

# Optimal Transport Reconstruction of Baryon Acoustic Oscillations

Farnik Nikakhtar\*

*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104 – USA*

Ravi K. Sheth

*Center for Particle Cosmology, University of Pennsylvania, Philadelphia, PA 19104 – USA and  
The Abdus Salam International Center for Theoretical Physics, Strada Costiera 11, Trieste 34151 – Italy*

Bruno Lévy

*Universités de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France*

Roya Mohayaee

*Sorbonne Université, CNRS, Institut d’Astrophysique de Paris, 98bis Bld Arago, 75014 Paris, France*

(Dated: March 4, 2022)

A model-independent, weighted semi-discrete, fast optimal transport algorithm to reconstruct the Lagrangian positions of proto-halos from their evolved Eulerian positions is presented. Tests with state-of-art cosmological simulations show that the positions of proto-halos are reconstructed accurately, without having to assume a background cosmology. The algorithm, which makes use of a mass estimate of the biased tracers, but is robust to errors in this estimation, recovers the shape and amplitude of the initial pair correlation function of the tracers, enabling sub-percent precision in the BAO distance scale that is not tied to a cosmological model. In principle, our algorithm also allows direct and independent determinations of the bias factor and the smearing scale, potentially providing new methods for breaking the degeneracy between the bias factor  $b$  and  $\sigma_8$ .

The Baryon Acoustic Oscillations (BAO) are frozen sound waves from the pre-decoupling era that leave a peak in the 2-point correlation function (2PCF) of the linear matter distribution [1] on scales of order 140 Mpc. This ‘BAO peak’ provides an important distance scale for measuring the expansion history of the Universe [2]. However, gravitational evolution shifts and smears this peak in the late-time baryon plus dark matter distribution and hence in the distribution of biased tracers [3–5]. This potentially biases the inferred distance scale, so most BAO analyses aim to sharpen the peak and remove its shift by undoing the effects of the gravity [6, 7]. This ‘reconstruction’ of the density field is feasible because on scales of order 140 Mpc, perturbation theory based analyses should be accurate. However, it is complicated by the fact that we only observe biased tracers of the full density field. Therefore, most density field reconstruction methods make assumptions about the background cosmology as well as the nature of the bias between the observed tracers and the dark matter [8–10]. In what follows, we describe and test a method in which only the bias must be specified.

Optimal transport (hereafter OT) is a powerful mathematical framework which has recently found applications in diverse branches of science [11–14]. In astrophysics, OT reconstruction is a deterministic algorithm that recovers the initial Lagrangian positions  $\mathbf{q}$  of a given final Eulerian distribution of particles  $\mathbf{x}$  by solving the Monge-Ampère-Kantorovich (MAK) problem [15, 16]. In semi-discrete OT reconstruction [17], the final particle distribution is considered to have evolved from a smooth,

uniform, continuous initial field (rather than a discrete cartesian grid). In this setting, instead of a single point, a patch of Lagrangian space, a *Laguerre cell*, is assigned to each evolved object. The map that assigns the Laguerre cells to final Eulerian positions is the unique solution to a semi-discrete optimization problem as well [18]. The Laguerre tessellation is specified by a set of  $\psi_i$  values that is the unique maximizer of what is known as the Kantorovich dual [19–21]:

$$K(\psi) = \sum_i \int_{V_i^\psi} \left[ \frac{1}{2} |\mathbf{x}_i - \mathbf{q}|^2 - \psi_i \right] d^3 q + \sum_i v_i \psi_i, \quad (1)$$

subject to  $V_i^\psi \neq \emptyset$  for all  $i$ . A Laguerre diagram reduces to the Voronoi diagram [22] for identical  $\psi_i$ . Strictly speaking, the above expression assumes that the Lagrangian density field is uniform, so  $v_i$  is the volume of the  $i$ th Laguerre cell, imposed as a constraint. The  $\psi_i$  coefficients correspond to the Lagrange multiplier associated with the constraint. Clearly, for this to be valid, the sum over all  $v_i$  needs to be equal to the total volume. Previous semi-discrete OT reconstructions were limited to using the same  $v_i$  for all cells [17]. In contrast, the weighted semi-discrete OT we use here allows a different  $v_i$  for each Laguerre cell (see Figure 1).

The objective function  $K(\psi)$  is concave – this guarantees the existence and uniqueness of  $\psi$  – and smooth ( $C^2$ ). Therefore, the exhaustive combinatorial computation of the discrete-discrete Monge problem can be replaced with an efficient and convergent Newton method [23–25]. Hence, our semi-discrete approach scales as  $\mathcal{O}(N_{LC} \log N_{LC})$  [17], where  $N_{LC}$  is the number of La-

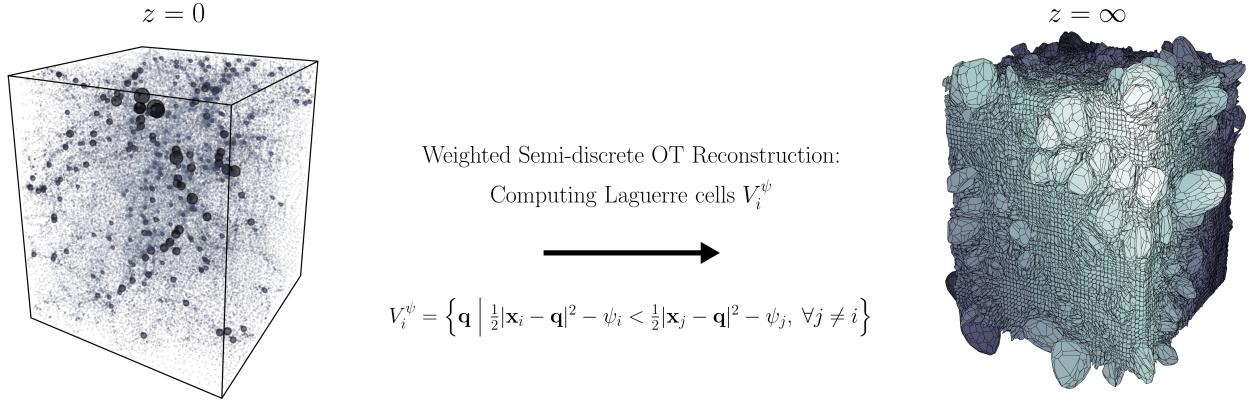


FIG. 1: Schema for our fast weighted and model-independent semi-discrete algorithm: Reconstruction of the Laguerre cells from the present distribution of biased tracers (halos) and dark matter field particles (dust). The volume of each Laguerre cell represents the mass of the object (halo or dust) to which it corresponds. The initial power spectrum/correlation function is obtained directly from the distribution of the barycenters of the Laguerre cells; there is no need to make any additional Lagrangian approximation or any other model-dependent intermediary.

guerre cells: a *significant* improvement over previous algorithms.

The dark matter field is comprised of equal mass particles. The assumption of uniform initial conditions implies that all the  $v_i$  in equation 1 are the same. Ref. [17] show that the associated OT algorithm accurately recovers the set of displacements  $\mathbf{x} - \mathbf{q}$  which maps the  $z = 0$  position  $\mathbf{x}$  of each (equal-mass) dark matter particle to the barycenter of its associated Laguerre cell  $\mathbf{q}$ .

There are two reasons why treating biased tracers, and halos in particular, is more complicated.

- i. The first is that halos typically span a range of masses.

Accommodating this turns out to be straightforward. The assumption of uniform initial conditions means that one simply sets  $v_i \propto m_i$  in equation (1), where  $m_i = n_i m_p$  is the mass of each object and  $m_p$  is the particle mass.

- ii. The second is more pernicious: Whereas the initial dark matter field was uniform, the initial proto-halo distribution was not [26, 27]. Moreover, the Lagrangian-space clustering of the proto-halo patches depends on their mass. This means that one may *not* assume that the (mass-weighted) barycenters of the Laguerre cells of the biased tracers are uniformly distributed. In principle, this seriously complicates the OT approach.

One way to address this is by including a model for the ‘dust’ – the mass that is not associated with the biased tracers – since the initial distribution of (mass-weighted) tracers plus dust should be uniform. Typically (e.g. for BAO surveys), the biased tracers account for only  $\sim 20\%$  of the total mass density. Fortunately, our OT algorithm

is fast, so the additional computational load required to reconstruct dust as well as the biased tracers is offset by the simplicity of the initial condition.

Therefore, in addition to the masses of the Eulerian objects which are observed, our weighted OT algorithm also requires a guess regarding the Eulerian spatial distribution of the dust. In what follows, we use a random subset of the dark matter particles which are not assigned to the biased tracers to model the dust. This will be more complicated in real data but we defer discussion of how to model the dust (using results from, e.g., [28–30]), to a follow-up study.

Since the OT reconstructed field is uniform, in all previous (discrete or semi-discrete) OT work, the reconstructed positions were evolved forward in time by using  $\mathbf{q} + D(z)(\mathbf{x} - \mathbf{q})$  for some  $D(z) \ll 1$ , before spatial statistics were measured. Since  $D(z)$  is like a linear theory growth factor (indeed, this evolution is like the linear Lagrangian ‘Zel’dovich’ approximation), this made the OT-reconstruction appear to be cosmology dependent, even though OT itself makes no assumptions other than convexity of the potential flow of matter. In contrast, for biased tracers, the Lagrangian distribution is non-trivial, so it should be possible to extract the BAO feature directly from the OT-reconstructed positions  $\mathbf{q}$  of the Laguerre cells, without needing cosmological perturbation theory.<sup>1</sup> In what follows, we demonstrate that our fast, weighted semi-discrete algorithm (a) works well for biased tracers when a good proxy for the dust is available,

---

<sup>1</sup> We show elsewhere that this can be done even for the dark matter if one estimates the Lagrangian correlation function by weighting each Laguerre cell by the divergence of  $\mathbf{x} - \mathbf{q}$ , or by plotting the Laplacian of the correlation function of the displacement field.

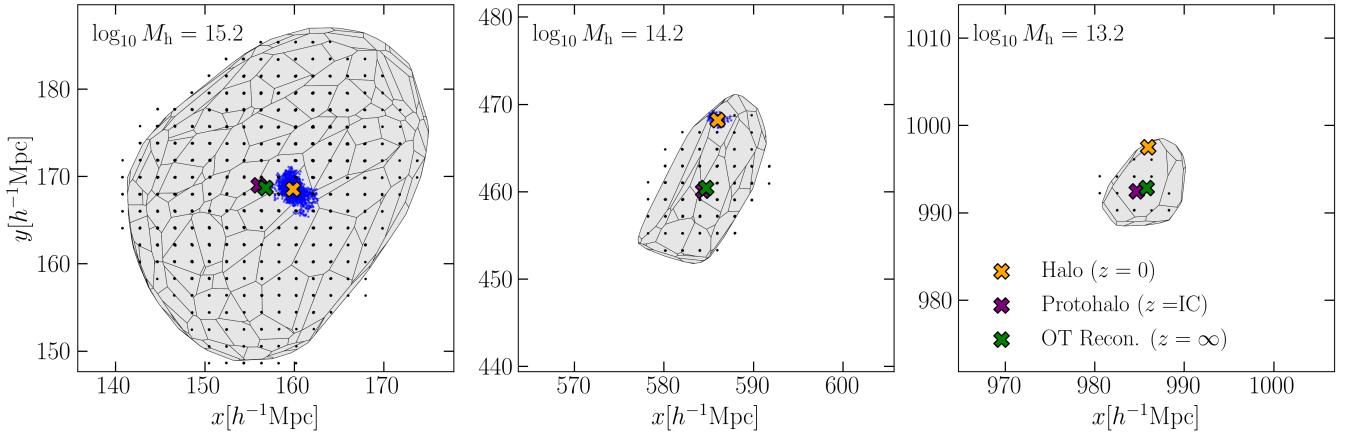


FIG. 2: Evolved, Eulerian distribution of particles that make up a halo at  $z = 0$  (blue dots), and the initial, Lagrangian (i.e.  $z = \infty$ ) distribution of the same particles (grid-like black dots). Grey shaded region shows the Laguerre cell associated with each halo, as determined by the weighted-OT reconstruction. Each panel shows results for a single halo: the halo mass (in solar units) is shown in the upper left corner. In each panel, yellow and red crosses show the Eulerian ( $z = 0$ ) and Lagrangian ( $z = \infty$ ) center-of-mass positions. The displacement between the two is what smears and shifts the BAO feature in the two-point correlation function of these biased tracers. Green cross shows the barycenter of the Laguerre cell associated with each halo, as determined by the weighted-OT reconstruction. It is extremely close to the red cross, indicating that the weighted-OT algorithm describes the displacements extremely well.

and (b) requires *no* assumption on cosmology – it is fully model-independent.

We demonstrate our results using halos identified in 20 realizations of the HADES simulations of [31]. Each simulation follows the gravitational evolution of  $512^3$  identical particles, each of mass  $m_p = 6.566 \times 10^{11} h^{-1} M_\odot$ , in a periodic cube of side  $L = 1 h^{-1} \text{Gpc}$  in which the background cosmology is flat,  $\Omega_\Lambda = 1 - \Omega_m$ , with  $(\Omega_m, \Omega_b) = (0.3175, 0.049)$ , and Hubble constant today  $H_0 = 100h \text{ km s}^{-1} \text{Mpc}^{-1}$ , with  $h = 0.6711$ . The initial fluctuation field is Gaussian so it is completely specified by its power spectrum,  $P_{\text{Lin}}(k)$ , which is taken from CLASS, with shape  $n_s = 0.9624$  and amplitude set by  $\sigma_8 = 0.833$ .

The biased tracers we consider in what follows are ‘halos’ identified in the  $z = 0$  output of each box using a friends-of-friends algorithm with linking length  $b = 0.2$ . We only use halos with more than 20 particles; this corresponds to a minimum halo mass of  $1.3 \times 10^{13} h^{-1} M_\odot$ . About 75% percent of the particles in the box are not bound up in such halos: these make up the ‘dust’.

Our OT algorithm takes as input the list of halo masses and positions and some knowledge about the spatial distribution of the dust. In practice, we randomly select a subset of the dust particles: If we choose a fraction  $p$ , then we would replace the mass of each by  $m_p/p$ . Since halos have at least 20 particles, we only consider  $p \geq 0.1$ , as this ensures that each dust particle is less massive than the least massive halo. In the results which follow, we set  $p = 1$ , although our results are not particularly sensitive to this choice.

The list of  $n_i$  particles which makes up a halo of mass

$m_i = n_i m_p$  defines a protohalo patch in the initial conditions. We refer to the center of mass of this halo at the time it was identified ( $z = 0$ ) as its Eulerian position  $\mathbf{x}_i$ , and that of the protohalo as the Lagrangian position  $\mathbf{q}_i$  ( $z = \infty$ ; strictly speaking the initial conditions are at  $z_{\text{IC}} = 99$ , but the displacement between  $z_{\text{IC}}$  and  $\infty$  is negligible). Figure 2 shows three representative halos (black dots) and the protohalo patches (blue dots) from which they evolved. Yellow and green crosses show the center of mass positions  $\mathbf{x}_i$  and  $\mathbf{q}_i$ . The offset between the two,  $\mathbf{S}_i \equiv \mathbf{x}_i - \mathbf{q}_i$  is what causes both the shifting and the smearing of the BAO feature in the halo pair correlation function. The grey shaded region shows the Laguerre cell associated with each halo, as determined by the weighted-OT reconstruction; it reconstructs the shape of the protohalo patch quite well. This is non-trivial, because the long axis of the protohalo becomes the short axis of the final halo – the two are not simply rescaled versions of one-another [32, 33]. We discuss the shapes elsewhere; here, we focus on the positions. The green cross in each panel shows the barycenter of the Laguerre cell; it is extremely close to the red cross, indicating that the weighted-OT algorithm describes, hence undoes, the displacements extremely well. Therefore, we expect it to reconstruct the BAO feature very well.

Figure 3 compares pair correlation functions  $\xi$  averaged over 20 simulation boxes. In each panel, the solid orange curve shows the average halo correlation function  $\xi_{\text{Eul}}(r)$  (i.e. using the Eulerian positions  $|\mathbf{x}_i - \mathbf{x}_j|$ ), and the yellow bands around it show the standard deviation. The purple curve shows a similar analysis of the Lagrangian protohalos,  $\xi_{\text{Lag}}$  (which is built from the  $|\mathbf{q}_i - \mathbf{q}_j|$  pairs),

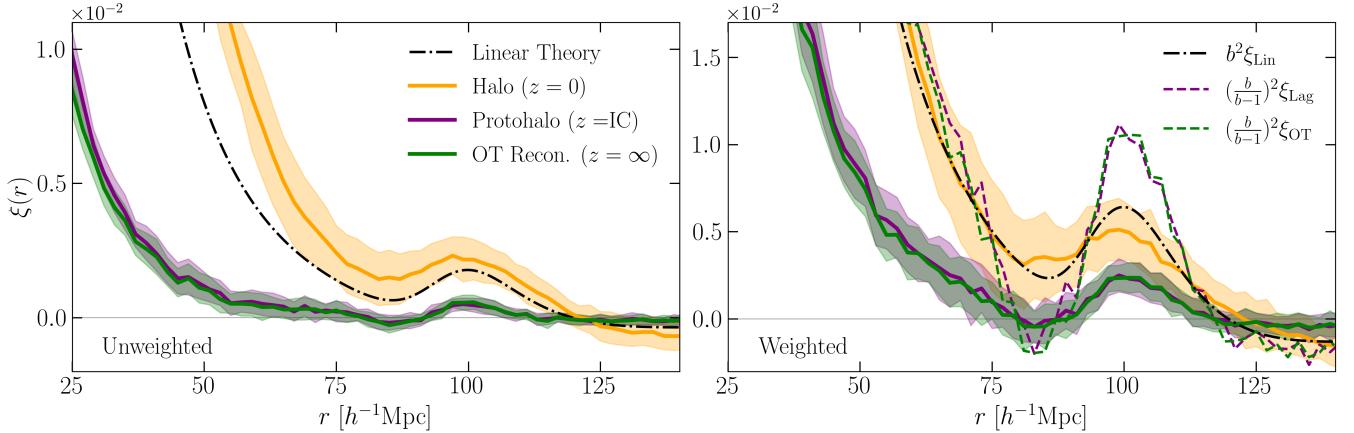


FIG. 3: Left: Unweighted pair correlation functions measured in non-overlapping bins of width  $2h^{-1}\text{Mpc}$ . Thick orange curve shows  $\xi_{\text{Eul}}$  of the  $z = 0$  halo centers averaged over 20 simulations, and yellow band shows the rms scatter around this mean. Thick purple curve and error band shows  $\xi_{\text{Lag}}$  of the corresponding Lagrangian (i.e.  $z = \infty$ ) protohalo centers. Thick green curve and error band – which is very similar to  $\xi_{\text{Lag}}$  – shows  $\xi_{\text{OT}}$  measured in our OT reconstructed field. Smooth dot-dashed curve shows  $\xi_{\text{Lin}}$ , the correlation function of the dark matter in linear theory. The ‘linear bias factor’  $b = \sqrt{\xi_{\text{Eul}}/\xi_{\text{Lin}}}$  is  $\approx 1.4$ . Right: Same as left, but now objects are weighted by their mass, so  $b = 1.9$ ; dot-dashed curve now shows  $b^2 \xi_{\text{Lin}}$ , and the short and long dashed curves show  $(1 - 1/b)^{-2}$  times  $\xi_{\text{Lag}}$  and  $\xi_{\text{OT}}$  respectively. The difference in shape between  $\xi_{\text{Eul}}$  and  $b^2 \xi_{\text{Lin}}$  is why reconstruction is necessary. The difference in shape between  $\xi_{\text{Lag}}$  and  $b^2 \xi_{\text{Lin}}$  is due to scale dependent Lagrangian bias.

and the green curve shows  $\xi_{\text{OT}}$ , which is built from the OT-reconstructed Laguerre cell barycenters.

To make the point that our results do not depend on the halo sample, the right hand panel shows results when we weight each halo by its mass when computing the pair counts (this weight is the same for the corresponding protohalos and Laguerre cells, of course). Comparison with the left hand panel shows that mass-weighting increases the amplitudes of all the correlation functions, as expected [27]. (We have also checked that the mass-weighted correlation function of all the particles – i.e. including the ‘dust’ – is zero. This confirms that, once all the mass is included, the OT algorithm has indeed converged to a uniform density initial condition.)

The dot-dashed curve in the left hand panel shows  $\xi_{\text{Lin}}$ , the correlation function of the dark matter field in linear theory. The ‘linear’ ‘scale-independent’ bias factor  $b \equiv \sqrt{\xi_{\text{Eul}}/\xi_{\text{Lin}}}$  in this panel is about 1.4, whereas it is 1.9 in the right hand panel. The most important point we wish to make is that, whatever the value of  $b$ ,  $\xi_{\text{OT}}$  reproduces the shape of  $\xi_{\text{Lag}}$  very well, despite the fact that  $\xi_{\text{Lag}}$  is rather different from  $\xi_{\text{Eul}}$ , in both panels.

We are most interested in whether the various  $\xi$  have different shapes, since differences in amplitude alone will not bias cosmological constraints. Therefore, the dot-dashed curve in the right hand panel shows  $b^2 \xi_{\text{Lin}}$  (with  $b = 1.9$ ) rather than  $\xi_{\text{Lin}}$  itself. This shows clearly that, compared to the linear theory shape, the BAO feature is smeared out in  $\xi_{\text{Eul}}$ . The reason why is well understood [3–5], and is why reconstruction is necessary.

Similarly, the difference in amplitude between  $\xi_{\text{Lag}}$  and  $\xi_{\text{Eul}}$  is striking but well understood: for tracers of a fixed

number density, such as Eulerian halos, their Lagrangian protohalos and their OT-reconstructed Laguerre cells, the correlation functions are related:

$$\xi_{\text{Eul}} \approx b^2 \xi_{\text{Lin}} \quad \text{then} \quad \xi_{\text{Lag}} \approx (b - 1)^2 \xi_{\text{Lin}} \quad (2)$$

[26, 27, 34, 35]. To test this expectation, the short and long dashed curves in the right hand panel show the result of multiplying  $\xi_{\text{Lag}}$  and  $\xi_{\text{OT}}$  by a scale-independent factor of  $(1 - 1/b)^{-2}$ . This indeed brings the curves into good agreement on scales of  $\sim 70h^{-1}\text{Mpc}$ . However, compared to the linear theory shape, the BAO feature in  $\xi_{\text{Lag}}$  is significantly enhanced. Although the reason for this ‘scale-dependent bias’ is also well understood [36–38], this shape difference is ignored – or simply not reproduced – by essentially all other reconstruction schemes [e.g. 7, 9, 10, 39]. In contrast,  $\xi_{\text{OT}}$  reproduces  $\xi_{\text{Lag}}$  on these scales exquisitely.

The results we have shown so far assume that the mass of each tracer is known perfectly. Our results are robust to factor-of-two or three errors in the mass estimate. The fidelity and robustness of the OT-reconstructed  $\xi_{\text{OT}}$  potentially enables a number of novel analyses which we now discuss. We caution, however, that before they can be applied to real data, we must be able to treat two additional complications. The first is that, so far, all our results are in configuration space; we have yet to show that our OT reconstruction remains accurate when redshift space distorted positions are used. The second is that we have yet to build a model for the ‘dust’ which plays an important role in our algorithm. This is why some of the discussion below is qualitative: a more quan-

titative analysis would only be justified after these two complications have been more fully addressed.

The agreement between  $\xi_{\text{OT}}$  and  $\xi_{\text{Lag}}$  is a significant success of the OT reconstruction. However, since  $\xi_{\text{OT}}$  (like  $\xi_{\text{Lag}}$ ) differs in shape from  $\xi_{\text{Lin}}$ , we now address how this compromises the original goal, which is to estimate the BAO scale.

There are three approaches to this problem. The traditional approach fits a fiducial  $\xi_{\text{Lin}}$  to the reconstructed correlation function (in this case,  $\xi_{\text{OT}}$ ), but as  $\xi_{\text{OT}}$  and  $\xi_{\text{Lin}}$  have different shapes, this might lead to a bias. We believe the reason for the shape difference is reasonably well understood [36–38], so we are in the process of determining a better fiducial template shape to fit. The second is motivated by [7, 38], and seeks to correct the shape of  $\xi_{\text{OT}}$  so that it becomes more like  $\xi_{\text{Lin}}$  by using the OT displacements to distort a uniform grid, and then cross-correlating the result with the OT field. We discuss both these efforts elsewhere. The third, as we describe below, is a much simpler test that is *not* tied to the detailed shape of  $\xi_{\text{Lin}}$ .

The midpoint between the peak and dip scales in  $\xi_{\text{Eul}}$ ,  $r_{\text{LP}} \equiv (r_{\text{peak}} + r_{\text{dip}})/2$ , is close in scale to  $r_{\text{LP}}$  in  $\xi_{\text{Lin}}$  [40]. In addition, Ref.[40] argued that  $r_{\text{LP}}$  should be much less sensitive to the scale-dependent bias which causes the shapes of  $\xi_{\text{Lag}}$  and  $\xi_{\text{Lin}}$  to differ. The right hand panel of Figure 3 shows that although  $\xi_{\text{Eul}}$ ,  $\xi_{\text{Lin}}$  and  $\xi_{\text{OT}}$  have different amplitudes around  $r_{\text{peak}}$  and  $r_{\text{dip}}$ , they all intersect at a scale that is approximately half-way in between. I.e.,  $r_{\text{LP}}$  is indeed relatively immune to both scale dependent bias and evolution, so we will use it to quantify the gain from using  $\xi_{\text{OT}}$  rather than  $\xi_{\text{Eul}}$ .

To estimate  $r_{\text{LP}}$  we fit polynomials, or more carefully chosen basis functions, to  $\xi$  over the range  $r = [60, 120]h^{-1}\text{Mpc}$ , and then differentiate the fit [41–43]. The peak and LP scales in  $\xi_{\text{Eul}}$ ,  $(r_{\text{peak}}, r_{\text{LP}})/h^{-1}\text{Mpc} = (98.5, 91.1)$ , are shifted from their values in  $\xi_{\text{Lin}}$ ,  $(99.8, 92.7)$ ; this 1% shift is why reconstruction was necessary. In contrast, they are  $(100.7, 92.6)$  in both  $\xi_{\text{Lag}}$  and  $\xi_{\text{OT}}$ . This  $r_{\text{LP}}$  is very similar to its value in  $\xi_{\text{Lin}}$ , showing that  $r_{\text{LP}}$  is indeed insensitive to scale-dependent bias, and so OT reconstruction works well:  $r_{\text{LP}}$  in  $\xi_{\text{OT}}$  provides a nearly unbiased constraint on the cosmological distance scale.

Distance scale estimates use the shape of  $\xi_{\text{OT}}$  but not its amplitude. In most analyses, this amplitude is proportional to the product of a bias factor  $b$  and the amplitude of fluctuations in the dark matter field, which is sometimes referred to as  $\sigma_8$ . As we describe below, the fact that the OT-reconstructed field is so similar to the Lagrangian one provides three separate ways to break this  $b\sigma_8$  degeneracy.

The first is more standard: ratios of the two- and three-point correlation functions yield determinations of bias factors that are independent of  $\sigma_8$  [44]. In the Eulerian field these ratios involve terms coming from perturbation

theory, as well as from bias. However, since our OT-reconstructions are so faithful to the Lagrangian field, measuring these ratios in the OT field should be both feasible and easy to interpret. Moreover, combining them with similar measurements in the evolved field should provide useful constraints on the terms which arise from evolution.

Our OT reconstructions enable two other novel methods for determining  $b$ . First, because  $\xi_{\text{Eul}}/\xi_{\text{OT}} \approx (1 - 1/b)^{-2}$  on  $\sim 50\text{--}70h^{-1}\text{Mpc}$  scales, (Figure 3), one could combine measurements over a range of  $r$  by fitting for that multiplicative factor  $A$  which brings  $\xi_{\text{Eul}}$  and  $\xi_{\text{OT}}$  into agreement over this range of scales. Since  $A = (1 - 1/b)^2$  this furnishes a direct estimate of  $b$  with *no* assumptions about the shape or amplitude of  $\xi_{\text{Lin}}$ .

In addition to the shape of  $\xi_{\text{OT}}$ , the OT algorithm also returns a distribution of displacements. Let  $\Sigma_{\text{OT}}$  denote the rms of  $|\mathbf{x} - \mathbf{q}_{\text{OT}}|$ .<sup>2</sup> If we ignore redshift space distortions, this rms is approximately  $\int dk P_{\text{Lin}}(k)/2\pi^2$ , which is independent of  $b$ . Following Ref. [43], let  $\Sigma_{\text{obs}}^2$  denote the result of performing this integral with the observed power spectrum  $P_{\text{Eul}}$  in place of  $P_{\text{Lin}}$ . Since  $P_{\text{Eul}} \approx b^2 P_{\text{Lin}}$ , the ratio  $\Sigma_{\text{obs}}/\Sigma_{\text{OT}} \approx b$ .

In summary: Our weighted OT algorithm provides a cosmology independent reconstruction of the BAO feature in the correlation function of biased tracers and potentially provides three distinct ways of determining the bias parameter  $b$  independently of  $\sigma_8$ , each with very different systematics. In forthcoming work, we will extend our algorithm to work with redshift-space distorted positions (e.g., by incorporating results of [17]), and to model the dust (e.g., [28–30]), before applying it to real observational data.

---

\* Electronic address: farnik@sas.upenn.edu

- [1] P. J. E. Peebles and J. T. Yu, *Astrophys. J.* **162**, 815 (1970).
- [2] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H.-J. Seo, M. Tegmark, Z. Zheng, et al., *Astrophys. J.* **633**, 560 (2005), astro-ph/0501171.
- [3] S. Bharadwaj, *Astrophys. J.* **472**, 1 (1996), astro-ph/9606121.
- [4] M. Crocce and R. Scoccimarro, *Phys. Rev. D* **77**, 023533 (2008), 0704.2783.
- [5] D. J. Eisenstein, H.-J. Seo, and M. White, *Astrophys. J.* **664**, 660 (2007), astro-ph/0604361.
- [6] D. J. Eisenstein, H.-J. Seo, E. Sirk, and D. N. Spergel, *Astrophys. J.* **664**, 675 (2007), astro-ph/0604362.

---

<sup>2</sup> We note in passing that  $\Sigma_{\text{OT}}$  can be used directly in the Laguerre reconstruction algorithm of Ref. [42, 43] to provide a consistency check of the BAO distance scale.

- [7] 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), 1202.0090.
- [8] E. Sarpa, C. Schimd, E. Branchini, and S. Matarrese, Mon. Not. R. Astron. Soc. **484**, 3818 (2019), 1809.10738.
- [9] M. Schmittfull, T. Baldauf, and M. Zaldarriaga, Phys. Rev. D **96**, 023505 (2017), 1704.06634.
- [10] R. Hada and D. J. Eisenstein, Mon. Not. R. Astron. Soc. **478**, 1866 (2018), 1804.04738.
- [11] J. Benamou and Y. Brenier, Numerische Mathematik **84**, 375 (2000), URL <https://doi.org/10.1007/s002110050002>.
- [12] C. Léonard, Discrete & Continuous Dynamical Systems **34**, 1533 (2014).
- [13] A. Mondino and S. Suhr, arXiv e-prints arXiv:1810.13309 (2018), 1810.13309.
- [14] G. Peyré and M. Cuturi, arXiv e-prints arXiv:1803.00567 (2018), 1803.00567.
- [15] U. Frisch, S. Matarrese, R. Mohayaee, and A. Sobolevskii, Nature **417** (2002).
- [16] Y. Brenier, U. Frisch, M. Hénon, G. Loeper, S. Matarrese, R. Mohayaee, and A. Sobolevskii, Mon. Not. R. Astron. Soc. **346** (2003).
- [17] B. Levy, R. Mohayaee, and S. von Hausegger, Mon. Not. R. Astron. Soc. **506**, 1165 (2021), 2012.09074.
- [18] Y. Brenier, Communications on Pure and Applied Mathematics **44**, 375 (1991).
- [19] C. Villani, *Optimal transport : old and new*, Grundlehren der mathematischen Wissenschaften (Springer, Berlin, 2009), ISBN 978-3-540-71049-3, URL <http://opac.inria.fr/record=b1129524>.
- [20] F. Santambrogio, *Optimal transport for applied mathematicians*, vol. 87 of *Progress in Nonlinear Differential Equations and their Applications* (Birkhäuser/Springer, Cham, 2015), ISBN 978-3-319-20827-5; 978-3-319-20828-2, calculus of variations, PDEs, and modeling.
- [21] B. Lévy and E. L. Schwindt, Comput. Graph. **72**, 135 (2018), URL <https://doi.org/10.1016/j.cag.2018.01.009>.
- [22] R. van de Weygaert, A&A **283**, 361 (1994).
- [23] J. Kitagawa, Q. Mérigot, and B. Thibert, CoRR (2016), 1603.05579, URL <http://arxiv.org/abs/1603.05579>.
- [24] F. Aurenhammer, F. Hoffmann, and B. Aronov, in *Symposium on Computational Geometry* (1992), pp. 350–357.
- [25] B. Lévy, ESAIM M2AN (Mathematical Modeling and Analysis) (2015).
- [26] H. J. Mo and S. D. M. White, Mon. Not. R. Astron. Soc. **282**, 347 (1996), astro-ph/9512127.
- [27] R. K. Sheth and G. Tormen, Mon. Not. R. Astron. Soc. **308**, 119 (1999), astro-ph/9901122.
- [28] Y.-C. Cai, G. Bernstein, and R. K. Sheth, Mon. Not. R. Astron. Soc. **412**, 995 (2011), 1007.3500.
- [29] H. Wang, H. J. Mo, X. Yang, Y. P. Jing, and W. P. Lin, The Astrophysical Journal **794**, 94 (2014), URL <https://doi.org/10.1088/0004-637x/794/1/94>.
- [30] A. Paranjape and S. Alam, Mon. Not. R. Astron. Soc. **495**, 3233 (2020), 2001.08760.
- [31] F. Villaescusa-Navarro, A. Banerjee, N. Dalal, E. Castorina, R. Scoccimarro, R. Angulo, and D. N. Spergel, The Astrophysical Journal **861**, 53 (2018), URL <https://doi.org/10.3847/1538-4357/aac6bf>.
- [32] G. Despali, G. Tormen, and R. K. Sheth, Mon. Not. R. Astron. Soc. **431**, 1143 (2013), 1212.4157.
- [33] A. D. Ludlow, M. Borzyszkowski, and C. Porciani, Mon. Not. R. Astron. Soc. **445**, 4110 (2014).
- [34] A. Nusser and M. Davis, Astrophys. J. Lett. **421**, L1 (1994), astro-ph/9309009.
- [35] J. N. Fry, Astrophys. J. Lett. **461**, L65 (1996).
- [36] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304**, 15 (1986).
- [37] V. Desjacques, M. Crocce, R. Scoccimarro, and R. K. Sheth, Phys. Rev. D **82**, 103529 (2010), 1009.3449.
- [38] T. Baldauf and V. Desjacques, Phys. Rev. D **95**, 043535 (2017), 1612.04521.
- [39] E. Sarpa, C. Schimd, E. Branchini, and S. Matarrese, Mon. Not. R. Astron. Soc. **484**, 3818 (2019), 1809.10738.
- [40] S. Anselmi, G. D. Starkman, and R. K. Sheth, Mon. Not. R. Astron. Soc. **455**, 2474 (2016), 1508.01170.
- [41] S. Anselmi, P.-S. Corasaniti, G. D. Starkman, R. K. Sheth, and I. Zehavi, Phys. Rev. D **98**, 023527 (2018), 1711.09063.
- [42] F. Nikakhtar, R. K. Sheth, and I. Zehavi, Phys. Rev. D **104**, 043530 (2021), 2101.08376.
- [43] F. Nikakhtar, R. K. Sheth, and I. Zehavi, Phys. Rev. D **043536** (2022), 2110.03591.
- [44] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, Phys. Rept. **367**, 1 (2002), astro-ph/0112551.