^{-3} Gpc^3). The method returns distributions that are centered at $\Sigma_{eff} = 10.50$, 10.41, and 10.42 h^{-1} Mpc for the smallest to largest volumes. These are all close to the theory value of $\Sigma_{eff} = 10.34$ h^{-1} Mpc. The rms scatter around the mean is about 20%, 8.5%, and 6% of the mean value; while this is significantly worse than for the dark matter at $z = 1$, it is still a sub-10% determination for DESI-like volumes ($\sim$50 h^{-3} Gpc^3).

Perhaps more importantly, these values for the rms are similar to (although slightly larger than) those returned by the fitting procedure when fitting a single simulation. Since the distribution of Σ shown in Fig. 3 is approximately Gaussian, one can approximate $p(\Sigma)$ by assuming it is Gaussian and treating the uncertainty on Σ determined from a single realization as its rms. The dashed curve in Fig. 3 shows an example; it was obtained by propagating the errors on the fit, shown as a black dashed curve, to the symbols in Fig. 4 (which show one member of the ξ_{50}

___
[2]This "scale-independent bias" model is only an approximation. In Ref. [4], we noted that "higher-order" and/or "derivative" bias schemes are easily incorporated into the Laguerre reconstruction methodology, but we ignore these complications here for the same reason we ignore the mode-coupling contribution ξ_{MC} to Eq. (2): none of our main points are changed by including them.
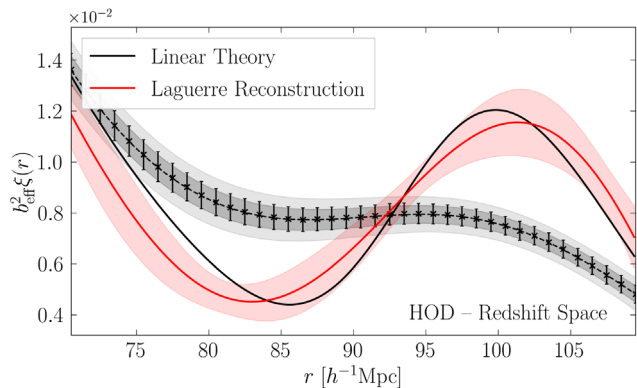
FIG. 4. Measured redshift-space monopole for a mock galaxy catalog in an effective volume of 50 h^{-1} Gpc (symbols with error bars), the ninth-order Laguerre fit to it [black dashed curve, which shows Eq. (5)], and associated reconstruction [red curve, which shows Eq. (4) with the a_k determined by the black dashed curve]. This particular realization happens to have $\Sigma_{\mathrm{eff}} = 10.3$ h^{-1} Mpc. Gray bands show the 68% and 95% ranges covered by 300 such realizations of the observed redshift-space monopole, and the pink band shows the 68% range covered by the 300 associated reconstructed shapes ξ_{Lag}. Solid black curve shows the linear theory $b_{\mathrm{eff}}^2 \xi_{\mathrm{L}}$.

ensemble). It is worth emphasizing that, because the Laguerre methodology is agnostic about the shape or amplitude of ξ_{L}, the $p(\Sigma)$ that it returns—essentially the shape shown in Fig. 3—is obtained without assuming anything about the linear theory power spectrum.

III. APPLICATIONS

The previous section showed that the Laguerre methodology is able to provide useful estimates of the smearing scale that are not tied to a fiducial cosmology. We now discuss what additional science such estimates enable.

A. Accuracy and precision of reconstructed distance scale

Figure 3 shows the distribution of Σ_{eff} values at which Eq. (5) best fits the measured ξ_0. However, the best fit also determines a set of coefficients a_k which, when inserted in Eq. (4), determine the reconstructed shape ξ_{Lag}. To illustrate, the black symbols (with error bars) in Fig. 4 show a single realization of ξ_0 in a 50 h^{-3} Gpc^3 volume, and the dashed black curve shows the best fit of Eq. (5) to it. This fit determines the coefficients a_k as well as Σ_{eff}. This particular realization has $\Sigma_{\mathrm{eff}} = 10.3$ h^{-1} Mpc. The gray bands show the regions that enclose 68% and 95% of 300 realizations of ξ_0 in the same effective volume. Notice that the measurement (symbols) shows almost no peak or dip because, by $z = 0$, particles have moved far from their initial positions, and their speeds (which give rise to redshift-space distortions) are also large. This is also true of the ensemble average bracketed by the gray regions;
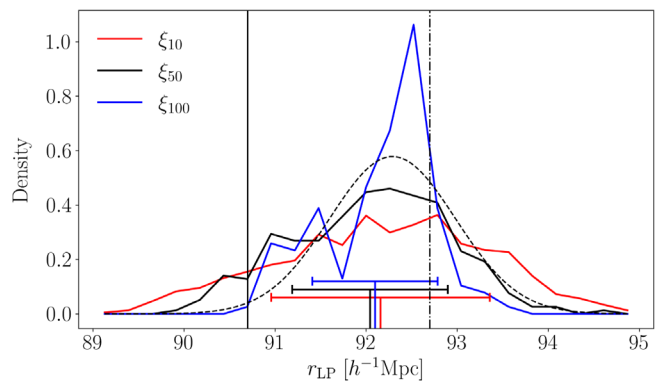

FIG. 5. Distribution of reconstructed r_{LP} from our Laguerre-based analysis [fit Eq. (5) to ξ_0, the monopole of the measured redshift-space correlation function; then, insert the a_k coefficients of the best fit in Eq. (4) to define ξ_{Lag}. Finally, determine r_{LP} from the peak and dip scales of ξ_{Lag}]. Three histograms show results for three different effective volumes as indicated. Vertical bars near the bottom mark the mean values of the distributions; horizontal error bars show the mean $\pm$ the rms. Smooth black dashed curve shows a Gaussian distribution with mean and rms determined from the single ξ_{Lag} shown in Fig. 4. Vertical solid line shows the mean r_{LP} scale measured in ξ_0 prior to reconstruction (it is similar for all three effective volumes), and dotted-dashed line shows the scale in linear theory.

in fact, in about 8% of the simulations, there is no discernable peak or dip.

Despite this extreme smearing, the Laguerre reconstructed shape ξ_{Lag} [red curve shows Eq. (4) with the best fit a_k] is reasonably close to that of linear theory ξ_{L} (black solid).[3] The pink bands show the region that encompasses 68% of the 300 reconstructed ξ_{Lag} curves, i.e., the reconstructions of the curves which resulted in the gray bands.

To turn each of these reconstructed shapes into an estimate of the cosmological distance scale (at the survey redshift), we use the "linear point" r_{LP}, which is defined as the midpoint between the peak and dip scales in the correlation function,

$$r_{\mathrm{LP}} \equiv \frac{r_{\mathrm{peak}} + r_{\mathrm{dip}}}{2}. \qquad (9)$$

For the background cosmology of our mock catalogs, $r_{\mathrm{LP}} = 92.7$ h^{-1} Mpc in linear theory. Reference [19] shows that this scale is interesting because r_{LP} in the evolved correlation function is approximately the same as in ξ_{L}. Distance scale estimates boil down to estimating r_{LP} in each $z = 0$ mock and providing a realistic error bar for it.

Although r_{LP} evolves less than either the peak or dip scales [4,19], it does shift to slightly smaller scales at later times. In the $z = 0$ mocks we are studying here, the

---

[3]Note that because we are ignoring both mode-coupling and scale-dependent bias the reconstructed shape is not as good as it could possibly be. However, because our main concern here is to illustrate the qualitative effects of Σ, we continue to ignore them.

smearing is so large that there is no discernable peak or dip in about 8% of the simulations having effective volumes of 50 h^{-3} Gpc^3. (For the 10 and 100 h^{-3} Gpc^3 volumes, this fraction is 9% and 6%, and they are not used in the reconstruction process.) But in the others, the values of r_{LP} estimated from the Laguerre fit [i.e., by finding the maximum and minimum of Eq. (5) using the best-fitting values of a_k and Σ] are centered on 90.7 h^{-1} Mpc, with an rms scatter of about ± 1.3 h^{-1} Mpc. This is a significant offset from its linear theory value, so the question is if the r_{LP} estimates from reconstructed ξ_{Lag} are closer to the linear theory value of 92.7 h^{-1} Mpc, and what is the associated uncertainty on this "reconstructed" value.

The solid black curve in Fig. 5 shows the distribution of reconstructed r_{LP} values determined from the ξ_{Lag} that were reconstructed from the same ~300 ξ_{50} measurements which led to Fig. 3. The black vertical bar near the bottom marks the mean value of the distribution. and the horizontal error bar shows the mean $\pm$ the rms. The smooth dashed curve shows a Gaussian distribution with mean and variance determined from propagating the errors on r_{LP} for a single member of this ensemble—the one shown by the symbols in Fig. 4, that is best fit by the black dashed curve there, and whose reconstructed ξ_{Lag} is the dashed red curve there (whereas the mean and variance come from standard error analysis, using a Gaussian shape is an extra assumption; this is reasonable as the distribution defined by the ensemble is not too non-Gaussian.) The mean of this single realization is consistent with the ensemble mean reconstructed value which is centered on ~92.1 h^{-1} with an rms scatter of ± 0.85 h^{-1} Mpc. Evidently, even though the reconstructed ξ_{Lag} is not as close to ξ_L as in Refs. [4,7]—presumably because the smearing scale at $z = 0$ is so large—Laguerre reconstruction improves the accuracy *and* precision of the distance scale estimate.[4]

The other histograms in Fig. 5 show that, for the other effective volumes as well, the reconstructed values are centered on ~92.1 h^{-1} with an rms scatter that is slightly smaller for the larger volumes. In all cases, accuracy and precision are both improved compared to r_{LP} in the original ξ_0 (i.e., prior to Laguerre reconstruction).

It is interesting to contrast this with the accuracy and precision that result from fixing Σ to a fiducial value (in this case, 10.34 h^{-1} Mpc) and only determining the a_k from the fits. Doing so does not change the mean r_{LP}, but the error bar is about 30% smaller (± 0.6 h^{-1} Mpc rather than ± 0.85 h^{-1} Mpc). This is not surprising; as Ref. [4] discusses, fixing Σ to a fiducial value in this way is a little like performing the reconstruction step with a prior on the background cosmological model, and this artificially

reduces the estimated error bar. In effect, determining Σ (in addition to the a_k) from the fit and using the best-fit Σ to reconstruct frees one from this dependence on a fiducial model. In this sense, the estimate of r_{LP} that results (the one shown in Fig. 5) has been marginalized over the *a priori* unknown value of Σ.

Evidently, even in this extreme smearing scenario, the Laguerre reconstruction methodology—which makes no assumption about the expected shape of the baryon acoustic oscillations (BAO) signal—returns a distance scale estimate that is accurate to sub-percent precision for volumes that are larger than ~10 h^{-3} Gpc^3. Since future surveys target similar comoving volumes but at higher redshifts where the smearing is smaller, we expect our methodology to return sub-percent precision on the estimated distance scale. Of course, for smaller survey volumes, other datasets may provide better constraints on Σ and hence on the prior distribution one should use when marginalizing. We discuss how one might proceed in such cases in Sec. IV.

B. Other uses of the estimated smearing scale

Our Laguerre-based estimate of the smearing scale is particularly interesting as Σ potentially provides an estimate of the amplitude of $P_L(k, z)$, and hence the linear theory growth factor, that is not degenerate with the bias of the tracers (but see Ref. [20] for why this may not be exactly true). Crudely speaking, this is because on the scales which dominate the integrand in Eq. (1), the power spectrum has approximately the same shape as P_L; only its amplitude is different: $P_{obs}(k) \approx b^2 P_L(k)$. Therefore, if one defines Σ_{obs}^2 by inserting P_{obs} in place of P_L in Eq. (1), then the ratio $\Sigma_{obs}/\Sigma \approx b$.

In practice, we must apply this methodology to the monopole of the redshift space clustering signal, for which

$$\frac{\Sigma_{obs}}{\Sigma_{eff}} \approx b_{eff} \left[1 + \frac{f(2+f)}{3} \frac{1 + 6\beta/5 + 3\beta^2/7}{1 + 2\beta/3 + \beta^2/5} \right]^{-1/2} \quad (10)$$

depends on both f and β rather than b alone. If β is determined from the angular dependence of the clustering signal (e.g., the ratio of the monopole to the quadrupole), then this can be combined with Eq. (10) to estimate f. We intend to explore this estimate of f in future work.

IV. THE HERMITE LIMIT

The previous section noted that it may be interesting to search for other parametrizations of ξ_L which return tighter constraints on the smearing scale. However, if the tracers are sufficiently sparse that the uncertainties on Σ become large, then the constraint on Σ may not be sharp enough to be interesting (either for constraining b or for reconstructing the shape of ξ_L). If one must use constraints on the smearing scale from other datasets to perform a BAO reconstruction, then it is interesting to ask if alternative

---

[4]Figure 6 in [7] suggests that ignoring the mode-coupling term as we have done here leads to a slight underestimate of r_{LP}, so this may be why our reconstructed value is still biased slightly low.

parametrizations (to Laguerre) simplify the process of marginalizing over the value of the smearing scale when reconstructing ξ_L. As the Introduction notes, this provides additional motivation for exploring other parametrizations of ξ_L.

To address this, we begin with Eq. (3), and consider the limit in which $r \gg \Sigma$ and $s \gg \Sigma$ (Fig. 1 and 2 show that the smearing scale is indeed much smaller than the BAO scales of interest). Then, $2\sinh(rs/\Sigma^2) \approx e^{rs/\Sigma^2}$ making

$$\xi_{NL}(s) \approx \int_0^\infty \frac{dr\, r^2}{\Sigma^3} \frac{e^{-(r-s)^2/(2\Sigma^2)}}{\sqrt{2\pi}rs/\Sigma^2} \xi_L(r), \qquad (11)$$

so

$$s\xi_{NL}(s) \approx \int_{-\infty}^\infty dr \frac{e^{-(r-s)^2/(2\Sigma^2)}}{\sqrt{2\pi}\Sigma} r\xi_L(r). \qquad (12)$$

Note that we can extend the lower limit of integration down to $-\infty$ only if $r - s \ll \Sigma$. If we parametrize ξ_L using a simple polynomial [i.e., Eq. (4)], then the integral above can be done analytically. It is easy to check that each a_k multiplies $(\Sigma/\mathcal{R})^k$ times a polynomial in s/Σ. This polynomial is the same as that which appears in the Laguerre expansion $\mu_k(s/\Sigma)$, when one takes the $s \gg \Sigma$ limit ($E_1 \to 1$ and $E_2 \to 0$ in Eq. (A3) of Ref. [4]).

There is no *a priori* reason for parametrizing ξ_L with a simple polynomial. If we parametrize $r\xi_L(r)$ using the probabilist's Hermite polynomials instead,

$$r\xi_L(r) = \sum_{k=0}^n a_k H_k\left(\frac{r - r_{fid}}{\mathcal{R}}\right), \qquad (13)$$

then Appendix B shows that the integral in Eq. (12) can still be done analytically. For each k, the result is $(s - r_{fid})^k/\mathcal{R}^k$ plus additional terms which are lower-order polynomials in s multiplied by terms proportional to $(\mathcal{R}/\Sigma)^2$. By carefully grouping these other terms, it is possible to find that $r\xi_L(r)$ which, when inserted in Eq. (12), produces a simple polynomial in s. Namely, if

$$r\xi_L(r) = \sum_{k=0}^n a_k \mathcal{H}_k\left(\frac{r - r_{fid}}{\mathcal{R}}, \frac{\Sigma}{\mathcal{R}}\right), \qquad (14)$$

where the $\mathcal{H}_k$ are given in Appendix B, then

$$s\xi_{NL}(s) = \sum_{k=0}^n a_k \left(\frac{s - r_{fid}}{\mathcal{R}}\right)^k. \qquad (15)$$

Therefore, if we determine the a_k by fitting the simple polynomial of Eq. (15) to the observed $s\xi_{NL}(s)$, then the reconstructed/deconvolved $r\xi_L(r)$ is given by Eq. (14), provided we first assume a value for $\Sigma/\mathcal{R}$ (which we discuss shortly). Note that whereas "Laguerre reconstruction" has a simple polynomial as the reconstructed

shape of $\xi_L(r)$, this "Hermite reconstruction" of $r\xi_L(r)$ has a simple polynomial as the nonlinear shape of $s\xi_{NL}(s)$.

A. Dependence on smearing scale

In practice, we will not know the correct value of Σ, so it is interesting to study the sensitivity of Hermite reconstruction to incorrect choices of Σ. In this regard, the structure of this Hermite reconstruction problem has two surprising consequences. First, because Σ does not appear in Eq. (15), fitting it to the measurements yields *no* information about Σ. That is, if we fit $s\xi_{NL}(s)$ to a simple polynomial, then we cannot make a plot like Fig. 2! Second, for the same fitted a_k, varying Σ only changes the reconstructed shape [Eq. (14)]. In contrast, for Laguerre reconstruction, each choice of Σ requires a new fit [Σ appears in Eq. (5) for ξ_{NL}]. In this sense, Hermite reconstruction is more efficient than Laguerre: one only needs to determine the a_k once.

Although $s\xi_{NL}(s)$ does not depend on Σ, the Hermite reconstruction of ξ_L, which we refer to as ξ_{Her}, does. So we now turn to the value of Σ. There are two natural choices. One is to treat the Laguerre-based analysis (e.g., Fig. 2) as providing a prior on the value of Σ. But if this is rather broad (e.g., for sparse tracers), it may be that an alternative approach is more constraining. Following [7], this second approach exploits the fact that, for all biased tracers, the smearing scale is expected to be well approximated by Eq. (1). Suppose we use Σ_{obs}^2 to denote the result of inserting the observed $P_{obs}(k)$ in Eq. (1). On the large scales (small k) which dominate the integral (in ΛCDM models), $P_{obs}(k) \approx b_{10}^2 P_L(k)$, making $\Sigma \approx \Sigma_{obs}/b_{10}$. In this approximation, the uncertainty in what to use for Σ boils down to what to use for b_{10}.

We use b_{fid} to denote our best guess for this value, and so we define $\Sigma_{fid} \equiv \Sigma_{obs}/b_{fid}$. This suggests that if we fit the observed correlation function to Eq. (15), with $\mathcal{R} = \Sigma_{fid}$, then the effect of varying b from its fiducial value yields reconstructed

$$r\xi_{Her}(r) = \sum_{k=0}^n a_k \mathcal{H}_k\left(\frac{r - r_{fid}}{\Sigma_{fid}}, \frac{b_{fid}}{b_{10}}\right). \qquad (16)$$

Note that when $b_{10} = b_{fid}$, then $\mathcal{H}_k \to H_k$; the reconstructed shape is a simple sum of Hermite polynomials (of course, this is only true if $\Sigma \approx \Sigma_{obs}/b_{10}$). This allows a straightforward estimate of the uncertainty on the reconstructed ξ_{Her}: fitting to ξ_{NL} yields the covariance matrix of the fitted a_k. For a given choice of b_{fid}/b_{10}, this can be used to produce uncertainty bands around the shape given by Eq. (16), and further marginalizing over the value of b_{fid}/b_{10} gives the full uncertainty on the reconstructed shape (see [4] for a detailed discussion).

We must also decide what to use for r_{fid}. We are particularly interested in the BAO scale, around which
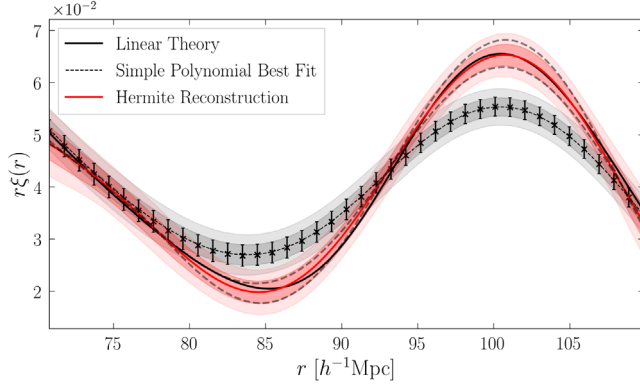
FIG. 6. Comparison of Hermite reconstruction [Eq. (14)] of the shape of the $z = 1$ dark matter correlation function (solid red curve; pink bands show 1- and 2-standard deviation uncertainties) with the linear theory shape (solid black) when $\Sigma = 5.1$ h^{-1} Mpc, the linear theory value, is used. Dashed gray curves show reconstructions when Σ is assumed to be larger or smaller by 10%. Symbols with error bars show the measured (i.e., evolved) dark matter correlation function. Dashed black curve shows the best fit of Eq. (15) to these measurements, which was used to determine the a_k coefficients used in the reconstruction; gray bands show the 1- and 2-standard deviation uncertainties.

the correlation function exhibits a peak and a dip. Therefore, a reasonable way to determine $r_{\rm fid}$ is as follows. Initially choose $r_{\rm fid}$ arbitrarily—a reasonable choice would use the expected value of $r_{\rm LP}$ of Eq. (9) in the current best-fitting cosmological model. Then, fit Eq. (15), with $\mathcal{R} = \Sigma_{\rm fid}$, to the observed $s\xi_{\rm NL}(s)$. Find those scales $r_{\rm peak}$ and $r_{\rm dip}$ where $d\xi_{\rm NL}/ds = 0$. In practice, since $d[s\xi]/d\ln s = s\xi + sd\xi/d\ln s$, the peak and dip scales in the evolved correlation function are at those s where

$$\sum_{k=0}^{n} a_k \left(\frac{s - r_{\rm fid}}{\Sigma_{\rm fid}}\right)^{k-1} \left[\frac{ks - (s - r_{\rm fid})}{\Sigma_{\rm fid}}\right] = 0. \quad (17)$$

Now, set $r_{\rm fid}$ equal to $(s_{\rm peak} + s_{\rm dip})/2$ and refit. Doing this ensures that the higher-order polynomials contribute less and less in the vicinity of $r_{\rm LP}$ (see Fig. 10 in [4] for an explicit demonstration). The a_k which result can then be inserted in Eq. (16). The peak and dip scales in the reconstruction are where $d[r\xi_{\rm Her}]/d\ln r = r\xi_{\rm Her}$,

$$\sum_{k=0}^{n} a_k \left[\frac{kr}{\Sigma_{\rm fid}}\mathcal{H}_{k-1}\left(\frac{r - r_{\rm fid}}{\Sigma_{\rm fid}}, \frac{b_{\rm fid}}{b_{10}}\right) - \mathcal{H}_k\left(\frac{r - r_{\rm fid}}{\Sigma_{\rm fid}}, \frac{b_{\rm fid}}{b_{10}}\right)\right] = 0,$$

$$(18)$$

where we have used the fact that $d\mathcal{H}_j(x)/dx = j\mathcal{H}_{j-1}(x)$. This can be used to determine how the value of $r_{\rm LP}$ in the reconstruction depends on b. One can, of course, weight each of these values by a prior on the value of b.

B. The Hermite-reconstructed shape

Figure 6 illustrates the various steps associated with Hermite reconstruction. The symbols with error bars show the measured dark matter correlation function at $z = 1$ in an effective volume of 50 h^{-3} Gpc^3. The dashed black curve shows the best fit of Eq. (15) with $n = 9$ (i.e., a ninth-order simple polynomial) to these measurements; associated gray bands are the 1- and 2-standard deviation uncertainties. This best fit determines the coefficients a_k. The solid red curve with pink error bands shows the associated Hermite reconstruction [Eq. (14) with the a_k determined from the fit to the symbols, and $\Sigma = 5.1$ h^{-1} Mpc or, equivalently, Eq. (16) with $b_{\rm fid} = b_{10} = 1$] and corresponding uncertainties. This red curve should be compared with the solid black one, which shows linear theory. Evidently, when the correct smearing scale is assumed, then Hermite reconstruction works quite well. (It appears to work much better than the Laguerre reconstructions shown in Fig. 4 only because here we are working with the dark matter at $z = 1$ rather than redshift-space distorted mock galaxies at $z = 0$. In fact, for the $z = 1$ dark matter, the reconstructed shapes, $\xi_{\rm Lag}$ and $\xi_{\rm Her}$, are very similar.)

Dashed gray curves show the Hermite reconstructions when the smearing scale is assumed to be larger or smaller by 10% (the a_k are the same, of course). Comparison with Fig. 4 in Ref. [4] shows that, in all cases, the linear theory shape is recovered at least as well as it is for Laguerre reconstruction. We have also checked that $r_{\rm LP}$ of Eq. (9) in the Hermite reconstructions depends on a smearing scale similarly to the Laguerre reconstructions: weakly (see, e.g., Fig. 6 of Ref. [4]). Therefore, reconstructed distance scale estimates and their uncertainties from the Hermite reconstructions are comparable to those from Laguerre reconstruction. This is reassuring because Laguerre reconstruction of the distance scale is accurate, precise, and fast [4,7]. Moreover, as we noted above, Hermite reconstruction is even more efficient, requiring only a single determination of the coefficients a_k.

C. Discussion

One might have thought that how one chooses to parametrize the nonlinear correlation function is of little consequence—provided the goodness of fit is acceptable. Our analysis has shown that some parametrizations are more useful than others. The Laguerre parametrization of $\xi_{\rm NL}$ [Eq. (5)] has smearing scale Σ dependence in $\xi_{\rm NL}$ but none in ξ_L, whereas our modified Hermites [Eq. (13)] have no Σ dependence in $s\xi_{\rm NL}(s)$ but some in $r\xi_L(r)$. As a result, the Laguerre-based parametrization of $\xi_{\rm NL}$ constrains Σ (Fig. 2), whereas the simple polynomial parametrization of $s\xi_{\rm NL}(s)$ associated with Eq. (13) does not. (Some of this is a consequence of ignoring the mode-coupling term. Had we included it, then the Laguerre approach would have no additional Σ dependence in $\xi_{\rm NL}$, whereas the modified

Hermites approach would. However, in practice, the mode-coupling contribution is too small to matter.)

This raises the question of whether or not there is a parametrization of ξ_{NL} that yields tighter constraints on Σ. While one is allowed to fit data with a model whether or not the model makes good physical sense, our results suggest that a parametrization that is closer to the physics will fare better: the modified Hermites have the shape of the linear theory correlation function depending on time (because Σ depends on time), which is unphysical. In contrast, the Laguerre parametrization is consistent with the physics.

While a detailed investigation of physically reasonable parametrizations is beyond the scope of this work, we have performed the following test. We parametrize ξ_L using Eq. (13), for which the associated ξ_{NL} is analytic and depends on $\mathcal{R}/\Sigma$. This is similar to the Laguerre case, for which ξ_L is a simple polynomial that depends on $\mathcal{R}$ but not Σ—so it is as physically reasonable. Because of this similarity, we can study how this parametrization constrains Σ. We have found that the analogue of Fig. 2 is almost identical; there is no significant difference between parameterizing $\xi_L(r)$ using simple polynomials or $r\xi_L(r)$ using Hermites. Of course, this does not exclude the possibility that there are other parametrizations which will better constrain Σ.

Finally, in the context of parametrizations, it is worth noting that BAO analyses that assume a fiducial (typically ΛCDM-motivated) model when estimating the distance scale attempt to account for the fact that the fiducial model —either for the background cosmology or for the bias between observed tracer and the underlying dark matter field—may not be correct by adding $A_1 r^{-1} + A_2 r^{-2}$ and then marginalizing over the values of A_1 and A_2 [21]. Since such terms are not present in our approach, one might wonder if their inclusion would bias our results. We checked this by explicitly adding $A_1/s + A_2/s^2$ to Eq. (5) or $A_1 + A_2/s$ to Eq. (15) prior to fitting. We found that these extra terms have almost no impact on our analysis.

V. CONCLUSIONS

To a good approximation on BAO scales, the evolved correlation function ξ_{NL} is related to the initial one, ξ_L, by a convolution [Eq. (3)]. Fitting a series of half-integer Laguerre polynomials [Eq. (5)] to the evolved two-point correlation function allows one to constrain the smearing scale Σ of the convolution kernel (Fig. 2) even when neither the amplitude nor shape of ξ_L are known. In addition, when applied to ξ_{NL} at different redshifts, the method correctly returns the fact that the smearing scale is larger at later times (compare Figs. 2 and 3). Our Laguerre approach shows that to constrain the value of Σ it is enough to endow the parametrization of the $\xi_{NL} - \xi_L$ relation with the correct structure (i.e., one that reflects the fact that the two are related by a convolution).

In configuration space, the smearing is expected to be approximately independent of the nature of the observed tracers—i.e., of halo or galaxy bias. However, in redshift space, the effective smearing is expected to depend weakly on bias [Eq. (8)]; our Laguerre-based estimates of Σ_{eff} in redshift-space distorted mock galaxy catalogs are consistent with this expectation (Fig. 3).

In the Laguerre framework, knowledge of the smearing scale allows one to deconvolve and hence reconstruct the shape of ξ_L from measurements of ξ_{NL}, without any prior assumptions about the shape or amplitude of ξ_L. As the shape and amplitude of the reconstructed correlation function can be used to constrain cosmological parameters, the Laguerre methodology can be used to provide more realistic estimates of the precision of the constraints. In particular, the estimated accuracy and precision of the Laguerre-reconstructed constraints do not depend on choosing a fiducial cosmological model (also see discussion at the end of Sec. IV C). We demonstrated this for the linear point feature [Eq. (9)] in the reconstructed ξ_L (Figs. 4 and 5).

In practice, such constraints will depend on the nature of the biased tracers and the volume of the survey. For small survey volumes (e.g., BOSS), the constraint on the smearing scale is not tight, so marginalizing over its value can significantly weaken constraints on cosmological parameters. When this occurs, it may be preferable to use tighter constraints on Σ which come from other datasets, and then, we argued that the full Laguerre-based analysis (which may be justified in a DESI-like survey) may not be necessary. Provided that Σ is a small fraction of the scales of interest, the simple polynomial-modified Hermites [Eq. (16) with Appendix B) combination for $s\xi_{NL}(s)$ and $r\xi_L(r)$ provides a more efficient way of marginalizing over the value of Σ when quantifying the accuracy and precision on the distance scale estimate (Fig. 6 and associated discussion).

We illustrated many of our points using mock galaxy catalogs at $z = 0$, where the smearing is so large that the BAO feature in ξ_{NL} is nearly completely smeared out. As our Laguerre-based results were promising, nevertheless, we are in the process of implementing the ideas presented here in realistic mock galaxy catalogs which are more relevant to the next generation of cosmological surveys.

Finally, although we have focussed on the BAO smearing scale, recent work has highlighted the benefits of combining full-shape analyses of galaxy power spectra with BAO distance scale estimates to constrain cosmological parameters [22]. Since both Hermite and Laguerre reconstructions reproduce the full shape of ξ_L over a rather broad range of scales, their speed and simplicity enable the development of a similar program in configuration rather than Fourier space. We hope this feature of our reconstructions is exploited in future work.

APPENDIX A: DEPENDENCE OF χ^2 ON EFFECTIVE VOLUME

Sufficiently close to the best fit (the minimum of the χ^2), we expect the χ^2 curve to be approximated by a quadratic function. We can expand it around the best-fit coefficients $\hat{\mathbf{a}}$ as

$$\chi^2(\mathbf{a}) = \chi^2(\hat{\mathbf{a}}) + \mathcal{D} \cdot (\mathbf{a} - \hat{\mathbf{a}}) + \frac{1}{2}(\mathbf{a} - \hat{\mathbf{a}})^T \cdot \mathcal{M} \cdot (\mathbf{a} - \hat{\mathbf{a}}),$$

$$(A1)$$

where $\mathcal{D}_i = \partial \chi^2 / \partial a_i$ is the first derivative vector, and $\mathcal{M}_{ij} = \partial^2 \chi^2 / \partial a_i \partial a_j$ is the Hessian matrix. The best-fit coefficients $\hat{\mathbf{a}}$ will be determined from the fact that the first-order derivatives $\partial \chi^2 / \partial a_i$ are zero for all a_i at the minimum $\chi^2_{\min} = \chi^2(\hat{\mathbf{a}})$, so we can write the above equation as

$$\Delta \chi^2 \equiv \chi^2(\mathbf{a}) - \chi^2_{\min} = (\mathbf{a} - \hat{\mathbf{a}})^T \cdot (\mathcal{M}/2) \cdot (\mathbf{a} - \hat{\mathbf{a}}). \quad (A2)$$

The χ^2 function can be expressed as the following weighted inner product:

$$\chi^2 = (\xi - \hat{\xi})^T \cdot \mathcal{C}^{-1} \cdot (\xi - \hat{\xi}), \quad (A3)$$

where $\mathcal{C}^{-1}$ is the inverse of the covariance matrix, ξ the measurement, and $\hat{\xi}$ our best fit. As we discussed in the text, if the value of the smearing scale Σ is known, then our fitting model is linear in coefficients. For a linear model, the best fit can be written in matrix form as $\hat{\xi} = \mathcal{A} \cdot \hat{\mathbf{a}}$, where $\mathcal{A}$ is the design matrix. In our case, the elements of the design matrix are $\mathcal{A}_{ij} = \mathcal{H}_j([r_i - r_{\mathrm{fid}}]/\Sigma_{\mathrm{fid}}, b_{\mathrm{fid}}/b_{10})$. Inserting this linear model in Eq. (A3) and calculating the associated Hessian matrix yields

$$\Delta \chi^2 = (\mathbf{a} - \hat{\mathbf{a}})^T \cdot (\mathcal{A}^T \mathcal{C}^{-1} \mathcal{A}) \cdot (\mathbf{a} - \hat{\mathbf{a}}). \quad (A4)$$

The covariance matrix $\mathcal{C}$ scales as the inverse of the effective survey volume, so $\mathcal{C}^{-1}$, and consequently, $\chi^2(\mathbf{a}) - \chi^2_{\min}$ scale as the effective volume.

Let $\delta \mathbf{a} = \mathbf{a} - \hat{\mathbf{a}}$ be a change in the fitting coefficients whose first element is arbitrary δa_1, but the rest of whose elements are selected to minimize the $\Delta \chi^2 = \chi^2(\mathbf{a}) - \chi^2_{\min}$. Then, δa_1, the uncertainty on the value of a_1, is

$$\delta a_1 = \pm \sqrt{\frac{\Delta \chi^2}{[\mathcal{C}^{-1}]_{11}}}, \quad (A5)$$

which scales as the inverse of the square root of the effective volume.

For the nonlinear case (when Σ is not known), Eq. (A3) is still valid. By calculating the second derivatives in the general form, $\Delta \chi^2$ of the nonlinear model can be written as

$$\Delta \chi^2 = \delta \mathbf{a}_k \left[\frac{\partial \hat{\xi}_i^T}{\partial a_k} [\mathcal{C}^{-1}]_{ij} \frac{\partial \hat{\xi}_j}{\partial a_l} - (\xi - \hat{\xi})_i^T [\mathcal{C}^{-1}]_{ij} \frac{\partial^2 \hat{\xi}_j}{\partial a_k \partial a_l} \right]_{kl} \delta \mathbf{a}_l,$$

$$(A6)$$

where all the derivatives are evaluated at the best fit values of the fitting parameters.

In the linear model, the second derivatives are all zero, and this expression reduces to Eq. (A4). In a nonlinear model, finding the best fit $\hat{\xi}$ must proceed iteratively, and then, we only need to insert the solution into Eq. (A6). Since the inverse of the covariance matrix appears in both the first and second terms in the square brackets above, the scaling with the effective volume for the nonlinear case is the same as for the linear one.

As noted in the main text, to estimate the best-fitting Σ, we solve a linear least square problem for a number of choices of Σ. We then compute the χ^2 values and choose the Σ that minimizes the χ^2. For this step and measuring the uncertainty on the best-fitting Σ for each realization, we fit a quadratic function $A_0(\Sigma - A_1)^2$ to the χ^2 values as a function of Σ. The A_1 parameter gives us the best-fitting Σ, and then, by setting the confidence level of $\Delta \chi^2 = 1$, the uncertainty on Σ can be determined by $1/\sqrt{A_0}$.

For the error bar on r_{LP}, we need to propagate the uncertainty from the fitted parameters of the correlation function to the positions of the peak and the dip and finally to the linear point. In this nonlinear case that Σ is unknown, we should write the linear point position as a function of the polynomial coefficients $\{a_k\}$ and Σ and then expand the result around the best-fit parameters. The error bar on r_{LP} can be written as

$$\sigma_{\mathrm{LP}} = \left\{ \sum_{i,j} \frac{\partial r_{\mathrm{LP}}}{\partial \beta_i} [\mathrm{Cov}(\bar{\boldsymbol{\beta}})]_{ij} \frac{\partial r_{\mathrm{LP}}}{\partial \beta_j} \right\}^{1/2}, \quad (A7)$$

where $\beta = \{a_0, a_1, \ldots, a_k, \Sigma\}$, and $\bar{\beta}$ denotes the best fitting values.

APPENDIX B: MODIFIED HERMITE POLYNOMIALS

We are interested in integrals of the form

$$I_k(y, \beta) \equiv \int dx \frac{e^{-(y-x)^2/2}}{\sqrt{2\pi}} H_k(\beta x) \tag{B1}$$

$$= (1 - \beta^2)^{k/2} H_k\left(\frac{y\beta}{\sqrt{1 - \beta^2}}\right), \tag{B2}$$

where the $H_k(x)$ are the probabilist's Hermite polynomials, and $\beta > 0$. We could use this to write the $s\xi_{\text{NL}}(s)$ that is associated with Eq. (13) of the main text. This expression would depend on β, making it the Hermite analog of the Laguerre reconstruction we discussed in the main text. However, one of our goals in the main text was to show that, with the right combination of Hermite polynomials, it is possible to write ξ_{NL} as a simple polynomial in $s/R = y\beta$.

To achieve this, start with the fact that when $\beta = 1$, then

$$I_k(y, 1) = \int dx \frac{e^{-(y-x)^2/2}}{\sqrt{2\pi}} H_k(x) = y^k. \tag{B3}$$

We then note that

$$H_n(\beta x) = n! \sum_{i=0}^{[n/2]} \frac{\beta^{n-2i}(\beta^2 - 1)^i}{2^i} \frac{H_{n-2i}(x)}{i!(n - 2i)!}. \tag{B4}$$

This with the definition of $I_k(y, 1)$ show that when $\beta \neq 1$, then $I_k(y, \beta)$ is a sum of polynomials that are each multiplied by different powers of β. For example, the term of highest order in y is $(\beta y)^k$. We can remove all the other terms by subtracting appropriate combinations of β and

$H_j(\beta x)$. Doing so defines the functions called $\mathcal{H}_k(x, \beta)$ in the main text. Explicitly, they are

$$\mathcal{H}_0(\beta x) = H_0(\beta x),$$

$$\mathcal{H}_1(\beta x) = H_1(\beta x),$$

$$\mathcal{H}_2(\beta x) = H_2(\beta x) - AH_0(\beta x),$$

$$\mathcal{H}_3(\beta x) = H_3(\beta x) - 3AH_1(\beta x),$$

$$\mathcal{H}_4(\beta x) = H_4(\beta x) - 6AH_2(\beta x) + 3A^2 H_0(\beta x),$$

$$\mathcal{H}_5(\beta x) = H_5(\beta x) - 10AH_3(\beta x) + 15A^2 H_1(\beta x),$$

$$\mathcal{H}_6(\beta x) = H_6(\beta x) - 15AH_4(\beta x) + 45A^2 H_2(\beta x) - 15A^3 H_0(\beta x),$$

$$\mathcal{H}_7(\beta x) = H_7(\beta x) - 21AH_5(\beta x) + 105A^2 H_3(\beta x) - 105A^3 H_1(\beta x),$$

$$\mathcal{H}_8(\beta x) = H_8(\beta x) - 28AH_6(\beta x) + 210A^2 H_4(\beta x) - 420A^3 H_2(\beta x) + 105A^4 H_0(\beta x),$$

$$\mathcal{H}_9(\beta x) = H_9(\beta x) - 36AH_7(\beta x) + 378A^2 H_5(\beta x) - 1260A^3 H_3(\beta x) + 945A^4 H_1(\beta x), \tag{B5}$$

where $A \equiv \beta^2 - 1$. This shows that $\mathcal{H}_k \to H_k$ as $\beta \to 1$. Moreover, note the similarity of the structure to that for the H_k themselves: the numerical coefficients are the same as for H_k, with $x^j \to H_j$ and with each extra term receiving an additional power of A. That is,

$$\mathcal{H}_n(\beta x) \equiv n! \sum_{m=0}^{[n/2]} \frac{H_{n-2m}(\beta x)}{m!(n - 2m)!} \frac{(1 - \beta^2)^m}{2^m}. \tag{B6}$$

This, with Eq. (B2) for $I_k(y, \beta)$, makes it easy to check that Eq. (14) leads to Eq. (15). This also makes it easy to see that $d\mathcal{H}_j(y)/dy = j\mathcal{H}_{j-1}(y)$, which mirrors the fact that $dH_j(x)/dx = jH_{j-1}(x)$.

[1] M. Crocce and R. Scoccimarro, Phys. Rev. D **77,** 023533 (2008).

[2] V. Desjacques, M. Crocce, R. Scoccimarro, and R. K. Sheth, Phys. Rev. D **82,** 103529 (2010).

[3] S. Bharadwaj, Astrophys. J. **472,** 1 (1996).

[4] F. Nikakhtar, R. K. Sheth, and I. Zehavi, Phys. Rev. D **104,** 043530 (2021).

[5] M. R. Blanton, D. W. Hogg, N. A. Bahcall, J. Brinkmann, M. Britton, A. J. Connolly, I. Csabai, M. Fukugita, J. Loveday, A. Meiksin et al., Astrophys. J. **592,** 819 (2003).

[6] R. Easther, G. Guralnik, and S. Hahn, Phys. Rev. D **63,** 085017 (2001).

[7] F. Nikakhtar, R. K. Sheth, and I. Zehavi, Phys. Rev. D **104,** 063504 (2021).

[8] N. Padmanabhan, X. Xu, D. J. Eisenstein, R. Scalzo, A. J. Cuesta, K. T. Mehta, and E. Kazin, Mon. Not. R. Astron. Soc. **427,** 2132 (2012).

[9] M. Schmittfull, T. Baldauf, and M. Zaldarriaga, Phys. Rev. D **96,** 023505 (2017).

[10] E. Sarpa, C. Schimd, E. Branchini, and S. Matarrese, Mon. Not. R. Astron. Soc. **484,** 3818 (2019).

[11] S. Anselmi, P.-S. Corasaniti, A. G. Sanchez, G. D. Starkman, R. K. Sheth, and I. Zehavi, Phys. Rev. D **99,** 123515 (2019).

[12] B. D. Sherwin and M. White, J. Cosmol. Astropart. Phys. 02 (2019) 027.

[13] M. Hansen, A. Krolewski, and Z. Slepian, arXiv:2112 .06438.

[14] F. Villaescusa-Navarro, C. Hahn, E. Massara, A. Banerjee, A. M. Delgado, D. K. Ramanah, T. Charnock, E. Giusarma, Y. Li, E. Allys *et al.*, Astrophys. J. Suppl. Ser. **250,** 2 (2020).

[15] R. E. Smith, R. Scoccimarro, and R. K. Sheth, Phys. Rev. D **77,** 043525 (2008).

[16] G. Parimbelli, S. Anselmi, M. Viel, C. Carbone, F. Villaescusa-Navarro, P. S. Corasaniti, Y. Rasera, R. Sheth,

G. D. Starkman, and I. Zehavi, J. Cosmol. Astropart. Phys. 01 (2021) 009.

[17] C. Hahn and F. Villaescusa-Navarro, J. Cosmol. Astropart. Phys. 04 (2021) 029.

[18] Z. Zheng, A. L. Coil, and I. Zehavi, Astrophys. J. **667,** 760 (2007).

[19] S. Anselmi, G. D. Starkman, and R. K. Sheth, Mon. Not. R. Astron. Soc. **455,** 2474 (2016).

[20] V. Desjacques and R. K. Sheth, Phys. Rev. D **81,** 023526 (2010).

[21] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden, and M. Manera, Mon. Not. R. Astron. Soc. **449,** 835 (2015).

[22] O. H. E. Philcox, M. M. Ivanov, M. Simonović, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 05 (2020) 032.